Introduction to Cryptography

Jean-Sébastien Coron

January 2007

Outline

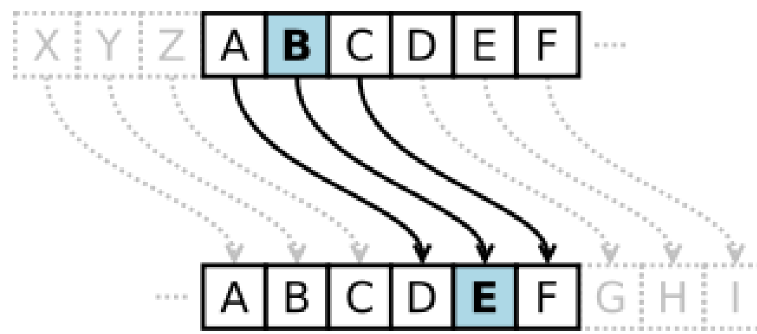
- History
  - From the Greeks to Shannon
- Modern Cryptography
  - Goals: confidentiality, integrity, authenticity
  - Schemes: secret-key encryption, public-key encryption, digital signature…
  - Tools: block-ciphers, discrete-log hard groups, trapdoor permutations…

Early history

- Hiding the content of a message
  - Scytale: 500 B.C., used by the Spartan military.

Caesar's cipher (50 B.C.)

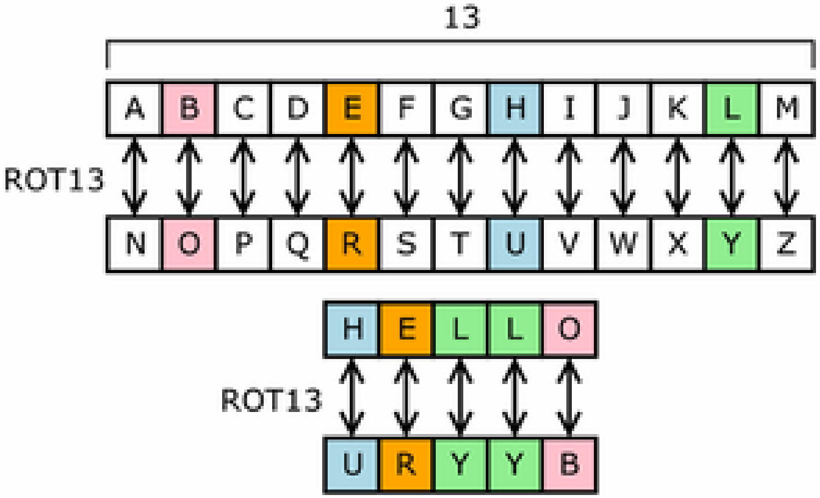
- Used by Caesar to communicate with his generals
- Each letter is shifted by a constant position in the alphabet.



  - With n=3: VENI VIDI VICIT => YHQL YLGL YLFL
- Only 25 possibilities for n => WEAK

ROT13

- Caesar cipher with n=13



- Sometimes used in online forums
- WEAK

Mono-alphabetic Cipher

- Each letter is replaced with another letter, according to a fixed substitution

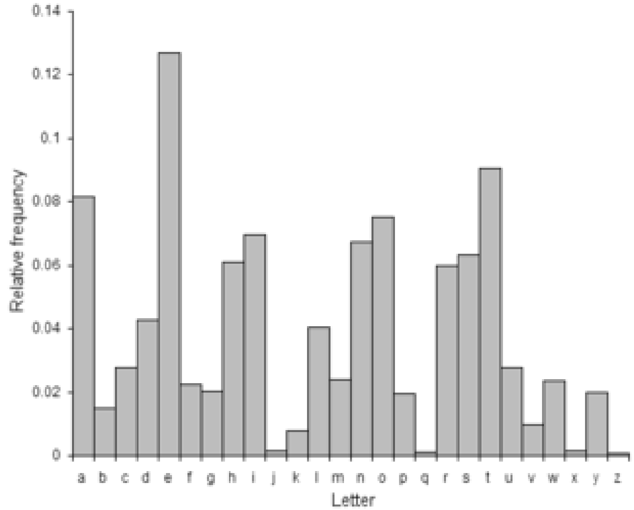
Plaintext :	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
Ciphertext:	C	G	H	U	Z	J	T	E	L	Y	X	I	F	O	P	K	J	W	V	A	B	D	M	S	N	Q

Then `HELLO WORLD` enciphers to `EZIIP MPWIU`

Number of possible keys is large:
26!=2^88.4 or 88 bits, but…

Frequency analysis

- Frequency of letters in English:



- Cryptanalysis of mono-alphabetic cipher:
  - The most frequent letter in the ciphertext is likely E,T or A.
  - Substitute and continue with less frequent letters.
  - WEAK

The Vigenere Cipher (1586)

- Consists of several Caesar ciphers in sequence with different shift values
- Encrypt using a repeated keyword

```
CRYPTOCRYPTOCR
ATTACKTOMORROW
--------------
CWRPVYVFKDKFQN
```

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
A	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
B	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A
C	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B
D	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C
E	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
F	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E
G	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F
H	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G
I	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H
J	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I
K	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J
L	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
M	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L
N	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M
O	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N
P	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O
Q	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
R	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q
S	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
T	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
U	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T
V	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U
W	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V
X	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W
Y	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
Z	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y

Security of Vigenere

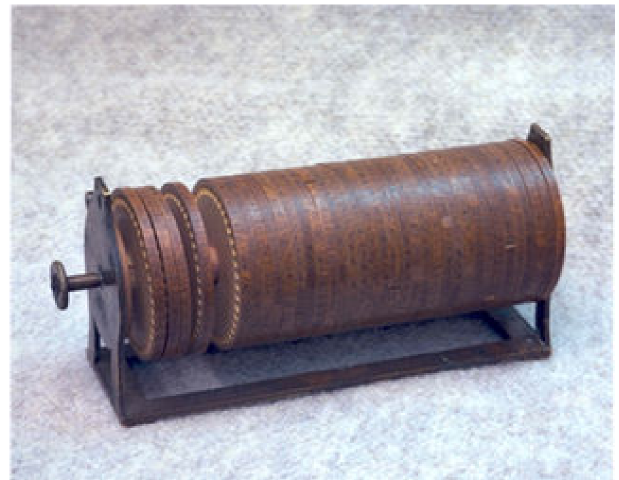
- Simple frequency analysis is defeated
- Critical weakness:
  - once the keyword's length is known, can be separated in a sequence of Caesar cipher, which can be individually broken.
  - Kasiski's attack (1863): consists in looking for repeated sequences in the ciphertext.

    SECRETESECRETESECRETESECRET
    CRYPTOLOGUESOU**CRYPTO**GRAPHES

    UVAGXHPGKWVWHY**UVAGXH**KJERYIL
  - Enables to discover the keyword length
- WEAK

Jefferson's disk (1795)

- Used by the US army between 1923 and 1942.
- A set of 20 to 30 numbered wheels, each with the letters of the alphabet arranged randomly around them.



- Plaintext is made on one row and ciphertext is read on another row.
- Sender and recipient have to agree on the ordering of the wheels.

Security of Jefferson's disk

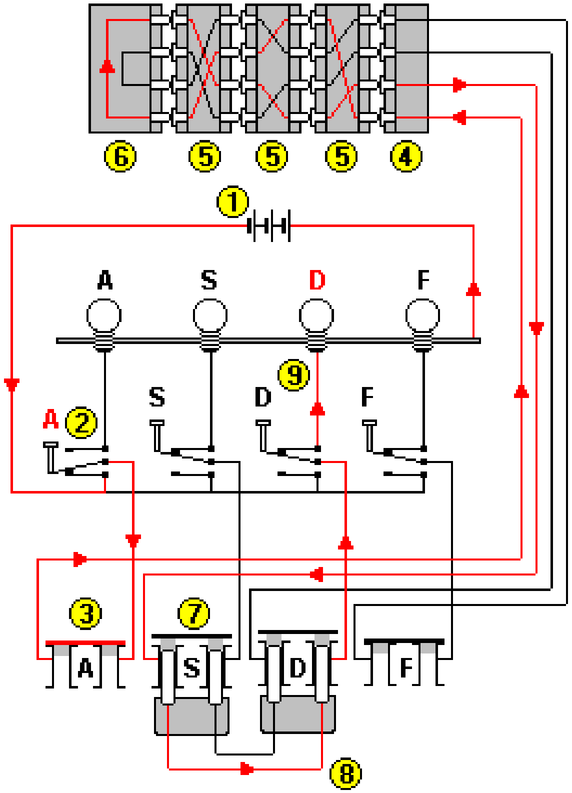
- Weakness: the offset between plaintext letter and ciphertext letter is the same for all disks.
- Assume that the attacker has a set of Jefferson's disk but doesn't know the ordering.
  - Assume that he knows some part of the plaintext (e.g., a plaintext always start with « hello »).
  - Then he can determine the possible offsets for each disk and each letter, and determine the common offset.

- WEAK

Enigma

- A cipher machine with rotors.
- Developed in the 20s and used by the German army during WW2.
- Encryption:
  - Set the initial rotor position and plug board
  - Types plaintext
  - Corresponding letter lights up and rotors moves one step.

Enigma



- Current flows from the battery (1) through the pressed switch A (2), the plugboard (3), the rotors (4)(5), the reflector (6), the rotors (5)(4), the plugboard (7)(8) to light-up the D lamp (9).
- The rotors move with every key press, which changes the electrical path.
- Implements polyalphabetic substitution with a long period.

Security of Enigma

- Very secure for that time
- Polish Cipher Bureau: algebraic cryptanalysis from Marian Rejewski.
- Bletchley Park: during WWII, Alan Turing and Gordon Welchman developed the « bombe » to search for the correct rotor positions.

One-time pad (1917)

- Plaintext is xored with the key to produce the ciphertext

```
 011001011001
 111010010010
─────────────
 100011001011
```

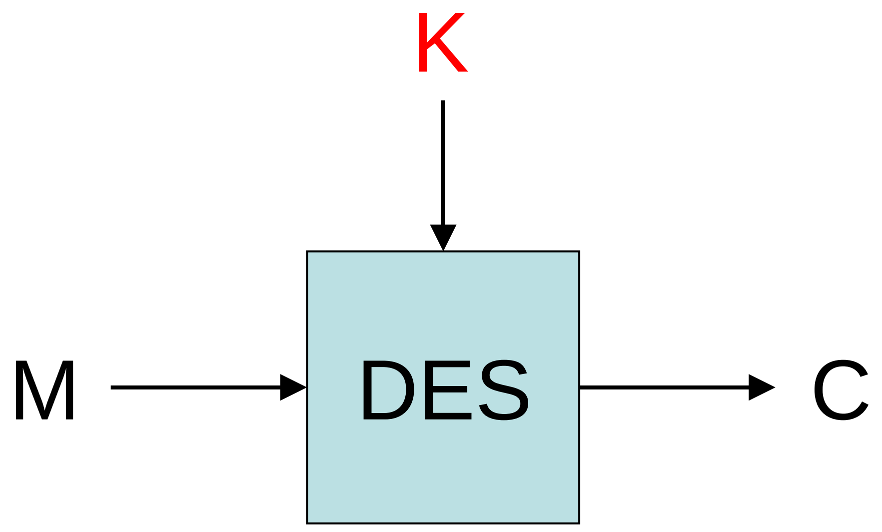
- Proved unbreakable by Shannon (1949) if key is random and as long as the plaintext.
- Issue: key as long as the plaintext.
- Used for the hotline between Washington and Moscow during the cold war.

DES (1976)

- Data Encryption Standard (DES), published as FIPS PUB 46.
- Developed by NBS (National Bureau of Standards), now NIST (National Institute of Standards and Technology), following an algorithm from IBM.
- De facto world-wide standard since 1976.
- Superseded by the AES, but remains in widespread use.

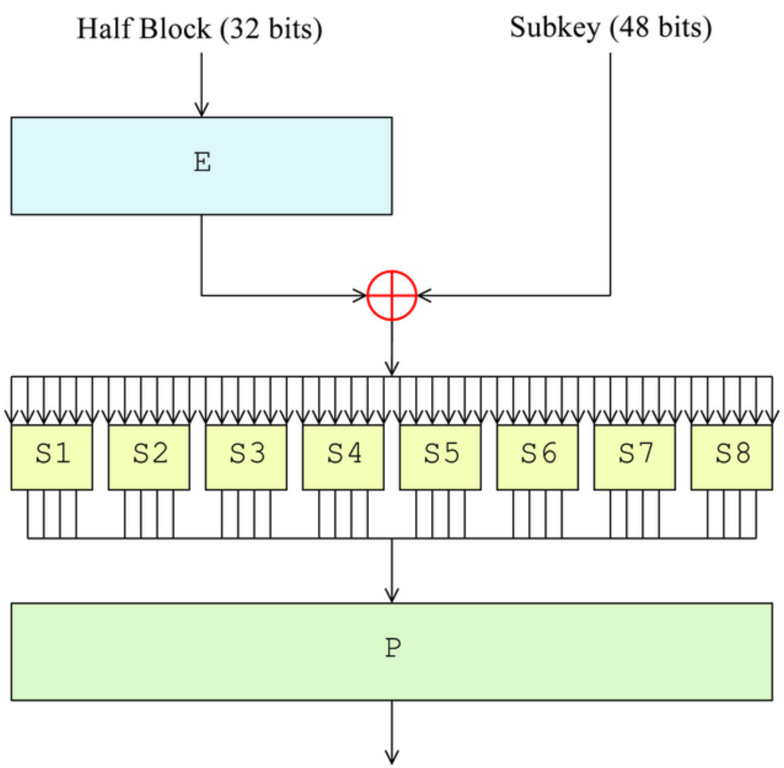
DES block-cipher

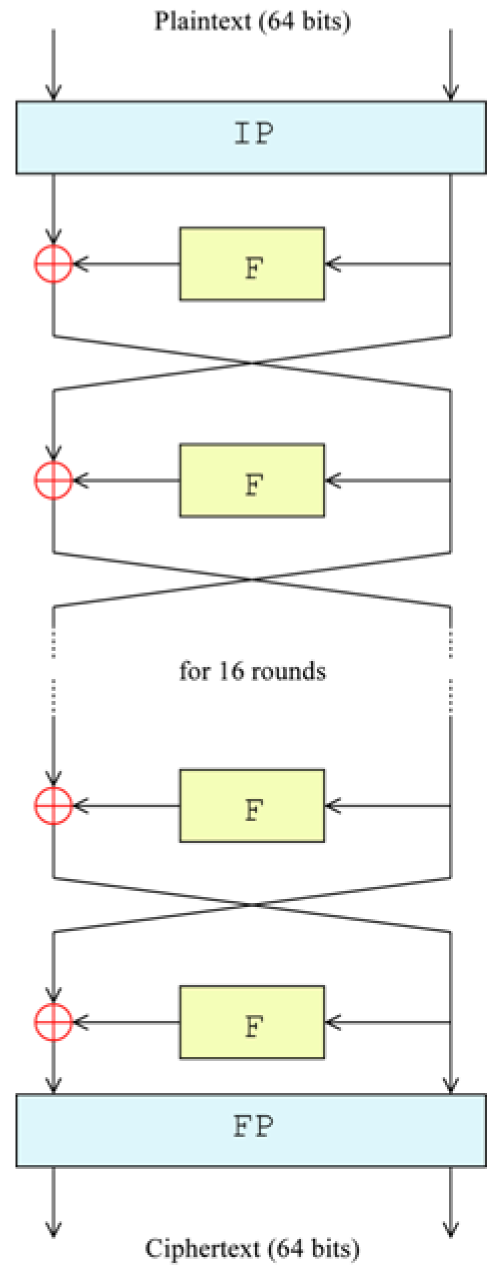
- Input length: 64 bits.
- Output length: 64 bits.
- Key length: 56 bits.



DES

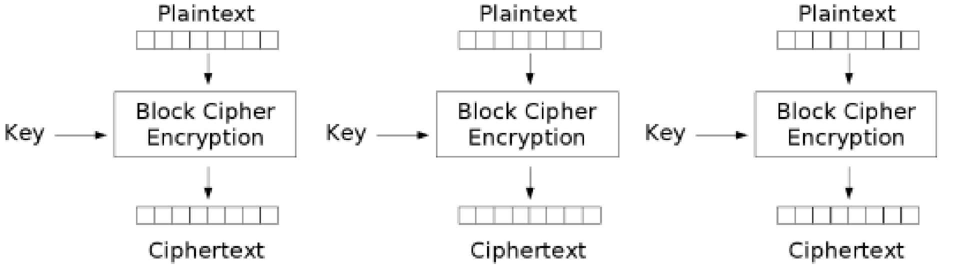
- Feistel Cipher
- F function:





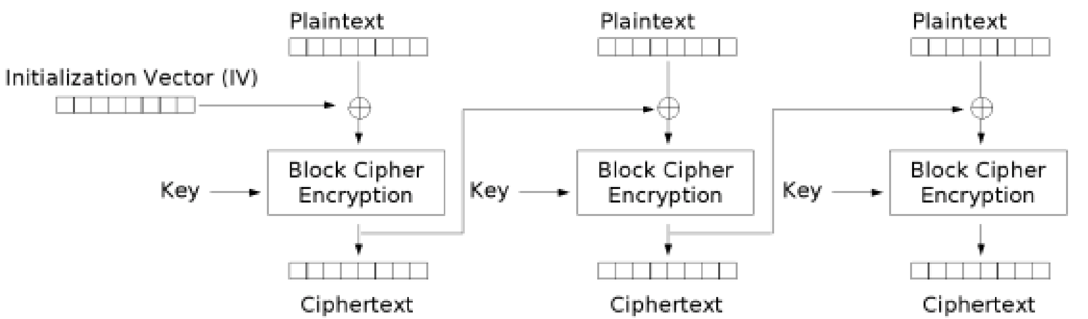
DES modes of operation

- Encrypting longer messages (>64 bits)
- FIPS-81: DES modes of operation
- ECB: WEAK



Electronic Codebook (ECB) mode encryption

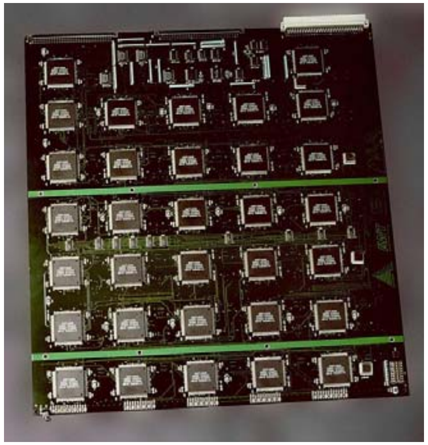
- CBC: OK



Cipher Block Chaining (CBC) mode encryption

Security of DES

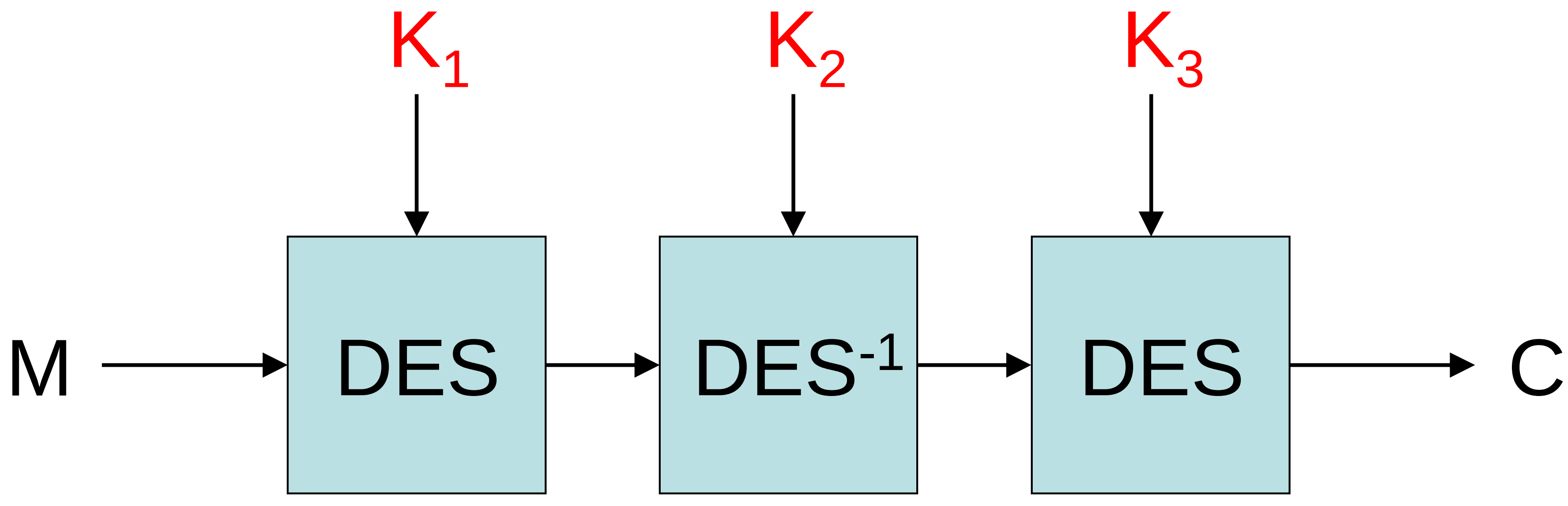
- Problem: key is too short (56 bits). Exhaustive search has become feasible
  - DES cracker from Electronic Frontier Foundation (EFF). Breaks DES in 2 days (1998).



- Other attacks
  - Differential cryptanalysis (Biham and Shamir). 2^47 chosen plaintexts
  - Linear cryptanalysis (Matsui, 1993). 2^43 known plaintexts.

TRIPLE DES

- Block cipher
  - 64-bit input and output, 168-bit key

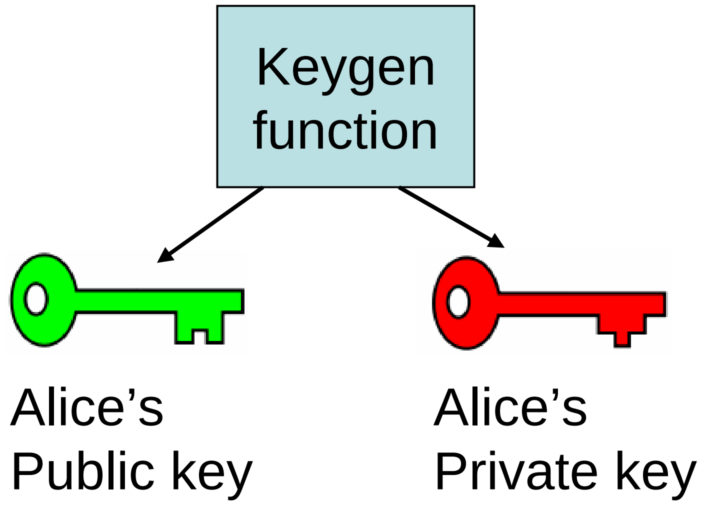


- Slowly disappearing, replaced by AES (6 times faster in software).

Public-key cryptography

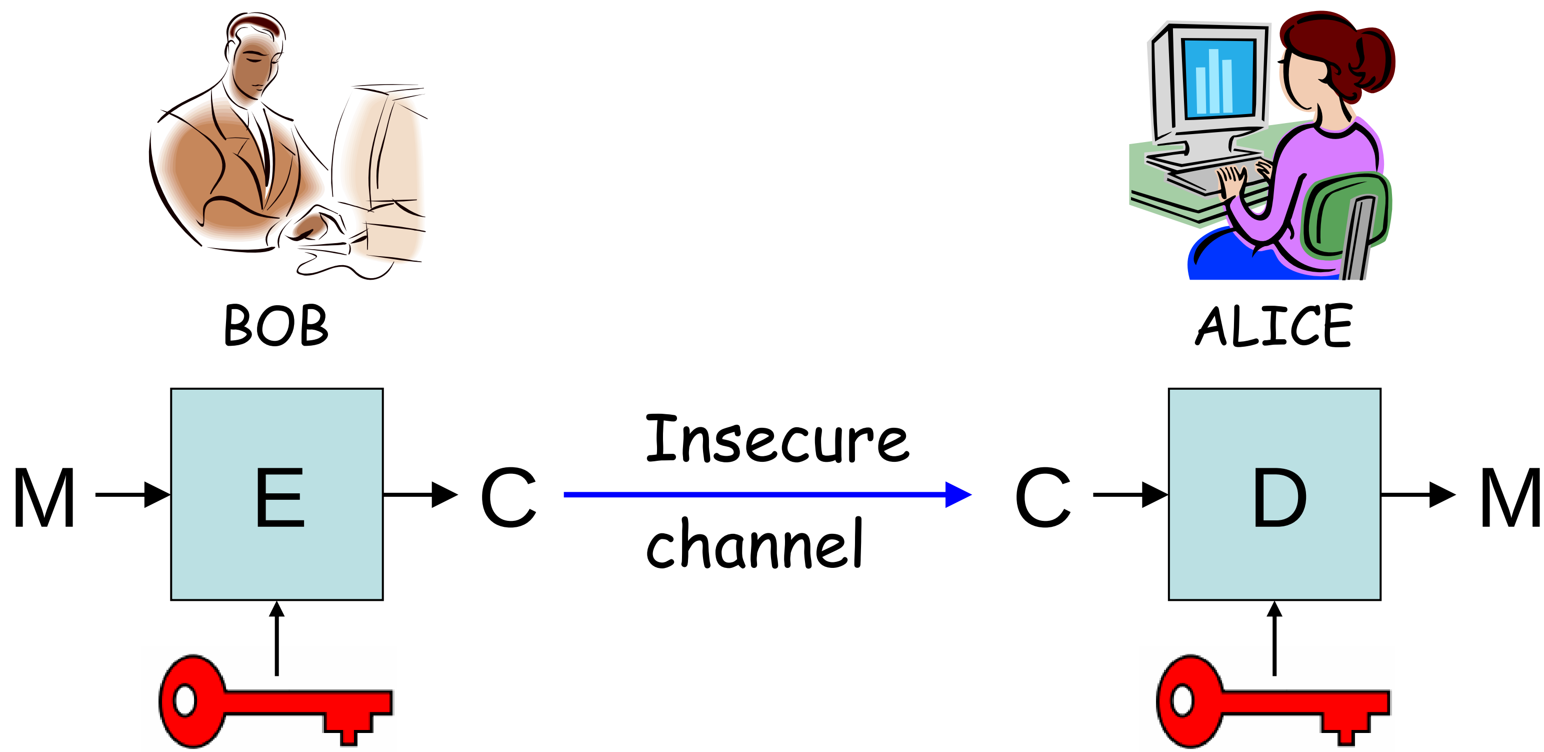
- Invented by Diffie and Hellman in 1976. Revolutionized the field.
- Each user now has two keys
  - A public key
  - A private key
- Should be hard to compute the private key from the public key.
- Enables:
  - Asymmetric encryption
  - Digital signatures
  - Key exchange
  - Identification,

  and many other protocols.



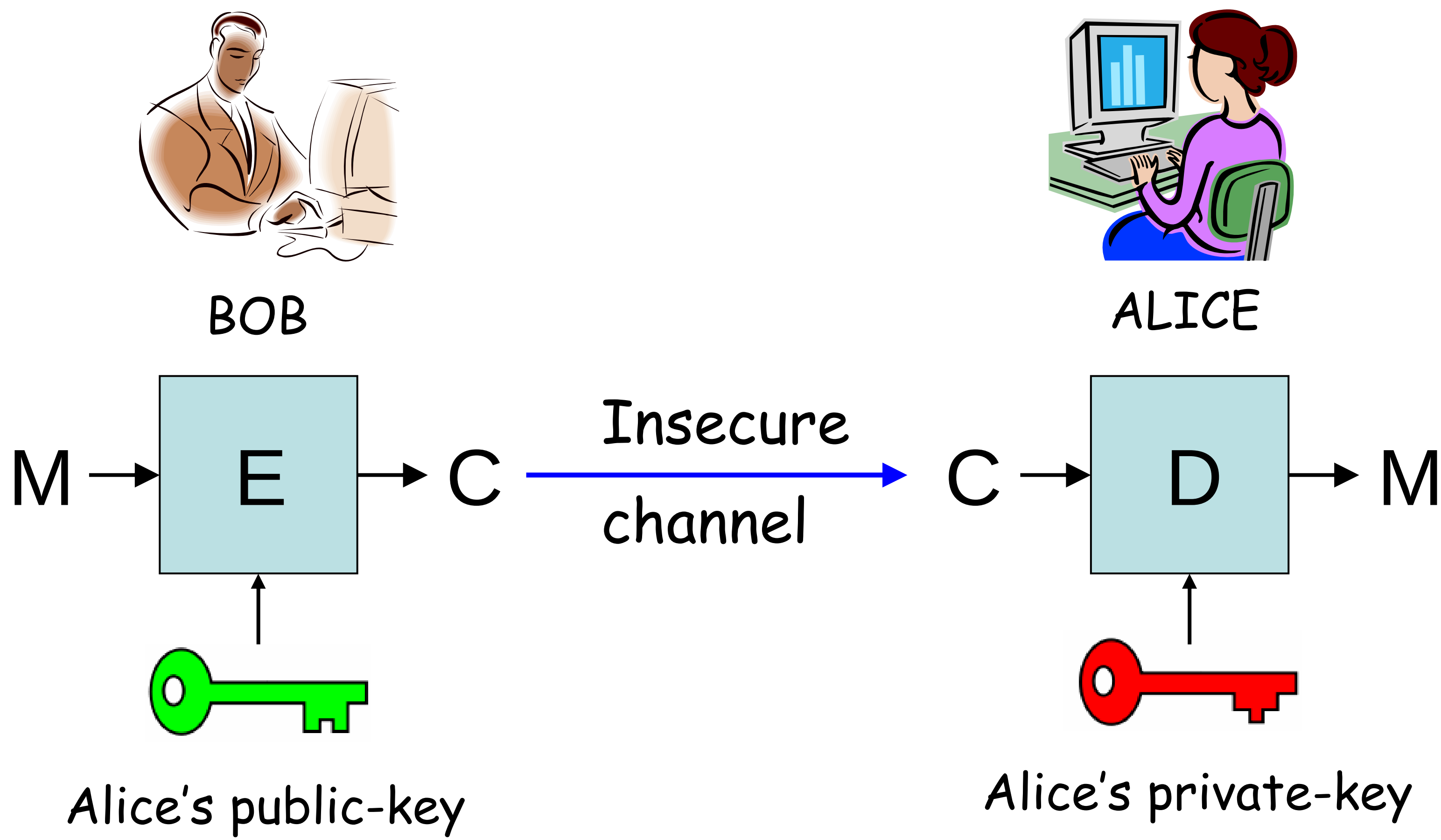
Key distribution issue

- Symmetric cryptography
  - How to initially distribute the key to establish a secure channel ?



Asymmetric encryption

- Solves the key distribution issue

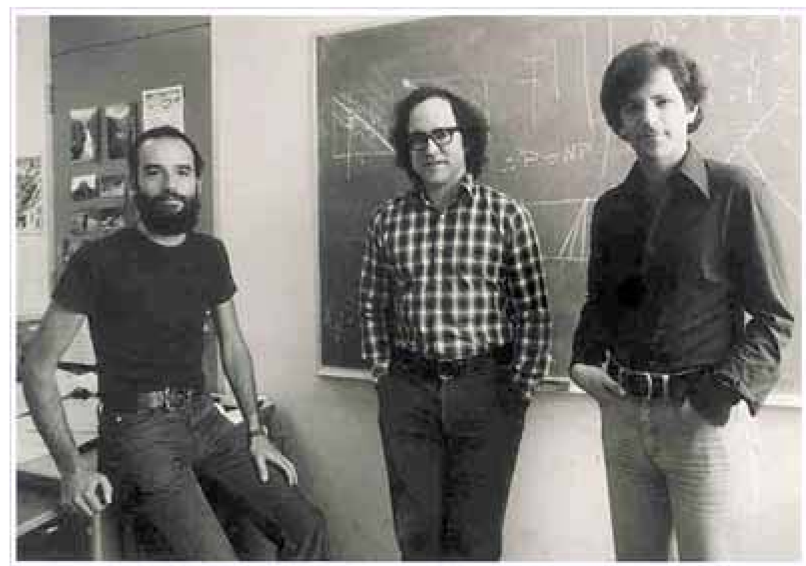


Analogy: the mailbox

- Bob wants to send a letter to Alice
  - Bob obtains Alice's adress
  - Bob puts his letter in Alice's mailbox
  - Alice opens her mailbox and read Bob's letter.
- Properties of the mailbox
  - Anybody can put a letter in the mailbox
  - Only Alice can open her mailbox

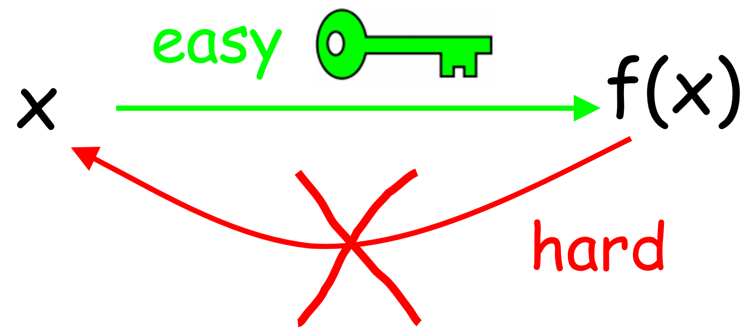
RSA (1977)

- Invented by Rivest, Shamir and Adleman
- First realization of asymmetric encryption.
- Implements a trapdoor one-way permutation.
- Still the most widely PK algorithm in use.

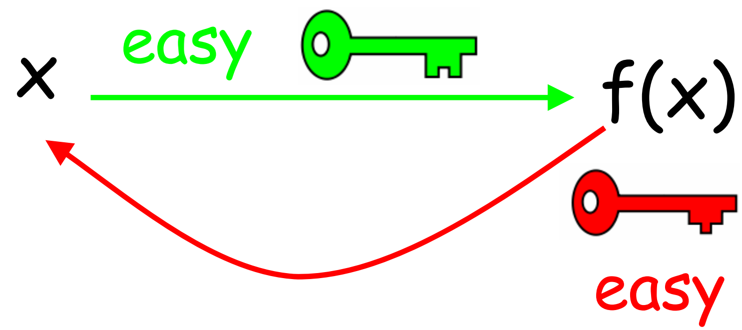


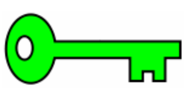
Trapdoor one-way permutation

- Trapdoor unknown:



- Trapdoor known:



- Asymmetric encryption:
  - Everybody can encrypt to Alice using 
  - Only Alice can decrypt using 

RSA

- Public key: $n=p.q$ and e
  - Primes p and q remain secret.
- Private key: d such that

  $$e.d=1 \bmod (p-1)(q-1)$$
- Encryption using public n,e:

  $$c=m^e \bmod n$$
- Decryption using private d:

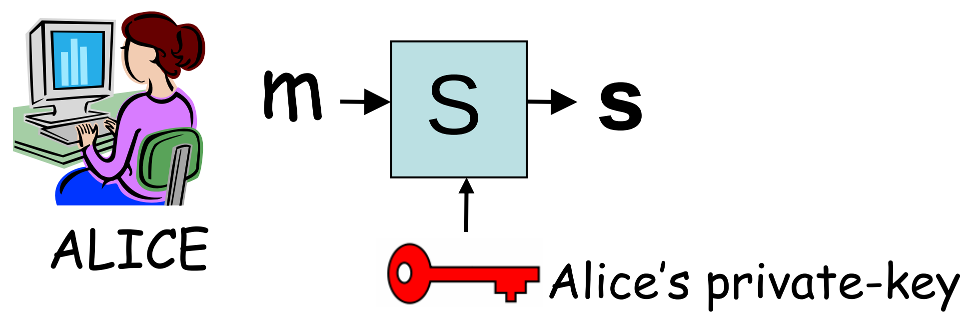
  $$m=c^d \bmod n$$
- PKCS#1 v2.1

RSA

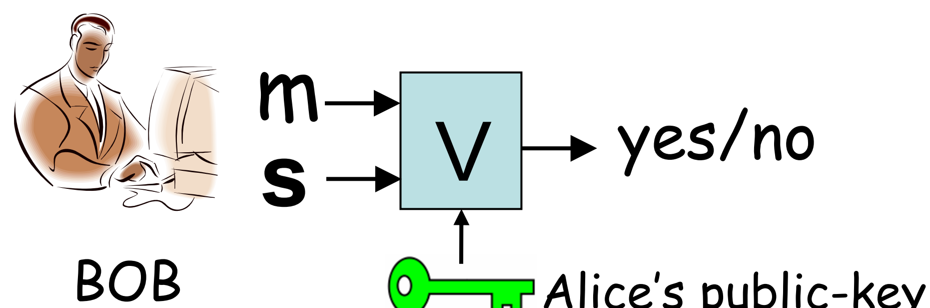
- Decryption works because $m = c^d = (m^e)^d = m^{e.d} = m$ because $e.d = 1 \bmod \mathbf{f}$
- Security is based on the hardness of factorization
  - Given $n = p.q$, no known efficient algorithm to recover p and q.
  - Factorization record: 663 bits (2005)
- Public modulus n must be large enough
  - At least 1024 bits. 2048 bits is better.

Digital signature

- A bit string that depends on the message m and the user's public-key
  - Only Alice can sign a message using her private-key



  - Anybody can verify Alice's signature of m given her public-key



Digital signature

- A digital signature provides:
  - Authenticity: only Alice can produce a signature of a message valid under her public-key.
  - Integrity: the signed message cannot be modified.
  - Non-repudiation: Alice cannot later claim that she did not sign the message

Signing with RSA

- Public key: $n=p.q$ and e
- Private key: d such that

  $$e.d=1 \bmod (p-1)(q-1)$$

- Signing using private d:

  $$\mathbf{s}=m^{d} \bmod n$$

- Verifying using public n,e:

  check that $m=\mathbf{s}^{e} \bmod n$
- ISO 9796-2, PKCS#1 v2.1

Other signature schemes

- Digital Signature Algorithm (DSA) (1993)
  - Digital Signature Standard (DSS) proposed by NIST, specified in FIPS 186.
  - Security based on the hardness of discrete log.
  - ECDSA: a variant of DSA for elliptic-curves
- Rabin signature scheme
  - Similar to RSA but with $e=2$
- El-Gamal signature scheme (1984)
  - Based on the discrete-log problem

Diffie-Hellman key exchange (1976)

- Public parameters: g and p

BOB

ALICE

$B = g^b$ $\xrightarrow{\quad B \quad}$ $A = g^a$

$\xleftarrow{\quad A \quad}$

$K_B = A^b = (g^a)^b = g^{a.b}$

$K_A = B^a = (g^b)^a = g^{b.a}$

$K_B = K_A$

Security of Diffie-Hellman

- Based on the hardness of the discrete-log problem:
  - Given $A=g^a$ mod p, find a
  - No efficient algorithm for large p.
- No authentication
  - Vulnerable to the man in the middle attack
- Authenticated key exchange
  - Using a PKI. Alice and Bob can sign A and B
  - Password-authenticated key-exchange IEEE P1363.2

Lessons from the past

- Cryptography is a permanent race between construction and attacks
  - but somehow this has changed with modern cryptography and security proofs.
- Security should rely on the secrecy of the key and not of the algorithm
  - Open algorithms enables open scrutiny.

Modern cryptography

- New functionalities
  - Identity-based encryption, voting, electronic money, auction…
- Formalization of security notions
  - What is a secure encryption scheme ? a secure signature scheme ?
- Construction of schemes or protocols that provably achieve these security notions
  - Based on some hardness assumption (e.g., factoring is hard).
- Modern cryptography is about security proofs.
  - A scheme without security proof is useless.