



Cybersecurity

Foundations of Cryptology

Mauro Barni

University of Siena

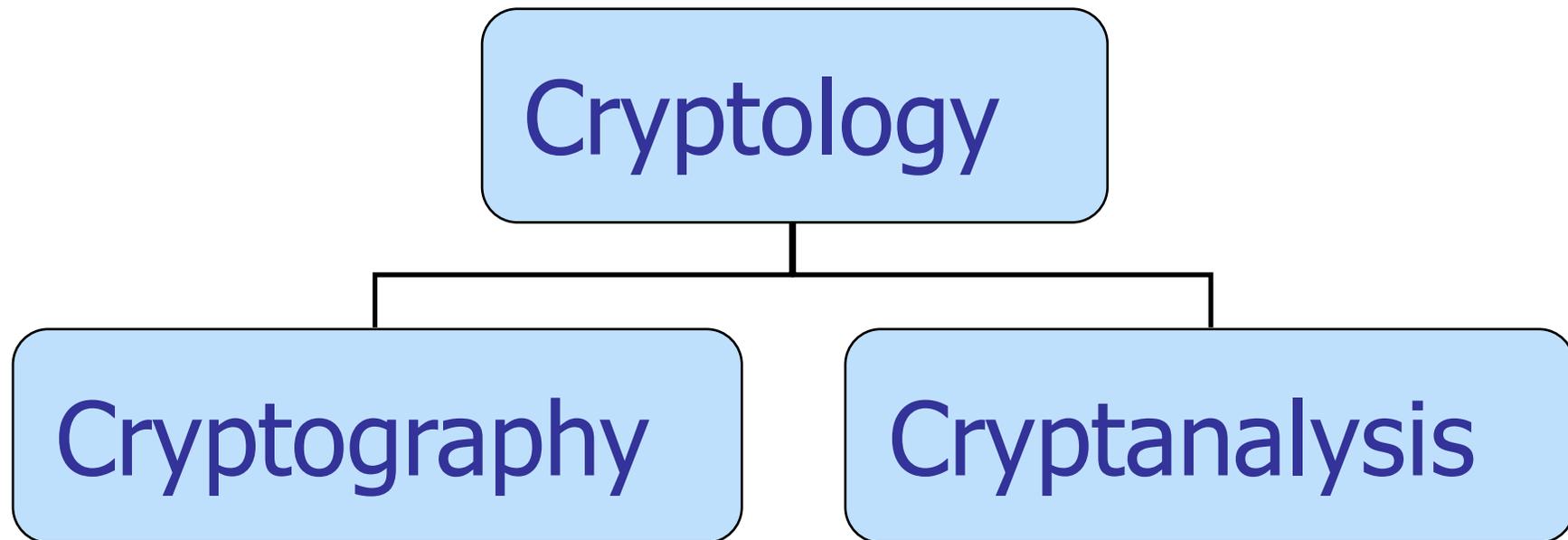


Cryptography

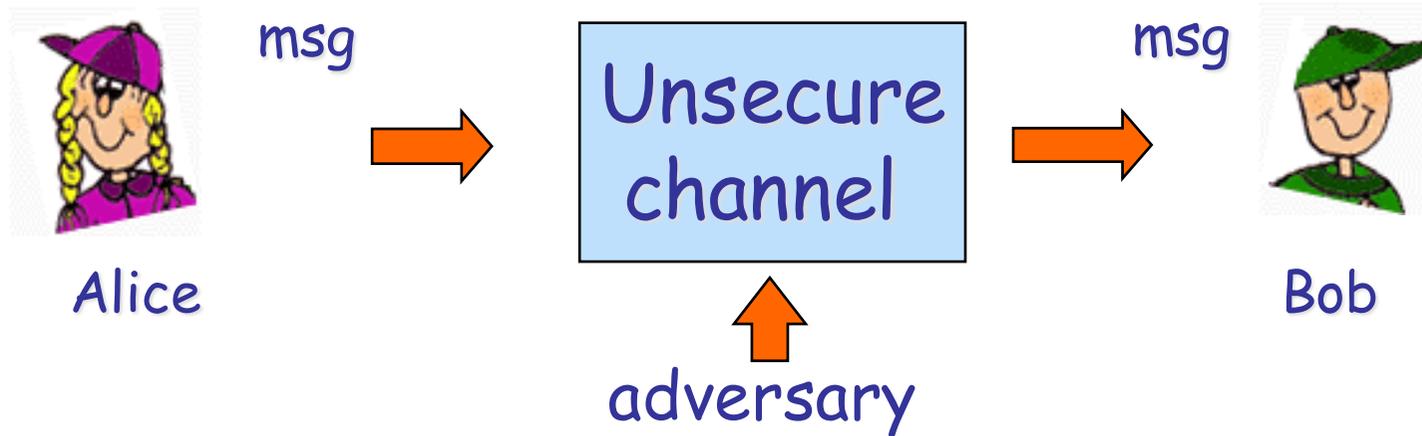
- Cryptography is the art or science of keeping messages secret;
 - the word cryptography is derived from Greek and literally means **secret** (crypto-) **writing** (-graphy)
- It allows to protect the content against disclosure or modification both during data transmission or storage.
- It deals with all aspects of secure messaging, authentication, digital signatures, electronic money ...



What is Cryptology?



Framework

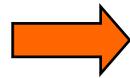


- A *sender* is an entity in a two-party communication which is the legitimate transmitter of information.
- A *receiver* is an entity which is the intended recipient of information.
- An *adversary* is an entity which is neither the sender nor receiver, and which tries to break the security (often the confidentiality) of the transmission.



Basic Terminology

"The quick
brown fox
jumps over
the lazy dog"



Cipher



"AxCv;5bmEseTfid3)
fGsmWe#4^,sdgfMw
ir3:dkJeTsY8R\s@!q
3%"

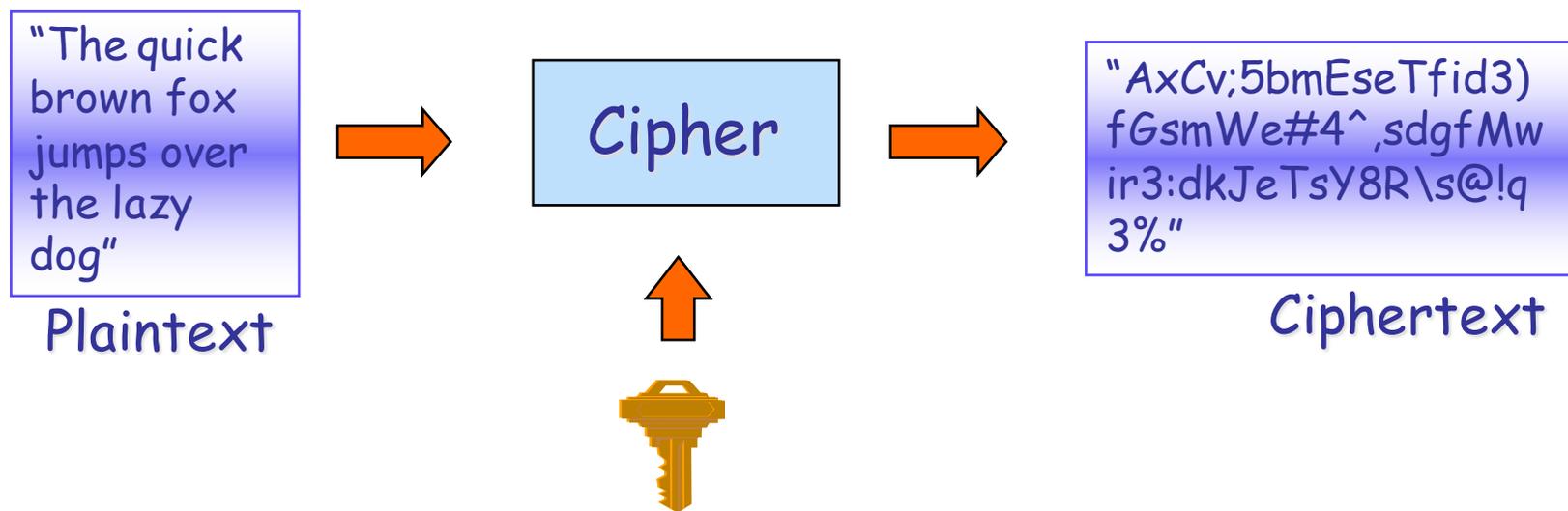
Plaintext
the original
message, i.e. the
stuff to secure.

Cipher
algorithm for
transforming plaintext
to ciphertext and
viceversa, composed by
encipher and decipher

Ciphertext
the coded message,
apparently random
sequence of data and
thus unreadable



Basic Terminology

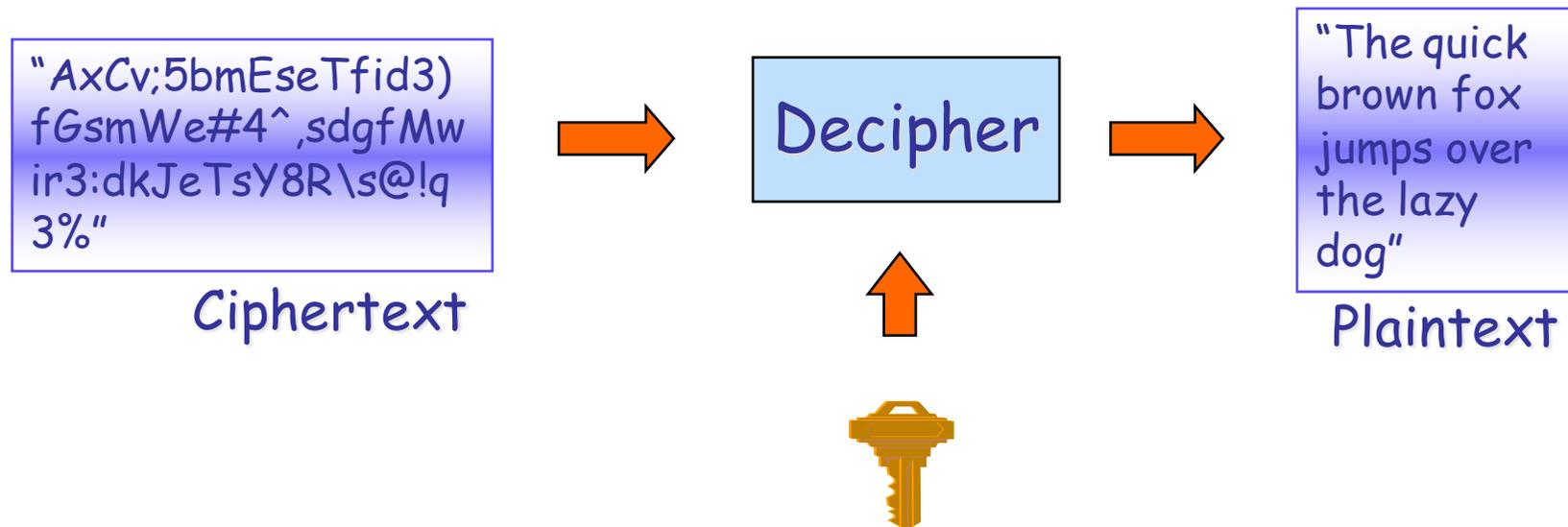


Key

some critical information used by the cipher, **independent** from plaintext. The cipher, given an input, will produce different outputs for different keys.



Basic Terminology



Decipher

(decryption) the process of converting ciphertext back into plaintext using a cipher and a key



Basic Terminology

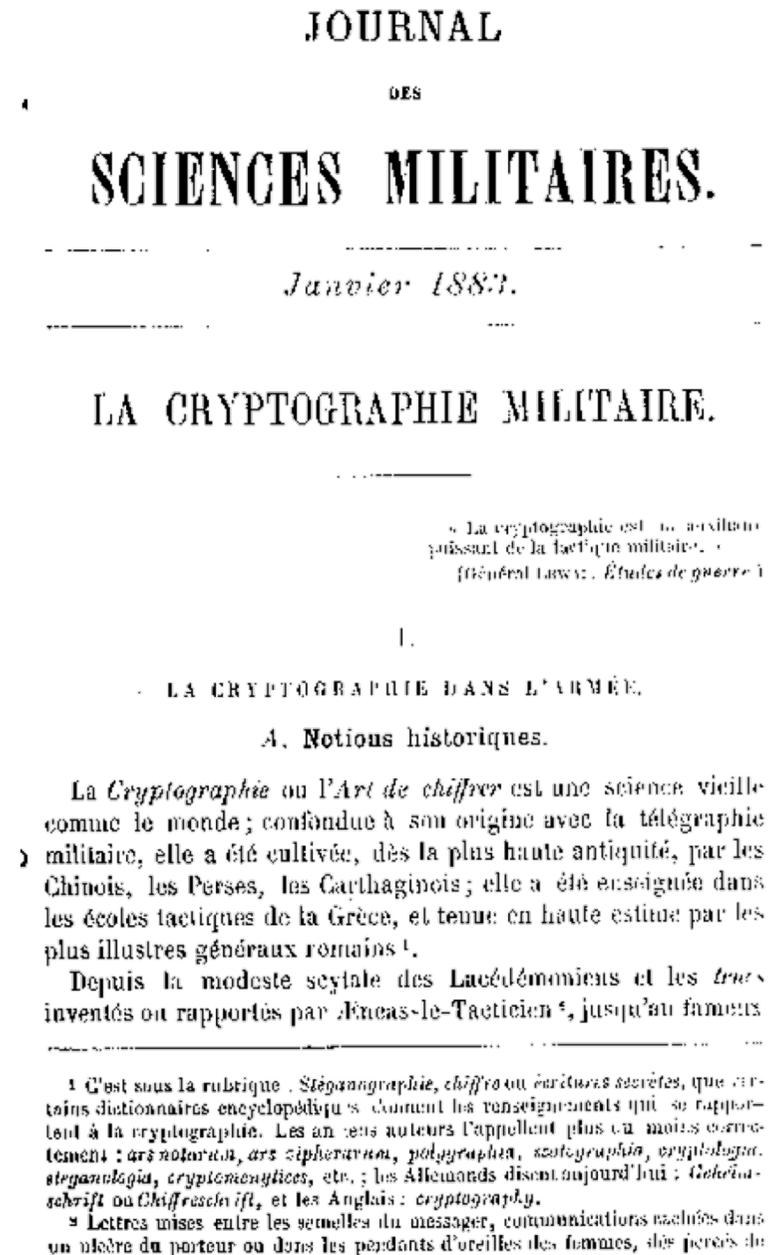
- **Kerckhoffs' principle**, stated by Auguste Kerckhoffs in the 19th century
- *“a cryptosystem should be secure even if everything about the system, except the key, is public knowledge”*





Basic Terminology

- In contrast to security through obscurity
- History has shown that maintaining the secrecy of the transformations is very difficult indeed.
 - Auguste Kerckhoffs, *La cryptographie militaire*, Journal des sciences militaires, vol. IX, pp. 5–83, Jan. 1883.





Basic Terminology

We can characterize algorithms by:

- Number of keys used
 - single-key or secret or symmetric
 - two-key or public or asymmetric
- Encryption operations used to transform x in y
 - Substitution
 - Transposition
 - Product
- Way in which plaintext is processed
 - block
 - stream



Symmetric Cryptography

Plain-text input

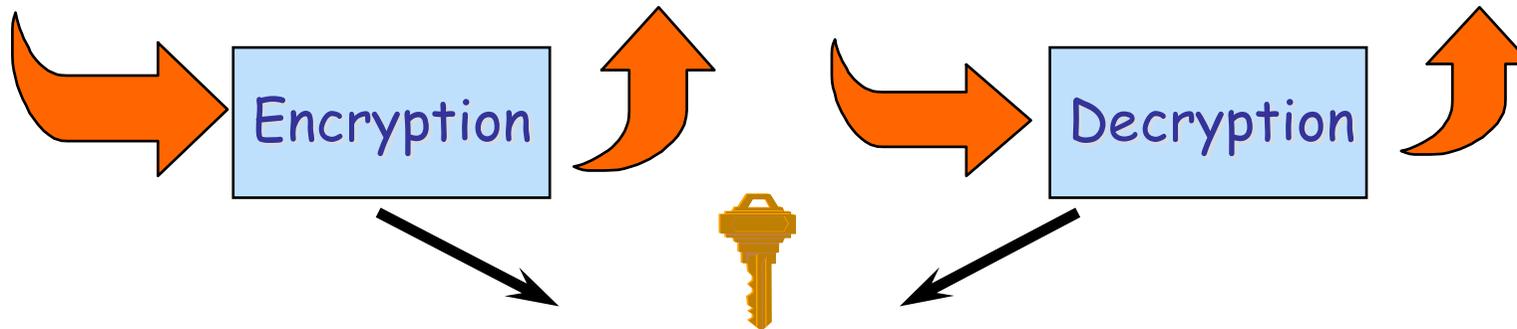
"The quick
brown fox
jumps over
the lazy dog"

Cipher-text

"AxCv;5bmEseTfid3)fGsmW
e#4^,sdgfMwir3:dkJeTsY8R
\s@!q3%"

Plain-text output

"The quick
brown fox
jumps over
the lazy dog"



Same key (shared secret), i.e $d=e$

- all classical encryption algorithms are private-key
- it was only type prior to invention of public-key in 1970's
- it implies a secure channel to distribute the key

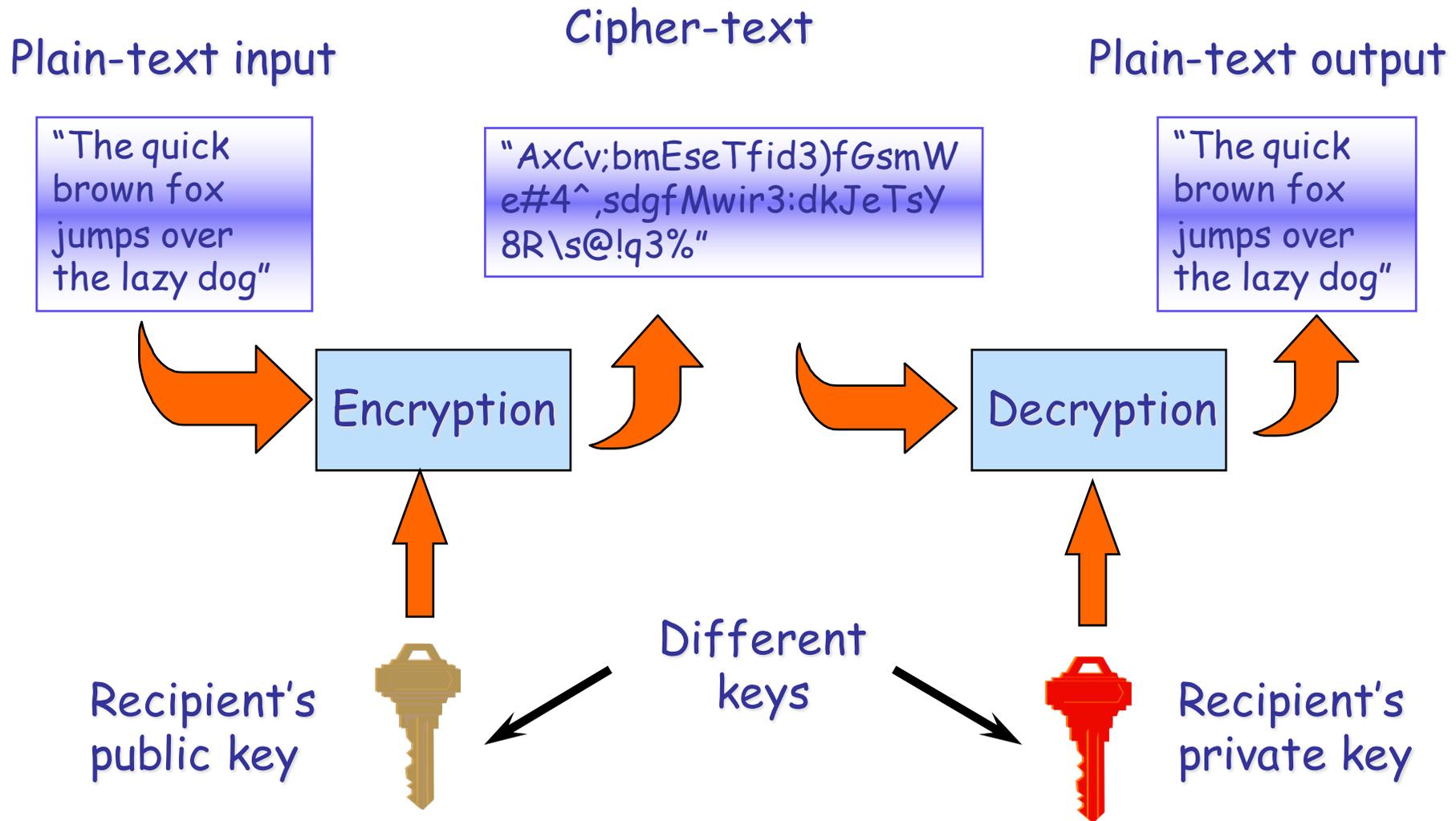


Example

- For instance, the key **e** is chosen to be the permutation which maps each letter to the one which is 3 positions to its right, as shown below:
- 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
- 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
- The key **d** is the inverse permutation that maps each letter to the one which is 3 positions on the left.



Asymmetric Cryptography





Asymmetric Cryptography

- Probably most significant advance in the 3000 year history of cryptography
- Uses **two** keys – a public & a private key
- **Asymmetric** since parties are **not** equal
- Based on number theory



Advantages of symmetric cryptography

- Very high data throughput (in HW achieves encrypt rates of ~ 100 MB/s, in SW MB/s).
- Relatively short keys
- Employed as primitive to construct various cryptographic mechanisms (e.g. PNRGs, hash functions, digital signature schemes)
- Can be composed to produce stronger ciphers.
- Sometimes symmetric-key encryption is perceived to be more secure because of its long history



Disadvantages of symmetric crypto

- Most disadvantages are linked to key management
 - In a two-party communication, the key must remain secret at both ends.
 - In a two-party communication between A and B, sound cryptographic practice dictates that the key be changed frequently, and perhaps for each communication session



Advantages of public-key cryptography

- Only the private key must be kept secret
 - authenticity of public keys must, however, be guaranteed
- A private key/public key pair may remain unchanged for long periods, e.g., many sessions
- In a large network, the number of keys may be considerably smaller than in the symmetric-key.



Symmetric vs. Asymmetric

- Number of keys needed for n users:
 - in public key: $2n$ keys (2 for each user);
 - in private key: $n*(n-1)/2$ keys (1 for each link between 2 users);

users	Keys in public-key system	Keys in private-key system
10	20	45
100	200	4.950
1.000	2.000	499.500
10.000	20.000	49.995.000
100.000	200.000	4.999.950.000
1.000.000	2.000.000	499.999.500.000



Disadvantages of public-key crypto

- Throughput is several orders of magnitude smaller than that of the best known symmetric-key schemes
- Key size is typically much larger than that required for symmetric-key encryption.
- No public-key scheme has ever been proven to be secure
 - The security of most schemes is based on the presumed difficulty of a small set of number-theoretic problems
- Public-key cryptography does not have an extensive history as symmetric-key (it was discovered in 1970's)



Cryptanalysis

- To attack a cryptographic algorithm, two kinds of attacks are possible:
 - Brute force search
 - Cryptanalysis



Brute Force Search

- (or Exhaustive key search), is the basic technique of trying every possible key in turn until the correct key is identified.
 - most basic attack, proportional to key size: we assume it needs to test one half of all keys
 - assume either *know / recognise* plaintext



Brute Force Search

- Consider the following example
 - I know that the key e is chosen to be a shift on the alphabet A (it maps each letter to the one which is n positions to its right).
 - I know the ciphertext $c = \text{WKL VFLSKHU}$
- I wish to find the plaintext (in this case **THISCIPHER**), and the value of the key e (that is a mapping of each letter to the one which is 3 positions to its right)



Brute Force Search

- We can find the key $d=e^{-1}$ by trying all possible shifts of each letter to the one which is n positions to its left, where $n = 1, 2, 3, \dots, 25$.
 - ciphertext $c = \text{WKLVFLSKHU}$
 - For $n=1$: $D_d(c) = \text{VJKUEKRJGT}$
 - For $n=2$: $D_d(c) = \text{UIJTDJQIFS}$
 - For $n=3$: $D_d(c) = \text{THISCIPHER}$
- Only for $n=3$ we obtain a meaningful message !!!
- The key is $e=3$



Brute Force Search

- 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

Key Size (bits)	Number of Alternative Keys	Time required at 1 encryption/ μ s	Time required at 10^6 encryptions/ μ s
32	$2^{32} = 4.3 \times 10^9$	$2^{31} \mu\text{s} = 35.8$ minutes	2.15 milliseconds
DES 56	$2^{56} = 7.2 \times 10^{16}$	$2^{55} \mu\text{s} = 1142$ years	10.01 hours
AES 128	$2^{128} = 3.4 \times 10^{38}$	$2^{127} \mu\text{s} = 5.4 \times 10^{24}$ years	5.4×10^{18} years
TripleDES 168	$2^{168} = 3.7 \times 10^{50}$	$2^{167} \mu\text{s} = 5.9 \times 10^{36}$ years	5.9×10^{30} years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu\text{s} = 6.4 \times 10^{12}$ years	6.4×10^6 years



Brute Force Search

Reference	Magnitude
Seconds in a year	$\approx 3 \times 10^7$
Age of our solar system (years)	$\approx 6 \times 10^9$
Seconds since creation of solar system	$\approx 2 \times 10^{17}$
Clock cycles per year, 50 MHz computer	$\approx 1.6 \times 10^{15}$
Binary strings of length 64	$2^{64} \approx 1.8 \times 10^{19}$
Binary strings of length 128	$2^{128} \approx 3.4 \times 10^{38}$
Binary strings of length 256	$2^{256} \approx 1.2 \times 10^{77}$
Number of 75-digit prime numbers	$\approx 5.2 \times 10^{72}$
Electrons in the universe	$\approx 8.37 \times 10^{77}$

Reference numbers comparing relative magnitudes.

An important question then is “How large is large?”.

In order to gain some perspective on the magnitude of numbers, the Table lists various items along with an associated magnitude.



Cryptanalysis

- It is the art (science) of cracking codes.
 - It exploits the characteristics of the cryptosystem, and, depending on the type of attack, the knowledge of the plaintext/ciphertext characteristics
 - **The aim is to find the key and, in some cases, the plaintext**
 - In order to design a robust encryption algorithm or cryptographic protocol, one should use cryptanalysis to find and correct any weaknesses



Types of Cryptanalytic Attacks

We assume that the algorithm and the ciphertext to be decrypted are known. Then we have:

- ciphertext only
- known plaintext
- chosen plaintext
- chosen ciphertext



Ciphertext only

- **Only ciphertext to be decrypted is known**
 - Easiest attack to cope with
 - The attacker has only access to some enciphered messages and tries to derive the key
 - Attacker does not try all the possible keys
 - probably they are too many...



Ciphertext only

- Without any knowledge about the plaintext, only statistical attacks can be used:
 - some hypothesis about the statistical distribution of the letters in the alphabet
 - some hypothesis about the header format (e.g. email begins with “dear ..”, or it is a .ps file)



Known plaintext

- **One or more pairs of plaintext & ciphertext known**
- The attacker knows (or strongly suspects) some plaintext-ciphertext pairs (blocks)
 - Suppose he/she knows that “bcmjhvbc” corresponds to the encryption of “caligula”
 - this knowledge can be used to decrypt the other parts of the cyphertext ...
 - ... or to discover the key used for encryption



Chosen plaintext

- **The opponent has obtained temporary access to the encryption machinery**
- Hence, he is able to **select a plaintext and to obtain the corresponding ciphertext**
 - the aim is to obtain the key used for encryption
 - the attacker uses knowledge of algorithm structure in attack to choose properly formatted plaintexts to be encrypted
 - This type of attack is generally most applicable to public-key cryptosystems



Chosen ciphertext

- **The opponent has obtained temporary access to the decryption machinery**
- **Then, he is able to select a ciphertext and obtain the corresponding plaintext** (with the exclusion of the one he wants to decrypt)
 - allows further knowledge of algorithm structure to be used for the attack
 - This type of attack is generally most applicable to public-key cryptosystems



Security definition

- One of the most important properties of a cryptographic system is a proof of security.
- However, every design involves a trade-off between the strength of the security and further important qualities of a cryptosystem, such as efficiency and practicality.
- The most popular security models currently used in cryptographic research include **computational security** and **unconditional security**.



Computational security

- It is based on the amount of computation required to break a system by the best currently known cryptanalytic method
- A proposed scheme is *computationally secure if the level of computation required to defeat it (using the best attack known) exceeds, by a comfortable margin, the computational resources of the hypothesized adversary*



Computational security

- Most of the currently used public-key cryptosystems (RSA, Diffie-Hellman) as well as private-key systems (DES, IDEA, RC5) fall into this category.
- In principle, all of them can be broken by trying the possible keys in sequence.
- But in practice, such attacks are considered to be unfeasible because they would take from months to millions of years on the fastest of today's computers.



Computational security

- Sometimes computational security can be proven
- A cryptographic method is provably secure if the difficulty of defeating it can be shown to be essentially as difficult as solving a well-known and supposedly difficult (typically number-theoretic) problem.
- Factoring large numbers is a typical example.
- It is believed that there is essentially no better way for solving these problems than trying all possible solutions in sequence, but nobody has found a proof for this yet.



Unconditional security

- It is based on information theory and imposes no limits on the adversary's computational power
- Unconditionally secure systems can not be broken even if all possible keys could be tried within short time
- The first definition of information-theoretic secrecy was given by Shannon
 - C.E Shannon, “Communication Theory of Secrecy Systems”, *Bell System Technical Journal*, vol. 28(4), pp. 656–715, 1949.



Unconditional security

- Cryptosystems are unconditionally secure against a ciphertext-only attack if plaintext X and ciphertext Y are statistically independent
 - The attacker has the same probability of obtaining X whether he knows Y or not
- This is equivalent to saying that the cryptanalyst can do no better than guessing the plaintext without knowledge of the encrypted data, no matter how much time and computing power is used



Unconditional security

- If we denote the *a priori* probability that plaintext x occurs by $p_M(x)$,
- A cryptosystem has perfect secrecy if $p_M(x|y) = p_M(x)$ for all $x \in \mathbf{M}$ and $y \in \mathbf{C}$
- This is equivalent to requiring that $I(X;Y) = 0$
- Shannon proved that unconditional security can be achieved only when the key-length is at least equal to the message entropy, which precludes the applications of these schemes in practice



References

- W. Stallings, *Cryptography and Network Security*, Mc Graw Hill, 4-th edition
- A.J. Menezes, P.C. van Oorschot, and S.A. Vanstone, *Handbook of Applied Cryptography*, CRC Press
- RSA Laboratories' FAQ About Today's Cryptography,
– <http://www.rsasecurity.com/rsalabs/node.asp?id=2152>