Departamento de ENGENHARIA INFORMÁTICA

Instituto Superior de Engenharia do Porto

# Systems and Information Security SEGSI

## Topic 3

## Data Security

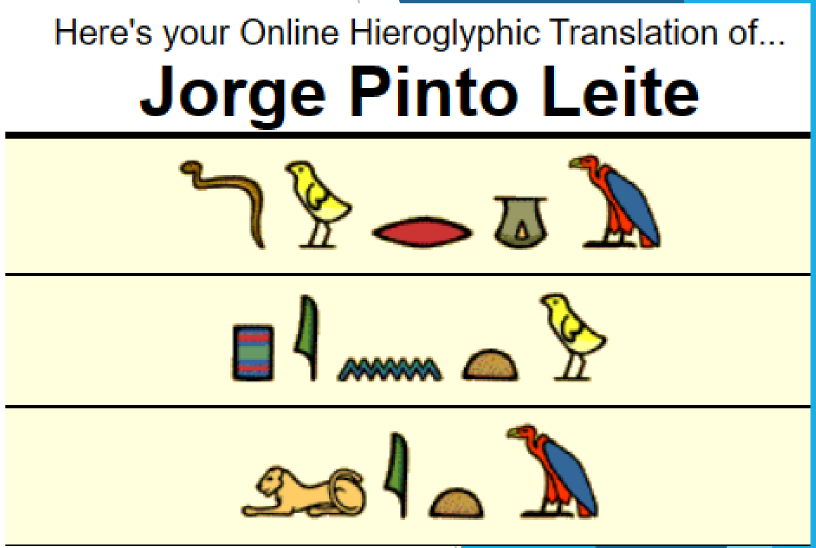
Pinto Leite, Jorge (jpl@isep.ipp.pt)

# Data Security

- Security of data at rest or in transit must be achieved due to ethical, social, political and legal concerns
- It can be achieved with the help of cryptography
- As a must, no loss of information can occur
- Cryptography is a very old practice with the intention of
  - Hide true intentions
  - Take advantage of direct competition
  - Reduce vulnerabilities (interpret here *vulnerabilities* as a broad meaning)

# Data Security

- Cryptography appeared almost 3000 years ago
- What is the difference nowadays?
  - Technology
  - Message transport
- The roots of cryptography began to emerge around 2000 BC
- Hebrews developed the *atbash* method



Here's your Online Hieroglyphic Translation of...

**Jorge Pinto Leite**

Source: http://www.quizland.com/hiero.htm

ABCDEFGHI JK LMNOPQ R STU VW XYZ

ZYXWVUTSR QP ONMLKJ I HGF ED CBA

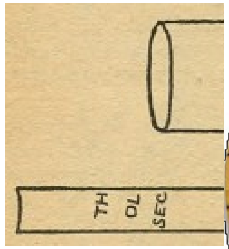
atbash

# Data Security

- These methods are called monoalphabetic
- A monoalphabetic cipher is a cipher where a character of the plain text is always mapped to the same character on the cipher text
- The monoalphabetic approach is simple, fast to encrypt and decrypt, but not applicable on modern times
- Polialphabetic cryptography became the more usual technique to apply cryptography

# Data Security

- Around 200 BC Spartans created a new method, *scytale*
- What is the parameter that works as the encryption key?
- The Roman Empire (27 BC – 476 AC) created a method (Ceaser Cipher) simila to scytale by applying a 3 letter offset to plain text
  - All “A” become “D” when encrypting, all “J” become “G” when decrypting
- It seems naive nowadays but note that it used the modulus operation that was stated hundreds years later
- During 1st World War but mainly used on 2nd, the Enigma machine was invented
  - 26,672,901,348,424,004,787,290,112 x 10^26 combinations
  - The secret reside in the initial configuration of the sprockets



Scytale (sourc

Source: cybersecuritynext.org
Source: app.emaze.com

# Data Security

- Cryptography has always been associated with important historical events
  - Mainly, military

- "*The Index of Coincidence and its Applications in Cryptography*" (William Friedman, 1922) is considered the beginning of modern cryptography

# Data Security

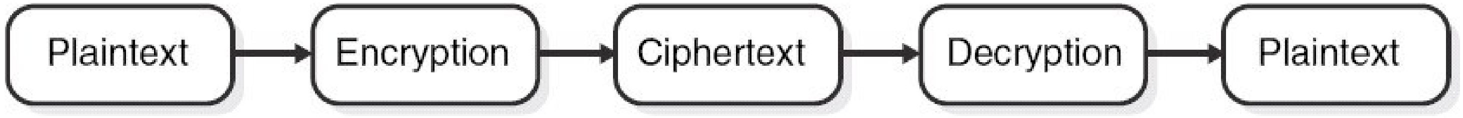
- At modern times, computers turned deciphering simpler and easier
- As response, there is an increased complexity of crypto-systems and new opportunities appeared to cryptographic system designers
- However, the more complex something is, more difficult it is to be confident that it has no flaws
  - ***There are two ways to design a system. One is to make it so simple there are obviously no deficiencies. The other is to make it so complex there are no obvious deficiencies.*** (C. A. R. Hoare, quoted in Kaufman et al., Network security, p. 441)

# Data Security

- Nowadays, almost everything uses encryption
- The bad guys get smarter and more resourceful, the good guys must increase efforts and strategy
  - Another instance of the game of "cat-and-mouse"
- Increasingly the effort required for good protection is greater
  - Good protection tends to last less and less ...

# Data Security – definitions

- **Crypto-Analysis** – study and discover the secret of cipher algorithms

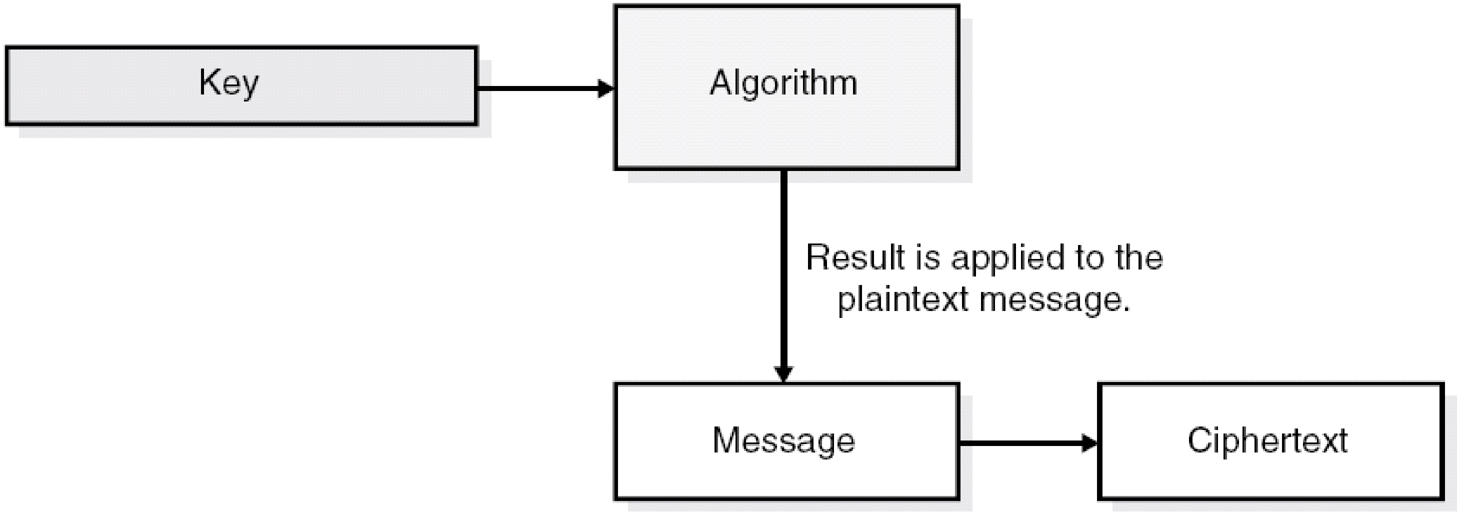


The process of encryption transforms plaintext into ciphertext and the process of decryption transforms ciphertext into plaintext.

- **Crypto-System** – a system that provides encryption and decryption capabilities in Hardware and / or Software
- Crypto-Systems use algorithms, more or less complex and
- Applying mathematical formulas in a certain sequence to the normal text (plaintext P)
- Usually they use a secret value called a Key (K) that is used together with the algorithm to encrypt and decipher information
    - It works as a parameter of the algorithm

# Data Security – definitions

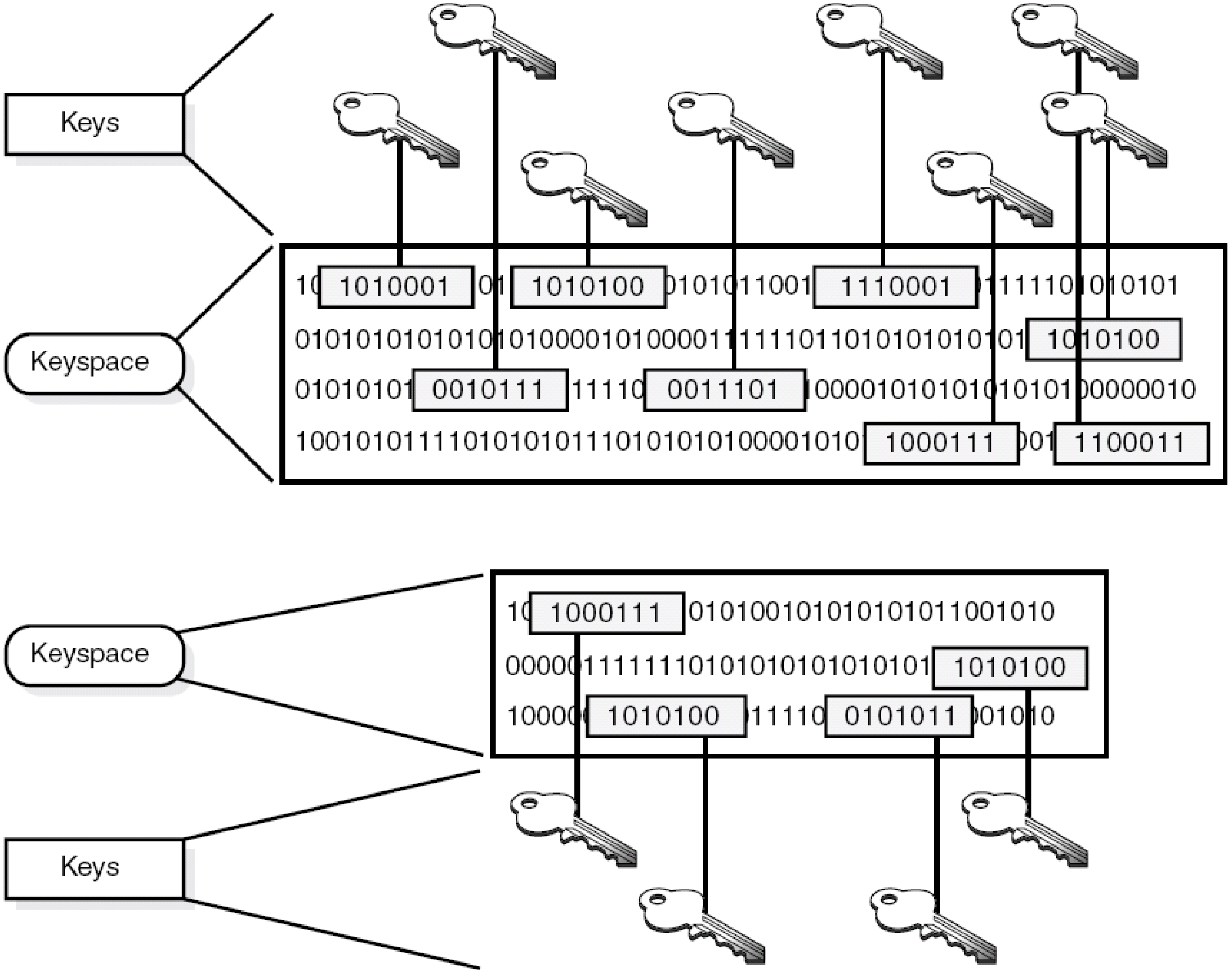
- **Algorithm**
  - Dictates the rules / sequence / mathematical formulas used
    - Are mostly public
    - It should not be the secret!



# Data Security – definitions

- The **Key** is the secret!
  - It has a size according to the possibilities space of the key (*keyspace*)
  - The higher the keyspace
    - More values can be represented
    - More random may be the keys
  - Examples
    - Keyspace with 5 vowels
    - Keyspace with 8 letters
    - Keyspace with 16 ASCII symbols (7 bits)
    - 1024-bit key keyspace

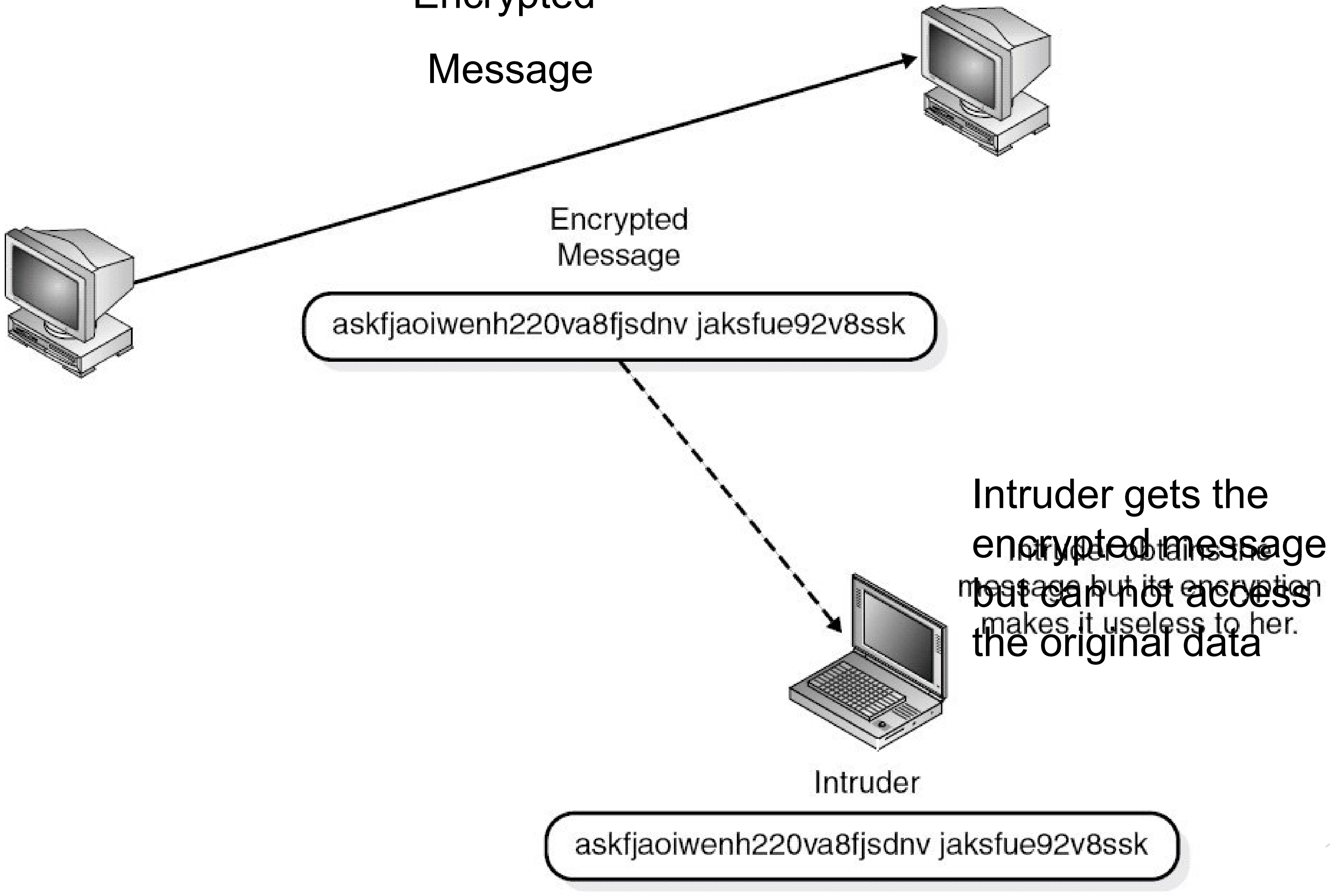
# Data Security – definitions



# Data Security

Encrypted

Message


Encrypted
Message
askfjaoiwenh220va8fjsdnv jaksfue92v8ssk
Intruder obtains the message but its encryption makes it useless to her.
Intruder
askfjaoiwenh220va8fjsdnv jaksfue92v8ssk

Intruder gets the encrypted message but can not access the original data

# Data Security – attacks

- Ciphertext-Only Attack
  - Attacker takes advantage of encrypted messages
  - Search for common parts in order to discover the secret key
- Know-Plaintext Attack
  - Attacker knows clear text and ciphertext (but not knowing if they are related)
  - Try to discover the secret key

# Data Security – attacks

- Chosen-Plaintext Attack
  - Attacker knows the plaintext and gets its ciphertext (knows that they are related)
- Chosen-Ciphertext Attack
  - Similar to the previous one, but choosing the ciphertext and getting the respective text in clear

# Data Security – attacks

- Man-in-the-Middle Attack
  - Steps
    1. X sends its public key to Z
    2. Y intercepts the public key of X and sends its own public key to Z
    3. Z sends its public key to X
    4. Y intercepts the public key of Z and sends its own public key to X
    5. From here Y managed to elude the two players and can see all messages between X and Z

# Data Security – attacks

- There are other types of attacks
  - Dictionary Attacks
  - Replay Attacks
  - Side Channel Attacks

# Data Security – strength

- The strength of a crypto-system comes from
  - Algorithm
  - Secrecy of the key
  - Key size
  - Boot Vectors
  - Combination of all previous parameters
- Is measured by the time required to break the key (get that key that matters)
  - *Brute Force Attack*
- It is crucial to protect the key

# Data Security – purpose

- Purpose of a crypto-system is to make it too expensive or too slow to break it
- Provides (or can provide)
  - Confidentiality
  - Authenticity
  - Integrity
  - Irrefutability (this is new)
    - Ability to avoid being denied authorship of an action actually taken
    - What entities need repudiation much?

# Data Security – types of cipher

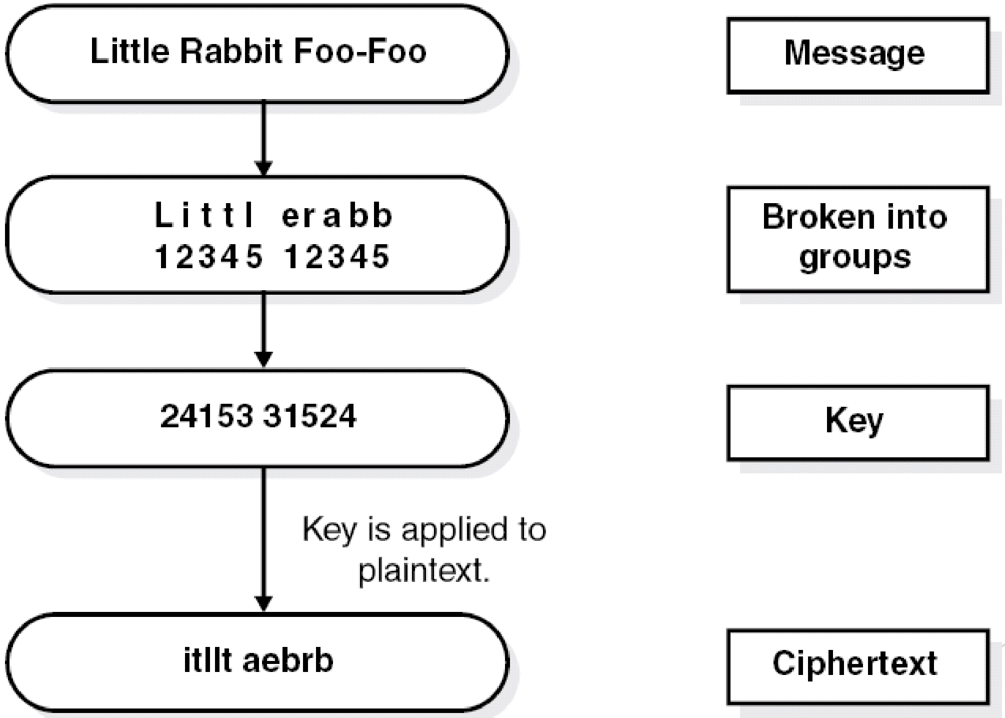
- Replacement Ciphers
- Transposition Ciphers
- Running keys
- Concealment Ciphers
- Steganography

# Data Security – replacement cipher

- Replace
  - Bits
  - Characters
  - Blocks
- Use key to calculate substitution
- Most known example
  - Cipher of the Roman Empire
- Still used today but much more complex

# Data Security – transposition cipher

- No substitution (symbols remain)
- Reorganizes the original text in order to hide the true meaning
- Example



# Data Security – running key cipher

- Coming from the world of spies
- Does not use electronic algorithms
- Depends on the physical world
- Example:
  - “1.49l6c7.2.99l3c7.9.11l5c8…”
  - 1º symbol: book 1, page 49, line 6, column 7
  - 2º symbol: book 2, page 99, line 3, column 7
  - 3º symbol: book 9, page 11, line 5, column 8 4º symbol: …

# Data Security – concealment cipher

- Also from the world of spies
- Message has an apparent meaning
- The trick is to choose only a few words to define a second message - the true one
- Examples:
  - All three words in each sentence
  - Every second word on every page
  - All the initial letters of the third sentence …
    - *Will eventually all risks ever gone or is not good to opt, mandatory as key event, a time elapsed so true* (poem from unknown author)

# Data Security – steganography

- Method that hides the data in another type of message medium
- In this way the data is "hidden"
- Usually images or digital audio / video
- Examples:
  - Least significant bit of each byte of an image is replaced by the bit of the message
  - Message bits inserted in an inaudible zone of an audio message

# Data Security – cipher methods

- Symmetrical
  - Both parties use the same key to encrypt and decrypt messages
  - A secret key is required to each pair of users
    - With N people, necessary keys is N(N – 1)/2
  - All security lies on the key
  - Does it provides confidentiality?
    - It depends!
    - How is the secret (key) shared?
  - Does it provides authenticity?
    - It depends!
    - Is it possible to guess the secret and hasn't it been shared with someone else?
- As an advantage, the speed of performance and the difficulty of be broken if well used

# Data Security – cipher methods

- Asymmetrical
  - There is a pair of keys
    - Private key / Public key
  - Keys are mathematically linked
  - The public key is known to all
  - The private key must be kept secret
  - From the public key it should not be possible to calculate the respective private key
  - From the private key it is simple to obtain the respective public key

# Data Security – cipher methods

- Asymmetrical
  - Encrypted messages with the private key can be decrypted with the public key
  - Encrypted messages with the public key can be decrypted with the private key
  - A message encrypted with a public/private key can only be deciphered with its related key
  - Provides
    - Confidentiality
      - Only the actors understand the message  and get guarantees
    - Authenticity
      - Each actor is authentic to the other
      - Irrefutability (non-repudiation)
      - Each actor can not deny the message

# Data Security – cipher methods

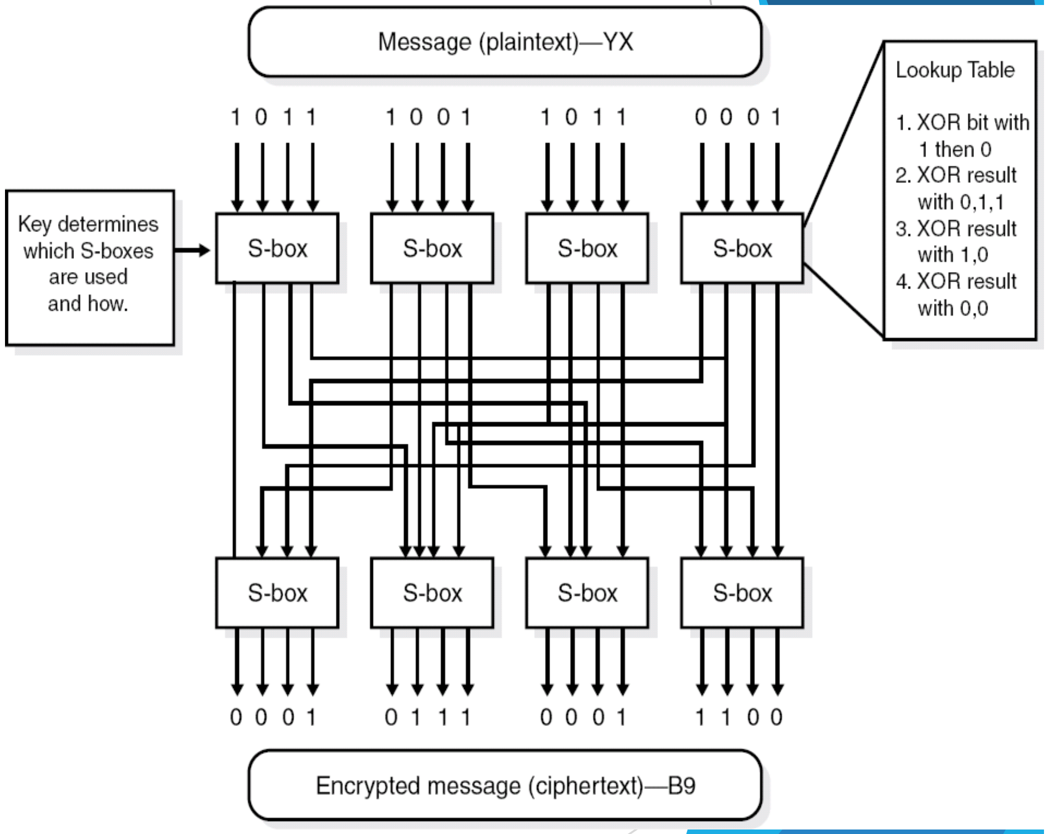
- Asymmetric cipher pros
  - Turns key distribution easy (in fact, there is no distribution, only the publication of the public key)
  - Is naturally scalable
  - Can provide authenticity and irrefutability
- And cons
  - Slow operation
    - More complex mathematical processing

# Data Security – symmetric algorithm types

- Block
- Stream
- Hybrid

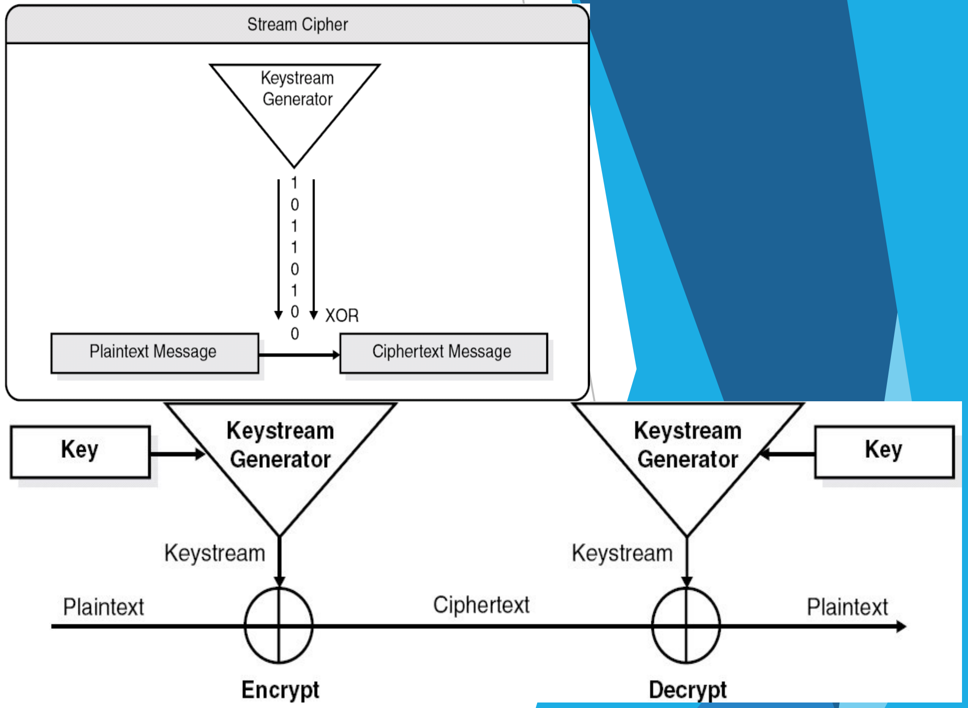
# Data Security – block type

- Goal
  - High diffusion
  - Immune to malicious injection
  - High error propagation

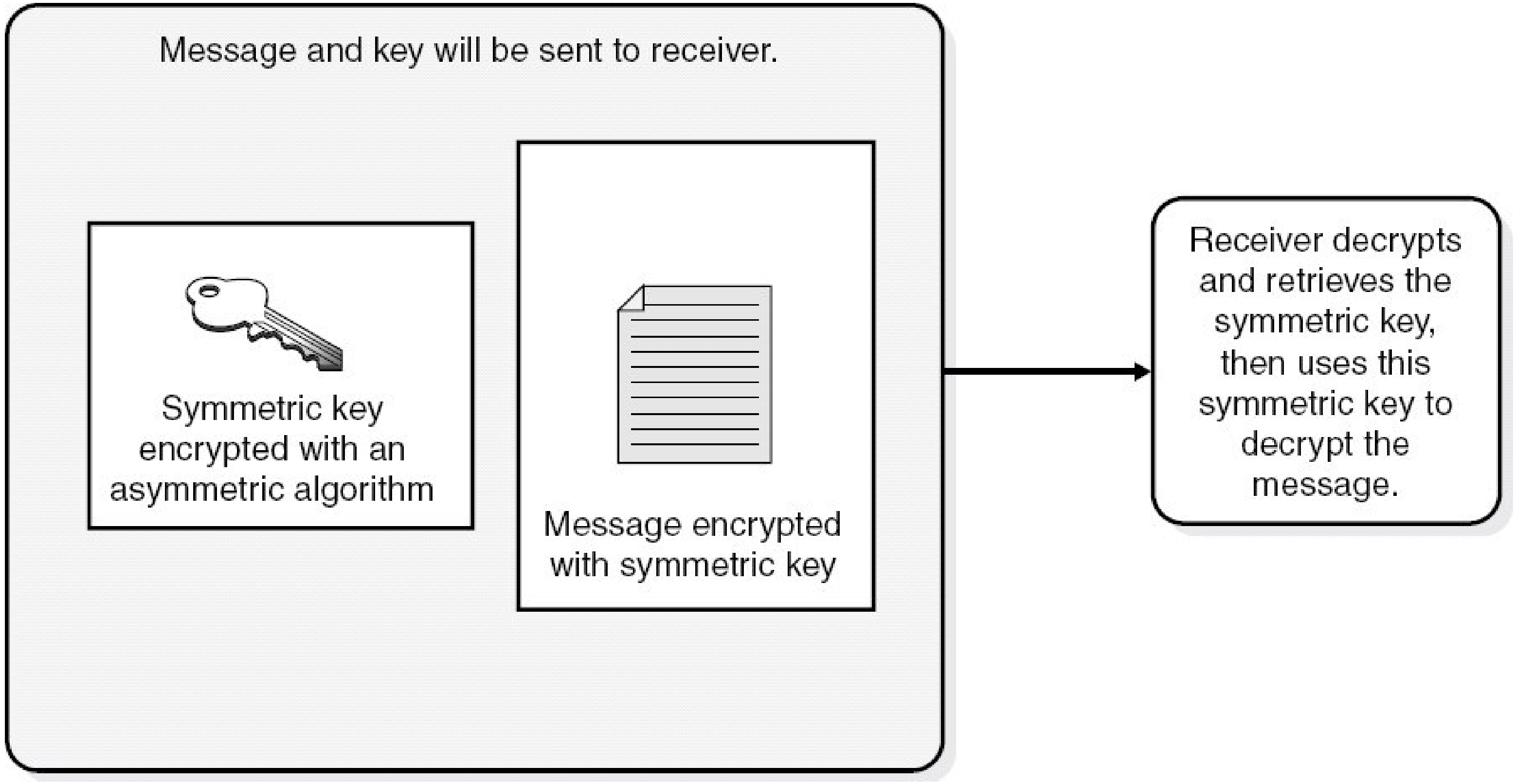


# Data Security – stream type

- Process each bit individually
  - Low diffusion
  - Susceptible to malicious injection
  - Low error propagation



# Data Security – hybrid type



Message and key will be sent to receiver.

Symmetric key encrypted with an asymmetric algorithm

Message encrypted with symmetric key

Receiver decrypts and retrieves the symmetric key, then uses this symmetric key to decrypt the message.

# Data Security

- On previous slides confidentiality was discussed
  - A two-way or bidirectional algorithms
- But that’s not the only concern, data integrity is also an important objective
  - For that purpose, one-way or unidirectional algorithms (hash functions) are used

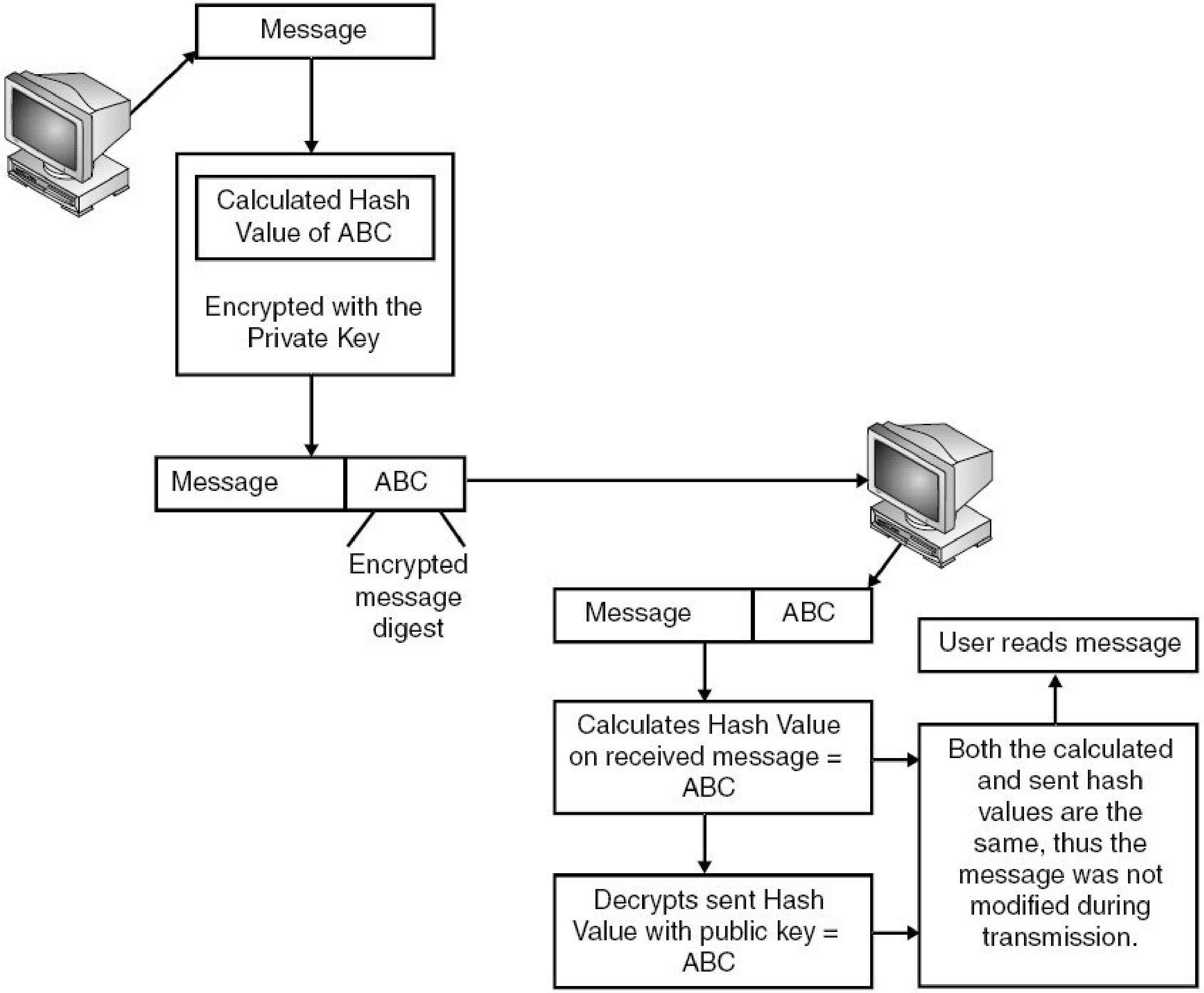
# Data Security

- Unidirectional algorithms
  - Receives strings of variable length
  - Produces fixed-length values (hash, H)
    - They represent the original data
    - Are like fingerprints

  - The data at rest or sent is in fact
    - Encrypt:
      - E(data+H(data),$K^+$)
    - Decrypt:
      - D(data+H(data),$K^-$) = data+H(data) (and the later can be checked against the calculated one)

# Data Security

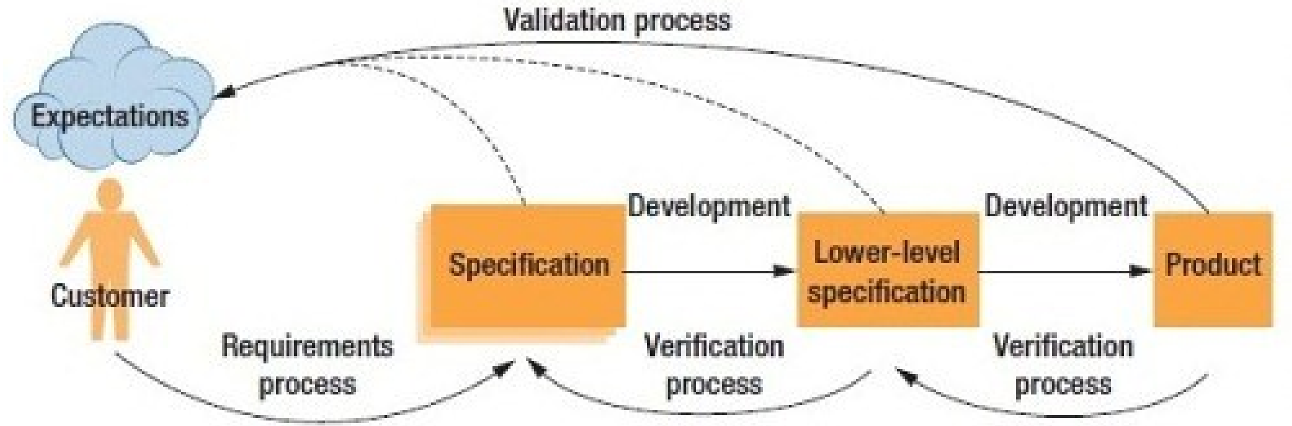
- With asymmetrical encryption irrefutability can be used
  - As with digital signature



Message
Calculated Hash Value of ABC
Encrypted with the Private Key
Message
ABC
Encrypted message digest
Message
ABC
Calculates Hash Value on received message = ABC
Decrypts sent Hash Value with public key = ABC
Both the calculated and sent hash values are the same, thus the message was not modified during transmission.
User reads message

# Data Security

- There are additional aspects about data security like
  - Security on applications programming
  - Application faults
- These topics are addressed on other classes of MEI so they are not covered here
- However, keep in mind the V&V (Verification and Validation) mindset



Source: Michael, J.B., Drusinsky, D., Otani, T.W., Shing, M. - *Verification and Validation for Trustworthy Software Systems*, 2012, IEEE Software