CSIT 495/595 - Introduction to Cryptography

Introduction

Bharath K. Samanthula
Department of Computer Science
Montclair State University

Need to Learn Cryptography

- Data (or information) is the most crucial thing for many companies and government agencies
- Strong need to protect data while at rest, being transmitted and shared with others
- Cryptography offers very useful tools to protect data
- Examples:
  - All your emails on Gmail are encrypted by Google
  - Where is your password stored and how it is your protected?

What is Cryptography?

- Cryptography is the art of writing or solving codes
- Cryptography deals with *the scientific study of techniques for securing digital information, transactions, and distributed computations* (Textbook)

Importance of Cryptography

- Historically, major consumers of cryptography are military and intelligence organizations
  - Breaking Japanese Naval code in the middle of the second world war
  - Breaking Enigma machine
- Breaking Blue-ray and DVD encryption
- Today, cryptography is everywhere!!
  - Security mechanisms that rely on cryptography are an integral part of almost any computer system

Importance of Cryptography

"Windtalkers" (2002) - about the secret code used by American military during world war II which baffled the Japanese.

"Breaking the Code" (1996) - about British mathematician, Alan Turing, "Father of computer science", who worked on breaking German military code during World War II. http://www.turing.org.uk/turing/

"A Beautiful Mind" (2001) - Oscar-winning movie about US mathematician and Nobel laureate John Nash. Nash worked for the US military on secret codes.

Others: **"U-571" (2000)**, **"Swordfish" (2001)**, **"Enigma" (2001)**, **"Hackers" (1995)**

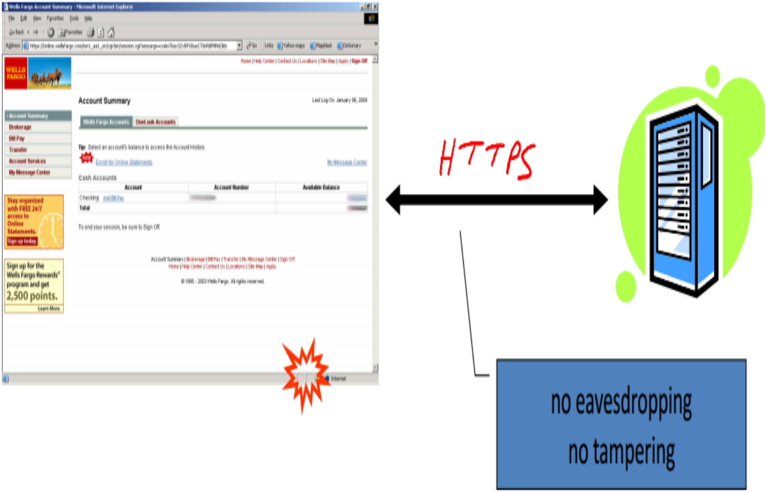
Check www.imdb.com for details.

Source: Cryptography and Network Security, Behrouz Forouzan

Some well-known Applications

- Secure communication (e.g., HTTPS)
    - Email, e-commerce, ATM machines or cellular phones
- Encrypting files on Disk (EFS, TrueCrypt)
- Content Protection (e.g., Blue-Ray, DVD): CSS, AACS
- User Authentication (e.g., verifying user password/identity)
- .... and many other electronic applications over Internet

Secure Communication



Data Confidentiality

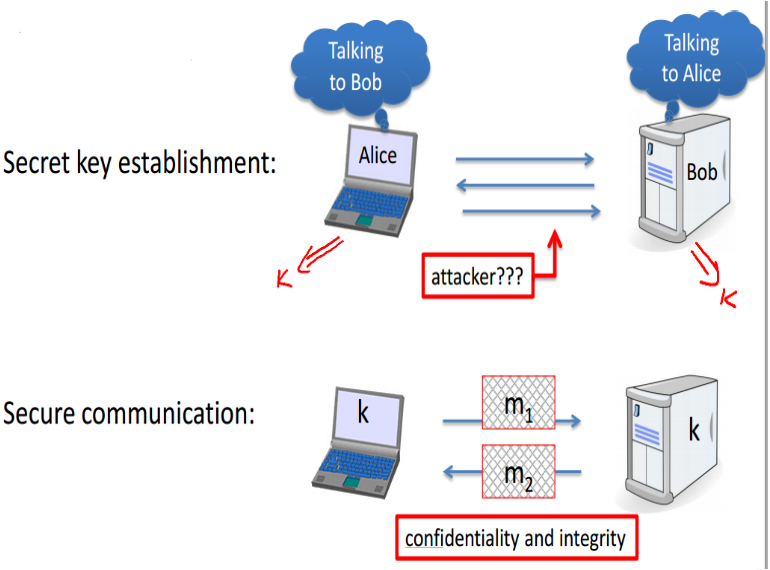
Objectives

- Ensure data is accessible only to intended parties
- Data, whether at rest or being transmitted, need to remain confidential
- Often related to privacy

How this is achieved

- Typically achieved by encryption the data
- Example: user passwords and bio-metric data need to be protected from other users/attackers

Data Confidentiality



Source: Cryptography, Dan Boneh

Data Integrity

Objectives

- Maintain the accuracy and trustworthiness of data
- Data at rest or in transit should not be changed by unauthorized people (e.g., in the case of hacking)

How this is achieved

- Checksums or digital signatures
- Example: when accessing your bank account details, how can the user ensure that his/her account information is accurate and trustworthy?

Data Authentication

- Only authenticated people can receive/send messages in communication
- User has to show proper identity to be verified by the verifier - data authentication allows a receiver to verify that the data really was sent by the claimed sender
- Examples: PIN, password, biometrics
- Key question: Can the user authenticate him/herself to the server without revealing his/her identity?
    - YES!!! using proper cryptographic techniques

Private-Key Encryption: Introduction

- Classical cryptosystems are mostly based on private-key encryption where the security depends on a secret, commonly called key
- General setting: two parties sharing a key can communicate and exchange messages using the key
- The eavesdropper can monitor all the communication between the two parties
- Basic Steps:
  - Party 1 encrypts a message (known as plaintext) using the shared key
  - Party 1 sends the encrypted message (known as ciphertext) to party 2
  - Party 2 decrypts ciphertext to get the actual message
- Also referred to as symmetric-key setting

Private-Key Encryption: Applications

Secure Communication

- Two different parties separated in space (e.g., a CIA agent in USA communicating with his colleague in Europe)
- Parties can exchange messages after establishing key
- Key Question: How the parties can share the key securely?

Disk Encryption

- Same party communicating with itself over time
- User encrypts his files and stores them on Disk
- At a later time, he uses the same key to decrypt the files using the same key
- Key Question: How the user can remember/store the key securely and reliably?

Private-Key Encryption: Formal Definition

- Encryption schemes are defined over the message space $\mathcal{M}$, the key space $\mathcal{K}$, and the ciphertext $\mathcal{C}$

Key Generation (Gen)

- A probabilistic algorithm that selects a key $k \in \mathcal{K}$

Encryption (Enc)

- Input: message $m \in \mathcal{M}$ and k
- Output: Ciphertext $c \leftarrow \mathrm{Enc}_k(m)$, where $c \in \mathcal{C}$

Decryption (Dec)

- Input: c and k
- Output: $m \leftarrow \mathrm{Dec}_k(c)$

Kerckhoff's Principle (1)

- An adversary (or eavesdropper) can decrypt a ciphertext if he/she knows the decryption function (Dec) and the key used (k)
- Therefore, k has to be shared securely between the participating parties
- Do we also need to protect Dec from the adversary?
- Hide all the information related to the encryption scheme - Is this a better idea??

Kerckhoff's Principle (2)

- Auguste Kerckhoff, a Dutch Cryptographer, argues the opposite in a paper he wrote in the late 19th century

The cipher method must not be required to be secret, and it must be able to fall into hands of the enemy without inconvenience

- Commonly known as Kerckhoff's Principle
- Encryption should be designed to be secure even if all the details of the encryption scheme are revealed to the adversary, except the key
- In short, security relies solely on the secrecy of the key

Kerckhoff's Principle (3)

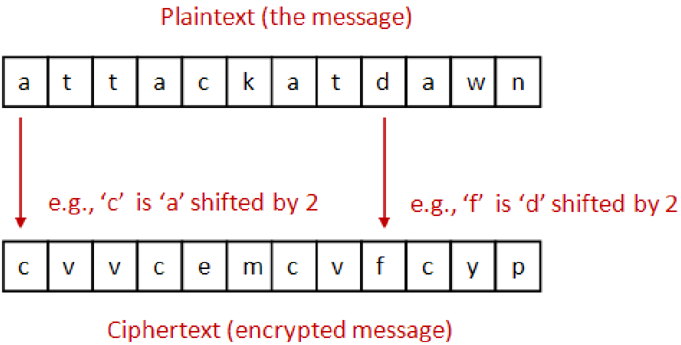
- But.. Why this is correct?
- Three reasons:
  1. It is often unrealistic to assume encryption/decryption function will remain secret
  2. Easier to replace a key than to replace the entire encryption system
  3. Feasible for large-scale deployment: it is better to use same encryption software/algorithm + users might use different keys

Caesar’s Cipher

- Named after Julius Caesar, who described it in 110 CE
- He encrypted the messages by shifting each alphabet 3 places forward

Alphabet shifted by 3 spaces.

Caesar’s Cipher - Example



- Note: punctuation, spaces, and numbers are removed
- What is the ciphertext when Caesar’s cipher is used?

Caesar's Cipher - Security

- Encryption method is fixed and there is no key
- It is very easy for anyone to decrypt ciphertexts
- ROT-13:
  - A variant of Caesar's cipher with 13 places shift instead of 3
  - Still used in various online forums

Shift Cipher

- A keyed variant of Caesar's cipher
- Key k is a number between 0 and 25
- Encryption: letters in plaintext are shifted by k places forward
- Decryption: letters in ciphertext are shifted by k places backward

Shift Cipher - Formal Definition

- Suppose English alphabet set is mapped to the set $\{0, \ldots, 25\}$ (i.e., $a = 0, b = 1$, and so on)
- Given a key k, encryption of a message $m = m_1 \cdots m_\ell$ is

$$\mathrm{Enc}_k(m) = c_1 \cdots c_\ell, \text{ where } c_i = (m_i + k) \bmod 26$$

- Decryption function is given by

$$\mathrm{Dec}_k(c_1 \cdots c_\ell) = m_1 \cdots m_\ell, \text{ where } m_i = (c_i - k) \bmod 26$$

Shift Cipher - Example

Let $k = 20$

Original	M	A	T	H	R	U	L	E	S
Number-fied	12	0	19	7	17	20	11	4	18
+key	32	20	39	27	37	40	31	24	38
mod 26	6	20	13	1	11	14	5	24	12
Letter-fied	G	U	N	B	L	O	F	Y	M

Shift Cipher - Attack Scenarios

- Is it possible to decrypt a ciphertext without knowing k?
- YES!! It is trivial
- Brute-force or exhaustive-search attack:
  - There are only 26 possible keys
  - Given a ciphertext, use every possible key and compute 26 possible plaintexts (one of them certainly being the correct one)
- In general, when the ciphertext is long enough, then the only one on the list that "makes sense" is the correct plaintext

Key Space - Important Observations

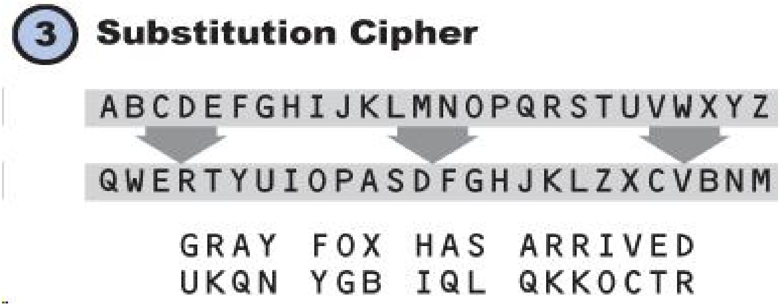
- An encryption scheme should not be vulnerable to brute-force attacks
- sufficient key-space principle: Any secure encryption scheme must have a key space that is sufficiently large to avoid brute-force attacks in-feasible
- Example: key size of at least 2^{80}
- Nevertheless, the sufficient key space is a necessary condition, but not the only requirement, to ensure proper security

Substitution Ciphers

- Mono-alphabetic
- Poly-alphabetic

Mono-alphabetic Substitution Cipher (1)

- The key defines a fixed substitution for individual letters in the plaintext
- Each letter is mapped to one of the remaining letters
- key space: consists of all the bijections or permutations



Mono-alphabetic Substitution Cipher (2)

- The key space is of size 26! ($\approx 2^{88}$)
- Brute-force attack becomes difficult
- Does it mean this cipher is secure? ... NO
- It is still easy to break this scheme using frequency cryptanalysis - the frequency of encrypted characters in the ciphertext is used to derive the plaintext

Mono-alphabetic Substitution Cipher (3)

- The attack can be carried based on the following facts (assuming English language text)
  1. Given any key, the mapping for each letter is fixed. For example, if a is mapped to q, then every a in the plaintext is mapped to q
  2. The frequency distribution of individual letters in English is known
- Count the frequencies of each letter in the ciphertext and compare to the English letter frequency distribution
- Additional knowledge can also be used (e.g., h is likely to appear between t and e)

Poly-alphabetic Substitution Cipher

- Key is defined on blocks of plaintext characters
- Example: for plaintext *abac* and 2-character block, the ciphertext might be *DZTY*
- Here we have two blocks *ab* and *ac*
- Observe that *a* is not mapped to a fixed ciphertext character

Vigenere Cipher

- A poly-alphabetic shift cipher
- Applies several independent instances of shift cipher in sequence
- Example 1:
  Plaintext: attackatdawn
  Key: lemonlemonle
  Ciphertext: lxfopvefrnhr
- Example 2:
  Plaintext: tellhimaboutme
  Key: cafecafecafeca
  CIphertext: ??

How can we define security

Two Important components:

- Security Guarantee: Defines what the scheme is intended to protect from the attacker
- Threat Model: Defines the attacker's capability

Security Guarantee - Some Thoughts

What is a good security guarantee??

- Impossible for an attacker to recover the key
- Impossible for an attacker to recover the entire plaintext
- Impossible for an attacker to recover any character of the plaintext
- Right Answer: Attacker should not know any information about the plaintext other than what he has already known

Threat Models

In order of increasing power of attacker:

- Ciphertext-only attack
- Known-plaintext attack
- Chosen-plaintext attack
- Chosen-ciphertext attack

Some Tips

- Never use your own or non-standardized crypto algorithms
- Understand the application requirements and use the existing schemes that are mathematically strong, provably secure and widely used

Summary

- Importance of Cryptography
- Data Confidentiality + Integrity + Authentication
- Kerckhoff’s principle and its broad impact
- Historical Ciphers and their cryptanalysis
- Security goals and threat models

Useful References

- Chapter 1, Introduction to Modern Cryptography by Jonathan Katz and Yehuda Lindell, 2nd Edition, CRC Press, 2015.
- `http://cseweb.ucsd.edu/~mihir/crypto-links.html`