# Introduction to Cryptography and Data Security

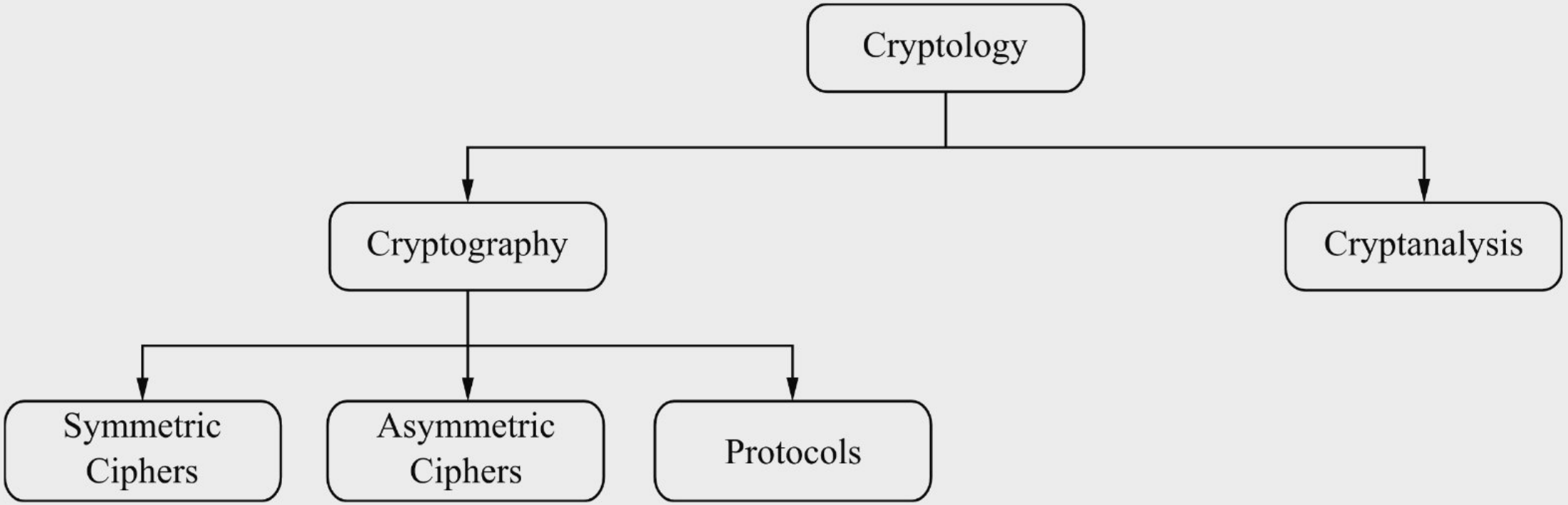
IT3122 Computer Security

2025.08.29

# Intended Learning Outcomes

- Define the basic concepts of cryptography.
- Present an overview of the main concepts of symmetric cryptography.
- Explain the difference between cryptanalysis and brute-force attack.
- Understand the operation of a monoalphabetic substitution cipher.

# Overview of the Field of Cryptology



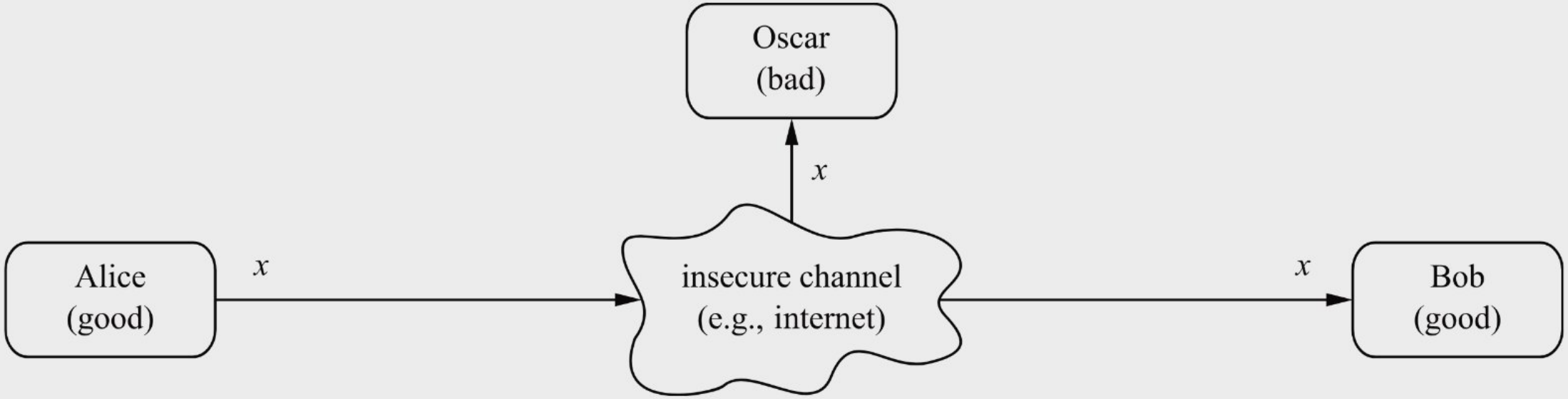
# Cryptology

1. **Cryptography** is the science of securing communication against an adversary.
2. **Cryptanalysis** is the science and sometimes art of **breaking** cryptosystems.

# Cryptography

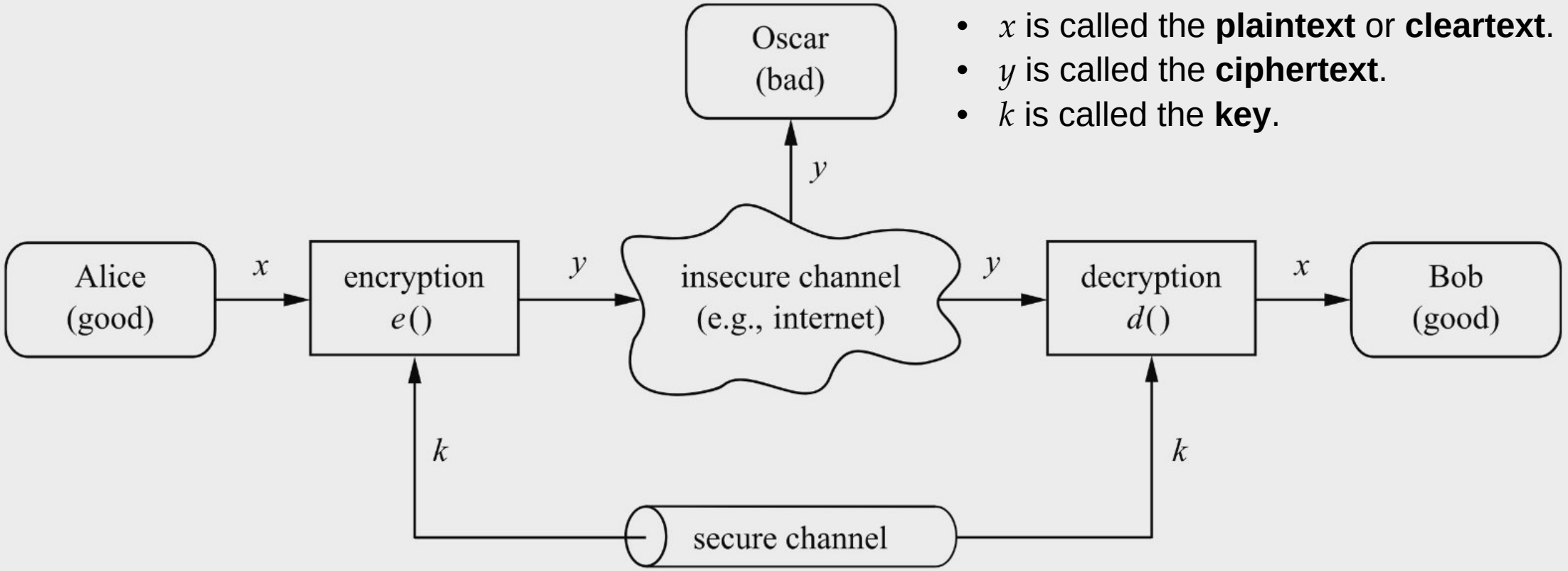
1. **Symmetric Algorithms**: All cryptography from ancient times until 1976 was exclusively based on symmetric methods.
2. **Asymmetric (Public-Key) Algorithms**: In 1976 asymmetric cipher was introduced by Whitfield Diffie, Martin Hellman and Ralph Merkle.
3. **Cryptographic Protocols**: Cryptographic protocols realize more complex security functions through the use of cryptographic algorithms. E.g., Transport Layer Security (TLS) scheme.

# Communication over an Insecure Channel




1. Alice and Bob want to communicate over an insecure **channel**.
2. The bad guy, Oscar has access to the channel.

# Symmetric-key Cryptosystem




- $x$ is called the **plaintext** or **cleartext**.
- $y$ is called the **ciphertext**.
- $k$ is called the **key**.

- Oscar obtains the ciphertext that will look like random bits.

# Symmetric Cipher Model

- Encryption equation

$$y = e_k(x) \text{ or } Y = \mathrm{E}(K,X)$$

- Decryption equation

$$x = d_k(y) \text{ or } X = \mathrm{D}(K,Y)$$

- Encryption and decryption are inverse operations if the same key $k$ is used on both sides:

$$d_k(y) = d_k(e_k(x)) = x$$

- **The problem of transmitting a message securely is reduced to the problems of transmitting a key secretly and of storing the key in a secure fashion.**

# Monoalphabetic Substitution Ciphers

- **Substitution (= replacement)** cipher.
- Substitute each letter of the alphabet with another one.

| Plaintext | Ciphertext |
|---|---|
| A | → k |
| B | → d |
| C | → w |
| ... | |

- E.g., `BABA` would be encrypted as `dkdk`.

# Example

```
iq ifcc vqqr fb rdq vfllcq na rdq cfjwhwz hr
bnnb hcc hwwhbsqvqbre hwq vhlq
```

- This does not seem to make too much sense and looks like decent cryptography.
- **However, the substitution cipher is not secure at all!**

# Basic Exhaustive Key Search or Brute-Force Attack

Let $(x,y)$ denote the pair of plaintext and ciphertext, and let $K = \{k_1, \ldots, k_\kappa\}$ be the key space of all possible keys $k_i$. A brute-force attack now checks for every $k_i \in K$ whether

$$d_{k_i}(y) \stackrel{?}{=} x.$$

If the equality holds, a possible correct key is found; if not, proceed with the next key.

# First Attack: Brute-Force Attack or Exhaustive Key Search

- Key space of the substitution cipher

$$n(K) = 26 \cdot 25 \cdots 3 \cdot 2 \cdot 1 = 26! \approx 2^{88}$$

- The key space has roughly a size of $2^{88}$, which is equal to the key space of a cipher that has a key consisting of $88$ bits.
- Even with hundreds of thousands of high-end PCs such a search would take several decades!

# Question

Can we now conclude that the substitution cipher is secure since a brute-force attack is not feasible?

# Answer

Can we now conclude that the substitution cipher is secure since a brute-force attack is not feasible?

No! We have to protect against **all** possible attacks...

# Letter Frequency Analysis

1. Determine the frequency of every ciphertext letter. E.g., in English E is the most frequent letter (about 13%), T is the second most frequent letter (about 9%), A is the third most frequent letter (about 8%), and so on.
2. Look at pairs or triples, or quadruples, and so on of ciphertext symbols. E.g., in English the letter Q is almost always followed by a U.
3. If we assume that word separators, which means “blanks”, have been found, one can often detect frequent short words such as THE, AND, etc.

# Relative Letter Frequencies of the English Language

| Letter | Frequency | Letter | Frequency |
|---|---|---|---|
| A | 0.0817 | N | 0.0675 |
| B | 0.0150 | O | 0.0751 |
| C | 0.0278 | P | 0.0193 |
| D | 0.0425 | Q | 0.0010 |
| E | 0.1270 | R | 0.0599 |
| F | 0.0223 | S | 0.0633 |
| G | 0.0202 | T | 0.0906 |
| H | 0.0609 | U | 0.0276 |
| I | 0.0697 | V | 0.0098 |
| J | 0.0015 | W | 0.0236 |
| K | 0.0077 | X | 0.0015 |
| L | 0.0403 | Y | 0.0197 |
| M | 0.0241 | Z | 0.0007 |

# Second Attack: Letter Frequency Analysis

- The letter `q` occurs most frequently in the text.

  i**q** ifcc v**qq**r fb rd**q** vfllc**q** na rd**q** cfjwhwz hr bnnb hcc hwwhbs**q**v**q**bre hw**q** vhl**q**

- `q` must be the substitution for one of the frequent letters in the English language.

  i**E** ifcc v**EE**r fb rd**E** vfllc**E** na rd**E** cfjwhwz hr bnnb hcc hwwhbs**E**v**E**bre hw**E** vhl**E**

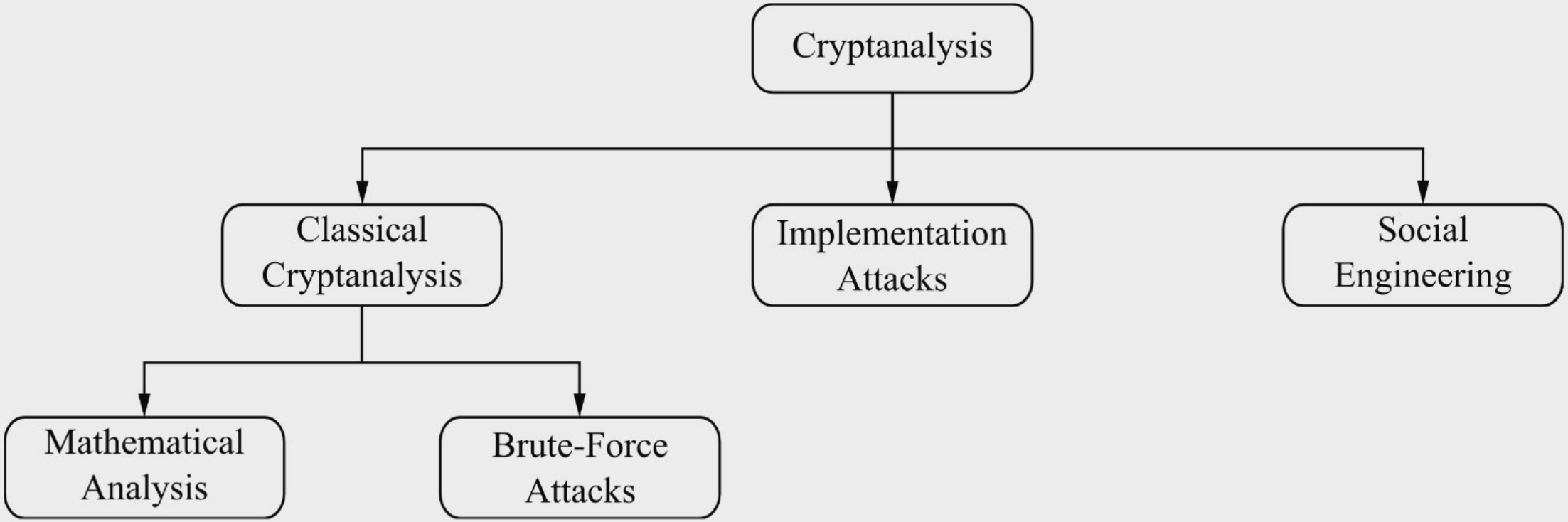
- By further guessing based on the frequency of the remaining letters we obtain the plaintext:

  **WE WILL MEET IN THE MIDDLE OF THE LIBRARY AT NOON ALL ARRANGEMENTS ARE MADE**

# Lesson Learned

- Good ciphers should hide the statistical properties of the encrypted plaintext.
- The ciphertext symbols should appear to be random.
- A large key space alone is not sufficient for a strong encryption function.

# Overview of Cryptanalysis



# Classical Cryptanalysis

- **Ciphertext-only attack**: The adversary has only access to the ciphertext.
- **Known-plaintext attack**: In addition to the ciphertext, the adversary also knows some pieces of the plaintext.
- **Chosen-plaintext attack**: The adversary can choose the plaintext that is being encrypted and also has access to the corresponding ciphertext.
- **Chosen-ciphertext attack**: The adversary can choose ciphertexts and also obtains the corresponding plaintexts.

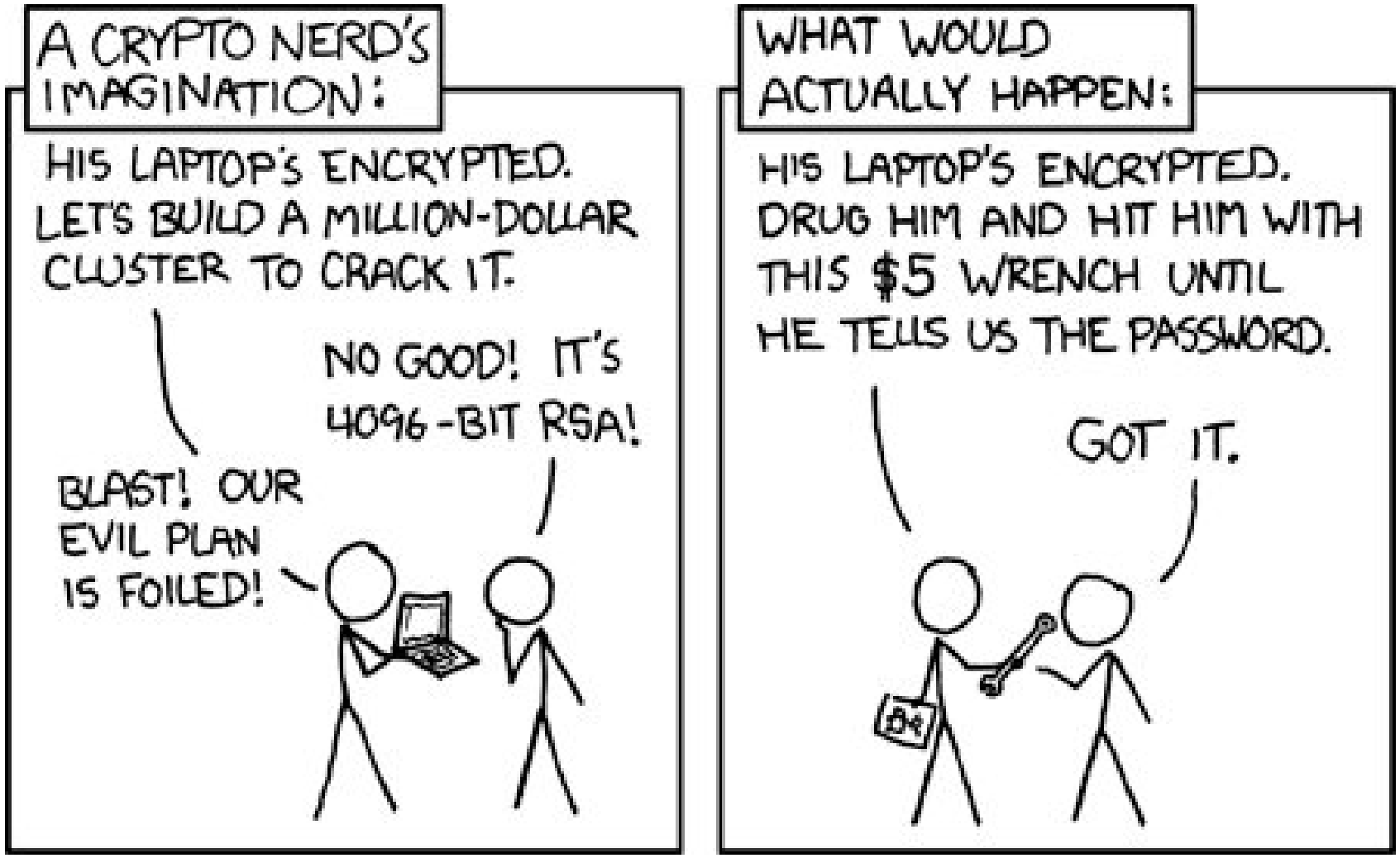
# Implementation Attacks

- Side-channel analysis can be used to extract a secret key by observing the behavior of a cryptographic **implementation**., e.g., an integrated circuit or a piece of software.

# Social Engineering Attacks

- Bribing, blackmailing, tricking or classical espionage can be used to obtain a secret key by involving humans.
- E.g., simply call the victim by phone and say: “This is the IT department of your company. For important software updates we need your password”.

# Security






https://xkcd.com/538

# Lesson Learned

- **An attacker always looks for the weakest link in your cryptosystem. That means we have to choose strong algorithms *and* we have to make sure that all other attacks such as social engineering and implementation attacks are not feasible.**

# Kerckhoffs’ Principle

- A cryptosystem should be secure even if the attacker (Oscar) knows all details about the system, with the exception of the secret key. In particular, the system should be secure when the attacker knows the encryption and decryption algorithms.

# Security by Obscurity

- Designing a system that appears to be more secure because we keep the details hidden is called **security by obscurity**.
- However, experience and military history has shown over time that such systems are almost always weak, and they are very often broken easily as soon as the secret design has been reverse-engineered or leaked out through other means.
- E.g., Content Scramble System (CSS).

# How Many Key Bits Are Enough?

- The discussion of key lengths for symmetric cryptographic algorithms is only relevant if a brute-force attack is the best known attack.
- The key lengths for symmetric and asymmetric algorithms are dramatically different. E.g., a 128-bit symmetric key provides roughly the same security as a 3072-bit RSA (RSA is a popular asymmetric algorithm) key.

# Time for successful brute-force attacks on symmetric ciphers

| Key length | Security estimation |
|---|---|
| 56–64 bits | short term: a few hours or days |
| 112–128 bits | long term: several decades in the absence of quantum computers |
| 256 bits | long term: several decades, even with quantum computers that run the currently known quantum computing algorithms |

# References


- C. Paar, J. Pelzl, and Tim Güneysu, “Introduction to Cryptography and Data Security,” in *Understanding Cryptography*, Springer, 2nd Edition, 2024, pp. 1–35.
- W. Stallings, “Monoalphabetic Ciphers,” in *Cryptography and Network Security, Principles and Practice*, 8th Edition, 2023, pp. 92–95.