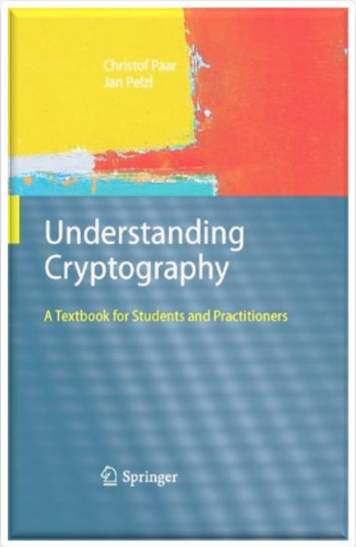
Introduction to Cryptography

Acknowledgement

- These slides are based on the following resources (but modified):
  - Slides accompanying the textbook '*Understanding Cryptography*' by Christof Paar and Jan Pelzl

    http://crypto-textbook.com/


Christof Paar
Jan Pelzl
Understanding Cryptography
A Textbook for Students and Practitioners
Springer

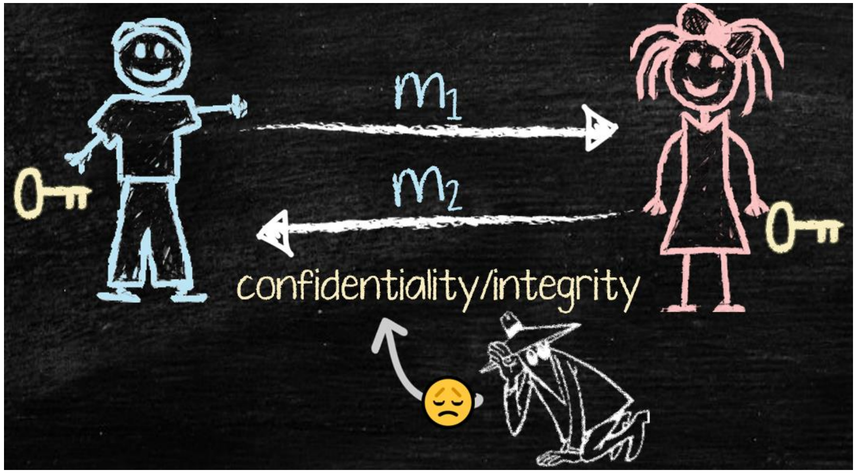
  - Slides from Dr. Ryan Riley

    https://vsecurity.info/

Outline

- Introduction to Cryptography
- Caesar Cipher
- Substitution Cipher
- One-Time Pad Cipher

Introduction to Cryptography



Definition

- The word Cryptography is Greek
  - Crypto: Secret + Graphy: Writing
  - Method to send secret messages using a key

- Basic goal is **Secure communication**
  - Send messages that no one but the expected recipient can read

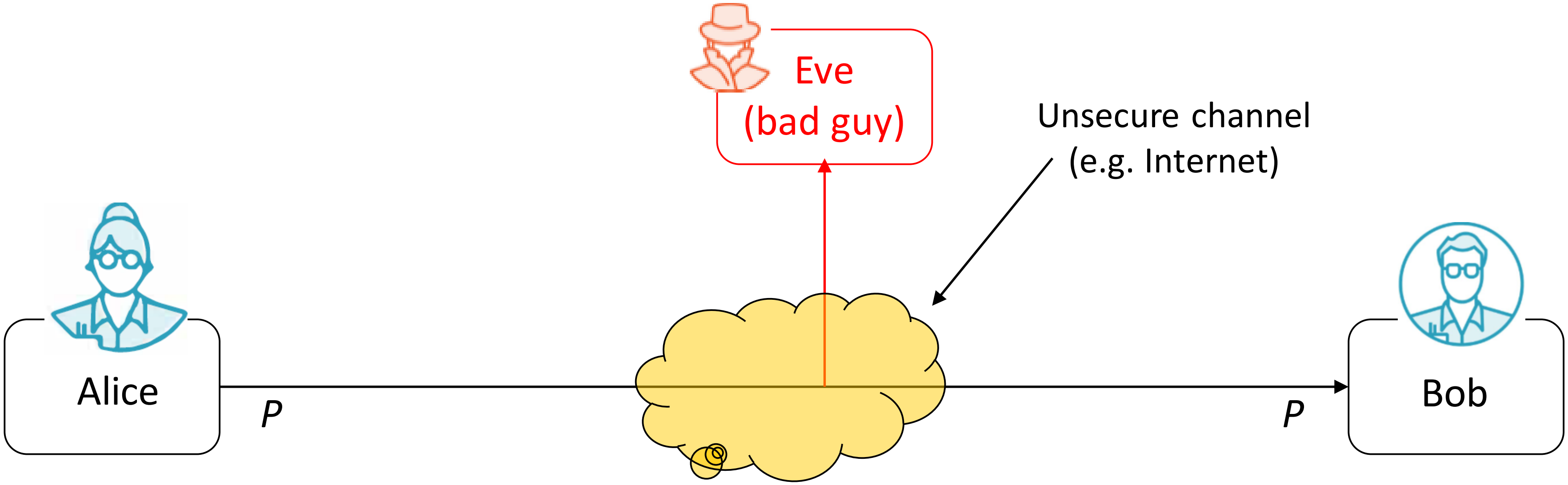
- Many other applications such as:
  - Cryptocurrency, Blockchains
  - Authentication
  - Digital signatures
  - …

Terminology

- Plaintext: A message in its original form
- Ciphertext: A message in encrypted form
- Encryption: Transforming PT to CT
- Decryption: Transforming CT to PT
- Encryption Algorithm / Cipher: The method used for encryption

Symmetric Cryptography

- Alternative names: **private-key**, **single-key** or **secret-key** cryptography.



- **Problem Statement:**

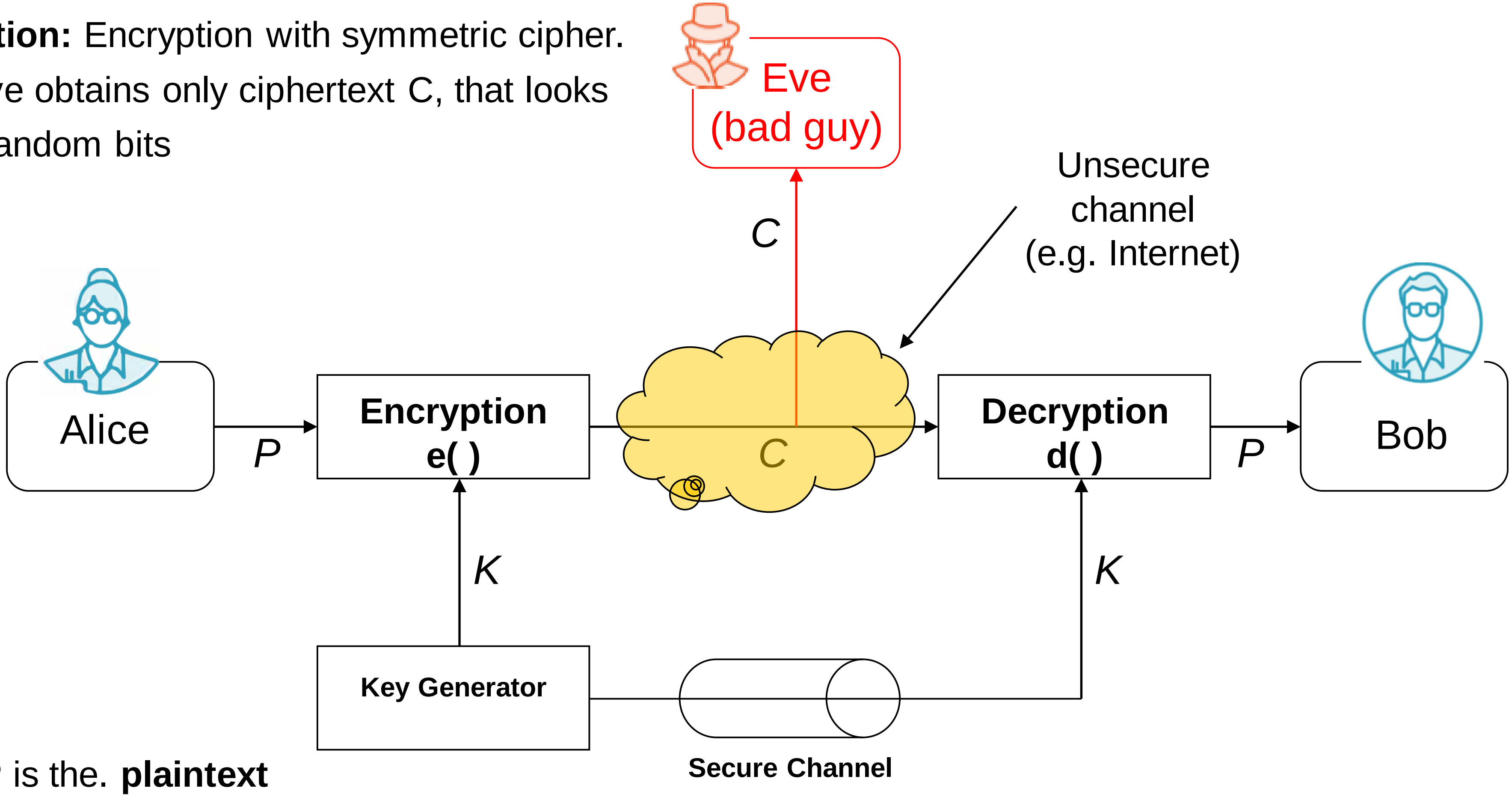
  1) Alice and Bob would like to communicate via an unsecure channel (e.g., Internet)

  2) A malicious third party Eve (the bad guy) has channel access but should not be able to understand the exchanges messages

Symmetric Cryptography

Solution: Encryption with symmetric cipher.

⇒ Eve obtains only ciphertext C, that looks like random bits



- P is the. **plaintext**
- C is the **ciphertext**
- *K* is the **key**
- Set of all keys $\{K_1, K_2, ..., K_n\}$ is the **key space**

Symmetric Cryptography

- Encryption equation **$C = e_K(P)$**
- Decryption equation **$P = d_K(C)$**

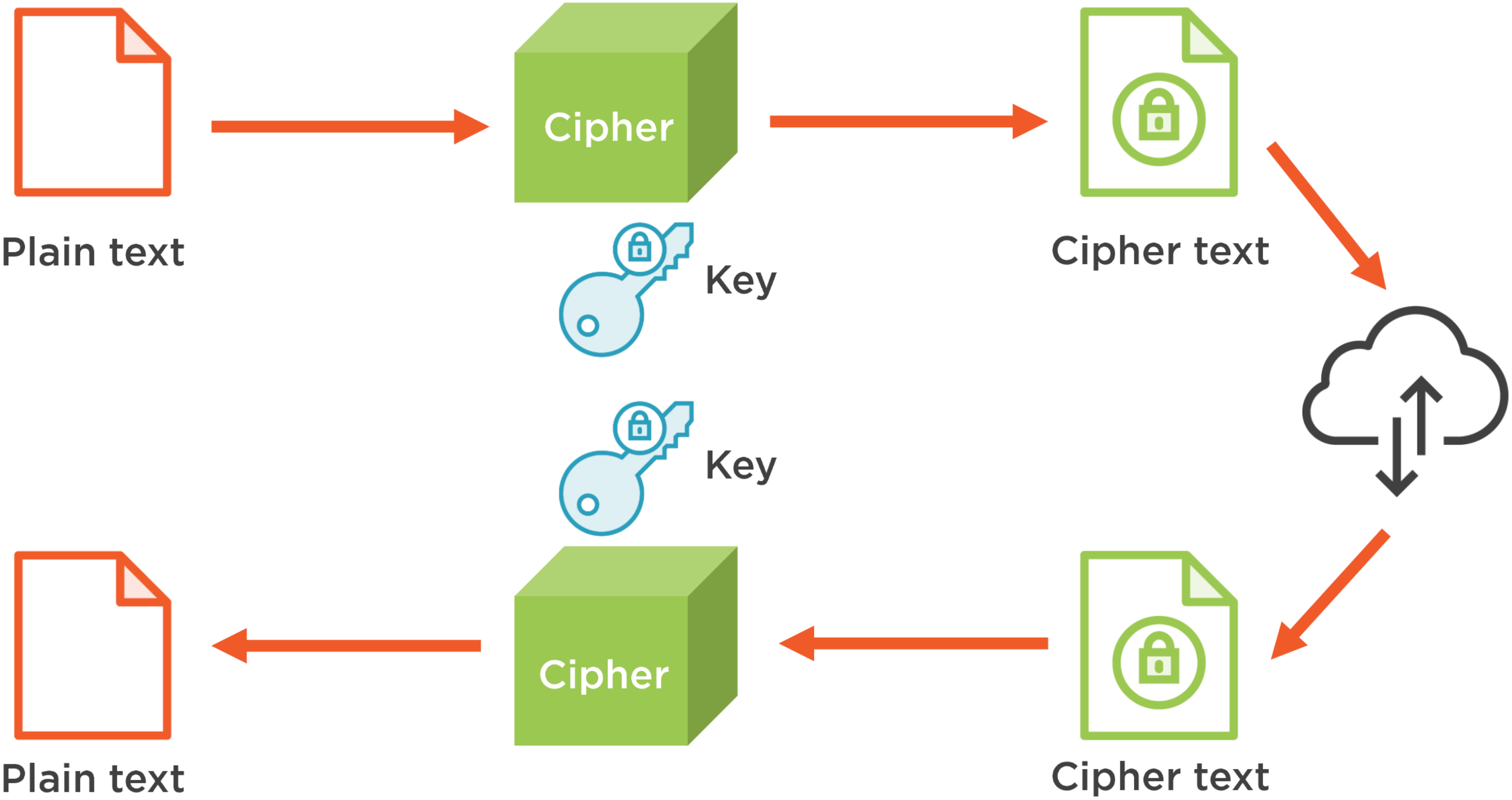
- Encryption and decryption are inverse operations if the same key K is used on both sides:

$$d_K(C) = d_K(e_K(P)) = P$$

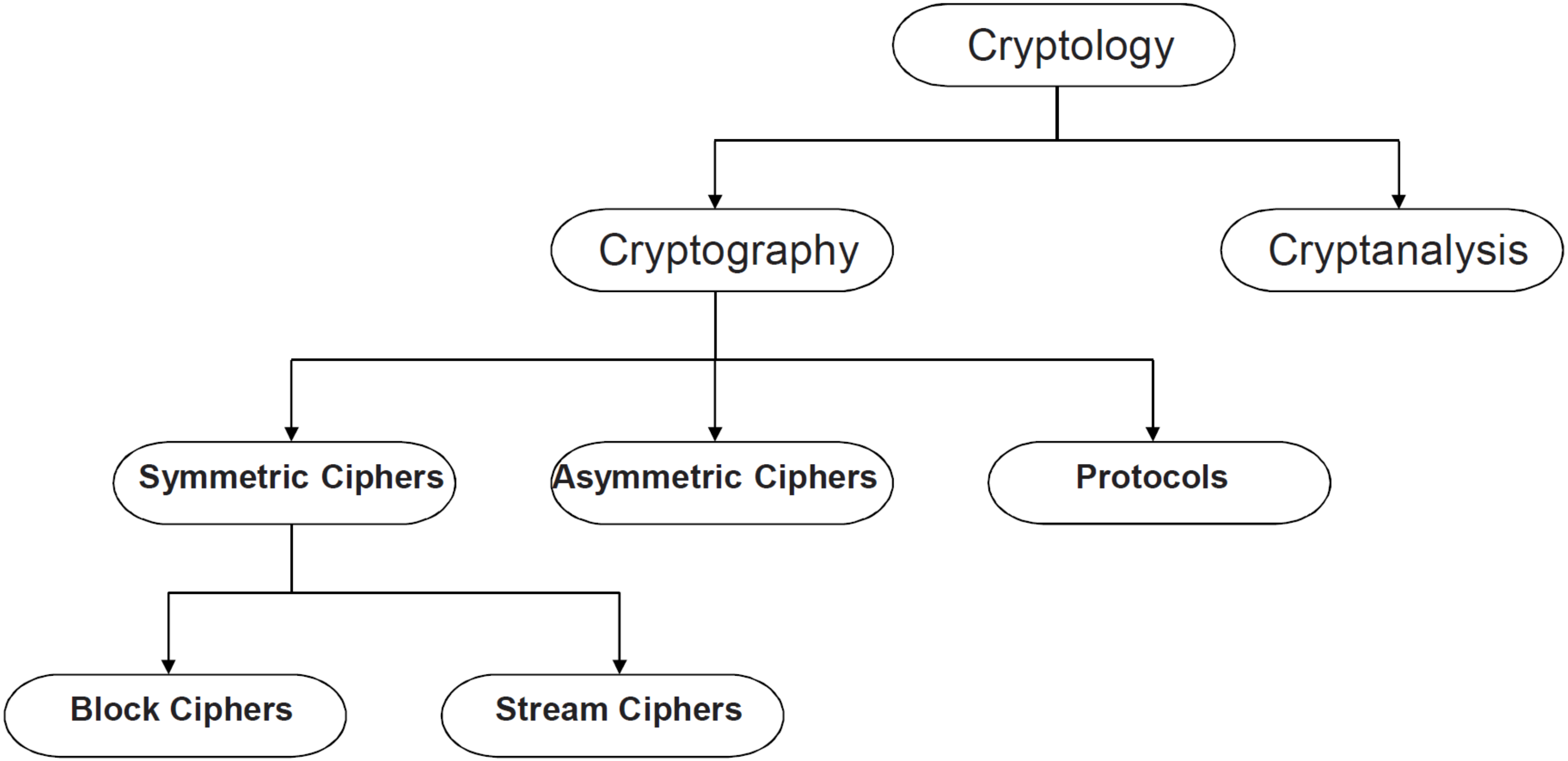
- Important: The key must be transmitted via a **secure channel** between Alice and Bob.
- The secure channel can be realized, e.g., by a human courier or a secure key exchange mechanism (this will be covered later)
- However, the system is only secure if an attacker does not learn the key K!

⇒ **The problem of secure communication is reduced to secure transmission and storage of the key K**

Symmetric Cryptography - Summary



Classification of the Field of Cryptology



Cryptanalysis = Trying to break the key and read enrypted messages

Why do we need Cryptanalysis?

- There is no *mathematical proof of security* for any practial cipher
- The only way to have assurance that a cipher is secure is to try to break it (and fail) !
  - We let lots of really smart people try to break it (cryptanalysis). If they can't, we assume it is secure
  - But… We might be wrong

Kerckhoff Principle is paramount in modern cryptography:

A cryptosystem should be secure even if the attacker knows all details about the system, with the exception of the secret key

- In order to achieve Kerckhoff's Principle in practice:

  Only use widely known ciphers that have been cryptanalyzed for several years by good cryptographers!
- **Remark:** It is tempting to assume that a cipher is "more secure" if its details are kept secret. However, history has shown time and again that secret ciphers can almost always been broken once they have been reversed engineered. (Example: Content Scrambling System (CSS) for DVD content protection.)

Kerckhoff’s Principle

Security of a Cryptographic Algorithm should rely ONLY on the **secrecy of the KEYS**, and **NOT on the secrecy of the METHOD** used

“Do not rely on security through obscurity”

Brute-Force Attack (or Exhaustive Key Search) against Symmetric Ciphers

- Treats the cipher as a black box
- Requires (at least) 1 plaintext-ciphertext pair (P_0, C_0)
- Check all possible keys until condition is fulfilled:

$$d_k(C_0) = P_0$$

Key length in bit	Key space	Security life time
64	2^{64}	Short term (few days or less)
128	2^{128}	Long-term (several decades in the absence of quantum computers)
256	2^{256}	Long-term (also resistant against quantum computers – note that QC do not exist at the moment)

Caesar Cipher

Back

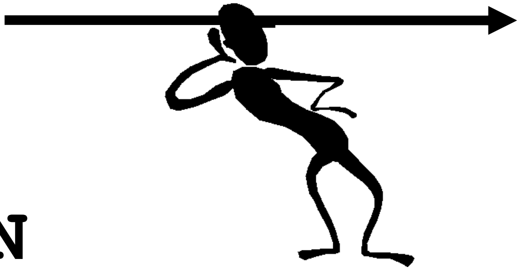
Simple Ciphers

- Originally, cryptography was performed by hand
- Goal was to protect messages sent by couriers
  - From people who might intercept the courier
  - From the courier himself
- War was a popular time to use them

Caesar Cipher

Julius Ceasar (100-44 BC)

DWWDFN DW GDZQ



ATTACK AT DAWN

ATTACK AT DAWN

Encrypt the message!

Decrypt the ciphertext!

- The sender and receiver must know something that the adversary doesn't.
- This is called a **cryptographic key**

Caesar Cipher

Secret key: A random number from {1,…,26}, say **3**

DWWDFN DW GDZQ →

Encryption

Message:	**ATTACK AT DAWN**
Key: + 3	↓↓↓↓↓↓ ↓↓ ↓↓↓↓
Ciphertext:	**DWWDFN DW GDZQ**

Caesar Cipher

Secret key: A random number from {1,…,26}, say **3**

DWWDFN DW GDZQ →

Decryption

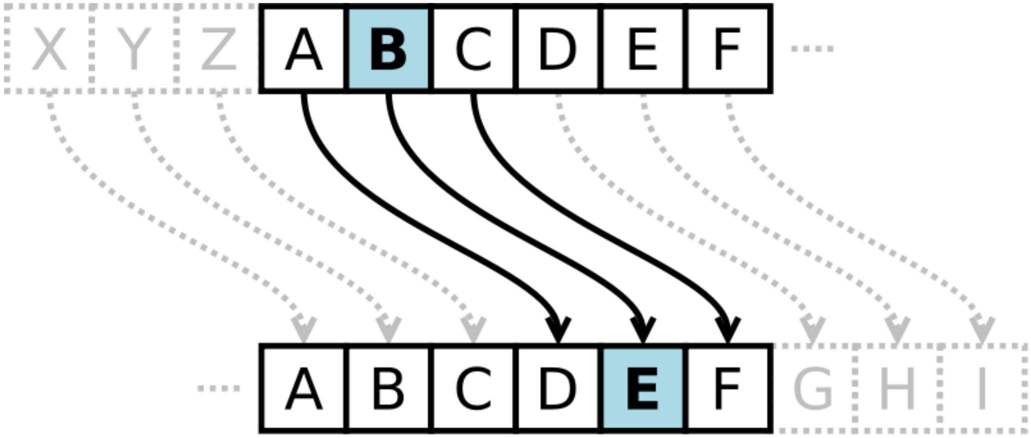
Ciphertext: **DWWDFN DW GDZQ**

Key: - 3 ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓

Message: **ATTACK AT DAWN**

Caesar Cipher

- Earliest documented cipher was used by Caesar in 50BC !
- Each letter in a message is substituted by another that is 3 letters away
  - **A** becomes **D**, **B** becomes **E**, etc.
  - Note that the letters ”wrap around” at the end of the alphabet, which can be mathematically expressed using **mod 26**



Caesar Cipher Example

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

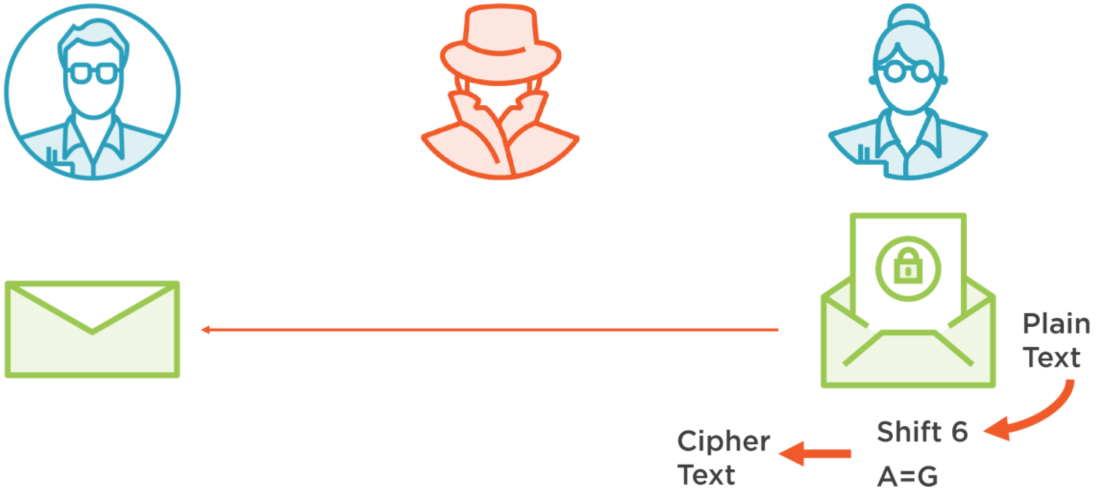
```
ATTACK AT DAWN
```

Shift Cipher

- Generic version of Caesar cipher
- Each letter is shifted by N. In Caesar, N=3

Let k, x, y ε {0,1, …, 25}

- Encryption: $y = e_k(x) \equiv x + k \bmod 26$
- Decryption: $x = d_k(x) \equiv y - k \bmod 26$



Shift Cipher: Example

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Let’s do one for N=10…

```
ATTACK AT DAWN
```

Shift Cipher Cryptanalysis

- How do we break this?
- Brute-force: Try all possible values for N
  - There are only 26
- Feasibility?
  - Easy by hand
  - Trivial by computer

Substitution Cipher

Substitution Cipher

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

- Generate a random set of substitutions for each letter
    - Always a 1:1 correspondence

Substitution Cipher Example

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

```
ATTACK AT DAWN
```

Substitution Cipher Cryptanalysis

- Brute-force: Try all possible letter combinations
    - There are (26!) = **403291461126605635584000000**

- Exhaustive key search will take a long time …
- Letter frequency analysis attack can be used against the substitution cipher

Substitution Cipher Cryptanalysis

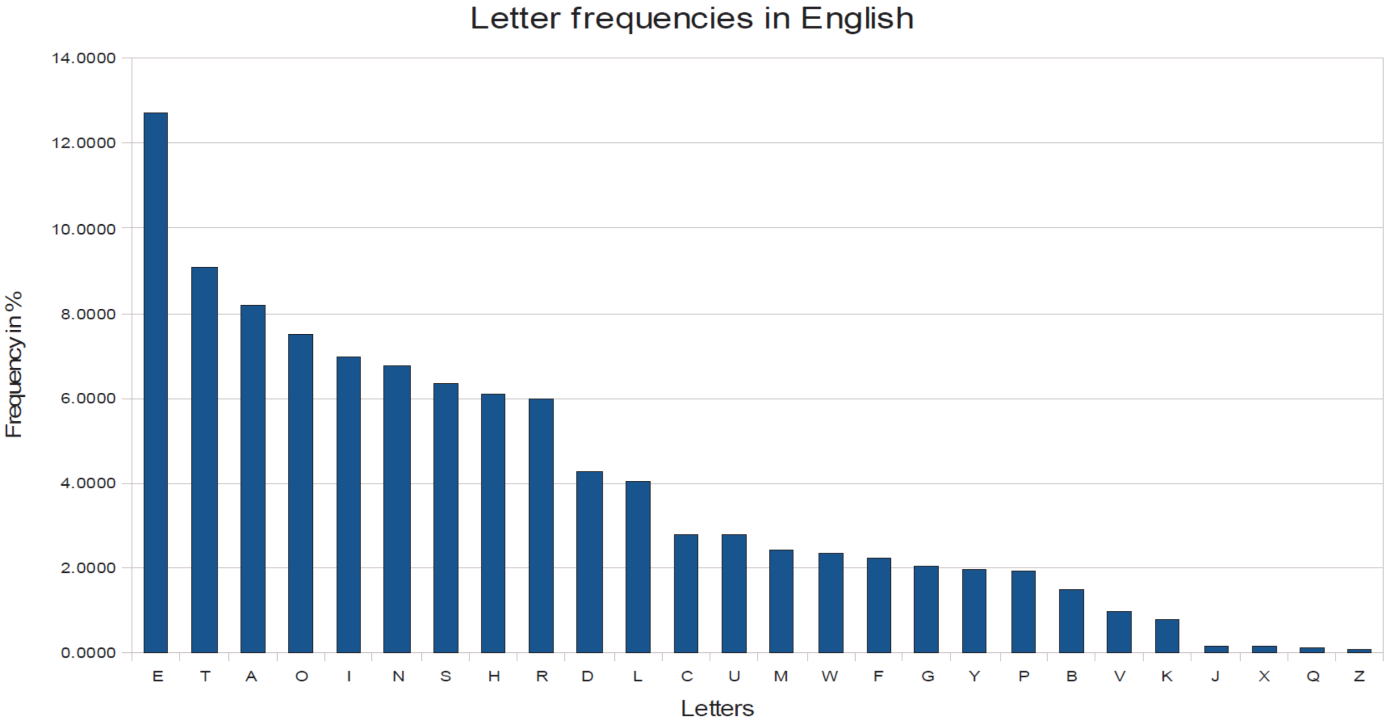
`QYYQAV QY WQID`

- Key observation: In a substitution cipher, **basic language features** are preserved
  - You can tell how often a letter occurs in the message
  - You can see when letters repeat
  - Etc.

- Use a technique called frequency analysis

Frequency Analysis

- Not all letters in a language occur with the same frequency. E.g., In English,
  - E is most common
  - Vowels are about 40%
  - Vowels tend to be separated by consonants
  - Q tends to be followed by U
  - Etc.



Breaking the Substitution Cipher with Letter Frequency Attack

- Let‘s take an example and identify the most frequent letter:

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

- We replace the ciphertext letter q by E and obtain:

  ```
  iE ifcc vEEr fb rdE vfllcE na rdE cfjwhwz hr
  bnnb  hcc hwwhbsEvEbre hwE vhlE
  ```

- 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
  ```

- In practice, not only frequencies of individual letters can be used for an attack, but also the frequency of letter pairs (i.e., “th“ is very common in English), letter triples, etc.

Vigenère Cipher
(1900-1950)

- Poly-alphabetic cipher
  - One plaintext letter can become *different* ciphertext letters
- Uses a text based key and modulo arithmetic to perform the encryption
- Frequency analysis is possible, but much more difficult

Vigenère Cipher: Example

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

- Let’s choose a key of “MONKEY”

```
ATTACK AT DAWN
MONKEY MO NKEY
MHGKGI MH QKAL
```

One-Time Pad Cipher

< Back

One-Time Pad

- Vigenère cipher with a randomly chosen key as long as the message
- Key needs to be shared between parties beforehand
- Key can **never** be re-used
- Provable unbreakable without the key
- This is the only perfect cryptography

One-Time Pad: Example

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

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
26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51

- Our random key is “FOWIFOZMQOAF”

```
ATTACK AT DAWN
FOWIFO ZM QOAF
```

One-Time Pad

Perfect Secrecy is guaranteed by the One-time Pad

Unconditionally secure cryptosystem:
A cryptosystem is unconditionally secure if it cannot be broken even with *infinite* computational resources

THE GOOD: Unbreakable regardless of the power of the adversary

THE BAD: Impractical! Needs very, very long shared keys

Crypto Components

- All of the previous techniques have two basic components:
  - **Algorithm** (What you do to the message)
  - **Key** (The secret that you need in order to encrypt/decrypt properly)

- When using these algorithms, the **key is secret**
- The algorithm is not

Summing Up

- We trust a cryptographic algorithm if lots of smart people can’t break it
- We looked at three types of simple ciphers:
  - Shift Cipher
  - Substitution Cipher
  - One-Time Pad Cipher
- They each have an algorithm and a key
- Long key is required for cryptographic algorithms in order to prevent exhaustive key-search attacks

Resources

- Cryptool - Software demonstrating many ancient and modern ciphers

https://www.cryptool.org/en/

- An excellent one-hour video summarizing the last 40 years of modern cryptography by Ron Rivest

- The International Association of Cryptographic Research is the professional organization of cryptographers.