Cryptography I

CS 166: Introduction to Computer Systems Security

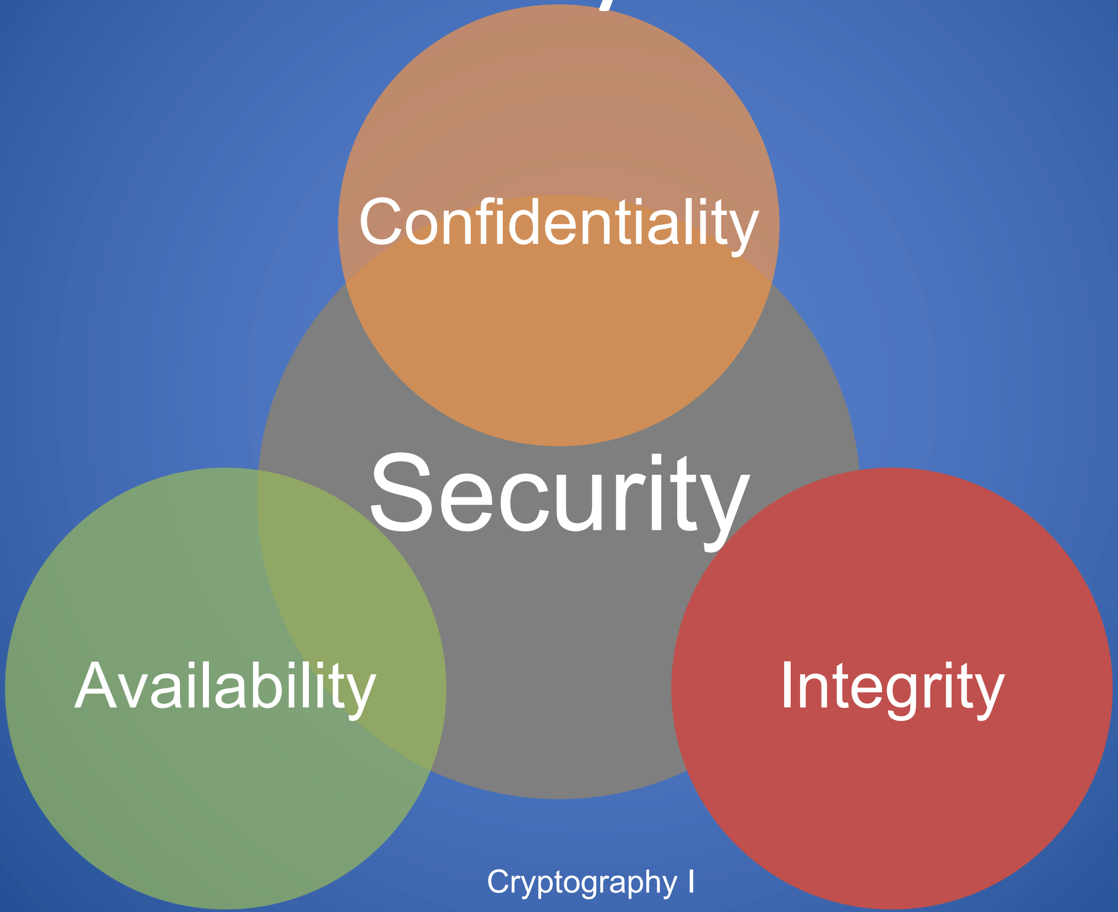
Bernardo Palazzi

- Brown experiences:
  - In 2007, started collaborating with Brown University.
  - Founder and Academic Director of Brown Executive Master in Cyber Security.
  - Founder and former DGS of Master of Science in Cyber Security.
  - Co-chair of the Strategic Planning Committee for the Cyber Master.

Other professional experiences:

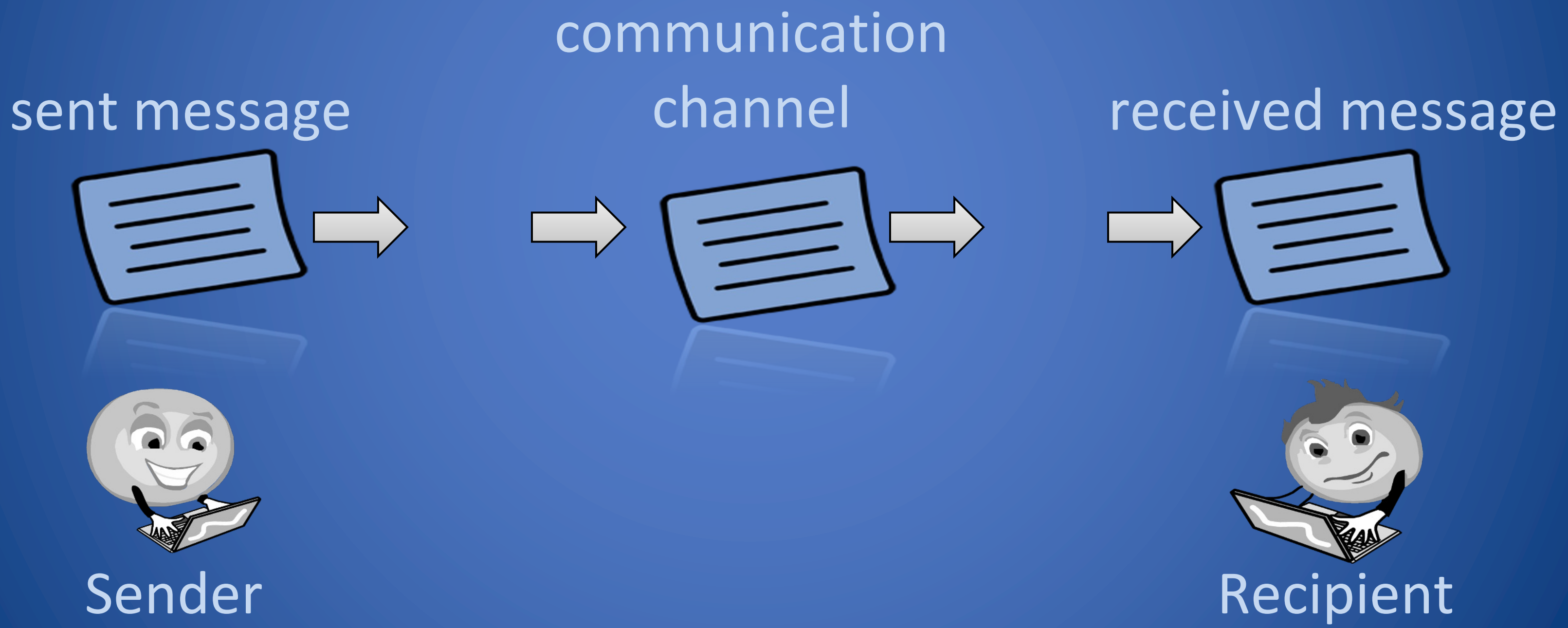
- Advisor for Capacity and Competence Development at the Italian National Cybersecurity Agency (ACN), similar to CISA and NSA in the US.
- As the First Data Protection Officer (DPO), oversaw privacy regulation for the whole population at the Italian National Institute of Statistics.
- Managed the computer security of the first online population census.
- Founder and CTO of a cloud data security startup based on an international patent based on my PhD thesis.

Security Goals

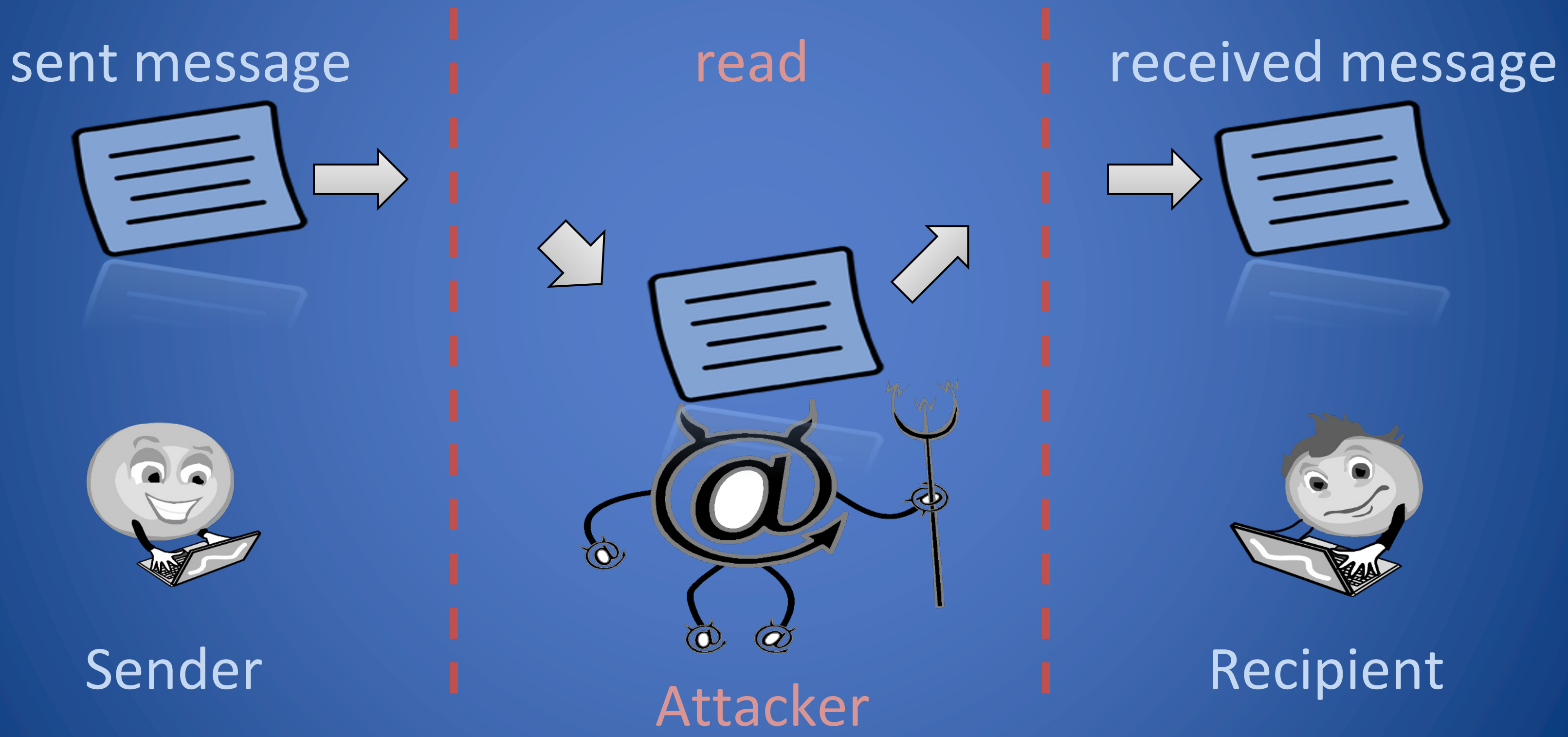


Attacks on Communication

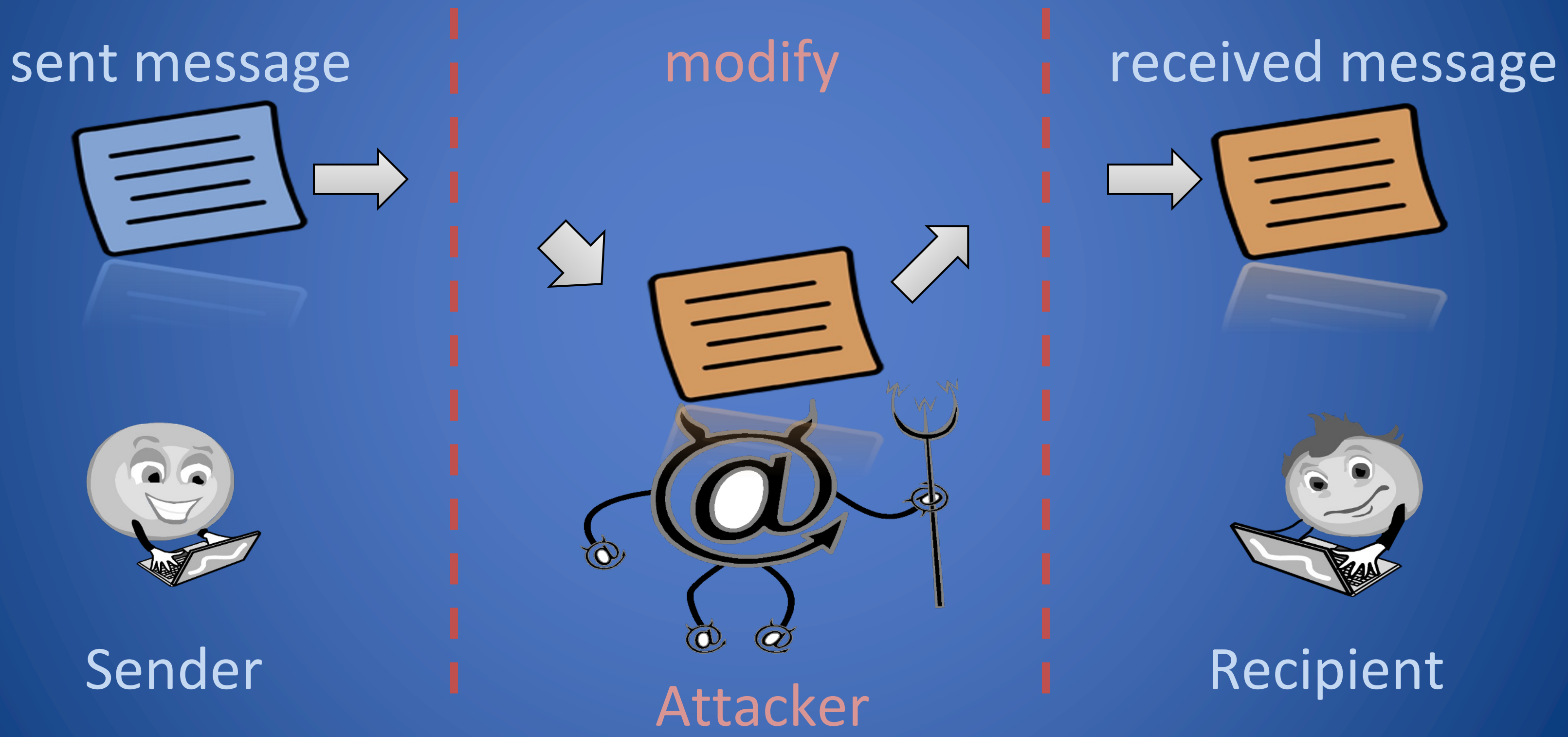
Standard Communication



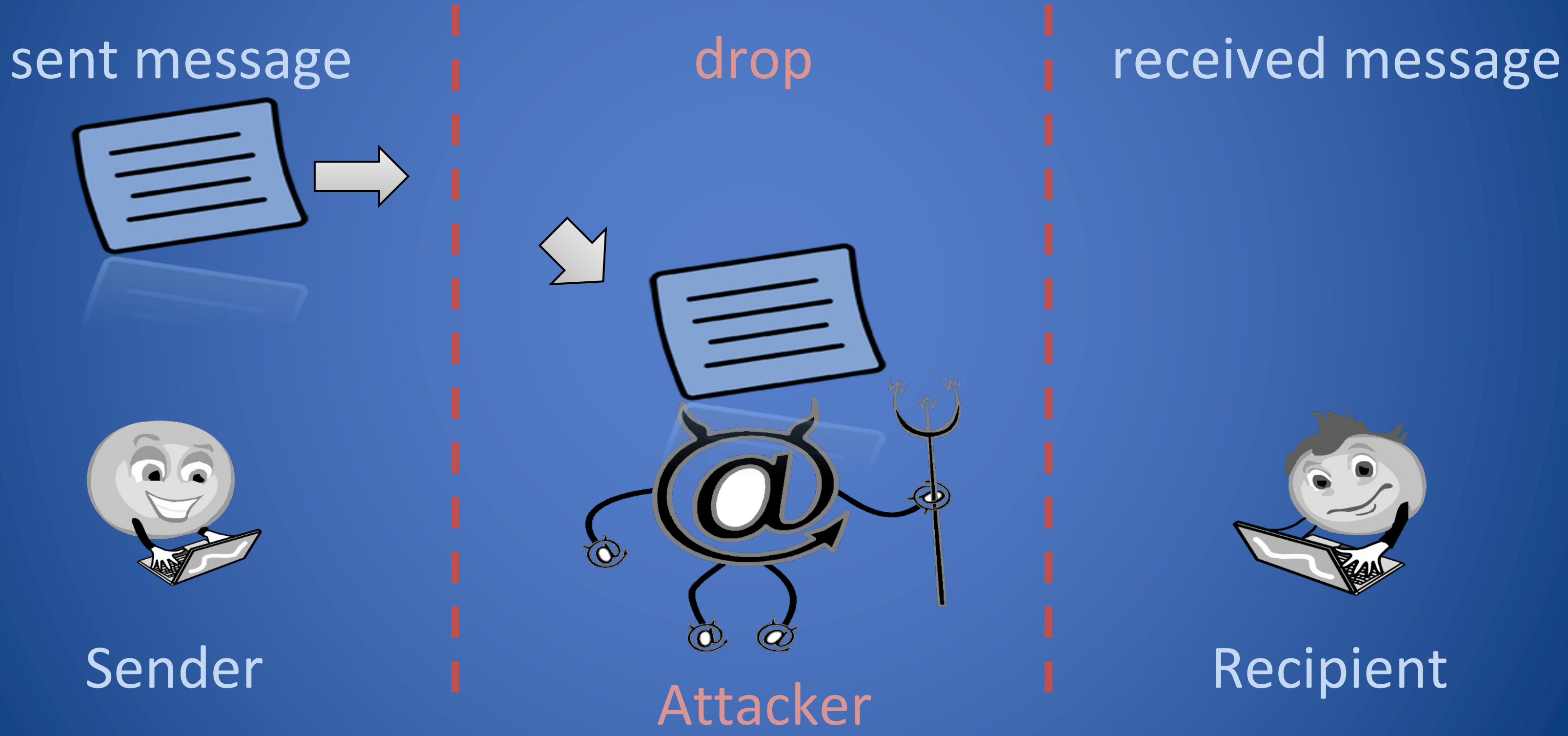
Eavesdropping



Tampering

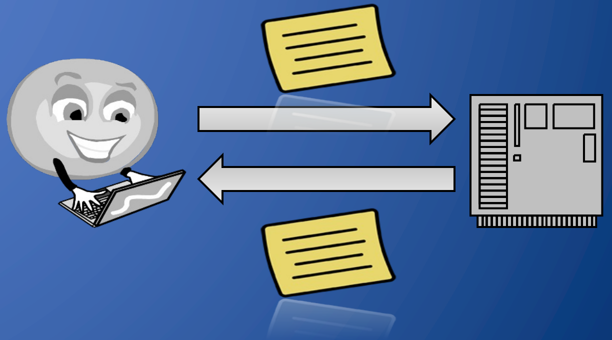

sent message
modify
received message
Sender
Attacker
Recipient

Blocking


sent message
drop
received message
Sender
Attacker
Recipient

Cryptography

- Cryptography provides methods for assuring the confidentiality and integrity of data that is
  - transmitted over communication channels (e.g., web pages and email messages)
  - stored on devices (e.g., files on a laptop or data center)

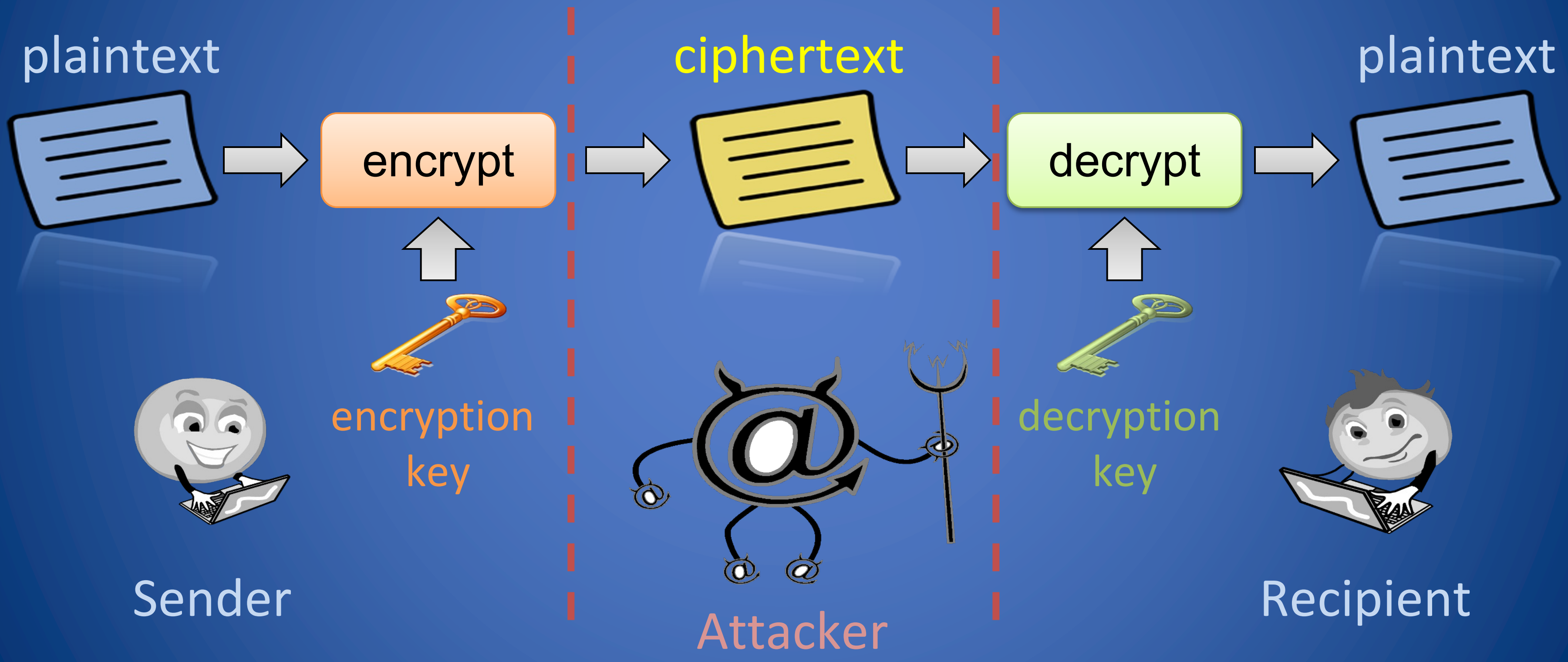


Open Design Principle

- Publicly available system architecture and algorithms
- Security relies solely on keeping keys secret
- Formulated by Auguste Kerckhoffs in 1883
- Opposite of “security by obscurity”
- Claude Shannon in 1949 said *"the enemy knows the system"*:
  - *"one ought to design systems under the assumption that the enemy will immediately gain full familiarity with them"*

Image source: https://en.wikipedia.org/wiki/Auguste_Kerckhoffs#/media/File:Auguste_Kerckhoffs.jpg

Encrypted Communication

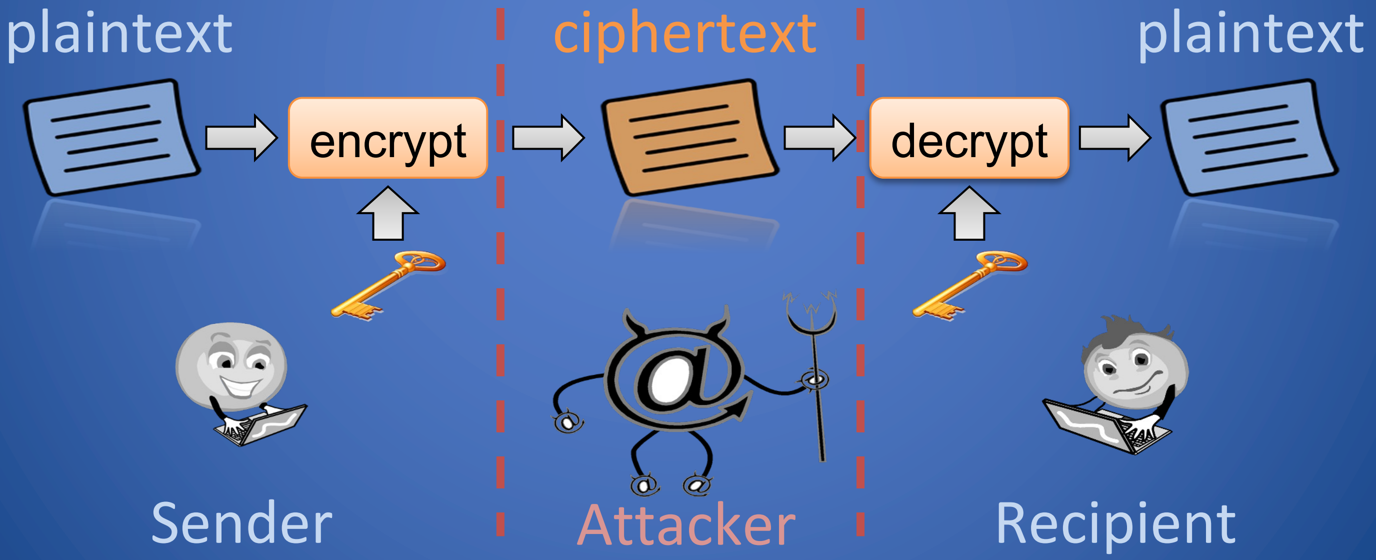


Encryption

- Encryption allows to secure communication
  - Originally focused on confidentiality alone
- The encryption algorithm combines the plaintext with the encryption key to produce the ciphertext 
  - The ciphertext is transmitted instead of the plaintext
- The decryption algorithm combines the ciphertext with the decryption key to return the plaintext 
  - Only the intended recipient should have the secret key
- Encryption and decryption should be computationally infeasible without the corresponding keys

Symmetric Encryption

- Same key is used for encryption and decryption
- Encryption and decryption algorithms are one the reverse of the other
- We need a secure channel to set up key



Classic Symmetric Encryption

Julius Caesar's Cipher

- Encryption
  - replace A with D
  - replace B with E
  - replace C with F
  - …
  - replace X with A
  - replace Y with B
  - replace Z with C
- Encryption key
  - Forward alphabet shift: +3
- Decryption key
  - Reverse alphabet shift: –3

AVE → DZH

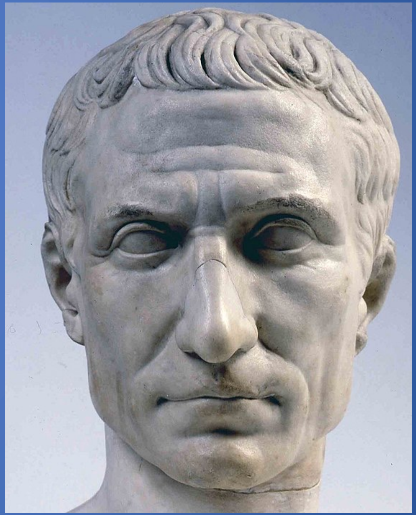


Image source: https://en.wikipedia.org/wiki/Julius_Caesar#/media/File:Gaius_Iulius_Caesar_(Vatican_Museum).jpg

Alphabet Shift Cipher

- Generalization of Caesar's cipher
- Replace each character c of the plaintext with the character k positions after c in the alphabet
- Key for encryption and decryption: number k
- Insecure encryption method
- Can be easily cracked by trying all possible values of k between 1 and the size of the alphabet

Substitution Cipher

- Arbitrary permutation of the characters
  - A → K
  - B → T
  - C → G
  - …

CAB → GKT

- Key: permutation of the alphabet characters (e.g., KTG …)
- Number of possible keys for a 26-character alphabet ≈ 4×10^{26}
- Unfeasible to try all possible keys but …
- Can be cracked by frequency analysis
  - most frequent letters in English: e, t, o, a, n, i, ...
  - most frequent digrams: th, in, er, re, an, ...
  - most frequent trigrams: the, ing, and, ion, ...
- Attack first described in a 9th century book by al-Kindi

Frequency Analysis

PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP **LBO** LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV **LBO** LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV **LBO** DJCMPV ZOICJO BYS, KXUYPD: "DJOXL EYPD, ICJ X LBCMKXPV XPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV **LBO** RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?"

OFYRCDMO, LXROK IJCS **LBO** LBCMKXPV XPV CPO PYDBLK

Example from

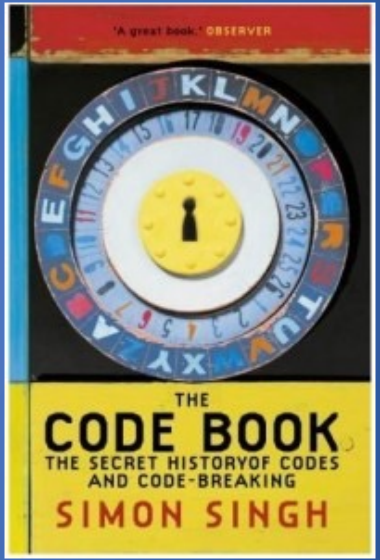
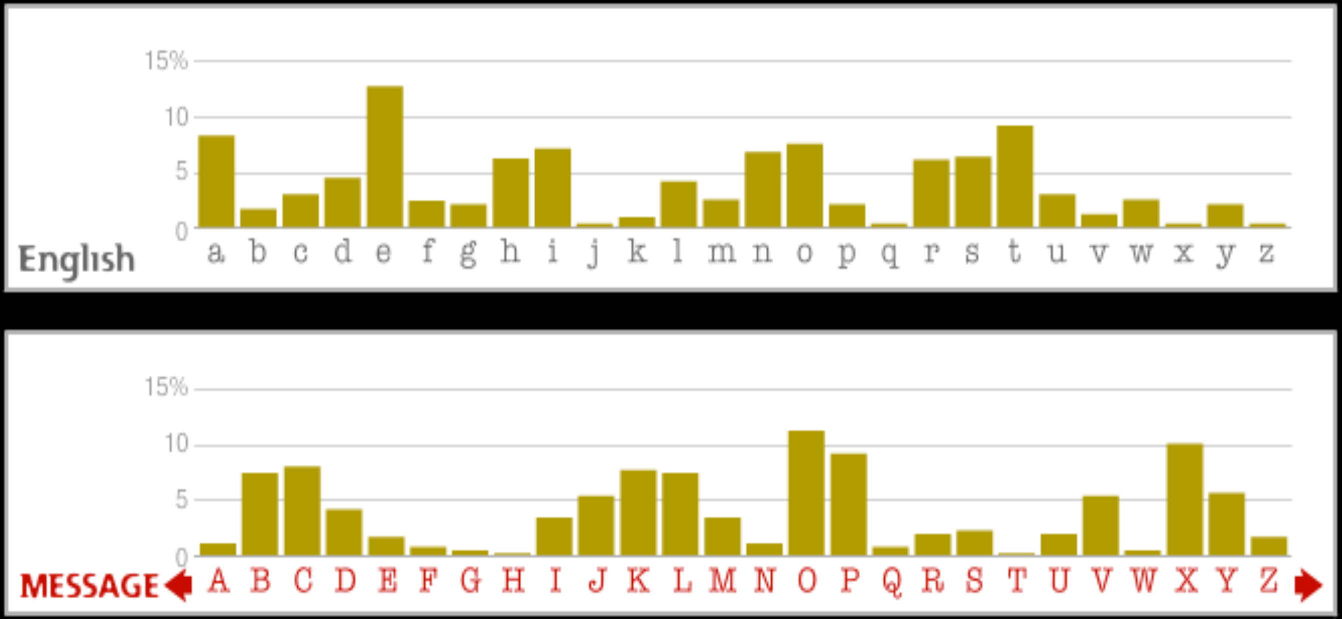


Image source: https://simonsingh.net

Letter Frequencies Graph



Common Digrams

In English: th, he, in, er, an

In MESSAGE: LB, PV, BO, XP, CM

Common Trigrams

In English: the, and, tha, ent, ing

In MESSAGE: XPV, YPD, LBO, EYP, LBC

First guess

- LBO → THE

Frequency Analysis (cont.)

PCQ VMJYPD THYK TYSE KHXHJXWXV HXV
ZCJPE EYPD KHXHJYUXJ THJEE KCPK. CP
THE THCMKXPV XPV IYJKT PYDHT, QHEP
KHE HXV EPVEV THE TXRE CI SX'XJMI, KHE
JCKE XPV EYKKEV THE DJCMPV ZEICJE HYS,
KXUYPD: "DJEXT EYPD, ICJ X THCMKXPV
XPV CPE PYDHTK Y HXNE ZEEP
JEACMPTYPD TC UCM THE IXZREK CI FXKT
XDEK XPV THE REDEPVK CI XPAYEPT EYPDK.
SXU Y SXEE KC ZCRV XK TC AJXNE X IXNCMJ
CI UCMJ SXGEKTU?"

EFYRCDME, TXREK IJCS THE THCMKXPV
XPV CPE PYDHTK

L → T

B → H

O → E

More guesses

J → R

K → S

X → A

Frequency Analysis (cont.)

PCQ VMRYPD THYS TYSE SHAHRAWAV HAV ZCRPE EYPD SHAHRYUAR THREE SCPS. CP THE THCMSAPV APV IYRST PYDHT, QHEP SHE HAV EPVEV THE TARE CI SA'ARMI, SHE RCSE APV EYSSEV THE DRCMPV ZEICRE HYS, SAUYPD: “DREAT EYPD, ICR A THCMSAPV APV CPE PYDHTS Y HANE ZEEP REACMPTYPD TC UCM THE IAZRES CI FAST ADES APV THE REDEPVS CI APAYEPT EYPDS. SAU Y SAEE SC ZCRV AS TC ARANE A IANCMR CI UCMR SAGESTU?”

EFYRCDME, TARES IRCS THE THCMSAPV APV CPE PYDHTS

L → T

B → H

O → E

J → R

K → S

X → A

Decryption

PCQ VMJYPD LBYK LYSO KBXBJXWXV BXV ZCJPO EYPD KBXBJYUXJ LBJOO KCPK. CP LBO LBCMKXPV XPV IYJKL PYDBL, QBOP KBO BXV OPVOV LBO LXRO CI SX'XJMI, KBO JCKO XPV EYKKOV LBO DJCMPV ZOICJO BYS, KXUYPD: “DJOXL EYPD, ICJ X LBCMKXPV XPV CPO PYDBLK Y BXNO ZOOP JOACMPLYPD LC UCM LBO IXZROK CI FXKL XDOK XPV LBO RODOPVK CI XPAYOPL EYPDK. SXU Y SXEO KC ZCRV XK LC AJXNO X IXNCMJ CI UCMJ SXGOKLU?”

OFYRCDMO, LXROK IJCS LBO LBCMKXPV XPV CPO PYDBLK

Now during this time Shahrazad had borne king Shahriyar three sons. On the thousand and first night, when she had ended the tale of Ma'aruf, she rose and kissed the ground before him, saying: “great king, for a thousand and one nights I have been recounting to you the fables of past ages and the legends of ancient kings. May I make so bold as to crave a favour of your majesty?”

Epilogue, Tales from the Thousand and One Nights

Clicker Question (TopHat: 821033)

Clicker Question

- Bob is experimenting with different symmetric encryption schemes to securely communicate with Alice
- To test his knowledge, he decides to encrypt the plaintext "HELLO WORLD" using an alphabet shift cipher, where k = 4
- Which of the following ciphertexts is correct?

a. KHOOR ZRUOG

b. MHPOS ARVPH

c. LIQQR WRVOH

d. LIPPS ASVPH

Clicker Question

Answer: D

H E L L O W O R L D

+4

L I P P S A S V P H

Symmetric Encryption: The modern era

Symmetric Encryption Algorithms

Two different approaches for symmetric key encryption

Stream cipher:

1. Generate a sequence of bits (keystream)
2. Combine the keystream with plaintext (e.g. **XOR**)
3. Create a ciphertext stream

- If the entire keystream is random and used only once , is a One-Time Pad (OTP)

Block cipher:

1. Take a fixed-length block of plaintext
2. Create a block of ciphertext of the same length

- Generally, the keys are reused
- It is more common

Stream cipher: One-Time Pad

Bitwise XOR

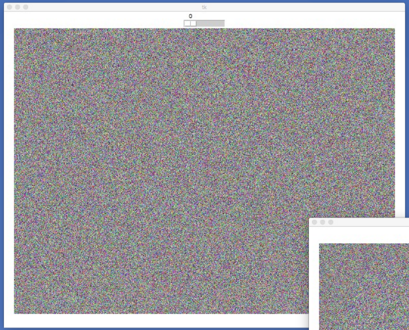
X	Y	$X \oplus Y$
0	0	0
0	1	1
1	0	1
1	1	0

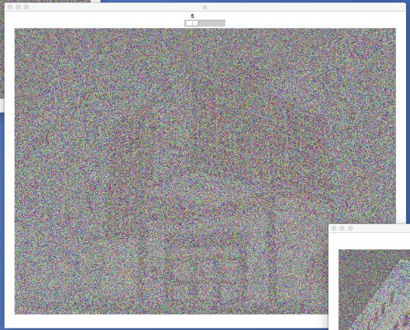
One-Time Pad (~Vernam Cipher)

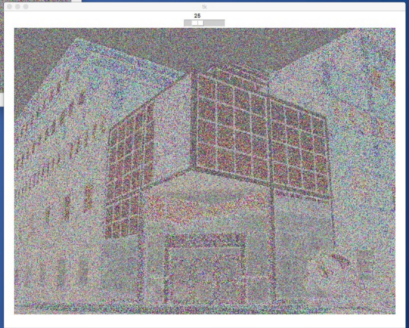
- Key
  - Sequence of random bits
  - Same length as plaintext
- Encryption
  - $C = K \oplus P$
  - Example
    - P = 01101001
    - K = 10110010
    - C = 11011011
- Decryption
  - $P = K \oplus C$
- Advantages
  - Each bit of the cyphertext is random
  - **Fully secure** if key used only once (e.g., Beale's treasure)
- Disadvantages
  - Key as large as plaintext
    - Difficult to generate and share
  - Key cannot be reused

Demo: Pitfalls with One-Time Pads

Imperfect Randomness

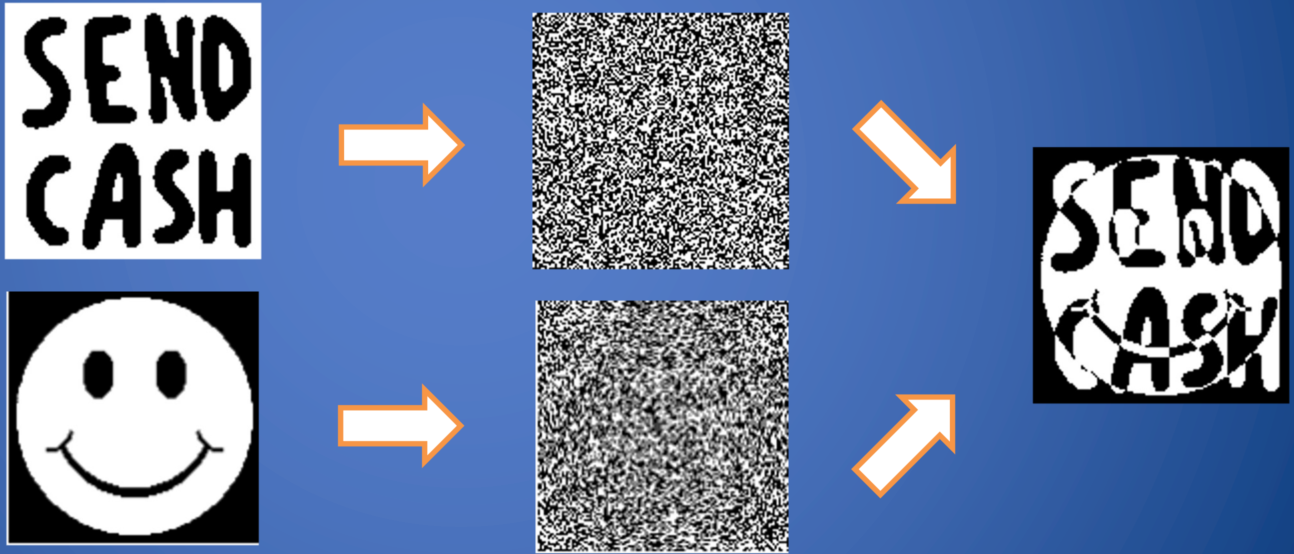






Source: Justin Bisignano and Joshua Liebow-Feeser

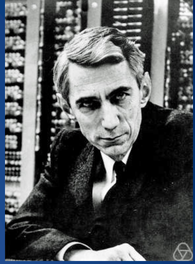
Key Reuse



Source: Cryptosmith and David Lowry-Duda, Cryptography Stack Exchange

Block ciphers

Confusion and Diffusion


Claude Shannon
wikimedia

Two properties of the operation of a secure cipher, defined by Claude Shannon in 1949 - Communication Theory of Secrecy Systems

- Confusion seeks to make the relationship between the key and the ciphertext as complex and difficult as possible
  - It typically involves substituting one element for another (e.g. , Caesar Cipher, Vigenère Cipher)
- Diffusion aims to dissipate the redundancy in the statistics of the plaintext in the statistics of the ciphertext
  - This ensures that changing one character of the plaintext results in multiple changes in the ciphertext (e.g. transposition, permutation)

Confusion: Vigenère Cipher (Polyalphabetic)

This is a type of substitution cipher

- Invented by Blaise de Vigenère in 19th
- The algorithm is polyalphabetic
  - Where the secret key is repeated along the length of plaintext/ciphertext
  - The same letter in plain text could be encrypted with different letters in cipher text

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

Plaintext:	C	Y	B	E	R	I	S	A	W	E	S	O	M	E
Keyword:	B	R	O	W	N	B	R	O	W	N	B	R	O	W
Ciphertext:	D	P	P	A	E	J	J	O	S	R	T	F	A	A

Vigenèe Source:wikimedia

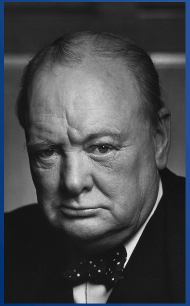
Symmetric Encryption at War

Vigenere Cipher
(American Civil War)

Navajo Code
(WW II US vs Japan)

Enigma machine[3]
(WW II Nazi vs. Allies)
A substitution cipher with a period of 16.900 characters

Alan Turing[4] decrypted under the project 'Ultra'



"It was thanks to Ultra that we won the war."

Winston Churchill[5] to King George VI

1: https://en.wikipedia.org/wiki/Vigen%C3%A8re_cipher#/media/File:Confederate_cipher_disk.png
2: https://www.wikitree.com/blog/wp-content/uploads/2019/08/24418587.jpeg.jpg
3: https://en.wikipedia.org/wiki/Enigma_machine#/media/File:Enigma_(crittografia)_-_Museo_scienza_e_tecnologia_Milano.jpg
4: https://www.npg.org.uk/collections/search/use-this-image/?mkey=mw165875
5: https://en.wikipedia.org/wiki/Winston_Churchill#/media/File:Sir_Winston_Churchill_-_19086236948.jpg

The Dawn of the Digital Era for the civilian sector

- In 1959, the integrated circuit was invented, and private organizations, particularly banks, started to use computers.
- Security has become more and more critical for relevant transactions between different.
- Different companies could use proprietary crypto schemes that the receivers should have implemented for decryption.
- Standardization was necessary to allow easy communication between different parties.
- In 1973, the National Bureau of Standards (NBS), now NIST, invited researchers to propose a cryptographic candidate for the protection of sensitive, unclassified electronic government data.

The call for a Data Encryption Standard (DES)

The algorithm must:

- provide a high level of security.
- be completely specified and easy to understand.
- be available to all users.
- be adaptable for use in diverse applications.
- be economically implementable in electronic devices.
- be efficient to use.
- be able to be validated.
- be exportable.

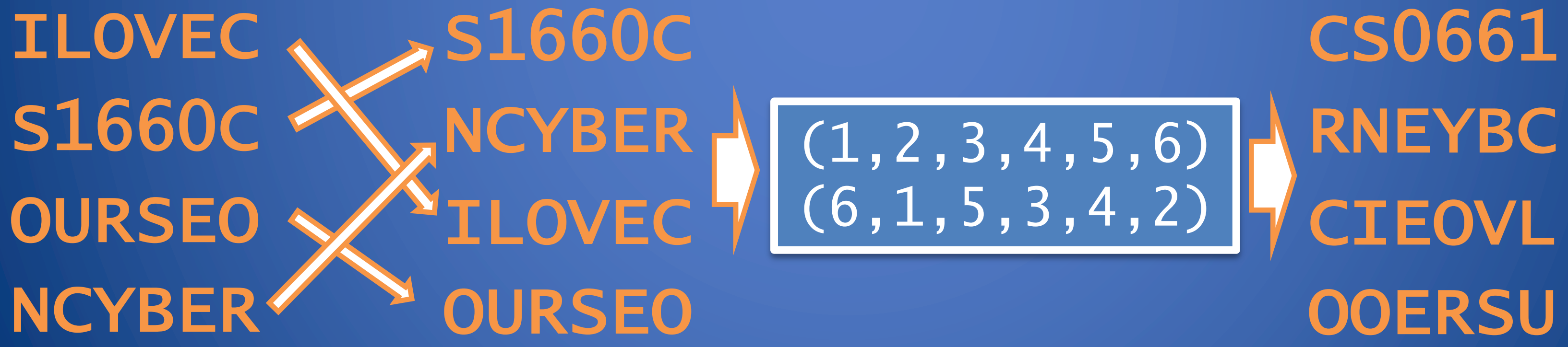
The security of the algorithm must reside in the key; the security should not depend on the secrecy of the algorithm.

Transposition Cipher

- Instead of replacing the characters with other characters, this cipher alters the order of the characters.
- The key determines the positions that the characters are moved to
  - Instead of a list of alphabetic substitutions, it is a mapping order
  - Such as (1,2,3,4,5,6) = (6, 1, 5,3,4,2)
  - Example: CS1660 -> 0C616S

Permutation

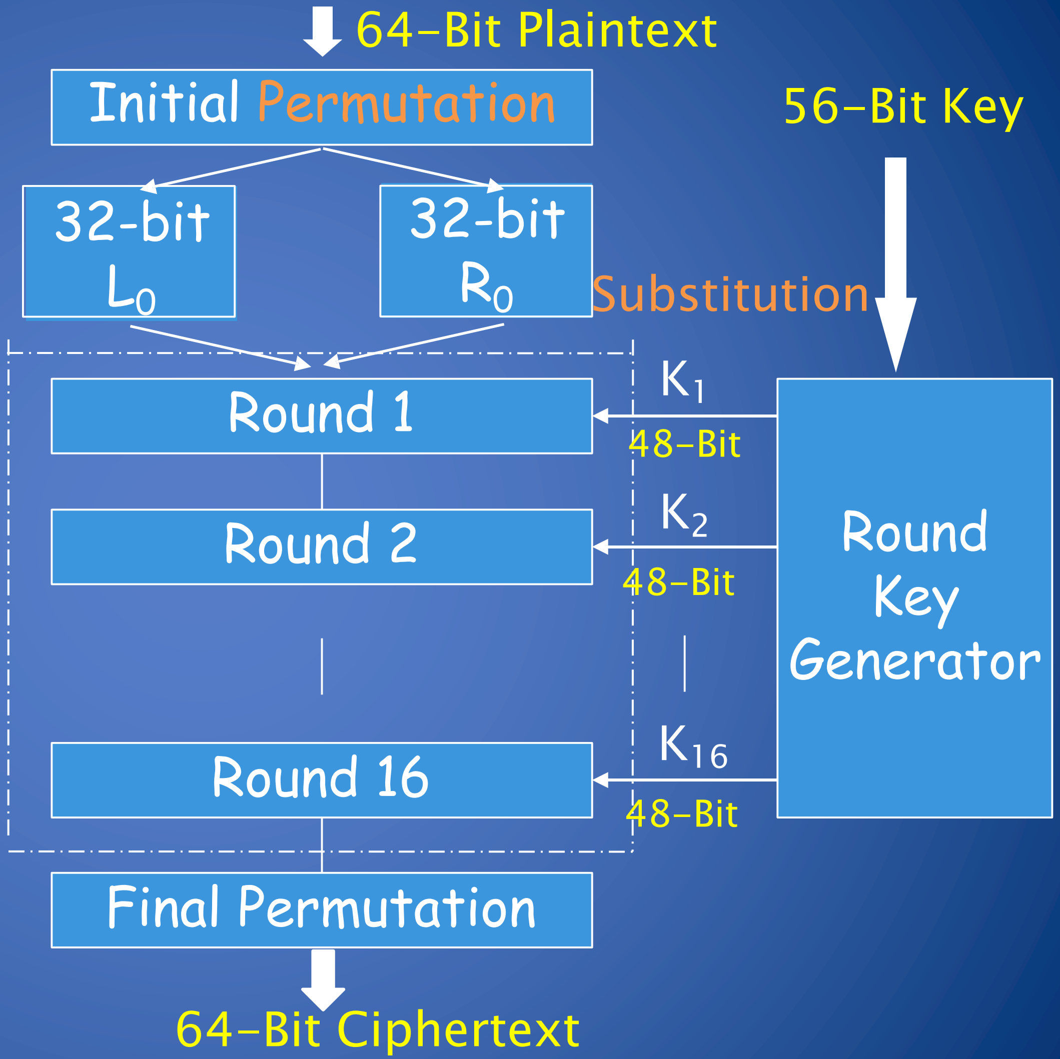
- The permutation of this cipher runs in the rows and then in the columns of a matrix.
- This means that the message is spread out into a matrix.
- Example: `I LOVE CS1660 COURSE ON CYBER`



DES Structure

- DES is a block cipher operating on 64-bit blocks
- Split in two parts
- The Key is 56-bit
  - total of 2^{56} possible keys
- Encryption process:
  - 16 rounds of permutation and substitution

ensuring data security through confusion and diffusion.

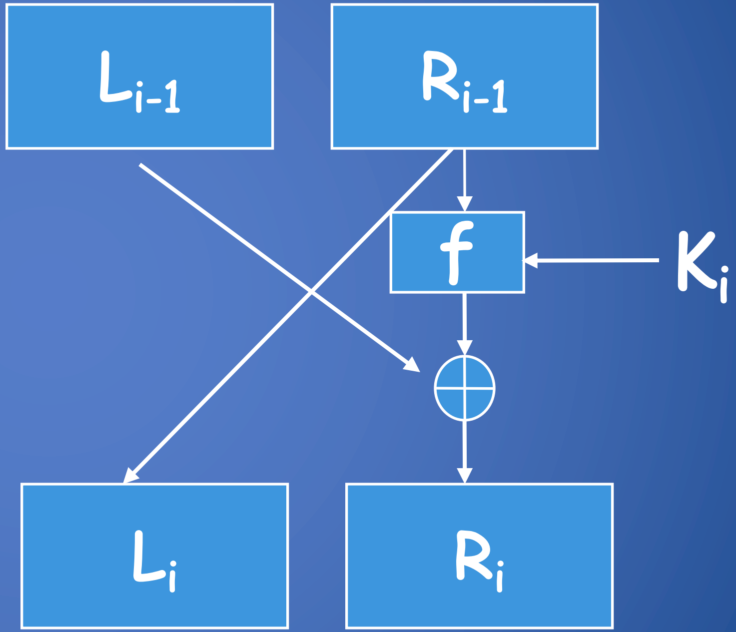


Single Round

- K_i is a subkey
- L_i , R_i(32 bit)

(Left and Right of a block)

- Each round has the same function f
  - –key transformation
  - –expansion permutation
  - –s-box substitution
  - –p-box permutation
  - – XOR and swapping



DES Challenge

DES developed by IBM with suggestions by NSA:

- Originally, the key was 64-bit instead of 56-bit
- The S-Box was changed by the NSA and not made publicly available

The challenge was proposed by RSA to test the strength of DES against brute-force attacks

- Electronic Frontier Foundation (EFF) and others participated
- Using specially designed hardware or collaborative computing
  - **DES I (1997):** First successful brute-force attack against DES (prize of 10k $)
  - **DES II (1998):** Demonstrated the decreasing cost and time to break DES
  - **DES III (1999):** Final challenge, broken in just 22 hours

Proved that DES was vulnerable to brute-force attacks and led to a stronger encryption standard like AES (Advanced Encryption Standard) with a public call

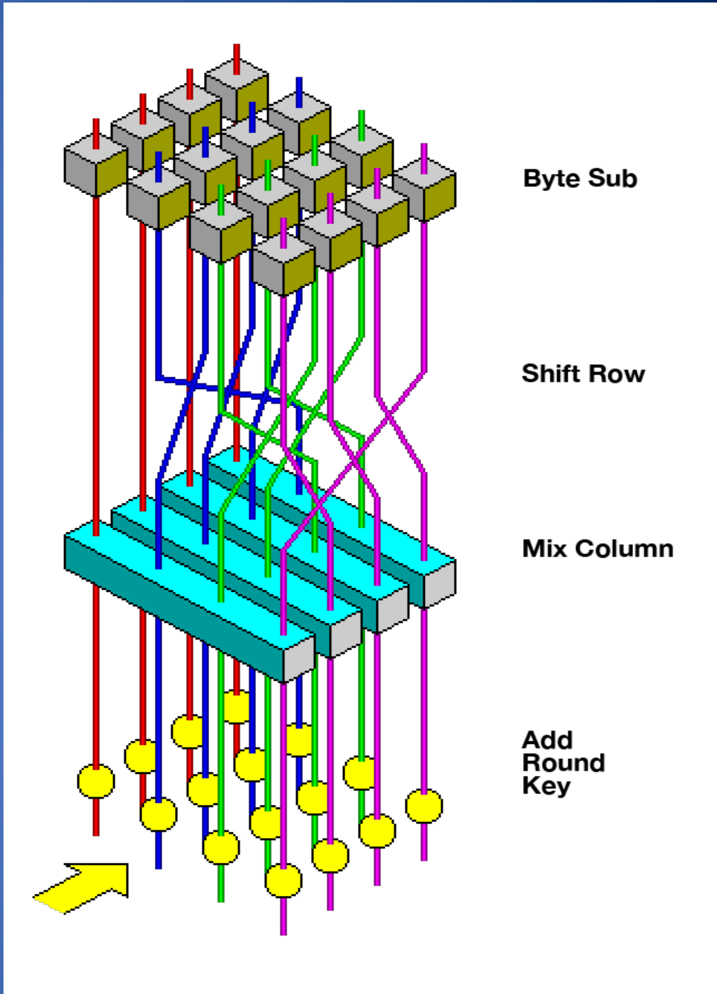
Image source: https://www.nsa.gov/resources/everyone/digital-media-center/image-galleries/places/
https://pt.wikipedia.org/wiki/EFF_DES_cracker#/media/Ficheiro:Chip300.jpg

Advanced Encryption Standard (AES)

NIST competition started in 1997:

- an unclassified, publicly disclosed encryption algorithm capable of protecting sensitive government information well into the next century
- AES shall be available on a worldwide, non-exclusive, royalty-free basis
- Mainly an academic competition

AES supports keys of length 128, 192, and 256 bits



Source:wikimedia

Big Numbers in the real world

- Odds for all 5 numbers + Powerball
  - $292 \times 10^{6} \Rightarrow 2^{38}$
- The Age of the Universe in Seconds
  - $4.3 \times 10^{17} \Rightarrow 2^{58}$ https://81018.com/universeclock/
- # of cycles in a century of a 4 GHz CPU => 2^{64}
- # of arrangements of a Rubik's cube $4.3 \times 10^{19} \Rightarrow 2^{65}$
- Atoms in the Earth $1.33 \times 10^{50} \Rightarrow 2^{166}$
- Electrons in the universe $10^{80} \Rightarrow 2^{266}$

Modern Symmetric Encryption

Algorithm	Year of Introduction	Key Space
DES (Data Encryption Standard)	1977	2^{56}
Blowfish	1993	2^32 to 2^448
Twofish	1998	2^128 to 2^256
Serpent	1999	2^128 to 2^256
RC4 (Stream cipher)	1987	2^40 to 2^2048
RC5	1994	up to 2^2040
RC6	1998	2^128 up to 2^2040
CAST-128 (GPG and PGP)	1996	2^40 to 2^128
AES 128, AES 192, AES 256 (Advanced Encryption Standard)	2001	2^128 2^192 2^256

What We Have Learned

- Security goals and attacks on communication
- Frequency analysis defeats classic encryption
- Modern symmetric encryption
  - Stream cipher: one-time pads and the importance of randomness
  - Block cipher: Confusion and Diffusion, Vigenère, Transposition, DES
- Use AES (not DES) for symmetric encryption