FUNDAMENTALS OF INFORMATION SCIENCE:

PART 4: SECURITY

Shandong University
2025 Spring

Lecture 4.1: What is Cryptography?

Cryptography is Everywhere

Secure communication:

- web traffic: HTTPS
- wireless traffic: 802.11i WPA2 (and WEP), GSM, Bluetooth

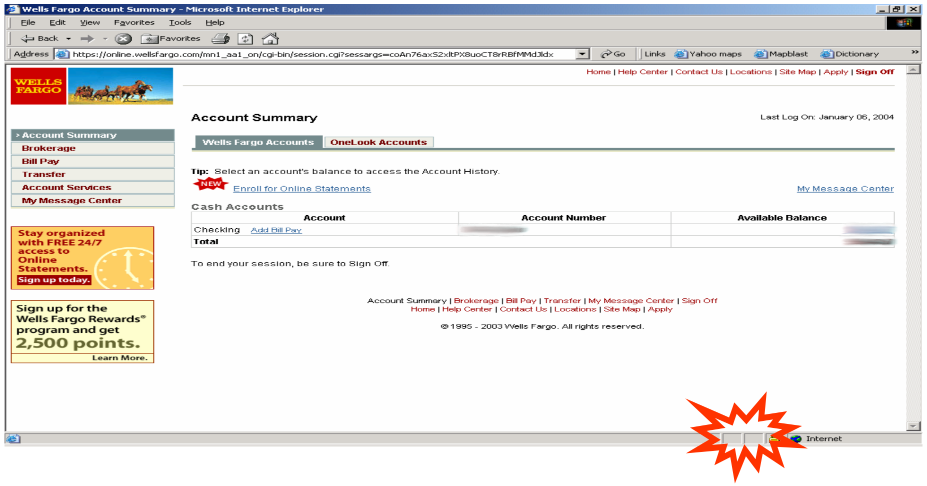
Encrypting files on disk: EFS, TrueCrypt

Content protection (e.g. DVD, Blu-ray): CSS, AACS

User authentication

… and much much more

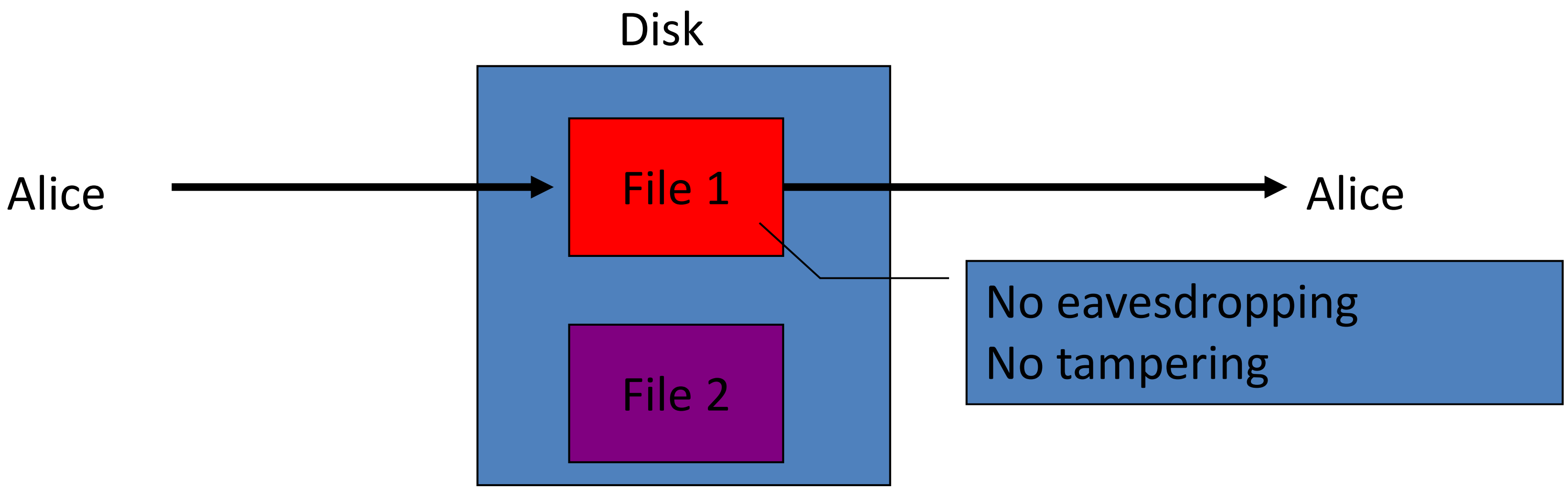
Secure Communication



HTTP

no eavesdropping
no tampering

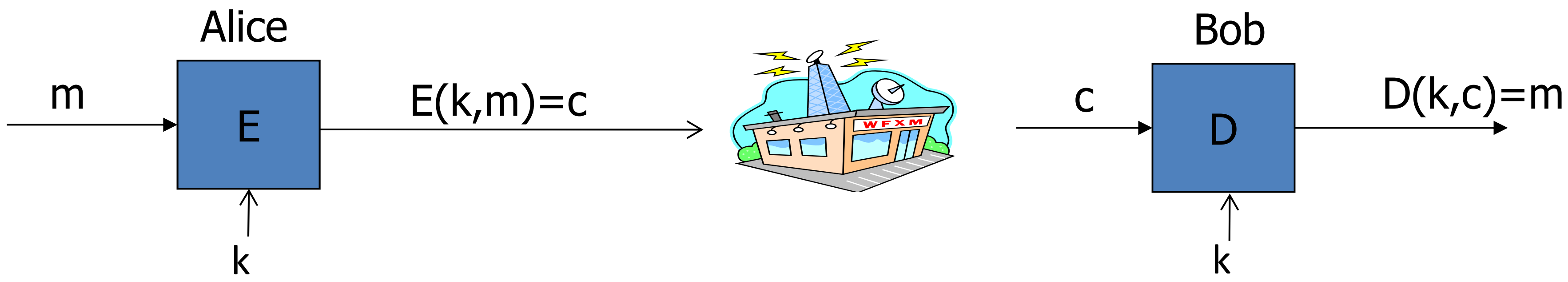
Protect Files on Disk



Analogous to secure communication:

Alice today sends a message to Alice tomorrow

Building Block: Symmetric Encryption



E, D: cipher   k: secret key (e.g. 128 bits)

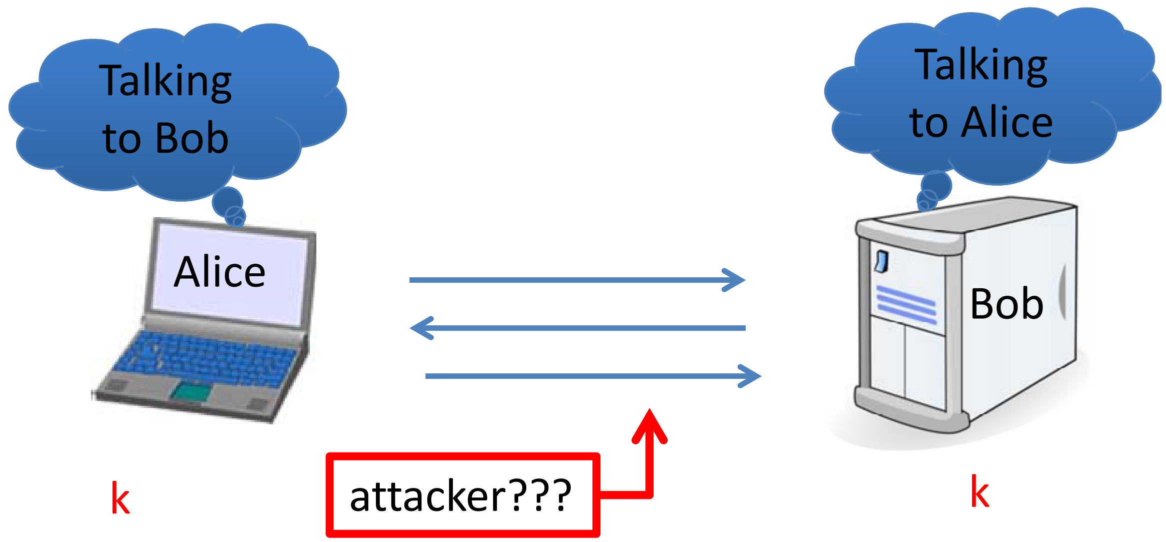
m, c: plaintext, ciphertext

Encryption algorithm is publicly known

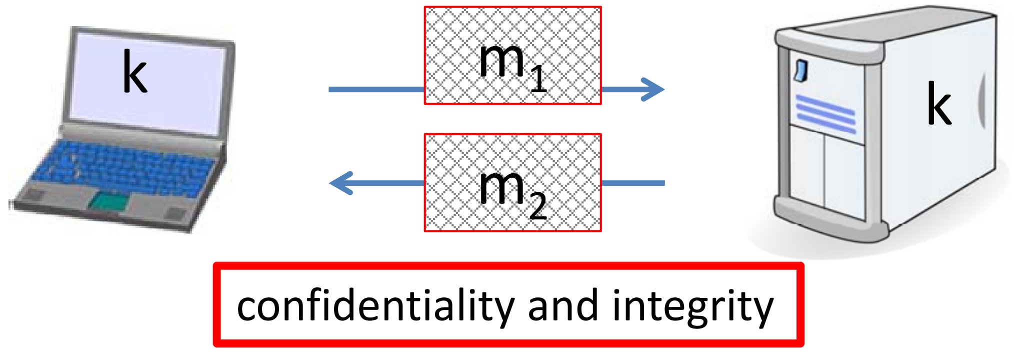
- Never use a proprietary cipher

Crypto Core

Secret key establishment:



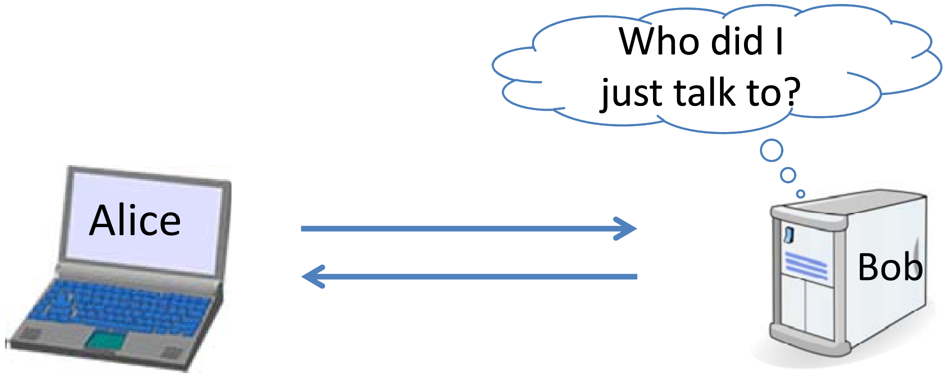
Secure communication:



But Crypto can Do Much More

Digital signatures

Anonymous communication



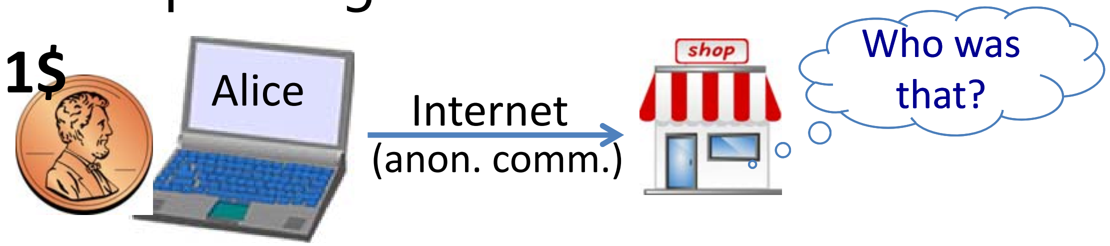
But Crypto can Do Much More

Digital signatures

Anonymous communication

Anonymous **digital** cash

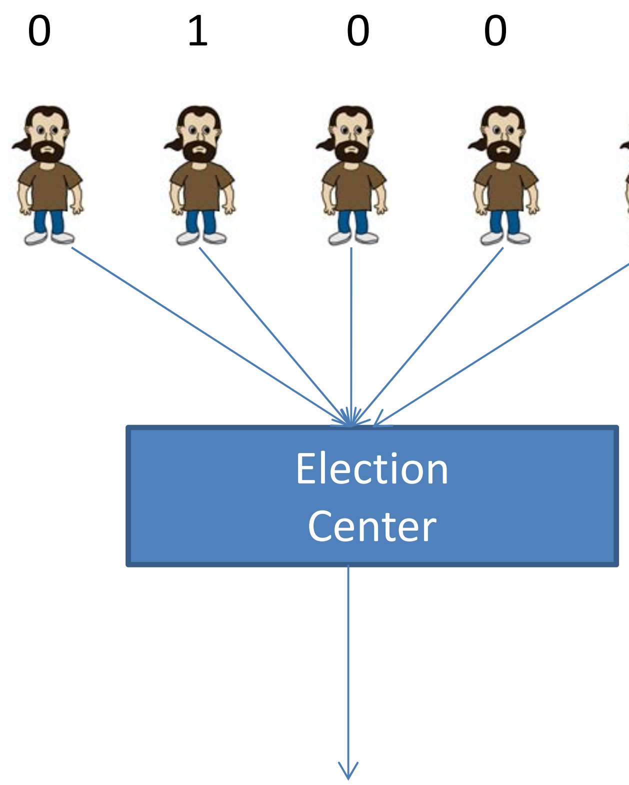
- Can I spend a "digital coin" without anyone knowing who I am?
- How to prevent double spending?



Protocol

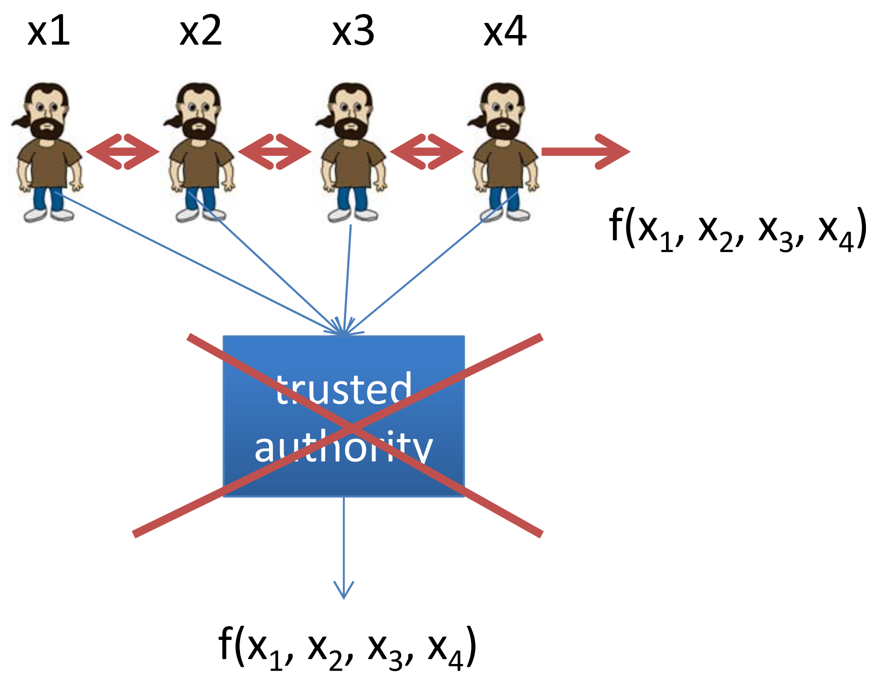
Elections

Winner = MAJ [votes]



Protocol

Elections

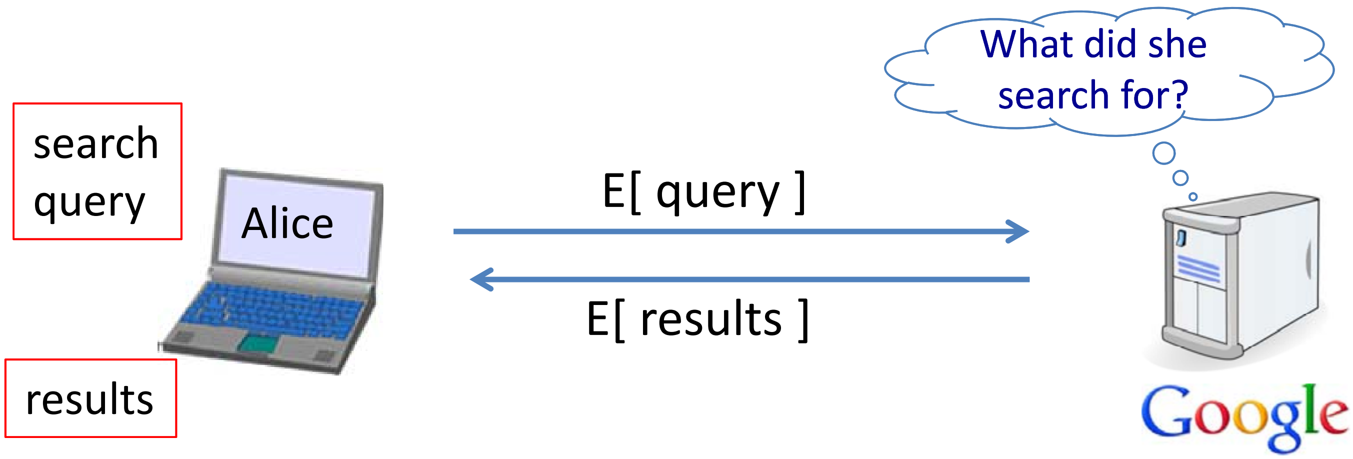


Goal: compute $f(x_1, x_2, x_3, x_4)$

"Thm:" anything that can done with trusted auth. can also be done without

Crypto Magic

Privately outsourcing computation



A rigorous science

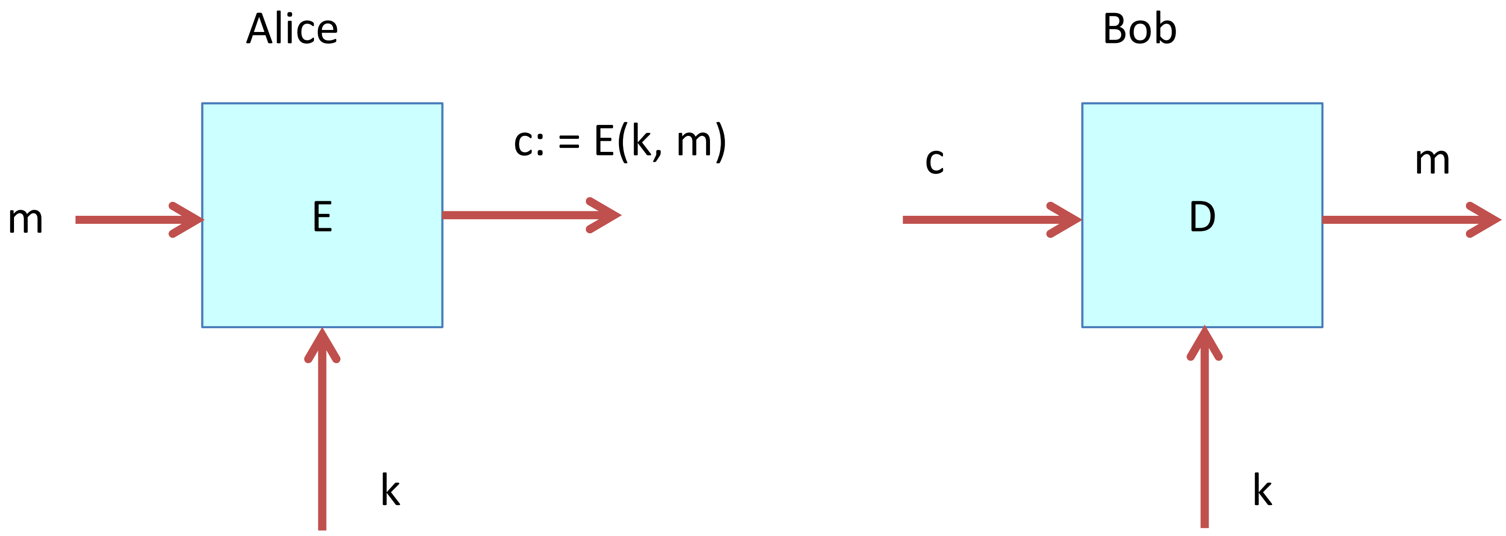
The three steps in cryptography:

Precisely specify threat model

Propose a construction

Prove that breaking construction under
threat mode will solve an underlying hard problem

Symmetric Cipher



the same key

Few Historic Examples (all badly broken)

1. Substitution cipher

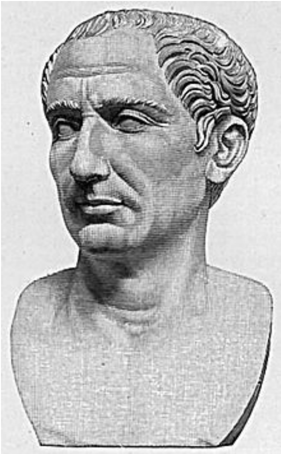
C: = E(k, “bcza”) = “wnac”

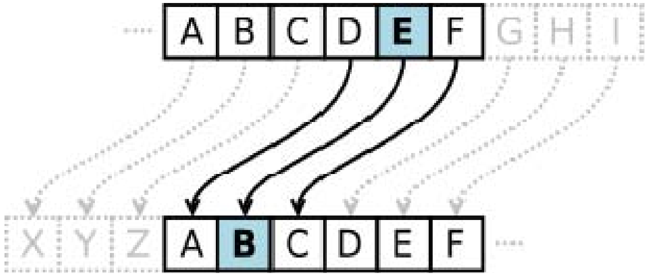
D(k, c) = “bcza”

k :=

a → c
b → w
c → n
…
z → a

Caesar Cipher (no key)





Shift by 3

How to break a substitution cipher?

What is the most common letter in English text?

"X"

"L"

"E"

"H"

How to break a substitution cipher?

(1) Use frequency of English letters

e: 12.7% **t: 9.1%** **a:8.1%**

(2) Use frequency of pairs of letters (digrams)

he **an** **in** **th**

An Example

UKBYBIPOUZBCUFEEBORUKBYBHOBBRFESPVKBWFOFERVNBCVBZPRUBOFERVNBCVBPCYYFVUFO
FEIKNWFRFIKJNUPWRFIPOUNVNIPUBRNCUKBEFWWFDNCHXCYBOHOPYXPUBNCUBOYNRVNIWN
CPOJIOFHOPZRVFZIXUBORJRUBZRBCHNCBBONCHRJZSFWNVRJRUBZRPCYZPUKBZPUNVPWPCYVF
ZIXUPUNFCPWRVNBCVBRPYYNUNFCPWWJUKBYBIPOUZBCUIPOUNVNIPUBRNCHOPYXPUBNCUB
OYNRVNIWNCPOJIOFHOPZRNCRVNBCUNENVVFZIXUNCHPCYVFZIXUPUNFCPWZPUKBZPUNVR

B	**36**	→ E
N	**34**	
U	**33**	→ T
P	**32**	→ A
C	**26**	

NC	**11**	→ IN
PU	**10**	→ AT
UB	**10**	
UN	**9**	

digrams

UKB	**6**	→ THE
RVN	**6**	
FZI	**4**	

trigrams

2. Data Encryption Standard (1974)

DES: # keys = 2^{56} , block size = 64 bits

Today: AES (2001), Salsa20 (2008) (and many others)

Lecture 4.2: One Time Pad

Symmetric Ciphers: Definition

Def: a **cipher** defined over (K, M, C)

is a pair of "efficient" algs (***E***, ***D***) where

$$E: K \times M \rightarrow C \qquad \mathrm{D}: K \times C \rightarrow M$$

s.t

$$\forall\, m \in M, k \in K \qquad D\big(k, E(k, m)\big) = m$$

E is often randomized. ***D*** is always deterministic.

The One-Time Pad **(Vernam 1917)**

First example of a "secure" cipher

$$M = C = \{0,1\}^n \qquad K = \{0,1\}^k$$

key = (random bit string as long the message)

The One-Time Pad (Vernam 1917)

msg: 0 1 1 0 1 1 1

key: 1 0 1 1 0 1 0 $\oplus$

CT:

$$c := E(k, m) = k \oplus m$$

$$D(k, c) = k \oplus \mathrm{c}$$

Indeed:

$$D(k, E(k, m)) = k \oplus k \oplus m = m$$

The One-Time Pad (Vernam 1917)

You are given a message (m) and its OTP encryption (c).

Can you compute the OTP key from m and c ?

No, I cannot compute the key.

Yes, the key is $k = m \oplus c$.

I can only compute half the bits of the key.

Yes, the key is $k = m \oplus m$.

The One-Time Pad (Vernam 1917)

Very fast enc/dec !!

… but long keys (as long as plaintext)

Is the OTP secure? What is a secure cipher?

What is a secure cipher?

Attacker's abilities: **CT only attack** (for now)

Possible security requirements:

attempt #1: **attacker cannot recover secret key**

attempt #2: **attacker cannot recover all of plaintext**

Shannon's idea:

CT should reveal no "info" about PT

Information Theoretic Security (Shannon 1949)

Def: A cipher ***(E,D)*** over (K,M,C) has **perfect secrecy** if

$\forall m_0, m_1 \in M \quad (|m_0| = |m_1|) \quad \text{and} \quad \forall c \in C$

$$Pr[\ E(k,m_0)=c\ ] \ = \ Pr[\ E(k,m_1)=c\ ]$$

where k is uniform in K, denoted by $k \longleftarrow K$

---

1. Given CT cannot tell if message is m_0 or m_1 (for all m_0, m_1)
2. Most powerful adversary learns nothing about PT from CT
3. No CT only attack!

Information Theoretic Security (Shannon 1949)

<u>Lemma</u>: OTP has perfect secrecy.

Proof:

$$\forall m, c, \quad \boldsymbol{Pr[\, E(k,m)=c \,] = \frac{\#\text{ keys } k \in K, \text{ s.t. } E(k,m)=c}{|K|}}$$

$$\boldsymbol{\#\, \{k \in K, \text{ s.t. } E(k,m)=c\,\} \quad = \quad \text{const.}}$$

→ cipher has perfect security

Information Theoretic Security (Shannon 1949)

Lemma: OTP has perfect secrecy.

Proof:

$$\forall m, c, \quad \Pr[\, E(k,m)=c \,] = \frac{\#\text{ keys } k \in K, \text{ s.t. } E(k,m)=c}{|K|}$$

$$\#\{k \in K, \text{ s.t. } E(k,m)=c\} \quad = \quad \text{const.} = 1$$

→ cipher has perfect security

The bad news

Thm: perfect secrecy $\Rightarrow$ $|\mathcal{K}| \geq |\mathcal{M}|$

i.e. key length ≥ message length

had to use in practice

Another Definition

perfect secrecy $\Rightarrow$

The mutual information between ciphertexts and plaintexts is almost 0:

$$I\,[m, c] \to 0$$

i.e. ciphertexts do not imply any information about plaintexts.

What is secure cipher?

Attacker's abilities: **obtains one ciphertext** (for now)

Possible security requirements:

attempt #1: **attacker cannot recover secret key**

attempt #2: **attacker cannot recover all of plaintext**

Recall Shannon's idea:

CT should reveal no "info" about PT

Recall Shannon's perfect secrecy

Let (E,D) be a cipher over (K,M,C)

(E,D) has perfect secrecy if $\forall\ m_0, m_1 \in M \quad (\ |m_0| = |m_1|\ )$

$$\{\ E(k,m_0)\ \} \quad = \quad \{\ E(k,m_1)\ \} \qquad \text{where} \quad k \longleftarrow K$$

(E,D) has semantic secrecy if $\forall\ m_0, m_1 \in M \quad (\ |m_0| = |m_1|\ )$

$$\{\ E(k,m_0)\ \} \ \approx_p\ \{\ E(k,m_1)\ \} \qquad \text{where} \quad k \longleftarrow K$$

... but also need adversary to exhibit $m_0, m_1 \in M$ explicitly