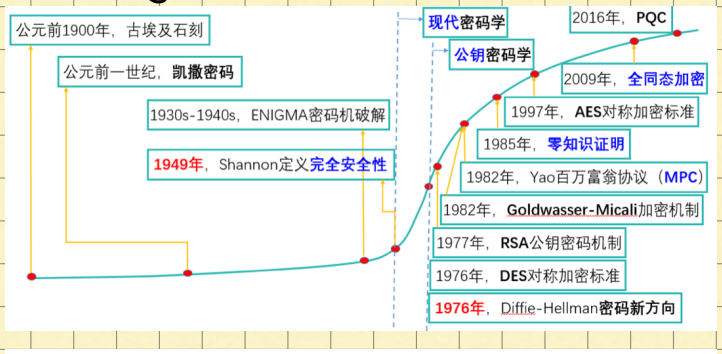
History



- History (until 1970's)
  - ◇ Secret code invented
  - ◇ Typically claimed "unbreakable" by inventor
  - ◇ Used by spies, ambassadors, kings, generals for crucial tasks
  - ◇ Broken by adversaries using cryptanalysis

"*invent-break-tweak*" cycle

古典密码学的密码算法总是在发明-破解这个循环中.

Modern Cryptography (Post 1970's)

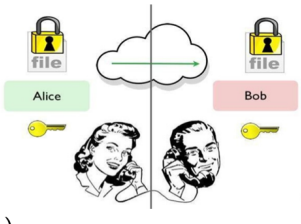
可证明安全

- Provable security
  - ◇ Perfect security (Shannon) and its limitations
  - ◇ Computational security
  - ◇ Pseudorandom generators, one-way functions
- Beyond encryption
  - ◇ Public-key encryption based on factoring, RSA problem
  - ◇ Digital signatures, hash functions
  - ◇ Zero-knowledge proofs
- Advanced topics*
  - ◇ The SSL protocol
  - ◇ Multi-party secure computation
  - ◇ Post-quantum cryptography
  - ◇ Fully homomorphic encryption (Gentry 2009), ...

Encryption Schemes

- Alice wants to send Bob a secret message

They agree in advance on 3 components:
- ◇ Encryption algo.: E
- ◇ Decryption algo.: D
- ◇ Secret key: k



Encryption: Alice $\xrightarrow{c = \mathrm{E}(m,k)}$ Bob

Decryption: Bob computes $m' = \mathrm{D}(c,k)$

A scheme is valid if $m' = m$.
Intuitively, a scheme is secure if an eavesdropper cannot learn m from c.

若窃听者不能从密文 m 中得出明文 C，则这个模式是安全的.

Caesar Cipher 凯撒密码

- Key: $k = 0, 1, \ldots, 25$

  Encryption: encode i as $(i + k)$ mod 26

  Decryption: decode j as $(j - k)$ mod 26

  plaintext: SEND REINFORCEMENT
  Key: 2
  ciphertext: UGPF TGKPHQTEGOGPV

  Problem: only 26 possibilities for keys!

  Kerchoff's Principle (1883): System should be secure even if algorithms are known, as long as key is secret.

字母向前/后位移.

Kerchoff原理:
当密钥是隐秘时,即便算法已知,系统也是安全的.

Substition Cipher 替换密码

- Key: table mapping each letter to another letter

A	B	C																						Z
V	R	E																						D

  Encryption & Decryption: letter by letter according to table

  # of possible keys: $26! \approx 4 \times 10^{26}$

  However, substitution cipher is still insecure!

  Key observation: can recover plaintext using *statistics* on *letter frequencies*.

使用概率分析法

Example

Table 1: Relative frequencies of the letters of the English language

Letter	Relative Frequency (%)	Letter	Relative Frequency (%)
a	8.167	n	6.749
b	1.492	o	7.507
c	2.782	p	1.929
d	4.253	q	0.095
e	12.702	r	5.987
f	2.228	s	6.327
g	2.015	t	9.056
h	6.094	u	2.758
i	6.966	v	0.978
j	0.153	w	2.360
k	0.772	x	0.150
l	4.025	y	1.974
m	2.406	z	0.074

Table 2: Number of Diagraphs Expected in 2,000 Letters of English Text

th	50	at	25	st	20
er	40	en	25	io	18
on	39	es	25	le	18
an	38	of	25	is	17
re	36	or	25	ou	17
he	33	nt	24	ar	16
in	31	ea	22	as	16
ed	30	ti	22	de	16
ne	30	to	22	rt	16
ha	26	it	20	ve	16

Table 3: The 15 Most Common Trigraphs in the English Language

1 - the	6 - tio	11 - edt
2 - and	7 - for	12 - tis
3 - tha	8 - nde	13 - oft
4 - ent	9 - has	14 - sth
5 - ion	10 - nce	15 - men

- LIVITCSWPIYVEWHEVSRIQMXLEYVEOIEWHRXEXIPFEMVEWHKVSTYLXZIXLIKIIXPIJVSZEYPERRGERIMWQLMGLMXQERIWGPSRIHMXQEREKI

  I – *most common letter* — I = e
  LI – *most common pair* — L = h
  XLI – *most common triple* — X = t

  LIVI = he?e — V = r, E = a, Y = g

  HereUpOnLeGrandAroseWithAGraveAndStatelyAirAndBroughtMeTheBeetleFromAGlassCaseInWhichItWasEnclosedIt-WasABe

Vigenere Cipher 维吉尼亚密码

- "*Multi-Caesar Cipher*" – stateful

  Key: $\mathbf{k} = (k_1, k_2, \ldots, k_m)$ – list of m numbers in [0..25]

  Encryption: encode i as $(i + k_j)$ mod 26, if the location of i is j mod m

  Decryption: decode j as $(j - k_i)$ mod 26, if the location of j is i mod m

  Important: Cannot break using letter frequencies alone. Because the same letter e may be mapped to $e + k_1, e + k_2, \ldots, e + k_m$ depending on different locations.

  Considered as "unbreakable" for 300 years (broken by Babbage, Kasiski 1850's)

对每一位的加密都用不一样的key的凯撒密码.

Example

- Breaking Vigenere:

LIVITCSWPIYVEWHEVSRIQMXLEYVEOIEWHRXEXIPFEMVEWHK

Step 1: **Guess** the length of the key m

Step 2: Group together positions $\{1, m+1, 2m+1, \ldots\}$, $\{2, m+2, 2m+2, \ldots\}$, $\cdots$, $\{m-1, 2m-1, 3m-1, \ldots\}$

Step 3: Frequency-analyze each group independently.

LIVITC
SWPIYV
EWHEVS
RIQMXL
EYVEOI
EWHRXE
XIPFEM
VEWHKV

首先猜出循环长度m.

36 5

Questions

$\mathcal{Q}$: Can Bob send Alice the secret key over the channel?

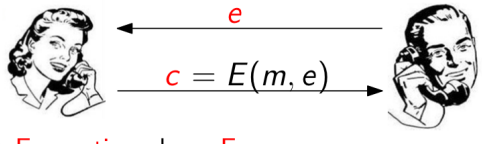
$\mathcal{A}$: Of course not! Eve could decrypt c.

信道上不能传密钥.

$\mathcal{Q}$: What if Bob could send Alice a "special key" useful only for encryption but no help for decryption?

但可传加密密钥

- Alice wants to send Bob a secret message



Bob将加密密钥e明文发给Alice, Alice将信息用e加密后发给Bob, Bob用解密密钥解密.

◇ Encryption algo.: E
◇ Decryption algo.: D
◇ Key: Bob chooses two keys: secret key d for decryption and public key e for encryption.

Decryption: Bob computes $m' = \mathrm{D}(c, d)$

A scheme is valid if $m' = m$.
Intuitively, a scheme is secure if eavesdropper cannot learn m from c (even if Eve knows the key e).

不对称加密,即加密密钥与解密密钥不同.

Cryptography Wonders

- *Digital Signatures*. Electronically sign documents

  Zero-knowledge Proofs. Alice proves to Bob that she earns $< \$50k$ without Bob learning her income.

  Privacy-perserving data mining. Bob holds DB. Alice gets answer to one query, without Bob knowing what she asked.

  Playing poker over the net. Alice, Bob, Carol and David can play Poker over the net without trusting each other or any central server. (*E-Voting*)

  Electronic Auctions. Can run auctions s.t. no one (even not seller) learns anything other than winning party and bid.

  Fully Homomorphic Encryption. Encrypt $E(m)$ in a way that allows to compute $E(f(m))$.

Principles of Modern Cryptography

- Principle 1 – *Formal Definitions* 形式化定义
  - Precise, mathematical model and definition of what security means
- Principle 2 – *Precise Assumptions* 精确假设
  - Clearly stated and unambiguous
- Principle 3 – *Proofs of Security* 安全性证明
  - Move away from "design-break-tweak"