Basics of Cryptography



What we have learned

- Principles of computer security
 - Confidentiality
 - Integrity
 - Availability
 - Authenticity
 - Accountability
 - Non-repudiation
- Trust is inevitable somewhere
- Human factor also key in computer security
- Broader issues related to computer security, such as ethics, law, and privacy

Check all answers that apply

I have a file in a UNIX file system with **READ-ONLY** access right. If someone steals my credential to log into the file system, what security principles can be violated?

- A: Confidentiality
- B: Availability
- C: Integrity
- D: Authenticity



Check all answers that apply

I have a file in a UNIX file system with **READ-ONLY** access right. If someone steals my credential to log into the file system, what security principles can be violated?

A: Confidentiality

B: Availability

C: Integrity

D: Authenticity



Crypto Basics



Crypto Terms

- Crypto from Greek "krypto" = hide
- Recently, crypto means Cryptocurrency





- Cryptography making "secret codes"
- Cryptanalysis breaking "secret codes"

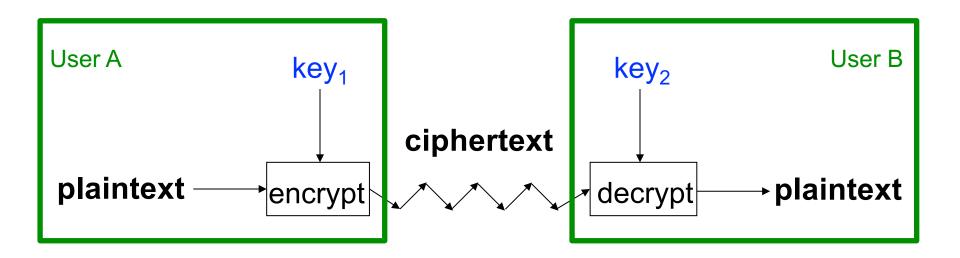


Crypto Terms

- A cipher is used to encrypt the plaintext
 - The *algorithm* to encrypt plaintext
- The result of encryption is ciphertext
- We decrypt ciphertext to recover plaintext
- A key is used to configure a cryptosystem
 - The configuration of the algorithm
- A symmetric key cryptosystem uses the same key to encrypt as to decrypt
- An asymmetric key cryptosystem uses different keys for encryption and decryption



Illustration of a cryptosystem

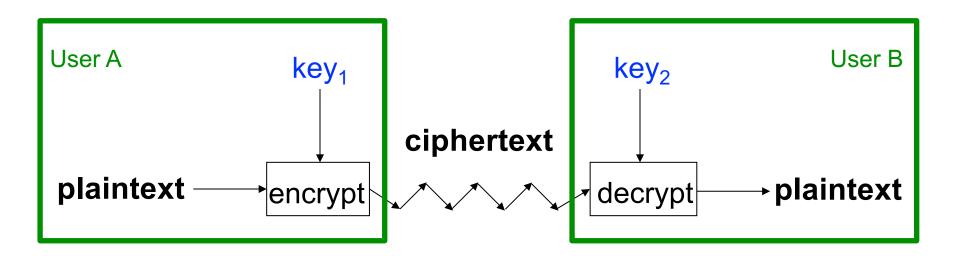


 $Key_1 = Key_2$: Symmetric

 $Key_1 \neq Key_2$: Asymmetric



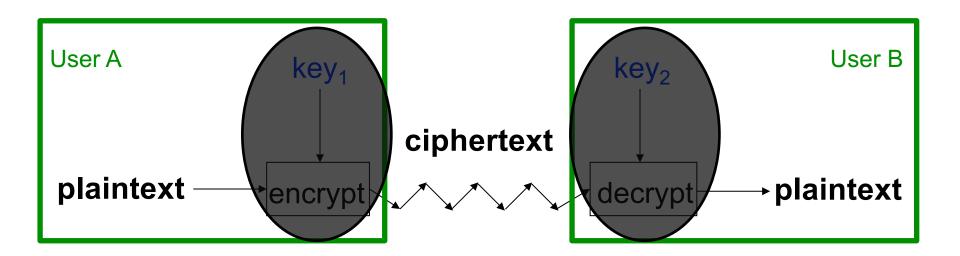
Illustration of a cryptosystem



How to protect the security of this cryptosystem?

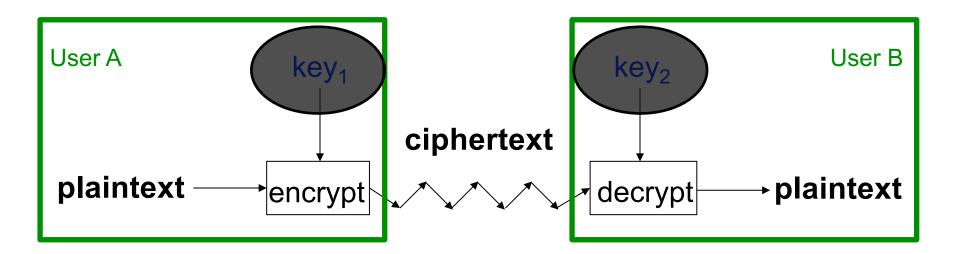


Case 1: both key and cipher are secret





Case 2: key is secret





Kerckhoffs' Principle

- **Kerckhoff's principle:** the cipher method must not be required to be secret, and it must be able to fall into the hands of the enemy without inconvenience.
- Kerckhoff's principle demands that <u>security must rely solely</u> on the <u>secrecy of the keys</u>.
- In other words,
 - Algorithm must **not** be required to be secret
 - Configuration must be secret



Why Kerckhoffs' Principle?

- Easier to maintain key than algorithm
 - Key is short, algorithm is long
- Easier to replace key than algorithm
 - In case that the key is exposed, it is much easier to replace the key than to replace the algorithm (and the software that implements it)
- Easier to use different key with same algorithm
 - In case many pairs of people need to encrypt their communications, it's much easier for all parties to use the same algorithm/program but different keys, than for everyone to use a different algorithm/program
- Algorithm can be reverse engineered
- Algorithm can be improved under many eyeballs



Is Kerckhoffs' Principle enough?

- Is it secure to rely solely on keeping the keys secret when the algorithm is exposed to everyone?
- A fast machine can verify 2⁴⁰ keys per second
 - A keyspace of size 2⁵⁶
 - Keyspace: the total number of possible keys
 - 2¹⁶ seconds = 18 hours for exhaustive key search
 - A keyspace of size 2¹²⁸ (128 bits)
 - More than 9 quintillion(9 $*10^{18}$) years
 - Impractical for exhaustive key search

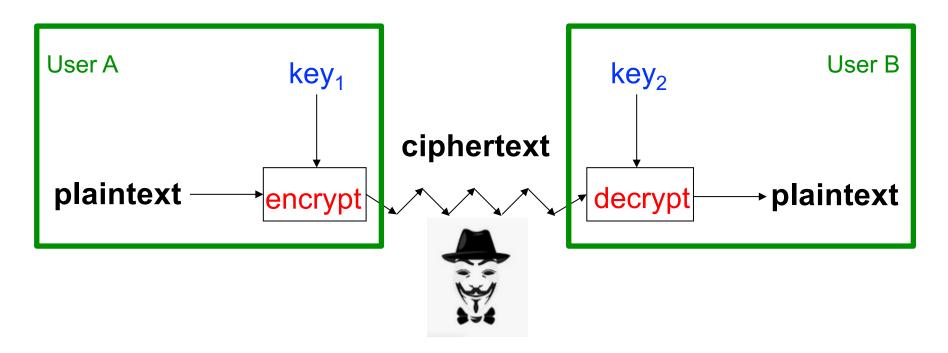


Cryptanalysis: Terminology

- A cryptosystem is secure if best known attack is to try all keys
- A cryptosystem is insecure if any shortcut attack is known
- Secure Cipher: mathematically proved secure
 - Very few, not practical
- Which is more difficult to break?
 - A secure cipher with a small number of keys
 - An insecure cipher with a large number of keys
- Insecure cipher might be harder to break than a secure cipher!



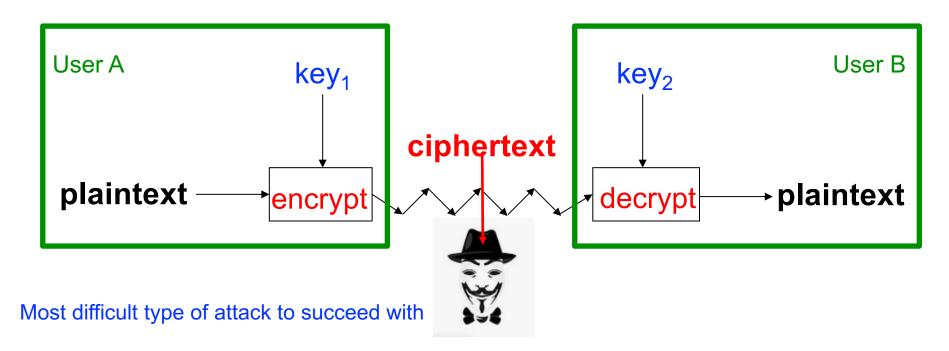
Attack scenarios against a cipher by adversary



Adversary



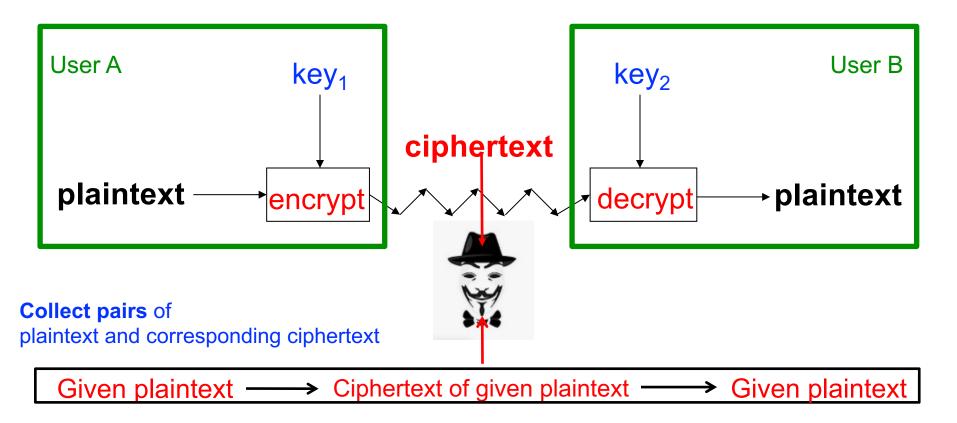
1. Ciphertext Only Attack



Adversary

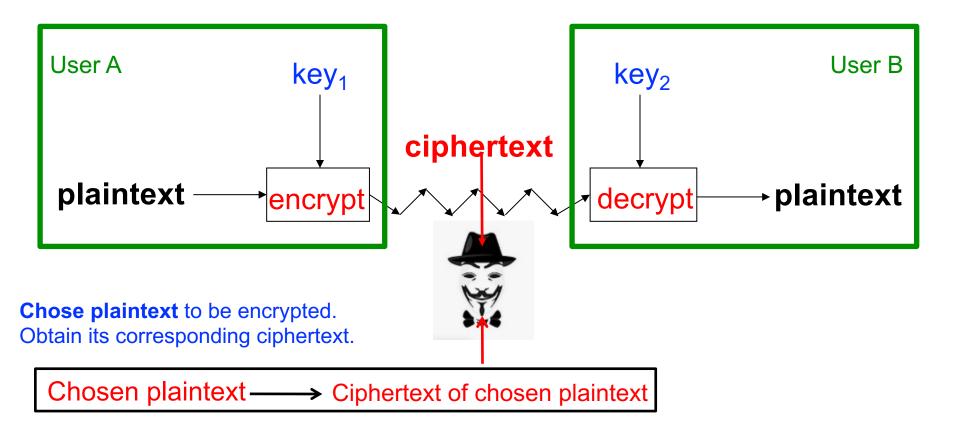


2. Known Plaintext Attack



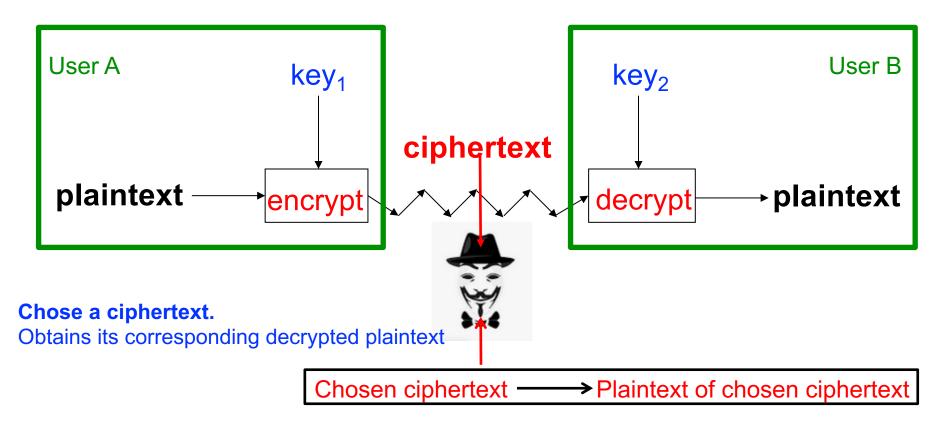


3. Chosen Plaintext Attack



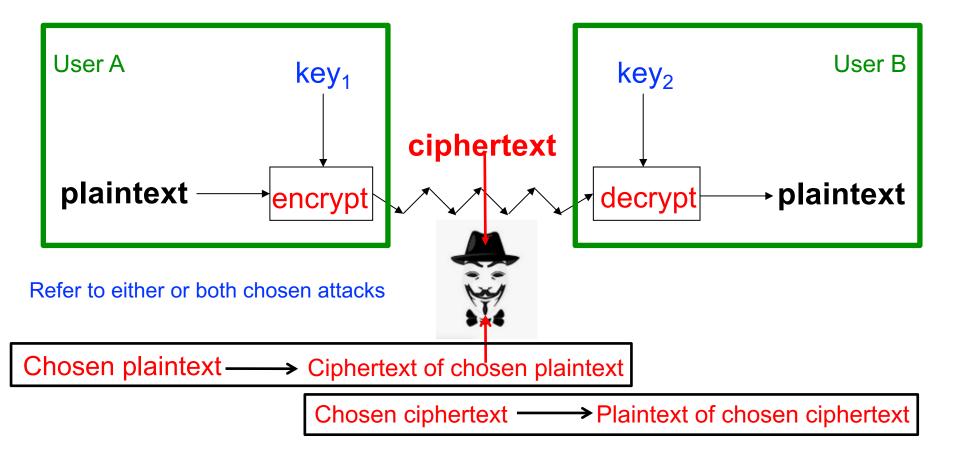


4. Chosen Ciphertext Attack





5. Chosen Text Attack





passive vs. active attacks

- Passive: adversary just receives some ciphertext
- Active: adversary can adaptively ask for encryption/decryption

Among all the attack scenarios, which are active attacks?

- A: Ciphertext only attack
- B: Known plaintext attack
- C: Chosen plaintext attack
- D: Chosen ciphertext attack
- E: Chosen text attack



passive vs. active attacks

- Passive: adversary just receives some ciphertext
- Active: adversary can adaptively ask for encryption/decryption

Among all the attack scenarios, which are active attacks?

- A: Ciphertext only attack
- B: Known plaintext attack
- C: Chosen plaintext attack
- D: Chosen ciphertext attack
- E: Chosen text attack



Historical ciphers



Simple Substitution Cipher

- Plaintext: fourscoreandsevenyearsago
- Encryption: substituting the letter of the alphabet n places ahead of the current letter
 - Shift by n
- Key: 3

Plaintext 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

Ciphertext 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

- Ciphertext: IRXUVFRUHDQGVHYHQBHDUVDJR
- Shift by 3 is "Caesar's cipher"



Caesar's Cipher Decryption

Suppose we know a Caesar's cipher is being used:

Given ciphertext: EHJLQWKHDWWDFNQRZ

Plaintext: begintheattacknow



Not-so-Simple Substitution

- Shift by n for some $n \in \{0, 1, 2, ..., 25\}$
- Then key is n
- Example: key n = 13

Plaintext Ciphertext

a	b	С	d	e	f	9	h	i	j	k	I	m	n	0	p	q	r	S	†	u	٧	W	X	у	Z
Ν	0	Р	Ø	R	5	Т	C	\	W	X	У	Z	A	В	С	D	Ε	F	G	Н	Ι	J	Κ	L	W

This is also called ROT13



Cryptanalysis I: Try Them All

- A simple substitution (shift by n) is used
 - Algorithm is known
 - But the key(n) is unknown
- Given ciphertext: CSYEVIXIVQMREXIH
 - Ciphertext only attack
- How to find the key?
- Only 26 possible keys try them all!
 - Keyspace is 26
- Exhaustive key search
- Solution: key is n = 4
 - youareterminated



Simple Substitution – More Keys

- Is there a way to increase the size of keyspace for simple substitution?
 - Simple substitution key can be any permutation of letters
 - Not necessarily a shift of the alphabet
- For example

Plaintext	a	b	С	d	e	f	9	h	i	j	k	1	m	n	0	р	q	r	S	†	u	٧	W	×	У	z
Ciphertext	J	I	С	Α	X	5	Ε	Y	٧	٥	K	W	В	Q	Т	Z	R	Н	F	M	Ρ	2	J	L	G	0

Then $26! > 2^{88}$ possible keys!

Using previous machine (2^{40}) , it takes around 9 million years to exhaustive key search!

Seems secure to me!



Cryptanalysis II: Be Clever

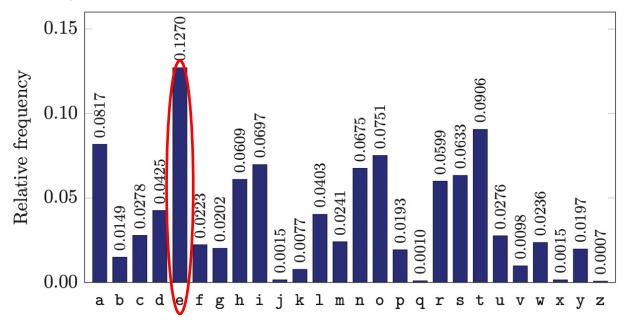
- We know that a simple substitution is used
 - We know the algorithm
- But not necessarily a shift by n
- Find the key given the ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQWAX BVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLXQWAEBIPBFXFQVXGTVJ VWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPFHXZHVF AGFOTHFEFBQUFTDHZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWAQJJTO DXQHFOQPWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTHPB QPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACFCCFHQW AUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQAITIXPFHXAF QHEFZQWGFLVWPTOFFA



Cryptanalysis II – Statistical Attack

- Cannot try all 2⁸⁸ simple substitution keys
- Can we be more clever?
- English letter frequency counts...



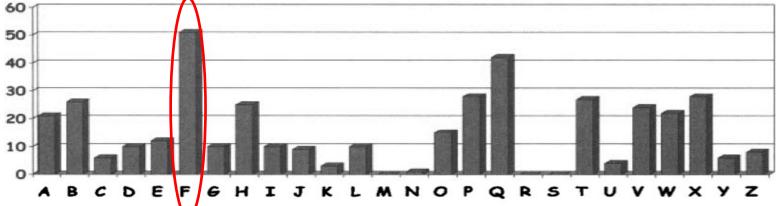


Cryptanalysis II – Statistical Attack

Ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXTOXBTFXQWAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVWLXQWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFEVWLXGDPEQVPQGVPPBFTIXPFHXZHVFAGFOTHFEFBQUFTDHZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQPWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTHPBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIWFFACFCCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWTEVOITDHFHFQAITIXPFHXAFQHEFZQWGFLVWPTOFFA







Cryptanalysis II – Statistical Attack

Ciphertext:

PBFPVYFBQXZTYFPBFEQJHDXXQVAPTPQJKTOYQWIPBVWLXT
OXBTFXQWAXBVCXQWAXFQJVWLEQNTOZQGGQLFXQWAKVW
LXQWAEBIPBFXFQVXGTVJVWLBTPQWAEBFPBFHCVLXBQUFE
VWLXGDPEQVPQGVPPBFTIXPFHXZHVFAGFOTHFEFBQUFTDH
ZBQPOTHXTYFTODXQHFTDPTOGHFQPBQWAQJJTODXQHFOQ
PWTBDHHIXQVAPBFZQHCFWPFHPBFIPBQWKFABVYYDZBOTH
PBQPQJTQOTOGHFQAPBFEQJHDXXQVAVXEBQPEFZBVFOJIW
FFACFCCFHQWAUVWFLQHGFXVAFXQHFUFHILTTAVWAFFAWT
EVOITDHFHFQAITIXPFHXAFQHEFZQWGFLVWPTOFFA

'F' should be 'E'

'the' is a popular starting word
'P' and 'B' relatively frequent in ciphertext
't' and 'h' relatively frequent in plaintext

Guess

'P' is 't'
B' is 'h'



Vigenere Cipher

- Simple substitution is monoalphabetic
 - Each letter is always substitute by another fixed letter
 - 'a' is always substitute by 'D' in Cesar's Cipher
 - A fixed plaintext-ciphertext table can be generated
- Vigenere cipher is simple example of a polyalphabetic substitution
 - Caesars ciphers, based on a keyword
 - Shift based on keyword
- For example, keyword CAT indicates shift by 2, shift by 0, shift by 19
 - Then repeat as needed



Vigenere Example

Suppose that we want to encrypt attackatdawn

Encryption:

keyword: CATCATCAT

plaintext: attackatdawn

ciphertext: ctmccdctwcwg

- Ciphertext is ctmccdctwcwg
- How to decrypt?
 - Reverse the process, shift back according to the keyword
- This obscures the statistical information from the ciphertext



Cryptanalysis III - Vigenere

Plaintext: attackatdawn
Keyword: CATCATCAT
Ciphertext: ctmccdctwcwg

- Suppose we know the keyword length and the ciphertext
 - Length is 3
- We can divide the ciphertext into 3 groups
 - ctmccdctwcwg
 - [C]ccc, [A]tctw, [T]mdwg
- The letter in the same group shift by the same 'n'
 - The same keyword letter is used for that group
- This makes the Vigenere a multiple Caesar's ciphers
- Decrypt each group with the 26 possible keys
 - get frequency of each decrypted group, compare it with the English frequence

Cryptanalysis III - Vigenere

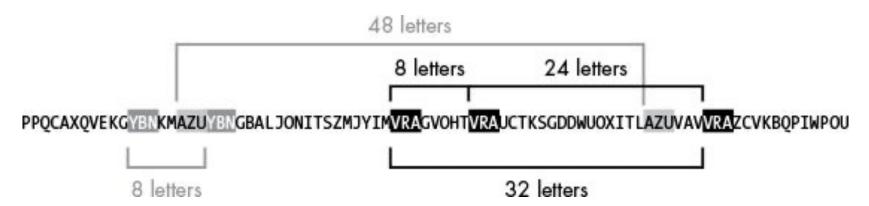
• How to get the keyword length just based on ciphertext?

Plaintext: attackatdawn
Keyword: CATCATCAT
Ciphertext: ctmccdctwcwg

- Found repeat pattern in ciphertext
 - ct
- Guess the length based on pattern
 - **6**
- Consider the factors of length as well
 - **2**, 3
- Try each possible key length



Cryptanalysis III - Vigenere



Length	Factors
8	2, 4, 8
24	2, 4, 6, 8, 12, 24
32	2, 4, 8, 16
48	2, 4, 6, 8, 12, 24, 48

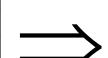
The key is most likely to be the most frequently occurring factors 2, 4, 8 in this example



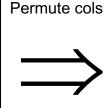
Double Transposition

- Plaintext: attackxatxdawnx
 - 5 x 3 matrix

	col 1	col 2	col 3
row 1	а	t	t
row 2	а	С	k
row 3	Х	а	t
row 4	х	d	а
row 5	W	n	х



		col 1	col 2	col 3
Permute rows	row 3	Х	а	t
	row 5	W	n	х
	row 1	а	t	t
	row 4	X	а	d
	row 2	а	С	k



	col 1	col 3	col 2
row 3	X	t	а
row 5	w	х	n
row 1	а	t	t
row 4	х	а	d
row 2	а	k	С

- Ciphertext: xtawxnattxadakc
- Key is matrix size and permutations: (3, 5, 1, 4, 2) and (1, 3, 2)



Double Transposition - Decryption

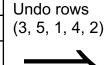
- Ciphertext: xtawxnattxadakc
 - 5 x 3 matrix

	col 1	col 3	col 2
row 3	Х	t	а
row 5	W	х	n
row 1	а	t	t
row 4	Х	а	d
row 2	а	k	С





	col 1	col 2	col 3
row 3	Х	а	t
row 5	w	n	x
row 1	а	t	t
row 4	Х	а	d
row 2	а	С	k



•	•
	_
	/

	col 1	col 2	col 3
row 1	а	t	t
row 2	а	С	k
row 3	х	а	t
row 4	х	d	а
row 5	W	n	Х

- Plaintext: attackxatxdawnx
- Does not disguise the letters



XOR

- XOR (Exclusive OR) is a logical operation that outputs true only when the inputs differ
- Symbol: ⊕
- Truth table

Input A	Input B	Output
0	0	0
0	1	1
1	0	1
1	1	0

XOR properties

$$A \oplus 0 = A$$

$$A \oplus A = 0$$
,

$$(A \oplus B) \oplus C = A \oplus (B \oplus C),$$

$$(B \oplus A) \oplus A = B \oplus 0 = B,$$

■ E.g., 1010 ⊕ 1100 = 0110



```
Alphabet
```

```
e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111
```

"xor"

Encryption: Plaintext \oplus Key = Ciphertext

 h
 e
 i
 l
 h
 i
 l
 l
 e
 r

 Plaintext:
 001
 000
 010
 100
 001
 010
 111
 100
 000
 101
 110
 000

 Key:
 111
 101
 110
 101
 110
 110
 101
 110
 101
 101
 101

 Ciphertext:
 110
 101
 100
 001
 110
 110
 111
 001
 110
 101

 s
 r
 l
 h
 s
 s
 t
 h
 s
 r

Encryption



```
Alphabet
```

```
e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111
```

Decryption: Ciphertext ⊕ Key = Plaintext

Plaintext ⊕ Key ⊕ Key = Plaintext h S S Ciphertext: 110 101 100 001 110 110 111 001 110 101 111 100 000 101 Plaintext. 001 000 010 100 001 010 111 100 h h

Decryption



Double agent claims sender used following "key"

```
S r l h S S t h S r

Ciphertext: 110 101 100 001 110 110 111 001 110 101

"key": 101 111 000 101 111 100 000 101 110 000

"Plaintext": 011 010 100 100 001 010 111 100 000 101

k i l l h i t l e r

e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111
```



Or sender is captured and claims the key is...



One-Time Pad Summary

- Provably secure...
 - Secure cipher
 - Ciphertext provides no info about plaintext
 - All plaintexts are equally likely
- ...but, only when be used correctly
 - Pad must be random, used only once
 - Pad is known only to sender and receiver
- pad (key) is same size as message
 - Impractical for some real-world cases
- You can already distribute secure pad, why not distribute message itself instead?



Real-World One-Time Pad

- Project <u>VENONA</u>
 - Encrypted spy messages sent from U.S. to Moscow in 30's, 40's, and 50's
 - Nuclear etc.
 - Thousands of messages
- Spy carried one-time pad into U.S.
- Spy used pad to encrypt secret messages
- Repeats within the "one-time" pads made cryptanalysis possible



VENONA Decrypt (1944)

[C% Ruth] learned that her husband [v] was called up by the army but he was not sent to the front. He is a mechanical engineer and is now working at the ENORMOUS [ENORMOZ] [vi] plant in SANTA FE, New Mexico. [45 groups unrecoverable]

detain VOLOK [vii] who is working in a plant on ENORMOUS. He is a FELLOWCOUNTRYMAN [ZEMLYaK] [viii]. Yesterday he learned that they had dismissed him from his work. His active work in progressive organizations in the past was cause of his dismissal. In the FELLOWCOUNTRYMAN line LIBERAL is in touch with CHESTER [ix]. They meet once a month for the payment of dues. CHESTER is interested in whether we are satisfied with the collaboration and whether there are not any misunderstandings. He does not inquire about specific items of work [KONKRETNAYa RABOTA]. In as much as CHESTER knows about the role of LIBERAL's group we beg consent to ask C. through LIBERAL about leads from among people who are working on ENOURMOUS and in other technical fields.

"Enormous" == the atomic bomb



Codebook Cipher

- Literally, a book filled with "codewords"
- Zimmerman Telegram encrypted via codebook

Februar 13605

fest 13732

finanzielle 13850

folgender 13918

Frieden 17142

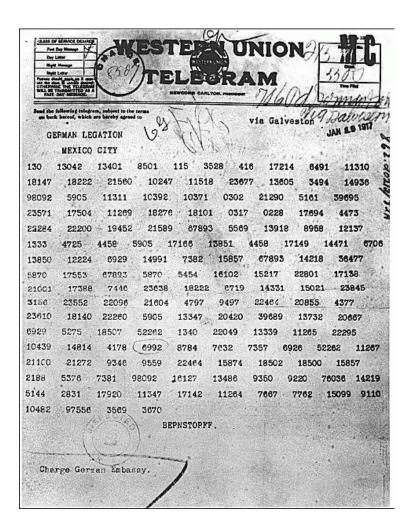
Friedenschluss 17149



Zimmerman Telegram

 Perhaps most famous codebook ciphertext ever

 A major factor in U.S. entry into World War I





History behind Zimmerman Telegram

- Zimmerman (then the German Foreign Minister) sent an encrypted message to German Ambassador in Mexico City.
- Zimmermann told his ambassador that Germany should try to recruit Mexico as an ally to fight against the United States. The incentive for Mexico was that it would "reconquer the lost territory in Texas, New Mexico and Arizona."
- At that time, the British and French at war with German but US was neutral



Zimmerman Telegram Decrypted

- British had recovered partial codebook, then was able to fill in missing parts
- After the decrypted
 Zimmerman telegram
 was released to U.S.,
 public opinion turned
 against Germany and,
 the U.S. declared war.



"We intend to begin on the first of February unrestricted submarine warfare. We shall endeavor in spite of this to keep the United States of america neutral. In the event of this not succeeding, we make Mexico a proposal of alliance on the following basis: make war together, make peace together, generous financial support and an understanding on our part that Mexico is to reconquer the lost territory in Texas, New Mexico, and arizona. The settlement in detail is left to you. You will inform the President of the above most . secretly as soon as the outbreak of war with the United States of America is certain and add the suggestion that he should, on his own initiative. Japan to immediate adherence and at the same time mediate between Japan and ourselves. Please call the President's attention to the fact that the ruthless employment of our submarines now offers the prospect of compelling England in a few months to make peace." Signed, ZIMERHARM.



Codebook Cipher - Additive

- Codebooks also (usually) use additive
- Additive book of "random" numbers
- Procedure of using a codebook and addtive
 - Encrypt message with codebook
 - Then choose position in additive book
 - Add additives to get ciphertext
 - Send ciphertext and additive position
 - Recipient subtracts additives before decrypting
- Both the codebook and additive are needed for encryption and decryption
- Why use an additive sequence?
 - Reduce frequency analysis, increase the life of codebook



Codebook Cipher - Illustration

Additive codebook

Original codebook

Word	Code
The	001
good	002
staff	003
dog	004
cat	005

Plaintext: good dog → 002 004

Random additive: 3 3 is sent in the clear at the start of transmission

Ciphertext: 094 226



- Basics of cryptography
 - Crypto terms
 - Crypto
 - Cryptography
 - Cryptanalysis
 - Kerckhoffs' principles



- Basics of cryptography
 - Crypto terms
 - Kerckhoffs' principles
 - Shift crypto



- Basics of cryptography
 - Crypto terms
 - Kerckhoffs' principles
 - Shift crypto
 - Vegenere



- Basics of cryptography
 - Crypto terms
 - Kerckhoffs' principles
 - Shift crypto
 - Vegenere
 - Double transposition



- Basics of cryptography
 - Crypto terms
 - Kerckhoffs' principles
 - Shift crypto
 - Vegenere
 - Double transposition
 - One-time pad



Refresher: One-Time Pad, Encryption

```
e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111
```

Encryption: Plaintext \oplus Key = Ciphertext



Refresher: One-Time Pad, Decryption

```
e=000 h=001 i=010 k=011 l=100 r=101 s=110 t=111
```

Decryption: Ciphertext ⊕ Key = Plaintext



- Basics of cryptography
 - Crypto terms
 - Kerckhoffs' principles
 - Shift crypto
 - Double transposition
 - One-time pad
 - Codebook cipher



Codebook Cipher

Original codebook

Word	Code
The	001
good	002
staff	003
dog	004
cat	005

Plaintext: good dog → 002 004

Random additive: 3

Ciphertext: 094 226

Additive codebook

