Cryptanalysis 101

How Hackers Think — and How We Stay One Step Ahead

CЯ4CK1NG-THΞ-C0D3

Today’s Roadmap

- Basics of Cryptanalysis
- Types of Cryptanalytic Attacks
- Hands-on Cryptanalysis, Group solving and CTF

Firstly, what is cryptanalysis?

Cryptanalysis

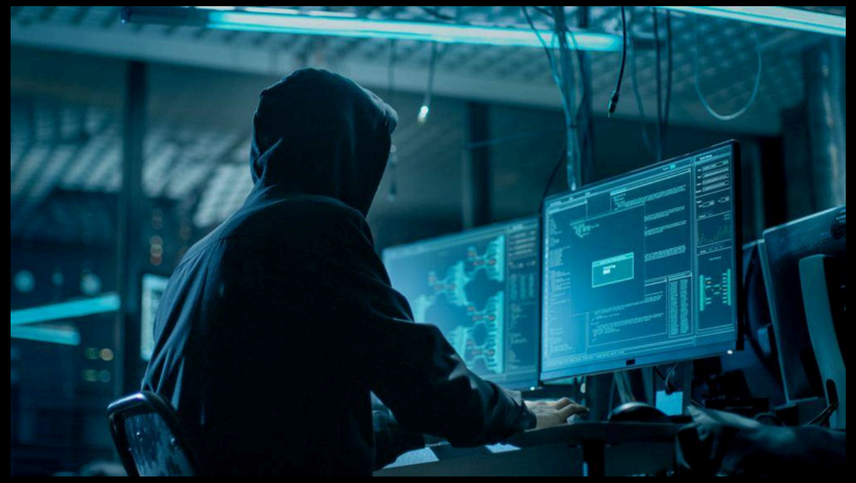
“It is the art of trying to decrypt the encrypted messages without the use of the key”

Lets formally define it:

- **Cryptanalysis** which is the study of the cryptographic algorithm and the breaking of those secret codes.
- The person practicing Cryptanalysis is called a **Cryptanalyst**
- To determine the weak points of a cryptographic system, it is important to attack the system. This attacks are called **Cryptanalytic attacks**.

Why learn cryptanalysis?

- To think like an attacker
- To test and strengthen cryptosystems
- To understand weaknesses before they’re exploited



“To defend, you must first understand how to attack.”

What is a secure (cryptographic) system?

“What cipher or encryption system is the safest to use?”

- A cryptosystem is **secure** if the best know attack is to try all possible keys
- Cryptosystem is **insecure** if **any** shortcut attack is known
- By this definition, an insecure system might be harder to break than a secure system!

How can hacker attack a cipher?

Method 1: Exhaustive key search

- An attacker can simply try all possible keys and test each to see if it is correct
- To prevent an exhaustive key search, a cryptosystem must have a large **keyspace***
- *Must be too many keys for hacker*
- *to try them all in any reasonable amount of time*

What is a key space?

Oops..You forgot your code for the number lock :(

Can we use this concept to break a cipher?

- Suppose that a cipher has a 100 bit key
- Then keyspace is of size 2^100
- On average, for exhaustive search hacker tests 2^100/2 = 2^99 keys
- Suppose hacker can test 2^30 keys/second
- Then she can find the key in about 37.4 trillion years

This in more known terms is called-
A Bruteforce attack

Brute-force attack

To formally define it:

- The attacker tries every possible key on a piece of cipher- text until an intelligible translation into plaintext is obtained. On average, half of all possible keys must be tried to achieve success.

Method 2: Letter Frequency analysis

- Every language has patterns — some letters appear more often than others.
- In English, for example, E, T, A, O, and N are the most common letters.
- By counting how often each letter appears in the ciphertext and comparing it to normal English frequencies, we can guess which ciphertext letters map to which plaintext letters.
- This method helps crack simple substitution ciphers without knowing the key.

Method 2: Letter Frequency analysis

Abu al-Kindī (c. 801–873 AD) is credited with developing a method whereby variations in the frequency of the occurrence of letters could be analyzed and exploited to break ciphers

- Characters do not occur with the same frequency in writings
  - For example, in English text E makes up 12.7% of the characters, while M makes up 2.4%
- You can use the frequency of letters in a ciphertext to make guesses about likely mappings from plaintext to ciphertext

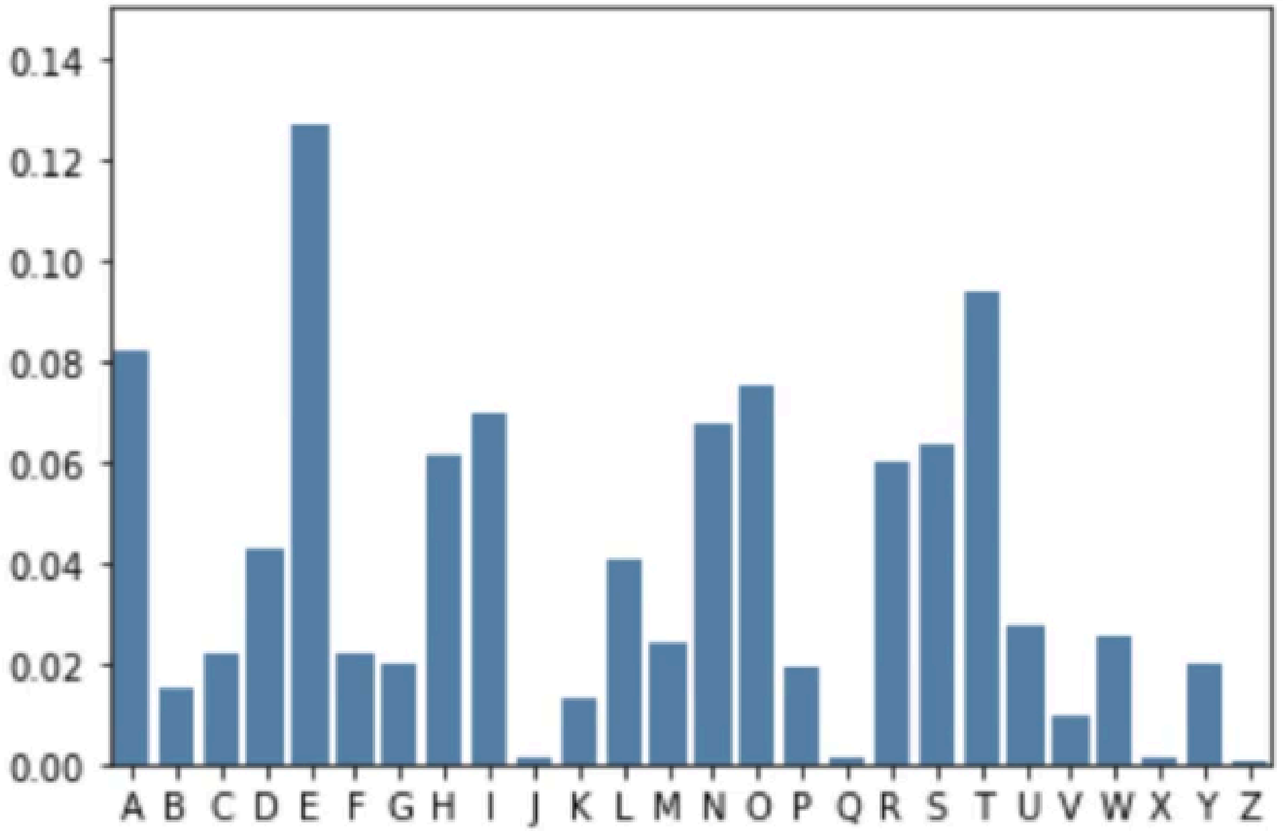
The first page of al-Kindi's manuscript "On Deciphering Cryptographic Messages", containing the oldest known description of cryptanalysis by frequency analysis.

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

English Language

Letter	Relative frequency in the English language
a	8.167%
b	1.492%
c	2.202%
d	4.253%
e	12.702%
f	2.228%
g	2.015%
h	6.094%
i	6.966%
j	0.153%
k	1.292%
l	4.025%
m	2.406%
n	6.749%
o	7.507%
p	1.929%
q	0.095%
r	5.987%
s	6.327%
t	9.356%
u	2.758%
v	0.978%
w	2.560%
x	0.150%
y	1.994%
z	0.077%

Letter frequency - Wikipedia



Lets try the methods on simple ciphers to understand better

Cryptanalysis of Caesar Cipher

- Think of how we could break the caesar cipher?

- Shift Cipher is not Secure.

- Brute-force cryptanalysis easily performed on the shift cipher by trying all 25 possible keys.

Ciphertext in Caesar cipher

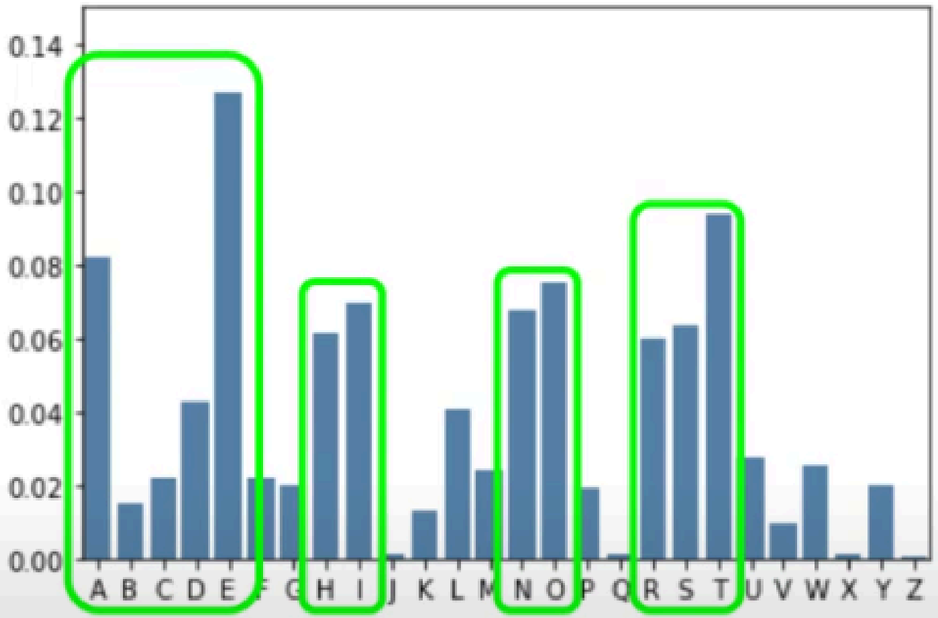
KEPPME IWX SQRMW HMZMWE MR TEVXIW XVIW

Only 25 different decryptions are possible.
The correct decryption is obvious, even if you don't know Latin.

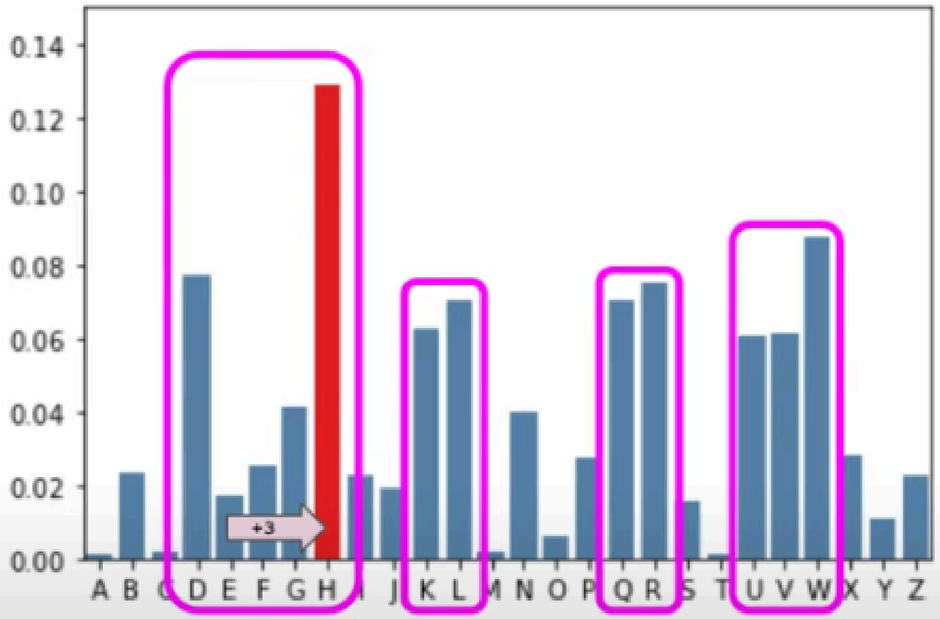
+1	LFQQNF JXY TRSNX INANXF NS UFWYJX YWJX
+2	MGRROG KYZ USTOY JOBOYG OT VGXZKY ZXKY
+3	NHSSPH LZA VTUPZ KPCPZH PU WHYALZ AYLZ
+4	OITTQI MAB WUVQA LQDQAI QV XIZBMA BZMA
+5	PJUURJ NBC XVWRB MRERBJ RW YJACNB CANB
+6	QKVVSK OCD YWXSC NSFSCK SX ZKBDOC DBOC
+7	RLWWTL PDE ZXYTD OTGTDL TY ALCEPD ECPD
+8	SMXXUM QEF AYZUE PUHUEM UZ BMDFQE FDQE
+9	TNYYVN RFG BZAVF QVIVFN VA CNEGRF GERF
+10	UOZZWO SGH CABWG RWJWGO WB DOFHSG HFSG
+11	VPAAXP THI DBCXH SXKXHP XC EPGITH IGTH
+12	WQBBYQ UIJ ECDYI TYLYIQ YD FQHJUI JHUI
+13	XRCCZR VJK FDEZJ UZMZJR ZE GRIKVJ KIVJ
+14	YSDDAS WKL GEFAK VANAKS AF HSJLWK LJWK
+15	ZTEEBT XLM HFGBL WBOBLT BG ITKMXL MKXL
+16	AUFFCU YMN IGHCM XCPCMU CH JULNYM NLYM
+17	BVGGDV ZNO JHIDN YDQDNV DI KVMOZN OMZN
+18	CWHHEW AOP KIJEO ZEREOW EJ LWNPAO PNAO
+19	DXIIFX BPQ LJKFP AFSFPX FK MXOQBP QOBP
+20	EYJJGY CQR MKLGQ BGTGQY GL NYPRCQ RPCQ
+21	FZKKHZ DRS NLMHR CHUHRZ HM OZQSDR SQDR
+22	**GALLIA EST OMNIS DIVISA IN PARTES TRES**
+23	HBMMJB FTU PNOJT EJWJTB JO QBSUFT USFT
+24	ICNNKC GUV QOPKU FKXKUC KP RCTVGU VTGU
+25	JDOOLD HVW RPQLV GLYLVD LQ SDUWHV WUHV

Letter analysis on Caesar Cipher

Caesar Cipher Patterns



English Language



Ciphertext
Caesar: k = 3

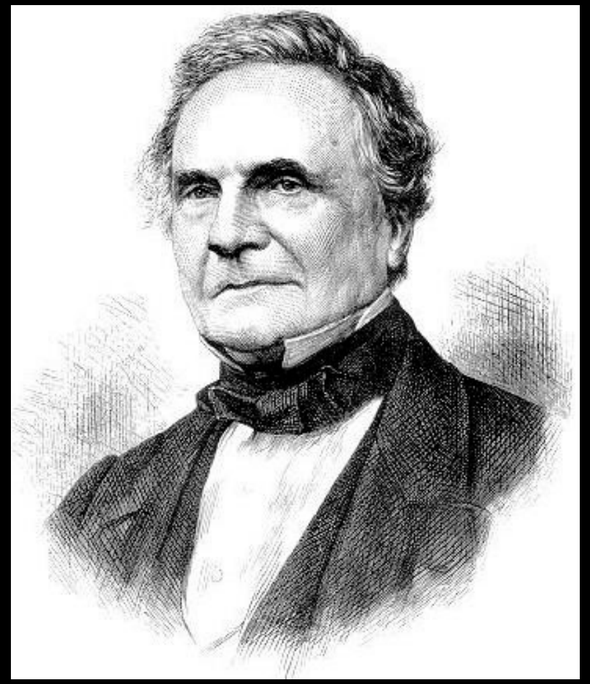
How about Vigenère Cipher?

Cryptanalysis of Vigenère Cipher

- Think of how we could break the Vigenère cipher?

- A shift cipher involves replacing each letter in the message by a letter that is some fixed number of positions further along in the alphabet.
  - Vigenere masks the frequency with which a character appears in a language: one letter in the ciphertext corresponds to multiple letters in the plaintext. Makes the use of frequency analysis more difficult.
  - Any message encrypted by a Vigenere cipher is a collection of as many shift ciphers as there are letters in the key.

Vigenère Cipher - Kasiski Examination



- Friedrich Kasiski (1805 – 1881) is credited with breaking the Vigenère cipher in 1863.
- It takes advantage of the fact that repeated words may, by chance, sometimes be encrypted using the same key letters, leading to repeated groups in the ciphertext.
- Example, Consider the following encryption using the keyword ABCD:
- Key: ABCDABCDABCDABCDABCDABCDABCD
- Plaintext: CRYPTOISSHORTFORCRYPTOGRAPHY
- Ciphertext: CSASTPKVSIQUTGQUCSASTPIUAQJB

Lets make sense of that with an example

Think of a common trigraph – say, the – and assume that it appears twice in the plaintext message.

If the is not encrypted by the same three alphabets at both locations, the two ciphertexts of the would be different.

```
GALOISGALOISGALOISGALOISGALOIS...GALOISGALOISGALOIS
the                                 the
ZHP                                 TSS
```

But, if we are lucky and the is encrypted by the same three alphabets, we would see a duplicate trigraph.

```
GALOISGALOISGALOISGALOISGALOIS...GALOISGALOISGALOIS
the                                 the
ZHP                                 ZHP
```

What is important to notice is that the distance between the beginnings of the ZHP trigraphs is a multiple of the length of the keyword. This provides information about the length of the keyword.

So does that make any sense?

Lets do better and try it with a question

Frequency Analysis on Monoalphabetic Substitution Cipher

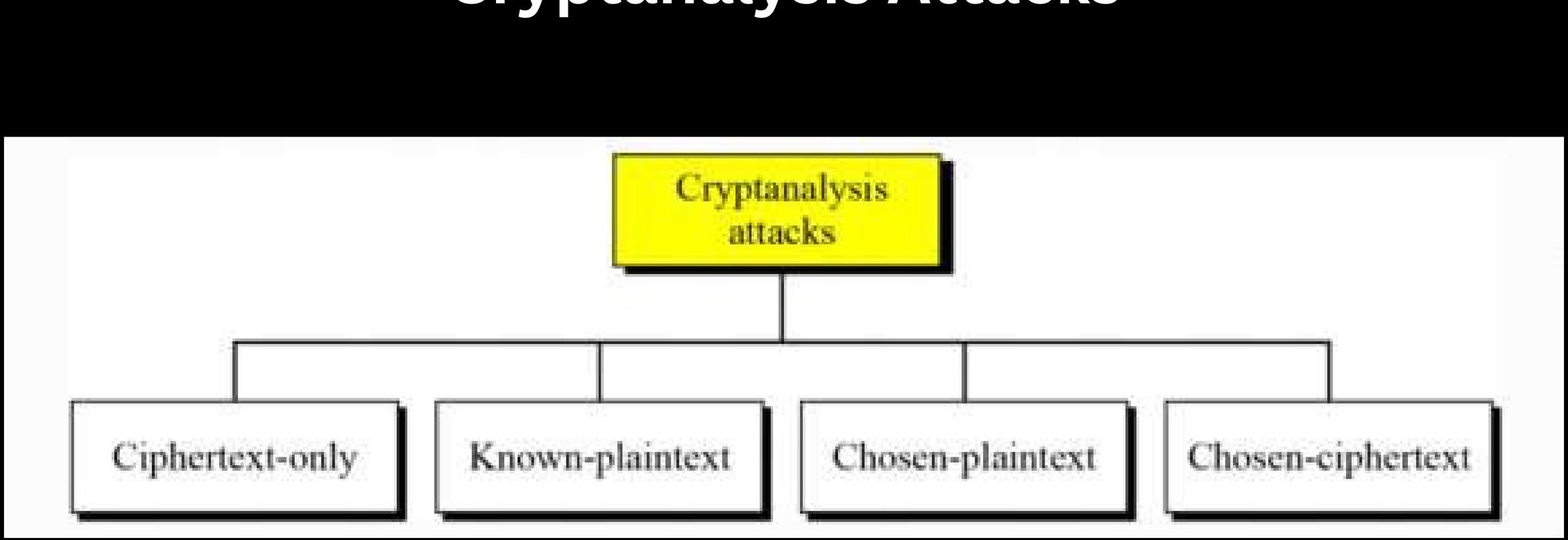
- You are given the following ciphertext, which was produced using a Monoalphabetic Substitution Cipher:

BFPXTRWTOW CEN BRTEWTA JNFVU E NJLNWFWJWFGV BFPXTR CXTRT TEBX QTWWTR MEPN WG EVGWXTR VGWFBT WXT PEWWTRV GZ WXT CGRAN WXT BGMMGV QTWWTRN GZ WXT TVUQFNX EQPXELTW XTQP FV ATBGAFVU E NMEQQ UJTNN EW ZFRNW WTNWFVU WXTM FV WXT WTOW EVA WXT RTNJQW LTBGMTN BQTERTR WXT PRGBTNN RTDTEQN WXT NTBRTW GZ WXT BGAT GVBT WXT MEPPFVU FN BGRRTBW WXT MTNNEUT XFAATV FN WXEW NJLNWFWJWFGV BFPXTRN ERT NFMPQT LJW ZRTIJTVBK EVEQKNFN MESTN WXTM CTES

Tasks

- 1. Perform letter frequency analysis on the ciphertext. Create a table of letter counts and percentages.
- 2. Compare the letter frequencies with standard English frequencies.
- 3. Make educated guesses for the most common ciphertext letters (e.g., likely e, t, a, o).
- 4. Look for common patterns:
  - Single-letter words (a, I)
  - Two-letter words (to, of, in, ...)
  - Three-letter words (the, and, for, ...)
- 5. Iteratively apply substitutions and reveal more plaintext.
- 6. Fully decrypt the message.

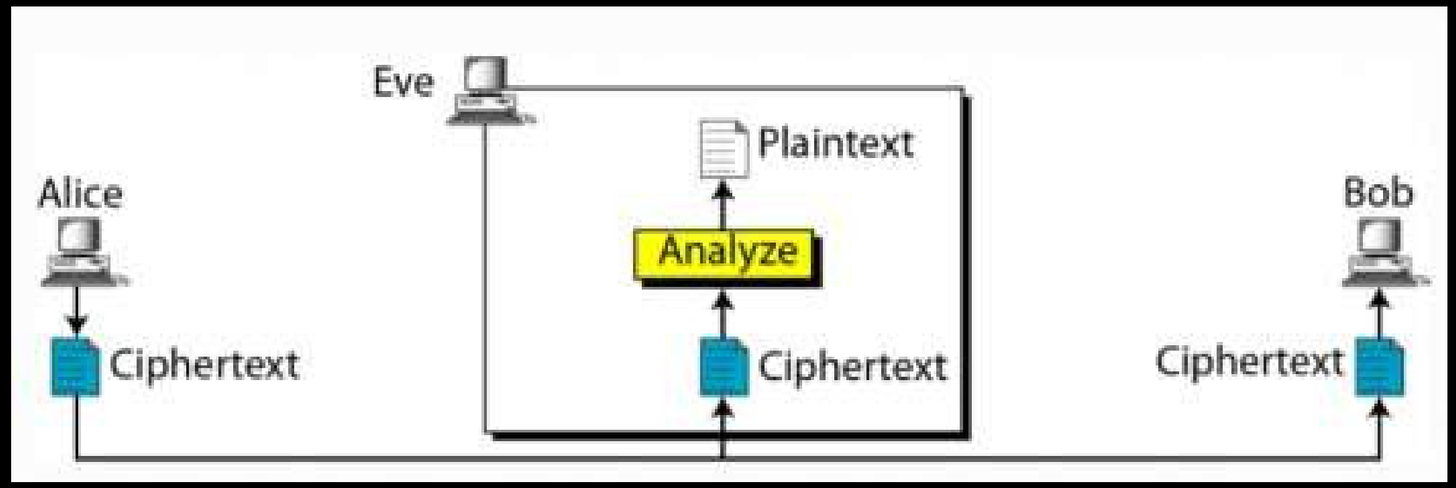
Cryptanalysis Attacks



Type of Attack	Known to Cryptanalyst
Cipher text only	• Encryption algorithm • Cipher text to be decoded
Known plain text	• Encryption algorithm • Cipher text to be decoded • One or more plain text-cipher text pairs formed with the secret key
Chosen plain text	• Encryption algorithm • Cipher text to be decoded • Plain text message chosen by cryptanalyst, together with its corresponding cipher text generated with the secret key
Chosen cipher text	• Encryption algorithm • Cipher text to be decoded • The purported cipher text chosen by cryptanalyst, together with its corresponding decrypted plain text generated with the secret key
Chosen text	• Encryption algorithm • Cipher text to be decoded • Plain text message chosen by cryptanalyst, together with its corresponding cipher text generated with the secret key • The purported cipher text chosen by cryptanalyst, together with its corresponding decrypted plain text generated with the secret key

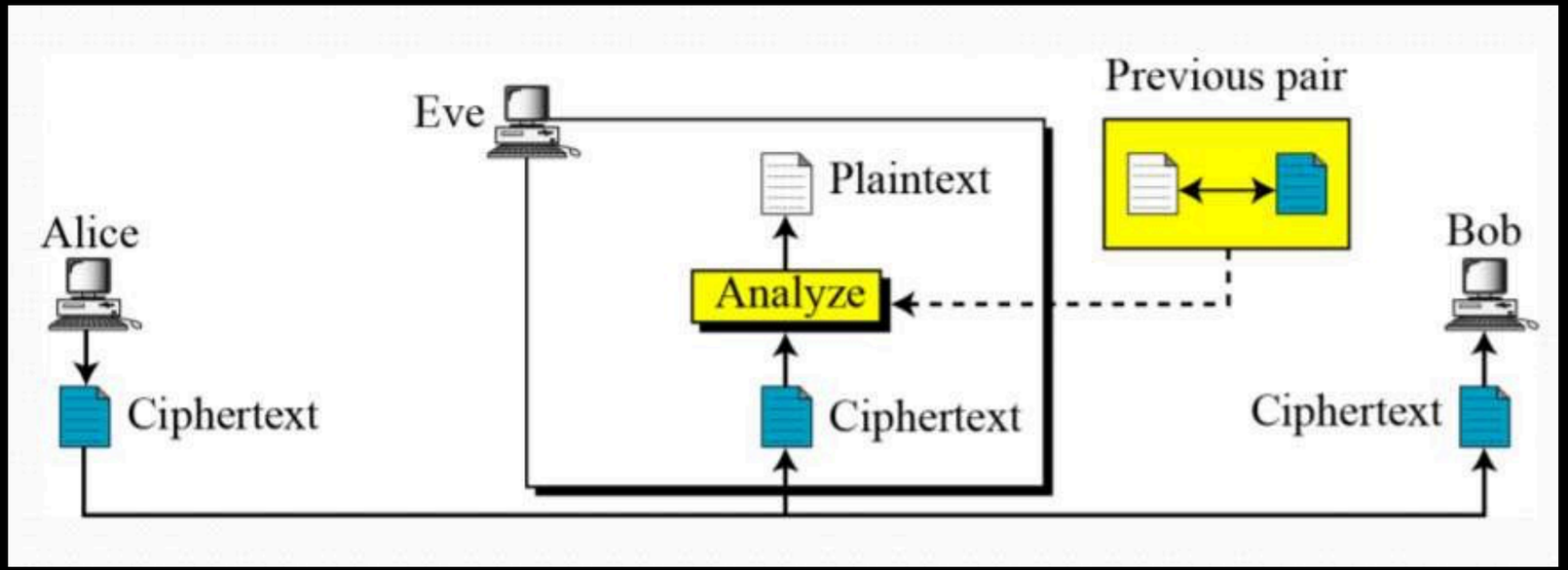
Ciphertext-Only Analysis (COA)

- The attacker only has the ciphertext (the encrypted message).
- They try to guess the plaintext or key using patterns, frequency analysis, or educated guesses.
- Example: Cracking a Caesar cipher by analyzing letter frequency.



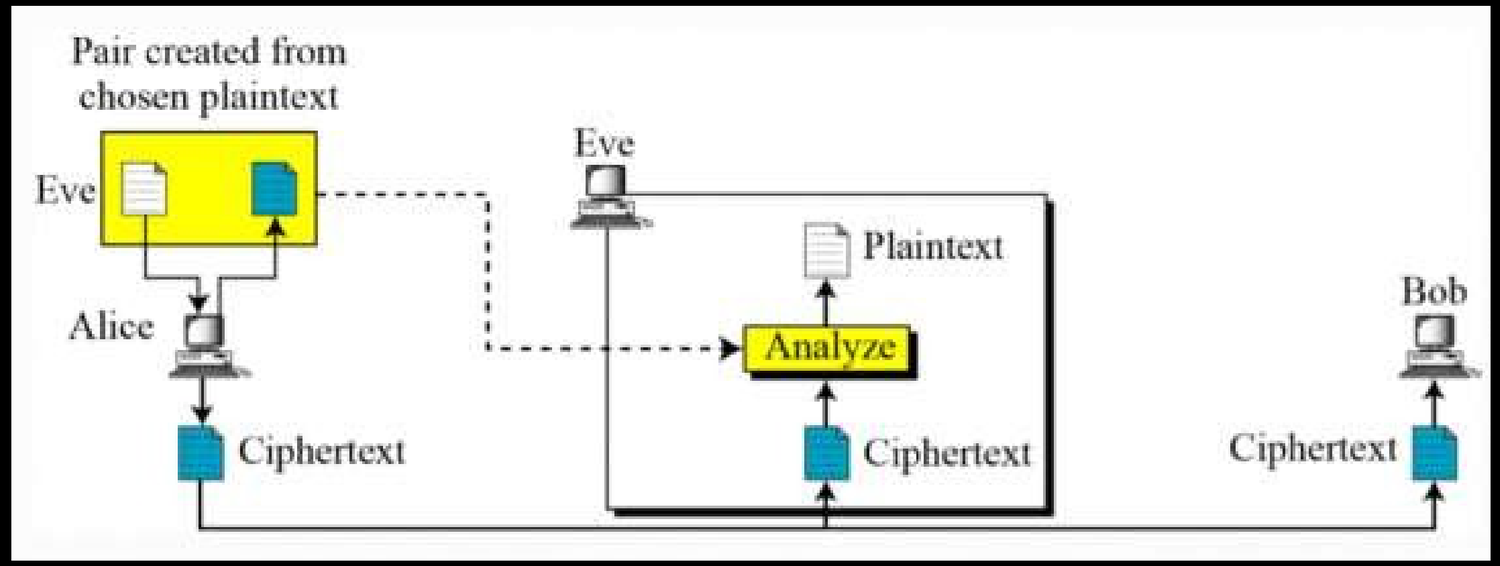
Known-Plaintext Analysis (KPA)

- The attacker has both the plaintext and its corresponding ciphertext for some messages.
- They use this information to figure out the key or the encryption method.
- Example: Knowing “HELLO” → “KHOOR” reveals a Caesar shift of 3.



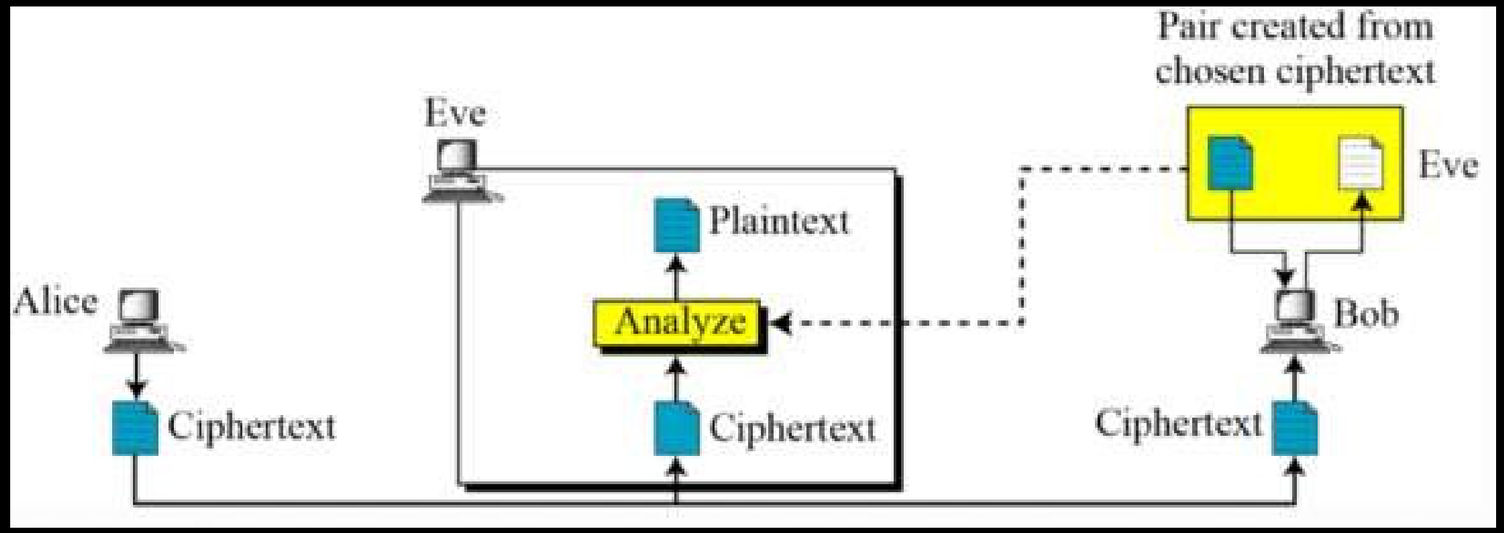
Chosen-Plaintext Analysis (CPA)

- The attacker can choose a plaintext and get its ciphertext encrypted by the system.
- This helps them learn how the system behaves and possibly find the key.
- Example: Sending “AAAA” to see what ciphertext comes out and spotting patterns.



Chosen-Ciphertext Analysis (CCA)

- The attacker can choose ciphertexts and get their plaintexts decrypted.
- This is very powerful for finding system weaknesses.
- Example: Sending modified ciphertexts to a server and analyzing responses to break RSA.



Chosen Text Analysis (CTA)

- A general case that includes both CPA and CCA.
- The attacker can choose both plaintexts and ciphertexts to study how encryption and decryption behave.
- Used to fully test and exploit the encryption scheme.

Hands on for implementing the same

Read carefully:

Among the endless symbols a secret lies,
digits of earth and sky disguise.
In the poster’s maze where @’s and &’s roam,
seek the numbers that point you home.
Two parts—north and west—form a pair;
enter them together and the flag waits there.

The poster you’ve been seeing all this while is not just art—it’s a cipher.