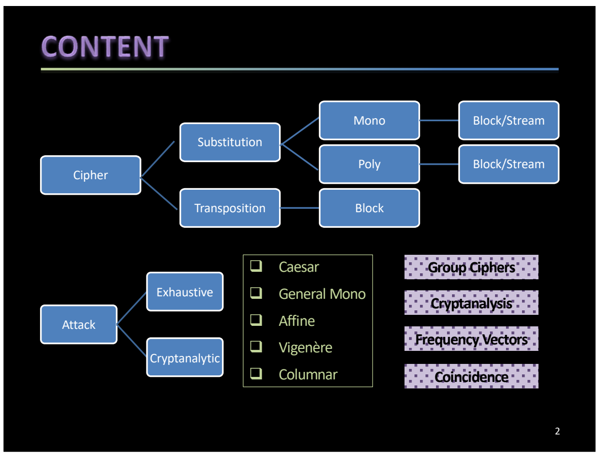
EECS 3481 APPLIED CRYPTOGRAPHY

CLASSICAL CRYPTO

PROF. H. ROUMANI
Dept. of Computer Science and Engineering, York University

1



2

CAESAR

Symmetric, Stream, Substitution, Mono-Alphabetic

The key is 3

Plaintext

THEKEYOFTHISCODESHIFTISTHREE

WKHNHBRIWKLVFRGHVKLIWLVWKUHH

Ciphertext

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
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25

3

3

Caesar Encryption

1. Read the plaintext file into an array of bytes pt
2. Clean pt keeping only letters and upper case them
3. Shift: ct[i] = [(pt[i] – 'A') + key] % 26 + 'A'
4. Write the ciphertext array ct to a file.

The key of this code shift is: three
THE KEY OF THIS CODE SHIFT IS THREE
THEKEYOFTHISCODESHIFTISTHREE
WKHNHBRIWKLVFRGHVKLIWLVWKUHH

4

4

Caesar Decryption

1. Read the ciphertext file into an array of bytes ct
2. Un-Shift: pt[i] = [(ct[i] – 'A') – key] **mod** 26 + 'A'
3. Write the ciphertext array pt to a file.

Note:
- *After subtracting 'A', all the array elements must be in [0,25]*
- *We should therefore work **modulus** 26*
- *Java's % gives the remainder, not the mod*
- *Hence, add an if statement to check for negative after % 26*

5

5

Questions About Caesar

- Why is Caesar symmetric?
- Is it a stream or a block cipher?
- Does it rely on substitution or transposition?
- Is it mono or poly alphabetic?
- Is it a group cipher?
- Describe its exhaustive and crypta attacks.
- Provide KPA and CPA examples.

6

6

Caesar Exhaustive Attack

- Try every possible key in the key space.
- How big is the key space?
- But how do you recognize success?
  - Dictionary Lookup via a Trie
  - Dot Product of Frequency Vectors
- Can you enlarge the key space?
  - Yes, can make it 26! ($\approx 10^{26} \approx 2^{88}$)

7

7

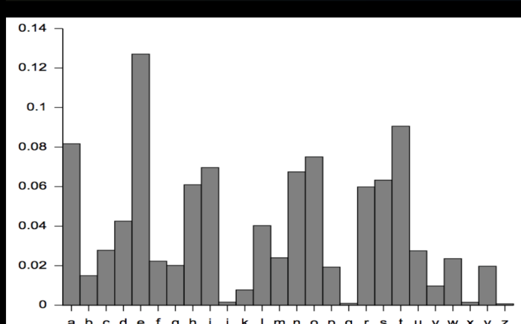
Monoalphabetic Cryptanalytic Attack

- Plaintext has certain patterns (regularities)
  - A Crib such as: *Date, From, GET, Dear …*
  - Language Statistics such as N-Gram Frequencies
- Do they die hard (survive the encryption)?
  - Compute the letter frequencies in ciphertext;
  - The largest is probably the shifted 'E' (or 'T');
  - Subtract to find the key.

8

8

Stats for English



Source https://en.wikipedia.org/wiki/Letter_frequency

9

9

N-Gram Stats for (No-Space) English

- Monogram
  E (13%), T (9%), A (8%)
  O, N, R, I, S, H — 6% ; D, L — 4%
  F, C, M, U, G, Y, P, W — 2% ; B, V, K — 1%
- Bigram
  TH, HE, IN, ER, AN, RE ...
- Same-letter Bigram
  LL, EE, SS, OO, TT, FF ...
- Trigram
  THE, AND, ING, ENT, ION, HER ...

10

10

Frequency Vectors

- Compute the frequencies of letters in the array
- Compare with the frequencies of English letters:
  - ❑ Think in frequency space (26 dimensions)
  - ❑ The computed frequencies form a vector
  - ❑ English frequencies form another vector
  - ❑ Are the two vectors "close"?
- For Mono, the two vectors have the same length
- Proximity measured by maximal dot product.

This technique is used for data mining to detect clusters; machine learning to detect similarity/patterns. Websites / Streaming services use it for recommendation

11

11

Obliterating Patterns

To defeat the cryptanalyst, we must prevent PT's patterns from appearing in CT; i.e. make CT as random as possible—maximize its entropy. How about these attempts:

- Compose two ciphers –Affine
- Different mappings for same PT letter –Vigenère
- Encrypt in blocks –Hill

12

12

The Affine Cipher

- A symmetric product* cipher
  $c \equiv \alpha\, p + \beta \pmod{26}$ where $\alpha \in [1,25]$ and $\beta \in [0,25]$
- Example
  Key = (α,β) = (3,5). P="CS" leads to C="LH"
- Decryption function
  $p \equiv (c - \beta) / \alpha \pmod{26}$
- Example
  For key (3,5), $1/\alpha \equiv 9$. Hence C="EM" leads to P=?

*Product = composition

A	0
B	1
C	2
D	3
E	4
F	5
G	6
H	7
I	8
J	9
K	10
L	11
M	12
N	13
O	14
P	15
Q	16
R	17
S	18
T	19
U	20
V	21
W	22
X	23
Y	24
Z	25

13

13

The Extended Euclid Algorithm

- Bézout [1730 AD]
  If a,b are co-prime integers, there exists integers x,y such that: ax + by = 1.
- Euclid [300 BC]
  His extended algorithm allows us to find x and y.
- Multiplicative Inverse
  Working with modulus a, y is nothing but 1/b
  Similarly, if we choose b as modulus then x = 1/a

14

14

Questions About Affine

- Is it symmetric or asymmetric?
- Is it a stream or block cipher?
- Is it substitution or transposition?
- Is it mono or ploy-alphabetic?
- Is Double-Affine better?

15

15

Affine Cryptanalytic Attacks

- Known Ciphertext … frequency based.
- Known Plaintext Attack … how many pairs?
- Chosen Plaintext Attack … how many pairs?
- Meet in the Middle Attack * … a KPA

**Sophisticated but an overkill here.*

16

16

Questions About Affine

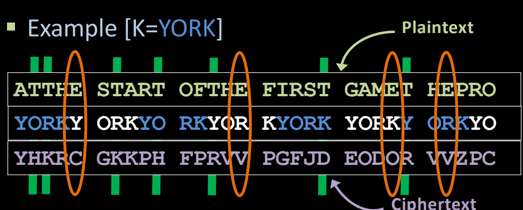
- Is it symmetric or asymmetric?
- Is it a stream or block cipher?
- Is it substitution or transposition?
- Is it mono or ploy-alphabetic?
- Is Double-Affine better?

17

17

The Vigenère Cipher

- A Polyalphabetic Cipher
  $c \equiv p + s_k \pmod{26}$ where k spans K circularly
- Example [K=YORK]



Exhaustive Attack: Key Space = $26^{|K|} \approx 2^{5|K|}$ → *hopeless!*

18

18

Vigenère Cryptanalysis: Key Length

- Crucial Observation
  Characters that are |K| apart are shifted equally!
  → Can easily answer: is the key of a given length?
- The Friedman Attack *(use this in this course)*
  Pick two letters from random locations and compute Index of Coincidence, IC = probability they are equal.

$$IC = \Sigma_i \, [f_i \times (f_i - 1)] / [n \times (n-1)]$$

→ *Can attack the key length exhaustively!*

19

19

Index of Coincidence

- From a random set of letters, select two randomly (i.e. from randomly chosen, not equal, locations).
- What is the probability that they are equal?
  - ❑ Direct Computation
  - ❑ Monte Carlo Sampling
- We call this the Index of Coincidence, IC
- What is IC for English?

20

20

Questions About Vigenère

- Is it symmetric or asymmetric?
- Is it a stream or block cipher?
- Is it substitution or transposition?
- Is it mono or ploy-alphabetic?
- Is Double-Vigenère better?

21

21

The Hill Cipher

- Algorithm
  `E: [C]≡[P]*[K] (mod 26)`, where K is an nxn matrix
  Must be able to invert the key matrix → GCD(det([K]),26)=1.
- Example with n=3
  K = {{1,2,3},{4,5,6},{11,9,8}}, K^{-1}={{22,5,1},{6,17,24},{15,13,1}}
  If P=ABC then C=AXW. Note that GCD=1 → |key space|< 26^9
- Key Characteristics
  – No more P-C positional correspondence (within n)
  – The K-C relationship is complex

22

22

Exercise on Hill's

Eve mounts a CPA with P="DONT", intercepts C="ELNI". Find the 2x2 Hill's key

[3 14] → [4 11], [13 19] → [13 8]

→{10 9}, {13, 23}

Repeat with P="DONT", C="ELNK".

→{10 19}, {13, 19}

One letter change in C changed a column in K.

23

23

Questions About Hill's

- Is it symmetric or asymmetric?
- Is it a stream or block cipher?
- Is it substitution or transposition?
- Is it mono or ploy-alphabetic?
- Is Double-Hill better?
- Attacks: KPA (crib dragging) or n-gram frequencies.

24

24

A Transposition Cipher

Plaintext in rows, ciphertext in columns
The key is: CFDBE (25314)

https://en.wikipedia.org/wiki/Scytale

THEKEYOFTHISCODESHIFTISTHREEZZ

Plaintext

THEKE
YOFTH
ISCOD
ESHIF
TISTH
REEZZ

Ciphertext

HOSSIEEHDFHZEFCHSETYIETRKTOITZ

25

25

Attacking Columnar Transposition

- Guess the key length
  Typically divisor of |C| or a dictionary word
- Exhaustive
  Parallel searches guided by anagrams
- Known / Chosen Plaintext
  Trivial to find the key

26

26

More on Transposition

- Describe a ciphertext-only attack.
- Is it symmetric or asymmetric?
- Is it a stream or block cipher?
- Is it substitution or transposition?
- Is it mono or ploy-alphabetic?
- Is it a group cipher?

27

27

Reflections

- Patterns die hard
- Confuse and Diffuse
- Provably, Computationally, and 'Hopefully' Secure
- Perfect Secrecy
  - Entropy and the One-Time Pad
  - Modern Stream Ciphers
- Imperfect Secrecy
  - Mix Substitution with Transposition (SPN)
  - Modern Block Ciphers

28

28