Classical Ciphers Analysis

COMPUTER SCIENCE AND ENGINEERING
INDIAN INSTITUTE OF INFORMATION TECHNOLOGY
SRI CITY, INDIA

Cryptosystem

A cryptosystem is a five tuple *(P, C, K, E, D)*, where the following conditions are satisfied:

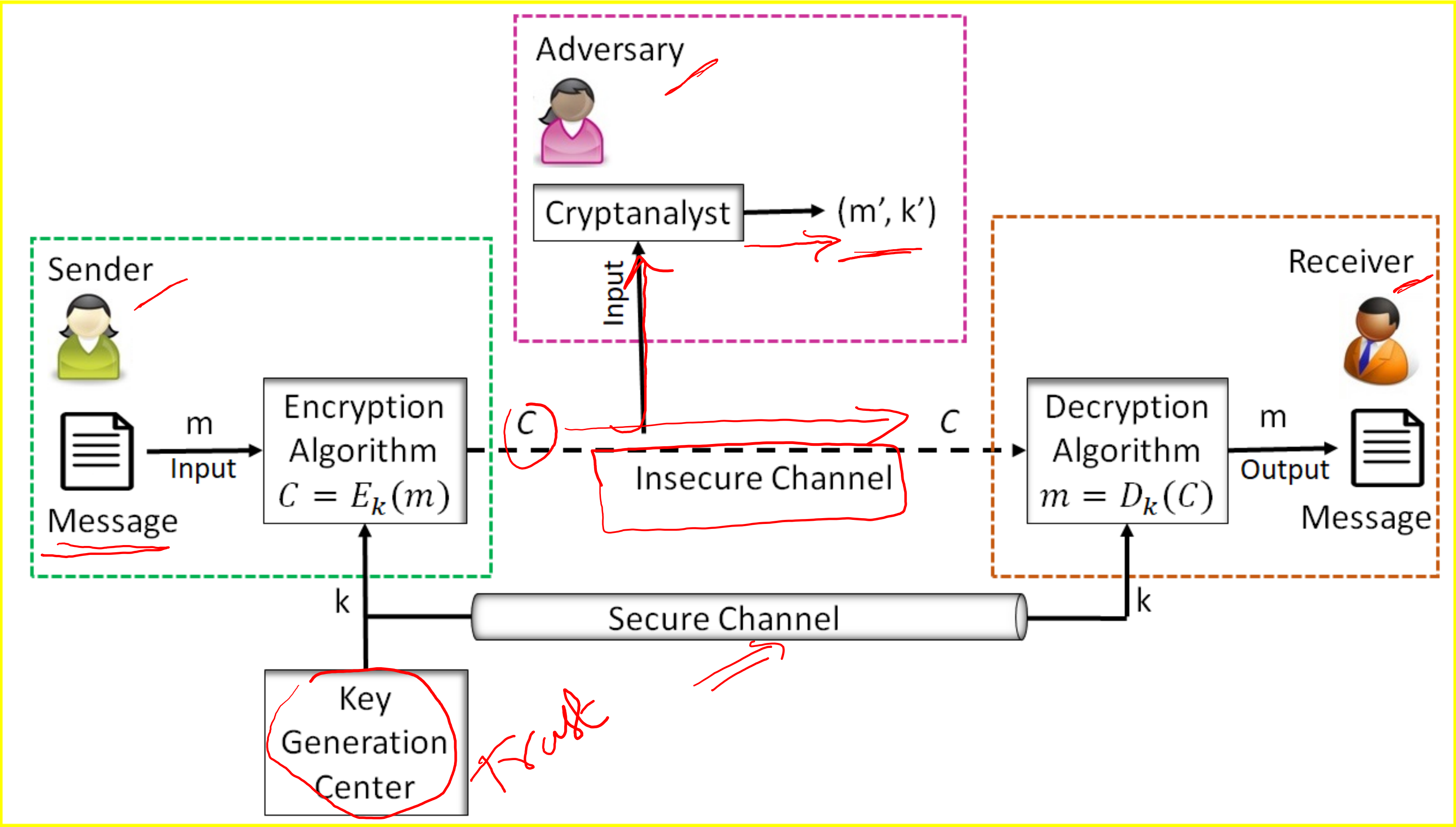
1. *P* is a finite set of possible plaintexts
2. *C* is a finite set of possible ciphertexts
3. *K*, the keyspace, is a finite set of possible keys
4. For each, $k \epsilon K$, there is an encryption rule $E_k \epsilon E$ and a corresponding decryption rule $D_k \epsilon D$.

   Each $E_k : P \rightarrow C$ and $\cancel{E}D_k : C \rightarrow P$ are functions such that

$$\boxed{D_k(E_k(m)) = m}$$

for every plaintext $m \epsilon P$.

Symmetric Cipher Model



Requirements

- Two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver

A

USA

U

Ind

public key

- Assume encryption algorithm is known
  - **Kerckhoff's Principle**: security in secrecy of key alone, not on the secrecy of the encryption algorithm

Symmetric

- Implies a secure channel to distribute key
  - Central problem in symmetric cryptography

Cryptography

Characterize cryptographic system by:

- Major types of encryption operations
  - Substitution
  - Transposition
- The way in which plaintext is processed
  - Block
  - Stream

Cryptanalysis

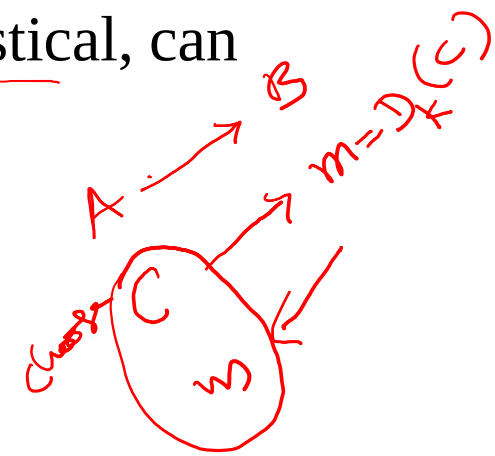
- Objective to recover key not just message
- General approaches:
  - Cryptanalytic attack
  - Brute-force attack

Cryptanalytic Attacks

hello

SKDCP

- **ciphertext only**
  - only know algorithm & ciphertext, is statistical, can identify plaintext
- **known plaintext**
  - know/suspect plaintext & ciphertext
- **chosen plaintext**
  - select plaintext and obtain ciphertext
- **chosen ciphertext**
  - select ciphertext and obtain plaintext
- **chosen text**
  - select plaintext or ciphertext to en/decrypt

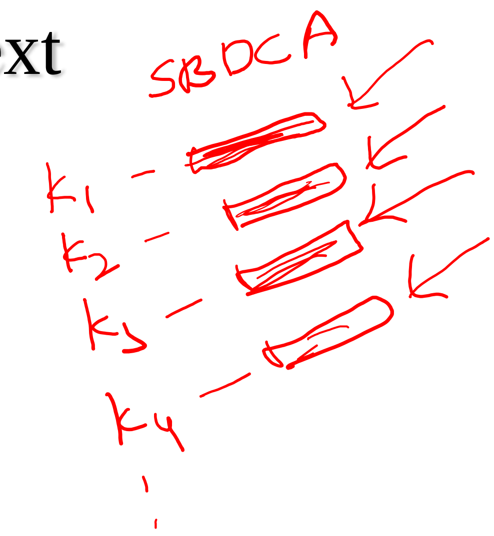


Cipher Strength

- **Unconditional security**
  - No matter how much computer power or time is available, the cipher cannot be broken since the ciphertext provides insufficient information to uniquely determine the corresponding plaintext
- **Computational security**
  - Given limited computing resources (e.g. time needed for calculations is greater than age of universe), the cipher cannot be broken

Brute Force Search

- Always possible to simply try every key
- Most basic attack, exponential in key length
- Assume either know / recognise plaintext



Classical Substitution Ciphers

- Letters of plaintext are replaced by other letters or by numbers or symbols

or

- If plaintext is viewed as a sequence of bits, then substitution involves replacing plaintext bit patterns with ciphertext bit patterns

Caesar Cipher

- Earliest known substitution cipher

  by Julius Caesar

- First attested use in military affairs
- Replaces each letter by 3rd letter on

$= (m+3) \bmod 26$

Example:

```
meet me after the toga party
PHHW PH DIWHU WKH WRJD SDUWB
```

Caesar Cipher

- Define transformation as:

```
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 = IN
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 = OUT
```

- Mathematically give each letter a number

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

- Caesar (rotation) cipher as:

$$c = E(k, p) = (p + k) \bmod (26)$$

$$p = D(k, c) = (c - k) \bmod (26)$$

$k \in \{0, 1, \cdots, 25\}$

$k \in \mathbb{Z}_{26}$

Cryptanalysis of Caesar Cipher

- Only have 26 possible ciphers
  - A maps to A,B,..Z
- So, simply try each in turn

  brute force search
- Given ciphertext, just try all shifts of letters

Break ciphertext "GCUA VQ DTGCM"

Affine Cipher

- Define affine transformation as:

  $$c = E(k, m) = (am + b) \bmod (26)$$

  $$m = D(k, c) = (a^{-1}(c - b)) \bmod (26)$$

- key $k=(a,b)$ and $(a, 26)=1$

$E_K(m)$

$D_K(c)$

$< 26 \times 26$

?

$\phi(26) = 12$ in Z_{26}

Affine Cipher - Example

- Example k=(17,3):

```
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 = IN
D U L C T K B S J A R I Z Q H Y P G X O F W N E V M = OUT
```

- Now how many keys are there?
  - 12 x 26 = 312

- Still can be brute force attacked!

Monoalphabetic Cipher

- Rather than just shifting the alphabet
- We could shuffle (permute) the letters arbitrarily
- Each plaintext letter maps to a different random ciphertext letter
- Hence, **key is 26 letters long**

a
P
{permute}

```
Plain:  abcdefghijklmnopqrstuvwxyz
Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN
```

```
Plaintext:  ifwewishtoreplaceletters
Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA
```

Monoalphabetic Cipher Security

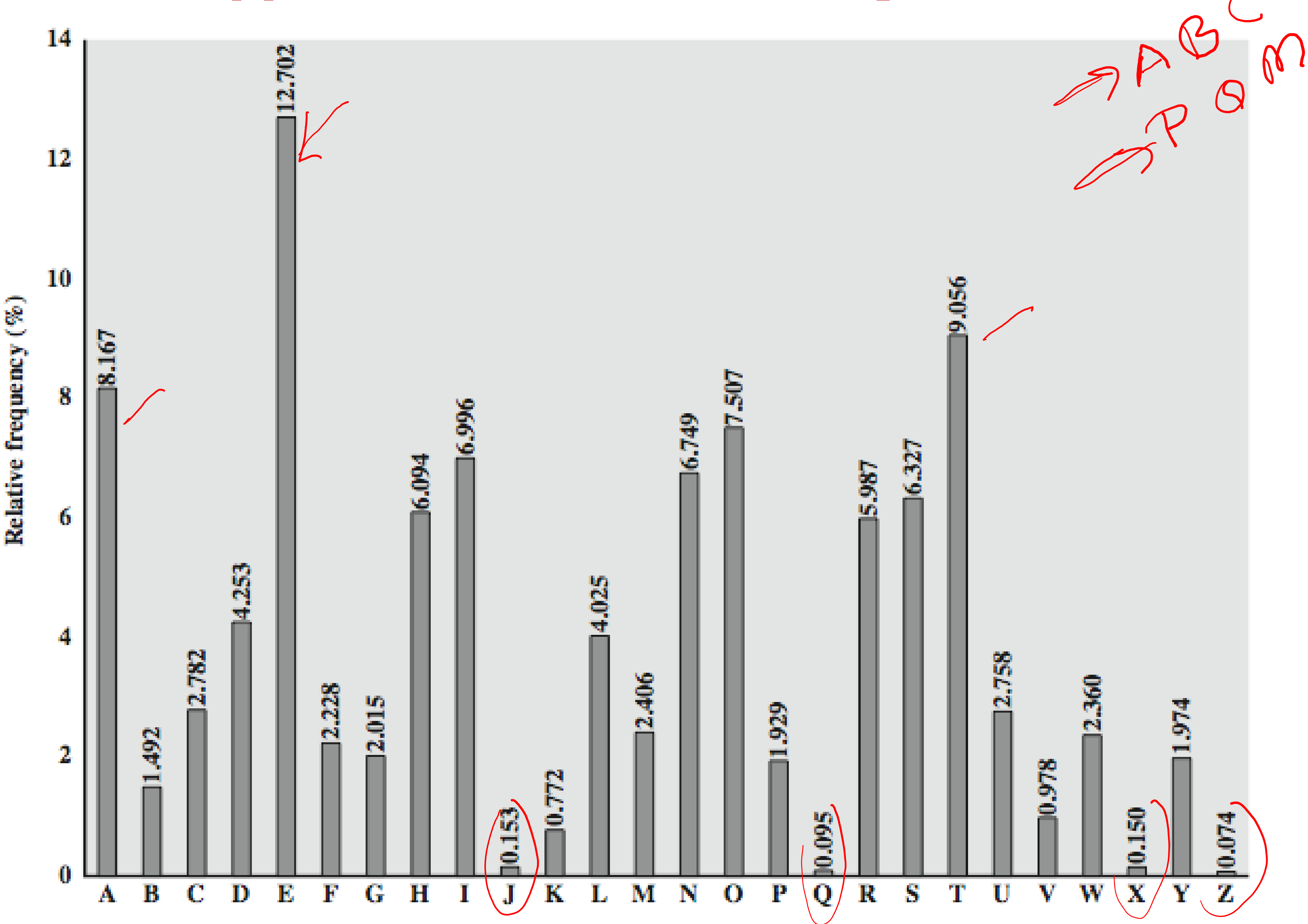
- key size is now 25 characters.
- Now, a total of $26! = 4 \times 10^{26}$ keys
- So many keys, might think is secure
- But, would be **!!!WRONG!!!**
- Problem is language characteristics

Brute force not possible

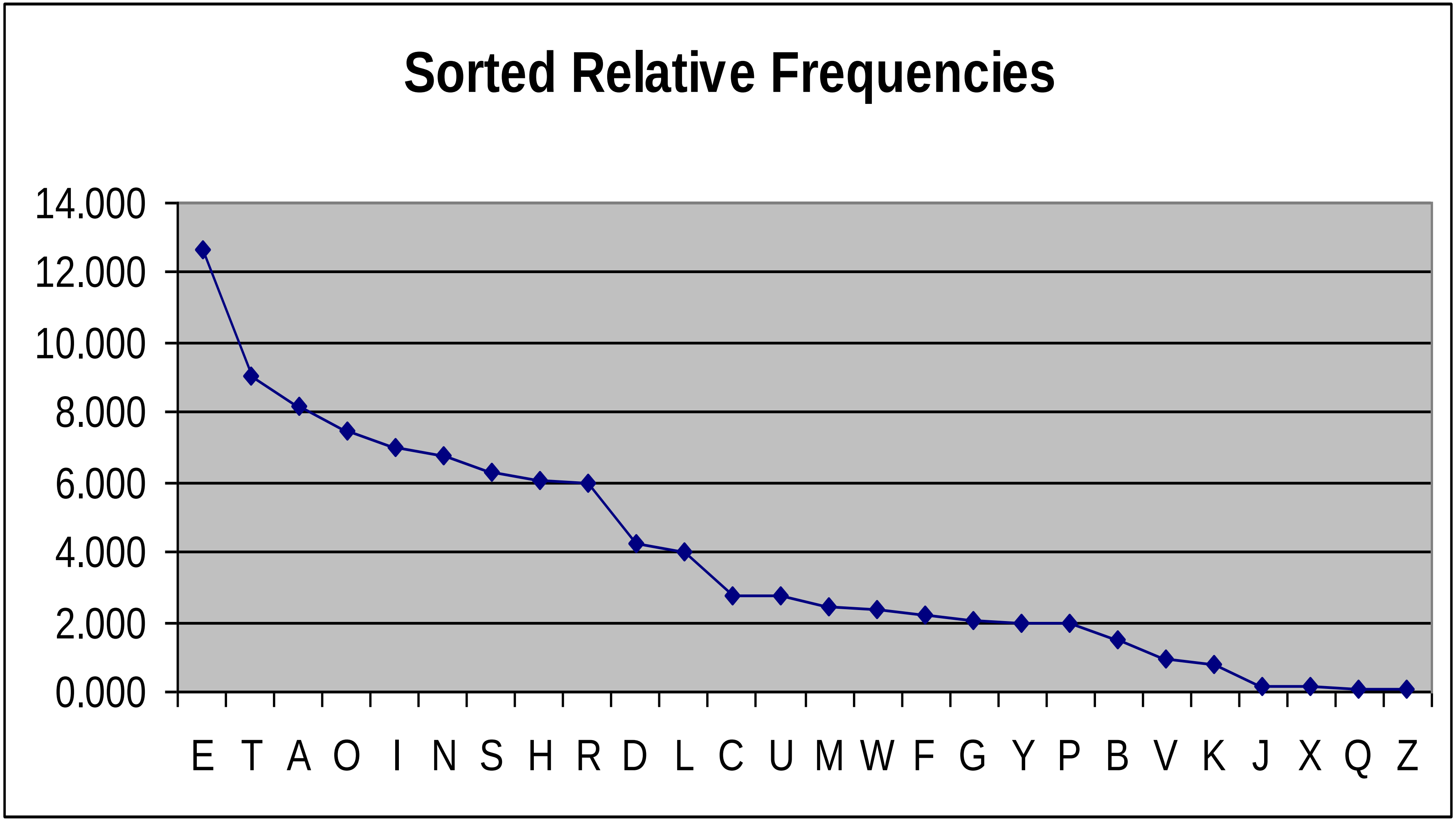
Language Redundancy and Cryptanalysis

- letters are not equally commonly used
- in English E is by far the most common letter
  - followed by T,R,N,I,O,A,S
- other letters like Z,J,K,Q,X are fairly rare

English Letter Frequencies



English Letter Frequencies



Example Cryptanalysis

- Given ciphertext:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- Count relative letter frequencies

Example Cryptanalysis

- Given ciphertext:

```
UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ
VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX
EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ
```

- guess P & Z are e and t
- guess ZW is th and hence ZWP is “the”
- proceeding with trial and error finally get:

```
it was disclosed yesterday that several informal but
direct contacts have been made with political
representatives of the viet cong in moscow
```

Polyalphabetic Ciphers

$a \to y$

$a \to t$

$a \to p$

- **polyalphabetic substitution ciphers**
- Improve security using multiple cipher alphabets
- Make cryptanalysis harder with more alphabets to guess and flatter frequency distribution
- Use a key to select which alphabet is used for each letter of the message
- Use each alphabet in turn, and repeat from start after end of key is reached

Vigenère Cipher

- Simplest polyalphabetic substitution cipher
- Effectively multiple caesar ciphers
- Key is multiple letters long $K = k_1 k_2 \ldots k_d$
- i^{th} letter specifies i^{th} alphabet to use
- Use each alphabet in turn, and repeat from start after d letters in message
- Decryption simply works in reverse

Example of Vigenère Cipher

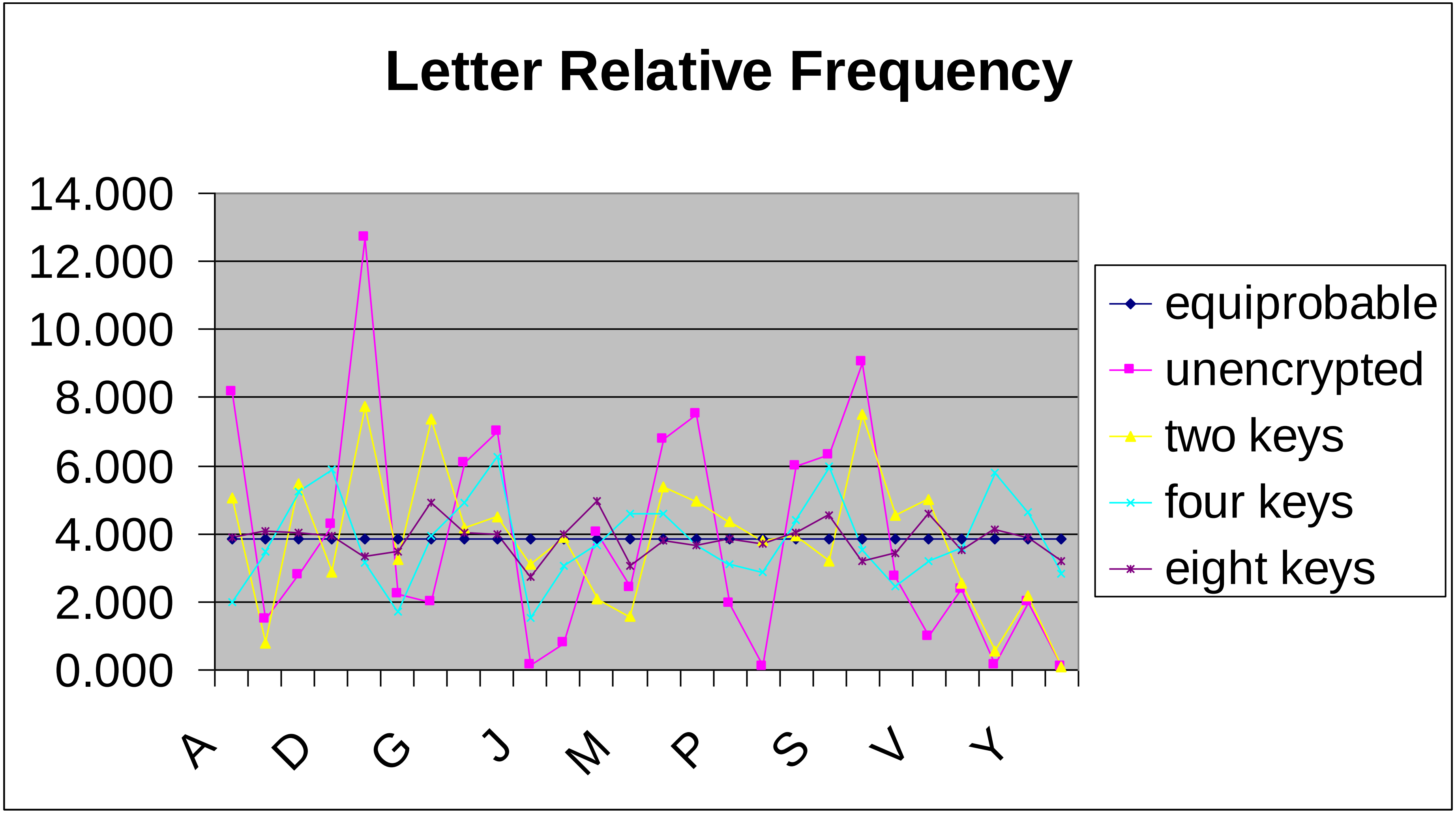
plaintext (m)	a	b	c	d	e	f	g	h	i	j	k	l	m	n
Assigned No.	0	1	2	3	4	5	6	7	8	9	10	11	12	13
plaintext (m)	o	p	q	r	s	t	u	v	w	x	y	z		
Assigned No.	14	15	16	17	18	19	20	21	22	23	24	25		

- Example: keyword *deceptive*

```
key:        deceptivedeceptivedeceptive
plaintext:  wearediscoveredsaveyourself
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

$3 + 22 = 25 \bmod 26 = 25$

Frequencies After Polyalphabetic Encryption



Frequencies After Polyalphabetic Encryption