CRYPTOGRAPHY AND NETWORK SECURITY
UNIT 2

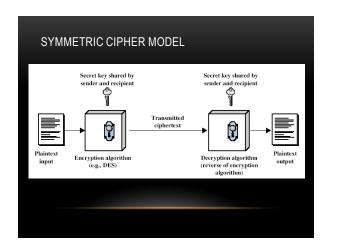
# UNIT 2 CLASSICAL ENCRYPTION TECHNIQUES

# SYMMETRIC ENCRYPTION

- or conventional / private-key / single-key
- sender and recipient share a common key
- all classical encryption algorithms are private-key
- was only type prior to invention of public-key in 1970's
- and by far most widely used

# SOME BASIC TERMINOLOGY

- plaintext original message
- ciphertext coded message
- cipher algorithm for transforming plaintext to ciphertext
- key info used in cipher known only to sender/receiver
- encipher (encrypt) converting plaintext to ciphertext
- decipher (decrypt) recovering ciphertext from plaintext
- cryptography study of encryption principles/methods
- cryptanalysis (codebreaking) study of principles/ methods of deciphering ciphertext without knowing key
- cryptology field of both cryptography and cryptanalysis



## SYMMETRIC CIPHER MODEL...

- plaintext original message
- encryption algorithm performs substitutions/transformations on plaintext
- secret key control exact substitutions/transformations used in encryption algorithm
- ciphertext scrambled message
- decryption algorithm inverse of encryption algorithm

# **REQUIREMENTS**

- two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver
  - We assume that it is impractical to decrypt a message on the basis of the cipher- text plus knowledge of the cryption/decryption algorithm, and do not need to keep the algorithm secret; rather we only need to keep the key secret.
  - This feature of symmetric encryption is what makes it feasible for widespread use.
  - It allows easy distribution of s/w and h/w implementations.

# **REQUIREMENTS...**

mathematically it can be considered a pair of functions with: plaintext X, ciphertext Y, key K, encryption algorithm  $\mathsf{E}_\mathsf{K}$ , decryption algorithm D<sub>K</sub>.

$$Y = E_K(X)$$

$$X = D_K(Y)$$

- assume encryption algorithm is known
- implies a secure channel to distribute key

# **CRYPTOGRAPHY**

characterize cryptographic system by:

- Type of encryption operations used
  - substitution / transposition / product
- · Number of keys used
  - single-key or private / two-key or public
- · Way in which plaintext is processed
  - block / stream

# **CRYPTANALYSIS**

- · Objective is to recover key not just message
- General approaches:
  - Cryptanalytic attack
    - Attacks rely on the nature of the algorithm plus perhaps some knowledge of the general characteristics of the plaintext or even some sample plaintext-ciphertext pairs.
  - Brute-force attack
    - try every possible key on a piece of ciphertext until an intelligible translation into plaintext is obtained.
    - On average, half of all possible keys must be tried to achieve success

# **CRYPTANALYTIC ATTACKS**

- · Ciphertext only
  - · Only know algorithm & ciphertext, is statistical, know or 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

# **MORE DEFINITIONS**

- 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 (eg 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, proportional to key size
- Assume either know / recognise plaintext

Key Size (bits)	Number of Alternative Keys	Time required at 1 decryption/µs	Time required at 10 <sup>6</sup> decryptions/μs
32	$2^{32} = 4.3 \times 10^9$	$2^{31} \mu s = 35.8$ minutes	2.15 milliseconds
56	$2^{56} = 7.2 \times 10^{16}$	$2^{55} \mu s = 1142 \text{ years}$	10.01 hours
128	$2^{128} = 3.4 \times 10^{38}$	$2^{127} \mu s = 5.4 \times 10^{24}$ years	5.4 × 1018 years
168	$2^{168} = 3.7 \times 10^{50}$	$2^{167} \mu s = 5.9 \times 10^{36}$ years	5.9 × 10 <sup>30</sup> years
26 characters (permutation)	$26! = 4 \times 10^{26}$	$2 \times 10^{26} \mu s$ = 6.4 × $10^{12}  years$	$6.4 \times 10^6$ years

# **CLASSICAL SUBSTITUTION CIPHERS**

- Where 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
- Example:

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

# **CAESAR CIPHER**

• can define transformation as:

abcdef ghij kl mnopq r s t u v w x y z DEFGH IJ K LMNOP QRS T U V W X Y Z A B C

• 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

• then have Caesar cipher as:

 $c = E(p) = (p + k) \mod (26)$ 

 $p = D(c) = (c - k) \mod (26)$ 

### CRYPTANALYSIS OF CAESAR CIPHER

- only have 26 possible ciphers
  - A maps to A,B,..Z
- · could simply try each in turn
- a brute force search
- given ciphertext, just try all shifts of letters
- do need to recognize when have plaintext
- eg. break ciphertext "GCUA VQ DTGCM"

# MONOALPHABETIC CIPHER

- Rather than just shifting the alphabet
- could shuffle (jumble) the letters arbitrarily
- each plaintext letter maps to a different random ciphertext letter
- hence key is 26 letters long

Plain: abcdefghijklmnopqrstuvwxyz Cipher: DKVQFIBJWPESCXHTMYAUOLRGZN

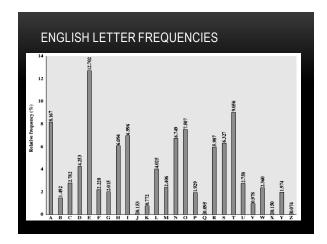
Plaintext: **ifwewishtoreplaceletters** Ciphertext: WIRFRWAJUHYFTSDVFSFUUFYA

# MONOALPHABETIC CIPHER SECURITY

- now have a total of 26! = 4 x 1026 keys
- · with so many keys, might think is secure
- but would be !!! WRONG!!!
- problem is language characteristics

# LANGUAGE REDUNDANCY AND CRYPTANALYSIS

- human languages are redundant
- eg "th Ird s m shphrd shll nt wnt"
- · 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
- have tables of single, double & triple letter frequencies for various languages



# **USE IN CRYPTANALYSIS**

- key concept monoalphabetic substitution ciphers do not change relative letter frequencies
- discovered by Arabian scientists in 9th century
- calculate letter frequencies for ciphertext
- · compare counts/plots against known values
- · if caesar cipher look for common peaks/troughs
  - peaks at: A-E-I triple, NO pair, RST triple
  - troughs at: JK, X-Z
- · for monoalphabetic must identify each letter
  - · tables of common double/triple letters help

# TRY TO FIND THE PLAINTEXT

• Given ciphertext:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESX UDBMETSXAIZVUEPHZHMDZSHZOWSFPAPP DTSVPQUZWYMXUZUHSXEPYEPOPDZSZUFP OMBZWPFUPZHMDJUDTMOHMQ

# **EXAMPLE CRYPTANALYSIS**

• Given ciphertext:

UZQSOVUOHXMOPVGPOZPEVSGZWSZOPFPESXUDBMETSXAIZ VUEPHZHMDZSHZOWSFPAPPDTSVPQUZWYMXUZUHSX

EPYEPOPDZSZUFPOMBZWPFUPZHMDJUDTMOHMQ

- · Count relative letter frequencies (see text)
- 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

# PLAYFAIR CIPHER

- Not even the large number of keys in a monoalphabetic cipher provides security
- · One approach to improving security was to encrypt multiple letters
- the Playfair Cipher is an example
- Invented by Charles Wheatstone in 1854, but named after his friend Baron Playfair

# PLAYFAIR KEY MATRIX

- The best-known multiple-letter encryption cipher
- a 5X5 matrix of letters based on a keyword
- fill in letters of keyword (sans duplicates)
- · fill rest of matrix with other letters
- · eg. using the keyword MONARCHY

M	0	N	Α	R
С	Н	Υ	В	D
Е	F	G	I/J	K
L	Р	Q	S	T
U	V	W	Х	Z

## ENCRYPTING AND DECRYPTING PLAYFAIR CIPHER

- Plaintext is encrypted two letters at a time
  - if a pair is a repeated letter, insert filler like 'X'
     eg. "balloon" encrypts as "ba lx lo on"
  - if both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
  - 3. if both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom)
  - otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

# SECURITY OF PLAYFAIR CIPHER

- security much improved over monoalphabetic
- since have 26 x 26 = 676 digrams
- would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic)
- · and correspondingly more ciphertext
- · was widely used for many years
  - eg. by US & British military in WW1
- it can be broken, given a few hundred letters
- · since still has much of plaintext structure

# POLYALPHABETIC CIPHERS

- 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
- · 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<sub>1</sub> k<sub>2</sub> ... k<sub>d</sub>
- ith letter specifies ith alphabet to use
- · use each alphabet in turn
- · repeat from start after d letters in message
- · decryption simply works in reverse

# EXAMPLE OF VIGENÈRE CIPHER

- · write the plaintext out
- write the keyword repeated above it
- use each key letter as a caesar cipher key
- encrypt the corresponding plaintext letter
- eg using keyword deceptive

key: deceptivedeceptive plaintext: wearediscoveredsaveyourself ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

# AIDS FOR EN/DECRYPTION

- simple aids can assist with en/decryption
- a Saint-Cyr Slide is a simple manual aid
  - a slide with repeated alphabet
  - line up plaintext 'A' with key letter, eg 'C'
  - · then read off any mapping for key letter
- · can bend round into a cipher disk
- or expand into a Vigenère Tableau

# SECURITY OF VIGENÈRE CIPHERS

- have multiple ciphertext letters for each plaintext letter
- · hence letter frequencies are obscured
- but not totally lost
- · start with letter frequencies
  - see if look monoalphabetic or not
- if not, then need to determine number of alphabets, since then can attach each

# KASISKI METHOD

- Method Developed By Babbage / Kasiski
- · Repetitions In Ciphertext Give Clues To Period
- So Find Same Plaintext An Exact Period Apart
- · Which Results In The Same Ciphertext
- · Of Course, Could Also Be Random Fluke

# **AUTOKEY CIPHER**

- Ideally Want A Key As Long As The Message
- Vigenère Proposed The Autokey Cipher
- · With Keyword Prefixed To Message As Key
- Knowing Keyword Can Recover The First Few Letters
- Use These In Turn On The Rest Of The Message
- But Still Have Frequency Characteristics To Attack
- Eg. Given Key Deceptive

Key: Deceptivewearediscoveredsav Plaintext: Wearediscoveredsaveyourself

Ciphertext:zicvtwqngkzeiigasxstslvvwla

# ONE-TIME PAD

- if a truly random key as long as the message is used, the cipher will be secure
- · called a One-Time pad
- is unbreakable since ciphertext bears no statistical relationship to the plaintext
- since for any plaintext & any ciphertext there exists a key mapping one to other
- · can only use the key once though
- problems in generation & safe distribution of key

# ONE TIME PAD

# OTP...

- The one-time pad offers complete security but, in practice, has two fundamental difficulties:
  - There is the practical problem of making large quantities of random keys.
  - And the problem of key distribution and protection, where for every message to be sent, a key of equal length is needed by both sender and receiver.
- Because of these difficulties, the one-time pad is of limited utility, and is useful primarily for low-bandwidth channels requiring very high security.

# TRANSPOSITION CIPHERS

- now consider classical transposition or permutation ciphers
- these hide the message by rearranging the letter order
- without altering the actual letters used
- can recognise these since have the same frequency distribution as the original text

# RAIL FENCE CIPHER

## **RAIL FENCE CIPHER**

- · write message letters out diagonally over a number of rows
- · then read off cipher row by row
- eg. write message out as:

```
mematrhtgpry etefeteoaat
```

· giving ciphertext

MEMATRHTGPRYETEFETEOAAT

# **ROW TRANSPOSITION CIPHERS**

- a more complex transposition
- · write letters of message out in rows over a specified number of columns
- then reorder the columns according to some key before reading off the rows

Key: 3421567
Plaintext: a t t a c k p
ost pone
duntilt
woamxyz

Ciphertext: TTNAAPTMTSUOAODWCOIXKNLYPETZ

# **PRODUCT CIPHERS**

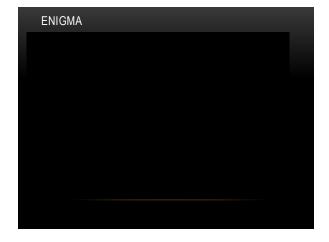
- ciphers using substitutions or transpositions are not secure because of language characteristics
- hence consider using several ciphers in succession to make harder, but:
  - two substitutions make a more complex substitution
  - two transpositions make more complex transposition
  - but a substitution followed by a transposition makes a new much harder cipher
- · this is bridge from classical to modern ciphers

# **ROTOR MACHINES**

- before modern ciphers, rotor machines were most common complex ciphers in use
- widely used in WW2
  - German Enigma, Allied Hagelin, Japanese Purple
- implemented a very complex, varying substitution cipher
- used a series of cylinders, each giving one substitution, which rotated and changed after each letter was encrypted
- with 3 cylinders have 263=17576 alphabets

# ROTOR MACHINE...

- A rotor machine consists of a set of independently rotating cylinders through which electrical pulses can flow.
- Each cylinder has 26 input pins and 26 output pins, with internal wiring that connects each input pin to a unique output pin. If we associate each input and output pin with a letter of the alphabet, then a single cylinder defines a monoalphabetic substitution.
- After each input key is depressed, the cylinder rotates one position, so that the internal connections are shifted accordingly.
- The power of the rotor machine is in the use of multiple cylinders, in which the output pins of one cylinder are connected to the input pins of the next, and with the cylinders rotating like an "odometer", leading to a very large number of substitution alphabets being used, eg with 3 cylinders have 26³=17576 alphabets used.
- They were extensively used in world war 2, and the history of their use and analysis is one of the great stories from WW2.



# HAGELIN ROTOR MACHINE

# STEGANOGRAPHY • an alternative to encryption • hides existence of message • using only a subset of letters/words in a longer message marked in some way • using invisible ink • hiding in LSB in graphic image or sound file • has drawbacks • high overhead to hide relatively few info bits

# HILL CIPHERS Created by L

- Created by Lester S. Hill in 1929
- Polygraphic Substitution Cipher
- Uses Linear Algebra to Encrypt and Decrypt

# POLYGRAPHIC SUBSTITUTION CIPHERS

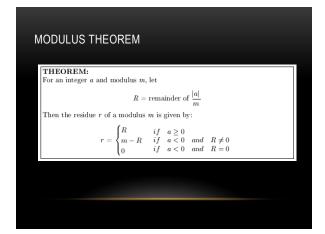
- Encrypts letters in groups
- Frequency analysis more difficult

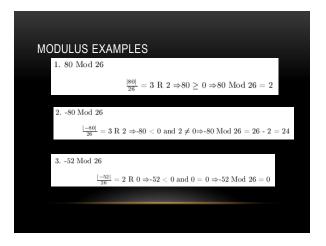
# HILL CIPHERS

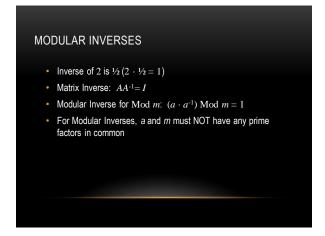
- Polygraphic substitution cipher
- Uses matrices to encrypt and decrypt
- Uses modular arithmetic (Mod 26)

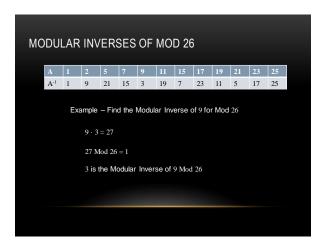
# MODULAR ARITHMETIC

- For a Mod b, divide a by b and take the remainder.
- $14 \div 10 = 1 \text{ R } 4$
- 14 Mod 10 = 4
- 24 Mod 10 = 4









## HILL CIPHER MATRICES

- One matrix to encrypt, one to decrypt
- Must be n x n, invertible matrices
- Decryption matrix must be modular inverse of encryption matrix in Mod 26

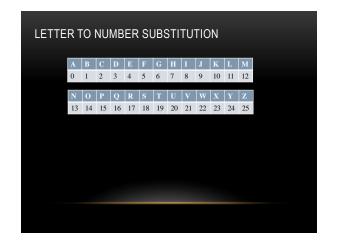
# MODULARLY INVERSE MATRICES

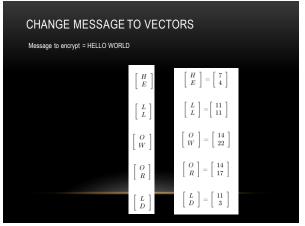
- · Calculate determinant of first matrix A, det A
- Make sure that det A has a modular inverse for Mod 26
- · Calculate the adjugate of A, adj A
- · Multiply adj A by modular inverse of det A
- · Calculate Mod 26 of the result to get B
- Use A to encrypt, B to decrypt

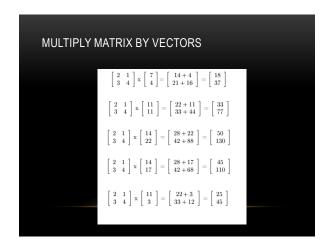
# 

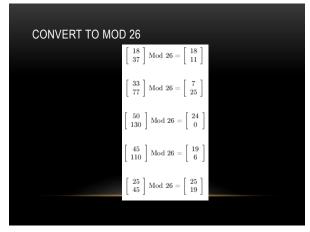
# **ENCRYPTION**

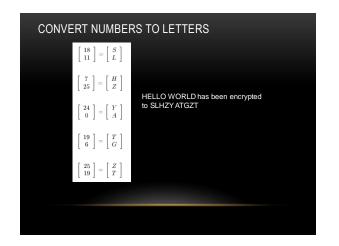
- Assign each letter in alphabet a number between 0 and 25
- Change message into 2 x 1 letter vectors
- Change each vector into 2 x 1 numeric vectors
- Multiply each numeric vector by encryption matrix
- Convert product vectors to letters



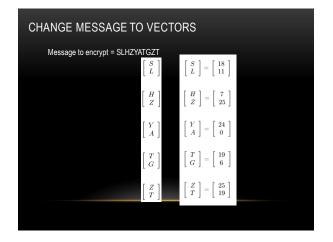


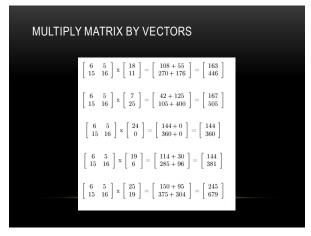


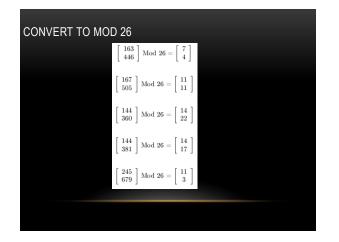


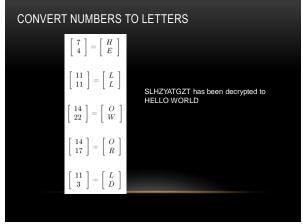












# CONCLUSION Creating valid encryption/decryption matrices is the most difficult part of Hill Ciphers. Otherwise, Hill Ciphers use simple linear algebra and modular arithmetic

