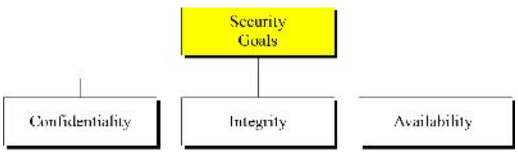
Cryptography
Introduction and classical ciphers

Anand Ballabh Joshi
Department of Mathematics
University of Lucknow, Lucknow, India

Applications

Security Goals

Taxonomy of security goals



1.3

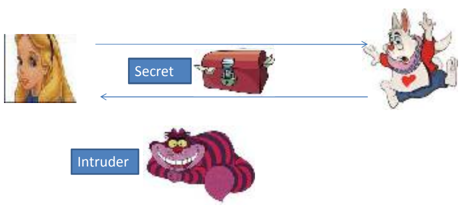
Security goals: Confidentiality

Confidentiality is probably the most common aspect of information security. We need to protect our confidential information. An organization needs to guard against those malicious actions that endanger the confidentiality of its information.

Confidentiality is related to the protection of information from disclosure to the unauthorized persons.

1.4

Confidentiality



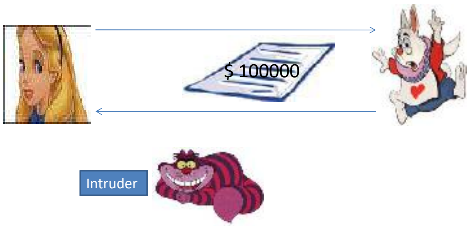
Security goals: Integrity

Information needs to be changed constantly. Integrity means that changes need to be done only by authorized entities and through authorized mechanisms.

Integrity is about protection of information from being modified by unauthorized persons.

1.6

Integrity of message



Security goals: Availability

The information created and stored by an organization needs to be available to authorized entities. Information needs to be constantly changed, which means it must be accessible to authorized entities.

The avialability of the information refers to ensuring that authorized parties are able to access the information when needed.

➢Imagine what would happen to a bank if the customer could not access their accounts for transactions.

1.8

Security goals

Confidentiality is related to the protection of information from disclosure to the unauthorized persons.

Integrity is about protection of information from being modified by unauthorized persons.

Non repudiation is the assurance that someone can not deny something. It is different form the other because it is related to the communicating parties.

The avialability of the information refers to ensuring that authorized parties are able to access the information when needed.

ATTACKS

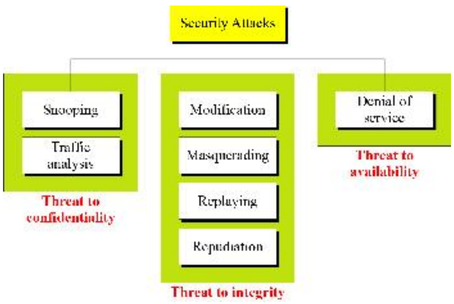
The three goals of security—confidentiality, integrity, and availability—can be threatened by security attacks.

1. Attacks Threatening Confidentiality
2. Attacks Threatening Integrity
3. Attacks Threatening Availability

1.10

Security attacks

Taxonomy of attacks with relation to security goals



1.11

Attacks Threatening Confidentiality

Snooping refers to unauthorized access to or interception of data.

Traffic analysis refers to obtaining some other type of information by monitoring online traffic.

To prevent snooping the data can be made unintelligible to the intercepter by using encipherment techniques.

1.12

Attacks Threatening Integrity

Modification means that the attacker intercepts the message and changes it.

Masquerading or spoofing happens when the attacker impersonates somebody else.

Replaying means the attacker obtains a copy of a message sent by a user and later tries to replay it.

Repudiation means that sender of the message might later deny that she has sent the message; the receiver of the message might later deny that he has received the message.

1.13

Attacks Threatening Availability

Denial of service (DoS) is a very common attack. It may slow down or totally interrupt the service of a system.

Attacker can use several strategies to achieve this:

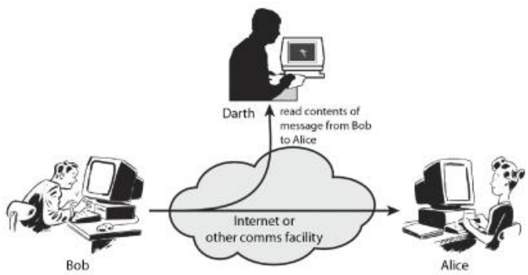
- Send so many bogus requests to a server that the server crashes because of the heavy load.
- Intercept and delete a server's response to a client, making the client to believe that the server is not responding.
- Intercept request from the clients, causing the clients to send requests many times and overload the system.

1.14

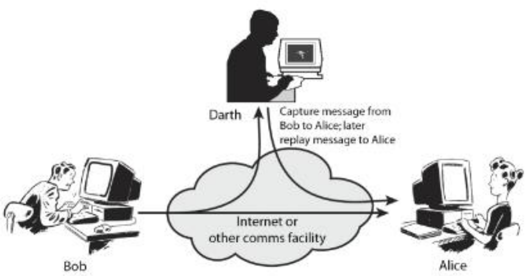
Security Attack

- any action that compromises the security of information owned by an organization
- information security is about how to prevent attacks, or failing that, to detect attacks on information-based systems
- often *threat* & *attack* used to mean same thing
- have a wide range of attacks
- can focus of generic types of attacks
  - passive
  - active

Passive Attacks



Active Attacks



Passive versus Active attacks

- In a passive attack, the attacker's goal is just to obtain information. This means that the attack does not modify data or harm the system. The system continue with its normal operation. However, the attack may harm the sender or receiver or the message. Attacks that threaten the confidentiality—snooping and traffic analysis-are passive attaks.
- An active attack may change the data or harm the system. Attacks that threaten the integrity and availability are active attacks.

Passive Versus Active Attacks

Table 1.1 Categorization of passive and active attacks

Attacks	Passive/Active	Threatening
Snooping Traffic analysis	Passive	Confidentiality
Modification Masquerading Replaying Repudiation	Active	Integrity
Denial of service	Active	Availability

SERVICES AND MECHANISMS

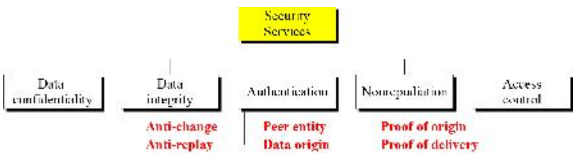
ITU-T provides some security services and some mechanisms to implement those services. Security services and mechanisms are closely related because a mechanism or combination of mechanisms are used to provide a service.

Topics discussed in this section:

1. Security Services
2. Security Mechanism
3. Relation between Services and Mechanisms

Security Services

Security services



Security Services

Data confidentiality is design to protect data from disclosure attack.

Data integrity is designed to protect data from modification , insertion, deletion, and replaying

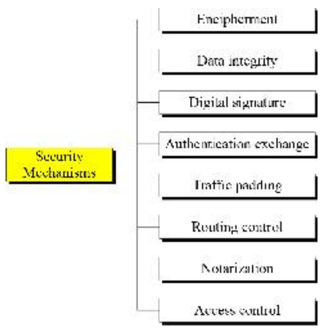
Authentication provide the authentication of the party at the other end of the line.

Nonrepudiation service protect against repudiation by either the sender or receiver of the data

Access control provide protection against unauthorized access to data.

Security Mechanism

Security mechanisms



Security mechanisms

- Encipherment, hiding or covering data, can provide confidentiality. Two techniques-cryptography and steganography.
- The data integrity mechanism appends to the data a short check value that has been created by a specifics process from the data itself. The receiver receives the data and the checkvalue. He creates a new chekvalue with the one received. If the two check values are same, the integrity of the data has been preserved.
- A digital signature is a means by which the sender can electronically sign the data and receiver can electronically verify the signature.
- In authentication exchange two entities exchange some message to prove their identity to each other.
- Traffic padding means inserting some bogus data into the data traffic to thwart the adversary's attempt to use the traffic analysis.
- Notarization means selecting a third trusted party to control the communication between two entities.

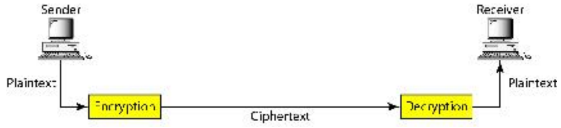
Relation between Services and Mechanisms

Table : Relation between security services and mechanisms

Security Service	*Security Mechanism*
Data confidentiality	Encipherment and routing control
Data integrity	Encipherment, digital signature, data integrity
Authentication	Encipherment, digital signature, authentication exchanges
Nonrepudiation	Digital signature, data integrity, and notarization
Access control	Access control mechanism

Mechanism for confidentiality

Mechanism for confidentiality



Authentication

Authenticity is the process of determining whether someone or something is, in fact, who or what it is declared to be.

Mechanism for authentication

1. Something that you know
   E.g. a PIN or a password
2. Something that you have
   E.g. a smart-card
3. Something that you are
   Biometric characteristics like voice, fingerprints, eyes, ...
4. Where you are located
   E.g. in a secure building

Strong authentication combines multiple factors:
E.g., Smart-Card + PIN

TECHNIQUES

Mechanisms discussed in the previous sections are only theoretical recipes to implement security. The actual implementation of security goals needs some techniques. Two techniques are prevalent today: cryptography and steganography.

1. Cryptography
2. Steganography

Cryptography

Cryptography, a word with Greek origins, means “secret writing.” However, we use the term to refer to the science and art of transforming messages to make them secure and immune to attacks.

Steganography

The word steganography, with origin in Greek, means “covered writing,” in contrast with cryptography, which means “secret writing.”

Example: covering data with text

This book is mostly about cryptography, not steganography.
□ □□□ □ □ □ □□
0 1 0 0 0 0 1

1.31

Continued

Example: using dictionary

A	**friend**	**called**	**a**	**doctor.**
0	10010	0001	0	01001

Example: covering data under color image

0101001**1**	1011110**0**	0101010**1**
0101111**0**	1011110**0**	0110010**1**
0111111**0**	0100101**0**	0001010**1**

1.32

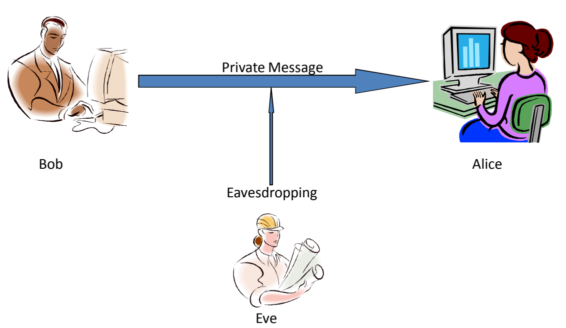
What is cryptography?

- kryptos – “hidden”
- grafo – “write”
- Keeping messages secret
  - Usually by making the message unintelligible to anyone that intercepts it

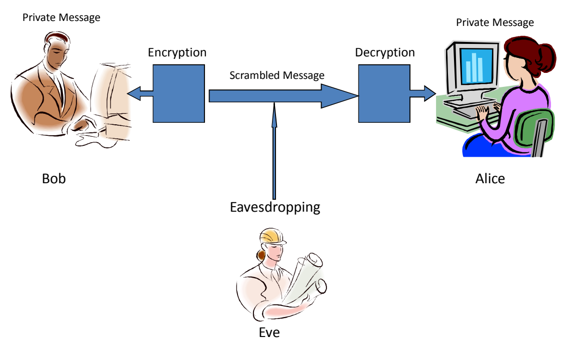
Some basic definitions

- 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 plaintext from ciphertext
- 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

The Problem



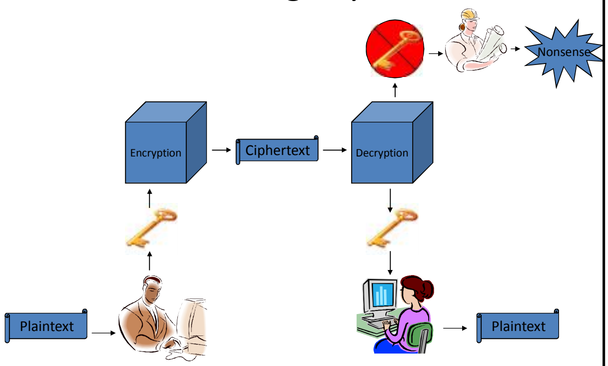
The Solution



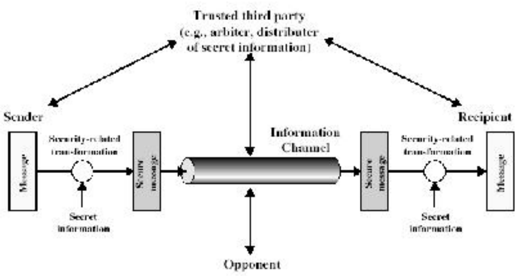
What do we need?

- Bob and Alice want to be able to encrypt/decrypt easily
- But no one else should be able to decrypt
- How do we do this?
  - Keys!

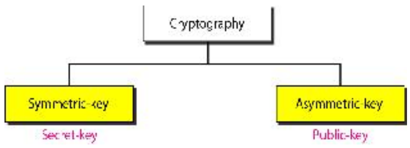
Using Keys



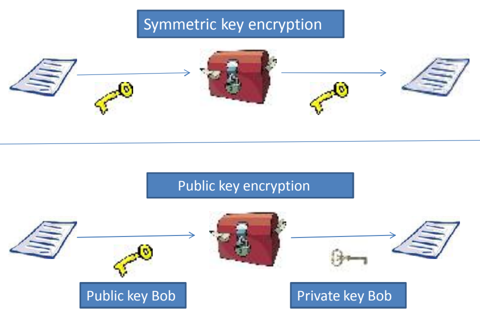
Model for Network Security



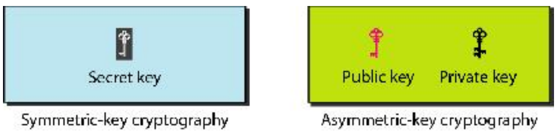
Category of cryptography



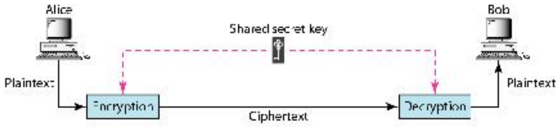
Symmetric and public key encryption



Keys in cryptography



Symmetric key cryptography

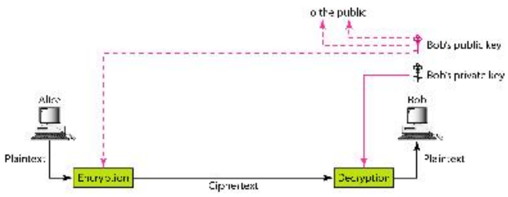


In symmetric-key cryptography, the same key is used by the sender (for encryption) and the receiver (for decryption). The key is shared.

Requirements for symmetric encryption

- two requirements for secure use of symmetric encryption:
  - a strong encryption algorithm
  - a secret key known only to sender / receiver
- mathematically have:
  $Y = E(K, X) = E_K(X) = \{X\}_K$
  $X = D(K, Y) = D_K(Y)$
- assume encryption algorithm is known
  - Kerckhoff's Principle: security in secrecy of key alone, not in obscurity of the encryption algorithm
- implies a secure channel to distribute key
  - Central problem in symmetric cryptography

Public (Asymmetric) key cryptography

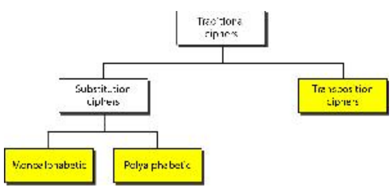


SYMMETRIC-KEY CRYPTOGRAPHY

Symmetric-key cryptography started thousands of years ago when people needed to exchange secrets (for example, in a war). We still mainly use symmetric-key cryptography in our network security.

2.46

Figure : *Traditional ciphers*



2.47

Note

A substitution cipher replaces one symbol with another.

2.48

Example

The following shows a plaintext and its corresponding ciphertext. Is the cipher monoalphabetic?

Plaintext: HELLO
Ciphertext: KHOOR

Solution

The cipher is probably monoalphabetic because both occurrences of L's are encrypted as O's.

2.49

Example 2.2

The following shows a plaintext and its corresponding ciphertext. Is the cipher monoalphabetic?

Plaintext: HELLO
Ciphertext: ABNZF

Solution

The cipher is not monoalphabetic because each occurrence of L is encrypted by a different character. The first L is encrypted as N; the second as Z.

2.50

Note

The shift cipher is sometimes referred to as the Caesar cipher.

2.51

Example 2.3

Use the shift cipher with key = 15 to encrypt the message "HELLO."

Solution

We encrypt one character at a time. Each character is shifted 15 characters down. Letter H is encrypted to W. Letter E is encrypted to T. The first L is encrypted to A. The second L is also encrypted to A. And O is encrypted to D. The cipher text is WTAAD.

2.52

Example 2.4

Use the shift cipher with key = 15 to decrypt the message "WTAAD."

Solution

We decrypt one character at a time. Each character is shifted 15 characters up. Letter W is decrypted to H. Letter T is decrypted to E. The first A is decrypted to L. The second A is decrypted to L. And, finally, D is decrypted to O. The plaintext is HELLO.

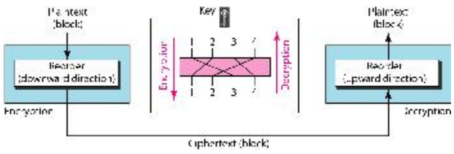
2.53

Note

A transposition cipher reorders (permutes) symbols in a block of symbols.

2.54

Figure : *Transposition cipher*



Example 2.5

Encrypt the message "HELLO MY DEAR," using the key shown in Figure .

Solution

We first remove the spaces in the message. We then divide the text into blocks of four characters. We add a bogus character Z at the end of the third block. The result is HELL OMYD EARZ. We create a three-block ciphertext ELHLMDOYAZER.

Example 2.6

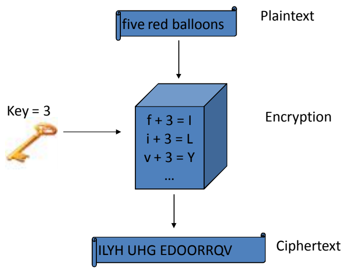
Using Example , decrypt the message "ELHLMDOYAZER".

Solution

The result is HELL OMYD EARZ. After removing the bogus character and combining the characters, we get the original message "HELLO MY DEAR."

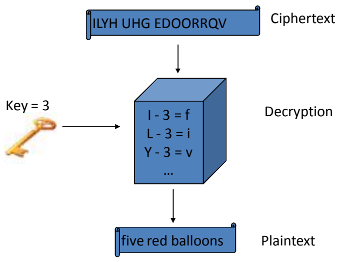
The Shift Cipher: Ceaser cipher

- We "shift" each letter over by a certain amount



The Shift Cipher cont.

- To decrypt, we just subtract the key



What's wrong with the shift cipher?

- Not enough keys!
- If we shift a letter 26 times, we get the same letter back
  - A shift of 27 is the same as a shift of 1, etc.
  - So we only have 25 keys (1 to 25)
- Eve just tries every key until she finds the right one

The Substitution Cipher

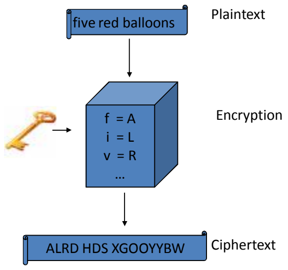
- Rather than having a fixed shift, change every plaintext letter to an arbitrary ciphertext letter

Plaintext	Ciphertext
a	G
b	X
c	N
d	S
e	D
…	…
z	Q

The Substitution Cipher cont.

Key =

a	G	n	B
b	X	o	Y
c	N	p	Z
d	S	q	P
e	D	r	H
f	A	s	W
g	F	t	I
h	V	u	J
i	L	v	R
j	M	w	U
k	C	x	K
l	O	y	T
m	E	z	Q



The Substitution Cipher cont.

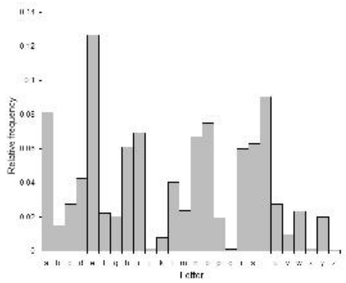
- To decrypt we just look up the ciphertext letter in the table and then write down the matching plaintext letter
- How many keys do we have now?
  - A key is just a permutation of the letters of the alphabet
  - There are 26! permutations
    - 403291461126605635584000000
- What's wrong with this substitution Cipher?

Frequency Analysis

- In English (or any language) certain letters are used more often than others
- If we look at a ciphertext, certain ciphertext letters are going to appear more often than others
- It would be a good guess that the letters that occur most often in the ciphertext are actually the most common English letters

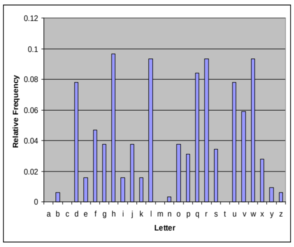
Letter Frequency

- This is the letter frequency for English
- The most common letter is 'e' by a large margin, followed by 't', 'a', and 'o'
- 'J', 'q', 'x', and 'z' hardly occur at all

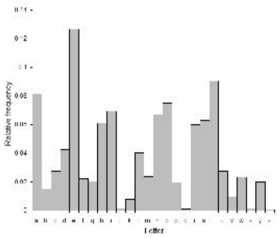


Frequency Analysis in Practice

- Suppose this is our ciphertext
  - dq lqwurgxfwlrq wr frpsxwlqj surylglqj d eurdg vxuyhb ri wkh glvflsolqh dqg dq lqwurgxfwlrq wr surjudpplqj. vxuyhb wrslfv zloo eh fkrvhq iurp: ruljlqv ri frpsxwhuv, gdwd uhsuhvhqwdwlrq dqg vwrudjh, errohdq dojheud, gljlwdo orjlf jdwhv, frpsxwhu dufklwhfwxuh, dvvhpeohuv dqg frpslohuv, rshudwlqj vbvwhpv, qhwzrunv dqg wkh lqwhuqhw, wkhrulhv ri frpsxwdwlrq, dqg duwlilfldo lqwhooljhqfh.



Ciphertext distribution



English distribution

In our ciphertext we have one letter that occurs more often than any other (h), and 6 that occur a good deal more than any others (d, l, q, r, u, and w)

There is a good chance that h corresponds to e, and d, l, q, r, u, and w correspond to the 6 next most common English letters

Frequency Analysis cont.

- If we replace 'e' with 'h' and the 6 next most common letters with their matches, the ciphertext becomes
  - an intro???tion to ?o?p?tin? pro?i?in? a ?roa? ??r?e? o? t?e ?i??ip?ine an? an intro???tion to pro?ra??in?. ??r?e? topi?? ?i?? ?e ??o?en ?ro?: ori?in? o? ?o?p?ter?, ?ata repre?entation an? ?tora?e, ?oo?ean a??e?ra, ?i?ita? ?o?i? ?ate?, ?o?p?ter ar??ite?t?re, a??e???er? an? ?o?pi?er?, operatin? ???te??, net?or?? an? t?e internet, t?eorie? o? ?o?p?tation, an? arti?i?ia? inte??i?en?e.

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

- 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	O	N	A	R
C	H	Y	B	D
E	F	G	I/J	K
L	P	Q	S	T
U	V	W	X	Z

Encrypting and Decrypting

- plaintext is encrypted two letters at a time
  1. if a pair is a repeated letter, insert filler like 'X'
  2. 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)
  4. 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 $K = k_1\ k_2\ ...\ k_d$
- i^{th} letter specifies i^{th} 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: deceptivedeceptivedeceptive
plaintext: wearediscoveredsaveyourself
ciphertext:ZICVTWQNGRZGVTWAVZHCQYGLMGJ

Aids

- 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
- eg repeated “VTW” in previous example
- suggests size of 3 or 9
- then attack each monoalphabetic cipher individually using same techniques as before

Autokey Cipher

- ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher
- with keyword is 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

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:

```
m e m a t r h t g p r y
 e t e f e t e o a a t
```

- 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:        3 4 2 1 5 6 7
Plaintext:  a t t a c k p
            o s t p o n e
            d u n t i l t
            w o a m x y z
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

Hill Cipher

- We have explored three simple substitution ciphers that generated ciphertext C from plaintext p by means of an arithmetic operation modulo 26.
- Caesar cipher: The Caesar cipher is an additive cipher. $C = p + k \pmod{26}$. The number of keys is 26 . Decryption is accomplished by adding the additive inverse of the key to ciphertext $p = C - k \pmod{26}$.
- Multiplicative cipher: $C = p \times k \pmod{26}$. The number of keys is 12 .Decryption is accomplished by multiplying ciphertext by the multiplicative inverse of the key $p = C \times inv(k) \bmod 26$.
- Affine cipher: The affine cipher composes the multiplicative cipher and the Caesar cipher. (We will do the multiplicative cipher first and the Caesar cipher second.) C=(multiplicative key)x p+(additive key) mod 26 . The number of keys is $12 \times 26 = 312$, one of which produces plaintext.
- Each of these can be attacked by frequency analysis – each ciphertext letter inherits all the frequency characteristics of the plaintext letter it replaces. It is easy to spot high frequency letters (e, t, a, o, i, n, s).
- One way to destroy the value of frequency analysis is to encrypt a string of letters as one block.

Hill Cipher

- Developed by the mathematician Lester Hill in 1929.
- The encryption algorithm takes m successive plain text and substitute for them m cipher text letters.
- Each character is assigned a numerical value (a=0,…z=25).

$$\begin{pmatrix} C_1 \\ C_2 \\ C_3 \end{pmatrix} = \begin{pmatrix} K_{11}K_{12}K_{13} \\ K_{21}K_{22}K_{23} \\ K_{31}K_{32}K_{33} \end{pmatrix} \begin{pmatrix} P_1 \\ P_2 \\ P_3 \end{pmatrix} \bmod 26$$

$$C = KP \bmod 26$$

$$P = K^{-1}C \bmod 26 = KK^{-1}P = P$$

Classical to Modern Cryptography

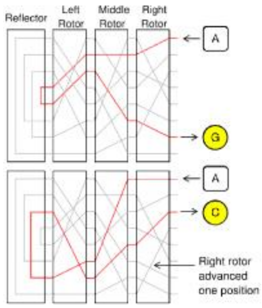
- Classical cryptography
  - Encryption/decryption done by hand
- Modern cryptography
  - Computers to encrypt and decrypt
  - Same principles, but automation allows ciphers to become much more complex

The Enigma Machine

- German encryption and decryption machine used in WWII
- Essentially a complex, automated substitution cipher

How did Enigma work?

- Rotors have different wiring connecting input to output
- Rotors move after each keypress
- The key is the initial position of the three rotors



Breaking the Enigma

- Britain set up its cryptanalysis team in Bletchley Park
- They consistently broke German codes throughout the war
- Important location in the history of computing
  - Alan Turing: British Cryptanalyst
  - COLOSSUS: used by British codebreakers for Cryptanalysis

Cryptography in the Computer Age

- Working with binary instead of letters
- We can do things many, many times
  - Think of an Enigma machine that has 2^{128} pairs of symbols on each rotor, and 20 rotors
- Other than that, the basic principles are the same as classical cryptography

Modern Ciphers

- We design one relatively simple scrambling method (called a round) and repeat it many times
  - Think of each round as a rotor on the Enigma
  - One round may be easy to break, but when you put them all together it becomes very hard
- Almost all ciphers follow one of two structures
  - SPN (Substitution Permutation Network)
  - Feistel Network (basis for DES)
  - These describe the basic structure of a round

Modern Ciphers in Practice

- Follow SPN/Feistel structure in general, but with added twists for security
- There are two important ciphers in the history of modern cryptography
  - DES (Data Encryption Standard)
  - AES (Advanced Encryption Standard)

Encryption:Stream and block cipher