CS302 Information Security and Cryptography – Lecture 3
B Tech III Div. A and B
(Jan-Apr 2023) – 19 Jan 2023

Dr Dhiren Patel

NIT Surat, India

Web - Universal Client (browser, apps)
Any thing…. Any where….. Convergence

Attacks on ICT: Motivation

- Theft of sensitive info.
- Disruption of service
- Illegal access to resources

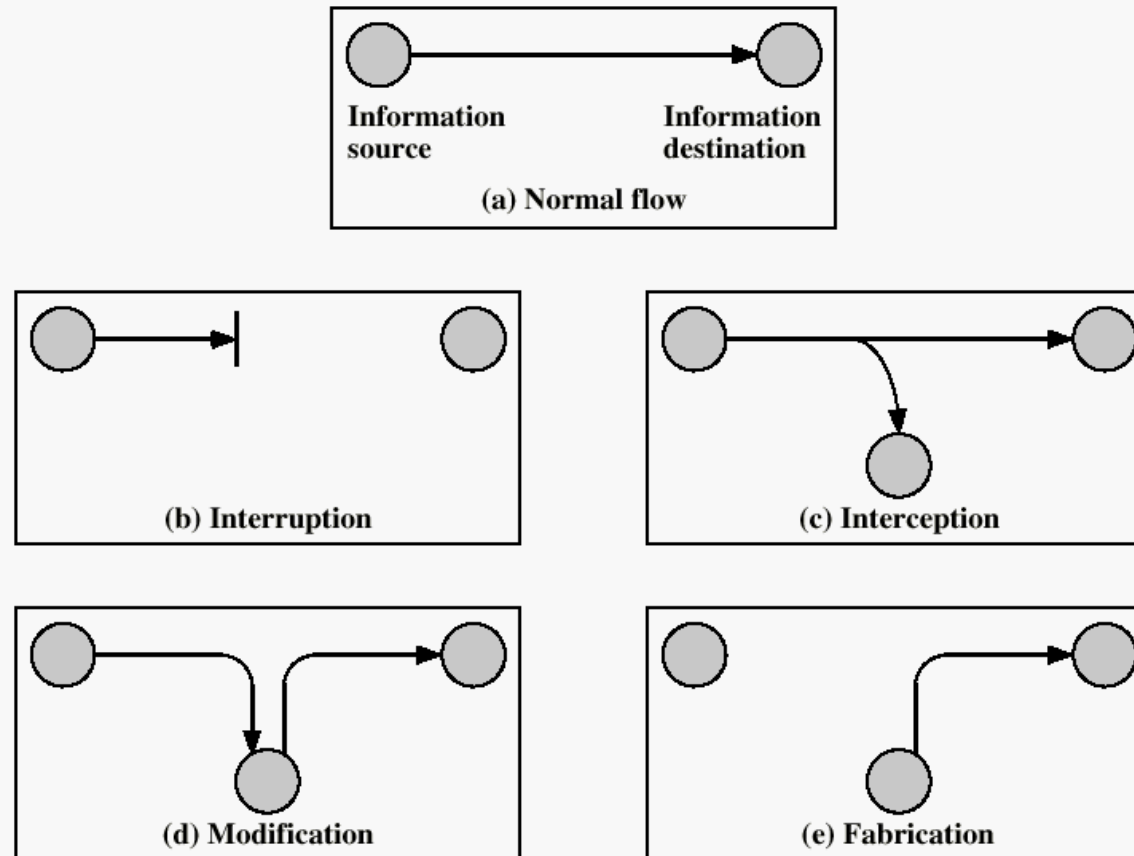# Attacks – on confidentiality, integrity, masquerading, non-repudiation, replay,
<example: going to Cinema - India>

Figure 1.1   Security Threats

Usable Apps – attack targets
(when launched! – Now Secure??)

Universal client

Changing landscape (and Security scope)

INTERNET

INDUSTRIAL IOT

NETWORK CENTRIC

SENSORS

INDUSTRIAL APPS

MONITORING TOOLS

INDUSTRY 4.0

QUALITY

AUTOMATION
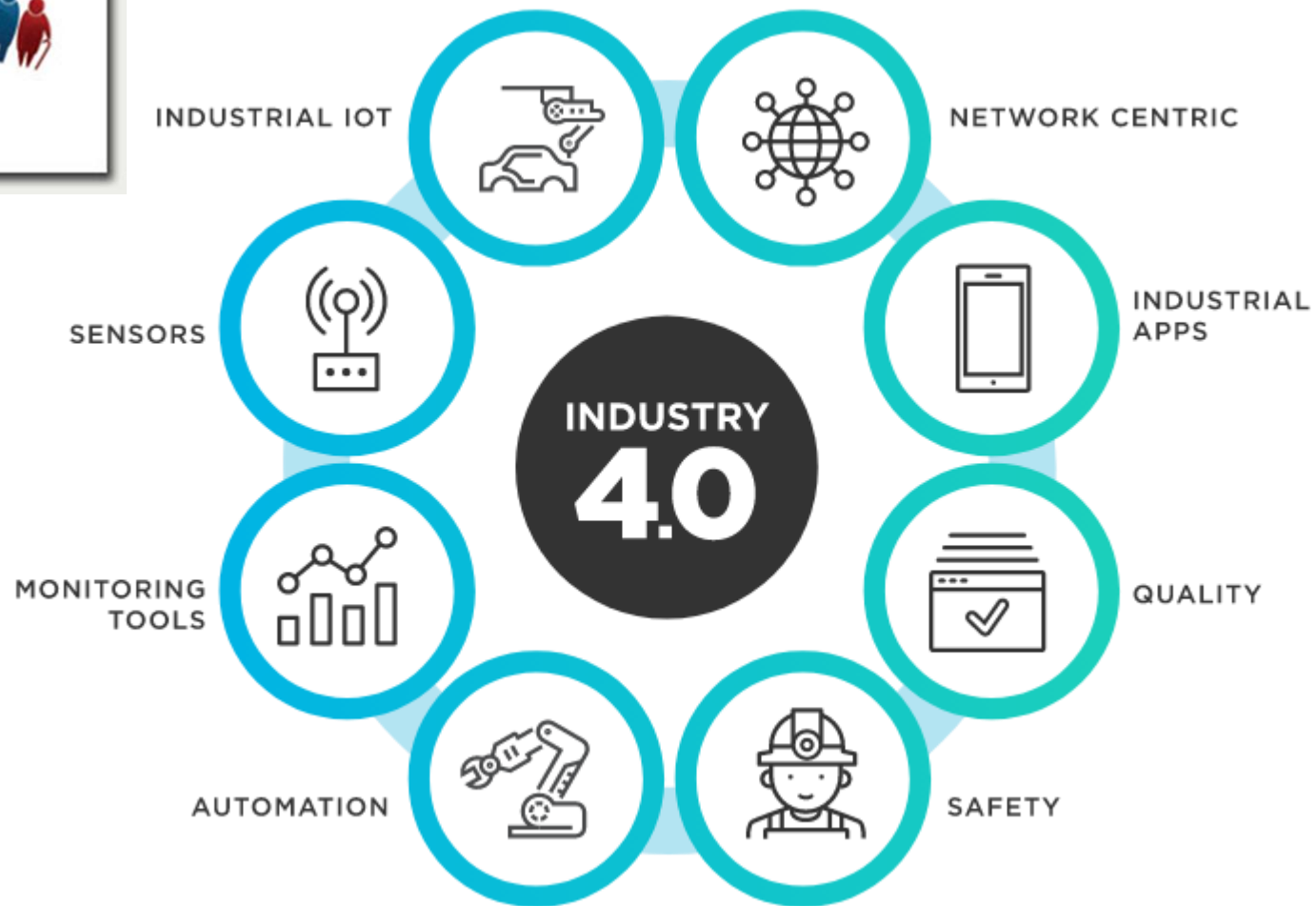
SAFETY

IIoT

- Digital transformation of industries toward a demand service model

- IIoT is used across a range of industries from manufacturing, logistics, oil and gas, transportation, mining, aviation, energy, and more.

- Its focus is to optimize operations--particularly the automation of processes and maintenance.

- Trans disciplinary (Inter disciplinary) systems engineering

Remote Monitoring and Diagnostics (maintenance and health audit of …..)

Smart Factories

- An interconnected network of machines, communication mechanisms, and computing power, the smart factory is a cyber-physical system.

- It uses (advanced) technologies such as artificial intelligence (AI) and machine learning to analyze data, drive automated processes, and learn as it goes.

Data Security, Privacy, Access Control

Security Primitives and Solutions

- Steganography (hide)
- Encryption – Symmetric, Asymmetric
- Hash functions (ability to detect change/temporing)
- Key exchange (prior to communicating encrypted data)
- Key life time (example - Data provenance)
- Biometric (system with errors – error margin)
- CAPTCHA (example – click advertisement, eye ball count)
- Embedded Security
- Authentication, Identity Management and Access control
- Multi-factor
- Blockchain (???) - secrecy v/s transparency

Information Security

- Safeguarding data at transit, at store (rest), on cloud, on-prim and off-prim, access control….
- Encrypting information – symmetric/asymmetric keys
- Replace encrypted data?? Decrypting key won't work – attack!!
- Data integrity, Hash function – message digest
- Message Authentication Code (MAC)
- Digital signatures

Systems security

- Systems – Server, Ports, OS, Applications, Web, Mail, DNS, Cloud ??? Mobile ??? GPS???
- CIA - Confidentiality – ciphers (e.g. AES)
- Integrity – hash functions (e.g. SHA3)
- Authentication – RBAC, Digital Certificate
- Attack on availability (lock account after three unsuccessful login attempts!)
- ACL – Access Control List (example – access control matrix for users and resources)
- Virus/Worm/Malware
- Security is a process not a product!

Network Security

- Firewalls, DMZ
- IPSEC, SSL
- Secure channel
- Security Associations (example – interbank communications – router configuration)
- Encrypted payload
- Secure Routing (example – QoS, strict path, trusted routers, group routers)
- Intrusion Detection System (IDS)
- Intrusion Prevention System (IPS)

Classical Cryptography

- Classical Cryptography had only one objective: *Confidentiality* → keep information (and communication) secret

- Kings' secret message services //msg. on head??

- two basic components of classical ciphers: <u>*substitution*</u> (letters are replaced by other letters)

- and <u>*transposition*</u> (hide the message contents by rearranging the order of the letters)

Classical Ciphers

- *classical ciphers* -- illustrate important basic principles and common pitfalls
- Ciphers are further classified as:
- *monoalphabetic* - only one substitution/ transposition is used
- *polyalphabetic* - where several substitutions/ transpositions are used
- several such ciphers may be concatenated together to form a *product cipher*
- Steganography (hiding / concealing information)
- Machine ciphers (Rotor machines)

Transposition Cipher – Example 1
(Rail fence)

- Text is written down as a sequence of diagonals (Rail fence) and then read of as a sequence of rows.

- *Plain text*: Meet at five pm behind P lab.

- Written as *Rail fence* of depth 2:

m		e		a		f		v		p		b		h		n		p		a	
	e		t		t		i		e		m		e		i		d		l		b

- Encrypted as: mea fvp bhn pae tti eme idl b

- Where to cut the halves? (message length should be even, append x at the end if required)

Transposition cipher - example 2 (Permutation)

- a simple transposition with t = 6 and e = (6 4 1 3 5 2) as a *permutation* on the set.

- The message m = CAESAR is encrypted to c = RSCEAA.

- Decryption uses the inverse permutation d = (3 6 4 2 5 1).

- A mnemonic keyword may be used in place of a key. For example, for n = 6, the keyword "CIPHER" could be used to specify the column ordering 1, 4, 5, 3, 2, 6 (by alphabetic priority).

- Sequential composition of two or more simple transpositions with respective periods t_1, t_2, ... , t_i is called a *compound transposition*.

Transposition cipher – example 3

- write a message in a rectangle (square <u>matrix</u>), row by row and read it off, column by column.
- Plain text message: meet at five pm behind p lab

m	e	e	t	a
t	f	i	v	e
p	m	b	e	h
i	n	d	p	l
a	b	x	x	x

- Encrypted text (5x5 matrix): <read column vise>
- mtpia efmnb eibdx tvepx aehlx
- xxx is appended to make message 25 character long.
- Block size (matrix size and dimension)

Transposition cipher (Cont.)

- Additionally, a key can also be defined to permute the order of the columns. Write text in a row by row..

m	e	e	t	a
t	f	i	v	e
p	m	b	e	h
i	n	d	p	l
a	b	x	x	x

- E.g. key (41523) defines -- read it off – 4th column first, 1st column second, 5th column third, followed by 2nd and 3rd column, and prepare encrypted text.

- Encrypted text is:

- tvepx mtpia aehlx efmnb eibdx

Substitution Ciphers

- Cipher text - FJKTGP –
- making any sense?
- Encrypted with shift cipher!
- Original letters are shifted by 2 positions – Right
- D → F ….
- Decrypt it – Try … F – 2 = D …
- P – 2 = N…

Simple!!

F	J	K	T	G	P						
D	H	I	R	E	N						
M	R	P	A	T	E	L					
O	T	R	C	V	G	N					

Substitution Ciphers

- Symbols of the text are substituted with another symbols of the same alphabet.
- DHIREN
- FJKTGP
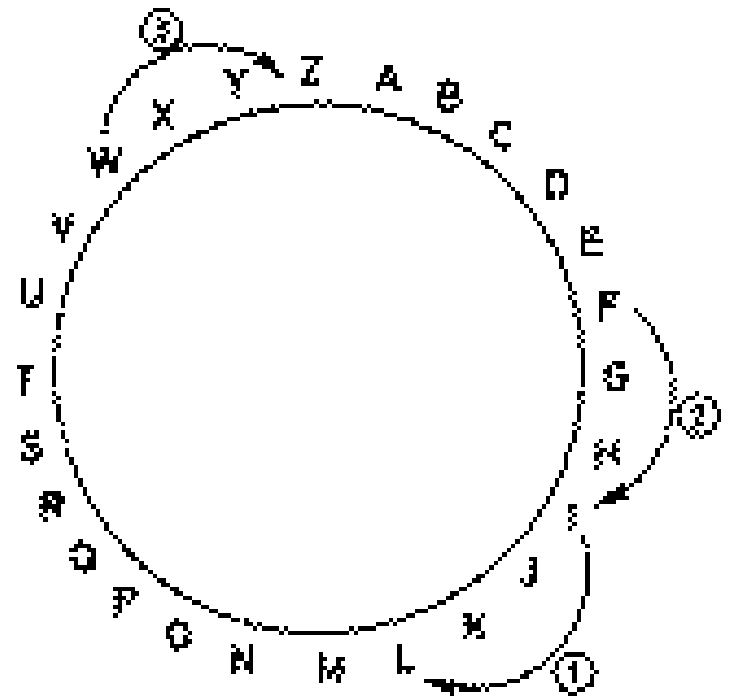
Cryptosystem: generic Definition

A CRYPTOSYSTEM is a 5-tuple (P,C,K ,E,D) satisfying

1. P is a finite set of possible plaintexts
2. C is a finite set of possible ciphertexts
3. K is a finite set of possible keys
4. E is a finite set of encryption rules indexed by K so for each K there is a function $e_K : P \rightarrow C$
5. D is a finite set of decryption rules indexed by K so for each K there is a function $d_K : C \rightarrow P$

Caesar cipher (Shift cipher)

- permutation is constrained to an alphabetic shift through k characters (for some fixed k)
- <u>Encryption</u> is defined by
- $c_i = e(m_i) = (p + k) \bmod s$
- <u>Decryption</u> is defined by
- $d(c_i) = (c - k) \bmod s$
- For English text, $s = 26$ and characters A through Z are associated with integers p=0 through p=25
- According to history, Julius Caesar used the key $k = 3$ (DRP is encrypted to GUS).

Shift Cipher - revisited

The **shift cipher** is the cryptosystem defined by taking

- $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{z}_{26}$

- $e_K(x) = (x + K) \bmod 26$

- $d_K(y) = (y - K) \bmod 26$

Letters are identified with numbers:
A=0, B=1, ..., Z=25

$\mathbb{Z}_{26}$ denotes the set $\{0, 1, ..., 25\}$ with addition and multiplication taken modulo 26

Breaking simple Substitution ciphers

Mono-alphabetic substitution

- letters of the plain text alphabet are mapped on to unique letters throughout the entire message text

- cipher can be trivially broken because

i. The language of the plain text is easily recognizable. (frequency distribution-unigram statistics - next slide)

ii. There are only $s = |A|$ keys (e.g. for Roman alphabet, only 25 keys – 1 to 25) to search exhaustively

- Exhaustive key search is always possible <make it practically infeasible is the goal>

Frequency of single characters in English text
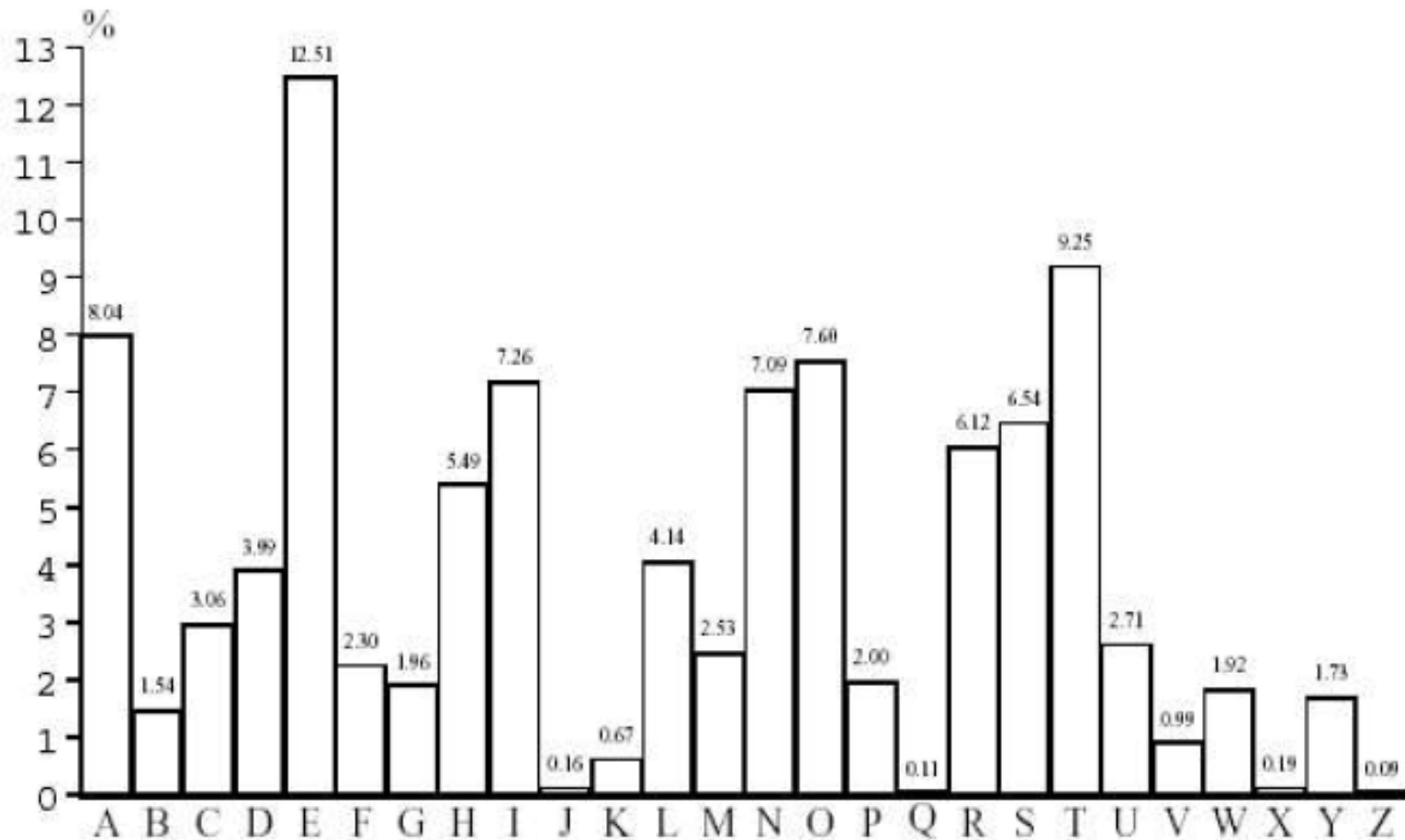
Polygram substitution

- groups of characters being substituted (simultaneously) by other groups of characters from the same alphabet

- sequences of two plaintext characters (*digrams*) be replaced by other digrams;

- key space is large (digrams - 26^2!)

- Applicable to trigrams to n-grams

Playfair cipher

- By Sir Charles Wheatstone in 1854 (named for his friend B. Playfair)

- A digram substitution defined by arranging the characters of a 25-letter alphabet (*I* and *J* are equated) in a 5 x 5 matrix as a key (example on next slide)

- A mnemonic aid (a meaningful keyphrase) may be used to easily remember the 5x5 square

- repeated letters are not considered again and the remaining characters included alphabetically at the end.

Example – Key Matrix

- The key phrase "PLAYFAIR IS A DIGRAM CIPHER" would define a square with rows PLAYF, IRSDG, MCHEB, KNOQT, UVWXZ.

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

Playfair – Encryption Rules

- Adjacent plaintext characters are paired (p1,p2).

  p1 → c3, p2 → c4; pair (c3,c4) is defining cipher text

- Rule 1: If p_1 and p_2 are in distinct rows and columns, they define the corners of a sub-matrix, with the remaining corners c_3 and c_4, c_3 is the character in the same row as p_1.

- Rule 2: If p_1 and p_2 are in a common row, c_3 is the character immediately to the right of p_1 and c_4 that immediately right of p_2 (the first column is viewed as being to the right of the last). For decryption, use left direction.

- Rule 3: If p_1 and p_2 are in the same column, the characters immediately (circularly) below them are c_3 and c_4. For decryption, use up direction.

- Rule 4: If $p_1 = p_2$, an infrequent plaintext character (such as X) is inserted between them and the plaintext is re-grouped. E.g. word *Balloon* would become *ba lx lo on*.

Playfair cipher: Example

Plain text: Meet at five pm behind P lab.

- The key phrase "PLAYFAIR IS A DIGRAM CIPHER"
- *Plain text Written as digrams*:

ME ET AT FI VE PM BE HI ND PL AB

- *Encrypted as*:

CB QB OF GP CX IK MB SM RQ LA HF

- *Decrypted back* to plain text as:

ME ET AT FI/J VE PM BE HI/J ND PL AB.

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

Playfair cipher - variant

- Another example of designing a Playfair cipher for 26-alphabet and 10 numerals, we can use 6 x 6 matrix that can accommodate all 36 symbols (without equating I and J)
- Take a key phrase: Course No 821 Winter 2K05 Sem 7 Inf Sec and Cry by Patel H9

C	O	U	R	S	E
N	8	2	1	W	I
T	K	0	5	M	7
F	A	D	Y	B	P
L	H	9	G	J	Q
V	X	Z	3	4	6

- To avoid the trailing characters always being from the end of the alphabet, a further shift cipher could be applied to the resulting character string.

Hill cipher

- The encryption takes *n* successive plaintext letters and substitutes for them *n* ciphertext letters determined by *n* linear equations. (Mathematician *Lester Hill* in 1929)

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

- n-gram substitution may be defined using an invertible *n* x *n* matrix **A** = a_{ij} as the key to map an *n*-character plaintext $m_1 m_2 \ldots m_n$ to a ciphertext *n*-gram
- expressed in terms of column vectors and matrices:
- ***Encryption C*** = $E_K(X)$ = ***KX*** mod 26
- ***Decryption X*** = $D_K(C)$ = ***K^{-1}C*** mod 26

Hill cipher - Example

- Encrypt "Meet B" using a 2 X 2 Hill Cipher
- with the keys k = $\begin{bmatrix} 3 & 1 \\ 5 & 2 \end{bmatrix}$ and decryption key $k^{-1} = \begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$
- $c_1 = (k_{11}x_1 + k_{12}x_2) \bmod 26$
- $c_2 = (k_{21}x_1 + k_{22}x_2) \bmod 26$
- Plain text : me et bx (x added to complete last (pair) digram)
- Numerical equivalent = 12 4 4 19 1 23, read as pairs x_1x_2, x_3x_4, x_5x_6.

- $c_1 = (36 + 4) \bmod 26 = 14$ (O), $c_2 = (60 + 8) \bmod 26 = 16$ (Q)
- Encrypted as → oq fg az

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

Hill Cipher Example (Decryption)

- Decryption key K^{-1} : $\begin{bmatrix} 2 & -1 \\ -5 & 3 \end{bmatrix}$

- ***Decryption X*** $= D_K(C) = \boldsymbol{K^{-1}C}$ mod 26
- $x_1 = (k_{11}c_1 + k_{12}c_2)$ mod 26
- $x_2 = (k_{21}c_1 + k_{22}c_2)$ mod 26
- oq fg az   <14, 16><5,6><0,25>
- $x_1 = (28 - 16)$ mod 26 = 12 = m
- $x_2 = (-70 + 48)$ mod 26 = -22 = 4 = e
- me et bx

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

Hill cipher ..

- 3 X 3 Hill cipher using tri-grams
- E.g. key matrix   **K** =   and **K^{-1}** =

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 1 & 2 \\ 1 & 0 & -2 \end{bmatrix}$$

$$\begin{bmatrix} 2 & -2 & -1 \\ -4 & 5 & 3 \\ 1 & -1 & -1 \end{bmatrix}$$

- **Implementation issues**
- Key generation and distribution
- Matrix and its inverse (elements (integer))
- Matrix Calculator (open source/on-line)

Polygram ciphers (attacks)

- polygram substitution ciphers (Playfair, Hill) are linear transformation, and fall under known-plaintext attack.
- Frequency analysis (digrams)
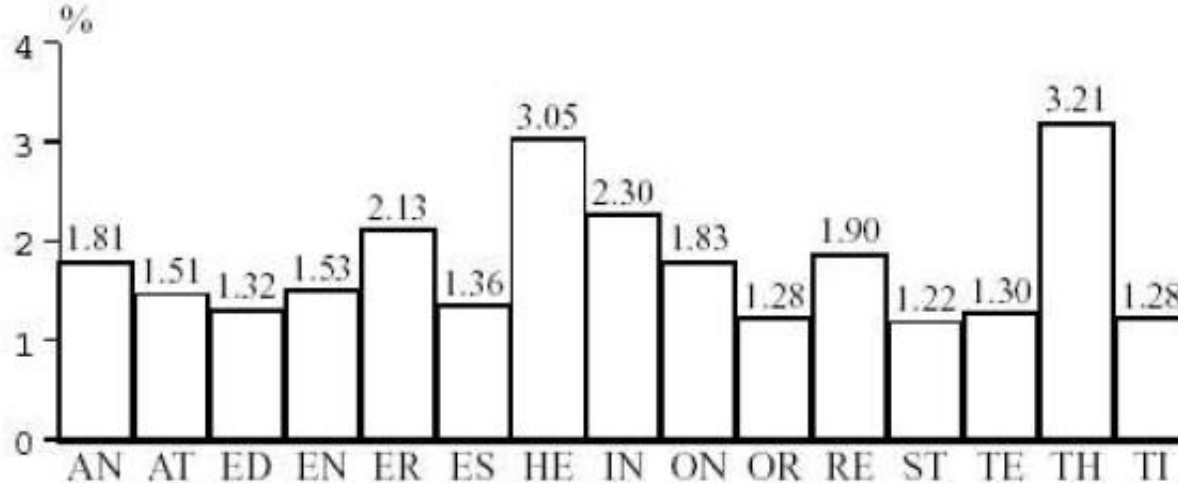
Similarly -- Trigrams - The most common trigrams (triples) in English language are: THE, ING, AND, HER, ERE, ENT, THA and NTH.

Poly-alphabetic substitutions

- the letters of the plain text alphabet are mapped into letters of cipher text space <u>depending on their position</u> in the text

- Under different alphabets, the same plaintext character is thus encrypted to different ciphertext characters, precluding simple frequency analysis of mono-alphabetic substitution

Polyalphabetic substitution

- Key is taken from some known text (with corresponding character value (0 to 25)) defining the shifting of underlying m.

- The mapping of plaintext $m = m_1\ m_2\ m_3$ ... to ciphertext $c = c_1\ c_2\ c_3$ ... is defined on individual characters by $c_i = (m_i + k_i)$ mod s, where subscript i in k_i is taken modulo t (the key is re-used)

Vigenère cipher - Vigenère table

- the Each of the 26 shift ciphers is laid out horizontally, with the key letter for each cipher to its left.
- Plain text runs across the top.
- The cipher text is at the intersection of the row labeled key letter "k" and the column labeled text letter "p".
- Eg. Plain text "baby" – encrypted using key "abcd" → bbdb

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

Vigenère cipher

- To encrypt a message, a key is needed that is as long as the message. Usually the key is a repeating keyword.

- To decrypt the cipher text, the key letter again defines the row, and the position of ciphertext letter in that row determines the column. The top letter on that column is decrypted letter.

- E.g. *Plain text*: Meet at five pm behind P lab.

- *Keyword phrase*: confidential

Plain text	m	e	e	t	a	t	f	i	v	e	p	m	b	e	h	i	n	d	p	l	a	b
Key	c	o	n	f	i	d	e	n	t	i	a	l	c	o	n	f	i	d	e	n	t	i
Cipher text	o	s	r	y	i	w	j	v	o	m	p	x	d	s	u	n	v	g	t	y	t	j
Recovered	m	e	e	t	a	t	f	i	v	e	p	m	b	e	h	i	n	d	p	l	a	b

Breaking Vigenère cipher

- The cryptanalyst looking only at the ciphertext would detect the repeated sequences of common digrams and trigrams, devise common factors in the displacements of the various sequences, and make good guess of the keyword length.

- If the keyword is of length N, then the cipher is consisting N mono-alphabetic substitution ciphers.

- One can use the known frequency characteristics of a plain text language to attack each of the mono-alphabetic ciphers separately.

- Refinement leads to eliminate periodic nature of the keyword (very long keyword would help).

Vigenère cipher

- Running-key *Vigenère*

  If the key stream k_i of a simple *Vigenère* is as long as the plaintext, the cipher is called a *running-key cipher*. For example, the key may be meaningful text from a known book (properly synchronized).

- If the key has redundancy – cryptanalyst can exploit statistical imbalances;

- E.g., when encrypting plaintext English characters using a meaningful text as a running key, cryptanalysis is possible based on the observation that a significant proportion of ciphertext characters results from the encryption of high-frequency running text characters with high-frequency plaintext characters. (unigram frequency analysis)

Vernam Cipher

- The system can be expressed as:

$c_i = m_i \oplus k_i$ – Encryption

$m_i = c_i \oplus k_i$ – Decryption

- where, $m_i = i^{\text{th}}$ binary digit of plain text
- $k_i = i^{\text{th}}$ binary digit of key material
- $c_i = i^{\text{th}}$ binary digit of cipher text
- $\oplus$ = exclusive-or (XOR) operation

Gilbert Vernam in 1918 proposed the use of a running tape that eventually repeated the key, so that the system can work with a very long (but repeating) keyword.

One-time pad – unbreakable cipher!

- A one-time pad (OTP) is a large non-repeating set of random key letters, written on sheets of paper, and glued together in a pad.

- The sender uses each key letter on the pad to encrypt exactly one plaintext character. (*Major Joseph Mauborgne* - an Army Signal Corp officer)

- The sender encrypts the message and then destroys the used pages of the pad.

- The receiver has an identical pad and uses each letter on the pad, in turn, to decrypt each letter of the cipher-text. The receiver destroys the same used pages of the pad after decrypting the message.

OTP - Vernam

- This idea can be easily extended to binary data by using a one-time pad of key bits and XOR operation (same as Vernam cipher – with non repeating random key).

- To decrypt, XOR the cipher-text with a string from an equivalent copy of the one-time pad. Everything else remains the same and security is just as perfect as there are no patterns or regularities that a cryptanalyst can use to attack.

- Problems associated with key generation (truly random), key distribution, and perfect synchronization between the sender and receiver(s).

Security of OTP

- A random key sequence "added" to a nonrandom plaintext message produces a completely random cipher-text message and no amount of computing power can break that.
- Conditional security v/s un-conditional security