Information Systems Security (SOEN321)
Introduction

Dr. Amr Youssef

Concordia Institute for Information Systems Engineering (CIISE)
Concordia University
Montreal, Canada

youssef@ciise.concordia.ca

Reference Materials

- **There is no mandatory textbook**
  - **Basically there is no single reference that covers all the topics covered in this course**
- **Handouts will be available through the course website and/or the copy-center**
- **Optional Textbook (Available at the book store)**
  - **Computer Security: Principles and Practice, by William Stallings and Lawrie Brown, Prentice Hall, 5th edition**

What should you learn in SOEN321?

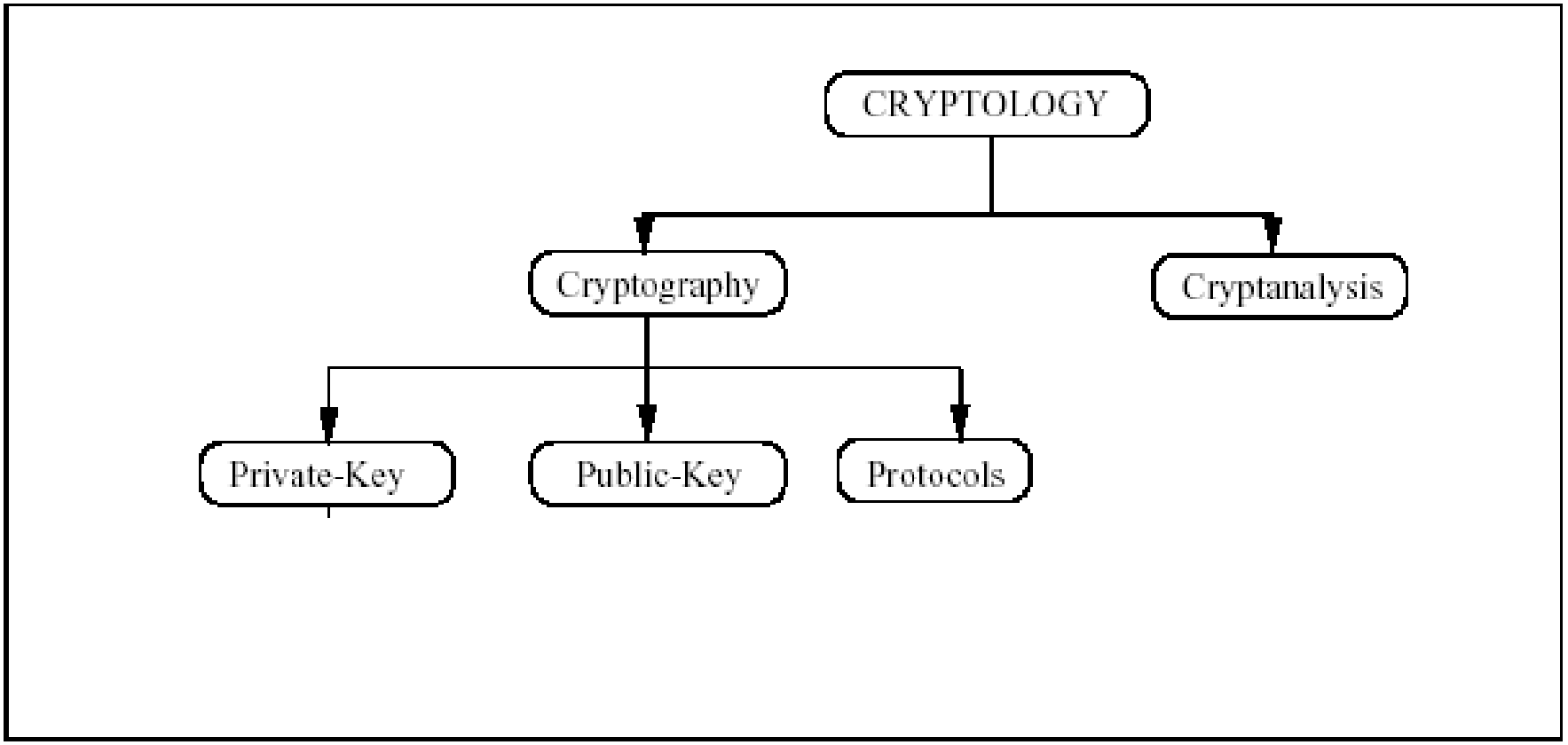
- Common Body of Knowledge (CBK) domains as specified by International Information Systems Security Certification Consortium $(ISC)^2$
- Commonly referred to as the ten domains of information security
  - Cryptography
  - Access Control
  - Application security
  - Business Continuity and Disaster Recovery Planning
  - Information Security and Risk Management
  - Legal, Regulations, Compliance and Investigations
  - Operations Security
  - Security Architecture and Design
  - Telecommunications and Network Security
  - Physical (Environmental) Security

Courses available at CIISE

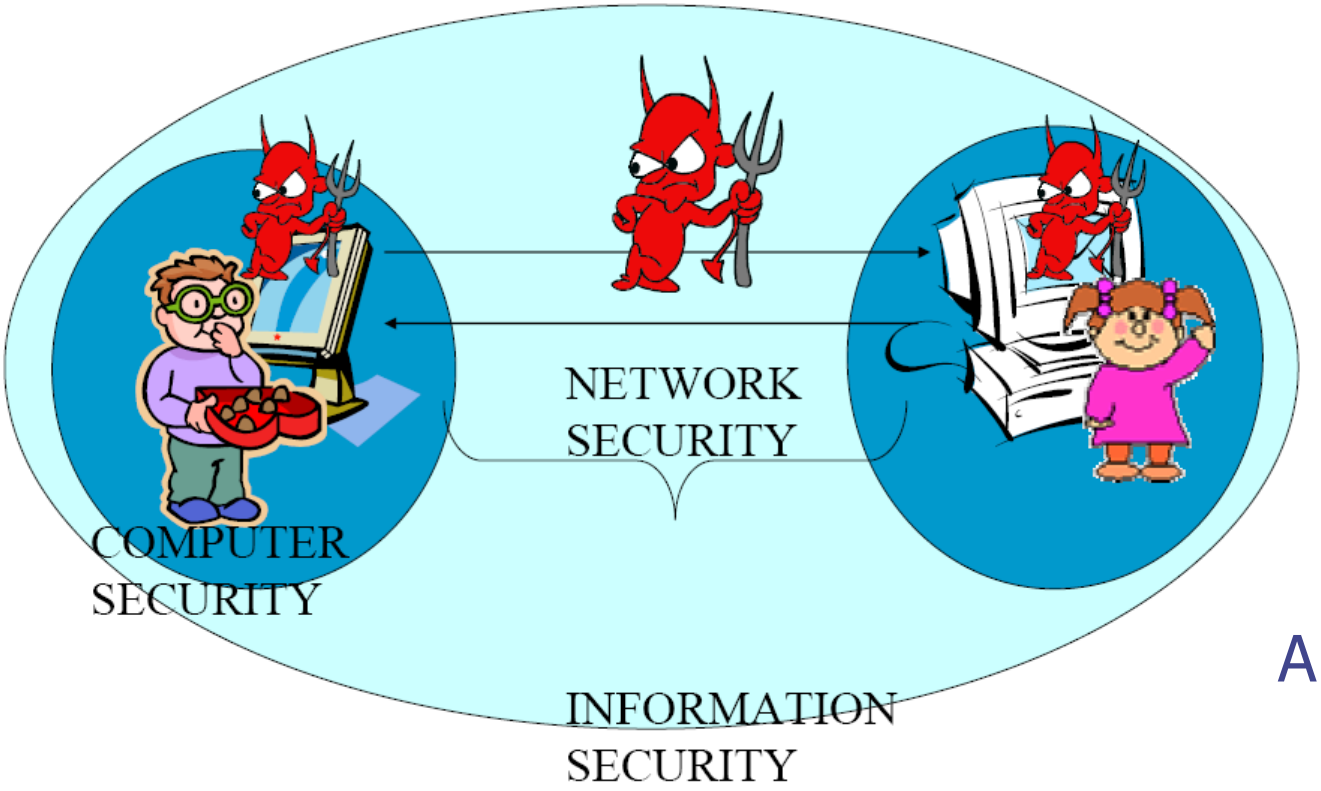
- **E69 - Topic Area: Information Systems Security**
  - INSE 6110 Foundation of Cryptography
  - INSE 6120 Crypto-Protocol and Network Security
  - INSE 6130 Operating Systems Security
  - INSE 6140 Malware Defenses and Application Security
  - INSE 6150 Security Evaluation Methodologies
  - INSE 6160 Database Security and Privacy
  - INSE 6180 Security and Privacy Implications of Data Mining
  - INSE 6190 Wireless Network Security
  - INSE 6610 Cybercrime Investigations
  - INSE 6620 Cloud Computing Security and Privacy
  - INSE 6630 Recent Developments in Information Systems Security
  - INSE 6640 Smart Grids and Control System Security
  - INSE 6650 Trusted Computing

Overview of Cryptology

- Greek: “krypto” = hide
- Cryptology – science of hiding
  = cryptography + cryptanalysis



Communication Model



Alice

Bob

Eve, Oscar, Carl

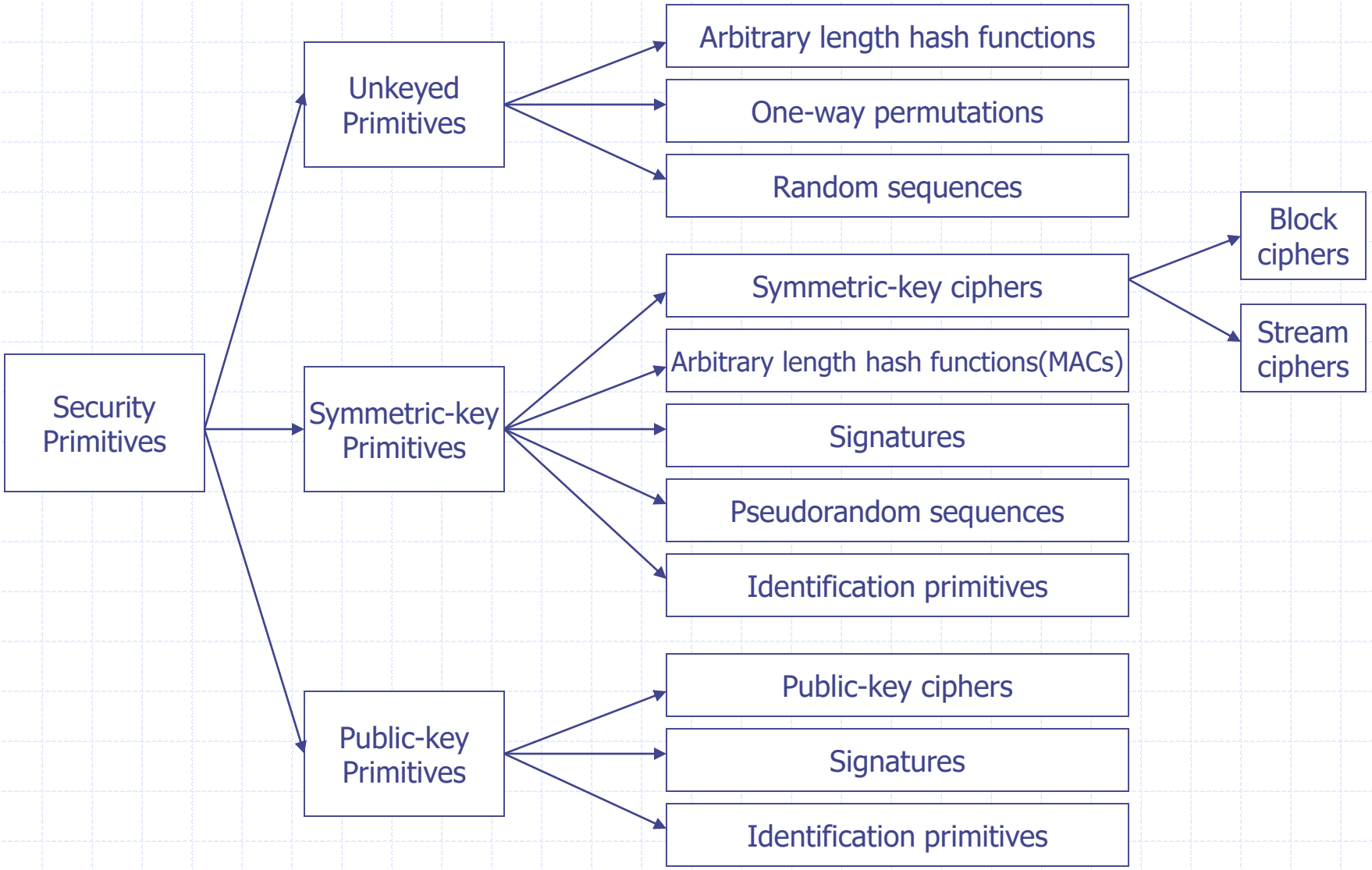
Goals of Cryptography

- Cryptography is the science that enables Alice and Bob to communicate securely in the presence of Eve
- Goals
  - Confidentiality
  - Data integrity
  - Authentication: Entity authentication (Identification) and Message authentication (Data origin authentication)
  - Non-repudiation
- Solutions: Protocols between Alice and Bob
- At least one of Alice or Bob needs to know more (or can do more) than Eve

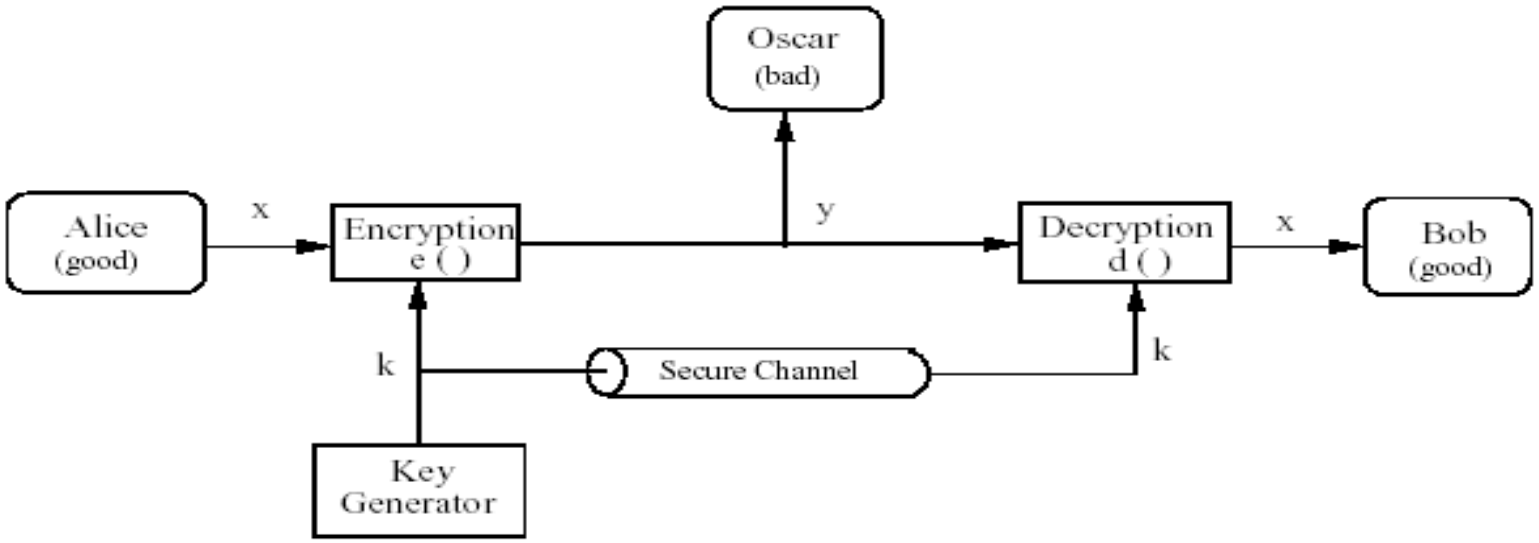
Attacker (Cryptanalyst) Model

- Passive
  - The attacker does not modify the data, only monitors the communication.
  - It threatens confidentiality.
  - Example: listen to the communication between Alice and Bob, and if it's encrypted try to decrypt it.
- Active
  - The attacker is actively involved in inserting, deleting, or modifying data.
  - It threatens authentication and confidentiality.

Taxonomy of cryptographic primitives



Symmetric Key Cryptosystem



- Advantages
  - high data throughput
  - relatively short size
- Disadvantages
  - the key must remain secret at both ends.
  - $O(n^2)$ keys to be managed.

- **Basic Terminology**
  - **plaintext** - the original message
  - **ciphertext** - the 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

Basic Terminology

- **plaintext** - the original message
- **ciphertext** - the 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

Kerckhoff's Principle

- A. Kerckhoffs was a 19th century Dutch cryptographer
- **Security should depend only on the key**
  - Don't assume enemy won't know algorithm
    - Can capture machines, disassemble programs, etc.
    - Too expensive to invent new algorithm if it might have been compromised
- **Security by obscurity doesn't work**
    - Look at history of examples
    - Better to have scrutiny by open experts
- "The enemy knows the system being used." (Claude Shannon)

Examples of Classical Ciphers

- Letters of plaintext are replaced by other letters or by numbers or symbols
- Capital/small letters
- Space

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

Shift Cipher

- Defined over Z_{26} as follows:
- Convert each letter in the plaintext P to it's
- corresponding number.
- • Key K, $0 \leq K \leq 25$.
- • Let $P = C = Z_{26}$
- • $e_k(P) = (P + K) \bmod 26$
- • $d_k(C) = (C - K) \bmod 26$

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

Example

- P = CRYPTOGRAPHYISFUN
- K = 11
- C = NCJAVZRCLASJTDQFY

- Steps
  - C → 2; 2+11 mod 26 = 13 → N
  - R → 17; 17+11 mod 26 = 2 → C
  - …
  - N → 13; 13+11 mod 26 = 24 → Y

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

Shift Cipher: Cryptanalysis

- $e_k(P) = (P + K) \bmod 26$
- Can an attacker find K?
  - YES: exhaustive search,
  - key space is small (26 possible keys).
  - Once K is found, very easy to decrypt
  - $d_k(C) = (C - K) \bmod 26$
- History: K = 3, Caesar's cipher

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

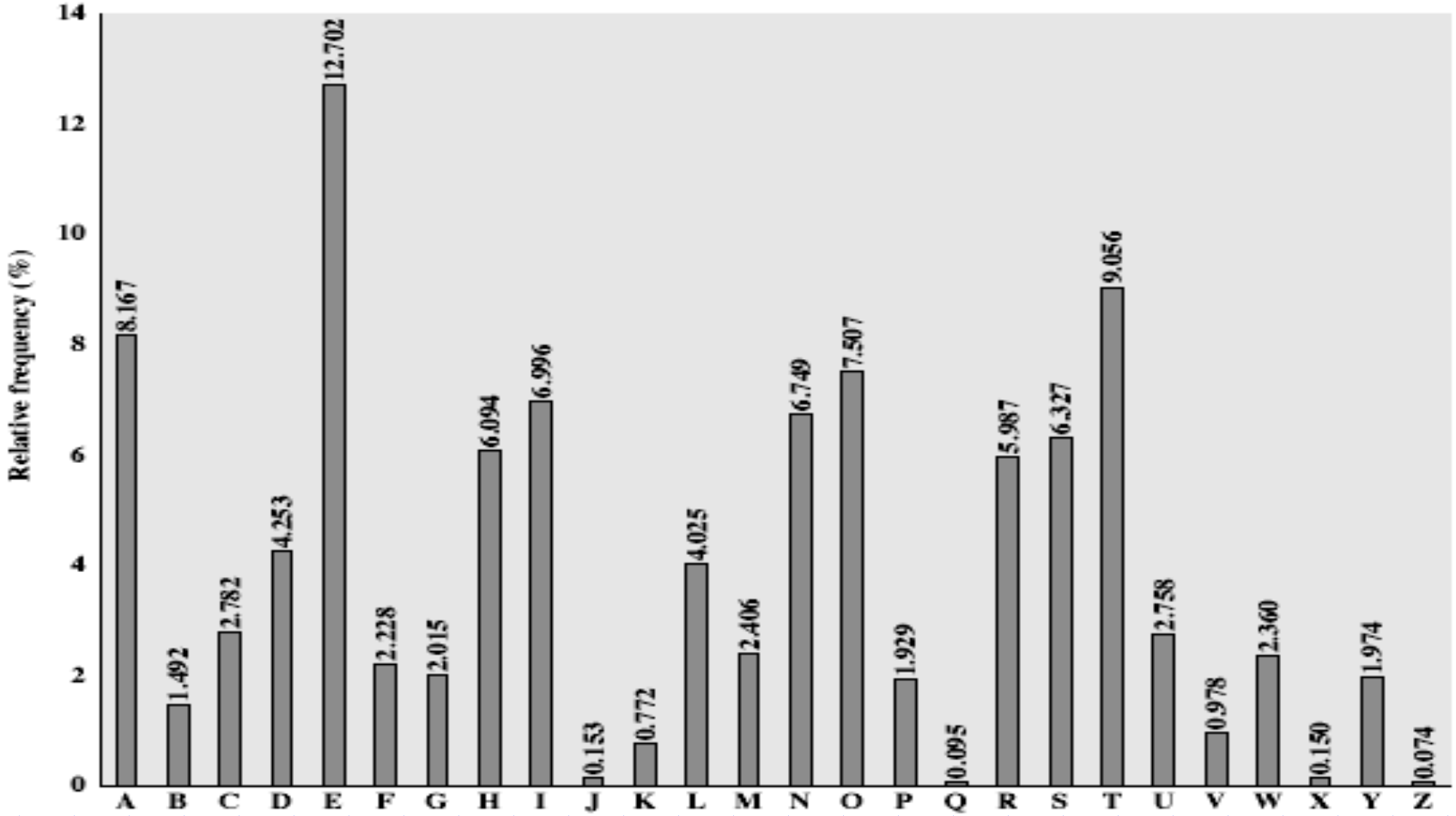
Substitution Ciphers

- Ciphertext, Plaintext $\in Z_{26}$
- • e_{Π}(Plaintext) = Π (Plaintext)
- • d_{Π}(Ciphertext) = Π^{-1} (Ciphertext)
- **Example:**
- 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
- π = B A D C E F G H I J K L M N O P Q R S T U V W X Y Z
- BECAUSE → AEDBUSE

Substitution Ciphers: Cryptanalysis

- Exhaustive search is infeasible
- – key space size is $26! \approx 4 \times 10^{26}$
- Each language has certain features: frequency of letters, or of groups of two or more letters.
- Substitution ciphers preserve these language statistics.
- Substitution ciphers are vulnerable to frequency analysis attacks.

English Letter Frequencies



Vigenere Cipher

- **Definition**:
- Given m, a positive integer, $P = C = (Z_{26})^m$, and
- K = (k1, k2, … , km) a key, we define:
- **Encryption**:
- $e_k(p1, p2 \ldots pm) = (p1+k1, p2+k2 \ldots pm+km) \pmod{26}$
- **Decryption**:
- $d_k(c1, c2 \ldots cm) = (c1-k1, c2-k2 \ldots cm- km) \pmod{26}$
- **Example:**
- Plaintext: C R Y P T O G R A P H Y
- Key: L U C K L U C K L U C K
- Ciphertext: N L A Z E I I B L J J I

One Time Pad

- The one-time pad, which is a provably secure cryptosystem, was developed by Gilbert Vernam in 1918.
- $X = Y = K = (Z_m)^n$
- $X = (x1\ x2\ \ldots\ xn)$
- $K = (k1\ k2\ \ldots\ kn)$
- $Y = (y1\ y2\ \ldots\ yn)$
- $e_k(X) = (x1+k1\ x2+k2\ \ldots\ xn+kn) \bmod m$
- $d_k(Y) = (y1-k1\ y2-k2\ \ldots\ yn-kn) \bmod m$

- Provides perfect secrecy
- Disadvantages:
  - The size of key must be at least the size of the message
  - Each key is used only once. Otherwise, vulnerable to know plaintext attack.

OTP (Binary Example)

- The message is represented as a binary string (a sequence of 0's and 1's using a coding mechanism such as ASCII coding.
- The key is a truly random sequence of 0's and 1's of the same length as the message.
- The encryption is done by adding the key to the message modulo 2, bit by bit. This process is often called *exclusive or*, and is denoted by *XOR.* The symbol $\oplus$ is used.

- **Example:**
  - Let the message be IF then its ASCII code be (1001001 1000110) and the key be (1010110 0110001).
  - The ciphertext can be found xoring message and key bits
  - *Encryption:*

1001001 1000110 plaintext
1010110 0110001 key
0011111 1110110 ciphertext

  - *Decryption:*

0011111 1110110 ciphertext
1010110 0110001 key
1001001 1000110 plaintext

p	k	$c = p \oplus k$
0	0	0
0	1	1
1	0	1
1	1	0

Provable security of the OTP

- The security depends on the randomness of the key.
- It is hard to define randomness.
- In cryptographic context, we seek two fundamental properties in a binary random key sequence:
- Unpredictability: Independent of the number of the bits of a sequence observed, the probability of guessing the next bit is not better than ½. Therefore, the probability of a certain bit being 1 or 0 is exactly equal to ½.
- Balanced (Equal Distribution): The number of 1's and 0's should be equal.

Mathematical Proof (sketch)

- the probability of a key bit being 1 or 0 is exactly equal to ½.
- The plaintext bits are not balanced. Let the probability of 0 be x and then the probability of 1 turns out to be 1-x.

m_i	prob.	k_i	prob.	c_i	prob.
0	x	0	½	0	½ x
0	x	1	½	1	½ x
1	1-x	0	½	1	½ (1-x)
1	1-x	1	½	0	½ (1-x)

- Let us calculate the probability of ciphertext bits.
- We find out the probability of a ciphertext bit being 1 or 0 is equal to $(½)x + (½)(1-x) = ½$. Ciphertext looks like a random sequence.

Affine Cipher

- $e_k(x) = y = (\alpha \ \ x + \beta) \bmod 26$.
- The key $k = (\alpha, \beta)$ and $\alpha, \beta \in Z_{26}$
- $d_k(x) = x = \alpha\text{-}1 \cdot y + \beta$

- **Example:** $k = (\alpha, \beta) = (13, 4)$

- INPUT = (8, 13, 15, 20, 19) $\Rightarrow$ ERRER
- ALTER = (0, 11, 19, 4, 17) $\Rightarrow$ ERRER

- There is no one-to-one map btw plaintext and
- ciphertext space. What went wrong?

Affine Cipher: Valid Key Space

- β can be any number in Z_{26} . 26 possibilities
- Since α^{-1} has to exist we can only select integers in Z_{26}
- s.t. gcd(α, 26) = 1. Candidates are {1, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, 25}
- Therefore, the key space has $12 \cdot 26 = 312$ candidates.

Attacks on encryption schemes

- Ciphertext-only attack
  - deduce the decryption key or plaintext by only observing ciphertext.
- Known-plaintext attack
  - using a quantity of plaintext and corresponding ciphertext.
- Chosen-plaintext attack
  - chooses plaintext and is then given corresponding ciphertext.
- Adaptive chosen-plaintext attack
  - chosen-plaintext attack where the choice of plaintext may depend on the ciphertext received from previous requests.
- Chosen-ciphertext attack
  - selects the ciphertext and is then given the corresponding plaintext.
- Adaptive chosen-ciphertext attack
  - chosen-ciphertext attack where the choice of ciphertext may depend on the plaintext received from previous requests.

Affine Cipher: Cryptanalysis

- **Attack types:**
  - *Ciphertext only:* exhaustive search or frequency analysis
  - *Known plaintext:* two letters in the plaintext and corresponding ciphertext letters would suffice to find the key.
    Example : plaintext: IF=(8, 5) and ciphertext PQ=(15, 16)
    $8 \cdot \alpha + \beta \equiv 15 \bmod 26$
    $5 \cdot \alpha + \beta \equiv 16 \bmod 26 \quad \Rightarrow \quad \alpha = 17$ and $\beta = 9$
  - *Chosen plaintext:* Chose A and B as the plaintext. The first character of the ciphertext will be equal to $0 \cdot \alpha + \beta = \beta$ and the second will be $\alpha + \beta$.

Hill Cipher

- Block Cipher
- Let n=3 and the key matrix be $M = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 11 & 9 & 8 \end{pmatrix}$

- and the plaintext be ABC = (0, 1, 2) then the encryption operation is a vector-matrix multiplication

$$(0,1,2) \times \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 11 & 9 & 8 \end{pmatrix} \equiv (0,23,22) \bmod 26 \Rightarrow \text{AXW (ciphertext)}$$

- In order to decrypt we need the inverse of key matrix M, which is

$$N = \begin{pmatrix} 22 & 5 & 1 \\ 6 & 17 & 24 \\ 15 & 13 & 1 \end{pmatrix}$$

Hill Cipher

- If we change one letter in the plaintext, all the letters of the ciphertext will be affected.
- Let the plaintext be BBC instead of ABC then the ciphertext

$$(1,1,2)\times\begin{pmatrix}1 & 2 & 3\\ 4 & 5 & 6\\ 11 & 9 & 8\end{pmatrix}\equiv(1,25,25)\bmod 26\Rightarrow \mathrm{BZZ\,(ciphertext)}$$

- Claude Shannon, in *Communication theory of secrecy systems* Bell Systems Technical Journal 28, (1949), 656-715, introduced properties that a good cryptosystems should have:
- **Diffusion:** one character change in the plaintext should effect as many ciphertext characters as possible, and v.v.
- **Confusion:** The key/plaintext should not relate to the ciphertext in a simple way.

Number Theory (Refreshments)

- **Modulo Operation:**
- **Question:** What is 12 mod 9?
- **Answer:** $12 \bmod 9 \equiv 3$ or $12 \equiv 3 \bmod 9$

- **Definition:** Let $a, r, m \in Z$ (where Z is a set of all integers) and $m > 0$. We write

- $a \equiv r \bmod m$ if m divides $r - a$.

- m is called the *modulus and* r is called the *remainder*

- $a = q \cdot m + r$ $\qquad\qquad$ $0 \leq r < m$

Number Theory (Cont.)

- **Example:** $a = 42$ and $m=9$
- $42 = 4 \cdot 9 + 6$ therefore $42 \equiv 6 \bmod 9$

- **Ring:**
- **Definition:** The ring Z_m consists of
- The set $Z_m = \{0, 1, 2, \ldots, m\text{-}1\}$
- Two operations "+" and "×" for all $a, b \in Z_m$ such that
  - $a + b \equiv c \bmod m \; (c \in Zm)$
  - $a \times b \equiv d \bmod m \; (d \in Zm)$
- **Example:** $m = 9$ $\quad Z_9 = \{0, 1, 2, 3, 4, 5, 6, 7, 8\}$

  $6 + 8 = 14 \equiv 5 \bmod 9$

  $6 \times 8 = 48 \equiv 3 \bmod 9$

Properties of the ring $Z_m = \{0, 1, \ldots, m-1\}$

- The additive identity "0": $a + 0 = a$
- The additive inverse of a: $-a = m - a$ s.t. $a + (-a) \equiv 0 \bmod m$
- Addition is closed i.e if $a, b \in Z_m$ then $a + b \in Z_m$
- Addition is commutative $a + b = b + a$
- Addition is associative $(a + b) + c = a + (b + c)$
- Multiplicative identity "1": $a \times 1 \equiv a \bmod m$
- The multiplicative inverse of a exists if $\gcd(a, m) = 1$ and denoted as a^{-1} s.t. $a^{-1} \times a \equiv 1 \bmod m$
- Multiplication is closed i.e if $a, b \in Z_m$ then $a \times b \in Z_m$
- Multiplication is commutative $a \times b = b \times a$
- Multiplication is associative $(a \times b) \times c = a \times (b \times c)$

Some Remarks on Z_m

- Roughly speaking a ring is a mathematical structure in which we can add, subtract, multiply, and even sometimes divide.
  - **Example:** Is the division 4/15 mod 26 possible?
  - In fact, $4/15 \bmod 26 = 4 \times 15^{-1} \bmod 26$
  - Does $15^{-1} \bmod 26$ exist ?
  - It exists only if gcd(15, 26) = 1.
  - $15^{-1} \bmod 26 = 7$
  - therefore, $4/15 \bmod 26 = 4 \times 7 \bmod 26 = 28 \equiv 2 \bmod 26$

- The modulo operation can be applied whenever we want
  $(a + b) \bmod m = [(a \bmod m) + (b \bmod m)] \bmod m$
  $(a \times b) \bmod m = [(a \bmod m) \times (b \bmod m)] \bmod m$