Principles of Cryptography: Part 1

Gaurav S. Kasbekar
Dept. of Electrical Engineering
IIT Bombay

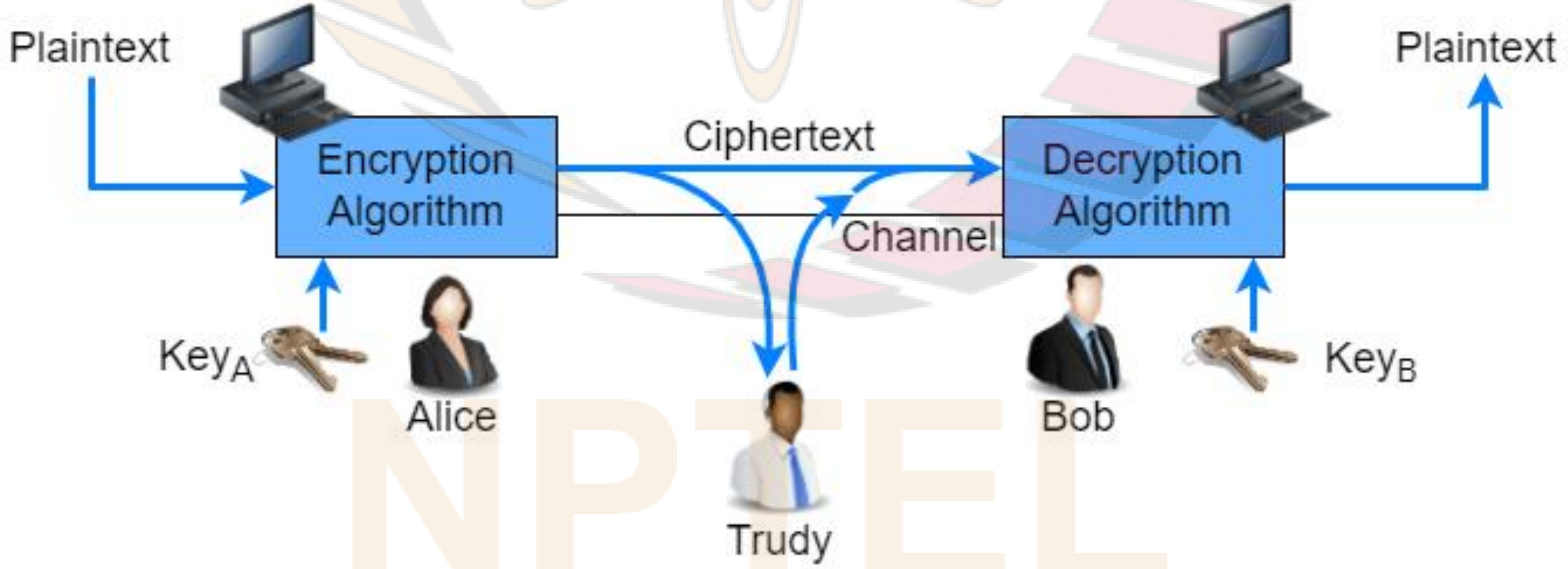
References

- J. Kurose, K. Ross, "*Computer Networking: A Top Down Approach*", Sixth Edition, Pearson Education, 2013
- C. Kaufman, R. Perlman, M. Speciner, "*Network Security: Private Communication in a Public World*", Pearson Education, 2nd edition, 2002

NPTEL

Cryptography

- Cryptography allows a sender to replace the original text (called **plaintext**) with its disguised (encrypted) version (called **ciphertext**)
- Intruder who intercepts message (*e.g.*, using a sniffer) gains no information from it
- Has been used since ancient times by spies, military personnel, diplomats, etc.
- For now, we focus on use of cryptography for achieving confidentiality
    - ❑ later we will see that it is also useful for achieving other functions (*e.g.*, authentication, message integrity)

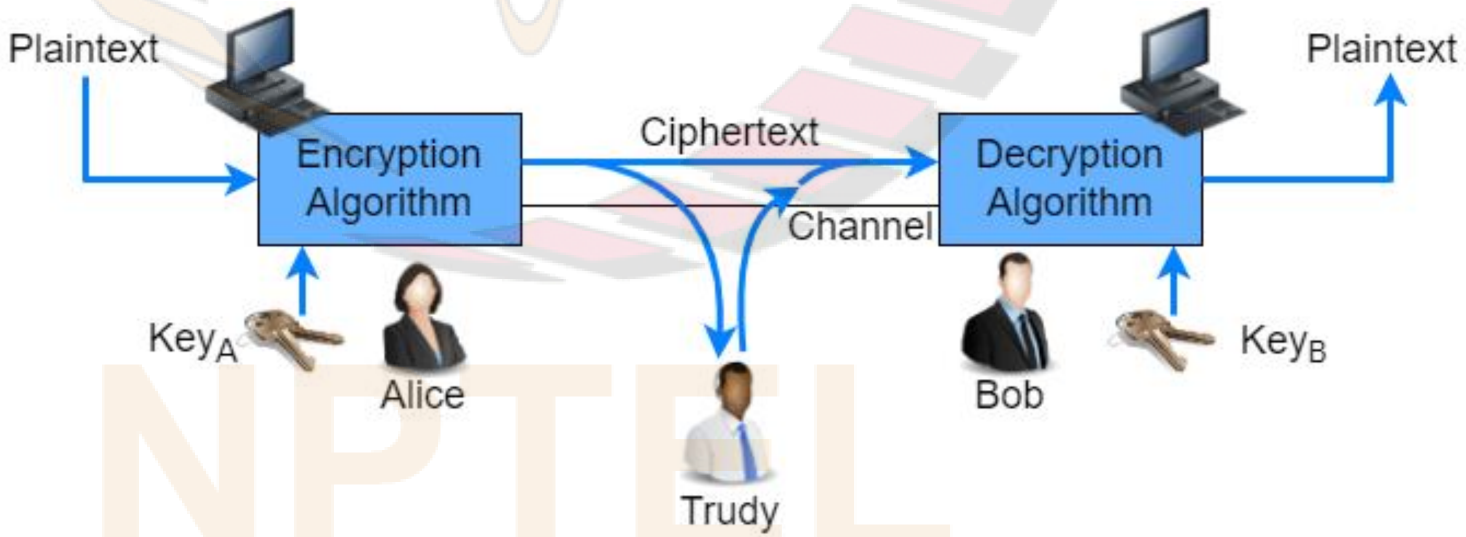


Example: Caeser Cipher

- Was used by Julius Caeser to encrypt his messages
- Each letter in plaintext replaced with the letter that is k letters later in the alphabet (with wraparound)
- E.g., if $k = 3$, then we replace:
  - a with d, b with e, …, z with c
- E.g.: suppose $k = 23$ (equivalent to left shift of 3)
  - Plaintext: THE QUICK BROWN FOX JUMPS OVER THE LAZY DOG
  - Ciphertext: QEB NRFZH YOLTK CLU GRJMP LSBO QEB IXWV ALD
- Suppose interceptor knew that Caeser cipher used, but value of k unknown. Then to break the cipher:
  - needs to only try out all 25 possible values
- Caeser cipher highly insecure

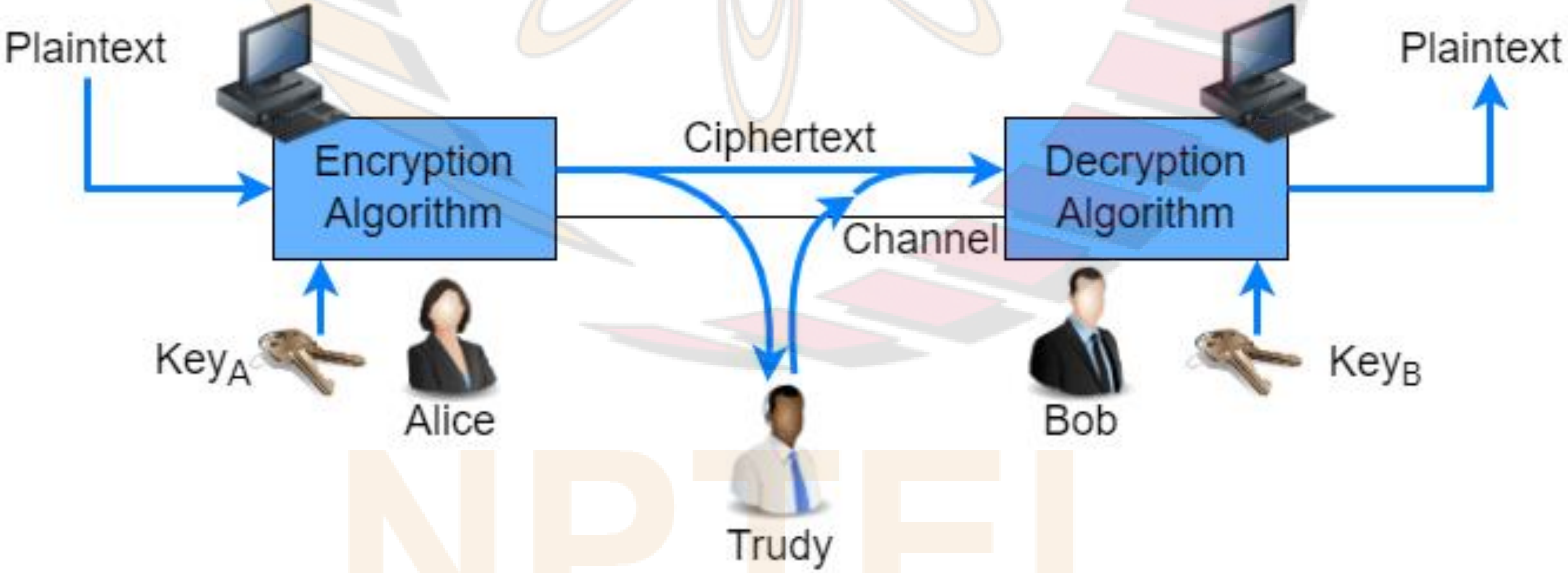
Encryption and Decryption Algorithms, Keys

- Algorithm used to transform:
  - ❑ plaintext into ciphertext called *encryption algorithm*
  - ❑ ciphertext into plaintext called *decryption algorithm*
- In many cryptographic systems, including those used in Internet, encryption and decryption algorithms are standardized and published
  - ❑ available to everyone including potential intruders
- Clearly there must be some secret information that prevents intruder from decrypting message
  - ❑ "**key**" (a string of letters or numbers or bits)
  - ❑ *e.g.*, key in Caeser cipher:
    - o shift value k



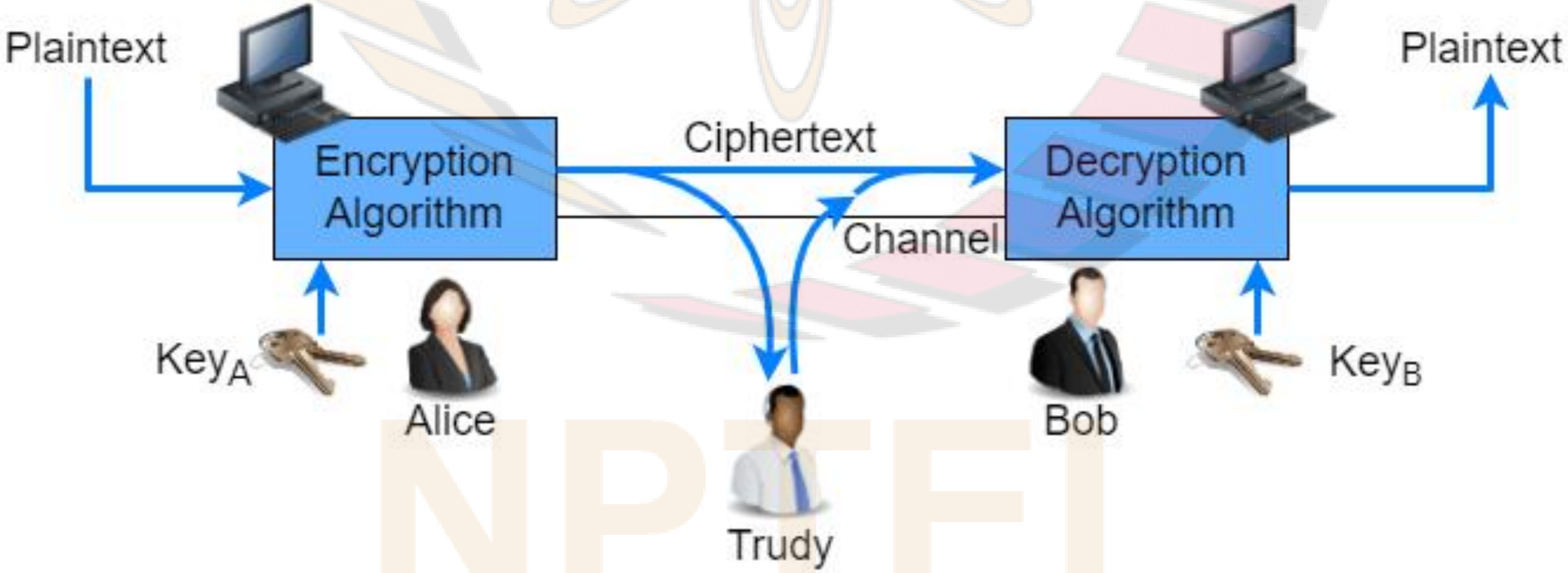
Notation

- Suppose Alice sends an encrypted message to Bob
- Let K_A and K_B denote keys of Alice and Bob, resp.
- If m is plaintext message, then let:
  - ❑ $K_A(m)$: ciphertext obtained by encrypting m with K_A
- $K_B(.)$ similarly defined; $K_B(K_A(m)) = m$



Types of Cryptography

- **Symmetric Key Cryptography**
  - ❑ Keys of sender and receiver are same, *i.e.*, $K_A = K_B$
  - ❑ Key known to only sender and receiver
  - ❑ *E.g.*: Caeser cipher
- **Public Key Cryptography**
  - ❑ Keys of sender and receiver are different, *i.e.*, $K_A \neq K_B$
  - ❑ Sender's key (called *public key*) is available to everyone, *e.g.*, may be included in a database similar to telephone directory
  - ❑ Receiver's key (called *private key*) known to only receiver



NPTEL

Monoalphabetic Cipher

- Recall: Caeser cipher, which is symmetric key based, easy to break
- *Monoalphabetic Cipher*: generalization of Caeser cipher
  - replaces each letter in plaintext with another letter (*e.g.* in figure below)
  - the substitute for a given letter is fixed throughout the message, e.g., each time a occurs in message, it is replaced with m
- E.g.:
  - *Plaintext*: attack at noon
  - *Ciphertext*: muumbf mu jkkj
- Key in monoalphabetic cipher:
  - The string of ciphertext letters for the 26 alphabets
- No. of possible keys:
  - $26! \approx 4 \times 10^{26}$
- Brute force approach for breaking cipher:
  - try out all possible keys
  - time consuming
- However, possible to break cipher efficiently by *frequency analysis*

Plaintext Letter	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
Ciphertext Letter	m	n	b	v	c	x	z	a	s	d	f	g	h	j	k	l	p	o	i	u	y	t	r	e	w	q

Frequency Analysis

- Table shows frequencies of different letters in typical English text
- So if message is sufficiently long, can measure the frequencies of different letters in ciphertext
- Used to form guesses
  - *e.g.*, if P is most frequent letter in ciphertext, then it is probably the substitute for e, t or a
- Also, several two or three letter sequences of letters (e.g., “in”, “it”, “the”, “ion”, “ing”) appear frequently; this fact used
- After some replacements of ciphertext letters with plaintext guesses made, more guesses can often be made
  - *e.g.*, suppose after substitutions for “e” and “t”, the pattern “tKe” appears frequently; then “K” is probably the substitute for “h”
- Above process repeated
- See following website for an example:
  - http://crypto.interactive-maths.com/frequency-analysis-breaking-the-code.html

Letter	Frequency
e	12.7
t	9.1
a	8.2
o	7.5
i	7.0
n	6.7
s	6.3
h	6.1
r	6.0
d	4.3
l	4.0
c	2.8
u	2.8
m	2.4
w	2.4
f	2.2
g	2.0
y	2.0
p	1.9
b	1.5
v	1.0
k	0.8
j	0.15
x	0.15
q	0.10
z	0.07

NPTEL

Polyalphabetic Cipher

- Multiple monoalphabetic ciphers used
- Each letter in plaintext replaced with corresponding letter from one of these ciphers
  - which cipher used depends on position of letter
- **E.g.**:
  - Two monoalphabetic ciphers used: a Caeser cipher with $k = 5$ and one with $k = 19$ (see fig.)
  - These two ciphers used in the repeating pattern:
    - C_1, C_2, C_2, C_1, C_2
  - *Plaintext*: attack at noon
  - *Ciphertext*: fmmfvp tm shtg
- Polyalphabetic cipher cannot be broken using frequency analysis; however, several polyalphabetic ciphers (e.g., Vigenère Cipher) have been broken using other techniques

Plaintext Letter	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
C_1 (k=5)	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
C_2 (k=19)	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