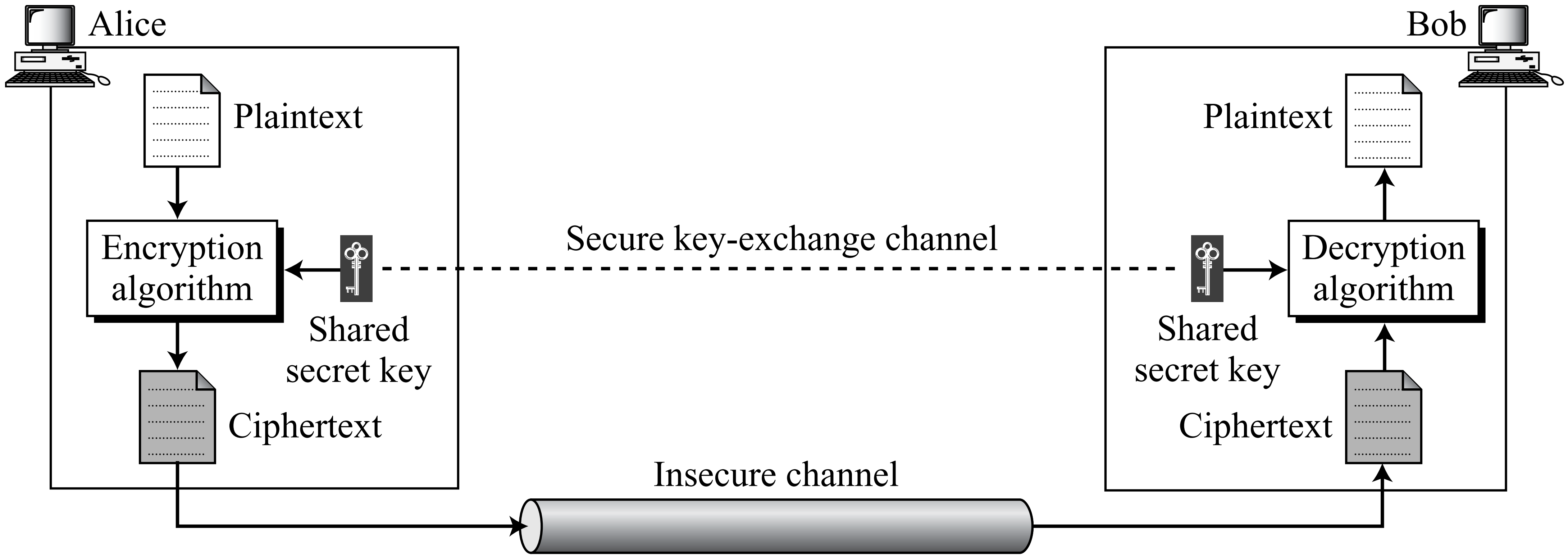
Symmetric Key Ciphers

Part II: Polyalphabetic Substitution Ciphers

THE UNIVERSITY OF UTAH

Priyank Kalla
Professor
Electrical & Computer Engineering



- MonoAlphabetic Ciphers (MAC): A character in the plaintext is always changed to the same character in the cipher text, regardless of its position in the text.
  - Additive ciphers, multiplicative ciphers and affine ciphers fall into this category
  - MAC don’t change the frequency of characters: so they are vulnerable to statistical attacks

Statistical Attacks

Frequency of occurrence of letters in an English text

Letter	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*	*Letter*	*Frequency*
E	12.7	H	6.1	W	2.3	K	0.08
T	9.1	R	6.0	F	2.2	J	0.02
A	8.2	D	4.3	G	2.0	Q	0.01
O	7.5	L	4.0	Y	2.0	X	0.01
I	7.0	C	2.8	P	1.9	Z	0.01
N	6.7	U	2.8	B	1.5		
S	6.3	M	2.4	V	1.0		

Intercepted Cipher:

XLILSYWIMWRSAJSVWEPIJSVJSYVQMPPMSRHSPPEVWMXMWASVX-LQSVILY-VVCFIJSVIXLIWIPPIVVIGIMZIWQSVISJJIVW

- The letter “I” appears 14 times: maximum occurrence of a character
- Probably: “I” in cipher text = “e” in plaintext
  - Key = 4: $C = P + k \pmod{26} = e + 4 = 4 + 4 = 8 = I$

Decipher:

the house is now for sale for four million dollars it is worth more hurry before the seller receives more offers

Monoalphabetic Substitution Cipher

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

- Key = Look-up Table
  - Key space = $26! = 4 \times 10^{26}$ (approx)
  - Brute-force attack is difficult, but statistical attacks are successful
- Additive, multiplicative and affine ciphers: Brute-force attacks can work, because the frequency of characters in the cipher text does not change
- Solution: Polyalphabetic Substitution Ciphers

PolyAlphabetic Substitution Ciphers

- Each occurrence of a character may have a different substitute in the cipher text
- Make each ciphertext character dependent on: 1) the corresponding plaintext character; and/or 2) the position of the plaintext character in the message
- Examples: Autokey Cipher, Vigenere Cipher, Hill Cipher
- Autokey Cipher:
  - Key = stream of subkeys, 1 subkey for 1 character
  - Only the first subkey is shared between Alice and Bob
  - Following subkeys = plaintext characters $P_1P_2\ldots$

Autokey Cipher:

$P = P_1P_2P_3 \ldots$ $C = C_1C_2C_3\ldots$ $k = (k_1, P_1, P_2, \ldots)$

Encryption: $C_i = (P_i + k_i) \bmod 26$ Decryption: $P_i = (C_i - k_i) \bmod 26$

Plaintext:	a	t	t	a	c	k	i	s	t	o	d	a	y
P's Values:	00	19	19	00	02	10	08	18	19	14	03	00	24
Key stream:	*12*	*00*	*19*	*19*	*00*	*02*	*10*	*08*	*18*	*19*	*14*	*03*	*00*
C's Values:	12	19	12	19	02	12	18	00	11	7	17	03	24
Ciphertext:	**M**	**T**	**M**	**T**	**C**	**M**	**S**	**A**	**L**	**H**	**R**	**D**	**Y**

- Autokey: Subkeys automatically created from P
- Hides single-letter frequency
- Easy to break: Brute-force attack on k_1

Vigenère Cipher

$P = P_1P_2P_3 \ldots$ $C = C_1C_2C_3 \ldots$ $K = [(k_1, k_2, \ldots, k_m), (k_1, k_2, \ldots, k_m), \ldots]$

Encryption: $C_i = P_i + k_i$ Decryption: $P_i = C_i - k_i$

- Key is a stream: $(k_1, k_2, \ldots, k_m),\ 1 \leq m \leq 26$, which repeats
- Example key stream = "PASCAL" = (15, 00, 18, 2, 0, 11)
- Key depends on the position of the the character in the Ciphertext, not on the plaintext character itself

Plaintext:	s	h	e	i	s	l	i	s	t	e	n	i	n	g
P's values:	18	07	04	08	18	11	08	18	19	04	13	08	13	06
Key stream:	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***	***18***	***02***	***00***	***11***	***15***	***00***
C's values:	**07**	**07**	**22**	**10**	**18**	**22**	**23**	**18**	**11**	**6**	**13**	**19**	**02**	**06**
Ciphertext:	**H**	**H**	**W**	**K**	**S**	**W**	**X**	**S**	**L**	**G**	**N**	**T**	**C**	**G**

Vigenère Cipher with $m = 1$ = additive cipher

Hill Cipher

- Divide plaintext into equal-size ***blocks (block cipher)***
  - Sometimes need to add bogus characters
- Blocks are encrypted 1 at a time
- Each Character in the block contributes to the encryption of other characters in the block
- Key = square $m \times m$ matrix

$$\mathbf{K} = \begin{bmatrix} k_{11} & k_{12} & \ldots & k_{1m} \\ k_{21} & k_{22} & \ldots & k_{2m} \\ \vdots & \vdots & & \vdots \\ k_{m1} & k_{m2} & \ldots & k_{mm} \end{bmatrix}$$

Hill Cipher

$$C_1 = P_1 k_{11} + P_2 k_{21} + \cdots + P_m k_{m1}$$
$$C_2 = P_1 k_{12} + P_2 k_{22} + \cdots + P_m k_{m2}$$
$$\ldots$$
$$C_m = P_1 k_{1m} + P_2 k_{2m} + \cdots + P_m k_{mm}$$

- Represent C, P, K as matrices: $C = P \cdot K$ and $P = C \cdot K^{-1}$
- P = “code is ready” (11 chars) = “code is ready**z**”

$$\overset{\mathbf{C}}{\begin{bmatrix} 14 & 07 & 10 & 13 \\ 08 & 07 & 06 & 11 \\ 11 & 08 & 18 & 18 \end{bmatrix}} = \overset{\mathbf{P}}{\begin{bmatrix} 02 & 14 & 03 & 04 \\ 08 & 18 & 17 & 04 \\ 00 & 03 & 24 & 25 \end{bmatrix}} \overset{\mathbf{K}}{\begin{bmatrix} 09 & 07 & 11 & 13 \\ 04 & 07 & 05 & 06 \\ 02 & 21 & 14 & 09 \\ 03 & 23 & 21 & 08 \end{bmatrix}}$$

a. Encryption

$$\pmod{26}$$

$$\overset{\mathbf{P}}{\begin{bmatrix} 02 & 14 & 03 & 04 \\ 08 & 18 & 17 & 04 \\ 00 & 03 & 24 & 25 \end{bmatrix}} = \overset{\mathbf{C}}{\begin{bmatrix} 14 & 07 & 10 & 13 \\ 08 & 07 & 06 & 11 \\ 11 & 08 & 18 & 18 \end{bmatrix}} \overset{\mathbf{K}^{-1}}{\begin{bmatrix} 02 & 15 & 22 & 03 \\ 15 & 00 & 19 & 03 \\ 09 & 09 & 03 & 11 \\ 17 & 00 & 04 & 07 \end{bmatrix}}$$

b. Decryption

Hill Cipher: Cryptanalysis

$$
\begin{aligned}
&\mathrm{C}_1 = \mathrm{P}_1\, k_{11} + \mathrm{P}_2\, k_{21} + \cdots + \ \mathrm{P}_m\, k_{m1} \\
&\mathrm{C}_2 = \mathrm{P}_1\, k_{12} + \mathrm{P}_2\, k_{22} + \cdots + \ \mathrm{P}_m\, k_{m2} \\
&\cdots \\
&\mathrm{C}_\mathrm{m} = \mathrm{P}_1\, k_{1m} + \mathrm{P}_2\, k_{2m} + \cdots + \ \mathrm{P}_m\, k_{mm}
\end{aligned}
$$

- Cryptanalysis is hard: Key space $26^{m \times m}$, also need to know m
- Frequency analysis of characters also not helpful as this is a polyalphabetic cipher
- Try "known-plaintext" attack if you know m and generate (P,C) pairs for m blocks
- $C = P \cdot K \implies K = C \cdot P^{-1}$
- If P is non-invertible, generate more (P,C) pairs

Attack on Hill Cipher

- Assume Eve knows $m = 3$
- Intercept 3 (P,C) pairs, to create $m \times m$ matrices for C and P

$$\begin{bmatrix} 05 & 07 & 10 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 03 & 06 & 00 \end{bmatrix}$$

$$\begin{bmatrix} 13 & 17 & 07 \end{bmatrix} \longleftrightarrow \begin{bmatrix} 14 & 16 & 09 \end{bmatrix}$$

$$\underset{\mathbf{P}}{\begin{bmatrix} 00 & 05 & 04 \end{bmatrix}} \longleftrightarrow \underset{\mathbf{C}}{\begin{bmatrix} 03 & 17 & 11 \end{bmatrix}}$$

$$\underset{\mathbf{K}}{\begin{bmatrix} 02 & 03 & 07 \\ 05 & 07 & 09 \\ 01 & 02 & 11 \end{bmatrix}} = \underset{\mathbf{P}^{-1}}{\begin{bmatrix} 21 & 14 & 01 \\ 00 & 08 & 25 \\ 13 & 03 & 08 \end{bmatrix}} \underset{\mathbf{C}}{\begin{bmatrix} 03 & 06 & 00 \\ 14 & 16 & 09 \\ 03 & 17 & 11 \end{bmatrix}}$$

Next class: Transposition ciphers and modern block ciphers