

Algoritma Kriptografi Klasik (bagian 2)

IF5054 Kriptografi

Rinaldi Munir/IF5054 Kriptografi

1

Jenis-jenis *Cipher Substitusi*

1. **Cipher abjad-tunggal (monoalphabetic cipher)**
Satu huruf di plainteks diganti dengan satu huruf yang bersesuaian.

Jumlah kemungkinan susunan huruf-huruf cipherteks yang dapat dibuat adalah sebanyak

$$26! = 403.291.461.126.605.635.584.000.000$$

Contoh: *Caesar Cipher*

Rinaldi Munir/IF5054 Kriptografi

2

- Tabel substitusi dapat dibentuk secara acak:

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

- Atau dengan kalimat yang mudah diingat:

Contoh: we hope you enjoy this book

Buang duplikasi huruf: wehopyunj t isbk

Sambung dengan huruf lain yang belum ada:

wehopyunj t isbkacdfglmqrvxz

Tabel substitusi:

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

2. Cipher substitusi homofonik

(*Homophonic substitution cipher*)

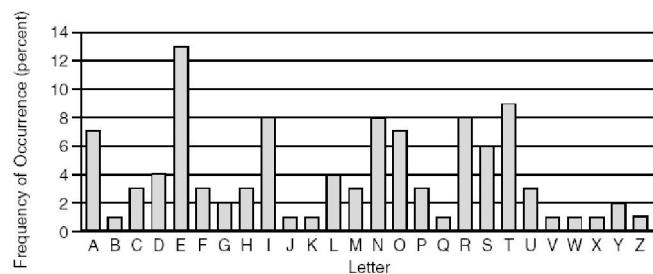
Setiap huruf plainteks dipetakan ke dalam salah satu huruf cipherteks yang mungkin.

Tujuan: menyembunyikan hubungan statistik antara plainteks dengan cipherteks

Fungsi *ciphering* memetakan satu-ke-banyak (*one-to-many*).

Misal: huruf E → **AB**, **TQ**, **YT**, **UX** (homofon)
 huruf B → **EK**, **MF**, **KY** (homofon)

- Contoh: Sebuah teks dengan frekuensi kemunculan huruf sbb:



- Huruf E muncul 13 % → dikodekan dengan 13 huruf homofon

Huruf													
Plainteks			Pilihan untuk unit cipherteks										
A	BU	CP	AV	AH	BT	BS	CQ						
B	AT												
C	DL	BK	AU										
D	BV	DY	DM	AI									
E	DK	CO	AW	BL	AA	CR	BM	CS	AF	AG	BO	BN	BE
F	BW	CM	CN										
G	DN	BJ											
H	AS	CL	CK										
I	DJ	BI	AX	CJ	AB	BP	CU	CT					
J	BX												
K	DI												
L	AR	BH	CI	AJ									
M	DH	BG	AY										
N	BY	DG	DF	CH	AC	BR	DU	DT					
O	DZ	BF	DX	AK	CG	BQ	DR						
P	BZ	DE	AZ										
Q	DD												
R	AQ	DC	DQ	AL	CE	CF	CV	DS					
S	AP	AN	AO	CD	DW	DV							
T	CB	DB	DP	CC	AD	CY	CW	CX	AE				
U	CA	AM	BA										
V	BB												
W	CZ												
X	BD												
Y	DO	DA											
Z	BC												

- Unit cipherteks mana yang dipilih diantara semua homofon ditentukan secara acak.
- Contoh:

Plainteks: KRIPTO

Cipherteks: DI CE AX AZ CC DX
- Enkripsi: satu-ke-banyak
- Dekripsi: satu-ke-satu
- Dekripsi menggunakan tabel homofon yang sama.

3. ***Cipher abjad-majemuk*** (*Polyalpabetic substitution cipher*)
- *Cipher* abjad-tunggal: satu kunci untuk semua huruf palinteks
 - *Cipher* substitusi-ganda: setiap huruf menggunakan kunci berbeda.
 - *Cipher* abjad-majemuk dibuat dari sejumlah *cipher* abjad-tunggal, masing-masing dengan kunci yang berbeda.
 - Kebanyakan *cipher* abjad-majemuk adalah *cipher* substitusi periodik yang didasarkan pada periode m .

- Plainteks:

$$P = p_1 p_2 \dots p_m p_{m+1} \dots p_{2m} \dots$$

- Cipherteks:

$$E_k(P) = f_1(p_1) f_2(p_2) \dots f_m(p_m) f_{m+1}(p_{m+1}) \dots f_{2m}(p_{2m}) \dots$$

- Untuk $m = 1$, *cipher*-nya ekivalen dengan *cipher* abjad-tunggal.

- Contoh *cipher* substitusi periodik adalah *cipher Vigenere*
Kunci: $K = k_1 k_2 \dots k_m$
 k_i untuk $1 \leq i \leq m$ menyatakan jumlah pergeseran pada huruf ke- i .

Karakter cipherteks: $c_i(p) = (p + k_i) \bmod 26$ (*)

Misalkan periode $m = 20$, maka 20 karakter pertama dienkripsi dengan persamaan (*), setiap karakter ke- i menggunakan kunci k_i .

Untuk 20 karakter berikutnya, kembali menggunakan pola enkripsi yang sama.

Contoh: (spasi dibuang)

P : KRIPTOGRAFI KLASIK DENGAN CIPHER ALFABET MAJEMUK

K : LAMPION LAMPION LAMPION LAMPION LAMPION LAMPION LAMPION LAMPION

C : VR...

Perhitungan:

$$(K + L) \bmod 26 = (10 + 11) \bmod 26 = 21 = V$$

$$(R + A) \bmod 26 = (17 + 0) \bmod 26 = 17 = A$$

dst

Contoh 2: (dengan spasi)

P: SHE SELLS SEA SHELLS BY THE SEASHORE

K: KEY KEYKE YKE YKEYKE YK EYK EYKEYKEY

C: CLC CIJVW QOE QRIJVW ZI XFO WCKWFYVC

4. Cipher substitusi poligram (*Polygram substitution cipher*)

- Blok huruf plainteks disubstitusi dengan blok cipherteks.
- Misalnya AS diganti dengan **RT**, BY diganti dengan **SL**
- Jika unit huruf plainteks/cipherteks panjangnya 2 huruf, maka ia disebut digram (*biigram*), jika 3 huruf disebut ternari-gram, dst
- Tujuannya: distribusi kemunculan poligram menjadi *flat* (datar), dan hal ini menyulitkan analisis frekuensi

Cipher Transposisi

- Ciphereteks diperoleh dengan mengubah posisi huruf di dalam plaintekls.
- Dengan kata lain, algoritma ini melakukan *transpose* terhadap rangkaian huruf di dalam plainteks.
- Nama lain untuk metode ini adalah **permutasi**, karena *transpose* setiap karakter di dalam teks sama dengan mempermutasikan karakter-karakter tersebut.

Contoh: Misalkan plainteks adalah

DEPARTEMEN TEKNIK INFORMATIKA ITB

Enkripsi:

DEPART
EMENTE
KNIKIN
FORMAT
IKAITB

Cipherteks: (baca secara vertikal)

DEKFIEMNOKPEIRAANKMIRTIATTENTB

DEKF IEMN OKPE IRAA NKMI RTIA TTEN TB

Dekripsi: Bagi panjang cipherteks dengan kunci.
 (Pada contoh ini, $30 / 6 = 5$)

DEKFI

EMNOK

PEIRA

ANKMI

RTIAT

TENTB

Plainteks: (baca secara vertikal)

DEPARTEMEN TEKNIK INFORMATIKA ITB

- Contoh lain: Plainteks: ITB GANESHA SEPULUH
- Bagi menjadi blok-blok 8-huruf. Jika < 8, tambahkan huruf palsu.

1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8	1	2	3	4	5	6	7	8
I	T	B	G	A	N	E	S	H	A	S	E	P	U	L	U	H	A	B	C	D	E	F	G
S	T	B	A	G	N	E	I	U	A	S	P	E	U	L	H	G	A	B	D	C	E	F	H

- Cipherteks: **STBAGNEIUASPEULHGABDCEFH**

Contoh lain. Misalkan plainteks adalah

CRYPTOGRAPHY AND DATA SECURITY

Plainteks disusun menjadi 3 baris ($k = 3$) seperti di bawah ini:

C	T	A	A	A	E	I
R	P	O	R	P	N	T
Y	G	H	D	A	S	C

maka cipherteksnya adalah

CTAAAEIRPORPYNDTSCRTYGHDAUY

Super-enkripsi

- Menggabungkan *cipher* substitusi dengan *cipher* transposisi.
- **Contoh.** Plainteks HELLO WORLD
- dienkripsi dengan *caesar cipher* menjadi KHOOR ZRUOG kemudian hasil enkripsi ini dienkripsi lagi dengan *cipher* transposisi ($k = 4$):
 - KHOO
 - RZRU
 - OGZZ

Cipherteks akhir adalah: KROHZGORZOUZ

Lebih Jauh dengan *Cipher Abjad-tunggal*

- Jumlah kemungkinan kunci = 26!
- Tidak dapat menyembunyikan hubungan antara plainteks dengan cipherteks.
- Huruf yang sama dienkripsi menjadi huruf cipherteks yang sama
- Huruf yang sering muncul di dalam palinteks, sering muncul pula di dalam cipherteksnya.

- Oleh karena itu, cipherteks dapat didekripsi tanpa mengetahui kunci (*ciphertext-only attack*)
- Metode yang digunakan:
 1. Terkaan
 2. Statistik (analisis frekuensi)
- Informasi yang dibutuhkan:
 1. Mengetahui bahasa yang digunakan untuk plainteks
 2. Konteks plainteks

Metode Terkaan

Asumsi: - bahasa plainteks adalah B. Inggris
- spasi tidak dibuang

Tujuan: mereduksi jumlah kunci

Contoh 1. Cipherteks: **G WR W RWL**

Plainteks: I AM A MA*
I AM A MAN

Jumlah kunci berkurang dari 26! menjadi 22!

Contoh 2.

Cipherteks: **HKC**

Plainteks:

- lebih sukar ditentukan,
- tetapi tidak mungkin
 - Z diganti dengan **H**,
 - Q dengan **K**,
 - K dengan **C**,

karena tidak ada kata “ZQC” dalam Bahasa Inggris

Contoh 3.

Cipherteks: **HATTPT**

Plainteks: salah satu dari **T** atau **P**
merepresentasikan huruf vokal, misal

CHEESE

MISSES

CANNON

Contoh 4.

Cipherteks: **HATTPT**

Plainteks: diketahui informasi bahwa pesan tersebut adalah nama negara.

→ GREECE

- Proses menerka dapat menjadi lebih sulit jika cipherteks dikelompokkan ke dalam blok-blok huruf.
- Contoh:

CTBMN BYCTC BTJDS QXBNS GSTJC BTSWX CTQTZ CQVUJ QJSGS
TJQZZ MNQJS VLNSX VSZJU JDSTS JQUUS JUBXJ DSKSU JSNTK
BGAQJ ZBGYQ TLCTZ BNYBN QJSW

- Jika diberikan informasi bahwa cipherteks tersebut berasal dari perusahaan yang bergerak di bidang keuangan, maka proses menerka dapat lebih mudah
- Kata keuangan dalam Bahasa Inggris adalah FINANCIAL

- Ada dua buah huruf I yang berulang, dengan empat buah huruf lain di antara keduanya (NANC)
- Cari huruf berulang dengan pola seperti itu di dalam cipherteks (tidak termasuk spasi). Ditemukan pada posisi 6, 15, 27, 31, 42, 48, 58, 66, 70, 71, 76, dan 82
- Hanya dua diantaranya, yaitu 31 dan 42 yang mempunyai huruf berikutnya yang berulang (berkoresponden dengan N)
- Dan dari keduanya hanya pada posisi 31 huruf A berada pada posisi yang tepat
- Jadi ditemukan FINANCIAL pada posisi 30, yaitu untuk kriptogram XCTQTZCQV

- Diperoleh pemetaan:

X	→	F		C	→	I
T	→	N		Q	→	A
Z	→	C		V	→	L

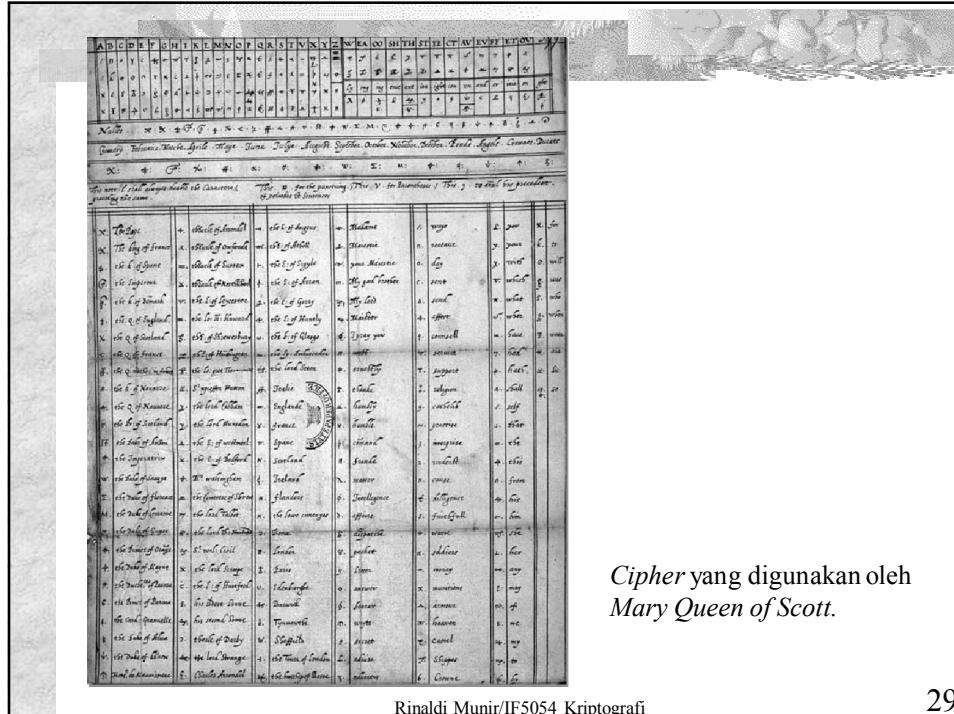
- Ganti semua huruf X, C, T, Q, Z, V dengan F, I, N, A, C, L:

CTBMN BYCTC BTJDS QXBNS GSTJC BTSWX CTQTZ CQVUJ
 QJSGS TJQZZ MNQJS VLNSX VSZJU JDSTS JQUUS JUBXJ
 DSKSU JSNTK BGAQJ ZBGYQ TLCTZ BNYBN QJSW

inBMN BYini BnJDS cfBNS GSnJi BnSWf inanc ialUJ
 aJSGS nJacc MNaJS VLNSf VScJU JDSnS JaUUS JUBfJ
 DSKSU JSNnK BGAAJ cBGYa nLinc BNYBN aJSW

- Jumlah kunci berkurang menjadi 20! Deduksi dapat diteruskan.

- Peristiwa yang menimpa Queen Mary of Scotland pada abad 18 karena menggunakan *cipher* abjad-tunggal yang mudah diterka → mudah dipecahkan.



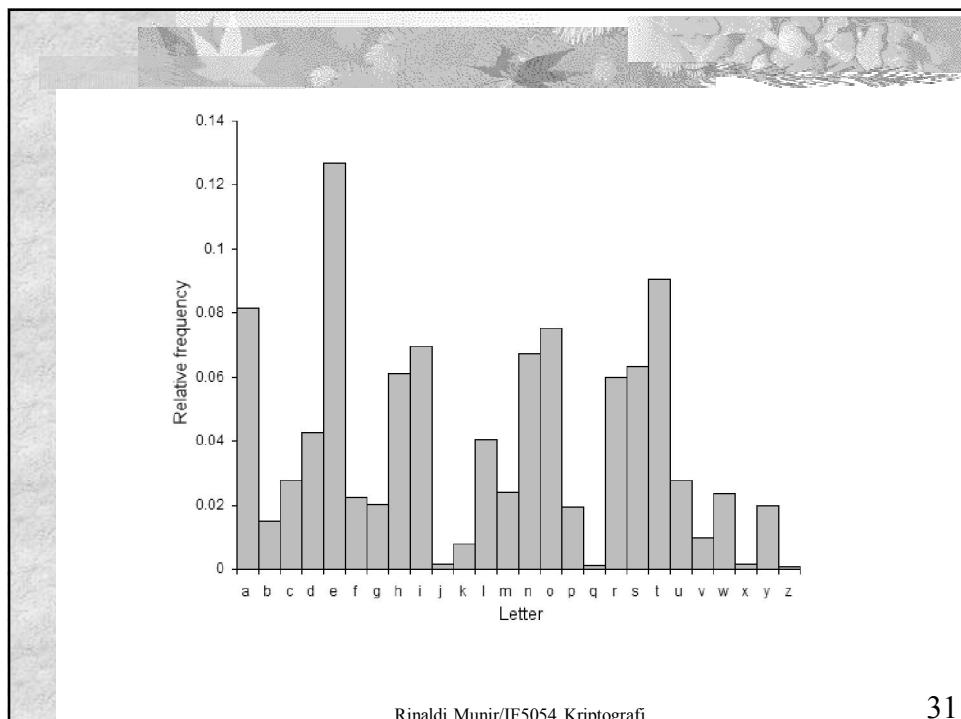
Cipher yang digunakan oleh Mary Queen of Scott.

Rinaldi Munir/IF5054 Kriptografi

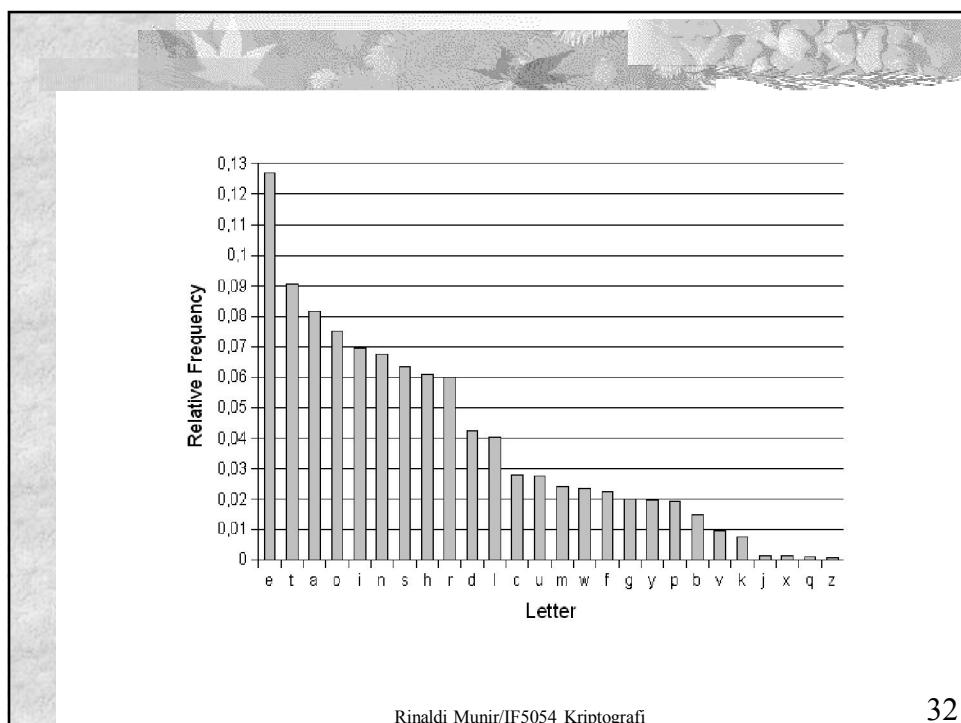
Metode Analisis Frekuensi

Tabel 2. Frekuensi kemunculan (relatif) huruf-huruf dalam teks Bahasa Inggris (sampel mencapai 300.000 karakter di dalam sejumlah novel dan suratkabar)

Huruf	%	Huruf	%
A	8,2	N	6,7
B	1,5	O	7,5
C	2,8	P	1,9
D	4,2	Q	0,1
E	12,7	R	6,0
F	2,2	S	6,3
G	2,0	T	9,0
H	6,1	U	2,8
I	7,0	V	1,0
J	0,1	W	2,4
K	0,8	X	2,0
L	4,0	Y	0,1
M	2,4	Z	0,1



31



32

- *Top 10* huruf yang sering muncul dalam teks Bahasa Inggris: E, T, A, O, I, N, S, H, R, D, L, U
- Top 10 huruf *bigram* yang sering muncul dalam teks B. Inggris: TH, HE, IN, EN, NT, RE, ER, AN, TI, dan ES
- Top 10 huruf *trigram* yang sering muncul dalam teks B. Inggris: THE, AND, THA, ENT, ING, ION, TIO, FOR, NDE, dan HAS

- Kriptanalisis menggunakan tabel frekuensi kemunculan huruf dalam B. Inggris sebagai kakas bantu melakukan dekripsi.
- Kemunculan huruf-huruf di dalam sembarang plainteks tercermin pada tabel tersebut.
- Misalnya, jika huruf “R” paling sering muncul di dalam cipherteks, maka kemungkinan besar itu adalah huruf “E” di dalam plainteksnya.

Teknik analisis frekuensi dilakukan sebagai berikut:

1. Misalkan plainteks ditulis dalam Bahasa Inggris (plainteks dalam bahasa lain secara prinsip tidak jauh berbeda).
2. Asumsikan plainteks dienkripsi dengan *cipher* alfabet-tunggal.
3. Hitung frekuensi kemunculan relatif huruf-huruf di dalam cipherteks.
4. Bandingkan hasil langkah 3 dengan Tabel 4.3. Catatlah bahwa huruf yang paling sering muncul dalam teks Bahasa Inggris adalah huruf E. Jadi, huruf yang paling sering muncul di dalam cipherteks kemungkinan besar adalah huruf E di dalam plainteksnya.
5. Langkah 4 diulangi untuk huruf dengan frekuensi terbanyak berikutnya.

■ Contoh: Diberikan cipherteks berikut ini:

```
UZ QSO VUOHXMOPV GPOZPEVSG ZWSZ OPFPESX
UDBMETSX AIZ VUEPHZ HMDZSHZO WSFP APPD
TSVP QUZW YMXUZUHSX EPYEPOPDZSZUFPO MB
ZWP FUPZ HMDJ UD TMOHMQ
```

Lakukan kriptanalisis dengan teknik analisis frekuensi untuk memperoleh plainteks. Asumsi: bahasa yang digunakan adalah Bahasa Inggris dan *cipher* yang digunakan adalah *cipher* abjad-tunggal.

- Frekuensi kemunculan huruf didalam cipherteks tersebut:

Huruf	%	Huruf	%
P	13,33	Q	2,50
Z	11,67	T	2,50
S	8,33	A	1,67
U	8,33	B	1,67
O	7,50	G	1,67
M	6,67	Y	1,67
H	5,83	I	0,83
D	5,00	J	0,83
E	5,00	C	0,00
V	4,17	K	0,00
X	4,17	L	0,00
F	3,33	N	0,00
W	3,33	R	0,00

- Huruf yang paling sering muncul di dalam cipherteks: huruf P dan Z.
- Huruf yang paling sering muncul di dalam B. Inggris: huruf E dan T.
- Kemungkinan besar,
 - P adalah pemetaan dari E
 - Z adalah pemetaan dari T
- Tetapi kita belum dapat memastikannya sebab masih diperlukan cara *trial and error* dan pengetahuan tentang Bahasa Inggris.
- Tetapi ini adalah langkah awal yang sudah bagus.

Iterasi 1:

UZ QSO VUOHXMOPV GPOZPEVSG ZWSZ OPFPESX UDBMETSX AIZ
t e e te t t e e t

VUEPHZ HMDZSHZO WSFP APPD TSVP QUZW YMUXUHHSX
e t t t e ee e t t

EPYEPOPDZSZUFPO MB ZWP FUPZ HMDJ UD TMOHMQ
e e e t t e t e et

- **ZWP** dan **ZWSZ** dipetakan menjadi **t***e dan **t****t
- Kemungkinan besar **W** adalah pemetaan dari **H** sehingga kata yang mungkin untuk **ZWP** dan **ZWSZ** adalah **the** dan **that**

- Diperoleh pemetaan:

$$\begin{aligned} P &\rightarrow e \\ Z &\rightarrow t \\ W &\rightarrow h \\ S &\rightarrow a \end{aligned}$$

- Iterasi 2:

UZ QSO VUOHXMOPV GPOZPEVSG ZWSZ OPFPESX UDBMETSX AIZ
t a e e te a that e e a a t

VUEPHZ HMDZSHZO WSFP APPD TSVP QUZW YMUXUHHSX
e t ta t ha e ee a e th t a

EPYEPOPDZSZUFPO MB ZWP FUPZ HMDJ UD TMOHMQ
e e e tat e the et

- WSFP dipetakan menjadi ha*e.
- Dalam Bahasa Inggris, kata yang mungkin untuk ha*e hanyalah have, hate, hale, dan haze
- Dengan mencoba mengganti semua Z di dalam cipherteks dengan v, t, l, dan z, maka huruf yang cocok adalah v sehingga WSFP dipetakan menjadi have
- Dengan mengganti F menjadi v pada kriptogram **EPYEPOPDZSZUFPO** sehingga menjadi *e*e*e*tat*ve*, maka kata yang cocok untuk ini adalah representatives

- Diperoleh pemetaan:
 - $E \rightarrow r$
 - $Y \rightarrow p$
 - $U \rightarrow i$
 - $O \rightarrow s$
 - $D \rightarrow n$
- Hasil akhir bila diselesaikan):

It was disclosed yesterday that several informal but direct contacts have been made with political representatives of the viet cong in moscow

- Analisis frekuensi tetap bisa dilakukan meskipun spasi dihilangkan:

LIVITCSWPIYVEWHEVSRIQMXXLEYVEOIEWHRXEXIP
 FEMVEWHKVSTYLVZIXLIKIIXPVJSZEYPERRGERI
 MWQLMGLMXQERIWGPSRIHMXQEREKIEETXMJTPRGEV
 EKEITREWHEXXLEXXMZITWAWSQWXSWEXTVEPMRXR
 SJGSTVRIEYVIEXVMUIMWERGMIWXMJMGCMSWXSJ
 OMIQXLIVIQIVIXQSVSTWHKPEGARCSXRWIEVSWII
 BXVIZMXFSJXLIKEGAEWHESWYSWIWIEVXLISXLI
 VXLIRGEPIRQIVIIBGIIHMWYPFLEVHEWHYPSRRFQ
 MXLEPPXLIECCIEVEWGJSJKTVWMRLIHYSPHXLIQI
 MYLXSJXLIMWRIGXQEROIVFVIZEVAEKPIEWHXEAM
 WYEPPXLMWYRMWXSGSWMRHIVEXMSWMGSTPHLEVHP
 FKPEZINTCMXIVJSVLMRSCMWMSWVIRCIGXMWYMX

- Hasil perhitungan frekuensi kemunculan huruf:
 - huruf I paling sering muncul,
 - XL adalah bigram yang paling sering muncul,
 - XLI adalah trigram yang paling sering muncul.

Ketiga data terbanyak ini menghasilkan dugaan bahwa

I	berkoresponden dengan huruf plainteks e,
XLI	berkoresponden dengan the,
XL	berkoresponden dengan th

Pemetaan:

$$\begin{aligned} I &\rightarrow e \\ X &\rightarrow t \\ L &\rightarrow h \end{aligned}$$

- XLEX dipetakan menjadi th*t.
- Kata yang cocok untuk th*t. adalah that.
- Jadi kita memperoleh: E → a
- Hasil iterasi pertama:

heVeTCSWPeYVaWHaVSReQMthaYVaOeaWHRtatePFaMVaWHKVSTYhtZet
 heKeetPeJVSZaYPaRRGaReMWQhMGhMtQaReWGPSReH MtQaRaKeaTtMJT
 PRGaVaKaeTRaWHatthat t MZeTWAWSQWtSWatTVaPMRtRSJGSTVReaYVe
 atCVMUeMWaRGMeWtMjMGCsmWtSJOMeQtheVeQeVetQSVSTWHKPaGARCS
 tRWeaVSWeeBtVeZMtFSJtheKaGAaWHaPSWYSWeWeaVtheStheVtheRGa
 PeRQeVeeBGeeHMWYPFhaVHaWHYPSRRFQMthapPt heaCCeaVaWGeSJKT
 WMRheHYSPHtheQeMYhtSJt heMWRGtQaROeVFVeZaVAaKPeaWhtaAMWY
 aPPthMWYRMWtSGSWRMHeVatMSWMGSTPHhaVHPFKPaZeNTCMteVJSvhMR
 SCMWMSSVVeRCeGtMWYMT

- Selanjutnya,
 Rstate mungkin adalah state,
 at that tMZE mungkin adalah at that time,
 heVe mungkin adalah here.

- Jadi, kita memperoleh pemetaan baru:

$$\begin{aligned} R &\rightarrow s \\ M &\rightarrow i \\ Z &\rightarrow m \\ V &\rightarrow r \end{aligned}$$

■ Hasil iterasi ke-2:

hereTCSWPYraWHarSseQi thaYraOeaWHstatePFairaWHKrSTYhtmetheK
eetPeJrSmaYPassGaseiWQhiGhitQaseWGPSseHitQasaKeaTt i JTPsGara
KaeTsaWHat that timeTWAWSQWtSWatTraPistsSJGSTrseaYreatCriUeiW
asGieWt i J i GCSiWtSJOieQthereQeretQSrSTWHKPaGAsCSt sWearSWeeBt
remitFSJtheKaGAaWHaPSWYSWeWeartheStherthesGaPesQereBGeheeHiW
YPFharHaWHYPSssFQithaPPtheaCCearaWGeSJKTrWi sheHYSPhtheQeiYh
tSJtheiWseGtQasOerFremarAakPeaWhtaAiWYaPPthiWYsiWtSGSWsHer
atiSWiGSTPHharHPFKPameNTCiterJSrhisiSCiWiSWresCeGtiWYit

- Tersukan, dengan menerka kata-kata yang sudah dikenal, misalnya remarA mungkin remark , dsb