Bahan kuliah IF4020 Kriptografi

Kriptanalisis Sederhana
(Bagian 2)

Oleh: Dr. Rinaldi Munir

Prodi Informatika

Sekolah Teknik Elektro dan Informatika

2019

Metode Kasiski

- Kembali ke *Vigenere cipher*…

- Friedrich Kasiski adalah orang yang pertama kali memecahkan *Vigènere cipher* pada Tahun 1863 .

Friedrich Kasiski
Born: November 29, 1805 @ Schlochau, Kingdom of Prussia
Died: May 22, 1881 (aged 75) @ Neustettin, German Empire
Nationality: German

- Metode Kasiski tidak secara langsung menemukan kunci Vigenere Cipher, tetapi membantu menemukan panjang kunci *Vigenere cipher*.

- Metode Kasiski memanfaatkan keuntungan bahwa bahasa Inggris tidak hanya mengandung perulangan huruf,

- tetapi juga perulangan pasangan huruf atau tripel huruf, seperti `TH`, `THE`, `EN`, dsb.

- Perulangan kelompok huruf ini ada kemungkinan menghasilkan kriptogram yang berulang.

Contoh 1:

Plainteks : **crypto**isshortfor**crypto**graphy

Kunci : abcdabcdabcdabcdabcdabcdabcdabcd

Cipherteks : **CSASTP**KVSIQUTGQU**CSASTP**IUAQJB

- Pada contoh ini, crypto dienkripsi menjadi kriptogram yang sama, yaitu **CSATP**.

- Tetapi kasus seperti ini tidak selalu demikian, misalnya pada contoh berikut ini....

Contoh 2:

Plainteks : **crypto**isshortfor**crypto**graphy

Kunci : abcdefabcdefabcdefabcdefabcd

Cipherteks : **CSASXT**ITUKWSTGQU**CWYQVR**KWAQJB

- Pada contoh di atas, crypto tidak dienkripsi menjadi kriptogram yang sama.

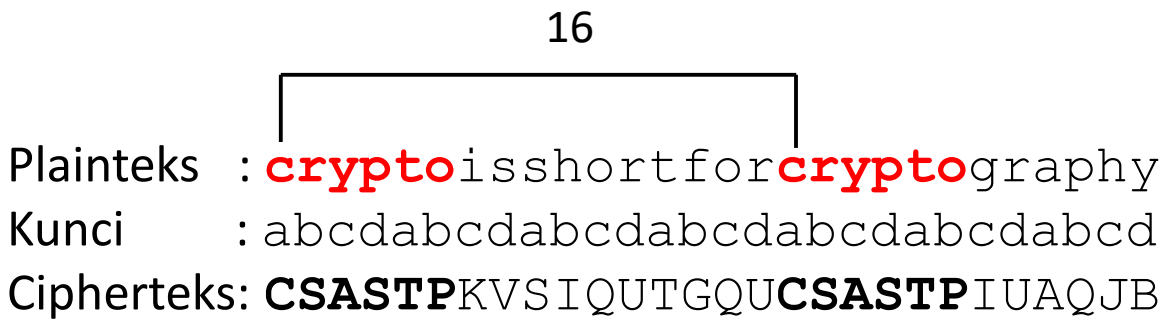
- Mengapa bisa demikian?

- Secara intuitif: jika jarak antara dua buah *string* yang berulang di dalam plainteks merupakan kelipatan dari panjang kunci,

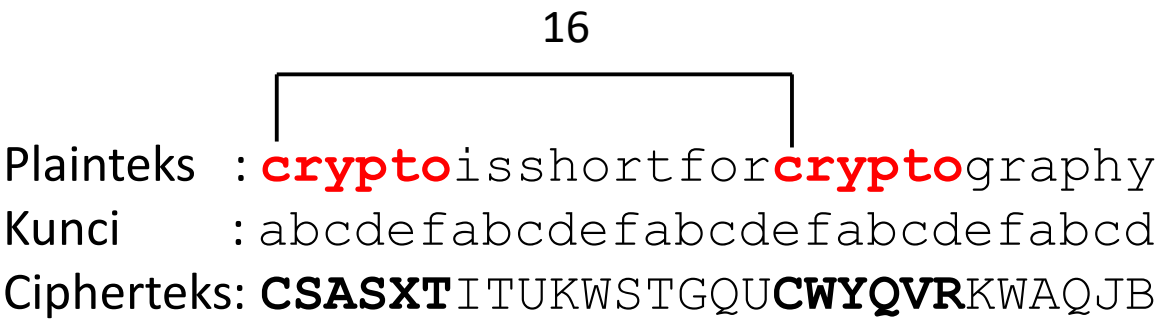
- maka *string* yang sama tersebut akan muncul menjadi kriptogram yang sama pula di dalam cipherteks.

- Pada Contoh 1,

  - kunci = `abcd`
  - panjang kunci = 4
  - jarak antara dua `crypto` yang berulang = 16
  - 16 = kelipatan 4



∴ `crypto` dienkripsi menjadi kriptogram yang sama



- Pada Contoh 2,
  - kunci = `abcdef`
  - panjang kunci = 6
  - jarak antara dua `crypto` yang berulang = 16
  - 16 bukan kelipatan 6

∴ `crypto` tidak dienkripsi menjadi kriptogram yang sama

- Goal metode Kasiski: mencari dua atau lebih kriptogram yang berulang untuk menentukan panjang kunci.

Langkah-langkah metode Kasiski:

1. Temukan semua kriptogram yang berulang di dalam cipherteks (pesan yang panjang biasanya mengandung kriptogram yang berulang).
2. Hitung jarak antara kriptogram yang berulang
3. Hitung semua faktor (pembagi) dari jarak tersebut (faktor pembagi menyatakan panjang kunci yang mungkin).
4. Tentukan irisan dari himpunan faktor pembagi tersebut. Nilai yang muncul di dalam irisan menyatakan angka yang muncul pada semua faktor pembagi dari jarak-jarak tersebut . Nilai tersebut mungkin adalah panjang kunci.

- Contoh:

  DYDUXRMHTVDV**NQD**QNW**DYDUXRMH**ARTJGW**NQD**

  Kriptogram yang berulang: **DYUDUXRMH** dan **NQD**.

  Jarak antara dua buah perulangan **DYUDUXRMH** = 18.

  Semua faktor pembagi 18 : {18, 9, 6, 3, 2}

  Jarak antara dua buah perulangan **NQD** =20.

  Semua faktor pembagi 20 : {20, 10, 5, 4, 2}.

  Irisan dari kedua buah himpunan tersebut adalah 2

  Panjang kunci kemungkinan besar adalah 2.

- Setelah panjang kunci diketahui, maka langkah berikutnya menentukan kata kunci

- Kata kunci dapat ditentukan dengan menggunakan *exhaustive key serach*

- Jika panjang kunci = *p*, maka jumlah kunci yang harsu dicoba adalah 26^p

- Namun lebih mangkus menggunakan metode analisis frekuensi.

Langkah-langkahnya sbb:

1. Misalkan panjang kunci yang sudah berhasil dideduksi adalah *n.* Setiap huruf kelipatan ke-*n* pasti dienkripsi dengan huruf kunci yang sama. Kelompokkan setiap huruf ke-*n* bersama-sama sehingga kriptanalis memiliki *n* buah "pesan", masing-masing dienkripsi dengan substitusi alfabet-tunggal (dalam hal ini *Caesar cipher*).

2. Tiap-tiap pesan dari hasil langkah 1 dapat dipecahkan dengan metode analisis frekuensi.

3. Dari hasil langkah 2 kriptanalis dapat menyusun huruf-huruf kunci. Atau, kriptanalis dapat menerka kata yang membantu untuk memecahkan cipherteks

- Contoh:

1 2 3 4

LJVBQ STNEZ LQMED **LJV**MA MPKAU FAVAT **LJV**DA YYVNF JQLNP **LJV**HK VTRNF
LJVCM LKETA **LJV**HU YJVSF KRFTT WEFUX VHZNP

5 6

Kriptogram yang berulang adalah **LJV**.

Jarak **LJV** ke-1 dengan **LJV** ke-2 = 15

Jarak **LJV** ke-2 dengan **LJV** ke-3 = 15

Jarak **LJV** ke-3 dengan **LJV** ke-4 = 15

Jarak **LJV** ke-4 dengan **LJV** ke-5 = 10

Jarak **LJV** ke-5 dengan **LJV** ke-6 = 10

Faktor pembagi 15 = {3, 5, 15}

Faktor pembagi 10 = {2, 5, 10}

Irisan kedua himpunan ini = 5. Jadi, panjang kunci diperkirakan = 5

- Kelompokkan “pesan” setiap kelipatan ke-5, dimulai dari huruf cipherteks pertama, kedua, dan seterusnya.

LJVBQ STNEZ LQMED **LJV**MA MPKAU FAVAT **LJV**DA YYVNF JQLNP **LJV**HK
VTRNF **LJV**CM LKETA **LJV**HU YJVSF KRFTT WEFUX VHZNP

Kelompok	Pesan	Huruf paling sering muncul
1	LSLLM FLYJL VLLLY KWV	L
2	JTQJP AJYQJ TJKJJ REH	J
3	VNMVK VVVLV RVEVV FFZ	V
4	BEEMA ADNNH NCTHS TUN	N
5	QZDAU TAFPK FMAUF TXP	A

- Dalam Bahasa Inggris, 10 huruf yang yang paling sering muncul adalah E, T, A, O, I, N, S, H, R, dan D,

- Triplet yang paling sering muncul adalah THE. Karena **LJV** paling sering muncul di dalam cipherteks, maka dari 10 huruf tsb semua kemungkinan kata 3-huruf dibentuk dan kata yang yang cocok untuk **LJV** adalah THE.

- Jadi, kita dapat menerka bahwa **LJV** mungkin adalah THE.

- Dari sini kita buat tabel yang memetakan huruf plainteks dengan cipherteks dan huruf-huruf kuncinya (ingatlah bahwa setiap nilai numerik dari huruf kunci menyatakan jumlah pergeseran huruf pada *Caesar cipher*):

Kelompok	Huruf plainteks	Huruf cipherteks	Huruf kunci
1	T	L	S (=18)
2	H	J	C (=2)
3	E	V	R (=17)
4	N	N	A (=0)
5	O	A	M (=12)

Jadi, kuncinya adalah SCRAM

- Dengan menggunakan kunci SCRAM cipherteks berhasil didekripsi menjadi:

```
THEBE ARWEN TOVER THEMO UNTAI NYEAH
THEDO GWENT ROUND THEHY DRANT THECA
TINTO THEHI GHEST SPOTH ECOUL DFIND
```

- atau dalam kalimat yang lebih jelas:

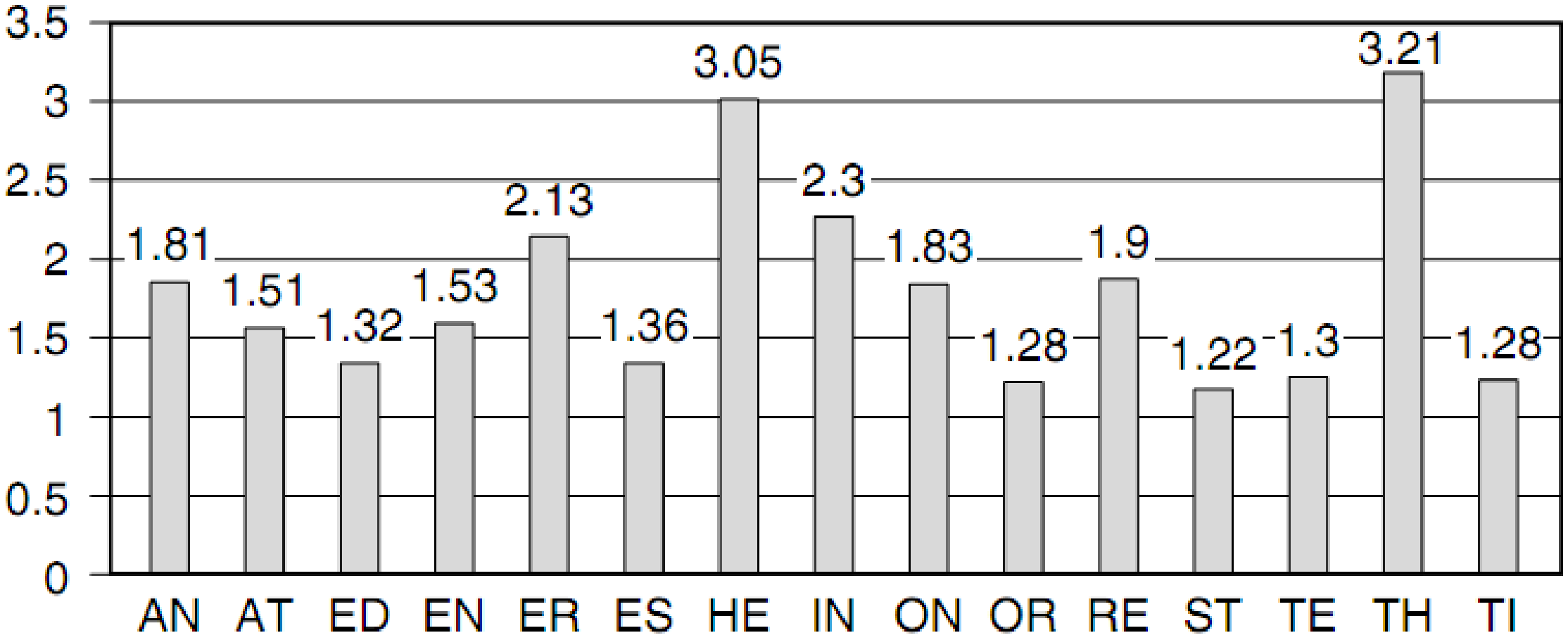
```
THE BEAR WENT OVER THE MOUNTAIN YEAH
THE DOG WENT ROUND THE HYDRANT
THE CAT INTO THE HIGHEST SPOT HE COULD FIND
```

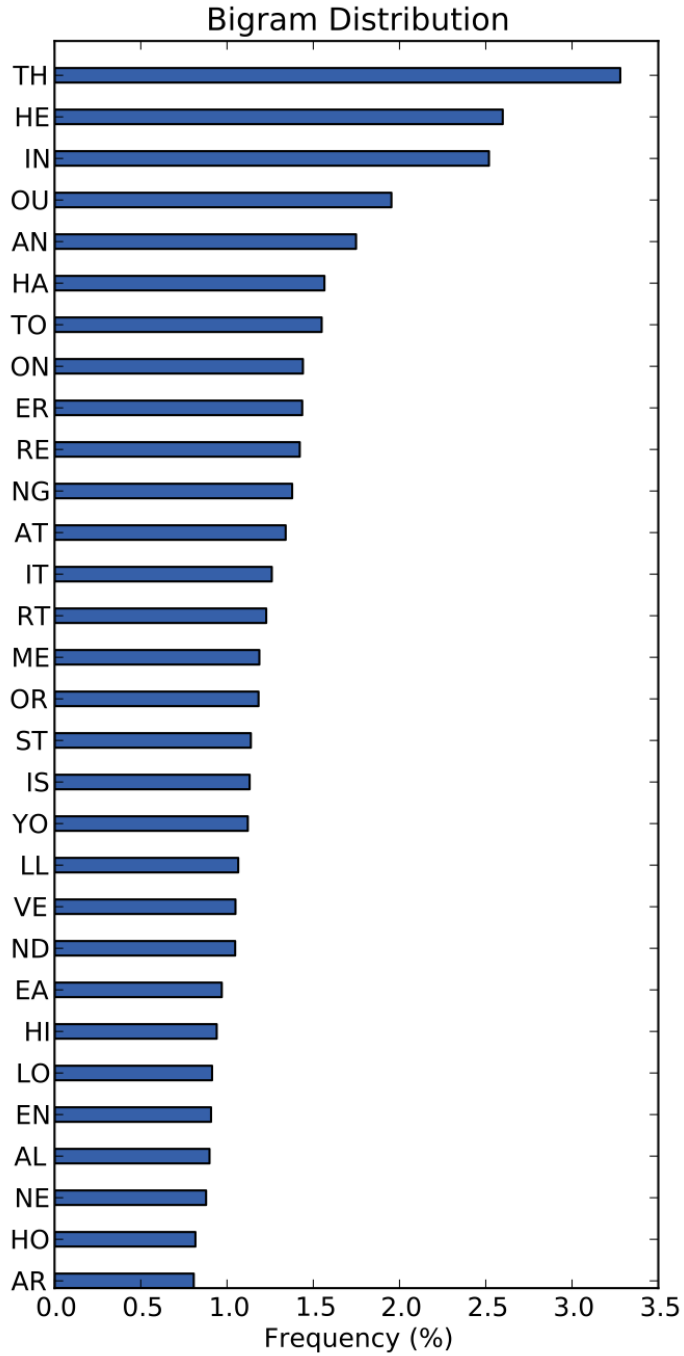

Kriptanalisis Playfair Cipher

- Sayangnya ukuran poligram di dalam *Playfair cipher* tidak cukup besar, hanya dua huruf sehingga *Playfair cipher* tidak aman.

- *Playfair cipher* dapat dipecahkan dengan analisis frekuensi pasangan huruf, karena terdapat tabel frekuensi kemunculan pasangan huruf dalam Bahasa Inggris.

- Dalam Bahasa Inggris pasangan huruf TH dan HE paling sering muncul.


3.5
3
2.5
2
1.5
1
0.5
0
1.81
1.51
1.32
1.53
2.13
1.36
3.05
2.3
1.83
1.28
1.9
1.22
1.3
3.21
1.28
AN
AT
ED
EN
ER
ES
HE
IN
ON
OR
RE
ST
TE
TH
TI


Bigram Distribution
TH
HE
IN
OU
AN
HA
TO
ON
ER
RE
NG
AT
IT
RT
ME
OR
ST
IS
YO
LL
VE
ND
EA
HI
LO
EN
AL
NE
HO
AR
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
Frequency (%)

- Dengan menggunakan tabel frekuensi kemunculan pasangan huruf di dalam Bahasa Inggris dan cipherteks yang cukup banyak, *Playfair cipher* dapat dipecahkan.

- Kelemahan lainnya, bigram dan kebalikannya (misal AB dan BA) akan didekripsi menjadi pola huruf plainteks yang sama (misal RE dan ER). Di dalam bahasa Inggris terdapat banyak kata yang mengandung bigram terbalik seperti REceivER dan DEpartED.