Information Security

CS3002

Lecture 3

31th August 2023

Dr. Rana Asif Rehman

Email: r.asif@nu.edu.pk

3. Playfair Cipher

- Not even the large number of keys in a homophonic cipher provides security.
- One approach to improving security was to encrypt multiple letters.
- The **Playfair cipher** is an example, invented by Charles Wheatstone in 1854.

Playfair Key Matrix

- A 5x5 matrix of letters based on a keyword
- Fill in letters of keyword (minus duplicates)
- Fill rest of matrix with other letters

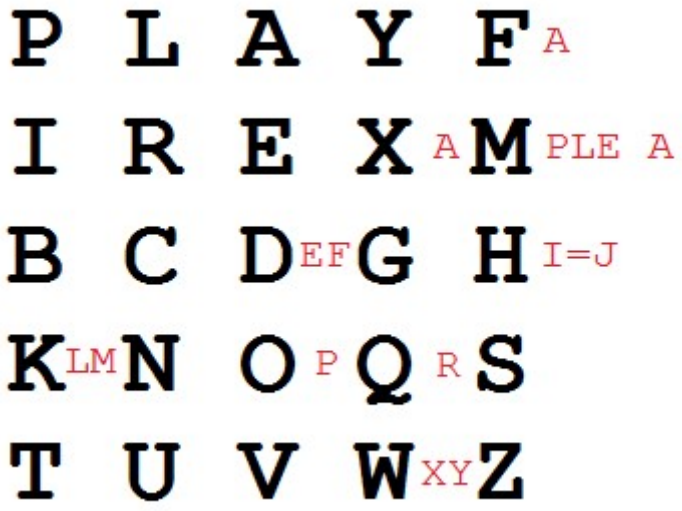
e.g. using the keyword MONARCHY

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

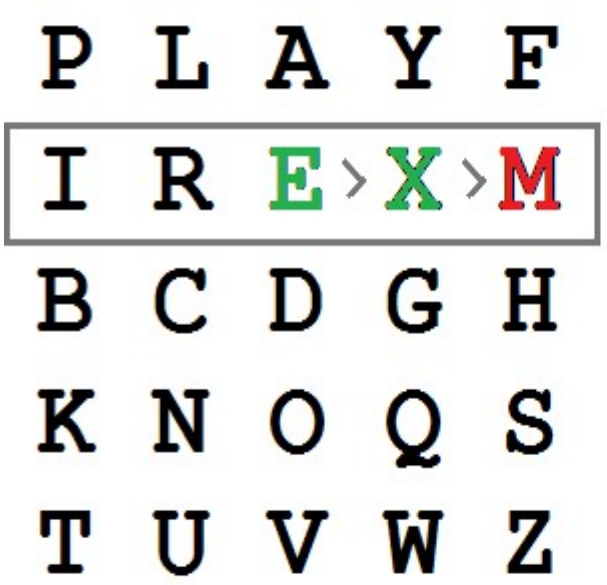
Encrypting and Decrypting

- Plaintext is encrypted two letters at a time
  1. For odd combination append ‘X’ to make it even.
  2. If a pair is a repeated letter, insert filler like 'X’
  3. If both letters fall in the same row, replace each with letter to right (wrapping back to start from end)
  4. If both letters fall in the same column, replace each with the letter below it (again wrapping to top from bottom)
  5. Otherwise each letter is replaced by the letter in the same row and in the column of the other letter of the pair

How the key formed (Keyword: PLAYFAIREXAMPLE)



Formation of encryption and decryption an example Row rule



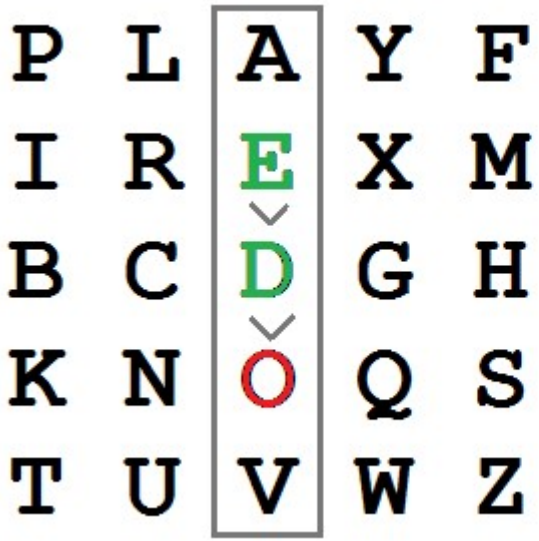
EX

Shape: Row

Rule: Pick Items to Right of Each Letter, Wrap to Left if Needed

XM

Formation of encryption and decryption an example Column rule

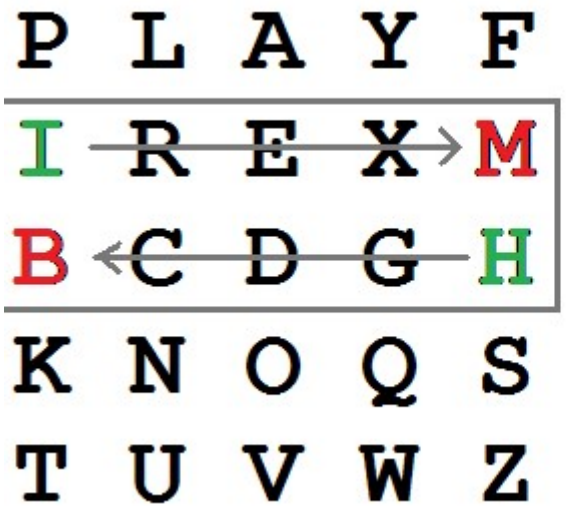


DE

Shape: Column
Rule: Pick Items Below Each Letter, Wrap to Top if Needed

OD

Formation of encryption and decryption an example Rectangle rule



HI

Shape: Rectangle
Rule: Pick Same Rows, Opposite Corners

BM

Example

- Encrypting the message

"Hide the gold in the tree stump"

HI DE TH EG OL DI NT HE TR EX ES TU MP

Example

1. The pair HI forms a rectangle, replace it with BM
2. The pair DE is in a column, replace it with OD
3. The pair TH forms a rectangle, replace it with ZB
4. The pair EG forms a rectangle, replace it with XD
5. The pair OL forms a rectangle, replace it with NA
6. The pair DI forms a rectangle, replace it with BE
7. The pair NT forms a rectangle, replace it with KU
8. The pair HE forms a rectangle, replace it with DM
9. The pair TR forms a rectangle, replace it with UI
10. The pair EX (X inserted to split EE) is in a row, replace it with XM
11. The pair ES forms a rectangle, replace it with MO
12. The pair TU is in a row, replace it with UV
13. The pair MP forms a rectangle, replace it with IF

Example

- Encrypting the message

"Hide the gold in the tree stump":

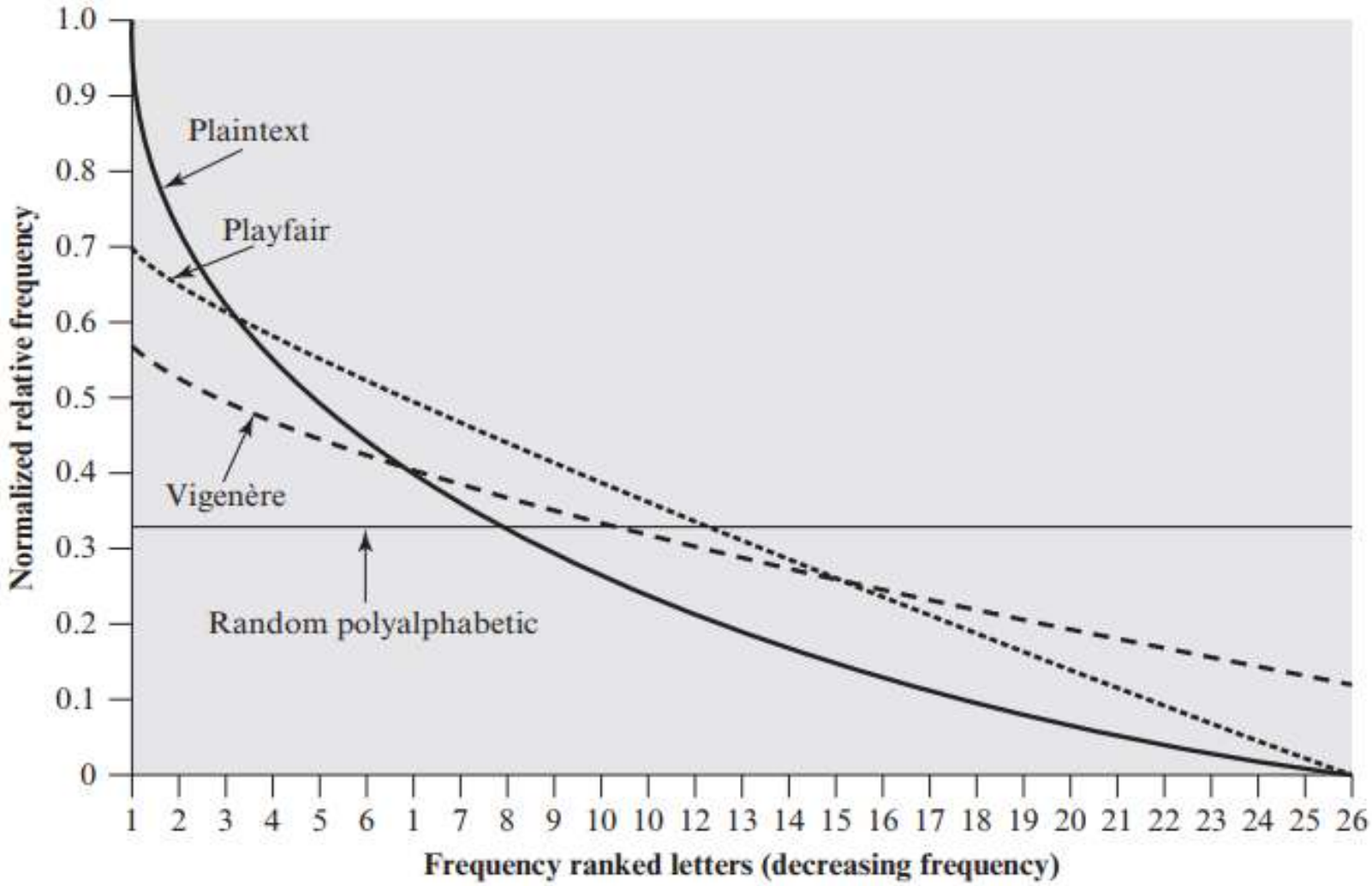
HI DE TH EG OL DI NT HE TR EX ES TU MP

BM OD ZB XD NA BE KU DM UI XM MO UV IF

Security of Playfair Cipher

- Security much improved over monoalphabetic.
- Since have 26 x 26 = 676 digrams.
- Would need a 676 entry frequency table to analyse (verses 26 for a monoalphabetic).
- And correspondingly more ciphertext.
- Was widely used for many years
  - e.g. By US & british military in WW1
  - World war II by US Army and other Allied
- It **can** be broken, given a few hundred letters.
- Since still has much of plaintext structure.

Relative Frequency of Occurrence of Letters



Plot Mechanism

of occurrence of each letter in the text

/

# of occurrences of the most frequently used letter

4. Hill Cipher (Self Study)

- Developed by the mathematician Lester Hill in 1929
- Strength is that it completely hides single-letter frequencies
- The use of a larger matrix hides more frequency information
- A 3 x 3 Hill cipher hides not only single-letter but also two-letter frequency information
- Strong against a ciphertext-only attack but easily broken with a known plaintext attack

Encryption Decryption

Hill Cipher

- **Encryption :**
  Cipher Text =(Plain Text x Key) Mod 26

- **Decryption:**
  Plain Text =(Cipher Text x Key^{-1}) Mod 26

5. Polyalphabetic Substitution

- Another way to improve on the simple monoalphabetic technique is to use different monoalphabetic substitutions as one proceeds through the plaintext message. The general name for this approach is **polyalphabetic substitution cipher**. All these techniques have the following features in common:
  - **1.** A set of related monoalphabetic substitution rules is used.
  - **2.** A key determines which particular rule is chosen for a given transformation.

Polyalphabetic Ciphers

- Polyalphabetic substitution ciphers improve security using multiple cipher alphabets.
- Make cryptanalysis harder with more alphabets to guess and flatter frequency distribution.
- Use a key to select which alphabet is used for each letter of the message.
- Use each alphabet in turn repeat from start after end of key is reached.

5.1. Vigenère Cipher

- Simplest polyalphabetic substitution cipher.
- Effectively multiple caesar ciphers.
- Key is multiple letters long $K = k_1\ k_2\ \ldots\ k_d$
- i^{th} letter specifies i^{th} alphabet to use.
- Use each alphabet in turn.
- Repeat from start after d letters in message.
- Decryption simply works in reverse.

Example of Vigenère Cipher

$$C_i = (p_i + k_{i \bmod m}) \bmod 26$$

```
key:          deceptivedeceptivedeceptive
plaintext:    wearediscoveredsaveyourself
ciphertext:   ZICVTWQNGRZGVTWAVZHCQYGLMGJ
```

Expressed numerically, we have the following result.

key	3	4	2	4	15	19	8	21	4	3	4	2	4	15
plaintext	22	4	0	17	4	3	8	18	2	14	21	4	17	4
ciphertext	25	8	2	21	19	22	16	13	6	17	25	6	21	19

key	19	8	21	4	3	4	2	4	15	19	8	21	4
plaintext	3	18	0	21	4	24	14	20	17	18	4	11	5
ciphertext	22	0	21	25	7	2	16	24	6	11	12	6	9

$$p_i = (C_i - k_{i \bmod m}) \bmod 26$$

Security of Vigenère Ciphers

- Have multiple ciphertext letters for each plaintext letter.
- Hence letter frequencies are obscured.
- But not totally lost (see frequency distribution)

Kasiski Method

- Method developed by Babbage / Kasiski.
- Repetitions in ciphertext give clues to period.
- Find same plaintext an exact period apart which results in the same ciphertext eg repeated “VTW” in previous example suggests size of 3 or 9
- Then attack each monoalphabetic cipher individually using same techniques as before.

Autokey Cipher

- Ideally want a key as long as the message
- Vigenère proposed the **autokey** cipher with keyword is prefixed to message as key knowing keyword can recover the first few letters use these in turn on the rest of the message.
- But still have frequency characteristics to attack
- E.g. given key *deceptive*

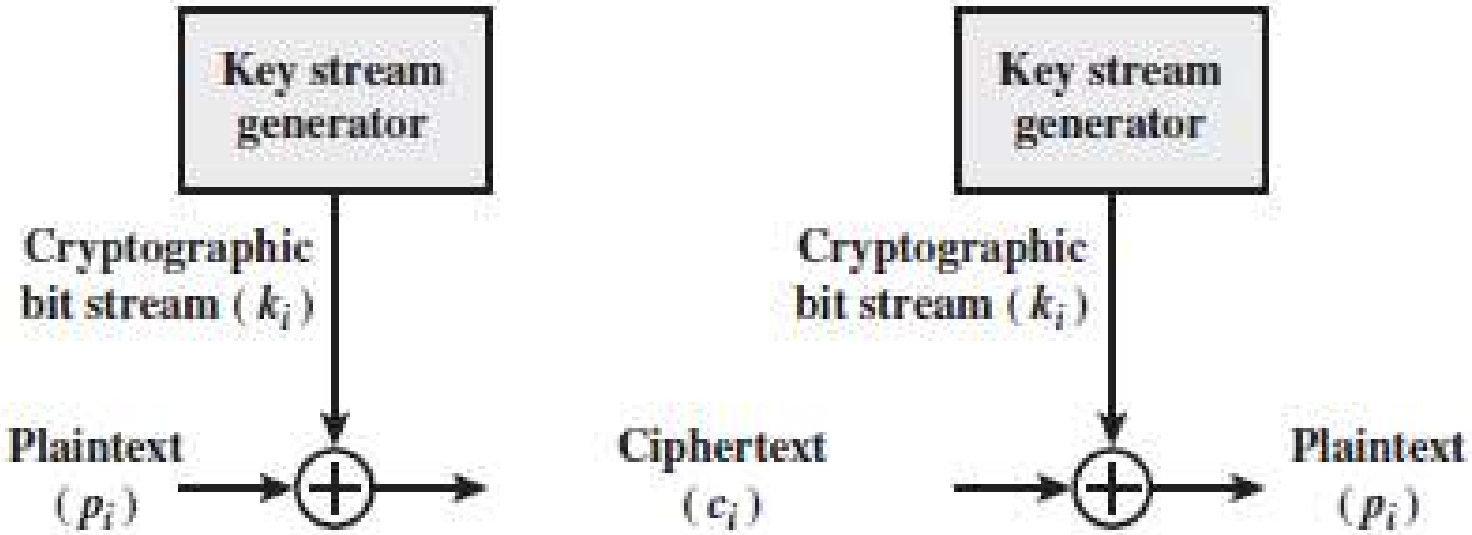
```
key:        deceptivewearediscoveredsav
plaintext:  wearediscoveredsaveyourself
ciphertext:ZICVTWQNGKZEIIGASXSTSLVVWLA
```

Autokey Cipher (cont.)

- Even this scheme is vulnerable to cryptanalysis. Because the key and the plaintext may share the same frequency distribution of letters, a statistical technique can be applied. For example, e enciphered by *e* can be expected to occur with a frequency of $(0.127)^2$ 0.016, whereas t enciphered by *t* would occur only about half as often. These regularities can be exploited to achieve successful cryptanalysis.

5.2. Vernam Cipher

The ultimate defense against such a cryptanalysis is to choose a keyword that is as long as the plaintext and has no statistical relationship to it. [coming forward from Vigenère and Autokey ciphers]



Vernam Cipher (cont.)

- SENDING
- message: 001011010111 ...
- pad: 100111001011 ...
- XOR ----------
- cipher: 101100011100 ...

- RECEIVING
- cipher: 1 0 1 1 0 0 0 1 1 1 0 0 ...
- pad: 1 0 0 1 1 1 0 0 1 0 1 1 ...
- XOR ----------
- message: 0 0 1 0 1 1 0 1 0 1 1 1 ...

$$c_i = p_i \oplus k_i$$

p_i = ith binary digit of plaintext
k_i = ith binary digit of key
c_i = ith binary digit of ciphertext
$\oplus$ = exclusive-or (XOR) operation

$$p_i = c_i \oplus k_i$$

Very long but repeating keywords – Still exploitable with sufficient long ciphertext.

6. One-Time Pad

- If a truly random key as long as the message is used, the cipher will be secure called a One-Time pad.
- It is unbreakable since ciphertext bears no statistical relationship to the plaintext.
- Since for **any plaintext** & **any ciphertext** there exists a key mapping one to other.
- Can only use the key **once** though.
- Problems in generation & safe distribution of key.
- Cannot use extensively and bears perfect secrecy.

One-Time Pad Security

ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS

We now show two different decryptions using two different keys:

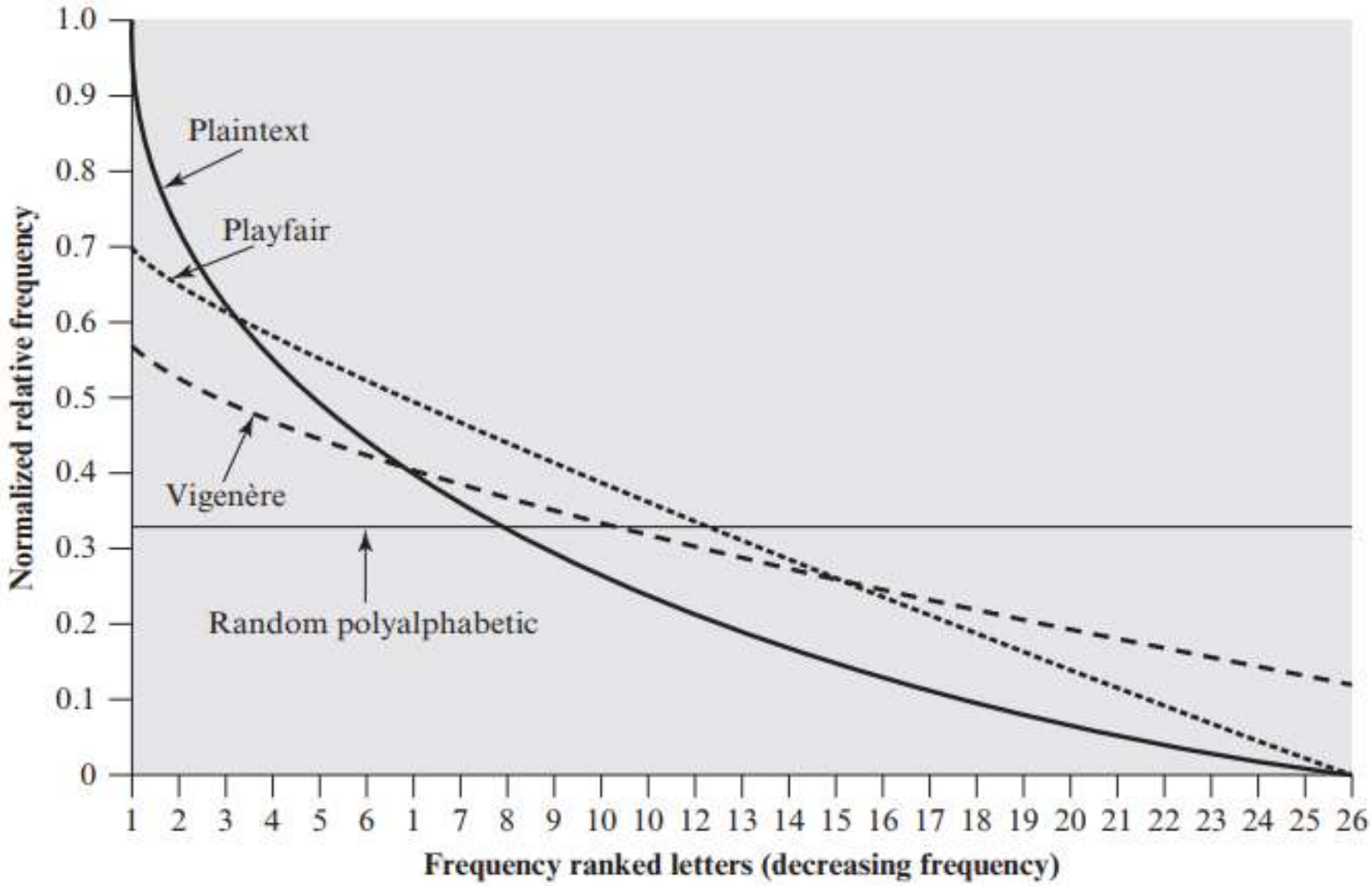
```
ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        pxlmvmsydofuyrvzwc tnlebnecvgdupahfzzlmnyih
plaintext:  mr mustard with the candlestick in the hall

ciphertext: ANKYODKYUREPFJBYOJDSPLREYIUNOFDOIUERFPLUYTS
key:        pftgpmiydgaxgoufhkllmhsqdqogtewbqfgyovuhwt
plaintext:  miss scarlet with the knife in the library
```

Why not practical??

- Large number random key formation
- Distribution & protection among parties

Relative Frequency of Occurrence of Letters



Classical Symmetric Ciphers

- Classical Transposition Ciphers
  - Rail Fence Cipher
  - Row Transposition Cipher
- Product Ciphers

Transposition Ciphers

- Now consider classical **transposition** or **permutation** ciphers.
- A very different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. This technique is referred to as a transposition cipher.

Transposition Ciphers.

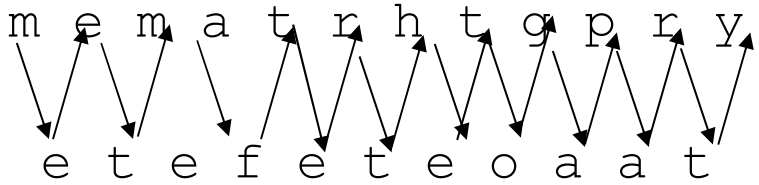
- Hide the message by rearranging the letter order without altering the actual letters used.
- Can recognise these since have the same frequency distribution as the original text.

Transposition Ciphers

- Now consider classical **transposition** or **permutation** ciphers.
- A very different kind of mapping is achieved by performing some sort of permutation on the plaintext letters. This technique is referred to as a transposition cipher.
- These hide the message by rearranging the letter order without altering the actual letters used.
- Can recognise these since have the same frequency distribution as the original text.

1. Rail Fence Cipher

- Write message letters out diagonally over a number of rows
- Then read off cipher row by row
- The message ‘meet me after toga party’
- e. g. write message out as:



- Giving ciphertext

MEMATRHTGPRYETEFETEOAAT

2. Row Transposition Ciphers

- A more complex transposition
- Write letters of message out in rows over a specified number of columns
- Then reorder the columns according to some key before reading off the rows

```
Key:          4 3 1 2 5 6 7
Plaintext:    a t t a c k p
              o s t p o n e
              d u n t i l t
              w o a m x y z
Ciphertext:   TTNAAPTMTSUOAODWCOIXKNLYPETZ
```

Problem with Transposition approach

- A pure transposition cipher is easily recognized because it has the same letter frequencies as the original plaintext.
- The transposition cipher can be made significantly more secure by performing more than one stage of transposition.

```
Key:      4 3 1 2 5 6 7
Input:    t t n a a p t
          m t s u o a o
          d w c o i x k
          n l y p e t z
Output:   NSCYAUOPTTWLTMDNAOIEPAXTTOKZ
```

To visualize the result of this double transposition, designate the letters in the original plaintext message by the numbers designating their position. Thus, with 28 letters in the message, the original sequence of letters is

```
01 02 03 04 05 06 07 08 09 10 11 12 13 14
15 16 17 18 19 20 21 22 23 24 25 26 27 28
```

After the first transposition, we have

```
03 10 17 24 04 11 18 25 02 09 16 23 01 08
15 22 05 12 19 26 06 13 20 27 07 14 21 28
```

which has a somewhat regular structure. But after the second transposition, we have

```
17 09 05 27 24 16 12 07 10 02 22 20 03 25
15 13 04 23 19 14 11 01 26 21 18 08 06 28
```

This is a much less structured permutation and is much more difficult to cryptanalyze.

Product Ciphers (combining two approaches)

- Ciphers using substitutions or transpositions are not secure because of language characteristics.
- Hence consider using several ciphers in succession to make harder, but:
  - Two substitutions make a more complex substitution
  - Two transpositions make more complex transposition
  - But a substitution followed by a transposition makes a new much harder cipher
- This is bridge from classical to modern ciphers.