Polyalphabetic Ciphers

- 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.
- The best known, and one of the simplest, polyalphabetic ciphers is the ***Vigenère cipher***.

The Vigenere Cipher

- Let each letter of the encryption key denote a ***shifted Caesar cipher***, the shift corresponding to the key.
- Since, in general, the encryption key will be shorter than the message to be encrypted, for the Vigenere cipher the key is repeated as required. For example, the key here is the string "***abracadabra***".
- Now a plaintext message may be encrypted as follows

key:	a b r a c a d a b r a a b r a c a d a b r a
plaintext:	c a n y o u m e e t m e a t m i d n i g h t
ciphertext:	C B E Y Q U P E F K M E B K . . . . .

The Vigenere Cipher: Example

If the keyword is "***deceptive***", the message "*we are discovered save yourself*" is encrypted as:

key: d ec e p t i v e de c e p t i v e de c e p t i ve
plaintext: wea r e d i s c ov e r e d s a ve y o u r s e l f
ciphertext: ZICVTWQNGRZGVTWAVZHCQYGLMGJ

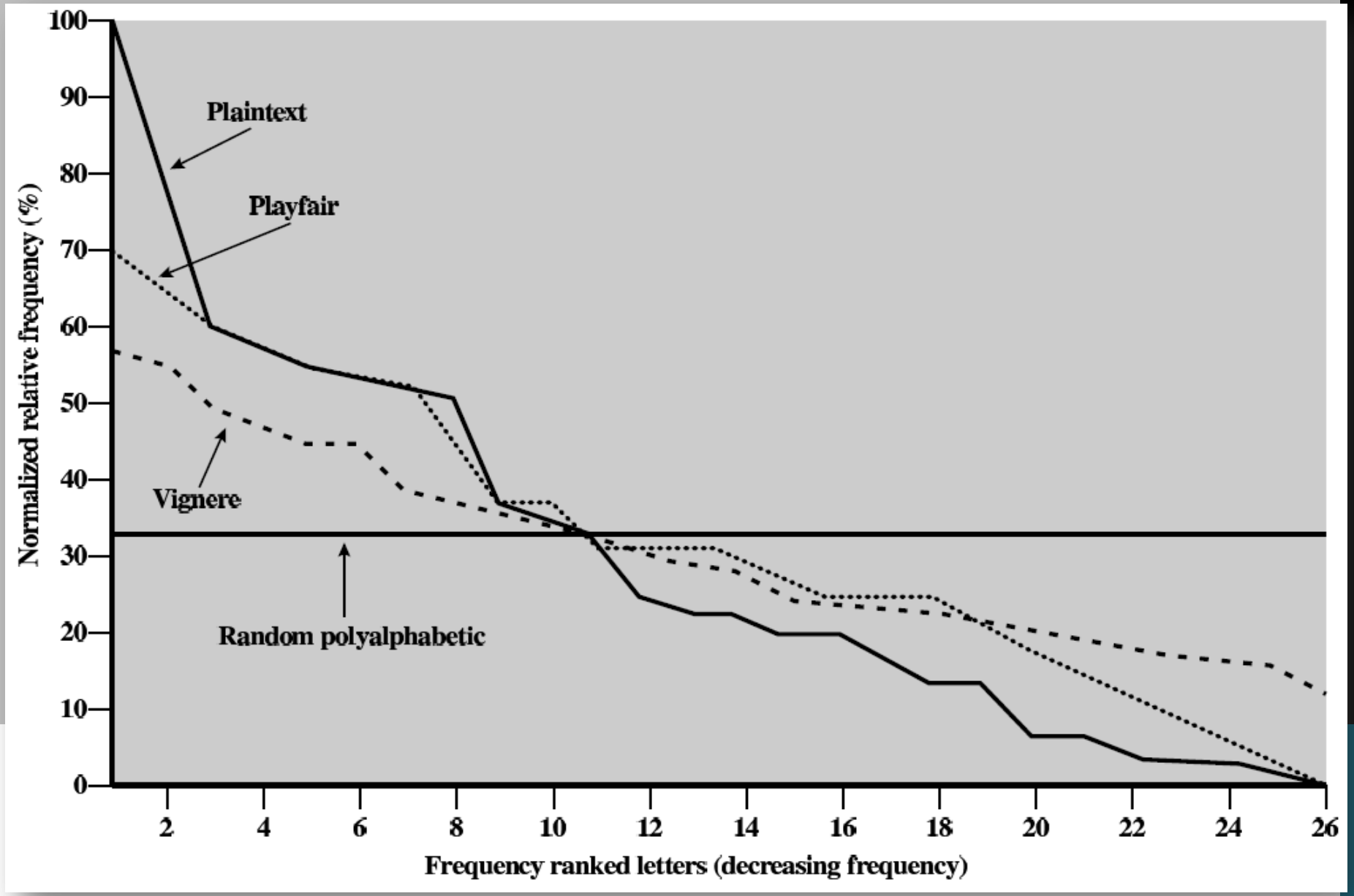
Expressed numerically, we have the following result.

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

How Secure is the Vigenere Cipher? (1)

- Since there exist in the output multiple ciphertext letters for each plaintext letter, you would expect that the relative frequency distribution would be effectively destroyed. But as can be seen in the plots on next page, a great deal of the input statistical distribution still shows up in the output.
  - The plot shown for Vigenere cipher is for an encryption key that is 9 letters long.
- Obviously, the longer the encryption key, the greater the masking of the structure of the plaintext.
- The best possible key is as long as the plaintext message and consists of a purely random permutation of the 26 letters of the alphabet. This would yield the ideal plot shown in the figure. The ideal plot is labeled "Random polyalphabetic" in the figure.



RELATIVE FREQUENCY OF OCCURRENCE OF LETTERS

How Secure is the Vigenere Cipher? (2)

- In general, to break the Vigenere cipher, you first try to estimate the ***length of the encryption key***. This length can be estimated by using the logic that plaintext words separated by multiples of the length of the key will get encoded in the same way.
- If the estimated length of the key is ***N***, then the cipher consists of ***N*** *monoalphabetic substitution* ciphers and the plaintext letters at positions 1, ***N***, 2***N***, 3***N***, etc., will be encoded by the same monoalphabetic cipher.
- This insight can be useful in the decoding of the monoalphabetic ciphers involved.

Vernam Cipher (1)

- The ultimate defense against cryptanalysis is to choose a keyword that is as long as the plaintext and has no statistical relationship to it. Such a system was introduced by Gilbert Vernam in 1918. His system works on binary data (bits) rather than letters. The system can be expressed as follows.

$$c_i = p_i \oplus k_i$$

where

p_i = ith binary digit of plaintext

k_i = ith binary digit of key

c_i = ith binary digit of ciphertext

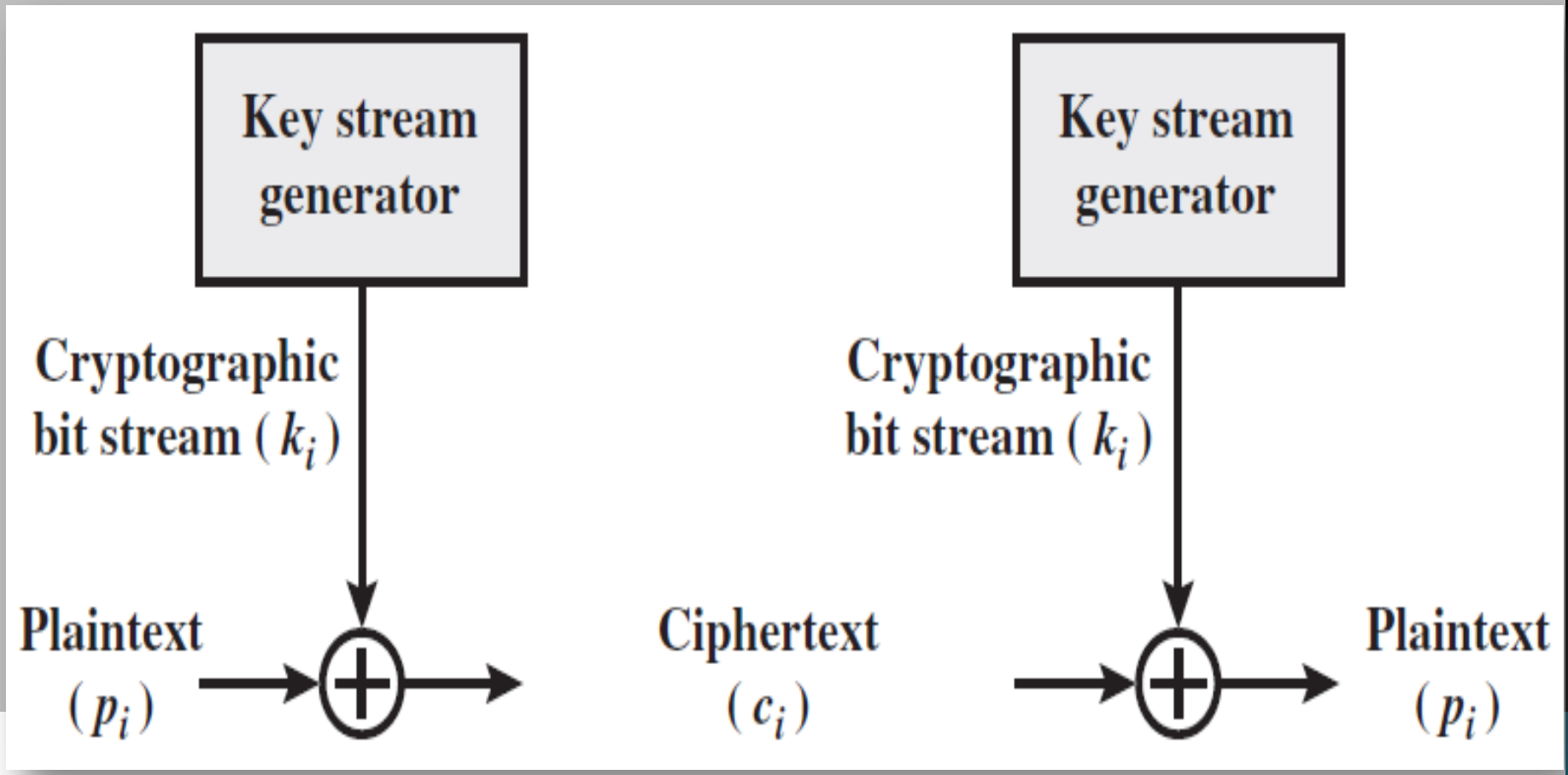
$\oplus$ = exclusive-or (XOR) operation

Vernam Cipher (2)

- Thus, the ciphertext is generated by performing the bitwise XOR of the plaintext and the key. Because of the properties of the XOR, decryption simply involves the same bitwise operation:

$$p_i = c_i \oplus k_i$$

- The essence of this technique is the means of ***construction of the key***.
  - Vernam proposed the use of a running loop of tape that eventually repeated the key, so that in fact the system worked with a very long but repeating keyword.
  - Although such a scheme, with a long key, presents formidable cryptanalytic difficulties, it can be broken with sufficient ciphertext, the use of known or probable plaintext sequences, or both.



Vernam Cipher Encryption and Decryption

One-Time Pad (1)

- Joseph Mauborgne proposed an improvement to Vernam cipher that yields the ultimate in security. He suggested:
  1. using of a random key
  2. key that is truly as long as the message
  3. and with no repetitions.
- Such a scheme, known as ***one-time pad***, is ***unbreakable***.
- It produces random output that bears no statistical relationship to the plaintext. Because the ciphertext contains no information whatsoever about the plaintext, there is no way to break the code.

One-Time Pad (2)

- In theory, we need look no further for a cipher. The one-time pad offers complete security but, in practice, has ***two fundamental difficulties*** that make the one-time pad is of limited utility. These are:

  1. There is the practical problem of making large quantities of random keys.
  2. Even more daunting is the problem of key distribution and protection. For every message to be sent, a key of equal length is needed by both sender and receiver.

Transposition Techniques

- All of our discussion so far has dealt with substitution ciphers. We have talked about monoalphabetic substitutions, polyalphabetic substitutions, etc.
- We will now talk about a different notion in classical cryptography: ***permuting the plaintext***.
- This is how a pure permutation cipher could work:
  1. You write your plaintext message along the ***rows*** of a matrix of some size.
  2. You generate ciphertext by reading along the ***columns***.
  3. The *order in which you read* the columns is determined by the ***encryption Key***.
- The cipher can be made more secure by performing ***multiple rounds of such permutations.***

Columnar Transposition Cipher Example

Encrypt the message: "*meet me at midnight for the godies*" with the following key:

```
key:                4 1 3 6 2 5

plaintext:          m e e t m e
                    a t m i d n
                    i g h t f o
                    r t h e g o
                    d i e s x y

ciphertext:         ETGTIMDFGXEMHHEMAIRDENOOYTITES
```

Steganography

- A plaintext message may be hidden in one of two ways.
  1. The methods of ***steganography*** conceal the existence of the message,
  2. whereas the methods of ***cryptography*** render the message unintelligible to outsiders by various transformations of the text.
- A simple form of steganography is one in which an arrangement of words or letters within an apparently innocuous text spells out the real message. For example, the sequence of first letters of each word of the overall message spells out the hidden message.
- As a contemporary example consider hiding a message by using the least significant bits of frames on a CD. The least significant bit of each 24-bit pixel can be changed without greatly affecting the quality of the image.

Finally . . .

- **Acknowledgment:** These lecture notes are based on the textbook by William Stallings and notes prepared by Avinash Kak, Purdue University. My sincere thanks are devoted to them and to all other people who offered the material on the web.

- Students are advised to study and solve the problems and answer the questions in **Assignment-3**.