Cryptography II

Block ciphers and modes of operations

Review

- Security Goals
  - Confidentiality (secrecy, privacy)
    - Assure that data is accessible to only one who are authorized to know
  - Integrity
    - Assure that data is only modified by authorized parties and in authorized ways
  - Availability
    - Assure that resource is available for authorized users

Review

- Secret-key cryptography
  - symmetric cryptography
  - same key for both encryption & decryption (Z=Z')
- Public-key cryptography
  - asymmetric cryptography
  - encryption key different from decryption key and

Review

- Shift cipher (additive cipher)
  - Key Space: [1 .. 25]
  - Encryption given a key K:
    - Y=X + K (mod 26)
  - Decryption given K:
    - X = Y – K (mod 26)
  - Cryptanalysis
    - exhaustive search (<= 26 possible keys).

Review

- Mono-alphabetical Substitution Cipher
  - The key space: all permutations of $\Sigma = \{A, B, C, \ldots, Z\}$
  - Encryption given a key π:
    - each letter X in the plaintext P is replaced with $\pi(X)$
  - Decryption given a key π :
    - each letter Y in the ciphertext P is replaced with $\pi^{-1}(Y)$
  - Cryptanalysis
    - frequency analysis attacks

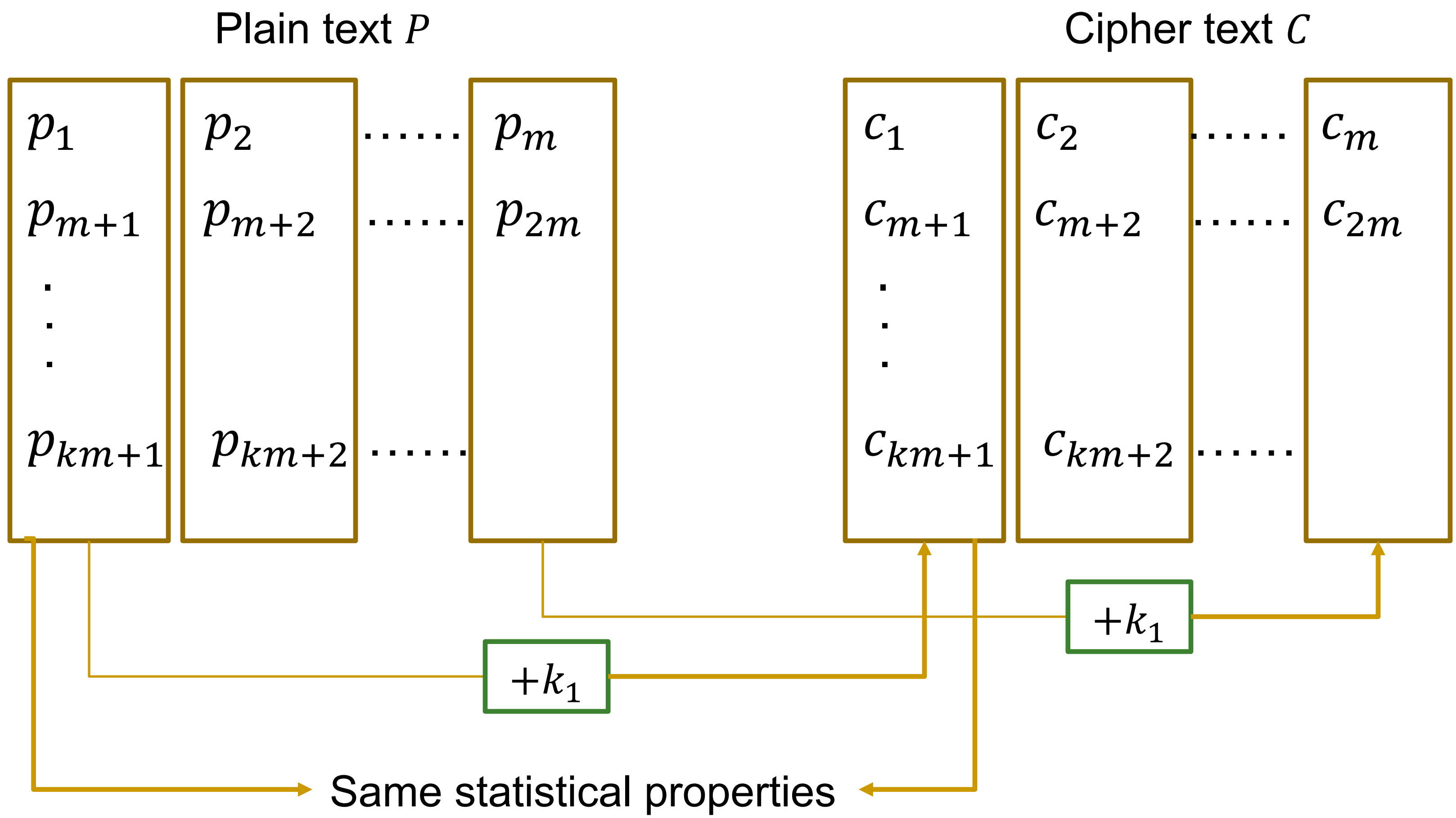
Review

- Polyalphabetic Substitution Ciphers (Vigenère cipher - published in 1586)
  - Definition:
    - Given m, a positive integer, $P = C = (Z_{26})^n$, and $K = (k_1, k_2, \ldots, k_m)$ a key, we define:
  - Encryption:
    - $e_k\ (p_1, p_2 \ldots\ p_m) = (p_1+k_1, p_2+k_2 \ldots p_m+k_m)\ (\mathrm{mod}\ 26)$
  - Decryption:
    - $d_k\ (c_1, c_2 \ldots\ c_m) = (c_1\text{-}k_1, c_2\text{-}k_2\ \ldots\ c_m\text{-}\ k_m)\ (\mathrm{mod}\ 26)$
  - Example:

Plaintext:	C R Y P T O G R A P H Y
Key:	L U C K L U C K L U C K
Ciphertext:	N L A Z E I I B L J J I

Vigenère cipher

- Cryptanalysis



Can be broken by the statistical method once the key length is determined

Vigenère cipher

- How to determine the key length
  - The frequency of letters in $\{p_j, p_{m+j}, \dots, p_{km+j}\}$ is approximately the same as that in the plain text P
  - The frequency of letters in $\{c_j, c_{m+j}, \dots, c_{km+j}\}$ is the same as that in $\{p_j, p_{m+j}, \dots, p_{km+j}\}$
- The index of coincidence (IC)
  - Suppose $x = x_1 x_2 \dots . x_n$ is a string of alphabetic characters → $IC(x)$ is the probability that two random elements of x are identical

Vigenère cipher

- The index of coincidence (IC)
  - Suppose the frequencies of $A, B, \ldots, Z$ in x are $f_0, f_1 \ldots, f_{25}$
  - $IC(x) = \frac{\sum_{i=0}^{25}\binom{f_i}{2}}{\binom{n}{2}} = \sum_{i=0}^{25} \frac{f_i}{n}\frac{f_i - 1}{n-1} \approx \sum_{i=0}^{25}(p_i)^2$

p_i: the frequency of the i-th letter

letter	probability
A	.082
B	.015
C	.028
D	.043
E	.127
F	.022
…	
Z	.001

For an English text
$IC(x) \approx 0.065$

For a totally random string
$$IC(x) \approx \sum_{i=0}^{25} \frac{1}{26} = 0.038$$

Vigenère cipher

- The index of coincidence (IC)
  - Let $P_j = \{p_j, p_{m+j}, \dots, p_{km+j}\}$; $C_j = \{c_j, c_{m+j}, \dots, c_{km+j}\}$
    - $IC(C_j) = IC(P_j) \approx 0.065$
- Cryptanalysis algorithm
  1. Set $m = 1$
  2. Check if m is indeed the key length
     - Divide the cipher into m letter group and compute the IC of each
     - If they are quite the same and approximately equals to 0.065 then m is the key length
     - If they are quite different and smaller than 0.065, then the key length should be greater
  3. Increase m by 1 and go to step 1

Vigenère cipher

- Kasiski method: a hint to find the key length
  - Observation: two identical segments of plaintext will be encrypted to the same cipher text wherever their occurrence in the plain text is δ position apart, $\delta \equiv 0 (mod\ m)$

$$p_1 \;\cdots \boxed{p_j p_{j+1} \cdots p_{j+u}} \cdots\cdots\; p_m$$

$$\vdots$$

$$p_{vm+1} \;\cdots \boxed{p_{vm+j} p_{vm+j+1} \cdots p_{vmj+u}} \cdots\cdots\; p_{vm+m}$$

If these are the same

$$c_1 \;\cdots \boxed{c_j c_{j+1} \cdots c_{j+u}} \cdots\cdots\; c_m$$

$$\vdots$$

$$c_{vm+1} \;\cdots \boxed{c_{vm+j} c_{vm+j+1} \cdots c_{vmj+u}} \cdots\cdots\; c_{vm+m}$$

Then, these will be the same

Vigenère cipher

- Kasiski method
  - Search the cipher text for pairs of identical segments and record the distance between their starting positions
    - Suppose the obtained distances are $\delta_1, \ldots, \delta_k$
  - Then, m should divides the greatest common divisor of $\delta_1, \ldots, \delta_k$

Vigenère cipher

- Example

```
CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP
WQAIIWXNRMGWOIIFKEE
```

Vigenère cipher

- Example

```
CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP
WQAIIWXNRMGWOIIFKEE
```

Kasiski method: CHR's occurrence positions: 1, 166, 236, 276 and 286
→ Distances: 165, 235, 275 and 285
→ Gcd(165, 235, 275, 285) = 5
→ The key length should divides 5

Vigenère cipher

- Example

```
CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQEQERBW
RVXUOAKXAOSXXWEAHBWGJMMQMNKGRFVGXWTRZXWIAK
LXFPSKAUTEMNDCMGTSXMXBTUIADNGMGPSRELXNJELX
VRVPRTULHDNQWTWDTYGBPHXTFALJHASVBFXNGLLCHR
ZBWELEKMSJIKNBHWRJGNMGJSGLXFEYPHAGNRBIEQJT
AMRVLCRREMNDGLXRRIMGNSNRWCHRQHAEYEVTAQEBBI
PEEWEVKAKOEWADREMXMTBHHCHRTKDNVRZCHRCLQOHP
WQAIIWXNRMGWOIIFKEE
```

Confirmation of Kasiski method

M = 1 → IC = 0.045

M = 2 → ICs = 0.046 and 0.041

M = 3 → ICs = 0.043, 0.050, 0.047

M = 4 → ICs = 0.042, 0.039, 0.046, 0.040

M = 5 → ICs = 0.063, 0.068, 0.069, 0.061 and 0.072

Vigenère cipher

i	value of $M_g(\mathbf{y}_i)$								
1	.035	.031	.036	.037	.035	.039	.028	.028	.048
	$\boxed{.061}$	.039	.032	.040	.038	.038	.045	.036	.030
	.042	.043	.036	.033	.049	.043	.042	.036	
2	$\boxed{.069}$	.044	.032	.035	.044	.034	.036	.033	.029
	.031	.042	.045	.040	.045	.046	.042	.037	.032
	.034	.037	.032	.034	.043	.032	.026	.047	
3	.048	.029	.042	.043	.044	.034	.038	.035	.032
	.049	.035	.031	.035	$\boxed{.066}$	.035	.038	.036	.045
	.027	.035	.034	.034	.036	.035	.046	.040	
4	.045	.032	.033	.038	$\boxed{.060}$	.034	.034	.034	.050
	.033	.033	.043	.040	.033	.029	.036	.040	.044
	.037	.050	.034	.034	.039	.044	.038	.035	
5	.034	.031	.035	.044	.047	.037	.043	.038	.042
	.037	.033	.032	.036	.037	.036	.045	.032	.029
	.044	$\boxed{.072}$	.037	.027	.031	.048	.036	.037	

$$M_g = \sum_{i=0}^{25} \frac{p_i f_{i+g}}{n'}$$

If $g \neq k_i$, then $M_g \ll 0.065$

$$K = (9, 0, 13, 4, 19) = JANET$$

The almond tree was in tentative blossom. The days were longer, often ending with magnificent evenings of corrugated pink skies. The hunting season was over, with hounds and guns put away for six months. The vineyards were busy again as the well-organized farmers treated their vines and the more lackadaisical neighbors hurried to do the pruning they should have done in November.

Exercises

- Decode the following cipher texts
  - Encrypted by shift cipher:
    - JBCRCLQRWCRVNBJENBWRWN
  - Encrypted by substitution cipher:
    - Pjmu mu b amtjfo rfsr. Mr jbu cffi fiaowtrfg cw rjf uvcurmrvrmqi amtjfo. Wqv bof xfow nvahw. Rjf amtjfo jbu cffi coqhfi
    - YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ
      NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ
      NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ
      XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
    - Hints:
      - The letters in the English alphabet can be divided into 5 groups of similar frequencies
        - e
        - t,a,o,i,n,s,h,r
        - d,l
        - c,u,m,w,f,g,y,p,b
        - v,k,j,x,q,z
      - Some frequently appearing bigrams or trigrams
        - Th, he, in, an, re, ed, on, es, st, en at, to
        - The, ing, and, hex, ent, tha, nth, was eth, for, dth.

Exercises

- Decode the following cipher texts
  - Encrypted by substitution cipher:

YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR

letter	frequency	letter	frequency
A	0	*N*	9
B	1	*O*	0
C	15	*P*	1
D	13	*Q*	4
E	7	*R*	10
F	11	*S*	3
G	1	*T*	2
H	4	*U*	5
I	5	*V*	5
J	11	*W*	8
K	1	*X*	6
L	0	*Y*	10
M	16	*Z*	20

DZ and *ZW:* four times each
NZ and *ZU:* three times each
RZ, HZ, YZ, FZ, ZR, ZV, ZC, ZD, ZJ: twice each