Classical Cryptography
Outline

[1] Introduction: Some Simple Cryptosystems
  <1> The Shift Cipher
  <2> The Substitution Cipher
  <3> The Affine Cipher
  <4> The Vigenère Cipher
  <5> The Hill Cipher
  <6> The Permutation Cipher

[2] Cryptanalysis
  <1> Cryptanalysis of the Affine Cipher
  <2> Cryptanalysis of the Substitution Cipher
  <3> Cryptanalysis of the Vigenère Cipher
  <4> Cryptanalysis of the Hill Cipher
Classical Cryptography

[1] Introduction

- Alice
- encrypter
- secure channel
- key source
- decrypter
- Bob
- Oscar

Key source: $K$

Encrypted message: $x$

Decrypted message: $y$
Definition 1.1: A cryptosystem is a five-tuple \( (\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D}) \) satisfies

- \( \mathcal{P} \) is a finite set of possible plaintexts
- \( \mathcal{C} \) is a finite set of possible ciphertexts
- \( \mathcal{K} \), the keyspace, is a finite set of possible keys
- For each \( K \in \mathcal{K} \) there is an encryption rule \( e_K \in \mathcal{E} \) and a corresponding decryption rule \( d_K \in \mathcal{D} \)

\[ e_K : \mathcal{P} \rightarrow \mathcal{C} \]

\[ d_K : \mathcal{C} \rightarrow \mathcal{P} \]

\[ d_K(e_K(x)) = x \text{ for every plaintext } x \in \mathcal{P} \]
Classical Cryptography

- **Definition 1.2**: $a$ and $b$ are integers, $m$ is a positive integer.
  - congruence: $a \equiv b \pmod{m}$ if $m$ divides $b-a$

- $\mathbb{Z}_m$: the set \{0, 1, ..., $m-1$\}
  - with 2 operations $+$ and $\times$
  - $10 + 20 = 4$ in $\mathbb{Z}_{26}$ ($10 + 20 \mod 26 = 4$)
  - $10 \times 20 = 18$ in $\mathbb{Z}_{26}$ ($10 \times 20 \mod 26 = 18$)
Classical Cryptography

- **<1> Shift Cipher**
  - **Cryptosystem 1.1: Shift Cipher**
    - $P = C = K = \mathbb{Z}_{26}$
    - $K, x, y \in \mathbb{Z}_{26}$
    - $e_K(x) = (x + K) \mod 26$
    - $d_K(y) = (y - K) \mod 26$

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
<th>C</th>
<th>D</th>
<th>E</th>
<th>F</th>
<th>G</th>
<th>H</th>
<th>I</th>
<th>J</th>
<th>K</th>
<th>L</th>
<th>M</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>A</td>
<td>B</td>
<td>C</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>G</td>
<td>H</td>
<td>I</td>
<td>J</td>
<td>K</td>
<td>L</td>
<td>M</td>
</tr>
<tr>
<td>5</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>9</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>11</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>N</td>
<td>O</td>
<td>P</td>
<td>Q</td>
<td>R</td>
<td>S</td>
<td>T</td>
<td>U</td>
<td>V</td>
<td>W</td>
<td>X</td>
<td>Y</td>
<td>Z</td>
</tr>
<tr>
<td>14</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>15</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>17</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>18</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Classical Cryptography

- eg.: Suppose $K=11$
  - Plaintext: student
  - Ciphertext: DEFOPZE

<table>
<thead>
<tr>
<th>plaintext</th>
<th>s</th>
<th>t</th>
<th>u</th>
<th>d</th>
<th>e</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>3</td>
<td>4</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>$+K$</td>
<td>3</td>
<td>4</td>
<td>5</td>
<td>14</td>
<td>15</td>
<td>25</td>
<td>4</td>
</tr>
<tr>
<td>ciphertext</td>
<td>D</td>
<td>E</td>
<td>F</td>
<td>O</td>
<td>P</td>
<td>Z</td>
<td>E</td>
</tr>
</tbody>
</table>
Classical Cryptography

<2> Substitution Cipher

- Cryptosystem 1.2: Substitution Cipher
  - $P = C = \mathbb{Z}_{26}$
  - $K$: all possible permutations of the 26 symbols
  - For each $\pi \in K$
    - $e_\pi(x) = \pi(x)$
    - $d_\pi(y) = \pi^{-1}(y)$

where $\pi^{-1}$ is the inverse permutation to $\pi$
**Classical Cryptography**

- **eg.**:

<table>
<thead>
<tr>
<th>$x$</th>
<th>a</th>
<th>b</th>
<th>C</th>
<th>d</th>
<th>e</th>
<th>f</th>
<th>g</th>
<th>h</th>
<th>i</th>
<th>j</th>
<th>k</th>
<th>l</th>
<th>m</th>
</tr>
</thead>
<tbody>
<tr>
<td>$e_{\pi}(x)$</td>
<td>X</td>
<td>N</td>
<td>Y</td>
<td>A</td>
<td>H</td>
<td>P</td>
<td>O</td>
<td>G</td>
<td>Z</td>
<td>Q</td>
<td>W</td>
<td>B</td>
<td>T</td>
</tr>
<tr>
<td>$x$</td>
<td>n</td>
<td>o</td>
<td>p</td>
<td>q</td>
<td>r</td>
<td>s</td>
<td>t</td>
<td>u</td>
<td>v</td>
<td>w</td>
<td>x</td>
<td>y</td>
<td>z</td>
</tr>
<tr>
<td>$e_{\pi}(x)$</td>
<td>S</td>
<td>F</td>
<td>L</td>
<td>R</td>
<td>C</td>
<td>V</td>
<td>M</td>
<td>U</td>
<td>E</td>
<td>K</td>
<td>J</td>
<td>D</td>
<td>I</td>
</tr>
</tbody>
</table>

- **Plaintext:** student
- **Ciphertext:** VMUSHSM
Classical Cryptography

- <3> Affine Cipher
  - Theorem 1.1: \( ax \equiv b \pmod{m} \) has a unique solution \( x \in \mathbb{Z}_m \) for every \( b \in \mathbb{Z}_m \) iff \( \gcd(a, m) = 1 \)
  - Definition 1.3: Suppose \( a \geq 1 \) and \( m \geq 2 \) are integers
    - \( a \) and \( m \) are relatively prime if \( \gcd(a, m) = 1 \)
    - \( \phi(m) \): the number of integers in \( \mathbb{Z}_m \) that are relatively prime to \( m \)
  - Theorem 1.2: Suppose
    \[
    m = \prod_{i=1}^{n} p_i^{e_i}
    \]
    \[
    \phi(m) = \prod_{i=1}^{n} (p_i^{e_i} - p_i^{e_i-1})
    \]
**Classical Cryptography**

- **Definition 1.4:** Suppose $a \in \mathbb{Z}_m$
  - $a^{-1} \mod m$: the multiplicative inverse of $a$ modulo $m$
  - $aa^{-1} \equiv a^{-1}a \equiv 1 \pmod{m}$

- **Cryptosystem 1.3: Affine Cipher**
  - $\mathcal{P} = \mathcal{C} = \mathbb{Z}_{26}$
  - $\mathcal{K} = \{(a,b) \in \mathbb{Z}_{26} \times \mathbb{Z}_{26} : \gcd(a,26) = 1\}$
  - For $K = (a,b) \in \mathcal{K}$; $x, y \in \mathbb{Z}_{26}$
    - $e_K(x) = (ax+b) \mod 26$
    - $d_K(y) = a^{-1}(y-b) \mod 26$
**Classical Cryptography**

- **e.g.:** Suppose $K=(7, 3)$
  - $7^{-1} \mod 26 = 15$
  - Plaintext: student
  - Ciphertext: ZGNYFQG

<table>
<thead>
<tr>
<th>plaintext</th>
<th>s</th>
<th>t</th>
<th>u</th>
<th>d</th>
<th>e</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>3</td>
<td>4</td>
<td>13</td>
<td>19</td>
</tr>
</tbody>
</table>

$e_K(x) = (7x + 3) \mod 26$

<table>
<thead>
<tr>
<th>ciphertext</th>
<th>Z</th>
<th>G</th>
<th>N</th>
<th>Y</th>
<th>F</th>
<th>Q</th>
<th>G</th>
</tr>
</thead>
</table>

$d_K(y) = 15(y - 3) \mod 26$
Classical Cryptography

- **<4> Vigenère Cipher**
  - **Cryptosystem 1.4: Vigenère Cipher**
    - **m**: a positive integer
    - **$\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m$$**
    - For a key $K=(k_1,k_2,...,k_m)$
      - $e_K(x_1,x_2,...,x_m) = (x_1+k_1,x_2+k_2,...,x_m+k_m)$
      - $d_K(y_1,y_2,...,y_m) = (y_1-k_1,y_2-k_2,...,y_m-k_m)$
Classical Cryptography

- e.g.: Suppose $m=4$ and $K=(2,8,15,7)$
  - Plaintext: student
  - Ciphertext: UBJKGVI

<table>
<thead>
<tr>
<th>plaintext</th>
<th>s</th>
<th>t</th>
<th>u</th>
<th>d</th>
<th>e</th>
<th>n</th>
<th>t</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>18</td>
<td>19</td>
<td>20</td>
<td>3</td>
<td>4</td>
<td>13</td>
<td>19</td>
</tr>
<tr>
<td>+K</td>
<td>2</td>
<td>8</td>
<td>15</td>
<td>7</td>
<td>2</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>ciphertext</td>
<td>20</td>
<td>1</td>
<td>9</td>
<td>10</td>
<td>6</td>
<td>21</td>
<td>8</td>
</tr>
</tbody>
</table>
Classical Cryptography

- <5> Hill Cipher
  - Definition 1.5: Suppose \( A = (a_{i,j}) \) is an \( m \times m \) matrix
    - \( A_{i,j} \): the matrix obtained from \( A \) by deleting the \( i \)th row and the \( j \)th column
    - \( \det A \): the determinant of \( A \)
      - \( m=1 \): \( \det A = a_{1,1} \)
      - \( m>1 \): for any fixed \( i \)
        \[
        \det A = \sum_{j=1}^{m} (-1)^{i+j} a_{i,j} \det A_{i,j}
        \]
    - \( A^* = (a^*_{i,j}) \): the adjoint matrix of \( A \)
      - \( a^*_{i,j} = (-1)^{i+j} \det A_{j,i} \)
Classical Cryptography

- Theorem 1.3: Suppose \( K \) is an \( m \times m \) invertible matrix over \( \mathbb{Z}_n \)
  - \( K^{-1} = (\det K)^{-1} K^* \)

  - e.g.:
    - \( K = \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} \)
    - \( K_{1,2} = 3 \)  \( \implies \begin{pmatrix} 11 & 8 \\ 3 & 7 \end{pmatrix} \)
    - \( K^* = \begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} \)
    - \( \det K = 11 \times 7 - 8 \times 3 \mod 26 = 1 \)
    - \( K^{-1} = (\det K)^{-1} K^* = \begin{pmatrix} 7 & 18 \\ 23 & 11 \end{pmatrix} \)
Classical Cryptography

- **Cryptosystem 1.5: Hill Cipher**
  - $m \geq 2$ is an integer
  - $P = C = (\mathbb{Z}_{26})^m$
  - $K = \{m \times m$ invertible matrices over $\mathbb{Z}_{26}\}$
  - For a key $K$
    - $e_K(x) = xK$
    - $d_K(y) = yK^{-1}$

  *where $K^{-1}$ is the inverse of $K$*
Classical Cryptography

- e.g.: 
  \[ K = \begin{pmatrix} 10 & 5 & 12 \\ 3 & 14 & 21 \\ 8 & 9 & 11 \end{pmatrix} , \quad K^{-1} = \begin{pmatrix} 21 & 15 & 17 \\ 23 & 2 & 16 \\ 25 & 4 & 3 \end{pmatrix} \]

- Plaintext: GOD  \( (6, 14, 3) \)
- Ciphertext: WTJ  \( (22, 19, 9) \)

\[
\begin{pmatrix} 6 & 14 & 3 \\ 10 & 5 & 12 \\ 8 & 9 & 11 \end{pmatrix} \begin{pmatrix} 3 & 14 & 21 \end{pmatrix} = \begin{pmatrix} 22 & 19 & 9 \end{pmatrix}
\]
Classical Cryptography

- <6> Permutation Cipher
  - Cryptosystem 1.6: Permutation Cipher
    - \( m \) is a positive integer
    - \( \mathcal{P} = C = (\mathbb{Z}_{26})^m \)
    - \( \mathcal{K} \) consist of all permutations of \( \{1,\ldots,m\} \)
    - For a key (a permutation) \( \pi \)
      - \( e_{\pi}(x_1,\ldots,x_m) = (x_{\pi(1)},\ldots,x_{\pi(m)}) \)
      - \( d_{\pi}(y_1,\ldots,y_m) = (y_{\pi^{-1}(1)},\ldots,y_{\pi^{-1}(m)}) \)

where \( \pi^{-1} \) is the inverse permutation to \( \pi \)
Classical Cryptography

- e.g.: Suppose $m=6$
  - Plaintext: CYBERFORMULA
  - Ciphertext: BRCFEYMLOAUR

\[
\begin{array}{ccccccc}
 x & 1 & 2 & 3 & 4 & 5 & 6 \\
\pi(x) & 3 & 5 & 1 & 6 & 4 & 2 \\
\end{array}
\]

<table>
<thead>
<tr>
<th>plaintext</th>
<th>C</th>
<th>Y</th>
<th>B</th>
<th>E</th>
<th>R</th>
<th>F</th>
<th>O</th>
<th>R</th>
<th>M</th>
<th>U</th>
<th>L</th>
<th>A</th>
</tr>
</thead>
<tbody>
<tr>
<td>ciphertext</td>
<td>B</td>
<td>R</td>
<td>C</td>
<td>F</td>
<td>E</td>
<td>Y</td>
<td>M</td>
<td>L</td>
<td>O</td>
<td>A</td>
<td>U</td>
<td>R</td>
</tr>
</tbody>
</table>
Classical Cryptography

[2] Cryptanalysis

Assumption: (Kerckhoffs' principle)

The opponent knows the cryptosystem being used

Attack models:

- ciphertext only attack
- known plaintext attack
- chosen plaintext attack
- chosen ciphertext attack
Classical Cryptography

- Statistical properties of the English language: (see Table 1.1)
  - E: probability about 0.120
  - T, A, O, I, N, S, H, R: between 0.06 and 0.09
  - D, L: 0.04
  - C, U, M, W, F, G, Y, P, B: between 0.015 and 0.028
  - V, K, J, X, Q, Z: 0.01

- Most common digrams:
  - TH, HE, IN, ER, AN, ND, ...

- Most common trigrams:
  - THE, ING, AND, END, ...
## Classical Cryptography

Table 1.1

<table>
<thead>
<tr>
<th>letter</th>
<th>probability</th>
<th>letter</th>
<th>probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>.082</td>
<td>N</td>
<td>.067</td>
</tr>
<tr>
<td>B</td>
<td>.015</td>
<td>O</td>
<td>.075</td>
</tr>
<tr>
<td>C</td>
<td>.028</td>
<td>P</td>
<td>.019</td>
</tr>
<tr>
<td>D</td>
<td>.043</td>
<td>Q</td>
<td>.001</td>
</tr>
<tr>
<td>E</td>
<td>.127</td>
<td>R</td>
<td>.060</td>
</tr>
<tr>
<td>F</td>
<td>.022</td>
<td>S</td>
<td>.063</td>
</tr>
<tr>
<td>G</td>
<td>.020</td>
<td>T</td>
<td>.091</td>
</tr>
<tr>
<td>H</td>
<td>.061</td>
<td>U</td>
<td>.028</td>
</tr>
<tr>
<td>I</td>
<td>.070</td>
<td>V</td>
<td>.010</td>
</tr>
<tr>
<td>J</td>
<td>.002</td>
<td>W</td>
<td>.023</td>
</tr>
<tr>
<td>K</td>
<td>.008</td>
<td>X</td>
<td>.001</td>
</tr>
<tr>
<td>L</td>
<td>.040</td>
<td>Y</td>
<td>.020</td>
</tr>
<tr>
<td>M</td>
<td>.024</td>
<td>Z</td>
<td>.001</td>
</tr>
</tbody>
</table>
Classical Cryptography

1 Cryptanalysis of the Affine Cipher

- Ciphertext obtained from an Affine Cipher:
  - FMXVEDKAPHFERBNDKRXRSREFMORUDSDKDVSH
  - VUFEDKAPRKDLYEVLRHHRH
- Frequency analysis: Table 1.2
- Most frequent ciphertext characters:
  - R: 8 occurrences
  - D: 7 occurrences
  - E, H, K: 5 occurrences
- We now guess the mapping and solve the equation $e_K(x)=ax+b \mod 26$
**Classical Cryptography**

<table>
<thead>
<tr>
<th>letter</th>
<th>frequency</th>
<th>letter</th>
<th>frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>2</td>
<td>N</td>
<td>1</td>
</tr>
<tr>
<td>B</td>
<td>1</td>
<td>O</td>
<td>1</td>
</tr>
<tr>
<td>C</td>
<td>0</td>
<td>P</td>
<td>2</td>
</tr>
<tr>
<td>D</td>
<td>7</td>
<td>Q</td>
<td>0</td>
</tr>
<tr>
<td>E</td>
<td>5</td>
<td>R</td>
<td>8</td>
</tr>
<tr>
<td>F</td>
<td>4</td>
<td>S</td>
<td>3</td>
</tr>
<tr>
<td>G</td>
<td>0</td>
<td>T</td>
<td>0</td>
</tr>
<tr>
<td>H</td>
<td>5</td>
<td>U</td>
<td>2</td>
</tr>
<tr>
<td>I</td>
<td>0</td>
<td>V</td>
<td>4</td>
</tr>
<tr>
<td>J</td>
<td>0</td>
<td>W</td>
<td>0</td>
</tr>
<tr>
<td>K</td>
<td>5</td>
<td>X</td>
<td>2</td>
</tr>
<tr>
<td>L</td>
<td>2</td>
<td>Y</td>
<td>1</td>
</tr>
<tr>
<td>M</td>
<td>2</td>
<td>Z</td>
<td>0</td>
</tr>
</tbody>
</table>
Classical Cryptography

- **Guess** e→R, t→D
  - \(e^K(4) = 17, e^K(19) = 3\)
  - \(a = 6, b = 19\)
  - **ILLEGAL** \((\gcd(a,26) > 1)\)

- **Guess** e→R, t→E
  - \(e^K(4) = 17, e^K(19) = 4\)
  - \(a = 13, b = 17\)
  - **ILLEGAL** \((\gcd(a,26) > 1)\)

- **Guess** e→R, t→H
  - \(e^K(4) = 17, e^K(19) = 7\)
  - \(a = 8, b = 11\)
  - **ILLEGAL** \((\gcd(a,26) > 1)\)
Classical Cryptography

- **Guess** \( e \rightarrow R, t \rightarrow K \)
  - \( e_K(4) = 17, e_K(19) = 10 \)
  - \( a = 3, b = 5 \)
  - LEGAL
  - \( d_K(y) = 9y - 19 \)

- **Plaintext:**
  - Algorithms are quite general definitions of arithmetic processes
Classical Cryptography

- **<2> Crytanalysis of the Substitution Cipher**
  - Ciphertext obtained from a Substitution Cipher
    - YIFQFMZRQFWQYVE CFMDZPCVMRZWNMDZVEJBTXCDU MJDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZ UNMXZNZUCDRJXYYYSMTMEYIFZWDYVYZVFZUMRZCR WNZDZJJXZWGCWRSRNMHDNCFQCHZJMSXJZWIEJYUCFWDJNZZDIR

- **Frequency analysis: Table 1.3**
  - Z occurs most: guess $d_K(Z) = e$
  - occur at least 10 times: C,D,F,J,M,R,Y
    - These are encryptions of \{t,a,o,i,n,s,h,r\}
  - But the frequencies do not vary enough to guess
## Classical Cryptography

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

Table 1.3
Classical Cryptography

- We now look at digrams: -Z or Z-
  - 4 times: DZ, ZW
    - Guess $d_k(W) = d$: ed $\rightarrow$ ZW
  - 3 times: NZ, ZU
    - Guess $d_k(N) = h$: he $\rightarrow$ NZ
  - We have ZRW: guess $d_k(R) = n$, end $\rightarrow$ ZRW
  - We have CRW: guess $d_k(C) = a$, and $\rightarrow$ CRW
- We have RNM, which decrypts to nh-
  - Suggest h- begins a word: M should be a vowel
  - We have CM: guess $dK(M) = i$
    (ai is more likely than ao)
| - | - | - | i | e | n | d | - | - | - | a | i | e | a |
| Y | I | F | Q | F | M | Z | R | W | Q | F | Y | V | E | C | F | M | D | Z | P | C |
| - | i | n | e | d | h | i | - | e | - | - | - | - | - | a | - - | - | i | - |
| V | M | R | Z | W | N | M | D | Z | V | E | J | B | T | X | C | D | D | U | M | J |
| h | - | - | - | - | i | e | a | i | e | a | - | - | - | a | - |
| N | D | I | F | E | F | M | D | Z | C | D | M | Q | Z | K | C | E | Y | F | C | J |
| i | n | h | a | d | - | a | - | e | n | - | - | a | - | e | h | i | e |
| M | Y | R | N | C | W | J | C | S | Z | R | E | X | C | H | Z | U | N | M | X | Z |
| h | e | a | n | - | - | - | - | i | n | - | i | - | - | - | e | d |
| N | Z | U | C | D | R | J | X | Y | Y | S | M | R | T | M | E | Y | I | F | Z | W |
| - | - | e | - | - | e | - | i | n | e | a | n | d | h | e | - | - |
| D | Y | V | Z | V | Y | F | Z | U | M | R | Z | C | R | W | N | Z | D | Z | J | J |
| - | e | d | - | a | - | i | n | h | i | - | - | h | a | i | - | - | a | - | e |
| X | Z | W | G | C | H | S | M | R | N | M | D | H | N | C | M | F | Q | C | H | Z |
| - | i | - | e | d | - | - | - | a | d | - | h | e | - | - | n |
Classical Cryptography

- We have DZ(4 times) and ZD(2 times)
  - Guess $d_k(D) \in \{r,s,t\}$
- Since o is a common letter
  - Guess $e_k(o) \in \{F,J,Y\}$
  - We have CFM and CJM: guess $d_k(Y) = o$
    (aoi is impossible)
- Guess NMD → his: $d_k(D) = s$
- Guess HNCMF → chair: $d_k(H) = c$, $d_k(F) = r$
- $d_k(J) = t$: the → JNZ
Classical Cryptography

- Now easy to determine the others
  - $d_K(I) = u$  \quad $d_K(Q) = f$
  - $d_K(V) = m$  \quad $d_K(E) = p$
  - $d_K(P) = x$  \quad $d_K(B) = y$
  - $d_K(T) = g$  \quad $d_K(X) = l$
  - $d_K(U) = w$  \quad $d_K(K) = v$
  - $d_K(S) = k$  \quad $d_K(G) = b$
<table>
<thead>
<tr>
<th>Our friend from Paris exa</th>
<th>Y IF Q FMZ RWQ FY VE CF MDZPC</th>
</tr>
</thead>
<tbody>
<tr>
<td>mined his empty glass wit</td>
<td>VMRZWNMDZVEJBTXCDCDUMJ</td>
</tr>
<tr>
<td>his surprise as if evaporat</td>
<td>NJDFEFMDZCDMQZKCEYFCJ</td>
</tr>
<tr>
<td>ion had taken place while</td>
<td>MYRNCDWJCSZREXCHZUNMXZ</td>
</tr>
<tr>
<td>he wasn't looking it poured</td>
<td>NZUCDRJXYYSMRTMEYIFZW</td>
</tr>
<tr>
<td>some more wine and he settt</td>
<td>DYZVYFZUMRZCRCRWNZDZJJ</td>
</tr>
<tr>
<td>led back in his chair face</td>
<td>XZWGCHSMRNMDHDHCNCMFQCHZ</td>
</tr>
<tr>
<td>tilted up towards the sun</td>
<td>JMXJZWIEJYUCFDWDJNZDIR</td>
</tr>
</tbody>
</table>
Classical Cryptography

- <3> Cryptanalysis of the Vigenère Cipher
  - Kasaski test (1863) (Find m only):
    - Search the ciphertext for pairs of identical segments (length at least 3)
    - Record the distance between the starting positions of the 2 segments
    - If we obtain several such distances \( \delta_1, \delta_2, \ldots \), we would conjecture that the key length \( m \) divides all of the \( \delta_i \)'s
    - \( m \) divides the gcd of the \( \delta_i \)'s
Classical Cryptography

- Friedman test (1925)
- Definition 1.7:

  - Suppose $x = x_1x_2...x_n$ is a string of $n$ alphabetic characters
  - *Index of coincidence of $x$, denoted $I_C(x)$: the probability that 2 random elements of $X$ are identical*
  - We denote the frequencies of A,B,..,Z in $x$ by $f_0,f_1,...,f_{25}$

  $$I_C(X) = \frac{\sum_{i=0}^{25} \binom{f_i}{2}}{\binom{n}{2}} = \frac{\sum_{i=0}^{25} f_i(f_i - 1)}{n(n-1)}$$
Classical Cryptography

- Using the expected probabilities in Table 1.1
  - \( I_c(X) \approx \sum_{i=0}^{25} p_i^2 = 0.065 \)

  \( p_0, \ldots, p_{25} \): the expected probability of A, ..., Z

- Suppose a ciphertext \( Y = y_1 y_2 \cdots y_n \)
- Define \( m \) substrings of \( Y_1, \ldots, Y_m \) of \( Y \)
  - \( Y_1 = y_1 y_{m+1} y_{2m+1} \cdots \)
  - \( Y_2 = y_2 y_{m+2} y_{2m+2} \cdots \)
  - \( \vdots \)
  - \( Y_m = y_m y_{2m} y_{3m} \cdots \)

- Each value \( I_c(Y_i) \) should be roughly equal to 0.065
Classical Cryptography

- If $m$ is not the keyword length
  - $Y_i$ will look much more random
  - A completely random string will have

$$I_C \approx 26 \left( \frac{1}{26} \right)^2 = \frac{1}{26} = 0.038$$
Classical Cryptography

- Ciphertext obtained from a Vigenere Cipher
  
  - CHREEVOAHMAERATBIAXXWTNXBEEOPHBSBQMQ
  - EQERBWRVXUOAKXAOSXXWEAHBWGJMMQMNKG
  - RFVGXWTRZXWIAKLXFPSKAUTEMNDCMGTSXMXB
  - TUIADNGMGPSRELXNJELXVRVPRTLHDNQWTWD
  - TYGBPHXTFALJHASVBFXNGLLCHRZBWELEKMSJIK
  - NBHWRJGNMGJSGLXFEYPHAGNRIEQJTIMAVLC
  - RREMNDGXLRRIMGNSNRWCHRQHAVETAQEBB
  - IPEEEWEVKAKOEWADREMXMTBHHCHRRTDNVRZC
  - HRCLQOHPWQAIIWXNRMGWIOIFKEE

- CHR occurs in 5 places: 1,166,236,276,286
- The distances from the 1st one: 165,235,275,285
- g.c.d. is 5: we guess m=5 (by Kasaski test)
Classical Cryptography

- We check the indices of coincidences:
  - $m=1$: $I_C(Y) = 0.045$
  - $m=2$: $I_C(Y_1) = 0.046$, $I_C(Y_2) = 0.041$
  - $m=3$: $I_C = 0.043$, $0.050$, $0.047$
  - $m=4$: $I_C = 0.042$, $0.039$, $0.046$, $0.040$
  - $m=5$: $I_C = 0.063$, $0.068$, $0.069$, $0.061$, $0.072$

- By Friedman test, $m=5$
Classical Cryptography

- Now we want to determine the key $K=(k_1, k_2, \ldots, k_m)$
- $f_0, f_1, \ldots, f_{25}$: the frequencies of $A, B, \ldots, Z$
- $n'=n/m$: the length of the string $Y_i$
- The probability distribution of the 26 letters in $Y_i$:
  \[
  \frac{f_0}{n'}, \ldots, \frac{f_{25}}{n'}
  \]
- $Y_i$ is obtained by shift encryption using a shift $k_i$
  \[\Rightarrow\] We hope that the shifted probability distribution would be close to $p_0, \ldots, p_{25}$

  \[
  \frac{f_{k_i}}{n'}, \ldots, \frac{f_{25+k_i}}{n'}
  \]
Classical Cryptography

- Define the quantity $M_g$: $M_g = \sum_{i=0}^{25} \frac{p_i f_{i+g}}{n'}$
  
  - If $g=k_i$, $M_g \approx \sum_{i=0}^{25} p_i^2 = 0.065$
  
  - If $g \neq k_i$, $M_g$ will smaller than 0.065

- Return to the previous example:
  - Computes the values $M_g$, for $1 \leq i \leq 5$ (Table 1.4)
  - For each $i$, look for a value of $M_g$ close to 0.065
  - From Table 1.4: $K=(9,0,13,4,19)$
  - The keyword is JANET
<table>
<thead>
<tr>
<th>i</th>
<th>Value of $M_g(Y_i)$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.35 0.31 0.36 0.37 0.35 0.39 0.28 0.28 0.48</td>
</tr>
<tr>
<td></td>
<td><strong>0.61</strong> 0.39 0.32 0.40 0.38 0.38 0.44 0.36 0.30</td>
</tr>
<tr>
<td></td>
<td>0.42 0.43 0.36 0.33 0.49 0.43 0.41 0.36</td>
</tr>
<tr>
<td>2</td>
<td>0.69 0.44 0.32 0.35 0.44 0.34 0.36 0.33 0.30</td>
</tr>
<tr>
<td></td>
<td>0.31 0.42 0.45 0.40 0.45 0.46 0.42 0.37 0.32</td>
</tr>
<tr>
<td></td>
<td>0.34 0.37 0.32 0.34 0.43 0.32 0.26 0.47</td>
</tr>
<tr>
<td>3</td>
<td>0.48 0.29 0.42 0.43 0.44 0.34 0.38 0.35 0.32</td>
</tr>
<tr>
<td></td>
<td>0.49 0.35 0.31 0.35 <strong>0.65</strong> 0.35 0.38 0.36 0.45</td>
</tr>
<tr>
<td></td>
<td>0.27 0.35 0.34 0.34 0.37 0.35 0.46 0.40</td>
</tr>
<tr>
<td>4</td>
<td>0.45 0.32 0.33 0.38 <strong>0.60</strong> 0.34 0.34 0.34 0.50</td>
</tr>
<tr>
<td></td>
<td>0.33 0.33 0.43 0.40 0.33 0.28 0.36 0.40 0.44</td>
</tr>
<tr>
<td></td>
<td>0.37 0.50 0.34 0.34 0.39 0.44 0.38 0.35</td>
</tr>
<tr>
<td>5</td>
<td>0.34 0.31 0.35 0.44 0.47 0.37 0.43 0.38 0.42</td>
</tr>
<tr>
<td></td>
<td>0.37 0.33 0.32 0.35 0.37 0.36 0.45 0.32 0.29</td>
</tr>
<tr>
<td></td>
<td><strong>0.72</strong> 0.36 0.27 0.30 0.48 0.36 0.37</td>
</tr>
</tbody>
</table>
Classical Cryptography

4 Cryptanalysis of the Hill Cipher

- Hill Cipher is difficult to break with a ciphertext-only attack

⇒ We use a known plaintext attack

- Suppose the unknown key is an $m \times m$ matrix and we have at least $m$ distinct plaintext-ciphertext pairs

\[ x_j = (x_{1,j}, x_{2,j}, \ldots, x_{m,j}) \]
\[ y_j = (y_{1,j}, y_{2,j}, \ldots, y_{m,j}) \]
\[ y_j = e_K(x_j), \text{ for } 1 \leq j \leq m \]
Classical Cryptography

- We define 2 \( m \times m \) matrices \( X = (x_{i,j}) \) and \( Y = (y_{i,j}) \)
  \( \Rightarrow Y = XK \)
  \( \Rightarrow K = X^{-1}Y \)

- e.g.: \( m = 2 \), plaintext: friday, ciphertext: PQCFKU
  - \( e_k(5,17) = (15,16) \)
  - \( e_k(8,3) = (2,5) \)
  - \( e_k(0,24) = (10,20) \)
Classical Cryptography

- e.g. (cont.)

\[
\begin{pmatrix}
15 & 16 \\
2 & 5
\end{pmatrix}
= \begin{pmatrix}
5 & 17 \\
8 & 3
\end{pmatrix}K
\]

\[
K = \begin{pmatrix}
5 & 17 \\
8 & 3
\end{pmatrix}^{-1} \begin{pmatrix}
15 & 16 \\
2 & 5
\end{pmatrix}
\]

\[
= \begin{pmatrix}
9 & 1 \\
2 & 15
\end{pmatrix} \begin{pmatrix}
15 & 16 \\
2 & 5
\end{pmatrix}
= \begin{pmatrix}
7 & 19 \\
8 & 3
\end{pmatrix}
\]