Classial Ciphers: Affine Block Ciphers

Çetin Kaya Koç

http://cs.ucsb.edu/~koc
koc@cs.ucsb.edu

Input/Output Alphabet and Encoding

- Input/output alphabet is $\{a, b, \ldots, z\}$ with encoding $\{0, 1, \ldots, 25\}$
- However, other encodings can also be used, for example, we can increase the input size by adding capital letters, punctuation symbols, etc
- In general, we will assume that our alphabet consists of m symbols, represented using the integers $\mathcal{Z}_m = \{0, 1, 2, \ldots, m-1\}$
- Furthermore, we will perform the addition and multiplication operations mod m
- The set and the operations together is called *the ring of integers modulo m*, represented as the triple $(\mathcal{Z}_m, +, \times)$

The Affine Block Cipher

- Encryption function:

$$v = \mathcal{A}\, u + w \pmod{m}$$

such that u and v are $d \times 1$ input (plaintext) and output (ciphertext) vectors, $\mathcal{A}$ is a fixed $d \times d$ key matrix and w is a $d \times 1$ fixed key vector

- Decryption function:

$$u = \mathcal{A}^{-1}(v - w) \pmod{m}$$

such that $\mathcal{A}^{-1}$ is the inverse of $\mathcal{A}$ in the ring $(\mathcal{Z}_m, +, \times)$

- All elements of these vectors and matrices are from $\mathcal{Z}_m$ and the arithmetic is performed in the ring $(\mathcal{Z}_m, +, \times)$, i.e., modulo m arithmetic

The Affine Block Cipher

- Encryption keys: $\mathcal{A}$ and w
- Decryption keys: $\mathcal{A}^{-1}$ and w
- Key space: The number of distinct invertible $\mathcal{A}$ matrices times the number of distinct w vectors
- Observation: The Hill Cipher is an Affine Block Cipher such that $\mathcal{A}$ is the Hill matrix, w is a zero vector, and $m = 26$

$$\begin{aligned} v &= \mathcal{A}\, u \pmod{26} \\ u &= \mathcal{A}^{-1}\, v \pmod{26} \end{aligned}$$

The Vigenère Cipher

- Another well known cipher is the Vigenère Cipher which was incorrectly attributed to Blaise de Vigenère (1523-1596), a French diplomat and cryptographer
- It seems that the Vigenère Cipher was reinvented several times!
- The Vigenère Cipher makes use of repeated applications of the Shift Cipher with different keys — it is a poly-alphabetic cipher
- The Vigenère Cipher is easy to understand and implement, and seems unbreakable to beginners, which explains its popularity!
- It has earned a special name: *le chiffre indéchiffrable*

The Vigenère Cipher - Informal Description

- Select a key word or key phrase: `herbalist`
- Write key word under the plaintext message and perform mod 26 addition on letter encodings in order to obtain the plaintext

```
physicists at ucsb are studying quantum entanglement
herbalisth er bali sth erbalist herbali stherbalisth
wlptinqkmz ek vcdj skl wkvdjqfz xyrotfu wgaeehlpuwga
```

- For example, to find "p" + "h" we add their encodings 15 and 7 modulo 26, and thus

$$15 + 7 = 22 \pmod{26}$$

obtain 22 which is the encoding of "w"

- The key word length (in our example $d = 9$) is the dimension of the Affine Block Cipher representing the Vigenère Cipher
- The key word itself is represented as $d \times 1$ vector with elements from $\mathcal{Z}_{26}$
- In our example, `herbalist` implies $w = [7, 4, 17, 1, 0, 11, 8, 18, 19]^T$
- The encryption function is given simply as $v = u + w \pmod{26}$ where u and v are the plaintext and ciphertext vectors of dimension 9×1
- In other words, the Vigenère Cipher is an Affine Block Cipher with $\mathcal{A} = I$, the unit matrix, that is $v = \mathcal{A}v + w = v + w \pmod{26}$
- The decryption function is obtained as $u = v - w \pmod{26}$

The Vigenère Cipher - Affine Block Cipher

- As an example, let us obtain the encryption of the plaintext `"physicist"` which is encoded as $u = [15, 7, 24, 18, 8, 2, 8, 18, 19]^T$
- Since $w = [7, 4, 17, 1, 0, 11, 9, 18, 19]^T$, we obtain the ciphertext

$$v = u + w = \begin{bmatrix} 15 \\ 7 \\ 24 \\ 18 \\ 8 \\ 2 \\ 8 \\ 18 \\ 19 \end{bmatrix} + \begin{bmatrix} 7 \\ 4 \\ 17 \\ 1 \\ 0 \\ 11 \\ 8 \\ 18 \\ 19 \end{bmatrix} = \begin{bmatrix} 22 \\ 11 \\ 15 \\ 19 \\ 8 \\ 13 \\ 17 \\ 10 \\ 12 \end{bmatrix} = \begin{bmatrix} \texttt{"w"} \\ \texttt{"l"} \\ \texttt{"p"} \\ \texttt{"t"} \\ \texttt{"i"} \\ \texttt{"n"} \\ \texttt{"q"} \\ \texttt{"k"} \\ \texttt{"m"} \end{bmatrix}$$

Known (or Chosen) Text Analysis

- Now we show how to obtain the key ($\mathcal{A}$ and w) of an Affine Block Cipher using a set of known or chosen texts
- Consider the encryption function of the d-dimensional Affine Block Cipher:

$$v = \mathcal{A}\, u + w \pmod{m}$$

such that u, v, w are $d \times 1$ vectors and $\mathcal{A}$ is a $d \times d$ matrix

- Assume that we have $d + 1$ pairs of (known or chosen) plaintext and ciphertext vectors:

$$(u_i, v_i) \quad \text{for} \quad i = 0, 1, 2, \ldots, d$$

- Since each vector has d elements, this means we have $d(d + 1)$ plaintext and ciphertext letters

Known (or Chosen) Text Analysis

- This means each pair (u_i, v_i) satisfies the equation

$$v_i = \mathcal{A}\, u_i + w \pmod{m}$$

for $i = 0, 1, 2, \ldots, d$, and particularly, $v_0 = \mathcal{A}\, u_0 + w \pmod{m}$

- This implies

$$\begin{aligned} v_i - v_0 &= \mathcal{A}\, u_i + w - (\mathcal{A}\, u_0 + w) \pmod{m} \\ &= \mathcal{A}\, u_i - \mathcal{A}\, u_0 \pmod{m} \\ &= \mathcal{A}(u_i - u_0) \pmod{m} \end{aligned}$$

where the vector $(u_i - u_0)$ is of dimension $d \times 1$

Known (or Chosen) Text Analysis

- Assemble the $d \times 1$ column vectors $(u_i - u_0)$ and $(v_i - v_0)$ into respective matrices of dimension $d \times d$ as

$$\begin{aligned} \mathcal{U} &= [u_1 - u_0,\ u_2 - u_0,\ u_3 - u_0,\ \cdots, u_d - u_0] \\ \mathcal{V} &= [v_1 - v_0,\ v_2 - v_0,\ v_3 - v_0,\ \cdots, v_d - v_0] \end{aligned}$$

- This way we can write all d equations as follows:

$$\mathcal{V} = \mathcal{A}\mathcal{U} \pmod{m}$$

- By finding the inverse of the $d \times d$ matrix $\mathcal{U}$, and multiplying both sides of the above equation, we find

$$\mathcal{A} = \mathcal{V}\mathcal{U}^{-1} \pmod{m}$$

Known (or Chosen) Text Analysis

- Once we have the key matrix $\mathcal{A}$, we easily obtain the key vector w as

$$w = v_0 - \mathcal{A}\, u_0 \pmod{m}$$

- This analysis requires $d+1$ knows plaintext and ciphertext vectors: (u_i, v_i) for $i = 0, 1, 2, \ldots, d$ — since each vector has d entries, we need $d(d+1)$ plaintext and ciphertext letters
- Considering that d is the dimension of the system, and is probably a small integer, this attack is very powerful
- For example, the 5-dimensional Hill cipher had $10^{115.8}$ keys, making the exhaustive key search an impossible task — however, we can break it using only $5 \cdot 6 = 30$ plaintext and ciphertext pairs