Encoding and decoding secret messages Hill ciphers

Iulia Crivat

CALIFORNIA STATE UNIVERSITY LONG BEACH

INTRODUCTION

“Cryptography” (from the Greek *kryptos,* “hidden, secret” and *grapho, “I write”)* is the practice and study of hiding information. In the language of cryptography, codes= *ciphers,* encoded messages=*plaintext,* coded messages= *ciphertext*

Letter-by-letter substitution ciphers preserve the frequencies of individual letters, making it relatively easy to break the code by statistical methods. Polygraphic ciphers, by contrast, in which each list of *n* consecutive letters of the plaintext (an n-graph) is replaced by another n-graph according to some key, can be more challenging to break. The first systematic yet simple polygraphic ciphers using more than two letters per group are the Hill ciphers, first described by Lester Hill in 1929.

METHODS

Each plaintext and ciphertext letter except Z is assigned the numerical value that specifies his position in the standard alphabet. Z is assigned a value of 0

A	B	C	D	E	F	G	H	I	J	K	L	M
1	2	3	4	5	6	7	8	9	10	11	12	13

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
14	15	16	17	18	19	20	21	22	23	24	25	0

1. *Enciphering*

The process of converting from plaintext to ciphertext is called *enciphering.*

1.a) Choose an invertible modulo 26 matrix with integer entries:

1.b) Group successive plaintext letters into sets of n letters, adding a “dummy letter” to fill out the last pair (if necessary), then replace each plaintext letter by its numerical value.

1.c) Successively convert each plaintext pair $p_1p_2\ldots p_n$ into a column vector and form the product Ap=c. We will call *p* a plaintext vector and *c* the corresponding ciphertext vector.

1.d) Convert each ciphertext vector into its alphabetic equivalent.

2. *Deciphering*

The reverse process of converting from ciphertext to plaintext is called *deciphering*.

In the case of a hill cipher, decipherment uses the inverse (mod 26) of the enciphering matrix.

We saw that Ap=c, where A is the invertible modulo 26 enciphering matrix, *p* is the plaintext vector and *c* is the corresponding ciphertext vector. Because A is invertible, each plaintext vector can be recovered from the corresponding ciphertext vector by multiplying it on the left by A^{-1} (mod 26).

$p=A^{-1}$ c (mod 26)

3. *Modular Arithmetic*

Because of its importance in cryptography, we will focus for a few moments on some of the main ideas in the area of modular arithmetic.

Definition:

If m is a positive integer and a and b are any integers, then we say that a is equivalent to b modulo m, written

a=b(mod m)

if a-b is an integer multiple of m.

Example:

14=5 (mod 9) 44=4 (mod 8)

Every integer a is equivalent, modulo m, to one of the integers 0,1,2…m-1. This integer is called the *residue of a modulo m* and Z= {0,1,2…m-1} is called *the set of residues modulo m*.

Definition

If a is a number in Z_m, then a number a^{-1} is called a reciprocal or multiplicative inverse of a modulo m if

$aa^{-1} = a^{-1} a = 1$(mod m).

Theorem

A square matrix A with entries in Z_m is invertible modulo m if and only if the residue of det(A) modulo m has a reciprocal modulo m

Corollary

A square matrix A with entries in Z_m is invertible modulo m if and only if m and the residue of det(A) modulo m have no common prime factors.

Corollary

A square matrix A with entries in Z_m is invertible modulo 26 if and only if the residue of det(A) is not divisible by 2 or 13.

Reciprocals Modulo 26

a	1	3	5	7	9	11	15	17	19	21	23	25
a^{-1}	1	9	21	15	3	19	7	23	11	5	17	25

RESULTS

In order to show the results of the Hill’s coding/ decoding process we will use two examples:

Example 1 : Use the matrix $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

to obtain the hill 2-cipher for the plaintext message ATTACK

<u>Solution:</u>

First we group the plaintext into pairs. We obtain

AT TA CK

Then we replace each letter with its numerical value

1 20 20 1 3 11

To encipher the pairs AT, TA and CK we form the matrix products

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 1 \\ 20 \end{bmatrix} = \begin{bmatrix} 61 \\ 22 \end{bmatrix} = \begin{bmatrix} 9 \\ 22 \end{bmatrix} \qquad \begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 20 \\ 1 \end{bmatrix} = \begin{bmatrix} 23 \\ 41 \end{bmatrix} = \begin{bmatrix} 23 \\ 15 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}\begin{bmatrix} 3 \\ 11 \end{bmatrix} = \begin{bmatrix} 36 \\ 17 \end{bmatrix} = \begin{bmatrix} 10 \\ 17 \end{bmatrix}$$

These correspond to the ciphertext pairs IV, WO, JQ.

So, the ciphertext message is IVWOJQ.

Example 2 : Decode the Hill 2-cipher IVWOJQ which was enciphered by the matrix A= $\begin{bmatrix} 1 & 3 \\ 2 & 1 \end{bmatrix}$

First we find the inverse of A modulo 26.

det(A)= 1- 6 = -5 = 21 (mod 26)

From table 21^{-1} =5 (mod 26)

$$A^{-1} = 21\begin{bmatrix} 1 & -3 \\ -2 & 1 \end{bmatrix} = \begin{bmatrix} 5 & -15 \\ -10 & 5 \end{bmatrix} = \begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix}$$

The numerical equivalent of the ciphertext is

9 22 23 15 10 17

$$\begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix}\begin{bmatrix} 9 \\ 22 \end{bmatrix} = \begin{bmatrix} 287 \\ 254 \end{bmatrix} = \begin{bmatrix} 1 \\ 20 \end{bmatrix} \qquad \begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix}\begin{bmatrix} 23 \\ 15 \end{bmatrix} = \begin{bmatrix} 280 \\ 443 \end{bmatrix} = \begin{bmatrix} 20 \\ 3 \end{bmatrix}$$

$$\begin{bmatrix} 5 & 11 \\ 16 & 5 \end{bmatrix}\begin{bmatrix} 10 \\ 17 \end{bmatrix} = \begin{bmatrix} 237 \\ 245 \end{bmatrix} = \begin{bmatrix} 3 \\ 11 \end{bmatrix}$$

The alphabet equivalents of these vectors are AT TA CK which yields the message ATTACK.

DISCUSSIONS

Unfortunately, the basic Hill cipher is vulnerable because is completely linear. An opponent who intercepts n^2 plaintext/ ciphertext pairs can set up a linear system that can be solved.

CONCLUSIONS

Even if Hill ciphers hardly have the currency of secret-key block ciphers, they do have interest at least for historical reasons: they constitute the first general method for successfully applying linear algebra to polygraphic ciphers.

REFERENCES

Howard Anton and Chris Rorres- *Elementary Linear Algebra: Applications Version*

Lester S. Hill- *Cryptography in an algebraic alphabet,* Amer. Math. Monthly 36 (1929)