Criptografía Clásica

Curso 2019
Laboratorio 1

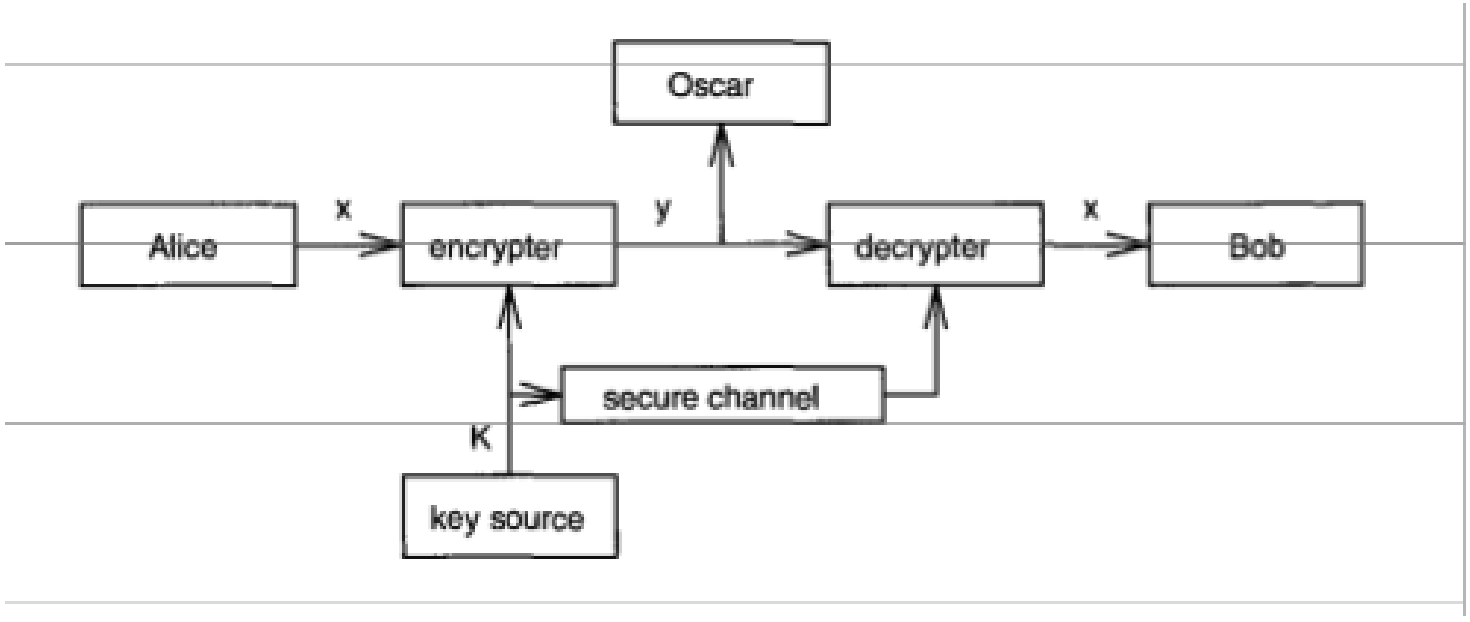
Instituto de Computación
Facultad de Ingeniería
Universidad de la República

Contenido

Objectivo Fundamental de la Criptografía

- es permitir a dos personas (Alcie y Bob) comunicarse por un canal inseguro, en el sentido que un adversario (Oscar) no puede entender lo que se está diciendo

Objetivo Fundamental de la Criptografía



Criptosistema

Definition 1.1: A *cryptosystem* is a five-tuple $(\mathcal{P}, \mathcal{C}, \mathcal{K}, \mathcal{E}, \mathcal{D})$, where the following conditions are satisfied:

1. $\mathcal{P}$ is a finite set of possible *plaintexts*;
2. $\mathcal{C}$ is a finite set of possible *ciphertexts*;
3. $\mathcal{K}$, the *keyspace*, is a finite set of possible *keys*;
4. For each $K \in \mathcal{K}$, there is an *encryption rule* $e_K \in \mathcal{E}$ and a corresponding *decryption rule* $d_K \in \mathcal{D}$. Each $e_K : \mathcal{P} \to \mathcal{C}$ and $d_K : \mathcal{C} \to \mathcal{P}$ are functions such that $d_K(e_K(x)) = x$ for every plaintext element $x \in \mathcal{P}$.

Shift Cipher - Aritmetica Modular

Definition 1.2: Suppose a and b are integers, and m is a positive integer. Then we write $a \equiv b \pmod{m}$ if m divides $b-a$. The phrase $a \equiv b \pmod{m}$ is called a *congruence*, and it is read as "a is *congruent* to b modulo m." The integer m is called the *modulus*.

Shift Cipher - Criptosistema

Cryptosystem 1.1: *Shift Cipher*

Let $\mathcal{P} = \mathcal{C} = \mathcal{K} = \mathbb{Z}_{26}$. For $0 \leq K \leq 25$, define

$$e_K(x) = (x + K) \bmod 26$$

and

$$d_K(y) = (y - K) \bmod 26$$

$(x, y \in \mathbb{Z}_{26})$.

Shift Cipher - Ejemplo

A	B	C	D	E	F	G	H	I	J	K	L	M
0	1	2	3	4	5	6	7	8	9	10	11	12

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
13	14	15	16	17	18	19	20	21	22	23	24	25

A small example will illustrate.

Example 1.1 Suppose the key for a *Shift Cipher* is $K = 11$, and the plaintext is

`wewillmeetatmidnight.`

Shift Cipher - Ejemplo

We first convert the plaintext to a sequence of integers using the specified correspondence, obtaining the following:

22	4	22	8	11	11	12	4	4	19
0	19	12	8	3	13	8	6	7	19

Next, we add 11 to each value, reducing each sum modulo 26:

7	15	7	19	22	22	23	15	15	4
11	4	23	19	14	24	19	17	18	4

Finally, we convert the sequence of integers to alphabetic characters, obtaining the ciphertext:

HPHTWWXPPELEXTOYTRSE.

Shift Cipher - Búsqueda Exhaustiva de Clave

JBCRCLQRWCRVNBJENBWRWN,

we successively try the decryption keys d_0, d_1, etc. The following is obtained:

jbcrclqrwcrvnbjenbwrwn
iabqbkpqvbqumaidmavqvm
hzapajopuaptlzhclzupul
gyzozinotzoskygbkytotk
fxynyhmnsynrjxfajxsnsj
ewxmxglmrxmqiweziwrmri
dvwlwfklqwlphvdyhvqlqh
cuvkvejkpvkogucxgupkpg
btujudijoujnftbwftojof
astitchintimesavesnine

Cifrado de Sustitución

Cryptosystem 1.2: *Substitution Cipher*

Let $\mathcal{P} = \mathcal{C} = \mathbb{Z}_{26}$. $\mathcal{K}$ consists of all possible permutations of the 26 symbols $0, 1, \ldots, 25$. For each permutation $\pi \in \mathcal{K}$, define

$$e_\pi(x) = \pi(x),$$

and define

$$d_\pi(y) = \pi^{-1}(y),$$

where π^{-1} is the inverse permutation to π.

Cifrado de Sustitución - Ejemplo

a	b	c	d	e	f	g	h	i	j	k	l	m
X	N	Y	A	H	P	O	G	Z	Q	W	B	T

n	o	p	q	r	s	t	u	v	w	x	y	z
S	F	L	R	C	V	M	U	E	K	J	D	I

Thus, $e_\pi(a) = X$, $e_\pi(b) = N$, etc. The decryption function is the inverse permutation. This is formed by writing the second lines first, and then sorting in alphabetical order. The following is obtained:

A	B	C	D	E	F	G	H	I	J	K	L	M
d	l	r	y	v	o	h	e	z	x	w	p	t

N	O	P	Q	R	S	T	U	V	W	X	Y	Z
b	g	f	j	q	n	m	u	s	k	a	c	i

Hence, $d_\pi(A) = d$, $d_\pi(B) = l$, etc.

Cifrado de Sustitución - Ejercicio

As an exercise, the reader might decrypt the following ciphertext using this decryption function:

MGZVYZLGHCMHJMYXSSFMNHAHYCDLMHA

- cantidad de permutaciones posibles distintas?
- criptoanalizable por otros métodos

Cifrado Affine

$$e(x) = (ax + b) \bmod 26,$$

$a, b \in \mathbb{Z}_{26}$. These functions are called *affine functions*, hence the name *Affine Cipher*. (Observe that when $a = 1$, we have a *Shift Cipher*.)

- cuando la función es inyectiva?

$$ax + b \equiv y \pmod{26}$$

to have a unique solution for x. This congruence is equivalent to

$$ax \equiv y - b \pmod{26}.$$

Cifrado Affine

We claim that this congruence has a unique solution for every y if and only if $\gcd(a, 26) = 1$ (where the gcd function denotes the greatest common divisor of its arguments). First, suppose that $\gcd(a, 26) = d > 1$. Then the congruence $ax \equiv 0 \pmod{26}$ has (at least) two distinct solutions in $\mathbb{Z}_{26}$, namely $x = 0$ and $x = 26/d$. In this case $e(x) = (ax + b) \bmod 26$ is not an injective function and hence not a valid encryption function.

For example, since $\gcd(4, 26) = 2$, it follows that $4x + 7$ is not a valid encryption function: x and $x + 13$ will encrypt to the same value, for any $x \in \mathbb{Z}_{26}$.

Let's next suppose that $\gcd(a, 26) = 1$. Suppose for some x_1 and x_2 that

$$ax_1 \equiv ax_2 \pmod{26}.$$

Then

$$a(x_1 - x_2) \equiv 0 \pmod{26},$$

Cifrado Affine

and thus

$$26 \mid a(x_1 - x_2).$$

We now make use of a fundamental property of integer division: if $\gcd(a, b) = 1$ and $a \mid bc$, then $a \mid c$. Since $26 \mid a(x_1 - x_2)$ and $\gcd(a, 26) = 1$, we must therefore have that

$$26 \mid (x_1 - x_2),$$

i.e., $x_1 \equiv x_2 \pmod{26}$.

Cifrado Affine

- cuantas claves posibles tiene el Cifrado Affine modulo 26?

Cifrado Affine

THEOREM 1.1 *The congruence* $ax \equiv b \pmod{m}$ *has a unique solution* $x \in \mathbb{Z}_m$ *for every* $b \in \mathbb{Z}_m$ *if and only if* $\gcd(a, m) = 1$.

Since $26 = 2 \times 13$, the values of $a \in \mathbb{Z}_{26}$ such that $\gcd(a, 26) = 1$ are $a = 1$, 3, 5, 7, 9, 11, 15, 17, 19, 21, 23, and 25. The parameter b can be any element in $\mathbb{Z}_{26}$. Hence the *Affine Cipher* has $12 \times 26 = 312$ possible keys. (Of course, this is much too small to be secure.)

Let's now consider the general setting where the modulus is m. We need another definition from number theory.

Definition 1.3: Suppose $a \geq 1$ and $m \geq 2$ are integers. If $\gcd(a, m) = 1$, then we say that a and m are *relatively prime*. The number of integers in $\mathbb{Z}_m$ that are relatively prime to m is often denoted by $\phi(m)$ (this function is called the *Euler phi-function*).

Cifrado Affine

THEOREM 1.2 *Suppose*

$$m = \prod_{i=1}^{n} p_i^{e_i},$$

where the p_i's are distinct primes and $e_i > 0$, $1 \leq i \leq n$. Then

$$\phi(m) = \prod_{i=1}^{n} (p_i^{e_i} - p_i^{e_i - 1}).$$

Cifrado Affine

- cuál es la función de desencriptado?

Cifrado Affine

Definition 1.4: Suppose $a \in \mathbb{Z}_m$. The *multiplicative inverse* of a modulo m, denoted $a^{-1} \bmod m$, is an element $a' \in \mathbb{Z}_m$ such that $aa' \equiv a'a \equiv 1 \pmod{m}$. If m is fixed, we sometimes write a^{-1} for $a^{-1} \bmod m$.

By similar arguments to those used above, it can be shown that a has a multiplicative inverse modulo m if and only if $\gcd(a, m) = 1$; and if a multiplicative inverse exists, it is unique modulo m. Also, observe that if $b = a^{-1}$, then $a = b^{-1}$.

Cifrado Affine - Criptosistema

Cryptosystem 1.3: *Affine Cipher*

Let $\mathcal{P} = \mathcal{C} = \mathbb{Z}_{26}$ and let

$$\mathcal{K} = \{(a, b) \in \mathbb{Z}_{26} \times \mathbb{Z}_{26} : \gcd(a, 26) = 1\}.$$

For $K = (a, b) \in \mathcal{K}$, define

$$e_K(x) = (ax + b) \bmod 26$$

and

$$d_K(y) = a^{-1}(y - b) \bmod 26$$

$(x, y \in \mathbb{Z}_{26})$.

Cifrado Vigenere

Using the correspondence $A \leftrightarrow 0, B \leftrightarrow 1, \ldots, Z \leftrightarrow 25$ described earlier, we can associate each key K with an alphabetic string of length m, called a *keyword*. The *Vigenère Cipher* encrypts m alphabetic characters at a time: each plaintext element is equivalent to m alphabetic characters.

Let's do a small example.

Example 1.4 Suppose $m = 6$ and the keyword is $CIPHER$. This corresponds to the numerical equivalent $K = (2, 8, 15, 7, 4, 17)$. Suppose the plaintext is the string

`thiscryptosystemisnotsecure`.

We convert the plaintext elements to residues modulo 26, write them in groups of six, and then "add" the keyword modulo 26, as follows:

19	7	8	18	2	17	24	15	19	14	18	24
2	8	15	7	4	17	2	8	15	7	4	17
21	15	23	25	6	8	0	23	8	21	22	15

Cifrado Vigenere - Criptosistema

Cryptosystem 1.4: *Vigenère Cipher*

Let m be a positive integer. Define $\mathcal{P} = \mathcal{C} = \mathcal{K} = (\mathbb{Z}_{26})^m$. For a key $K = (k_1, k_2, \ldots, k_m)$, we define

$$e_K(x_1, x_2, \ldots, x_m) = (x_1 + k_1, x_2 + k_2, \ldots, x_m + k_m)$$

and

$$d_K(y_1, y_2, \ldots, y_m) = (y_1 - k_1, y_2 - k_2, \ldots, y_m - k_m),$$

where all operations are performed in $\mathbb{Z}_{26}$.

Cifrado Permutación

All of the cryptosystems we have discussed so far involve substitution: plaintext characters are replaced by different ciphertext characters. The idea of a permutation cipher is to keep the plaintext characters unchanged, but to alter their positions by rearranging them using a permutation.

A *permutation* of a finite set X is a bijective function $\pi : X \to X$. In other words, the function π is one-to-one (injective) and onto (*surjective*). It follows that, for every $x \in X$, there is a unique element $x' \in X$ such that $\pi(x') = x$. This allows us to define the *inverse permutation*, $\pi^{-1} : X \to X$ by the rule

$$\pi^{-1}(x) = x' \quad \text{if and only if} \quad \pi(x') = x.$$

Then π^{-1} is also a permutation of X.

Cifrado Permutación - Ejemplo

Example 1.7 Suppose $m = 6$ and the key is the following permutation π:

x	1	2	3	4	5	6
$\pi(x)$	3	5	1	6	4	2

Note that the first row of the above diagram lists the values of x, $1 \leq x \leq 6$, and the second row lists the corresponding values of $\pi(x)$. Then the inverse permuta-

Cifrado Permutación - Ejemplo

tion π^{-1} can be constructed by interchanging the two rows, and rearranging the columns so that the first row is in increasing order. Carrying out these operations, we see that the permutation π^{-1} is the following:

x	1	2	3	4	5	6
$\pi^{-1}(x)$	3	6	1	5	2	4

Now, suppose we are given the plaintext

`shesellsseashellsbytheseashore.`

We first partition the plaintext into groups of six letters:

`shesel` | `lsseas` | `hellsb` | `ythese` | `ashore`

Now each group of six letters is rearranged according to the permutation π, yielding the following:

`EESLSH` | `SALSES` | `LSHBLE` | `HSYEET` | `HRAEOS`

Cifrado Permutación - Criptosistema

Cryptosystem 1.6: *Permutation Cipher*

Let m be a positive integer. Let $\mathcal{P} = \mathcal{C} = (\mathbb{Z}_{26})^m$ and let $\mathcal{K}$ consist of all permutations of $\{1, \ldots, m\}$. For a key (i.e., a permutation) π, we define

$$e_\pi(x_1, \ldots, x_m) = (x_{\pi(1)}, \ldots, x_{\pi(m)})$$

and

$$d_\pi(y_1, \ldots, y_m) = (y_{\pi^{-1}(1)}, \ldots, y_{\pi^{-1}(m)}),$$

where π^{-1} is the inverse permutation to π.

Criptoanálisis

ciphertext only attack

The opponent possesses a string of ciphertext, **y**.

known plaintext attack

The opponent possesses a string of plaintext, **x**, and the corresponding ciphertext, **y**.

chosen plaintext attack

The opponent has obtained temporary access to the encryption machinery. Hence he can choose a plaintext string, **x**, and construct the corresponding ciphertext string, **y**.

chosen ciphertext attack

The opponent has obtained temporary access to the decryption machinery. Hence he can choose a ciphertext string, **y**, and construct the corresponding plaintext string, **x**.

Criptoanálisis

TABLE 1.1
Probabilities of occurrence of the 26 letters

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

Criptoanálisis

1. E, having probability about 0.120
2. T, A, O, I, N, S, H, R, each having probability between 0.06 and 0.09
3. D, L, each having probability around 0.04
4. $C, U, M, W, F, G, Y, P, B$, each having probability between 0.015 and 0.028
5. V, K, J, X, Q, Z, each having probability less than 0.01.

It is also useful to consider sequences of two or three consecutive letters, called *digrams* and *trigrams*, respectively. The 30 most common digrams are (in decreasing order):

$$TH, HE, IN, ER, AN, RE, ED, ON, ES, ST,$$
$$EN, AT, TO, NT, HA, ND, OU, EA, NG, AS,$$
$$OR, TI, IS, ET, IT, AR, TE, SE, HI, OF.$$

The twelve most common trigrams are:

$$THE, ING, AND, HER, ERE, ENT,$$
$$THA, NTH, WAS, ETH, FOR, DTH.$$

Criptoanálisis de Affine

Example 1.10 Ciphertext obtained from an *Affine Cipher*

FMXVEDKAPHFERBNDKRXRSREFMORUDSDKDVSHVUFEDK
APRKDLYEVLRHHRH

Criptoanálisis de Affine

TABLE 1.2
Frequency of occurrence of the 26 ciphertext letters

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

Criptoanálisis de Affine

There are only 57 characters of ciphertext, but this is usually sufficient to cryptanalyze an *Affine Cipher*. The most frequent ciphertext characters are: R (8 occurrences), D (7 occurrences), E, H, K (5 occurrences each), and F, S, V (4 occurrences each). As a first guess, we might hypothesize that R is the encryption of e and D is the encryption of t, since e and t are (respectively) the two most common letters. Expressed numerically, we have $e_K(4) = 17$ and $e_K(19) = 3$. Recall that $e_K(x) = ax + b$, where a and b are unknowns. So we get two linear equations in two unknowns:

$$4a + b = 17$$

$$19a + b = 3.$$

This system has the unique solution $a = 6$, $b = 19$ (in $\mathbb{Z}_{26}$). But this is an illegal key, since $\gcd(a, 26) = 2 > 1$. So our hypothesis must be incorrect.

Criptoanálisis de Affine

next possibility, that R is the encryption of e and H is the encryption of t. This yields $a = 8$, again impossible. Continuing, we suppose that R is the encryption of e and K is the encryption of t. This produces $a = 3$, $b = 5$, which is at least a legal key. It remains to compute the decryption function corresponding to $K = (3, 5)$, and then to decrypt the ciphertext to see if we get a meaningful string of English, or nonsense. This will confirm the validity of $(3, 5)$.

If we perform these operations, we obtain $d_K(y) = 9y - 19$ and the given ciphertext decrypts to yield:

```
algorithmsarequitegeneraldefinitionsofarit
hmeticprocesses
```

Criptoanálisis de Sustitución

Example 1.11 Ciphertext obtained from a *Substitution Cipher*

```
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

The frequency analysis of this ciphertext is given in Table 1.3.

Since Z occurs significantly more often than any other ciphertext character, we might conjecture that $d_K(Z) = e$. The remaining ciphertext characters that occur at least ten times (each) are C, D, F, J, M, R, Y. We might expect that these letters are encryptions of (a subset of) t, a, o, i, n, s, h, r, but the frequencies really do not vary enough to tell us what the correspondence might be.

At this stage we might look at digrams, especially those of the form $-Z$ or $Z-$, since we conjecture that Z decrypts to e. We find that the most common digrams of this type are DZ and ZW (four times each); NZ and ZU (three times each); and RZ, HZ, XZ, FZ, ZR, ZV, ZC, ZD, and ZJ (twice each). Since ZW occurs four times and WZ not at all, and W occurs less often than many other characters, we might guess that $d_K(W) = d$. Since DZ occurs four times and

Criptoanálisis de Sustitución

TABLE 1.3
Frequency of occurrence of the 26 ciphertext letters

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

Criptoanálisis de Sustitución

ZD occurs twice, we would think that $d_K(D) \in \{r, s, t\}$, but it is not clear which of the three possibilities is the correct one.

If we proceed on the assumption that $d_K(Z) = e$ and $d_K(W) = d$, we might look back at the ciphertext and notice that we have ZRW occurring near the beginning of the ciphertext, and RW occurs again later on. Since R occurs frequently in the ciphertext and nd is a common digram, we might try $d_K(R) = n$ as the most likely possibility.

At this point, we have the following:

```
------end---------e----ned---e------------
YIFQFMZRWQFYVECFMDZPCVMRZWNMDZVEJBTXCDDUMJ

--------e----e----------n--d---en----e----e
NDIFEFMDZCDMQZKCEYFCJMYRNCWJCSZREXCHZUNMXZ

-e---n------n------ed---e---e--ne-nd-e-e--
NZUCDRJXYYSMRTMEYIFZWDYVZVYFZUMRZCRWNZDZJJ

-ed-----n----------e----ed-------d---e--n
XZWGCHSMRNMDHNCMFQCHZJMXJZWIEJYUCFWDJNZDIR
```

Criptoanálisis de Vigenere

- Primero debemos determinar el largo de la clave m : Kasiski e Indice de Coincidencia Mutua
- Realizamos un ataque estádistico

Criptoanálisis de Vigenere - Kasiski

The *Kasiski test* was described by Friedrich Kasiski in 1863; however, it was apparently discovered earlier, around 1854, by Charles Babbage. It is based on the observation that two identical segments of plaintext will be encrypted to the same ciphertext whenever their occurrence in the plaintext is δ positions apart, where $\delta \equiv 0 \pmod{m}$. Conversely, if we observe two identical segments of ciphertext, each of length at least three, say, then there is a good chance that they correspond to identical segments of plaintext.

The Kasiski test works as follows. We search the ciphertext for pairs of identical segments of length at least three, and record the distance between the starting positions of the two segments. If we obtain several such distances, say $\delta_1, \delta_2, \ldots$, then we would conjecture that m divides all of the δ_i's, and hence m divides the greatest common divisor of the δ_i's.

Criptoanálisis de Vigenere - Indice de Coincidencia Mutua

Definition 1.7: Suppose $\mathbf{x} = x_1 x_2 \cdots x_n$ is a string of n alphabetic characters. The *index of coincidence* of $\mathbf{x}$, denoted $I_c(\mathbf{x})$, is defined to be the probability that two random elements of $\mathbf{x}$ are identical.

Suppose we denote the frequencies of $A, B, C, \ldots, Z$ in $\mathbf{x}$ by $f_0, f_1, \ldots, f_{25}$ (respectively). We can choose two elements of $\mathbf{x}$ in $\binom{n}{2}$ ways.[3] For each i, $0 \leq i \leq 25$, there are $\binom{f_i}{2}$ ways of choosing both elements to be i. Hence, we have the formula

$$I_c(\mathbf{x}) = \frac{\sum_{i=0}^{25} \binom{f_i}{2}}{\binom{n}{2}} = \frac{\sum_{i=0}^{25} f_i(f_i - 1)}{n(n-1)}.$$

Suppose $\mathbf{x}$ is a string of English language text. Denote the expected probabilities of occurrence of the letters $A, B, \ldots, Z$ in Table 1.1 by $p_0, \ldots, p_{25}$, respectively. Then, we would expect that

$$I_c(\mathbf{x}) \approx \sum_{i=0}^{25} {p_i}^2 = 0.065,$$

Criptoanálisis de Vigenere - Indice de Coincidencia Mutua

Now, suppose we start with a ciphertext string $\mathbf{y} = y_1 y_2 \cdots y_n$ that has been constructed by using a *Vigenère Cipher*. Define m substrings of $\mathbf{y}$, denoted $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m$, by writing out the ciphertext, in columns, in a rectangular array of dimensions $m \times (n/m)$. The rows of this matrix are the substrings $\mathbf{y}_i$, $1 \leq i \leq m$. In other words, we have that

$$\begin{aligned}
\mathbf{y}_1 &= y_1 y_{m+1} y_{2m+1} \cdots, \\
\mathbf{y}_2 &= y_2 y_{m+2} y_{2m+2} \cdots, \\
&\vdots \\
\mathbf{y}_m &= y_m y_{2m} y_{3m} \cdots.
\end{aligned}$$

If $\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_m$ are constructed in this way, and m is indeed the keyword length, then each value $I_c(\mathbf{y}_i)$ should be roughly equal to 0.065. On the other hand, if m is not the keyword length, then the substrings $\mathbf{y}_i$ will look much more random, since they will have been obtained by shift encryption with different keys. Observe that a completely random string will have

$$I_c \approx 26 \left(\frac{1}{26}\right)^2 = \frac{1}{26} = 0.038.$$

Criptoanálisis de Vigenere - Ejemplo

Example 1.12 Ciphertext obtained from a *Vigenère Cipher*

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

First, let's try the Kasiski test. The ciphertext string CHR occurs in five places in the ciphertext, beginning at positions 1, 166, 236, 276 and 286. The distances from the first occurrence to the other four occurrences are (respectively) 165, 235, 275 and 285. The greatest common divisor of these four integers is 5, so that is very likely the keyword length.

Criptoanálisis de Vigenere - Ejemplo

Let's see if computation of indices of coincidence gives the same conclusion. With $m = 1$, the index of coincidence is 0.045. With $m = 2$, the two indices are 0.046 and 0.041. With $m = 3$, we get 0.043, 0.050, 0.047. With $m = 4$, we have indices 0.042, 0.039, 0.045, 0.040. Then, trying $m = 5$, we obtain the values 0.063, 0.068, 0.069, 0.061 and 0.072. This also provides strong evidence that the keyword length is five. ∎

Criptoanálisis de Vigenere - Ejemplo

TABLE 1.4
Values of M_g

i	value of $M_g(\mathbf{y}_i)$								
1	.035	.031	.036	.037	.035	.039	.028	.028	.048
	.061	.039	.032	.040	.038	.038	.045	.036	.030
	.042	.043	.036	.033	.049	.043	.042	.036	
2	.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	.066	.035	.038	.036	.045
	.027	.035	.034	.034	.036	.035	.046	.040	
4	.045	.032	.033	.038	.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	.072	.037	.027	.031	.048	.036	.037	

Criptoanálisis de Vigenere - Ejemplo

Example 1.12 ***(Cont.)*** We have hypothesized that the keyword length is 5. We now compute the values M_g as described above, for $1 \leq i \leq 5$. These values are tabulated in Table 1.4. For each i, we look for a value of M_g that is close to 0.065. These g's determine the shifts $k_1, \ldots, k_5$.

From the data in Table 1.4, we see that the key is likely to be $K = (9, 0, 13, 4, 19)$, and hence the keyword likely is $JANET$. This is correct, and the complete decryption of the ciphertext is the following:

Criptoanálisis de Vigenere - Ejemplo

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.[4]

The end.