Vigenere

Roberto Gómez Cárdenas
rogomez@itesm.mx
http://cryptomex.org
@cryptomex

Antes de Vigenere

Leon Battista Alberti
Italiano (1404 – 1472)
Tractate on Ciphers

Johannes Trithemius
Aleman (1462- 1516)

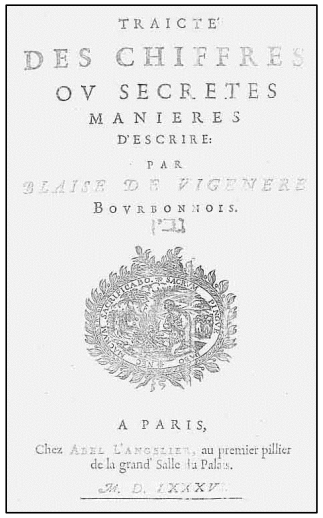
Giovan Battista Bellaso
Italiano (1505- ?????)
La Cifra del Sig. Giovan Battista Bellaso

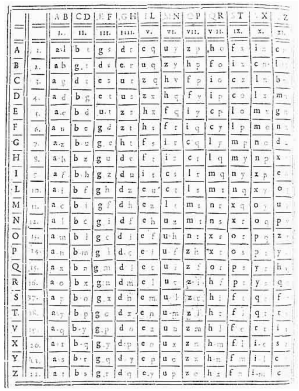
Giovanni Battista della Porta.
Italiano (1535 – 1615)

Blaise Vigenère

Francia (1523 – 1596)

- Cifrado se conocer como el criptosistema de Vigenere.
- Fue descrito por primera vez por Giobsn Battista Bellaso en 1553
- En el siglo XIX le fue atribuido, por error, a Blaise Vigenere
- El tan solo le propuso este criptosistema a Henry III
- Paternidad del Criptosistema volvió a Giovan Batista Belaso.
- Traicté des Chiffres (1585)
  - Le chiffre indéchiffrable
  - Resistió casi dos siglos sin que nadie la rompiera.
- Cifrado de autollave


TRAICTE'
DES CHIFFRES
OV SECRETES
MANIERES
D'ESCRIRE:
PAR
BLAISE DE VIGENERE
BOVRBONNOIS.
A PARIS,
Chez ABEL L'ANGELIER, au premier pillier de la grand' Salle du Palais.
M. D. LXXXV.



Ejemplo cifrado

- Texto plano: EL MUNDO ES UNA TOMBOLA
- Llave: SIUX

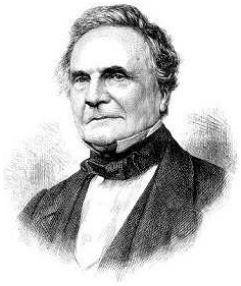
E	L	M	U	N	D	O	E	S	U	N	A	T	O	M	B	O	L	A
S	I	U	X	S	I	U	X	S	I	U	X	S	I	U	X	S	I	U
W	T	G	R	F	L	I	B	K	C	H	X	L	W	G	Y	G	T	U

- Criptograma: WTGRF LIBKC HXLWG YGTU
- Formula:

$$C_i = M_i + K_i \bmod 27$$

Babagge, Kasiski y Friedman

- Charles Babagge
  - En 1854 logra romper Vigenere
  - En realidad rompe el criptosistema llamado autoclave
  - No hace público sus avances en el criptoanálisis de sistemas polialfabéticos, ni tampoco los resultados a los que llega.
- Friedrich W. Kasiski
  - En 1863 publica su metodología de criptoanálisis: Die Geheimschriften und die Dechiffrierkunst
  - primera publicación sobre criptoanálisis aplicado a los cifrados de sustitución polialfabéticos
- William F. Friedman
  - En 1922, (casi 60 años después del ataque de Kasiski), publica: The Index of Coincidence and Its Applications in Cryptanalysis
  - Documento de alto secreto y que se mantiene clasificado como confidencial durante 50 años.

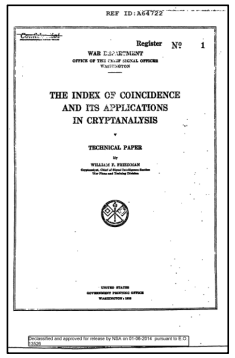


Charle Babagge
UK (1791-1871)

Friedrich Kasiski
Alemania (1805 – 1881)

William Friedman
USA (1891-1969)

Escritura Secreta y el Arte de Descifrar



El Índice de Coincidencia y sus Aplicaciones en el Criptoanálisis

Detalles del cifrado

- Llave se repite a lo largo del texto plano
- Si la llave es de longitud *n*, cada *n* letras el cifrado se lleva a cabo con el mismo alfabeto
- Ejemplo, llave: SIUX

E	L	M	U	N	D	O	E	S	U	N	A	T	O	M	B	O	L	A
S	I	U	X	S	I	U	X	S	I	U	X	S	I	U	X	S	I	U
W	T	G	R	F	L	I	B	K	C	H	X	L	W	G	Y	G	T	U

- Letras E,N,S,T y O son cifradas con alfabeto que empieza con S
- Letras L,D,U,O y L son cifradas con alfabeto que empieza con I
- Letras M,O,N,M y A son cifradas con alfabeto que empieza con U
- Letras U,E,A y B son cifradas con alfabeto que empieza con X

Cadenas típicas

- Existen trigramas o cuatrigramas que son comunes en los diferentes lenguajes.
- Ejemplo trigramas en español:
  - ADO, IDO,
- Ejemplo cuatrigramas en español:
  - ANDO, CION
- Ejemplo pentagramas
  - MENTE, IENDO,
- Estos se repetirán en un texto en claro lo suficientemente extenso, y cabe la posibilidad de que ese conjunto de n-gramas frecuentes se cifre más de una vez con la misma parte de la llave.

Ejemplo

SHANNON ES RECONOCIDO POR HABER FUNDADO EL CAMPO DE LA TEORÍA DE LA INFORMACIÓN. MIENTRAS REALIZABA SU MAESTRÍA EN EL MASSACHUSETTS INSTITUTE OF TECHNOLOGY, DEMOSTRÓ CON SU TESIS QUE LAS APLICACIONES ELECTRÓNICAS DE ÁLGEBRA BOOLEANA PODRÍAN CONSTRUIR CUALQUIER RELACIÓN LÓGICO-NUMÉRICA. CONTRIBUYÓ ASIMISMO AL CAMPO DEL CRIPTOANÁLISIS PARA LA DEFENSA DE ESTADOS UNIDOS DURANTE LA SEGUNDA GUERRA MUNDIAL, CON TRABAJOS SOBRE EL DESCIFRADO DE CÓDIGOS Y LA SEGURIDAD EN LAS TELECOMUNICACIONES.

Ejemplo

SHANNON ES RECONOCIDO POR HABER FUNDADO EL CAMPO DE LA TEORÍA DE LA INFORMACIÓN. MIENTRAS REALIZABA SU MAESTRÍA EN EL MASSACHUSETTS INSTITUTE OF TECHNOLOGY, DEMOSTRÓ CON SU TESIS QUE LAS APLICACIONES ELECTRÓNICAS DE ÁLGEBRA BOOLEANA PODRÍAN CONSTRUIR CUALQUIER RELACIÓN LÓGICO-NUMÉRICA. CONTRIBUYÓ ASIMISMO AL CAMPO DEL CRIPTOANÁLISIS PARA LA DEFENSA DE ESTADOS UNIDOS DURANTE LA SEGUNDA GUERRA MUNDIAL, CON TRABAJOS SOBRE EL DESCIFRADO DE CÓDIGOS Y LA SEGURIDAD EN LAS TELECOMUNICACIONES.

- Repetición n-gramas
  - ACION - aparece 4 veces, separadas 99, 164 y 183 espacios
  - ONES - aparece 3 veces, separadas 158 y 247 espacio
  - DELA - aparece 2 veces, separadas 10 espacios
  - NICA - aparece 2 veces, separadas 230 espacios
  - IDO - aparece 2 veces, separadas 293 espacios
  - CON -aparece 6 veces

Rompiendo Vigenere.

- Para descifrar un criptograma que se ha cifrado polialfabéticamente con una llave periódica, como es el caso de Vigenère, es necesario:
  - Buscar repeticiones de tres, cuatro, cinco o más caracteres en el criptograma.
  - Anotar los espacios que separan entre sí a dichas repeticiones, lógicamente cadenas de texto iguales.
  - Encontradas esas repeticiones y apuntadas esas distancias, calcular el máximo común denominador entre ellas.
  - Este será el valor será candidato a ser la longitud L de la clave buscada.
- Es decir, si las distancias entre las cadenas repetidas observadas han sido d1, d2, d3,… dn, el período (longitud) L de la llave puede ser:
  - L = mcd (d1, d2, d3,… dn)
- No es recomendable buscar repeticiones de dos caracteres en el criptograma (incluso a veces de tres caracteres) porque es posible que éstas se origen por azar.

Ejemplo repeticiones

- Sea el criptograma:

```
VOPVC  FRTCU  MBGHG  KGSCT  YKTNJ  MPKIB  VGYQE  YSSRK  YYEEJ  GWGVQ  RGPST
LNDIF  VIOMM  TGGYU  SLGZN  CZCJE  YGGSJ  VGJIJ  DSLRV  ARWCY  ECOTP  VWYUS
CEIQL  SQLIP  DMLGW  RJEQJ  IAZFG  JIJRO  RRILV  OFGWÑ  ZXYCY  QHWYE  NNKIB
VPYUV  GUHNE  HCVWR  RFYZQ  EJIQR  HNLVY  KWSWV  GJYLR  GAZHC  EXCUY  PRQRV
YLNMY  AIBVG  YTIPZ  ECEFN  LWSRQ  YIYCC  IÑJST  GGNMQ  YWVYT  XSJEC  EOYTE
BVVY
```

```
VOPVC  FRTCU  MBGHG  KGSCT  YKTNJ  MPKIB  VGYQE  YSSRK  YYEEJ  GWGVQ  RGPST
LNDIF  VIOMM  TGGYU  SLGZN  CZCJE  YGGSJ  VGJIJ  DSLRV  ARWCY  ECOTP  VWYUS
CEIQL  SQLIP  DMLGW  RJEQJ  IAZFG  JIJRO  RRILV  OFGWÑ  ZXYCY  QHWYE  NNKIB
VPYUV  GUHNE  HCVWR  RFYZQ  EJIQR  HNLVY  KWSWV  GJYLR  GAZHC  EXCUY  PRQRV
YLNMY  AIBVG  YTIPZ  ECEFN  LWSRQ  YIYCC  IÑJST  GGNMQ  YWVYT  XSJEC  EOYTE
BVVY
```

Tomando de Criptored: http://www.criptored.upm.es/crypt4you/temas/criptografiaclasica/leccion9.html

Deduciendo la longitud de la llave

```
VOPVC  FRTCU  MBGHG  KGSCT  YKTNJ  MPKIB  VGYQE  YSSRK  YYEEJ  GWGVQ  RGPST
LNDIF  VIOMM  TGGYU  SLGZN  CZCJE  YGGSJ  VGJIJ  DSLRV  ARWCY  ECOTP  VWYUS
CEIQL  SQLIP  DMLGW  RJEQJ  IAZFG  JIJRO  RRILV  OFGWÑ  ZXYCY  QHWYE  NNKIB
VPYUV  GUHNE  HCVWR  RFYZQ  EJIQR  HNLVY  KWSWV  GJYLR  GAZHC  EXCUY  PRQRV
YLNMY  AIBVG  YTIPZ  ECEFN  LWSRQ  YIYCC  IÑJST  GGNMQ  YWVYT  XSJEC  EOYTE
BVVY
```

- Dos cadenas de 4 letras repetidas una vez: KIBV y GJIJ
- Cuatro cadenas de 3 letreas repetidas una vez: TGG, YUS, VGJ y ECE
- Distancias cadenas
  - KIBV = 135; GJIJ = 48; TGG = 189; YUS = 39; VGJ = 114; ECE = 33
- Longitud llave: L = mcd (135, 48, 189, 39, 114, 33) = 3
- Tres líneas cifradas con el mismo alfabeto.

Separando los criptogramas

VOPVC FRTCU MBGHG KGSCT YKTNJ MPKIB VGYQE YSSRK YYEEJ GWGVQ RGPST
LNDIF VIOMM TGGYU SLGZN CZCJE YGGSJ VGJIJ DSLRV ARWCY ECOTP VWYUS
CEIQL SQLIP DMLGW RJEQJ IAZFG JIJRO RRILV OFGWÑ ZXYCY QHWYE NNKIB
VPYUV GUHNE HCVWR RFYZQ EJIQR HNLVY KWSWV GJYLR GAZHC EXCUY PRQRV
YLNMY AIBVG YTIPZ ECEFN LWSRQ YIYCC IÑJST GGNMQ YWVYT XSJEC EOYTE
BVVY

- Criptogramas monoalfabéticos:
  - C_1: VVRUG KCKJK VQSKE GVGTD VMGUG CJGJJ DRRYO VUELL
    DGJJZ JRRVG ZCHEK VUUEV RZJRL KWJRZ EURVN AVTZE
    LRICJ GMWTJ ETV
  - C_2: OCTMH GTTMI GESYE WQPLI IMGSZ ZEGVI SVWET WSISI
    MWEIF IOIOW XYWNI PVHHW FQIHV WVYGH XYQYM IGIEF
    WQYIS GQVXE OEV
  - C_3: PFCBG SYNPB YYRYJ GRSNF OTYLN CYSGJ LACCP YCQQP
    LRQAG JRLFÑ YQYNB YGNCR YEQNY SGLAC CPRLY BYPCN
    SYCÑT NYYSC YBY

La regla AEO

- Una vez definidos los subcriptogramas, se buscan las posiciones relativas que ocupan las letras más frecuentes del idioma original.
- En el caso del idioma español estas letras son la letra A, la letra E y la letra O
  - Deben presentar una mayor frecuencia en el subcriptograma.
- Deben cumplir con la estructura del alfabeto
  - De la letra A (código 0) hasta la letra E (código 4) hay 4 espacios
  - De la letra E (código 4) hasta la letra O (código 15) hay 11 espacios
  - Se podría utilizar también la S (código 19), cuarta letra más frecuente.

A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26

Aplicando la regla AEO

- Frecuencias de las letras en los criptogramas.
  - Letra A en color amarillo
  - Letra E en color verde
  - Letra O en color azul celeste

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	Ñ	O	P	Q	R	S	T	U	V	W	X	Y	Z
C_1	1	0	4	3	7	0	9	1	1	11	6	4	2	1	0	1	0	1	10	1	4	6	12	2	0	1	5
							O												A				E				
C_2	0	0	1	0	8	3	7	5	14	0	0	1	5	1	0	4	2	5	0	6	4	0	7	9	3	6	2
					A				E											O							
C_3	3	5	11	0	1	3	6	0	0	3	0	6	0	8	2	1	6	5	6	6	2	0	0	0	0	19	0
			E											O												A	

- Leyendo las celdas en color amarillo, las llave es: REY

Texto plano

- El texto plano es:

```
ELREY HAPED IDODI SCULP ASPOR IRSED ECAZA ABOTS UANAL OSIEN TOMUC
HOMEH EEQUI VOCAD ONOVO LVERA AOCUR RIREL MONAR CASEH AEXPR ESADO
ENEST OSTER MINOS TRASR ECIBI RELAL TAENE LHOSP ITALU SPSAN JOSED
EMADR IDDON DEEST ABAIN GRESA DOTRA SSUFR IRUNA CCIDE NTEDU RANTE
UNVIA JEDEC ACERI AENBO TSUAN AQUEL EPROV OCOUN AFRAC TURAE NLACA
DERA
```

- Acomodando las palabras:

```
EL REY HA PEDIDO DISCULPAS POR IRSE DE CAZA A BOTSUANA LO SIENTO MUCHO ME HE
EQUIVOCADO NO VOLVERA A OCURRIR EL MONARCA SE HA EXPRESADO EN ESTOS TERMINOS TRAS
RECIBIR EL ALTA EN EL HOSPITAL USP SAN JOSE DE MADRID DONDE ESTABA INGRESADO TRAS
SUFRIR UN ACCIDENTE DURANTE UN VIAJE DE CACERIA EN BOTSUANA QUE LE PROVOCO UNA
FRACTURA EN LA CADERA
```

- Añadiendo signos de puntación:

EL REY HA PEDIDO DISCULPAS POR IRSE DE CAZA A BOTSUANA: "LO SIENTO MUCHO. ME HE EQUIVOCADO. NO VOLVERÁ A OCURRIR". EL MONARCA SE HA EXPRESADO EN ESTOS TÉRMINOS TRAS RECIBIR EL ALTA EN EL HOSPITAL USP SAN JOSÉ DE MADRID, DONDE ESTABA INGRESADO TRAS SUFRIR UN ACCIDENTE DURANTE UN VIAJE DE CACERÍA EN BOTSUANA, QUE LE PROVOCÓ UNA FRACTURA EN LA CADERA.