Epsilones

Descifrado por análisis de frecuencias

SBLY AFL FWY IL UYC VYDQGYC VEC ZYILBYCYC AFL SYWIFSLW EU ÑYVPBL
EU EBDL K E UE SQLWSQE LC LU ILCLY IL LGEIQBCL IL UE GQIE SYDQIQEWE
SYW CF ECZLBLME IYUYBYCE K CF GESQY ILCLCZLBEWDL IL LCSEZEB E UEC
SEILWEC IL ILCLYC CQLVZBL SEVPQEWDLC EUPLBD LQWCDLQW

¿Qué? ¿Se entiende? ¿No? Bueno, es lógico, porque se trata de un criptograma, es decir, un mensaje secreto. Sin embargo, vamos a intentar descifrarlo mediante algunos trucos que aprendí del viejo Poe. Para aplicarlos hace falta conocer algunas técnicas, echarle algo de imaginación y bastante paciencia para probar las distintas conjeturas que se nos vayan ocurriendo.

La idea central, lo que se conoce como análisis de frecuencias, es que no todas las letras aparecen con la misma frecuencia, sino que unas lo hacen mucho más a menudo que otras. Según aparece en la web *Criptología*, y según un estudio sobre textos del diario El País, la frecuencia de las letras en castellano es aproximadamente la que sigue (en lo sucesivo seguiré la convención criptográfica por la cual las letras del texto plano, es decir, el texto original, se escriben en minúsculas, y las del texto cifrado, en mayúsculas):

e- 16,78%	r - 4,94%	y - 1,54%	j - 0,30%
a - 11,96%	u - 4,80%	q - 1,53%	ñ - 0,29%
o - 8,69%	i - 4,15%	b - 0,92%	z - 0,15%
l - 8,37%	t - 3,31%	h - 0,89%	x - 0,06%
s - 7,88%	c - 2,92%	g - 0,73%	k - 0,00%
n - 7,01%	p - 2,776%	f - 0,52%	w - 0,00%
d - 6,87%	m - 2,12%	v - 0,39%	

Lo primero que llama la atencion es que entre las seis primeras letras suman más de dos tercios del total.

Contando los caracteres del texto cifrado, tenemos para cada letra:

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
2	11	22	7	26	6	4	0	15	0	2	35	1	0	1	0	3	12	0	11	0	9	5	11	0	18	5

que ordenadas por frecuencias dan la siguiente tabla:

L	E	C	Y	I	Q	B	S	W	U	D	F	V	Z	G	P	A	K	M	Ñ	H	J	N	O	R	T	X
35	26	22	18	15	12	11	11	11	9	7	6	5	5	4	3	2	2	1	1	0	0	0	0	0	0	0

Estamos hablando de valores estadísticos. Para textos muy grandes la cosa funciona muy bien, pero cuando el texto cifrado es pequeño, como en nuestro caso, debemos avanzar poco a poco. Por eso probaremos solo a sustituir los dos caracteres más frecuentes en el texto cifrado por los dos caracteres más frecuentes en castellano. Entonces **L = e** y **E = a** y tenemos:

> SBeY AFe FWY Ie UYC VYDQGYC VaC ZYIeBYCYC AFe SYWIFSeW aU ÑYVPBe aU aBDe K a Ua SQeWSQa eC eU IeCeY Ie eGaIQBCe Ie Ua GQIa SYDQIQaWa SYW CF aCZeBeMa IYUYBYCa K CF GaSQY IeCeCZeBaWDe Ie eCSaZaB a UaC SaIeWaC Ie IeCeYC CQeVZBe SaVPQaWDeC aUPeBD eQWCDeQW

Otro truco es analizar palabras cortas de dos y tres letras, en las que es frecuente encontrar la **l** (por los artículos determinados) y la **s** (formas plurales, partículas reflexivas...). Encontramos las siguientes:

- UYC; eC; VaC; C4 y varios finales de palabra con C
- aU; Ua; eU; UaC

que hacen pensar en que **U = l** y **C = s**:

> SBeY AFe FWY Ie lYs VYDQGYs Vas ZYIeBYsYs AFe SYWIFSeW al ÑYVPBe al aBDe K a la SQeWSQa es el IeseY Ie eGaIQBse Ie la GQIa SYDQIQaWa SYW sF asZeBeMa IYlYBYsa K sF GaSQY IesesZeBaWDe Ie esSaZaB a las SaIeWas Ie IeseYs sQeVZBe SaVPQaWDes alPeBD eQWsDeQW

Volviendo a la tabla de frecuencias, vemos que entre las más frecuentes tenemos la **o** del texto plano y la **Y** del cifrado. Al sustituir, varias palabras parecen confirmarlo:

> SBeo AFe FWo Ie los VoDQGos Vas ZoIeBosos AFe SoWIFSeW al ÑoVPBe al aBDe K a la SQeWSQa es el Ieseo Ie eGaIQBse Ie la GQIa SoDQIQaWa SoW sF asZeBeMa IoloBosa K sF GaSQo IesesZeBaWDe Ie esSaZaB a las SaIeWas Ie Ieseos sQeVZBe SaVPQaWDes alPeBD eQWsDeQW

Además, la **I** aparece como candidata para ser la **d** y la **B** como **r:**

> SBeo AFe FWo **Ie** los VoDQGos Vas ZoIeBosos AFe SoWIFSeW al ÑoVPBe al aBDe K a la SQeWSQa es el **Ieseo Ie** eGaIQBse **Ie** la GQIa SoDQIQaWa SoW sF asZeBeMa **IoloBosa** K sF GaSQo IesesZeBaWDe **Ie** esSaZaB a las SaIeWas **Ie Ieseos** sQeVZBe SaVPQaWDes alPeBD eQWsDeQW

Quedaría:

> Sreo AFe FWo de los VoDQGos Vas Zoderosos AFe SoWdFSeW al ÑoVPre al arDe K a la SQeWSQa es el deseo de eGadQrse de la GQda SoDQdQaWa SoW sF asZereMa dolorosa K sF GaSQo desesZeraWDe de esSaZar a las SadeWas de deseos sQeVZre SaVPQaWDes alPerD eQWsDeQW

Esto ya casi está: se ve que la **Z** es la **p** y que la **S** la **c**:

> **Sreo** AFe FWo de los VoDQGos Vas **Zoderosos** AFe SoWdFSeW al ÑoVPre al arDe K a la SQeWSQa es el deseo de eGadQrse de la GQda SoDQdQaWa SoW sF asZereMa dolorosa K sF GaSQo **desesZeraWDe** de **esSaZar** a las SadeWas de deseos sQeVZre SaVPQaWDes alPerD eQWsDeQW

Quedaría:

> creo AFe FWo de los VoDQGos Vas poderosos AFe coWdFceW al ÑoVPre al arDe K a la cQeWcQa es el deseo de eGadQrse de la GQda coDQdQaWa coW sF aspereMa dolorosa K sF GacQo desesperaWDe de escapar a las cadeWas de deseos sQeVpre caVPQaWDes alPerD eQWsDeQW

Lo que tenemos hasta ahora es los siguiente:

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
E		S	I	L							U				Y	Z		B	C							

Si el criptoanalista me conociese sabría que puedo ser lo suficientemente torpe como para elegir la palabra **EPSILONES** como clave. En tal caso completar la tabla sería trivial y el texto quedaría descifrado.

De todas maneras, tal intuición no nos es necesaria: basta mirar el texto para deducir las nuevas letras **q** y **u** de la reveladora secuencia **AF**, así como que **W = n**. Sustituyendo:

> creo que uno de los VoDQGos Vas poderosos que conducen al ÑoVPre al arDe K a la cQencQa es el deseo de eGadQrse de la GQda coDQdQana con su aspereMa dolorosa K su GacQo desesperanDe de escapar a las cadenas de deseos sQeVpre caVPQanDes alPerD eQnsDeQn

Resumiendo lo obtenido hasta ahora, tenemos:

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
E		S	I	L							U		W		Y	Z	A	B	C		F					

Observando la tabla de cifrado, la cosa parece bastante clara, y podemos deducir nuevas letras:

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
E		S	I	L							U	**V**	W	**X**	Y	Z	A	B	C	**D**	F					

que dan:

> creo que uno de los motQGos mas poderosos que conducen al ÑomPre al arte K a la cQencQa es el deseo de eGadQrse de la GQda cotQdQana con su aspereMa dolorosa K su GacQo desesperante de escapar a las cadenas de deseos sQempre camPQantes **alPert eQnsteQn**

Del texto anterior salta a la vista que **Q = i**. Además, si nos fijamos en las dos últimas palabras, parece claro que un viejo conocido nos dice que **P = b**:

> creo que uno de los motiGos mas poderosos que conducen al Ñombre al arte K a la ciencia es el deseo de eGadirse de la Gida cotidiana con su aspereMa dolorosa K su Gacio desesperante de escapar a las cadenas de deseos siempre cambiantes albert einstein

El resto de letras es evidente. Terminamos:

> **creo que uno de los motivos mas poderosos que conducen al hombre al arte y a la ciencia es el deseo de evadirse de la vida cotidiana con su aspereza dolorosa y su vacio desesperante de escapar a las cadenas de deseos siempre cambiantes albert einstein**

La tabla de cifrado quedaría entonces así:

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
E	P	S	I	L	O	N	Ñ	Q	R	T	U	V	W	X	Y	Z	A	B	C	D	F	G	H	J	K	M

siendo la clave, efectivamente, **EPSILONES** (las letras repetidas se ignoran).

Hay que indicar que en el ejemplo que se acaba de descifrar se han hecho varias suposiciones no necesariamente ciertas: se ha supuesto que estaba en castellano, que es mucho suponer; se ha supuesto que estaba cifrado siguiendo un sistema monoalfabético sin repeticiones; se ha dado por hecho que los espacios del texto cifrado correpondían a espacios del texto plano... ¿Que hubiese pasado si alguna de estas suposiciones no hubiese resultado cierta: pues que tras darle vueltas y vueltas hubiésemos hubiésemos empezado a sospechar que así era, hubiésemos entonces cambiado alguna de las suposiciones y vuelta a empezar. Por eso en criptografía es importante cualquier información añadida acerca del mensaje y del criptógrafo que lo ha generado.

La facilidad con que se descifra este tipo de criptografía explica que no se utilice desde hace mucho tiempo cuando se quieren unas comunicaciones realmente seguras. Sin embargo, es revelador saber que durante muchos siglos fue un sistema considerado seguro. Y lo fue, hasta que a alguien se le ocurrió el análisis de frecuencias.

Comentarios Inicio página

Epsilones.
Sitio + o - matemático de
Alberto Rodríguez Santos.
Correo: alberto@epsilones.com.
En la red desde el 4-7-2002 (ya hace).
Última actualización: ver Novedades.

Con esto se termina la página: