Cómo descifrar el código

Un informe de investigación de muestra

«*To keep your secret is wisdom; but to expect others to keep it is folly.*»
(«Saber guardar secretos es una muestra de sabiduría, pero esperar que otros los guarden es una locura.»)

Samuel Johnson

«*Secrets are made to be found out with time*»
(«Los secretos están hechos para ser descubiertos con el tiempo»)

Charles Sanford

Los militares llevan miles de años utilizando códigos para mantener sus secretos fuera del alcance del enemigo. En la era de la información, en la que mucha gente utiliza Internet para realizar transacciones bancarias y para comprar, no queremos que la información que introducimos esté a disposición de personas no autorizadas. Por ello, en los sitios web seguros dicha información se codifica.

C El alumno podría haber relacionado esto con sus intereses personales.

Este informe analiza algunos de los códigos que se han utilizado a lo largo de los años y el papel que juegan las matemáticas para elaborar dichos códigos y para descifrarlos.

A La introducción no incluye las bases o fundamentos ni el objetivo general del trabajo.

El cifrado César por desplazamiento

El primer uso documentado de códigos para fines militares lo instauró Julio César (100 – 44 a. C.) El tipo de código que utilizó se conoce comúnmente como «desplazamiento de César». En la siguiente tabla, la fila central contiene el texto plano (original, sin codificar) y la fila inferior da el correspondiente texto cifrado para un desplazamiento de César de 2 lugares. Normalmente se escribe el texto plano sin codificar en minúsculas, y el texto codificado en mayúsculas.

Posición	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
Texto	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
Código	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B

Este código se puede representar mediante la aplicación $x \to x + 2$ donde x es la posición que ocupa la letra en el abecedario. De este modo la letra *k*, que ocupa la posición número 10, se convierte en la M, que ocupa la posición 12. Hacia el final del abecedario, la letra *y* pasaría a convertirse en la letra que ocupa la posición número 24 + 2 = 26; sin embargo, 25 es el número de posición máximo que hay, puesto que sólo hay 26 letras en el abecedario y utilizamos la posición 0 para la letra *a*. Lo que hacemos es restar 26 del resultado, y obtenemos 0. A esto se le denomina suma módulo 26.

Se han utilizado las posiciones de 0 a 25 (en vez de utilizar las comprendidas entre 1 y 26) porque la aritmética modular se puede interpretar también como el resto de una división. 26 es equivalente a 0 módulo 26 porque $26 \div 26 = 1$ y el resto de la división es 0.

Si utilizamos este código, la segunda cita (en inglés) que aparece al inicio de este informe se codificaría como sigue:

secretsaremadetobefoundoutwithtime
UGETGVUCTGOCFGVQDGHQWPFQWVYKVJVKOG

C Ejemplo del alumno relacionado con el tema

Se han eliminado los espacios entre palabras para que sea aún más difícil descifrar el código, aunque si alguien encontrase este mensaje y supiera que se trata de cifrado por desplazamiento de César no le llevaría mucho tiempo descifrarlo. Sólo existen 25 desplazamientos de César posibles, o 26 si incluimos el desplazamiento «cero» que no modifica las letras,... ¡pero ése no es muy secreto, que digamos!

Es posible escribir las 25 posibilidades utilizando algo tan sencillo como un bolígrafo y un papel, y una vez que obtenemos un mensaje que tiene sentido, ya sabemos que hemos logrado descifrarlo. Si utilizamos un computador para calcular los 25 descifres posibles, realmente se podría descifrar el código en muy poco tiempo.

Cifrados por sustitución

En un cifrado por sustitución cada letra es reemplazada por otra letra. Un desplazamiento de César es un tipo particular de cifrado por sustitución, pero el patrón de sustituciones en un desplazamiento de César es tal que resulta fácil descifrarlo. En un cifrado por sustitución genérico se reordenan las 26 letras del abecedario y se escribe esta permutación aleatoria debajo del abecedario normal. De este modo, habría $26! \approx 4 \times 10^{26}$ códigos posibles. ¡Analizar todos los posibles códigos hasta encontrar el que realmente se ha utilizado llevaría muchísimo tiempo! Sin embargo, algunos de los posibles códigos de sustitución se descifran demasiado fácilmente; por ejemplo, aquél en el que cada letra se codifica con la misma letra no sirve para nada, como tampoco sirven de mucho aquellos códigos en los que hay muchas letras que no cambian.

B Notación de aproximación bien aplicada.

Aquellas permutaciones donde ningún elemento se queda en su lugar original se denominan «desarreglos». El número de desarreglos posibles con un grupo de 26 letras es igual al subfactorial de 26, que se escribe !26. Para ver cómo se calcula este valor, consideremos un abecedario mucho más corto, de sólo 4 letras. El número total de permutaciones es 4! = 24; todas ellas se muestran a continuación.

abcd	abdc	acbd	acdb	adbc	adcb
bacd	**badc**	bcad	**bcda**	**bdac**	bdca
cabd	**cadb**	cbad	cbda	**cdab**	**cdba**
dabc	dacb	dbac	dbca	**dcab**	**dcba**

Los 9 posibles desarreglos aparecen en negrita, con un borde alrededor. Para calcular el número de desarreglos, comencemos por el número total de permutaciones y vayamos quitando aquellas permutaciones que tienen al menos una letra en su posición original.

Hay 6 ordenamientos en los que la *a* está en su posición original: estos se muestran sobre fondo gris. Esto se debe a que cuando la *a* está en su posición original, las otras 3 letras se pueden ordenar de 3! = 6 formas distintas. Del mismo modo, hay 3! permutaciones posibles si se deja a *b* en su posición original, y lo mismo es aplicable a *c* y a *d*. Sin embargo, restar $4 \times 3!$ es demasiado, puesto que los ordenamientos en los que tanto *a* como *b* se mantienen en su posición original los hemos contado dos veces.

Así, $4! - 4 \times 3!$ (1.1) es un valor demasiado pequeño.

B *La explicación matemática está presentada correctamente.*

Si *a* y *b* se mantienen ambas en su posición original nos quedan 2 letras para permutar, las cuales se pueden ordenar de 2! maneras posibles. Lo mismo se cumple para otros pares de letras. Hay ${}^4C_2$ pares de letras, por lo que tenemos que sumar ${}^4C_2 \times 2!$ a (1.1) para así obtener:

$4! - 4 \times 3! + {}^4C_2 \times 2!$ (1.2)

Sin embargo, los ordenamientos que tienen 3 letras en su posición original los hemos contado más de una vez. De hecho, no es posible que 3 letras mantengan su posición original y la 4ª letra esté en una posición diferente, por lo que la única combinación de este tipo que hemos contado demasiadas veces es aquella que tiene las cuatro letras en sus posiciones originales. Así, *abcd* se ha contado 4 veces, cuando debería haberse contado sólo una vez.

El número de desarreglos es:

$4! - 4 \times 3! + {}^4C_2 \times 2! - 4 + 1$ (1.3)

Esto se puede escribir de este modo:

$$!4 = 4! - \frac{4!}{1!} + \frac{4!}{2!} - \frac{4!}{3!} + \frac{4!}{4!} = 4!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!}\right) \quad (1.4)$$

C *Buena demostración del aprendizaje y descripción de aspectos matemáticos desconocidos.*

En general:

$$!n = n!\left(1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots\ldots + (-1)^n \frac{1}{n!}\right) \quad (1.5)$$

B *Buena definición matemática de los términos.*

E *El trabajo presentado es superior al nivel del programa de estudios.*

La serie de Maclaurin para e^x es

$$e^x = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots\ldots + \frac{x^r}{r!} + \ldots\ldots$$

De modo que $1 - \frac{1}{1!} + \frac{1}{2!} - \frac{1}{3!} + \ldots\ldots + (-1)^n \frac{1}{n!} \approx e^{-1}$

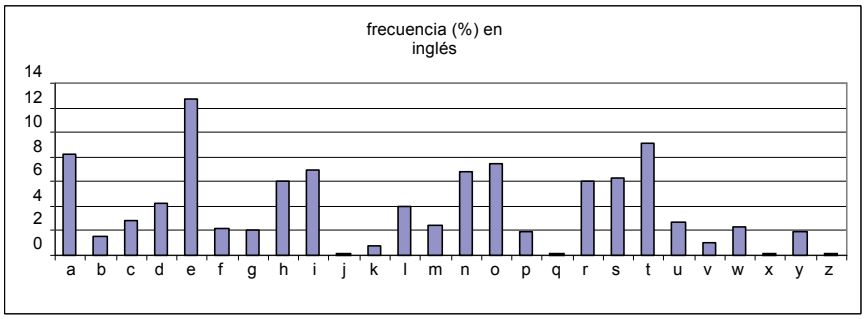
E *Se demuestra una buena comprensión de las matemáticas.*

$$!n \approx \frac{n!}{e}$$

De hecho, si redondeamos $\frac{n!}{e}$ obtenemos $!n$, con lo que el número de desarreglos posibles con el abecedario es igual a $!26 \approx \frac{26!}{e} \approx 1{,}5 \times 10^{26}$. Éstas son todavía muchísimas más posibilidades de las que se podrían probar en un intervalo de tiempo razonable; sin embargo, los códigos por sustitución se pueden descifrar rápidamente y con facilidad.

Cómo descifrar un código obtenido mediante cifrado por sustitución

En inglés algunas letras se utilizan con mayor frecuencia que otras. El siguiente gráfico muestra la frecuencia con la que se usa cada letra.



Si se ha cifrado el mensaje con un código de sustitución, la letra utilizada para codificar la *e* será la letra que aparezca con mayor frecuencia. Utilizando un computador, contar la frecuencia con la que aparecen las distintas letras es un proceso rápido y preciso; por ello, los códigos de sustitución no son seguros. La descripción más antigua que se conoce del análisis de frecuencias como método para descifrar códigos data del siglo IX, y su autor es el erudito árabe Al-Kindi. Además de analizar la frecuencia de las letras individuales, se puede comparar la frecuencia de pares de letras con ciertos pares que se emplean a menudo en inglés, como *th*, *er*, *on*, etc.

C Oportunidad para que el alumno demuestre su compromiso personal mediante el uso de un ejemplo.

El cifrado de Vigenère

En el siglo XVI, Blaise de Vigenère se basó en el trabajo iniciado por otros para desarrollar un cifrado que no fuese vulnerable al análisis de frecuencias y que resultase sencillo de utilizar. Este cifrado se basa en los desplazamientos de César pero no utiliza siempre el mismo desplazamiento. Para utilizar el cifrado, se necesita una contraseña o palabra clave. La persona que codifica el mensaje y la persona que recibe el mensaje tienen que conocer la contraseña, pero eso es todo lo que necesitan recordar. Por el contrario, si se utiliza un cifrado por sustitución que no esté basado en desplazamientos de César, es necesario o bien aprenderse de memoria o bien tener por escrito toda la tabla de sustituciones.

Vamos a ilustrar este método de cifrado utilizando la contraseña STELLA. Así, utilizaremos los alfabetos con desplazamiento de César que empiezan por las letras S, T, E, L y A.

Texto	a	b	c	d	e	f	g	h	i	j	k	l	m	n	o	p	q	r	s	t	u	v	w	x	y	z
Código	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
Código	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
Código	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D
Código	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z	A	B	C	D	E	F	G	H	I	J	K
Código	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z

Para codificar la primera cita de este informe, hay que escribir sobre el texto la contraseña, tantas veces como sea necesario:

```
stellastellastellastellastellastellastellastellastellast
tokeepyoursecretiswisdombuttoexpectotherstokeepitisfolly
```

C Uso de un ejemplo propio del alumno.

La primera letra del mensaje se codifica utilizando el alfabeto de César que comienza por la letra *S*, de modo que *t* se convierte en *L*. La segunda letra del mensaje se codifica utilizando el alfabeto que comienza por *T*, de modo que *o* se convierte en *H*. El mensaje, ya codificado, es el siguiente:

LHOPPPQHYCDEUKIETSOBWOZMTNXEZEPIINEOLAICDTGDIPAILBWQZLDR

El cifrado de Vigenère no es susceptible de ser descifrado mediante análisis de frecuencias, puesto que cada letra se codifica de diferentes formas. Sin embargo, este cifrado no se utilizó mucho porque se tarda bastante en codificar un mensaje manualmente, puesto que primero hay que ver qué alfabeto hay que utilizar y luego hay que encontrar la letra codificada correspondiente. Si se pudiese mecanizar el proceso resultaría mucho más sencillo, pero el uso de la tecnología fue también lo que permitió descifrar el código unos 400 años después de su invención.

Cómo descifrar el código obtenido mediante cifrado de Vigenère

Charles Babbage (autor también de un prototipo primitivo de computador) se dio cuenta de que para descifrar el código de Vigenère lo primero que había que hacer era decidir cuál es la longitud de la contraseña.

Consideremos el siguiente texto codificado:

JERFIWFLXFIWFLXVHXEAMCNWVHXHIWFLX ← C Ejemplo adicional.

Se han señalado en rojo, en el mensaje codificado, los grupos de letras que se repiten. Es probable que estos grupos provengan de grupos iguales de letras del mensaje original que se hayan codificado de la misma forma. En este caso, es probable que la contraseña sea de 3 letras, porque la distancia entre estas repeticiones es un múltiplo de 3.

```
123123123123123123123123123123123
JERFIWFLXFIWFLXVHXEAMCNWVHXHIWFLX
```

Si la palabra clave tiene 3 letras, entonces sabemos que todas las letras sobre las cuales hay un 1 se han codificado con el mismo desplazamiento de César. Lo mismo se puede aplicar a las letras que tienen un 2 asignado y a las que tienen un 3. Así, para descifrar el código, podemos utilizar ahora el análisis de frecuencias por separado para cada uno de estos conjuntos de letras (aquellas que tienen el mismo número asignado). Hay una herramienta informática muy útil para descifrar los códigos de Vigenère en la siguiente página web:
http://www.simonsingh.net/The_Black_Chamber/cracking_tool.html

La máquina Enigma

Para conseguir una adaptación del cifrado de Vigenère que sea más segura, es necesario utilizar una contraseña más larga y que, además, no sea una palabra real, para que no se pueda recurrir a conjeturas o a suposiciones para averiguarla. Esto hace que el cifrado de Vigenère resulte mucho más complicado de utilizar, a no ser que el proceso se pueda automatizar. En 1918, un inventor alemán llamado Arthur Scherbius inventó la máquina Enigma. Posteriormente se mejoró y fue utilizada por el ejército alemán durante la Segunda Guerra Mundial.

La máquina Enigma disponía de un teclado. Con ayuda del teclado, el operador escribía un mensaje y la máquina lo codificaba. El sistema se basa en un rotor que se utiliza para ir avanzando y codificar la siguiente letra utilizando un abecedario por sustitución distinto. Con un rotor y 26 posiciones tendríamos 26 posiciones iniciales posibles y, por tanto, 26 códigos posibles. Sin embargo, se utilizaron 3 rotores, cada uno de los cuales se configuró con una posición inicial distinta cada día, lo que da lugar a $26 \times 26 \times 26 = 17.576$ códigos posibles.

Es más, los operadores elegían 3 rotores de entre 5 rotores posibles; es decir, había ${}^{5}P_{3} = 60$ formas de colocar los rotores, lo que da lugar a un total de $60 \times 17.576 = 1.054.560$ códigos posibles.

Además de los 3 rotores, en el teclado se podían conectar pares de letras. Si la *a* está conectada a la *b*, el cifrado de la *a* se intercambia con el cifrado de la *b*. Se realizaron 10 conexiones: para ello, se eligieron 10 pares de letras de entre las 26 que había en el teclado.

El número de maneras de elegir el primer par de letras, luego el segundo par y así sucesivamente es:

$$
\begin{aligned}
{}^{26}C_{2} \times {}^{24}C_{2} \times {}^{22}C_{2} \times \ldots \times {}^{8}C_{2} &= \frac{26!}{24! \times 2!} \times \frac{24!}{22! \times 2!} \times \frac{22!}{20! \times 2!} \times \ldots \frac{8!}{6! \times 2!} \\
&= \frac{26!}{6! \times 2^{10}}
\end{aligned}
\qquad (1.6)
$$

E Buena demostración de la comprensión de las matemáticas

Sin embargo, el orden en el que se eligen los pares de letras no importa, por lo que concluimos que cada conjunto de 10 pares se ha contado demasiadas veces. Por ello, el valor de (1.6) se tiene que dividir por 10!.

$$\frac{26!}{6! \times 2^{10} \times 10!} \approx 1{,}5 \times 10^{14}$$

Si a esto le unimos las posibilidades que ofrecen los rotores, obtenemos un total de $1{,}5 \times 10^{14} \times 1.054.560 \approx 1{,}59 \times 10^{20}$ códigos posibles. Teniendo en cuenta que no era posible utilizar el análisis de frecuencias, el código Enigma se consideraba que era indescifrable.

Cómo descifrar el código Enigma

A los operadores de las máquinas Enigma se les entregaban unos libros que contenían los parámetros de configuración para ese día concreto. Se les decía qué letras tenían que conectar en el teclado, qué rotores tenían que utilizar y cuál tenía que ser la configuración inicial de cada rotor. Si los «descifradores de códigos» hubiesen tenido una máquina Enigma y hubiesen podido averiguar cuál era la configuración para ese día concreto, habrían sido capaces de descifrar todos los mensajes de ese día.

Los mensajes comenzaban con la máquina Enigma en esta posición y luego se transmitía la configuración de los rotores para ese mensaje; a continuación se ajustaban los rotores antes de transmitir el resto del mensaje. De esto modo, incluso si los «descifradores de códigos» lograban interceptar todos los mensajes de ese día, como en cada uno de ellos se utilizaba un código distinto (exceptuando las 3 primeras letras de cada mensaje), los «descifradores de códigos» no tenían muchos datos de los que agarrarse.

Para descifrar el código Enigma, los «descifradores de códigos» utilizaron una copia de una máquina Enigma y elaboraron unas máquinas, denominadas bombes, que podían ir probando un número enorme de posibles códigos. Estas máquinas no podían, ni de lejos, probar todos los posibles códigos en un sólo día, por lo que se necesitaba tener información adicional y realizar algún tipo de deducción.

La máquina Enigma nunca codificaba una letra como ella misma. Esto ayudó a descifrar algunos mensajes. Por ejemplo, un día interceptaron un mensaje que no contenía ninguna letra L. El operador había enviado un mensaje de prueba y se había limitado a pulsar todo el rato la letra L; por ello, era esta letra la única que no aparecía en el mensaje codificado. Los «descifradores de códigos» sabían que los mensajes que se enviaban a determinadas horas del día eran sobre el pronóstico del tiempo; esto les permitía adivinar una parte del mensaje. Además, el saber que una letra no se podía codificar como ella misma les permitió ir descartando posibilidades y, con eso, las «bombes» ya podían tratar de averiguar cuál era la configuración que se estaba utilizando ese día.

Supongamos que supiésemos que la palabra «weather» (tiempo meteorológico en inglés) forma parte del texto plano original correspondiente al siguiente mensaje codificado: Podemos probar con distintas posiciones, a ver si encontramos la palabra:

Mensaje codificado: `QZWPM ZHVVG YGQOZ UQZX`
Posición en el texto original:
`weath er`
`weat her`
`wea ther`

C Otro ejemplo personal.

La palabra «*weather*» no puede encontrarse en la posición indicada en esta tercera opción, puesto que la *w* no se puede codificar como *W* ni la *h* se puede codificar como *H.*

D El alumno podría haber reflexionado acerca de los efectos de descifrar el código.

Criptografía de clave pública

Todos los anteriores sistemas de cifrado requieren que la clave o contraseña se mantenga en secreto, de forma que ninguna persona no autorizada que pudiera tratar de descodificar el mensaje tenga acceso a ella. Esta clave podría ser el desplazamiento utilizado en el cifrado de César, la contraseña utilizada en el cifrado de Vigenère, o la configuración inicial de la máquina Enigma, pero una vez que alguien la conoce, ya puede descodificar los mensajes.

En la década de 1970, Rivest, Shamir y Adleman idearon una manera de codificar mensajes que no requería que la clave se mantuviera en secreto. El sistema que inventaron recibe el nombre de RSA. El sistema codifica números en vez de letras, pero también es posible construir un sistema equivalente que convierta letras o palabras en números. La siguiente tabla muestra los pasos de los que consta el proceso, utilizando para ello un ejemplo sencillo.

Paso	**Ejemplo sencillo**
Elegir 2 números primos, p y q. Por motivos de seguridad, tienen que ser números altos.	$p = 2$, $q = 11$
$m = pq$	$m = 22$

Calcular $A = (p-1)(q-1)$ y elegir un número E que sea menor que éste y que no tenga ningún divisor común con él.	$A = (p-1)(q-1) = 1\times 10 = 10$ $E = 3$
Hallar un número D tal que $DE - 1$ sea un múltiplo de A	$3D - 1$ ha de ser múltiplo de 10 $D = 7$
Para codificar un número M, hay que calcular $C = M^E \pmod m$	A continuación se muestran varios ejemplos
Para descifrar un número C, hay que calcular $M = C^D \pmod m$	
No es necesario mantener E y m en secreto, siempre y cuando D se mantenga en secreto.	

C Ejemplo propio.

Codificación

M	C
1	$1^3 \pmod{22} = 1$
2	$2^3 \pmod{22} = 8 \pmod{22} = 8$
3	$3^3 \pmod{22} = 27 \pmod{22} = 5$
4	$4^3 \pmod{22} = 64 \pmod{22} = 20$

Recordemos del apartado sobre cifrado de César que 64 (mod 22) es igual al resto que se obtienen de dividir 64 entre 22. Sin embargo, el trabajar con potencias de números puede derivar en cálculos con números muy grandes. Afortunadamente, las matemáticas pueden ayudarnos.

Imaginemos que queremos calcular ab (módulo 15). Supongamos que $a = r$ (módulo 15) y que $b = R$ (módulo 15).

De este modo, $a = 15x + r$ y $b = 15y + R$, donde x e y son números enteros.

Por lo tanto, $ab = (15x + r)(15y + R) = 15^2 xy + 15xR + 15r + rR$
Cuando se divide ab entre 15, el resto es el mismo que se obtiene cuando se divide rR entre 15. Así, para calcular cuánto es un número grande módulo 15, lo más fácil es escribir dicho número grande como producto de factores, calcular cada uno de ellos módulo 15 y multiplicar luego estos resultados. Esto se puede resumir del siguiente modo:

ab (módulo m) = ((a módulo m)(b módulo m)) módulo m .

De modo que $15^3 \pmod{22} = (15^2 \times 15) \pmod{22} = 225 \pmod{22} \times 15 \pmod{22}$
$= (5 \times 15) \bmod 22 = 75 \bmod 22 = 9$

E Se demuestra un uso correcto de las matemáticas aquí y en la página siguiente.

Descifre

M	C
1	$1^7 \pmod{22} = 1$
8	$8^7 \pmod{22} = (8^2 \bmod 22) \times (8^2 \bmod 22) \times (8^2 \bmod 22) \times (8 \bmod 22) \pmod{22}$ $= ((64 \bmod 22)^3 \times 8) \bmod 22$ $= (20^3 \times 8) \bmod 22$ $= 64.000 \bmod 22$ $= 2$
5	$5^7 \pmod{22} = (5^2 \bmod 22) \times (5^2 \bmod 22) \times (5^2 \bmod 22) \times (5 \bmod 22) \pmod{22}$ $= ((3^3 \bmod 22) \times 5) \bmod 22$ $= ((27 \bmod 22) \times 5) \bmod 22$ $= (5 \times 5) \bmod 22$ $= 3$
20	$20^7 \equiv 20^2 \times 20^2 \times 20^2 \times 20$ $\equiv 4 \times 4 \times 4 \times 20$ $\equiv 64 \times 20$ $\equiv 20 \times 20$ $\equiv 400 \equiv 4$

Ya que el escribir tantas veces «mod 22» hace que el cálculo parezca mucho más complicado de lo necesario, para descifrar el 20 se ha utilizado una notación distinta. ≡ significa «se obtiene el mismo resto cuando se divide entre 22».

En las tablas anteriores queda demostrado que el proceso de descifre te devuelve al número original, para el caso particular de los ejemplos que hemos probado. Este proceso siempre es válido. La demostración del mismo se basa en el teorema de Euler, que establece que:

$M^{\phi(m)} \equiv 1 \bmod m$

$\phi(m)$ es igual al número de enteros menores que m que no tienen ningún divisor común con m. Si $m = pq$, donde p y q son primos, $\varphi(m) = (p-1)(q-1)$, que es justamente lo que hemos denominado A en la descripción del algoritmo RSA.

Así pues, el teorema de Euler implica que $M^A \equiv 1 \bmod m$

Utilizando el símbolo ≡ para indicar «tiene el mismo resto cuando se lo divide entre m» y recordando que la fórmula para descifrar el código es $M = C^D \pmod{m}$

$C \equiv M^E$

$C^D \equiv (M^E)^D$

$\equiv M^{ED} = M^{ED-1} \times M$

$ED - 1$ es un múltiplo de A, de modo que $ED - 1 = kA$, donde k es un número entero. Así,

$C^D \equiv M^{kA} \times M = (M^A)^k \times M$

$\equiv 1^k \times M = M$

Cómo descifrar el código RSA

$m = pq$ es conocido; para descifrar el código necesitamos saber cuáles son estos números primos p y q. $m = 22$ no da lugar a un código seguro, puesto que resulta obvio que estos números primos son 2 y 11. Es necesario que estos números primos sean mucho más altos para que el código resultante sea seguro.

655.427 es el producto de dos números primos. Para hallar estos primos podríamos utilizar una hoja de cálculo. Éste es el comienzo de una hoja de cálculo que ilustra una de las formas de hallar los primos:

n	655427/n	Parte entera de	¿Factor?
1	655427	655.427	Sí
2	327713,5	327.713	No
3	218475,6667	218.475	No
4	163856,75	163.856	No
5	131085,4	131.085	No
6	109237,8333	109.237	No
7	93632,42857	93.632	No

En vez de limitarnos a dividir entre números primos, decidimos probar todos los números enteros. Así vemos que 655.427 es divisible entre 1, pero esto no nos ayuda demasiado. Los factores primos se encuentran un poco más abajo en la tabla:

438	1496,408676	1.496	No
439	1493	1.493	Sí
440	1489,606818	1.489	No

Tardé aproximadamente 1 minuto en construir la hoja de cálculo y hallar los factores, por lo que se puede concluir que 655.427 tampoco sirve para elaborar un código seguro.

C El alumno registró el tiempo que le llevó preparar la hoja de cálculo.

D Reflexión sobre la importancia del tiempo empleado.

La descomposición en factores se puede hacer de manera más rápida utilizando el método de Fermat de descomposición en factores. Este método se basa en el siguiente resultado:

$$m = pq = \left(\frac{p+q}{2}\right)^2 - \left(\frac{p-q}{2}\right)^2$$

Si desarrollamos los paréntesis del lado derecho obtenemos:

$$\frac{p^2+q^2+2pq}{4} - \frac{p^2+q^2-2pq}{4} = pq$$

Si *p* y *q* son números primos grandes ambos serán impares y, por lo tanto, $\frac{p+q}{2}$ y $\frac{p-q}{2}$ serán números enteros.

Supongamos que $m = x^2 - y^2$; por tanto, $x^2 - m = y^2$. Esto implica que $x^2 \geq m$ dado

que el cuadrado de un número no puede ser negativo, con lo que el valor más pequeño que puede tener m es $\sqrt{m}$.

Estamos tratando de descomponer en factores el número $m = 655.427$.
$\sqrt{655\ 427} = 809,58...$ con lo que el valor de x más pequeño posible es 810.

Para cada posible valor de x calculamos x^2 − m. Cuando el resultado sea igual al cuadrado de un número (y^2) los factores serán $(x - y)(x + y)$. Si hacemos estos cálculos en una hoja de cálculo obtenemos lo siguiente:

n	810+n	(810+n)^2-655.427=Y	√(Y)	Factor	Factor
1	811	2.294	47,8957	no	no
2	812	3.917	62,5859	no	no
3	813	5.542	74,4446	no	no
4	814	7.169	84,6699	no	no
5	815	8.798	93,7977	no	no
6	816	10.429	102,122	no	no

155	965	275.798	525,165	no	no
156	966	277.729	527	1.493	439
157	967	279.662	528,831	no	no

En este caso encontramos los factores tras probar con 156 números, mientras que con el método anterior tuvimos que probar con 439 números. Es tan rápido encontrar los factores para un valor de m de 6 cifras que realmente este método no supone una gran diferencia, pero demuestra que la velocidad de cálculo de un computador y el empleo de métodos eficientes de descomposición en factores pueden hacer que sea posible descifrar el código RSA, a no ser que los números primos que se utilicen sean muy altos. El record actual de descomposición de un número en factores está en 200 cifras.

«En los años ochenta en general se creía que bastaba con utilizar números primos de cincuenta y tantas cifras. Sin embargo, los avances llegaron mucho más rápido de lo previsto, y resulta complicado aventurarse a hacer previsiones cuantitativas en este campo. Cuando Rivest retó al mundo en 1977 a descomponer en factores RSA-129, un número de 129 cifras (de una lista especial), estimó que sobre la base de los métodos computacionales y los sistemas informáticos contemporáneos el cálculo llevaría unos 10^{16} años de tiempo de cálculo computacional. Diecisiete años más tarde, y gracias a un esfuerzo cooperativo a nivel mundial, sólo se tardaron ocho meses en completar la tarea.»

A *No hay conclusión.*

http://www.cwi.nl/en/RSA

D *Muchas oportunidades de reflexión: ¿qué pasará en el futuro? ¿Los computadores más rápidos finalmente eliminarán la capacidad de codificar mensajes? ¿Qué otros efectos podrían surgir?*

Bibliografía

"Cryptology timeline" ["*Cronología de la criptología*"] [en línea]
www.math.cornell.edu/~morris/135/timeline.html

"*An introduction to cryptography"* ["*Introducción a la criptografía*"] [en línea]
http://www.math.sunysb.edu/~scott/papers/MSTP/crypto/crypto.html

SINGH, Simon, "*The Code Book"* ["*El libro de los códigos*"], Fourth Estate, 2000

"*Derangements"* ["*Desarreglos*"] [en línea]
http://www.mathlab.mtu.edu/~eewestlu/ma3210_lecture19.pdf

"*Codes and ciphers*" ["*Códigos y cifrados*"] [en línea]
http://www.bletchleypark.org.uk/edu/teachers/ccresources.rhtm

[en línea] http://en.wikipedia.org/wiki/Cryptanalysis_of_the_Enigma

"Euler function and theorem" ["*La función y el teorema de Euler*"] [en línea]
http://www.cut-the-knot.org/blue/Euler.shtml

"Mathematical formulae" ["*Fórmulas matemáticas*"] [en línea]
http://www.po28.dial.pipex.com/maths/formulae.htm

[en línea] http://www.cwi.nl/en/RSA