María Isabel González Vasco

Las matemáticas de la criptología

SECRETOS DEMOSTRABLES Y DEMOSTRACIONES SECRETAS

colección
MIRADAS MATEMÁTICAS

DISEÑO DE CUBIERTA: ESTUDIO SÁNCHEZ/LACASTA

LAS MATEMÁTICAS DE LA CRIPTOLOGÍA.
SECRETOS DEMOSTRABLES Y DEMOSTRACIONES SECRETAS

ISBN: 978-84-9097-505-3
DEPÓSITO LEGAL: M-19.825-2018
IBIC: PDZ/GPJ

A Carlos, María y Nacho, mis usuarios honestos,
pero infinitamente curiosos.

A Palmira, depositaria de todos los secretos.

Índice

Introducción

Como casi todas las cosas importantes en la vida, la criptografía me encontró en un pupitre del colegio. Mi amiga Sara, sentada dos filas delante, tenía algo que contarme y escribía pequeños renglones en fragmentos de papel cuadriculado que intentaba hacerme llegar escondidos en su goma o en la caña de un bolígrafo. Sin duda, burlar la vigilancia de al menos dos potenciales entidades interesadas en nuestros mensajes (a la sazón, mi compañero de pupitre y don Primitivo, nuestro profesor de cuarto) era la principal razón de ser de este intercambio de información. Nuestra técnica fue, por tanto, ganando en sofisticación: una vez vimos claro que no bastaba con intentar esconder la existencia de la transmisión que nos ocupaba, comenzamos a pensar en ocultar no la presencia de los mensajes, sino su contenido.

Este tipo de procesos se repiten cada día en miles de aulas de educación primaria a lo largo del mundo. Esto no es sorprendente, pues la criptología, definida como "ciencia y práctica del diseño de sistemas de comunicación que son seguros en presencia de adversarios"[1], surge de modo inseparable de la necesidad de comunicación del ser humano. Existen

1. "Cryptology is the science and practice of designing computation and communication systems which are secure in the presence of adversaries", International Association for Cryptologic Research, www.iacr.org

evidencias históricas de métodos para la ocultación de la transmisión de la información (técnicas esteganográficas)[2] y esquemas de cifrado elementales contemporáneos de los lenguajes más antiguos conocidos. Por ejemplo, la escritura cuneiforme de los sumerios o el lenguaje jeroglífico de los egipcios son métodos de comunicación cuyo fin es transmitir información no de manera universal, sino a ciertos receptores autorizados. Un *esquema de cifrado*, la herramienta criptográfica más conocida y antigua, persigue exactamente eso: no el ocultar que existe transmisión de información, sino limitar el acceso a la información transmitida por medio de instrumentos matemáticos.

La criptología tiene dos vertientes bien diferenciadas: una constructiva y otra crítica o destructiva. La primera, llamada criptografía, se ocupa del diseño de herramientas, mientras que la segunda, el criptoanálisis, es el estudio crítico de las mismas. Criptógrafos y criptoanalistas llevan siglos enzarzados en una pugna cuyo resultado son construcciones cada vez más seguras para todo tipo de aplicaciones relacionadas con la gestión, transmisión y almacenamiento de información. Ambos buscan sus armas en un arsenal inagotable y precioso: las matemáticas. El papel de las matemáticas en criptología es central desde dos puntos de vista. Por un lado, como fuente de problemas cuya dificultad mantendrá bajo control a adversarios externos y usuarios maliciosos de un cierto sistema. Por otro, las matemáticas proporcionan el único lenguaje formal adecuado para la cimentación de demostraciones rigurosas e irrefutables de seguridad.

Hasta finales del siglo pasado, las construcciones criptográficas eran fundamentalmente esquemas de cifrado, diseñados para conseguir transmitir de manera segura información entre dos usuarios. En los últimos cuarenta años, sin embargo, las reglas del juego han cambiado radicalmente. La forma en

2. El término *esteganografía* proviene de las palabras griegas *steamos* ('oculto') y *graphos* ('escritura'). Se aplica a los métodos para el envío y entrega de mensajes camuflados dentro de un objeto o contenedor, de forma que no se detecte su presencia y consigan pasar desapercibidos.

que hoy compartimos, gestionamos y almacenamos la información plantea escenarios de aplicación fascinantes que suponen un reto constante para criptógrafos y criptoanalistas.

Este libro es una introducción a la criptología desde una perspectiva moderna. En cada capítulo presentaremos al lector un tipo de construcción criptográfica, profundizando en las herramientas matemáticas que pueden usarse para su implementación. Queremos acercar al lector a la criptología de manera amena y divulgativa, y presentarle además las ideas y conceptos matemáticos que subyacen en diferentes construcciones criptográficas. Nuestra ambición es proporcionar al lector un beneficio doble: aprender matemáticas a través de la criptología y desarrollar la inquietud por la criptología moderna desde el placer del formalismo matemático. Así, este libro proporciona a los profesores de educación secundaria algunos ejemplos novedosos con los que motivar a sus alumnos en el aprendizaje. Con ese fin plantearemos distintos retos o ejercicios sencillos, que pueden ser abordados con éxito por estudiantes de este nivel.

Capítulo 1

Criptografía simétrica. Al César lo que es del César

Volvamos atrás, quizá no tantos años, hacia un pasado sin internet, sin mensajería instantánea, sin móviles. Un pasado en el que los mensajes eran complejos en fondo y forma, redactados con paciencia y pulcritud. Entonces, los destinatarios eran pocos y selectos, y en la mayoría de los casos no se enviaba más de una carta cada mes, precedida de varios encuentros cara a cara. Este escenario, que parece rescatado de una infancia en blanco y negro, era el marco habitual de la comunicación hasta hace menos de treinta años. Es en cierta manera el ideal al que se aspira: cada día se intenta simular cientos de veces a través de la comunicación digital. La razón es sencilla; querríamos confiar en la autenticidad y confidencialidad de un archivo de datos recibido por *e-mail* igual que hacemos al recibir un sobre inmaculado (sin trazas de manipulación) por correo ordinario.

Este escenario en blanco y negro es el hábitat natural de la llamada *criptografía simétrica* (también llamada criptografía clásica o de clave privada). Parte del supuesto de que los interlocutores involucrados comparten un secreto *grande*, acordado en una fase previa completamente confiable, como podría ser un encuentro en persona. En este contexto, *grande* significa difícil de predecir o, en términos técnicos, de *alta entropía*. Actualmente, se considera que una cadena de ceros y unos (bits) es de alta entropía si su longitud es, al menos, de 180

bits. En contraposición, cadenas de 30 bits o menos suelen considerarse de baja entropía; ese es, por ejemplo, el caso de las contraseñas que somos capaces de memorizar. La entropía puede definirse rigurosamente como una medida para cuantificar la incertidumbre: en física, por ejemplo, sirve para medir el desorden, la desorganización de un sistema.

El secreto que comparten los dos usuarios, emisor y receptor, es la consigna que servirá para enviar mensajes entre ellos. Esta clave servirá tanto para cifrar (transformar el mensaje original en otro que, en caso de ser interceptado, sea inteligible si no se dispone de la clave) como para descifrar los mensajes (recuperar el mensaje original, a partir del mensaje cifrado, para poder leerlo).

En lenguaje matemático

Formalicemos matemáticamente esta idea. Nombramos, como es habitual en los libros de criptología, Alice a nuestra emisora y Bob al receptor. Todos los objetos involucrados en la transmisión viven en una estructura algebraica X, es decir, un conjunto en el que está definida una cierta operación * (por ejemplo, los números enteros y la operación suma). Tanto el conjunto de posibles mensajes sin cifrar (al que llamaremos M) como el de mensajes ya cifrados (llamémosle C) están contenidos en X. Distinguiremos un subconjunto especial, K, dentro de X, donde estarán las llamadas claves simétricas. Estas claves son los elementos indispensables para transformar un texto claro en un texto cifrado, y también para revertir esta transformación.

Formalmente, un esquema de cifrado se definirá haciendo explícitos tres procesos o *algoritmos*[3]:

3. Un algoritmo es un procedimiento descrito con exactitud para realizar una tarea en pasos bien diferenciados. Habitualmente, hay un valor o valores de entrada que determinan el desarrollo de dichos pasos. Tras el último paso, el algoritmo devuelve algún valor de salida. La notación A(*input*) = *output* se suele utilizar para expresar que el algoritmo A con valores de entrada *input* proporciona como salida los valores *output*.

- Un *algoritmo de generación de clave*, denotado KeyGen, que sirva para seleccionar cada una de las claves utilizadas a partir de un valor prefijado, llamado parámetro de seguridad. Este suele ser una medida de la fortaleza del esquema en cuestión.
- Un *algoritmo de cifrado*, que denotaremos Enc, que recibe como entrada una clave y un mensaje claro y da como salida un texto cifrado.
- Un *algoritmo de descifrado*, denotado por Dec, que recibe como entrada una clave y un texto cifrado y da como salida un mensaje claro.

Así, un esquema de cifrado Π = (KeyGen, Enc, Dec) será correcto si el algoritmo de descifrado permite recuperar el texto claro m que subyace a un texto cifrado c, siempre que se le dé como entrada la clave k que sirvió para construir c. En símbolos:

$$\text{Enc}(k, m) = c \Rightarrow \text{Dec}(k, c) = m$$

para toda clave k que haya sido seleccionada por KeyGen.

Alice será la encargada de ejecutar Enc, mientras que Bob tendrá que usar el algoritmo Dec para descifrar. En principio, si ambos usan la misma clave y nadie más la conoce, todo irá bien.

Un poco de historia

Un ejemplo clásico de cifrado simétrico usando la notación anterior es el *cifrado de César*. Gracias a los escritos de Suetonio[4], datados en el siglo segundo antes de Cristo, sabemos que el emperador romano Julio César empleaba un sencillo sistema para cifrar sus misivas más delicadas. Simplemente, César sustituía cada letra del alfabeto latino por la que ocupaba

4. La obra fundamental de Suetonio es *Vida de los Césares* (*De vita Caesarum*), que contiene las biografías de doce emperadores romanos, de Julio César a Domiciano.

tres posiciones más adelante, volviendo a las primeras letras desde el final para cifrar las tres últimas. Así, podemos decir que efectuaba una simple sustitución utilizando una tabla como esta:

Tabla 1
Cifrado de César

A	B	C	D	....	W	X	Y	Z
D	E	F	G		X	A	B	C

La primera fila contendría las letras del texto claro y la segunda el resultado tras cifrar. Así, el mensaje "TAMBIÉN TÚ, BRUTO" —obviando la coma y los acentos— quedaría cifrado como "WDPELHQ WX EUXWR".

Este método es muy sencillo, pero presenta ciertos problemas. El primero es que no hay variabilidad, es decir, al cifrar el mismo texto claro siempre obtendremos el mismo texto cifrado como resultado. Esta característica, llamada *determinismo*, es evidentemente una debilidad. Todo aquel que conozca el descifrado de un texto concreto podrá identificar las palabras contenidas en este que se repitan en envíos posteriores. Por ejemplo, la cadena "QR" en los textos cifrados siempre quiere decir "NO". Un espía que sepa que una respuesta "QR" de César es negativa, sabrá reconocer la palabra "NO" contenida en cualquier mensaje que se envíe, pues el cifrado de esta siempre será "QR".

En lenguaje matemático

Pasemos a describir formalmente el esquema de César para así entender de manera más general dónde está el problema. El conjunto X que nos sirve de estructura básica será el alfabeto latino, que codificaremos con los números del 0 al 25.

La operación de cifrado es, por tanto, sumar 3 al número que representa cada letra del texto claro. Pero al llegar a la última letra, se vuelve a empezar desde el comienzo. Así, la X

se cifraría con la A, la Y con la B y la Z con la C (como pasa con el reloj, al llegar al 12, se vuelve a pasar al 1). Esta operación se llama *suma módulo 26*. Operamos, por tanto, identificando cada número con el resto que da al dividirse por 26. En el caso concreto de sumar tres, cuando el resultado de la suma sea superior a 25, restamos a este 26 (para obtener, de nuevo, un número en el rango adecuado). Para manejarnos con este tipo de aritmética es útil pensar que el 26 se transforma en un cero. Así, al sumar o restar —las veces que queramos— 26 a un número X, nos quedamos en el mismo número.

Tabla 2

Codificación para el cifrado de César

A	B	C	D		X	Y	Z
0	1	2	3	....	23	24	25

Para descifrar, de nuevo, utilizamos la codificación numérica descrita en la tabla 2 y restamos a cada número recibido el 3, ajustando (sumando 26, si obtenemos un número negativo) para tener siempre números entre 0 y 25.

Así, nuestro conjunto X es el de los enteros {0, 1, …, 25}, y la operación de cifrado puede describirse como

$$\text{Enc}(m) = m + 3 \bmod 26$$

siendo el descifrado

$$\text{Dec}(m) = m - 3 \bmod 26$$

donde la expresión "mod 26" se traduce como "divide entre 26, y quédate con el resto". En efecto, estamos usando, en realidad, una estructura algebraica clásica, el grupo aditivo que se define en el conjunto de números enteros {0, …, n – 1} denotado por Z_n con la operación suma mod n anteriormente descrita (para n = 26).

Ejercicio 1. Ya hemos razonado que al trabajar con aritmética modular, el módulo (en el texto, el número 26) desempeña el papel del neutro para la suma. Utiliza esta idea para argumentar, sin hacer la división, que el resto de 52.002 al dividirse entre 26 es 2.

Y, en este caso, ¿cuál es la clave? Dicho de otra forma: ¿qué valor de $X = Z_{26}$ es imprescindible conocer para cifrar y descifrar? La respuesta es sencilla: el número 3. El tamaño del "salto" con el que ciframos y desciframos determina la clave y, además, se mantiene estable. Esto nos permite deducir que Julio César, o bien no tenía demasiadas ocasiones para acordar claves de cifrado con sus homólogos, o no tenía un elevado concepto del talento matemático de sus adversarios.

En lenguaje matemático

En términos modernos, una descripción de este método que nos dejaría más tranquilos sería una terna de algoritmos descritos como sigue:

- $\text{KeyGen}(n) = k \in Z_n$
- $\text{Enc}(k, m) = m + k \bmod n$
- $\text{Dec}(k, c) = c - k \bmod n$

El parámetro de seguridad, n, fijaría el tamaño del alfabeto utilizado en la comunicación y la seguridad del método dependería esencialmente del diseño del algoritmo KeyGen. Si para un n fijado (por ejemplo, n = 26) la salida de KeyGen fuese siempre el mismo número (otra vez, por no llevar la contraria al César, el 3), estaríamos de nuevo con una construcción determinista. Yéndonos al extremo contrario, forzaríamos que cualquier valor entre 0 y 25 tuviese la misma probabilidad de resultar como salida de KeyGen. Esta es, evidentemente, la mejor estrategia para ganar seguridad en un diseño de este tipo[5].

5. Malas noticias: en cualquier caso, ese esquema no es muy seguro.

Demos a la construcción recién descrita una vuelta más, concretamente, cifrando cada letra con un método como el anterior, pero con clave distinta. Esta evolución del cifrado de César se conoce con el nombre de *cifrado de Vigenère*, más conocido como el código indescifrable (*le chiffre indéchiffrable*, en francés).

Blaise de Vigenère

Blaise de Vigenère fue un diplomático, criptógrafo y químico francés del siglo XVI, que ejerció de secretario de la cámara del rey Enrique III de Francia. Curiosamente, él no inventó el cifrado que ha pasado a la historia con su nombre; este fue descrito, en realidad, por un criptógrafo italiano: Giovan Battista Belasso.

En lenguaje matemático

Veamos cómo funciona. Para cifrar mensajes de longitud t, esta sería la descripción de los algoritmos involucrados:

- KeyGen(n, t) = k = $(k_1, \ldots, k_t)$ con $k_i \in Z_n$ para i = 1, …, t
- Enc(k, m) = c
  siendo m = $(m_1, \ldots, m_t)$ y c = $(c_1, \ldots, c_t)$
  donde para i = 1, …, t
  $c_i = m_i + k_i \bmod n$
- Dec(k, c) = m,
  donde para i = 1, …, t,
  $m_i = c_i - k_i \bmod n$

Este método, del que se tiene constancia desde el siglo XVI, heredó la debilidad del método usado por Julio César. De nuevo, el algoritmo KeyGen utilizado era esencialmente determinista, pues la clave quedaba fijada a través de una tabla que, además, en muchos casos, se resumía utilizando una palabra clave corta. La tabla 3 ejemplifica este método, donde las claves asociadas $k = (k_1,\ldots, k_t)$ van de 0 a 25 en la primera columna de la izquierda. La tabla 4 representa la misma clave, pero sin codificación numérica[6].

Tabla 3

Tabla Vigenère numérica

Clave K	A	B	C	D		W	X	Y	Z
0	0	1	2	3		22	23	24	25
1	1	2	3	4		23	24	25	0
2	2	3	4	5		24	25	0	1
3	3	4	5	6		25	0	1	2
4	4	5	6	7	...	0	1	2	3
...						...			
23	23	24	25			19	20	21	22
24	24	25	0			20	21	22	23
25	25	0	1			21	22	23	24

Tabla 4

Tabla Vigenère alfabética

Clave K	A	B	C	D		W	X	Y	Z
0	A	B	C	D		W	X	Y	Z
1	B	C	D	E		X	y	X	A
2	C	D	E	F		Y	Z	A	B
3	D	E	F	G		Z	A	B	C
4	E	F	G	H	...	A	B	C	D
...						...			
23	X	Y	Z			T	U	V	W
24	Y	Z	A			U	V	W	X
25	Z	A	B			V	W	X	Y

6. Eliminamos la letra ñ de estas tablas.

Ejemplo 1

Así, para cifrar la palabra "FÁCIL", si utilizáramos la secuencia de clave (0, 1, 2, 3, 4), tendríamos como resultado el texto cifrado "FBELP". Si queremos que la secuencia de clave no sea siempre de la forma (0, 1, ..., t), podemos usar una palabra clave que señale qué filas de la tabla Vigenère se usarán para cada letra. Por ejemplo, si usamos la palabra clave "MISTERIO", estamos señalando a las filas 12, 8, 18, 19, 4, 17, 8 y 14. Así, por ejemplo, para cifrar la palabra "ALMA" haremos:

$m = (A, L, M, A)$, $k = (12, 8, 18, 19)$ y así, el texto cifrado $c = (c_1, ..., c_t)$ se construirá con la fórmula

$$c_i = m_i + k_i \bmod 26$$

de donde obtenemos, codificando numéricamente el texto claro m como (0, 11, 12, 0):

- $c_1 = 0 + 12 \bmod 26 = 12$, y en letras $c_1 = M$
- $c_2 = 11 + 8 \bmod 26 = 19$, y en letras $c_2 = T$
- $c_3 = 12 + 18 \bmod 26 = 4$, y en letras $c_3 = E$
- $c_4 = 0 + 19 \bmod 26 = 19$, y en letras $c_4 = T$

Así, el resultado del cifrado es la palabra "MTET". En este caso, el texto que ciframos tenía menor longitud que la palabra clave. Si la longitud del texto claro es mayor que la palabra clave, habitualmente se repite la palabra clave el número de veces necesario (es decir, después de la O de "MISTERIO", empezamos de nuevo por la M).

Ejercicio 2. Utiliza las tablas 3 y 4 para cifrar la palabra SECRETO con la palabra código LUNA.

Regreso al futuro

Ahora que conocemos dos ejemplos de cifrado simétrico históricos, avancemos unos siglos. ¿Qué aspecto tienen los esquemas actuales? En realidad, su estructura formal sigue

encajando en la definición que hemos dado como una terna de algoritmos. Existen, sin embargo, enormes diferencias, que tienen mucho que ver con los avances en computación del siglo XX y con el (apasionante) desarrollo del criptoanálisis durante la Segunda Guerra Mundial.

Una de las piezas esenciales de muchos de los esquemas de cifrado simétrico moderno son las llamadas *redes de Feistel*. Horst Feistel fue un criptógrafo de IBM que pasó a la historia como diseñador de uno de los métodos de cifrado más utilizados mundialmente, el cifrado DES[7]. No nos detendremos en las entrañas de este diseño, sino en su esqueleto básico, que es el armazón que sustenta otros muchos esquemas (vivos y muertos), como Lucifer, Gost, Simon o Camellia.

Una red de Feistel sirve para cifrar bloques de texto claro, es decir, cadenas de bits de un cierto tamaño t, transformándolas en bloques de texto cifrado del mismo tamaño. Su principal ventaja es la llamada *reversibilidad*: las operaciones de cifrado y descifrado siguen exactamente la misma estructura, realizan las mismas operaciones, pero involucrando la clave en "orden inverso".

En lenguaje matemático

Veamos una descripción formal (aunque simplificada) del proceso: tenemos como antes el texto claro, la clave y el texto cifrado dividido en bloques. Vamos a ver cómo se opera en un solo bloque de texto claro M, que va a ser una cadena de r bits (con r par). Tenemos también una clave K fragmentada en n claves parciales (K_1, …, K_n), cada una de r/2 bits y, además, una misteriosa función f que servirá para "mezclar" trozos de M con claves parciales. Podemos considerar que tanto n como r forman el parámetro de seguridad, y la clave $K = (K_1, \ldots, K_n)$ es el resultado de ejecutar el algoritmo KeyGen.

7. Las siglas DES proceden de la denominación en inglés del cifrador, llamado *Data Encryption Standard*, por ser declarado el estándar para cifrado simétrico en Estados Unidos de 1977 a 1999. Su sucesor, el algoritmo AES (*Advanced Encryption Standard*), no utiliza una red de Feistel como base.

En lo que sigue, denotaremos por $\oplus$ la suma de cadenas binarias en base dos, es decir, atendiendo a las reglas

$$1 \oplus 1 = 0,\ 0 \oplus 0 = 0,\ 1 \oplus 0 = 1,\ 0 \oplus 1 = 1$$

Vamos a cifrar el texto M con la clave K; es decir, veamos cómo se construye la salida del algoritmo Enc si la entrada es el par (K, M):

- Primero, se divide M en dos mitades (cada una de r/2 bits), denotadas L_0 y R_0 (parte izquierda y parte derecha).
- Después, para i = 1 hasta n se define:
  $L_i := R_{i-1}$
  $R_i := L_{i-1} \oplus f(R_{i-1}, K_i)$
- Por último, se construye la salida:
  $C := (L_n, R_n)$

Lo mejor es que, independientemente de cómo se defina f, el algoritmo de descifrado hace esencialmente lo mismo que Enc, pero ordenando las claves al revés, es decir, Dec(K, C):

- Divide C en dos mitades (cada una de r/2 bits), denotadas L_n y R_n (parte izquierda y parte derecha).
- Después, define para i = n hasta 1 (con paso decreciente) los valores:
  $R_{i-1} := L_i$
  $L_{i-1} := R_i \oplus f(L_i, K_i)$
- Proporciona la salida: $M = (L_0, R_0)$

Ejercicio 3. Escribe cómo se cifran y descifran mensajes con una red de Feistel en la que f(X,Y):= X $\oplus$ Y. Considera r = 4 y n = 3. Cifra y descifra el mensaje 0111 con clave K = (11 00 10).

Cada paso (designado con la letra i en la instrucción "para" anterior) en el cifrado de una red de Feistel suele

llamarse ronda, y por eso la función f (que en muchos diseños no es la misma para todas las rondas) suele llamarse *función de ronda*. La seguridad de los cifrados basados en redes de Feistel depende, en gran medida, de la estructura de la función de ronda subyacente.

¿Es mi esquema seguro?

Ahora que sabemos un poco sobre esquemas de cifrado simétrico, es momento de pasar a una pregunta importante: ¿cómo se evalúa su seguridad? Si puedo elegir, ¿qué sistema utilizo para enviar mis mensajes? Para responder a esta pregunta, lo primero es que pensemos detenidamente en la razón por la que utilizamos un esquema de cifrado. Sara, en el colegio, lo tenía muy claro: para que nadie más que yo (la legítima receptora de sus envíos) pudiera leer el contenido de los mensajes. Ni siquiera nuestro estupendo profesor de cuarto, don Primitivo, que, desde luego, sabía muchas más matemáticas que nosotras. Pero ¿es aceptable que alguien que intercepte el mensaje pueda obtener alguna información sobre su contenido, aunque sea pequeña? Y ¿podemos suponer que el presunto espía no conoce qué método concreto estamos utilizando para cifrar? Sin dar una respuesta precisa a estas preguntas, nuestro análisis de seguridad será incompleto (y, en muchas ocasiones, peligroso). Volveremos sobre estas preguntas y sus posibles respuestas en numerosas ocasiones.

El principio de Kerckhoffs

A finales del siglo XIX, el lingüista y criptógrafo holandés Auguste Kerckhoffs publicaba en su artículo "La cryptographie militaire" los seis principios fundamentales que debía cumplir todo sistema criptográfico. El segundo de ellos ha

pasado a la historia como *principio de Kerckhoffs,* y señala que la seguridad de un sistema criptográfico nunca debe depender de que su diseño se mantenga en secreto. En definitiva, esto significa que siempre hemos de suponer que un adversario conoce absolutamente todo lo relativo a un sistema criptográfico, con excepción de aquellos valores señalados explícitamente como *claves secretas.*

EJEMPLO 2

Pero retrocedamos de nuevo a los tiempos de Julio César. Pensemos cuál sería nuestra estrategia, como adversarios, si interceptamos un texto cifrado con este método cuya clave desconocemos (pues Julio César, en algún momento, ha abandonado su querido 3). Este es nuestro texto cifrado objetivo:

LS WLYYV KL ZHU YVXBL UV APLUL YHIV WVYXBL YHTVU YVKYPNBLG ZL SV OH JVYAHKV

Si sabemos que el texto claro está en castellano, podemos intentar averiguar la clave analizando la frecuencia de aparición de las distintas letras en el texto. En castellano, las letras más frecuentes son la e y la a. En este texto, las letras que más se repiten son la L y la V (ambas aparecen 9 veces). Si asumimos que la L cifra la E, la clave utilizada sería K = 7, y en caso de pensar que es la V quien cifra la E, deduciríamos que K vale 17. Al probar con las dos claves, inmediatamente vemos que K = 17 da un texto inteligible, mientras que con K = 7 obtenemos la frase (obviando los acentos):

EL PERRO DE SAN ROQUE NO TIENE RABO PORQUE RAMÓN RODRÍGUEZ SE LO HA CORTADO

Ejercicio 4. Intenta tú ahora descifrar de manera similar el texto:
BR JURLN H KXK BDYRNAJW VJB MN LARYCXPAJORJ, BNPDAJVNWCN ZDNMJARJW YJAJ CXVJA DW LJON H QJKUJA

Ahora bien, ¿por qué es tan efectivo este sencillo análisis de frecuencias? ¿Ocurriría lo mismo con el cifrado de Vigenère? Evidentemente, el éxito de nuestro método radica en el determinismo del cifrado de César: el cifrado de una letra siempre da como resultado la misma. Vigenère no tiene esta debilidad, pero, desde luego, va a distar mucho de ser un cifrado indescifrable[8]; por ejemplo, un análisis de frecuencias puede ayudarnos a detectar la longitud de la palabra clave con la que se ha cifrado el texto. Conocido esto, solo tenemos que separar los caracteres cifrados con cada fila de la tabla 4 para luego averiguar, como hicimos en el texto anterior, con qué desplazamiento concreto se corresponde cada una.

Así, el análisis de frecuencias parece una herramienta poderosa; pero solo lo es contra los esquemas clásicos con cierto grado de determinismo. Un cifrado construido con una red de Feistel robusta es mucho más difícil de romper, aunque, por supuesto, todo depende de la frecuencia con la que "refresquemos" las claves criptográficas. Pensemos cuál sería el escenario ideal, por ejemplo, en el cifrado Vigenère. La respuesta es sencilla: lo ideal es no reutilizar ninguna clave y, por tanto, cifrar cada letra con una fila distinta de la tabla. Esto solo es posible si nuestros mensajes son, a lo sumo, del mismo tamaño (t) que la clave asociada.

Hemos descubierto uno de los principios que subyacen a la seguridad: reutilizar *lo menos posible* cada clave secreta, que además se elige uniformemente al azar[9].

El llamado *cifrado Vernam* o *one-time pad* es el primer esquema de cifrado simétrico para el que existe una demostración formal de seguridad. Se basa justamente en utilizar claves de un solo uso del mismo tamaño del texto que se va a

8. Con gran tristeza, es obligado contradecir aquí a Lewis Carroll, quien, junto a otros muchos matemáticos, científicos y criptógrafos, utilizó el término *cifrado indescifrable* para referirse a Vigenère.
9. Informalmente, elegir un elemento uniformemente al azar en un conjunto implica que todos los elementos del conjunto tengan la misma probabilidad de ser elegidos. Así, tirando una moneda perfecta al aire elegimos uniformemente al azar entre cara y cruz.

cifrar. Las construcciones actuales basadas en redes de Feistel intentan, en cierto sentido, emular la utilización de claves de un solo uso. Otro recurso para esquivar el poderoso análisis de frecuencias es utilizar una codificación de mensajes cuyos patrones no sean tan transparentes como el alfabeto latino en castellano. A lo largo de la historia, se han utilizado lenguas muertas o muy poco conocidas para codificar mensajes que luego se cifrarían. Por ejemplo, el idioma navajo fue utilizado por los americanos durante la Segunda Guerra Mundial (como puede verse en la película *Windtalkers*, dirigida por John Woo). En cierto sentido, confiar en que este tipo de estrategia robustezca una construcción criptográfica va contra el principio de Kerckhoffs.

Pero ¿acaso no hay alguna forma mejor de esquivar el fastidioso análisis de frecuencias? ¿Podríamos de alguna manera refrescar continuamente las claves sin tener que tatuarlas en el cuero cabelludo de cientos de soldados etruscos u obligar a valientes agentes de inteligencia a cruzar la Alsacia en tren noche tras noche? La respuesta, que supuso el fin de muchos quebraderos de cabeza para esforzados criptógrafos (y el comienzo de otros muchos), aguarda en el próximo capítulo.

Proyectos para saber más

Recopila información sobre los siguientes métodos criptográficos y escribe un informe analizando informalmente su seguridad:

- Cifra general de Felipe II.
- Cifrado utilizando la máquina Enigma.