ACADEMIA DE CIENCIAS DE MORELOS, A.C.

¿Comentarios y sugerencias?, ¿Preguntas sobre temas científicos? CONTÁCTANOS:
editorial @acmor.org.mx

Breve Historia de la Criptografía: Segunda Parte

L. Enrique Sucar
Instituto Nacional de Astrofísica, Óptica y Electrónica
Miembro de la Academia de Ciencias de Morelos
Edgar A. Sucar
Estudiante de Preparatoria, ITESM-Puebla

La criptografía consiste en alterar la representación de un mensaje para que no pueda ser entendido a menos que se conozca la llave secreta. En la primera parte de esta historia (La Unión de Morelos: 28/3/2011 - http://www.acmor.org.mx/descargas/11_mar_28_criptografia.pdf) vimos el *cifrado por sustitución*, que consiste en reemplazar cada letra en el mensaje por otra, de acuerdo a cierta clave en la que se indica la correspondencia entre cada letra en el mensaje original con otra letra en el mensaje cifrado; y así es como codificamos un mensaje secreto: "UIUOPYTU OP ITPZITUE OP YADPXAE". Vamos ahora a descifrarlo como si no tuviéramos acceso a dicha clave.

El método para descifrar este tipo de códigos se basa en la frecuencia de las letras en cierto idioma, en este caso el español. Comparemos entonces la frecuencia de las letras en español contra la frecuencia de las letras en el mensaje, como se muestra en la Tabla 1. Como este mensaje es muy corto, la frecuencia de las letras no es necesariamente la misma que la del español; pero para poder ilustrar el método de decodificación, supondremos que es parecida, lo cual es más probable mientras más mensajes tengamos a la mano (cifrados de la misma forma) o que éstos sean más largos. Las dos letras más comunes en el español son las vocales "e" y "a", por lo que podríamos asumir que corresponden con las dos más frecuentes en el código, "P" y "U", respectivamente. Reemplazando estas dos letras en el mensaje codificado obtenemos: "alaOeYTa Oe ITeZITaE Oe YADeXAE" (usaremos minúsculas para distinguir las letras decodificadas de las originales que mostraremos en mayúsculas).

Español	e	a	o	s	r	n	i	d	l	c
Mensaje	P	U	T	I	O	Y	E	A	Z	D

Tabla 1: Las 10 letras más comunes en el español y en el mensaje ordenadas de izquierda a derecha por frecuencia.

Una de las palabras de dos letras más comunes en español es "de", por lo que suponemos que "Oe" es "de", y entonces la "O" corresponde a "d". Sustituyendo ahora en el mensaje obtenemos: "aladeYTa de ITeZITaE de YADeXAE". La palabra "aladeYTa" parece conocida (normalmente entendemos una palabra o frase aunque varias letras sean desconocidas), podemos adivinar que es "academia". Si estamos en lo correcto entonces "I" es "c", "Y" es "m" y "T" es "l". Reemplazando ahora estas letras en el mensaje: "academia de cieZciaE de mADeXAE". La tercera palabra del mensaje es fácilmente reconocible, "ciencias", entonces "Z" es "n", "E" es "s"; ahora el mensaje es: "academia de ciencias de mADeXAs". Por el contexto, podemos adivinar que la última palabra es "Morelos", con lo que "A" es "o", "D" es "r", y "X" es "l". Con lo que finalmente hemos descifrado el mensaje: "Academia de Ciencias de Morelos". La llave completa se muestra en la Tabla 2. Para que sea más fácil recordar la llave comúnmente se usa una palabra al inicio, en este caso "union" (eliminando las letras repetidas), y el resto de las letras se ponen simplemente en orden alfabético a partir de la última letra de la palabra inicial.

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

Tabla 2: La llave del código secreto para el ejemplo. El primer renglón muestra las letras en español del mensaje original; en el segundo renglón están las letras correspondientes que se utilizan en el mensaje cifrado de acuerdo a la llave.

Como muestra el ejemplo anterior, el *cifrado por sustitución* no es muy seguro, por lo que se desarrollaron códigos más complejos que no eran tan fáciles de descifrar; pero el punto débil seguía siendo el que todos los métodos dependen de una llave secreta, quien conozca la llave puede descifrar el mensaje. Este problema se volvió más crítico con el desarrollo de Internet, ya que muchas aplicaciones dependen de comunicaciones seguras, como la banca y comercio electrónicos. En un principio se usaban servicios de mensajería para enviar las llaves secretas, pero esto se volvió impráctico y muy costoso. La única solución era algo que parecía imposible: ¡Un método de criptografía con el cual no se pudiera descifrar el mensaje aunque se conociera la llave! Esto se conoce como *criptografía de llave pública*.

El desarrollo de la criptografía de llave pública es una fascinante historia de perseverancia e ingenio de dos grupos de investigadores, uno en Estados Unidos y otro en Gran Bretaña. Este importante tipo de criptografía basado en la teoría de números constituye un excelente ejemplo de cómo ciertas ramas de las matemáticas que aparentemente no tienen aplicación práctica, pueden volverse fundamentales para el desarrollo de tecnologías que ahora nos son indispensables.

En la actualidad los mensajes se codifican utilizando las computadoras, donde cada letra se representa mediante un número (utilizando el código ASCII, por ejemplo), así que un mensaje es simplemente una concatenación de números; es decir, un número más grande. Por ejemplo, la palabra "ADIOS" es equivalente al número "6568737982". Convertidos en números, los mensajes se podrían cifrar mediante operaciones aritméticas. La llave en este caso está definida por un número y ciertas operaciones aritméticas, supongamos que el número es "2" y la operación es multiplicación. Así podríamos cifrar "ADIOS" multiplicándolo por 2 (6568737982 x 2 = 13137475964). El receptor simplemente dividiría entre 2 para descifrar el mensaje.

La criptografía de llave pública se basa en dos ideas principales. La primera es la idea de la llave asimétrica: la llave para codificar el mensaje es diferente de la llave para decodificarlo. De esta forma puedes hacer pública la llave para cifrar los mensajes secretos que te vayan a enviar, pero sólo tú conoces la llave secreta para descifrarlos. La segunda idea clave es el uso de ciertas funciones matemáticas que son fáciles de aplicar en un sentido pero difíciles en el sentido opuesto, llamadas funciones de un solo sentido. Por ejemplo, el producto de dos números *primos* (sólo divisibles entre uno y el mismo número) es fácil de obtener, pero dado el producto, es difícil obtener los factores. La criptografía de llave pública también hace uso de la *aritmética modular*, en la cual se representa el número resultante de una operación aritmética mediante el residuo que se obtiene al dividirlo entre un número conocido como módulo. Todos usamos aritmética modular cuando hacemos operaciones con las horas que señala un reloj, para las cuales él módulo es 12. Por ejemplo, si son las 10 y vamos a reunirnos 5 horas después, nos veremos a las 3, pues (10 + 5) modulo 12 = 3. Es decir, sumamos 10 con 5 y nos quedamos con el residuo que resultaría de dividir entre 12.

La criptografía consiste en alterar la representación de un mensaje para que no pueda ser entendido a menos que se conozca la llave secreta.

El primer sistema de criptografía de llave pública es conocido como RSA (por las iniciales de sus autores: Rivest, Shamir y Adleman). El sistema RSA se basa en el uso de llaves asimétricas y funciones de un solo sentido, utilizando operaciones matemáticas para el cifrado y descifrado del mensaje. A continuación presentamos una descripción general del método, sin pretender detallar los fundamentos matemáticos que requieren un conocimiento más profundo de una rama de las matemáticas conocida como teoría de números.

Para cifrar los mensajes, el dueño de la llave debe seleccionar dos números primos (de preferencia muy grandes), *p* y *q*, los cuáles deben mantenerse secretos. El producto, $n = p \times q$, es parte de la llave pública y por lo tanto no es secreto. Se asume que el mensaje a ser cifrado se representa como un número, si este número es mayor que *n*, se debe dividir en pedazos o bloques, cada uno menor a *n*. Además de

ACADEMIA DE CIENCIAS DE MORELOS, A.C.

¿Comentarios y sugerencias?, ¿Preguntas sobre temas científicos? CONTÁCTANOS:
editorial @acmor.org.mx

n, se debe seleccionar otro número, e (el cual no debe tener factores comunes con *(p-1) x (q-1)*). Cada bloque, *b*, se codifica aplicando las siguientes operaciones: $E(b)=b^e$ modulo *n*, donde *E(b)* es el bloque codificado.

Por ejemplo, si seleccionamos *p=149 y q=157, n = p x q =23,393*; y *e* puede ser 5. Entonces, si suponemos que nuestro mensaje es corto y tiene un solo bloque que es "*20,232*", el mensaje cifrado sería: $E(20,232)=20,232^5$ *modulo* 23,393 = 20,036.

Para descifrar el mensaje se aplica la siguiente fórmula: $D(a)=a^d$ modulo *n*, donde *n* es el producto de *p* y *q* que ya conocemos. Pero para obtener *d* necesitamos conocer *p* y *q*, y luego aplicar otra fórmula: $d=e^{-1}$ modulo *(p-1) x (q-1)*. Continuando con el ejemplo, $d=5^{-1}$ modulo 23,088=13,853, y por lo tanto, b = $20,035^{13,853}$ modulo 23,393=20,232, que coincide con nuestro mensaje original.

Bajo este enfoque, quien va a recibir los mensajes secretos, por ejemplo la Academia, establece la llave privada (*p*, *q*) y da a conocer a todos la llave pública (*n*, *e*). Cualquiera puede cifrar mensajes con la llave pública, pero sólo la Academia, que conoce la llave privada, puedes descifrarlos. Pero que tal si algún espía logra obtener los factores de *n*, *p* y *q*, utilizando una computadora muy poderosa, ¡entonces lograría romper nuestro código y descifrar el mensaje!

La seguridad del método RSA se basa en que si los números p y q son suficientemente grandes, le tomaría a la computadora más rápida del mundo muchísimo tiempo para obtener los factores de *n*. Entonces, aunque en principio es posible obtener la llave privada, en la práctica es muy, muy difícil. Así que gracias a esta técnica podemos actualmente transmitir en forma práctica y segura información delicada, como números de tarjetas de crédito, a través de Internet.

Como mencionamos en la primera parte, la historia de la criptografía es básicamente una competencia entre los que diseñan los códigos y los que intentan descifrarlos. Actualmente tienen la ventaja quienes diseñan los códigos, mientras no se encuentre una forma eficiente de encontrar los factores de un número.

Finalmente, la criptografía ha mostrado cómo la teoría de números, que por muchos años se consideró sin aplicación práctica alguna, ha resultado indispensable para resolver un problema tecnológico de gran trascendencia.

Para los lectores interesados en conocer más sobre la criptografía y su fascinante historia recomendamos el libro por Simon Singh, "The Code Book", *Anchor Books*, 2000.

Para actividades recientes de la Academia y artículos anteriores puede consultar: www.acmor.

CARTELERA CINES

LISTA DE FOTOGRAFIAS DE CANDIDATOS DEL PARTIDO VERDE

DIANA
CASA DE MI PADRE 11:45 / 13:50 / 15:35 / 17:35 / 19:20 / 21:20
EL CUERVO 11:15 / 13:35 / 15:50 / 18:15 / 20:35 / 22:50
SIN SALIDA 12:15 / 14:20 / 16:15 / 18:10 / 20:15 / 22:25
BATALLA NAVAL ING 12:50 / 15:40 / 18:20 / 21:10
THE AVENGERS 3D ING 11:00 / 14:00 / 17:00 / 19:55 / 22:55
THE AVENGERS ESP DIGITAL 10:30 / 13:30 / 16:30 / 19:30 / 22:30
THE AVENGERS ESP 12:30 / 15:25 / 18:35 / 21:45
THE AVENGERS 3D ESP 12:00 / 15:05 / 17:55 / 21:00
BATALLA NAVAL ESP DP 11:30 / 14:10 / 16:45 / 19:35 / 22:15
TRAVESIA DEL DESIERTO 11:05 / 13:40 / 16:10 / 18:50 / 21:25
CUANDO TE ENCUENTRE 12:20 / 14:40 / 16:55 / 19:05 / 21:30
PLAN PERFECTO 16:35 / 18:45 / 20:55 / 23:05
CRISTIADA 2P 10:50 / 13:45

JACARANDAS
THE AVENGERS 3D ESP 12:00 / 15:00 / 18:00 / 21:00
CASA DE MI PADRE 11:40 / 13:40 / 15:40 / 17:25 / 19:40 / 21:35
THE AVENGERS ESP DIGITAL 10:55 / 14:00 / 17:00 / 20:00 / 23:00
BATALLA NAVAL ESP 11:00 / 13:50 / 16:40 / 19:30
BATALLA NAVAL ING 1U 22:20
BATALLA NAVAL ESP (ILOCK S4) 12:25 / 15:15 / 18:05 / 20:55
CRISTIADA 1A Y 3A 11:45 / 16:25
APARTAMENTO 143 79 14:35 / 19:15 / 21:10 / 22:55
THE AVENGERS ESP 13:00 / 16:00 / 19:00 / 22:00
CUANDO TE ENCUENTRE 11:10 / 13:20 / 15:30 / 17:45 / 19:50 / 22:10
EL CUERVO 11:15 / 13:30 / 15:45 / 18:10 / 20:20 / 22:40
SIN SALIDA 12:35 / 14:25 / 16:30 / 18:35 / 20:30 / 22:30

CINEMEX CUAUTLA
PLAN PERFECTO 11:35 / 13:40 / 15:45 / 17:50 / 19:55 / 22:00
THE AVENGERS 3D ESP 11:45 / 14:35 / 17:25 / 20:15 / 23:00
TRAVESIA DEL DESIERTO 12:10 / 15:05 / 18:00 / 21:05
BATALLA NAVAL ESP DIF 11:25 / 14:05 / 16:45 / 19:25
BATALLA NAVAL ING 1U 22:15
BATALLA NAVAL ESP (ILOCK S4) 12:45 / 15:25 / 18:05 / 20:45
THE AVENGERS ESP (ILOCK S7) 11:00 / 13:50 / 16:40 / 19:30 / 22:20
THE AVENGERS ESP DIF 12:25 / 15:15 / 18:05 20:55
THE AVENGERS 3D ESP 13:10 / 16:00 / 18:50 / 21:40
CUANDO TE ENCUENTRE 11:15 / 13:20 / 15:35 / 17:40 / 19:45 / 21:50
SIN SALIDA 12:30 / 14:20 / 16:10 / 18:15 / 20:05 / 21:55
CASA DE MI PADRE 11:10 / 13:00 / 14:50 / 16:35 / 18:30 / 20:20
APARTAMENTO 143 79 1U 22:30
EL CUERVO DP 12:00 / 14:15 / 16:25 / 18:40 / 20:50