TECNOLÓGICO DE MONTERREY

Introducción al Criptoanálisis

Roberto Gómez Cárdenas
rogomez@itesm.mx

TECNOLÓGICO DE MONTERREY

Un punto de vista…

Breaking a cipher doesn't necessarily mean finding a practical way for an eavesdropper to recover the plaintext from just the ciphertext. In academic cryptography, the rules are relaxed considerably. Breaking a cipher simply means finding a weakness in the cipher that can be exploited with a complexity less than brute-force. Never mind that brute-force might require 2^{128} encryptions; an atack requiring 2^{110} encryptions would be considered a break. Breaks might also require unrealistic amounts of known or chosen plaintext -- 2^{56} blocks or unrealistic amounts of storage: 2^{80}.Simply put, a break can just be a "certificational weakness": evidence that the cipher does not perform as advertised.

--Bruce Schneier;
from his "Self-Study Course in Block Cipher Cryptanalysis"

Tipos ataques criptográficos

Clasificación de acuerdo a los datos que se requieren para el ataque.

- Ciphertext only attack
- Known-Plaintext attack
- Chosen text attack
  - Chosen plaintext Attack
  - Chosen ciphertext Attack
  - Adaptive Chosen Plaintext Attack
  - Adaptive Chosen Ciphertext Attack

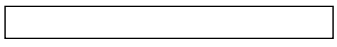
Lámina 3 Dr. Roberto Gómez Cárdenas

Ciphertext only attack

Dado:

criptograma 

Se busca por

texto claro
o llave 

Ejemplo

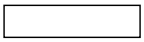
Análisis de frecuencia en el criptograma

Lámina 4 Dr. Roberto Gómez Cárdenas

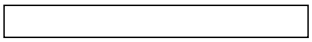
Known-Plaintext Attack

Dado:

criptograma 

un fragmento del texto claro 

Se busca por

resto texto claro o llave 

Ejemplo

Búsqueda exhaustiva de llave (ataque de fuerza bruta)

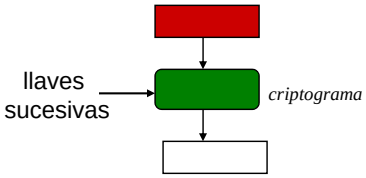


Lámina 5 Dr. Roberto Gómez Cárdenas

Ataques contraseñas

- Ataques por diccionario
- Ataques por fuerza bruta

Lámina 6 Dr. Roberto Gómez Cárdenas

Una alternativa

- Calcular todos los valores hash generados al cifrar todas las posibles contraseñas y almacenarlos en una tabla.
- Una vez construida, un atacante solo tendría que consultar la tabla para encontrar la contraseña asociada a un hash.

Lámina 7 Dr. Roberto Gómez Cárdenas

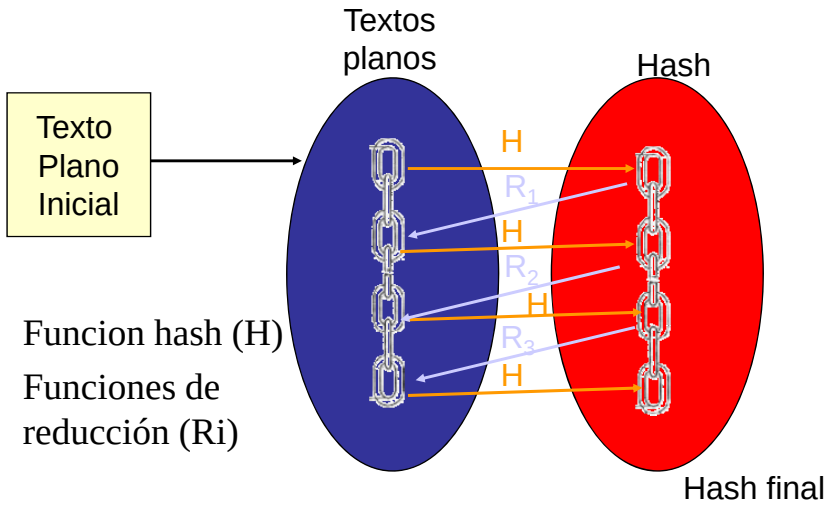
Dos “problemitas”

- El tiempo que toma llevar a cabo todo el cálculo
  - Computo paralelo o distribuido
- El espacio de almacenamiento
  - Almacenar solo una parte de toda la información, de tal forma que a partir de una almacenada se deduzca una no almacenada

Lámina 8 Dr. Roberto Gómez Cárdenas

TECNOLÓGICO DE MONTERREY

Generando cadenas



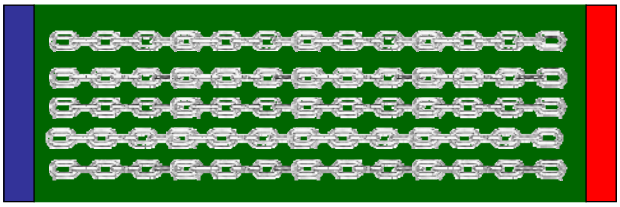
- Funcion hash (H)
- Funciones de reducción (Ri)

Lámina 9
Dr. Roberto Gómez Cárdenas

TECNOLÓGICO DE MONTERREY

Al final

Textos Planos Iniciales

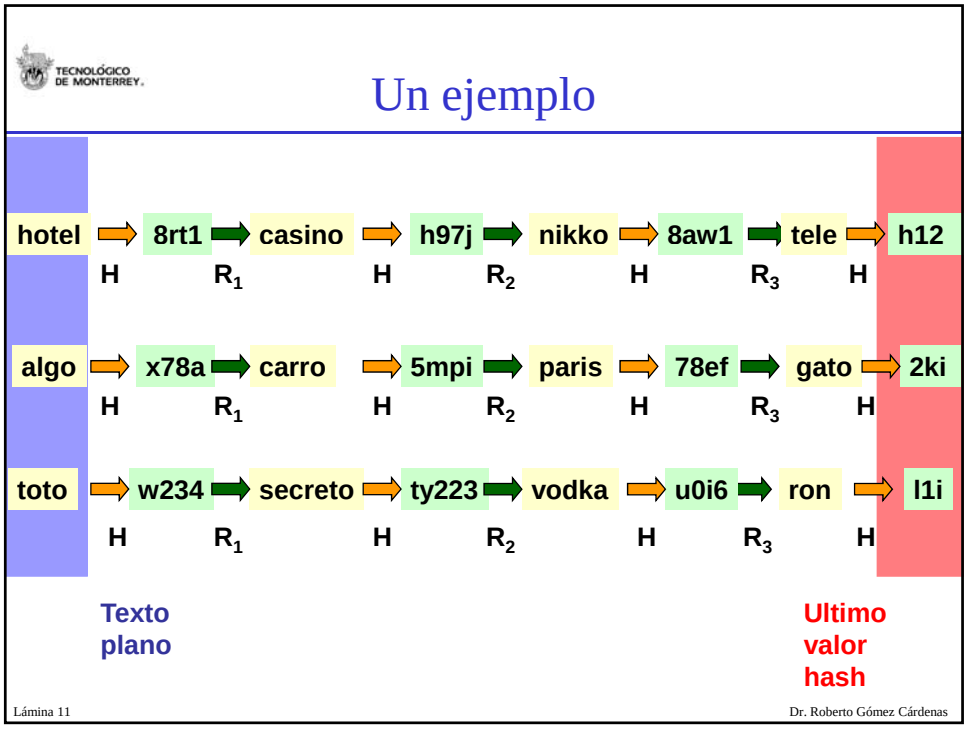


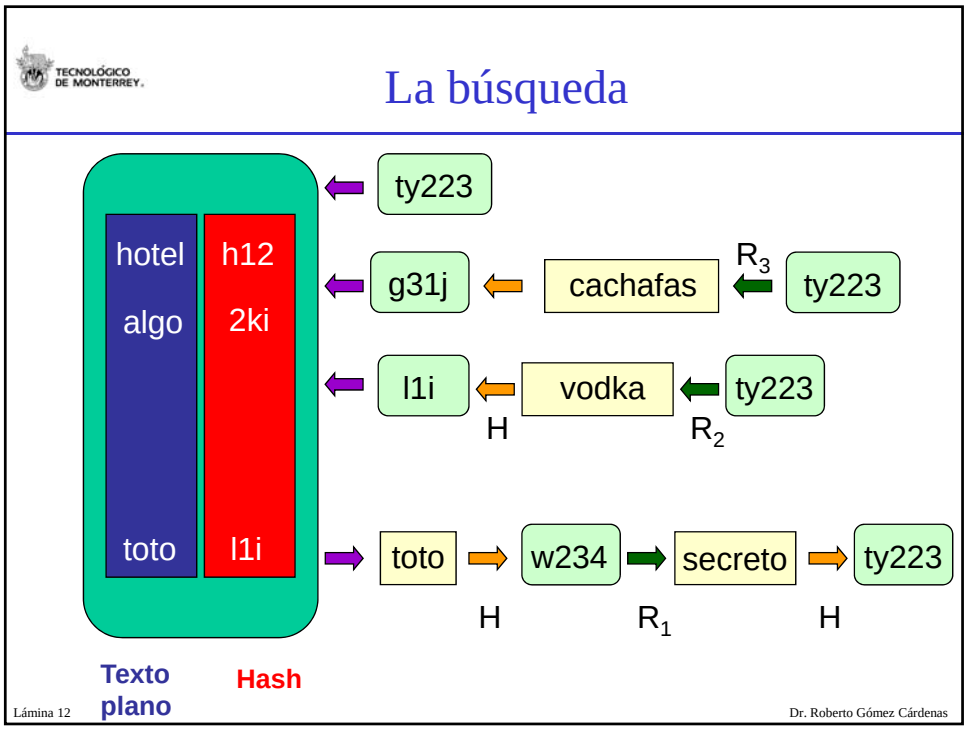
Hashes finales

estos hashes no son almacenados
pero pueden ser generados y buscados

iaisudhiu -> 4259cc34599c530b1e4a8f225d665802
oxcvioix -> c744b1716cbf8d4dd0ff4ce31a177151
9da8dasf -> 3cd696a8571a843cda453a229d741843
[...]
sodifo8sf -> 7ad7d6fa6bb4fd28ab98b3dd33261e8f

Lámina 10
Dr. Roberto Gómez Cárdenas


TECNOLÓGICO DE MONTERREY
Un ejemplo
hotel
8rt1
casino
h97j
nikko
8aw1
tele
h12
H
R1
H
R2
H
R3
H
algo
x78a
carro
5mpi
paris
78ef
gato
2ki
H
R1
H
R2
H
R3
H
toto
w234
secreto
ty223
vodka
u0i6
ron
l1i
H
R1
H
R2
H
R3
H
Texto plano
Ultimo valor hash
Lámina 11
Dr. Roberto Gómez Cárdenas


TECNOLÓGICO DE MONTERREY
La búsqueda
hotel
algo
toto
h12
2ki
l1i
ty223
g31j
cachafas
R3
ty223
l1i
vodka
ty223
H
R2
toto
w234
secreto
ty223
H
R1
H
Texto plano
Hash
Lámina 12
Dr. Roberto Gómez Cárdenas

Funciones reducción y hash

- Uno de los secretos de esta técnica se encuentra en la funciones de reducción.
- Recordemos que esta función convierte una cadena de caracteres en un conjunto de bits que representa un valor hash.

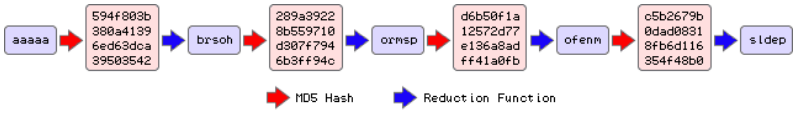


Lámina 13 Dr. Roberto Gómez Cárdenas

Implementaciones

- El proyecto RainbowCrack
  - http://www.antsight.com/zsl/rainbowcrack
- La herramienta Ophcrack
  - SourceForge
  - http://ophcrack.sourceforge.net/es.index.php

Lámina 14 Dr. Roberto Gómez Cárdenas

¿Por qué arcoiris?

- Cada una de las columnas usa una función de reducción diferente.
- Si cada función reducción fuera de un color diferente y el texto plano se pone en la parte superior y el hash abajo.
  - Se vería como un arcoiris

Lámina 15
Dr. Roberto Gómez Cárdenas

Observaciones

- Aun se requiere de espacio
  - Pero no tanto como el de fuerza bruta
- Solo sirve para contraseñas calculadas sin salto
  - El salto es un parámetro adicional a la contraseña para calcular el hash.
  - No hay forma de diferenciar entre una contraseña con salto y una contraseña sin salto, ya que las tablas fueran calculadas sin tomar en cuenta los saltos.

Lámina 16
Dr. Roberto Gómez Cárdenas

Chosen plaintext attack

Dado:

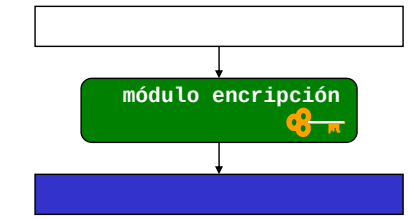
Capacidad de encriptar un fragmento arbitrario, elegido, del texto en claro

Se busca por

llave

Ejemplo

Criptoanálisis diferencial



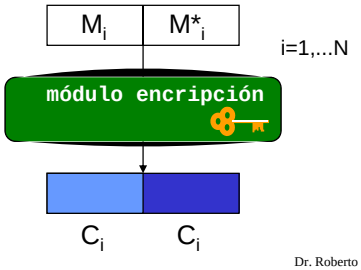


Lámina 17 Dr. Roberto Gómez Cárdenas

Diagrama ataque diferencial

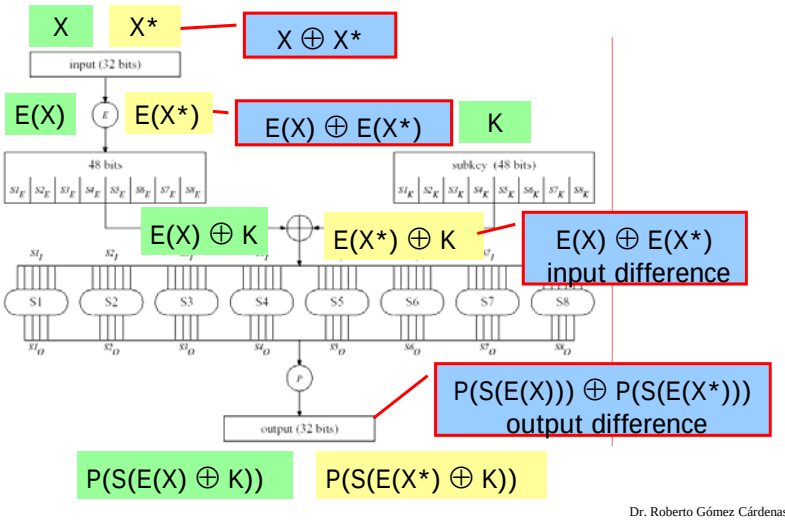


Lámina 18 Dr. Roberto Gómez Cárdenas

Timing attack

- Descubierto por Paul Kocher
- Atacante cuenta con medidas de tiempo (obtenidas a través de fisgoneo) y elige varios criptogramas c1, c2, ... cn
- Objetivo
  - deducir llave secreta K
  - por ejemplo: RSA, atacante hace que objetivo calcule c^e mod n, para descubrir e
- Toma ventaja de la lentitud de calculo de la llave pública
- Velocidad de los cálculos depende de la llave de encripción y de los datos de entrada
- Para algunos algoritmos, algunos bits de la llave pueden ser adivinados observando el tiempo del procesador cuando esta haciendo los cálculos de la llave secreta

Lámina 19 Dr. Roberto Gómez Cárdenas

Otros ataques

- Chosen-key attack
  - similar a criptoanálisis diferencial, pero examina diferencias entre las llaves
  - atacante elige una relación entre llaves, pero no conoce la llave
  - independiente del número de iteraciones de encripción
  - no es un ataque práctico, pero es interesante
- Rubber-hose cryptoanalysis
  - threats, blackmails, torturas hasta que se obtiene la llave
  - soborno, también conocido como *purchase-key attack*
  - mejor forma de romper un algoritmo y la más efectiva

Lámina 20 Dr. Roberto Gómez Cárdenas

Ataques sobre funciones de un solo sentido

- Existen dos ataques de fuerza bruta sobre una función de sólo un sentido:
  - dado el hash de un mensaje, H(M), el adversario quiere ser capaz de crear otro documento M' tal que H(M) = H(M´).
  - el adversario quisiera encontrar dos mensajes al azar, M y M' tal que H(M) = H(M'), a esto se le concoce como colisión.

Lámina 21 Dr. Roberto Gómez Cárdenas

Birthday attack

- Es un problema de tipo estadístico.
- ¿Cuál es el valor mínimo de k, para que la probabilidad de que al menos una persona, en un grupo de k gentes, cumpla años el mismo día que usted, sea mayor a 0.5?
  - Respuesta: 253
- ¿Cuál es el valor mínimo de k, para que la probabilidad de que al menos dos personas, en un grupo de k gentes, cumplan años el mismo día, sea mayor a 0.5?
  - Respuesta: 23

Lámina 22 Dr. Roberto Gómez Cárdenas

Analogía con funciones un solo sentido

- Encontrar a alguien con un día de nacimiento es análogo a encontrar un mensaje que coincida con un valor de hash conocido.
- Encontrar a dos gentes con el mismo día de cumpleaños al azar es análogo a encontrar una colisión de mensajes. Este es conocido como el *birthday attack*.

Lámina 23 Dr. Roberto Gómez Cárdenas

¿Es complejo el ataque?

- Asumir función hash produce una salida de m bits y se almacena en x.
- Encontrar un mensaje cuyo valor hash sea igual a x, requiere aplicar la función a 2^m mensajes.
- Encontrar dos mensajes cuyo valor hash sea igual a x, requiere aplicar la función a $2^{m/2}$ mensajes.
- Una máquina que procese un millón de mensajes por segundo tomaría 600,000 años para encontrar un mensaje que colisione con un valor hash de 64 bits.

Lámina 24 Dr. Roberto Gómez Cárdenas

Sin embargo

- Algunas vulnerabilidades se han encontrado
  - Conferencia Crypto 2004
- Multicollisions in iterated hash functions. Application to cascaded constructions , Antoine Joux
- Collisions for Hash Functions MD4, MD5, HAVAL-128 and RIPEMD, Xiaoyun Wang 1 , Dengguo Feng 2 , Xuejia Lai 3 , Hongbo Yu 1
- Near-Collisions of SHA-0 Eli Biham and Rafi Chen

Lámina 25 Dr. Roberto Gómez Cárdenas

Ejemplo de colisión en MD5

- Archivos datos 1 (en hexadecimal)

```
d1 31 dd 02 c5 e6 ee c4 69 3d 9a 06 98 af f9 5c
2f ca b5 87 12 46 7e ab 40 04 58 3e b8 fb 7f 89
55 ad 34 06 09 f4 b3 02 83 e4 88 83 25 71 41 5a
08 51 25 e8 f7 cd c9 9f d9 1d bd f2 80 37 3c 5b
d8 82 3e 31 56 34 8f 5b ae 6d ac d4 36 c9 19 c6
dd 53 e2 b4 87 da 03 fd 02 39 63 06 d2 48 cd a0
e9 9f 33 42 0f 57 7e e8 ce 54 b6 70 80 a8 0d 1e
c6 98 21 bc b6 a8 83 93 96 f9 65 2b 6f f7 2a 70
```

Hay hasta 24 bits diferentes

Los primeros 4 de 8 bits son distintos.

- Archivo datos 2 (en hexadecimal)

```
d1 31 dd 02 c5 e6 ee c4 69 3d 9a 06 98 af f9 5c
2f ca b5 07 12 46 7e ab 40 04 58 3e b8 fb 7f 89
55 ad 34 06 09 f4 b3 02 83 e4 88 83 25 f1 41 5a
08 51 25 e8 f7 cd c9 9f d9 1d bd 72 80 37 3c 5b
d8 82 3e 31 56 34 8f 5b ae 6d ac d4 36 c9 19 c6
dd 53 e2 34 87 da 03 fd 02 39 63 06 d2 48 cd a0
e9 9f 33 42 0f 57 7e e8 ce 54 b6 70 80 28 0d 1e
c6 98 21 bc b6 a8 83 93 96 f9 65 ab 6f f7 2a 70
```

Y la función hash MD5 es:

MD5 = 79 05 40 25 25 5f b1 a2 6e 4b c4 22 ae f5 4e b4

Lámina 26 Dr. Roberto Gómez Cárdenas

Caso anecdótico de la multa de tráfico

- Australia, agosto de 2005: fotografía de un radar tomada por la policía a un automóvil por exceso de velocidad.
- El abogado defensor del infractor recurre la denuncia argumentando que la imagen puede ser de otro automóvil (no de su cliente) o se puede haber alterado el original.
- La policía responde que usa MD5 para comprobar la integridad de las fotografías de dicho aparato electrónico, … pero luego ¡no encuentra ningún perito que se atreva a demostrar ante el juez la infalibilidad de ese algoritmo!
- Conclusión: se retira la denuncia y se anula la multa.
- ¡Una jurisprudencia muy peligrosa si se lleva esta situación por ejemplo al escenario de una certificación X.509!
- Referencia:
  - http://www.hispasec.com/unaaldia/2489

Lámina 27 Dr. Roberto Gómez Cárdenas

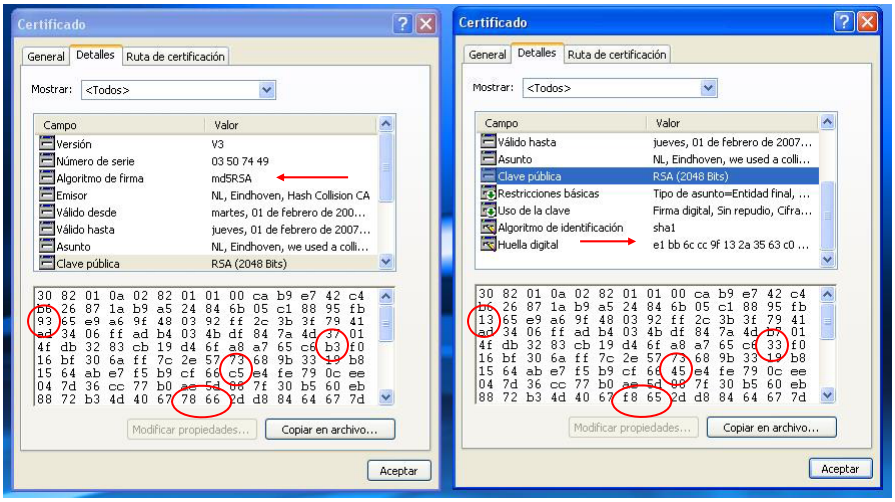
Colisión MD5 en certificados X.509

- Seguimiento con OpenSSL de dos certificados X.509 distintos y con igual firma, mostrado por Lenstra, Wang y Weger en el artículo Colliding X.509 Certificates.
- Dos certificados X.509 válidos, firmados por la misma autoridad, con igual número de serie, pero distintos en sus datos del módulo RSA y por consiguiente con diferente huella digital o fingerprint, entregan la misma firma digital RSA sobre MD5.
- En este ejemplo, el módulo n de RSA con 2.048 bits tiene 6 bytes distintos.
- Referencia:
  - http://www.win.tue.nl/~bdeweger/CollidingCertificates

Lámina 28 Dr. Roberto Gómez Cárdenas

TECNOLÓGICO DE MONTERREY

Certificados C1 y C2 X.509



TECNOLÓGICO DE MONTERREY

"Integridad" del código ejecutable

```
C:\TEMP> md5sum hello.exe
cdc47d670159eef60916ca03a9d4a007
C:\TEMP> .\hello.exe
Hello, world!

(press enter to quit)
C:\TEMP>
```

```
C:\TEMP> md5sum erase.exe
cdc47d670159eef60916ca03a9d4a007
C:\TEMP> .\erase.exe
This program is evil!!!
Erasing hard drive...1Gb...2Gb... just kidding!
Nothing was erased.

(press enter to quit)
C:\TEMP>
```

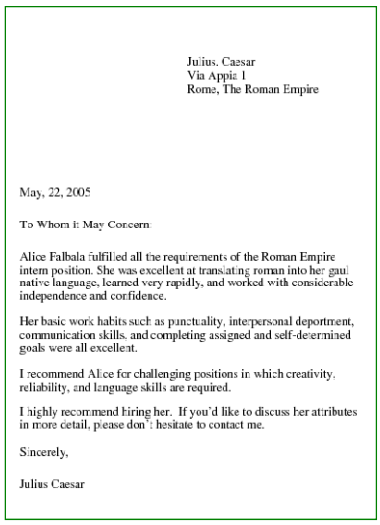
```
C:\OpenSSL\bin>openssl md5 hello.exe
MD5(hello.exe)= cdc47d670159eef60916ca03a9d4a007

C:\OpenSSL\bin>openssl md5 erase.exe
MD5(erase.exe)= cdc47d670159eef60916ca03a9d4a007
```

Cálculo de hash con OpenSSL

Artículo de Peter Selinger
http://www.mathstat.dal.ca/~selinger/md5collision/

Alicia y su jefe César



Julius. Caesar
Via Appia 1
Rome, The Roman Empire

May, 22, 2005

To Whom It May Concern

Alice Falbala fulfilled all the requirements of the Roman Empire intern position. She was excellent at translating roman into her gaul native language, learned very rapidly, and worked with considerable independence and confidence.

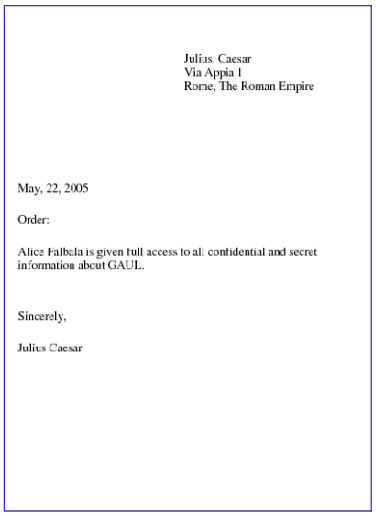
Her basic work habits such as punctuality, interpersonal deportment, communication skills, and completing assigned and self-determined goals were all excellent.

I recommend Alice for challenging positions in which creativity, reliability, and language skills are required.

I highly recommend hiring her. If you'd like to discuss her attributes in more detail, please don't hesitate to contact me.

Sincerely,

Julius Caesar



Julius. Caesar
Via Appia 1
Rome, The Roman Empire

May, 22, 2005

Order:

Alice Falbala is given full access to all confidential and secret information about GAUL.

Sincerely,

Julius Caesar

MD5 = a25f7f0b29ee0b3968c860738533a4b9

Lámina 31 Dr. Roberto Gómez Cárdenas

NIST nuevo estándar hash

- Debido a los ataques a las funciones hash, varias voces reconocidas venían pidiendo un concurso similar al que se realizó con el algoritmo AES para encontrar un nuevo estándar para estas funciones.
  - además de seguro, sea "duradero".
- Finalmente el NIST publica una nota de prensa el 23 de enero 2007 con un borrador para nuevo estándar en hash.
- El 30 de abril se cerró el plazo para envío de comentarios.
- Se espera que esté operativo para 2009, aunque otros lo ponen más lejos... y los más escépticos dicen que habrá que cambiar de estándar cada 5 años...

Lámina 32 Dr. Roberto Gómez Cárdenas

Páginas relacionadas

- Artículo colisiones de Arturo Quirantes en Kriptópolis
  - http://www.kriptopolis.org/sha-1-y-las-colisiones-de-cumpleanos
- Implementación del trabajo de Xiaoyun Wang, et al.
  - http://www.stachliu.com/collisions.html
- Base de datos de huellas digitales
  - http://www.ahazu.com/
- NIST's Plan for New Cryptographic Hash Functions
  - http://csrc.nist.gov/pki/HashWorkshop/

Lámina 33 Dr. Roberto Gómez Cárdenas