La semana pasada...

El operador módulo en lenguajes de programación...

```
n * (a / n) + a % n = a
```

División entera

Si `7/-15 = -1` entonces `7%-15 = -8`

Si `7/-15 = 0` entonces `7%-15 = 7`

Si nos mantenemos **siempre** en módulo -15 entonces todo funciona bien.

¿Si salimos de ese mundo?

```
56 / (7 % -15) + 8 = 1
                     16
```

Al menos $1 \equiv 16 \mod -15$

¿Si aplicamos otros módulos?

```
                    -6
10 % (7 % -15) =
                     3
```

Aquí tenemos:

$3 \not\equiv -6 \mod -15$ $\quad$ $3 \not\equiv -6 \mod 7$ $\quad$ $3 \not\equiv -6 \mod -8$

ONE-TIME PAD (OTP)

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
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	26

Para enviar un mensaje de largo ℓ
necesitaremos una llave de largo ℓ

HOLAMUNDO		7	15	11	0	12	21	13	3	15
SECRETKEY	$+$	19	4	2	18	4	20	10	4	25
	$\bmod 27$	26	19	13	18	16	41	23	7	40
ZSNRPÑWHN		26	19	13	18	16	14	23	7	13

texto cifrado

¿Cómo decriptar?

ZSNRPÑWHN		26	19	13	18	16	14	23	7	13
SECRETKEY	$-$	19	4	2	18	4	20	10	4	25
	$\bmod 27$	7	15	11	0	12	−6	13	3	−12
HOLAMUNDO		7	15	11	0	12	21	13	3	15

¿Cómo decriptar?

ZSNRPÑWHN		26	19	13	18	16	14	23	7	13
SECRETKEY	−	19	4	2	18	4	20	10	4	25
	mod 27	7	15	11	0	12	-6	13	3	-12
HOLAMUNDO		7	15	11	0	12	21	13	3	15

Para formalizarlo necesitamos convertir mensajes, llaves y textos cifrados en arreglos de enteros

$m = \texttt{HOLAMUNDO}$ $\qquad$ $\bar{m} = (7, 15, 11, 0, 12, 21, 13, 3, 15)$

$k = \texttt{SECRETKEY}$ $\qquad$ $\bar{k} = (19, 4, 2, 18, 4, 20, 10, 4, 25)$

$c = \texttt{ZSNRPÑWHN}$ $\qquad$ $\bar{c} = (26, 19, 13, 18, 16, 14, 23, 7, 13)$

De la misma forma necesitamos hacer
la conversión en la otra dirección

$a = (4, 9, 4, 12, 16, 11, 15)$ $\qquad \bar{a} = \texttt{EJEMPLO}$

Naturalmente, siempre se cumple que $\overline{\overline{s}} = s$

Con esto definimos OTP en base a

$$Enc_k(m) = \overline{(\bar{m} + \bar{k}) \mod 27}$$

$$Dec_k(c) = \overline{(\bar{c} - \bar{k}) \mod 27}$$

Desde ahora supondremos que nuestros mensajes y llaves **son** arreglos de números

Definiremos OTP simplemente usando

$$Enc_k(m) = (m + k) \mod 27$$

$$Dec_k(c) = (c - k) \mod 27$$

$$Dec_k(Enc_k(m)) = ((m + k) \mod 27 - k) \mod 27$$

$$= (m + k - k) \mod 27 = m \mod 27 = m$$

Esperamos que para un esquema criptográfico (Gen, Enc, Dec) se cumpla

$$\forall k \in \mathcal{K}\, \forall m \in \mathcal{M} : Dec_k(Enc_k(m)) = m$$

En este caso diremos que el esquema es *perfectamente correcto*

¿Por qué *perfectamente*?

OTP: sobre $\{0,1,\ldots,N-1\}$ y mensajes de largo ℓ ($\mathrm{OTP}^{N,\ell}$)

$$\mathcal{K}=\mathcal{M}=\mathcal{C}=\{0,\ldots,N-1\}^{\ell}$$

Gen es la distribución uniforme sobre $\{0,\ldots,N-1\}^{\ell}$

$$Enc_k(m)=(m+k) \mod N$$

$$Dec_k(c)=(c-k) \mod N$$

¿Qué tan bueno es OTP?

¿Qué pasa si veo un mensaje cifrado c pasar?

Aquí un ejemplo:

```
c = YFTGXEIWIWEHAGQGESLPKRVLMYGXSJIQZVIYHVBRJGNTR
m = ESTEMENSAJEESLITERALMENTEIMPOSIBLEDEDECRIPTAR
k = UNACLAVEINADIVINABLEYNISIQUIERAPORFUERZABRUTA
```

Perfect Secrecy

La clase pasada...

Esquema criptográfico

$$\mathcal{K}, \mathcal{M}, \mathcal{C}$$

Espacio de llaves, mensajes y textos cifrados

Un esquema es un triple (Gen, Enc, Dec)

Gen es una distribución de probabilidades sobre $\mathcal{K}$

Es decir, $Gen : \mathcal{K} \to [0, 1]$ tal que $\sum_{k \in \mathcal{K}} Gen(k) = 1$

Esquema criptográfico

$$\mathcal{K}, \mathcal{M}, \mathcal{C}$$

Espacio de llaves, mensajes y textos cifrados

Un esquema es un triple (Gen, Enc, Dec)

$Enc = \{Enc_k \mid k \in \mathcal{K}\}$ es una familia de algoritmos para encriptar

Para cada $k \in \mathcal{K}$, se tiene que $Enc_k : \mathcal{M} \to \mathcal{C}$

Esquema criptográfico

$$\mathcal{K}, \mathcal{M}, \mathcal{C}$$

Espacio de llaves, mensajes y textos cifrados

Un esquema es un triple (Gen, Enc, Dec)

$Dec = \{Dec_k \mid k \in \mathcal{K}\}$ es una familia de algoritmos para decriptar

Para cada $k \in \mathcal{K}$, se tiene que $Dec_k : \mathcal{C} \to \mathcal{M}$

¿Cuándo decimos que un esquema criptográfico es *perfectamente secreto?*

Pensemos en la idea de que si un atacante ve un texto cifrado *no gana información.*

Podríamos decir algo como lo siguiente:

"Al ver un texto cifrado c_0 pasar, para el atacante el mensaje original m podría haber sido cualquiera"

¿Cómo formalizamos esto?

Dado un texto cifrado c_0 se cumple que

$$\forall m_0 \in \mathcal{M} : \quad \Pr_{k \sim Gen} [Enc_k(m_0) = c_0] = \frac{1}{|\mathcal{M}|}$$

$$\forall m_0 \in \mathcal{M} : \Pr_{k \sim Gen} [Enc_k(m_0) = c_0] = \frac{1}{|\mathcal{M}|}$$

¿Cómo se calcula esta probabilidad?

Es simplemente la probabilidad de haber elegido una llave que encripte m_0 como c_0

$$\sum_{k \in \mathcal{K} \,:\, Enc_k(m_0) = c_0} Gen(k)$$

¿Qué pasa si el atacante tenía información previa sobre el mensaje?

Por ejemplo sabe que el mensaje puede ser *"atacar ahora"* o *"emprender retirada"*

Podría incluso estimar que atacarán con probabilidad $1/3$

¿Cómo modelamos esto matemáticamente?

Supondremos que el atacante tiene una distribución de probabilidad $\mathbb{D}$ sobre $\mathcal{M}$

Para cada distribución de probabilidad $\mathbb{D}$ sobre $\mathcal{M}$ y cada texto cifrado $c_0 \in \mathcal{C}$ se cumple que

$$\forall m_0 \in \mathcal{M} : \quad \Pr_{\substack{m \sim \mathbb{D} \\ k \sim Gen}} [m = m_0 \mid Enc_k(m) = c_0] \quad = \quad \Pr_{m \sim \mathbb{D}}[m = m_0]$$

Recordemos que $\Pr(A|B) = \frac{\Pr(A \cap B)}{\Pr(B)}$

$$\frac{\mathbb{D}(m_0) \sum_{k \in \mathcal{K} \,:\, Enc_k(m_0) = c_0} Gen(k)}{\sum_{m \in \mathcal{M}} \mathbb{D}(m) \sum_{k \in \mathcal{K} \,:\, Enc_k(m) = c_0} Gen(k)} \quad \stackrel{?}{=} \quad \mathbb{D}(m_0)$$

¿Es $\mathrm{OTP}^{N,\ell}$ perfectamente secreto?

1. Gen es la distribución uniforme $1/N^{\ell}$
2. Para cada c_0 y cada m_0 existe una única llave k tal que $Enc_k(m_0) = c_0$

Sea $c_0 \in \mathcal{C}$ un texto cifrado y m_0 un mensaje.

$$\frac{\mathbb{D}(m_0) \sum_{k \in \mathcal{K} \,:\, Enc_k(m_0) = c_0} Gen(k)}{\sum_{m \in \mathcal{M}} \mathbb{D}(m) \sum_{k \in \mathcal{K} \,:\, Enc_k(m) = c_0} Gen(k)}$$

$$\frac{\mathbb{D}(m_0) \cdot \cancel{1/N^{\ell}}}{\cancelto{1}{\sum_{m \in \mathcal{M}} \mathbb{D}(m)} \cdot \cancel{1/N^{\ell}}} = \mathbb{D}(m_0)$$

¿Definiciones alternativas?

1. La probabilidad de ver cualquier texto cifrado sin conocimiento previo es la misma que la probabilidad de ver dicho texto cifrado conociendo el mensaje de antemano.

2. La distribución de probabilidad sobre los mensajes es independiente de la distribución de probabilidad sobre los textos cifrados.

Ejercicio: formalizar estas nociones

Un poco de historia

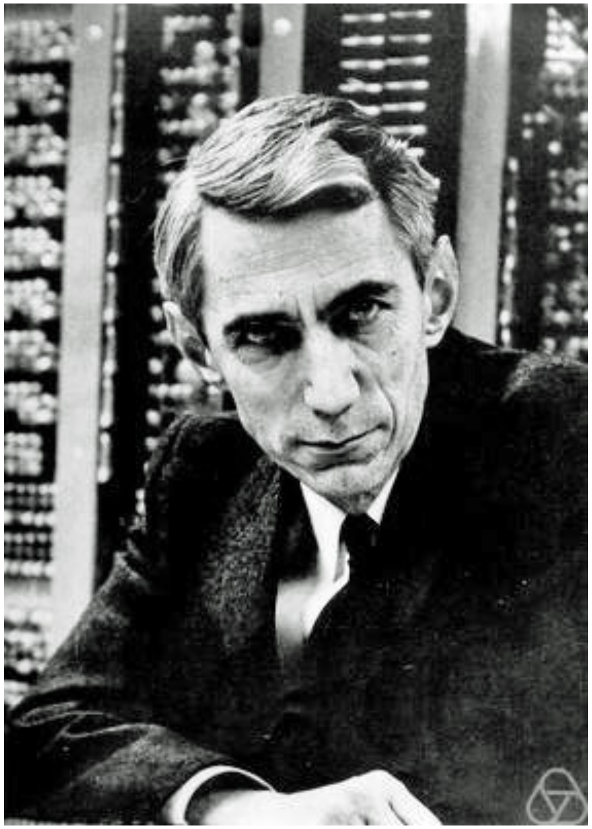
Julio César

100 - 44 a.C.

Shannon (1945)

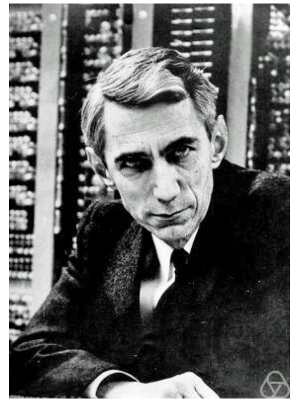
A Mathematical Theory of Cryptography

Primer desarrollo de fundamentos matemáticos de la criptografía



Julio César

Shannon



Perfect
secrecy

100 - 44 a.C.

1945

Pero perfect secrecy es una condición muy fuerte

Lamentablemente...

Hemos discutido en clases que pareciera *molesto* y/o *poco razonable* que la llave tenga que ser tan larga como el mensaje.

¿Cómo modificamos OTP para tener $|\mathcal{K}| < |\mathcal{M}|$ y seguir teniendo un esquema criptográfico *perfectamente secreto?*

No podemos 🤦

Teorema

Sean $\mathcal{M}, \mathcal{K}, \mathcal{C}$ espacios de mensajes, llaves y textos cifrados, respectivamente.

Si $|\mathcal{K}| < |\mathcal{M}|$, entonces no existe un esquema (Gen, Enc, Dec) que sea perfectamente secreto.

Demostración

Supongamos que $|\mathcal{K}| < |\mathcal{M}| \leq |\mathcal{C}|$ y sea (Gen, Enc, Dec) un esquema criptográfico

Sea $\mathbb{D}$ una distribución sobre $\mathcal{M}$ y $m_0 \in \mathcal{M}$ un mensaje tal que $\mathbb{D}(m_0) > 0$

Como $|\mathcal{K}| < |\mathcal{M}| \leq |\mathcal{C}|$, debe existir $c_0 \in \mathcal{C}$ para el cual **ninguna** llave $k \in \mathcal{K}$ satisface $Enc_k(m_0) = c_0$

$$\cancelto{0}{\Pr_{\substack{m \sim \mathbb{D} \\ k \sim Gen}} [m = m_0 \mid Enc_k(m) = c_0]} \quad < \quad \mathbb{D}(m_0) = \Pr_{m \sim \mathbb{D}}[m = m_0]$$

Adiós perfect secrecy...


Life is a series of closing doors, isn't it?

Back to reality

OTP y la noción de Perfect Secrecy son fundamentales para entender lo que viene

Pero en la práctica vamos a buscar otras propiedades...

¿Bajo qué modelo de ataque es OTP seguro?

1. Texto cifrado (sólo veo $c_0, c_1, \ldots$)
2. Texto plano (veo $(m_0, c_0), (m_1, c_1), \ldots$)
3. Texto plano elegido (mando a encriptar m's)
4. Texto cifrado elegido (mando a encriptar m's y a decriptar c's)

¿Qué pasa si usamos llaves más cortas que el mensaje?

Pensemos por ejemplo en repetir la llave varias veces
para encriptar un mensaje más largo que la llave

HOLAMUNDOCOMOTEHAIDO →	7	15	11	0	12	21	13	3	15	3	15	12	15	20	4	7	0	8	3	15
SECRETKEYSECRETKEYSE $+$	19	4	2	18	4	20	10	4	25	19	4	2	18	4	20	10	4	25	19	4
	26	19	13	18	16	41	23	7	40	22	19	14	33	24	24	17	4	33	22	19
$\bmod 27$																				
ZSNRPÑWHNVSÑGXXQEGVS ←	26	19	13	18	16	14	23	7	13	22	19	14	6	24	24	17	4	6	22	19

texto cifrado

¿Podemos quebrarlo?