Criptografía y Seguridad Computacional **2016-01**

Clase 1: 02/03/2016

Profesor: Fernando Krell *Notas: Pedro Pablo Aste Kompen*

1. Esquemas perfectamente secretos

Empezaremos el curso definiendo que entendemos por un esquema de cifrado. Por el momento nos enfocaremos en esquemas simétricos, en donde la llave para cifrar el mensaje es la misma que lo descifra. Un esquema de cifrado está compuesto por los siguientes 3 algoritmos:

- Gen: Es un algoritmo aleatorizado que genera una llave de cifrado k.
- $\mathsf{Enc}(k, m)$: Es el algoritmo de encriptación. Toma como entrada una llave k, generada por Gen y un mensaje m, y genera en su salida un mensaje c. m pertenece a un conjunto $\mathcal{M}$ que denominaremos espacio de textos planos. c pertenece a un conjunto $\mathcal{C}$ que denominaremos espacio de textos cifrados. Enc puede ser determinístico o aleatorizado.
- $\mathsf{Dec}(k, c)$: Es el algoritmo de desencriptación. Dada una llave k y el texto cifrado c retorna un mensaje de texto plano m si, y sólo si, $c \in \mathsf{Enc}(k, m)$.[1]

Gen induce un conjunto $\mathcal{K}$ de posibles llaves definido por $\mathcal{K} = \{k | \Pr[\mathsf{Gen} \to k] > 0\}$.

Hemos definido los algoritmos de un esquema de cifrado, pero no hemos descrito las propiedades deben cumplir estos para que el esquema sea seguro.

A continuación describiremos 4 definiciones de seguridad equivalentes. Intuitivamente, queremos que si el adversario logra obtener un texto cifrado c, este no sirva para obtener información alguna sobre el respectivo texto plano (a menos que tenga la llave k), más allá de la información a-priori que el adversario tenga. Para modelar situaciones en las cuales el adversario puede tener información a-priori sobre el texto plano, asociaremos el espacio de textos planos $\mathcal{M}$ a una distribución de probabilidad. Es decir, asumiremos que cada mensaje $m \in \mathcal{M}$ tiene asociada una probabilidad de ser

[1] Si Enc es aleatorizado, $\mathsf{Enc}(k, m)$ define un conjunto. En cambio, si Enc es determinístico, existe un único c para k y m, y sería correcto escribir $c = \mathsf{Enc}(k, m)$.

escogido, y esta probabilidad es públicamente conocida. Por ejemplo, $\Pr[\texttt{atacar}] = 0{,}2$ y $\Pr[\texttt{no atacar}] = 0{,}8$.

Abusando la notación, denominaremos la distribución de probabilidad sobre textos planos como $\mathcal{M}$. Por lo tanto, denotaremos la probabilidad de que una variable aleatoria M distribuida acorde a $\mathcal{M}$ tenga valor m como: $\Pr_{M\sim\mathcal{M}}[M = m]$, o sencillamente $\Pr[M = m]$. Denotaremos C a la variable aleatoria que define el procedimiento de obtener la llave K y después computar $\mathsf{Enc}(K, M)$

Definición 1. *Sea* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *un esquema de cifrado. Diremos que* Π *es* perfectamente secreto *si* $\forall m \in \mathcal{M}\ \forall c \in \mathcal{C}$

$$\Pr[M = m|C = c] = \Pr[M = m]$$

La definición anterior captura precisamente nuestra intuición anterior. La probabilidad de que M tome un valor específico m. dado que se sabe que su texto cifrado respectivo es c, es exactamente igual a la probabilidad de que M sea m sin saber el valor que toma $C = \mathsf{Enc}_{K,m}$. En otras palabras, c no revela información alguna sobre m.

Lema 2. Π *es perfectamente secreto si, y sólo si,* $\forall c \in \mathcal{C}, \forall m \in \mathcal{M}$

$$\Pr[C = c|M = m] = \Pr[C = c]$$

El lema anterior nos da una definición alternativa para esquemas perfectamente secretos. Intuitivamente, la probabilidad de que el texto cifrado tenga un valor específico c no debería cambiar por el hecho que se sepa cuál es el texto plano. Demostramos formalmente el lema a continuación.

Demostración. Usaremos Bayes. ($\Rightarrow$)

$$\begin{aligned}\Pr[C = c|M = m] &\overset{\texttt{bayes}}{=} \frac{\Pr[M = m|C = c] \cdot \Pr[C = c]}{\Pr[M = m]} \\ &\overset{\texttt{def 1}}{=} \Pr[C = c]\end{aligned}$$

($\Leftarrow$)Para el otro sentido, la demostración es análoga. □

Lema 3. Π *es perfectamente secreto si, y sólo si,* $\forall c \in \mathcal{C}, \forall m_0, m_1 \in \mathcal{M}$

$$\Pr[\mathsf{Enc}(k, m_0) = c] = \Pr[\mathsf{Enc}(k, m_1) = c]$$

Demostración. Si Π es perfectamente secreto, entonces por lema 2 tenemos que

$$\Pr[C = c|M = m] = \Pr[C = c]$$

para todo m y para todo c. Por lo tanto

$$\Pr[C = c|M = m_0] = \Pr[C = c] = \Pr[C = c|M = m_1]$$

.

En el otro sentido, tenemos que $\Pr[C = c|M = m_0] = \Pr[C = c|M = m_1]$ para todo c, m_0 y m_1 Por lo tanto podemos definir la constante $\delta \stackrel{\texttt{def}}{=} \Pr[C = c|M = m]$. Ahora demostraremos que para todo c, m, $\Pr[C = c] = \Pr[C = c|M = m](= \delta)$ y por lo tanto el esquema es perfectamente secreto.

$$\begin{aligned}
\Pr[C = c] &= \sum_{m \in \mathcal{M}} \Pr[C = c \wedge M = m] \\
&= \sum_{m \in \mathcal{M}} \Pr[C = c|M = m] \Pr[M = m] \\
&= \delta \sum \Pr[M = m] \\
&= \delta
\end{aligned}$$

□

1.1. Definición basada en experimentos

La última definición equivalente que daremos para esquemas de cifrado perfectamente secretos introduce el concepto de adversario y experimento.

El experimento consiste en un juego contra el adversario. Generamos una llave k y dejaremos al adversario elegir 2 mensajes, m_0 y m_1. Luego elegimos un bit b uniformemente aleatorio ($\Pr[b = 1] = \Pr[b = 2] = 1/2$) y daremos al adversario el texto cifrado $\mathsf{Enc}(k, m_b)$. El adversario gana en este experimento si puede adivinar b. Claramente, cualquier adversario puede ganar con probabilidad 1/2 (por ejemplo, puede elegir b' = 0 sea cuál sea el texto cifrado). Lo que queremos del esquema, en términos de seguridad, es que la probabilidad de que el adversario gane *no* sea más que 1/2. En otras palabras, el texto cifrado no le ayuda al adversario a saber cuál fue el texto plano elegido, a pesar de que este haya elegido los mensajes arbitrariamente.

Definición 4. Π *es perfectamente indistinguible si para todo adversario* $\mathcal{A}$ $\Pr[\mathsf{Exp}^{ind}_{\Pi,A} =$ *"*$\mathcal{A}$ *gana"*$] = 1/2$

Lema 5. Π *es perfectamente secreto si y sólo si* Π *es perfectamente indistinguible.*

$\underline{\mathsf{Exp}^{ind}_{\Pi,\mathcal{A}}}$:

1. $k \leftarrow \mathsf{Gen}$
2. $m_0, m_1 \leftarrow \mathcal{A}$
3. $b \xleftarrow{\$} \{0,1\}$
4. $b' \leftarrow A(\mathsf{Enc}(k, m_b))$

output "$\mathcal{A}$ gana" sii $b' = b$

Figura 1: Experimento $\mathsf{Exp}^{ind}_{\Pi,\mathcal{A}}$

2. El One-Time Pad

En esta sección construiremos un esquema de cifrados perfectamente secreto, llamado one-time pad (OTP). Usaremos la función "o exclusivo" la cual usualmente se denota con el símbolo $\oplus$.

$\oplus$	0	1
0	0	1
1	1	0

La propiedad que usaremos es que si b es uniformemente aleatorio en $\{0,1\}$, entonces $\Pr[b \oplus b' = 0] = \Pr[b \oplus b' = 1] = 1/2$, para $b' \in \{0,1\}$. Describimos el esquema a continuación para mensajes en $\mathcal{M} = \{0,1\}^n$

- Gen: elegir una llave k uniformemente aleatoria $\{0,1\}^n$
- $\mathsf{Enc}(k, m)$: output $c = m \oplus k$ ($\oplus$ bit a bit)
- $\mathsf{Dec}(k, c)$: output $m = c \oplus k$

Primero vemos que el esquema es correcto: si $c = m \oplus k$, entonces $c \oplus k = m \oplus k \oplus k = m$.

Ahora analizamos su seguridad.

Teorema 6. *El One-Time Pad es un cifrador simétrico perfectamente secreto.*

Demostración. Utilizaremos lema 2 para demostrar que OTP es perfectamente secreto. Esto es, mostraremos que $\Pr[C = c | M = m] = \Pr[C = c]$. □

$$\begin{aligned}\Pr[C = c|M = m] &= \Pr[M \oplus K = c|M = m] \\ &= \Pr[m \oplus K = c] \\ &= \Pr[K = c \oplus m] = 2^{-n}\end{aligned}$$

En donde la última igualdad utilizamos el hecho de que K es uniformemente distribuido en $\{0,1\}^n$. Ahora demostraremos que $\Pr[C = c]$ también es 2^{-n}

$$\begin{aligned}\Pr[C = c] &= \Pr[M \oplus K = c] \\ &= \Pr[K = c \oplus M] \\ &= \sum_{m \in \mathcal{M}} \Pr[K = c \oplus m] \cdot \Pr[M = m] \\ &= 2^{-n} \sum_{m \in \mathcal{M}} \cdot \Pr[M = m] \\ &= 2^{-n}\end{aligned}$$

Este cifrador es llamado one-time pad debido a que el pad (la llave k) puede ser utilizado una sola vez. Si la misma llave es utilizada para más mensajes, entonces el esquema deja de ser seguro en la práctica. En particular, si un adversario obtiene $c_0 = m_0 \oplus k$ y $c_1 = m_1 \oplus k$, este puede computar el valor $m_0 \oplus m_1$ lo cual es inseguro si el adversario tiene información a-priori sobre el espacio de textos planos y su distribución.

Una segunda desventaja, que esta relacionada con la anterior, es que la llave k tiene que ser tan larga como el mensaje m. Esto implica que en la práctica el OTP no puede ser usado regularmente.

Esta segunda desventaja es intrínseca para cualquier cifrador perfectamente seguro:

Teorema 7. *Sea* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *un cifrador,* $\mathcal{K}$ *el espacio de llaves generado por* Gen*, y* $\mathcal{M}$ *el espacio de mensajes. Si* Π *es perfectamente secreto, entonces* $|\mathcal{K}| \geq |\mathcal{M}|$*.* [2]

Demostración. Demostraremos que si $|\mathcal{K}| < |\mathcal{M}|$ entonces existe c y m tales que $\Pr[M = m|C = c] \neq \Pr[M = m]$ y por lo tanto Π no es perfectamente secreto.

Sea c arbitrario. Definimos el siguiente conjunto que contiene todos los mensajes para los cuales existe una llave k que cifra m hacia c:

$$\mathcal{M}(c) = \{m | \exists k \in \mathcal{K} \text{ tal que } \mathsf{Enc}(k, m) = c\}$$

[2] en donde $|S|$ denota la cantidad de elementos de un conjunto S.

Claramente $|\mathcal{M}(c)| \leq |\mathcal{K}|$, como asumimos por contradicción que $|\mathcal{K}| < |\mathcal{M}|$ sabemos que $|\mathcal{M}(c)| < |\mathcal{M}|$. Es decir, existe m en $\mathcal{M}$ tal que $m \notin \mathcal{M}(c)$. Por lo tanto,

$$\Pr[M = m|C = c] = 0 \neq \Pr[M = m]$$

□