Criptografía y Seguridad Computacional - IIC3253
Tarea 1
Solución pregunta 3

Considere el juego que define una PRP con una ronda sobre OTP. Demuestre que todo adversario tiene una probabilidad de ganar de exactamente 1/2.

Bonus: ¿Es cierto lo anterior si consideramos cualquier protocolo que satisface la propiedad de perfect secrecy? Demuestre o de un contraejemplo.

Solución. Trabajaremos sobre strings binarios de largo n, es decir $\mathcal{C} = \mathcal{M} = \mathcal{K} = \{0,1\}^n$. Comenzamos describiendo el juego para definir la notación. Al iniciar el juego, el Verificador elige $b = 0$ o $b = 1$. Si $b = 0$, el Verificador también elige una llave k y define $f : \mathcal{M} \to \mathcal{C}$ como $f(m) = Enc_k(m) = m \oplus k$. De lo contrario, selecciona una permutación π al azar y define $f(m) = \pi(m)$. Notar que tanto la elección de b como la elección de k siguen la distribución uniforme.

A continuación, el Adversario envía un mensaje m al Verificador. El verificador responde con $r = f(m)$. Finalmente, el Adversario debe retornar 0 o 1, y gana si el valor retornado es igual a b.

Teniendo el setup anterior, para demostrar que la probabilidad de ganar de cualquier adversario es exactamente 1/2, pensaremos el problema de la siguiente forma: el adversario responde $b = 0$ ó $b = 1$ basándose la respuesta $r = f(m)$ que envía el adversario. La pregunta entonces es: ¿Cuál es la probabilidad de que b haya sido cero o uno, dada la respuesta del verificador? Esto lo podemos calcular de la siguiente forma

$$\Pr_{b\sim\mathbb{U},\, k\sim\mathbb{U}}(b = 0 \mid r = f(m))$$

Donde $\mathbb{U}$ representa la distribución uniforme en los respectivos espacios. Como las variables aleatorias seguirán siendo b y k las omitiremos. Aplicando teorema de Bayes obtenemos:

$$\Pr(b = 0 \mid r = f(m)) = \frac{\Pr(r = f(m) \mid b = 0) \cdot \Pr(b = 0)}{\Pr(r = f(m))}$$

Tomamos la probabilidad del denominador y la dividimos en dos casos, $b = 0$ y $b = 1$ (como hicimos varias veces en clases), obteniendo que la expresión anterior es igual a:

$$\frac{\Pr(r = f(m) \mid b = 0) \cdot \Pr(b = 0)}{\Pr(r = f(m) \mid b = 0) \cdot \Pr(b = 0) + \Pr(r = f(m) \mid b = 1) \cdot \Pr(b = 1)}$$

Como $\Pr(b = 0) = \Pr(b = 1) = 1/2$, podemos simplificar lo anterior a

$$\frac{\Pr(r = f(m) \mid b = 0)}{\Pr(r = f(m) \mid b = 0) + \Pr(r = f(m) \mid b = 1)}$$

Ahora calculamos por separado los dos términos que aparecen:

- $\Pr(r = f(m) \mid b = 0)$ es la probabilidad de que la respuesta del verificador haya sido r dado que f es la función de encriptación correspondiente a OTP. Es decir, la probabilidad de que $m \oplus k = r$. Dado que k se elige de manera uniforme y existe exactamente una llave k de las 2^n llaves posibles que cumple con $m \oplus k = r$, tenemos que $\Pr(r = f(m) \mid b = 0) = 1/2^n$.

- $\Pr(r = f(m) \mid b = 1)$ es la probabilidad de que la respuesta del verificador haya sido r dado que f es una permutación aleatoria. Esta probabilidad la hemos calculado antes en clases, y es simplemente la probabilidad de que al sacar un string de $\{0,1\}^n$ al azar obtengamos exactamente r, es decir $1/2^n$.

Tenemos entonces

$$\frac{\Pr(r = f(m) \mid b = 0)}{\Pr(r = f(m) \mid b = 0) + \Pr(r = f(m) \mid b = 1)} = \frac{\frac{1}{2^n}}{\frac{1}{2^n} + \frac{1}{2^n}} = \frac{1}{2}$$

Recapitulando, lo que hicimos fue demostrar que

$$\Pr_{b \sim \mathbb{U},\, k \sim \mathbb{U}}(b = 0 \mid r = f(m)) = \frac{1}{2}$$

Es decir, conociendo la respuesta que dio el verificador, para el adversario la probabilidad de que b haya sido 0 es exactamente $1/2$. Obviamente esto implica que la probabilidad de que b haya sido 1 es también $1/2$. Por lo tanto, si el adversario responde 0 o responde 1, su probabilidad de ganar es exactamente $1/2$.