Tarea 2

Criptografía y seguridad 2017-2

Fecha de entrega: 6 de marzo.
Calificación máxima: 16.

1. (2 puntos) Supongamos que usamos one-time pad para mensajes de ℓ bits. Si se usa la clave 0^ℓ, el mensaje cifrado es el mismo que el mensaje claro, ya que $m \oplus 0^\ell = m$. ¿Se mantiene la seguridad perfecta de one-time pad si usamos solamente claves distintas de 0^ℓ? Explica.

2. (2 puntos) Demuestra que en un esquema de cifrado perfectamente indistinguible se satisface $|\mathcal{K}| \geq |\mathcal{M}|$. Procede por contradicción, primero supón que $|\mathcal{K}| < |\mathcal{M}|$ y exhibe un adversario $\mathcal{A}$ (posiblemente con operaciones aleatorias) para el que $\Pr[\mathsf{PrivK}_{\mathcal{A}} = 1] > 1/2$.

3. (3 puntos) Demuestra si cada uno de los siguientes esquemas es perfectamente seguro, en caso contrario explica por qué.

   a) Sobre el espacio de mensajes $\mathcal{M} = \{m \in \{0,1\}^9 : m \text{ empieza con } 1\}$. Las claves se eligen aleatoriamente del conjunto $\{0,1\}^8$, y $\mathsf{Enc}_k(m) = (0||k) \oplus m$, $\mathsf{Dec}_k(c) = (0||k) \oplus c$ (el símbolo $||$ es concatenación).

   b) Cifrado de César para mensajes de tamaño uno.

   c) Cifrado de Vigenère para mensajes de tamaño ℓ usando claves de tamaño ℓ.

4. (3 puntos) Implementa el cifrado y descifrado de CESAR, AFIN, MEZCLADO y VIGENERE. El programa se llamará `cifrado` y se ejecutará

   ```
   $ cifrado [c|d] [cesar | ... | vigenere] archivoClave archivoEntrada
   ```

   donde `c|d` indica si es para cifrar o descifrar, `archivoClave` es un archivo que contiene únicamente la clave que se usará, y `archivoEntrada` pues...

   Las claves serán de esta forma:

   - CESAR. Un entero entre 0 y 255.
   - AFIN. Una pareja de enteros a, b separados por una coma. Ambos enteros están entre 0 y 255, pero si a no es válido para el cifrado afín, se deberá indicar un error.
   - MEZCLADO. Dos líneas de la misma longitud de caracteres ASCII imprimibles. En cada línea no pueden repetirse caracteres. En la primera estarán los caracteres que van a cambiarse, en la segunda estarán los caracteres nuevos. Por ejemplo:

     ```
     jBis9w8/)&%-@
     qwe23mDndjf7W
     ```

     indica que `j` se cambiará por `q`, `B` por `w`, `@` por `W`, etc.

- VIGENERE. Una cadena de caracteres. Si esta cadena es más grande que la longitud ℓ del archivo de entrada, únicamente se usarán los primeros ℓ caracteres.

El resultado cifrado o descifrado será guardado en archivos con extensión .cifrado o .descifrado, respectivamente.

Ejercicios extra

5. (2 puntos) Verifica si la función *random* de tu lenguaje favorito pasa el siguiente test. La probabilidad de que dos números enteros aleatorios sean primos relativos es $6/\pi^2$, es decir, si $a \xleftarrow{R} \mathbb{Z}$ y $b \xleftarrow{R} \mathbb{Z}$, tenemos $\Pr[\text{mcd}(a, b) = 1] = 6/\pi^2$. Genera varios números aleatorios y comprueba si se cumple la condición anterior. Escribe cuál es la función *random* que usaste, si se aproximó a 3.1415 el valor de π, y cuántos pares de números fueron necesarios para llegar a 3.14. También entrega tu programa.

6. (4 puntos) En el esquema de one-time pad se requiere que una clave no sea reusada, y esta condición también es necesaria para cualquier esquema de cifrado perfectamente seguro. Podemos verlo analizando la información liberada al cifrar dos mensajes con una misma clave. Consideremos una distribución sobre $\mathcal{M} \times \mathcal{M}$ y variables aleatorias M_1, M_2 para denotar al primero y segundo mensajes. Si escogemos una clave k aleatoria y elegimos dos mensajes (m_1, m_2) con la distribución dada, luego ciframos $c_1 = \mathsf{Enc}_k(m_1)$, $c_2 = \mathsf{Enc}_k(m_2)$, obtenemos una distribución sobre $\mathcal{C} \times \mathcal{C}$. Denotemos por C_1, C_2 a las variables aleatorias correspondientes a los mensajes cifrados.

   a) Extendemos la definición de seguridad perfecta al cifrado de dos mensajes bajo la misma clave:

   Un esquema es perfectamente seguro sobre parejas de mensajes, si para cualquier distribución sobre $\mathcal{M} \times \mathcal{M}$, cualesquiera $m_1, m_2 \in \mathcal{M}$ y $c_1, c_2 \in \mathcal{C}$ se cumple

   $$\Pr[M_1 = m_1 \text{ y } M_2 = m_2 \mid C_1 = c_1 \text{ y } C_2 = c_2] = \Pr[M_1 = m_1 \text{ y } M_2 = m_2]$$

   Demuestra que no existe ningún esquema perfectamente seguro sobre parejas de mensajes. Considera el caso en que $c_1 = c_2$.

   b) ¿Qué pasa si solo nos fijamos en las distribuciones sobre $\mathcal{M} \times \mathcal{M}$ donde ambos mensajes son distintos? Es decir, tenemos lo mismo que en el inciso anterior pero con la garantía de que $m_1 \neq m_2$. Muestra un esquema de cifrado que es perfectamente seguro sobre parejas de mensajes distintos.

7. (4 puntos) Haz un programa que automatice el criptoanálisis del cifrado de sustitución monoalfabética (alfabeto mezclado). Los mensajes claros son textos en español en mayúsculas con espacios, y además los mensajes tienen un tamaño mayor a 100 caracteres. Así que el alfabeto será el siguiente

   `ABCDEFGHIJKLMNÑOPQRSTUVWXYZespacio`

El programa devolverá cinco textos, que corresponden a cinco posibles mensajes descifrados, y se espera que entre ellos esté el correcto. Por ejemplo, el programa puede recibir la cadena

UOKUÑUOÑTYÑMUWTUIQÑUZULMAQÑRUÑLTUOKHS

y la salida será algo parecido a esto

```
1. NKTNSNKSGISWNHGNQMSNJNYWLMSUNSYGNKTOV
2. ESTE ES UN PEQUEÑO EJEMPLO DE MUESTRA
3. RHORBRHBKLBQRMKRYIBRERAQJIB RBAKRHOWÑ
4. BYIBTBYTQDTJBAQBEZTBRBPJNZTCBTPQBYILG
5. JCBJNJCNÑONGJFÑJKTNJQJHGWTNZJNHÑJCBXP
```

El programa se ejecutará así

```
$ criptoanalisis archivoCifrado
```

y mostrará en salida estándar los 5 posibles mensajes (solo los primeros 60 caracteres de cada uno).