Criptografía de Llave Pública

Introducción a la Criptografía y a la Seguridad de la Información

Iván Castellanos

Departamento de ingeniería de sistemas e industrial
Universidad Nacional de Colombia

7 de noviembre de 2019

Motivación

- Los sistemas de clave privada son muy eficentes computacionalmente
- tienen una gran desventaja: *emisor y receptor deben compartir una misma información (llave) en común*

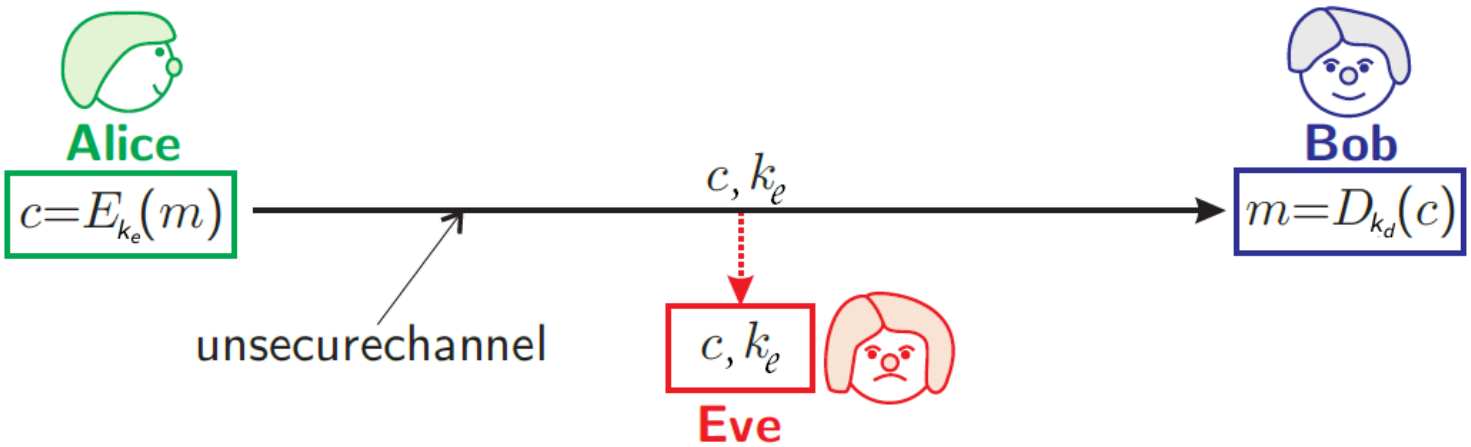
Nota

Para resolver este problema se desarrolló la criptografía de llave pública

Criptografía de llave pública

Definición (Criptografía de llave pública)

Un sistema de *cifrado publico* maneja 2 llaves diferentes k_e y k_d donde la llave de cifrado k_e es pública, la otra se mantiene secreta



Nota

La criptografía de llave pública es conocida también como criptografía *asimétrica*

Criptografía de llave pública

Un sistema de cifrado es *perfectamente seguro* si el texto cifrado no le da al adversario ninguna *información* sobre el texto limpio.

Definición (Seguridad)

En sistema es perfectamente seguro si

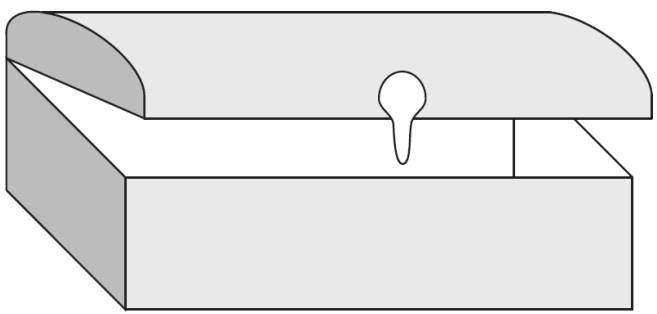
$$\forall m^* \in M,\ P(m = m^* | E_{k_e}(m) = c) = P(m = m^*)$$

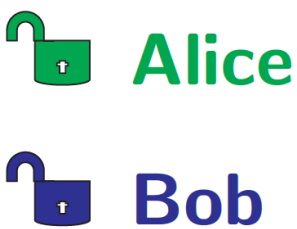
Nota

Note que en este caso no obtener ninguna información sobre el texto limpio implica que tampoco debemos obtener información sobre la llave de descifrado

Analogía

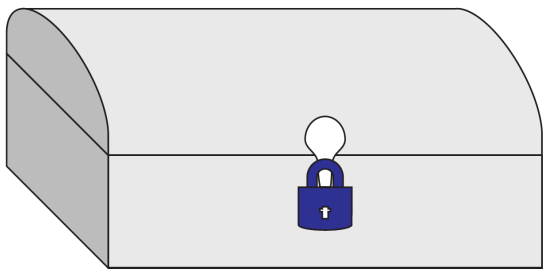
- Supongamos que Alice quiere mandarle un mensaje secreto a Bob
- Para mandar el mensaje utilizamos una caja con anillos para el cierre
- Alice y Bob tienen candados abiertos disponibles en la oficina de correos





Analogia

- Alice utiliza el candado de Bob para cerrar la caja con el mensaje y se lo manda
- Independiente de quienes tengan la caja en sus manos durante el envío el único que puede abrirla es bob con us llave



- Para responder Bob utiliza el candado de Alice y envía la caja

Nota

Note que si Alice cerrara la caja con su candado Bob no podría abrirla

Dificultad del conocimiento de la llave privada

Ejemplo

1. ¿Cuanto es 314159265358979^2?
2. ¿Cuál es la raiz cuadrada de 98696044010893382709735922441?

- La pregunta 2 es la *inversa* de la pregunta 1
- La pregunta 2 es mucho mas compleja computacionalmente que la pregunta 1

Nota

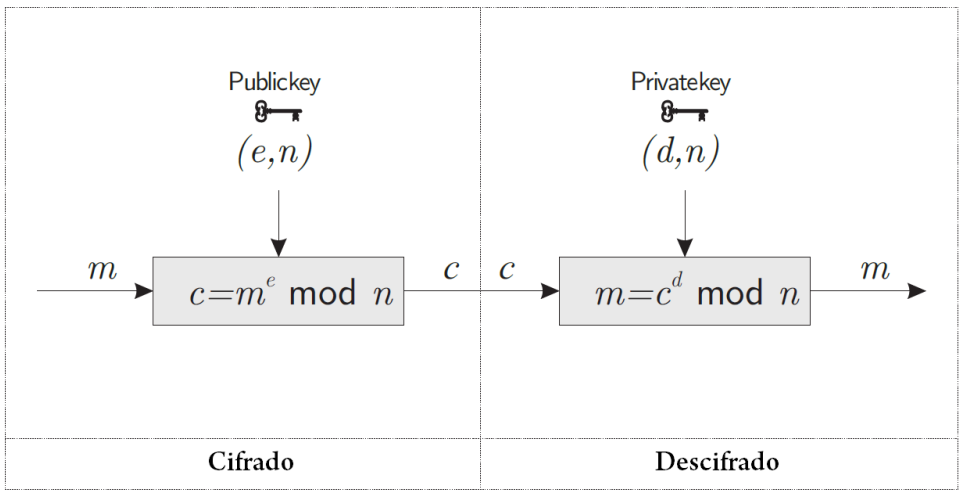
Esto mismo sucede entre la multiplicación y factorización o entre la exponenciación y el logaritmo

RSA

Definición (RSA)

El sistema de cifrado RSA es un sistema de cifrado público creado por Ronald Rivest, Adi Shamir y Leonard Adleman en 1977

- $M = C = \mathbb{Z}_n$
- $K = \mathbb{Z}^*_{\phi(n)} \times \{n\}$
- Sea $k_e = (e, n)$, $E_{k_e}(m) = m^e \bmod n$
- Sea $k_d = (d, n)$, $D_{k_d}(c) = c^d \bmod n$



Nota

La escogencia de las llaves (y los posibles mensajes) depende fuertemente del valor n

Generación de llaves

El proceso de escogencia de n y de las llaves es el siguiente:

1. Generar 2 primos grandes p y q de manera aleatoria $(p \neq q)$
2. Calcular $n = pq$, luego $\phi(n) = (p-1)(q-1)$
3. Selecionar $e \in \mathbb{Z}^*_{\phi(n)}$ aleatoramente $(e \neq 1)$
4. Calcular $d = e^{-1} (\text{mod } \phi(n))$
5. Publicar (e, n) como la llave pública
6. Conservar secreta (d, n), la llave privada

Generación de llaves

El proceso de escogencia de n y de las llaves es el siguiente:

1. Generar 2 primos grandes p y q de manera aleatoria $(p \neq q)$
2. Calcular $n = pq$, luego $\phi(n) = (p-1)(q-1)$
3. Selecionar $e \in \mathbb{Z}^*_{\phi(n)}$ aleatoramente $(e \neq 1)$
4. Calcular $d = e^{-1} (\text{mod } \phi(n))$
5. Publicar (e, n) como la llave pública
6. Conservar secreta (d, n), la llave privada

Ejercicio

Sean $p = 41$ y $q = 53$ calcule una llave publica (e, n) y su correspondiente llave privada (d, n)

RSA

Nota

La seguridad del sistema de llaves en RSA está basada en la dificultad de factorizar $n = pq$

- Para cifrar un mensaje cualquiera m podemos partir $m = m_1 m_2 m_3 ... m_t$ donde $m_i < n \ \forall i \in \{1, ..., t\}$
- para cifrar se calcula $c_i = m_i^e \mod n$ y para descifrar se calcula $m_i = c_i^d \mod$ n

Ejemplo

Considere un sistema que cifra la información de tarjetas de credito de sus usuarios, cifre $m = 6882\ 3268\ 7966\ 6683$

RSA

Ejemplo

Sean $n = 3337 = 47 * 71$, $\phi(n) = 46 * 70 = 3220$

Se escoge aleatoreamente 79, $gcd(3220, 79) = 1$ y se computa $d = e^{-1}(mod\ 3220)$, $d = 1019$

$k_e = (79, 3337)$ pública y $k_d = (1019, 3337)$

$m = 6882\ 3268\ 7966\ 6683 = m_1\, m_2\, m_3\, m_4\, m_5\, m_6$, $m_1 = 688$, $m_2 = 232$, $m_3 = 687$, $m_4 = 966$, $m_5 = 668$ y $m_6 = 3$

i	0	1	2	3	4	5	6	c
b_i	1	1	1	1	0	0	1	
m_1	688	2827	3151	1226	1426	1243	18	**1570**
m_2	232	432	3089	1438	2241	3233	805	**2756**
m_3	687	1452	2657	1894	3298	1521	900	**2091**
m_4	966	2133	1358	2140	1236	2687	2038	**2276**
m_5	668	2403	1399	1719	1716	1422	3199	**2423**
m_6	3	9	81	3224	2758	1541	2074	**158**

luego $c = 1570\ 2756\ 2091\ 2276\ 2423\ 158$

Correctitud RSA

Teorema

RSA es un sistema de cifrado correcto

Demostración.

Sea $k_e = (e, n)$ y $k_d = (d, n)$, queremos probar que $\forall m \in M$ $D_{k_d}(E_{k_e}(m)) = m$.

Para el RSA esto es $\forall m \in \mathbb{Z}_n$, $(m^e)^d \equiv m (\text{mod } n)$, que es lo mismo que probar que $m^{ed-1} \equiv 1 (\text{mod } n)$, como $d \equiv e^{-1} (\text{mod } \phi(n))$ entonces $\phi(n) \mid ed - 1$, luego $m^{ed-1} \equiv m^{\phi(n)*k} \equiv (m^{\phi(n)})^k (\text{mod } n)$ para algún $k \in \mathbb{Z}$

Cuando $gcd(m, n) = 1$ se tiene el resultado $\square$

Correctitud RSA

Teorema

RSA es un sistema de cifrado correcto

Demostración.

Sea $k_e = (e, n)$ y $k_d = (d, n)$, queremos probar que $\forall m \in M$ $D_{k_d}(E_{k_e}(m)) = m$.

Para el RSA esto es $\forall m \in \mathbb{Z}_n$, $(m^e)^d \equiv m (\text{mod } n)$, que es lo mismo que probar que $m^{ed-1} \equiv 1 (\text{mod } n)$, como $d \equiv e^{-1} (\text{mod } \phi(n))$ entonces $\phi(n) \mid ed - 1$, luego $m^{ed-1} \equiv m^{\phi(n)*k} \equiv (m^{\phi(n)})^k (\text{mod } n)$ para algún $k \in \mathbb{Z}$

Cuando $gcd(m, n) = 1$ se tiene el resultado □

Ejercicio

Demuestre que 1. $m^{ed} \equiv m (\text{mod } p)$ y $m^{ed} \equiv m (\text{mod } q)$ para probar que 2. $m^{ed} \equiv m (\text{mod } n)$ cuando $gcd(m, n) \neq 1$

ElGamal

Definición (ElGamal)

El cifrado ElGamal descrito por Taher ElGamal en 1984 se basa en la dificultad del problema del logaritmo discreto

Sea p primo

- $M = \mathbb{Z}_p^*$
- $C = \mathbb{Z}_p^* \times \mathbb{Z}_p^*$
- $K = \{(p, \alpha, x, \beta) : \beta \equiv \alpha^x (\text{mod } p)\}$ donde $\alpha, x, \beta \in \mathbb{Z}_p^*$, α debe ser un elemento primitivo
- Los valores p, α y β son públicos, x se mantiene secreto
- Sea $y \in \mathbb{Z}_p^*$ y $k = (p, \alpha, x, \beta)$ escogido de manera aleatoria $E_k(m) = (\alpha^y \text{ mod } p, m\beta^y \text{ mod } p)$
- sea $c = (\gamma, \delta)$, $D_k(c) = \delta(\gamma^x)^{-1} \text{ mod } p$

Generación de llaves

- Generar un primo grande p y α un elemento primitivo (generador) de $\mathbb{Z}_p$
- Seleccionar un entero grande x $(1 < x \leq p-2)$
- Calcular $\beta = \alpha^x \bmod p$
- Publicar (p, α, β) como la llave pública
- Conservar x de manera secreta

Generación de llaves

- Generar un primo grande p y α un elemento primitivo (generador) de $\mathbb{Z}_p$
- Seleccionar un entero grande x $(1 < x \leq p-2)$
- Calcular $\beta = \alpha^x \bmod p$
- Publicar (p, α, β) como la llave pública
- Conservar x de manera secreta

Ejercicio

Calcule una llave pública y una llave privada de ElGamal para $p = 59$

ElGamal

Nota

$x = \log_\alpha \beta (\text{mod } p)$, luego si hubiera un algoritmo eficiente que calcule el logaritmo discreto romperiamos el sistema de cifrado

- Para cifrar un mensaje cualquiera m podemos partir $m = m_1 m_2 ... m_t$ tal que $m_i \in \mathbb{Z}_p$
- En el cifrado se selecciona un entero aleatorio y_i, $1 < y_i \leq p-2$ y se calcula $c_i = (\alpha^{y_i} \text{ mod } p, m\beta^{y_i} \text{ mod } p)$
- para descifrar $c_i = (\gamma, \delta)$ se calcula $m_i = \delta(\gamma^x)^{-1} \text{ mod } p$

ElGamal

Ejemplo

Cifrar $m = 1299$ con ElGamal usando $p = 2579$

Se selecciona $\alpha = 2$ elemento primitivo (generador) de $\mathbb{Z}_{2579}$

Como llave privada se escoge $x = 765$ y se calcula
$\beta = 2^{765} \bmod 2579$

i	0	1	2	3	4	5	6	7	8	9	β
b_i	1	0	1	1	1	1	1	1	0	1	
α	2	4	16	256	1061	1277	801	2009	2525	337	**949**

Luego la llave pública $(2579, 2, 949)$ y la privada es 765

ElGamal

Ejemplo

Aleatoreamente se escoge $y = 853$

Se calcula $\gamma = 2^{853} \mod 2579$

i	0	1	2	3	4	5	6	7	8	9	γ
b_i	1	0	1	0	1	0	1	0	1	1	
α	2	4	16	256	1061	1277	801	2009	2525	337	**435**

Se calcula $\delta' = 949^{853} \mod 2579$

i	0	1	2	3	4	5	6	7	8	9	δ'
b_i	1	0	1	0	1	0	1	0	1	1	
α	949	530	2368	678	622	34	1156	414	1182	1885	**2424**

Se calcula $\delta = 1299 * 2424 \mod 2579 = 2396$

Luego $c = (435, 2396)$

ElGamal

Nota

Note que y_i no se utiliza en el descifrado ni en la comunicación, por lo que ese valor se desecha luego del cálculo

Nota

Si se logra calcular $y_i = \log_\alpha \alpha^{y_i} (\text{mod } p)$ o $x = \log_\alpha \beta (\text{mod } p)$ se rompe el cifrado completamente

Nota

Podemos generalizar ElGamal para que funcione con cualquier cuerpo finito $\mathbb{F}$ en lugar de $\mathbb{Z}_p$

Rabin

- Sean p, q primos y $n = pq$
- El *cifrado de Rabin* se basa en la dificultad de calcular raices cuadradas de un número en $\mathbb{Z}_n$ sin conocer p y q
- La llave pública es n, la privada es (p, q)
- Para el cifrado $E_{k_e}(m) = m^2 \bmod n$
- Calculamos $a, b \in \mathbb{Z}$ tal que $ap + bq = 1$, $r = c^{\frac{p+1}{4}} \bmod p$ y $s = c^{\frac{q+1}{4}} \bmod q$, para el descifrado se calcula $x = (aps + bqr) \bmod n$ y $y = (aps - bqr) \bmod n$. m es igual a x, $-x$, y o $-y$ (pues hay 4 raices módulo n)

Merkle-Hellman

- El *cifrado de Merkle-Hellman* se basa en la dificultad de resolver el *problema de la mochila* (cada objeto una única vez)
- Se toma una instancia sencilla de resolver del problema de la mochila, $b_1, b_2, ..., b_n$ (supercreciente) y $s = \sum_{i=1}^{n} x_i b_i$ donde $x_i \in \{0, 1\}$
- Se tienen 2 valores $k > b_1 + b_2 + ... + b_n$ y $w \in \mathbb{Z}_k^*$ y una permutación π de $\{1, 2, ..., n\}$
- Se calcula la secuencia $a_i = w * b_{\pi(i)} \mod k$ para $i = 1, ..., n$
- La llave pública es $(a_1, a_2, ..., a_n)$ la llave privada es $(k, w, (b_1, b_2, ..., b_n))$
- Sea $m = m_1 m_2 ... m_n$ una cadena de n bits, $E_{k_e}(m) = \sum_{i=0}^{n} a_i m_i$
- El descifrado corresponde a la solución del problema de la mochila tomando la secuencia $d_i = w^{-1} a_i \mod k$
- Note que la secuencia $d_1, d_2, ..., d_n$ es la misma $b_i, b_2, ..., b_n$ en otro orden

Cifrado probabilístico

Nota

Cifrados como RSA, Rabin y Merkle-Hellman son determinísticos en el sentido que para $m_1 = m_2$, $E_k(m_1) = E_k(m_2)$, es decir, un mismo mensaje siempre se cifra al mismo texto cifrado

Definición

Un cifrado probabilístico utiliza la *aleatorización* para que lo anterior no suceda, es decir, que un mismo mensaje pueda ser cifrado a diferentes textos sin romper la correctitud del cifrado

Ejemplo

ElGamal es un cifrado probabilístico, otros ejemplos de cifrados probabilísticos son Goldwasser-Micali y Blum-Goldwasser