Clase 17: cifrado híbrido, el modelo KEM-DEM

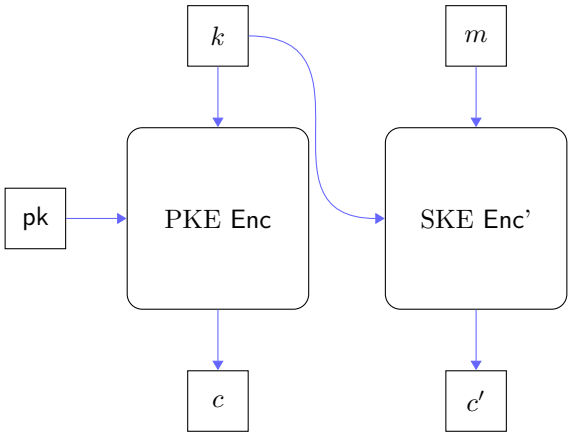
Fernando Virdia, versión: 0.0.1, junio 2024

9.1 Cifrado híbrido

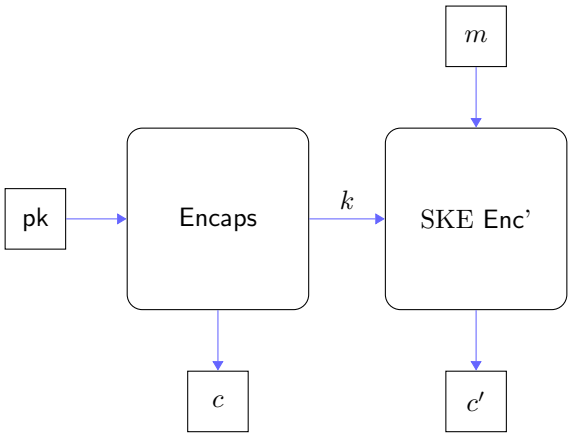
Es interesante observar que en practica, el costo de un cifrado a clave publica es generalmente mas alto del de un cifrado simétrico. Lanzando `openssl speed rsa aes` en mi laptop:

- En 10s, 6852 descifrados y 344141 cifrados de RSA-3072 ("128 bits de seguridad"), mensaje de ≈ 3000 bits
- En 10s, 941746 cifrados o descifrados de AES-128 CBC-mode, mensaje de ≈ 131072 bits

Por este motivo, aunque es técnicamente posible usar sistemas como ElGamal para mensajes arbitrariamente largos, en la practica PKE y KEX ambos utilizan una primera fase a clave publica para generar claves para un SKES. En el caso de PKE, este método se llama de cifrado híbrido ("hybrid encryption"). Por ejemplo, sean $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ un PKES, $\Pi' = (\mathsf{Gen}', \mathsf{Enc}', \mathsf{Dec}')$ un SKES. Podemos cifrar una clave k de Π' utilizando Π, y luego cifrar un mensaje arbitrariamente largo utilizando $\mathsf{Enc}'(k, \cdot)$. Transmitiendo ambos cifrados, es posible recuperar el mensaje.



Una técnica mas común hoy en día es remplazar el PKES con un mecanismo de encapsulamiento de claves (KEM, "key-encapsulation mechanism"). Se puede pensar un KEM cono un sistema de cifrado de "mensajes al azar". Esta restricción generalmente le permite a un KEM ser mas compacto y eficiente que un PKES que ofrece similar seguridad.



Un sistema híbrido que utiliza un KEM se lo llama también KEM-DEM, donde el cifrado simétrico tiene el rol de mecanismo de encapsulamiento de datos (DEM, "data-encapsulation mechanism").

En la próxima clase, vamos a definir que es un KEM, y vamos a proponer definiciones de seguridad PKE y KEM cercanas a AE, o sea "IND-CCA". Vamos a ver un KEM que ofrece OT-IND-CCA, y utilizarlo para construir un PKE "KEM-DEM" que ofrece seguridad IND-CCA.

Así como en el caso de la criptografía simétrica, la noción de IND-CPA no nos protege de un adversario capaz de manipular cifrados. Por ejemplo, si $c = (c_0, c_1) = (g^y, g^{xy} \cdot m_1)$ cifra un elemento $m_1 \in \mathbb{G}$, un adversario puede generar $c' = (c_0, c_1 \cdot m_2)$ para obtener un cifrado de $m_1 \cdot m_2$, de manera parecida a como se modifican cifrados simétricos CTR.

En el caso de SKES, definimos la noción de cifrado autenticado (AE), para capturar este tipo de ataque. En este caso, definimos una nueva noción, de indistinguibilidad bajo ataques de cifrado elegido (IND-CCA), donde le otorgamos al adversario un nuevo oráculo $D(c)$ que dado un cifrado c retorna $\mathsf{Dec}(\mathsf{sk}, c)$ al adversario.

Una similar noción es posible con SKES también, pero innecesaria dado que AE $\Rightarrow$ IND-CCA. Cifrados elegidos son una mayor preocupación en el caso de PKE, dado que cualquiera con la clave publica puede cifrar, y que las claves publicas pueden quedar en uso mas tiempo de las claves secretas de los cifrados simétricos.

En la practica, un oráculo de descifrado podría suceder en sistemas de cifrado de email, donde un proveedor de correo electrónico "Daniel" podría modificar un correo "`De Alicia para Bob:` $\mathsf{Enc}(\mathsf{pk}_{\text{Bob}}, m)$" en "`De Dan para Bob:` $\mathsf{Enc}(\mathsf{pk}_{\text{Bob}}, m)$". Al responderle a Dan, Bob podría citar los contenidos originales de m, resultando en un oráculo de descifrado.

Del punto de vista mas teórico, otorgar un oráculo $D(c)$ y de cualquier manera obtener seguridad IND-CCA captura la noción de que $\mathcal{A}$ no puede obtener $c = \mathsf{Enc}(\mathsf{pk}, m_b)$, modificarlo en $c' \neq c$ y recuperar $m_b \leftarrow D(c')$, por lo que aunque el cifrado no está "autenticado" de la misma manera de que un cifrado simétrico AE si lo está, modificar cifrados en transito no le otorga al adversario ninguna ventaja respecto a extrapolar información sobre el mensaje cifrado.

Definición 31 (OT- IND-CCA) *Sea* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *un PKES.*

$\mathsf{Exp}^{\text{OT-IND-CCA}}(\mathcal{A}, b)$

1: $(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{Gen}()$
2: $\mathsf{disabled} \leftarrow 0$
3: $Q \leftarrow \{\}$
4: $b' \leftarrow \mathcal{A}^{\mathcal{O}(\cdot,\cdot), D(\cdot)}(\mathsf{pk})$
5: **return** b'

$\mathcal{O}(m_0, m_1)$

1: **if** $\mathsf{disabled} = 1$:
2: **return** $\perp$
3: $\mathsf{disabled} \leftarrow 1$
4: $c \leftarrow \mathsf{Enc}(\mathsf{pk}, m_b)$
5: $Q \leftarrow Q \cup \{c\}$
6: **return** c

$D(c)$

1: **if** $c \in Q$:
2: **return** $\perp$
3: $m \leftarrow \mathsf{Dec}(\mathsf{sk}, c)$
4: **return** m

Decimos que Π *otorga* $(\varepsilon, t, q_e, q_d)$*-indistinguibilidad bajo ataques de cifrado elegido (IND-CCA, "indistinguishibility under chosen-ciphertext attacks"), si cualquier adversario* $\mathcal{A}$ *que corre en tiempo* $\leq t$*, realiza* $\leq q_e$ *queries a* $\mathcal{O}$ *(one-time:* $q_e = 1$*) y realiza* $\leq q_d$ *queries a* D*, tiene ventaja*

$$\mathsf{Adv}(\mathsf{Exp}^{\text{OT-IND-CCA}}, \mathcal{A}) := \left|\Pr[\mathsf{Exp}^{\text{OT-IND-CCA}}(\mathcal{A}, 0) \Rightarrow 1] - \Pr[\mathsf{Exp}^{\text{OT-IND-CCA}}(\mathcal{A}, 1) \Rightarrow 1]\right| \leq \varepsilon.$$

Lemma 18 (OT-IND-CCA $\Rightarrow$ IND-CCA) *Sea* $\Pi = (\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *un PKES que ofrece* $(\varepsilon, t, 1, q_d)$*-IND-CCA (i.e. OT-IND-CCA). Entonces ofrece* $(q_e \cdot \varepsilon, \approx t, q_e, q_d)$*-IND-CCA con* $q_e \in \mathbb{Z}_{>0}$*.*

Demostración. La demostración es idéntica a la de Lemma 17. □

Para obtener PKE que otorgue seguridad IND-CCA, vamos a utilizar una construcción KEM-DEM.

Definición 32 (KEM, mecanismo de encapsulado de claves) *Sean* $\mathcal{K}_p$*,* $\mathcal{K}_s$*,* $\mathcal{M}$*,* $\mathcal{C}$ *conjuntos. Decimos que* $\Pi = (\mathsf{Gen}, \mathsf{Encaps}, \mathsf{Decaps})$ *forman un mecanismo de encapsulamiento de claves ("key-encapsulation mechanism", KEM) con espacio de claves publicas* $\mathcal{K}_p$*, de claves privadas* $\mathcal{K}_s$*, de claves a encapsular* $\mathcal{M}$*, y de cifrados o "encapsulados"* $\mathcal{C}$*, si*

- Gen *es un algoritmo aleatorio,* $\mathsf{Gen}\colon \emptyset \to \mathcal{K}_s \times \mathcal{K}_p$ *que retorna una tupla (*sk*,* pk*);*
- Encaps *es un algoritmo aleatorio,* $\mathsf{Encaps}\colon \mathcal{K}_p \to \mathcal{M} \times \mathcal{C}$ *que retorna una clave* k *y un encapsulamiento* c *de* k*;*
- Decaps *es un algoritmo determinista* $\mathsf{Dec}\colon \mathcal{K}_s \times \mathcal{C} \to \mathcal{M}$*, que decapsula* c*;*

donde Gen*,* Encaps *y* Decaps *son "eficientes". Decimos que* Π *es* δ*-correcto si*

$$\Pr_{\substack{(\mathsf{sk},\mathsf{pk})\sim\mathsf{Gen}(),\\ \$\ en\ \mathsf{Encaps}}}[(k, c) \leftarrow \mathsf{Enc}(\mathsf{pk}; \$) : \mathsf{Decaps}(\mathsf{sk}, c) \neq k] \leq \delta.$$

Comentario 55 *En el resto del curso, consideramos solamente KEMs* 0*-correctos.*

Ahora podemos definir una noción de "KEM IND-CCA".

Definición 33 (OT-IND-CCA KEM) *Sea* $\Pi = (\mathsf{Gen}, \mathsf{Encaps}, \mathsf{Decaps})$ *un KEM.*

$\mathsf{Exp}^{\text{KEM-OT-CCA}}(\mathcal{A}, b)$	$D(c)$
1 : $(\mathsf{sk}, \mathsf{pk}) \leftarrow \mathsf{Gen}()$	1 : **if** $c = c^*$:
2 : $(k_0, c^*) \leftarrow \mathsf{Encaps}(pk)$	2 : **return** $\perp$
3 : $k_1 \xleftarrow{\$} \mathcal{M}$	3 : $k \leftarrow \mathsf{Decaps}(\mathsf{sk}, c)$
4 : $b' \leftarrow \mathcal{A}^D(\mathsf{pk}, k_b, c^*)$	4 : **return** k
5 : **return** b'	

Decimos que Π *otorga "one-time"* (ε, t, q_d)*-indistinguibilidad bajo ataques de cifrado elegido (KEM IND-CCA, "indistinguishibility under chosen-plaintext attacks"), si cualquier adversario* $\mathcal{A}$ *que corre en tiempo* $\leq t$*, y realiza* $\leq q_d$ *queries a* D*, tiene ventaja*

$$\mathsf{Adv}(\mathsf{Exp}^{\text{KEM-OT-CCA}}, \mathcal{A}) \coloneqq \left| \Pr[\mathsf{Exp}^{\text{KEM-OT-CCA}}(\mathcal{A}, 0) \Rightarrow 1] - \Pr[\mathsf{Exp}^{\text{KEM-OT-CCA}}(\mathcal{A}, 1) \Rightarrow 1] \right| \leq \varepsilon.$$