

Bezpečnosť asymetrického šifrovania

Cryptology (1)

Martin Stanek

2025

KI FMFI UK Bratislava

- asymetrické šifrovanie
- formálny prístup k definícii bezpečnosti šifrovacích schém
- IND-CPA a IND-CCA2
- dôkazy redukciou na podkladový problém
- ElGamalova schéma
- Fujisakiho-Okamoto transformácia

šifrovacia schéma $\Pi = \langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$

– **Gen** – znáhodnený polynomiálny algoritmus

- vstup: 1^k (k je bezpečnostný parameter schémy)
- výstup: (pk, sk) , kde pk je verejný kľúč a sk je súkromný kľúč
- priestor správ $\mathcal{M}$ je obvykle určený vygenerovanými kľúčmi

– **Enc** – znáhodnený polynomiálny algoritmus

- vstup: $pk, m \in \mathcal{M}$
- výstup: šifrový text $c \leftarrow \text{Enc}_{pk}(m)$

– **Dec** – deterministický polynomiálny algoritmus

- vstup: sk, c
- výstup: otvorený text $m = \text{Dec}_{sk}(c)$, resp. chybový symbol $\perp$

– korektnosť schémy: $\forall k \in \mathbb{N}; \forall (pk, sk) \leftarrow \text{Gen}(1^k) \forall m \in \mathcal{M} : \text{Dec}_{sk}(\text{Enc}_{pk}(m)) = m$

- definícia bezpečnosti – kombinácia cieľa a schopností útočníka
- najsilnejšia definícia = najslabší cieľ + najsilnejšie schopnosti
- **ciele:**
 - získať súkromný kľúč (úplné rozbitie schémy)
 - získať otvorený text k danému šifrovému textu
 - vypočítať netriviálnu informáciu o otvorenom texte (sémantická bezpečnosť)
 - rozlíšiť správny otvorený text prislúchajúci k šifrovému textu (nerozlíšiteľnosť)
 - modifikovať šifrový text tak, že (neznáme) otvorené texty sú známym spôsobom závislé (nepoddajnosť)
 - skonštruovať korektný šifrový text bez znalosti prislúchajúceho otvoreného textu
- **schopnosti:**
 - znalosť verejného kľúča (teda aj možnosť čokoľvek zašifrovať)
 - prístup k dešifrovaciemu orákulu (potenciálne s rôznymi obmedzeniami)

Definícia bezpečnosti – experiment (IND-CPA)

- výpočtovo obmedzený (polynomiálny) útočník A
- útočník má prístup k verejnému kľúču
- cieľ: rozlíšiť, z ktorého z dvoch otvorených textov vznikol šifrový text

Ind_{A,Π}^{cpa}(k):

$(pk, sk) \leftarrow \text{Gen}(1^k)$

$(m_0, m_1, s) \leftarrow A(pk)$, kde $|m_0| = |m_1|$ a s reprezentuje stav výpočtu

$c \leftarrow \text{Enc}_{pk}(m_b)$, kde $b \in_R \{0, 1\}$

$b_A \leftarrow A(s, c)$

return $b_A \stackrel{?}{=} b$ (1 ak A určil b správne, inak 0)

Definícia. Šifrovacia schéma $\Pi = \langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$ je IND-CPA bezpečná, ak pre ľubovoľného PPT útočníka A existuje zanedbateľná funkcia $\text{negl}(\cdot)$ taká, že platí:

$$\Pr[\text{Ind}_{A,\Pi}^{\text{cpa}}(k) = 1] \leq \frac{1}{2} + \text{negl}(k).$$

– $f : \mathbb{N} \rightarrow \mathbb{R}_0^+$ je (asymptoticky) *zanedbateľná*, ak pre ľubovoľný kladný polynóm p :

$$\exists n_0 \in \mathbb{N} \ \forall n > n_0 : \quad f(n) < \frac{1}{p(n)}$$

– *nejakú* zanedbateľnú funkciu označíme $\text{negl}(\cdot)$

- asymptotická definícia
- IND-CPA nie je *najsilnejšia* definícia
- IND-CPA a iné varianty je možné definovať aj pre symetrické šifrovacie schémy
 - CPA: prístup k šifrovaciemu orákulu
 - používané napríklad pri analýze módov blokových šifier
- prečo nemôže byť definícia IND-CPA “nejako perfektná”, teda
 - s výpočtovo neobmedzeným útočníkom $\rightarrow$ dokáže dešifrovať s pravd. 1
 - s $\Pr[\cdot] = 1/2 \rightarrow$ aj polynomiálny útočník sa dokáže odchýliť od $1/2$
- deterministické schémy (s deterministickým Enc) nemôžu byť IND-CPA bezpečné
 - A vie šifrovať a výpočtom $\text{Enc}_{pk}(m_0)$, $\text{Enc}_{pk}(m_1)$ určí b s pravdepodobnosťou 1
- nerozlíšiteľnosť ekvivalentná pojmu *sémantická bezpečnosť* (nemožnosť vypočítať netriviálnu informáciu o otvorenom texte)

- $\text{Gen}(1^k)$:
 - vygenerujeme grupu $(G, \cdot)$ prvočíselného rádu q (kde $|q| = k$) a jej generátor g
 - $\text{sk} = x \in_R \mathbb{Z}_q$; $\text{pk} = y$, kde $y = g^x$
 - parametre grupy (G, q, g) sú všeobecne známe, prípadne sú súčasť sk a pk
 - teda napr. $\text{pk} = (G, q, g, y)$
 - priestor správ $\mathcal{M} = G$
- $\text{Enc}_{\text{pk}}(m) = (g^t, y^t \cdot m)$, kde $t \in_R \mathbb{Z}_q$ (znáhodnené šifrovanie)
- $\text{Dec}_{\text{sk}}(r, s) = s \cdot r^{-x}$
- korektnosť schémy: $s \cdot r^{-x} = y^t \cdot m \cdot r^{-x} = g^{xt} \cdot m \cdot g^{-tx} = m$

- parametre: G, q, g ako v schéme
- problém diskretného logaritmu (DLOG):
 - pre dané $y \in_R G$ vypočítať $x \in \mathbb{Z}_q$: $g^x = y$
- výpočtový Diffieho-Hellmanov problém (CDH):
 - pre dané g^a, g^b , kde $a, b \in_R \mathbb{Z}_q$ vypočítať g^{ab}
- rozhodovací Diffieho-Hellmanov problém (DDH):
 - pre danú trojicu prvkov z G rozhodnúť jej *typ*
 - skutočná DH trojica: (g^a, g^b, g^{ab}) , kde $a, b \in_R \mathbb{Z}_q$
 - náhodná trojica: (g^a, g^b, g^c) , kde $a, b, c \in_R \mathbb{Z}_q$

- formalizácia predpokladu, že DDH problém je v danej grupe ťažký
- definujme výhodu útočníka/rozlišovača D :

$$\text{Adv}_{G,q,g}^{\text{ddh}}(D) = \left| \Pr[a, b \in_R \mathbb{Z}_q : D(g^a, g^b, g^{ab}) = 1] \right. \\ \left. - \Pr[a, b, c \in_R \mathbb{Z}_q : D(g^a, g^b, g^c) = 1] \right|$$

- (DDH predpoklad) Pre ľubovoľného PPT útočníka D existuje negl:

$$\text{Adv}_{G,q,g}^{\text{ddh}}(D) \leq \text{negl}(k)$$

Veta. ElGamalova schéma je IND-CPA bezpečná, ak pre príslušnú grupu platí DDH predpoklad.

Dôkaz (1) – modifikovaná schéma Π^*

- nech Π je ElGamalova schéma, A je ľubovoľný PPT útočník
- označme $\varepsilon(k) = \Pr[\text{Ind}_{A,\Pi}^{\text{cpa}}(k) = 1]$
- vytvorme modifikovanú schému Π^*
 - rovnaký algoritmus Gen
 - upravený $\text{Enc}^*(m) = (g^t, g^u \cdot m)$, kde $t, u \in_R \mathbb{Z}_q$
 - na dešifrovacom algoritme nezáleží (ani nie je možné dešifrovať)
 - Ind^{cpa} experiment závisí len na Gen a Enc
- pozorovania pre šifrový text v Π^* , teda $(g^t, g^u \cdot m)$:
 - prvý komponent nezávisí na m a je uniformne distribuovaný v G
 - pre ľubovoľné m je druhý komponent uniformne distribuovaný v G (lebo g^u je)
 - teda $g^u \cdot m$ nenesie žiadnu informáciu o m (*one-time pad* nad G)
 - preto $\Pr[\text{Ind}_{A,\Pi^*}^{\text{cpa}}(k) = 1] = \frac{1}{2}$

- skonštruujeme PPT rozlišovač D pre DDH problém
 - vstup: G, q, g a trojica prvkov z G , označme ich g_1, g_2, g_3
- D použije útočníka A (bude ho simulovať):
 1. položí $pk = (G, q, g, g_1)$, teda $y = g_1$
 2. simuluje: $(m_0, m_1, s) \leftarrow A(pk)$
 3. zvolí náhodné $b \in_R \{0, 1\}$ a vytvorí šifrový text $c = (g_2, g_3 \cdot m_b)$
 4. simuluje: $b_A \leftarrow A(s, c)$
 5. return $b_A = b$ (teda 1 pri zhode, 0 inak)

1. na vstupe D je náhodná trojica (g^x, g^t, g^u)
 - šifrový text zodpovedá šifrovaniu v schéme Π^* , $c = (g^t, g^u \cdot m_b)$
 - D vráti 1 práve vtedy, keď A uspeje v Ind^{cpa} experimente pre Π^*
 - teda $\Pr[D(g^x, g^t, g^u) = 1] = \Pr[\text{Ind}_{A, \Pi^*}^{\text{cpa}}(k) = 1] = \frac{1}{2}$
2. na vstupe D je DH trojica (g^x, g^t, g^{xt})
 - šifrový text zodpovedá šifrovaniu v schéme Π , $c = (g^t, g^{xt} \cdot m_b)$
 - D vráti 1 práve vtedy, keď A uspeje v Ind^{cpa} experimente pre Π
 - teda $\Pr[D(g^x, g^t, g^{xt}) = 1] = \Pr[\text{Ind}_{A, \Pi}^{\text{cpa}}(k) = 1] = \varepsilon(k)$

- keďže platí DDH predpoklad rozdiel pravdepodobností oboch prípadov pre D musí byť zanedbateľný:

$$\left| \Pr[D(g^x, g^t, g^{xt}) = 1] - \Pr[D(g^x, g^t, g^u) = 1] \right| = \left| \varepsilon(k) - \frac{1}{2} \right| \leq \text{negl}(k)$$

- odtiaľ dostávame: $\varepsilon(k) \leq \frac{1}{2} + \text{negl}(k)$

- alternatívny spôsob dokazovania – *game hopping*
- definujeme postupnosť hier (experimentov, ako napr. Ind^{cpa})
 - začíname s hrou podľa definície (G_0)
 - postupne upravujeme niektoré časti (voľba parametrov, výpočet a pod.)
 - končíme s hrou, v ktorej útočník nemôže vyhrať (G_n)
- ukážeme, že rozdiel pravdepodobností výhry medzi G_i a G_{i+1} je zanedbateľný, napr.
 - nerozlíšiteľnosť na základe nejakého problému
 - identické hry až na chybovú udalosť so zanedbateľnou pravdepodobnosťou
 - matematicky identické hry (len inak, ekvivalentne zapísané)
- V. Shoup: *Sequences of Games: A Tool for Taming Complexity in Security Proofs*, 2006

Definícia bezpečnosti – experiment (IND-CCA2)

- výpočtovo obmedzený (polynomiálny) útočník A
- útočník má prístup k verejnému kľúču a k dešifrovaciemu orákulu $\text{Dec}_{\text{sk}}(\cdot)$
- cieľ: rozlíšiť, z ktorého z dvoch otvorených textov vznikol šifrový text

$\text{Ind}_{A,\Pi}^{\text{cca2}}(k)$:

$(\text{pk}, \text{sk}) \leftarrow \text{Gen}(1^k)$

$(m_0, m_1, s) \leftarrow A^{\text{Dec}_{\text{sk}}}(\text{pk})$, kde $|m_0| = |m_1|$ a s reprezentuje stav výpočtu

$c \leftarrow \text{Enc}_{\text{pk}}(m_b)$, kde $b \in_R \{0, 1\}$

$b_A \leftarrow A^{\text{Dec}_{\text{sk}}}(s, c)$, pričom sa A nesmie pýtať priamo $\text{Dec}_{\text{sk}}(c)$

return $b_A \stackrel{?}{=} b$ (1 ak A určil b správne, inak 0)

Definícia. Šifrovacia schéma $\Pi = \langle \text{Gen}, \text{Enc}, \text{Dec} \rangle$ je IND-CCA2 bezpečná, ak pre ľubovoľného PPT útočníka A existuje zanedbateľná funkcia $\text{negl}(\cdot)$ taká, že platí:

$$\Pr[\text{Ind}_{A,\Pi}^{\text{cca2}}(k) = 1] \leq \frac{1}{2} + \text{negl}(k).$$

- variant experimentu a definície IND-CCA1: A po získaní c nemá vôbec prístup k Dec_{sk} orákulu

- A postupuje takto (využije poddajnosť schémy):
 1. zvolí ľubovoľné $m_0 \neq m_1 \in G$
 2. získa šifrový text $c = (r, s)$
 3. využije prístup k dešifrovaciemu orákulu: $\text{Dec}_{sk}(r, s \cdot m')$, kde $m' \in G$ je ľubovoľné
 4. získaný otvorený text $m_b \cdot m'$ použije na určenie b

- RSA-OAEP
 - model s náhodným orákulom
 - RSA predpoklad (nie tesná redukcia)
 - iné konštrukcie ako RSA použité s OAEP vyžadujú *partial-domain one-wayness*
- Cramer-Shoup
 - *štandardné* kryptografické predpoklady
 - DDH predpoklad
 - hašovacia funkcia s vlastnosťou *target collision resistance / universal one-wayness* ($\approx$ hľadanie druhého vzoru vopred zvolenej správy pre rodinu hašovacích funkcií)
- ideálne modely (náhodné orákulá, ideálne šifry a pod.) vs. štandardné predpoklady
 - *neinštancovateľné* schémy

- Fujisaki, Okamoto (1999)
- ako dosiahnuť IND-CCA2 bezpečnú schému skombinovaním slabších schém
 - hybridná konštrukcia (šifrovanie)
 - generická konštrukcia v modeli s náhodným orákulom
 - slabšia asymetrická schéma (v podstate OW-CPA, jednosmernosť)
 - γ -uniformita pre asymetrickú schému (pravd. ľub. dvojice OT, ŠT je $\leq \gamma$)
 - slabšia symetrická konštrukcia (Find-Guess $\sim$ IND, útočník bez šifr. orákula)
- viaceré vylepšenia a modifikácie neskôr
- redukcia nie je tesná

- Gen – ako predtým, m je z priestoru správ symetrickej šifry
- H_1, H_2 náhodné orákulá s vhodnými definičnými obormi aj obormi hodnôt
- v ElGamalovej schéme šifrujeme náhodne zvolenú správu σ

$\text{Enc}_{\text{pk}}(m) \mapsto (r, s, c) :$

$$\sigma \in_R G$$

$$t = H_2(\sigma, m)$$

$$(r, s) = (g^t, \sigma \cdot y^t)$$

$$c = m \oplus H_1(\sigma) \quad \text{resp: } E_{H_1(\sigma)}^{\text{sym}}(m)$$

$\text{Dec}_{\text{sk}}(r, s, c) :$

$$\sigma' = s \cdot r^{-x}$$

$$m' = c \oplus H_1(\sigma') \quad \text{resp: } D_{H_1(\sigma')}^{\text{sym}}(c)$$

$$\text{if } s = \sigma' \cdot y^{H_2(\sigma', m')} : m = m'$$

else: $m = \perp$ (explicitné zamietnutie)

return m

- dokázateľná bezpečnosť $\nRightarrow$ bezpečná schéma v praxi
- dôkazy môžu byť chybné (pôvodný dôkaz pre RSA-OAEP)
- definície bezpečnosti nemusia zodpovedať praktickým potrebám
- inštancie schém nezodpovedajúce pravdepodobnostným odhadom (nie tesné redukcie)
- predpoklady bezpečnosti nemusia byť naplnené (náhodné orákulá)
- nevhodná implementácia (útoky postrannými kanálmi, ukladanie kľúčov, entropia, ...)
- obídenie kryptografie (ľudský faktor)
- *napriek všetkému to nie je zbytočné*

1. Zanedbateľné funkcie

- a) Ktoré z nasledujúcich funkcií sú zanedbateľné? 2^{-n} , $\frac{1}{\sqrt{n}}$, $n^{-\ln n}$, $2^{-\pi \cdot \ln n}$
- b) Ak $f(n)$ je zanedbateľná, tak aj $f(n) \cdot n^3$ je zanedbateľná?

2. Skonstruujme šifrovaciu schému z RSA: $c = (r^e \bmod n, \text{lsb}(r) \oplus m)$, kde $m \in \{0, 1\}$ a $r \in_R \mathbb{Z}_n$ (lsb označuje najmenej významný bit). Zdôvodnite, že táto schéma je IND-CPA bezpečná. Aký predpoklad je potrebný? Je schéma aj IND-CCA2 bezpečná?

3. Ak by sme pri dešifrovaní v FO transformácii pre ElGamalovu schému a OTP kontrolovali, či $r = g^{H_2(\sigma', m')}$ namiesto testu $s = \sigma' \cdot y^{H_2(\sigma', m')}$, stalo by sa niečo zlé? Zdôvodnite.