25. 10. 2017

Minule jsme si ukázali aplikace pseudonáhodných generátorů (značíme *PRG* z anglického Pseudo Random Generator). Cílem této přednášky je ukázat, že *PRG* je možné vytvořit z jednosměrných funkcí (tedy za poměrně slabého předpokladu). Jednosměrné funkce budeme značit *OWF* z anglického One Way Functions. Ve skutečnosti si ale ukážeme slabší výsledek a to, že *PRG* lze vytvořit z jednosměrných permutací (*OWP* – One Way Permutations). Poznamenejme však, že Johan Håstad, Russell Impagliazzo, Leonid A. Levin, Michael Luby (A pseudorandom generator from any one way function) dokázali existenci *PRG* i za slabšího předpokladu jednosměrných funkcí.

Hardcore bit

Připomeňme, že z definice *OWF*, je sice těžké takové funkce invertovat. To nám ale nedává, že neumíme z jejího výstupu získat nějaký konkrétní bit vstupu. Například pro jednosměrnou funkci f vytvoříme funkci f' takovou, že f' dostává na vstupu jeden bit navíc a ten pouze připojíme na výstup. Nová funkce f' je stále jednosměrná, ale poslední bit vstupu jsme schopni určit.

Proto používáme následující definici hardcore bitu zachycující obtížnost inverze jednosměrné funkce. Zhruba řečeno ta část vstupu, kterou jednosměrná funkce ochrání. Například si můžeme představit RSA a least significant bit.

Definice: Říkáme, že $b\colon \{0,1\}^* \to \{0,1\}$ je *hardcore bit* pro jednosměrnou funkci f, pokud

1. b je počitatelná v polynomiálním čase
2. Pro každý PPT algoritmus A existuje negligible funkce ε taková, že

$$\Pr[A(f(x)) = b(x)] \leq \frac{1}{2} + \varepsilon(n),$$

pro každou délku vstupu n, kde pravděpodobnost je přes volbu $x \leftarrow \{0,1\}^n$ a náhodné bity algoritmu A.

Nyní si ukážeme, jak zkonstruovat hardcore bit pro libovolnou jednosměrnou funkci. Tuto konstrukci vymysleli Odded Goldreich a Leonid Levin (nalézt ji můžete třeba v knize Computational Complexity: A modern approach – Sanjeev Arora a Boaz Barak).

Tvrzení: Nechť f je *OWF*, definujme $f'(x,r) = (f(x), r)$ pro $|r| = |x|$. Potom

$$\langle x, r\rangle = \sum_{i\in\{1,\ldots,n\}} x_i r_i \mod 2 = \oplus_{i\in\{1,\ldots,n\}\ r_i=1}\ x_i$$

je hardcore bit pro f'.

Důkaz: Mějme PPT A, který předvídá $\langle x, r\rangle$ z $(f(x), r)$ s nonnegligible výhodou $\frac{1}{p(n)}$. Chceme zkonstruovat PPT B, který za použití A invertuje funkci f na náhodném $x \leftarrow \{0,1\}^n$. Dělíme na následující případy:

- *Jednoduchý případ*

  Předpokládejme, že A předvídá $\langle x, r\rangle$ vždy (s pravděpodobností 1). Potom B na vstupu y:

  1. Pro všechna $i \in \{1, \ldots, n\}$ spočte $w_i = A\,(y, e^i)$, kde e^i je vektor jehož i-tá souřadnice je 1 a zbytek 0.
  2. Vrátí $w_1 w_2 \ldots w_n$.

  Jelikož A předvídá $\langle x, r\rangle$ vždy, každé w_i je skutečně i-tý bit vstupu a proto B vždy invertuje vstup. Obecně je ale adversary A horší.

- *Středně těžký případ*

  Nechť A je úspěšný s pravděpodobností alespoň $\frac{3}{4} + \frac{1}{p(n)}$. Všimněme si, že

$$\begin{aligned}\langle x, r\rangle \oplus \left\langle x, r \oplus e^i\right\rangle &= \sum_{j\in\{1,\ldots,n\}} x_j r_j \mod 2 \ \oplus \sum_{j\in\{1,\ldots,n\}} x_j(r_j \oplus e^i_j) \mod 2\\ &= \sum_{j\in\{1,\ldots,n\}} x_j(r_j \oplus r_j \oplus e^i_j) \mod 2\\ &= \sum_{j\in\{1,\ldots,n\}} x_j e^i_j \mod 2\\ &= \left\langle x, e^i\right\rangle.\end{aligned}$$

  Navíc je-li $r \leftarrow \{0,1\}^n$ rovnoměrně náhodně rozdělené, potom i $r \oplus e^i$ je rovnoměrně náhodně rozdělené.

  Nový adversary B na vstupu y a s volbou $t = \Theta(\frac{n}{(\frac{1}{p(n)})^2}) = \Theta(n \cdot p(n)^2)$:

  1. Pro každé i, $1 \le i \le n$: volíme náhodně nezávisle $r^i_1, r^i_2, \ldots, r^i_t \leftarrow \{0,1\}^n$
  2. Spočti $w_i = MAJ\left(\left\{(A(y, r^i_j) \oplus A(y, r^i_j \oplus e^i)\right\}_{j\in\{1,\ldots,t\}}\right)$, kde MAJ (z anglického majority) vrací hodnotu, která se vyskytuje nejčastěji.
  3. Vrať $w_1 w_2 \ldots w_n$

  Tvrzení: B invertuje f s nonnegligible pravděpodobností

  Nejprve připomeňme následující dvě tvrzení z pravděpodobnosti, které se nám budou hodit v důkazu.

  Tvrzení: *(Union Bound)* Nechť $\{A_i\}_{i\in\{1,\ldots,n\}}$ je konečně mnoho jevů z nichž každý nastane s pravděpodobností $\Pr[A_i]$. Potom $\Pr\left[\bigcup_{i\in\{1,\ldots,n\}} A_i\right] \le \sum_{i\in\{1,\ldots,n\}} \Pr[A_i]$.

  Tvrzení: *(Chernoff bound)* Pro $\varepsilon > 0$ a $b \in \{0,1\}$ nechť $\{X_i\}_{i\in\{1,\ldots,t\}}$ jsou nezávislé binární náhodné veličiny takové, že pro všechna i: $\Pr[X_i = b] = \frac{1}{2} + \varepsilon$. Potom

$$\Pr\left[MAJ(\{X_i\}_{i\in\{1,\ldots,t\}}) \neq b\right] \le \exp\left(-\frac{\varepsilon^2 t}{2}\right).$$

  Definujeme množinu $GOOD$: $|GOOD| \ge \frac{1}{2}\frac{1}{p(n)}2^n$ taková, že

$$\forall x \in GOOD\colon \Pr[A(f(x), r) = \langle x, r\rangle] \ge \frac{3}{4} + \frac{1}{2}\frac{1}{p(n)},$$

  kde pravděpodobnost je přes $x \leftarrow \{0,1\}^n$ a náhodné bity algoritmu A. Všimněme si, že právě definovaná množina $GOOD$ vždy existuje. Jinak

$$\begin{aligned}\Pr[A(f(x), r) = \langle x, r\rangle] &= \Pr[A(f(x), r) = \langle x, r\rangle \mid x \notin GOOD] \cdot \Pr[x \notin GOOD] +\\ &+ \Pr[A(f(x), r) = \langle x, r\rangle \mid x \in GOOD] \cdot \Pr[x \in GOOD]\\ &\le \Pr[A(f(x), r) = \langle x, r\rangle \mid x \notin GOOD] + \Pr[x \in GOOD]\\ &< \left(\frac{3}{4} + \frac{1}{2}\frac{1}{p(n)}\right) + \frac{1}{2}\frac{1}{p(n)} = \frac{3}{4} + \frac{1}{p(n)}\end{aligned}$$

  a tudíž není splněn předpoklad, že A je úspěšný s pravděpodobností alespoň $\frac{3}{4} + \frac{1}{p(n)}$.

  Stačí uvažovat x, která jsou z množiny $GOOD$ a ukázat, že na množině $GOOD$ umíme invertovat s nonnegligible pravděpodobností. Potom umíme i na množině 2^n invertovat s nonnegligible

pravděpodobností, protože množina $GOOD$ obsahuje $\frac{1}{poly(n)}$ zlomek všech možných vstupů. Rozeberme si tedy s jakou pravděpodobností B invertuje na vstupu $y = f(x)$ pro $x \in GOOD$. Potom pro všechna $x \in GOOD$ a pro všechna $i \in \{1, \ldots, n\}$

$$\Pr[A(f(x), r) \oplus A(f(x), r \oplus e^i) \neq x_i] \leq \frac{1}{2} - \frac{1}{p(n)},$$

kde pravděpodobnost je přes náhodné bity A a nerovnost platí za použití *Union bound*. Dále z *Chernoffovy* nerovnosti dostáváme

$$\Pr[MAJ(\{A(f(x), r_j^i) \oplus A(f(x), r_j^i \oplus e^i)\}_{j \in \{1,\ldots,t\}}) \neq x_i] \leq \exp(-\frac{n}{2}),$$

kde pravděpodobnost je zase přes náhodné bity algoritmu A a x je libovolné z množiny $GOOD$. A konečně použijeme *Union bound* a získáme, že B invertuje pro x z množiny $GOOD$ s pravděpodobností

$$\Pr[B \text{ určil všechny bity správně}] \geq 1 - n \cdot \exp(-\frac{n}{2}) \geq \frac{1}{2}$$

- *Nejtěžší případ*

  Pro adversary A pracující s nonnegligible pravděpodobností. Na přednášce se neprobral, lze jej dohledat třeba v již dříve zmiňované knize Arora, Barak: Computational Complexity: A Modern Approach.

Jak jsme již zmiňovali dříve, nedokážeme existenci PRG za předpokladu existence libovolné OWF, ale dokážeme ji za předpokladu existence OWP (jednosměrné permutace). Poznamenejme, že tento předpoklad je silnější díky známé blackbox separaci mezi OWP a OWF (viz A Dual Version of Reimer's Inequality and a Proof of Rudich's Conjecture – Jeff Kahn, Michael Saks, Cliff Smyth).
Definice: *Jednosměrná permutace OWP* je jednosměrná funkce, která je zároveň bijekce.

Nyní ukážeme, že za předpokladu jednosměrných permutací, umíme vytvořit pseudonáhodný generátor s expanzí 1. Na příští přednášce toto tvrzení zesílíme. Za použití techniky hybridního argumentu (v kryptografii velmi rozšířená technika) dokážeme, že z libovolného PRG s expanzí 1 lze vytvořit PRG s neomezenou (čti libovolně, avšak polynomiálně velkou) expanzí.
Tvrzení: Nechť f je jednosměrná permutace a b je její hardcore bit, potom

$$G(s) = f(s)||b(s),$$

kde $||$ značí konkatenaci, je pseudonáhodný generátor s expanzí 1.
Důkaz: Chceme ukázat, že pro libovolného distinguishera D rozlišujícího výstup pseudonáhodného generátoru G od náhodných bitů platí:

$$\Pr[D(U_{n+1}) = 1] - \Pr[D(G(U_n)) = 1] \leq \varepsilon(n),$$

kde $\varepsilon(n)$ je negligible funkce. Nejprve si uvědomme, že první pravděpodobnost můžeme přepsat do následující podoby:

$$\begin{aligned}\Pr[D(U_{n+1}) = 1] = \Pr[D(U_n||U_1) = 1] = \Pr[D(f(U_n)||U_1) = 1] = \\ = \frac{1}{2}\Pr[D(f(U_n)||b(U_n)) = 1] + \frac{1}{2}\Pr[D(f(U_n)||\overline{b(U_n)}) = 1],\end{aligned}$$

pro druhou rovnost využíváme faktu, že f je jednosměrná permutace. Nyní nahradíme $\Pr[D(U_{n+1}) = 1]$ za právě získaný výraz $\frac{1}{2}\Pr[D(f(U_n)||b(U_n)) = 1] + \frac{1}{2}\Pr[D(f(U_n)||\overline{b(U_n)}) = 1]$ a dostáváme, že chceme dokázat:

$$\frac{1}{2}\Pr[D(f(U_n)||\overline{b(U_n)}) = 1] - \frac{1}{2}\Pr[D(f(U_n)||b(U_n)) = 1] \leq \varepsilon(n)$$

Vytvořme adversary $\mathcal{A}$, který dostane $y \leftarrow f(U_n)$ a předpovídá $b(U_n)$ a to tak, že volí $r' \leftarrow \{0,1\}$ náhodně a pokud $D(y||r') = 0$ vystoupí r' jinak vystoupí $\overline{r'}$. Potom

$$\begin{aligned}
\Pr[\mathcal{A}(f(U_n)) = b(U_n)] &= \frac{1}{2}\Pr[\mathcal{A}(f(U_n)) = b(U_n) \mid r' = b(U_n)] + \frac{1}{2}\Pr[\mathcal{A}(f(U_n)) = b(U_n) \mid r' \neq b(U_n)] = \\
&= \frac{1}{2}\Big(\Pr[D(f(U_n||b(U_n))) = 0] + \Pr[D(f(U_n))||\overline{b(U_n)} = 1]\Big) = \\
&= \frac{1}{2}\Big(1 - \Pr[D(f(U_n||b(U_n))) = 1] + \Pr[D(f(U_n))||\overline{b(U_n)} = 1]\Big) = \\
&= \frac{1}{2} - \frac{1}{2}\Pr[D(f(U_n||b(U_n))) = 1] + \frac{1}{2}\Pr[D(f(U_n))||\overline{b(U_n)} = 1] \leq \frac{1}{2} + \varepsilon(n)
\end{aligned}$$

Poslední nerovnost plyne z toho, že f je jednosměrná funkce a tudíž $\mathcal{A}$ na vstupu $f(U_n)$ předpovídá $b(U_n)$ s negligible pravděpodobností. Dostáváme, že

$$\frac{1}{2}\Pr[D(f(U_n))||\overline{b(U_n)} = 1] - \frac{1}{2}\Pr[D(f(U_n||b(U_n))) = 1] \leq \varepsilon(n)$$