

Průměrné šíření difference

K odvození věty o průměrném šíření difference budeme potřebovat pojem autokorelace binární booleovské funkce.

Definice. *Cross-korelace* binárních booleovských funkcí $f, g \in R_n$ je zobrazení

$$\begin{aligned}\widehat{c}_{f,g} : \mathbb{F}_2^n &\rightarrow \mathbb{R} \\ u &\mapsto C(f, g \circ \tau_u)\end{aligned}$$

kde $\tau_u : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$ je posunutí o vektor u .

Pro cross-korelaci platí:

$$\begin{aligned}\widehat{c}_{f,g}(u) &= \frac{(\widehat{f}, \widehat{g \circ \tau_u})}{\|\widehat{f}\| \cdot \|\widehat{g \circ \tau_u}\|} \\ &= \frac{1}{2^n} \cdot \sum_{v \in \mathbb{F}_2^n} \widehat{f}(v) \cdot \widehat{g \circ \tau_u}(v) \\ &= \frac{1}{2^n} \cdot \sum_{v \in \mathbb{F}_2^n} \widehat{f}(v) \cdot \widehat{g}(u + v) \\ &= \frac{1}{2^n} \cdot (\widehat{f} \otimes \widehat{g})(u)\end{aligned}$$

Získali jsme rovnost

$$\widehat{c}_{f,g} = \frac{1}{2^n} \cdot \widehat{f} \otimes \widehat{g}$$

Z toho například vidíme, že $\widehat{c}_{f,g} = \widehat{c}_{g,f}$.

Lemma. Pro $f, g \in R_n$ platí

$$W(\widehat{c}_{f,g}) = W(\widehat{f}) \cdot W(\widehat{g})$$

Důkaz. Pravá strana rovnosti je bodový součin spekter f a g . Pro $u \in \mathbb{F}_2^n$:

$$\begin{aligned} (W(\widehat{f}) \cdot W(\widehat{g}))(u) &= W(\widehat{f})(u) \cdot W(\widehat{g})(u) \\ &= \frac{(\widehat{f}, \widehat{p}_u)}{2^n} \cdot \frac{(\widehat{g}, \widehat{p}_u)}{2^n} \\ &= \frac{1}{2^n} \cdot \frac{1}{2^n} \cdot \left(\sum_{v \in \mathbb{F}_2^n} \widehat{f}(v) \cdot \widehat{p}_u(v) \right) \cdot \left(\sum_{w \in \mathbb{F}_2^n} \widehat{g}(w) \cdot \widehat{p}_u(w) \right) \\ &= \frac{1}{2^n} \cdot \frac{1}{2^n} \cdot \sum_{v \in \mathbb{F}_2^n} \sum_{w \in \mathbb{F}_2^n} \widehat{f}(v) \cdot \widehat{g}(w) \cdot \widehat{p}_u(v + w)\end{aligned}$$

$$\begin{aligned}
& \stackrel{(z=v+w)}{=} \frac{1}{2^n} \cdot \frac{1}{2^n} \cdot \sum_{v \in \mathbb{F}_2^n} \sum_{z \in \mathbb{F}_2^n} \hat{f}(v) \cdot \hat{g}(z+v) \cdot \hat{p}_u(z) \\
&= \frac{1}{2^n} \cdot \sum_{z \in \mathbb{F}_2^n} \left(\frac{1}{2^n} \cdot \sum_{v \in \mathbb{F}_2^n} \hat{f}(v) \cdot \hat{g}(z+v) \right) \cdot \hat{p}_u(z) \\
&= \frac{1}{2^n} \cdot \sum_{z \in \mathbb{F}_2^n} \left(\frac{1}{2^n} \cdot (\hat{f} \otimes \hat{g})(z) \right) \cdot \hat{p}_u(z) \\
&= \frac{1}{2^n} \cdot \sum_{z \in \mathbb{F}_2^n} \hat{c}_{f,g}(z) \cdot \hat{p}_u(z) \\
&= \frac{1}{2^n} \cdot (\hat{c}_{f,g}, \hat{p}_u) \\
&= W(\hat{c}_{f,g})(u)
\end{aligned}$$

□

Definice. Zobrazení $\hat{r}_f = \hat{c}_{f,f}$ se nazývá *autokorelace* binární booleovské funkce f .

Věta (*Wiener-Khintchine*). $W(\hat{r}_f) = W^2(\hat{f})$.

Nyní můžeme popsat vztah matice šíření difference dané booleovské funkce a matice korelačních potenciálů. Tuto závislost pak využijeme k odvození věty o průměrném šíření difference pro klíč alternující šifru.

Definice. Pro booleovskou funkci $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ definujeme *matici korelačních potenciálů* CP^f jako matici typu $2^m \times 2^n$ se vstupy $CP_{u,v}^f = C^2(p_u \circ f, p_v)$, kde $v \in \mathbb{F}_2^n$, $u \in \mathbb{F}_2^m$. Hodnotu $CP_{u,v}^f$ jsme již dříve definovali jako *korelační potenciál* binární booleovské funkce $p_u \circ f$ ve v .

Následující věta říká, že pro $n = m$ jsou matice šíření difference a matice korelačních potenciálů podobné.

Věta. Pro booleovskou funkci $f : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^m$ platí

$$DP^f = \mathcal{T}_m^{-1} \cdot CP^f \cdot \mathcal{T}_n$$

kde $\mathcal{T}_k$ je matice přechodu od ortogonální báze znaménkových funkcí parit ke kanonické bázi v prostoru $\mathbb{R}^{2^k}$.

Důkaz. Označíme-li B_k ortogonální bázi znaménkových funkcí parit v $\mathbb{R}^{2^k}$, pak musí platit $\mathcal{T}_k \cdot \{w\}_{B_k} = w$ pro každý vektor $w \in \mathbb{R}^{2^k}$. Tedy sloupce matice $\mathcal{T}_k$ musí být vektory báze B_k , tj. ve sloupci $v \in \mathbb{F}_2^k$ je vektor $\hat{p}_v$. Z ortogonality B_k pak plyne: $\mathcal{T}_k^\top \cdot \mathcal{T}_k$ je diagonální, hodnoty na diagonále jsou rovny 2^k , tedy $\mathcal{T}_k^{-1} = 2^{-k} \cdot \mathcal{T}_k^\top$.

Pokud platí znění věty, pak vstupy matice šíření difference musí být v následujícím tvaru:

$$\begin{aligned}
DP_{u,v}^f &= (2^{-m} \cdot \mathcal{T}_m^\top \cdot CP^f \cdot \mathcal{T}_n)_{u,v} \\
&= 2^{-m} \sum_{a \in \mathbb{F}_2^m} (\mathcal{T}_m^\top)_{u,a} \cdot (CP^f \cdot \mathcal{T}_n)_{a,v} \\
&= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \hat{p}_u(a) \cdot \sum_{b \in \mathbb{F}_2^n} CP_{a,b}^f \cdot (\mathcal{T}_n)_{b,v} \\
&= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \sum_{b \in \mathbb{F}_2^n} \hat{p}_u(a) \cdot \hat{p}_v(b) \cdot CP_{a,b}^f
\end{aligned}$$

Pokusme se tento tvar $DP_{u,v}^f$ odvodit pomocí vlastností autokorelace. Pokud označíme $d_{w,v} = f(w) + f(w+v)$, můžeme šíření difference zapsat ve tvaru:

$$DP_{u,v}^f = 2^{-n} \sum_{w \in \mathbb{F}_2^n} 2^{-m} \cdot (\hat{p}_{d_{w,v}}, \hat{p}_u)$$

Totíž skalární součin $(\hat{p}_{d_{w,v}}, \hat{p}_u)$ je nenulový pouze v případě $d_{w,v} = u$. V tomto případě se jeho hodnota rovná 2^m a předchozí suma tedy udává počet vektorů $w \in \mathbb{F}_2^n$, pro které platí $f(w) + f(w+v) = u$. Rozepsáním skalárního součinu dostaneme:

$$\begin{aligned} DP_{u,v}^f &= 2^{-n} 2^{-m} \sum_{w \in \mathbb{F}_2^n} \sum_{a \in \mathbb{F}_2^m} \hat{p}_{d_{w,v}}(a) \cdot \hat{p}_u(a) \\ &= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \hat{p}_u(a) \cdot (2^{-n} \sum_{w \in \mathbb{F}_2^n} \hat{p}_{d_{w,v}}(a)) \end{aligned}$$

Úpravou části zápisu získáme:

$$\begin{aligned} 2^{-n} \sum_{w \in \mathbb{F}_2^n} \hat{p}_{d_{w,v}}(a) &= 2^{-n} \sum_{w \in \mathbb{F}_2^n} \hat{p}_{f(w)}(a) \cdot \hat{p}_{f(w+v)}(a) \\ &= 2^{-n} \sum_{w \in \mathbb{F}_2^n} \hat{p}_a(f(w)) \cdot \hat{p}_a(f \circ \tau_v(w)) \\ &= 2^{-n} \sum_{w \in \mathbb{F}_2^n} \widehat{p_a \circ f}(w) \cdot \widehat{p_a \circ f \circ \tau_v}(w) \\ &= \frac{(\widehat{p_a \circ f}, \widehat{p_a \circ f \circ \tau_v})}{\|\widehat{p_a \circ f}\| \cdot \|\widehat{p_a \circ f \circ \tau_v}\|} \\ &= C(p_a \circ f, p_a \circ f \circ \tau_v) \\ &= \widehat{r}_{p_a \circ f}(v) \end{aligned}$$

kde τ_v je posunutí (translace) o vektor v . Výsledek odvození je hodnota autokorelace $\widehat{r}_{p_a \circ f}$ ve vektoru v . Pro šíření difference tedy dostáváme:

$$DP_{u,v}^f = 2^{-m} \sum_{a \in \mathbb{F}_2^m} \hat{p}_u(a) \cdot \widehat{r}_{p_a \circ f}(v)$$

Zapišeme autokorelaci v ortogonální bázi znaménkových funkcí parit: $\sum_{b \in \mathbb{F}_2^n} W(\widehat{r}_{p_a \circ f})(b) \cdot \hat{p}_b$. Podle Wiener-Khintchinovy věty pro Walsh-Hadamardovu transformaci autokorelace platí $W(\widehat{r}_{p_a \circ f}) = W^2(\widehat{p_a \circ f})$, tedy:

$$\begin{aligned} DP_{u,v}^f &= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \hat{p}_u(a) \cdot (\sum_{b \in \mathbb{F}_2^n} W^2(\widehat{p_a \circ f})(b) \cdot \hat{p}_b)(v) \\ &= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \hat{p}_u(a) \cdot \sum_{b \in \mathbb{F}_2^n} W^2(\widehat{p_a \circ f})(b) \cdot \hat{p}_b(v) \\ &= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \sum_{b \in \mathbb{F}_2^n} \hat{p}_u(a) \cdot \hat{p}_b(v) \cdot C^2(p_a \circ f, p_b) \\ &= 2^{-m} \sum_{a \in \mathbb{F}_2^m} \sum_{b \in \mathbb{F}_2^n} \hat{p}_u(a) \cdot \hat{p}_v(b) \cdot CP_{a,b}^f \end{aligned}$$

□

Ted' máme vše potřebné k důkazu věty o průměrném šíření difference. Před jejím vyslovením připomeneme již dokázané tvrzení o průměrném korelačním potenciálu. Ve znění těchto tvrzení chápeme klíč alternující šifru $f = f^{(N)} \circ \dots \circ f^{(1)}$ jako produkt kolových kryptosystémů, zatímco po dosazení konkrétní $(N + 1)$ -tice kolových klíčů dostáváme booleovské funkce $f, f^{(1)}, \dots, f^{(N)} : \mathbb{F}_2^n \rightarrow \mathbb{F}_2^n$.

Věta (*průměrný korelační potenciál*). Ať $f = f^{(N)} \circ \dots \circ f^{(1)}$ je klíč alternující šifra. Pak pro vektory $u, v \in \mathbb{F}_2^n$ platí

$$E(CP_{u,v}^f) = \sum_{\substack{U \text{ lin. cesta nad } f \\ \text{od } v \text{ k } u}} C_p^2(U)$$

kde $E(CP_{u,v}^f)$ je průměrný korelační potenciál, tedy aritmetický průměr hodnot $CP_{u,v}^f$ na množině všech $(N + 1)$ -tic kolových klíčů. □

Následující lemma je pouze přeformulací definice násobení (čtvercových) matic.

Lemma. Ať jsou $M = M^{(N)} \cdot \dots \cdot M^{(1)}$ matice řádu k nad libovolným okruhem. Pak pro $i, j \in \{0, \dots, k - 1\}$ platí

$$M_{i,j} = \sum_{\substack{U \in \{0, \dots, k-1\}^{N+1} \\ U = (u^{(0)}, \dots, u^{(N)}) \\ u^{(0)} = j, u^{(N)} = i}} M_{u^{(N)}, u^{(N-1)}}^{(N)} \cdot \dots \cdot M_{u^{(1)}, u^{(0)}}^{(1)}$$

Důkaz. Pro $N = 1$ lemma platí. Předpokládejme, že platí pro $N - 1$ a odvoďme tvrzení pro součin délky N :

$$\begin{aligned} M_{i,j} &= (M^{(N)} \cdot \dots \cdot M^{(1)})_{i,j} \\ &= \sum_{l=0}^{k-1} M_{i,l}^{(N)} \cdot (M^{(N-1)} \cdot \dots \cdot M^{(1)})_{l,j} \\ &= \sum_{l=0}^{k-1} M_{i,l}^{(N)} \cdot \sum_{\substack{U \in \{0, \dots, k-1\}^N \\ U = (u^{(0)}, \dots, u^{(N-1)}) \\ u^{(0)} = j, u^{(N-1)} = l}} M_{u^{(N-1)}, u^{(N-2)}}^{(N-1)} \cdot \dots \cdot M_{u^{(1)}, u^{(0)}}^{(1)} \\ &= \sum_{l=0}^{k-1} \sum_{\substack{U \in \{0, \dots, k-1\}^N \\ U = (u^{(0)}, \dots, u^{(N-1)}) \\ u^{(0)} = j, u^{(N-1)} = l}} M_{i,l}^{(N)} \cdot M_{u^{(N-1)}, u^{(N-2)}}^{(N-1)} \cdot \dots \cdot M_{u^{(1)}, u^{(0)}}^{(1)} \\ &= \sum_{\substack{U \in \{0, \dots, k-1\}^{N+1} \\ U = (u^{(0)}, \dots, u^{(N)}) \\ u^{(0)} = j, u^{(N)} = i}} M_{u^{(N)}, u^{(N-1)}}^{(N)} \cdot M_{u^{(N-1)}, u^{(N-2)}}^{(N-1)} \cdot \dots \cdot M_{u^{(1)}, u^{(0)}}^{(1)} \end{aligned}$$

□

Věta (*průměrné šíření difference*). Ať $f = f^{(N)} \circ \dots \circ f^{(1)}$ je klíč alternující šifra. Pak pro vektory $u, v \in \mathbb{F}_2^n$ platí

$$E(DP_{u,v}^f) = \sum_{\substack{U \text{ dif. cesta nad } f \\ \text{od } v \text{ k } u}} 2^{-w(U)}$$

kde $E(DP_{u,v}^f)$ je průměrné šíření difference, tedy aritmetický průměr hodnot $DP_{u,v}^f$ na množině všech $(N+1)$ -tic kolových klíčů.

Důkaz. Označme M matici typu $2^n \times 2^n$ se vstupy $M_{u,v}$, $u, v \in \mathbb{F}_2^n$:

$$M_{u,v} = \sum_{\substack{U \text{ dif. cesta nad } f \\ \text{od } v \text{ k } u}} 2^{-w(U)} = \sum_{\substack{U \text{ dif. cesta nad } f \\ U=(u^{(0)}, \dots, u^{(N)}) \\ u^{(0)}=v, u^{(N)}=u}} DP_{u^{(N)}, u^{(N-1)}}^{f^{(N)}} \cdot \dots \cdot DP_{u^{(1)}, u^{(0)}}^{f^{(1)}}$$

Podle předchozího lemmatu platí $M = DP^{f^{(N)}} \cdot \dots \cdot DP^{f^{(1)}}$. Pro aplikaci lemmatu některé cesty v součtu chybí - v definici diferenční cesty požadujeme kompatibilitu sousedních diferencí. Příslušné sčítance jsou ale pro vynechané cesty nulové.

Budeme-li $E(DP^f)$ chápat jako matici typu $2^n \times 2^n$ se vstupy $E(DP^f)_{u,v} = E(DP_{u,v}^f)$, pak nám zbývá ukázat:

$$E(DP^f) = DP^{f^{(N)}} \cdot \dots \cdot DP^{f^{(1)}}$$

Označme $f_{K_1, \dots, K_{N+1}}$ booleovskou funkci získanou dosazením kolových klíčů do šifry f . Můžeme psát:

$$E(DP^f) = \frac{1}{2^{(N+1) \cdot n}} \cdot \sum_{K_1, \dots, K_{N+1} \in \mathbb{F}_2^n} DP^{f_{K_1, \dots, K_{N+1}}}$$

Podobně pro průměrný korelační potenciál:

$$E(CP^f) = \frac{1}{2^{(N+1) \cdot n}} \cdot \sum_{K_1, \dots, K_{N+1} \in \mathbb{F}_2^n} CP^{f_{K_1, \dots, K_{N+1}}}$$

Víme, že pro průměrný korelační potenciál platí

$$E(CP_{u,v}^f) = \sum_{\substack{U \text{ lin. cesta nad } f \\ \text{od } v \text{ k } u}} C_p^2(U) = \sum_{\substack{U \text{ lin. cesta nad } f \\ U=(u^{(0)}, \dots, u^{(N)}) \\ u^{(0)}=v, u^{(N)}=u}} CP_{u^{(N)}, u^{(N-1)}}^{f^{(N)}} \cdot \dots \cdot CP_{u^{(1)}, u^{(0)}}^{f^{(1)}}$$

Z předchozího lemmatu musí být $E(CP^f) = CP^{f^{(N)}} \cdot \dots \cdot CP^{f^{(1)}}$. Konjugací obou stran této rovnosti maticí přechodu $\mathcal{T}_n$ od ortogonální báze znaménkových funkcí parit ke kanonické bázi postupně dostaneme:

$$\begin{aligned} \mathcal{T}_n^{-1} \cdot E(CP^f) \cdot \mathcal{T}_n &= \mathcal{T}_n^{-1} \cdot \left(\frac{1}{2^{(N+1) \cdot n}} \cdot \sum_{K_1, \dots, K_{N+1} \in \mathbb{F}_2^n} CP^{f_{K_1, \dots, K_{N+1}}} \right) \cdot \mathcal{T}_n \\ &= \frac{1}{2^{(N+1) \cdot n}} \cdot \sum_{K_1, \dots, K_{N+1} \in \mathbb{F}_2^n} \mathcal{T}_n^{-1} \cdot CP^{f_{K_1, \dots, K_{N+1}}} \cdot \mathcal{T}_n \\ &= \frac{1}{2^{(N+1) \cdot n}} \cdot \sum_{K_1, \dots, K_{N+1} \in \mathbb{F}_2^n} DP^{f_{K_1, \dots, K_{N+1}}} = E(DP^f) \end{aligned}$$

A také:

$$\begin{aligned}\mathcal{T}_n^{-1} \cdot \text{CP}^{f^{(N)}} \cdot \dots \cdot \text{CP}^{f^{(1)}} \cdot \mathcal{T}_n &= (\mathcal{T}_n^{-1} \text{CP}^{f^{(N)}} \mathcal{T}_n) \cdot (\mathcal{T}_n^{-1} \text{CP}^{f^{(N-1)}} \mathcal{T}_n) \cdot \dots \cdot (\mathcal{T}_n^{-1} \text{CP}^{f^{(1)}} \mathcal{T}_n) \\ &= \text{DP}^{f^{(N)}} \cdot \text{DP}^{f^{(N-1)}} \cdot \dots \cdot \text{DP}^{f^{(1)}}\end{aligned}$$

□

Poznámka. V důkazu jsme počítali s maticemi $\text{DP}^{f^{(i)}}$ a $\text{CP}^{f^{(i)}}$ pro kolové funkce $f^{(i)}$, i když booleovské funkce z kolových funkcí získáme až dosazením konkrétních klíčů. Postup je v pořádku, protože hodnoty šíření difference i korelačního potenciálu na kolech klíč alternující šifry nezávisí na klíčích.

Poznámka. Jelikož suma ze znění předchozí věty je součet kladných čísel, pro každou cestu U od v k u platí $E(\text{DP}_{u,v}^f) \geq 2^{-w(U)}$. Podle předchozí věty tedy můžeme říct, že pokud nad klíč alternující šifrou nalezneme diferenční cestu od v k u s malou vahou, pak pro průměrnou (náhodnou) volbu kolových klíčů bude hodnota šíření difference $\text{DP}_{u,v}^f$ významná. To lze chápat jako zdůvodnění úspěšnosti metody diferenční kryptoanalýzy a oprávněnosti požadavku na odolnost současných symetrických algoritmů vůči konstrukci dobrých diferenčních cest.