MI-PB-4

Lineární kryptoanalýza, lineární aproximace S-boxů, lineární aproximační funkce, extrakce bitů klíče.

Slabý klíč: klíč, jehož zvláštní matematické vlastnosti umožňují snadné prolomení šifry

Poloslabý klíč: $k_1 \neq k_2$ poloslabé $\Leftrightarrow E_{k_1}(E_{k_2}(OT)) = OT$

Kryptoanalýza: věda o zkoumání a prolamování šifer bez znalosti klíče

Lineární kryptoanalýza: hledá lineární souvislosti (aproximace) k jednotlivým šifrám

Vlastnosti LK:

- Použití pro kryptoanalýzu **blokových šifer**
- Využívá vysokou pravděpodobnost výskytu **lineárních vyjádření zahrnujících OT, bity ŠT a bity podklíčů** pro danou rundu
- Hledá **lineárná závislosti** mezi vstupy a výstupy SBOXů
- Útok se znalostí OT a odpovídajícího ŠT, nelze je ale zvolit

Základní myšlenka:
Aproximovat operace částí šifry výrazem, který je lineární:

$$X_1 \oplus X_2 \oplus ... \oplus X_u \oplus Y_1 \oplus Y_2 \oplus ... \oplus Y_v = 0$$

kde X_i je i-tý bit vstupu $X = [X_1, X_2, ..., X_u]$ a Y_j je j-tý bit výstupu $Y = [Y_1, Y_2, ..., Y_v]$

Cíl LK:
Nalézt taková vyjádření, která jsou ve výše popsaném tvaru a mají vysokou či naopak nízkou pravděpodobnost výskytu

Odchylka lineární pravděpodobnosti

Pokud výraz platí s pravděpodobností p, **odchylka (LPB)** je $p - \frac{1}{2}$

Velikost odchylky je potom $\left|p - \frac{1}{2}\right|$

Čím **větší LPB**, tím **lépe** lze analyzovat danou šifru

Pokud $p = 1 \Rightarrow$ je nalezený lineární výraz perfektní reprezentací šifry a šífra má **katastrofické slabiny**
Pokud $p = 0 \Rightarrow$ jde o **afinní výraz** a šífra má též katastrofické slabiny

Piling-up princip:

- **Předpoklady:**
  2 náhodné proměnné X_1, X_2.
  Lineární výraz $X_1 \oplus X_2 = 0$, afinní výraz $X_1 \oplus X_2 = 1$.
  $P(X_1 = 0) = p_1$, $P(X_1 = 1) = 1 - p_1$
  $P(X_2 = 0) = p_2$, $P(X_2 = 1) = 1 - p_2$
- Pokud X_1, X_2 nezávislé, potom:
  $P(X_1 = 0, X_2 = 0) = p_1 p_2$
  $P(X_1 = 0, X_2 = 1) = p_1(1 - p_2)$
  $P(X_1 = 1, X_2 = 0) = (1 - p_1)p_2$
  $P(X_1 = 1, X_2 = 1) = (1 - p_1)(1 - p_2)$
- Potom:
  $P(X_1 \oplus X_2 = 0) = P(X_1 = X_2) = P(X_1 = 0, X_2 = 0) + P(X_1 = 1, X_2 = 1) = p_1 p_2 + (1 - p_1)(1 - p_2)$
- Pokud $p_1 = \frac{1}{2} + \epsilon_1$ a $p_2 = \frac{1}{2} + \epsilon_2$, kde ϵ_1, ϵ_2 jsou pravděpodobnostní odchylky, **platí:**
  $P(X_1 \oplus X_2 = 0) = \frac{1}{2} + 2\epsilon_1\epsilon_2$

Piling-up věta:
Pro n nezávislých náhodných binárních proměnných $X_1, ..., X_n$ platí:

$$P(X_1 \oplus X_2 \oplus ... \oplus X_n = 0) = \frac{1}{2} + 2^{n-1} \prod_{i=1}^{n} \epsilon_i$$

Konstrukce silně lineárního výazu

Nelineární součást šifry: SBOXy

Pokud **lineární vlastnosti SBOXu zjistitelné,** je možné vytvořit **lineární aproximaci mezi vstupem a výstupem**

Lineární aproximace potom možno **zřetězit** tak, že se **vyruší mezilehlé bity** (bity prostupující mezi SBOXy)

$\Rightarrow$ lineární výraz popisující chování šifry obsahuje jen **bity OT a bity poslední rundy** a má velký LPB

Bity podklíčů jednotlivých rund se přesunou na pravou stranu výrazu: v sumě mohou mít buď hodnotu 1, nebo 0 (pouze změna zneménka LPB)

Průběh LK

Aproximace SBOXů:

- **Určit aproximované části:** nelineární (SBOXy)
- **Konstrukce lineární aproximační tabulky:** obsahuje LPB všech možných dvojic vstup-výstup SBOXu
  - např. položka na pozici $[6, B]$ je počet případů ($\Rightarrow \mathrm{LPB}$), kdy $X_2 \oplus X_3 = Y_1 \oplus Y_3 \oplus Y_4$, tedy $6 = 11_{10} = B_{16}$ pro vstup do SBOXu X_1, X_2, X_3, X_4 a výstup Y_1, Y_2, Y_3, Y_4
- Podle tabulky **nalézt výrazy,** kterými budou jednotlivé **SBOXy aproximovány:** cílem je co největší LPB
- Vytvořit **lineární aproximační výraz** pro $R - 1$ rund
  - Výraz je ve tvaru $U_i \oplus U_j \oplus ... \oplus U_n \oplus P_k \oplus P_l \oplus ... \oplus P_m = 0$
    kde P jsou bity OT, U bity vstupu po posledního SBOXu
  - Tvorba výrazu postupným dosazováním aproximačních výrazů zvolených v tabulce
- Z lineárního aproximačního výrazu (zvolených aproximací SBOXů) lze pomocí Piling-up věty spočítat **teoretický LPB** aproximace (cíl -- co největší)

Extrakce bitů klíče:

- Pro každou dvojici OT a příslušný ŠT (N párů):
  - Pro každou možnou hodnotu části podklíče z ŠT **zpětnou substitucí** získat $U_i, U_j, ..., U_n$ a pro odpovídající PT **vyhodnotit lineární aproximační výraz**
    - Zpětná substituce: $SBOX^{-1}(K \oplus \check{S}T)$
    - Hledaná část podklíče: pouze ta, které se týká výstup z aktivních SBOXů poslední rundy
  - Pokud je lineární aproximační výraz pro zpětně získané $U_i, ..., U_n$ pravdivý, **inkrementovat čítač pro podklíč**
- Pro každou hodnotu podklíče **spočítat** $|\mathrm{bias}| = \frac{|\text{čítač} - \frac{N}{2}|}{N}$
- Největší $\mathrm{bias} \Rightarrow$ **kandidát na podklíč**

(Největší bias by měl být blízký teoretickému LPB spočítanému při aproximaci. Odchylka může být způsobena nedostatečným počtem dvojic OT-ŠT nebo neúplnou nezávislostí vzorků)