Cryptanalyse de LFSRs combinés

Yann Laigle-Chapuy

INRIA - Équipe SECRET/LIP6 - Équipe SPIRAL

Yann.Laigle-Chapuy@inria.fr

Frédéric Didier

ÉPFL - alg⊕lma

frederic.didier@epfl.ch

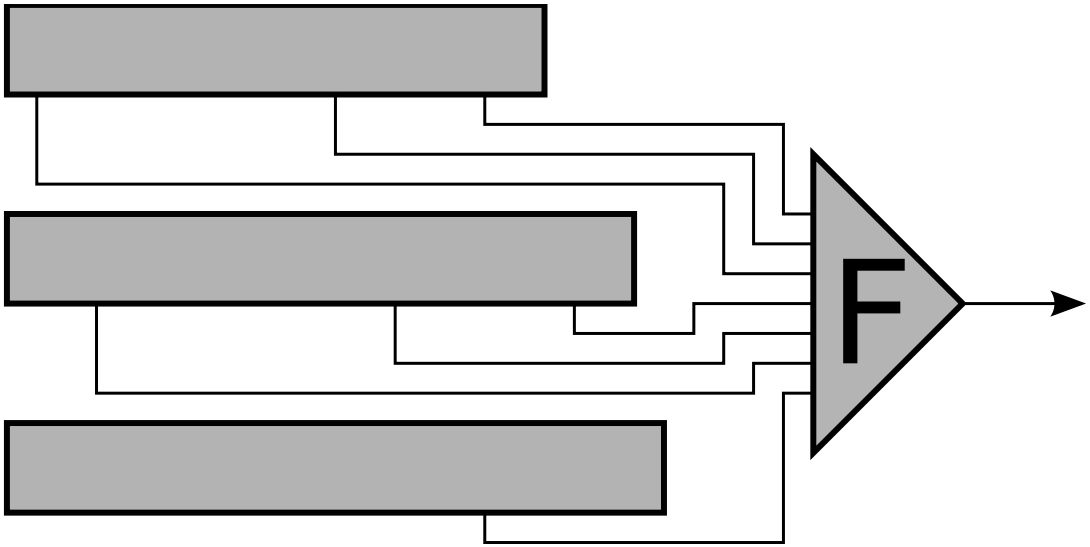
27 mars 2008

Plan de l'exposé

- Modèle d'attaque par corrélation
- Attaque sur les LFSRs combinés
  - Trouver les équations de parité
  - Rechercher l'état initial
- Résultats expérimentaux
- Conclusion

Le modèle

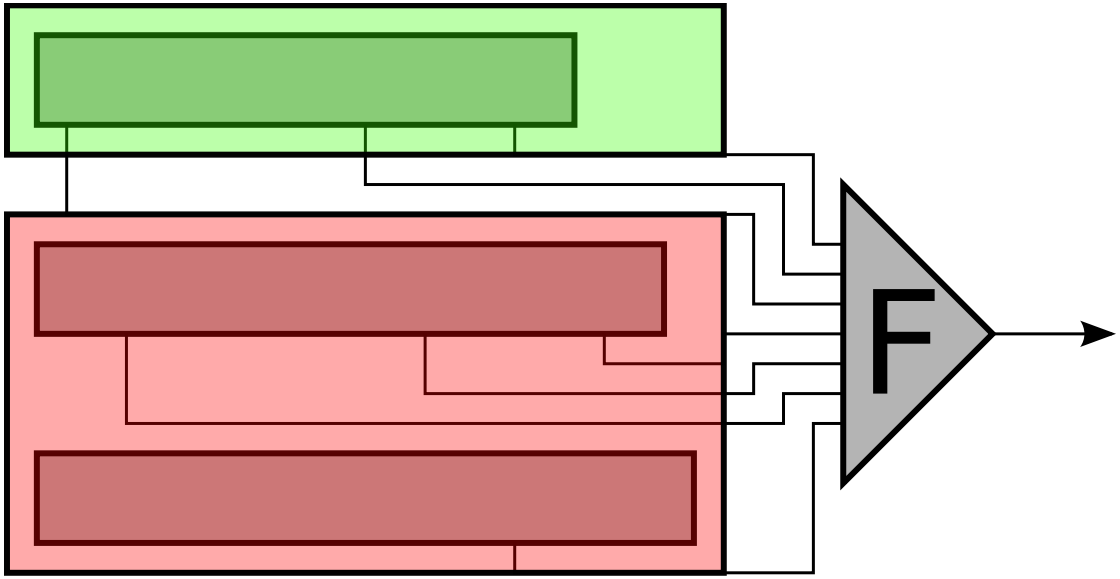
un chiffrement à flot



- registres combinés par une fonction booléenne

$$F : \{0,1\}^n \to \{0,1\}$$

Le modèle d'attaque



- registres combinés par une fonction booléenne

$$F : \{0,1\}^{v+r} \rightarrow \{0,1\}$$

But : retrouver l'initialisation du registre vert

Le modèle d'attaque

un biais

$$Pr\left(F(\vec{x}_1)+\cdots+F(\vec{x}_{2k})=0 \middle| \vec{x}_1+\cdots+\vec{x}_{2k}=0\right) \geq \frac{1}{2}\left(1+2^{-n(k-1)}\right)$$

- Provient des coefficients de Walsh de F
- Ne dépend que du nombre d'entrées de F

Comment l'utiliser ?

Le modèle d'attaque

un biais

$$Pr\left(F(\vec{x}_1)+\cdots+F(\vec{x}_{2k})=0 \middle| \vec{x}_1+\cdots+\vec{x}_{2k}=0\right) \geq \frac{1}{2}\left(1+2^{-n(k-1)}\right)$$

- Provient des coefficients de Walsh de F
- Ne dépend que du nombre d'entrées de F

Comment l'utiliser ?

- besoin d'instants tels que $\vec{x}_{t_1}+\cdots+\vec{x}_{t_{2k}}=0$

Le modèle d'attaque

un distingueur

$$Pr\Big(F(\vec{x}_1) + \cdots + F(\vec{x}_{2k}) = 0 \Big| \vec{x}_1 + \cdots + \vec{x}_{2k} = 0\Big) = P_0$$

Le modèle d'attaque

un distingueur

$$Pr\left(\vec{x}_1 + \cdots + \vec{x}_{2k} = 0 \middle| \vec{F}(x_1) + \cdots + \vec{F}(x_{2k}) = 0\right)$$

$$=$$

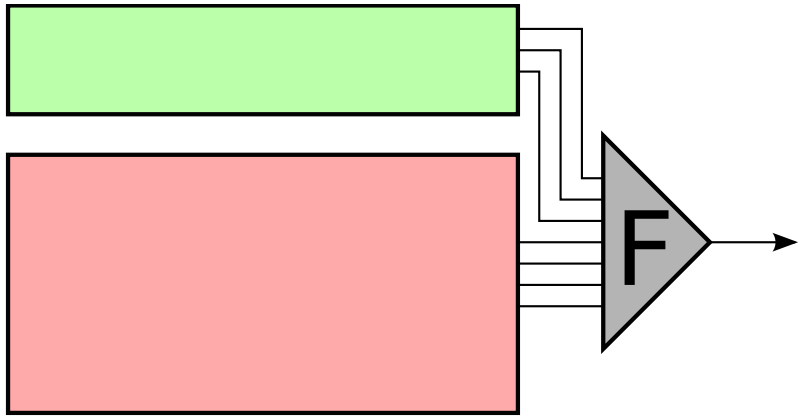
$$\frac{P(\vec{x}_1 + \cdots + \vec{x}_{2k} = 0)}{P(\vec{F}(x_1) + \cdots + \vec{F}(x_{2k}) = 0)} P_0$$

$$\geq$$

$$2^{-n} + 2^{-nk} \quad \ldots \text{ bof}$$

Le modèle d'attaque

hypothèse



- On sait trouver des équations de parités partielles, i.e. des $2k$-uplets d'instants tels que

$$\vec{x}_{t_1} + \cdots + \vec{x}_{t_{2k}} = \left(\frac{\star}{0}\right)$$

Le modèle d'attaque

un distingueur

$$Pr\left(\vec{x}_1+\cdots+\vec{x}_{2k}=0 \,\middle|\, \begin{array}{rcl}\vec{F}(x_1)+\cdots+\vec{F}(x_{2k}) & = & 0\\ \vec{x}_1+\cdots+\vec{x}_{2k} & = & \binom{\star}{0}\end{array}\right)$$

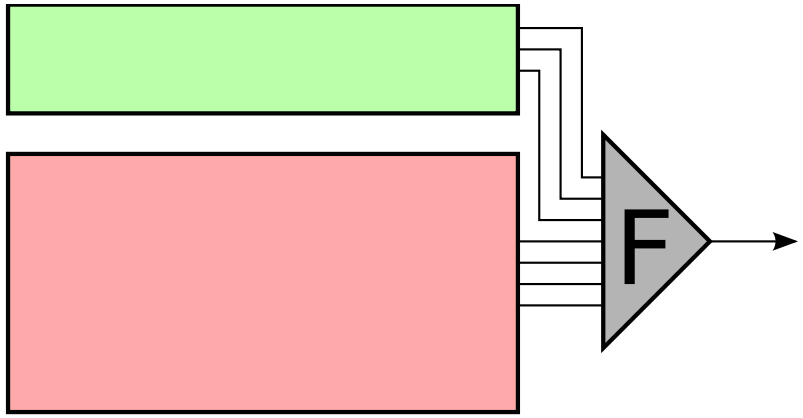
$$=$$

$$\frac{P(\vec{x}_1+\cdots+\vec{x}_{2k}=0 | \vec{x}_1+\cdots+\vec{x}_{2k}=\binom{\star}{0})}{P(\vec{F}(x_1)+\cdots+\vec{F}(x_{2k})=0)}\,P_0$$

$$\geq$$

$$2^{-v}+\varepsilon \quad \text{avec } \varepsilon = 2^{-n(k-1)-v}$$

Le modèle d'attaque



- Calculer les équations de parités
- Pour chaque initialisation
  - calculer le biais
  - s'il a la valeur souhaitée, on a trouvée la bonne initialisation

Remarque : ici, F peut être inconnue ou/et dépendre de la clé

Attaque sur les LFSRs combinés

recherche des équations de parités

$\longrightarrow$ multiples de poids faibles du polynôme de rétroaction

Attaque sur les LFSRs combinés

recherche des équations de parités

LFSR avec pour polynôme de rétroaction

$$P(x) = x^7 + x^5 + x^4 + x^3 + x^2 + x + 1$$

0 1 0 1 0 0 1 0 1 0 1 0 1 1 1 0 0 0 1 1 1 1 1 0

$Q(x)$

$Q(x) = x^{19} + x + 1$ un multiple de $P(x)$.

Attaque sur les LFSRs combinés

recherche des équations de parités

$$1 + \sum_{i=1}^{2k-1} x^{p_i} = 1 + \sum_{i=1}^{k} x^{p_i} + \sum_{i=k+1}^{2k-1} x^{p_i}$$

Méthode classique : [Chose-Joux-Mitton]

- Calculer les

$$\sum_{i=1}^{k} x^{p_i} \pmod{P}, \quad 0 \leq p_i \leq D$$

- Chercher des collisions

Temps $\mathcal{O}(D^k)$ Mémoire $\mathcal{O}(D^{\lceil k/2 \rceil})$

Attaque sur les LFSRs combinés

LFSRs combinés $\Rightarrow$ Groupe produit

$$1 + \sum_{i=1}^{2k-1} x^{p_i} = 1 + \sum_{i=1}^{k} x^{p_i} + x^{k+1} \left(\sum_{i=k+1}^{2k} x^{p_i - p_{k+1}} \right)$$

Méthode alternative : [DL]

- Calculer les $\log \left(\sum_{i=1}^{k} x^{p_i} \right) \pmod P, \quad 0 \leq p_i \leq D$
- Chercher des "presque" collisions

Temps $\mathcal{O}(D^{k-1}L)$ Mémoire $\mathcal{O}(D^{k-1})$

Attaque sur les LFSRs combinés

compter. . .

- 2^I initialisations
- $E \sim \varepsilon^{-2}$ équations

$$\rightarrow 2^I(\varepsilon^{-2} + D) \quad \text{trop...}$$

Attaque sur les LFSRs combinés

compter. . .

- 2^I initialisations
- $E \sim \varepsilon^{-2}$ équations
- le calcul de la i-ème itération d'un LFSR se fait rapidement

$$\longrightarrow 2^I \varepsilon^{-2} \quad \text{mieux...}$$

Attaque sur les LFSRs combinés

compter plus vite

$$S(i) = \sum_{j=1}^{\varepsilon^{-2}} \mathbf{1}_{(\overset{\star}{0})}(x_{t_1^j} + \cdots + x_{t_{2k}^j}), \quad i \in [0, 2^I - 1]$$

LFSRs $\longrightarrow$ on peut trouver des vecteurs E_ℓ^j tels que

$$< x_{t_1^j} + \cdots + x_{t_{2k}^j} | \vec{e}_\ell > = < i | E_\ell^j >$$

Attaque sur les LFSRs combinés

compter plus vite

$$\mathcal{F}[\mathbf{1}_{\{E_\ell^j\}}](i) = \sum_{E_\ell^j} < i|E_\ell^j >$$

mais on veut compter

$$(< i|E_1^j >= 0)\,\&\,\ldots\,\&\,(< i|E_v^j >= 0)$$

Attaque sur les LFSRs combinés

compter plus vite

$$\mathcal{F}[\mathbf{1}_{\{E_\ell^j\}}](i) = \sum_{E_\ell^j} < i|E_\ell^j >$$

heureusement

$$\prod_{\ell=1}^{v} \frac{1+(-1)^{<i|E_\ell^j>}}{2} = \frac{1}{2^v} \sum_{u=0}^{2^v-1} (-1)^{<i|E_u^j>}$$

où $E_u^j = \bigoplus_{\substack{\ell\in[1,v]\\ u_\ell=1}} E_\ell^j$

et linéarité de la transformée de Walsh

Attaque sur les LFSRs combinés

compter plus vite

- 2^I initialisations
- $E \sim 2^{2n(k-1)+2v}$ équations

$$\longrightarrow 2^{2n(k-1)+3v} + I2^I$$

Résumé de l'attaque

On attaque les registres un par un.

- Précalcul :
  - trouver $\sim 2^{2n(k-1)+2v}$ multiples de poids $2k$ des polynômes de rétroaction des LFSRs restants
  - Calculer les vecteurs E_ℓ^j correspondants
- Phase active :
  - effectuer la transformée de Walsh pour retrouver l'initialisation
- Recommencer sur les registres suivants

Résultats expérimentaux

- 3 LFSRs de longueurs 29, 31 et 37
- une fonction F à 9 entrées ; 3 par LFSR

- On prend $k = 2$, i.e. des équations de parité de poids 4.
- Biais réel : $2^{-8.08}$.

On recherche l'initialisation des registres dans l'ordre 2, 1, 3

- suite chiffrante nécessaire : 2^{24}
- phase active : $\sim 2^{35}$
- temps réel : ~ 10min.

Résultats expérimentaux

- 3 LFSRs de longueurs 29, 31 et 37
- une fonction F à 9 entrées ; 3 par LFSR

- On prend $k = 2$, i.e. des équations de parité de poids 4.
- Biais réel : $2^{-8.08}$.

On recherche l'initialisation des registres dans l'ordre 3, 1, 2

- suite chiffrante nécessaire : $2^{22.6}$
- phase active : $\sim 2^{42}$
- temps réel : ~ 20h.

Conclusion

- On a proposé une attaque par corrélation efficace
- Utilisable avec une fonction de filtrage dépendant de la clé
- Pourrait éventuellement s'utiliser avec des registres non linéaires

Questions

? ? ?