Analyse de la résistance aux attaques algébriques des fonctions de filtrage augmentées

Stage de recherche du MPRI

Stéphane JACOB, encadré par Anne CANTEAUT

Équipe-projet SECRET,
Centre de recherche INRIA Paris-Rocquencourt

24 novembre 2008

Plan

Plan

Introduction

Étude d'une attaque :

- proposée par Simon Fischer et Willi Meier en 2007 ;
- appartient à la famille des attaques algébriques ;
- repose sur les fonctions augmentées ;
- s'applique à des chiffrements jusqu'alors immunisés contre les attaques algébriques.

Introduction

Étude d'une attaque :

- proposée par Simon Fischer et Willi Meier en 2007 ;
- appartient à la famille des attaques algébriques ;
- repose sur les fonctions augmentées ;
- s'applique à des chiffrements jusqu'alors immunisés contre les attaques algébriques.

Introduction

Étude d'une attaque :

- proposée par Simon Fischer et Willi Meier en 2007 ;
- appartient à la famille des attaques algébriques ;
- repose sur les fonctions augmentées ;
- s'applique à des chiffrements jusqu'alors immunisés contre les attaques algébriques.

Introduction

Étude d'une attaque :

- proposée par Simon Fischer et Willi Meier en 2007 ;
- appartient à la famille des attaques algébriques ;
- repose sur les fonctions augmentées ;
- s'applique à des chiffrements jusqu'alors immunisés contre les attaques algébriques.

Introduction

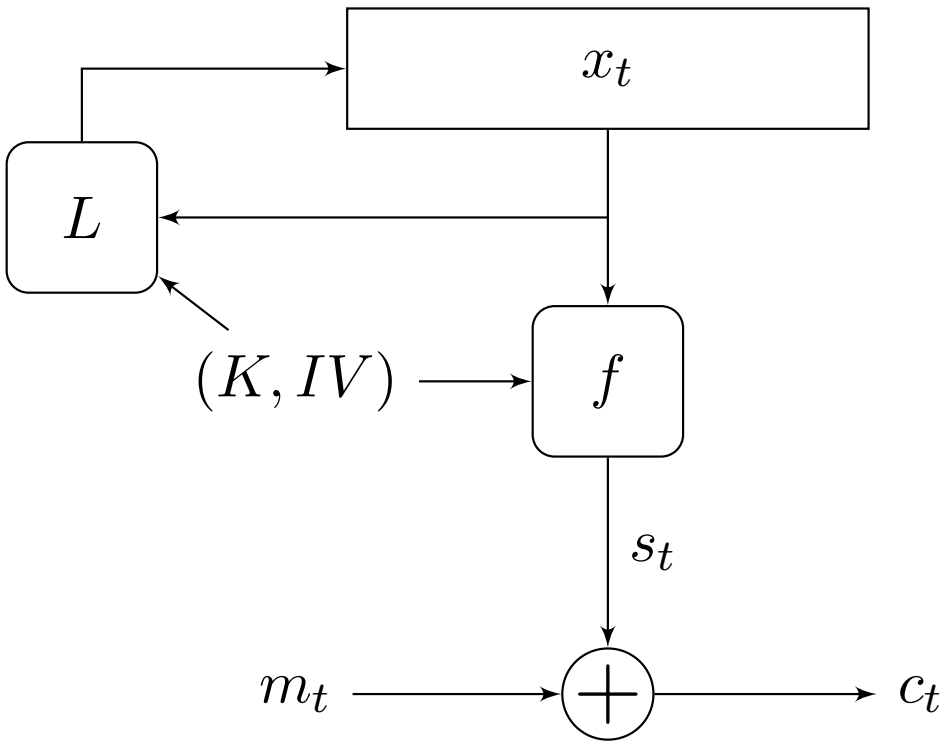
Étude d'une attaque :

- proposée par Simon Fischer et Willi Meier en 2007 ;
- appartient à la famille des attaques algébriques ;
- repose sur les fonctions augmentées ;
- s'applique à des chiffrements jusqu'alors immunisés contre les attaques algébriques.

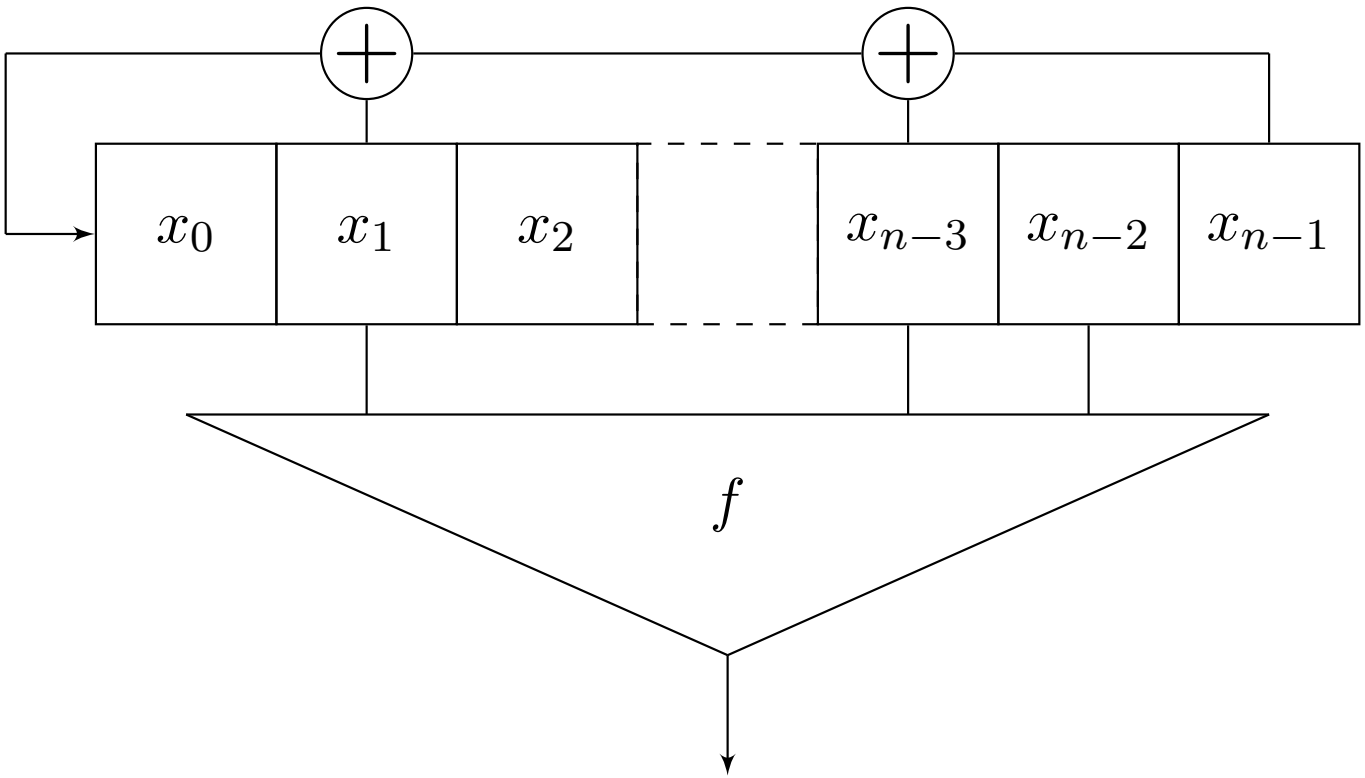
Plan

Plan

Chiffrement à flot additif



LFSR filtré



Attaques algébriques

Attaques algébriques pour les chiffrements à flot :

- Courtois et Meier (03) ;
- mise en système d'équations multivariées liant les bits du chiffré, les bits du clair et ceux de la clef ;
- exploitation de l'existence de multiples de petit degré de la fonction de filtrage.

Attaques algébriques

Attaques algébriques pour les chiffrements à flot :

- Courtois et Meier (03) ;
- mise en système d'équations multivariées liant les bits du chiffré, les bits du clair et ceux de la clef ;
- exploitation de l'existence de multiples de petit degré de la fonction de filtrage.

Attaques algébriques

Attaques algébriques pour les chiffrements à flot :

- Courtois et Meier (03) ;
- mise en système d'équations multivariées liant les bits du chiffré, les bits du clair et ceux de la clef ;
- exploitation de l'existence de multiples de petit degré de la fonction de filtrage.

Attaques algébriques

Attaques algébriques pour les chiffrements à flot :

- Courtois et Meier (03) ;
- mise en système d'équations multivariées liant les bits du chiffré, les bits du clair et ceux de la clef ;
- exploitation de l'existence de multiples de petit degré de la fonction de filtrage.

Plan

Fonctions augmentées

Définition

Considérons un chiffrement à flot d'état interne x de longueur n bits, une fonction de mise à jour L et une fonction de filtrage f qui sort un bit de suite chiffrante par itération. La *fonction augmentée* S_m est définie comme suit :

$$\begin{array}{rccl} S_m : & \mathbb{F}_2^n & \longrightarrow & \mathbb{F}_2^m \\ & x & \mapsto & y = \left(f(x), f(L(x)), \ldots, f(L^{m-1}(x)\right). \end{array}$$

Relations algébriques et relations algébriques conditionnées

Définition

Soit F une fonction de $\mathbb{F}_2^n$ dans $\mathbb{F}_2^m$. On appelle *relation algébrique pour* F toute relation du type

$$\sum_{\alpha \in \mathbb{F}_2^n} \sum_{\beta \in \mathbb{F}_2^m} c_{\alpha,\beta} x^\alpha y^\beta = 0$$

satisfaite pour tout couple $(x, y) \in \mathbb{F}_2^n \times \mathbb{F}_2^m$ où $y = F(x)$, avec $c_{\alpha,\beta} \in \mathbb{F}_2$ et la notation $x^\alpha := (x_1^{\alpha_1} \cdots x_n^{\alpha_n})$.

Relations algébriques et relations algébriques conditionnées

Définition

On appelle *relation algébrique conditionnée* par y_0 toute relation du type

$$\sum_{\alpha \in \mathbb{F}_2^n} c_\alpha x^\alpha = 0$$

satisfaite pour tout $x \ \text{tq} \ F(x) = y_0$, avec $c_\alpha \in \mathbb{F}_2$ et la notation $x^\alpha := (x_1^{\alpha_1} \cdots x_n^{\alpha_n})$.

Principe de l'attaque de Fischer et Meier

Principe

Données : les fonctions de filtrage et de mise à jour, certains bits de suite chiffrante.

But : retrouver l'état interne initial.

Précalcul : trouver des relations algébriques pour S_m de petit degré en les bits de l'état initial ;

Attaque : s'il y a suffisamment de relations, résoudre le système polynomial ainsi formé.

Principe de l'attaque de Fischer et Meier

Principe

Données : les fonctions de filtrage et de mise à jour, certains bits de suite chiffrante.

But : retrouver l'état interne initial.

Précalcul : trouver des relations algébriques pour S_m de petit degré en les bits de l'état initial ;

Attaque : s'il y a suffisamment de relations, résoudre le système polynomial ainsi formé.

Principe de l'attaque de Fischer et Meier

Principe

Données : les fonctions de filtrage et de mise à jour, certains bits de suite chiffrante.

But : retrouver l'état interne initial.

Précalcul : trouver des relations algébriques pour S_m de petit degré en les bits de l'état initial ;

Attaque : s'il y a suffisamment de relations, résoudre le système polynomial ainsi formé.

Principe de l'attaque de Fischer et Meier

Principe

Données : les fonctions de filtrage et de mise à jour, certains bits de suite chiffrante.

But : retrouver l'état interne initial.

Précalcul : trouver des relations algébriques pour S_m de petit degré en les bits de l'état initial ;

Attaque : s'il y a suffisamment de relations, résoudre le système polynomial ainsi formé.

Principe de l'attaque de Fischer et Meier

Principe

Données : les fonctions de filtrage et de mise à jour, certains bits de suite chiffrante.

But : retrouver l'état interne initial.

Précalcul : trouver des relations algébriques pour S_m de petit degré en les bits de l'état initial ;

Attaque : s'il y a suffisamment de relations, résoudre le système polynomial ainsi formé.

Nombre de bits de suite chiffrante nécessaires

Théorème

Soit f une fonction booléenne de filtrage à n variables, et S_m sa fonction augmentée à m sorties. Alors, si $m \geq m_0$ avec $m_0 = n + 1 - \lfloor \log_2(D) \rfloor$ où $D = \sum_{i=0}^{d} \binom{n}{i}$, la fonction S_m possède des relations algébriques de degré inférieur ou égal à d quelle que soit la fonction de mise à jour linéaire L utilisée.

Fonction de filtrage	Premier m tel qu'il y ait une relation linéaire
CanFil1, 2 et 3	13
CanFil4	8-10
CanFil5	6-11
CanFil6	9-13
CanFil7	8-9

Plan

Plan

Liens entre les relations générales et conditionnées

Théorème

Soit F une fonction de $\mathbb{F}_2^n$ dans $\mathbb{F}_2^m$. Alors le degré minimal d atteignable pour une relation algébrique pour F correspond à la plus petite valeur d pour laquelle il existe une relation conditionnée de degré d.

Algorithme avec une unique matrice M

Entrées : une fonction de mise à jour L et une de filtrage f.
Sorties : les relations $\sum c_{\alpha,\beta} x^\alpha S_m(x)^\beta = 0$ de degré $\leq d$ en x.

{**Calcul de la matrice génératrice G du LFSR**}
{**Calcul de M**}
pour tout x dans $\mathbb{F}_2^n$ **faire**
  calculer $S_m(x) \in \mathbb{F}_2^m$
  pour tout $\alpha \in \mathbb{F}_2^n$ tq $w(\alpha) \leq d$ **faire**
    stocker $x^\alpha S_m(x)^\beta, \forall \beta \in \mathbb{F}_2^m$
  fin pour
fin pour
{**Recherche des relations**}
Faire un pivot de Gauss sur M pour trouver son rang et les relations algébriques pour S_m

Algorithme avec plusieurs matrices M_y

Entrées : une fonction de mise à jour L et une de filtrage f.

Sorties : pour chaque y, les relations conditionnées par y de la forme $\sum_\alpha c_\alpha x^\alpha = 0$ pour tout $x \in S_m^{-1}(y)$ de degré au plus d

{**Calcul de la matrice génératrice G du LFSR**}

pour tout y dans $\mathbb{F}_2^m$ **faire**

{**Calcul de M_y**}

pour tout x dans $\mathbb{F}_2^n$ **faire**

calculer $S_m(x) \in \mathbb{F}_2^m$

si $S_m(x) = y$ **alors**

stocker x^α dans M_y, $\forall \alpha \in \mathbb{F}_2^n$ tq $w(\alpha) \leq d$

fin si

fin pour

{**Recherche des relations**}

Trouver le rang de M_y et son noyau par pivot de Gauss.

fin pour

Algorithme hybride

Entrées : une fonction de mise à jour L et une de filtrage f.
Sorties : pour chaque y, les relations conditionnées par y, de la forme $\sum_\alpha c_\alpha x^\alpha = 0$ pour tout $x \in S_m^{-1}(y)$, de degré au plus d.

{**Calcul de la matrice génératrice G du LFSR**}
{**Calcul de M'**}
pour tout x dans $\mathbb{F}_2^n$ **faire**
  calculer $S_m(x) \in \mathbb{F}_2^m$
  pour tout $\alpha \in \mathbb{F}_2^n$ tq $w(\alpha) \leq d$ **faire**
    stocker x^α
    Ajouter $S_m(x)$ à la fin de la ligne courante dans M'
  fin pour
  incrémenter le nombre d'antécédents de $S_m(x)$
fin pour

Algorithme hybride

{**Préparation de** M'}
trier M' en fonction de la valeur des m derniers bits de chaque ligne
{**Recherche des relations**}
pour tout y dans $\mathbb{F}_2^m$ **faire**
faire un pivot de Gauss sur la sous-matrice de M' dont les m derniers bits forment y et qu'on a enlevés.
fin pour

Complexité du précalcul

Comparatif des complexités dans le cas linéaire

	Temps de précalcul	Mémoire
Algorithme avec M	$2^{n+2m}n^2$	$2^{n+m}n$
Algorithme avec les M_y	$2^n(2^m m + n^2)$	$2^{n-m}n$
Algorithme hybride	$n^2 2^n$	$2^n(n+m)$

Plan

Se contenter de très petits m

Maintenir m *très* petit

TAB.: Nombre de x_i passant dans S_m

m	x_i passant dans S_m
1	v
2	$< 2v + wt(P) - 1$
3	$\sim \frac{3}{4}n$
$\vdots$	$\vdots$
6	$\sim n$

Restreindre l'espace des états initiaux parcourus

Ne pas parcourir tous les x :

- suggéré par Fischer et Meier ;
- mais impliquant une complexité multipliée par 10^9 pour $\varepsilon = 10^{-1}$;
- possible en utilisant le décodage par ensemble d'information.

Restreindre l'espace des états initiaux parcourus

Ne pas parcourir tous les x :

- suggéré par Fischer et Meier ;
- mais impliquant une complexité multipliée par 10^9 pour $\varepsilon = 10^{-1}$;
- possible en utilisant le décodage par ensemble d'information.

Restreindre l'espace des états initiaux parcourus

Ne pas parcourir tous les x :

- suggéré par Fischer et Meier ;
- mais impliquant une complexité multipliée par 10^9 pour $\varepsilon = 10^{-1}$;
- possible en utilisant le décodage par ensemble d'information.

Une attaque par décodage

Utilisation du décodage pour l'attaque

- on se ramène à un problème de décodage d'un code linéaire de longueur $\nu \geq \frac{n}{\mathrm{C}(\varepsilon)}$ et de dimension n pour le canal binaire symétrique de probabilité d'erreur ε ;
- on utilise ensuite un décodage par ensemble d'information dont la complexité en donnée est en $\mathcal{O}\left(n^3\left(1-(\ln 2)^{-1}\varepsilon(1-\ln\varepsilon)+o\left(\varepsilon\right)\right)e^{2n(\varepsilon+o(\varepsilon))}\right)$ quand ε est petit.

Complexité deux attaques quand ε varie, $n = 256$

Complexité en données : ($\log_2$ (nb de bits de suite ch.) - m)

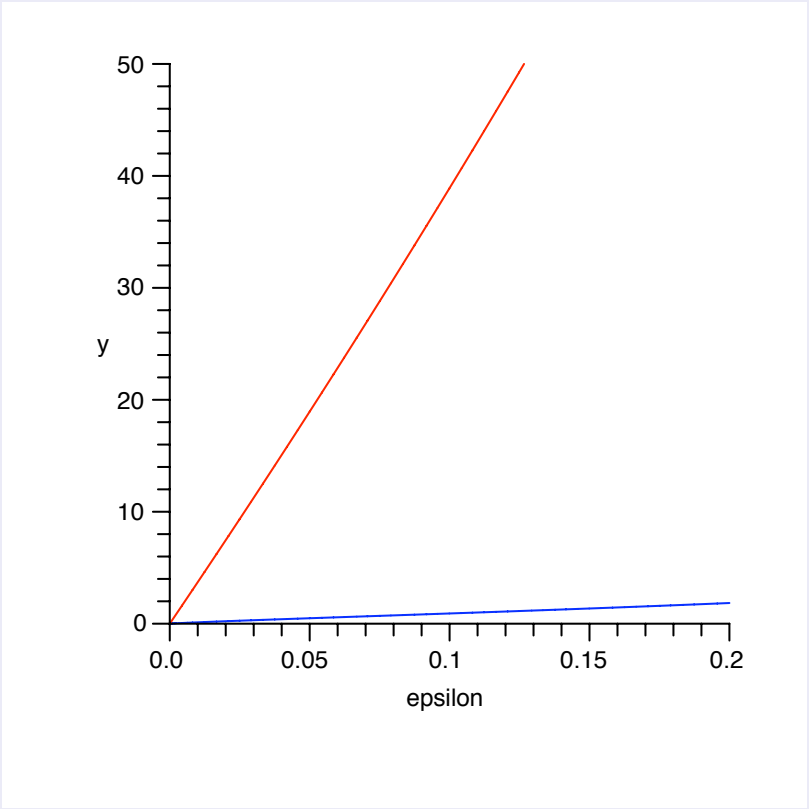


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Complexité deux attaques quand ε varie, $n = 256$

Complexité en temps : ($\log_2$ (nombre d'opérations) - m)

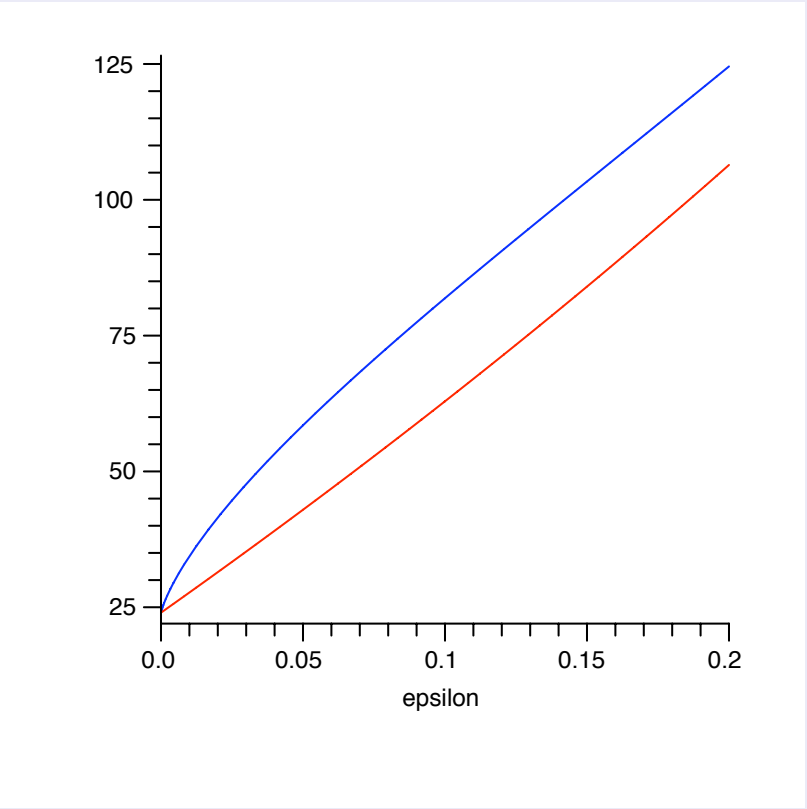


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Plan

Conclusion

Bilan

Nous avons proposé des améliorations :

- une variante de l'algorithme de recherche des relations ;
- une nouvelle attaque permettant de garder m petit et/ou de restreindre le nombre de valeurs de l'état initial parcourues pour rechercher les relations.

Perspectives

- Déterminer précisément l'influence respective des fonctions de filtrage et de mise à jour.
- Exhiber les caractéristiques de celles-ci qui rendent un chiffrement les utilisant vulnérable à cette attaque.

Conclusion

Bilan

Nous avons proposé des améliorations :

- une variante de l'algorithme de recherche des relations ;
- une nouvelle attaque permettant de garder m petit et/ou de restreindre le nombre de valeurs de l'état initial parcourues pour rechercher les relations.

Perspectives

- Déterminer précisément l'influence respective des fonctions de filtrage et de mise à jour.
- Exhiber les caractéristiques de celles-ci qui rendent un chiffrement les utilisant vulnérable à cette attaque.

Conclusion

Bilan

Nous avons proposé des améliorations :

- une variante de l'algorithme de recherche des relations ;
- une nouvelle attaque permettant de garder m petit et/ou de restreindre le nombre de valeurs de l'état initial parcourues pour rechercher les relations.

Perspectives

- Déterminer précisément l'influence respective des fonctions de filtrage et de mise à jour.
- Exhiber les caractéristiques de celles-ci qui rendent un chiffrement les utilisant vulnérable à cette attaque.

Conclusion

Bilan

Nous avons proposé des améliorations :

- une variante de l'algorithme de recherche des relations ;
- une nouvelle attaque permettant de garder m petit et/ou de restreindre le nombre de valeurs de l'état initial parcourues pour rechercher les relations.

Perspectives

- Déterminer précisément l'influence respective des fonctions de filtrage et de mise à jour.
- Exhiber les caractéristiques de celles-ci qui rendent un chiffrement les utilisant vulnérable à cette attaque.

Conclusion

Bilan

Nous avons proposé des améliorations :

- une variante de l'algorithme de recherche des relations ;
- une nouvelle attaque permettant de garder m petit et/ou de restreindre le nombre de valeurs de l'état initial parcourues pour rechercher les relations.

Perspectives

- Déterminer précisément l'influence respective des fonctions de filtrage et de mise à jour.
- Exhiber les caractéristiques de celles-ci qui rendent un chiffrement les utilisant vulnérable à cette attaque.

Plan

Plan

Complexité deux attaques quand ε varie, $n = 512$

Complexité en données : ($\log_2$ (nb de bits de suite ch.) - m)

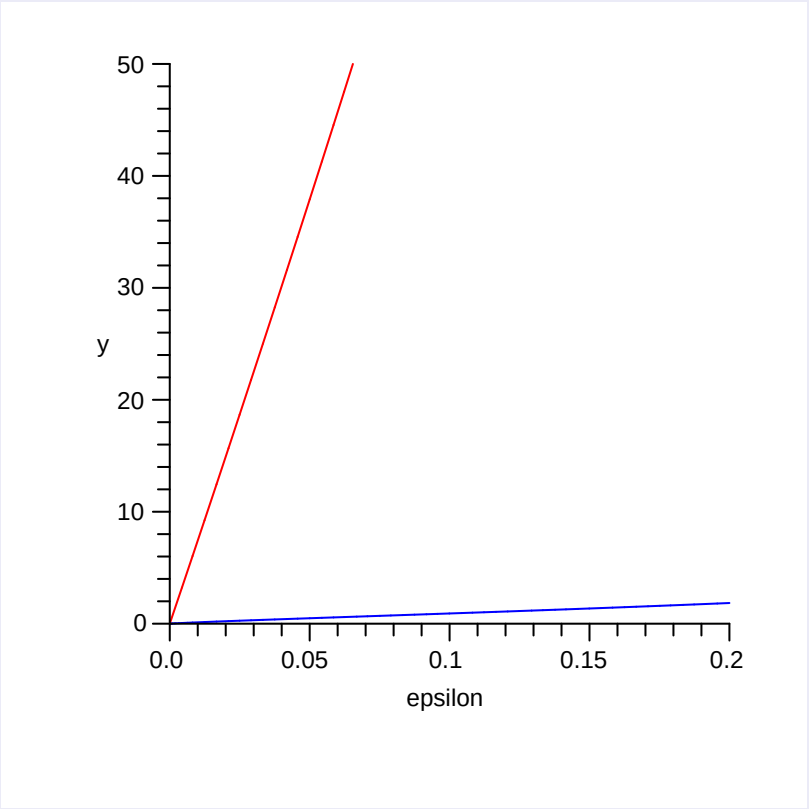


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Complexité deux attaques quand ε varie, $n = 512$

Complexité en temps : ($\log_2$ (nombre d'opérations) - m)

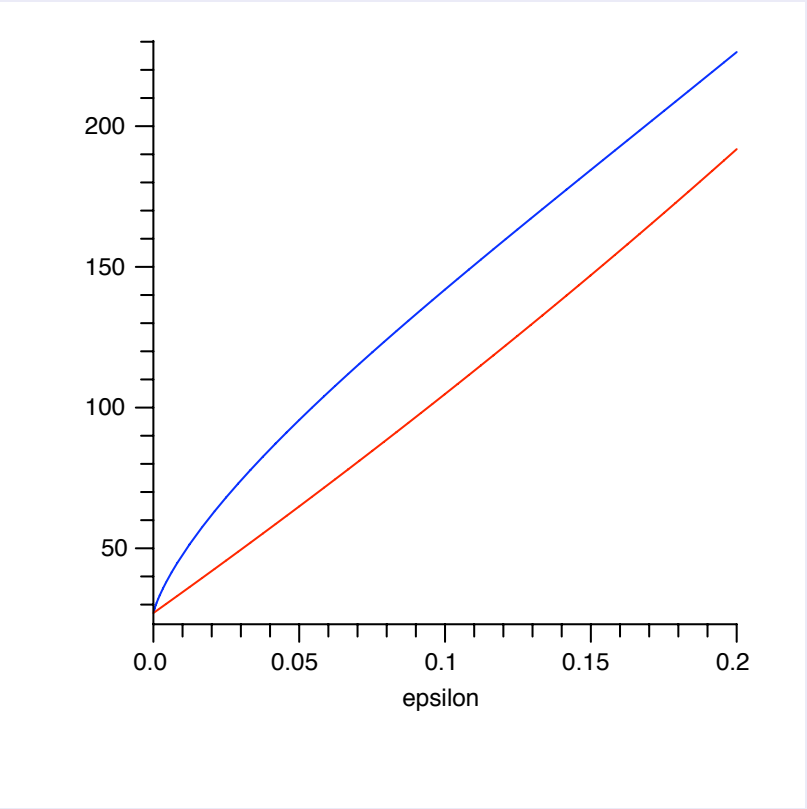


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Complexité deux attaques quand ε varie, $n = 1024$

Complexité en données : ($\log_2$ (nb de bits de suite ch.) - m)

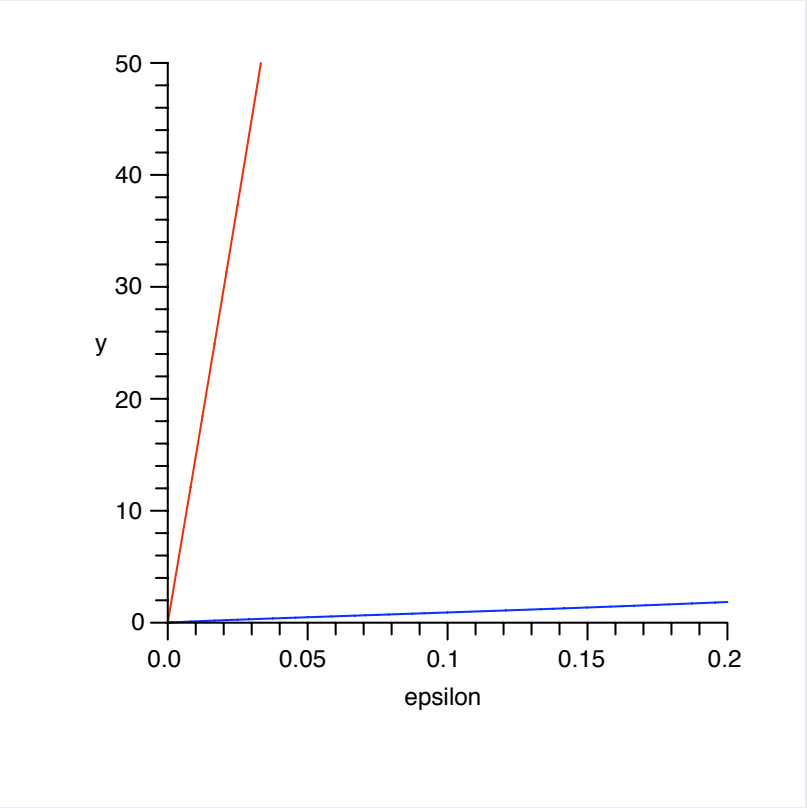


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Complexité deux attaques quand ε varie, $n = 1024$

Complexité en temps : ($\log_2$ (nombre d'opérations) - m)

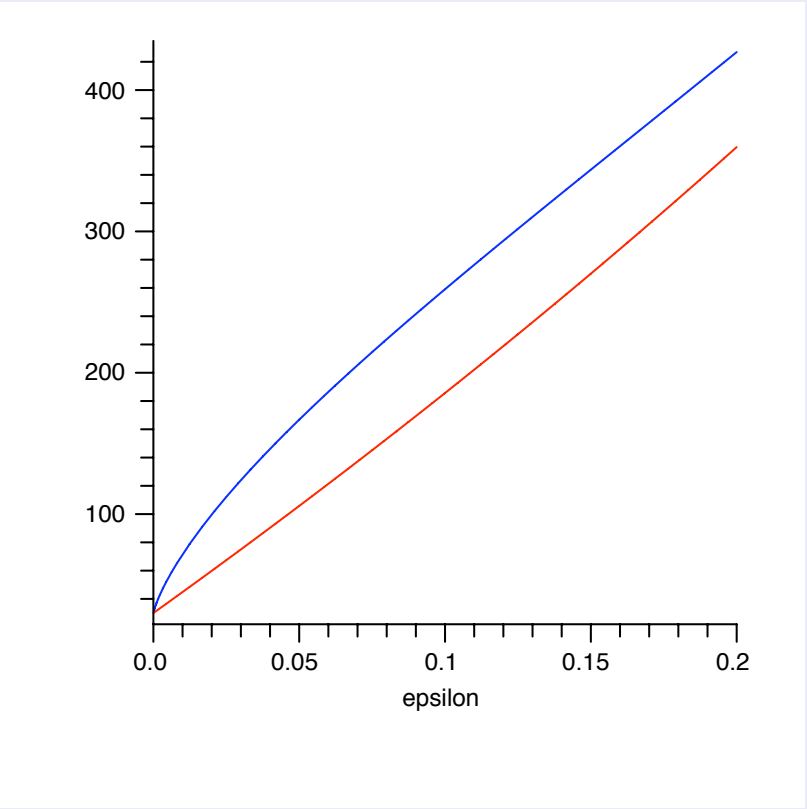


FIG.: Rouge : attaque de Fischer et Meier, bleue : notre variante

Plan

Décodage par ensemble d'information

Entrées :

- G, une matrice génératrice du code (n, k).
- x, un mot de code bruité avec une probabilité d'erreur ε

Sorties : un mot de code à distance de l'ordre de εn de x

tant que un vecteur d'erreurs de poids environ égal à εn n'a pas été trouvé **faire**

1. Choisir aléatoirement un ensemble d'information I.
2. Mettre la matrice G sous la forme systématique $U^{-1}G = (Id, Z)_I$ et décomposer le vecteur x sous la forme $x = (x_I, x_J)_I$.
3. Calculer $w(x_J + x_I Z)$.

si ce poids est de l'ordre de εn **alors**

$(x_I, x_I Z)$ est un mot de code à distance $w\left(x_J + x_I Z\right)$ de x

fin si

fin tant que

Plan

Étape commune à tous les algorithmes

{**Calcul de la matrice génératrice G du LFSR**}
À partir de L, la fonction de mise à jour, calculer la matrice génératrice G telle que :

$$(x_0 \dots x_{n-1}) \begin{pmatrix} 1 & & 0 & \cdot & \dots & \cdot \\ & \ddots & & \cdot & \dots & \cdot \\ 0 & & 1 & \cdot & \dots & \cdot \end{pmatrix} = (x_0 \dots x_{n-1+m})$$