Базисы Грёбнера и алгоритм Бухбергера в криптографии

(интенсив Математического центра Новосибирска)

Сергей Агиевич

НИИ прикладных проблем математики и информатики

Белорусский государственный университет

agievich@{bsu.by|gmail.com}

[Минск-Новосибирск, 2020-07-21]

1. План

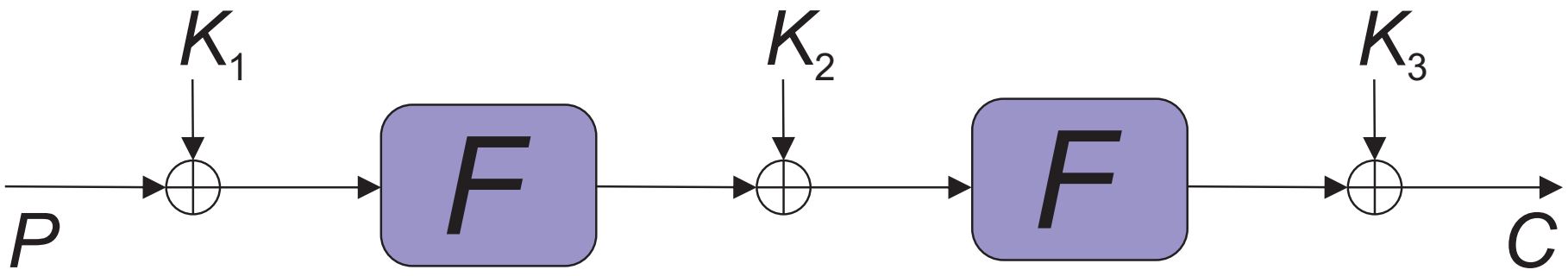
- Учебная криптосистема
- Демонстрация алгебраической атаки
- Попутное введение в алгебраический контекст
- Базисы Грёбнера
- Алгоритм Бухбергера
- Оптимизация алгоритма Бухберегера

2. Реализация плана

Библиотека GF2 (`https://github.com/agievich/GF2`)

- Модуль test/test.cpp
  - Функция testEM() [≈ 60 строк плюс комментарии]

3. Схема Эвена — Мансура (Even-Mansour, EM)



F — биекция на $\{0,1\}^n \sim \mathbb{F}_2^n$

$K_i \in \{0,1\}^n$ — тактовые ключи

$P \in \{0,1\}^n$ — открытый текст

$C \in \{0,1\}^n$ — шифртекст

⚠ Речь идет о 2-тактовой схеме

4. Учебная криптосистема

Инстанциирование EM:

- $n = 3$,
- $F(a, b, c) = (c \oplus a \wedge b, a \oplus b \vee \neg c, b \oplus a \vee c)$
  (циклический сдвиг выхода S-блока Bash-f)

F через многочлены факторкольца $\mathbb{F}_2[a, b, c]/(a^2 - a, b^2 - b, c^2 - c)$:
$F(a, b, c) = (c + ab, a + bc + c + 1, b + ac + a + c)$

Элемент факторкольца:
ANF (алгебраическая нормальная форма), многочлен Жегалкина

Многочлены $a^2 - a$, $b^2 - b$,...:
многочлены поля, описывают принадлежность $\mathbb{F}_2$

5. Алгебраическая атака

Модель:
при известном открытом тексте (KPA = Known Plaintext Attack)

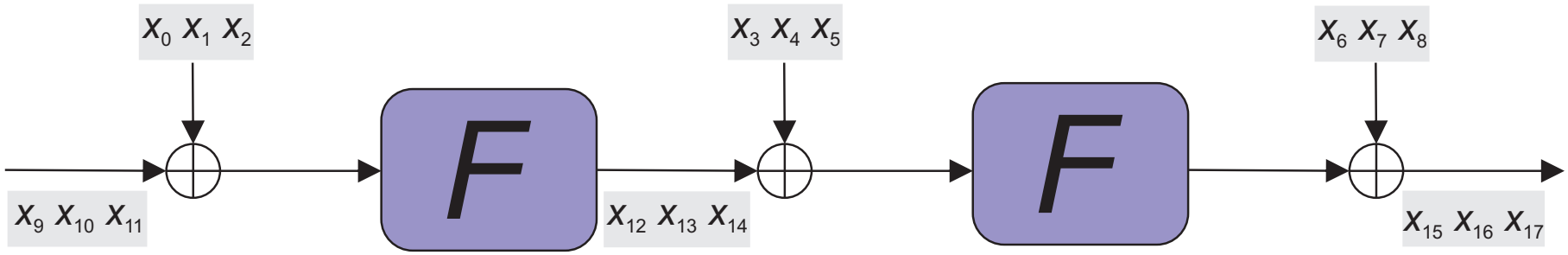
Шифрматериал:
несколько пар (P, C)

Цель:
найти $K = (K_1, K_2, K_3)$ (KR = Key Recovery)

Метод:
составление и решение системы алгебраических уравнений

Система:
- набор многочленов Жегалкина,
- многочлены описывают связь между P и C,
- многочлены описывают шифрматериал

6. Назначение переменных



Введено дополнительное слово $T = F(P \oplus K_1) = x_{12}x_{13}x_{14}$

7. Алгебраическое описание

```
{
  x9 x10 + x1 x9 + x0 x10 + x0 x1 + x12 + x11 + x2,
  x9 x10 + x1 x9 + x0 x10 + x0 x1 + x13 + x11 + x9 + x2 + x0 + 1,
  x9 x11 + x2 x9 + x0 x11 + x0 x2 + x14 + x11 + x10 + x9 + x2 + x1 + x0,
  x12 x13 + x4 x12 + x3 x13 + x3 x4 + x15 + x14 + x6 + x5,
  x12 x14 + x5 x12 + x3 x14 + x3 x5 + x17 + x14 + x13 + x12 + x8 + x5 + x4 + x3,
  x13 x14 + x5 x13 + x4 x14 + x4 x5 + x16 + x14 + x12 + x7 + x5 + x3 + 1
}
```

Первая тройка многочленов описывает равенство $T = F(P \oplus K_1)$

Вторая тройка описывает равенство $C \oplus K_3 = F(T \oplus K_2)$

⚠ Неявно присутствуют многочлены поля

8. Описание шифрматериала

Равенство $x_i = c_i$ задается многочленом $x_i + c_i$

$P = 000$: $\{x_9, x_{10}, x_{11}\}$

$C = 111$: $\{x_{12} + 1, x_{13} + 1, x_{14} + 1\}$

```
{
  x9,
  x10,
  x11,
  x15 + 1,
  x16 + 1,
  x17 + 1,
  x9 x10 + x0 x10 + x1 x9 + x0 x1 + x12 + x11 + x2,
  x9 x10 + x1 x9 + x0 x10 + x0 x1 + x13 + x11 + x9 + x2 + x0 + 1,
  x9 x11 + x2 x9 + x0 x11 + x0 x2 + x14 + x11 + x10 + x9 + x2 + x1 + x0,
  x12 x13 + x4 x12 + x3 x13 + x3 x4 + x15 + x14 + x6 + x5,
  x12 x14 + x5 x12 + x3 x14 + x3 x5 + x17 + x14 + x13 + x12 + x8 + x5 + x4 + x3,
  x13 x14 + x5 x13 + x4 x14 + x4 x5 + x16 + x14 + x12 + x7 + x5 + x3 + 1
}
```

Как решать?

9. Алгебра: мономы

Идет речь о кольце $R = \mathbb{F}_2[x_0, x_1, \ldots]$

⚠ Число переменных конечно

⚠ Пока без многочленов поля

Многочлен $f \in R$ представляет собой сумму мономов вида
$m = x_0^{\alpha_0} x_1^{\alpha_1} \ldots$, $\alpha_i \in \mathbb{Z}_{\geq 0}$

$\mathbb{M}$ — множество всех мономов

Мультистепень m: $\mathrm{multideg}(m) = (\alpha_0, \alpha_1, \ldots)$

Степень m: $\deg(m) = \alpha_0 + \alpha_1 + \ldots$

Степень f **(**$f \neq 0$**)**: максимальная из степеней его мономов

10. Алгебра: мономиальный порядок

Определение. Порядок $<$ на $\mathbb{M}$ называется мономиальным, если:
1) $m < m' \Rightarrow mt < m't$ для любого $t \in \mathbb{M}$;
2) $1 \leq m$ для любого $m \in \mathbb{M}$

Порядок lex: $m < m'$, если самая правая ненулевая координата $\mathrm{multideg}(m) - \mathrm{multideg}(m')$ отрицательна ($x_0 < x_1 < x_2 < \ldots$)

⚠ Соглашения GF2. Часто самая левая ($x_0 > x_1 > x_2 > \ldots$).

Grlex: $m < m'$, если $\deg(m) < \deg(m')$ или $\deg(m) = \deg(m')$ и $m <_{\mathrm{lex}} m'$

Grevlex: $m < m'$, если $\deg(m) < \deg(m')$ или $\deg(m) = \deg(m')$ и самая левая ненулевая координата $\mathrm{multideg}(m) - \mathrm{multideg}(m')$ положительна

Старший моном f (при выбранном порядке): $\mathrm{LM}(f)$

11. Алгебра: деление

Пока некоторый моном m многочлена f делится на $\mathrm{LM}(g)$:
$f \leftarrow f - \frac{m}{\mathrm{LM}(g)} g$

Пример 1. Деление на $x_i + c$ — замена x_i на c.

Пример 2. Деление на $x_i^2 - x_i$ — замена x_i^2 на x_i, x_i^3 на x_i^2, ...

⚠ Системы $\{f, g\}$ до деления и после эквивалентны!

Деление f на $G = \{g_1, g_2, \ldots\}$ (редукция):
делить пока делится на некоторый g_i;
запись редукции: $f \xrightarrow{G} h$; h — нормальная форма

⚠ Нормальная форма h определяется неоднозначно: зависит от выбора подходящего многочлена $g_i \in G$ и подходящего монома $m \in f$ на шагах деления

Однозначный результат:
$\bar{f}^G$ (при определенной стратегии выбора)

12. Саморедукция

Идея: упростить систему G, деля $g \in G$ на $G \setminus \{g\}$ пока делится

Результат (grlex):

```
{
  x9,
  x10,
  x11,
  x13 + x12 + x0 + 1,
  x15 + 1,
  x16 + 1,
  x17 + 1,
  x0 x1 + x12 + x2,
  x0 x2 + x14 + x2 + x1 + x0,
  x4 x12 + x3 x12 + x0 x12 + x3 x4 + x0 x3 + x14 + x6 + x5 + x3 + 1,
  x4 x14 + x3 x14 + x0 x14 + x4 x5 + x3 x5 + x0 x5 + x14 + x12 + x8 + x7 + x5 + x4 + x0,
  x12 x14 + x3 x14 + x5 x12 + x3 x5 + x14 + x8 + x5 + x4 + x3 + x0
}
```

Система упростилась, но несущественно.

⚠ Каждая из переменных $x_9, x_{10}, x_{11}, x_{13}, x_{15}, x_{16}, x_{17}$ входит только в один многочлен (исключена из других).

13. Алгебра: базис Грёбнера

Система $G = \{g_1, \ldots, g_r\} \subset R$ определяет **идеал**
$I = \langle G \rangle = \{h_1 g_1 + \ldots + h_r g_r : h_i \in R\}$

G является **базисом** I. У идеала I может быть несколько базисов.

Базис Грёбнера: для любого $f \in I$ найдется $g_i \in G$ такой, что $\mathrm{LM}(g_i) \mid \mathrm{LM}(f)$.

⚠ Базисы Грёбнера ввел Бруно Бухбергер, ученик Вольфанга Грёбнера (1965).

Минимальный базис Грёбнера: $\mathrm{LM}(g_i)$ не делится на $\mathrm{LM}(g_j)$ $(i \neq j)$.

Приведенный базис Грёбнера: ни один из мономов g_i не делится на $\mathrm{LM}(g_j)$ $(i \neq j)$.

14. Базис Грёбнера: факты

Факт 1. Если G — базис Грёбнера идеала I, то $f \xrightarrow{G} 0$ для любого $f \in I$ (независимо от стратегии деления).

Факт 2. Многочлены поля первоначальной системы G можно перенести в базис Грёбнера. Тогда остальные многочлены базиса будут многочленами Жегалкина.

Факт 3. Если система имеет единственное решение $x_i = c_i$, то приведенный базис Грёбнера состоит из многочленов $x_i + c_i$.

Факт 4. Если решений нет, то базис Грёбнера содержит многочлен 1.

Факт 5. Число решений — это число мономов $m \in M$, которые не делятся на старшие мономы многочленов базиса Грёбнера (можно эффективно вычислить по базису).

15. Вернемся к атаке...

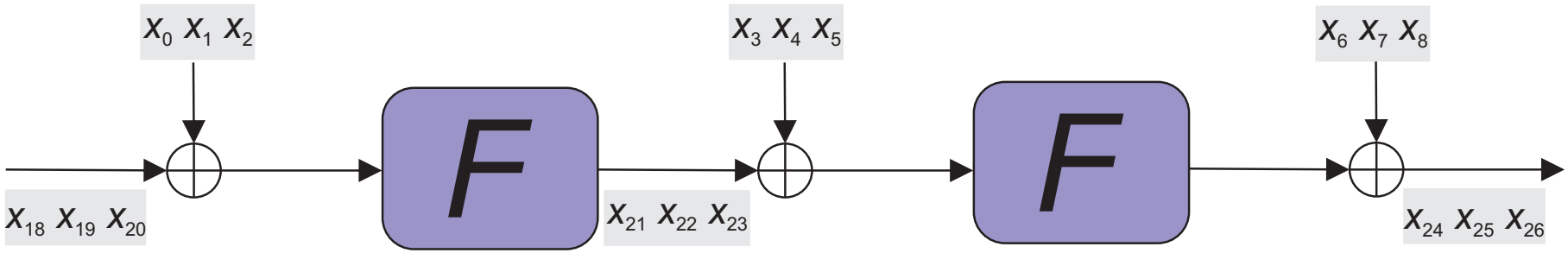
Базис Грёбнера (grlex, без многочленов поля):

```
{
  x9,
  x10,
  x11,
  x13 + x12 + x0 + 1,
  x15 + 1,
  x16 + 1,
  x17 + 1,
  x0 x1 + x12 + x2,
  ...
  x2 x5 x12 + x2 x3 x5 + x1 x2 x3 + x2 x8 + x2 x5 + x2 x4 + x14 + x2 + x1 + x0,
  x3 x5 x12 + x1 x2 x3 + x3 x8 + x3 x4 + x2 x3 + x0 x3 + x3,
  x1 x2 x3 x5 + x1 x2 x8 + x2 x3 x5 + x1 x2 x4 + x2 x12 + x2 x8 + x2 x4 + x14 + x1 + x0
}
```

58 многочленов, степени от 1 до 4

Число решений: 64

16. Дополнительный шифрматериал



$P = 001$: $\{x_{18}, x_{19}, x_{20} + 1\}$

$C = 010$: $\{x_{24}, x_{25} + 1, x_{26}\}$

Базис Грёбнера:

24 многочлена степени 1 (линейная система!);

8 решений

17. Завершение атаки

$(P = 100, C = 100)$: 2 решения
$(P = 101, C = 001)$: 1 решение

Базис Грёбнера:

```
{
  x0 + 1,
  x1,
  x2 + 1,
  x3 + 1,
  x4 + 1,
  x5,
  x6,
  x7 + 1,
  x8,
  ...
}
```

Ключ: $K_1 = x_0x_1x_2 = 101$, $K_2 = x_3x_4x_5 = 110$, $K_3 = x_6x_7x_8 = 010$

18. Базис Грёбнера: S-многочлены

S-многочлен пары (g_i, g_j):

$$S(g_i, g_j) = \frac{[\mathrm{LM}(g_i), \mathrm{LM}(g_j)]}{\mathrm{LM}(g_i)} g_i - \frac{[\mathrm{LM}(g_i), \mathrm{LM}(g_j)]}{\mathrm{LM}(g_j)} g_j,$$

$[\mathrm{LM}(g_i), \mathrm{LM}(g_j)]$ — н.о.к. мономов $\mathrm{LM}(g_i)$ и $\mathrm{LM}(g_j)$.

Критерий Бухбергера. G — базис Грёбнера тогда и только тогда, когда $S(g_i, g_j) \xrightarrow{G} 0$ для любых $g_i, g_j \in G$.

⚠ Неформально, $S(g_i, g_j)$ — это “следствие” “теорем” g_i и g_j. Следствие может тривиально вытекать из предыдущих теорем или давать новую теорему. В базисе Грёбнера все следствия тривиальны. Достигнута система “аксиом”.

19. Алгоритм Бухбергера: схема

Параметры: описание R, порядок $<$.

Вход: $F \subset R$.

Выход: $G \subset R$ — базис Грёбнера идеала $\langle F \rangle$ относительно $<$.

Шаги:

1. $G \leftarrow \varnothing$, $B \leftarrow \varnothing$, $B^* \leftarrow \varnothing$.
2. Для всех $f \in F$:
   (a) $g \leftarrow \bar{f}^G$;
   (b) если $g \neq 0$, то $(G, B) \leftarrow \texttt{Update}(G, B, B^*, g)$.
3. Пока $B \neq \varnothing$:
   (a) выбрать $(g_i, g_j) \in B$,
   (b) $B \leftarrow B \setminus \{(g_i, g_j)\}$, $B^* \leftarrow B^* \cup \{(g_i, g_j)\}$;
   (c) $g \leftarrow \overline{S(g_i, g_j)}^G$;
   (d) если $g \neq 0$, то $(G, B) \leftarrow \texttt{Update}(G, B, B^*, g)$.
4. Возвратить G.

(g_i, g_j) — критические пары

B / B^* — критические пары для обработки / обработанные

20. Алгоритм Бухбергера: обновление

Алгоритм Update

Вход: $G \subset R$, $B \subset R \times R$, $B^* \subset R \times R$, $g \in R \setminus \{0\}$.

Выход: обновленные G и B.

Шаги:

1. $G \leftarrow G \cup \{g\}$.
2. $B \leftarrow B \cup \{(f, g) : f \in G\}$.
3. Возвратить (G, B).

---

21. Алгоритм Бухбергера: бесполезные редукции

⚠ Более 90% вычислений в алгоритме Бухбергера бесполезны: $\overline{S(g_i, g_j)}^{G} = 0$.

Первый критерий Бухбергера: если $[\mathrm{LM}(g_i), \mathrm{LM}(g_j)] = \mathrm{LM}(g_i) \cdot \mathrm{LM}(g_j)$ (нет зацепления), то $S(g_i, g_j) \overset{\{g_i, g_j\}}{\to} 0$ и (g_i, g_j) можно не включать в B.

Второй критерий Бухбергера: если $(g_i, g_k) \in B^*$ и $(g_j, g_k) \in B^*$ (пары уже обработаны) и $\mathrm{LM}(g_k) \mid [\mathrm{LM}(g_i), \mathrm{LM}(g_j)]$, то $S(g_i, g_j) \overset{\{S(g_i, g_k), S(g_j, g_k)\}}{\to} 0$ и (g_i, g_j) можно не включать в B.

Общая стратегия: раннее обнаружение редукций $\overset{G}{\to} 0$ без полноценного деления.

Важный пример: GMI (Gebauer — Möller Installation, 1987) — критерии исключения критических пар.

22. Другие направления оптимизации

- Ускорение редукции:
  одновременно несколько многочленов;
  матричные техники

- Выбор мономиального порядка
  `grevlex` считается лучшим

- Стратегии выбора критических пар
  нормальная стратегия: min [LM(g_i), LM(g_j)]

- Подробная история вычислений
  не только B^*

23. Алгоритмы F4 и F5

Автор: Жан-Шарль Фожер (Jean-Charles Faugère)

F4 (1999): матричное приведение сразу многих многочленов

F5 (2002): без бесполезных редукций, история (сигнатуры)

Атаки с помощью F5: HFE и C* (Multivariate-Based Crypto)

Трудности с F5:

только скетч алгоритма от автора;

только закрытая реализация от автора.