УДК 621.391:519.2

МЕТОД ПОШУКУ АЛГЕБРАЇЧНО ВИРОДЖЕНИХ НАБЛИЖЕНЬ БУЛЕВИХ ФУНКЦІЙ ДЛЯ ПОБУДОВИ СТАТИСТИЧНИХ АТАК НА СИНХРОННІ ПОТОКОВІ ШИФРИ

Олексійчук Антон; Конюшок Сергій; Сторожук Артем
ІСЗЗІ НТУУ "КПІ"

A METHOD FOR FINDING ALGEBRAIC DEGENERATE APPROXIMATIONS OF BOOLEAN FUNCTIONS FOR CONSTRUCTION OF STATISTICAL ATTACKS ON SINCHRONOUS STREAM CIPHERS

Alekseychuk Anton; Konushok Sergey; Storozhuk Artem
ІСЗЗІ НТУУ "КПІ"

Анотація: **Запропоновано метод знаходження (або обґрунтування відсутності) певних алгебраїчно вироджених наближень булевих функцій, заданих за допомогою оракулів. Зазначений метод відрізняється за сутністю від раніше відомих та базується на отриманих необхідних умовах існування шуканих наближень. Для розв'язання окремих задач на кожному етапі методу запропоновано поліноміальні ймовірнісні алгоритми, які можуть бути застосовані на практиці до функцій-оракулів від декількох десятків чи сотень змінних. Отримані результати надають можливість будувати більш ефективні статистичні атаки на синхронні потокові шифри, а також обґрунтовувати практичну стійкість таких шифрів відносно зазначених атак.**
Ключові слова: **Синхронний потоковий шифр, кореляційний криптоаналіз, алгебраїчно вироджене наближення булевої функції, ймовірнісний алгоритм, Grain-128.**
Summary: **A method for finding (or providing absence justification) of certain algebraic degenerate approximations of Boolean functions defined by oracles is proposed. This method differs by its nature from previously known and is based on obtained necessary conditions for the existence of searched approximations. To solve particular tasks on each step of the method there are proposed polynomial probabilistic algorithms that can be applied to functions-oracles of tens or hundreds of variables. Obtained results allow us to construct more effective statistical attacks on synchronous stream ciphers and also to prove practical resistance of such ciphers against mentioned attacks.**
Keywords: **Synchronous stream cipher, correlation cryptanalysis, algebraically degenerate approximation of Boolean functions, probabilistic algorithm, Grain-128**

Вступ

В [1] описано статистичну атаку на синхронні потокові шифри (СПШ), яка узагальнює раніше відомі атаки на основі підібраних векторів ініціалізації [2], [3]. Згідно з [1], оцінювання стійкості СПШ відносно зазначеної атаки зводиться до знаходження або обґрунтування відсутності певних алгебраїчно вироджених наближень булевих функцій, заданих за допомогою оракулів. Ця задача є обчислювально складною, а ефективні алгоритми її розв'язання відомі лише для окремих випадків [2], [4] – [7].

В даній статті пропонується метод розв'язання поставленої задачі, який відрізняється за сутністю від відомих [2], [4] – [7] та базується на отриманих нижче умовах, яким задовольняють шукані наближення. На відміну від [4] – [6], запропонований метод не має передумовою виконання будь-яких обмежень стосовно відстані, на якій треба відшукати наближення, а на відміну від [2], [7] він дозволяє знаходити наближення з більш широкого класу булевих функцій. Крім того, за певних умов запропонований метод надає можливість переконуватися у відсутності зазначених наближень, що дозволяє використовувати його для обґрунтування практичної стійкості СПШ відносно статистичних атак [1], [2].

Решта статті має таку структуру. В п. II наведені означення основних понять та певні допоміжні результати, що використовуються далі. В п. III сформульовано точну постановку задачі та

основні етапи методу її розв'язання. Детальному викладенню методу присвячено пп. IV, V, а в п. VI наведено результати його практичного застосування до криптоаналізу СПШ Grain-128. Нарешті, в останньому пункті статті сформульовано стислі висновки.

Означення основних понять та допоміжні результати

Позначимо V_n множину двійкових векторів довжини n. Ця множина є векторним простором вимірності n над полем $F_2 = \{0,1\}$. Сума векторів $\alpha = (\alpha_1,...,\alpha_n)$, $x = (x_1,...,x_n) \in V_n$ визначається за формулою $\alpha \oplus x = (\alpha_1 \oplus x_1, \dots, \alpha_n \oplus x_n)$, а булев скалярний добуток – за формулою $\alpha x = \alpha_1 x_1 \oplus \cdots \oplus \alpha_n x_n$ (тут і далі символ $\oplus$ позначає операцію додавання як елементів поля F_2, так і векторів над цим полем).

Нижче використовуються такі позначення:

$\#U$ – потужність множини U;

$<U>$ – підпростір векторного простору V_n, породжений множиною $U \subseteq V_n$;

$U^{\perp}$ – підпростір, дуальний до множини $U \subseteq V_n$: $U^{\perp} = \{\alpha \in V_n \mid \forall x \in U : \alpha x = 0\}$;

$F_2^{m \times n}$ – множина матриць розміру $m \times n$ над полем F_2;

I_m – одинична матриця порядку m над полем F_2;

$\overline{a,b} = \{i \in Z : a \le i \le b\}$, $a, b \in Z$.

Позначимо B_n множину булевих функцій від n змінних. Відносна відстань між функціями $f, g \in B_n$ визначається за формулою $d(f,g) = 2^{-n} \#\{x \in V_n : f(x) \ne g(x)\}$, а відносна відстань між функцією $f \in B_n$ та множиною $U \subseteq B_n$ – за формулою $d(f,U) = \min_{g \in U} d(f,g)$. Число $wt(f) = 2^{-n} \#\{x \in V_n : f(x) = 1\}$ називається відносною вагою функції $f \in B_n$.

Для будь-якої функції $F \in B_n$ покладемо

$$\hat{F}(\alpha) = 2^{-n} \sum_{x \in V_n} (-1)^{F(x) \oplus \alpha x}, \alpha \in V_n \qquad (1)$$

$$D_\alpha F(x) = F(x \oplus \alpha) \oplus F(x), x \in V_n \qquad (2)$$

Числа (1) називаються нормованими коефіцієнтами Уолша-Адамара функції F, а функція (2) – її похідною за напрямом $\alpha \in V_n$ [8].

Для будь-якого $m \in \overline{0,n}$ позначимо $L_{n,m}$ множину всіх m-вимірних підпросторів векторного простору V_n. Покладемо

$$\omega_F(H) = \sum_{\beta \in H} \hat{F}(\beta)^2, H \in L_{n,m} \qquad (3)$$

Справедливі такі рівності (див., наприклад, теорему 2.89 в [8] та "факт 11" в [4]):

$$\omega_F(H) = 1 - 2 \cdot \left(2^{-(n-m)} \sum_{\alpha \in H^{\perp}} wt(D_\alpha F) \right), \quad H \in L_{n,m}. \qquad (4)$$

$$wt(D_\alpha F) = \sum_{\beta \in V_n : \alpha\beta = 1} \hat{F}(\beta)^2, \alpha \in V_n \qquad (5)$$

Функція $g \in B_n$ називається m-вимірною, $m \in \overline{0,n}$, якщо вона допускає представлення у вигляді

$$g(x) = \phi(xM), x \in V_n \qquad (6)$$

де $\phi \in B_m$, $M \in F_2^{n \times m}$. Булева функція від n змінних, яка є m-вимірною для деякого $m \le n-1$, називається *алгебраїчно виродженою* [4] – [7].

Для будь-якого $H \in L_{n,m}$ позначимо $B_n(H)$ множину всіх функцій вигляду (6) таких, що $\phi \in B_m$, $M \in F_2^{n \times m}$ і стовпці матриці M належать підпростору H. Зрозуміло, що множина m-вимірних функцій від n змінних є об'єднанням множин $B_n(H)$ за всіма $H \in L_{n,m}$. Наступне твердження є прямим наслідком лем 3 і 4 в [9].

Твердження 1 [9]**.** Для будь-яких $F \in B_n$, $H \in L_{n,m}$ справедливі такі співвідношення:

$$d(F, B_n(H)) = 1/2 \cdot (1 - l_F(H)), \qquad (7)$$

$$1/2 \cdot (1 - \sqrt{\omega_F(H)}) \le d(F, B_n(H)) \le 1/2 \cdot (1 - \omega_F(H)), \qquad (8)$$

де

ISSN 2074-9481
Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, вип. 1 (31), 2016р.

$$l_F(H) = 2^{-m} \sum_{y \in V_m} \left| 2^{-(n-m)} \sum_{x \in L_y} (-1)^{F(x)} \right| \qquad (9)$$

$$L_y = \{x \in V_n : xM = y\}, y \in V_m$$

Підпростір $H \in I_{n,m}$ називається θ-допустимим для функції $F \in B_n$, якщо виконується умова $d(F, B_n(H)) \le 1/2 \cdot (1-\theta)$, $\theta \in (0,1)$ [9]. З твердження 1 випливає, що H є θ-допустимим підпростором для функції F тоді й тільки тоді, коли $l_F(H) \ge \theta$; крім того, H є θ-допустимим підпростором за умови $\omega_F(H) \ge \theta$.

Постановка задачі та основні етапи методу її розв'язання

Нехай $F = F(k, c)$ – булева функція від $n = l_0 + l_1$ змінних $k \in V_{l_0}$, $l \in V_{l_1}$, задана за допомогою оракула. Треба з'ясувати, чи існують для заданих чисел $s \in \overline{1, l_0 - 3}$, $\theta \in (0,1)$ функція $\phi \in B_{s+l_1}$ та матриця $M_0 \in F_2^{l_0 \times s}$ рангу s такі, що відносна відстань між функціями F та $\phi(kM_0, c)$, $k \in V_{l_0}$, $l \in V_{l_1}$ не перевищує $1/2 \cdot (1-\theta)$. У випадку існування зазначеної пари (M_0, ϕ) треба побудувати в явному вигляді матрицю M_0 та визначити алгоритм, що реалізує функцію ϕ.

Назвемо функцію $g(k,c) = \phi(kM_0, c)$, $k \in V_{l_0}$, $l \in V_{l_1}$, θ-допустимим наближенням функції F, якщо $\phi \in B_{s+l_1}$, $M_0 \in F_2^{l_0 \times s}$, $\operatorname{rank}(M_0) = s$ і $d(F, g) \le 1/2 \cdot (1-\theta)$. Отже, задача полягає в перевірці існування та побудові (у випадку позитивного результату перевірки) хоча б одного $(s+l_1)$-вимірного θ-допустимого наближення функції F.

Введемо такі підпростори векторного простору V_n:

$$H_0 = \{(0,c) : c \in V_{l_1}\}, \qquad (10)$$

$$L_0 = H_0^{\perp} = \{(k,0) : k \in V_{l_0}\}. \qquad (11)$$

Метод розв'язання поставленої задачі, що пропонується, базується на такому твердженні, справедливість якого випливає безпосередньо з наведених означень.

Твердження 2. Нехай $F \in B_n$, $n = l_0 + l_1$, $s \in \overline{1, l_0 - 3}$, $\theta \in (0,1)$ і $g(k,c) = \phi(kM_0, c)$, $k \in V_{l_0}$, $l \in V_{l_1}$ є $(s+l_1)$-вимірним θ-допустимим наближенням функції F. Тоді стовпці матриці

$$M = \begin{pmatrix} M_0 & 0 \\ 0 & I_{l_1} \end{pmatrix} \qquad (12)$$

утворюють базис θ-допустимого для функції F підпростіру H такого, що $\dim H = s + l_1$ і $H \supseteq H_0$.

Навпаки, якщо $H \supseteq H_0$ є $(s+l_1)$-вимірним θ-допустимим для F підпростором, то існують базис цього підпростору, який складається зі стовпців матриці M вигляду (12), де $\operatorname{rank}(M_0) = s$, а також функція $\phi \in B_{s+l_1}$ такі, що $g(k,c) = \phi(kM_0, c)$, $k \in V_{l_0}$, $l \in V_{l_1}$ є $(s+l_1)$-вимірним θ-допустимим наближенням функції F.

Отже, на підставі твердження 2 задачу пошуку зазначених наближень функції F можна розв'язувати в два етапи. На першому етапі здійснюється пошук $(s+l_1)$-вимірних θ-допустимих для функції F підпросторів $H \supseteq H_0$. Потім, на другому етапі визначається алгоритм, що реалізує шукану функцію $\phi \in B_{s+l_1}$.

Перейдемо до детального викладення алгоритмів, які виконуються на кожному з зазначених етапів.

Перший етап: пошук допустимих підпросторів для функції F

Встановимо необхідні умови існування $(s+l_1)$-вимірних θ-допустимих для функції F підпросторів $H \supseteq H_0$.

Для будь-якого підпростору $L \subseteq V_n$ покладемо $\xi_L(\alpha) = wt(D_\alpha F)$, $\alpha \in L$. Задамо на підпросторі L рівномірний розподіл ймовірностей. Тоді ξ_L є випадковою величиною. Наступна лема встановлює основні властивості цієї випадкової

величини.

Лема 1. Справедливі такі твердження:

1) ξ_L є сумою попарно незалежних випадкових величин:

$$\xi_L(\alpha)=\sum_{\beta\in V_n\setminus L^\perp}\hat{F}(\beta)^2 I_\beta(\alpha),$$ де $I_\beta(\alpha)$ – індикатор події $\{\alpha\beta=1\}$, $\alpha\in L$;

2) $E\xi_L=1/2\cdot(1-\omega_F(L^\perp))$;

3) $D\xi_L\le 1/4\cdot\sum_{\beta\in V_n}\hat{F}(\beta)^4$;

4) за умови $\sum_{\beta\in V_n}\hat{F}(\beta)^4\le\mu^2<1$ для будь-якого $\varepsilon>0$ має місце нерівність

$$P\{|\xi_L-E\xi_L|\ge\varepsilon\}\le\left(\frac{\mu}{2\varepsilon}\right)^2.$$

Доведення. Справедливість п. 1) випливає з формули (5) та незалежності подій $\{\alpha\beta_1=1\}$ і $\{\alpha\beta_2=1\}$ для будь-яких різних векторів $\beta_1,\beta_2\in V_n\setminus L^\perp$; п. 2) є безпосереднім наслідком формул (3), (4), а п. 3) випливає з п. 1). Нарешті, п. 4) є наслідком п. 3) та нерівності Чебишова. Лему доведено.

Наступне твердження відіграє ключову роль для побудови методу пошуку допустимих підпросторів.

Твердження 3. Нехай H є $(s+l_1)$-вимірним θ-допустимим для функції F підпростором і $L=H^\perp$. Тоді для кожного $\alpha\in L$ справедлива нерівність $wt(D_\alpha F)\le 1-\theta$. Крім того, за умови $\sum_{\beta\in V_n}\hat{F}(\beta)^4\le\mu^2<1$ для будь-якого $\varepsilon>0$ існує не менше ніж $2^{l_0-s}\left(1-\left(\frac{\mu}{2\varepsilon}\right)^2\right)$ векторів $\alpha\in L$, кожен з яких задовольняє нерівності $wt(D_\alpha F)\le 1/2\cdot(1-\theta^2)+\varepsilon$.

Доведення. Справедливість першої частини твердження випливає з теореми 2 в [9]. Для доведення другої частини розглянемо зазначену вище випадкову величину ξ_L. З п. 2) леми 1, умови $d(F,B_n(H))\le 1/2\cdot(1-\theta)$ та нижньої межі (8) випливає, що $E\xi_L\le 1/2\cdot(1-\theta^2)$. Звідси, використовуючи п. 4) леми 1 отримаємо, що при випадковому рівноймовірному виборі вектора α з підпростору L справедливі такі співвідношення:

$$P\{wt(D_\alpha F)>1/2\cdot(1-\theta^2)+\varepsilon\}=$$
$$=P\{\xi_L-E\xi_L>1/2\cdot(1-\theta^2)-E\xi_L+\varepsilon\}\le$$
$$\le P\{\xi_L-E\xi_L>\varepsilon\}\le\left(\frac{\mu}{2\varepsilon}\right)^2.$$

Розглянемо множину $D=\{\alpha\in L: wt(D_\alpha F)\le 1/2\cdot(1-\theta^2)+\varepsilon\}$. На підставі вищевикладеного отримаємо, що

$$|D|=2^{\dim L}(1-P\{wt(D_\alpha F)>$$
$$>1/2\cdot(1-\theta^2)+\varepsilon\})\ge 2^{l_0-s}\left(1-\left(\frac{\mu}{2\varepsilon}\right)^2\right).$$

Отже, твердження повністю доведено.

Покладемо $\mathcal{D}(F,\varepsilon)=\{\alpha\in L_0: wt(D_\alpha F)\le\varepsilon\}$, де підпростір L_0 визначається за формулою (11), $\varepsilon\in(0,1)$. Для будь-яких $\theta,\mu\in(0,1)$ позначимо

$$w(\theta,\mu)=\min\{1-\theta,1/2\cdot(1-\theta^2)+\mu\} \quad (13)$$

Наслідок 1. Нехай за умови твердження 3 підпростір H містить підпростір H_0 вигляду (10). Тоді існує множина $D\subseteq\mathcal{D}(F,w(\theta,\mu))$ потужності $|D|\ge 3/4\cdot 2^{l_0-s}$ така, що $H\subseteq D^\perp$.

Отже, будь-який шуканий підпростір H складається с векторів, кожен з яких є ортогональним певній множині $D\subseteq\mathcal{D}(F,w(\theta,\mu))$ потужності не менше ніж $3/4\cdot 2^{l_0-s}$. Метод пошуку допустимих підпросторів, що пропонується, полягає в знаходженні зазначених векторів, побудові за ними матриці M вигляду (12) та оцінюванні відносної відстані між функцією F та множиною функцій вигляду (6).

Більш докладно, **метод, що пропонується, складається з таких кроків.**

1. Обчислити значення верхньої межі μ^2 параметра $\sum_{\beta\in V_n}\hat{F}(\beta)^4$.

2. Перевірити існування лінійно

ISSN 2074-9481
Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, вип. 1 (31), 2016р.

незалежних векторів $\alpha_1, \dots, \alpha_{l_0-s}$, що належать множині $\mathcal{D}(F, w(\theta,\mu))$, та побудувати ці вектори (у випадку позитивного результату перевірки).

3. Задати M як матрицю вигляду (12), стовпці якої утворюють базис векторного простору $<\alpha_1, \dots, \alpha_{l_0-s}>^{\perp}$.

4. Оцінити значення параметра $d(F, B_n(H))$ для підпростору H, що породжується стовпцями матриці M. Якщо $d(F, B_n(H)) \le 1/2 \cdot (1-\theta)$, вважати H шуканим підпростором; в протилежному випадку зробити висновок про відсутність θ-допустимих для функції F підпросторів вимірності $s+l_1$.

Опишемо зараз (ймовірнісні) алгоритми, які використовуються на кроках 1, 2 і 4 та з'ясуємо, за яких умов запропонований метод дозволяє розв'язувати поставлену задачу з заданою достовірністю.

Для побудови алгоритму оцінювання параметра $\sum_{\beta \in V_n} \hat{F}(\beta)^4$ скористаємося відомою формулою [10]

$$\sum_{\beta \in V_n} \hat{F}(\beta)^4 = 2^{-3n} \sum_{x,y,z \in V_n} (-1)^{F(x) \oplus F(x \oplus y) \oplus F(x \oplus z) \oplus F(x \oplus y \oplus z)} \qquad (14)$$

а також наступною лемою, яка неодноразово використовується далі.

Лема 2. [11]**.** Нехай $\zeta_1, \dots, \zeta_t$ є незалежними випадковими величинами, такими, що $\alpha_j \le \zeta_j \le \beta_j$, $\alpha_j, \beta_j \in R$, $j \in \overline{1,t}$. Тоді для будь-якого $x > 0$

$$P\left\{ t^{-1} \sum_{l=1}^{t} \zeta_l - E\left(t^{-1} \sum_{l=1}^{t} \zeta_l \right) \ge x \right\} \le$$

$$\le \exp\left\{ - \frac{2t^2 x^2}{\sum_{l=1}^{t} (\beta_l - \alpha_l)^2} \right\}.$$

Алгоритм 1.

Вхідні дані:

– функція $F \in B_n$, задана за допомогою оракула;

– числа $\varepsilon, \delta \in (0,1)$.

1. Покласти $t = \left\lceil 2\varepsilon^{-2} \ln(\delta^{-1}) \right\rceil$.

2. Згенерувати t незалежних випадкових елементів (X_i, Y_i, Z_i), де X_i, Y_i та Z_i є незалежними випадковими векторами з рівномірним розподілом ймовірностей на множині V_n, $i \in \overline{1,t}$.

3. Обчислити значення

$$\mu^2 = t^{-1} \sum_{i=1}^{t} (-1)^{F(X_i) \oplus F(X_i \oplus Y_i) \oplus F(X_i \oplus Z_i) \oplus F(X_i \oplus Y_i \oplus Z_i)} + \varepsilon.$$

Безпосередньо з опису алгоритму, формули (14) та леми 2 випливає такий результат.

Твердження 4. Для будь-яких $\varepsilon, \delta \in (0,1)$ справедлива нерівність $P\{ \sum_{\beta \in V_n} \hat{F}(\beta)^4 \le \mu^2 \} \ge 1 - \delta$. При цьому часова складність алгоритму 1 дорівнює $t = O(\varepsilon^{-2} \ln \delta^{-1})$.

Опишемо зараз алгоритм, який виконується на другому кроці запропонованого методу.

Алгоритм 2.

Вхідні дані:

– функція $F \in B_n$ задана за допомогою оракула ($n = l_0 + l_1$);

– числа $s \in \overline{1, l_0 - 3}$, $\theta \in (0,1)$, $\mu \in (0,1)$, $\varepsilon \in (0, \theta)$, $\delta \in (0,1)$.

1. Обчислити $w(\theta,\mu)$ за формулою (13); покласти $t = \left\lceil 3^{-1} 2^{s+3} \ln(2\delta^{-1}) \right\rceil$, $l = \left\lceil 2^{-1} \varepsilon^{-2} \ln(2^{-s+1} t \delta^{-1}) \right\rceil$.

2. Покласти α рівним нульовому вектору довжини l_0;

3. Для кожного $i \in \overline{1,t}$:

– згенерувати випадковий вектор ξ_i з рівномірним розподілом на множині $L_0 \setminus \{0\}$, де L_0 визначається за формулою (11);

– згенерувати незалежні випадкові вектори $X_{i,1}, \dots, X_{i,l}$ з рівномірним розподілом на множині V_n та обчислити значення $\Delta_l(\xi_i) = l^{-1} \sum_{j=1}^{l} D_{\xi_i} F(X_{i,j})$;

– якщо $\Delta_l(\xi_i) \le w(\theta,\mu) + \varepsilon$, покласти $\alpha = \alpha_i$ та закінчити роботу.

Результатом виконання алгоритму 2 є ненульовий вектор α, що дорівнює значенню ξ_i з найменшим номером $i \in \overline{1,t}$, для якого $\Delta_l(\xi_i) \le w(\theta,\mu)+\varepsilon$, або нульовий вектор, якщо таких i не знайшлося.

Покажемо, що наведений алгоритм дозволяє надійно знаходити вектори з множини $\mathcal{D}(F, w(\theta,\mu))\setminus\{0\}$ за умови існування θ-допустимих підпросторів для функції F.

Твердження 5. Нехай існує θ-допустимий для функції F підпростір H вимірності $s+l_1$. Тоді випадковий вектор α, отриманий за допомогою алгоритму 2, належить множині $\mathcal{D}(F, w(\theta,\mu))\setminus\{0\}$ з ймовірністю не менше ніж $1-\delta$. Крім того, часова складність алгоритму 2 дорівнює

$$lt = O(2^s \varepsilon^{-2} \ln(\delta^{-1}) \ln(\delta^{-1} \ln \delta^{-1})). \qquad (15)$$

Доведення. На підставі наслідку 1 множина $\mathcal{D}(F, w(\theta,\mu))$ містить множину D потужності $\lceil 3/4 \cdot 2^{l_0-s} \rceil$. При цьому, якщо $\alpha \notin \mathcal{D}(F, w(\theta,\mu))\setminus\{0\}$, то відбувається одна з двох подій: або жоден з випадкових векторів $\xi_1, \dots, \xi_t$ не потрапляє до множини D, або існує $i \in \overline{1,t}$ таке, що $\xi_i \in D$ і $\Delta_l(\xi_i) > w(\theta,\mu)+\varepsilon$. Отже, $P\{\alpha \notin \mathcal{D}(F, w(\theta,\mu))\setminus\{0\}\} \le$ $\le p_1 + p_2$, де

$$p_1 = P\{\xi_1, \dots, \xi_t \notin D\},$$

$$p_2 = P\left\{\bigcup_{i=1}^{t} \{\xi_i \in D, \Delta_l(\xi_i) > w(\theta,\mu)+\varepsilon\}\right\}.$$

В силу означення випадкових векторів $\xi_1, \dots, \xi_t$ і параметра t, а також умови $s \in \overline{1, l_0-3}$ справедливі такі співвідношення:

$$p_1 = (1-\mathbf{P}\{\xi_1 \in D\})^t \le$$

$$\le \left(1 - \frac{|D|-1}{2^{l_0}-1}\right)^t \le \left(1 - \frac{3/4 \cdot 2^{l_0-s}-1}{2^{l_0}-1}\right)^t \le$$

$$\le (1-3\cdot 2^{-s-3})^t \le \exp\{-3\cdot 2^{-s-3} t\} \le \delta/2.$$

Далі, використовуючи незалежність випадкових векторів ξ_i, $X_{i,j}$ ($i \in \overline{1,t}$, $j \in \overline{1,l}$), лему 2, означення параметра l та умову $s \in \overline{1, l_0-3}$ отримаємо, що

$$p_2 \le \sum_{i=1}^{t} P\{\xi_i \in D, \Delta_l(\xi_i) > w(\theta,\mu)+\varepsilon\} \le$$

$$\le \sum_{i=1}^{t} \sum_{a \in D\setminus\{0\}} P\{\xi_i = a, \Delta_l(a) > w(\theta,\mu)+\varepsilon\} =$$

$$= t \sum_{a \in D\setminus\{0\}} \frac{1}{2^{l_0}-1} P\{\Delta_l(a) > w(\theta,\mu)+\varepsilon\} \le$$

$$\le \frac{3/4 \cdot 2^{l_0-s}+1}{2^{l_0}-1} t \exp\{-2l\varepsilon^2\} \le$$

$$\le 2^{-s} t \exp\{-2l\varepsilon^2\} \le \delta/2.$$

Таким чином, на підставі отриманих співвідношень $P\{\alpha \in \mathcal{D}(F, w(\theta,\mu))\setminus\{0\}\} \ge$ $\ge 1-(p_1+p_2) \ge 1-\delta$, що і треба було довести. Нарешті, формула (15) випливає безпосередньо з опису алгоритму 2.

Твердження доведено.

Наслідок 2. Нехай $\alpha_1, \dots, \alpha_r$ є випадковими векторами, отриманими в результаті r-кратного застосування алгоритму 2 до вхідних даних F, s, θ, μ, ε, $\delta = r^{-1}\delta'$, де $\delta' \in (0,1)$. Тоді за умови твердження 5 справедлива нерівність $P\{\alpha_1, \dots, \alpha_r \in \mathcal{D}(F, w(\theta,\mu))\setminus\{0\}\} \ge 1-\delta'$.

Опишемо, нарешті, алгоритми оцінювання параметра $d(F, B_n(H))$ на кроці 4 запропонованого методу. Перший алгоритм базується на рівності (7) і дозволяє отримувати двосторонні статистичні оцінки зазначеного параметра. Другий алгоритм базується на верхній межі (8) і дозволяє отримувати тільки верхні оцінки значення $d(F, B_n(H))$, проте вимагає значно менше часу обчислень.

Алгоритм 3.

Вхідні дані:

– функція $F \in B_n$, задана за допомогою оракула ($n = l_0 + l_1$);

– матриця $M \in F_2^{n\times m}$, стовпці якої утворюють базис підпростору H ($m = s + l_1$, $s \in \overline{1, l_0-3}$);

– числа $\varepsilon, \delta \in (0,1)$.

1. Покласти $l = \lceil 2\varepsilon^{-2} \ln(4\delta^{-1}) \rceil$,

ISSN 2074-9481
Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, вип. 1 (31), 2016р.

$t = \left\lceil 8\varepsilon^{-2}\ln(4l\delta^{-1})\right\rceil$.

2. Згенерувати незалежні випадкові вектори $\xi_1, \xi_2, \ldots, \xi_l$ з рівномірним розподілом на множині V_m; для кожного $i \in \overline{1,l}$ згенерувати незалежні випадкові вектори $\eta_{i1}, \ldots, \eta_{it}$ з рівномірним розподілом на множині $L_{\xi_i} = \{x \in V_n : xM = \xi_i\}$ та обчислити значення

$$\lambda_{l,t} = l^{-1}\sum_{i=1}^{l}\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})}\right|.$$

Твердження 6. Для будь-яких $\varepsilon, \delta \in (0,1)$ справедлива нерівність

$$P\{|\, d_F(B_{n,m}(H)) - 1/2\cdot(1-\lambda_{l,t})\,| \ge \varepsilon\} \le \delta. \qquad (16)$$

При цьому часова складність алгоритму 3 дорівнює $lt = 16\cdot C\varepsilon^{-4}\ln(4\delta^{-1})\ln(8\varepsilon^{-2}\delta^{-1}\ln(4\delta^{-1}))$, де $C = \text{const} \ge 1$.

Доведення. Помітимо, що

$$|\lambda_{l,t} - l_F(H)| \le$$

$$\le \left| l^{-1}\sum_{i=1}^{l}\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})}\right| - l^{-1}\sum_{i=1}^{l}\left| 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right|\right| +$$

$$+ \left| l^{-1}\sum_{i=1}^{l}\left| 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| - 2^{-m}\sum_{y\in V_m}\left| 2^{-(n-m)}\sum_{x\in L_y}(-1)^{F(x)}\right|\right| =$$

$$= \nu_1 + \nu_2.$$

Отже,

$$P\{|\, d_F(B_{n,m}(H)) - 1/2\cdot(1-\lambda_{l,t})\,| \ge \varepsilon\} =$$
$$= P\{|\,\lambda_{l,t} - l_F(H)\,| \ge \varepsilon\} \le \qquad (17)$$
$$\le P\{\nu_1 \ge \varepsilon/2\} + P\{\nu_2 \ge \varepsilon/2\}.$$

Оцінимо зверху другий доданок у правій частині нерівності (17). Позначимо $\varsigma_i = \left| 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right|$, $i \in \overline{1,l}$. Випадкові величини $\varsigma_1, \varsigma_2, \ldots, \varsigma_l$ є незалежними, однаково розподіленими та приймають значення в проміжку $[0,1]$. При цьому $E\varsigma_i = 2^{-m}\sum_{y\in V_m}\left| 2^{-(n-m)}\sum_{x\in L_y}(-1)^{F(x)}\right| = l_F(H), i \in \overline{1,l}$ і $\nu_2 = \left| l^{-1}\sum_{i=1}^{l}\varsigma_i - l^{-1}\sum_{i=1}^{l}E\varsigma_i\right|$. Отже, на підставі леми 2 та означення параметра l справедлива нерівність

$$P\{\nu_2 \ge \varepsilon/2\} \le 2\exp\{-2l(\varepsilon/2)^2\} \le \delta/2 \qquad (18)$$

Оцінимо зараз перший доданок у правій частині нерівності (17). Помітимо, що

$$P\{\nu_1 \ge \varepsilon/2\} =$$

$$= P\left\{\left| l^{-1}\sum_{i=1}^{l}\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})}\right| - l^{-1}\sum_{i=1}^{l}\left| 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right|\right| \ge \right.$$
$$\left. \ge \varepsilon/2\right\} \le$$

$$\le P\left\{\left| l^{-1}\sum_{i=1}^{l}\left(t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right)\right| \ge \right.$$
$$\left. \ge \varepsilon/2\right\} \le$$

$$\le P\left\{\max_{1\le i\le l}\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| \ge \right.$$
$$\left. \ge \varepsilon/2\right\} \le$$

$$\le P\left\{\bigcup_{i=1}^{l}\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| \ge \right\}\right.$$
$$\left. \ge \varepsilon/2\right\} \le$$

$$\le \sum_{i=1}^{l}P\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| \ge \right.$$
$$\left. \ge \varepsilon/2\right\}.$$

Далі, згідно з а формулою повної ймовірності,

$$P\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| \ge \varepsilon/2\right\} =$$

$$= 2^{-m}\sum_{y\in V_m}P\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_{ij})} - 2^{-(n-m)}\sum_{x\in L_{\xi_i}}(-1)^{F(x)}\right| \ge \right.$$
$$\left. \ge \varepsilon/2\,\middle|\,\xi_i = y\right\} =$$

$$= 2^{-m}\sum_{y\in V_m}\mathbf{P}\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_j)} - 2^{-(n-m)}\sum_{x\in L_y}(-1)^{F(x)}\right| \ge \right.$$
$$\left. \ge \varepsilon/2\right\},$$

де $\eta_1,...,\eta_t$ є незалежними випадковими величинами з рівномірним розподілом ймовірностей на множині L_y. Отже, на підставі леми 2 маємо

$$P\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_j)} - 2^{-(n-m)}\sum_{x\in L_y}(-1)^{F(x)}\right| \ge \varepsilon/2\right\} =$$

$$= P\left\{\left| t^{-1}\sum_{j=1}^{t}(-1)^{F(\eta_j)} - t^{-1}\sum_{j=1}^{t}E(-1)^{F(\eta_j)}\right| \ge \varepsilon/2\right\} \le$$

$$\le 2\exp\{-1/2\cdot t(\varepsilon/2)^2\},$$

звідки в силу означення параметра t випливає, що

$$P\{\nu_1 \ge \varepsilon/2\} = 2l\exp\{-1/2\cdot t(\varepsilon/2)^2\} \le \delta/2. \quad (19)$$

Таким чином, на підставі формул (17) – (19) справедлива нерівність (16). Твердження доведено.

Опишемо другий алгоритм статистичного оцінювання верхніх меж параметра $d(F, B_n(H))$. Позначимо M' $(n-m)\times n$-матрицю, рядки якої утворюють базис дуального до H підпростору $H^{\perp}$. На підставі формул (4), (8) справедлива нерівність

$$d(F, B_n(H)) \le$$

$$\le 2^{-(n-m)}\ 2^{-n}\sum_{(\alpha,\,x)\,\in\, H^{\perp}\times V_n}(F(x\oplus\alpha)\oplus F(x)). \quad (20)$$

Алгоритм 4.

Вхідні дані:

– функція $F\in B_n$, задана за допомогою оракула ($n = l_0 + l_1$);

– $(n-m)\times n$-матриця M', рядки якої утворюють базис підпростору $H^{\perp}$ ($m = s + l_1$, $s\in\overline{1, l_0-3}$);

– числа $\varepsilon,\delta\in(0,1)$.

1. Покласти $t = \left\lceil 2^{-1}\varepsilon^{-2}\ln(\delta^{-1})\right\rceil$.

2. Згенерувати t незалежних випадкових елементів (X_i, Y_i), де X_i та Y_i є незалежними випадковими векторами з рівномірними розподілами ймовірностей на множинах V_n та V_{n-m} відповідно, $i\in\overline{1,t}$.

3. Обчислити значення $\Delta_t = t^{-1}\sum_{i=1}^{t}(F(X_i\oplus Y_iM')\oplus F(X_i))$.

Безпосередньо з опису алгоритму 4, нерівності (20) та леми 2 випливає такий результат.

Твердження 7. Для будь-яких $\varepsilon,\delta\in(0,1)$ справедлива нерівність $P\{d_F(B_{n,m}(H))\le\Delta_t+\varepsilon\}\ge 1-\delta$. При цьому часова складність алгоритму 4 дорівнює $t = O(\varepsilon^{-2}\ln\delta^{-1})$.

Зауважимо, що загальна трудомісткість описаного методу пошуку θ-допустимих для функції F підпросторів суттєво залежить від кількості випадкових векторів, які формуються за допомогою алгоритму 2, і може бути надто великою, якщо множина $\mathcal{D}(F, w(\theta,\mu))$ має велику потужність. Поряд з тим, якщо результатом алгоритму 2 є нульовий вектор, то на підставі твердження 5 функція F не має $(s+l_1)$-вимірних θ-допустимих наближень з достовірністю не менше ніж $1-\delta$. В цьому випадку трудомісткість методу визначається за формулою (15).

Другий етап: визначення функції ϕ

Нехай $F\in B_n$ – булева функція, задана за допомогою оракула, $M\in F_2^{n\times m}$ – матриця рангу $m\in\overline{1,n-1}$. Треба розробити алгоритм, який дозволяє ефективно обчислювати значення функції $\phi\in B_m$, для якої відносна відстань між функціями F та $g(x) = \phi(xM)$, $x\in V_n$ є, за можливістю, мінімальною.

Позначимо H підпростір векторного простору V_n, породжений стовпцями матриці M. В [1] запропоновано поліноміальний ймовірнісний алгоритм розв'язання поставленої задачі у випадку, коли $d(F, B_n(H))\le 1/2\cdot(1-\theta)$, $\theta\in(0,1)$. А саме, для будь-якого $\tau\in(0, 4\theta)$ цей алгоритм обчислює значення випадкової функції $\psi\in B_m$, для якої відносна відстань між функціями F та $\psi(xM)$, $x\in V_n$ не перевищує $1/2\cdot(1-\theta)+\tau$ з ймовірністю не менше $1-\exp\{-2^{m-1}\tau^2\}$, використовуючи $T_{\psi} = O(T_F m(n-m)\tau^{-2}\ln\tau^{-1})$ операцій, де T_F

ISSN 2074-9481
Правове, нормативне та метрологічне забезпечення системи захисту інформації в Україні, вип. 1 (31), 2016р.

– часова складність обчислення значення функції F. Якщо $\tau = \theta/4 \geq \left(2^{-(m-1)} \ln \delta^{-1}\right)^{1/2}$, де δ – верхня межа ймовірності помилки алгоритму, то його складність дорівнює $T_\psi = O(T_F m(n-m)\theta^{-2} \ln \theta^{-1})$.

Нижче пропонується інший алгоритм, який не гарантує такої високої якості наближень, як алгоритм з [1], але має суттєво меншу часову складність. Цей алгоритм базується на розвитку ідеї [2] використання в ролі наближень функції F її підфункцій.

Твердження 8. Нехай $F \in B_n$, $M \in F_2^{n\times m}$, $\operatorname{rank}(M) = m \in \overline{1, n-1}$, H – підпростір, породжений стовпцями матриці M і U – оборотна $n\times n$-матриця така, що $UM = \begin{pmatrix} I_m \\ 0 \end{pmatrix}$. Для будь-якого $z \in V_{n-m}$ покладемо

$$\phi_z(y) = F((y, z)U), y \in V_m \qquad (21)$$

Тоді при випадковому рівноймовірному виборі вектора z середнє значення відносної відстані між функціями F та $g_z(x) = \phi_z(xM)$, $y \in V_m$ не перевищує $1-\omega_F(H)$, де $\omega_F(H)$ визначається за формулою (3). Зокрема, за умови $\omega_F(H) \geq \theta$ справедлива нерівність $E_z d(F, g_z) \leq 1-\theta$.

Доведення. Згідно з означенням функції g_z

$$E_z d(F, g_z) = 2^{-(n-m)} \sum_{z\in V_{n-m}} 2^{-n} \sum_{x\in V_n} (F((xM, z)U) \oplus F(x)) =$$

$$= 2^{-(n-m)} \sum_{z\in V_{n-m}} 2^{-n} \sum_{x\in V_n} (F((xUM, z)U) \oplus F(xU)) =$$

$$= 2^{-(n-m)} \sum_{z\in V_{n-m}} 2^{-n} \sum_{x\in V_n} (F((x\begin{pmatrix} I_m \\ 0 \end{pmatrix}, z)U) \oplus F(xU)) =$$

$$= 2^{-(n-m)} \sum_{z\in V_{n-m}} 2^{-n} \sum_{\substack{x_1\in V_m,\\ x_2\in V_{n-m}}} (F((x_1, z)U) \oplus F((x_1, x_2)U)) \leq$$

$$\leq a + b,$$

де

$$a = 2^{-n} 2^{-(n-m)} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} (F((x_1, z)U) \oplus F((x_1, z\oplus x_2)U)),$$

$$b = 2^{-n} 2^{-(n-m)} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} (F((x_1, z\oplus x_2)U) \oplus F((x_1, x_2)U)).$$

Далі,

$$a = 2^{-n} 2^{-(n-m)} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} (F((x_1, z)U) \oplus F((x_1, z)U \oplus (0, x_2)U)) =$$

$$= 2^{-(n-m)} 2^{-n} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} D_{(0,x_2)U} F((x_1, z)U) = 2^{-(n-m)} \sum_{x_2\in V_{n-m}} wt(D_{(0,x_2)U} F)$$

$$b = 2^{-n} 2^{-(n-m)} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} (F((x_1, x_2)U) \oplus F((x_1, x_2)U \oplus (0, z)U)) =$$

$$= 2^{-(n-m)} 2^{-n} \sum_{\substack{z\in V_{n-m},\\ x_1\in V_m, x_2\in V_{n-m}}} D_{(0,z)U} F((x_1, x_2)U) = 2^{-(n-m)} \sum_{z\in V_{n-m}} wt(D_{(0,z)U} F).$$

Отже,

$$E_z d(F, g_z) \leq 2 \cdot 2^{-(n-m)} \sum_{x_2\in V_{n-m}} wt(DF_{(0,x_2)U}) =$$

$$= 2 \cdot 2^{-(n-m)} \sum_{\alpha\in H^\perp} wt(D_\alpha F),$$

де остання рівність випливає з означень підпростору H та матриці U. Нарешті, використовуючи формулу (4), отримаємо, що $E_z d(F, g_z) \leq 1-\omega_F(H)$.

Твердження доведено.

Отже, алгоритм обчислення значень функції $\phi \in B_m$ має такий вигляд.

Алгоритм 5.

Вхідні дані:

– функція $F \in B_n$, задана за допомогою оракула ($n = l_0 + l_1$);

– матриця $M \in F_2^{n\times m}$ вигляду (12), де $\operatorname{rank}(M_0) = s$; оборотна матриця $U \in F_2^{n\times n}$ така, що $UM = \begin{pmatrix} I_m \\ 0 \end{pmatrix}$ ($m = s + l_1$, $s \in \overline{1, l_0 - 3}$);

1. Згенерувати випадковий рівноймовірний вектор $z \in V_{n-m}$.

2. Для будь-якого $y \in V_m$ покласти $\phi(y) = F((y, z)U)$, $y \in V_m$.

З твердження 8 випливає, що середня (за всіма $z \in V_{n-m}$) відносна відстань між функцію F та її наближенням вигляду $\phi(xM)$, $x \in V_n$ не перевищує $1 - \sum_{y\in V_m} \hat{F}(My)^2$. При цьому часова складність обчислення значення функції ϕ є $T_\phi = O(T_F n^2)$, де T_F – складність обчислення значення функції F.

Приклад застосування запропонованого методу

Застосуємо отримані результати до

побудови статистичної атаки на редуковану версію шифру Grain-128 [14]. Нагадаємо, що цей СПШ є одним з трьох потокових шифрів, орієнтованих на апаратну реалізацію, що рекомендовані для застосувань за підсумками дослідного проекту eSTREAM. Шифр має довжину ключа $l_0 = 128$ біт і довжину вектора ініціалізації $l_1 = 96$ біт. При цьому алгоритм формування початкового стану генератора гами шифру являє собою ітераційну процедуру, яка складається з 256 однотипних кроків (раундів).

У [2] (приклад 9) описано атаку на редуковану версію Grain-128, де замість 256 раундів алгоритму формування початкового стану виконується тільки 180. Для побудови цієї атаки використовується певна функція $F = F_{\mathrm{FKM}}(k, c)$, $k = (k_0, \dots, k_{127}) \in V_{128}$, $c = (c_0, \dots, c_{95}) \in V_{96}$, а також її наближення, яке отримується шляхом фіксації певних її 18 змінних k_i, $i \in \overline{0,127}$, нульовими значеннями. Згідно з [2], складність відповідної атаки на шифр дорівнює 2^{124}.

Для побудови більш ефективної атаки знайдемо за допомогою запропонованого методу θ-допустимий для функції F підпростір вимірності $m = s + l_1$ при $s = 52$, $\theta = 0,57$.

Перш за все, обчислимо значення верхньої межі μ^2 параметра $\sum_{\beta \in V_n} \hat{F}(\beta)^4$, застосовуючи десять разів до функції F алгоритм 1 при $\varepsilon = 0,01$, $\delta = 0,1$ (табл. 1). Середнє арифметичне значення μ^2 за десятьма проведеними експериментами дорівнює $0,5105$. Час однократного виконання алгоритму 1 складає 43 хвилини.

Таблиця 1

Результати виконання алгоритму 1

№ експерименту	μ^2
1	0,5119
2	0,5096
3	0,5060
4	0,5087
5	0,5107
6	0,5132
7	0,5093
8	0,5117
9	0,5105
10	0,5136

Далі застосуємо $r = 600$ разів алгоритм 2 до вхідних даних F, $s = 52$, $\theta = 0,57$, $\mu = \sqrt{0,5105} = 0,7145$, $\varepsilon = 0,05$, $\delta = r^{-1}\delta'$, де $\delta' = 0,1$. В результаті отримаємо список випадкових векторів $\alpha_1, \dots, \alpha_r$, які задовольняють умові $P\{\alpha_1, \dots, \alpha_r \in \mathcal{D}(F, w(\theta, \mu)) \setminus \{0\}\} \geq 1 - \delta'$ (див. наслідок 2). Параметр $w(\theta, \mu)$ в цьому випадку складає $0,43$. В табл. 2 показано десять отриманих векторів та відповідні їм значення параметра $\Delta_l(\alpha)$ (див. опис алгоритму 2). Середній час однократного застосування алгоритму 2 складає приблизно 2 хвилини.

Таблиця 2

Перші десять векторів,отриманих за допомогою алгоритму 2

№ вектора α у списку	Шістнадцяткове значення α	$\Delta_l(\alpha)$	$\Delta_l(\alpha)+\varepsilon$
1	d2a380839522b94c506001c1351c7241	0,2170	0,2670
2	691e3c0a3bc85309cbb891213c367e4d	0,2196	0,2696
3	d7c4a397c0bc0400b0429993113a7b80	0,2410	0,2910
4	2925dec4f8ef5a593e9dfdc140c4c868	0,2178	0,2678
5	fe0a5f9c09e2c7ed255dca3a19eeda70	0,2479	0,2979
6	0a47ac5858e0f8e87e5a139a572477e8	0,2226	0,2726
7	8826bd9d0ca203826729064b5872056c	0,2068	0,2568
8	f4954e03dfd98b0c3026e332635ef7c8	0,2252	0,2752
9	c8afe96cc49e3d28542e6fc909c43378	0,2375	0,2875
10	f81a2ab7d2de82c4a85420710dc4dc4d	0,2192	0,2692

Далі, використовуючи перші $l_0 - s = 76$ векторів з отриманого списку, побудуємо матрицю M вигляду (12) як зазначено на кроці 3 запропонованого методу. На рис. 1 показано (у транспонованому вигляді) підматрицю M_0 отриманої матриці M.

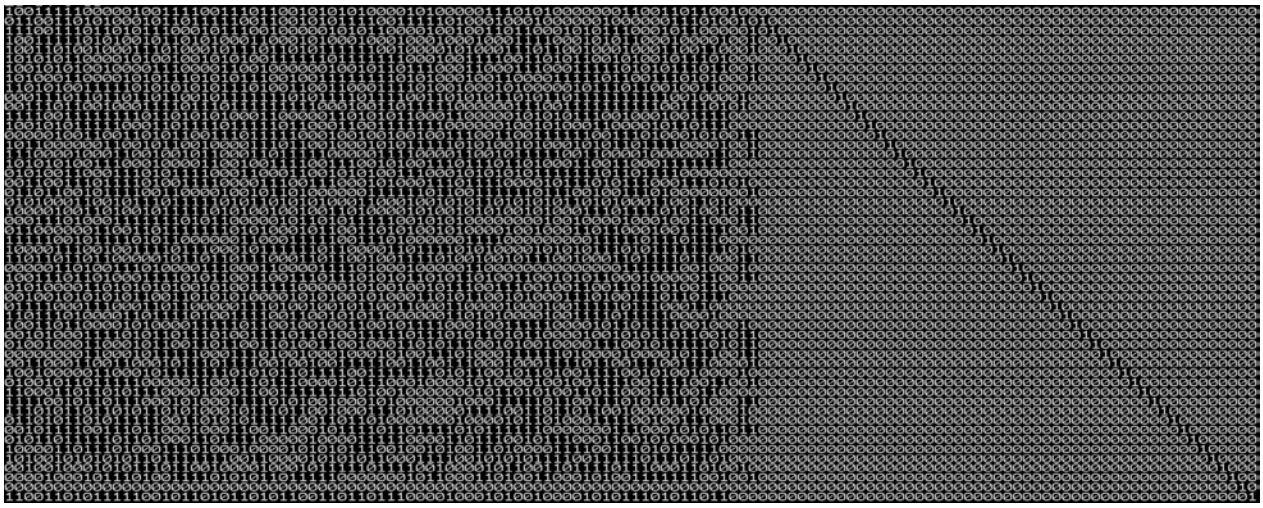

Рис. 1. Матриця, транспонована до M_0

Оцінимо тепер значення параметра $d(F, B_n(H))$, де H є підпростором, породженим стовпцями матриці M, застосовуючи алгоритми 3 і 4 до вхідних даних F, M, $\varepsilon = 0,2$, $\delta = 0,1$ (табл. 3).

Середні арифметичні значення оцінок параметра $d(F, B_n(H))$, отримані за допомогою алгоритмів 3 та 4 дорівнюють $0,3497$ та $0,4307$ відповідно. Звідси на підставі описів зазначених алгоритмів і тверджень 6 та 7 випливає, що $d(F, B_n(H)) \le 0,3497$ з достовірністю не менше ніж $1 - \delta/2 - 2(\delta/4)^{10} \ge 0,946$ і $d(F, B_n(H)) \le 0,4307$ з достовірністю не менше ніж $1 - \delta^{10} = 1 - 10^{-10}$.

Таблиця 3

Результати виконання алгоритмів 3 і 4

№ експерименту	Відносна відстань між функцією F та множиною $B_n(H)$					
	Алгоритм 3			Алгоритм 4		
	d	$d+\varepsilon$	T, хв	d	$d+\varepsilon$	T, с
1	0,1557	0,3557	150	0,1724	0,3724	2
2	0,1486	0,3486	150	0,2759	0,4759	2
3	0,1665	0,3665	150	0,1379	0,3379	2
4	0,1583	0,3583	150	0,2414	0,4414	2
5	0,1505	0,3505	150	0,2414	0,4414	2
6	0,1577	0,3577	150	0,3448	0,5448	2
7	0,1358	0,3358	150	0,1379	0,3379	2
8	0,1477	0,3477	150	0,2069	0,4036	2
9	0,1369	0,3369	150	0,3793	0,5793	2
10	0,1396	0,3396	150	0,1724	0,3724	2

Нарешті, скористаємося алгоритмом 5 для побудови наближення $g(x)=\phi(xM)$, $x\in V_n$ функції F. В табл. 4 (з правого боку) показано результати оцінювання відносної відстані між функціями F та g з точністю $\varepsilon=0,005$ і достовірністю $1-\delta=0,9$. Для порівняння (з лівого боку) наведено верхні оцінки параметра $d(F,B_n(H))$, отримані за допомогою алгоритму 4 при тих самих значеннях ε і δ.

Таблиця 4

Статистичні оцінки параметра $d(F,B_n(H))$

№ експерименту	Відносна відстань між функцією F та множиною $B_n(H)$ (алгоритм 4)			Відносна відстань між функцією F та її запропонованим наближенням g		
	d	$d+\varepsilon$	T, хв.	d	$d+\varepsilon$	T, хв.
1	0,2367	0,2417	41	0,1876	0,1926	40
2	0,2391	0,2441	41	0,1652	0,1702	40
3	0,2371	0,2421	41	0,2103	0,2153	40
4	0,2384	0,2434	41	0,3286	0,3336	40
5	0,2344	0,2394	41	0,2056	0,2106	40
6	0,2395	0,2445	41	0,2762	0,2812	40
7	0,2374	0,2424	41	0,2634	0,2684	40
8	0,2350	0,2400	41	0,1844	0,1894	40
9	0,2355	0,2405	41	0,1786	0,1836	40
10	0,2348	0,2398	41	0,1779	0,1829	40

Результати табл. 4 дозволяють отримати більш точні (порівняно з табл. 3) оцінки параметра $d_F(B_{n,m}(H))$. Так, середнє арифметичне значення оцінок цього параметра, обчислених за допомогою алгоритму 4, складає $0,2418$, а найменше (з десяти отриманих) значення параметра $d(F,g)$ дорівнює $0,1702$. Таким чином, $d(F,g)\le 0,1702$ з достовірністю не менше ніж $1-\delta=0,9$. При цьому двійкове значення вектора z, який генерується на першому кроці алгоритму 5 та залишається незмінним протягом обчислень значень функції наближення, є таким:

11010111011111100000001110010101111110000000110001100100100010101100100101 00

Відзначимо, що алгоритми 1 – 5 реалізовано на платформі .NET Framework 4.0 (C#). Для виконання обчислень використано ЕОМ на базі 64-розрядної операційної системи Windows 7 Service Pack 1 з процесором Intel(R) Core(TM) i5-2300 CPU @ 2.80GHz та оперативною пам'яттю обсягом 4,00 ГБ. Використовуючи отримане наближення функції F можна побудувати узагальнену статистичну атаку на редуковану версію шифру Grain-128 [1]. Трудомісткість цієї атаки визначається за формулою (6) в [1]:

$$T(l_0,s,r)=O\left(2^s r(T_F+T_\phi)+s(l_0-s)2^{l_0-s}T_{\mathcal{A}}\right),$$

де T_F, T_ϕ і $T_{\mathcal{A}}$ є, відповідно, часові складності обчислення значень функцій F, ϕ і алгоритма $\mathcal{A}$ опробування ключів на другому етапі атаки, $r=\left\lceil 2^{2\nu+3}\ln(2^s\delta^{-1})\right\rceil$, де число $\nu>1$ визначається з нерівності $d(f,g)\le 1/2-2^{-\nu}$. Вважаючи $T_F=128$, $T_\phi=T_F n^2$, $T_{\mathcal{A}}=3\cdot 128$, та використовуючи оцінку $d(F,g)\le 0,1791$, отримаємо, що (за умови теореми 1 в [1]) трудомісткість узагальненої статистичної атаки на зазначену версію шифру Grain-128 не перевищує 2^{97}, що є в 2^{27} разів менше ніж трудомісткість раніше відомої атаки [2].

Висновки

Основним результатом статті є метод знаходження (або обгрунтування відсутності) певних алгебраїчно вироджених наближень булевих функцій, заданих за допомогою оракулів. Наявність таких наближень є необхідною умовою застосовності узагальненої статистичної атаки на СПШ [1], в той час як їх відсутність гарантує практичну стійкість СПШ відносно цієї атаки.

Запропонований метод відрізняється за сутністю від раніше відомих [2], [4] – [7] та базується на отриманих необхідних умовах існування шуканих наближень. Метод складається з двох етапів. На першому з яких здійснюється пошук певних підпросторів, що є допустимими для вхідної функції-оракула (при цьому відсутність зазначених просторів свідчить про відсутність відповідних наближень). На другому етапі за знайденим підпростором будується булева функція від меншої кількості змінних, яка (поряд з базисом знайденого підпростору) задає наближення вхідної функції.

Для розв'язання окремих задач на кожному етапі методу запропоновано поліноміальні ймовірнісні алгоритми, які можуть бути застосовані на практиці до функцій-оракулів від декількох десятків чи сотень змінних.

Застосування отриманих результатів до редукованої версії СПШ Grain-128 показує, що (за умови теореми 1 в [1]) складність узагальненої статистичної атаки на шифр не перевищує 2^{97}, в той час як складність раніше відомої атаки [2] дорівнює 2^{124}.

Перелік посилань

[1] А. Н. Алексейчук, С. Н. Конюшок, А. Ю. Сторожук *Обобщенная статистическая атака на синхронные поточные шифры* // Захист інформації.– 2015. – Т. 17. – № 3. – С. 54–65.

[2] S. Fischer, S. Khazaei, W. Meier, *Chosen IV statistical analysis for key recovery attacks on stream ciphers* // AFRICACRYPT 2008. – Proceedings. – Springer-Verlag. – 2008. P. 236–245.

[3] I. Dinur, A. Shamir *Cube attacks on tweakable black box polynomials* // Advances in Cryptology – EUROCRYPT'09. – Proceedings. – Springer-Verlag. – 2009. – P.278–299.

[4] P. Gopalan, R. O'Donnel, A. Servedio, A. Shpilka, K. Wimmer *Testing Fourier dimensionality and sparsity* // SIAM J. on Computing. – 2011. – Vol. 40(4). – P. 1075–1100.

[5] P. A. Gopalan *Fourier-analytic approach to Reed-Muller decoding* // Annual IEEE Symp. on Foundation in Computer Science. – FOCS 2010. – Proceedings. – Springer-Verlag. – 2010. – P. 685–694.

[6] А. М. Олексійчук, С. М. Конюшок, А. Ю. Сторожук *Швидкі алгоритми побудови k-вимірних наближень наближень булевих функцій* // Захист інформації.– 2015. – Т. 17. – № 1. – С. 43–52.

[7] Е. К. Алексеев *О некоторых мерах нелинейности булевых функций* // Прикладная дискретная математика. – 2011. – № 2(12). – С. 5–16.

[8] О. А. Логачев , А. А. Сальников , В. В. Ященко *Булевы функции в теории кодирования и криптологии.* – М. : МЦНМО, 2004. – 470 с.

[9] А. Н. Алексейчук, С. Н. Конюшок *Алгебраически вырожденные приближения булевых функций* // Кибернетика и системный анализ.– 2014. – Т. 50. – № 6. – С. 3–14.

[10] T. Gowers *A new proof of Szemere'di's theorem* // Geom Funct. Anal. – 2001. – Vol. 11(3). – P. 465–588.

[11] W. Hoeffding *Probability inequalities for sums of bounded random variables* // J. Amer. Statist. Assoc. – 1963. – Vol. 58. – № 301. – P. 13–30.

[12] M. Hell, T. Johansson, A. Maximov, W. Meier *A stream cipher proposal: Grain-128* // eStream: the ECRYPT Stream Cipher Project. http://www.ecrypt.eu.org/stream/ index.html.

References

[1] A. N. Alekseichuk, S. N. Koniushok, A. Yu. Storozhuk *Obobshchennaia statystycheskaia ataka na synkhronnыe potochnыe shyfrы* // Zakhyst informatsii.– 2015. – T. 17. – # 3. – S. 54–65.

[2] S. Fischer, S. Khazaei, W. Meier, *Chosen IV statistical analysis for key recovery attacks on stream ciphers* // AFRICACRYPT 2008. – Proceedings. – Springer-Verlag. – 2008. P. 236–245.

[3] I. Dinur, A. Shamir *Cube attacks on tweakable black box polynomials* // Advances in Cryptology – EUROCRYPT09. – Proceedings. – Springer-Verlag. – 2009. – P.278–299.

[4] P. Gopalan, R. ODonnel, A. Servedio, A. Shpilka, K. Wimmer *Testing Fourier dimensionality and sparsity* // SIAM J. on Computing. – 2011. – Vol. 40(4). – P. 1075–1100.

[5] P. A. Gopalan *Fourier-analytic approach to Reed-Muller decoding* // Annual IEEE Symp. on Foundation in Computer Science. – FOCS 2010. – Proceedings. – Springer-Verlag. – 2010. – P. 685–694.

[6] A. M. Oleksiichuk, S. M. Koniushok, A. Yu. Storozhuk *Shvydki alhorytmy pobudovy k-vymirnykh nablyzhen nablyzhen bulevykh funktsii* // Zakhyst informatsii.– 2015. – T. 17. – # 1. – S. 43–52.

[7] E. K. Alekseev *O nekotorыkh merakh nelyneinosty bulevыkh funktsyi* // Prykladnaia dyskretnaia matematyka. – 2011. – # 2(12). – S. 5–16.

[8] O. A. Lohachev, A. A. Salnykov, V. V. Yashchenko *Bulevы funktsyy v teoryy*

kodyrovanyia y kryptolohyy. – M. : MTsNMO, 2004. – 470 s.
[9] A. N. Alekseichuk, S. N. Koniushok *Alhebraychesky vыrozhdennыe pryblyzhenyia bulevыkh funktsyi* // Kybernetyka y systemnыi analyz.– 2014. – T. 50. – # 6. – S. 3–14.
[10] T. Gowers *A new proof of Szemeredis theorem* // Geom Funct. Anal. – 2001. – Vol. 11(3). – P. 465–588.
[11] W. Hoeffding *Probability inequalities for sums of bounded random variables* // J. Amer. Statist. Assoc. – 1963. – Vol. 58. – # 301. – P. 13–30.
[12] M. Nell, T. Johansson, A. Maximov, W. Meier *A stream cipher proposal: Grain-128* // eStream: the ECRYPT Stream Cipher Project. http://www.ecrypt.eu.org/stream/ index.html.

Реферат

Олексійчук Антон, Конюшок Сергій, Сторожук Артем

Метод пошуку алгебраїчно вироджених наближень булевих функцій для побудови статистичних атак на синхронні потокові шифри

В 2013 році авторами цієї статті запропоновано статистичну атаку на синхронні потокові шифри (СПШ), яка узагальнює деякі раніше відомі атаки на основі підібраних векторів ініціалізації. Оцінювання стійкості СПШ відносно зазначеної атаки зводиться до знаходження або обґрунтування відсутності певних алгебраїчно вироджених наближень булевих функцій, заданих за допомогою оракулів. Ця задача є обчислювально складною, а ефективні алгоритми її розв'язання відомі лише для окремих випадків.

В даній статті пропонується метод розв'язання поставленої задачі, який відрізняється за сутністю від відомих та базується на отриманих умовах, яким задовольняють шукані наближення. На відміну від інших, запропонований метод не має передумовою виконання будь-яких обмежень стосовно відстані, на якій треба відшукати наближення; він також дозволяє знаходити наближення з більш широкого класу булевих функцій. Крім того, за певних умов запропонований метод надає можливість переконуватися у відсутності зазначених наближень, що дозволяє використовувати його для обґрунтування практичної стійкості СПШ відносно статистичних атак.

Метод складається з двох етапів, на першому з яких здійснюється пошук так званих допустимих підпросторів для вхідної функції-оракула (при цьому відсутність зазначених підпросторів свідчить про відсутність шуканих наближень). Після цього, на другому етапі за знайденим допустимом підпростором будується булева функція від меншої кількості змінних, яка (разом з базисом знайденого підпростору) задає шукане наближення вхідної функції.

Для вирішення окремих задач на кожному етапі методу запропоновано поліноміальні ймовірнісні алгоритми, які можуть бути застосовані на практиці до функцій-оракулів від декількох десятків або сотень змінних.

Застосування отриманих результатів до редукованої версій СПШ Grain-128 показує, що за певних умов складність узагальненої статистичної атаки на шифр не перевищує 2^{97}, в той час як складність раніше відомої статистичної атаки складає 2^{124}.

В цілому, викладені в статті результати надають більше можливостей як для побудови статистичних атак на синхронні потокові шифри, так і для обґрунтування стійкості таких шифрів відносно зазначених атак.

Алексейчук Антон, Конюшок Сергей, Сторожук Артем

Метод поиска алгебраически вырожденных приближений булевых функций для построения статистических атак на синхронные поточные шифры

В 2013 году авторами этой статьи предложена статистическая атака на синхронные поточные шифры (СПШ), которая обобщает некоторые ранее известные атаки на основе подобранных векторов инициализации. Оценивание стойкости СПШ относительно указанной атаки сводится к нахождению или обоснованию отсутствия определенных алгебраически вырожденных приближений булевых функций, заданных при помощи оракулов. Эта задача является вычислительно сложной, а эффективные алгоритмы ее решения известны только для частных случаев.

В данной статье предлагается метод решения поставленной задачи, который отличается по сути от известных и базируется на полученных условиях, которым удовлетворяют искомые приближения. В отличие от других, предложенный метод не предполагает выполнения каких-либо ограничений относительно расстояния, на котором требуется отыскать приближения; он также позволяет находить приближения из более широкого класса булевых функций. Кроме этого, при определенных условиях предложенный метод дает возможность убеждаться в отсутствии указанных приближений, что позволяет использовать его для обоснования практической

стойкости СПШ относительно статистических атак.

Метод состоит из двух этапов, на первом из которых осуществляется поиск так называемых допустимых подпространств для исходной функции-оракула (при этом отсутствие указанных подпространств свидетельствует об отсутствии искомых приближений). Затем, на втором этапе по найденному допустимому подпространству строится булева функция от меньшего числа переменных, которая (вместе с базисом найденного подпространства) задает искомое приближение исходной функции.

Для решения частных задач на каждом этапе метода предложены полиномиальные вероятностные алгоритмы, которые могут быть применены на практике к функциям-оракулам от нескольких десятков или сотен переменных.

Применение полученных результатов к редуцированной версии СПШ Grain-128 показывает, что при определенных условиях сложность обобщенной статистической атаки на шифр не превосходит 2^{97}, в то время как сложность ранее известной статистической атаки составляет 2^{124}.

В целом, изложенные в статье результаты предоставляют больше возможностей как для построения статистических атак на синхронные поточные шифры, так для обоснования стойкости таких шифров относительно указанных атак.

Alekseychuk Anton, Konushok Sergey, Storozhuk Artem

A method for finding algebraic degenerate approximations of boolean functions for construction of statistical attacks on sinchronous stream ciphers

In 2013, the authors of this paper proposed a statistical attack on synchronous stream ciphers (SSC). This attack generalize some previously known chosen initialization vectors attacks. The security evaluation of SSC against mentioned attacks reduces to finding or absence justification of some algebraic degenerate approximations to Boolean functions specified by oracles. This task is computationally hard and effective algorithms of its solving for special cases are only known.

In this paper, we propose a method for solving of this task. This method differs by its nature from previously known and is based on conditions for the existence of searched approximations. Unlike others, the proposed method doesn't require any restrictions on distance for the approximations that have to be found; it also allows finding approximations from wider class of Boolean functions. Except this, under certain conditions the proposed method gives the possibility to ensure that specified approximations are absent, what allows using it for proving the practical resistance of SSC against statistical attacks.

The method proceeds in two stages, on the first one it is required to find a so-called admissible subspace for the given oracle (wherein the absence of specified subspaces points on the absence of searched approximations). Then, on the second stage a Boolean function of lower number of variables is built. This function with a found subspace provides a searched approximation of the given oracle.

In order to solve particular tasks on each phase of the method we propose polynomial probabilistic algorithms that can be applied in practice to oracles of dozens or hundreds variables.

Applying obtained results to reduced version of Grain-128 shows that under certain conditions the complexity of generalized statistical attack on the cipher doesn't exceed 2^{97}, while the complexity of previously known statistical attack is 2^{124}.

In general, the obtained results provide more opportunities for both constructing statistical attacks on synchronous stream ciphers and proving of their resistance against such attacks.

Відомості про авторів

Олексійчук Антон Миколайович
Освіта: прикладна математика (1992).
Місце роботи: Державний заклад «Інститут спеціального зв'язку та захисту інформації НТУУ «КПІ», доцент, д.т.н. (2009).
Область знань: криптографія.
Наукові інтереси: теоретичні дослідження.
Email: *alex-dtn@ukr.net*

Конюшок Сергій Миколайович
Освіта: безпека інформації в спеціальних інформаційних системах (2002).
Місце роботи: Державний заклад «Інститут спеціального зв'язку та захисту інформації НТУУ «КПІ», доцент, к.т.н. (2006).
Область знань: криптографія.
Наукові інтереси: криптографічні властивості булевих функцій.
Email: *3tooth@gmail.com*

Сторожук Артем Юрійович
Освіта: безпека державних інформаційних ресурсів (2012).
Місце роботи: Державний заклад «Інститут спеціального зв'язку та захисту інформації НТУУ «КПІ».
Область знань: криптографія.
Наукові інтереси: прикладні та експериментальні дослідження.
Email: *storojs72@gmail.com*