ВЫСОКОВЕРОЯТНЫЕ *k*-МЕРНЫЕ ПРИБЛИЖЕНИЯ БУЛЕВЫХ ФУНКЦИЙ

А. Н. Алексейчук

Институт специальной связи и защиты информации НТУУ «КПИ»
Киев, Украина
E-mail: alex-crypto@mail.ru

Булева функция от n переменных называется *k*-мерной, если она линейно эквивалентна функции от k переменных, $0 \le k \le n$. Основным результатом, представленным в докладе, является теорема о строении *k*-мерных функций степени d, на расстоянии не более $2^{n-d}(1-\varepsilon)$, $\varepsilon \in (0,1)$, от заданной булевой функции n переменных. Эта теорема существенно усиливает известный результат П. Гопалана и позволяет предложить эффективный алгоритм построения всех указанных *k*-мерных булевых функций.

Ключевые слова: корреляционный криптоанализ, преобразования Уолша-Адамара, *k*-мерное приближение булевой функции.

Основные понятия и обозначения

Обозначим V_n множество двоичных векторов длины n. Это множество является векторным пространством размерности n над полем $F = \mathbf{GF}(2)$ (при $n = 0$ полагаем $V_0 = \{0\}$). Сумма векторов $\alpha = (\alpha_1, \dots, \alpha_n)$, $x = (x_1, \dots, x_n) \in V_n$ определяется по формуле $\alpha \oplus x = (\alpha_1 \oplus x_1, \dots, \alpha_n \oplus x_n)$, а булево скалярное произведение – по формуле $\alpha x = \alpha_1 x_1 \oplus \cdots \oplus \alpha_n x_n$ (здесь и ниже символ $\oplus$ обозначает операцию сложения как элементов поля F, так и векторов над этим полем).

Обозначим B_n множество булевых функций от n переменных. Относительное расстояние между функциями $f, g \in B_n$ определяется по формуле $d(f,g) = 2^{-n} |\{x \in V_n : f(x) \ne g(x)\}|$, а относительное расстояние между функцией $f \in B_n$ и множеством $U \subseteq B_n$ – по формуле $d(f,U) = \min\limits_{g \in U} d(f,g)$.

Для любой функции $f \in B_n$ положим $wt(f) = 2^{-n} |\{x \in V_n : f(x) = 1\}|$,

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

$$I_f = \{\alpha \in V_n \mid \forall x \in V_n : f(x \oplus \alpha) = f(x)\}. \tag{2}$$

Числа (1) называются нормированными коэффициентами Уолша-Адамара функции f. Множество (2) является подпространством векторного пространства V_n, дуальное к которому подпространство имеет следующий вид (см., например, [1], задача 2.112):

$$I_f{}^{\perp} = \left\langle \{\alpha \in V_n : \hat{f}(\alpha) \ne 0\} \right\rangle. \tag{3}$$

Функция $g \in B_n$ называется k-мерной [2, 3], $k \in \overline{0,n}$, если $\dim I_g^{\perp} \le k$ или, что равносильно, если существуют функция $\varphi \in B_k$ и векторы $\alpha_1, \ldots, \alpha_k \in V_n$ такие, что

$$g(x) = \varphi(\alpha_1 x, \ldots, \alpha_k x), \ x \in V_n . \qquad (4)$$

Обозначим $B_{n,k}$ множество всех k-мерных функций от n переменных, положим $B_{n,-1} = \varnothing$. Справедливы соотношения

$$B_{n,0} \subseteq B_{n,1} \subseteq \ldots \subseteq B_{n,n-1} \subseteq B_{n,n};$$

при этом множество $B_{n,0}$ состоит из двух функций-констант, множество $B_{n,1}$ совпадает с классом аффинных функций, а множество $B_{n,n}$ – с совокупностью всех булевых функций от n переменных. Функции из множества $B_{n,n-1}$ называются алгебраически вырожденными, а функции из множества $B_n \setminus B_{n,n-1}$ – невырожденными [4, 5].

Краткий обзор публикаций, посвященных алгебраически выраженным функциям

Первые результаты о корреляционных свойствах алгебраически вырожденных булевых функций относятся к 70-м годам прошлого века [6]. В последнее время интерес к исследованию этих функций во многом обусловлен задачами криптоанализа и теории кодирования. Отметим работы [7–9], в которых описан ряд атак на генераторы гаммы поточных шифров, функции усложнения которых алгебраически вырожденны или близки к таковым.

В [5] исследованы приближения булевых функций функциями из множества $B_{n,n-1}$, в частности, получено выражение для расстояния между произвольной функцией $f \in B_n$ и множеством всех алгебраически вырожденных функций от n переменных, указан способ нахождения функций, ближайших к f во множестве $B_{n,n-1}$ и получены оценки их порядков (в [4, 5] порядком вырожденности функции $g \in B_{n,n-1}$ называется число $n-k$ такое, что $g \in B_{n,k} \setminus B_{n,k-1}$, $k \in \overline{0,n-1}$).

Изучению k-мерных приближений булевых функций при всех возможных значениях k посвящены работы [2, 3, 10, 11], причем в [3] рассматриваются функции над произвольным конечным полем. В [2] предложен вероятностный алгоритм распознавания свойства k-мерности. Для любой функции $f \in B_n$, заданной с помощью оракула, и чисел $k \in \overline{0,n-1}$, $\varepsilon \in (0,1)$ этот алгоритм позволяет проверить гипотезу H_0: $f \in B_{n,k}$ против альтернативы H_1: $d(f, B_{n,k}) \ge \varepsilon$ за $O(n 2^{2k} k\, \varepsilon^{-1})$ двоичных операций. Более эффективный тест k-мерности, трудоемкость которого составляет $O(n\, 2^k k^2\, \varepsilon^{-1})$ двоичных операций, предложен в [10].

Для построения эффективных атак на симметричные криптосистемы необходимо находить k-мерные функции, достаточно близкие к заданной булевой функции $f \in B_n$. При этом наибольший интерес представляет случай, в котором функция f

задается с помощью оракула, число n велико (например, $n \geq 64$), а k – мало (фиксировано или медленно растет с ростом n). Эффективность решения этой задачи существенно зависит от расстояния между функцией f и ее искомыми приближениями.

Пусть g – k-мерная функция от n переменных, удовлетворяющая условию $d(f,g) \leq 2^{-(k+1)}(1-\varepsilon)$, $\varepsilon \in (0,1)$ (отметим, что функция g определяется этим условием однозначно). В [2] предложен вероятностный алгоритм, позволяющий восстанавливать g по заданным f, k и ε с вероятностью не менее $1-\delta$, $\delta \in (0,1)$, за $O(2^{4k} n^2 \varepsilon^{-2} \log(2^{2k} n \delta^{-1}))$ двоичных операций. В [11] представлен другой алгоритм, двоичная сложность которого равна $O(2^{2k} k^{-2} n^3 \varepsilon^{-2} \delta^{-1} \log(2^{2k} k^{-1} n \delta^{-1} \varepsilon^{-1}))$.

Задача нахождения всех функций $g \in B_{n,k}$ таких, что $d(f,g) \leq 2^{-k}(1-\varepsilon)$, $\varepsilon \in (0,1)$, является более трудной. Существенный вклад в ее решение сделан в работе [3], посвященной изучению k-мерных приближений функций от n переменных над произвольным конечным полем. Один из основных результатов этой работы (применительно к булевым функциям) состоит в следующем.

Теорема 1 [3]**.** Пусть $f \in B_n$, $g \in B_{n,k}$, $\deg g \leq d$ и $d(f,g) \leq 2^{-d}(1-\varepsilon)$, где $1 \leq d \leq k$, $\varepsilon \in (0,1)$. Тогда

$$I_g^{\perp} = \left\langle \{\alpha \in I_g^{\perp} : |\hat{f}(\alpha)| \geq \frac{1}{8\sqrt{2}} 2^{-k/2-d} \varepsilon^2\} \right\rangle.$$

Из теоремы 1 следует, что каждая функция g, удовлетворяющая ее условию, допускает такое представление (4), в котором векторы $\alpha_1, \ldots, \alpha_k$ принадлежат множеству $S_f(\mu) = \{\alpha \in V_n : |\hat{f}(\alpha)| \geq \mu\}$ при $\mu = \frac{1}{8\sqrt{2}} 2^{-k/2-d} \varepsilon^2$. Поскольку $|S_f(\mu)| \leq \mu^{-2}$ для любых $f \in B_n$, $\mu \in (0,1)$, то число функций g ограничено сверху величиной

$$\mu^{-2k} N(k,d) = 2^{k^2 + k(2d+7)} \varepsilon^{-4k} N(k,d), \qquad (5)$$

где $N(k,d) = 2^{\sum_{i=0}^{d} \binom{k}{i}}$ – число булевых функций степени не выше d от k переменных. Таким образом, для нахождения всех указанных функций достаточно перебрать всевозможные наборы $(\alpha_1, \ldots, \alpha_k, \varphi)$, где $\alpha_1, \ldots, \alpha_k \in S_f(\mu)$, $\varphi \in B_k$, $\deg \varphi \leq d$, задать функцию g по формуле (4) и проверить условие $d(f,g) \leq 2^{-d}(1-\varepsilon)$. Если каждую такую проверку принять за одну операцию, то трудоемкость описанного алгоритма построения всех функций g по заданному множеству $S_f(\mu)$ [3] оценивается сверху по формуле (5).

Основные результаты

Для любой функции $g \in B_n$ положим $\Delta(g) = 1/2 \cdot \min\limits_{\alpha \notin I_g} wt(D_\alpha g)$, где $D_\alpha g(x) = g(x \oplus \alpha) \oplus g(x)$, $x \in V_n$ – производная функции g по направлению $\alpha \in V_n$.

Следующая теорема обобщает и одновременно усиливает теорему 1.

Теорема 2. Пусть $f \in B_n$, $g \in B_{n,k} \setminus B_{n,k-1}$, $\deg g \le d$ и $d(f,g) \le \Delta(g)(1-\varepsilon)$, где $1 \le d \le k$, $\varepsilon \in (0,1)$. Тогда

$$I_g^{\perp} = \left\langle \left\{ \alpha \in I_g^{\perp} : |\hat{f}(\alpha)| \ge \max\left\{ 2^{1-k}\varepsilon,\ \frac{4}{3\sqrt{3}} 2^{-k/2} \varepsilon^{3/2} \Delta(g)^{1/2} \right\} \right\} \right\rangle.$$

Следствие 1. Пусть $f \in B_n$, $g \in B_{n,k}$, $\deg g \le d$ и $d(f,g) \le 2^{-d}(1-\varepsilon)$, где $1 \le d \le k$, $\varepsilon \in (0,1)$. Тогда функция g допускает такое представление (4), в котором векторы $\alpha_1, \ldots, \alpha_k$ принадлежат множеству $S_f(\mu_0)$, где

$$\mu_0 = \max\left\{ 2^{1-k}\varepsilon,\ \frac{4}{3\sqrt{3}} 2^{-k/2-d/2} \varepsilon^{3/2} \right\}. \tag{6}$$

В частности, число указанных функций g ограничено сверху величиной

$$\mu_0^{-2k} N(k,d) < \min\left\{ 2^{2k^2-2k}\varepsilon^{-2k},\ 2^{k^2+k(d+1)}\varepsilon^{-3k} \right\} N(k,d), \tag{7}$$

где $N(k,d) = 2^{\sum_{i=0}^{d} \binom{k}{i}}$ – число булевых функций степени не выше d от k переменных.

Как видно из формул (5) и (7), полученная оценка количества k-мерных функций, удовлетворяющих условию теоремы 1, заметно лучше оценки из [3]. Более того, справедливо следующее утверждение.

Следствие 2. Пусть $f \in B_n$, $k \in \overline{1, n-1}$. Тогда каждая функция $g \in B_{n,k} \setminus B_{n,k-1}$, удовлетворяющая условию $d(f,g) \le 2^{-k}(1-\varepsilon)$, $\varepsilon \in (0,1)$, допускает такое представление (4), в котором векторы $\alpha_1, \ldots, \alpha_k$ принадлежат множеству $S_f\left(2^{1-k}\varepsilon\right)$. При этом число указанных функций не превосходит $2^{2k^2-2k}\varepsilon^{-2k}\left(2^k+1\right)$.

Отметим, что импликация

$$\left(d(f,g) \le 2^{-k}(1-\varepsilon),\ g(x) = \varphi(\alpha_1 x, \ldots, \alpha_k x),\ x \in V_n \right) \Rightarrow \left(\alpha_1, \ldots, \alpha_k \in S_f\left(2^{1-k}\varepsilon\right) \right),$$

справедливая при выполнении условия следствия 2, обращается в равносильное утверждение при $k = 1$: аффинная функция с вектором коэффициентов $\alpha \in V_n \setminus \{0\}$ тогда и только тогда находится от функции $f \in B_n$ на относительном расстоянии не более $1/2 \cdot (1-\varepsilon)$, $\varepsilon \in (0,1)$, когда $|\hat{f}(\alpha)| \ge \varepsilon$. При этом оценка из теоремы 1 позволяет лишь утверждать, что $|\hat{f}(\alpha)| \ge 1/16 \cdot \varepsilon^2$.

В целом полученные результаты (наряду с другими естественными усовершенствованиями) дают возможность заметно повысить эффективность предложенного в [3] алгоритма построения k-мерных приближений булевых функций.

Библиографические ссылки

1. *Логачев О. А., Сальников А. А., Ященко В. В.* Булевы функции в теории кодирования и криптологии. М. : МЦНМО, 2004.
2. *Gopalan P., O'Donnel R., Servedio A., Shpilka Amir, Wimmer Karl*. Testing Fourier dimensionality and sparsity // SIAM J. on Computing. 2011. V. 40(4). P. 1075–1100.
3. *Gopalan P.* A Fourier-analytic approach to Reed-Muller decoding // Annual IEEE Symp. on Foundation in Computer Science. FOCS 2010, Proceedings. Berlin. Springer-Verlag. 2010. P. 685–694.
4. *Dawson E., Wu C. K.* Construction of correlation immune Boolean functions // Information and Communication Security, Proceedings. Berlin. Springer-Verlag. 1997. P. 170–180.
5. *Алексеев Е. К.* О некоторых мерах нелинейности булевых функций // Прикладная дискретная математика. 2011. № 2(12). С. 5–16.
6. *Lechner R. L.* Harmonic analysis of switching functions // Recent Developments in Switching Theory. New-York: Academic Press, 1971. P. 122–228.
7. *Daemen J., Govaerts R., Vandewalle J.* Resynchronization weaknesses in synchronous stream ciphers // Advances in Cryptology – EUROCRYPT'93, Proceedings. Berlin: Springer-Verlag, 1993. P. 159–167.
8. *Golić J., Morgari G.* On the resynchronization attack // Fast Software Encryption. FSE'03. Proceedings. Berlin: Springer-Verlag, 2003. P. 100–110.
9. *Алексеев Е.* Об атаке на фильтрующий генератор с функцией усложнения, близкой к алгебраически вырожденной // Сборник статей молодых ученых факультета МВК МГУ. 2011. В. 8. С. 114–123.
10. *Алексейчук А. Н., Конюшок С. Н.* Усовершенствованный тест *k*-мерности для булевых функций // Кибернетика и системный анализ. 2013. Т. 49. № 2. С. 27–35.
11. *Alekseychuk A. N., Konyushok S. N.* Fast algorithm for reconstruction of high-probable low-dimensional approximations for Boolean functions // Modern Stochastics: Theory and Applications III, Proceedings. Kyiv: Taras Shevchenko National University, 2012. P. 32.