K. N. Pankov, An upper bound for the number of functions satisfying the strict avalanche criterion,
Diskr. Mat., 2005, Volume 17, Issue 2, 95–101

Дискретная математика

том 17 выпуск 2 * 2005

УДК 519.7

Верхняя граница для числа функций, удовлетворяющих строгому лавинному критерию

© 2005 г. К. Н. Панков

Строгий лавинный критерий был предложен при изучении критериев построения некоторых криптографических функций. Двоичная функция $f(x)$, $x \in V_n$, удовлетворяет этому критерию, если при замене любой координаты вектора x ее дополнением значение $f(x)$ изменяется ровно в половине случаев. В данной работе представлена верхняя граница для числа таких функций при достаточно больших n.

Пусть V_n — множество двоичных векторов размерности n. Двоичная функция $f(x)$, $x \in V_n$, удовлетворяет строгому лавинному критерию, если при замене любой координаты вектора x дополнением значение $f(x)$ изменяется ровно в половине случаев. Это эквивалентно тому, что для любого $i = 1, 2, \ldots$

$$\| f(x) \oplus f(x \oplus e_i) \| = \sum_{\beta \in V_n} (f(\beta) \oplus f(\beta \oplus e_i)) = 2^{n-1},$$

где $e_i \in V_n$ — вектор веса 1 с единицей на i-м месте.

Данный критерий был введен в [1]. Как и для любого другого критерия, для строгого лавинного критерия представляет интерес подсчет числа удовлетворяющих ему функций от n переменных для произвольного n. Данное число обозначим через S^n. К настоящему времени получены лишь оценки этого числа. Нижние оценки, представленные в [2, 3, 4] определяются путем построения классов функций, удовлетворяющих изучаемому критерию, и вычисления мощностей этих классов. В [5] с помощью представления элементов V_n как вершин n-мерного гиперкуба доказана самая сильная на нынешний момент нижняя оценка для S^n, справедливая для достаточно больших n:

$$S^n \geqslant \frac{\exp_2(2^n - n^2/2 + n)}{\pi^{n/2}},$$

где $\exp_2(x) = 2^x$.

Нетривиальная верхняя граница для S^n была построена в [6].

Пусть на множестве V_n задано равномерное распределение вероятностей, тогда $f(x)$ можно рассматривать как случайную величину. Будем говорить, что двоичная функция $f(x)$, $x = (x_1, \ldots, x_n) \in V_n$, зависит на 50% от входа x_i, если вероятность того, что она принимает разные значения на векторах, различающихся только i-й координатой, равна $1/2$.

Обозначим через $S(n,k)$ число функций от n переменных, зависящих на 50% от переменных $x_1,\dots,x_k$. Функция $f(x)$ зависит на 50% от переменных $x_1,\dots,x_k$ тогда и только тогда, когда

$$\sum_{\beta\in V_n}\begin{pmatrix} f(\beta)\oplus f(\beta\oplus e_1)\\ \dots\\ f(\beta)\oplus f(\beta\oplus e_k)\end{pmatrix}=\begin{pmatrix}2^{n-1}\\ \dots\\ 2^{n-1}\end{pmatrix}$$

При этом $S^n = S(n,n)$.

В [6] найдены точные значения для $S(n,1)$ и $S(n,2)$:

$$S(n,1)=\binom{2^{n-1}}{2^{n-2}}2^{n-1},$$

$$S(n,2)=\sum_{i=0}^{2^{n-3}}\binom{2^{n-2}}{2i}8^{2^{n-2}-2i}2^{2i}\sum_{j=0}^{i}\binom{2i}{2j}\binom{2j}{j}\binom{2i-2j}{i-j}.$$

Очевидно, что $S^n \leqslant S(n,k)$ для любого $k \leqslant n$, однако вычисление точных значений $S(n,k)$ при $k \geqslant 2$ представляет значительную сложность.

Обозначим через A_i событие, заключающееся в том, что случайная двоичная функция не является зависимой на 50% от переменной x_i. Тогда

$$\mathbf{P}(A_i)=1-\frac{S(n,1)}{\exp_2(2^n)}<1,$$

$$\mathbf{P}(A_i\cap A_j)=1-\frac{S(n,2)}{\exp_2(2^n)}<1,\qquad i\neq j.$$

В [6] предложена следующая верхняя оценка для числа функций, удовлетворяющих строгому лавинному критерию:

$$S^n \leqslant 2^{2^n}-\frac{2^{2^n}n^2(\mathbf{P}(A_1))^2}{n\mathbf{P}(A_1)+n(n-1)\mathbf{P}(A_1\cap A_2)}. \tag{1}$$

В данной работе доказывается асимптотическая формула для $S(n,k)/\exp_2(2^n)$ в случае произвольного натурального фиксированного k и из нее в качестве следствия выводится верхняя оценка S^n, улучшающая оценку (1).

Теорема 1. *Пусть $n\to\infty$, k — фиксированное натуральное число. Тогда*

$$\frac{S(n,k)}{\exp_2(2^n)}=2^{2k-1}\pi^{-k/2}2^{-nk/2}+o(2^{-n(k+1)/2}). \tag{2}$$

Доказательство. Рассмотрим сумму

$$\Sigma_n=\sum_{\beta\in V_n}\begin{pmatrix} f(\beta)\oplus f(\beta\oplus e_1)\\ \dots\\ f(\beta)\oplus f(\beta\oplus e_k)\end{pmatrix}$$

и разобьем область суммирования на 2^{n-k} областей, соответствующих векторам β с оди-

наковыми последними $n-k$ координатами, набор которых будем обозначать α. Тогда

$$\Sigma_n = \sum_{\alpha\in V_{n-k}} \left(\sum_{\gamma\in V_k} \begin{pmatrix} f(\gamma\alpha)\oplus f(\gamma\alpha\oplus e_1) \\ \dots \\ f(\gamma\alpha)\oplus f(\gamma\alpha\oplus e_k) \end{pmatrix} \right)$$

$$= \sum_{\alpha\in V_{n-k}} 2 \begin{pmatrix} \sum_{t\in V_{k-1}} (f(1\,t\,\alpha)\oplus f(0t\alpha)) \\ \dots \\ \sum_{t\in V_{k-1}} (f(t_1\dots t_{i-1}\,1\,t_i\dots t_k\alpha)\oplus f(t_1\dots t_{i-1}\,0\,t_i\dots t_{k-1}\alpha)) \\ \dots \\ \sum_{t\in V_{k-1}} (f(t\,1\,\alpha)\oplus f(t\,0\,\alpha)) \end{pmatrix},$$

здесь $t=(t_1,\dots,t_{k-1})\in V_{k-1}$, $(1\,t\,\alpha)$ означает конкатенацию векторов 1, t и α, $\sum$ означает сложение в поле действительных чисел, а $\oplus$ — сложение по модулю 2.

Пусть функция f выбирается случайно и равновероятно из множества B_n всех двоичных функций от n переменных. Это эквивалентно независимому равновероятному выбору значений $f(\beta)$ из $\{0,1\}$ для всех β из множества V_n. Очевидно, что

$$\mathbf{P}\left(\sum_{\beta\in V_n} \begin{pmatrix} f(\beta)\oplus f(\beta\oplus e_1) \\ \dots \\ f(\beta)\oplus f(\beta\oplus e_k) \end{pmatrix} = \begin{pmatrix} 2^{n-1} \\ \dots \\ 2^{n-1} \end{pmatrix} \right) = \frac{S(n,k)}{2^{2^n}}.$$

Введем случайный вектор

$$\eta_\alpha = \begin{pmatrix} \eta_\alpha^1 \\ \dots \\ \eta_\alpha^k \end{pmatrix} = \begin{pmatrix} \sum_{t\in V_{k-1}} (f(1\,t\,\alpha)\oplus f(0\,t\,\alpha)) \\ \dots \\ \sum_{t\in V_{k-1}} (f(t\,1\,\alpha)\oplus f(t\,0\,\alpha)) \end{pmatrix}.$$

Слагаемые в каждой компоненте вектора η_α независимы и принимают значения 0 и 1 с вероятностью $1/2$. Поэтому каждая компонента вектора η_α имеет биномиальное распределение с параметрами $(2^{k-1}, 1/2)$.

Очевидно, что все случайные векторы η_α независимы, одинаково распределены и

$$\mathbf{M}\eta_\alpha = \begin{pmatrix} 2^{k-2} \\ \dots \\ 2^{k-2} \end{pmatrix},$$

$$\mathbf{P}\left(\sum_{\beta\in V_n} \begin{pmatrix} f(\beta)\oplus f(\beta\oplus e_1) \\ \dots \\ f(\beta)\oplus f(\beta\oplus e_k) \end{pmatrix} = \begin{pmatrix} 2^{n-1} \\ \dots \\ 2^{n-1} \end{pmatrix} \right) = \mathbf{P}\left(\sum_{\alpha\in V_{n-k}} \eta_\alpha = \begin{pmatrix} 2^{n-2} \\ \dots \\ 2^{n-2} \end{pmatrix} \right).$$

Лемма 1. *Справедливо равенство*

$$\mathbf{M}\eta_\alpha^i\eta_\alpha^j = 2^{2(k-2)} + 2^{k-3} I\{i=j\}, \qquad 1\leqslant i,j\leqslant k. \tag{3}$$

Доказательство. Справедливо равенство

$$\eta_\alpha^i = \frac{1}{2}\sum_{t\in V_k} (f(t\alpha)\oplus f(t\alpha\oplus e_i)).$$

Если $i=j$, то $\mathbf{M}(\eta_\alpha^i)^2$ — второй момент случайной величины с биномиальным распределением с параметрами $(2^{k-1}, 1/2)$ (суммы 2^{k-1} независимых равновероятных индикаторов).

Если $i \neq j$, то

$$\mathbf{M}\eta_\alpha^i \eta_\alpha^j = \frac{1}{4}\mathbf{M} \sum_{t_1,t_2 \in V_k} (f(t_1\alpha) \oplus f(t_1\alpha \oplus e_i))(f(t_2\alpha) \oplus f(t_2\alpha \oplus e_j)).$$

Значения f в разных точках независимы и равновероятно распределены на $V = \{0, 1\}$. При любых t_1, t_2 и $i \neq j$, каждый сомножитель содержит слагаемое, не зависящее от второго, следовательно, сомножители независимы.

Лемма 1 доказана.

Продолжим доказательство теоремы. Из равенств (3) следует, что $\mathbf{D}\eta_\alpha^i = 2^{k-3}$ и $\operatorname{cov} \eta_\alpha^i \eta_\alpha^j = 0$ при $i \neq j$.

Следовательно, ковариационная матрица случайного вектора η_α имеет вид $\operatorname{Cov} \eta_\alpha = 2^{k-3} I$, где I — единица в кольце квадратных матриц размера k.

Рассмотрим решетку Z_k всех целочисленных векторов размерности k — дискретную подгруппу в множестве всех вещественных векторов размерности k. Ясно, что

$$\mathbf{P}(\eta_\alpha \in Z_k) = 1.$$

Следовательно, η_α — решетчатый случайный вектор.

Согласно [7] решетка M называется минимальной для случайного вектора X, если она удовлетворяет следующим двум условиям:

(1) $\mathbf{P}(X \in x + M) = 1$ для любого x такого, что $\mathbf{P}(X = x) > 0$,

(2) если существует подгруппа C множества всех вещественных векторов размерности k такая, что $\mathbf{P}(X \in y + C) = 1$ при некотором y, то $M \subseteq C$.

Пусть

$$\bar{e} = \sum_{j=1}^{k} e_j, \quad e_i \in V_k, \quad i = 1, \ldots, k.$$

Рассмотрим множество L такое, что

$$L = 2Z_k \cup \{2Z_k + \bar{e}\}.$$

Лемма 2. *Множество L является минимальной решеткой для случайного вектора η_α.*

Доказательство. Очевидно, что L является решеткой.

Рассмотрим распределение случайного вектора η_α. Многочлен Жегалкина функции f можно единственным образом представить в виде

$$f(x_1, \ldots, x_n) = \bigoplus_{\nu \in V_k} x^\nu f_\nu(x_{k+1}, \ldots, x_n),$$

где $x = (x_1, \ldots, x_k)$, $\nu = (\nu_1, \ldots, \nu_k)$, $x^\nu = x_1^{\nu_1} \ldots x_k^{\nu_k}$, а знак $\bigoplus$ означает сложение по модулю 2.

Случайному и равновероятному выбору f из множества B_n соответствует случайный, независимый и равновероятный выбор из множества B_{n-k} набора из 2^k функций f_ν, $\nu \in V_k$, значения которых можно рассматривать как независимые случайные величины, равномерно распределенные на V.

Несложно убедиться, что случайную величину η_α^i можно представить в виде

$$\eta_\alpha^i = \sum_{m=0}^{k-1} \sum_{J \subset \{1,\dots,k\} \setminus \{i\}} \bigoplus_{\mu \in V_m} f_{e_i \oplus \mu_1 e_{j_1}} \oplus \dots \oplus \mu_m e_{j_m}(\alpha),$$

где $|J| = m$, $J = \{j_1, \dots, j_m\}$, $\mu = (\mu_1, \dots, \mu_m)$.

Если в представлении η_α^i перейти от суммирования по модулю 2 к обычному суммированию в поле действительных чисел, использовав то, что

$$f_1(\alpha) \oplus f_2(\alpha) = (f_1(\alpha) - f_2(\alpha))^2 = f_1(\alpha) + f_2(\alpha) - 2f_1(\alpha)f_2(\alpha)$$

нетрудно увидеть, что

$$\eta_\alpha = \begin{pmatrix} 2G_1(f_{(1t)}(\alpha),\ t \in V_{k-1}, t \neq \bar{e}) + f_{(11\dots11)}(\alpha) \\ \dots \\ 2G_1(f_{(t_1\dots t_{i-1}1\ t_i\dots t_{k-1})}(\alpha), t = (t_1 \dots t_{k-1}) \in V_{k-1}, t \neq \bar{e}) + f_{(11\dots11)}(\alpha) \\ \dots \\ 2G_k(f_{(t1)}(\alpha), t \in V_{k-1}, t \neq \bar{e}) + f_{(11\dots11)}(\alpha) \end{pmatrix}$$

где $G_1(x_i, i \in \{1, 2^{k-1} - 1\}), \dots, G_k(x_i, i \in \{1, 2^{k-1} - 1\})$ — некоторые многочлены от $2^{k-1} - 1$ переменных. Следовательно, случайный вектор η_α принимает значения либо из множества $\{2Z_k\}$, либо из множества $\{2Z_k + \bar{e}\}$. Получаем, что $\mathbf{P}(\eta_\alpha \in L) = 1$.

Нетрудно убедиться, что базисом L является набор векторов вида

$$\{2e_1, 2e_2, \dots, 2e_{k-1}, \bar{e}\}.$$

Если случайная величина $f_{\bar{e} \oplus e_i}(\alpha)$ принимает значение 1, а f_ν принимает значение 0 для прочих $\nu \in V_k$, то нетрудно убедиться, что η_α принимает значение $2\bar{e} - 2e_i$.

Аналогично, если $f_{\bar{e}}(\alpha) = 1$, а $f_\nu = 0$ для прочих $\nu \in V_k$, то η_α принимает значение $\bar{e}$. Для любого i, $0 < i < k$, справедливо равенство $2e_i = 2\bar{e} - (2\bar{e} - 2e_i)$, следовательно, можно сделать вывод, что если $\mathbf{P}(\eta_\alpha \in M) = 1$, где M — некоторая подгруппа множества всех вещественных векторов размерности k, то $L \subseteq M$. Следовательно, L — минимальная решетка для η_α.

Лемма 2 доказана.

Вернемся к доказательству теоремы и рассмотрим случайный вектор $X_\alpha = T\eta_\alpha$, где $T = 2^{-(k-3)/2} I$.

Очевидно, что

$$\mathbf{M}X_\alpha = 2^{(k-1)/2}\bar{e}, \qquad \operatorname{Cov} X_\alpha = I.$$

Нетрудно убедиться, что $\mathbf{M}\|X_\alpha - \mathbf{M}X_\alpha\|^3 < \infty$.

Повторив доказательство леммы 2 с точностью до умножения на матрицу T, можно заметить, что минимальная решетка L' для случайного вектора X_α имеет базис

$$\{2^{(5-k)/2}e_1, \dots, 2^{(5-k)/2}e_{k-1}, 2^{(3-k)/2}\bar{e}\},$$

и следовательно, модуль определителя матрицы, составленной из базисных векторов (определитель решетки, не зависящий от выбора базиса), равен

$$\det L' = 2^{-(k^2-5k+2)/2},$$

и

$$\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}\eta_\alpha=\begin{pmatrix}2^{n-2}\\ \dots\\ 2^{n-2}\end{pmatrix}\right)=\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}X_\alpha=T\begin{pmatrix}2^{n-1}\\ \dots\\ 2^{n-1}\end{pmatrix}\right)$$
$$=\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}X_\alpha=\begin{pmatrix}2^{(2n-k-1)/2}\\ \dots\\ 2^{(2n-k-1)/2}\end{pmatrix}\right).$$

Таким образом, при $n\to\infty$ для суммы 2^{n-k} независимых одинаково распределенных случайных векторов X_α выполнены условия локальной предельной теоремы из [7] для сумм независимых одинаково распределенных случайных векторов и, следовательно,

$$(1+\|y_{\theta,2^{n-k}}\|^3)\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}X_\alpha=\theta\right)=\frac{1}{2^{(n-k)k/2}(2\pi)^{-k/2}}\det L'\exp\left(-\frac{1}{2}\|y_{\theta,2^{n-k}}\|^2\right)$$
$$+\frac{1}{2^{(n-k)(k+1)/2}(2\pi)^{-k/2}}\det L'\exp\left(-\frac{1}{2}\|y_{\theta,2^{n-k}}\|^2\right)$$
$$\times\left\{-\frac{1}{6}(\chi_{(3,0,\dots,0)}(-y_1^3+3y_1)+\dots+\chi_{(0,0,\dots,3)}(-y_k^3+3y_k))\right.$$
$$-\frac{1}{2}(\chi_{(2,1,\dots,0)}(-y_1^2y_2+y_2)+\dots+\chi_{(0,\dots,1,2)}(-y_k^2y_{k-1}+y_{k-1}))$$
$$\left.-\chi_{(1,1,1,0,\dots,0)}(-y_1y_2y_3)+\chi_{(0,\dots,1,1,1)}(-y_{k-2}y_{k-1}y_k)\right\}+o(2^{-n(k+1)/2}),$$

где $y_{\theta,2^{n-k}}=(y_1,\dots,y_k)=2^{-(n-k)/2}(\theta-2^{n-k}\mathbf{M}X_\alpha)$, а $\chi_{(\gamma_1,\dots\gamma_k)}=\chi_\gamma$ — семиинвариант случайного вектора X_α порядка γ.

В нашем случае

$$\theta=2^{(2n-k-1)/2}\bar e,\qquad y_{\theta,2^{n-k}}=2^{-(n-k)/2}(2^{(2n-k-1)/2}\bar e-2^{n-k}2^{(k-1)/2}\bar e)=\bar 0,$$

где $\bar 0\in V_k$ — вектор веса 0, следовательно,

$$\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}X_\alpha=\begin{pmatrix}2^{(2n-k-1)/2}\\ \dots\\ 2^{(2n-k-1)/2}\end{pmatrix}\right)$$
$$=2^{-(k^2-5k+2)/2}2^{-(n-k)k/2}(2\pi)^{-k/2}+o(2^{-n(k+1)/2}),$$
$$\frac{1}{2^n}S(n,k)=\mathbf{P}\left(\sum_{\beta\in V_n}\begin{pmatrix}f(\beta)\oplus f(\beta\oplus e_1)\\ \dots\\ f(\beta)\oplus f(\beta\oplus e_k)\end{pmatrix}=\begin{pmatrix}2^{n-1}\\ \dots\\ 2^{n-1}\end{pmatrix}\right)$$
$$=\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}\eta_\alpha=\begin{pmatrix}2^{n-2}\\ \dots\\ 2^{n-2}\end{pmatrix}\right)=\mathbf{P}\left(\sum_{\alpha\in V_{n-k}}X_\alpha=\begin{pmatrix}2^{(2n-k-1)/2}\\ \dots\\ 2^{(2n-k-1)/2}\end{pmatrix}\right)$$
$$=2^{2k-1}\pi^{-k/2}2^{-nk/2}+o(2^{-n(k+1)/2}).$$

Теорема 1 доказана.

Данная теорема при $k=1$ предлагает асимптотическую формулу для $S(n,1)$, совпадающую с формулой, доказанной в [2].

Следствие 1. *Пусть k — некоторое натуральное число, ε — положительное вещественное число. Тогда существует натуральное число N такое, что для любых $n > N$*

$$S^n < (1+\varepsilon)\pi^{-k/2}2^{2^n-nk/2+2k-1}.$$

Доказательство. Воспользуемся неравенством $S^n < S(n,k)$. Согласно теореме 1

$$\lim_{n\to\infty} S(n,k)/2^{2^n}2^{2k-1}\pi^{-k/2}2^{-nk/2} = 1.$$

Следовательно, для любого $\varepsilon > 0$ найдется N такое, что для всех $n > N$

$$S(n,k) < (1+\varepsilon)2^{2k-1}\pi^{-k/2}2^{2^n-nk/2}.$$

Следствие 1 доказано.

Теперь, используя результаты данной работы и работы [5], можно выписать следующую двустороннюю оценку для S^n при произвольном положительном ε и фиксированном k при всех достаточно больших n:

$$\frac{\exp_2(2^n - n^2/2 + n)}{\pi^{n/2}} \leqslant S^n < \frac{(1+\varepsilon)2^{2k-1}\exp(2^n - nk/2)}{\pi^{k/2}}.$$

Список литературы

1. Webster A. F., Tavares S. E., On the design of S-boxes. *Lecture Notes Comput. Sci.* (1986) **218**, 523–534.
2. Cusick T. W., Bounds on the number of functions satisfying the strict avalanche criterion. *Inform. Processing Lett.* (1996) **57**, 261–263.
3. Cusick T. W., Stanica P., Bounds on the number of functions satisfying the strict avalanche criterion. *Inform. Processing Lett.* (1996) **60**, 215–219.
4. Youssef A. M., Tavares S. E., Comment on bounds on the number of functions satisfying the strict avalanche criterion. *Inform. Processing Lett.* (1996) **60**, 271–275.
5. Bliss D. K., A lower bound on the number of functions satisfying the strict avalanche criterion. *Discrete Math.* (1998) **185**, 29–39.
6. O'Connor L., An upper bound on the number of functions satisfying the strict avalanche criterion. *Inform. Processing Lett.* (1994) **52**, 325–327.
7. Бхаттачария Р. Н., Ранга Рао Р., *Аппроксимация нормальным распределением и асимптотические разложения.* Наука, Москва, 1982.

Статья поступила 05.10.2004.