О СВОЙСТВАХ БУЛЕВЫХ ПОЛИНОМОВ, АКТУАЛЬНЫХ ДЛЯ КРИПТОСИСТЕМ

Гордеев Э.Н.[1], Леонтьев В.К.[2], Медведев Н.В.[3]

Булевы функции вообще и булевы полиномы (полиномы Жегалкина, АНФ – алгебраические нормальные формы), в частности, – предмет теоретических и прикладных исследований в различных областях информатики. Свойства булевых полиномов – это классические разделы дискретной математики и комбинаторного анализа. Теоретические основы информационной безопасности включают изучение свойств булевых полиномов в связи с вопросами криптографии. Например, в ряде популярных криптосистем с открытым ключом используются коды Рида–Маллера, а их представление, алгоритмы кодирования и декодирования базируются на булевых полиномах, спектральные свойства определяются количеством нулей полиномов и исследуются с помощью леммы о рандомизации. Известно, что в общем случае задача нахождения числа нулей Z_g полинома $g(x)$ является NP–трудной. Поэтому алгоритмы, учитывающие «комбинаторную структуру полинома», хотя и переборные, представляют прикладной интерес. В работе предлагается такой алгоритма на основе свойств матрицы мономов. Представлена формула для нахождения числа нулей полинома. Приведены формулы для матожидания числа нулей для нескольких классов булевых полиномов, позволяющие путем вариации параметров получить результаты сколь угодно далекие от свойства «сбалансированности». Теоретические результаты работы могут быть основой методик оценки применимости полиномов в различных задачах защиты информации.

***Ключевые слова**: криптосистемы с открытым ключом, булев полином, корни полиномов, NP– трудные задачи, сбалансированность, матожидание.*

DOI: 10.21681/2311-3456-2017-3-63-69

1. Введение

Пусть $B=\{0,1\}$, B^n – n-мерный булев куб, $g(x)$ – булева функция: $B^n \to B$ в базисе $\{1,\wedge,\oplus\}$. Как обычно, $||x||$- норма булевого вектора – это его вес Хэмминга, т.е. число единиц в этом векторе. Весом матрицы из нулей и единиц $C=||c_{ij}||$ - это вес Хэмминга одной строки, которая является суммой строк этой матрицы.

Функцию $g(x)$ можно представить в виде (здесь и далее всюду сложение – это сложение по mod2):

$$g(x)=\sum_{w\in B^n} c_w x^w, \qquad (1)$$

где

$$c_w \in B, w=(w_1,\ldots,w_n)\in B^n, x^w=x_1^{w_1},\ldots,x_n^{w_n}, x_k^{w_k}=\begin{cases} x_k, если & w_k=1 \\ 1, если & w_k=0 \end{cases}$$

Представление (1) называется булевым полиномом, полиномом Жегалкина, алгебраической нормальной формой (АНФ). Очевидно, что число булевых полиномов от n переменных равно 2^{2^n}.

Конъюнкция x^w называется мономом, а число $\deg x^w=\sum_{k=1}^{n} w_k$ называется степенью этого монома. Степенью всего полинома $g(x)$ называется число $\deg g(x)=\max_{c_w=1}\{\deg x^w\}$.

Исследование свойств булевых полиномов является классической задачей дискретной математики и комбинаторного анализа. Булевы полиномы широко применяются, в частности, в криптографии и криптологии. (См., например, [1]-[4]). В 1994 г. Появилась известная криптосистема Сидельникова [1] на основе двоичных кодов Рида-Маллера. Она представляет собой модернизацию системы Мак-Элиса, которая к тому времени уже вскрывалась с полиномиальной трудоемкостью. Появились и другие модернизации и криптосистемы на основе кодов Рида-Маллера. В 2013 году (см. [2]) уже была выявлена уязвимость и этого подхода.

Наиболее удобный способ задания кодов Рида-Маллера (см., например, гл. 13 в классической книге по теории кодирования [5]) – это их представление в виде булевых полиномов определенного вида. Нахождение числа нулей этих полиномов, а также полиномов для вспомогательных конструкций необходимо для исследования спектральных характеристик кодов (см., например, в [5] теорему 5, стр.425). А здесь важным инструментом служит лемма о рандомизации ([5], стр. 360).

При этом булевы полиномы применяются в самых разных криптосистемах и актуальным явля-

1 Гордеев Эдуард Николаевич, д.ф.-м.н., профессор, МГТУ им. Н.Э.Баумана, Москва, Россия. E-mail: werhorn@yandex.ru
2 Леонтьев Владимир Константинович, д.ф.-м.н., профессор, МГТУ им. Н.Э.Баумана, Москва, Россия. E-mail: vkleontiev@yandex.ru
3 Медведев Николай Викторович, к.т.н, доцент, МГТУ им. Н.Э.Баумана, Москва, Россия. E-mail: medvedevnick54@eandex.ru

ется исследование их свойств с точки зрения степени пригодности для прикладных криптосистем. (Выделение нужных подклассов полиномов.) (См., например, [3],[4].)

Можно называть две основные причины актуальности исследования свойств булевых полиномов для криптографии:

1. Сами функции представляются в виде булевых полиномов. (В прикладных пакетах программ обязательно присутствуют программы построения полиномов Жегалкина, преобразования Мебиуса и т.п.).
2. Теоретическая аргументация применения булевых с точки зрения их надежности для криптосистем включает несколько важных критериев (сводку результатов можно найти, например, в [3]): «нелинейность», «сбалансированность» («уравновешенность»), «корреляционная имунность», «r-устойчивость», «алгебраическая имунность» и др. Ряд этих критериев прямо или косвенно связан со свойствами множеств значений булевой функции и свойствами ее АНФ.

По ходу изложения мы прокомментируем связь полученных результатов с упомянутыми критериями.

Полученные в работе результаты имеют и другие применения. В качестве примеров задач дискретного моделирования, где возникает проблема нахождения числа его нулей, можно привести также работы [6]-[8].

В следующем разделе мы приведем сводку некоторых результатов, связанных с числом нулей булева полинома и предложим алгоритм нахождения числа нулей в общем случае.

В третьем разделе описан алгоритма нахождения числа нулей полинома. В четвертом представлены формулы для среднего числа нулей некоторых классов булевых полиномов.

Хотя данная работа носит преимущественно теоретический характер, явные формулы и описанный алгоритм могут быть легко запрограммированы и применены.

2. Базовые понятия и утверждения

Сначала приведем несколько известных базовых результатов. Они будут либо прямо использованы нами, либо необходимы для понимания проблематики.

Утверждение 1. Если степень булева полинома больше двух, то задача нахождения числа его нулей является NP-трудной.

Доказательство можно найти, например, в [6].

Простой переборный алгоритм нахождения числа нулей булевого полинома – перебор по всему булеву кубу – имеет сложность $O(n2^n)$.

Пусть $f(x)=\sum_{w\in B^n} c_w x^w$. Рассмотрим классический частичный порядок на векторах булевого куба:

$$x=(x_1,\dots,x_n)\le y=(y_1,\dots,y_n)\Leftrightarrow x_i\le y_i, i=1,\dots,n.$$

Из того, что $x^w=1$ очевидным образом следует, что $w_i\le x_i$, $i=1,\dots,n$. Отсюда получаем формулу: $f(x)=\sum_{w\le x} c_w$.

Используя формулу Мебиуса (см., например, [6]), получаем классический результат.

Утверждение 2. Справедлива явная формула для c_w:

$$c_w=\sum_{x\le w} f(x). \qquad (2)$$

Пример 1. Пусть $f(x_1,\dots,x_n)=x_1\vee x_2\vee\dots\vee x_n$.

Тогда

$$c_w=\sum_{x\le w} f(x)=\sum_{x\in B^{\|w\|}} f(x)=1\ . \qquad (3)$$

Определение: Многочлен от нескольких переменных называется симметрическим, если он не изменяется ни при какой перестановке неизвестных.

Пример: $f(x_1+x_2+x_3)=x_1+x_2+x_3$, т.к. $(x_1+x_2+x_3)=x_1+x_2+x_3$ не изменился при перестановке переменных.

Следующие n многочленов от n неизвестных называются элементарными симметрическими: $G_1=G_1(x_1,\dots,x_n)=x_1+x_2+\dots+x_n$, $G_2=G_2(x_1,\dots,x_n)=x_1x_2+x_1x_3+\dots+x_{n-1}x_n$, …, $G_n=G_n(x_1,\dots,x_n)=x_1x_2\dots x_{n-1}x_n$.

Пусть $G_k=G_k(x_1,\dots,x_n)$ – k-й элементарный симметрический полином.

Так как в сумме (3) ровно $(2^{\|w\|}-1)$ слагаемых равны единице и, по условию $c_{0\dots0}=0$, то булев полином для нашей функции имеет вид:

$$f(x_1,\dots,x_n)=\sum_{i=1}^{n} x_i\oplus\sum_{1\le i<j\le n} x_ix_j\oplus\dots\oplus x_1x_2\dots x_n=$$

$$=G_1+G_2+\dots+G_n\ .$$

Утверждение 3. Каждый полином степени 1 имеет ровно 2^{n-1} нулей.

Доказательство просто следует из определений полинома, степени полинома, а также из следующего простого факта: для любого $1\le k\le n$ ровно половина всех точек B^n имеет среди своих компонент $\{x_{i_1},x_{i_2},\dots,x_{i_k}\}$ четное число единиц, а половина – нечетное.

Пусть $S_n^k=\sum_{i=0}^{k} C_n^k$ -мощность шара радиуса k из B^n.

Утверждение 4. Число булевых полиномов от n переменных, имеющих степень k, равно $2^{S_n^k} - 2^{S_n^{k-1}}$.

Утверждение 5. Число булевых полиномов, имеющих ровно k мономов , равно $C_{2^n}^k$.

Утверждение 6. Число булевых полиномов, обращающихся в ноль в заданной фиксированной точке булева куба , равно $2^{2^{n-1}}$.

Утверждение 7. Число булевых полиномов степени не больше k, обращающихся в ноль в заданной фиксированной точке булева куба , равно $2^{S_n^k - 1}$.

Утверждение 8. Множество минимальной мощности в B^n такое, что любой булев полином степени не больше k, имеет ноль в этом множестве равно $S_n^k = \sum_{i=0}^{k} C_n^k$.

3. Алгоритм нахождения числа нулей на основе матрицы мономов

Пусть полином $g(x_1,\ldots,x_n)= y_1(x_1,\ldots,x_n)+\ldots+ y_m(x_1,\ldots,x_n)$, где $y_1(x_1,\ldots,x_n),\ldots, y_m(x_1,\ldots,x_n)$ – мономы этого полинома. Заметим, что представление булевой функции в виде булева полинома однозначно. Любая переменная либо входит в заданный моном, либо не входит, поэтому каждому моному можно сопоставить n-мерный характеристический двоичный вектор с единицами на местах тех переменных , которые входят в моном. Взяв эти вектора в виде строк длины n можно построить двоичную матрицу A_g (матрицу мономов) размеров mxn.

Положим по определению, что моному «1» сопоставляется матрица из нулей размером 1xn.

Пример 2. Если $g(x_1,\ldots,x_n)= x_1x_2\ldots x_n$ то $A_g =||11,\ldots,1||$. Если $g(x_1,\ldots,x_n)= x_1+x_2+\ldots+x_n$ то $A_g = \begin{Vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & 0\ldots & 0 \\ . & . & . & . \\ 0 & 0\ldots & 0 & 1 \end{Vmatrix}$. Если $g(x_1,\ldots,x_n)= x_1x_2+ x_2x_3+\ldots+ x_{n-1}x_n$ то $A_g = \begin{Vmatrix} 11 & 0 & \ldots & 0 \\ 0 & 11 & 0\ldots & 0 \\ . & . & . & . \\ 0 & 0\ldots & 0 & 11 \end{Vmatrix}$.

Определение. Весом матрицы называется число ненулевых столбцов этой матрицы.

Определение. Подматрицей исходной матрицы назовем подмножество ее строк.

С матрицей A_g связана матрица $C_A = C_{A(g)} = C_g =||c_{ij}||$, где c_{ij} - число (ixn) подматриц матрицы A_g таких, что любая из этих подматриц имеет вес j. Другими словами, матрица C_g устроена следующим образом c_{ij} – число таких совокупностей из i-мономов множества $y_1(x_1,\ldots,x_n),\ldots, y_m(x_1,\ldots,x_n)$, что объединение мономов каждой совокупности содержит ровно j переменных из множества $x_1,\ldots, x_n$. Эта матрица в дальнейшем называется матрицей весов матрицы A_g. (Вариабельность обозначений введена для удобства минимизации индексации. Везде из контекста ясно, о чем идет речь.)

Пример 3. Если $g(x_1,\ldots,x_n)= x_1+x_2+\ldots+x_n$ то $A_g = \begin{Vmatrix} 1 & 0 & \ldots & 0 \\ 0 & 1 & 0\ldots & 0 \\ . & . & . & . \\ 0 & 0\ldots & 0 & 1 \end{Vmatrix}$. В этом простейшем случае число c_{ij} равно 0 при i≠j, а при i=j оно равно числу мономов, содержащий ровно j переменных, т.е. $C_g = \begin{Vmatrix} C_n^1 & 0 & \ldots & 0 \\ 0 & C_n^2 & 0\ldots & 0 \\ . & . & . & . \\ 0 & 0\ldots & 0 & C_n^n \end{Vmatrix}$. Если $g(x_1,\ldots,x_n)= x_{1+}x_1x_2\ldots x_{n-1}x_n$ то $A_g == \begin{Vmatrix} 10\ldots & 0 \\ 11\ldots & 1 \end{Vmatrix}$, а $C_g == \begin{Vmatrix} 10\ldots & 1 \\ 00\ldots & 1 \end{Vmatrix}$.

Если A_g состоит из одних единиц, то c_{ij} равно 0 при j≠n, а в остальных случаях $c_{in} = C_n^i$. Поэтому $C_g = \begin{Vmatrix} 0 & 0 & \ldots & C_n^1 \\ 0 & 0 & 0\ldots & C_n^2 \\ . & . & . & . \\ 0 & 0\ldots & 0 & C_n^n \end{Vmatrix}$.

Напомним, что Z_g - число нулей булева полинома g(x), а ||g(x)|| - число единиц полинома g(x).

Приведем следующее утверждение, проливающее свет на смысл матрицы C_g.

Теорема 1. Справедлива формула для числа нулей булева полинома

$$Z_g = 2^n + \sum_{s=1}^{r} (-1)^s \sum_{r=1}^{n} c_{sr} 2^{n+s-r-1}. \qquad (4)$$

Доказательство теоремы приведено в [7].

Но в теореме речь идет о нулях. А для единиц полинома очевидно, что справедливо ее следствие.

Следствие 1.

$$\|g(x)\| = \sum_{s=1}^{r} (-1)^{s+1} \sum_{r=1}^{n} c_{sr} 2^{n+s-r-1}. \qquad (5)$$

Очевидно, что

$$\sum_{r=1}^{n} c_{sr} = C_k^s, s = 1,\ldots,k. \qquad (6)$$

Теперь мы подошли к практическому применению вышеприведенных теоретических построений.

Начнем с формулы (4). С ее помощью можно построить алгоритм для нахождения числа нулей булевого полинома.

Для начала вспомним проверочную матрицу H_k размеров $kx(2^k-1)$ кода Хэмминга, в которой столбцы – это двоичные записи чисел от 1 до 2^k-1.

Определение. Назовем $\varphi(A)$ φ-преобразованием бинарной матрицы A размеров kxn матрицу $C_A^1 = H_k^T A^1$, где A^1 получается из A приписыванием слева столбца из единиц.

Лемма 1. Строки матрицы $\varphi(A)$, имеющие первым элементом число r, порождают r-ю строку матрицы C_A в том смысле, что элемент c_{rj} равен числу строк $\varphi(A)$, имеющих вес, равный j, т.е. ровно с j ненулевыми элементами. При этом элементы первого столбца C_A^1 не учитываются.

Доказательство. Из определения матриц H_k и $\varphi(A)$ следует, что строки матрицы $C_A^1 = H_k^T A^1$ - это суммы подмножеств строк A^1, взятых по одной, по две, по три и т.д. Отсюда следует, что для той строки C_A^1, которая образована суммой r строк матрицы A^1, первым элементом будет число r.

Лемма доказана.

Пример 4. Пусть $A = \begin{Vmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{Vmatrix}$. Тогда k=3 и

$$\begin{Vmatrix} 11 & 0 & 0 \\ 11 & 1 & 0 \\ 11 & 1 & 1 \end{Vmatrix}, \tilde{N}_A^1 = H_3^T A^1 = \begin{Vmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \\ 2 & 2 & 2 & 1 \\ 1 & 1 & 0 & 0 \\ 2 & 2 & 1 & 1 \\ 2 & 2 & 1 & 0 \\ 3 & 3 & 2 & 1 \end{Vmatrix}.$$

Первая, вторая и четвертая строки начинаются с 1. Без учета первого столбца имеем среди них по одной строке веса 1, 2 и 3. Третья, пятая и шестая строки начинаются с 2. Без учета первого столбца имеем среди них одну строку веса 2 и две строки веса 3. Седьмая строка начинаются с 3. Без учета первого столбца имеем одну строку веса 3. Отсюда следует $C_A = \begin{Vmatrix} 1 & 1 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{Vmatrix}$.

Из теоремы 1 и леммы 1 следует алгоритм нахождения числа нулей булевого полинома.

АЛГОРИТМ 1

Шаг 1. По $g(x)$ заданному полиному $g(x)$ строим матрицу мономов A_g. Трудоемкость $O(mn)$.

Шаг 2. По матрице A= Ag строим $C_A^1 = H_k^T A^1$. Трудоемкость $O(nm2^m)$.

Шаг 3. По формуле (4) находим число нулей полинома. Трудоемкость $O(n^2)$.

Теорема 2. Алгоритм 1 находит число нулей булева полинома за время $O(n^2+nm2^m)$.

Доказательство просто следует из описания алгоритма.

По сравнению с тривиальным алгоритмом полного перебора по всем векторам булева куба алгоритм имеет преимущества для случая небольшого числа мономов $m=o(n)$ или даже $m<cn$ при малых константах с за счет учета комбинаторики задачи при большой доле нулей в матрице мономов.

Обратим внимание на то, что вычисление упомянутых выше примеров характеристик булевых функций: «нелинейность», «сбалансированность» («уравновешенность»), «корреляционная имунность », «r-устойчивость», «алгебраическая имунность» - это не только NP-трудные задачи, но и задачи, лежащие за пределами класса NP. На этом фоне описанный алгоритм представляется сравнительно эффективным.

Подчеркнем два методологических аспекта, связанных с применением булевых функций в криптосистемах.

Булева функция с нужными свойствами конструируется. Для этого вычисляются подклассы «хороших» функций (например, известные бент-функции).

Булева функция выбирается с помощью некоторой эвристики, а затем вычисляются ее характеристики.

Все было бы хорошо, если бы не противоречивость критериев. Высокая нелинейность означает несбалансированность, а сбалансированность низкую нелинейность. Удобное и эффективное описание «хорошего» подкласса уменьшает его мощность и повышает вероятность распознавания, а обширные подклассы описывать неудобно и приходится проверять предполагаемые (с большой вероятностью) характеристики эвристически выбранной функции.

В свете этого замечания предложенный алгоритм непосредственно может быть использован для проверки «сбалансированности», а также применяться для анализа «корреляционной имунности » и «r-устойчивости».

В следующих двух разделах мы представим инструменты как для описания подклассов «хороших» функций, так и для описания эвристик их выбора.

4. Формулы для среднего числа нулей булевых полиномов

Пусть $w(f_n)$ – число единиц булевой функции $f(x)=f(x_1,\ldots,x_n)$. Если $w(f_n) = 2^{n-1}$, то функция f называется уравновешенной или равновероятной, так как она при случайном выборе аргумента x принимает оба значения с одинаковой вероятностью: $Pr[f(x)=1] = Pr[f(x)=0]$, где по определению $Pr[f(x)=1] = w(f_n)/ 2^n$ и $Pr[f(x)=0] = 1 - Pr[f(x)=1]$. В криптографических приложениях очень часто рассматриваются именно уравновешенные функции.

Таким образом, булев полином при $Z_g = 2^n / 2$ является уравновешенной (сбалансированной) функцией.

Для исследования числа нулей Z_g булева полинома g(x) рассмотрим сумму характеров \

$$S_g(x) = \sum_{x\in B^n} (-1)^{g(x)} .$$

Смысл этого рассмотрения в том, что по ней автоматически можно вычислить Z_g. Действительно

$$S_g(x) = \sum_{g(x)=0, x\in B^n} (-1)^{g(x)} + \sum_{g(x)=1, x\in B^n} (-1)^{g(x)} =$$

$$= Z_g - (2^n - Z_g) = 2Z_g - 2^n .$$

То есть

$$Z_g = \frac{1}{2}(2^n + S_g) . \qquad (7)$$

Но не это главное. Дело в том, что сумму характеров вычислять, как правило, легче, чем число нулей полинома.

Ну а что можно сказать о среднем числе корней $\overline{Z}_g(n)$ булева полинома от n переменных? Очевидно, что $\overline{Z}_g(n) = 2^n / 2$. Действительно, для среднего числа единиц (так удобнее) $\overline{N}_g(n)$ имеем:

$$\overline{N}_g(n) = \frac{1}{2^{2^n}} \sum_g N_g = \frac{1}{2^{2^n}} \sum_{x\in B^n} \sum_g g(x) .$$

Внутренняя сумма равна числу полиномов, обращающихся в единицу в фиксированной точке x булева куба, а это число решений линейного равнения

$$c_0 + \sum_{i=1}^{n} c_i x_i + \sum_{i<j} c_{ij} x_i x_j + \ldots = 1$$

с 2^n неизвестными $\{c_{i_1},\ldots,c_{i_k}\}$, k=0,…,n. А это уравнение имеет $2^{2^n} / 2$ решений (по аналогии с Утверждением 3). Отсюда и получаем $\overline{Z}_g(n) = 2^n / 2$.

Но, может быть степень полинома поможет получить что-то интересное? Ведь нелинейность и сбалансированность противоречит друг другу. Пусть среднее числе корней $\overline{Z}_g(n,k)$ булева полинома степени k от n переменных.

Число полиномов степени не выше k, обращающихся в единицу в фиксированной точке x булева куба, равно числу решений линейного равнения

$$c_0 + \sum_{i=1}^{n} c_i x_i + \ldots + \sum_{i_1<i_2,\ldots,i_k} c_{i_1 i_2 \ldots i_k} x_{i_1} x_{i_2} \ldots x_{i_k} = 1$$

с $\sum_{i=0}^{k} C_n^i$ неизвестными . А это уравнение имеет $2^{\sum_{i=0}^{k} C_n^i} / 2$ решений. Отсюда получаем, что число полиномов степени ровно k, обращающихся в единицу в фиксированной точке x булева куба, равно числу $2^{\sum_{i=0}^{k} C_n^i} / 2 - 2^{\sum_{i=0}^{k-1} C_n^i} / 2$. А так как общее число полиномов степени k (Утверждение 4) равно $2^{\sum_{i=0}^{k} C_n^i} - 2^{\sum_{i=0}^{k-1} C_n^i}$, то вновь получаем, что $\overline{Z}_g(n,k) = 2^n / 2$.

Теперь мы представим два класса булевых полиномов, которые позволят обосновать эвристики подбора несбалансированных булевых функций.

Пусть $L_{n,k,p}$ – класс булевых полиномов от n переменных ровно с k мономами, каждый из которых имеет степень p.

Пусть теперь полином из этого класса $g(x_1,\ldots, x_n)= y_1(x_1,\ldots,x_n)+\ldots+ y_k(x_1,\ldots,x_n)$ выбирается случайным образом путем выбора мономов степени p из множества мономов степени p с равномерным распределением. Рассмотрим случайную величину

$$\zeta(x) = \sum_{x\in B^n} (-1)^{y_1+y_2+\ldots+y_k} \qquad (8)$$

И ее матожидание $\mu_\zeta(k)$. Пусть также $\overline{Z}(n,k,p)$ - среднее число нулей полинома из класса $L_{n,k,p}$. То есть

$$\overline{Z}(n,k,p) = \frac{1}{(C_n^p)^k} \sum_{g\in L(n,k,p)} Z_g .$$

Теорема 3. Справедлива формула

$$\overline{\zeta}(x) = \sum_{m=p}^{n} C_n^m (1 - 2C_m^p / C_n^p)^k + 2^{n-1} - \sum_{m=p}^{n} C_n^m . \qquad (9)$$

Доказательство приведено в [9].

Следствие 2. Справедлива формула

$$\overline{Z}(n,k,p) = \frac{1}{2}\sum_{m=p}^{n} C_n^m (1 - 2C_m^p / C_n^p)^k + 2^{n-1} - \frac{1}{2}\sum_{m=p}^{n} C_n^m . \qquad (10)$$

Мы получили, что матожидание числа нулей отличается от $\overline{Z}_g(n,p) = 2^n / 2$.

Мы можем варьировать это отклонение от «сбалансированности» подбором параметров p и k.

Но теперь построим еще более обширный класс полиномов как с более тонкими настройками параметров, так и с большими затруднениями для угадывания выбранного полинома.

Рассмотрим общую ситуацию. Пусть теперь $L(n,p_0,p_1,\ldots,p_n)$ – класс булевых полиномов от n переменных ровно с p_0 мономами степени 0, ровно с p_1 мономами степени 1,…, ровно с p_n мономами степени n. То есть здесь $k= p_0+p_1+\ldots+p_n$.

Пусть теперь полином из этого класса $g(x_1,\ldots, x_n)= y_1(x_1,\ldots,x_n)+\ldots+ y_k(x_1,\ldots,x_n)$ выбирается случайным образом путем выбора мономов степени нужной степени p из множества мономов степени p с равномерным распределением. Рассмотрим ту же случайную величину

$$\zeta(x) = \sum_{x\in B^n} (-1)^{y_1+y_2+\ldots+y_k}$$

и ее матожидание.

В [9] доказана следующая теорема.

Теорема 4. Справедлива формула

$$\overline{\zeta}(x) = \sum_{m=0}^{n} C_n^m (1-2C_m^0 / C_n^0)^{p_0} (1-2C_m^1 / C_n^1)^{p_1} \ldots (1-2C_m^n / C_n^n)^{p_n}. \quad (11)$$

Пусть $\overline{Z}(p_0,\ldots,p_n)$ - матожидание числа нулей полинома из класса $L(n,p_0,p_1,\ldots,p_n)$. Из теоремы 4 и формулы (7) следует.

Следствие 3. Справедлива формула

$$\overline{Z}(p_0,\ldots,p_n) = \frac{1}{2}(2^n + \sum_{m=0}^{n} C_n^m (1-2C_m^0 / C_n^0)^{p_0} (1-$$

$$-2C_m^1 / C_n^1)^{p_1} \ldots (1-2C_m^n / C_n^n)^{p_n}) \ . \quad (12)$$

Здесь мы вновь получили, что матожидание числа нулей отличное от $\overline{Z}_g(n,p) = 2^n / 2$. Мы можем варьировать это отклонение от «сбалансированности» подбором параметров $p_0,\ldots,p_n$.

Пример 5. Пусть p=n, k=1. То есть мы имеем единственный полином степени n $g(x_1,\ldots,x_n)= x_1x_2\ldots x_n$. Очевидно, что число его нулей равно 2^n-1. Именно это и дают наши соотношения.

Мы имеем

$$\overline{\zeta}(x) = \sum_{m=p}^{n} C_n^m (1-2C_m^n / C_n^n)^1 + 2^n - \sum_{m=n}^{n} C_n^m = 2^n - 2,$$

а $\overline{Z}(n,k,p)=2^n$-1.

Замечание. Обратим внимание на сходство полученных нами формул (10) и (12) с некоторыми классическими соотношениями анализа булевых функций в криптографии. Одним из наиболее распространенных инструментов здесь является нахождение спектра Уолша-Адамара булевой функции $f(x_1,\ldots,x_n)$. Преобразование Уолша-Адамара для вектора y из B^n:

$$\hat{f}(y) = \sum_{x\in B^n} (-1)^{f(x)\oplus(x,y)}$$

дает 2^n чисел – компонент спектра функции. (Ясно, что и время, и память нахождения спектра одной функции экспоненциальны по n в общем случае). (Обозначим через $W(f)$ максимальную по модулю компоненту спектра.) Через эти числа выражаются многие криптографические характеристики функций (кстати, в большинстве случаев, см., например, [3], с ними работают как с булевыми полиномами): сбалансированность, порядок корреляционной имунности, нелинейность и др. Например, нелинейность N_f (расстояние до класса аффинных функций):

$$N_f = 2^n / 2 - W(f) / 2. \quad (14)$$

Замечание. Известны простые оценки для числа нулей, связанные еще с одним широко применяющемся в криптографии объектом: аннигилятором – ненулевым полиномом, при умножении на который, функция становится тождественным нулем. Если d- минимальная степень аннигилятора, то

$$\sum_{i=0}^{d-1} C_n^i \le 2^n - Z_f \le \sum_{i=0}^{n-d} C_n^i .$$

5. Выводы

В работе дана явная формула для числа нулей булевых полиномов в общем случае.

Обоснован и построен алгоритм нахождения числа нулей булевых полиномов в общем случае, позволяющий использовать «комбинаторную структуру полинома» для вычисления количества его нулей.

Описаны два класса булевых полиномов, для которых матожидание числа нулей может быть как угодно отличным от 2^{n-1}. Приведены явные формулы для среднего числа нулей в обоих случаях.

Эти формулы и алгоритмы могут быть использованы для анализа криптосистем с открытым ключом на базе кодов Рида-Маллера, а также в других задачах криптографии, которые используют булевы полиномы. В работе проиллюстрирована эта связь на примерах известных криптографических характеристик булевых функций: сбалансированности, нелинейности, алгебраической имунности, корреляционной имунности.

***Рецензент**: Цирлов Валентин Леонидович, кандидат технических наук, доцент МГТУ им. Н.Э. Баумана. E-mail: v.tsirlov@cnpo.ru*

Литература:

1. Сидельников В.М. Открытое шифрование на основе двоичных кодов Рида-Маллера // Дискретная математика. 1994. № 2 (4). С. 3-20.
2. Чижов И.В., Бородин М.А. Уязвимость криптосистемы Мак-Элиса, построенной на основе двоичных кодов Рида-Маллера // Прикладная дискретная математика. Приложение. 2013. № 6. С.48-49.
3. Панкратова И.А. Булевы функции в криптографии: учебное пособие. Томск.: Томский Университет, 2014. 88 с.
4. Сизоненко А.Б. Параллельная реализация криптографических блоков подстановок и перестановок арифметическими полиномами // Доклады ТУСУРа. 2012. Т.26. № 2. С. 140-144.
5. Мак-Вильямс Дж., Слоэн Н. Теория кодов, исправляющих ошибки. М.: Связь, 1979. 745 с.
6. Леонтьев В.К. Комбинаторика и информация. Часть 1. Комбинаторный анализ. М.: МФТИ, 2015. 174 с.
7. Леонтьев В.К., Морено О. О нулях булевых полиномов // Ж. вычисл. матем. и матем. физ. 1998. Т. 38. № 9. С. 1608–1615.
8. Леонтьев В.К. Симметрические булевы полиномы. // Ж. вычисл. матем. и матем. физ. 2010. Т. 50. № 8. С. 1520-1531.
9. Леонтьев В.К., Гордеев Э.Н. О числе нулей булевых полиномов. // Ж. вычисл. матем. и матем. физ. 2017. В печати.

THE PROPERTIES OF BOOLEAN POLYNOMIALS THAT ARE RELEVANT TO CRYPTOSYSTEMS

E. Gordeev [4], V. Leontiev[5], N. Medvedev [6]

Boolean functions and Boolean polynomials (polynomials Zhegalkin, ANF – algebraic normal forms), in particular, is the subject of theoretical and applied research in various areas of computer science. Properties of Boolean polynomials is a classical topics of discrete mathematics and combinatorial analysis. Theoretical foundations of information security include the study of properties of Boolean polynomials in connection with the issue of cryptography. For example, in some popular public key cryptosystems are used Reed–Muller codes and their representation, the algorithms of encoding and decoding are based on Boolean polynomials, spectral properties are determined by the number of zeros of polynomials and investigated with the help of the Lemma on the randomness. It is known that in General case the problem of finding the number of zeros of the polynomial Z_g of the polynomial $g(x)$ is NP–hard.. Therefore, algorithms that take into account «the combinatorial structure of the polynomial», though exhaustive, are of applied interest. This paper proposes such an algorithm based on the properties of the matrix of monomials. Presents a formula for finding the number of zeros of the polynomial. Formulas for the expected value of the number of zeros for several classes of Boolean polynomials. The theoretical results can be the basis of methods to assess the applicability of the polynomials in various tasks of information security.

***Keywords**: cryptosystem with a public key, Boolean polynomial, roots of polynomial, NP – hard problem, expected value of the number of zeros.*

References

1. Sidelnikov V. M. Public key cryptosystems based on binary reed-Muller codes // Discrete mathematics. 1994. No. 2 (4). P. 3-20.
2. Chizhov I. V., Borodin M. A. The vulnerability of the cryptosystem Mac-Elys based on binary reed-Muller codes // Applied discrete mathematics. App. 2013. No. 6. P. 48-49.
3. Pankratova I.A. Boolean functions in cryptology. Tomsk: Tomskiy Universitet, 2014. 88 p.
4. Sizonenko A. B. Parallel implementation of cryptographic units of substitutions and permutations arithmetic with polynomials // Proceedings of TUSUR. 2012. T. 26. No. 2. P. 140-144.
5. Mac Williams J., Sloan, N. The theory of error-correcting codes. M.: Sviaz, 1979. 745 p.
6. Leont'ev V. K. Combinatorics and information. Part 1. Combinatorial analysis. Moscow: MIPT, 2015. 174 p.
7. Leont'ev V. K., Moreno O. On the zeros of Boolean polynomials // Zh. Vychisl. matem. and matem. Fiz.. Vol. 38. No. 9. P. 1608-1615.
8. Leont'ev V. K. Symmetric Boolean polynomials. // Zh. Vychisl. matem. and matem. Fiz. 2010. V. 50. No. 8. P. 1520-1531.
9. Leont'ev V. K., Gordeev E.N. Number of zeroes of Boolean polynomials. // Zh. Vychisl. matem. and matem. Fiz. 2017. In preparatin.

4 Eduard Gordeev , Dr.Sc. (in Math,), professor, BMSTU, Moscow, Russia, E-mail: werhorn@yandex.ru

5 Vladimir Leontiev, Dr.Sc. (in Math,), professor, BMSTU, Moscow, Russia, E-mail: vkleontiev@yandex.ru

6 Nikolay Medvedev, Ph.D. (in Tech.), professor, BMSTU, Moscow, Russia, E-mail: medvedevnick54@yandex.ru