Общероссийский математический портал

А. Ю. Нестеренко, Об одном семействе универсальных функций хеширования, *Матем. вопр. криптогр.*, 2015, том 6, выпуск 3, 135–151

DOI: <https://doi.org/10.4213/mvk164>

Использование Общероссийского математического портала Math-Net.Ru подразумевает, что вы прочитали и согласны с пользовательским соглашением
<http://www.mathnet.ru/rus/agreement>

Параметры загрузки:

IP: 18.209.158.208

26 апреля 2023 г., 05:05:44

УДК: 519.719.2

Об одном семействе универсальных функций хеширования

А. Ю. Нестеренко

Национальный исследовательский университет «Высшая школа экономики»,
Москва

Получено 19.VI.2014

Построено новое семейство сжимающих отображений с помощью нескольких биективных преобразований и отображений с заданными свойствами. Доказывается, что все функции из данного семейства являются универсальными функциями хеширования. Приводятся явные примеры функций из рассматриваемого семейства, которые могут использоваться при построении прикладных криптографических алгоритмов.

Ключевые слова: сжимающие отображения, универсальные функции хеширования

On a family of universal hash functions

A. Yu. Nesterenko

National Research University «Higher School of Economics», Moscow

Abstract. We construct a new family of compressing mappings by means of superposition of several bijective mappings and mappings with specified properties. All functions in this family are proved to be universal hash functions. Concrete examples of functions from the family which are suitable for cryptographic applications are supplied.

Key words: compressing mappings, universal hash functions

Citation: *Mathematical Aspects of Cryptography*, 2015, vol. 6, no. 3, pp. 135–151 (Russian).

Рассмотрим конечные множества: множество сообщений $\mathcal{S}$, множество ключей $\mathcal{K}$ и множество $\mathcal{A}$, элементы которого назовем кодами аутентичности. Будем считать, что мощности множеств $\mathcal{S}$ и $\mathcal{K}$ не меньше мощности множества $\mathcal{A}$, т. е. $|\mathcal{S}| \geq |\mathcal{A}|$, $|\mathcal{K}| \geq |\mathcal{A}|$. Ключевой функцией хеширования будем называть отображение

$$h(k, x) : \mathcal{K} \times \mathcal{S} \rightarrow \mathcal{A}, \quad k \in \mathcal{K}, x \in \mathcal{S}. \quad (1)$$

В последние десятилетия ключевые функции хеширования активно применяются в средствах защиты информации, в частности в криптографических протоколах распределения ключей, для подтверждения целостности и аутентичности передаваемых сообщений, для выработки псевдослучайных значений, для выработки одноразовых ключей и т. д.

Известно несколько подходов к построению ключевых функций хеширования: на основе алгоритмов блочного шифрования [9, 16], применения криптографических функций хеширования [7, 13] и применения так называемых универсальных функций хеширования [1, 19]. В настоящей работе автором рассматривается подход, основанный на применении универсальных функций хеширования и ведущий свое начало от работ Картера и Вегмана [11, 22], а также Стинсона [20].

Для нас важно, что универсальные функции хеширования обладают двумя свойствами: вероятность выбора случайного сообщения, на котором функция принимает заданное значение, постоянна и не зависит от выбранного ключа. Аналогично, вероятность выбора случайного ключа, на котором функция принимает заданное значение, постоянна и не зависит от выбора сообщения. Указанные свойства не позволяют реализовывать атаки, основанные на случайном опробовании сообщений и соответствующих им кодов аутентичности. Это соображение позволяет применять универсальные функции хеширования в качестве основы для построения преобразований, удовлетворяющих криптографическим требованиям, накладываемым на ключевые функции хеширования, см. [6].

Настоящая статья построена следующим образом. Вначале даются необходимые определения и примеры универсальных функций хеширования, которые можно отнести к одному классу. Далее вводится новое семейство универсальных функций хеширования из данного класса, построенных при помощи некоторого произвольного биективного отображения f и отображения g с заданными свойствами. Мы исследуем свойства введенного семейства функций и в конце статьи в явном виде предъявляем биективное

отображение f и предлагаем конкретную универсальную функцию хеширования, которая может быть использована при реализации криптографических приложений.

1. Универсальные функции хеширования

Дадим необходимые определения.

Определение 1. (см. [11, 20]) Отображение $h(k, x)$ называют ε -почти универсальной функцией хеширования (ε -almost universal, ε -AU), если найдется такое действительное число ε , что $0 < \varepsilon \leq 1$, и для любых двух различных элементов $x, y \in \mathcal{S}$ выполнена оценка

$$|\{k \in \mathcal{K} : h(k, x) = h(k, y)\}| \leq \varepsilon |\mathcal{K}|.$$

Если выполнено равенство $\varepsilon = |\mathcal{A}|^{-1}$, то отображение $h(k, x)$ называют универсальной функцией хеширования.

Из данного определения следует, что для универсальной функции хеширования при случайном выборе ключа k вероятность появления коллизии на двух различных сообщениях x, y не превосходит величины $|\mathcal{A}|^{-1}$.

Определение 2. (см. [11, 20]) Отображение $h(k, x)$ называют ε -почти строго универсальным (ε -almost strong universal, ε -ASU), если выполнены следующие условия:

- 1) для любых $x \in \mathcal{S}$ и $a \in \mathcal{A}$ выполнено равенство

$$|\{k \in \mathcal{K} : h(k, x) = a\}| = \frac{|\mathcal{K}|}{|\mathcal{A}|},$$

- 2) для любых различных $x, y \in \mathcal{S}$ и для любых $a, b \in \mathcal{A}$ выполнена оценка

$$|\{k \in \mathcal{K} : h(k, x) = a, h(k, y) = b\}| \leq \varepsilon \frac{|\mathcal{K}|}{|\mathcal{A}|}.$$

Если $\varepsilon = |\mathcal{A}|^{-1}$, то отображение $h(k, x)$ называют строго универсальной функцией хеширования.

Для ε -ASU функций хеширования значение ε мажорирует величину условной вероятности появления события $h(k, y) = b$ при выполнении события $h(k, x) = a$.

Остановим внимание на нескольких универсальных функциях хеширования вида (1), которые отнесем к одному классу. Для каждой из этих функций из данного класса в соответствующих работах доказано, что она является ε -ASU при некотором значении ε . Будем считать, что множества $\mathcal{S}$ и $\mathcal{K}$ совпадают и состоят из векторов размерности t , т. е.

$$\mathcal{S} = \{(x_1, \dots, x_t)\}, \quad \mathcal{K} = \{(k_1, \dots, k_t)\},$$

где координаты векторов принадлежат некоторому кольцу $\mathcal{R}$.

1. ММН, см. [14]. Положим $\mathcal{R} = \mathcal{A} = \mathbb{Z}_{2^{32}}$, $p = 2^{32} + 15$ — простое число. Тогда

$$h(k, x) \equiv \left(\left(\sum_{n=1}^t k_n x_n \pmod{2^{64}} \right) \pmod{p} \right) \pmod{2^{32}}.$$

2. Square Hash, см. [12]. Пусть $w \geq 2$ — натуральное число, p — наибольшее простое число, меньшее 2^w . Положим $\mathcal{R} = \mathcal{A} = \mathbb{Z}_{2^w}$. Тогда

$$h(k, x) \equiv \sum_{n=1}^t (k_i + x_i)^2 \pmod{p}.$$

3. NMH, см. [14, 22]. Пусть $t = 2l$, $l \in \mathbb{N}$. Положим $\mathcal{R} = \mathcal{A} = \mathbb{Z}_{2^{32}}$; $p = 2^{32} + 15$ — простое число. Тогда

$$h(k, x) \equiv \left(\sum_{n=1}^l (k_{2n-1} + x_{2n-1})(k_{2n} + x_{2n}) \pmod{p} \right) \pmod{2^{32}}.$$

4. NH (алгоритм UMAC), см. [8, 17]. Пусть $w \in \mathbb{N}$, $w \geq 2$, $t = 2l$, $l \in \mathbb{N}$. Положим $\mathcal{R} = \mathbb{Z}_{2^w}$, $\mathcal{A} = \mathbb{Z}_{2^{2w}}$. Тогда

$$h(k, x) \equiv \sum_{n=1}^l z_{2n-1} z_{2n} \pmod{2^{2w}},$$

где $z_n \equiv k_n + x_n \pmod{2^w}$.

5. Badger, см. [10]. Пусть $t = 4l$, $l \in \mathbb{N}$. Положим $\mathcal{R} = \mathbb{Z}_{2^{32}}$, $\mathcal{A} = \mathbb{Z}_{2^{64}}$. Тогда

$$h(k, x) \equiv \sum_{n=1}^l (z_{4n-3} z_{4n-2} + v_n) \pmod{2^{64}},$$

где $z_n \equiv k_n + x_n \pmod{2^{32}}$, $v_n = x_{4n-1} + 2^{32} x_{4n}$.

Отметим, что все перечисленные отображения используют операцию умножения, которая обладает рядом свойств, снижающих криптографические свойства функции $h(k, x)$. Проиллюстрируем это на примере функции NH (алгоритм UMAC). Зафиксируем произвольную пару значений $x_{2n-1}, x_{2n} \in \mathbb{Z}_{2^w}$, вычет $a \in \mathbb{Z}_{2^{2w}}$ и будем считать, что для некоторого индекса n выполнено равенство

$$(k_{2n-1} + x_{2n-1} \pmod{2^w})(k_{2n} + x_{2n} \pmod{2^w}) \equiv a \pmod{2^{2w}}. \quad (2)$$

Тогда имеют место следующие свойства.

1. *Симметричность.* Если ключи k_{2n-1}, k_{2n} удовлетворяют равенству (2), то ключи

$$\begin{aligned} k'_{2n-1} &\equiv k_{2n} + x_{2n} - x_{2n-1} \pmod{2^w}, \\ k'_{2n} &\equiv k_{2n-1} + x_{2n-1} - x_{2n} \pmod{2^w} \end{aligned}$$

также удовлетворяют равенству (2). Очевидно, что если величины x_{2n-1} и x_{2n} различны, то пары ключей (k_{2n-1}, k_{2n}) и (k'_{2n-1}, k'_{2n}) также различны.

2. *Неоднозначность разложения числа a на множители.* Пусть a есть составное число и p, q есть какая-либо пара делителей числа $a = pq$, удовлетворяющая условиям $p < 2^w, q < 2^w$. Тогда, с учетом свойства 1, каждой такой паре делителей p, q будут соответствовать две пары ключей k_{2n-1}, k_{2n} , для которых выполнено равенство (2).

Например, для $w = 4, x_{2n-1} = 0, x_{2n} = 0$ и $a = 12$ ключевыми парами, удовлетворяющими сравнению

$$(k_{2n-1} + x_{2n-1})(k_{2n} + x_{2n}) \equiv 12 \pmod{2^8 = 256},$$

будут $(1, 12), (12, 1), (2, 6), (6, 2), (3, 4), (4, 3)$.

3. *Существование запретных значений.* Если величина a является простым числом в интервале $2^w < a < 2^{2w}$, то она не может являться произведением двух чисел, меньших 2^w .
4. *Повышенное число нулей (проблема нуля).* Значение $a = 0$ достигается на значительно большем числе пар (k_{2n-1}, k_{2n}) , чем все остальные значения a . Действительно, если один из сомножителей в произведении (2)

равен нулю, то значение второго сомножителя может принимать любое допустимое значение. Так, в приведенном выше примере равенство

$$(k_{2n-1} + x_{2n-1})(k_{2n} + x_{2n}) \equiv 0 \pmod{2^8 = 256}$$

достигается на следующих ключевых парах

$$\underbrace{(0, 0), (0, 1), \dots, (0, 15), (1, 0), (2, 0), \dots, (15, 0)}_{31 \text{ пара}}.$$

Данное свойство впервые использовалось в работе [15] для определения множества так называемых слабых ключей. Позднее автором в работе [3] был предложен алгоритм, использующий данное свойство и позволяющий строить коллизии с меньшей трудоемкостью, чем у метода, основанного на «парадоксе дней рождений».

2. Новое семейство универсальных функций

Далее строится семейство универсальных функций хеширования, свободное от перечисленных выше «слабостей».

Пусть w, l, s — произвольные натуральные числа. Определим множество сообщений $\mathcal{S}$ и множество ключей $\mathcal{K}$:

$$\mathcal{S} = \{(x_1, \dots, x_{2l})\}, \quad \mathcal{K} = \{(k_1, \dots, k_{2sl})\},$$

где координаты векторов $x_1, \dots, x_{2l}, k_1, \dots, k_{2sl} \in \mathbb{Z}_{2^w}$. Определим множество кодов аутентичности $\mathcal{A}$ равенством $\mathcal{A} = \mathbb{Z}_{2^{2w}}$.

Рассмотрим отображение

$$g(k_1, \dots, k_s, x_1) : \mathbb{Z}_{2^w} \times \dots \times \mathbb{Z}_{2^w} \rightarrow \mathbb{Z}_{2^w}, \quad (3)$$

удовлетворяющее следующим условиям.

1. При фиксированном наборе значений $k_1, \dots, k_s \in \mathbb{Z}_{2^w}$ отображение

$$g(k_1, \dots, k_s, x_1) = g(x_1) : \mathbb{Z}_{2^w} \rightarrow \mathbb{Z}_{2^w}$$

является биективным отображением $\mathbb{Z}_{2^w}$ в себя.

2. При фиксированном значении $x_1 \in \mathbb{Z}_{2^w}$ и любом значении $z \in \mathbb{Z}_{2^w}$ уравнение $g(k_1, \dots, k_s, x_1) = z$ имеет ровно $2^{w(s-1)}$ различных решений относительно неизвестных $k_1, \dots, k_s$.

3. Зафиксируем произвольную пару $x_1, y_1 \in \mathbb{Z}_{2^w}$ и будем считать, что вычеты $z_1, u_1 \in \mathbb{Z}_{2^w}$ пробегают множество всех возможных значений. Тогда суммарное число решений системы сравнений

$$\begin{cases} g(k_1, \dots, k_s, x_1) \equiv z_1 \pmod{2^w}, \\ g(k_1, \dots, k_s, y_1) \equiv u_1 \pmod{2^w} \end{cases} \quad (4)$$

относительно неизвестных $k_1, \dots, k_s$ в точности равно 2^{ws} .

Отметим, что при фиксированных значениях $z_1, u_1 \in \mathbb{Z}_{2^w}$ система сравнений (4) может иметь от нуля до $2^{w(s-1)}$ решений относительно неизвестных $k_1, \dots, k_s$.

Можно привести несколько примеров отображения (3), удовлетворяющих приведенным требованиям. Так, при $s = 1$ таким является отображение

$$g(k_1, x_1) = k_1 + x_1 \pmod{2^w},$$

а при $s = 2$ — отображение

$$g(k_1, k_2, x_1) = k_1 + k_2 \oplus x_1 \pmod{2^w}, \quad (5)$$

где операция $\oplus$ означает побитовое сложение векторов, задающих¹ двоичное представление вычетов x_1 и k_2 .

Рассмотрим также биективное отображение

$$f(z_1, z_2) : \mathbb{Z}_{2^w} \times \mathbb{Z}_{2^w} \rightarrow \mathbb{Z}_{2^{2w}}, \quad z_1, z_2 \in \mathbb{Z}_{2^w}. \quad (6)$$

Пусть $x = (x_1, \dots, x_{2l}) \in \mathcal{S}$, $k = (k_1, \dots, k_{2sl}) \in \mathcal{K}$. Определим функцию

$$h(k, x) \equiv \sum_{n=1}^l f(z_{2n-1}, z_{2n}) \pmod{2^{2w}}, \quad (7)$$

где $z_n = g(k_{s(n-1)+1}, \dots, k_{sn}, x_n)$, $n = 1, \dots, 2l$. Далее исследуем свойства функции (7) и покажем, что она также является ε -ASU функцией хеширования. Верна следующая теорема.

¹Напомним, что вектор $(x_0, \dots, x_{w-1})$ задает двоичное представление вычета $x \in \mathbb{Z}_{2^w}$, если $x \equiv \sum_{n=0}^{w-1} x_n 2^n \pmod{2^w}$.

Теорема 1. Пусть функция $h(k, x)$ определена сравнением (7). Тогда для любого $a \in \mathbb{Z}_{2^{2w}}$ и любого $x = (x_1, \dots, k_{2l}) \in \mathcal{S}$ найдется ровно $2^{2w(s-1)}$ таких элементов $k = (k_1, \dots, k_{2sl}) \in \mathcal{K}$, что $h(k, x) = a$.

Доказательство. Проведем доказательство индукцией по значению величины l . Пусть $l = 1$, тогда уравнению $f(z_1, z_2) = a$ в силу биективности отображения f удовлетворяет ровно одна пара значений $z_1, z_2 \in \mathbb{Z}_{2^w}$. Из свойств отображения g следует, что уравнение $g(k_1, \dots, k_s, x_1) = z_1$ имеет ровно $2^{w(s-1)}$ решений относительно неизвестных $k_1, \dots, k_s$. Аналогично, уравнение $g(k_{s+1}, \dots, k_{2s}, x_2) = z_2$ имеет ровно $2^{w(s-1)}$ решений относительно неизвестных $k_{s+1}, \dots, k_{2s}$. Таким образом, общее число векторов $k = (k_1, \dots, k_{2s})$, для которых выполнено равенство $h(k, x) = a$, есть $2^{2w(s-1)}$. Для $l = 1$ утверждение теоремы выполнено.

Теперь предположим, что утверждение теоремы выполнено для всех значений $1, 2, \dots, l-1$. Покажем, что оно выполнено и для величины l . Запишем уравнение $h(k, x) = a$ в виде

$$f(z_{2l-1}, z_{2l}) \equiv a - \sum_{n=1}^{l-1} f(z_{2n-1}, z_{2n}) \pmod{2^{2w}},$$

где $z_n = g(k_{s(n-1)+1}, \dots, k_{sn}, x_n)$, $n = 1, \dots, 2l$.

Согласно индуктивному предположению для любого $a_1 \in \mathbb{Z}_{2^{2w}}$ найдется $2^{2w((l-1)s-1)}$ такой вектор $k = (k_1, \dots, k_{2s(l-1)})$, что

$$\sum_{n=1}^{l-1} f(z_{2n-1}, z_{2n}) \equiv a_1 \pmod{2^{2w}}.$$

Пусть вычет a_1 пробегает все возможные значения $0, 1, \dots, 2^{2w} - 1$, тогда вычет $b \equiv a - a_1 \pmod{2^{2w}}$ также пробегает то же множество значений. Для каждого такого значения b найдется ровно одна пара вычетов $z_{2l-1}, z_{2l} \in \mathbb{Z}_{2^w}$, удовлетворяющая равенству $f(z_{2l-1}, z_{2l}) = b$. Аналогично сказанному выше, каждой такой паре вычетов соответствует $2^{2w(s-1)}$ наборов $k_{2s(l-1)+1}, \dots, k_{2ls}$, для которых выполнено равенство $f(z_{2l-1}, z_{2l}) = b$.

Таким образом, общее число векторов $k = (k_1, \dots, k_{2ls})$, для которых выполнено равенство $h(k, x) \equiv a \pmod{2^{2w}}$, есть

$$2^{2w(s-1)} \cdot 2^{2w} \cdot 2^{2w((l-1)s-1)} = 2^{2w(sl-1)}.$$

Теорема доказана. □

Из доказанной теоремы следует, что число таких ключей $k \in \mathcal{K}$, что выполнено равенство $h(k, x) = a$, равно $2^{2w(ls-1)} = \frac{2^{2wls}}{2^{2w}} = \frac{|\mathcal{K}|}{|\mathcal{A}|}$. Таким образом, функция $h(k, x)$ удовлетворяет первому свойству определения 2.

Докажем утверждение о числе сообщений, удовлетворяющих уравнению $h(k, x) = a$.

Теорема 2. Пусть функция $h(k, x)$ определена сравнением (7). Тогда для любого $a \in \mathbb{Z}_{2^{2w}}$ и любого $k = (k_1, \dots, k_{2sl}) \in \mathcal{K}$ найдется ровно $2^{2w(l-1)}$ таких элементов $x = (x_1, \dots, x_{2l}) \in \mathcal{K}$, что $h(k, x) = a$.

Доказательство. Зафиксируем какое-либо значение $k = (k_1, \dots, k_{2s})$ и рассмотрим уравнение $f(z_1, z_2) = a$ относительно неизвестных x_1, x_2 . В силу биективности отображения f найдется ровно одна пара вычетов $z_1, z_2 \in \mathbb{Z}_{2^w}$, удовлетворяющих данному уравнению.

Далее, в силу биективности отображения g при фиксированных значениях $k_1, \dots, k_{2s}$ и z_1, z_2 найдется ровно одна пара таких значений $x_1, x_2 \in \mathbb{Z}_{2^w}$, что

$$\begin{cases} g(k_1, \dots, k_s, x_1) \equiv z_1 \pmod{2^w}, \\ g(k_{s+1}, \dots, k_{2s}, x_2) \equiv z_1 \pmod{2^w}, \end{cases}$$

т.е. уравнение $f(z_1, z_2) = a$ имеет единственное решение относительно неизвестных x_1, x_2 .

Теперь рассмотрим сравнение $\sum_{n=1}^l a_n \equiv a \pmod{2^{2w}}$. Этому сравнению удовлетворяют $2^{2w(l-1)}$ векторов вида

$$\left(a_1, \dots, a_{l-1}, a - \sum_{n=1}^{l-1} a_n \pmod{2^{2w}} \right), \quad (8)$$

где $a_1, \dots, a_{l-1}$ принимают произвольные значения из $\mathbb{Z}_{2^{2w}}$. Как отмечено выше, каждому вектору (8) соответствует единственное решение $x_1, \dots, x_{2l}$, где

$$f(z_{2n-1}, z_{2n}) = a_n, \quad n = 1, \dots, l,$$

и $z_i = g(k_{s(i-1)+1}, \dots, k_{si}, x_i)$, $i = 1, \dots, 2l$. Теорема доказана. $\square$

Пусть $\mathcal{S}' \subset \mathcal{S}$ — множество сообщений x , удовлетворяющих при фиксированном $k \in \mathcal{K}$ уравнению $h(k, x) = a$. Тогда из доказанной теоремы следует, что

$$\frac{|\mathcal{S}'|}{|\mathcal{S}|} = \frac{2^{2w(l-1)}}{2^{2wl}} = \frac{1}{2^{2w}} = |\mathcal{A}|^{-1}.$$

Таким образом, вероятность выбора случайного сообщения $x \in \mathcal{S}$ с заданным значением функции $h(k, x) = a$ не зависит от выбора ключа $k \in \mathcal{K}$, от значения кода аутентичности $a \in \mathcal{A}$ и равна 2^{-2w} .

Теперь исследуем вопрос о том, является ли введенная нами функция (7) строго универсальной функцией хеширования. Верна следующая лемма.

Лемма 1. Пусть $f : \mathbb{Z}_{2^w} \times \mathbb{Z}_{2^w} \rightarrow \mathbb{Z}_{2^{2w}}$ — введенное ранее равенством (6) биективное отображение. Пусть $x = (x_1, x_2)$ и $y = (y_1, y_2)$ — два произвольных сообщения из множества $\mathcal{S}$. Пусть $a, b \in \mathbb{Z}_{2^{2w}}$ — два произвольных вычета. Тогда система уравнений

$$\begin{cases} f(z_1, z_2) = a, \\ f(u_1, u_2) = b, \end{cases} \quad (9)$$

где $z_n = g(k_{s(n-1)+1}, \dots, k_{sn}, x_n)$, $u_n = g(k_{s(n-1)+1}, \dots, k_{sn}, y_n)$, $n = 1, 2$, имеет не более $2^{w(s-1)}$ решений относительно неизвестных $k_1, \dots, k_{2s}$.

Доказательство. Уравнению $f(z_1, z_2) = a$ в силу биективности отображения f удовлетворяет ровно одна пара значений $z_1, z_2 \in \mathbb{Z}_{2^w}$. Из свойств отображения g следует, что уравнение $g(k_1, \dots, k_s, x_1) = z_1$ имеет ровно $2^{w(s-1)}$ решений относительно неизвестных $k_1, \dots, k_s$. Каждое из этих решений может являться решением уравнения $g(k_1, \dots, k_s, y_1) = u_1$. Следовательно, общее число решений системы (4)

$$\begin{cases} g(k_1, \dots, k_s, x_1) \equiv z_1 \pmod{2^w}, \\ g(k_1, \dots, k_s, y_1) \equiv u_2 \pmod{2^w} \end{cases}$$

не превосходит $2^{w(s-1)}$. Аналогичным образом выводится оценка числа неизвестных $k_{s+1}, \dots, k_{2s}$. Так как решения $k_1, \dots, k_s$ и $k_{s+1}, \dots, k_{2s}$ независимы, то число решений системы (4) относительно неизвестных $k_1, \dots, k_{2s}$ не превосходит $2^{2w(s-1)}$. $\square$

Отметим, что полученная в лемме 1 оценка числа решений существенно зависит от конкретных значений x, y, a, b и может достигать указанной величины $2^{2w(s-1)}$. Так, для $w = 3$, $s = 2$ и функции g , определенной равенством (5), система сравнений

$$\begin{cases} k_1 + k_2 \oplus 3 \equiv 2 \pmod{8}, \\ k_1 + k_2 \oplus 7 \equiv 6 \pmod{8} \end{cases}$$

имеет $8 = 2^{3(2-1)}$ пар решений

$$(0, 1), (1, 2), (2, 3), (3, 4), (4, 5), (5, 6), (6, 7), (7, 0).$$

Аналогично, система

$$\begin{cases} k_1 + k_2 \oplus 4 \equiv 1 \pmod{8}, \\ k_1 + k_2 \oplus 0 \equiv 5 \pmod{8} \end{cases}$$

также имеет 8 решений $(0, 5), (1, 4), (2, 3), (3, 2), (4, 1), (5, 0), (6, 7), (7, 6)$. Тогда для таких $x = (3, 4), y = (7, 0)$ и $a, b \in \mathbb{Z}_{64}$, что

$$\begin{cases} f(2, 1) = a, \\ f(6, 5) = b, \end{cases}$$

найдется ровно 64 набора решений $k_1, \dots, k_4$.

Лемма 2. Пусть в условиях леммы 1 вычеты $a, b \in \mathbb{Z}_{2^{2w}}$ пробегают множество всех возможных значений, тогда суммарное число решений системы (9) относительно неизвестных $k_1, \dots, k_{2s}$ равно 2^{2ws} .

Доказательство. Пусть вычеты $z_1, z_2, u_1, u_2 \in \mathbb{Z}_{2^w}$ удовлетворяют системе уравнений (9)

$$\begin{cases} f(z_1, z_2) = a, \\ f(u_1, u_2) = b. \end{cases}$$

Пусть a, b пробегают все возможные значения, тогда в силу биективности отображения f пары (z_1, z_2) и (u_1, u_2) пробегают множество всех возможных значений, как и пары (z_1, u_1) и (z_2, u_2) .

В этом случае, согласно третьему свойству отображения g , система уравнений (4) имеет ровно 2^{ws} решений относительно неизвестных $k_1, \dots, k_s$. Аналогично, система

$$\begin{cases} g(k_{s+1}, \dots, k_{2s}, x_2) \equiv z_2 \pmod{2^w}, \\ g(k_{s+1}, \dots, k_{2s}, y_2) \equiv u_2 \pmod{2^w} \end{cases}$$

суммарно имеет 2^{ws} решений относительно неизвестных $k_{s+1}, \dots, k_{2s}$. Так как решения $k_1, \dots, k_s$ и $k_{s+1}, \dots, k_{2s}$ независимы, то лемма доказана. $\square$

Заметим, что, в отличие от леммы 1, полученная в лемме 2 оценка числа решений не зависит от выбора значений $x, y \in \mathcal{S}$. Теперь мы можем доказать следующее утверждение.

Теорема 3. Пусть функция $h(k, x)$ определена сравнением (7). Тогда для любых вычетов $a, b \in \mathbb{Z}_{2^{2w}}$ и любых $x = (x_1, \dots, x_{2l}), y = (y_1, \dots, y_{2l})$ из множества $\mathcal{S}$ найдется не более $2^{2w(sl-1)}$ таких элементов $k = (k_1, \dots, k_{2sl})$, что

$$\begin{cases} h(k, x) = a, \\ h(k, y) = b. \end{cases} \quad (10)$$

Доказательство. Рассмотрим систему сравнений

$$\begin{cases} \sum_{n=1}^l a_n \equiv a \pmod{2^{2w}}, \\ \sum_{n=1}^l b_n \equiv b \pmod{2^{2w}} \end{cases}$$

относительно неизвестных $a_1, \dots, a_l, b_1, \dots, b_l \in \mathbb{Z}_{2^{2w}}$. Хорошо известно, см. например [4], что решениями данной системы являются векторы, в которых $a_l \equiv a - \sum_{n=1}^{l-1} a_n \pmod{2^{2w}}$, $b_l \equiv b - \sum_{n=1}^{l-1} b_n \pmod{2^{2w}}$, а вычеты $a_1, \dots, a_{l-1}, b_1, \dots, b_{l-1}$ принимают все возможные значения.

Для произвольного индекса $n = 1, \dots, l-1$ рассмотрим систему

$$\begin{cases} f(z_{2n-1}, z_{2n}) = a_n, \\ f(u_{2n-1}, u_{2n}) = b_n, \end{cases}$$

где $z_i = g(k_{s(i-1)+1}, \dots, k_{si}, x_i)$, $u_i = g(k_{s(i-1)+1}, \dots, k_{si}, y_i)$, $i = 2n-1, 2n$, и будем считать, что вычеты $a_n, b_n \in \mathbb{Z}_{2^{2w}}$ пробегают множество всех возможных значений.

В силу леммы 2 для каждого индекса $n = 1, \dots, l-1$ найдется 2^{2ws} наборов $k_{s(2n-1)+1}, \dots, k_{2sn}$, являющихся решениями указанной системы. В силу леммы 1 найдется не более $2^{2w(s-1)}$ наборов $k_{s(2l-1)+1}, \dots, k_{2sl}$, удовлетворяющих системе

$$\begin{cases} f(z_{2l-1}, z_{2l}) \equiv a - \sum_{n=1}^{l-1} a_n \pmod{2^{2w}}, \\ f(u_{2l-1}, u_{2l}) \equiv b - \sum_{n=1}^{l-1} b_n \pmod{2^{2w}}. \end{cases}$$

Таким образом, число решений, удовлетворяющих системе (10), не превосходит

$$2^{2ws(l-1)} \cdot 2^{2w(s-1)} = 2^{2w(ls-1)}.$$

Теорема доказана. $\square$

Заметим, что из замечания к лемме 1 следует, что найдутся такие значения $x, y \in \mathcal{S}$, что полученная нами оценка будет достижима.

Пусть $\mathcal{K}' \subset \mathcal{K}$ — множество ключей, для которых разрешима система (10). Тогда из утверждения теоремы следует, что

$$|\mathcal{K}'| \leq 2^{2w(ls-1)} = \frac{2^{2wls}}{2^{2w}} = \frac{|\mathcal{K}|}{|\mathcal{A}|},$$

и определенная нами функция $h(k, x)$ удовлетворяет определению 2 со значением $\varepsilon = 1$.

3. Построение биективного отображения

Предъявим в явном виде биективное отображение (6)

$$f(z_1, z_2) : \mathbb{Z}_{2^w} \times \mathbb{Z}_{2^w} \rightarrow \mathbb{Z}_{2^{2w}},$$

использованное в предыдущих рассуждениях. Легко видеть, что отображения

$$f(z_1, z_2) = z_1 + z_2 2^w, \quad \bar{f}(z_1, z_2) = z_2 + z_1 2^w \quad (11)$$

задают взаимно-однозначное соответствие между кольцом $\mathbb{Z}_{2^{2w}}$ и множеством пар $z_1, z_2 \in \mathbb{Z}_{2^w}$, т. е. могут быть использованы в наших рассуждениях.

С другой стороны, в [18] было показано, что для сообщения $x = (x_1, \dots, x_{2l}) \in \mathcal{S}$ универсальная функция хеширования должна содержать в себе не менее l нелинейных операций. В противном случае существует алгоритм, позволяющий строить коллизии с трудоемкостью меньшей, чем у метода, основанного на «парадоксе дней рождений». Так как каждое из отображений (11) является линейной формой своих аргументов, то функция $h(k, x)$, заданная сравнением (7), не содержит в себе достаточного количества нелинейных операций.

Построим отображение (6), являющееся нелинейной функцией своих аргументов. Сперва рассмотрим вспомогательное отображение с необходимыми для нас свойствами. Определим многочлен

$$\phi(u, v) = \frac{1}{2} \left((u + v)^2 - 3u - v + 2 \right), \quad \phi(u, v) \in \mathbb{Q}[u, v]. \quad (12)$$

Лемма 3. *Отображение $\phi(u, v) : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$ биективно.*

Доказательство. Рассмотрим на двумерной плоскости решетку, узлами которой являются точки (u, v) с натуральными координатами. Чтобы задать отображение $\phi : \mathbb{N} \times \mathbb{N} \rightarrow \mathbb{N}$, поместим в каждый узел решетки натуральное число следующим образом.

Разобьем все узлы решетки на диагонали, то есть на множества, координаты точек которых удовлетворяют равенству $u + v = m + 1$, $m = 1, 2, \dots$; такая диагональ содержит ровно m узлов решетки.

Для всех $m = 1, 2, 3, \dots$ поместим в узлы решетки, расположенные на одной диагонали, m последовательно идущих чисел натурального ряда. Первое из m чисел поместим в узел $(m, 1)$, а последнее — в узел $(1, m)$. Например,

В силу построения многочлена f_1 выполнено еще одно следствие из леммы 3.

Следствие 2. Если пары целых чисел (z_1, z_2) , $z_1 \geq 0$, $z_2 \geq 0$, пробегают все множество различных значений, удовлетворяющих неравенству $z_1 + z_2 < m$, то порядок аргументов многочлена $f(z_1, z_2)$ не влияет на множество принимаемых им значений, т. е. множества $\{f_1(z_1, z_2), z_1 + z_2 < m\}$ и $\{f_1(z_2, z_1), z_1 + z_2 < m\}$ совпадают.

Зафиксируем некоторое натуральное число m , определим многочлен $f_2(z_1, z_2) \in \mathbb{Q}[z_1, z_2]$ равенством

$$\begin{aligned} f_2(z_1, z_2) &= m^2 - 1 - f_1(m - 1 - z_1, m - 1 - z_2) = \\ &= \frac{1}{2} \left[(4m - 3)z_1 + (4m - 1)z_2 - (z_1 + z_2)^2 \right] - (m - 1)^2 \end{aligned} \quad (14)$$

и отображение $f : \mathbb{Z}_m \times \mathbb{Z}_m \rightarrow \mathbb{Z}_{m^2}$ равенством

$$f(z_1, z_2) = \begin{cases} f_1(z_1, z_2), & \text{если } z_1 + z_2 < m, \\ f_2(z_2, z_1) & \text{в противном случае.} \end{cases} \quad (15)$$

Теорема 4. Отображение $f : \mathbb{Z}_m \times \mathbb{Z}_m \rightarrow \mathbb{Z}_{m^2}$ биективно.

Доказательство. Согласно первому следствию леммы 3 при выполнении неравенств $0 \leq z_1 + z_2 < m$ значения многочлена $f_1(z_1, z_2)$ пробегают все целые значения от 0 до $\frac{m(m+1)}{2} - 1$.

Аналогично, при выполнении неравенств $m \leq z_1 + z_2 \leq 2(m - 1)$ значения многочлена $f_1(m - 1 - z_1, m - 1 - z_2)$ пробегают все целые числа от 0 до $\frac{m(m-1)}{2} - 1$. Следовательно, значения многочлена $f_2(z_1, z_2)$ пробегают все целые числа от $m^2 - 1$ до $m^2 - 1 - \left(\frac{m(m-1)}{2} - 1\right) = \frac{m(m+1)}{2}$.

Так как при заданных ограничениях на переменные z_1, z_2 множества значений $\{f_1(z_1, z_2)\}$ и $\{f_1(z_2, z_1)\}$ совпадают, то $f_2(z_2, z_1)$ пробегает те же значения, что $f_2(z_1, z_2)$. Теорема доказана. $\square$

Если $m = 2^w$, то отображение, заданное равенством (15), удовлетворяет равенству (6). Так как отображение (15) нелинейно, то оно может быть использовано при реализации универсальной функции хеширования $h(k, x)$, заданной сравнением (7).

Поскольку отображение f биективно, при заданном $a \in \mathbb{Z}_{2^{2w}}$ решение уравнения $f(z_1, z_2) = a$ относительно неизвестных $z_1, z_2 \in \mathbb{Z}_{2^w}$ единственно и может быть найдено следующим образом.

1. Если $a = 0$, то $z_1 = z_2 = 0$, если $a = 2^{2\omega} - 1$, то $z_1 = z_2 = 2^\omega - 1$.
2. Если $0 < a \leq \frac{2^\omega(2^\omega+1)}{2} - 1$, то определим наименьшее натуральное число n , для которого $\frac{(n-1)n}{2} - 1 < a \leq \frac{n(n+1)}{2} - 1$, и положим

$$z_1 = \frac{n(n+1)}{2} - 1 - a, \quad z_2 = a - \frac{n(n-1)}{2}. \quad (16)$$

Легко проверить, что данные значения удовлетворяют равенству $f(z_1, z_2) = a$. Действительно, при заданных ограничениях на величину a выполнено равенство $f(z_1, z_2) = f_1(u, v)$. Поскольку из (16) следует $z_1 + z_2 = n - 1$, то получаем

$$\begin{aligned} f_1(z_1, z_2) &= \\ &= \frac{1}{2} \left((n-1)^2 + \frac{n(n+1)}{2} - 1 - a + 3 \left(a - \frac{n(n-1)}{2} \right) \right) = a. \end{aligned}$$

3. Если $\frac{2^\omega(2^\omega+1)}{2} \leq a < 2^{2\omega} - 1$, то воспользуемся равенством (14) и определим наименьшее натуральное число n , для которого $\frac{(n-1)n}{2} - 1 < 2^{2\omega} - 1 - a \leq \frac{n(n+1)}{2} - 1$. Тогда

$$z_1 = a + \frac{n(n-1)}{2} - 2^\omega(2^\omega - 1), \quad z_2 = 2^\omega(2^\omega + 1) - \frac{n(n+1)}{2} - 1 - a.$$

Подставив эти значения в (14), получим $f(z_1, z_2) = f_2(z_2, z_1) = a$.

Список литературы

- [1] Алферов А. П., Зубов А. Ю., Кузьмин А. С., Черемушкин А. В., *Основы криптографии: учебное пособие*, М.: Гелиос АРВ, 2001, 480 с.
- [2] Ильясов И. И., “К распределению простых чисел в многочленах второй степени с целыми коэффициентами”, *Чебышевский сборник*, **14**:1 (2013), 56–60.
- [3] Лебедев П. А., Нестеренко А. Ю., “Режим шифрования с возможностью аутентификации”, *Системы высокой доступности*, **9**:3 (2013), 6–13.
- [4] Нестеренко Ю. В., *Теория чисел*, М.: Академия, 2008, 272 с.
- [5] Серпинский В., *О решении уравнений в целых числах*, М.: Изд-во физ.-мат. лит., 1961, 88 с.
- [6] *Словарь криптографических терминов / под ред. Б. А. Погорелова и В. Н. Сачкова*, М.: МЦМНО, 2006, 94 с.
- [7] Bellare M., Canetti R., Krawczyk H., “HMAC: Keyed-hashing for message authentication”, *Internet Engineering Task Force, Request for Comments (RFC) 2104* — February 1997.

- [8] Black J., Halevi S., Krawczyk H., Krovetz T., Rogaway P., “UMAC: Fast and secure message authentication”, CRYPTO’99, *Lect. Notes Comput. Sci.*, **1666** (1999), 216–233.
- [9] Black J., Rogaway P., “CBC MACs for arbitrary-length messages: The three-key constructions”, CRYPTO 2000, *Lect. Notes Comput. Sci.*, **1880** (2000), 197–215.
- [10] Boesgaard M., Scavenius O., Pedersen T., Christensen T., Zenner E., “Badger — a fast and provably secure MAC”, Appl. cryptogr. network secur. 3rd Int. Conf., ACNS 2005, *Lect. Notes Comput. Sci.*, **3531** (2005), 176–191.
- [11] Carter J. L., Wegman M. N., “Universal classes of hash functions”, *J. Comput. Syst. Sci.*, **18** (1979), 143–154.
- [12] Etzel M., Patel S., Ramzan Z. Square Hash: Fast message authentication via optimized universal hash functions, CRYPTO’99, *Lect. Notes Comput. Sci.*, **1666** (1999), 234–251.
- [13] *FIPS PUB 198-1. Computer Security. Cryptography. The Keyed-Hash Message Authentication Code (HMAC)*, 2008, 13 pp.
- [14] Halevi S., Krawczyk H., “MMH: software message authentication in the Gbit/second rates”, FSE’97, *Lect. Notes Comput. Sci.*, **1267** (1997), 172–189.
- [15] Handschuh H., Preneel B., “Key-recovery attacks on universal hash function based MAC algorithms”, CRYPTO 2008, *Lect. Notes Comput. Sci.*, **5157** (2008), 144–161.
- [16] Iwata T., Kurosawa K., “OMAC: One-key CBC MAC”, FSE 2003, *Lect. Notes Comput. Sci.*, **2887** (2003), 129–153.
- [17] Krovetz T., “Message authentication on 64-bit architectures”, SAC 2006, *Lect. Notes Comput. Sci.*, **4356** (2007), 327–341.
- [18] Nandi M., “On the minimum number of multiplications necessary for universal hash constructions”, *IACR Cryptology ePrint Archive*, 2013, № 574.
- [19] Preneel B., *Analysis and design of cryptographic hash functions*, Doct. diss., Katholieke Univ. Leuven, 1993.
- [20] Stinson D. R., “Universal hashing and authentication codes”, CRYPTO 91, 1991, 74–85.
- [21] Stinson D. R., “Universal hashing and message authentication codes”, *Des., Codes, and Cryptogr.*, **4:4** (1994), 369–380.
- [22] Wegman M. N., Carter J. L., “New hash functions and their use in authentication and set equality”, *J. Comput. Syst. Sci.*, **22:3** (1981), 265–279.