УДК 681.3.06

Г.З. Халимов

Харьковский национальный университет радиоэлектроники, Харьков

УНИВЕРСАЛЬНОЕ ХЕШИРОВАНИЕ СО СЛАБО СМЕЩЕННЫМИ МАССИВАМИ

Представлено решение задачи построения строго универсального хеширования по алгебраическим кривым и слабо смещенным массивам.

Ключевые слова: *кривые Ферма, универсальное хеширование, слабо смещенные массивы.*

Введение

Криптографические применения слабо смещенных массивов предложены в [1]. Распределения значений элементов массива со слабым отклонением от равновероятного определяют конструкции МАС кодов с коллизионными оценками близкими к теоретическим [2]. Универсальное хеширование со слабо смещенными массивами впервые представлено в [2,3]. Безусловная аутентификация в представлении Стинсона определяется каскадной конструкцией универсального и строго универсального хеширования [4]. Наилучший результат по скорости универсального хеширования достигается на алгебраических кривых с большим числом точек в простом поле. Исследования по кривым Ферма широко представлены в работах [5,6]. В работе [5] рассмотрены условия максимальности кривых Ферма, в [6] – оценки параметров кривых Ферма для универсального хеширования в простом и в квадратичном поле. Актуальным является построение безусловной аутентификации с использованием строго универсального хеширования со слабо смещенными массивами над простым полем вычислений.

Целью статьи является оценка параметров и построение строго универсального хеширования со слабо смещенными массивами по кривым Ферма в простом поле вычислений. В разделе 1 рассмотрены основные свойства кривых Ферма и универсальное хеширование по кривым Ферма в простом поле. В разделе 2 представлены определения строго универсального хеширования со слабо смещенными массивами, в разделе 3 - строго универсальное хеширование по кривым Ферма.

1. Универсальное хеширование по кривым Ферма в простом поле

Кривые Ферма определяются выражением

$$X^m + Y^m + Z^m = 0, \qquad (1)$$

имеют частные производные вида $F_X = mX^{m-1}$, $F_Y = mY^{m-1}$, $F_Z = mZ^{m-1}$. Основные свойства кривых Ферма для случая простого поля F_q представлены утверждением 1.

Утверждение 1 [7]**.** Пусть кривая Ферма определена над простым полем F_q. Справедливо следующее:

1) является неприводимой, несингулярной кривой степени m без особенностей, рода $g = (m-1)(m-2)/2$;

2) если m взаимно просто с $q-1$, тогда $X^m + Y^m + Z^m = 0$ изоморфна $X + Y + Z = 0$ и имеет число точек $N = q+1$;

3) кривая $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ имеет $N = 2(q-1)^2/9$ и $g = (q-4)(q-7)/18$;

4) кривая $X^{(q-1)/2} + Y^{(q-1)/2} + Z^{(q-1)/2} = 0$ имеет $N = 3(q-1)/2$ и $g = (q-3)(q-5)/8$;

5) кривая $X^{(q-1)/6} + Y^{(q-1)/6} + Z^{(q-1)/6} = 0$ имеет число точек $N = (q-1)/2 + (q-1)^2/18$ и род $g = (q-7)(q-13)/72$;

6) кривая $X^{(q-1)/2^m} + Y^{(q-1)/2^m} + Z^{(q-1)/2^m} = 0$ имеет род $g = \left(q - 2^m - 1\right)\left(q - 2^m - 2\right)/2^{2m+1}$ и число точек $N = 3(q-1)/2^m$, если $\left|2^{2^{m-1}}\right| \neq 1 (\mathrm{mod}\, q)$ и $N = 3(q-1)/2^m + 3(q-1)^2/2^{2m}$, если

$$2^{2^{m-1}} = -1 (\mathrm{mod}\, q).$$

Замечание 1.

1. Результаты 1 и 2 являются очевидными и известными. Результаты 3÷6 получаются методом подсчета числа решений для уравнений Ферма на основе свойства суммы элементов мультипликативной подгруппы второго, третьего и шестого порядков. Значение рода кривой определяется формулой Римана-Роха.

2. Точных решений для вычисления числа точек кривой Ферма, когда степень уравнения являет-

ся произвольным делителем порядка конечного поля F_q, в настоящее время нет. Теорема 1 определяет оценки для числа точек кривой Ферма общего вида.

Теорема 1 [7]**.** Пусть кривая $X^m + Y^m + Z^m = 0$ определена над простым полем F_q, где *m* есть делитель $q-1$. Оценка для числа точек кривой Ферма при $m > 2$ равна

$$N \approx 2\left\lceil \frac{(q-1)}{2m^2} \right\rceil m^2 \qquad (2)$$

где $\lceil x \rceil$ округление числа до большего целого.

Замечание 2.

1. Вычисления оценочных значений числа точек и точные вычисления дают хорошее совпадение [7]. Расхождения проявляются и могут быть существенными, когда степень уравнения становится меньше $\sqrt{q}$. Относительная погрешность вероятностной оценки уменьшается с ростом q.

2. Асимптотические результаты по кривым Ферма над простым полем определяются теоремой 2.

Теорема 2 [7]**.** Асимптотическая граница для отношения максимального числа точек $N_g(q)$ к её роду g для кривой Ферма в простом поле определяется выражением

$$\limsup_{g \to \infty} \frac{N_g(q)}{g} = 10. \qquad (3)$$

Замечание 3.

1. Точные вычисления $N_g(q)/g$ согласуются с оценкой (3). Так для $q = 257$ и $m = 16$ имеем $N = 816$ и $N_g(q)/g = 7.76$, а для $q = 2^{16} + 1 = 65537$ и $m = 2048$ имеем $N = 12589056$ и $N_g(q)/g = 6.01$.

2. В простом поле не существует максимальных кривых Ферма. При большом роде проигрыш границе Хассе-Вейля пропорционален $1/\sqrt{q}$. С уменьшением рода кривой значение числа точек приближается к границей Хассе-Вейля и при $g = 0,1$ имеем тривиальный случай $N = q + 1$.

3. Результаты для кривых Ферма представлены для простого поля и, частично, могут быть отнесены к свойствам кривых в расширениях конечного поля. Расширенные конечные поля имеют большее многообразие по комбинаторным свойствам, что влияет на оценки параметров кривых.

4. Наилучший результат по числу точек в простом поле достигается на кривой Ферма $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ [7]. Кривые имеют наилучшее отношение числа точек кривой к роду $N/g \approx 4$.

Утверждение 2 [8]. Хеширование по рациональным функциям кривой

$$X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$$

над полем F_q определяет универсальный хеш класс $\varepsilon - U(2(q-1)^2/9, q^k, q)$, где $2(q-1)^2/9$ - число хеш функций (объём ключевого пространства), q^k - объём пространства сообщений, q - объём пространства хеш кодов. Вероятность коллизии ε, если $k < g$ определяется соотношением

$$\varepsilon = 3\left\lceil (2k + 1/4)^{1/2} - 1/2 \right\rceil / (2(q-1)),$$

где g - род кривой, $\lceil \cdot \rceil$ есть округление значения до наибольшего целого.

2. Строго универсальное хеширование со слабо смещенными массивами

Безусловная аутентификация определяется строго универсальным хешированием $\varepsilon - SU(N; n, m)$ и почти строго универсальным хешированием $\varepsilon - ASU(N; n, m)$.

Определение 1 [9]. $(N; n, m)$ хеш семейство является $\varepsilon-$ строго универсальным $\varepsilon - SU(N; n, m)$, если для каждого $x \in A$ и $y \in B$ число функций $h \in H$, таких, что $h(x) = y$ равно N/m, а для любых двух различных элементов $x_1, x_2 \in A$, и не обязательно различных $y_1, y_2 \in B$ число функций $h \in H$ таких, что $h(x_1) = y_1$, $h(x_2) = y_2$ не превышает $\nu \le \varepsilon \cdot N/m$. Аббревиатура $\varepsilon - SU$ используется для обозначения $\varepsilon-$ строго универсальных хеш-функций.

Замечание 4.

1. Массив значений MAC кодов состоит из N строк, n столбцов, элементы принимают одно из m значений. Каждая функция $h \in H$ определяется значением используемого ключа, связывается со строкой и определяет правило отображения элементов множества A (номеров столбцов массива) в элементы B (собственные значения элементов массива).

2. Строгая универсальность определена для $\varepsilon = 1/m$. При смягчении требования $\varepsilon > 1/m$ класс функций определяется как почти строго универсальный $\varepsilon - ASU$.

3. Строго (почти строго) универсальное хеширование определяет безусловную аутентификацию и было представлено Стинсоном [4,9].

Коллизионные свойства почти строго универсальных MAC кодов представлены следующими утверждениями.

Утверждение 3. Пусть $(N;n,m)$ семейство хеш функций является $\varepsilon-$ строго универсальным ($\varepsilon-SU(N;n,m)$). Тогда $N \ge m^2$, вероятность имитации по МАС коду $P_{èì} = 1/m$ и вероятность подмены $P_{ïîä} = 1/m$.

Утверждение 4. Пусть $\varepsilon-ASU(N;n,m)$ семейство почти строго универсальных хеш функций. При равновероятном выборе хеш функции вероятность успеха имитационной атаки равна $P_{èì} = 1/m$ и вероятность подмены $P_{ïîä} \le \varepsilon$.

Вероятностные оценки в утверждениях 2,3 следуют из определений хеш классов. Основной результат конструкции Стинсона для строго универсального хеширования определяется теоремой 3.

Теорема 3 [9]. Композиция из универсального класса хеш-функций $\varepsilon_1 - U(N_1,n,u)$ и строго универсального класса хеш-функций $\varepsilon_2 - SU(N_2,n,m)$ является строго универсальным классом с параметрами $\varepsilon - SU(N_1N_2,n,m)$, где $\varepsilon = \varepsilon_1 + \varepsilon_2 - \varepsilon_1\varepsilon_2$.

Замечание 5.

1. Первый каскад определяет универсальное хеширование по алгебраическим кривым.

2. Второй каскад определяется строго универсальным или почти строго универсальным хешированием.

Для построения строго универсального хеширования применяется метод со слабо смещенными массивами [1 – 3].

Определение 2 [2,3]**.** Пусть p - простое число, $u = (u_1,u_2,...,u_n) \in F_p^n$. Для $\forall i \in F_p$, $v_i(u)$ есть частота появления элемента i в последовательности u $v_i(u) = \frac{n}{p} + \delta_i(u)$, где $\delta_i(u)$ - есть отклонение частоты $v_i(u)$ от среднего значения и $\sum_{i\in F_p} \delta_i(u) = 0$.

Пусть ξ комплексный корень p - степени из единицы, тогда смещение вектора u определяется как

$$\mathrm{bias}(u) = \frac{1}{n}\left|\sum_{i\in F_p} \delta_i(u)\xi^i\right| = \frac{1}{n}\left|\sum_{i\in F_p} v_i(u)\xi^i\right|.$$

Смещение $\mathrm{bias}(u)$ имеет следующие свойства.

Утверждение 5 [3]**.** Для произвольного вектора u $0 \le \mathrm{bias}(u) \le 1$ и $\mathrm{bias}(u) = 1$ только тогда, когда $u = \mathrm{const}$.

Определение 3 [2]**.** Пусть $(n,k)_p$-массив, содержащий n строк, k столбцов и записи из набора p элементов и $0 \le \varepsilon \le 1$. Массив $(n,k)_p$ является ε -смещённым (ε-biased), если любая нетривиальная линейная комбинация столбцов имеет смещение $\mathrm{bias} \le \varepsilon$.

Замечание 6.

1. Смещение массива является свойством F_p - линейного кода, построенного с помощью столбцов порождающей матрицы.

2. Для двоичных массивов параметр ε смещения прямо связывается с вероятностями появления 0 и 1 в столбцах массива.

3. Для строго универсального класса, массив хеш значений определяется $(n,k)_p$ массивом со смещением равным нулю [3].

Утверждение 6 [3]**.** Пусть смещение массива $(n,k)_p$ равно 0. Тогда $\varepsilon - ASU(n,p^k,p)$ определяет почти строго универсальный хеш класс аутентификаторов.

Число записей в любой нетривиальной линейной комбинации столбцов $(n,k)_p$ массива для каждого хеш значения строго равно $\frac{n}{p}$, по определению 1 для строго универсального хеш класса. Тогда $v_j(u) = \frac{n}{p}$, $j \in F_p$ и отсюда имеем

$$\mathrm{bias}(u) = \frac{1}{n}\left|\sum_{i\in F_p} v_i(u)\xi^i\right| = \frac{1}{n}\left|\sum_{i\in F_p} \frac{n}{p}\xi^i\right| = \frac{1}{p}\left|\sum_{i\in F_p} \xi^i\right| = 0. \square$$

Соотношение между смещением и зависимостью установлено в [1].

Теорема 4 [1]. Если массив является t - связным и ε – смещенным, он является также и t - связным и ε' - зависимым, причём, $\varepsilon' < \varepsilon$.

Фундаментальное значение этой теоремы заключается в том, что она определяет возможность применения слабо смещённых массивов в схемах аутентификации.

Метод сумм экспонент Вейля- Карлитца- Ушиямы (ВКУ) определяет массив

$$\left(p^f, f*(n - n/p)\right)_p$$

со смещением $\mathrm{bias} \le (n-1)p^{-f/2}$, с записями вида

$$\mathrm{Tr}\left(a_j\alpha^i\right),$$

где a_j - базис поля $F_{p^f} \mid F_p$, $i \le n$ и i не кратно p, $\mathrm{Tr}: F_{p^f} \to F_p$ - след элемента $a_j\alpha^i$ [2, 3].

Замечание 7.

1. Пусть $f = 2$, $n = 1$, тогда имеем $\left(p^2,2\right)_p$. Строки массива индексируются элементами $\alpha \in F_{p^2}$, столбцы – функциями: $X, \alpha X$, записи - $\mathrm{Tr}(\beta) = \beta + \beta^p$. Значение смещения столбца

$\text{bias} \le (n-1)p^{-f/2}$ будет равно 0. Можно показать, что если $f^*(n-n/p)$ чуть меньше 2, верхняя граница смещения массива $\left(p^2,2\right)_p$ $\text{bias} \le p^{-1}$.

2. Линейная комбинация столбцов массива $\left(p^2,2\right)_p$ $Y = \sum_{j=1}^{2} \gamma_j Y_j$, $\gamma_j \in F_p$ имеет смещение $\text{bias} = 0$ и значение $Y+\eta$ в строке индексированной α,η, $\alpha \in F_{p^2}$, $\eta \in F_p$ определяет строго универсальный класс $\frac{1}{p} - SU(p^3,p^2,p)$.

3. Пусть $f=2$, $n=2$, тогда имеем $\left(p^2,4\right)_p$. Строки массива индексируются элементами $\alpha \in F_{p^2}$, столбцы – функциями: $X, \alpha X, X^2, \alpha X^2$, записи - $\text{Tr}(\beta) = \beta + \beta^p$. Если $f^*(n-n/p)$ строго равняется 4, значение смещения будет точно равно $\text{bias} = p^{-1}$. Можно показать, что если $f^*(n-n/p)$ чуть меньше 4, верхняя граница смещения массива $\left(p^2,4\right)_p$ $\text{bias} \le 2/p$. Линейная комбинация столбцов массива $\left(p^2,4\right)_p$ $Y = \sum_{j=1}^{4} \gamma_j Y_j$, $\gamma_j \in F_p$ имеет смещение $\text{bias} \le 1/p$ и значение $Y+\eta$ в строке индексированной α,η, $\alpha \in F_{p^2}$, $\eta \in F_p$ определяет почти строго универсальный класс $\frac{1}{p} - ASU(p^3,p^4,p)$.

4. По теореме 3 применение несмещенного массива $\left(q^2,2\right)_q$ приводит к композиционной конструкции Стинсона вида

$$\varepsilon - SU(N_1 q^3, N, q), \qquad (4)$$

где $\varepsilon \le \varepsilon_1$, ε_1 - вероятность коллизии каскада с универсальным хешированием

$$\varepsilon_1 - U(N_1, N, q^2).$$

5. По теореме 3 применение слабо смещенного массива $\left(q^2,4\right)_q$ приводит к композиционной конструкции вида

$$\varepsilon - ASU(N_1 q^3, N, q), \qquad (5)$$

где $\varepsilon < \varepsilon_1 + q^{-1}$, ε_1 - вероятность коллизии каскада с универсальным хешированием

$$\varepsilon_1 - U(N_1, N, q^2).$$

3. Строго универсальное хеширование по кривым Ферма

Параметры строго универсальной композиционной конструкции для хеширования в простом поле представлены утверждениями 7,8.

Утверждение 7. Композиционное хеширование по рациональным функциям кривой Ферма $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ над полем F_q и слабо смещенным масивам $\left(q^2,2\right)_q$, с отображением $\phi: F_q \to F_q$ (во втором каскаде) определяет строго универсальный хеш класс $\varepsilon - SU(2q^3(q-1)^2/9, q^k, q)$ с вероятностью коллизии

$$\varepsilon = 3\left\lceil (k+1/4)^{1/2} - 1/2 \right\rceil / (2(q-1)) + q^{-1}, \qquad (6)$$

где q^k - объём пространства сообщений, q - объём пространства хеш кодов, $\lceil \cdot \rceil$ есть округление значения до наибольшего целого.

Доказательство. В первом каскаде хеширования по кривой Ферма получим хеш результат $h_{x,y}(m) \in F_q$, который определяется вычислением

$$h_{x,y}(m) = \sum_{i\ge 0, j\ge 0, (i+j)(q-1)/3 \le \rho_k} m_{i,j} \cdot x^i \cdot y^j$$

по базисным функциям $\{x^i \cdot y^j : (i+j)(q-1)/3 \le \rho_k\}$ векторного пространства $L(\rho_k P_\infty)$, где $m_{i,j} \in F_q$ - слова сообщения m, значение k определяет число слов данных. Хеширование по кривой Ферма является почти универсальным с вычислением над полем F_q. Второй каскад хеширования со слабо смещенным массивом $\left(q^2,2\right)_q$ предполагает вычисление $Y = \sum_{j=1}^{2} \gamma_j Y_j$, $\gamma_j \in F_q$ и индексирование результирующего хеш значения $Y+\eta$ строкой α,η, $\alpha \in F_{q^2}$, $\eta \in F_q$. Ключевое пространство увеличивается в q^3 раз по сравнению с хешированием по кривой Ферма. Хеширование по массиву $\left(q^2,2\right)_q$ требует подстановки двух значений хешей γ_1 и γ_2 от первого каскада. Универсальное хеширование по кривой Ферма выполняется над последовательными строками данных длиной $k/2$ слов, что приводит к вероятности коллизии $\varepsilon = 3\left\lceil (k+1/4)^{1/2} - 1/2 \right\rceil / (2(q-1))$ и параметрам хеширования $\varepsilon - U(2(q-1)^2/9, q^k, q^2)$. По теореме 3 получим требуемый результат (1) с верхней границей для ε. ◊

Утверждение 8. Композиционное хеширование по рациональным функциям кривой Ферма

$$X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$$

над полем F_q и слабо смещенным масивам $\left(q^2, 4\right)_q$, с отображением $\phi : F_q \to F_q$ (во втором каскаде) определяет почти строго универсальный хеш класс

$$\varepsilon - \mathrm{ASU}(2q^3(q-1)^2/9, q^k, q)$$

с вероятностью коллизии

$$\varepsilon = 3\left\lceil (k/2+1/4)^{1/2} - 1/2 \right\rceil / (2(q-1)) + q^{-1}, \quad (7)$$

где q^k - объём пространства сообщений, q - объём пространства хеш кодов, $\lceil \cdot \rceil$ есть округление значения до наибольшего целого.

Доказательство утверждения аналогично предыдущему. Второй каскад хеширования использует хеш вычисление со слабо смещенным масивом $\left(q^2, 4\right)_q$, с отображением

$$\phi : F_q \to F_q .$$

Универсальное хеширование по кривой Ферма выполняется над четырьмя последовательными строками данных длиной $k/4$ слов и подстановкой четырех хешей во второй каскад. Применение теоремы 3 приводит к требуемому результату (7).

Выводы

1. Задача построения безусловной аутентификации может быть решена в теории слабо смещённых массивов, которая учитывает неравновероятность в распределениях MAC кодов.

2. Применение слабо смещенных массивов позволяет в полтора и два раза снизить вероятность коллизии строго универсального хеширования и приводит к двух с половиной кратному увеличению размера ключа по сравнению с универсальным хешированием по кривой Ферма.

Список литературы

1. *Kurosawa K. Almost k – wise independent sample spaces and their cryptologic applications / K. Kurosawa, T. Johansson, D. Stinson // Lecture Notes in Computer Science. – 1997. – № 1233. – P. 409 – 421.*

2. *Bierbrauer J. Weakly biased arrays, almost independent arrays and error – correcting codes / J. Bierbrauer, H. Schellwat //Publication in Proceedings of AMS – DIMACS, 2000. – P. 33.*

3. *Халимов Г.З. Безусловная аутентификация с использованием слабо смещенных массивов / Г.З. Халимов // Радиотехника. Всеукр. межвед. науч. – техн. сб. Тем. выпуск «Информационная безопасность». – 2003. – № 134. – С. 165 – 171.*

4. *Stinson D. Universal hashing and authentication codes / D. Stinson // Design, Codes and Cryptography. – 1994. – V. 4. – P. 369 – 380.*

5. *Халимов Г.З. Оценка параметров кривых Ферма в расширенном поле для универсального хеширования / Г.З. Халимов, А.В. Леншин // Защита информации, Сборник научных трудов НАУ, Киев. – 2010. – Вып. 17, – С. 116 – 120*

6. *Халимов Г.З. Оценка параметров кривых Ферма для универсального хеширования в простом поле / Г.З. Халимов // Научно – техническая конференция с международным участием. Компьютерное моделирование в наукоемких технологиях (часть 2). КМНТ Харьков, 18 – 21 мая 2010. – С. 266*

7. *Халимов Г.З. Оценка параметров кривых Ферма для универсального хеширования / Г.З. Халимов // Журнал «Радіоелектроніка, інформатика, управління» Запоріжжя: ЗТТУ, 2011. –№ 1(24). – С.82 – 86.*

8. *Халимов Г.З. Универсальное хеширование по алгебраическим кривым в простом поле / Г.З. Халимов // Журнал «Системи управління, навігації та зв'язку» Міністерство промислової політики України, ДП «Центральний науково – дослідний інститут навігації і управління» Київ. – 2011. – Вип. 1(17). – С. 156 – 161.*

9. *Stinson D.R. Combinatorial techniques for universal hashing / D.R. Stinson //Journal of Computer and Systems Science. – 1994. – V. 48. – P. 337 – 346.*

Поступила в редколлегию 6.02.2013

Рецензент: д-р техн. наук проф. В.И. Долгов, Харьковский национальный университет радиоэлектроники, Харьков.

УНІВЕРСАЛЬНЕ ГЕШУВАННЯ ЗІ СЛАБО ЗМІЩЕНИМИ МАСИВАМИ

Г.З. Халімов

Представлено рішення задачі побудови строго універсального гешування за алгебричними кривими та слабо зміщеними масивами.

Ключові слова: *криві Ферма, універсальне гешування, слабо зміщені масиви.*

UNIVERSAL HASHING WITH WEAKLY BIASED ARRAYS

G.Z. Khalimov

The solution of the problem of constructing a universal hashing strictly algebraic curves and weakly biased arrays.

Keywords: *Fermat curves, universal hashing, weakly biased arrays.*