УДК 681.3.06

Г. З. ХАЛИМОВ, д-р техн. наук

КОМПОЗИЦИОННОЕ УНИВЕРСАЛЬНОЕ ХЕШИРОВАНИЕ ПО КРИВЫМ ФЕРМА В ПРОСТОМ ПОЛЕ

Безусловная аутентификация определяется строго универсальным хешированием и почти строго универсальным хешированием. Для построения строго универсального хеширования применяется метод ортогональных массивов [1, 2] и метод сумм экспонент Вейля-Карлитца-Ушиямы [3, 4]. В общем случае строго универсальные семейства хеш функций определяют t кратную аутентификацию. Практическим ограничением этих методов является большой размер ключевых данных, который в несколько раз превышает размер сообщений. Стинсон рассмотрел композиционное хеширование со снятием ограничения на размер ключевых данных для строго универсального хеширования [5]. Конструкция определяется каскадной схемой хеширования по почти универсальному семейству хеш функций и по ортогональным массивам. Скоростное универсальное хеширование определяется над простым конечным полем. Практические схемы таких вычислений использованы в UMAC алгоритме.

В работе предлагается композиционное универсальное хеширование в простом поле по наилучшим кривым Ферма. С этой целью в разд. 1 рассмотрены определение и свойства композиционного хеширования. В разд. 2 – универсальное хеширование по ортогональным массивам. В разд. 3 представлено композиционное универсальное хеширование по кривым Ферма и проективной прямой в простом поле.

1. Определение и свойства композиционного хеширования

Композиционные схемы являются эффективным механизмом построения класса хеш функций с заданными комбинаторными свойствами. Семейства универсальных хеш функций в представлении массивов аутентификаторов имеют определения 1, 2.

Определение 1 [6]. $(N;n,m)$ хеш семейство является $\varepsilon-$ универсальным, если для любых двух различных элементов $x_1, x_2 \in A$, существует самое большее εN функций $h \in H$ таких, что $h(x_1) = h(x_2)$. Аббревиатура $\varepsilon - U$ используется для обозначения ε -универсальных хеш функций.

Замечание 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).

Утверждение 1 [7]. Пусть h выбирается случайно из заданного $\varepsilon - U(N;n,m)$ хеш семейства, тогда вероятность коллизии хеш значений для двух разных входных сообщений $x_1, x_2 \in A$ не превышает ε.

Замечание 2.

1. Первоначальное определение универсальных хеш функций было предложено Картером и Вегманом для $\varepsilon = 1/m$ [6].

2. Вероятность коллизии для универсальных хеш функций Картера и Вегмана является наименьшей и определяется мощностью пространства хеш значений $P_{кол} = 1/|B|$.

Определение 2 [6]. H является $\varepsilon-$ почти универсальным семейством хеш функций ($\varepsilon - AU(N;n,m)$), если $P_{кол} = \Pr_{h \in H}\left[h(x_1) = h(x_2)\right] \le \varepsilon$ для $x_1, x_2 \in A$, $x_1 \ne x_2$, $1/m < \varepsilon \le 1$.

Замечание 3. Для почти универсальных семейств несколько ослабляются требования к вероятности коллизии.

Для уменьшения размера хеш кода используется композиционная конструкция Стинсона [5]. Композиционная конструкция Стинсона имеет определение 3 и свойства представлены утверждением 2.

Определение 3 [5]. Пусть $H_1 = \left\{h : \{0,1\}^a \to \{0,1\}^b\right\}$ и $H_2 = \left\{h : \{0,1\}^a \to \{0,1\}^c\right\}$ есть класс хеш функций. Композиционный класс хеш функций $H_1 H_2 = \left\{h : \{0,1\}^a \to \{0,1\}^c\right\}$ имеет определение

$$(h_1, h_2)(x) = h_2(h_1(x)).$$

Утверждение 2 [5]**.** Если H_1 есть $\varepsilon_1 - U$ универсальный класс и H_2 есть $\varepsilon_2 - U$, тогда $H_1 \cap H_2$ есть $(\varepsilon_1 + \varepsilon_2) - U$.

Замечание 4. Композиционное включение хеш функций с целью уменьшения размера хеш кода приводит к увеличению вероятности коллизии и увеличению сложности вычислений и возможно размера ключевых данных.

Безусловная аутентификация определяется строго универсальным хешированием $\varepsilon - SU(N; n, m)$ (определение 4, утверждение 3) и почти строго универсальным хешированием $\varepsilon - ASU(N; n, m)$ (утверждение 4).

Определение 4 [5]. $(N; n, m)$ 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 5.

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

2. Строго (почти строго) универсальное хеширование определяет безусловную аутентификацию и было представлено Стинсоном [5,8].

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 $N \ge m^2$, $P_{им} = 1/m$ и $P_{под} = 1/m$.

Доказательство. По определению строгой универсальности число функций $h \in H$ таких, что $h(x_1) = y_1$, $h(x_2) = y_2$ не превышает $\varepsilon \cdot N/m$. Возьмём нижнюю границу $\nu = 1$ и так как $\varepsilon = 1/m$ имеем $N \ge m^2$. Прямое вычисление вероятности имитационной атаки по ключу дает $N \ge m^2$ $P_{им.кл} = (N/m)/M = 1/m$, что соответствует нижней границе для вероятности имитации по MAC коду, следовательно $P_{им} = 1/m$. Вероятность подмены определяется условной вероятностью. Так как число h для которых $h(x) = y$ равно N/m, а число h для которых $h(x) = y$, $h(x') = y'$ равно $\nu \le \varepsilon \cdot N/m = N/m^2$ получим $P_{под} = 1/m$.

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

Доказательство аналогично предыдущему.

Основной результат конструкции Стинсона для строго универсального хеширования определяется теоремой 1.

Теорема 1. [6] 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 $\varepsilon - SU(N_1 N_2, n, m)$, где $\varepsilon = \varepsilon_1 + \varepsilon_2 - \varepsilon_1 \varepsilon_2$.

Замечание 6.

1. Первый каскад определяет универсальное хеширование по алгебраическим кривым.

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

3. Для построения строго универсального хеширования применяется метод ортогональных массивов [1,2].

2. Универсальное хеширование по ортогональным массивам

Определение 5. [1] Пусть X, Y являются множествами из k и υ элементов соответственно, и H есть множество функций осуществляющих отображение $f : X \to Y$. *Ортогональным массивом* $OA_\lambda(t, k, \upsilon)$ называется массив элементов $y_i \in Y$, со столбцами соответствующими элементам множества X и строками, определяемыми элементами множества m, в котором для любой выборки из t элементов $y_1, y_2, ..., y_t$ из Y существует только λ функций $f \in m$ для которых справедливо $f(x_i) = y_i$, $i = 1,2,\ldots,t$.

Основная конструкция OA массивов определена теоремой 2.

Теорема 2 [2]. Пусть q простое число, m, n, t – целые числа, $n \geq m$, $2 \leq t \leq q^n$. Зафиксируем сюрективное F_q – линейное отображение $\varphi : F_q^n \to F_q^m$. Для каждого t набора $(z, a_1, a_2, ..., a_{t-1})$, где $z \in F_{q^m}$, $a_j \in F_{q^n}$, $i = 1,2,\ldots,t-1$, определим отображение $f = f(z, a_1, a_2, ..., a_{t-1}) : F_q^n \to F_q^m$ вида

$$f(x) = \varphi\left(\sum_{j=1}^{t-1} a_j x^j\right) + z . \qquad (1)$$

Тогда массив, составленный из отображений вида (1), является ортогональным с параметрами $OA_{q^{(t-1)(n-m)}}\left(t, q^n, q^m\right)$.

Следствие 1. Пусть q – простое число, $m = n$, $t = 2$. Тогда $OA_{\lambda=1}\left(2, q^m, q^m\right)$ называется простым, каждая строка повторяется только (точно) один раз и определяется линейным отображением $\varphi : F_{q^m} \to F_{q^m}$ с функцией $f(x) = \varphi(ax) + z$, где $a, z \in F_{q^m}$.

Следствие 2. Пусть q_1 и q_2 – простые числа, $q_1 > q_2$, $m = n = 1$, $t = 2$. Тогда $OA_{\lambda=\lceil q_1/q_2 \rceil}(2, q_1, q_2)$ – массив каждая строка которого повторяется самое большее $\lambda = \lceil q_1 / q_2 \rceil$ раз и определяется линейным отображением $\varphi : F_{q_1} \to F_{q_2}$ с функцией $f(x) = \varphi(ax) + z$, где $a \in F_{q_1}$, $z \in F_{q_2}$, $\lceil \cdot \rceil$ – округление к наибольшему целому.

Замечание 7.

1. Теорема 2 определяет строго универсальное хеширование $1/q^m - SU\left(q^{n+m}, q^n, q^m\right)$ над расширенным полем (утверждение 3 [7]). В композиционной конструкции Стинсона по теореме 1 получим

$$\varepsilon - SU(N_1 q^{n+m}, N, q^m),$$

где $\varepsilon < \varepsilon_1 + 1/q^m$, ε_1 – вероятность коллизии каскада с универсальным хешированием $\varepsilon_1 - U(N_1, N, q^n)$.

По следствию 1 получим

$$\varepsilon - SU(N_1 q^{2m}, N, q^m), \tag{2}$$

где $\varepsilon < \varepsilon_1 + 1/q^m$, $m \geq 1$.

2. Линейное отображение $\varphi : F_{q_1} \to F_{q_2}$ с функцией $f(x) = \varphi(ax) + z$ приводит к почти строго универсальному хешированию $2/q_2 - ASU(q_1 q_2, q_1, q_2)$. Вычисление $f(x) = \varphi(ax) + z$ определяется модульными вычислениями в простом конечном поле F_{q_2}.

3. Для композиционной конструкции Стинсона с $2/q_2 - ASU(q_1 q_2, q_1, q_2)$ хешированием во втором каскаде получим следующие параметры хеширования

$$\varepsilon - ASU(N_1 q_1 q_2, N, q_2), \tag{3}$$

где $\varepsilon < \varepsilon_1 + 2/q_2$, ε_1 – вероятность коллизии каскада с универсальным хешированием $\varepsilon_1 - U(N_1, N, q_1)$.

4. Линейное отображение $\varphi : F_q^n \to F_q^m$ определяет умножение элементов в F_{q^n}, проектирование m координат $F_{q^n} \to F_{q^m}$ и сложение в F_{q^m}.

3. Композиционное универсальное хеширование по кривым Ферма

Хеширование по алгебраическим кривым на функциональном пространстве $L(\rho_k P_\infty)$ над простым полем F_q определяет универсальный хеш класс $\varepsilon - U(N, q^k, q)$, где N – число точек алгебраической кривой (объём ключевого пространства), q^k – объём пространства сообщений, q – объём пространства хеш кодов. Вероятность коллизии ε определяется соотношением $\varepsilon = \rho_k / N$.

Универсальное хеширование в простом поле определено по проективной прямой и кривой Ферма.

Известные результаты.

1. Наилучший результат по числу точек в простом поле достигается на кривой Ферма

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

Кривые имеют наилучшее отношение числа точек кривой к роду $N/g \approx 4$ [9].

2. При большом роде проигрыш границе Хассе – Вейля в простом поле для кривых Ферма и Гурвица пропорционален $1/\sqrt{q}$. С уменьшением рода кривой значение числа точек приближается к границе Хассе – Вейля.

3. Кривая $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ имеет $N = 2(q-1)^2/9$ F_q- рациональных точек, род $g = (q-4)(q-7)/18$. Точками кривой являются $P_{a,b} = (a : b : 1)$, где $a, b \in F_q$, $a \neq 0$, $b \neq 0$ и $a^{(q-1)/3} + b^{(q-1)/3} + 1 = 0$. Базис пространства $L(mP_\infty)$ задается функциями вида $\{x^i \cdot y^j : (i+j)(q-1)/3 \leq m\}$ и универсальное хеширование для сообщения $m = (m_0, m_1, ..., m_k)$, $m_i \in F_q$ определяется выражением

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

где $m_{i,j} \in F_q$ – слова сообщения m, параметр k определяет число слов данных.

Утверждение 5 [10]. 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 определяется соотношениями

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

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

5. Для кривых Ферма в простом поле для рода $g = 0{,}1$ имеем случай проективной прямой $X + Y + Z = 0$ с числом точек $N = q+1$. Точками прямой являются $P_{a,b} = (a : b : 1)$, где $a \in F_q$ и $a + b + 1 = 0$. Базис пространства $L(kP)$, задается функциями вида $\{x^i : i \leq k\}$ и универсальное хеширование для сообщения $m = (m_0, m_1, ..., m_k)$, $m_i \in F_q$ определяется выражением

$$h_x(m) = \sum_{i=0}^{k-1} m_i \cdot x^i ,$$

где x точки проективной прямой.

Утверждение 6 [11]. Хеширование по проективной прямой $X + Y + Z = 0$ над полем F_q определяет универсальный хеш класс $\varepsilon - U(q, q^k, q)$, где q – число хеш функций (объём ключевого пространства), q^k – объём пространства сообщений, q – объём пространства хеш кодов. Вероятность коллизии ε определяется соотношением $\varepsilon = k/q$.

Параметры строго универсальной композиционной конструкции для хеширования в простом поле представлены утверждениями 7 – 10.

Утверждение 7. Композиционное хеширование по рациональным функциям кривой Ферма $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ (в первом каскаде конструкции Стинсона) над полем F_q и по ортогональным массивам (во втором каскаде) с отображением $f(x) = \varphi(ax) + z$, $\varphi : F_q \to F_q$, $a, z \in F_q$ определяет строго универсальный хеш класс $\varepsilon - SU(2q(q-1)^2/9, q^k, q)$ с вероятностью коллизии

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

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 $h_{x,y}(m) \in F_q$. Второй каскад хеширования с отображением $f(s) = \varphi(as) + z$, $\varphi : F_q \to F_q$, $a, z \in F_q$, $s = h_{x,y}(m)$ определяет строго универсальный класс по теореме 2 с параметрами $q^{-1} - SU\left(q^2, q, q\right)$. По теореме 1 получим требуемый результат (4) с верхней границей для ε.

Утверждение 8. Композиционное хеширование по рациональным функциям кривой Ферма $X^{(q-1)/3} + Y^{(q-1)/3} + Z^{(q-1)/3} = 0$ в первом каскаде конструкции Стинсона над полем

F_q и по ортогональным массивам (во втором каскаде) с отображением $f(x)=\varphi(ax)+z$, $\varphi: F_{q_1} \to F_{q_2}$, $a \in F_{q_1}$, $z \in F_{q_2}$, определяет почти строго универсальный хеш класс $\varepsilon - ASU(2q_2(q_1-1)^3/9, {q_1}^k, q_2)$ с вероятностью коллизии:

$$\varepsilon = 3\left\lceil (2k+1/4)^{1/2} - 1/2 \right\rceil /(2(q_1-1)) + {q_2}^{-1},$$

где ${q_1}^k$ – объём пространства сообщений, q_2 – объём пространства хеш кодов, $\lceil \cdot \rceil$ есть округление значения до наибольшего целого.

Доказательство утверждения аналогично предыдущему. Второй каскад хеширования использует отображение одного простого поля на меньшее $\varphi: F_{q_1} \to F_{q_2}$. В силу линейности хеш преобразований на первом и втором каскадах результатом является почти строго универсальное хеширование.

Хеширование по проективной прямой приводит к строго универсальным классам с параметрами которые определяются утверждениями 9,10.

Утверждение 9. Композиционное хеширование по проективной прямой с ортогональным массивом во втором каскаде на основе отображения $f(x)=\varphi(ax)+z$, $\varphi: F_q \to F_q$, $a, z \in F_q$ определяет строго универсальный хеш класс $k/q - SU(q^3, q^k, q)$, где q^k – объём пространства сообщений, q – объём пространства хеш кодов.

Утверждение 10. Композиционное хеширование по проективной прямой с ортогональным массивом (во втором каскаде) на основе отображения $f(x)=\varphi(ax)+z$, $\varphi: F_{q_1} \to F_{q_2}$, $a \in F_{q_1}$, $z \in F_{q_2}$, определяет почти строго универсальный хеш класс $\varepsilon - ASU(q_1^2 q_2, q_1^k, q_2)$, $\varepsilon = k/q_1 + 2/q_2$, где q_1^k – объём пространства сообщений, q_2 – объём пространства хеш кодов.

Замечание 8.

1. Результаты утверждений 9, 10 определяются соотношениями (2), (3).

2. Отображение $f(x)=\varphi(ax)+z$, $\varphi: F_{q_1} \to F_{q_2}$, $a \in F_{q_1}$, $z \in F_{q_2}$ не приводит к класическому ортогональному массиву, так как каждая строка массива аутентификатора повторяется самое большее $\lambda = \lceil q_1/q_2 \rceil$ раз (см. следствие 2).

Выводы

1. Композиционное строгое (почти строгое) универсальное хеширование над простым полем по проективной прямой кривой Ферма определяется трёх кратным увеличением размера ключа. Это является минимальной платой за безусловную аутентификацию.

2. Несколько снижает ключевые затраты вычисления $f(x)=\varphi(ax)+z$ по разным модулям $q_2 < q_1$. Значение q_2 определяет коллизионные оценки второго каскада хеширования. Выбор q_2 можно определить приблизительным равенством $\varepsilon_1 \approx 1/q_2$. Вклад по коллизии первого и второго каскадов композиционной схемы будет приблизительно равным. Хеширование по разным модулям снимает избыточность размера хеша по коллизионной оценке. Практические оценки по вероятности коллизии определяются оценками первого каскада.

Список литературы: 1. *Mukhopadhyay A.L.* Construction of some series of ortogonal array / A.L. Mukhopadhyay // Sankya B43. – 1981. – P.81-92. 2. *Bierbrauer J.* Bounds on orthogonal arrays and resilient functions / J.Bierbrauer // Journal of Combinatorial Designs. – 1995. – N.3. – P.179-183.1. Carter J. L. Universal classes of hash functions / J. L.Carter, M.N.Wegman // Journal of Computer and Systems Science. – 1979. – V.18. – P.143-154. 3. *Carlitz L.* Bounds for exponential sums / L.Carlitz, S.Uchiyama //Duke Mathematical Journal. – 1957. – N.24. – P.37-41. 4. *Халимов Г.З.* Безусловная аутентификация с использованием слабосмещенных массивов / Г.З.Халимов // Радиотехника. – 2003. – №134. –С.165-171. 5. *Stinson D.R.* Com-

binatorial techniques for universal hashing / D.R.Stinson // Journal of Computer and Systems Science. – 1994. – V.48. – P.337-346. 6. *Carter J. L.* Universal classes of hash functions / J. L.Carter, M.N.Wegman // Journal of Computer and Systems Science. – 1979. – V.18 . -P.143-154. 7. *Халимов Г.З.* Аутентификация и универсальное хеширование / Г.З.Халимов, А.А.Кузнецов // Радиотехника. – 2001. – Вып.119. – С. 88-94. 8. *Stinson D.R.* Universal hashing and authentication codes / D.R.Stinson // Designs, Codes and Cryptography. – 1994. – N.4. – P.369–380. 9. *Халимов Г.З.* Оценка параметров кривых Ферма для универсального хеширования / Г.З.Халимов // Радіоелектроніка, інформатика, управління. – Запоріжжя : ЗТТУ, 2011. – №1(24). – С.82-86. 10. *Халимов Г.З.* Универсальное хеширование по алгебраическим кривым в простом поле / Г.З.Халимов // 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. 11. *Халимов Г.З.* Универсальное хеширование по рациональным функциям кривой Эрмита / Г.З.Халимов, А.Ю.Иохов // Междунар. науч.-практ. конф. «Застосування інформаційніх технологій у підготовці та діяльності сил охорони правопорядку» / Академія внутрішніх військ МВС України 17.03.2011. Зб тези доповідей, 2011. – С.48-51.

Харьковский национальный
университет радиоэлектроники

Поступила в редколлегию 11.11.2012