УДК 681.3.06

А.Г. КОРЧЕНКО, Е.В. КОТУХ, А. А. БОЙКО

МНОГОКАСКАДНОЕ УНИВЕРСАЛЬНОЕ ХЕШИРОВАНИЕ ПО РАЦИОНАЛЬНЫМ ФУНКЦИЯМ КРИВОЙ СУЗУКИ

Введение

Применение для аутентификации каскадного хеширования позволяет получить наименьшую вероятность коллизии, сложность вычислений хешей для фиксированного объёма данных и размерности конечного поля. Основные результаты представлены в работах [1, 2]. Наилучший результат универсального хеширования достигается на максимальных кривых, число точек которых лежит на границе Хассе–Вейля. Свойства каскадного универсального хеширования по проективной прямой и кривой Эрмита представлены в [1]. Оценки вероятности коллизии для многократного универсального хеширования получены в [2]. Максимальные кривые ассоциированные с группой Сузуки и группой Ри имеют максимально возможное значение рода и соответственно число точек [3]. Актуальным является построение многокаскадного универсального хеширования по рациональным функциям кривой Сузуки.

В данной работе предлагается метод каскадного универсального хеширования по рациональным функциям кривой Сузуки. С этой целью в разд. 1 рассмотрено каскадное универсальное хеширование по рациональным функциям алгебраических кривых. В разд. 2 представлено многокаскадное универсальное хеширование на основе произведения функциональных полей кривой Сузуки и хеширования по проективной прямой и кривой Эрмита.

1. Каскадное универсальное хеширование по алгебраическим кривым

Каскадное универсальное хеширование по алгебраическим кривым рассмотрено в [1] и представляется двумя основными схемами: каскадное хеширование со связкой хеша и текста и каскадное хеширование на основе произведения функциональных полей.

Определение каскадной схемы универсального хеширования со связкой хеша и текста имеет следующее представление.

Определение 1. [1]. Пусть F_q – конечное поле, M – сообщение и $M = M_1 \| M_2$. Каскадное универсальное хеширование по рациональным функциям алгебраических кривых определяется выражением $Ch(M) = AGh_2(AGh_1(M_1) \| M_2)$, где AGh_1, AGh_2 есть универсальные схемы хеширования по рациональным функциям алгебраических кривых, $Ch(M)$ определяет универсальное семейство хеш функций $\varepsilon - AU$, где $\varepsilon = \max(\varepsilon_1, \varepsilon_2) + 1/|H^2|$, $\varepsilon_1, \varepsilon_2$ – соответственно, вероятности коллизий для AGh_1 и AGh_2 хеширования.

Утверждение 1. [1] Пусть AGh_1, AGh_2 соответственно $\varepsilon_1 - U$ и $\varepsilon_2 - U$ универсальные классы хеш функций. Каскадная конструкция $Ch(M)$ имеет наименьшую вероятность коллизии, если $\varepsilon_1 = \varepsilon_2$.

Замечание 1.

1. Каскадное хеширование $Ch(M)$ при фиксированном поле вычислений предполагает разбиение данных на блоки приблизительно равной длины. Вероятность коллизии каскадного хеширования имеет ограничение по наименьшему полю вычисления хеша одного из каскадов.

2. Размер ключевых данных увеличивается пропорционально числу каскадов, с учетом поля вычисления и универсального хеширования каскада.

3. Каскадное хеширование позволяет эффективно увеличить общую длину хешируемых данных и зафиксировать вероятность коллизии на уровне хеша первого каскада, если на втором и последующих каскадах увеличить поле вычислений. Примером является алгоритм хе-

ширования UMAC(2000), каскадная схема применяется с подъёмом поля вычисления, сначала 32, затем 64 и 128 бит.

Свойства каскадного хеширования по алгебраическим кривым на основе произведения функциональных полей имеют следующее определение.

Определение 2. Пусть F_q – конечное поле, M – сообщение и $M = M_1 \| M_2 \| ... \| M_t$. Алгоритм вычисления хеш кода в каскадной конструкции определяется выражением

$$Ch_t(M) = AGh_2\left(AGh_1(M_1) \| AGh_1(M_2) \| ... \| AGh_1(M_t)\right), \quad (1)$$

где AGh_1, AGh_2 – универсальные схемы хеширования по алгебраическим кривым, $Ch_t(M)$ определяет универсальное семейство хеш функций $\varepsilon - AU$, где $\varepsilon = \max(\varepsilon_1, \varepsilon_2) + 1/|H^2|$, $\varepsilon_1, \varepsilon_2$ – соответственно вероятности коллизий для AGh_1 и AGh_2 хеширования.

Замечание 2.

1. Коллизионные свойства каскадного хеширования следуют из утверждения 1.

2. Каскадное хеширование $Ch_t(M)$ при фиксированном поле вычислений предполагает разбиение данных на t блоков равной длины. Для первого каскада вероятность коллизии определяется размером блока данных, для второго – значением t – числа блоков данных.

3. Вероятность коллизии каскадного хеширования имеет ограничение по наименьшему полю вычисления хеша одного из каскадов.

4. Размер ключевых данных определяется произведением пространства ключей первого и второго каскадов, с учетом поля вычисления и универсального хеширования каскада.

5. Каскадное хеширование позволяет эффективно увеличить общую длину хешируемых данных и зафиксировать вероятность коллизии на уровне хеша одного из каскадов.

6. Двух каскадная конструкция $Ch_t(M)$ легко распространяется на многокаскадную $l - Ch_t(M)$, где l – число вложенных каскадов универсальных схем хеширования по алгебраическим кривым AGh_1, AGh_2,..., AGh_l.

2. Многокаскадное универсальное хеширование по рациональным функциям кривой Сузуки

Основные результаты по кривой Сузуки. Кривые Сузуки S являются F_q изоморфными плоской кривой $y^q - y = x^{q_0}(x^q - x)$, где $q = 2q_0^2$ и $q_0 = 2^s$. Род кривой $g = q_0(q-1)$, число F_q рациональных точек равно $q^2 + 1$.

Точками кривой являются особая точка на бесконечности $P_0 = (0:1:0)$ кратности q_0 и рациональные точки $P_{a,b} = (a:b:1)$, где $a, b \in F_{q^2}$ и $b^q - b = a^{q_0}(a^q - a)$.

Подгруппа Вейерштрасса функционального поля кривой содержит подгруппу $H(P_\infty) = \{q, q+q_0, q+2q_0, q+2q_0+1\}$. Кривая Сузуки определяется полной линейной серией $D = |(q+2q_0+1)P_0|$ размерности $\dim = 4$. Базис пространства $L(\rho_t P_0)$ задается функциями вида $\{w^i \cdot v^t \cdot y^j \cdot x^r : i(q+2q_0) + j(q+2q_0+1) + t(q+q_0) + r \cdot q \le \rho_t\}$, что следует из подгруппы Вейерштрасса $H(P_0)$, представленной порядками полюсов функций $x = X/Z$, $y = Y/Z$, $v = x^{2q_0+1} + y^{2q_0}$, $w := xy^{2q_0} + x^{2q+2q_0} + y^{2q}$. Порядки полюсов $div_\infty(x) = qP_0$, $div_\infty(y) = (q+q_0)P_0$, $div_\infty(v) = (q+2q_0)P_0$, $div_\infty(w) = (q+2q_0+1)P_0$.

Универсальный класс хеш функций, построенный по рациональным функциям кривой Сузуки, представлен в [4, 5]. Основные результаты следующие.

Определение 3. [4] Хеш функция $h_{x,y}(m) \in F_q$, $q = 2q_0^2$, $q_0 = 2^s$ для сообщения m по рациональным функциям кривой $Y^q Z^{q_0} - YZ^{q+q_0-1} = X^{q+q_0} - X^{q_0+1} Z^{q-q_0-1}$ в точке x, y определяется выражением $h_{x,y}(m) = \sum m_{i,j,t,r} \cdot w^i \cdot v^j \cdot y^t \cdot x^r$, где ρ_k – полюс подгруппы Вейерштрасса $H(P_\infty)$, $m_{i,j,t,r} \in F_q$ – слова сообщения m, $i \geq 0$, $0 \leq j \leq 2q_0 - 1$, $0 \leq t \leq 1$, $0 \leq r \leq q_0$. $i(q+2q_0) + j(q+2q_0+1) + t(q+q_0) + rq \leq \rho_k$. $x = X/Z$, $y = Y/Z$, $v = x^{2q_0+1} + y^{2q_0}$). $w := xy^{2q_0} + x^{2q+2q_0} + y^{2q}$.

Утверждение 2. [4] Хеширование по рациональным функциям кривой $y^q - y = x^{q_0}(x^q - x)$, где $q = 2q_0^2$ и $q_0 = 2^s$ над полем F_q, определяет универсальный хеш класс $\varepsilon - U(q^2, q^k, q)$, где q^2 – число хеш функций (объём ключевого пространства), q^k – объём пространства сообщений, q – объём пространства хеш кодов. Вероятность коллизии ε определяется соотношениями

$$\varepsilon = (i(q+2q_0) + j(q+2q_0+1) + t(q+q_0) + rq)/q^2, \text{ если } k < q_0(q-1), \quad (2)$$

$$\varepsilon = (k + q_0(q-1))/q^2, \text{ если } k \geq q_0(q-1), \quad (3)$$

где i, j, t, r определяются леммой.

Лемма [4]. Пусть $k < q_0(q-1)$. Для кривой Сузуки имеет место $i = p - j$, $j = \Delta - t \cdot q_0 - 1$. $r = s - s_2 + d \cdot q_0$, $t = t_1 \bmod 2$, где $s' = \lfloor (3k)^{1/3} \rfloor$, $\Sigma = s'(s'+1)(2s'+1)/6$, $s = s' + \lfloor k/\Sigma \rfloor$. $s_1 = s - q_0 - 1$, $\Sigma_{s-1} = s(s-1)(2s-1)/6 - s_1(s_1-1)(2s_1-1)/3 - s_1(s_1-1)$, $k' = k - \Sigma_{s-1}$, $k_1 = \lceil k'/2 \rceil$. $d = \lfloor s_2/(s-t) \rfloor$, $k_2 = k_1 + s_1(s_1+1)/2$, $s_2 = \lfloor (2k_2 + 1/4)^{1/2} - 1/2 \rfloor$, $s_3 = s_2 - q_0 - 1$, $\Delta = k' - 2k_3$. $t_1 = \lfloor \Delta / q_0 \rfloor$, $k_3 = (s_2 - 1)s_2/2 - (s_1 - 1)(s_1 + 1)/2 - s_3(s_3 + 1)/2$, $p = s_2 - (s_2 - t)d$, $\lceil \cdot \rceil$ – округление к большему целому числу, $\lfloor \cdot \rfloor$ – округление к меньшему целому числу, $[\cdot]$ – округление к ближайшему целому числу.

Предложение [5]. Сложность универсального хеширования по кривым $y^q - y = x^{q_0}(x^q - x)$, где $q = 2q_0^2$ и $q_0 = 2^s$ над полем F_q определяется выражением

$$N_{опер} = k + s^3/3 + s^2/2 + 2s - 1, \text{ если } s \leq q_0. \quad (4$$

$$N_{опер} = k + q_0^3/3 + q_0^2/2 + (s - q_0)(2q_0 - 1) + 2s - 1, \text{ если } s > q_0. \quad (5$$

где $s = (3k)^{1/3}$.

Замечание 1. Асимптотическая оценка для вероятности коллизии (2) имеет значение $\varepsilon = (3k)^{1/3}/q$ для $k < q_0(q-1)$.

Каскадное хеширование $Ch_t(M)$ по функциональному полю кривой Сузуки имеет следующее представление.

Определение 4. Пусть F_q, $q = p^2$ – расширенное конечное поле, M – сообщение и $M = M_1 \| M_2 \| \ldots \| M_t$. Алгоритм вычисления хеш кода в каскадной конструкции определяется выражением $Ch_t(M) = Sh_q(PLh_q(M_1) \| PLh_q(M_2) \| \ldots \| PLh_q(M_t))$, где Sh_q, PLh_q – универсальные схемы хеширования по кривой Сузуки и проективной прямой.

Утверждение 3. Пусть F_q – конечное поле, M – сообщение и $M = M_1 \| M_2 \| \ldots \| M_t$. $Ch_t(M)$ – хеширование вида (1), где $AGh_1 = PLh_q$, $AGh_2 = Sh_q$ есть универсальные схемы

хеширования по проективной прямой $\varepsilon_{PL}-U(q,q^{k_1},q)$, $\varepsilon_{PL}=k_1/q$ и кривой Сузуки $\varepsilon_S-U(q^2,q',q)$, $\varepsilon_S=(3t)^{1/3}/q$ соответственно. Тогда $Ch_t(M)$ определяет универсальное семейство хеш функций $\varepsilon-U\left(q^3,q^k,q\right)$, $\varepsilon=(3k)^{1/4}/q$, $k=k_1t$ – число слов данных, $0<k\le gq$, $g=q_0(q-1)$ – род кривой Сузуки.

Доказательство. Хеширование на каждом каскаде является универсальным, и каскадная хеш функция также является универсальной. Пространство ключей определяется произведением числа ключей первого и второго каскадов и равно $q^2\cdot q=q^3$. Пусть k число слов данных и k' – размер блока данных, $t=k/k'$. Наименьшая вероятность коллизии в силу утверждения 1 реализуется в случае $\varepsilon_S=\varepsilon_{PL}$. Подставим в выражения для ε_S и ε_{PL} значения k' и $t=k/k'$, получим $(3t)^{1/3}=k'$, $k'=(3k)^{1/4}$ и оценку для вероятности $\varepsilon=(3k)^{1/4}/q$. Так как ε_{PL} имеет оценку сверху $\varepsilon_{PL}<1$ для $k'\le q$, а $\varepsilon_S\le(3k)^{1/3}/q$ для $k\le g$, получим $0<k\le gq$, где g - род кривой Сузуки.

Замечание 4.

1. Результаты каскадного хеширования PLh_q-Sh_q являются наилучшими среди схем, где на первом каскаде используется PLh_q хеширование, и обеспечивают хеширование наибольшей длины данных.

2. Недостатком каскадного хеширования PLh_q-Sh_q является повышенная сложность хеширования. Вычисления по PLh_q каскаду выполняются на одной рациональной функции со сложностью $\sim k'$ и по Sh_q каскаду на четырёх рациональных функциях – со сложностью $\sim 2t+1.04t^{2/3}+2\sqrt[3]{3}t^{1/3}$ (4). С учетом $t=k/k'$ и $k'=(3k)^{1/4}$ получим оценку для числа вычислений $k+2t+1.04t^{2/3}+2\sqrt[3]{3}t^{1/3}=k+1.52k^{3/4}+0.87k^{1/2}+2.63k^{1/4}$. Относительное увеличение сложности по сравнению с PLh_q хешированием составляет $\sim 1+1.52k^{-1/4}$.

Возможны комбинации других универсальных хеш функций в двух каскадной схеме хеширования, свойства представлены утверждением 4.

Утверждение 4. Пусть F_q – конечное поле, M – сообщение и $M=M_1\|M_2\|...\|M_r$, $Ch_t(M)$ – хеширование вида (1), где $AGh_1=Hh_q$, $AGh_2=Sh_q$ – универсальные схемы хеширования по кривой Эрмита и кривой Сузуки соответственно. Тогда $Ch_t(M)$ определяет универсальное семейство хеш функций $\varepsilon-U\left(q^3\sqrt{q},q^k,q\right)$, $\varepsilon=1.43k^{1/5}/q$, $0<k\le q^2\sqrt{q}$.

Замечание 5.

1. Доказательство подобно доказательству утверждения 3. Результаты каскадного хеширования Hh_q-Sh_q являются абсолютно наилучшими среди двух каскадных схем и обеспечивают хеширование наибольшей длины данных.

2. Недостатком каскадного хеширования Hh_q-Sh_q является увеличенная сложность хеширования. Вычисления по Hh_q каскаду выполняются на двух рациональных функциях со сложностью $\sim t(k'+\sqrt{k'})$ и по Sh_q каскаду на четырёх рациональных функциях – со сложностью $\sim 2t+1.04t^{2/3}+2\sqrt[3]{3}t^{1/3}$ (4). С учетом $t=k/k'$ и $k'=1.02k^{2/5}$ получим оценку для числа вычислений $t(k'+\sqrt{k'})+2t+1.04t^{2/3}+2\sqrt[3]{3}t^{1/3}=k+k^{4/5}+2k^{3/5}+k^{2/5}+2.88k^{1/5}$. Относительное увеличение сложности по сравнению с PLh_q хешированием составляет $\sim 1+k^{-1/5}$.

Оценки многокаскадного универсального хеширования $l-Ch_t(M)$ представлены утверждением 5.

Утверждение 5. Пусть F_q, $q=2^{2s+1}$ – расширенное конечное поле, M – сообщение и $M=M_1\|M_2\|...\|M_t$, $l-Ch_t(M)$ – l-каскадное универсальное хеширование по кривой Сузуки. Тогда $l-Ch_t(M)$ определяет универсальное семейство хеш функций $\varepsilon-U\left(q^{2l},q^k,q\right)$. $\varepsilon=\sqrt[3]{3}k^{1/3l}/q$, $0<k\le q^{l+l/2}$, со сложностью вычислений $2k+1.04k^{(3l-1)/3l}+2.88t^{(3l-2)/3l}+2k^{(l-1)/l}+1.04k^{(3l-4)/3l}+2.88t^{(3l-5)/3l}+...$

Доказательство. Хеширование на каждом каскаде является универсальным и каскадная хеш функция также является универсальной. Пространство ключей определяется произведением числа ключей всех каскадов и равно q^{2l}. Наименьшая вероятность коллизии в силу утверждения 1 реализуется, если на каждом каскаде значение вероятности коллизии является наименьшим. Это достигается, если размер данных хеширования k' на каждом каскаде является наименьшим $k'=k^{1/l}$. Подставим в выражения для ε_S значения k', получим $\varepsilon_S=\sqrt[3]{3}k^{1/3l}/q$ и оценку для $0<k\le q^{l+l/2}$. Оценка сложности вычислений определяется тем, что на каждом каскаде число вычислений уменьшается в $k'=k^{1/l}$.

Суммирование по всем каскадам дает результирующее выражение

$$2k+1.04k^{(3l-1)/3l}+2.88t^{(3l-2)/3l}+2k^{(l-1)/l}+1.04k^{(3l-4)/3l}+2.88t^{(3l-5)/3l}+....$$

Параметры многокаскадного универсального хеширования по кривой Сузуки представлены в табл.1.

Таблица 1

Параметры многокаскадного универсального хеширования по кривой Сузуки

Схемы каскадного включения	Параметры универсального хеширования	Оценки сложности вычислений
$Ch_t(M)$, PSh_q-Sh_q	$\varepsilon-U\left(q^3,q^k,q\right)$, $\varepsilon=(3k)^{1/4}/q$, $0<k\le q^2$	$k+1.52k^{3/4}+0.87k^{1/2}+2.63k^{1/4}$
$Ch_t(M)$, Hh_q-Sh_q	$\varepsilon-U\left(q^3\sqrt{q},q^k,q\right)$, $\varepsilon=1.43k^{1/5}/q$, $0<k\le q^2\sqrt{q}$	$k+k^{4/5}+2k^{3/5}+k^{2/5}+2.88k^{1/5}$
$l-Ch_t(M)$, $Sh_q-Sh_q-...$	$\varepsilon-U\left(q^{2l},q^k,q\right)$, $\varepsilon=\sqrt[3]{3}k^{1/3l}/q$, $0<k\le q^{l+l/2}$	$2k+1.04k^{(3l-1)/3l}+2.88t^{(3l-2)/3l}+2k^{(l-1)/l}+1.04k^{(3l-4)/3l}+2.88t^{(3l-5)/3l}+...$

Оценки вероятности коллизии и сложности вычислений для многокаскадного хеширования в конечном поле представлены в табл. 2.

Выводы

1. Каскадное хеширование эффективно увеличивает размер хешируемых данных и выравнивает вероятность коллизии с изменением длины данных.

2. Применение многокаскадного универсального хеширование $l-Ch_t(M)$ с одной и той же функцией хеширования на всех каскадах в $\sqrt[l]{k}$ раз уменьшает вероятность коллизии. Наименьшая вероятность коллизии реализуется в схеме Sh_q-Sh_q. Вычисления в поле ~64 бит обеспечивает доказуемую стойкость $P_{кол}<2^{-57}$ для данных длиной до нескольких *Гбт*.

3. Двухкаскадные схемы хеширования $PLh_q - Sh_q$ и $Hh_q - Sh_q$ по вероятности коллизии являются эквивалентными. С увеличением размерности поля отличие от $PLh_q - Sh_q$ и $Hh_q - Sh_q$ хеширования становится меньше.

Таблица 2

Оценки вероятности коллизии и сложности вычислений для многокаскадного хеширования по кривой Сузуки над полем F_q

Схемы каскадов	F_q	Вероятность коллизии для данных размером L /сложность вычислений			Размер ключей, бит	Размер хеш кода, бит
		1 *Кбт*	1 *Мбт*	1 *Гбт*		
$Ch_t(M)$, $PSh_q - Sh_q$	$q = 2^{31}$	$2^{-28,7}/2^8+2^6$	$2^{-26,2}/2^{18}+2^{16}$	$2^{-3,7}/2^{28}+2^{20}$	93	31
	$q = 2^{63}$	$2^{-61}/2^7+2^5$	$2^{-58,5}/2^{17}+2^{10}$	$2^{-56}/2^{27}+2^{20}$	189	63
$Ch_t(M)$, $Hh_q - Sh_q$	$q = 2^{31}$	$2^{-29}/2^8+2^7$	$2^{-27}/2^{18}+2^{15}$	$2^{-25}/2^{28}+2^{23}$	109	31
	$q = 2^{63}$	$2^{-61}/2^7+2^6$	$2^{-59}/2^{17}+2^{14}$	$2^{-57}/2^{27}+2^{22}$	221	63
$Ch_t(M)$ $Sh_q - Sh_q$	$q = 2^{31}$	$2^{-29}/2^9+2^7$	$2^{-27,5}/2^{19}+2^{14}$	$2^{-26}/2^{29}+2^{24}$	124	31
	$q = 2^{63}$	$2^{-61}/2^8+2^6$	$2^{-59,5}/2^{18}+2^{13}$	$2^{-58}/2^{28}+2^{23}$	148	62

Список литературы: 1. *Халимов Г.З.* Каскадное универсальное хеширование с использованием АГК кодов / *Халимов Г.З., Иохов А.Ю.* // Восточно-европейский журнал передовых технологий. – Х., 2005. – Вып. 2/2(14). – С. 111–119. 2. *Халимов Г.З.* Багатократне універсальне хешування // *Халимов Г.З.* // Спеціальні телекомунікаційні системи та захист інформації : Зб. наук. праць. – Київ : ДССЗ та ЗІ, 2010. – Вип. 2(18). – С.43-49. 3. *Torres F.* The Deligne-Lusztig curve associated to the Suzuki group / *F. Torres* // arXiv:alg-geom/9706012v1 26Jun. – 1997. 4. *Халимов Г.З.* Универсальное хеширование по кривой Сузуки / *Г.З. Халимов, Е.В. Котух* // Прикладная радиоэлектроника. – Харьков : ХНУРЭ. – 2011. – Т.10, № 2. – С.80-86. 5. *Халимов Г.З.* Универсальное хеширование по кривой Сузуки / *Г.З. Халимов, Е.В. Котух* // Восточно-европейский журнал передовых технологий. – Х., 2011. – Вып. 3/9(51). – С. 10–15.

Харьковский национальный университет радиоэлектроники

Поступила в редколлегию 25.08.2011