УДК 621.391

О.Г. Король[1], Л.Т. Пархуць[2], С.П. Евсеев[1]

[1] *Харьковский национальный экономический университет, Харьков*
[2] *Национальный университет "Львовская политехника", Львов*

ИССЛЕДОВАНИЕ СВОЙСТВ МОДУЛЯРНЫХ ПРЕОБРАЗОВАНИЙ И МЕТОДОВ ХЕШИРОВАНИЯ ИНФОРМАЦИИ НА ИХ ОСНОВЕ

Исследованы свойства модулярных преобразований и построенных на их основе методов бесключевого хеширования информации (MASH-1 и MASH-2), а также методов ключевого хеширования, построенных на основе алгоритмов MASH-1 и MASH-2 при смене вектора инициализации в качестве секретных ключевых данных. Исследованы различные виды цикловых функций в схеме итеративного хеширования, построенные с использованием модулярных преобразований, задача инвертирования которых эквивалентна решению одной из известных теоретико-сложностных задач.

Ключевые слова: *модулярные преобразования, цикловые функции, ключевое хеширование.*

Вступление

Применение многослойных схем ключевого хеширования позволяет строить эффективные механизмы контроля целостности и аутентичности информации в телекоммуникацион-ных системах и сетях. Однако известные многослойные конструкции (на примере алгоритма UMAC) наряду с высокими показателями быстродействия и криптографической стойкости за счет применения криптографического слоя преобразования (с использованием блочного симметричного шифра) теряют свойства универсального хеширования, что приводит к ухудшению коллизионных свойств формируемых кодов аутентификации сообщений [1; 3]. Перспективным направлением исследований в этом смысле является разработка и теоретическое обоснование новых схем ключевого хеширования, позволяющих обеспечить как высокие коллизионные свойства (с сохранением свойств универсального хеширования), так и высокие показатели безопасности.

Целью работы является исследование свойств модулярных преобразований и построенных на их основе методов бесключевого хеширования информации (MASH-1 и MASH-2), а также методов ключевого хеширования, построенных на основе алгоритмов MASH-1 и MASH-2 при смене вектора инициализации в качестве секретных ключевых данных, различных видов цикловых функций в схеме итеративного хеширования, построенных с использованием модулярных преобразований, задача инвертирования которых эквивалентна решению одной из известных теоретико-сложностных задач.

Основная часть

Исследование свойств модулярных преобразований и методов хеширования информации на их основе. Модулярные преобразования широко используются при построении криптографических алгоритмов преобразования информации, в том числе при построении ассиметричных средств защиты информации и протоколов распространения ключевых данных [1; 2; 4], для формирования псевдослучайных последовательностей [1; 2], методов хеширования и других механизмов защиты информации [1 – 4].

Проведенный анализ [1 – 4] показывает, что модулярные преобразования применяются на сегодняшний день при построении бесключевых хеш-функций. Так в четвертой части международного стандарта ISO/IEC 10118-4 определены две бесключевые функции хеширования MASH-1 и MASH-2, которые используют модулярную арифметику, а именно модульное возведение в степень для построения хеш-кода [4]. Само название функций MASH-1 и MASH-2 происходит от скрашенного Modular Arithmetic Secure Hash (безопасное хеширование на основе модулярной арифметики), подчеркивающего применение модулярных преобразований при формировании хеш-образа.

В основе построения хеш-функций MASH-1 и MASH-2 лежит использование итеративной цикловой функции, которая определяется через модулярное возведение в степень (в простейшем случае через модулярное возведение в квадрат). В данном случае используются RSA-подобные модули N, длина которых обеспечивает необходимую стойкость. Число N должно быть трудно разложимым на множители, на чем и основывается стойкость алгоритма. Размер модуля N определяет длину блоков обрабатываемого сообщения, а также размер хеш-кода (например, 1025-битний модуль обеспечивает формирование 1024-битного хеш-кода).

Наиболее эффективные схемы итеративного хеширования с использованием модулярной арифметики базируются на модулярном возведении в

квадрат, т.е. на применении цикловой функции вида:

$$f(x_i, H_{i-1}) = (x_i \oplus H_{i-1})^2 \bmod N, \qquad (1)$$

или

$$f(x_i, H_{i-1}) = \left((x_i)^2 \oplus H_{i-1}\right)^2 \bmod N. \qquad (2)$$

В определенных международным стандартом ISO/IEC 10118-4 хеш-функциях MASH-1 и MASH-2 использованы следующие цикловые функции:

$$f(x_i, H_{i-1}) = \left(\left(\left((x_i \oplus H_{i-1}) \vee A\right)^2 \bmod N\right) \perp n\right) \oplus H_{i-1} \qquad (3)$$

и

$$f(x_i, H_{i-1}) = \left(\left(\left((x_i \oplus H_{i-1}) \vee A\right)^{2^8+1} \bmod N\right) \perp n\right) \oplus H_{i-1} \qquad (4)$$

соответственно, где $\vee$ – операция побитного логического ИЛИ; $\oplus$ – суммирование по модулю 2 (XOR); $\perp n$ – сохранение младших n-разрядов m-разрядного результата.

Алгоритм вычисления значения хеш-функции MASH-1 имеет следующий вид [1; 2].

Вход. Двоичная строка x длиной $0 \le b \le 2^{n/2}$.

Выход. n-розрядний хеш-код от строки x, длиной приблизительно равный длине модуля N.

1. Системные установки и определение констант. Установить RSA-подобный модуль N = pq длиной m бит, где p и q случайно выбранные большие простые числа, которые сохраняются в секрете. Определить двоичную длину n хеш-кода, как наибольшее произведение числа 16, т.е. длина хеш-кода определяется из условия $n = 16 \times s < m$, где s – наибольшее целое, удовлетворяющее указанному ограничению.

Как вектор инициализации выбрать $H_0 = 0$. Определить n-битное целое число как константу $A = f00...00_x$.

2. Предобработка. Дополнить, если необходимо, строку x нулевыми битами, для того, чтобы получить двоичную строку длиной $t \times n/2$ для наименьшего $t \ge 1$. Разделить дополненный текст на n/2-розрядные блоки x_1 ..., x_t и прибавить последний блок x_{t+1}, который содержит n/2-розрядное представление числа b.

3. Расширение. Расширить каждый x_i блок в n-розрядный блок y_i путем вставки между 4- разрядными полубайтами блока x_i комбинации из четырех единиц (1111). Последний блок y_{t+1} формируется аналогичным способом за исключением того, что вставляется комбинация 1010.

4. Цикловая функция. Для всех $1 < i \le t + 1$ отобразить два n-розрядних входных блока (H_{i-1}, y_i) в один n-розрядный блок в соответствии с выражением (3).

5. Окончание. В качестве хеш-кода принимается n-розрядный блок H_{t+1}.

Алгоритм MASH-2 отличается от алгоритма MASH-1 только показателем степени в цикловой функции, которая имеет вид (4).

В табл. 1 приведены результаты сравнительного анализа показателей эффективности некоторых бесключевых функций хеширования, в том числе и хеш-функции на модулярной арифметике MASH-1 и MASH-2. В табл. 1 указаны длины соответствующих хеш-кодов, применяемые преобразования, оценки скорости обработки (хеширования) информации, а также обеспечиваемый уровень стойкости, соответствующий одной из моделей безопасности, введенных в проекте NESSIE [2].

Таблица 1

Результаты сравнительного анализа некоторых бесключевых функций хеширования

Хеш-функция	Длина хеш-кода	Применяемые преобразования	Скорость обработки данных	Модель безопасности (по NESSIE)
SHA-2	256, 384, 512	логические и арифметические	$10^8..10^9$ бит/с	Практическая секретность (Practical Security)
Whirlpool	512	В конечных полях Галуа	$10^7..10^8$ бит/с	Практическая секретность (Practical Security)
ГОСТ 34311-95	256	Блочное симметричное шифрование	$10^7..10^8$ бит/с	Практическая секретность (Practical Security)
RIPEMD-160	160	Логические и арифметические	$10^8..10^9$ бит/с	Практическая секретность (Practical Security)
MASH-1	*	Модулярное возведение в квадрат	$10^5..10^6$ бит/с	Доказуемая безопасность** (“Provable” Security)
MASH-2	*	Модулярное возведение в степень $2^8+1 = 257$	$10^4..10^5$ бит/с	Доказуемая безопасность** (“Provable” Security)

** Определяется размерностью модуля преобразований*

*** Если параметры модульного возведения в степень соответствуют ограничениям на RSA-подобные системы*

Проведенный анализ показал, что основным недостатком функций хеширования MASH-1 и MASH-2 является низкая скорость формирования хеш-кода. Фактически она определяется скоростью RSA-подобного шифрования, которое на 2 – 3 порядка ниже скорости шифрования современными блочно-симметричными шифрами. Тем не менее, по причине наличия возможности использования существующих программных и аппаратных средств модулярной арифметики, применяемых в несимметричных RSA-подобных криптосистемах, а также по причине возможности обеспечения доказуемого уровня безопасности (по классификации моделей безопасности NESSIE) рассматриваемые бесключевые хеш-функции MASH-1 и MASH-2 были стандартизированы [2; 4].

Следует, однако, отметить, что алгоритмы хеширования MASH-1 и MASH-2 не в полной мере соответствуют ограничениям на параметры модульного возведения в степень, которые установлены для RSA-систем (а соответственно и обеспечиваемой модели доказуемой безопасности).

Действительно, по спецификации криптографической RSA-системы, обеспечивающей доказуемую безопасность (по модели безопасности NESSIE) значение модульной экспоненты e должно быть выбрано из условия

$$\gcd(e,\varphi(N)) = 1, \qquad (5)$$

где gcd(x,y) – наибольший общий делитель чисел x и y.

Т.е. значение экспоненты e не должно содержать общих делителей с числом (значением функции Эйлера) φ(N):

$$\varphi(N) = (p - 1)(q - 1),\ N = pq.$$

По спецификации алгоритмов MASH-1 и MASH-2 это условие может не выполняться. Так, например, в алгоритме MASH-1 показатель степени установлен равным e = 2, что при нечетных значениях простых чисел p и q всегда нарушает условие (5).

В алгоритме MASH-2 показатель степени установлен равным $e = 2^8+1 = 257$, что также не всегда влечет выполнение условия (5): значение функции Эйлера может, например, делиться на показатель степени $e = 2^8+1 = 257$, и в этом случае

$$\gcd(e,\varphi(N)) = 257.$$

Таким образом, модель доказуемой безопасности (по классификации моделей безопасности NESSIE) может быть применена к алгоритмам MASH-1 и MASH-2 только условно. Полного соответствия задачи нахождения прообраза или секретного ключа в схеме хеширования и теоретико-сложностной задачи факторизации (или задачи RSA) не наблюдается.

Рассмотрим цикловые функции MASH-1 и MASH-2 на предмет построения ключевых универсальных хеширующих функций. Рассмотрим вариант хеширования, когда начальное состояние (вектор инициализации) задается некоторым ключевым правилом, т.е. выберем H_0 = Key. В этом случае имеем некоторый класс хеш-функций, зависящих от параметра Key.

Для экспериментальной проверки свойств универсального хеширования была разработана программная реализация алгоритмов хеширования MASH-1 и MASH-2 при смене значений вектора инициализации секретным ключом. Листинг исходного кода программной реализации приведен в приложении Г.

Для проведения экспериментальных исследований выбраны следующие параметры: $p = 17$, $q = 19$, $N = 323$. Исследования состояли в проверке условий универсального хеширования при полном переборе всех значений векторов инициализации ($Key = 0, ..., 2^m - 1$, $m = 8$) по выборке из генеральной совокупности значений информационных блоков. Полученные результаты сведены в табл. 2.

Таблица 2

Результаты исследований коллизионных свойств ключевого хеширования, построенных на основе алгоритмов MASH-1 и MASH-2 при смене значений вектора инициализации секретным ключом

	На основе алгоритма MASH-1	На основе алгоритма MASH-2
$\tilde{m}(n_1)$	41,42	0
$\tilde{D}(n_1)$	42,74	0
$P_д = P(\lvert\tilde{m}(n_1) - m(n_1)\rvert < 5)$	0,98	≈ 1
$\tilde{m}(n_2)$	3,99	1
$\tilde{D}(n_2)$	0,01	0
$P_д = P(\lvert\tilde{m}(n_2) - m(n_2)\rvert < 0,025)$	0,99	≈ 1
$\tilde{m}(n_3)$	0,26	0,31
$\tilde{D}(n_3)$	0,21	0,22
$P_д = P(\lvert\tilde{m}(n_3) - m(n_3)\rvert < 0,1)$	0,97	0,97

Исследования проводились над выборкой, объема N = 100, для формирования каждого элемента выборки рассчитывался максимум по множеству из M = 100 кортежей элементов. Таким образом, общий объем формируемых наборов составил NM = 10^4. Для каждого проведенных N = 100 экспериментов оценивались математические ожидания $m(n_1)$, $m(n_2)$ и $m(n_3)$, дисперсии $D(n_1)$, $D(n_2)$ и $D(n_3)$, а также для фиксированной точности ε рассчитывались соответствующие доверительные вероятности $P(|\tilde{m}(n_i) - m(n_i)| < \varepsilon)$, $i = 1, 2, 3$. Подробно предложенная методика статистических исследований коллизионных свойств описана в [5].

Анализ данных, приведенных в табл. 2, позволяет утверждать об адекватности полученных экспериментальных результатов. Для фиксированной точности ε получены высокие значения доверительной вероятности, что свидетельствует об обоснованности и достоверности полученных результатов, соответствии их статистическим свойствам всей генеральной совокупности данных.

Проанализируем полученные результаты статистических исследований, сопоставим их с теоретическим оценками: числом $P_{кол} \cdot |H| = 1$ (для первого критерия), с числом $|H| / |B| = 1$ (для второго критерия) и числом $P_{кол} \cdot |H| = 1$ (для третьего критерия). Как видно из приведенных в табл. 2 данных реализация схемы ключевого хеширования на основе алгоритма MASH-1 при смене значений вектора инициализации секретным ключом не позволяет обеспечить высокие коллизионные свойства. Число возникающих коллизий существенно выше верхней теоретической границы, как по первому, так и по второму критерию, следовательно, такая конструкция не является схемой универсального и, тем более, строго универсального хеширования. Этот результат получен с высокой доверительной вероятностью $P_д = P(|\tilde{m}(n_i) - m(n_i)| < \varepsilon) > 0,9$, $i = 1, 2, 3$ для высокой точности. Так для первого критерия доверительный интервал составил $41,42 \pm 5$ (доверительная вероятность 0,98), для второго критерия доверительный интервал составил $3,99 \pm 0,025$ (доверительная вероятность 0,99) и для третьего критерия доверительный интервал составил $0,26 \pm 0,1$ (доверительная вероятность 0,97). Рассматриваемая схема ключевого хеширования на основе алгоритма MASH-1 при смене значений вектора инициализации секретным ключом удовлетворяет только третьему критерию ($\tilde{m}(n_3) = 0,26$).

Использование ключевого хеширования на основе алгоритма MASH-2 при смене значений вектора инициализации секретным ключом напротив, обеспечивает высокие коллизионные характеристики универсального хеширования. По всем трем критериям полученные оценки лежат ниже верхней теоретической границы $\tilde{m}(n_i) < 1$, $i = 1, 2, 3$. Это положение подтверждено, практически, со 100% вероятностью. Так для первого и второго критерия значения дисперсий $D(n_1)$ и $D(n_2)$, характеризующих рассеивание значений числа правил хеширования (правил формирования МАС), при которых выполняются равенства, относительно их математических ожиданий $m(n_1)$ и $m(n_2)$, соответственно, равны нулю, что означает идентичность полученных результатов во всех проведенных опытах и, практически достоверно имеем $m(n_1) = 0$, $m(n_2) = 0$. Полученная оценка по третьему критерию также лежит ниже верхней теоретической оценки ($\tilde{m}(n_3) = 0,31$) и это значение подтверждено с высокой доверительной вероятностью

$$P_д = P(|\tilde{m}(n_3) - m(n_3)| < 0,1) = 0,97$$

для фиксирован-ной точности (доверительный интервал равен $0,31 \pm 0,1$).

Объяснение такого поведения модулярных преобразований в схемах MASH-1 и MASH-2 лежит в выбранных параметрах модульной экспоненты. Так, для алгоритма MASH-1 цикловая функция (3) предполагает значение модульной экспоненты равным e = 2, что всегда нарушает условие (5). В алгоритме MASH-2 показатель степени установлен равным e = 2^8+1 = 257, что для выбранных параметров $p = 17$, $q = 19$, $N = 323$ удовлетворяет ограничению (5): $\gcd(e, \phi(N)) = \gcd(257, 288) = 1$.

Следовательно, ключевое хеширование, построенное на основе модулярных преобразований, в некоторых случаях позволяет обеспечить свойства универсального и строго универсального хеширования. Для выполнения этих свойств необходимо выполнение условия (5), что и демонстрирует при выбранных параметрах схема на основе алгоритма MASH-2.

Таким образом, проведенные исследования показали, что применение преобразований с использованием модулярной арифметики позволяет строить универсальные и строго универсальные классы хеширующих функций, которые с одной стороны позволяют обеспечить высокие коллизионные свойства, с другой стороны, при выполнении определенных ограничений на значение модулярной экспоненты обеспечивают высокие показатели безопасности и применимость модели доказуемой стойкости.

Основными недостатками подобных конструкций являются:

очень высокая сложность преобразований, которая обусловлена использованием в качестве цикловой функции модулярного возведения в степень.

Фактически сложность применяемых преобразований выше сложности блочного симметричного шифрования на 2 – 3 порядка, что и обуславливает соответствующее повышение времени формирования кодов аутентификации сообщений (см. табл.1);

формирование кодов аутентификации сообщений с использованием ключевого хеширования, построенного на основе алгоритма MASH-1 с изменяемыми векторами инициализации, не позволяет строить универсальные и строго универсальные классы хеш-функций (см. табл. 2). Это обусловлено использованием в качестве показателя степени цикловой функции значения e = 2, что при нечетных значениях простых чисел p и q всегда нарушает условие (5);

формирование кодов аутентификации сообщений с использованием ключевого хеширования, построенного на основе алгоритма MASH-2 с изменяемыми векторами инициализации, в некоторых случаях (при выполнении условия (5)) позволяет строить универсальные и строго универсальные классы хеш-функций (см. табл. 2). Однако не для всех значений начальных параметров (простых чисел p и q) это условие выполнимо;

модель доказуемой безопасности (по классификации моделей безопасности NESSIE) к рассмотренным методам хеширования может быть применена только при выполнении определенных ограничений на значения экспоненты и модуля преобразования, т.е. при выполнении ограничений для RSA-подобных систем на параметры модульного возведения в степень. Полного соответствия задачи нахождения прообраза или секретного ключа в схеме хеширования и теоретико-сложностной задачи дискретного логарифмирования (или задачи RSA) для рассмотренных методов хеширования не наблюдается.

Выводы

Таким образом, как показывают проведенные исследования, применение модулярных преобразований потенциально может решить задачу построения коллизионно стойких универсальных хеширующих функций с доказуемым уровнем безопасности, однако для этого необходимо устранить выявленные противоречия. Перспективным направлением дальнейших исследований является разработка метода ключевого универсального хеширования доказуемой стойкости на основе модулярных преобразований с учетом выявленных закономерностей.

Список литературы

1. Кузнецов О.О. Захист інформації в інформаційних системах / О.О. Кузнецов, С.П. Євсеєв, О.Г. Король. – Х.: Вид. ХНЕУ, 2011. – 504 с.

2. . Final report of European project number IST-1999-12324, named New European Schemes for Signatures, Integrity, and Encryption, April 19, 2004 – Version 0.15 (beta), Springer-Verlag.

3. Столлингс В. Криптография и защита сетей: принципы и практика, 2-е изд. : пер. с англ. / В. Столлингс – М.: издательский дом «Вильям», 2001. – 672 с.

4. Король О.Г. Исследование методов обеспечения аутентичности и целостности данных на основе односторонних хеш-функций/ / О.Г. Король, С.П. Евсеев // Захист інформації: науково-технічний журнал. –Спецвипуск (40). – 2008. – С. 50 – 55.

5. Кузнецов А.А. Исследование коллизионных свойств кодов аутентификации сообщений UMAC / А.А. Кузнецов, О.Г. Король, С.П. Евсеев // Прикладная радиоэлектроника. – Х.: ХНУРЭ, 2012. – Том 11 № 2. – С. 171-183.

Поступила в редколлегию 22.04.2013

Рецензент: д-р техн. наук, проф. В.А. Хорошко, Национальный авиационный университет, Киев.

ДОСЛІДЖЕННЯ ВЛАСТИВОСТЕЙ МОДУЛЯРНИХ ПЕРЕТВОРЕНЬ І МЕТОДІВ ГЕШУВАННЯ ІНФОРМАЦІЇ НА ЇХ ОСНОВІ

О.Г. Король, Л.Т. Пархуць, С.П. Євсеєв

Досліджено властивості модулярних перетворень і побудованих на їх основі методів безключового гешування інформації (MASH-1 і MASH-2), а також методів ключового гешування, побудованих на основі алгоритмів MASH-1 і MASH-2 при зміні вектору ініціалізації в якості секретних ключових даних. Досліджено різні види циклових функцій у схемі ітеративного гешування, які побудовані з використанням модулярних перетворень, завдання інвертування яких еквівалентне вирішенню однієї з відомих теоретико-складних завдань.

Ключові слова: модулярні перетворення, циклові функції, ключове гешування.

INVESTIGATION OF MODULAR TRANSFORMATION AND HASHING METHODS INFORMATION BASED ON THEM

O.G. Korol, L.T. Parkhuts , S.P. Evseev

The properties of modular transformations and built on them hashing methods Keyless information (MASH-1 and MASH-2), as well as key hashing methods that are based on algorithms MASH-1 and MASH-2 by changing the initialization vector as the secret key data. Investigated various kinds of cycle functions in the scheme of iterative hash, built using modular transformations, the problem of inverting which is equivalent to solving one of the well-known complexity-theoretic problems.

Keywords: modular transformation, cyclic function key hashing.