УДК 621.322

В.Е. ЧЕВАРДИН, Я.В. ЯНСОНС

Полтавский военный институт связи, Украина

ПРИМЕНЕНИЕ МОДУЛЯРНОЙ АРИФМЕТИКИ ДЛЯ ФОРМИРОВАНИЯ КЛЮЧЕВЫХ ХЕШ-ФУНКЦИЙ

Представлены результаты анализа подходов к построению универсальных хеш-функций. Рассмотрены их особенности, достоинства и недостатки. Предложен способ ключевого хеширования на основе применения модулярной арифметики. С использованием модели оценки параметров универсальных хеш-функций проведен сравнительный анализ параметров предложенного способа ключевого хеширования.

модулярная арифметика, универсальный класс хеш-функций, вероятность коллизий, вычислительная стойкость

Постановка задачи

Активное внедрение новых информационно-телекоммуникационных технологий и электронно-вычислительных средств во все сферы жизнедеятельности нашего государства, их открытость и широкая доступность обусловили рост нарушений целостности и аутентичности информации в системах управления критического применения (СУКП). Известно, что обеспечение целостности и аутентичности данных является одной из главных задач подсистемы защиты информации современных СУКП. В связи с этим **актуальным** направлением в криптографии является разработка и исследование механизмов аутентификации данных.

Как показали исследования [1, 2], особое место среди механизмов обеспечения целостности и аутентичности данных, рассматриваемых в теории аутентификации и частично в теории секретных систем, занимают ключевые итерационные хеш-функции. В терминах введенных в [3], они называются схемами аутентификации и делятся на три категории в зависимости от обеспечиваемой стойкости: вычислительно, доказуемо и безусловно стойкие. Построение безусловно стойких схем аутентификации [3] связано с необходимостью передачи по закрытому каналу связи ключевых данных существенно превосходящих по размеру информационные данные. Аналогичным примером в теории секретных систем является шифр Вернама. Такие схемы не получили широкого применения. В сравнении с вычислительно стойкими схемами аутентификации (пример – блочно-симметричные схемы) доказуемая стойкость схем аутентификации сводится к одной из существующих теоретико-сложностных задач (RSA-подобные схемы и т.д.). Примером доказуемо стойких схем аутентификации являются стандарты ключевых хеш-функций MASH-1, MASH-2 [4]. Основным достоинством MASH-1, MASH-2 является сведение задачи взлома секретного ключа к решению задачи факторизации числа. Применение хеш-функций MASH-1, MASH-2 позволило использовать существующие библиотеки криптографических функций для схем цифровой подписи, обмена ключами и механизмов шифрования. Вычислительная стойкость схем хеширования MASH-1, MASH-2 зависит от размера числа p, по модулю которого выполняются криптопреобразования. Недостатком хеш-функций MASH-1, MASH-2 является невозможность теоретически обосновать верхнюю границу вероятности коллизий.

В отличие от хеш-функций MASH-1, MASH-2 обосновать верхнюю границу вероятности коллизий для хеш-функции позволяет доказательство ее универсальных свойств. Детальные исследования уни-

версальных ключевых хеш-функций представлены в работах [6, 7]. В связи с этим, **целью данной статьи** является применение теоретико-сложностной задачи для построения универсальных ключевых хеш-функций. Рассмотрим существующие подходы к построению универсальных хеш-функций.

Анализ подходов к построению универсальных хеш-функций

Пользуясь известными определениями [5] для хеш-функций, проведем анализ известных представителей классов $\varepsilon - U(N, n, r)$ и $\varepsilon - SU(N, n, r)$, где N – количество функций (правил, ключей) отображения множества открытых текстов $\sum^n$ мощности n в множество хеш-кодов $\sum^r$ мощности r.

Примерами отечественных разработок являются универсальные функции хеширования на основе РС-кодов $\frac{k}{q} - U(q, q^k, q)$, на основе ортогональных массивов $\frac{1}{q^b} - SU\left(q^{a+b}, q^a, q^b\right)$, детальное исследование которых представлено в [6 – 8]. Достоинством полиномиальных схем и схем на РС-кодах является высокая скорость хеширования, недостатком – линейность, что позволяет, решая систему линейных уравнений, вскрывать ключевые данные. В современных работах [9] был предложен новый подход к построению ключевых хеш-функций, позволяющий избежать уязвимости рассмотренных схем. Это происходит за счет композиции полиномиальных схем с блочно-симметричным шифром и более сложной реализации схем генерации ключевых последовательностей для хеширования (UMAC, HMAC, TTMAC и др.). При этом отсутствие в открытой печати доказательств универсальности самой схемы хеширования делает ее сомнительной. Заметим, что задача раскрытия блочно-симметричного шифра не является теоретико-сложностной, поэтому такие схемы позволяет обеспечить лишь вычислительную стойкость аутентификации.

Распространенным методом выработки кодов аутентификации сообщений являются схемы UMAC [9]. В формализованном виде UMAC можно представить выражением:

$$UMAC(K, M, Nonce) =$$

$$UHASH(K, M) \oplus PDF(K, Nonce).$$

Функции UHASH и PDF ключевые: UHASH использует ключ для выбора конкретной хеш-функции из семейства универсальных хеш-функций; PDF использует ключ для внутреннего шифрования блочным шифром AES. Число подключей должно быть сгенерировано с использованием K и это является значением ключевой дифференцированной функции. UHASH состоит из трех раздельных уровней, где каждый из них основан на различных семействах хеш-функций: сжатие L1HASH32, хеширование фиксированной длины L2HASH32, усиление и свертка L3HASH32.

В общем виде структурная схема UMAC представлена на рис. 1.

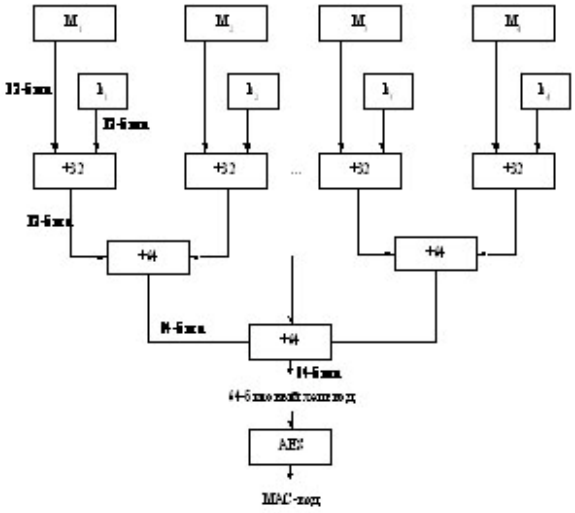

Рис. 1. Схема UMAC

Достоинством этой схемы является высокая скорость хеширования, и нелинейные свойства последовательностей на выходе (после блока AES). Стойкость UMAC зависит от криптографической функции AES, которая используется для дифференциации ключевой последовательности, а также для вычисления значения *pad*. Основным требованием к такой схеме является то, что вероятность нахождения образа должна быть значительно выше установ-

ленной вероятности коллизий. Потенциальная сложность атаки на UMAC эквивалентна сложности атак на AES.

Заметим, что UMAC эффективен для аутентификации потоков данных большого объема со сжатием данных в 128 раз и выше. Недостатком является сложность процедуры выработки ключей, поэтому UMAC не применяется для приложений, где необходима частая смена ключей (т.е. где каждый ключ используется для аутентификации малого объема данных). Учитывая рассмотренные особенности схемы UMAC можно предположить, что для современных СУКП, особенностями которых являются сравнительно небольшой размер обрабатываемых кодограмм, транзакций (служебных сообщений) и высокие требования к стойкости схем аутентификации и вероятности коллизий 2^{-64}, 2^{-96}, 2^{-128}, 2^{-160}, применение схем подобных UMAC не позволит обеспечить возросшие вероятностно временные требования к механизмам обеспечения аутентичности данных на основе ключевого хеширования.

Одним из возможных подходов к построению универсальных хеширующих функций является применение модулярной арифметики, что позволит обеспечить доказуемую аутентификацию данных и обосновать универсальные свойства хеш-функции.

Применение криптопреобразований в простом поле в целях ключевого хеширования

Поставленная цель достигается за счет применения теоретико-сложностной задачи в основном шаге хеш-функции. Это реализуется путем представления информационных данных в виде последовательности

$$I = \{M_1, M_2, .., M_n\}$$

ключевых данных большим секретным числом k и хеширования информационных данных по итерационному правилу

$$\begin{cases} h_i = m_i^{\,h_{i-1}+k} \bmod p; \\ h_0 = k, \end{cases} \qquad (1)$$

где h_i – значение хеш-функции на i-й итерации; h_{i-1} – значение хеш-функции на (i–1)-й итерации; k – секретное число; M – значение i-го блока хешируемых данных; p – простое число.

На рис. 2. представлена схема предлагаемого способа ключевого хеширования.

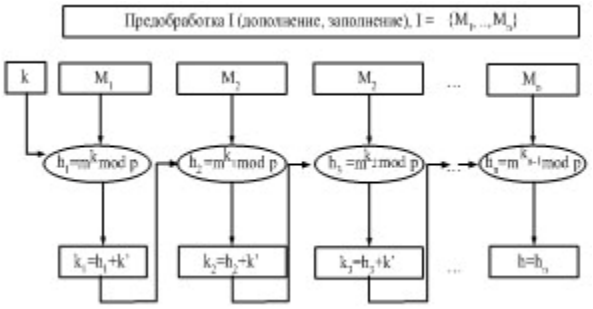

Рис. 2. Схема ключевого хеширования на основе модулярной арифметики

Сущность предложенной схемы хеширования заключается в итерационном возведении чисел $M_i \in \Sigma^n$, в степень $h_{i-1} \in \Sigma^r$ ($h_0 = k$). Ядром хеш-функции является операция возведения чисел в степень, реализующая отображение вида $\Sigma^n \to \Sigma^r$ (нелинейная по своей природе).

Для примера рассмотрим распределение значений предложенной хеш-функции для случая $n = r$, над полем GF(5) (исключим из рассмотрения текст либо ключ, равный $0 = \alpha^{\infty}$ или $1 = \alpha^0$).

Матрица 1

H\M$_i$	1	2	3	4
$f_1 = 1$	1	2	3	4
$f_2 = 2$	2	4	1	3
$f_3 = 3$	3	1	4	2
$f_4 = 4$	4	3	2	1

Для класса хеш-функций $\varepsilon - U(N, n, r)$ с $n = r$, предложенная схема позволяет получить строгое отображение $\Sigma^n \to \Sigma^r$ для любого k, причем при $k_i \neq k_j$ $P_{[H_{k_i}(M)=H_{k_j}(M)]} = 0$ и при $M_i \neq M_j$ $P_{[H_k(M_i)=H_k(M_j)]} = 0$.

Полученное распределение можно представить $\varepsilon - SU(N, n, r)$ классом хеш-функций, где $N = n = r = 4$, а $\varepsilon = 0$.

Матрицу 1 можно представить как совокупность векторов в базисе q –1 либо в базисе q с исключением 0. Из матрицы 4×4 путем перестановки элементов векторов можно получить две матрицы 4×12 матрица A и B. В представленных матрицах A и B

вектора совпадают лишь 1 позиции, т.е. отличаются друг от друга в $q-1=3$ позициях. Если матрицы объединить, число совпадений удваивается.

Матрица А

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	3	4	3	4	1	2	4	1	1	2	3
3	3	4	2	1	3	4	4	1	2	3	1	2
4	4	2	3	4	1	3	1	2	4	2	3	1

Матрица В

	1	2	3	4	5	6	7	8	9	10	11	12
1	1	1	1	2	2	2	3	3	3	4	4	4
2	2	3	4	3	4	1	2	4	1	1	2	3
3	4	2	3	4	1	3	1	2	4	2	3	1
4	3	4	2	1	3	4	4	1	2	3	1	2

Количество полученных векторов в данном случае равно для матрицы А и В отдельно 12, а для обеих $q! = 4! = 4\cdot3\cdot2 = 24$.

Рассмотрим распределение хеш-значений предложенной схемы ключевого хеширования аналогичным образом в виде матрицы, столбцы которой представляются в виде векторов.

Матрица 2

H\M$_i$	**1**	**2**	**3**	**4**	**11**	**12**	**13**	**14**	**21**	**22**	**23**	**24**
f_1 **= 1**	1	2	3	4	2	4	1	3	3	1	4	2
f_2 **= 2**	2	4	1	3	3	1	4	2	1	2	3	4
f_3 **= 3**	3	1	4	2	4	3	2	1	2	4	1	3
f_4 **= 4**	4	3	2	1	1	2	3	4	4	3	2	1
№$_{мат}$	A	B	B	A	B	A	A	B	A	B	B	A
H\M$_i$	**31**	**32**	**33**	**34**	**41**	**42**	**43**	**44**				
f_1 **= 1**	4	3	2	1	1	2	3	4				
f_2 **= 2**	2	4	1	3	4	3	2	1				
f_3 **= 3**	1	2	3	4	3	1	4	2				
f_4 **= 4**	3	1	4	2	2	4	1	3				
№$_{мат}$	A	B	B	A	B	A	A	B				

Из полученной матрицы мы видим, что общее число векторов равно $4!-4 = 4\cdot3\cdot2 = 4^2 + 4 = 24 - 4 = 20$. Так как из множества ключей были исключены $0 = \alpha^{\infty}$ или $1 = \alpha^0$, число векторов в общем случае равно

$$n = (q-1)^2 + q - 1. \qquad (2)$$

Заметим, что в каждой «четверке» векторов (столбцов матрицы) пара векторов принадлежат одной из матриц А или В [10]. Это значит, если вектора одной из матриц А и В совпадают в одной позиции, то для полученной схемы совпадений вдвое больше. С использованием выражений результатов исследования семейства универсальных хеш-функций (при разработке модели оценки параметров универсальной хеш-функции $\frac{t}{q} - \left(q, \frac{q!}{(q-(t+1))!}, q\right)$ [10], предложенный способ ключевого хеширования можно представить в виде семейства универсальных хеш-функций:

$$\frac{2\cdot1}{4} - U(4,\ 20,\ 4) \Rightarrow \frac{1}{2} - U(4,\ 20,\ 4). \qquad (3)$$

Согласно выражения (3), для данного случая универсальный класс хеш-функций с теоретически возможными параметрами имеет вид

$$\frac{2}{4} - U(4,\ \frac{4!}{(4-(2+1))!},\ 4) \Rightarrow \frac{1}{2} - U(4,\ 24,\ 4).$$

Проведенные исследования распределения хеш-значений над полями GF(q) подтвердили полученные выражения. К примеру, над GF(7) выражение (3) примет вид

$$\frac{1}{3} - U(6, 42,\ 6).$$

В общем случае предложенная схема ключевого хеширования соответствует классу

$$\frac{2t}{p} - U(p,\ p^t + p,\ p), \qquad (4)$$

где p – порядок мультипликативной группы целых чисел; t – переменная.

Проведена оценка параметров предложенной схемы ключевого хеширования в сравнении с существующими универсальными классами ключевого хеширования. На рис. 3 представлена зависимость

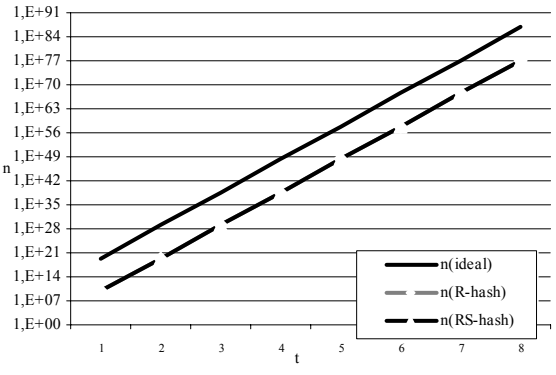

Рис. 3. Сравнительная оценка объема открытых текстов n от t

объема хешируемых данных (длина хешируемой последовательности) от значения ε, при $n = const$ для

классов: на основе РС-кодов $\frac{k}{q} - U(q, q^k, q)$, предложенного класса $\frac{2(t+1)}{q} - U(q, q^t, q)$ и класса $\frac{t}{q} - \left(q, \frac{q!}{(q-(t+1))!}, q\right)$ с теоретически возможными параметрами хеширования.

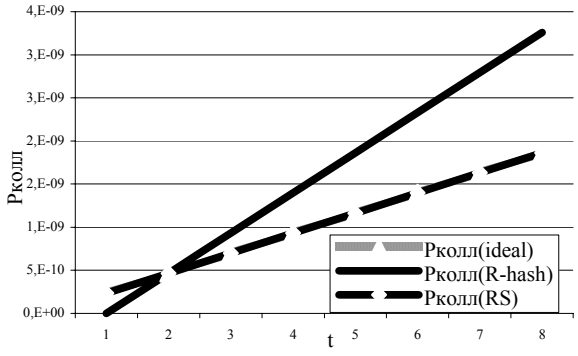

Рис. 4. Сравнительная оценка объема открытых текстов *n* от *t*

Результаты исследований способов ключевого хеширования показали, что предложенная схема ключевого хеширования (4) эффективна для малых объемов открытых текстов. Это подтверждает участок сужения прямых графиков (рис. 4).

Выводы

Таким образом, в статье проведен анализ известной схемы UMAC, выявлены ее достоинства и недостатки. Предложен новый подход к построению ключевых схем хеширования с использованием модулярной арифметики в цикловой функции схемы хеширования. Предложенная схема позволяет обеспечить универсальность хеширования и одновременно доказуемо стойкую аутентификацию данных за счет сведения задачи вскрытия ключевых данных к решению задачи дискретного логарифмирования в простом поле. Полученная схема ключевого хеширования является нелинейной, что позволяет избежать уязвимостей к линейному криптоанализу и избавиться от необходимости внесения дополнительных криптопреобразований. В сравнении со схемой UMAC предложенная схема не требует дополнительных криптопреобразований для выработки подключей хеширования и позволяет обеспечить вероятность коллизий 2^{-64}, 2^{-96}, 2^{-128}, 2^{-160} при переносе вычислительной сложности криптопреобразований из пространственной области во временную. В дальнейшем планируется более глубоко исследовать стойкость предложенной хеш-функции.

Литература

1. Шипли Г. Основы безопасности ИТ // Сети и системы связи. – М., 2003. – №4. – С. 78-82.

2. Сереченко Д. MPLS и безопасность // Сети и системы связи. – М., 2004. – №13 (119). – С. 89-91.

3. Simmons G.I. Authentication theory/coding theory, presented at Crypto'84. – Santa Barbara, CA. – 1984. – P. 19-22.

4. ISO/IEC 10118-4:1998 Information technology – Security techniques – Hash-functions – Part 4: hash-functions using modular arithmetic.

5. Carter J.L., Wegman M.N. Universal classes of hash functions // J. Computer and System Sci. – 1979. – № 18. – P. 143-154.

6. Халимов Г.З., Кузнецов А.А. Аутентификация и универсальное хеширование // Радиотехника. – 2001. – № 119. – С. 95-102.

7. Wegman M., Carter L. New hash functions and their use in authentication and set equality // J. of Computer and System Science. – 1981. – Vol. 22. – P. 265-279.

8. Halevi S., Krawczyk H. MMH: Software Message Authentication in the Gbit/second Rates // J. of Computing. – Vol. 16, No. 2. – P. 133-140.

9. Black J., Halevi S., Krawczyk H., Krovetz T., Rogaway P. UMAC security bound // November 14, 2005 [Электрон. ресурс]. – Режим доступа: www.cs.ucdavis.edu/~rogavay/umac.

10. Кузнецов А.А., Чевардин В.Е. Модель оценки параметров универсальной хеширующей функции. // Радіоелектронні і комп'ютерні системи. – 2006. – № 5 (17). – С. 108-111.

Поступила в редакцию 12.02.2007

Рецензент: д-р техн. наук, проф. В.С. Харченко, Национальный аэрокосмический университет им. Н.Е. Жуковского «ХАИ», Харьков.