УДК 681.3.06 (0.43)

О.Г. Король, С.П. Евсеев, Д.С. Захаров

Харьковский национальный экономический университет, Харьков

ИССЛЕДОВАНИЕ СТАТИСТИЧЕСКИХ КОЛЛИЗИОННЫХ СВОЙСТВ MAC-КОДОВ АУТЕНТИЧНОСТИ ДАННЫХ И ОБОСНОВАНИЕ ПРЕДЛОЖЕНИЙ ПО ИХ СОВЕРШЕНСТВОВАНИЮ

Проводятся исследования статистических коллизионных свойств кодов аутентичности данных UMAC на основе методики статистического исследования коллизионных свойств MAC, на основе полученных результатов экспериментальных исследований разрабатываются предложения по совершенствованию методов каскадного формирования MAC-кодов и обеспечению высоких коллизионных свойств.

Ключевые слова: *статистические коллизионные свойства, коды аутентичности данных, каскадная схема MAC-кодов UMAC.*

Постановка проблемы

Проведенный анализ показал, что наибольшей вычислительной эффективностью обладает отобранный при проведении европейского конкурса NESSIE алгоритм UMAC (Message Authentication Code using Universal Hashing) [1 – 7], для формирования кодов контроля целостности и аутентичности в котором используются семейства универсальных хеширующих функций [8 - 11]. Число коллизий (столкновений) формируемых хеш-образов для каждого введенного ключа универсального хеширования не превышает некоторой заранее заданной величины, а криптостойкость UMAC обеспечивается на уровне выбранного криптоалгоритма (по спецификации рекомендован алгоритм шифрования AES). Однако влияние используемого криптоалгоритма на коллизионные свойства кодов подлинности сообщений UMAC на сегодняшний день не исследовано, обеспечение свойств универсального хеширования в такой многослойной конструкции не обосновано/

Целью статьи является исследования коллизионных свойств кодов контроля целостности и аутентичности данных UMAC, на основе разработанной методики статистического исследования коллизионных свойств MAC и на основе полученных результатов экспериментальных исследований разрабатываются предложения по совершенствованию методов каскадного формирования MAC и обеспечению высоких коллизионных свойств.

1. Методика статистического исследования коллизионных свойств

Проведение экспериментальных исследований коллизионных свойств кодов аутентификации сообщений UMAC проведем по соответствующим слоям преобразования:

1. На первом этапе исследуем коллизионные свойства мини-версии универсального хеширования. Для этого необходимо подтвердить в ходе эксперимента теоретические оценки числа возникающих коллизий формируемых хеш-кодов Y_{mini};

2. На втором этапе проведем экспериментальные исследования коллизионных свойств псевдослучайных подложек Pad_{mini} на основе анализа свойств уменьшенной модели шифра Baby-Rijndael. Подобные исследования в доступной литературе не описаны и, по всей видимости, проводятся нами впервые;

3. На третьем этапе проведем экспериментальные исследования коллизионных свойств формируемых с использованием mini-UMAC кодов аутентификации сообщений $Tag_{mini} = Y_{mini} \oplus Pad_{mini}$. Это наиболее важная часть проводимых исследований, поскольку она позволит ответить на вопрос о сохранении свойств универсального хеширования после применения слоя криптографического преобразования информации.

Оценку числа коллизий формируемых элементов будем проводить, ориентируясь на коллизионные свойства универсального хеширования. Собственно говоря, нам требуется подтвердить или опровергнуть гипотезу о сохранении коллизионных свойств универсального хеширования на всех этапах формирования кодов аутентификации сообщений mini-UMAC.

Идея универсального хеширования заключается в определении такого набора элементов конечного множества H хеш-функций $h: A \to B$, $|A| = a$, $|B| = b$ чтобы случайный выбор функции $h \in H$ обеспечивал бы низкую вероятность коллизии, т.е. для любых различных входов x_1 и x_2 вероятность того, что $h(x_1) = h(x_2)$ (вероятность коллизии, столкновения) не может должна превосходить некоторой заранее заданной величины ε:

$$P_{кол} = P(h(x_1) = h(x_2)) \le \varepsilon,$$

причем вероятность коллизии может быть рассчитана как

$$P_{кол} = \frac{\delta_H(x_1, x_2)}{|H|},$$

где $\delta_H(x_1, x_2)$ – количество таких хеш-функций в H, при которых значения $x_1, x_2 \in A$, $x_1 \neq x_2$ вызывают коллизию, т.е. $h(x_1) = h(x_2)$.

Приведем два определения универсального хеширования [8, 9].

1. Пусть $0 < \varepsilon < 1$. H является ε – универсальным хеш-классом (сокращенно $\varepsilon - U(H, A, B)$), если для двух различных элементов $x_1, x_2 \in A$ существует не больше, чем $|H| \cdot \varepsilon$ функций $f \in H$ таких, что $h(x_1) = h(x_2)$, если $\delta_H(x_1, x_2) \leq \varepsilon |H|$ для всех $x_1, x_2 \in A$, $x_1 \neq x_2$.

2. Пусть $0 < \varepsilon < 1$. H является ε – строго универсальным хеш-классом (сокращенно $\varepsilon - SU(H, A, B)$) если выполняются следующие условия:

– для каждого $x_1 \in A$ и для каждого $y_1 \in B$,

$$\left|\{h \in H : h(x_1) = y_1\}\right| = |H| / |B|;$$

– для каждого $x_1, x_2 \in A$, $x_1 \neq x_2$ и для каждого $y_1, y_2 \in B$,

$$\left|\{h \in H : h(x_1) = y_1, h(x_2) = y_2\}\right| \leq \varepsilon |H|.$$

Определение универсального класса хеш-функций эквивалентно определению такого алгоритма формирования кода аутентификации, при котором число различных правил формирования кода аутентификации (число ключей), при которых существует коллизия (совпадение кодов аутентификации) для двух произвольных входных последовательностей, ограничено. Число таких ключей не может превосходить значение $P_{кол} \cdot |H|$, где $P_{кол}$ – вероятность коллизии, $|H|$ – число всех правил (ключей).

Определение строго универсального класса хеш-функций эквивалентно определению такого алгоритма формирования кодов аутентификации, при котором будут выполняться следующие правила:

1. Число правил формирования кода аутентификации (число ключей), при которых для произвольной входной последовательности значение кода аутентификации не изменяется, ограничено. Число таких ключей не может превосходить значения $|H| / |B|$, где $|H|$ – число всех ключей, $|B|$ – число возможных состояний кода аутентификации;

2. Число правил формирования кода аутентификации (число ключей), при которых для двух произвольных входных последовательностей соответствующие им значения кода аутентификации не изменяются, ограничено. Число таких ключей не может превосходить значения $P_{кол} |H|$, где $P_{кол}$ – вероятность коллизии, $|H|$ – число всех ключей.

Вероятность коллизии кодов аутентификации в схеме со строго универсальным хешированием определяется как $P_{кол} \leq \varepsilon$.

В основе предлагаемой методики статистического исследования коллизионных свойств формируемых элементов $h(x)$ лежит эмпирическая оценка максимумов числа ключей (правил хеширования) при которых:

1. Для произвольных $x_1, x_2 \in A$, $x_1 \neq x_2$ выполняется равенство

$$h(x_1) = h(x_2); \quad (1)$$

2. Для произвольных $x_1 \in A$ и $y_1 \in B$ выполняется равенство

$$h(x_1) = y_1; \quad (2)$$

3. Для произвольных $x_1, x_2 \in A$, $x_1 \neq x_2$ и $y_1, y_2 \in B$ выполняются равенства

$$h(x_1) = y_1, h(x_2) = y_2. \quad (3)$$

Оценка по первому критерию соответствует проверке выполнимости условия для универсального класса хеш-функций, оценка по второму и третьему критерию – условий для строго универсального класса хеш-функций.

Введем следующие обозначения:

$$n_1(x_1, x_2) = \left|\{h \in H : h(x_1) = h(x_2)\}\right|, \; x_1, x_2 \in A, \; x_1 \neq x_2;$$

$$n_2(x_1, y_1) = \left|\{h \in H : h(x_1) = y_1\}\right|, \; x_1 \in A, \; y_1 \in B;$$

$$n_3(x_1, x_2, y_1, y_2) = \left|\{h \in H : h(x_1) = y_1, h(x_2) = y_2\}\right|, \; x_1, x_2 \in A, \; x_1 \neq x_2, \; y_1, y_2 \in B.$$

Первый показатель $n_1(x_1, x_2)$ характеризует число правил хеширования, при которых для заданных $x_1, x_2 \in A$, $x_1 \neq x_2$ выполняется равенство (1), т.е. число ключей, при которых существует коллизия (совпадение хеш-кодов) для двух входных последовательностей x_1 и x_2.

Второй показатель $n_2(x_1, y_1)$ характеризует число правил хеширования, при которых для заданных $x_1 \in A$, $y_1 \in B$ выполняется равенство (2), т.е. число ключей, при которых для входной последовательности x_1 значение хеш-кода y_1 не изменяется.

Третий показатель $n_3(x_1, x_2, y_1, y_2)$ характеризует число правил хеширования, при которых для заданных $x_1, x_2 \in A$, $x_1 \neq x_2$, $y_1, y_2 \in B$ выполняется равенство (3), т.е. число ключей, при которых для двух входных последовательностей x_1 и x_2 соответствующие им значения хеш-кодов y_1 и y_2 не изменяются.

Поскольку число ключей, при которых могут выполняться равенства (1), (2) и (3), не должно превосходить соответствующих им значений $P_{кол} \cdot |H|$,

$|H|/|B|$ и $P_{кол}|H|/|B|$ проведем оценку максимального числа таких ключей для каждого из рассматриваемого набора элементов.

Ограничимся изучением статистических характеристик максимумов этих величин, а затем сравним полученные результаты с числом $P_{кол} \cdot H$ (для первого критерия), с числом $|H|/|B|$ (для второго критерия) и числом $P_{кол} \cdot H$ (для третьего критерия).

Таким образом, в качестве статистических показателей оценки коллизионных свойств, по которым будем проводить экспериментальные исследования, предлагается использовать:

– математические ожидания $m(n_1)$, $m(n_2)$ и $m(n_3)$ максимумов числа правил хеширования, при которых выполняются равенства (1), (2) и (3), соответственно;

– дисперсии $D(n_1)$, $D(n_2)$ и $D(n_3)$, характеризующие рассеивание значений числа правил хеширования, при которых выполняются равенства (1), (2) и (3), относительно их математических ожиданий $m(n_1)$, $m(n_2)$ и $m(n_3)$, соответственно.

Оценку коллизионных свойств по приведенным критериям будем производить в среднестатистическом смысле. Другими словами, при постановке эксперимента будем использовать ограниченный набор элементов $x_1, x_2 \in A$, $x_1 \neq x_2$ и соответствующих им хеш-образов $y_1, y_2 \in B$, рассматривая соответствующие результаты как выборку из генеральной совокупности.

Естественной оценкой для математического ожидания m случайной величины X является среднее арифметическое ее наблюдаемых значений X_i (или статистическое среднее) [11]

$$\tilde{m} = \frac{1}{N}\sum_{i=1}^{N} X_i ,$$

где N – количество реализаций случайной величины X.

Оценка дисперсии случайной величины X определяется выражением

$$\tilde{D} = \frac{1}{N-1}\sum_{i=1}^{N} (X_i - \tilde{m})^2 .$$

В силу центральной предельной теоремы теории вероятностей при больших значениях количества реализаций N среднее арифметическое будет иметь распределение, близкое к нормальному закону [11] с математическим ожиданием

$$m[\tilde{m}] \approx \tilde{m}$$

и средним квадратическим отклонением

$$\sigma[\tilde{m}] \approx \frac{\sigma}{\sqrt{N}} ,$$

где σ – среднее квадратическое отклонение оцениваемого параметра.

При этом вероятность того, что оценка $\tilde{m}$ отклонится от своего математического ожидания меньше, чем на ε (доверительная вероятность), равна [11]

$$P(|\tilde{m} - m| < \varepsilon) \approx 2\Phi\left(\frac{\varepsilon}{\sigma[\tilde{m}]}\right), \qquad (4)$$

где $\Phi(x)$ – функция Лапласа, определяется выражением

$$\Phi(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-\frac{t^2}{2}} dt . \qquad (5)$$

Таким образом, при проведении экспериментальных исследований коллизионных свойств будем использовать методы статистической проверки гипотез и математической статистики.

1. Из генеральной совокупности случайной величины X сформируем выборку следующим образом:

– для среднестатистической оценки математического ожидания $m(n_1)$ и дисперсии $D(n_1)$ в качестве случайной величины выступает максимум $n_1(x_1, x_2)$ при которых выполняется равенство $h(x_1) = h(x_2)$, следовательно, выборку объема N: X_1, X_2, ..., X_N сформируем отбором N множеств, в каждом из которых содержится M пар элементов $x_1, x_2 \in A$, $x_1 \neq x_2$ и оценивается $n_1(x_1, x_2)$, т.е. общий объем формируемых пар элементов $x_1, x_2 \in A$, $x_1 \neq x_2$ составит NM;

– для среднестатистической оценки $m(n_2)$ и $D(n_2)$ в качестве случайной величины выступает максимум $n_2(x_1, y_1)$ при которых выполняется равенство $y_1 = h(x_1)$, следовательно, выборку объема N: X_1, X_2, ..., X_N сформируем отбором N множеств, в каждом из которых содержится M пар элементов $x_1 \in A$, $y_1 \in B$ и оценивается $n_2(x_1, y_1)$. Общий объем формируемых пар элементов $x_1 \in A$, $y_1 \in B$ составит NM;

– для среднестатистической оценки $m(n_3)$ и $D(n_3)$ в качестве случайной величины выступает максимум $n_3(x_1, x_2, y_1, y_2)$ при которых выполняются равенства $y_1 = h(x_1)$ и $y_2 = h(x_2)$, следовательно, выборку объема N: X_1, X_2, ..., X_N сформируем отбором N множеств, в каждом из которых содержится M четверок элементов $x_1, x_2 \in A$, $x_1 \neq x_2$, $y_1, y_2 \in B$ и оценивается $n_3(x_1, x_2, y_1, y_2)$, общий объем формируемых четверок составит NM.

2. При экспериментальных исследованиях коллизионных свойств хеширования будем оценивать среднее арифметическое $\tilde{m}(n_i)$ наблюдаемых значений максимумов n_i и дисперсию $\tilde{D}(n_i)$, $i = 1, 2, 3$.

3. Достоверность полученных среднестатистических оценок обоснуем следующим образом. Зафиксируем точность ε и рассчитаем значения функции Лапласа, которые, в соответствии с выражением (4), дадут соответствующие доверительные вероятности:

$$P(|\tilde{m}(n_i) - m(n_i)| < \varepsilon) \approx 2\Phi\left(\frac{\varepsilon}{\sigma[\tilde{m}(n_i)]}\right),$$

$$\sigma[\tilde{m}(n_i)] \approx \frac{\sqrt{\tilde{D}(n_i)}}{\sqrt{N}}.$$

При обратной постановке задачи, т.е. для фиксированной доверительной вероятности $P_д$ при объеме выборки N доверительный интервал определим следующим образом:

$$\tilde{m}(n_i) - t_\rho \cdot \sigma[\tilde{m}(n_i)] < m(n_i) < \tilde{m}(n_i) + t_\rho \cdot \sigma[\tilde{m}(n_i)] \quad (6)$$

где t_ρ – корень уравнения $2\Phi(t_\rho) = P_д$.

Таким образом, предлагаемая методика, используя уменьшенные модели отдельных слоев преобразований, на основе оценки распределения столкновений формируемых образов позволяет экспериментально исследовать коллизионные свойства кодов аутентификации сообщений.

2 Результаты экспериментальных исследований коллизионных свойств кодов контроля целостности и аутентичности данных

С использованием разработанной уменьшенной модели UMAC (mini-UMAC) и методики статистического исследования коллизионных свойств кодов аутентификации сообщений проведем экспериментальную оценку распределения числа столкновений (коллизий) формируемых образов.

Поскольку в рассмотренной выше схеме UMAC на первом слое (при формировании хеш-кода Y_{mini}) используются семейства универсальных хеширующих функций, подробно исследуемые в работах [1-7], статистические исследования проведем только на втором слое (при формировании псевдослучайной подложки Pad_{mini}) и на заключительном этапе формирования кодов аутентификации (после выполнения суммирования $Tag_{mini} = Y_{mini} \oplus Pad_{mini}$). Именно на этих этапах, по нашему предположению и нарушаются свойства универсальности формируемых кодов аутентификации.

При проведении статистических исследований коллизионных свойств формируемых значений Pad_{mini} и Tag_{mini} для каждого эксперимента 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 = 0,1 рассчитывались соответствующие доверительные вероятности

$$P(|\tilde{m}(n_i) - m(n_i)| < \varepsilon).$$

Исследования проводились над выборкой, объема N = 100, для формирования каждого элемента выборки рассчитывался максимум по множеству из M = 1000 кортежей элементов. Таким образом, общий объем формируемых наборов составил $NM = 10^5$. Полученные результаты экспериментальных исследований сведены в табл. 1.

Таблица 1

Результаты экспериментальных исследований коллизионных свойств кодов аутентификации

	mini-AES, Pad_{mini}	mini-UMAC, Tag_{mini}
$\tilde{m}(n_1)$	-	4,23
$\tilde{D}(n_1)$	-	0,18
$P_д = P(\|\tilde{m}(n_1) - m(n_1)\| < \varepsilon)$	-	0,98
$\tilde{m}(n_2)$	6,68	4,78
$\tilde{D}(n_2)$	0,42	0,42
$P_д = P(\|\tilde{m}(n_2) - m(n_2)\| < \varepsilon)$	0,88	0,88
$\tilde{m}(n_3)$	0,19	5,31
$\tilde{D}(n_3)$	0,15	0,24
$P_д = P(\|\tilde{m}(n_3) - m(n_3)\| < \varepsilon)$	0,99	0,96

Анализ приведенных в табл. 2 данных позволяет утверждать об адекватности полученных результатов и соответствии их статистическим свойствам всей генеральной совокупности данных. Для фиксированной точности ε = 0,1 получены высокие значения доверительной вероятности, что свидетельствует об обоснованности и достоверности полученных экспериментальных результатов.

Проанализируем полученные результаты статистических исследований коллизионных свойств кодов аутентификации сообщений, сравним полученные результаты среднестатистических оценок математических ожиданий $m(n_1)$, $m(n_2)$ и $m(n_3)$ числа правил хеширования, при которых выполняются равенства (1), (2) и (3), соответственно, с теоретическим оценками: числом $P_{кол} \cdot |H|$ (для первого критерия), с числом $|H|/|B|$ (для второго критерия) и числом $P_{кол} \cdot H$ (для третьего критерия).

Рассмотрим первый критерий, по которому оценивается число правил хеширования, при которых существует коллизия (совпадение кодов аутентификации) для двух произвольных входных последовательностей. В соответствии с теоретическими оценками эта величина ограничена сверху числом $P_{кол} \cdot |H|$. Конкретизируем эту (теоретическую) оценку для кодов аутентификации, сформированных с использованием mini-AES и mini-UMAC.

Мощность ключевого множества для mini-AES и mini-UMAC составляет $|H| = 2^{16}$, мощность множества формируемых кодов аутентификации также составляет $|B| = 2^{16}$. Если использовать верхнюю оценку вероятности коллизий как обратную величину мощности формируемых кодов аутентификации $P_{кол} = 2^{-16}$ получим $n_1(x_1, x_2) \le P_{кол} \cdot |H| = 1$. Для мини-версии шифра AES это условие выполняется (обосновывается биективностью шифрующего преобразования), однако коллизионные свойства mini-UMAC существенно уступают этой верхней теоретической оценке. Фактически, число коллизий выше теоретической границы более чем в четыре раза и это положение подтверждено с высокой доверительной вероятностью $P_д = P(|\tilde{m}(n_1) - m(n_1)| < 0{,}1) > 0{,}98$.

Рассмотрим второй критерий, по которому оценивается правил хеширования, при которых для произвольной входной последовательности значение кода аутентификации не изменяется. В соответствии с теоретическими оценками эта величина для кодов аутентификации, сформированных с использованием mini-AES и mini-UMAC, ограничена сверху числом. $|H|/|B| = 1$. Полученные экспериментальные результаты свидетельствуют, что коллизионные свойства кодов аутентификации, сформированных с использованием mini-AES и mini-UMAC, не удовлетворяют второму критерию, число ключей, при которых для произвольной входной последовательности значение кода аутентификации не изменяется в несколько раз превышает теоретическую оценку для универсального хеширования.

В соответствии с третьим критерием оценивается число правил хеширования, при которых для двух произвольных входных последовательностей соответствующие им значения кода аутентификации не изменяются. Теоретическая оценка этой величины для универсального хеширования ограничена сверху числом $P_{кол}|H|$, что при использовании верхней оценки вероятности коллизий $P_{кол} = 2^{-16}$ дает $n_3(x_1, x_2, y_1, y_2) \le P_{кол} \cdot |H| = 1$. Значения, в табл.1, свидетельствуют о том, что коллизионные свойства кодов аутентификации, удовлетворяют третьему критерию. В тоже время число ключей mini-UMAC, при которых для двух произвольных входных последовательностей соответствующие им значения кода аутентификации не изменяются, более чем в пять раз выше верхней теоретической оценки.

Таким образом, из полученных результатов статистических исследований коллизионных свойств кодов аутентификации сообщений, сформированных с использованием mini-AES и mini-UMAC, можно сделать следующие важные в прикладном отношении выводы:

криптографический слой формирования кодов аутентификации сообщений (mini-AES) удовлетворяет свойствам универсального хеширования, вероятность коллизии формируемых хеш-образов не превосходит наперед заданной величины (первый критерий). Это объясняется, прежде всего, тем, что шифрование неповторяющегося (уникального) для всех информационных сообщений значения приводит к формированию множества неповторяющихся (уникальных) для всех информационных сообщений псевдослучайных подложек. Другими словами, формирование псевдослучайных подложек осуществляется в результате биективного отображения множества неповторяющихся (уникальных) для всех информационных сообщений значений, в результате чего коллизии (столкновения) подложек отсутствуют по определению. Однако, данный слой преобразований не удовлетворяет свойствам строго универсального хеширования (не выполняется второй критерий) (табл. 1). Кроме того, обеспечение свойств универсального хеширования на этом слое предполагает формирование и передачу неповторяющегося для каждого сообщения значения, что требует дополнительных временных и программно-аппаратных затрат;

результат формирования кодов контроля целостности и аутентичности по схеме mini-UMAC не удовлетворяет свойствам как универсального хеширования, так и, тем более, свойствами строго универсального хеширования. Это объясняется тем, что схема с простым суммированием по модулю два (XOR) двух результатов универсального хеширования не всегда сохраняет свойства универсального хеширования.

Поясним последний вывод на примере. Пусть первый и второй слой схемы формирования кодов контроля целостности и аутентичности обладают свойствами универсального хеширования. Условно обозначим процесс такого хеширования в виде табл. 2 и 3, где столбцами обозначены информационные сообщения M_1, M_2, …, M_n, а строками – правила хеширования (h_i и g_j, соответственно), заданные (параметризированные) соответствующими секретными ключами. В ячейках таблиц содержатся результаты хеширования, т.е. искомые хеш-коды.

Таким образом, общее число правил хеширования не изменилось (по сравнению с числом правил хеширования для функций h(x) и g(x), соответственно), оно определяется мощностью множества используемых секретных ключевых данных. Каждый секретный ключ K_i задает (параметризирует) два правила h_i и g_i, которые применяются к каждому информационному сообщению, подлежащему хешированию. Результат преобразования представлен в соответствующих ячейках табл. 4 как результат суммирования по модулю 2 (XOR) значений $h_i(M_j)$ и $g_i(M_j)$.

Таблица 2

Процесс хеширования (слой 1)

	M_1	M_2	M_3	…	M_n
h_1	$h_1(M_1)$	$h_1(M_2)$	$h_1(M_3)$	…	$h_1(M_n)$
h_2	$h_2(M_1)$	$h_2(M_2)$	$h_2(M_3)$	…	$h_2(M_n)$
h_3	$h_3(M_1)$	$h_3(M_2)$	$h_3(M_3)$	…	$h_3(M_n)$
…	…	…	…	…	…
h_k	$h_k(M_1)$	$h_k(M_2)$	$h_k(M_3)$	…	$h_k(M_n)$

Таблица 3

Процесс хеширования (слой 2)

	M_1	M_2	M_3	…	M_n
g_1	$g_1(M_1)$	$g_1(M_2)$	$g_1(M_3)$	…	$g_1(M_n)$
g_2	$g_2(M_1)$	$g_2(M_2)$	$g_2(M_3)$	…	$g_2(M_n)$
g_3	$g_3(M_1)$	$g_3(M_2)$	$g_3(M_3)$	…	$g_3(M_n)$
	…	…	…	…	…
g_k	$g_k(M_1)$	$g_k(M_2)$	$g_k(M_3)$	…	$g_k(M_n)$

Таблица 4

Результат преобразования

		M_1	M_2	M_3	…	M_n
K_1	h_1, g_1	$h_1(M_1)\oplus g_1(M_1)$	$h_1(M_2)\oplus g_1(M_2)$	$h_1(M_3)\oplus g_1(M_3)$	…	$h_1(M_n)\oplus g_1(M_n)$
K_2	h_2, g_2	$h_2(M_1)\oplus g_2(M_1)$	$h_2(M_2)\oplus g_2(M_2)$	$h_2(M_3)\oplus g_2(M_3)$	…	$h_2(M_n)\oplus g_2(M_n)$
K_3	h_3, g_3	$h_3(M_1)\oplus g_3(M_1)$	$h_3(M_2)\oplus g_3(M_2)$	$h_3(M_3)\oplus g_3(M_3)$	…	$h_3(M_n)\oplus g_3(M_n)$
…	…	…	…	…	…	…
K_k	h_k, g_k	$h_k(M_1)\oplus g_k(M_1)$	$h_k(M_2)\oplus g_k(M_2)$	$h_k(M_3)\oplus g_k(M_3)$	…	$h_k(M_n)\oplus g_k(M_n)$

Коллизия хеш-кодов будет наблюдаться для всех сообщений M_i и M_j, для которых выполняется равенство:

$$h_w(M_i)\oplus g_w(M_i) = h_w(M_j)\oplus g_w(M_j). \quad (1)$$

Даже если функции h_w и g_w для сообщений M_i и M_j не вызывают коллизию, т.е., если

$$h_w(M_i) \neq h_w(M_j)$$

и

$$g_w(M_i) \neq g_w(M_j)$$

равенство (3.7) все равно может выполняться, и число правил (и число соответствующих ключей), вызывающих коллизию в результирующей схеме, возрастет. Это событие будет достоверным (произойдет наверняка), например, в случае, если

$$h_w(M_i) = g_w(M_j)$$

и

$$g_w(M_i) = h_w(M_j).$$

Таким образом, схема с простым суммированием по модулю два (XOR) двух результатов универсального хеширования в общем случае не обеспечивает сохранение свойств универсального хеширования. Коллизионные свойства кодов аутентификации сообщений снижаются и, как показывает анализ табл. 1, не удовлетворяют поставленным требованиям. Следовательно, нарушение коллизионных свойств универсального хеширования в схеме mini-UMAC (после применения криптографического слоя преобразования) следует считать экспериментально доказанным. Проанализируем возможности по совершенствованию рассмотренной схемы формирования кодов аутентификации сообщений с обеспечением свойств универсального хеширования, обоснуем предложения по обеспечению высоких коллизионных свойств в усовершенствованной схеме UMAC.

3. Обоснование предложений по совершенствованию методов каскадного ключевого хеширования

Для обоснования предложений по совершенствованию схемы UMAC рассмотрим процесс формирования кодов аутентификации сообщений с суммированием по модулю два (XOR) двух результатов универсального хеширования, причем таким образом, что обе схемы универсального хеширования функционировали под управлением собственных независимо сформированных секретных ключевых данных.

Обозначим, как и прежде, процесс универсального хеширования в виде таблиц 2 и 3, где каждая пара правил хеширования h_i и g_i задается соответствующими независимо сформированными секретными ключами $K^h_i \in \{K^h_1, K^h_2, \ldots, K^h_{kh}\}$ и $K^g_i \in \{K^g_1, K^g_2, \ldots, K^g_{kg}\}$.

В этом случае результирующая схема с простым суммированием по модулю два (XOR) двух результатов универсального хеширования может быть представлена табл. 5. Очевидно, что число правил хеширования такой схемы будет определяться мощностью множества используемых ключевых данных, т.е. произведением $kh \cdot kg$ соответствующих мощностей множеств $\{K^h_1, K^h_2, \ldots, K^h_{kh}\}$ и $\{K^g_1, K^g_2, \ldots, K^g_{kg}\}$, соответственно.

Каждая строка табл.5 задается парой секретных ключей $\{K^h_i, K^g_i\}$ и соответствует выполнению двух соответствующих случайно и независимо выбранных правил h_i и g_i. Результат такого преобразования представлен в соответствующих ячейках табл. 5.

Рассмотрим случайное событие, состоящее в возникновении коллизии (столкновения) хеш-образов для двух произвольно выбранных информационных сообщений M_i и M_j:

$$h_u(M_i) \oplus g_w(M_i) = h_u(M_j) \oplus g_w(M_j). \quad (2)$$

Если каждая из функций h_u и g_w, параметризированная некоторым случайным ключом K^h_u и K^g_w, соответственно, реализует универсальное хеширование, тогда, по аналогии с приведенными выше рассуждениями для случая, описанного в табл. 5, число коллизий (число-ключей K^h_u и K^g_w, приводящих к равенству (2)) будет таким же, как и число ключей K_w в табл. 4, приводящих к равенству (1).

Однако общее число ключей в схеме, описываемой табл. 5, значительно больше (в квадратичной зависимости), по сравнению со схемой, описываемой табл. 4.

Следовательно, коллизионные свойства такой схемы, как отношение числа коллизий к мощности ключевого пространства очевидно улучшаться.

Однако общее число ключей в схеме, описываемой табл. 5, значительно больше (в квадратичной зависимости), по сравнению со схемой, описываемой табл. 4. Следовательно, коллизионные свойства такой схемы, как отношение числа коллизий к мощности ключевого пространства очевидно улучшаться.

Таблица 5

Описание схемы

		M_1	M_2	M_3	…	M_n
K^h_1,K^g_1	h_1, g_1	$h_1(M_1)\oplus g_1(M_1)$	$h_1(M_2)\oplus g_1(M_2)$	$h_1(M_3)\oplus g_1(M_3)$	…	$h_1(M_n)\oplus g_1(M_n)$
K^h_1,K^g_2	h_1, g_2	$h_1(M_1)\oplus g_2(M_1)$	$h_1(M_2)\oplus g_2(M_2)$	$h_1(M_3)\oplus g_2(M_3)$	…	$h_1(M_n)\oplus g_2(M_n)$
K^h_1,K^g_3	h_1, g_3	$h_1(M_1)\oplus g_3(M_1)$	$h_1(M_2)\oplus g_3(M_1)$	$h_1(M_3)\oplus g_3(M_3)$	…	$h_1(M_n)\oplus g_3(M_n)$
…	…	…	…	…	…	…
K^h_1,K^g_{kg}	h_1, g_{kg}	$h_1(M_1)\oplus g_{kg}(M_1)$	$h_1(M_2)\oplus g_{kg}(M_2)$	$h_1(M_3)\oplus g_{kg}(M_3)$	…	$h_1(M_n)\oplus g_{kg}(M_n)$
K^h_2,K^g_1	h_2, g_1	$h_2(M_1)\oplus g_1(M_1)$	$h_2(M_2)\oplus g_1(M_2)$	$h_2(M_3)\oplus g_1(M_3)$	…	$h_2(M_n)\oplus g_1(M_n)$
K^h_2,K^g_2	h_2, g_2	$h_2(M_1)\oplus g_2(M_1)$	$h_2(M_2)\oplus g_2(M_2)$	$h_2(M_3)\oplus g_2(M_3)$	…	$h_2(M_n)\oplus g_2(M_n)$
K^h_2,K^g_3	h_2, g_3	$h_2(M_1)\oplus g_3(M_1)$	$h_2(M_2)\oplus g_3(M_2)$	$h_2(M_3)\oplus g_3(M_3)$	…	$h_2(M_n)\oplus g_3(M_n)$
…	…	…	…	…	…	…
K^h_2,K^g_{kg}	h_2, g_{kg}	$h_2(M_1)\oplus g_{kg}(M_1)$	$h_2(M_2)\oplus g_{kg}(M_2)$	$h_2(M_3)\oplus g_{kg}(M_3)$	…	$h_2(M_n)\oplus g_{kg}(M_n)$
…	…	…	…	…	…	…
K^h_{kh},K^g_1	h_{kh}, g_1	$h_{kh}(M_1)\oplus g_1(M_1)$	$h_{kh}(M_2)\oplus g_1(M_2)$	$h_{kh}(M_3)\oplus g_1(M_3)$	…	$h_{kh}(M_n)\oplus g_1(M_n)$
K^h_{kh},K^g_2	h_{kh}, g_2	$h_{kh}(M_1)\oplus g_2(M_1)$	$h_{kh}(M_2)\oplus g_2(M_2)$	$h_{kh}(M_3)\oplus g_2(M_3)$	…	$h_{kh}(M_n)\oplus g_2(M_n)$
K^h_{kh},K^g_3	h_{kh}, g_3	$h_{kh}(M_1)\oplus g_3(M_1)$	$h_{kh}(M_2)\oplus g_3(M_2)$	$h_{kh}(M_3)\oplus g_3(M_3)$	…	$h_{kh}(M_n)\oplus g_3(M_n)$
…	…	…	…	…	…	…
K^h_{kh},K^g_{kg}	h_{kh}, g_{kg}	$h_{kh}(M_1)\oplus g_{kg}(M_1)$	$h_{kh}(M_2)\oplus g_{kg}(M_2)$	$h_{kh}(M_3)\oplus g_{kg}(M_3)$	…	$h_{kh}(M_n)\oplus g_{kg}(M_n)$

Для подтверждения обоснованности выдвинутого предложения проведем исследования коллизионных свойств мини-версии схемы аутентификации по рассмотренной выше методике. При проведении статистических исследований коллизионных свойств формируемые значения на первом и втором слое будут параметризированы случайно и независимо сформированными ключами. В результате исследований коллизионных свойств формируемых аутентификаторов $Tag_{min\,i}$ будем оценивать 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 $\varepsilon = 0{,}1$ рассчитывать соответствующие доверительные вероятности $P(|\tilde{m}(n_i) - m(n_i)| < \varepsilon)$, $i = 1, 2, 3$. Исследования проведем над выборкой, объема $N = 100$, для формирования каждого элемента выборки рассчитаем максимум по множеству из $M = 1000$ кортежей элементов, общий объем формируемых наборов составил $NM = 10^5$. Полученные результаты экспериментальных исследований сведены в табл. 6.

Во второй колонке табл. 6 приведены результаты экспериментальных исследований коллизионных свойств мини-версии алгоритма UMAC (соответствуют данным последней колонке табл.1). Третья колонка табл. 6 соответствует данным, полученным в результате модификации каскадного ключевого хеширования для обеспечения высоких коллизионных свойств, т.е. случаю использования случайно и независимо формируемых подключей на первом и втором слое схемы аутентификации.

Приведенные в табл. 6 данные позволяют утверждать, что применение схемы UMAC с независимым формированием подключей существенно улучшает коллизионные свойства кодов контроля целостности и аутентичности данных.

Таблица 6

Результаты экспериментальных исследований коллизионных свойств кодов аутентификации mini-UMAC с использованием связанных и независимых подключей

	связанные подключи	независимые подключи
$\tilde{m}(n_1)$	4,23	2,952
$\tilde{D}(n_1)$	0,18	0,32
$P_д = P(\lvert\tilde{m}(n_1) - m(n_1)\rvert < \varepsilon)$	0,98	0,0045
$\tilde{m}(n_2)$	4,78	3,98
$\tilde{D}(n_2)$	0,42	0,41
$P_д = P(\lvert\tilde{m}(n_2) - m(n_2)\rvert < \varepsilon)$	0,88	0,004
$\tilde{m}(n_3)$	5,31	3,52
$\tilde{D}(n_3)$	0,24	0,41
$P_д = P(\lvert\tilde{m}(n_3) - m(n_3)\rvert < \varepsilon)$	0,96	0,0039

Выводы

Таким образом, на основании приведенных выше рассуждений сформируем следующие предложения по обеспечению свойств универсального хеширования в схеме формирования кодов аутентификации сообщений UMAC: используемые секретные ключевые данные $K^h_i \in \{K^h_1, K^h_2, \ldots, K^h_{kh}\}$ и $K^g_i \in \{K^g_1, K^g_2, \ldots, K^g_{kg}\}$ и соответствующие им правила хеширования h_i и g_i на отдельных слоях преобразований должны формироваться случайно и независимо друг от друга. В этом случае отношение числа правил аутентификации, вызывающих коллизию к общему их числу не превысит наперед заданной величины и коллизионные свойства не снизятся даже после применения криптографического преобразования на последнем этапе формирования аутентификаторов.

Перспективным направлением исследований является разработка и теоретическое обоснование новых схем ключевого хеширования, использующих многослойную конструкцию и позволяющих обеспечить как высокие коллизионные свойства (с сохранением свойств универсального хеширования), так и высокие показатели безопасности, что позволит строить эффективные механизмы обеспечения целостности и аутентичности информации в телекоммуникационных системах и сетях.

Список литературы

1. Black J. "UMAC: Fast and provably secure message authentication", Advances in Cryptology / J. Black, S. Halevi H., Krawczyk, T. Krovetz, P. Rogaway. – CRYPTO '99, LNCS vol. 1666, PP. 216-233, Springer-Verlag, 1999.

2. Krovetz T. Fast universal hashing with small keys and no preprocessing", work in progress, 2000. / T. Krovetz, P. Rogaway. – URL: http://www.cs.ucdavis.edu/~rogaway/umac

3. Krovetz T. UMAC -Message authentication code using universal hashing. IETF Internet Draft / T. Krovetz, J. Black, S. Halevi, A. Hevia, H. Krawczyk, P. Rogaway. – draft-krovetz-umac-00.txt. – URL: www.cs.ucdavis.edu/~rogaway/umac, 2000.

4. Krovetz T. UMAC -Message authentication code using universal hashing. IETF Internet Draft, draft-krovetz-umac-02.txt. – URL: www.cs.ucdavis.edu/~rogaway/umac, 2004.

5. 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.

6. Krovetz T. UMAC - Message authentication code using universal hashing, 2006. – URL: http://www.cs.ucdavis.edu.

7. Krovetz T. Software-Optimized Universal Hashing and Message Authentication. Dissertation submitted in partial satisfaction of the requirements for the degree of doctor of philosophy. University Of California Davis. September 2000. – 269 p.

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

9. Wegman M. N. New hash functions and their use in authentication and set equality / M. N. Wegman, J. L. Carter / Computer and System Scince – 1981 – № 22 – P. 265-279.

10. Кузнецов А.А. Исследование коллизионных свойств кодов аутентификации сообщений UMAC / А.А. Кузнецов, О.Г. Король, С.П. Евсеев. Прикладная радиоэлектроника. – Харьков:Изд-во ХНУРЭ, 2012. Том 11 N 2. – С.171-183

11. Кузнецов А.А. Методика исследования коллизионных свойств кодов аутентификации сообщений/ А.А Кузнецов., О.Г. Король, В.В. Босько, С.П. Евсеев Військово-технічний збірник / Академія сухопутних військ, Львів: АСВ. 2011, Вип.2(5) 2011. – С.23 –30.

Поступила в редколлегию 21.09.2012

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

ДОСЛІДЖЕННЯ СТАТИСТИЧНИХ КОЛІЗІЙНИХ ВЛАСТИВОСТЕЙ MAC-КОДІВ АВТЕНТИЧНОСТІ ДАНИХ І ОБҐРУНТУВАННЯ ПРОПОЗИЦІЙ З ЇХ ВДОСКОНАЛЕННЯ

О.Г. Король, С.П. Євсеєв, Д.С. Захаров

Проводяться дослідження статистичних колізійних властивостей кодів автентичності даних UMAC на основі методики статистичного дослідження колізійних властивостей MAC, на основі отриманих результатів експериментальних досліджень розробляються пропозиції з вдосконалення методів каскадного формування MAC-кодів і забезпеченню високих колізійних властивостей.

***Ключові слова**: статистичні колізійні властивості, коди автентичності даних, каскадна схема MAC-кодів UMAC*

STUDY STATISTICAL COLLISIONS CHARACTERISTIC MAC-CODES TO AUTHENTICITY DATA AND MOTIVATION OF THE OFFERS UPON THEIR IMPROVEMENT

O.G. Korol, S.P. Evseev, D.S. Zakharov

They are conducted studies statistical collisions characteristic of the codes to authenticity given UMAC on base of the methods of the statistical study collisions characteristic MAC, on base got result of the experimental studies are developed offers on improvement of the methods of the cascade shaping MAC-codes and provision high collisions characteristic.

***Keywords**: statistical conflict characteristic, codes to authenticity data, cascade scheme MAC-codes UMAC.*