Захист інформації

УДК 004.021+681.3.05

А.В. Антонов, В.Б. Бзот

Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков

ВАРИАНТ ЭФФЕКТИВНОЙ РЕАЛИЗАЦИИ МЕТОДА ПОСТРОЕНИЯ ХЕШ-ФУНКЦИИ НА ОСНОВЕ ХАОТИЧЕСКИХ ОТОБРАЖЕНИЙ С ПЕРЕМЕННЫМИ ПАРАМЕТРАМИ И ПАРАЛЛЕЛЬНОЙ ОРГАНИЗАЦИЕЙ ВЫЧИСЛЕНИЙ

Предложен вариант хеш-функции на основе усовершенствованного метода построения хеш-функций с применением хаотических отображений с переменными параметрами и параллельной организацией вычислений. Подтверждены ожидаемые характеристики статистической безопасности и вычислительной эффективности. Обобщены результаты исследований соответствующего метода, приведены рекомендации по его практическому применению.

Ключевые слова: *хеш-функция, хаотическое отображение, распараллеливание, безопасность.*

Введение

Задача конструирования надежных и безопасных хеш-функций лежит на стыке нескольких областей наук в сфере информационных технологий и систем. Так, одним из аспектов функциональной безопасности и живучести информационных и информационно-управляющих систем является обеспечение достоверности информации, недопущение как предумышленного, так и непреднамеренного ее разрушения или искажения. В системах защиты информации высокую значимость имеют задачи обеспечения целостности, подлинности и неотказуемости информации. Эти задачи в современных информационных системах, как правило, решаются комплексно, в том числе и с помощью хеш-функций.

В последнее время в качестве одного из альтернативных подходов к конструированию сжимающих функций алгоритмов хеширования все чаще предлагается использовать достижения теории динамического хаоса. Однако широкому распространению таких функций (алгоритмов) препятствует низкая вычислительная эффективность расчета траекторий хаотических отображений. Поэтому особый интерес представляет предложенный в публикации [1] метод построения хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организации вычислений (ПХФ ХОПП и ПОВ). Внимание к этому методу обусловлено его способностью в значительной мере компенсировать низкую вычислительную эффективность путем распараллеливания вычислений, что в целом соответствует современным тенденциям в развитии как вычислительных систем, так и структуры построения алгоритмов хеширования.

В тоже время детальное исследование, проведенное в работах [2] и [3], показало уязвимость метода к поиску и обнаружению близких коллизий, а так же коллизий первого и второго рода. В работе [4] был предложен вариант дальнейшего развития и усовершенствования метода, позволяющий устранить выявленные в работах [2, 3] недостатки, а также повысить эффективность вычислений. В работе [5] бы выполнен анализ (экспресс-оценка) статистической безопасности усовершенствованного метода с помощью пакета статистических тестов NIST STS [6].

Тестирование в целом показало правильность предпринятых мер по усовершенствованию метода ПХФ ХОПП и ПОВ, предложенных в работе [4]. В то же время тесты показали наличие отклонений от случайности практически для всех выборок, сформированных хеш-функцией на основе усовершенствованного метода ПХФ ХОПП и ПОВ при различных ее параметрах. Было указано, что вариант реализации метода, использовавшийся при тестировании, является недостаточно надежным и безопасным, а для каждой конкретной хеш-функции, построенной на основе усовершенствованного метода, требуются отдельные исследования по выбору оптимальных (рациональных) инициализирующих значений (для заданных базовых преобразований) и наиболее эффективного механизма (способа) их коррекции на каждой итерации блочной функции сжатия. В целом выбор оптимальных параметров метода при построении хеш-функции должен быть направлен на получение уникальных траекторий отображений максимальной длины.

Обобщая цикл работ [1 – 5] с целью оценки практической ценности усовершенствованного метода ПХФ ХОПП и ПОВ, ниже будет предложен

вариант хеш-функции на его основе. Будет дана оценка статистической безопасности, вычислительной эффективности и практической ценности предложенной реализации метода (варианта хеш-функции) и метода в целом.

Вариант (версия) хеш-функции на основе усовершенствованного метода ПХФ ХОПП и ПОВ

Опираясь на результаты исследований и рекомендации, приведенные в работе [5], предлагается использовать при построении хеш-функции на основе усовершенствованного метода ПХФ ХОПП и ПОВ следующие значения его параметров (вариант).

В качестве базовых преобразований используются представители класса кусочно-линейных отображений: палаточное и кусочно-линейное отображение второго порядка, которое в [1] называют просто кусочно-линейным. Рекуррентное уравнение для палаточного отображения задается в виде:

$$x_{i+1} = \begin{cases} \dfrac{x_i}{a}, & 0 \le x_i < a, \\ \dfrac{1-x_i}{1-a}, & a \le x_i \le 1, \end{cases} \qquad (1)$$

где $x_i \in (0,1)$ – точки траектории; $a \in (0,1)$ – управляющий параметр отображения (часто называемый ключевым). Рекуррентное уравнение для кусочно-линейного отображения второго порядка задается следующим выражением:

$$x_{i+1} = \begin{cases} x_i/\beta & 0 \le x_i < \beta, \\ (x_i-\beta)/(0.5-\beta) & \beta \le x_i < 0.5, \\ (1-x_i-\beta)/(0.5-\beta) & 0.5 \le x_i < 1-\beta, \\ (1-x_i)/\beta & 1-\beta \le x_i < 1, \end{cases} \qquad (2)$$

где $x_i \in (0,1)$ – точки траектории; $\beta \in (0, 0.5)$ – управляющий параметр отображения (ключевой).

В качестве инициализирующих значений на первой итерации блочной функции сжатия предлагается использовать пары значений $x'_0 = 0.11 = 0x1(C28F5)$ и $a = 0.3 = 0x4(C)$ для отображения (1), $x''_0 = 0.13 = 0x2(147AE)$ и $\beta = 0.15 = 0x2(6)$ для отображения (2).

На следующих итерациях блочной функции сжатия предлагается отбирать последние элементы векторов $\vec{x}'$ и $\vec{x}''$ (см. [4]), сформированных на предыдущей итерации, и исходные значения управляющих параметров a и β. Далее для усиления хеш-функции предлагается корректировать отобранные инициализирующие значения, комбинируя их, со значением промежуточного хеша H. Опытным путем было установлено, что коррекция управляющих параметров является нецелесообразной, а в отобранные начальные значения x'_0 и x''_0 следует вносить лишь малые возмущения (не более половины младших значащих бит в бинарном представлении). Для этого промежуточный хеш H разбивается на блоки кратные половине точности вычислений p, которые объединяются между собой и младшими битами x'_0, x''_0 бинарной операцией «исключающее или».

В соответствии с рекомендациями, приведенными в работе [5], значение точности вычислений должно составлять $p \ge 16$ бит. При меньших значениях (8 бит) приемлемые характеристики статистической безопасности практически недостижимы.

Также было обнаружено, что при некоторых вариантах исходного текста (длинные последовательности нулей) параллельная функция сжатия (см. [4]) имеет достаточно слабый лавинный эффект, поскольку длины (размер) рассчитываемых траекторий слишком малы. Поэтому для определения длин рассчитываемых траекторий хаотических отображений на каждой итерации параллельной функции сжатия предлагается использовать следующие выражения. Так, порядковые номера отбираемых точек траектории отображений для функции $f_{atm}(\ )$ вычисляются как $z_j = \lfloor (j/L) \cdot (m_{i,j} \wedge 0x0F + 4) \rfloor + 1$ при прямом проходе, а при обратном – как $z_j = \lfloor (j/L) \cdot (4 + (m_{i,j} \wedge 0xF0)/0x0F) \rfloor + 1$. Для функции $f_{plm}(\ )$ при прямом проходе номер точки определяется как $z_j = \lfloor (1 - j/L) \cdot (m_{i,j} \wedge 0x0F + 4) \rfloor + 1$, а при обратном $z_j = \lfloor (1 - j/L) \cdot (4 + (m_{i,j} \wedge 0xF0)/0x0F) \rfloor + 1$. В данных выражениях $j = 1, 2, \ldots, L$ – номер итерации параллельной функции сжатия, L – размер хеша в битах, $i = 1, 2, \ldots, d$, d – глубина распараллеливания, $m_{i,j}$ – блоки (байты) обрабатываемого сообщения.

Кроме того, способ вычисления управляющих параметров отображений на каждой итерации параллельной функции сжатия следует также отнести к параметрам метода, обусловленным конкретной его реализацией, т.к. он тесно связан с выбранными базовыми преобразованиями.

Установлено, что при использовании на некоторых итерациях значений управляющих параметров близких к границам области определения ($a \in (0,1)$ для отображения (1) и $\beta \in (0, 0.5)$ для отображения (2)) проявления лавинного эффекта отображений крайне незначительны. Поэтому пред-

лагается для вычисления управляющих параметров отображений на каждой итерации параллельной функции использовать выражения $a_{i,j} = (2*i/d + 2*j/L)/5$, $\beta_{i,j} = a_{i,j}/2$.

Результаты тестирования предложенного варианта хеш-функции пакетом статистических тестов NIST STS

Полное описание пакета, тестов, рекомендации по выбору методики исследований и интерпретации результатов приведены в публикации [6]. В работе [5] описана методика исследований применительно к рассматриваемому методу и его реализациям.

На основе предложенных выше рекомендаций по выбору параметров и конструированию хеш-функции на основе усовершенствованного метода была разработана программная ее реализация. Для исследования статистической безопасности хеш-функции было сформировано при различных ее параметрах 12 выборок по сто последовательностей в каждой (размером более одного миллиона бит в каждой последовательности). В целом методика и порядок исследований практически полностью идентичны предложенным в работе [5]. Не исследовались, как не представляющие практического интереса, только варианты формирования последовательностей при точности вычислений в восемь бит и при использовании в качестве исходного текста псевдослучайных последовательностей. Это обусловлено тем, что при восьмибитной точности вычислений добиться приемлемых характеристик хеш-функции не представляется возможным, а при использовании псевдослучайного исходного текста, сформированного линейным конгруэнтным генератором, хорошие результаты тестирования обусловливаются в первую очередь свойствами самого генератора. Обобщенные результаты тестирования представлены в табл. 1.

Таблица 1

Результаты тестирования варианта хеш-функции на основе усовершенствованного метода

Вариант последовательности			Результат тестирования	
Длина хеша (бит)	Точность вычислений (бит)	Глубина распараллеливания	Количество и % успешно пройденных тестов	Средний % «хороших» последовательностей
128	16	4	131 (69,7%)	91,8%
		2	186 (98,9%)	98,9%
		1	184 (97,9%)	97,7%
	24	4	187 (99,5%)	99,0%
		2	188 (100%)	98,8%
		1	188 (100%)	98,9%
256	16	3	140 (74,5%)	93,6%
		2	181 (96,3%)	97,4%
		1	182 (96,8%)	97,0%
	24	3	188 (100%)	99,1%
		2	187 (99,5%)	99,2%
		1	188 (100%)	99,1%

Напомним, что принятие решение о «качестве» генератора последовательностей осуществляется с применением двух основных подходов [5, 6]: первый основан на оценке равномерности распределения значений вероятности P (сила доказательств против нулевой гипотезы) для последовательностей из выборки, и второй – на оценке процента прошедших тест последовательностей из выборки. Хотя в табл. 1 представлены результаты тестирования только на основании второго похода, они отражают общую картину тестирования и по первому.

В целом результаты тестирования, приведенные в табл. 1, указывают на практически полное отсутствие отклонения характера последовательностей, сгенерированных предложенной версией хеш-функции, от случайных. Исключение составляют варианты генерирования последовательностей при 16-ти битной точности вычислений и с глубиной распараллеливания, при которой осуществляется только одна итерация блочной функции сжатия ($d = 4$ для 128 битного хеша и $d = 3$ для 256 битного при длине исходного текста в 256 байт). Однако как было указано в работе [5], такой выбор параметров хеш-функций на основе усовершенствованного метода не рекомендуется использовать на практике и тестировался он только для подтверждения данной рекомендации. Для остальных параметров хеш-функции было показано либо полное прохождение всех тестов, либо незначительные отклонения, соизмеримые с результатами тестирования хорошо изученных «эталонных» генераторов, в том числе и на основе хеш-функций, например SHA-1 [5, 7]. При этом даже для не пройденных тестов отклонения выборок от случайности имеют в целом незначительный характер, и процент прохождения (94 – 95%) соответствующих тестов выборками очень близок к пороговому значению (96%).

В целом результаты тестирования полностью соответствуют ожидаемым, подтверждают правильность сделанных в работах [4, 5] выводов и реко-

мендаций, свидетельствуют о практической возможности конструирования надежных и статистически безопасных хеш-функций в рамках усовершенствованного метода ПХФ ХОПП и ПОВ.

Эффективность вычислений в предложенном варианте хеш-функции пакетом статистических тестов NIST STS

Как было указано в работе [4], меры по усовершенствованию метода ПХФ ХОПП и ПОВ позволяют не только повысить надежность и безопасность хеш-функций на его основе, но и повысить их эффективность вычислений. Выигрыш в производительности (числе элементарных вычислительных операций или процессорном времени, затрачиваемом для обработки данных) по сравнению с исходной версией в работе [4] оценивался приблизительно в восьмикратном размере.

При подготовке выборок для тестирования пакетом NIST STS была осуществлена практическая оценка прироста производительности в усовершенствованном методе ПХФ ХОПП и ПОВ. При этом все вычисления производились на одной и той же программно-аппаратной платформе, программная реализация (в частности алгоритмы вычисления траекторий отображений) исходной и усовершенствованной версий хеш-функции максимально унифицированы. Кроме того, для максимально точной оценки производительности физическое распараллеливание не производилось (только логическое в рамках структуры вычислений) и все вычисления организовывались на одном специально выделенном ядре мультиядерной аппаратной платформы.

По результатам оценивания при различных параметрах хеш-функций прирост производительности усовершенствованной ее версии составил 3.3…5.5 раз. Наименьший прирост производительности зафиксирован при вычислении хешей исходного текста состоящего преимущественно из длинных последовательностей нулей. Наибольший прирост – при псевдослучайном исходном тексте. Такое расхождение объясняется спецификой организации вычислений в параллельной функции сжатия, число вычислительных операций в которой напрямую зависит от характера обрабатываемых данных. Кроме того, некоторое уменьшение производительности по сравнению с предварительной оценкой в работе [4] вызвано увеличением длительности рассчитываемых траекторий на каждой итерации параллельной функции сжатия (см. выше), направленным на усиление лавинного эффекта, а также накладными расходами вычислительных ресурсов на инициализацию/финализацию блочной и параллельной функций сжатия. Тем не менее, с учетом возможности распараллеливания вычислений можно утверждать о достаточно высоком соотношении уровня надежности и безопасности предложенной версии хеш-функции на основе усовершенствованного метода ПХФ ХОПП и ПОВ к затратам вычислительных ресурсов.

Выводы

Таким образом, обобщая цикл работ [1 – 5] и результаты, полученные в данной работе, применительно к методу ПХФ ХОПП и ПОВ можно сделать следующие общие выводы:

– в работе [1] был предложен оригинальный и перспективный метод ПХФ ХОПП и ПОВ. Однако, в работах [2, 3] было показано, что его реализации не являются стойкими к близким коллизиям, коллизиям первого и второго рода. При этом было указано, что, несмотря на выявленные недостатки, сам метод представляет интерес для дальнейшего изучения и развития;

– в работе [4] был предложен вариант усовершенствования метода и устранения выявленных недостатков, повышения вычислительной эффективности. В частности, можно утверждать, что усовершенствованный метод является стойким к выявленным в работах [2, 3] уязвимостям;

– «экспресс-оценка» статической безопасности усовершенствованного метода ПХФ ХОПП и ПОВ, выполненная в публикации [5], в целом показала правильность предпринятых по его усовершенствованию мер. Однако было отмечено, что для каждой конкретной реализации метода (построенной на его основе хеш-функции) требуются отдельные исследования по выбору допустимых и оптимальных (рациональных) параметров (базовых преобразований, инициализирующих значений, механизмов (способа) их коррекции и т.д.);

– в данной работе был предложен вариант практической реализации усовершенствованного метода ПХФ ХОПП и ПОВ. Подтверждены высокие ожидаемые характеристики статистической безопасности и вычислительной эффективности соответствующей хеш-функции;

– основной областью применения предложенного варианта хеш-функции и метода ее построения, является решение задач проверки и подтверждения целостности и достоверности данных в сфере обеспечения функциональной безопасности и живучести информационных и информационно-управляющих систем, а также (в совокупности с методами криптографической защиты информации) – обеспечение и проверка их подлинности в сфере информационной безопасности;

– в то же время следует отметить ряд ограничений на параметры и сферы применения предложенного варианта хеш-функции, обусловленные спецификой метода, на основе которого она скон-

струирована. Так, глубина распараллеливания не должна превышать число блоков (векторов) данных, на которые разбивается исходное сообщение (с учетом дополнения). Также есть предостережения по выбору в качестве значения глубины распараллеливания $d = 1$, поскольку в данном случае наблюдается недостаточно сильное связывание векторов данных между собой. Эти правила несколько ограничивают область применения, как варианта хеш-функции, так и метода их построения, поскольку требуют некоторую минимальную длину обрабатываемого сообщения (длина сообщения должна быть не менее $L \cdot d$ байт, где L – длина хеша в битах), хотя дополнением сообщения можно устранить эти ограничения. Кроме того, выбор точности организации вычислений должен осуществляться исходя из требований к безопасности хеш-функции, обусловленных сферой ее применения. Так, для задач проверки/подтверждения целостности и достоверности данных в сфере обеспечения функциональной безопасности и живучести информационных систем приемлемым значением точности вычислений является 16 бит. В сфере информационной безопасности рекомендуется при организации вычислений использовать не менее 24 бит;

– следует отметить, что большая гибкость и вариативность параметров метода (например, возможность выбора различных хаотических отображений в его реализациях, возможность задания требуемой длины хеша и т.д.) закладывают необходимый потенциал для дальнейшей его модернизации и развития.

Список литературы

1. Yantao Li. Parallel Hash function construction based on chaotic maps with changeable parameters [Text] / Yantao Li, Di Xiao, Shaojiang Deng, Qi Han, Gang Zhou // Neural Computing and Applications. – 2011. – Vol. 20, №8. – P. 1305-1312.

2. Антонов А.В. Анализ уязвимостей структуры хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организации вычислений [Текст] / А.В. Антонов // Системи управління, навігації та зв'язку. – К.: ДП «ЦНДІ НІУ», 2012. – Вип. 2(22). – С. 157-162.

3. Антонов А.В. Анализ уязвимостей реализации в цифровых вычислительных системах хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организации вычислений [Текст] / А.В. Антонов // Збірник наукових праць Харківського університету Повітряних Сил. – Х.: ХУ ПС, 2012. – Вип. 4(33). – С. 123-129.

4. Антонов А.В. Развитие метода построения хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организацией вычислений [Текст] / А.В. Антонов // Системи озброєння та військової техніки. – 2012. – №2(30). – С. 111-117.

5. Антонов А.В. Анализ статистической безопасности метода построения хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организацией вычислений (с помощью пакета тестирования NIST STS) [Текст] / А.В. Антонов, В.В. Слободянюк // Системи обробки інформації. – Х.: ХУ ПС, 2012. – Вип. 7(105). – С. 27-33.

6. A Statistical Test Suite for Random and Pseudorandom Number Generators for Cryptographic Applications. NIST Special Publication 800-22, Rev. 1a [Text] / Technology Administration, U.S. Department of Commerce. – Washington: National Institute of Standards and Technology. – 2010. – P. 131.

7. Кузнецов А.А. Исследование статистической безопасности генераторов псевдослучайных чисел / А.А. Кузнецов, Р.В. Королев, Ю.Н. Рябуха // Системи обробки інформації. – Х.: ХУ ПС, 2008. – Вип. 3(70). – С. 79-82.

Поступила в редколлегию 17.12.2012

Рецензент: д-р техн. наук, проф. П.Ю. Костенко, Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков.

ВАРІАНТ ЕФЕКТИВНОЇ РЕАЛІЗАЦІЇ МЕТОДУ ПОБУДОВИ ҐЕШ-ФУНКЦІЇ НА ОСНОВІ ХАОТИЧНИХ ВІДОБРАЖЕНЬ ЗІ ЗМІННИМИ ПАРАМЕТРАМИ Й ПАРАЛЕЛЬНОЮ ОРГАНІЗАЦІЄЮ ОБЧИСЛЕНЬ

А.В. Антонов, В.Б. Бзот

Запропоновано варіант ґеш-функції на основі вдосконаленого методу побудови ґеш-функцій із застосуванням хаотичних відображень зі змінними параметрами й паралельною організацією обчислень. Підтверджено очікувані характеристики статистичної безпеки й обчислювальної ефективності. Узагальнено результати досліджень відповідного методу, наведені рекомендації з його практичного застосування.

***Ключові слова**: геш-функція, хаотичне відображення, розпаралелювання, безпека.*

VARIANT OF EFFECTIVE IMPLEMENTATION OF HASH FUNCTION CONSTRUCTING METHOD BASED ON CHAOTIC MAPS WITH CHANGEABLE PARAMETERS AND PARALLEL CALCULATIONS

A.V. Antonov, V.B. Bzot

Variant of effective implementation of hash function constructing improved method based on chaotic maps with changeable parameters and parallel calculations were proposed. Expected performance of its statistical security and computational efficiency were confirmed. Generalized results of research on the appropriate method, provides guidelines for its practical application.

***Keywords**: hash function, the chaotic map, paralleling, security.*