Захист інформації

УДК 004.021+681.3.05

А.В. Антонов, И.Е. Кужель, Н.В. Шигимага

Харьковский университет Воздушных Сил им. И. Кожедуба, Харьков

ОБЩИЕ ПРИНЦИПЫ ПРИМЕНЕНИЯ ДИСКРЕТНЫХ АППРОКСИМАЦИЙ ХАОТИЧЕСКИХ ОТОБРАЖЕНИЙ ДЛЯ КОНСТРУИРОВАНИЯ ХЕШ-ФУНКЦИЙ

Анализируются общие подходы и тенденции построения хеш-функций на основе хаотических отображений. Указано, что зачастую реальные свойства таких хеш-функций не соответствуют ожидаемым. Выявлены причины возникновения уязвимостей и изъянов в хеш-функциях на основе хаотических отображений. Выработаны принципы и рекомендации по использованию дискретных аппроксимаций хаотических отображений для их построения.

Ключевые слова: *хеш-функция, хаотическое отображение, эффективность, коллизия.*

Введение

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. В информационно-коммуникационных системах важной задачей является контроль целостности данных при их передаче по каналам связи. Эти задачи в современных информационных системах, как правило, решаются комплексно, в том числе и с помощью хеш-функций. Кроме этого, хеш-функции позволяют также решать задачи построения ассоциативных массивов, реализации эффективных алгоритмов поиска в больших выборках данных, построения уникальных идентификаторов для наборов данных, и т.д. Таким образом, хеш-функции являются достаточно универсальным примитивом, востребованным при решении широкого круга задач в системах обработки информации.

Требования к хеш-функциям

В зависимости от сферы применения хеш-функций требования к ним могут различаться. Тем не менее, можно выделить две основных группы достаточно универсальных свойств, которыми в той или иной мере должны обладать все хеш-функции [1]. К первой группе можно отнести свойства, описывающие устойчивость хеш-функций к поиску и возникновению коллизий, в частности:

– необратимость или стойкость к восстановлению прообраза;

– стойкость к коллизиям первого рода или восстановлению вторых прообразов;

– стойкость к коллизиям второго рода.

Данные свойства не являются независимыми. Так, обратимая функция нестойка к коллизиям первого и второго рода, а функция, нестойкая к коллизиям первого рода, нестойка и к коллизиям второго рода (обратное утверждение неверно). Помимо основных свойств к хеш-функциям часто выдвигаются дополнительные требования, такие как:

– устойчивость к близким коллизиям (near-collision);

– устойчивость к псевдоколлизиям;

– отсутствие корреляции между входом и выходом хеш-функции;

– устойчивость к нахождению частичного прообраза и другие.

Вторая группа свойств описывает эффективность практической реализации хеш-функций в современных вычислительных средствах, в частности, к хеш-функциям выдвигаются требования простоты реализации и высокой скорости вычислений на стандартных вычислительных платформах (как при программной, так и аппаратной реализации). Хотя само определение хеш-функции подразумевает необходимость вычисления ее значения за полиномиальное от длины входа время, значительный рост объемов обрабатываемых данных в современных информационных системах приводит к существенному усилению требований к вычислительной эффективности. В частности, это: ориентированность на специфику организации вычислений в современных аппаратных средствах, минимальные затраты памяти, машинного времени, ориентированность на мультипоточные (параллельные) вычисления/среды, и т.д.

Идея создания универсальной хеш-функции, удовлетворяющей всем выдвинутым критериям, является заманчивой, но на современном этапе развития информационных технологий – нереализуемой. Это вызвано определенными пробелами в теории сложности (отсутствие доказательства сущест-

© А.В. Антонов, И.Е. Кужель, Н.В. Шигимага

вования односторонних функций), а также акцентами на разных свойствах хеш-функций в зависимости от сферы их применения. Так, в информационно-коммуникационных системах при контроле целостности информации в сетях передачи данных основной упор делается на производительность алгоритмов хеширования (возможно в ущерб стойкости). В системах защиты информации внимание разработчиков фокусируется на стойкости к преднамеренному поиску коллизий (зачастую в ущерб производительности). В информационных системах предпочтительным является баланс характеристик производительности и стойкости к появлению непреднамеренных коллизий. Тем не менее, вне зависимости от сферы применения, задача создания вычислительно эффективной и стойкой к поиску/возникновению коллизий хеш-функции является актуальной.

Принципы построения хеш-функций

Общепринятым принципом построения хеш-функций стало использование итеративных (как правило, последовательных, на основе модели Меркла-Дамгарда) схем, ядром которых является сжимающая функция. В простейшем случае структура хеш-функции на основе модели Меркла-Дамгарда [1, 2] выглядит следующим образом (рис. 1).

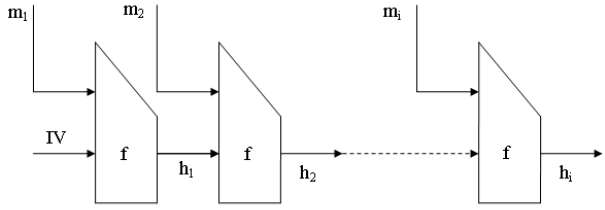

Рис. 1. Структура хеш-функции
на основе модели Меркла-Дамгарда

Здесь f – сжимающая функция; IV –инициализирующий вектор; m – информационные блоки (элементы исходного сообщения), h – выходы сжимающей функции (хеши).

Можно выделить два основных подхода к конструированию сжимающих функций [1], это:

– использование схем на основе перемешивающих и рассеивающих битовых операций над блоками входных данных. Эти конструкции могут быть как специально сконструированными, так и позаимствованными из других криптографических примитивов, например блочных шифров;

– использование доказуемо безопасных сжимающих функций на основе известных и хорошо изученных математических задач.

Первый подход является наиболее распространенным (SHA-1/2, MD-4/5/6, и др.), он ориентирован на эффективную реализацию в современных вычислительных средствах и в целом позволяет удовлетворить выдвигаемые к стойкости хеш-функций требования. Однако безопасность таких хеш-функций не может быть строго доказана, т.к. они не строятся на основе математических моделей.

С другой стороны, хеш-функции на основе доказуемо безопасных сжимающих функций (DaCoTa, FSB, RFSB и др.) основаны на сложности решения хорошо изученных математических задач, их стойкость потенциально может быть доказана и на данный момент общепризнанна. Однако, неподтвержденность ряда гипотез в теории сложности вычислений, и как следствие отсутствие строгого доказательства нижних пределов сложности решения соответствующих задач, развитие методов их решения и регресс оценок их сложности делают такие хеш-функции потенциально уязвимыми. Также такие хеш-функции, как правило, очень требовательны к вычислительным ресурсам. Эти факторы обусловили недостаточное их распространение.

В качестве альтернативного подхода к конструированию сжимающих функций в последнее время все чаще предлагается использовать достижения теории динамического хаоса. Примерами хеш-функций на основе хаотических отображений являются: CHA-1, CBHF и другие. Однако большинство таких функций имеют высокие требования к вычислительным ресурсам, что обусловило крайне низкую их востребованность при решении прикладных задач.

Методы хаотической динамики и конструирование хеш-функций

С момента начала разработки теории нелинейных динамических систем начались попытки имплементации ее достижений в различных сферах деятельности человека, в том числе и в задачах конструирования хеш-функций. В данном случае привлекательными для разработчиков оказались такие свойства хаотических систем как: высокая чувствительность к начальным значениям параметров (состояниям) хаотической системы (лавинный эффект), непредсказуемость на больших интервалах наблюдения, достаточно простые математические модели систем, облегчающие их анализ и изучение, возможность использования для исследований, как аналитических методов, так и методов теории вероятности и т.д. На примере логистического отображения рассмотрим общие принципы построения хеш-функций на основе хаотических отображений. Рекуррентное уравнение для логистического отображения задается в виде:

$$x_{i+1} = \lambda x_i (1 - x_i) = f^{(i+1)}(x_0) = f(f(f(...f(x_0)))), \quad (1)$$

где $f(x)$ – логистическое отображение; $f^{(i)}(x)$ – i-я итерация отображения; $x_i \in (0,1)$ – точки траектории; x_0 – начальное значение; $\lambda \in (0,4)$ – управляющий параметр отображения (часто называемый ключевым). Несмотря на свою простоту, отображение (1) обладает

всеми фундаментальными свойствами хаотических отображений: чувствительность к начальным условиям, достаточно быстрая разбегаемость, неповторяемость и некоррелированность траекторий, и т.д. Пример разбегания траекторий отображения (1) при малых ошибках в задании начальных условий показан на рис. 2. Динамика отображения проиллюстрирована на рис. 3 с помощью диаграммы Ламерея.

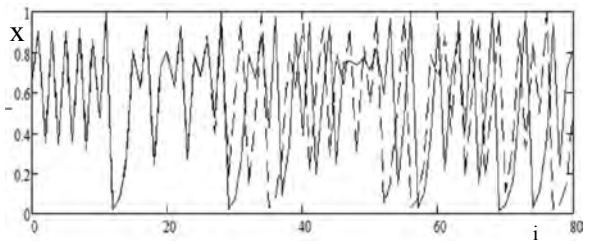

Рис. 2. Разбегание траекторий логистического отображения при малых возмущениях начальных значений

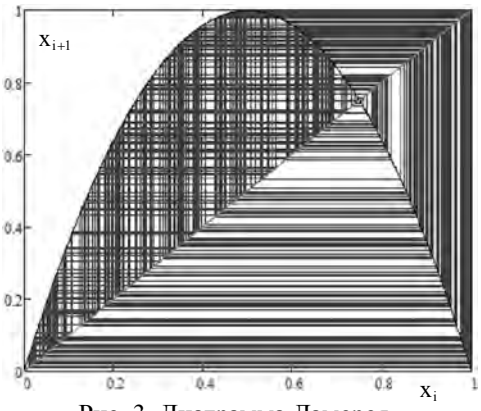

Рис. 3. Диаграмма Ламерея логистического отображения

Как правило, информационный блок вносится в сжимающую функцию (рис. 1) на основе хаотического отображения в виде возмущений управляющего параметра(ов), начального значения отображения, числа итераций отображения, либо некой комбинации вариантов. Каждый из этих подходов имеет свои достоинства и недостатки.

Так, при конструировании хеш-функций разработчики зачастую достаточно ограничены в выборе значений параметров отображений. Например, для отображения (1) областью определения параметра $\lambda \in (0,4)$. Однако только для значений $\lambda \in (3.57,4)$ отображение (1) потенциально может демонстрировать хаотическое поведение. Но и в этом интервале присутствуют т.н. окна периодичности, в которых даже при бесконечной точности представления данных и вычислений траектории отображения будут цикличными, что показано на бифуркационной диаграмме на рис. 4.

Также следует помнить, что зачастую управляющий параметр достаточно сильно влияет на плотность распределения отображения, поэтому для соблюдения статистических свойств хеш-функции значительное варьирование его значениями может оказаться неприемлемым. Поэтому разработчики, как правило, используют лишь малые возмущения управляющего параметра.

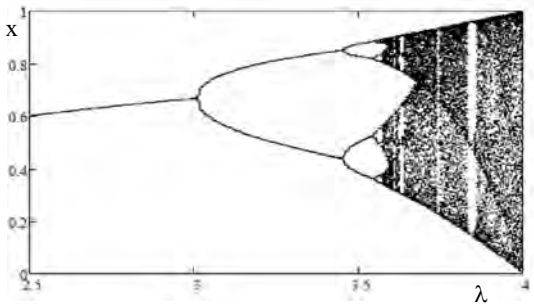

Рис. 4. Поведение хаотических систем на конечном множестве состояний

Использование информационного блока для выбора начальных точек отображения также не всегда является оправданным. С одной стороны, всегда существует вероятность попадания на достаточно короткую циклическую орбиту. Это обуславливается спецификой реализации хаотических отображений на конечных множествах и будет в общих чертах рассмотрено ниже. Отметим лишь, что более подробно задача выбора начальных значений и их влияние на свойства отображений были рассмотрены в публикации [3], а корректное решение этой задачи в дискретных аппроксимациях хаотических отображений имеет не меньшее значение, чем выбор значений управляющих параметров. С другой стороны, связь между раундами сжимающих функций в модели Меркла-Дамгарда происходит передачей выхода одного раунда на вход другого, что необходимо для противодействия атакам дополнением (в т.ч. вставкой) исходного сообщения. Таким выходом является, как правило, текущее состояние отображения, и разработчики зачастую используют лишь малые возмущения к этим состояниям на основе нового информационного блока.

Использование информационного блока для вычисления числа итераций отображения в сжимающей функции также имеет некоторые предостережения, которые, с одной стороны, обусловлены наличием относительно длинных участков расхождения траекторий, а с другой – падением производительности и скатыванию траекторий в цикл при достаточно большом числе итераций. Так, если в раундах хеш-функции число итераций отображений невелико, то расхождение их траекторий при близких начальных значениях может быть незначительным (см. рис. 2), что будет способствовать возникновению близких коллизий. В то же время большое число итераций может привести к значительному падению производительности, схождению траектории в цикл и, как следствие, к возникновению коллизий.

Фундаментальные ограничения дискретных аппроксимаций хаотических отображений в хеш-функциях

Детерминированные системы с нелинейными преобразованиями информации, обладающими хорошими перемешивающими свойствами, поведение которых слабо поддается предсказанию на больших интервалах наблюдения, казалось, как нельзя лучше подходят для построения хеш-функций. Однако впоследствии, в силу ряда фундаментальных проблем и ограничений при реализации хаотических систем на ограниченном множестве их состояний в дискретных вычислительных системах, было показано, что подавляющее большинство предложенных на их основе хеш-функций обладают значительными изъянами. Рассмотрим детальнее причины возникновения этих изъянов и уязвимостей.

Как правило, областью определения любых хаотических отображений (систем) является ограниченное некоторым интервалом множество вещественных чисел, являющееся бесконечным. Например, для отображения (1) областью определения является множество вещественных чисел на интервале $X \in (0,1)$. Именно в пределах своей области определения хаотические системы, в частности, задаваемые рекуррентными уравнениями, демонстрируют хаотическое поведение и проявляют присущие им свойства, такие как: бесконечность, неповторяемость, непредсказуемость и некоррелированность значений (точек) траекторий, высокая чувствительность к начальным параметрам и т.д. В то же время современные цифровые вычислительные средства при обработке данных ограничены разрядностью своих процессоров и объемами памяти, т.е. работают только с конечными подмножествами рациональных чисел, а точность представления данных зависит от характеристик вычислительных систем.

Таким образом, при реализации хаотических динамических систем на конечном множестве состояний в современных вычислительных средствах их траектории становятся цикличными (конечными), а длина цикла k зависит от точности представления данных и не может быть выше $k \le 2^L$, где L – точность представления данных в двоичной системе. Исследованию поведения и свойств хаотических динамических систем, задаваемых рекуррентными уравнениями, на конечном множестве состояний посвящено ряд работ, например, [4, 5]. В этих же работах можно найти эвристические и экспериментальные оценки средней длины цикла для различных отображений при разной точности вычислений. Длина циклов, как правило, значительно меньше предельного значения 2^L, а вхождению в цикл предшествует некоторый (как правило, относительно небольшой) уникальный участок траектории. Длина уникального участка (длительность вхождения в цикл) и длина самого цикла зависят от структуры рекуррентного уравнения, выбранных начальных значений, управляющих параметров и точности представления данных (вычислений). Иллюстрация поведения хаотических систем на конечном множестве состояний в цифровых вычислительных системах схематически приведена на рис. 5.

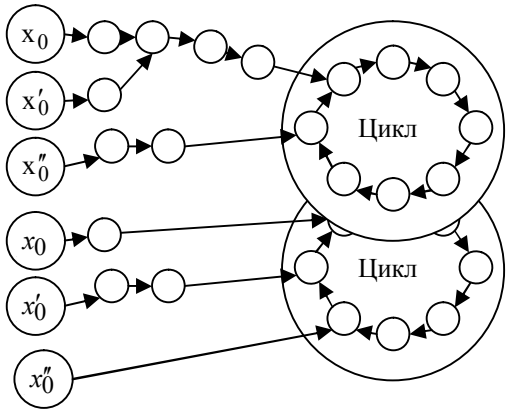

Рис. 5. Поведение хаотических систем на конечном множестве состояний

Пример диаграммы Ламерея отображения (1) на дискретном множестве состояний (например, при точности представления данных/вычислений 0,025) показан на рис. 6. На этом рисунке пунктиром обозначены участки вхождения в цикл, сплошной линией – цикличная орбита.

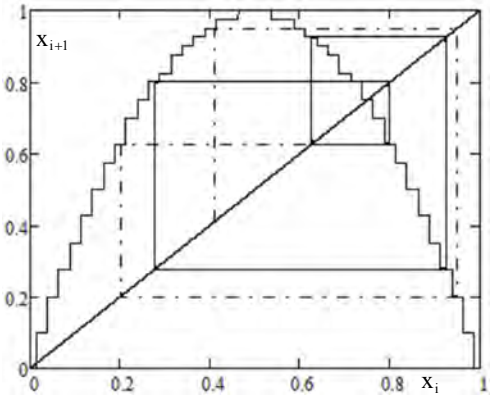

Рис. 6. Диаграмма Ламерея логистического отображения на конечном множестве состояний

Таким образом, из рис. 5, 6 видно, что дискретность представления данных и ограниченная точность вычислений меняют природу хаотических систем и приводят к деградации хаотических режимов поведения к псевдослучайному или набору циклических орбит/траекторий. При этом теряется значительная часть свойств хаотических отображений, на которые опираются при построении хеш-функций.

Способы устранения недостатков дискретных аппроксимаций хаотических отображений в хеш-функциях

На данный момент выработано несколько подходов к преодолению ограничений дискретных аппроксимаций отображений в хеш-функциях:

– использование более сложных схем преобразования информации в сжимающих функциях с дополнительным перемешиванием/перестановками (напр. рис. 7 из [6]). Такой подход, несмотря на свою достаточно высокую эффективность, явно противоречит самой идее использования хаотических отображений для построения хеш-функций. В частности, одним из достоинств применения хаотических отображений для построения хеш-функций являются достаточно простые математические модели, облегчающие их анализ и изучение. При этом предполагается, что стойкость хеш-функций будет определяться именно свойствами отображений. Однако в данном подходе свойства хаотических отображений уже не являются определяющими, и стойкость хеш-функции во многом определяется уже ее структурой, а не сутью преобразований. В рамках такого подхода без потери характеристик могут использоваться и другие нелинейные операции в сжимающей хеш-функции, в том числе, вычислительно более эффективные (например, битовые перестановки/подстановки). Т.е. такие хеш-функции теряют прямую ассоциацию с хаотическими отображениями, а их анализ и изучение становится нетривиальной задачей;

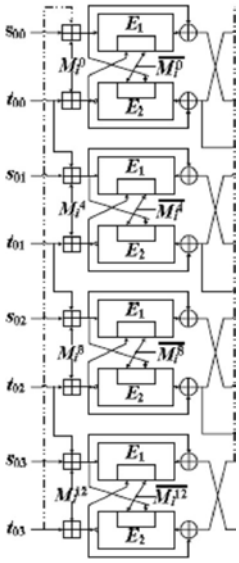

Рис. 7. Пример сжимающей функции на основе хаотического отображения (здесь E – хаотическое отображение) со «сложной структурой»

– использование комбинации вариантов внесения информации в сжимающую функцию на основе хаотического отображения (в виде возмущений управляющего параметра(ов), начального значения отображения, а также числа итераций отображения). Примером такого подхода могут служить [7 – 9]. Он позволяет в большей мере сохранить принципы построения хеш-функции на основе хаотических отображений (хотя тоже несколько усложняет их анализ) и при этом достичь более точного приближения дискретных аппроксимаций хаотических отображений к их реальным аналогам за счет постоянного перемешивания/переключения между разными траекториями/орбитами;

– использование более сложных (дву-, трех- и более мерных) отображений, например [10]. Хотя разработчики утверждают, что данный подход позволяет добиться более высоких характеристик стойкости хеш-функций за счет высоких характеристик более сложных отображений (как то Чебышева, Якоби, и т.д.) – с системной точки зрения отображения более высокой размерности «не более хаотичны», чем их одномерные (и более простые) аналоги (такие как логистическое или даже палаточное отображения). Фактически этот подход – не более чем один из способов увеличения точности вычислений за счет расширения набора рассчитываемых параметров отображения. Естественно, это достигается более сложной структурой вычислений, а значит и увеличением ресурсоемкости хеш-функций на их основе. Хотя в целом этот подход является приемлемым и может найти определенную сферу применения, апеллирование к «высоким» свойствам многомерных отображений при обосновании стойкости конструируемых функций является несколько «надуманным». Скорее в данном случае следует говорить о расширении размерностей множества, на котором определено отображение, а значит о более качественной аппроксимации в вычислительных системах.

Общие принципы применения дискретных аппроксимаций хаотических отображений для конструирования хеш-функций

На основе проведенного в работе анализа и обобщения подходов и специфики конструирования хеш-функций на основе хаотических отображений можно выделить основные принципы применения дискретных аппроксимаций хаотических отображений для конструирования хеш-функций:

– комбинация параллельных вычислений с итеративными схемами на основе модели Меркла-Дамгарда. Распараллеливание вычислений – как общий тренд построения хеш-функций – позволит в должной мере компенсировать недостатки хеш-функции на основе хаотических отображений связанные с ресурсоемкостью вычислений. А использование схемы Меркла-Дамгарда – сконцентрироваться на разработке сжимающих функций и упростить процесс конструирования и анализа;

– применение комбинации вариантов внесения информации в сжимающую функцию на основе хаотического отображения (в виде возмущений управ-

ляющего параметра(ов), начального значения отображения, а также числа итераций отображения), что позволит в значительной мере нивелировать недостатки реализации хаотических отображений в дискретных вычислительных системах;

– анализ, изучение и выработка рекомендаций по выбору инициализирующих значений вычислений, а также точности их представления и организации вычислений. В качестве одной из мер приближения дискретных аппроксимаций отображений к эталонным моделям возможно применение дву- и более мерных отображений;

– контроль расходимости траекторий, а также их скатывания в предельно малые циклы.

Выводы

В работе выполнен анализ общих подходов и тенденций построения хеш-функций на основе хаотических отображений. Было отмечено, что применение хаотических отображений для построения хеш-функций является перспективным и многообещающим направлением, в рамках которого, возможно, удастся построить вычислительно эффективные и стойкие к поиску/возникновению коллизий хеширующие функции/алгоритмы. Такое предположение опирается на свойства хаотических систем: как динамические (случайноподобность), так и структурные (детерминированность и относительная простота моделей).

В то же время значительная часть предложенных разработчиками практических реализаций этого подхода являются уязвимыми, либо обладают иными изъянами (высокая требовательность к вычислительным ресурсам, сложность моделей и т.д.). В работе были проанализированы причины несоответствия реально получаемых свойств хеш-функций ожидаемым в рамках имплементации подхода по использованию хаотических отображений для построения хеш-функций. Исходя из выявленных фундаментальных ограничений, были выработаны общие принципы использования дискретных аппроксимаций хаотических отображений в хеш-функциях, которые, возможно, позволят построить эффективные и стойкие хеширующие функции/алгоритмы.

Список литературы

1. Al-Kuwari, S. Cryptographic Hash Functions: Recent Design Trends and Security Notions [Text] / Saif Al-Kuwari, James Davenport, Russell Bradford // Proceedings of Inscrypt '10. – Science Press of China, 2010. – P. 133-150 (Режим доступа к полной он-лайн версии доклада: http : www/eprint.iacr.org/2011/565).

2. Merkle R.C. A Certified Digital Signature. In Advances in Cryptology [Text] / R.C. Merkle // CRYPTO '89 Proceedings, Lecture Notes in Computer Science. – G. Brassard, ed, Springer-Verlag. – 1989. – V. 435. – P. 218-238.

3. Костенко П.Ю. Оценка структурной скрытности хаотических сигналов / П.Ю. Костенко, С.Н. Симоненко, А.Н. Барсуков, А.В. Антонов // Известия ВУЗов. Сер. Радиоэлектроника: Науч.-техн. журн. – 2012. – Том 55, № 11. – С. 3-10.

4. Wang S.H. Periodicity of chaotic trajectories in realizations of finite computer precisions and its implication in chaos communications [Text] / S.H. Wang, W.R. Liu, H.P. Lu, J.Y. Kuang, G. Hu // International Journal of Modern Physics. – 2004. – B 18(17-19). – P. 2617-2622.

5. Goitia C.B. A model for computational collapse of chaotic 1D maps [Text] / C.B. Goitia // International Symposium on Signals, Circuits and Systems, 2005: Proceedings of ISSCS 2005 (14-15 July 2005). – 2005. – Vol. 2. – P. 741-744.

6. Dong Liao. Parallel hash function using DM-based integer-valued chaotic maps network [Electronic resource] / Liao Dong, Wang Xiaomin. Access mode: http : www/paper.edu.cn/en_releasepaper/content/4500967.

7. 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.

8. 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.

9. Антонов А.В. Вариант эффективной реализации метода построения хеш-функции на основе хаотических отображений с переменными параметрами и параллельной организацией вычислений / А.В. Антонов, В.Б. Бзот // Системи обробки інформації. – Х.: Харківський університет Повітряних Сил, 2013. – Вип. 1(108). – С. 187-191.

10. Akhavan A. A novel parallel hash function based on 3D chaotic map [Electronic resource] / Amir Akhavan, Azman Samsudin and Afshin Akhshani // EURASIP Journal on Advances in Signal Processing, 2013. – Access mode: http : www/asp.eurasipjournals.com/content/2013/1/126.

Поступила в редколлегию 19.08.2013

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

ЗАГАЛЬНІ ПРИНЦИПИ ЗАСТОСУВАННЯ ДИСКРЕТНИХ АПРОКСИМАЦІЙ ХАОТИЧНИХ ВІДОБРАЖЕНЬ ДЛЯ КОНСТРУЮВАННЯ ҐЕШ-ФУНКЦІЙ

А.В. Антонов, І.Є. Кужель, Н.В. Шигімага

Аналізуються загальні підходи і тенденції побудови ґеш-функцій на основі хаотичних відображень. Зазначено, що часто реальні властивості таких ґеш-функцій не відповідають очікуваним. Виявлено причини виникнення уразливостей і вад в ґеш-функціях на основі хаотичних відображень. Вироблені принципи і рекомендації з використання дискретних апроксимацій хаотичних відображень для їх побудови.

***Ключові слова**: ґеш-функція, хаотичне відображення, ефективність, колізія.*

GENERAL PRINCIPLES OF APPLICATION OF CHAOTIC MAPS DISCRETE APPROXIMATIONS FOR THE HASH FUNCTIONS CONSTRUCTION

A.V. Antonov, I.Ye. Kuzhel, N.V. Shigimaga

Common approaches and trends in construction of hash functions based on chaotic maps were analyzed. Were pointed out that often the real properties of such hash functions are not as expected. Causes of vulnerabilities and flaws in a hash function based on chaotic maps were identified. Principles and guidelines on the use of discrete approximations of chaotic maps for their construction were developed.

***Keywords**: hash function, the chaotic map, effectiveness, collision.*