ПРИКЛАДНАЯ ДИСКРЕТНАЯ МАТЕМАТИКА

№ 9 ПРИЛОЖЕНИЕ Сентябрь 2016

Секция 8

ВЫЧИСЛИТЕЛЬНЫЕ МЕТОДЫ В ДИСКРЕТНОЙ МАТЕМАТИКЕ

УДК 519.7 DOI 10.17223/2226308X/9/51

ПРИМЕНЕНИЕ АЛГОРИТМОВ РЕШЕНИЯ ПРОБЛЕМЫ БУЛЕВОЙ ВЫПОЛНИМОСТИ К ПОСТРОЕНИЮ РАЗНОСТНЫХ ПУТЕЙ В ЗАДАЧАХ ПОИСКА КОЛЛИЗИЙ КРИПТОГРАФИЧЕСКИХ ХЕШ-ФУНКЦИЙ СЕМЕЙСТВА MD[1]

И. А. Грибанова

Представлен новый метод построения разностных путей в задаче поиска коллизий криптографических хеш-функций, основанный на использовании алгоритмов решения проблемы булевой выполнимости. На начальном этапе метода строится пропозициональная кодировка задачи поиска коллизий рассматриваемой хеш-функции. Затем в полученную кодировку добавляются (в виде КНФ) дополнительные ограничения. Как правило, это значения целочисленных разностей на сообщения, дающие коллизию, а также на отдельные фрагменты дифференциального пути. В качестве отправной точки поиска можно использовать некоторый известный либо случайный дифференциальный путь (во втором случае наличие коллизий, удовлетворяющих такому пути, не гарантируется). Задача варьирования значений разности переменных, кодирующих заполнение хеш-регистров, сводится к SAT. Для эффективного решения соответствующих серий SAT-задач написана MPI-программа, работающая на вычислительном кластере. Основным результатом стало построение дифференциального пути для задачи поиска коллизий криптографической хеш-функции MD4, отличного от известных.

Ключевые слова: *криптографическая хеш-функция, коллизия, разностные атаки, задача о булевой выполнимости (SAT), хеш-функции семейства MD.*

Криптографические хеш-функции являются важным примитивом, используемым в многочисленных приложениях современной криптографии. Хеш-функция — это всюду определённая алгоритмически вычислимая функция вида $\chi : \{0,1\}^* \mapsto \{0,1\}^C$, где C — некоторая константа, называемая длиной хеша. При $n > C$ в $\{0,1\}^*$ обязательно найдутся такие $x_1, x_2 \in \{0,1\}^n$, $x_1 \neq x_2$, что имеет место $\chi(x_1) = \chi(x_2)$. В этом случае говорят, что сообщения x_1, x_2 дают коллизию. К криптографическим хеш-функциям предъявляется ряд дополнительных требований, кратко суммировать которые можно так: задачи, так или иначе связанные с обращением хеш-функции, должны быть вычислительно трудными. В частности, высокую вычислительную трудность должна иметь и задача поиска коллизий.

Многие современные криптографические хеш-функции имеют в своей основе конструкцию Меркля — Дамгарда [1, 2], в соответствии с которой процесс построения хеш-образа исходного сообщения выглядит как последовательность итераций, на каждой из которых происходит обновление содержимого специальных регистров, далее назы-

[1]Работа выполнена при частичной финансовой поддержке РФФИ, проект № 14-07-00403а.

ваемых хеш-регистрами. На начальном шаге в хеш-регистры записываются константы, определённые в спецификации алгоритма. Затем содержимое хеш-регистров итеративно обновляется за счёт последовательного замешивания в них подписываемого сообщения. Длина хеш-регистра для различных алгоритмов может варьироваться от 128 до 512 битов, а число итераций, называемых далее шагами, от 48 до 80. Данные в хеш-регистрах разделены на слова длиной 32 или 64 бита. К наиболее популярным криптографическим хеш-функциям, основанным на конструкции Меркля — Дамгарда, относятся функции семейства SHA, а также функции семейства MD (MD4, MD5, RIPEMD).

В 2004–2005 гг. коллективом китайских криптографов, возглавляемым К. Ванг (X. Wang), был предложен подход к поиску коллизий криптографических хеш-функций семейства MD [3, 4]. Данная атака получила название разностной, или дифференциальной из-за сходства подхода, лежащего в её основе, с дифференциальным криптоанализом. Основная идея «атаки Ванг» заключается в попытке дополнить уравнения, задающие хеш-функцию, дополнительными ограничениями, которые резко сужают пространство поиска коллизий. Далее мы описываем основу метода Ванг, подразумевая, что рассматривается задача поиска коллизии хеш-функции MD4. Рассмотрим два процесса вычисления хеш-значений MD4 для двух 512-битных сообщений, традиционно обозначаемых через M, M'. Обозначим через H и H' хеш-регистры, заполняемые при вычислении хеш-значений M, M'. На шаге с фиксированным номером k в H и H' записаны некоторые 32-битные слова. На эти слова, рассматриваемые как целые числа, накладываются дополнительные ограничения, имеющие вид целочисленных разностей, обозначаемых через δ_k. Разность между значениями хеш-регистров на последнем шаге должна быть равна нулю, что означает условие равенства хешей: $\chi(M) = \chi(M')$, то есть сообщения M, M' дают коллизию. Множество значений δ_k для всех или части шагов (в MD4 число шагов равно 48) задаёт так называемый дифференциальный (разностный) путь. На сообщения M, M' также накладывается дополнительное ограничение, имеющее вид разности для соответствующих целых 512-битных чисел.

В [5] для задачи поиска коллизий хеш-функций семейства MD успешно применён SAT-подход; более точно, реализована SAT-версия разностной «атаки Ванг». Как результат, при помощи SAT-решателя `minisat` были найдены одноблоковые коллизии для хеш-функции MD4 и представлены реалистичные оценки времени поиска двухблоковых коллизий для MD5. Во всех случаях использовались разностные пути, представленные в [3, 4]. В [6] аналогичные задачи решены с более высокой эффективностью и построено семейство специального вида двухблоковых коллизий хеш-функции MD5; для сведения задач поиска коллизий MD4 и MD5 к SAT использован программный комплекс `Transalg` [7, 8].

В ряде последующих работ [9 – 11] представлены попытки построения новых дифференциальных путей, отличных от приведённых в [3, 4]. Наиболее успешными в этом плане можно считать результаты [11]. Особо отметим, что во всех этих случаях новые дифференциальные соотношения являлись результатом громоздких построений, фактически выполняемых авторами «вручную». Мы предлагаем автоматизировать данный процесс за счёт использования современных технологий решения SAT, имея в виду работу [5] как пример автоматизации «атаки Ванг» посредством SAT.

На данном этапе новые разностные соотношения для значений хеш-регистров, позволяющие быстро находить коллизии, построены для функции MD4. Остановимся кратко на основах разработанного метода. Начальным этапом является построение

пропозициональной кодировки алгоритма вычисления хеш-функции; для этой цели используется комплекс `Transalg`. Результатом является построение двух КНФ (так называемых «шаблонов» [8]) C и C' над множествами переменных X и X', $X \cap X' = \varnothing$. Возможности `Transalg` позволяют отслеживать связь между переменными, входящими в C, C', и содержимым хеш-регистров, получаемым в процессе вычисления хеш-образов сообщений M и M'. На переменные в C, C', кодирующие заполнение хеш-регистров на последнем шаге, накладывается условие их логической эквивалентности. Полученная в результате КНФ $\tilde{C}$ выполнима тогда и только тогда, когда существуют сообщения M и M', дающие коллизию. К КНФ C можно конъюнктивно приписывать различные дополнительные условия. В частности, к $\tilde{C}$ может быть приписана КНФ C_Δ, принимающая значение 1 тогда и только тогда, когда разность целых чисел, заданных 512-битными векторами M и M', равна некоторой константе Δ.

Пусть $C_{\delta_k=c}$ — КНФ, принимающая значение 1 тогда и только тогда, когда переменная δ_k принимает значение c. Поскольку c — 32-битная константа, не представляет труда перебрать все возможные её значения и рассмотреть проблемы выполнимости соответствующих КНФ вида $\tilde{C} \wedge C_\Delta \wedge C_{\delta_k=c}$. Экспериментально установлено, что для подавляющего большинства возможных значений c современные SAT-решатели доказывают невыполнимость $\tilde{C} \wedge C_\Delta \wedge C_{\delta_k=c}$ очень быстро (за доли секунды). Задачи о выполнимости КНФ вида $\tilde{C} \wedge C_\Delta \wedge C_{\delta_k=c}$, для которых не удаётся получить ответ за относительно небольшое время (несколько секунд), прерываются. Пусть c' — значение δ_k, при котором решение SAT-задачи для КНФ $\tilde{C} \wedge C_\Delta \wedge C_{\delta_k=c'}$ прервано, и при этом c' не совпадает с соответствующим значением в пути Ванг. Тогда c' можно рассматривать как кандидата на значение разности δ_k в новом дифференциальном пути. Описанный алгоритм реализован в виде параллельной MPI-программы и запускался на вычислительном кластере «Академик В. М. Матросов» ИНЦ СО РАН. В результате его работы построен новый дифференциальный путь для задачи поиска коллизий криптографической хеш-функции MD4. Данный путь отличается от пути Ванг [3] разностными соотношениями для шагов с номерами $k \in \{13, 17, 20, 21\}$. Введя в пропозициональную кодировку MD4 новый путь и используя значение разности между сообщениями M, M', взятое из [3], мы построили КНФ, к которой применили SAT-решатель `cryptominisat` [12], имеющий функцию перечисления множеств решений. Для построенной КНФ данный решатель нашел 1000 различных коллизий за 416 с. При применении к аналогичной КНФ, в которой записан оригинальный путь Ванг, `cryptominisat` находит 1000 коллизий за 520 с. На наш взгляд, полученные результаты убедительно демонстрируют применимость SAT-подхода к анализу стойкости криптографических хеш-функций к разностным атакам.

ЛИТЕРАТУРА

1. *Merkle R. A.* Certified digital signature // LNCS. 1990. V. 435. P. 218–238.
2. *Damgard I. A.* A design principle for hash functions // LNCS. 1990. V. 435. P. 416–427.
3. *Wang X., Lai X., Feng D., et al.* Cryptanalysis of the hash functions MD4 and RIPEMD // LNCS. 2005. V. 3494. P. 1–18.
4. *Wang X. and Yu H.* How to break MD5 and other hash functions // LNCS. 2005. V. 3494. P. 19–35.
5. *Mironov I. and Zhang L.* Applications of SAT solvers to cryptanalysis of hash functions // LNCS. 2006. V. 4121. P. 102–115.

6. *Богачкова И. А., Заикин О. С., Кочемазов С. Е. и др.* Задачи поиска коллизий для криптографических хеш-функций семейства MD как варианты задачи о булевой выполнимости // Вычислительные методы и программирование. 2015. Т. 16. С. 61–77.
7. *Отпущенников И. В., Семёнов А. А.* Технология трансляции комбинаторных проблем в булевы уравнения // Прикладная дискретная математика. 2011. № 1. С. 96–115.
8. *Otpuschennikov I., Semenov A., and Kochemazov S.* Transalg: a tool for translating procedural descriptions of discrete functions to SAT // WCSE 2015-IPCE: Proc. 5th Intern. Workshop on Computer Science and Engineering: Information Processing and Control Engineering. 2015. P. 289–294.
9. *Hawkes P., Paddon M., and Rose G.* Musings on the Wang et al. MD5 Collision. IACR Eprint archive. `http://eprint.iacr.org/2004/264`. 2004.
10. *Hirshman G.* Further Musings on the Wang et al. MD5 Collision: Improvements and Corrections on the Work of Hawkes, Paddon, and Rose. Cryptology ePrint Archive, Report 2007/375. 2007.
11. *Stevens M.* Attacks on Hash Functions and Applications. PhD Thesis. Amsterdam: Ipskamp Drukkers, 2012. 258 p.

УДК 519.688 DOI 10.17223/2226308X/9/52

О ВЫЧИСЛЕНИИ ФУНКЦИЙ РОСТА КОНЕЧНЫХ ДВУПОРОЖДЁННЫХ БЕРНСАЙДОВЫХ ГРУПП ПЕРИОДА 5[1]

А. А. Кузнецов, С. С. Карчевский

Пусть $B_0(2,5) = \langle a_1, a_2 \rangle$ — наибольшая конечная двупорождённая бернсайдова группа периода 5, порядок которой равен 5^{34}. Для каждого элемента данной группы существует уникальное коммутаторное представление вида $a_1^{\alpha_1} \cdot a_2^{\alpha_2} \cdot \ldots \cdot a_{34}^{\alpha_{34}}$, где $\alpha_i \in \mathbb{Z}_5$, $i = 1, 2, \ldots, 34$. Здесь a_1 и a_2 — порождающие элементы $B_0(2,5)$; $a_3, \ldots, a_{34}$ — коммутаторы, которые вычисляются рекурсивно через a_1 и a_2. Определим фактор-группу группы $B_0(2,5)$ следующего вида: $B_k = B_0(2,5)/\langle a_{k+1}, \ldots, a_{34} \rangle$. Очевидно, что $|B_k| = 5^k$. В настоящей работе вычислены функции роста B_k относительно порождающих множеств $\{a_1, a_2\}$ и $\{a_1, a_1^{-1}, a_2, a_2^{-1}\}$ для $k = 15$, 16, 17.

Ключевые слова: *функция роста группы, группа Бернсайда.*

Одним из важных инструментов для определения строения группы является изучение её роста относительно фиксированного порождающего множества. Пусть $G = \langle X \rangle$. *Шаром* K_s радиуса s группы G будем называть множество всех её элементов, которые могут быть представлены в виде групповых слов в алфавите X длиной не больше s. Для каждого целого неотрицательного s можно определить функцию роста группы $F(s)$, которая равна числу элементов группы G относительно X, представимых в виде несократимых групповых слов длиной s. Таким образом,

$$F(0) = |K_0| = 1,\ F(s) = |K_s| - |K_{s-1}| \text{ при } s \in \mathbb{N}.$$

Как правило, функцию роста конечной группы представляют в виде таблицы, в которую записывают ненулевые значения $F(s)$.

Пусть $F(s_0) > 0$, но $F(s_0 + 1) = 0$, тогда s_0 является диаметром графа Кэли группы G в алфавите порождающих X, который будем обозначать $D_X(G)$.

[1]Работа поддержана грантом Президента РФ (проект МД-3952.2015.9).