Merkle-Damgård vs Sponge: сравнительный анализ двух конструкций функций хэширования

Григорий Маршалко, Василий Шишкин

ФСБ России

26 марта 2014

Последние "новинки модельного ряда"

- Хэш-функции семейства *Стрибог* в 2012 г. приняты в качестве национального стандарта ГОСТ Р 34.11-2012
- Хэш-функции семейства *Keccak* в 2012 г. победили в конкурсе SHA-3 и планируются в скором времени к стандартизации в США

Разные принципы синтеза, как они соотносятся, что выбирать?

Криптографические функции хэширования

Определение

Криптографической функцией хэширования называется отображение $H : V^* \to V_n$, где $n \in \mathbb{N}$ – натуральное число, V^* – множество всех двоичных векторов конечной размерности (включая пустую строку), V_n – множество всех n-мерных двоичных векторов.

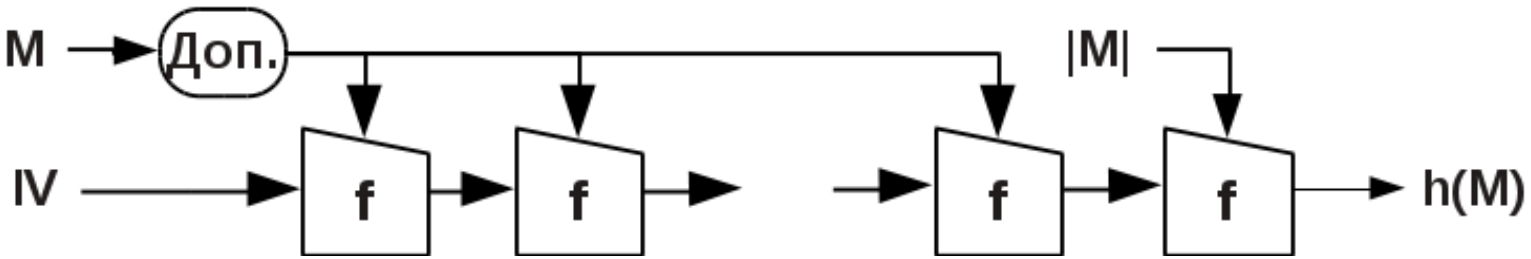
Задачи анализа

- **Обращение хэш-функции** (Preimage attack). Необходимо по известному y найти такой x, что $h(x) = y$. Сложность $\approx 2^n$.
- **Построение коллизии.** Необходимо найти такие различные сообщения x_1, x_2, что $h(x_1) = h(x_2)$. Сложность $\approx 2^{\frac{n}{2}}$.
- **Построение второго прообраза** (Second preimage attack). Необходимо по известным y и x найти такой x_1, отличный от x, что $h(x) = h(x_1)$. Сложность $\approx 2^n$.
- **Расширение сообщения.** По заданным значениям $|x|$, $H(x)$ найти некоторое значение $y \in V^*$, для которого вычислить $H(x\|y)$. Сложность $\approx 2^n$

- мульти-, псевдо-, почти- и т.д.
- **Построение различителя.** Построить алгоритм позволяющий в некоторой модели отличить хэш-функцию h от случайного отображения.

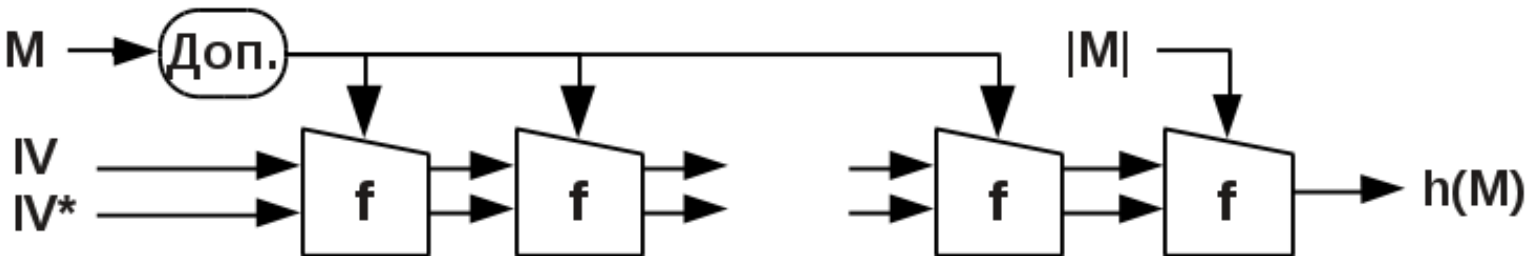
25-летие МД-конструкции

- Независимо предложена в 1989 г. Мерклем и Дамгордом
- Теоретические результаты по стойкости в рамках ряда формальных моделей (т.н. доказуемая стойкость)
- Начало 2000-х годов – новые методы анализа (мультиколлизии, расширение сообщения, прообраз для длинных сообщений...) продемонстрировали неидеальность МД-конструкции с теоретической точки зрения



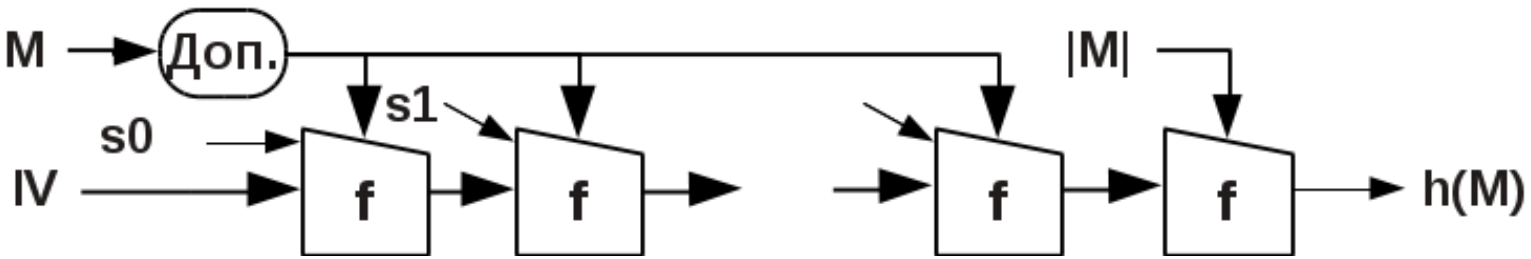
Пути противодействия атакам

- Wide-pipe (Double-pipe) (2004 г.) – увеличение внутреннего состояния (ограничения по памяти)



Пути противодействия атакам

- Wide-pipe (Double-pipe) (2004 г.) – увеличение внутреннего состояния (ограничения по памяти)
- Рандомизированное хэширование (2006 г.) – добавление "соли" (увеличение трудоемкости хэширования)

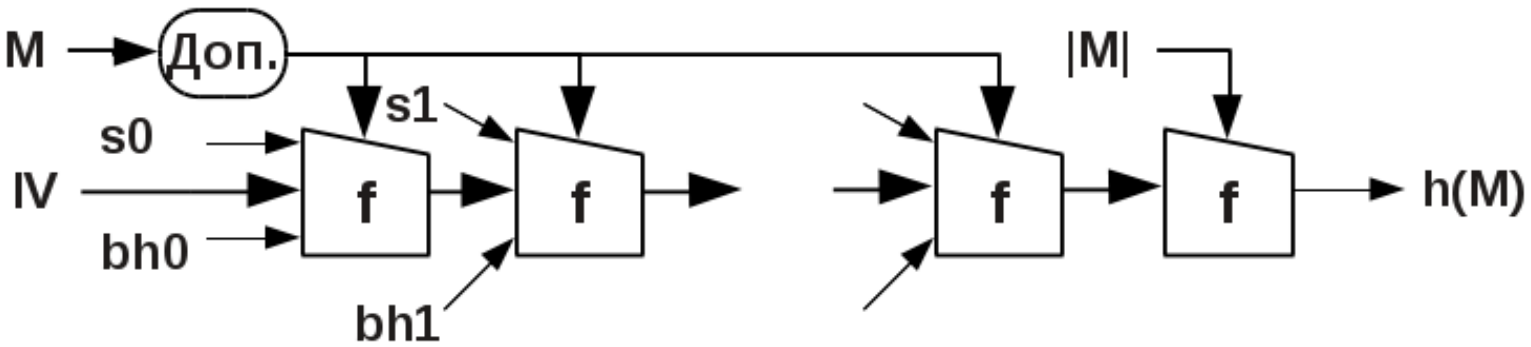


Пути противодействия атакам

- Wide-pipe (Double-pipe) (2004 г.) – увеличение внутреннего состояния (ограничения по памяти)
- Рандомизированное хэширование (2006 г.) – добавление "соли" (увеличение трудоемкости хэширования)
- Ряд других способов, но как оказалось, и они не спасают от недостатков

Пути противодействия атакам

- Wide-pipe (Double-pipe) (2004 г.) – увеличение внутреннего состояния (ограничения по памяти)
- Рандомизированное хэширование (2006 г.) – добавление "соли" (увеличение трудоемкости хэширования)
- Ряд других способов, но как оказалось, и они не спасают от недостатков
- Наиболее взвешенное решение – конструкция HAIFA (2006 г.) – добавление на очередной итерации числа хэшированных бит и, возможно, "соли"



Хэш-функции *Стрибог*

- Основана на МД-конструкции – наследование известных результатов о стойкости
- Модификация МД-конструкции – исключение возможности применения расширения сообщения, ограничение возможности применения мультиколлизий
- Основная идея: взять лучшее от МД и устранить возможные недостатки при сохранении приемлемых эксплуатационных характеристик

Результаты независимого анализа хэш-функции *Стрибог*

- AlTawy R., Kircanski A., Youssef A. M., Rebound attacks on Stribog. ICISC'2013 – применение атаки "отражения"
  - Условная псевдо-коллизия для 7.75 из 12 итераций функции сжатия
  - Условная псевдо-почти-коллизия для 9.75 из 12 итераций функции сжатия
- Wang Z., Yu H., Wang X., Cryptanalysis of GOST R hash function. Cryptology ePrint Archive – аналогичный подход
  - Коллизии для 9.5 из 12 итераций функции сжатия
  - Ограниченный различитель (по подмножествам входа и выхода) на 10 из 12 итераций функции сжатия
  - Указано на возможность построения k-коллизий (указанное ранее свойство МД-конструкции),
- AlTawy R., Youssef A. M., Integral distinguishers for reduced-round Stribog. Cryptology ePrint Archive
  - Различитель для 7 из 12 итераций функции сжатия

Короткое резюме

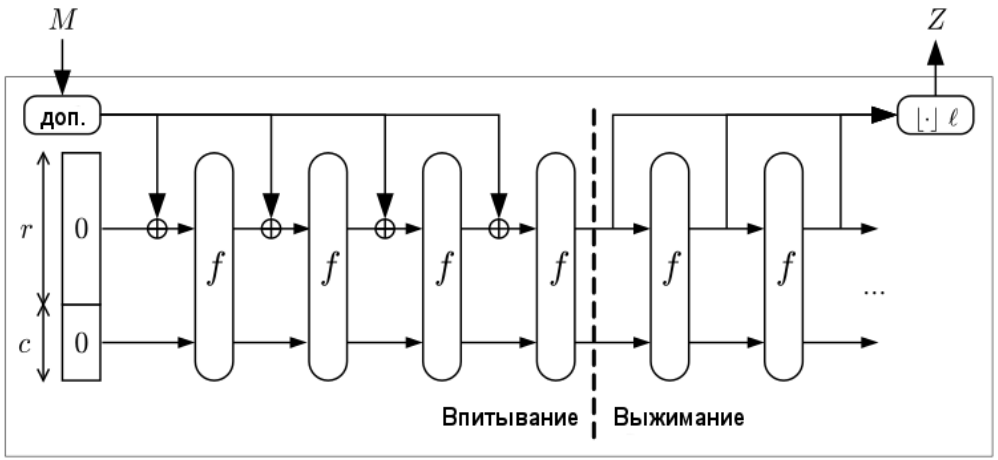
- Коллизии и различители для усеченной функции сжатия
- k-коллизии в соответствии с методом поиска мультиколлизий + контрольная сумма (противодействие – wide-pipe $\Rightarrow$, что приводит к ухудшению эксплуатационных свойств)

Sponge-конструкция

- Предложена в 2007 г. и воплощена в алгоритме *Keccak*
- Эволюция: *Panama* (1998 г.) $\rightarrow$ *RadioGatun* (2006 г.) $\rightarrow$ *Keccak* (2007 г.)
- Основана на итеративном использовании подстановки (отображения) большой размерности $\approx$ аналог wide-pipe

Sponge-конструкция

- Размер подстановки $b = r + c$
  - Скорость r – длина очередного хэшируемого блока, который поступает на вход функции сжатия (длина выходного блока функции сжатия)
  - Емкость c – длина незадействованного входа подстановки
- Дополнение: 100.....
- Длина хэш-кода n
- Две фазы функционирования
  - Впитывание – подмешивание входного сообщения к внутреннему состоянию хэш-функции
  - Выжимание – выработка хэш-кода



Анализ Sponge-конструкции

- G. Bertoni, J. Daemen, M. Peeters, G. Van Assche. Cryptographic Sponge functions. sponge.noekeon.org
  - Доказательство стойкости в формальной модели (отличимость от случайного оракула)
    - Отличие от случайного оракула – наличие внутренних коллизий
    - Авторы предлагают изменить парадигму идеальной хэш-функции – вместо повсеместно используемого случайного оракула использовать "случайную губку" – тонкий момент!
    - "Стратегия герметичной губки" – неотличимость реального алгоритма
  - Стойкость относительно классических атак – на основе большой размерности подстановки
- E. Andreeva, B. Mennink, B. Preneel. The Parazoa Family: Generalizing the Sponge Hash Functions. Cryptology ePrint Archive,

Стойкость Sponge-конструкции

- **Обращение хэш-функции.** Сложность $\approx min\left\{2^{min\{n,b\}}, max\left\{2^{min\{n,b\}-r}, 2^{\frac{c}{2}}\right\}\right\}$.
- **Построение коллизии.** Сложность $\approx min\left\{2^{\frac{n}{2}}, 2^{\frac{c}{2}}\right\}$.
- **Построение второго прообраза.** Сложность $\approx min\left\{2^{n}, 2^{\frac{c}{2}}\right\}$.
- Остальные атаки – точных оценок нет, авторы обосновывают стойкость на качественном уровне (с учетом большой длины подстановки)

Специфическое свойство – для обеспечения стойкости значение емкости c должно принимать достаточно большие значения (попытка NIST уменьшить емкость была встречена международным криптографическим сообществом крайне негативно)

Хэш-функция *Keccak*

- Размер подстановки – $b = 1600$ бит
- Емкость – (?) бит – в зависимости от длины хэш-кода
- 5 типов преобразований внутреннего состояния
  - Добавление констант
  - Нелинейное преобразование степени 2
  - Три линейных перемешивающих преобразования

Некоторые результаты анализа хэш-функции *Keccak*

- C. Boura, A. Canteaut, A zero-sum property for the Keccak-f permutation with 18 rounds, comment on the NIST Hash Competition, 2010 – Различитель, использующий малую степень нелинейности итерационного преобразования, – авторы *Keccak* увеличивают число итераций с 18 до 24
- M. Duan, X. Lai, Improved zero-sum distinguisher for full round Keccak-f permutation, Cryptology ePrint Archive Report – Аналогичный различитель уже на 24 итерации – однако авторы *Keccak* в угоду производительности отказались от "стратегии герметичной губки"и не увеличивают число итераций
- D. J. Bernstein, Second preimages for 6 (7? (8??)) rounds of Keccak?, NIST hash forum mailing list – второй прообраз для 8 итераций
- I. Dinur, O. Dunkelman, A. Shamir, New attacks on Keccak-224 and Keccak-256, FSE 2012 – коллизии за реальное время для 4 итераций

Короткое резюме

- Коллизии и прообразы только для усеченной функции сжатия
- Различитель для полной функции сжатия – свойство нелинейного преобразования (противодействие – увеличение числа итераций $\Rightarrow$ ухудшение эксплуатационных свойств. Учитывая то, что пока эти результаты не приводят к построению коллизий и прообразов, смысла нет)

Сравнение Sponge- и МД-конструкций

- Две различные конструкции: разные подходы к обоснованию стойкости и анализу
  - Для МД-конструкции за 25 лет получены существенные результаты анализа, использующие специфические свойства конструкции
  - Новая Sponge-конструкция сравнительно мало анализировалась независимыми исследователями
- Необходимо использовать разные математические модели при обосновании стойкости (хотя используются одинаковые)
  - МД-конструкция – использование блочного шифра в функции сжатия позволяет моделировать ее с помощью случайного отображения
  - Sponge-конструкция – функция сжатия – фиксированная подстановка (требует, по всей видимости, новых подходов к обоснованию стойкости)
- Различные подходы к синтезу функции сжатия
  - МД-конструкция – понятные, устоявшиеся подходы на основе блочных шифров
  - Sponge-конструкция – результаты показывают, что нужны новые способы синтеза больших подстановок

Выводы

- Два различных подхода к синтезу хэш-функций
- В каждом случае специфические эксплуатационные и криптографические свойства
- Рассмотренные в докладе примеры хэш-функций превосходят существующие стандартизированные (международными и национальными организациями) аналоги по стойкости относительно известных методов анализа