УДК 003.26

Миронкин В.О.

Национальный исследовательский университет «Высшая школа экономики», Москва

ОБ ОДНОЙ ТЕОРЕТИКО-ВЕРОЯТНОСТНОЙ МОДЕЛИ SPONGE-КОНСТРУКЦИИ

Исследованы граф внутренних состояний Sponge-конструкции и взаимосвязь между внутренними состояниями и элементами выходной последовательности. Предложены методы построения коллизий, использующие особенности цикловой структуры подстановки Sponge-конструкции. Описан общий вид соответствующих коллизий.

ХЭШ-ФУНКЦИЯ, SPONGE-КОНСТРУКЦИЯ, КОЛЛИЗИЯ, ВТОРОЙ ПРООБРАЗ, ГРАФ, ПОДСТАНОВКА, ОТОБРАЖЕНИЕ, ЦИКЛОВАЯ СТРУКТУРА.

Mironkin V.O.

National Research University Higher School of Economics, Moscow

ON SOME THEORETIC-PROBABILISTIC MODEL OF SPONGE-CONSTRUCTION

The graph of internal states of Sponge construction and the relationship between internal states and elements of the output sequence are investigated. The methods of constructing collisions which use features of the cyclic structure of Sponge construction's substitution are proposed. The general form of the corresponding collisions is described.

A HASH FUNCTION, SPONGE CONSTRUCTION, A COLLISION, THE SECOND PREIMAGE, A GRAPH, A SUBSTITUTION, A MAPPING, A CYCLIC STRUCTURE.

Введение

Sponge-конструкция [6, 7, 14] (далее Sponge), используемая в стандарте хэширования данных SHA-3 [5, 8, 10-13], представляет собой итерационный алгоритм выработки последовательности конечной длины на основе сообщений произвольной длины. Данный алгоритм использует некоторое подстановочное преобразование и состоит из процедур впитывания (absorbing), выжимания (squeezing), а также процедуры усечения выходной последовательности.

Замечание 1. *В настоящей статье изучается процесс функционирования Sponge без применения операции усечения.*

Введем ряд используемых обозначений:

r	размер входного и выходного блоков Sponge;
m	число выходных блоков Sponge;
x_i	входной блок Sponge, $x_i \in V_r$;
z_i	выходной блок Sponge, $z_i \in V_r$;
$\|\|$	конкатенация блоков;
x	входное сообщение Sponge, $x = x_0 \| \| \ldots \| \| x_{n-1}$, $n \in \mathbb{N}$;
z	выходная последовательность Sponge, $z = z_0 \| \| \ldots \| \| z_{m-1}$;
ψ_m	преобразование Sponge с m тактами процедуры выжимания;
b	размер внутренних состояний Sponge;
V^*	множество всех двоичных векторов конечной длины, включая пустую строку;
V_l	множество всех двоичных векторов длины $l \in \mathbb{N} \cup \{0\}$;
S_{2^b}	множество всех подстановок на V_b;
f	подстановка Sponge, используемая для преобразования внутренних состояний, $f \in S_{2^b}$;
s_i	внутреннее состояние Sponge, $s_i \in V_b$, $i = \overline{0,n}$;
c	размер недоступной для наблюдения и изменения части внутреннего состояния Sponge, $c = b - r$;

$\lceil x \rceil_l$	первые l бит двоичного вектора x;
$\lfloor x \rfloor_l$	последние l бит двоичного вектора x;
0^v, 1^v	нулевой и единичный векторы длины $v \in \mathbb{N}$.

На рис. 1. изображена общая схема реализации преобразования ψ_m на основе входного сообщения $x \in V^*$.

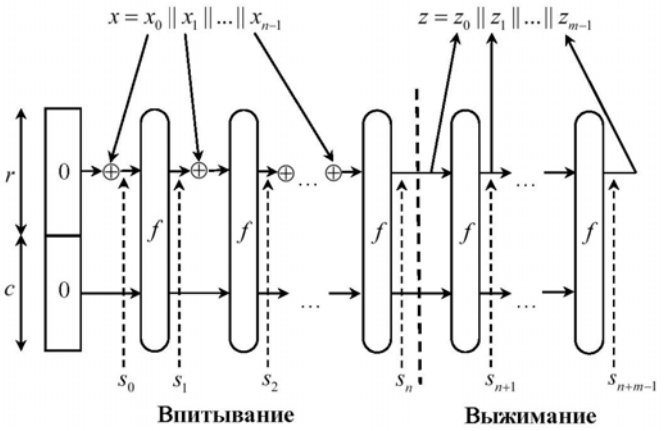

Рис. 1. Схема выработки хэш-кода

Под элементарной операцией в процессе функционирования Sponge будем понимать одно вычисление значения подстановки $f \in S_{2^b}$.

§ 1. Модель функционирования Sponge

Рассмотрим подробнее процесс преобразования внутренних состояний Sponge на основе произвольного сообщения $x = x_0 \| \ldots \| x_{n-1}$.

При впитывании первого блока x_0 вырабатывается сразу два внутренних состояния Sponge s_0 и s_1 (см. рис. 1). При дальнейшем впитывании блоков x_{i-1}, $i > 1$, вырабатывается по одному внутреннему состоянию s_i:

- $s_0 = 0^b \oplus \left(x_0 \| 0^c \right) = x_0 \| 0^c$;
- $s_1 = f\left(0^b \oplus \left(x_0 \| 0^c \right) \right)$;
- $s_i = f\left(s_{i-1} \oplus \left(x_{i-1} \| 0^c \right) \right)$ для любого $i \in \overline{2,n}$;
- $s_i = f\left(s_{i-1} \right)$ для любого $i \in \overline{n+1, n+m-1}$.

При этом для элементов выходной последовательности $\{z_i\}_{i=0}^{m-1}$ выполняется следующее соотношение:

$$z_i = s_{n+i},\ i = \overline{0,m-1}.$$

Замечание 3. *На практике [8] параметр $b = r + c$ принимает значения порядка 1600 бит, что в настоящее время не позволяет полностью описать цикловую структуру подстановки f. Поэтому будем считать, что f является случайной равновероятной подстановкой из S_{2^b}.*

Для произвольных $x \in V_b$ и $\alpha \in V_r$ рассмотрим подстановку $\varphi_\alpha : V_b \to V_b$:

$$\varphi_\alpha(x) = f\left(x \oplus \left(\alpha \,||\, 0^c\right)\right).$$

С учетом замечания 3 при случайном выборе блоков сообщения $x_i \in V_r$, $i \in \overline{0,2^r - 1}$, подстановки φ_{x_i} являются независимыми и случайными равновероятными [1-4].

На основе входного сообщения $x = x_0 \,||\, ... \,||\, x_{n-1}$ в результате впитывания вырабатывается последовательность внутренних состояний $\{s_i\}_{i=0}^{n}$:

$$s_i = \varphi_x(s_0) = \varphi_{x_{i-1}} \circ ... \circ \varphi_{x_0}(s_0),\ i \geq 1,$$

где $s_0 = x_0 \,||\, 0^c$. При этом для вырожденных случаев, когда $x_{i-1} = 0^r$, $s_{i-1} = 0^b$, справедливо равенство:

$$s_i = \begin{cases} f\left(x_{i-1} \,||\, 0^b\right), & s_{i-1} = 0^b, \\ f\left(s_{i-1}\right), & x_{i-1} = 0^r. \end{cases}$$

Т.е. при поступлении на вход Sponge нулевого блока внутреннее состояние формируется только на основе подстановки f.

§ 2. Структурные коллизии Sponge

Через G_f обозначим граф подстановки f [2]. Множество вершин графа G_f, лежащих на циклах длины $l \geq 1$, обозначим $C_l\left(G_f\right)$ [4].

Предложение 1. *Пусть* $x = x_0 \| \ldots \| x_{i-1} \in V_{ri}$, $i \in \mathbb{N}$, *и* $s_i \in C_l\left(G_f\right)$, $l > m$. *Тогда для входных сообщений* x *и* $x \| 0^{rlp}$, $p \in \mathbb{N}$, *внутренние состояния* s_i *и* s_{i+lp} *совпадают. Если при этом* $i = 1$, *то совпадают и состояния* s_0 *и* s_{lp}.

Доказательство. Для $i > 1$ имеем цепочку равенств:

$$s_{i+lp} = \varphi_{x_0,\ldots,x_{i-1},\underbrace{0,\ldots,0}_{lp}}\left(s_0\right) = \varphi_{\underbrace{0,\ldots,0}_{lp}} \circ \varphi_{x_0,\ldots,x_{i-1}}\left(s_0\right) = \varphi_{\underbrace{0,\ldots,0}_{lp}}\left(s_i\right) = f^{lp}\left(s_i\right) = s_i.$$

Для $i = 1$ утверждение следует из равенства $s_0 = f^{-1}\left(s_1\right)$. □

Следствие 1. *Если при условии предложения 1 выполняется равенство* $i = n$, *то выходные последовательности* $\psi_m\left(x\right)$ *и* $\psi_m\left(x \| 0^{lrp}\right)$ *совпадают и имеют период* l *(рис. 2).*

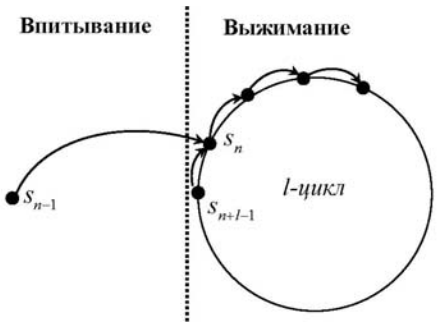

Рис. 2. Попадание внутреннего состояния Sponge на l -цикл

Предложение 2. *Пусть* $\alpha, \beta \in V_r$ *и вершины* $\alpha \| 0^c, \beta \| 0^c$ *лежат на одном цикле длины* l *в графе* G_f *на расстоянии* d *друг от друга. Тогда для произвольного* $\gamma \in V_{rp}$, $p \in \mathbb{N}$, *выполняются равенства:*

- *если* $d > 1$,

$$\psi_m\left(\alpha \| 0^{r(d-1+lp_1)} \| \gamma\right) = \psi_m\left(\beta \| 0^{r(l-1+lp_2)} \| \gamma\right) \text{ и } \psi_m\left(\alpha \| 0^{r(d+lp_1)} \| \gamma\right) = \psi_m\left(\beta \| 0^{rlp_2} \| \gamma\right),$$

- *если* $d = 0$,

$$\psi_m\left(\alpha \| 0^{r(l-1+lp_1)} \| \gamma\right) = \psi_m\left(\alpha \| 0^{r(l-1+lp_2)} \| \gamma\right),$$

где $p_1, p_2 \in \mathbb{N} \cup \{0\}$.

Заметим, что для применения результата предложения 2 достаточно сформировать хотя бы один цикл подстановки f на основе вершины из

множества $\left\{x \| 0^c : x \in V_r\right\}$. При этом трудоемкость данной процедуры будет не превосходить $\max\left(l_1,...,l_n\right)$ элементарных операций, где $\left[1^{l_1},...,n^{l_n}\right]$ - цикловая структура подстановки f.

Предложение 3. *Пусть* $x = x_0 \| ... \| x_{i-1} \in V_{ri}$, $y = y_0 \| ... \| x_{j-1} \in V_{rj}$, *и пусть* $\gamma \in V_{rp}$, $p \in \mathbb{N}$. *Тогда если* $s_i \in C_{l_1}\left(G_f\right)$, $s_j \in C_{l_2}\left(G_f\right)$ *и* $\left\lfloor s_i \oplus s_j \right\rfloor_c = 0^c$, *выполняется равенство:*

$$\psi_m\left(x_0 \| ... \| x_{i-1} \| 0^{rlp_1} \| x^* \| \gamma\right) = \psi_m\left(y_0 \| ... \| y_{j-1} \| 0^{rlp_2} \| \gamma\right),$$

где $x^* = \left\lceil s_i \oplus s_j \right\rceil_r$ *и* $p_1, p_2 \in \mathbb{N}$.

Доказательство. Для входных подвекторов $x \| 0^{rp_1} \| x^*$ и $y \| 0^{lrp_2}$ справедлива цепочка равенств:

$$\varphi_{x,\underbrace{0...,0}_{lp_1},x^*}\left(s_0\right) = \varphi_{\underbrace{0...,0}_{lp_1},x^*} \circ \varphi_{x_0,...,x_{i-1}}\left(s_0\right) = \varphi_{\underbrace{0...,0}_{lp_1},x^*}\left(s_i\right) = \varphi_{x^*}\left(f^{lp_1}\left(s_i\right)\right) = \varphi_{x^*}\left(s_i\right) =$$

$$= f\left(s_i \oplus \left(x^* \| 0^c\right)\right) = f\left(s_j\right) = \varphi_{\underbrace{0,...,0}_{lp_2}}\left(s_j\right) = \varphi_{\underbrace{0,...,0}_{lp_2}} \circ \varphi_{y_0,...,y_{j-1}}\left(s_0\right) = \varphi_{y,\underbrace{0,...,0}_{lp_2}}\left(s_0\right).$$

Таким образом, после впитывания сообщений $x \| 0^{rp_1} \| x^*$ и $y \| 0^{lrp_2}$ формируется одно и то же внутреннее состояние Sponge, а, следовательно, для любого $\gamma \in V_{rp}$ выполняется искомое равенство. □

Выводы

Из полученных результатов следует, что знание цикловой структуры подстановки f или ее части позволяет строить коллизии Sponge. В частности, конструктивным решением, исключающим возможность принудительного влияния на траекторию внутренних состояний Sponge за счёт вставки серий нулевых и специально подобранных блоков, является использование DuplexSponge [9] – модификации Sponge, в которой при впитывании очередного блока сообщения сразу вырабатывается выходной блок (рис. 3).

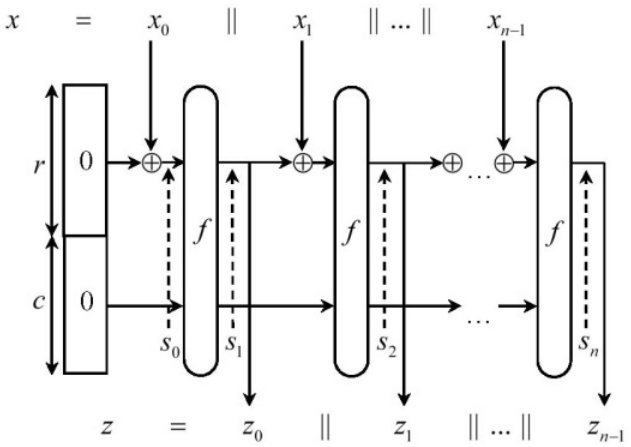

Рис. 3. Схема функционирования DuplexSponge

Однако такой способ хэширования накладывает существенные ограничения на длину хэш-кода, т.к. при каждом выжимании должен оставаться хотя бы один бит выходных данных.

Список литературы

1. *Колчин В.Ф.* Случайные отображения. – М., Наука, 1984.
2. *Колчин В.Ф.* Случайные графы (2-е изд.). – М.: ФИЗМАТЛИТ, 2004.
3. *Колчин В.Ф., Севастьянов Б.А., Чистяков В.П.* Случайные размещения. – М.: Наука, 1976.
4. *Сачков В.Н.* Вероятностные методы в комбинаторном анализе. – М.: Наука, 1978.
5. *Bernstein D. J.* Second preimages for 6 rounds of Keccak. NIST hash forum mailing list, November 27, 2010.
6. *Bertoni G., Daemen J., Peeters M., Van Assche G.* Sponge Functions, http://noekeon.org.
7. *Bertoni G., Daemen J., Peeters M., Van Assche G.* On the Indifferentiability of the Sponge Construction. www.iacr.org/cryptodb/data/paper.php?pubkey=15560.
8. *Bertoni G., Daemen J., Peeters M., Van Assche G.* The Keccak reference, http: // noekeon.org.
9. *Bertoni G., Daemen J., Peeters M., Van Assche G.* Duplexing the sponge: single-pass authenticated encryption and other applications. ePrint.iacr.org/2011/499.
10. *Dinur I., Dunkelman O., Shamir A.* New attacks on Keccak-224 and Keccak-256. Fast Software Encryption 2012.

11. *Gligoroski D., Odegard S.-R., Jensen R.-E.* OBSERVATION: An explicit form for a class of second preimages for any message M for the SHA-3 candidate Keccak. Eprint.iacr.org/2011/261.
12. *Morawiecki P., Pieprzyk J., Srebrny M.* Rotational cryptanalysis of round-reduced Keccak. ePrint.iacr.org/2012/546.
13. *Naya-Plasencia M., Rock A., Meier W.* Practical Analysis of Reduced-Round Keccak. Indocrypt 2011.
14. *Perrin L., Khovratovich D.* Collision Spectrum, Entropy Loss, T-Sponges, and Cryptanalysis of GLUON-64. ePrint.iacr.org/2014/223.