View metadata, citation and similar papers at core.ac.uk
brought to you by CORE
provided by Crossref

Article
Chinese Science Bulletin
Computer Science & Technology
February 2012 Vol.57 No.6: 694–697
doi: 10.1007/s11434-011-4909-x

Improved zero-sum distinguisher for full round *Keccak-f* permutation

DUAN Ming[1,2*] & LAI XueJia[1]

[1] *Department of Computer Science and Engineering, Shanghai Jiao Tong University, Shanghai 200240, China;*
[2] *Basic Courses Department, University of Foreign Language, Luoyang 471003, China*

Received August 2, 2011; accepted September 30, 2011

Keccak is one of the five hash functions selected for the final round of the SHA-3 competition, and its inner primitive is a permutation called *Keccak-f*. In this paper, we observe that for the inverse of the only nonlinear transformation in *Keccak-f*, the algebraic degree of any output coordinate and the one of the product of any two output coordinates are both 3, which is 2 less than its size of 5. Combining this observation with a proposition on the upper bound of the degree of iterated permutations, we improve the zero-sum distinguisher for the *Keccak-f* permutation with full 24 rounds by lowering the size of the zero-sum partition from 2^{1590} to 2^{1575}.

hash functions, higher order differentials, algebraic degree, zero-sum, SHA-3

Citation: Duan M, Lai X J. Improved zero-sum distinguisher for full round *Keccak-f* permutation. Chin Sci Bull, 2012, 57: 694–697, doi: 10.1007/s11434-011-4909-x

Zero-sum distinguishers, introduced by Aumasson and Meier and presented at the rump session of CHES 2009, are a method for generating zero-sum structures for iterated permutations, which combine higher order differential technique with inside-out technique and are mainly decided by the algebraic degree of the permutation. In the public comment on the NIST Hash competition 2010, zero-sum distinguishers are shown to be deterministic and valid, although they generate zero-sum structures with only a small advantage over the generic method. Zero-sum distinguishers can also be used to create partitions of inputs into many different zero-sum structures for the permutation [1].

Keccak is a family of cryptographic sponge functions and is one of the five hash functions selected for the third (and final) round of the SHA-3 competition. Its core component is a permutation named *Keccak-f*, which is composed of several iterations of five transformations. A first zero-sum distinguisher for the *Keccak-f* permutation with 16 rounds was given in 2009. Since then, zero-sum distinguishers for the *Keccak-f* permutation with a greater number of rounds were obtained [1,2], with the smallest known zero-sum partition for the *Keccak-f* permutation with full 24 rounds having size 2^{1590}.

In this paper, we study the properties of the inverse of the nonlinear transformation in *Keccak-f*, and observe that the algebraic degree of the product of any two output coordinates of this inverse is 2 less than its size. This enables us construct a zero-sum partition for the *Keccak-f* permutation with full 24 rounds of size 2^{1575}.

1 Zero-sum distinguishers

We first introduce the notions of higher order derivatives related to zero-sum distinguishers.

1.1 Higher order derivatives

Higher order derivatives were introduced into cryptography by Lai in [3].

*Corresponding author (email: dodoxixi@gmail.com)

© The Author(s) 2012. This article is published with open access at Springerlink.com

Definition 1. Let $f(x)$ be a Boolean function from F_2^n to F_2. The derivative of f at point $a \in F_2^n$ is defined by

$$\Delta_a f(x) = f(x+a) + f(x) .$$

The i-th (i>1) derivative of the function f at points $\{a_1, a_2, \ldots, a_i\}$ is defined by

$$\Delta_{a_1,\ldots,a_i}^{(i)} f(x) = \Delta_{a_i} \left(\Delta_{a_1,\ldots,a_{i-1}}^{(i-1)} f(x) \right),$$

where $\Delta_{a_1,\ldots,a_{i-1}}^{(i-1)} f(x)$ is the (i–1)-th derivative of f at points $\{a_1, a_2, \ldots, a_{i-1}\}$. The 0-th derivative of f is defined to be $f(x)$ itself.

Higher order derivatives should be computed at points that are linearly independent, otherwise the derivative will trivially be zero. Note that the degree of the derivative of a function is at least 1 less than the degree of the function. This implies that the (d+1)-th derivative of an n-variable Boolean function of degree d is zero, and this is used in many cryptanalysis methods including zero-sum distinguishers.

1.2 Zero-sum properties

Note that the permutation used in a hash function does not depend on a secret parameter, and this property of the permutation can be exploited from the middle. The zero-sum property introduced by Aumasson and Meier is based on higher order differential technique and inside-out technique. The main idea is to take higher order derivatives at initial states inverted from an intermediate internal state subspace, which differs from traditional higher order differential distinguishers that take derivatives directly from the initial state subspace. So zero-sum distinguishers lower the degree of higher order derivatives by nearly half with the added cost of some inverted computations.

We now give the definitions of zero-sum and zero-sum partitions. Further details can be found in [1].

Definition 2. Let F be a function from F_2^n into F_2^m. A *zero-sum* for F of size K is a subset $\{x_1, x_2, \ldots, x_K\} \subset F_2^n$ of elements which sum to zero and for which the corresponding images by F also sum to zero. That is,

$$\sum_{i=1}^{K} x_i = \sum_{i=1}^{K} F(x_i) = 0 .$$

Definition 3. Let P be a permutation from F_2^n into F_2^n. A *zero-sum partition* for F of size K=2^k is a collection of 2^{n-k} disjoint zero-sums $X_i=\{x_{i,1}, x_{i,1}, \ldots, x_{i,2^k}\} \subset F_2^n$. That is,

$$\bigcup_{i=1}^{2^{n-k}} X_i = F_2^n \quad and \quad \sum_{j=1}^{2^k} x_{i,j} = \sum_{j=1}^{2^k} P\left(x_{i,j}\right) = 0, \forall 1 \leqslant i \leqslant 2^{n-k} .$$

2 Description of the *Keccak-f* permutation

The size of *Keccak-f* is 1600, and the state can be represented by a 3-dimensional binary matrix of size 5×5×64. The five transformations are respectively called θ, ρ, π, ι and χ. Only the transformation χ is nonlinear, its degree being 2 while the degree of its inverse is 3. The Boolean components of χ are listed in Table 1. More details of the *Keccak-f* permutation are available in the website of the NIST Hash competition.

3 Generalized and intuitive upper bound of the degree of iterated permutations

High algebraic degree is an important design principle for cryptographic algorithms. It is difficult to determine the algebraic degree when the number of rounds in the algorithm is too big. Estimating the upper bound on the algebraic degree is relatively feasible. In [4], Canteaut and Videau gave an upper bound on the degree of composition of nonlinear functions and used it to estimate the algebraic degree of the whole algorithm. In the rump session of Crypto 2010, Boura et al. [2] proposed an improved upper bound for iterated permutations with a nonlinear layer composed of parallel applications of small balanced S-boxes. We next discuss this latter upper bound and give a proposition for a visualized bound in some cases.

Theorem 1[2]. Let f be a function from F_2^n into F_2^n corresponding to the concatenation of m smaller balanced S-boxes, $S_1, \ldots, S_m$, defined over F_2^n. Let δ_k be the maximal degree of the product of any k coordinates from any one of these smaller S-boxes. Then, for any function G from F_2^n into F_2^l, we have

$$\deg\left(G \circ F\right) \leqslant n - \frac{n - \deg(G)}{\gamma},$$

where

$$\gamma = \max_{1 \leqslant i \leqslant n_0 - 1} \frac{n_0 - i}{n_0 - \delta_i} .$$

Most notably, we have

Table 1 Boolean components of χ

Output	Corresponding Boolean function
χ_0	$x_0+x_2+x_1x_2$
χ_1	$x_1+x_3+x_2x_3$
χ_2	$x_2+x_4+x_3x_4$
χ_3	$x_0+x_3+x_0x_4$
χ_4	$x_1+x_4+x_0x_1$

$$\deg\left(G \circ F\right) \leqslant n - \frac{n - \deg(G)}{n_0 - 1}.$$

Moreover, if $n_0 \geqslant 3$ and all S-boxes have degree at most $n_0 - 2$, we have

$$\deg\left(G \circ F\right) \leqslant n - \frac{n - \deg(G)}{n_0 - 2}.$$

Lemma 1. Let f be a function from F_2^n into F_2^n corresponding to the concatenation of m smaller S-boxes, S_1, ..., S_m, defined over $F_2^{n_0}$. Let δ_k be the maximal degree of the product of any k coordinates of any one of these smaller S-boxes. If $n_0 \geqslant 2k-1 (k \geqslant 1)$ and $\delta_i \leqslant n_0 - 1$ for any i from 1 to $n_0 - 1$, and $\delta_i \leqslant n_0 - 2$ for any i from 1 to $k-1$ $(k \geqslant 2)$, then

$$(n_0 - k)(n_0 - \delta_i) - (n_0 - i) \geqslant 0$$

for any i from 1 to $n_0 - 1$.

Proof: When $k = 1$, then we have

$$\begin{aligned}(n_0 - k)(n_0 - \delta_i) - (n_0 - i) &= (n_0 - 1)(n_0 - \delta_i) - (n_0 - i) \\ &\geqslant (n_0 - 1) - (n_0 - i) \\ &\geqslant (n_0 - 1) - (n_0 - 1) = 0.\end{aligned}$$

When $k \geqslant 2$, then we have

$$\begin{aligned}(n_0 - k)(n_0 - \delta_i) - (n_0 - i) &\geqslant 2(n_0 - k) - (n_0 - i) \\ &\geqslant 2(n_0 - k) - (n_0 - 1) \\ &\geqslant n_0 - 2k + 1 \geqslant 0.\end{aligned}$$

First, from Theorem 1, we know that the condition that the S-boxes are balanced in confirms that the inequality $\delta_i \leqslant n_0 - 1$ is satisfied for any i from 1 to $n_0 - 1$. This is not a necessary condition, however, and it is not necessary to limit the condition to balanced S-boxes. That is, the condition in Theorem 1 can be generalized.

Second, note that the parameter γ in the theorem is the maximum value of $\frac{n_0 - i}{n_0 - \delta_i}$ for i from 1 to $n_0 - 1$. Lemma 1 tells us that the positive integer $n_0 - k$ also suffices in some cases. That is, the inequality $n_0 - k \geqslant \frac{n_0 - i}{n_0 - \delta_i}$ always holds under the conditions of the lemma, so we get $\gamma \leqslant (n_0 - k)$.

With the more generalized condition and more determinate parameter, we have the following intuitive upper bound by combining the theorem and the lemma.

Proposition 1. Let f be a function from F_2^n into F_2^n corresponding to the concatenation of m smaller S-boxes, S_1, ..., S_m, defined over $F_2^{n_0}$. Let δ_k be the maximal degree of the product of any k coordinates of any one of these smaller S-boxes. If $n_0 \geqslant 2k-1 (k \geqslant 1)$ and $\delta_i \leqslant n_0 - 1$ for any i from 1 to $n_0 - 1$, and $\delta_i \leqslant n_0 - 2$ for any i from 1 to $k-1$ $(k \geqslant 2)$, then, for any function G from F_2^n into F_2^l, we have

$$\deg(G \circ F) \leqslant n - \frac{n - \deg(G)}{n_0 - k}.$$

Actually, when the conditions of Proposition 1 are satisfied and n_0 is an even number, then the parameter γ can be improved to $n_0 - k - 1/2$, but a discussion of this is not relevant to this paper.

4 Improved zero-sum distinguisher for *Keccak-f*

4.1 An observation about *Keccak-f*

We give the Boolean components of χ^{-1} and the product of any two output coordinates of the transformation in Tables 2 and 3 respectively.

From Table 3, an interesting observation about the inverse of the nonlinear layer of *Keccak-f* can be obtained.

Observation: For the inverse of the only nonlinear transformation in *Keccak-f*, the algebraic degree of any output coordinate and the one of the product of any two output coordinates are both 3, which is 2 less than its size of 5.

4.2 Improved zero-sum partition for full 24-rounds *Keccak-f* permutation

Let R denote the *Keccak-f* round permutation. Note that χ

Table 2 Boolean components of χ^{-1}

Output	Corresponding Boolean function
χ_0^{-1}	$x_0+x_2+x_4+x_1x_2+x_1x_4+x_3x_4+x_1x_3x_4$
χ_1^{-1}	$x_0+x_1+x_3+x_0x_2+x_0x_4+x_2x_3+x_0x_2x_4$
χ_2^{-1}	$x_1+x_2+x_4+x_0x_1+x_1x_3+x_3x_4+x_0x_1x_3$
χ_3^{-1}	$x_0+x_2+x_3+x_0x_4+x_1x_2+x_2x_4+x_1x_2x_4$
χ_4^{-1}	$x_1+x_3+x_4+x_0x_1+x_0x_3+x_2x_3+x_0x_2x_3$

Table 3 Product of any two output coordinates of χ^{-1}

Output	Corresponding Boolean function
$\chi_0^{-1}\chi_1^{-1}$	$x_0+x_0x_1+x_0x_2+x_0x_3+x_0x_4+x_0x_2x_3+x_0x_2x_4$
$\chi_0^{-1}\chi_2^{-1}$	$x_2+x_4+x_0x_2+x_0x_4+x_1x_2+x_1x_4+x_3x_4+x_0x_3x_4+x_1x_3x_4$
$\chi_0^{-1}\chi_3^{-1}$	$x_0+x_2+x_0x_3+x_0x_4+x_1x_2+x_2x_3+x_2x_4+x_3x_4+x_1x_2x_3+x_1x_2x_4$
$\chi_0^{-1}\chi_4^{-1}$	$x_4+x_0x_3+x_0x_4+x_1x_4+x_2x_4+x_1x_2x_4+x_1x_3x_4$
$\chi_1^{-1}\chi_2^{-1}$	$x_1+x_0x_1+x_1x_2+x_1x_3+x_1x_4+x_0x_1x_3+x_1x_3x_4$
$\chi_1^{-1}\chi_3^{-1}$	$x_0+x_3+x_0x_1+x_0x_2+x_0x_4+x_1x_3+x_2x_3+x_0x_1x_4+x_0x_2x_4$
$\chi_1^{-1}\chi_4^{-1}$	$x_1+x_3+x_0x_1+x_0x_3+x_1x_4+x_2x_3+x_3x_4+x_0x_2x_3+x_2x_3x_4$
$\chi_2^{-1}\chi_3^{-1}$	$x_2+x_0x_2+x_1x_2+x_2x_3+x_2x_4+x_0x_2x_4+x_1x_2x_4$
$\chi_2^{-1}\chi_4^{-1}$	$x_1+x_4+x_0x_1+x_1x_2+x_1x_3+x_2x_4+x_3x_4+x_0x_1x_2+x_0x_1x_3+x_0x_3x_4+x_2x_3x_4$
$\chi_3^{-1}\chi_4^{-1}$	$x_3+x_0x_3+x_1x_3+x_2x_3+x_3x_4+x_0x_1x_3+x_0x_2x_3$

is the only nonlinear transformation in R. Combining our earlier observation and Proposition 1, we have

$$\deg\left(G\circ R\right)=\deg\left(G\circ\chi\right)\leqslant n-\frac{n-\deg(G)}{3}$$

and

$$\deg\left(G\circ R^{-1}\right)=\deg\left(G\circ\chi^{-1}\right)\leqslant n-\frac{n-\deg(G)}{2},$$

where G is any function from F_2^5 into F_2^l. Our upper bounds on the degree of the inverse of *Keccak-f* are less than the bounds in [2] when the number of rounds is more than seven. The comparisons are listed in Table 4.

Combining these upper bounds on deg(R^{γ}) with those in [2] and our lowered upper bounds on deg($R^{-\gamma}$), we have a zero-sum partition of size 2^{1575} for the full *Keccak-f* permutation. This is smaller than the original size of 2^{1590}, as confirmed in the updated version of [2] appearing in the Preproceedings of FSE 2011. Indeed, one can consider the intermediate states after the three linear layers θ, ρ and π, in the 12-th round of *Keccak-f* in any subspace V corresponding to a collection of 315 rows, because the upper bound of the backward 11 rounds is 1572 and that of the forward 12 rounds is 1536 [2].

Table 4 Comparison of the upper bounds on deg($R^{-\gamma}$)

Round	Bound in [2]	Our bound
1	3	3
2	9	9
3	27	27
4	81	81
5	243	243
6	729	729
7	1309	1164
8	1503	1382
9	1567	1491
10	1589	1545
11	1596	1572
12	1598	1586
13	1599	1593
14	1599	1596
15	1599	1598
16	1599	1599

5 Discussion

In this paper, we lower the size of a zero-sum partition for the *Keccak-f* permutation with full 24 rounds based on an interesting observation about the inverse of the nonlinear transformation in the permutation. One can verify that some of the products of three output coordinates also have a degree of only 3. This property may be used for more practical cryptanalysis of *Keccak* in the future.

The authors express their great thanks to C. Boura and the anonymous reviewers for their helpful comments. This work was supported by the National Natural Science Foundation of China (*60573032, 60773092 and 61073149*)*, and Research Fund for the Doctoral Program of Higher Education of China* (*20090073110027*).

1 Boura C, Canteaut A. Zero-sum distinguishers for iterated permutations and application to *Keccak-f* and Hamsi-256. In: Proceedings of the 17th International Workshop on Selected Areas in Cryptography 2010 Aug 12–13, Waterloo, Ontario, Canada. Waterloo: LNCS Springer Press, 2010. 1–17

2 Boura C, Canteaut A, Cannière C D. Higher-order differential properties of *Keccak* and Luffa, In: Proceedings of the 18th International Workshop on Fast Software Encryption 2011 Feb 14–16, Lyngby, Denmark. Lyngby: LNCS Springer Press, 2011. 252–269

3 Lai X J. Higher order derivatives and differential cryptanalysis. In: Communications and Cryptography: Two Sides of One Tapestry 1994, Switzerland. Switzerland: Kluwer Adademic Publishers, 1994. 227–233

4 Canteaut A,Videau M. Degree of composition of highly nonlinear functions and applications to higher order differential cryptanalysis. In: Advances in Cryptology-EUROCRYPT 2002, International Conference on the Theory and Applications of Cryptographic Techniques, Apr 28–May 2, Amsterdam, The Netherlands. Amsterdam: LNCS Springer Press, 2002. 518–533

Open Access This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.