Cryptogr. Commun. (2013) 5:1–6
DOI 10.1007/s12095-012-0063-9

Comments on "A design of Boolean functions resistant to (fast) algebraic cryptanalysis with efficient implementation"

Wenhao Wang · Meicheng Liu · Yin Zhang

Received: 31 December 2011 / Accepted: 4 July 2012 / Published online: 19 July 2012

Abstract In this correspondence, it is shown that the Boolean functions constructed by Pasalic (Cryptogr Commun 4(1):25–45, 2012) do not always have the high degree product of order $n-1$ as expected.

Keywords Boolean functions · High degree product · Fast algebraic attacks

Mathematics Subject Classifications (2010) 11T55 · 11T71

Introduction A Boolean function on n variables is a mapping from $\mathbb{F}_2^n$ to $\mathbb{F}_2$. Denote the set of all n-variable Boolean functions by $\mathcal{B}_n$. Any $f \in \mathcal{B}_n$ can be uniquely represented as a multivariate polynomial over $\mathbb{F}_2$, called the algebraic normal form (ANF), as

$$f(x_1, \cdots, x_n) = \sum_{u \in \mathbb{F}_2^n} \lambda_u \prod_{i=1}^{n} x_i^{u_i}, \ \lambda_u \in \mathbb{F}_2, u = (u_1, \cdots, u_n).$$

The algebraic degree of f, denoted by $\deg(f)$, is the maximal value of the Hamming weight of u such that $\lambda_u \neq 0$.

Supported by the National 973 Program of China under Grant 2011CB302400, the National Natural Science Foundation of China under Grants 10971246, 60970152, and 61173134, the Grand Project of Institute of Software of CAS under Grant YOCX285056.

W. Wang · M. Liu (✉) · Y. Zhang
The State Key Laboratory of Information Security, Institute of Information Engineering, Chinese Academy of Sciences, Beijing 100195, People's Republic of China
e-mail: meicheng.liu@gmail.com

W. Wang
e-mail: wangwenhao@is.iscas.ac.cn

Y. Zhang
e-mail: zhangy@is.iscas.ac.cn

A preprocessing of fast algebraic attacks on linear feedback shift register based stream ciphers, which use a Boolean function f as the filter or combination generator, is to find a function g of small degree such that the multiple gf has reasonable degree. In [1], Pasalic introduced the notion of high degree product ($\mathcal{HDP}$) to scale the ability of Boolean functions resistant to fast algebraic attacks. A Boolean function $f \in \mathcal{B}_n$ satisfies the $\mathcal{HDP}$ of order n if for any non-annihilating function g of degree e, $1 \le e \le \lceil \frac{n}{2} \rceil - 1$, we necessarily have that $d = \deg(gf)$ satisfies $e + d \ge n$. The author presented an iterative construction of Boolean functions with almost optimal $\mathcal{HDP}$, that is, the $\mathcal{HDP}$ of order $n-1$. In this letter, it is shown that the constructed functions do not always achieve desired properties. First we point out that there is a flaw in the proof of [1, Theorem 4], which is used to construct functions with almost optimal $\mathcal{HDP}$. Then we examine the example given by the author, and it turns out that some of the constructed functions do not satisfy the $\mathcal{HDP}$ of order $n-1$ as claimed.

Review of Pasalic's construction A Boolean function $f \in \mathcal{B}_{n+2}$ can be considered as a concatenation of four functions, denoted by $f = f_1||f_2||f_3||f_4$ with $f_i \in \mathcal{B}_n$. The ANF of f is given by

$$f = x_{n+1}x_{n+2}(f_1 + f_2 + f_3 + f_4) + x_{n+1}(f_1 + f_2) + x_{n+2}(f_1 + f_3) + f_1. \tag{1}$$

The iterative construction is described as follows,

$$\begin{aligned} f_1^i &= f_1^{i-1}||f_2^{i-1}||1 + f_1^{i-1}||f_3^{i-1}, \\ f_2^i &= f_2^{i-1}||1 + f_3^{i-1}||f_1^{i-1}||1 + f_2^{i-1}, \\ f_3^i &= 1 + f_3^{i-1}||f_1^{i-1}||f_2^{i-1}||f_3^{i-1}, \end{aligned} \tag{2}$$

where $f_1^0, f_2^0, f_3^0 \in \mathcal{B}_n$ are initial functions and $f_1^i, f_2^i, f_3^i \in \mathcal{B}_{n+2i}$ the constructed functions.

Statement 1 [1, Theorem 4][1] *Let $f_1^0, f_2^0, f_3^0 \in \mathcal{B}_n$ and for any $g = g_1^0||g_2^0||g_3^0||g_4^0 \in \mathcal{B}_{n+2}$ of degree $e \in [1, \lceil \frac{n}{2} \rceil - 1]$ the following is satisfied,*

$$\deg\left[f_1^0\left(\sum_{j=1}^{4} b_j g_j^0\right) + f_2^0\left(\sum_{j=1}^{4} c_j g_j^0\right) + f_3^0\left(\sum_{j=1}^{4} d_j g_j^0\right)\right] \ge n - e - 1, b_j, c_j, d_j \in \mathbb{F}_2. \tag{3}$$

Then the functions $f_j^i \in \mathcal{B}_{n+2i}$, $i \ge 0$ and $j = 1, 2, 3$, defined by (2), have almost optimal $\mathcal{HDP}$, that is satisfying $e + d \ge n + 2i - 1$ for $e \in [1, \lceil \frac{n}{2} \rceil + i - 1]$.

Let $g^{i+1} = g_1^i||g_2^i||g_3^i||g_4^i \in \mathcal{B}_{n+2i+2}$, $\deg(g^{i+1}) = e$ and

$$\mu_e^i = \deg\left[f_1^i\left(\sum_{j=1}^{4} b_j g_j^i\right) + f_2^i\left(\sum_{j=1}^{4} c_j g_j^i\right) + f_3^i\left(\sum_{j=1}^{4} d_j g_j^i\right)\right], b_j, c_j, d_j \in \mathbb{F}_2.$$

[1] Here is omitted from [1] that the functions f_1^0, f_2^0, f_3^0 achieve maximum algebraic immunity since it does not influence the $\mathcal{HDP}$ properties of the constructed functions.

In [1], the proof of the above statement was presented by induction for

$$\mu_e^i \geq n + 2i - e - 1, e \in \left[1, \left\lceil \frac{n}{2} \right\rceil + i - 1\right] \tag{4}$$

which implies the functions f_j^i have almost optimal $\mathcal{HDP}$. The case $i = 0$ follows directly from (3). Suppose the conditions are satisfied for all $k < i$, that is, for any $g^{k+1} = g_1^k||g_2^k||g_3^k||g_4^k \in \mathcal{B}_{n+2k+2}$ of degree $e \in [1, \lceil \frac{n}{2} \rceil + k - 1]$, it holds that $\mu_e^k \geq n + 2k - e - 1$ (which was misprinted in [1] as $\mu_e^{k-1} \geq n + 2k - e - 1$). Then it needs to show the conditions hold for $k + 1$ as well. Considering the function $f_1^{k+1} = f_1^k||f_2^k||1 + f_1^k||f_3^k \in \mathcal{B}_{n+2k+2}$ and a degree e function $g^{k+1} \in \mathcal{B}_{n+2k+2}$, it is necessary that $\deg(f_1^{k+1}g^{k+1}) \geq n + 2k - e + 1$ for any $e \in [1, \lceil \frac{n}{2} \rceil + k]$. The author focused on the following term in the product $f_1^{k+1}g^{k+1}$,

$$x_{n+2k+1}x_{n+2k+2}\left[g_3^k + f_4^k g_4^k + f_1^k(g_1^k + g_3^k) + f_2^k g_2^k\right],$$

and claimed that

$$\deg\left[f_4^k g_4^k + f_1^k\left(g_1^k + g_3^k\right) + f_2^k g_2^k\right] \geq n + 2k - e - 1 \tag{5}$$

according to (4). Note that (4) holds for $e \in [1, \lceil \frac{n}{2} \rceil + k - 1]$ but not necessarily for $e = \lceil \frac{n}{2} \rceil + k$. Therefore (5) may not hold, then the function $f_j^i \in \mathcal{B}_{n+2i}$ may admit a function g of degree $\lceil \frac{n}{2} \rceil + k$ for $k < i$ such that $\deg(gf_j^i) \leq n + 2i - \lceil \frac{n}{2} \rceil - k - 2$, i.e., the function may not achieve almost optimal $\mathcal{HDP}$. In particular, the function f_j^i may admit a function g of degree $\lceil \frac{n}{2} \rceil$ such that $\deg(gf_j^i) \leq n + 2i - \lceil \frac{n}{2} \rceil - 2$. For example, when $n = 4$, the 10-variable function f_2^3 may admit a function g of degree 2 such that $\deg(gf_2^3) \leq 6$.

Observation on the constructed functions For $i \geq 2$, according to (2) it holds that

$$\begin{aligned}
f_1^{i-1} &= f_1^{i-2}||f_2^{i-2}||1 + f_1^{i-2}||f_3^{i-2},\\
f_2^{i-1} &= f_2^{i-2}||1 + f_3^{i-2}||f_1^{i-2}||1 + f_2^{i-2},\\
f_3^{i-1} &= 1 + f_3^{i-2}||f_1^{i-2}||f_2^{i-2}||f_3^{i-2},
\end{aligned}$$

and therefore by (1) we have

$$\begin{aligned}
f_1^{i-1} &= x_{n+2i-2}x_{n+2i-3}(f_2^{i-2} + f_3^{i-2} + 1) + x_{n+2i-3}(f_1^{i-2} + f_2^{i-2}) + x_{n+2i-2} + f_1^{i-2},\\
f_2^{i-1} &= x_{n+2i-2}x_{n+2i-3}(f_1^{i-2} + f_3^{i-2}) + x_{n+2i-3}(f_2^{i-2} + f_3^{i-2} + 1)\\
&\quad + x_{n+2i-2}(f_1^{i-2} + f_2^{i-2}) + f_2^{i-2},\\
f_3^{i-1} &= x_{n+2i-2}x_{n+2i-3}(f_1^{i-2} + f_2^{i-2} + 1) + x_{n+2i-3}(f_1^{i-2} + f_3^{i-2} + 1)\\
&\quad + x_{n+2i-2}(f_2^{i-2} + f_3^{i-2} + 1) + f_3^{i-2} + 1.
\end{aligned}$$

Furthermore we represent f_2^i by f_1^{i-2}, f_2^{i-2} and f_3^{i-2},

$$\begin{aligned}
f_2^i &= x_{n+2i}x_{n+2i-1}(f_1^{i-1} + f_3^{i-1}) + x_{n+2i-1}(f_2^{i-1} + f_3^{i-1} + 1) \\
&\quad + x_{n+2i}(f_1^{i-1} + f_2^{i-1}) + f_2^{i-1} \\
&= x_{n+2i}x_{n+2i-1}x_{n+2i-2}x_{n+2i-3}(f_1^{i-2} + f_3^{i-2}) \\
&\quad + x_{n+2i}x_{n+2i-1}x_{n+2i-2}(f_2^{i-2} + f_3^{i-2}) \\
&\quad + x_{n+2i}x_{n+2i-1}x_{n+2i-3}(f_2^{i-2} + f_3^{i-2} + 1) \\
&\quad + x_{n+2i}x_{n+2i-1}(f_1^{i-2} + f_3^{i-2} + 1) \\
&\quad + x_{n+2i}x_{n+2i-2}x_{n+2i-3}(f_1^{i-2} + f_2^{i-2} + 1) \\
&\quad + x_{n+2i}x_{n+2i-2}(f_1^{i-2} + f_2^{i-2} + 1) \\
&\quad + x_{n+2i}x_{n+2i-3}(f_1^{i-2} + f_3^{i-2} + 1) \\
&\quad + x_{n+2i}(f_1^{i-2} + f_2^{i-2}) \\
&\quad + x_{n+2i-1}x_{n+2i-2}x_{n+2i-3}(f_2^{i-2} + f_3^{i-2} + 1) \\
&\quad + x_{n+2i-1}x_{n+2i-2}(f_1^{i-2} + f_3^{i-2} + 1) \\
&\quad + x_{n+2i-1}x_{n+2i-3}(f_1^{i-2} + f_2^{i-2}) \\
&\quad + x_{n+2i-1}(f_2^{i-2} + f_3^{i-2}) \\
&\quad + x_{n+2i-2}x_{n+2i-3}(f_1^{i-2} + f_3^{i-2}) \\
&\quad + x_{n+2i-2}(f_1^{i-2} + f_2^{i-2}) \\
&\quad + x_{n+2i-3}(f_2^{i-2} + f_3^{i-2} + 1) \\
&\quad + f_2^{i-2}.
\end{aligned}$$

Let

$$g = (x_{n+2i-3} + x_{n+2i-1})(x_{n+2i-2} + x_{n+2i}),$$

then we calculate that[2]

$$\begin{aligned}
g(f_2^i + f_2^{i-2}) &= x_{n+2i}x_{n+2i-1}x_{n+2i-2}x_{n+2i-3} + x_{n+2i}x_{n+2i-1}x_{n+2i-3} \\
&\quad + x_{n+2i}x_{n+2i-1} + x_{n+2i}x_{n+2i-2}x_{n+2i-3} + x_{n+2i-1}x_{n+2i-2} + x_{n+2i-2}x_{n+2i-3},
\end{aligned} \tag{6}$$

which has degree 4. Therefore we have

$$\begin{aligned}
gf_2^i &= gf_2^{i-2} + x_{n+2i}x_{n+2i-1}x_{n+2i-2}x_{n+2i-3} + x_{n+2i}x_{n+2i-1}x_{n+2i-3} \\
&\quad + x_{n+2i}x_{n+2i-1} + x_{n+2i}x_{n+2i-2}x_{n+2i-3} + x_{n+2i-1}x_{n+2i-2} + x_{n+2i-2}x_{n+2i-3},
\end{aligned} \tag{7}$$

[2]This can be examined in Magma, see also Appendix for the Magma source codes.

and

$$e = \deg(g) = 2,$$

$$d = \deg(gf_2^i) = \max\{\deg(gf_2^{i-2}), 4\} = \max\{\deg(f_2^{i-2}) + 2, 4\}.$$

For $n + 2i \geq 7$, if f_2^{i-2} is a balanced function, which implies $\deg(f_2^{i-2}) \leq n + 2i - 5$, then $e + d \leq n + 2i - 1$ and f_2^i never achieves the optimal $\mathcal{HDP}$. For $n + 2i \geq 8$, if $\deg(f_2^{i-2}) \leq n + 2i - 6$, then $e + d \leq n + 2i - 2$ and f_2^i does not have almost optimal $\mathcal{HDP}$.

For $i \geq 2$, let

$$g' = x_{n+2i-3}(x_{n+2i-2} + x_{n+2i-1} + x_{n+2i} + 1)$$

and

$$g'' = (x_{n+2i-3} + 1)(x_{n+2i-1} + 1).$$

Similarly to (7), we can obtain that

$$\begin{aligned} g' f_2^i &= g' f_1^{i-2} + x_{n+2i}x_{n+2i-1}x_{n+2i-3} + x_{n+2i}x_{n+2i-2}x_{n+2i-3} + x_{n+2i}x_{n+2i-3} \\ &\quad + x_{n+2i-1}x_{n+2i-3} + x_{n+2i-2}x_{n+2i-3} + x_{n+2i-3}, \\ g'' f_1^i &= g'' f_3^{i-2} + x_{n+2i}x_{n+2i-1}x_{n+2i-3} + x_{n+2i}x_{n+2i-1} + x_{n+2i}x_{n+2i-3} + x_{n+2i} \\ &\quad + x_{n+2i-1}x_{n+2i-2}x_{n+2i-3} + x_{n+2i-1}x_{n+2i-2} + x_{n+2i-2}x_{n+2i-3} + x_{n+2i-2}. \end{aligned}$$

The above equations and (7) show that f_1^i or f_2^i has not almost optimal $\mathcal{HDP}$ if one of the functions f_1^{i-2}, f_2^{i-2}, f_3^{i-2} has degree at most $n + 2i - 6$.

Example Hereinafter is an example in [1] of the initial functions with $n = 4$.

$$\begin{aligned} f_1^0 &= x_1 + x_1x_2 + x_3x_4 + x_1x_2x_3 + x_1x_2x_3x_4, \\ f_2^0 &= x_2 + x_4 + x_1x_2 + x_2x_4 + x_3x_4 + x_1x_2x_3 + x_1x_3x_4 + x_2x_3x_4 + x_1x_2x_3x_4, \\ f_3^0 &= x_2 + x_3 + x_1x_2 + x_2x_3 + x_3x_4 + x_1x_2x_3 + x_1x_2x_3x_4. \end{aligned}$$

We verify that the above functions satisfy relation (3). From (1) and (2), we have

$$\begin{aligned} f_2^1 &= x_5x_6(f_1^0 + f_3^0) + x_5(1 + f_2^0 + f_3^0) + x_6(f_1^0 + f_2^0) + f_2^0 \\ &= x_1x_2x_3x_4 + x_1x_2x_3 + x_1x_2 + x_1x_3x_4x_5 + x_1x_3x_4x_6 + x_1x_3x_4 + x_1x_5x_6 \\ &\quad + x_1x_6 + x_2x_3x_4x_5 + x_2x_3x_4x_6 + x_2x_3x_4 + x_2x_3x_5x_6 + x_2x_3x_5 + x_2x_4x_5 \\ &\quad + x_2x_4x_6 + x_2x_4 + x_2x_5x_6 + x_2x_6 + x_2 + x_3x_4 + x_3x_5x_6 + x_3x_5 + x_4x_5 \\ &\quad + x_4x_6 + x_4 + x_5, \end{aligned}$$

and therefore $\deg(f_2^1) = 4$. Let $g = (x_7 + x_9)(x_8 + x_{10})$, then it follows from (7) that

$$gf_2^3 = gf_2^1 + x_7x_8x_9x_{10} + x_7x_8x_{10} + x_7x_8 + x_7x_9x_{10} + x_8x_9 + x_9x_{10},$$

where g has degree 2 and gf_2^3 has degree 6. This shows that the 10-variable function f_2^3 has not the $\mathcal{HDP}$ of order 9.

As a matter of fact, the degree of the $2i$-variable function f_2^{i-2} equals to $2i-2$ by our computational experiment for $3 \leq i \leq 12$. Then, as mentioned previously, the function f_2^i $(3 \leq i \leq 12)$ has not almost optimal $\mathcal{HDP}$. We also examine the functions f_1^i and f_3^i with i up to 6. It turns out that $f_3^5 \in \mathcal{B}_{14}$ admits $e+d=12$ for $e=4$ and $f_1^6, f_3^6 \in \mathcal{B}_{16}$ admit $e+d=14$ for $e=4$.

Conclusion The functions constructed by (2) are not always balanced functions with the $\mathcal{HDP}$ of order n whatever initial functions are. Yet the constructed functions do not always achieve the $\mathcal{HDP}$ of order $n-1$ even though the initial functions satisfy the condition (3). This raises the question[3] whether these functions have the $\mathcal{HDP}$ of order $n-2$. We check the constructed functions on 8, 10, 12, 14 variables for dozens of initial functions which satisfy (3), and no function is found to have the $\mathcal{HDP}$ of order $< n-2$. Iterative construction of (almost) optimal Boolean functions resistant to fast algebraic attacks seems to be a challenge, since it seems very difficult to ensure the lower bound of $e+d$ from $\mathcal{B}_n$ to $\mathcal{B}_{n+2}$ for every n.

Acknowledgements The authors thank the anonymous referees for their valuable comments and suggestions on improving the manuscript. The authors also thank Tao Shi and Tianze Wang for helpful discussions, and seminar participants at SKLOIS: Shaoyu Du, Lin Jiao, Yao Lu, Wenlun Pan, et. al.

Appendix: Magma codes

```
P<[x]>:=PolynomialRing(GF(2),7);
Q<x1,x2,x3,x4,f1,f2,f3>:=quo<P|[x[i]^2-x[i]:i in [1..7]]>;
x:=[x1,x2,x3,x4];
f:=[f1,f2,f3];
for i:=1 to 2 do
    n:=2*i;
    tp1:=(f[2]+f[3]+1)*x[n-1]*x[n]
         +(f[1]+f[2])*x[n-1]+x[n]+f[1];
    tp2:=(f[1]+f[3])*x[n-1]*x[n]+(f[2]+f[3]+1)*x[n-1]
         +(f[1]+f[2])*x[n]+f[2];
    tp3:=(f[2]+f[1]+1)*x[n-1]*x[n]+(f[1]+f[3]+1)
         *x[n-1]+(f[2]+f[3]+1)*x[n]+f[3]+1;
    f:=[tp1,tp2,tp3];
end for;
(x[1]+x[3])*(x[2]+x[4])*(f2+f[2]);
x[1]*(x[2]+x[3]+x[4]+1)*(f3+f[2]);
(x[1]+1)*(x[3]+1)*(f1+f[1]);
```

Reference

1. Pasalic, E.: A design of Boolean functions resistant to (fast) algebraic cryptanalysis with efficient implementation. Cryptogr. Commun. **4**(1), 25–45 (2012)

[3]This question is suggested by one of the anonymous reviewers.