June 2003

Construction of highly non-linear cubic homogeneous Boolean functions on GF2n+l (2)

Jing Wu
University of Wollongong, jw91@uow.edu.au

Tianbing Xia
University of Wollongong, txia@uow.edu.au

Jennifer Seberry
University of Wollongong, jennie@uow.edu.au

Construction of highly non-linear cubic homogeneous Boolean functions on GF2n+l (2)

Abstract

The work studies highly nonlinear Boolean functions in GF2n+1(2), i.e. for the dimensions where bent functions do not exist. We prove that for every n > 2 there exist homogeneous Boolean functions on GF(2)2n+1 with non-linearity greater than or equal to 22n — 2n and without linear structures.

Keywords

Homogeneous, Bent, High non-linearity.

Disciplines

Physical Sciences and Mathematics

Publication Details

This article was originally published as Construction of highly non-linear cubic homogeneous Boolean functions on GF2n+l (2), in Arabnia, HR, Mun, Y and Aissi, S (eds), Proceedings of the 2003 International Conference on Security and management (SAM'03), Las Vegas, 23-26 June 2003, 241-247.

Construction of highly non-linear cubic homogeneous Boolean functions on $GF^{2n+1}(2)$

Jing Wu, Tianbing Xia * and Jennifer Seberry *
* School of IT and CS
University of Wollongong
Wollongong, NSW 2522
Australia

Abstract

The work studies highly nonlinear Boolean functions in $GF^{2n+1}(2)$, i.e. for the dimensions where bent functions do not exist. We prove that for every $n \geq 2$ there exist homogeneous Boolean functions on $GF(2)^{2n+1}$ with non-linearity greater than or equal to $2^{2n} - 2^n$ and without linear structures.

Key Words: Homogeneous, Bent, High non-linearity.

1 Introduction

Homogeneity is an important property when cryptographic algorithms use the Feistel structure such as these applied in MD4 and MD5 hashing algorithms or in the DES encryption algorithm. It was showed in [2] that homogeneous functions permit to re-use evaluations from previous iterations if the Feistel structure contains $n \geq 4$ wires (inputs/outputs) and the rotation is used as the diffusion P-box. These Boolean functions create a class of rotation-symmetric functions. An important property of rotation-symmetric functions is that they can be decomposed into one or more homogeneous parts. To keep a round function $f(x)$ short, one would prefer a homogeneous rotation-symmetric function.

It turns out [5] that homogeneous bent functions of degree n do not exist on $GF^{2n}(2)$ when $n > 3$. However, cubic homogeneous bent functions on $GF^{2n}(2)$ exist for all $n \geq 3$ and $n \neq 4$ (see [4]).

In this paper we give constructions for homogeneous function on $GF^{2n+1}(2)$, $n \geq 2$ and $n \neq 4$. The constructed functions attain high nonlinearity and contain no linear structures. These features are common with bent functions making the functions attractive for cryptographic applications.

2 Background

Let $V_n = GF^n(2)$ be the set of vectors with n binary co-ordinates. V_n contains 2^n different vectors from $\alpha_0(0,0,\cdots,0)$ to $\alpha_{2^n-1}(1,1,\cdots,1)$. A *Boolean function* $f : V_n \to GF(2)$ assigns binary values to vectors from V_n. Let $x(x_1,\cdots,x_n)$ and $y(y_1,\cdots,y_n)$ be two vectors in $GF(2)^n$. Throughout the paper we use the following notations:

- *The inner product* of x and y

$$\begin{aligned}\langle x,y\rangle &= x\odot y = x_1y_1\oplus\cdots\oplus x_ny_n\\ &= \sum_{i1}^{n} x_iy_i \pmod 2;\end{aligned}$$

- *The inner addition* of x and y

$$x\oplus y = (x_1\oplus y_1,\cdots,x_n\oplus y_n).$$

  Note that inner addition is equivalent to bit-by-bit XOR addition;

- *The Hadamard product* of vector $a = (a_1,\cdots,a_n)'$ and vector $b = (b_1,\cdots,b_n)'$

$$a * b = (a_1b_1,\cdots,a_nb_n)'$$

  where the symbol "$'$" means transpose of the vector or matrix.

A function $f(x)$ on V_n is called an affine function if $f(x) = a_1x_1\oplus\cdots a_nx_n\oplus c$ where $a_i \in GF(2)$, $i = 1,\cdots,n$, $c \in GF(2)$. When $c = 0$, $f(x)$ is called a linear Boolean function. The sequence of an affine (or linear) function is called affine (or linear) sequence.

Definition 1 *(Hamming weight and distance) The Hamming weight of a vector $\alpha \in V_n$, denoted by $W(\alpha)$, is the number of ones in the vector. The Hamming weight of a Boolean function $f(x)$, denoted by $W(f)$, is the number of ones in its truth table.*
The distance between two vectors α and β, denoted by $d(\alpha,\beta)$, is the number of co-ordinates which are different. Clearly, $d(\alpha,\beta) = W(\alpha\oplus\beta)$. The distance between two Boolean functions $f(x)$ and $g(x)$, denoted by $d(f,g)$, is equal to $W(f(x)\oplus g(x))$.

Definition 2 *(Propagation Criterion) A function $f(x)$ on V_n satisfies the propagation criterion with respect to $\alpha \neq 0$ if the derivative $f(x)\oplus f(x\oplus\alpha)$ is balanced, i.e. contains the same number of ones and zeros in its truth table.*

Definition 3 *(Homogeneous Boolean Function) A Boolean function $f : V_n \to GF(2)$ is homogeneous of degree k if it can be represented as*

$$f(x) = \bigoplus_{1\le i_1\le\cdots\le i_k\le n} a_{i_1\cdots i_k}x_{i_1}\cdots x_{i_k}. \quad (1)$$

where $x(x_1,\ldots,x_n)$. Each term $x_{i_1}\cdots x_{i_k}$, $a_{i_1\cdots i_k} \in GF(2)$ is a product of precisely k co-ordinates.

Let N_f denote the non-linearity of a Boolean function $f(x)$. The N_f is defined as follows:

$$N_f = \min\{d(f,\varphi) \mid \varphi \text{ is an affine function}\}.$$

Following the definition of nonlinearity of boolean functions, on even size boolean

spaces there is a special class of boolean functions which are called bent boolean functions that have Hamming distances to any affine function either $2^{n-1} - 2^{\frac{n}{2}-1}$ or $2^{n-1} + 2^{\frac{n}{2}-1}$. Bent functions have maximum nonlinearity, $2^{n-1} - 2^{\frac{n}{2}-1}$, on the space they exist.

The following well known results are given for completeness.

Lemma 1 *Let $f(x)$ and $g(x)$ be two Boolean functions on V_n, and let ξ and η be their sequences. Then*

$$d(f,g) = 2^{n-1} - \frac{1}{2}\langle \xi, \eta \rangle$$

Lemma 2 *Let $f(x)$ be any Boolean function on V_n. Then the non-linearity of $f(x)$,*

$$N_f \leq 2^{n-1} - 2^{\frac{1}{2}n-1}$$

Definition 4 *(Linear Structure) Let $f(x)$ be a Boolean function on $GF(2)^n$ and α be a non-zero vector from V_n. Then α is called a linear structure of $f(x)$ if its derivative $f(x \oplus \alpha) \oplus f(x)$ is constant.*

Lemma 3 *Let A be a nonsingular $n \times n$ matrix, α be a vector from V_n, and $\varphi(x)$ be an affine function on V_n. Then an arbitrary function $f(x) : V_n \to GF(2)$ and $g(x) = f(xA \oplus \alpha) \oplus \varphi(x)$ share the same nonlinearity, or*

$$N_f = N_g.$$

Theorem 1 *Suppose $f(x)$ is a Boolean function on V_{2m+1} with $N_f \geq 2^{2m} - 2^m$ which has no linear structure. Then there exists a Boolean function $F(x)$ on V_{2n+1} that has no linear structure and its nonlinearity $N_F \geq 2^{2n} - 2^n$ for every $n \geq m$.*

Corollary 1 *Let $f(x)$ be a Boolean function on V_{2n+1} with $N_f \geq 2^{2n} - 2^n$ and $g(y)$ be a bent function on V_{2m}. Then $F(z) = f(x) \oplus g(y)$ is the Boolean function on $V_{2(n+m)+1}$ with $N_F \geq 2^{2(n+m)} - 2^{n+m}$.*

3 Homogeneous Boolean Functions with High Nonlinearity

Definition 5 *Let $m \geq 1$, $F \in \Re_m$. The ranks $r_i(F)$, $1 \leq i \leq m$ are defined inductively on $deg(F)$:*

1. *When $deg(F) \leq 1$*

$$r_i(F) = rank(B_{1,i-1}^{(i,m)}(F(x))). \quad (2)$$

where $1 \leq i \leq m$.

2. *When $deg(F) = t > 1$, let $r_i(F)$, $1 \leq i \leq m$ be given by (2). Write $F \sim f(x_1, \cdots, x_r) \oplus g(x_1, \cdots, x_m)$ where $deg(f) = t$, $r = r_t(F)$, $deg(g) < t$, and*

$$r_i(F) = r_i(g(0, \cdots, 0, x_{r+1}, \cdots, x_m)), \quad (3)$$

$$1 \leq i < t.$$

The meaning of the ranks is this: If $deg(F) = t$, then $r_i(F) = 0$ for $i > t$, and $r_t(F)$ is the least number of independent linear combinations of $x_1, \cdots, x_m$ needed in the degree t part of F. Setting these

linear combinations equal to 0, the resulting function is used to define $r_i(F)$ for $i < t$(see Hou [1]).

When we construct highly non-linear cubic homogeneous Boolean functions, we only consider the case $r_3(f(x)) > n$ and $r_2(f(x)) = 0$.

Notation 1 *Let $T = (t_1, \cdots, t_n)$ be an $m \times n$ matrix, where t_i, $1 \leq i \leq n$ are the vectors of the columns with $m-$dimensions. We define*

$$T^* = (t_1 * t_2, \cdots, t_1 * t_n, t_2 * t_3, \cdots, t_2 * t_n, \cdots, t_{n-1} * t_n) \quad (4)$$

to be an $m \times \begin{pmatrix} n \\ 2 \end{pmatrix}$ matrix.

Lemma 4 *Let A be an $n \times n$ matrix, with all its entries in $GF(2)$. Then $xAx^{'} = \langle \beta, x \rangle$ for some $\beta \in V_n$ if and only if $A = A^{'}$.*

Theorem 2 *Let $r = r_3(f(x))$,$F(x) = f(x) \oplus g(x)$ where*

$$\begin{aligned} f(x) = & \bigoplus_{(i,j,k)\in E} x_i x_j x_k, \\ & 1 \leq i, j, k \leq r \leq n, \\ g(x) = & \bigoplus_{(i,j)\in S} x_i x_j, \\ & 1 \leq j \leq r,\ j \leq i \leq n. \end{aligned} \quad (5)$$

$F(x)$ is equivalent($\cong$) to a cubic homogeneous Boolean function iff there exists a nonsingular $n \times n$ matrix T and a constant vector $\alpha = (\alpha_1, \cdots, \alpha_n)$ such that:

$$\begin{aligned} & \left(T^*_{(1)} C \oplus TQ \oplus R\right) T^{'}_{(1)} \\ & = T_{(1)} \left(T^*_{(1)} C \oplus TQ \oplus R\right)^{'}, \end{aligned} \quad (6)$$

where $C^{'} = B^{(3,r)}_{1,2}(f(x))$, $T = (t_1, \cdots, t_n)$ with t_i, $1 \leq i \leq n$, are column vectors of n coordinates, $T_{(1)} = (t_1, \cdots, t_r)$, $T_{(2)} = (t_{r+1}, \cdots, t_n)$, and

$$\begin{aligned} Q = & \quad (q_{ij}) \quad q_{ij} = \begin{cases} 1, & (i,j) \in S,\ i > j \\ 0, & otherwise \end{cases} \\ R = & \left(\bigoplus_{(j,k)\in E_1, k>1} \alpha_j t_k, \cdots, \right. \\ & \left. \bigoplus_{(j,k)\in E_{r-1}, k>r-1} \alpha_j t_k, 0\right). \end{aligned} \quad (7)$$

$1 \leq i \leq n,\ \ 1 \leq j \leq r,$

Proof.

Let $(\alpha, \beta, c) \in G_n$. We have $(\alpha, \beta, c)(F(x)) = F(xT \oplus \alpha) \oplus \langle \beta, x \rangle \oplus c$, where

$$F(xT \oplus \alpha) = f(xT \oplus \alpha) \oplus g(xT \oplus \alpha), \quad (8)$$

$$\begin{aligned} f(xT \oplus \alpha) &= \bigoplus_{(i,j,k)\in E} (xt_i \oplus \alpha_i)(xt_j \oplus \alpha_j) \\ &\quad (xt_k \oplus \alpha_k) \\ &= \bigoplus_{(i,j,k)\in E} \{xt_i xt_j xt_k \oplus \alpha_k xt_i xt_j \\ &\quad \oplus \alpha_j xt_i xt_k \oplus \alpha_i xt_j xt_k\} \oplus \varphi_1(x) \\ &= \bigoplus_{(i,j,k)\in E} xt_i xt_j xt_k \oplus xRT^{'}_{(1)} x^{'} \\ &\quad \oplus \varphi_1(x) \\ &= H(x) \oplus xT^*_{(1)} CT^{'}_{(1)} x^{'} \\ &\quad \oplus xRT^{'}_{(1)} x^{'} \oplus \varphi(x), \end{aligned} \quad (9)$$

where H is a cubic homogeneous function, and $\varphi_1(x)$ is an affine function.

$$\begin{aligned} g(xT \oplus \alpha) &= (xT \oplus \alpha) Q (xT_{(1)} \oplus \alpha_{(1)})^{'} \\ &\quad xTQT^{'}_{(1)} x^{'} \oplus \varphi_2(x), \end{aligned}$$

where $\varphi_2(x)$ is an affine function, and $\alpha_{(1)} = (\alpha_1, \cdots, \alpha_r)$. Now

$$(\alpha, \beta, c) \quad (F(x)) = H(x) \oplus \varphi_3(x) \oplus x(T^*_{(1)} C \oplus TQ \oplus R) T'_{(1)} x', \quad (10)$$

where
$\varphi_3(x)$ is an affine function. $(\alpha, \beta, c)(F(x))$ is a cubic homogeneous function if and only if $x(T^*_{(1)} C \oplus TQ \oplus R) T'_{(1)} x'$ is an affine function. From lemma 4 we get $(T^*_{(1)} C \oplus TQ \oplus R) T'_{(1)}$ is a symmetric matrix so Equation (6) holds. □

Theorem 3 *There exist cubic homogeneous functions on V_{2n+1} without linear structure, except possibly $n = 4$, and the non-linearity of such functions are no less than $2^{2n} - 2^n$ for $n > 1$,*

Proof. On V_5, let $F(x) = x_1x_2x_3 \oplus x_1x_4 \oplus x_2x_5$. We know that $F(x)$ has no linear structure, and its non-linearity 12. Now $r = r_3(F(x)) = 3$, and take

$$T = (t_1, t_2, t_3, t_4, t_5) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \end{pmatrix}.$$

We compute

$$\begin{aligned} T^*_{(1)} C &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \\ &= (t_4, t_5, 0) = TQ \\ &= T \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \end{aligned}$$

and (6) holds.

On
V_7, let $f(x) = x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_1x_4 \oplus x_2x_5 \oplus x_2x_7 \oplus x_3x_7$. It is easy to check $f(x)$ has no linear structure and $N_f = 56$. $r = r_3(f(x)) = 6$. Take

$$T = (t_1, \cdots, t_7) = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

We have

$$\begin{aligned} T^*_{(1)} C &= \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix} \\ &= (0, t_1 \oplus t_5 \oplus t_7, t_3 \oplus t_7, t_1 \oplus t_6, 0, 0) \end{aligned}$$

and

$$\begin{aligned} TQ &= (t_4, t_5 \oplus t_7, t_7, 0, 0, 0), \\ R &= (t_2, 0, 0, t_6, 0, 0). \end{aligned}$$

Now

$$\begin{aligned} &(T^*_{(1)} C \oplus TQ \oplus R) T'_{(1)} \\ &= (t_2 \oplus t_4) t'_1 \oplus t_1 (t_2 \oplus t_4)' \oplus t_3 t'_3. \end{aligned}$$

and it is a symmetric matrix. In this case $F(xT \oplus \alpha) \oplus x(t_3 \oplus t_7)$ is a cubic homogeneous function with $N_F = N_f = 56$ and has no linear structure.

For any $m \geq 3$ and $m \neq 4$, we know, from [4], that there exist cubic homogeneous bent functions on V_{2m}. For any $n > 4$, let $m = n - \ell \geq 3 \neq 4$, $F(z) = f(x) \oplus g(y)$ where $f(x)$ is a cubic homogeneous cubic function with $N_f \geq 2^{2l} - 2^{\ell}$, $g(y)$ is a cubic homogeneous bent function on $V_{2(n-\ell)}$, $z(x, y)$, $x \in V_{2\ell}+1$, $y \in V_{2(n-\ell)}$. From Theorem 1, Corollary 1 and Theorem 2([4]), we get cubic homogeneous functions with $N_F \geq 2^{2n} - 2^n$ and with no linear structure. □

4 Explicit Solutions

In this section we give transformation matrix solutions for each of the remaining seven inequivalent cases for V_7 and give results for V_n, n odd, $n \geq 11$. We discuss the properties of these solutions.

4.1 Homogeneous Cubic Functions on V_5

Let

$$f(x) = x_{i_1}x_{i_2}x_{i_3} \oplus x_{i_1}x_{i_4} \oplus x_{i_2}x_{i_5}, \quad (11)$$

where $\{i_1, i_2, i_3, i_4, i_5\}\{1, 2, 3, 4, 5\}$. Assume the following linear transformation T:

$$\begin{array}{l} x_{i_1} \to x_{i_1} \oplus x_{i_5}, \quad x_{i_2} \to x_{i_2} \oplus x_{i_4}, \\ x_{i_3} \to x_{i_3} \oplus x_{i_4} \oplus x_{i_5}, \\ x_{i_4} \to x_{i_4}, \quad x_{i_5} \to x_{i_5}. \end{array} \quad (12)$$

Clearly, matrix T is nonsingular and $f(xT) = \bigoplus_{1 \leq u,v,w \leq 5} x_u x_v x_w \oplus x_{i_1}x_{i_3}x_{i_5} \oplus x_{i_2}x_{i_3}x_{i_4}$ is a cubic homogeneous function with non-linearity $N_f = 12$.

There are 60 different functions of the form (11). They are equivalent under the action of T as defined by (12). There are 15 different cubic homogeneous Boolean functions with $N_f = 12$ which are equivalent under the action of (12).

4.2 Homogeneous Cubic Functions on V_7

Including the Boolean function on V_7 which given in theorem 3, we found 8 cubic functions of 7 variables. Every one of them is equivalent to a cubic homogeneous function under the action of G_7. But they are not equivalent to each other. We continue to search for a generalised construction for the matrix T and the vector α so that $f(xT \oplus \alpha) = H(x) \oplus \varphi(x)$ where $H(x)$ is cubic homogeneous function, and $\varphi(x)$ is an affine function. The cases just given are positive evidence for such a construction.

4.3 On V_{2n+1}, $n > 4$

The cubic homogeneous Boolean functions, $f(x)$, exist on V_5 with high non-linearity $N_f = 12$, and cubic homogeneous bent functions, $g(y)$, exist on V_6. From theorem 1 and corollary 1 we can construct a cubic homogeneous Boolean function $F(z) = f(x) \oplus g(y)$ on V_{11} with high non-linearity $N_F = 2^{10} - 2^5 = 992$.

Example 1 *Let*

$$\begin{aligned} f(x) &= x_1x_2x_3 \oplus x_1x_2x_4 \oplus x_1x_2x_5 \\ &\quad \oplus x_1x_3x_4 \oplus x_1x_3x_5 \oplus x_2x_3x_4 \\ &\quad \oplus x_2x_4x_5 \oplus x_3x_4x_5, \end{aligned}$$

and

$$\begin{aligned} g(y) &= y_1y_2y_3 \oplus y_1y_2y_4 \oplus y_1y_2y_5 \oplus y_1y_3y_4 \\ &\quad \oplus y_1y_3y_6 \oplus y_1y_4y_5 \oplus y_1y_4y_6 \\ &\quad \oplus y_1y_5y_6 \oplus y_2y_3y_5 \oplus y_2y_3y_6 \\ &\quad \oplus y_2y_4y_5 \oplus y_2y_4y_6 \oplus y_2y_5y_6 \\ &\quad \oplus y_3y_4y_5 \oplus y_3y_4y_6 \oplus y_3y_5y_6, \end{aligned}$$

then $F(z) = f(x) \oplus g(y)$ *is a cubic homogeneous Boolean function and* $N_F = 992$.

Using this method, we can construct cubic homogeneous Boolean functions F on V_{2n+1} for $n > 4$ which high non-linearity $N_F \geq 2^{2n} - 2^n$.

References

[1] Xiang-dong Hou, Cubic bent functions, *Discrete Mathematics*, 189, (1998), pp. 149-161.

[2] Josef Pieprzyk, Chenxin Qu, Rotate symmetric functions and fast hashing, *Information Security and Privacy, ACISP'98, Lecture Notes in Computer Science*, vol. 1438, Springer-Verlag Berlin, Heidelberg, New York, (1998). pp. 169-180.

[3] O.S.Rothaus, On "bent" functions, *Journal of Combinatorial Theory*, Ser. A, 20, (1976), pp. 300-305.

[4] J. Seberry, T. Xia and J. Pieprzyk, Construction cubic homogeneous bent functions, *The Australiasian Journal of Combinatorics*, vol. 22, pp. 233-245, 2000.

[5] T. Xia, J. Seberry, J. Pieprzyk, and Chris Charnes, Homogeneous Bent Functions of degree n in $2n$ variables do not exist for $n > 3$, (in print).