Front. Electr. Electron. Eng. China (2006) 3: 277–282
DOI 10.1007/s11460-006-0038-8

RESEARCH ARTICLE

CHEN Wei, YANG Yi-xian, NIU Xin-xin

Construction of optimized Boolean functions

Abstract Considering connections of characteristics, this paper is aimed at the construction of optimized Boolean functions. A new method based on the Bent function, discrete Walsh spectrum and characteristics matrices are presented by concatenating, breaking, and revising output sequences conditionally. This new construction can be used to construct different kinds of functions satisfying different design criteria.

Keywords function construction method, balance, nonlinearity, SAC, correlation-immune

1 Introduction

The main design criteria of a Boolean function are balance, nonlinearity, algebra degree and term distribution, completeness, strict avalanche criteria (SAC), correlation immunity (CI), etc.. Boolean function construction, which satisfies multiple characteristics requisition, is always one of the hot research topics in this field. Scientists in China and eleswhere present a lot of constructions from all kinds of fields [1, 2], which can be divided into three kinds: convolute (concatenate) construction, algebraic(add, multiply) construction and matrix vector construction. All of them have their own emphasis and shortcomings, and their consideration of characteristics is not more than 3 or 4. Construction of a function that can optimize multiple characteristics (5 or more) is still a practical problem with relative difficulty. On the basis of induction and generalization of spectrum trait and matrix trait of characteristics, this paper presents a construction method that can optimize multiple characteristics, and have much more real meaning on block and stream cipher design, such as the construction of key stream generator.

Translated from *Journal on Communications*, 2006, 27(2): 22–28 (in Chinese)

CHEN Wei (✉), YANG Yi-xian，NIU Xin-xin
Information Security Center,
Beijing University of Posts and Telecommunications,
Beijing 100876, China
E-mail: cyberella2000@163.com

2 Spectrum trait and matrix trait of general characteristics

The balance of n-dimension function $\boldsymbol{f}(\boldsymbol{x})$ means the equivalence of 0's and 1's in its true value table, and nonlinearity $N_f = \min(d_H(\boldsymbol{f}, l))$ reflects the level of approximation by its affine function $l(\boldsymbol{x})$. On the viewpoint of cryptology, nonlinearity of $\boldsymbol{f}$ should be much better. The number of $i(0 \le i \le n-1)$ degree term is named the "i-term number of $\boldsymbol{f}$". Term distribution is the distribution of all degree terms of $\boldsymbol{f}$, and algebra degree is equal to the highest degree. When the term number is low, interpolation attack will be successful; and when the algebra degree is low, high order attack and differential attack can be applied. A good function in cryptology is highly non-linear 0–1 balanced, with high algebraic order and uniform term distribution. $\boldsymbol{f}$ is called m-CI when it is statistic independent to any m input of n dimension, which means $\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}$ is balanced to any $\boldsymbol{w} = (0, \cdots, w_{i_1}, \cdots, w_{i_m}, \cdots, 0), \in F_2^n$, $1 \le W_H(\boldsymbol{w}) \le m$. $\boldsymbol{f}(\boldsymbol{x})$ is SAC if and only if the self correlation function $C_f(\boldsymbol{e}_i) = \sum_{\boldsymbol{x} \in F_2^n} (-1)^{\boldsymbol{f}(\boldsymbol{x}) \oplus \boldsymbol{f}(\boldsymbol{x} \oplus \boldsymbol{e}_i)} = 0$ to any n-dimension vector $\boldsymbol{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0)$. SAC ensures realization of completeness and non-degeneracy, and such function is also called the "n-dimension H-function". The second non-normalized loop Walsh spectrum of $\boldsymbol{f}$ is defined as $S_{(f)}(\boldsymbol{w}) = \sum_{w \in F_2^n} (-1)^{\boldsymbol{f}(\boldsymbol{x}) + \boldsymbol{w} \cdot \boldsymbol{x}}$, which has a relation with nonlinearity that $N_f = 2^{n-1} - \max_{\boldsymbol{w} \in F_2^n} \left| S_{(f)}(\boldsymbol{w}) \right| / 2$ [1], so it can be used to describe nonlinearity. N-dimension vector $\boldsymbol{x}$ is called the "trait vector of $\boldsymbol{f}$" when $\boldsymbol{f}(\boldsymbol{x}) = 1$ is satisfied, so matrix $\boldsymbol{A}$ which uses the trait vector as its row

vector is called "trait matrix of $\boldsymbol{f}$". Walsh spectrum and trait matrix are useful tools to study function characteristics. Using them to describe function characteristics, we can draw some conclusions, as follows:

Lemma 1 [3] Balanced Boolean function $\boldsymbol{f}(\boldsymbol{x})$ has the maximum degree $n-1$, as well as the spectrum trait of $S_{(f)}(0)=0$ [1], the upper bound of its nonlinearity follows inequality as:

$$N_{f\mathrm{B}} \le N_{f\mathrm{Bmax}} = \begin{cases} 2^{n-1}-(2^{\frac{n}{2-1}}+2^{\frac{n}{4-1}}+L+2^3+2^2), \\ \text{if } n=2^m \\ 2^{n-1}-(2^{\frac{n}{2-1}}+2^{\frac{n}{4-1}}+L+2^{2t}+2^t), \\ \text{if } n=2^s(2t+1) \end{cases} \tag{1}$$

Let $\mathrm{RB}_f = \max\limits_{\boldsymbol{w}\in F_2^n} \left|S_{(f)}(\boldsymbol{w})\right|$, so if we treat the first instance of Eq. (1) as a special result of the second instance when $t=0$, then we can get the equation that $\mathrm{RB}_{f\max} = 2^{n/2} + 2^{n/4} + \cdots + 2^{2t+1} + 2^{t+1}$, when $n = 2^s(2t+1)$.

Lemma 2 [4, 5] If the weight of $\boldsymbol{f}(\boldsymbol{x})$ is $2k$, then $\boldsymbol{f}(\boldsymbol{x})$ is an H-function if and only if in a sub matrix composed by any $n-1$ column of matrix, there are $k-2^{n-3}$ couples ($2k-2^{n-2}$ total) of row vectors that are the same with each other and other row vectors that are different from each other, namely: there exist $k-2^{n-3}$ $x \in F_2^n$, which have $f(\boldsymbol{x}|\boldsymbol{x}_i = 0) = f(\boldsymbol{x}|\boldsymbol{x}_i = 1) = 1$.

We name such vector couples as a "conjugate vector couple".

Lemma 3 [1] **(Xiao-Massey Theorem)** $\boldsymbol{f}(\boldsymbol{x})$ is m-CI if and only if $S_{(f)}(\boldsymbol{w}) = 0$ for any $\boldsymbol{w} = (0,\cdots,w_{i_1},\cdots,w_{i_m},\cdots,0) \in F_2^n$, $1 \le W_H(\boldsymbol{w}) \le m$.

Because there is a relation between nonlinearity N_f and the number of points in which spectrum value is zero $N_{S_{(f)}}$, that is $N_f / 2^{n-1} + 1/\sqrt{2^n - N_{S_{(f)}}} \le 1$. So, by spectrum trait of CI, we can get the inequality as follows:

$$\frac{N_f}{2^{n-1}} + \frac{1}{\sqrt{2^n - \sum\limits_{i=1}^{m} C_n^i}} \le 1 \tag{2}$$

3 Normal function constructions

For diverse applications, scientists in China and abroad present a lot of constructions. In general, they can be divided into some categories, as follows:

1) Convolute (concatenate) construction: constructions presented by Zhang [6] and Qiu [7] can be generalized into convolute construction, which connects end-to-end $2^m (m \ge 1)$ n-dimension function output satisfying some requisition to build $n+m$ dimension function. Rough analysis indicates that $\deg(\boldsymbol{f}) \le m + \max(\deg(f_i)) \le n+m-1$, $N_f \ge \sum\limits_{\boldsymbol{a}\in F_2^m} N_{f_{\boldsymbol{a}}}$. If at least one maximum degree term exists in function collection F with odd times, then collection F is easily obtained. But, when $n+m$ is large, collection F is not so easy to obtain, and neither is the nonlinearity of $\boldsymbol{f}$ easy to control.

2) Algebra construction: Qin [8] and Ji [9] presented add construction, which was generalized by Zeng [10] as fallows: suppose $\boldsymbol{f}_i(\boldsymbol{x}^{(i)})$ is an n_i ($1 \le i \le k$) dimension Boolean function satisfied to some requisition(such as PC (1), CI(n), etc.), and there exists function $\boldsymbol{h}_i(\boldsymbol{x}^{(i)})$, which makes $\boldsymbol{f}_i(\boldsymbol{x}^{(i)}) + \boldsymbol{h}_i(\boldsymbol{x}^{(i)})$ satisfy the same requisition, and then to any k dimension function $\boldsymbol{g}(z)$, $(n_1+\cdots+n_k)$ dimensional function $\sum\limits_{i=1}^{k} \boldsymbol{f}_i(\boldsymbol{x}^{(i)}) + \boldsymbol{g}(\boldsymbol{h}_1(\boldsymbol{x}^{(1)}),\cdots,\boldsymbol{h}_k(\boldsymbol{x}^k))$ will satisfy the same requisition. We can adjust $\boldsymbol{g}(z)$ to balance the resuling function.

Rough analysis indicates that $\deg(\sum\limits_{i=1}^{k} \boldsymbol{f}_i(\boldsymbol{x}^{(i)})) = \max(\deg(\boldsymbol{f}_i(\boldsymbol{x}^{(i)}))) < \max(n_i)$, and the algebra degree of $\boldsymbol{g}(\boldsymbol{h}_1(\boldsymbol{x}^{(1)}),\cdots,\boldsymbol{h}_k(\boldsymbol{x}^k))$ is less than $\sum \deg(\boldsymbol{h}_i(\boldsymbol{x})) \le \sum(n_i - 1) \le \sum n_i - k$. The algebra degree of the resulting function is bigger than one of these two values. The nonlinearity of the function is $N_f = \frac{1}{2}[\prod\limits_{i=1}^{k} 2^{n_i} - \prod\limits_{i=1}^{k}(2^{n_i} - 2N_{f_i})]$.

Kurosawa [11] presents a similar construction, namely: to any $s \times t$ matrix $\boldsymbol{Q}$ and s-dimension function $\boldsymbol{g}(\boldsymbol{x})$, let $\boldsymbol{f}(x_1,\cdots,x_s,y_1,\cdots,y_t) = [x_1,\cdots,x_s]\boldsymbol{Q}_{s\times t}[y_1,\cdots,y_t]^{\mathrm{T}} \oplus \boldsymbol{g}(x_1,\cdots,x_s)$, if every row and every column of $\boldsymbol{Q}$ have at lease one '1', $\boldsymbol{f}$ is an H-function. We can adjust $\boldsymbol{g}(\boldsymbol{x})$ to balance the resulting function. If s is even, $\deg(\boldsymbol{f}) = s/2$, $N_f \ge 2^{t+s-1} - 2^{t+s/2-1}$; else $\deg(f) = s - 1/2$, $N_f \ge 2^{t+s-1} - 2^{t+(s-1)/2-1}$. Its shortcoming is that it has a low algebra degree and low nonlinearity. When $t = s$, it evolves in to a Bent function constructioned mentioned in Ref. [1].

3) Revising the true value to balance the unbalanced function: References [12, 13] studied how to construct a highly nonlinear balanced function by concatenate, split and revised Bent sequence, but its aim was only a highly nonlinear balanced function, and not involved in other criteria, so the SAC structure of the Bent sequence is destroyed.

Constructions mentioned above have their own advantages and disadvantages. Using any will bring us some limitations. On the construction of a Boolean function, we consider that on the basis of highly nonlinear H-function, we revise the output sequence to satisfy other requisitions.

4 Optimized function construction

If $S_{(f)}(\boldsymbol{w}) = \pm 2^{n/2}$ is right for any $\boldsymbol{w} \in F_2^n$, n-dimension function $\boldsymbol{f}(\boldsymbol{x})$ is named a Bent function. It has a weight of $2^{n-1} \pm 2^{n/2-1}$, and has maximum nonlinearity. On the viewpoint of cryptology, the Bent function has many shortcomings, such as unbalance, limitation on n of being even only, and being not more than $n/2$ degree. By spectrum trait, it is not a CI function. So, we can revise the Bent function to build an optimized function in multiple characteristics.

We put much research emphasis on how to use methods mentioned above to revise the Bent function, playing attention to the different characteristic requisition. Because nonlinearity will decrease in an exponential manner by CI grade, they should compromise with each other suitably in application. Next, we discuss on the example of how to build a high degree, highly nonlinear balanced H-function and high degree, highly nonlinear balanced 1-CI function.

4.1 Convolute (concatenate) and split Bent function

Sun et al. [14] and Wen et al. [15] presented a convolute (concatenate) construction that uses Bent function to build a highly nonlinear function. Suppose $\boldsymbol{f}_1(\boldsymbol{x}) =: \boldsymbol{f}_1(x_1, x_2, \cdots, x_m)$ and $\boldsymbol{f}_2(\boldsymbol{x}) =: \boldsymbol{f}_2(x_1, x_2, \cdots, x_m)$ are two m-dimension Bent functions. If we have $\boldsymbol{f}(\boldsymbol{x}) = x_{m+1}\boldsymbol{f}_1(\boldsymbol{x}) + (x_{m+1}+1)\boldsymbol{f}_2(\boldsymbol{x})$, then its nonlinearity $N_f \geq 2^m - 2^{m/2}$. So, the following theorem can be obtained.

Theorem 1 If $\boldsymbol{f}_1(\boldsymbol{x})$ has a weight of $2^{m-1} \pm 2^{m/2-1}$ and $\boldsymbol{f}_2(\boldsymbol{x})$ has a weight of $2^{m-1} \mp 2^{m/2-1}$, and $d_H(\boldsymbol{f}_1, \boldsymbol{f}_2) = 2^{m-1}$, then $\boldsymbol{f}(\boldsymbol{x}) = x_{m+1}\boldsymbol{f}_1(\boldsymbol{x}) + \overline{x}_{m+1}\boldsymbol{f}_2(\boldsymbol{x})$ will be a balanced H-function with the highest nonlinearity, its algebra degree $\deg(\boldsymbol{f}) = \max(\deg(\boldsymbol{f}_1), \deg(\boldsymbol{f}_2)) + 1$.

Proof

1) Balance : for convolute (concatenate) construction, we have: $W_H(\boldsymbol{f}) = W_H(\boldsymbol{f}_1) + W_H(\boldsymbol{f}_2) = 2^{m-1} \pm 2^{m/2-1} + 2^{m-1} \mp 2^{m/2-1} = 2^m$, so $\boldsymbol{f}(\boldsymbol{x})$ is m+1 dimension balanced function.

2) Nonlinearity: from Eq. (1), when $n = 2k+1$, $m = 2k$, there is $N_f \leq 2^{n-1} - 2^{(n-1)/2} = 2^m - 2^{m/2}$ for balanced function. We already have $N_f \geq 2^m - 2^{m/2}$ from above, so N_f get its maximal value.

3) H-function: Bent function is H-function. For any $\boldsymbol{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0) \in F_2^m$, $1 \leq i \leq m$, $\boldsymbol{f}_1(\boldsymbol{x}) \oplus \boldsymbol{f}_1(\boldsymbol{x} \oplus \boldsymbol{e}_i)$ and $\boldsymbol{f}_2(\boldsymbol{x}) \oplus \boldsymbol{f}_2(\boldsymbol{x} \oplus \boldsymbol{e}_i)$ will be both balanced functions, then $\boldsymbol{f}(\boldsymbol{x}) \oplus \boldsymbol{f}(\boldsymbol{x} \oplus \boldsymbol{e}_i) = x_{m+1}(\boldsymbol{f}_1(\boldsymbol{x}) \oplus \boldsymbol{f}_1(x \oplus \boldsymbol{e}_i)) + (x_{m+1}+1)(\boldsymbol{f}_2(\boldsymbol{x}) \oplus \boldsymbol{f}_2(\boldsymbol{x} \oplus \boldsymbol{e}_i))$ is also a balanced function. For $\boldsymbol{e}_{m+1} = (0, \cdots, 0, 1)$:

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{x}) \oplus \boldsymbol{f}(\boldsymbol{x} \oplus \boldsymbol{e}_{m+1}) &= x_{m+1}(\boldsymbol{f}_1(\boldsymbol{x}) \oplus \boldsymbol{f}_2(\boldsymbol{x})) \\ &\quad + (x_{m+1}+1)(\boldsymbol{f}_2(\boldsymbol{x}) \oplus \boldsymbol{f}_1(\boldsymbol{x})) \\ &= \boldsymbol{f}_1(\boldsymbol{x}) \oplus \boldsymbol{f}_2(\boldsymbol{x}) = 2d_H(\boldsymbol{f}_1, \boldsymbol{f}_2) \\ &= 2^m = 2^{n-1}\end{aligned}$$

is balanced too. So, for every $\boldsymbol{e}_i = (0, \cdots, 0, 1, 0, \cdots, 0) \in F_2^n$, $1 \leq i \leq n$, $\boldsymbol{f}(\boldsymbol{x}) \oplus \boldsymbol{f}(\boldsymbol{x} \oplus \boldsymbol{e}_i)$ is balanced.

4) Algebra degree：we conclude the theorem conclusion from the expression.

Reference [13] discussed the method of split Bent sequence to an build an odd dimension highly nonlinear balanced H-function. For $2k$ dimension Bent function $\boldsymbol{f}$, between two $2k-1$ dimension sequence as $\boldsymbol{f}(x_1, x_2, \cdots, x_{i-2}, 0, x_i, \cdots x_{2k})$ and $\boldsymbol{f}(x_1, x_2, \cdots, x_{i-2}, 1, x_i, \cdots x_{2k})$ ($1 \leq i \leq 2k$), there must be one balanced function with its nonlinearity $N_f = 2^{2k-2} - 2^{k-1}$. By characteristics of the Bent function, it must be an H-function too.

4.2 Revise function output

It is easy to build a high degree, highly nonlinear balanced SAC function by convolute and split construction. But, because the Bent function is not a CI function, the method of revised function output must be applied to satisfy 1-CI. And, concatenate and split construction can build an odd dimension balanced H-function with maximal nonlinearity from even dimension Bent function, but not vice versa. By Eq. (1), we can calculate the following table:

Table 1 Maximum nonlinearity of balanced function

N	7	8	9	10	11	12
$N_{f\mathrm{B}\max}$	56	116	240	492	992	2 012

From Table 1, we know that an even dimension function composed by odd function convolution cannot get maximal nonlinearity, and nonlinearity of $m+k$ dimension function from m-dimension Bent function convolute is far more from maximal nonlinearity. Furthermore, we can only assure the lower bound of the resulting function, but not the exact nonlinearity. So, we cannot create a high dimension function with optimized characteristic by heaping low dimension functions together. A revised function output to improve function characteristic is necessary. Two situations are discussed as follows.

4.2.1 Revise output, the weight of the function is changed

It is fit to revise a Bent function to improve multi-characteristics.

Suppose the origin function $\boldsymbol{f}_0$ is a Bent function of weight $2^{n-1} + 2^{n/2-1}$, and its trait matrix is $\boldsymbol{A}_b$. To balance it, we must extract $2^{n/2-1}$ row vector and removei it from

$\boldsymbol{A}_b$. Consider the following relations:

$$\boldsymbol{f}_i=\begin{cases}0, & \text{if } \boldsymbol{x}=x_i\\ \boldsymbol{f}_{i-1}, & \text{else}\end{cases},\qquad \boldsymbol{f}_{i-1}=\begin{cases}1, & \text{if } \boldsymbol{x}=x_i\\ \text{any}, & \text{else}\end{cases},$$

$$\boldsymbol{g}_i=\begin{cases}1, & \text{if } \boldsymbol{x}=x_i\\ 0, & \text{else}\end{cases}$$

where $\boldsymbol{f}_i$, $\boldsymbol{f}_{i-1}$, $\boldsymbol{g}_i$ are all n-dimension functions, $\boldsymbol{f}_i=\boldsymbol{f}_{i-1}\oplus\boldsymbol{g}_i$, $1\le i\le 2^{n/2-1}$. The resulting function $\boldsymbol{f}_{2^{n/2-1}}$ is balanced. We suppose that the extracted trait vector compose matrix T_e, which has a corresponding sparse function $\boldsymbol{g}(\boldsymbol{x})=\boldsymbol{g}_1(\boldsymbol{x})\oplus\boldsymbol{g}_2(\boldsymbol{x})\oplus\cdots\oplus\boldsymbol{g}_{2^{n/2-1}}(\boldsymbol{x})$.

First, let us study SAC. We delete the ith column $(1\le i\le n)$ of $\boldsymbol{A}_b$, then divide the row vectors into two sets based on whether they belongs to some conjugate vector couple of sub matrix or not. $\boldsymbol{C}(i)_r$ includes these conjugate vector couples, and $\boldsymbol{C}(i)_{ur}$ the rest. By Lemma 2, any $n-1$ column sub matrix from the trait matrix of Bent function include $2^{n-3}\pm 2^{n/2-2}$ conjugate vector couple, $|\boldsymbol{C}(i)_r|=2^{n-2}+2^{n/2-1}$, $|\boldsymbol{C}(i)_{ur}|=2^{n-2}$, and any $n-1$ column sub matrix from the trait matrix of balanced H-function include 2^{n-3} conjugate vector couple, $|\boldsymbol{C}(i)_r|=|\boldsymbol{C}(i)_{ur}|=2^{n-2}$. So, we must extract $2^{n/2-2}$ conjugate vector couple in every $\boldsymbol{C}(i)_r$ $(1\le i\le n)$ of $\boldsymbol{A}_b$. Thus the following theorem:

Theorem 2 For all $1\le i\le n$, $\boldsymbol{C}(i)_r$ must choose $2^{n/2-2}$ different conjugate vector couples, from each of which at least one trait vectors of $\boldsymbol{T}_e$ is obtained.

Secondly, let us consider the nonlinearity. From definition of loop Walsh spectrum, we known that $S_{(\boldsymbol{f}_i)}(\boldsymbol{w})=S_{(\boldsymbol{f}_{i-1})}(\boldsymbol{w})+S_{(\boldsymbol{g}_i)}(\boldsymbol{w})$, $S_{(\boldsymbol{g}_i)}(\boldsymbol{w})=\pm 2$. Let $R_{\boldsymbol{f}_{i-1}}$ be the maximal spectrum absolute value of $\boldsymbol{f}_{i-1}$, $R^{-}_{\boldsymbol{f}_{i-1}}$ be the sub maximum. By Lemma 1, when $n=2^s(2t+1)$, $R_{f_{2^{n/2-1}}}=\mathrm{RB}_{f_{2^{n/2-1}}\max}=2^{n/2}+2^{n/4}+\cdots+2^{2t+1}+2^{t+1}$. So we have:

$$R_{\boldsymbol{f}_i}\le 2^{n/2}+2^{n/4}+\cdots+2^{2t+1}+2^{t+1}+(2^{n/2}-2i) \tag{3}$$

From Eq. (3), we can draw the following conclusion:

Theorem 3 For $\boldsymbol{f}_{i-1}(\boldsymbol{x})=1$, $0\le x_i\le 2^n-1$, if we can find $\boldsymbol{w}'$ and $\boldsymbol{w}''$, which make $|R_{\boldsymbol{f}_{i-1}}|=2^{n/2}+2^{n/4}+\cdots+2^{2t+1}+2^{t+1}+(2^{n/2}-2i+2)$, $|R^{-}_{\boldsymbol{f}_{i-1}}|=|R_{\boldsymbol{f}_{i-1}}|-2$, then if $S_{(\boldsymbol{g}_i)}(\boldsymbol{w}')$ and $S_{(\boldsymbol{f}_{i-1})}(\boldsymbol{w}')$ have opposite signs, and $S_{(\boldsymbol{g}_i)}(\boldsymbol{w}'')$ and $S_{(\boldsymbol{f}_{i-1})}(\boldsymbol{w}'')$ is either, then Eq. (3) holds.

Thirdly, let us consider CI, and make 1-CI as the example. By Lemma 3, when $W_H(\boldsymbol{w})=1$, there must be $S_{(\boldsymbol{f}_{2^{n/2-1}})}(\boldsymbol{w})=0$, so the request of CI to the spectrum value is: for any $1\le i\le 2^{n/2-1}$, we must keep $|S_{(\boldsymbol{f}_i)}(\boldsymbol{w})|\le 2^{n/2}-2i$. If they are equal, $S_{(\boldsymbol{g}_i)}(\boldsymbol{w})$ and $S_{(\boldsymbol{f}_0)}(\boldsymbol{w})$ must be opposite signs.

4.2.2 Revise output, the weight of the function is not changed

It is fit to revise a balanced function, which is not CI, to make it CI.

Suppose we change l "1" to "0", the origin function is $\boldsymbol{f}_0$ and the resulting function is $\boldsymbol{f}_l$, $\boldsymbol{f}_i=\boldsymbol{f}_{i-1}\oplus\boldsymbol{g}_i$, $1\le i\le l$, where $\boldsymbol{f}_i$, $\boldsymbol{f}_{i-1}$, $\boldsymbol{g}_i$ are all n-dimension functions, the trait matrix of $\boldsymbol{f}_i$ is $\boldsymbol{A}_b$, the trait matrix of $\boldsymbol{g}_i$ is $\boldsymbol{T}_g$. Consider the following relation:

$$\boldsymbol{f}_i=\begin{cases}0, & \text{if } \boldsymbol{x}=x_{i_1}\\ 1, & \text{if } \boldsymbol{x}=x_{i_2}\\ f_{i-1}, & \text{else}\end{cases},\quad \boldsymbol{f}_{i-1}=\begin{cases}1, & \text{if } \boldsymbol{x}=x_{i_1}\\ 0, & \text{if } \boldsymbol{x}=x_{i_2}\\ f_{i-1}, & \text{else}\end{cases},$$

$$\boldsymbol{g}_i=\begin{cases}1, & \text{if } \boldsymbol{x}=x_{i_1} \text{ or } \boldsymbol{x}=x_{i_2}\\ 0, & \text{else}\end{cases}$$

From consideration of CI, we can get the lower bound of l. Let us make 1-CI as an example. For $W_H(\boldsymbol{g}_i)=2$, $S_{(\boldsymbol{g})}(\boldsymbol{w})=\pm 4,0$, by Lemma 3, when $W_H(\boldsymbol{w})=1$, there must be $S_{(\boldsymbol{f}_{2^{n/2-1}})}(\boldsymbol{w})=0$, so $l\ge\max(|S_{(\boldsymbol{f}_i)}(\boldsymbol{w})|)/4$. As discussed above, the request of the CI to the spectrum value is: for any w $(W_H(\boldsymbol{w})=1)$ and $1\le i\le l$, we must keep $|S_{(\boldsymbol{f}_i)}(\boldsymbol{w})|\le 4(l-i+1)$.

If the original function is SAC and the resuling function is not to destroy it, then we should make $\boldsymbol{A}_b$ keep $|\boldsymbol{C}(i)_r|=|\boldsymbol{C}(i)_{ur}|=2^{n-2}$ $(1\le i\le n)$. For $W_H(\boldsymbol{g}_i)=2$, T_g is composed of two row vectors, supposing "$\boldsymbol{a}$" and "$\boldsymbol{b}$", in which there must be only one coming from $\boldsymbol{A}_b$, assuming $\boldsymbol{T}_g\cap\boldsymbol{A}_b=\boldsymbol{a}$. If $\boldsymbol{a}\in\boldsymbol{C}(i)_r$, then $\boldsymbol{b}$ must be coupled with one of the row vectors in $\boldsymbol{C}(i)_{ur}$ to make a conjugate vector couple. Thus one loss of the conjugate vector couple by deleting "$\boldsymbol{a}$" from $\boldsymbol{A}_b$ is compensated. Similarly, if $\boldsymbol{a}\in\boldsymbol{C}(i)_{ur}$, then $\boldsymbol{b}$ must be different from any row vector in $\boldsymbol{C}(i)_{ur}$ to compensate the loss of non-conjugate vector by deleting "$\boldsymbol{a}$" from $\boldsymbol{A}_b$. The restriction above must be held by all i $(1\le i\le n)$.

5 Example of function construction

Now, we present an example of the construction of eight dimension-balanced H-function that is highly nonlinear and has even term distribution. Bent function output is obtained by spread Walsh-Hadamard matrix with some columns complemented. The complemented matrix is:

$$\boldsymbol{H}=\begin{bmatrix}
1&1&1&0&1&0&1&0&1&1&1&0&1&0&1&1\\
1&0&1&1&1&1&1&1&1&0&1&1&1&1&1&0\\
1&1&0&1&1&0&0&1&1&1&0&1&1&0&0&0\\
1&0&0&0&1&1&0&0&1&0&0&0&1&1&0&1\\
1&1&1&0&0&1&0&1&1&1&1&0&0&1&0&0\\
1&0&1&1&0&0&0&0&1&0&1&1&0&0&0&1\\
1&1&0&1&0&1&1&0&1&1&0&1&0&1&1&1\\
1&0&0&0&0&0&1&1&1&0&0&0&0&0&1&0\\
1&1&1&0&1&0&1&0&0&0&0&1&0&1&0&0\\
1&0&1&1&1&1&1&1&0&1&0&0&0&0&0&1\\
1&1&0&1&1&0&0&1&0&0&1&0&0&1&1&1\\
1&0&0&0&1&1&0&0&0&1&1&1&0&0&1&0\\
1&1&1&0&0&1&0&1&0&0&0&1&1&0&1&1\\
1&0&1&1&0&0&0&0&0&1&0&0&1&1&1&0\\
1&1&0&1&0&1&1&0&0&0&1&0&1&0&0&0\\
1&0&0&0&0&0&1&1&0&1&1&1&1&1&0&1
\end{bmatrix}$$

The Table 2 is the 16×16 matrix serial number and conjugate group numbers from which we get vectors of $\boldsymbol{T}_{\mathrm{e}}$ matrix (if the table items in the upper 8 rows is zero, it belongs to non-conjugate vector collection $\boldsymbol{C}_{ur}(1)$; if the table items in the upper 8 rows is non zero, it indicates the conjugate group number of $\boldsymbol{C}_r(1)$):

Table 2 Revising position of original function output

$\boldsymbol{C}_r(1)$ or $\boldsymbol{C}_{ur}(1)$	1	2	3	0	9	0	0	0
$\boldsymbol{C}_r(2)$ or $\boldsymbol{C}_{ur}(2)$	1	2	3	0	0	0	0	35
$\boldsymbol{C}_r(3)$ or $\boldsymbol{C}_{ur}(3)$	1	2	0	0	8	0	24	0
$\boldsymbol{C}_r(4)$ or $\boldsymbol{C}_{ur}(4)$	1	0	2	0	3	0	0	27
$\boldsymbol{C}_r(5)$ or $\boldsymbol{C}_{ur}(5)$	1	2	3	0	9	0	0	0
$\boldsymbol{C}_r(6)$ or $\boldsymbol{C}_{ur}(6)$	1	0	2	0	5	0	0	30
$\boldsymbol{C}_r(7)$ or $\boldsymbol{C}_{ur}(7)$	1	0	1	0	6	0	27	0
$\boldsymbol{C}_r(8)$ or $\boldsymbol{C}_{ur}(8)$	1	1	0	3	5	0	0	0
Row number of matrix	0	1	2	15	4	15	13	10
Column number of matrix	0	0	0	0	1	3	10	11

And the resulting function output is:

$$\begin{aligned}\mathrm{True}(\boldsymbol{f})=\{&\underline{0},\underline{0},\underline{0},0,1,0,1,0,1,1,1,0,1,0,1,\underline{0},\\
&1,0,1,1,\underline{0},1,1,1,1,0,1,1,1,1,1,0,\\
&1,1,0,1,1,0,0,1,1,1,0,1,1,0,0,0,\\
&1,0,0,0,1,1,0,0,1,0,0,0,1,1,0,\underline{0},\\
&1,1,1,0,0,1,0,1,1,1,1,0,0,1,0,0,\\
&1,0,1,1,0,0,0,0,1,0,1,1,0,0,0,1,\\
&1,1,0,1,0,1,1,0,1,1,0,1,1,1,1,1,\\
&1,0,0,0,0,0,1,1,1,0,0,0,0,0,1,0,\\
&1,1,1,0,1,0,1,0,0,0,0,1,0,1,0,0,\\
&1,0,0,0,0,0,1,1,0,1,1,1,1,1,0,1\}\end{aligned}$$

in which the underlined make up the original function output from '1' to '0'. The resulting function is SAC, and its nonlinearity has the upper bound 116, algebra degree is 7 and term distribution is even. Its algebra expression is:

$$\begin{aligned}\boldsymbol{f}(\boldsymbol{x})=&1+x_0x_1+x_2+x_0x_1x_2+x_3+x_0x_1x_3+x_2x_3+x_4\\
&+x_0x_1x_4+x_0x_2x_4+x_1x_2x_4+x_3x_4+x_0x_1x_3x_4+x_0x_2x_3x_4\\
&+x_1x_2x_3x_4+x_0x_1x_2x_3x_4+x_5+x_0x_1x_5+x_2x_5+x_0x_1x_2x_5\\
&+x_3x_5+x_0x_1x_3x_5+x_2x_3x_5+x_4x_5+x_0x_1x_4x_5+x_1x_2x_4x_5\\
&+x_3x_4x_5+x_0x_1x_3x_4x_5+x_0x_2x_3x_4x_5+x_1x_2x_3x_4x_5+x_6\\
&+x_0x_1x_6+x_2x_6+x_0x_1x_2x_6+x_3x_6+x_0x_1x_3x_6+x_2x_3x_6\\
&+x_4x_6+x_0x_1x_4x_6+x_0x_2x_4x_6+x_1x_2x_4x_6+x_3x_4x_6\\
&+x_0x_1x_3x_4x_6+x_0x_2x_3x_4x_6+x_1x_2x_3x_4x_6+x_0x_1x_2x_3x_4x_6\\
&+x_5x_6+x_0x_1x_5x_6+x_2x_5x_6+x_0x_1x_2x_5x_6+x_3x_5x_6\\
&+x_0x_1x_3x_5x_6+x_2x_3x_5x_6+x_4x_5x_6+x_0x_1x_4x_5x_6\\
&+x_0x_2x_4x_5x_6+x_1x_2x_4x_5x_6+x_3x_4x_5x_6+x_0x_1x_3x_4x_5x_6\\
&+x_0x_2x_3x_4x_5x_6+x_1x_2x_3x_4x_5x_6+x_7+x_0x_1x_7+x_2x_7\\
&+x_0x_1x_2x_7+x_3x_7+x_0x_1x_3x_7+x_2x_3x_7+x_4x_7+x_0x_1x_4x_7\\
&+x_0x_2x_4x_7+x_1x_2x_4x_7+x_3x_4x_7+x_0x_1x_3x_4x_7\\
&+x_0x_2x_3x_4x_7+x_1x_2x_3x_4x_7+x_0x_1x_2x_3x_4x_7+x_5x_7\\
&+x_0x_1x_5x_7+x_2x_5x_7++x_0x_1x_2x_5x_7+x_3x_5x_7+x_0x_1x_3x_5x_7\\
&+x_2x_3x_5x_7+x_0x_2x_3x_5x_7+x_0x_1x_2x_3x_5x_7+x_4x_5x_7\\
&+x_0x_1x_4x_5x_7+x_0x_2x_4x_5x_7+x_1x_2x_4x_5x_7+x_3x_4x_5x_7\\
&+x_1x_3x_4x_5x_7+x_6x_7+x_0x_1x_6x_7+x_2x_6x_7+x_0x_1x_2x_6x_7\\
&+x_3x_6x_7+x_0x_1x_3x_6x_7+x_2x_3x_6x_7+x_4x_6x_7+x_0x_1x_4x_6x_7\\
&+x_0x_2x_4x_6x_7+x_1x_2x_4x_6x_7+x_3x_4x_6x_7+x_0x_1x_3x_4x_6x_7\\
&+x_0x_2x_3x_4x_6x_7+x_1x_2x_3x_4x_6x_7+x_0x_1x_2x_3x_4x_6x_7+x_5x_6x_7\\
&+x_0x_1x_5x_6x_7+x_2x_5x_6x_7+x_0x_1x_2x_5x_6x_7+x_3x_5x_6x_7\\
&+x_0x_1x_3x_5x_6x_7+x_2x_3x_5x_6x_7+x_0x_2x_3x_5x_6x_7\\
&+x_0x_1x_2x_3x_5x_6x_7+x_4x_5x_6x_7+x_0x_1x_4x_5x_6x_7\\
&+x_0x_2x_4x_5x_6x_7+x_1x_2x_4x_5x_6x_7+x_3x_4x_5x_6x_7\\
&+x_1x_3x_4x_5x_6x_7\end{aligned}$$

6 Conclusions

There exists a strong restraint relation between function characteristics now. Some constructions are only algebraic and do not have much satisfying characteristics. When constructing of a high dimension function from a low dimension function, simple iteration is usually done. By

studying constructions in the past, this paper presents a set of new thoughts that systematically use three methods of convolute (concatenate), split and revising function output to build an optimized function. For super high dimension functions, the methods above should be assisted by multi-grade function construction. Their compound usage can make a solid base for the designation of a good cryptology system.

References

1. Feng Deng-guo, Pei Ding-yi, Guide of Cryptology, Beijing: Science Publishing House, 1999 (in chinese)
2. Feng Deng-guo, Wu Wen-ling, Design and Analysis of Block Cipher, Beijing: Tsinghua University Pubilshing House, 2000 (in Chinese)
3. Dobbertin H., Construction of Bent functions and balanced Boolean functions with high nonlinearity, IEE Proceedings, 1989: 1436
4. Yang Yi-xian, N dimension H-Boolean function, Journal of Beijing University of Posts and Telecommunications, 1988 11(3): 1–9 (in Chinese)
5. Yang Yi-xian, Xing Yu-sen, N dimension H-Boolean function (II), Journal of Electronic Science, 1997, 19(2): 214–216 (in Chinese)
6. Zhang Wen-zheng, Study of some design criterion of Boolean function, Communication Secrecy, 1994(2): 68–84 (in Chinese)
7. Qiu Xian-jie, Some study on Bent function construction, Journal of Xiangtan University in Natural Science, 2002, 24(2): 16–19 (in Chinese)
8. Qin Jing, Construction of odds dimension Boolean function and their cryp ology property, Journal of Shandong University (Science), 2002，32(2): 127–130 (in Chinese)
9. Ji Qing-bing, Zhang Zhi-rang, Remark on construction of high nonlinearity balance Boolean function, Journal of Chongqing College of Posts and Telecommunications, 2004, 16(1): 79–82 (in Chinese)
10. Zeng Ben-xing, Li shi-qu, Li Kun, A decomposition formula of Walsh spectrum of a class of Boolean functions and its applications, Chinacrypt'98, Beijng: Science Publishing House, 1998: 217–220 (in Chinese)
11. Kurosawa K., Satoh T., Design of SAC/PC(l) of order k Boolean functions and three other cryptographic criteria, Advances in Cryptology-Eurocrypto'97, Springer-Verlag, 1998
12. Johnson T., Pasalic E., A construction of resilient functions with high nonlinearity, Lectures Notes in Computer Science, Springer-Verlag, 2000
13. Zhu Ling, He Min, Nonlinear balance Boolean function and its construction. Communication Secrecy, 1998, (2): 60–63 (in Chinese)
14. Sun Lin-hong, Ye Din-feng, Lü Wang-shu., Construction of high nonlinearity Boolean function, Journal of Graduated School in CAS, 2003, 20(4): 441–445 (in Chinese)
15. Wen Qiao-yan, Niu Xin-xin, Yang Yi-xian, Boolean function in modern Crypology, Beijing: Science Publishing House, 2000 (in Chinese)