Bent functions at the minimal distance and algorithms of constructing linear codes for CDMA [1]

Andrey V. Pavlov apavlov.nsk@gmail.com
Novosibirsk State University, Novosibirsk, RUSSIA

1 Introduction

In this paper we study linear codes for CDMA (Code Division Multiple Access). This is the standard for the 3rd Generation cellular communications systems. In this standard bent functions are used for constructing codes of constant amplitude. This application allows to decrease PAPR (peak-to-average power ratio) coefficient as much as possible. Such codes consist of vectors of values for bent functions.

Let us give the paper structure. In this section we consider some definitions and facts about bent functions. In the second section we give some information about affine equivalent bent functions. In the third section we give a simple method for constructing bent functions at the minimal distance from the given one. In the fourth section we briefly discuss CDMA. In the fifth section we give some known facts on linear codes for CDMA and present a new algorithm for constructing such codes. Our codes of small lengths (obtained with the algorithm) have the best known parameters.

Let us give some definitions and known facts connected to bent functions. Denote by $\mathrm{dist}(f,g) = |\{x : f(x) \neq g(x), x \in E^n\}|$ Hamming distance between Boolean functions f and g. Denote by E^n a *n-ary binary cube*. Let $\mathcal{F}_n$ be a set of all Boolean functions in n variables. By $\oplus$ denote the sum modulo 2.

Definition 1. *A Boolean function f with even number of variables is a bent function if it is on the maximal possible distance from all affine functions.*

Denote by $\mathfrak{B}_n$ the class of bent functions in n variables. For two Boolean functions f and g in n variables, denote by $D(f,g)$ the set of vectors for which they differ.

[1] The author is supported by the RF President grant for young Russian scientists (MK-1250.2009.1)

Definition 2. *A Boolean function f in n variables is called affine on the set $D \subseteq E^n$ if there exist $w_0 \in E^n$, $c \in E$ such that for any $x \in D$ we have $f(x) = \langle w_0, x \rangle \oplus c$.*

Remind that the set $L \subseteq E^n$ is called an *affine subspace* if $L = x_0 \oplus U$, where x_0 is a vector from E^n and U is a subspace in E^n. In [1] it is proved
Theorem 1 *Let $f, g \in \mathcal{F}_n$, $f \in \mathfrak{B}_n$, $dist(f, g) = 2^{n/2}$. Then $g \in \mathfrak{B}_n$ if and only if the set $D(f, g)$ is an affine subspace and function f is affine on $D(f, g)$.*

2 Affine equivalent bent functions

Boolean functions f and g in n variables are called *affine equivalent* if there exist a nonsingular matrix A of order n, vector b of length n and affine function ℓ in n variables such that it is true $g(x) = f(Ax \oplus b) \oplus \ell(x)$ for any x.

Definition 3. *Let $f \in \mathfrak{B}_n$. Then vector r in $\mathbb{Z}^{2^n}$ is called a* distance spectrum *for bent function f, if the component i of vector r is equal to the number of bent functions at distance i from the function f.*

Statement 1. *Distance spectra for affine equivalent bent functions are the same.*

Proof: The following equalities take a place: $\mathrm{dist}(f(x), g(x)) = \mathrm{dist}(f(Ax \oplus b), g(Ax \oplus b))$, $\det(A) \neq 0$; $\forall h \in \mathcal{F}_n \ \ \mathrm{dist}(f, g) = \mathrm{dist}(f \oplus h, g \oplus h)$. As far as the class of bent functions is closed under affine transforms of variables and under addition of affine functions, then from these equalities the statement follows. □

Hence, to obtain various distance spectra for bent functions it is enough to find distance spectrum for one bent function from each class of affine equivalence.

3 Exhaustive search of bent functions at the minimal distance from the given one

From Theorem 1 it follows that if we want to find all bent functions at the minimal distance from the given one it is enough to find all affine subspaces of dimension $n/2$, on which the given bent function becomes affine.

Definition 4. *Basis of subspace with dimension $n/2$ consists of $n/2$ rows of an echelon matrix over $\mathbb{Z}_2$ such that:*
1) every next line ends with number of zeros running in succession smaller then

has the previous line;
2) under each leading one (first one in a line) there are zeros in a column;
3) other elements are any.

We give examples of such a matrix and bent functions at the minimal distance.

Example 1. Let us take a bent function in 6 variables: $f_1(x) = x_1x_2x_3 \oplus x_1x_4 \oplus x_2x_5 \oplus x_3x_6$. For constructing bent function at the minimal distance we find affine subspace $L = v \oplus U$ such that f_1 becomes affine on it. By algorithm given further, we construct bases of all subspaces U and all v such that f_1 is affine on $L = v \oplus U$. Let us give an example for v, for a basis matrix of subspace U and for the basis matrix of orthogonal subspace. $v = \begin{pmatrix}0 & 1 & 0 & 0 & 0 & 0\end{pmatrix}$,

$$A_U = \begin{pmatrix}1 & 0 & 1 & 0 & 0 & 0\\ 1 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\end{pmatrix}, A_{U^\perp} = (a_{ij}) = \begin{pmatrix}0 & 0 & 0 & 0 & 0 & 1\\ 0 & 1 & 0 & 0 & 0 & 0\\ 1 & 0 & 1 & 1 & 0 & 0\end{pmatrix}.$$

Now we construct an indicator. It is a conjunction of $n/2$ items of the following form. The item number i is constructed from i-th row of matrix $A_{U^\perp}$ as follows: $(a_{i1}x_1 \oplus a_{i2}x_2 \oplus ... \oplus a_{in}x_n \oplus 1)$. Hence, $I_L(x) = I_U(x \oplus v) = I_U(x_1, x_2 \oplus 1, x_3, x_4, x_5, x_6) = (x_6 \oplus 1)(x_2 \oplus 1 \oplus 1)(x_1 \oplus x_3 \oplus x_4 \oplus 1) = x_1x_2x_6 \oplus x_2x_3x_6 \oplus x_2x_4x_6 \oplus x_2x_6 \oplus x_1x_2 \oplus x_2x_3 \oplus x_2x_4 \oplus x_2$. Thus we have bent function $f_2(x) = f_1(x) \oplus I_L(x) = x_1x_2x_3 \oplus x_1x_4 \oplus x_2x_5 \oplus x_3x_6 \oplus x_1x_2x_6 \oplus x_2x_3x_6 \oplus x_2x_4x_6 \oplus x_2x_6 \oplus x_1x_2 \oplus x_2x_3 \oplus x_2x_4 \oplus x_2$ at the minimal distance $2^{n/2} = 8$ from f_1.

Algorithm 1. (Exhaustive search of all affine subspaces of dimension $n/2$ such that Boolean function becomes affine on them)
1) Input: Boolean function $f(x)$ (in our case it is a bent function).
2) for any v in E^n we take a new function: $g(x) = f_v(x) = f(x \oplus v)$.
3) after that $g(x)$ is normalised: $g'(x) = g(x) \oplus g(0)$.
Now we search subspaces such that the function g' becomes linear on them. For each function g' we recurrently construct all basis matrices of size $n/2 \times n$.
4) let A_1 be a matrix with one row.
5) Suppose that we have constructed the matrix A_{i-1} of size $(s-1) \times n$, where $s - 1 < n/2$. Let us construct a matrix A_i of size $s \times n$. For this we add one more vector to the basis.

We run through all vectors in lexicographical order and put each vector u to the matrix A_{i-1} in such a way that matrix A_i satisfies to conditions $1-3$ from definition 4.

Then for each such vector u we check two conditions: (a) $g(x \oplus u) = g(x) \oplus g(u)$; (b) vector v lexicographically precedes to the vector $u \oplus x \oplus v$, where x runs through the subspace generated by rows of matrix A_{i-1}.

We put u to the matrix A_i only if all these conditions are satisfied. If there are no suitable vectors then it is a deadlock of recursion brunch.

4 CDMA

In this section we briefly discuss general model for communication (CDMA). For more details see [3]. In this model a binary data vector $c = (c_0, c_1, ..., c_{m-1})$ is input to orthogonal transform, where $m = 2^n$. The output is $S_c(t) = \sum_{i=1}^{m-1}(-1)^{c_i} f_i(t)$, where $f_i(t)$ are orthogonal functions of time t, $1 \leqslant t \leqslant m$. For CDMA the orthogonal transform is a Walsh-Hadamard transform:

$$WH(m) = \begin{pmatrix} WH(m-1) & WH(m-1) \\ WH(m-1) & -WH(m-1) \end{pmatrix}, WH(1) = 1.$$

Thus $S_c(t) = \sum_{i=1}^{m-1}(-1)^{c_i} WH(m)_{it}$. *Peak-to-average power ratio* is defined as $PAPR(c) = \frac{1}{m} max_t |S_c(t))|$, $1 \leqslant PAPR(c) \leqslant m$. So we can see that $PAPR(c)$ can be as large as m in a communication system using an orthogonal transform. This results in more expensive and inefficiently used components. Thus, there is a task to decrease PAPR as much as possible.

5 Linear codes for CDMA

There is a problem how we can control the PAPR of transmissions? We can use coding. A subset C in E^m is called a *binary code* of length m. The elements of code are called *codewords*. *The code distance* is equal to the minimal Hamming distance between different codewords. *Linear* $[m, k, d]$*-code* is a linear subspace $C \subseteq E^m$ of dimension k with the code distance d. We can construct code $C \subset E^m$ in which every word has small PAPR. One of such codes is the code based on vectors of values of bent functions. Such a code is called *constant amplitude code*. It has the smallest PAPR equal to 1.

Further we do not distinguish a function and its vector of values.

We introduce the following definition.

Definition 5. *For bent function f in n variables a code C of length 2^n we call SPB-code (i.e. saving property "bent") if function $f \oplus c$ is bent for any its codeword c.*

Statement 2. *Maximal sizes of SPB-codes for affine equivalent bent functions are the same.*

Proof: Suppose we have a bent function f and SPB-code C of maximal size for it. Let us construct all bent functions of the form $g = f \oplus c$, where c in C, and apply affine transform A for them. As far as the class of bent functions is closed under affine transforms we have bent functions $g'(x) = (f \oplus c)(Ax \oplus b) = f(Ax \oplus b) \oplus c(Ax \oplus b)$. Let us add $f(Ax \oplus b)$ to g' and we get SPB-code C' for bent function g'. It is easy to notice that this code has a maximal size. □
Further we study only linear SPB-codes.

Earlier (implicitly) linear SPB-codes were already considered. For constructing linear SPB-codes in [3] it is suggested to use construction of McFarland [2] $f(x, y) = \langle x, \pi(y)\rangle \oplus g(y)$, where $x, y \in E^{n/2}$, $g(y)$ is a Boolean function in $n/2$ variables, π is a permutation on $E^{n/2}$. Let us consider a linear SPB-code of length 2^n. This code consists of vectors of values of functions $h(x, y) = g(y)$ and all affine functions in n variables. Dimension of this code is equal to $k = 2^{n/2} + n/2$, code distance $d = 2^{n/2}$. For example for any bent function in McFarland class in 6 variables we have linear SPB-code with parameters $[2^6, 11, 8]$ and for bent function in 8 variables we have linear SPB-code with parameters $[2^8, 20, 16]$. Further we give an algorithm for constructing linear SPB-codes of more large size.
Algorithm 2. 1) Input: bent function f;
2) We add f to the list of functions ($functionList$);
3) We construct all affine subspaces of dimension $n/2$ on which the given bent function becomes affine (for that we use algorithm 1). We add them to $list$;
4) Further we call recursive function **findCode**(f, $list$, $functionList$);

findCode(f, $list$, $functionList$)
1) Input: bent function f, list of affine subspaces on which f becomes affine, list of bent functions;
2) For each affine subspace from $list$ we construct bent function g at the minimal distance from f;
3) If g is linear independent to all functions in $functionList$ we add g to $functionList$;
4) Now we form a $newList$ of subspaces. For each affine subspace from $list$ we check the following condition. If function g becomes affine on this subspace, then we add it to $newList$;
5) Then we call recursive function **findCode**(g, $newList$, $functionList$);
We use this algorithm for constructing linear SPB-codes for some bent functions in 6 and 8 variables. Note that our linear SPB-codes are not necessarily optimal. Further we give the table with affine nonequivalent bent functions and dimensions of the corresponding linear SPB-codes.

n	bent function	dim
6	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_1x_4 \oplus x_2x_6 \oplus x_3x_4 \oplus$ $\oplus x_3x_5 \oplus x_3x_6 \oplus x_4x_5 \oplus x_4x_6$	15
6	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_1x_2 \oplus x_1x_4 \oplus x_2x_6 \oplus x_3x_5 \oplus x_4x_5$	15
6	$x_1x_2x_3 \oplus x_1x_4 \oplus x_2x_5 \oplus x_3x_6$	15
6	$x_1x_2 \oplus x_3x_4 \oplus x_5x_6$	15
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_1x_4x_7 \oplus x_3x_5 \oplus$ $\oplus x_2x_7 \oplus x_1x_5 \oplus x_1x_6 \oplus x_4x_8$	29
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_3x_5 \oplus x_2x_6 \oplus$ $\oplus x_2x_5 \oplus x_1x_7 \oplus x_4x_8$	28
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_3x_5 \oplus x_1x_3 \oplus$ $\oplus x_1x_4 \oplus x_2x_7 \oplus x_6x_8$	30
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_3x_5 \oplus x_2x_6 \oplus$ $\oplus x_2x_5 \oplus x_1x_2 \oplus x_1x_3 \oplus x_1x_4 \oplus x_7x_8$	30
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4x_6 \oplus x_3x_5 \oplus x_1x_6 \oplus x_2x_7 \oplus x_4x_8$	28
8	$x_1x_2x_7 \oplus x_3x_4x_7 \oplus x_5x_6x_7 \oplus x_1x_4 \oplus x_3x_6 \oplus$ $\oplus x_2x_5 \oplus x_4x_5 \oplus x_7x_8$	29
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_3x_4 \oplus x_2x_6 \oplus x_1x_7 \oplus x_5x_8$	28
8	$x_1x_2x_3 \oplus x_2x_4x_5 \oplus x_1x_3 \oplus x_1x_5 \oplus x_2x_6 \oplus x_3x_4 \oplus x_7x_8$	30
8	$x_1x_2x_3 \oplus x_1x_4 \oplus x_2x_5 \oplus x_3x_6 \oplus x_7x_8$	29
8	$x_1x_2 \oplus x_3x_4 \oplus x_5x_6 \oplus x_7x_8$	28

Code distances of these linear SPB-codes are equal to the minimal possible distance $2^{n/2}$ between bent functions in n variables. In contrast to the method based on construction of McFarland, our method substantially depends on the concrete form of a bent function.

References

[1] N.A. Kolomeec, A.V. Pavlov Properties of bent functions at the minimal distance // Prikladnaya Diskretnaya Matematika. 2009. V. 2. No 4. P. 5–20.(in Russian)

[2] R.L. McFarland A family of difference sets in non-cyclic groups // J. Combin. Theory. Ser. A. 1973. V. 15. No. 1. P. 1–10.

[3] K.G. Paterson Sequences For OFDM and Multi-code CDMA: two problems in algebraic Coding Theory // Sequences and their applications. - Seta 2001. Second Int. Conference (Bergen, Norway, May 13-17, 2001). Proc. Berlin: Springer, 2002. P. 46–71.