Enumeration of Balanced Symmetric Functions over $GF(p)$

Shaojing Fu[1], Chao Li[1] and Longjiang Qu[1] Ping Li[1]

Department of Mathematics and System Science, Science College of National University of Defence Technology, Changsha, China, 410073
shaojing1984@yahoo.cn

Abstract. It is proved that the construction and enumeration of the number of balanced symmetric functions over $GF(p)$ are equivalent to solving an equation system and enumerating the solutions. Furthermore, we give an lower bound on number of balanced symmetric functions over $GF(p)$, and the lower bound provides best known results.
Key words: symmetric functions, balanced functions, Permutation.

1 Introduction

Since symmetry guarantees that all of the input bits have equal status in a very strong sense, symmetric Boolean functions display some interesting properties. A lot of research about symmetry in characteristic 2 has been previously done. Y.X.Yang and B.Guo [1] studied the balanced symmetric functions and correlation immune symmetric functions. S.Maitra and P.Sarkar [2] studied the maximum nonlinearity of symmetric Boolean function on odd number of variables. A.Canteaut and M.Videau [3] established the link between the periodicity of simplified value vector of an symmetric Boolean functions and its degree. A.Braeken, B.Preneel [4] studied Algebraic immunity of Symmetric Boolean function.

On the other hand, it is natural to extend various cryptographic ideas from GF(2) to other finite fields of characteristic $p > 2$, $GF(p)$ or GF(p^n), p being a prime number. For example, [5] and [6] studied the correlation immune and resilient functions on $GF(p)$. Also, [7] and [8] investigated the generalized bent functions on GF(p^2). Li and Cusick [9] first introduced the strict avalanche criterion over $GF(p)$. In [10], they generalized most results of [11] and determined all the linear structures of symmetric functions over $GF(p)$. Recently, Cusick and Li [12] give a lower bound for the number of balanced symmetric functions over $GF(p)$, and they show the existence of nonlinear balanced symmetric functions. In [13], Pinhui Ke improved the lower bound, then he showed that the number of n-variable balanced symmetric functions over GF(p) is not less than the number of solutions of an given equation systems over $\mathbb{Z}^*$.

In this paper, we prove that enumeration of the number of balanced symmetric functions over $GF(p)$ is equivalent to solve the equation system in [13] and we give formulas to count balanced symmetric functions over $GF(p)$. Then based on our formulas and Cusick's method, the lower bound on the number of n-variable balanced symmetric functions over finite fields $GF(p)$ presented in [13] is improved.

2 Preliminaries

In this paper, $\mathbb{Z}$ is the set of positive integers, $\mathbb{Z}^*$ is the set of positive integers, and p is a odd prime number. Let $GF(p)$ be the finite field of p elements, and $\mathrm{GF}(p)^n$ be the vector space of dimension n over $GF(p)$. An n-variable function $f(x_1, x_2, \cdots, x_n)$ can be seen as a multivariate polynomial over $GF(p)$, that is,

$$f(x_1, x_2, \cdots, x_n) = \sum_{k_1,k_2,\cdots,k_n=0}^{n} a_{k_1,k_2,\cdots,k_n} x_1^{k_1} x_2^{k_2} \cdots x_n^{k_n}$$

where the coefficients $a_{k_1,k_2,\cdots,k_n}$ are a constant in $GF(p)$. This representation of f is called the algebraic normal form (ANF) of f. The number $k_1 + k_2 + \cdots + k_n$ is defined as the degree of term with nonzero coefficient. The greatest degree of all the terms of f is called the Algebraic degree of f, denoted by $\deg(f)$.

The function $f(x)$ is called an affine function if $f(x) = a_1x_1 + a_2x_2 + \cdots + a_nx_n + a_0$. If $a_0 = 0$, $f(x)$ is also called a linear function.We will denote by F_n the set of all functions of n variables and by L_n the set of affine ones. We will call a function nonlinear if it is not in L_n.

Definition 1. *f: $GF(p)^n \to GF(p)$ is balanced if $\#\{x \in GF(p)^n | f(x) = k\} = p^{n-1}$ for any $k = 1, 2, \cdots, p-1$.*

Definition 2. *f:$GF(p)^n \to GF(p)$ is a symmetric if for any permutation π on $\{1, 2, \cdots, n\}$, we have $f(x_1, x_2, \cdots, x_n) = f(x_{\pi(1)}, x_{\pi(n)}, \cdots, x_{\pi(n)})$.*

Denote the set of permutations on $\{1, 2, \cdots, n\}$ by S_n. Then we define the following equivalence relation on $\mathrm{GF}(p)^n$: for any $x = (x_1, \cdots, x_n)$, $y = (y_1, \cdots, y_n) \in GF(p)^n$. we say x and y are equivalent, if there exists a permutation $\pi \in S_n$ such that $(y_1, \cdots, y_n) = (x_{\pi(1)}, \cdots, x_{\pi(n)})$. (by abuse of notation we write $y = \pi(x)$). Let $\tilde{x} = \{y | \exists \pi \in S_n, \pi(x) = y\}$, Let $\bar{x} = (\bar{x}_1, \bar{x}_2, \cdots, \bar{x}_n)$ be the representative of $\tilde{x}$, where $0 \le \bar{x}_1 \le \bar{x}_2 \le \cdots \le \bar{x}_n \le p-1$. Obviously, we have $\tilde{x} = \tilde{y} \Leftrightarrow \bar{x} = \bar{y}$. It is easy to show that f: $\mathrm{GF}(p)^n \to \mathrm{GF}(p)$ is symmetric if $f(x) = f(y)$ whenever $\tilde{x} = \tilde{y}$.

Definition 3. *Let $\bar{x} = (\underbrace{0, \cdots, 0}_{i_0}, \underbrace{1, \cdots, 1}_{i_1}, \cdots, \underbrace{p-1, \cdots, p-1}_{i_{p-1}})$ be representative of the classes $\tilde{x}$, then the vector $(i_0, i_1, \cdots, i_{p-1})$ is called the corresponding vector of $\tilde{x}$.*

Let the corresponding vector of $\tilde{x}$ be $(i_0, i_1, \cdots, i_{p-1})$, then $\#\tilde{x} = \frac{n!}{i_0! i_1! \cdots i_{p-1}!}$. The number of different equivalence classes $\tilde{x}$ is the number of solutions of the linear equation $i_0 + i_1 + \cdots + i_{p-1} = n$, where i_k is the number of times k appears in $\bar{x}$. We know that the number of solutions to this linear diophantine equation is the same as the number of n-combinations of a set with p elements, that is $C(n + p - 1, n)$.

3 Enumeration of Balanced Symmetric Functions

From the definition of symmetric function, we know that symmetric function has the same value for any n-tuple in the same equivalence classes. So in order to get symmetric function,we should partition the vectors in the $C(n+p-1,n)$ equivalence classes into p groups, and the vectors in the same equivalence class must be in the same group. We start with the definition of similar equivalence class.

Definition 4. *$\tilde{x}$ and $\tilde{y}$ are called the similar equivalence class, if there exists $\pi \in S_n$, such that $\tilde{x}$'s corresponding vector $(i_0, i_1, \cdots, i_{p-1})$ and $\tilde{y}$'s corresponding vector $(j_0, j_1, \cdots, j_{p-1})$ satisfies $(i_0, i_1, \cdots, i_{p-1}) = (j_{\pi(0)}, j_{\pi(1)}, \cdots, j_{\pi(p-1)})$. The class-sets of $\tilde{x}$ is the set constituted by all the similar equivalence class of $\tilde{x}$.*

Let all the $C(n+p-1,n)$ classes are divided into N class-sets $(\Omega_1, \Omega_2, \cdots, \Omega_N)$ by using the similar equivalence relation, $M_i = \#\Omega_i$ and $T_i = \#\{\tilde{x} | \forall \tilde{x} \in \Omega_i\}$(Because any equivalence classes in a class-sets have same cardinality, T_i is denoted without misapprehending).

For a fixed solution $\tilde{x} \in \Omega_i$, let $(i_0, i_1, \cdots, i_{p-1})$ be the corresponding vector of $\tilde{x}$, and m_l be the number of times that l appears in $\{i_0, i_1, \cdots, i_{p-1}\}$. Then it is easy to show that $M_i = \frac{p!}{m_0! m_1! \cdots m_n!}$, $T_i = \frac{n!}{i_0! i_1! \cdots i_{p-1}!}$.

Consider the equation system Φ:

$$\Phi : \begin{cases} \sum_{i=0}^{N} x_{i,j} \cdot T_i = p^{n-1}, & 1 \le j \le p \\ \sum_{i=0}^{p} x_{i,j} = M_i, & 1 \le i \le N \\ x_{i,j} \in \mathbb{Z},\ x_{i,j} \ge 0. \end{cases}$$

In [13], Pinhui Ke presented the following result.

Theorem 1. *[13]The number of n-variable balanced symmetric functions over GF(p) is not less than the number of solutions of equation system Φ.*

From theorem 1, we know that Pinhui Ke only prove that the number of solutions of equation system Φ is a lower bound. Now we will give a method to construct balanced symmetric functions by using the solutions of Φ, and give formulas to count balanced symmetric functions over $GF(p)$.

First, let M_Φ be the number of solutions of Φ, and the solutions be $X_1, X_2, \cdots, X_{M_\Phi}$, where

$$X_k = \begin{pmatrix} x_{1,1}^{(k)} & x_{1,2}^{(k)} & \cdots & x_{1,p}^{(k)} \\ x_{1,1}^{(k)} & x_{1,2}^{(k)} & \cdots & x_{1,p}^{(k)} \\ \vdots & \vdots & \ddots & \vdots \\ x_{N,1}^{(k)} & x_{N,2}^{(k)} & \cdots & x_{N,p}^{(k)} \end{pmatrix}, 1 \le k \le M_\Phi$$

Theorem 2. *Let N_n be the number of n-variable balanced symmetric functions over GF(p), then $N_n = \sum_{k=1}^{M_\Phi} \prod_{i=1}^{N} \frac{M_i!}{\prod_{j=1}^{p} x_{i,j}^{(k)}!}$.*

Proof. To construct balanced symmetric functions, basically, we try to divide p^n vectors into p groups (on each group, the function has the same value), such that each group has p^{n-1} vectors. Of course, the vectors in the same equivalence class must be in the same group since the function is symmetric. For a fixed solution X_k, we construct p groups $A_1, A_2, \cdots, A_p$ as follow,

(1) Select $x_{1,1}^{(k)}$ equivalence classes from Ω_1, $x_{2,1}^{(k)}$ equivalence classes from Ω_2, $\cdots$, $x_{N,1}^{(k)}$ equivalence classes from Ω_N, regard the vectors in these equivalence classes as the vectors of A_1.

(2) Select $x_{1,2}^{(k)}$ equivalence classes from the rest $M_1 - x_{1,1}^{(k)}$ equivalence classes of Ω_1, $x_{2,2}^{(k)}$ equivalence classes from the rest $M_2 - x_{2,1}^{(k)}$ equivalence classes of Ω_2, $\cdots$, $x_{N,2}^{(k)}$ equivalence classes from the rest $M_N - x_{N,1}^{(k)}$ equivalence classes of Ω_N, regard the vectors in these equivalence classes as the vectors of A_2.

$\cdots$

(t) Select $x_{1,t}^{(k)}$ equivalence classes from the rest $M_1 - \sum_{j=1}^{t-1} x_{1,j}^{(k)}$ equivalence classes in Ω_1, $x_{2,t}^{(k)}$ equivalence classes from the rest $M_2 - \sum_{j=1}^{t-1} x_{2,j}^{(k)}$ equivalence classes in Ω_2, $\cdots$, $x_{N,t}^{(k)}$ equivalence classes from the rest $M_N - \sum_{j=1}^{t-1} x_{2,j}^{(k)}$ equivalence classes in Ω_N, regard the vectors in these equivalence classes as the vectors of A_t.

$\cdots$

$(p-1)$ Select $x_{1,p-1}^{(k)}$ equivalence classes from the rest $M_1 - \sum_{j=1}^{p-2} x_{1,j}^{(k)}$ classes in Ω_1, $x_{2,p-1}^{(k)}$ equivalence classes from the rest $M_2 - \sum_{j=1}^{p-2} x_{2,j}^{(k)}$ equivalence classes in Ω_2, $\cdots$, $x_{N,p-1}^{(k)}$ equivalence classes from the rest $M_N - \sum_{j=1}^{p-2} x_{N,j}^{(k)}$ equivalence classes in Ω_N, regard the vectors in these equivalence classes as the vectors of A_{p-1}.

(p) there are $x_{1,p}^{(k)}$ equivalence classes in the rest of Ω_1, $x_{2,p}^{(k)}$ equivalence classes in the rest of Ω_2, $\cdots$, $x_{N,p}^{(k)}$ equivalence classes in the rest of Ω_N, regard the vectors in these equivalence classes as the vectors of A_p.

Let $f(x)$: $\{x | f(x) = j-1\} = A_j$, $1 \le j \le p$.

It is obvious that $f(x)$ is a balanced symmetric function, and the number of balanced symmetric function constructed by the solution equals the ways to select $A_1, A_2, \cdots, A_p$, that is $\prod_{i=1}^{N} \frac{M_i!}{\prod_{j=1}^{p} x_{i,j}^{(k)}!}$.

Given two solutions $X_{k_1} \ne X_{k_2}$, without lost of generality, Let $x_{1,1}^{(k_1)} \ne x_{1,1}^{(k_2)}$, the corresponding A_1 constructed byX_{k_1}, X_{k_2} are different, so the balanced symmetric functions are different. Hence, the total number of symmetric functions constructed by the construction above is $\sum_{k=1}^{M_\Phi} \prod_{i=1}^{N} \frac{M_i!}{\prod_{j=1}^{p} x_{i,j}^{(k)}!}$

Now we show that any n-variable balanced symmetric functions can be constructed by the construction above. Given a balanced symmetric function $f(x)$, let $A_j = \{x | f(x) = j-1\}$ and $A_j^* = \{\tilde{x} | x \in A_j\}$, and $x_{i,j} = \#\{A_j^* \bigcap \Omega_i\}$, then It is easy to show that $x_{i,j} (1 \le i \le N, 1 \le j \le p)$ is a solution of the equation system Φ.

So we get the count.

Example 1. Let $n = 3$, $p = 5$, we can get $N = 3$, $M_1 = 5$, $M_2 = 20$, $M_3 = 10$, and $M_\Phi = 281$. Then the number of 3-variable balanced symmetric functions over $GF(5)$ is $1.24419789850356 \times 10^{20}$.

When p and n become larger, it is hard to solve the equation system Φ. In this case, we solve another easier equation, and we can obtain an improved lower bound of [13] by solving this equation when p is not a proper divisor of n.

Consider the equation with restricted conditions:

$$\Theta : \sum_{i=0}^{N} z_i T_i = 0, z_i \in \mathbb{Z}, |z_i| \leq \frac{M_i}{p}$$

Let $\Psi = \{(y_1^{(1)}, y_2^{(1)}, \cdots, y_N^{(1)}), (y_1^{(2)}, y_2^{(2)}, \cdots, y_N^{(2)}), \cdots (y_1^{(W)}, y_2^{(W)}, \cdots, y_N^{(W)})\}$ be the set of the solutions of Θ whose most right nonzero component is positive. Then we can get the following lower bound on the number of n-variable balanced symmetric Functions.

Theorem 3. *If p is not a proper divisor of n, then* $N_n \geq \prod_{i=1}^{N} \frac{M_i!}{((\frac{M_i}{p})!)^p} +$

$$\sum_{t=1}^{\frac{p-1}{2}} A(p, 2t) \sum_{1 \leq k_1, k_2, \cdots, k_t \leq W} \prod_{i=1}^{N} \frac{M_i!}{((\frac{M_i}{p})!)^{p-2t} \prod_{j=1}^{t} (\frac{M_i}{p} - y_1^{(k_j)})! (\frac{M_i}{p} + y_1^{(k_j)})!}$$

where $A(p, 2t) = \frac{n!}{(n-k)!}$.

Proof. If $p \nmid n$, let m_l be the number of times that l appears in $\{i_0, i_1, \cdots, i_{p-1}\}$. Then

$$\begin{cases} m_0 + m_1 + \cdots + m_n = p \\ 0 \times m_0 + 1 \times m_1 + \cdot + n \times m_n = n \end{cases}$$

so p is a proper divisor of $M_i = \frac{p!}{m_0! m_1! \cdots m_n!}$ for any $1 \leq i \leq N$. Then It is easy to show that the matrix

$$\triangle = \begin{pmatrix} \frac{M_1}{p} & \frac{M_1}{p} & \cdots & \frac{M_1}{p} \\ \frac{M_2}{p} & \frac{M_2}{p} & \cdots & \frac{M_2}{p} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{M_N}{p} & \frac{M_N}{p} & \cdots & \frac{M_N}{p} \end{pmatrix}$$

a solution of Φ.

For any $k = 1, 2, \cdots, W$

$$\begin{pmatrix} \frac{M_1}{p} & \cdots & \frac{M_1}{p} & \frac{M_1}{p} - y_1^{(k)} & \frac{M_1}{p} & \cdots & \frac{M_1}{p} & \frac{M_1}{p} + y_1^{(k)} & \frac{M_1}{p} & \cdots & \frac{M_1}{p} \\ \frac{M_2}{p} & \cdots & \frac{M_2}{p} & \frac{M_2}{p} - y_2^{(k)} & \frac{M_2}{p} & \cdots & \frac{M_2}{p} & \frac{M_2}{p} + y_2^{(k)} & \frac{M_2}{p} & \cdots & \frac{M_2}{p} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \frac{M_N}{p} & \cdots & \frac{M_N}{p} & \frac{M_N}{p} - y_N^{(k)} & \frac{M_N}{p} & \cdots & \frac{M_N}{p} & \frac{M_N}{p} + y_N^{(k)} & \frac{M_N}{p} & \cdots & \frac{M_N}{p} \end{pmatrix}$$

is also a solution of Φ.

Select t vectors $(y_1^{(k_1)}, y_2^{(k_1)}, \cdots, y_N^{(k_1)})$, $(y_1^{(k_2)}, y_2^{(k_2)}, \cdots, y_N^{(k_2)})$,$\cdots$ $(y_1^{(k_t)}, y_2^{(k_t)}$, $\cdots$,$y_N^{(k_t)})$ from Ψ, then add these t vectors to any t columns of the matrix $\triangle$ respectively. At the same time, subtract $(y_1^{(k_1)}, y_2^{(k_1)}, \cdots, y_N^{(k_1)})$, $\cdots, y_N^{(k_2)})$, $\cdots, (y_1^{(k_t)}, y_2^{(k_t)}, \cdots, y_N^{(k_t)})$ to another t columns of the matrix $\triangle$. By the selecting we can obtain $A(p, 2) \times \#\Psi \times A(p-2, 2) \times \#\Psi$ new matrices, these matrices are also solutions of Φ.

Now We distinguish $\frac{p+1}{2}$ case.

(0)If we change 0 columns of the matrix $\triangle$, then we can construct $\prod_{i=1}^{N} \frac{M_i!}{((\frac{M_i}{p})!)^p}$ balanced symmetric functions.

(1)If we change 2 columns of the matrix $\triangle$, then we can construct

$$\sum_{1 \leq k_1 \leq W} p(p-1) \prod_{i=1}^{N} \frac{M_i!}{((\frac{M_i}{p})!)^{p-2t}(\frac{M_i}{p} - y_1^{(k_1)})!(\frac{M_i}{p} + y_1^{(k_1)})!}$$

balanced symmetric functions.

$\cdots$

(t)If we change 2t columns of the matrix $\triangle$, then we can construct

$$\sum_{1 \leq k_1, k_2, \cdots, k_t \leq W} A(p, 2t) \prod_{i=1}^{N} \frac{M_i!}{((\frac{M_i}{p})!)^{p-2t} \prod_{j=1}^{t}(\frac{M_i}{p} - y_1^{(k_j)})!(\frac{M_i}{p} + y_1^{(k_j)})!}$$

balanced symmetric functions.

$\cdots$

$(\frac{p+1}{2})$ If we change $p-1$ columns of the matrix $\triangle$, then we can construct

$$\sum_{1 \leq k_1, k_2, \cdots, k_{\frac{p-1}{2}} \leq W} A(p, 2t) \prod_{i=1}^{N} \frac{M_i!}{(\frac{M_i}{p})! \prod_{j=1}^{\frac{p-1}{2}}(\frac{M_i}{p} - y_1^{(k_j)})!(\frac{M_i}{p} + y_1^{(k_j)})!}$$

balanced symmetric functions.

Hence we get the lower bound.

At the end of this section, for $n = 3$ and $p = 5$, we compare the lower bound obtained by [12], [13], and Theorem 3 in the following table:

Table 1. the number of 3-variable balanced symmetric functions over GF(5)

	[12]	[13]	Theorem 3
upper bound	$4.15779151788 \times 10^{18}$	$2.653066968552 \times 10^{19}$	$9.7809924135924 \times 10^{19}$

4 Conclusion

In this paper, we obtain some counting results about balanced symmetric functions over finite field GF(p), and we get a lower bound by finding solutions of an equation system. With an example, we show that this bound is better than the known results, but when p is a proper divisor of n, it is still an open problem to get a lower bound on the number of balanced symmetric functions. Besides, Finding more solutions to improve this lower bound may be another interesting problem for the future research.

Acknowledgments

The work in this paper is supported by the National Natural Science Foundation of China (NO:60573028).

References

1. Y.X.Yang, B.Guo. Further enumerating Boolean functions of cryptographic significance. J.Cryptology, 8(3),pp.115-122, 1995
2. S.Maitra, P.Sarkar. Maximum nonlinearity of symmetric Boolean functions on odd number of variables. IEEE Trans.on Infor.theory, 48(9), pp.2626-2630, 2002.
3. A.Canteaut, M.Videau. Symmetric Boolean functions. IEEE Trans.on Infor. theory, 51(8), pp.2791-2811, 2005.
4. A.Braeken, B.Preneel. On the algebraic immunity of symmetric Boolean functions. in INDOCRYPT2005,LNCS 3797, Berlin, Germay:Springer-Verlag , pp.35-48, 2005.
5. Mulan Liu, Peizhong Lu and G.L. Mullen. Correlation-Immune Functions over Finite Fields. IEEE Trans. on Information Theory 44 (1998), 1273-1276.
6. Yupu Hu and Guozhen Xiao. Resilient Functions Over Finite Fields. IEEE Trans. on Information Theory 49 (2003), 2040-2046.
7. Keqin Feng and Fengmei Liu. New Results On The Nonexistence of Generalized Bent Functions. IEEE Trans. on Information Theory 49 (2003), 3066-3071.
8. P.V. Kumar, R.A. Scholtz, and L.R. Welch. Generalized Bent Functions and Their Properties. J. Combinatorial Theory (A) 40 (1985), 90-107.
9. Yuan Li and T.W. Cusick. Strict Avalanche Criterion over Finite Fields. J. Math. Crypt. 1(2007): 65-78.
10. Yuan Li and T.W. Cusick. Linear Structures of Symmetric Functions over Finite Fields. Information Processing Letters 97 (2006), 124-127.
11. Ed Dawson and Chuan-kun Wu. On the Linear Structure of Symmetric Boolean Functions. Australasian Journal of Combinatorics 16 (1997), 239-243.
12. T.W. Cusick, Yuan Li and P. Stanica. Balanced symmetric functions over GF(p). IEEE Trans. on Information Theory 54(3), pp. 1304-1307, 2008.
13. P.H.Ke. Improve lower bound on the number of balanced symmetric functions over GF(p). http://eprint.iarc.org/2008/165.pdf.