The Analysis of affinely Equivalent Boolean Functions*

Qing-shu Meng Min Yang Huan-guo Zhang
Yuzhen Liu

October 21, 2005

Abstract

By Walsh transform, autocorrelation function, decomposition, derivation and modification of truth table, some new invariants are obtained. Based on invariant theory, we get two results: first a general algorithm which can be used to judge if two boolean functions are affinely equivalent and to obtain the affine equivalence relationship if they are equivalent. For example, all 8-variable homogenous bent functions of degree 3 are classified into 2 classes; second, the classification of the Reed-Muller code $R(4,6)/R(1,6), R(3,7)/R(1,7)$, which can be used to almost enumeration of 8-variable bent functions.
Keywords: boolean functions, affine equivalent, affine general group

1 Introduction

Boolean functions are used widely in science and engineering, like in circuit design, cryptography and error-correction coding. The linearly classification of boolean functions is meaningful for the following two reasons: first, equivalent functions have similar properties (like Hamming weight distribution in error-correction coding, same nonlinearity in cryptography). second, the number of representatives is much less than the number of boolean functions. Out of the need of circuit design, the classification of boolean functions under affine linear group was discussed much often in 60s in 20th century[1, 2, 3]. Recently the analysis of linearly equivalence of boolean functions was discussed in several papers[4, 5, 6, 7, 8]. Fuller and Millan disclosed the linear equivalence between the output functions of the AES s-box by getting the linear equivalence relationship, but the method is not efficient in the case of bent function. Method in paper[8] are not efficient too in bent functions case though it improve the efficiency of Fuller-Millan algorithm. In eurocrypt'03, a toolbox is developed to analyze linear equivalence between bijective s-box or s-box with small $n-m$, where n, m are number of inputs and outputs respectively, and

*supported by the National Natural Science Foundation of China (60373087,90104005)

thus the toolbox cant deal with boolean functions, where $m = 1$. In attacking HFE problem(hidden fields equation), Geiseleman, Meier and Steinwandt gave an collum-wise method, but the method is not efficient in boolean function with uneven truth table. Other papers on classify boolean functions can be found in papers[9, 10, 11]. To authors' knowledge, to judge if two functions are equivalent and how to get the equivalent relationship if they are equivalent is not known in general case.

In this paper, an algorithm is given which can efficiently solve the two above problems in general case. The basic tools we use is Walsh transform, autocorrelation function, derivation function, decomposition, and modification of truth table.

2 Preliminary

For each subset $s \subseteq \{1, 2, \cdots, n\}$, there exists a corresponding vector $(s_1, s_2, \cdots, s_n)$ of dimension n by letting $s_i = 1$ if element i is in s else letting $s_i = 0$. And the vector $(s_1, s_2, \cdots, s_n), s_i \in \{0, 1\}$ for $i = 1, 2, \cdots, n$ can be denoted by an integer s whose 2-adic expansion is just the vector $(s_1, s_2, \cdots, s_n)$. Obviously, the set, the vector and the integer are isomorphic. In this paper, if confusion is not caused, we will use the three notations for description convenience. Denote by F_2 the Galois field with two elements $\{0, 1\}$ and denote by F_2^n the vector space over F_2. Denote by $p_n = F_2[x_1, x_2, \cdots, x_n]/(x_1^2 - x_1, \cdots, x_n^2 - x_n)$ the algebra of all functions $F_2^n \to F_2$. For each subset $s \subseteq \{1, 2, \cdots, n\}$, denote $\prod_{i \in s} x_i \in p_n$ by x^s. The algebraic normal form of a Boolean function $F_2^n \to F_2$ can be written as $f(x) = \sum_{s=0}^{2^n-1} a_s x^s$, where $a_s \in F_2$. The degree of $f(x)$ is defined by

$$\max_{s \in \{0,1,\cdots,2^n-1\}, a_s \neq 0} H(s),$$

where $H(s)$ is the Hamming weight of vector s. The set of functions whose degree less than or equal to 1 is called affine functions. The set $\{f(x) | deg(f) \leq r\}$ is denoted by $R(r, n)$. Denote by $R(r, n)/R(s, n)$ the set $\{f(x) + R(s, n) | s < deg(f) \leq r\}$.

Denote by $GL(n, 2)$ the set of all nonsingular matrix of order n, i.e. the general linear group. Denote by $AGL(n, 2)$ the group $\{(A, b) | A \in GL(n, 2), b \in F_2^n\}$. The group operation is defined by

$$(A, u)(B, w) = (AB, A(w) + u)$$

$$(A, u)^{-1} = (A^{-1}, A^{-1}(u),$$

where $(A, u), (B, w) \in AGL(n, 2)$.

The action of group $AGL(n, 2)$ on Boolean functions is defined by:

$$\begin{array}{rc} c: & p_n \to p_n \\ by: & f(x) \to f(xA + b) \end{array},$$

where $c = (A, b) \in AGL(n, 2)$.

Two functions $f(x), g(x) \in R(r,n)/R(s,n)$ are called equivalent if there exists $(A,b) \in AGL(n,2)$ such that $f(x) = g(xA+b) \mathrm{mod} R(s,n)$. An invariant of $R(r,n)/R(s,n)$ is a mapping M from $R(r,n)/R(s,n)$ to a set such that for any two equivalent functions $f(x), g(x) \in R(r,n)/R(s,n)$, $M(f) = M(g)$ holds.

If $s = 1$, we get $f(x) = g(xA+b) + lx$, in this paper we will mainly discuss how to get (A,b) and l when the two functions are given.

3 Basic Transforms

3.1 Walsh Transform and Autocorrelation Function

Definition 1: Define

$$s_{(f)}(w) = \sum_{x \in F_2^n} (-1)^{f(x)} (-1)^{w \cdot x}$$

be the Walsh spectrum of $f(x)$ at vector w, where $f(x) \in p_n, w \in F_2^n$.

The transform is called the Walsh transform.

Definition 2: Define the functions $c_f(s) = \sum_{x=0}^{2^n-1} (-1)^{f(x)} (-1)^{f(x+s)}$ be the autocorrelation function of $f(x)$, where $f(x) \in p_n, s \in F_2^n$.

The following two propositions are well known. And the fact that the distributions of absolute Walsh spectra and autocorrelation function are invariant is due to Preneel's work[12].

Proposition 1: Let $f(x), g(x) \in p_n$ be two functions such that $g(x) = f(xA+b)+lx$, then for any $w \in F_2^n$, $s_{(g)}(w) = (-1)^{(l+w) \cdot bA^{(-1)}} s_{(f)}((l+w)A^{-1T}$

Corollary 1: The Walsh spectrum of $f(x)$ at i is equal to the Walsh spectrum of $g(x)$ at j, where $j = l + iA^T$. Therefore the deficiency of the rank of vectors with same spectrum between two equivalent functions is at most 1. The distribution of absolute value of Walsh spectra of $f(x)$ is same to that of $g(x)$.

Proposition 2: Let $f(x), g(x) \in p_n$ be two functions such that $g(x) = f(xA+b) + lx$, then for any given $s \in F_2^n$, $c_g(s) = (-1)^{l \cdot s} c_f(sA)$.

Corollary 2: The autocorrelation function of $f(x)$ at j is equal to the autocorrelation function of $g(x)$ at i, where $j = iA$. Therefore the ranks of vectors with same absolute autocorrelation function value are same between two equivalent functions. The distribution of absolute value of autocorrelation function of $f(x)$ is same to that of $g(x)$.

3.2 Derivation

For any boolean function $f(x) \in R(r,n)$, define its derivation function as $D_a(f) = f(x) + f(x+a)$. Similarly we can define two-order derivative function as $D_{a,b}(f) = f(x) + f(x+a) + f(x+b) + f(x+a+b)$. By definition, it is easy to get following properties:

Property 1: $D_{a,b}(f) = D_a(f) + D_b(f) + D_{a+b}(f)$.

Property 2: $D_a(f \circ B) = D_{aA}(f) \circ B$, where $B \in AGL(n,2)$. similarly, $D_{a,b}(f \circ B) = D_{aA,bA}(f) \circ B$, where $B \in AGL(n,2)$.

Proposition 3: If $f(x) \in R(r,n)/R(s,n)$, then $D_a(f \circ B) = (D_{aA}(f)) \circ B$ mod $R(s-1,n)$, where $B = (A,b) \in AGL(n,2)$. If M is an invariant of $R(r-1,n)/R(s-1,n)$, then $M(D_a(f \circ B)) = M((D_{aA}(f)) \circ B)$, so $\{M(D_a(f)|a \in F_2^n\}$ is an invariant of $R(r,n)/R(s,n)$.

Remark: The derivation function is used by hou [10] in classification of $R(3,7)/R(2,7)$ and by Brier and Langevin[13] in classification of $R(3,9)/R(2,9)$. Proposition 5 is an extension of their result.

3.3 Decomposition

Let $f(x), g(x) \in R(r,n)$ be two functions such that $g(x) = f(xA+b)$ mod $R(s,n)$. If $f(x) = (x_1+1)f_0(x') + x_1 f_1(x')$, where $x' = (x_2, \cdots, x_n)$, then $g(x) = (x \cdot C_1 + b_1 + 1)f_0(x'') + (x \cdot C_1 + b_1)f_1(x'')$, where $C_1, C_2, \cdots, C_n$ are the columns of the matrix A, and $x'' = (x \cdot C_2 + b_2, \cdots, x \cdot C_n + b_n)$. Obviously, $f_0(x'), f_1(x')$ are affinely equivalent to $f_0(x''), f_1(x'')$ module $R(s, n-1)$ respectively. Similar result hods for two-vector based decomposition.

Proposition 4:If $f(x)$ is decomposed into two subfunctions at vector a, then $g(x)$ can be decomposed into two subfunctions at vector d, where $d = aA^T$, such that the two subfunctions of $f(x)$ are equivalent to those of $g(x)$ module $R(s, n-1)$.

Proof decomposing $f(x)$ at vector a means to decompose $f(x)$ according to $a \cdot x = 0$ or $a \cdot x = 1$. It means to decompose $g(x)$ according to $a \cdot (xA+b) = 0$ or $a \cdot (xA+b) = 1$. Let $a = (a_1, \cdots, a_n)$, $x = (x_1, \cdots, x_n)$, $A = (C_1, \cdots, C_n)$, where $C_i, i = 1, \cdots, n$ are the columns of matrix A, then

$$\begin{aligned}
a \cdot (xA+b) =& a \cdot (x \cdot C_1, \cdots, x \cdot C_n) + a \cdot b \\
=& a_1(x \cdot C_1) + \cdots + a_n(x \cdot C_n) + a \cdot b \\
=& a_1(x_1 C_{1,1} + \cdots + x_n C_{n,1}) + \cdots + a_n(x_1 C_{1,n} + \cdots + x_n C_{n,n}) + a \cdot b \\
=& x_1(a_1 C_{1,1} + \cdots + a_n C_{1,n}) + \cdots + x_n(a_1 C_{n,1} + \cdots + a_n C_{n,n}) + a \cdot b \\
=& x_1(a \cdot R_1) + \cdots + x_n(a \cdot R_n) + a \cdot b \\
=& x \cdot (a \cdot R_1, \cdots, a \cdot R_n) + a \cdot b \\
=& x \cdot (aA^T) + a \cdot b
\end{aligned}$$

As $a \cdot b$ is a constant, that is ,we decompose $g(x)$ at vector $d = aA^T$.

Proposition 5: If M is an invariant of $R(r,n-1)/R(s,n-1)$, then the set $\{\{M(f_{ax=0}), M(f_{ax=1})\}|a \in F_2^n\}$ is an invariant of $R(r,n)/R(s,n)$.

Remarks: The basic idea of the decomposition of a function can be found early in Maiorana's paper[9], which made the classification of R(6,6)/R(1,6) possible early in 90s in 20th century. And recently it is used by Brier and Langevin[13] to classify R(3,9)/R(2,9).

3.4 The Modification of Truth Table

Definition 3[14]: For a function $f(x)$, define its 1-local connection functions as

$$\{f_i(x)|f_i(x) = \{ \begin{array}{cc} f(x) & x \neq i \\ f(x)+1 & x = i \end{array}, i = 0, 1, \cdots, 2^n - 1\}.$$

similarly 2-local connection functions can be defined.

Proposition 6[7]: Let $f(x), g(x) \in R(r,n)$ be such that $g(x) = f(xA+b) + lx$, then $g_j(x) = f_i(xA+b) + lx$, where $jA = (i+b), i = 0, 1, \cdots, 2^n - 1$. Similar result holds for two-local connection functions.

Proposition 7: If M is an invariant of $R(n,n)/R(1,n)$, then $\{M(f_i(x)|i \in F_2^n\}$ is an invariant of $R(r,n)/R(1,n)$.

4 The Analysis of affinely Equivalent Boolean Functions

4.1 Algorithm

input: two functions $f(x), g(x) \in R(n,n)$,

output: A, b, and l, if $g(x) = f(xA+b) + lx$ else output they are not equivalent.

1. Calculate the Walsh spectra and autocorrelation function of $f(x), g(x)$ respectively. Compare the distribution of absolute value of Walsh spectra and absolute autocorrelation function of $f(x)$ with those of $g(x)$ respectively. If the two functions have two same distributions, then goto step 2 else they are not linearly equivalent, exit.

2. Denote the autocorrelation value of $g(x)$ at unit vector e_i by $c_g(e_i)$. By proposition 2, there exists at least one element $v \in \{v|abs(c_f(v)) = abs(c_g(e_i))\}$ such that $v = e_iA$ holds. Let $i = 1, 2, \cdots, n$, there are n equations.

3. Decompose $f(x)$ at unit vector e_i, and calculate the invariant of the two subfunctions, denote it by $de_{e_i}(f)$. by proposition 4, there exists at least one element $v \in \{v|de_v(g) = de_{e_i}(f)\}$ such that $v = e_iA^T$ holds. Let $i = 1, 2, \cdots, n$, we get another n equations.

4. Calculate the invariant of the derivation function of $g(x)$ at unit vector e_i, and denote it by $d_{e_i}(g)$. By proposition 3, there exists at least one element $v \in \{v|d_v(f) = d_{e_i}(g)\}$ such that $v = e_iA$ holds. Let $i = 1, 2, \cdots, n$, we get another n equations. These n equations should be consistent to the n equations obtained in step 2 and 3, else the two functions are not equivalent.

5. Denote by $g_{e_i}(x)$ the local connection function of $g(x)$ at unit vector e_i, and denote its invariant by $lv_{e_i}(g)$. By proposition 6, there exists at least one element $v \in \{v|lv_v(f) = lv_{e_i}(g)\}$ such that $v = e_iA+b$ holds. Let $i = 1, 2, \cdots, n$, we get another n equations.

6. Denote by $s_{(f)}(e_i)$ the absolute value of Walsh spectrum of $f(x)$ at unit vector e_i. By corollary 1, there exists at least one element $v \in \{v|abs(s_{(g)}(v)) = s_{(f)}(e_i)\}$ such that $v = e_iA^T + l$ holds. Let $i = 1, 2, \cdots, n$, we get n equations.

7. By step $2 \sim 4$, we get matrix A. By step 5, we can obtain b. By step 6, we can get l. With all these parameters(usually there are many choice for some parameters), we can use the method in [5] to filter some impossible choice. Finally we can verify them by checking if the equation $g(x) = f(xA+b) + lx$ holds.

4.2 Analysis of the Algorithm

Walsh transform, autocorrelation function ,derivation, decomposition and modification of truth table are the basic transforms to a boolean functions. Walsh transform and autocorrelation functions can be done by fast Hadamard transform. Derivation transform lows the degree of the two functions, and decomposition transform lows the number of variables. Thus these two transforms low the complexity of our problem. Modification of truth table gives us more equations with same affine equivalence and thus it is more possible to obtain the affine equivalence. By the step 5 and 6, it is unnecessary to enumerate parameters b, l. By above analysis, we say our algorithm is more efficient. However it is not easy to analyze the computation complexity.

By step 3 and 4, it is easy to address the bent functions case, and by step 2,3 and 4, we can deal with functions with uneven truth table. Therefore we say our algorithm is more general.

4.3 application

In paper [15], all 293760 homogeneous bent functions in 8-variable of degree 3 are classified into 14 classes under the action of permutation group(see [15]). However there exists no efficient algorithm for further classification under the action of $AGL(8,2)$. By above algorithm, they are classified into two classes. Functions 1, 10,14 are equivalent to each other, and functions 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13 are equivalent to each other. Here we give some relations. $f_1(x) = f_{10}(xA)$, $A = (18, 27, 188, 17, 124, 51, 248, 252)$, where the integers are the row vectors of the matrix. $f_1(x) = f_{14}(xA)$, $A = (206, 209, 94, 241, 76, 79, 245, 249)$. $f_2(x) = f_3(xA)$, $A = (103, 165, 219, 37, 222, 47, 202, 250)$. $f_2(x) = f_4(xA)$, $A = (179, 221, 35, 222, 55, 43, 156, 252)$.

5 Classification of Reed-Muller Code

Invariant is a good tool to classify set. If we know N, the number of equivalent classes under some equivalent relationship, and an invariant just takes N different values, then the set is already classified.

5.1 Classification of R(4,6)/R(1,6)

The number of orbits of R(4,6)/R(1,6) under the action of $AGL(6,2)$ is 2499 by hou's work[11]. The classification of R(4,6)/R(1,6) can be done as follows:

1. It is easy to get the four oribits of R(2,6)/R(1,6). By hou's work[10], their complementary functions are the four orbits of R(4,6)/R(3,6), denoted by $f_i + R(3,6), i = 0, 1, 2, 3$, where $f_0(x) = 0, f_1(x) = x_3x_4x_5x_6$, $f_2(x) = x_1x_2x_5x_6 + x_3x_4x_5x_6$, $f_3(x) = x_1x_2x_3x_4 + x_1x_2x_5x_6 + x_3x_4x_5x_6$.

2. By proposition 3, classify the four cosets $f_i + R(3,6), i = 0, \cdots, 3$ into 6,10,12,6 cosets of form $g_j + R(2,6), 2 < deg(g_j(x) \leq 4$ respectively. The invari-

ant of R(3,6)/R(1,6) used in proposition 3 is the distribution of absolute Walsh spectra. The basic time complexity of this step is $O(4 \times 2^{20})$.

3. By proposition 5 and 7,classify the 34 cosets $g_i + R(2,6), i = 0, 1, \cdots, 33$ into 2499 cosets of form $h_i(x) + R(1,6), 1 < deg(h_i(x)) \leq 4, i = 0, 1, \cdots, 2498$. The invariant of R(4,5)/R(1,5) used in proposition 5 is the distribution of absolute Walsh spectra and absolute autocorrelation function. The invariant of R(6,6)/R(1,6) used in proposition 7 is the distribution of absolute Walsh spectra and absolute autocorrelation function. For any combination of invariants given in this paper except the invariant in proposition 9, we can't get 2499 orbits. The basic complexity is $O(34 \times 2^{15})$.

5.2 Classification of R(3,7)/R(1,7)

The number of orbits of R(3,7)/R(1,7) under the action of $AGL(7, 2)$ is 179 by Hou's work[11]. All these 179 orbits can be obtained as follows:

1. By Hou's work[10], we can get 12 representatives of R(3,7)/R(2,7):$f_i(x)+R(2, 7)$.

2. By proposition 5 the coset $f_i(x) + R(2, 7), i = 0, 1, \cdots, 11$ can be classifed into 4,8,19,10,20,6,7,29,12,39,10,15 cosets of form $g_i(x) + R(1, 7)$ respectively. these are all 179 representatives. The invariant of R(3,6)/R(1,6) used in proposition 5 is the distribution of absolute Walsh spectra and absolute autocorrelation function.

By above two examples, it is very efficient to classify Reed-muller code for some parameters by invariant theory.

6 Conclusion

Based on some basic transforms, we give an algorithm which can be used to judge if two functions are equivalent and to get the equivalent relationship if they are equivalent in general case. This result also can be used for IP(isomorphism of polynomials) problem with one secret over F_2; second, $R(4, 6)/R(1, 6)$ and $R(3, 7)/R(1, 7)$ is classified efficiently by invariant theory. The direct application of this classification is the semi-enumeration of 8-variable bent functions[16].

Except transforms in this paper, finding other transforms is a useful work.

References

[1] Harrison. M.A, Counting theorems and their applications to classifications of switching functions, in A. Mukhopadhyay ed. Recent developments in switching theory, Acdemic press, new york, lodon,1971,pp 85-120.

[2] Harrison.M.A, On the classification of Boolean functions by the general linear and affine groups, J.soc.indust.appl.math. 12, 285-299,1964.

[3] Berlekamp.e.r, welch.l.r, weight distributions of the cosets of the (32,6) reed-muller code, ieee trans. Inform. Theory V.18,203-207,1972.

[4] A.biryukov, c.d.canniere,A. braeken,B. preneel, Toolbox for cryptanlysis:linear and affine equivalence algorithms, eurocrypt'03,lncs2656,33-50.

[5] Geiselmann. W., Meier. W., Steinwandt. R.. An attack on the isomorphisms of polynomials problem with one secret. International Journal of Information Security, 2003, 2(1): 59-64

[6] An braeken, yuri borissov,s. nikova, b.preneel, Classification of Boolean functions of 6 variables or less with respect to cryptographic properties, http://eprint.iacr.org,2004/248.

[7] Fuller. J., Millan. W.. Linear redundancy in S-box. In: Fast Software Encryption, LNCS 2887, Springer-Verlag, 2003, 74-86.

[8] Meng qingshu,zhang huanguo, The analysis of linear equivalence of boolean functions and its applications,chinese journal of computers,2004,11.1528-1532.(in Chinese)

[9] Maiorana.j.a, A classification of the cosets of the reed-muller code R(1,6), math. Comp.57,403-414,1991.

[10] Hou xiang-dong, GL(m,2) acting on R(r,m)/R(r-1,m), discrete mathematics,149(1996)99-122

[11] Hou xiang-dong, AGL(m,2) acting on R(r,m)/R(s,m), journal of algebra, 171(1995)921-938.

[12] B. Preneel, Analysis and Design of Cryptographic Hash Functions, Ph.D Thesis, KU Leuven(Belgium),February 1993.

[13] Eric brier, philippe langevin, Classification of Boolean cubic forms in nine variables, 2003 ieee information theory workshop 179-182.

[14] Millan w.,clark a.,dawson E.. Smart hill climbing finds better boolean functions. in Proceeding of the workshop on selected areas in cryptology 1997,ottawa,canada,1997,50-63.

[15] Qingshu Meng, Min Yang, Huanguo Zhang, et.al. A novel algorithm enumerating bent functions. http://eprint.iacr.org,2004/274

[16] Qingshu Meng, Huanguo Zhang, Jingsong Cui,et al. Semi-Enumeration of 8-Variable Bent Functions. http://eprint.iacr.org. 2005/100.