文章编号：1004-4353(2014)02-0130-04

源于密码学的组合数猜想问题中的数据特征

邓宇龙， 张晓朋， 郑玉军
（湖南科技学院数学与计算科学系 计算数学研究所，湖南 永州 425199）

摘要：通过在源于密码学的组合数猜想问题中提出权值表构造函数，给出了权值表构造算法和组合数分布的机器证明算法. 实验数据分析表明，旋转集合和反射集合内的元素具有同等的组合分布.

关键词：Boolean 函数；权值表构造函数；旋转；反射
中图分类号：O178　　**文献标识码**：A

On some data characteristics of combinatorial conjecture

DENG Yulong， ZHANG Xiaopeng， ZHENG Yujun
（*Institute of Computational Mathematics*, *Department of Mathematics and Computational Science*, *Hunan University of Science and Engineering*, *Yongzhou* 425199, *China*）

Abstract: By presenting the weight table constructed function of combination conjecture in cryptography, the weight table constructed algorithm and machine algorithm of combination distribution are given. We also conclude that the elements within the rotation set and reflection set is of the same combination distribution by the analysis of experimental data.

Key words: Boolean function; weight table constructed function; rotation; reflection

0 引言

Tu 和 Deng[1] 提出了源于密码学的组合数猜想：

猜想 1　设 $S_{t,k}=\{(a,b)\mid a,b\in \mathbf{Z}_{2^k-1}, a+b\equiv t(\bmod 2^k-1), w(a)+w(b)\leqslant k-1\}$，其中 $1\leqslant t\leqslant 2^k-2$，$k\geqslant 2$，$w(t)$ 是 t 的 Hamming 权，则集合 $S_{t,k}$ 的基数 $\sharp S_{t,k}\leqslant 2^{k-1}$.

根据文献[1]的猜想，Tu 和 Deng 获得了两类具有最优代数免疫的 Boolean 函数类：一类仍然是 Bent 函数类，另一类是具有最优代数次数和目前所知最好的非线性度(这个结果优于文献[2]给出的函数类非线性度)的平衡 Boolean 函数类. 但是，对于猜想 1 的证明，文献[1]仅仅只是给出了 transfer-matrix 算法，然后利用计算机辅助，机械地验证了 $k\leqslant 29$ 的情况，这与猜想 1 的结论还有很大的差距.

观察猜想 1 中集合 $S_{t,k}$ 的构造发现，b 由 a 唯一确定，因此设想猜想 1 的结论可以通过对整数 a 的个数的计算来获得. 源于这种想法，文献[3]把 a 划分成两类：第一类为 $a=0,1,\cdots,t$，$b=t-a$；第二类为 $a=t+v$，$b=2^k-1-v$，$v=1,2,\cdots,2^k-t-2$. 根据这两类划分，文献[3]证明了当 t 的权 $w(t)$ 较小时($w(t)\leqslant 3$)猜想 1 的结论成立，并且还指出 a 的个数的统计很大程度依赖于

收稿日期：2013-11-17　　**作者简介**：邓宇龙(1980—)，男，讲师，研究方向为调和分析、密码学.
基金项目：湖南科技学院科研项目(13XKYTA001)；国家自科科学基金资助项目(61202463)；湖南科技学院重点学科建设项目(13XKYTB002)

$$N_r^{(i_1,i_2,\cdots,i_s)}=\#\{x\mid 0\leqslant x\leqslant 2^k-1, w(2^{i_1}+2^{i_2}+\cdots+2^{i_s}+x)=r+w(x)\},$$

其中 $0\leqslant i_1<i_2<\cdots<i_s\leqslant k-1$. 显然，当 $r>s$ 时 $N_r^{(i_1,i_2,\cdots,i_s)}=0$ 且 $N_s^{(i_1,i_2,\cdots,i_s)}=2^{k-s}$，其他情况时 $N_r^{(i_1,i_2,\cdots,i_s)}$ 的获得较为困难. 文献[4]给出了 $N_r^{(i_1,i_2,\cdots,i_s)}$ 的机器算法，但最后的整理仍很困难. 有关 Tu-Deng 猜想问题的研究还可参见文献[5-6]. 设 $a\in \mathbf{Z}_{2^k-1}$，$a\neq 0$，并且令 $r(a,t)=w(a)+w(t)-w(a+t)$，则 $S_{t,k}=\{a\in \mathbf{Z}_{2^k-1}\mid r(a,t)>w(t)\}$. 于是，本文得到 Tu-Deng 猜想 1 的一个等价命题：

猜想 2[5]　设 $P_{t,k}=\dfrac{\#\mid S_{t,k}\mid}{2^k}$，则 $P_{t,k}\leqslant\dfrac{1}{2}$.

1　背景知识

设非负整数 x 的二进展开表达式为 $x=x_0+x_12+x_22^2+\cdots+x_{k-1}2^{k-1}=\sum_{i=0}^{k-1}x_i2^i$，其中 $x_i\in F_2=\{0,1\}$，k 为 x 的二进长度. 记 x 的 k 二进表示为 $x=(x_0x_1\cdots x_{k-1})$，则 $w(x)=\sum_i x_i$ 称为 x 的 Hamming 权. 容易证明：

引理 1[3]　$w(2^k-1-x)=k-w(x)$，$0\leqslant x\leqslant 2^k-1$；当 $x_i=1$ 时，$w(x+2^i)\leqslant w(x)$；当且仅当对任意的 i，$x_i+y_i\leqslant 1$ 时，有 $w(x+y)\leqslant w(x)+w(y)$；以及

$$w(x)=w(x-1)-i+1,\ x\equiv 2^i(\bmod 2^{i+1}),\ i=0,1,2,\cdots.$$

根据引理 1 有：

定义 1(权值表构造函数)　设二进长度为 k 的非负整数 x，y 对应的和 $x+y$ 的取值范围为 $0\sim 2^k-1$，$x=0$ 或 2^i，$i=0,1,\cdots,k-1$，$0\leqslant y<x$，则 $x+y$ 的权值表构造函数为

$$w(x+y)=\begin{cases}0, & x=0,\ y=0;\\ 1, & x=2^i,\ y=0;\\ w(x)+w(y)=1+w(y), & x=2^i,\ 0<y<x.\end{cases}$$

定义 2(旋转)　k 二进表示为 $(x_0x_1\cdots x_{k-1})$ 的 x 的旋转 $\rho_k^i(x)$ 定义为

$$\rho_k^i(x)=\rho_k^i(x_0x_1\cdots x_{k-1})=(x_{(0+i)\bmod k},\cdots,x_{(k-1+i)\bmod k}),$$

其中 $1\leqslant i\leqslant k$. 由 k 二进表示为 $(x_0x_1\cdots x_{k-1})$ 的 x 旋转形成的集合为

$$R_k(x)=R_k(x_0x_1\cdots x_{k-1})=\{\rho_k^i(x)=\rho_k^i(x_0x_1\cdots x_{k-1})\mid 1\leqslant i\leqslant k\}.$$

显然，$\rho_k^i(x)$ 只是 $(x_0x_1\cdots x_{k-1})$ 的一个重排，不会改变 Hamming 权，即 $w(x)=w(\rho_k^i(x))$.

定义 3(反射)　k 二进表示为 $(x_0x_1\cdots x_{k-1})$ 的 x 的反射 $\gamma(x)$ 定义为 $\gamma(x)=\gamma(x_0x_1\cdots x_{k-1})=(x_{k-1}x_{k-2}\cdots x_0)$. 显然，$w(\gamma(x))=w(x)$.

定义 4(二进元素补)　对于 k 二进表示为 $(x_0x_1\cdots x_{k-1})$ 的 x，如果存在 k 二进表示为 $(y_0y_1\cdots y_{k-1})$ 的 y，使得 $x+y$ 按二进制加法表示为 $x+y=(\underbrace{11\cdots1}_{k})=2^k-1$，则称 y 为 x 的二进元素补，记为 $\bar{x}$. 于是 $\bar{x}=2^k-1-x$，且 $w(\bar{x})=w(2^k-1-x)=k-w(x)$.

2　算法设计

本文采用 C++ 语言，通过计算机算法实现符合条件的元素个数的计算. 由于要搜寻的元素数据量呈 $2^i(1\leqslant i\leqslant k)$ 增加，算法中内存空间的分配需求较大，而计算机配置较低，因此本文只统计了二进长度较小的元素个数. 算法的实现由以下两个步骤完成：

步骤 1　权值表的构造. 首先，根据权值表构造函数，由于非负整数 x 的取值设定为 $2^i(0\leqslant i\leqslant$

$k-1$),因此在循环控制变量 i 小于二进长度 k 时,总有 $w(x)=1$. 其次,初始化二进长度为 k 的变量 x, $y=1$, y 由二进长度为 k 非负整数 x 控制;如果 $y<x$,根据权值表构造函数,由 $w(x+y)=w(x)+w(y)$ 计算 $x+y$ 的权值,否则选取 $x=2x$. 最后,当循环变量 i 的取值大于二进长度 k 时,完成 $w(x+y)$ 的计算,得到权值表.

步骤 2　元素个数的分类统计. 根据文献[3],把二进长度为 k 的非负整数 x 划分成两类:第一类为 $x=0,1,\cdots,t$, $y=t-x$;第二类为 $x=t+v$, $y=2^k-1-v$, $v=1,2,\cdots,2^k-t-2$. 首先,初始化循环控制变量 $x=0$, $y=t$. 如果 $x\leqslant t$,则计算 $w(x)+w(y)$ 的值. 假设 $w(x)+w(y)$ 的值为 i,则有 $S_{t,i}$ 的元素个数 $\sharp S_{t,i}=\sharp S_{t,i}+1$;否则 $x=x+1$, $y=y-1$,循环执行得到元素个数的统计分布表.

3　实验数据分析

本文选取 k 二进长度为 5 和 6 的数据进行分析,实验数据统计如表 1 和表 2 所示. 注意到 $1\leqslant t\leqslant 2^k-2$, $w(t)$ 是 t 的 Hamming 权, $w(a)+w(b)$ 是可能的权值, t 与 $w(a)+w(b)$ 所对应的数字是相应的取值个数统计.

表 1　$k=5$ 的组合分布情况

k	$t(w(t))$	$w(a)+w(b)$									同分布的其他可能 t 值			
		1	2	3	4	5	6	7	8	9				
5	1(1)	2	1	2	4	8	16	0	0	0	2	4	8	16
5	3(2)	0	4	2	5	10	4	8	0	0	6	12	17	24
5	5(2)	0	4	4	5	4	8	8	0	0	9	10	18	20
5	7(3)	0	0	8	4	10	5	2	4	0	14	19	25	28
5	11(3)	0	0	8	8	4	5	4	4	0	13	21	22	26
5	15(4)	0	0	0	16	8	4	2	1	2	23	27	29	30

表 2　$k=6$ 的组合分布情况

k	$t(w(t))$	$w(a)+w(b)$											同分布的其他可能 t 值										
		1	2	3	4	5	6	7	8	9	10	11											
6	1(1)	2	1	2	4	8	16	32	0	0	0	0	2	4	8	16	32						
6	3(2)	0	4	2	5	10	20	8	16	0	0	0	6	12	24	33	48						
6	5(2)	0	4	4	5	12	8	16	16	0	0	0	10	17	20	34	40						
6	7(3)	0	0	8	4	10	21	10	4	8	0	0	14	28	35	49	56						
6	9(2)	0	4	4	9	4	12	16	16	0	0	0	18	36									
6	11(3)	0	0	8	8	12	9	12	8	8	0	0	13	19	22	25	26	37	38	41	44	50	52
6	15(4)	0	0	0	16	8	20	10	5	2	4	0	30	39	51	57	60						
6	21(3)	0	0	8	12	6	13	6	12	8	0	0	42										
6	23(4)	0	0	0	16	16	8	12	5	4	4	0	29	43	46	53	58						
6	27(4)	0	0	0	16	16	12	4	9	4	4	0	45	54									
6	31(5)	0	0	0	0	32	16	8	4	2	1	2	47	55	59	61	62						

通过对表 1 和表 2 数据的分析,可得到 Tu-Deng 猜想问题中的如下一些数据分布特征:

定理 1(旋转的数据特征)　设 $\rho_k^i(t)=t_1$, $1\leqslant i\leqslant k-1$,则 $\rho_k^i(S_{t,j})=S_{t_1,j}$, $j=1,2,\cdots,2k-1$.

证明　因为 $t_1=\rho_k^i(t)=\rho_k^i((a+b)\bmod(2^k-1))=(\rho_k^i(a)+\rho_k^i(b))\bmod(2^k-1)$,所以

$$\rho_k^i(S_{t,j})=\rho_k^i(\{(a,b)\mid(a+b)\bmod(2^k-1)=t\text{ 且 }w(a)+w(b)=j\})=$$

$$\{(\rho_k^i(a),\rho_k^i(b))\mid(\rho_k^i(a)+\rho_k^i(b))\bmod(2^k-1)=\rho_k^i(t)\text{ 且 }w(\rho_k^i(a))+w(\rho_k^i(b))=j\}.$$

注意到 $\rho_k^i(t)=t_1$ 及 $w(\rho_k^i(a))=w(a)$，于是有 $\rho_k^i(S_{t,j})=S_{t_1,j}$.

定理 2(反射的数据特征)　设 $\gamma(t)=t_2$，则 $S_{t_2,j}=\gamma(S_{t,j})$.

证明　因为 $t_2=\gamma(t)=\gamma((a+b)\bmod(2^k-1))=(\gamma(a)+\gamma(b))\bmod(2^k-1)$，所以

$$\gamma(S_{t,j})=\gamma(\{(a,b)\mid(a+b)\bmod(2^k-1)=t\text{ 且 }w(a)+w(b)=j\})=$$
$$\{(\gamma(a),\gamma(b))\mid(\gamma(a)+\gamma(b))\bmod(2^k-1)=\gamma(t)\text{ 且 }w(\gamma(a))+w(\gamma(b))=j\}.$$

注意到 $\gamma(t)=t_2$ 及 $w(\gamma(a))=w(a)$，于是有 $\gamma(S_{t,j})=S_{t_1,j}$.

定理 3(二进元素补的数据特征)　设 $\bar{t}=2^k-1-t$，则 $\overline{S_{t,j}}=S_{t,2k-j}$.

证明　由于 $\bar{t}=\overline{(a+b)\bmod(2^k-1)}=(\bar{a}+\bar{b})\bmod(2^k-1)$ 且

$$w(\bar{a})+w(\bar{b})=w(2^k-1-a)+w(2^k-1-b)=k-w(a)+k-w(b)=2k-(w(a)+w(b)),$$

于是 $\overline{S_{t,j}}=\overline{\{(a,b)\mid(a+b)\bmod(2^k-1)=t\text{ 且 }w(a)+w(b)=j\}}=\{(\bar{a},\bar{b})\mid(\bar{a}+\bar{b})\bmod(2^k-1)=\bar{t}$ 且 $w(\bar{a})+w(\bar{b})=2k-(w(a)+w(b))=2k-j\}=S_{t,2k-j}$.

参考文献：

[1] Tu Ziren, Deng Yingpu. A conjecture on binary string and its application on constructing Boolean functions of optimal algebraic immunity[J/OL]. [2013-11-17]. http://eprint.iacr.org/2009/272.pdf.

[2] Carlet Claude, Feng Keqin. An infinite class of balanced functions with optimal algebraic immunity, good immunity to fast algebraic attacks and good nonlinearity[C]//Advances in Cryptology-ASIACRYPT 2008. Springer-Verlag, 2008:425-440.

[3] Thomas W Cusick, Li Yuan, Pantelimon. On a combinatorial conjecture[J/OL]. [2013-11-17]. http://eprint.iacr.org/2009/554.pdf.

[4] Zhang Xiaopeng, Deng Yulong. On a combinatorial conjecture[J]. Journal of Hunan University of Science and Engineering, 2013,34(12):58-61.

[5] Flori Jean Pierre, Randriam Hugues, Cohen Gerard, et al. On a conjecture about binary strings distribution [J/OL]. [2013-11-17]. http://eprint.iacr.org/2010/170.pdf.

[6] Flori Jean Pierre, Randriam Hugues. On the number of carries occurring in an addition mod 2^k-1[J]. Citation Information, 2012,12(4):601-647.