计数一阶相关免疫布尔函数 [1]

张建州 游志胜

(四川大学计算机学院图像图形研究所 成都 610064)

摘 要 基于列平衡矩阵的概念，利用组合数学中的容斥原理和图论中的子图概念，给出了在流密码学中有重要应用的一阶相关免疫布尔函数的一般计数公式。

关键词 布尔函数，相关免疫，计数，容斥原理，子图

中图号 O157.1, O153.2

1 引 言

自从 Siegenthaler[1,2] 在流密码学中引入相关免疫布尔函数的概念以后，人们对相关免疫函数的结构、构造和计数问题进行了广泛地研究，其中计数问题得到极大关注，但始终未能解决，文献 [3] 将其列为 20 个有待研究的问题之一。而计数各种不同类型的布尔函数也是组合数学研究的重要内容。

目前有关一阶相关免疫布尔函数计数问题的结果主要是一些上界和下界及递推关系。文献 [4] 给出了几个上界，由于从实用角度看下界比上界更为重要，因此，大量的文献研究它的下界 [4−10] 。在下界的研究中，一个主要的研究方法就是构造相关免疫函数类，但目前所构造的相关免疫函数绝大多数是线性结构函数，事实上，存在很多非线性结构的相关免疫函数，所以，通过研究下界很难给出一阶相关免疫函数的一般计数公式。利用递推关系研究一阶相关免疫函数的计数问题是另一个重要方面，文献 [8] 利用列平衡矩阵 [11] 来研究递推关系是一个很好的想法；文献 [12] 给出了构造列平衡矩阵的一般方法。文献 [13] 从 n 维立方体 Q_n 顶点着色的角度研究了相应的等价问题，利用 Polya 方法给出了一个一般的递推关系，但其递推式中含有难以确定的系数且没有一致性；同时，又利用叠置 (superposition) 方法给出了相关免疫布尔函数的计数母函数，但由于其难于计算而很难使用。这说明利用递推关系也难给出一个满意的结果。

本文基于列平衡矩阵的概念，利用组合数学中容斥原理 [14] 和图论 [15] 中子图概念，给出了一阶相关免疫布尔函数的一般计数公式。

2 相关免疫函数和列平衡矩阵

相关免疫布尔函数的定义有多种等价形式，这里只给出两种。

定义 1 n 元布尔函数 $f(x_1,\cdots,x_n)$ 称为一个相关免疫函数，如果对任意 $i(1\le i\le n)$ 和 $a(a=0$ 或 $1)$ 都有如下等式：

$$W(f(x_1,\cdots,x_n))=2W(f(x_1,\cdots,x_n)|x_i=a).$$

这里 $W(f)$ 和 $W(f|x_i=a)$ 分别表示 n 元和 $n-1$ 元布尔函数 $f(x_1,\cdots,x_n)$ 和 $f(x_1,\cdots,x_{i-1},a,x_{i+1},\cdots,x_n)$ 的 Hamming 重量，即 $W(f)=\sum_x f(x)$ 。

为了解决一阶相关免疫布尔函数的计数问题，下面引入布尔函数定序特征矩阵和列平衡矩阵的概念。

定义 2[8] 设 $f(x_1,\cdots,x_n)$ 是 n 元布尔函数，记 $w=W(f)$ ，则称向量集合

[1] 1997-10-27 收到， 1999-05-20 定稿
中国博士后科学基金和国家自然科学基金重点项目 (69732010) 支持课题

$$D=\{(d_1,\cdots,d_n)|f(d_1,\cdots,d_n)=1\}$$

为 $f(\bullet)$ 的特征集合。又记 $c_1=(c_{11},\cdots,c_{1n}),\cdots,c_w=(c_{w1},\cdots,c_{wn})$ 为集合 D 中按字典序排列的 $f(\bullet)$ 的一切特征向量，下面的 $w\times n$ 阶 0, 1 矩阵：

$$C=\begin{bmatrix}c_1\\c_2\\\vdots\\c_w\end{bmatrix}=\begin{bmatrix}c_{11}&c_{12}&\cdots&c_{1n}\\c_{21}&c_{22}&\cdots&c_{2n}\\\vdots&\vdots&\cdots&\vdots\\c_{w1}&c_{w2}&\cdots&c_{wn}\end{bmatrix}.$$

称为 $f(\bullet)$ 的定序特征矩阵。

显然布尔函数和其定序特征矩阵是相互唯一确定的。

定义 3[11] $2k\times n$ 阶 0, 1 矩阵 $A=[a_{ij}]$ 称为列平衡矩阵，如果矩阵 A 各行互异且 A 的各列中刚好含有 k 个 0 和 k 个 1 。

由定义 1 和相关免疫函数的 Hamming 重量恒为偶数，及文献 [7] 推论 2 可知相关免疫函数还可等价地定义为：

定义 4 n 元布尔函数 $f(x_1,\cdots,x_n)$ 是相关免疫函数，如果 $W(f)$ 为偶数 (不妨记 $W(f)=2k$) 且 $f(\bullet)$ 的定序特征矩阵是列平衡矩阵。

由于相关免疫函数的 Hamming 重量为偶数，因此若对任意 $k(0\le k\le 2^{n-1})$ 求出了重量为 $2k$ 的相关免疫函数的精确个数 $N(2k,n)$ ，则 n 元相关免疫函数的精确值 $N(n)$ 就可表示为

$$N(n)=\sum_{k=0}^{2^{n-1}}N(2k,n).$$

容易证明

$$\begin{aligned}&N(2k,n)=N(2^n-2k,n),\\&N(0,n)=N(2^n,n)=1,\\&N(2,n)=N(2^n-2,n)=2^{n-1}.\end{aligned}$$

文献 [8] 得到

$$N(4,n)=N(2^n-4,n)=(6^n-3\times 2^n)/4!$$

及当 $n\ge 3$ 时

$$N(6,n)=N(2^n-6,n)=\left[\sum_{i=3}^{n-1}8^{n-i-1}g(i)+45\times 8^{n-1}\right]/6!,$$

其中 $g(i)=12\times 20^i+75\times 2^{n+4}-45\times 4^{n+2}$ ，并且 $N(6,1)=0$, $N(6,2)=0$ 。

根据定义 4 、定义 2 和定义 3 易知：若记 $M(2k,n)$ 为 $2k\times n$ 阶列平衡矩阵的个数，则重量为 $2k$ 的相关免疫函数的数目是：

$$N(2k,n)=M(2k,n)/(2k)!.$$

下节利用组合数学中的容斥原理 [14] 和图论 [15] 中的子图概念，给出 $M(2k,n)$ 的一般表达式.

3 列平衡矩阵的计数

为了用组合数学中的容斥原理给出 $M(2k,n)$ 的一般表达式，记 $\{0,1\}^{2k\times n}$ 为 $2k\times n$ 阶 (0,1)- 矩阵的全体构成的集合，设

$$A(2k,n)=\{X\in\{0,1\}^{2k\times n}|X\text{是列平衡矩阵}\},$$
$$B(2k,n)=\{X\in\{0,1\}^{2k\times n}|X\text{的每列中含有}k\text{个0和}k\text{个1}\},$$
$$C(i,j)=\{X\in B(2k,n)|X\text{的第}i\text{行与第}j\text{行相同}\},$$

其中 $1\le i<j\le 2k$. 不难知道：当 $2k>2^n$ 时，$|A(2k,n)|=0$；并且 $|B(2k,n)|=\begin{bmatrix}2k\\k\end{bmatrix}^n$.

根据容斥原理 [14] 有

$$M(2k,n)=|A(2k,n)|=|B(2k,n)|+\sum_{F\subseteq P_2,F\neq\varnothing}(-1)^{|F|}\cdot\left|\bigcap_{\{i,j\}\in F}C(i,j)\right|,\tag{1}$$

其中 P_2 是 $\{1,2,\cdots,2k\}$ 的全体 2- 元子集构成的集合，P_2 可以看成是 $2k$ 阶完全图的边集；F 是 P_2 的非空子集，F 可以看成是 $2k$ 阶完全图的非空边子集.

$\bigcap_{\{i,j\}\in F}C(i,j)$ 中矩阵的性质和 F 的边生成子图 $<F>$ 的结构有着密切的关系. 根据 $C(i,j)$ 的定义和 $\bigcap_{\{i,j\}\in F}C(i,j)$ 的含义，易知以下事实：$\bigcap_{\{i,j\}\in F}C(i,j)$ 中矩阵的行标号若在 $<F>$ 的同一连通支中，则对应的矩阵的行完全相同. 若用 r 表示 $<F>$ 连通支的数目，$<F>$ 的 r 个连通支的点数分别记为 $q_1,q_2,\cdots,q_r$，显然 $\max_{1\le i\le r}q_i\ge 2$. 由于 $B(2k,n)$ 中不存在 $k+1$ 行完全相同的矩阵，所以，当 $\max_{1\le i\le r}q_i\ge k+1$ 时，$\bigcap_{\{i,j\}\in F}C(i,j)=\varnothing$，即 $|\bigcap_{\{i,j\}\in F}C(i,j)|=0$.

根据以上分析，可以知道：$\bigcap_{\{i,j\}\in F}C(i,j)$ 中矩阵的每列有 $\sum_{\alpha\in\{0,1\}^r}\begin{bmatrix}2k-(I,q(F))\\k-(\alpha,q(F))\end{bmatrix}$ 种选取方法，其中 $I=(1,1,\cdots,1)$ 是 r 维向量，$q(F)=(q_1,q_2,\cdots,q_r)$，$\alpha$ 是 r 维 (0,1)- 向量，$(I,q(F))$ 和 $(\alpha,q(F))$ 都是向量的内积运算. 并且约定：当 $k-(\alpha,q(F))<0$ 或 $2k-(I,q(F))<k-(\alpha,q(F))$ 时，$\begin{bmatrix}2k-(I,q(F))\\k-(\alpha,q(F))\end{bmatrix}=0$；并且 $\begin{bmatrix}0\\0\end{bmatrix}=1$；另外当 $2k-(I,q(F))\ge 0$ 时，$\begin{bmatrix}2k-(I,q(F))\\0\end{bmatrix}=1$. 所以

$$\left|\bigcap_{\{i,j\}\in F}C(i,j)\right|=\prod_{l=1}^{n}\sum_{\alpha\in\{0,1\}^r}\begin{bmatrix}2k-(I,q(F))\\k-(\alpha,q(F))\end{bmatrix}=\left[\sum_{\alpha\in\{0,1\}^r}\begin{bmatrix}2k-(I,q(F))\\k-(\alpha,q(F))\end{bmatrix}\right]^n.\tag{2}$$

特别是在 $\max_{1\le i\le r}q_i\ge k+1$ 时，不妨设 $q_1\ge k+1$. 这时，若 α 的第一个分量是 1，则 $k-(\alpha,q(F))<0$，于是 $\begin{bmatrix}2k-(I,q(F))\\k-(\alpha,q(F))\end{bmatrix}=0$；当 α 的第一个分量是 0 时，由于

$$2k-(I,q(F))=2k-\sum_{i=1}^{r}q_i=2k-q_1-\sum_{i=2}^{r}q_i\le k-1-\sum_{i=2}^{r}q_i<k-\sum_{i=2}^{r}q_i\le k-(\alpha,q(F))$$

从而 $\begin{bmatrix} 2k-(I,q(F)) \\ k-(\alpha,q(F)) \end{bmatrix}=0$. 这说明当 $\max_{1\le i\le r} q_i \ge k+1$ 时, $\sum_{\alpha\in\{0,1\}^r} \begin{bmatrix} 2k-(I,q(F)) \\ k-(\alpha,q(F)) \end{bmatrix}=0$,

因此 $|\bigcap_{\{i,j\}\in F} C(i,j)|=0$, 这与前面根据 $B(2k,n)$ 的含义得到 $|\bigcap_{\{i,j\}\in F} C(i,j)|=0$ 一致。

由此可设 $\max_{1\le i\le r} q_i \le k$, 即 $q_i \le k(i=1,2,\cdots,r)$; 另外又因 $\sum_{i=1}^r q_i \le 2k$, 于是

$$|F| \le \sum_{i=1}^{r} \begin{bmatrix} q_i \\ 2 \end{bmatrix} = \frac{\sum_{i=1}^r q_i^2 - \sum_{i=1}^r q_i}{2} \le \frac{k\sum_{i=1}^r q_i - \sum_{i=1}^r q_i}{2} = \frac{(k-1)\sum_{i=1}^r q_i}{2} \le k(k-1).$$

所以, 当 $|F|>k(k-1)$ 时, 必有 $\max_{1\le i\le r} q_i \ge k+1$ 从而 $|\bigcap_{\{i,j\}\in F} C(i,j)|=0$.

综合以上分析, 可将 (1) 式和 (2) 式合并写成

$$M(2k,n)=\begin{bmatrix} 2k \\ k \end{bmatrix}^n + \sum_{F\subseteq P_2, 1\le|F|\le k(k-1)} (-1)^{|F|}\cdot\left[\sum_{\alpha\in\{0,1\}^r} \begin{bmatrix} 2k-(I,q(F)) \\ k-(\alpha,q(F)) \end{bmatrix}\right]^n . \tag{3}$$

4 计算实例

本节利用 (3) 式计算 $M(4,n)$. 这时 $k=2$, F 应满足 $1\le|F|\le 2$, 且 $\max_{1\le i\le r} q_i \le 2$. 当 $|F|=1$ 时, $<F>$ 是两点完全图 K_2 , r=1 , $q(F)=(2)$, 在 K_4 中 F 有 6 种选法, 同时

$$\begin{aligned} |\bigcap_{\{i,j\}\in F} C(i,j)| &= \left[\sum_{\alpha\in\{0,1\}} \begin{bmatrix} 4-(I,q(F)) \\ 2-(\alpha,q(F)) \end{bmatrix}\right]^n \\ &= \left[\sum_{\alpha\in\{0,1\}} \begin{bmatrix} 4-2 \\ 2-(\alpha,q(F)) \end{bmatrix}\right]^n = \left[\begin{bmatrix} 2 \\ 0 \end{bmatrix}+\begin{bmatrix} 2 \\ 2 \end{bmatrix}\right]^n = 2^n; \end{aligned}$$

$|F|=2$ 时, $<F>$ 的连通支是完全图 K_2 , r=2, $q(F)=(2,2)$, 在 K_4 中 F 有 3 种选法, 同时

$$|\bigcap_{\{i,j\}\in F} C(i,j)| = \left[\begin{bmatrix} 0 \\ 2 \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \end{bmatrix}+\begin{bmatrix} 0 \\ 0 \end{bmatrix}\begin{bmatrix} 0 \\ -2 \end{bmatrix}\right]^n = 2^n.$$

于是

$$M(4,n)=\begin{bmatrix} 4 \\ 2 \end{bmatrix}^n - 6\times 2^n + 3\times 2^n = 6^n - 3\times 2^n.$$

这与文献 [8] 的结果相同, 但求解过程简单。

同理, 用 (3) 式可计算得

$$M(6,n)=2^n\times(10^n - 15\times 4^n + 45\times 2^n - 40).$$

比文献 [8] 结果 (见本文第 2 节) 简明.

5 结 束 语

本文依据布尔函数的列平衡矩阵表示，利用组合数学中的容斥原理和图论中的子图概念，给出了一阶相关免疫函数的一个简单且宜于理解的计数公式. 与文献 [13] 的结果相比，本文公式更实用. 另一方面，从目前相关免疫布尔函数计数问题的研究结果来看，要给出一个更为简明且利于应用的公式，还需要对相关免疫布尔函数的列平衡矩阵的结构作深入细致的分析.

参 考 文 献

[1] Siegenthaler T. Correlation-immunity of nonlinear combining functions for cryptographic applications. IEEE Trans. on Information Theory, 1984, IT-30(5): 776–780.

[2] Siegenthaler T. Decrypting a class of stream ciphers using ciphertext only. IEEE Trans. on Computer, 1985, C-34(1): 81–85.

[3] 丁存生，肖国镇. 流密码学及其应用. 北京：国防工业出版社， 1994, 160–180.

[4] 杨义先，胡正名. 用于序列密码的布尔函数计数问题. 通信学报， 1992, 13(4): 18–24.

[5] Pichier F. On the Walsh-Fourier analysis of correlation-immune switching functions. Proc. of Eurocrypt'86, Linkoping: 1986, 237–241.

[6] Mitchell C. Enumerating Boolean functions of cryptographic significance. J. of Cryptology, 1990, 2(3): 155–170.

[7] 单炜娟. 相关免疫函数的结构与构造. 应用数学学报， 1991, 14(3): 331–336

[8] 杨义先. 相关免疫布尔函数的计数. 电子科学学刊， 1993, 15(2): 140– 146.

[9] 郭宝安. 非线性序列的分析和综合： [博士论文]. 北京：北京邮电大学， 1993.

[10] 王建宇. 线性结构函数与一阶相关免疫函数的计数. 通信学报， 1996, 17(1): 87–91.

[11] 杨义先. 高维 Hadamard 矩阵的几个猜想之证明. 科学通报， 1986, 31(2): 85–89.

[12] 张建州. 非相关布尔函数个数的精确值. 电子科技大学学报， 1994, 23(1): 89–94.

[13] Palmer E M, Read R C, Robinson R W. Balancing the n-cube: A census of colorings. J of Algebraic Combinatorics, 1992, 1(4): 257–273.

[14] 柯召，魏万迪. 组合论. 北京：科学出版社， 1984, 85–88.

[15] Bondy J A, Murty U S R. Graph Theory with Applications. London: The MacMillan Press Ltd., 1976, 8–10.

THE ENUMERATION OF THE FIRST ORDER CORRELATION IMMUNE BOOLEAN FUNCTIONS

Zhang Jianzhou　　You Zhisheng

(*Institute of Image & Graphics, College of Computer, Sichuan University, Chengdu* 610064)

Abstract　A enumerative formula of Boolean functions with the first-order correlation-immunity, which are important in stream ciphers, is given with help of column-balance matrixes, inclusion-exclusioin principle in combinatorics and edge-induced subgraphs in graph theory.

Key words　Boolean functions, Correlation immunity, Enumeration, Inclusion-exclusion principle, Subgraph

张建州：　男， 1962 年生，教授，博士，研究方向：计算机视觉和图像处理，神经计算，多传感器数据融合及目标跟踪，组合数学和编码，图论.

游志胜：　男， 1945 年生，教授，博士生导师，研究方向：图像处理和仿真技术，多传感器数据融合及目标跟踪.