第25卷第1期 数学研究与评论 Vol.25, No.1
2005年2月 JOURNAL OF MATHEMATICAL RESEARCH AND EXPOSITION Feb., 2005

文章编号: 1000-341X(2005)01-0176-07 **文献标识码**: A

对有限域上复合变换的线性逼近

金晨辉 [1], 李世取 [2]

(1. 解放军信息工程大学电子技术学院, 河南 郑州 450004;
2. 解放军信息工程大学信息工程学院, 河南 郑州 450002)
(E-mail: Jinchenhui@126.com)

摘 要: 本文解决了有限交换群上复合函数的特征谱的计算问题, 定义了有限交换群上函数的相关系数的概念, 并解决了有限域上复合函数与线性映射的相关系数的计算问题, 从而建立了对密码算法中的复合变换进行线性逼近的理论基础.

关键词: 线性逼近; 复合变换; 相关系数; 谱; 有限域; 有限交换群.
MSC(2000): 94A60, 11T71
中图分类: TN918.1

对密码函数进行线性逼近是密码分析的一个重要手段. 对分组密码的线性密码分析 [1] 就是一种线性逼近攻击. 当线性函数是由模 2 加运算定义的线性函数 $A(x) = w_1x_1 \oplus \cdots \oplus w_nx_n$ 时, 可用 $|p(f(x) = A(x)) - p(f(x) \neq A(x))|$ 来刻划 Boole 函数 $f(x)$ 与线性函数 $A(x) = wx$ 的接近程度, 它恰好就是 $f(x)$ 在 w 点的 Walsh 谱的绝对值, 因而此时线性逼近问题完全可归结为 Walsh 谱理论进行研究.

密码算法通常由许多变换复合而成, 因而研究复合变换的线性逼近问题具有重要的理论价值和实用价值. 文献 [2] 给出了复合变换的 Walsh 谱计算定理, 从而建立了对二元域上复合变换进行线性逼近的理论基础, 解决了线性密码分析 [1] 的理论基础问题.

随着密码算法设计理论研究的日益深入, 越来越多的密码算法直接采取有限域上的特定变换作为基本的密码变换, 如美国 21 世纪的数据加密标准 AES 算法 (即 Rijndael 算法) 就是采用 256 元域上的线性变换和乘法逆变换作为其基本的密码变换. 在此背景下, 研究新的线性逼近技术, 例如研究有限域 $GF(2^n)$ 和剩余类环 $Z/(2^n)$ 上的线性逼近问题就既具有理论意义, 又具有现实意义.

但是, 当将二元域改为一般的有限域或剩余类环时, 如何对复合变换进行线性逼近这个问题还没有解决. 为此, 本文首先从概率论的角度出发定义多值函数之间的相关系数, 并用它刻划多值函数之间的接近程度; 其次, 本文将解决有限交换群上复合函数的特征谱的计算问题, 并借助于此解决有限域上复合函数与线性函数的相关系数的计算问题. 最后, 我们指出, 相应的结果对剩余类环上的复合函数一般不成立.

定义 1 设 G 和 H 都是有限交换群, $f, g : G \to H, k \in H$, 则称

$$\rho(f, g; k) = |H|p(f(x) - g(x) = k) - 1 = |G|^{-1}|H||\{x \in G : f(x) - g(x) = k\}| - 1$$

为函数 f 与函数 g 在 k 点的相关系数.

收稿日期: 2002-05-31
基金项目: 河南省杰出青年科学基金 (0312001800)

如果 x 在 G 中服从均匀分布，则函数 f 与函数 g 在 k 点的相关系数反映了 $f(x)-g(x)$ 取 k 值的概率与平均值的偏离程度，因而 $\{\rho(f,g;k),k\in H\}$ 完全反映了 f 与 g 的接近程度. 从这个意义上讲，函数 $f(x)$ 与有限域或剩余类环上的仿射函数 $ax+b$ 的相关程度和 $f(x)$ 与线性函数 ax 的相关程度是一致的，因而仿射逼近可完全归结为线性逼近进行研究. 显然，有

命题 1 设 $f,g:\{0,1\}^n\to\{0,1\}$, 则

$$\rho(f,g;0)=\frac{1}{2^n}\sum_{x\in\{0,1\}^n}(-1)^{f(x)\oplus g(x)}.$$

可见，二元域上 Boole 函数 f 与线性函数 wx 在 0 点的相关系数和 f 在 w 点的 Walsh 谱是一致的.

设 G 为有限交换群，则 G 至复数域之乘法群的同态称为 G 上的一个特征. 记 G^* 为 G 的特征全体，规定 G^* 中运算为： $\forall\tau,\sigma\in G^*,(\tau+\sigma)(x)=\tau(x)+\sigma(x)$, 则 $(G^*,+)$ 构成群且与 G 同构. 记 $G^*=\{\chi_w:w\in G\}$ 且 $\varphi:w\to\chi_w$ 是 G 至 G^* 的一个群同构.

定义 2[3] 设 G 和 H 均是有限交换群， $f:G\to H,(\alpha,\beta)\in G\times H$, 则称

$$S_{(f)}(\alpha,\beta)=|G|^{-1}\sum_{x\in G}\chi_\beta(f(x))\chi_\alpha(-x)$$

为 f 在点 (α,β) 的特征谱.

特别地， $\alpha,x\in[Z/(n)]^k$, 记 $\alpha x=(\alpha_1x_1+\cdots+\alpha_kx_k)\mathrm{mod}n$, 则

$$\{e^{\frac{2\pi\sqrt{-1}}{n}\alpha x}:\alpha\in[Z/(n)]^k\}$$

是 $[Z/(n)]^k$ 的特征全体. 又设 $f:[Z/(n)]^m\to[Z/(n)]^k$, 则此时

$$S_{(f)}(\alpha,\beta)=n^{-m}\sum_{x\in[Z/(n)]^m}e^{\frac{2\pi\sqrt{-1}}{n}(\beta f(x)-\alpha x)}$$

就是模 n 剩余类环上函数 $\beta f(x)$ 在 α 点的 Chrestenson 谱 [4]; $\alpha,x\in[GF(p^t)]^k$, 记 $\alpha x=\alpha_1x_1+\cdots+\alpha_kx_k,Tr_t(y)$ 为 $GF(p^t)$ 至素域 $GF(p)$ 的迹函数，则

$$\{e^{\frac{2\pi\sqrt{-1}}{p}Tr_t(\alpha x)}:\alpha\in[GF(p^t)]^k\}$$

是 $[GF(p^t)]^k$ 的特征全体. 又设 $f:[GF(p^t)]^m\to[GF(p^t)]^k$, 则此时

$$S_{(f)}(\alpha,\beta)=p^{-mt}\sum_{x\in[GF(p^t)]^m}e^{\frac{2\pi\sqrt{-1}}{p}(Tr_t(\beta f(x)-\alpha x))}$$

就是有限域 $GF(p^t)$ 上函数 $\beta f(x)$ 在 α 点的广义 Chrestenson 循环谱 [5].

定理 1 设 G,H 和 L 均是有限交换群， $g:G\to H,f:H\to L,(\alpha,\beta)\in G\times L$, 则有

$$S_{(fg)}(\alpha,\beta)=\sum_{\gamma\in H}S_{(g)}(\alpha,\gamma)S_{(f)}(\gamma,\beta).$$

证明

$$\begin{aligned}&\sum_{\gamma\in H}S_{(g)}(\alpha,\gamma)S_{(f)}(\gamma,\beta)\\&=|G|^{-1}|H|^{-1}\sum_{\gamma\in H}\sum_{x\in G}\chi_\gamma(g(x))\chi_\alpha(-x)\sum_{y\in H}\chi_\beta(f(y))\chi_\gamma(-y)\\&=|G|^{-1}|H|^{-1}\sum_{x\in G}\sum_{y\in H}\chi_\beta(f(y))\chi_\alpha(-x)\sum_{\gamma\in H}\chi_\gamma(g(x)-y)\\&=|G|^{-1}\sum_{x\in G}\chi_\beta(fg(x))\chi_\alpha(-x)=S_{(fg)}(\alpha,\beta).\end{aligned}$$

推论[2] 设 $g:\{0,1\}^m\to\{0,1\}^n, f:\{0,1\}^n\to\{0,1\}^t$, 则 $\forall(\alpha,\beta)\in\{0,1\}^m\times\{0,1\}^t$, 有

$$S_{(\beta fg)}(\alpha)=\sum_{\gamma\in\{0,1\}^n}S_{(\gamma g)}(\alpha)S_{(\beta f)}(\gamma),$$

这里 $S_{(\beta f)}(\gamma)=2^{-n}\sum_x(-1)^{\alpha f(x)\oplus\gamma x}$.

引理 1 设 G 是有限交换群, f 为 G 上的复数值函数, $a\in G$, 令

$$S_f\langle\alpha\rangle=|G|^{-1}\sum_{x\in G}f(x)\chi_\alpha(-x),$$

则 $x\in G$, 有 $f(x)=\sum_{a\in G}S_f\langle\alpha\rangle\chi_x(\alpha)$.

定义 3 设 G 是含幺有限交换环, 如果 $\forall\alpha,x\in G$, 都有 $\chi_\alpha(x)=\chi_1(\alpha x)$, 则称 G 满足 χ 条件.

显然, 当 G 满足 χ 条件时, 有 $\chi_a(x)=\chi_x(\alpha)$, 而且有限域和剩余类环都是满足 χ 条件的含幺有限交换环.

设 F 是满足 χ 条件的含幺有限交换环, $F^n=\{(x_1,\cdots,x_n):x_i\in F\}.\forall\alpha,x\in F^n,d\in F$, 记 $\alpha x=\alpha_1x_1+\cdots+\alpha_nx_n\in F$, 并约定 $dx=(dx_1,\cdots,dx_n)$. 显然, $\forall\alpha,x\in F^n$ 及 $d\in F$, 有 $(d\alpha)x=d(\alpha x)=\alpha(dx)$, 且由有限交换群上特征的性质知, $\forall w,x\in F^n,\forall d\in F$, 有

$$\begin{aligned}\chi_{dw}(x)&=\prod_{i=1}^n\chi_{dw_i}(x_i)=\prod_{i=1}^n\chi_1((dw_i)x_i)=\prod_{i=1}^n\chi_d(w_ix_i)\\&=\chi_d(w_1x_1+\cdots+w_nx_n)=\chi_d(wx).\end{aligned}$$

定理 2 设 F 是满足 χ 条件的含幺有限交换环, $g:F^m\to F^k, f:F^k\to F^t$, 则以下两结论等价:

(1) $\forall(\alpha,\beta)\in F^m\times F^t,\forall d\in F$, 有

$$p(\beta fg(x)-\alpha x=d)=\sum_{\gamma\in F^k}\sum_{s\in F}p(\beta f(y)-\gamma y=s)p(\gamma g(x)-\alpha x=d-s)+|F|^{-1}-|F|^{k-1};$$

(2) $\forall(\alpha,\beta)\in F^m\times F^t, d\in F$ 且 $d\neq0$, 有

$$S_{(fg)}(d\alpha,d\beta)=\sum_{\gamma\in H}S_{(g)}(d\alpha,d\gamma)S_{(f)}(d\gamma,d\beta).$$

证明 设 $w \in F$ 且 $w \neq 0$, 则

$$\begin{aligned}
&\sum_{d\in F} p(\beta f g(x) - \alpha x = d)\chi_w(d) = \frac{1}{|F|^m}\sum_{x\in F^m}\chi_w(\beta f g(x) - \alpha x)\\
&= \frac{1}{|F|^m}\sum_{x\in F^m}\chi_{w\beta}(fg(x))\chi_{w\alpha}(-x)\\
&= S_{(fg)}(w\alpha, w\beta)\\
&\sum_{d\in F}[\sum_{\gamma\in F^k}\sum_{s\in F} p(\beta f(y) - \gamma y = s)p(\gamma g(x) - \alpha x = d - s)]\chi_w(d)\\
&= \sum_{\gamma\in F^k}[\sum_{s\in F} p(\beta f(y) - \gamma y = s)\chi_w(s)][\sum_{d\in F} p(\gamma g(x) - \alpha x = d - s)]\chi_x(d - s)]\\
&= \sum_{\gamma\in F^k}[|F|^{-k}\sum_{y\in F^k}\chi_w(\beta f(y) - \gamma y)] \times [|F|^{-m}\sum_{x\in F^m}\chi_w(\gamma g(x) - \alpha x)]\\
&= \sum_{\gamma\in F^k} S_{(f)}(w\gamma, w\beta)S_{(g)}(w\alpha, w\gamma).
\end{aligned}$$

故由 $\sum_{d\in F}\chi_w(d) = 0$ 知 (1) 蕴涵 (2). 反之，设 (2) 成立，则由引理 1 知

$$\begin{aligned}
&p(\beta f g(x) - \alpha x = d)\\
&= |F|^{-1}S_{(fg)}(0,0) + |F|^{-1}\sum_{z\in F\setminus\{0\}} S_{(fg)}(z\alpha, z\beta)\chi_z(-d)\\
&= F^{-1} + F^{-1}\sum_{z\in F\setminus\{0\}}\sum_{\gamma\in F^k} S_{(f)}(z\gamma, z\beta)S_{(g)}(z\alpha, z\gamma)\chi_z(-d)\\
&= F^{-1} - F^{k-1} + F^{-1}\sum_{z\in F}\sum_{\gamma\in F^k} S_{(f)}(z\gamma, z\beta)S_{(g)}(z\alpha, z\gamma)\chi_z(-d)\\
&= \sum_{\gamma\in F^k}\sum_{s\in F} p(\alpha f(y) - \gamma y = s)p(\gamma g(x) - \beta x = d - s) + F^{-1} - F^{k-1}.
\end{aligned}$$

这说明 (1) 成立.

定理 3 设 F 是有限域， $g: F^m \to F^k, f: F^k \to F^t$, 则 $\forall(\alpha,\beta) \in F^m \times F^t, d \in F$, 有

$$p(\beta f g(x) - \alpha x = d) = \sum_{\gamma\in F^k}\sum_{s\in F} p(\beta f(y) - \gamma y = s)p(\gamma g(x) - \alpha x = d - s) + |F|^{-1} - |F|^{k-1}.$$

证明 设 $(\alpha,\beta) \in F^m \times F^t, d \in F$ 且 $d \neq 0$, 则由 F 是有限域知，映射 $\psi(x) = dx$ 是 F^k 到自身的双射，即 $F^k = dF^k = \{dx : x \in F^k\}$, 故由定理 1 知

$$\begin{aligned}
S_{(fg)}(d\alpha, d\beta) &= \sum_{y\in F^k} S_{(g)}(d\alpha, y)S_{(f)}(y, d\beta) = \sum_{x\in dF^k} S_{(g)}(d\alpha, x)S_{(f)}(x, d\beta)\\
&= \sum_{x\in F^k} S_{(g)}(d\alpha, dx)S_{(f)}(dx, d\beta),
\end{aligned}$$

故由定理 2 知本定理成立.

定理 4 设 F 是有限域，$g:F^m\to F^k, f:F^k\to F^t$，则 $\forall(\alpha,\beta)\in F^m\times F^t, d\in F$，有

$$\rho(\beta fg,\alpha x;d)=\frac{1}{|F|}\sum_{\gamma\in F^k}\sum_{s\in F}\rho(\beta f,\gamma y;s)\rho(\gamma g,\alpha x;d-s).$$

证明 由定理 3 知

$$\begin{aligned}\rho(\beta fg,\alpha x;d)&=|F|p(\beta fg(x)-\alpha x=d)-1\\&=|F|\sum_{\gamma\in F^k}\sum_{s\in F}p(\beta f(y)-\gamma y=s)p(\gamma g(x)-\alpha x=d-s)-|F|^k,\end{aligned}$$

又因

$$\begin{aligned}&\frac{1}{|F|}\sum_{\gamma\in F^k}\sum_{s\in F}\rho(\beta f,\gamma y;s)\rho(\gamma g,\alpha x;d-s)\\&=\frac{1}{|F|}\sum_{\gamma\in F^k}\sum_{s\in F}[|F|p(\beta f(y)-\gamma y=s)-1][|F|p(\gamma g(x)-\alpha x=d-s)-1]\\&=|F|\sum_{\gamma\in F^k}\sum_{s\in F}p(\beta f(y)-\gamma y=s)p(\gamma g(x)-\alpha x=d-s)+\frac{1}{|F|}\sum_{\gamma\in F^k}\sum_{s\in F}1-\\&\quad\sum_{\gamma\in F^k}[\sum_{s\in F}p(\beta f(y)-\gamma y=s)]-\sum_{y\in F^k}[\sum_{s\in F}p(\gamma g(x)-\alpha x=d-s)]\\&=|F|\sum_{\gamma\in F^k}\sum_{s\in F}p(\beta f(y)-\gamma y=s)p(\gamma g(x)-\alpha x=d-s)+|F|^k-\sum_{\gamma\in F^k}1-\sum_{\gamma\in F^k}1\\&=|F|\sum_{\gamma\in F^k}\sum_{s\in F}p(\beta f(y)-\gamma y=s)p(\gamma g(x)-\alpha x=d-s)-|F|^k,\end{aligned}$$

故本定理成立.

推论 设 F 是有限域，

$$f_i:F^{n_i}\to F^{n_{i+1}}, f=f_kf_{k-1}\cdots f_2f_1, \alpha_1\in F^{n_1}, \alpha_{k+1}\in F^{n_{k+1}},$$

则 $\forall d\in F$，有

$$\rho(\alpha_{k+1}f,\alpha_1x;d)=\frac{1}{|F|^{k-1}}\sum_{\alpha_2\in F^{n_2},\cdots,\alpha_k\in F^{n_k}}\sum_{s_1+\cdots+s_k=d}\prod_{i=1}\rho(\alpha_{i+1}f_i,\alpha_ix;s_i).$$

证明 利用定理 4 和归纳法即证.

定理 4 的推论中的

$$\frac{1}{|F|^{k-1}}\sum_{s_1+\cdots+s_k=d}\prod_{i=1}^{k}\rho(\alpha_{i+1}f_i,\alpha_ix;s_i)$$

反映了对有限域 F 上多输出复合函数 $f=f_kf_{k-1}\cdots f_2f_1$ 进行线性逼近时，由线性逼近路径 $\alpha_1\to\alpha_2\to\cdots\to\alpha_{k+1}$ 所获得的相关系数 $\rho(\alpha_{k+1}f,\alpha_1x;d)$ 的信息，称为该线性逼近路径的相关系数. 定理 4 推论说明 $\rho(\alpha_{k+1}f,\alpha_1x;d)$ 是所有这些线性逼近路径的相关系数之和. 当其中某条线性逼近路径的相关系数的绝对值显著大于其它线性逼近路径的相关系数时，通常就以该线

性逼近路径的相关系数作为 $\rho(\alpha_{k+1}f,\alpha_1 x;d)$ 的近似值，这就是线性密码分析方法的理论基础. 定理 4 推论说明对有限域上的复合函数进行线性逼近时仍然可以采用这种逼近方法，从而解决了对有限域上的复合函数进行线性逼近的方法及其理论基础问题.

从定理 4 推论还可看出，如果满足 $s_1+\cdots+s_k=d$ 的一组 $(s_1,\cdots,s_k)$ 的对应的

$$\prod_{i=1}^{k}\rho(\alpha_{i+1}f_i,\alpha_i x;s_i)$$

与平均值

$$\frac{1}{|F|^{k-1}}\sum_{s_1+\cdots+s_k=d}\prod_{i=1}^{k}\rho(\alpha_{i+1}f_i,\alpha_i x;s_i)$$

较为接近，则通常还可用 $\prod_{i=1}^{k}\rho(\alpha_{i+1}f_i,\alpha_i x;s_i)$ 作为 $\frac{1}{|F|^{k-1}}\sum_{s_1+\cdots+s_k=d}\prod_{i=1}^{k}\rho(\alpha_{i+1}f_i,\alpha_i x;s_i)$ 的近似值，进而作为 $\rho(\alpha_{k+1}f,\alpha_1 x;d)$ 的近似值，从而简化对相关系数 $\rho(\alpha_{k+1}f_i,\alpha_1 x;s_i)$ 的估计.

然而，一般而言，定理 4 及定理 2 中的 (1) 和 (2) 对模 n 剩余类环上的多输出函数并不成立.

例 1 设 $f=(f_n,f_{n-1},\cdots,f_1):Z/(2^n)\to Z/(2^n)$ 是平衡函数，记 $x=(x_n,x_{n-1},\cdots,x_1)$. 如果 $f_1(x)\oplus x_1$ 是平衡函数，则存在 $Z/(2^n)$ 至 $Z/(2^n)$ 的平衡函数 g，使定理 4 对 f 和 g 不成立.

事实上，令 $d=2^{n-1}.\forall\gamma\in Z/(2^n)$，记 $\gamma=(\gamma_n,\gamma_{n-1},\cdots,\gamma_1)$，则由 $f_1(x)\oplus x_1$ 和 $f_1(x)$ 都是平衡函数知

$$\begin{aligned}S_{(f)}(d\gamma,d)&=\frac{1}{2^n}\sum_{x\in Z/(2^n)}e^{\frac{2\pi\sqrt{-1}}{2^n}(df(x)-d\gamma x)}=\frac{1}{2^n}\sum_{x\in Z/(2^n)}e^{\pi\sqrt{-1}(f_1(x)+\gamma_1 x_1)}\\&=\frac{1}{2^n}\sum_{x\in Z/(2^n)}(-1)^{f_1(x)+\gamma_1 x_1}=0.\end{aligned}$$

如果定理 4 对 f 和 $g=f^{-1}$ 成立，则由定理 4 与定理 2 之 (1) 和 (2) 等价知

$$S_{(fg)}(d,d)=\sum_{\gamma\in Z/(2^n)}S_{(g)}(d,d\gamma)S_{(f)}(d\gamma,d)=0.$$

但由 $fg(x)=x$ 易证 $S_{(fg)}(d,d)=1$，该矛盾说明定理 4 对剩余类环上的函数未必成立.

本文解决了有限域上复合变换与线性变换的相关系数的计算问题，从而为利用有限域方法对密码算法进行线性逼近提供了理论基础. 由于类似结果对剩余类环上复合变换并不成立，因此如何对剩余类环上复合函数进行线性逼近，是有待进一步解决的问题.

参考文献:

[1] MATSUI M. *Linear Cryptanalysis Method for DES Cipher* [C]. In: T Helleseth, ed. Advances in Cryptology — Eurocrypt'93, LNCS 765, Springer-Verlag, 1993, 386–397.

[2] DAEMEN J. et al. *Correlation Matrices* [C]. In: Bart Preneel, ed.Fast Software Encryption. LNCS 1008. Berlin: Springer-Verlag, 1995, 275–285

[3] 陈卫红. 谱理论在密码学逻辑函数的分析和设计中的应用 [D]. 解放军信息工程学院博士论文, 1998.
CHEN Wei-hong. *Applications of spectral theory in the analysis and design of logic functions in cryptology* [D]. The Dissertation for doctor degree in PLA information Engineering Institute, 1998.

[4] 丁存生, 肖国镇. 流密码学及其应用 [M]. 北京: 国防工业出版社, 1994.
DING Cun-sheng, XIAO Guo-zhen. *Stream Ciphers and Its Applications* [M]. Beijing: The Defence Industry Press, 1994.

[5] 冯登国, 肖国镇. 有限域上的函数的相关免疫性和线性结构的谱特征 [J]. 通信学报, 1997, **18**(1): 40–45.
FENG Deng-guo, XIAO Guo-zhen. *The correlation-immunity and spectral characteristics for functions over finite fields* [J]. Journal of China Institute of Communications, 1997, **18**(1): 40–45.

Linear Approximation to Composition Transformations over Finite Fields

JIN Chen-hui[1], *LI Shi-qu*[2]
(1. Institute of Electronic Technology, PLA Information Engineering University, Zhengzhou 450004, China;
2. Institute of Information Engineering, PLA Information Engineering University, Zhengzhou 450002, China)

Abstract: In this paper, the problem of how to compute the characteristic spectrum for composition transformations over finite Abelian groups is solved, the concept of correlation coefficient of functions over finite Abelian groups is defined, and the problem of how to compute the correlation coefficient between a composition function and a linear function over finite field is solved, so the theoretic foundation of linear approximation to composition transformations in cryptography algorithm is established.

Key words: linear approximation; composition function; correlation coefficient; spectrum; finite field; finite Abelian group.