计算机科学
COMPUTER SCIENCE

优良平衡布尔函数的Rank排序混合遗传搜索算法

赵海霞，李鑫，韦永壮

引用本文

赵海霞，李鑫，韦永壮. 优良平衡布尔函数的Rank排序混合遗传搜索算法[J]. 计算机科学，2025，52(12)：351-357.

ZHAO Haixia, LI Xin, WEI Yongzhuang. Rank-sorting Hybrid Genetic Algorithm for Search High Quality Balanced Boolean Functions [J]. Computer Science, 2025, 52(12): 351-357.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.241200039

优良平衡布尔函数的 Rank 排序混合遗传搜索算法

赵海霞[1,3]　李　鑫[1]　韦永壮[2]
1 桂林电子科技大学数学与计算科学学院　广西 桂林 541004
2 桂林电子科技大学计算机与信息安全学院　广西 桂林 541004
3 广西应用数学中心(GUET)　广西 桂林 541004
(guetzhx@163.com)

摘　要　对称密码算法通常采用安全指标良好的平衡布尔函数作为核心部件,以保障整个算法的安全性。使用启发式算法进行搜索是获得优良平衡布尔函数的一个重要途径。对此,设计了 Rank 排序混合遗传算法,用于搜索高非线性度、低自相关绝对值指标的平衡布尔函数。与传统的遗传算法相比,Rank 排序混合遗传算法在交叉阶段设计了交叉保护策略,以保障子代的平衡性;在选择步骤,采用基于适应度函数值的精英选择策略,以防止优秀个体流失;在进入下一轮迭代之前,设计了 Rank 排序环节,以增大下一轮进行交叉的个体间的差异,提高生成优秀子代的可能性,降低算法陷入局部最优解的风险。实验结果表明,以 6 至 14 元的偶变元数的布尔函数为搜索对象,使用 Rank 排序混合遗传算法均可搜索得到非线性度严格几乎最优、自相关绝对值指标低的平衡布尔函数。

关键词:平衡布尔函数;混合遗传算法;非线性度;自相关绝对值指标
中图分类号　TP306

Rank-sorting Hybrid Genetic Algorithm for Search High Quality Balanced Boolean Functions

ZHAO Haixia[1,3], LI Xin[1] and WEI Yongzhuang[2]
1 School of Mathematics and Computing Science, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China
2 School of Computer and Information Security, Guilin University of Electronic Technology, Guilin, Guangxi 541004, China
3 Center for Applied Mathematics of Guangxi(GUET), Guilin, Guangxi 541004, China

Abstract　The balanced Boolean functions with favorable security indicators are always used as core component in symmetric cipher, which can guarantee the overall security of cipher. One of important approaches to get high quality Boolean functions is searching by using heuristic algorithm. This paper designs Rank-sorting hybrid genetic algorithm to search balanced Boolean functions with high nonlinearity and low absolute autocorrelation values. Compared to traditional genetic algorithms, the following strategies and methods are designed and used in Rank-sorting hybrid genetic algorithm. Firstly, the crossover protection strategy is designed and used in the crossover phase, which can assure the balance of offspring. Secondly, elite selection strategy based on the value of fitness function is utilized in the selection step, in order to prevent the loss of excellent individuals. In particular, a sorting method named Rank-sorting algorithm is proposed and implemented on the selected offspring before they entering the next iteration, the result of using Rank-sorting algorithm is that the differences between individuals for the next crossover are increased, the possibility of generating excellent offspring is enhanced and the risk of the whole algorithm getting stuck in local optimal solutions is reduced. Experimental results show that for the Boolean functions with even number(6 to 14) of variables, balanced Boolean functions with almost optimal nonlinearity and low autocorrelation can be searched by using Rank-sorting hybrid genetic algorithm.

Keywords　Balanced Boolean function, Hybrid genetic algorithm, Nonlinearity, Absolute autocorrelation values

1　引言

对称密码算法是确保数据存储安全和通信信息安全的重要环节。布尔函数是对称密码算法的核心部件,其密码学指标,如非线性度、自相关性和弹性阶等,决定了密码算法能否有效抵御线性攻击、差分攻击及相关攻击等经典密码分析。布尔函数的各个密码学指标之间往往存在着制约关系,如何在各个指标之间进行权衡,获得兼顾多个安全指标的布尔函

到稿日期:2025-05-21　返修日期:2025-09-03
基金项目:国家自然科学基金(62402132)
This work was supported by the National Natural Science Foundation of China(62402132).
通信作者:韦永壮(walker_wyz@guet.edu.cn)

数，是密码学领域的重点研究问题。

获得安全布尔函数的方法大致可分为代数构造法与启发式搜索方法两类。用代数构造法通常能获得单一性质最佳的布尔函数，如用 MM 构造法[1]、谱不叠加方法[2]可构造得到 bent 函数（非线性度最高的布尔函数）。用代数构造法获得兼顾多个安全指标的布尔函数是一项极具挑战性的工作。2008 年，Carlet 等[3]首次构造出一类具备多项安全指标（最优代数免疫度、最大代数次数，非线性度至少为 $2^{n-1}-(n2^{n/2}\ln 4)/\pi$）的布尔函数；Tu 等[4]受 PS 类 bent 函数的启发，构造了一类非线性度至少为 $2^{n-1}-2^{n/2-1}-3\times(n/2)\times 2^{n/4}\ln 2-7$ 的一阶弹性函数，其代数次数能达到 $n-2$；Zhang 等[5]对 PS^- 类 bent 函数进行适当修改，提出了一种构造 1 阶弹性函数的方法，其构造的函数非线性度可达 $2^{n-1}-2^{n/2-1}-2^{\lceil n/4\rceil}$。代数构造法通常先针对某个密码学指标来构造函数，再在一定条件下对函数进行修改以兼顾其他密码学指标。在兼顾多项密码学指标方面，启发式搜索方法更具优势，该方法以多项密码学指标为目标，通过搜索局部解空间，选出最有可能成为最优解的候选解，搜索得到综合性能良好的布尔函数。经典的启发式搜索方法有爬山算法、模拟退火法、进化算法等。Millan 等[6]于 2007 年使用爬山算法来提高布尔函数的非线性度。随后 Millan 等[7]使用改进后的爬山算法获得了非线性度和自相关绝对值指标达到局部最优的布尔函数。Kuznetsov 等[8]将优化后的爬山算法与其所设计的成本函数相结合，有效减少了迭代次数，能够快速获得高非线性度向量布尔函数。Clark 等[9]指出通过选择恰当的代价函数，采用可避免局部最优的局部搜索算法（如模拟退火法），可获得高非线性度和低自相关绝对值指标的布尔函数。Clark 等[10]利用经典的密码学定理来改进模拟退火法，并使用改进后的算法搜索能兼顾代数次数、非线性度、自相关绝对值指标的布尔函数。Yang 等[11]以 1 阶弹性函数需满足的必要条件作为提前终止条件，使用模拟退火法搜索高非线性度的布尔函数，再使用文献[12]中的线性变换法获得高非线性度的 1 阶弹性函数。Asthana 等[13]使用代数次数、相关免疫度和非线性度 3 个指标定义了一个新的适应度函数，用于搜索 6 至 10 元的布尔函数。该方案能够在较少的迭代次数内，成功搜索到非线性度为 488 的 10 元平衡布尔函数。Mariot 等[14]将粒子群算法和爬山算法相结合，对 7 至 12 元的布尔函数进行了搜索；同时，采用局部单峰采样和连续遗传算法来优化粒子群的速度方程。通过这些方法，搜索到了非线性度达到 1972 的 12 元平衡布尔函数。爬山算法与模拟退火法大多只能在局部解空间中搜索局部最优解。在解决离散优化问题时，进化算法具有搜索范围广、并行性高、平衡冲突目标等优势，在搜索适用于对称密码算法的多指标良好布尔函数、逐重量完美平衡函数方面发挥着重要作用[15]。遗传算法作为进化算法的一个重要分支，是密码学学者搜索安全性能良好的布尔函数的常用手段。Picek 等[16]通过改进遗传算法中的初始化过程、变异算子及适应度函数，设计了对布尔函数的不同安全指标组合均行之有效的搜索方法。Picek 等[17]将代数构造法与遗传规划、遗传算法相结合，尝试解决非线性度为 118 的 8 元平衡布尔函数的搜索问题，发现用 bent 函数做初始种群可达到最佳效果。Manzoni 等[18]研究了在遗传算法中使用平衡交叉算子对搜索效果的影响，发现与使用单点交叉算子相比，使用平衡交叉算子能提升遗传算法的搜索效果。在 Zeki 等[19]设计的遗传算法中，将与某个 bent 函数汉明距离最近的平衡布尔函数作为初始种群，采用与 bent 函数 walsh 谱值相关的适应度函数，搜索得到了 8 至 26 元的高非线性度的平衡布尔函数。Carlet 等[20]将搜索空间限定在旋转对称布尔函数范围内，使用遗传算法成功演化出了一个非线性度为 241 的 9 元布尔函数。除了上述经典的启发式搜索方法之外，引力搜索法[21]、混合禁忌算法[22]在搜索综合指标良好的布尔函数方面也有着各自的优势。上述使用遗传算法搜索综合指标良好的布尔函数的相关研究结果表明，算法的初始化过程、遗传算子及适应度函数等环节均对搜索效果与效率有着至关重要的影响。

本文设计了 Rank 排序混合遗传算法，用于搜索非线性度高、自相关绝对值指标低的平衡布尔函数。首先，选择非线性度最高、自相关绝对值指标最低的 bent 函数作为初始种群。在交叉阶段，采用各轮随机选点的两点交叉方法将两个父代相应的基因段进行互换，在交叉之后，设计了交叉保护策略以保障子代的平衡性。在变异环节，对每个个体选择相同位置的基因段，随机生成各基因段的移位数，各基因段再进行相应位数的循环移位操作。由此，在不改变个体平衡性的情况下产生新个体。在选择步骤，采用基于适应度函数值的精英选择策略选出新的父代种群。将父代种群、交叉平衡之后的种群和变异后的种群合并，使用适应度函数从中选出 1/3 的个体作为子代，以防止优秀个体流失。在进入下一轮迭代之前，设计了 Rank 排序环节：先根据非线性度对选择出的个体划分等级，再根据自相关绝对值指标对同一个等级内的个体进行排序。在等级内部计算各个个体的拥挤度，根据拥挤度对等级内部的个体进行二次排序。经过 Rank 排序，可增大下一轮进行交叉的两个个体的绝对值指标之间的差异，提高生成优秀子代的可能性，同时避免出现过多的个体聚集在同一区域的局面，降低了整个算法陷入局部最优解的风险。

使用 Rank 排序混合遗传算法对 6－14 元的偶变元数的布尔函数进行搜索。不同变元数下，搜索得到的非线性度 N_f 与自相关绝对值指标 Δ_f 最优的布尔函数情况如表 1 所列。研究结果表明，对于不同变元数，使用 Rank 排序混合遗传算法均可搜索得到非线性度严格几乎最优、自相关绝对值指标低的平衡布尔函数。

表 1 Rank 排序混合遗传算法搜索得到的布尔函数指标（N_f，Δ_f）

Table 1 Indicators (N_f, Δ_f) of Boolean functions searched by using Rank-sorting hybrid genetic algorithm

	n					平衡
	6	8	10	12	14	
bent 函数	(28,0)	(120,0)	(496,0)	(2016,0)	(8128,0)	否
本文	(26,16)	(116,24)	(488,40)	(2000,72)	(8102,120) (8098,112)	是

本文第 2 章介绍布尔函数的基础知识；第 3 章阐述 Rank 混合遗传算法的设计细节、实施步骤及算法性能的理论分析；第 4 章分析实验结果；最后总结全文并展望未来。

2 预备知识

记二元有限域为 F_2，F_2 中的元素记为 0，1，加法运算符记为“$\oplus$”，n 维向量空间记为 F_2^n。

定义 1[23] 设 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in F_2^n$，$f(\boldsymbol{x})$ 是 F_2^n 到 F_2 的映射，称 $f(\boldsymbol{x})$ 为 n 元布尔函数，全体 n 元布尔函数所构成的集合记作 $\mathcal{B}_n$。

若将 F_2^n 中的元素简记为 $0,1,\cdots,2^n-1$，则称由 $f(\boldsymbol{x})$ 所确定的长为 2^n 的比特串 $f(0)f(1)\cdots f(2^n-1)$ 为 $f(\boldsymbol{x})$ 的真值表。真值表中“1”的个数称为 $f(\boldsymbol{x})$ 的汉明重量，记作 $wt(f)$，汉明重量为 2^{n-1} 的函数称为平衡函数。设 $g(\boldsymbol{x})\in\mathcal{B}_n$，定义 f 与 g 的汉明距离为 $wt(f\oplus g)$，记作 $d_{\mathrm{H}}(f,g)$。

除了真值表之外，代数正规形（ANF）也是布尔函数常用的表示方法之一。

定义 2[23] 设 $f(\boldsymbol{x})\in\mathcal{B}_n$，$\boldsymbol{x}=(x_1,x_2,\cdots,x_n)\in F_2^n$，$f(\boldsymbol{x})$ 的代数正规形为：

$$f(\boldsymbol{x})=\bigoplus_{\boldsymbol{u}\in F_2^n}\lambda_u\left(\prod_{j=1}^{n}x_j^{u_j}\right)\tag{1}$$

其中，$\boldsymbol{u}=(u_1,u_2,\cdots,u_n)\in F_2^n$，$\lambda_u\in F_2^n$。$f$ 的代数正规形中，单项式代数次数的最大值称为 f 的代数次数，记作 $\deg(f)$：

$$\deg(f)=\max\{wt(\boldsymbol{u})\mid\lambda_u\neq 0\}\tag{2}$$

代数次数为 1 的布尔函数称为仿射函数，全体仿射函数所构成的集合记为 $A_n[\boldsymbol{x}]$。

布尔函数的 Walsh 变换是研究布尔函数性质的有力工具，布尔函数的安全指标，如非线性度、绝对值指标等均与其 Walsh 谱值存在关联，这意味着可通过 Walsh 谱值研究布尔函数的性质。

定义 3[24] 设 $f(\boldsymbol{x})\in\mathcal{B}_n$，定义 F_2^n 到 R 的映射 W_f：对于 $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in F_2^n$，有：

$$W_f(\boldsymbol{\alpha})=\sum_{x\in F_2^n}(-1)^{f(x)\oplus x\cdot\alpha}\tag{3}$$

其中，$\boldsymbol{x}\cdot\boldsymbol{\alpha}=x_1\alpha_1\oplus x_2\alpha_2\oplus\cdots\oplus x_n\alpha_n$，$\Sigma$ 是实数域 R 上的求和运算。W_f 为 f 的 Walsh 变换，$W_f(\boldsymbol{\alpha})$ 为 f 在 $\boldsymbol{\alpha}$ 处的 Walsh 谱值。

定理 1[24]（Parseval 恒等式） 设 $f(\boldsymbol{x})\in\mathcal{B}_n$，$W_f$ 为 f 的 Walsh 变换，则有：

$$\sum_{\boldsymbol{\alpha}\in F_2^n}W_f^2(\boldsymbol{\alpha})=2^{2n}\tag{4}$$

定义 4[24] 设 $f(x)\in\mathcal{B}_n$，f 与 n 元仿射函数的汉明距离的最小值称为 f 的非线性度，记作 N_f，即 $N_f=\min\limits_{l\in A_n[x]}d_{\mathrm{H}}(f,l)$。$N_f$ 与 f 的 Walsh 谱值存在如下关系：

$$N_f=2^{n-1}-2^{-1}\max_{a\in F_2^n}|W_f(\boldsymbol{\alpha})|\tag{5}$$

由式(4)与式(5)可得，对任意 $f(\boldsymbol{x})\in\mathcal{B}_n$，有：

$$N_f\leqslant 2^{n-1}-2^{n/2-1}\tag{6}$$

非线性度达到 $2^{n-1}-2^{n/2-1}$ 的布尔函数称为 bent 函数，其 Walsh 谱值只可能为 $2^{n/2}$ 或 $-2^{n/2}$。非线性度小于 $2^{n-1}-2^{n/2-1}$ 且大于 $2^{n-1}-2^{\lfloor n/2\rfloor}$ 的布尔函数称为是严格几乎最优的。

定义 5[24] 设 $f(\boldsymbol{x})\in\mathcal{B}_n$，定义 F_2^n 到 R 的映射 C_f：对于 $\boldsymbol{\alpha}=(\alpha_1,\alpha_2,\cdots,\alpha_n)\in F_2^n$，有：

$$C_f(\boldsymbol{\alpha})=\sum_{x\in F_2^n}(-1)^{f(x)\oplus f(x\oplus\alpha)}\tag{7}$$

其中，Σ 是实数域 R 上的求和运算，C_f 为 f 的自相关函数。

绝对值指标 Δ_f 与平方和指标 σ_f 是度量 f 自相关性的重要指标，其定义如下：

$$\Delta_f=\max_{\boldsymbol{\alpha}\neq 0,\boldsymbol{\alpha}\in F_2^n}|C_f(\boldsymbol{\alpha})|\tag{8}$$

$$\sigma_f=\sum_{\boldsymbol{\alpha}\in F_2^n}C_f^2(\boldsymbol{\alpha})\tag{9}$$

Δ_f 与 σ_f 越小，f 的自相关性越好，这两个指标的上、下界分别为 $0\leqslant\Delta_f\leqslant 2^n$，$2^{2n}\leqslant\sigma_f\leqslant 2^{3n}$。对于 $n(n\geqslant 3)$ 元平衡函数，这两个指标的下界均有所提升，为 $8\leqslant\Delta_f$，$2^{2n}+2^{n+3}\leqslant\sigma_f$。

定理 2[25] 设 $\boldsymbol{b}\in F_2^n$，C_f 是 f 的自相关函数，W_f 是 f 的 Walsh 变换，则：

$$W_f(\boldsymbol{b})=\sum_{\boldsymbol{\alpha}\in F_2^n}C_f(\boldsymbol{\alpha})(-1)^{\boldsymbol{b}\cdot\boldsymbol{\alpha}}\tag{10}$$

由式(4)与式(10)可得：

$$\sum_{\boldsymbol{\alpha}\in F_2^n}W_f^4(\boldsymbol{\alpha})=2^n\sigma_f\tag{11}$$

由式(11)可得，bent 函数的平方和指标为 2^{2n}，相应地，绝对值指标为 0。综上，bent 函数具有最高的非线性度与最佳的自相关性。

3 Rank 排序混合遗传算法

遗传算法的基本原理最早是由 Holland[26] 提出的，他借鉴了达尔文的繁殖、适者生存和自然选择的遗传操作原则，通过遗传操作将种群中的个体转化为新一代的种群。在遗传算法中，首先产生初始种群，然后根据适者生存和优胜劣汰原则，逐代演化产生出近似最优解。具体地，在每一代中，依据适应度值对个体进行选择，再对选出的个体进行交叉和变异来生成新的种群，经过多代进化后，在末代种群中选出最优个体作为问题的最优解。

遗传算法在解决离散优化问题时具有搜索范围广、并行性高、鲁棒性强等优势，因此可用于搜索具备多项安全指标的布尔函数。然而，在搜索过程中会遇到因过早收敛而陷入局部最优解的问题。导致该问题的主要原因有：1）选择操作、交叉操作和变异操作选择不恰当；2）当布尔函数的变元个数 n 较大时，无法快速搜索到近似最优解。

本文基于遗传算法搜索高非线性度、低自相关度的平衡布尔函数，设计了 Rank 排序算法以在较小区域内快速找到局部最优解，将其与遗传算法相结合提出了一种新的混合遗传算法。算法的具体流程如图 1 所示。

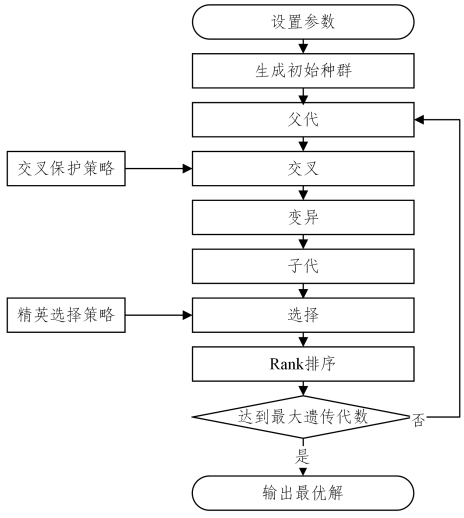

图 1 Rank 排序混合遗传算法流程图

Fig. 1 Flow diagram of Rank-sorting hybrid genetic algorithm

3.1 算法细节设计

布尔函数可由其真值表唯一确定，根据真值表可计算得

到布尔函数的各项安全指标，故用真值表作为染色体的基因序列。n 元布尔函数的真值表是长为 2^n 的比特串，因此采用二进制编码。如表 2 所列，随着变元数 n 的增大，搜索问题的解空间的基数呈指数增长。

表 2 n 元布尔函数的解空间的基数

Table 2 Cardinality of solution space of Boolean function with n variables

n	解空间基数
6	2^{64}
8	2^{256}
10	2^{1024}
12	2^{4096}
14	2^{16384}

3.1.1 初始种群的选取

初始种群的选取往往会影响搜索效果，若以高非线性度、低自相关绝对值指标的平衡布尔函数为搜索目标，则选取非线性度高、自相关绝对值指标低的布尔函数作为初始种群能获得好的搜索效果。bent 函数是非线性度最高且自相关绝对值指标最低的函数，因此选择 bent 函数作为初始种群。同时，考虑到 bent 函数的不平衡性，设计了交叉保护策略，在交叉操作之后，使用该策略来保障子代的平衡性。

3.1.2 交叉操作

交叉操作是用两个父代来产生两个子代，常用的交叉机制有一点交叉、两点交叉、均匀交叉、部分匹配交叉。平衡性是布尔函数最基本的一项安全指标，为保证解的平衡性，在交叉环节设计了交叉保护策略，并将其与两点交叉机制相结合来获得平衡子代。

传统的两点交叉可能无法保证解的平衡性，而结合交叉保护策略可以在保持多样性的同时，确保子代继承了父代的某些关键特征，并且保证了平衡性。

两点交叉是指从父代种群中依次不放回地选择两个父代进行交叉，产生新的后代个体，从而增加种群的多样性、扩大搜索空间。将进行交叉的两个父代个体分别记为 y_1 和 y_2，交叉后产生的子代个体分别记为 c_1 和 c_2。两点交叉的具体流程如下：

1)随机选取两个交叉点 r_1 和 r_2，$0\leqslant r_1<r_2\leqslant 2^n$；

2)互换两个父代在 r_1 和 r_2 之间的部分(含 r_1，r_2)生成子代 c_1 和 c_2。

交叉保护策略的实施过程为：统计 $c_i(i=1,2)$ 中 0 和 1 的数量，若 0 的数量比 1 的数量多 k，则在 c_i 中随机选取 $k/2$ 个 0 变为 1；反之亦然。

3.1.3 变异操作

在使用遗传算法解决排列优化问题时，变异操作是引入种群多样性的关键步骤，可防止解的过早收敛。本文采用了一种基于移位变异的变异操作，旨在保持布尔函数在变异过程中的平衡性。除了移位变异，常见的变异机制还包括翻转变异、交换变异、反转变异。与传统的翻转变异、交换变异、反转变异不同，本文算法采用的移位变异能够更快地定位到最优解所在的区域，在减少无效搜索的同时还能保持布尔函数的平衡性，从而提高解的多样性和算法的全局搜索能力。本文所采用的移位变异的具体流程如下：

1)随机选取两个变异点 z_1 和 z_2，$0\leqslant z_1<z_2\leqslant 2^n$；

2)随机生成移位数 $p(0\leqslant p\leqslant z_2-z_1+1)$；

3)将子代 c_i 染色体的第 z_1 和 z_2 之间的比特循环右移 p 位，得到变异后的子代。

3.1.4 选择操作

选择操作的目的是选出种群中的优秀个体，使得种群朝着更优解的方向演化。本文采用基于适应度函数值的精英选择策略来选出新的父代种群。精英选择策略的具体实施流程为：首先将父代、变异后的子代，以及交叉平衡之后的子代合并成扩张种群；然后根据个体的适应度值对扩张种群中的个体以升序排序；最后选出前三分之一的个体作为新父代。采用精英选择策略生成的父代既不会遗漏在上一轮遗传过程中的优秀个体，又能保留本轮遗传过程中产生的优秀新个体。

本文的搜索目标是具有高非线性度和低自相关绝对值指标的布尔函数。由式(5)可知，函数 Walsh 谱值绝对值的最大值越小，函数的非线性度就越高；由式(11)可知，函数 Walsh 谱值 4 次方和越小，函数的平方和指标就越低，进而自相关绝对值指标就可能越低。根据这两个观察结果，定义适应度函数为：

$$fit(f)=\sum_{x\in F_2^n}2^{-4n}(w_f(\boldsymbol{x}))^4 \tag{12}$$

3.1.5 Rank 排序算法

Rank 排序算法对新父代种群中全部个体进行排序，具体步骤如下。

步骤 1 计算新父代种群各个个体的非线性度 N_{f_i} 和自相关绝对值指标 Δ_{f_i}，f_i 为第 i 个个体。

步骤 2 非线性度相等的个体被认定在同一个 Rank 等级，非线性度最大的个体放在第一个等级 Rank_1 中，非线性度次大的个体放在第二个等级 Rank_2 中，以此类推。之后对各个 Rank_i 中的个体进行等级内部排序，具体地，按其自相关绝对值指标以升序排序。

步骤 3 计算种群中所有个体的拥挤距离 $Cd(f_i)$。对于各个 Rank_i，若该等级中个体的数量小于或等于 2，则该等级中所有个体的拥挤距离均设置为 10 000；若该等级中个体的数量大于 2，则将第一个和最后一个个体的拥挤距离均设置为 10 000，其余个体的拥挤距离计算式为 $Cd(f_i)=\Delta_{f_{i+1}}-\Delta_{f_{i-1}}$。

步骤 4 对各个 Rank_i 中的个体进行等级内部排序，具体地，按其拥挤距离以降序排序。

拥挤距离反映了基因个体与其相邻个体之间的拥挤程度，拥挤距离越大，表明种群中个体的分布越分散。在同一等级的个体按其拥挤距离以降序排序后，拥挤距离大的个体将排在前面。如图 1 所示，若未达到最大遗传代数，则所有等级中的个体将成为下一轮的父代。如图 2 所示，在下一轮的交叉中，由上而下，相邻的两个父代进行交叉。Rank 排序算法可使得下一轮进行交叉的两个个体的绝对值指标之间的差异较大，从而提高生成优秀子代的可能性。Rank 排序算法可避免出现过多的个体聚集在同一区域的局面，降低了算法陷入局部最优解的风险。此外，在计算拥挤距离时仅需考虑自相关绝对值指标，故排序速度快。

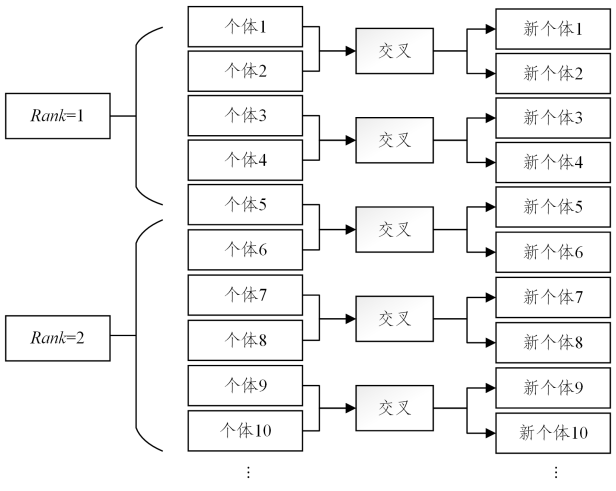

图 2 Rank 等级排序及下一轮交叉方式

Fig. 2 Rank level ranking and the cross-over method in the next iteration

通过量化不同变元情况下算法的运行时间，可说明所提算法的运行效率与性能，算法执行一轮的耗时如表 3 所列。

表 3 算法一轮循环时间

Table 3 Per-iteration loop time of the algorithm

n	运行时间/s
6	1
8	1
10	40
12	177
14	963

3.2 算法具体步骤

Rank 排序混合遗传算法的具体步骤如下。

步骤 1 设置参数：设置初始种群数量 POP、父代数量 PAP、布尔函数的变元个数 n、个体染色体长度 2^n，以及最大迭代次数 $MAXITER$。

步骤 2 生成初始种群：采用 Maiorana-McFarland 构造法随机生成 POP 个 n 元 *bent* 函数，以其真值表作为个体染色体。

步骤 3 选择父代：将初始种群的前 PAP 个个体作为父代。

步骤 4 父代进行交叉，交叉后实施交叉保护策略：从父代种群中依次不放回地选择两个个体，对所选出的两个个体进行两点交叉，得到两个新个体；对交叉环节后得到的全部新个体实施交叉保护策略。

步骤 5 变异操作：对每个个体执行移位变异操作（随机选取每个个体的变异点，随机生成每个个体移位数 p）。

步骤 6 采用精英选择策略进行选择：将父代、变异后的子代及交叉平衡后的子代合并为扩张种群，记为 HB；计算 HB 中各个个体的适应度值 $fit(f_i)$，个体按适应度值以升序排序；选出前 PAP 个个体作为新父代。

步骤 7 Rank 排序：对个体进行 Rank 排序（详见 3.1.5 节算法步骤），得到下一轮的父代，迭代次数 $ITER$ 加 1。

步骤 8 判断：判断当前迭代次数 $ITER$ 是否大于最大迭代次数 $MAXITER$，若 $ITER$ 大于 $MAXITER$，则跳到步骤 9 出最优解，否则跳到步骤 4。

步骤 9 输出：输出当前种群中的最优个体。

3.3 算法性能的理论分析

Rank 排序混合遗传算法以布尔函数的相关概念、性质及结论为依据，紧紧围绕高非线性度、低自相关绝对值指标以及平衡性这 3 个目标设计各个环节。

遗传算法本身的设计理念决定了初始种群的选取将影响搜索效果，采用与目标性能接近的函数作为初始种群，能提高获得标的函数的可能性。本文旨在搜索高非线性度、低自相关绝对值指标的平衡布尔函数，因此选择 bent 函数作为初始种群（此类函数具有最高的非线性度和最低的自相关绝对值指标，但不平衡），这为获得标的函数提供了保障。

平衡函数的真值表中 0 和 1 各占一半，与之对应的基因序列亦然。然而，作为初始种群的 bent 函数并不具备该特征。因此，基于两点交叉机制设计了交叉保护策略：统计子代个体 c 中 0 和 1 的数量，若 0 的数量比 1 的数量多 k，则在 c 中随机选取 $k/2$ 个 0 变为 1，反之亦然。由此，在确保子代继承父代的某些关键特征（高非线性度、低自相关绝对值指标）的同时也获得了平衡性。进一步地，采用变异点随机选取、移位数随机生成的循环右移的变异操作，既能保持布尔函数的平衡性，又能快速定位到最优解所在的区域，减少无效搜索。

选择操作依据适应度函数执行，好的适应度函数能准确捕捉最优解的特征，以驱动种群向目标进化，避免过早收敛，提高搜索效率。因此，定义适应度函数是遗传算法设计的核心环节。由式(5)与式(11)可知：布尔函数的非线性度与其 Walsh 谱值绝对值的最大值成反比；布尔函数的自相关绝对值指标与其 Walsh 谱值 4 次方和成正比。根据“高非线性度、低自相关绝对值指标布尔函数”的搜索目标，定义适应度函数为 $fit(f)=\sum_{x\in F_2^n}2^{-4n}(w_f(\boldsymbol{x}))^4$，适应度值低的个体作为优秀个体被留存。同时，还采用了“父代、变异后的子代，以及交叉平衡之后的子代合并成扩张种群，按升序排列后取前 1/3 作为新父代”的精英选择策略，以保障上一轮的优秀个体与本轮产生的优秀新个体均被留存。

4 实验及结果分析

本文实验环境在 windows 10，主频 3.6 GHz 的 Intel(R) Core(R) Gold 6256 处理器，内存为 128 GB 的台式机上进行，编程软件采用 vs code。

布尔函数是对称密码算法唯一的非线性部件，对密码算法的安全性起着至关重要的作用。例如，数据加密标准(DES)采用了 8 个独特的 S-Box，每个 S-Box 实现了一个从 6 位输入到 4 位输出的布尔函数映射，从而将 6 位的输入转换为 4 位的输出。高级加密标准(AES)则使用了一个 8 位输入到 8 位输出的 S-Box，以实现更加复杂的非线性变换。同样，PRESENT 算法采用了 16 个 4 位到 4 位的 S-Box，每个 S-Box 负责将 4 位输入映射为 4 位输出。为了深入探索布尔函数的密码学特性，本文分别对 $n=6,8,10,12,14$ 元的布尔函数进行搜索。

搜索得到的非线性度最高的函数的非线性度情况如表 4

所列，可见，搜索得到的布尔函数均为严格几乎最优函数。

表 4 搜索得到的布尔函数非线性度情况

Table 4 Nonlinearities of the Boolean functions searched by using the algorithm

n	最优($2^{n-1}-2^{n/2-1}$)	严格几乎最优($>2^{n-1}-2^{\lfloor n/2 \rfloor}$)	本文方法
6	28	24	26
8	120	112	116
10	496	480	488
12	2016	1984	2000
14	8128	8064	8102

对 10 元布尔函数的 30 次仿真结果的非线性度进行了统计(按算法设计，一次仿真只保留一个最优结果)，其分布如图 3 所示。在 30 次结果中，非线性度最低为 480，最高为 488；非线性度 484 及以上的占 1/3。由此可见，使用该算法可搜索得到非线性度优良的布尔函数。

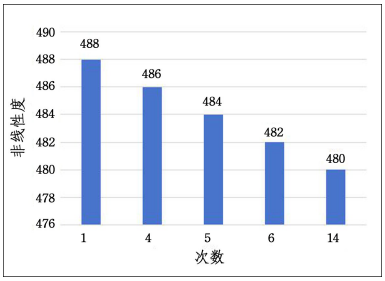

图 3 30 次仿真结果的非线性度分布情况

Fig. 3 Distribution of nonlinearity for the 30 simulation results

搜索得到的两项指标最好的布尔函数的非线性度 N_f 与自相关绝对值指标 Δ_f 的情况如表 5 所列，并将其与其他文献的结果进行了对比。在非线性度指标上，本文方法的结果优于文献[4]、文献[11](10 元及以上)、文献[22](12 元及以上)；在自相关绝对值指标上，对于 10 元及以上的布尔函数，本文方法的结果优于其他文献。综合两项指标来看，本文方法的结果优于文献[4]、文献[11]，以及文献[22]；8 元和 10 元的结果与文献[21]一致，但文献[21]未搜索 12 元和 14 元的布尔函数。与文献[5]相比，本文方法在 10 元及以下的结果均更优；搜索 12 元时，非线性度略低于文献[5]，但自相关度显著低于文献[5]；搜索 14 元时，非线性度略低于文献[5]，但文献[5]未给出相应的自相关绝对值指标。

算法搜索得到的某些布尔函数同时具备一阶弹性，例如，可搜索得到(N_f, Δ_f)=(24, 24)的 6 元一阶弹性布尔函数，其真值表为 6895 9766 877A 3C68；可搜索得到(N_f, Δ_f)=(116, 32)的 8 元一阶弹性布尔函数，其真值表为 7EB8 CAC4 B4CC A434 CB75 B1A5 C961 0F35 E48A 6A26 9F53 9937 A082 7D57 00FF 0F36。由于非线性度、自相关绝对值指标与一阶弹性之间存在一定的制约关系，因此具备一阶弹性的布尔函数的非线性度和自相关绝对值指标无法达到最优的结果。

16 元布尔函数的真值表长度为 65 536，这使得在执行交叉和变异两个环节时的计算量远超 14 元布尔函数时的计算量(14 元布尔函数真值表长度为 16 384)。16 元布尔函数的搜索空间大小为 2^{65536}，这使得算法搜索过程中可能会遇到效率低下和收敛速度慢的问题。

表 5 不同方法非线性度与自相关绝对值指标(N_f, Δ_f)对比

Table 5 Comparison of nonlinearities and absolute autocorrelation values (N_f, Δf) for different methods

n	文献[4]	文献[5]	文献[11]	文献[21]	文献[22]	本文
6	(24,—)	—	—	—	—	(26,16)
8	(112,—)	(116,32)	(116,48)	(116,24)	(116,24)	(116,24)
10	(484,—)	(488,—)	(484,96)	(488,40)	(488,72)	(488,40)
12	(1996,—)	(2008,96)	(1988,184)	—	(1996,152)	(2000,72)
14	(8100,—)	(8112,—)	(8072,368)	—	(8080,320)	(8102,120) (8098,112)

结束语 本文提出了一种 Rank 排序混合遗传算法来搜索具有高非线性度、低自相关绝对值指标的平衡布尔函数。该算法在搜索 6 至 14 元偶变元数布尔函数时，表现出了良好的性能。该算法也存在一定的局限性。首先，随着布尔函数变元个数的增加，搜索空间的大小呈指数级增长，导致算法的计算量和搜索难度显著提高。当前算法在处理高变元布尔函数时可能面临效率低下和收敛速度慢的问题。其次，虽然本文算法在一定程度上维持了解的多样性和平衡性，但在面对更高变元的布尔函数时，可能需要更高效的交叉和变异策略来保持算法的有效性。未来工作将针对以上局限性，进一步优化 Rank 排序混合遗传算法，以搜索更高变元的综合指标良好的布尔函数。

参 考 文 献

[1] MCFARLAND R L. A Family of Difference Sets in Non-Cyclic Groups[J]. Journal of Combinatorial Theory, Series A, 1973, 15(1): 1-10.

[2] ZHAO H, WEI Y, ZHANG F, et al. Two Secondary Constructions of Bent Functions without Initial Conditions[J]. Designs, Codes and Cryptography, 2022, 90(3): 653-679.

[3] CARLET C, FENG K. An Infinite Class of Balanced Functions with Optimal Algebraic Immunity, Good Immunity to Fast Algebraic Attacks and Good Nonlinearity[C]// Proceedings of International Conference on the Theory and Application of Cryptology and Information Security. Berlin: Springer, 2008: 425-440.

[4] TU Z R, DENG Y P. Boolean Functions Optimizing Most of the Cryptographic Criteria[J]. Discrete Applied Mathematics, 2012, 160(4/5): 427-435.

[5] ZHANG W G, PASALIC E. Improving the Lower Bound on the Maximum Nonlinearity of 1-Resilient Boolean Functions and Designing Functions Satisfying All Cryptographic Criteria[J]. Information Sciences, 2017, 376: 21-30.

[6] MILLAN W, CLARK A, DAWSON E. Smart Hill Climbing Finds Better Boolean Functions[EB/OL]. https://api.semanticscholar.org/CorpusID:17217144.

[7] MILLAN W,CLARK A,DAWSON E. Boolean Function Design Using Hill Climbing Methods[C]// Proceedings of Information Security and Privacy: 4th Australasian Conference. Berlin: Springer,1999:1-11.

[8] KUZNETSOV A,POLUYANENKO N,PELIUKH K O. A New Cost Function for Heuristic Search of Nonlinear Substitutions[J]. Expert Systems with Application,2024,237:1-14.

[9] CLARK J A,JACOB J L. Two-Stage Optimisation in the Design of Boolean Functions[C]// Proceedings of Australasian Conference on Information Security and Privacy. Berlin: Springer, 2000:242-254.

[10] CLARK J A,JACOB J L,STEPNEY S,et al. Evolving Boolean Functions Satisfying Multiple Criteria[C]// Proceedings of Cryptology — INDOCRYPT 2002. Berlin: Springer, 2002: 246-259.

[11] YANG J P,ZHANG W G. Generating Highly Nonlinear Resilient Boolean Functions Resistance Against Algebraic and Fast Algebraic Attacks[J]. Security and Communication Networks, 2015,8(7):1256-1264.

[12] MAITRA S,PASALIC E. Further Constructions of Resilient Boolean Functions with Very High Nonlinearity[M]. London: Springer,2002:17-35.

[13] ASTHANA R,VERMA N,RATAN R. Generation of Boolean Functions Using Genetic Algorithm for Cryptographic Applications[C]// Proceedings of IEEE International Advance Computing Conference (IACC). New York: Piscataway, 2014: 1361-1366.

[14] MARIOT L,LEPORATI A,MANZONI L. A Discrete Particle Swarm Optimizer for the Design of Cryptographic Boolean Functions[J]. arXiv:2401.04567,2024.

[15] MANDUJANO S,LARA A,CAUICH J K. Using Evolutionary Algorithms for the Search of 16-Variable Weight-Wise Perfectly Balanced Boolean Functions with High Non-linearity[C]// Proceedings of International Conference on Parallel Problem Solving from Nature. Berlin:Springer,2024:416-428.

[16] PICEK S,JAKOBOVIC D,GOLUB M. Evolving Cryptographically Sound Boolean Functions[C]// Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation. New York:ACM,2013:191-192.

[17] PICEK S,MARCHIORI E,BATINA L,et al. Combining Evolutionary Computation and Algebraic Constructions to Find Cryptography-Relevant Boolean Functions[C]// Proceedings of Parallel Problem Solving from Nature-PPSN XIII:13th International Conference. Berlin:Springer,2014:822-831.

[18] MANZONI L,MARIOT L,TUBA E. Does Constraining the Search Space of Ga Always Help? The Case of Balanced Crossover Operators[C]// Proceedings of the Genetic and Evolutionary Computation Conference Companion. New York: ACM, 2019:151-152.

[19] ZEKI E,KAVUT S,KUTUCU H. Genetic Approach to Improve Cryptographic Properties of Balanced Boolean Functions Using Bent Functions[J]. Computers,2023,12(8):1-14.

[20] CARLET C,URASEVIC M,JAKOBOVIC D,et al. A Systematic Evaluation of Evolving Highly Nonlinear Boolean Functions in Odd Sizes[C]// Proceedings of Genetic Programming: 28th European Conference. Berlin:Springer,2025:18-34.

[21] JIA S S,ZHANG F R. Boolean Function Generation Algorithm Based on Gravitational Search[J]. Application Research of Computers,2021,38(2):430-434.

[22] WANG W Q,XU H J,CUI M,et al. Mixed Tabu Search Algorithm for Excellent Boolean Functions[J]. Journal on Communications,2022,43(5):133-143.

[23] CARLET C. Boolean Functions for Cryptography and Error Correcting Codes[C]// Boolean Models and Methods in Mathematics,Computer Science,and Engineering. Cambridge University Press,2006:257-397.

[24] ZHOU Y,HU Y P,DONG X F. Design and Analysis of Boolean Function[M]// Beijing: National Defense Industry Press,2015: 20-87.

[25] CARLET C. Partially-bent Functions [J]. Designs,Codes and Cryptography,1993,3:135-145.

[26] HOLLAND J H. Genetic Algorithms[J]. Scientific American, 1992,267(1):66-73.

ZHAO Haixia, born in 1981, Ph.D, associate professor. Her main research interests include Boolean function and design and analysis of symmetric ciphers.

WEI Yongzhuang, born in 1976, Ph.D, professor. His main research interest is cryptographic algorithm design and analysis.

(责任编辑:何杨)