计算机科学
COMPUTER SCIENCE

基于 DQN 增强遗传算法的 Plateaued 函数高效构造研究

吴严生, 曹心怡, 樊卫北

引用本文

吴严生, 曹心怡, 樊卫北. 基于 DQN 增强遗传算法的 Plateaued 函数高效构造研究[J]. 计算机科学, 2026, 53(4): 57-65.

WU Yansheng, CAO Xinyi, FAN Weibei. Research on Efficient Construction of Plateaued Functions Based on DQN-enhanced Genetic Algorithm [J]. Computer Science, 2026, 53(4): 57-65.

计算机科学 COMPUTER SCIENCE
http://www.jsjkx.com
DOI:10.11896/jsjkx.251100083

基于 DQN 增强遗传算法的 Plateaued 函数高效构造研究

吴严生　曹心怡　樊卫北
南京邮电大学计算机学院、软件学院、网络空间安全学院　南京 210023

摘　要　作为 Bent 函数的重要推广，Plateaued 函数继承了很多 Bent 函数的优良密码学性质，具有重要的应用价值。由于传统构造 Plateaued 函数的方法存在计算复杂度高、灵活性不足等问题，因此提出一种基于深度 Q 网络(Deep Q-Network，DQN)增强的自适应遗传算法。该算法深度融合 DQN 与遗传算法，构建多维状态空间感知种群进化特征，通过群体共识机制智能选择 6 种交叉与变异策略组合，实现遗传参数的自适应调控。实验结果表明，该算法的适应度提升幅度达 0.20～0.35，收敛速度更快，稳定性更高，平均可生成 230～300 个有效 Plateaued 函数真值序列，显著优于标准遗传算法和基础 Q-learning 遗传算法。算法能智能调节变异率(0.235～0.276)与交叉操作使用率(70%～90%)，在优化 Walsh 谱分布的同时保持种群多样性。尽管计算开销略有增加，但所提算法在解的质量、收敛性能和策略自适应能力上具有显著优势，验证了深度强化学习在密码学函数构造中的有效性，为布尔函数智能化设计提供了新方案。

关键词：Plateaued 函数；真值序列；Q-learning；深度 Q 网络；遗传算法；Walsh 谱；非线性度
中图分类号　TP181

Research on Efficient Construction of Plateaued Functions Based on DQN-enhanced Genetic Algorithm

WU Yansheng，CAO Xinyi and FAN Weibei
College of Computer Science，Nanjing University of Posts and Telecommunications，Nanjing 210023，China

Abstract　Plateaued functions，as an important generalization of Bent functions，inherit many of the desirable cryptographic properties of Bent functions and hold significant application value. However，traditional methods for constructing plateaued functions suffer from issues such as high computational complexity and limited flexibility. To address these challenges，this paper proposes an adaptive genetic algorithm enhanced by a deep Q-Network(DQN). This algorithm deeply integrates DQN with the genetic algorithm，constructing a multi-dimensional state space to perceive population evolutionary characteristics. Through a group consensus mechanism，it intelligently selects from six combinations of crossover and mutation strategies，enabling adaptive control of genetic parameters. Experimental results demonstrate that the proposed algorithm achieves a fitness improvement of 0.20～0.35，exhibits faster convergence speed and higher stability，and can generate an average of 230～300 valid Plateaued function truth sequences，significantly outperforming both the standard genetic algorithm and the basic Q-learning-enhanced genetic algorithm. The algorithm intelligently adjusts the mutation rate within the range of 0.235～0.276 and maintains the crossover operation usage rate between 70% and 90%，effectively optimizing the Walsh spectrum distribution while preserving population diversity. Although computational overhead increases slightly，its significant advantages in solution quality，convergence performance，and strategic adaptability validate the effectiveness of deep reinforcement learning in the construction of cryptographic functions，providing a novel approach for the intelligent design of Boolean functions.

Keywords　Plateaued function，Truth sequence，Q-learning，Deep Q-Network，Genetic algorithm，Walsh spectrum，Nonlinearity

1　引言

在数字信息技术迅猛发展的背景下，信道编码理论面临新的技术挑战，推动着纠错编码研究持续发展。从 Shannon[1] 提出有噪信道编码理论奠定基础，到 Hamming[2] 提出实用编码方法，再到有限域等数学工具[3-4]的引入推动理论突破，纠错编码技术经历了从理论奠基到实践应用的重要发展历程。近年来，Plateaued 函数[5-6]等复杂数学工具的引入，进一步丰富了编码设计手段。其特殊的 Walsh 谱特性，使其在密码学系统和编码构造中展现出独特优势，为现代编码理论

到稿日期：2025-11-17　返修日期：2026-01-16
基金项目：国家自然科学基金(62372247)；江苏省研究生科研与实践创新计划项目(SJCX24_0326)
This work was supported by the National Natural Science Foundation of China(62372247) and Jiangsu Province Postgraduate Research and Practice Innovation Program(SJCX24_0326).
通信作者：吴严生(yanshengwu@njupt.edu.cn)

的发展提供了新的研究方向。

Plateaued 函数因具有特殊的 Walsh 谱特性而备受关注，其谱值仅取{0，±λ}3 个值的性质，使其在密码学系统和编码构造中具有重要价值。然而，传统的构造方法[7-9]存在计算复杂度高、灵活性不足等固有局限，难以适应高维复杂优化问题的求解需求，这促使研究者寻求更高效的智能优化方法来解决 Plateaued 函数的构造难题。

遗传算法虽在各领域广泛应用，但其性能高度依赖参数设置，而传统的固定参数或简单自适应策略难以实现基于搜索状态的智能调整。Q-learning 作为强化学习的重要分支，通过与环境交互学习最优策略的特性，为遗传算法参数的动态调整提供了新思路。尽管现有研究已尝试将强化学习与遗传算法结合，实际应用于生产调度[10-12]、公共卫生与应急管理[13-14]、组合优化[15]、路径规划[16]等多个领域，但在理论体系的完整性和多目标协同优化等方面仍存在深入探索的空间。

为突破传统方法的局限性，遗传算法被引入 Plateaued 函数的构造过程中[17]，通过模拟自然进化机制在解空间中进行全局搜索。但算法的搜索性能高度依赖于交叉率、变异率等关键参数的设置，传统静态参数策略难以适应搜索过程的动态特性。本研究通过引入 DQN，对遗传算法进行了关键改进。两者深度协同，构建了一个动态平衡探索与利用的自适应系统，最终显著提升了 Plateaued 函数的演化效率与质量。

2 相关工作

Plateaued 函数的构造方法已形成三大主流研究方向。一是基于 Maiorana-McFarland(MM)类[18]的构造，这是最经典且成熟的方法。该方法借助 MM 类函数“级联线性函数与置换”的核心结构特点，通过严格限制置换的性质，实现对 Walsh 谱分布的精确调控，能够构造出高代数度与高非线性度的 Plateaued 函数。二是谱方法与特征函数构造[19]。该方法从 Walsh 谱特性出发，通过分析 Walsh 谱的支撑集构造规律反推函数的布尔真值表。三是递归构造与二次构造[20]。该方法利用已知的低维 Plateaued 函数，通过直接和、间接和等组合方式生成高维 Plateaued 函数。这些方法在理论上具有严密性，但通常计算复杂度高，灵活性有限，难以适应高维或动态优化场景。

近年来，随着智能优化算法的兴起，研究者开始尝试将遗传算法、模拟退火、粒子群优化等启发式方法引入密码学函数的构造中。Mariot 等[21]将遗传算法用于演化 Plateaued 布尔函数，通过适应度函数引导搜索，显著提升了构造效率；Picek 等[22]则系统研究了进化算法在多种密码学布尔函数设计中的应用，验证了其在高非线性度、平衡性等指标优化上的潜力。这一方向的研究不断深化，衍生出多种混合启发式策略与自适应优化框架[23-24]，以同时优化布尔函数的多个密码学性质，并尝试解决高维搜索空间下的效率与收敛问题。

与此同时，强化学习与进化算法的融合为优化过程的自适应调控提供了新思路。在遗传算法中，交叉率、变异率等参数对搜索性能具有关键影响，传统方法多采用静态设置或简单自适应策略，难以实现基于搜索状态的动态调整。Q-learning 等强化学习技术通过与环境的交互学习，实现参数的在线优化，已在作业调度、路径规划和多目标优化等领域得到成功应用。然而，在密码学函数构造这一特定问题上，如何将深度强化学习与遗传算法深度结合，实现状态感知、策略选择与奖励机制的协同优化，仍是一个有待深入探索的课题。在此背景下，本研究提出一种基于 DQN 增强的自适应遗传算法框架。该框架引入了多维状态感知与群体共识决策机制，实现 Plateaued 函数构造过程中交叉与变异策略的智能动态选择，从而在解质量、收敛速度与稳定性方面取得显著提升。

3 预备知识

3.1 Plateaued 函数

设 $\mathbb{F}_2=\{0,1\}$为二元有限域(其中元素的加法为异或运算⊕，乘法为逻辑与运算 ·)，$\mathbb{F}_2^n$表示由所有长度为 n 的二元向量构成的 n 维向量空间。若函数满足 $f:\mathbb{F}_2^n\to\mathbb{F}_2$，则称 f 为 n 元布尔函数。其中，$\mathbb{F}_2^n$中向量 $\boldsymbol{x}=(x_1,x_2,\cdots,x_n)$ 称为函数的输入；$f(\boldsymbol{x})\in\{0,1\}$ 称为函数在输入 $\boldsymbol{x}$ 处的输出值。任一 n 元布尔函数 f 均可唯一地表示为如下形式的 $\mathbb{F}_2$上多项式：

$$f(x_1,\cdots,x_n)=\bigoplus_{I\subseteq\{1,\cdots,n\}}a_I\prod_{i\in I}x_i \tag{1}$$

此表示称为函数 f 的代数正规型(ANF)。布尔函数可以通过多种方式表示。虽然从数学角度，这些形式具有等价性，但每种方式在实际应用中都存在各自的优缺点[25]。

真值表是有限域$\mathbb{F}_2^n$上布尔函数 $f:\mathbb{F}_2^n\to\mathbb{F}_2$的标准表示形式，其本质是通过有序排列函数所有输入对应的输出值，形成的结构化二元序列。对于定义域$\mathbb{F}_2^n$，其全部2^n个输入向量可按固定顺序枚举为$\boldsymbol{x}_0,\boldsymbol{x}_1,\cdots,\boldsymbol{x}_{2^n-1}$，其中全 0 真值序列$\boldsymbol{x}_0=(0,0,\cdots,0)$，全 1 值序列$\boldsymbol{x}_{2^n-1}=(1,1,\cdots,1)$，其余向量按二元递增规则排列。布尔函数 f 的真值表是长度为2^n的二元序列。$TruthTable(f)=(f(\boldsymbol{x}_0),f(\boldsymbol{x}_1),\cdots,f(\boldsymbol{x}_{2^n-1}))$ (2)

该序列中每个元素对应输入$\boldsymbol{x}_i\in\mathbb{F}_2^n$经函数 f 映射后的输出真值序列。为深入分析布尔函数的密码学性质，引入其 Walsh 变换的概念。函数 f 在任意 $\boldsymbol{b}\in\mathbb{F}_2^n$处的 Walsh 变换定义为：

$$W_f(\boldsymbol{b})=\sum_{\boldsymbol{x}\in\mathbb{F}_2^n}(-1)^{f(x)+x\cdot b} \tag{3}$$

其中，$\boldsymbol{x}\cdot\boldsymbol{b}$ 表示向量 $\boldsymbol{x}$ 与 $\boldsymbol{b}$ 在 $\mathbb{F}_2$上的标准点积，谱半径为函数 f 的 Walsh 谱中所有谱值绝对值的最大值。

布尔函数 f 被称为 s-Plateaued 函数，当且仅当存在一个满足 $0\leqslant s\leqslant n$ 的固定整数 s，使得 f 的 Walsh 谱值集合$\{W_f(\boldsymbol{b})\mid\boldsymbol{b}\in\mathbb{F}_2^n\}$是集合$\{0,\pm2^{\frac{n+s}{2}}\}$的子集；特别地，当 $s=0$，n 为偶数时，该布尔函数 f 即为 bent 函数。

布尔函数的 Walsh 谱表是基于其真值表与 Walsh 变换定义形成的结构化表示，核心是按二元递增顺序枚举有限域$\mathbb{F}_2^n$的所有2^n个输入向量$\boldsymbol{b}_i$，$i\in\boldsymbol{Z}$ 且 $0\leqslant i\leqslant2^n-1$ 与真值表中输入向量 $\boldsymbol{x}$ 的排序规则一致，并对应列出每个 $\boldsymbol{b}_i$ 对应的 Walsh 谱值$\mathrm{W_f}(\boldsymbol{b}_i)$如下：

$$Table(W_f)=(W_f(\boldsymbol{b}_0),W_f(\boldsymbol{b}_1),\cdots,W_f(\boldsymbol{b}_{2^n-1})) \tag{4}$$

3.2 遗传算法

遗传算法(Genetic Algorithm)是一类模拟生物自然选择与遗传变异机制的随机化智能优化算法，其核心思想源于达尔文进化论的“适者生存”原则，通过对“种群”中“个体”的迭

代优化，逐步逼近复杂问题的最优解。

将布尔函数的真值表转换为二元序列，并以该序列作为遗传算法中的个体 $\boldsymbol{p}$，以支持种群迭代运算。在此基础上，个体的生成和更新分两阶段进行。1）初始化阶段，通过调用 initialize_population 方法随机生成 $|\boldsymbol{P}|$ 个二元序列，以保证初始种群的多样性，避免过早陷入局部最优。此阶段无先验信息引导。2）进化阶段，新个体序列并非随机选择，而是通过遗传操作定向演化而来。先经锦标赛选择法筛选高适应度父代，再通过 DQN 指导的动作进行交叉变异，以当前所选择的交叉动作 $a_{c,t}$ 对选定的父代个体对（$\boldsymbol{p}_1$，$\boldsymbol{p}_2$）实施基因重组，采用单点交叉，即选取交叉点 $c\in[1,N-1]$，进行拼接，生成新子代 $\boldsymbol{p}_{\text{child}}$。然后是变异操作，即以变异概率 μ_t 对子代个体的每个基因位进行随机翻转（0→1 或 1→0），以此引入随机扰动，维持种群多样性并避免早熟收敛。最后是种群更新，即用生成的新子代个体替换当前种群中的部分或全部个体，保留每代中适应度最高的若干个体直接进入下一代，形成新一代种群 $\boldsymbol{P}_{t+1}$。循环上述过程，直至满足预设的终止条件。

3.3 Q-learning

Q-learning[26] 是基于值迭代的无模型强化学习算法，其核心目标是在未知环境中通过与环境的交互学习最优动作策略。该算法通过量化“状态-动作对”的长期价值，优化复杂决策过程，无需预先知晓环境的状态转移概率模型。

Q-learning 通过迭代修正 Q 函数逼近最优值，其核心为：

$$\boldsymbol{Q}(s_t,a_t)\leftarrow\boldsymbol{Q}(s_t,a_t)+\alpha[r_{t+1}+\gamma\max_a\boldsymbol{Q}(s_{t+1},a)-\boldsymbol{Q}(s_t,a_t)] \tag{5}$$

其中，t 为当前迭代数，s_t 为当前状态，a_t 为当前动作，r_{t+1} 为执行 a_t 后的即时奖励，s_{t+1} 为转移后的下一状态；$\alpha\in(0,1]$ 为学习率，用来控制每次更新的幅度，值越大收敛越快，但可能振荡，值越小越稳定，但收敛慢；折扣因子 γ 用于衡量对“未来奖励”的重视程度，更关注长期收益，以规避短视调整；$\max_a\boldsymbol{Q}(s_{t+1},a)$ 表示基于当前 Q 函数估计的下一状态的最优动作价值。

4 理论框架

本章将详细阐述所提模型架构，并说明具体构建方法。核心问题是通过“DQN 增强的遗传算法”求解基于二元特征选择的组合优化问题。从给定特征矩阵中筛选最优子集，由二元序列指示选择，使该子集经 Walsh 函数相关计算后获得最高适应度；算法以遗传算法为基础框架，通过种群初始化、选择、单点交叉、变异操作实现种群迭代进化，同时引入 DQN 强化学习代理，其通过接收当前状态，采用 ε-贪婪策略自适应选择多种变异强度，平衡种群探索新特征与利用优良特征的能力，最终提升遗传算法的优化效率与收敛效果。

4.1 初始种群的序列转换构造

在二元域框架下，布尔函数首先通过其真值表被表示为二元序列作为个体 $\boldsymbol{p}$，以支持种群的迭代运算。序列中每个元素取值为 0 或 1，分别对应布尔函数在对应输入向量 $\boldsymbol{x}_i$（$\boldsymbol{x}_i\in\mathbb{F}_2^n$，$i\in\{0,1,\cdots,2^n-1\}$）上的输出值；接着，将生成的二元序列个体 $\boldsymbol{p}$ 调用 Walsh 变换，计算该函数的完整 Walsh 谱值集合，记为 Table($\boldsymbol{W}_p$)，以目标 Plateaued 函数的谱值约束为筛选标准，为后续的选择、交叉、变异等进化操作提供基础（之后的函数 f 对应的真值序列即为个体 $\boldsymbol{p}$，进行遗传操作）。

在此基础上，个体的生成和更新分两阶段进行。1）初始化阶段，通过调用 initialize_population 方法随机生成 $|\boldsymbol{P}|$ 个二元序列，以保证初始种群的多样性，避免过早陷入局部最优。此阶段无先验信息引导。2）进化阶段，新个体序列并非随机选择，而是通过遗传操作定向演化而来，先经锦标赛选择法筛选高适应度父代，再通过单点交叉重组父代基因，最后以 DQN 指导的可控变异微调基因，核心是保留优良特征并引入适度探索性变异。

4.2 适应度函数

为提升 Plateaued 函数的组合优化效率与求解精度，须设定如下目标函数，以引导 DQN 增强的遗传算法通过自适应选择“交叉类型 + 变异强度”的组合策略，高效搜索最优特征子集，最终实现函数值的优化。

由于非线性度与谱半径存在直接量化关系，因此有：

$$nl(\boldsymbol{p})=2^{n-1}-\max(Table(\boldsymbol{W}_p)) \tag{6}$$

其中，$Table(\boldsymbol{W}_p)$ 表示真值序列 $\boldsymbol{p}$ 的所有 Walsh 值，$\max(Table(\boldsymbol{W}_p))$ 即为 Walsh 谱的谱半径。谱半径越小，非线性度越高，函数抵御线性攻击的能力也越强。因此，为了构造高非线性度的函数，须最小化 Walsh 谱的最大绝对值。通过交叉与变异操作机制，不断地混合和微调候选真值序列的基因编码，以期在迭代过程中发现那些具有更小谱半径从而具备更高非线性度的优良个体。每个个体的适应度函数为：

$$fit(\boldsymbol{p})=\frac{1}{|count(Table(\boldsymbol{W}_p))-3|^2+\sum\limits_{w_i,w_j>0,w_i,w_j\in Table(\boldsymbol{W}_p)}|w_i-w_j|^2+0.1} \tag{7}$$

其中，$Table(\boldsymbol{W}_p)$ 表示真值序列个体 $\boldsymbol{p}$ 所对应的所有 Walsh 值，$count(Table(\boldsymbol{W}_p))$ 统计 $Table(\boldsymbol{W}_p)$ 中不重复项的数量。通过该函数，可计算得到初始种群中每个 Plateaued 函数对应序列的适应度值，形成初始适应度集合 $\boldsymbol{F}_0$。

为评估种群演化状态，须引入两项关键适应度指标。

1）种群平均适应度 $\overleftarrow{fit}_t$：第 t 代种群 $\boldsymbol{P}_t$ 中所有个体适应度的算术平均值，当前种群大小是 $|\boldsymbol{P}_t|$ 个。

$$\overleftarrow{fit}_t=\frac{1}{|\boldsymbol{P}_t|}\sum_{\boldsymbol{p}\in\boldsymbol{P}_t}fit(\boldsymbol{p}) \tag{8}$$

其值存储于 avg_fitness_history 数组，用于反映种群的整体适应水平。

2）种群最佳适应度 $fit_{\max,t}$：第 t 代种群中个体适应度的最大值。

$$fit_{\max,t}=\max_{\boldsymbol{p}\in\boldsymbol{P}_t}fit(\boldsymbol{p}) \tag{9}$$

其值存储于 best_fitness_history 数组，用于追踪算法在迭代过程中的寻优能力演进。此外，为量化种群基因型差异，定义种群多样性指标 d_t：

$$d_t=\frac{2}{|\boldsymbol{P}_t|(|\boldsymbol{P}_t|-1)}\sum_{i<j}Hamming(\boldsymbol{P}_t[i],\boldsymbol{P}_t[j])/2^n \tag{10}$$

基于汉明距离来度量两个不同真值序列个体的差异，其

中 $Hamming(\boldsymbol{P}_t[i],\boldsymbol{P}_t[j])$ 表示第 t 代种群中第 i 个个体与第 j 个个体之间的汉明距离，即两个不同个体的染色体编码中对应位置取值不同的数量。

4.3 DQN 框架

深度 Q 网络[27-28]作为 Q-learning 的深度强化学习延伸，用神经网络替代传统 Q 表来逼近动作价值函数(Q 值函数)，使其能够有效处理高维状态空间；并借助经验回放机制与目标网络设计，使得学习过程的稳定性与效率得到显著增强。

在针对 Plateaued 函数的组合优化算法中，以马尔可夫决策过程为理论框架，通过 DQN 动态调控遗传操作中的交叉与变异策略——利用强化学习中基于状态感知与奖励信号选择最优动作的机制，实现对搜索过程的自适应引导，最终提升 Plateaued 函数组合优化的效率与精度。

下面构建的状态空间 $\boldsymbol{s}$ 是一个多维特征向量，旨在为深度 Q 网络提供全面的进化环境感知能力。该状态空间[29]包含 5 个核心维度。

1)基因序列采样，从当前个体 $\boldsymbol{p}\in\boldsymbol{P}_t$ 中提取局部基因片段，提供微观结构信息；

2)相对谱值适应度 $\frac{fit_p}{fit_{\max,t}+\delta}$，反映个体在种群中的竞争地位；

3)进化进程 $\frac{t}{T}$，标记搜索过程的相对阶段；

4)种群收敛度 $\overleftarrow{fit}_t$，表征种群的整体适应度水平；

5)基因多样性 d_t，度量解空间探索的广度。

动作空间 $\boldsymbol{A}$ 定义了在每个状态下可以执行的所有可能动作，其中变异幅度集合定义为 $a_m\in A_m=\{0.5,1.0,1.5\}$，0.5,1.0,1.5 分别对应轻度、标准和强化 3 种变异强度；交叉操作集合定义为 $a_c\in A_c=\{0,1\}$，其中 0 表示不执行交叉操作，1 表示执行单点交叉操作。

动作选择策略则基于学习到的 Q 值进行智能决策。特别地，通过 $\boldsymbol{A}=A_m\times A_c$ 的组合构建 6 种完整的遗传操作策略 $\boldsymbol{a}_j\in\boldsymbol{A}$。

1)动作 $\boldsymbol{a}_0$：$a_c=0$，$a_m=0.5$，即不交叉配合轻度变异；

2)动作 $\boldsymbol{a}_1$：$a_c=0$，$a_m=1.0$，即不交叉配合标准变异；

3)动作 $\boldsymbol{a}_2$：$a_c=0$，$a_m=1.5$，即不交叉配合强化变异；

4)动作 $\boldsymbol{a}_3$：$a_c=1$，$a_m=0.5$，即单点交叉配合轻度变异；

5)动作 $\boldsymbol{a}_4$：$a_c=1$，$a_m=1.0$，即单点交叉配合标准变异；

6)动作 $\boldsymbol{a}_5$：$a_c=1$，$a_m=1.5$，即单点交叉配合强化变异。

DQN 作为“决策器”，为全连接三层神经网络，用于学习选择“交叉类型+变异强度”的组合策略，以最大化适应度提升。调整后的 DQN 类仍为强化学习的“决策器”，用于学习选择“交叉类型+变异强度”的组合策略，以最大化适应度提升，其结构为全连接三层神经网络。输入层为当前状态向量。隐藏层包含两层，均使用 ReLU 激活函数，分别实现非线性特征提取和高阶特征提炼。输出层维度对应上述 6 种组合动作，直接输出 6 种动作的 Q 值，即每种组合策略的预期累积奖励，为动作选择提供依据，使 DQN 能自适应学习最优的交叉与变异组合策略。

在决策机制上，采用基于群体共识的 Q 值平均策略[30]：从当前种群 $\boldsymbol{P}_t$ 中随机选取 $\lfloor|\boldsymbol{P}_t|/2\rfloor$ 个个体，分别计算其状态向量对应的 Q 值分布，通过对这些分布逐元素求平均得到群体的 Q 值期望，最终执行操作 arg max，确定最优动作。这种设计实现了从个体决策到群体智能的升华，使算法能够在考虑全局进化状态的前提下，选择最具潜力的遗传操作策略。随机选取的子集记为 $\boldsymbol{P}_t^{\text{sub}}$，满足 $|\boldsymbol{P}_t^{\text{sub}}|=\lfloor|\boldsymbol{P}_t|/2\rfloor$。对于子集中每个个体 $\boldsymbol{p}_i\in\boldsymbol{P}_t^{\text{sub}}$，其状态向量为 $\boldsymbol{s}_{p_i}$，通过深度 Q 网络得到对应的 Q 值向量：

$$\boldsymbol{Q}(\boldsymbol{s}_{p_i})=[\boldsymbol{Q}(\boldsymbol{s}_{p_i},\boldsymbol{a}_1),\boldsymbol{Q}(\boldsymbol{s}_{p_i},\boldsymbol{a}_2),\cdots,\boldsymbol{Q}(\boldsymbol{s}_{p_i},\boldsymbol{a}_6)] \tag{11}$$

计算群体单个动作 $\boldsymbol{a}_j\in\boldsymbol{A}$ 的平均 Q 值：

$$\overleftarrow{Q}_{s_t,a_j}=\frac{1}{|\boldsymbol{P}_t^{\text{sub}}|}\sum_{\boldsymbol{p}_i\in\boldsymbol{P}_t^{\text{sub}}}Q(s_{p_i},\boldsymbol{a}_j) \tag{12}$$

基于当前种群得到 $\boldsymbol{Q}(\boldsymbol{s}_t,\boldsymbol{a})=[\overleftarrow{Q}_{s_t,a_1},\overleftarrow{Q}_{s_t,a_2},\cdots,\overleftarrow{Q}_{s_t,a_6}]$，每个分量 $\overleftarrow{Q}_{s_t,a_j}$ 表示动作对 $\boldsymbol{a}_j$ 在当前群体层面的期望累积奖励。选择其中的最大值作为最优动作，这一决策机制可表示为：

$$\pi(\boldsymbol{P}_t)=\arg\max_{\boldsymbol{a}_j\in\boldsymbol{A}}\left[\frac{1}{\lfloor|\boldsymbol{P}_t|/2\rfloor}\sum_{\boldsymbol{p}_i\in\boldsymbol{P}_t^{\text{sub}}}\boldsymbol{Q}(\boldsymbol{s}_{p_i},\boldsymbol{a}_j)\right] \tag{13}$$

该策略 π 依托群体共识机制，筛选出在当前进化环境下对群体中多数个体具有最优潜力的遗传操作，实现了从个体层面优化到群体协同进化的范式跃迁。通过对状态 $\boldsymbol{s}_t$ 下的动作价值函数 $\boldsymbol{Q}(\boldsymbol{s}_t,\boldsymbol{a})$ 取最大值，可得到当前状态-动作对的价值估计 $\boldsymbol{Q}(\boldsymbol{s}_t,\boldsymbol{a}_t,\boldsymbol{s}_{t+1})$；基于该最大值对应的动作完成决策选择，记为 $\boldsymbol{a}_t$，并通过执行该动作得到下一状态 $\boldsymbol{s}_{t+1}$。

动作选择由 act 方法完成，采用 ε-贪婪策略。以概率 ε 随机选择动作以探索新的交叉-变异组合，以概率 $1-\varepsilon$ 基于当前 Q 网络选择最优动作 $\max\boldsymbol{Q}(\boldsymbol{s}_t,\boldsymbol{a})$。此外，通过经验存储 remember 方法，将“状态-动作对-奖励-下一状态”的样本存入经验池(最大容量为 2000)，为后续网络训练提供数据支撑。

4.4 网络更新

DQN 采用结构相同的评估网络 θ_t 和目标网络 θ^{TD} 双网络架构，其核心设计目的是通过稳定学习目标来提升训练过程的稳定性和收敛速度。评估网络负责交互环境、选择动作并执行梯度更新；而目标网络则专门用于计算目标中的下一状态 Q 值，其参数定期从评估网络复制，而非实时更新。

这种解耦机制使得 TD 目标在短期内保持相对固定，有效避免了“追逐移动目标”导致的价值估计振荡与发散，如同为优化过程提供了一个稳定的参考锚点，从而确保了学习过程的平稳与高效。

利用时序差分(TD)学习思想来指导 DQN 网络学习，目标 Q 值表示为：

$$y_t=r_t+\gamma\cdot\max_{\boldsymbol{a}_{t+1}\in\boldsymbol{A}}\boldsymbol{Q}(\boldsymbol{s}_{t+1},\boldsymbol{a}_{t+1};\theta^{\text{TD}}) \tag{14}$$

其中，r_t 为当前动作的即时奖励，γ 为折扣因子(用于权衡即时奖励与未来奖励的权重)，θ^{TD} 表示目标网络的参数集合；$\max_{\boldsymbol{a}_{t+1}\in\boldsymbol{A}}\boldsymbol{Q}(\boldsymbol{s}_{t+1},\boldsymbol{a}_{t+1};\theta^{\text{TD}})$ 为目标 Q 值，其通过从经验回放池中随机抽取小批量样本，并经目标网络基于策略 π 计算得到，反映下一状态下最优动作的预期累积奖励。损失函数采用均方误差(MSE)，其核心作用是量化并最小化当前 Q 值与上述目标 Q 值之间的偏差，以此引导当前 Q 网络的参数优化过程。

$$Loss=(\boldsymbol{Q}(\boldsymbol{s}_t,\boldsymbol{a}_t;\theta_t)-y_t)^2 \tag{15}$$

通过 Adam 优化器定期最小化该损失，进行参数更新，使 DQN 网络逐步掌握在何种序列状态下应采用何种交叉-变异组合，才能最有效地引导 Walsh 谱收敛至 Plateaued 特性，为

Plateaued 函数的高效构造提供了智能决策支持。

该深度 Q 网络通过学习方式（见图 1），实现了对遗传算法操作参数的自适应调控，能够根据进化状态智能选择最优的交叉-变异组合策略，为 Plateaued 函数的高效构造提供了可靠的决策支持。网络训练过程中定期更新目标网络参数，确保学习过程的稳定性，最终使算法能够有效引导种群向具有理想 Walsh 谱特性的方向进化。

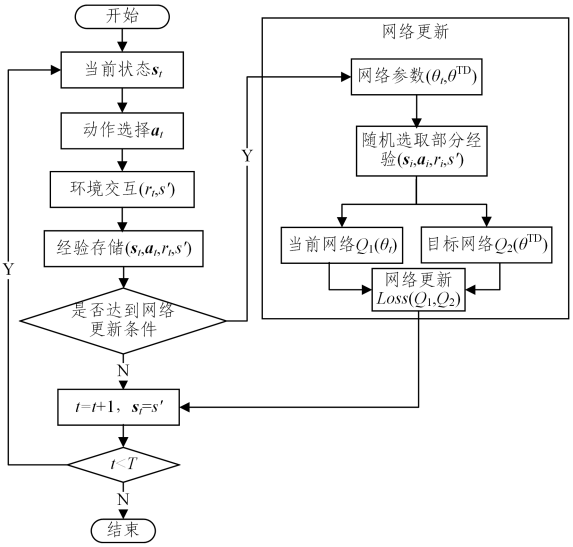


图 1 DQN 算法的训练流程

Fig. 1 Training flow of deep Q-Network algorithm

4.5 选择

选择策略的目标是从当前种群中筛选出最有可能产生 Plateaued 后代的潜力真值序列，其核心机制为锦标赛选择：每次从种群中随机选取 3 个个体并选择 Walsh 谱适应度最高者，重复此过程，直至选出 R 个优质父本（记为 *best_parents*）及其对应适应度集合 *best_parents_fitness*。

新产生的适应度集合的高低直接反映了 $\boldsymbol{p}$ 所得到的 Walsh 谱接近理想三值分布的程度。因此，选择操作通过这种方式施加进化压力，优先保留那些选择模式能产生更优 Walsh 谱特性的序列。例如，若某真值序列个体 $\boldsymbol{p}$ 能够使得其 Walsh 谱中非零值数量减少且幅度趋于目标，其适应度会更高，从而更易被选为父本。同时，锦标赛选择在保证精英导向的基础上，通过随机抽样维持了种群多样性，可有效避免算法过早收敛到局部最优。

4.6 交叉

为定向优化谱值这一核心特性，交叉操作采用 DQN 策略控制机制，仅当输出动作 $a_{c,t}=1$ 时执行。对选定的父本 $\boldsymbol{p}_1$ 和 $\boldsymbol{p}_2$ 进行配对，先通过分析二者的 Walsh 谱 $Table(W_{\boldsymbol{p}_1})$ 和 $Table(W_{\boldsymbol{p}_2})$ 定位高谱半径对应的关键真值序列个体基因位置，即对高谱值贡献显著的真值序列个体基因进行索引，再从这些位置中随机选取候选交叉点 $c\in[1,2^n-1]$。通过序列拼接生成子代个体 $\boldsymbol{p}_{\text{child}}=\boldsymbol{p}_1[1:c]+\boldsymbol{p}_2[c+1:2^n]$。随后生成的子代序列 *child* 中，前 c 个真值序列个体基因继承自 $\boldsymbol{p}_1$，后 2^n-c 个基因则来自 $\boldsymbol{p}_2$，即 $\boldsymbol{p}_{\text{child}}=(\boldsymbol{p}_1^1,\boldsymbol{p}_1^2,\cdots,\boldsymbol{p}_1^c,\boldsymbol{p}_2^{c+1},\cdots,\boldsymbol{p}_2^{2^n})$。

交叉本质上是对父代选择向量所定义的列集合进行混合。若父代 $\boldsymbol{p}_1$ 对应列集 $parent_1$，父代 $\boldsymbol{p}_2$ 对应列集 $parent_2$，交叉可产生子代对应的集合，这相当于探索了全新的真值序列 $\boldsymbol{p}_{\text{child}}$，其 Walsh 谱特性可能与两个父代均存在差异。作为算法在选择空间中开展全局探索的核心步骤，交叉操作为发现新的真值序列个体结构提供了重要支撑。

4.7 变异

变异操作通过动态变异率 μ_t 对个体 $\boldsymbol{p}$ 实施真值序列个体基因位翻转突变。若已执行交叉，则针对子代个体实施真值序列位翻转突变；若未执行交叉，则直接对父本实施真值序列位翻转突变。

$$\boldsymbol{p}[i]=\begin{cases}1-\boldsymbol{p}[i], & random(0,1)\leqslant\mu_t\\ \boldsymbol{p}[i], & \text{其他}\end{cases}\tag{16}$$

该操作通过随机扰动序列中的位来微调选择模式，或帮助算法跳出局部最优，而 DQN 的引入使其具备自适应特性。其核心机制是在传统位翻转变异上，以概率 μ_t 改变序列 $\boldsymbol{p}$ 中某一位的值，由 DQN 动态调整变异率；先通过分析父本个体的 Walsh 谱，定位高谱半径分量对应的基因索引，即对高谱值贡献显著的基因位置，将其作为候选变异区域；随后，根据 DQN 输出的动作空间 $a_{m,t}\in\{0.5,1.0,1.5\}$ 动态调整基础变异率，且对上述关键基因位置优先执行变异操作；若候选基因位置已覆盖种群的多数基因，则在全序列中随机选择变异位点，但关键基因位置的变异概率仍提升 50%，确保对谱特性的定向优化；最终生成子代个体。

不同变异强度各有侧重：轻度变异适用于序列 $\boldsymbol{p}$ 已具备较好 Plateaued 特性的场景，通过微小调整在优良解邻域进行局部搜索，以优化精度；重度变异则在序列适应度低或算法早期探索阶段触发，通过大幅修改选择模式跳出局部最优，转向更具潜力的新方向；标准变异则作为平衡策略，兼顾探索与利用的需求。经 DQN 增强，智能地调整对选择模式的扰动强度，实现了对优良模式的精细局部开发与对平庸模式的激进全局探索之间的自适应平衡。

4.8 奖励计算

在奖励计算阶段，通过采用多目标综合评价体系[31]量化当前动作 $(a_{m,t},a_{c,t})$ 的优化效果。构建即时奖励函数：

$$r_t=10\cdot\Delta fit+5\cdot d_t\tag{17}$$

该函数集成多个评估维度，Δfit 为当前种群的整体适应度提升项。

$$\Delta fit=\overleftarrow{fit}_t-\overleftarrow{fit}_{t-1}\tag{18}$$

该函数针对 Plateaued 函数的特殊谱特性需求进行优化设计，其中 Δfit 表征种群 Walsh 谱向理想三值分布的逼近程度。而 d_t 项则反映了当前种群的基因多样性水平。这种双目标奖励机制通过平衡谱特性优化与解空间探索，有效引导 DQN 智能体学习在何种进化状态下应采用何种交叉-变异组合策略，从而高效构造出具有理想密码学特性的 Plateaued 函数真值序列。

4.9 迭代过程

该遗传算法的迭代过程通过深度 Q 网络实现自适应调控：在每一代中，首先计算种群平均适应度、最佳适应度和基因多样性三项核心指标，以量化搜索状态。随后，为做出稳健的群体级决策，从种群中随机抽取半数个体，基于其个体状态与全局指标构建多维状态向量，输入 DQN 网络分别评估 6 种“交叉-变异”组合动作的 Q 值，并通过求平均获得群体共识的期望 Q 值。采用 ε-贪婪策略，基于此共识选择本代的最

优动作。决策既定，便依次执行锦标赛选择优质父本；决定是否执行基于 Walsh 谱指导的单点交叉；再以动态变异率实施变异操作。新子代经适应度评估后，通过精英保留策略更新种群。最后，算法根据适应度提升与多样性计算即时奖励，并将本代经验存入回放池，定期采样训练 DQN 网络，优化其决策能力，从而智能引导种群向具有理想 Walsh 谱特性的 Plateaued 函数进化。上述完整流程可形式化地表述为算法 1。

算法 1 基于 DQN 调控的自适应遗传算法

输入：$(T, t, \mathbf{P}_t, \text{Table}(W_p), \mu_{base}, \mathbf{F}, \mathbf{S}, \mathbf{A}=A_m \times A_c, r_t, \alpha, \gamma, \varepsilon, \theta, C, \mathbf{M})$

输出：(result，图像可视化)

初始化：$t=1, \mathbf{P}_1, \mathbf{F}_1; \mathbf{Q}(;\theta); \mathbf{M}$

for t←1 to T do:

  $\overleftarrow{fit}_t, fit_{max,t}, d_t \leftarrow (\mathbf{F}_t, \mathbf{P}_t)$

  for i←1 to $\lfloor |P_t|/2 \rfloor$ do:

    $\mathbf{s}_{p_i} = (\mathbf{p}_i, \overleftarrow{fit}_t, fit_{max,t}, t, d_t)$，动作 $\mathbf{A}$

    $\mathbf{Q}(\mathbf{s}_{p_i}) = [Q(\mathbf{s}_{p_i}, \mathbf{a}_1), \cdots, \mathbf{Q}(\mathbf{s}_{p_i}, \mathbf{a}_6)]$

  end for

  $\overleftarrow{Q}_{s_i, a_j} \leftarrow \text{Average}(\mathbf{Q}(\mathbf{s}_{p_i}), \lfloor |P_t|/2 \rfloor)$

  $\mathbf{a}_t \leftarrow (\overleftarrow{Q}_{s_i, a_j}, \text{random}, \varepsilon)$

  parents←锦标赛选择$(\mathbf{P}_t, \mathbf{F}_t, 3)$

  child←交叉(parents, $a_{c,t}$, Table(W_p))

  $\mu_t \leftarrow \mu_{base} \times a_{m,t} \times (0.995)^t$

  child←变异(child, μ_t, Table(W_p))

  $\mathbf{P}_{t+1}$←精英保留$(\mathbf{P}_t, \text{child}, \text{Table}(W_{\mathbf{P}_t}))$

  $\mathbf{F}_{t+1}$←计算适应度$(\mathbf{P}_{t+1})$

  result←筛选$(\mathbf{P}_{t+1}, \mathbf{F}_{t+1})$

  $r_t \leftarrow (\overleftarrow{fit}_{t+1}, \overleftarrow{fit}_t, d_t)$

  $\mathbf{s}_{t+1} \leftarrow (\mathbf{s}_t, \mathbf{a}_t, r_t)$

  $\mathbf{M} \leftarrow \mathbf{M} \cup \{(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1})\}$

  if t mod C==0 then:

    从 $\mathbf{M}$ 中采样小批量经验

    $y_i = r_i + \gamma \cdot \max_{\mathbf{a}_{i+1} \in \mathbf{A}} \mathbf{Q}(\mathbf{s}_{i+1}, \mathbf{a}_{i+1}; \theta^{TD})$

    $\text{Loss} = (\mathbf{Q}(\mathbf{s}_i, \mathbf{a}_i; \theta_t) - y_i)^2$

    Adam 更新网络$\theta_{t+1} \leftarrow \theta^{TD}$

  end if

  i←i+1

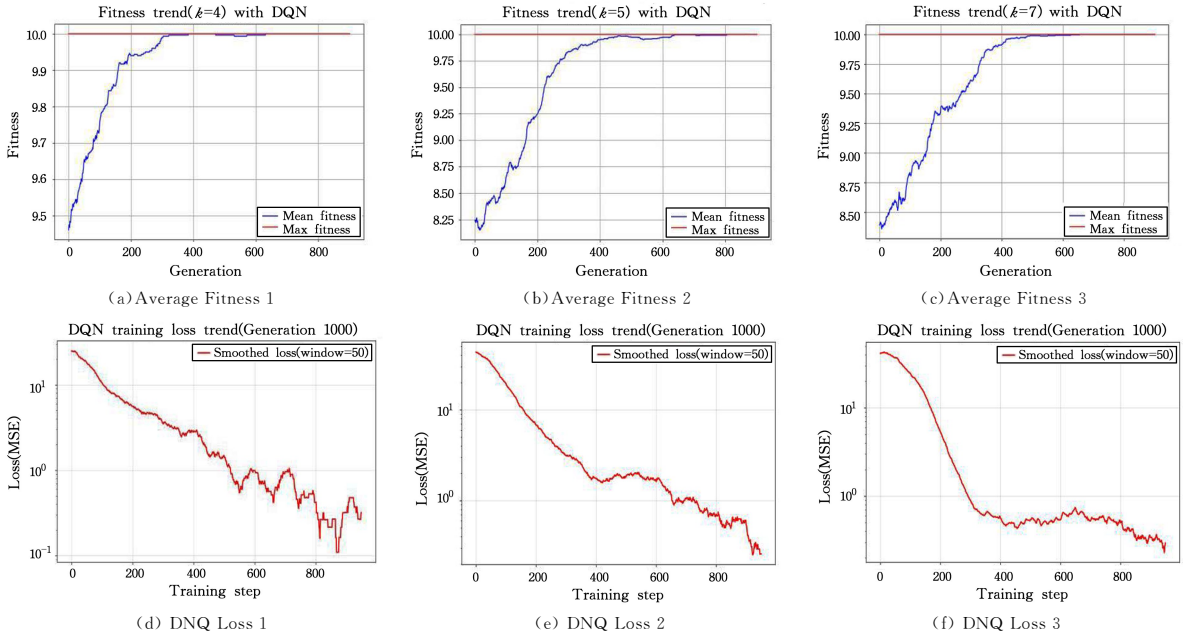
end for

5 实例分析

5.1 基于 DQN 的自适应遗传算法的性能分析

本实验采用 DQN 与遗传算法的混合架构，通过强化学习机制实现遗传操作策略的自适应优化。在遗传算法参数配置中，设置最大迭代次数 $T=1000$，以保障充分收敛；真值序列长度 $n=2^k$($k=4,5,7$)；种群规模 $|\boldsymbol{P}_t|=30$，平衡探索效率与计算复杂度；基础变异率 $\mu_{base}=0.248$，通过预实验确定最优值，并采用精英保留策略维持优良基因型。在 DQN 参数设计中，学习率 $\alpha=0.001$，确保梯度下降的稳定性；折扣因子 $\gamma=0.95$，权衡远期收益；探索率 $\varepsilon=0.1$，配合衰减机制，实现从探索到利用的平滑过渡；目标网络同步频率 $C=100$ 和经验回放缓冲区容量 $|\boldsymbol{M}|=2000$，以共同保障训练过程的稳定性。基于上述参数体系的实验数据表明，该混合算法在 Plateaued 函数的真值序列构造任务中展现出显著的优化效率与解质量提升。

依据上述参数得到了图 2 所示的数据(上侧是算法适应度进化趋势，下侧是 DQN 训练损失变化趋势，从左至右分别为 $k=4,5,7$ 的结果)。该算法通过 DQN 智能体对进化状态的连续感知，动态生成交叉算子选择与变异强度调节的联合决策，有效引导种群 Walsh 谱向理想三值分布特性收敛。



(a) Average Fitness 1　(b) Average Fitness 2　(c) Average Fitness 3

(d) DNQ Loss 1　(e) DNQ Loss 2　(f) DNQ Loss 3

图 2 基于 DQN 的自适应遗传算法的性能分析

Fig. 2 Performanceanalysis of adaptive genetic algorithm based on DQN

5.2 与基准算法的性能对比分析

采用 3 种算法进行对比研究：标准遗传算法(Standard GA)[32]、基础 Q-learning 遗传算法(Basic QL-GA)[33]和 DQN 遗传算法(DQN-GA)。参数设置：种群规模$|\boldsymbol{P}_t|=30$，染色体

编码长度 $n=2^k$（$k=5$），繁殖父本数量 $R=5$，基础变异率 $\mu_{base}=0.252$，最大迭代次数 $T=1\,000$，采用精英保留策略。对于 Basic QL-GA 算法，配置学习率 $\alpha=0.12$、折扣因子 $\gamma=0.91$、探索率 $\varepsilon=0.1$，DQN-GA 引入目标网络更新频率 $C=100$、经验回放缓冲区容量 $|\boldsymbol{M}|=2\,000$。依据上述参数实验，得到了如图 3 和表 1 所示的结果（3 种算法的最佳适应度重合）。

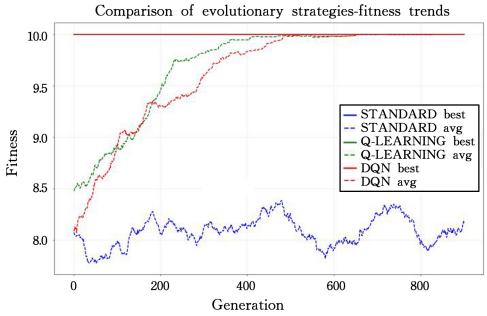


图 3 不同算法下平均适应度随迭代次数变化的对比（电子版为彩色）

Fig. 3 Comparison of the change in average fitness with the number of iterations under different algorithms

表 1 3 种算法结果的对比

Table 1 Comparison of results of three algorithms

性能指标	标准遗传算法	Q 学习遗传算法	DQN 遗传算法
适应度提升幅度	基准	+0.20～0.35	+0.15～0.40
收敛效果	不收敛	较快	较快
变异率调整范围	0.252～0.008（固定衰减）	0.238～0.268（动态调整）	0.235～0.276（智能调整）
交叉操作使用率/%	100（始终开启）	60～80（适时关闭）	70～90（智能开关）
时间消耗/h	0.85～1	1.1～1.45	1.3～1.6
有效解数量/个	100～150	160～220	230～300

表 1 中的指标多维度构建了算法性能的综合评估体系：适应度提升幅度直接量化算法优化能力，体现解的优劣程度相对于初始状态或基准算法的提升水平；收敛效果聚焦算法逼近最优解的速度与稳定性，评估是否能快速、平稳收敛至全局最优解，而非陷入局部最优；变异率调整范围反映算法对变异概率的自适应调控能力，其动态区间体现算法平衡探索与利用的水平；交叉操作使用率表征算法执行基因重组的频率，反映对优质基因传递与融合的依赖程度及自适应选择能力；时间消耗衡量算法完成一次完整优化过程的计算成本；有效解数量则体现算法输出满足问题要求的合格解的规模，直接反映算法的实用价值与寻优广度。六者协同，构成全面、严谨的算法性能评判标准。

基于图 3 和表 1 中的数据，在 Plateaued 函数真值序列构造任务的性能对比中，3 种进化算法展现出显著差异。标准遗传算法作为基准，其适应度提升有限且难以完全收敛，变异率采用固定衰减模式，始终开启交叉操作，虽时间消耗最短（0.85～1 h），但仅能找到少量有效解。基础 Q 学习遗传算法在适应度提升（+0.20～0.35）和收敛效果上有所改善，通过动态调整变异率和适时关闭交叉操作，增加了有效解数量，但也增加了时间消耗。DQN 遗传算法表现最优，在保持相同适应度提升幅度的同时，智能调整变异率和交叉操作，尽管时间消耗最多，但能找到 230～300 个有效 Plateaued 函数的真值序列，充分证明了深度强化学习在复杂密码学函数构造任务中的显著优势。

6 结论

系统对比所提 DQN-GA 算法与基准方法的综合性能，本文从算法内在特性与输出结果两个维度进行了总结。表 2 从参数调整机制、学习能力、状态感知方式、探索策略、收敛性能及有效解数量等方面，对标准遗传算法、基础 Q 学习遗传算法与 DQN 遗传算法进行了横向比较，清晰展示了 DQN-GA 在智能性、自适应性和稳定性方面的优势。表 3 则展示了 DQN-GA 在 $k=5$ 时输出的部分 Plateaued 函数真值序列示例，直观呈现了算法生成的有效解结构与规模，进一步验证了其在高维布尔函数构造任务中的实用性与高效性。

表 2 3 种算法特性的比较

Table 2 Comparison of characteristics of three algorithms

特性	标准遗传算法	Q 学习遗传算法	DQN 遗传算法
参数调整	固定	动态调整	智能动态调整
学习能力	无	基础 Q-learning	DQN
状态感知	无	简单状态划分	多维度状态评估
探索策略	随机变异	ε-greedy+Q 表	ε-greedy+Q 网络
收敛性能	依赖参数设置	自适应优化	稳定高效收敛
有效解数量/个	100～150	160～220	230～300

表 3 DQN-GA 算法输出真值序列示例

Table 3 Example of the output truth sequence of the DQN-GA algorithm

算法	DQN 遗传算法（$k=5$）
真值序列（部分）	[0,0,1,0,1,1,1,0,1,0,1,1,1,0,1,1,0,0,1,1,0,0,0,0,0,1,0,1,1,0,1,0]
	[0,0,1,0,1,0,0,1,1,0,0,0,1,1,0,0,1,1,1,0,1,0,1,0,1,1,1,0,0,1,0,1]
	[0,0,1,0,1,0,1,0,1,1,1,0,0,0,0,0,0,1,1,1,1,0,0,1,1,0,1,0,1,0,1,1]
	[0,1,0,0,0,0,0,0,0,1,0,1,1,0,1,0,1,1,0,1,0,1,1,1,0,1,1,0,0,1,1,0,1]
	[0,1,0,0,1,1,1,1,1,1,0,1,1,0,0,1,1,1,1,1,1,1,1,1,0,1,1,0,0,1,1,0,0,0]
	[0,0,0,1,1,1,0,0,1,1,1,1,1,0,0,0,1,1,0,1,0,1,1,0,1,1,0,1,1,1,1,1,1,1]
	[0,1,0,1,1,1,1,1,0,0,0,0,0,1,1,0,1,0,1,1,0,1,0,0,0,0,0,1,1,1,1,0,1,1]
	[0,0,1,1,1,0,1,1,0,1,1,1,1,0,0,1,1,0,1,1,1,1,1,1,0,0,1,1,1,1,1,0,0,0,1]
	……
	[0,0,1,0,1,1,1,0,1,1,1,1,0,1,0,1,1,1,0,0,1,1,1,1,1,0,0,1,0,0,1,1,1]
	[0,1,1,1,1,0,1,1,1,0,0,1,1,1,0,0,0,0,1,1,0,1,0,1,0,1,0,1,0,0,0,1,0]
	[0,0,0,0,1,0,0,0,0,0,1,1,0,0,1,1,1,1,1,1,1,0,1,0,1,0,1,0,1,1,1,1,0]
	[0,1,1,1,0,1,1,1,1,0,1,0,1,1,0,0,0,0,0,1,1,1,1,1,0,0,0,0,0,1,0,0,1]
	[0,1,0,0,1,0,1,0,1,1,0,1,0,0,0,0,0,1,1,0,0,0,1,0,1,0,1,0,0,1,0,0,1]
	[0,1,1,1,1,0,0,1,1,0,0,0,0,1,0,1,0,1,1,1,0,1,1,0,0,0,0,0,1,1,1,0,0]
	[0,0,0,1,1,1,0,0,0,0,1,0,0,0,1,1,0,0,1,0,0,1,0,0,1,1,0,1,0,1,1,0,1]
	[0,1,1,1,0,1,1,0,1,1,0,0,0,1,0,0,1,0,0,0,0,0,1,1,0,0,1,0,1,0,1,1,0]
总计个数	276

在 Plateaued 函数构造中取得了突破性进展。该算法聚焦于基于二元特征选择的组合优化问题，目标是从给定矩阵中筛选出最优序列，通过 Walsh 函数相关准则定义适应度，最终找到满足特定数学特性的最优序列，实现 Plateaued 函数真值序列的高效组合优化。这不仅显著提升了遗传参数调整的统计可靠性，更通过学习机制实现了从探索到利用的平滑过渡，为密码函数构造提供了一种兼具理论严谨性与实践效能的新方法。

结束语 本文通过 DQN 增强遗传算法，并将其用于高效构造密码学中的 Plateaued 函数。该方法通过强化学习与进化计算的深度融合，实现了遗传参数的智能动态调控，在解质量、收敛速度和稳定性方面均显著优于传统算法。然而，本文仍存在网络参数依赖经验、高维场景开销较大等问题，未来将聚焦网络结构自适应优化，探索并行计算在高维场景的应用，拓展算法的多目标优化能力，以进一步提升算法的实用性与泛化能力。

参 考 文 献

[1] SHANNON C E. A mathematical theory of communication[J]. The Bell System Technical Journal, 1948, 27(3): 379-423.

[2] HAMMING R W. Error detecting and error correcting codes [J]. The Bell System Technical Journal, 1950, 29(2): 147-160.

[3] BOSE R C, RAY-CHAUDHURI D K. On a class of error correcting binary group codes[J]. Information and Control, 1960, 3(1): 68-79.

[4] MACWILLIAMS F J, SLOANE N J A. The theory of error-correcting codes[M]. Elsevier, 1977.

[5] CARLET C. Partially-bent functions [J]. Designs, Codes and Cryptography, 1993, 3(2): 135-145.

[6] ZHENG Y, ZHANG X M. Plateaued functions[C]// International Conference on Information and Communications Security. Berlin, Springer, 1999: 284-300.

[7] HYUN J Y, LEE J, LEE Y. Explicit criteria for construction of plateaued functions[J]. IEEE Transactions on Information Theory, 2016, 62(12): 7555-7565.

[8] STANKOVIĆ M, MORAGA C, STANKOVIĆ R. Construction of ternary plateaued functions from quadratic forms for ternary bent functions[C]// 2021 IEEE 51st International Symposium on Multiple-Valued Logic(ISMVL). IEEE, 2021: 1-6.

[9] SUN T F, HU B, YANG Y. Research on the Construction of Plateaued Functions[J]. Journal of Electronics & Information Technology, 2018, 40(10): 2352-2357.

[10] CHEN R, YANG B, LI S, et al. A self-learning genetic algorithm based on reinforcement learning for flexible job-shop scheduling problem[J]. Computers & Industrial Engineering, 2020, 149: 106778.

[11] CHENG L, TANG Q, ZHANG L, et al. Scheduling flexible manufacturing cell with no-idle flow-lines and job-shop via Q-learning-based genetic algorithm[J]. Computers & Industrial Engineering, 2022, 169: 108293.

[12] YANG Q. Research on Adaptive Scheduling of Automated Warehousing System Based on Q-learning[D]. Hangzhou: Hangzhou Dianzi University, 2025.

[13] MIRALLES-PECHUÁN L, JIMÉNEZ F, PONCE H, et al. A methodology based on deep q-learning/genetic algorithms for optimizing covid-19 pandemic government actions[C]// Proceedings of the 29th ACM International Conference on Information & Knowledge Management. 2020: 1135-1144.

[14] KUKKER A, SHARMA R. A genetic algorithm assisted fuzzy Q-learning epileptic seizure classifier[J]. Computers & Electrical Engineering, 2021, 92: 107154.

[15] LIU F, ZENG G. Study of genetic algorithm with reinforcement learning to solve the TSP[J]. Expert Systems with Applications, 2009, 36(3): 6995-7001.

[16] ZHANG Z Y, WANG L, CAI J C, et al. Application research of improved Q-learning genetic algorithm in path planning[J]. CAAI Transactions on Intelligent Systems, 2025, 20(6): 1493-1504.

[17] ASTHANA R, VERMA N, RATAN R. Generation of Boolean functions using Genetic Algorithm for cryptographic applications [C]// 2014 IEEE International Advance Computing Conference (IACC). IEEE, 2014: 1361-1366.

[18] CARLET C. On the confusion and diffusion properties of Maiorana-McFarland's and extended Maiorana-McFarland's functions[J]. Journal of Complexity, 2004, 20(2/3): 182-204.

[19] HODŽIĆ S, PASALIC E, WEI Y, et al. Designing plateaued Boolean functions in spectral domain and their classification[J]. IEEE Transactions on Information Theory, 2019, 65(9): 5865-5879.

[20] HODŽIĆ S, PASALIC E, WEI Y. A general framework for secondary constructions of bent and plateaued functions[J]. Designs, Codes and Cryptography, 2020, 88(10): 2007-2035.

[21] MARIOT L, LEPORATI A. A genetic algorithm for evolving plateaued cryptographic boolean functions[C]// International Conference on Theory and Practice of Natural Computing. Cham: Springer, 2015: 33-45.

[22] PICEK S, CARLET C, GUILLEY S, et al. Evolutionary algorithms for boolean functions in diverse domains of cryptography [J]. Evolutionary Computation, 2016, 24(4): 667-694.

[23] JEONG J, LEE Y. Algorithms for constructing balanced plateaued functions with maximal algebraic degrees [J]. IEEE Transactions on Information Theory, 2023, 70(2): 1408-1421.

[24] BEHERA P K, GANGOPADHYAY S. An improved hybrid genetic algorithm to construct balanced Boolean function with optimal cryptographic properties [J]. Evolutionary Intelligence, 2022, 15(1): 639-653.

[25] PICEK S, MARCHIORI E, BATINA L, et al. Combining evolutionary computation and algebraic constructions to find cryptography-relevant boolean functions [C]// International Conference on Parallel Problem Solving from Nature. Cham: Springer, 2014: 822-831.

[26] CLIFTON J, LABER E. Q-learning: Theory and applications [J]. Annual Review of Statistics and Its Application, 2020, 7(1): 279-301.

[27] FAN J, WANG Z, XIE Y, et al. A theoretical analysis of deep Q-

learning[C]// Learning for Dynamics and Control. PMLR，2020：486-489.

[28] GU S，LILLICRAP T，SUTSKEVER I，et al. Continuous deep q-learning with model-based acceleration [C] // International Conference on Machine Learning. PMLR，2016：2829-2838.

[29] WU Y N，ZHANG J. An Improved NSGA-II Algorithm for Bullet Distribution [J]. Ship Electronic Engineering，2023，43(4)：29-33.

[30] PANDA S K，LIU R，XIANG Y. Asymptotic Analysis of Sample-averaged Q-learning[J]. IEEE Transactions on Information Theory，2025，71(7)：5601-5619.

[31] HE K，CHEN C，CHEN S，et al. Reinforcement Learning for Multi-Objective Optimization：A Review[J/OL]. Archives of Computational Methods in Engineering，2025：1-30. https://doi.org/10.1007/s11831-025-10389-3.

[32] MIRJALILI S. Evolutionary algorithms and neural networks [J]. Studies in Computational Intelligence，2019，780(1)：43-53.

[33] ZOJAJI Z，KAZEMI A. Adaptive reinforcement - based genetic algorithm for combinatorial optimization[J]. Journal of Computing and Security，2022，9(1)：71-84.

WU Yansheng，born in 1989，postdoctoral，associate professor. His main research interests include coding theory and cryptographic functions.

（责任编辑：柯颖）

CCF 高质量国际学术会议培育计划研讨会

2026 年 3 月 7 日，CCF 高质量国际学术会议培育计划研讨会在北京中国科技会堂举办。本次会议由 CCF 专委工委主任武成岗主持，CCF 理事长孙凝晖、秘书长唐卫清及各专委代表、相关工作人员齐聚一堂，围绕培育计划的核心目标与关键问题展开深入研讨，为推动我国国际学术会议高质量发展凝聚智慧、汇聚力量。

会议伊始，CCF 理事长孙凝晖发表讲话。他强调了发展自有国际学术会议的重要性。他表示，我国计算机领域基础研究当前正处在从量的积累迈向质的突破的关键阶段，要着力提升自有国际学术会议的质量。CCF 对培育计划中的会议给予合理支持，不搞拔苗助长，依托投稿量与学术影响力，稳步推动培育计划的会议成长为有国际影响力的会议，为科技强国提供有力支撑。

随后，武成岗汇报了培育计划会议的发展情况。他介绍，过去一年培育计划取得了一定进展，但部分会议在申请升级过程中，与 B 类会议在参会人数、论文数量、录取率等核心指标上仍存在明显差距，亟待针对性改进。

Internetware、APPT、JCC 等会议代表分享了各自在会议组织、国际化推进、论文质量把控等方面的成功经验，为其他会议提供了有益借鉴。各专委代表围绕年度目标与措施进行了详细汇报，结合自身会议特点，提出了优化组织架构、扩大宣传推广、加强国际合作等具体规划。

CCF 副秘书长王新霞就合规要求进行了专题宣贯，明确了财务收支、经费使用、法律风俗遵守等方面的具体规范，为会议的合规有序举办筑牢基础。

分组研讨与集中交流环节成为本次会议的亮点。与会人员分为两组，围绕“扩大会议规模”“吸纳高质量论文”“提升国际化参与度”及“机制建设”四大主题展开热烈讨论，提出了一系列富有建设性的思路与举措。例如，建立统一的培育计划宣传网站、与 A 刊捆绑推荐论文、合理选择海外办会地点、利用国际宣传渠道扩大影响力等。大家一致认为，评价指标应结合不同领域特点制定，避免"一刀切"；要平衡培育与淘汰机制，既要给予会议成长空间，也要形成良性竞争氛围。

CCF 秘书长唐卫清在总结讲话中强调，培育计划是 CCF 推动学术交流与创新的重要举措。他明确了不同类别会议的发展要求，鼓励各专委加强与亚太地区相关机构的合作，注重邀请国外知名人士实质性参与会议工作，合理选择办会地点以提升会议效果。他表示，CCF 将全力支持各专委的工作，推动各项研讨成果落地见效。

本次研讨会的成功举办，为各专委搭建了交流合作的平台，进一步明确了培育计划的发展方向与重点任务。未来，CCF 将持续聚力赋能，与各专委携手共进，不断提升我国国际学术会议的质量与影响力，为我国基础研究水平提升和科技强国建设贡献更大力量。

据 CCF 微信公众号