NOT PEER-REVIEWED

Commentary

Revealing temporal sequence patterns in constrained multiobjective optimization

Miao Rong, Yunjie Wang, Chen Peng*

School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China.

***Correspondence to:** Prof. Chen Peng, School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China. E-mail: c.peng@shu.edu.cn

Received: 13 May 2026 | Approved: 25 May 2026 | Online: 25 May 2026

INTRODUCTION

Constrained multi-objective optimization problems (CMOPs) are ubiquitous in real-world engineering applications, ranging from Internet of Things resource scheduling and autonomous vehicle path planning to analog circuit sizing and short-term crude oil scheduling. The central challenge in solving these problems lies in simultaneously balancing convergence, diversity, and feasibility—three mutually constraining objectives that must be carefully managed throughout the evolutionary process.

In recent years, constrained multi-objective evolutionary algorithms (CMOEAs) have made steady progress through adaptive designs of constraint-handling techniques (CHTs) and genetic operators[1,2]. However, most existing approaches overlook a key fact: the requirements for constraint-handling strategies change dynamically across generations. Early generations may demand strong exploration to traverse infeasible regions, whereas later generations require exploitation power to converge onto the constrained Pareto front.

DOI: 10.20517/scierxiv202605.0346.v1

In a recent study published in *IEEE Transactions on Evolutionary Computation*, Peng *et al.*[3] propose a deep reinforcement learning framework that conceptualizes the selections of CHTs and genetic operators as a temporal sequence, offering a principled way to handle this generation-varying adaptation problem.

CORE CONTRIBUTIONS

The most significant conceptual contribution of this work is the formalization of the "temporal sequence of constrained handling selection". By modeling each generation's choice of CHT and genetic operator as a discrete time step, the authors establish a unified framework in which most existing CMOEAs[2]—whether multiple-population, multi-stage, penalty-based, or learning-based—can be viewed as special cases with predetermined or degenerate selection sequences. This perspective elevates algorithm design from ad hoc engineering to a structured sequential decision problem.

To discover systematic patterns within this temporal sequence, the authors design a deep Q-network that learns from historical performance feedback. The state representation comprises seven carefully crafted features characterizing feasibility, convergence, and diversity of the current population. The action space consists of nine discrete combinations formed by pairing three representative CHTs (ICV, ε-method, and CDP) with three genetic operators (SBX, DE/rand/1/bin, and DE/rand/2/bin). A two-phase adaptive reward function dynamically shifts emphasis between exploration and exploitation as evolution progresses[3,4].

The experimental validation is thorough. The proposed CMOEA-TS is compared against nine peer algorithms on 37 benchmark instances from three test suites (MW, LIR-CMOP, and DAS-CMOP). Statistical results from Wilcoxon rank-sum tests and Friedman ranking consistently demonstrate CMOEA-TS's superiority[3]. Ablation studies further verify that: (i) learning systematic patterns from the sequence significantly outperforms random selections; (ii) comprehensive selection of both CHTs and genetic operators yields better adaptability than using either component in isolation[5]; and (iii) the proposed credit assignment function more effectively balances objective optimization and constraint satisfaction than simpler reward schemes.

DISCUSSION AND FUTURE DIRECTIONS

Despite its compelling results, several aspects warrant further investigation. The deep Q-network's training relies on accumulated experience samples in a replay buffer. During early generations when samples are scarce and potentially biased, Q-value estimates may suffer from high variance[3]. Although the ε-greedy strategy mitigates cold-start issues, it remains unclear how the framework performs under extremely small population sizes or very short evolutionary horizons.

Second, the action space is discretized into nine fixed combinations of CHTs and genetic operators. While this discretization guarantees interpretability, it also constrains the algorithm's expressive power. Future work could explore continuous action spaces or finer-grained parameterization[3,4]—for instance, using policy gradient methods to directly optimize crossover probabilities, mutation step sizes, or penalty coefficients as continuous variables.

Third, the computational overhead of the deep Q-network merits attention in resource-constrained scenarios. The theoretical time complexity is $O(T_{\max}(mN^3+((l-1)h^2+(r+o)h)\omega b))$, which, although remaining in the same order as the evolutionary component, introduces non-negligible constants through neural network forward passes and backpropagation. Extending this framework to large-scale CMOPs with high-dimensional decision variables[3] or dynamic environments where constraints and objectives change over time[2]presents both practical challenges and exciting research opportunities.

All in all, this work makes a thought-provoking contribution by recasting CMOEA design as a sequential decision-making problem solvable through deep reinforcement learning[3]. It reminds the community that evolutionary algorithm behavior across generations is not an unstructured random walk but contains learnable systematic patterns. This insight opens the door for more data-driven, autonomous algorithm configuration paradigms in evolutionary computation.

DECLARATIONS

Authors' contributions

Made substantial contributions to conception and writing of the manuscript: M.R, YJ.W, P.C;

Performed manuscript revision and provided supervision: P.C.

Availability of data and materials

Not applicable.

AI and AI-assisted tools Statement

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

REFERENCES

[1]Tian Y, Zhang Y, Su X, et al. Balancing objective optimization and constraint satisfaction in constrained evolutionary multiobjective optimization. *IEEE Transactions on Cybernetics 52.9* (2021)*:* 9559-72.[DOI: 10.1109/TCYB.2020.3021138]

[2]Fan Z, Li W, Cai X, et al. Push and pull search for solving constrained multi-objective optimization problems. *Swarm and evolutionary computation 44* (2019)*:* 665-79.[DOI: 10.1016/j.swevo.2018.08.017]

[3]Peng C, Yan S, Zhong C, et al. Learning-based temporal sequence of constrained handling selection for constrained multi-objective evolutionary optimization. *IEEE Transactions on Evolutionary Computation* (2025).[DOI: 10.1109/TEVC.2025.3584207]

[4]Zou M, Gong D, Wang Y, et al. Process knowledge-guided autonomous evolutionary optimization for constrained multiobjective problems. *IEEE Transactions on Evolutionary Computation 28.1* (2023): 193-207.[DOI: 10.1109/TEVC.2023.3243109]

[5]Li K, Chen R, Fu G, et al. Two-archive evolutionary algorithm for constrained multiobjective optimization. *IEEE Transactions on Evolutionary Computation 23.2* (2018): 303-15.[DOI: 10.1109/TEVC.2018.2855411]