نوزدهمین کنفرانس بین‌المللی مهندسی صنایع
2023
چرخش‌های تحول آفرین، هوشمندسازی و داده‌محوری
۲۶ , ۲۷ اردیبهشت ماه ۱۴۰۳
دانشگاه صنعتی امیرکبیر

Route Optimization for Order Picking in Distribution Centers using Reinforcement Learning based Genetic Algorithm

Matineh Rangzan[a], Hossein Akbaripour[b,*]

[a] *B.Sc. Student, Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran*

[b] *Assistant Professor, Department of Industrial Engineering and Management Systems, Amirkabir University of Technology, Tehran, Iran*

*** Corresponding author: Hossein Akbaripour, akbaripour@aut.ac.ir**

Abstract:

This research aims to use a hybrid approach combining **Multi-Agent Q-Learning (MAQL)** and **Genetic Algorithms (GAs)** to optimize pathfinding for Order Picking in distribution centers.

This approach investigates the use of **MAQL-GA** for finding the optimal path in the shortest time. By proposing a strategy for a centralized network of reinforcement learning agents that share information via a Genetic Algorithm and pass it to the next generation.

The objective of this research is to reduce the time required to identify the optimal path, as compared to both a non-centralized multi-agent algorithm and a single-agent q-learning algorithm.

The Environment:

The environment is an **11x11** distribution center. There are **4** types of cells in the environment: (The agent can move in four directions: up, down, left, and right.)

Cell Type	Description	Reward
Orders	The agent can move to these cells and pick up the order	**+100**
Aisles	The agent can use them to travel throughout the warehouse	**-1**
Shelves	The agent can not move to these cells and these locations are for storing items	**-100**
Crowded Area	It is better that the agent does not pass through this area	**-10, -50**

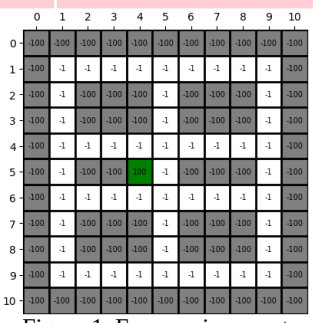

Figure 1. Easy environment

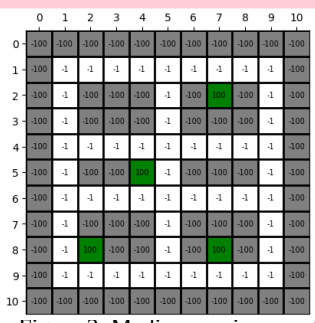

Figure 2. Medium environment

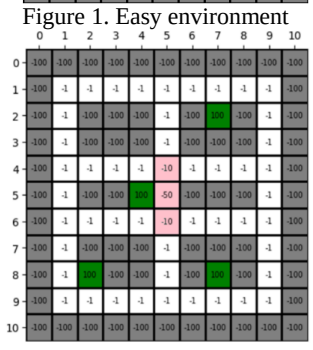

Figure 3. Hard environment

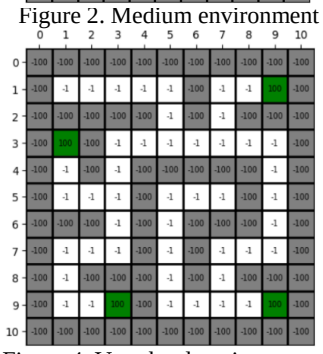

Figure 4. Very hard environment

Proposed Approach:

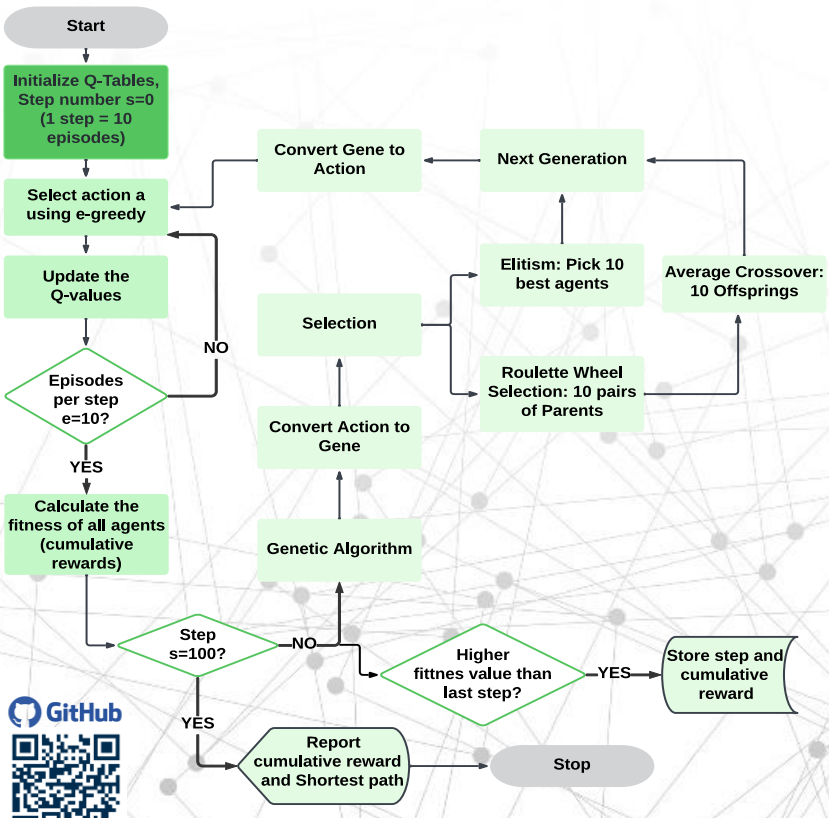

Figure 5. MAQL-GA Flowchart

Results and Discussion:

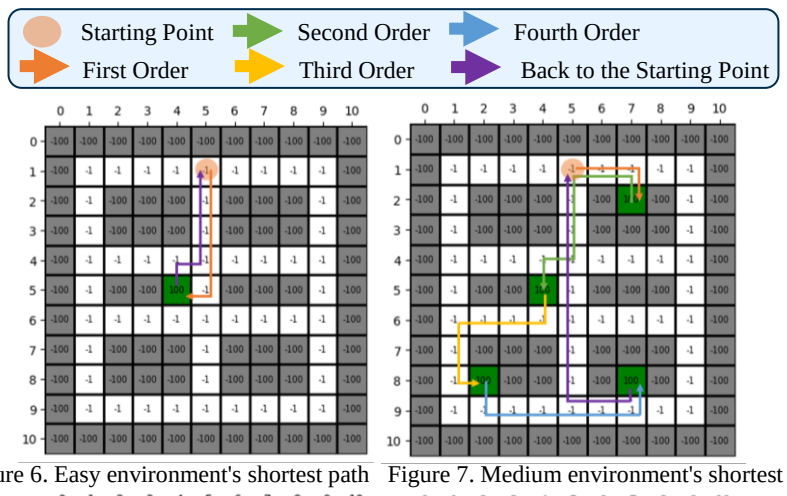

Figure 6. Easy environment's shortest path

Figure 7. Medium environment's shortest path

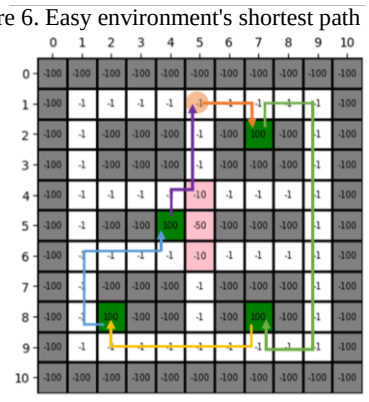

Figure 8. Hard environment's shortest path

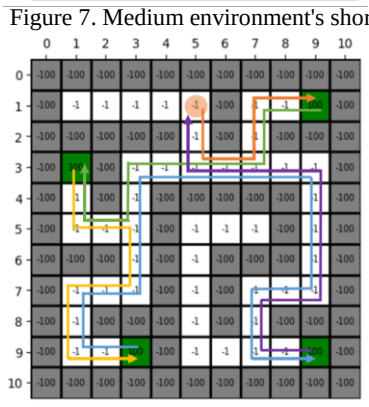

Figure 9. Very hard environment's shortest path

Our algorithm shows the fastest convergence amongst all other methods and across all environments.

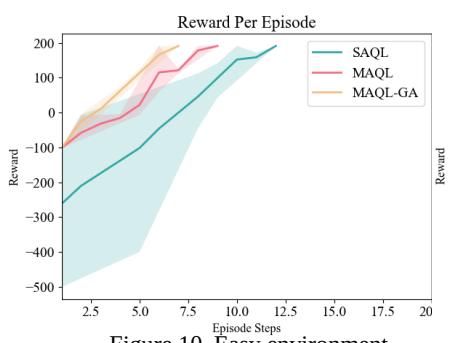

Figure 10. Easy environment

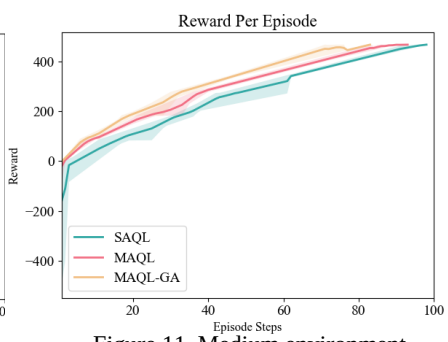

Figure 11. Medium environment

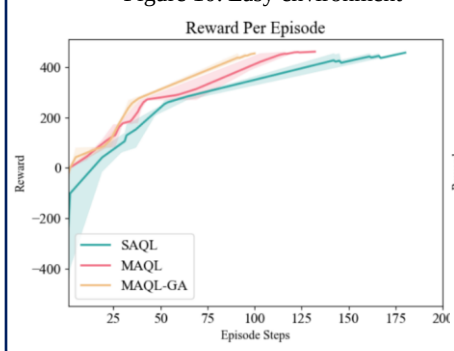

Figure 12. Hard environment

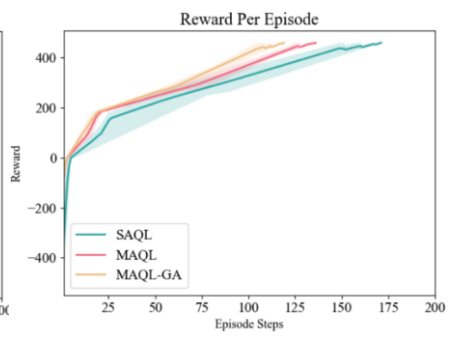

Figure 13. Very hard environment

Conclusion:

This study explored the integration of multi-agent reinforcement learning (MARL) and genetic algorithms (GAs) to enhance order-picking paths in distribution centers, demonstrating faster optimal path identification than Q-learning and non-centralized multi-agent Q-learning. Future research could focus on dynamic environments with multiple robots to further optimize efficiency and assess scalability in larger, complex layouts.

مسیریابی بهینه جمع آوری سفارشات در مراکز توزیع با استفاده از الگوریتم ژنتیک مبتنی بر یادگیری تقویتی

چکیده:

این تحقیق به دنبال استفاده از رویکرد ترکیبی شامل **یادگیری تقویتی چند عامله (MAQL)** و **الگوریتم‌های ژنتیک (GAs)** برای بهینه‌سازی مسیریابی جمع آوری سفارشات در مراکز توزیع است.

این رویکرد، استفاده از **MAQL-GA** را برای یافتن بهترین مسیر در کوتاه‌ترین زمان ممکن بررسی می‌کند. با پیشنهاد یک استراتژی برای شبکه متمرکزی از عوامل یادگیری تقویتی که اطلاعات را از طریق الگوریتم ژنتیک به اشتراک می‌گذارند و آن را به نسل بعد منتقل می‌کنند. هدف این تحقیق کاهش زمان مورد نیاز برای شناسایی مسیر بهینه در مقایسه با یک الگوریتم چندعاملی غیرمتمرکز و یک الگوریتم یادگیری تقویتی تک عاملی است.