# 第 2 章 文獻探討

## 2.1 引言

本章旨在系統性地回顧與機器人移動履行系統（Robotic Mobile Fulfillment System, RMFS）相關的交通管理研究。全文將循序漸進地展開：首先，深入闡述 RMFS 中交通壅塞問題的根源及其對系統效能的深遠影響；其次，依序評估傳統的啟發式控制方法、主流的深度強化學習（Deep Reinforcement Learning, DRL）方法，以及新興的神經演化（Neuroevolution）方法在解決此問題上的應用與局限。透過此一脈絡化的分析，本章旨在精準定位出現有研究的不足之處（Research Gap），並闡明本研究提出基於神經演化強化學習（Neuroevolution Reinforcement Learning, NERL）的交通控制策略之必要性與創新性。

## 2.2 機器人移動履行系統 (RMFS) 的運作與挑戰

RMFS 作為一種「貨到人」（Goods-to-Person）的倉儲自動化典範，其核心是透過大量自主移動機器人（Autonomous Mobile Robots, AMR）搬運可移動貨架（Pods），以取代傳統的人工揀選路徑，從而顯著提升訂單履行效率 [12]。自 Kiva Systems 被 Amazon 收購並大規模部署以來 [10]，RMFS 已成為現代電子商務物流的基礎設施，研究表明其產能可達傳統倉儲的 2 至 4 倍 [15]。隨著 Amazon 等企業部署的機器人數量突破 75 萬台，系統的運作密度與複雜性也達到了前所未有的高度 [22]。

然而，高密度部署帶來了新的瓶頸：**交通壅塞（Traffic Congestion）**。當大量機器人在有限的通道網路中執行任務時，尤其是在路網的交叉點（Intersections），極易因路權協調不當而產生衝突與等待。這不僅會引發訂單延遲、降低系統吞tu量，近期研究更指出，頻繁的啟停等待會顯著增加能耗，並加速電池老化 [24]。因此，設計高效且能自我適應的**交叉路口管理策略**，已成為提升整體 RMFS 效能、兼顧效率與能源可持續性的核心挑戰。

RMFS 內的決策問題本質上是多層次的，通常可分為策略、戰術與作業三個層級 [11]。其中，任務指派、路徑規劃與交通控制等**作業層（Operational Level）**決策，因其高動態性與對系統即時表現的直接影響，成為了近年來的研究熱點 [16]。本研究即聚焦於作業層中的交叉路口交通控制問題。

**【表格建議：表 2.2.1 - RMFS 決策問題層級與範例】** 為清晰展示 RMFS 決策的複雜性，建議此處插入一個表格。該表格應包含三欄：「決策層級」、「典型問題」與「代表性研究引用」，將 [11]、[14]、[19] 等文獻的研究內容歸納進去，以突顯本研究聚焦於「作業層」的定位。

## 2.3 交叉路口交通控制方法

為解決 RMFS 中的交叉路口交通問題，學界已提出多種控制策略，其發展脈絡大致可分為啟發式方法與基於學習的方法兩大類。

### 2.3.1 啟發式與規則式控制器 (Heuristic and Rule-based Controllers)

早期的解決方案多依賴於人工設計的啟發式規則，其優點是實現簡單、計算開銷低，常作為評估複雜演算法效能的基準。一類是**靜態時制控制 (Static Time-based Control)**，此方法借鑒了傳統城市交通號誌，為交叉路口設定固定的通行權切換週期 [14]。這種方法的實現對應了本研究中的**時間基礎控制器 (TimeBasedController)**，但其主要缺陷在於無法感知即時交通流量，當流量波動劇烈時，極易導致不必要的等待或通行資源的浪費。

為彌補靜態方法的不足，另一類**動態反應式控制 (Dynamic Reactive Control)** 應運而生。研究者提出了基於即時狀態的反應式規則，例如，根據等待隊列長度、任務優先級等指標動態分配路權 [14]。這類方法構成了本研究中**佇列基礎控制器 (QueueBasedController)** 的設計基礎。儘管反應更為靈活，但這些啟發式規則的設計高度依賴專家經驗，且難以捕捉複雜交通模式下的最優協同策略。

### 2.3.2 基於深度強化學習 (DRL) 的控制器

為了克服人工規則的局限性，深度強化學習（DRL）被廣泛應用於交通控制問題，使得智能體（控制器）能從與環境的互動中自主學習複雜的決策策略。深度 Q 網路（Deep Q-Network, DQN）及其變體是該領域最常用的演算法，大量研究已證實其在城市交通信號控制中，於減少車輛延誤和油耗方面超越了傳統方法 [30]。

近年來，研究者開始將此範式遷移至 RMFS 場景，用於資源分配 [27] 和交通協同。這些工作通常將交叉路口的局部交通資訊（如佇列長度、等待時間）作為**狀態 (State)**，將切換通行相位作為**動作 (Action)**，並以最大化通行量或最小化等待時間作為**獎勵 (Reward)**。儘管基於 DQN 的方法展現了巨大潛力，但它也面臨一些固有的挑戰。其性能高度依賴於密集且設計精良的即時獎勵信號，不良的設計可能導致非預期的次優行為。同時，在複雜的狀態空間中，基於梯度的優化方法有時會陷入局部最優解，並且可能需要龐大的訓練樣本才能達到收斂 [26]。鑑於 DQN 在該領域的代表性與成熟度，本研究將其作為一個關鍵的**比較基線 (Baseline)**，以評估新方法所能帶來的潛在改進。

## 2.4 探索新範式：神經演化強化學習 (NERL)

## 為應對傳統 DRL 方法的挑戰，本研究引入了神經演化強化學習（Neuroevolution Reinforcement Learning, NERL）。與 DQN 學習一個價值函數不同，NERL 直接在策略網路的參數空間 (Parameter Space) 中進行搜索。它將神經網路的權重視為個體的基因，透過演化演算法（如選擇、交叉、變異）來優化整個族群，最終得到高性能的控制策略。

## NERL 在解決複雜控制問題上展現出獨特優勢。首先，演化算法屬於無梯度（gradient-free）的黑箱優化，不要求獎勵函數連續可微。這使得它天然適用於處理稀疏、延遲或基於全局性能的獎勵信號。例如，Rizqi and Chou (2024) 成功地將 NERL 應用於多階供應鏈的庫存優化問題，他們直接使用整個模擬週期的總成本作為適應度函數，有效地解決了包含不確定折扣和多種運送選項的複雜決策問題。這種直接優化全局目標的能力，與本研究的全局獎勵 (Global Reward) 設計思想高度契合。

## 其次，演化過程中的交叉和變異操作有助於在參數空間中進行更廣泛的探索，降低了陷入局部最優的風險。最後，族群中每個個體的評估過程是相互獨立的，可以大規模並行化，有效利用現代計算資源縮短訓練時間。Rizqi and Chou (2024) 的研究亦證實，NERL 在庫存優化上的表現顯著優於傳統的 (s,Q) 模型，尤其在問題複雜度提高時，其優勢更為凸顯。

## 受到此研究的啟發，我們相信 NERL 的這些特性使其非常適合解決 RMFS 交通控制這一同樣具有高度動態性和不確定性的問題。儘管神經演化已在機器人控制等領域取得成功 [28]，但將其專門應用於 RMFS 交叉路口交通控制，並系統性地與啟發式方法和主流 DRL 方法（如 DQN）進行比較的研究，目前尚屬空白。

## 2.5 本章總結與研究缺口

綜合上述分析，現有研究的發展脈絡清晰地指向了幾個關鍵的待解議題。RMFS 交通控制策略正從簡單的靜態規則，演進到動態啟發式，再到基於 DRL 的自適應方法。然而，現有的 DRL 方法，主要是 DQN，在獎勵設計和避免局部最優方面仍存在挑戰。與此同時，多數交通控制研究主要關注吞吐量和時間延遲，而將**能源消耗**作為核心優化目標之一的研究相對匱乏。更重要的是，神經演化作為一種具備更強全局探索能力、且對稀疏全局獎勵更具魯棒性的方法，其在 RMFS 交通控制問題上的應用潛力尚未得到系統性的驗證。

因此，本研究的核心**研究缺口**在於：**缺乏一套能夠有效優化 RMFS 交通效率與能源消耗，並能克服傳統 DRL 局限的智能控制策略。**

為填補此缺口，本論文提出了一套基於神經演化強化學習（NERL）的交叉路口控制器，並將其與傳統啟發式方法和標準的 DQN 方法在一個考慮了能耗模型的高擬真度模擬平台上進行全面的比較。此研究不僅旨在提出一種性能更有潛力的控制策略，也為評估不同學習範式在解決複雜多智能體協同問題上的優劣提供了實證依據。