**Lab 5 結報**

姓名：賴昱凱 學號：111511141

1. 請簡述Resource Allocation的目的與重要性

**目的：**

Resource Allocation的主要目的是在有限的系統資源下，根據不同裝置的需求與通訊條件，合理且有效地分配可用資源（如頻寬、時間、空間、傳輸功率等），以達到系統整體效能最大化、提高服務品質（Quality of Service, QoS），並確保系統公平性與效率。

**重要性：**

1. **應對不同需求：**

移動通訊系統中，設備之間的服務需求多樣（例如語音、影像串流、即時通訊），資源分配確保各種服務能符合其特定的延遲、頻寬或可靠性要求。

1. **適應通道狀況：**

每台裝置的通道條件（如訊號強度、干擾情形）不同，資源分配機制會根據實際通道狀況進行調整，提升傳輸效率與穩定性。

1. **最大化系統容量：**

合理分配空間、時間、頻率與功率等資源，可在有限頻譜下提升總用戶數量與傳輸速率，增加整體系統容量。

1. **節能與延長裝置壽命：**

適當的資源分配能有效降低不必要的傳輸功率消耗，有助於裝置節能並延長電池使用時間。

1. **維持公平性：**

資源分配機制也要考慮不同用戶之間的公平性，避免資源過度集中於某些用戶而犧牲其他用戶的服務品質。

1. 請簡述Q-Learning的功用

Q-Learning是一種強化學習 (Reinforcement Learning, RL) 的方法。

其原理是讓一個Agent透過與環境互動來學習一個策略，以最大化累計的獎勵。它主要依賴一個稱為Q-table的表格。

1. **State**：

Q-Learning將系統目前的狀況表示為一個狀態 s。在本實驗中（無線通訊系統），狀態可以是對系統總傳輸速率 (Sum Rate) 的量化表示，是一個離散變數。

1. **Action**：

Agent在觀察到當前state後，需要選擇一個行動a。在本實驗中，action是指功率分配 (power allocation) 與使用者關聯 (association assignment) 的決策。選擇Action的方式採用-greedy策略，即以的機率隨機選擇行動，以的機率選擇Q-table中當前狀態下具有最大Q值 (期望的未來累計獎勵) 的行動。

1. **Reward**：

在執行一個行動後，Agent會收到來自環境的獎勵r。在本實驗中，這個reward就是系統的總傳輸速率 (throughput)，這個數值越高代表越好，因此很適合作為RL的reward。

1. **Q-table update**：

Q-Learning的核心是更新Q-table中的值。當Agent從狀態st採取行動at後轉移到狀態st+1並獲得獎勵r，它會根據**Bellman equation**來更新Q-table中Q(st, at)的值。更新公式為

這裡的η是學習率 (learning rate)，決定更新數值的幅度，δ是折扣因子 (discount factor)，決定是過去Q的影響大小，兩者都在0 ~ 1之間。這個更新過程結合了當前獲得的獎勵 r 與對未來狀態st+1的最大期望獎勵max Q(st+1, at+1)。

而為了處理系統中的限制條件（功率必須小於最大功率），還定義一個**懲罰項 (penalty)** 來懲罰違反限制的行動，因次在實驗中penalty必須是負數，才會達到懲罰的目的。

**功用：**

用一句話來說，上述的步驟就是為了找到在特定狀態下應該採取哪個行動，可以最大化獎勵，而對我們來說，獎勵最大化就代表這個系統表現的最佳化。

而Q-Learning在無線通訊系統中的功用，本實驗應用於網路資源分配 (Network Resource Allocation)。其目的是在考慮到不同裝置的通道條件、並在多個基地台 (AP) 和多個使用者 (UE)下，利用Q-Learning來選取較好的通道並分配功率進行下行資料傳輸，以達到最佳的傳輸速率 (throughput)。

1. 心得

本次實驗是最後一個module 2實驗，教了一個比較特別的AI訓練方式：增強式學習（Reinforcement learning, RL），相較於之前學習的DNN、CNN等NN模型，利用supervised和Gradient Descent的方式進行訓練不同，Reinforcement learning可以達到不需要有label過的解答，模型依然可以學會我們希望他做的事情，可以利用加分、扣分的方式讓模型意識到做什麼是對的什麼是錯的，這個結果我覺的十分有趣。

這次不是使用NN model來進行Reinforcement learning，讓我覺得有一點小可惜，因為我一直對如何在NN模型上套用reinforcement learning十分感興趣，但一直沒有課程有上到，幾乎都在使用supervised learning，不過多學到一種非模型的AI訓練方式也是有所收穫，也期待下一個module可以學到更多有趣的AI應用。謝謝助教這五個禮拜的教學，助教人都很好而且願意與我們詳細的一對一指導，讓我收穫良多。