

**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

-----



**Nguyễn Mạnh Thắng**

**NGHIÊN CỨU GIẢI PHÁP QUẢN LÝ TÀI NGUYÊN  
MẠNG O-RAN SỬ DỤNG HỌC SÂU**

**Chuyên ngành: Kỹ thuật Viễn thông**

**Mã số: 8.52.02.08**

**TÓM TẮT ĐỀ ÁN TỐT NGHIỆP THẠC SĨ**

**HÀ NỘI - 2024**

Đề án tốt nghiệp được hoàn thành tại:  
**HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG**

Người hướng dẫn khoa học: TS. Nguyễn Thị Thu Hiền

Phản biện 1: TS. Nguyễn Quang Hưng

Phản biện 2: PGS.TS. Đặng Thế Ngọc

Đề án tốt nghiệp sẽ được bảo vệ trước Hội đồng chấm đề án tốt nghiệp thạc sĩ tại Học viện Công nghệ Bưu chính Viễn thông

Vào lúc: 10 giờ 40 ngày 12 tháng 6 năm 2024

Có thể tìm hiểu đề án tốt nghiệp tại:  
- Thư viện của Học viện Công nghệ Bưu chính Viễn thông.

## MỞ ĐẦU

Trong những năm trở lại đây, ngành viễn thông đã có những bước tiến to lớn trong công cuộc truyền tải thông tin. Mạng truy nhập vô tuyến đã, đang và vẫn sẽ là một thành phần quan trọng trong mạng. Đặc biệt với sự phát triển của các mạng cảm biến không dây và nhu cầu truy cập mạng tốc độ cao mọi lúc mọi nơi càng thúc đẩy sự cần thiết của các công nghệ mới và giải pháp cải tiến các công nghệ đang có.

Mạng truy nhập vô tuyến mở (O-RAN) là một công nghệ tiềm năng nhằm giải quyết vấn đề độc quyền, không hỗ trợ lẫn nhau giữa các hãng sử dụng ảo hóa đem lại rất nhiều các ưu điểm so với mạng truy nhập vô tuyến truyền thông. Việc sử dụng học máy (ML) nói chung và học sâu (DL) nói riêng vào các công nghệ để phục vụ việc tối ưu, phân bổ tài nguyên hay xử lý dữ liệu cũng rất được quan tâm và nghiên cứu phát triển.

Với mục đích đưa ra các phương án phát triển cho những công nghệ tiềm năng để phục vụ cho cuộc sống, tôi xin chọn đề tài nghiên cứu **“Nghiên cứu giải pháp quản lý tài nguyên mạng O-RAN sử dụng học sâu”** với 3 chương như sau:

### CHƯƠNG 1: TỔNG QUAN VỀ MẠNG TRUY NHẬP VÔ TUYẾN MỞ

Chương 1 Giới thiệu tổng quan về mạng truy nhập vô tuyến mở, lịch sử hình thành và phát triển của các thế hệ mạng truy nhập vô tuyến. Kiến trúc và tiêu chuẩn đang được nghiên cứu phát triển cho mạng vô tuyến.

### CHƯƠNG 2: GIẢI PHÁP QUẢN LÝ TÀI NGUYÊN MẠNG O-RAN SỬ DỤNG HỌC SÂU

Chương 2 trình bày các về học sâu và các thuật toán trong quản lý tài nguyên mạng. Từ đó đưa ra phương pháp sử dụng học sâu trong quản lý tài nguyên của mạng truy nhập vô tuyến mở.

### CHƯƠNG 3: MÔ PHỎNG VÀ ĐÁNH GIÁ HIỆU NĂNG

Chương 3 đưa ra một kịch bản với các tham số chọn lọc từ đó đánh giá về hiệu quả của việc sử dụng học sâu trong quản lý tài nguyên mạng truy nhập vô tuyến mở.

# CHƯƠNG 1 – TỔNG QUAN VỀ MẠNG TRUY NHẬP VÔ TUYẾN MỞ

## 1.1 Giới thiệu chung về mạng truy nhập vô tuyến mở

### 1.1.1 Lịch sử phát triển và hình thành của mạng truy nhập vô tuyến mở

Mạng truy nhập vô tuyến có sự phát triển mạnh bắt đầu từ những năm 80 của thế kỷ 20. Với nhu cầu trong việc dễ dàng quản lý, giám sát cũng như thuận tiện trong cấu hình mạng vô tuyến. Cùng với sự phát triển của các công nghệ ảo hóa như SDN, NFV, có thể chia mạng truy nhập vô tuyến phát triển qua các giai đoạn với 4 mốc công nghệ như sau:

- Mạng truy nhập vô tuyến phân tán D-RAN
- Mạng truy nhập vô tuyến tập trung C-RAN
- Mạng truy nhập vô tuyến ảo hóa v-RAN
- Mạng truy nhập vô tuyến mở O-RAN

Nhờ có Liên minh O-RAN (một tổ chức gồm nhiều các ông lớn trong ngành cung cấp dịch vụ viễn thông cùng nhau hợp tác) mà mạng truy nhập vô tuyến mở đang dần được phổ biến và nghiên cứu rộng rãi

### 1.1.2 Các nguyên tắc trong mạng truy nhập vô tuyến mở

Tựu chung lại, để xây dựng một mạng O-RAN phù hợp với các tiêu chí đặt ra cũng như đáp ứng sự mở đối với các hãng khác, cần có một số nguyên tắc phải được tuân thủ như sau:

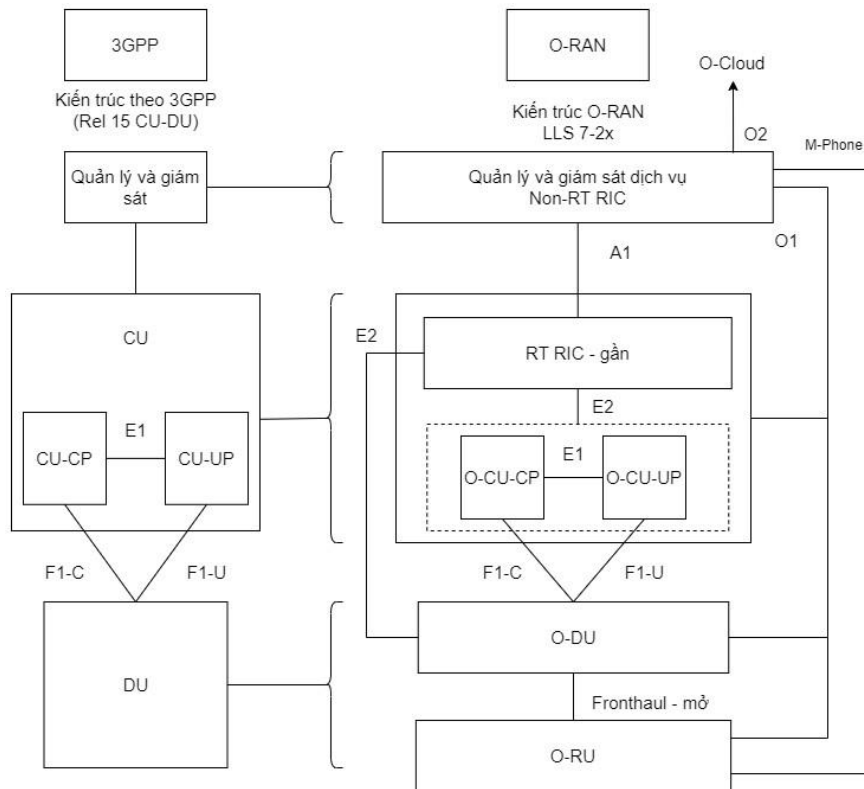
- Tính cởi mở
- Phần mềm hóa và ảo hóa
- Thông minh
- Khả năng lập trình
- Kiến trúc linh hoạt

## 1.2 Kiến trúc và tiêu chuẩn của mạng truy nhập vô tuyến mở

### 1.2.1 Kiến trúc của O-RAN

Kiến trúc cơ bản của O-RAN sẽ bao gồm các khối

- Khối vô tuyến RU để truyền tín hiệu vô tuyến
- Khối phân tán DU để xử lý dữ liệu, quản lý tài nguyên vô tuyến



**Hình 1. 1: Mô hình kiến trúc của O-RAN**

- Khối tập trung CU để quản lý kết nối, điều khiển di động và quản lý dữ liệu. Thường được chia làm mặt phẳng điều khiển CP và mặt phẳng người dùng UP để tách biệt chức năng điều khiển và dữ liệu
- Bộ điều khiển thông minh RAN (RIC): để điều khiển RAN thông qua near-real-time RIC và non-real-time-RIC

Bên cạnh đó cũng cần chú ý đến các giao diện trong các kết nối giữa các khối.

### 1.2.2 Sự phát triển của các tiêu chuẩn O-RAN

Trong những năm trở lại đây, sự phát triển của O-RAN không chỉ trong các công nghệ mà còn ở các tiêu chuẩn đã được Liên minh O-RAN đề xuất kiến nghị, điều này cho thấy triển vọng phát triển cho công nghệ này. Dưới đây là một số tiên tiến trong các tiêu chuẩn của O-RAN

- Kiến trúc dựa trên dịch vụ cho SMO: Kiến trúc SMO tách rời trong O-RAN trình bày SMO như một tập hợp các chức năng SMO cung cấp các dịch vụ SMO, đây là một tập hợp các khả năng quản lý, điều phối và tự động hóa được chuẩn hóa.
- Kiến trúc dựa trên dịch vụ cho Non-real-time RIC và giao diện R1: Kiến trúc Non-real-time RIC trong O-RAN WG2 trình bày một kiến trúc dựa trên dịch vụ, trong đó khung

Non-real-time RIC/SMO cung cấp một gói các dịch vụ sẽ được cung cấp bởi các chức năng khung Non-real-time RIC/SMO (hoạt động như nhà sản xuất dịch vụ).

- c. Mô hình dịch vụ và định nghĩa kiểu cho giao diện E2 và A1: O-RAN đã chuẩn hóa các mô hình dịch vụ mới, chẳng hạn như Mô hình dịch vụ E2 - Điều khiển RAN (E2SM-RC), Mô hình dịch vụ E2 - Giám sát hiệu suất chính (E2SM-KPM), Mô hình dịch vụ E2 - Cấu hình và điều khiển ô (E2SM-CCC) để triển khai các dịch vụ Near-real-time RIC qua giao diện E2 hướng tới việc thực hiện các trường hợp sử dụng O-RAN như định tuyến lưu lượng, QoS, phân đoạn mạng, MIMO quy mô lớn...
- d. Đám mây hóa và điều phối qua giao diện O2: O2 là một giao diện logic mở trong kiến trúc O-RAN để giao tiếp giữa SMO và O-Cloud để quản lý cơ sở hạ tầng O-Cloud và quản lý vòng đời triển khai của các chức năng mạng đám mây O-RAN chạy trên O-Cloud.
- e. M-plane phân cấp và lai cho giao diện open fronthaul: O-RAN đã chuẩn hóa hai tùy chọn cho quản lý OAM của O-RU. Trong M-plane phân cấp, quản lý OAM cho O-RU diễn ra thông qua O-DU. SMO sử dụng giao diện O1 để thực hiện các hoạt động FCAPS liên quan đến O-RU trên O-DU, sau đó sử dụng giao diện fronthaul M-plane với O-RU. Trong M-plane lai, quản lý OAM cho O-RU được thực hiện trực tiếp từ SMO thông qua giao diện M-plane open fronthaul tới O-DU.
- f. API Near-real-time RIC: O-RAN gần đây đã chuẩn hóa các API để cho phép khả năng tương tác và tích hợp giữa xApp của bên thứ ba, chịu trách nhiệm về RRM chi tiết của các chức năng C-plane, U-plane và M-plane của các chức năng mạng O-RAN qua giao diện E2 với độ chi tiết gần thời gian thực sử dụng các vòng lặp điều khiển độ trễ thấp và các chức năng nền tảng Near-real-time RIC.
- g. Giao diện mới để tiếp xúc thông tin phân tích RAN: Phân tích RAN chi tiết được cung cấp bởi nền tảng Near-real-time RIC qua giao diện dựa trên dịch vụ Y1 mới cho người tiêu dùng như các chức năng lõi gói tin hoặc máy chủ ứng dụng hoặc máy chủ cạnh có nhiều trường hợp sử dụng khác nhau, chẳng hạn như sau:
- h. Chức năng quản lý mạng lưới phân đoạn trong SMO

### 1.3 Kết luận chương 1

Chương này của đề án đã trình bày về sự lịch sử hình thành và phát triển của các thế hệ mạng truy nhập vô tuyến và tính thiết yếu của O-RAN khi các công nghệ ảo hóa lớp mạng ngày càng hoàn thiện hơn. Các nguyên tắc chung cần phải có cho mạng O-RAN cũng như kiến trúc đề xuất của Liên minh O-RAN cho mô hình vật lý của một mạng truy nhập mở từ

đó nắm được những lợi thế là hạn chế còn tồn tại trong quá trình nghiên cứu và phát triển sau này. Có thể thấy một trong những xu hướng đang rất được quan tâm là tích hợp AI hay ML vào trong giám sát và quản lý cho mạng O-RAN.

## **CHƯƠNG 2 – GIẢI PHÁP QUẢN LÝ TÀI NGUYÊN MẠNG O-RAN**

### **SỬ DỤNG HỌC SÂU**

#### **2.1 Giới thiệu về học sâu và những thuật toán**

##### **2.1.1 Giới thiệu về học sâu**

Những năm gần đây, Trí tuệ nhân tạo (AI) và cụ thể hơn là Học máy (ML) nổi lên như một minh chứng của cuộc cách mạng công nghiệp lần thứ tư (1 - cơ động hơi nước, 2 - năng lượng điện, 3 - công nghệ thông tin). Dù chỉ mới xuất hiện không lâu nhưng AI đã đang và sẽ hiện diện trong mọi lĩnh vực của đời sống con người, từ kinh tế, giáo dục, y khoa cho đến những công việc nhà, giải trí hay thậm chí là trong quân sự. Những ứng dụng nổi bật trong việc phát triển AI đến từ nhiều lĩnh vực để giải quyết nhiều vấn đề khác nhau. Nhưng những đột phá phần nhiều đến từ Học sâu (DL) - một mảng nhỏ đang mở rộng dần dần đến từng loại công việc, từ đơn giản đến phức tạp.

##### **2.2.2 Các thuật toán và đặc điểm**

Đối với DL, một số thuật toán nổi bật có thể kể đến:

###### **a. Mạng nơ-ron tích chập (CNN)**

Gồm có:

- Lớp tích chập (Convolutional Layer): Lớp này sử dụng các bộ lọc để quét qua đầu vào và phát hiện các đặc trưng cục bộ như cạnh và họa tiết.
- Lớp tổng hợp (Pooling Layer): Giảm kích thước không gian của biểu diễn đầu vào (thường dùng Max Pooling hoặc Average Pooling) để giảm tính toán và kiểm soát sự quá khớp (overfitting).
- Lớp kết nối đầy đủ (Fully Connected Layer): Kết nối tất cả các nơ-ron từ lớp trước đến mỗi nơ-ron trong lớp này, thường dùng cho phân loại cuối cùng.

Ứng dụng: Nhận diện hình ảnh, phân loại ảnh, phát hiện đối tượng, xử lý video.

###### **b. Mạng Nơ-ron hồi tiếp (RNN)**

Gồm có:

- Lớp hồi quy (Recurrent Layer): Các nơ-ron trong RNN có kết nối vòng lặp, cho phép thông tin từ các bước thời gian trước được sử dụng trong các bước tiếp theo.



- Bộ nhớ dài - ngắn hạn (LSTM): Một loại RNN đặc biệt với các cổng (gates) giúp kiểm soát luồng thông tin, khắc phục vấn đề về độ dốc biến mất (vanishing gradient).
- Đơn vị hồi tiếp có cổng (GRU): Tương tự như LSTM nhưng có cấu trúc đơn giản hơn với ít cổng hơn.

Ứng dụng: Xử lý chuỗi thời gian, xử lý ngôn ngữ tự nhiên (NLP), dịch máy, nhận diện giọng nói.

#### c. Mạng Nơ-ron đối nghịch tạo sinh (GAN)

Gồm có:

- Bộ tạo sinh (Generator): Mạng sinh tạo ra dữ liệu giả từ nhiễu ngẫu nhiên.
- Bộ phân biệt (Discriminator): Mạng phân biệt cố gắng phân biệt giữa dữ liệu thật và dữ liệu giả.

Ứng dụng: Tạo hình ảnh mới, tạo nhạc, sinh dữ liệu giả lập, cải thiện chất lượng ảnh.

#### d. Mạng Nơ-ron biến đổi (Transformer Networks)

Gồm có:

- Cơ chế tự tập trung (Self-Attention Mechanism): Cho phép mô hình tập trung vào các phần quan trọng của chuỗi dữ liệu.
- Kiến trúc bộ mã hóa – giải mã (Encoder-Decoder Architecture): Bao gồm một bộ mã hóa để xử lý đầu vào và một bộ giải mã để sinh đầu ra.

Ứng dụng: Xử lý ngôn ngữ tự nhiên, dịch máy, tóm tắt văn bản, phân loại văn bản.

#### e. Bộ mã hóa tự động (Autocoders)

Gồm có:

- Bộ mã hóa (Encoder): Nén dữ liệu đầu vào thành một biểu diễn mã hóa (latent representation).
- Bộ giải mã (Decoder): Giải mã biểu diễn mã hóa trở lại thành dữ liệu gốc.

Ứng dụng: Giảm chiều dữ liệu, nén dữ liệu, phát hiện bất thường, sinh dữ liệu.

#### f. Mạng Nơ-ron chuyển tiếp (FNN)

Gồm có:

- Lớp đầu vào: Nhận dữ liệu đầu vào.
- Các lớp ẩn: Một hoặc nhiều lớp ẩn với các nơ-ron liên kết hoàn toàn.
- Lớp đầu ra: Sinh ra kết quả cuối cùng.

Ứng dụng: Phân loại, hồi quy, nhận dạng mẫu.

g. Mạng Nơ-ron đồ thị (GNN)

Gồm có:

- Lớp tích chập đồ thị (Graph Convolutional Layer): Thực hiện tích chập trên đồ thị, truyền thông tin qua các nút và cạnh.
- Lớp tổng hợp đồ thị (Graph Pooling Layer): Giảm kích thước của đồ thị, tương tự như pooling trong CNNs.

Ứng dụng: Phân tích dữ liệu đồ thị, mạng xã hội, phân tích phân tử, dự đoán liên kết.

## **2.2 Giải pháp quản lý tài nguyên mạng O-RAN sử dụng học sâu**

### **2.2.1 Các ưu nhược điểm**

Học sâu có thể được áp dụng trong quản lý tài nguyên mạng truy nhập vô tuyến mở để cải thiện hiệu suất mạng, tối ưu hóa quy trình vận hành, và tăng cường trải nghiệm người dùng. Dưới đây là một số ứng dụng cụ thể:

- Tối ưu hóa tài nguyên mạng, dự đoán tài nguyên.
- Quản lý sự cố và bảo trì.
- Tối ưu hóa tài nguyên mạng và hiệu suất.
- Tăng cường Bảo mật và An toàn.
- Tối ưu hóa vận hành và tái cấu trúc mạng.

Từ đây, có thể đưa một số các ưu nhược điểm trong việc áp dụng học sâu vào trong mạng O-RAN. Về ưu điểm, có thể kể đến:

- Tăng cường hiệu suất mạng.
- Tự động hóa quy trình vận hành.
- Dự đoán sự cố và bảo trì.
- Tăng cường bảo mật và an toàn.
- Tối ưu hóa linh hoạt và tái cấu trúc mạng.

Ngược lại, việc sử dụng học sâu cũng có những nhược điểm nhất định:

- Yêu cầu dữ liệu lớn.
- Phức tạp trong triển khai.
- Tính toán và tài nguyên.
- Khó khăn trong diễn giải.
- Nhận thức về quyền riêng tư.

### ***2.2.2 Sử dụng học sâu trong quản lý tài nguyên mạng truy nhập vô tuyến mở***

Trong đề án này, thuật toán được sử dụng là Q-Learning, với các bước hành động, trạng thái và phần thưởng để huấn luyện một mô hình mạng truy nhập vô tuyến sử dụng O-RAN đơn giản.

Đối với Q-Learning, ngoài các bước trên còn cần lưu ý đối với giá trị của bảng Q. Bảng Q là biểu đồ tra cứu lưu trữ giá trị Q. Những giá trị Q đó được tính từ phần thưởng trong giai đoạn huấn luyện. Sau khi bảng Q được điền, có thể tìm kiếm hành động tối ưu để thực hiện dựa trên giá trị Q được ánh xạ từ trạng thái trong thử nghiệm hoặc thực tiễn thực tế. Tuy nhiên, có một số hạn chế đã biết khi có hoặc sử dụng bảng Q. Ví dụ: bảng Q yêu cầu nhiều bộ nhớ lưu trữ; nếu không, nó sẽ nhanh chóng vượt quá giới hạn. Ngoài ra, nếu tốc độ học cao, có thể bảng sẽ trống ở hầu hết các phần. Nói cách khác, bảng có thể có hầu hết các phần chưa được xem. Do đó, một số phương pháp giảm thứ tự được áp dụng để giữ cho kích thước bảng Q hợp lý và hoạt động tốt.

## **2.3 Kết luận chương 2**

Trong chương 2, đề án đã khái quát về học sâu, các thuật toán nổi tiếng của học sâu và ứng dụng của chúng. Việc áp dụng học sâu vào mạng O-RAN với các ưu nhược điểm ứng với từng ứng dụng. Đồng thời, phần sau của chương 2 đã giới thiệu chi tiết thuật toán Q-Learning, một thuật toán hiệu quả được sử dụng trong quản lý tài nguyên mạng. Đây cũng là cơ sở để thực hiện mô phỏng sẽ được học viên trình bày cụ thể và đưa ra đánh giá cụ thể trong chương 3.

## CHƯƠNG 3 – MÔ PHỎNG VÀ ĐÁNH GIÁ HIỆU NĂNG

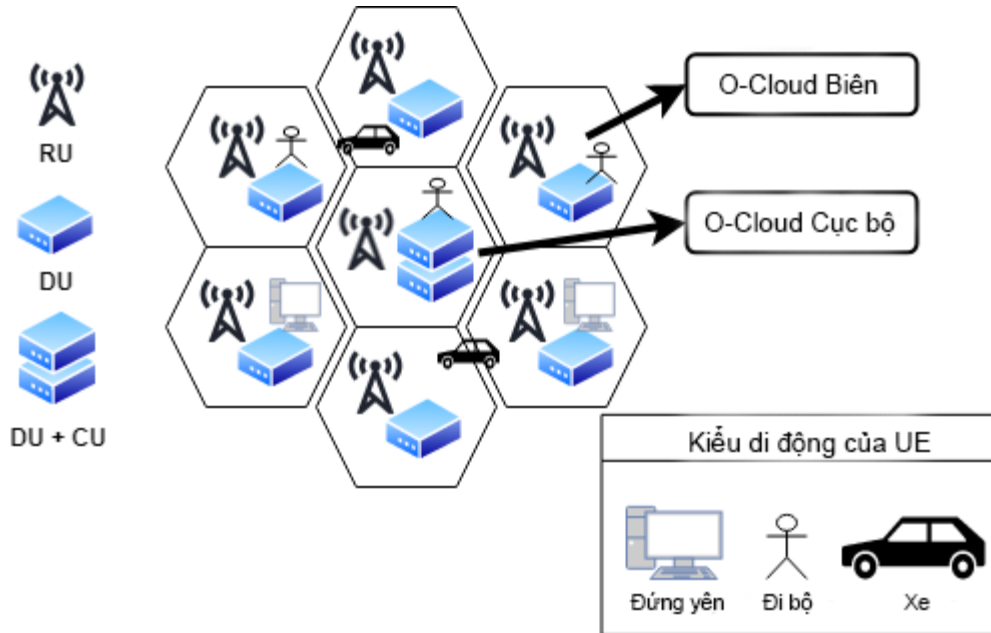
### 3.1 Kịch bản mô phỏng

Từ nội dung các chương 1 và 2 đã được nêu, chương 3 sẽ tập trung vào việc xây dựng một kịch bản mô phỏng khả năng quản lý tài nguyên mạng, cụ thể là tài nguyên băng thông cho mạng O-RAN khi sử dụng học sâu.

#### 3.1.1 Thông số cho O-Cloud

Một Edge O-Cloud có K lát mạng. Mỗi lát được gán cho  $b_j, k$  băng thông theo tỷ lệ phần trăm. Các lát mạng được triển khai là eMBB, URLLC và Voice.

Cấu trúc liên kết bao gồm một hình lục giác tổ ong 7 ô với sáu O-Cloud Biên có hình lục giác nhỏ xung quanh O-Cloud Cục bộ ở trung tâm với Non-RT RIC được hiển thị trong Hình 3.1. Edge O-Cloud bao gồm một cặp RU-DU. O-Cloud Cục bộ bổ sung CU cùng với cặp RU-DU. Do đó, cấu trúc liên kết có tổng cộng 7 RU, 7 DU và 1 CU là cấu trúc liên kết mạng tập trung.



Hình 3. 1: Mô hình mạng O-RAN sử dụng để mô phỏng

#### 3.1.2 Cấu hình cho UE

Phân phối Poisson được sử dụng trên miền thời gian để tạo sự kiện UE mới nhằm đưa tính ngẫu nhiên vào thuật toán RL. Một UE sẽ được gán ngẫu nhiên cho một lát mạng và một loại chuyển động độc lập. Trong quá trình mô phỏng, UE di chuyển tự do trong mạng O-RAN và yêu cầu dữ liệu của một lát mạng nhất định từ O-Cloud Biên gần đó. O-Cloud Biên gần đó

được chọn để thiết lập kết nối dựa trên thứ hạng khoảng cách và số lượng người dùng được kết nối.

**Bảng 3. 1: Loại UE và tốc độ tương ứng**

Loại UE	Tốc độ (km/h)
Đứng yên	0
Đi bộ	0 – 10
Xe	10 – 120

**a. Tính di động của UE**

Các UE ban đầu được phân phối ngẫu nhiên với phân bố Gaussian là vị trí ban đầu của chúng. Các UE sẽ di đến địa điểm mới mỗi giây dựa trên tốc độ nhất định được xác định trong Bảng 3.2.

Tốc độ di chuyển sau đây cho từng loại UE trong Bảng 3.2 được cho bởi. Đơn vị được trình bày trong yêu cầu là km/h. Trong quá trình thực hiện, tốc độ di chuyển được chuyển đổi thành mét trên giây. Quy ước nhằm điều chỉnh tốt hơn thang thời gian được sử dụng trong bước hành động và bước đáp ứng, tất cả đều được thực hiện hoặc tính toán mỗi giây.

**b. Tốc độ dữ liệu của UE**

Trong khi UE di chuyển trong O-Cloud, tốc độ dữ liệu của UE thay đổi không chỉ bởi băng thông được chỉ định mà còn bởi khoảng cách của nó so với tâm của Edge O-Cloud được kết nối nơi đặt RU. Mỗi quan hệ giữa tốc độ dữ liệu và khoảng cách nhận được bị ảnh hưởng bởi Mật độ công suất tần số vô tuyến (RF) (Công thức 3.1) trong [13] và Định lý Shannon-Hartley (Công thức 3.2) trong [14].

$$S = \frac{P \times G}{4\pi \times R^2} \quad (3.1)$$

$$C = B \times \log_2 \left( 1 + \frac{S}{N} \right) \quad (3.2)$$

Công suất tín hiệu thu  $S$  được tính bằng công suất máy phát  $P$ ; mức tăng công suất  $G$  trên khoảng cách  $R^2$ . Mật độ năng lượng có dạng hình cầu ba chiều ( $4\pi R^2$ ). Dung lượng kênh  $C$  trong (5.2) là tốc độ dữ liệu UE lý thuyết.  $B$  là băng thông của kênh tính bằng hertz.  $S/N$  là Tỷ lệ tín hiệu trên tạp âm (SNR).  $S$  là công suất tín hiệu thu được và  $N$  là công suất nhiễu. Do đó, tốc độ dữ liệu di của  $UE_i$  dựa trên khoảng cách  $R$  của nó tới O-Cloud được kết nối có thể được tính toán dễ dàng.

### 3.1.3 Cấu hình thuật toán Q-Learning

**Bảng 3. 2: Thông số cho Q-Learning**

Thông số	
Tốc độ học ( $\alpha$ )	0.1
Hệ số suy giảm ( $\gamma$ )	0.9
Hệ số phân rã $\varepsilon$	0.01
Giá trị $\varepsilon$ tối thiểu ( $\varepsilon_{min}$ )	0.01

Cài đặt Q-Learning được hiển thị trong Bảng 3.3. Mô hình sử dụng tốc độ học chậm ( $\alpha$ ) là 0,1 để giá trị bảng Q thay đổi rất nhỏ qua từng tập và hệ số suy giảm điển hình ( $\gamma$ ) là 0,9 để kết quả từng tập được đánh giá có giá trị gần tương đương cho các tập trước đó cho phương trình Q-Learning (3.3) [22]. Phương trình này nhằm tối ưu hành động tiếp theo dựa trên giá trị của bước hiện tại. Cách tiếp cận tham lam động được triển khai trong phương trình (3.4) [23] để tính xác suất thăm dò ( $\varepsilon$ ). Giá trị lớn hơn giữa giá trị  $\varepsilon$  hoặc  $\varepsilon$  tối thiểu nhất định của tập được sử dụng. Giá trị  $\varepsilon$  trên mỗi tập được tính bằng cách lấy giá trị hàm mũ của hệ số phân rã âm nhân với tập thứ  $i$ . Do đó, giá trị  $\varepsilon$  được sử dụng sẽ nằm trong khoảng từ 1 đến  $\varepsilon_{min}$  (Phương trình (3.5) [23]). Phương pháp khám phá này sẽ bắt đầu quá trình đào tạo với nhiều giá trị hơn và kết thúc bằng việc có được nhiều kết quả hơn. Các giá trị của  $\varepsilon$  được chọn là 0.01 nhằm đảm bảo tính ngẫu nhiên được lựa chọn ít từ đó tập trung kết quả dựa theo giá trị của bảng Q trong các tập trước đó.

$$Q(s_t, a_t) < (1 - \alpha)Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_A Q(s_{t+1}, A)] \quad (3.3)$$

$$\varepsilon = \max(\varepsilon_{min}, e^{-(\text{decay} * \text{episode})}) \quad (3.4)$$

$$s.t. \ \varepsilon_{min} < \varepsilon < 1 \quad (3.5)$$

#### 3.1.4 Thiết lập thử nghiệm

**Bảng 3. 3: Thiết lập UE**

Thiết lập số UE	Trung bình	Cao điểm
UE cho eMBB	[100, 200, 300, 400, 500]	
UE cho URLLC	500	1000
UE cho thoại	1000	500

Để chứng minh rằng mô hình Q-Learning có thể sử dụng băng thông trong nhiều tình huống khác nhau, số lượng UE tăng lên trong mạng giữa các bộ thử nghiệm khác nhau được

hiển thị trong Bảng 3.4. Sự gia tăng của UE có thể cho thấy hiệu suất của mô hình từ hoạt động mạng nhàn rỗi đến hoạt động mạng bận rộn. Xem xét các nhóm thử nghiệm và được kiểm soát, các UE của lát mạng được nhắm mục tiêu, eMBB, tăng thêm 100 UE từ 100 UE lên 500 UE giữa các bộ thử nghiệm khác nhau. Các lát cắt mạng, URLLC và thoại được kiểm soát có số lượng UE cố định. Cấu hình UE với 500 UE URLLC và 1000 UE thoại được cấu hình cho trường hợp lưu lượng trung bình. Cấu hình UE với 1000 UE URLLC và 500 UE thoại được đặt cho trường hợp lưu lượng cao điểm. Để đánh giá mô hình và hiệu suất ở các mức độ cao điểm, đề án đã tiến hành chạy mô phỏng gồm 1000 tập trên tất cả năm bộ eMBB UE.

### 3.1.5 Đường cơ sở

**Bảng 3. 4: Thông số đường cơ sở**

<b>Băng thông (%)</b>	<b>eMBB</b>	<b>URLLC</b>	<b>Thoại</b>
Cân bằng	40	30	30
Tập trung vào eMBB	90	5	5

So với mô hình Q-Learning, hai bộ đường cơ sở phân bổ băng thông được tạo trong Bảng 3.5. Cấu hình Cân bằng phân bổ băng thông một cách công bằng giữa các lát mạng và chỉ định thêm 10% băng thông cho lát eMBB vì eMBB được biết đến với yêu cầu lưu lượng truy cập cao hơn. Đường cơ sở còn lại, Tập trung vào eMBB, cung cấp 90% băng thông cho phần eMBB và hai phần mạng còn lại chia sẻ 10% băng thông còn lại để đáp ứng yêu cầu tối thiểu.

### 3.2 Công cụ mô phỏng

Đề án này sẽ xây dựng mô phỏng dựa trên Python để tái tạo một môi trường O-RAN bao gồm các thiết bị người dùng như hình 3.1. Sử dụng thuật toán Q-Learning chạy bằng Python để tương tác trực tiếp với môi trường mô phỏng. Ứng dụng để chạy mô phỏng

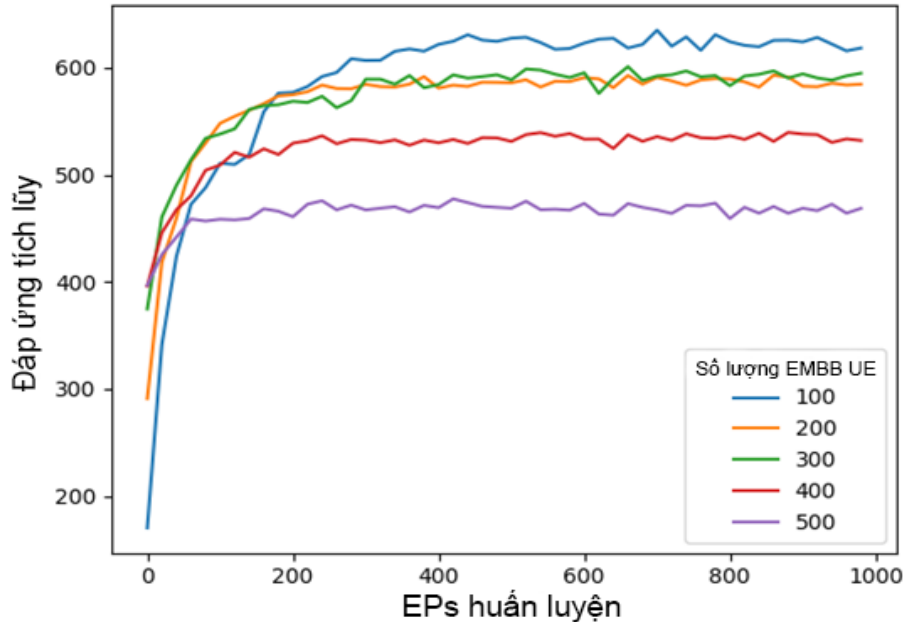
- SimPy [19]: Đây là một thư viện mở cho phép tạo và biểu diễn các biểu thức toán học phức tạp, giải các bài toán toán học, thực hiện các tính toán ký pháp số và phân tích các biểu thức toán học. Một môi trường lý tưởng để xử lý tuần tự các yêu cầu từ UE gửi đến các O-Cloud biên.
- OpenAI [20] tạo ra môi trường MDP trong đó tác nhân thực hiện hành động dựa trên các trạng thái của môi trường và môi trường sẽ cung cấp các trạng thái mới và

đáp ứng trả lại cho tác nhân. Để tùy chỉnh các tham số Q-Learning, gói GYM [21] từ OpenAI được chọn vì nó cũng được viết bằng Python. Gói GYM cung cấp khả năng tùy chỉnh từ các trạng thái, hành động và đáp ứng ban đầu để tạo lại mô hình Q-Learning MDP theo ý muốn.

- Mô hình kết hợp: Các gói SimPy, GYM và OpenAI có thể xếp chồng và tương tác tốt với nhau vì chúng đều tuân theo thực hành Lập trình hướng đối tượng (OOP) được viết bằng Python. SimPy tạo môi trường mạng. GYM xác định trạng thái, hành động và đáp ứng ở mỗi dấu thời gian. OpenAI cung cấp môi trường thực hành Q-Learning cũng như cơ sở của các thuật toán học sâu khác, có thể được tích hợp trong tương lai.
- Matplotlib: Là một thư viện cài thêm để có thể xuất kết quả dưới dạng hình ảnh bởi kết quả sau khi chạy Q-Learning thường sẽ cho ra các bảng số liệu theo từng tập

### 3.3 Thực hiện mô phỏng và đánh giá

#### 3.3.1 Thực hiện mô phỏng



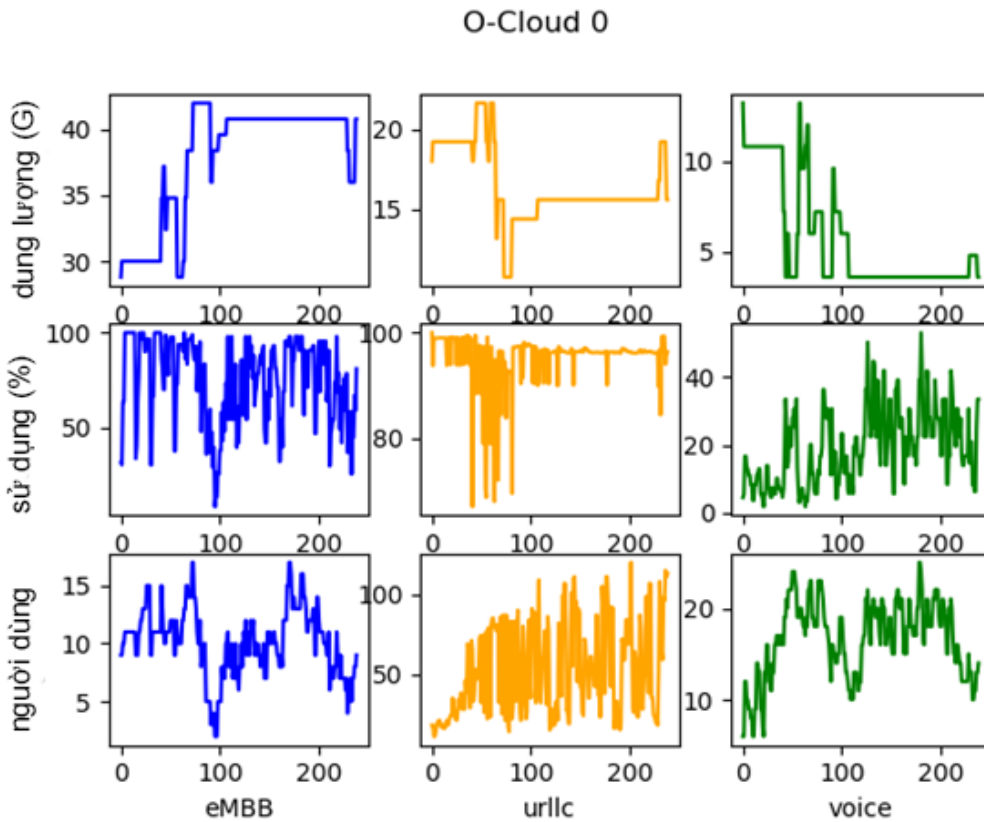
*Hình 3. 2: Sự hội tụ đáp ứng trong quá trình học*

Hình 3.2 cho thấy đường cong học tập của các kịch bản eMBB khác nhau. Đường cong học tập được vẽ bằng cách tích lũy đáp ứng theo thời gian cho 1000 tập. Trong giai đoạn huấn luyện, đáp ứng tích lũy được ghi lại sau mỗi lần lặp và mất trung bình 50 lần lặp cho mỗi điểm dữ liệu. Năm đường cong học tập được vẽ để thể hiện kịch bản của từng người dùng



eMBB. Giá trị  $Q$  tối ưu được tìm thấy trong khoảng 400 tập được hiển thị trong Hình 3.2, nơi các đường cong học tập hội tụ. Sự hội tụ của đáp ứng tích lũy chứng tỏ mô hình có thể học hỏi từ môi trường O-RAN mô phỏng. Vì vậy, mô hình đã sẵn sàng để tiến tới giai đoạn thử nghiệm. Giai đoạn thử nghiệm so sánh kết quả Q-Learning với đường cơ sở của cùng số lượng người dùng eMBB trong cùng một kịch bản. Như đã đề cập trong Bảng 3.5, đường cơ sở có băng thông tĩnh được phân bổ cho từng lát mạng. Thời lượng mô phỏng được đặt thành 120 giây để thu thập kết quả từ đường cơ sở và mô hình Q-Learning.

Để đánh giá xem mô hình Q-Learning được đề xuất có thể điều chỉnh theo các luồng lưu lượng thay đổi trong môi trường mạng mô phỏng hay không, Hình 3.3 là ảnh chụp nhanh từ mô phỏng đang chạy được tạo từ một trong các O-Cloud ở biên. Hình này hiển thị băng thông được phân bổ giữa các lát khác nhau, mức sử dụng hoặc nhu cầu băng thông thực tế của từng lát và số lượng người dùng trên mỗi lát. Số liệu thống kê của lát eMBB có màu xanh lam ở cột bên trái. URLLC có màu vàng ở cột giữa. Phần thoại có màu xanh lục ở cột bên phải.



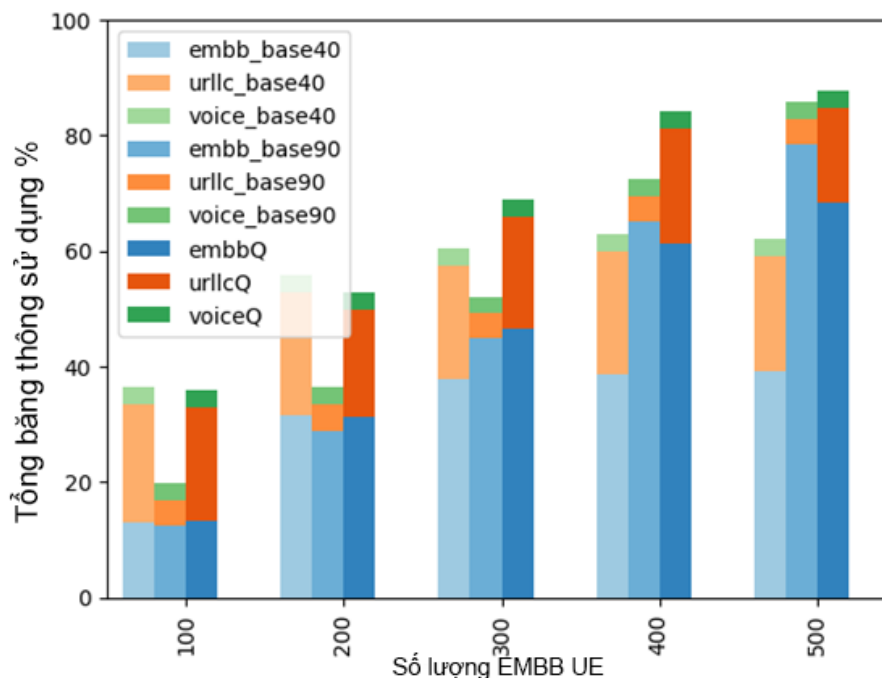
**Hình 3. 3: Phân bổ băng thông trong quá trình huấn luyện**

Hàng đầu tiên hiển thị dung lượng băng thông được phân bổ cho mỗi lát mạng. Phần eMBB được phân bổ nhiều băng thông hơn theo thời gian, trong khi phần URLLC và phần

giọng nói nhận được ít băng thông hơn theo thời gian. Hàng thứ hai của hình hiển thị phần trăm mức sử dụng dựa trên băng thông được phân bổ. Việc tăng băng thông không làm giảm nhiều tỷ lệ sử dụng eMBB, nghĩa là thông lượng tăng lên cho mỗi UE. Hàng thứ ba là số lượng người dùng được kết nối với O-Cloud cho mỗi lát, phản ánh nhu cầu lưu lượng của từng lát. Tỷ lệ phần trăm sử dụng tăng hoặc giảm tương ứng với sự thay đổi về số lượng người dùng.

### 3.3.2 Kết quả và phân tích

Trong kịch bản Lưu lượng trung bình, các tham số được kiểm soát là 1000 UE sử dụng Voice slice và 500 UE URLLC được đặt trong Bảng 3.4. Không có phần nào trong số hai phần mạng này dự kiến sẽ tiêu thụ nhiều băng thông hơn phần eMBB. Trong các hình dưới đây, chúng ta có lược đồ màu tiêu chuẩn cho từng lát mạng. eMBB có màu xanh lam, URLLC sử dụng màu cam và phần giọng nói hiển thị màu xanh lục. Trong phần sử dụng băng thông Hình 3.4, kết quả được nhóm theo số lượng người dùng eMBB từ trái sang phải. Trong mỗi nhóm, ba thanh từ trái sang phải biểu thị kết quả cơ bản cân bằng, kết quả eMBB Focus và kết quả Q-Learning.

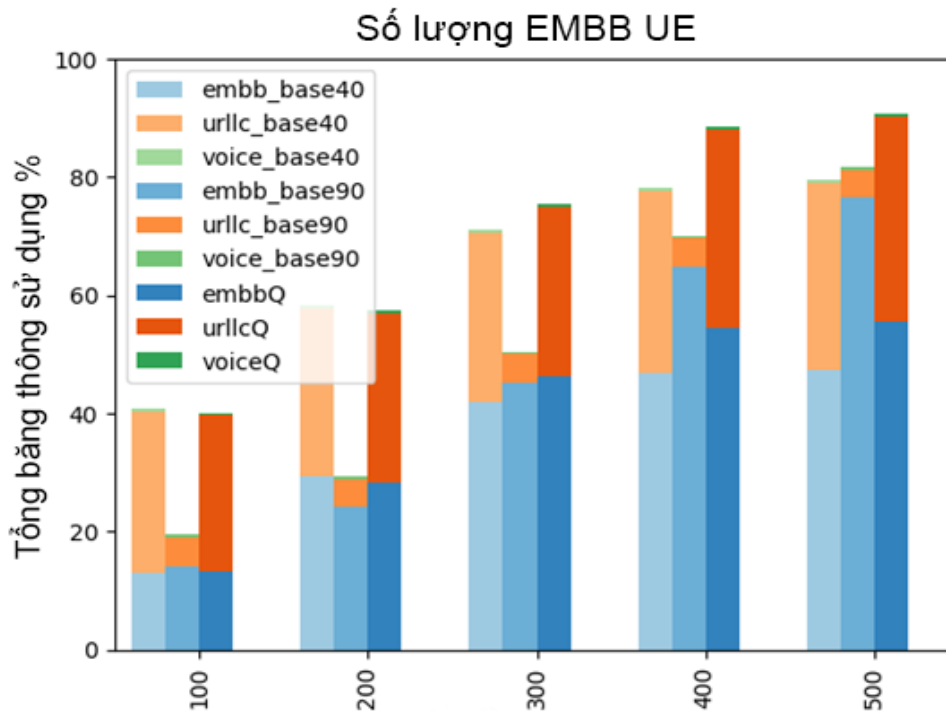


**Hình 3. 4: Băng thông sử dụng khi trung bình**

Đường cơ sở Cân bằng cho thấy rằng phân bổ tĩnh giới hạn băng thông eMBB ở giá trị nhất định trong khi một lượng lớn băng thông ở các phần khác không được sử dụng. Hình

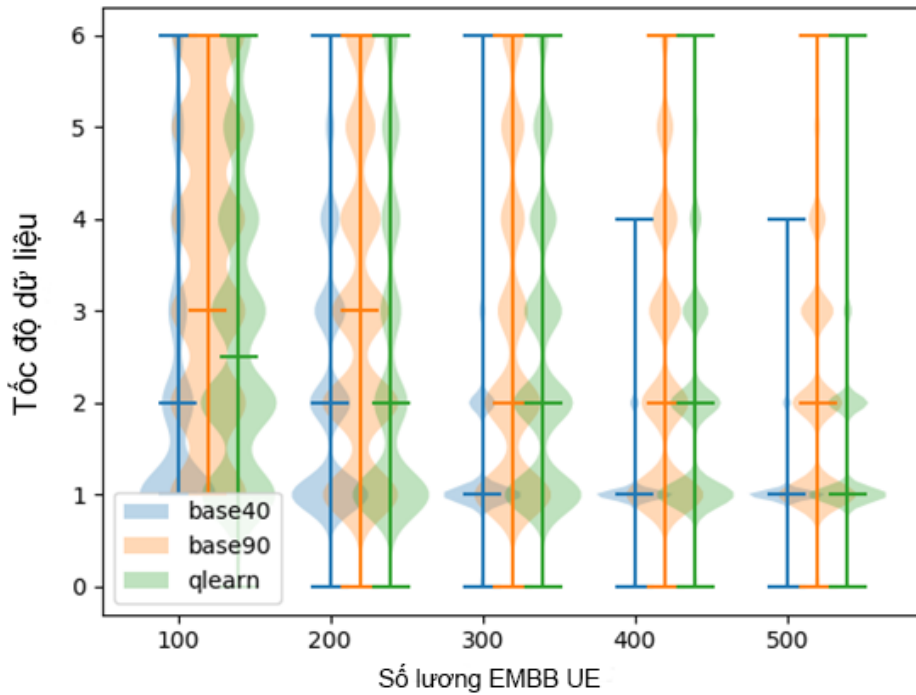
3.4 minh họa rằng đường cơ sở cân bằng cung cấp đủ băng thông cho người dùng 100, 200 và 300 eMBB trong cấu hình Lưu lượng cao. Tuy nhiên, cách tiếp cận băng thông cân bằng tĩnh này không phù hợp với người dùng eMBB tăng cao hơn. Thông lượng eMBB đã đạt đến giới hạn băng thông 40% và không thể tăng thêm.

Sau khi thấy phần eMBB bị giới hạn băng thông tĩnh trong đường cơ sở cân bằng, đường cơ sở tập trung eMBB sẽ cải thiện và chỉ định 90% băng thông cho eMBB vì các yêu cầu eMBB được cho là sẽ đòi hỏi băng thông trong mạng. Theo kết quả của kịch bản cơ sở này trong Hình 3.4, lát eMBB nhận đủ băng thông trong khi người dùng tăng số lượng eMBB. Tuy nhiên, sự đánh đổi là các lát mạng khác không thể yêu cầu thêm băng thông từ mạng, mặc dù lát eMBB có băng thông chưa được sử dụng trong các tình huống ít yêu cầu về lưu lượng hơn ở phía bên trái của Hình 3.4.



**Hình 3. 5: Băng thông sử dụng khi cao điểm**

Mặt khác, mô hình Q-Learning có thể tự động gán các phần không sử dụng bổ sung từ các lát mạng khác cho lát eMBB trong khi yêu cầu eMBB tăng từ 100 người dùng lên 500 người dùng. Đồng thời, các lát khác có thể nhận đủ băng thông được phân bổ trong khi băng thông chưa sử dụng có sẵn trong mạng. Trong kịch bản có tối đa 500 người dùng eMBB, mô hình Q-Learning có thể cung cấp phân bổ băng thông hợp lý cho từng lát mạng bằng cách hy sinh một phần băng thông eMBB và chia sẻ nó với những người khác.



**Hình 3. 6: Tốc độ dữ liệu trung bình trường hợp lưu lượng trung bình**

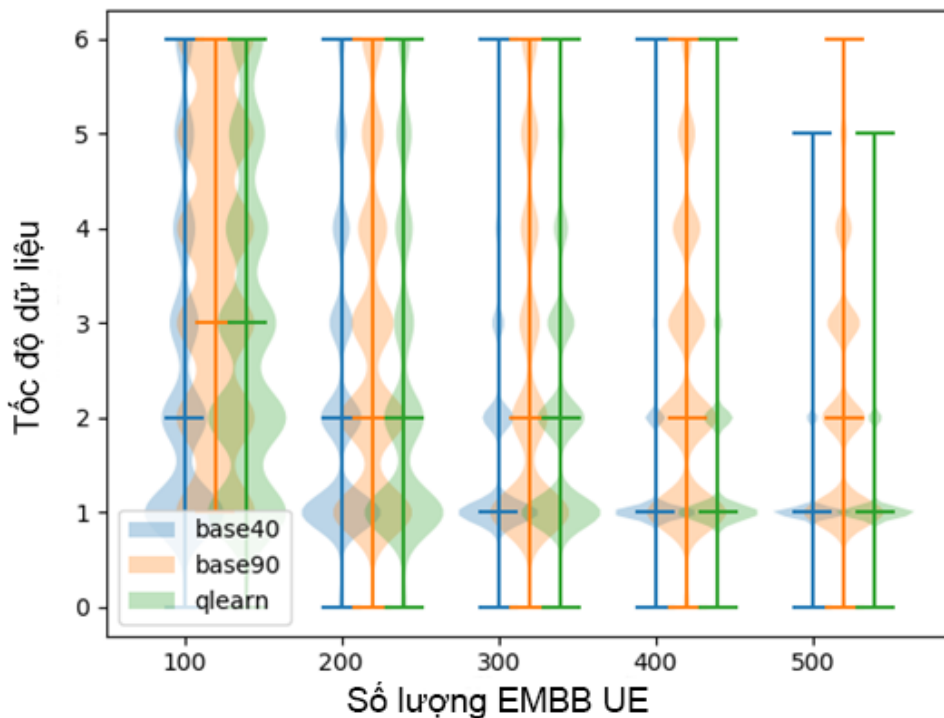
Khi số lượng eMBB UE thấp, đường cơ sở cân bằng, base40, có thể quản lý sự gia tăng nhu cầu URLLC cũng như mô hình Q-Learning. Tuy nhiên, đường cơ sở tập trung eMBB, base90, cung cấp phân bổ băng thông bảo hòa cho lát URLLC. Đúng như mong đợi, băng thông 5% được phân bổ cho URLLC được người dùng URLLC cao hơn sử dụng hoàn toàn trong cấu hình lưu lượng truy cập cao chỉ với 100 eMBB UE được hiển thị ở phía bên trái của Hình 3.5. Di chuyển về phía bên phải của Hình 3.5, đường cơ sở cân bằng không đủ để cung cấp đủ băng thông cho lát eMBB. Đường cơ sở tập trung eMBB không bao giờ nỗ lực nhiều hơn để chia sẻ băng thông tới lát URLLC. Mô hình Q-Learning cố gắng tìm ra sự công bằng trong việc chia sẻ băng thông giữa các lát URLLC và eMBB bằng cách hy sinh một số băng thông eMBB nhiều hơn cấu hình Lưu lượng trung bình.

Nhìn chung, kết quả từ trường hợp Lưu lượng truy cập cao đã nêu bật những khác biệt đã biết giữa hai đường cơ sở và mô hình Q-Learning. Mô hình Q-Learning có xu hướng coi trọng yếu tố cân bằng trong khi vẫn duy trì phân bổ băng thông cao trên lát eMBB mục tiêu.

Tốc độ dữ liệu UE và thời gian truyền (thời gian kết thúc) đối với các yêu cầu dữ liệu tĩnh là hai chỉ báo tốt thể hiện QoS và QoE của UE do giải pháp Q-Learning đề xuất của chúng tôi cung cấp. Trong phần sau, kết quả so sánh giữa hai đường cơ sở, đường cơ sở Cân

bằng và tập trung eMBB cũng như phương pháp Q-Learning được đề xuất về mặt trải nghiệm của UE.

Hình 3.6 hiển thị phân bố tốc độ dữ liệu người dùng của mạng trong cấu hình Lưu lượng giữa. Biểu đồ vĩ cảm màu xanh từ đường cơ sở cân bằng cho thấy dân số ít hơn ở tốc độ dữ liệu cao hơn cả đường cơ sở tập trung eMBB và kết quả Q-Learning. Mặt khác, các đường kẻ của mô hình Q-Learning và đường cơ sở tập trung eMBB có kết quả tương tự về cả tốc độ dữ liệu tối đa (tốc độ cao nhất) và các quần thể phụ. Như phân tích được trình bày trong phần tốc độ dữ liệu UE ở trên, đường cơ sở tập trung eMBB, base90, có lượng dữ liệu tốc độ cao lớn hơn một chút so với mô hình Q-Learning vì mô hình Q-Learning chia sẻ một số băng thông với lát URLLC. Một điều cần lưu ý trong Hình 3.6 là thanh ngang nhỏ trong biểu đồ là đường trung bình của phân bố dữ liệu. Đường cơ sở tập trung eMBB luôn có giá trị trung bình cao nhất so với hai mô hình còn lại. Giá trị trung bình của mô hình Q-Learning khác nhau giữa giá trị trung bình của hai đường cơ sở.

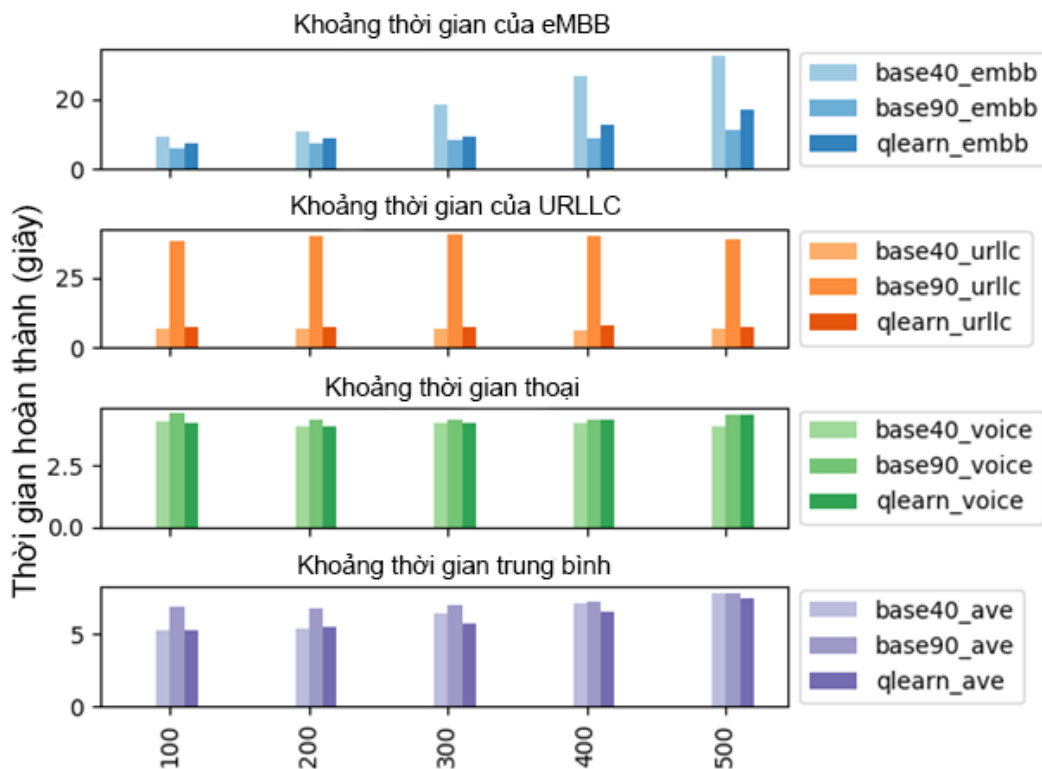


**Hình 3. 7: Tốc độ dữ liệu trung bình trường hợp cao điểm**

Kết quả Lưu lượng truy cập cao điểm được hiển thị trong Hình 3.7. Các kết quả cho thấy xu hướng tương tự là đường cơ sở cân bằng có hầu hết quần thể tốc độ dữ liệu ở phía dưới. Cơ sở tập trung eMBB cho thấy khả năng cung cấp QoS tốt cho người dùng eMBB

bằng cách duy trì tốc độ cao nhất lên tới 500 người dùng eMBB và mức trung bình cao nhất trong số hai mô hình còn lại. Với nhu cầu URLLC cao hơn, tốc độ cao nhất và tốc độ trung bình trên Q-Learning hiện giảm xuống khi người dùng eMBB tăng lên.

Thời gian truyền là một chỉ số khác của QoE. Nói chung, nếu một mô hình có thể hoàn thành quá trình truyền nhanh hơn với cùng lượng dữ liệu mà UE yêu cầu thì mô hình đó sẽ cung cấp trải nghiệm tải xuống tốt hơn cho người dùng. Trong kết quả sau, thời gian truyền được so sánh riêng lẻ trên các lát khác nhau và cũng được cộng lại thành tổng mức trung bình. Từ trái sang phải, chúng ta có 5 nhóm giống nhau bằng cách tăng 100 người dùng eMBB, bắt đầu từ 100 người dùng, như mô tả trong Bảng 3.4. Mỗi lát sử dụng cùng một màu như trong hình trên. eMBB có màu xanh lam, URLLC có màu cam và giọng nói có màu xanh lục. Màu tím mới được giới thiệu để thể hiện tổng thời gian truyền trung bình. Các thanh màu giống nhau từ sáng đến tối biểu thị đường cơ sở cân bằng, đường cơ sở tập trung eMBB và kết quả Q-Learning. Để so sánh kết quả một cách chính xác, yêu cầu dữ liệu sử dụng các giá trị tính khác nhau tùy thuộc vào loại lát mạng của nó.

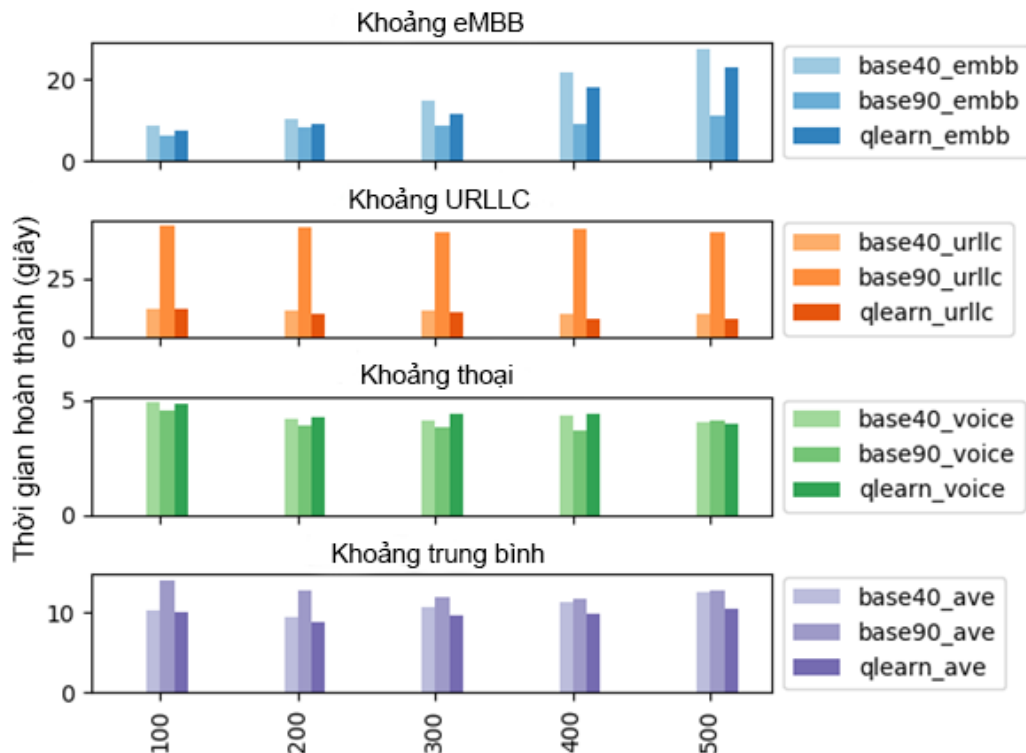


Hình 3. 8: Thời gian truyền của UE khi lưu lượng trung bình

Trong cấu hình lưu lượng trung bình, thời lượng yêu cầu eMBB có màu xanh lam tăng lên khi có nhiều người dùng eMBB tham gia mạng ở hàng trên cùng trong Hình 3.8. Thời gian truyền tăng chậm hơn trên đường cơ sở tập trung eMBB và mô hình Q-Learning so với

đường cơ sở Cân bằng khi tăng số lượng người dùng eMBB giữa các đợt kiểm tra. Như đã nêu trước đó, Q-Learning được biết là chia sẻ nhiều băng thông chưa sử dụng đến lát URLLC hơn eMBB. Vì Q-Learning có ít băng thông được phân bổ cho eMBB hơn nên thời gian truyền trên mô hình Q-learning cao hơn một chút so với đường cơ sở tập trung eMBB.

So sánh thời gian truyền của lát URLLC được hiển thị bằng màu cam, đường cơ sở cân bằng thực sự mất thời gian hoàn thành ngắn hơn so với Q-learning trong Hình 3.8. Ngược lại, đường cơ sở tập trung eMBB có hiệu suất kém nhất đáng kể trong việc hoàn thành quá trình truyền đối với cùng một lượng dữ liệu mà các UE URLLC yêu cầu so với hai mô hình còn lại. Hiệu suất không mong muốn này là do băng thông tối thiểu 5% được phân bổ cho lát URLLC.



**Hình 3. 9: Thời gian truyền của UE khi lưu lượng cao điểm**

Trong cấu hình Lưu lượng cao điểm, thời gian truyền từ lát eMBB đến tổng thời gian trung bình trong Hình 3.9 theo cùng thứ tự trong Hình 3.8. Có thể thấy, đường cơ sở tập trung eMBB có thời gian truyền ngắn nhất do băng thông được phân bổ nhiều nhất cho việc sử dụng eMBB. Thời gian truyền eMBB tăng đáng kể đối với đường cơ sở cân bằng. Trong mô hình Q-Learning có màu xanh đậm, thời gian truyền eMBB tăng chậm hơn so với đường cơ sở Cân bằng nhưng nhanh hơn đường cơ sở tập trung eMBB. Ngược lại, đường cơ sở tập trung eMBB có thời gian truyền URLLC cao đáng kể với chỉ 100 eMBB UE trong mạng.

Đường cơ sở cân bằng có thời gian truyền URLLC tĩnh giữa số lượng eMBB UE khác nhau. Mô hình Q-Learning có thể cung cấp thời gian truyền URLLC ngắn nhất. Kết quả này minh họa sự cân bằng giữa thông lượng eMBB và thời gian truyền URLLC. Các thanh màu xanh biểu thị thời gian truyền của đoạn thoại. Các mô hình nhìn chung có kết quả tương tự. Cuối cùng, tổng thời lượng trung bình có màu tím cho thấy đường cơ sở cân bằng tăng chậm theo thời gian. Đường cơ sở tập trung eMBB hoạt động kém nhất với ít nhu cầu eMBB hơn. Mô hình học Q lại cung cấp thời gian truyền ngắn nhất.

### 3.3 Kết luận chương 3

Trong đề án, một mô hình Q-Learning được đề xuất để phân bổ động băng thông đường trung gian được chia sẻ giữa các lát mạng với khái niệm phân chia chức năng O-RAN hiện có. Để chứng minh điều này, một trình mô phỏng dựa trên Python được phát triển để tùy chỉnh môi trường Q-Learning và mạng di động biên 5G. Chuyển động của UE với các tốc độ khác nhau. Để ngăn UE vượt quá giới hạn và tăng thêm độ phức tạp cho việc mô phỏng và đánh giá, một cấu trúc đơn giản được đề xuất với liên kết nhiều O-Cloud, tạo thành dạng tổ ong gồm 7 O-Cloud và 6 O-Cloud biên bao quanh một O-Cloud cục bộ. Trình mô phỏng được viết bằng Python, vì hầu hết các thư viện cho DL mới nhất cũng được phát triển bằng Python. Môi trường để chạy thuật toán Q-Learning tùy chỉnh được triển khai trong gói python OpenAI và gym. Mô hình đã tuân thủ chặt chẽ các quy tắc của MDP. Bảng Q chứa số lượng UE và phần trăm băng thông của các lát mạng dưới dạng trạng thái và sự thay đổi phần trăm băng thông dưới dạng hành động. Đáp ứng được xây dựng dựa trên các khái niệm về băng thông, mức sử dụng và ngưỡng tối thiểu. Bên cạnh đó, xét các yếu tố như tốc độ trung bình dữ liệu hay thời gian truyền cũng củng cố sự hiệu quả của việc sử dụng



## KẾT LUẬN VÀ KIẾN NGHỊ

Sau thời gian nghiên cứu và thực hiện đề án tốt nghiệp, các kết quả học viên đã đạt được như sau: Thứ nhất, giới thiệu tổng quan về lịch sử hình thành và phát triển của mạng truy nhập vô tuyến băng, sự ra đời của O-RAN như một lễ tất yếu phục vụ cho 5G khi các công nghệ ảo hóa thành phần và chức năng mạng ngày càng phát triển, hoàn thiện; Thứ hai, sự phát triển của trí tuệ nhân tạo với các thuật toán học máy, học sâu áp dụng trong các bài toán tối ưu, quản lý và giám sát tài nguyên mạng là một trong những điều rất được quan tâm; Thứ ba, đề án đã xây dựng một mô hình mạng O-RAN cơ bản để thực hiện mô phỏng trong hai trường hợp trung bình và cao điểm giúp thấy rõ khả năng phân bổ băng thông linh hoạt, đảm bảo cân bằng và tối ưu hơn khi có sự thay đổi số lượng thiết bị người dùng trong mạng.

Để thực hiện mô phỏng, đề án đã sử dụng Pycharm cùng các mã nguồn mở, các gói GYM và SymPy cũng như Open AI để hỗ trợ tạo môi trường mô phỏng đúng với yêu cầu đặt ra ban đầu. Bên cạnh đó là thư viện Matplotlib để thể hiện kết quả bằng hình ảnh trực quan thay cho các trường số liệu.

Hướng phát triển của đề án là việc sử dụng thêm nhiều các đường cơ sở khác để có thêm sự so sánh và đánh giá khả năng phân bổ băng thông có tối ưu hơn so với nhiều các phương pháp đang có hiện nay không. Bên cạnh đó, các yếu tố để đánh giá sự hiệu quả theo QoS hay QoE là rất nhiều nên có thể tích hợp cùng để xét trong nhiều các trường hợp phức tạp hơn... Với mã nguồn mở và dễ dàng trong cài đặt các gói hỗ trợ có thể mở rộng đề án trong các bài toán cho các tài nguyên khác hoặc kết hợp xét nhiều yếu tố trong cùng một kịch bản. Ngoài ra, sử dụng kết hợp nhiều các thuật toán khác nhau trong học máy cũng đem lại hiệu quả trong các quyết định cho hệ thống trong quá trình triển khai sau này.