

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ

BÁO CÁO THU HOẠCH HỌC PHẦN CÁC VẤN ĐỀ HIỆN ĐẠI CỦA
TRUYỀN THÔNG VÀ MẠNG MÁY TÍNH

Mô phỏng điều hướng lưu lượng trong O-RAN với xApp dùng Deep Reinforcement Learning

Họ và Tên:

Bùi Minh Thắng - 23020646

Nguyễn Vũ Minh - 23020629

Ma Đức Minh - 23020626

Nguyễn Hoàng Tùng Dương - 21020182

Người hướng dẫn:

TS. Nguyễn Ngọc Tân

CN. Nguyễn Thái Dương

Hà Nội, 2025

Lời cam đoan

Nhóm em xin cam đoan: Báo cáo nghiên cứu khoa học với đề tài “Mô phỏng điều hướng lưu lượng trong O-RAN với xApp dùng Deep Reinforcement Learning” này là của nhóm em. Những gì nhóm em viết ra không có sự sao chép từ các tài liệu, không sử dụng kết quả của người khác mà không trích dẫn cụ thể. Đây là công trình nghiên cứu tập thể nhóm em tự phát triển, không sao chép mã nguồn của người khác. Nếu vi phạm những điều trên, nhóm em xin chấp nhận tất cả những truy cứu về trách nhiệm theo quy định.

Sinh viên

Bùi Minh Thắng
Nguyễn Vũ Minh
Ma Đức Minh
Nguyễn Hoàng Tùng Dương

Lời cảm ơn

Lời đầu tiên, em xin được gửi lời cảm ơn chân thành tới Khoa Công nghệ Thông tin – Trường Đại học Công nghệ – Đại học Quốc gia Hà Nội đã tạo điều kiện thuận lợi để em được học tập, nghiên cứu và thực hiện đề tài này.

Em xin bày tỏ lòng biết ơn sâu sắc tới thầy Nguyễn Ngọc Tân và thầy Nguyễn Thái Dương đã tận tình hướng dẫn, hỗ trợ em trong suốt quá trình nghiên cứu và triển khai đề tài.

Bên cạnh đó, em xin được bày tỏ lòng biết ơn tới các thầy cô trong khoa đã tận tâm giảng dạy và trang bị cho em những kiến thức quý báu trong suốt quá trình học tập tại trường.

Cuối cùng, em xin chúc các thầy cô, các bạn luôn mạnh khỏe, hạnh phúc và gặt hái nhiều thành công trong cuộc sống.

Tóm tắt

Báo cáo này trình bày quá trình phát triển và mô phỏng một xApp trong kiến trúc Open RAN, tập trung vào việc điều hướng lưu lượng mạng sử dụng Deep Reinforcement Learning (DRL). Chúng tôi đã xây dựng một xApp để tối ưu hóa việc phân phối lưu lượng mạng trong môi trường O-RAN, sử dụng mô hình DRL để cải thiện hiệu suất mạng. Bằng cách sử dụng mô phỏng ns-3, chúng tôi đã kiểm tra hiệu quả của xApp trong các tình huống thực tế, cho thấy khả năng cải thiện đáng kể trong việc quản lý lưu lượng và tối ưu hóa tài nguyên mạng. Báo cáo cũng cung cấp hướng dẫn chi tiết về cách phát triển xApp và tích hợp nó vào hệ thống O-RAN.

Từ khoá: Open RAN architecture, O-RAN RIC, xApp development, O-RAN use cases, ns-3 O-RAN simulation, DRL for traffic steering, RAN Intelligent Controller.

Mục lục

1	Đặt vấn đề	5
2	Kiến thức cơ sở	6
2.1	So sánh kiến trúc RAN truyền thống và O-RAN	6
2.2	Chức năng của Near-Real-Time RIC và Non-Real-Time RIC	7
2.3	Giao diện mở (E2, A1, O1) và vai trò của chúng	8
3	Phương pháp tối ưu hoá ABR	9
4	Cài đặt phương pháp và thực nghiệm	10
5	Kết luận	11
	Tài liệu tham khảo	12

Chương 1

Đặt vấn đề

Chương 2

Kiến thức cơ sở

2.1 So sánh kiến trúc RAN truyền thống và O-RAN

Các mạng di động truyền thống trong lịch sử chủ yếu áp dụng kiến trúc RAN phân tán (Distributed RAN - D-RAN), trong đó mỗi trạm gốc tích hợp cả chức năng xử lý băng tần gốc (baseband) và chức năng vô tuyến. Mô hình này phù hợp với các thế hệ di động đầu tiên (2G/3G/4G), nhưng tồn tại nhiều hạn chế về khả năng mở rộng, tính tương tác giữa các nhà cung cấp và khả năng quản lý tập trung. Với sự ra đời của 5G và tương lai là 6G, ngành công nghiệp viễn thông đang chuyển dịch sang các giải pháp linh hoạt và định nghĩa bằng phần mềm. Open RAN (O-RAN) đưa ra một kiến trúc mở, mô-đun và tách rời bằng cách phân tách các chức năng RAN thành ba thành phần riêng biệt: CU, DU và RU, cùng với các giao diện tiêu chuẩn hóa [6]. Bài viết này phân tích các điểm khác biệt về kiến trúc và những hàm ý giữa D-RAN và O-RAN.

Trong D-RAN, mỗi trạm thu phát (cell site) chứa đồng thời một Bộ xử lý băng tần gốc (BBU) và một Bộ vô tuyến từ xa (RRU). BBU thực hiện toàn bộ xử lý tầng baseband, bao gồm PHY, MAC, RLC và các tầng giao thức cao hơn, trong khi RRU đảm nhận các chức năng đầu vào/ra RF tương tự. Sự tích hợp này giúp đơn giản hóa việc đồng bộ hóa và giảm độ trễ đường truyền fronthaul, nhưng lại dẫn đến hệ thống cứng nhắc và phụ thuộc vào nhà cung cấp. Do phần cứng và phần mềm gắn kết chặt chẽ, việc mở rộng hoặc nâng cấp hệ thống thường đòi hỏi phải thay thế toàn bộ các thành phần độc quyền. Ngoài ra, khả năng tối ưu hóa mạng và chia sẻ tài nguyên giữa các trạm bị hạn chế do xử lý vẫn diễn ra độc lập tại từng vị trí [1].

O-RAN khác biệt với cách tiếp cận nguyên khối của D-RAN bằng việc phân tách chức năng rõ ràng giữa RU, DU và CU. RU đảm nhận xử lý PHY thấp và RF, DU chịu trách nhiệm xử lý PHY cao, MAC và RLC, trong khi CU điều khiển các tầng SDAP, PDCP và RRC. Các thành phần này giao tiếp thông qua các giao diện mở: giao diện F1 giữa CU và DU, và giao diện fronthaul mở (thường là chia tách 7.2x) giữa DU và RU [4]. Tính mô-đun này cho phép các thiết bị từ nhiều nhà cung cấp khác nhau tương tác, đồng thời hỗ trợ ảo hóa các chức năng CU/DU trên phần cứng thương mại thông dụng [3].

Một điểm khác biệt lớn giữa D-RAN và O-RAN nằm ở tính linh hoạt khi triển khai. Trong khi D-RAN đồng vị tất cả xử lý tại trạm thu phát, thì O-RAN hỗ trợ việc tập trung hóa CU tại các trung tâm dữ liệu khu vực và phân phối DU gần biên mạng. Sự phân tách này cho phép phân bổ tài nguyên động, giảm chi phí đầu tư (CAPEX) và nâng cao khả năng mở rộng. Hơn nữa, O-RAN còn hỗ trợ các chức năng tiên tiến như bộ điều khiển RAN thông minh (RIC), cho phép tối ưu hóa dựa trên AI/ML gần như theo thời gian thực [6].

Tuy có nhiều ưu điểm, O-RAN cũng đặt ra thách thức trong việc tích hợp do môi trường đa nhà cung cấp và yêu cầu nghiêm ngặt về đường truyền fronthaul. Các liên kết có tốc độ cao và độ trễ thấp là điều kiện bắt buộc để đảm bảo hiệu năng trong kiến trúc chia tách. Ngược lại, D-RAN đơn giản hơn trong triển khai và đã được kiểm chứng về độ tin cậy, nhưng thiếu sự linh hoạt và tính mở cần thiết cho các trường hợp sử dụng mới như chia sẻ mạng (network slicing)

và mạng 5G riêng (private 5G) [3].

Kiến trúc D-RAN truyền thống đã đóng vai trò là nền tảng của mạng di động trong nhiều thập kỷ, mang lại sự đơn giản và độ ổn định cho các thể hệ mạng di động trước. Tuy nhiên, nhu cầu ngày càng tăng về tính linh hoạt, hiệu quả chi phí và phân biệt dịch vụ đang thúc đẩy quá trình chuyển đổi sang hệ thống mở và tách rời. O-RAN, với việc phân tách chức năng và các giao diện tiêu chuẩn, đại diện cho một sự chuyển đổi mô hình hướng tới giải pháp RAN gốc đám mây và trung lập với nhà cung cấp. Mặc dù vẫn còn tồn tại các thách thức về tích hợp và tối ưu hóa hiệu năng, nhưng những lợi ích tiềm năng của O-RAN đang định vị nó như một yếu tố then chốt trong việc phát triển mạng di động tương lai [4].

2.2 Chức năng của Near-Real-Time RIC và Non-Real-Time RIC

Trong kiến trúc mạng truy cập vô tuyến mở (Open RAN), Bộ điều khiển RAN thông minh (RAN Intelligent Controller – RIC) được phân tách thành hai phần: RIC thời gian gần thực (Near-Real-Time RIC) và RIC thời gian không thực (Non-Real-Time RIC). Hai thành phần này phối hợp điều khiển và tối ưu mạng ở những quy mô thời gian khác nhau nhằm nâng cao hiệu năng của RAN. Cụ thể, Near-RT RIC đảm nhiệm việc điều khiển RAN với độ trễ thấp (từ khoảng 10 mili-giây đến <1 giây) [5], còn Non-RT RIC phụ trách các tác vụ ở quy mô thời gian dài hơn (>1 giây, thường tính bằng giây, phút hoặc lâu hơn) [5]. Sự phân chia này cho phép tối ưu mạng ở cả thời gian thực ngắn hạn lẫn hoạch định dài hạn, tạo nên hệ thống điều khiển nhiều tầng cho RAN.

Near-Real-Time RIC (Near-RT RIC): Đây là thành phần RIC hoạt động gần thời gian thực, thường được triển khai trên hạ tầng điện toán biên hoặc cụm mạng khu vực gần với các nút RAN. Near-RT RIC có chức năng thu thập thông tin trạng thái mạng và thực thi các hành động điều khiển nhanh lên mạng vô tuyến với độ trễ yêu cầu dưới 1 giây [5]. Theo đặc tả O-RAN, Near-RT RIC là một chức năng logic cho phép điều khiển và tối ưu tài nguyên RAN ở mức độ nhanh, thông qua việc thu thập dữ liệu chi tiết và tác động hành động lên các nút RAN qua giao diện E2 docs.o-ran-sc.org. Near-RT RIC thường xử lý các tác vụ như điều khiển truy cập vô tuyến và tài nguyên vô tuyến theo thời gian thực gần, ví dụ: điều phối lịch truyền dẫn, cân bằng tải giữa các cell, điều chỉnh tham số handover, điều khiển can nhiễu,... nhằm tối ưu hiệu suất thông lượng và chất lượng dịch vụ tức thời cho người dùng. Thành phần này tương tác trực tiếp với các nút mạng RAN (như O-DU, O-CU) qua giao diện E2 để nhận số liệu tình trạng (telemetry) và gửi chỉ thị điều khiển một cách liên tục. Đặc điểm quan trọng của Near-RT RIC là khả năng mở rộng chức năng qua các xApp – những ứng dụng plug-and-play chạy trên nền tảng Near-RT RIC để thực hiện các thuật toán điều khiển radio chuyên biệt [5]. Near-RT RIC cũng có cơ chế phối hợp và tránh xung đột giữa nhiều xApp khác nhau cùng tác động lên RAN (ví dụ: cơ chế quản lý message bus, lớp dữ liệu chia sẻ, và logic phân giải xung đột) để đảm bảo các quyết định điều khiển không mâu thuẫn [5].

Non-Real-Time RIC (Non-RT RIC): Đây là thành phần RIC hoạt động ngoài thời gian thực chặt chẽ, nằm trong khối dịch vụ quản lý và điều hành (Service Management and Orchestration – SMO) ở trung tâm mạng hoặc đám mây. Non-RT RIC chịu trách nhiệm thực hiện các tác vụ quản lý, tối ưu RAN ở quy mô dài hạn hơn (trên 1 giây) [5], bao gồm quản lý chính sách dịch vụ, phân tích hiệu năng, tối ưu cấu hình và hoạch định tài nguyên chiến lược cho mạng. Theo O-RAN Alliance, Non-RT RIC là một chức năng logic trong SMO hỗ trợ điều khiển/tối ưu RAN phi-thời-gian-thực, cung cấp khung AI/ML để huấn luyện và cập nhật mô hình, và truyền tải các hướng dẫn chính sách tới RIC gần thực [2] [7]. Non-RT RIC được cấu thành bởi framework Non-RT RIC (nền tảng) và các ứng dụng rApp (các ứng dụng chạy trên Non-RT RIC). Nền tảng Non-RT RIC thực hiện việc kết thúc (terminate) giao diện A1 với Near-RT RIC, đồng thời

phoi bày dịch vụ quản lý dữ liệu và ML cho các rApp thông qua giao diện nội bộ R1 [7]. Các rApp (RAN applications) là những ứng dụng mô-đun chạy trên Non-RT RIC, sử dụng các dịch vụ mà nền tảng cung cấp để tạo ra các giá trị gia tăng cho vận hành RAN [7]. Nhiệm vụ của rApp rất đa dạng, bao gồm: đề xuất và điều chỉnh chính sách điều khiển RAN, phân tích dữ liệu hiệu năng dài hạn, tối ưu cấu hình tham số, cũng như cung cấp thông tin bổ sung (enrichment information) cho các ứng dụng khác [7]. Non-RT RIC gửi hướng dẫn chính sách và mục tiêu đến Near-RT RIC thông qua giao diện A1 (ví dụ: chính sách về phân bổ tài nguyên, mục tiêu QoS cần đạt, tham số ngưỡng sự kiện, v.v.), nhờ đó ảnh hưởng gián tiếp đến hành vi của các xApp trên Near-RT RIC [7]. Ngược lại, Non-RT RIC cũng thu thập phản hồi từ mạng (thông qua dữ liệu O1 hoặc qua báo cáo từ Near-RT RIC) để đánh giá và điều chỉnh các chiến lược tối ưu. Có thể xem Non-RT RIC như “bộ não” ở tầng trên, vạch ra chiến lược dài hạn cho mạng, trong khi Near-RT RIC là “cánh tay tác động” ở tầng dưới thực thi các điều chỉnh nhanh theo chiến lược đó.

2.3 Giao diện mở (E2, A1, O1) và vai trò của chúng

Chương 3

Phương pháp tối ưu hoá ABR

Chương 4

Cài đặt phương pháp và thực nghiệm

Chương 5

Kết luận

Tài liệu tham khảo

- [1] 3GPP. *Study on new radio access technology: Radio access architecture and interfaces*. 2017. URL: <https://www.3gpp.org/DynaReport/38801.htm>.
- [2] O-RAN Software Community. *O-RAN Architecture Overview*. 2025. URL: <https://docs.o-ran-sc.org/en/latest/architecture/architecture.html>.
- [3] Ericsson. *Open RAN: Flexibility and Interoperability in 5G Networks*. White Paper. 2023. URL: <https://www.ericsson.com/en/reports-and-papers/white-papers/open-ran>.
- [4] A. Khurshid **and others**. “Disaggregated RAN: O-RAN Functional Split and Deployment Insights”. **in** *IEEE Communications Magazine*: 62.2 (2024), **pages** 112–118.
- [5] Michele Polese, Leonardo Bonati, Salvatore D’Oro, Stefano Basagni, Tommaso Melodia. “Understanding O-RAN: Architecture, Interfaces, Algorithms, Security, and Research Challenges”. **in** (2022): URL: <https://scispace.com/pdf/understanding-o-ran-architecture-interfaces-algorithms-2jhd9ql.pdf#:~:text=The%20O,standardizing%20a%20virtualization%20platform%20for>.
- [6] O-RAN Alliance. *O-RAN Architecture Description*. O-RAN WG6 White Paper. 2020. URL: <https://www.o-ran.org/specifications>.
- [7] WG1. “O-RAN Architecture Description”. **in** (2024): URL: https://www.etsi.org/deliver/etsi_ts/103900_103999/103982/08.00.00_60/ts_103982v080000p.pdf.