

PERFORMANCE IMPROVEMENT OF EDGE USERS BY UTILIZING JOINT SCHEDULING AND DEEP Q NETWORK

Luu Bach Hung, Lam Sinh Cong*, Nguyen Nam Hoang

VNU – University of Engineering and Technology

ARTICLE INFO		ABSTRACT
Received:	13/02/2023	Improving the user performance, especially for users at cell edge areas or edge users, is one of the most important requirements of 5G and 6G cellular networks which have a high densification of base stations. This paper presents the analysis and optimization of edge user performance under the dual-slope path loss model and Rayleigh fading by utilizing advanced techniques, called Joint Scheduling (JC), and Deep Q Network (DQN) utilizing Q-learning. The dual-slope path loss model is widely used to analyze the network performance of 5G and beyond since it is able to capture the attenuation properties of the real wireless transmission environment. In a JC system, a user can proactively select the base station with the highest downlink signal to communicate with. Meanwhile, the DQN dynamically adjusts the base station transmission power to adapt to change in the wireless environment. The analytical results indicate that the deployment of JC with DQN can increase the edge user performance up to 35% or 200% more than that of the system which applies DQN only or Maximum power (MP) respectively.
Revised:	27/4/2023	
Published:	28/4/2023	
KEYWORDS		
Edge user		
5G		
6G		
Joint Scheduling		
Deep Q Network		

NÂNG CAO HIỆU NĂNG NGƯỜI DÙNG BIÊN BẰNG KỸ THUẬT LẬP LỊCH ĐỒNG THỜI VÀ HỌC SÂU

Luu Bách Hưng, Lâm Sinh Công*, Nguyễn Nam Hoàng

Trường Đại học Công nghệ - ĐH Quốc gia Hà Nội

THÔNG TIN BÀI BÁO		TÓM TẮT
Ngày nhận bài:	13/02/2023	Nâng cao hiệu năng người dùng, đặc biệt là những người dùng ở biên của vùng phủ sóng, là một trong những yêu cầu cấp thiết của mạng di động có mật độ trạm phát dày đặc như mạng 5G, 6G. Do đó, bài báo này tập trung phân tích và tối ưu hiệu năng của người dùng biên trong môi trường suy hao dual – slope và ảnh hưởng của Rayleigh pha đỉnh bằng cách sử dụng kỹ thuật tiên tiến của di động là lập lịch đồng thời (JC) và mạng học sâu (DQN) sử dụng thuật toán Q-learning. Mô hình suy hao dual-slope được sử dụng rộng rãi để đánh giá hiệu năng mạng vì nó có khả năng mô tả đặc điểm suy hao công suất của môi trường vô tuyến thực tế. Trong hệ thống sử dụng kỹ thuật JC, người dùng chủ động lựa chọn trạm có tín hiệu mạnh nhất để kết nối. Việc sử dụng mạng DQN cho phép các trạm cơ sở lựa chọn công suất phát phù hợp với môi trường. Các kết quả phân tích chỉ ra rằng việc sử dụng kỹ thuật lập lịch đồng thời kết hợp với DQN có thể nâng cao hiệu năng của người dùng biên tương ứng lên 35% và 200% so với trường hợp chỉ sử dụng DQN và thuật toán công suất tối đa (MP).
Ngày hoàn thiện:	27/4/2023	
Ngày đăng:	28/4/2023	
TỪ KHÓA		
Người dùng biên		
5G		
6G		
Lập lịch đồng thời		
Học sâu		

DOI: <https://doi.org/10.34238/tnu-jst.7316>

* Corresponding author. Email: congls@vnu.edu.vn

1. Giới thiệu

Mạng thông tin di động thế hệ thứ 5 (5G) đang được triển khai tại nhiều nước trên thế giới với kỳ vọng cung cấp các loại dịch vụ băng thông rộng, thời gian thực và độ ổn định cao, đồng thời cung cấp lượng kết nối lớn cho người dùng [1]. Để đạt được các kỳ vọng trên, mạng 5G được xây dựng với mật độ trạm dày đặc và vùng phủ sóng của một trạm di động cũng nhỏ hơn so với các thế hệ mạng trước đó [2]. Khi đó, xác suất người dùng nằm tại vị trí biên hay vùng giáp ranh (gọi là người dùng biên) giữa các trạm là rất lớn. Tại các vị trí này, tín hiệu từ các trạm phục vụ tương đối yếu nhưng tín hiệu nhiễu từ các trạm lân cận lại lớn. Do đó, hiệu năng của người dùng biên thường rất thấp.

Việc đánh giá và nâng cao hiệu năng của người dùng biên đã được quan tâm trong rất nhiều các công trình nghiên cứu trước đây. Cụ thể, các nghiên cứu thường tập trung đánh giá ảnh hưởng của các kỹ thuật tiên tiến lên hiệu năng của người dùng biên. Một trong các kỹ thuật đó là kỹ thuật truyền thông hợp tác đa điểm (Coordinated Multipoint Point Processing (CoMP)) [3]. Kỹ thuật này cho phép người dùng đồng thời kết nối tới nhiều hơn 1 trạm phát. Trong các tài liệu của 3GPP, kỹ thuật CoMP có thể được phân chia thành 3 loại: Truyền đồng thời (Joint Transmission), Lập lịch đồng thời (Joint Scheduling) và Lập lịch Beamforming đồng thời giữa các trạm (Joint Beamforming). Kỹ thuật lập lịch theo Beamforming được coi là phức tạp nhất vì cần có sự kết hợp và đồng bộ giữa antenna của các trạm phát. Kỹ thuật truyền đồng thời cho phép người dùng nhận dữ liệu cùng lúc từ một số trạm lân cận nhưng kỹ thuật này yêu cầu sự đồng bộ cũng như tính toán và sự phức tạp cao. Trong khi đó, kỹ thuật lập lịch đồng thời hoàn toàn có thể do người dùng chủ động bằng cách dò tìm trạm phát có tín hiệu mạnh nhất. Đây là kỹ thuật có khả năng thực thi cao hơn cả. Do đó, bài báo này tập trung nghiên cứu và đánh giá ảnh hưởng của kỹ thuật lập lịch đồng thời lên người dùng biên.

Việc nghiên cứu ảnh hưởng của kỹ thuật lập lịch đồng thời đối với người dùng biên đã được xem xét trong [4]. Kết quả phân tích và mô phỏng cho thấy rằng kỹ thuật lập lịch đồng thời có thể cải thiện hiệu năng người dùng có điều kiện xấu nhất, là người dùng tại vị trí giao của các trạm. Tuy nhiên, bài báo này sử dụng mô hình suy hao truyền thống (single-slope) $L = r^{-\alpha}$, trong đó r là khoảng cách truyền tin, α là hệ số suy hao. Trong hệ thống thông tin di động sử dụng tần số cao như 5G và 6G, công suất tín hiệu nhận của người dùng biến thiên một cách ngẫu nhiên và phụ thuộc vào rất nhiều yếu tố khác nhau như đặc tính của môi trường truyền tin, mật độ của các vật cản. 3GPP khuyến nghị sử dụng mô hình suy hao dual-slope, trong đó xác định xác suất đường truyền giữa người dùng là đường truyền trong tầm nhìn thẳng LoS (Line-of-Sight) hay đường truyền không trong tầm nhìn thẳng (non Line-of-Sight) [5]. Với mỗi trường hợp, 3GPP đều đưa ra biểu thức tính suy hao đường truyền tương ứng.

Bên cạnh kỹ thuật lập lịch đồng thời, bài báo còn đề xuất áp dụng phương pháp mạng học sâu sử dụng thuật toán Q-learning để tối ưu công suất của các trạm phát, từ đó nâng cao hiệu năng của người dùng. Việc sử dụng thuật toán Q-learning để tối ưu các bài toán trong mạng viễn thông đã trở nên phổ biến nhờ sự thích nghi của thuật toán đối với tính ngẫu nhiên của mạng [6] - [8]. Cụ thể, trong nghiên cứu của mình, nhóm tác giả Zhong và các cộng sự đã xây dựng thuật toán lập lịch phù hợp với các loại hình dịch vụ đa phương tiện của 5G. Trong [6], thuật toán Q-learning đã được sử dụng để tối ưu công suất của người dùng được phân bố ngẫu nhiên trong mạng. Các kết quả mô phỏng cho thấy rằng thuật toán Q-learning có thể cải thiện đáng kể hiệu năng của người dùng.

Tuy mô hình dual-slope đã được 3GPP khuyến nghị để mô hình kênh truyền trong mạng di động, việc sử dụng mô hình này để đánh giá hiệu năng của người dùng biên còn rất hạn chế. Cụ thể, các công trình nghiên cứu thường sử dụng mô hình suy hao truyền thống $L = r^{-\alpha}$ hay mô hình suy hao Stretched Path Loss model [4], [9], [10]. Bên cạnh đó, các bài toán tối ưu hiệu năng người dùng chủ yếu xem xét người dùng có vị trí ngẫu nhiên trong mạng, và bỏ qua người dùng có vị trí cố định tại biên. Những điểm trên đã thúc đẩy nhóm tác giả tập trung nghiên cứu và tối

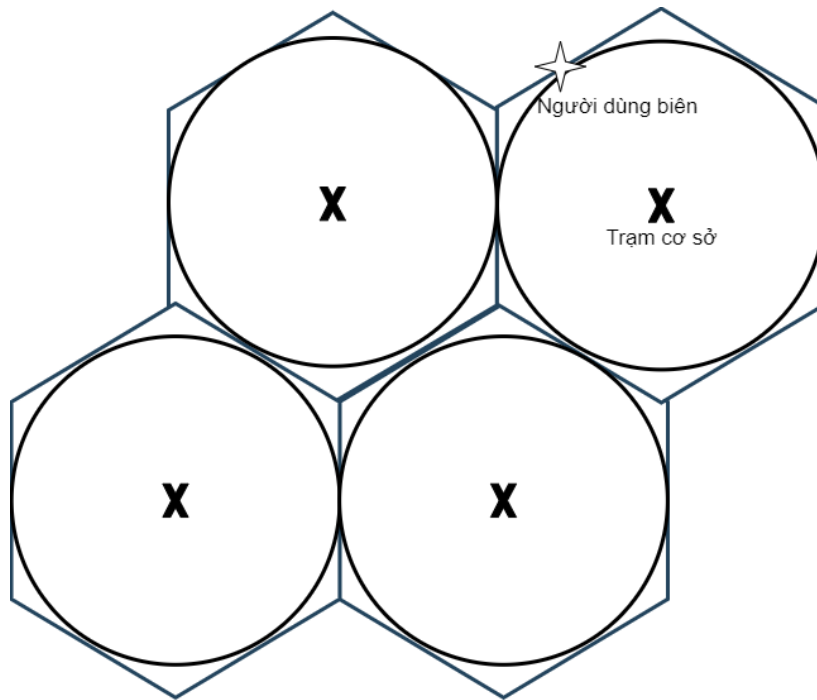
ưu hiệu năng của người dùng biên sử dụng mô hình suy hao dual-slope bằng kỹ thuật lập lịch đồng thời và Q-learning. Nghiên cứu này được thực hiện nhằm các mục tiêu sau:

- Trình bày các kết quả phân tích và đánh giá hiệu năng của người dùng biên bằng việc sử dụng mô hình dual-slope theo khuyến nghị của 3GPP.
- Đề xuất việc kết hợp kỹ thuật lập lịch đồng thời với mạng học sâu sử dụng thuật toán Q-learning để nâng cao hiệu năng người dùng biên.
- So sánh hiệu năng của phương pháp đề xuất với các phương pháp khác để cho thấy rằng việc áp dụng hai kỹ thuật nêu trên là hướng khả thi để tăng dung lượng người dùng biên.

2. Phương pháp nghiên cứu

2.1. Mô hình hệ thống

Bài báo nghiên cứu mạng thông tin di động có vùng phủ sóng được phân chia thành N tế bào lục giác đồng nhất. Trong đó, mỗi tế bào được phục vụ bởi 1 trạm phát đặt tại trung tâm của tế bào như Hình 1. Người dùng được giả thiết phân bố một cách ngẫu nhiên và được coi là người dùng biên nếu người dùng này nằm trên rìa phủ sóng của trạm phát.



Hình 1. Mô hình mạng di động gồm 4 trạm cơ sở

Tín hiệu vô tuyến được giả thiết chịu ảnh hưởng bởi cả pha-đỉnh nhanh (fast fading) và pha-đỉnh chậm (slow fading). Trong đó, pha-đỉnh nhanh được mô hình hóa bằng phân bố Rayleigh với độ lợi công suất của kênh truyền tuân theo phân bố mũ. Giả thiết phân bố của độ lợi công suất là $f(g) = e^{-g}$. Pha-đỉnh chậm được đặc trưng thông qua suy hao công suất và được tính theo mô hình dual-slope. Cụ thể, xác suất đường truyền giữa người dùng và trạm phát k là đường truyền trong tầm nhìn thẳng (LoS) và đường truyền không trong tầm nhìn thẳng (nLoS) tương ứng là p_k^l và $1 - p_k^l$. Đối với trường hợp LoS thì công suất suy hao theo khoảng cách d là $p_k^l(r_k)$ và đối với trường hợp nLoS là p_k^{nl} . Khi đó, công suất suy hao trung bình của tín hiệu khi khoảng cách giữa người dùng và trạm phát ở vị trí r_k được tính như sau:

$$L(r_k) = p_k^l p_k^l(r_k) + (1 - p_k^l) p_k^{nl}(r_k) \quad (1)$$

2.1.1. Hệ thống không sử dụng kỹ thuật Lập lịch đồng thời (JC)

Trong hệ thống không sử dụng kỹ thuật lập lịch đồng thời, mỗi người dùng chỉ nhận tín hiệu từ một trạm phát duy nhất. Khi đó, tín hiệu đường xuống tại người dùng ở khoảng cách r từ trạm phục vụ được tính bằng công thức sau:

$$S = P g L(r) \quad (2)$$

trong đó P là công suất của trạm phục vụ.

Bên cạnh nhiễu Gaussian với công suất σ^2 , người dùng trong mạng thông tin di động thường phải chịu ảnh hưởng nhiễu có công suất I từ các trạm lân cận. Trong trường hợp xấu nhất, mỗi người dùng chịu ảnh hưởng từ tất cả các trạm lân cận. Khi đó, tổng công suất nhiễu mỗi người dùng phải chịu là:

$$I = \sum_{k=1}^{N-1} P_k g_k L(r_k) \quad (3)$$

trong đó P_k là công suất của trạm nhiễu thứ k , g_k và $L(r_k)$ lần lượt là độ lợi công suất và suy hao công suất của đường truyền từ người dùng tới trạm nhiễu k .

Khi đó, tỉ số Tín hiệu - Nhiễu cộng Ôn (SINR) được tính như sau:

$$\text{SINR} = \frac{P g L(r)}{\sum_{k=1}^{N-1} P_k g_k L(r_k) + \sigma^2} \quad (4)$$

2.1.2. Hệ thống sử dụng kỹ thuật Lập lịch đồng thời (JC)

Trong hệ thống mạng thông tin di động sử dụng kỹ thuật truyền thông hợp tác, cụ thể là kỹ thuật lập lịch đồng thời (Joint Scheduling), một người dùng có thể kết nối với nhiều hơn một trạm phát nhưng tại một thời điểm người dùng đó chỉ nhận tín hiệu từ trạm phát có tín hiệu tốt nhất.

Giả sử người dùng được kết nối với M trạm phát cùng lúc. Khi đó, trạm m được coi là trạm phục vụ của người dùng nếu nó thỏa mãn điều kiện sau:

$$P_m g_m L(r_m) = \max_{k=1, \dots, M} P_k g_k L(r_k) \quad (5)$$

Trong nội dung bài báo này, giả thiết rằng khi trạm m được chỉ định phục vụ 1 người dùng nào đó thì các trạm phát khác trong nhóm M trạm liên kết hoàn toàn có thể phục vụ các người dùng khác và gây nhiễu lên chính người dùng đang xét. Khi đó, công suất nhiễu tại người dùng được tính bằng tổng công suất của tất cả tín hiệu tại người dùng đó trừ đi công suất của tín hiệu phục vụ. Cụ thể là:

$$I = \sum_{n=1}^N P_n g_n L(r_n) - \max_{n=1, \dots, M} P_k g_k L(r_k) \quad (6)$$

Khi đó tỉ số SINR được tính như sau:

$$\text{SINR} = \frac{\max_{n=1, \dots, M} P_k g_k L(r_k)}{\left(\sum_{n=1}^N P_n g_n L(r_n) - \max_{n=1, \dots, M} P_k g_k L(r_k) \right) + \sigma^2} \quad (7)$$

2.2. Dung lượng người dùng

Theo định lý Shannon, dung lượng của người dùng với băng thông chuẩn hóa được tính bằng công thức sau:

$$C = \log_2(1 + \text{SINR}) \quad (8)$$

Từ biểu thức 4 và biểu thức 8, ta thấy dung lượng của người dùng là hàm số của công suất phát gồm cả trạm nhiễu và trạm phục vụ, số lượng trạm nhiễu và khoảng cách giữa người dùng

tới các trạm. Trong mô hình mạng 5G, các trạm phát có thể trao đổi thông tin với nhau thông qua giao diện vô tuyến Xn [11]. Tuy nhiên, việc phân cấp nguồn tài nguyên diễn ra thường xuyên theo một hoặc một vài khe thời gian, nên việc trao đổi thông tin về cấp nguồn tài nguyên hoàn toàn có thể gây ra quá tải trên giao diện vô tuyến Xn. Do đó, các trạm thường cấp nguồn tài nguyên một cách độc lập với nhau. Hay nói cách khác, khó có thể thay đổi số lượng trạm nhiều của người dùng. Ngoài ra, khoảng cách người dùng tới các trạm phát sóng cũng không thể thay đổi. Do đó, điều chỉnh công suất trạm phát là một phương án tối ưu để nâng cao dung lượng người dùng.

Vì vậy, bài báo này tập trung xây dựng bài toán tối ưu công suất của trạm phát để nâng cao hiệu năng của người dùng. Theo khuyến nghị của 3GPP, công suất của trạm cơ sở có thể dao động trong 1 khoảng nhất định từ (P_{min}, P_{max}) [11]. Khi đó, bài toán tối ưu công suất được mô tả như sau:

$$\max_P \sum_n C_n \quad (9)$$

$$s.t. P_{min} \leq P \leq P_{max}$$

Trong bài báo này, mạng học sâu sử dụng thuật toán Q-learning sẽ được sử dụng để tối ưu công suất phát của các trạm cơ sở.

2.3. Mạng học sâu sử dụng thuật toán Q-learning

Học sâu sử dụng thuật toán Q-learning là một trong những thuật toán học tăng cường phổ biến có bản chất là thử đi thử lại và rút ra kinh nghiệm sau mỗi lần như vậy. Một trong những lợi thế của thuật toán Q-learning là nó có thể học 1 giá trị của một hành động nào đó mà không cần yêu cầu một mô hình cụ thể [7]. Các bước chính của thuật toán Q-learning được mô tả như sau:

- **Trạng thái - State:** Tại bước này, các trạng thái ban đầu của mạng, cụ thể là số lượng và vị trí người dùng, chất lượng kênh truyền, vị trí của các trạm và giới hạn công suất của chúng được thiết lập. Giả thiết rằng, các trạm cơ sở có khả năng xác định tương đối chính xác vị trí của người dùng thông qua các bản tin tham chiếu (Positioning Reference Signal) [12]. Bên cạnh đó, các thông số của môi trường suy hao và pha-định nhanh cũng được xác định và lưu trữ tại các trạm cơ sở.
- **Hành động - Action:** Theo các khuyến nghị của 3GPP [13], trạm cơ sở có thể lựa chọn bất kỳ mức công suất nào trong dải công suất đã được định trước (P_{min}, P_{max}). Trong thuật toán Q-learning, dải công suất này sẽ được chia thành nhiều mức khác nhau. Mỗi bước của thuật toán Q-learning sẽ lựa chọn giá trị công suất phù hợp để gán cho từng trạm.
- **Giải thưởng - Reward:** Bài toán tập trung tối ưu dung lượng của người dùng nên hàm tính dung lượng trong Công thức 8 được dùng làm hàm giải thưởng cho thuật toán Q-learning.

3. Kết quả mô phỏng

Nội dung phần này tập trung đánh giá sự ổn định của thuật toán Q-learning và so sánh ảnh hưởng của thuật toán Q-learning và kỹ thuật JC lên hiệu năng người dùng biên. Thuật toán công suất cực đại (Maximum Power - MP), cho phép tất cả các trạm phát hoạt động với công suất tối đa, được sử dụng làm thuật toán tham chiếu. Đối với hệ thống sử dụng DQN và thuật toán MP, giả thiết rằng người dùng kết nối tới trạm gần nhất. Đây cũng là một phương án khả thi trong mạng 5G do vị trí của người dùng hoàn toàn có thể được xác định một cách gần đúng thông qua các tín hiệu tham chiếu vị trí đường xuống (Downlink Positioning Reference Signal) [12]. Trong khi đó, với hệ thống sử dụng JC - DQN, người dùng sẽ chủ động lựa chọn trạm phát cung cấp tín hiệu có công suất lớn nhất để kết nối.

3.1. Mô hình mô phỏng

Để đánh giá ảnh hưởng của thuật toán DQN và kỹ thuật lập lịch đồng thời, trong phần này mô hình mạng lục giác gồm 25 trạm phát sóng được mô phỏng và đánh giá. Người dùng được phân bố một cách ngẫu nhiên trên đường biên của vùng phủ sóng. Trong đó, tất cả các trạm phát có bán kính phủ sóng là 250 m và sử dụng chung băng tần. Do tính độc lập của các trạm phát trong việc cấp phát băng thông cho người dùng, tất cả các trạm có thể phát trên cùng một băng tần tại cùng một thời điểm. Do đó, chương trình mô phỏng giả thiết mỗi người dùng chịu ảnh hưởng nhiều từ tất cả các trạm lân cận. Khi đó, SINR và tốc độ tối đa của người dùng được tính theo Công thức 8 và 9. Thông thường, mỗi người dùng được phục vụ trên dải tần riêng. Trong mô phỏng, chúng tôi chỉ xem xét một đơn vị băng thông cho một người dùng.

Tín hiệu truyền trong mạng được giả thiết chịu ảnh hưởng của cả pha-đỉnh nhanh và pha-đỉnh chậm (suy hao công suất). Trong đó, pha-đỉnh nhanh có hàm phân bố Rayleigh, suy hao công suất theo mô hình dual-slope. Theo 3GPP [5], xác suất đường truyền giữa người dùng và trạm phát sóng là đường truyền trong tầm nhìn thẳng là:

$$p^l = \begin{cases} 1 & \text{nếu } d < 2,5m \\ e^{-\frac{d-5}{20,8}} & \text{nếu } 5m \leq d \leq 49m \\ e^{-\frac{d-49}{211}} & \text{nếu } 49m \leq d \end{cases} \quad (10)$$

Xác suất đường truyền trong tầm nhìn không thẳng là $p^{nl} = 1 - p^l$. Công suất suy hao đối với đường truyền trong tầm nhìn thẳng và không trong tầm nhìn thẳng tương ứng như sau:

$$P^l(d) = 32,4 + 17,3 \log(3D) + 20 \log_{10}(f_c) \\ P^{nl}(d) = \max(P^l(d), P^{l'}(d)) \quad (11)$$

Trong đó $P^{l'} = 38,3 \log(3D) + 17,3 + 24,9 \log_{10}(f_c)$. Các tham số mô phỏng khác được mô tả bởi bảng I. Trong đó, số episode thể hiện số lượng các trạng thái và hành động của thuật toán. Khi episode đủ lớn, thuật toán Q-learning sẽ có tính hội tụ và ổn định. Trong mỗi trạng thái, giả thiết rằng pha-đỉnh là bất biến và quá trình truyền tin được thực hiện trong 100 khe thời gian - mỗi khe thời gian có độ lớn là 20 ms. Khi số lượng khe thời gian lớn thì quá trình truyền tin diễn ra lâu hơn, khi đó quá trình học sẽ được diễn ra dài hơn. Do đó, tương tự như số episode, số lượng khe thời gian lớn thì kết quả cũng ổn định hơn.

Bảng 1. Tham số DQN

Tham số mô phỏng	Giá trị
Số episode	5000
Số test_episode	500
Số khe thời gian/episode	100
Thời gian của một khe thời gian	20 ms
Tốc độ học ζ	0,001
Hệ số chiết khấu (discount factor)	0

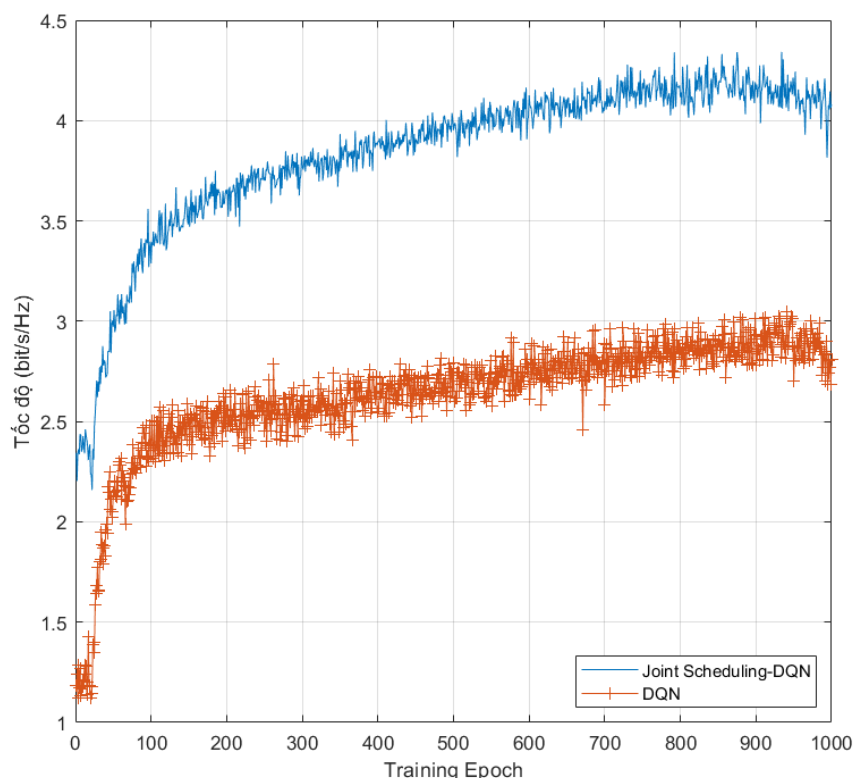
Thuật toán tối ưu Adam được sử dụng trong Q-learning với tốc độ học là ζ . Khi giá trị ζ càng nhỏ thì thuật toán càng tối ưu. Tuy nhiên, điều này cũng kéo theo thời gian thực thi của chương trình cũng tăng theo. Do đó, bài báo này sử dụng giá trị ζ được khuyến nghị chung, đủ để đảm bảo tính ổn định của thuật toán Q-learning là $\zeta = 0,001$.

3.2. Đánh giá hiệu năng theo dữ liệu huấn luyện (training)

Hình 2 so sánh hiệu năng của người dùng trong hệ thống sử dụng JC - DQN và DQN. Lưu ý rằng với thuật toán MP thì công suất được gán cố định và không bị ảnh hưởng bởi số lần training nên kết quả mô phỏng không đề cập thuật toán MP trong phần này. Đối với thuật toán JC - DQN và DQN quá trình huấn luyện được mô tả như sau:

- Với mỗi giá trị epsilon, chương trình mô phỏng phát một mô hình mạng với các thông số như vị trí người dùng, pha-đỉnh, xác suất đường truyền thẳng và không thẳng.
- Với mỗi mô hình mạng được phát, quá trình truyền tin được thực hiện trong 100 khe thời gian, thời gian của mỗi khe thời gian là 20 ms.
- Trong quá trình truyền tin, thuật toán DQN thực hiện quá trình học để tối ưu công suất phát cho từng trạm phát và tính toán dung lượng của người dùng. Thông qua việc đánh giá dung lượng người dùng, thuật toán DQN thực hiện tối ưu các giá trị công suất. Lưu ý rằng việc lựa chọn các thông số như hàm lỗi, hệ số chiết khấu (discount factor) ảnh hưởng lớn tới sự hội tụ của thuật toán.

Hình 2 mô tả tốc độ trung bình của quá trình huấn luyện với các giá trị epoch khác nhau. Tốc độ biểu thị trong Hình 2 là tốc độ trung bình của 100 khe thời gian và được vẽ bằng kỹ thuật cửa sổ với kích thước cửa sổ là 5.



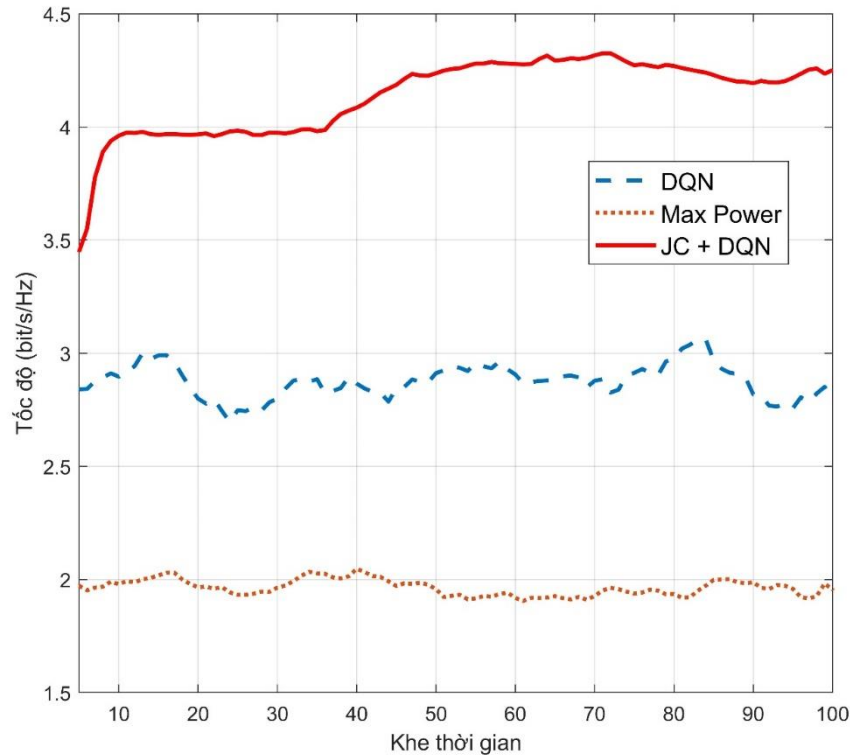
Hình 2. So sánh hiệu năng JC - DQN, DQN với Training Epoch

Hình 2 cho thấy hệ thống JC - DQN hoàn toàn vượt trội so với hệ thống chỉ sử dụng DQN. Trong các giá trị epoch ban đầu, các thuật toán thực hiện việc xây dựng mạng cũng như các tính chất vô tuyến. Do đó, đóng góp của thuật toán Q-learning trong giai đoạn này là không đáng kể. Sự khác biệt khoảng 2,1 bit/s/Hz về tốc độ người dùng ở đây là do kỹ thuật JC. Cũng theo hình vẽ trên, ta thấy sự khác biệt trên được duy trì trong suốt quá trình huấn luyện (training). Kết thúc pha huấn luyện, chúng ta giả thiết rằng thuật toán DQN đã học được cách ước lượng công suất tối ưu cho từng trạm phát với các mô hình và điều kiện truyền tin cụ thể.

3.3. Đánh giá hiệu năng theo khe thời gian tại pha kiểm thử

Trong phần này, chúng tôi đánh giá hiệu năng của 3 thuật toán JC – DQN, DQN và MP trong pha kiểm thử (testing phase) với giả thiết rằng việc lựa chọn công suất tối ưu cho từng trạm phát trong mô hình mạng cụ thể đã được xác định trong pha huấn luyện. Tương tự với pha huấn luyện,

quá trình kiểm thử cũng được thực hiện với 100 khe thời gian và thời gian của mỗi khe thời gian là 20 ms. Hình 3 cho thấy sự ổn định và tương quan giữa tốc độ của người dùng biên trong hệ thống sử dụng JC - DQN, chỉ sử dụng DQN và sử dụng thuật toán MP.

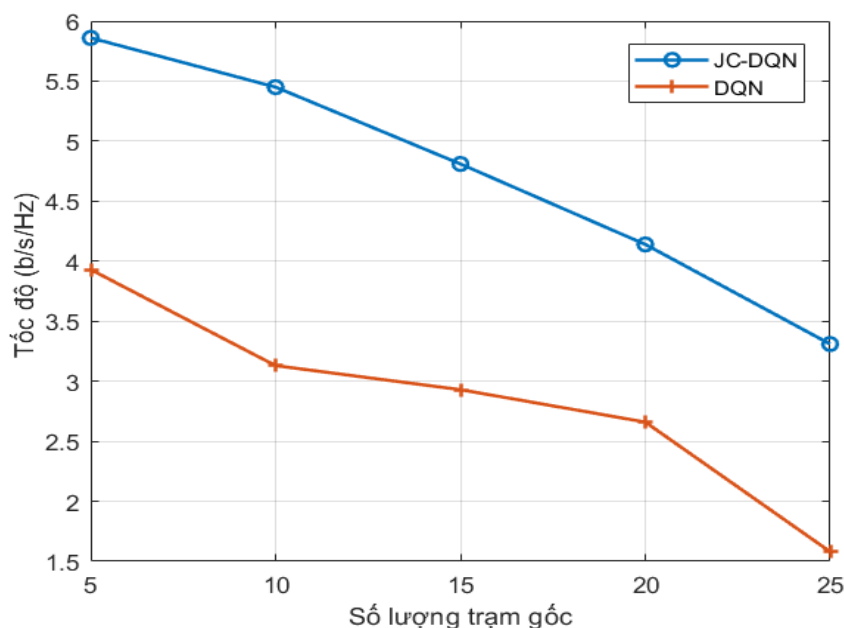


Hình 3. So sánh hiệu năng trong trường hợp sử dụng JC - DQN với DQN và MP

Hình 3 cho thấy rằng hệ thống sử dụng JC - DQN có thể cho hiệu năng vượt trội so với hệ thống chỉ sử dụng DQN và MP. Cụ thể, người dùng biên trong hệ thống sử dụng JC - DQN có thể đạt tốc độ trung bình là 4,2 bit/s/Hz, cao hơn khoảng 35% so với trường hợp chỉ sử dụng DQN và gấp đôi so với trường hợp sử dụng thuật toán MP. Thuật toán MP đạt hiệu năng thấp nhất do việc sử dụng công suất tối đa của các trạm phát cũng đồng nghĩa với việc công suất nhiễu từ các trạm lân cận gây ra cho người dùng biên ở mức tối đa. Điều này trực tiếp làm giảm hiệu năng của mạng. Ngoài ra, thuật toán MP gán cố định công suất cho từng trạm nên không có tính linh động khi môi trường thay đổi. Mạng học sâu DQN sử dụng thuật toán Q-learning có tính linh động cao và khắc phục được nhược điểm của thuật toán MP bằng cách tính toán và cấp công suất phù hợp cho từng trạm, trong từng khe thời gian cụ thể. Tuy nhiên, hệ thống sử dụng DQN vẫn chỉ cho người dùng kết nối tới 1 trạm, trong trường hợp này là trạm gần nhất. Do sự biến động của kênh truyền như giá trị độ lợi công suất tức thời và suy hao, trạm gần nhất có thể không có tín hiệu mạnh nhất. Do đó, hệ thống này chưa tối ưu tối đa được hiệu năng người dùng biên. Kỹ thuật kết hợp JC và DQN khắc phục được nhược điểm này bằng cách cho phép người dùng kết nối với trạm có công suất lớn nhất trong khe thời gian đó. Đồng thời, mạng DQN thực hiện tối ưu công suất phát của các trạm. Do đó, hệ thống kết hợp JC và DQN cho hiệu năng lớn nhất.

3.4. So sánh hiệu năng JC – DQN và DQN với số lượng các trạm phát khác nhau

Trong phần này, hiệu năng của kỹ thuật JC – DQN và DQN được so sánh trong các mô hình mạng có số lượng trạm phát sóng khác nhau. Cụ thể, số lượng trạm gốc được xem xét thay đổi trong khoảng từ 5 đến 25. Giá trị tốc độ trung bình trong Hình 4 được đo bằng cách lấy trung bình của 500 lần kiểm thử.



Hình 4. So sánh hiệu năng JC-DQN, DQN khi tăng số lượng trạm gốc

Hình 4 cho thấy giá trị tốc độ trung bình của người dùng trong cả 2 hệ thống sử dụng JC-DQN, và DQN giảm xuống khi mật độ trạm gốc trở nên dày đặc hơn. Nguyên nhân là do sự gia tăng số lượng trạm gốc đồng nghĩa với sự gia tăng nhiễu tác động lên người dùng. Khi đó, người dùng sẽ chịu ảnh hưởng nhiều từ nhiều trạm lân cận và thu được tín hiệu kém hơn. Hệ quả là hiệu năng người dùng giảm. Ngoài ra, hình 4 còn cho thấy hiệu năng của người dùng trong hệ thống JC-DQN vượt trội hơn hệ thống DQN với khác biệt khoảng 2 bit/s/Hz. Việc đạt được hiệu năng vượt trội vẫn được duy trì khi số lượng trạm gốc tăng lên. Sự vượt trội này là do kỹ thuật JC.

4. Kết luận

Trong bài báo này, các kỹ thuật tiên tiến của mạng di động là kỹ thuật lập lịch đồng thời và mạng học sâu sử dụng thuật toán Q-learning đã được xem xét để cải thiện dung lượng của người dùng biên trong môi trường suy hao công suất dual-slope. Mô hình dual-slope được khuyến nghị bởi 3GPP để mô hình hóa sự mất mát công suất trong môi trường trong nhà, trong đó kênh truyền giữa người dùng và trạm phát có thể là kênh truyền trong tầm nhìn thẳng (LoS) hoặc kênh truyền không trong tầm nhìn thẳng (nLoS). Các kết quả phân tích đều cho thấy rằng việc sử dụng kỹ thuật lập lịch đồng thời và mạng học sâu có khả năng nâng cao đáng kể hiệu năng của người dùng biên.

TÀI LIỆU THAM KHẢO/ REFERENCES

- [1] M. Agiwal, A. Roy, and N. Saxena, "Next generation 5g wireless networks: A comprehensive survey," *IEEE Communications Surveys Tutorials*, vol. 18, no. 3, pp. 1617–1655, 2016.
- [2] T. N. H. Holma and A. Toskala, *5G Technology: 3GPP New Radio*, Wiley, December 2019.
- [3] L. Daewon, S. Hanbyul, B. Clerckx, E. Hardouin, D. Mazzaresse, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in LTE-advanced: deployment scenarios and operational challenges," *IEEE Commun. Mag.*, vol. 50, no. 2, pp. 148–155, 2012.
- [4] S. Y. Jung, H. K. Lee, and S. L. Kim, "Worst-Case User Analysis in Poisson Voronoi Cells," *IEEE Commun. Lett.*, vol. 17, no. 8, pp. 1580–1583, August 2013.
- [5] 3GPP, "Evolved universal terrestrial radio access (e-utra); further advancements for e-utra physical layer aspects," *Technical Specification (TS) 36.300 version 9.4.0 Release 9, 3rd Generation Partnership Project (3GPP)*, 2018, pp.1 – 178.

-
- [6] F. Meng, P. Chen, L. Wu, and J. Cheng, "Power allocation in multi-user cellular networks: Deep reinforcement learning approaches," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6255–6267, 2020.
 - [7] B. Jang, M. Kim, G. Harerimana, and J. W. Kim, "Q-learning algorithms: A comprehensive classification and applications," *IEEE Access*, vol. 7, pp. 133653–133667, 2019.
 - [8] L. Zhong, X. Ji, Z. Wang, J. Qin, and G.-M. Muntean, "A q-learning driven energy-aware multipath transmission solution for 5g media services," *IEEE Transactions on Broadcasting*, vol. 68, no. 2, pp. 559–571, 2022.
 - [9] W. S. Afifi, A. A. El-Moursy, M. Saad, S. M. Nassar, and H. M. El-Hennawy, "A novel scheduling technique for improving cell-edge performance in 4G/5G systems," *Ain Shams Engineering Journal*, vol. 12, no. 1, pp. 487-495, 2021.
 - [10] H. Mariam, I. Ahmed, S. Ali, M. I. Aslam, and I. U. Rehman, "Performance of Millimeter Wave Dense Cellular Network Using Stretched Exponential Path Loss Model," *Electronics*, vol. 11, no. 24, 2022, Art. no. 4226.
 - [11] 3GPP, "5G; Study on channel model for frequencies from 0.5 to 100 GHz," *Technical Report (TR) 38.901 version 14.3.0 Release 14, 3rd Generation Partnership Project (3GPP)*, 2018, pp. 1 – 103.
 - [12] 3GPP, "Study on nr positioning support," *Technical Report (TR) 38.85 version 16, 3rd Generation Partnership Project (3GPP)*, 2018, pp. 1 – 197.
 - [13] 3GPP, "5G; NR; Base Station (BS) radio transmission and reception," *Technical Specification (TS) 38.104 version 15.3.0 Release 15, 3rd Generation Partnership Project (3GPP)*, 2018, pp. 1 – 160.