

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG



PHAN HOÀNG TIỀN

KHẢO SÁT ĐIỀU KHIỂN TỰ CHẤP NHẬN DỊCH VỤ TRONG MẠNG 5G

CHUYÊN NGÀNH: HỆ THỐNG THÔNG TIN

MÃ SỐ : 8.48.01.04

TÓM TẮT LUẬN VĂN THẠC SĨ KỸ THUẬT

(Theo định hướng ứng dụng)

Hà Nội - 2019

Luận văn được hoàn thành tại:

HỌC VIỆN CÔNG NGHỆ BƯU CHÍNH VIỄN THÔNG

Người hướng dẫn khoa học: TS. NGUYỄN VĂN THỦY

(Ghi rõ học hàm, học vị)

Phản biện 1:

Phản biện 2:

Luận văn sẽ được bảo vệ trước Hội đồng chấm luận văn thạc sĩ tại Học viện
Công nghệ Bưu chính Viễn thông

Vào lúc: giờ ngày tháng năm

Có thể tìm hiểu luận văn tại:

- Thư viện của Học viện Công nghệ Bưu chính Viễn thông

MỞ ĐẦU

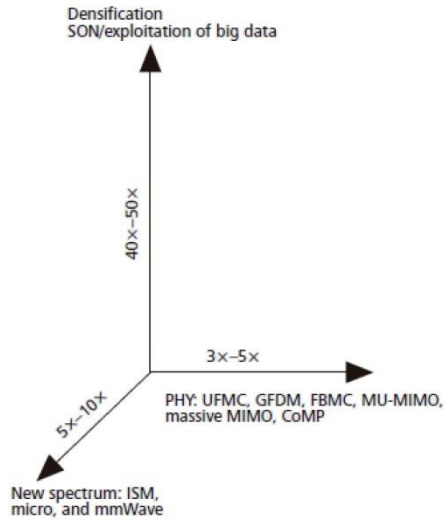
1. Lý do chọn đề tài

Với nhu cầu sử dụng mạng không dây không ngừng gia tăng như hiện nay các nhà mạng viễn thông luôn cải tiến và áp dụng các mô hình mạng tiên tiến để đáp ứng nhu cầu sử dụng mạng không dây của người dùng. Nếu mạng 1G là hệ thống thông tin di động không dây cơ bản đầu tiên trên thế giới vào khoảng thập niên 80, đã làm thay đổi thế giới trong việc kết nối thông tin giữa người với người, thì mạng 5G là thế hệ mới của hệ thống thông tin di động được kỳ vọng áp dụng vào năm 2020 với khoảng 8 tỷ lượt sử dụng trên thế giới [1]. Ngày nay, nhờ sự tăng trưởng vượt bậc của các thiết bị thông minh và sự ra đời của Internet vạn vật (IoT), mạng 5G được hướng đến kết nối mọi vật với nhau. Không như các mạng di động thế hệ trước, mạng 5G được thiết kế dựa trên tập hợp các công nghệ mới có hiệu năng cao và đề tối ưu theo các chỉ số hiệu suất (KPIs). Cụ thể, theo góc nhìn của nhà vận hành, các chỉ số KPIs được xem xét là: hiệu năng mạng, chất lượng dịch vụ (QoS), chi tiêu vốn (CAPEX) và chi phí hoạt động (OPEX). Mặt khác, theo góc nhìn của người tiêu dùng các chỉ số KPIs chủ yếu gồm: kết nối liền mạch, độ trễ không đáng kể và chi phí dịch vụ thấp.

Đề tài “Khảo sát điều khiển tự chấp nhận dịch vụ trong mạng 5G” nhằm nghiên cứu và thử nghiệm các kỹ thuật, thuật toán tự tối ưu chấp nhận dịch vụ trong việc nâng cao hiệu năng trong mạng 5G.

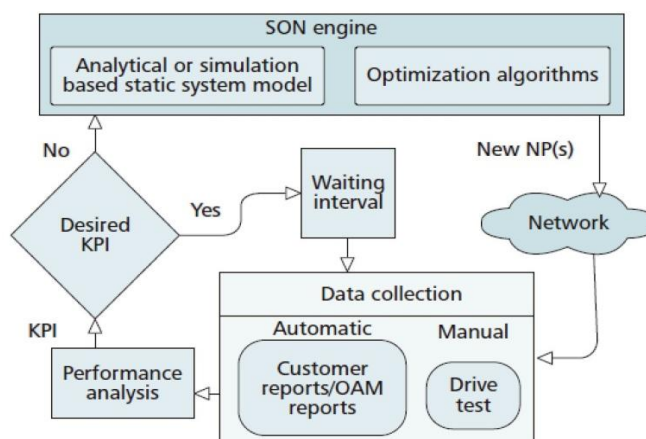
2. Tổng quan về vấn đề nghiên cứu

Để giải quyết nhu cầu dung lượng mạng gia tăng không ngừng, có ba giải pháp được đề xuất: cải thiện hiệu suất băng tần của công nghệ không dây (3-5 lần), cho phép phân bổ nhiều băng tần (5-10 lần), triển khai nhiều nút mạng hơn (40-50 lần) hoặc bằng cách khai thác cả ba cách tiếp cận nói trên cho phép tổng dung lượng tăng 600-2.500 lần như minh họa kích thước tăng trưởng trong mạng 5G bên dưới. Sự phức tạp trong hoạt động của mạng lưới 5G trong tương lai sẽ tăng tuyến tính với mật độ mạng tăng và độ phức tạp của thiết kế phần cứng tăng[10].



Minh họa kích thước tăng trưởng công suất dự kiến trong 5G [10]

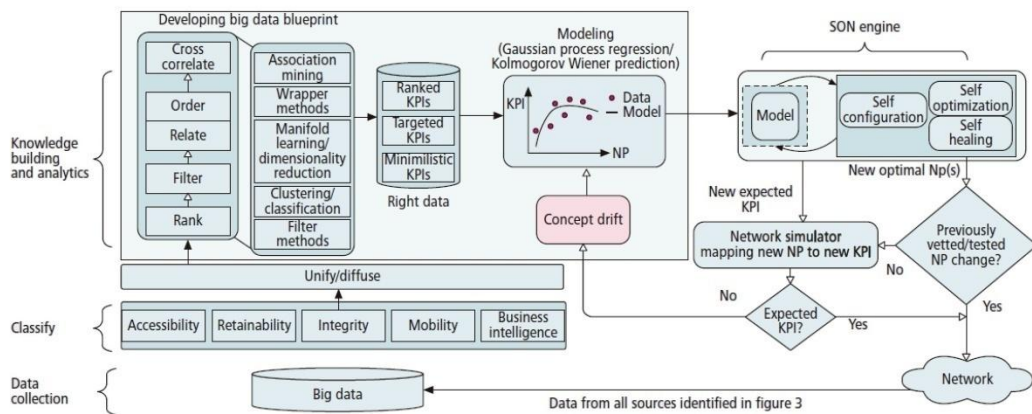
Trong bối cảnh mạng không dây phát triển, hệ thống mạng 5G được yêu cầu phải có tính năng tự tổ chức để đảm bảo lợi nhuận của các mô hình kinh doanh khác nhau. Trong bối cảnh này, việc giới thiệu một mạng tự tổ chức (SON) sẽ cho phép khai thác các thuật toán thông minh nhân tạo (AI) để quản lý hiệu quả tài nguyên mạng, cho phép người dùng cảm nhận được kết nối liền mạch và trong suốt. Do đó, SONs cho phép giảm OPEX (Chi phí hoạt động) với tính năng tự cấu hình, tự tối ưu hóa và tự phục hồi dịch vụ. Các hệ thống mạng hiện tại: mạng 4G, 3G và thậm chí 2G thường tuân theo hình minh họa dưới đây. Tuy nhiên, cách tiếp cận SONs này chưa được tối ưu, vì nó không cung cấp các mô hình động để dự đoán hành vi hệ thống theo kiểu hoạt động trực tiếp nhằm đáp ứng các yêu cầu về độ trễ nghiêm ngặt của thế hệ di động sắp tới [10].



Mô hình SON trong mạng 2G, 3G và 4G [10]

Một mô hình làm việc hoạt động cho mạng 5G bởi SON được mô tả dưới đây. Có thể thấy dữ liệu lớn, được định nghĩa ngắn gọn là lượng thông tin khổng lồ có sẵn từ các nguồn khác nhau của mạng di động, là tính năng chính giúp SONs trong tương lai khác biệt với các

hệ thống di động trước. Các nguồn dữ liệu lớn cho mạng 5G SONS có thể được chia thành ba lớp chính: dữ liệu cấp thuê bao (ví dụ: tỷ lệ thành công cuộc gọi, tỷ lệ cuộc gọi hỏng, chất lượng âm thanh, lưu lượng IP), dữ liệu cấp trạm gốc (ví dụ: công suất nhiễu nhiệt, công suất cơ sở kênh) và dữ liệu cấp độ mạng lõi (ví dụ: nhật ký báo động, bản ghi cấu hình thiết bị, xác thực). Bên cạnh việc thu thập dữ liệu, việc giới thiệu công cụ học máy và phân tích dữ liệu cho phép chuyển đổi tự động từ dữ liệu lớn (thô) sang dữ liệu đúng (có ý nghĩa). Khi dữ liệu phù hợp có giá trị, các mô hình hành vi của hệ thống và người dùng có thể được trích xuất và được gửi đến bộ máy SON để thực hiện các chức năng SON thích hợp.



Khuôn làm việc dự kiến trong tương lai cho mạng 5G bởi SON [10]

SON cho phép khai thác các kỹ thuật dựa trên trí thông minh nhân tạo (ví dụ: học máy, thuật toán lấy cảm hứng từ sinh học, Q-Learning...) để xử lý một cách hiệu quả các vấn đề của hệ thống phức tạp quy mô lớn.

Trong đó, kỹ thuật trí tuệ nhân tạo được chọn để phát triển một thuật toán kiểm soát tự tổ chức nhập cell cho các mạng 5G. Mặt khác, phương pháp được chọn không thể là phương pháp tối ưu trong số tất cả các giải pháp được đề cử, nhưng đã được lựa chọn theo các thuộc tính dễ áp dụng của nó để phù hợp với hầu hết các quy trình tự tối ưu hóa.

3. Mục đích nghiên cứu

- Tìm hiểu tổng quan lý thuyết về SON và kỹ thuật dựa trên AI để tự tối ưu hóa trên các mạng không đồng nhất (HetNets)
- Phân tích lý thuyết về kiểm soát đầu vào cho hệ thống truy cập vô tuyến nhiều người dùng

- Từ đó, xác định thuật toán tự tối ưu hóa phù hợp nhất cho đề tài nghiên cứu và thực hiện
- Phân tích kết quả của thuật toán AI và nghiên cứu tính khả thi của AI trong triển khai SON giả thuyết
- Nắm được các kỹ thuật và phương pháp xử lý thuật toán trong đề tài.

4. Đối tượng và phạm vi nghiên cứu

Đối tượng nghiên cứu: Hệ thống mạng tự tổ chức SON, sử dụng trí tuệ nhân tạo cho mạng không đồng nhất HetNets. Cách thức điều khiển tự chấp nhận dịch vụ trong mạng 5G

Phạm vi nghiên cứu: Giới hạn trong môi trường mô phỏng.

5. Phương pháp nghiên cứu

Phương pháp nghiên cứu là kết hợp nghiên cứu lý thuyết và thực nghiệm.

CHƯƠNG 1: CƠ SỞ LÝ LUẬN

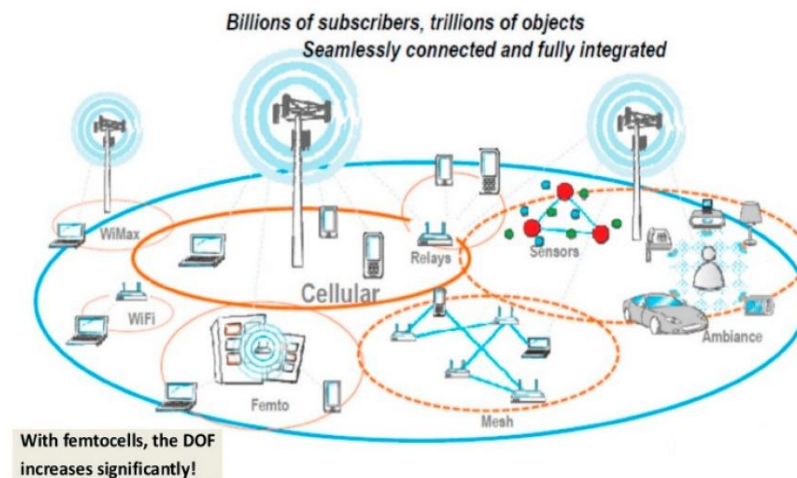
1.1. Hệ thống mạng tự tổ chức SON

1.1.1. Định nghĩa

Hệ thống mạng tự tổ chức được định nghĩa là một công nghệ tự động hóa được thiết kế để giúp cho việc lập kế hoạch, cấu hình, quản lý, tối ưu hóa và bảo trì các mạng truy cập vô tuyến đơn giản và nhanh hơn.

Với mục đích làm rõ nhu cầu áp dụng SONs cho mạng 5G trong tương lai, thuật ngữ mức độ tự do (DoF) được giới thiệu. DoF của hệ thống được định nghĩa là số lượng thông số có thể tinh chỉnh trong bất kỳ mạng không dây nào. Tuy nhiên một số DoF có tỷ lệ phụ thuộc lẫn nhau cao.

HetNets được định nghĩa là một mạng không dây chứa rất nhiều RAT, định dạng ô và nhiều khía cạnh khác, nhằm kết hợp chúng để vận hành một cách liên mạch minh họa hình 1.1. Do đó, DoF tăng đáng kể với các HetNets đang phát triển và do đó, xác suất mọi thứ trở nên sai (ví dụ như sự cố ngừng hoạt động, lỗi chuyển giao).



Hình 1.1: Minh Họa mạng HetNets [6]

SONs nhằm giảm thiểu hiệu quả của DoF trong HetNets và cải thiện khả năng mở rộng của toàn bộ hệ thống, bằng cách giảm chi phí vòng đời (O / CAPEX) và tối ưu hóa hiệu suất mạng vô tuyến trong quá trình hoạt động.

1.1.2. Cấu trúc

- (1) Địa phương hóa
- (2) Phân tán
- (3) Tập trung
- (4) Hỗn hợp

1.1.3. Các chức năng tự làm việc

- Tự cấu hình
- Tự tối ưu hóa
- Tự phục hồi

1.1.4. Đặc điểm

1.2 Kỹ thuật trí thông minh nhân tạo cho mạng HetNet

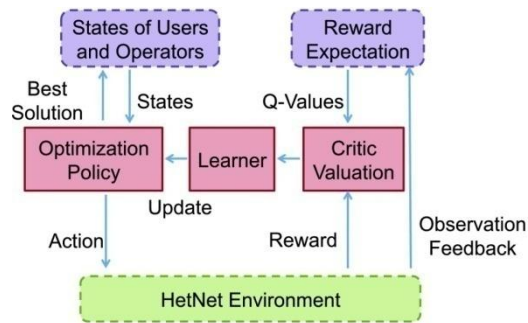
HetNets đang trở nên khá khó khăn để giải quyết khi số lượng tài nguyên mạng liên tục tăng lên. Kỹ thuật trí thông minh nhân tạo (AI) nhằm khắc phục những hạn chế của các hệ thống quy mô lớn và do đó, việc thực hiện chúng sẽ bổ sung trí thông minh cho HetNets hiện tại và tương lai để giảm sự tham gia của con người, đó là một trong những mục tiêu chính của SON. Sau đó, các kỹ thuật dựa trên AI có thể giảm đáng kể chi phí hoạt động và vốn (O / CAPEX) và tối ưu hóa dung lượng mạng, độ phủ sóng và Chất lượng dịch vụ (QoS) trong HetNets, theo các tính năng tự làm việc [20].

Các kỹ thuật trí thông minh nhân tạo góp phần là mục tiêu chính của việc biến HetNets thông minh hơn, nhưng chúng có thể khá khác nhau về mô hình hoạt động. Một số được lấy cảm hứng từ các phát hiện trong thiên nhiên (ví dụ: Thuật toán lấy cảm hứng từ Sinh học), một vài trong số chúng được thúc đẩy bởi các cách thức lý luận của con người (ví dụ: Hệ thống mờ) và một số khác dựa trên tương tác cục bộ và học tập dựa trên phản hồi để quy (ví dụ: Machine Learning). Phần này tiếp tục nghiên cứu về các kỹ thuật AI và tính khả thi, tìm hiểu được ưu nhược điểm của từng loại trong các ứng dụng mạng nhất định. Các kỹ thuật dựa trên AI có liên quan nhất đang được nghiên cứu cho triển khai trong HetNets sẽ được trình bày với các phương pháp đã được chọn để thực hiện AC tự tổ chức cho thuật toán mạng 5G nhiều người dùng.

1.2.1. Học máy

Học máy (ML) là một lĩnh vực của trí tuệ nhân tạo liên quan đến việc nghiên cứu và xây dựng các kỹ thuật cho phép các hệ thống học tự động từ dữ liệu để giải quyết những vấn đề cụ thể. Ngày nay, có rất nhiều ứng dụng dựa trên ML thành công trong các lĩnh vực khác nhau. Ví dụ, các công ty bán lẻ thu thập dữ liệu mua hàng trong quá khứ để phân tích hành vi của khách hàng, nhờ đó cải thiện dịch vụ. Các công ty tài chính xem xét các giao dịch trong quá khứ để dự phóng rủi ro tín dụng của khách hàng. Hầu hết các ứng dụng email, không xem xét mức độ phổ biến của chúng, sử dụng ML để quyết định xem thư đến có nên được coi là

spam hay không. Trong thông tin sinh học, số lượng dữ liệu khổng lồ có sẵn chỉ có thể được phân tích và trích xuất bằng cách sử dụng khai thác dữ liệu (data mining) [1].



Hình 1.2: Lựa chọn tối ưu hóa dựa trên học máy trong HetNets [20]

1.2.2. Thuật toán lấy cảm hứng từ sinh học

1.3. Kết luận chương

CHƯƠNG 2: PHƯƠNG PHÁP TỰ CHẤP NHẬN DỊCH VỤ TRONG MẠNG 5G

2.1. Điều khiển dịch vụ cho các mạng truy cập vô tuyến

2.1.1. Kiểm soát điều khiển nhập cell nhiều người dùng

2.1.2. Kết quả thuật toán

2.1.3. Đánh giá hiệu quả

2.2. Điều khiển tự chấp nhận dịch vụ trong mạng 5G

Trong phần này, một số kỹ thuật AI sẽ được đề xuất để phát triển một AC tự tổ chức cho các mạng 5G nhiều người dùng. Cụ thể hơn, mục tiêu chính là tập trung vào tự học giá trị tối ưu của thuật ngữ $\Delta C(s, n)$ mà trước đây đã được định nghĩa. Các phương pháp được đề xuất tương ứng với AI nổi tiếng của ML, trong đó hai thuật toán học tập chính phát sinh:

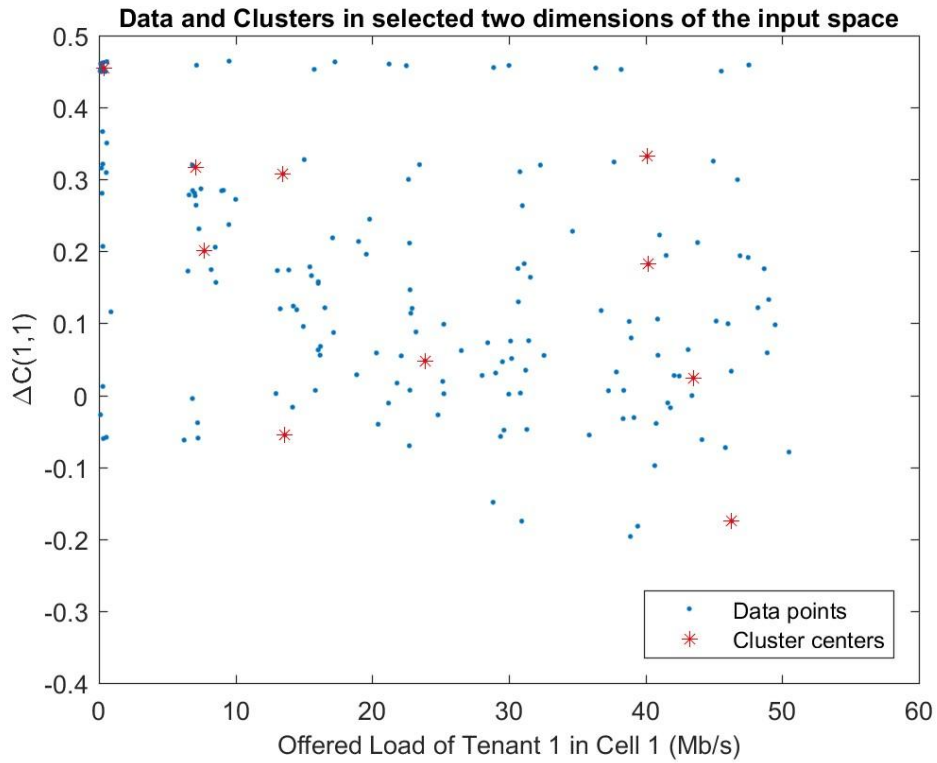
- **Học tập được giám sát:** chủ yếu được sử dụng khi cả hai biến đầu vào và đầu ra của một hệ thống được xác định, do đó một hàm ánh xạ có thể được xác định ($Y = f(x)$). Tuy nhiên, mục tiêu chính là ước tính hàm ánh xạ theo cách mà các biến đầu ra có thể được dự đoán chính xác bất cứ khi nào có dữ liệu đầu vào mới (tức là dữ liệu chưa được thực hiện).
- **Học tập không được giám sát:** không giống như học tập được giám sát, học tập không được giám sát thiếu bất kỳ một chức năng ánh xạ, vì dữ liệu đầu ra không rõ hoặc khó có được. Do đó, các thuật toán không giám sát được thiết kế để khám phá cấu trúc tối ưu hoặc mối quan hệ giữa các đầu vào/đầu ra khác nhau.

2.2.1. Học tập được giám sát

Việc thực hiện phương pháp học có giám sát để khai thác kiến thức về $\Delta C(s, n)$ có hai mục tiêu chính. Trước hết, kiến thức cho phép xác định các giới hạn của $\Delta C(s, n)$ và hành vi tổng thể của nó trong các tình huống tải lưu lượng khác nhau. Điều này dẫn đến mục tiêu thứ hai, nhằm mục đích tạo thuận lợi cho việc thực hiện phương pháp học AC không giám sát. Vì có 4 biến để tự tối ưu hóa ($\Delta C(1,1), \Delta C(1,2), \Delta C(2,1), \Delta C(2,2)$), và thời gian hội tụ giải pháp tối ưu có thể quá cao, việc học có giám sát sẽ được áp dụng trong một số biến $\Delta C(s, n)$ vì mục đích đơn giản, và các biến bên trái sẽ được chọn để tối ưu hóa.

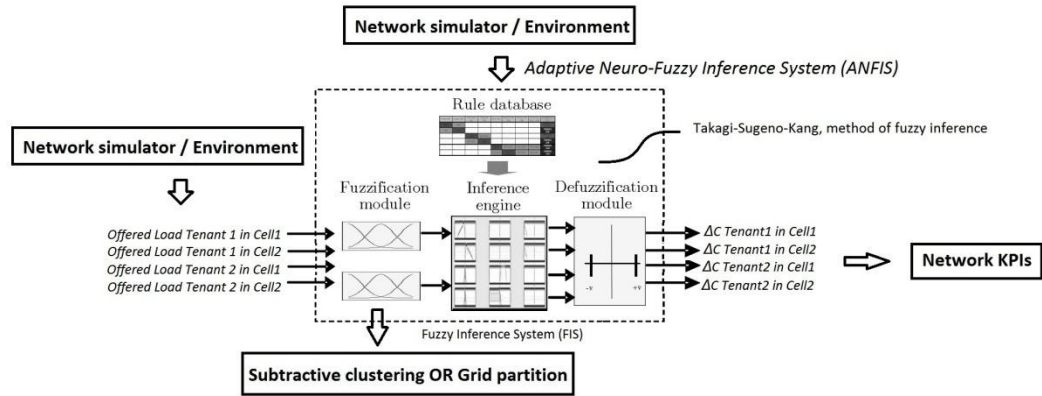
Đối với phạm vi của bài nghiên cứu, một hệ thống suy luận thần kinh thích nghi mờ (ANFIS) được đề xuất như là một kỹ thuật học được giám sát. ANFIS là một loại ANN kết

hợp động cơ suy luận mờ Takagi-Sugeno, chỉ tạo ra một đầu ra duy nhất sau giai đoạn làm mờ (tức là một trong các $\Delta C(s, n)$). Mặt khác, trong 4 đầu vào được xem xét, tương ứng với tải trọng được cung cấp của từng đối tượng thuê trong mỗi ô, được lấy trực tiếp từ môi trường/hệ thống mô phỏng mạng. Hơn nữa, hệ thống suy luận mờ (FIS) có thể kết hợp hai kiểu phân cụm dữ liệu: phân vùng lưới và phân cụm trù. Mỗi đầu vào có nhiều chức năng như số lượng các cụm được xác định. Ví dụ, 10 cụm được xác định trong hình 2.7, mặc dù có thể tìm thấy số lượng cụm nhiều hơn hoặc ít hơn, cần xem xét sự đánh đổi giữa lỗi đào tạo và thời gian đào tạo.



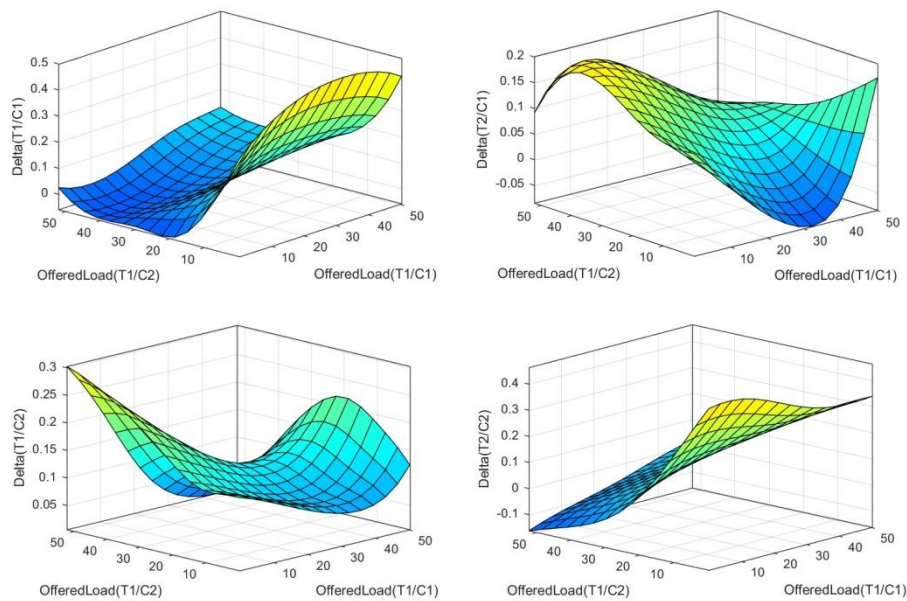
Hình 2.1: Kỹ thuật phân cụm trù (Bán kính cụm ảnh hưởng = 0,65)

Tổng quan đề án học tập đề xuất được minh họa dưới đây hình 2.8.



Hình 2.2: Đề án học máy có giám sát nhằm khai thác kiến thức $\Delta C(s,n)$

Cuối cùng, khi dữ liệu đầu vào/đầu ra đã được xử lý thông qua lược đồ nói trên, kiến thức về $\Delta C(s, n)$ có sẵn để chính xác ($> 99\%$) khai thác dữ liệu đầu vào chưa được xử lý. Một vài biểu diễn của $\Delta C(s, n)$ như là một hàm của các tải được cung cấp khác nhau được thể hiện trong hình 2.9.



Hình 2.3: $\Delta C(s,n)$ như là một chức năng của T1 được cung cấp (Mb/s) trong cả hai ô

Bước tiếp theo là lấy giá trị tối ưu của $\Delta C(s, n)$ với điều kiện tải lưu lượng cho trước trong mỗi lần lặp tối ưu hóa của chương trình học không giám sát, để lại một đặc tính $\Delta C(s, n)$ tự tối ưu hóa, với mục đích đơn giản hóa.

2.2.2. Học tập không được giám sát (Fuzzy Q-Learning)

Để đạt được việc tự tối ưu hóa, mỗi tác nhân phân tán cần biết tham số cần được điều chỉnh theo trạng thái hoạt động hiện tại. Các khái niệm cơ bản về FQL sẽ được trình bày dưới đây:

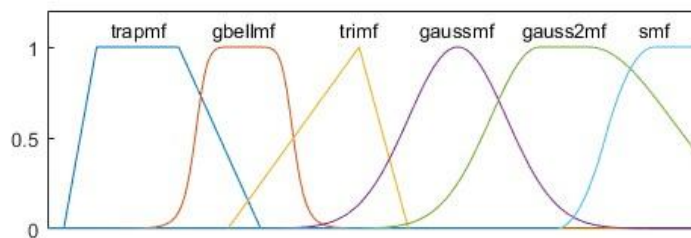
a. Hệ thống mờ

“Sự mơ hồ là một phần phổ biến của kinh nghiệm của con người. Ngôn ngữ của con người là một công cụ không chính xác. Nhận thức của con người đầy sự thiếu chính xác. Những kỷ niệm thật thoáng qua và dễ chỉnh sửa. Thế giới thực không phải là trừu tượng, nó không được nhận thức rõ ràng, xác định rõ ràng và tính toán chính xác”

Lý thuyết mờ đã được phát triển để xử lý thông tin không chính xác. Nó bắt đầu với khái niệm về tập mờ, có chức năng là ánh xạ (tức là làm mờ) tập hợp các phần tử đầu vào thành một hàm cho biết mức độ chân lý thuộc về tập hợp.

$$\mu_A: X \rightarrow [0,1] \quad (2.6)$$

Công thức 2.6 biểu diễn mức độ chân lý μ của tập mờ A lấy biến đầu vào x nằm trong khoảng từ 0 (tức là x không thuộc A) đến 1 (tức là cách khác). Tuy nhiên, bên cạnh tập hợp cổ điển đặc biệt mà một phần tử có thể thuộc về hay không, logic mờ cho phép biến đầu vào được ánh xạ trong một tập hợp nhất định theo nghĩa rộng hơn. Con người làm loại lý luận này mọi lúc, nhưng nó là một khái niệm khá mới mẻ đối với máy tính.



Hình 2.4: Hình dạng hàm thành viên (Matlab Fuzzy Logic Toolbox™)

Ngoài ra, logic mờ cho phép thực hiện kiến thức của con người dưới dạng các quy tắc suy luận if-then. Một quy tắc if-then mờ duy nhất có dạng sau:

If x is A , then y is B

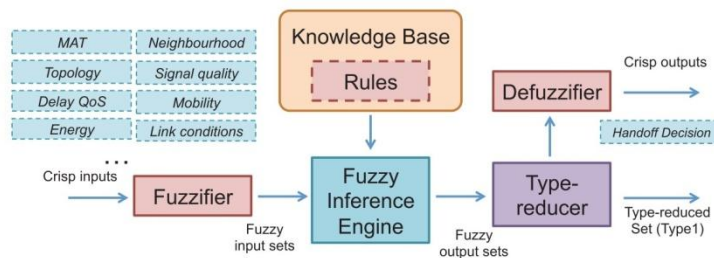
trong đó A và B là các giá trị ngôn ngữ (ví dụ: thấp, trung bình và cao) được xác định bởi các tập mờ X , Y tương ứng. Giá trị đầu vào và đầu ra ngôn ngữ sắc nét (ví dụ: chất lượng tín hiệu và quyết định chuyển giao) là x và y , tương ứng.

Phần “if” của quy tắc “ x là A ” còn được gọi là tiền tố của quy tắc, trong khi phần “then” của quy tắc “ y là B ” được gọi là hậu tố. Đối với một quy tắc if-then, tiền tố, p , ngụ ý hậu tố, q . Trong logic nhị phân, nếu p là đúng, thì q cũng đúng ($p \rightarrow q$). Tuy nhiên, trong logic mờ, nếu p là đúng với một số mức độ, thì q cũng đúng với cùng mức độ [13].

Hơn nữa, có thể nhận thấy rằng các quy tắc dựa trên con người trong logic mờ có thể không tối ưu và do đó, các kỹ thuật tối ưu hóa cần được thực hiện để xây dựng một cơ sở kiến thức chính xác.

Cuối cùng, bước cuối cùng của quá trình suy luận mờ là làm mờ, một phương pháp xác định một giá trị sắc nét duy nhất từ tập hợp đầu ra.

Phương pháp logic mờ phù hợp với sự xử lý thiếu chính xác của các mạng di động không dây [19]. Trên thực tế, kỹ thuật hệ thống mờ đã được đề xuất gần đây để xử lý các thuật toán quyết định. Ví dụ, nghiên cứu trong báo cáo đề xuất một thuật toán quyết định dựa trên logic mờ dạng 2, có tính đến một loạt các mạng truy cập và thuộc tính người dùng, và chọn mạng có giá trị thỏa mãn tối đa. Một minh họa về việc xử lý logic mờ trong HetNets được thể hiện trong hình 2.11.



Hình 2.5: Minh họa Logic mờ cho HetNets [20]

b. Tăng cường học máy (Reinforcement Learning)

c. Q-Learning (QL)

Trong một số mô hình RL, nó đòi hỏi những thông tin rất chính xác về các xác suất truyền dẫn trạng thái và những kết quả. Tuy nhiên, trong một số mô hình hệ thống, thông tin này không có sẵn hoặc chỉ tương đối. Khi đó sẽ có vấn đề khi xây dựng mô hình RL cho các hệ thống này. QL là một trường hợp đặc biệt của RL có thể giải quyết các vấn đề khi các mô hình hệ thống này là không có sẵn. Thay vào đó, nó dựa trên sự khác biệt tạm thời để từng bước giải quyết các vấn đề học máy. QL đạt được mục tiêu này bằng cách ước lượng một hàm giá trị của mỗi cặp trạng thái hành động được gọi là giá trị Q. Chức năng này ước tính

đến giá trị kết quả dự kiến thì lấy một hoạt động a thuộc A ra khỏi tập các hành động A từ khi bắt đầu trạng thái s đến khi đã có một trạng thái cố định π . Mỗi hoạt động chuyển các nhân tố từ trạng thái s_i sang trạng thái s_{i+1} nhận được một kết quả r_{i+1} . Mục tiêu là để mở rộng tập kết quả. Các chức năng giá trị được định nghĩa theo phương trình 2.7:

$$Q^\pi(s, a) = E_\pi \left[\sum_{i=0}^{\infty} \gamma^i r(s_i, a_i) | s_0 = s, a_0 = a \right] \quad (2.7)$$

Và cũng có thể được ước tính bằng cách sử dụng phương pháp cập nhật sự khác biệt tạm thời một cách lặp đi lặp lại:

$$Q_{i+1}(s_i, a_i) = Q_i(s_i, a_i) + \beta [r_{i+1} + \gamma \max_a Q_i(s_{i+1}, a) - Q_i(s_i, a_i)] \quad (2.8)$$

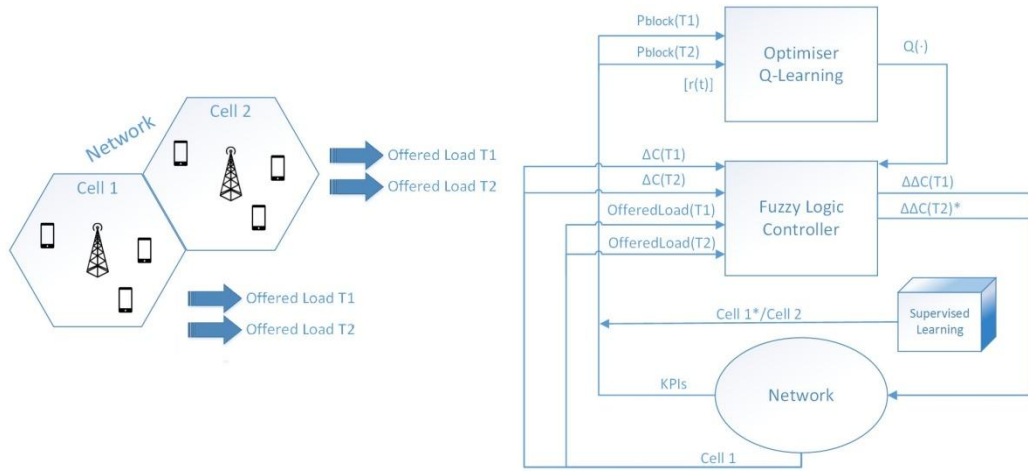
Trong đó: β là tỷ lệ học tập ($0 \leq \beta \leq 1$), các tham số tỷ lệ học tập để xác định việc học tập diễn ra nhanh hay chậm. Sự điều khiển thuật toán QL thì có thể thay đổi giá trị Q với sự biến thiên của các trạng thái và hoạt động. Nếu tỷ lệ học tập là quá nhỏ, quá trình học tập sẽ diễn ra rất chậm. Nếu β là quá cao thì các thuật toán có thể không hội tụ. γ là hệ số chiết khấu ($0 \leq \gamma \leq 1$), nếu $\gamma = 0$ các tác nhân chỉ xem xét những kết quả trước mắt, hệ số γ càng gần 1 thì càng có nhiều nhân tố được xác định trước.

QL là một phương pháp RL, trong đó người học được xây từng bước hàm Q , cố gắng để ước tính chi phí trong tương lai cho một hành động trong trạng thái hiện tại của nhân tố. Kết quả của hàm Q được gọi là giá trị Q . Giá trị Q trong QL là một ước tính về giá trị của chi phí trong tương lai nếu các chi nhánh có một hành động đặc trưng a khi nó đang ở trong trạng thái s . Bằng cách khảo sát môi trường, các nhân tố đã lập ra một bảng giá trị Q cho mỗi trạng thái và mỗi hành động.

Trong một khoảng thời gian ngắn, QL là một kỹ thuật RL với mục tiêu tối đa hóa một giá trị tích lũy bằng cách thực hiện các hành động trong một môi trường. QL tích lũy dần dần một hàm Q , được biểu thị bằng $Q(s, a)$, bằng cách ước tính giá trị tương lai được chiết khấu để thực hiện các hành động từ trạng thái s đã cho. Một phiên bản mờ của QL được xem xét trong công việc này để kế thừa những lợi ích của lý thuyết mờ. Về cơ bản, FQL cho phép giải phóng các trạng thái và không gian hành động để tránh đối phó với các không gian phức tạp và liên tục.

Cấu trúc của quy trình tự tối ưu hóa được thể hiện trong hình 2.12, được phân phối rõ ràng. Bên cạnh khối QL của trình tối ưu hóa, cập nhật hàm Q cho phù hợp với giá trị thu được, bộ điều khiển logic mờ quản lý tập hợp các trạng thái môi trường đầu vào của nó (nghĩa là tải lưu lượng truy cập và $\Delta C(s, 1)$) và tập các kết quả đầu ra của nó (mức độ tăng của

$\Delta C(s, 1)$). Ban đầu, nó được kế hoạch để tự tối ưu hóa cả hai $\Delta C(s, 1)$ và để lại $\Delta C(s, 2)$ được tối ưu hóa thông qua việc học có giám sát. Tuy nhiên, khi thời gian tối ưu hóa là quá dài, $\Delta C(1,1)$ là biến duy nhất được tự tối ưu hóa.



Hình 2.6: Kiến trúc của thủ tục tự tối ưu hóa đề xuất

Các thuật toán mô phỏng linh hoạt đã xuất hiện như một ứng viên tiềm năng cho việc giải quyết các vấn đề trên. Nó là một giải pháp phỏng đoán nghiêm cứu, xác suất bằng cách mô phỏng các quá trình vật lý của luyện kim nghĩa là một chất đang dần nguội để đạt được một trạng thái năng lượng tối thiểu. Phương pháp này đã được áp dụng rộng rãi trong việc quy hoạch mạng lưới và tối ưu hóa. Ví dụ để giải quyết vấn đề bao phủ và khả năng tối ưu hóa dung lượng. Nó cho thấy rằng các thuật toán mô phỏng mềm dẻo có thể được phát triển và áp dụng trong việc giải quyết các vấn đề phức tạp trong tối ưu hóa mạng. Đó là khởi nguồn của thuật toán FQL. Nó được phát triển dựa trên ý tưởng mô phỏng luyện kim. QL là một hình thức thực tế của RL, đó là một lĩnh vực quan trọng của việc học tập máy tính. RL là một kiểu học máy mà tác nhân đó đạt được mục tiêu cuối cùng bằng cách tương tác với môi trường xung quanh và bằng cách xem xét các kinh nghiệm quá khứ bắt nguồn từ hành động trước đó. Trong một tình huống khi chúng ta xử lý cả tính bền vững và tính chính xác, FL đã được mở rộng để xử lý các khái niệm về sự thật, giá trị của nó có thể dao động từ hoàn toàn đúng sự thật và hoàn toàn sai.

d. Fuzzy Q-Learning

QL phải duy trì một giá trị Q cho mỗi cặp trạng thái hoạt động, do đó nó trở nên rất phức tạp và đôi khi không thực tế nếu không gian trạng thái hay không gian hoạt động là liên

tục. Vì vậy các FL được cung cấp để rời rạc các biến liên tục. FQL là một kỹ thuật kết hợp FL với QL nhằm khắc phục những thiếu sót của nhau.

Trong FQL trạng thái và hoạt động của hệ thống được xác định bằng cách sử dụng hàm thành phần mờ. Các biến trạng thái liên tục chuyển đổi thành một số hữu hạn các hàm thành phần biến mờ. Quá trình này được gọi là fuzzification. Từ các biến mờ kết quả tương ứng được tính toán dựa trên hệ thống suy luận mờ (FIS). Cuối cùng kết quả mờ của FIS là ánh xạ của các biến đầu ra liên tục thông qua một quá trình gọi là defuzzification.

e. Hệ thống suy luận mờ

2.3. Kết luận chương

CHƯƠNG 3: ĐÁNH GIÁ, CÀI ĐẶT, THỬ NGHIỆM

3.1. Đánh giá

3.2. Cài đặt MatLab

3.3. Thử nghiệm

3.3.1. Mô phỏng giả lập mạng

3.3.2. Thuật toán Fuzzy Q-Learning

Trước hết, cần xác định khái niệm về giá trị q . Đối với mỗi quy tắc của FIS, $a[i, j]$ được định nghĩa là hành động thứ j của quy tắc i và $q[i, j]$ như là chất lượng giá trị liên quan của nó (giá trị q). Do đó, giá trị $q[i, j]$ càng cao, độ tin cậy của hành động điều chỉnh tương ứng được chọn càng cao.

Để khởi tạo giá trị q trong thuật toán, tiêu chí đơn giản sau được sử dụng:

$$q[i, j] = 0, 1 \leq i \leq N \text{ and } 1 \leq j \leq A \quad (3.1)$$

trong đó $q[i, j]$ là giá trị q của quy tắc i và hành động j . N là tổng số quy tắc và A là số hành động có sẵn cho mỗi quy tắc.

Đối với mỗi quy tắc được kích hoạt (những quy tắc có mức độ chân lý khác không), một hành động được chọn theo chính sách thăm dò/khai thác. Chủ thể nên chọn các hành động tạo ra giá trị cao nhất trong quá khứ. Tuy nhiên, chủ thể học được từ kết quả của những hành động bằng cách thử các hành động chưa được chọn trước đó. Sau đó, ngoài giai đoạn khai thác, cần xem xét chính sách thăm dò để theo dõi các hành động chưa được khám phá mang lại giá trị dài hạn tối đa. Cụ thể:

$$a_i = \begin{cases} \text{random}\{a_k, k = 1, 2, \dots, A\}, & \text{with probability } \mathfrak{T} \\ \text{argmax}_k q[i, k], & \text{with probability } 1 - \mathfrak{T} \end{cases} \quad (3.2)$$

trong đó a_i là hành động cụ thể cho quy tắc i và \mathfrak{T} là tỷ lệ học tập của chính sách thăm dò/khai thác. Thông thường, \mathfrak{T} không cố định trong quá trình tối ưu hóa. Thay vào đó, nó dần dần giảm xuống các giá trị gần bằng không, có nghĩa là việc khám phá các hành động tiềm năng cũng giảm.

Tiếp theo, hành động tổng thể được xác định bởi:

$$a(t) = \sum_{i=1}^N \alpha_i(s(t)) \cdot a_i(t) \quad (3.3)$$

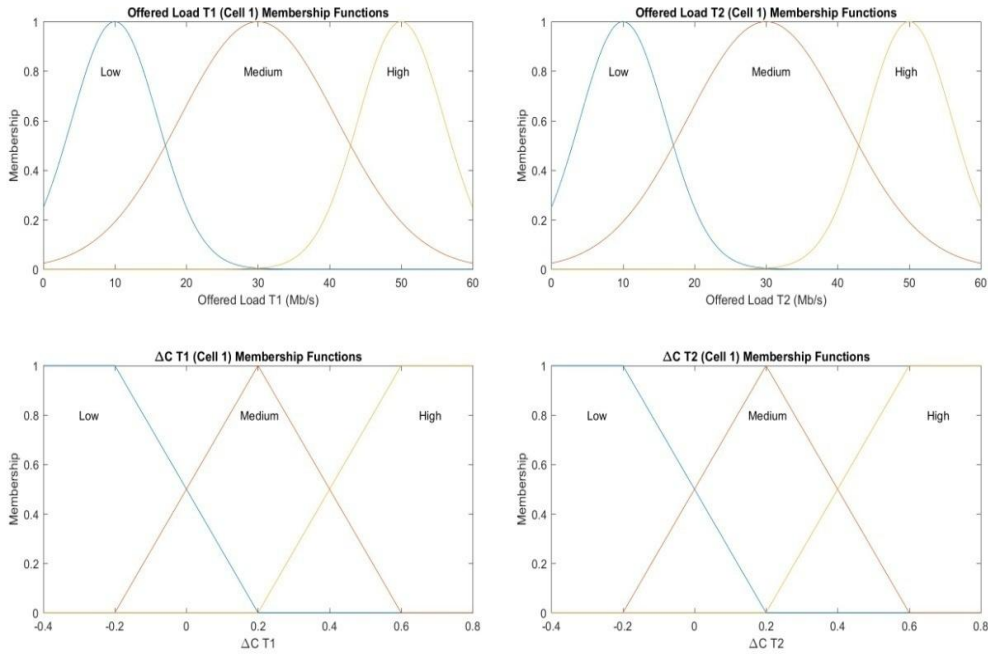
trong đó a là hành động điều chỉnh tham số và $\alpha_i(s(t))$ là hàm kích hoạt cho quy tắc i . Nói cách khác, $\alpha_i(s(t))$ đại diện cho mức độ chân lý của một trạng thái đầu vào $s(t)$ trong lần lặp thứ t :

$$\alpha_i(s(t)) = \prod_{j=1}^M \mu_{ij}(x_j(t)) \quad (3.4)$$

trong đó M là số đầu vào FIS và $\mu_{ij}(x_j(t))$ là giá trị hàm cho đầu vào thứ j và quy tắc thứ i . Ví dụ, xem xét quy tắc đầu tiên trong đó bốn đầu vào được dán nhãn là thấp (L), hàm kích hoạt được cho bởi:

$$\alpha_1(s(t)) = \mu_{11}(x_1(t)) \cdot \mu_{12}(x_2(t)) \cdot \mu_{13}(x_3(t)) \cdot \mu_{14}(x_4(t)) \quad (3.5)$$

Các hình dạng của các hàm được minh họa trong hình 3.2. Đối với tải lưu lượng truy cập cung cấp cho cả hai đối tượng thuê, ba hàm thành viên gaussian được chọn, được gán nhãn là Thấp (L), Trung bình (M) và Cao (H), tương ứng. Đối với $\Delta C(s, 1)$, sử dụng hai hàm thành viên hình thang và một hàm hình tam giác. Tuy nhiên có nhiều tùy chọn khi chọn hình dạng thành viên phù hợp.



Hình 3.1: Chức năng thành phần mờ

trong đó $Q(s(t), a(t))$ là giá trị của hàm Q cho trạng thái và hành động a . Hàm Q có thể được tính toán từ các hàm kích hoạt và các giá trị q của các quy tắc khác nhau:

$$Q(s(t), a(t)) = \sum_{i=1}^N \alpha_i(s(t)) \cdot q[i, \alpha_i] \quad (3.6)$$

trong đó $Q(s(t), a(t))$ là giá trị của hàm Q trong trạng thái s và hành động a .

Bước tiếp theo để hệ thống tự phát triển sang trạng thái tiếp theo $s(t+1)$.

Tại thời điểm này, tín hiệu tăng cường $r(t+1)$ được quan sát. Trong bài nghiên cứu này, tín hiệu tăng cường sau đây được xem xét, tương tự như được đề xuất trong [14]:

$$r(t) = r_1(t) + r_2(t) + k_1; \quad (3.7)$$

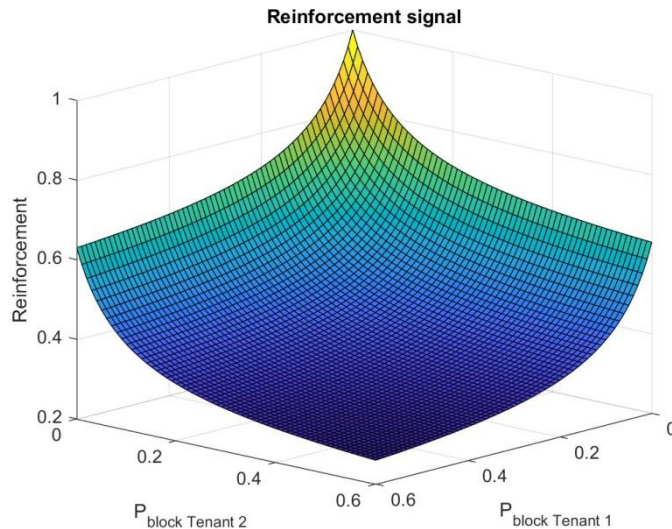
trong đó $r(t)$ là tín hiệu tăng cường tổng thể, $r_1(t)$ và $r_2(t)$ là các yếu tố đóng góp tín hiệu của cả hai đối tượng thuê dọc theo hai ô, và k_1 là một hằng số. Cụ thể, tín hiệu $r_i(t)$ được tính như sau:

$$r_1(t) = k_2 \cdot \log\left(\frac{1}{(P_{block}(T_i) + k_3) \cdot 100} + 1\right) \quad (3.8)$$

trong đó k_2 và k_3 là các tham số không đổi và $P_{block}(T_i)$ là xác suất chặn của đối tượng thuê T_i trong toàn bộ trường hợp. Các thông số được sử dụng để tính toán tín hiệu tăng cường có thể được tìm thấy trong bảng 3.1. Bên cạnh đó, một minh họa của tín hiệu tăng cường được thể hiện trong hình 3.3. Có thể quan sát thấy rằng khi xác suất chặn của cả hai người dùng là bằng không, thì cốt lõi hoặc giá trị thu được là tối đa (tức là bằng 1).

Bảng 3.1: Thông số tín hiệu cốt thép

Tham số	Giá trị
k_1	0.1357
k_2	100
k_3	0.1



Hình 3.2: Tín hiệu tăng cường

Khi tín hiệu tăng cường của trạng thái tiếp theo $r(t + 1)$ đã được quan sát, giá trị của trạng thái mới được biểu thị bằng $V_t(s(t + 1))$ có thể được tính như sau:

$$V_t(s(t + 1)) = \sum_{i=1}^N \alpha_i(s(t + 1)) \cdot \max_k q[i, a_k] \quad (3.9)$$

Tín hiệu lỗi giữa các hàm Q liên tiếp sẽ hữu ích để cập nhật các giá trị q . Nó được tính bởi:

$$\Delta Q = r(t + 1) + \gamma V_t(s(t + 1)) - Q(s(t), a(t)) \quad (3.10)$$

trong đó ΔQ là tín hiệu lỗi, $r(t + 1)$ là tín hiệu tăng cường, γ là hệ số chiết khấu và $Q(s(t), a(t))$ là hàm Q của trạng thái trước đó. γ được đặt ở mức 0,7, do chú trọng hơn giá trị dài hạn.

Cuối cùng, các giá trị q có thể được cập nhật bằng một phương thức gốc gradient thông thường:

$$q[i, a_i] = q[i, a_i] + \eta \cdot \Delta Q \cdot \alpha_i(s(t)) \quad (3.11)$$

trong đó η là tỷ lệ học tập, có giá trị được đặt ở mức 0,5, có nghĩa là thông tin cũ được coi trọng ngang với thông tin mới.

Quá trình nói trên được lặp lại từ việc lựa chọn hành động cho đến khi đạt được sự hội tụ.

Bản tóm tắt của thuật toán FQL [14] được mô tả ở được trình bày dưới đây:

1. Khởi tạo giá trị q :

$$q[i, j] = 0, 1 \leq i \leq N \text{ and } 1 \leq j \leq A$$

2. Chọn một hành động cho mỗi quy tắc kích hoạt:

$$a_i = \begin{cases} \text{random}\{a_k, k = 1, 2, \dots, A\}, & \text{with probability } \mathfrak{T} \\ \text{argmax}_k q[i, k], & \text{with probability } 1 - \mathfrak{T} \end{cases}$$

3. Tính hành động tổng thể:

$$a(t) = \sum_{i=1}^N \alpha_i(s(t)) \cdot a_i(t)$$

4. Hàm Q được tính từ giá trị q hiện tại và mức độ chuẩn của quy tắc:

$$Q(s(t), a(t)) = \sum_{i=1}^N \alpha_i(s(t)) \cdot q[i, a_i]$$

5. Chờ hệ thống giải quyết đến bước trạng thái tiếp theo, $s(t+1)$

6. Quan sát tín hiệu tăng cường, $r(t+1)$, và tính toán giá trị của trạng thái mới được biểu thị bằng $V_t(s(t+1))$:

$$V_t(s(t+1)) = \sum_{i=1}^N \alpha_i(s(t+1)) \cdot \max_k q[i, a_k]$$

7. Tính giá trị tín hiệu lỗi:

$$\Delta Q = r(t+1) + \gamma V_t(s(t+1)) - Q(s(t), a(t))$$

8. Cập nhật giá trị q bởi phương thức gốc gradient thông thường:

$$q[i, a_i] = q[i, a_i] + \eta \cdot \Delta Q \cdot \alpha_i(s(t))$$

9. Lặp lại quá trình được mô tả ở trên bắt đầu từ bước 2. cho trạng thái hiện tại mới cho đến khi đạt được sự hội tụ.

3.4. Kết luận chương

KẾT LUẬN

1. Các kết quả đạt được

Luận văn tập trung nghiên cứu hệ thống mạng tự tổ chức SON, mạng không đồng nhất HetNets. Cụ thể luận văn đạt được các kết quả sau:

- Nghiên cứu tổng quan về hệ thống mạng tự tổ chức SON, định nghĩa, cấu trúc, các chức năng tự làm việc.
- Các kỹ thuật dựa trên trí tuệ nhân tạo AI để tự tối ưu hóa trên các mạng không đồng nhất HetNets: kỹ thuật học máy, các thuật toán lấy cảm hứng từ sinh học, hệ thống mờ.
- Phân tích lý thuyết về kiểm soát nhập cell cho hệ thống truy cập vô tuyến nhiều người thuê.
- Từ các lý thuyết và các kỹ thuật xử lý thuật toán đã nghiên cứu đưa ra đề xuất việc sử dụng kỹ thuật học máy trong đó hai thuật toán là học tập được giám sát và học tập không được giám sát (Fuzzy Q-Learning) trong triển khai giả thuyết điều khiển tự chấp nhận dịch vụ trong mạng 5G.
- Cài đặt MatLab, khảo sát kết quả thuật toán Fuzzy Q-Learning.

2. Hướng phát triển

Luận văn có thể được phát triển theo hướng nghiên cứu áp dụng mô hình thử nghiệm triển khai trong môi trường thực. Lấy kết quả kiểm thử, triển khai lắp đặt hạ tầng mạng 5G của một nhà mạng cụ thể là của Tổng Công ty viễn thông MobiFone để đánh giá chính xác hơn các lập luận lý thuyết đã được nghiên cứu.