



(12) BẢN MÔ TẢ SÁNG CHẾ THUỘC BẰNG ĐỘC QUYỀN SÁNG CHẾ

(19) Cộng hòa xã hội chủ nghĩa Việt Nam (VN)
CỤC SỞ HỮU TRÍ TUỆ

(11)



1-0036565

(51)^{2020.01} H04B 17/00

(13) B

(21) 1-2020-07502

(22) 24/12/2020

(45) 25/08/2023 425

(43) 26/04/2021 397

(73) TẬP ĐOÀN CÔNG NGHIỆP - VIỄN THÔNG QUÂN ĐỘI (VN)

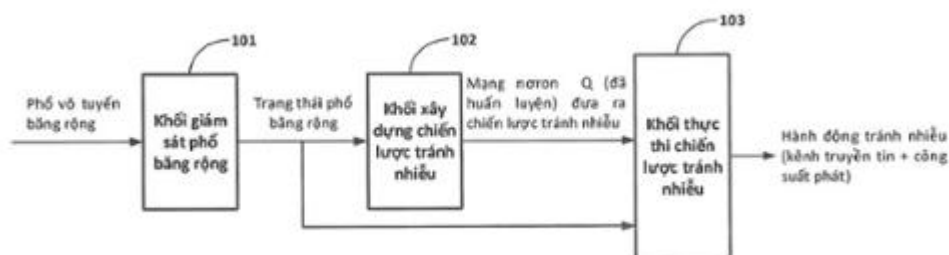
Lô D26 Khu đô thị mới Cầu Giấy, phường Yên Hoà, quận Cầu Giấy, thành phố Hà Nội

(72) NGUYỄN PHAN KHÁNH HÀ (VN); NGUYỄN VIỆT HÙNG (VN); ĐỖ VĂN LỘNG (VN); ĐỖ TRƯỜNG XUÂN (VN).

(74) Công ty TNHH Tư vấn Quốc Dân (NACILAW)

(54) HỆ THỐNG VÀ PHƯƠNG PHÁP TỰ ĐỘNG TRÁNH NHIỀU THÍCH NGHI TRONG BĂNG RỘNG

(57) Hệ thống và phương pháp tự động tránh nhiễu thích nghi trong băng rộng sử dụng thuật toán học sâu tăng cường dựa trên kiến trúc mạng nơron Double-Q. Cụ thể là, hệ thống và phương pháp tự động tránh nhiễu thích nghi trong băng rộng sử dụng các khối: khối giám sát phổ băng rộng, khối xây dựng chiến lược tránh nhiễu và khối thực thi chiến lược tránh nhiễu; được thể hiện qua các bước: bước 1 tính toán mật độ phổ công suất và ước lượng ngưỡng phát hiện tín hiệu trong băng rộng; bước 2: xác định trạng thái và tính toán công suất các kênh trong băng rộng; bước 3: xây dựng vectơ trạng thái và thông tin trạng thái phổ băng rộng; bước 4: xây dựng chiến lược tránh nhiễu tối ưu bằng cách huấn luyện mạng nơron Q; bước 5: thực thi chiến lược tránh nhiễu đã xây dựng tại mỗi khung thời gian sau quá trình huấn luyện mạng nơron Q ở bước 4.



Lĩnh vực kỹ thuật được đề cập

Sáng chế đề cập tới hệ thống và phương pháp tự động tránh nhiễu thích nghi trong băng rộng. Cụ thể, hệ thống và phương pháp tự động tránh nhiễu thích nghi trong băng rộng được đề cập trong sáng chế nhằm đưa ra chiến lược tránh nhiễu tối ưu và thực thi hành động tránh nhiễu dựa trên phương pháp học sâu tăng cường với kiến trúc mạng nơ-ron Double Q để đối phó với các loại thiết bị gây nhiễu, chế áp thông minh trong môi trường vô tuyến băng rộng phức tạp.

Tình trạng kỹ thuật của sáng chế

Vô tuyến nhận thức (Cognitive Radio) trong những năm gần đây đã nổi lên như một công nghệ tiên tiến cho phép các thiết bị thông tin vô tuyến có khả năng nhận thức về môi trường vô tuyến xung quanh và hoạt động thích nghi nhằm tối ưu hiệu quả sử dụng phổ, nâng cao hiệu quả truyền tin... Do đặc tính chia sẻ của phổ tần số băng rộng cũng như tính chất dễ bị tác động của đường truyền vô tuyến và sự phát triển ngày càng tinh vi của các thiết bị chế áp thông minh, các hình thức tấn công gây nhiễu trở thành nguy cơ lớn nhất làm suy giảm nghiêm trọng hiệu năng của các hệ thống vô tuyến nhận thức. Vì vậy, tự động tránh nhiễu thích nghi trong băng rộng trở thành nội dung cốt lõi trong nền tảng công nghệ vô tuyến nhận thức để bảo đảm thiết bị thông tin có khả năng thích ứng nhanh khi bị đối phương chế áp với khả năng tự động phân tích, đánh giá môi trường vô tuyến xung quanh và lựa chọn được kênh truyền tốt, tham số hoạt động tối ưu để đối phó hiệu quả với việc bị gây nhiễu.

Các phương pháp tránh nhiễu thích nghi là chủ đề đã được quan tâm nghiên cứu trong thời gian dài. Các phương pháp truyền thống bao gồm các kỹ thuật chống nhiễu dựa trên trải phổ như trải phổ nhảy tần (Frequency Hopping Spread Spectrum - FHSS), trải phổ chuỗi trực tiếp (Direct-Sequence Spread Spectrum - DSSS) và trải phổ kết hợp FHSS và DSSS. Các phương pháp này đòi hỏi sử dụng phổ với băng thông rộng nên hiệu suất sử dụng phổ không cao. Bên cạnh đó, để thực hiện các kỹ thuật trải phổ, thiết bị thông tin cần truyền tin với công suất cao, dẫn đến tiêu tốn năng

lượng của thiết bị. Ngoài ra, các thiết bị đầu cuối có khả năng thực hiện trải phổ thường có thiết kế, cấu trúc thiết bị phức tạp.

Một số phương pháp khác dựa trên lý thuyết trò chơi (game-theory) - là công cụ toán học để mô hình hóa và phân tích sự tương tác giữa thiết bị gây nhiễu và thiết bị truyền tin cũng được đề xuất để giải quyết vấn đề tránh nhiễu thích nghi trong băng rộng ví dụ như dựa trên cơ chế trò chơi tổng bằng 0 ngẫu nhiên (stochastic zero-sum game) hoặc lý thuyết trò chơi Stackelberg... Điểm hạn chế của các phương pháp này là chủ yếu dựa vào các mô hình toán học, thống kê cụ thể, có độ phức tạp cao và hiệu quả bị hạn chế trong các môi trường truyền tin bất định với số lượng trạng thái và hành động lớn.

Các phương pháp học tăng cường dựa trên thuật toán Q -learning và các biến thể của nó như Minimax- Q , Nash- Q , Friend-or-Foe Q và WoLF- Q ... cũng được nghiên cứu và đề xuất cho việc xây dựng các chiến lược tránh nhiễu tối ưu. Tuy nhiên, điểm hạn chế của các giải pháp này là chỉ hoạt động hiệu quả với số lượng trạng thái và hành động nhỏ (hữu hạn); trong thực tế, hiện nay môi trường vô tuyến băng rộng ngày càng phức tạp với không gian trạng thái và hành động rất lớn, thậm chí vô hạn thì hiệu quả của các phương pháp này là không cao.

Bản chất kỹ thuật của sáng chế

Mục đích thứ nhất của sáng chế là đề xuất một hệ thống tự động tránh nhiễu thích nghi trong băng rộng, để đạt được mục đích trên, hệ thống được đề xuất gồm các khối: khối giám sát phổ băng rộng, khối xây dựng chiến lược tránh nhiễu và khối thực thi chiến lược tránh nhiễu. Trong đó:

- Khối giám sát phổ băng rộng: xác định trạng thái phổ băng rộng (chứa thông tin về trạng thái và công suất chuẩn hóa của tất cả các kênh trong băng rộng) ở một khung thời gian.
- Khối xây dựng chiến lược tránh nhiễu: có chức năng huấn luyện mạng nơ-ron Q qua một số khung thời gian huấn luyện để xây dựng chiến lược tránh nhiễu tối ưu.

- Khối thực thi chiến lược tránh nhiễu: có chức năng lựa chọn hành động (kênh truyền và công suất phát) cho thiết bị dựa trên chiến lược tránh nhiễu đã xây dựng bởi khối xây dựng chiến lược tránh nhiễu.

Mục đích thứ hai của sáng chế là đề xuất một phương pháp tự động tránh nhiễu thích nghi trong băng rộng, phương pháp này bao gồm các bước như sau:

- Bước 1: xác định mật độ phổ công suất và ước lượng ngưỡng phát hiện tín hiệu trong băng rộng;
- Bước 2: xác định trạng thái và công suất các kênh trong băng rộng;
- Bước 3: xây dựng vectơ trạng thái và thông tin trạng thái phổ băng rộng;
- Bước 4: xây dựng chiến lược tránh nhiễu tối ưu bằng cách huấn luyện mạng nơron Q ;
- Bước 5: thực thi chiến lược tránh nhiễu; tại bước này, tại mỗi khung thời gian, với trạng thái phổ đầu vào, thiết bị truyền tin lựa chọn kênh truyền và công suất phát dựa trên đầu ra của mạng nơron Q đã được huấn luyện.

Trong sáng chế này, do sử dụng kiến trúc mạng nơron để xấp xỉ hàm Q tối ưu thay vì sử dụng hình thức bảng tra cứu (look-up table) như các thuật toán Q -learning truyền thống, giải pháp đề xuất có thể hoạt động hiệu quả trong điều kiện môi trường vô tuyến phức tạp, bất định với số lượng trạng thái rất lớn và mô hình chuyển đổi trạng thái là không biết trước. Bên cạnh đó, việc sử dụng mạng nơron theo kiến trúc Double- Q giúp cho quá trình huấn luyện để xây dựng chiến lược tránh nhiễu tối ưu đạt được hiệu quả ổn định, khắc phục các sai số do vấn đề đánh giá cực đại quá cao (maximum over-estimation) thường gặp phải bởi các mạng nơron Deep- Q thông thường.

Trong sáng chế này, giải pháp đề xuất có tính tổng quát, có thể áp dụng mở rộng cho các bài toán ra quyết định trong các môi trường phức tạp với số lượng trạng thái lớn và không xác định trước.

Mô tả vắn tắt các hình vẽ

Hình 1: hình ảnh minh họa kiến trúc hệ thống tự động tránh nhiễu thích nghi trong băng rộng sử dụng thuật toán học sâu tăng cường Double- Q .

Hình 2: hình ảnh minh họa kiến trúc khối giám sát phổ băng rộng.

Hình 3: hình ảnh minh họa kiến trúc khối xây dựng vectơ trạng thái phổ.

Hình 4: hình ảnh minh họa vectơ trạng thái phổ ở một khung thời gian.

Hình 5: hình ảnh minh họa trạng thái phổ băng rộng ở một khung thời gian.

Hình 6: hình ảnh minh họa kiến trúc khối xây dựng chiến lược tránh nhiễu.

Hình 7: hình ảnh minh họa kiến trúc khối thực thi chiến lược tránh nhiễu.

Mô tả chi tiết sáng chế

Nội dung sáng chế được mô tả chi tiết dưới đây với những thuật ngữ chuyên môn được hiểu như sau:

Kênh: là một phạm vi trên băng tần số, được giới hạn bởi tần số bắt đầu và tần số kết thúc. Do tài nguyên tần số là giới hạn nên các thiết bị thông tin cần phải truyền trên các kênh được cấp phát với độ rộng nhất định, để tránh ảnh hưởng đến các thiết bị thông tin khác. Ví dụ, một thiết bị thông tin hoạt động ở dải tần 50MHz, trên một băng rộng 20MHz, sẽ được cấp phát kênh truyền với độ rộng 1MHz, như vậy băng rộng 20MHz có thể chia thành hai mươi kênh khác nhau (mỗi kênh 1MHz). Tại một thời điểm, thiết bị thông tin chỉ có thể sử dụng một kênh cụ thể với độ rộng 1MHz. Đối với các thiết bị gây nhiễu, do chủ đích là gây nhiễu, phá hoại nên có thể truyền đồng thời trên nhiều kênh khác nhau, thậm chí trên toàn bộ băng rộng.

Phổ băng rộng: là cách biểu diễn tín hiệu trong miền tần số của tín hiệu. Phổ tín hiệu cho biết sự tồn tại và mức độ mạnh yếu của các tín hiệu (công suất) tại các tần số khác nhau trong phạm vi biểu diễn.

Phổ vô tuyến băng rộng: là biểu diễn phổ tín hiệu trong miền tần số trong toàn bộ một phạm vi băng rộng.

Giá trị phần thưởng: là thuật ngữ được sử dụng cho các mạng và thuật toán học sâu tăng cường (Reinforcement Learning) mà mạng nơron Double-Q sử dụng. Môi trường thường chứa đựng một nhiệm vụ được xác định rõ ràng và có thể cung cấp cho tác nhân (mạng nơron) một phần thưởng như một đáp ứng trực tiếp cho các hành động

của tác nhân (mạng nơron). Phần thưởng này là phản hồi về mức độ hiệu quả của hành động cuối cùng của tác nhân trong việc nó góp phần đạt được nhiệm vụ. Hàm chịu trách nhiệm về ánh xạ này được gọi là hàm phần thưởng (reward function) hoặc xác suất phần thưởng (reward probabilities). Mục tiêu của tác nhân trong quá trình được huấn luyện là tối đa hóa phần thưởng tổng thể mà nó nhận được và do đó, phần thưởng là động lực mà tác nhân (mạng nơron) cần để thực hiện hành vi mong muốn.

Công suất của các kênh: được tính dựa trên phổ bằng cách tính tích lũy (cộng) các giá trị điểm trong phổ tại các vị trí tần số tương ứng của phổ nằm trong kênh đó. Dựa trên việc tính toán, quan sát và phân tích phổ băng rộng, có thể xác định được các vị trí tần số hoặc kênh nào đang có tín hiệu/nhiều tồn tại.

Giả thiết rằng toàn bộ phổ băng rộng được phân chia thành N kênh truyền có cùng băng thông và không chồng lấn nhau. Thiết bị truyền tin và thiết bị gây nhiễu hoạt động trên cùng phổ tần chia sẻ này và truyền tin/gây nhiễu trong các khung thời gian (timeslot) bằng nhau. Thiết bị truyền tin tại mỗi khung thời gian chỉ có thể truyền tin trên một kênh, trong khi thiết bị gây nhiễu có thể gây nhiễu đồng thời trên nhiều kênh khác nhau. Thiết bị truyền tin và thiết bị gây nhiễu có thể điều chỉnh công suất phát của mình ở các mức công suất khác nhau tại các khung thời gian khác nhau; tuy nhiên, trong một khung thời gian, công suất phát là cố định. Quá trình truyền tin của thiết bị truyền tin được xem là thành công (tránh được nhiễu) khi và chỉ khi truyền trên kênh không bị gây nhiễu hoặc truyền trên kênh bị gây nhiễu nhưng có công suất phát đủ lớn hơn công suất gây nhiễu một ngưỡng xác định.

Mục đích của sáng chế là đề xuất hệ thống và phương pháp cho phép thiết bị truyền tin đưa ra chiến lược tránh nhiễu tối ưu (gồm lựa chọn kênh truyền truy nhập và công suất phát nhỏ nhất có thể) để đạt được hiệu quả truyền tin cao nhất.

Để đạt được mục đích thứ nhất của sáng chế là đề xuất một hệ thống tự động tránh nhiễu thích nghi trong băng rộng sử dụng thuật toán học sâu tăng cường Double- Q , hệ thống này gồm các khối: khối giám sát phổ băng rộng, khối xây dựng chiến lược tránh nhiễu và khối thực thi chiến lược tránh nhiễu như trong Hình 1, cụ thể:

- Khối giám sát phổ băng rộng 101:

Chức năng của khối giám sát phổ băng rộng là cho phép thiết bị truyền tin có khả năng phân tích, đánh giá môi trường vô tuyến xung quanh, cung cấp thông tin đầu vào trong quá trình xây dựng cũng như thực thi chiến lược tránh nhiễu. Dữ liệu đầu vào của khối này là phổ vô tuyến băng rộng tại mỗi khung thời gian. Khối giám sát phổ băng rộng 101 thực hiện tính toán và đưa ra thông tin về trạng thái phổ là đầu vào cho các khối xây dựng chiến lược tránh nhiễu 102 và khối thực thi chiến lược tránh nhiễu 103.

Tham chiếu Hình 2, khối giám sát phổ băng rộng 101 gồm có ba khối con:

(a) Khối xây dựng vectơ trạng thái phổ 201: nhận tín hiệu vô tuyến băng rộng đã được lấy mẫu và số hóa bởi bộ thu vô tuyến của thiết bị (sử dụng ăngten dải rộng, phần cứng thu cao tần và xử lý số), thực hiện xử lý để đưa ra vectơ trạng thái phổ băng rộng của khung thời gian hiện tại V_i chứa thông tin về trạng thái (kênh rảnh, không bị chiếm dụng hoặc đang bị chiếm dụng) và công suất của các kênh ở khung thời gian hiện tại. Trong đó, vectơ trạng thái phổ này sẽ được lưu vào khối lưu trữ vectơ trạng thái phổ 202 phục vụ cho việc xử lý ở các khung thời gian tiếp theo, đồng thời là đầu vào cho khối xây dựng trạng thái phổ 203. Thành phần của khối xây dựng vectơ trạng thái phổ 201 gồm các khối như tham chiếu trong Hình 3 gồm:

- + Khối tính toán mật độ phổ công suất 301.
- + Khối ước lượng ngưỡng phát hiện tín hiệu 302.
- + Khối tính toán công suất 303.
- + Khối lưu trữ hành động của thiết bị 304.
- + Khối xác định trạng thái kênh 305.
- + Khối chuẩn hóa công suất 306.

Cấu tạo cụ thể của các khối này như sau:

+ Khối tính toán mật độ phổ công suất 301: đầu vào của khối này là khối dữ liệu IQ của tín hiệu được lấy mẫu từ phổ vô tuyến băng rộng, đầu ra của khối này là khối dữ liệu biểu diễn mật độ phổ công suất (Power Spectrum Density - PSD) băng rộng đầu vào. Giá trị mật độ phổ công suất tính toán được sẽ là đầu vào cho khối ước lượng ngưỡng phát hiện tín hiệu 302 và khối tính toán công suất 303. Khối tính toán mật độ phổ công suất 301 thực hiện tính toán phép biến đổi Fourier nhanh (Fast Fourier Transform - FFT) cho dữ liệu đầu vào $s[n]$, từ kết quả phép biến đổi này, giá

trị mật độ phổ công suất sẽ được suy ra. Phép biến đổi Fourier nhanh được sử dụng theo công thức (CT1) dưới đây (công thức này được đề xuất trong bài báo “Một thuật toán cho tính toán các chuỗi Fourier phức bằng máy tính” (“An Algorithm for the Machine Calculation of Complex Fourier Series”) của J. W. Cooley và J. W. Tukey công bố trên tạp chí “Toán học của sự tính toán” (“Mathematics of Computation”), quyển 19, số 90, xuất bản năm 1965 (từ trang 297 đến trang 301):

$$Power_{linear}[k] = \left| \sum_{n=0}^{L-1} s[n] \times e^{-\frac{j2\pi kn}{L}} \right|^2$$

$$PSD_{linear}[k] = \frac{Power_{linear}[k]}{(F_s \times L)}$$

$$PSD_{dBm}[k] = 10 \times \log_{10}(PSD_{linear}[k]) + 30 \text{ với } k = 0, 1, \dots, L-1 \text{ (CT1)}$$

trong đó, F_s là tần số lấy mẫu phổ băng rộng, L là số điểm tính toán FFT.

+ Khối ước lượng ngưỡng phát hiện tín hiệu 302: khối này có chức năng tính toán ngưỡng phát tín hiệu làm cơ sở cho việc xác định trạng thái kênh truyền (đang bị chiếm dụng bởi các nguồn phát vô tuyến hoặc rảnh, không bị chiếm dụng) dựa trên áp dụng giải pháp được đề xuất trong bài báo: “Tự động ước lượng nền nhiễu của phổ trong sự hiện diện của các tín hiệu” (“Automatic Noise Floor Spectrum Estimation in the Presence of Signals”) của M. J. Ready, M. L. Downey và L. J. Corbalis công bố trên kỷ yếu “Hội nghị ASILOMAR” xuất bản năm 1997 (từ trang 877 đến trang 881) để ước lượng nền nhiễu.

Nền nhiễu được định nghĩa là tổng các nguồn nhiễu và tín hiệu không mong muốn đo được trong một hệ thống thu. Do việc ước lượng nền nhiễu thường có sai lệch và các tín hiệu thu bị thăng giáng nhanh, thay vì trực tiếp dùng nền nhiễu để phát hiện tín hiệu, người ta thường định nghĩa ngưỡng phát hiện tín hiệu bằng nền nhiễu cộng thêm một giá trị dịch (offset) theo thang dB. Các thành phần phổ có mật độ phổ công suất lớn hơn hoặc bằng giá trị ngưỡng phát hiện sẽ được xem là tín hiệu. Ngưỡng phát hiện tín hiệu là đầu ra của khối ước lượng ngưỡng phát hiện tín hiệu 302.

+ Khối tính toán công suất 303: khối này nhận đầu vào là dữ liệu mật độ phổ công suất băng rộng và thực hiện tính toán công suất của từng kênh truyền trong

phổ băng rộng bằng cách cộng tích lũy các giá trị điểm mật độ phổ công suất (PSD bin) tương ứng ở từng kênh truyền. Đầu ra của khối này là N giá trị công suất ứng với N kênh trong phổ băng rộng.

+ Khối chuẩn hóa công suất 304: khối này có chức năng thực hiện chuẩn hóa giá trị công suất của các kênh truyền đã được tính toán bởi khối tính toán công suất 303 về trong cùng khoảng giá trị $[-1; 1]$ để thuận lợi cho các quá trình huấn luyện, cũng như thực thi tiếp theo của mạng nơron Q .

+ Khối lưu trữ hành động của thiết bị 305: chức năng của khối này là lưu lại thông tin về lựa chọn kênh truyền của thiết bị trong các khung thời gian trước đó. Thông tin đầu vào được thu nhận từ khối xây dựng chiến lược tránh nhiễu 102 trong giai đoạn xây dựng chiến lược tránh nhiễu hoặc từ khối thực thi chiến lược tránh nhiễu 103 trong giai đoạn thực thi. Đầu ra của khối này là thông tin về kênh truyền của thiết bị ở khung thời gian trước đó $(t - 1)$, được gửi đến khối xác định trạng thái kênh 306 để xây dựng vectơ trạng thái phổ băng rộng V_t cho khung thời gian hiện tại t .

+ Khối xác định trạng thái kênh 306: đầu vào của khối này là công suất tính toán của các kênh và ngưỡng phát hiện tín hiệu tại khung thời gian hiện tại t phản ánh tình trạng của các kênh trong phổ băng rộng ở khung thời gian t . Ngoài ra, khối xác định trạng thái kênh 306 cũng nhận thông tin về lựa chọn kênh truyền của thiết bị ở khung thời gian trước đó $t - 1$ từ khối lưu trữ hành động của thiết bị 305. Đầu ra của khối là thông tin mô tả trạng thái của một kênh i tại khung thời gian t , $cs_{t,i}$.

Tham chiếu Hình 4, vectơ trạng thái phổ băng rộng tại khung thời gian hiện tại V_t được xây dựng dựa trên thông tin đầu ra của khối xác định trạng thái kênh 306 và khối chuẩn hóa công suất 304 mang thông tin về trạng thái $(cs_{t,i})$ và công suất chuẩn hóa $(cp_{t,i})$ của tất cả các kênh truyền trong phổ băng rộng tại khung thời gian t .

(b) Khối lưu trữ vectơ trạng thái phổ 202 lưu trữ các vectơ trạng thái phổ băng rộng của các khung thời gian liên tiếp nhau. Khối này có thể được tổ chức dưới dạng hàng đợi kiểu FIFO (First-In, First-Out) và có khả năng lưu trữ N_{ts} vectơ trạng thái phổ băng rộng ứng với N_{ts} khung thời gian liên tiếp gần nhất trước đó.

(c) Khối xây dựng trạng thái phổ 203 tạo ra trạng thái phổ băng rộng ở khung thời gian hiện tại s_t bằng cách ghép nối vectơ trạng thái phổ băng rộng của khung thời gian hiện tại V_t và các vectơ trạng thái phổ của N_{ts} khung thời gian gần nhất trước đó như minh họa trong Hình 5.

- Khối xây dựng chiến lược tránh nhiễu 102:

Tham chiếu Hình 6, khối xây dựng chiến lược tránh nhiễu 102 thực hiện huấn luyện mạng nơron Q để tìm ra chiến lược tránh nhiễu tối ưu thông qua quá trình tương tác với môi trường vô tuyến xung quanh trong một số khung thời gian huấn luyện. Số lượng khung thời gian dùng cho quá trình huấn luyện cần được lựa chọn đủ lớn để đảm bảo sự hội tụ của thuật toán và tìm được chiến lược tránh nhiễu tối ưu trước khi tiến hành thực thi ở khối thực thi chiến lược tránh nhiễu 103. Khối này gồm có các khối con: khối xử lý các thông tin tương tác với môi trường 601, khối lưu trữ các thông tin kinh nghiệm 602, mạng nơron Q dự đoán 603, mạng nơron Q mục tiêu 604 và khối cập nhật bộ trọng số cho mạng nơron Q 604.

Khối xây dựng chiến lược tránh nhiễu 102 sử dụng kiến trúc mạng nơron Double- Q gồm hai mạng nơron Q khác nhau gồm một mạng nơron Q dự đoán 603 được sử dụng cho quá trình lựa chọn hành động và một mạng nơron Q mục tiêu 604 được sử dụng cho đánh giá hành động. Hai mạng nơron này sử dụng chung một kiểu kiến trúc mạng nơron tích chập (Convolution Neural Network - CNN). Kiến trúc mạng CNN này gồm có các lớp: lớp đầu vào có kích thước bằng với kích thước trạng thái phổ băng rộng ở một khung thời gian; hai lớp tích chập (convolution) đi kèm với hai lớp tổng hợp cực đại (max pooling) để trích chọn các đặc trưng; tiếp theo là hai lớp kết nối đầy đủ (dense); đầu ra của mạng là một vectơ biểu diễn xấp xỉ giá trị hàm Q , có kích thước bằng số lượng hành động có thể có của thiết bị, hành động tương ứng có giá trị hàm Q lớn nhất được xem là hành động tối ưu cho việc tránh nhiễu.

Mục tiêu của việc huấn luyện mạng nơron Q dự đoán 603 trong mỗi khung thời gian t là lựa chọn một hành động a_t trong không gian hành động của thiết bị A ($a_t \in A$) sao cho tối đa hóa phần thưởng tích lũy (dự kiến) đạt được trong tương lai

$E_{\pi}[\sum_{t=0}^{\infty} \gamma^t r_t | s_0 = s]$ cho tất cả các trạng thái ban đầu $s \in S$, trong đó π là một chiến lược (policy) lựa chọn hành động. Mục tiêu này có thể đạt được thông qua việc tìm hàm giá trị hành động - trạng thái tối ưu $Q^*(s, a)$. Chiến lược tối ưu có thể được tìm thấy thông qua phép xấp xỉ của một hàm Q tối ưu được biểu diễn bởi mạng nơron Q dự đoán.

Tại mỗi khung thời gian, dựa trên trạng thái phổ băng rộng hiện tại s_{t-1} , mạng nơron Q dự đoán 603 tính toán đưa ra hành động tương ứng a_t (kênh truyền và công suất phát) dựa trên chính sách tham lam với xác suất epsilon (ϵ -greedy policy): với một xác suất ϵ , lựa chọn ngẫu nhiên một hành động trong không gian hành động A , ngược lại, chọn một hành động trong không gian hành động tương ứng có giá trị hành động - trạng thái $Q^*(s, a)$ lớn nhất.

Thiết bị thực hiện hành động (truyền tin trên kênh truyền với công suất lựa chọn) và tạo ra sự tương tác với môi trường vô tuyến băng rộng xung quanh. Dựa trên sự tương tác này, khối xử lý các thông tin tương tác với môi trường 601 tính toán được giá trị phần thưởng (reward) r_t mà thiết bị có thể đạt được theo hàm phần thưởng được định nghĩa như sau:

Nếu ($\exists$ kênh truyền i sao cho $c_{t+1,i} = 0$) thì

Nếu ($u = i$) thì

Nếu ($p_u = p_l | l = 0$) thì $r_t = 1,0$

Nếu ($p_u = p_l | l = 1$) thì $r_t = 0,5$

Nếu ($p_u = p_l | 1 < l < L - 1$) thì $r_t = 0,05$

Nếu ($u \neq i$) thì $r_t = 0$

Nếu ($\nexists$ kênh truyền i sao cho $c_{t+1,i} = 0$) thì

Nếu ($p_u = p_l | l = l_{min}$) sao cho $(p_u - p_j) \geq \beta_{th}$ thì $r_t = 1,0$, ngược lại $r_t = 0$

trong đó, thiết bị truyền tin có thể lựa chọn mức công suất phát l trong tập L mức công suất có thể phát của thiết bị; p_u và p_j lần lượt là công suất phát của thiết bị truyền tin và thiết bị gây nhiễu.

Mục đích của hàm phần thưởng là tối đa hóa xác suất truyền tin thành công của thiết bị và duy trì được mức tiêu tốn công suất phát thiết bị nhỏ nhất có thể (truyền được vào các kênh không bị chiếm dụng nếu tồn tại kênh này hoặc truyền vào các kênh đang bị gây nhiễu nhưng với một mức công suất vừa đủ lớn hơn công suất nhiễu bảo đảm truyền tin thành công).

Với tương tác này, khối xử lý các thông tin tương tác với môi trường 601 cũng nhận được trạng thái phổ băng rộng ở khung thời gian tiếp theo s_t . Bộ tham số (s_{t-1}, a_t, r_t, s_t) được gọi là kinh nghiệm (experience) sẽ được lưu vào một bộ nhớ (replay memory) tại khối lưu trữ các thông tin kinh nghiệm 602 sau mỗi khung thời gian huấn luyện. Khi số lượng kinh nghiệm trong bộ nhớ này đủ lớn, một tập con (minibatch) B gồm k kinh nghiệm $\{(s_{t-1}, a_t, r_t, s_t)\}$ được lấy mẫu ngẫu nhiên từ bộ nhớ. Các trọng số (weights) θ_t của mạng nơron Q dự đoán sẽ được cập nhật với tốc độ học (learning rate) α theo nguyên lý “đi ngược đạo hàm” (gradient-descent) để tối thiểu hóa hàm lỗi (loss function) $L(\theta_t)$ được định nghĩa theo công thức (CT2):

$$L(\theta_t) = \frac{1}{k} \sum_{t \in B} (r_t + \gamma Q(s_t, \operatorname{argmax}_a Q(s_t, a, \theta_t), \theta_t^-) - Q(s_{t-1}, a_t, \theta_t))^2 \quad (\text{CT2})$$

trong đó, γ là hệ số giảm (discount factor) đặc trưng cho mức độ mức độ quan trọng của giá trị phần thưởng trong tương lai, γ nhận giá trị trong khoảng $(0, 1)$ và γ càng nhỏ thì độ quan trọng của giá trị phần thưởng trong tương lai càng thấp.

Để đồng bộ hai mạng nơron Q , sau mỗi chu kỳ (gồm U khung thời gian trong quá trình huấn luyện), bộ trọng số θ đã được huấn luyện của mạng nơron Q dự đoán sẽ được sao chép sang thành bộ trọng số θ^- của mạng nơron Q mục tiêu.

Kết thúc quá trình huấn luyện, đầu ra của khối xây dựng chiến lược chống nhiễu là kiến trúc mạng nơron Q dự đoán với bộ trọng số θ sau huấn luyện bảo đảm tối thiểu hóa hàm lỗi $L(\theta_t)$ - cũng có nghĩa bảo đảm lựa chọn được hành động tối ưu để tránh nhiễu thích nghi trong băng rộng.

- Khối thực thi chiến lược tránh nhiễu 103.

Tham chiếu Hình 7, khối thực thi chiến lược tránh nhiễu 103 là mạng nơron Q dự đoán với bộ trọng số θ đã được huấn luyện bởi khối xây dựng chiến lược tránh

nhiều 102. Khối thực thi chiến lược tránh nhiễu 103 có chức năng đề xuất hành động tối ưu cho thiết bị gồm kênh truyền tin và công suất phát dựa theo chiến lược tránh nhiễu đã xây dựng nhằm tránh nhiễu, tối đa hóa hiệu suất truyền tin thành công và tối thiểu hóa mức tiêu tốn công suất phát.

Để đạt được mục đích thứ hai của sáng chế là đề xuất phương pháp tự động tránh nhiễu thích nghi trong băng rộng, phương pháp này bao gồm các bước như sau: bước 1: tính toán mật độ phổ công suất và ước lượng ngưỡng phát hiện tín hiệu trong băng rộng; bước 2: xác định trạng thái và công suất các kênh trong băng rộng; bước 3: xây dựng vectơ trạng thái và thông tin trạng thái phổ băng rộng; bước 4: thực hiện huấn luyện mạng nơron Q trong một số khung thời gian huấn luyện; bước 5: lựa chọn hành động (kênh và công suất truyền) để tránh nhiễu dựa trên kết quả đầu ra của mạng nơron Q đã được huấn luyện.

Phương pháp tự động tránh nhiễu thích nghi trong băng rộng sử dụng thuật toán học sâu tăng cường Double- Q được đề xuất dựa trên sử dụng hệ thống nói trên cụ thể như sau:

- Bước 1: tính toán mật độ phổ công suất và ước lượng ngưỡng phát hiện tín hiệu trong băng rộng.

Dữ liệu IQ của phổ vô tuyến băng rộng được đưa vào khối tính toán mật độ phổ công suất 301 để tính toán mật độ phổ công suất băng rộng (PSD) dựa trên việc thực hiện phép biến đổi Fourier nhanh (FFT) theo (CT1).

Dữ liệu mật độ phổ công suất tính toán sau đó được đưa vào khối ước lượng ngưỡng phát hiện tín hiệu 302. Khối này áp dụng thuật toán ước lượng nền nhiễu dựa trên các kỹ thuật xử lý ảnh để ước lượng nền nhiễu băng rộng. Dựa trên nền nhiễu, ngưỡng phát hiện tín hiệu được xác định bằng cách cộng giá trị nền nhiễu với một giá trị dịch (offset) nhất định được cấu hình ở khối ước lượng ngưỡng phát hiện tín hiệu 302.

- Bước 2: xác định trạng thái và công suất các kênh trong băng rộng.

Công suất của các kênh trong băng rộng được tính toán tại khối tính toán công suất 303 dựa trên đầu vào là mật độ phổ công suất băng rộng, cụ thể giá trị công suất

của một kênh được tính bằng cách cộng tích lũy giá trị mật độ phổ công suất của tất cả các điểm phổ (*bin*) tương ứng với kênh đó.

Công suất của từng kênh sau tính toán sẽ được thực hiện chuẩn hóa bởi khối chuẩn hóa công suất 304 về cùng dải giá trị $[-1, 1]$ sử dụng các giá trị công suất lớn nhất và nhỏ nhất có thể thu được trong một kênh.

Trạng thái của tất cả các kênh trong phổ băng rộng được xác định bởi khối xác định trạng thái kênh 306. Khối này quyết định trạng thái kênh dựa trên các thông tin đầu vào gồm công suất các kênh trước chuẩn hóa, ngưỡng phát hiện tín hiệu và kênh lựa chọn cho truyền tin của thiết bị ở khung thời gian ngay trước đó. Kênh truyền được đánh giá là không bị chiếm dụng (rảnh) hoặc bị chiếm dụng ở khung thời gian hiện tại nhờ vào việc thực hiện so sánh giữa công suất tính toán của kênh với ngưỡng phát hiện tín hiệu. Thông tin tổng hợp mô tả trạng thái của một kênh i tại khung thời gian t , $cs_{t,i}$ - được xây dựng theo công thức (CT3):

$$cs_{t,i} = \begin{cases} 0, & \text{nếu kênh } i \text{ rảnh ở } t \text{ và không được thiết bị chọn truyền ở } t-1 \\ 1, & \text{nếu kênh } i \text{ bị chiếm dụng ở } t \text{ và không được thiết bị chọn truyền ở } t-1 \\ 2, & \text{nếu kênh } i \text{ được thiết bị chọn truyền ở } t-1 \end{cases}$$

với $i = 0, 1, \dots, N-1$ (CT3)

- Bước 3: xây dựng vectơ trạng thái và thông tin trạng thái phổ băng rộng.

Khối xây dựng vectơ trạng thái phổ 201 tạo ra vectơ trạng thái phổ băng rộng tại khung thời gian t bằng cách tổng hợp thông tin trạng thái kênh và công suất kênh chuẩn hóa tại khung thời gian t của tất cả N kênh trong băng rộng và gửi dữ liệu này đến khối xây dựng trạng thái phổ 203.

Khối xây dựng trạng thái phổ 203 nhận dữ liệu vectơ trạng thái phổ băng rộng tại khung thời gian hiện tại t , đồng thời tổng hợp dữ liệu các vectơ trạng thái phổ băng rộng tại các khung thời gian trước đó lấy từ khối lưu trữ vectơ trạng thái phổ 202 (khối này về bản chất là một bộ nhớ dạng FIFO) để đưa ra thông tin trạng thái phổ băng rộng ở khung thời gian hiện tại - cũng là đầu ra của khối giám sát phổ băng rộng 101.

- Bước 4: thực hiện huấn luyện mạng nơron Q trong một số khung thời gian huấn luyện.

Quá trình huấn luyện mạng nơron Q được thực hiện bởi khối xây dựng chiến lược tránh nhiễu 102. Tại thời điểm ban đầu của quá trình huấn luyện, bộ nhớ trong

khối lưu trữ các thông tin kinh nghiệm được khởi tạo rỗng; bộ trọng số (weight) của các mạng nơron Q dự đoán và nơron Q mục tiêu được khởi tạo ngẫu nhiên; vectơ trạng thái phổ băng rộng và trạng thái phổ băng rộng ban đầu được khởi tạo 0. Tại mỗi khung thời gian trong quá trình huấn luyện, thông qua khối xử lý các thông tin tương tác với môi trường 601, trạng thái phổ băng rộng hiện tại sẽ được thu nhận, dựa trên thông tin này, hành động tương ứng được đưa ra bởi mạng nơron Q dự đoán 603. Giá trị phần thưởng (reward) và trạng thái phổ băng rộng tiếp theo được tính toán, thu nhận bởi khối xử lý thông tin tương tác với môi trường 601. Sau đó, toàn bộ thông tin kinh nghiệm trong khung thời gian huấn luyện đó được lưu vào bộ nhớ của khối lưu trữ các thông tin kinh nghiệm 602.

Khi số lượng kinh nghiệm lưu trong bộ nhớ đủ lớn, một tập con các kinh nghiệm được lấy mẫu ngẫu nhiên, làm đầu vào cho quá trình huấn luyện. Các mạng nơron Q dự đoán và mạng nơron Q mục tiêu tính toán các giá trị hàm Q và đưa vào khối cập nhật bộ trọng số mạng nơron Q 605. Tại khối này, đầu tiên giá trị hàm lỗi được tính toán theo công thức (CT2). Tiếp theo, sử dụng thuật toán “lan truyền ngược” (back-propagation) tham khảo trong bài báo: “Một nền tảng lý thuyết cho thuật toán lan truyền ngược” (“A Theoretical Framework for Back-propagation”) của Yann Le Cun công bố trên kỷ yếu “1988 Connectionist Models Summer School” xuất bản năm 1988 (từ trang 21 đến trang 28) và nguyên lý “đi ngược đạo hàm” (gradient descent) để tiến hành cập nhật lại bộ trọng số cho mạng nơron Q dự đoán với mục tiêu tối thiểu hóa hàm lỗi. Sau một số chu kỳ, bộ trọng số của mạng nơron Q dự đoán được sao chép sang mạng Q mục tiêu để đồng bộ giữa hai mạng.

Sau một số khung thời gian huấn luyện đủ lớn để bảo đảm thuật toán hội tụ, quá trình huấn luyện kết thúc, bộ trọng số sau huấn luyện của mạng nơron Q dự đoán sẽ được sử dụng ở bước 5 cho quá trình tính toán đưa ra hành động tránh nhiễu tối ưu cho thiết bị.

- Bước 5: lựa chọn hành động (kênh và công suất truyền) để tránh nhiễu dựa trên kết quả đầu ra của mạng nơron Q đã được huấn luyện.

Quá trình thực thi được thực hiện bởi khối thực thi chiến lược tránh nhiễu 103 sau khi kết thúc quá trình huấn luyện. Tại thời điểm bắt đầu của mỗi khung thời gian s_{t-1} , thông tin trạng thái phổ băng rộng sẽ được đưa vào khối thực thi chiến lược

tránh nhiễu 103, khối này tính toán hành động tránh nhiễu tối ưu dựa trên việc lựa chọn hành động tương ứng làm cho giá trị hàm Q - được xấp xỉ bởi đầu ra của mạng nơron Q dự đoán - đạt giá trị cực đại.

Thiết bị thực hiện theo đúng hành động này (truyền tin trên kênh với công suất phát tương ứng) trong khung thời gian để đạt hiệu quả tránh nhiễu tối ưu.

Yêu cầu bảo hộ

1. Hệ thống tự động tránh nhiễu thích nghi trong băng rộng, gồm các khối:

khối giám sát phổ băng rộng; khối này gồm có các khối con có chức năng nhận dữ liệu đầu vào là phổ vô tuyến băng rộng, thực hiện tính toán mật độ phổ công suất, ước lượng ngưỡng phát hiện tín hiệu, tính toán công suất các kênh trong phổ băng rộng và vectơ trạng thái phổ băng rộng - chứa thông tin về trạng thái (không bị chiếm dụng hoặc bị chiếm dụng) và công suất chuẩn hóa của tất cả các kênh trong băng rộng; từ dữ liệu vectơ trạng thái phổ băng rộng, khối này đưa ra thông tin về trạng thái phổ băng rộng ở khung thời gian thời gian hiện tại - là tập hợp của các vectơ trạng thái phổ băng rộng ở khung thời gian hiện tại và một số khung thời gian liên tiếp trước đó;

khối xây dựng chiến lược tránh nhiễu; khối này có chức năng huấn luyện mạng nơron Q qua một số khung thời gian huấn luyện để xây dựng chiến lược tránh nhiễu tối ưu; khối này gồm có các khối con là: khối xử lý các thông tin tương tác với môi trường để thu nhận các trạng thái phổ băng rộng ở khung thời gian hiện tại và khung thời gian tiếp theo từ khối giám sát phổ băng rộng, xử lý thực hiện hành động đưa ra bởi mạng nơron Q , tính toán giá trị phần thưởng thu được để tạo ra thông tin kinh nghiệm lưu vào khối lưu trữ kinh nghiệm; khối mạng nơron có kiến trúc Double- Q gồm một mạng nơron Q dự đoán cho lựa chọn hành động và một mạng nơron Q mục tiêu cho đánh giá hành động, hai mạng nơron này có cùng kiến trúc gồm các lớp: lớp đầu vào (input) có kích thước bằng với kích thước trạng thái phổ băng rộng ở một khung thời gian, hai lớp tích chập (convolution) đi kèm với với hai lớp tổng hợp cực đại (max pooling), hai lớp kết nối đầy đủ (dense) và lớp đầu ra (output) là một vectơ biểu diễn xấp xỉ hàm Q , có kích thước bằng số lượng hành động có thể có của thiết bị;

khối cập nhật bộ trọng số mạng nơron Q thực hiện quá trình huấn luyện bằng cách tính hàm lỗi và lần lượt cập nhật bộ trọng số cho mạng nơron Q với mục tiêu tối thiểu hóa hàm lỗi này qua từng khung thời gian;

khối thực thi chiến lược tránh nhiễu; khối này có chức năng lựa chọn hành động (kênh truyền và công suất phát) cho thiết bị dựa trên chiến lược tránh nhiễu đã được xây dựng, bản chất của khối này là một mạng nơron Q dự đoán với bộ trọng số

đã được huấn luyện; đầu vào của mạng là trạng thái phổ băng rộng ở một khung thời gian, mạng nơron tính toán và đưa ra một vectơ biểu diễn xấp xỉ hàm Q cho toàn bộ các hành động có thể thực hiện trong không gian hành động của thiết bị, hành động tương ứng có giá trị hàm Q lớn nhất được xem là hành động tối ưu và sẽ được lựa chọn để thực hiện cho thiết bị.

2. Hệ thống tự động tránh nhiễu thích nghi trong băng rộng theo điểm 1, trong đó khối xây dựng vectơ trạng thái phổ trong khối giám sát phổ băng rộng bao gồm các khối con:

khối tính toán mật độ phổ công suất; đầu vào của khối là dữ liệu phổ vô tuyến băng rộng, khối này thực hiện tính toán mật độ phổ công suất băng rộng dựa trên việc thực hiện tính toán phép biến đổi Fourier nhanh (FFT), đầu ra là mật độ phổ công suất băng rộng được tính toán;

khối ước lượng ngưỡng phát hiện tín hiệu; đầu vào của khối là mật độ phổ công suất băng rộng được tính toán bởi khối tính toán mật độ phổ công suất, ngưỡng phát hiện tín hiệu được ước lượng bằng một bộ xử lý thực hiện thuật toán tự động ước lượng nền nhiễu dựa trên các kỹ thuật xử lý ảnh phổ, đầu ra là ngưỡng phát hiện tín hiệu trong băng rộng;

khối tính toán công suất; đầu vào của khối là mật độ phổ công suất băng rộng được tính toán bởi khối tính toán mật độ phổ công suất, công suất của các kênh trong băng rộng được tính toán bởi một bộ xử lý thực hiện cộng tích lũy các giá trị mật độ phổ công suất tương ứng với từng kênh trong phổ, đầu ra là công suất được tính toán cho toàn bộ các kênh truyền;

khối chuẩn hóa công suất; đầu vào của khối là công suất của toàn bộ các kênh trong băng rộng, khối này thực hiện chuẩn hóa công suất cho các kênh về chung một dải giá trị $[-1,1]$ bằng một bộ xử lý tính toán chuẩn hóa;

khối lưu trữ hành động của thiết bị; khối này là một bộ nhớ lưu trữ lại thông tin về kênh truyền của thiết bị lựa chọn trong khung thời gian trước đó, nhận dữ liệu đầu vào từ khối xây dựng chiến lược tránh nhiễu trong quá trình huấn luyện hoặc từ khối

thực thi chiến lược tránh nhiễu trong quá trình thực thi, đầu ra của khối là thông tin về kênh truyền của thiết bị ở khung thời gian trước đó;

khối xác định trạng thái kênh; khối này đưa ra thông tin về trạng thái kênh (không bị chiếm dụng hoặc bị chiếm dụng) dựa trên các thông tin đầu vào gồm: công suất các kênh trước chuẩn hóa, ngưỡng phát hiện tín hiệu và thông tin về kênh truyền của thiết bị ở khung thời gian trước đó nhận từ khối tính toán công suất, khối ước lượng ngưỡng phát hiện tín hiệu và khối lưu trữ hành động của thiết bị.

3. Phương pháp tự động tránh nhiễu thích nghi trong băng rộng, bao gồm các bước:

bước 1: tính toán mật độ phổ công suất và ước lượng ngưỡng phát hiện tín hiệu trong băng rộng; cụ thể, đưa phổ vô tuyến băng rộng vào khối tính toán mật độ phổ công suất để thực hiện phép biến đổi Fourier nhanh và tính toán mật độ phổ công suất; sau đó, dựa trên mật độ phổ công suất, ước lượng ngưỡng phát hiện tín hiệu bằng thuật toán ước lượng nền nhiễu trong khối ước lượng ngưỡng phát hiện tín hiệu;

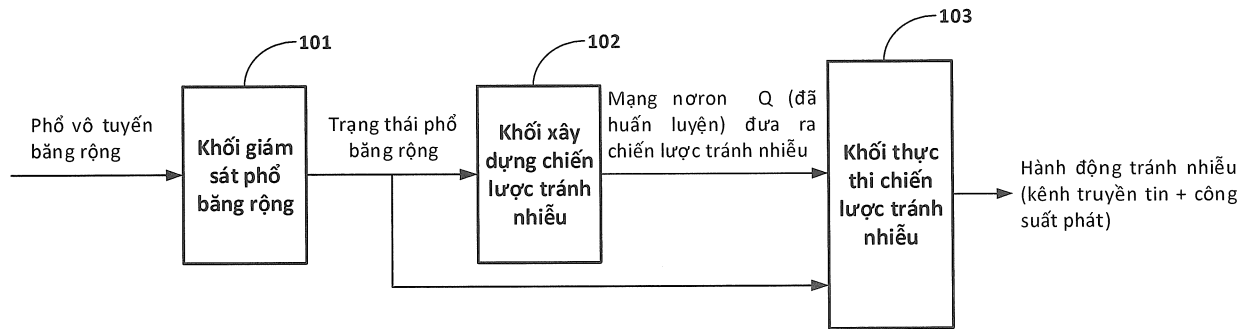
bước 2: xác định trạng thái và công suất các kênh trong băng rộng tại một khung thời gian; cụ thể dựa trên mật độ phổ công suất, tính toán công suất các kênh trong băng rộng và thực hiện chuẩn hóa công suất bởi khối tính toán công suất và khối chuẩn hóa công suất; trạng thái các kênh trong băng rộng được xác định bởi khối xác định trạng thái kênh sử dụng các thông tin đầu vào là công suất các kênh, ngưỡng phát hiện tín hiệu và kênh truyền của thiết bị ở khung thời gian trước đó;

bước 3: xây dựng vectơ trạng thái và thông tin trạng thái phổ băng rộng; cụ thể, vectơ trạng thái phổ băng rộng tại một khung thời gian được khối xây dựng vectơ trạng thái phổ tạo ra bằng cách tổng hợp thông tin trạng thái kênh và công suất kênh chuẩn hóa của tất cả các kênh; dựa trên thông tin các vectơ trạng thái phổ băng rộng, thông tin trạng thái phổ băng rộng ở khung thời gian hiện tại được tạo ra gồm vectơ trạng thái phổ băng rộng ở khung thời gian hiện tại cùng với các vectơ trạng thái phổ băng rộng ở một số khung thời gian trước đó;

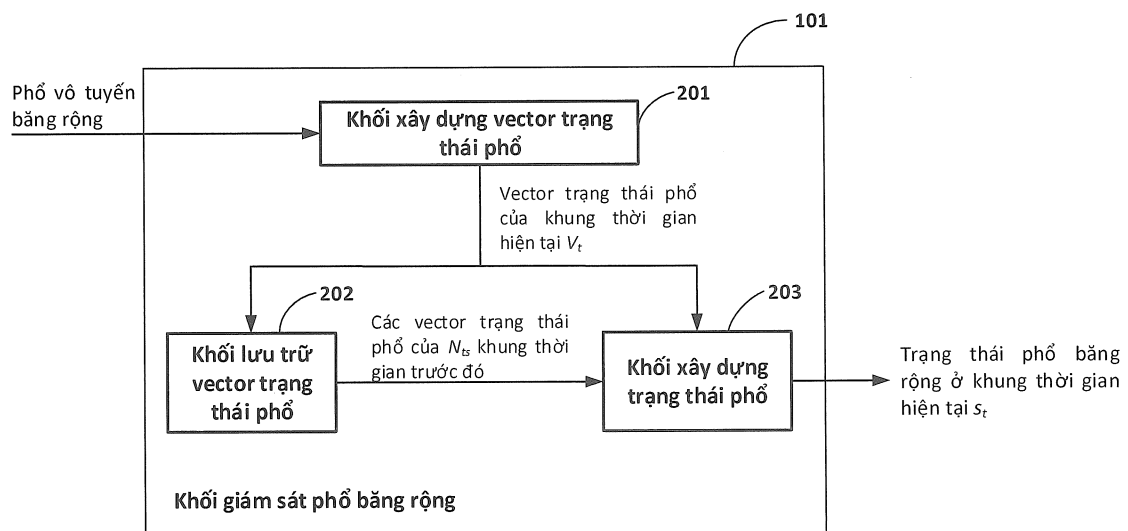
bước 4: thực hiện huấn luyện mạng nơron Q trong một số khung thời gian huấn luyện; cụ thể, tại mỗi khung thời gian huấn luyện, khối xây dựng chiến lược tránh nhiễu thu thập thông tin kinh nghiệm (gồm trạng thái phổ băng rộng ở khung

thời gian hiện tại, hành động, phần thưởng đạt được và trạng thái phổ băng rộng ở khung thời gian tiếp theo), lưu vào bộ nhớ khối lưu trữ các thông tin kinh nghiệm. Khi bộ nhớ đủ lớn, một tập con các kinh nghiệm được lấy mẫu ngẫu nhiên và sử dụng cho quá trình huấn luyện mạng nơron Q mục tiêu bởi khối cập nhật bộ trọng số mạng nơron Q . Khối này tính toán hàm lỗi và cập nhật bộ trọng số cho mạng nơron Q theo hướng tối thiểu hóa được hàm lỗi;

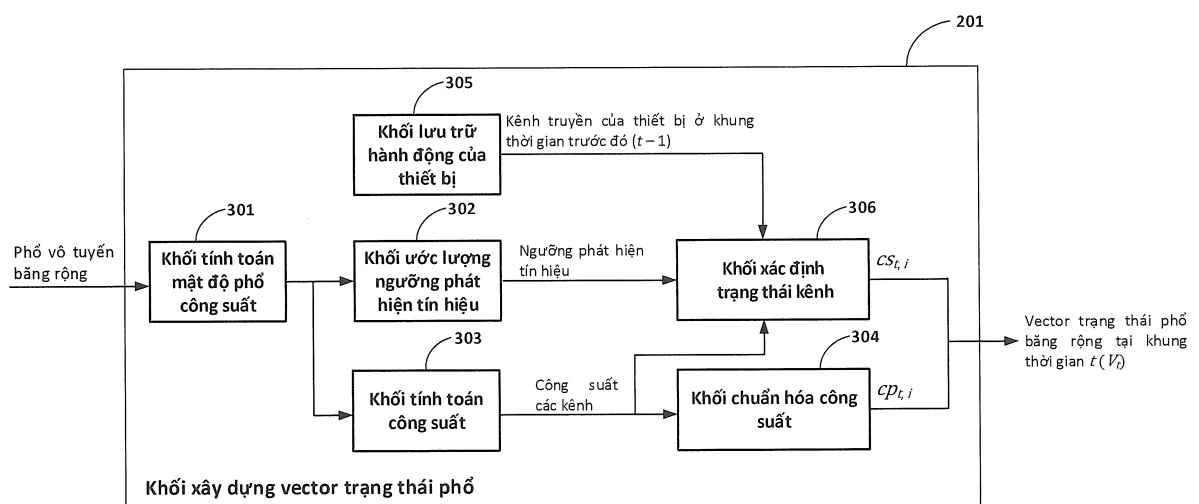
bước 5: lựa chọn hành động (kênh và công suất truyền) tránh nhiễu dựa trên kết quả đầu ra của mạng nơron Q đã được huấn luyện; cụ thể, trong quá trình thực thi chiến lược tránh nhiễu sau khi kết thúc quá trình huấn luyện ở bước 4, tại thời điểm bắt đầu mỗi khung thời gian, thông tin trạng thái phổ băng rộng được đưa vào khối thực thi chiến lược tránh nhiễu, khối này tính toán hành động tránh nhiễu tối ưu dựa trên việc lựa chọn hành động tương ứng làm cho giá trị hàm Q - được xấp xỉ bởi đầu ra của mạng nơron dự đoán Q - đạt giá trị cực đại; thiết bị thực hiện truyền tin trên kênh lựa chọn với công suất phát tương ứng theo hành động được đề xuất trong khung thời gian này để đạt được hiệu quả tránh nhiễu tối ưu.



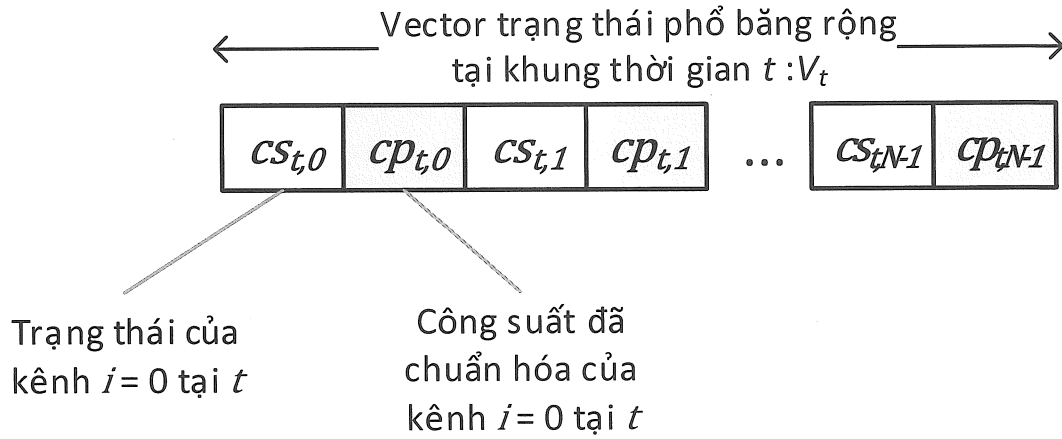
Hình 1



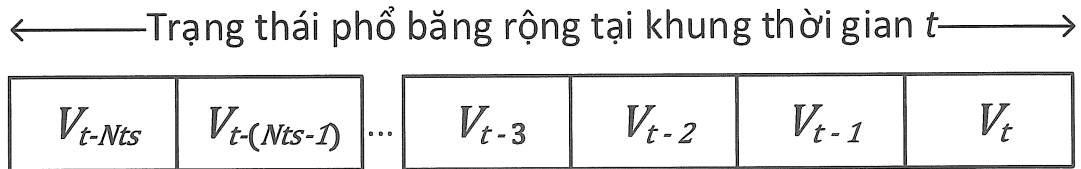
Hình 2



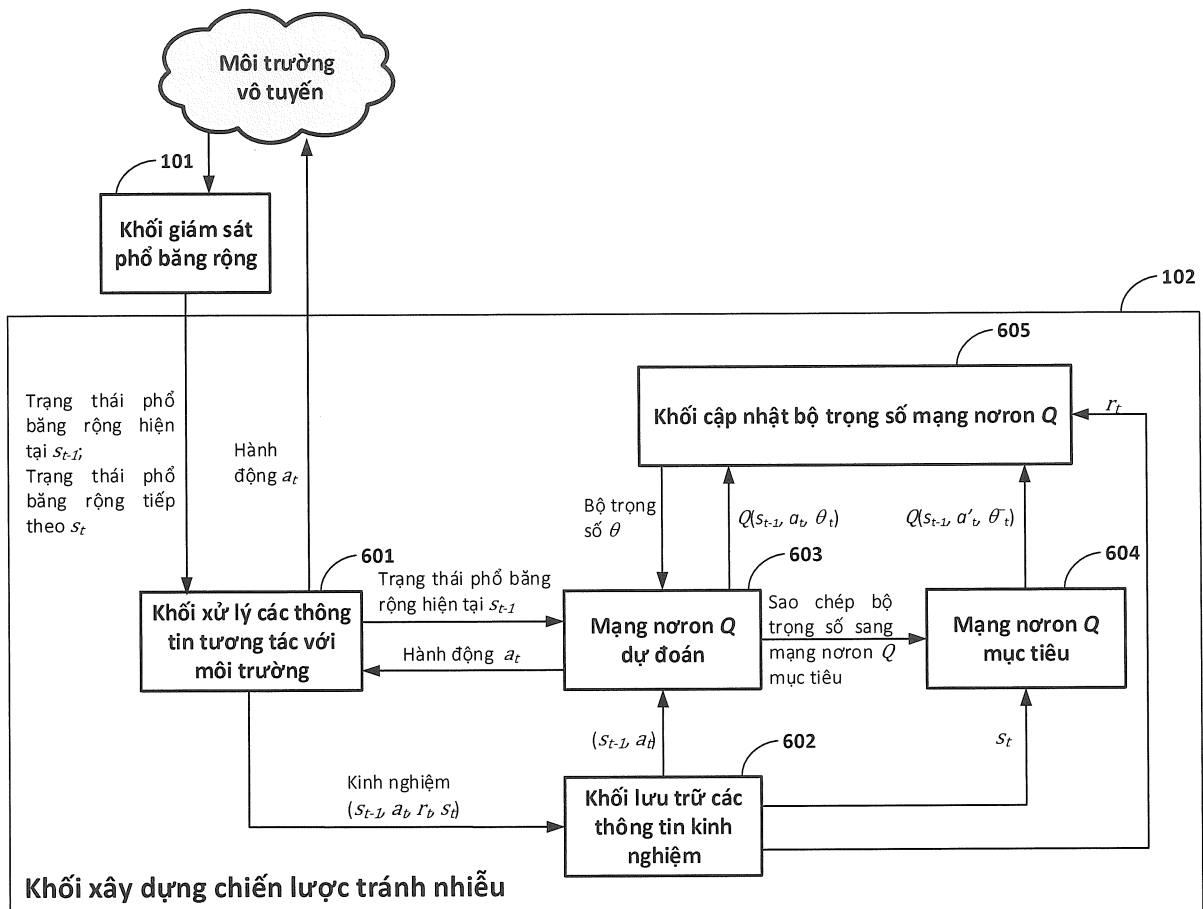
Hình 3



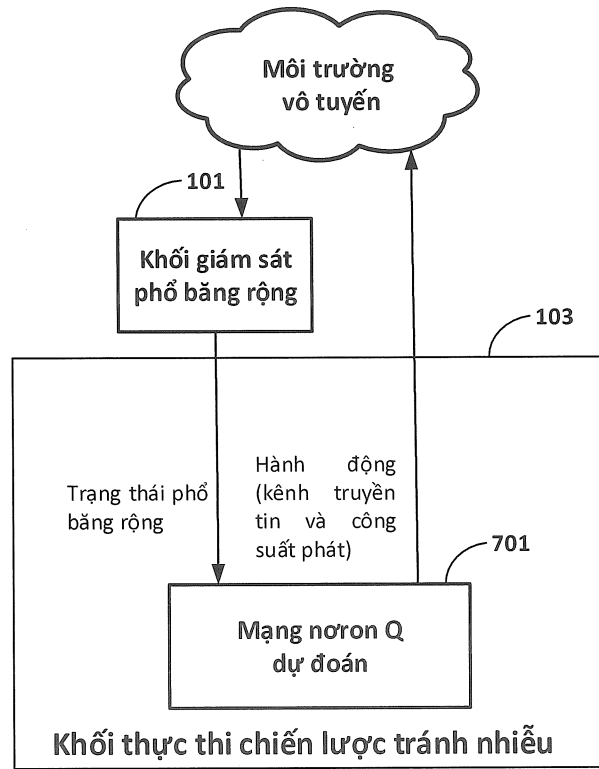
Hình 4



Hình 5



Hình 6



Hình 7