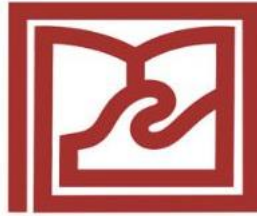


**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DUY TÂN**



**TỐI ƯU LƯU TRỮ VÀ TRUYỀN VIDEO CỘNG TÁC
TRONG MẠNG 5G SIÊU DÀY ĐẶC**

**CHUYÊN NGÀNH: Khoa học Máy tính
MÃ SỐ: 9.48.01.01**

TÓM TẮT LUẬN ÁN TIẾN SĨ

Đà Nẵng – 2021

**BỘ GIÁO DỤC VÀ ĐÀO TẠO
TRƯỜNG ĐẠI HỌC DUY TÂN**



NCS. BÙI MINH PHỤNG

**TỐI ƯU LƯU TRỮ VÀ TRUYỀN VIDEO CỘNG TÁC
TRONG MẠNG 5G SIÊU DÀY ĐẶC**

**CHUYÊN NGÀNH: Khoa học Máy tính
MÃ SỐ: 9.48.01.01**

TÓM TẮT LUẬN ÁN TIẾN SĨ

**NGƯỜI HƯỚNG DẪN KHOA HỌC
TS. VÕ NGUYỄN SƠN**

Đà Nẵng – 2021

MỞ ĐẦU

1. Lý do chọn đề tài

Thế giới đang bước vào kỷ nguyên của công nghiệp 4.0, song song đó, theo dự đoán sẽ có khoảng 50 tỉ thiết bị được kết nối để truyền thông với nhau và truy cập các dịch vụ/tiện ích đa phương tiện vào năm 2021 [1]. Trong bối cảnh này, mạng thế hệ thứ 5 sẽ đối diện với thách thức của một lượng lớn người dùng di động (MU) yêu cầu các dịch vụ và ứng dụng có dung lượng dữ liệu và tốc độ cao. Điều này sẽ khiến cho mạng 5G trở nên suy yếu bởi vấn đề tắc nghẽn do xung đột lưu lượng xảy ra tại các tuyến trục (backhaul link) của các trạm cơ sở lớn (MBS) cũng như của các trạm cơ sở nhỏ (SBS). Trong khi việc đầu tư phát triển các tuyến trục tốc độ cao tốn rất nhiều chi phí và thậm chí phải thay đổi kiến trúc mạng thì những giải pháp và kỹ thuật mới có thể thay thế một cách hiệu quả hơn.

Mạng siêu dày đặc (UDN) được xem như là một kiến trúc hứa hẹn có khả năng đáp ứng được các yêu cầu của mạng 5G như tăng dung lượng hệ thống gấp 1000 lần và độ trễ truy xuất dữ liệu nhỏ hơn 1ms [2]. Tuy nhiên, việc phát triển UDN còn phải yêu cầu nhiều hơn nữa các công nghệ, kỹ thuật và các thiết kế tối ưu đột phá nhằm cung cấp cho lượng lớn MU những dịch vụ và ứng dụng có dung lượng dữ liệu và tốc độ cao, ví dụ như dịch vụ và ứng dụng truyền video (VAS), với chất lượng dịch vụ (QoS) và hiệu quả sử dụng tài nguyên tốt nhất. Để đạt được điều này, nhiều công nghệ, kỹ thuật và thiết kế tối ưu cho UDN đã được nghiên cứu tập trung vào việc làm thế nào để tận dụng các nguồn tài nguyên về không gian, thời gian, mã, phổ tần, băng thông, năng lượng và dung lượng lưu trữ, cũng như làm thế nào để đưa các dịch vụ tiên tiến đến gần với các MU hơn.

Gần đây, kỹ thuật lưu trữ đã thu hút sự chú ý của các nhà nghiên cứu trong cả lĩnh vực học thuật và công nghiệp để mang lại lợi ích cho các nhà cung cấp dịch vụ Internet (ISP) và nhà cung cấp nội dung cũng như đáp ứng nhu cầu cao của người dùng cuối. Trong lĩnh vực này, các nhà nghiên cứu Việt Nam đã có cơ hội hợp tác với các chuyên gia hàng đầu thế giới nhằm phát triển các dự án và công bố các công trình nghiên cứu ở các tạp chí uy tín [14, 19, 21-23, 29-31, 35, 38, 44, 50]. Tuy vậy, các nhóm nghiên cứu này vẫn chưa triển khai một cách tổng thể mô hình, phân tích và thiết kế tối ưu lưu trữ và truyền video hợp tác trong 5G UDN. Do vậy, vẫn còn nhiều vấn đề thách thức chưa được giải quyết nhằm nâng cao năng lực cho 5G UDN trong việc cung cấp cho MU các dịch vụ tiên tiến có QoS và hiệu suất sử dụng tài nguyên cao. Ở các nước phát triển, các công trình nghiên cứu trong lĩnh vực này nhiều hơn về số lượng và tốt hơn về chất lượng, cụ thể như: lưu trữ tại các thiết bị [15, 20, 33, 42, 45], lưu trữ tại trạm cơ sở nhỏ femto [26, 27, 34], lưu trữ tại các trạm cơ sở nhỏ [17, 32, 37, 41, 43, 46], lưu trữ tại trạm cơ sở lớn [39, 40], và lưu trữ đa tầng [16, 18, 24, 25, 28, 36, 47, 48].

Mặc dù các công trình nghiên cứu được thực hiện tại các quốc gia phát triển nhiều hơn về số lượng và có những đóng góp quan trọng hơn so với các công trình được thực hiện bởi các nhà nghiên cứu trong nước, vẫn còn nhiều vấn đề chưa được giải quyết và cần phải có những kỹ thuật, mô hình, phân tích, thiết kế tối ưu và các tiêu chuẩn đột phá nhằm đáp ứng các yêu cầu cao và phức tạp của lưu trữ và truyền/phân phối video trong 5G UDN.

2. Mục tiêu, đối tượng, phạm vi và phương pháp nghiên cứu

2.1. Mục tiêu nghiên cứu

Mục tiêu nghiên cứu của luận án là đề xuất các cơ chế lưu trữ và truyền video hợp tác đa tầng trong 5G UDN nhằm cung cấp cho người dùng di động các dịch vụ tiên tiến có chất lượng dịch vụ và hiệu suất sử dụng tài nguyên cao.

2.2. Đối tượng nghiên cứu

- 5G UDN: tập trung vào nghiên cứu các mô hình, các đặc điểm của 5G UDN có nhiều tầng thiết bị thu phát gồm MBS, SBS (như microcell, picocell, femtocell) và truyền thông D2D.
- Video: các loại video chuẩn được dùng phổ biến, mô hình mối quan hệ giữa méo và tốc độ mã hóa của video, các cơ chế lưu trữ và truyền video trong 5G UDN.
- Mô hình: mối quan hệ xã hội, QoS và tài nguyên trong 5G UDN.
- Thuật giải: các thuật giải nhằm giải bài toán tối ưu quá trình lưu trữ, chia sẻ tài nguyên và truyền video từ MBS và SBS đến các MU và giữa các MU với nhau.

2.3. Phạm vi nghiên cứu

Trong luận án này, NCS giới hạn các phạm vi nghiên cứu liên quan đến kỹ thuật, người dùng, mạng và dịch vụ cũng như toán và các công cụ hỗ trợ, cụ thể như sau:

- Về kỹ thuật: luận án tập trung vào kỹ thuật lưu trữ và chia sẻ tài nguyên trong 5G UDN.
- Về người dùng, mạng và dịch vụ: luận án tập trung vào dịch vụ và ứng dụng truyền video (VAS) trong 5G UDN, các tiêu chí của QoS, hiệu suất sử dụng tài nguyên, hành vi và mối quan hệ xã hội của người dùng dựa trên mô hình Buffet Ấn Độ (IBM – Indian Buffet Model).
- Về toán và các công cụ hỗ trợ: luận án tập trung vào các thuật giải tìm kiếm, một số công cụ phân tích, công cụ mã hóa video và mô hình đường cong đặc tính méo – tốc độ (Rate – Distortion) của video.

2.4. Phương pháp nghiên cứu

Để đạt kết quả phù hợp với các mục tiêu nghiên cứu, dựa trên các đối tượng cần nghiên cứu, 02 phương pháp nghiên cứu được áp dụng gồm 1) Phương pháp phân tích và tổng hợp và 2) Phương pháp nghiên cứu định lượng, cụ thể như sau:

- Phương pháp phân tích và tổng hợp: Phân tích và đánh giá các mô hình và kết quả nghiên cứu được công bố liên quan, từ đó, xác định được các vấn đề còn tồn đọng, những câu hỏi và ý tưởng, đặt ra các giả thuyết nghiên cứu cũng như khẳng định sự cần thiết phải có mô hình và giải pháp mới tốt hơn. Sau đó, các kết quả phân tích và đánh giá sẽ được tổng hợp, liên kết, kết hợp và tổ chức lại một cách có hệ thống nhằm đề xuất các mô hình và giải pháp tốt hơn dựa trên giả thuyết nghiên cứu được đặt ra.
- Phương pháp nghiên cứu định lượng: Trong phương pháp định lượng, các mô hình đề xuất và các yếu tố liên quan được lượng hóa bởi các biểu thức tính toán. Các biểu thức tính toán mô tả bản chất của mô hình hệ thống được kiểm định tính đúng đắn thông qua việc thực hiện mô phỏng và quan sát đáp ứng của hệ thống. Tính ưu việt của giải pháp đề xuất được kiểm chứng thông qua việc đối sánh với các giải pháp của các nghiên cứu khác liên quan.

3. Nhiệm vụ nghiên cứu và kết quả đạt được

3.1. Nhiệm vụ nghiên cứu

- Nhiệm vụ nghiên cứu 1: Đề xuất cơ chế lưu trữ và chia sẻ tài nguyên tối ưu theo mối quan hệ xã hội người dùng nâng cao dung lượng truyền video trong 5G UDN, mô hình toán các thông số của mô hình đề xuất, xây dựng và giải bài toán tối ưu bằng các giải thuật tìm kiếm, mô phỏng tính toán và đánh giá hiệu quả của mô hình.
- Nhiệm vụ nghiên cứu 2: Đề xuất cơ chế lưu trữ đa phân giải và chia sẻ tài nguyên tối ưu theo nhu cầu người dùng nâng cao chất lượng truyền video trong 5G UDN, mô hình toán các thông số của mô hình đề xuất, xây dựng và giải bài toán tối ưu bằng các giải thuật tìm kiếm, mô phỏng tính toán và đánh giá hiệu quả của mô hình.

3.2. Kết quả đạt được

Trong quá trình nghiên cứu và hoàn thành luận án, NCS đã công bố 2 bài báo ISI và 1 bài báo hội nghị quốc tế. Kết quả nghiên cứu của luận án góp phần vào việc tối ưu lưu trữ và phân phối tài nguyên là video nhằm cung cấp cho người dùng dịch vụ và chất lượng video tốt hơn.

- Kết quả 1: Đề xuất cơ chế lưu trữ và truyền video cộng tác đa tầng có quan tâm đến hành vi và mối quan hệ xã hội của người dùng (SCS) thông qua việc tận dụng tài nguyên lưu trữ và tài nguyên phổ tần sẵn có của các phần tử trong 5G UDN.
- Kết quả 2: Đề xuất cơ chế lưu trữ đa phân giải và truyền video cộng tác có quan tâm đến nhu cầu người dùng (CRS) thông qua việc tận dụng hiệu quả tài nguyên lưu trữ và tài nguyên phổ tần sẵn có của các phần tử trong 5G UDN.

4. Bố cục luận án

Mở đầu

Chương 1: Tổng quan về lưu trữ và truyền video trong 5G UDN

Chương 2: Lưu trữ và chia sẻ tài nguyên tối ưu theo mối quan hệ xã hội người dùng nâng cao dung lượng truyền video trong 5G UDN

Chương 3: Lưu trữ đa phân giải và chia sẻ tài nguyên tối ưu theo nhu cầu người dùng nâng cao chất lượng truyền video trong 5G UDN

Kết luận

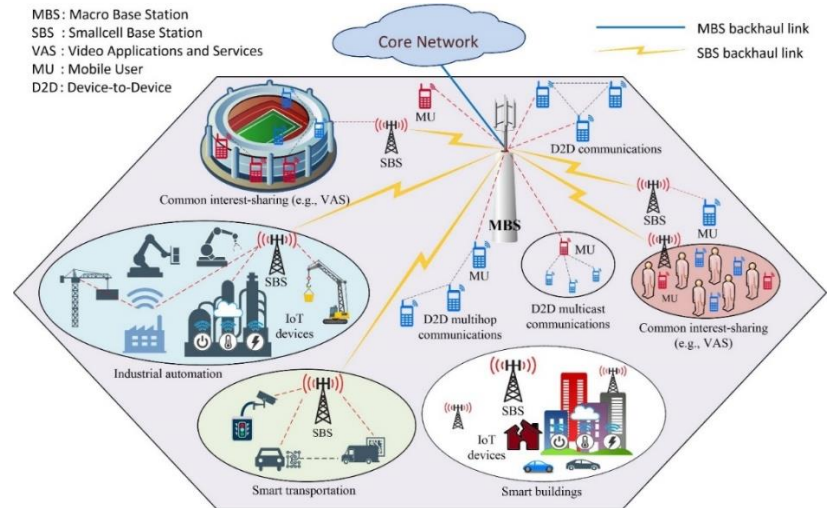
CHƯƠNG 1. TỔNG QUAN VỀ LƯU TRỮ VÀ TRUYỀN VIDEO TRONG 5G UDN

1.1. Giới thiệu về 5G UDN

Mô hình của 5G UDN gồm nhiều tầng, cung cấp cho người dùng di động (MU – Mobile User) nhiều dịch vụ và ứng dụng khác nhau, được mô tả trong Hình 1-1. Trong mô hình này, tầng trên sử dụng dải phổ tần cao hơn, tín hiệu mạnh hơn với tốc độ truyền cao hơn, trong khi đó, không gian dịch vụ và diện tích phủ sóng sẽ được mở rộng hơn nhờ vào các tầng trung gian và tầng thấp [2,52]. 5G UDN sẽ cho phép kết hợp linh hoạt các kỹ thuật truyền thông dựa vào việc tận dụng và chia sẻ tài nguyên giữa các phần tử/nhóm phần tử trong mạng nhằm đảm bảo kết nối, đáp ứng nhu cầu ngày càng cao của người dùng.

Tuy nhiên, 5G UDN vẫn còn những thách thức về vật lý cũng như quá trình sử dụng và khai thác dữ liệu của MU như:

- Mật độ kết nối siêu dày đặc từ những phạm vi truy cập và lưu lượng truy cập khác nhau tại các vị trí khác nhau. Điều đó có thể gây tắc nghẽn cục bộ, làm mất cân bằng hiệu suất và quyền hạn truy cập của các thiết bị.
- Những giới hạn truy cập riêng hoặc chung ở các tầng khác nhau sẽ tạo ra nhiều mức can nhiễu khác nhau. Ví dụ: các phiên truyền thông D2D khác nhau có thể ảnh hưởng can nhiễu cho nhau và thậm chí làm gián đoạn quá trình truy cập.
- Vấn đề ưu tiên truy cập các kênh ở các tần số khác nhau và ưu tiên trong các chiến lược phân bổ tài nguyên.



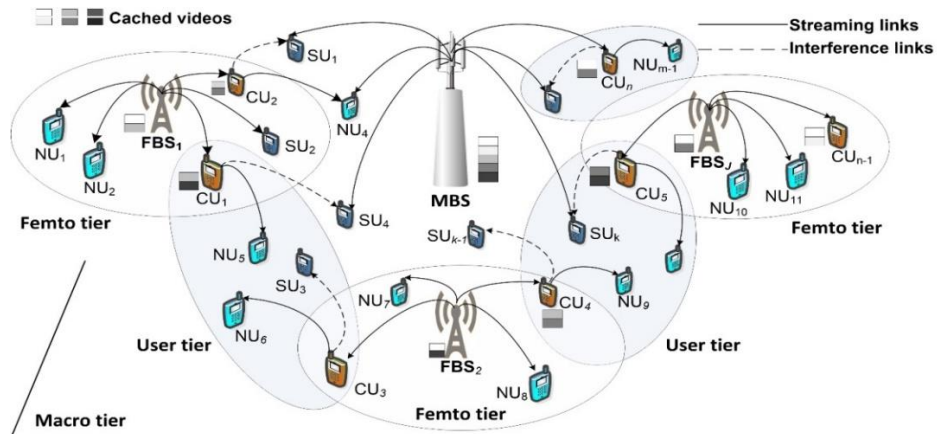
Hình 1-1. Mô tả kết nối giữa các thiết bị, các công nghệ và ứng dụng trong 5G UDN [2]

Như vậy, những nghiên cứu nhằm đưa ra các giải pháp khắc phục các vấn đề thách thức nêu trên trong 5G UDN là cấp thiết để sử dụng tối ưu nguồn tài nguyên, nâng cao chất lượng mạng di động và cải thiện sự hài lòng của người dùng.

1.2. Mô hình lưu trữ và truyền video trong 5G UDN

Kỹ thuật lưu trữ để mang các dịch vụ/ứng dụng/nội dung, đặc biệt là các dịch vụ và ứng dụng có dung lượng lớn và tốc độ dữ liệu cao như truyền video (VAS – Video Applications and Service), đến gần với MU hơn, là giải pháp nhằm giải quyết bài toán tắc nghẽn nhưng lại không làm thay đổi kiến trúc hệ thống.

Kỹ thuật lưu trữ thường được kết hợp với các cơ chế chia sẻ tài nguyên phổ tần nhằm mang lại lợi ích nhiều hơn cho nhà cung cấp dịch vụ Internet và nhà cung cấp nội dung cũng như đáp ứng nhu cầu cao của MU trong việc truy xuất các dịch vụ tiên tiến tốc độ cao như truyền video.



Hình 1-2. Mô hình lưu trữ và chia sẻ tài nguyên cho các dịch vụ và ứng dụng video trong 5G UDN

Kỹ thuật lưu trữ gồm lưu trữ và truyền video đơn tầng, lưu trữ và truyền video đa tầng. Lưu trữ và truyền video đơn tầng là kỹ thuật cho phép hệ thống lưu trữ các video tại tầng MBS hoặc tại tầng SBS hoặc tại tầng người dùng di động (user tier) có khả năng lưu trữ (CU – Caching User). Lưu trữ và truyền video cộng tác đa tầng là kỹ thuật lưu trữ video được thực hiện kết hợp tại nhiều hơn 1 tầng trong 5G UDN (Hình 1-2).

Trong lưu trữ đa tầng, mặc dù kỹ thuật sẽ phức tạp hơn so với đơn tầng, nhưng các MU được chọn lựa truy xuất video một cách linh động từ tầng nào được xem là tốt nhất. Các nghiên cứu liên quan đến lưu trữ đa tầng chủ yếu triển khai trên 2 tầng (sẽ được trình bày chi tiết trong Mục 1.3). Trong luận án này, lưu trữ đa tầng được hiểu là 3 tầng có kết hợp cơ chế chia sẻ tài nguyên phổ tần kênh truyền xuống của các SU (Spectrum owning User) cho truyền thông video từ thiết bị đến thiết bị (D2D – Device-to-Device). Các tầng sẽ phối hợp với nhau trong quá trình lưu trữ và chia sẻ tài nguyên để phục vụ MU một cách tốt nhất với hiệu quả sử dụng tài nguyên cao nhất.

1.3. Các thuật giải ứng dụng trong bài toán tối ưu lưu trữ và truyền video

Lưu trữ kết hợp chia sẻ tài nguyên tối ưu là kỹ thuật đã và đang được nhiều nhà nghiên cứu thực hiện ứng dụng trong truyền video qua 5G UDN đạt hiệu quả cao với chi phí thấp. Các bài toán tối ưu lưu trữ này thường được giải bằng cách tìm các chỉ số xác định vị trí lưu trữ và chọn lựa/phân bổ tài nguyên trong hệ thống dưới dạng vector hoặc ma trận và các chỉ số thường được gán giá trị 1 (hoặc 0) có nghĩa là hệ thống có (hoặc không) lưu trữ hoặc/và có (hoặc không) chia sẻ tài nguyên. Để tìm kết quả tối ưu, các thuật giải tìm kiếm tối ưu đã được đề xuất để giải các bài toán trên như tìm kiếm vét cạn [18, 54-57], quy hoạch động [50], thuật giải học ngẫu nhiên (Stochastic Learning), thuật giải học ngẫu nhiên [40], thuật giải trò chơi [43], thuật giải tham lam [58], và một số thuật giải heuristic [25, 59]. Mỗi thuật giải đều có những ưu và nhược điểm riêng về độ chính xác, độ phức tạp về thời gian và dung lượng tính toán. Trong luận án này, thuật giải di truyền (GA) sẽ được chọn lựa để áp dụng cho việc giải các bài toán tối ưu lưu trữ và chia sẻ tài nguyên trong các chương tiếp theo. Đồng thời, thuật giải đàn doi (BA – Bat Algorithm) cũng được thực hiện để khẳng định tính hiệu quả của thuật giải GA. Ngoài ra, thuật giải EA cũng được sử dụng vì đơn giản, cho kết quả chính xác, phù hợp cho bài toán có không gian tìm kiếm nhỏ, cũng như được dùng để làm thước đo độ chính xác cho các thuật giải GA và BA.

1.4. Tổng kết chương 1

Chương 1 đã giới thiệu tổng quan về 5G UDN với những ưu điểm nổi bật cũng như những thách thức cần nghiên cứu. Ngoài ra, cơ chế lưu trữ và chia sẻ tài nguyên trong 5G UDN cũng được trình bày một cách tổng quát. Các ưu điểm và nhược điểm của các mô hình, cơ chế lưu trữ và chia sẻ trong các công trình nghiên cứu liên quan cũng được phân tích và đánh giá chi tiết. Chương 1 cũng giới thiệu các thuật giải như toán vét cạn, thuật giải di truyền, thuật giải đàn doi, là những thuật giải sẽ được triển khai, so sánh và áp dụng để giải các bài toán tối ưu trong Chương 2 và Chương 3. Từ đó, cho thấy rõ được điểm khác biệt và những đóng góp của luận án.

CHƯƠNG 2. LƯU TRỮ VÀ CHIA SẺ TÀI NGUYÊN TỐI ƯU THEO MỐI QUAN HỆ XÃ HỘI NGƯỜI DÙNG NÂNG CAO DUNG LƯỢNG TRUYỀN VIDEO TRONG 5G UDN

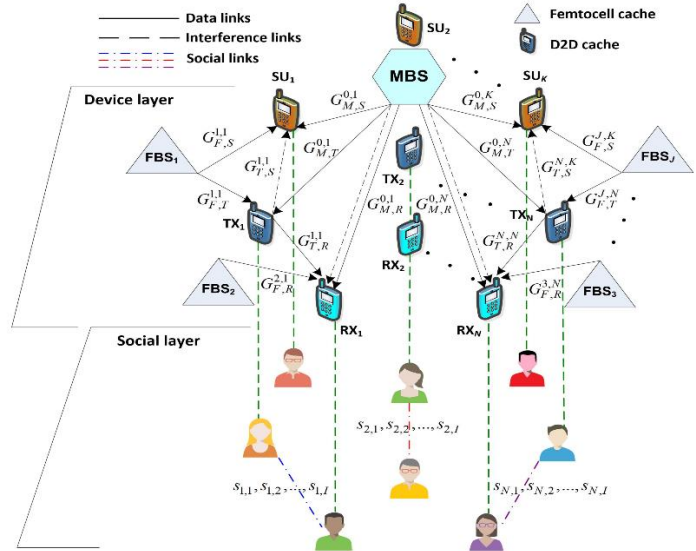
2.1. Giới thiệu cơ chế SCS

Trong mạng 5G, lượng lớn kết nối của các dịch vụ yêu cầu tốc độ và dung lượng dữ liệu cao như các dịch vụ truyền video sẽ gây tắc nghẽn tại các tuyến trục của các MBS. Mặc dù 5G UDN được coi là một kiến trúc đầy hứa hẹn để giải quyết các vấn đề về số lượng kết nối, nhưng tắc nghẽn mạng vẫn có thể xảy ra làm hạn chế dung lượng phân phối video đến MU. Trong Chương 2, cơ chế lưu trữ và chia sẻ tài nguyên tối ưu theo mối quan hệ xã hội người dùng (SCS) được đề xuất nhằm tối đa dung lượng phân phối video trong 5G UDN. Cơ chế SCS sẽ làm giảm tắc nghẽn tại MBS bằng cách cho phép MU truy xuất video một cách linh động từ 3 tầng của 5G UDN gồm: tầng MBS, tầng FBS và tầng người dùng từ các MU khác. Trong đó, các MU bao gồm những cặp Transmitter (TX) và Receiver (RX) trong truyền thông D2D (TX-RX) và các MU chia sẻ tài nguyên phổ tần kênh truyền xuống (SU) để cho phép thiết lập các phiên truyền thông video D2D. Vấn đề được đặt ra là làm thế nào để cơ chế SCS phân phối các video đến các MU với dung lượng cực đại.

Để giải quyết vấn đề trên, bài toán SCS được xây dựng và giải để chỉ định FBS nào lưu trữ video nào và chỉ định SU nào chia sẻ tài nguyên phổ tần kênh truyền xuống cho cặp truyền thông D2D nào, sao cho dung lượng phân phối trung bình đến các MU là cực đại. Nhằm nâng cao hiệu quả của cơ chế SCS, bài toán tối ưu có xét đến mối quan hệ xã hội của từng cặp TX-RX và có xét đến ràng buộc về dung lượng lưu trữ tại các FBS. Hơn nữa, do SU chia sẻ tài nguyên phổ tần kênh truyền xuống của nó cho truyền thông D2D từ TX đến RX, việc TX dùng lại phổ tần kênh truyền xuống này sẽ gây can nhiễu lên SU. Vì vậy, bài toán SCS cũng xem xét đến ràng buộc về giới hạn dưới của SINR tại SU để đảm bảo chất lượng cho SU. Những ưu điểm của cơ chế SCS được kiểm chứng thông qua các kết quả mô phỏng và so sánh với các cơ chế thông thường khác. Các công trình công bố liên quan đến Chương 2 gồm [C1] trong phần “Danh mục các công trình đã công bố của nghiên cứu sinh” ở cuối luận án.

2.2. Mô hình hệ thống với cơ chế SCS

Trong chương này, mô hình truyền video trong 5G UDN với cơ chế SCS được mô tả như Hình 2-1. Mô hình được phân ra thành lớp thiết bị và lớp mối quan hệ xã hội. Lớp thiết bị gồm 1 MBS, J FBS, $(K + 2N)$ MU gồm K SU và N cặp TX-RX, và I video. Lớp xã hội mô tả mối quan hệ xã hội của các cặp TX-RX trong truyền thông D2D. Trong mô hình này, khi MBS dự đoán được sẽ có lượng lớn MU yêu cầu truy xuất các video, nó sẽ thực thi cơ chế SCS gồm 3 bước như sau:



Hình 2-1. Mô hình hệ thống của SCS

Bước 1 – Cập nhật các thông số: Trong bước này, MBS sẽ cập nhật các thông số của hệ thống như số FBS (J) để lưu trữ video, số SU (K) để chia sẻ tài nguyên phổ tần kênh truyền xuống, số TX-RX (N) cho truyền thông D2D, số video (I), mối quan hệ xã hội của các cặp TX-RX, băng thông hệ thống, đặc tính kênh truyền và các thông số khác như được liệt kê trong Bảng 2-1.

Bước 2 – Xây dựng và giải bài toán tối ưu SCS: Dựa trên các thông số ở Bước 1, MBS sẽ xây dựng bài toán tối ưu SCS và giải bài toán này nhằm tìm ra số lượng bản sao lưu tối ưu của mỗi video để cực đại số bản sao lưu trữ trung bình trong hệ thống (tăng tỷ lệ truy suất thành công video); các chỉ số lưu trữ tối ưu $u_{j,i}$, với $j = 1, 2, \dots, J$ và $i = 1, 2, \dots, I$, và các chỉ số chia sẻ tài nguyên phổ tần kênh truyền xuống tối ưu $v_{k,n}$, $k = 1, 2, \dots, K$ và $n = 1, 2, \dots, N$; nhằm cực đại dung lượng phân phối của hệ thống. Trong đó:

$u_{j,i} = 1$ nghĩa là FBS thứ j quyết định lưu trữ video thứ i , ngược lại thì $u_{j,i} = 0$ và

$v_{k,n} = 1$ nghĩa là SU thứ k quyết định chia sẻ tài nguyên kênh truyền xuống của nó cho cặp TX-RX thứ n , ngược lại thì $v_{k,n} = 0$

Bước 3 – Triển khai SCS: Cuối cùng, dựa vào kết quả tối ưu sau khi giải bài toán SCS, MBS sẽ cộng tác với các FBS và các TX để phân phối các video được lưu trữ đến các SU và TX (phân phối bởi MBS và FBS) cũng như đến các RX (phân phối bởi TX) với dung lượng cực đại.

2.3. Tính toán các thông số của hệ thống với cơ chế SCS

Nhằm thuận tiện cho việc tính toán các thông số của hệ thống, các ký hiệu của các thông số được mô tả trong Bảng 2-1. Từ hệ thống với mục tiêu nêu trên, ta có thể nhận thấy các thông số cần được mô hình và tính toán cho bài toán tối ưu SCS bao gồm: mối quan hệ xã hội của các cặp TX-RX và tỷ số tín hiệu trên nhiễu (SNR – Signal-to-Noise Ratio) hoặc tỷ số tín hiệu trên can nhiễu và nhiễu trắng (SINR) của các kênh truyền không dây

để làm cơ sở tính toán dung lượng kênh truyền. Từ đó, ta có thể tính được dung lượng phân phối trung bình của hệ thống – là hàm mục tiêu của bài toán tối ưu SCS.

Bảng 2-1. Các ký hiệu sử dụng cho bài toán SCS

Ký hiệu	Diễn giải
I	Số video
J	Số trạm cơ sở nhỏ femto (FBS)
K	Số người dùng di động chấp nhận chia sẻ tài nguyên phổ tần kênh truyền xuống (SU)
N	Số cặp truyền thông D2D, mỗi cặp gồm một người dùng di động lưu trữ video (TX – Transmitter) và một người dùng nhận video (RX – Receiver).
T_{min}^i	Thời gian tối thiểu để truyền video thứ i , $i = 1, 2, \dots, I$ (tùy thuộc vào độ dài của mỗi video)
$s_{n,i}$	Xác suất dựa trên mối quan hệ xã hội để xem xét khả năng cặp TX-RX thứ n có truyền thành công video thứ i hay không
$u_{j,i}$	Chỉ số lưu trữ video tại FBS, $u_{j,i} = 1$ nghĩa là FBS thứ j quyết định lưu video thứ i và ngược lại.
$v_{k,n}$	Chỉ số chia sẻ tài nguyên phổ tần kênh truyền xuống: $v_{k,n} = 1$ nghĩa là SU thứ k đồng ý chia sẻ phổ tần kênh truyền xuống của nó cho cặp D2D thứ n và ngược lại.
P_M^0	Công suất phát của trạm cơ sở lớn (MBS)
P_F^j	Công suất phát của trạm cơ sở nhỏ femto (FBS) thứ j
P_T^n	Công suất phát của TX thứ n trong cặp truyền thông D2D thứ n .
$G_{X,Y}^{x,y}$	Độ lợi kênh giữa X và Y , trong đó x và y là các chỉ số của X và Y
N_0	Công suất nhiễu trắng Gaussian (AWGN – Additive White Gaussian Noise)
W	Băng thông hệ thống
r_i	Tỷ lệ truy xuất (hay độ nổi tiếng) của video thứ i
β_n	Phần trăm dung lượng lưu trữ còn trống của TX thứ n
$p_{n,i}$	Xác suất để TX thứ n quyết định lưu video thứ i
γ_0	Giới hạn dưới tỉ số tín hiệu trên can nhiễu và nhiễu trắng (SINR) tại SU
R_S	Dung lượng phân phối từ MBS và FBS đến SU
R_T	Dung lượng phân phối từ MBS và FBS đến TX
R_R	Dung lượng phân phối từ MBS, FBS và TX đến TR

2.3.1. Mối quan hệ xã hội giữa các cặp D2Ds

Mối quan hệ xã hội giữa TX và RX của cặp D2D thứ n được dùng để tính xác suất truyền thành công video thứ i trong thời gian T_{min}^i .

$$s_{n,i} = 1 - \int_0^{\delta T_{min}^i} f(u; \kappa_n, \theta_n) du = 1 - \frac{\gamma \left(\kappa_n, \frac{\delta T_{min}^i}{\theta_n} \right)}{\Gamma(\kappa_n)} \quad (2.5)$$

2.3.2. Kênh truyền không dây

Đặc trưng của kênh truyền không dây là độ lợi kênh $G_{X,Y}^{x,y} = h_{X,Y}^{x,y} g_{X,Y}^{x,y}$, trong đó:

$X \in \{M, F, T\}$ là viết tắt của {MBS, FBS, TX} và $Y \in \{S, T, R\}$ là viết tắt của {SU, TX, RX}.

Với $h_{X,Y}^{x,y}$ là hệ số fading phân phối mũ có trị trung bình là 1 ($\sim \exp(1)$)

Và $g_{X,Y}^{x,y} = \|d\|^{-\eta}$ với η là hệ số suy hao công suất theo số mũ và d là khoảng cách giữa X và Y .

2.3.3. Dung lượng phân phối hệ thống

2.3.3.1. Dung lượng phân phối đến SU

SNR tại SU thứ k từ FBS thứ j được thể hiện như sau:

$$\gamma_{F,S}^{j,k,i} = \frac{u_{j,i} P_F^j G_{F,C}^{j,k}}{N_0} \quad (2.7)$$

SINR tại SU thứ k từ MBS được tính như sau:

$$\gamma_{M,S}^{0,k,i} = \frac{\prod_{j=1}^J (1 - u_{j,i}) P_M^0 G_{M,S}^{0,k}}{N_0 + \sum_{n=1}^N s_{n,i} v_{k,n} p_{n,i} P_T^n G_{T,S}^{n,k}} \quad (2.8)$$

$p_{n,i}$ là xác suất để TX thứ n lưu trữ video thứ i , được xác định bởi

$$p_{n,i} = ar_i + b\beta_n \quad (2.9)$$

với $a, b \in [0,1]$, $a + b = 1$ và độ nổi tiếng của video thứ i được tính theo phân bố Zipf-like [88], như sau:

$$r_i = \frac{i^{-\alpha}}{\sum_{i=1}^I i^{-\alpha}} \quad (2.10)$$

với $\alpha \geq 0$ thể hiện cho độ lệch nổi tiếng giữa các video khác nhau. Nghĩa là nếu $\alpha = 0$ thì tất cả các video có cùng độ nổi tiếng là $\frac{1}{I}$, khi giá trị α cao hơn có nghĩa là độ lệch nổi tiếng giữa các video càng cao.

Với băng thông hệ thống là W và số video là I , dựa vào định luật Shannon, ta tính được tổng dung lượng phân phối hệ thống đến các SU như sau:

$$R_S = W \left\{ \sum_{i=1}^I r_i \sum_{k=1}^K \left[\log_2(1 + \gamma_{M,S}^{0,k,i}) + \sum_{j=1}^J \log_2(1 + \gamma_{F,S}^{j,k,i}) \right] \right\} \quad (2.11)$$

2.3.3.2. Dung lượng phân phối đến TX

Tương tự như với các SU, dung lượng phân phối từ FBS thứ j và MBS tới TX thứ n được tính thông qua SNR tại TX thứ n như sau:

$$\gamma_{F,T}^{j,n,i} = \frac{u_{j,i} P_F^j G_{F,T}^{j,n}}{N_0} \quad (2.12)$$

$$\gamma_{M,T}^{0,n,i} = \frac{\prod_{j=1}^J (1 - u_{j,i}) P_M^0 G_{M,T}^{0,n}}{N_0} \quad (2.13)$$

Từ (2.12) và (2.13), dung lượng phân phối từ FBS thứ j và MBS tới TX thứ n như sau:

$$R_T = W \left\{ \sum_{i=1}^I r_i \sum_{n=1}^N \left[\log_2(1 + \gamma_{M,T}^{0,n,i}) + \sum_{j=1}^J \log_2(1 + \gamma_{F,T}^{j,n,i}) \right] \right\} \quad (2.14)$$

2.3.3.3. Dung lượng phân phối đến RX

Dung lượng phân phối đến RX từ MBS, FBS và từ TX qua truyền thông D2D được tính thông qua SINR và SNR tại RX, cụ thể như sau:

SINR tại RX thứ n được tính trên hệ thống từ TX thứ n :

$$\gamma_{T,R}^{n,k,i} = \frac{s_{n,i} v_{k,n} p_{n,i} P_T^n G_{T,R}^{n,n}}{N_0 + P_M^0 G_{M,R}^{0,n} + \sum_{l=1, l \neq n}^N s_{l,i} v_{k,l} p_{l,i} P_T^l G_{T,R}^{l,l}} \quad (2.15)$$

SNR tại RX thứ n được tính trên hệ thống từ FBS thứ j như sau:

$$\gamma_{F,R}^{j,n,k,i} = \frac{u_{j,i} (1 - s_{n,i} v_{k,n} p_{n,i}) P_F^j G_{F,R}^{j,n}}{N_0} \quad (2.16)$$

SNR tại RX thứ n được tính trên hệ thống từ MBS như sau:

$$\gamma_{M,R}^{0,n,k,i} = \frac{\prod_{j=1}^J (1 - u_{j,i}) (1 - s_{n,i} v_{k,n} p_{n,i}) P_M^0 G_{M,R}^{0,n}}{N_0} \quad (2.17)$$

Như vậy, tổng dung lượng từ MBS, từ FBS và từ TX đến RX được tính như sau:

$$R_R = W \left\{ \sum_{i=1}^I r_i \sum_{n=1}^N \sum_{k=1}^K \left[\log_2(1 + \gamma_{M,R}^{0,n,k,i}) + \log_2(1 + \gamma_{T,R}^{n,k,i}) + \sum_{j=1}^J \log_2(1 + \gamma_{F,R}^{j,n,k,i}) \right] \right\} \quad (2.18)$$

Cuối cùng, từ các công thức (2.11), (2.14) và (2.18), ta có tổng dung lượng trung bình hệ thống cho mỗi MU được tính như sau:

$$R = \frac{R_s + R_T + R_R}{K + 2N} \quad (2.19)$$

2.4. Bài toán tối ưu SCS và thuật giải vét cạn

Bài toán tính tối ưu SCS sẽ bao gồm hàm mục tiêu (2.19) và thỏa 2 ràng buộc gồm 1) dung lượng lưu trữ c_i của FBS và 2) ngưỡng SINR để đảm bảo chất lượng cho SU. Bài toán được trình bày như sau:

$$\begin{aligned} & \max_{u_{j,i}, v_{k,n}} R \\ \text{s. t. } & \begin{cases} \sum_{j=1}^J u_{j,i} \leq c_i, i = 1, 2, \dots, I \\ \sum_{n=1}^N s_{n,i} v_{k,n} p_{n,i} P_T^n G_{T,C}^{n,k} \leq \frac{P_M^0 G_{M,S}^{0,k}}{\gamma_0} - N_0, k = 1, 2, \dots, K, i = 1, 2, \dots, I \end{cases} \end{aligned} \quad (2.20)$$

Trong đó, c_i là số bản sao tối ưu của video thứ i được lưu trữ trong hệ thống sao cho tổng số bản sao trung bình của tất cả các video là cực đại bằng cách giải bài toán quy hoạch tuyến tính sau:

$$\max_{c_i} \sum_{i=1}^I r_i c_i \quad (2.22)$$

$$\text{s. t. } \begin{cases} 0 \leq c_i \leq J, i = 1, 2, \dots, I \\ \sum_{i=1}^I c_i \leq C^*, I \leq C^* \leq I \times J \end{cases} \quad (2.23)$$

Bài toán này được giải bằng cách sử dụng thuật giải vét cạn [54] để tìm nghiệm của U_{JI}^* và V_{KN}^* được trình bày trong Thuật giải 2.1.

Thuật giải 2.1: Tìm kiếm vét cạn

- 1 **Input:** tham số hệ thống được liệt kê ở Bảng 2-2
Khởi tạo 2 không gian tìm kiếm ma trận khả thi $\mathcal{U}^* \in \mathcal{U}$ và $\mathcal{V}^* \in \mathcal{V}$ thỏa điều kiện (2.23)
 - 2 **Output:**
 \mathcal{R}^* : cực đại dung lượng phân phối trung bình đến các MU
 $\{U_{JI}^*, V_{KN}^*\}$
 - 3 **For** mỗi ma trận U_{JI} trong \mathcal{U}^* **do**
 - 4 **For** mỗi ma trận V_{KN} trong \mathcal{V}^* **do**
 - 5 $R(U_{JI}, V_{KN}) = R$, tính công thức (2.19)
 - 6 $\mathcal{R} \leftarrow \mathcal{R} \cup R(U_{JI}, V_{KN})$
 - 7 **End for**
 - 8 **End for**
 - 9 $R^* = \max \mathcal{R}$
 - 10 $\{U_{JI}^*, V_{KN}^*\} = \text{argmax } \mathcal{R}$
-

2.5. Đánh giá kết quả cơ chế SCS

2.5.1. Thiết lập thông số hệ thống

Để thực hiện mô phỏng đánh giá hệ thống, các tham số đầu vào được thể hiện trong Bảng 2-2. Nhằm thuận tiện trong tính toán nhưng không làm mất đi tính tổng quát, tôi bỏ qua ảnh hưởng của hệ số fading (fading coefficient) và chỉ quan tâm tới suy hao theo khoảng cách đường truyền với hệ số mũ suy hao là $\eta = 4$ và khoảng

cách từ MBS đến MU trong phạm vi 300m đến 1500m, từ FBS đến MU trong phạm vi 50m đến 250m, từ SU đến TX trong phạm vi 50m đến 100m, và từ TX đến RX trong phạm vi 1m đến 50m. Trong mỗi đánh giá, cơ chế SCS được so sánh với 4 cơ chế khác gồm: 1) Không xét đến yếu tố xã hội (None-SOA), 2) Không chia sẻ tài nguyên phổ tần kênh truyền xuống (None-DRS), 3) Trung bình của dung lượng phân phối (AVE) và 4) Cực tiểu của dung lượng phân phối (MIN).

Bảng 2-2. Các tham số đầu vào cho bài toán SCS

Tham số	Giá trị
I, J, K, N	5 video, 3 FBS, 3 SU, 5 cặp TX-RX
(θ_n)	(5, 10, 20, 15, 25) [46]
(κ_n)	(1, 4, 3, 2, 5) [46]
(T_{min}^i)	(1, 15, 10, 5, 20) s
(β_n)	(0,1; 0,5; 0,9; 0,3; 0,7)
$\delta, \alpha, \gamma_0, N_0$	10, 2, 5dB, 10^{-13} W
W, P_M^0, P_F^j, P_T^n	5MHz, 5W, 1W, 0.1W
a, b, C^*	0,5; 0,5; $0,7 \times I \times J$

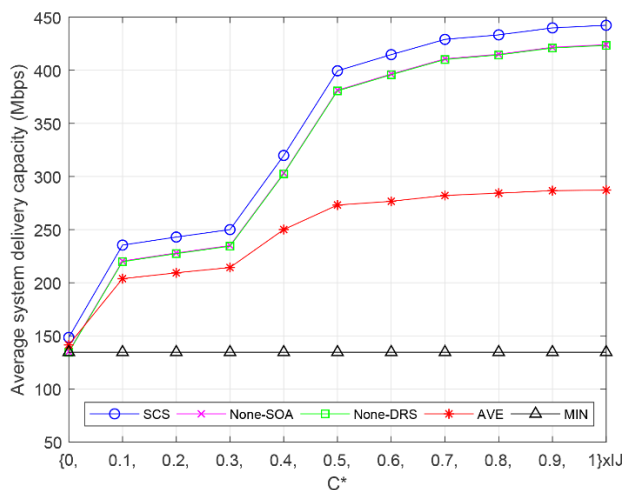
2.5.2. Đánh giá cơ chế SCS

2.5.2.1. Đánh giá SCS theo C^*

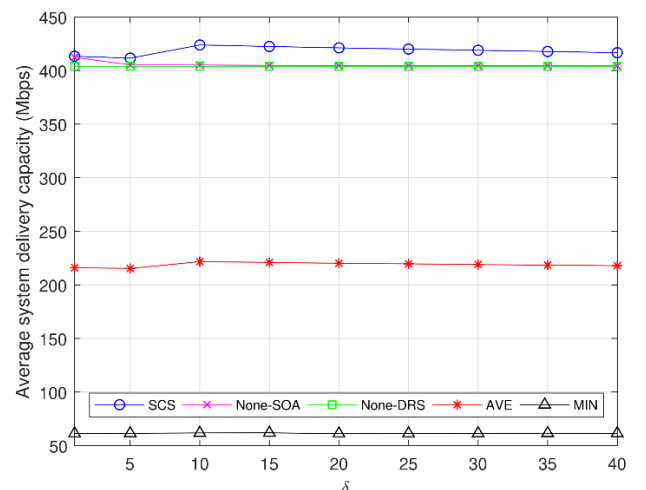
Trước tiên, cơ chế SCS được đánh giá theo dung lượng lưu trữ của FBS bằng cách đặt $C^* = \{0; 0,1; 0,2; \dots; 1\} \times I \times J$, và so sánh với các trường hợp None-SOA, None-DRS, AVE và MIN. Như được hiển thị trong Hình 2-2, chọn $C^* = 0,7 \times I \times J$ phù hợp nhất. Khi tăng C^* , cơ chế SCS luôn đạt hiệu suất cao hơn so với các cơ chế None-SOA, None-DRS, AVE, MIN. Tương tự, trường hợp None-SOA và None-DRS cho dung lượng tốt hơn so với AVE và MIN, trong khi MIN là trường hợp xấu nhất. Kết quả trường hợp None-SOA và None-DRS tương đương nhau cho thấy rằng có quan tâm đến mối quan hệ xã hội người dùng và có chia sẻ tài nguyên đóng vai trò quan trọng như nhau trong việc nâng cao hiệu suất hệ thống.

2.5.2.2. Đánh giá SCS theo δ

Hình 2-3 mô tả đánh giá dung lượng phân phối hệ thống của cơ chế SCS, None-SOA, None-DRS, AVE và MIN theo độ dài (thời lượng) của video bằng cách thay đổi δ trong công thức (2.5) từ 1 đến 40. Kết quả Hình 2-3 cho thấy cơ chế SCS cho kết quả tốt hơn các cơ chế khác và dung lượng hệ thống có xu hướng giảm xuống bằng None-SOA và None-DRS khi δ quá thấp (hoặc quá cao). Điều thú vị là ta có thể điều chỉnh thời lượng của video để đáp ứng các mối quan hệ xã hội của các cặp TX-RX nhằm đạt được dung lượng hệ thống cao nhất (ví dụ: tại $\delta = 10$).



Hình 2-2. Đánh giá dung lượng hệ thống theo khả năng lưu trữ của FBS

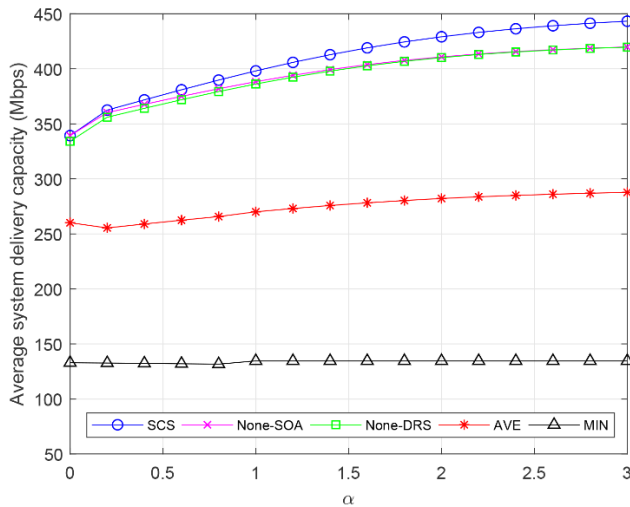


Hình 2-3. Dung lượng hệ thống theo hệ số thời lượng video

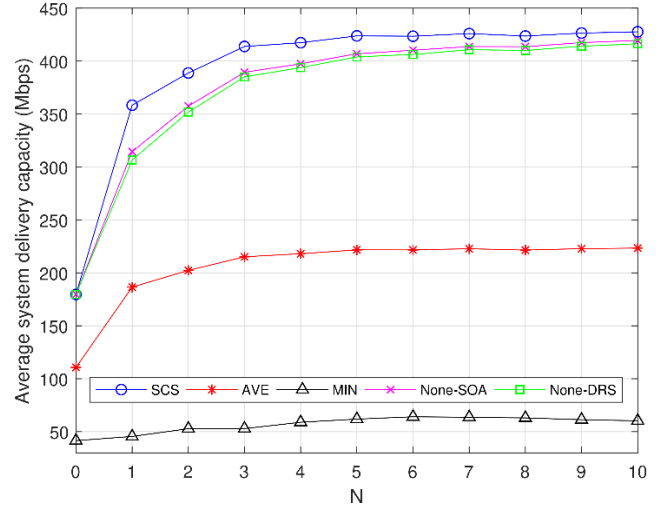
2.5.2.3. Đánh giá SCS theo α

α là hệ số mũ liên quan đến độ lệch về sự nổi tiếng giữa các video. Trong đánh giá này, tôi cho α thay đổi giá trị từ 0 đến 3. Kết quả được thể hiện như Hình 2-4. Trong hình này, khi tăng α , kết quả cho thấy rằng việc

khai thác độ lệch nổi tiếng giữa các video có thể cải thiện hiệu năng hệ thống. Nếu tất cả các video có cùng độ nổi tiếng, tức là $\alpha = 0$, cơ chế SCS không có được lợi ích cao khi so với None-SOA và None-DRS. Tuy nhiên, SCS đạt dung lượng phân phối hệ thống cao hơn khi tăng α . Ngoài ra, việc tập trung phục vụ một số video có độ nổi tiếng cao hơn (khi α tăng) sẽ mang lại dung lượng hệ thống cao hơn. Và qua đánh giá này, cơ chế SCS vẫn cho dung lượng hệ thống tốt nhất so với None-SOA, None-DRS, AVE và MIN.



Hình 2-4. Dung lượng hệ thống theo α

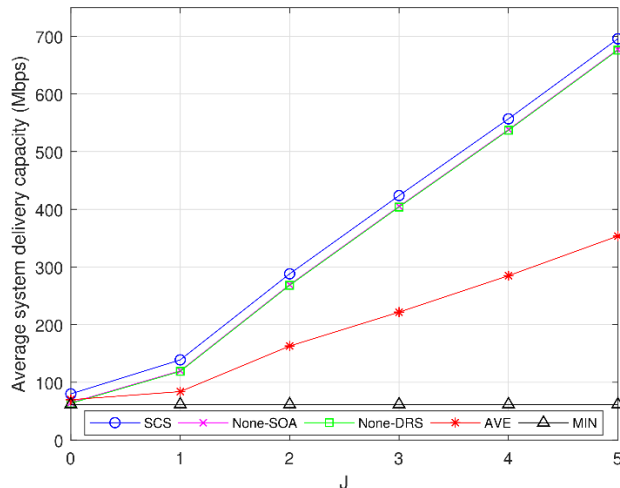


Hình 2-5. Dung lượng hệ thống theo số cặp D2D

2.5.2.4. Đánh giá SCS theo N

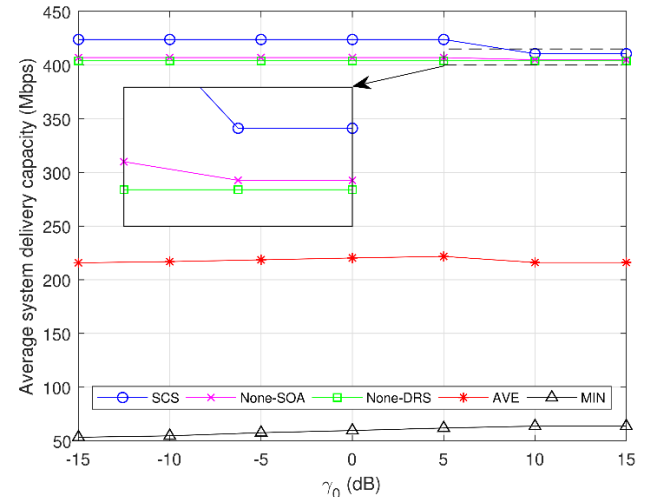
Đánh giá dung lượng hệ thống bằng cách thay đổi số lượng cặp TX-RX N từ 0 đến 10 (Hình 2-5). Dễ dàng thấy rằng nếu $N = 0$, hệ thống không thu được bất kỳ lợi ích nào từ việc chia sẻ tài nguyên phổ tần kênh truyền xuống và môi quan hệ xã hội vì không có truyền thông D2D. Do đó, SCS, None-SOA và None-DRS có cùng kết quả và dung lượng phân phối hệ thống thấp. Nếu N tăng, dung lượng phân phối hệ thống sẽ cao hơn và khi N tăng đến một mức nào đó thì dung lượng phân phối hệ thống trở nên bão hòa. Tình trạng bão hòa là do các ràng buộc đối với SINR của SU làm giới hạn số lượng các cặp truyền thông của D2D. Trong trường hợp có nhiều cặp D2D dày đặc của 5G UDN, chúng ta cần thận trọng chọn số lượng cặp TX-RX sao cho SINR của SU được đảm bảo và dung lượng phân phối hệ thống đủ cao trước khi đạt đến tình huống bão hòa. Kết quả chứng minh rằng SCS vẫn cho dung lượng hệ thống tốt nhất so với None-SOA, None-DRS, AVE và MIN.

2.5.2.5. Đánh giá hệ thống theo J



Hình 2-6. Dung lượng hệ thống theo số lượng FBS

Hình 2-6 biểu thị dung lượng phân phối hệ thống theo số lượng FBS J. Kết quả cho thấy số lượng FBS ảnh hưởng đáng kể đến dung lượng phân phối hệ thống. Không giống như khi tăng số lượng cặp D2D, việc tăng số lượng FBS trong 5G UDN làm cho dung lượng phân phối hệ thống được cải thiện nhanh chóng mà không bị bão hòa vì các điều kiện đối với can nhiễu liên tầng và đồng tầng không được xem xét nhờ sử dụng mô hình phân kênh và F-ALOHA [47, 48].



Hình 2-7. Dung lượng hệ thống theo Gamma

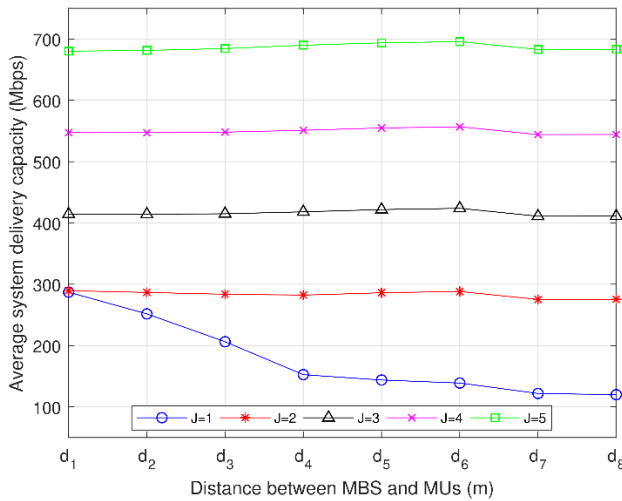
2.5.3. Đánh giá hệ thống theo γ_0

Trong phần này, tôi kiểm chứng hiệu năng hệ thống dưới tác động bởi ngưỡng SINR (γ_0) của SU như được mô tả trong Hình 2-7. Như kết quả được mô tả trong Hình 2-7, khi γ_0 tăng, dung lượng phân phối hệ thống của SCS và None-SOA giảm và giảm dần bằng với dung lượng phân phối trong trường hợp None-DRS. Lý do là nếu γ_0 thấp, nhiều cặp TX-RX được thiết lập nhờ tài nguyên kênh truyền xuống được chia sẻ bởi SU nên dung lượng phân phối hệ thống cao hơn. Ngược lại, nếu γ_0 cao, số cặp TX-RX sẽ ít đi để giảm can nhiễu lên SU nên dung lượng phân phối hệ thống thấp hơn. Dung lượng phân phối hệ thống của MIN tăng vì khi γ_0 tăng, nhiễu ma trận ứng viên gây ra tác động nhiễu cao cho SU được loại khỏi không gian tìm kiếm \mathcal{V} . Rõ ràng thấy rằng, kết quả của None-DRS không bị ảnh hưởng bởi γ_0 vì không có hiệu ứng nhiễu từ truyền thông D2D trên SU. Trong trường hợp này, cơ chế SCS cũng cho kết quả tốt hơn None-SOA, None-DRS, AVE và MIN.

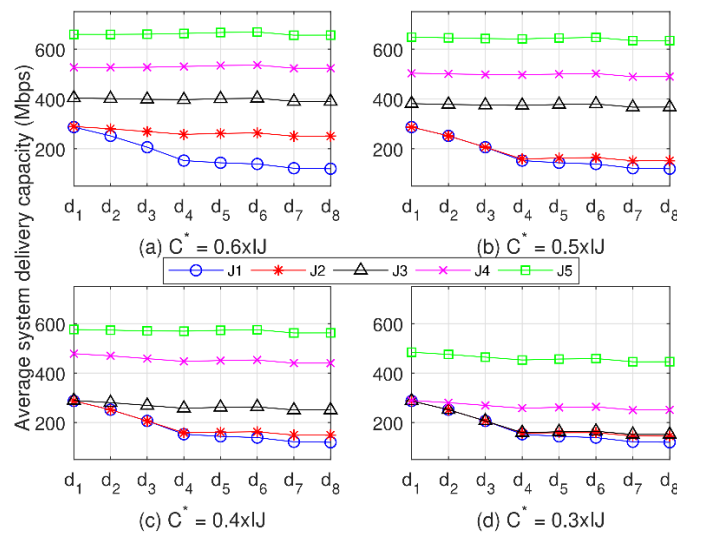
2.5.4. Đánh giá hệ thống theo khoảng cách giữa MBS và các MU

Cuối cùng, tôi đánh giá dung lượng phân phối hệ thống theo khoảng cách giữa MBS và MU: $d_1 \in [5;10]m$, $d_2 \in [10;20]m$, $d_3 \in [20;100]m$, $d_4 \in [100;500]m$, $d_5 \in [200;1000]m$, $d_6 \in [300;1500]m$, $d_7 \in [400;2000]m$ và $d_8 \in [500;2500]m$. Trong đánh giá này, tôi chỉ xem xét dung lượng hệ thống của SCS với số lượng FBS J khác nhau từ 1 đến 5. Kết quả trong Hình 2-8 cho thấy rằng nếu số lượng FBS thấp (nghĩa là $J = 1$ và $J = 2$), các MU có nhiều khả năng chỉ được MBS phục vụ cho nên dung lượng phân phối hệ thống thấp. Dễ dàng thấy rằng nếu các MU quá gần MBS, khi đó tất cả các MU hầu như chỉ được phục vụ bởi MBS nên các MU đều nhận cùng dung lượng phân phối của hệ thống. Tuy nhiên, nếu số lượng FBS tăng (tức là, $J = 3$, $J = 4$ và $J = 5$), các MU có nhiều khả năng được phục vụ thêm bởi FBS, nên khi đó dung lượng phân phối hệ thống cao hơn. Do ta thiết lập khả năng lưu trữ của FBS khá cao ($C^* = 0.7 \times I \times J$), nên dù các MU ở rất gần với MBS nhưng MBS không phục vụ các MU đó mà chúng đã được các FBS phục vụ. Một điều thú vị ở đây là, tồn tại khu vực tốt nhất (tức là, $d_6 \in [300, 1500]m$) trong đó các MU được phục vụ với dung lượng phân phối hệ thống cao nhất khi chúng không ở quá gần và không quá xa MBS. Nói cách khác, các MU trong khu vực này được phục vụ tối ưu bởi sự cộng tác của tất cả MBS, FBS và TX.

Để hiểu rõ hơn, tôi kiểm chứng thêm các kịch bản mô phỏng khi dung lượng lưu trữ của FBS thấp hơn, cụ thể là $C^* = \{0.6; 0.5; 0.4; 0.3\} \times I \times J$ như trong Hình 2-9.



Hình 2-8. Dung lượng hệ thống theo khoảng cách giữa MBS và các MU



Hình 2-9. Dung lượng hệ thống theo khoảng cách giữa MBS và các MU với C^* thấp hơn

Có thể thấy rằng nếu dung lượng lưu trữ của FBS quá thấp ($C^* = 0.3 \times I \times J$) (Hình 2-9d), thì các MU chủ yếu được MBS phục vụ khi chúng ở gần MBS hơn. Đặc biệt, trong trường hợp $d_1 \in [5, 10]m$, hiệu suất của SCS là như nhau đối với tất cả các trường hợp $J = 1$, $J = 2$, $J = 3$ và $J = 4$ vì tất cả các MU chỉ được phục vụ bởi MBS. Rõ ràng, nếu C^* tăng, khả năng các MU được phục vụ bởi MBS sẽ giảm. Điều quan trọng là sự gia tăng của cả C^* và J trong 5G UDN có thể giảm tải lưu lượng tuyến trực của MBS.

2.6. Kết luận chương 2

Trong Chương 2, tôi đã đề xuất một cơ chế lưu trữ và chia sẻ tài nguyên (SCS) tối ưu có xét đến mối quan hệ xã hội của người dùng để phân phối video với dung lượng cao trong 5G UDN. Cụ thể, việc lưu trữ nội dung video tại MBS, FBS và TX được sử dụng để đưa video đến gần hơn với các MU và tài nguyên phổ tần kênh truyền của SU được chia sẻ với các cặp TX-RX trong truyền thông D2D. Để hiệu quả hơn, mối quan hệ xã hội của các cặp TX-RX và độ phổ biến của các video được khai thác trong cơ chế SCS. Cơ chế SCS được phân tích cẩn thận bằng cách tính đến ngưỡng SINR của các SU để đảm bảo QoS của chúng. Ngoài ra, kết quả cũng cho thấy rằng: 1) tồn tại một giá trị thời lượng trung bình cho nhóm các video được quan tâm và 2) tồn tại khu vực tốt nhất mà ở đó hệ thống đạt được dung lượng phân phối cao nhất. Những phát hiện này có thể giúp cả nhà cung cấp dịch vụ Internet và nhà cung cấp nội dung hiểu rõ hơn về đặc điểm của 5G UDN để phục vụ các MU hiệu quả hơn.

Tuy nhiên, cơ chế SCS vẫn còn những hạn chế như: 1) chưa quan tâm đến độ phân giải của thiết bị người dùng, từ đó, xác định nhu cầu người dùng để lưu trữ video với các phiên bản khác nhau (chất lượng/độ phân giải khác nhau) nhằm đồng thời sử dụng hiệu quả dung lượng lưu trữ và đáp ứng được nhu cầu người dùng; 2) các cặp thu phát video D2D được định sẵn, nghĩa là, chưa cho phép thiết bị thu chọn lựa thiết bị phát có lưu trữ video được yêu cầu; và 3) thông số đánh giá hệ thống là dung lượng phân phối video (được tính bằng bps) không đặc tả được chất lượng video tại đầu thu một cách tường minh. Những hạn chế này sẽ được khắc phục trong Chương 3.

CHƯƠNG 3. LƯU TRỮ ĐA PHÂN GIẢI VÀ CHIA SẺ TÀI NGUYÊN TỐI ƯU THEO NHU CẦU NGƯỜI DÙNG NÂNG CAO CHẤT LƯỢNG TRUYỀN VIDEO TRONG 5G UDN

3.1. Giới thiệu cơ chế CRS

Chương 3 trình bày cơ chế lưu trữ đa phân giải và chia sẻ tài nguyên tối ưu (CRS) theo nhu cầu người dùng nhằm nâng cao chất lượng truyền video trong mạng 5G siêu dày đặc (UDN). Cơ chế CRS một mặt vẫn cho phép người dùng di động (MU) truy xuất video linh hoạt từ 3 tầng gồm: trạm cơ sở chính (MBS), trạm cơ sở nhỏ femto (FBS) và từ người dùng khác có lưu trữ video (CU) thông qua truyền thông từ thiết bị đến thiết bị (D2D), mặt khác khắc phục những vấn đề của cơ chế SCS trong Chương 2 liên quan đến cách thức lưu trữ, chia sẻ tài nguyên và thông số đánh giá hệ thống.

Về cách thức lưu trữ, cơ chế CRS có xem xét đến độ phân giải của thiết bị người dùng, từ đó, xác định nhu cầu người dùng và lưu trữ video với các phiên bản có độ phân giải (chất lượng) khác nhau nhằm sử dụng hiệu quả dung lượng lưu trữ đáp ứng nhu cầu người dùng. Về cách thức chia sẻ tài nguyên, cơ chế CRS cho phép người dùng thông thường (NU – Normal User) chọn lựa CU và chọn lựa người dùng sở hữu phổ tần (SU) để thiết lập bộ 3 (SU, CU, NU) cho phiên truyền thông video D2D, thay vì thiết lập bộ 2 giữa SU và cặp (CU, NU) được định sẵn như trong cơ chế SCS. Việc thiết lập bộ 3 sẽ cho phép mở rộng sự lựa chọn để có được phiên truyền thông video D2D tốt nhất. Về thông số đánh giá hệ thống, trong Chương 3, hiệu suất hệ thống được đánh giá tường minh hơn bằng cách chuyển đổi dung lượng phân phối video sang tỉ số tín hiệu cực đại trên nhiễu (PSNR) được đo lường bằng dB, thông qua mối liên hệ giữa dung lượng và chất lượng video và các bước tính toán xác suất đạt dung lượng trung gian.

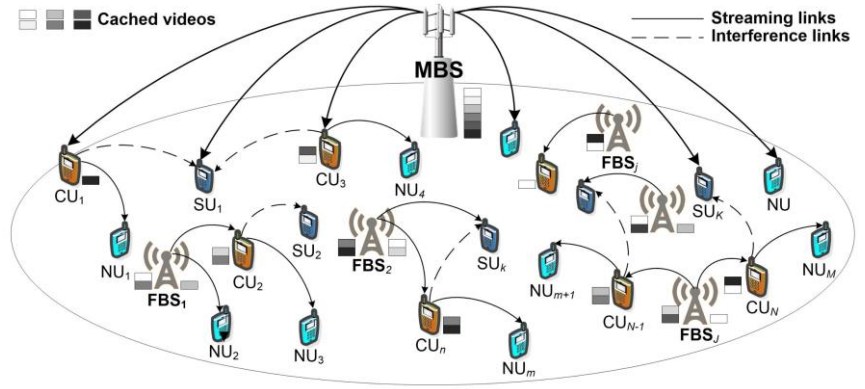
Từ những cải tiến nêu trên, bài toán CRS được xây dựng để xác định đồng thời: 1) video được phân phối từ tầng nào, 2) video với độ phân giải nào sẽ được lưu trữ trong FBS nào, và 3) các bộ 3 (SU, CU, NU) nào được chọn? Mục tiêu của cơ chế CRS là cực đại chất lượng phát lại (playback) video và sử dụng hiệu quả dung lượng lưu trữ tại FBS theo nhu cầu người dùng. Công trình nghiên cứu liên quan đến Chương 3 là [C2] và [C3] được trình bày tại “Danh mục các công trình đã công bố của nghiên cứu sinh” ở cuối luận án.

3.2. Mô hình và tính toán các thông số hệ thống

3.2.1. Mô hình hệ thống

Trong Chương 3, mô hình truyền video trong 5G UDN với cơ chế CRS được đề xuất như Hình 3-1. Các thông số tính toán liên quan đến mô hình được liệt kê và diễn giải trong Bảng 3-1.

Trong mô hình này, có một MBS, J FBS, và các MU có nhu cầu truy xuất nội dung video bao gồm K người dùng sẵn sàng chia sẻ tài nguyên phổ tần kênh truyền xuống (SU), N người dùng đã lưu trữ một số video có độ phân giải khác nhau (CU), M người dùng thông thường (NU) và I video. Video thứ i , $i = 1, 2, \dots, I$, có V_i mức phân giải được mã hóa ở các tốc độ khác nhau tương ứng với các mức chất lượng khác nhau.



Hình 3-1. Mô hình hệ thống của CRS

Các video này được gửi từ MBS đến các MU thông qua kênh truyền di động thông thường, từ FBS đến các MU bằng cách sử dụng các cơ chế phân kênh và F-ALOHA để tránh can nhiễu [48] và từ CU đến NU qua truyền thông từ thiết bị đến thiết bị (D2D) bằng cách tái sử dụng tài nguyên phổ tần kênh truyền xuống được chia sẻ bởi SU.

Bảng 3-1. Các ký hiệu sử dụng cho bài toán CRS

Ký hiệu	Diễn giải
I	Số video
J	Số trạm cơ sở nhỏ femto (FBS)
K	Số người dùng di động chấp nhận chia sẻ tài nguyên phổ tần kênh truyền xuống (SU)
N	Số người dùng di động đã lưu trữ các video (CU) có độ phân giải khác nhau
M	Số người dùng di động thông thường (NU)
V_i	Số mức phân giải của video thứ i , $i=1, 2, \dots, I$
$w_{n,m}^k$	Chỉ số chia sẻ tài nguyên phổ tần kênh truyền xuống, khi $w_{n,m}^k = 1$ thì SU thứ k sẽ chia sẻ kênh truyền xuống của nó cho CU thứ n và NU thứ m thực hiện truyền video qua truyền thông D2D
$u_j^{v_i}$	Chỉ số lưu trữ video tại FBS, khi $u_j^{v_i} = 1$ thì FBS thứ j lưu phiên bản video v_i và ngược lại
P_M^0	Công suất phát của trạm cơ sở lớn (MBS)
P_F^j	Công suất phát của trạm cơ sở nhỏ femto (FBS) thứ j
P_T^n	Công suất phát của thiết bị CU thứ n trong truyền thông D2D.
G_0^k	Độ lợi kênh từ MBS tới SU k
N_0	Công suất nhiễu trắng Gaussian (AWGN – Additive White Gaussian Noise)
W	Băng thông hệ thống
L_F^i	Tổng dung lượng cần có để lưu video thứ i với các mức phân giải khác nhau trong các FBS theo tính toán
L_{max}^i	Giới hạn trên về dung lượng bộ nhớ của các FBS để lưu trữ video thứ i với các mức phân giải khác nhau, cụ thể $L_{max}^i = r_i \times I \times J \times \max\{L_i^{v_i}, v_i = 1, 2, \dots, V_i\},$ trong đó $r_i = i^{-\alpha} (\sum_{i=1}^I i^{-\alpha})^{-1}$ là độ nổi tiếng của video với hệ số mũ phản ánh độ lệch nổi tiếng α được mô hình hóa dưới dạng phân phối Zipf-like [45] và $L_i^{v_i}$ là kích thước của video thứ i với mức phân giải v_i
C	Tổng dung lượng (bps) được yêu cầu bởi các MU theo tính toán
C^*	Giới hạn trên về tổng dung lượng được yêu cầu bởi các MU
μ và δ	μ và δ với $0 < \mu, \delta \leq 1$ là hệ số được sử dụng để điều chỉnh giới hạn trên L_{max}^i và C^*

Trong mô hình truyền video với cơ chế CRS, giả sử các tham số hệ thống không thay đổi trong quá trình truyền phiên bản video dài nhất trong số tất cả các phiên bản được yêu cầu. Cũng cần nhấn mạnh thêm rằng cơ chế CRS sẽ phục vụ hiệu quả cho các MU trong khu vực đông đúc như, sân vận động, phòng hòa nhạc, hội trường,

bảo tàng, tòa nhà văn phòng làm việc, bệnh viện, khuôn viên trường học, v.v.... Bất cứ khi nào MBS dự đoán có sự gia tăng yêu cầu truy cập các video nào đó, thì nó sẽ thực thi cơ chế CRS theo 3 bước: i) cập nhật các tham số hệ thống, ii) thiết lập và giải bài toán tối ưu CRS, iii) truyền phát video, cụ thể như sau:

- Đầu tiên, MBS thu thập các thông số của hệ thống như đặc tính kênh truyền; dung lượng lưu trữ của FBS và CU; thông tin video (các mức phân giải, kích thước, và mức độ phổ biến của từng video); dung lượng yêu cầu bởi các MU; và tỉ số tín hiệu trên can nhiễu và nhiễu trắng (SINR) của SU cùng các thông số liên quan khác được trình bày trong Bảng 3-1.
- Tiếp theo, dựa vào các thông số của hệ thống, các biểu thức liên quan đến hàm mục tiêu và các ràng buộc được tính toán để xây dựng bài toán tối ưu có ràng buộc CRS. Bài toán CRS sẽ được giải để tìm ra chỉ số $u_j^{v_i}$ tối ưu nhằm xác định các phiên bản video phù hợp và vị trí tối ưu để lưu trữ chúng trong các FBS ($u_j^{v_i} = 1$ nghĩa là video thứ i với mức phân giải v_i được lưu trữ tại FBS j ; và ngược lại thì $u_j^{v_i} = 0$). Đồng thời, việc giải bài toán CRS cũng xác định được chiến lược chia sẻ tài nguyên phổ tần kênh truyền xuống tối ưu, cụ thể là tìm ra chỉ số $w_{n,m}^k$ tối ưu để xác định các bộ 3 (SU, CU, NU) cho truyền thông video D2D ($w_{n,m}^k = 1$ nghĩa là SU k đồng ý chia sẻ phổ tần kênh truyền xuống của nó cho CU n thực hiện truyền video đến NU m qua truyền thông D2D; ngược lại thì $w_{n,m}^k = 0$). Phân bố của các CU được chia sẻ tài nguyên phổ tần kênh truyền xuống từ SU k , $k = 1, 2, \dots, K$, được mô hình bởi quá trình Poisson đồng nhất (Poisson Point Process) với mật độ λ_c^k [49]. Bài toán tối ưu CRS cũng tính đến các ràng buộc về dung lượng mà các MU yêu cầu để phát lại (playback) video, về dung lượng lưu trữ của FBS để sử dụng tài nguyên lưu trữ một cách hiệu quả và về tỉ số tín hiệu trên can nhiễu và giao thoa (SINR) để đảm bảo chất lượng dịch vụ (QoS) của SU.
- Cuối cùng, tất cả các video với mức phân giải khác nhau được lưu trong FBS cũng như tất cả các bộ 3 (SU, CU, NU) được chỉ định để sẵn sàng truyền từ MBS, FBS và CU đến các MU với chất lượng phát lại cao đồng thời đáp ứng được nhu cầu về dung lượng của các MU, sử dụng hiệu quả dung lượng lưu trữ của FBS và đảm bảo tỉ số tín hiệu trên can nhiễu và giao thoa (SINR) của SU (hay nói cách khác là đảm bảo chất lượng/sự công bằng cho SU).

3.3. Tính toán các thông số hệ thống với cơ chế CRS

3.3.1. Xác suất đạt dung lượng tại SU

Để xác định được xác suất đạt dung lượng tại SU thứ k , ta cần tính dung lượng đạt tại SU thứ k từ FBS thứ j :

$$C_j^{k,v_i} = W \log_2 \left(1 + \frac{u_j^{v_i} P_j G_j^k}{N_0} \right) \quad (3.1)$$

và từ MBS:

$$C_0^{k,v_i} = W \log_2 \left(1 + \frac{\prod_{j=1}^J (1 - u_j^{v_i}) P_0 G_0^k}{N_0 + I_{C,S}^{k,v_i}} \right) \quad (3.2)$$

trong đó $I_{C,S}^{k,v_i} = \sum_{n=1}^N \sum_{m=1}^M w_{n,m}^k p_n^{v_i} P_C^k G_n^k$ và $p_n^{v_i}$ là xác suất mà CU thứ n lưu trữ video thứ i với mức phân giải v_i , được tính như sau:

$$p_n^{v_i} = a r_i + b \theta_n^{v_i} \quad (3.3)$$

với $a, b \in [0, 1]$ và r_i là độ nổi tiếng của video thứ i như được trình bày trong công thức (2.10).

Khi đó, xác suất đạt dung lượng tại SU thứ k từ FBS thứ j là

$$p_j^{k,v_i} = Pr\{C_j^{k,v_i} \geq C_{th}^{k,v_i}\} = \exp\left(\frac{-\xi_j^{k,v_i} N_0}{u_j^{v_i} P_j}\right) \quad (3.6)$$

và xác suất đạt dung lượng tại SU thứ k từ MBS là

$$p_0^{k,v_i} = Pr\{C_0^{k,v_i} \geq C_{th}^{v_i}\} = exp\left\{-\xi_0^{k,v_i} \left[\lambda_C^{k,v_i} \left(\frac{P_C^k}{\prod_{j=1}^J (1-u_j^{v_i}) P_0}\right)^{\frac{2}{\eta}}\right]\right\} \quad (3.7)$$

với

- $\xi_j^{k,v_i} = (d_j^k)^\eta \left(2^{\frac{C_{th}^{v_i}}{W}} - 1\right)$, $\xi_0^{k,v_i} = \pi(d_0^k)^2 \Gamma\left(1 + \frac{2}{\eta}\right) \Gamma\left(1 - \frac{2}{\eta}\right) \left(2^{\frac{C_{th}^{v_i}}{W}} - 1\right)^{2/\eta}$

- d_j^k và d_0^k là khoảng cách từ FBS thứ j và MBS đến SU thứ k

- $\lambda_C^{k,v_i} = \sum_{n=1}^N \sum_{m=1}^M w_{n,m}^k p_n^{v_i}$ là mật độ CU trong bán kính của tế bào mạng

Như vậy, xác suất đạt dung lượng để truyền video thứ i với mức phân giải v_i từ FBS thứ j hoặc từ MBS đến SU thứ k được tính như sau:

$$p_{0,j}^{k,v_i} = 1 - (1 - p_j^{k,v_i})(1 - p_0^{k,v_i}) \quad (3.8)$$

3.3.2. Xác suất đạt dung lượng tại CU

Dung lượng tại CU thứ n từ FBS j và MBS được tính bằng:

$$C_{j,n}^{v_i} = W \log_2 \left(1 + \frac{u_j^{v_i} P_j G_j^n}{N_0}\right) \quad (3.9)$$

và

$$C_{0,n}^{v_i} = W \log_2 \left(1 + \frac{\prod_{j=1}^J (1 - u_j^{v_i}) P_0 G_0^n}{N_0}\right) \quad (3.10)$$

Như vậy, ta có được xác suất đạt dung lượng tương ứng tại CU thứ n như sau:

- Từ FBS thứ j tới CU thứ n :

$$p_{j,n}^{v_i} = Pr\{C_{j,n}^{v_i} \geq C_{th}^{v_i}\} = exp\left(\frac{-\xi_j^{n,v_i} N_0}{u_j^{v_i} P_j}\right) \quad (3.11)$$

- Và từ MBS tới CU thứ n :

$$p_{0,n}^{v_i} = Pr\{C_{0,n}^{v_i} \geq C_{th}^{v_i}\} = exp\left(\frac{-\xi_0^{n,v_i} N_0}{\prod_{j=1}^J (1 - u_j^{v_i}) P_0}\right) \quad (3.12)$$

với

- $\xi_j^{n,v_i} = (d_j^n)^\eta \left(2^{\frac{C_{th}^{v_i}}{W}} - 1\right)$ và $\xi_0^{n,v_i} = \pi(d_0^n)^\eta \left(2^{\frac{C_{th}^{v_i}}{W}} - 1\right)$

- d_j^n và d_0^n là khoảng cách từ FBS thứ j và MBS tới CU thứ n .

Khi đó, xác suất đạt dung lượng để truyền video thứ i với mức phân giải v_i từ FBS thứ j hoặc MBS đến CU thứ n là:

$$p_{0,j,n}^{v_i} = 1 - (1 - p_{j,n}^{v_i})(1 - p_{0,n}^{v_i}) \quad (3.13)$$

3.3.3. Xác suất đạt dung lượng tại các NU

Khác với SU và CU, dung lượng tại NU thứ m , đến từ các CU thứ n , FBS thứ j và MBS, được tính tương ứng là:

$$C_{n,m}^{k,v_i} = W \log_2 \left(1 + \frac{w_{n,m}^k p_n^{v_i} P_C^k G_n^m}{N_0 + P_0 G_0^m + I_{C,C}^{k,v_i}} \right) \quad (3.14)$$

$$C_{j,m}^{k,v_i} = W \log_2 \left(1 + \frac{u_j^{v_i} (1 - w_{n,m}^k p_n^{v_i}) P_j G_j^m}{N_0} \right) \quad (3.15)$$

và

$$C_{0,m}^{k,v_i} = W \log_2 \left(1 + \frac{\prod_{j=1}^J (1 - u_j^{v_i}) (1 - w_{n,m}^k p_n^{v_i}) P_0 G_0^m}{N_0} \right) \quad (3.16)$$

$$\text{với } I_{C,C}^{k,v_i} = \sum_{\substack{n'=1 \\ n' \neq n}}^N \sum_{\substack{m'=1 \\ m' \neq m}}^M w_{n',m'}^k p_{n'}^{v_i} P_C^k G_{n',m'}^k$$

và xác suất đạt dung lượng tương ứng là:

$$p_{n,m}^{k,v_i} = Pr\{C_{n,m}^{k,v_i} \geq C_{th}^{v_i}\} = \exp \left\{ -\xi_{n,m}^{v_i} \left[\lambda_M \left(\frac{P_0}{w_{n,m}^k p_n^{v_i} P_C^k} \right)^{\frac{2}{\eta}} + \lambda_C^{k,v_i} \right] \right\} \quad (3.17)$$

$$p_{j,m}^{k,v_i} = Pr\{C_{j,m}^{k,v_i} \geq C_{th}^{v_i}\} = \exp \left[\frac{-\xi_{j,m}^{v_i} N_0}{u_j^{v_i} (1 - w_{n,m}^k p_n^{v_i}) P_j} \right] \quad (3.18)$$

và

$$p_{0,m}^{k,v_i} = Pr\{C_{0,m}^{k,v_i} \geq C_{th}^{v_i}\} = \exp \left[\frac{-\xi_{0,m}^{v_i} N_0}{\prod_{j=1}^J (1 - u_j^{v_i}) (1 - w_{n,m}^k p_n^{v_i}) P_0} \right] \quad (3.19)$$

trong đó

- $\xi_{j,m}^{v_i} = (d_j^m)^\eta \left(2^{\frac{C_{th}^{v_i}}{W}} - 1 \right)$
- $\xi_{0,m}^{v_i} = (d_0^m)^\eta \left(2^{\frac{C_{th}^{v_i}}{W}} - 1 \right)$
- $\xi_{n,m}^{v_i} = \pi (d_n^m)^2 \Gamma \left(1 + \frac{2}{\eta} \right) \Gamma \left(1 - \frac{2}{\eta} \right) \left(2^{\frac{C_{th}^{v_i}}{W}} - 1 \right)^{2/\eta}$
- d_n^m, d_j^m và d_0^m là khoảng cách từ CU thứ n , FBS thứ j và MBS đến NU thứ m
- $\lambda_C^{k,v_i} = \sum_{\substack{n'=1 \\ n' \neq n}}^N \sum_{\substack{m'=1 \\ m' \neq m}}^M w_{n',m'}^k p_{n'}^{v_i}$ trong phạm vi bán kính của tế bào mạng

Cuối cùng, xác suất đạt dung lượng tại NU m được tính như sau:

$$p_{0,j,n,m}^{k,v_i} = 1 - (1 - p_{n,m}^{k,v_i}) (1 - p_{j,m}^{k,v_i}) (1 - p_{0,m}^{k,v_i}) \quad (3.20)$$

3.3.4. Chất lượng trung bình của video truyền

Nếu video thứ i của phiên bản v_i được phát lại ở tốc độ (hoặc dung lượng) $C_{th}^{v_i}$, độ méo (distortion) tương ứng được tính bởi:

$$D_i(C_{th}^{v_i}) = \gamma_i (C_{th}^{v_i})^{\beta_i} \quad (3.21)$$

trong đó γ_i và β_i là các tham số phụ thuộc được chọn sao cho công thức (3.21) đáp ứng các đường cong RD thực nghiệm. Khi đó, ta có thể tính toán các giá trị chất lượng trung bình của các video với các mức phân giải khác nhau nhận tại các MU (SU, CU, NU) như sau:

$$\bar{Q} = \frac{\sum_{j=1}^J (Q_S^j + Q_C^j + Q_N^j)}{3J} \quad (3.22)$$

với

$$Q_S^j = \frac{1}{K} \sum_{k=1}^K \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} p_{0,j}^{k,v_i} Q(D_i(C_{th}^{v_i})) \quad (3.23)$$

$$Q_C^j = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} p_{0,j,n}^{v_i} Q(D_i(C_{th}^{v_i})) \quad (3.24)$$

$$Q_N^j = \frac{1}{KMN} \sum_{k=1}^K \sum_{N=1}^N \sum_{m=1}^M \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} p_{0,j,n,m}^{k,v_i} Q(D_i(C_{th}^{v_i})) \quad (3.25)$$

Trong đó $Q(D_i(C_{th}^{v_i})) = 10 \log_{10} \frac{255^2}{D_i(C_{th}^{v_i})}$ là tỉ số tín hiệu cực đại trên nhiễu (PSNR).

3.4. Giải bài toán tối ưu cho CRS và giải thuật di truyền

3.4.1. Bài toán tối ưu CRS

Để xây dựng bài toán tối ưu CRS, ta tính toán thêm dung lượng lưu trữ trên FBS (L_F^i) được sử dụng để lưu trữ video i với các mức phân giải khác nhau và tổng dung lượng (C) mà hệ thống cung cấp cho các MU và xem xét chúng như các ràng buộc trong quá trình giải bài toán tối ưu CRS. Cụ thể, L_F^i và C tương ứng được tính bởi:

$$L_F^i = \sum_{j=1}^J \sum_{v_i=1}^{V_i} u_j^{v_i} L_i^{v_i} \quad (3.26)$$

$$C = C_S + C_C + C_N \quad (3.27)$$

trong đó

$$C_S = \sum_{j=1}^J \sum_{k=1}^K \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} (C_j^{k,v_i} + C_0^{k,v_i}) \quad (3.28)$$

$$C_C = \sum_{j=1}^J \sum_{n=1}^N \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} (C_{j,n}^{v_i} + C_{0,n}^{v_i}) \quad (3.29)$$

$$C_N = \sum_{j=1}^J \sum_{k=1}^K \sum_{N=1}^N \sum_{m=1}^M \sum_{i=1}^I \frac{r_i}{V_i} \sum_{v_i=1}^{V_i} (C_{n,m}^{k,v_i} + C_{j,m}^{k,v_i} + C_{0,m}^{k,v_i}) \quad (3.30)$$

Cuối cùng, bài toán tối ưu CRS được mô tả như sau:

$$\max_{u_j^{v_i}, w_{n,m}^k} \bar{Q} \quad (3.31)$$

$$s. t. \begin{cases} \sum_{v_i=1}^{V_i} u_j^{v_i} \leq 1, i = 1, 2, \dots, I, j = 1, 2, \dots, J, \\ \sum_{m=1}^M w_{n,m}^k \leq 1, k = 1, 2, \dots, K, n = 1, 2, \dots, N, \\ \sum_{n=1}^N w_{n,m}^k \leq 1, k = 1, 2, \dots, K, m = 1, 2, \dots, M, \\ \sum_{k=1}^K w_{n,m}^k \leq 1, n = 1, 2, \dots, N, m = 1, 2, \dots, M, \\ L_F^i \leq \mu L_{max}^i, i = 1, 2, \dots, I, \\ C \leq \delta C^* \\ I_{C,S}^{k,v_i} \leq \frac{P_0 G_0^k}{\gamma_0} - N_0, k = 1, 2, \dots, K, i = 1, 2, \dots, I, v_i = 1, 2, \dots, V_i \end{cases} \quad (3.32)$$

3.4.2. Giải thuật di truyền cho bài toán CRS

Để giải bài toán tối ưu CRS, trong chương này, tôi áp dụng công cụ “Giải thuật di truyền trong Matlab” được trình bày trong tài liệu [62]. Tuy nhiên, công cụ này có hạn chế là chỉ có thể giải các bài toán có ràng buộc cận trên và cận dưới đơn giản của biến chứ không thể giải các bài toán có ràng buộc phức tạp như ràng buộc trong bài toán CRS nêu trên. Do đó, giải pháp là chuyển bài toán tối ưu có ràng buộc thành bài toán tối ưu không có ràng buộc bằng phương pháp hàm phạt (penalty function) [99-102].

Thuật giải 3.1. Thuật giải di truyền

- 1 **Input:** Tham số hệ thống được liệt kê ở Bảng 3-2
 $N_P = 20.000$; số cá thể trong quần thể
 $PRECI = J * I * V_i + K * N * M$: chiều dài của nhiễm sắc thể.
 $P_G = 0,9$: khoảng cách giữa thế hệ
 $P_C = 0,9$: xác suất lai ghép
 $P_m = 10^{-6}$: xác suất đột biến
 Khởi tạo thế hệ đầu tiên gồm tập N_P cá thể ngẫu nhiên $\{X_z\}$, $z = 1, 2, \dots, N_P$
 $Gen = 0$: số thế hệ
 TC : điều kiện dừng vòng lặp
 - 2 **Output:** X^* và \bar{Q}_F^*
 - 3 **While** TC chưa thỏa do
 - 4 $Gen = Gen + 1$
 - 6 Đưa $\{X_z\}$ và $\bar{Q}_F(X_z)$ vào nhóm sẵn sàng lai ghép để xếp hạng.
 - 6 Chọn $N_{PG} \times P_G$ cá thể tốt nhất trong tập $\{X_z\}$ có $\bar{Q}_F(X_z)$ cao để chuẩn bị lai ghép cho ra thế hệ kế tiếp.
 - 7 Chọn các cặp cha mẹ theo phương pháp lai ghép 2 điểm (Double Point Crossover) để lai ghép ra thế hệ con kế tiếp. Việc lai ghép không áp dụng với tất cả các cặp cha mẹ mà chỉ áp dụng với xác suất P_C .
 - 8 Thực hiện đột biến theo phương pháp bù nhị phân (Inversion Mutation) ở thế hệ con được tạo ra với xác suất đột biến là P_M .
 - 9 Đánh giá giá trị thích nghi (có được sau khi đã thay thế các giá trị chuỗi nhị phân của từng cá thể vào hàm thích nghi) cho các thế hệ con được tạo ra sẵn sàng cho các bước lai ghép tiếp theo.
 - 10 **End While**
 - 11 Chọn cá thể X^* tương ứng với giá trị cực đại \bar{Q}_F^* trong thế hệ cuối cùng.
-

Theo đó, các ràng buộc trong (3.32) của bài toán CRS trên sẽ được chuyển đổi như sau:

$$\left\{ \begin{array}{l} \Delta u_{i,j} = 1 - \sum_{v_i=1}^{V_i} u_j^{v_i} \geq 0, i = 1, 2, \dots, I, j = 1, 2, \dots, J, \\ \Delta w_{k,n} = 1 - \sum_{m=1}^M w_{n,m}^k \geq 0, k = 1, 2, \dots, K, n = 1, 2, \dots, N, \\ \Delta w_{k,m} = 1 - \sum_{n=1}^N w_{n,m}^k \geq 0, k = 1, 2, \dots, K, m = 1, 2, \dots, M \\ \Delta w_{n,m} = 1 - \sum_{k=1}^K w_{n,m}^k \geq 0, n = 1, 2, \dots, N, m = 1, 2, \dots, M \\ \Delta L_F^i = \mu L_{max}^i - L_F^i \geq 0, i = 1, 2, \dots, I, \\ \Delta C = \delta C^* - C \geq 0, \\ \Delta I_{C,S}^{k,i,v_i} = \frac{P_0 G_0^k}{\gamma_0} - N_0 - I_{C,S}^{k,v_i} \geq 0, k = 1, 2, \dots, K, i = 1, 2, \dots, I, v_i = 1, 2, \dots, V_i. \end{array} \right. \quad (3.33)$$

Từ đây, ta có hàm phạt bao gồm các ràng buộc ở (3.33)

$$\begin{aligned}
F = & \lambda_1 \sum_{i=1}^I \sum_{j=1}^J (\min\{0, \Delta u_{i,j}\})^2 + \lambda_2 \sum_{k=1}^K \sum_{n=1}^N (\min\{0, \Delta w_{k,n}\})^2 \\
& + \lambda_3 \sum_{k=1}^K \sum_{m=1}^M (\min\{0, \Delta w_{k,m}\})^2 + \lambda_4 \sum_{n=1}^N \sum_{m=1}^M (\min\{0, \Delta w_{n,m}\})^2 \\
& + \lambda_5 \sum_{i=1}^I (\min\{0, \Delta L_F^i\})^2 + \lambda_6 (\min\{0, \Delta C\})^2 \\
& + \lambda_7 \sum_{k=1}^K \sum_{i=1}^I \sum_{v_i=1}^{V_i} (\min\{0, \Delta L_{C,S}^{k,i,v_i}\})^2
\end{aligned} \tag{3.34}$$

Và như vậy, thuật giải di truyền như được mô tả trong Thuật giải 3.1 có thể được áp dụng để giải bài toán tối ưu CRS dưới dạng không có ràng buộc như sau:

$$\max_{u_j^i, w_{n,m}^k} \bar{Q}_F = \bar{Q} - F. \tag{3.35}$$

3.5. Đánh giá hiệu suất cơ chế CRS và thuật giải di truyền

3.5.1. Thiết lập thông số hệ thống và cấu hình máy tính

Để đánh giá hiệu suất cơ chế CRS và hiệu quả thuật giải di truyền, mô hình hệ thống được triển khai với các tham số được liệt kê trong Bảng 3-2. Giả sử MBS bao phủ 1 vùng có bán kính cực đại là 1500m và khoảng cách giữa MBS và MU, FBS và MU, CU và SU, CU và NU được phân phối đều ngẫu nhiên trong các phạm vi lần lượt tương ứng là [500, 1500]m, [20, 100]m, [50, 150]m, [1, 20]m. Ngoài ra, 3 video được xem xét gồm: $V_1^{v_1}$ (Basketballpass), $V_2^{v_2}$ (Racehorses) và $V_3^{v_3}$ (Foreman). Các video này được phân tích để thu được các đường cong RD thực nghiệm của chúng bằng cách sử dụng phần mềm HM phiên bản 12.0 [52], từ đó ta có thể mã hóa mỗi video với 3 mức phân giải khác nhau ($V_i = 3$) cũng có được các giá trị $L_i^{v_i}$, $C_{th}^{v_i}$, γ_i và β_i như trong Bảng 3-2.

Bảng 3-2. Thông số hệ thống

Ký hiệu	Trị số	Ký hiệu	Trị số	
J, K, N, M	5 FBS, 5 SU, 5 CU, 10 NU	r_n	1, nghĩa là CU thứ n còn trống 100% dung lượng để lưu video (nhằm tăng ưu thế cho truyền thông D2D).	
I	3 videos	$L_i^{v_i}$	[11867 23734 35600 198680 264906 351000 33382 66763 113496 1172340 1758510 2344680 160410 320820 453960] Kbit	
V_i	3, mỗi video có 3 mức phân giải		$C_{th}^{v_i}$	[1000 2000 3000 3000 4000 5300 50000 100000 170000 10000 15000 20000 10000 20000 28300]Kbps
W, P_0, P_j, P_C^k	5MHz, 5W, 1W, 1mW			
N_0, η, γ_0	10^{-12} W, 4, 10dB			
$\{a, b\}$	{0.5, 0.5}			
γ_i	[9806 76520 1644000]			
β_i	[-0.9972 -1.1530 -1.0920]			
μ, δ	0.5, 1			
C^*	10 Gbps, nghĩa là mỗi MU được phục vụ với dung lượng trung bình 0.5 Gbps			

Cơ chế CRS được đánh giá bằng cách so sánh với 3 cơ chế khác bao gồm cơ chế chỉ lưu trữ (OCC – Only Caching), cơ chế chỉ chia sẻ tài nguyên phổ tần kênh truyền xuống (ORS – Only Downlink Resource Sharing) và cơ chế không lưu trữ cũng như không chia sẻ tài nguyên phổ tần kênh truyền xuống (NCS – No Caching Nor Downlink Resource Sharing). Một máy tính thông thường với các thông số chính được mô tả trong Bảng 3-3 được sử dụng để triển khai tính toán, mô phỏng và đánh giá kết quả.

Bảng 3-3. Cấu hình máy tính

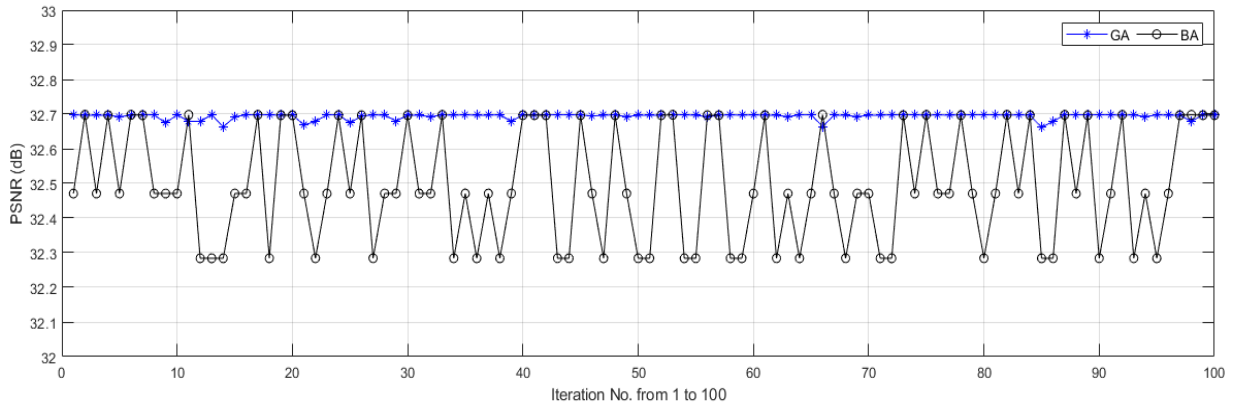
Thông số cấu hình máy tính	
Processor	Intel(R) Core(TM) i7-8700 CPU @ 3.20GHz
Processor type	x64 Family 6 Model 14 Stepping 10, GenuineIntel
PHY processor packages	1
Processor cores	6
Logical processors	12
Total PHY memory	16,599,444KB
Operating System	Windows 10

3.5.2. Đánh giá hiệu quả của thuật giải di truyền

Để đánh giá hiệu quả của thuật giải di truyền (GA), thuật giải đàn dơi nhị phân (BBA) và thuật toán vét cạn (EA) đồng thời được triển khai để làm cơ sở so sánh về độ chính xác và thời gian thực thi cũng như làm cơ sở để chọn lựa các thông số cho thuật giải di truyền khi giải bài toán CRS. BBA được trình bày trong Chương 1, các thông số ($L_i^0 = 0,25$; $E_i^0 = 0,1$; $f_{min} = 0$; $f_{max} = 2$; $\varepsilon = 0,9$; $\zeta = 0,9$) để triển khai BBA được tham khảo từ [61]. Tương tự như trong thuật toán GA, phương pháp hàm phạt cũng được áp dụng để triển khai thuật toán BBA. Do sự không đồng nhất về các thông số đầu vào và các toán tử được sử dụng trong GA và BBA, trong luận án này, GA và BBA được đánh giá dựa trên: 1) sự tương đương về thời gian thực thi, ví dụ: $N_{PG} = 20.000$ trong GA với thời gian thực thi trung bình là 1003,68 giây và $N_{PB} = 10.000$ trong BBA với thời gian thực thi trung bình là 957,12 giây sẽ được chọn để đánh giá GA và BBA, 2) độ chính xác so với EA và 3) độ ổn định (EA được triển khai tương tự như trong Chương 2). Trong trường hợp này, các tham số kích thước hệ thống gồm (J, N, M và K) được thiết lập với những giá trị thấp tương ứng là (3, 3, 5 và 3) để thuật giải vét cạn thực thi được. Bài toán CRS sẽ được giải bằng cách tìm giá trị cực đại của chất lượng trung bình (được đo bằng dB) theo hệ số mũ phản ánh độ nổi tiếng của video α [88].

Bảng 3-4. So sánh độ chính xác giữa GA, BBA và EA

α	0	0,5	1,0	1,5	2,0	2,5	3,0
EA (dB)	29,0286	31,2176	32,6980	34,1178	36,0832	37,2044	38,0584
BBA (dB)	28,9100	31,2176	32,4708	33,6306	36,0832	36,5090	37,6689
Accuracy	99,59%	100,00%	99,31%	98,57%	100,00%	98,13%	98,98%
GA (dB)	29,0286	31,2176	32,6924	34,1178	36,0716	37,1935	38,0476
Accuracy	100,00%	100,00%	99,98%	100,00%	99,97%	99,97%	99,97%



Hình 3-2. Đánh giá độ ổn định của GA và BBA

Kết quả được trình bày trong Bảng 3-4 cho thấy, độ chính xác khi sử dụng GA có thể lên tới 100% so với EA. Với thời gian thực thi tương đương nhau, GA cho kết quả chính xác hơn so với BBA. Hơn nữa, khi thực thi giải thuật 100 lần (với $\alpha = 1$), GA cho kết quả ổn định hơn BBA như được mô tả trong Hình 3-2. Từ những kết quả này, GA được chọn để áp dụng cho việc giải bài toán CRS với kích thước hệ thống lớn hơn.

Ngoài ra, tôi cũng so sánh việc đánh đổi giữa thời gian thực thi và độ chính xác của của thuật giải di truyền bằng cách thay đổi số lượng cá thể N_p trong khi vẫn giữ $\alpha = 1$. Kết quả từ Bảng 3-5 cho thấy, với mỗi số lượng cá thể N_p khác nhau, bài toán được thực thi 100 lần và chọn kết quả xấu nhất trong 100 lần thực thi đó, ta có:

- Khi chấp nhận độ chính xác thấp là 97.34%, thì thời gian thực thi chỉ là 19.58 giây với số lượng cá thể $N_p = 1.000$.
- Khi tăng số lượng cá thể lên, thì độ chính xác cũng tăng lên và thời gian thực thi cũng lâu hơn.

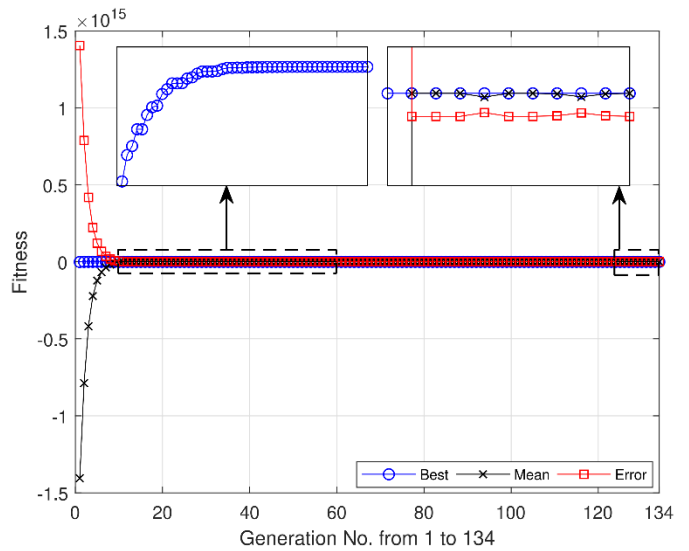
Tuy nhiên, khi tăng số lượng cá thể N_p lên thì độ chính xác sẽ tăng lên rất ít nhưng thời gian thực thi tăng lên rất nhiều. Ở Bảng 3-5 thì khi so sánh giữa $N_p = 1.000$ và $N_p = 10.000$, thì độ chính xác từ 97.34% lên 99.23% (tăng 1.89%), nhưng khi tăng tiếp số lượng cá thể từ $N_p = 10.000$ lên $N_p = 20.000$, thì độ chính xác chỉ tăng từ 99.23% lên 99.89% (tăng chỉ 0.66%). Như vậy, ở $N_p = 10.000$ thì thuật giải di truyền cho kết quả chấp nhận được

với mức độ chính xác là 99.23% và thời gian thực thi là 317.42 giây so với 5,664.45 giây của thuật giải vét cạn. Dựa trên kết quả từ Bảng 3-5, tôi chọn $N_P = 20.000$ để thực thi thuật giải di truyền với các tham số kích thước hệ thống (J, N, M và K) lớn hơn như trong Bảng 3-2 để đảm bảo độ chính xác tốt nhất với thời gian thực thi chấp nhận được.

Bảng 3-5. Thực thi GA theo số N_P với kết quả xấu nhất

Metrics	$N_P (\alpha = 1)$					EA
	1,000	5,000	10,000	15,000	20,000	
Time (s)	19.58	142.50	317.42	556.01	795.63	5,664.45
PSNR (dB)	30.8283	32.2547	32.4474	32.6486	32.6631	32.6980
Accuracy	97.34%	98.64%	99.23%	99.85%	99.89%	100%

Để đánh giá kết quả thực hiện thuật giải di truyền, thuật giải di truyền được thực thi với các tiêu chí hội tụ (hay tiêu chí để dừng thuật giải) được trình bày trong Thuật giải 3-1. Kết quả thực thi theo số thế hệ được mô tả trong Hình 3-3, trong đó, “Best” là giá trị $\bar{Q}_F(X_Z)$ cao nhất có được từ cá thể tốt nhất trong từng thế hệ, “Mean” là giá trị trung bình của $\bar{Q}_F(X_Z)$ được tính cho tất cả các cá thể trong từng thế hệ, và “Error” chính là giá trị của hàm phạt F trong (3.34) và (3.35). Kết quả cho thấy, thuật giải di truyền bắt đầu hội tụ sau khoảng 20 thế hệ và đến thế hệ thứ 125 thì đáp ứng được tiêu chí hội tụ. Khi đó, “Error” tiến về 0 ($F = 0$) đồng nghĩa với các ràng buộc đã thỏa; giá trị “Best” không đổi và trùng với “Mean” đồng nghĩa với việc các cá thể đều tốt như nhau ở những thế hệ cuối cùng. Như vậy, từ các kết quả được phân tích trong Bảng 3-4, Bảng 3-5, Hình 3-2 và Hình 3-3, thuật giải di truyền cho thấy tính linh động và khả thi để giải bài toán tối ưu CRS trong phạm vi lớn hơn của 5G UDN.

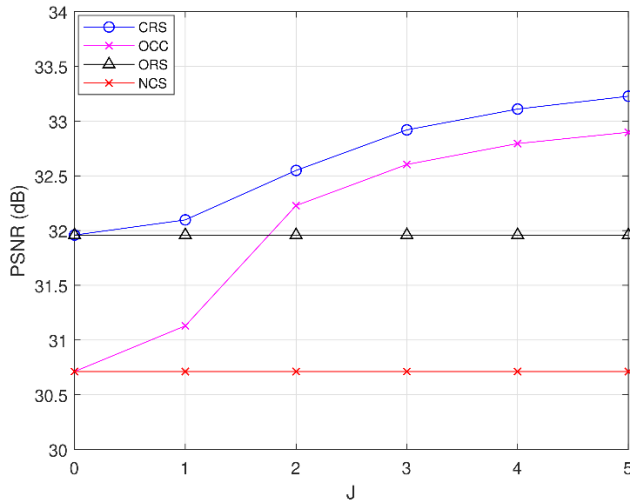


Hình 3-3. Mức độ hội tụ của thuật giải di truyền

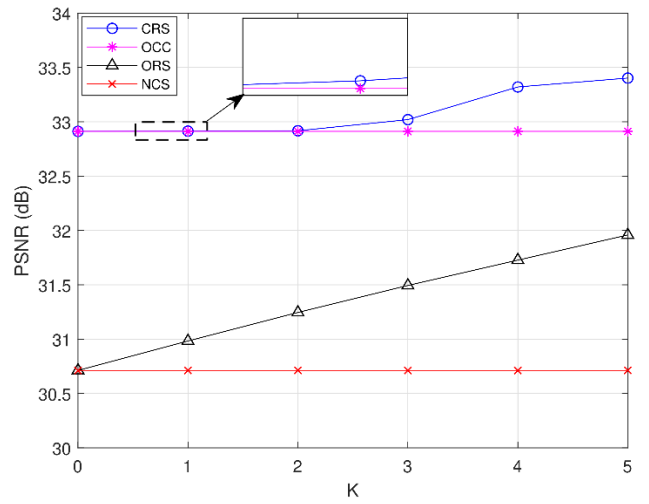
3.5.3. Đánh giá hiệu suất cơ chế CRS

Đầu tiên, CRS được đánh giá so với 3 trường hợp khác là: chỉ có lưu trữ (OCC), chỉ có chia sẻ phổ tần kênh truyền xuống (ORS) và không có lưu trữ cũng không có chia sẻ (NCS) khi thay đổi số lượng FBS Hình 3-4. Nếu không có FBS để lưu trữ video ($J=0$), thì cơ chế OCC trở thành cơ chế NCS và cả 2 đều cho chất lượng thấp nhất vì không có lưu trữ cũng không có chia sẻ. Đồng thời, cơ chế CRS không có lưu trữ thì trở thành cơ chế ORS, nhưng vẫn cho chất lượng cao hơn so với OCC và NCS vì CRS có được sự chia sẻ tài nguyên phổ tần kênh truyền xuống. Khi tăng số lượng FBS, thì chất lượng của ORS và NCS không tăng lên, trong khi CRS và OCC tăng lên theo số lượng của FBS. Tuy nhiên, CRS và OCC sẽ có khuynh hướng bão hòa khi tiếp tục tăng số lượng FBS. Điều này cho thấy rằng, với một lượng người dùng nhất định, ta không cần thiết phải lưu tất cả các video trong nhiều FBS. Nói một cách khác, số lượng FBS để lưu trữ các video phải được tính toán cẩn thận để đảm bảo dung lượng lưu trữ hợp lý cũng như để giảm mức độ phức tạp trong tính toán.

Tiếp theo, cơ chế CRS được so sánh với các cơ chế OCC, ORS và NCS theo số lượng SU đồng ý chia sẻ tài nguyên (Hình 3-5). Tương tự như với kết quả đánh giá theo số lượng FBS, nếu không có SU nào chia sẻ tài nguyên kênh truyền, thì cơ chế CRS trở thành OCC và cơ chế ORS trở thành NCS. Khi tăng số lượng SU lên đồng nghĩa với việc mở rộng chọn lựa chia sẻ tài nguyên kênh truyền thì chất lượng của CRS và ORS cũng tăng lên theo, trong khi OCC và NCS không thay đổi. Hai cơ chế CRS và ORS sẽ bão hòa khi tiếp tục tăng số lượng SU vì khi đó các SU được thêm vào không thể tốt hơn các SU trước đó trong quá trình tìm kiếm bộ 3 (SU, CU, NU) tốt nhất.

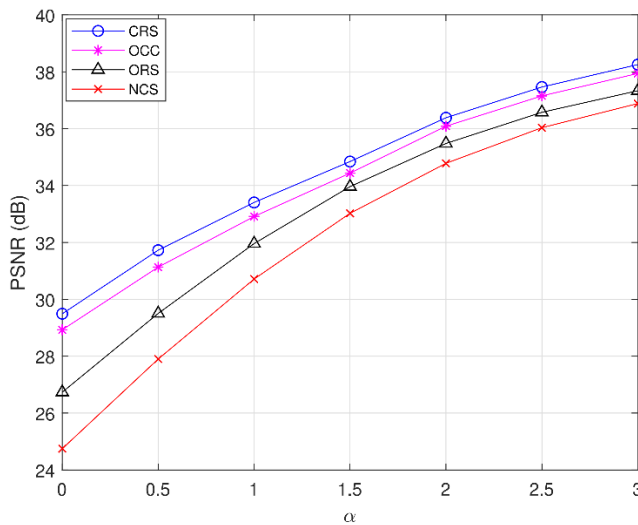


Hình 3-4. Kết quả CRS so với OCC, ORS và NCS theo số lượng FBS

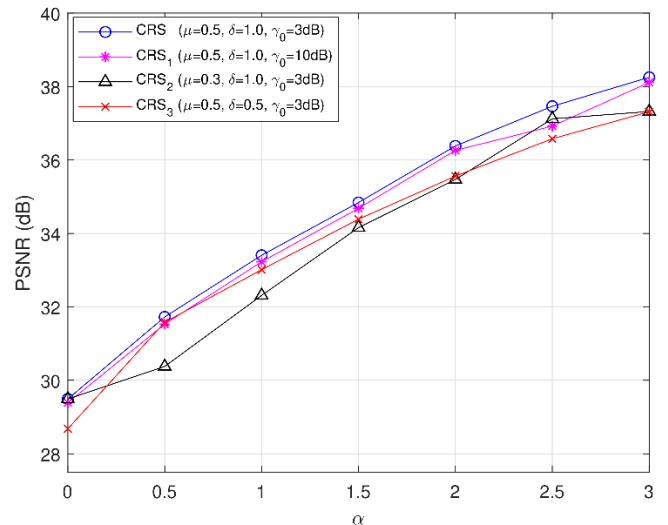


Hình 3-5. Kết quả CRS so với OCC, ORS và NCS theo số lượng SU

Hình 3-6 mô tả kết quả đánh giá hiệu suất giữa CRS, OCC, ORS và NCS theo hệ số độ lệch nổi tiếng của video α . Kết quả chỉ ra rằng hệ thống đạt hiệu suất cao hơn nếu nó tập trung vào việc phục vụ MU các video có mức độ phổ biến cao, tức là giá trị của α cao hơn. Và kết quả đánh giá cũng cho thấy rằng, cơ chế đề xuất CRS vượt trội hơn so với cơ chế OCC, ORS và đương nhiên tốt hơn nhiều so với cơ chế NCS. Và trường hợp OCC tốt hơn trường hợp ORS vì nó cung cấp nhiều khả năng lưu trữ và truyền qua các kênh tốt hơn từ FBS. Trường hợp xấu nhất là NCS vì nó không có sự hỗ trợ của cơ chế lưu trữ và chia sẻ tài nguyên phổ tần kênh truyền xuống.



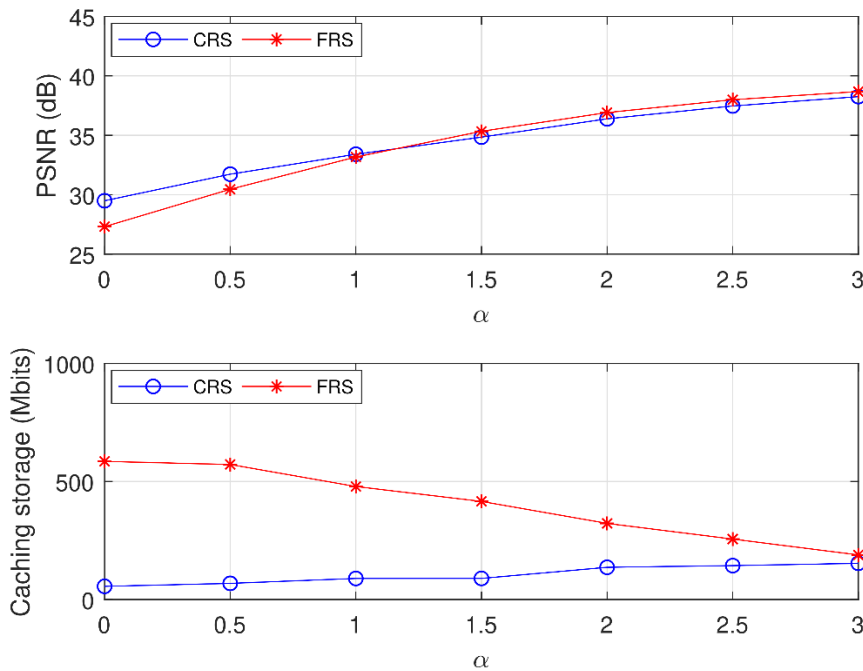
Hình 3-6. Chất lượng hệ thống của CRS, OCC, ORS, và NCS theo α



Hình 3-7. Chất lượng của hệ thống giữa CRS với các tham số hệ thống khác nhau theo α

Tiếp theo, ảnh hưởng của các ràng buộc đến hiệu suất của cơ chế CRS được đánh giá trong Hình 3-7. Để làm như vậy, dung lượng yêu cầu bởi MU (δ) được giảm từ 1 xuống 0.5, ràng buộc về dung lượng lưu trữ của FBS (μ) được giảm từ 0.5 xuống 0.3 và tăng tỉ số tín hiệu trên can nhiễu và nhiễu trắng của SU (γ_0) từ 3dB lên 5dB. Kết quả cho thấy, nếu ràng buộc dung lượng δ được yêu cầu bởi MU giảm (từ 1 xuống 0.5), tức là MU chỉ yêu cầu các video có mức phân giải thấp, các kênh truyền cung cấp dung lượng cao sẽ không được chọn để truyền phát nhằm tiết kiệm tài nguyên băng thông của hệ thống. Điều đó có nghĩa là hệ thống chỉ cần phục vụ ở mức chất lượng thấp vì đã đáp ứng đủ yêu cầu của MU. Trong trường hợp giảm ràng buộc về dung lượng lưu trữ của FBS μ (từ 0.5 xuống 0.3), hệ thống có ít khả năng để lưu trữ video và do đó sẽ hạn chế cung cấp kênh truyền chất lượng cao từ FBS đến MU dẫn đến chất lượng video được thu cũng giảm. Ngoài ra, khi tăng tỉ số tín hiệu trên can nhiễu và nhiễu trắng γ_0 (từ 3dB lên 5dB) để đảm bảo QoS cao hơn cho SU, thì số lượng cặp truyền thông D2D giữa CU và NU sẽ giảm để giảm tác động can nhiễu lên SU. Điều này cũng làm giảm hiệu suất

hệ thống vì hệ thống không thể khai thác hiệu quả của việc truyền thông D2D để truyền phát video trong khoảng cách gần.



Hình 3-8. So sánh CRS và FRS

Điều này có thể giải thích là, khi α có giá trị thấp, hệ thống sẽ phục vụ MU với nhiều video có độ phổ biến không lệch nhau quá cao. Trong trường hợp này, việc chọn lựa video có mức phân giải phù hợp để lưu trữ và phục vụ người dùng sẽ làm tăng hiệu quả hệ thống (thỏa mãn người dùng hơn với dung lượng lưu trữ thấp hơn). Ngược lại, khi α có giá trị cao hơn, hệ thống tập trung vào việc phục vụ các video có độ phổ biến cao, nghĩa là sẽ tập trung phục vụ một số ít video hơn. Trong trường hợp này, trong khi hiệu suất của CRS bị hạn chế do bị ràng buộc về dung lượng bộ nhớ lưu trữ và các ràng buộc thông lượng cần thiết, thì các ràng buộc này hoàn toàn không được xét đến trong FRS. Khi xem xét mức tiêu thụ dung lượng bộ nhớ đệm, FRS rõ ràng yêu cầu sử dụng bộ nhớ lưu trữ cao hơn CRS. Và có thể thấy rằng, dung lượng cần cho bộ nhớ đệm của cả CRS và FRS để lưu trữ video sẽ hội tụ về một giá trị nếu phiên bản video trong mô hình CRS và phiên video trong FRS là như nhau.

3.6. Kết luận chương 3

Chương 3 đã trình bày cơ chế lưu trữ đa phân giải và chia sẻ tài nguyên tối ưu (CRS – Multi-Resolution Caching and Resource Sharing Optimization) theo nhu cầu người dùng nhằm nâng cao chất lượng truyền video trong 5G UDN. Trong chương này, các hạn chế của Chương 2 đã được cải tiến như: 1) QoS hệ thống được đánh giá tường minh hơn, 2) sử dụng hiệu quả tài nguyên dung lượng lưu trữ tại FBS và tài nguyên phổ tần của SU hơn và 3) thuật giải di truyền được áp dụng để giải bài toán CRS một cách hiệu quả.

KẾT LUẬN

1. Kết quả đạt được

Luận án tập trung vào việc nghiên cứu và đề xuất cơ chế lưu trữ và chia sẻ tài nguyên tối ưu cho các ứng dụng truyền video trong 5G UDN nhằm nâng cao chất lượng lượng dịch vụ (QoS) và sử dụng tài nguyên hiệu quả. Kết quả đạt được cụ thể như sau:

- 1- Đề xuất cơ chế lưu trữ và truyền video cộng tác đa tầng có quan tâm đến hành vi và mối quan hệ xã hội của người dùng (SCS) thông qua việc tận dụng hiệu quả tài nguyên lưu trữ và tài nguyên phổ tần sẵn có của các phần tử trong 5G UDN. Kết quả này được công bố trên tạp chí Springer Mobile Networks & Applications

[C1] và được trình bày tại hội nghị Heterogeneous Networking for Quality, Reliability, Security and Robustness (Qshine 2018).

- 2- Đề xuất cơ chế lưu trữ đa phân giải và truyền video cộng tác có quan tâm đến nhu cầu người dùng (CRS) thông qua việc tận dụng hiệu quả tài nguyên lưu trữ và tài nguyên phổ tần sẵn có của các phần tử trong 5G UDN. Đặc biệt, CRS cải tiến cơ chế SCS bằng cách 1) thay đổi thông số đánh giá QoS từ dung lượng phân phối thành chất lượng video thu được, 2) tận dụng hiệu quả hơn tài nguyên lưu trữ và phổ tần và 3) đề xuất thuật giải di truyền để giải bài toán tối ưu lưu trữ và truyền video một cách hiệu quả. Kết quả này được công bố trên tạp chí IEEE Communications Letters [C2] và được trình bày tại hội nghị Recent Advances on Signal Processing, Telecommunications & Computing (SigTelCom 2020) [C3].

Các công bố liên quan đến luận án là những kết quả nghiên cứu quan trọng, cụ thể là đề xuất mô hình lưu trữ đa tầng kết hợp với chia sẻ tài nguyên tối ưu theo định hướng người dùng nhằm nâng cao hiệu quả truyền video trong 5G UDN. Đóng góp này mang giá trị tham khảo hữu ích cho cộng đồng nghiên cứu thuộc lĩnh vực lưu trữ và chia sẻ tài nguyên cho các ứng dụng truyền video trong 5G UDN.

2. Định hướng nghiên cứu

Nội dung nghiên cứu trong Chương 2 và Chương 3 có sự nối tiếp nhau. Nghiên cứu trong Chương 3 có kế thừa và bổ sung thêm cho các nghiên cứu đã thực hiện trong Chương 2. Tuy nhiên, luận án vẫn còn những hạn chế cần được phát triển trong tương lai liên quan đến mô hình, tiêu chí đánh giá và thuật giải, cụ thể như sau:

- 1- Cải tiến mô hình hệ thống: Một hệ thống hoàn thiện hơn sẽ là hệ thống được kết hợp bởi kết quả của Chương 2 và Chương 3. Nghĩa là phải 1) xem xét đến các video với các mức phân giải khác nhau để lưu trữ hiệu quả hơn tại FBS, 2) tận dụng tài nguyên phổ tần hiệu quả hơn cho truyền thông D2D và 3) xem xét đồng thời đến cư xử, mối quan hệ xã hội và nhu cầu người dùng cũng như yếu tố di chuyển của người dùng. Ngoài ra, một tầng lưu trữ nữa cần được xem xét đến đó là lưu trữ tại các thiết bị bay không người lái (UAV – Unmanned Aerial Vehicle).
- 2- Về tiêu chí đánh giá: Bộ thông số đánh giá chất lượng trải nghiệm người dùng (QoE – Quality of Experience) với những tiêu chí khắt khe hơn cần được xây dựng. Bộ tiêu chí này không đơn thuần là dung lượng phân phối hay chất lượng phát lại video, mà còn là tỷ lệ truy xuất thành công video, tính liên tục cũng như sự ổn định chất lượng trong quá trình phát lại.
- 3- Về thuật giải tối ưu: Với tính phức tạp về hệ thống được mở rộng hoàn thiện hơn và bộ thông số QoE đánh giá hiệu suất hệ thống phức tạp hơn, thuật giải GA cần được nghiên cứu để có những cải tiến hơn về độ chính xác và thời gian thực thi. Ngoài ra, cũng cần tìm hiểu thêm các thuật giải phù hợp khác theo xu thế hiện nay như máy học, học sâu,... để giải quyết bài toán tối ưu và so sánh hiệu quả với thuật giải GA đã áp dụng nhằm chọn lựa thuật giải tốt nhất để triển khai hệ thống.

DANH MỤC CÁC CÔNG TRÌNH ĐÃ CÔNG BỐ CỦA NGHIÊN CỨU SINH

- [C1] **Minh-Phung Bui**, Nguyen-Son Vo, Sang Quang Nguyen, and Quang-Nhat Tran, "Social-Aware Caching and Resource Sharing Maximized Video Delivery Capacity in 5G Ultra-Dense Networks," *ACM/Springer Mobile Networks & Applications*, pp. 1-13, July 2019;
- [C2] Nguyen-Son Vo, **Minh-Phung Bui**, Phuc Quang Truong, Cheng Yin, and Antonino Masaracchia, "Multi-tier Caching and Resource Sharing for Video Streaming in 5G Ultra-dense Networks," *IEEE Communications Letters*, vol. 24, no. 7, pp. 1500-1504, July 2020;
- [C3] **Minh-Phung Bui**, Nguyen-Son Vo, Tien-Vu Truong, Thanh-Hieu Nguyen, Nam Van Nguyen and Cheng Yin, "Genetic Algorithms for Multi-tier Caching and Resource Sharing Optimized Video Streaming in 5G Ultra-dense Networks," in *Proc. of International Conference on Recent Advances on Signal Processing, Telecommunications & Computing (SigTelCom'20)*, Ha Noi, Vietnam, Aug. 2020, pp. 66-71;