

**ĐẠI HỌC QUỐC GIA HÀ NỘI**  
**VIỆN CÔNG NGHỆ THÔNG TIN**

**LÊ MINH TUẤN**

**NGHIÊN CỨU ĐỀ XUẤT CẢI TIẾN**  
**THUẬT TOÁN LẬP LỊCH VÀ ỨNG DỤNG**

**LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN**

Hà Nội - 2022

ĐẠI HỌC QUỐC GIA HÀ NỘI  
VIỆN CÔNG NGHỆ THÔNG TIN

LÊ MINH TUẤN

NGHIÊN CỨU ĐỀ XUẤT CẢI TIẾN  
THUẬT TOÁN LẬP LỊCH VÀ ỨNG DỤNG

Chuyên ngành: Quản lý Hệ thống Thông tin

Mã số: 9480205.01 QTD

LUẬN ÁN TIẾN SĨ CÔNG NGHỆ THÔNG TIN

NGƯỜI HƯỚNG DẪN KHOA HỌC

1. PGS.TS. Lê Hoàng Sơn
2. TS. Vũ Như Lâm

## LỜI CAM ĐOAN

Tôi xin cam đoan luận án “**Nghiên cứu đề xuất cải tiến thuật toán lập lịch và ứng dụng**” là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể cán bộ hướng dẫn. Luận án có sử dụng thông tin trích dẫn từ nhiều nguồn tham khảo khác nhau và các thông tin trích dẫn được ghi rõ nguồn gốc. Các kết quả nghiên cứu của tôi được viết chung với các tác giả khác đã được sự nhất trí của đồng tác giả khi đưa vào luận án. Các số liệu, kết quả được trình bày trong luận án là hoàn toàn trung thực và chưa từng được công bố trong bất kỳ một công trình nào khác.

Luận án được hoàn thành trong thời gian tôi làm Nghiên cứu sinh tại Viện Công nghệ Thông tin, Đại học Quốc gia Hà Nội.

Tác giả:

---

Hà Nội:

---

# LỜI CẢM ƠN

Lời đầu tiên, em xin bày tỏ lòng biết ơn chân thành và sâu sắc tới tập thể hướng dẫn khoa học PGS.TS Lê Hoàng Sơn và TS. Vũ Như Lâm, những người đã định hướng khoa học, tận tâm giúp đỡ và chỉ bảo trong suốt quá trình em hoàn thành luận án này tại viện Công nghệ Thông tin - Đại học Quốc gia Hà Nội. Một vinh dự rất lớn cho em có cơ hội được học tập, nghiên cứu dưới sự hướng dẫn tận tâm của các Thầy.

Xin chân thành cảm ơn các nhà khoa học, tác giả các công trình nghiên cứu được trích dẫn, tham khảo trong luận án này, đây là những kiến thức cơ sở để tôi phát triển và hoàn thiện các công bố của mình.

Xin chân thành cảm ơn Đảng ủy, Ban giám hiệu trường Đại học Nội vụ Hà Nội, tập thể giảng viên, chuyên viên nơi tôi công tác đã luôn động viên, khuyến khích và tạo điều kiện thuận lợi nhất trong suốt quá trình tôi thực hiện luận án.

Xin chân thành cảm ơn Ban Lãnh đạo Viện Công nghệ Thông tin, các phòng chức năng, các Giảng viên, các đồng nghiệp làm việc trong Lab nghiên cứu tại Phòng Công nghệ đa phương tiện và thực tại ảo, Viện Công nghệ Thông tin - Đại học Quốc gia Hà Nội đã luôn quan tâm, giúp đỡ và tạo điều kiện về nhiều mặt, chỉ bảo tận tình trong suốt quá trình tôi thực hiện luận án.

Xin bày tỏ lòng biết ơn vô hạn đối với cha mẹ, vợ con, anh chị em và gia đình, những người đã kiên trì chia sẻ, động viên cả về vật chất lẫn tinh thần, ủng hộ và yêu thương vô điều kiện.

Cuối cùng, xin kính chúc các Thầy, Cô, các bạn đồng nghiệp, anh chị em, bạn bè luôn mạnh khỏe, đạt được nhiều thành tựu trong công tác, học tập và nghiên cứu khoa học!

NCS. Lê Minh Tuấn

# Mục lục

|  |           |
|--|-----------|
| Lời cam đoan   | ii        |
| Lời cảm ơn   | iii       |
| MỤC LỤC  | iv        |
| Danh mục các từ viết tắt   | vii       |
| Danh mục các bảng  | xii       |
| Danh mục các hình vẽ   | xiii      |
| <br>   |           |
| <b>MỞ ĐẦU</b>  | <b>1</b>  |
| Tính cấp thiết của luận án . . . . .                                     | 1         |
| Một số vấn đề tồn tại trong các nghiên cứu liên quan . . . . .           | 8         |
| Mục tiêu và đối tượng nghiên cứu . . . . .                               | 9         |
| Nội dung nghiên cứu . . . . .  | 9         |
| Phương pháp nghiên cứu . . . . .   | 10        |
| Đóng góp chính và tính mới của luận án . . . . .                         | 10        |
| Phạm vi và giới hạn của đề tài nghiên cứu . . . . .                      | 11        |
| Môi trường mô phỏng và công cụ đánh giá . . . . .                        | 12        |
| Cấu trúc luận án . . . . .   | 12        |
| <br>   |           |
| <b>Chương 1. TỔNG QUAN NGHIÊN CỨU VÀ CƠ SỞ LÝ THUYẾT</b>                 | <b>14</b> |
| 1.1 Sơ lược về hệ thống tính toán phân tán và bối cảnh bài toán lập lịch | 14        |
| 1.2 Tổng quan về lập lịch tài nguyên trong mạng di động đa dịch vụ . .   | 18        |
| 1.2.1 Các dịch vụ của mạng LTE . . . . .                                 | 19        |
| 1.2.2 Bài toán lập lịch trong mạng LTE . . . . .                         | 19        |
| 1.2.3 Các nghiên cứu liên quan về lập lịch trong mạng LTE . . . . .      | 24        |
| 1.2.4 Một số nhận xét . . . . .  | 34        |
| 1.3 Tổng quan về lập lịch trong lưới tính toán di động . . . . .         | 35        |

|   |   |           |
|---|---|-----------|
| 1.3.1   | Sơ lược về điện toán lưới và lưới tính toán di động . . . . .   | 36        |
| 1.3.2   | Bài toán lập lịch trong lưới tính toán di động . . . . .  | 38        |
| 1.3.3   | Các mục tiêu lập lịch trong lưới tính toán di động . . . . .  | 40        |
| 1.3.4   | Các nghiên cứu liên quan về lập lịch trong lưới tính toán di động   | 44        |
| 1.3.5   | Một số nhận xét . . . . .   | 49        |
| 1.4   | Kết luận chương . . . . .   | 50        |
| <b>Chương 2. ĐỀ XUẤT GIẢI PHÁP LẬP LỊCH TRONG MẠNG<br/>DI ĐỘNG ĐA DỊCH VỤ</b> |   | <b>52</b> |
| 2.1   | Giới thiệu . . . . .  | 54        |
| 2.2   | Mô hình mạng và các giả thiết . . . . .   | 58        |
| 2.3   | Mô hình hóa toán học . . . . .  | 59        |
| 2.4   | Thuật toán lập lịch kết hợp giữa miền thời gian và miền tần số cho<br>đường tải xuống - ITFDS . . . . .         | 62        |
| 2.5   | Thực nghiệm đánh giá kết quả . . . . .  | 70        |
| 2.5.1   | Thực nghiệm tỷ lệ trễ gói với lượng người dùng khác nhau . .  | 71        |
| 2.5.2   | Thực nghiệm tỷ lệ mất gói với lượng người dùng khác nhau .  | 72        |
| 2.5.3   | Thực nghiệm thông lượng với lượng người dùng khác nhau . .  | 73        |
| 2.5.4   | Thực nghiệm chỉ số công bằng với lượng người dùng khác nhau   | 74        |
| 2.6   | Phân tích độ phức tạp thuật toán . . . . .  | 75        |
| 2.7   | Ưu và nhược điểm của phương pháp đề xuất . . . . .  | 75        |
| 2.8   | Kết luận chương . . . . .   | 76        |
| <b>Chương 3. ĐỀ XUẤT GIẢI PHÁP LẬP LỊCH CHO LƯỚI TÍNH<br/>TOÁN DI ĐỘNG</b>    |   | <b>78</b> |
| 3.1   | Giới thiệu . . . . .  | 79        |
| 3.2   | Mô hình mạng và các giả thiết . . . . .   | 84        |
| 3.3   | Mô hình hóa . . . . .   | 85        |
| 3.4   | Thuật toán lập lịch công việc trong lưới tính toán di động dựa trên<br>tìm kiếm meta-heuristic - HGLA . . . . . | 91        |
| 3.5   | Ví dụ minh họa . . . . .  | 95        |
| 3.6   | Thực nghiệm đánh giá kết quả . . . . .  | 98        |
| 3.6.1   | Thử nghiệm tỉ lệ địa phương hóa trong trường hợp lưới có nút<br>di động và không có nút di động . . . . .       | 99        |
| 3.6.2   | Thử nghiệm chỉ lệ tăng tốc với trường hợp lưới có số lượng nút<br>thay đổi . . . . .                            | 101       |

|  |            |
|--|------------|
| 3.6.3 Thử nghiệm thông lượng với trường hợp lưới có số lượng nút<br>thay đổi . . . . . | 102        |
| 3.7 Ưu và nhược điểm của thuật toán đề xuất . . . . .                                  | 103        |
| 3.8 Kết luận chương . . . . .  | 104        |
| <b>KẾT LUẬN</b>  | <b>106</b> |
| <b>DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ</b>                                     | <b>110</b> |
| <b>TÀI LIỆU THAM KHẢO</b>  | <b>110</b> |

## DANH MỤC CÁC TỪ VIẾT TẮT

| STT | Từ viết tắt | Từ tiếng Anh                   | Diễn giải/Tạm dịch                   |
|-----|-------------|--------------------------------|--------------------------------------|
| 1   | 3G          | Third Generation               | Thế hệ thứ ba                        |
| 2   | 3GPP        | 3G Partnership Program         | Dự án đối tác thế hệ thứ ba          |
| 3   | ABC         | Artificial Bee Colony          | Thuật toán bầy ong nhân tạo          |
| 4   | ACO         | Ant Colony Optimization        | Thuật toán tối ưu đàn kiến           |
| 5   | AI          | Artificial Intelligence        | Trí tuệ nhân tạo                     |
| 6   | AMC         | Adaptive Modulation and Coding | Bộ điều chế mã hóa thích ứng         |
| 7   | BET         | Blind Equal Throughput         | Thuật toán lập lịch BET              |
| 8   | CG          | Computational Grid             | Lưới tính toán                       |
| 9   | CPS         | Cyber Physical System          | Hệ thống vật lý ảo                   |
| 10  | CQI         | Channel Quality Indicator      | Bộ chỉ thị chất lượng kênh truyền    |
| 11  | DAG         | Directed Acyclic Graph         | Đồ thị có hướng không có chu trình   |
| 12  | DCI         | Downlink Control Information   | Thông tin điều khiển đường xuống     |
| 13  | DRX         | Discontinuous Reception        | Phương pháp tiếp nhận không liên tục |

*Danh mục các từ viết tắt*

---

|    |         |  |  |
|----|---------|--|--|
| 14 | EDF     | Earliest Deadline First                            | Thuật toán lập lịch EDF                                |
| 15 | EPC     | Evolved Package Core                               | Mạng lõi chuyển mạch gói                               |
| 16 | E-UTRAN | Evolved-Universal Terrestrial Radio Access Network | Mạng truy cập vô tuyến phổ quát tiến hóa               |
| 17 | EXP/PF  | Exponential/PF                                     | Thuật toán lập lịch EXP/PF                             |
| 18 | FDPS    | Frequency Domain Package Scheduler                 | Bộ lập lịch gói theo miền tần số                       |
| 19 | FIFO    | First In First Out                                 | Hàng đợi FIFO  |
| 20 | GA      | Genetic Algorithm                                  | Giải thuật di truyền                                   |
| 21 | GBR     | Guaranteed Bit Rate                                | Đảm bảo tốc độ bit                                     |
| 22 | GC      | Grid Computing                                     | Điện toán lưới   |
| 23 | GPF     | Generalized Proportional Fair                      | Thuật toán lập lịch cân bằng tổng quát                 |
| 24 | GPS     | Global Positioning System                          | Hệ thống định vị toàn cầu                              |
| 25 | HARQ    | Hybrid Automatic Retransmission Request            | Yêu cầu tự động truyền lại kết hợp                     |
| 26 | HEFT    | Heterogeneous Earliest Finish Time                 | Thuật toán thời gian kết thúc sớm nhất không đồng nhất |
| 27 | HPC     | High Performance Computing                         | Tính toán hiệu năng cao                                |
| 28 | HSS     | Home Subscriber Server                             | Máy chủ thuê bao thường trú                            |

*Danh mục các từ viết tắt*

|    |         |   |  |
|----|---------|---|--|
| 29 | IMS     | IP Multimedia Subsystem                                 | Hệ thống mạng con đa phương tiện trên giao thức IP                     |
| 30 | IP      | Internet Protocol                                       | Giao thức IP   |
| 31 | ITFDS   | Integrated Time and Frequency-based Downlink Scheduling | Lập lịch đường xuống dựa theo miền thời gian và miền tần số            |
| 32 | LA      | Lion Algorithm  | Giải thuật tối ưu dựa trên ý tưởng săn mồi của sư tử                   |
| 33 | LTE     | Long Term Evolution                                     | Tiến hóa dài hạn   |
| 34 | LWDF    | Largest Weighted Delay First                            | Thuật toán LWDF  |
| 35 | M2M     | Machine 2 Machine                                       | Giao tiếp trực tiếp giữa các thiết bị sử dụng bất kỳ kênh liên lạc nào |
| 36 | MA      | Memetic Algorithm                                       | Thuật toán Memetic   |
| 37 | MAX-MIN | MAX-MIN   | Thuật toán MAX-MIN   |
| 38 | MANET   | Mobile-Adhoc Network                                    | Mạng di động tùy biến  |
| 39 | MCG     | Mobile Computational Grid                               | Lưới tính toán di động   |
| 40 | MCPS    | Medical Cyber Physical System                           | Hệ thống vật lý ảo trong y tế  |
| 41 | MIN-MIN | MIN-MIN   | Thuật toán MIN-MIN   |
| 42 | M-LWDF  | Modified Largest Weighted Delay First                   | Thuật toán M-LWDF  |
| 43 | MME     | Mobility Management Entity                              | Thực thể quản lý tính di động  |

*Danh mục các từ viết tắt*

|    |         |   |  |
|----|---------|---|--|
| 44 | MT      | Maximum Throughput                            | Thuật toán thông lượng tối đa  |
| 45 | NFC     | Near Field Communication                      | Giao tiếp tầm gần  |
| 46 | Non-GBR | Non-Guaranteed Bit Rate                       | Không đảm bảo tốc độ bit   |
| 47 | OFDM    | Orthogonal Frequency-Division Multiplexing    | Ghép kênh phân chia theo tần số trực giao  |
| 48 | OFDMA   | Orthogonal Frequency-Division Multiple Access | Đa truy cập phân chia theo tần số trực giao  |
| 49 | OTT     | Over The Top                                  | Thuật ngữ chỉ các ứng dụng và các nội dung như âm thanh, video, v.v được cung cấp trên nền tảng Internet |
| 50 | PCRF    | Policy and Charging Rules Function            | Chức năng chính sách và tính phí   |
| 51 | PDA     | Personal Digital Assistant                    | Thiết bị điện toán hỗ trợ cá nhân  |
| 52 | PDCCH   | Physical Downlink Control Channel             | Kênh điều khiển đường xuống vật lý   |
| 53 | PF      | Proportional Fair                             | Thuật toán cân bằng  |
| 54 | P-GW    | PDN-Gateway                                   | Cổng mạng dữ liệu gói  |
| 55 | PLR     | Package Loss Rate                             | Tỷ lệ mất gói  |
| 56 | PSO     | Particle Swarm Optimization                   | Tối ưu bầy đàn   |

*Danh mục các từ viết tắt*

|    |      |   |   |
|----|------|---|---|
| 57 | PSS  | Priority Set Scheduling Algorithm       | Thuật toán lập lịch ưu tiên   |
| 58 | QAM  | Quadrature Amplitude Modulation         | Điều chế biên độ cầu phương   |
| 59 | QCI  | QoS class Identifier                    | Tham số xác định cách đối xử của các nút mạng với những gói tin IP nhận được trên mỗi kênh mang |
| 60 | QuAS | QoS-aware Downlink Scheduling Algorithm | Thuật toán lập lịch QuAS  |
| 61 | RB   | Resource Block                          | Khối tài nguyên   |
| 62 | RFID | Radio Frequency Identification          | Nhận dạng qua tần số vô tuyến   |
| 63 | RR   | Round Robin                             | Thuật toán lập lịch xoay vòng   |
| 64 | RRM  | Radio Resource Management               | Bộ quản lý tài nguyên vô tuyến  |
| 65 | S-GW | Serving Gateway                         | Cổng phục vụ  |
| 66 | SLA  | Service-Level Agreement                 | Thỏa thuận cấp độ dịch vụ   |
| 67 | TDPS | Time Domain Package Scheduler           | Bộ lập lịch gói theo miền thời gian   |
| 68 | TTI  | Time Transmission Interval              | Khoảng thời gian truyền   |
| 69 | VoIP | Voice over Internet Protocol            | Công nghệ truyền tiếng nói của con người qua mạng máy tính sử dụng bộ giao thức TCP/IP          |
| 70 | UE   | User Equipment                          | Thiết bị người dùng   |

# DANH MỤC CÁC BẢNG

|   |    |
|---|----|
| 1.1 Các ký hiệu sử dụng cho mô hình lập lịch trong mạng di động đa dịch vụ và mô tả . . . . . | 24 |
| 1.2 Các ký hiệu sử dụng cho mô hình lập lịch trong lưới tính toán di động và mô tả . . . . .  | 39 |
| 2.1 Bảng các giá trị QCI chuẩn [98] . . . . .   | 63 |
| 2.2 Bảng các giá trị CQI cho các sơ đồ điều chế [38] . . . . .                                | 64 |
| 2.3 Các tham số sử dụng trong mô phỏng của thuật toán đề xuất . . . . .                       | 70 |
| 3.1 Các ký hiệu sử dụng cho thuật toán lập lịch đề xuất . . . . .                             | 86 |
| 3.2 Các tham số được sử dụng mô phỏng thuật toán HGLA [92] . . . . .                          | 99 |

# DANH MỤC CÁC HÌNH VẼ

|     |  |     |
|-----|--|-----|
| 1   | Cấu trúc của luận án . . . . .   | 13  |
| 1.1 | Các thành phần của hệ thống vật lý ảo . . . . .  | 15  |
| 1.2 | Các thành phần cơ bản của hệ thống MCPS . . . . .  | 16  |
| 1.3 | Mạng di động đa dịch vụ . . . . .  | 18  |
| 1.4 | Môi trường lưới tính toán di động . . . . .  | 35  |
| 1.5 | Các mục tiêu lập lịch trong MCG . . . . .  | 40  |
| 2.1 | Kiến trúc dịch vụ trong mạng LTE . . . . .   | 54  |
| 2.2 | Mô hình lập lịch trong LTE . . . . .   | 58  |
| 2.3 | Các bước thực hiện của bộ lập lịch đề xuất ITFDS . . . . .   | 62  |
| 2.4 | Trễ gói với số lượng người dùng khác nhau . . . . .  | 72  |
| 2.5 | Tỷ lệ mất gói với số lượng người dùng khác nhau . . . . .  | 72  |
| 2.6 | Thông lượng với số lượng người dùng khác nhau . . . . .  | 73  |
| 2.7 | Chỉ số công bằng với các người dùng khác nhau . . . . .  | 74  |
| 3.1 | Môi trường điện toán lưới . . . . .  | 80  |
| 3.2 | Môi trường lưới tính toán di động . . . . .  | 81  |
| 3.3 | Mô hình mạng cho bài toán lập lịch trong lưới tính toán di động . . . . .                          | 85  |
| 3.4 | Lưới tính toán di động và năng lực các máy được sử dụng trong ví dụ . . . . .                      | 96  |
| 3.5 | Lưới tính toán di động và năng lực các máy được cấu trúc lại sau khi có máy vào, ra lưới . . . . . | 97  |
| 3.6 | Thử nghiệm với tỷ lệ địa phương hóa . . . . .  | 100 |
| 3.7 | Thử nghiệm với tỷ lệ tăng tốc . . . . .  | 101 |
| 3.8 | Thử nghiệm thông lượng của thuật toán với số nút từ 100 tới 1100 . . . . .                         | 102 |
| 3.9 | Thử nghiệm thông lượng của thuật toán với số nút từ 700 đến 1800 . . . . .                         | 103 |

# MỞ ĐẦU

## 1. Tính cấp thiết của luận án

Gần đây, với sự phát triển của công nghệ, các hệ thống tính toán phân tán không đồng nhất đang nổi lên như một xu hướng tất yếu để giải quyết các bài toán khoa học và tính toán chuyên sâu. Các tài nguyên của hệ thống được quản lý bởi các trung tâm dữ liệu và cấp phát động cho người dùng dựa trên tính khả dụng và nhu cầu người dùng dựa trên các tham số chất lượng dịch vụ.

Dựa trên hạ tầng phân tán, điện toán đám mây là một trong những nền tảng ảo mới, đáng tin cậy, có thể mở rộng dễ dàng cung cấp cơ sở hạ tầng và các dịch vụ cho người dùng cuối. Theo cách tiếp cận này, lợi ích mà điện toán đám mây mang lại cho các ứng dụng lớn và phức tạp như bài toán luồng công việc [103]. Điện toán đám mây cung cấp một số lợi thế so với các hệ thống tính toán hiệu năng cao (HPC) truyền thống như: (1) cung cấp tài nguyên theo yêu cầu; (2) giảm chi phí phần cứng do tài nguyên được chia sẻ dựa trên mô hình “trả phí cho mỗi lần sử dụng”; (3) cải thiện việc sử dụng tài nguyên với việc sử dụng các máy ảo [97]; (4) sự hài lòng của người dùng tăng lên khi tuân theo các thỏa thuận dịch vụ (SLA), là một thỏa hiệp giữa người dùng và nhà cung cấp dịch vụ.

Việc đạt được chất lượng dịch vụ (QoS) trong môi trường điện toán ảo hóa với các nguồn lực sẵn có vẫn là một thách thức không nhỏ. Mục tiêu chính là lập

lịch cho các nguồn lực sẵn có để có thể cải thiện chất lượng dịch vụ (QoS) tổng thể do môi trường ảo hóa này cung cấp [44].

Bài toán lập lịch nói chung được mô tả theo nhiều cách khác nhau và thường là tái hiện lại quan điểm cổ điển về trình tự công việc trong quản lý sản xuất. Một trình tự công việc là một loạt các bước nhằm đơn giản hóa sự phức tạp của việc thực thi và quản lý của ứng dụng. Các lời giải cho một trình tự công việc tạo thành một mô hình chung để mô tả một loạt các ứng dụng khoa học trong các hệ thống phân tán [15]. Bài toán lập lịch đóng vai trò quan trọng trong nhiều lĩnh vực khoa học như hóa học, vật lý, sinh học v.v., thường được biểu diễn bằng đồ thị có hướng không có chu trình (DAG), trong đó, các nút đại diện cho các tác vụ tính toán và các cạnh biểu thị mức độ ưu tiên và ràng buộc luồng giữa các tác vụ. Một phân loại cho lớp các bài toán quản lý luồng công việc trong điện toán lưới được nêu chi tiết trong [116].

Thực thi các quy trình phức tạp trong môi trường máy tính phân tán không đồng nhất là mục tiêu chính trong việc áp dụng thông tin và công nghệ truyền thông trong công nghiệp và khoa học. Thuật ngữ “lập lịch luồng công việc” đề cập đến việc lập kế hoạch tài nguyên, tức là, ánh xạ không gian và thời gian của các nhiệm vụ luồng công việc lên tài nguyên hiện có. Lập lịch cho các tác vụ tính toán và luồng công việc trên lưới là một bài toán tối ưu hóa phức tạp, thường đòi hỏi một số tiêu chí lập lịch như thời gian thực thi, chi phí, tính công bằng, sử dụng tài nguyên v.v phải được xem xét [110].

Do đặc điểm không đồng nhất của các thiết bị, các đặc tính khác nhau của các công việc, bài toán lập lịch công việc là bài toán thuộc lớp NP-đầy đủ, không có một phương pháp cụ thể nào để giải bài toán này trong thời gian đa thức. Do đó, các kỹ thuật dựa trên meta-heuristic được coi là một phương pháp nhằm tìm kiếm lời giải gần tối ưu [99]. Meta-heuristic là phương pháp được phát triển gần

đây nhất trong các phương pháp tìm kiếm gần đúng để giải quyết các vấn đề tối ưu nảy sinh trong kỹ thuật, công nghiệp, và nhiều lĩnh vực khác. Meta-heuristic sử dụng các kỹ thuật có nguồn gốc từ trí tuệ nhân tạo, sinh học, toán học, tự nhiên để cải thiện hiệu suất tìm kiếm [66]. Nhiều nghiên cứu đã áp dụng các thuật toán meta-heuristic trong lập lịch công việc [20].

Điện toán lưới (GC) là một tập hợp các máy tính phân tán được kết nối mạng nhằm chia sẻ tài nguyên và làm việc cùng nhau để hoàn thành nhiệm vụ chung [12]. Sự gia tăng nhanh chóng của các thiết bị điện toán hỗ trợ cá nhân (PDA), thiết bị không dây di động đã thu hút sự quan tâm của các nhà khoa học nhằm tích hợp các thiết bị này vào lưới tính toán (CG) [73].

Tích hợp các thiết bị không dây di động vào lưới nhằm mở rộng hệ thống tính toán để chia sẻ một lượng tài nguyên dư thừa của các thiết bị không dây di động còn có nhiều thách thức [73]. Mặc dù đã có nhiều tiêu chuẩn và công nghệ hỗ trợ cho điện toán lưới bằng cách sử dụng các máy tính cố định tại một nơi, tuy nhiên các công nghệ này không thể dễ dàng điều chỉnh cho các thiết bị di động do nhiều giả định khi thiết kế lưới không có sự xuất hiện của chúng.

Một vấn đề gặp phải khác khi làm việc với các thiết bị di động đó là tính thiếu ổn định trong kết nối mạng. Vấn đề này chủ yếu bắt nguồn từ việc một thiết bị di động di chuyển từ vị trí này tới vị trí khác. Với sự phát triển của công nghệ, phạm vi phủ sóng của các trạm cơ sở đã được mở rộng nhưng vẫn tồn tại các vị trí mà ở đó các thiết bị không nhận được tín hiệu. Mặc dù còn rất nhiều khó khăn khi tích hợp các thiết bị di động vào lưới, nhưng nếu tích hợp thành công, lưới tính toán sẽ đem lại nhiều lợi ích to lớn: gia tăng số lượng các thiết bị di động với khả năng tính toán cao trong mạng; các thiết bị di động luôn sẵn sàng gia nhập lưới và tạo thành một lưới di động có mặt ở khắp nơi cho phép lưới tăng khả năng xử lý các bài toán lớn [73].

Một ví dụ về lưới các thiết bị di động: những cảm biến được nhúng trong các thiết bị di động, camera kết hợp với micro và GPS được sử dụng để thu thập thông tin về môi trường. Các cảm biến trong thiết bị di động có thể hỗ trợ đo lường giao thông trên đường cao tốc hoặc hỗ trợ lập bản đồ tiếng ồn của một thành phố [67]. Với sự hỗ trợ của trí tuệ nhân tạo (AI), tính toán phân tán, giao tiếp trực tiếp giữa các thiết bị sử dụng bất kỳ kênh liên lạc nào (M2M) giúp cuộc cách mạng công nghiệp 4.0 trở thành hiện thực. Một trong những vấn đề quan trọng của cuộc cách mạng công nghiệp 4.0 là làm thế nào để tăng hiệu suất truyền thông của các thiết bị di động.

Do nhu cầu ngày càng tăng đối với các dịch vụ băng thông rộng di động với tỷ lệ dữ liệu, tốc độ và chất lượng dịch vụ cao đã thúc đẩy dự án đối tác thế hệ thứ ba (3GPP) phát triển công nghệ truyền thông thế hệ thứ tư (4G) tiến hóa dài hạn (LTE). Bản phát hành 8 của LTE cho phép băng thông có thể mở rộng từ 5Mhz đến 20MHz với tốc độ dữ liệu cao nhất lên đến 300 Mb/giây cho đường xuống và 75 Mb/giây cho đường lên [34]. Các thiết bị trong mạng di động tùy biến (MANET) và lưới tính toán di động (MCG) giao tiếp thông qua phương tiện truyền dẫn không dây và truyền thông di động như LTE/5G [13, 86, 114].

Công nghệ OFDMA cho phép nhiều quyền truy cập tài nguyên vô tuyến được chia sẻ giữa nhiều người dùng bằng cách chia băng thông khả dụng thành nhiều sóng mang con băng hẹp và phân bổ một nhóm sóng mang con cho người dùng dựa trên: yêu cầu; tải hiện tại của hệ thống; và cấu hình hệ thống. Mạng LTE được kỳ vọng sẽ hỗ trợ một loạt các dịch vụ đa phương tiện và Internet ngay cả trong tình huống di động cao. Trong LTE, khối quản lý tài nguyên vô tuyến (RRM) khai thác kết hợp các chức năng vật lý và MAC nâng cao như chia sẻ tài nguyên, chỉ thị chất lượng kênh truyền (CQI), điều chế mã hóa thích ứng (AMC) và yêu cầu truyền lại tự động kết hợp (HARQ) để đạt được mục tiêu này.

Việc thiết kế một chiến lược phân bổ nguồn lực tài nguyên một cách hiệu quả là bài toán hết sức quan trọng. Trong thực tế, sử dụng hiệu quả tài nguyên vô tuyến là cần thiết nhằm đáp ứng các nhu cầu người dùng theo các yêu cầu chất lượng dịch vụ (QoS) cụ thể. Các ứng dụng đa phương tiện chiếm đa số ứng dụng của mạng LTE bao gồm tương tác trực tiếp, hội nghị truyền hình và vô tuyến. Lập lịch phân bổ tài nguyên trong LTE tới nhiều người dùng được thực hiện bởi eNodeB phụ thuộc vào chất lượng kênh truyền và chất lượng dịch vụ [54]. Sự phổ biến của các ứng dụng đa phương tiện là do tốc độ truyền của chúng cao hơn bất kỳ định dạng dữ liệu nào khác. Do đó, để lưu lượng truy cập cho thông tin di động được cao hơn, bộ lập lịch LTE cần đảm bảo chất lượng dịch vụ về độ trễ, tính chính xác của dữ liệu nhận được, tỷ lệ mất dữ liệu cho các ứng dụng dạng này [62].

Trong kiến trúc LTE, một thành phần quan trọng chịu trách nhiệm lập lịch phân bổ tài nguyên tới người dùng là trạm cơ sở (eNodeB). Trong lập lịch LTE, một số tham số được sử dụng để đánh giá hiệu suất của hệ thống như hiệu suất phổ, độ trễ, tính công bằng và thông lượng hệ thống. Việc đáp ứng tất cả các tham số này là điều khó có thể đạt được do bài toán lập lịch phân bổ tài nguyên là bài toán thuộc lớp *NP*-đầy đủ. Sự đa dạng các tham số dẫn đến việc tạo ra nhiều chiến lược lập lịch và nhiều thuật toán lập lịch khác nhau. Các tham số này cần phải được chỉ rõ trong bài toán lập lịch và được mạng LTE phân biệt giữa các luồng dữ liệu:

*Lớp hội thoại:* Gồm các ứng dụng như hội nghị truyền hình, điện thoại. Các ứng dụng dạng này không chấp nhận trễ.

*Lớp truyền hình trực tuyến:* Gồm ứng dụng như truyền video. Giống như lớp hội thoại nhưng nó giả định rằng chỉ có một người ở cuối kết nối, do đó, nó ít đòi hỏi về thời gian và độ trễ hơn.

*Lớp tương tác:* Các ví dụ về lớp này là: duyệt web, truy cập cơ sở dữ liệu v.v. Không giống như các lớp đã đề cập trước đó, dữ liệu cần được phân phối trong một khoảng thời gian xác định và nhấn mạnh đến tỷ lệ mất dữ liệu.

*Lớp nền:* Không đòi hỏi chất lượng dịch vụ; ứng dụng dạng này chấp nhận sự chậm trễ, mất gói. Ví dụ về lớp này: FTP, E-mail.

Cùng với sự phát triển nhanh chóng và đa dạng của điện thoại thông minh hỗ trợ 4G-LTE, chúng đã kết hợp với nhau và với mạng di động băng rộng 4G-LTE cung cấp dịch vụ tốt hơn cho người sử dụng. Các yêu cầu về tốc độ, chất lượng và sự sẵn sàng luôn đòi hỏi các nhà cung cấp dịch vụ di động phải thay đổi phương thức cung cấp dịch vụ. **Mạng di động đa dịch vụ** được xem là mạng các thiết bị di động băng rộng kết nối qua mạng 4G-LTE hay 5G nhằm cung cấp lượng lớn các dịch vụ đa phương tiện tốc độ cao tới người dùng.

Lập lịch trong các hệ thống phân tán là một bài toán có ý nghĩa thực tiễn lớn giúp tăng chất lượng dịch vụ, giảm chi phí cho cả người cung cấp dịch vụ và người sử dụng dịch vụ. Mặc dù vậy, hầu hết các thuật toán lập lịch hiện nay sử dụng cho các hệ thống lớn còn khá đơn giản, chưa mang lại hiệu quả theo hướng tối ưu hóa chi phí. Nguyên nhân chính là do các vấn đề lập lịch trong thực tế của các hệ thống này thường chứa đựng nhiều tác nhân phức tạp mà hầu hết các thuật toán lập lịch chưa giải quyết được triệt để do dựa nhiều và các giả thuyết. Chính vì vậy, nghiên cứu bài toán lập lịch theo hướng tăng độ hiệu quả chính xác các thuật toán này trong các hệ thống phân tán có ý nghĩa quan trọng về tính khoa học và thực tiễn. Cụ thể, luận án quan tâm đến hai bài toán sau:

(1). **Bài toán lập lịch tài nguyên trong mạng di động đa dịch vụ:** Bài toán được chia làm hai loại: lập lịch cho đường tải xuống và lập lịch cho đường tải lên. Trong phạm vi luận án, luận án tập trung vào bài toán lập lịch cho đường tải xuống. Tải xuống là truyền dữ liệu từ trạm cơ sở đến thiết bị di động. Bộ lập lịch tải xuống là thành phần chính của trạm cơ sở thực hiện gán các khối tài nguyên vật lý chia sẻ tới các người dùng khác nhau trong mạng. Bài toán được phát biểu như sau: có  $k$  hàng đợi  $q_1, q_2, \dots, q_k$  chứa dữ liệu lập lịch. Các hàng đợi này gửi thông tin lập lịch đến trạm cơ sở. Dựa trên chất lượng kênh truyền, trạm cơ sở phân phối tài nguyên tới các người dùng khác nhau trong mạng theo cả miền thời gian và miền tần số. Theo miền thời gian, kênh tải xuống được chia thành các khe có độ lớn 10ms. Theo miền tần số, băng thông được chia thành các kênh phụ có độ lớn 180KHz. Tài nguyên được cấp phát tới người dùng là các khối tài nguyên có độ lớn 1ms theo miền thời gian và 180KHz theo miền tần số. Thuật toán lập lịch phân phối các khối tài nguyên tới người dùng phải đảm bảo các yếu tố chất lượng như **giảm độ trễ gói, giảm tỷ lệ mất gói, tăng băng thông hệ thống, tăng tỷ lệ cân bằng** giữa các người dùng.

(2). **Bài toán lập lịch trong lưới tính toán di động:** Là bài toán phân bổ các công việc vào các lưới tính toán gồm lưới di động và lưới cố định sao trong cùng một khoảng thời gian, số lượng công việc được thực hiện là nhiều nhất. Bài toán được phát biểu như sau: có  $n$  công việc, mỗi công việc có  $m_i$  công việc phụ cần lập lịch phân bổ tài nguyên để thực hiện. Bộ lập lịch toàn cục phân phối các công việc của các công việc vào từng lưới phụ trong mạng theo chiến lược heuristic. Bộ lập lịch địa phương sau khi nhận được các công việc phụ sẽ phân phối chúng tới các máy trong lưới để thực hiện. Một chu trình lập lịch là tổng thời gian các công việc được gửi tới lưới phụ, thời gian chờ tại lưới phụ, thời gian thực thi các công việc. Bộ lập lịch phải đảm bảo các yêu cầu người dùng như **tăng tỷ**

lệ địa phương hóa, tăng thông lượng hệ thống và tăng tỷ lệ tăng tốc cũng như giảm thời gian của một chu trình.

## 2. Một số vấn đề tồn tại trong các nghiên cứu liên quan

Một số vấn đề còn tồn tại với bài toán lập lịch chưa được giải quyết:

1. Với bài toán lập lịch đường tải xuống trong mạng di động đa dịch vụ, các thuật toán lập lịch như thuật toán thông lượng tối đa (MT) tập trung vào tối đa thông lượng mà chưa quan tâm đến tính cân bằng, thuật toán cân bằng (PF) tập trung vào tính cân bằng mà bỏ qua thông lượng, thuật toán EXP/PF cố gắng tìm kiếm giải pháp cân bằng giữa thông lượng và tính cân bằng nhưng mới tập trung vào ứng dụng thời gian thực. Do đó cần phải có một thuật toán lập lịch đường tải xuống đảm bảo cân bằng các yếu tố về độ trễ gói, tỷ lệ mất gói, băng thông và tính cân bằng cho cả ứng dụng thời gian thực và không thời gian thực;
2. Với bài toán lập lịch trong lưới tính toán di động, các thuật toán heuristic hiện tại như thuật toán MIN-MIN, MAX-MIN, thuật toán thời gian kết thúc sớm nhất không đồng nhất (HEFT) mới chỉ tập trung vào môi trường mạng đồng nhất mà chưa tập trung vào môi trường mạng hiện diện cả nút tĩnh và nút động. Do đó, cần thiết có một thuật toán lập lịch cho lưới tính toán di động có thể áp dụng cho lưới có cả nút tĩnh và động, đồng thời đảm bảo được các tiêu chí chất lượng dịch vụ tỷ lệ địa phương hóa, tỷ lệ tăng tốc và thông lượng.

Chính vì vậy, luận án đã chọn đề tài “*Nghiên cứu đề xuất cải tiến thuật toán lập lịch và ứng dụng*” làm đề tài nghiên cứu và giới hạn tập trung vào

*“Nghiên cứu phát triển thuật toán lập lịch cho mạng di động đa dịch vụ và lưới tính toán di động”.*

### 3. Mục tiêu và đối tượng nghiên cứu

#### Mục tiêu nghiên cứu

Từ việc nghiên cứu các hệ thống phân tán và mạng di động đa dịch vụ, có hai vấn đề cần quan tâm là phân bổ tài nguyên và truyền thông giữa các thành phần trong hệ thống. Từ đó, luận án đề xuất hai mục tiêu nghiên cứu như sau:

a) *Mục tiêu 1:* Nghiên cứu, đề xuất thuật toán lập lịch cho đường tải xuống trong mạng di động đa dịch vụ kết hợp giữa miền thời gian và miền tần số.

b) *Mục tiêu 2:* Nghiên cứu, đề xuất thuật toán lập lịch trong lưới tính toán di động dựa trên tìm kiếm heuristic.

**Đối tượng nghiên cứu:** Các thuật toán lập lịch công việc dựa trên tìm kiếm heuristic và meta-heuristic, các thuật toán lập lịch trong lưới tính toán di động, các thuật toán lập lịch trong mạng di động đa dịch vụ.

### 4. Nội dung nghiên cứu

Dựa vào mục tiêu nghiên cứu của luận án, các nội dung nghiên cứu của đề tài được trình bày như sau:

a) *Nội dung 1:* Nghiên cứu các hệ thống tính toán phân tán, mạng di động đa dịch vụ. Các thuật toán lập lịch cho đường tải xuống trong mạng di động đa dịch vụ và cho lưới tính toán di động.

b) *Nội dung 2*: Nghiên cứu phát triển thuật toán lập lịch tích hợp giữa miền thời gian và miền tần số cho đường tải xuống trong mạng di động đa dịch vụ.

c) *Nội dung 3*: Nghiên cứu phát triển thuật toán lập lịch trong lưới tính toán di động.

## 5. Phương pháp nghiên cứu

Từ các nội dung nghiên cứu ở trên, các phương pháp nghiên cứu được đề xuất và thực hiện để hoàn thiện đề tài nghiên cứu như sau:

*Nghiên cứu lý thuyết*: Khảo sát các cơ sở lý thuyết liên quan đến lập lịch, phân tích tính đúng đắn của thuật toán đề xuất dựa trên cơ sở lý thuyết.

*Mô phỏng*: Sử dụng công cụ, ngôn ngữ lập trình để thực nghiệm mô phỏng nhằm kiểm tra lại kết quả nghiên cứu lý thuyết.

*Nghiên cứu ứng dụng*: Áp dụng thuật toán đề xuất cho bài toán phân phối tài nguyên cho đường tải xuống trong mạng di động đa dịch vụ và cho lưới tính toán di động.

## 6. Đóng góp chính và tính mới của luận án

### Đóng góp chính của luận án

Luận án tham gia vào nghiên cứu các hệ tính toán phân tán, các thuật toán lập lịch trong mạng di động đa dịch vụ và trong lưới tính toán di động. Từ các tồn tại của bài toán lập lịch, luận án tập trung nghiên cứu đề xuất cải tiến thuật toán lập lịch trong mạng di động đa dịch vụ và trong lưới tính toán di động.

Luận án có hai đóng góp chính:

a) *Đóng góp 1*: Thuật toán lập lịch tài nguyên cho đường tải xuống trong mạng di động đa dịch vụ kết hợp giữa miền thời gian và miền tần số

b) *Đóng góp 2*: Thuật toán lập lịch tài nguyên cho lưới tính toán di động dựa theo tìm kiếm heuristic.

### **Tính mới của luận án**

Với đóng góp 1, thuật toán lập lịch đề xuất phân bổ tài nguyên tới người dùng kết hợp cả miền thời gian và miền tần số nhằm cải thiện tỷ lệ trễ gói, tỷ lệ mất gói, băng thông hệ thống và tỷ lệ cân bằng. Đóng góp này được công bố trong [CT2, CT3]. Với đóng góp 2, thuật toán đề xuất hoạt động trong môi trường mạng không đồng nhất gồm cả các nút tính và các nút di động. Cấu trúc mạng thay đổi theo thời gian khi các nút di động tham gia hoặc rời mạng. Đóng góp này được công bố trong [CT4].

## **7. Phạm vi và giới hạn của đề tài nghiên cứu**

Từ những mục tiêu, nội dung và phương pháp nghiên cứu, phạm vi và giới hạn của đề tài nghiên cứu được đề xuất như sau:

a) *Lý thuyết*: Nghiên cứu các hệ thống tính toán phân tán, mạng di động đa dịch vụ. Một số thuật toán lập lịch cho đường tải xuống trong mạng di động đa dịch vụ và một số thuật toán lập lịch cho lưới tính toán di động dựa theo heuristic. Đề xuất thuật toán lập lịch cho đường tải xuống kết hợp giữa miền thời gian và miền tần số trong mạng di động đa dịch vụ và thuật toán lập lịch tài nguyên cho lưới tính toán di động.

b) *Ứng dụng*: Áp dụng thuật toán đề xuất vào môi trường mạng không đồng nhất gồm các nút di động và các nút cố định. Kiến trúc mạng luôn thay đổi do có nút mạng vào và ra khỏi mạng.

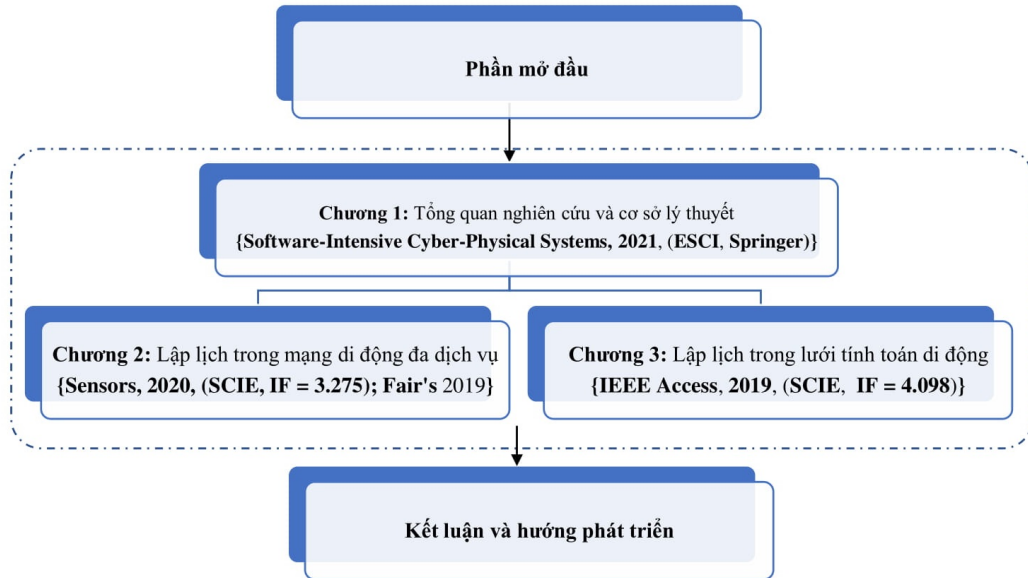
## 8. Môi trường mô phỏng và công cụ đánh giá

Để kiểm chứng kết quả các thuật toán đề xuất, luận án đã sử dụng môi trường mô phỏng và công cụ đánh giá sau: 1) với bài toán lập lịch trong mạng di động đa dịch vụ, luận án sử dụng MATLAB và NS3 để thực hiện mô phỏng và thực hiện các đánh giá. Các tiêu chí đánh giá gồm tỷ lệ trễ gói, tỷ lệ mất gói, thông lượng hệ thống và chỉ số công bằng; 2) với bài toán lập lịch trong lưới tính toán di động, luận án sử dụng Eclipse và GridSim để thực hiện mô phỏng và thực hiện các đánh giá. Các tiêu chí đánh giá gồm tỷ lệ địa phương hóa, chỉ số tăng tốc và thông lượng hệ thống. Các phân tích về kết quả thực nghiệm được trình bày cụ thể tại mục các thực nghiệm và kết quả tại chương 2 và chương 3.

## 9. Cấu trúc luận án

Luận án “*Nghiên cứu phát triển thuật toán lập lịch cho đường tải xuống trong mạng di động đa dịch vụ và cho lưới tính toán di động*” bao gồm 3 chương. Ngoài phần *Mở đầu* trình bày về tính cấp thiết của đề tài, lý do chọn đề tài, đối tượng và nội dung nghiên cứu của luận án. Các chương còn lại được tổ chức như trong hình (1), cụ thể như sau:

Chương 1, *Tổng quan nghiên cứu và cơ sở lý thuyết*. Chương này trình bày các kiến thức nền tảng lý thuyết và bài toán lập lịch trong mạng di động đa dịch



**Hình 1.** Cấu trúc của luận án

vụ và lưới tính toán di động làm cơ sở cho các chương tiếp theo của luận án [CT1].

Chương 2, *Lập lịch cho mạng di động đa dịch vụ*. Chương này trình bày mô hình mạng vào các giả thiết, từ đó dẫn đến mô hình toán học làm cơ sở cho thuật toán lập lịch cho đường tải xuống trong mạng đa dịch vụ [CT2] và [CT3].

Chương 3, *Lập lịch cho lưới tính toán di động*. Chương này trình bày mô hình mạng và các giả thiết, từ đó dẫn đến mô hình toán học làm cơ sở cho thuật toán lập lịch trong lưới tính toán di động [CT4].

Cuối cùng, phần *Kết luận và hướng phát triển*, trình bày các kết quả đã đạt được và thảo luận về các hướng nghiên cứu trong tương lai.

## Chương 1

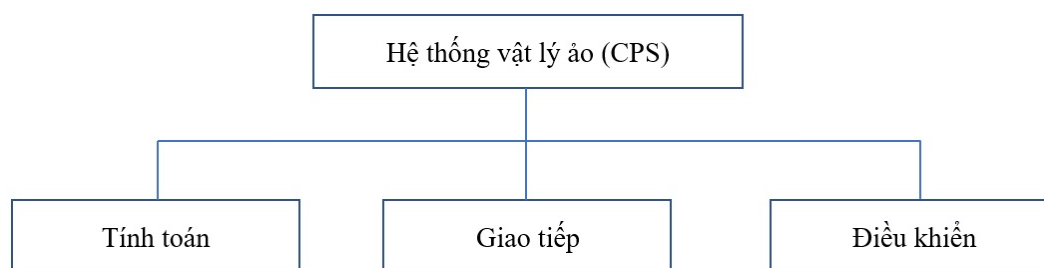
# TỔNG QUAN NGHIÊN CỨU VÀ CƠ SỞ LÝ THUYẾT

Chương này, luận án sẽ trình bày những kiến thức cơ sở được sử dụng trong các chương tiếp theo. Mở đầu, mục 1.1, luận án sẽ trình bày tổng quan về hệ thống tính toán phân tán, tập trung vào hệ thống vật lý ảo (CPS), mở rộng với ví dụ về hệ thống vật lý ảo trong y tế (MCPS) và bối cảnh bài toán lập lịch. Mục 1.2, luận án trình bày tổng quan về lập lịch trong mạng di động đa dịch vụ và mục 1.3, luận án trình bày tổng quan về lập lịch cho lưới tính toán di động.

### 1.1 Sơ lược về hệ thống tính toán phân tán và bối cảnh bài toán lập lịch

Hệ thống tính toán phân tán là hệ thống gồm các máy tính phân tán được điều khiển hoặc được giám sát bằng các thuật toán dựa trên máy tính. Trong hệ thống phân tán, các thành phần vật lý và phần mềm kết hợp với nhau một cách chặt chẽ, hoạt động theo không gian và thời gian, thể hiện nhiều phương thức riêng biệt và tương tác với nhau theo ngữ cảnh [26]. Hệ thống vật lý ảo được xây

dựng theo cách tiếp cận đa ngành, là sự hợp nhất của điều khiển học, cơ điện tử, khoa học thiết kế và các quá trình. Hệ thống thực ảo chia sẻ cùng một kiến trúc cơ bản tương tự hệ thống IoT nhưng CPS thể hiện một sự kết hợp chặt chẽ hơn giữa các yếu tố vật lý và tính toán. Hệ thống CPS được trình bày như trong hình (1.1) gồm: Tính toán, giao tiếp và điều khiển. Điều này có nghĩa CPS là một hệ

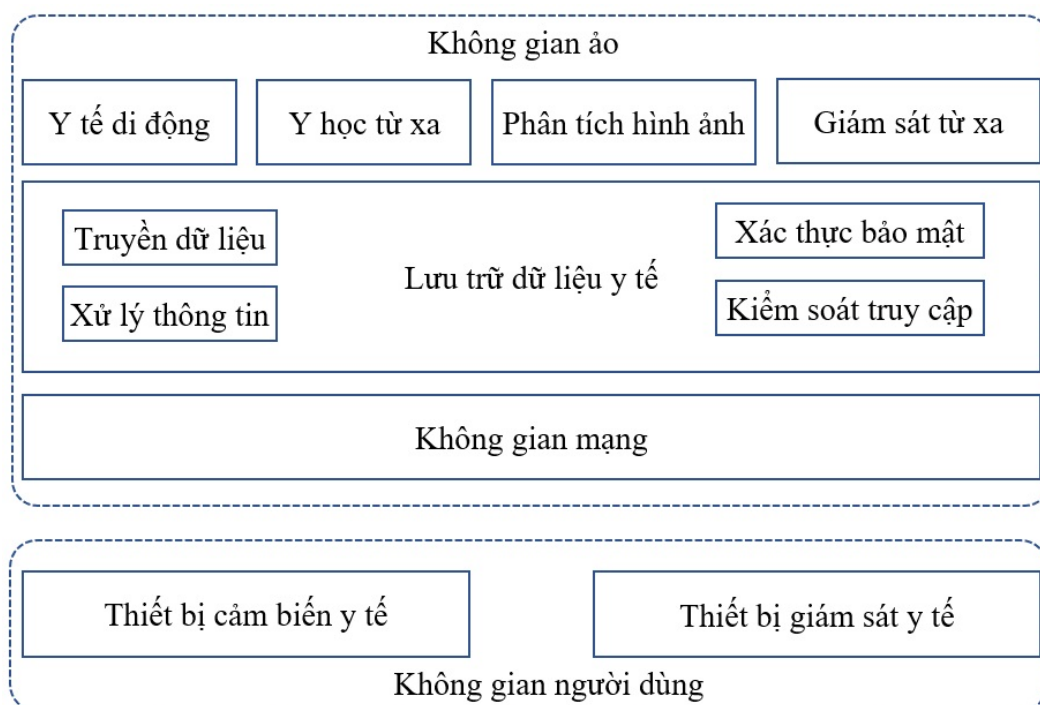


**Hình 1.1.** Các thành phần của hệ thống vật lý ảo

thống phức tạp, tích hợp sâu các thiết bị tính toán, truyền thông và điều khiển. Truyền thông là trung tâm cốt lõi để đảm bảo sự tương tác hiệu quả giữa quá trình tính toán rời rạc trong thế giới mạng và quá trình vật lý liên tục trong thế giới thực [112].

Một ví dụ của CPS là Hệ thống vật lý ảo trong y tế (Medical CPS - MCPS) gồm các hệ thống nhúng điều khiển và hệ thống mạng độc lập [47]. Một hệ thống MCPS cơ bản gồm không gian mạng ảo (bao gồm không gian mạng) và không gian vật lý (bao gồm cả không gian người dùng), như thể hiện trong hình (1.2).

Không gian vật lý, bao gồm các thiết bị cảm biến sức khỏe theo thời gian thực và thiết bị chuẩn đoán sức khỏe cung cấp thông tin cảm biến cho không gian mạng thông qua các thiết bị cảm biến và nhận thông tin điều khiển từ không gian mạng để điều khiển các thiết bị vật lý. Không gian mạng là thành phần cốt lõi của MCPS, chịu trách nhiệm xử lý, lưu trữ và quản lý bảo mật truy cập của người dùng và thông tin sức khỏe. Là trung tâm của hệ thống của MCPS, không gian mạng nhận thông tin cảm biến từ không gian vật lý thông qua hệ thống



**Hình 1.2.** Các thành phần cơ bản của hệ thống MCPS

truyền dẫn mạng, xác thực, lưu trữ, phân tích và xử lý chúng đồng thời tạo ra thông tin điều khiển phản hồi được gửi đến không gian vật lý thông qua mạng hệ thống truyền dẫn [79].

Kiến trúc MCPS truyền thống tập trung vào việc theo dõi bệnh nhân và phản hồi dữ liệu. Một số nhà nghiên cứu tập trung vào sự tương tác giữa không gian vật lý và không gian mạng. Các hệ thống dạng này được xây dựng dựa trên CPS kết hợp dữ liệu lớn và điện toán đám mây. Dựa trên kiến trúc IoT, nhiều nhà nghiên cứu khác sử dụng các thiết bị đeo cảm biến được để cảm biến dữ liệu bệnh nhân trong các ứng dụng của mình. Bệnh nhân, bác sĩ, người quản lý là các người dùng chính, hoàn thành giao tiếp của lớp mạng thông qua IPv4 hoặc IPv6 bên trong bệnh viện và truyền dữ liệu của lớp cảm biến đến lớp ứng dụng cho việc phân tích hoặc trình bày trực quan.

Kiến trúc của hệ thống có thể mô tả chính xác các thành phần và mối quan hệ giữa chúng. Kiến trúc hệ thống MCPS có thể được phân loại dựa trên một số khía cạnh:

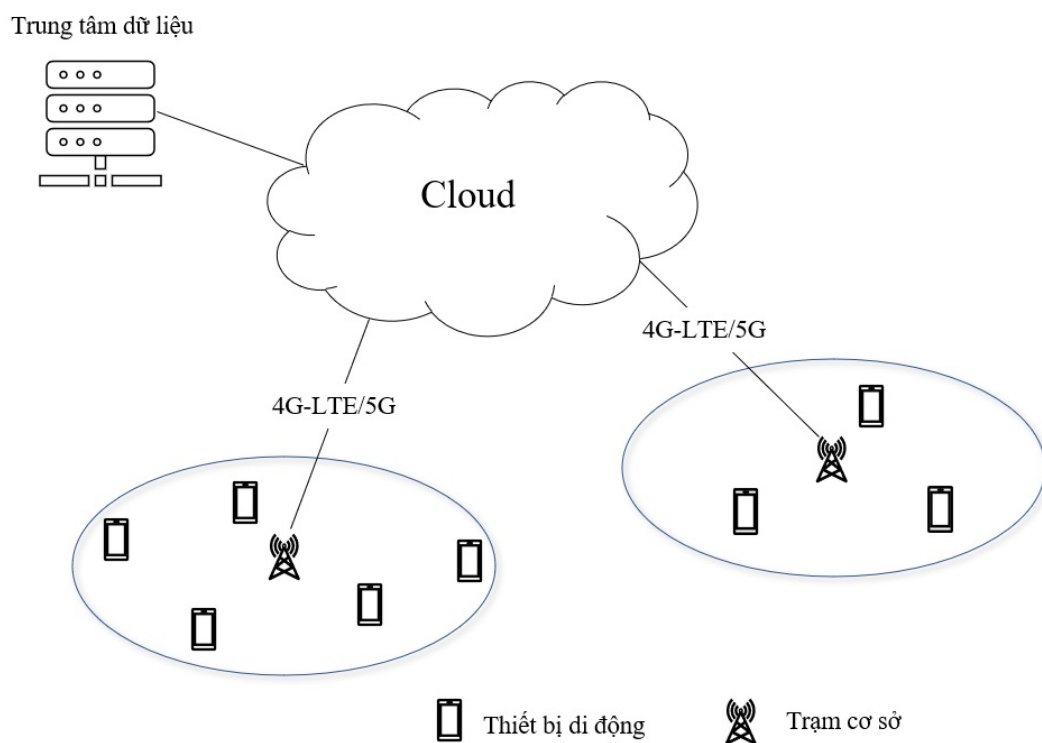
- *Chức năng và hành vi*: Nhiệm vụ chính của các hệ thống MCPS là thu thập, truyền dẫn, phân tích, xử lý, lưu trữ và truy cập dữ liệu sức khỏe cá nhân. Theo chức năng và hành vi khác nhau, MCPS được chia làm 2 loại: Kiến trúc truyền dẫn [1, 49] và kiến trúc xử lý dữ liệu [61].
- *Cấu trúc và phân cấp*: Mô hình MCPS gồm 3 lớp: Thu thập năng lượng và tạo dữ liệu, tiền xử lý dữ liệu, xử lý dữ liệu và ứng dụng.
- *Truyền thông*: Theo các phương thức truyền thông khác nhau, MCPS được chia thành 5 loại: Mô hình khách/chủ, điểm/điểm, đối tượng phân tán, thành phần phân tán và các dịch vụ được phân phối dựa trên sự kiện [85].

Là một sự mở rộng của CPS trong lĩnh vực y tế, MCPS được kế thừa 3 đặc điểm cơ bản của CPS, do đó, các vấn đề về truyền thông cần phải được xem xét đầy đủ trong hệ thống MCPS.

- *Công nghệ truyền thông không dây (Lớp 1/2)*: Nhiều thiết bị cảm biến khác nhau cùng tồn tại trong một hệ thống MCPS. Để đảm bảo thông tin được truyền theo thời gian thực, nhiều công nghệ truyền thông cùng được sử dụng, ví dụ, công nghệ ZigBee để truyền dữ liệu sức khỏe của bệnh nhân được thu thập bởi các cảm biến y tế, nhận dạng và xử lý thông tin thuốc dựa trên công nghệ nhận dạng qua tần số vô tuyến (RFID), truyền thông tin của các thiết bị điện toán hỗ trợ cá nhân (PDA) và điện thoại di động dựa trên công nghệ Bluetooth năng lượng thấp. Ngoài ra, Z-Wave, 4G-LTE, 5G, Wi-Fi và công nghệ giao tiếp tầm ngắn (NFC) cũng là những công nghệ giao tiếp không dây khả dụng trong các ngữ cảnh khác nhau của MCPS [11].

- Công nghệ IPv6 (Lớp 3): IPv6 có một không gian địa chỉ lớn, có thể đáp ứng nhu cầu triển khai mạng cảm biến quy mô lớn và mật độ cao [105]. Tuy nhiên, tại lớp dưới cùng của hệ thống, các nút cảm biến thường có đặc điểm tiêu thụ điện năng và tài nguyên hạn chế. Từ góc độ phần cứng, chúng khó có thể chạy trực tiếp trên IPv6. Hơn nữa, có nhiều công nghệ truyền thông không dây trong MCPS, mỗi công nghệ sẽ chạy trên không gian riêng giúp tăng hiệu quả truyền thông trong hệ thống.

## 1.2 Tổng quan về lập lịch tài nguyên trong mạng di động đa dịch vụ



Hình 1.3. Mạng di động đa dịch vụ

Trong CPS, truyền thông giao tiếp là vấn đề cần được xem xét toàn diện. Với truyền thông, các thành phần hệ thống thực hiện truyền thông bên trong hệ

thống hoặc truyền thông với đám mây. Hình (1.3) là một ví dụ về mạng di động đa dịch vụ. Các thiết bị trong mạng, một phần của hệ thống CPS, truyền thông với đám mây qua mạng 4G-LTE/5G.

Chương này sẽ giới thiệu một số dịch vụ của mạng tiến hóa dài hạn (LTE), bài toán lập lịch truyền thông trong mạng LTE, các nghiên cứu liên quan đến lập lịch truyền thông trong mạng tiến hóa dài hạn, đồng thời đưa ra một số nhận xét. Từ đó, làm động lực để phát triển thuật toán lập lịch tích hợp giữa miền thời gian và miền tần số trong mạng di động đa dịch vụ được trình bày ở chương 2.

### 1.2.1 Các dịch vụ của mạng LTE

Với khả năng truyền tốc độ cao, băng thông linh hoạt, hiệu suất phổ cao và giảm độ trễ gói, LTE mang đến nhiều dịch vụ đa dạng tới người dùng. Một số dịch vụ phổ biến của LTE như dịch vụ cung cấp nội dung như hình ảnh, tin nhắn tức thời, gọi điện VoIP, tivi và video cho người dùng dựa trên việc tận dụng không gian rộng lớn Internet, gọi chung là các ứng dụng OTT. Với LTE, khách hàng có thêm nhiều ứng dụng đòi hỏi băng thông dữ liệu lớn, người sử dụng có thể thực hiện các hoạt động của mình trên thiết bị di động mọi lúc, mọi nơi chỉ cần có kết nối Internet qua mạng không dây hoặc mạng di động.

### 1.2.2 Bài toán lập lịch trong mạng LTE

Lập lịch đa người dùng là một trong những chức năng chính của hệ thống LTE với nhiệm vụ phân bổ tài nguyên sẵn có giữa những người dùng đang hoạt động nhằm đáp ứng các yêu cầu dịch vụ.

Trong LTE, kênh dữ liệu được chia sẻ giữa những người dùng, có nghĩa là các phần của phổ phải được phân phối tại mọi khoảng thời gian truyền (TTI).

Lập lịch gói được triển khai tại eNodeB và thực hiện tại mỗi TTI với một khối tài nguyên (RB) theo cả miền thời gian và miền tần số.

Việc phân bổ tài nguyên cho mỗi người dùng thường dựa trên việc so sánh dữ liệu trên mỗi khối tài nguyên. Khối tài nguyên thứ  $k$  được cấp phát cho người dùng thứ  $i$  nếu giá trị  $m_{i,k}$  của nó là lớn nhất, tức là thỏa mãn phương trình:

$$m_{i,k} = \max_i \{m_{i,k}\} \quad (1.1)$$

Các chỉ số này là các chỉ số ưu tiên truyền tải của mỗi người dùng trên khối tài nguyên cụ thể. Dựa trên yêu cầu hiệu suất mong muốn, việc tính toán thường dựa trên các thông tin liên quan tới mỗi luồng dữ liệu để đưa ra quyết định phân bổ tài nguyên:

- *Trạng thái hàng đợi truyền:* Trạng thái hàng đợi truyền tại thiết bị người dùng (UE) có thể được sử dụng để giảm thiểu độ trễ phân phối tài nguyên (hàng đợi càng dài thì giá trị  $m_{i,k}$  càng cao).
- *Chất lượng kênh:* Các giá trị chỉ thị chất lượng kênh truyền (CQI) phản hồi có thể được sử dụng để phân bổ tài nguyên cho người dùng có điều kiện kênh tốt hơn (thông lượng mong muốn càng cao thì giá trị  $m_{i,k}$  càng cao).
- *Lịch sử phân bổ tài nguyên:* Thông tin nhận được về hiệu năng trong quá khứ có thể được sử dụng để cải thiện tính công bằng (thông lượng nhận được trong quá khứ càng thấp thì giá trị  $m_{i,k}$  càng cao).
- *Trạng thái bộ đệm:* Trạng thái bộ đệm phía người dùng có thể được sử dụng để tránh tràn bộ đệm (không gian khả dụng trong bộ đệm phía người dùng càng cao thì giá trị  $m_{i,k}$  càng cao).

Tại mỗi TTI, bộ lập lịch quyết định mức phân bổ hợp lệ cho TTI kế tiếp và gửi thông tin đến người dùng sử dụng kênh điều khiển đường xuống vật lý

(PDCCH). Các bản tin điều khiển đường xuống trong PDCCH thông báo tới các thiết bị người dùng về các khối tài nguyên đã cấp phát để truyền trên PDCCH.

Một khái niệm quan trọng liên quan đến bộ lập lịch là “độ nhảy kênh”. Ý tưởng là lập lịch truyền cho các UE tại thời điểm hiện tại và trên một tần số nhất định có điều kiện kênh “tốt” dựa trên số liệu đã chọn. Cách tiếp cận này còn được gọi là lập lịch gói theo miền tần số (FDPS), chống lại tính chất chọn lọc tần số theo thời gian của kênh.

Do là một thành phần quan trọng của hệ thống LTE, bộ lập lịch cần phải được thiết kế có tính đến:

- *Độ phức tạp và khả năng mở rộng*: Bộ lập lịch gói trong mạng LTE hoạt động với thời gian mỗi 1ms và phải đưa ra quyết định phân bổ tài nguyên tại mỗi TTI. Do đó, khả năng mở rộng thấp để hạn chế thời gian xử lý và sử dụng bộ nhớ. Việc tìm ra quyết định phân bổ tài nguyên tốt nhất thông qua các bài toán tối ưu phi tuyến có thể tốn kém về chi phí và thời gian tính toán [46]. Do đó, các quyết định phân bổ tài nguyên của bộ lập lịch gói theo miền tần số thường dựa trên việc tính toán các chỉ số trên mỗi khối tài nguyên (RB) cho mỗi người dùng. Theo cách này, độ phức tạp của thuật toán giảm xuống do mỗi RB được phân bổ cho người dùng có chỉ số cao nhất độc lập với các RB khác. Giả sử, gọi  $N$  và  $R$  tương ứng là số người dùng đang hoạt động trong TTI hiện tại và số RB khả dụng. Bộ lập lịch sẽ tính giá trị  $M = N * R$  tại mỗi TTI và do đó đảm bảo được các yêu cầu mở rộng vì giá trị đo phụ thuộc tuyến tính vào số lượng người dùng hoạt động và số khối tài nguyên khả dụng.
- *Hiệu suất phổ tần*: Trong lập lịch LTE, một trong những mục tiêu chính cần đạt được là tối ưu hóa sử dụng tài nguyên vô tuyến, vì thế, một số chỉ số

hiệu năng được sử dụng, ví dụ, một chính sách cụ thể có thể làm tối đa hóa số lượng người dùng được phục vụ trong một khoảng thời gian nhất định hoặc phổ biến hơn là hiệu suất phổ tần (bit/s/Hz) bằng cách luôn phục vụ những người dùng đang có điều kiện kênh truyền tốt nhất. Một trong những chỉ số hiệu quả được sử dụng nhiều nhất là thông lượng của người dùng, đó là thước đo tốc độ dữ liệu truyền thực tế không bao gồm chi phí truyền lại gói do lỗi vật lý.

- *Cân bằng*: Tối đa hóa thông lượng toàn mạng chắc chắn cho phép sử dụng kênh hiệu quả về mặt phổ nhưng sẽ dẫn đến việc chia sẻ tài nguyên mất công bằng giữa những người dùng. Do đó tính công bằng là một yêu cầu chính cần được tính đến để đảm bảo hiệu suất tối thiểu cho những người có điều kiện kênh truyền không tốt. Vấn đề này có thể được khắc phục bằng cách xem xét giá trị đo và thông lượng mà người dùng nhận được trong quá khứ, từ đó đảm bảo tính công bằng cho người dùng.
- *Thỏa thuận chất lượng dịch vụ (SLA)*: Việc đáp ứng chất lượng dịch vụ cho người dùng là rất quan trọng trong các mạng di động thế hệ tiếp theo. Đây là một tính năng chính trong kiến trúc IP. LTE ánh xạ các luồng ràng buộc ràng buộc chất lượng dịch vụ tới các thiết bị mang sóng vô tuyến chuyên dụng, tùy thuộc vào tham số xác định cách đối xử của nút mạng với những gói tin IP nhận được trên mỗi kênh mang (QCI), cho phép thực hiện các quy trình của bộ quản lý tài nguyên vô tuyến (RRM). Các ràng buộc chất lượng dịch vụ có thể khác nhau tùy thuộc vào ứng dụng và chúng thường được ánh xạ thành một số tham số: tốc độ bit, độ trễ phân bổ tài nguyên và tỷ lệ mất gói.

Tuy nhiên các hệ thống LTE trong thực tế vẫn phải đối mặt với một số nhược điểm:

- *Kiểm soát chi phí:* Do kênh điều khiển đường xuống vật lý mang các bản tin điều khiển có thể chiếm một số ký hiệu trong OFDM (có thể lên đến 3) làm cho số lượng tài nguyên dành riêng cho PDCCH bị giảm xuống dẫn đến giảm độ tự do cho các bộ lập lịch. Chi phí PDCCH có thể được giảm thiểu bằng các thủ tục RRM bằng cách giảm số lượng người dùng được phân bổ tài nguyên trước (giảm số bản tin DCI) hoặc sử dụng bản tin DCI tốc độ bit thấp. Giải pháp thứ hai đặt ra một số ràng buộc đối với việc xây dựng mặt nạ cấp phát RB như cho phép chỉ gán các RB liền kề cho cùng một UE.
- *Đa dạng người dùng:* Đối với đường xuống, UE chịu trách nhiệm gửi CQI tới eNodeB. Lược đồ CQI được chọn sẽ ảnh hưởng đến độ đa dạng người dùng, do nó phản ánh thông tin chất lượng kênh khả dụng của bộ lập lịch [72]. Trong trường hợp lược đồ là CQI băng rộng (một giá trị CQI duy nhất được truyền cho toàn bộ phổ) thì độ nhảy kênh sẽ trở nên hoàn toàn vô dụng và không có lợi cho sự đa dạng của nhiều người dùng.
- *Năng lượng tiêu thụ:* Tiết kiệm năng lượng là một trong những tính năng bắt buộc của các thiết bị di động đầu cuối, nó đạt được thông qua các phương pháp tiếp nhận không liên tục (DRX) trong LTE. Ý tưởng của phương pháp này là cho phép thiết bị người dùng tiết kiệm năng lượng bằng cách tắt thiết bị vô tuyến của nó khi không truyền dữ liệu. Về nguyên tắc, nếu mạng có thể dự đoán khi nào một người dùng cần truyền/ nhận dữ liệu, thì một UE sẽ chỉ cần thức tại các khoảng thời gian cụ thể. Tuy nhiên, trong trường hợp không có sự kết hợp chính xác giữa bộ lập lịch và DRX, có thể xảy ra trường hợp tài nguyên được phân bổ cho một UE trong thời gian nó không hoạt động dẫn đến mất dữ liệu và lãng phí băng thông [8].

Trong phần tiếp theo, luận án sẽ trình bày các nghiên cứu liên quan đến lập lịch trong mạng LTE làm cơ sở để đề xuất thuật toán mới khắc phục các nhược

điểm mà các thuật toán hiện nay đang gặp phải.

### 1.2.3 Các nghiên cứu liên quan về lập lịch trong mạng LTE

**Bảng 1.1.** Các ký hiệu sử dụng cho mô hình lập lịch trong mạng di động đa dịch vụ và mô tả

| STT | Ký hiệu                   | Giải nghĩa   |
|-----|---------------------------|--|
| 1   | $m_{i,k}$                 | Giá trị đo của người dùng thứ $i$ tại khối tài nguyên thứ $k$                                |
| 2   | $r^i(t)$                  | Tỷ lệ dữ liệu nhận được bởi người dùng thứ $i$ tại thời điểm $t$                             |
| 3   | $\overline{R^i}(t)$       | Trung bình thông lượng nhận được của người dùng thứ $i$ trong quá khứ tính đến thời điểm $t$ |
| 4   | $\overline{R_{sch}^i}(t)$ | Thông lượng trung bình nhận được bởi luồng dữ liệu của người dùng thứ $i$ khi được lập lịch  |
| 5   | $D_{HOL,i}$               | Head of Line Delay, độ trễ của gói đầu tiên được gửi tới người dùng thứ $i$                  |
| 6   | $\tau_i$                  | Ngưỡng độ trễ của người dùng thứ $i$   |
| 7   | $\delta_i$                | Tỷ lệ lỗi chấp nhận được của người dùng thứ $i$  |
| 8   | $d^i(t)$                  | Tỷ lệ dữ liệu băng rộng mong đợi của người dùng thứ $i$ tại thời điểm $t$                    |
| 9   | $d_k^i(t)$                | Tỷ lệ dữ liệu mong đợi của người dùng thứ $i$ tại thời điểm $t$ trên khối tài nguyên thứ $k$ |
| 10  | $\Gamma_k^i$              | Phổ hiệu quả của người dùng thứ $i$ với khối tài nguyên thứ $k$                              |

Phần này sẽ trình bày một số chiến lược lập lịch cho mạng di động đa dịch vụ, làm nổi bật các ưu và nhược điểm của từng chiến lược. Các thuật toán lập lịch trong mạng LTE được phân loại thành (1) không nhận biết kênh; (2) nhận biết kênh/ không nhận biết chất lượng dịch vụ; (3) nhận biết kênh/ nhận biết

chất lượng dịch vụ. Các thuật toán lập lịch thường khác nhau về tham số đầu vào, mục tiêu lập lịch và dịch vụ đích.

Trước tiên, luận án trình bày một số phương pháp lập lịch không nhận biết kênh. Những thuật toán dạng này đã được sử dụng để xử lý vấn đề với tính cân bằng. Một số ký hiệu sử dụng trong luận án được trình bày trong bảng (1.1).

### A. Một số chiến lược lập lịch không nhận biết kênh

Được giới thiệu lần đầu cho mạng có dây với giả định về phương tiện truyền dẫn bất biến theo thời gian và không có lỗi, chúng thường được kết hợp với các chiến lược lập lịch nhận biết kênh để cải thiện hiệu suất hệ thống. Một số thuật toán không nhận biết kênh:

- i) *First In First Out (FIFO)*: Là thuật toán đơn giản nhất, thuật toán phân bổ tài nguyên tới người dùng theo thứ tự yêu cầu giống như hàng đợi FIFO. Khi đó, giá trị đo của người dùng thứ  $i$  tại khối tài nguyên thứ  $k$  là  $m_{i,k}^{FIFO}$  được mô tả:

$$m_{i,k}^{FIFO} = t - T_i \quad (1.2)$$

ở đó,  $t$  là thời điểm hiện tại,  $T_i$  là thời điểm mà người dùng thứ  $i$  gửi yêu cầu tài nguyên. Đây là một kỹ thuật đơn giản nhưng không hiệu quả và không cân bằng do nếu có một người dùng yêu cầu ít tài nguyên nằm ở cuối hàng đợi gồm những người dùng yêu cầu nhiều tài nguyên thì sẽ phải chờ đợi lâu.

- ii) *Round Robin (RR)*: Thuật toán Round Robin chia sẻ tài nguyên cân bằng giữa những người dùng. Giá trị đo cho người dùng thứ  $i$  được tính tương tự như trong thuật toán FIFO. Sự khác biệt ở chỗ giá trị  $T_i$  được coi là thời điểm mà lần cuối cùng người dùng được phục vụ. Thuật toán đảm bảo tính cân bằng cho người dùng nhưng không cân bằng về mặt

thông lượng do các người dùng yêu cầu các ứng dụng với tốc độ bit khác nhau và do đó, việc phân bổ cùng một lượng thời gian cho các người dùng là không hiệu quả.

- iii) *Blind Equal Throughput (BET)*: Thuật toán này sử dụng giá trị thông lượng trung bình đạt được trong quá khứ của người dùng làm thước đo [40] và do đó nó đạt được tính công bằng về thông lượng. Giá trị đo  $m_{i,k}^{BET}$  của thuật toán được tính:

$$m_{i,k}^{BET} = \frac{1}{\overline{R^i(t)}} \quad (1.3)$$

với

$$\overline{R^i(t)} = \beta \overline{R^i(t-1)} + (1-\beta)r^i(t) \quad (1.4)$$

trong đó, trọng số  $\beta$  nhận giá trị trong khoảng  $0 \leq \beta \leq 1$ ,  $\overline{R^i(t-1)}$  là thông lượng trung bình nhận được tại thời điểm  $t-1$ ,  $r^i(t)$  là tỷ lệ dữ liệu nhận được của người dùng thứ  $i$  tại thời điểm  $t$ .

Có thể thấy, BET cấp phát tài nguyên tới luồng đã được phục vụ có thông lượng trung bình thấp hơn trong quá khứ. Theo chính sách phân bổ này, người dùng với thông lượng trung bình trong quá khứ thấp hơn thực hiện quyền ưu tiên tài nguyên thì sẽ được cấp phát tài nguyên nếu người dùng này vẫn chưa đạt được thông lượng như những người dùng khác trong mạng. Nhân tố  $\overline{R^i}$  được cập nhật sau mỗi TTI cho phép đảm bảo rằng việc cấp phát tài nguyên được công bằng.

- iv) *Weighted Fair Queueing (WFT)*: Đưa ra các ưu tiên nhằm tránh tình trạng “đói tài nguyên” bằng sử dụng tiếp cận hàng đợi cân bằng có trọng số (WFQ). Theo cách này, trọng số  $w_i$  của người dùng thứ  $i$  (hoặc

lớp người dùng) được sử dụng để tính lại giá trị đo của RR:

$$m_{i,k}^{WFQ} = w_i * m_{i,k}^{RR} \quad (1.5)$$

trong đó,  $m_{i,k}^{RR}$  là giá trị đo của người dùng thứ  $i$  theo RR. Theo cách tiếp cận này, các nguồn lực tài nguyên được chia sẻ giữa các người dùng tương ứng với tỷ lệ giữa các trọng số (trọng số càng cao thì tài nguyên được phân bổ càng cao). Do RR kiểm soát thời gian chờ của người dùng nên hạn chế được tình trạng đói tài nguyên. Nói chung, cách tiếp cận này sẽ đúng nếu cả chỉ số BET được tính đến.

- v) *Guaranteed Delay*: Để tránh mất gói, dịch vụ đảm bảo trì hoãn yêu cầu mỗi gói phải được nhận trong một thời hạn nhất định. Mục tiêu này có thể đạt được bằng cách thêm tham số thời gian gồm thời gian tức thì khi gói được tạo và thời hạn của nó. Thời hạn sớm nhất trước tiên (EDF) và thời gian trễ có trọng số lớn nhất đầu tiên (LWDF) là hai thuật toán dạng này cho mạng có dây [50]. Thuật toán EDF, như tên gọi của nó, lập lịch gói tin với thời hạn cuối gần nhất. Giá trị đo  $m_{i,k}^{EDF}$  khi đó được tính:

$$m_{i,k}^{EDF} = \frac{1}{(\tau_i - D_{HOL,i})} \quad (1.6)$$

Có thể thấy, độ trễ của gói đầu tiên càng gần đến thời gian hết hạn thì giá trị đo càng tăng. Trong khi đó, giá trị đo của LWDF được tính dựa trên tham số hệ thống  $\delta_i$ , đại diện cho tham số chấp nhận được của người dùng thứ  $i$  mà ở đó một gói bị mất do hết hạn. Giá trị đo này được tính:

$$m_{i,k}^{LWDF} = \alpha_i * D_{HOL,i} \quad (1.7)$$

ở đó,  $\alpha_i$  được tính:

$$\alpha_i = -\frac{\log \delta_i}{\tau_i} \quad (1.8)$$

Giả sử hai luồng có trễ gói đầu bằng nhau, tham số  $\alpha_i$  sẽ làm tăng giá trị đo sao cho người dùng có yêu cầu cao nhất về tỷ lệ mất gói và thời gian hết hạn chấp nhận được sẽ được lên lịch phân bổ tài nguyên.

### B. Một số chiến lược lập lịch nhận biết kênh/ không nhận biết chất lượng dịch vụ

Bằng cách sử dụng các bản tin điều khiển đặc biệt, bộ lập lịch có thể tính toán chất lượng kênh mà mỗi UE nhận được dựa trên các phản hồi CQI được gửi định kỳ từ UE đến eNodeB. Do đó, bộ lập lịch có thể dự đoán thông lượng tối đa mà hệ thống nhận được. Gọi  $d^i(t)$  và  $d_k^i(t)$  tương ứng là các thông lượng có thể đạt được của người dùng thứ  $i$  tại khoảng thời gian truyền thứ  $t$  và tại khối tài nguyên thứ  $k$ . Các thuật toán nhóm này được mô tả như sau:

- i) *Maximum Throughput (MT)*: Với mục đích tối đa hóa dung lượng tổng thể, thuật toán MT gán từng RB cho người dùng có khả năng đạt thông lượng tối đa tại mỗi TTI. Giá trị đo  $m_{i,k}^{MT}$  của thuật toán được tính:

$$m_{i,k}^{MT} = d_k^i(t) \quad (1.9)$$

Dễ dàng nhận thấy, thuật toán MT cho phép hệ thống đạt thông lượng tối đa. Tuy nhiên, thuật toán thực hiện phép phân bổ tài nguyên không cân bằng đối với các người dùng có điều kiện kênh truyền thấp dẫn đến tình trạng đói tài nguyên. Bộ lập lịch thực tế phải kết hợp giữa MT và BET để đảm bảo phân phối tài nguyên cân bằng giữa những người dùng.

- ii) *Proportional Fair (PF)*: Thuật toán lập lịch PF đảm bảo tính công bằng giữa những người dùng. Giá trị đo  $m_{i,k}^{PF}$  của thuật toán được tính:

$$m_{i,k}^{PF} = m_{i,k}^{MT} * m_{i,k}^{BET} = d_k^i(t) * \frac{1}{R^i(t-1)} = \frac{d_k^i(t)}{R^i(t-1)} \quad (1.10)$$

Thuật toán sử dụng thông lượng trung bình trong quá khứ như một tham số để tính tỷ lệ dữ liệu dự kiến. Theo cách này, người dùng có điều kiện kênh không tốt vẫn được cấp phát tài nguyên trong một thời gian xác định. Lúc này, trọng số  $\beta$  trong công thức  $\overline{R}^i(t) = \beta \overline{R}^i(t-1) + (1-\beta)r^i(t)$  (công thức 1.4) trở nên quan trọng do có liên quan đến cửa sổ thời gian  $T_f$ , theo đó, thông lượng hệ thống được cải thiện. Theo đó:

$$T_f = \frac{1}{(1-\beta)} \quad (1.11)$$

Rõ ràng thấy, với  $\beta = 0$ , thông lượng trung bình trong quá khứ bằng tỷ lệ tức thời cuối cùng. Lúc này  $T_f = 1$  TTI. Mặt khác, khi  $\beta$  tiến gần tới 1, tỷ lệ nhận được cuối cùng sẽ không bao giờ được đưa vào để tính thông lượng trong quá khứ và khi đó,  $T_f$  về mặt lý thuyết sẽ tiến tới vô hạn.

Một mở rộng của thuật toán PF là thuật toán lập lịch cân bằng tổng quát (GPF) được giới thiệu trong [109], với giá trị đo  $m_{i,k}^{GPF}$  của thuật toán được tính:

$$m_{i,k}^{GPF} = \frac{[d_k^i(t)]^\xi}{[\overline{R}^i(t-1)]^\psi} \quad (1.12)$$

Trong công thức (1.12), vai trò của  $\xi$  và  $\psi$  là sửa đổi chính sách phân bổ dữ liệu tức thời và thông lượng trong quá khứ. Trong trường hợp này, nếu  $\xi = 0$  giá trị đo của thuật toán GPF bằng giá trị đo của thuật toán BET, có nghĩa là tính cân bằng đối với người dùng có thể nhận được phụ thuộc vào chất lượng kênh truyền. Mặt khác, nếu  $\psi = 0$  thì thuật toán GPF trở thành thuật toán MT.

- iii) *Throughput to Average (TTA)*: Thuật toán TTA được xem là trung gian giữa thuật toán MT và thuật toán PF [40]. Giá trị đo  $m_{i,k}^{TTA}$  của thuật toán được tính:

$$m_{i,k}^{TTA} = \frac{d_k^i(t)}{d^i(t)} \quad (1.13)$$

Ở đây, thông lượng có thể đạt được trong TTI hiện tại được sử dụng làm hệ số chuẩn hóa của thông lượng có thể đạt được trên RB đang xét. Thuật toán TTA mang đến tính công bằng cao trong mỗi TTI. Từ giá trị đo của nó, có thể thấy rằng thông lượng mong đợi tổng của người dùng càng cao thì chỉ số của nó trên một khối tài nguyên càng thấp. Điều này có nghĩa là bộ lập lịch khai thác các điều kiện về kênh để đảm bảo mức dịch vụ tối thiểu cho mọi người dùng.

iv) *Joint Time và Frequency Domain*: Một kỹ thuật hai bước được giới thiệu trong [75]:

- Trước tiên, bộ lập lịch gói theo miền thời gian (TDPS) lựa chọn một tập người dùng trong TTI hiện tại;
- Sau đó, các khối tài nguyên được phân bổ tới người dùng bằng bộ lập lịch theo miền tần số (FDPS).

Quyết định phân bổ tài nguyên cuối cùng là kết quả của hai bộ lập lịch (trong miền thời gian và trong miền tần số). Do số lượng người dùng được lựa chọn ít hơn số người dùng thực tế nên độ phức tạp tính toán tại FDPS được giảm xuống. Tại mỗi thời điểm, thuật toán FDPS có thể sử dụng RR hoặc PF để tăng sự công bằng về tài nguyên thời gian công bằng về hiệu quả phổ giữa các người dùng.

### C. Một số chiến lược lập lịch nhận biết kênh/ nhận biết chất lượng dịch vụ

Sự khác biệt chất lượng dịch vụ được quản lý bằng cách kết hợp một tập các tham số chất lượng dịch vụ tại mỗi luồng. Biết được giá trị của các tham số này, bộ lập lịch có thể xử lý dữ liệu sao cho đảm bảo một số yêu cầu về hiệu suất tối thiểu, về tốc độ dữ liệu hoặc về độ trễ phân phối. Tuy nhiên, nhận biết chất lượng dịch vụ nhưng không nhất thiết có nghĩa là cung cấp chất

lượng dịch vụ do việc quyết định phân bổ tài nguyên tùy thuộc vào yêu cầu của từng luồng, mà không nhất thiết phải đảm bảo đáp ứng các yêu cầu đó.

i) *Bộ lập lịch đảm bảo tốc độ dữ liệu*: Một giải pháp nhận biết chất lượng dịch vụ yêu cầu đảm bảo tỷ lệ dữ liệu được giới thiệu trong [60]. Thuật toán làm việc trong cả miền thời gian và miền tần số. Theo miền thời gian, bộ lập lịch Priority Set Scheduler (PSS) lựa chọn người dùng có tốc độ bit thấp tạo thành tập hợp có độ ưu tiên cao, những người dùng còn lại, có độ ưu tiên thấp hơn. Tập những người dùng này sẽ được lên lịch bởi BET và PF theo thứ tự. Khi một tập người dùng được lựa chọn bởi TDPS, FDPS cấp phát tài nguyên khả dụng theo PF ( $PF_{sch}$ ). Giá trị đo  $m_{i,k}^{PF_{sch}}$  của thuật toán được tính:

$$m_{i,k}^{PF_{sch}} = \frac{d_k^i(t)}{R_{sch}^i(t-1)} \quad (1.14)$$

ở đó  $\overline{R_{sch}^i}(t-1)$  là thông lượng trung bình trong quá khứ tương tự trong công thức (1.4) với sự khác biệt ở chỗ giá trị này chỉ được cập nhật khi người dùng thứ  $i$  được phục vụ.

Một cách tiếp cận tương tự được đưa ra trong [119]. Tập người dùng ưu tiên được chọn dựa trên giá trị QCI và được phân loại thành tập đảm bảo tốc độ bit (GBR) và không đảm bảo tốc độ bit (non-GBR). Sau bước này, FDPS theo thứ tự gán các khối tài nguyên tốt nhất đến người dùng trong nhóm GBR và cập nhật tốc độ bit nhận được. Khi tất cả người dùng trong danh sách đạt được tốc độ bit mong muốn và các khối tài nguyên vẫn còn, bộ lập lịch sẽ gán chúng cho người dùng trong tập non-GBR sử dụng thuật toán PF.

Các ưu tiên được sử dụng trong [95] được tính cho người dùng thứ  $i$ :

$$P_i = \frac{D_{HOL,i}}{\tau_i} \quad (1.15)$$

Trong trường hợp này, với giá trị  $P_i$  càng cao thì việc truyền gói tin HOL trở nên cấp thiết hơn. Tài nguyên sau đó được phân bổ tới người dùng có mức độ ưu tiên cao nhất để đạt đến tốc độ bit được đảm bảo. Việc phân bổ này sẽ được thực hiện với người dùng có mức độ ưu tiên cao thứ hai, v.v. Vào cuối quá trình lập lịch, mọi tài nguyên còn lại được phân bổ cho người dùng theo thứ tự ưu tiên được lặp lại.

Rõ ràng nhận thấy, cách tiếp cận sử dụng các danh sách có thứ tự các ưu tiên với các luồng trễ nhất giúp nâng cao hiệu quả phổ trong miền tần số.

- ii) *Bộ lập lịch đảm bảo trễ*: Các chiến lược dạng này nhằm đảm bảo các gói tin được gửi tới người dùng trong thời gian cho phép. Chúng thường được áp dụng cho các ứng dụng thời gian thực như Video streaming hay truyền tiếng nói của con người qua mạng máy tính sử dụng bộ giao thức TCP/IP (VoIP).

Thuật toán Modified-LWDF (M-LWDF) [3] là một mở rộng của thuật toán LWDF. Giá trị đo của M-LWDF được tính theo công thức  $m_{i,k} = \max_i \{m_{i,k}\}$  (công thức 1.1). Việc xử lý các luồng không đòi hỏi thời gian thực và luồng đòi hỏi thời gian thực được thực hiện khác nhau. Với luồng không đòi hỏi thời gian thực có thể được lập lịch bởi bộ lập lịch PF còn với luồng đòi hỏi thời gian thực, giá trị đo  $m_{i,k}^{M-LWDF}$  của thuật toán được tính:

$$\begin{aligned} m_{i,k}^{M-LWDF} &= \alpha_i * D_{HOL,i} * m_{i,k}^{PF} \\ &= \alpha_i * D_{HOL,i} * \frac{d_k^i(t)}{R^i(t-1)} \end{aligned} \quad (1.16)$$

trong đó,  $D_{HOL,i}$  là độ trễ do nghẽn đầu luồng của gói tin và  $\alpha_i$  được tính theo công thức (1.8).

Có thể nhận thấy, M-LWDF sử dụng thông tin độ trễ tích lũy để xác định hành vi của PF, do đó, đảm bảo sự cân bằng tốt giữa hiệu quả phổ tần số và tính công bằng.

Trong [7] các tác giả điều chỉnh một quy tắc hàm mũ nổi tiếng trong các hệ thống OFDMA là Exponential/PF (EXP/PF). Như tên gọi của nó, EXP/PF tính đến các đặc tính của PF và hàm mũ của độ trễ đầu-cuối. Tương tự như M-LWDF, EXP/PF phân biệt giữa ứng dụng thời gian thực và luồng tốt nhất. Đối với các luồng thời gian thực, giá trị đo  $m_{i,k}^{EXP/PF}$  của thuật toán được tính:

$$m_{i,k}^{EXP/PF} = \exp\left(\frac{\alpha_i * D_{HOL,i} - \chi}{1 + \sqrt{\chi}}\right) * \frac{d_k^i(t)}{R^i(t-1)} \quad (1.17)$$

ở đó:

$$\chi = \frac{1}{N_{rt}} \sum_{i=1}^{N_{rt}} \alpha_i D_{HOL,i} \quad (1.18)$$

trong đó,  $N_{rt}$  là tổng số luồng tải xuống theo thời gian thực.

Trong trường hợp này, PF quản lý các luồng không theo thời gian thực. Cả hai thuật toán MLWDF và EXP/PF đều dựa trên giả định rằng xác suất loại bỏ các gói tin hoạt động là có thể chấp nhận được. Hơn nữa, thuật toán có tính đến chỉ số PF của mỗi luồng, điều này cho phép đảm bảo thông lượng và tính công bằng giữa các người dùng.

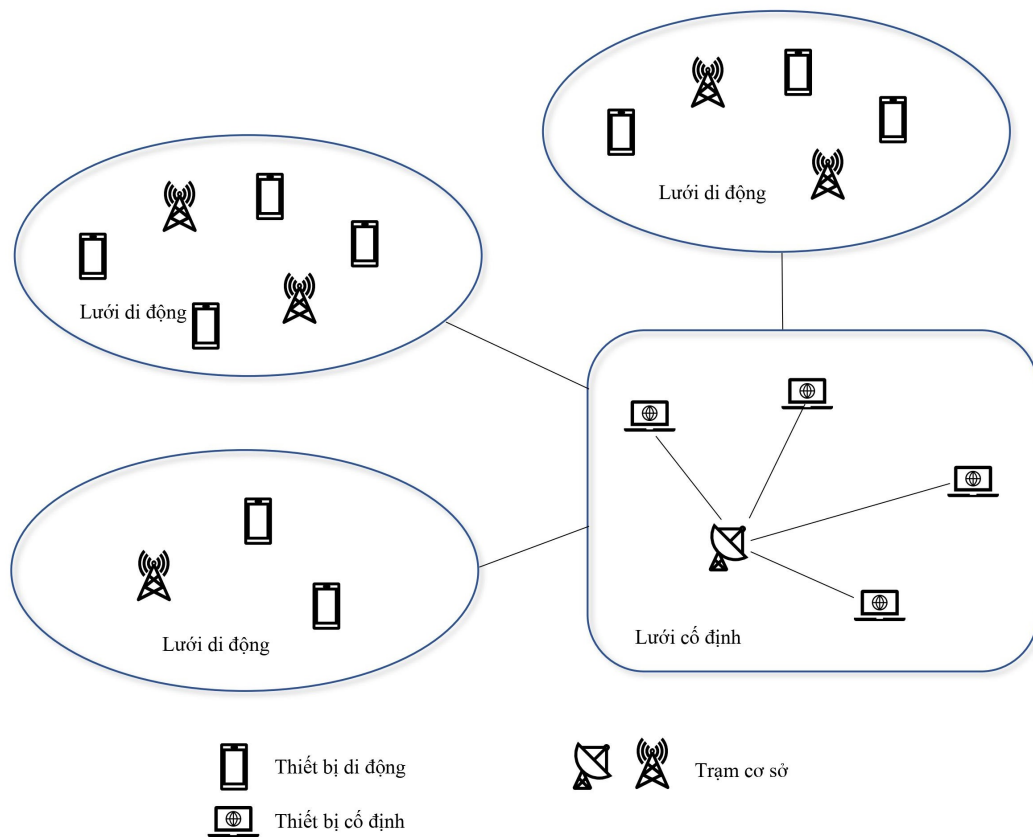
#### 1.2.4 Một số nhận xét

Các thuật toán lập lịch trong mạng LTE được phân loại thành: Không nhận biết kênh; Nhận biết kênh/ Không nhận biết chất lượng dịch vụ; Nhận biết kênh/ Nhận biết chất lượng dịch vụ. Các thuật toán lập lịch thuộc nhóm không nhận biết kênh thường đơn giản và không hiệu quả. Chúng thường được kết hợp với các thuật toán khác thuộc nhóm nhận biết kênh để tăng hiệu suất hệ thống. Nhận biết kênh là một khái niệm cơ bản nhằm đạt tới hiệu suất cao trong môi trường không dây, bằng cách khai thác các tính năng RRM như báo cáo CQI và thích ứng liên kết. Nếu chất lượng kênh nhận được của mỗi người dùng có thể ước lượng được thì tốc độ phân bổ tài nguyên có thể tăng lên đáng kể. Tham số thông lượng mong đợi  $d_k^i(t)$  được tính theo định lý Shannon  $d_k^i(t) = B * \log[1 + SINR_k^i(t)]$  là tham số quan trọng để đạt mục tiêu này. Tuy nhiên, hiệu suất phổ tần không phải là mục tiêu duy nhất đối với nhà khai thác mạng di động. Bên cạnh đó, thuật toán PF là cách tiếp cận tiêu biểu để làm tăng tính công bằng giữa những người dùng.

Việc phân phối các gói trước thời điểm hết hạn là một đặc điểm cơ bản đối với bộ lập lịch gói nhận biết QoS, đặc biệt khi xem xét trên các luồng đa phương tiện trong mạng di động đa dịch vụ. Đã có nhiều nghiên cứu được đề xuất, tuy nhiên chúng mới chỉ dừng ở bước thay đổi các thiết lập tham số mà mức độ tối ưu của nó phụ thuộc vào tình huống cụ thể. Hiệu năng của thuật toán được cải thiện khi có nhiều tham số tham gia vào quá trình lập lịch, tuy nhiên khi đó độ phức tạp của thuật toán sẽ tăng lên.

### 1.3 Tổng quan về lập lịch trong lưới tính toán di động

Ngày nay, lưới tính toán được sử dụng để giải quyết các bài toán đòi hỏi tính toán thông lượng cao với chi phí thấp. Lưới tính toán có khả năng tích lũy sức mạnh của các nguồn lực phân tán và không đồng nhất về mặt địa lý để tạo thành một nguồn lực liên kết để thực hiện giải quyết các bài toán ở mức độ cao hơn. Các ứng dụng đa chiều, đa lĩnh vực ngày nay cần nhiều tài nguyên với các đặc điểm đa dạng, tức là sức mạnh tính toán lớn, dung lượng lưu trữ lớn và khả năng truy cập dữ liệu nhanh.



Hình 1.4. Môi trường lưới tính toán di động

Trong phần này, luận án sẽ giới thiệu sơ lược về lưới tính toán di động, bài toán lập lịch trong lưới tính toán di động, các nghiên cứu liên quan và đưa ra

một số nhận xét làm động cơ nghiên cứu đề xuất thuật toán lập lịch mới trong lưới tính toán di động được trình bày ở chương 3. Môi trường lưới tính toán di động được mô tả như hình (1.4).

### 1.3.1 Sơ lược về điện toán lưới và lưới tính toán di động

Điện toán lưới (GC) chủ yếu dựa vào việc sử dụng các công nghệ điện toán hiện có và duy trì mức độ tương quan cao giữa người dùng lưới và các công nghệ lưới. Lưới được thừa hưởng số thuộc tính chuẩn hóa từ các hệ thống song song và phân tán truyền thống, đồng thời bổ sung nhiều đặc điểm mới là đặc trưng riêng của lưới. Chính vì lẽ đó, lưới duy trì mối quan hệ với các hệ thống phân tán truyền thống, nhưng vẫn bảo đảm sự vượt trội cho tính toán hiệu năng cao. Khái niệm điện toán lưới đã chuyển đổi phương thức tính toán so với các hệ thống phân tán truyền thống, tức là cách thức thực thi các ứng dụng và cách sử dụng tài nguyên.

Vào những năm 90 của thế kỷ trước, một tiêu chuẩn cho các ứng dụng tính toán hiệu năng cao xuất hiện, cùng thời điểm đó, các siêu máy tính có khả năng chia sẻ tài nguyên, đây được coi là thế hệ đầu tiên của lưới [45]. Ban đầu, lưới được giới thiệu dưới dạng nền tảng siêu máy tính phân tán để quản lý các vấn đề toán học chuyên sâu. Thuật ngữ “lưới” được sử dụng để tượng trưng cho một cơ sở hạ tầng máy tính phân tán được đề xuất cho các dự án khoa học và kỹ thuật tiên tiến [58].

Lưới đề cập đến các hệ thống và ứng dụng tích hợp tài nguyên và các dịch vụ được phân phối trên nhiều miền kiểm soát [32]. Nếu Internet là mạng truyền thông, thì điện toán lưới có thể được coi là mạng của tính toán. Lưới cung cấp một nền tảng thế hệ tiếp theo cho tính toán hiệu suất cao, khả năng truy cập

nhất quán, phổ biến, đáng tin cậy và minh bạch [5].

Lưới với sức mạnh của công nghệ Web có xu hướng cung cấp khả năng truy cập minh bạch và dễ dàng quản lý tài nguyên với khái niệm tổ chức ảo. Lưới ngày nay được phát triển để kết hợp khái niệm tính toán mềm nhằm làm cho hệ thống lưới trở nên thông minh và tự động hơn.

Ngày nay, các thiết bị di động trong lưới không chỉ thỏa mãn một số vai trò của lưới truyền thống như chia sẻ tài nguyên, quản lý tài nguyên [30] hay quản lý quy trình làm việc [29] mà còn đóng vai trò là các nút cảm biến hoặc thu thập dữ liệu mà các vai trò này chỉ có thể thực hiện được bởi các thiết bị di động.

Ban đầu, các thiết bị di động được tích hợp vào lưới như một giao diện người dùng để quản lý và giám sát tài nguyên lưới. Do các thiết bị di động đầu tiên có khả năng kết nối không dây với tài nguyên rất hạn chế, nên chúng không thể chia sẻ tài nguyên với lưới, và nếu có thì đóng góp của chúng cũng không nhiều.

Với mục đích quản lý và giám sát tài nguyên nên khi một thiết bị di động rời khỏi lưới thì tài nguyên lưới không bị suy giảm. Tuy nhiên, việc kết nối các thiết bị di động vào lưới cho phép người dùng thu thập tài nguyên lưới ở bất kỳ đâu có kết nối mạng. Ngoài ra, việc truy cập tài nguyên lưới từ các thiết bị di động để thực hiện các tác vụ tính toán rất khó thực hiện do thiếu tài nguyên. Mặc dù việc thực thi tác vụ không được thực hiện trên thiết bị di động, nhưng phần mềm trung gian của lưới tính toán đã che giấu các tiến trình khiến người dùng có cảm giác rằng tác vụ đang chạy trên thiết bị di động.

Theo thời gian, khả năng của thiết bị di động đã được cải thiện với tốc độ theo cấp số nhân [69]. Mạng không dây cũng đã trở nên nhanh hơn, đáng tin cậy hơn và khả dụng hơn. Ví dụ, mạng 4G luôn kết nối với Internet trong khi di chuyển qua các khu vực rộng lớn, chẳng hạn như thành phố hoặc đường cao tốc.

Ngoài ra, nhiều nơi công cộng, chẳng hạn như sân bay, khách sạn và nhà hàng được trang bị mạng WiFi riêng, rẻ hơn và nhanh hơn mạng 4G. Những thực tế đó thúc đẩy các thiết bị di động kết nối với hệ thống lưới. Các thiết bị di động hiện đại có một lượng tài nguyên sẵn có quan trọng, chúng có thể chia sẻ tài nguyên, chẳng hạn như CPU, bộ nhớ hoặc không gian lưu trữ hoặc chúng có thể cung cấp dịch vụ.

Tuy nhiên, các thiết bị di động không chỉ hạn chế về kết nối mà còn hạn chế về năng lượng. Do nguồn được cung cấp bị giới hạn bởi pin của chúng, các thiết bị di động kém tin cậy hơn đối với lưới so với các thiết bị cố định. Khi đó, giả định lưới truyền thông luôn có một thiết bị cụ thể kết nối từ cùng một vị trí cố định không còn đúng nữa. Do đó, các giải pháp công nghệ hoạt động cho các thiết bị cố định có thể không còn đúng cho các thiết bị di động.

Dù có nhiều lợi ích khi tích hợp các thiết bị di động vào lưới, nhưng những vấn đề gặp phải của lưới tính toán di động có thể kể đến như:

- Độ tin cậy của mạng và các thuật toán định tuyến [42].
- Lãng phí tài nguyên do khả năng khôi phục lỗi của các thiết bị di động [36].
- Tài nguyên như tốc độ CPU, bộ nhớ bị giới hạn [68].
- Ràng buộc về năng lượng [48].

### 1.3.2 Bài toán lập lịch trong lưới tính toán di động

Lập lịch công việc đóng một vai trò quan trọng trong các hệ thống quản lý luồng công việc. Đây là quá trình gán các nguồn lực phù hợp cho các tác vụ phụ thuộc của một ứng dụng đáp ứng một số mục tiêu do người dùng mong muốn. Mô hình lập lịch luồng công việc trên đám mây có thể được chia thành hai phần: (i) ánh xạ tác vụ và (ii) lập lịch.

- i) *Ánh xạ tác vụ*: Có  $m$  công việc cần lập lịch trong mạng biểu thị  $R = \{r_1, r_2, \dots, r_m\}$  và một ánh xạ  $MF = \Upsilon \rightarrow R$ , ở đó  $MF(i), 1 \leq i \leq n$  định nghĩa tài nguyên được gán cho  $v_i$ . Tài nguyên máy tính là tập hợp các máy ảo mà các máy vật lý được ảo hóa trên đó.
- ii) *Lập lịch*: Quá trình lập lịch được thực hiện bởi bộ lập lịch. Bộ lập lịch chọn tài nguyên phù hợp để thực thi các công việc. Hàm lập lịch cho một công việc được định nghĩa:  $S_j : \Upsilon \rightarrow \{0, 1, \dots, n\}$  ở đó  $S_j(i)$  với  $0 \leq i \leq n$  và  $1 \leq j \leq m$  cho biết thứ tự thực hiện công việc  $v_i$  trên tài nguyên  $r_j$ . Do đó, một hàm ánh xạ và một tập hàm lập lịch  $\{S_1, S_2, \dots, S_m\}$  cho biết ánh xạ các công việc tới các máy ảo trong điện toán đám mây.

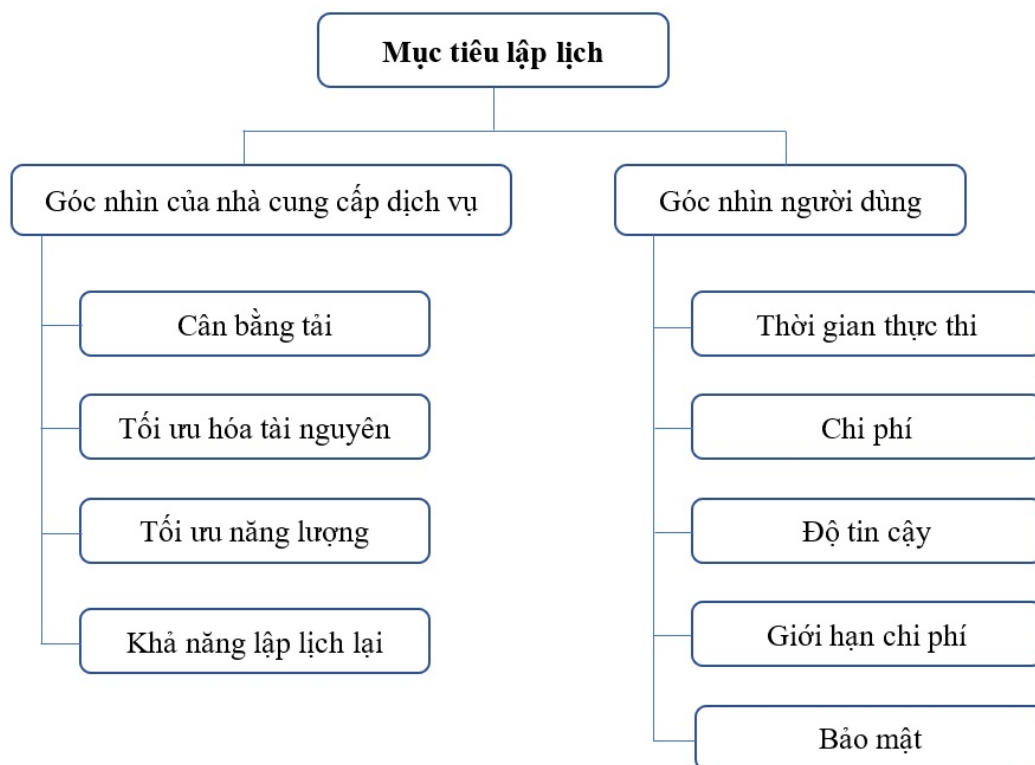
**Bảng 1.2.** Các ký hiệu sử dụng cho mô hình lập lịch trong lưới tính toán di động và mô tả

| STT | Ký hiệu         | Giải nghĩa   |
|-----|-----------------|--|
| 1   | $W(V, E)$       | Tập $ V $ công việc và tập $E$ cạnh có hướng   |
| 2   | $v_i$           | Công việc $v_i$ của ứng dụng   |
| 3   | $e_{ij}$        | Quan hệ giữa nhiệm vụ $v_i$ và nhiệm vụ $v_j$ trong đồ thị có hướng không có chu trình (DAG) |
| 4   | $D$             | Thời hạn thực thi luồng công việc  |
| 5   | $S_j(i)$        | Hàm lập lịch cho biết thứ tự thực thi của công việc $v_i$ trên tài nguyên $r_j$              |
| 6   | $C_{ij}^{Comp}$ | Chi phí tính toán  |
| 7   | $C_{ij}^T$      | Chi phí truyền dữ liệu giữa công việc $v_i$ và $v_j$   |
| 8   | $C_{ij}^{Stor}$ | Chi phí lưu trữ công việc $v_i$ trên tài nguyên $r_j$  |
| 9   | $Data_{(i,j)}$  | Số phần tử dữ liệu mà công việc $v_i$ cần để truyền tới công việc $v_j$                      |
| 10  | $FT(v_i)$       | Thời gian kết thúc công việc $v_i$   |
| 11  | $EFT(v_i)$      | Thời gian kết thúc sớm nhất của công việc $v_i$  |

Để làm được điều này, một số tiêu chí tối ưu hóa được tích hợp với quá trình lập lịch xem xét các tham số QoS được tuân theo. Độ dài lịch trình là một trong

những tham số thường được các nhà nghiên cứu sử dụng. Các ký hiệu sử dụng trong mô hình lập lịch ở phần này được đưa ra trong bảng (1.2).

### 1.3.3 Các mục tiêu lập lịch trong lưới tính toán di động



**Hình 1.5.** Các mục tiêu lập lịch trong MCG

Quá trình lập lịch phân bổ các công việc vào các tài nguyên phù hợp dựa trên các tiêu chí lập lịch nhất định. Các tham số lập lịch như thời gian thực thi, chi phí, độ tin cậy, giới hạn chi phí, bảo mật [84] có ảnh hưởng đến sự thành công của bài toán lập lịch. Các mục tiêu lập lịch này được phân loại theo khía cạnh định hướng dịch vụ thành hai quan điểm: nhà cung cấp dịch vụ và người sử dụng dịch vụ như hình (1.5). Có hai thực thể trong kịch bản lập lịch: cung cấp tài nguyên cho thuê và theo yêu cầu; và người dùng gửi nhiệm vụ để được xử lý

bởi các tài nguyên này. Phân loại các mục tiêu lập lịch được luận án trình bày cụ thể như sau:

***Dưới góc nhìn của nhà cung cấp dịch vụ***

- i) *Cân bằng tải*: Máy ảo là các phần tử xử lý chính trong môi trường điện toán đám mây. Trong bài toán lập lịch, có thể có nhiều hơn một công việc được gán cho các máy ảo để thực thi. Điều này dẫn đến mất cân bằng tải trên các máy ảo. Bộ lập lịch cần phải phân phối khối lượng công việc cho các tài nguyên có sẵn theo cách mà tài nguyên không bị quá tải. Cân bằng tải cải thiện việc sử dụng tài nguyên và do đó cải thiện hiệu suất lập lịch tổng thể. Có hai phương pháp cân bằng tải: cân bằng tĩnh và cân bằng động. Với cân bằng tải tĩnh, bộ lập lịch cần biết trước thông tin về các công việc cần lập lịch, trong khi đó, bộ lập lịch không cần biết trước các thông tin về công việc với cân bằng tải động. Cân bằng tải động rất hữu ích khi mục tiêu lập lịch là tối đa hóa việc sử dụng tài nguyên thay vì giảm thời gian thực hiện [90]. Rất nhiều phương pháp cân bằng tải đã được các nhà nghiên cứu thảo luận trong [33] dựa trên cách tiếp cận meta-heuristic. Ngoài ra, các mục tiêu lập kế hoạch khác như thời gian thực hiện, tối ưu năng lượng và tối ưu tài nguyên cũng được xem xét cùng với cân bằng tải [21, 93].
- ii) *Tối ưu tài nguyên*: Sử dụng tài nguyên hợp lệ và tối ưu đem lại nhiều lợi nhuận cho nhà cung cấp dịch vụ. Lập lịch công việc và tài nguyên nhằm tối đa hóa sử dụng tài nguyên được nêu ra trong [10].
- iii) *Tối ưu năng lượng*: Năng lượng tiêu thụ bởi một công việc là năng lượng xử lý tại CPU và tại các tài nguyên khác của hệ thống. Nếu CPU không được sử dụng đúng cách vì nguồn năng lượng nhàn rỗi không được sử dụng hiệu quả sẽ dẫn đến mức tiêu thụ năng lượng của hệ thống tăng cao. Đôi khi việc tiêu thụ nhiều năng lượng do nhu cầu lớn về tài nguyên dẫn đến giảm hiệu

suất [106]. Bộ lập lịch cần tìm ra một trình tự thực hiện công việc một cách hiệu quả để giảm tiêu thụ năng lượng của các tài nguyên được cấp phát [81]. Một trong những phương pháp để kiểm soát mức tiêu thụ năng lượng được nêu trong [63]. Một số thuật toán lập lịch cũng có thể được thiết kế để hỗ trợ việc giảm tiêu thụ năng lượng [91].

iv) *Khả năng lập lịch lại*: Việc lập lịch lại được coi là một quá trình lập kế hoạch chi phí vì nó dẫn đến việc đánh giá lại lịch trình và chi phí di chuyển dữ liệu giữa các công việc phụ thuộc trên các máy khác nhau [87]. Việc đổi lịch có thể xảy ra khi máy chủ quá tải. Một số trường hợp khác như lỗi máy ảo, lỗi hệ thống [113]. Khi thực hiện lập lịch lại, không phải mọi công việc đều được chọn để lập lịch lại vì sẽ làm tăng thời gian thực hiện tổng thể và làm giảm hiệu suất hệ thống. Sự cân bằng giữa chi phí lập lịch lại và hiệu suất được đưa ra trong [87] theo cách mà chỉ những nhiệm vụ được chọn mới được lập lịch lại.

#### ***Dưới góc nhìn của người sử dụng dịch vụ***

i) *Thời gian thực thi*: Là thời gian cần thiết để thực hiện toàn bộ quy trình công việc bằng cách xem xét thời gian khi các nhiệm vụ đã hoàn thành việc thực thi và thời gian khi nó đã được gửi lên hệ thống để bắt đầu quá trình lập lịch [111]. Hầu hết các thuật toán lập lịch đều tập trung vào việc tối ưu hóa thời gian [28]. Giảm tổng thời gian thực thi cũng làm giảm chi phí thực thi trong khi ảnh xạ các tác vụ với các tài nguyên là một trong các nhiệm vụ của bộ lập lịch. Thời gian thực thi có thể được định nghĩa là:

$$MS = \max_{v_i \in V} \{FT(v_i)\} \quad (1.19)$$

trong đó  $FT(v_i)$  là thời gian kết thúc của công việc  $v_i$ .

- ii) *Chi phí*: Chi phí của một ứng dụng phụ thuộc vào chi phí trên hai tài nguyên: chi phí tính toán  $C^{Comp}$ , chi phí truyền dữ liệu  $C^T$  và lưu trữ  $C^{Stor}$ . Chi phí này có thể được biểu diễn:

$$C(W) = \sum_{j=1}^m C_{r_j}^{Comp} + \sum_{(v_i, v_j)} C_{ij}^T + C_{ij}^{Stor} \quad (1.20)$$

Khi một tài nguyên được giao cho một số nhiệm vụ, thì dữ liệu được chuyển giữa các công việc và chi phí chuyển giao đó được tính như sau:

$$C_{ij}^T = \frac{Data_{(i,j)}}{B_{i,j}} \quad (1.21)$$

ở đó,  $Data_{(i,j)}$  kích thước dữ liệu đầu ra cho bởi nhiệm vụ  $v_i$  và  $B_{i,j}$  là băng thông trung bình giữa các máy ảo. Khi công việc cha  $v_i$  và công việc phụ  $v_j$  được lập lịch trên cùng một máy ảo thì chi phí truyền lúc này là  $C_{ij}^T$  sẽ bằng không.

- iii) *Độ tin cậy*: Các lỗi xảy ra trong quá trình lập lịch do nhiều nguyên nhân như lỗi hệ thống, không có sẵn tài nguyên hay lỗi mạng. Bộ lập lịch cần phải tính đến lỗi tài nguyên và thực thi thành công trong trường hợp có lỗi xảy ra. Độ tin cậy là xác suất các công việc chạy thành công và hoàn thành một quy trình. Đây là một mục tiêu để đảm bảo giảm nguy cơ thất bại của một quy trình công việc được giao.
- iv) *Giới hạn chi phí*: Là một ràng buộc cơ bản được người dùng xác định khi sử dụng các dịch vụ từ nhà cung cấp, sử dụng các ràng buộc về chi phí, bộ lập lịch đưa ra các quyết định để giảm thiểu thời gian hoàn thành quy trình công việc trong giới hạn chi phí cho trước [117]. Do đó, tổng chi phí thực hiện  $C(W)$  là tổng chi phí của các máy ảo khác nhau được sử dụng để thực thi luồng công việc với  $C(W) \leq B$ , ở đó,  $B$  là giới hạn chi phí. [77].

v) *Bảo mật*: Trong điện toán đám mây, các tài nguyên không đồng nhất và phân tán, vì vậy bảo mật trở thành một vấn đề cần được xem xét trong quá trình lập lịch. Các vấn đề bảo mật như rò rỉ dữ liệu cần phải được xem xét trong quá trình chia sẻ tài nguyên. Một kỹ thuật ảo hóa hiệu quả và an toàn cần được giải quyết trong mối quan tâm của bảo mật được đưa ra trong [24].

#### 1.3.4 Các nghiên cứu liên quan về lập lịch trong lưới tính toán di động

Do đặc điểm không đồng nhất của các thiết bị trong lưới di động, Do đặc điểm không đồng nhất của các thiết bị, các đặc tính khác nhau của các công việc, bài toán lập lịch công việc là bài toán thuộc lớp NP-đầy đủ, không có một phương pháp cụ thể nào để giải bài toán này trong thời gian đa thức. Do đó, các kỹ thuật dựa trên meta-heuristic được coi là một phương pháp nhằm tìm kiếm lời giải gần tối ưu [99]. Meta-heuristic là phương pháp được phát triển gần đây nhất trong các phương pháp tìm kiếm gần đúng để giải quyết các vấn đề tối ưu nảy sinh trong kỹ thuật, công nghiệp, và nhiều lĩnh vực khác. Meta-heuristic sử dụng các kỹ thuật có nguồn gốc từ trí tuệ nhân tạo, sinh học, toán học, tự nhiên để cải thiện hiệu suất tìm kiếm [66]. Nhiều nghiên cứu đã áp dụng các thuật toán meta-heuristic trong lập lịch công việc [20].

“Meta-heuristic” được Glover đưa ra như một chiến lược mở rộng của phương pháp heuristic để tìm các lời giải tốt hơn trong bối cảnh với tính tối ưu của tìm kiếm địa phương. Hầu hết các bài toán tối ưu trong thế giới thực là phi tuyến và đa phương thức với nhiều ràng buộc khác nhau. Ngay cả với một mục tiêu duy nhất, đôi khi, các giải pháp tối ưu có thể không tồn tại. Nhìn chung, việc tìm ra một giải pháp tối ưu hoặc thậm chí là các giải pháp gần tối ưu không phải là một việc dễ dàng. Khi đó, các phương pháp meta-heuristic mang lại sự thay thế tốt hơn giữa chất lượng giải pháp và thời gian tính toán. Ngoài ra, các

cách tiếp cận này có thể điều chỉnh được vì nó có thể thích nghi với bất kỳ vấn đề tối ưu thực tế nào về hiệu suất và thời gian thực hiện [96].

### Một số phương pháp meta-heuristic dựa trên cảm hứng sinh học

Với khao khát bất chước tự nhiên và mong muốn đưa tự nhiên vào trong công nghệ, điện toán lấy cảm hứng từ sinh học đã mang lại một kỷ nguyên mới trong tính toán. Đã có nhiều ứng dụng sử dụng điện toán lấy cảm ứng từ sinh học được áp dụng trong các hệ thống điều khiển, kỹ thuật y sinh, xử lý song song và khai thác dữ liệu. Thuật toán di truyền (GA), thuật toán Memetic (MA), thuật toán tối ưu bầy sư tử (LA) là một số thuật toán lấy cảm hứng từ sinh học đã được sử dụng để lập lịch luồng công việc trong điện toán đám mây.

- i) *Thuật toán di truyền (GA)*: Dựa trên nguyên lý của Darwin về sự “sống sót của những người khỏe mạnh nhất”, trong đó thể lực được cải thiện với quá trình tiến hóa thông qua sinh sản [94].

Thuật toán di truyền khai thác các giải pháp tốt nhất từ các tìm kiếm đã được thực hiện trước đó kết hợp với việc phân tích các vùng mới của không gian lời giải. Bất kỳ lời giải nào cho một vấn đề đều được biểu diễn bằng nhiễm sắc thể, bao gồm tập hợp các phần tử được gọi là gen.

Thuật giải bắt đầu với quần thể ngẫu nhiên của các lời giải và chất lượng quần thể được tìm thấy bằng cách đánh giá nó với hàm mục tiêu còn được gọi là hàm thích nghi. Sau khi khởi tạo quần thể với các lời giải ngẫu nhiên, các nhiễm sắc thể con mới được tạo ra bằng cách sử dụng biến đổi nhờ lai tạo hoặc đột biến [23].

Các thế hệ mới được tạo ra một lần nữa được đánh giá và điều chỉnh để thích ứng với quần thể nếu nó tạo ra một lời giải tốt hơn trước đó. Quá trình tạo ra con cái được lặp lại cho đến khi tìm được lời giải tối ưu [78].

ii) *Thuật toán Memetic (MA)*: Lấy cảm hứng từ cả các nguyên tắc tiến hóa tự nhiên của Darwin và khái niệm của Dawkins về meme, thuật ngữ memetic đã được Pablo Moscato đưa ra trong báo cáo kỹ thuật của mình [64] vào năm 1989, nơi ông xem MA gần giống với một dạng thuật toán di truyền lai dựa trên quần thể kết hợp với một thủ tục học cá nhân có khả năng thực hiện các sàng lọc cục bộ.

Thuật toán MA được cho là một phiên bản sửa đổi của thuật toán di truyền với tìm kiếm địa phương được thực hiện bằng cách sử dụng thuật toán leo đồi hoặc tìm kiếm tabu hoặc mô phỏng luyện thép. Về cơ bản, nó là kỹ thuật tối ưu toàn cục dựa trên quần thể mà ở đó tìm kiếm địa phương được thực hiện bởi mọi cá thể. Tuy nhiên, hoạt động của thuật toán MA khá giống với thuật toán di truyền, nhưng lời giải cho bài toán được đại diện bằng các chuỗi mã hóa dựa trên mô hình gọi là nhiễm sắc thể thay vì gen được sử dụng trong thuật toán di truyền.

Trong thuật toán memetic, tất cả các nhiễm sắc thể và con được truyền qua tìm kiếm địa phương trước khi tham gia vào quá trình tiến hóa [57]. Hiệu suất của thuật toán memetic phụ thuộc vào sơ đồ mã hóa được sử dụng. Ban đầu, các quần thể được khởi tạo ngẫu nhiên, và sau đó quần thể mới được hình thành dựa trên các cá nhân có chất lượng tốt hơn. Trong cơ chế tìm kiếm địa phương, các nhiễm sắc thể tốt nhất được chọn từ bộ nhiễm sắc thể ban đầu sau đó thực hiện lai ghép và đột biến để tạo ra thế hệ kế tiếp. Những con cái này sau đó thực hiện tìm kiếm cục bộ và cải tiến tối ưu cục bộ để đạt kết quả tốt hơn.

iii) *Thuật toán tối ưu bầy sư tử*: Vấn đề tối ưu hóa có thể được giải quyết bằng cách sử dụng hành vi xã hội của loài sư tử. “Đàn” là một thuật ngữ chỉ hệ thống xã hội của sư tử, ở đó sư tử con được sinh ra bởi sư tử bố, mẹ ở trong

đàn. Hành vi bảo vệ lãnh thổ và chiếm lĩnh lãnh thổ của sư tử đã được áp dụng để giải quyết vấn đề tối ưu hóa [82].

Bảo vệ lãnh thổ được thực hiện bởi các sư tử đực trong đàn và kẻ du mục. Quá trình này sẽ thay thế các sư tử đực trong đàn bằng giải pháp mới nếu nó tốt hơn giải pháp tồn tại. Trong hành vi chiếm lĩnh lãnh thổ, những sư tử đực và sư tử cái tốt nhất được giữ lại. Trong cả hai trường hợp, các giải pháp hiện có đều biến mất.

Giống như thuật toán di truyền, thủ tục lựa chọn được chọn để thực hiện hành vi chiếm lĩnh lãnh thổ, trong đó các giải pháp tốt hơn được giữ lại. Với quá trình bảo vệ lãnh thổ, thuật toán này bắt đầu sinh ra thế hệ mới để tạo ra lãnh thổ gồm cả con đực và con cái. Sau đó, các lời giải mới được đưa ra dựa trên các toán tử lai tạo, đột biến và phân nhóm giới tính. Cuối cùng, các lời giải hiện có được thay thế bằng lời giải mới dựa trên hành vi của sư tử là bảo vệ lãnh thổ và tiếp quản lãnh thổ.

Trí thông minh bầy đàn là một cách tiếp cận tương đối mới để giải quyết các vấn đề tối ưu hóa, được lấy cảm hứng từ trí thông minh tập thể của các quần thể sinh vật và của các động vật khác. Trí thông minh bầy đàn là một mô hình tính toán và hành vi giải quyết vấn đề phân tán dựa trên sự tương tác và xử lý thông tin của các cá nhân riêng lẻ.

Bầy đàn là từ chỉ sự đa dạng, phân phối và ngẫu nhiên. Thuật ngữ thông minh cho biết khả năng nhận thức, học tập và sáng tạo được sử dụng để đưa ra lời giải cho một vấn đề. Với các thuật toán này, thông tin thường được chia sẻ giữa nhiều bầy đàn để cùng tiến hóa và học cách tìm kiếm không gian lời giải. Một số thuật toán dựa trên trí thông minh bầy đàn như tối ưu đàn kiến (ACO), tối ưu bầy đàn (PSO) hay tối ưu bầy ong nhân tạo (ABC).

- i) *Tối ưu đàn kiến (ACO)*: Lấy cảm hứng từ hành vi tìm kiếm thức ăn của một số loài kiến, thuật toán được Marco Dorigo giới thiệu trong luận án tiến sĩ với tên gọi ban đầu là hệ thống kiến [19]. Những con kiến để lại mùi trên mặt đất để chỉ đạo những con kiến khác đi đúng đường. Hành vi như vậy có thể được sử dụng để giải quyết các vấn đề tối ưu hóa. Những con kiến nhân tạo đóng vai trò là tác nhân và chúng tìm ra giải pháp tối ưu bằng cách di chuyển qua một không gian các lời giải. Mùi là giá trị được sử dụng để tìm không gian lời giải này, những con kiến này ghi lại vị trí và chất lượng lời giải của chúng để nhằm tìm kiếm lời giải tối ưu. Cách tiếp cận dựa trên đàn kiến được sử dụng cho rất nhiều bài toán tối ưu như lập lịch job-shop hay bài toán chiếc ba lô đa chiều.
- ii) *Tối ưu hóa bầy đàn (PSO)*: Tối ưu hóa bầy đàn được lấy cảm hứng từ hành vi xã hội của các đàn như đàn chim, được phát triển bởi Eberhart và Kennedy [76] để giải các bài toán tối ưu hóa. Các đàn thay đổi đường đi của mình dựa trên vị trí tốt nhất của chính mình và vị trí tốt nhất của cá thể trong đàn của mọi thế hệ.

Quá trình này hội tụ nhanh chóng đến cực tiểu toàn cục với một lời giải tốt. Lời giải cho bất kỳ vấn đề tối ưu hóa nào được biểu diễn bằng một cá thể giống như một nhiễm sắc thể trong thuật toán di truyền. Kỹ thuật này bắt đầu với việc khởi tạo ngẫu nhiên quần thể các cá thể, trong đó mỗi cá thể được liên kết với vị trí và vận tốc riêng của chúng với mục đích xác định tốc độ và hướng bay.

Các cá thể phát triển hành vi xã hội của chúng và tiến tới mục tiêu bằng cách thay thế vị trí và vận tốc hiện tại của nó bằng vị trí và vận tốc của cá thể tốt nhất. Mỗi cá thể nhìn qua không gian tìm kiếm và tự cập nhật chúng từ những vị trí được biết đến nhiều nhất. Quá trình này lặp lại cho đến khi đáp ứng được lời giải mong muốn.

Thuật toán tối ưu bầy đàn được áp dụng trong bài toán lập lịch tác vụ trong môi trường lưới [51, 120]

- iii) *Tối ưu bầy ong nhân tạo*: Thuật toán bầy ong nhân tạo được lấy cảm hứng từ hành vi kiếm ăn thông minh của ong mật nhằm kiếm nguồn thức ăn. Cách tiếp cận này được Dervis Karaboga đưa ra vào năm 2005 để giải quyết các vấn đề tối ưu trong thế giới thực [100].

Trong đàn ong, tất cả các ong thợ cùng hợp tác để tìm kiếm nguồn thức ăn và thông tin này được sử dụng để đưa ra quyết định khám phá không gian tìm kiếm. Sự di chuyển của đàn ong được chia sẻ bởi mỗi con ong bằng cách chia chúng thành ba nhóm: ong thợ, ong quan sát và ong trinh sát như trong tự nhiên. Mỗi nhóm có một chức năng duy nhất.

Một con ong thợ tiếp tục tìm kiếm nguồn thức ăn mới trong không gian lân cận cho đến khi nó tìm được một nguồn thức ăn nhiều hơn nguồn hiện tại. Những con ong quan sát chia sẻ lượng mật hoa của nguồn mới với những con ong khác bằng cách thực hiện điệu nhảy lắc lư và đánh giá chất lượng nguồn thức ăn bằng phương pháp xác suất, ví dụ, kỹ thuật bánh xe roulette. Dựa trên chất lượng nguồn thức ăn mới được đánh giá này, ong quan sát sẽ cố gắng tìm kiếm nguồn thức ăn mới với lượng mật hoa cao hơn. Với bất kỳ bài toán tối ưu nào, nguồn thức ăn được sử dụng đại diện cho lời giải của bài toán, lượng mật hoa của nguồn sẽ đại diện cho chất lượng của giải pháp. Thuật toán này được sử dụng để giải quyết nhiều vấn đề kết hợp như bài toán lập lịch lưu lượng [53], lập lịch các dự án [121], người bán hàng [35].

### 1.3.5 Một số nhận xét

Bài toán lập lịch công việc là bài toán thuộc lớp NP-Đầy đủ, không có một phương pháp cụ thể nào để giải bài toán này trong thời gian đa thức. Các

kỹ thuật dựa trên meta-heuristic được coi là giải pháp tối ưu để giải quyết các bài toán dạng này. Các tham số chất lượng dịch vụ được sử dụng làm tiêu chí lập lịch gồm: thời gian thực thi; chi phí; cân bằng tải; tối ưu tài nguyên và một số tham số khác.

Việc sử dụng các kỹ thuật meta-heuristic vẫn còn ở giai đoạn đầu vì các nghiên cứu gồm GA, PSO, ABC và ACO là các kỹ thuật được sử dụng chủ yếu không hiệu quả nhiều về mặt mô tả sự phù hợp của các thao tác lập lịch công việc trên diện toán lưới và diện toán đám mây. Lựa chọn một phương pháp meta-heuristic cho bài toán lập lịch thường không dễ dàng, chúng phụ thuộc nhiều vào sự lựa chọn tiêu chí và kiến thức chuyên môn của người dùng.

#### 1.4 Kết luận chương

Trong chương này, luận án đã tập trung trình bày các tổng quan nghiên cứu và cơ sở lý thuyết hệ thống tính toán phân tán và bối cảnh lập lịch. Trong hệ các hệ thống tính toán phân tán, hai thành phần quan trọng đảm bảo hiệu năng của hệ thống là quản trị dữ liệu và truyền thông. Dữ liệu được thu thập bởi các thành phần cố định hoặc di động và gửi về trung tâm dữ liệu để phân tích và ra quyết định. Truyền thông được thực hiện ở mọi thành phần của hệ thống và giao tiếp với đám mây thông qua mạng.

Phần này đã trình bày tổng quan về bài toán lập lịch trong mạng di động đa dịch vụ và các nghiên cứu liên quan. Từ đó rút ra các nhận xét làm động cơ nghiên cứu đề xuất một thuật toán lập lịch mới tích hợp giữa miền thời gian và miền tần số được trình bày trong chương 2. Tổng quan về diện toán lưới và lưới tính toán di động cũng được trình bày trong phần này. Từ các nghiên cứu về lưới tính toán di động và bài toán lập lịch đã nghiên cứu, tác giả đã rút ra các nhận

xét làm động cơ nghiên cứu đề xuất một thuật toán lập lịch mới trong lưới tính toán di động dựa trên phương pháp meta-heuristic được trình bày ở chương 3.

Kết quả nghiên cứu của chương đã được công bố tại [CT1] trong danh mục các công trình khoa học của tác giả có liên quan đến luận án.

## Chương 2

# ĐỀ XUẤT GIẢI PHÁP LẬP LỊCH TRONG MẠNG DI ĐỘNG ĐA DỊCH VỤ

Ngành công nghiệp 4.0 đã trở thành hiện thực với sự hỗ trợ của các kỹ thuật tiên tiến như trí tuệ nhân tạo (AI), điện toán phân tán, giao tiếp trực tiếp giữa các thiết bị (M2M). Trong mạng di động tùy biến (MANET) và lưới tính toán di động (MCG), các thiết bị giao tiếp qua các phương tiện di động truyền thông không dây [13, 86, 114].

Dự án đối tác thế hệ thứ 3 (3GPP) đã giới thiệu công nghệ truyền thông thế hệ thứ tư (4G) – LTE cho phép hỗ trợ nhiều người dùng trong đường xuống với điều chế lên đến 64 QAM [107]. Đây là phiên bản đa người dùng của sơ đồ điều chế kỹ thuật số ghép kênh phân chia theo tần số trực giao (OFDM) phổ biến được gọi là đa truy cập phân chia theo tần số trực giao (OFDMA).

OFDMA được sử dụng trong LTE có ưu điểm là băng thông có thể mở rộng và khả năng chống lại hiện tượng đa đường. Với tốc độ dữ liệu đường xuống có thể lên đến 300 Mb/s bằng cách sử dụng băng thông trong khoảng từ 5 đến 20 MHz [34], mạng di động đa dịch vụ cung cấp một lượng lớn ứng dụng với băng

thông lớn, độ trễ thấp đáp ứng ngày càng cao các yêu cầu chất lượng dịch vụ của người dùng.

Công nghệ 5G là thế hệ tiếp theo của 4G được đề xuất làm tiêu chuẩn cho mạng di động thế hệ mới. Mạng 5G cung cấp băng thông rộng, cho phép tải nhanh hơn, và cung cấp nền tảng mạng tốt hơn trong ngành công nghiệp 4.0, tuy nhiên, hiện nay mới chỉ có một số các quốc gia đã triển khai thành công mạng 5G, trong khi số còn lại vẫn đang sử dụng 4G hoặc phiên bản tiền nhiệm của nó, do đó, 4G LTE vẫn có cơ hội lớn trong ngành công nghiệp 4.0 [101].

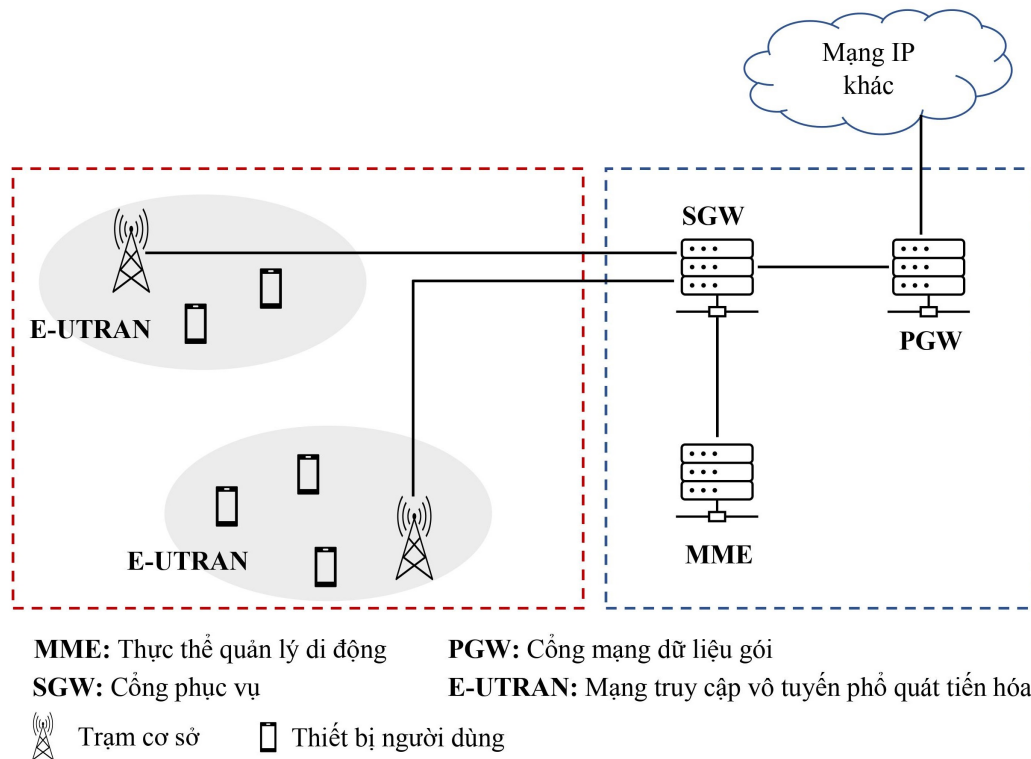
Chương này trình bày thuật toán đề xuất kết hợp giữa miền thời gian và miền tần số nhằm cải thiện chất lượng dịch vụ như giảm tỷ lệ mất gói, giảm tỷ lệ trễ gói, tăng tính công bằng, tăng băng thông hệ thống. Kết quả của thuật toán đề xuất được so sánh với một số phương pháp liên quan như QuAS và PSS để kiểm chứng. Phương pháp được đề xuất sẽ là cơ sở hạ tầng cho việc phát triển các phương pháp mới áp dụng cho mạng 5G trong tương lai.

Long Term Evolution viết tắt là LTE hay tiến hóa dài hạn, có thể nói là một thế hệ trung gian giữa 3G và 4G, được thiết kế làm tiền đề nhằm hiểu rõ hơn về 4G và các tính năng phải được phát triển trong 4G. Các tính năng có thể kể đến là tăng tốc độ dữ liệu, tránh nhiễu, cải thiện khả năng kết nối cuộc gọi và cung cấp dịch vụ tới nhiều người dùng hơn. Các tính năng trên thường được sử dụng làm tham số chất lượng dịch vụ.

Phân bổ tài nguyên là một yếu tố quan trọng góp phần làm tăng chất lượng dịch vụ của hệ thống. Phân bổ tài nguyên là quá trình cung cấp các tài nguyên cần thiết cho những người sử dụng có nhu cầu. Hiệu quả phân bổ tài nguyên là một yếu tố đánh giá chất lượng dịch vụ. Hiệu quả ở đây được định nghĩa theo tốc độ, hệ số tiện ích và công suất hệ thống. Nhiều thuật toán đã được phát triển

dựa trên các thuật toán gốc nhằm giúp phân bổ tài nguyên nhanh hơn. Một số thuật toán quan trọng là Max Rate, Round Robin, Proportional Fair, Modified Weighted Largest Delay First, Exp-Rule [39] (xem tại Mục 1, Chương 1).

## 2.1 Giới thiệu



**Hình 2.1.** Kiến trúc dịch vụ trong mạng LTE

Trong kiến trúc dịch vụ mạng LTE được trình bày tại hình (2.1), gồm các thành phần: (1) thiết bị người dùng; (2) mạng truy cập vô tuyến mặt đất phổ quát tiến hóa; (3) mạng lõi chuyển mạch gói. Chức năng của các thành phần này như sau:

- i) *Thiết bị người dùng (UE)*: Thiết bị di động là điểm đầu cuối chính trong các mạng di động, tương tác với các trạm gốc eNodeB thông qua tín hiệu vô tuyến, nhằm mục đích gửi và nhận thông tin. Một UE bao gồm: UICC chứa

các khóa mã bí mật được chia sẻ trước với nhà khai thác di động (Mobile Network Operator - MNO) trước khi nó được cấp tới người dùng và IMSI là định danh cố định mà nhà mạng sử dụng để xác định thuê bao di động.

- ii) *E-UTRAN* là một khối gồm nhiều eNodeB liên kết với nhau qua giao diện (giao diện X2), eNodeB được giới thiệu trong LTE nhằm giảm độ trễ sóng vô tuyến và giao tiếp với mạng lõi chuyển mạch gói thông qua giao diện S1.
- iii) *Mạng lõi chuyển mạch gói (EPC)* là mạng kết hợp công phục vụ, công mạng dữ liệu gói, thực thể quản lý di động và máy chủ thuê bao thường trú. EPC có vai trò là cầu nối giữa người dùng và mạng ngoài bao gồm hệ thống mạng con đa phương tiện dựa trên IP (IMS).

- + Máy chủ quản lý thuê bao thường trú (HSS): Hoạt động như một sổ đăng ký thông tin gồm tất cả các thông tin tỷ lệ dữ liệu yêu cầu bởi người dùng, khóa bảo mật của người dùng. HSS cũng hỗ trợ quản lý tính di động người dùng.
- + Cổng dịch vụ (SGW): Đây là phần tử mạng cung cấp kết nối eNodeB và mạng lõi chuyển mạch gói. SGW phục vụ thiết bị người dùng bằng cách định tuyến các gói IP (giao thức Internet) đến và đi. SGW hoạt động như một điểm neo cho thiết bị di động nội bộ bên trong LTE, tức là cho quá trình chuyển giao diễn ra giữa hai eNodeB. SGW kết nối với một cổng khác là cổng PDN.
- + Cổng mạng dữ liệu gói (PDN): Cổng PDN hoạt động như một kết nối giữa mạng lõi chuyển mạch gói và mạng IP bên ngoài. Cổng PDN cung cấp cho các gói tin một đường dẫn đến và đi từ các mạng dữ liệu công cộng. PDN cũng thực hiện phân bổ địa chỉ IP cho người dùng và tính phí cho người dùng tùy theo sử dụng.

- + Thực thể quản lý di động (MME): Thực thể mạng này xử lý tín hiệu liên quan đến tính di động của người dùng. MME cung cấp tính năng bảo mật cho truy cập E-UTRAN. MME hoạt động như một bộ theo dõi để theo dõi thiết bị của người dùng và phân trang nó ở chế độ nhàn rỗi.

Đường tải xuống liên quan đến dữ liệu được gửi đi trạm gốc đến máy thu hoặc thiết bị người dùng. E-UTRAN được giới thiệu trong LTE hoạt động như một giao diện giữa trạm cơ sở và thiết bị người dùng. Trạm cơ sở (eNodeB) chịu trách nhiệm phân bổ tài nguyên cho người dùng sử dụng dịch vụ. Các yếu tố chất lượng dịch vụ như độ trễ thời gian, tốc độ bit, v.v. phụ thuộc vào hiệu quả của thuật toán được sử dụng để lập lịch cho người dùng sử dụng tài nguyên tại eNodeB. Kiến trúc dịch vụ trong LTE được mô tả như hình (2.1).

Một số thuật toán đã được đề xuất và triển khai để lập lịch phân bổ tài nguyên. Một thuật toán lập lịch hiệu quả phải đáp ứng được các yêu cầu sau [83]:

- i) Cấp phát tài nguyên tới người dùng với độ trễ tối thiểu.
- ii) Dịch vụ được cung cấp tới người dùng ngay cả khi họ ở biên vùng phủ sóng của trạm cơ sở.
- iii) Cung cấp dịch vụ cho tất cả người dùng mà không có bất kỳ sự suy giảm nào về chất lượng.
- iv) Tối thiểu lãng phí tài nguyên.

Chất lượng dịch vụ (QoS) đề cập đến việc đo lường hiệu suất tổng thể của một dịch vụ mà người dùng mạng trải nghiệm. Chất lượng dịch vụ được thể hiện qua các tham số như thông lượng, tính công bằng, độ trễ gói, tỷ lệ mất gói v.v. ở đó:

- i) *Thông lượng*: Thông lượng mạng đề cập đến số lượng dữ liệu có thể được truyền từ nguồn tới đích trong một khoảng thời gian, được tính bằng tốc độ truyền gói tin thành công và tốc độ tối đa của kênh.

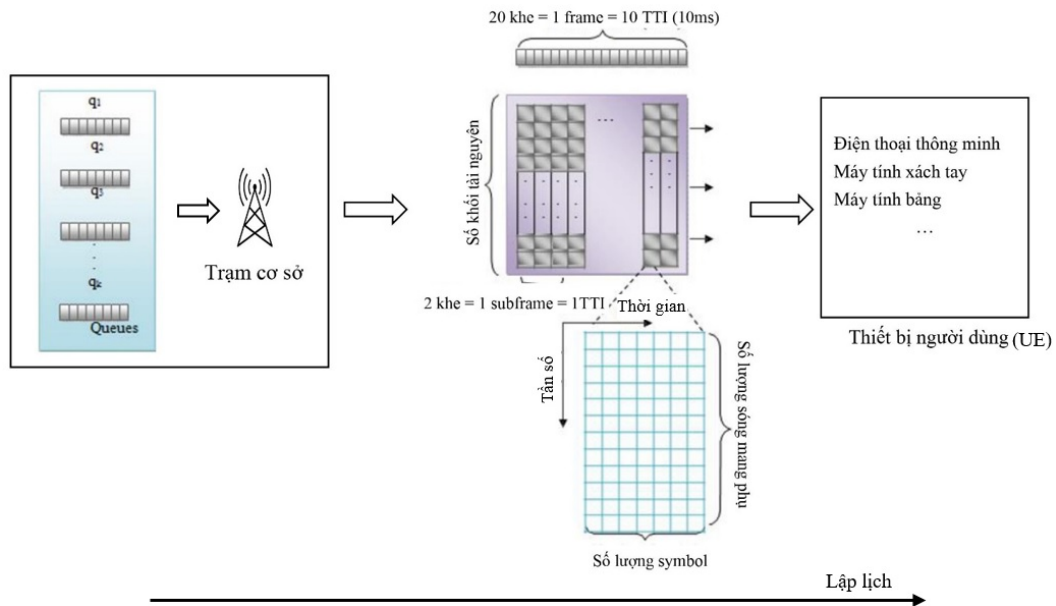
- ii) *Tính công bằng*: Là thước đo được sử dụng trong kỹ thuật mạng để xác định xem người dùng hoặc ứng dụng có đang nhận được một phần tài nguyên hệ thống một cách công bằng hay không.
- iii) *Độ trễ gói tin*: Là thời gian gói tin đến được người dùng kể từ thời điểm gói tin được gửi từ trạm gốc.
- iv) *Tỷ lệ mất gói*: Là tỷ số giữa số gói tin mà người dùng bị mất với số gói được gửi đi.

Trong thực tế, những người dùng ở biên vùng phủ sóng của trạm cơ sở thường có chất lượng kênh kém hơn vì những chướng ngại vật có thể gặp giữa trạm cơ sở và thiết bị người dùng. Điều này dẫn đến tình huống thiết bị không nhận được tài nguyên từ trạm cơ sở. **Thuật toán đề xuất được xây dựng theo cả miền thời gian và miền tần số nhằm tăng cường chất lượng dịch vụ cho người dùng mạng di động cả ở biên vùng phủ sóng.** Người dùng ở biên vùng phủ sóng sẽ nhận được tài nguyên công bằng với các thiết bị khác với độ trễ và tỷ lệ mất gói tối thiểu.

**Bộ lập lịch đề xuất có khả năng phân tích dữ liệu về độ trễ và kích thước của gói tin người dùng để thực hiện phân bổ các khối tài nguyên một cách tốt nhất.** Khi người dùng giao tiếp với một vài eNodeB, dữ liệu người dùng có thể được truyền bằng kênh có chất lượng cao hơn nhằm tăng chất lượng kênh.

Bộ lập lịch đề xuất phân nhóm các UE theo mức độ của lưu lượng truy cập thành hai nhóm: nhóm chính và nhóm phụ. Nếu kênh truyền có chất lượng tốt thì nhóm chính được chọn để truyền, nếu không thì nhóm phụ được sẽ được chọn. Để phân bổ băng thông được tốt hơn, truyền đa sóng mang và ghép kênh phân chia theo tần số được sử dụng là các tham số chính. Bộ lập lịch đề xuất hỗ trợ chọn và chia sẻ các tài nguyên vật lý giữa các kiểu người dùng khác nhau.

## 2.2 Mô hình mạng và các giả thiết



**Hình 2.2.** Mô hình lập lịch trong LTE

Mô hình mạng lập lịch được mô tả như hình (2.2), trong đó, trạm cơ sở (eNodeB) gồm bộ chọn người dùng và bộ cấp phát tài nguyên. Có  $k$  hàng đợi bao gồm  $q_1, q_2, \dots, q_k$  được sử dụng cho quá trình lập lịch. Các hàng đợi này gửi thông tin lập lịch đến trạm cơ sở, sau đó, 20 khe dữ liệu với mỗi khe chứa các khe phụ có tần suất  $1ms$  được lập lịch để gửi tới người dùng. Tín hiệu được gửi tới các thiết bị như điện thoại thông minh, laptop, máy tính bảng v.v. Quá trình lập lịch được thực hiện theo cả miền thời gian và miền tần số. Các thiết bị người dùng kết nối với trạm cơ sở trong môi trường mạng. Các ứng dụng người dùng bao gồm theo thời gian thực và không theo thời gian thực. Môi trường lập lịch giả định là lý tưởng, không có nhiễu tín hiệu.

### 2.3 Mô hình hóa toán học

Trong môi trường nhiều người dùng, việc phân phối tài nguyên giữa những người dùng được thực hiện bằng cách chia sẻ dung lượng kênh giữa chúng. Dung lượng kênh được tính bằng cách sử dụng định lý Shannon [70] như sau:

$$d_k^i(t) = B * \log[1 + SINR_k^i(t)] \quad (2.1)$$

Trong đó,  $d_k^i(t)$  là dung lượng kênh lớn nhất có thể đạt được tính bằng *bit/s* và  $B$  là băng thông tính bằng *Hz*, SINR là tỷ số tín hiệu trên nhiễu. Trên thực tế, giá trị này thường nhỏ hơn giới hạn thực sự đạt được của nó. Với mục đích làm tối đa băng thông hệ thống, dung lượng kênh tính theo thuật toán MT được mô tả như trong công thức  $m_{i,k}^{MT} = d_k^i(t)$ ,  $i = 1..n$ ,  $k = 1..q$  thỏa mãn đối với các ứng dụng không yêu cầu thời gian thực, với các ứng dụng yêu cầu thời gian thực thì dung lượng kênh của thuật toán được tính lại như sau:

$$m_{i,k}^{MT} = \arg \max_i d_k^i(t) \quad (2.2)$$

Dễ dàng nhận thấy, bộ lập lịch MT cho thông lượng tối đa, tuy nhiên hiện tượng “đói” tài nguyên xảy ra khi những người có điều kiện kênh truyền thấp sẽ không nhận được tài nguyên.

Thuật toán PF là sự kết hợp giữa thuật toán MT và BET nhằm cải thiện tính công bằng. Thuật toán này xem xét cả khía cạnh tỷ lệ truyền khả dụng và tỷ lệ truyền dữ liệu nhận được trong quá khứ.

$$\begin{aligned}
 m_{i,k}^{PF} &= \arg \max_i (m_{i,k}^{MT} m_{i,k}^{BET}) \\
 &= \arg \max_i (d_k^i(t) \frac{1}{\overline{R}^i(t-1)}) \\
 &= \arg \max_i \frac{d_k^i(t)}{\overline{R}^i(t-1)}
 \end{aligned} \tag{2.3}$$

Trong đó,  $\overline{R}^i(t-1)$  là thông lượng trung bình trong quá khứ. Với các ứng dụng thời gian thực, trọng số tiện ích theo miền thời gian  $W_i(t)$  và độ trễ được thêm vào khi tính toán:

$$m_{i,k}^{EXP/PF} = \arg \max_i \frac{d_k^i(t)}{\overline{R}^i(t-1)} \alpha_i W_i(t) \tag{2.4}$$

Trong đó,  $\alpha_i$  là độ trễ được tính theo công thức (1.8). Dễ dàng nhận thấy, công thức (2.4) là một phiên bản cải tiến của bộ lập lịch PF. Trong cùng khoảng thời gian truyền, quản lý hàng đợi được thực hiện theo miền thời gian. Để tăng thông lượng, độ trễ gói HOL do nghẽn đầu dòng được thêm vào cho tất cả người dùng theo công thức (2.5).

$$m_{i,k}^{M-LWDF} = \arg \max_i \left( \alpha_i \frac{d_k^i(t)}{\overline{R}^i(t-1)} \exp\left(\alpha_i \frac{W_i(t) - \overline{W}_i(t)}{1 + \sqrt{W_i(t)}}\right) \right) \tag{2.5}$$

**Trễ gói** là thời gian chờ đợi để gói tin được truyền nguồn phát tới người dùng tương ứng. Giá trị này bao gồm thời gian thiết lập kênh truyền, thời gian xử lý tại các nút trung gian và thời gian truyền trên mạng. Gọi  $PD$  là tỷ lệ trễ gói và  $d_i$  là trễ gói của gói tin thứ  $i$ ,  $P_s$  là tổng số gói tin đã gửi, thì hàm mục tiêu:

$$J_1 = PD = \sum_{i=1}^{P_s} d_i \rightarrow \min \tag{2.6}$$

**Mất gói** xảy ra khi một hoặc nhiều gói dữ liệu đi qua mạng không đến được

người dùng. Mất gói xảy ra trong quá trình truyền dữ liệu, thường là qua các mạng không dây hoặc do tắc nghẽn mạng. Mất gói được tính bằng tỷ lệ gói mất và gói được gửi.

$$J_2 = PL = \frac{\sum P_s - \sum P_r}{\sum P_s} \rightarrow \min \quad (2.7)$$

Trong đó,  $PL$  là tỷ lệ mất gói,  $P_s$  là tổng số gói tin đã gửi,  $P_r$  là tổng số gói tin nhận được.

**Băng thông** được định nghĩa là số lượng bit tối đa được gửi qua một kết nối mạng, đơn vị cơ bản của băng thông là bit/s. Thông lượng là số lượng bit thực tế được gửi qua một kết nối mạng trong một khoảng thời gian nhất định. Thông lượng thường nhỏ hơn hoặc bằng băng thông và không bao giờ vượt quá băng thông hệ thống. Băng thông hệ thống phụ thuộc vào nhiều yếu tố như: tắc nghẽn mạng, số người dùng cùng truy cập vào cùng một máy chủ, tài nguyên CPU, RAM của các thiết bị mạng bị giới hạn. Trong mạng LTE, **tốc độ tải xuống** được tính theo công thức:

$$J_3 = DS = ((SBw * 1000)/8)/1024 \rightarrow \max \quad (2.8)$$

Trong đó,  $DS$  là tốc độ tải xuống có đơn vị đo là  $kbps$ ,  $SBw$  là băng thông hệ thống có đơn vị đo là  $kbps$ .

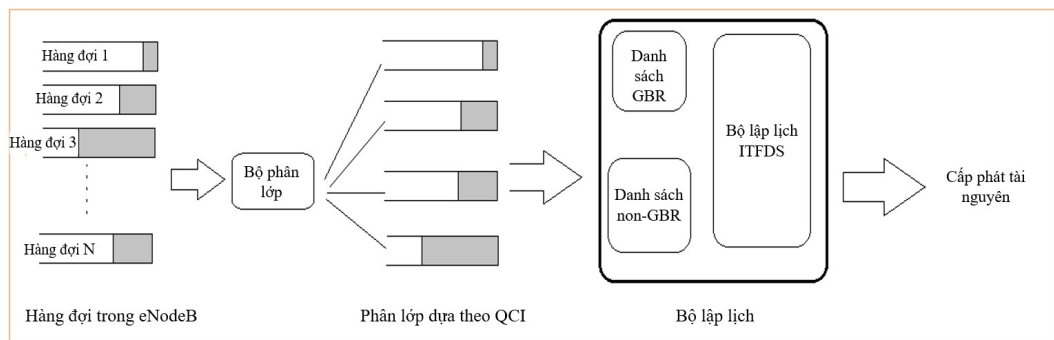
**Chỉ số công bằng** có giá trị là 1 nghĩa là tất cả các người dùng trong mạng có quyền truy cập tài nguyên một cách bình đẳng vào các tài nguyên mạng được chỉ định [17]. Chỉ số công bằng thường được đo bằng Jain's fairness index để điều khiển chỉ số công bằng  $F$  trong mạng [89].

$$J_4 = F = \frac{(\sum_i^N Th_i)^2}{N * \sum_i^N Th_i^2} \quad (2.9)$$

Trong đó,  $N$  là tổng số người dùng,  $Th_i$  là thông lượng của người dùng thứ  $i$  được đo bằng kbps.

## 2.4 Thuật toán lập lịch kết hợp giữa miền thời gian và miền tần số cho đường tải xuống - ITFDS

Chất lượng dịch vụ của các ứng dụng thời gian thực phụ thuộc vào lập lịch các khối tài nguyên dựa trên độ ưu tiên, dung lượng của kênh và tỷ lệ truyền trung bình của tất cả các người dùng trước đó. Trong đó, tỷ lệ truyền được tính bằng số lượng bit truyền đi trong một khoảng thời gian xác định. Trạm cơ sở thực hiện thu nhận thông tin CQI và QCI từ các người dùng, phụ thuộc và thông tin nhận được, trạm cơ sở cấp phát tài nguyên tới các người dùng khác nhau. Hình (2.3) mô tả quy trình thực hiện của thuật toán đề xuất ITFDS.



**Hình 2.3.** Các bước thực hiện của bộ lập lịch đề xuất ITFDS

Trong lập lịch dựa trên thời gian, bộ lập lịch trước tiên sẽ chọn người dùng theo yêu cầu chất lượng dịch vụ của họ cùng với sự công bằng và quyết định tỷ lệ khối tài nguyên tương ứng của nó theo thông lượng trung bình đạt được trong quá khứ và độ dài hàng đợi để đáp ứng các yêu cầu chất lượng dịch vụ. Thuật toán LWDF và tỷ lệ mất gói tin trong quá khứ được sử dụng để xem xét yêu cầu chất lượng dịch vụ và tính công bằng của hệ thống. Đối với mỗi người dùng

được chọn trong miền thời gian, bộ lập lịch theo miền tần số phân bổ các khối tài nguyên cần thiết trong kênh theo điều kiện băng thông của kênh hiện tại. Giá trị QCI cho các ứng dụng khác nhau được đưa ra trong bảng (2.1), các giá trị CQI tiêu chuẩn cho các ứng dụng khác nhau được đưa ra trong bảng (2.2).

**Bảng 2.1.** Bảng các giá trị QCI chuẩn [98]

| QCI | Kiểu tài nguyên (00-GBR, 01-non-GBR) | Giá trị QCI ưu tiên | Chi phí trì hoãn (ms) | Tỷ lệ lỗi gói | Kiểu lưu lượng                      |
|-----|--------------------------------------|---------------------|-----------------------|---------------|-------------------------------------|
| 1   | 00                                   | 2                   | 100                   | $10^{-2}$     | Đàm thoại giọng nói                 |
| 2   | 00                                   | 4                   | 150                   | $10^{-3}$     | Đàm thoại Video                     |
| 3   | 00                                   | 3                   | 50                    | $10^{-3}$     | Chơi game thời gian thực            |
| 4   | 00                                   | 5                   | 300                   | $10^{-6}$     | Phát trực tuyến trong bộ đệm        |
| 5   | 01                                   | 1                   | 100                   | $10^{-6}$     | Tín hiệu IMS                        |
| 6   | 01                                   | 6                   | 300                   | $10^{-6}$     | Phát trực tuyến trong bộ đệm, email |
| 7   | 01                                   | 7                   | 100                   | $10^{-3}$     | Chơi game tương tác                 |
| 8   | 01                                   | 8                   | 300                   | $10^{-6}$     | Các dịch vụ dựa trên TCP            |

Thuật toán đề xuất sử dụng dung lượng kênh và tỷ lệ truyền trung bình trong quá khứ của tất cả các người dùng trước đó để tính độ ưu tiên của gói tin. Bằng cách thu thập tín hiệu CQI và QCI từ người dùng, thuật toán lập lịch cấp phát tài nguyên theo các tỷ lệ trong bảng (2.1), bảng (2.2). Để tăng tỷ lệ truyền, độ trễ  $HOL$  do nghẽn đầu dòng của gói tin được xét đến. Khi đó:

$$m_{i,k}^{ITFDS} = \arg \max_i \left( \alpha_i \frac{d_k^i(t)}{R^i(t-1)} \exp\left(\alpha_i \frac{W_i(t) + \Delta_{i,k}(t) - \overline{W}_i(t)}{1 + \sqrt{W_i(t) + \Delta_{i,k}(t)}}\right) \right) \quad (2.10)$$

Các bước của thuật toán lập lịch đề xuất như sau:

**Đầu vào:** Tập tài nguyên người dùng.

**Bảng 2.2.** Bảng các giá trị CQI cho các sơ đồ điều chế [38]

| CQI | Loại điều chế<br>(00-QPSK,01-16<br>QAM,02-64 QAM) | Tỷ lệ mã | Hiệu quả (số lượng<br>bit thông tin trên<br>mỗi symbol) |
|-----|---|----------|---|
| 1   | 00  | 0,0762   | 0,1523  |
| 2   | 00  | 0,1172   | 0,2344  |
| 3   | 00  | 0,1885   | 0,3770  |
| 4   | 00  | 0,3008   | 0,6016  |
| 5   | 00  | 0,4385   | 0,8770  |
| 6   | 00  | 0,5879   | 1,1758  |
| 7   | 01  | 0,3691   | 1,4766  |
| 8   | 01  | 0,4785   | 1,9141  |
| 9   | 01  | 0,6016   | 2,4063  |
| 10  | 02  | 0,4551   | 2,7305  |
| 11  | 02  | 0,5537   | 3,3223  |
| 12  | 02  | 0,6504   | 3,9023  |
| 13  | 02  | 0,7539   | 4,5234  |
| 14  | 02  | 0,8525   | 5,1152  |
| 15  | 02  | 0,9258   | 5,5547  |

**Đầu ra:** Tỷ lệ mất gói, độ trễ gói, thông lượng và chỉ số công bằng

- *Bước 1:* Tính toán dung lượng kênh theo phương trình

$$d_k^i(t) = B * \log[1 + SINR_k^i(t)]$$

- *Bước 2:* Tính giá trị CQI trong bảng (2.2) theo tỷ lệ truyền cho các ứng dụng khác nhau.
- *Bước 3:* Tùy thuộc vào giá trị QCI, chọn kiểu tài nguyên tương ứng được liệt kê bảng (2.1).
- *Bước 4:* Áp dụng ITFDS bằng cách sử dụng công thức

$$m_{i,k}^{ITFDS} = \arg \max_i \left( \alpha_i \frac{d_k^i(t)}{R^i(t-1)} \exp\left(\alpha_i \frac{W_i(t) + \Delta_{i,k}(t) - \overline{W_i(t)}}{1 + \sqrt{W_i(t) + \Delta_{i,k}(t)}}\right)\right)$$

- *Bước 5:* Trong cùng TTI, hàng đợi được quản lý trong miền thời gian bằng cách kiểm tra:

$$W_i(t) + \Delta_{i,k}(t) > W_i(t)$$

với  $\Delta_{i,k}(t)$  là gia số thời gian.

- *Bước 6:* Tính kích thước bộ nhớ đệm trong UE bằng công thức:

$$Q_{i,k} = Q_{l_{i,k}} - Q_{th_{i,k}}$$

- *Bước 7:* Nếu  $Q_{i,k} \neq \emptyset$ , lặp lại bước 4 đến 7.
- *Bước 8:* Tính tỷ lệ mất gói, độ trễ gói, thông lượng và chỉ số công bằng.

Giải mã của ITFDS được trình bày như trong thuật toán (2.1).

---

### Thuật toán 2.1 Thuật toán ITFDS

---

**Đầu vào:** Tập tài nguyên người dùng

**Đầu ra:** Tỷ lệ mất gói, độ trễ gói, thông lượng và chỉ số công bằng.

---

- 1: Tính dung lượng kênh theo công thức (2.1)
  - 2: Tính giá trị CQI trong bảng (2.2) theo tỷ lệ truyền cho các ứng dụng khác nhau.
  - 3: Chọn kiểu tài nguyên trong bảng (2.1) theo giá trị QCI.
  - 4: **while**  $Q_{i,k} \neq \emptyset$  **do**
  - 5:    Tính giá trị đo  $m_{i,k}^{ITFDS}$  theo công thức (2.10).
  - 6:    **for**  $i \leftarrow 1$  to  $n_{TTI}$  **do**
  - 7:        Hàng đợi được quản lý bằng cách kiểm tra  $W_i(t) + \Delta_{i,k}(t) > W_i(t)$
  - 8:    **end for**
  - 9:    Tính kích thước bộ đệm trong thiết bị người dùng theo công thức  $Q_{i,k} = Q_{l_{i,k}} - Q_{th_{i,k}}$
  - 10: **end while**
  - 11: Tính độ trễ gói  $P_D$  theo công thức (2.6)
  - 12: Tính tỷ lệ mất gói  $P_L$  theo công thức (2.7)
  - 13: Tính thông lượng  $BW$  theo công thức (2.8)
  - 14: Tính chỉ số công bằng  $F_I$  theo công thức (2.9)
- 

Thuật toán lập lịch được đề xuất là thuật toán lập lịch nhận biết kênh/nhận biết chất lượng dịch vụ. Thuật toán này sử dụng điều kiện kênh truyền và các tham số chất lượng dịch vụ làm các tham số lập lịch để phân phối tài nguyên cho tất cả người dùng khả dụng.

Trong mỗi TTI, trạm cơ sở thu thập thông tin từ các hàng đợi khác nhau được kết nối với nó bằng cách sử dụng  $Q_{i,k} = Q_{l_i,k} - Q_{th_{i,k}}$ . Thiết bị người dùng tính toán CQI bằng cách sử dụng tín hiệu tham chiếu từ trạm cơ sở và gửi lại cho trạm cơ sở. CQI là dạng lượng tử hóa của tỷ lệ tín hiệu trên nhiễu (SINR) [14]. Trong hình (2.3), bộ phân lớp được sử dụng để phân loại lưu lượng của các hàng đợi khác nhau theo giá trị QCI. Trong nghiên cứu, thuật toán học có giám sát SVM được sử dụng để phân loại người dùng kể từ khi tham gia hệ thống cho đến khi kết nối chấm dứt. Dữ liệu yêu cầu bởi UE được chia thành nhóm GBR và non-GBR dựa trên các ràng buộc chất lượng dịch vụ [74]. QCI được sử dụng để giúp xác định loại dựa trên dữ liệu ứng dụng của người dùng.

Các hàng đợi lập lịch trong thuật toán đề xuất được lập lịch tùy theo loại dữ liệu của chúng, được thực hiện trên cả thời gian và tần số. Thuật toán được thực hiện trong cả hai chế độ: hoạt động và nền. Việc lập lịch được thực hiện trong mỗi TTI cho tất cả các khối tài nguyên, do đó, làm tăng độ phức tạp của bộ lập lịch.

Có thể thấy, trong thuật toán ITFDS, hàng đợi được quản lý bằng cách kiểm tra độ dài sau mỗi TTI. Một trong các vấn đề cần giải quyết là phải ước lượng được độ trễ  $\Delta$  tại mỗi TTI. Mục sau, luận án đề xuất một phương pháp mới có tên là DCE nhằm phân bổ tài nguyên dựa trên hàng đợi từ nhóm các TTI không được xác định để tăng băng thông hệ thống.

### **Thuật toán Dynamic Class-based Establishment (DCE)**

Các ký hiệu được sử dụng cho thuật toán đề xuất:

- $T, T_1, T_2$  là các ma trận
- $r, s$  là các phần tử của khoảng thời gian truyền (TTI)
- $\delta$  là xác suất trạng thái truyền

- $k$  biểu thị trạng thái truyền
- $\gamma$  là giá trị hằng số
- $R$  biểu thị độ dài kết quả
- $Q$  là giá trị hàng đợi lập lịch
- $R(Q)$  biểu diễn độ dài hàng đợi lập lịch
- $delay_i$  biểu thị giá trị trễ
- $P_c$  là tham số kênh truyền
- $w_k$  là giá trị trọng số truyền
- $T_w$  biểu thị khoảng thời gian truyền theo giá trị trọng số
- $W_i(t)$  biểu thị trọng số theo miền thời gian của mỗi TTI
- $Q_{i,k}$  biểu thị độ dài hàng đợi giữa người dùng thứ  $i$  và người dùng thứ  $k$
- $Q_{th_{i,k}}$  biểu diễn giá trị ngưỡng kích thước hàng đợi của người dùng thứ  $i$  và người dùng thứ  $k$
- $T_v$  là tỷ lệ truyền của người dùng thứ  $v$
- $T_{beginning}$  là phạm vi truyền ở đầu của quá trình thiết lập trong DCE

Khoảng thời gian truyền  $T$  có thể được tạo bằng cách sử dụng ma trận  $T_1$  và  $T_2$ :  $T = T_1 \cup T_2$ .

$$T_1 = \begin{pmatrix} 1-s & s & 0 & 0 & 0 & \dots \\ r(1-s) & rs + (1-r)(1-s) & s(1-r) & 0 & 0 & \dots \end{pmatrix} \quad (2.11)$$

$$T_2 = \begin{pmatrix} 0 & r(1-s) & rs + (1-r)(1-s) & r(1-s) & 0 & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots \end{pmatrix} \quad (2.12)$$

Khi đó, điều kiện truyền là:

$$\delta T = \delta \quad (2.13)$$

ở đó:

$$\delta = [\delta_0, \delta_1, \dots, \delta_k, \dots] \quad (2.14)$$

trong đó  $\delta_k$  là xác suất trạng thái truyền của trạng thái  $k$ . Khi đó phương trình (2.13) được viết lại thành:

$$[\delta_0, \delta_1, \dots, \delta_k, \dots][T] = [\delta_0, \delta_1, \dots, \delta_k, \dots] \quad (2.15)$$

Giải phương trình (2.15) ta được:

$$\delta_1 = \gamma\delta_0, \delta_2 = \gamma^2\delta_0, \dots, \delta_k = \gamma^k\delta_0 \quad (2.16)$$

ở đó:

$$\gamma = \frac{r(1-s)}{s(1-r)} \quad (2.17)$$

Với điều kiện:

$$\delta_0 + \delta_1 + \delta_2 + \dots + \delta_k + \dots + \delta_\infty = 1 \quad (2.18)$$

Giải phương trình (2.17) ta được:

$$\delta_0 = 1 - \gamma$$

$$\delta_0 = \frac{(s-r)}{s(1-r)} \quad (2.19)$$

Độ dài kết quả của lớp trong hàng đợi là:

$$R(Q) = \sum_{k=1}^{\infty} k\delta_k \quad (2.20)$$

Mục tiêu chính của thuật toán là cải thiện thông lượng của quá trình truyền. Dễ dàng nhận thấy, độ dài của hàng đợi có liên quan đến độ dài kết quả của lớp,

thỏa mãn điều kiện:

$$R(Q) \leq delay_i \quad (2.21)$$

Các ràng buộc của bài toán tối ưu là:

$$\left\{ \begin{array}{l} R(Q) \leq delay_i \\ s_i \geq s_i + 1 \\ s_i \geq r_i \\ s_1 + s_2 + \dots + s_n \leq 1 \end{array} \right. \quad (2.22)$$

Các bước thực hiện của DCE được mô tả trong thuật toán (2.2).

---

### Thuật toán 2.2 Thuật toán DCE

---

**Đầu vào:** Nhóm các TTI không xác định

**Đầu ra:** Phân bố truyền dựa trên hàng đợi

*Bước 1:* Khởi tạo tham số kênh  $P_c$ ,  $T_{beginning}$  là phạm vi truyền trong khi quá trình truyền bắt đầu trong mạng.

*Bước 2:* Tính tỷ lệ truyền cao nhất  $T_v$ .

*Bước 3:* Xác định kích thước của TTI.

$$T_w = \sum_{k=1}^n w_k * (P_c - T_v)^2 \quad (2.23)$$

*Bước 4:* Tính trọng số của phạm vi truyền bắt đầu cho mỗi lần lặp lại theo  $P_c$  và  $T_w$

$$T_{current} = \sum_{k=1}^n w_k * (P_c - T_w)^2 \quad (2.24)$$

*Bước 5:* Tính phạm vi truyền theo:

$$\begin{array}{l} \text{if } (T_{beginning} > T_{current}) \\ T_{beginning} = T_{current} \end{array}$$


---

Thuật toán DCE được sử dụng để phân bổ việc truyền dựa trên hàng đợi cho nhóm các TTI không xác định (thuật toán 2.2). Việc tối ưu hóa đạt được

bằng cách sử dụng thuật toán thiết lập dựa trên lớp động mà việc truyền dữ liệu được duy trì với phạm vi truyền. Cơ chế lập lịch DCE được thực hiện khi bắt đầu sự kiện lập lịch. Ở đây, ma trận trọng số  $w_k$  trên khối tài nguyên  $i \in K$  được xác định.

Mặc dù có các công thức khác nhau để xử lý các luồng dịch vụ khác nhau nhưng chỉ có một giá trị  $\delta$  được tính cho tất cả các luồng tại TTI hiện tại. Sau đó các luồng được so sánh với nhau để chọn ra luồng  $j \in K$  với giá trị đo lớn nhất, do đó băng thông của hệ thống được cải thiện.

## 2.5 Thực nghiệm đánh giá kết quả

Thuật toán lập lịch ITFDS đề xuất được so sánh với một số phương pháp liên quan là PSS [88] và QuAS [102], cho các ứng dụng thời gian thực và không theo thời gian thực. Mọi gói dữ liệu được thiết bị của người dùng yêu cầu được lên lịch đến hàng đợi để truyền lập lịch đường xuống LTE.

**Bảng 2.3.** Các tham số sử dụng trong mô phỏng của thuật toán đề xuất

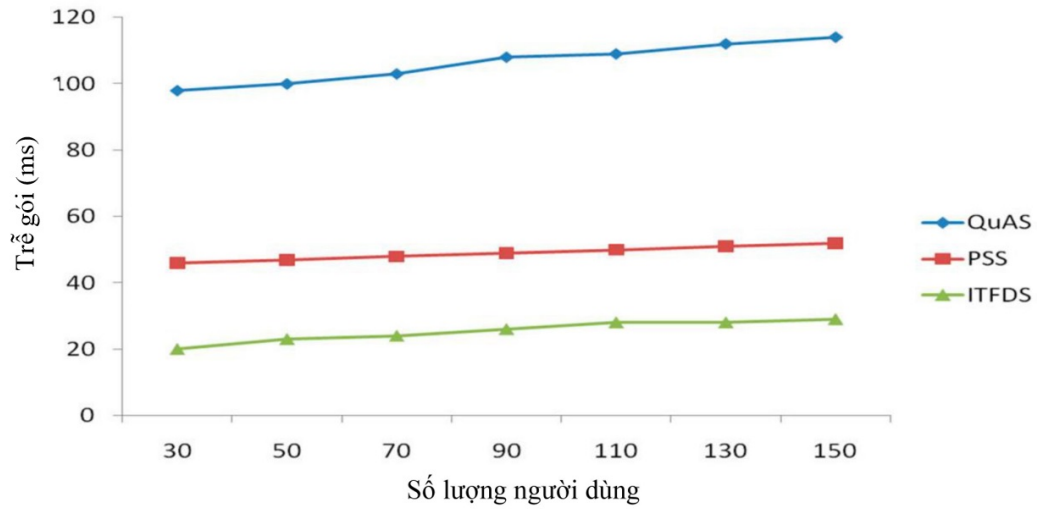
| Giải thích                   | Miền giá trị  |
|------------------------------|---|
| Băng thông khả dụng          | 20,0 MHz [55]   |
| Số khối tài nguyên           | 100 (12 sóng mang phụ cho mỗi khối tài nguyên vật lý [6]) |
| Bán kính phủ sóng            | 0-15 km [6]   |
| Số lượng người dùng          | 30, 50, 70, 90, 110, 130, 150 [9]                         |
| Tốc độ                       | 3 km/h [62]   |
| Luồng GBR                    | 8,4 kbps cho âm thanh và 242kbps cho Video (H264) [62]    |
| Thời gian mô phỏng           | 30s [62]  |
| Số lần lập cho mỗi lịch biểu | 10 [62]   |

Miền tần số được sử dụng để tối ưu hóa băng thông của thuật toán một cách hiệu quả. Các tham số mô phỏng được cho trong bảng (2.3). Các tiêu chí đánh giá trong thực nghiệm bao gồm tỷ lệ trễ gói theo công thức (2.6), tỷ lệ mất gói theo công thức (2.7) và thông lượng theo công thức (2.8). Trong đó, tỷ lệ trễ gói càng nhỏ càng tốt, tỷ lệ mất gói càng nhỏ càng tốt và thông lượng càng cao càng tốt.

### 2.5.1 Thực nghiệm tỷ lệ trễ gói với lượng người dùng khác nhau

Trễ gói được dự đoán bằng cách sử dụng kỹ thuật nâng cao thông qua hàng đợi trong mạng LTE. Các khối tài nguyên được sử dụng hiệu quả để cung cấp chiến lược phân bổ dựa trên miền thời gian và miền tần số. Hàng đợi ảo đã được tạo ra để phân tích tổng số lượng các gói tin đến và được sử dụng để sửa đổi các tham số truyền nhằm giảm thiểu độ trễ gói. Hình (2.4) biểu diễn kết quả hiệu suất trễ gói đối với số lượng các người dùng khác nhau trong mô phỏng. Kết quả của mô phỏng chỉ ra rằng, khi  $w_k$  được điều chỉnh thích hợp, bộ lập lịch đề xuất có thể đối phó tốt với cả lưu lượng theo thời gian thực và không theo thời gian thực đồng thời. Ngoài ra, bộ lập lịch được đề xuất cũng đảm bảo mức độ công bằng nhất định giữa tất cả người dùng vì nó dựa trên tiêu chí của thuật toán PF. Kết quả của mô phỏng cũng cho thấy độ trễ tăng khi số lượng người dùng càng tăng. Thuật toán đề xuất cho thấy hiệu suất tốt hơn QuAS và PSS, vì trọng số khác nhau được phân bổ cho các ứng dụng khác nhau QuAS và PSS chỉ cho phép phân bổ để cùng một trọng số được phân bổ cho tất cả các người dùng.

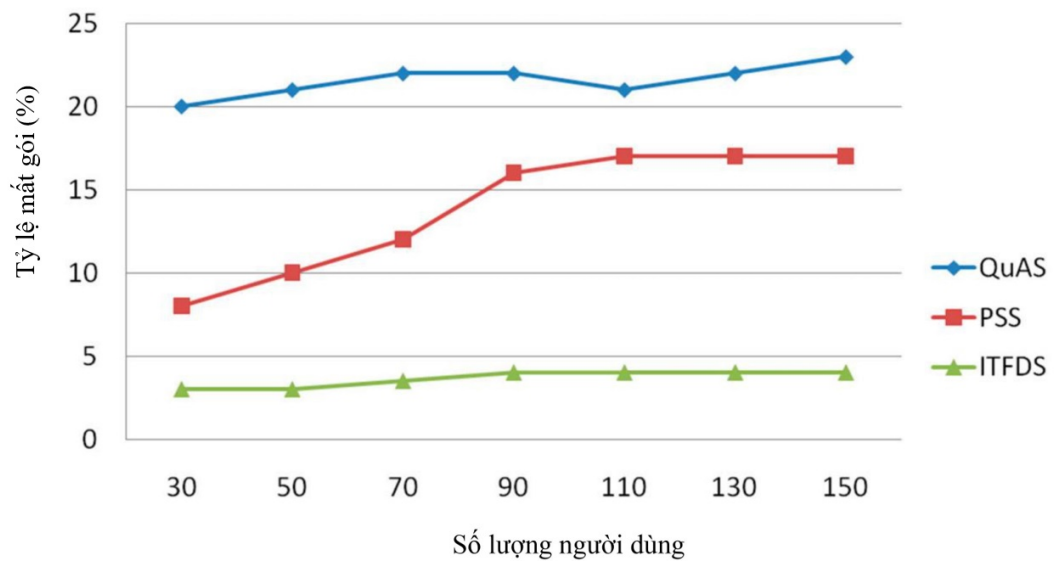
Thuật toán ITFDS được đề xuất phân phối các độ trễ khác nhau cho các ứng dụng khác nhau nhưng nó luôn cho hiệu suất tốt nhất so với các thuật toán khác vì sự kết hợp cả miền thời gian và miền tần số để lập lịch trình trong mô phỏng, việc thiết lập đúng giá trị trọng số  $w_k$  có thể giảm độ trễ một cách hiệu



Hình 2.4. Trễ gói với số lượng người dùng khác nhau

quả. Đối với tải hệ thống cao, độ trễ tối đa của gói theo thời gian thực trung bình luôn bằng 0,1 giây vì gói theo thời gian thực sẽ bị loại bỏ nếu nó vi phạm ngưỡng 20 vòng (0,1 giây)

### 2.5.2 Thực nghiệm tỷ lệ mất gói với lượng người dùng khác nhau

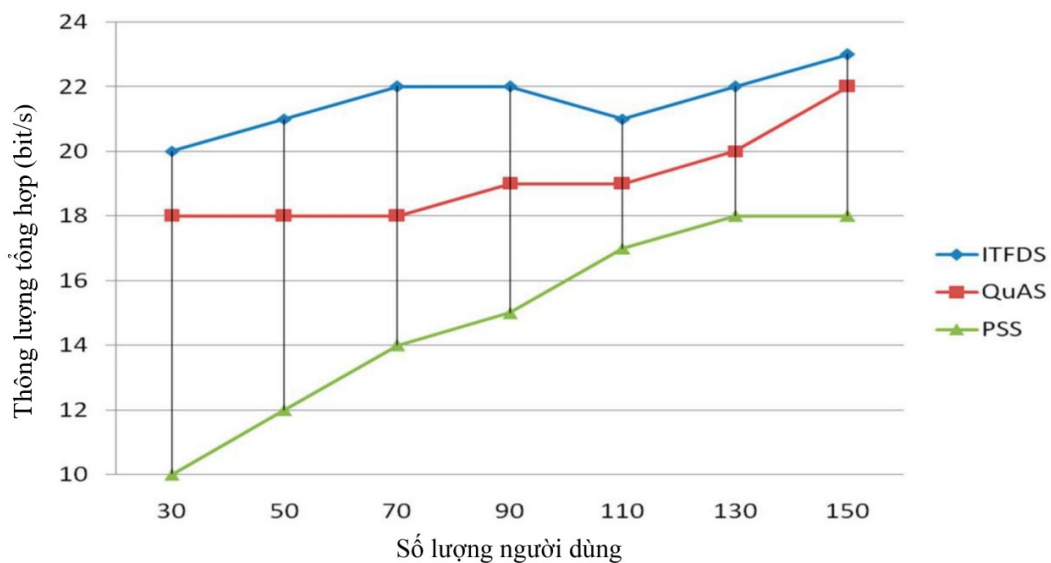


Hình 2.5. Tỷ lệ mất gói với số lượng người dùng khác nhau

Hình (2.5) biểu diễn tỷ lệ mất gói cho các thuật toán lập lịch khác nhau. Tỷ lệ mất gói (PLR) là một trong những tham số chính trong truyền thông không dây và ảnh hưởng trực tiếp đến chất lượng dịch vụ. Trong kết quả mô phỏng, khi số lượng người dùng tăng lên, tỷ lệ mất gói của QuAS và PSS cũng tăng theo, trong khi thuật toán lập lịch đề xuất đạt được tỷ lệ mất gói thấp hơn các thuật toán lập lịch được sử dụng để so sánh và duy trì cùng một tỷ lệ mất gói cho tất cả các người dùng khác nhau của UE.

Các phương thức hiện có không cung cấp mức độ ưu tiên cho các gói trong khi thực hiện quá trình lập lịch trình. Ngược lại, kỹ thuật được đề xuất có việc thực hiện ưu tiên gói để truyền tải sao cho tỷ lệ mất gói là giá trị nhỏ nhất cho kỹ thuật đề xuất.

### 2.5.3 Thực nghiệm thông lượng với lượng người dùng khác nhau



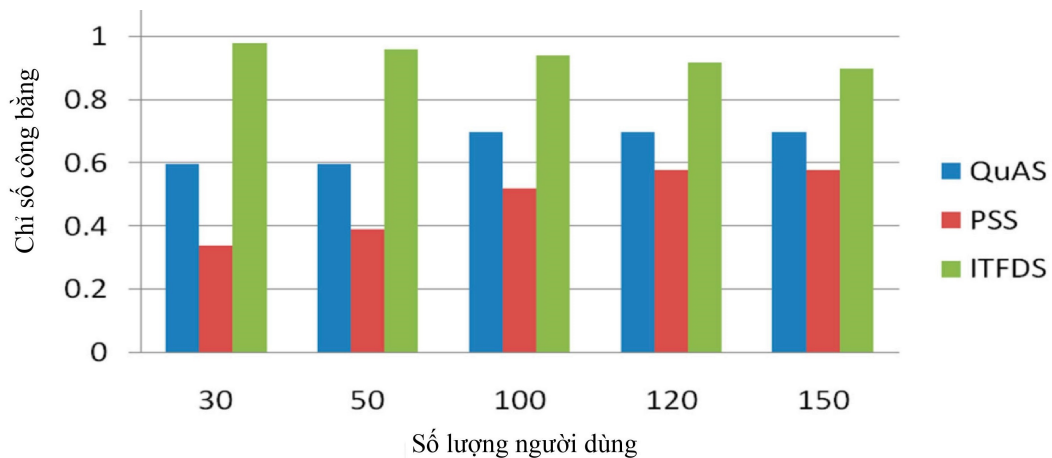
Hình 2.6. Thông lượng với số lượng người dùng khác nhau

Hình (2.6) biểu diễn thông lượng tổng hợp cho các ứng dụng khác nhau. Theo đó, qua kết quả mô phỏng, phương pháp được đề xuất không cho thấy hiệu

suất tốt hơn các phương pháp về thông lượng tổng hợp vì các ứng dụng là ứng dụng không theo thời gian thực, sự thay đổi trọng số không làm thay đổi kết quả thực nghiệm.

Sự lựa chọn luồng dữ liệu được cải thiện theo chỉ số chất lượng kênh (CQI) với nhu cầu lưu lượng thấp. Do quá trình truyền video phụ thuộc vào cả QCI và CQI, những thay đổi trong giá trị  $\alpha$  dẫn đến sự thay đổi trong quá trình truyền Video. Hình (2.6) cho thấy hiệu suất thông lượng cho các ứng dụng khác nhau bằng cách thay đổi giá trị của  $\alpha$ . Các cải tiến của phương pháp đề xuất sẽ được nghiên cứu thêm trong các công trình trong tương lai.

#### 2.5.4 Thực nghiệm chỉ số công bằng với lượng người dùng khác nhau



**Hình 2.7.** Chỉ số công bằng với các người dùng khác nhau

Hình (2.7) biểu diễn chỉ số công bằng cho các người dùng khác nhau. Từ các kết quả thu được, các thuật toán QuAS và PSS hầu hết hoạt động theo một cách tương tự. Sự khác biệt chính giữa hai phương pháp này là sự hội tụ. Với thuật toán lập lịch ITFDS đề xuất, độ phức tạp được giảm bớt và tính công bằng được cải thiện.

Chỉ số công bằng được tính bằng cách sử dụng các giá trị của thông lượng chuẩn hóa và tổng số kết nối động trong LTE. Đảm bảo tính công bằng là một tham số khác để tạo ra chất lượng dịch vụ, các yêu cầu dịch vụ của người dùng là rất cần thiết để chuyển một lượng lớn dữ liệu. Theo kết quả mô phỏng, khi số lượng người dùng tăng, giá trị công bằng đã được tăng lên, thuật toán đề xuất phân bổ nhiều tài nguyên hơn tới người dùng.

## 2.6 Phân tích độ phức tạp thuật toán

Độ phức tạp của phương pháp đề xuất được tính như sau. Gọi  $n$  là tổng số người dùng, người dùng có thể chọn tốc độ được cải thiện tối ưu tài nguyên truyền gói dữ liệu tại mọi thời điểm. Nếu thuật toán được đề xuất xác định  $p$  người dùng từ tổng số người dùng để truyền dữ liệu thì độ phức tạp của thuật toán là  $O(\log p)$ .

Tỉ lệ trung bình được sử dụng để chọn mỗi người dùng là  $O(1)$ . Thời gian đến trung bình được tính toán và phân tích gói được truyền cuối cùng trong hàng đợi hoạt động. Thời gian ước tính cũng được tính trong khoảng thời gian quy định. Xác suất thông qua việc truyền dữ liệu được tính toán để phân tích độ phức tạp thời gian trong hàng đợi. Quá trình lập lịch cũng được phân tích để truyền gói dữ liệu và độ phức tạp thời gian tổng thể được tính là  $O(\log p)$ .

## 2.7 Ưu và nhược điểm của phương pháp đề xuất

Thuật toán đề xuất ITFDS là thuật toán thuộc nhóm nhận biết kênh/nhận biết chất lượng dịch vụ. Qua thực nghiệm, thuật toán đề xuất cho kết quả tốt

hơn các thuật toán QuAS và PSS ở khía cạnh thông lượng, tính công bằng và tỷ lệ mất gói.

Mô phỏng được thực nghiệm với các ứng dụng theo thời gian thực và không theo thời gian thực. Thuật toán đề xuất cho kết quả tốt hơn về thông lượng với các ứng dụng theo thời gian thực và đảm bảo tính công bằng với các ứng dụng không đòi hỏi thời gian thực. Tỷ lệ mất gói giảm xuống trong thuật toán đề xuất so với thuật toán QuAS và PSS.

Mặc dù đã đạt được những cải tiến, nhưng thuật toán đề xuất có tỷ lệ dữ liệu truyền bị giảm khi khoảng cách giữa trạm cơ sở và người dùng lớn. Đôi khi, các vấn đề về kết nối xảy ra khi tín hiệu âm thanh chuyển thành gói, khi đó thời gian truyền sẽ tăng lên.

## 2.8 Kết luận chương

Chương này đề xuất một thuật toán lập lịch ITFDS thuộc nhóm nhận biết kênh/ nhận biết chất lượng dịch vụ mới tích hợp giữa miền thời gian và miền tần số nhằm cải thiện chất lượng lập lịch cho đường tải xuống trong mạng di động đa dịch vụ. Thuật toán đề xuất phân phối tài nguyên tới tất cả người dùng theo cả miền thời gian và miền tần số. Thuật toán ITFDS cho kết quả tốt hơn các thuật toán QuAS và PSS về thông lượng, tính công bằng, và tỷ lệ mất gói theo thực nghiệm. Thuật toán cho kết quả tốt với các ứng dụng yêu cầu thời gian thực đồng thời đảm bảo tính công bằng giữa các người dùng.

Mặc dù đã đạt được những cải tiến đáng kể so với các thuật toán QuAS và PSS, nhưng thuật toán đề xuất vẫn có nhược điểm như chưa cho kết quả tốt hơn tỷ lệ truyền với khoảng cách giữa trạm cơ sở và người dùng lớn do các vấn đề về kết nối xảy ra khi tín hiệu âm thanh chuyển thành gói. Bên cạnh đó thuật toán

đề xuất mới chỉ được cài đặt trên mạng 4G-LTE mà chưa thử nghiệm trên môi trường mạng 5G. Thuật toán đề xuất cũng mới so sánh kết quả với thuật toán QuAS và PSS mà chưa so sánh với các thuật toán cùng loại khác.

Nhược điểm của thuật toán đề xuất là hướng nghiên cứu tiếp theo nhằm mở rộng thuật toán trong bối cảnh khác và mạng 5G để đảm bảo thuật toán đề xuất đủ tốt để áp dụng vào ngành công nghiệp 4.0. Ngoài bài toán lập lịch trong mạng di động đa dịch vụ, lập lịch trong lưới tính toán di động cũng là một thành tố quan trọng trong các hệ thống tính toán phân tán. Trong chương 3, luận án sẽ trình bày nội dung nghiên cứu mới trên lưới tính toán di động, đề xuất một thuật toán lập lịch mới cho mạng này với mục tiêu làm tăng chất lượng dịch vụ.

Các kết quả của chương này đã được công bố tại [CT2] và [CT3] trong danh mục các công trình khoa học của tác giả có liên quan đến luận án.

## Chương 3

# ĐỀ XUẤT GIẢI PHÁP LẬP LỊCH CHO LƯỚI TÍNH TOÁN DI ĐỘNG

Để đáp ứng những đòi hỏi ngày càng cao về chất lượng dịch vụ của người dùng, nhiều bài toán phức tạp đòi hỏi các hệ thống tính toán phải xử lý trong thời gian cho phép. Tuy nhiên, để xử lý các vấn đề phức tạp này trong thời gian ngắn là một việc hết sức khó khăn. Các nhà khoa học trên thế giới vẫn đang nghiên cứu và phát triển các phương pháp để giải quyết vấn đề này bên cạnh các mô hình giải quyết đã có. Xử lý song song, điện toán phân tán, điện toán cụm, điện toán đám mây và điện toán lưới là một số ví dụ [25, 32].

Với sự phát triển nhanh chóng của công nghệ, những thiết bị di động ngày càng trở nên phổ biến hơn bởi khả năng thích ứng và tính toán của chúng. Các thiết bị di động thông minh với kích thước nhỏ nhưng được tích hợp với quy mô rất lớn sẽ tạo thành một lưới tính toán có thể thực hiện hàng triệu phép tính mỗi giây tạo thành một hệ thống tính toán với hiệu năng rất cao [31].

Chương này, sẽ trình bày lý thuyết về tính toán lưới, tập trung vào lưới có sự hiện diện của các thiết bị di động và đề xuất một **thuật toán lập lịch mới**

trong lưới tính toán di động dựa trên tìm kiếm theo heuristic nhằm cải thiện hiệu năng của hệ thống. Kết quả của thuật toán đề xuất được so sánh với một số thuật toán lập lịch truyền thống như MIN-MIN, MAX-MIN, HEFT để kiểm chứng.

### 3.1 Giới thiệu

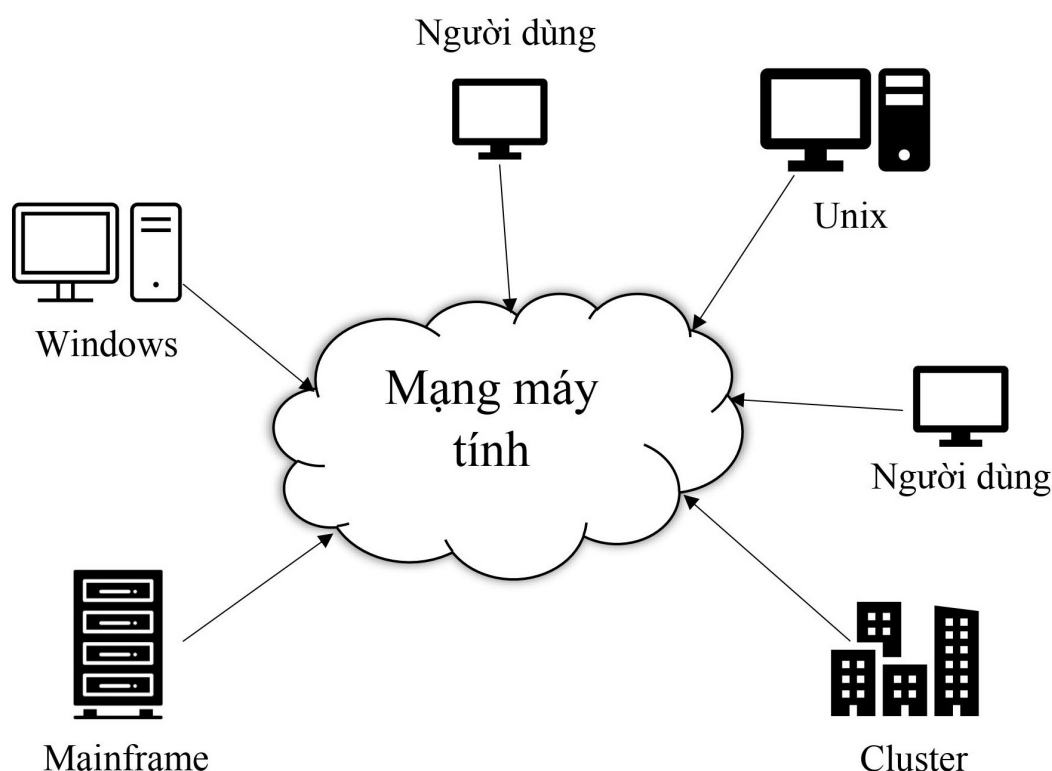
Trong những năm gần đây, các yêu cầu ngày càng tăng của các ứng dụng tiếp tục đặt ra những thách thức to lớn đối với các hệ thống tính toán song song và tính toán phân tán truyền thống, mặc dù tính khả dụng của tính song song, sức mạnh tính toán, các giải pháp đa lỗi và lưu trữ là rất lớn của chúng [18].

Các vấn đề phức tạp mới như vật lý năng lượng cao, mô hình phân tử và khoa học trái đất, v.v. luôn cần nhiều sức mạnh tính toán cao, dung lượng lưu trữ lớn, tốc độ xử lý nhanh hơn và khả năng truy cập dữ liệu nhanh hơn với chi phí thấp hơn. Nhu cầu của các ứng dụng cho các nguồn tài nguyên mạnh mẽ đã được đáp ứng với những tiến bộ đạt được trong công nghệ xử lý thông tin và truyền thông.

Các nhu cầu lớn của các ứng dụng chỉ có thể được giải quyết thông qua một nền tảng tiến hóa tạo ra bằng cách tích lũy khả năng của nhiều tài nguyên không đồng nhất, phi tập trung, cục bộ và từ xa, trải rộng khắp các vị trí địa lý. Vào giữa những năm 1990, một nền tảng tiến hóa được phát triển bằng cách mở rộng cộng tác, cùng với khả năng tính toán và mạng ngày càng tăng đã tạo ra một kỷ nguyên mới của điện toán, được gọi là “điện toán lưới” [32].

Điện toán lưới có thể được phát biểu như một sự trừu tượng hóa cho phép tích hợp rất nhiều hệ thống máy tính nối mạng với khả năng kết hợp tốc độ xử lý, không gian lưu trữ, v.v. của các nguồn tài nguyên được phân phối theo địa lý,

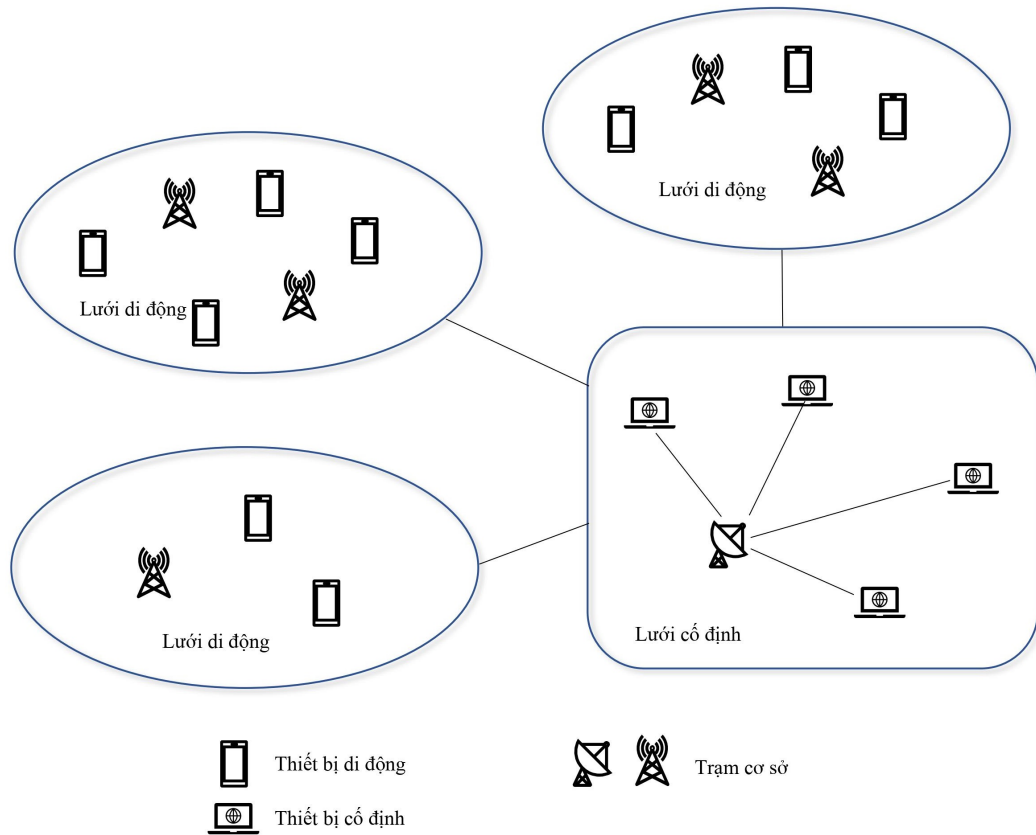
sẵn có và không đồng nhất để cung cấp một môi trường an toàn và thân thiện với người dùng nhằm tạo ra một môi trường thực thi các ứng dụng của người dùng một cách hiệu quả đồng thời che giấu nhiều chi tiết kỹ thuật và sự phức tạp của nó khỏi thế giới bên ngoài [59]. Môi trường điện toán lưới được mô tả như trong hình (3.1).



**Hình 3.1.** Môi trường điện toán lưới

Lưới tính toán di động (MCG) là một sự mở rộng của điện toán lưới [32] mà ở đó có sự kết hợp giữa các thiết bị cố định và thiết bị di động. Các thiết bị di động ngày càng thông minh, chúng phù hợp với các yêu cầu của người dùng, khi được tích hợp vào lưới, chúng có khả năng giải quyết các bài toán tính toán chuyên sâu [32]. Lưới tính toán di động được mô tả như hình (3.2).

Bài toán lập lịch công việc trong lưới tính toán di động là bài toán thuộc lớp NP-đầy đủ [71] và chưa có lời giải tối ưu do ở đó lưới hiện diện cả nút tĩnh và



**Hình 3.2.** Môi trường lưới tính toán di động

nút động. Với các nút động, kết quả tính toán cần phải được trả lại cho hệ thống nếu nút động di chuyển từ trạm làm việc này sang trạm làm việc khác hoặc hết năng lượng. Trong trường hợp này, quá trình địa phương hóa sẽ giúp tìm kiếm tài nguyên cục bộ nhằm thỏa mãn yêu cầu của một công việc [22]. Phương thức tìm kiếm dựa trên heuristic giúp định vị tài nguyên một cách chính xác nhằm thỏa mãn yêu cầu người dùng [104, 115].

Tài nguyên trong lưới tính toán di động là năng lực xử lý của của bộ xử lý, băng thông mạng, năng lực lưu trữ, phần mềm, v.v. Tất cả đều được liên kết với nhau tạo thành một thể thống nhất. Lớp các chương trình phần mềm trung gian có nhiệm vụ quản lý, giám sát công việc, quản trị tài nguyên, bảo mật v.v. Người dùng có thể yêu cầu tài nguyên trực tiếp hoặc gián tiếp tới lưới.

Do đặc điểm của lưới tính toán di động, cấu trúc của lưới luôn luôn thay đổi khi có nút ra nhập mạng hoặc rời mạng. Đó là do các nút di động thay đổi vị trí một cách thường xuyên, hoặc do năng lượng của các thiết bị cạn kiệt. Khi đó cấu trúc mạng ban đầu không được bảo toàn.

Khi nhận được yêu cầu, hệ thống quản lý tài nguyên của lưới sẽ lập lịch trình cho các tài nguyên hoạt động. Việc quản lý tài nguyên trong lưới trở nên phức tạp hơn do tính chất không đồng nhất của các tài nguyên và quyền sở hữu. Cùng với sự thay đổi liên tục trong cấu trúc liên kết mạng, bài toán lập lịch tài nguyên trong lưới là một trong các thách thức mà lưới tính toán di động này cần giải quyết [115]. Những thách thức lớn mà lưới phải đối mặt là khả năng mở rộng và khả năng thích ứng, sự hạn chế của thiết bị di động cũng như tính di động và sự sẵn có của các nguồn lực.

Do lưới tính toán di động gồm các nút tĩnh và nút động, lưới tính toán di động được chia gồm nhiều lưới phụ. Các lưới phụ có thể chỉ chứa các nút di động, hoặc chỉ có các nút cố định hoặc cả hai. Để tối ưu hóa hệ thống, bộ lập lịch trong lưới được chia thành: (1) bộ lập lịch địa phương; (2) bộ lập lịch toàn cục.

- 1) *Bộ lập lịch địa phương*: Là một thành phần quan trọng thực thi trong lưới tính toán di động. Bộ lập lịch địa phương là một dịch vụ chạy trên các thiết bị, một phần của tác nhân thông minh được chia sẻ để quản lý các dịch vụ di động cuối. Bộ lập lịch địa phương xử lý các tác vụ của trình quản lý thiết bị đầu cuối. Các tác vụ lập lịch khác như phần mềm hoặc hệ điều hành được xử lý bởi máy chủ. Bộ lập lịch địa phương được sử dụng để lên lịch các tác vụ trên thiết bị theo thời gian định sẵn. Bộ lập lịch địa phương cấp cho mỗi tác vụ một số ID. Các tập lệnh của bộ lập lịch địa phương có dải ID khác với các tập lệnh của bộ lập lịch địa phương mặc định mà trình quản lý đầu

cuối sử dụng [41, 56]. Theo mặc định, chỉ có thể có một tập lệnh lập lịch tùy chỉnh hoạt động trên mỗi thiết bị.

- 2) *Bộ lập lịch toàn cục*: Chịu trách nhiệm chọn lưới phù hợp và ánh xạ các công việc cho lưới được chỉ định này. Bộ lập lịch toàn cục thiết lập ưu tiên cho các công việc bằng cách sử dụng ánh xạ ưu tiên của công việc với ưu tiên toàn cục [41, 56].

Địa phương hóa là một phương pháp gán các tài nguyên cho bộ lập lịch địa phương. Các tài nguyên có thể ở dạng địa phương hoặc toàn cục hoặc cả hai trong lưới tính toán di động. Bộ lập lịch địa phương tối ưu hóa chất lượng của các tham số dịch vụ như thời gian phân bổ tài nguyên [52]. Một số kết quả của địa phương hóa đã được công bố trong thời gian gần đây, tiêu biểu như sau:

Trong [27], một thuật toán địa phương hóa được Han và cộng sự đề xuất sử dụng nút neo di động trong các mạng cảm biến không dây. Các thuật toán địa phương hóa hỗ trợ nút neo di động được phân loại: địa phương hóa dựa trên mô hình di động và địa phương hóa bằng cách sử dụng lược đồ lập lịch đường. Bài toán quan trọng nhất đối với thủ tục địa phương hóa hỗ trợ nút neo di động là xác định đường di chuyển của các nút neo nhằm tối ưu tỷ lệ địa phương hóa trong khu vực giám sát. Tiếp theo, thủ tục địa phương hóa hỗ trợ nút neo di động xác định quá trình địa phương hóa được sử dụng bởi các nút khác để tính toán vị trí của chúng tùy thuộc vào thông tin báo hiệu nhận được bằng cách di chuyển thiết bị neo đã nhận biết vị trí. Các thủ tục như vậy hoặc sử dụng các nút neo di chuyển hoặc nút tham chiếu và di chuyển nút neo để hỗ trợ các nút khác thực hiện địa phương hóa.

Chenglu [108] đã đưa ra một hệ thống để lập bản đồ di động cho mục đích sử dụng đặc biệt với các khu vực không có hệ thống vệ tinh định vị toàn cầu/

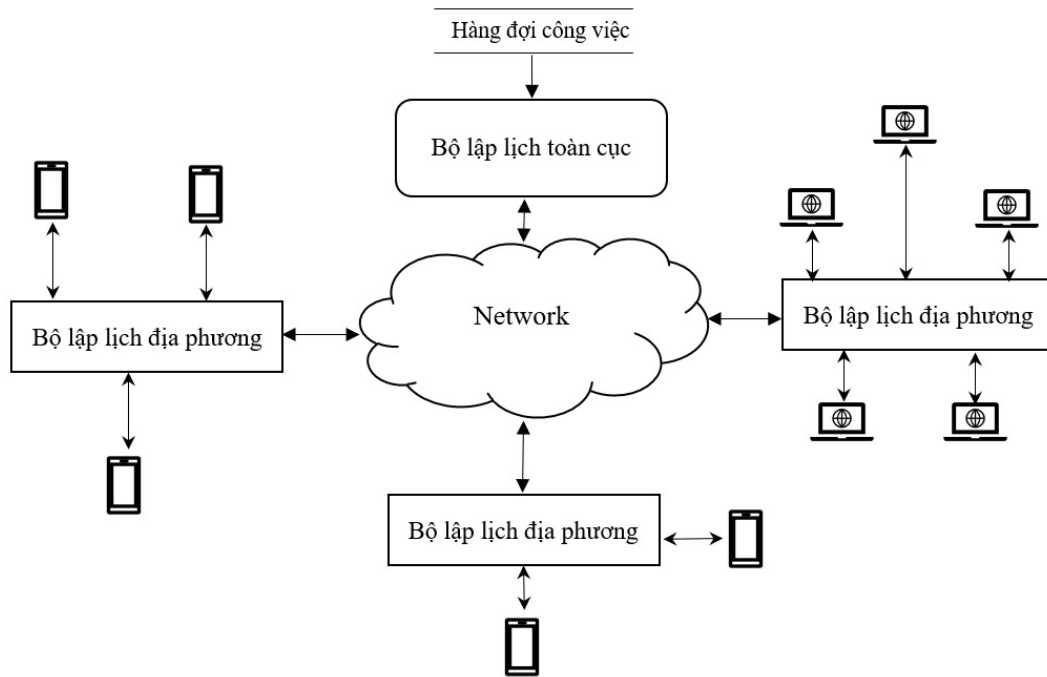
và hệ thống định vị toàn cầu. Bản đồ 2D và 3D có thể được tạo ra bởi hệ thống bằng cách sử dụng địa phương hóa 6 bậc tự do (6DOF).

Fatih [118] đã sử dụng dịch vụ dựa trên hệ thống định vị toàn cầu (GPS) để tìm vị trí của đồ vật bị mất. Để tính toán chi phí, các tham số mới đã được đưa vào để tính toán. Thuật toán GA và tìm kiếm heuristic được sử dụng để phân nhóm người dùng trên bộ dữ liệu mạng xã hội. Địa phương hóa được xác định bởi dịch vụ dựa trên GPS của các đồ vật bị mất. Những thiết bị người dùng được phân cụm và một mạng lưới báo hiệu được thành lập khi các đồ vật bị mất nhận được tín hiệu từ các thiết bị địa phương khác. Nhóm tác giả đã làm một lượng lớn mô phỏng trên bộ dữ liệu; hiệu quả của phương pháp này là các cụm được tạo ra với một số lượng nhỏ sự trao đổi giữa các người dùng, đảm bảo tạo ra một sự bảo vệ riêng tư của người dùng.

Liu [52] đề xuất phương pháp quản lý khả năng mở rộng địa phương hóa và độ chính xác của điện thoại sử dụng cảm biến cơ hội (một mô hình xử lý tín hiệu và thông tin trong đó một mạng lưới các hệ thống cảm biến có thể tự động phát hiện và lựa chọn để cảm biến dựa trên một kịch bản định trước). Phương pháp đề xuất sử dụng phương pháp địa phương hóa để giải quyết các vấn đề như: vị trí như chuyển động trong nhà cho người mù, tìm kiếm thực tế ảo trong các trò chơi hay chuyển động rô bốt. Mô hình sử dụng dữ liệu cảm biến đa phương thức để tăng cường khả năng mở rộng và nâng cao tính hiệu quả.

### 3.2 Mô hình mạng và các giả thiết

Để xây dựng thuật toán, mục này nghiên cứu các phân tích toán học cơ bản về lập lịch được đưa ra trong các tài liệu [32, 41, 115], trong đó, bài toán lập lịch công việc trong lưới tính toán di động sử dụng các giả định sau:



**Hình 3.3.** Mô hình mạng cho bài toán lập lịch trong lưới tính toán di động

- Mỗi máy có thể thực hiện một công việc phụ tại một thời điểm.
- Thời gian thực hiện công việc phụ của một việc được biết trước.
- Năng lực xử lý của máy lớn hơn kích thước công việc phụ.
- Kích thước các công việc phụ là như nhau.
- Phương pháp đề xuất thực nghiệm trong môi trường mô phỏng và mạng bị giới hạn, không bị lỗi mạng trong quá trình lập lịch.

Mô hình mạng cho bài toán lập lịch trong lưới tính toán di động được mô tả như hình (3.3).

### 3.3 Mô hình hóa

Các ký hiệu và mô tả sử dụng trong mô hình đề xuất được đề cập trong bảng (3.1). Mô hình đề xuất xem một máy là máy chính và các máy khác là máy

**Bảng 3.1.** Các ký hiệu sử dụng cho thuật toán lập lịch đề xuất

| STT | Ký hiệu      | Mô tả  |
|-----|--------------|--|
| 1   | $V_k$        | Công việc thứ $k$ trong tập công việc $V$                        |
| 2   | $v_{ik}$     | Công việc phụ thứ $k$ của công việc $i$                          |
| 3   | $TG(v_{ik})$ | Thời gian thực hiện công việc phụ $v_{ik}$                       |
| 4   | $TG(V_i)$    | Thời gian thực hiện công việc thứ $i$                            |
| 5   | $S_{ik}$     | Độ lớn của công việc phụ $v_{ik}$                                |
| 6   | $S_k$        | Tốc độ xử lý của nút $d_k$ trong lưới                            |
| 7   | $NS_k$       | Số lượng công việc phụ chưa được phân bổ trên nút do quá tải     |
| 8   | $NS_1$       | Tổng số công việc phụ chưa được thực hiện bởi lưới               |
| 9   | $NS'_k$      | Số công việc phụ có thể được phục vụ bởi nút dưới tải trong lưới |
| 10  | $NS_2$       | Tổng số công việc phụ còn có thể được thực hiện bởi lưới         |
| 11  | $\mu_i$      | Tỷ lệ phục vụ trung bình của máy $M_i$                           |
| 12  | $\lambda_i$  | Tỷ lệ công việc phụ đến máy $M_i$                                |

phụ. Máy chính nhận tất cả các công việc và phân phối chúng tới các lưới phụ dưới sự hỗ trợ của bộ lập lịch toàn cục. Tại các lưới, các máy phụ được lựa chọn để gán công việc dưới sự hỗ trợ của bộ lập lịch địa phương.

Ban đầu, mỗi máy thứ cấp đều khai báo tốc độ và khả năng tính toán của chúng. Sau đó, máy chính phân phối các công việc và các công việc phụ liên quan (với sự trợ giúp của bộ lập lịch địa phương và bộ lập lịch toàn cục) cho máy phụ phù hợp với yêu cầu công việc và khả năng xử lý của máy. Giả sử công việc  $V_k$  có  $n_k$  công việc con,  $v_{ik}$  là công việc phụ thứ  $k$  của việc thứ  $i$ . Nếu  $TG(v_{ik})$  đại diện cho thời gian thực hiện  $v_{ik}$  thì:

$$\begin{aligned}
 TG(V_1) &= \sum_{i=1}^{n_1} TG(v_{1i}) \\
 TG(V_2) &= \sum_{i=1}^{n_2} TG(v_{2i}) \\
 &\dots\dots\dots \\
 TG(V_n) &= \sum_{i=1}^{n_n} TG(v_{ni})
 \end{aligned} \tag{3.1}$$

Trong đó,  $TG(V_i), i = \overline{1..n}$  là thời gian thực thi công việc  $i$ . Để tối ưu hóa hiệu năng hệ thống, hàm mục tiêu thứ nhất được xây dựng như sau:

$$J_1 = \sum_{i=1}^n TG(V_i) \rightarrow \min \tag{3.2}$$

Với sự phát triển ngày càng cao, các thiết bị di động trong lưới có khả năng thực hiện hàng triệu phép tính mỗi giây (MIPS) nhưng những công việc có thể được xử lý của một nút trong lưới vẫn bị giới hạn cụ thể về số lượng bởi khả năng xử lý của nút đó. Giả sử công việc phụ thứ  $i$  của công việc  $k$  là  $v_{ik}$  có kích thước  $S_{ik}$  MIPS và tốc độ xử lý của nút  $d_k$  trong lưới là  $S_k$  MIPS thì thời gian hoàn thành công việc phụ được xác định như [16].

$$TG(v_{ik}) = \frac{S_{ik}}{S_k} \tag{3.3}$$

Trong trường hợp, một máy phụ được gán số lượng công việc phụ vượt quá khả năng xử lý của máy thì số lượng công việc phụ chưa được xử lý sẽ được phân phối lại trong số các máy phụ có số lượng công việc được yêu cầu xử lý nhỏ hơn khả năng của máy.

Gọi  $NS_k$  là số công việc phụ chưa được phân bổ trên nút  $d_k$  của lưới do quá tải thì:

$$NS_1 = \sum_{k=1}^m NS_k, \forall NS_k > 0 \quad (3.4)$$

Trong đó,  $m$  là số lượng máy phụ trong lưới,  $NS_1$  là tổng số công việc phụ chưa được thực hiện của lưới do quá tải.

Gọi  $NS_2$  là số lượng các công việc còn có thể thực hiện được bởi lưới. Khi đó:

$$NS_2 = \sum_{k=1}^m NS'_k, \forall NS'_k < 0, NS_k = |NS'_k| \quad (3.5)$$

Trong đó,  $NS'_k$  là số công việc phụ có thể được phục vụ bởi nút dưới tải trong lưới.

Do ban đầu, máy chính phân phối các công việc tới các lưới là ngẫu nhiên nên khi đó, có ba khả năng có thể xảy ra:

- 1)  $NS_1 = NS_2$ : Các công việc được phân phối tới lưới một cách suôn sẻ.
- 2)  $NS_1 < NS_2$ : Số lượng công việc được phân bổ tới lưới nhỏ hơn khả năng xử lý của lưới, khi đó, việc phân bổ này là thành công.
- 3)  $NS_1 > NS_2$ : Số lượng công việc được phân bổ tới lưới lớn hơn năng lực xử lý của lưới, như vậy, một số công việc không được phân bổ một cách chính xác tới các nút.

Sau thời điểm đầu tiên, các công việc chưa được cấp phát máy phụ để thực thi sẽ được trả lại cho máy chính. Lúc này máy chính đã biết được năng lực của từng lưới nên bộ lập lịch toàn cục sẽ phân bổ công việc tối ưu hơn.

Trong trường hợp có nút vào, ra mạng, lúc này bộ lập lịch toàn cục sẽ sử dụng phương pháp heuristic để phân bổ công việc tới các lưới và quá trình được lặp lại cho đến khi các công việc được thực thi hoàn toàn.

Gọi  $TG_{EX}$  là thời gian chờ mong đợi của việc  $v_j$  trên máy  $M_j$  thì:

$$TG_{EX} = \frac{\lambda_j}{\mu_j(\mu_j - \lambda_j)} \quad (3.6)$$

Gọi  $TG_{ES}$  là thời gian phục vụ trung bình của máy  $M_j$  thì:

$$TG_{ES} = \frac{1}{\mu_j} \quad (3.7)$$

Khi đó, tổng thời gian thực thi công việc  $i$  tại máy  $j$  là  $TG(i, j)$  được tính:

$$TG(i, j) = \sum_{i=1}^m (TG_{EX} + TG_{ES}) * \delta_{ij} * S_i \quad (3.8)$$

Trong đó,  $m$  là số công việc phụ của công việc  $i$ ,  $S_i$  là kích thước của công việc  $i$  (MISP).  $\delta_{ij}$  là một biến Boolean, ở đó:

$$\delta_{ij} = \begin{cases} 0 & \text{Nếu công việc } i \text{ được gán cho máy } j \\ 1 & \text{Nếu công việc } i \text{ không được gán cho máy } j \end{cases}$$

Khi đó, thời gian của lịch biểu được tính:

$$TG = \max_{1 \leq i \leq R} TG(i, j) \quad (3.9)$$

Tỷ lệ tăng tốc là tỷ lệ giữa các thời gian hoàn thành các công việc và thời gian của lịch biểu:

$$SP_u = \frac{\text{Thời gian hoàn thành các công việc}}{TG} \quad (3.10)$$

Mục tiêu thứ hai của bài toán là giảm thời gian tăng tốc, tức là thời gian chờ và thực thi giảm. Hàm mục tiêu:

$$J_2 = \frac{SP_u}{nCPU} \rightarrow \min \quad (3.11)$$

Trong đó,  $nCPU$  là số lượng CPU của các nút trong lưới. Công thức (3.11) là hàm mục tiêu thứ ba nhằm giảm tỷ lệ địa phương hóa.

Tiếp theo, một đồ thị không trọng số được sử dụng để xác định thông tin kết nối địa phương tới trạm cơ sở địa phương. Các đỉnh của đồ thị là các nút tính toán của lưới tính toán di động và các cạnh là các liên kết vô tuyến từ trạm cơ sở tới nút. Để xác định một nút thuộc lưới tính toán, cần có thông tin về địa phương và số lượng hàng xóm trung bình trên nút đó. Thông tin về địa phương là tỷ lệ địa phương hóa là một hàm theo  $e$  được tính:

$$J_3 = R_{LR} = \frac{R * e - local - Z_1}{e^{1-local*\pi*(\cos(t-t\sqrt{1-t^2}))}} \rightarrow \min \quad (3.12)$$

Trong đó  $local$  là tham số địa phương ngẫu nhiên trong khoảng  $[0, 1]$  và  $local = rand(0, 1)$ .  $R$  miền phủ sóng vô tuyến lớn nhất.  $Z_1$  khoảng cách lớn nhất từ trạm cơ sở tới nút trong lưới.

Thông lượng của toàn hệ thống là thông lượng lớn nhất mà trạm cơ sở gán các công việc tới các lưới và thông lượng của từng lưới:

$$J_4 = sBW = \max(BW_i, G_i) \rightarrow \max, \quad i \in \overline{1..n} \quad (3.13)$$

Trong đó,  $sBW$  là thông lượng của hệ thống,  $n$  là số lưới phụ trong mạng, giá trị

$BW_i$  là thông lượng từ trạm tổng tới các lưới.  $BW_i$  thường nhỏ hơn thông lượng thực sự mà hệ thống có thể cung cấp nhằm đảm bảo tốc độ của hệ thống.

$$BW_i = \max_{j \in \overline{1..n}} BW_j, \quad (3.14)$$

và  $G_i$  là thông lượng tại các lưới phụ, được tính:

$$G_i = \frac{n_i * S_{nk}}{\sum_{j=1}^m M_j} \quad (3.15)$$

Trong đó,  $n_i$  là số lượng công việc phụ được gửi tới lưới  $i$ ,  $S_{nk}$  là kích thước của công việc phụ  $n$  của công việc  $k$ ,  $m$  là số lượng máy của lưới  $i$ .

Như vậy, mục tiêu của bài toán lập lịch trong lưới tính toán di động là tối thiểu thời gian tăng tốc theo công thức (3.11), tối thiểu tỷ lệ địa phương hóa theo công thức (3.12) và tối đa thông lượng theo công thức (3.13).

### 3.4 Thuật toán lập lịch công việc trong lưới tính toán di động dựa trên tìm kiếm meta-heuristic - HGLA

Thuật toán HGLA đề xuất gồm hai pha: (1) bộ lập lịch toàn cục tại máy chính nhận tất cả các công việc với các công việc phụ được gửi bởi người dùng và phân bổ tới các lưới sử dụng heuristic dựa trên quá trình địa phương hóa với sự trợ giúp của bộ lập lịch địa phương; (2) bộ lập lịch địa phương sử dụng MIN-MIN, MAX-MIN hay HEFT để lựa chọn máy phụ để phân bổ các công việc phụ.

Phương pháp thực hiện của thuật toán được mô tả như sau:

Máy chính sử dụng định dạng của các máy phụ trong lưới tính toán và phân bổ ngẫu nhiên. Năng lực tính toán và tốc độ mỗi máy phụ trong lưới tính toán di động, thời gian thực hiện của mỗi công việc phụ cũng được biết trước. Trong quá

trình lập lịch, nếu một máy phụ nào đó được phân bổ số lượng công việc vượt quá khả năng của nó thì những công việc phụ đó sẽ được trả lại máy chính. Máy chính tìm các máy phụ đã được phân bổ số lượng công việc phụ nhỏ hơn công suất của chúng để gán các việc phụ chưa được thực thi này. Toàn bộ kết quả trên máy phụ sẽ được chuyển về máy chính.

Để cài đặt thuật toán, luận án đã sử dụng các cấu trúc dữ liệu khác nhau như sau:

- $R_{List}$  là mảng chứa tất cả các công việc phụ cần được phân bổ lại.
- $NM_c$  là tổng số máy thứ cấp sẵn sàng nhận nhiệm vụ để thực hiện.
- Mảng hai chiều  $Alloc$ , nếu  $Alloc[i, j] = k$  có nghĩa là công việc con thứ  $j$  của công việc thứ  $i$  được thực hiện trên máy  $k$ .
- Mảng một chiều  $CMc$  chứa số lượng công việc con có thể được thực hiện tại một thời điểm trên máy phụ.
- $List[i]$  là cấu trúc dữ liệu chứa các công việc phụ được gán cho máy  $M$ .
- $SMc$  là mảng được sử dụng để lưu trữ tốc độ xử lý của tất cả các máy thứ cấp.
- Số lượng công việc và các công việc phụ của nó được lưu trong  $Mjobs$  và  $Msjobs$ .
- Biến tạm  $Tmp$  lưu trữ số lượng công việc vượt quá khả năng của một máy cụ thể.

Các bước của thuật toán HGLA tại pha 1 trong MCG như sau:

- *Bước 1*: Khởi tạo tất cả các công việc và các công việc phụ của nó tới máy chính  $PM$ .

- *Bước 2:* Cho tất cả các phần tử của  $R_{List}$  trở tới NULL; khởi tạo tổng thời gian xoay vòng và tổng thời gian chờ ban đầu là 0.
- *Bước 3:* Máy chính phân phối các công việc con một cách ngẫu nhiên cho tất cả các máy thứ cấp trong lưới sử dụng phương pháp heuristic địa phương trong bộ lập lịch địa phương và heuristic toàn cục trong bộ lập lịch toàn cục.
- *Bước 4:* Nếu số lượng công việc con trên máy thứ cấp lớn hơn công suất của máy thì thêm các công việc gây quá tải cho máy vào  $R_{List}$ .
- *Bước 5:* Tìm các máy thứ cấp còn dư thừa năng lực tính toán.
- *Bước 6:* Phân phối các công việc con từ  $R_{List}$  cho các máy phụ được sử dụng không đúng bằng cách sử dụng phương pháp dựa theo heuristic địa phương trong bộ lập lịch địa phương và heuristic toàn cục trong bộ lập lịch toàn cục.
- *Bước 7:* Tính tỷ lệ tăng tốc, tỷ địa phương hóa và thông lượng theo các công thức (3.10), (3.13) và (3.12).

Giả mã của của HGLA pha 1 được trình bày như trong thuật toán (3.1).

Các bước của thuật toán HGLA tại pha 2 trong MCG như sau:

- *Bước 1:* Khởi tạo các công việc và các công việc phụ của nó và gán cho máy chính.
- *Bước 2:* Phần còn lại  $R_{List}$  được khởi tạo là *NULL*, tổng thời gian quay vòng và tổng thời gian chờ được khởi tạo là 0
- *Bước 3:* Phân phối các việc phụ ngẫu nhiên tới các máy phụ sử dụng MIN-MIN, MAX-MIN, HEFT.
- *Bước 4:* Nếu số lượng các công việc phụ trên máy phụ lớn hơn khả năng của máy thì chuyển các công việc không được thực hiện vào  $R_{List}$
- *Bước 5:* Tìm các máy thứ cấp còn dư thừa năng lực tính toán.

**Thuật toán 3.1** Thuật toán HGLA - Pha 1 (tại máy chính)

---

**Đầu vào:** Tập  $V$  công việc cần lập lịch

**Đầu ra:** Tỷ lệ địa phương hóa, thông lượng, và tỷ lệ tăng tốc

---

```
1: Gán tất cả các công việc tới máy chính  $PM$ 
2:  $R_{List} \leftarrow NULL$ 
3: Khởi tạo  $T = W = 0$ 
4: while  $R_{List} \langle \rangle NULL$  do
5:   for  $i \leftarrow 1$  to  $NMc$  do
6:      $Alloc[i, j] \leftarrow Random.Math * NMc$ 
7:     Xây dựng danh sách  $List[i]$  chứa các công việc phụ được gán cho máy
        $M[i]$ 
8:   end for
9:   for  $i \leftarrow 1$  to  $NMc$  do
10:    if  $List[i].Size = CMc[i]$  then
11:      for  $j \leftarrow CMc[i]$  to  $List[i]$  do
12:        Thêm các công việc phụ vào  $List[i]$ 
13:      end for
14:    end if
15:  end for
16:  for  $i \leftarrow 1$  to  $NMc$  do
17:    if  $List[i].Size < CMc[i]$  then
18:      Thêm các công việc con từ chưa được xử lý vào  $List[i]$ 
19:    end if
20:     $T \leftarrow T +$  thời gian quay vòng của công việc phụ đã hoàn thành
21:     $W \leftarrow W +$  tổng thời gian đã sử dụng để hoàn thành công việc đang chờ
       xử lý
22:  end for
23:  Tính tỷ lệ tăng tốc, tỷ lệ địa phương hóa và thông lượng
24: end while
```

---

— *Bước 6:* Phân phối các công việc con từ  $R_{List}$  tới các máy dưới tải sử dụng MIN-MIN, MAX-MIN, HEFT

— *Bước 7:* Tính tỷ lệ tăng tốc, tỷ lệ địa phương hóa và thông lượng theo các công thức (3.10), (3.12) và (3.13)

Giải mã của HGLA tại pha 2 được trình bày như trong thuật toán (3.2).

---

**Thuật toán 3.2** Thuật toán HGLA - Pha 2 (tại máy phụ)

---

**Đầu vào:** Tập  $V$  công việc cần lập lịch

**Đầu ra:** Tỷ lệ địa phương hóa

---

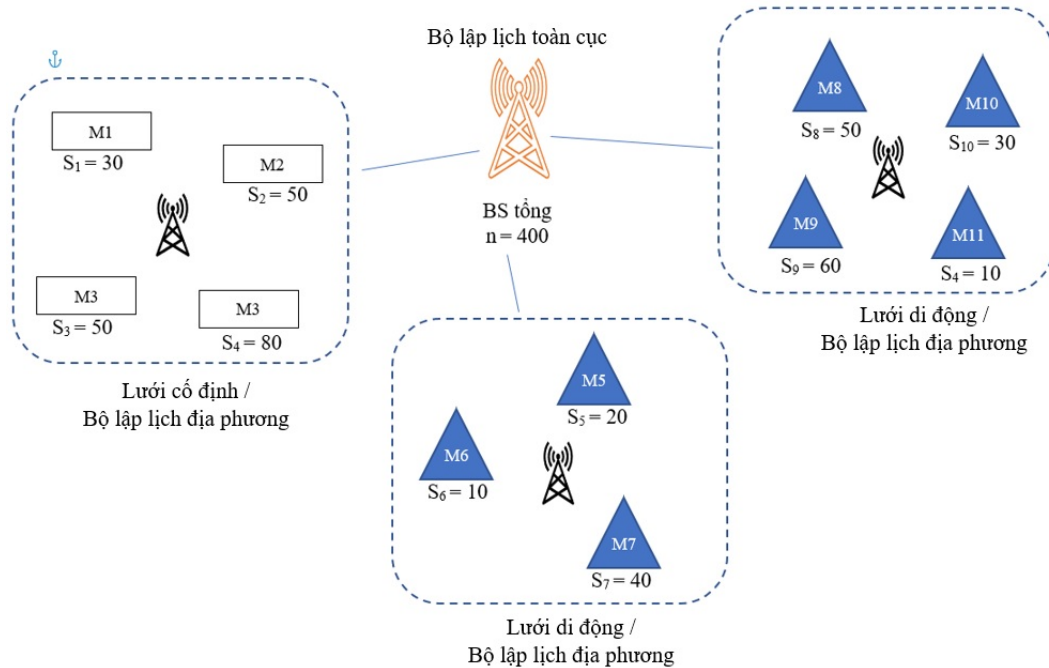
```
1: Gán tất cả các công việc tới máy chính  $PM$ 
2: Khởi tạo phần còn lại của  $R_{List} \leftarrow NULL$ 
3: Khởi tạo  $T = W = 0$ 
4: while  $R_{List} \langle \rangle NULL$  do
5:   for  $i \leftarrow 1$  to  $N_{sjob}[i]$  do
6:     Thêm tất cả công việc phụ vào  $ListA$ 
7:   end for
8:   if  $A.Size \neq NULL$  then
9:     for  $i \leftarrow 0$  to  $NMc$  do
10:       $Tmp \leftarrow 0$ 
11:    end for
12:    end if
13:    for  $i \leftarrow 1$  to  $NMc$  do
14:      if  $Njob$  phân bổ cho máy  $M[i] > CMc[i]$  then
15:         $Tmp[i] \leftarrow Tmp[i] + (Njob - CMc[i])$ 
16:      end if
17:      Tìm công việc phụ tốt nhất bằng MIN-MIN, MAX-MIN, HEFT và phân
        bổ cho máy  $M$  để thực hiện
18:       $T \leftarrow T +$  thời gian quay vòng của công việc phụ đã hoàn thành
19:       $W \leftarrow W +$  tổng thời gian đã sử dụng để hoàn thành công việc đang chờ
        xử lý
20:    end for
21:    Tính tỷ lệ tăng tốc, thông lượng và tỷ lệ địa phương hóa
22: end while
```

---

### 3.5 Ví dụ minh họa

Giả sử có 400 công việc cần lập lịch đến các lưới và các máy trong lưới. Lưới sử dụng để lập lịch và năng lực các máy trong lưới được mô tả như hình (3.4).

- Tại thời điểm  $t_0$  (bắt đầu), bộ lập lịch toàn cục phân phối ngẫu nhiên tài nguyên cho các lưới như sau:  $n_1 = 250$  công việc cho lưới cố định,  $n_2 = 50$  công việc cho lưới di động 1,  $n_3 = 100$  công việc cho lưới di động 2.



**Hình 3.4.** Lưới tính toán di động và năng lực các máy được sử dụng trong ví dụ

- Tại  $t_0$  phương án lập lịch sử dụng bộ lập lịch địa phương tại lưới cố định là: 30, 50, 50, 80 công việc.

- Phương án lập lịch sử dụng bộ lập lịch địa phương tại lưới di động 1 là: 5, 10, 35 công việc.

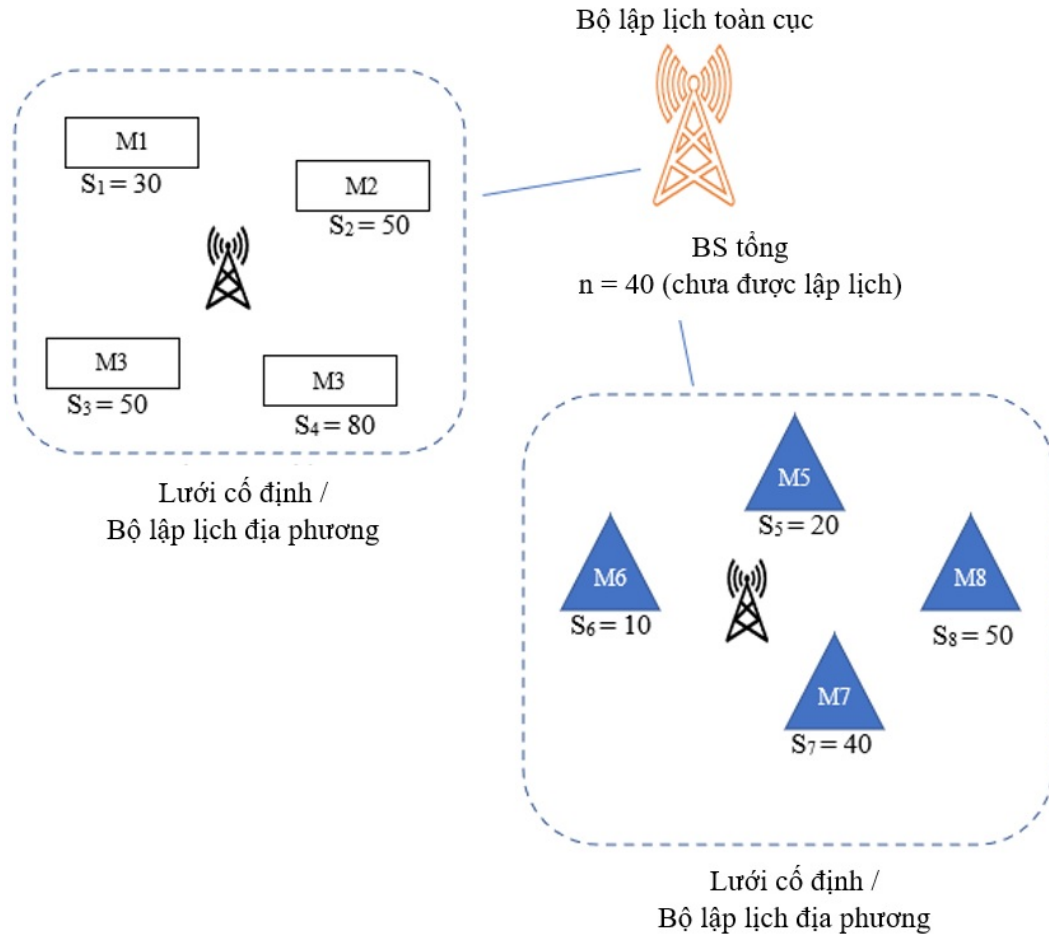
- Phương án lập lịch sử dụng bộ lập lịch địa phương tại lưới di động 2 là: 30, 30, 30, 10 công việc.

Tổng các công việc chưa được phân bổ  $K_1$ , tổng năng lực dư thừa  $K_2$  trong lưới được tính như sau:

- Tại lưới cố định:  $K_1 = 40, K_2 = 0$

- Tại lưới di động 1:  $K_1 = 0, K_2 = 20$

- Tại lưới di động 2:  $K_1 = 0, K_2 = 50$



**Hình 3.5.** Lưới tính toán di động và năng lực các máy được cấu trúc lại sau khi có máy vào, ra lưới

Nhận xét:

+ Với phương án phân bổ tài nguyên tại vòng 1, lưới cố định với năng lực 210, nhưng được phân bổ 250 công việc, lúc này tình trạng quá tải xảy ra. Lưới di động 1 với năng lực 70, được phân bổ 50 công việc dẫn đến tình trạng dư thừa năng lực tính toán. Lưới di động 2 với năng lực 150, được phân bổ 100 công việc và do đó, dẫn đến dư thừa năng lực tính toán.

Tại thời điểm  $t_1$ , lúc này số lượng công việc chưa được thực hiện được trả lại BS tổng là 40. Có 2 trường hợp xảy ra:

1) Trường hợp 1: Giả sử không có công việc mới nào được phát sinh, mạng ổn định sau vòng 1, tỷ lệ địa phương hóa tại các lưới được tính. Dựa trên tỷ lệ địa phương hóa đã được tính toán, BS tổng sẽ phân lại tài nguyên cho các lưới dựa trên bộ lập lịch toàn cục.

2) Trường hợp 2: Một số nút di động rời mạng, giả sử mạng còn lại hai lưới như hình (3.5), lúc này bộ lập lịch toàn cục sẽ phân bổ tài nguyên cho 2 lưới này và bắt đầu tính lại tỷ lệ địa phương hóa. Vòng lặp kết thúc khi số lượng công việc được thực hiện hết.

Thời gian thực thi một lịch biểu bằng thời gian thực thi lớn nhất của các công việc con được thực hiện tại các lưới.

Trong phần tiếp theo của luận án, các kết quả thực nghiệm và đánh giá được đưa ra thảo luận. Thuật toán đề xuất HGLA được thực nghiệm và so sánh với các phương pháp truyền thống MIN-MIN, MAX-MIN và HEFT theo các tiêu chí tăng tốc, tỷ lệ địa phương hóa và thông lượng hệ thống.

### 3.6 Thực nghiệm đánh giá kết quả

Mô hình đề xuất được cài đặt trên Eclipse và GridSim để kiểm chứng và phân tích hiệu năng. Kích thước công việc chính và công việc phụ cho xác định sự hội tụ của giải pháp đề xuất. Các tham số được sử dụng để mô phỏng trong thực nghiệm được đưa ra trong bảng (3.2). Thực nghiệm được thực hiện 40 lần chạy với ổ cứng 1TB và 16 GB RAM và giá trị trung bình được sử dụng để đánh giá cho mọi quan sát. Các tiêu chí đánh giá trong thử nghiệm bao gồm tỷ lệ tăng tốc theo công thức (3.10), tỷ lệ địa phương hóa theo công thức (3.12) và thông lượng theo công thức (3.13). Theo đó, tỷ lệ tăng tốc càng nhỏ càng tốt, tỷ lệ địa phương hóa càng nhỏ càng tốt, thông lượng càng lớn càng tốt.

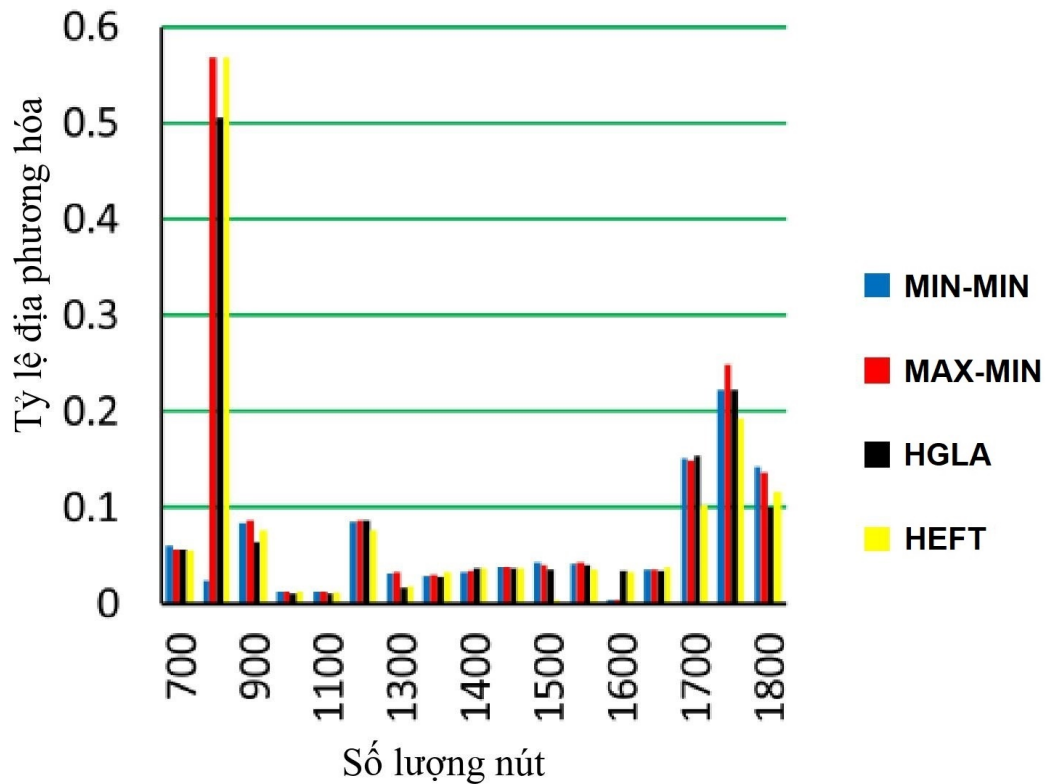
**Bảng 3.2.** Các tham số được sử dụng mô phỏng thuật toán HGLA [92]

| Tham số                                  | Giá trị          |
|--|------------------|
| Số lượng nút                             | 100-1800         |
| Số lượng công việc                       | 100-1000         |
| Số lượng công việc con của một công việc | 10-100           |
| Tỷ lệ đến                                | 1-100 MIPS       |
| Tốc độ thực thi                          | 10-500 MIPS      |
| Kích thước công việc                     | 1000-4000 MI     |
| Kích thước công việc phụ                 | 20-100 MI        |
| Kích thước bộ nhớ của một nút            | 4-32 GB          |
| Kích thước bộ nhớ đệm Cache              | 2 - 128 KB       |
| Số lượng thanh ghi                       | 16-256           |
| Số lượng Bus                             | 2-64             |
| Tốc độ lập lịch địa phương               | 2-2048 công việc |
| Tốc độ lập lịch toàn cục                 | 1-512 công việc  |

### 3.6.1 Thử nghiệm tỉ lệ địa phương hóa trong trường hợp lưới có nút di động và không có nút di động

Thử nghiệm này đã sử dụng số lượng nút thay đổi cho các lưới trong mạng tại mỗi thực nghiệm. Kết quả thử nghiệm về tỷ lệ địa phương hóa của thuật đề xuất được so sánh với các phương pháp truyền thống như MIN-MIN, MAX-MIN và HEFT. Trong thực nghiệm, với hầu hết các trường hợp, tỷ lệ địa phương hóa của thuật toán đề xuất cho kết tốt hơn các thuật toán được so sánh, do thuật

toán đề xuất có sử dụng các nút động và tĩnh trong lưới. Trong quá trình lập lịch, lưới được tái cấu trúc lại do có các nút động rời hoặc ra nhập mạng. Hình (3.6) là kết quả so sánh giữa tỷ lệ địa phương hóa của thuật toán đề xuất và các thuật toán khác.



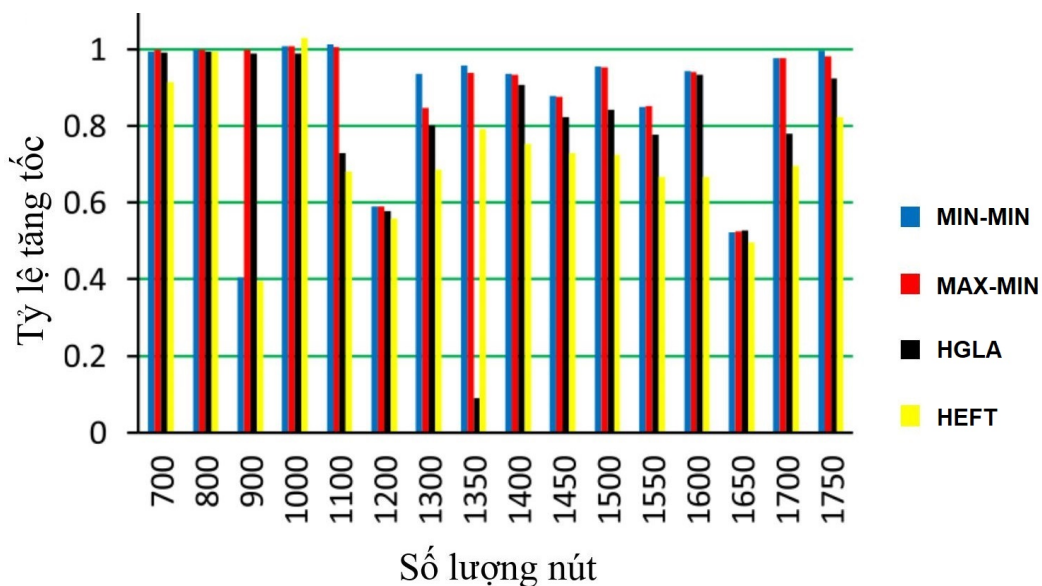
Hình 3.6. Thử nghiệm với tỷ lệ địa phương hóa

Tỉ lệ địa phương hóa đo kết quả việc thực hiện lập lịch được tiến hành ở các máy chủ cục bộ hay toàn cục. Việc này sẽ quyết định hiệu quả của việc phân phối công việc được xử lý ở cấp độ nào. Nếu toàn bộ tác vụ được tiến hành ở máy chủ cục bộ thì sẽ dẫn đến có những máy chủ cục bộ quá tải trong khi một số máy nhàn rỗi (phân công không đều). Nếu toàn bộ tác vụ được thực hiện ở máy chủ toàn cục thì còn kém hiệu quả hơn do nhiệm vụ chính của máy chủ toàn cục là phân phối công việc cho các máy chủ cục bộ và trong một số trường hợp hãn hữ thực hiện tác vụ. Bởi vậy, tỉ lệ địa phương hóa rất quan trọng trong đánh giá

tính hiệu quả của thuật toán. Khi số lượng nút càng lớn thì tỉ lệ địa phương hóa gần có xu hướng giảm, trung bình khoảng 10% tổng số tác vụ (đối với HGLA và HEFT). So với thuật toán MIN-MIN có xu hướng giảm tỉ lệ địa phương hóa thì thuật toán HGLA và HEFT có tỉ lệ cao hơn. Điều này phản ánh rõ đặc điểm của thuật toán đề xuất.

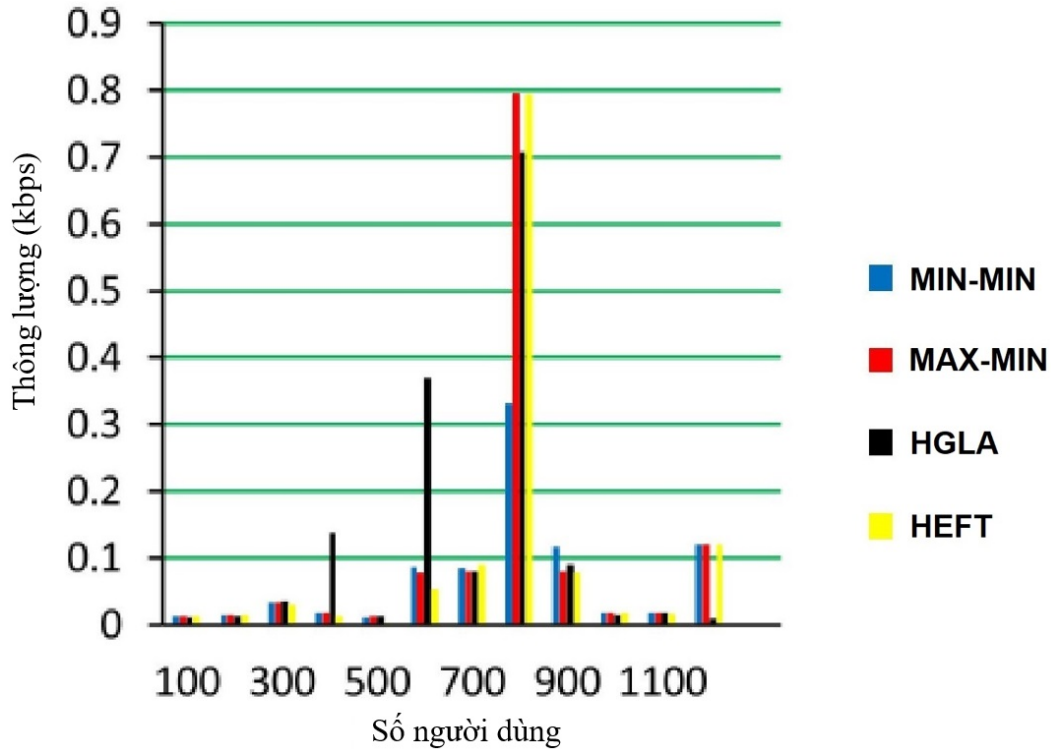
### 3.6.2 Thử nghiệm chỉ lệ tăng tốc với trường hợp lưới có số lượng nút thay đổi

Hình (3.7) mô tả kết quả thực nghiệm của thuật toán đề xuất được so sánh với các phương pháp truyền thống MIN-MIN, MAX-MIN và HEFT theo tham số tăng tốc. Trong hầu hết các trường hợp, tỷ lệ tăng tốc của thuật toán đề xuất cho kết quả tốt hơn các phương pháp được so sánh nhưng vẫn kém hơn so với MIN-MIN do số lượng công việc có thời gian thực hiện ngắn chiếm tỷ trọng lớn hơn trong danh sách các công việc được lập lịch.



Hình 3.7. Thử nghiệm với tỷ lệ tăng tốc

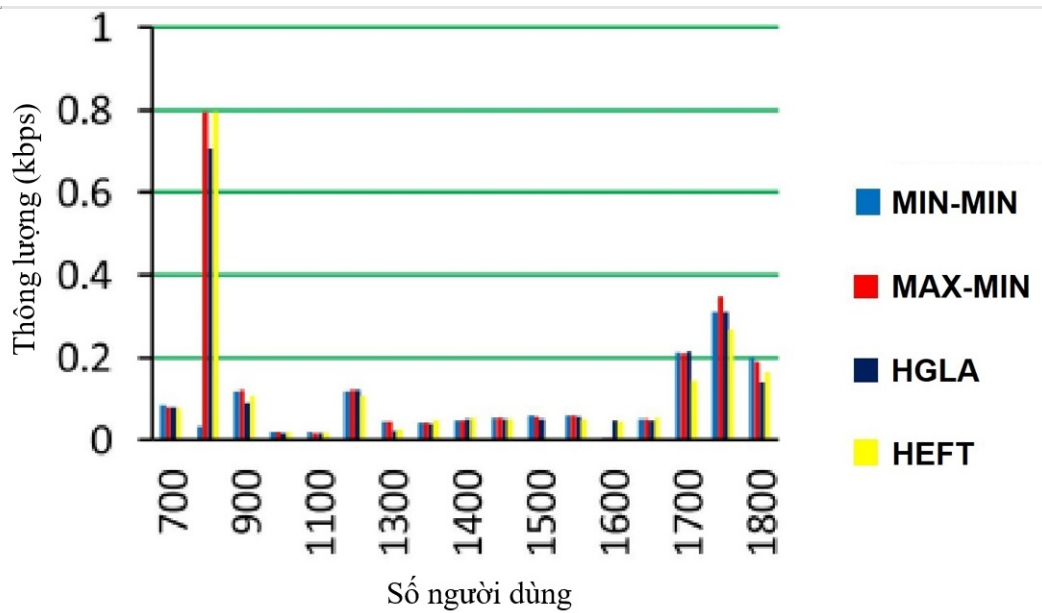
3.6.3 Thử nghiệm thông lượng với trường hợp lưới có số lượng nút thay đổi



Hình 3.8. Thử nghiệm thông lượng của thuật toán với số nút từ 100 tới 1100

Kết quả thử nghiệm về tham số thông lượng của thuật toán đề xuất được so sánh với các phương pháp truyền thống như MIN-MIN, MAX-MIN và HEFT. Hình (3.8) và hình (3.9) cho biết sự thay đổi của băng thông giữa thuật toán đề xuất và các thuật toán được so sánh.

Kết quả so sánh cho thấy sau một số lần lặp và với một số lượng công việc ổn định, thông lượng tăng lên một cách nhanh chóng nhưng sau đó lại bắt đầu giảm. Thông lượng tối đa đạt được khi số lượng nút trong lưới vào khoảng 800 nút, đây cũng là con số được thiết kế để tối ưu hóa trong trường hợp tham số thông lượng là một tham số được coi là mục tiêu đánh giá chất lượng dịch vụ trong lưới tính toán di động.



Hình 3.9. Thử nghiệm thông lượng của thuật toán với số nút từ 700 đến 1800

### 3.7 Ưu và nhược điểm của thuật toán đề xuất

Thuật toán đề xuất có một số ưu điểm sau:

- Thuật toán đề xuất HGLA loại bỏ vấn đề “đói” tài nguyên bằng cách phân phối các công việc đồng đều giữa các nút trong lưới tính toán di động.
- Thuật toán đề xuất HGLA xem xét phân phối công việc giữa các tài nguyên không đồng nhất giữa các mạng.
- Lưới tính toán di động được xét trong luận án gồm cả nút tĩnh và nút động.
- Việc phân phối lại các công việc được thực hiện theo cách đảm bảo cân bằng tải giữa các nút.

Mặc dù có nhiều đóng góp cải thiện nâng cao hiệu quả của bài toán lập lịch công việc. Tuy nhiên, thuật toán đề xuất vẫn còn có các hạn chế như thuật toán lập lịch đơn lẻ. Giống như các thuật toán lập lịch khác, thuật toán đề xuất tập

trung nhiều vào cải thiện hiệu năng bằng cách tăng thời gian thực thi, giảm thời gian chờ của các công việc mà chưa tập trung vào cải tiến năng lượng.

Do đặc thù của lưới tính toán di động là ở đó hiện diện cả nút tĩnh và nút động, trong đó các nút động là các thiết bị di động thông minh bị ràng buộc về mặt năng lượng do chỉ được trang bị một viên pin nhỏ. Nhược điểm này của thuật toán đề xuất trong luận án sẽ là sẽ là một nghiên cứu tiếp theo của tác giả.

### 3.8 Kết luận chương

Chương này đã đề xuất một phương pháp lập lịch mới trong lưới tính toán di động dựa trên heuristic (HGLA). Thuật toán đề xuất được thực hiện qua hai giai đoạn: giai đoạn một, công việc được phân bổ tới các lưới bởi bộ lập lịch toàn cục; giai đoạn hai, tại mỗi lưới, tài nguyên được phân phối tới mỗi nút xử lý theo sự phân bổ của bộ lập lịch địa phương. Kết quả của thuật toán đề xuất được so sánh với một số thuật toán khác như MIN-MIN, MAX-MIN, HEFT.

Ban đầu, bộ lập lịch cục bộ phân phối công việc một cách ngẫu nhiên vào các lưới nên sẽ có trường hợp lưới bị quá tải, khi đó các công việc chưa được cấp phát tài nguyên để thực hiện sẽ được trả lại máy chính. Nếu mạng không thay đổi, máy chính sẽ phân bổ lại các công việc chưa được thực hiện tới các lưới để hoàn thành. Tuy nhiên, trong thực tế, có thể mạng có thêm các nút mới gia nhập hoặc có nút rời hệ thống. Máy chính sẽ phải phân bổ lại tài nguyên dựa trên heuristic trước đó.

Mô phỏng được thực nghiệm trên số lượng nút khác nhau để chứng minh hiệu suất của mô hình. Các thực nghiệm cho thấy kết quả của phương pháp đề xuất cải thiện hơn một số phương pháp truyền thống về tỷ lệ địa phương hóa, thông lượng, khả năng tăng tốc trong một số trường hợp nhất định

Mặc dù đã đạt được một số cải tiến cho việc lập lịch công việc trong môi trường tính toán lưới di động, nhưng mô hình đề xuất chưa xét đến cải tiến về mặt tiết kiệm năng lượng. Đôi khi các nút được phát hiện là sử dụng năng lượng quá mức và do đó có rất nhiều nhiệt được tạo ra, điều này dẫn đến một lượng lớn chi phí được chi cho việc làm mát. Đây cũng là một hướng phát triển bằng cách tập trung cải thiện thuật toán để giảm năng lượng tiêu thụ, đồng thời tập trung giải quyết các thông số chất lượng dịch vụ khác như cân bằng tải, độ bảo mật, độ tin cậy, tính sẵn sàng, tính di động, khả năng chịu lỗi bằng cách sử dụng heuristic địa phương.

Kết quả của chương này đã được công bố tại [CT4] trong danh mục các công trình khoa học của tác giả có liên quan đến luận án.

# KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

## Kết luận của luận án

Lập lịch là bài toán cấp phát nguồn lực, tài nguyên để thực hiện công việc trên một chuỗi các tiến trình trên cơ sở thời gian, tài nguyên và các ràng buộc đã được định sẵn. Do đó, lập lịch được coi như tìm kiếm một giải pháp tối ưu trong các điều kiện hạn chế. Trong các hệ thống phân tán, lập lịch là một bài toán có ý nghĩa thực tiễn với mục đích làm giảm chi phí cho cả người sử dụng dịch vụ và người cung cấp dịch vụ. Các thuật toán lập lịch cho các hệ thống lớn hiện nay chứa đựng nhiều yếu tố phức tạp ảnh hưởng đến chất lượng lập lịch chưa được giải quyết dẫn đến chưa mang lại hiệu quả và tối ưu về chi phí. Chính vì vậy, việc nghiên cứu đề xuất cải tiến thuật toán lập lịch trở nên vô cùng cấp thiết và có giá trị thực tiễn.

Với mục tiêu nghiên cứu đề xuất cải tiến thuật toán lập lịch, các kết quả nghiên cứu tại chương 2, chương 3 đã đưa ra các đóng góp của đề tài như sau:

- Thuật toán lập lịch cho đường tải xuống trong mạng di động đa dịch vụ kết hợp giữa miền thời gian và miền tần số (ITFDS). Thuật toán lập lịch đề xuất phân phối tài nguyên cho tất cả người dùng theo miền thời gian và miền tần số. Kết quả mô phỏng phương pháp đề xuất được so sánh với các thuật toán QuAS, PSS về các tham số như độ trễ, tỷ lệ mất gói, tổng thông lượng và

chỉ số công bằng. Kết quả của thuật toán đề xuất đã được công bố tại [CT2, CT3] trong danh mục các công trình khoa học có liên quan đến luận án.

- Đề xuất thuật toán lập lịch phân phối tài nguyên cho lưới tính toán di động dựa trên tìm kiếm heuristic. Thực nghiệm trên phương pháp đề xuất được so sánh với MIN-MIN, MAX-MIN và HEFT trên các tham số dịch vụ như tỷ lệ tăng tốc, thông lượng và tỷ lệ địa phương hóa. Kết quả của phương pháp đề xuất đã được công bố tại [CT4] trong danh mục các công trình khoa học có liên quan đến luận án.

Bên cạnh các kết quả đạt được của các phương pháp đề xuất thì vẫn còn có những hạn chế nhất định của các kết quả nghiên cứu này. Với thuật toán ITFDS, thuật toán mới thử nghiệm trên môi trường mạng 4G-LTE mà chưa thử nghiệm với môi trường mạng 5G hay tỷ lệ dữ liệu truyền bị giảm với các người dùng ở xa trạm cơ sở, hay mô phỏng mới chỉ so sánh với phương pháp QuAS và PSS mà chưa thử nghiệm với các phương pháp cùng loại khác. Trong thuật toán HGLA, mặc dù đạt được kết quả về mặt hiệu suất, nhưng chưa có nhiều cải tiến về mặt năng lượng.

### **Hướng phát triển**

Ngoài những kết quả đạt được đã nêu trong luận án, một số hướng phát triển liên quan đến luận án mà tác giả hướng tới như tích hợp khả năng tính toán trong các thiết bị IoT và lấy mẫu nén dựa trên định tuyến và tái tạo dữ liệu nhằm tiết kiệm năng lượng cho mạng.

Tích hợp khả năng tính toán trong IoT sẽ mang tới một bước nhảy vọt trong các lĩnh vực như y tế, quân sự, gia đình, giải trí, v.v. IoT bao gồm sự kết hợp giữa các công nghệ khác nhau như điện toán đám mây, mạng cảm biến không dây, dữ liệu lớn và thông tin dữ liệu [2, 80], trong đó, nhiệm vụ chính của

các thành phần IoT (như cảm biến, điện thoại và thẻ RFID) là cảm nhận, thu thập và lưu trữ dữ liệu sau đó gửi chúng đến trạm cơ sở [16].

Việc tìm ra các kỹ thuật nhằm kéo dài tuổi thọ mạng là rất cần thiết [37]. Nhiều kỹ thuật khác nhau đã được đề xuất như giao thức định tuyến hoặc lược đồ nén dữ liệu [43, 65]. Kỹ thuật này có thể được sử dụng để giảm năng lượng tiêu thụ bởi giao tiếp các nút được coi là một trong những nguồn tiêu thụ điện năng chính trong IoT. Kết quả của nghiên cứu này đã được công bố tại [4] trong thời gian làm luận án.

# DANH MỤC CÁC CÔNG TRÌNH KHOA HỌC ĐÃ CÔNG BỐ

[CT1]. Ishaani Priyadarshini, Raghvendra Kumar, **Le Minh Tuan**, Le Hoang Son, Hoang Viet Long, Rohit Sharma, Sakshi Rai (2021), *A new enhanced cyber security framework for medical cyber physical systems*. SICS Software-Intensive Cyber-Physical Systems, 35(3), pp.159-183, (**ESCI, Springer, ISSN:2524-8510**)

[CT2]. **Lê Minh Tuấn**, Lê Hoàng Sơn, Phạm Thị Minh Phương, Vũ Như Lân, Đặng Thanh Hải, Đinh Thu Khánh (2019), *Nghiên cứu đề xuất mô hình mạng động cho bài toán lập lịch tài nguyên trong mạng Long Term Evolution (LTE)*, Kỷ yếu Hội nghị Quốc gia lần thứ XII về Nghiên cứu cơ bản và ứng dụng Công nghệ thông tin (FAIR), Huế, ngày 7-8/6/2019, tr.345-352.

[CT3]. **Le Minh Tuan**, Le Hoang Son, Hoang Viet Long, Rajaretnam Priya, Ruba Soundar, Harold Robinson, Raghvendra Kumar (2020), *ITFDS: channel-aware integrated time and frequency-based downlink LTE scheduling in MANET*, Sensors, 20(12), pp.3394, (**SCIE, 2019, IF=3.847, ISSN: 1424-8220**)

[CT4]. Dinesh Prasad Sahu, Karan Singh, Manisha Manju, David Taniar, **Le Minh Tuan**, Le Hoang Son, Mohamed Abdel Basset, Hoang Viet Long (2019), *Heuristic search based localization in mobile computational grid*. IEEE Access, 7, pp.78652-78664, (**SCIE, 2019, IF=3.476, ISSN: 2169-3536**)

## TÀI LIỆU THAM KHẢO

- [1] Abdellatif A. A., Mohamed A., Chiasserini C. F., Tlili M., and Erbad A. (2019), “Edge computing for smart health: Context-aware approaches, opportunities, and challenges”, *IEEE Network* 33(3), pp. 196–203.
- [2] Al-Kashoash H. A., Kharrufa H., Al-Nidawi Y., and Kemp A. (2019), “Congestion control in wireless sensor and 6LoWPAN networks: toward the Internet of Things”, *Wireless Networks* 25(8), pp. 4493–4522.
- [3] Andrews M., Kumaran K., Ramanan K., Stolyar A., Whiting P., and Vijayakumar R. (2001), “Providing quality of service over a shared wireless link”, *IEEE Communications magazine* 39(2), pp. 150–154.
- [4] Aziz A., Singh K., Osamy W., Khder A., Tuan L. M., Son L. H., Long H. V., and Rakhmonov D. (2021), “Compressive sensing based routing and data reconstruction scheme for IoT based WSNs”, *Journal of Intelligent & Fuzzy Systems* 41, pp. 1–17.
- [5] Baker M., Buyya R., and Laforenza D. (2002), “Grids and Grid technologies for wide-area distributed computing”, *Software: Practice and Experience* 32(15), pp. 1437–1466.
- [6] Basilashvili G., Gogokhia S., et al. (2018), “Customized packet scheduling algorithm for LTE network”, *American Academic Scientific Research Journal for Engineering, Technology, and Sciences* 40(1), pp. 203–216.
- [7] Basukala R., Ramli H. M., and Sandrasegaran K. (2009), “Performance analysis of EXP/PF and M-LWDF in downlink 3GPP LTE system”, In *2009 First Asian Himalayas International Conference on Internet* , pp. 1–5. IEEE.
- [8] Bontu C. S. and Illidge E. (2009), “DRX mechanism for power saving in LTE”, *IEEE Communications Magazine* 47(6), pp. 48–55.
- [9] Capozzi F., Laselva D., Frederiksen F., Wigard J., Kovacs I. Z., and Mogenssen P. E. (2009), “UTRAN LTE downlink system performance under realistic control channel constraints”, In *2009 IEEE 70th Vehicular Technology Conference Fall* , pp. 1–5. IEEE.

- [10] Chana I. et al. (2013), “Bacterial foraging based hyper-heuristic for resource scheduling in grid computing”, *Future Generation Computer Systems* 29(3), pp. 751–762.
- [11] Chen F., Tang Y., Wang C., Huang J., Huang C., Xie D., Wang T., and Zhao C. (2021), “Medical cyber-physical systems: A solution to smart health and the state of the art”, *IEEE Transactions on Computational Social Systems* , pp. 1–28.
- [12] Čibej U., Sulistio A., and Buyya R. (2009), *Grid Computing* , pp. 117–145, Springer.
- [13] Condoluci M. and Mahmoodi T. (2018), “Softwarization and virtualization in 5G mobile networks: Benefits, trends and challenges”, *Comput. Netw.* 146, pp. 65–87.
- [14] Dardouri S. and Bouallegue R. (2015), “Comparative study of downlink packet scheduling for LTE networks”, *Wireless Personal Communications* 82(3), pp. 1405–1418.
- [15] Deelman E., Gannon D., Shields M., and Taylor I. (2009), “Workflows and e-Science: An overview of workflow system features and capabilities”, *Future generation computer systems* 25(5), pp. 528–540.
- [16] Dhumane A. and Prasad R. (2019), “Multi-objective fractional gravitational search algorithm for energy efficient routing in IoT”, *Wireless networks* 25(1), pp. 399–413.
- [17] Dolińska I., Jakubowski M., Masiukiewicz A., and Szeszko M. (2016), “Fairness calculation on the base of the station media access time in WI-FI networks”, In *2016 International Conference on Information and Digital Technologies (IDT)* , pp. 76–82. IEEE.
- [18] Dongarra J. J., Meuer H. W., Strohmaier E., et al. (1997), “TOP500 supercomputer sites”, *Supercomputer* 13, pp. 89–111.
- [19] Dorigo M. and Blum C. (2005), “Ant colony optimization theory: A survey”, *Theoretical computer science* 344(2-3), pp. 243–278.
- [20] Dutta M., Aggarwal N., et al. (2016), “Meta-Heuristics Based Approach for Work flow Scheduling in Cloud Computing: a Survey”, , pp. 1331–1345. Springer, New Delhi.
- [21] Fan Z., Shen H., Wu Y., and Li Y. (2013), “Simulated-annealing load balancing for resource allocation in cloud environments”, In *2013 International Conference on Parallel and Distributed Computing, Applications and Technologies* , pp. 1–6. IEEE.

- [22] Foster I., Kesselman C., and Tuecke S. (2001), “The anatomy of the grid: Enabling scalable virtual organizations”, *The International Journal of High Performance Computing Applications* 15(3), pp. 200–222.
- [23] Golberg D. E. (1989), “Genetic algorithms in search, optimization, and machine learning”, *Addion wesley* 1989(102), pp. 36.
- [24] Gonzalez N., Miers C., Redigolo F., Simplicio M., Carvalho T., Näslund M., and Pourzandi M. (2012), “A quantitative analysis of current security concerns and solutions for cloud computing”, *Journal of Cloud Computing: Advances, Systems and Applications* 1(1), pp. 1–18.
- [25] Hajibaba M. and Gorgin S. (2014), “A review on modern distributed computing paradigms: Cloud computing, jungle computing and fog computing”, *Journal of computing and information technology* 22(2), pp. 69–84.
- [26] Halperin D., Heydt-Benjamin T. S., Fu K., Kohno T., and Maisel W. H. (2008), “Security and privacy for implantable medical devices”, *IEEE pervasive computing* 7(1), pp. 30–39.
- [27] Han G., Jiang J., Zhang C., Duong T. Q., Guizani M., and Karagiannidis G. K. (2016), “A survey on mobile anchor node assisted localization in wireless sensor networks”, *IEEE Communications Surveys & Tutorials* 18(3), pp. 2220–2243.
- [28] Huang J. (2014), “The Workflow Task Scheduling Algorithm Based on the GA Model in the Cloud Computing Environment.”, *JSW* 9(4), pp. 873–880.
- [29] Hummel K. A., Bohs G., Brezany P., and Janciak I. (2006), “Mobility extensions for knowledge discovery workflows in data mining grids”, In *17th International Workshop on Database and Expert Systems Applications (DEXA'06)*, pp. 246–250. IEEE.
- [30] Hummel K. A. and Meyer H. (2008), “Self-organizing fair job scheduling among mobile devices”, In *2008 Second IEEE International Conference on Self-Adaptive and Self-Organizing Systems Workshops*, pp. 230–235. IEEE.
- [31] Isaiadis S. and Getov V. (2005), “Integrating mobile devices into the grid: Design considerations and evaluation”, In *European Conference on Parallel Processing*, pp. 1080–1088. Springer.
- [32] Jacob B., Brown M., Fukui K., Trivedi N., et al. (2005), “Introduction to grid computing”, *IBM redbooks*, pp. 3–6.
- [33] Jianfang C., Junjie C., and Qingshan Z. (2014), “An optimized scheduling algorithm on a cloud workflow using a discrete particle swarm”, *Cybernetics and Information Technologies* 14(1), pp. 25–39.

- [34] Kanchi S., Sandilya S., Bhosale D., Pitkar A., and Gondhalekar M. (2013), “Overview of LTE-A technology”, In *2013 IEEE global high tech congress on electronics* , pp. 195–200. IEEE.
- [35] Karaboga D. and Gorkemli B. (2011), “A combinatorial artificial bee colony algorithm for traveling salesman problem”, In *2011 International Symposium on Innovations in Intelligent Systems and Applications* , pp. 50–53. IEEE.
- [36] Katsaros K. and Polyzos G. C. (2007), “Towards the realization of a mobile grid”, In *Proceedings of the 2007 ACM CoNEXT conference* , pp. 1–2 ACM, New York, NY, USA.
- [37] Kavitha M. and Geetha B. (2019), “An efficient city energy management system with secure routing communication using WSN”, *Cluster Computing* 22(6), pp. 13131–13142.
- [38] Kawser M., Hamid N. I. B., Hasan N., Alam S., and Rahman M. M. (2012), “Downlink SNR to CQI mapping for different multiple antenna techniques in LTE”, *International journal of information and electronics engineering* 2(5), pp. 757–760.
- [39] Kayali M. O., Shmeiss Z., Safa H., and El-Hajj W. (2017), “Downlink scheduling in LTE: Challenges, improvement, and analysis”, In *2017 13th International Wireless Communications and Mobile Computing Conference (IWCMC)* , pp. 323–328. IEEE.
- [40] Kela P., Puttonen J., Kolehmainen N., Ristaniemi T., Henttonen T., and Moisio M. (2008), “Dynamic packet scheduling performance in UTRA long term evolution downlink”, In *2008 3rd International Symposium on Wireless Pervasive Computing* , pp. 308–313. IEEE.
- [41] Khan I. U., Ma X., Taylor J., Javaid N., and Gamage K. A. (2018), “Heuristic algorithm based dynamic scheduling model of home appliances in smart grid”, In *2018 24th International Conference on Automation and Computing (ICAC)* , pp. 1–6. IEEE.
- [42] Kiess W. and Mauve M. (2007), “A survey on real-world implementations of mobile ad-hoc networks”, *Ad Hoc Networks* 5(3), pp. 324–339.
- [43] Krishnamoorthy D., Vaiyapuri P., Ayyanar A., Harold Robinson Y., Kumar R., Long H. V., and Son L. H. (2020), “An Effective Congestion Control Scheme for MANET with Relative Traffic Link Matrix Routing”, *Arabian Journal for Science and Engineering* 45(8), pp. 6171–6181.

- [44] Kumar N., Chilamkurti N., Zeadally S., and Jeong Y.-S. (2014), “Achieving quality of service (QoS) using resource allocation and adaptive scheduling in cloud computing with grid support”, *The Computer Journal* 57(2), pp. 281–290.
- [45] Kurdi H., Li M., and Al-Raweshidy H. (2008), “A classification of emerging and traditional grid systems”, *IEEE Distributed Systems Online* 9(3), pp. 1–1.
- [46] Kwan R., Leung C., and Zhang J. (2008), “Multiuser scheduling on the downlink of an LTE cellular system”, *Research Letters in Communications* , pp. 1–4.
- [47] Lee I. and Sokolsky O. (2010), “Medical cyber physical systems”, In *Design automation conference* , pp. 743–748. IEEE.
- [48] Li C. and Li L. (2009), “Utility-based scheduling for grid computing under constraints of energy budget and deadline”, *Computer Standards Interfaces* 31(6), pp. 1131–1142.
- [49] Liu C., Zhang C., Chen F., and Zhao C. (2018), “Towards IPv6-based Architecture for Big Data Processing of Community Medical Internet of Things”, In *2018 14th International Wireless Communications & Mobile Computing Conference (IWCMC)* , pp. 1333–1333. IEEE.
- [50] Liu D. and Lee Y.-H. (2005), “An efficient scheduling discipline for packet switching networks using earliest deadline first round robin”, *Telecommunication Systems* 28(3), pp. 453–474.
- [51] Liu H., Abraham A., and Hassanien A. E. (2010), “Scheduling jobs on computational grids using a fuzzy particle swarm optimization algorithm”, *Future Generation Computer Systems* 26(8), pp. 1336–1343.
- [52] Liu K. and Li X. (2018), “Enhancing localization scalability and accuracy via opportunistic sensing”, *IEEE/ACM Transactions on Networking* 26(3), pp. 1517–1530.
- [53] Liu Y.-F. and Liu S.-Y. (2013), “A hybrid discrete artificial bee colony algorithm for permutation flowshop scheduling problem”, *Applied Soft Computing* 13(3), pp. 1459–1463.
- [54] Maia A. M., Vieira D., de Castro M. F., and Ghamri-Doudane Y. (2016), “A fair QoS aware dynamic LTE scheduler for machine-to-machine communication”, *Comput. Commun.* 89, pp. 75–86.
- [55] Maskooki A., Sabatino G., and Mitton N. (2015), “Analysis and performance evaluation of the next generation wireless networks”, In *Modeling and Simulation of Computer Networks and Systems* , pp. 601–627. Elsevier.

- [56] Memeti S., Pillana S., Binotto A., Kołodziej J., and Brandic I. (2018), “A review of machine learning and meta-heuristic methods for scheduling parallel computing systems”, In *Proceedings of the International Conference on Learning and Optimization Algorithms: Theory and Applications* , pp. 1–6.
- [57] Merz P. and Freisleben B. (1997), “A genetic local search approach to the quadratic assignment problem”, In *Proceedings of the 7th international conference on genetic algorithms* , pp. 1–1.
- [58] metacomputing infrastructure toolkit G. A. (1997), *The International Journal of Supercomputer Applications and High Performance Computing* 11(2), pp. 115–128.
- [59] Mishra M. K., Mohanty P., and Mund G. (2012), “A Time-Minimization Dynamic Job Grouping-based Scheduling in Grid Computing”, *International Journal of Computer Applications* 40(16), pp. 16–25.
- [60] Monghal G., Pedersen K. I., Kovacs I. Z., and Mogensen P. E. (2008), “QoS oriented time and frequency domain packet schedulers for the UTRAN Long Term Evolution”, In *VTC Spring 2008-IEEE Vehicular Technology Conference* , pp. 2532–2536 IEEE.
- [61] Mowla N. I., Doh I., and Chae K. (2018), “On-device AI-based cognitive detection of bio-modality spoofing in medical cyber physical system”, *IEEE Access* 7, pp. 2126–2137.
- [62] Nasralla M. M. (2020), “A hybrid downlink scheduling approach for multi-traffic classes in LTE wireless systems”, *IEEE Access* 8, pp. 82173–82186.
- [63] Negru C., Pop F., Cristea V., Bessisy N., and Li J. (2013), “Energy efficient cloud storage service: key issues and challenges”, In *2013 Fourth International Conference on Emerging Intelligent Data and Web Technologies* , pp. 763–766. IEEE.
- [64] Neri F. and Cotta C. (2012), “Memetic algorithms and memetic computing optimization: A literature review”, *Swarm and Evolutionary Computation* 2, pp. 1–14.
- [65] Omar D. M., Khedr A. M., and Agrawal D. P. (2017), “Optimized clustering protocol for balancing energy in wireless sensor networks”, *International Journal of Communication Networks and Information Security* 9(3), pp. 367–375.
- [66] OSMAN I. (1996), “Meta-heuristics: an overview”, *Meta-Heuristics: Theory and Applications* , pp. 1–21.

- [67] Ozturk E. and Altılar T. (2007), “IMOGA: an architecture for integrating mobile devices into grid applications”, In *Annual International Conference on Mobile and Ubiquitous Systems* , pp. 1–8.
- [68] Palmer N., Kemp R., Kielmann T., and Bal H. (2009), “Ibis for mobility: solving challenges of mobile computing using grid techniques”, In *HotMobile '09: Proceedings of the 10th workshop on Mobile Computing Systems and Applications* , pp. 1–6 ACM, New York, NY, USA.
- [69] Paradiso J. A. and Starner T. (2005), “Energy scavenging for mobile and wireless electronics”, *IEEE Pervasive computing* 4(1), pp. 18–27.
- [70] Parker M. (2017), “Error-Correction Coding”, In *Digital Signal Processing 101: Everything you need to know to get started* chapter 12, pp. 129–147. Newnes.
- [71] Passerini F. and Tonello A. M. (2019), “Smart grid monitoring using power line modems: Effect of anomalies on signal propagation”, *IEEE Access* 7, pp. 27302–27312.
- [72] Pedersen K. I., Monghal G., Kovacs I. Z., Kolding T. E., Pokhariyal A., Frederiksen F., and Mogensen P. (2007), “Frequency domain scheduling for OFDMA with limited and noisy channel feedback”, In *2007 IEEE 66th Vehicular Technology Conference* , pp. 1792–1796. IEEE.
- [73] Phan T., Huang L., and Dulan C. (2002), “Challenge: Integrating mobile wireless devices into the computational grid”, In *Proceedings of the 8th annual international conference on Mobile computing and networking* , pp. 271–278.
- [74] Piro G., Grieco L. A., Boggia G., Fortuna R., and Camarda P. (2011), “Two-level downlink scheduling for real-time multimedia services in LTE networks”, *IEEE Transactions on Multimedia* 13(5), pp. 1052–1065.
- [75] Pokhariyal A., Pedersen K., Monghal G., Kovacs I., Rosa C., Kolding T., and Mogensen P. (2007), “HARQ aware frequency domain packet scheduler with different degrees of fairness for the UTRAN Long Term Evolution”, In *2007 IEEE 65th Vehicular Technology Conference-VTC2007-Spring* , pp. 2761–2765 IEEE.
- [76] Poli R., Kennedy J., and Blackwell T. (2007), “Particle swarm optimization”, *Swarm intelligence* 1(1), pp. 33–57.
- [77] Poola D., Garg S. K., Buyya R., Yang Y., and Ramamohanarao K. (2014), “Robust scheduling of scientific workflows with deadline and budget constraints in clouds”, In *2014 IEEE 28th international conference on advanced information networking and applications* , pp. 858–865. IEEE.

- [78] Pop F., Dobre C., and Cristea V. (2009), “Genetic algorithm for DAG scheduling in grid environments”, In *2009 IEEE 5th International Conference on Intelligent Computer Communication and Processing*, pp. 299–305. IEEE.
- [79] Priyadarshini I., Kumar R., Tuan L. M., Son L. H., Long H. V., Sharma R., and Rai S. (2021), “A new enhanced cyber security framework for medical cyber physical systems”, *SICS Software-Intensive Cyber-Physical Systems* 35(3), pp. 159–183.
- [80] Rahmani A. M., Gia T. N., Negash B., Anzanpour A., Azimi I., Jiang M., and Liljeberg P. (2018), “Exploiting smart e-Health gateways at the edge of healthcare Internet-of-Things: A fog computing approach”, *Future Generation Computer Systems* 78, pp. 641–658.
- [81] Raileanu S., Anton F., Iatan A., Borangiu T., Anton S., and Morariu O. (2017), “Resource scheduling based on energy consumption for sustainable manufacturing”, *Journal of Intelligent Manufacturing* 28(7), pp. 1519–1530.
- [82] Rajakumar B. (2012), “The Lion’s Algorithm: a new nature-inspired search algorithm”, *Procedia Technology* 6, pp. 126–135.
- [83] Ramesh T. and Konda S. K. (2016), “Energy-Efficient resource allocation in WDM networks”, In *2016 International Conference on Communication and Signal Processing (ICCSP)*, pp. 1697–1700. IEEE.
- [84] Ranaldo N. and Zimeo E. (2009), “Time and cost-driven scheduling of data parallel tasks in grid workflows”, *IEEE Systems Journal* 3(1), pp. 104–120.
- [85] Roehrs A., da Costa C. A., da Rosa Righi R., da Silva V. F., Goldim J. R., and Schmidt D. C. (2019), “Analyzing the performance of a blockchain-based personal health record implementation”, *Journal of biomedical informatics* 92, pp. 103140.
- [86] Sahu D. P., Singh K., Manju M., Taniar D., Abdel-Basset M., Long H. V., et al. (2019), “Heuristic search based localization in mobile computational grid”, *IEEE Access* 7, pp. 78652–78664.
- [87] Sakellariou R. and Zhao H. (2004), “A low-cost rescheduling policy for efficient mapping of workflows on grid systems”, *Scientific Programming* 12(4), pp. 253–262.
- [88] Salman M. I., Mansoor A. M., Jalab H. A., Sabri A. Q. M., and Ahmed R. (2018), “A joint evaluation of energy-efficient downlink scheduling and partial CQI feedback for LTE video transmission”, *Wireless Personal Communications* 98(1), pp. 189–211.

- [89] Sediq A. B., Gohary R., Schoenen R., and Yanikomeroğlu H. (2013), “Optimal tradeoff between sum-rate efficiency and Jain’s fairness index in resource allocation”, *IEEE Transactions on Wireless Communications* 12(7), pp. 3496–3509.
- [90] Shaw S. B. and Singh A. (2014), “A survey on scheduling and load balancing techniques in cloud computing environment”, In *2014 international conference on computer and communication technology (ICCCCT)* , pp. 87–95. IEEE.
- [91] Shu W., Wang W., and Wang Y. (2014), “A novel energy-efficient resource allocation algorithm based on immune clonal optimization for green cloud computing”, *EURASIP Journal on Wireless Communications and Networking* 2014(1), pp. 1–9.
- [92] Singh K. V. and Raza Z. (2018), “Resource Scheduling Approach Using a Quantum Variant of Gravitational Search Algorithm in Computational Mobile Grid”, In *2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence)* , pp. 641–645. IEEE.
- [93] Singhal U. and Jain S. (2014), “A new fuzzy logic and GSO based load balancing mechanism for public cloud”, *International Journal of Grid and Distributed Computing* 7(5), pp. 97–110.
- [94] Sivanandam S. and Deepa S. (2008), “Genetic algorithms”, In *Introduction to genetic algorithms* , pp. 15–37. Springer.
- [95] Skoutas D. N. and Rouskas A. N. (2010), “Scheduling with QoS provisioning in mobile broadband wireless systems”, In *2010 European Wireless Conference (EW)* , pp. 422–428 IEEE.
- [96] Sörensen K., Sevaux M., and Glover F. (2018), “A history of metaheuristics”, In *Handbook of heuristics* , pp. 791–808. Springer.
- [97] Stevens T., De Leenheer M., Devellder C., Dhoedt B., Christodoulopoulos K., Kokkinos P., and Varvarigos E. (2009), “Multi-cost job routing and scheduling in grid networks”, *Future Generation Computer Systems* 25(8), pp. 912–925.
- [98] Tabany M. R., Guy C. G., and Sherratt R. S. (2017), “A novel downlink semi-persistent packet scheduling scheme for VoLTE traffic over heterogeneous wireless networks”, *EURASIP Journal on Wireless Communications and Networking* 2017(1), pp. 1–14.
- [99] Talbi E. G. (2009), “Common Concepts for Metaheuristics”, In *Metaheuristics: from design to implementation* chapter 1, pp. 1–76. Wiley, London.

- [100] Tsai P. W., Pan J. S., Liao B. Y., Chu S. C., et al. (2009), “Enhanced artificial bee colony optimization”, *International Journal of Innovative Computing, Information and Control* 5(12), pp. 5081–5092.
- [101] Tuan L. M., Son L. H., Long H. V., Priya R., Soundar R., Robinson H., Kumar R., et al. (2020), “ITFDS: channel-aware integrated time and frequency-based downlink LTE scheduling in MANET”, *Sensors* 20(12), pp. 3394.
- [102] Uyan O. G. and Gungor V. C. (2019), “QoS-aware LTE-A downlink scheduling algorithm: A case study on edge users”, *International Journal of Communication Systems* 32(15), pp. 4066.
- [103] Vöckler J. S., Juve G., Deelman E., Rynge M., and Berriman B. (2011), “Experiences using cloud computing for a scientific workflow application”, In *Proceedings of the 2nd international workshop on Scientific cloud computing*, pp. 15–24.
- [104] Wang K., Li H., Maharjan S., Zhang Y., and Guo S. (2018), “Green energy scheduling for demand side management in the smart grid”, *IEEE Transactions on Green Communications and Networking* 2(2), pp. 596–611.
- [105] Wang W. (2015), “Next generation Internet and IPv6 transition”, *China Communications* 12(3), pp. 151–152.
- [106] Wang X., Wang Y., Zhu H., et al. (2012), “Energy-Efficient Multi-Job Scheduling Model for Cloud Computing and Its Genetic Algorithm”, *Mathematical Problems in Engineering* 2012, pp. 1–16.
- [107] Wang Y.-C. and Chuang C.-A. (2015), “Efficient eNodeB deployment strategy for heterogeneous cells in 4G LTE systems”, *Comput. Netw.* 79, pp. 297–312.
- [108] Wen C., Pan S., Wang C., and Li J. (2016), “An indoor backpack system for 2-D and 3-D mapping of building interiors”, *IEEE Geoscience and Remote Sensing Letters* 13(7), pp. 992–996.
- [109] Wengertter C., Ohlhorst J., and von Elbwart A. G. E. (2005), “Fairness and throughput analysis for generalized proportional fair frequency scheduling in OFDMA”, In *2005 IEEE 61st vehicular technology conferences volume 3*, pp. 1903–1907 IEEE.
- [110] Wiczorek M., Hoheisel A., and Prodan R. (2009), “Towards a general model of the multi-criteria workflow scheduling on the grid”, *Future Generation Computer Systems* 25(3), pp. 237–256.

- [111] Wu Q., Yun D., Lin X., Gu Y., Lin W., and Liu Y. (2012), “On workflow scheduling for end-to-end performance optimization in distributed network environments”, In *Workshop on Job Scheduling Strategies for Parallel Processing* , pp. 76–95. Springer.
- [112] Xie G. et al. (2016), “Survey of modeling methods in cyber-physical system”, *Journal on Communications* 37(5), pp. 165–175.
- [113] YAN J. and WU K.-g. (2010), “Scheduling algorithm for instance-intensive cloud workflow”, *Journal of Computer Applications* 30, pp. 2864–2866.
- [114] Yang S., Wang C., and Jiang C. (2018), “Centron: Cooperative neighbor discovery in mobile Ad-hoc networks”, *Comput. Netw.* 136, pp. 128–136.
- [115] Younis M. T. and Yang S. (2018), “Hybrid meta-heuristic algorithms for independent job scheduling in grid computing”, *Applied soft computing* 72, pp. 498–517.
- [116] Yu J. and Buyya R. (2005), “A taxonomy of workflow management systems for grid computing”, *Journal of Grid Computing* 3, pp. 171–200.
- [117] Yu J. and Buyya R. (2006), “Scheduling scientific workflow applications with deadline and budget constraints using genetic algorithms”, *Scientific Programming* 14(3), pp. 217–230.
- [118] Yucel F. and Bulut E. (2018), “Clustered crowd GPS for privacy valuing active localization”, *IEEE Access* 6, pp. 23213–23221.
- [119] Zaki Y., Weerawardane T., Gorg C., and Timm-Giel A. (2011), “Multi-QoS-aware fair scheduling for LTE”, In *2011 IEEE 73rd Vehicular Technology Conference (VTC Spring)* , pp. 1–5 IEEE.
- [120] Zhang L., Chen Y., Sun R., Jing S., and Yang B. (2008), “A task scheduling algorithm based on PSO for grid computing”, *International Journal of Computational Intelligence Research* 4(1), pp. 37–43.
- [121] Ziarati K., Akbari R., and Zeighami V. (2011), “On the performance of bee algorithms for resource-constrained project scheduling problem”, *Applied Soft Computing* 11(4), pp. 3720–3733.