

BỘ GIÁO DỤC  
VÀ ĐÀO TẠO

VIỆN HÀN LÂM KHOA HỌC  
VÀ CÔNG NGHỆ VIỆT NAM

HỌC VIỆN KHOA HỌC VÀ CÔNG NGHỆ

Nguyễn Thị Thanh Vân

**NGHIÊN CỨU NÂNG CAO CHẤT LƯỢNG TRUYỀN THÔNG  
VÀ ỨNG DỤNG THUẬT TOÁN HỌC SÂU TĂNG CƯỜNG  
TRONG ĐIỀU HƯỚNG TỐI ƯU CHO ROBOT DI ĐỘNG**

**LUẬN ÁN TIẾN SĨ KỸ THUẬT ĐIỆN, ĐIỆN TỬ VÀ VIỄN THÔNG**

**Ngành: Kỹ thuật điều khiển và tự động hóa**

**Mã số: 9 52 02 16**

Xác nhận của Học viện  
Khoa học và Công nghệ

Người hướng dẫn 1  
(Ký, ghi rõ họ tên)

Người hướng dẫn 2  
(Ký, ghi rõ họ tên)



**KI. GIÁM ĐỐC**  
**PHÓ GIÁM ĐỐC**

Nguyễn Thị Trung

TS. Ngô Mạnh Tiến

TS. Nguyễn Công Lượng

Hà Nội - 2024

## LỜI CAM ĐOAN

Tôi xin cam đoan luận án: "Nghiên cứu nâng cao chất lượng truyền thông và ứng dụng thuật toán học sâu tăng cường trong điều hướng tối ưu cho robot di động" là công trình nghiên cứu của chính mình dưới sự hướng dẫn khoa học của tập thể hướng dẫn. Luận án 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 công bố 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 ngoài các công trình công bố của tác giả. Luận án được hoàn thành trong thời gian tôi làm nghiên cứu sinh tại Học viện Khoa học và Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.

Hà Nội, ngày 14 tháng 11 năm 2024

Tác giả luận án



**Nguyễn Thị Thanh Vân**

## LỜI CẢM ƠN

Luận án này được hoàn thành với sự nỗ lực không ngừng của tác giả và sự giúp đỡ hết mình từ các thầy giáo hướng dẫn, bạn bè và người thân.

Đầu tiên, tác giả xin bày tỏ lời tri ân tới TS Ngô Mạnh Tiến và TS Nguyễn Công Lượng, cảm ơn các thầy đã tận tình hướng dẫn tác giả hoàn thành luận án này.

Tác giả xin gửi lời cảm ơn tới các thầy, cô giáo và cán bộ của Viện Công nghệ thông tin, Học viện Khoa học và Công nghệ (Viện Hàn lâm Khoa học và Công nghệ Việt Nam) đã nhiệt tình giúp đỡ và tạo ra môi trường nghiên cứu tốt để tác giả hoàn thành công trình của mình; cảm ơn các thầy, cô và các đồng nghiệp ở các nơi mà tác giả tham gia viết bài đã có những góp ý chính xác để tác giả có được những công bố như ngày hôm nay.

Tác giả xin cảm ơn tới Ban Giám hiệu trường Đại học Phenikaa, các đồng nghiệp khoa Điện Điện tử nơi tác giả công tác đã ủng hộ và hỗ trợ để luận án được hoàn thành.

Cuối cùng, tác giả xin gửi tới gia đình, bạn bè, người thân lời cảm ơn chân thành nhất vì đã đồng hành cùng tác giả trong suốt thời gian qua.

Hà Nội, ngày 14 tháng 11 năm 2024

Tác giả luận án



**Nguyễn Thị Thanh Vân**

## MỤC LỤC

<b>LỜI CAM ĐOAN</b>	<b>i</b>
<b>LỜI CẢM ƠN</b>	<b>ii</b>
<b>MỤC LỤC</b>	<b>iii</b>
<b>DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT</b>	<b>vi</b>
<b>DANH MỤC HÌNH VẼ</b>	<b>x</b>
<b>DANH MỤC BẢNG</b>	<b>xii</b>
<b>MỞ ĐẦU</b>	<b>1</b>
<b>1 TỔNG QUAN VỀ VIỆC LẬP QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG NHÀ MÁY</b>	<b>6</b>
1.1 Bài toán lập quỹ đạo cho robot di động . . . . .	6
1.2 Robot di động và vấn đề truyền thông không dây . . . . .	12
1.3 Tình hình nghiên cứu trong và ngoài nước . . . . .	16
1.3.1 Tình hình nghiên cứu trong nước . . . . .	16
1.3.2 Tình hình nghiên cứu ngoài nước . . . . .	18
1.4 Định hướng nghiên cứu của luận án . . . . .	23
1.5 Kết luận chương 1 . . . . .	24
<b>2 NÂNG CAO CHẤT LƯỢNG TRUYỀN THÔNG TRONG NHÀ MÁY PHỤC VỤ HOẠT ĐỘNG CỦA ROBOT DI ĐỘNG</b>	<b>26</b>
2.1 Vấn đề tối ưu hiệu suất truyền thông không dây tại các nhà máy . . . . .	26
2.2 Mô hình hệ thống . . . . .	31
2.2.1 Tín hiệu nhận tại các robot di động (IUs) . . . . .	33
2.2.2 Tín hiệu nhận tại các cảm biến không dây thu năng lượng (EUs) . . . . .	34

2.2.3	Tín hiệu nhận tại bộ thu D2D . . . . .	34
2.2.4	Bài toán tối ưu . . . . .	35
2.3	Thuật toán tối ưu . . . . .	37
2.3.1	Bài toán 1 với kịch bản N-OTA . . . . .	39
2.3.2	Bài toán 2 với kịch bản N-OTA . . . . .	42
2.3.3	Tính hội tụ, điểm khả thi và lựa chọn hệ số phạt $\eta$ . . . . .	46
2.4	Kịch bản phân chia thời gian . . . . .	46
2.4.1	Bài toán 1 với kịch bản OTA . . . . .	49
2.4.2	Bài toán 2 với kịch bản OTA . . . . .	51
2.5	Đánh giá hiệu quả mô hình và thuật toán . . . . .	53
2.6	Kết luận chương 2 . . . . .	58
<b>3</b>	<b>THUẬT TOÁN TỐI ƯU QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG MÔI TRƯỜNG TRUYỀN THÔNG LÝ TƯỞNG</b>	<b>60</b>
3.1	Tối ưu lỗi . . . . .	60
3.1.1	Bài toán tối ưu . . . . .	60
3.1.2	Bài toán tối ưu lỗi . . . . .	62
3.1.3	Thuật toán tối ưu lỗi . . . . .	63
3.1.4	Giới thiệu CVX . . . . .	65
3.2	Phát triển thuật toán tối ưu quỹ đạo cho robot di động . . . . .	66
3.2.1	Mục tiêu và các ràng buộc của bài toán . . . . .	66
3.2.2	Thuật toán tối ưu quỹ đạo . . . . .	69
3.2.3	Đánh giá kết quả thuật toán CAA . . . . .	71
3.3	Phát triển thuật toán tối thiểu năng lượng tiêu thụ . . . . .	73
3.3.1	Mô hình hóa năng lượng tiêu thụ của robot di động . . . . .	74
3.3.2	Các ràng buộc tránh va chạm . . . . .	75
3.3.3	Thuật toán tối ưu năng lượng tiêu thụ cho robot . . . . .	76
3.3.4	Bài toán tối ưu trong kịch bản môi trường có vật cản động . . . . .	78
3.3.5	Đánh giá kết quả . . . . .	82
3.4	Kết luận chương 3 . . . . .	85
<b>4</b>	<b>THUẬT TOÁN TỐI ƯU QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG MÔI TRƯỜNG TRUYỀN THÔNG CÓ STAR-IRS HỖ TRỢ</b>	<b>87</b>
4.1	Mô hình hệ thống điều hướng cho robot di động có STAR-IRS hỗ trợ . . . . .	87

4.1.1	STAR-IRS . . . . .	87
4.1.2	Mô hình hệ thống điều hướng cho robot di động trong môi trường truyền thông . . . . .	90
4.1.3	Mô hình tín hiệu . . . . .	93
4.1.4	Bài toán điều hướng tối ưu cho robot di động . . . . .	94
4.2	Mô hình quy trình quyết định MDP (Markov Decision Process) . . . . .	96
4.2.1	Học tăng cường RL (Reinforcement Learning) . . . . .	96
4.2.2	Mô hình MDP cho bài toán điều hướng tối ưu . . . . .	98
4.3	Thuật toán học sâu tăng cường DRL (Deep Reinforcement Learning) . . . . .	100
4.3.1	Thuật toán A2C (Advantage Actor-Critic) . . . . .	100
4.3.2	Thuật toán PPO (Proximal Policy Optimization) . . . . .	102
4.3.3	Độ hội tụ và độ phức tạp của thuật toán dựa trên actor-critic . . . . .	104
4.4	Đánh giá hiệu quả mô hình và thuật toán điều hướng cho robot di động	106
4.5	Kết luận chương 4 . . . . .	113
<b>KẾT LUẬN VÀ KIẾN NGHỊ</b>		<b>115</b>
<b>DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ</b>		<b>117</b>
<b>TÀI LIỆU THAM KHẢO</b>		<b>119</b>
<b>PHỤ LỤC</b>		<b>131</b>

## DANH MỤC CÁC KÝ HIỆU, CÁC CHỮ VIẾT TẮT

### 1. Danh mục các ký hiệu

Ký hiệu	Ý nghĩa
$\mathcal{S}$	Không gian trạng thái của robot
$\mathcal{A}$	Không gian hành động của robot
$\mathcal{R}$	Tập điểm thưởng của robot
$V(s)$	Hàm giá trị của trạng thái $s$
$Q(s, a)$	Hàm giá trị của hành động $a$
$\delta_t$	Lỗi TD tại thời điểm $t$
$\gamma$	hệ số chiết khấu (discount factor)
$\bar{\gamma}$	Ngưỡng tốc độ truyền thông tối thiểu của robot
$\hat{\mathbf{q}} = [x, y, \varphi]^T$	Véc tơ tư thế của robot trong hệ quy chiếu toàn cục
$\mathbf{p}$	công suất của thiết bị
$p^{\max}$	công suất lớn nhất của AP
$P_{B,\max}$	công suất lớn nhất của BS
$\mathbf{q} = [x, y]^T$	Véc tơ tọa độ của robot trong hệ quy chiếu toàn cục
$\mathbf{q}_s$	Véc tơ tọa độ vị trí điểm xuất phát của robot
$\mathbf{q}_g$	Véc tơ tọa độ vị trí điểm đích đến của robot
$v$	Vận tốc dài của robot
$\omega$	Vận tốc góc của robot
$m$	Khối lượng của robot
$I$	Momen quán tính của robot
$\theta$	Tham số của mạng actor (bao gồm trọng số và phương sai)
$\bar{\theta}_n$	độ dịch pha (phase shift) của phần tử thứ $n$ của IRS
$\mathbf{G}$	Kênh truyền trực tiếp từ BS/AP đến IRS
$\mathbf{h}_{(A,r)}$	Kênh truyền từ AP đến robot
$\mathbf{h}_{(I,r)}$	Kênh truyền từ IRS đến robot
$\beta_{l,n}^r$	Hệ số phản xạ của phần tử thứ $n$ của STAR-IRS tại khe thời gian $l$

$\beta_{l,n}^t$	Hệ số phản xạ của phần tử thứ $n$ của STAR-IRS tại khe thời gian $l$
$R_l^c$	Tốc độ truyền của robot tại khe thời gian $l$
$\mathbf{Q}$ [ $\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_L$ ]	Quy đạo rời rạc của robot
$\underline{\omega}$	Tham số của mạng critic (bao gồm trọng số và phương sai)
$s_t$	Trạng thái (state) tại thời điểm $t$
$a_t$	Hành động (action) tại thời điểm $t$
$r_t$	Điểm thưởng (reward) tại thời điểm $t$
$A(s_t, a_t)$	Hàm đánh giá cặp trạng thái-hành động ( $s_t, a_t$ )
$\pi_\theta$	Chính sách (policy) với tham số $\theta$
$V_\omega$	Hàm giá trị trạng thái với tham số $\omega$
$\bar{B}$	Kích thước lô dữ liệu (batch size)
$\bar{b}$	Mẫu thứ $\bar{b}$ trong lô dữ liệu
$\mathbf{PL}_{LoS}$	Mất mát đường truyền kênh LoS
$\mathbf{PL}_{NLoS}$	Mất mát đường truyền kênh NLoS
$f_c$	tần số sóng mang (carrier frequency)
$\bar{M}$	Số lượng ăng ten của BS/AP
$\hat{M}$	Tham số big-M
$M$	Số lượng vật cản trong môi trường
$\bar{\mathbf{w}}$	Bộ tạo chùm tia đến các thiết bị yêu cầu thông tin
$\bar{\mathbf{v}}$	Bộ tạo chùm tia đến các thiết bị yêu cầu năng lượng
$\tau$	Bước thời gian
$v$	Vận tốc của robot
$P_s$	Công suất nguồn hệ thống mạch điện trên robot

## 2. Danh mục các chữ viết tắt

Chữ viết tắt	Ý nghĩa tiếng Anh	Ý nghĩa tiếng Việt
NCS	PhD Student	Nghiên cứu sinh
CT1		Công trình công bố số 1 của NCS
AMR	Autonomous Mobile Robot	Robot tự hành



CAA	Convex Approximation Algorithm	Thuật toán xấp xỉ lồi
CCA	Conventional Computing Algorithm	Thuật toán tính toán thông thường
ANN	Artificial Neural Network	Mạng nơ ron nhân tạo
AP	Access Point	Điểm truy cập
BS	Base Station	Trạm gốc
IU	Information-demanded user	Thiết bị yêu cầu thông tin
EU	Energy-demanded user	Thiết bị yêu cầu năng lượng
D2D	Device-to-Device	các thiết bị truyền thông tin trực tiếp
DED2D	The integrated data and energy network and D2D communication coexistence	Hệ thống tích hợp mạng thông tin và năng lượng kết hợp truyền thông D2D
SWIPT	Simultaneous wireless information and power transfer	Đồng thời truyền thông tin và năng lượng
N-OTA	Non Orthogonal Time Allocation	Thuật toán với kịch bản không phân bổ thời gian cho truyền thông của mạng D2D
OTA	Orthogonal Time Allocation	Thuật toán với kịch bản phân bổ thời gian cho truyền thông của mạng D2D
LoS	Line-of-sight	Đường truyền trực tiếp
NLoS	non-LoS	Đường truyền không trực tiếp
IRS	Intelligent Reflecting Surface	Bề mặt phản xạ thông minh
STAR-IRS	Simultaneously Transmitting and Reflecting IRS	IRS đồng thời truyền và phản xạ
CVX	a Matlab-based modeling system for convex optimization	Hệ thống mô hình hóa dựa trên nền Matlab cho bài toán tối ưu lồi
SO	Static Obstacle	
MO	Moving Obstacle	

OSOW	Algorithm for scenario that has Only Static Obstacles in the robot's workspace	Thuật toán lập quỹ đạo cho robot di động trong môi trường chỉ có vật cản tĩnh
MSOW	Algorithm for scenario that has Moving and Static Obstacles in the robot's workspace	Thuật toán lập quỹ đạo cho robot di động trong môi trường có cả vật cản tĩnh và động
TD	Temporal-difference	Lỗi TD
DRL	Deep Reinforcement Learning	Học sâu tăng cường (Học tăng cường sâu)
RL	Reinforcement Learning	Học tăng cường
DQN	Deep Q-Network	
DDQN	Double Deep Q-Network	
PPO	Proximal Policy Optimization	
A2C	Advantage Actor-Critic	
PPO-STAR-IRS		Thuật toán PPO điều hướng cho robot có STAR-IRS hỗ trợ
PPO-IRS		Thuật toán PPO điều hướng cho robot có IRS hỗ trợ
PPO-non-IRS		Thuật toán PPO điều hướng cho robot không có IRS hỗ trợ
A2C-STAR-IRS		Thuật toán A2C điều hướng cho robot có STAR-IRS hỗ trợ
SGD	Stochastic Gradient Descent	
SGA	Stochastic Gradient Ascent	

## DANH MỤC HÌNH VẼ

1.1	Sơ đồ tổng quan hệ thống điều khiển robot di động, trong đó, $\mathbf{q}_r$ là vị trí của robot, $\mathbf{q}_e$ là sai số vị trí của robot, $[\tau_l, \tau_a]^T$ là mô men đặt vào động cơ, $\bar{\mathbf{v}}$ và $\bar{\mathbf{w}}$ lần lượt là vận tốc dài và vận tốc góc của robot [1]. . .	7
1.2	Số lượng các công trình nghiên cứu về lập quỹ đạo cho robot dựa trên số liệu từ Engineering Village [2]. . . . .	8
1.3	Hệ thống Wifi tại nhà máy SHI-V 1 thiết kế bởi Netmarks [3]. . . . .	14
1.4	Sơ đồ vị trí đặt AP tại nhà máy SHI-V 1 thiết kế bởi Netmarks [3]. . .	15
1.5	IRS giúp thay đổi vùng phủ sóng tương ứng với sự thay đổi layout của nhà máy. [4] . . . . .	21
1.6	IRS giúp truyền tín hiệu đến thiết bị nhận khi kênh truyền trực tiếp bị chặn. [5] . . . . .	22
1.7	Hệ thống điều hướng cho robot trong nhà máy có IRS hỗ trợ. [6] . . . .	23
2.1	Cấu trúc IRS [7, 8]. . . . .	28
2.2	Mô tả các ứng dụng của IRS với hệ thống mạng không dây trong tương lai [8]. . . . .	29
2.3	Mạng DED2D có IRS hỗ trợ. . . . .	32
2.4	a) Độ hội tụ của các thuật toán và b) thông lượng dữ liệu thay đổi với công suất định mức của BS $P_{B,max}$ . . . . .	55
2.5	Thông lượng đạt được với a) ngưỡng tốc độ truyền thông D2D $R_{k,min}$ và b) số lượng cặp D2D $K$ . . . . .	56
2.6	Thông lượng đạt được với a) số lượng các phần tử của IRS và b) số lượng ăng ten tại BS. . . . .	57
3.1	Mô hình robot di động hai bánh chủ động. . . . .	67
3.2	(a) Quỹ đạo của robot với thuật toán CAA, (b) Quỹ đạo của robot với thuật toán CCA, và (c) Góc quay của robot. . . . .	72
3.3	Quãng đường và thời gian di chuyển khi a) $v_{max}$ thay đổi b) $d_{max}^r$ thay đổi. . . . .	73

3.4	a) Quỹ đạo tối ưu bởi ba thuật toán b) Năng lượng tiêu thụ và thời gian di chuyển tương ứng với vận tốc lớn nhất của robot. . . . .	83
3.5	Năng lượng tiêu thụ và quãng đường di chuyển khi số lượng vật cản động trong môi trường làm việc thay đổi. . . . .	84
3.6	Quỹ đạo tối ưu của robot khi tối thiểu năng lượng tiêu thụ với a) kịch bản MSOW có 01 MO b) kịch bản MSOW có 02 MOs, c) ba kịch bản và d) Góc quay của robot tương ứng với các quỹ đạo tối ưu. . . . .	85
4.1	STAR-IRS. [9] . . . . .	88
4.2	Các chế độ hoạt động của STAR-IRS. [9] . . . . .	90
4.3	Mô hình hệ thống điều hướng robot di động trong nhà có STAR-IRS hỗ trợ truyền thông. . . . .	91
4.4	Tương tác giữa tác nhân và môi trường trong học tăng cường RL. [10] .	97
4.5	Tổng quan thuật toán PPO bài toán điều hướng tối ưu cho robot di động.	103
4.6	Quy trình quản lý tài nguyên thuật toán PPO áp dụng cho bài toán điều hướng tối ưu cho robot di động . . . . .	104
4.7	Bản đồ môi trường làm việc của robot di động. . . . .	106
4.8	a) Độ hội tụ của các thuật toán và b) Quỹ đạo tối ưu thu được bởi các thuật toán. . . . .	108
4.9	Ảnh hưởng của ngưỡng yêu cầu tốc độ truyền thông đến quỹ đạo tối ưu của robot di động. . . . .	110
4.10	Ảnh hưởng của ngưỡng yêu cầu tốc độ truyền thông lên a) năng lượng truyền thông và b) năng lượng tiêu thụ bởi robot di động. . . . .	110
4.11	Ảnh hưởng của số bit lượng tử ( $b$ ) đến quỹ đạo của robot di động. . . .	111
4.12	Ảnh hưởng của số bit lượng tử ( $b$ ) đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động. . . . .	111
4.13	Ảnh hưởng của số phần tử của STAR-IRS đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động. . . . .	112
4.14	Ảnh hưởng của công suất truyền tối đa của AP đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động. . . . .	113

## DANH MỤC BẢNG

1.1 Các yêu cầu đối với truyền thông không dây cho ứng dụng trong nhà máy tự động hóa tổng hợp bởi dự án KoI năm 2016 [11]. . . . .	13
2.1 Bảng tham số mô phỏng thuật toán N-OTA và OTA . . . . .	55
3.1 Bảng tham số mô phỏng thuật toán CAA và CCA . . . . .	72
3.2 Bảng tham số mô phỏng thuật toán OSOW và MOSW . . . . .	82
4.1 Bảng tham số mô phỏng thuật toán điều hướng tối ưu cho robot di động . .	107
4.2 Năng lượng tiêu thụ bởi các thuật toán . . . . .	109
4.3 Tổng quãng đường di chuyển của robot với các thuật toán . . . . .	109

## MỞ ĐẦU

### 1. Tính cấp thiết của luận án

Trong hơn một thập kỷ qua, sự phát triển của nền công nghiệp đã mang lại cho con người cuộc sống tiện nghi, hiện đại và thoải mái hơn nhờ sự trợ giúp của nhiều thiết bị công nghệ. Một trong số đó là các robot. Robot hiện nay là một chủ đề được quan tâm không chỉ ở các quốc gia lớn mạnh mà còn ở các quốc gia đang phát triển như Việt Nam. Ở Việt Nam, robot đã xuất hiện trong nhiều lĩnh vực như robot hút bụi, robot lễ tân, robot trong y tế, robot phục vụ trong các nhà máy... Robot trong các nhà máy đã có đóng góp lớn trong việc đảm bảo an toàn cho con người, nâng cao chất lượng cũng như hiệu quả sản xuất. Trong số các loại robot đang được sử dụng trong nhà máy, robot di động đang được một số lượng lớn các nhà khoa học và doanh nghiệp dành nhiều sự quan tâm bởi những ứng dụng và đóng góp không nhỏ của nó cho sự phát triển của các nhà máy cũng như nền công nghiệp.

Nhằm đáp ứng nhu cầu của xã hội, số lượng robot di động trong các nhà máy ngày càng tăng một cách đáng kể. Ngoài ra, số lượng thiết bị IoT (Internet of Things) trong nhà máy cũng tăng một cách nhanh chóng. Điều này dẫn đến vấn đề chia sẻ không gian làm việc giữa các robot, giữa robot với các thiết bị khác và giữa robot với con người. Để đạt hiệu quả làm việc tối ưu, các robot di động tại các nhà máy vừa cần hoàn thành nhiệm vụ di chuyển, vừa cần thực hiện trao đổi thông tin với các thiết bị khác như trạm chủ trung tâm, các robot khác trong hệ đa robot hay các thiết bị không dây khác nhằm đảm bảo việc quản lý cũng như vận hành trong toàn nhà máy. Hơn thế nữa, với đặc trưng di động, năng lượng phục vụ hoạt động của robot đều sử dụng nguồn pin rời gắn trên thân robot, vì vậy mức năng lượng sử dụng là hạn chế. Để đảm bảo các yêu cầu công việc phức tạp hiện nay, robot được kỳ vọng là sẽ tiêu tốn ít điện năng hơn và cho thời gian sử dụng lâu hơn.

Chính vì vậy, để đáp ứng các yêu cầu trong nhà máy, việc tích hợp truyền thông không dây cho robot di động là một giải pháp. Khi đó, robot di động đóng vai trò là một thành phần trong hệ thống IoT vừa giúp giảm bớt tiêu hao năng lượng tính toán

cho robot vừa giải quyết vấn đề truyền thông trong nhà máy. Để đảm bảo robot vừa hoàn thành nhiệm vụ vừa truyền thông tin cậy với hệ thống, việc giải bài toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông không dây tại nhà máy là cần thiết. Do đó, NCS chọn đề tài "***Nghiên cứu nâng cao chất lượng truyền thông và ứng dụng thuật toán học sâu tăng cường trong điều hướng tối ưu cho robot di động***".

Đây là bài toán bao gồm ba vấn đề chính: vấn đề điều hướng hay lập quỹ đạo cho robot (luận án sẽ sử dụng thuật ngữ lập quỹ đạo cho các chương đầu, và sử dụng thuật ngữ điều hướng cho chương 4), vấn đề truyền thông giữa robot với trạm điều khiển trung tâm của nhà máy và vấn đề tối ưu hệ thống.

Về vấn đề thứ nhất, về việc lập quỹ đạo cho robot di động. Đây là vấn đề được quan tâm từ lâu và vẫn đang dành được nhiều sự quan tâm của các nhà nghiên cứu và doanh nghiệp. Theo đó, đã có nhiều thuật toán lập quỹ đạo cho robot được đưa ra như các thuật toán Dijkstra, A\*, D\*... dựa trên một bản đồ về môi trường làm việc đã biết trước. Trong thực tế, robot thường làm việc trong môi trường chưa biết trước hoặc chỉ biết trước một phần. Do đó, một số thuật toán dựa trên học sâu tăng cường DRL (Deep Reinforcement Learning) đã được đề xuất trong những năm gần đây đã thể hiện được hiệu quả.

Về vấn đề thứ hai, về truyền thông cho robot di động. Đây là một vấn đề xuất phát từ các yêu cầu thực tế trong nhà máy, đặc biệt là do xu hướng phát triển IoT của nhà máy nói chung và xu hướng phát triển của hệ đa robot hay robot bầy đàn nói riêng. Theo đó, cũng đã có các nghiên cứu về vấn đề truyền thông không dây cho robot di động. Có hai vấn đề lớn được đặt ra đối với hệ thống truyền thông không dây trong nhà máy, một là vấn đề về độ tin cậy và độ trễ truyền thông, hai là vấn đề vùng chết tín hiệu (vùng không nhận được tín hiệu truyền đến) trong nhà máy.

Về vấn đề thứ ba, việc kết hợp hai vấn đề kể trên trong một bài toán tối ưu. Tùy vào độ phức tạp của bài toán với các yếu tố như hàm mục tiêu, các ràng buộc và số lượng các biến cần tối ưu, có thể chia phương pháp giải bài toán tối ưu thành hai dạng: phương pháp giải tường minh dựa trên các kỹ thuật toán học, và phương pháp giải dựa trên các kỹ thuật học máy.

Do đó, nội dung luận án sẽ tập trung vào việc nâng cao chất lượng truyền thông phục vụ hoạt động cho robot di động, từ đó phát triển thuật toán lập quỹ đạo cho robot trong môi trường truyền thông lý tưởng, và phát triển mô hình hệ thống và thuật

toán điều hướng cho robot di động trong môi trường truyền thông chưa đảm bảo với mục tiêu tối ưu năng lượng tiêu thụ cho robot di động, cuối cùng là thực hiện các mô phỏng để kiểm chứng tính hiệu quả của các mô hình hệ thống và thuật toán đề xuất.

## **2. Mục tiêu nghiên cứu của luận án**

- Nghiên cứu mô hình và thuật toán tối ưu hệ thống mạng nâng cao chất lượng truyền thông phục vụ hoạt động của robot di động.
- Nghiên cứu, phát triển các thuật toán tối ưu quỹ đạo với mục tiêu tối thiểu năng lượng tiêu thụ cho robot di động hoạt động trong môi trường truyền thông lý tưởng.
- Nghiên cứu, phát triển các hệ thống điều hướng cho robot di động kết hợp tối ưu quỹ đạo với truyền thông không dây và tối thiểu năng lượng tiêu thụ của hệ thống.
- Mô phỏng kiểm chứng và đánh giá hiệu quả của các thuật toán và mô hình hệ thống đề xuất.

## **3. Các nội dung nghiên cứu chính của luận án**

### **Về lý thuyết:**

- Nghiên cứu tổng quan về robot di động hoạt động trong môi trường truyền thông, tình hình nghiên cứu trong và ngoài nước, từ đó rút ra các hướng nghiên cứu thích hợp cho luận án.
- Nghiên cứu, đề xuất mô hình và thuật toán tối ưu mạng hỗ trợ giúp mở rộng vùng phủ sóng và nâng cao chất lượng truyền thông không dây trong nhà máy.
- Xây dựng bài toán tối ưu quỹ đạo di chuyển cho robot trong nhà máy nhằm tối thiểu năng lượng tiêu thụ trong điều kiện truyền thông lý tưởng, từ đó xây dựng thuật toán giải quyết bài toán đó.
- Nghiên cứu, đề xuất hệ thống điều hướng cho robot di động trong môi trường điều kiện truyền thông chưa đảm bảo.
- Xây dựng bài toán tối ưu quỹ đạo di chuyển cho robot trong nhà máy nhằm tối thiểu năng lượng tiêu thụ trong môi trường truyền thông chưa đảm bảo, từ đó xây dựng thuật toán giải quyết bài toán đó.



- Nghiên cứu sử dụng các công cụ tối ưu để thực hiện mô phỏng.

#### **Về thực nghiệm:**

- Xây dựng các mô phỏng kiểm chứng hiệu quả của mô hình và thuật toán tối ưu mạng truyền thông không dây hỗ trợ hoạt động của robot di động.
- Xây dựng các mô phỏng kiểm chứng thuật toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng.
- Xây dựng các mô phỏng kiểm chứng thuật toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông chưa đảm bảo.

#### **4. Đối tượng nghiên cứu và phạm vi nghiên cứu của luận án**

Đối tượng nghiên cứu của luận án là robot di động.

Phạm vi nghiên cứu: Điều hướng tối ưu cho robot di động trong môi trường truyền thông không dây.

**5. Cấu trúc của luận án** Luận án bao gồm phần "Mở đầu" trình bày lý do lựa chọn đề tài, mục tiêu và phạm vi nghiên cứu của luận án, và bốn chương chính với nội dung như sau:

**Chương 1:** "Tổng quan về việc lập quỹ đạo cho robot di động trong nhà máy", nghiên cứu tổng quan về việc lập quỹ đạo cho robot di động, tình hình nghiên cứu trong và ngoài nước, phân tích các công trình đã nghiên cứu trước đó theo nội dung đối tượng, phạm vi nghiên cứu của luận án, từ đó rút ra các hướng nghiên cứu thích hợp cho luận án.

**Chương 2:** "Nâng cao chất lượng truyền thông trong nhà máy phục vụ hoạt động của robot di động", nghiên cứu về vai trò của việc tối ưu hệ thống truyền thông không dây trong nhà máy, các vấn đề tồn tại và các giải pháp được nghiên cứu, từ đó đề xuất mô hình và thuật toán nâng cao chất lượng truyền thông không dây trong nhà máy phục vụ hoạt động của robot di động.

**Chương 3:** "Thuật toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng" nghiên cứu phương pháp xấp xỉ lồi để phát triển các thuật toán tối ưu quỹ đạo cho robot với mục tiêu tối thiểu năng lượng tiêu thụ cho robot di động.

**Chương 4:** "Thuật toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông có STAR-RIS hỗ trợ" nghiên cứu đề xuất sử dụng tấm phản xạ và truyền thông minh STAR-IRS (Simultaneously Transmitting and Reflecting - Intelligent Reflecting Surface) trong môi trường truyền thông đồng thời nghiên cứu, đề xuất thuật

toán dựa trên học sâu tăng cường DRL (Deep Reinforcement Learning) tối ưu quỹ đạo cho robot di động với mục tiêu tối thiểu năng lượng tiêu thụ cho robot di động đảm bảo ràng buộc về tốc độ truyền thông cho robot di động.

Phần **Kết luận và Kiến nghị** trình bày tóm tắt các đóng góp chính của luận án và hướng phát triển của luận án trong tương lai.

## **Chương 1. TỔNG QUAN VỀ VIỆC LẬP QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG NHÀ MÁY**

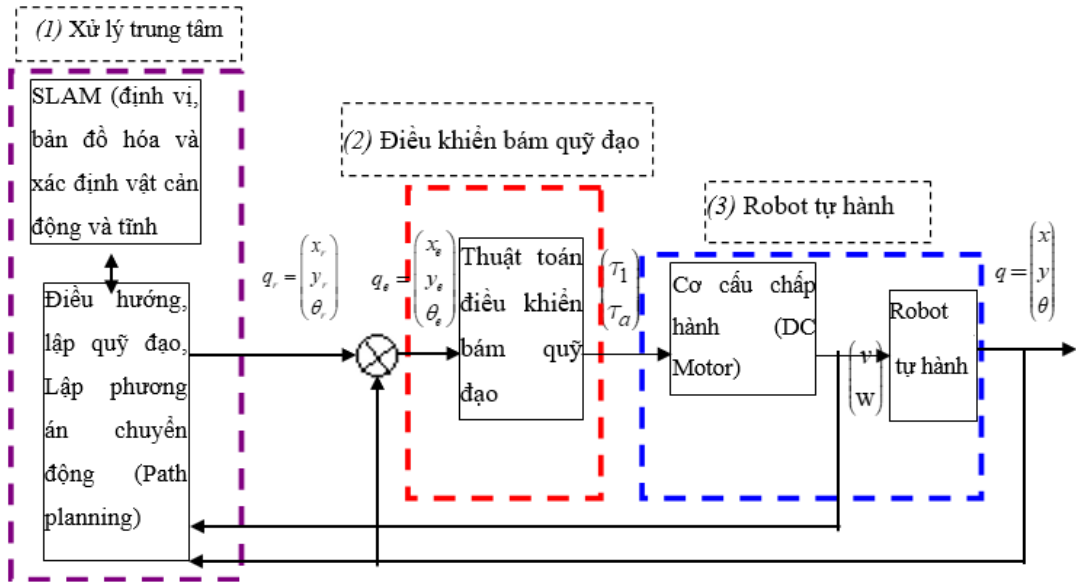
Robot di động (mobile robot) đã và đang đóng vai trò quan trọng trong các nhà máy và được đánh giá như một chìa khóa công nghệ trong sản xuất công nghiệp. Chương 1 trình bày tổng quan về robot di động và việc lập quỹ đạo cho robot di động trong môi trường nhà máy. Phần 1.2 mô tả vai trò của robot di động trong nhà máy. Bài toán lập quỹ đạo cho robot di động được trình bày ở phần 1.1. Phần 1.3 trình bày các nghiên cứu trong và ngoài nước về việc lập quỹ đạo cho robot di động trong môi trường truyền thông không dây tại nhà máy. Từ đó đưa ra những lập luận xây dựng hướng nghiên cứu của luận án ở phần 1.4. Cuối cùng là phần kết luận chương.

### **1.1 Bài toán lập quỹ đạo cho robot di động**

Trong suốt nhiều thập kỷ qua, lĩnh vực robot luôn được quan tâm, không ngừng phát triển và có ảnh hưởng lớn đến cuộc sống của con người [12]. Trong các nhà máy hiện đại, robot di động với khả năng di chuyển linh hoạt đóng một vai trò quan trọng trong các ứng dụng công nghiệp như robot vận chuyển hàng hóa, và các robot phục vụ các khâu chính trong dây chuyền sản xuất [13]. Chính vì vậy, robot di động đã và đang được nhiều nhà nghiên cứu cũng như các doanh nghiệp trong và ngoài nước quan tâm.

Nhìn chung, cấu trúc của điều khiển chuyển động cho một Robot di động có thể chia ra 3 giai đoạn: Định vị, lập quỹ đạo chuyển động (Localization and Path Planning), điều khiển bám quỹ đạo (Trajectory Tracking), và điều khiển động cơ robot (Motor Control) như Hình 1.1.

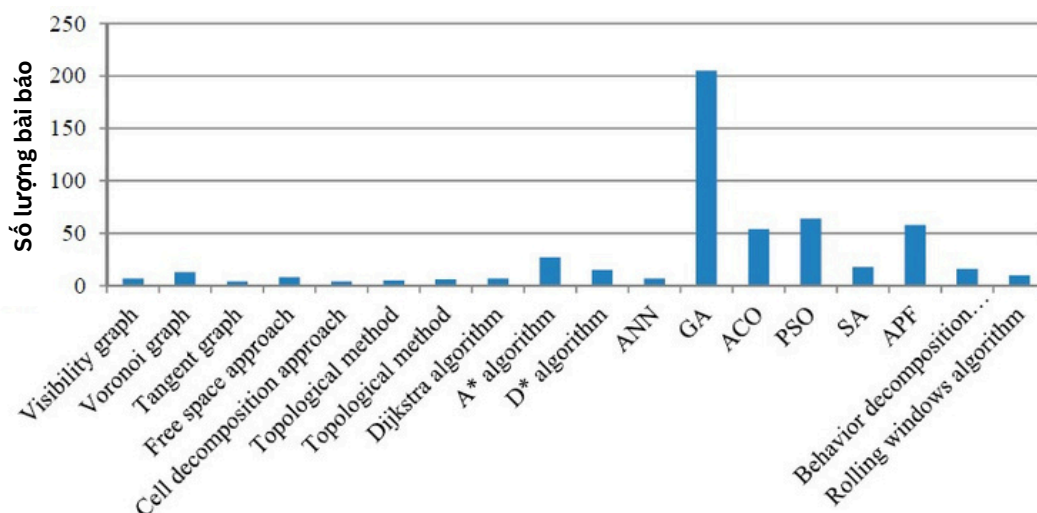
Dựa trên các tín hiệu trả về từ Lidar hoặc/và camera 3D, robot sẽ thực hiện việc định vị, bản đồ hóa và xác định các vật cản động và tĩnh trong môi trường làm việc. Dựa trên bản đồ đã biết trước cùng với các tín hiệu thu về từ các cảm biến, bộ xử lý trung tâm thực hiện lập quỹ đạo chuyển động cho robot bằng một thuật toán nào đó. Sau đó, robot sẽ thực hiện di chuyển theo quỹ đạo đã tính toán bằng các thuật toán điều khiển bám quỹ đạo. Các tín hiệu điều khiển bám quỹ đạo được đưa đến điều khiển các động cơ thực thi yêu cầu đảm bảo cho robot hoàn thành nhiệm vụ. Như vậy, có nhiều vấn đề cần quan tâm trong hệ thống điều khiển của robot AMR. Tuy nhiên, việc điều khiển chuyển động cho robot di động có hai vấn đề chính cần quan tâm đó là



Hình 1.1: Sơ đồ tổng quan hệ thống điều khiển robot di động, trong đó,  $\mathbf{q}_r$  là vị trí của robot,  $\mathbf{q}_e$  là sai số vị trí của robot,  $[\tau_1, \tau_\alpha]^T$  là mô men đặt vào động cơ,  $\bar{v}$  và  $\bar{w}$  lần lượt là vận tốc dài và vận tốc góc của robot [1].

lập quỹ đạo chuyển động và thiết kế bộ điều khiển bám quỹ đạo. Ở luận án này, NCS quan tâm đến bài toán lập quỹ đạo (path planning hay trajectory planning) cho robot di động. Đây là lĩnh vực được quan tâm với số lượng lớn công trình nghiên cứu trong nhiều thập kỷ qua.

Đầu tiên, việc lập quỹ đạo cho robot với mục đích là tìm ra quỹ đạo tốt nhất giúp robot có thể di chuyển đến đích mà không xảy ra va chạm. Hệ thống lập quỹ đạo cho robot bao gồm hai giai đoạn: lập quỹ đạo toàn cục (global planner) và lập quỹ đạo cục bộ (local planner). Quỹ đạo toàn cục làm nhiệm vụ xác định vị trí hiện tại của robot và mục tiêu cần di chuyển đến sau đó tìm ra quỹ đạo tốt nhất trên cơ sở xem xét các chướng ngại vật tĩnh từ bản đồ. Cụ thể, phần lập quỹ đạo toàn cục cho robot sử dụng thông tin về bản đồ môi trường, vị trí robot từ hệ thống nhận thức. Do đó tính chính xác khi lập quỹ đạo toàn cục phụ thuộc đáng kể vào hiệu năng của kỹ thuật SLAM (Simultaneous Localization and Mapping) được sử dụng. Trước tiên, robot cần xác định được vị trí của nó trong bản đồ tại thời điểm hiện tại. Lúc này, robot sẽ có những nhận thức về môi trường xung quanh, đặc biệt là bản đồ bao gồm những vật cản tĩnh đã xác định được ở hệ thống nhận thức. Tiếp đến, vị trí đích được xác định và được đặt trong vùng phạm vi đã xác định của bản đồ. Từ đó, các thuật toán lập quỹ đạo sẽ tính toán để đảm bảo robot có thể di chuyển tới vị trí đích với quãng đường ngắn nhất. Việc tối ưu quỹ đạo chuyển động cho robot được nghiên cứu



Hình 1.2: Số lượng các công trình nghiên cứu về lập quỹ đạo cho robot dựa trên số liệu từ Engineering Village [2].

nhằm nâng cao các chỉ số như quãng đường di chuyển, thời gian di chuyển, độ mượt mà khi di chuyển, hay năng lượng tiêu hao khi di chuyển của robot. Quỹ đạo toàn cục sẽ được đưa vào làm quỹ đạo tham chiếu cho phần lập quỹ đạo cục bộ để đảm bảo robot có thể di chuyển tới vị trí đích. Bởi vì, ngoài các chướng ngại vật tĩnh thì khi robot di chuyển còn có thể có những chướng ngại vật động xuất hiện. Khi đó cần kết hợp dữ liệu đọc từ cảm biến để lập quỹ đạo cục bộ nhằm tránh các chướng ngại vật động. Chính vì vậy việc lập quỹ đạo cục bộ còn có thể hiểu là thuật toán tránh vật cản động. Có nhiều thuật toán lập quỹ đạo chuyển động cho robot di động được đề xuất và áp dụng như thuật toán Dijkstra, A\*, D\*, thuật toán di truyền GA (Genetic Algorithm),...

Cần nhấn mạnh thêm rằng việc lập quỹ đạo hay điều hướng cho robot di động là một trong những vấn đề quan trọng của điều khiển robot. Bắt đầu từ giữa những năm 1960, vấn đề này đã nhận được sự quan tâm của các nhà nghiên cứu. Việc lập quỹ đạo cho robot trong một môi trường làm việc được mô tả là việc tìm quỹ đạo tối ưu hoặc cận tối ưu từ điểm ban đầu cho đến điểm đích đến để đạt hiệu quả về quãng đường, thời gian, năng lượng hay các yếu tố khác [14]. Một quỹ đạo tốt không chỉ tiết kiệm thời gian di chuyển mà còn giảm được các vấn đề về hao mòn và đầu tư chi phí cho robot. Do đó, đã có nhiều thuật toán lập quỹ đạo cho robot được nghiên cứu và đề xuất. Hình 1.2 chỉ ra số lượng công bố liên quan đến việc lập quỹ đạo cho robot di động, số liệu được lấy từ Engineering Village [2].

Có nhiều tiêu chí để đánh giá quỹ đạo của robot di động như thời gian di chuyển,

quãng đường di chuyển, tổng mức năng lượng tiêu thụ,... Đối với robot di động, việc di chuyển thường xuyên dẫn đến việc phải trang bị các nguồn pin tháo rời. Do đó, mức năng lượng tiêu thụ của nguồn pin gắn trên thân robot là một vấn đề cần quan tâm. Các robot di động cần nguồn năng lượng lớn cho hệ thống cảm biến, hệ thống điều khiển và hệ thống chuyển động. Do đó, trong bất kỳ ứng dụng nào, robot di động cũng đều được kỳ vọng có hiệu suất sử dụng pin cao để tăng thời gian làm việc của robot. Trên thực tế, yêu cầu về nguồn pin sử dụng cho robot cũng ngày càng tăng lên, dẫn đến việc tăng chi phí tại khâu này. Điều này gây ra một áp lực lớn cho các nhà thiết kế robot trong việc phát triển hệ thống nâng cao hiệu quả sử dụng nguồn pin và một áp lực khác lên người sử dụng khi vận hành hệ thống. Chính vì vậy, tối ưu năng lượng tiêu thụ là một vấn đề lớn cần quan tâm. Để tối thiểu mức năng lượng tiêu thụ và tăng cường hiệu quả sử dụng năng lượng cho robot, các vấn đề liên quan đến thiết kế phần cứng và phát triển phần mềm đều đã được đề xuất. Trong đó, việc tối ưu quỹ đạo là một trong những giải pháp.

Thuật toán Dijkstra được đề xuất bởi E.W. Dijkstra vào năm 1959 [15]. Đây là một trong những thuật toán cổ điển để giải quyết bài toán tìm đường đi ngắn nhất từ một điểm cho trước tới tất cả các điểm còn lại trong bản đồ có trọng số không âm. Trọng số không âm của các cạnh nối hai đỉnh gần nhau trong đồ thị có thể chính là khoảng cách giữa chúng hoặc có thể xem xét đến các yếu tố khác liên quan đến việc di chuyển dọc cạnh. Thuật toán Dijkstra ban đầu dựa trên một chiến lược tham lam để tìm đường đi ngắn nhất. Thuật toán chỉ quan tâm đến quãng đường ngắn nhất mà không quan tâm đến tính khả thi của giải pháp. Thuật toán Dijkstra là một giải pháp phù hợp cho môi trường tĩnh hoặc cho việc lập quỹ đạo toàn cục. Dijkstra cũng có thể được ứng dụng trong môi trường động. Trong trường hợp này, thông tin về môi trường không được biết hoàn toàn, do đó thông tin về các đỉnh và các vật cản động được tính toán trong quá trình di chuyển. Tuy nhiên, chỉ sử dụng thuật toán Dijkstra sẽ không thể thực hiện được trong môi trường động. [16].

Năm 1968, các tác giả của [17] đã đề xuất thuật toán A\*. Thuật toán A\* được phát triển dựa trên thuật toán Dijkstra. Bắt đầu từ một điểm xác định, trọng số đánh giá điểm hiện tại được cập nhật trên cơ sở lấy điểm có trọng số nhỏ nhất. Điểm quan trọng của thuật toán A\* là hàm đánh giá vị trí điểm  $n$

$$f(n) = g(n) + h(n), \quad (1.1)$$

trong đó  $g(n)$  thể hiện chi phí của đường đi cho đến điểm hiện tại  $n$ , và  $h(n)$  là hàm đánh giá chi phí nhỏ nhất đi từ điểm hiện tại đến đích. Mục tiêu của thuật toán  $A^*$  là thiết kế đánh giá được chi phí cho hàm  $f(n)$  từ vị trí ban đầu đến vị trí đích qua điểm  $n$ . Khoảng cách Euclidean giữa hai điểm thường được sử dụng làm giá trị của  $h(n)$ . Trong khi đó, giá trị của  $g(n)$  là một hằng số, do đó, giá trị của hàm  $f(n)$  chủ yếu phụ thuộc vào  $h(n)$ . Khi càng gần điểm đích thì giá trị  $h(n)$  càng nhỏ, theo đó, giá trị của  $f(n)$  cũng nhỏ dần. Kết quả là sẽ tạo ra quỹ đạo với khoảng cách nhỏ nhất. Thuật toán  $A^*$  yêu cầu thông tin vị trí đích và tìm kiếm dọc theo đường hướng đến đích. Do đó, việc tìm đường bằng thuật toán  $A^*$  sẽ có hiệu quả cao hơn so với thuật toán Dijkstra. Với sự phát triển của trí thông minh nhân tạo, thuật toán  $A^*$  được cải tiến và điều chỉnh cho phù hợp với nhiều ứng dụng khác nhau bao gồm cả thuật toán tìm đường cho robot, giao thông thông minh, lý thuyết đồ thị và điều khiển tự động [18].

Nhìn vào biểu thức (1.1) ta thấy có một sự đánh đổi giữa tốc độ tính toán và sai số của thuật toán  $A^*$ . Để tăng tốc độ tính toán cần có bộ nhớ đủ lớn, nếu không tốc độ tính toán sẽ giảm. Để khắc phục vấn đề này, thuật toán hierarchical  $A^*$  đã thực hiện chia việc tính toán thành nhiều chu trình, sau đó gộp các kết quả lại để có quỹ đạo toàn cục.

Thuật toán  $D^*$  (Dynamic  $A^*$ ) được Stentz [19] đề xuất thuật vào năm 1994, về cơ bản tương tự như thuật toán  $A^*$  nhưng có xét đến các yếu tố thay đổi trong môi trường khi robot di chuyển. Thuật toán  $D^*$  bao gồm hai giai đoạn chính. Giai đoạn một là sử dụng thuật toán  $A^*$  để tìm kiếm đường đi ban đầu. Giai đoạn hai là cập nhật đường đi khi môi trường thay đổi. Ở giai đoạn này, các thông tin về sự thay đổi trong môi trường để cập nhật các giá trị ước tính của khoảng cách và chi phí di chuyển giữa các điểm. Từ đó, sử dụng thuật toán  $A^*$  để tìm kiếm đường đi mới trên cơ sở sử dụng các giá trị ước tính mới được cập nhật. Việc này được lặp lại liên tục cho đến khi tới đích.

Thuật toán  $A^*$  cũng như Dijkstra có một nhược điểm chung là trên bản đồ cần xác định trước các điểm. Đây chính là lý do quỹ đạo sinh ra không được mượt mà. Để khắc phục vấn đề này, [20] đưa ra một hàm đánh giá như sau

$$f(n) = K_1g(n) + K_2h(n) + K_3p(n), \quad (1.2)$$

trong đó  $K_1$ ,  $K_2$ ,  $K_3$  là các trọng số dương,  $g(n)$  và  $h(n)$  có ý nghĩa giống như trong

thuật toán  $A^*$  ban đầu, và  $p(n)$  là một nhân tố phạt dựa trên vị trí thực tế.

Các thuật toán lập quỹ đạo kể trên đều là các thuật toán dựa trên việc rời rạc hóa môi trường làm việc của robot. Như vậy, robot cần biết trước bản đồ môi trường làm việc, rồi dựa vào việc tính toán với các vị trí trên bản đồ để đưa ra đường đi cho robot. Điều này dẫn đến yêu cầu tốc độ xử lý của CPU cao và dung lượng bộ nhớ lớn. Hơn thế nữa, cũng chính vì vấn đề rời rạc hóa môi trường nên quỹ đạo được sinh ra không đảm bảo được tính mượt mà. Thuật toán di truyền GA được đề xuất bởi Holland vào năm 1975 sẽ giúp giải quyết những hạn chế kể trên. Đây là một thuật toán lập quỹ đạo toàn cục được lấy cảm hứng từ cơ chế tiến hóa của sinh vật trong tự nhiên. Trong GA, tất cả các giải pháp có thể của vấn đề được mã hóa thành các nhiễm sắc thể, và tất cả các nhiễm sắc thể tạo thành một quần thể ban đầu. Một số hoạt động cơ bản được xây dựng: lai ghép, đột biến và lựa chọn. Quần thể ban đầu được tạo ra, sau đó giá trị thích hợp của mỗi cá thể được tính bằng các mục tiêu. Các cá thể được chọn cho hoạt động lai ghép, hoạt động đột biến và hoạt động lựa chọn được xác định bằng giá trị thích hợp. GAs có thể được áp dụng với ưu điểm là các thuật toán như vậy bao phủ một không gian tìm kiếm lớn và sử dụng tài nguyên bộ nhớ và CPU tối thiểu. Họ cũng có khả năng thích nghi với môi trường thay đổi. Một nhược điểm là giải pháp được tìm thấy cho vấn đề tối ưu hóa không phải lúc nào cũng là một giá trị tối ưu toàn cục (đường đi ngắn nhất tổng thể).

Vì thuật toán cũng có thể được áp dụng trong môi trường động, giải pháp tối ưu tìm ra một đường đi không va chạm giữa hai điểm cần được cập nhật liên tục theo sự thay đổi trong môi trường xung quanh. Trong trường hợp này, các phương pháp tiến hóa hội tụ vào một giải pháp tối ưu. Tất cả các giải pháp có thể được biểu diễn dưới dạng cá nhân trong một quần thể, với mỗi gen đại diện cho một tham số. Các cá nhân được hình thành bởi một bộ gen hoàn chỉnh.

Một thế hệ mới được hình thành bằng cách chọn lựa các cá nhân tốt nhất từ thế hệ cha mẹ và áp dụng các toán tử di truyền, như lai ghép và đột biến, để khám phá các giải pháp bổ sung. Mỗi con cái từ thế hệ mới được kiểm tra bằng một hàm thích nghi được lập ra cho vấn đề đó. Từ tất cả các con cái, các cá nhân tốt nhất được chọn làm cha mẹ của thế hệ tiếp theo. Các tác giả trong [21] mô tả cách sử dụng phương pháp GA để giải quyết vấn đề lập quỹ đạo chuyển động trong một môi trường tĩnh. Kết quả của họ cho thấy rằng phương pháp được sử dụng là hiệu quả trong việc xử lý vấn đề lập quỹ đạo chuyển động trong các môi trường tĩnh khác nhau.



lập quỹ đạo chuyển động là một loại ánh xạ từ không gian trạng thái đến không gian hành vi, và mạng nơ-ron nhân tạo ANN (Artificial Neural Network) có thể biểu diễn mối quan hệ ánh xạ này. Mạng nơ-ron được sử dụng để mô tả các ràng buộc giữa môi trường, và năng lượng được xác định dưới dạng hàm của điểm đường đi. Các tác giả ở [22] đã sử dụng ANN để giải quyết vấn đề lập quỹ đạo chuyển động của robot và thảo luận về cách mạng nơ-ron có thể góp phần làm tăng hiệu suất cho thuật toán lập quỹ đạo chuyển động của robot.

## 1.2 Robot di động và vấn đề truyền thông không dây

Theo xu hướng phát triển IoT (Internet of Things) trong công nghiệp IIoT (Industrial IoT), các thiết bị trong nhà máy có xu hướng tích hợp các giao thức truyền thông. Sự kết nối của hàng loạt các thiết bị và yêu cầu về khả năng thu thập/chia sẻ dữ liệu thúc đẩy yêu cầu về chất lượng truyền thông giữa các thiết bị. Theo đó, robot di động cần đáp ứng được các yêu cầu về truyền thông trong nhà máy. Tức là, các robot di động tại các nhà máy vừa cần hoàn thành nhiệm vụ di chuyển, vừa cần thực hiện trao đổi thông tin với các thiết bị khác như trạm chủ trung tâm, các robot khác trong hệ đa robot hay các thiết bị không dây khác.

Hiện nay, tại các nhà máy, các ứng dụng tự động hóa nhà máy đang bị chi phối mạnh mẽ bởi các công nghệ có dây như PROFIBUS/PROFINET, SERCOS, HART và CAN [23]. Việc truyền thông có dây dẫn đến nhiều vấn đề phức tạp kéo theo về thiết kế, thi công, vận hành cũng như quá trình bảo trì, bảo dưỡng. Ngoài ra, các thiết bị di động như robot di động việc sử dụng truyền thông có dây là rất khó thậm chí không thể thực hiện. Do vậy, việc thực hiện truyền thông không cho các robot di động là một giải pháp tối ưu. Tuy nhiên, trên thực tế hiện nay, công nghệ truyền thông không dây lại đang không hoặc rất ít được sử dụng trong các ứng dụng tự động hóa nhà máy. Nguyên nhân được chỉ ra trong [11]. Dự án KoI [11] đã tiến hành một cuộc khảo sát chi tiết dựa trên bảng câu hỏi để thu thập thông tin trực tiếp từ các nhà sản xuất hàng đầu, những người tham gia vào một loạt các quy trình tự động hóa nhà máy. Bảng 1.1 tóm tắt các kết quả chính từ cuộc khảo sát về các yêu cầu truyền thông cho các ứng dụng trong nhà máy, tùy thuộc vào các ứng dụng cụ thể, các yêu cầu này có thể khác nhau. Dựa trên dữ liệu này, có thể thấy rằng, trong số nhiều nguyên nhân được đưa ra, nguyên nhân cốt lõi là do các công nghệ không dây hiện tại không đáp ứng được yêu cầu khắt khe của các ứng dụng tự động hóa nhà máy với độ tin cậy cực kỳ cao từ

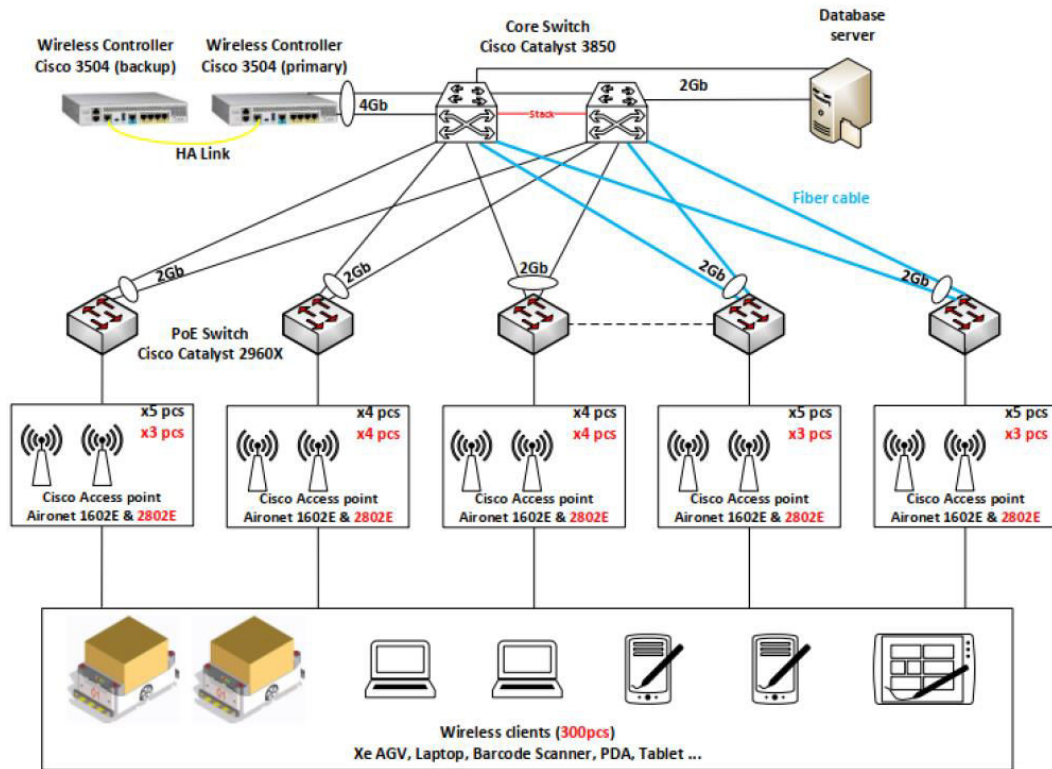
$1 - 10^{-9}$  cùng các ràng buộc về độ trễ rất thấp xuống đến 1 ms.

Bảng 1.1: Các yêu cầu đối với truyền thông không dây cho ứng dụng trong nhà máy tự động hóa tổng hợp bởi dự án KoI năm 2016 [11].

Độ trễ	Độ tin cậy	Kích thước dữ liệu	Vùng truyền thông giữa các thiết bị	Số lượng thiết bị trên một nhà máy	Độ di động của các thiết bị (trong nhà)
Kết quả tổng hợp					
1 – 50 ms	$1 - 10^{-6}$ đến $1 - 10^{-9}$	10 – 300 bytes	2 – 100 m	10 – 1000	0 – 10 m/s
Với các ứng dụng cho quy trình sản xuất					
<10 ms	$1 - 10^{-9}$	<50 bytes	<100 m	<1000	~ 1 m/s
Với các ứng dụng cho xe tự hành					
10 – 50 ms	$1 - 10^{-6}$ đến $1 - 10^{-9}$	<300 bytes	~ 2 m	<1000	<10 m/s

Vấn đề kể trên là một trong những nguyên nhân thúc đẩy sự phát triển của mạng di động thế hệ thứ 5 và các thế hệ sau đó (5G/B5G). Một điểm quan trọng của 5G là hỗ trợ thiết bị truyền thông siêu tin cậy và độ trễ thấp URLLC (ultra-reliable low-latency communication), được sử dụng cho các ứng dụng phẫu thuật từ xa và ứng dụng quan trọng trong tự động hóa công nghiệp. Ở phiên bản 5G hiện tại, độ trễ truyền thông trung bình khoảng 4 ms, giảm một nửa so với các mạng 4G và tốt hơn so với các mạng công cộng (khoảng 15 ms). Ở các phiên bản 5G URLLC thiết kế đặc biệt cho việc vận hành trong công nghiệp, độ trễ giảm xuống còn 1 ms trung bình và có thể so sánh được với các kết nối có dây [24]. Do đó, hệ thống mạng 5G/B5G hứa hẹn tạo ra một thế giới kết nối tin cậy, hiệu quả trong công nghiệp [11].

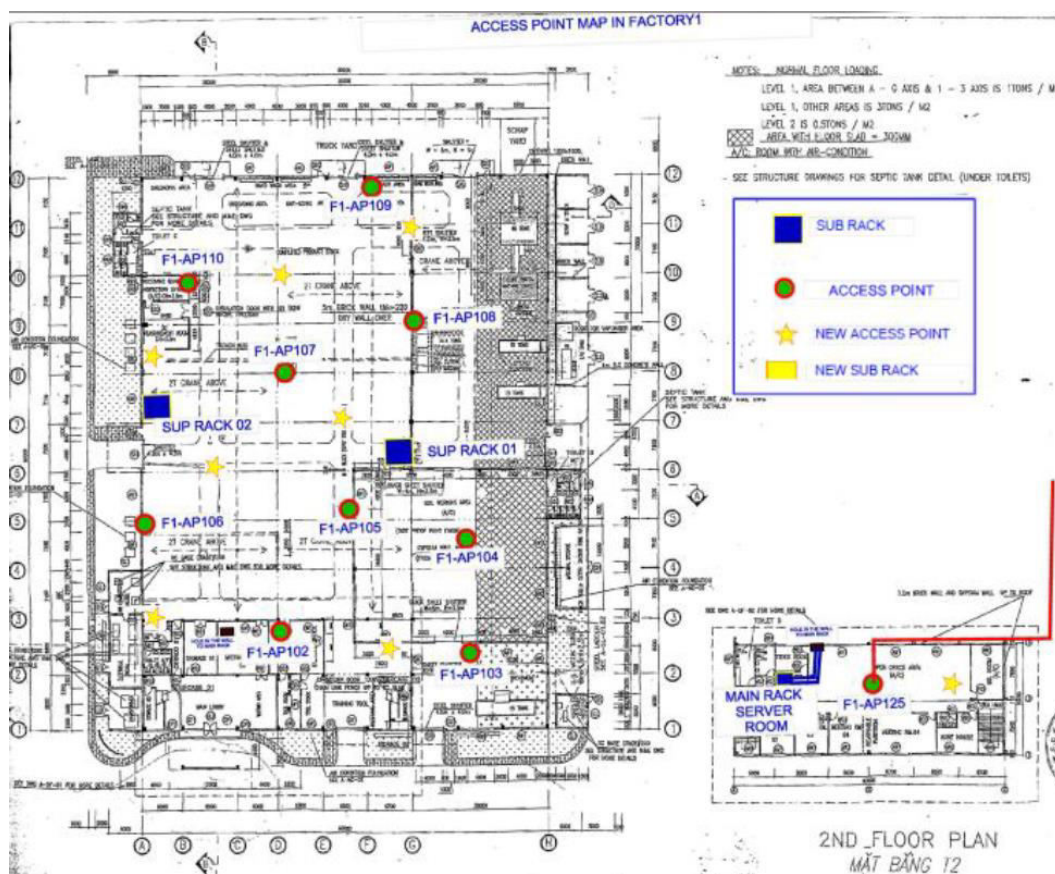
Theo [25], hiện nay có thể chia robot di động thành hai loại đó là robot di động tự hành AMR (Autonomous Mobile Robot) và robot di động kết nối CMR (Connected Mobile Robot). Được trang bị các cảm biến và nguồn lực tính toán, các robot tự hành AMR có thể tự thực hiện các nhiệm vụ của mình một cách độc lập. Vấn đề trở nên phức tạp khi các AMR phải đối mặt với vấn đề là chúng cần có bộ nhớ đủ lớn, cần có nguồn tài nguyên tính toán lớn, và cần có khả năng thực hiện được những thuật toán AI (Artificial Intelligent) phức tạp để thực hiện những nhiệm vụ khó khăn. Với sự phát triển mạnh mẽ của hệ thống mạng 5G/B5G, các robot di động kết nối mạng trở thành các thiết bị người dùng được phục vụ bởi các BS/AP được kỳ vọng là sẽ giải quyết được tất cả các vấn đề trên [25]. Khi đó, các robot không cần thiết phải trang bị nguồn lực tính toán cũng như bộ nhớ lớn. Khi kết nối trong hệ thống mạng, robot gửi các dữ liệu ngay khi nó nhận được thông qua các cảm biến (có thể là hình



Hình 1.3: Hệ thống Wifi tại nhà máy SHI-V 1 thiết kế bởi Netmarks [3].

ảnh hoặc video) đến cho trạm điều khiển (người vận hành), sau đó, trạm điều khiển hay người vận hành sẽ gửi các hướng dẫn dựa trên các dữ liệu nhận được đó đến cho CMR. Ngoài ra, khi số lượng robot trong các nhà máy tăng lên, cùng với các thiết bị IoT khác, các thiết bị trong nhà máy sẽ cùng hoạt động trong một không gian làm việc chung. Khi đó, các robot cần phối hợp với nhau để tạo thành hệ đa robot hay robot bầy đàn để đạt hiệu quả cao khi thực hiện các nhiệm vụ trong nhà máy. Do đó, một hệ thống các robot được kết nối mạng là một giải pháp hiệu quả đồng thời giảm bớt gánh nặng về bộ nhớ và tài nguyên tính toán cho từng robot. Với sự phát triển mạnh mẽ của hệ thống mạng 5G/B5G, hệ thống mạng được kỳ vọng rằng tốc độ truyền lớn, độ trễ thấp, và độ tin cậy cao, các CMR được kỳ vọng là sẽ trở thành một chìa khóa công nghệ trong tương lai.

Tuy nhiên, trong môi trường nhà máy, đường truyền có thể bị chặn bởi rất nhiều chướng ngại vật như các máy móc, thiết bị, tường, khu vực kho bãi có các vật liệu được xếp chồng lớn... Điều này có thể sẽ tạo ra các vùng chết tín hiệu, khi robot di chuyển trong vùng này, các đường truyền trực tiếp từ thiết bị phát sóng không dây BS/AP sẽ bị chặn lại, do đó, robot hay trạm trung tâm sẽ không nhận được tín hiệu của nhau. Đối mặt với vấn đề này, các nhà cung cấp dịch vụ đưa ra giải pháp là triển



Hình 1.4: Sơ đồ vị trí đặt AP tại nhà máy SHI-V 1 thiết kế bởi Netmarks [3].

khai nhiều BS hoặc AP. Một ví dụ cụ thể như hình 1.3 thể hiện một giải pháp mà công ty Netmarks đưa ra cho nhà máy Sumitomo tại Khu công nghiệp Thăng Long vào năm 2019.

Trước đó tại nhà máy SHI-V 1, hệ thống Wifi sử dụng AP loại Cisco Aironet 1602E với 9 APs. Tuy nhiên, cùng với sự phát triển của công ty, số lượng thiết bị trong hệ thống mạng đã tăng lên đồng thời đòi hỏi một tốc độ truyền thông lớn hơn đối với các thiết bị trong đó có các AGVs. Khi thiết kế lại, Netmarks đã nâng cấp AP từ Cisco Aironet 1602E lên 2802E để đạt tín hiệu và băng thông tốt hơn. Thêm vào đó, Netmarks cũng đã tăng thêm số lượng APs như trong hình 1.4 đã thể hiện.

Đây cũng được đánh giá là một giải pháp mang lại hiệu quả cao tại điểm đó. Hệ thống sau khi được triển khai đã hoạt động ổn định. AP truyền các thông tin về nhiệm vụ cho các robot, các robot nhận thông tin, thực hiện nhiệm vụ và gửi thông tin về tình trạng thực hiện nhiệm vụ, vị trí của robot, tình trạng nguồn năng lượng của robot... Tuy nhiên, tại nhà máy khi đó, các AGVs vẫn được hỗ trợ dẫn đường bởi QR code dán trên sàn nhà máy. Hiện tại, các robot được kỳ vọng là linh hoạt hơn và có khả năng tìm đường tối ưu cho từng nhiệm vụ, do đó các AGVs sẽ được thay thế

bởi các robot tự hành AMRs và được kết nối mạng trở thành các robot được kết nối mạng CMRs. Khi đó, có thể việc tăng một lượng lớn các AP có thể không chỉ đi kèm với chi phí tăng lên đáng kể mà còn không giải quyết hết vấn đề liên quan đến vùng chết tín hiệu.

Từ những vấn đề đã đề cập ở phần trên có thể thấy rằng, ứng dụng truyền thông không dây trong các nhà máy là một xu hướng tất yếu trong tiến trình hướng đến nhà máy trong tương lai. Hệ thống mạng 5G/B5G là một giải pháp hứa hẹn mang lại hiệu quả cao. Để đạt được hiệu quả như vậy, có một số vấn đề kéo theo như sau [8].

- Triển khai ngày càng nhiều nút BS/AP, trạm trung chuyển và tăng cường các ăng ten để đạt được phạm vi và khả năng phủ sóng mạng. Điều này gây ra lượng tiêu thụ năng lượng và chi phí triển khai và bảo trì cao hơn, cũng như vấn đề nhiễu mạng nghiêm trọng và phức tạp;
- Khi tăng một số lượng lớn ăng ten tại các BSs/APs/trạm trung chuyển để tận dụng các lợi ích từ hệ thống nhiều đầu vào - nhiều đầu ra M-MIMO (massive multiple-input-multiple-output), điều này đòi hỏi chi phí phần cứng và năng lượng tăng lên cùng với sự phức tạp trong xử lý tín hiệu;
- Chuyển sang các dải tần cao hơn như sóng millimet (mmWave) và thậm chí là tần số terahertz (THz) để tận dụng băng thông lớn và sẵn có của chúng, điều này tất yếu dẫn đến việc triển khai thêm nhiều nút hoạt động hơn và thậm chí nhiều ăng ten hơn để bù đắp cho mất mát truyền dẫn lớn từ khoảng cách.

Do đó, để đạt được sự tăng trưởng thông lượng bền vững của các mạng không dây trong nhà máy phục vụ hoạt động cho các robot di động với chi phí, độ phức tạp và tiêu thụ năng lượng thấp thì việc phát triển các công nghệ mới cùng với tối ưu hệ thống mạng là một vấn đề cấp bách.

### 1.3 Tình hình nghiên cứu trong và ngoài nước

#### 1.3.1 Tình hình nghiên cứu trong nước

Tại Việt Nam, việc lập quỹ đạo cho robot đã và đang được quan tâm trong công nghiệp và nghiên cứu. Trong công nghiệp, hiện nay, phần lớn robot trong hầu hết các nhà máy đều đang di chuyển với quỹ đạo đặt sẵn. Robot được dẫn đường nhờ các đường từ tính đặt trước dưới sàn di chuyển hoặc sử dụng các mã QR code trên sàn

hay dùng một số công cụ dẫn đường khác. Tuy nhiên, các nghiên cứu vẫn đang nỗ lực với kỳ vọng robot có khả năng tự đưa ra được các quyết định tối ưu cho các nhiệm vụ.

Các tác giả của [26] đã đề xuất thuật toán  $\eta^3 - spline$  với đa thức bậc 7 để thiết kế quỹ đạo tối ưu cho bài toán vận chuyển kệ hàng trong kho. Các thử nghiệm được thực hiện trên robot tự hành Pioneer 3-DX cho thấy tính hiệu quả của thuật toán. Trong khi đó, các giả ở [27] đã xây dựng được bản đồ sử dụng cảm biến Lidar và lập quỹ đạo cho robot sử dụng giải thuật  $A^*$  tránh vật cản cho mô hình robot di động 3 bánh. Kết quả thử nghiệm cho thấy tính khả thi của thuật toán sử dụng và có khả năng mở rộng cho các hướng nghiên cứu tiếp theo. Cũng sử dụng thuật toán  $A^*$  nhưng tác giả của [28] đã kết hợp với kỹ thuật dải đàn hồi thời gian TEB (timed elastic band) để giải quyết vấn đề lập quỹ đạo trong môi trường phức tạp, có vật cản động. Đầu tiên, thuật toán  $A^*$  được sử dụng để lập quỹ đạo toàn cục cho robot. Sau đó, kỹ thuật TEB được sử dụng để lập quỹ đạo cục bộ cho robot mỗi lần gặp vật cản. Thuật toán TEB sẽ tối ưu hóa đường đi cục bộ của robot để đảm bảo một quỹ đạo an toàn và hiệu quả khi vượt qua các chướng ngại vật. Các tác giả ở [29] trình bày một tiếp cận tránh vật cản cho robot di động trong nhà kho. Cụ thể, từ việc tính toán xác suất va chạm có thể xảy ra giữa robot và các đối tượng khi robot di chuyển, robot ra quyết định lựa chọn vùng an toàn để di chuyển nhằm tránh chướng ngại vật. Kết quả thử nghiệm và mô phỏng ban đầu cho thấy khả năng thực thi của thuật toán.

Ngoài ra, tác giả của [30] đã đề xuất hệ thống dải đàn hồi thời gian tích cực PTEB (proactive timed elastic band) để lập quỹ đạo cho robot di động trong môi trường động. Hệ thống PTEB kết hợp những điểm mạnh của kỹ thuật TEB và mô hình tạo môi trường vật cản động HRVO (hybrid reciprocal velocity obstacle). Hệ thống hoạt động tốt, robot được điều hướng có quỹ đạo an toàn trong môi trường vật cản động. Cũng cùng tác giả, một thuật toán điều hướng cho robot trong môi trường bằng mô hình cửa sổ động mở rộng EDWA (enhanced dynamic window approach) ở công trình [31]. Mô hình EDWA kết hợp lợi ích của kỹ thuật cửa sổ động DWA và mô hình HRVO. Kết quả mô phỏng và thực nghiệm đều cho thấy robot hoàn toàn tránh được vật cản động xung quanh robot.

### 1.3.2 Tình hình nghiên cứu ngoài nước

#### Vấn đề lập quỹ đạo cho robot di động

Đối với robot di động, lập quỹ đạo là một trong những vấn đề quan trọng giúp robot có thể thực hiện các nhiệm vụ một cách đáng tin cậy và an toàn. Đây là một vấn đề đã dành được sự quan tâm lớn của các nhà nghiên cứu trong vài thập kỷ qua. Để đảm bảo các robot không va chạm với vật cản trong quá trình di chuyển, các nghiên cứu trước đây về lập quỹ đạo cho robot đã đề xuất nhiều thuật toán cho các kịch bản ứng dụng khác nhau. Bằng cách chia không gian liên tục thành một mạng lưới với các điểm hữu hạn, các thuật toán lập quỹ đạo cho robot hiệu quả, bao gồm các thuật toán Dijkstra, A\*, và D\*, đã được phát triển để tìm đường đi ngắn nhất giữa hai vị trí trong môi trường tĩnh và động [32].

Với mục đích giảm năng lượng tiêu thụ và tăng hiệu quả sử dụng nguồn pin cho robot di động, các giải pháp thiết kế phần cứng cùng phần mềm đã được đề xuất trong nhiều công trình nghiên cứu.

Về các giải pháp phần cứng, các tác giả trong [33] đã thiết kế phần cứng cho tiêu hao năng lượng thấp để giảm mức năng lượng tiêu hao chung của các hệ thống trên robot. Ngoài ra, các tác giả trong [34] đã thiết kế một nguồn pin DC mới dành riêng cho robot công nghiệp giúp giảm năng lượng tiêu hao cho robot. Thêm vào đó, nhiều công bố đã đề xuất giải pháp quản lý mức năng lượng tiêu thụ, hệ thống truyền điện năng không dây và giải pháp chuyển đổi nguồn pin với mục đích giảm mức tiêu hao điện năng và tăng thời gian làm việc cho robot như ở [35]. Các tác giả ở [36] đã giới thiệu hai kỹ thuật chuyển đổi năng lượng là quản lý nguồn năng lượng động và bảng kế hoạch thời gian thực để tiết kiệm điện năng. Kết quả thực nghiệm được thực hiện trên robot Pioneer 3DX, và cho thấy các kỹ thuật đề xuất hứa hẹn sẽ làm giảm mức năng lượng tiêu thụ và kéo dài thời gian làm việc cho robot. Khác với các nghiên cứu kể trên, một số tác giả khác lại quan tâm đến việc tiết kiệm điện năng bằng cách tối ưu kế hoạch chuyển động cho robot như đề xuất trong [37]. Kết quả mô phỏng ở [37] chỉ ra rằng việc tối ưu kế hoạch chuyển động có thể tiết kiệm điện năng đến 10%. Việc lập quỹ đạo chuyển động sử dụng hiệu quả điện năng cho robot được đề xuất trong [38], có thể tiết kiệm lên đến 42% điện năng tiêu thụ.

Một quỹ đạo tốt có thể giúp giảm đáng kể năng lượng cần tiêu thụ của robot. Tuy nhiên, bài toán tối ưu quỹ đạo nhằm tối thiểu năng lượng tiêu thụ cho robot là

một thách thức bởi đây là một bài toán tối ưu không lồi. Do vậy, các tác giả của [39] đã kết hợp độ dài quãng đường di chuyển ngắn nhất vào trong hàm mục tiêu để tối thiểu năng lượng tiêu thụ sử dụng thuật toán Newton. Một cách tiếp cận khác là dựa trên thuật toán A\* như [40,41]. Các tác giả của [40] đề xuất một thuật toán A\* tăng hiệu suất sử dụng điện năng (EA\*) cho việc tối ưu đường đi cho robot. Trong công bố này, mặc dù các tác giả chỉ quan tâm đến điện năng tiêu thụ tại các điểm dừng và rẽ của robot, các kết quả mô phỏng cho thấy thuật toán EA\* thật sự hiệu quả trong việc tiết kiệm năng lượng. Các tác giả ở [41] cũng đã thiết kế đường đi cho robot bằng cách sử dụng thuật toán A\* với tiêu chí tối thiểu năng lượng được tích hợp trong hàm chi phí. Tuy nhiên, các nghiên cứu này mới dừng lại ở việc giải quyết bài toán trong môi trường chỉ có vật cản tĩnh. Ngoài ra, cũng có nhiều công bố đề xuất thuật toán tối ưu quỹ đạo cho robot có quan tâm đến môi trường có vật cản động như [42]. Mặc dù vậy, các công bố này lại chưa quan tâm đến năng lượng tiêu thụ cho robot. Thuật toán GA cũng được sử dụng để thực hiện tối ưu đường đi và giảm năng lượng tiêu thụ cho robot [43]. Công bố [44] đã phát triển thuật toán vận tốc tối ưu để tiết kiệm năng lượng cho robot hai bánh. Trong công bố này, các tác giả quan tâm đến mô hình năng lượng kết hợp giữa mô hình động học của robot và động cơ. Việc tối ưu tốc độ di chuyển cho robot cũng được sử dụng cho việc quản lý năng lượng ở [45].

Trong các kịch bản môi trường làm việc không biết trước hoặc chỉ biết trước một phần, thuật toán dựa trên DRL được đề xuất sử dụng để điều hướng cho robot. Các tác giả trong [46] đề xuất phương pháp sử dụng Q-learning kết hợp với APF (artificial potential field) lập quỹ đạo cho robot di động đảm bảo di chuyển từ vị trí ban đầu đến vị trí đích và tránh vật cản. Kết quả mô phỏng cho thấy thuật toán học QAPF đã cải thiện được 18,83% về độ dài đường đi, 169,75% về độ mượt của đường đi và 74,84% về thời gian huấn luyện so với phương pháp cổ điển. Cùng một mô hình hệ thống và bài toán tối ưu như [46], tuy nhiên, các tác giả trong [47] đề xuất một phương pháp học tăng cường dựa trên PPO (Proximal Policy Optimization) để giải quyết bài toán tối ưu của robot trong môi trường chưa biết trước. Cụ thể, robot di động không có thông tin về bản đồ của môi trường, chỉ có thông tin về môi trường xung quanh được thu được bởi 2D-Lidar. Các môi trường huấn luyện được thiết lập trong MuJoCo. Bằng cách kết hợp bộ điều khiển với thuật toán lập quỹ đạo chuyển động cho robot, kết quả mô phỏng đã cho thấy hiệu quả của phương pháp đề xuất. Cũng xem xét vấn đề lập quỹ đạo, các tác giả trong [48] đề xuất một phương pháp dựa trên DRL kết hợp



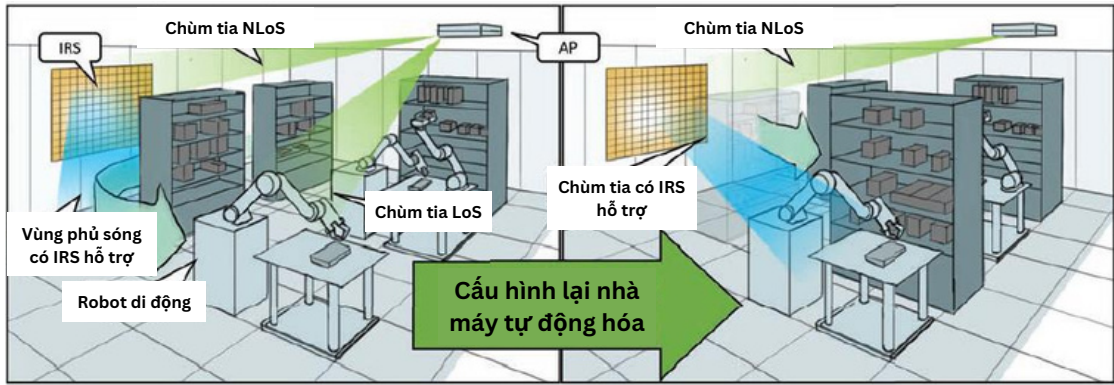
Double DQN với kiến trúc dueling cho robot di động trong một môi trường không biết. Môi trường mô phỏng được xây dựng trong Gazebo dựa trên ROS (Robot Operating System). Trong kịch bản mô phỏng, robot di động được trang bị một Lidar và một camera 3D để thu thập dữ liệu từ môi trường. Kết quả mô phỏng đã chứng minh hiệu quả của thuật toán đề xuất trong tình huống vật cản tĩnh.

### **Tích hợp truyền thông không dây cho robot di động**

Việc tích hợp truyền thông không dây cho robot di động đã được nhiều nhà khoa học quan tâm trong thời gian gần đây. Năm 2023, các tác giả [49] đã khảo sát trên 300 công trình nghiên cứu về hệ thống mạng robot rồi chỉ ra các ý tưởng, công nghệ, những thách thức, ứng dụng và hướng phát triển trong tương lai. Trong đó, các tác giả đã nói rằng sự phát triển của IoT đã mở đường cho một hướng đi mới của hệ thống robot mang tên IoRT (Internet of Robotic Things). Mới đây, trong năm 2024 này, các tác giả của [50] cũng đã nghiên cứu về việc lập quỹ đạo chuyển động cho robot trong môi trường công nghiệp sử dụng mạng 5G/B5G phục vụ truyền thông.

Trong [51], các tác giả nghiên cứu vấn đề tối ưu năng lượng tiêu thụ và tốc độ truyền thông của robot di động dựa trên hai mô hình: một là mô hình truyền thông position-critical (position-critical communication model) và hai là mô hình tốc độ truyền hằng số (constant bit-rate model). Về điều kiện truyền thông, các tác giả giả sử vùng phủ sóng của BS là vùng diện tích hình tròn với bán kính  $R$  tâm là BS. Kết hợp với việc đề xuất thuật toán xấp xỉ dựa trên Dijkstra, kết quả mô phỏng cho thấy mô hình 1 đã tiết kiệm được gần 50% mức năng lượng tiêu thụ, mô hình 2 tiết kiệm 22.18%. Ngoài ra, ở công trình [52], các tác giả quan tâm đến mô hình robot làm nhiệm vụ di chuyển qua một số điểm truyền dữ liệu trong một môi trường truyền thông không dây. Từ đó, các tác giả xây dựng một bài toán tối ưu quỹ đạo di chuyển của robot để tối thiểu mức năng lượng tiêu thụ của robot bao gồm cả năng lượng di chuyển và năng lượng dành cho việc truyền thông trong một khoảng thời gian cho trước. Bài toán đặt ra là một bài toán tuyến tính có biến là số nguyên MILP (Mixed Integer Linear Program). Các tác giả đã tìm quỹ đạo tối ưu bằng cách sử dụng xác suất để xử lý vấn đề ràng buộc liên quan đến truyền thông. Kết quả mô phỏng đã xác nhận tính đúng đắn của tính toán lý thuyết. Tuy nhiên, các mô hình trên vẫn chưa tính đến môi trường làm việc có vật cản.

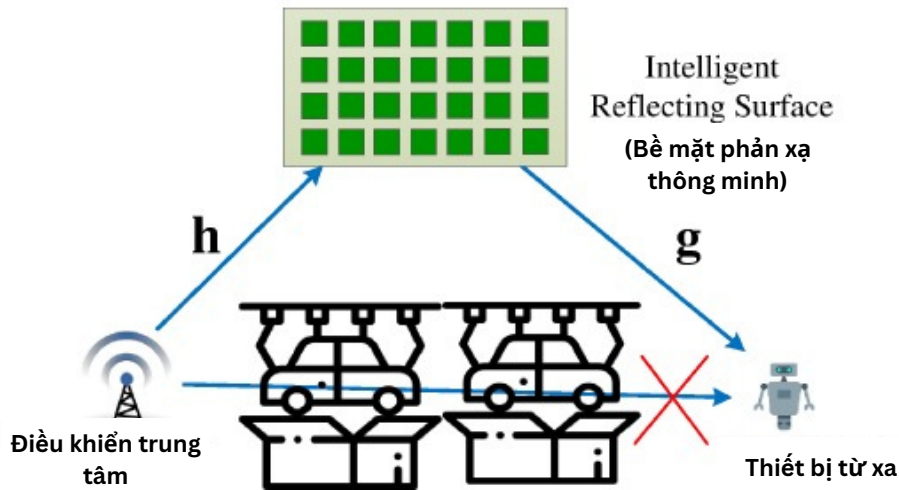
Trong môi trường có vật cản, một vấn đề là thách thức với hệ thống truyền



Hình 1.5: IRS giúp thay đổi vùng phủ sóng tương ứng với sự thay đổi layout của nhà máy. [4]

thông đó là vấn đề "vùng chết" tín hiệu. Một kỹ thuật gần đây được đề xuất sử dụng để giải quyết vấn đề này, đó là bề mặt phản xạ thông minh IRS (Intelligent Reflecting Surface). Các tác giả của [4] đã viết về tiến trình hướng đến nền công nghiệp 5.0 và vai trò của IRS trong các nhà máy. Theo các tác giả, nền công nghiệp 5.0 được hình dung bằng sự hợp tác chặt chẽ giữa con người và máy móc đòi hỏi việc truyền thông siêu đáng tin cậy và thời gian trễ cực thấp URLLC. Khi đó, IRS có tiềm năng đóng vai trò quan trọng trong việc triển khai URLLC không dây trong nhà máy. Cũng theo các tác giả, IRS được dự đoán sẽ là một yếu tố chính trong việc kích hoạt mạng truyền thông không dây 6G vì nó có thể cải thiện đáng kể hiệu suất của các mạng không dây bằng cách tạo ra một môi trường không dây có thể kiểm soát được. Ở hình 4 trong bài viết (hình 1.5 trong luận án này), các tác giả đã thể hiện một lợi ích vô cùng quan trọng của IRS, IRS giúp tự động thay đổi vùng phủ sóng thích nghi với vị trí mới khi nhà máy sắp xếp lại vị trí phục vụ các nhu cầu sản xuất khác nhau.

Bên cạnh đó, cũng đã có nhiều tác giả quan tâm đến môi trường truyền thông không dây sử dụng IRS cho robot. Cụ thể, các tác giả trong [5] xem xét một hệ thống truyền thông được hỗ trợ bởi IRS cho một kịch bản nhà máy tự động hóa, trong đó bộ điều khiển trung tâm như BS hoặc AP truyền thông tin đến một thiết bị từ xa như robot. Do mật độ các vật cản dày đặc giữa bên phát và bên thu, công suất tín hiệu truyền trực tiếp yếu. Để cải thiện chất lượng truyền thông, một IRS được triển khai giữa bộ phát và bộ thu, và một kênh truyền phản xạ (tín hiệu truyền từ AP đến IRS rồi phản xạ trở về thiết bị nhận) đã hỗ trợ truyền thông lúc này như trên hình 1.6. Mô phỏng được thực hiện với bảy trường hợp và cho thấy hiệu suất tăng khi số lượng phân tử của IRS tăng trong tất cả các trường hợp. Tuy nhiên, trong bài báo này, các

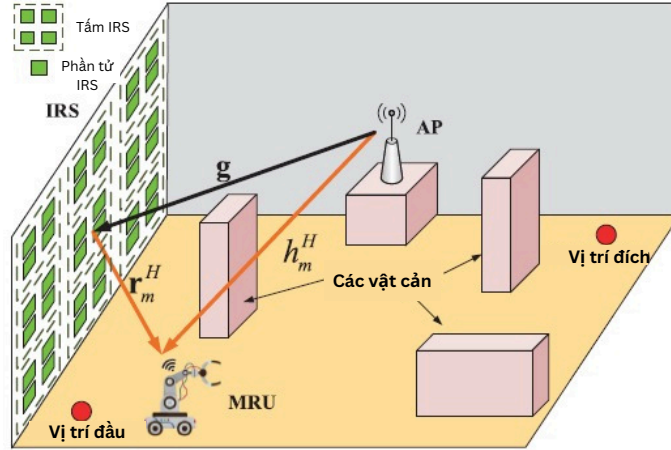


Hình 1.6: IRS giúp truyền tín hiệu đến thiết bị nhận khi kênh truyền trực tiếp bị chặn. [5]

tác giả không xem xét tính di động của các thiết bị từ xa như robot di động.

Trong [53], các tác giả quan tâm đến một robot di động làm việc trong nhà máy công nghiệp với một số chướng ngại vật tĩnh. Ở đây, các tác giả quan tâm kênh truyền từ robot di động đến AP được hỗ trợ bởi IRS. Sau đó, các tác giả xây dựng bài toán tối ưu với mục tiêu tối thiểu năng lượng di chuyển cho robot di động với ràng buộc về tổng thời gian di chuyển của robot và ngưỡng tốc độ truyền thông tối thiểu bằng cách tối ưu đồng thời quỹ đạo và bộ tạo chùm tia (beamformer). Sử dụng bản đồ vô tuyến và thông tin môi trường có sẵn, các tác giả đề xuất một phương pháp tối ưu liên tục để giải quyết vấn đề. Kết quả mô phỏng cho thấy IRS có thể cải thiện khả năng phủ sóng và hiệu quả năng lượng di chuyển của robot. Khác với [53], tác giả trong [54] xem xét lượng tiêu thụ năng lượng truyền của AP bằng cách tối ưu hóa đồng thời số lượng, vị trí và độ dịch pha (phase shift) của các IRS và bộ tạo chùm tia của AP. Bài toán tối ưu là bài toán không lồi và gặp khó khăn khi giải. Các tác giả đã chia bài toán gốc thành hai bài toán con để giải luân phiên. Điều này làm tăng hiệu quả tính toán và kết quả mô phỏng thể hiện được hiệu quả của thuật toán. Tuy nhiên, ở đây, các tác giả chưa quan tâm đến việc lập quỹ đạo cho robot di động.

Ngoài ra, các tác giả trong [6] đề xuất một hệ thống điều hướng cho robot trong nhà có IRS hỗ trợ. Hệ thống bao gồm một AP, một IRS và một robot di động làm việc trong môi trường nhà máy với nhiều chướng ngại vật tĩnh như trên hình 1 trong bài báo (hình 1.7 trong luận án này). Các tác giả đặt ra bài toán tối ưu với mục tiêu tối thiểu thời gian di chuyển của robot bằng cách tối ưu quỹ đạo của robot và ma trận hệ



Hình 1.7: Hệ thống điều hướng cho robot trong nhà máy có IRS hỗ trợ. [6]

số phản xạ của các phần tử của IRS, với ràng buộc về tốc độ truyền thông tối thiểu và vận tốc tối đa của robot. Bằng cách xây dựng bản đồ vô tuyến (radio), việc lập quỹ đạo cho robot sau đó được giải quyết bằng thuật toán Dijkstra. Kết quả mô phỏng cho thấy rằng vùng phủ sóng của AP được mở rộng đáng kể thông qua việc triển khai IRS và hệ thống có thể tìm thấy quỹ đạo tối ưu cho robot di động. Tuy nhiên, ở đây yêu cầu có sẵn bản đồ môi trường làm việc để xây dựng bản đồ vô tuyến. Đây là nghiên cứu đầu tiên đề xuất sử dụng IRS hỗ trợ hệ thống truyền thông không dây phục vụ cho robot di động ở nhà máy.

Gần đây, một khái niệm mới về tấm phản xạ thông minh ra đời, được gọi là bề mặt vừa truyền vừa phản xạ thông minh STAR-IRS (Simultaneously Transmitting and Reflecting IRS). Nhiều nghiên cứu đã chỉ ra hiệu quả của STAR-IRS trong việc nâng cao chất lượng truyền thông. Các tác giả [55] đã đề xuất sử dụng STAR-IRS giúp nâng cao chất lượng truyền thông cho hệ thống A2G (Air-to-Ground) gồm một UAV và các GUs (ground users). Tuy nhiên, nghiên cứu này đang quan tâm đến hệ UAV, chưa quan tâm đến hệ thống robot di động trong nhà máy.

#### 1.4 Định hướng nghiên cứu của luận án

Từ các phân tích ở phần 1.3 có thể nhận thấy rằng, bài toán lập quỹ đạo chuyển động cho robot đã được cộng đồng nghiên cứu và doanh nghiệp quan tâm. Do sự xuất hiện của ngày càng nhiều các thiết bị không dây cùng với nhu cầu quản lý và vận hành nhà máy, robot di động được kỳ vọng rằng sẽ hoạt động linh hoạt trong môi trường truyền thông tại nhà máy. Tuy nhiên, môi trường nhà máy ngày càng phức tạp do tồn

tại nhiều vùng chết tín hiệu và chưa đáp ứng được yêu cầu về độ tin cậy và độ trễ truyền thông. Để đảm bảo robot di động vừa hoàn thành nhiệm vụ vừa truyền thông tin cậy với hệ thống, việc giải bài toán tối ưu quỹ đạo cho robot di động trong môi trường truyền thông không dây tại nhà máy là cần thiết. Đây là bài toán bao gồm ba vấn đề: vấn đề lập quỹ đạo cho robot, vấn đề truyền thông cho robot và vấn đề tối ưu. Bài toán tối ưu cơ bản có dạng:

$$\begin{aligned} \mathbf{Q}^* = & \arg \min_{\mathbf{Q}} f_0(\mathbf{Q}) \\ \text{thỏa mãn: } & f_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (1.3)$$

trong đó  $f_0(\mathbf{Q})$  là hàm mục tiêu,  $\mathbf{Q}$  là quỹ đạo của robot là biến tối ưu, và  $f_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, m$  là các ràng buộc. Mục tiêu của bài toán có thể là tối thiểu quãng đường di chuyển, tối thiểu mức năng lượng tiêu thụ, tối thiểu thời gian di chuyển... Tương ứng với các mục tiêu khác nhau thì các biến tối ưu cũng có thể thay đổi, ngoài tối ưu quỹ đạo của robot có thể cần tối ưu công suất phát của BS/AP để thỏa mãn được các ràng buộc. Các ràng buộc ở đây có thể là ràng buộc về vị trí ban đầu, vị trí đích của robot, ràng buộc tránh vật cản hay ràng buộc về tốc độ truyền thông tối thiểu của robot.

Dựa trên các vấn đề từ bài toán tối ưu hệ thống phục vụ hoạt động của robot di động, trong luận án này, NCS sẽ nghiên cứu 03 vấn đề. Thứ nhất là nghiên cứu mô hình sử dụng IRS và thuật toán nâng cao chất lượng truyền thông không dây trong nhà máy phục vụ hoạt động của robot di động. Khi vấn đề về chất lượng truyền thông không dây đã được giải quyết, giả sử tại mọi điểm trong quá trình di chuyển của robot đều thỏa mãn ràng buộc tốc độ truyền thông, lúc này môi trường truyền thông được gọi là lý tưởng. Theo đó, vấn đề thứ hai là nghiên cứu đề xuất mô hình và các thuật toán lập quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng. Tuy nhiên, trong thực tế, rất khó để đạt được môi trường truyền thông lý tưởng. Do đó, vấn đề thứ ba là nghiên cứu hệ thống điều hướng cho robot trong môi trường truyền thông có STAR-IRS hỗ trợ. Nội dung của ba vấn đề này sẽ lần lượt trình bày ở các chương tiếp theo.

## 1.5 Kết luận chương 1

Chương 1 đã trình bày các vấn đề của việc lập quỹ đạo cho robot di động trong môi trường truyền thông không dây tại nhà máy. Qua nghiên cứu các công trình liên

quan trọng và ngoài nước, NCS đã đưa ra hướng nghiên cứu và các nội dung này sẽ được trình bày ở các chương tiếp theo của luận án. Cụ thể, chương 2 nghiên cứu mô hình sử dụng IRS và thuật toán nâng cao chất lượng truyền thông không dây trong nhà máy phục vụ hoạt động của robot di động. Theo đó, giả sử vấn đề về chất lượng truyền thông không dây đã được giải quyết, tức là mọi điểm trong quá trình di chuyển của robot đều thỏa mãn ràng buộc tốc độ truyền thông, lúc này môi trường truyền thông được gọi là lý tưởng. Tiếp theo, chương 3 nghiên cứu đề xuất thuật toán lập quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng. Tuy nhiên, trong thực tế, rất khó để đạt được môi trường truyền thông lý tưởng. Do đó, chương 4 nghiên cứu mô hình hệ thống và thuật toán điều hướng cho robot trong môi trường truyền thông có STAR-IRS hỗ trợ.

Các đóng góp của chương 1 được công bố ở công trình CT4 và CT5. Đây là những nghiên cứu ban đầu của NCS liên quan đến điều hướng cho robot. Công trình CT4 nghiên cứu cơ bản về định vị và điều hướng cho robot trong môi trường ROS (Robot Operating System). Công trình CT5 nghiên cứu ban đầu về vấn đề điều hướng cho robot trong ROS dựa trên nền học sâu tăng cường DRL (Deep Reinforcement Learning).

[CT4] **N. T. T. Van**, N. Manh Tien, N. M. Cuong, H. T. K. Duyen, B. T. T. Ha, and B. V. Tuan, "Building SLAM system and Intelligent Navigation for Autonomous Mobile Robot base on ROS," in Proceedings of the sixth Vietnam international conference and Exhibition on Control and Automation, VCCA-2021; ISBN 978-604-95-0875-2.

[CT5] **N. T. T. Van**, N. M. Tien, N. M. Cuong, H. T. K. Duyen, and N. D. Duy, "Constructing an Intelligent Navigation System for Autonomous Mobile Robot Based on Deep Reinforcement Learning". Cham: Springer International Publishing, 2021, pp. 251–261. [Online]. Available: [https://doi.org/10.1007/978-3-030-76620-7\\_22](https://doi.org/10.1007/978-3-030-76620-7_22).

## **Chương 2. NÂNG CAO CHẤT LƯỢNG TRUYỀN THÔNG TRONG NHÀ MÁY PHỤC VỤ HOẠT ĐỘNG CỦA ROBOT DI ĐỘNG**

Như đã trình bày ở chương 1, việc tối ưu hoạt động của robot di động trong môi trường truyền thông không dây tại nhà máy là cần thiết. Ở chương này, các vấn đề về truyền thông không dây và các kỹ thuật nâng cao chất lượng truyền thông không dây sẽ được trình bày. Cụ thể, phần 2.1 trình bày thực trạng của các mô hình truyền thông không dây từ đó đặt vấn đề về giải pháp tối ưu hiệu suất truyền thông không dây đặc biệt là trong các nhà máy. Mô hình hệ thống truyền thông có IRS hỗ trợ được đề xuất và bài toán tối ưu được trình bày ở phần 2.2. Phần 2.3, luận án trình bày thuật toán để giải bài toán tối ưu đặt ra ở phần 2.2. Phần 2.4, luận án đưa ra một kịch bản phân bổ thời gian trong hệ thống mạng có IRS hỗ trợ và trình bày thuật toán để giải bài toán tối ưu cho hệ thống mạng đó. Phần kết quả mô phỏng được trình bày và thảo luận ở phần 2.5, và cuối cùng là kết luận chương.

### **2.1 Vấn đề tối ưu hiệu suất truyền thông không dây tại các nhà máy**

Như đã trình bày ở chương 1, việc tối ưu hiệu suất truyền thông không dây trong nhà máy là một vấn đề cấp thiết với mục tiêu là đảm bảo yêu cầu về tốc độ thông tin và chất lượng dịch vụ QoS (Quality of Service) của thiết bị [56]. Trong môi trường nhà máy, ngoài robot di động còn có nhiều thiết bị tham gia kết nối mạng như các máy tính xách tay, máy tính bảng, điện thoại di động, các thiết bị IoT không dây như cảm biến, các thiết bị truyền bluetooth, và nhiều thiết bị truyền thông khác. Một vấn đề đặt ra là làm sao để nhiều thiết bị cùng sử dụng chung tài nguyên của một hệ thống mạng mà vẫn đảm bảo được các yêu cầu cho hoạt động của từng thiết bị. Đây là thách thức thứ nhất đối với hệ thống truyền thông không dây trong nhà máy. Hơn thế nữa, trong môi trường nhà máy, có nhiều thiết bị, máy móc và nhiều bức tường chắn... Điều này tạo ra nhiều vùng trong không gian nhà máy không nhận được tín hiệu truyền thông không dây, vùng này gọi là vùng chết (dead zone). Đây là thách thức thứ hai khi sử dụng truyền thông không dây trong nhà máy.

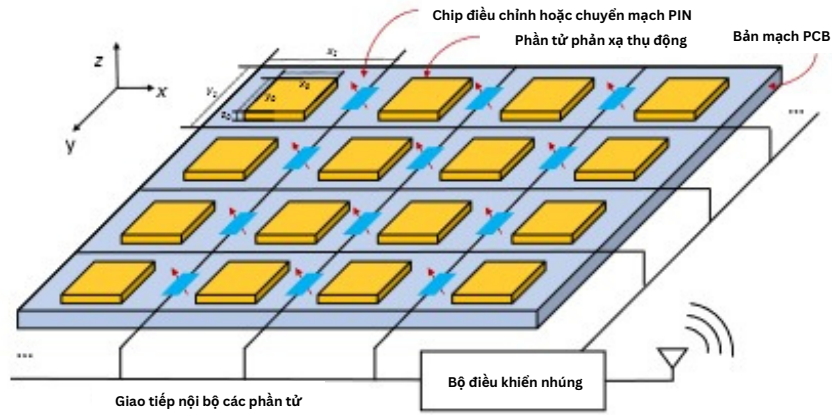
Để khắc phục vấn đề thứ nhất cần thực hiện các biện pháp tối ưu hiệu suất của hệ thống mạng truyền thông không dây trong nhà máy. Hiện nay, việc tối ưu hiệu suất của các mạng không dây tập trung vào phía thiết bị hoặc bộ điều khiển mạng như

trạm gốc BSs (Base Station), điểm truy cập AP (Access Point) và nhà mạng (network operator). Đối với các nhà mạng, để đáp ứng yêu cầu thông lượng ngày càng tăng, có thể triển khai nhiều cell nhỏ tiết kiệm năng lượng trong một hệ thống mạng hoặc sử dụng nhiều ăng-ten tại BS để tăng cường hiệu quả về tần số [57]. Bộ tạo chùm tia (bộ tạo chùm tia) của BS hoặc phân bổ công suất có thể được tối ưu để tương thích với sự thay đổi của kênh. Ở phía các thiết bị sử dụng mạng, nhiều loại thiết bị có thể tham gia vào cùng một hệ thống mạng. Tuy nhiên, việc tối ưu chỉ có thể trở nên khả thi khi việc trao đổi thông tin và phối hợp giữa các thiết bị sử dụng mạng và bộ điều khiển mạng được đảm bảo. Theo đó, nhiều công trình nghiên cứu đã đề xuất cách tối ưu hệ thống liên kết để cải thiện hiệu suất EE/SE của các mạng không dây bằng cách kết hợp các kỹ thuật khác nhau như ở [58]. Các phương pháp này có thể bao gồm tối ưu công suất truyền, hợp tác chuyển mạch (cooperative relaying), bộ tạo chùm tia và phân bổ tài nguyên,...

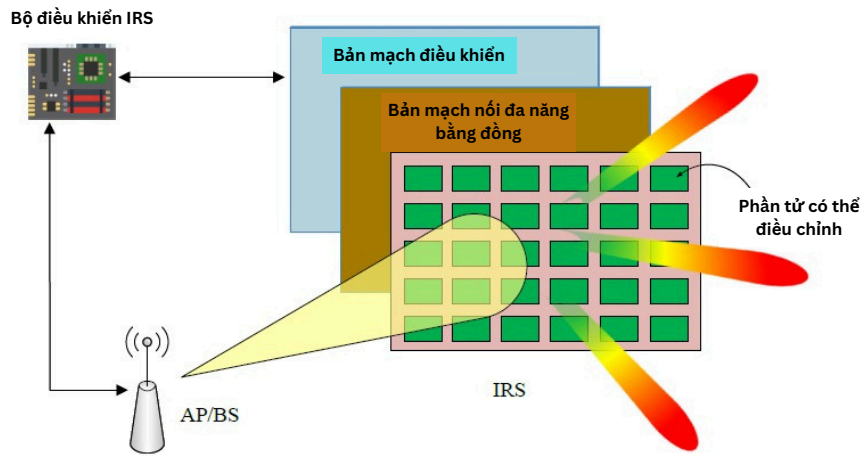
Tuy nhiên, các vấn đề tối ưu hệ thống mạng chỉ có thể khả thi khi tín hiệu truyền trong nhà máy luôn được đảm bảo, tức là cần phải khắc phục được thách thức thứ hai kể trên. Gần đây, một khái niệm mới về bề mặt phản xạ thông minh IRS (Intelligent Reflecting Surface) hay bề mặt cấu hình lại (Reconfigurable Intelligent Surface) đã được giới thiệu trong các nghiên cứu về truyền thông không dây [59] như là một giải pháp khắc phục vấn đề về vùng chết tín hiệu. IRS là một bề mặt hai chiều (2D) do con người tạo ra từ vật liệu điện từ EM (electromagnetic material), gọi là siêu bề mặt (metasurface), được tạo thành từ một mảng lớn các phần tử phản xạ thụ động với cấu trúc vật lý được thiết kế đặc biệt.

Như được mô tả trên hình 2.1, một IRS là một bề mặt nhân tạo mỏng bao gồm một số lượng lớn các phần tử phản xạ thụ động và chi phí thấp (ví dụ như các di ốt PIN), mỗi phần tử có thể phản xạ và tác động đến sự lan truyền của sóng điện từ truyền tới [60]. Kết quả là, IRSs có thể tạo ra một môi trường mạng không dây có thể lập trình được. Đối với môi trường nhà máy, nơi có nhiều vùng chết tín hiệu, tức là đường truyền tín hiệu trực tiếp từ thiết bị truyền đến thiết bị nhận bị chặn, IRS có thể được triển khai để cung cấp một đường truyền phản xạ để cải thiện cường độ tín hiệu nhận được. Hình 2.2 mô tả một hệ thống mạng không dây được hỗ trợ bởi IRS với nhiều ứng dụng khác nhau. Với vùng chết của dịch vụ mạng thông thường, khi triển khai IRS sẽ tạo ra kênh phản xạ hay còn gọi là kênh truyền trực tiếp LoS (line-of-sight) ảo giữa các thiết bị và BS hay AP thông qua IRS mặc dù có chướng ngại vật ở giữa.





(a)



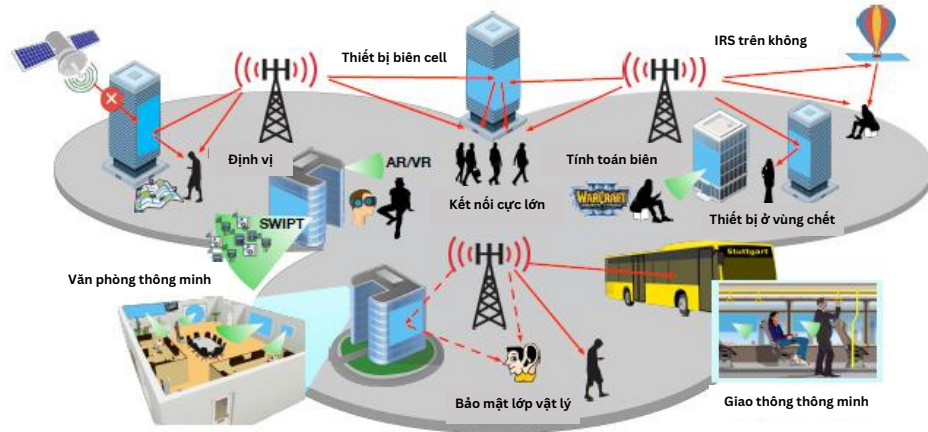
(b)

Hình 2.1: Cấu trúc IRS [7, 8].

Một IRS có thể phản xạ các tín hiệu truyền tới nó bằng cách điều chỉnh thông số về độ dịch pha để tạo ra một môi trường truyền thông có thể điều khiển được. Tín hiệu được phản xạ bởi IRS có thể được cộng với các tín hiệu LoS tại các bộ thu để tăng công suất tín hiệu nhận được, hoặc có thể được cộng lại để làm giảm nhiễu cùng kênh [61]. Ngoài ra IRS còn đặc biệt hữu ích trong việc mở rộng phạm vi phủ sóng của hệ thống mạng Terahertz hay mạng mmWave. Hơn nữa, việc triển khai IRS ở vùng biên của một cell không chỉ giúp cải thiện công suất tín hiệu truyền ở vùng biên mà còn hỗ trợ triệt tiêu nhiễu do tín hiệu ở cell bên cạnh truyền sang.

Những lợi ích mà IRS mang lại có thể kể ra như sau:

- Triển khai dễ dàng và hoạt động với độ ổn định cao.
- Cấu hình lại linh hoạt thông qua bộ tạo chùm tia thụ động.
- Nâng cao hiệu suất EE/SE.



Hình 2.2: Mô tả các ứng dụng của IRS với hệ thống mạng không dây trong tương lai [8].

- Mở ra những hướng nghiên cứu mới đầy hứa hẹn.

Do những nguyên nhân kể trên, IRS đã và đang nhận được sự quan tâm lớn từ cả các nhà nghiên cứu và doanh nghiệp. Tuy nhiên, một vấn đề của hệ thống sử dụng IRS là việc tối ưu các chùm tia phát và hệ số phản xạ cho IRS để mang lại tốc độ truyền cao do về mặt toán học, các tham số này đều là các số phức. Tối ưu một hệ thống có IRS hỗ trợ là một thách thức vì hai lý do sau: (i) cả tốc độ và công suất truyền đều là các hàm phức tạp với các biến bộ tạo chùm tia và hệ số phản xạ IRS; (ii) Hệ số phản xạ IRS có ràng buộc độ lớn đơn vị UMC (unit-modulus-constraint) là không lồi. Giải pháp tối ưu luân phiên giữa biến bộ tạo chùm tia và biến hệ số phản xạ IRS thường được sử dụng. Mỗi vòng tối ưu sẽ luân phiên giải hai bài toán con, một là tối ưu chùm tia với hệ số phản xạ giữ không đổi, hai là tối ưu hệ số phản xạ với chùm tia giữ cố định. Tuy nhiên, các bài toán con này đều là các bài toán không lồi và do đó vẫn còn thách thức lớn về mặt tính toán. Các tác giả ở [62] xây dựng bài toán tối ưu với ràng buộc hệ số phản xạ như một ma trận hạng 1, ràng buộc ma trận hạng 1 này sau đó được giải phóng (relax) để chuyển về một ràng buộc lồi. Ngoài ra, các tác giả [63] cũng đưa ra các giải pháp tối ưu xen kẽ và giải phóng ràng buộc hệ số phản xạ IRS để đạt hiệu quả tính toán. Tuy nhiên, hiệu quả tính toán vẫn không đạt được khi mà các bài toán vẫn là bài toán không lồi. Về mặt lý thuyết, kết quả tìm ra thậm chí có thể không phải là một điểm tối ưu cục bộ vì nó chỉ là điểm tối ưu trong một vùng cố định. Các tác giả ở [64] đã xây dựng vấn đề với hàm phạt một cách chính xác dựa trên việc kết hợp tính toán ràng buộc UMC vào hàm mục tiêu rồi xấp xỉ nó thành một bài toán lồi. Từ đó, các tác giả đưa ra một thuật toán với độ phức tạp tính toán

thấp với kết quả hội tụ tại một điểm tối ưu cục bộ.

Đã có một số công trình đã đề xuất sử dụng IRS để cải thiện thông lượng thông tin trong các mạng không dây. Một nghiên cứu khảo sát toàn diện về các công trình này được trình bày trong [65]. Gần đây, có một số công trình liên quan đến việc sử dụng IRS trong truyền thông D2D [61, 66]. Cụ thể, các tác giả trong [66] đã xây dựng bài toán tối đa hóa tổng tốc độ truyền thông tin của một cặp truyền nhận D2D và một thiết bị truyền thông tin trong hệ thống mạng hỗ trợ bởi IRS bằng cách tối ưu công suất phát tại bộ phát D2D và độ dịch pha (phase shift) của IRS. Để cải thiện cả thông lượng truyền thông tin và năng lượng thu cho các thiết bị, các nghiên cứu gần đây như [67, 68] đề xuất triển khai IRS trong các mạng tích hợp truyền thông tin và năng lượng. Đặc biệt, các tác giả trong [67] đề xuất sử dụng IRS cho hệ thống đồng thời truyền thông tin và năng lượng SWIPT (simultaneous wireless information and power transfer). Hệ thống SWIPT cho phép mỗi bộ thu chia tín hiệu nhận được thành hai thành phần, một phần dành cho thu năng lượng EH (energy harvesting) và một phần dành cho giải mã thông tin ID (information decoding) bằng cách sử dụng giao thức chia tỷ lệ công suất. Đối với phương thức này, hệ thống được gọi là SWIPT dựa trên PS (power-splitting). Ngoài ra, có một phương thức khác đó là dựa trên việc phân chia thời gian TS (time-splitting). Trong thực tế, việc tối ưu các mục tiêu đối lập của truyền thông tin và năng lượng đồng thời là khó khăn. Do đó, các thiết bị có thể được chia thành thiết bị truyền thông tin IUs (Information-demanded users) và thiết bị thu năng lượng EUs (Energy-demanded users), và sau đó các tác giả của [68] tối ưu đồng thời bộ tạo chùm tia cho IUs, bộ tạo chùm tia cho EUs, và độ dịch pha cho IRS. Trong [69], các tác giả đồng thời tối ưu bộ tạo chùm tia cho IUs và độ dịch pha cho IRS để tối đa hóa tổng tốc độ truyền của IUs đồng thời đáp ứng yêu cầu thu năng lượng của EUs. Trong khi đó, các tác giả trong [70] đề xuất tối ưu bộ tạo chùm tia cho IUs và EUs cũng như các độ dịch pha cho IRS để tối đa hóa tổng công suất nhận được tại EUs, dưới điều kiện ràng buộc SINR (Signal to Interference and Noise Ratio) của IUs. Các kết quả mô phỏng trong [70] cho thấy tổng công suất thu được bởi EUs cải thiện đáng kể so với trường hợp không có IRS. Tuy nhiên, [68] và [70] không xem xét tính công bằng giữa các thiết bị trong quá trình tối ưu. Trong [64], các tác giả xem xét một mạng với một BS phục vụ nhiều thiết bị có sự hỗ trợ của IRS. Trong mạng này, các tác giả xem xét bài toán tối ưu nhằm tối đa hóa tốc độ yếu nhất của thiết bị bằng cách tối ưu các hệ số phản xạ của IRS và bộ tạo chùm tia. Trong [71], các tác giả

xem xét các mạng tích hợp truyền thông tin và năng lượng, mỗi mạng bao gồm EUs và IUs. Các EUs thu năng lượng và các IUs nhận và giải mã thông tin được truyền bởi BS. Các tác giả xem xét một bài toán tối ưu nhằm tối đa hóa tốc độ truyền tại IU có tốc độ truyền yếu nhất bằng cách tối ưu đồng thời bộ tạo chùm tia của IUs và EUs, và phân bổ thời gian dành cho truyền thông tin và thu năng lượng.

## 2.2 Mô hình hệ thống

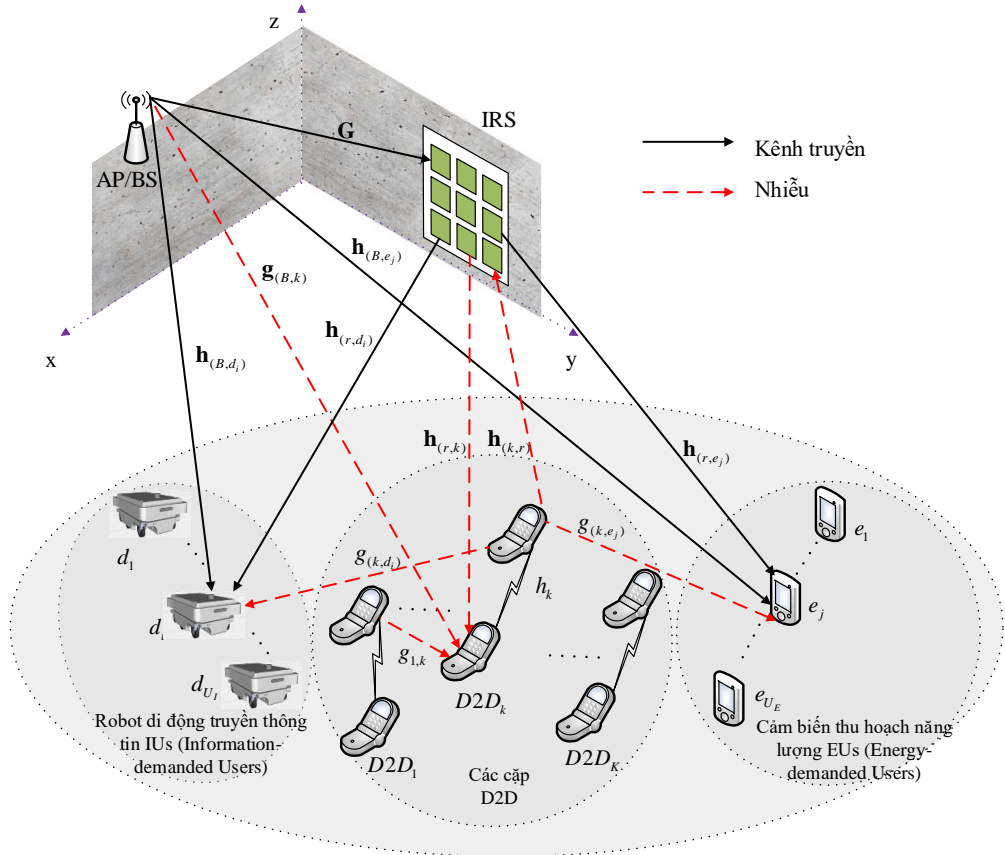
Như đã đề cập ở Chương 1, truyền thông và năng lượng là hai vấn đề của robot di động cần được quan tâm. Sự phát triển của các hệ thống mạng thế hệ mới như các hệ thống IoT đã làm gia tăng các thách thức về nhu cầu năng lượng và băng tần [72]. Để giải quyết vấn đề liên quan đến năng lượng, các hệ thống mạng tích hợp truyền thông tin và năng lượng gần đây đã được đề xuất, trong đó, một trạm gốc BS/điểm truy cập AP truyền thông tin cho các robot di động IUs và truyền năng lượng cho các cảm biến không dây EUs. Để giải quyết vấn đề liên quan đến băng tần và giảm năng lượng tiêu thụ cho hệ thống, truyền thông D2D (Device-to-Device) là một giải pháp hữu hiệu cho phép các thiết bị truyền D2D chia sẻ băng tần với mạng di động chính (cellular) [73]. Do đó, luận án đề xuất sử dụng mô hình mạng tích hợp truyền thông tin và năng lượng kết hợp với truyền thông D2D, gọi là hệ thống DED2D (the integrated data and energy network and D2D communication coexistence) để giải quyết các thách thức cho nhu cầu về truyền thông tin, thu năng lượng và băng tần ngày càng tăng trong hệ thống IoT trong nhà máy.

Tuy nhiên, hệ thống DED2D lại đặt ra một số thách thức. Vấn đề đầu tiên phát sinh từ sự hiện diện đồng thời của truyền thông D2D và mạng tích hợp truyền thông tin và năng lượng. Các bộ phát D2D chia sẻ cùng dải tần số với mạng chính, và do đó truyền thông D2D sẽ gây ra nhiễu đến các IUs và làm giảm đáng kể thông lượng thông tin truyền của các IUs [73]. Vấn đề thứ hai là các mất mát kênh truyền không dây, ví dụ như mất mát đường truyền và bóng mờ đa đường (fading) cũng làm giảm đáng kể thông lượng truyền thông tin của các IUs. Vấn đề thứ ba, các EUs thường yêu cầu một lượng năng lượng thu đáng kể, và năng lượng thu giảm đáng kể bởi khoảng cách từ BS/AP đến EUs. Để tăng thông lượng thông tin, một số công nghệ, ví dụ như hệ lớn với nhiều đầu vào nhiều đầu ra MIMO (massive multiple-input multiple-output), đã được đề xuất gần đây. Tuy nhiên, hệ này yêu cầu lượng năng lượng tiêu thụ cao và chi phí phần cứng lớn [64]. Vấn đề cuối cùng là vấn đề chung của tất cả các hệ thống

mạng chứ không của riêng hệ thống DED2D, đó là vấn đề liên quan đến vùng chết tín hiệu. Trong môi trường nhà máy luôn tồn tại các chướng ngại vật như các máy móc, thiết bị, hàng hóa... giữa thiết bị truyền và thiết bị nhận tín hiệu. Khi đó, các kênh truyền trực tiếp LoS bị chặn tạo ra các vùng chết tín hiệu mà tại đó các thiết bị thu không thể nhận được tín hiệu. Do đó, IRS là một giải pháp hứa hẹn để giải quyết các vấn đề đã đề cập, tức là vấn đề nhiễu cùng kênh, mất mát kênh truyền và các chướng ngại vật của hệ thống DED2D.

Trong luận án này, NCS quan tâm đến mạng DED2D có IRS hỗ trợ trong nhà máy như hình 2.3. BS/AP được trang bị  $\bar{M}$  ăng ten phục vụ một nhóm  $\mathcal{U}_I \triangleq \{(d_i), i = 1, \dots, U_I\}$  robot di động nhận thông tin IUs và một nhóm  $\mathcal{U}_E \triangleq \{(e_j), j = 1, \dots, U_E\}$  cảm biến không dây thu năng lượng EUs. Cả IUs và EUs đều được trang bị một ăng ten và hoạt động ở chế độ half-duplex. Có một nhóm  $\mathcal{K} = \{1, \dots, K\}$  cặp thiết bị D2D (điện thoại, máy tính, ...) chia sẻ cùng băng tần với mạng chính. IRS được trang bị với một bộ  $\mathcal{N} = \{1, \dots, N\}$  phần tử phản xạ.

Đặt  $\bar{\Theta} \in \mathbb{C}^{N \times N}$  là ma trận đường chéo hệ số phản xạ. Ta có  $\bar{\Theta} \triangleq \text{diag}[\bar{\theta}_1, \bar{\theta}_2, \dots, \bar{\theta}_N]$ , trong đó  $\bar{\theta}_n$  là hệ số phản xạ của phần tử  $n$  của IRS, thỏa mãn điều kiện ràng buộc



Hình 2.3: Mạng DED2D có IRS hỗ trợ.

sau:

$$|\bar{\theta}_n| = 1, \forall n \in \mathcal{N}. \quad (2.1)$$

Đặt  $\mathbf{G} \in \mathbb{C}^{N \times \bar{M}}$  là kênh truyền từ BS/AP đến IRS,  $\mathbf{h}_{(k,r)} \in \mathbb{C}^N$  là kênh từ bộ phát của cặp D2D thứ  $k$  đến IRS. Ngoài ra, đặt  $\mathbf{h}_{(B,d_i)} \in \mathbb{C}^{\bar{M}}$ ,  $\mathbf{h}_{(B,e_j)} \in \mathbb{C}^{\bar{M}}$ , và  $\mathbf{g}_{(B,k)} \in \mathbb{C}^{\bar{M}}$  lần lượt là các kênh trực tiếp từ BS đến IU  $d_i \in \mathcal{U}_I$ , EU  $e_j \in \mathcal{U}_E$ , và kênh từ BS đến bộ thu của cặp D2D thứ  $k \in \mathcal{K}$ . Thêm nữa, đặt  $\mathbf{h}_{(r,d_i)} \in \mathbb{C}^N$ ,  $\mathbf{h}_{(r,e_j)} \in \mathbb{C}^N$ , và  $\mathbf{h}_{(r,k)} \in \mathbb{C}^N$  lần lượt là các kênh từ IRS đến IU  $d_i \in \mathcal{U}_I$ , EU  $e_j \in \mathcal{U}_E$ , và bộ thu của cặp D2D thứ  $k \in \mathcal{K}$ .

Đối với truyền thông D2D, đặt  $h_k \in \mathbb{C}$  và  $g_{(l,k)} \in \mathbb{C}$  lần lượt là kênh truyền của cặp D2D thứ  $k$  và kênh từ bộ phát của cặp D2D thứ  $l$  đến bộ nhận của cặp D2D thứ  $k$ , đặt  $g_{(k,d_i)} \in \mathbb{C}$  và  $g_{(k,e_j)} \in \mathbb{C}$  lần lượt là kênh từ bộ phát từ cặp D2D thứ  $k$  đến IU  $d_i \in \mathcal{U}_I$  và EU  $e_j \in \mathcal{U}_E$ .

Để thực hiện việc truyền thông tin và truyền năng lượng của BS/AP, luận án áp dụng giao thức chuyển đổi thời gian (time-switching/time-splitting protocol) [74]. Đặt  $\bar{t}_i$  và  $\bar{t}_e$  lần lượt là khoảng thời gian để truyền thông tin và truyền năng lượng trong một khe thời gian, và ta có  $0 \leq \bar{t}_i + \bar{t}_e \leq 1$ . Lưu ý rằng, các bộ phát D2D truyền tín hiệu với các bộ thu của nó trong suốt khoảng thời gian  $\bar{t}_i$  và  $\bar{t}_e$  trong mỗi khe thời gian. Khi đó, tín hiệu nhận tại robot di động, EUs, và các bộ thu D2D được xác định như sau.

### 2.2.1 Tín hiệu nhận tại các robot di động (IUs)

Trong suốt khoảng thời gian  $\bar{t}_i$ , tín hiệu nhận tại IU  $d_i \in \mathcal{U}_I$  bao gồm i) tín hiệu truyền trực tiếp từ BS/AP, ii) tín hiệu truyền từ BS/AP đến IRS sau đó phản xạ lại, iii) tín hiệu truyền trực tiếp từ bộ phát của cặp D2D thứ  $k$  với  $k \in \mathcal{K}$ , và iv) tín hiệu truyền từ bộ phát của D2D đến IRS rồi phản xạ lại. Đặt  $y_{d_i}$  là tín hiệu nhận từ IU  $d_i$ , thì  $y_{d_i}$  được xác định bởi:

$$y_{d_i} = \sum_{l=1}^{U_I} \mathbf{h}_{(B,d_i)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{w}}_{d_l} s_{d_l} + \sum_{k=1}^K g_{(k,d_i)}(\bar{\boldsymbol{\theta}}) \sqrt{p_k} s_k + n_{d_i},$$

trong đó  $\bar{\boldsymbol{\theta}} \triangleq [\bar{\theta}_1, \dots, \bar{\theta}_N]$ ,  $\mathbf{h}_{(B,d_i)}(\bar{\boldsymbol{\theta}}) = \mathbf{h}_{(B,d_i)}^H + \mathbf{h}_{(r,d_i)}^H \bar{\boldsymbol{\Theta}} \mathbf{G}$ ,  $\mathbf{h}_{(B,d_i)}(\bar{\boldsymbol{\theta}}) \in \mathbb{C}^{1 \times \bar{M}}$ ,  $g_{(k,d_i)}(\bar{\boldsymbol{\theta}}) = g_{(k,d_i)} + \mathbf{h}_{(r,d_i)}^H \bar{\boldsymbol{\Theta}} \mathbf{h}_{(k,r)}$ ,  $\bar{\mathbf{w}}_{d_l} \in \mathbb{C}^{\bar{M}}$  và  $s_{d_l} \in \mathbb{C}$  lần lượt là bộ tạo chùm tia truyền thông tin và dữ liệu truyền đến IU  $d_l$ ,  $p_k$  và  $s_k$  lần lượt là công suất truyền của bộ phát của

cặp D2D thứ  $k$  và dữ liệu truyền đến bộ thu của cặp D2D thứ  $k$ , và  $n_{d_i} \sim \mathcal{CN}(0, 1)$  là nhiễu Gaussian (AWGN) tại IU  $d_i$ .

### 2.2.2 Tín hiệu nhận tại các cảm biến không dây thu năng lượng (EUs)

Trong suốt khoảng thời gian  $\bar{t}_e$ , tín hiệu nhận tại EU  $e_j \in \mathcal{U}_E$  bao gồm i) tín hiệu truyền trực tiếp tại BS/AP, ii) tín hiệu truyền từ BS/AP và phản xạ lại bởi IRS, iii) tín hiệu truyền trực tiếp từ bộ phát của cặp D2D thứ  $k$ , và iv) tín hiệu truyền từ bộ phát của D2D rồi phản xạ lại bởi IRS. Đặt  $y_{e_j}$  là tín hiệu nhận tại EU  $e_j$  thì  $y_{e_j}$  được xác định bởi

$$y_{e_j} = \sum_{l=1}^{U_E} \mathbf{h}_{(B,e_j)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{v}}_{e_l} s_{e_l} + \sum_{k=1}^K g_{(k,e_j)}(\bar{\boldsymbol{\theta}}) \sqrt{p_k} s_k^d + n_{e_j},$$

trong đó  $\mathbf{h}_{(B,e_j)}(\bar{\boldsymbol{\theta}}) = \mathbf{h}_{(B,e_j)}^H + \mathbf{h}_{(r,e_j)}^H \bar{\boldsymbol{\Theta}} \mathbf{G}$ ,  $\mathbf{h}_{(B,e_j)}(\bar{\boldsymbol{\theta}}) \in \mathbb{C}^{1 \times \bar{M}}$ ,  $g_{(k,e_j)}(\bar{\boldsymbol{\theta}}) = g_{(k,e_j)} + \mathbf{h}_{(r,e_j)}^H \bar{\boldsymbol{\Theta}} \mathbf{h}_{(k,r)}$ ,  $\bar{\mathbf{v}}_{e_l} \in \mathbb{C}^{\bar{M}}$  và  $s_{e_l} \in \mathbb{C}$  lần lượt là bộ tạo chùm tia truyền năng lượng và dữ liệu năng lượng truyền đến EU  $e_l$ , và  $n_{e_j} \sim \mathcal{CN}(0, 1)$  là nhiễu Gaussian tại EU  $e_j$ . ở đây, dữ liệu năng lượng không mang bất kỳ thông tin gì trong đó [75].

### 2.2.3 Tín hiệu nhận tại bộ thu D2D

Như đã đề cập trước đó, các cặp D2D làm việc trong suốt khoảng thời gian  $\bar{t}_i$  và  $\bar{t}_e$ . Cụ thể, trong suốt khoảng thời gian  $\bar{t}_i$ , tín hiệu tại bộ thu của cặp D2D thứ  $k \in \mathcal{K}$  bao gồm i) tín hiệu truyền trực tiếp từ bộ phát thứ  $k$ , ii) tín hiệu truyền từ bộ phát và phản xạ lại bởi IRS, iii) tín hiệu truyền từ bộ phát của cặp D2D thứ  $l, \forall l \in \mathcal{K} \setminus \{k\}$ , và iv) tín hiệu truyền từ BS/AP và phản xạ lại bởi IRS.

Tín hiệu tại bộ thu D2D trong suốt khoảng thời gian  $\bar{t}_e$  cũng tương tự như xét trong khoảng thời gian  $\bar{t}_i$ . Tuy nhiên, trong thời gian  $\bar{t}_e$ , BS/AP truyền năng lượng chứ không truyền dữ liệu. Đặt  $y_{\bar{t}_i,k}$  và  $y_{\bar{t}_e,k}$  là tín hiệu nhận từ cặp D2D thứ  $k$ . Khi đó,  $y_{\bar{t}_i,k}$  và  $y_{\bar{t}_e,k}$  được biểu diễn như sau:

$$y_{\bar{t}_i,k} = h_k(\bar{\boldsymbol{\theta}}) \sqrt{p_k} s_k + \sum_{l \in \mathcal{K} \setminus \{k\}} g_{l,k}(\bar{\boldsymbol{\theta}}) \sqrt{p_l} s_l + \sum_{i=1}^{U_I} \mathbf{g}_{(B,k)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{w}}_{d_i} s_{d_i} + n_k, \quad (2.2)$$

$$y_{\bar{t}_e,k} = h_k(\bar{\boldsymbol{\theta}}) \sqrt{p_k} s_k + \sum_{l \in \mathcal{K} \setminus \{k\}} g_{l,k}(\bar{\boldsymbol{\theta}}) \sqrt{p_l} s_l + \sum_{j=1}^{U_E} \mathbf{g}_{(B,k)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{v}}_{e_j} s_{e_j} + n_k, \quad (2.3)$$

trong đó,  $h_k(\bar{\boldsymbol{\theta}}) = h_k + \mathbf{h}_{(r,k)}^H \bar{\boldsymbol{\Theta}} \mathbf{h}_{(k,r)}$ ,  $g_{(l,k)}(\bar{\boldsymbol{\theta}}) = g_{(l,k)} + \mathbf{h}_{(r,k)}^H \bar{\boldsymbol{\Theta}} \mathbf{h}_{(l,r)}$ ,  $\mathbf{g}_{(B,k)}(\bar{\boldsymbol{\theta}}) = \mathbf{g}_{(B,k)}^H +$

$\mathbf{h}_{(r,k)}^H \bar{\boldsymbol{\Theta}} \mathbf{G}$ , và  $n_k \sim \mathcal{CN}(0, 1)$  là nhiễu Gaussian tại bộ thu của cặp D2D thứ  $k$ .

Đặt  $\bar{\mathbf{w}} \triangleq \{\bar{\mathbf{w}}_{d_i}, d_i \in \mathcal{U}_I\}$ ,  $\bar{\mathbf{v}} \triangleq \{\bar{\mathbf{v}}_{e_j}, e_j \in \mathcal{U}_E\}$ ,  $\bar{\mathbf{t}} \triangleq \{\bar{t}_i, \bar{t}_e\}$ , và  $\mathbf{p} \triangleq \{p_k, k \in \mathcal{K}\}$ .

Khi đó, ta sẽ xác định được thông lượng dữ liệu của robot di động, của cặp D2D và năng lượng thu được của EUs như sau.

Thông lượng dữ liệu của IU  $d_i$  (nats/s/Hz) trong suốt khoảng thời gian  $\bar{t}_i$  là

$$\bar{t}_i R_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}), \quad (2.4)$$

trong đó,  $R_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) = \ln \left( 1 + \frac{|\mathbf{h}_{(B, d_i)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{w}}_{d_i}|^2}{\psi_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}})} \right)$ ,

và  $\psi_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) = \sum_{l \neq i, d_l \in \mathcal{U}_I} |\mathbf{h}_{(B, d_l)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{w}}_{d_l}|^2 + \sum_{k=1}^K p_k |g_{(k, d_i)}(\bar{\boldsymbol{\theta}})|^2 + \sigma_{d_i}^2$ .

Đặt  $\rho \in [0, 1]$  là hiệu suất chuyển đổi năng lượng của EUs. Lượng năng lượng thu được của EU  $e_j$  trong khoảng thời gian  $\bar{t}_e$  là

$$\bar{t}_e \rho E_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}), \quad (2.5)$$

trong đó,  $E_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) = \sum_{l=1}^{U_E} |\mathbf{h}_{(B, e_j)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{v}}_{e_l}|^2 + \sum_{k=1}^K p_k |g_{(k, e_j)}(\bar{\boldsymbol{\theta}})|^2$ . Lưu ý, ở trong (2.5), năng lượng thu từ nhiễu nền bị bỏ qua do quá nhỏ.

Thông lượng dữ liệu đạt được bởi mỗi cặp D2D là tổng thông lượng đạt bởi mỗi cặp trong suốt khoảng thời gian  $\bar{t}_i$  và  $\bar{t}_e$ . Thông lượng của cặp D2D thứ  $k$  (nats/s/Hz) được xác định bởi

$$\bar{t}_i R_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) + \bar{t}_e R_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}), \quad (2.6)$$

trong đó,  $R_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) = \ln \left( 1 + \frac{p_k |h_k(\bar{\boldsymbol{\theta}})|^2}{\psi_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}})} \right)$  và  $R_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) = \ln \left( 1 + \frac{p_k |h_k(\bar{\boldsymbol{\theta}})|^2}{\psi_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}})} \right)$ .

Ở đây,  $\psi_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}})$  và  $\psi_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}})$  được định nghĩa bởi:

$$\begin{aligned} \psi_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) &= \sum_{l \in \mathcal{K} \setminus \{k\}} p_l |g_{l, k}(\bar{\boldsymbol{\theta}})|^2 + \sum_{d_i \in \mathcal{U}_I} |\mathbf{g}_{(B, k)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{w}}_{d_i}|^2 + \sigma_k^2, \\ \psi_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) &= \sum_{l \in \mathcal{K} \setminus \{k\}} p_l |g_{l, k}(\bar{\boldsymbol{\theta}})|^2 + \sum_{e_j \in \mathcal{U}_E} |\mathbf{g}_{(B, k)}(\bar{\boldsymbol{\theta}}) \bar{\mathbf{v}}_{e_j}|^2 + \sigma_k^2. \end{aligned} \quad (2.7)$$

## 2.2.4 Bài toán tối ưu

Trong mô hình DED2D, một BS/AP truyền thông tin đến nhiều robot di động và truyền năng lượng đến nhiều EUs thông qua hệ thống truyền thông tin và năng



lượng dựa trên thời gian (TFIET). Các bộ phát D2D có thể sử dụng cùng dải tần số như của mạng truyền thông tin và năng lượng. Một IRS được triển khai trong hệ thống với mục tiêu cải thiện thông lượng thông tin của các robot di động và truyền thông D2D cũng như quản lý nhiễu gây ra từ truyền thông D2D tới các robot di động. Do đó, luận án nghiên cứu một bài toán tối ưu kết hợp tối ưu i) bộ tạo chùm tia truyền thông tin cho robot di động, ii) bộ tạo chùm tia truyền năng lượng cho các cảm biến không dây, iii) tỷ lệ thời gian của hệ thống TFIET, iv) công suất phát của các bộ phát D2D, và v) hệ số phản xạ của IRS. Mục tiêu là tối đa thông lượng của robot di động có thông lượng thấp nhất trong khi đảm bảo yêu cầu năng lượng thu của các EUs và ngưỡng tốc độ truyền thông D2D. Do đó, với kịch bản không phân bổ thời gian cho truyền thông của mạng D2D (Non Orthogonal Time Allocation) hay gọi là kịch bản N-OTA, bài toán tối ưu được mô tả dưới dạng toán học như sau:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\mathbf{t}} = (\bar{t}_i, \bar{t}_e) \in \mathbb{R}_+^2, \bar{\boldsymbol{\theta}}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\mathbf{t}}, \bar{\boldsymbol{\theta}}) \triangleq \min_{d_i \in \mathcal{U}_I} \bar{t}_i R_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \quad (2.8a)$$

thoả mãn: (2.1),

$$\bar{t}_e \rho E_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.8b)$$

$$\bar{t}_i R_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) + \bar{t}_e R_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \geq R_{k, min}, \forall k \in \mathcal{K}, \quad (2.8c)$$

$$\bar{t}_i + \bar{t}_e \leq 1, \quad (2.8d)$$

$$\bar{t}_i \sum_{d_i \in \mathcal{U}_I} \|\bar{\mathbf{w}}_{d_i}\|^2 + \bar{t}_e \sum_{e_j \in \mathcal{U}_E} \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.8e)$$

$$\|\bar{\mathbf{w}}_{d_i}\|^2 \leq P_{B, max}; \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.8f)$$

$$p_k \leq P_{k, max}, \forall k \in \mathcal{K}, \quad (2.8g)$$

trong đó,  $e_{min}$  là ngưỡng năng lượng thu yêu cầu của EUs,  $R_{k, min}$  là ngưỡng tốc độ truyền yêu cầu của D2D,  $P_{k, max}$  và  $P_{B, max}$  lần lượt là công suất định mức của bộ phát của cặp D2D thứ  $k$  và BS/AP. Ràng buộc ở (2.1) là điều kiện mô đun đơn vị UMC của các phần tử IRS. Ràng buộc (2.8b) là mức năng lượng thu yêu cầu tối thiểu của EU, và ràng buộc (2.8c) là tốc độ truyền yêu cầu tối thiểu của các cặp D2D. Ràng buộc (2.8e), (2.8f), và (2.8g) là các ràng buộc về công suất định mức của BS/AP và các cặp D2D. Từ (2.8) có thể thấy rằng hàm mục tiêu (2.8a) và các ràng buộc (2.1), (2.8b), (2.8c) và (2.8e) là không lồi.

Đây là một bài toán phức tạp do cả hàm mục tiêu và các ràng buộc đều là các hàm phức tạp của các bộ tạo chùm tia và hệ số phản xạ của IRS. Đặc biệt, các hệ số

phản xạ của IRS bị ràng buộc bởi ràng buộc không lồi đơn vị. Như đã đề cập ở phần 2.1, việc giải bài toán tối ưu bộ tạo chùm tia và các hệ số phản xạ của IRS là một thách thức lớn. Các công trình gần đây, ví dụ như [64, 71], đã chứng minh hiệu quả của các thuật toán tối ưu luân phiên có thể giải quyết các bài toán tối ưu với độ phức tạp tính toán thấp và hội tụ ít nhất đến một nghiệm tối ưu cục bộ. Do đó, luận án phát triển một thuật toán tối ưu luân phiên để giải bài toán tối ưu (2.8). Chi tiết về thuật toán sẽ được trình bày ở phần 2.3.

### 2.3 Thuật toán tối ưu

Trước tiên, đặt  $\bar{\tau}_i = 1/\bar{t}_i$  và  $\bar{\tau}_e = 1/\bar{t}_e$ , bài toán (2.8) trở thành

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e) \in \mathbb{R}_+^2, \bar{\boldsymbol{\theta}}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\theta}}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i) R_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \quad (2.9a)$$

thỏa mãn: (2.1),

$$(1/\bar{\tau}_e) \rho E_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.9b)$$

$$(1/\bar{\tau}_i) R_{\bar{\tau}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) + (1/\bar{\tau}_e) R_{\bar{\tau}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) \quad (2.9c)$$

$$\geq R_{k, min}, \forall k \in \mathcal{K},$$

$$1/\bar{\tau}_i + 1/\bar{\tau}_e \leq 1, \quad (2.9d)$$

$$(1/\bar{\tau}_i) \sum_{d_i \in \mathcal{U}_I} \|\bar{\mathbf{w}}_{d_i}\|^2 + (1/\bar{\tau}_e) \sum_{e_j \in \mathcal{U}_E} \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.9e)$$

$$\|\bar{\mathbf{w}}_{d_i}\|^2 \leq P_{B, max}, \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.9f)$$

$$p_k \leq P_{k, max}, \forall k \in \mathcal{K}. \quad (2.9g)$$

Như đã đề cập trước đó, ràng buộc (2.1) là không lồi, do đó, trước khi giải bài toán tối ưu, luận án thực hiện giải phóng ràng buộc này. Cụ thể, ta thấy (2.1) tương đương với  $\sum_{n=1}^N |\bar{\theta}_n|^2 = N$ . Thật vậy, với  $\bar{\boldsymbol{\theta}} = \{(|\bar{\theta}_n| \in (0, 1)), \forall n \in \mathcal{N}\}$  thì  $|\bar{\theta}_n|^2 \leq |\bar{\theta}_n|, \forall n \in \mathcal{N}$  và  $|\bar{\theta}_n|^2 = |\bar{\theta}_n|$  khi ràng buộc (2.1) thỏa mãn. Với  $|\bar{\theta}_n|^2 \leq |\bar{\theta}_n|, \forall n \in \mathcal{N}$  thì  $\sum_{n=1}^N |\bar{\theta}_n|^2 \leq \sum_{n=1}^N |\bar{\theta}_n|$ . Do đó, ta có bất đẳng thức sau [76]

$$\frac{1}{\sum_{n=1}^N |\bar{\theta}_n|^2} \geq \frac{1}{\sum_{n=1}^N |\bar{\theta}_n|} = \frac{1}{N}. \quad (2.10)$$

Điều kiện đẳng thức ở (2.10) xảy ra khi ràng buộc (2.1) được thỏa mãn. Dựa trên

(2.10), luận án định nghĩa  $\Omega(\bar{\boldsymbol{\theta}})$  như sau [76]:

$$\Omega(\bar{\boldsymbol{\theta}}) = \frac{1}{N} - \frac{1}{\sum_{n=1}^N |\bar{\theta}_n|^2}. \quad (2.11)$$

$\Omega(\bar{\boldsymbol{\theta}})$  là căn cứ để đánh giá mức độ thỏa mãn của ràng buộc (2.12) và  $\Omega(\bar{\boldsymbol{\theta}}) \leq 0$  khi  $|\bar{\theta}_n| \leq 1$  và  $\Omega(\bar{\boldsymbol{\theta}}) = 0$  khi và chỉ khi  $\bar{\boldsymbol{\theta}} = \{(|\bar{\theta}_n| = 1), \forall n \in \mathcal{N}\}$ . Do đó, luận án giải phóng ràng buộc (2.1) trở thành [76]:

$$|\bar{\theta}_n|^2 \leq 1, \forall n \in \mathcal{N}. \quad (2.12)$$

Cùng với đó, luận án sử dụng  $\Omega(\bar{\boldsymbol{\theta}})$  ở (2.11) như một hàm phạt (penalty function) trong hàm mục tiêu (2.9a), và sử dụng ràng buộc (2.12) thay cho ràng buộc (2.1). Khi đó, bài toán (2.9) trở thành:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e) \in \mathbb{R}_+^2, \bar{\boldsymbol{\theta}}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\theta}}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i) R_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\boldsymbol{\theta}}) + \eta \Omega(\bar{\boldsymbol{\theta}}) \quad (2.13a)$$

$$\text{thỏa mãn: } (2.12), (2.9b) - (2.9g), \quad (2.13b)$$

trong đó,  $\eta$  là hệ số phạt.

Lúc này, bài toán (2.13) vẫn là bài toán không lồi bởi vì hàm mục tiêu (2.13a) và các ràng buộc (2.9b) và (2.9c) không lồi. Giải bài toán này vẫn là một thách thức, đặc biệt là việc tối ưu độ dịch pha. Trong khi đó, quan sát thấy rằng, nếu cố định  $(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e))$  hoặc độ dịch pha  $\bar{\boldsymbol{\theta}}$ , thì bài toán tối ưu (2.13) sẽ được giải dễ hơn. Do đó, luận án đề xuất sử dụng kỹ thuật tối ưu luân phiên [77] để giải bài toán (2.13). Cụ thể, luận án chia bài toán (2.13) thành hai bài toán con để giải luân phiên trong mỗi vòng lặp. Bài toán con đầu tiên gọi là bài toán 1 sẽ tối ưu  $(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e))$ , và bài toán con thứ hai gọi là bài toán 2 sẽ tối ưu  $\bar{\boldsymbol{\theta}}$ .

Đặt  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\boldsymbol{\theta}}^{(\kappa)})$  là một điểm khả thi của (2.13), điểm này được tìm thấy tại vòng lặp thứ  $(\kappa - 1)$ . Tại vòng lặp thứ  $\kappa$ , ta cố định  $\bar{\boldsymbol{\theta}} = \bar{\boldsymbol{\theta}}^{(\kappa)}$  rồi xác định  $\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}$ , và  $\bar{\tau}^{(\kappa+1)}$ , sau đó ta cố định  $\bar{\mathbf{w}} = \bar{w}^{(\kappa+1)}, \bar{\mathbf{v}} = \bar{v}^{(\kappa+1)}, \mathbf{p} = \bar{p}^{(\kappa+1)}$ , và  $\bar{\boldsymbol{\tau}} = \bar{\tau}^{(\kappa+1)}$  để tìm  $\bar{\boldsymbol{\theta}}^{(\kappa+1)}$  [78].

$$\begin{aligned}
& (1/\bar{\tau}_i)R_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \\
& \geq a_{\bar{\tau}_i,d_i}^{(\kappa)} + b_{\bar{\tau}_i,d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2}{|\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{w}}_{d_i}|^2} - \frac{\psi_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})}{\psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - c_{\bar{\tau}_i,d_i}^{(\kappa)} \bar{\tau}_i \quad (2.15) \\
& \geq a_{\bar{\tau}_i,d_i}^{(\kappa)} + b_{\bar{\tau}_i,d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2}{2\Re\{(\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{w}}_{d_i})(\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2} \right. \\
& \quad \left. - \frac{\psi_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})}{\psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - c_{\bar{\tau}_i,d_i}^{(\kappa)} \bar{\tau}_i \triangleq R_{\bar{\tau}_i,d_i}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}),
\end{aligned}$$


---

### 2.3.1 Bài toán 1 với kịch bản N-OTA

Trong bài toán con này, ta cố định  $\bar{\theta}^{(\kappa)}$ , khi đó ta có bài toán tối ưu sau:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau}=(\bar{\tau}_i, \bar{\tau}_e) \in \mathbb{R}_+^2} f_{sub-1}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i)R_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \quad (2.14a)$$

thoả mãn: (2.9d) – (2.9g),

$$(1/\bar{\tau}_e)\rho E_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.14b)$$

$$(1/\bar{\tau}_i)R_{\bar{\tau}_i,k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) + (1/\bar{\tau}_e)R_{\bar{\tau}_e,k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq R_{k,min}, \forall k \in \mathcal{K}. \quad (2.14c)$$

Để giải bài toán (2.14), ta sẽ thực hiện xấp xỉ hàm mục tiêu không lồi và các ràng buộc không lồi (2.14b) và (2.14c) thành dạng lồi. Đối với hàm mục tiêu, ta áp dụng bất đẳng thức (A.7) trong Phụ lục cho  $(1/\bar{\tau}_i)R_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$  với  $x = |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{w}}_{d_i}|^2$ ,  $y = \psi_{\bar{\tau}_i,d_i}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$ ,  $t = \bar{\tau}_i$ , và  $\bar{x} = |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2$ ,  $\bar{y} = \psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ ,  $\bar{t} = \bar{\tau}_i^{(\kappa)}$ . Khi đó, ta có bất đẳng thức (2.15) với

$$2\Re\{(\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{w}}_{d_i})(\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2 \geq 0, \quad (2.16)$$

và

$$\begin{aligned}
0 < a_{\bar{\tau}_i,d_i}^{(\kappa)} &= \frac{2}{\bar{\tau}_i^{(\kappa)}} \ln\left(1 + \frac{|\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2}{\psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\
0 < b_{\bar{\tau}_i,d_i}^{(\kappa)} &= \frac{|\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2 / \psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{\bar{\tau}_i^{(\kappa)} (1 + |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2 / \psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}, \\
0 < c_{\bar{\tau}_i,d_i}^{(\kappa)} &= \frac{\ln(1 + |\mathbf{h}_{(B,d_i)}(\bar{\theta}^{(\kappa)})\bar{w}_{d_i}^{(\kappa)}|^2 / \psi_{\bar{\tau}_i,d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}{(\bar{\tau}_i^{(\kappa)})^2}.
\end{aligned}$$

Hàm  $R_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$  lúc này đã trở thành hàm lõm. Tiếp theo, ta quan tâm đến ràng buộc (2.14b). Bằng cách sử dụng xấp xỉ Taylor bậc 1, ta có bất đẳng thức sau:

$$x^2 \geq 2x\bar{x} - \bar{x}^2, \forall x > 0, \bar{x} > 0. \quad (2.17)$$

Khi đó, ràng buộc không lồi (2.14b) có thể được xấp xỉ bởi ràng buộc lồi sau:

$$\begin{aligned} E_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) &\geq \sum_{l=1}^{U_E} [2\Re\{(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{v}}_{e_l})(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{v}}_{e_l}^*)\} \\ &\quad - |\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)})\bar{\mathbf{v}}_{e_l}^{(\kappa)}|^2] + \sum_{k=1}^K p_k |g_{(k, e_j)}(\bar{\theta}^{(\kappa)})|^2 \\ &\triangleq E_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq e_{\min} \bar{\tau}_e / \rho. \end{aligned} \quad (2.18)$$

Với ràng buộc không lồi (2.14c), ta áp dụng bất đẳng thức (A.7) trong Phụ lục cho  $(1/\bar{\tau}_i)R_{\bar{\tau}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$  với  $x = p_k |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $y = \psi_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$ ,  $t = \bar{\tau}_i$  và  $\bar{x} = \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ ,  $\bar{t} = \bar{\tau}_i^{(\kappa)}$ . Lúc này, ta có

$$\begin{aligned} (1/\bar{\tau}_i)R_{\bar{\tau}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) &\geq a_{\bar{\tau}_i, k}^{(\kappa)} + b_{\bar{\tau}_i, k}^{(\kappa)} \left( 2 - \frac{\bar{p}_k^{(\kappa)}}{p_k} - \frac{\psi_{\bar{t}_i, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)})}{\psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - c_{\bar{\tau}_i, k}^{(\kappa)} \bar{\tau}_i \\ &\triangleq R_{\bar{\tau}_i, k}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}), \end{aligned} \quad (2.19)$$

trong đó

$$\begin{aligned} 0 < a_{\bar{\tau}_i, k}^{(\kappa)} &= \frac{2}{\bar{\tau}_i^{(\kappa)}} \ln \left( 1 + \frac{\bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2}{\psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right), \\ 0 < b_{\bar{\tau}_i, k}^{(\kappa)} &= \frac{\bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{\bar{\tau}_i^{(\kappa)} (1 + \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}, \\ 0 < c_{\bar{\tau}_i, k}^{(\kappa)} &= \frac{\ln(1 + \bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}{(\bar{\tau}_i^{(\kappa)})^2}. \end{aligned}$$

Tương tự như vậy, áp dụng bất đẳng thức (A.7) trong Phụ lục cho  $(1/\bar{\tau}_e)R_{\bar{\tau}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$  với  $x = p_k |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $y = \psi_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)})$ ,  $t = \bar{\tau}_e$  và  $\bar{x} = \bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ ,  $\bar{t} = \bar{\tau}_e^{(\kappa)}$ . Khi đó, ta có

$$\begin{aligned} (1/\bar{\tau}_e)R_{\bar{\tau}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) &\geq a_{\bar{\tau}_e, k}^{(\kappa)} + b_{\bar{\tau}_e, k}^{(\kappa)} \left( 2 - \frac{\bar{p}_k^{(\kappa)}}{p_k} - \frac{\psi_{\bar{t}_e, k}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)})}{\psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - c_{\bar{\tau}_e, k}^{(\kappa)} \bar{\tau}_e \\ &\triangleq R_{\bar{\tau}_e, k}^{(\kappa)}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}), \end{aligned} \quad (2.20)$$

trong đó

$$\begin{aligned}
0 < a_{\bar{\tau}_e, k}^{(\kappa)} &= \frac{2}{\bar{\tau}_e^{(\kappa)}} \ln\left(1 + \frac{\bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2}{\psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\
0 < b_{\bar{\tau}_e, k}^{(\kappa)} &= \frac{\bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{\bar{\tau}_e^{(\kappa)} (1 + \bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}), \\
0 < c_{\bar{\tau}_e, k}^{(\kappa)} &= \frac{\ln(1 + \bar{p}_k^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}{(\bar{\tau}_e^{(\kappa)})^2}.
\end{aligned}$$

Dựa trên (2.19) và (2.20), ràng buộc không lồi (2.14c) được xấp xỉ bởi ràng buộc lồi sau:

$$R_{\bar{\tau}_i, k}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) + R_{\bar{\tau}_e, k}^{(\kappa)}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq R_{k, \min}, \forall k \in \mathcal{K}. \quad (2.21)$$

Từ (2.15), (2.18) và (2.21), bài toán 1 được xấp xỉ bởi bài toán tối ưu lồi sau:

$$\begin{aligned}
& \max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e) \in \mathbb{R}_+^2} f_{sub-1}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\theta}^{(\kappa)}) \\
& \triangleq \min_{d_i \in \mathcal{U}_I} R_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) + \eta \Omega(\bar{\theta}^{(\kappa)}) \quad (2.22a)
\end{aligned}$$

thoả mãn: (2.9d) – (2.9g), (2.16), (2.18) và (2.21).

Hàm  $f_{sub-1}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\theta}^{(\kappa)})$  là hàm lồi vì số hạng đầu tiên (2.22a) là hàm lồi (hàm nhỏ nhất của một hàm lồi là một hàm lồi [79]) và số hạng thứ hai cũng là hàm lồi. Do đó, thay vì giải bài toán (2.14), ta giải bài toán tối ưu lồi (2.22), và tìm ra  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  ở vòng lặp tiếp theo. Độ phức tạp tính toán của thuật toán để giải bài toán lồi (2.22) là [80]

$$\mathcal{O}(\alpha^2 \beta^{2.5} + \beta^{3.5}), \quad (2.23)$$

trong đó,  $\alpha = 2(\bar{M} + 1) + K$  là số lượng biến quyết định, và  $\beta = U_I + 2(U_E + K + 1)$  là số lượng ràng buộc. Lưu ý rằng  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)})$  là điểm khả thi của (2.22), nghĩa là  $f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}) = f_{sub-1}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)})$ . Trong khi đó,  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  là điểm tối ưu của (2.22), và do đó, ta có

$$f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) > f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}),$$

với  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}) \neq (\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)})$ . Lại quan tâm đến (2.15), ta có:

$$\begin{aligned} f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}) &= f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}) \\ &< f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) \\ &\leq f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}). \end{aligned} \quad (2.24)$$

Điều này cho thấy điểm tối ưu  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  của (2.22) thỏa mãn điều kiện hội tụ:  $f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) > f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ .

### 2.3.2 Bài toán 2 với kịch bản N-OTA

Trong bài toán này,  $\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}$  được cố định và tối ưu  $\bar{\theta}$ . Do đó, ta có bài toán sau:

$$\max_{\bar{\theta}} f_{sub-2}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i^{(\kappa)}) R_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) + \eta \Omega(\bar{\theta}) \quad (2.25a)$$

thỏa mãn: (2.12),

$$(1/\bar{\tau}_e^{(\kappa)}) \rho E_{\bar{\tau}_e, e_j}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.25b)$$

$$\begin{aligned} (1/\bar{\tau}_i^{(\kappa)}) R_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) + (1/\bar{\tau}_e^{(\kappa)}) R_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \\ \geq R_{k, min}, \forall k \in \mathcal{K}. \end{aligned} \quad (2.25c)$$

Tương tự như phần 2.3.1, để giải bài toán (2.25), ta xấp xỉ hàm mục tiêu (2.25a), và các ràng buộc không lồi (2.25b) và (2.25c) về dạng lồi. Để làm điều này, đầu tiên ta viết lại  $\mathbf{h}_{(B, d_i)}(\bar{\theta})$ ,  $\mathbf{h}_{(B, e_j)}(\bar{\theta})$ ,  $g_{(k, d_i)}(\bar{\theta})$ ,  $g_{(k, e_j)}(\bar{\theta})$ ,  $h_k(\bar{\theta})$ ,  $g_{l, k}(\bar{\theta})$  và  $\mathbf{g}_{(B, k)}(\bar{\theta})$  as  $\mathbf{h}_{(B, d_i)}(\bar{\theta}) = \mathbf{h}_{(B, d_i)}^H + \bar{\theta} \text{diag}(\mathbf{h}_{(r, d_i)}^H) \mathbf{G}$ ,  $\mathbf{h}_{(B, e_j)}(\bar{\theta}) = \mathbf{h}_{(B, e_j)}^H + \bar{\theta} \text{diag}(\mathbf{h}_{(r, e_j)}^H) \mathbf{G}$ ,  $g_{(k, d_i)}(\bar{\theta}) = g_{(k, d_i)} + \bar{\theta} \text{diag}(\mathbf{h}_{(r, d_i)}^H) \mathbf{h}_{k, r}$ ,  $g_{(k, e_j)}(\bar{\theta}) = g_{(k, e_j)} + \bar{\theta} \text{diag}(\mathbf{h}_{(r, e_j)}^H) \mathbf{h}_{k, r}$ ,  $h_k(\bar{\theta}) = h_k + \bar{\theta} \text{diag}(\mathbf{h}_{(r, k)}^H) \mathbf{h}_{k, r}$ ,  $g_{l, k}(\bar{\theta}) = g_{l, k} + \bar{\theta} \text{diag}(\mathbf{h}_{(r, k)}^H) \mathbf{h}_{l, r}$  và  $\mathbf{g}_{(B, k)}(\bar{\theta}) = \mathbf{g}_{(B, k)}^H + \bar{\theta} \text{diag}(\mathbf{h}_{(r, k)}^H) \mathbf{G}$ . Sau đó, ta xấp xỉ hàm mục tiêu và các ràng buộc không lồi của bài toán (2.25) như sau.

Áp dụng bất đẳng thức (A.2) trong Phụ lục cho  $R_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$  trong (2.25a) với  $x = |\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)}|^2$ ,  $y = \psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$ ,  $\bar{x} = |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2$ ,  $\bar{y} = \psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ . Khi đó, ta thu được bất đẳng thức (2.26) với

$$\begin{aligned}
& R_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \\
& \geq a_{\bar{\tau}_i, d_i}^{(\kappa)} + b_{\bar{\tau}_i, d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{|\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)}|^2} - \frac{\psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \\
& \geq a_{\bar{\tau}_i, d_i}^{(\kappa)} + b_{\bar{\tau}_i, d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2} \right. \\
& \quad \left. - \frac{\psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \triangleq R_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \tag{2.26}
\end{aligned}$$

$$2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 \geq 0, \tag{2.27}$$

trong đó

$$\begin{aligned}
0 < a_{\bar{\tau}_i, d_i}^{(\kappa)} &= \ln\left(1 + \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{\psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\
0 < b_{\bar{\tau}_i, d_i}^{(\kappa)} &= \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{1 + |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \psi_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}.
\end{aligned}$$

Hơn nữa, sử dụng bất đẳng thức (2.17), ta có

$$\Omega(\bar{\theta}) \geq \frac{1}{N} - \frac{1}{\sum_{n=1}^N (2\Re\{(\bar{\theta}_n^{(\kappa)})^* \bar{\theta}_n\} - |\bar{\theta}_n^{(\kappa)}|^2)} \triangleq \Omega^{(\kappa)}(\bar{\theta}) \tag{2.28}$$

với

$$\sum_{n=1}^N (2\Re\{(\bar{\theta}_n^{(\kappa)})^* \bar{\theta}_n\} - |\bar{\theta}_n^{(\kappa)}|^2) \geq 0. \tag{2.29}$$

Từ (2.26) và (2.28), ta có

$$\begin{aligned}
& f_{sub-2}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \geq f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \\
& \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i^{(\kappa)}) R_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) + \eta \Omega^{(\kappa)}(\bar{\theta}) \tag{2.30}
\end{aligned}$$

trong đó,  $f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta})$  đã là hàm lõm do (2.30) là hàm nhỏ nhất của một hàm lõm sẽ là một hàm lõm [79], và thành phần còn lại của hàm cũng đã là một hàm lõm.

Với ràng buộc (2.25b), bằng cách sử dụng bất đẳng thức (2.17), ràng buộc không



$$\begin{aligned}
R_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) &\geq a_{\bar{\tau}_i, k}^{(\kappa)} + b_{\bar{\tau}_i, k}^{(\kappa)} \left( 2 - \frac{|h_k(\bar{\theta}^{(\kappa)})|^2}{|h_k(\bar{\theta})|^2} - \frac{\psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \\
&\geq a_{\bar{\tau}_i, k}^{(\kappa)} + b_{\bar{\tau}_i, k}^{(\kappa)} \left( 2 - \frac{|h_k(\bar{\theta}^{(\kappa)})|^2}{2\Re\{h_k(\bar{\theta})h_k(\bar{\theta}^{(\kappa)})^*\} - |h_k(\bar{\theta}^{(\kappa)})|^2} \right. \\
&\quad \left. - \frac{\psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \triangleq R_{\bar{\tau}_i, k}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \tag{2.32}
\end{aligned}$$

lỗi này được xấp xỉ bởi ràng buộc lỗi sau:

$$\begin{aligned}
&E_{\bar{\tau}_e, e_j}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \\
&\geq \sum_{l=1}^{U_E} \left[ 2\Re\{(\mathbf{h}_{(B, e_j)}(\bar{\theta})\bar{v}_{e_l}^{(\kappa)})(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)})\bar{v}_{e_l}^{(\kappa)})^*\} - |\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)})\bar{v}_{e_l}^{(\kappa)}|^2 \right] \\
&+ \sum_{k=1}^K \bar{p}^{(\kappa)} \left[ 2\Re\{g_{(k, e_j)}(\bar{\theta})g_{(k, e_j)}(\bar{\theta}^{(\kappa)})^*\} - |g_{(k, e_j)}(\bar{\theta}^{(\kappa)})|^2 \right] \\
&\triangleq E_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \geq e_{\min} \bar{\tau}_e^{(\kappa)} / \rho. \tag{2.31}
\end{aligned}$$

Để xấp xỉ ràng buộc (2.25c) thành một ràng buộc lỗi, áp dụng bất đẳng thức (A.2) trong Phụ lục cho  $R_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$  trong (2.25c) với  $x = \bar{p}^{(\kappa)}|h_k(\bar{\theta})|^2$ ,  $y = \psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$ , và  $\bar{x} = \bar{p}^{(\kappa)}|h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ . Khi đó, ta thu được bất đẳng thức (2.32) với

$$2\Re\{h_k(\bar{\theta})h_k(\bar{\theta}^{(\kappa)})^*\} - |h_k(\bar{\theta}^{(\kappa)})|^2 \geq 0, \tag{2.33}$$

trong đó

$$\begin{aligned}
0 < a_{\bar{\tau}_i, k}^{(\kappa)} &= \ln\left(1 + \frac{\bar{p}^{(\kappa)}|h_k(\bar{\theta}^{(\kappa)})|^2}{\psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\
0 < b_{\bar{\tau}_i, k}^{(\kappa)} &= \frac{\bar{p}^{(\kappa)}|h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{1 + \bar{p}^{(\kappa)}|h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{\tau}_i, k}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}.
\end{aligned}$$

Tương tự với  $R_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$  trong (2.25c) với  $x = \bar{p}^{(\kappa)}|h_k(\bar{\theta})|^2$ ,  $y = \psi_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})$ , và  $\bar{x} = \bar{p}^{(\kappa)}|h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \psi_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ . Khi đó, ta có

$$\begin{aligned}
R_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) &\geq a_{\bar{\tau}_e, k}^{(\kappa)} + b_{\bar{\tau}_e, k}^{(\kappa)} \left( 2 - \frac{|h_k(\bar{\theta}^{(\kappa)})|^2}{|h_k(\bar{\theta})|^2} - \frac{\psi_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{\tau}_e, k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \\
&\geq a_{\bar{\tau}_e, k}^{(\kappa)} + b_{\bar{\tau}_e, k}^{(\kappa)} \left( 2 - \frac{|h_k(\bar{\theta}^{(\kappa)})|^2}{2\Re\{h_k(\bar{\theta})h_k(\bar{\theta}^{(\kappa)})^*\} - |h_k(\bar{\theta}^{(\kappa)})|^2} \right)
\end{aligned}$$

$$-\frac{\psi_{\bar{t}_e,k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta})}{\psi_{\bar{t}_e,k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \triangleq R_{\bar{\tau}_e,k}^{(\kappa)}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}), \quad (2.34)$$

trong đó

$$\begin{aligned} 0 < a_{\bar{\tau}_e,k}^{(\kappa)} &= \ln\left(1 + \frac{\bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2}{\psi_{\bar{t}_e,k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\ 0 < b_{\bar{\tau}_e,k}^{(\kappa)} &= \frac{\bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_e,k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{1 + \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \psi_{\bar{t}_e,k}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}. \end{aligned}$$

Dựa trên (2.32) và (2.34), ràng buộc không lồi (2.25c) được xấp xỉ bởi ràng buộc lồi sau:

$$(1/\bar{\tau}_i^{(\kappa)})R_{\bar{\tau}_i,k}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) + (1/\bar{\tau}_e^{(\kappa)})R_{\bar{\tau}_e,k}^{(\kappa)}(\bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}) \geq R_{k,min}, \forall k \in \mathcal{K}. \quad (2.35)$$

Từ (2.30), (2.31), và (2.35), bài toán không lồi (2.25) được xấp xỉ thành bài toán lồi sau:

$$\max_{\bar{\theta}} f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \quad (2.36)$$

thoả mãn: (2.12), (2.27), (2.29), (2.31), (2.33), và (2.35).

Nghiệm của (2.36) sẽ được sử dụng là điểm khả thi  $\bar{\theta}^{(\kappa+1)}$  ở vòng lặp tiếp theo. Độ phức tạp tính toán của thuật toán để giải bài toán tối ưu lồi (2.36) là  $\mathcal{O}(\alpha^2 \beta^{2.5} + \beta^{3.5})$ , trong đó  $\alpha = N$  và  $\beta = N + U_E + K$ . Tương tự như bài toán 1, ta có

$$\begin{aligned} f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) &= f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) \\ &< f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) \\ &\leq f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}), \quad (2.37) \end{aligned}$$

khi  $\bar{\theta}^{(\kappa+1)} \neq \bar{\theta}^{(\kappa)}$ .

Thuật toán tổng quát cho bài toán tối ưu thông lượng truyền dữ liệu (2.9) được trình bày ở Thuật toán 1. Thuật toán này được đặt tên là Thuật toán N-OTA, nghĩa là thuật toán với kịch bản không phân chia thời gian truyền thông của mạng D2D với mạng chính (non orthogonal time allocation). Để tăng cường hiệu quả tính toán của thuật toán, việc tìm điểm khả thi ban đầu là một việc quan trọng. Hơn nữa, cần xác định hệ số phạt  $\eta$ .

### 2.3.3 Tính hội tụ, điểm khả thi và lựa chọn hệ số phạt $\eta$

Từ (2.24) và (2.37) ta có

$$f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) > f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}). \quad (2.38)$$

Tức là điều kiện hội tụ được thỏa mãn và điểm  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})$  sẽ hội tụ về điểm tối ưu  $(\bar{\mathbf{w}}^*, \bar{\mathbf{v}}^*, \mathbf{p}^*, \tau^*, \bar{\boldsymbol{\theta}}^*)$ .

Cố định  $\tau^{(0)} = (\bar{\tau}_i^{(0)}, \bar{\tau}_e^{(0)})$ ,  $\bar{\theta}^{(0)}$  thỏa mãn các ràng buộc (2.1) và (2.9d), rồi lấy ngẫu nhiên các điểm  $(\bar{w}^{(0)}, \bar{v}^{(0)}, p^{(0)})$  khả thi cho (2.9d)-(2.9g). Sau đó, ta giải bài toán tối ưu

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}} \mu \quad (2.39a)$$

thỏa mãn: (2.9f) – (2.9g),

$$(1/\bar{\tau}_i^{(0)})R_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq \mu, \forall d_i \in \mathcal{U}_I, \quad (2.39b)$$

$$E_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq e_{\min} \bar{\tau}_e^{(0)} \mu / \rho, \forall e_j \in \mathcal{U}_E, \quad (2.39c)$$

$$R_{\bar{\tau}_i, k}^{(\kappa)}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) + R_{\bar{\tau}_e, k}^{(\kappa)}(\bar{\mathbf{v}}, \mathbf{p}, \bar{\theta}^{(\kappa)}) \geq R_{k, \min} \mu, \forall k \in \mathcal{K}, \quad (2.39d)$$

$$(1/\bar{\tau}_i^{(0)}) \sum_{d_i \in \mathcal{U}_I} \|\bar{\mathbf{w}}_{d_i}\|^2 + (1/\bar{\tau}_e^{(0)}) \sum_{e_j \in \mathcal{U}_E} \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, \max}, \quad (2.39e)$$

cho đến khi  $\mu \geq 1$ . Giải bài toán (2.39) để tạo ra một điểm  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \tau^{(0)} = (\bar{\tau}_i^{(0)}, \bar{\tau}_e^{(0)}), \bar{\theta}^{(0)})$  là điểm khả thi cho bài toán (2.9). Lúc này, ta có thể lựa chọn hệ số phạt  $\eta$  như sau [81]

$$\eta = - \left[ \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i^{(0)})R_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\theta}^{(0)}) \right] / \Omega(\bar{\theta}^{(0)}) \quad (2.40)$$

## 2.4 Kịch bản phân chia thời gian

Trong phần 2.3, luận án quan tâm đến hệ thống DED2D có IRS hỗ trợ, trong đó, các bộ phát D2D truyền tín hiệu trong suốt khoảng thời gian BS/AP truyền thông tin và truyền năng lượng. Kịch bản N-OTA giúp tăng hiệu quả sử dụng băng tần nhưng đồng thời cũng làm tăng nhiễu lên các phần tử của mạng. Trong phần này, luận án quan

---

**Thuật toán 1** Thuật toán N-OTA giải bài toán (2.9)
 

---

**Khởi tạo giá trị ban đầu:** Chọn giá trị bất kỳ  $\tau^{(0)} = (\bar{\tau}_i^{(0)}, \bar{\tau}_e^{(0)})$  và  $\bar{\theta}^{(0)}$  thỏa mãn các ràng buộc (2.1) và (2.9d), sau đó giải (2.39) để nhận được điểm khả thi  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \tau^{(0)} = (\bar{\tau}_i^{(0)}, \bar{\tau}_e^{(0)}), \bar{\theta}^{(0)})$  cho (2.9). Đặt  $\epsilon = 10^{-3}$  và  $\kappa = 0$ ;

1: Tính  $\eta$  theo (2.40);

2: **repeat**

3: Giải bài toán (2.22) với  $\bar{\theta} = \bar{\theta}^{(\kappa)}$  để thu được  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$ ;

4: Giải bài toán (2.36) với  $(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau}) = (\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  để thu được  $\bar{\theta}^{(\kappa+1)}$ ;

5:  $\hat{\epsilon} = \frac{|f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) - f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})|}{f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})}$

6:  $\kappa \leftarrow \kappa + 1$ ;

7: **until**  $\hat{\epsilon} \leq \epsilon$ .

8: Thu được nghiệm  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)} = (\bar{\tau}_i^{(\kappa)}, \bar{\tau}_e^{(\kappa)}), \bar{\theta}^{(\kappa)})$

---

tâm đến hệ thống DED2D có IRS hỗ trợ và có phân chia thời gian OTA (orthogonal time allocation) giữa thời gian truyền thông tin, truyền năng lượng và truyền thông D2D. Cụ thể, trong một khe thời gian, BS/AP sẽ chia thành các khoảng thời gian riêng cho các việc truyền thông tin, truyền năng lượng, và thời gian dành cho truyền thông D2D. Khi đó, sẽ không có nhiễu từ truyền thông D2D đến mạng chính nhưng thời gian dành cho việc truyền thông tin và năng lượng từ BS/AP sẽ bị giảm xuống. Do đó, việc thảo luận về hiệu quả của hai kịch bản N-OTA và OTA ở phần kết quả mô phỏng sẽ cho thấy một số điểm thú vị.

Trước tiên, luận án đặt một biến mới là  $\bar{t}_d$  để chỉ khoảng thời gian dành cho truyền thông D2D. Khi đó, thông lượng dữ liệu tại robot di động, năng lượng thu tại EUs, và thông lượng dữ liệu tại các cặp D2D được xác định như sau.

Thông lượng dữ liệu tại IU  $d_i$  (nats/s/Hz) trong khoảng thời gian  $\bar{t}_i$  là  $\bar{t}_i \tilde{R}_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta})$ , trong đó,  $R_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}) = \ln(1 + \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{\mathbf{w}}_{d_i}|^2}{\tilde{\psi}_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta})})$  với  $\tilde{\psi}_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}) = \sum_{l \neq i, d_l \in \mathcal{U}_I} |\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{\mathbf{w}}_{d_l}|^2 + \sigma_{d_i}^2$ . Năng lượng thu trong khoảng thời gian  $\bar{t}_e$  của EU  $e_j$  là  $\bar{t}_e \rho \tilde{E}_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \bar{\theta})$ , trong đó  $\tilde{E}_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \bar{\theta}) = \sum_{l=1}^{U_E} |\mathbf{h}_{(B, e_j)}(\bar{\theta}) \bar{\mathbf{v}}_{e_l}|^2$ . Thông lượng dữ liệu của cặp D2D  $k$  (nats/s/Hz) là  $\bar{t}_d \tilde{R}_{\bar{t}_d, k}(\mathbf{p}, \bar{\theta})$ , với  $\tilde{R}_{\bar{t}_d, k}(\mathbf{p}, \bar{\theta}) = \ln(1 + \frac{p_k |h_k(\bar{\theta})|^2}{\tilde{\psi}_{\bar{t}_d, k}(\bar{\mathbf{w}}, \mathbf{p}, \bar{\theta})})$ ,  $\tilde{\psi}_{\bar{t}_d, k}(\mathbf{p}, \bar{\theta}) = \sum_{l \in \mathcal{K} \setminus \{k\}} p_l |g_{l, k}(\bar{\theta})|^2 + \sigma_k^2$ .

Bài toán tối ưu đặt ra ở kịch bản OTA tương tự như (2.8), trong đó, biến  $\bar{t}_d$  thể hiện khoảng thời gian truyền thông D2D được tích hợp trong  $\bar{\mathbf{t}}$ , tức là,  $\bar{\mathbf{t}} = (\bar{t}_i, \bar{t}_e, \bar{t}_d)$ . Về mặt toán học, bài toán tối ưu được mô tả như sau:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\mathbf{t}} = (\bar{t}_i, \bar{t}_e, \bar{t}_d) \in \mathbb{R}_+^3, \bar{\theta}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\mathbf{t}}, \bar{\theta}) \triangleq \min_{d_i \in \mathcal{U}_I} \bar{t}_i \tilde{R}_{\bar{t}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}) \quad (2.41a)$$

thỏa mãn: (2.1),

$$\bar{t}_e \rho \tilde{E}_{\bar{t}_e, e_j}(\bar{\mathbf{v}}, \bar{\boldsymbol{\theta}}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.41b)$$

$$\bar{t}_d \tilde{R}_{\bar{t}_d, k}(\mathbf{p}, \bar{\boldsymbol{\theta}}) \geq R_{k, min}, \forall k \in \mathcal{K}, \quad (2.41c)$$

$$\bar{t}_i + \bar{t}_e + \bar{t}_d \leq 1, \quad (2.41d)$$

$$\bar{t}_i \sum_{d_i \in \mathcal{U}_I} \|\bar{\mathbf{w}}_{d_i}\|^2 + \bar{t}_e \sum_{e_j \in \mathcal{U}_E} \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}(1 - \bar{t}_d), \quad (2.41e)$$

$$\|\bar{\mathbf{w}}_{d_i}\|^2 \leq P_{B, max}, \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.41f)$$

$$\bar{t}_d p_k \leq P_{k, max}, \forall k \in \mathcal{K}. \quad (2.41g)$$

Để giải bài toán không lồi (2.41), đặt biến  $\bar{\tau}_i = 1/\bar{t}_i$ ,  $\bar{\tau}_e = 1/\bar{t}_e$ , và  $\bar{\tau}_d = 1/\bar{t}_d$ . Khi đó, bài toán (2.41) trở thành

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e, \bar{\tau}_d) \in \mathbb{R}_+^3, \bar{\boldsymbol{\theta}}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\theta}}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\boldsymbol{\theta}}) \quad (2.42a)$$

thoả mãn: (2.1),

$$(1/\bar{\tau}_e) \rho \tilde{E}_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \bar{\boldsymbol{\theta}}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.42b)$$

$$(1/\bar{\tau}_d) \tilde{R}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\boldsymbol{\theta}}) \geq R_{k, min}, \forall k \in \mathcal{K}, \quad (2.42c)$$

$$1/\bar{\tau}_i + 1/\bar{\tau}_e + 1/\bar{\tau}_d \leq 1, \quad (2.42d)$$

$$(1/\bar{\tau}_i) \sum_{d_i \in \mathcal{U}_I} \|\bar{\mathbf{w}}_{d_i}\|^2 + (1/\bar{\tau}_e) \sum_{e_j \in \mathcal{U}_E} \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}(1 - 1/\bar{\tau}_d), \quad (2.42e)$$

$$\|\bar{\mathbf{w}}_{d_i}\|^2 \leq P_{B, max}, \|\bar{\mathbf{v}}_{e_j}\|^2 \leq P_{B, max}, \quad (2.42f)$$

$$p_k \leq P_{k, max} \bar{\tau}_d, \forall k \in \mathcal{K}. \quad (2.42g)$$

Tương tự như phân tích ở phần trên, luận án cũng đưa hàm phạt  $\Omega(\bar{\boldsymbol{\theta}})$  vào trong bài toán tối ưu (2.42) khi giải phóng ràng buộc (2.1) và bài toán (2.42) tương đương với bài toán sau:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e, \bar{\tau}_d) \in \mathbb{R}_+^3, \bar{\boldsymbol{\theta}}} f(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\boldsymbol{\theta}}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\boldsymbol{\theta}}) + \eta \Omega(\bar{\boldsymbol{\theta}}) \quad (2.43a)$$

$$\text{thoả mãn: } (2.12), (2.42b) - (2.42g). \quad (2.43b)$$

Tương tự như phần 2.3, luận án chia bài toán (2.43) thành hai bài toán con để giải luân phiên trong mỗi vòng lặp. Bài toán 1 tối ưu  $\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e, \bar{\tau}_d)$ , và bài toán 2 tối ưu  $\bar{\boldsymbol{\theta}}$ .

$$\begin{aligned}
& (1/\bar{\tau}_i) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)}) \\
& \geq \tilde{a}_{\bar{\tau}_i, d_i}^{(\kappa)} + \tilde{b}_{\bar{\tau}_i, d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{\mathbf{w}}_{d_i}|^2} - \frac{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)})}{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - \tilde{c}_{\bar{\tau}_i, d_i}^{(\kappa)} \bar{\tau}_i \\
& \geq \tilde{a}_{\bar{\tau}_i, d_i}^{(\kappa)} + \tilde{b}_{\bar{\tau}_i, d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{\mathbf{w}}_{d_i})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2} \right. \\
& \quad \left. - \frac{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)})}{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - \tilde{c}_{\bar{\tau}_i, d_i}^{(\kappa)} \bar{\tau}_i \triangleq \tilde{R}_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)}) \tag{2.45}
\end{aligned}$$

#### 2.4.1 Bài toán 1 với kịch bản OTA

Ở đây,  $\bar{\theta}^{(\kappa)}$  được cố định và giải bài toán tối ưu sau:

$$\begin{aligned}
& \max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau} = (\bar{\tau}_i, \bar{\tau}_e, \bar{\tau}_d) \in \mathbb{R}_+^3} f_{sub-1}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau}, \bar{\theta}^{(\kappa)}) \\
& \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)}) + \eta \Omega(\bar{\theta}^{(\kappa)}) \tag{2.44a}
\end{aligned}$$

thoả mãn: (2.42d) – (2.42g),

$$(1/\bar{\tau}_e) \rho \tilde{E}_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \bar{\theta}^{(\kappa)}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \tag{2.44b}$$

$$(1/\bar{\tau}_d) \tilde{R}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\theta}^{(\kappa)}) \geq R_{k, min}, \forall k \in \mathcal{K}. \tag{2.44c}$$

Giả sử  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)})$  là điểm khả thi của (2.44) được tìm thấy ở vòng lặp  $(\kappa - 1)$ . Để giải bài toán (2.44), ta xấp xỉ hàm mục tiêu và các ràng buộc không lồi (2.44b) và (2.44c) thành dạng lồi như sau. Áp dụng bất đẳng thức (A.7) ở Phụ lục cho  $(1/\bar{\tau}_i) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)})$  trong hàm mục tiêu với  $x = |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{\mathbf{w}}_{d_i}|^2$ ,  $y = \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)})$ ,  $t = \bar{\tau}_i$  và  $\bar{x} = |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2$ ,  $\bar{y} = \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ ,  $\bar{t} = \bar{\tau}_i^{(\kappa)}$ . Khi đó, ta thu được bất đẳng thức (2.45) với

$$2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{\mathbf{w}}_{d_i})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 \geq 0, \tag{2.46}$$

trong đó,

$$0 < \tilde{a}_{\bar{\tau}_i, d_i}^{(\kappa)} = \frac{2}{\bar{\tau}_i^{(\kappa)}} \ln\left(1 + \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right),$$

$$\begin{aligned}
0 < \tilde{b}_{\bar{\tau}_i, d_i}^{(\kappa)} &= \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{\bar{\tau}_i^{(\kappa)} (1 + |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}, \\
0 < \tilde{c}_{\bar{\tau}_i, d_i}^{(\kappa)} &= \frac{\ln(1 + |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}{(\bar{\tau}_i^{(\kappa)})^2}.
\end{aligned}$$

Trong khi đó, ràng buộc (2.44b) có thể được xấp xỉ bởi

$$\begin{aligned}
\tilde{E}_{\bar{\tau}_e, e_j}(\bar{\mathbf{v}}, \bar{\theta}^{(\kappa)}) &\geq \sum_{l=1}^{U_E} \left[ 2\Re\{(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)}) \bar{\mathbf{v}}_{e_l})(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)}) \bar{v}_{e_l}^{(\kappa)})^*\} - |\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)}) \bar{v}_{e_l}^{(\kappa)}|^2 \right] \\
&\triangleq \tilde{E}_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{\mathbf{v}}, \bar{\theta}^{(\kappa)}) \geq e_{\min} \bar{\tau}_e / \rho.
\end{aligned} \tag{2.47}$$

Tương tự như vậy, áp dụng bất đẳng thức (A.7) ở Phụ lục cho  $(1/\bar{\tau}_d) \tilde{R}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\theta}^{(\kappa)})$  trong (2.44c) với  $x = p_k |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $y = \tilde{\psi}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\theta}^{(\kappa)})$ ,  $t = \bar{\tau}_d$ ,  $\bar{x} = \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ ,  $\bar{t} = \bar{\tau}_d^{(\kappa)}$ . Khi đó, ta có:

$$\begin{aligned}
(1/\bar{\tau}_d) \tilde{R}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\theta}^{(\kappa)}) &\geq \tilde{a}_{\bar{\tau}_d, k}^{(\kappa)} + \tilde{b}_{\bar{\tau}_d, k}^{(\kappa)} \left( 2 - \frac{\bar{p}_k^{(\kappa)}}{p_k} - \frac{\tilde{\psi}_{\bar{\tau}_d, k}(\mathbf{p}, \bar{\theta}^{(\kappa)})}{\tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) - \tilde{c}_{\bar{\tau}_d, k}^{(\kappa)} \bar{\tau}_d \\
&\triangleq \tilde{R}_{\bar{\tau}_d, k}^{(\kappa)}(\mathbf{p}, \bar{\theta}^{(\kappa)}),
\end{aligned} \tag{2.48}$$

trong đó,

$$\begin{aligned}
0 < \tilde{a}_{\bar{\tau}_d, k}^{(\kappa)} &= \frac{2}{\bar{\tau}_d^{(\kappa)}} \ln\left(1 + \frac{\bar{p}_k^{(\kappa)} |h_k|^2}{\tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right), \\
0 < \tilde{b}_{\bar{\tau}_d, k}^{(\kappa)} &= \frac{\bar{p}_k^{(\kappa)} |h_k|^2 / \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{\bar{\tau}_d^{(\kappa)} (1 + \bar{p}^{(\kappa)} |h_k|^2 / \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}, \\
0 < \tilde{c}_{\bar{\tau}_d, k}^{(\kappa)} &= \frac{\ln(1 + \bar{p}_k^{(\kappa)} |h_k|^2 / \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}{(\bar{\tau}_d^{(\kappa)})^2}.
\end{aligned}$$

Lúc này, bài toán 1 được xấp xỉ thành bài toán tối ưu lồi sau:

$$\max_{\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}} = (\bar{\tau}_i, \bar{\tau}_e, \bar{\tau}_d) \in \mathbb{R}_+^3} f_{\text{sub-1}}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\boldsymbol{\tau}}, \bar{\theta}^{(\kappa)}) \triangleq \min_{d_i \in \mathcal{U}_I} \tilde{R}_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\theta}^{(\kappa)}) + \eta \Omega(\bar{\theta}^{(\kappa)}) \tag{2.49a}$$

thoả mãn: (2.42d) – (2.42g), (2.46),

$$\tilde{E}_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{\mathbf{v}}, \bar{\theta}^{(\kappa)}) \geq e_{\min} \bar{\tau}_e / \rho, \forall e_j \in \mathcal{U}_E, \tag{2.49b}$$

$$\tilde{R}_{\bar{\tau}_d, k}^{(\kappa)}(\mathbf{p}, \bar{\theta}^{(\kappa)}) \geq R_{k, \min}, \forall k \in \mathcal{K}. \tag{2.49c}$$

Từ (2.49) có thể thấy rằng  $f_{sub-1}^{(\kappa)}(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \mathbf{p}, \bar{\tau}, \bar{\theta}^{(\kappa)})$  là hàm lõm. Do vậy, thay vì giải bài toán (2.22), ta đi giải bài toán xấp xỉ lồi (2.49), rồi nghiệm của nó được sử dụng làm điểm khả thi  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  tại vòng lặp sau. Độ phức tạp tính toán của thuật toán giải bài toán (2.49) được xác định bởi (2.23), trong đó,  $\alpha = 2\bar{M} + K + 3$  và  $\beta = U_I + 2(U_E + K + 1)$ . Ta có:

$$\begin{aligned} f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}) &= f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}) \\ &< f_{sub-1}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) \\ &\leq f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}). \end{aligned} \quad (2.50)$$

Điều này cho thấy điểm tối ưu  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  của (2.22) thỏa mãn điều kiện hội tụ:  $f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) > f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ .

#### 2.4.2 Bài toán 2 với kịch bản OTA

Bài toán 2 được định nghĩa như sau:

$$\max_{\bar{\theta}} f_{sub-2}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_i^{(\kappa)}) \tilde{R}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}) + \eta \Omega(\bar{\theta}) \quad (2.51a)$$

thỏa mãn: (2.12),

$$(1/\bar{\tau}_e^{(\kappa)}) \rho \tilde{E}_{\bar{\tau}_e, e_j}(\bar{v}^{(\kappa)}, \bar{\theta}) \geq e_{min}, \forall e_j \in \mathcal{U}_E, \quad (2.51b)$$

$$(1/\bar{\tau}_d^{(\kappa)}) \tilde{R}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}) \geq R_{k, min}, \forall k \in \mathcal{K}. \quad (2.51c)$$

Tương tự như phần 2.4.1, để giải (2.51), ta xấp xỉ hàm mục tiêu và các ràng buộc trong (2.51b) và (2.51c) như sau. Áp dụng bất đẳng thức (A.2) trong Phụ lục cho  $\tilde{R}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta})$  trong hàm mục tiêu với  $x = |\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)}|^2$ ,  $y = \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta})$ ,  $\bar{x} = |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2$ ,  $\bar{y} = \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ . Khi đó, ta thu được bất đẳng thức 2.52 với

$$2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 \geq 0,$$

trong đó,

$$0 < \tilde{a}_{\bar{\tau}_i, d_i}^{(\kappa)} = \ln \left( 1 + \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right),$$



$$\begin{aligned}
& \tilde{R}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}) \\
& \geq \tilde{a}_{\bar{\tau}_i, d_i}^{(\kappa)} + \tilde{b}_{\bar{\tau}_i, d_i}^{(\kappa)} \left( 2 - \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2}{2\Re\{(\mathbf{h}_{(B, d_i)}(\bar{\theta}) \bar{w}_{d_i}^{(\kappa)})(\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)})^*\} - |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2} \right. \\
& \quad \left. - \frac{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta})}{\tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \triangleq \tilde{R}_{\bar{\tau}_i, d_i}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{\theta}) \tag{2.52}
\end{aligned}$$

$$0 < \tilde{b}_{\bar{\tau}_i, d_i}^{(\kappa)} = \frac{|\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{(1 + |\mathbf{h}_{(B, d_i)}(\bar{\theta}^{(\kappa)}) \bar{w}_{d_i}^{(\kappa)}|^2 / \tilde{\psi}_{\bar{\tau}_i, d_i}(\bar{w}^{(\kappa)}, \bar{\theta}^{(\kappa)}))}.$$

Thêm vào đó, theo (2.28), ta có thể xấp xỉ  $\Omega(\bar{\theta})$  như sau

$$\Omega(\bar{\theta}) \geq \frac{1}{N} - \frac{1}{\sum_{n=1}^N (2\Re\{(\bar{\theta}_n^{(\kappa)})^* \bar{\theta}_n\} - |\bar{\theta}_n^{(\kappa)}|^2)} \triangleq \Omega^{(\kappa)}(\bar{\theta}).$$

Ràng buộc (2.51b) có thể được xấp xỉ bởi

$$\begin{aligned}
\tilde{E}_{\bar{\tau}_e, e_j}(\bar{v}^{(\kappa)}, \bar{\theta}) & \geq \sum_{l=1}^{U_E} \left[ 2\Re\{(\mathbf{h}_{(B, e_j)}(\bar{\theta}) \bar{v}_{e_l}^{(\kappa)})(\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)}) \bar{v}_{e_l}^{(\kappa)})^*\} \right. \\
& \quad \left. - |\mathbf{h}_{(B, e_j)}(\bar{\theta}^{(\kappa)}) \bar{v}_{e_l}^{(\kappa)}|^2 \right] \triangleq \tilde{E}_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{v}^{(\kappa)}, \bar{\theta}) \geq e_{\min} \bar{\tau}_e^{(\kappa)} / \rho. \tag{2.53}
\end{aligned}$$

Với ràng buộc (2.51c), áp dụng bất đẳng thức (A.2) trong Phụ lục cho  $\tilde{R}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta})$  with  $x = \bar{p}^{(\kappa)} |h_k(\bar{\theta})|^2$ ,  $y = \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta})$  và  $\bar{x} = \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2$ ,  $\bar{y} = \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})$ , ta có:

$$\begin{aligned}
\tilde{R}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}) & \geq \tilde{a}_{\bar{\tau}_d, k}^{(\kappa)} + \tilde{b}_{\bar{\tau}_d, k}^{(\kappa)} \left( 2 - \frac{|h_k(\bar{\theta}^{(\kappa)})|^2}{2\Re\{(h_k(\bar{\theta})) (h_k(\bar{\theta}^{(\kappa)}))^*\} - |h_k(\bar{\theta}^{(\kappa)})|^2} - \frac{\tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta})}{\tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})} \right) \\
& \triangleq \tilde{R}_{\bar{\tau}_d, k}^{(\kappa)}(\bar{p}^{(\kappa)}, \bar{\theta}) \tag{2.54}
\end{aligned}$$

với

$$2\Re\{(h_k(\bar{\theta})) (h_k(\bar{\theta}^{(\kappa)}))^*\} - |h_k(\bar{\theta}^{(\kappa)})|^2 \geq 0, \tag{2.55}$$

trong đó,

$$0 < \tilde{a}_{\bar{\tau}_d, k}^{(\kappa)} = \ln\left(1 + \frac{\bar{p}^{(\kappa)} |h_k|^2}{\tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}\right),$$

$$0 < \tilde{b}_{\bar{\tau}_d, k}^{(\kappa)} = \frac{\bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}{1 + \bar{p}^{(\kappa)} |h_k(\bar{\theta}^{(\kappa)})|^2 / \tilde{\psi}_{\bar{\tau}_d, k}(\bar{p}^{(\kappa)}, \bar{\theta}^{(\kappa)})}.$$

Từ các suy luận trên, bài toán 2 có thể được xấp xỉ thành bài toán tối ưu lồi sau:

$$\max_{\bar{\theta}} f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}) \triangleq \min_{d_i \in \mathcal{U}_I} (1/\bar{\tau}_d^{(\kappa)}) \tilde{R}_{\bar{\tau}_d, d_i}^{(\kappa)}(\bar{w}^{(\kappa)}, \bar{\theta}) + \eta \Omega^{(\kappa)}(\bar{\theta}) \quad (2.56a)$$

thoả mãn: (2.12), (2.53), (2.55),

$$\tilde{E}_{\bar{\tau}_e, e_j}^{(\kappa)}(\bar{v}^{(\kappa)}, \bar{\theta}) \geq e_{min} \bar{\tau}_e^{(\kappa)} / \rho, \forall e_j \in \mathcal{U}_E, \quad (2.56b)$$

$$(1/\bar{\tau}_d^{(\kappa)}) \tilde{R}_{\bar{\tau}_d, k}^{(\kappa)}(\bar{p}^{(\kappa)}, \bar{\theta}) \geq R_{k, min}, \forall k \in \mathcal{K}. \quad (2.56c)$$

Độ phức tạp tính toán của thuật toán giải bài toán tối ưu trên được xác định bởi (2.23), trong đó,  $\alpha = N$  và  $\beta = N + U_E + K$ . Ta có

$$\begin{aligned} f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) &= f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa)}) \\ &< f_{sub-2}^{(\kappa)}(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) \\ &\leq f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}), \end{aligned} \quad (2.57)$$

khi  $\bar{\theta}^{(\kappa+1)} \neq \bar{\theta}^{(\kappa)}$ .

Từ (2.24) và (2.37) ta có

$$f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) > f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)}). \quad (2.58)$$

Tức là điều kiện hội tụ được thỏa mãn và điểm  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})$  sẽ hội tụ về điểm tối ưu  $(\bar{w}^*, \bar{v}^*, \mathbf{p}^*, \tau^*, \bar{\theta}^*)$ .

Phương pháp giải bài toán tối ưu (2.42) được trình bày trong Thuật toán 2. Việc tìm điểm khả thi ban đầu và lựa chọn hệ số phạt  $\eta$  cho Thuật toán 2 tương tự như đã trình bày với Thuật toán 1.

## 2.5 Đánh giá hiệu quả mô hình và thuật toán

Trong phần này, luận án trình bày và đưa ra thảo luận về kết quả mô phỏng thu được bởi các thuật toán đề xuất. Để thuận tiện cho việc trình bày kết quả, luận án

---

**Thuật toán 2** Thuật toán OTA giải bài toán (2.42)
 

---

**Khởi tạo giá trị ban đầu:** Khởi tạo giá trị ban đầu: Tìm điểm khả thi  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(0)} = (\bar{\tau}_i^{(0)}, \bar{\tau}_e^{(0)}, \bar{\tau}_d^{(0)}, \bar{\theta}^{(0)})$  cho bài toán (2.42). Đặt  $\epsilon = 10^{-3}$  và  $\kappa = 0$ ;

- 1: Tính  $\eta$  theo (2.40) từ giá trị của điểm khả thi tìm được ban đầu;
- 2: **repeat**
- 3:   Giải bài toán tối ưu lồi (2.49) với  $\bar{\theta} = \bar{\theta}^{(\kappa)}$  để thu được  $(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$ ;
- 4:   Giải bài toán tối ưu lồi (2.56) với  $(\bar{\mathbf{w}}, \bar{\mathbf{v}}, \bar{\mathbf{p}}, \bar{\tau}) = (\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)})$  để thu được  $\bar{\theta}^{(\kappa+1)}$ ;
- 5:    $\hat{\epsilon} = \frac{|f(\bar{w}^{(\kappa+1)}, \bar{v}^{(\kappa+1)}, \bar{p}^{(\kappa+1)}, \bar{\tau}^{(\kappa+1)}, \bar{\theta}^{(\kappa+1)}) - f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})|}{f(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}, \bar{\theta}^{(\kappa)})}$
- 6:    $\kappa \leftarrow \kappa + 1$ ;
- 7: **until**  $\hat{\epsilon} \leq \epsilon$ .
- 8: Thu được nghiệm  $(\bar{w}^{(\kappa)}, \bar{v}^{(\kappa)}, \bar{p}^{(\kappa)}, \bar{\tau}^{(\kappa)}) = (\bar{\tau}_i^{(\kappa)}, \bar{\tau}_e^{(\kappa)}, \bar{\theta}^{(\kappa)})$

---

quan tâm đến kịch bản hệ thống bao gồm một BS/AP, một IRS và 3 robot di động, 2 cảm biến thu năng lượng EUs và 3 cặp D2D. Vị trí đặt của BS/AP và IRS lần lượt là  $(20, 0, 10)$  và  $(0, 30, 40)$ . Trong khi đó, các robot di động và các cặp D2D được đặt ngẫu nhiên trong vùng diện tích  $60\text{m} \times 60\text{m}$  về phía bên tay phải so với các vật cản và IRS, còn các cảm biến thu năng lượng EUs nằm ở vị trí ngẫu nhiên trong vùng bán kính 10m từ BS/AP. Mỗi kênh truyền bao gồm cả thành phần large-scale và small-scale fading. Large-scale fading là thành phần phụ thuộc vào khoảng cách truyền và được mô hình hóa bởi  $C\bar{d}^{-\alpha}$ , trong đó,  $\bar{d}$  là khoảng cách từ bên phát đến bên thu với đơn vị là mét (m),  $C = -30$  dBm là mất mát đường truyền (path loss) tại khoảng cách tham chiếu là 1 m, và  $\alpha$  là hệ số mất mát đường truyền. Các hệ số mất mát đường truyền tương ứng với các kênh truyền BS-IRS, IRS-IUs, IRS-D2Ds, IRS-EUs, BS-IUs, BS-D2Ds, và BS-EUs được đặt lần lượt là  $\alpha_{B-I} = 2.2$ ,  $\alpha_{I-IU} = 2.2$ ,  $\alpha_{I-D2D} = 2.2$ ,  $\alpha_{I-EU} = 2.2$ ,  $\alpha_{B-IU} = 3.67$ ,  $\alpha_{B-D2D} = 3.67$ , và  $\alpha_{B-EU} = 2$  [82]. Hệ số mất mát kênh truyền giữa bộ phát và bộ thu D2D, giữa bộ phát D2D với IUs và giữa bộ phát D2D với EUs được đặt lần lượt là  $\alpha_{D2D} = 2.8$ ,  $\alpha_{D-IU} = 2.8$ , và  $\alpha_{D-EU} = 3.67$  [61]. Ngoài ra, độ lợi ăng ten (antenna gain) của BS và độ lợi của IRS (gain of the IRS) là 5 dBi.

Giả sử ma trận kênh truyền trực tiếp từ BS/AP đến IRS có dạng

$$[\mathbf{G}]_{n,m} = e^{j\pi((n-1)\sin\bar{\theta}_{LoS}(n)\sin\bar{\phi}_{LoS}(n)+(m-1)\sin\bar{\theta}_{LoS}(n)\sin\phi_{LoS}(n))},$$

trong đó,  $\bar{\theta}_{LoS}(n)$  và  $\phi_{LoS}(n)$  theo phân phối đồng đều với  $\bar{\theta}_{LoS}(n) \sim \mathcal{U}(0, \pi)$  và  $\phi_{LoS}(n) \sim \mathcal{U}(0, 2\pi)$ , và  $\bar{\theta}_{LoS}(n) = \pi - \bar{\theta}_{LoS}(n)$  và  $\bar{\phi}_{LoS}(n) = \pi + \phi_{LoS}(n)$ . Ma trận tương quan không gian được định nghĩa bởi  $[H_{IRS-user}]_{n,n'} = e^{j\pi(n-n')\sin\bar{\theta}_{user}\sin\tilde{\phi}_{user}}$ , trong đó  $\bar{\theta}_{user}$  và  $\tilde{\phi}_{user}$  lần lượt là góc phương vị và góc độ cao tương ứng với mỗi thiết bị. Các kênh small-scale fading từ BS đến IUs và từ BS đến D2Ds được xem xét là có

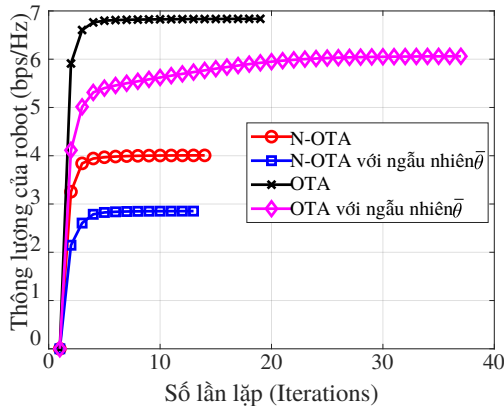
phân phối Rayleigh độc lập, trong khi các thành phần small-scale fading của các kênh khác tuân theo phân phối Rician độc lập với một hệ số K-factor là 3. Cụ thể, kênh D2D  $h_k$  được xác định bởi  $h_k = \sqrt{\frac{\kappa}{\kappa+1}}h_k^{LoS} + \sqrt{\frac{1}{\kappa+1}}h_k^{NLoS}$ , trong đó,  $\kappa$  là hệ số Rician,  $h_k^{LoS}$  là thành phần LoS (line-of-sight) và  $h_k^{NLoS}$  là thành phần NLoS (non-LoS). Các tham số mô phỏng khác được cung cấp trong Bảng 2.1.

Bảng 2.1: Bảng tham số mô phỏng thuật toán N-OTA và OTA

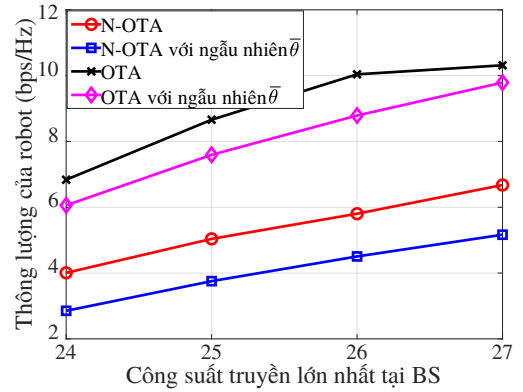
Tham số	Giá trị	Tham số	Giá trị
$U_I$	3	$U_E$	2
$\bar{M}$	10	$N$	200
$B$	10 MHz	$P_{B,max}$	24 dBm
$\rho$	0.5	$e_{min}$	-20 dBm
$K$	3	$R_{k,min}$	0.1 bps/Hz
$P_{k,max}$	24 dBm	$\sigma_{d_i}^2, \sigma_k^2$	-174 dBm/Hz

Nhằm mục đích so sánh, luận án đưa ra các thuật toán sau:

- **N-OTA:** Đây là thuật toán mô tả trong phần 2.3 (Thuật toán 1), thuật toán này được đề xuất sử dụng để giải bài toán tối ưu (2.9).
- **N-OTA với ngẫu nhiên  $\bar{\theta}$ :** Thuật toán này tương tự như thuật toán N-OTA trong đó giá trị độ dịch pha của IRS được lấy ngẫu nhiên chứ không tối ưu.
- **OTA:** Đây là thuật toán mô tả trong phần 2.4 (Thuật toán 2), thuật toán này được đề xuất sử dụng để giải bài toán tối ưu (2.42).
- **OTA với ngẫu nhiên  $\bar{\theta}$ :** Thuật toán này tương tự như thuật toán OTA trong đó giá trị độ dịch pha của IRS được lấy ngẫu nhiên chứ không tối ưu.



(a)

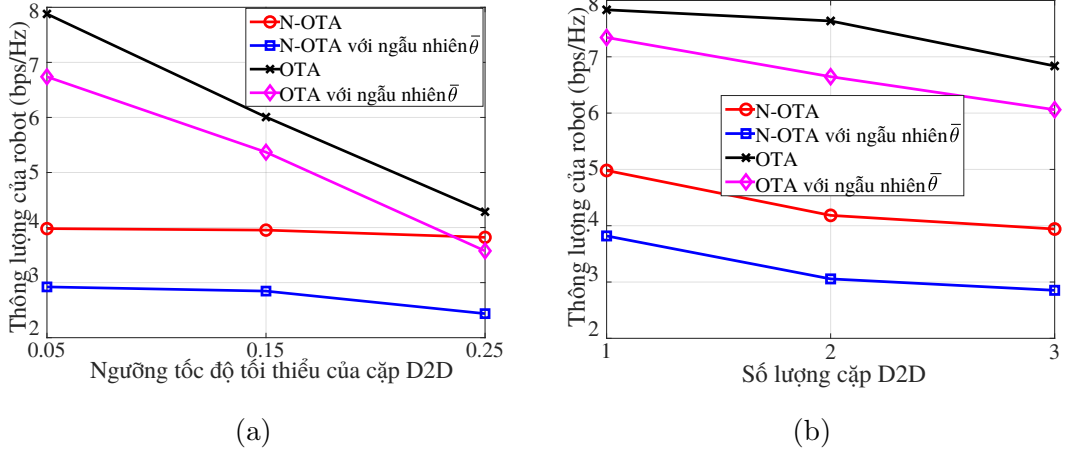


(b)

Hình 2.4: a) Độ hội tụ của các thuật toán và b) thông lượng dữ liệu thay đổi với công suất định mức của BS  $P_{B,max}$ .

Đầu tiên, luận án thảo luận về sự hội tụ của các thuật toán và thông lượng truyền dữ liệu tại IUs đạt được bởi các thuật toán. Như được thể hiện trong Hình 2.4(a), tất cả các thuật toán đều có khả năng hội tụ nhanh chóng đến các giá trị ổn định. Hơn nữa, thông lượng đạt được bởi các IUs với thuật toán OTA cao hơn so với thuật toán N-OTA. Điều này là do sự tồn tại của nhiễu gây ra bởi truyền thông D2D đến các IUs trong kịch bản N-OTA làm giảm SINR tại các IUs và giảm thông lượng IUs. Cũng có thể thấy từ Hình 2.4(a) rằng thông lượng đạt được bởi OTA và N-OTA cao hơn nhiều so với các thông lượng đạt được bởi OTA và N-OTA với ngẫu nhiên  $\bar{\theta}$ , tương ứng.

Tiếp theo, luận án thảo luận về tác động của công suất phát tối đa, tức là  $P_{B,max}$ , của BS đối với thông lượng đạt được tại các IUs. Như được thể hiện trong Hình 2.4(b), khi  $P_{B,max}$  tăng lên, thông lượng của các IUs cũng tăng lên. Điều này là do thông lượng dữ liệu của mỗi IU tỉ lệ thuận với  $P_{B,max}$ . Hơn nữa, có thể thấy rằng qua các giá trị của  $P_{B,max}$ , thông lượng đạt được bởi các thuật toán được đề xuất với tối ưu độ dịch pha luôn cao hơn rất nhiều so với thông lượng đạt được bởi các thuật toán cơ sở với ngẫu nhiên  $\bar{\theta}$ . Thuật toán N-OTA cải thiện thông lượng lên đến 10% so với thuật toán N-OTA với ngẫu nhiên  $\bar{\theta}$ .

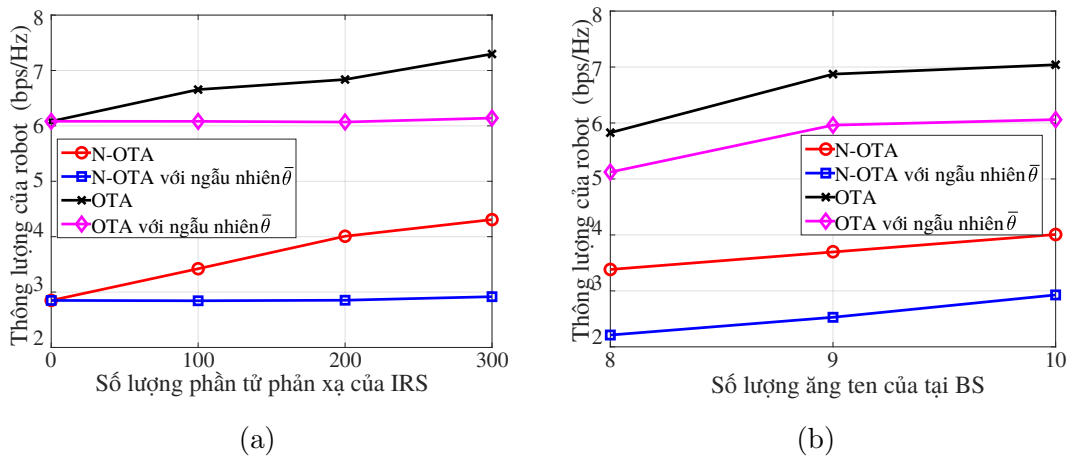


Hình 2.5: Thông lượng đạt được với a) ngưỡng tốc độ truyền thông D2D  $R_{k,min}$  và b) số lượng cặp D2D  $K$

Do mạng tích hợp truyền thông tin và năng lượng chia sẻ băng tần với các cặp truyền thông D2D, việc xem xét ảnh hưởng của ngưỡng tốc độ D2D, tức là  $R_{k,min}$ , đến thông lượng đạt được bởi các IUs là rất quan trọng. Như được thể hiện trong Hình 2.5(a), khi  $R_{k,min}$  tăng lên, thông lượng đạt được bởi tất cả các thuật toán đều giảm. Kết quả này có thể được giải thích như sau. Với thuật toán N-OTA, khi  $R_{k,min}$  tăng lên, các bộ phát D2D cần phát tín hiệu của chúng với công suất cao hơn. Điều

này làm tăng nhiều từ truyền thông D2D đến các IUs, làm giảm SINR tại các IUs và giảm thông lượng của chúng. Với thuật toán OTA, khi  $R_{k,min}$  tăng lên, thêm thời gian được phân bổ cho truyền thông D2D, điều này có thể làm giảm thời gian truyền thông tin đến các IUs và làm giảm thông lượng của chúng. Đáng chú ý là khi  $R_{k,min}$  tăng lên, thông lượng đạt được bởi thuật toán OTA giảm nhanh hơn so với thuật toán N-OTA. Lý do có thể là việc tối ưu độ dịch pha của các thuật toán N-OTA giúp giảm nhiễu gây ra bởi truyền thông D2D đến các IUs. Do đó, khi  $R_{k,min}$  tăng lên, sự tăng nhiễu gây ra bởi truyền thông D2D có thể không đáng kể, và thông lượng đạt được bởi thuật toán N-OTA giảm dần. Hơn nữa, thông lượng đạt được bởi các thuật toán OTA luôn cao hơn rất nhiều so với thông lượng đạt được bởi các thuật toán N-OTA.

Lúc này, luận án sẽ thảo luận về tác động của số cặp D2D  $K$  đối với thông lượng của các IUs, và kết quả được thể hiện trong Hình 2.5(b). Như đã thấy, khi  $K$  tăng lên, thông lượng đạt được bởi cả hai thuật toán N-OTA và OTA đều giảm. Đặc biệt đối với thuật toán N-OTA, nguyên nhân là khi  $K$  tăng lên, nhiễu gây ra bởi truyền thông D2D đến các IUs cũng tăng lên. Đối với thuật toán OTA, khi  $K$  tăng lên, SINR tại mỗi bộ nhận D2D giảm do sự tăng lên của nhiễu từ các bộ phát D2D. Để đảm bảo được một ngưỡng tốc độ D2D xác định, cần thêm thời gian phân bổ cho truyền thông D2D. Kết quả là điều này làm giảm thời gian phân bổ cho các IUs và thông lượng của chúng theo đó cũng giảm. Lưu ý rằng qua các giá trị của  $K$ , thông lượng với tối ưu độ dịch pha luôn cao hơn so với thông lượng khi để  $\bar{\theta}$  ngẫu nhiên.



Hình 2.6: Thông lượng đạt được với a) số lượng các phần tử của IRS và b) số lượng ăng ten tại BS.

Một điều đáng quan tâm là kích thước của IRS, tức là  $N$ , cũng có thể ảnh hưởng đến thông lượng đạt được bởi các IUs. Như được thể hiện trong Hình 2.6(a),

khi  $N$  tăng lên, thông lượng đạt được bởi cả hai thuật toán N-OTA và OTA đều tăng lên. Nguyên nhân là SINR tại các IUs tăng khi  $N$  tăng lên. Tuy nhiên, tốc độ tăng của các thuật toán N-OTA nhanh hơn so với thuật toán OTA. Điều này có thể được giải thích dựa trên định nghĩa của thông lượng của các IUs (trong các phần III và IV). Thực sự, trong cả hai tình huống N-OTA và OTA, thông lượng của mỗi IU có dạng  $\bar{t}_i R_{\bar{t}_i, d_i}$ , trong đó  $\bar{t}_i$  là thời gian được phân bổ cho IU  $i$  và  $R_{\bar{t}_i, d_i}$  là tốc độ dữ liệu được đạt được bởi IU. Trong kịch bản N-OTA, không có thời gian được phân bổ cho truyền thông D2D, và do đó  $\bar{t}_i$  lớn hơn so với kịch bản OTA. Do đó, khi  $N$  tăng lên,  $R_{\bar{t}_i, d_i}$  được đạt được bởi thuật toán N-OTA tăng nhanh hơn so với thuật toán OTA. Thông lượng của các IUs cũng tăng khi số ăng-ten của BS tăng lên như được thể hiện trong Hình 2.6(b). Điều này là rõ ràng do SINR tại các IUs tăng lên.

## 2.6 Kết luận chương 2

Chương 2 đã trình bày đề xuất mô hình và thuật toán để tối ưu hệ thống mạng trong các nhà máy, cụ thể, các đóng góp của chương bao gồm:

- Đề xuất sử dụng IRS cho hệ thống mạng tích hợp truyền thông tin của robot di động và thu năng lượng của cảm biến không dây kết hợp với truyền thông D2D.
- Đề xuất thuật toán N-OTA và OTA nâng cao chất lượng truyền thông của robot di động trong mô hình tương ứng với kịch bản trong đó không phân bổ và có phân bổ thời gian cho truyền thông D2D với truyền thông tin và năng lượng.

Như vậy, dựa trên đề xuất mô hình và thuật toán, kết hợp với việc thiết kế hệ thống thực tế, môi trường truyền thông tại các nhà máy có thể trở nên lý tưởng, tức là lúc này mọi vị trí trong nhà máy đều được phủ sóng và đạt được tốc độ truyền thông theo yêu cầu. Khi đó, môi trường truyền thông được gọi là lý tưởng. Khi đó, bài toán lập quỹ đạo cho robot sẽ không bị ảnh hưởng bởi ràng buộc về tốc độ truyền thông. Ở chương tiếp theo, chương 3, thuật toán lập quỹ đạo cho robot trong môi trường truyền thông lý tưởng sẽ được trình bày.

Nội dung của chương 2 được công bố trong công trình CT1. Công trình CT1 trình bày đề xuất một mô hình mạng không dây DED2D và thuật toán tối ưu mạng cho phép các robot di động truyền thông tin, các thiết bị IoT thu năng lượng và các thiết bị sử dụng phương thức giao tiếp trực tiếp trong nhà máy có thể chia sẻ cùng băng thông mạng mà vẫn đạt được hiệu quả về truyền thông cho từng loại thiết bị.

[CT1] **N. T. T. Van**, H. T. Nguyen, N. C. Luong, N. M. Tien, D. Niyato, and D. I. Kim, “Intelligence Reflecting Surface-Aided Integrated Data and Energy Networking Coexisting D2D Communications,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2022.



### Chương 3. THUẬT TOÁN TỐI ƯU QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG MÔI TRƯỜNG TRUYỀN THÔNG LÝ TƯỞNG

Giả sử bằng cách áp dụng mô hình và thuật toán đề xuất ở chương 1, môi trường truyền thông trong nhà máy đã đảm bảo chất lượng và ổn định tại mọi vị trí làm việc của robot. Ở chương này, NCS sẽ xây dựng nhiệm vụ lập quỹ đạo cho robot thành một bài toán tối ưu. Bài toán tối ưu này là không lồi nên gặp khó khăn khi giải. Do đó, NCS đã đề xuất thuật toán dựa trên xấp xỉ lồi để giải quyết. Các kết quả mô phỏng đã chứng minh hiệu quả của thuật toán đề xuất. Cụ thể, ở phần 3.1 trình bày cơ sở lý thuyết về bài toán tối ưu lồi và thuật toán lập xấp xỉ lồi để giải bài toán tối ưu bằng công cụ CVX trên Matlab. Sau đó, luận án trình bày thuật toán thiết kế quỹ đạo với mục tiêu tối thiểu khoảng cách di chuyển cho robot ở phần 3.2. Thuật toán thiết kế quỹ đạo với mục tiêu tối thiểu năng lượng tiêu thụ cho robot sẽ trình bày ở phần 3.3. Cuối cùng là kết luận chương.

#### 3.1 Tối ưu lồi

##### 3.1.1 Bài toán tối ưu

Một bài toán tối ưu hóa nói chung, có thể không có hoặc có các ràng buộc. Tuy nhiên, trên thực tế, các bài toán tối ưu thường có rất nhiều ràng buộc. Dưới đây là hai dạng tổng quát của bài toán tối ưu có ràng buộc

$$\begin{aligned} \mathbf{x}^* &= \arg \min_{\mathbf{x}} f_0(\mathbf{x}) \\ \text{thỏa mãn: } & f_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, p \end{aligned} \quad (3.1)$$

hoặc

$$\begin{aligned} \mathbf{x}^* &= \arg \max_{\mathbf{x}} g_0(\mathbf{x}) \\ \text{thỏa mãn: } & g_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, m \\ & h_j(\mathbf{x}) = 0, \quad j = 1, 2, \dots, p \end{aligned} \quad (3.2)$$

Trong đó, vector  $\mathbf{x} = [x_1, x_2, \dots, x_n]^T$  được gọi là *biến tối ưu* (optimization variable). Hàm số  $f_0 : \mathbb{R}^n \rightarrow \mathbb{R}$  được gọi là *hàm mục tiêu* (objective function). Các

bất phương trình  $f_i(\mathbf{x}) \leq 0, i = 1, 2, \dots, m$  được gọi là *bất phương trình ràng buộc* (inequality constraint), và các hàm tương ứng  $f_i(\mathbf{x}), i = 1, 2, \dots, m$  được gọi là *hàm bất phương trình ràng buộc* (inequality constraint function). Các phương trình  $h_j(\mathbf{x}) = 0, j = 1, 2, \dots, p$  được gọi là các *phương trình ràng buộc* (equality constraint), các hàm tương ứng là các *hàm phương trình ràng buộc* (equality constraint function).

Tập hợp các điểm mà trên đó hàm số  $f_0(\mathbf{x})$  xác định được gọi là *tập xác định* (domain) hay  $\mathbf{dom}f$ . Tập xác định của một bài toán tối ưu là giao của tập xác định tất cả các hàm liên quan [79]:

$$\mathcal{D} = \bigcap_{i=0}^m \mathbf{dom}f_i \cap \bigcap_{j=1}^p \mathbf{dom}h_j \quad (3.3)$$

Một điểm  $\mathbf{x}$  thuộc tập xác định được gọi là *điểm khả thi* (feasible point) nếu thỏa mãn tất cả các ràng buộc. Tập hợp các điểm khả thi được gọi là *tập khả thi* (feasible set).

$$\mathcal{X} = \{\mathbf{x} \in \mathbb{R}^n \mid f_i(\mathbf{x}) \leq 0, h_j(\mathbf{x}) = 0, i = 1, 2, \dots, m, j = 1, 2, \dots, p\} \quad (3.4)$$

Bài toán (3.1) được gọi là *khả thi* nếu tập khả thi  $\mathcal{X}$  khác tập rỗng.

Bài toán tối ưu dạng (3.1) được phát biểu bằng lời như sau: Tìm giá trị tối ưu của biến  $\mathbf{x}$  để tối thiểu hàm  $f_0(\mathbf{x})$  trong số những giá trị  $\mathbf{x}$  thỏa mãn các điều kiện ràng buộc. Tương tự cho bài toán tối ưu dạng (3.2): Tìm giá trị tối ưu của biến  $\mathbf{x}$  để tối đa hàm  $f_0(\mathbf{x})$  trong số những giá trị  $\mathbf{x}$  thỏa mãn các điều kiện ràng buộc.

*Giá trị tối ưu* (optimal value) của bài toán tối ưu (3.1) được định nghĩa là:

$$v^* = \inf \{f_0(\mathbf{x}) \mid f_i(\mathbf{x}) \leq 0, i = 1, \dots, m; h_j(\mathbf{x}) = 0, j = 1, \dots, p\}$$

### Điểm tối ưu và tối ưu cục bộ

Một điểm  $\mathbf{x}^*$  được gọi là *điểm tối ưu* (optimal point) của bài toán (3.1) nếu  $\mathbf{x}^*$  là một điểm khả thi và  $f_0(\mathbf{x}^*) = v^*$ . Tập hợp tất cả các điểm tối ưu được gọi là *tập tối ưu* (optimal set). Nếu tập tối ưu khác rỗng, ta nói bài toán (3.1) *giải được* (solvable). Ngược lại, nếu tập tối ưu rỗng, ta nói giá trị tối ưu không thể đạt được.

Để làm rõ vấn đề, xét hàm mục tiêu  $f(x) = 1/x$  với ràng buộc  $x > 0$ . Giá trị tối ưu của bài toán này là  $v^* = 0$  nhưng tập tối ưu là một tập rỗng vì không có giá trị

nào của  $x$  để hàm mục tiêu đạt giá trị  $v^*$ .

Với hàm một biến, một điểm là *tối ưu cục bộ* của hàm số nếu tại đó hàm số đạt giá trị nhỏ nhất trong một lân cận (và lân cận này thuộc tập xác định của hàm số). Trong không gian một chiều, lân cận của một điểm được hiểu là tập các điểm cách điểm đó một khoảng rất nhỏ. Trong không gian nhiều chiều, ta gọi một điểm  $\mathbf{x}$  là *tối ưu cục bộ* nếu tồn tại một giá trị  $R > 0$  sao cho:

$$\begin{aligned} f_0(\mathbf{x}) = & \inf\{f_0(\mathbf{z}) \mid f_i(\mathbf{z}) \leq 0, i = 1, \dots, m, \\ & h_j(\mathbf{z}) = 0, j = 1, \dots, p, \|\mathbf{z} - \mathbf{x}\|_2 \leq R, \mathbf{x} \in \mathcal{D}\} \end{aligned} \quad (3.5)$$

Nhìn chung, không có cách giải quyết tổng quát cho các bài toán tối ưu, thậm chí nhiều bài toán tối ưu chưa có lời giải hiệu quả. Hầu hết các phương pháp không chứng minh được nghiệm tìm được có phải là điểm tối ưu toàn cục hay không. Thay vào đó, nghiệm thường là các nghiệm tối ưu cục bộ. Trong nhiều trường hợp, các nghiệm tối ưu cục bộ cũng mang lại những kết quả tốt. Trong luận án này, NCS sẽ trình bày về các vấn đề liên quan đến phương pháp tối ưu lồi.

### 3.1.2 Bài toán tối ưu lồi

Bài toán tối ưu lồi là bài toán tối ưu có dạng

$$\begin{aligned} \mathbf{x}^* = \arg \min_{\mathbf{x}} f_0(\mathbf{x}) \\ \text{thỏa mãn: } f_i(\mathbf{x}) \leq 0, \quad i = 1, 2, \dots, m \end{aligned} \quad (3.6)$$

trong đó hàm mục tiêu  $f_0(\mathbf{x})$  là một *hàm lồi* (convex function), tập khả thi là một *tập lồi* (convex set), các hàm  $f_i(\mathbf{x}), i = 1, 2, \dots, m$  là các hàm lồi.

**Trong bài toán tối ưu lồi, nghiệm tối ưu cục bộ cũng chính là nghiệm tối ưu toàn cục [79, 83]**

Tính chất quan trọng nhất của bài toán tối ưu lồi chính là mọi điểm tối ưu cục bộ đều là điểm tối ưu toàn cục. Điều này có thể chứng minh bằng phản chứng. Gọi  $\mathbf{x}_0$  là một điểm nghiệm tối ưu cục bộ:

$$f_0(\mathbf{x}_0) = \inf\{f_0(\mathbf{x}) \mid \mathbf{x} \in \text{tập khả thi}, \|\mathbf{x} - \mathbf{x}_0\|_2 \leq R\}$$

với  $R > 0$  nào đó. Giả sử  $\mathbf{x}_0$  không phải là một nghiệm tối ưu toàn cục, tức tồn tại một điểm khả thi  $\mathbf{y}$  sao cho  $f(\mathbf{y}) < f(\mathbf{x}_0)$  (hiển nhiên  $\mathbf{y}$  không nằm trong lân cận đang xét). Ta có thể tìm được  $\theta \in [0, 1]$  sao cho  $\mathbf{z} = (1 - \theta)\mathbf{x}_0 + \theta\mathbf{y}$  nằm trong lân cận của  $\mathbf{x}_0$ , tức  $\|\mathbf{z} - \mathbf{x}_0\|_2 < R$ . Điều này có được vì tập khả thi là một tập lồi. Hơn nữa, vì hàm mục tiêu  $f_0$  là một hàm lồi, ta có

$$f_0(\mathbf{z}) = f_0((1 - \theta)\mathbf{x}_0 + \theta\mathbf{y}) \quad (3.7)$$

$$\leq (1 - \theta)f_0(\mathbf{x}_0) + \theta f_0(\mathbf{y}) \quad (3.8)$$

$$< (1 - \theta)f_0(\mathbf{x}_0) + \theta f_0(\mathbf{x}_0) = f_0(\mathbf{x}_0) \quad (3.9)$$

Điều này mâu thuẫn với giả thiết  $\mathbf{x}_0$  là một nghiệm tối ưu cục bộ và  $\mathbf{z}$  nằm trong lân cận của  $\mathbf{x}_0$ . Vậy giả thiết phản chứng là sai, tức  $\mathbf{x}_0$  chính là một nghiệm tối ưu toàn cục.

Nói cách khác, giả sử một nghiệm tối ưu cục bộ không phải là nghiệm tối ưu toàn cục. Vì hàm mục tiêu và tập khả thi đều lồi, ta luôn tìm được một điểm khác trong lân cận của điểm tối ưu cục bộ đó sao cho giá trị của hàm mục tiêu tại điểm mới này nhỏ hơn giá trị của hàm mục tiêu tại điểm tối ưu cục bộ. Sự mâu thuẫn này chỉ ra rằng với một bài toán tối ưu lồi, nghiệm tối ưu cục bộ phải là nghiệm tối ưu toàn cục.

### Điều kiện tối ưu cho hàm mục tiêu khả vi

Nếu hàm mục tiêu  $f_0$  là khả vi, theo điều kiện bậc nhất, với mọi  $\mathbf{x}, \mathbf{y} \in \mathbf{dom} f_0$ , ta có:

$$f_0(\mathbf{x}) \geq f_0(\mathbf{x}_0) + \nabla f_0(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0). \quad (3.10)$$

Đặt  $\mathcal{X}$  là tập khả thi. Điều kiện cần và đủ để một điểm  $\mathbf{x}_0 \in \mathcal{X}$  là điểm tối ưu là:

$$\nabla f_0(\mathbf{x}_0)^T(\mathbf{x} - \mathbf{x}_0) \geq 0, \quad \forall \mathbf{x} \in \mathcal{X}. \quad (3.11)$$

Điều này chỉ ra rằng nếu  $\nabla f_0(\mathbf{x}_0) = 0$  thì  $\mathbf{x}_0$  chính là một điểm tối ưu của bài toán. Nếu  $\nabla f_0(\mathbf{x}_0) \neq 0$ , nghiệm của bài toán sẽ phải nằm trên biên của tập khả thi.

### 3.1.3 Thuật toán tối ưu lồi

Trên thực tế, có nhiều thuật toán dựa trên tối ưu lồi. Tuy nhiên, luận án quan tâm đến phương pháp điểm trong (interior point) cụ thể là sử dụng phương pháp điểm

trong trên nền kỹ thuật hàm Lagrange [84] để giải bài toán tối ưu lồi (3.6).

Đặt  $\bar{\mathbf{x}} = [x_{n+1}, x_{n+2}, \dots, x_{n+m}]^T$  với mọi phần tử của  $\bar{\mathbf{x}}$  đều không âm. Khi đó, bài toán (3.6) có thể viết lại thành

$$\hat{\mathbf{x}}^* = \arg \min f_0(\hat{\mathbf{x}}) \quad (3.12a)$$

$$\text{thỏa mãn: } \hat{f}_i(\hat{\mathbf{x}}) = 0, \quad \forall i = 1, 2, \dots, m, \quad (3.12b)$$

$$x_{n+i} \geq 0, \quad \forall i = 1, 2, \dots, m, \quad (3.12c)$$

với  $\hat{\mathbf{x}} = \text{col}(\mathbf{x}^T, \bar{\mathbf{x}})$ ,  $\hat{f}_i(\hat{\mathbf{x}}) = f_i(\mathbf{x}) + x_{n+i}$ .

Để loại bỏ điều kiện ràng buộc (3.12c), kỹ thuật hàm chặn được sử dụng. Đầu tiên, cần xác định một hàm chặn lồi (barrier function), ký hiệu là  $\hat{B}(\bar{\mathbf{x}})$ . Sau đó, đặt  $\hat{f}_0(\hat{\mathbf{x}}, \hat{\mu}) = f_0(\hat{\mathbf{x}}) + \hat{\mu}\hat{B}(\bar{\mathbf{x}})$  với  $\hat{\mu} \geq 0$  là một số dương rất nhỏ được chọn trước. Khi đó, bài toán tối ưu lồi (3.12) trở thành

$$\hat{\mathbf{x}}^* = \arg \min \hat{f}_0(\hat{\mathbf{x}}, \hat{\mu}) \quad (3.13a)$$

$$\text{thỏa mãn: } \hat{f}_i(\hat{\mathbf{x}}) = 0, \quad \forall i = 1, 2, \dots, m, \quad (3.13b)$$

Bài toán (3.13) là bài toán tối ưu lồi có ràng buộc là các phương trình. Do đó, nghiệm của bài toán (3.13) cũng chính là nghiệm của

$$f(\hat{\mathbf{x}}, \hat{\lambda}) = 0 \quad (3.14)$$

trong đó,

$$f(\hat{\mathbf{x}}, \hat{\lambda}) = \nabla \hat{\mathcal{L}}(\hat{\mathbf{x}}, \hat{\lambda}), \quad (3.15)$$

và

$$\hat{\mathcal{L}}(\hat{\mathbf{x}}, \hat{\lambda}) = \hat{f}_0(\hat{\mathbf{x}}, \hat{\mu}) + \hat{\lambda}^T \hat{\mathbf{f}}(\hat{\mathbf{x}}) = f_0(\hat{\mathbf{x}}) + \hat{\mu}\hat{B}(\bar{\mathbf{x}}) + \hat{\lambda}^T(\hat{\mathbf{x}}) \quad (3.16)$$

với  $\hat{\mathbf{f}}(\hat{\mathbf{x}}) = \left[ \hat{f}_1(\hat{\mathbf{x}}), \hat{f}_2(\hat{\mathbf{x}}), \dots, \hat{f}_m(\hat{\mathbf{x}}) \right]^T$  là hàm Lagrange của bài toán (3.13). Từ đây, ta có phương pháp tối ưu lồi sử dụng phương pháp điểm trong với kỹ thuật hàm Lagrange được mô tả ở Thuật toán 3.

Phần tiếp theo sẽ giới thiệu công cụ thực hiện thuật toán 3.

---

**Thuật toán 3** Thuật toán tối ưu lồi dùng phương pháp điểm trong với kỹ thuật hàm Lagrange [84]

---

- 1: Chọn hàm chặn  $\hat{B}(\bar{\mathbf{x}})$  có tính lồi, trong đó  $\bar{\mathbf{x}} = [x_{n+1}, x_{n+2}, \dots, x_{n+m}]^T$ ;
  - 2: Chọn  $\hat{\mu} > 0$  đủ nhỏ, xây dựng hàm Lagrange  $\hat{\mathcal{L}}(\hat{\mathbf{x}}, \hat{\lambda})$  theo (3.16), trong đó,  $\hat{\mathbf{f}}(\hat{\mathbf{x}}) = \begin{bmatrix} \hat{f}_1(\hat{\mathbf{x}}), \hat{f}_2(\hat{\mathbf{x}}), \dots, \hat{f}_m(\hat{\mathbf{x}}) \end{bmatrix}^T$  và  $\hat{\mathbf{x}} = [\mathbf{x}^T, \bar{\mathbf{x}}^T]^T$ ;
  - 3: Tính vector hàm  $f(\hat{\mathbf{x}}, \hat{\lambda})$  theo (3.15);
  - 4: Chọn điểm bắt đầu  $(\hat{\mathbf{x}}_0, \hat{\lambda}_0)$ . Đặt  $k = 0$ , và sai số hội tụ rất nhỏ  $\hat{\varepsilon} > 0$ ;
  - 5: **repeat**
  - 6:    Tính  $H_k = \frac{\partial f(\hat{\mathbf{x}}, \hat{\lambda})}{\partial (\hat{\mathbf{x}}, \hat{\lambda})} \Big|_{\hat{\mathbf{x}}_k, \hat{\lambda}_k}$ , là ma trận Hesse của  $\hat{\mathcal{L}}(\hat{\mathbf{x}}, \hat{\lambda})$  tại  $(\hat{\mathbf{x}}_k, \hat{\lambda}_k)$ ;
  - 7:    Tính  $(\hat{\mathbf{x}}_{k+1}, \hat{\lambda}_{k+1}) = (\hat{\mathbf{x}}_k, \hat{\lambda}_k) + H_k^{-1} f(\hat{\mathbf{x}}_k, \hat{\lambda}_k)$ ;
  - 8:    Đặt  $\kappa \leftarrow \kappa + 1$
  - 9: **until** Hội tụ
  - 10: Điểm tối ưu thu được  $(\hat{\mathbf{x}}_k, \hat{\lambda}_k)$
- 

### 3.1.4 Giới thiệu CVX

CVX là một hệ thống mô hình hóa để xây dựng và giải các bài toán tối ưu lồi có định dạng theo quy định (disciplined convex programming DCPs). DCP là một phương thức xây dựng các bài toán tối ưu lồi được đề xuất bởi Michael Grant, Stephen Boyd và Yinyu Ye [85, 86]. Điều này nhằm hỗ trợ việc định dạng và xây dựng các bài toán tối ưu mà được xác định là các bài toán tối ưu lồi.

CVX hỗ trợ một lượng lớn các kiểu bài toán chuẩn bao gồm quy hoạch tuyến tính (linear programming LP), quy hoạch toàn phương (quadratic programming QP), quy hoạch nón bậc hai (second-order cone programming SOCP), và quy hoạch bán xác định (semidefinite programming SDP). CVX cũng có thể giải nhiều bài toán tối ưu lồi phức tạp hơn bao gồm nhiều hàm không khả vi như hàm chuẩn  $l_1$ . CVX được sử dụng để xây dựng và giải các bài toán tối thiểu hàm chuẩn, tối đa hàm entropy, tối đa hàm determinant có ràng buộc và nhiều bài toán tối ưu lồi khác. Từ bản 2.0, CVX đã có thể giải các bài toán có lẫn biến số nguyên (mixed integer disciplined convex programs MIDCPs) cùng với việc tích hợp các trình giải (solver) tương thích.

CVX được thực hiện trên Matlab, và sử dụng ngôn ngữ của Matlab cho việc mô hình hóa bài toán tối ưu một cách hiệu quả. Các đặc điểm kỹ thuật của mô hình được xây dựng thông thường dựa trên các toán tử và hàm thông dụng của Matlab và ngôn ngữ lập trình tiêu chuẩn của Matlab có thể được kết hợp với những đặc điểm kỹ thuật này. Sự kết hợp này làm đơn giản hóa các phép tính để định dạng bài toán tối ưu cũng như xử lý kết quả tối ưu và nghiệm thu được.

CVX cung cấp các mô hình đặc biệt để đơn giản hóa việc xây dựng các bài

toán. Ở chế độ quy hoạch bán xác định (SDP mode), CVX áp dụng giải tích ma trận cho toán tử bất đẳng thức nên các bất đẳng thức ma trận tuyến tính (linear matrix inequalities LMIs) và các bài toán bán xác định SDPs có thể được biểu diễn dưới dạng tự nhiên hơn. Ở chế độ quy hoạch hình học (geometric programming GP mode), CVX chấp nhận tất cả các hàm đặc biệt và các luật kết hợp của GP bao gồm đơn thức, đa thức, đa thức tổng quát, và chuyển dạng các bài toán đó sang dạng lồi để giải.

Phiên bản trước của CVX hỗ trợ hai trình giải miễn phí là SeDuMi và SDPT3. Từ bản 2.0, CVX hỗ trợ thêm hai trình giải thương mại, đó là Gurobi và MOSEK. Các trình giải này đã mở rộng khả năng của CVX trong việc giải các bài toán. Cụ thể, trình giải MOSEK đã hỗ trợ trên việc giải các hàm mũ, điều này cho phép giải các bài toán có chứa hàm logarit, hàm mũ và hàm entropy mà không cần phải xấp xỉ.

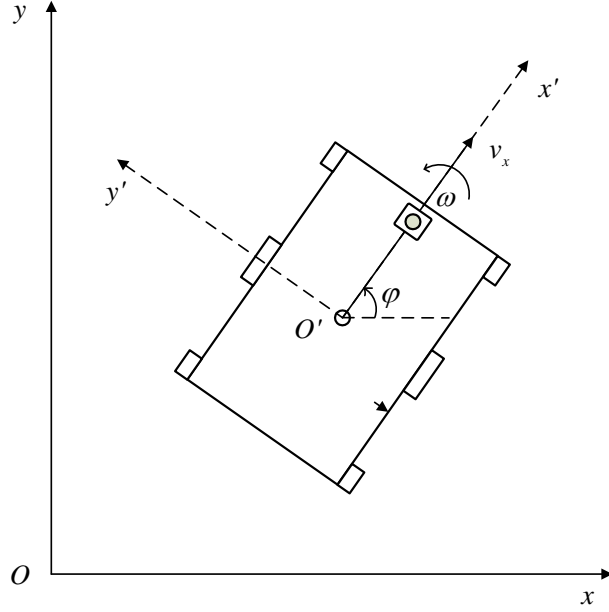
Những tính chất đặc biệt về nghiệm tối ưu cục bộ và toàn cục của một hàm lồi khiến tối ưu lồi trở nên cực kỳ quan trọng. Từ đó dẫn đến nhiều thuật toán tối ưu lồi ra đời [87]. Tiếp theo, NCS đề xuất thuật toán dựa trên xấp xỉ lồi để giải bài toán tối ưu quỹ đạo cho robot di động.

## 3.2 Phát triển thuật toán tối ưu quỹ đạo cho robot di động

### 3.2.1 Mục tiêu và các ràng buộc của bài toán

Trong luận án, NCS quan tâm đến robot di động dạng bánh xe vi sai như mô tả ở hình 3.1. Đây là loại robot di động có 2 bánh chủ động ở chính giữa và có 4 bánh chịu tải được đặt ở các góc xung quanh robot để tăng khả năng chịu tải cũng như sự cân bằng cho robot trong quá trình hoạt động. Chuyển động và hướng được thỏa mãn bởi các cơ cấu chấp hành độc lập.

Đối với loại robot mô tả trên Hình 3.1, robot sẽ chuyển động tiến thẳng về phía trước khi vận tốc của hai bánh xe bằng nhau, robot chuyển động quay phải khi tốc độ quay của bánh xe bên trái lớn hơn tốc độ quay của bánh xe bên phải, và robot chuyển động quay trái khi tốc độ quay của bánh xe bên phải lớn hơn tốc độ quay của bánh xe bên trái. Ngoài ra, robot có thể quay tròn tại chỗ khi hai bánh xe quay với vận tốc có độ lớn bằng nhau nhưng ngược chiều nhau [88]. Bốn bánh phụ ở bốn góc được sử dụng để đảm bảo sự ổn định và thăng bằng cho robot. Robot được đặt trong một hệ quy chiếu quán tính  $Oxy$  đặt tại một vị trí cố định trong không gian hoạt động của robot. Ngoài ra, khi robot di chuyển, không chỉ vị trí của robot thay đổi mà hướng của



Hình 3.1: Mô hình robot di động hai bánh chủ động.

robot cũng thay đổi. Để xác định hướng của robot, chúng ta quan tâm đến hệ tọa độ di động  $O'x'y'$  với  $O'$  gắn tại tâm của robot. Khi đó, tư thế của robot được cụ thể hóa bởi vectơ  $\hat{\mathbf{q}} = [x \ y \ \varphi]^T$ . Trong đó,  $\mathbf{q} = [x \ y]^T$  là tọa độ trọng tâm của robot trong hệ tọa độ  $Oxy$ , còn  $\varphi$  là góc giữa trục  $Ox$  và trục  $O'x'$  của hai hệ quy chiếu. Chuyển động của robot được chia làm hai dạng chuyển động: chuyển động tịnh tiến và chuyển động quay. Chuyển động của robot được đặc trưng bởi vectơ  $\mathbf{v} = [v \ \omega]^T$ , trong đó  $v$  và  $\omega$  lần lượt là vận tốc dài (tịnh tiến) và vận tốc quay gắn với hệ tọa độ di động của robot có phương và chiều như trên Hình 3.1. Gọi  $\hat{\mathbf{q}}_s = [x_s \ y_s \ \varphi_s]^T$  và  $\hat{\mathbf{q}}_g = [x_g \ y_g \ \varphi_g]^T$  lần lượt là tư thế điểm đầu và điểm đích của robot. Gọi  $\hat{\mathbf{q}}(t) = [x(t) \ y(t) \ \varphi(t)]^T$  và  $\mathbf{v}(t), t \in [1, T]$  lần lượt là tư thế và vận tốc của robot tại thời gian  $t$ , với  $T$  là khoảng thời gian di chuyển từ điểm ban đầu đến điểm đích của robot. Theo đó, mối liên hệ giữa vận tốc trong hệ tọa độ cố định  $\dot{\mathbf{p}}(t)$  với vận tốc trong hệ tọa độ di động  $\mathbf{v}(t)$  được biểu diễn như sau [89]

$$\dot{\mathbf{p}} = \mathbf{S}(\mathbf{p})\mathbf{v}(t), \quad (3.17)$$

trong đó,  $\mathbf{S}(\mathbf{p})$  là một ma trận Jacobian chuyển hệ trục tọa độ. Xét mối quan hệ hình học trên Hình 3.1, ta có:

$$\mathbf{S}(\mathbf{p}) = \begin{bmatrix} \cos \varphi & 0 \\ \sin \varphi & 0 \\ 0 & 1 \end{bmatrix} \quad (3.18)$$



Theo đó, phương trình động học của robot có dạng [89]

$$\dot{\mathbf{p}} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\varphi} \end{bmatrix} = \begin{bmatrix} \cos \varphi & 0 \\ \sin \varphi & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \omega \end{bmatrix}, \quad (3.19)$$

trong đó,  $v \leq v_{\max}$  và  $\omega \leq \omega_{\max}$ , với  $v_{\max}$  và  $\omega_{\max}$  lần lượt là vận tốc dài và vận tốc góc lớn nhất của robot. Khi đó, (3.19) chính là phương trình động học của robot di động hai bánh chủ động đang xét.

Quy đạo của robot được kỳ vọng rằng sẽ di chuyển đến đích với quãng đường di chuyển nhỏ nhất. Khi đó, để đơn giản cho việc khảo sát, bài toán được rời rạc hóa theo thời gian. Robot sẽ tính toán nhiều lần, mỗi lần sẽ di chuyển  $h$  bước về phía đích. Do đó, hàm mục tiêu có dạng

$$J = w_1 \sum_{k=1}^{h-1} \|\mathbf{q}(k+1) - \mathbf{q}(k)\|^2 + w_2 \|\mathbf{q}(h) - \mathbf{q}_g\|^2, \quad (3.20)$$

trong đó,  $\mathbf{q}_g$  là vị trí điểm đích đến của robot,  $w_1$  và  $w_2$  là trọng số của hàm mục tiêu. Đặt  $\mathbf{Q} = [\mathbf{q}(1), \mathbf{q}(2), \dots, \mathbf{q}(h)]$  là quỹ đạo di chuyển của robot trong  $h$  bước.

Giả sử robot làm việc trong môi trường công nghiệp, trong đó số lượng các chướng ngại vật được biểu diễn bởi một tập  $\mathcal{M}$ . Không mất tính tổng quát, giả sử các chướng ngại vật đều có dạng hình tròn. Khi đó, các ràng buộc đảm bảo tránh va chạm giữa robot và chướng ngại vật được biểu diễn dưới dạng

$$\|\mathbf{q}(k) - \mathbf{q}_{o_i}^O\| \geq r_{o_i}^O + d_{\max}^r + \delta, \forall k \in [1, h], o_i \in \mathcal{M}. \quad (3.21)$$

trong đó,  $\mathbf{q}_{o_i}^O = (x_{o_i}^O, y_{o_i}^O)$  và  $r_{o_i}^O$  là vị trí và bán kính của chướng ngại vật  $o_i$ ,  $d_{\max}^r$  là tham số liên quan đến kích thước của robot, và  $\delta$  là một giá trị khoảng cách an toàn giữa robot và chướng ngại vật. Đặt  $d_{r,o_i} = r_{o_i}^O + d_{\max}^r + \delta, \forall k \in [1, h], o_i \in \mathcal{M}$ . Khi đó, ràng buộc (3.21) có thể viết lại thành

$$\|\mathbf{q}(k) - \mathbf{q}_{o_i}^O\| \geq d_{r,o_i}, \forall k \in [1, h], o_i \in \mathcal{M}. \quad (3.22)$$

Ngoài ra, trong quá trình làm nhiệm vụ, các robot di động cần thu thập dữ liệu gửi về trạm trung tâm hoặc nhận tín hiệu điều khiển từ trạm trung tâm. Khi đó,

tín hiệu truyền thông không dây giữa robot với trạm chủ cần phải được đảm bảo. Gọi  $R(k)$  là tốc độ truyền thông từ trạm trung tâm đến robot di động tại thời điểm  $k$ ,  $\bar{\gamma}$  là tốc độ truyền thông tối thiểu mà hệ thống yêu cầu để đảm bảo thông tin liên tục được cập nhật. Khi đó, ràng buộc sau cần được thỏa mãn

$$R(k) \geq \bar{\gamma}, \forall k \in [1, h]. \quad (3.23)$$

Tuy nhiên, ở chương này, NCS giả sử ràng buộc truyền thông đã được giải quyết bởi mô hình mạng được trình bày ở Chương 2. Khi đó, ràng buộc truyền thông được bỏ qua.

### 3.2.2 Thuật toán tối ưu quỹ đạo

Như vậy bài toán tối ưu có dạng

$$\min_{\mathbf{q}} J \quad (3.24a)$$

$$\text{thỏa mãn: } \mathbf{q}(1) = \mathbf{q}_s, \quad (3.24b)$$

$$\|\mathbf{q}(k+1) - \mathbf{q}(k)\| \leq \tau v_{\max}, \forall k \in [1, h], \quad (3.24c)$$

$$\|\mathbf{q}(k) - \mathbf{q}_{o_i}^O\| \geq d_{r,o_i}^r, \forall k \in [1, h], o_i \in \mathcal{M}, \quad (3.24d)$$

trong đó,  $\mathbf{q}_s$  là vị trí điểm bắt đầu của robot, ràng buộc (3.24b) là điểm bắt đầu SP (Start Point), ràng buộc (3.24c) là giới hạn vận tốc của robot, và (3.24d) là ràng buộc tránh vật cản. Bài toán này là bài toán không lồi bởi vì ràng buộc (3.24d) là không lồi. Do đó, thuật toán xấp xỉ lồi được trình bày dưới đây để xấp xỉ bài toán (3.24) trở thành bài toán lồi để giải.

Ràng buộc (3.24d) tương đương với

$$(x(k) - x_{o_i}^O)^2 + (y(k) - y_{o_i}^O)^2 \geq d_{r,o_i}^2, \forall k \in [1, h], o_i \in \mathcal{M}. \quad (3.25)$$

Đặt vế phải của (3.25) là  $f(\mathbf{q}(k))$ . Có thể thấy rằng  $f(k)$  là một hàm lồi. Chính vì hàm  $f(\mathbf{q}(k))$  là hàm lồi, nên đạo hàm của nó, tức là,  $\nabla f(\mathbf{q}^{(\kappa)}(k))$  là một siêu đạo hàm của nó [90], nên

$$f(\mathbf{q}(k)) \geq f(\mathbf{q}_s) + \nabla f(\mathbf{q}_s)(\mathbf{q}(k) - \mathbf{q}_s)$$

$$\begin{aligned}
&= (x_s - x_{o_i}^O)^2 + (y_s - y_{o_i}^O)^2 + 2(x_s - x_{o_i}^O)(x(k) - x_s) + 2(y_s - y_{o_i}^O)(y(k) - y_s) \\
&\triangleq f_s(\mathbf{q}(k))
\end{aligned} \tag{3.26}$$

Dựa vào (3.25) và (3.27), ràng buộc không lồi (3.24d) được xấp xỉ bởi ràng buộc lồi sau

$$f_s(\mathbf{q}(k)) \geq d_{r,o_i}^2. \tag{3.27}$$

Từ (3.27), thay vì giải bài toán (3.24), chúng ta sẽ giải bài toán lồi sau

$$\min_{\mathbf{q}} J \tag{3.28a}$$

$$\text{thoả mãn: (3.24b), (3.24c) và (3.27)}. \tag{3.28b}$$

Thuật toán 4 mô tả khái quát các bước giải bài toán tối ưu lồi (3.28) để tìm được quỹ đạo tối ưu cho robot.

---

#### Thuật toán 4 Thuật toán CAA

---

**Khởi tạo:** Đặt giá trị cho  $\mathbf{q}_s, \mathbf{q}_g, h, v_{\max}, \mathbf{q}_{o_i}^O, r_{o_i}^O, d_{\max}^r$  với  $o_i \in \mathcal{M}$ . Đặt sai số vị trí đích đến  $\epsilon_f = 1, \epsilon = 10^{-3}$ .

1: **repeat**

2: Giải bài toán tối ưu lồi (3.28) để thu được quỹ đạo tối ưu  $(\mathbf{Q}^*, \tau^*)$

3: Đặt  $\mathbf{q}_s \leftarrow \mathbf{q}_h$

4:  $\epsilon_f = \|\mathbf{q}_s - \mathbf{q}_g\|$

5: **until**  $\epsilon_f \leq \epsilon$

---

Nhằm mục đích so sánh, luận án trình bày thêm một thuật toán tính toán quỹ đạo cho robot dựa trên mô hình động học của robot, thuật toán này được đặt tên là CCA (Conventional Computing Algorithm). Trong thuật toán này, vận tốc của robot được giả sử là một hằng số và đúng bằng giá trị vận tốc lớn nhất của robot  $v_{\max}$ . Giả sử hướng của robot ban đầu là hướng về vị trí đích, tức là

$$\varphi = \arctan \frac{y_g - y_s}{x_g - x_s}. \tag{3.29}$$

Từ phương trình động học (3.19), vị trí tiếp theo của robot được tính toán dựa

---

**Thuật toán 5** Thuật toán CCA
 

---

**Khởi tạo:** Đặt các giá trị khởi tạo cho  $\mathbf{q}_s, \mathbf{q}_g, h, v_{\max}, \mathbf{q}_{o_i}^O, r_{o_i}^O, d_{\max}^r$  với  $o_i \in \mathcal{M}$ . Cài đặt  $\epsilon_f = 1, \epsilon = 10^{-3}$ .

```

1: repeat
2:   Tính toán  $\mathbf{Q}$  bởi (3.29) và (3.30)
3:   if tính toán có xảy ra va chạm then
4:     Tính lại  $\mathbf{Q}$  bởi (3.31)
5:   end if
6:    $\epsilon_f = \|\mathbf{q}_s - \mathbf{q}_g\|$ 
7: until  $\epsilon_f \leq \epsilon$ 

```

---

trên

$$\begin{cases} \dot{x} = v \cos \varphi \\ \dot{y} = v \sin \varphi \\ \dot{\varphi} = \omega \end{cases} \quad (3.30)$$

Sau quá trình tính toán, robot sẽ kiểm tra xem vị trí này có bị va chạm với các chướng ngại vật hay không. Nếu kiểm tra chỉ ra rằng không có va chạm nào xảy ra thì đây chính là vị trí tiếp theo của robot. Nếu tính toán chỉ ra có xảy ra va chạm thì robot cần tính toán lại vị trí này như sau [91]

$$\begin{cases} x_{new}(t+1) = x(t+1) + (d_{r,o_i} - d_{Obs}) \cos \alpha, \\ y_{new}(t+1) = y(t+1) + (d_{r,o_i} - d_{Obs}) \sin \alpha, \end{cases} \quad (3.31)$$

trong đó,  $d_{Obs}$  và  $\alpha$  là khoảng cách và góc giữa điểm chính giữa của vật cản  $m$  và điểm  $\mathbf{q}(t+1)$ . Thuật toán 5 mô tả thuật toán CCA.

### 3.2.3 Đánh giá kết quả thuật toán CAA

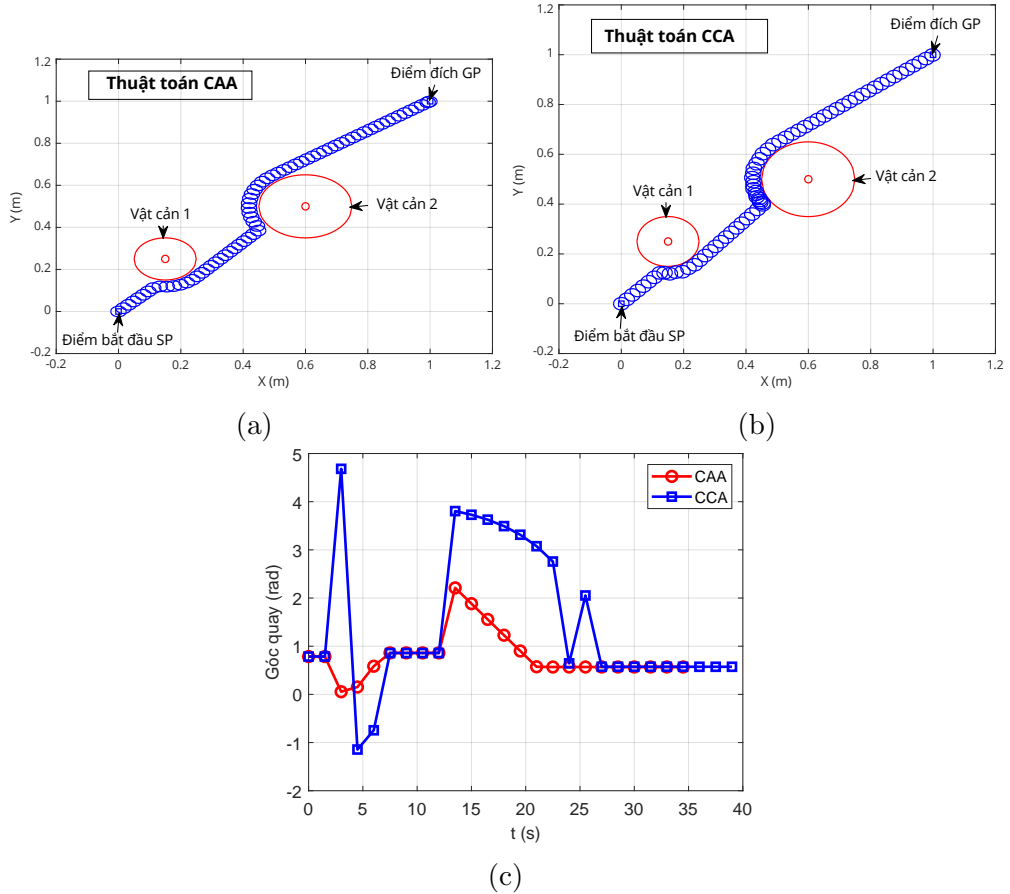
Trong phần này, các kết quả mô phỏng kiểm chứng hiệu quả của thuật toán đề xuất được trình bày. Môi trường mô phỏng được xem xét gồm có 2 chướng ngại vật và số bước di chuyển của robot mỗi lần là  $h = 10$ . Ngoài ra, các tham số mô phỏng được liệt kê ở Bảng 3.1. Các kết quả mô phỏng được thực hiện trên MATLAB, và bài toán tối ưu lồi được giải bằng gói CVX cùng với trình giải Mosek 9.1.9.

Hình 3.2 và Hình 3.3 thực hiện so sánh các kết quả thu được bởi hai thuật toán CAA (Thuật toán 4) và CCA (Thuật toán 5).

Với thuật toán xấp xỉ lồi CAA, tổng quãng đường di chuyển của robot là 1.5625m trong khoảng thời gian là 35.5s. Trong quá trình di chuyển, robot đi với vận tốc luôn

Bảng 3.1: Bảng tham số mô phỏng thuật toán CAA và CCA

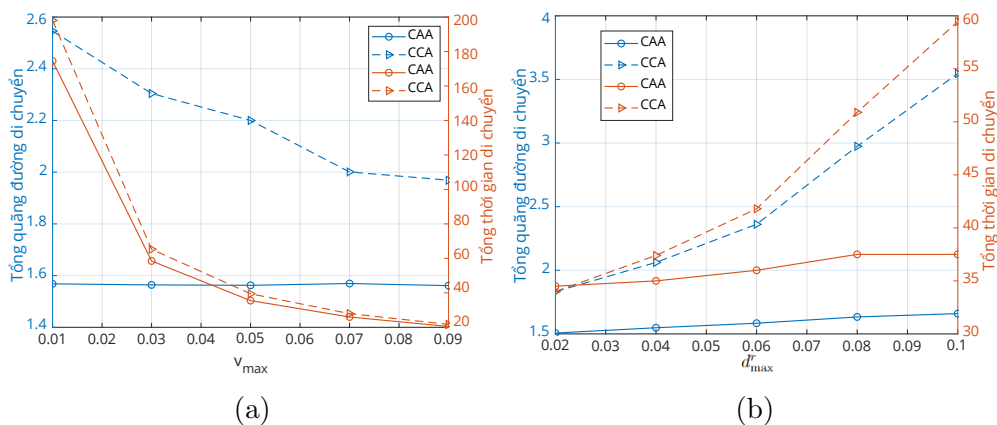
Tham số	Giá trị	Tham số	Giá trị
$d_{\max}^r$	0.05 m	$\delta$	0.005 m
$v_{\max}$	0.05 m/s	$\mathbf{q}_s$	[0, 0]
$\tau$	0.05 s	$\mathbf{q}_g$	[1, 1]
$w_1$	0.8	$\mathbf{q}_{o1}^O$	[0.15, 0.25]
$w_2$	0.2	$\mathbf{q}_{o2}^O$	[0.6, 0.5]



Hình 3.2: (a) Quỹ đạo của robot với thuật toán CAA, (b) Quỹ đạo của robot với thuật toán CCA, và (c) Góc quay của robot.

nhỏ hơn vận tốc  $v_{\max}$ . Nhìn vào Hình 3.2(a) ta cũng thấy rằng robot di chuyển khá mượt mà trong suốt quãng đường đi từ điểm bắt đầu cho đến điểm đích. Đối với thuật toán CCA, quãng đường di chuyển của robot là 2.2m với thời gian là 39.45s. Trong quá trình di chuyển, khi không có vật cản robot đi với vận tốc lớn nhất  $v_{\max}$ . Nhìn vào hình 3.2(b) ta cũng thấy rằng robot di chuyển khá khó khăn ở những vị trí gần vật cản. Hình 3.2(c) thể hiện độ mượt mà khi di chuyển tương ứng với hai thuật toán CAA và CCA. Cụ thể, thuật toán CAA thiết kế quỹ đạo giúp robot di chuyển mượt mà hơn với sự thay đổi góc quay nhỏ hơn ở mỗi bước di chuyển theo thời gian so với quỹ đạo được thiết kế bởi thuật toán CCA.

Tiếp theo, tham số khoảng cách giữa hai bánh xe robot (tương ứng là kích thước



Hình 3.3: Quãng đường và thời gian di chuyển khi a)  $v_{\max}$  thay đổi b)  $d_{\max}^r$  thay đổi.

của robot cũng thay đổi) thay đổi và quan sát sự thay đổi của quỹ đạo cũng như thời gian di chuyển của robot, đây là hai chỉ số quan trọng đánh giá chất lượng của quỹ đạo di chuyển của robot. Nhìn vào Hình 3.3(a), chúng ta thấy rằng, nhìn chung, cả độ dài quãng đường và thời gian di chuyển đối với cả hai thuật toán CAA và CCA đều tăng khi  $d$  tăng. Tuy nhiên, cả hai giá trị này ở thuật toán CAA đều tăng nhẹ còn ở thuật toán CCA thì tăng khá đáng kể. Có thể nói, quỹ đạo chuyển động của robot AGV khi sử dụng thuật toán CAA không bị ảnh hưởng quá nhiều vào kích thước của robot trong khi nó bị ảnh hưởng rất lớn khi sử dụng thuật toán CCA. Ngoài ra, rất dễ dàng nhận thấy rằng, cả về độ dài quãng đường hay thời gian di chuyển ở thuật toán CAA đều nhỏ hơn rất nhiều so với thuật toán CCA.

Cuối cùng, chúng ta xem xét việc thay đổi hai chỉ số quan trọng của bài toán tối ưu quỹ đạo chuyển động khi thay đổi vận tốc lớn nhất của robot. Nhìn vào hình 3.3(b), chúng ta thấy đối với thuật toán CCA, khi vận tốc cho phép lớn nhất của robot  $v_{\max}$  tăng lên thì cả quãng đường và thời gian di chuyển đều có xu hướng giảm nhưng vẫn còn lớn hơn so với thuật toán CAA. Trong khi đó, nhìn chung quỹ đạo chuyển động của robot ở thuật toán CAA khá ổn định khi  $v_{\max}$  tăng, còn thời gian di chuyển giảm xuống khá đáng kể.

Như vậy, có thể thấy rằng, xét về một bài toán tối ưu quãng đường di chuyển, hay xét đến thời gian di chuyển của robot hoặc xét đến độ mượt mà khi di chuyển, xét về mặt nào thì thuật toán CAA vẫn cho một quỹ đạo di chuyển tốt.

### 3.3 Phát triển thuật toán tối thiểu năng lượng tiêu thụ

Ở phần này, luận án quan tâm đến hàm mục tiêu là năng lượng tiêu thụ của robot. Tức là xem xét đến bài toán tối ưu quỹ đạo cho robot với mục tiêu là tối thiểu

năng lượng tiêu thụ cho robot.

### 3.3.1 Mô hình hóa năng lượng tiêu thụ của robot di động

Như đã đề cập ở Chương 1, do đặc trưng về chức năng di chuyển nên robot di động sử dụng các nguồn pin rời. Nhìn chung, năng lượng tiêu thụ bởi robot chia thành hai phần chính: năng lượng chuyển đổi sang dạng động học và năng lượng vượt qua lực cản tác động vào robot, như lực ma sát [41]. Ở phần này, luận án bỏ qua phần mất mát trong quá trình biến đổi động học quay của robot, mất mát do kết cấu cơ khí, và mất mát do tổn hao nhiệt ở cuộn dây phần ứng của động cơ như trong nghiên cứu [92]. Khi đó, mất mát năng lượng do biến đổi động học được biểu diễn như sau

$$\begin{aligned} E_k &= \frac{1}{2}mv^2(t) \\ &= \int_{t=0}^T \left( d\left(\frac{1}{2}mv^2(t)\right) \right), \end{aligned} \quad (3.32)$$

trong đó  $m$  là khối lượng của robot. Năng lượng mất mát do ma sát trong quá trình robot di chuyển được mô tả dưới dạng [93]

$$E_{friction} = 2\mu mg \int_{t=0}^T |v(t)|dt, \quad (3.33)$$

trong đó,  $g$  là gia tốc trọng trường, và  $\mu$  là hệ số ma sát trượt phụ thuộc vào dạng bề mặt môi trường robot làm việc. Do đó, năng lượng tiêu thụ cho hệ thống chuyển động của robot được mô tả dạng [41]

$$\begin{aligned} E_{motion} &= E_k + E_{friction} \\ &= \int_{t=0}^T \left( d\left(\frac{1}{2}mv^2(t)\right) \right) + 2\mu mg \int_{t=0}^T |v(t)|dt. \end{aligned} \quad (3.34)$$

Ngoài ra, robot còn tiêu hao một phần không nhỏ năng lượng cho máy tính, các mạch điện tử, cảm biến. Các thành phần này được mô tả bởi [41, 93]

$$E_s = P_s \int_{t=0}^T dt, \quad (3.35)$$

trong đó,  $P_s$  là công suất nguồn của hệ thống.

Từ (3.34) và (3.35) cùng với việc rời rạc hóa theo thời gian, tổng năng lượng

tiêu thụ của robot có thể được viết lại dưới dạng sau:

$$\begin{aligned}
E &= E_{motion} + E_s = E_k + E_{friction} + E_s \\
&= \frac{1}{2} \sum_{l=0}^{L-1} mv_l^2 + 2\mu mg \sum_{l=0}^{L-1} |\mathbf{q}_{l+1} - \mathbf{q}_l| + P_s \sum_{l=0}^{L-1} \tau, \\
&= \frac{1}{2} \sum_{l=0}^{L-1} m \left( \frac{|\mathbf{q}_{l+1} - \mathbf{q}_l|}{\tau} \right)^2 + 2\mu mg \sum_{l=0}^{L-1} |\mathbf{q}_{l+1} - \mathbf{q}_l| + P_s(L-1)\tau, \tag{3.36}
\end{aligned}$$

trong đó,  $L$  là số bước rời rạc hóa theo yêu cầu,  $l \in [0, L]$  là các bước theo thời gian,  $\mathbf{Q} = [\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_L]$ , và  $\mathbf{v} = [v_1, v_2, \dots, v_L]$  lần lượt là quỹ đạo và vận tốc tương ứng với các bước thời gian trên, và  $\tau$  là khoảng thời gian cho một bước.

### 3.3.2 Các ràng buộc tránh va chạm

Trong môi trường làm việc của robot, chúng ta giả sử các chướng ngại vật có dạng đa giác  $H$  cạnh. Cụ thể, giả sử có một tập  $\mathcal{M}$  vật cản trong môi trường làm việc của robot, mỗi vật cản được mô tả bởi [94]

$$\mathcal{P} = \{\mathbf{z} | \mathbf{A}\mathbf{z} < \mathbf{b}\}, \tag{3.37}$$

trong đó,  $\mathbf{A} \in \mathbb{R}^{H \times 2}$  và  $\mathbf{b} \in \mathbb{R}^H$ . Một điểm  $\mathbf{z}$  được gọi là nằm ngoài hình  $\mathcal{P}$  khi ít nhất một trong các bất phương trình  $\mathbf{A}\mathbf{q} \geq \mathbf{b}$  thỏa mãn, tức là,

$$\mathbf{A}\mathbf{z} \geq \mathbf{b} + (\mathbf{e} - \mathbf{1})\hat{M}, \tag{3.38}$$

$$\sum_{h=1}^H e_h = 1, \tag{3.39}$$

trong đó,  $\mathbf{e} = (e_1, e_2, \dots, e_H)^T$  là một véc tơ nhị phân,  $e_h \in \{0, 1\}$ , và  $\mathbf{1} = \mathbf{1}^H$  là véc tơ đơn vị,  $\hat{M}$  là một hằng số đủ lớn được sử dụng theo phương pháp Big-M để xác định một điểm nằm ngoài một đa giác [95]. Các ràng buộc trong (3.38) và (3.39) đảm bảo rằng có ít nhất một thành phần của  $\mathbf{e}$  bằng 1 và điểm  $\mathbf{z}$  nằm ngoài vật cản. Gọi  $d_r$  là tham số đại diện cho kích thước của robot. Để tránh va chạm xảy ra giữa robot và vật cản, chúng ta định nghĩa một khoảng cách an toàn  $\delta > 0$ . Do đó, quỹ đạo của robot cần đảm bảo các ràng buộc sau

$$\mathbf{A}_{o_i}\mathbf{q}_l \geq \mathbf{b}_{o_i} + d_r + \delta + (\mathbf{e}_{o_i} - \mathbf{1})\hat{M}, \quad \forall l \in [0, L], \forall o_i \in \mathcal{M}, \tag{3.40}$$



$$\mathbf{E} \in \{0, 1\}^{M \times H \times (D+1)}, \quad (3.41)$$

$$\sum_{h=1}^H e_{o_i, h, l} = 1, \forall o_i \in \mathcal{M}, l \in [0, L], \quad (3.42)$$

trong đó,  $\mathbf{e}_{o_i, h, l} \in \{0, 1\}^H$  là một véc tơ nhị phân cho vật cản  $o_i$  và  $\mathbf{E} = \{\mathbf{e}_{o_i}, \forall o_i \in \mathcal{M}\}$  là dành cho toàn bộ các vật cản tĩnh trong môi trường làm việc của robot.

### 3.3.3 Thuật toán tối ưu năng lượng tiêu thụ cho robot

Nhìn chung, chúng ta có thể thiết kế quỹ đạo cho robot giúp giảm tổng quãng đường di chuyển nhằm giảm năng lượng động học tiêu hao cho robot. Tuy nhiên, cần đảm bảo không xảy ra va chạm trong quá trình robot di chuyển. Thêm vào đó, chúng ta có thể giảm khoảng thời gian mỗi bước di chuyển để giảm năng lượng tiêu hao trên máy tính, các mạch điện tử và cảm biến. Tuy nhiên, điều này yêu cầu robot cần di chuyển nhanh hơn và có thể vượt ra khỏi khả năng của robot. Do đó, trong phần này, luận án xây dựng bài toán tối thiểu năng lượng tiêu thụ của robot bằng cách tối ưu quỹ đạo và thời gian mỗi bước di chuyển trong khi vẫn đảm bảo yêu cầu tránh vật cản. Bài toán tối ưu được đưa về dạng bài toán tối ưu có biến số nguyên như sau:

$$\min_{\mathbf{Q}, \tau, \mathbf{e}} E \quad (3.43a)$$

thoả mãn: (3.40) - (3.42),

$$\mathbf{q}_0 = \mathbf{q}_s, \quad (3.43b)$$

$$\mathbf{q}_D = \mathbf{q}_g, \quad (3.43c)$$

$$v_d \leq v_{\max}, \forall d \in \{1, 2, \dots, D\}, \quad (3.43d)$$

$$\tau_{\min} \leq \tau \leq \tau_{\max}, \quad (3.43e)$$

trong đó  $v_{\max}$  là vận tốc dài lớn nhất của robot,  $\tau_{\min}$  và  $\tau_{\max}$  lần lượt là thời gian tối thiểu và thời gian tối đa cho một bước di chuyển.

Bài toán (3.43) là một bài toán không lồi do hàm mục tiêu (3.43a) là một hàm không lồi. Tiếp theo, luận án sẽ trình bày việc phát triển thuật toán xấp xỉ lồi [96] để giải bài toán tối ưu trên.

Đặt  $(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)}, \mathbf{e}^{(\kappa)})$  là một điểm khả thi của bài toán (3.43). Điểm khả thi này được tìm thấy ở lần lặp thứ  $\langle \kappa - 1 \rangle$ . Tại lần lặp thứ  $\kappa$ , điểm khả thi tiếp theo, tức là,  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}, \mathbf{e}^{(\kappa+1)})$  sẽ được xác định. Áp dụng bất đẳng thức (A.2) ở Phụ lục A

cho hàm  $E_k$ , ta có

$$\begin{aligned} E_k &\leq \frac{1}{2}m \sum_{l=0}^{L-1} \left( \frac{|\mathbf{q}_{l+1}^{(\kappa)} - \mathbf{q}_l^{(\kappa)}|}{4\tau^{(\kappa)}} \left( \frac{|\mathbf{q}_{l+1} - \mathbf{q}_l|}{|\mathbf{q}_{l+1}^{(\kappa)} - \mathbf{q}_l^{(\kappa)}|} + \frac{\tau^{(\kappa)}}{\tau} \right)^2 \right)^2 \\ &= \frac{m}{32} \sum_{l=0}^{L-1} \frac{|\mathbf{q}_{l+1}^{(\kappa)} - \mathbf{q}_l^{(\kappa)}|^2}{(\tau^{(\kappa)})^2} \left( \frac{|\mathbf{q}_{l+1} - \mathbf{q}_l|}{|\mathbf{q}_{l+1}^{(\kappa)} - \mathbf{q}_l^{(\kappa)}|} + \frac{\tau^{(\kappa)}}{\tau} \right)^4 \triangleq E_k^{(\kappa)}. \end{aligned} \quad (3.44)$$

Nhận thấy rằng, biểu thức trong ngoặc lớn của  $E_k^{(\kappa)}$  là tổng của một hàm chuẩn, tức là,  $\frac{|\mathbf{q}_{l+1} - \mathbf{q}_l|}{|\mathbf{q}_{l+1}^{(\kappa)} - \mathbf{q}_l^{(\kappa)}|}$  và một hàm  $\frac{\tau^{(\kappa)}}{\tau}$ . Bằng các tính chất của hàm lồi đã đề cập đến ở phần phía trên, có thể dễ dàng chứng minh được các hàm này đều là hàm lồi. Do đó, hàm  $E_k^{(\kappa)}$  là một hàm lồi. Hơn nữa, hàm  $E_{friction}$  đã là một hàm lồi. Do đó, có thể xấp xỉ hàm mục tiêu (3.43a) bằng hàm lồi sau

$$E \leq E_k^{(\kappa)} + E_{friction} + E_s \triangleq E^{(\kappa)}. \quad (3.45)$$

Dựa trên (3.45), bài toán (3.43) được xấp xỉ bằng bài toán tối ưu lồi sau

$$\min_{\mathbf{Q}, \mathbf{t}, \mathbf{e}} E^{(\kappa)} \quad (3.46a)$$

$$\text{thoả mãn: (3.40) - (3.42), và (3.43b) - (3.43e),} \quad (3.46b)$$

Giải bài toán (3.46) sẽ thu được  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$  cho lần lặp tiếp theo. Độ phức tạp tính toán của thuật toán giải bài toán tối ưu lồi (3.46) là  $\mathcal{O}(\alpha^2 \beta^{2.5} + \beta^{3.5})$  [97], trong đó  $\alpha = 2L + KH(L + 1) + 3$  và  $\beta = 2(L + 1)KH + (L + 1)K + L + 3$  lần lượt là số lượng các biến quyết định và số lượng các ràng buộc lồi của bài toán (3.46). Chú ý rằng,  $(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)})$  và  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$  lần lượt là điểm khả thi và điểm tối ưu của bài toán (3.46). Khi đó,

$$E^{(\kappa)}(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}) > E^{(\kappa)}(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)}), \quad (3.47)$$

với mọi  $(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)}) \neq (\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$ . Điều này đơn giản thể hiện  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$  là một điểm tốt hơn điểm  $(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)})$  cho bài toán (3.43). Quan tâm đến (3.43a), ta có [98]:

$$\begin{aligned} E(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)}) &= E^{(\kappa)}(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)}) \\ &< E^{(\kappa)}(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}) \leq E(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}). \end{aligned} \quad (3.48)$$

---

**Thuật toán 6** Thuật toán OSOW giải bài toán (3.43)
 

---

**Khởi tạo:** Đặt  $\epsilon = 10^{-3}$  và  $\kappa = 1$ ;  
 1: **repeat**  
 2:   Giải bài toán lồi (3.46) để thu được  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}) = (\mathbf{Q}^*, \bar{\tau}^*)$ ;  
 3:   Cập nhật  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}) := (\mathbf{Q}^*, \tau^*)$   
 4:    $\hat{\epsilon} = \frac{|E^{(\kappa+1)} - E^{(\kappa)}|}{E^{(\kappa)}}$   
 5:   Đặt  $\kappa \leftarrow \kappa + 1$   
 6: **until**  $\hat{\epsilon} \leq \epsilon$ .  
 7: Điểm tối ưu thu được  $(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)})$

---

Điều này rõ ràng cho thấy điểm tối ưu  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$  của bài toán (3.43) đảm bảo điều kiện hội tụ:  $E(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)}) > E(\mathbf{Q}^{(\kappa)}, \tau^{(\kappa)})$ . Điểm  $(\mathbf{Q}^{(\kappa+1)}, \tau^{(\kappa+1)})$  hội tụ đến một điểm yên ngựa  $(\tilde{\mathbf{Q}}, \tilde{\tau})$  sau một số lần lặp giới hạn [98]. Thuật toán giải bài toán tối ưu (3.43) được chỉ ra trong Thuật toán 6. Thuật toán này được đặt tên là Thuật toán OSOW để chỉ thuật toán giải quyết bài toán tối ưu quỹ đạo robot trong môi trường chỉ có các vật cản tĩnh (Only Static Obstacles in the robot's Workspace).

### 3.3.4 Bài toán tối ưu trong kịch bản môi trường có vật cản động

Trong phần 3.3.3, chúng ta quan tâm đến kịch bản môi trường làm việc của robot chỉ có vật cản tĩnh SOs (Static Obstacles). Tuy nhiên, trong thực tế, môi trường làm việc của robot không chỉ có các vật cản tĩnh, mà còn có các vật cản có thể di chuyển được (vật cản động) như con người, các xe hàng... Do đó, trong phần này, luận án sẽ quan tâm đến kịch bản vật cản động xuất hiện trong quá trình robot di chuyển. Cụ thể, đối với vật cản tĩnh, các giả sử vẫn giữ nguyên như trong phần 3.3.2. Trước tiên, dựa vào thông tin các vật cản tĩnh, robot sẽ tính toán quỹ đạo tối ưu dựa trên Thuật toán 6, tức là Thuật toán OSOW giải bài toán (3.43). Sau đó, robot sẽ di chuyển dựa trên quỹ đạo này. Quỹ đạo này, tức là,  $\mathbf{Q} = [\mathbf{q}_0, \mathbf{q}_1, \dots, \mathbf{q}_L]$ , được gọi là quỹ đạo toàn cục. Trong quá trình robot di chuyển, tại bước  $i \in [0, L]$ , giả sử robot phát hiện có  $\mathcal{N}_i = \{1, \dots, N_i\}$  vật cản động MOs (moving obstacles). Mỗi vật cản động MO được giả sử có dạng hình tròn. Cụ thể, robot quan sát thấy được MO  $n_i$  có tâm tại  $\mathbf{q}_{n,i}^O = (x_{n,i}^O, y_{n,i}^O)$  di chuyển với vận tốc  $\mathbf{v}_{n_i} = (v_{nx_i}, v_{ny_i})$ . Do yêu cầu tránh vật cản, quỹ đạo di chuyển tiếp theo của robot  $\mathbf{q}_{i,l_i} = (x_{i,l_i}, y_{i,l_i})$ , trong đó  $l_i \in [0, L_i]$  với  $L_i$  là số bước di chuyển cần thiết để đi qua  $N_i$  vật cản. Gọi  $\mathbf{Q}_i = [\mathbf{q}_{i,0}, \mathbf{q}_{i,1}, \dots, \mathbf{q}_{i,L_i}]$ , và  $\mathbf{v}_i = [v_{i,1}, v_{i,2}, \dots, v_{i,L_i}]$  là tập các vị trí và vận tốc cho robot trong quá trình tránh vật cản. Theo đó,  $\mathbf{Q}_i$  được gọi là quỹ đạo cục bộ. Như vậy, điều kiện tránh vật cản lúc này

được mô tả như sau

$$\|\mathbf{q}_{i,l_i} - \mathbf{q}_{n,l_i}^O\| \geq r_{n,i}^O + d_r + \delta, \forall n \in \mathcal{N}, l_i \in [0, L_i], \quad (3.49)$$

trong đó,  $r_{n,i}^O$  là bán kính của vật cản  $n$ , và  $\mathbf{q}_{n,l_i}^O = (x_{n,l_i}^O, y_{n,l_i}^O)$  là vị trí của vật cản  $n$  tại bước  $l_i$ . Đặt  $d_{i,r,n} = r_{n,i}^O + d_r + \delta$ , ràng buộc (3.49) có thể được viết lại thành

$$\|\mathbf{q}_{i,l_i} - \mathbf{q}_{n,l_i}^O\| \geq d_{i,r,n}, \forall n \in \mathcal{N}, l_i \in [0, L_i]. \quad (3.50)$$

Bài toán tối thiểu năng lượng tiêu thụ của robot trong kịch bản tránh vật cản động sẽ có thêm các ràng buộc tránh vật cản động (3.49) như sau:

$$\min_{\mathbf{Q}_i, \mathbf{t}_i, \mathbf{e}} E_i \quad (3.51a)$$

thoả mãn: (3.40) - (3.42), và (3.50),

$$\mathbf{q}_{i,0} = \mathbf{q}_{i,s}, \quad (3.51b)$$

$$\mathbf{q}_{i,L_i} = \mathbf{q}_{i,g}, \quad (3.51c)$$

$$v_{i,l_i} \leq v_{\max}, \forall l_i \in [0, L_i], \quad (3.51d)$$

$$\tau_{\min} \leq \tau_i \leq \tau_{\max}, \quad (3.51e)$$

trong đó,  $\tau_i$  là bước thời gian,  $\mathbf{q}_{i,s}, \mathbf{q}_{i,g}$  lần lượt là vị trí robot phát hiện các vật cản động và vị trí robot vượt qua được các vật cản động và trở về quỹ đạo toàn cục  $\mathbf{Q}$ , và

$$\begin{aligned} E_i &= E_{k_i} + E_{friction_i} + E_{s_i} \\ &= \frac{1}{2} \sum_{l_i=0}^{L_i-1} m \left( \frac{|\mathbf{q}_{i,l_i+1} - \mathbf{q}_{i,l_i}|}{\tau_i} \right)^2 + 2\mu mg \sum_{l_i=0}^{L_i-1} |\mathbf{q}_{i,l_i+1} - \mathbf{q}_{i,l_i}| + P_s(L_i - 1)\tau_i \end{aligned} \quad (3.52)$$

là tổng năng lượng tiêu thụ để tránh được  $N_i$  MOs.

Bài toán (3.51) là bài toán không lồi vì hàm mục tiêu (3.51a) và ràng buộc (3.50) là không lồi. Tương tự như phần 3.3.3, bằng cách áp dụng bất đẳng thức (A.2) trong Phụ lục A cho  $E_{k_i}$ , ta có

$$E_{k_i} \leq \frac{1}{2} m \sum_{l_i=0}^{L_i-1} \left( \frac{|\mathbf{q}_{i,l_i+1}^{(\kappa)} - \mathbf{q}_{i,l_i}^{(\kappa)}|}{4\tau_i^{(\kappa)}} \left( \frac{|\mathbf{q}_{i,l_i+1} - \mathbf{q}_{i,l_i}|}{|\mathbf{q}_{i,l_i+1}^{(\kappa)} - \mathbf{q}_{i,l_i}^{(\kappa)}|} + \frac{\tau_i^{(\kappa)}}{\tau_i} \right)^2 \right)^2$$

$$= \frac{m}{32} \sum_{l_i=0}^{L_i-1} \frac{|\mathbf{q}_{i,l_i+1}^{(\kappa)} - \mathbf{q}_{i,l_i}^{(\kappa)}|^2}{(\tau_i^{(\kappa)})^2} \left( \frac{|\mathbf{q}_{i,l_i+1}^{(\kappa)} - \mathbf{q}_{i,l_i}^{(\kappa)}|}{|\mathbf{q}_{i,l_i+1}^{(\kappa)} - \mathbf{q}_{i,l_i}^{(\kappa)}|} + \frac{\tau_i^{(\kappa)}}{\tau_i} \right)^4 \triangleq E_{k_i}^{(\kappa)}. \quad (3.53)$$

Do đó, hàm mục tiêu (3.51a) không lời được xấp xỉ bằng hàm lời sau

$$E_i \leq E_{k_i}^{(\kappa)} + E_{friction_i} + E_{s_i} \triangleq E_i^{(\kappa)}. \quad (3.54)$$

Tiếp theo, ta quan tâm đến ràng buộc (3.50). Ràng buộc này tương đương với

$$(x_{i,l_i} - x_{n,l_i}^O)^2 + (y_{i,l_i} - y_{n,l_i}^O)^2 \geq d_{r,n}^2, \quad (3.55)$$

$$\forall n \in \mathcal{N}, l_i \in [0, L_i],$$

trong đó,  $x_{n,l_i}^O = x_{n,i}^O + v_{nx} l_i \tau_i$ ,  $y_{n,l_i}^O = y_{n,i}^O + v_{ny} l_i \tau_i$ . Đặt vế phải của (3.55) là  $f(\mathbf{u}_{n,l_i})$ , trong đó,  $\mathbf{u}_{i,l_i} = (\mathbf{q}_{i,l_i}, \mathbf{q}_{n,l_i})$ . Lấy vi phân từng phần bậc hai của hàm  $f(\mathbf{u}_{n,l_i})$  theo các biến  $x_{i,l_i}, y_{i,l_i}$  và  $\tau_i$  ta được một ma trận Hessian. Nhận thấy ma trận này là bán xác định dương. Do đó,  $f(\mathbf{u}_{n,l_i})$  là hàm lồi. Vì hàm  $f(\mathbf{u}_{n,l_i})$  là hàm lồi, nên đạo hàm của nó, tức là,  $\nabla f(\mathbf{u}_{n,l_i}^{(\kappa)})$ , là siêu đạo hàm [90]. Do đó, ta có

$$\begin{aligned} f(\mathbf{u}_{n,l_i}) &\geq f(\mathbf{u}_{n,l_i}^{(\kappa)}) + \nabla f(\mathbf{u}_{n,l_i}^{(\kappa)})(\mathbf{u}_{n,l_i} - \mathbf{u}_{n,l_i}^{(\kappa)}) \\ &= (x_{i,l_i}^{(\kappa)} - x_{n,l_i}^{O(\kappa)})^2 + (y_{i,l_i}^{(\kappa)} - y_{n,l_i}^{O(\kappa)})^2 \\ &\quad + 2(x_{i,l_i}^{(\kappa)} - x_{n,l_i}^{O(\kappa)})(x_{i,l_i} - x_{i,l_i}^{(\kappa)}) + 2(y_{i,l_i}^{(\kappa)} - y_{n,l_i}^{O(\kappa)})(y_{i,l_i} - y_{i,l_i}^{(\kappa)}) \\ &\quad - 2l_i(v_{nx}(x_{i,l_i}^{(\kappa)} - x_{n,l_i}^{O(\kappa)}) + v_{ny}(y_{i,l_i}^{(\kappa)} - y_{n,l_i}^{O(\kappa)}))(\tau_i - \tau_i^{(\kappa)}) \triangleq f^{(\kappa)}(\mathbf{u}_{n,l_i}). \end{aligned} \quad (3.56)$$

Từ (3.55) và (3.56), ràng buộc không lời (3.49) được xấp xỉ thành ràng buộc lời sau

$$f^{(\kappa)}(\mathbf{u}_{n,l_i}) \geq d_{i,r,n}^2. \quad (3.57)$$

Từ (3.54) và (3.57), thay vì giải bài toán không lời (3.51), ta giải bài toán lời sau

$$\min_{\mathbf{Q}_i, \bar{\tau}_i, \mathbf{e}} E_i^{(\kappa)} \quad (3.58a)$$

$$\text{thoả mãn: (3.40) - (3.42), (3.51b) - (3.51e) và (3.57),} \quad (3.58b)$$

để thu được  $(\mathbf{Q}_i^{(\kappa+1)}, \tau_i^{(\kappa+1)})$  ở vòng lặp tiếp theo. Độ phức tạp tính toán của thuật toán

---

**Thuật toán 7** Thuật toán MSOW
 

---

- 1: Sử dụng Thuật toán 6 để tạo ra quỹ đạo toàn cục  $(\mathbf{Q}^{(s)}, \tau^{(s)})$  cho môi trường chỉ quan tâm đến vật cản tĩnh.
  - 2: Đặt  $(\mathbf{Q}^{(g)}, \tau^{(g)}) := (\mathbf{Q}^{(s)}, \tau^{(s)})$
  - 3: **repeat**
  - 4:   Robot di chuyển theo quỹ đạo toàn cục  $(\mathbf{Q}^{(g)}, \tau^{(g)})$ .
  - 5:   **if** Robot phát hiện ra  $N_i$  MOs tại bước  $i$  **then**
  - 6:     Khởi tạo: Khởi tạo giá trị ban đầu bất kỳ  $(\mathbf{Q}_i^{(0)}, \tau_i^{(0)})$  thỏa mãn các ràng buộc (3.51b), (3.51c) và (3.51e), rồi giải bài toán (3.59) để tìm điểm khả thi  $(\mathbf{Q}_i^{(\kappa)}, \tau_i^{(\kappa)})$  cho bài toán (3.58). Đặt  $\kappa = 0$ ;
  - 7:     **repeat**
  - 8:       Giải bài toán lồi (3.58) để thu được  $(\mathbf{Q}_i^{(\kappa+1)}, \tau_i^{(\kappa+1)}) = (\mathbf{Q}_i^*, \bar{\tau}_i^*)$ ;
  - 9:       Cập nhật  $(\mathbf{Q}_i^{(\kappa+1)}, \tau_i^{(\kappa+1)}) := (\mathbf{Q}_i^*, \tau_i^*)$
  - 10:        $\hat{\epsilon} = \frac{|E_i^{(\kappa+1)} - E_i^{(\kappa)}|}{E_i^{(\kappa)}}$
  - 11:       Đặt  $\kappa \leftarrow \kappa + 1$
  - 12:       **until**  $\hat{\epsilon} \leq \epsilon$ .
  - 13:     Thu được điểm tối ưu  $(\mathbf{Q}_i^{(\kappa)}, \tau_i^{(\kappa)})$ .
  - 14:     Sử dụng  $(\mathbf{Q}_i^{(\kappa)}, \tau_i^{(\kappa)})$  để cập nhật cho  $(\mathbf{Q}^{(g)}, \tau^{(g)})$ .
  - 15:   **end if**
  - 16: **until** Robot di chuyển đến đích.
- 

giải bài toán tối ưu lồi (3.46) là  $\mathcal{O}(\alpha^2\beta^{2.5} + \beta^{3.5})$  [97], trong đó,  $\alpha = 2L_i + KH(L_i + 1) + 3$  và  $\beta = 2(L_i + 1)KH + (L_i + 1)K + L_i + L_iN_i + 3$  lần lượt là số biến quyết định và số ràng buộc lồi khi giải bài toán (3.58). Tương tự như phần 3.3.3, chuỗi  $(\mathbf{Q}_i^{(\kappa+1)}, \tau_i^{(\kappa+1)})$  hội tụ tại điểm yên ngựa  $(\tilde{\mathbf{Q}}_i, \tilde{\tau}_i)$  sau một số vòng lặp giới hạn [98]. Để nâng cao hiệu quả tính toán của thuật toán, việc tìm điểm khả thi ban đầu là rất quan trọng.

Chọn một điểm khả thi bất kỳ  $(\mathbf{Q}_i^{(0)}, \tau_i^{(0)})$  cho các ràng buộc (3.43b), (3.43c) và (3.43e). Một điểm được gọi là khả thi cho bài toán (3.58)  $(\mathbf{Q}_i^{(\kappa)}, \tau_i^{(\kappa)})$  cần phải thỏa mãn ràng buộc (3.43d) và (3.57). Do đó, bài toán tối ưu cần giải có dạng

$$\max_{\mathbf{Q}_i, \mathbf{v}_i, \mathbf{e}} \eta \quad (3.59a)$$

$$\text{thỏa mãn: } (3.40) - (3.42), (3.51b) - (3.51c), (3.51e),$$

$$v_{\max} \geq v_{l_i}\eta, \forall l_i \in \{1, 2, \dots, L\}, \quad (3.59b)$$

$$f^{(\kappa)}(\mathbf{u}_{n, l_i}) \geq d_{i, r, n}^2 \eta, \quad (3.59c)$$

với  $\kappa = 0, 1, \dots$  cho đến khi giá trị của hàm mục tiêu (3.59) lớn hơn hoặc bằng 1. Khi đó, nghiệm tìm được  $(\mathbf{Q}_i^{(\kappa)}, \tau_i^{(\kappa)})$  sẽ là điểm khả thi cho bài toán (3.58).

Như đã đề cập ở đoạn đầu của phần này (phần 3.3.4), robot sẽ di chuyển theo quỹ đạo cục bộ, tức là,  $\mathbf{Q}_i$ , để tránh các vật cản động, sau đó, robot sẽ quay trở lại quỹ đạo toàn cục, tức là,  $\mathbf{Q}$ , để di chuyển đến đích. Do đó, thuật toán tổng thể cho bài

toán lập quỹ đạo cho robot trong môi trường có cả các SOs và MOs được trình bày ở Thuật toán 7. Thuật toán này được đặt tên là Thuật toán MSOW, nghĩa là thuật toán lập quỹ đạo cho robot trong môi trường chứa cả vật cản tĩnh và vật cản động (both Moving and Static Obstacles in the robot's Workspace).

### 3.3.5 Đánh giá kết quả

Phần này trình bày kết quả mô phỏng cho các thuật toán OSOW và MSOW. Đầu tiên là kết quả mô phỏng cho thuật toán OSOW. Sau đó, kết quả mô phỏng cho thuật toán MSOW. Luận án quan tâm đến khu vực làm việc của robot trong vùng  $10\text{m} \times 10\text{m}$  [41] bao gồm 5 chướng ngại vật. Các tham số mô phỏng được trình bày trong Bảng 3.2. Mô phỏng được thực hiện trên MATLAB cùng với gói CVX 2.2 [86, 99–101] cùng với trình giải Mosek 9.1.9 với độ chính xác mặc định.

Bảng 3.2: Bảng tham số mô phỏng thuật toán OSOW và MOSW

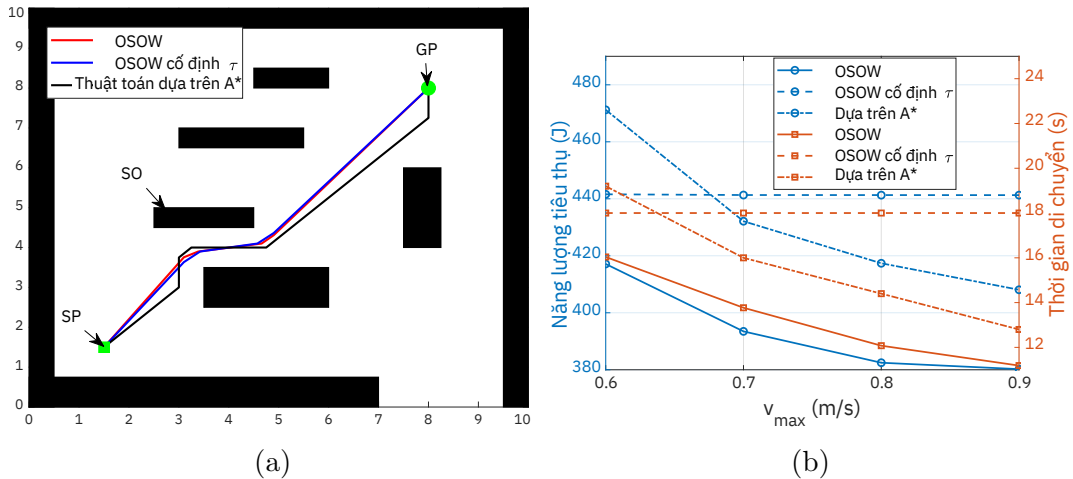
Tham số	Giá trị	Tham số	Giá trị
$L$	30	$\delta$	0.1 m
$v_{\max}$	0.7 m/s	$\mu$	0.05
$\tau_{\max}$	1 s	$\tau_{\min}$	0.01 s
$g$	$9.8\text{m/s}^2$	$m$	9 kg [41]
$\hat{M}$	100 [94]	$P_s$	17.8 W
$\mathbf{q}_s$	[1.5, 1.5]	$\mathbf{q}_g$	[8, 8]

Để thuận tiện cho việc so sánh, ở đây, chúng ta quan tâm đến các thuật toán sau:

- **OSOW:** Đây là thuật toán xấp xỉ lỗi được mô tả ở phần 3.3.3 (Thuật toán 6). Thuật toán 6 được đề xuất để giải bài toán tối ưu (3.43).
- **OSOW cố định  $\tau$ :** Đây là thuật toán tương tự như thuật toán OSOW trong đó giá trị khoảng thời gian mỗi bước  $\tau$  là cố định.
- **MSOW:** Đây là thuật toán xấp xỉ lỗi được mô tả ở phần 3.3.4 (Thuật toán 7). Thuật toán 7 được đề xuất để giải bài toán tối ưu (3.51).
- **MSOW cố định  $\tau$ :** Đây là thuật toán tương tự như thuật toán MSOW trong đó giá trị khoảng thời gian mỗi bước  $\tau$  là cố định.

Đầu tiên, chúng ta quan tâm đến kịch bản chỉ có các vật cản tĩnh (SOs). Để đánh giá thuật toán đề xuất, chúng ta xem xét đến thuật toán dựa trên nền thuật toán

$A^*$  [41]. Chú ý rằng thuật toán dựa trên thuật toán  $A^*$  chỉ được sử dụng cho kịch bản OSOW, nên việc so sánh chỉ được thực hiện với kịch bản này. Hình 3.4(a) thể hiện quỹ đạo tối ưu khi sử dụng ba thuật toán: Thuật toán OSOW, Thuật toán OSOW với  $\tau$  cố định, và Thuật toán dựa trên thuật toán  $A^*$ . Quan sát thấy rằng, quỹ đạo sinh ra bởi Thuật toán OSOW và Thuật toán OSOW với  $\tau$  cố định gần tương tự nhau và mượt mà hơn so với thuật toán dựa trên nền  $A^*$ . Ngoài ra, Hình 3.4(b) thể hiện sự thay đổi của năng lượng tiêu thụ và thời gian di chuyển của robot khi thay đổi giá trị vận tốc lớn nhất của robot khi sử dụng ba thuật toán. Nhìn vào hình ta thấy, năng lượng tiêu thụ và thời gian di chuyển của robot thu được khi sử dụng thuật toán OSOW luôn là thấp nhất tại mỗi giá trị  $v_{max}$ . Kết quả này thể hiện rõ ràng sự quan trọng và cần thiết của việc tối ưu khoảng thời gian mỗi bước  $\tau$  ở thuật toán OSOW. Điều này cũng thể hiện tính hiệu quả của thuật toán đề xuất.

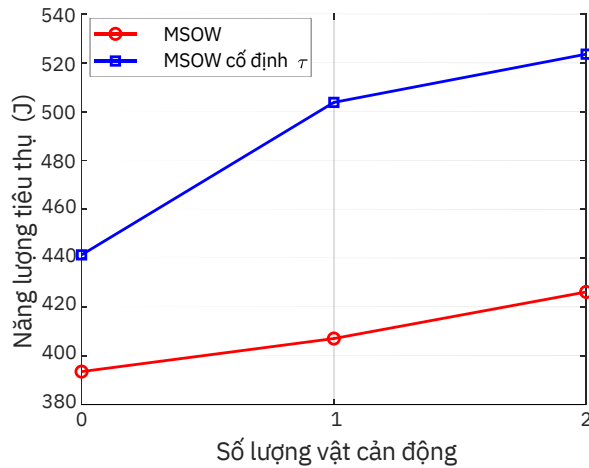


Hình 3.4: a) Quỹ đạo tối ưu bởi ba thuật toán b) Năng lượng tiêu thụ và thời gian di chuyển tương ứng với vận tốc lớn nhất của robot.

Tiếp theo, chúng ta xem xét đến ảnh hưởng của giá trị vận tốc lớn nhất của robot, tức là,  $v_{max}$ , đến mức năng lượng tiêu thụ của ba thuật toán trong kịch bản chỉ có vật cản tĩnh. Như nhìn thấy trong hình 3.4, khi  $v_{max}$  tăng, năng lượng tiêu thụ và thời gian di chuyển của robot khi sử dụng thuật toán OSOW và thuật toán dựa trên  $A^*$  giảm, còn khi sử dụng thuật toán OSOW với  $\tau$  cố định thì các đại lượng này gần như không đổi. Các kết quả này có thể được giải thích như sau. Với thuật toán OSOW, khi  $v_{max}$  tăng thì tổng thời gian di chuyển của robot sẽ giảm, từ đó sẽ giảm năng lượng tiêu thụ trên máy tính, các mạch điện tử và cảm biến, tức là giảm  $E_s$ . Mặt khác, khi  $v_{max}$  tăng, robot di chuyển với tốc độ cao hơn lại làm tăng năng lượng tiêu thụ cho hệ thống chuyển động của robot, tức là  $E_{motion}$ . Tuy nhiên, điều thú vị là tổng năng lượng



tiêu thụ của robot đã giảm. Điều này có thể lý giải do mức độ giảm của  $E_s$  nhiều hơn so với mức độ tăng của  $E_{motion}$ . Đây cũng chính là lý do để giải thích cho việc tổng năng lượng tiêu thụ và quãng đường di chuyển của robot giảm dần khi  $v_{max}$  tăng. Với thuật toán OSOW có  $\tau$  cố định, khi  $v_{max}$  tăng, thời gian di chuyển của robot không đổi. Nguyên nhân là do vị trí của robot tại mỗi bước không thay đổi khi  $v_{max}$  thay đổi. Hơn nữa, vì khoảng thời gian giữa các bước  $\tau$  đã được cố định nên vận tốc thực tế của robot cũng không đổi. Kết quả là, tổng năng lượng tiêu thụ gần như không thay đổi.

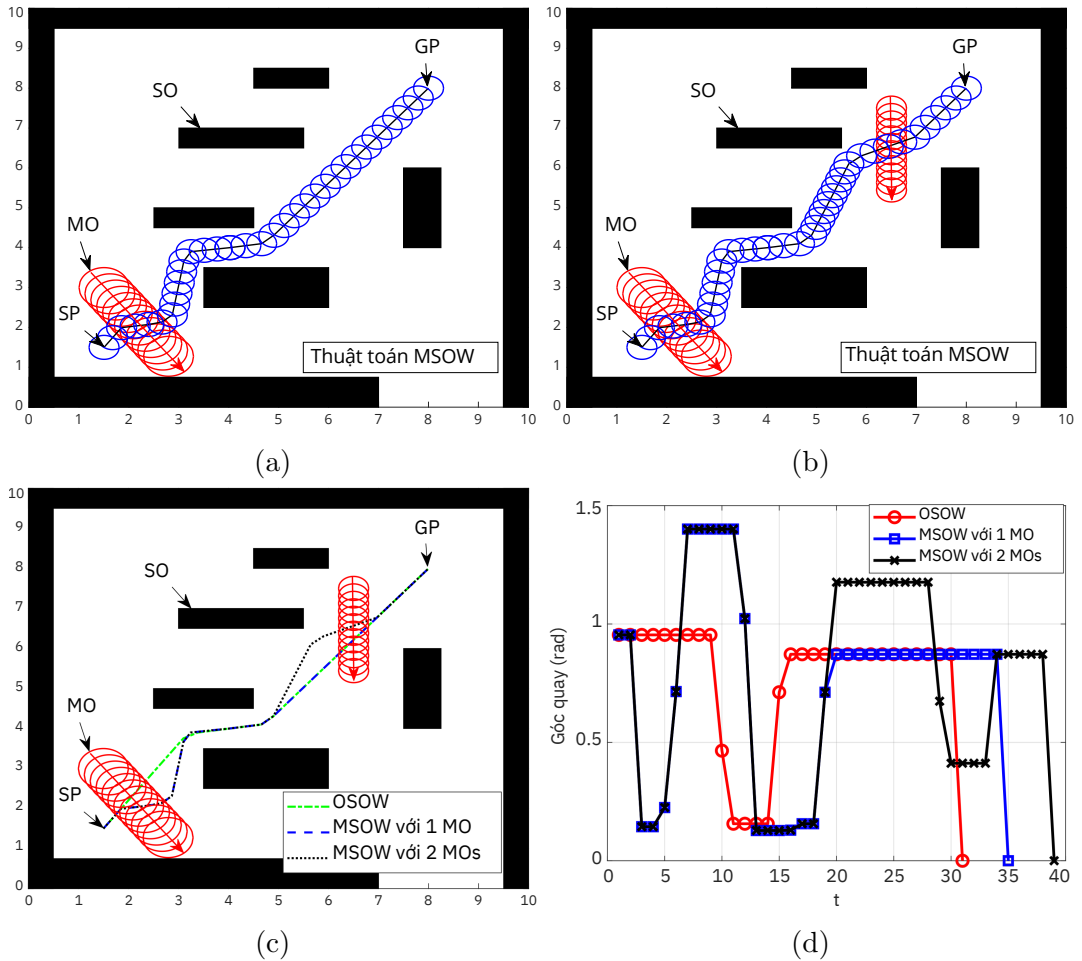


(a)

Hình 3.5: Năng lượng tiêu thụ và quãng đường di chuyển khi số lượng vật cản động trong môi trường làm việc thay đổi.

Bây giờ, chúng ta xét đến kịch bản trong môi trường làm việc của robot có cả vật cản tĩnh và vật cản động. Lúc này, việc chỉ ra sự ảnh hưởng của số lượng các vật cản động (MOs) xuất hiện tác động lên năng lượng tiêu thụ của robot là một vấn đề quan trọng. Trong các mô phỏng, giả sử robot phát hiện ra một MO tại bước số 3, lúc đó vị trí của MO là (1.5, 3) với bán kính 0.5 m. Sau đó, robot phát hiện thêm một MO khác ở bước 20, khi đó vị trí của MO là (6.5, 7.5) với bán kính 0.3 m. Như thấy trên Hình 3.5(a), khi số lượng MOs tăng lên, năng lượng tiêu thụ của robot tăng lên. Điều này hoàn toàn hợp lý, vì khi số lượng MOs tăng lên, để không xảy ra va chạm, robot cần tránh vật cản dẫn đến quãng đường và thời gian di chuyển tăng lên dẫn đến năng lượng tiêu thụ tăng lên. Ngoài ra, ta còn thấy năng lượng tiêu thụ khi sử dụng thuật toán MSOW nhỏ hơn khi sử dụng thuật toán MSOW với  $\tau$  cố định. Điều này thể hiện tính hiệu quả của thuật toán đề xuất.

Cuối cùng, chúng ta xem xét đến quỹ đạo được sinh ra bởi các thuật toán đề xuất ở Hình 3.6(a), (b), (c), và (d). Trên hình, mỗi đường tròn màu xanh và các đường



Hình 3.6: Quỹ đạo tối ưu của robot khi tối thiểu năng lượng tiêu thụ với a) kịch bản MSOW có 01 MO b) kịch bản MSOW có 02 MOs, c) ba kịch bản và d) Góc quay của robot tương ứng với các quỹ đạo tối ưu.

tròn màu đỏ lần lượt thể hiện cho robot và các vật cản động tại mỗi bước. Hình 3.6(a) và (b) lần lượt thể hiện quỹ đạo của robot ở kịch bản có một và hai vật cản động. Có thể thấy rằng quỹ đạo của robot đã thay đổi so với quỹ đạo toàn cục (tức là quỹ đạo sinh ra bởi thuật toán OSOW) nhằm mục đích tránh vật cản động. Hình 3.6(c) cũng đã thể hiện rõ sự thay đổi này trong mỗi trường hợp tương ứng với từng thuật toán đề xuất. Góc quay của robot theo thời gian trên hình 3.6(c) cho thấy quỹ đạo của robot vẫn mượt mà và yêu cầu tránh vật cản hoàn toàn được đảm bảo với mức năng lượng tiêu thụ nhỏ nhất có thể.

### 3.4 Kết luận chương 3

Với mục tiêu lập quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng, những đóng góp chính của chương 3 như sau:

- Đề xuất thuật toán CAA để giải quyết bài toán tối ưu quỹ đạo cho robot di động

trong môi trường truyền thông lý tưởng.

- Đề xuất thuật toán OSOW và MSOW để giải quyết bài toán lập quỹ đạo nhằm mục tiêu tối thiểu năng lượng tiêu thụ tương ứng với trường hợp vật cản tĩnh và vật cản động.

Các thuật toán trình bày ở Chương 3 đã giải quyết tốt bài toán với giả định rằng môi trường truyền thông là lý tưởng. Tuy nhiên, trên thực tế, rất khó để đạt được trạng thái truyền thông lý tưởng trong các nhà máy. Khi hoạt động trong môi trường truyền thông không lý tưởng, bài toán sẽ có thêm ràng buộc về tốc độ truyền thông trong quá trình robot di chuyển. Mô hình và thuật toán giải quyết bài toán này sẽ được trình bày ở chương 4.

Đóng góp của chương 3 được công bố ở công trình CT2 và CT3. Công trình CT2 trình bày về các thuật toán dựa trên tối ưu lồi tối ưu quỹ đạo cho robot trong môi trường nhà máy có vật cản động và tĩnh nhằm tối thiểu năng lượng tiêu thụ của robot. Công trình CT3 trình bày thuật toán dựa trên tối ưu lồi tối ưu quỹ đạo cho robot trong môi trường nhà máy nhằm tối thiểu quãng đường di chuyển cho robot.

[CT2] **N. T. T. Van**, N. M. Tien, N. C. Luong, and H. T. K. Duyen, “Energy consumption minimization for autonomous mobile robot: A convex approximation approach,” *Journal of Robotics and Control (JRC)*, vol. 4, no. 3, pp. 403–412, 2023.

[CT3] **N. T. T. Van**, N. Manh Tien, H. T. K. Duyen, N. M. Cuong, and D. H. Duân, “A Convex Approximation Method to Optimize Trajectory for AGV Robot Motion Planning,” in *Proceedings of the sixth Vietnam international conference and Exhibition on Control and Automation VCCA-2021*; ISBN 978-604-95-0875-2.

## **Chương 4. THUẬT TOÁN TỐI ƯU QUỸ ĐẠO CHO ROBOT DI ĐỘNG TRONG MÔI TRƯỜNG TRUYỀN THÔNG CÓ STAR-IRS HỖ TRỢ**

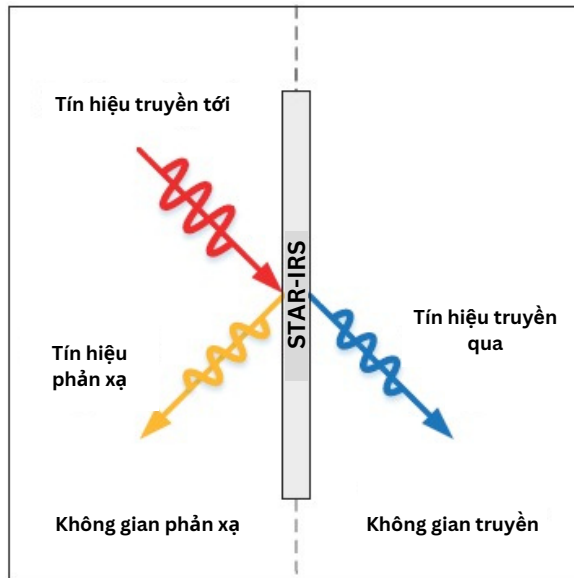
Chương 4 trình bày về mô hình hệ thống điều hướng cho robot di động có STAR-IRS hỗ trợ, từ đó xây dựng bài toán tối ưu quỹ đạo cho robot di động có xét đến ràng buộc về truyền thông. Bài toán tối ưu là không lỗi. Hơn thế nữa, khi robot di động di chuyển vị trí của robot thay đổi, dẫn đến kênh truyền thay đổi, dẫn đến bài toán tối ưu quỹ đạo cho robot lúc này là một bài toán động. Giải bài toán động là một thách thức đối với các thuật toán truyền thống như thuật toán xấp xỉ lỗi. Vì vậy, ở chương này, luận án đề xuất sử dụng thuật toán dựa trên kỹ thuật học sâu tăng cường DRL (Deep Reinforcement Learning) tối ưu quỹ đạo cho robot di động trong môi trường truyền thông có STAR-IRS hỗ trợ. Phần 4.1 đề xuất mô hình hệ thống điều hướng cho robot di động có sử dụng STAR-IRS hỗ trợ truyền thông. Sau đó, phần 4.2 xây dựng mô hình quy trình quyết định Markov, phần 4.3 trình bày thuật toán dựa trên kỹ thuật DRL để giải quyết bài toán tối ưu quỹ đạo cho mục tiêu tối thiểu năng lượng tiêu thụ của hệ thống. Đánh giá hiệu quả mô hình và thuật toán điều hướng cho robot di động được trình bày ở phần 4.4. Cuối cùng là phần kết luận chương.

### **4.1 Mô hình hệ thống điều hướng cho robot di động có STAR-IRS hỗ trợ**

Trong môi trường truyền thông, robot di động không chỉ làm việc riêng lẻ mà còn cần trao đổi thông tin với các thiết bị khác như các trạm chủ trung tâm, các robot di động khác trong hệ đa robot di động, hay các thiết bị không dây khác. Tuy nhiên, môi trường nhà máy với đặc tính nhiều vật cản cùng với hạn chế của hệ thống mạng không dây hiện tại dẫn đến rất khó để đạt được trạng thái môi trường truyền thông lý tưởng. Do đó, để hỗ trợ cho hoạt động của robot di động trong nhà máy, ở phần này, NCS đề xuất sử dụng một mô hình hệ thống điều hướng cho robot di động có STAR-IRS hỗ trợ.

#### **4.1.1 STAR-IRS**

Như đã trình bày ở phần 1.3.2, các mô hình hệ thống đề xuất sử dụng IRS đã nói ở trên đã cho kết quả rất tốt. Tuy nhiên, khả năng hỗ trợ của các IRS chỉ được một nửa không gian phía trước, tức là vùng làm việc là  $180^0$  vì các phần tử trên IRS chỉ



Hình 4.1: STAR-IRS. [9]

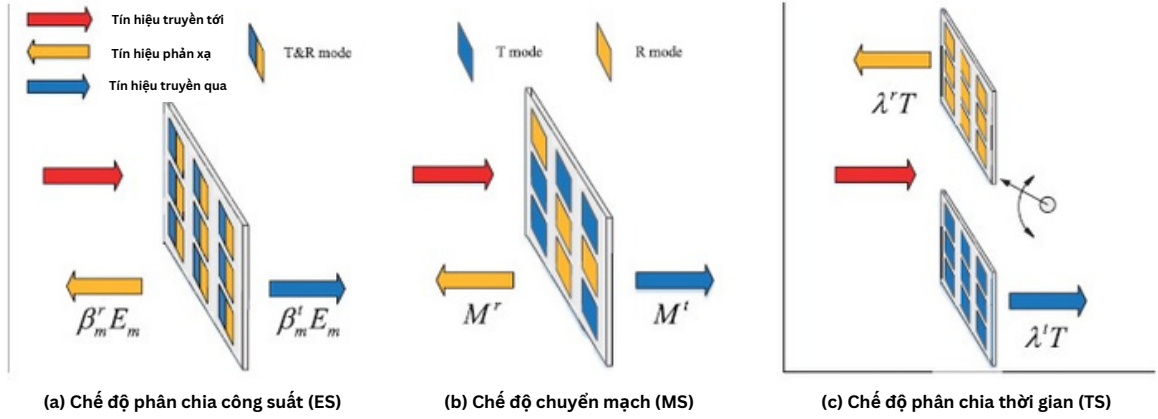
có khả năng phản xạ tín hiệu trở lại không gian của tín hiệu truyền tới. Điều này dẫn đến ràng buộc về mặt vị trí của bên phát và bên nhận tín hiệu phải cùng thuộc một phía so với IRS. Ràng buộc địa lý này rất khó áp dụng trong thực tế đặc biệt là đối với robot di động, và hơn nữa, nó hạn chế tính linh hoạt và hiệu quả của IRS, vì thông thường các thiết bị có thể được đặt ở cả hai bên của một IRS. Gần đây, khái niệm về IRS có khả năng vừa truyền vừa phản xạ STAR-IRS (Simultaneously Transmitting and Reflecting IRS) đã được đề xuất trong [102] và [103]. Cụ thể, như được thể hiện trong Hình 4.1, tín hiệu có thể truyền đến một phần tử của STAR-IRS từ cả hai hướng của bề mặt [104]. Một phần tín hiệu được phản xạ vào cùng phía không gian với tín hiệu đến, gọi là không gian phản xạ, và phần còn lại được truyền đến không gian đối diện với tín hiệu đến gọi là không gian truyền. Bằng cách điều khiển cả dòng điện và từ tính của một phần tử STAR-IRS, tín hiệu truyền và phản xạ có thể được cấu hình lại thông qua hai hệ số thường độc lập, tức là hệ số truyền và hệ số phản xạ [102]. Do đó, một môi trường vô tuyến thông minh (Smart Radio Environments) toàn diện được tạo ra [9]. Khi đó, với STAR-IRS, vùng phủ sóng của tín hiệu được mở rộng  $360^\circ$  xung quanh IRS. Vùng phủ sóng tăng lên, chất lượng truyền thông được cải thiện sẽ là một thế mạnh của STAR-IRS. Các tác giả ở [9] xem xét một hệ thống truyền thông downlink được hỗ trợ bởi STAR-IRS bao gồm một BS, một STAR-IRS và một số thiết bị. Bài toán tối ưu kết hợp bộ tạo chùm tia active và passive, độ dịch pha và hệ số phản xạ/truyền của STAR-IRS đã được đặt ra để tối thiểu năng lượng tiêu thụ của BS trong khi đảm bảo các yêu cầu QoS của thiết bị cho cả truyền dữ liệu unicast

và multicast. Kết quả mô phỏng cho thấy việc sử dụng STAR-IRS có thể giảm đáng kể lượng tiêu thụ năng lượng của BS so với việc sử dụng IRS thông thường. Ngoài ra, sử dụng STAR-IRS cho UAV, các tác giả trong [55] xem xét một hệ thống truyền thông A2G (Air-to-Ground) bao gồm một UAV hoạt động như một BS và một số GUs (ground users). Các tác giả đề xuất tối đa hóa tổng tốc độ truyền thông trên thời gian di chuyển của UAV bằng cách tối ưu đồng thời quỹ đạo của UAV, bộ tạo chùm tia của UAV và bộ tạo chùm tia của STAR-IRS, độ dịch pha và hệ số phản xạ/truyền của STAR-IRS. Một thuật toán dựa trên học sâu tăng cường DRL (Deep Reinforcement Learning) mới đã được đề xuất để giải quyết bài toán tối ưu. Kết quả mô phỏng chứng minh rằng STAR-IRS đạt được sự tăng tổng tốc độ truyền thông đáng kể so với kịch bản IRS thông thường.

Dựa trên những phân tích trên về lợi ích mà STAR-IRS mang lại, luận án đề xuất sử dụng STAR-IRS hỗ trợ hoạt động cho robot di động trong môi trường truyền thông không dây. Như quan sát thấy trên hình 4.1, tín hiệu truyền đến một phần tử của STAR-IRS có thể chia thành 2 thành phần: một phần tín hiệu truyền đi về không gian phía sau IRS và một tín hiệu phản xạ lại không gian phía trước IRS. Để giải thích về đặc tính của STAR-IRS này, gọi  $s_n$  là tín hiệu truyền trên phần tử thứ  $n$  của STAR-IRS, trong đó  $n \in \mathcal{N} \triangleq \{1, 2, \dots, N\}$  và  $N$  là tổng số lượng phần tử của STAR-IRS. Tín hiệu truyền và phản xạ bởi phần tử  $n$  được mô hình hóa lần lượt là  $t_n = (\sqrt{\beta_n^t} e^{j\bar{\theta}_n^t}) s_n$  và  $r_n = (\sqrt{\beta_n^r} e^{j\bar{\theta}_n^r}) s_n$  [102], trong đó  $\sqrt{\beta_n^t} \in [0, 1]$ ,  $\bar{\theta}_n^t \in [0, 2\pi)$  và  $\sqrt{\beta_n^r} \in [0, 1]$ ,  $\bar{\theta}_n^r \in [0, 2\pi)$  lần lượt là độ lớn và độ dịch pha của hệ số truyền và phản xạ của phần tử  $n$ . Lưu ý rằng, độ dịch pha truyền ( $\bar{\theta}_n^t$  và độ dịch pha phản xạ ( $\bar{\theta}_n^r$ ) có thể được điều chỉnh độc lập với nhau [102]. Tuy nhiên, việc điều chỉnh về độ lớn của hệ số truyền và phản xạ ( $\beta_n^t$  và  $\beta_n^r$ ) phải đảm bảo thỏa mãn ràng buộc về tổng năng lượng. Điều này có nghĩa là, đối với mỗi phần tử, tổng năng lượng của tín hiệu truyền và phản xạ từ STAR-IRS phải bằng năng lượng của tín hiệu truyền tới STAR-IRS, tức là  $|t_n|^2 + |r_n|^2 = |s_n|^2$ . Điều này dẫn đến ràng buộc về độ lớn hệ số truyền và phản xạ từ IRS trên mỗi phần tử như sau [102]:

$$\beta_n^t + \beta_n^r = 1, \forall n \in \mathcal{N}. \quad (4.1)$$

Từ (4.1) có thể thấy rằng bằng cách điều chỉnh tham số về độ lớn hệ số truyền và phản xạ, mỗi phần tử của STAR-IRS sẽ có 3 chế độ hoạt động: chế độ chỉ truyền



Hình 4.2: Các chế độ hoạt động của STAR-IRS. [9]

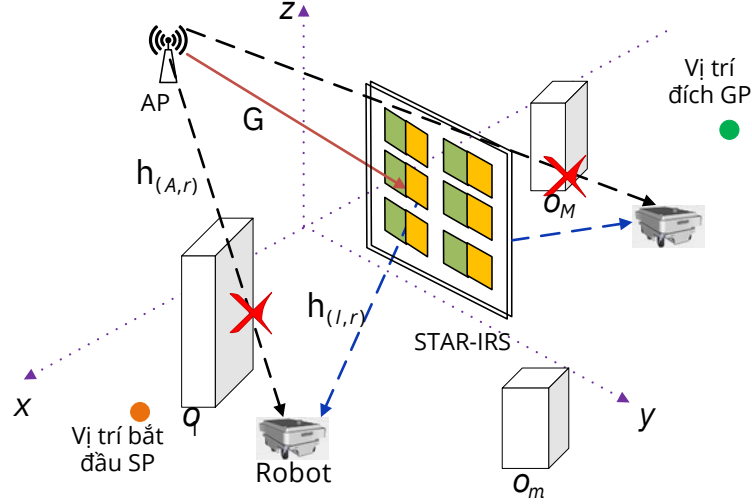
tín hiệu từ STAR-IRS (tức là  $\beta_n^t = 1$  và  $\beta_n^r = 0$ , gọi là chế độ T), chế độ chỉ phản xạ tín hiệu từ STAR-IRS (tức là  $\beta_n^t = 0$  và  $\beta_n^r = 1$ , gọi là chế độ R), và chế độ vừa truyền vừa phản xạ từ STAR-IRS (tức là  $\beta_n^t, \beta_n^r \in [0, 1]$ , gọi là chế độ T&R).

Nhìn chung, STAR-IRS có 3 kiểu hoạt động: phân chia năng lượng ES (Energy Splitting), chuyển mạch MS (Mode Switching) và phân chia thời gian TS (Time Switching) như trình bày trên hình 4.2. Với ES, tất cả các phần tử của STAR-IRS được giả sử là hoạt động ở chế độ T&R, tức là năng lượng từ tín hiệu truyền đến STAR-IRS được chia thành hai phần, một phần truyền từ IRS đi và một phần phản xạ lại với một tỷ lệ  $\beta_n^t : \beta_n^r$  nhất định nào đó. Với MS, các phần tử trên STAR-IRS được chia ra thành hai nhóm, một nhóm chứa  $N^t$  phần tử hoạt động ở chế độ T, nhóm còn lại gồm  $N^r$  phần tử hoạt động ở chế độ R, với  $N^t + N^r = N$ . Còn với TS, các phần tử của STAR-IRS sẽ được phân chia khoảng thời gian ở chế độ T và chế độ R trong mỗi khe thời gian.

#### 4.1.2 Mô hình hệ thống điều hướng cho robot di động trong môi trường truyền thông

Luận án đề xuất sử dụng một mô hình hệ thống điều hướng cho robot di động trong môi trường truyền thông bao gồm một AP một ăng-ten, một tấm STAR-IRS, và một robot di động có gắn một ăng-ten như mô tả trên Hình 4.3. Do môi trường phức tạp trong nhà máy, robot di động có thể gặp phải các chướng ngại vật, các kênh trực tiếp giữa AP và robot di động có thể không đủ ổn định hoặc thậm chí bị chặn. Để giải quyết vấn đề này, luận án đề xuất một hệ thống truyền thông được STAR-IRS hỗ trợ gồm  $N$  phần tử truyền và phản xạ tín hiệu, tập hợp các phần tử của STAR-IRS được

ký hiệu là  $\mathcal{N} = \{1, \dots, N\}$ . Về việc lắp đặt, STAR-IRS có thể được treo trên trần của nhà máy để có kênh truyền giữa AP và robot di động chất lượng cao trong toàn bộ không gian  $360^\circ$  quanh IRS. Đặt  $\hat{T}$  là thời gian di chuyển của robot di động. Để đơn giản,  $\hat{T}$  được chia thành  $T$  bước thời gian  $\tau$ , tức là  $\hat{T} = T\tau$ , trong đó,  $\tau$  là khoảng thời gian cho một bước thời gian.



Hình 4.3: Mô hình hệ thống điều hướng robot di động trong nhà có STAR-IRS hỗ trợ truyền thông.

robot di động di chuyển trong nhà máy với vận tốc không đổi  $v$  và thay đổi hướng của mình tại mỗi khe thời gian để điều hướng di chuyển. Quỹ đạo của robot di động sau đó có thể được biểu diễn bằng  $\mathbf{q}_t = [x_t, y_t], t \in \mathcal{T} = \{1, \dots, T\}$ , và robot di động được gán một ăng ten với độ cao  $z_0$ , và góc hướng tại khe thời gian  $t$  là  $\varphi(\mathbf{q}_t)$  với

$$\varphi(\mathbf{q}_t) = \arccos \frac{x_{t+1} - x_t}{\sqrt{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}}. \quad (4.2)$$

Trong thực tế, quỹ đạo của robot di động cần đảm bảo các ràng buộc về điểm đầu và điểm cuối như sau

$$\mathbf{q}_1 = \mathbf{q}_s, \mathbf{q}_T = \mathbf{q}_g, \quad (4.3)$$

trong đó,  $\mathbf{q}_s$  và  $\mathbf{q}_g$  lần lượt là vị trí điểm đầu (start position) và điểm cuối (goal position) của robot di động.

Xét môi trường nhà máy, NCS giả định rằng có một tập hợp  $\mathcal{M}$  các vật cản trong không gian làm việc của robot di động. Thông qua các cảm biến được trang bị, robot di động nhận biết được các thông tin môi trường xung quanh như vị trí của các



vật cản. NCS giả định rằng các cảm biến có phạm vi phát hiện là trong khoảng  $R$  m từ robot di động, điều này có nghĩa là robot di động không có thông tin về các vật cản trừ khi chúng nằm trong phạm vi phát hiện. Giả sử yêu cầu về khoảng cách an toàn giữa robot di động và các vật cản là  $d_{\min}$ , và  $d_{\min} < R$ . Một vùng cấm được xác định xung quanh vật cản với bán kính  $d_{\min}$ , và robot di động không được phép di chuyển qua vùng này để giữ khoảng cách an toàn. Các vật cản xuất hiện trong quá trình di chuyển của robot di động được gọi là  $o_i \in \mathcal{M}$ . Để đảm bảo an toàn khi di chuyển, các ràng buộc sau phải được thỏa mãn:

$$\|\mathbf{q}_t - \mathbf{q}_{o_i}\| \geq d_{\min}, \forall o_i \in \mathcal{M}, \quad (4.4)$$

trong đó,  $\mathbf{q}_{o_i}$  là vị trí của vật cản  $o_i$ .

Với sự hỗ trợ của các dòng từ bề mặt, cả độ lớn và độ dịch pha của các tín hiệu truyền và phản xạ của mỗi phần tử của STAR-IRS có thể được điều chỉnh một cách độc lập. Như thấy trong hình 4.3, vị trí của STAR-IRS được đặt tại trục Oy. Do đó, vùng làm việc có vị trí  $x > 0$  là vùng phản xạ tín hiệu từ STAR-IRS, trong khi vùng  $x < 0$  là vùng truyền tín hiệu xuyên qua STAR-IRS. Gọi hệ số truyền và phản xạ của STAR-IRS tại khe thời gian  $t$  lần lượt là  $\bar{\theta}_t^{re} = \beta_t^{re} e^{j\phi_t^{re}}$ , và  $\bar{\theta}_t^{tr} = \beta_t^{tr} e^{j\phi_t^{tr}}$ , trong đó,  $\beta_t^{re}, \beta_t^{tr} \in [0, 1]$ , và  $\phi_t^{re}, \phi_t^{tr} \in [0, 2\pi)$  lần lượt là độ lớn và pha truyền và phản xạ từ STAR-IRS. Gọi  $\bar{\Theta}_t^{re} \in \mathbb{C}^{N \times N}$  và  $\bar{\Theta}_t^{tr} \in \mathbb{C}^{N \times N}$  lần lượt là ma trận đường chéo của hệ số truyền và phản xạ. Ta có  $\bar{\Theta}_t^{re} \triangleq \text{diag}(\bar{\theta}_t^{re})$ , với  $\bar{\theta}_t^{re} = [\bar{\theta}_{t,1}^{re}, \bar{\theta}_{t,2}^{re}, \dots, \bar{\theta}_{t,N}^{re}]$ , trong đó,  $\bar{\theta}_{t,n}^{re}$  là hệ số phản xạ của phần tử  $n$  của STAR-IRS, và  $\bar{\Theta}_t^{tr} \triangleq \text{diag}(\bar{\theta}_t^{tr})$ , với  $\bar{\theta}_t^{tr} = [\bar{\theta}_{t,1}^{tr}, \bar{\theta}_{t,2}^{tr}, \dots, \bar{\theta}_{t,N}^{tr}]$ , trong đó  $\bar{\theta}_{t,n}^{tr}$  là hệ số truyền của phần tử  $n$  của STAR-IRS thỏa mãn ràng buộc sau:

$$(\beta_{t,n}^{re})^2 + (\beta_{t,n}^{tr})^2 = 1, \forall n \in \mathcal{N}, t \in \mathcal{T}. \quad (4.5)$$

Thêm vào đó, giả sử rằng các hệ số dịch pha (phase-shift coefficients) cho việc truyền tín hiệu và phản xạ tín hiệu được điều chỉnh riêng biệt với độ dịch pha đã được rời rạc hóa. Để đơn giản hóa, NCS quan tâm đến dạng độ dịch pha trong thực tế, tức là độ dịch pha được thiết kế bằng cách sử dụng  $b$  bit để lượng tử hóa khoảng dịch pha  $[0, 2\pi)$ . Theo đó,  $b$  được gọi là số lượng bit lượng tử hóa pha, và có  $2^b$  giá trị có thể có

của độ dịch pha. Để đạt được điều này, ràng buộc sau đây cần được đảm bảo

$$\phi_{t,n}^c = \frac{\xi\pi}{2^{b-1}}, \xi \in \{0, 1, \dots, 2^b - 1\}, \forall n \in \mathcal{N}, t \in \mathcal{T}, \quad (4.6)$$

trong đó  $c \in \{re, tr\}$ . Cần chú ý thêm rằng trong hệ thống đề xuất, mỗi STAR-IRS có thể hoạt động ở chế độ phản xạ hoàn toàn hoặc truyền hoàn toàn tín hiệu bằng cách điều chỉnh tỷ lệ phân chia năng lượng (ES).

### 4.1.3 Mô hình tín hiệu

Vị trí của robot di động sẽ xác định robot di động sẽ nhận được tín hiệu phản xạ hay tín hiệu truyền từ STAR-IRS. Đặt  $\mathbf{g} \in \mathbb{C}^{N \times 1}$  là kênh từ AP đến IRS,  $h_{(A,r)}(\mathbf{q}_t) \in \mathbb{C}$  và  $\mathbf{h}_{(I,r)}(\mathbf{q}_t) \in \mathbb{C}^N$  lần lượt là kênh trực tiếp từ AP đến robot di động và kênh từ STAR-IRS đến robot di động tại vị trí  $\mathbf{q}_t$ . Mỗi kênh bao gồm hai thành phần là large-scale component và small-scale fading.

Thành phần small-scale fading từ AP đến STAR-IRS, từ STAR-IRS đến robot di động, và từ AP đến robot di động được mô hình hóa với phân bố Rician với hệ số K-factor lần lượt là  $\kappa_{AI}$ ,  $\kappa_{Ir}(\mathbf{q}_t)$ , và  $\kappa_{Ar}(\mathbf{q}_t)$ . Cụ thể, các kênh truyền giữa AP-STAR-IRS, STAR-IRS-robot, và AP-robot được biểu diễn như sau [105]

$$\mathbf{g} = \frac{\mathbf{PL}_{AI}}{\sqrt{\kappa_{AI} + 1}} \left( \sqrt{\kappa_{AI}} \mathbf{g}^{LoS} + \mathbf{g}^{NLoS} \right), \quad (4.7)$$

$$\mathbf{h}_{(I,r)}(\mathbf{q}_t) = \frac{\mathbf{PL}_{Ir}}{\sqrt{\kappa_{Ir}(\mathbf{q}_t) + 1}} \left( \sqrt{\kappa_{Ir}(\mathbf{q}_t)} \mathbf{h}_{(I,r)}^{LoS}(\mathbf{q}_t) + \mathbf{h}_{(I,r)}^{NLoS} \right), \quad (4.8)$$

$$h_{(A,r)}(\mathbf{q}_t) = \frac{\mathbf{PL}_{Ar}}{\sqrt{\kappa_{Ar}(\mathbf{q}_t) + 1}} \left( \sqrt{\kappa_{Ar}(\mathbf{q}_t)} h_{(A,r)}^{LoS}(\mathbf{q}_t) + h_{(A,r)}^{NLoS} \right), \quad (4.9)$$

trong đó,  $\mathbf{PL}_{AI}$ ,  $\mathbf{PL}_{Ir}$ , và  $\mathbf{PL}_{Ar}$  lần lượt là tổn hao đường truyền (path loss) của các kênh truyền AP-IRS, IRS-robot di động, và AP-IRS,  $\mathbf{g}^{LoS}$ ,  $\mathbf{h}_{(I,r)}^{LoS}$ , và  $h_{(A,r)}^{LoS}$  là thành phần line-of-sight (LoS), và  $\mathbf{g}^{NLoS}$ ,  $\mathbf{h}_{(I,r)}^{NLoS}$ , và  $h_{(A,r)}^{NLoS}$  là thành phần non-LoS (NLoS). Ngoài ra, nếu tín hiệu truyền giữa robot di động tại vị trí  $\mathbf{q}_t$  và AP/IRS bị chặn bởi các vật cản thì các kênh truyền tương ứng sẽ chỉ còn thành phần NLoS, tức là khi đó  $\kappa_{Ar/Ir}(\mathbf{q}_t) = 0$ . Ngược lại, tín hiệu là LoS sẽ có hệ số K-factor  $\kappa_{Ar/Ir}(\mathbf{q}_t)$ .

Tín hiệu nhận được của robot di động được mô hình hóa như sau:

$$y_t^c = (h_{A,r}(\mathbf{q}_t) + \mathbf{h}_{(I,r)}^H(\mathbf{q}_t) \bar{\Theta}_t^c \mathbf{g}) \sqrt{p_t} s_t + n_t, \quad (4.10)$$

trong đó,  $p_t$  là công suất truyền từ AP đến robot di động,  $s_t$  là tín hiệu truyền đến robot di động và  $n_t \sim \mathcal{CN}(0, 1)$  là nhiễu Gaussian (additive white Gaussian noise - AWGN) tại robot di động với công suất trung bình  $\sigma^2$  tại khe thời gian  $t$ . Khi đó, SINR tại robot di động là

$$\zeta_t^c = \frac{|h_{A,r}(\mathbf{q}_t) + \mathbf{h}_{(I,r)}^H(\mathbf{q}_t)\bar{\Theta}_t^c \mathbf{g}|^2 p_t}{\sigma^2}. \quad (4.11)$$

Do đó, tốc độ truyền thông đạt được tại robot di động là

$$R_t^c = \log_2 \left( 1 + \frac{|h_{A,r}(\mathbf{q}_t) + \mathbf{h}_{(I,r)}^H(\mathbf{q}_t)\bar{\Theta}_t^c \mathbf{g}|^2 p_t}{\sigma^2} \right). \quad (4.12)$$

#### 4.1.4 Bài toán điều hướng tối ưu cho robot di động

Như đã trình bày ở Chương 1, năng lượng đối với robot di động là một vấn đề quan trọng do robot di động sử dụng nguồn rời. Hơn thế nữa, vấn đề năng lượng hiện nay là một vấn đề được quan tâm không chỉ từ các nhà nghiên cứu mà còn từ các doanh nghiệp và cá nhân mỗi thành viên sống trên Trái đất. Do đó, ở chương này, luận án quan tâm đến việc điều hướng cho robot di động nhằm tối thiểu năng lượng tiêu thụ của hệ thống. Năng lượng tiêu thụ của hệ thống bao gồm năng lượng tiêu thụ bởi robot di động và năng lượng tiêu thụ bởi AP. Đầu tiên, xét đến năng lượng tiêu thụ bởi robot di động. Năng lượng tiêu thụ bởi robot di động tương tự như đã trình bày ở phần 3.3.1 và có tính toán đến năng lượng động học cho việc quay của robot di động. Cụ thể, năng lượng động học của robot di động được biểu diễn như sau [41]

$$E_{k,t} = \frac{1}{2}mv^2 + \frac{1}{2}I\omega_t^2, \quad (4.13)$$

trong đó,  $m$  và  $I$  là khối lượng và mô men quán tính của robot di động,  $\omega_t = \frac{\varphi(\mathbf{q}_{t+1}) - \varphi(\mathbf{q}_t)}{\tau}$  là vận tốc góc của robot di động với

$$\varphi(\mathbf{q}_{t+1}) - \varphi(\mathbf{q}_t) = \arccos \frac{(x_{t+1} - x_t)(x_{t+2} - x_{t+1}) + (y_{t+1} - y_t)(y_{t+2} - y_{t+1})}{(x_{t+1} - x_t)^2 + (y_{t+1} - y_t)^2}. \quad (4.14)$$

Và tổng động năng của robot di động là

$$E_k = \sum_{t=1}^{T-1} E_{k,t} = \frac{1}{2} \sum_{l=1}^{T-1} mv^2 + \frac{1}{2} \sum_{t=1}^{T-1} I\omega_t^2. \quad (4.15)$$

Năng lượng mất mát do ma sát trong quá trình robot di động di chuyển được mô tả dưới dạng [41]

$$E_{friction} = 2\mu mg \sum_{t=1}^{T-1} |\mathbf{q}_{t+1} - \mathbf{q}_t|, \quad (4.16)$$

trong đó,  $g$  là gia tốc trọng trường, và  $\mu$  là hệ số ma sát trượt phụ thuộc vào dạng bề mặt môi trường robot di động làm việc. Do đó, năng lượng tiêu thụ cho hệ thống chuyển động của robot di động được mô tả dạng [41]

$$\begin{aligned} E_{motion} &= E_k + E_{friction} \\ &= \frac{1}{2} \sum_{t=1}^{T-1} mv^2 + \frac{1}{2} \sum_{t=1}^{T-1} I\omega_t^2 + 2\mu mg \sum_{t=1}^{T-1} |\mathbf{q}_{t+1} - \mathbf{q}_t|, \end{aligned} \quad (4.17)$$

Ngoài ra, robot di động còn tiêu hao một phần không nhỏ năng lượng cho máy tính, các mạch điện tử, cảm biến. Các thành phần này được mô tả bởi [41]

$$E_s = P_s L \tau, \quad (4.18)$$

trong đó,  $P_s$  là điện áp nguồn của hệ thống.

Tổng năng lượng tiêu thụ của robot di động có thể được viết lại dưới dạng sau:

$$\begin{aligned} E_R &= E_{motion} + E_s \\ &= \frac{1}{2} \sum_{t=1}^{T-1} mv^2 + \frac{1}{2} \sum_{t=1}^{T-1} I\omega_t^2 + 2\mu mg \sum_{t=1}^{T-1} |\mathbf{q}_{t+1} - \mathbf{q}_t| + P_s L \tau, \end{aligned} \quad (4.19)$$

Năng lượng tiêu thụ của AP được mô tả dạng

$$E_A = \sum_{l=1}^T p_l \tau. \quad (4.20)$$

Đặt  $\mathbf{Q} = [\mathbf{q}_1, \mathbf{q}_2, \dots, \mathbf{q}_t]$  và  $\mathbf{p} = [p_1, p_2, \dots, p_t]$  là quỹ đạo của robot di động và công suất truyền từ AP đến robot di động tương ứng với quỹ đạo đó. Với mục tiêu tối thiểu năng lượng tiêu thụ của hệ thống, bài toán tối ưu được xây dựng như sau:

$$\min_{\mathbf{Q}, \mathbf{p}, \Theta^c} W_1 E_R + W_2 E_A \quad (4.21a)$$

thoả mãn: (4.3), (4.4), (4.5),

$$R_t^c \geq \bar{\gamma}, \forall t \in \mathcal{T}, \quad (4.21b)$$

$$\hat{T} \leq \hat{T}^{\max}, \quad (4.21c)$$

$$p_t \leq p^{\max}, \forall t \in \mathcal{T}, \quad (4.21d)$$

$$\phi_{t,n}^c = \frac{\xi \pi}{2^{b-1}}, \xi \in \{0, 1, \dots, 2^b - 1\}, \forall n \in \mathcal{N}, t \in \mathcal{T}, c \in \{tr, re\} \quad (4.21e)$$

trong đó  $W_1$  và  $W_2$  là các trọng số hàm mục tiêu lần lượt cho năng lượng tiêu thụ bởi robot di động và AP,  $\bar{\gamma}$  là tốc độ truyền thông tối thiểu tại robot di động,  $p^{\max}$  là công suất truyền tối đa của AP. Theo đó, ràng buộc (4.5) là điều kiện uni-modulus cho các thành phần của STAR-IRS. Ràng buộc (4.21b) là yêu cầu đảm bảo về chất lượng truyền thông cho robot di động. Ràng buộc (4.4) là ràng buộc tránh vật cản. Ràng buộc (4.21d) là ràng buộc công suất của AP. Ràng buộc (4.3) đảm bảo vị trí điểm đầu và điểm cuối của robot di động. Ràng buộc (4.21e) là đảm bảo về ngưỡng khả thi về pha của các thành phần của STAR-IRS.

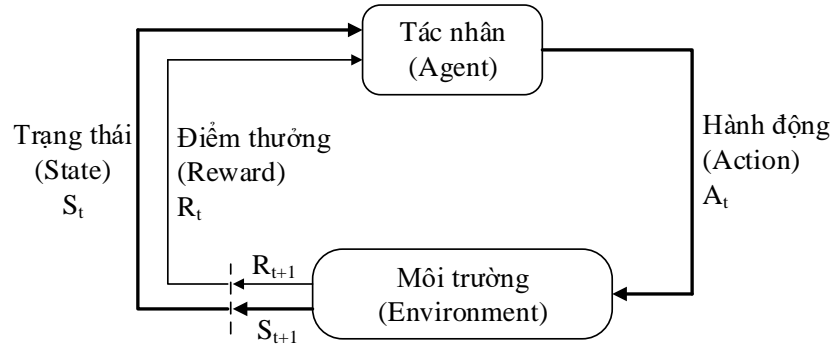
Bài toán (4.21) là bài toán không lồi. Hơn thế nữa, đây là một bài toán động do có nhiều tham số thay đổi theo thời gian như vị trí của robot di động, kênh truyền tương ứng,... cùng với các hàm về tốc độ truyền và kênh truyền phức tạp. Việc giải bài toán là một thách thức lớn đối với những thuật toán truyền thống như tối ưu lồi. Do đó, một thuật toán dựa trên kỹ thuật học sâu tăng cường. Phần tiếp theo sẽ trình bày tổng quan về kỹ thuật học tăng cường, sau đó thực hiện mô hình hóa bài toán (4.21) dưới dạng quy trình quyết định Markov MDP.

## 4.2 Mô hình quy trình quyết định MDP (Markov Decision Process)

### 4.2.1 Học tăng cường RL (Reinforcement Learning)

Theo các tài liệu học thuật về kỹ thuật học tăng cường RL (Reinforcement Learning), ý tưởng về RL đến từ chính bản chất về việc học một cách tự nhiên của con người: Chúng ta học bằng cách phản ứng lại với môi trường xung quanh chúng ta [10]. Các đặc tính đầy đủ của một bài toán RL nằm ở việc điều khiển tối ưu quy trình quyết định Markov MDPs (Markov Decision Processes).

Trong RL, đối tượng học và đưa ra các quyết định được gọi là tác nhân (agent). Tại thời điểm  $t$ , sự nhận thức của tác nhân về môi trường được thể hiện qua trạng thái (state)  $s_t \in \mathcal{S}$ , trong đó  $\mathcal{S}$  là tập hợp các trạng thái có thể có. Dựa vào trạng thái này, tác nhân sẽ lựa chọn một hành động (action)  $a_t \in \mathcal{A}(s_t)$ , trong đó  $\mathcal{A}(s_t)$  là tập hợp



Hình 4.4: Tương tác giữa tác nhân và môi trường trong học tăng cường RL. [10]

các hành động có sẵn ở trạng thái  $s_t$ . Kết quả của hành động này là tác nhân sẽ nhận được một điểm thưởng (reward)  $r_{t+1} \in \mathcal{R} \subset \mathbb{R}$  và đồng thời tác nhân sẽ chuyển sang trạng thái  $s_{t+1}$ . Hình 4.4 thể hiện sự tương tác giữa tác nhân với môi trường [10]. Việc quyết định hành động nào là do chính sách (policy). Chính sách  $\pi(a|s)$  là khả năng lựa chọn hành động  $a \in \mathcal{A}$  tại trạng thái  $s \in \mathcal{S}$ . Chính sách  $\pi$  thường được biểu diễn bởi một hàm không lồi với tham số là  $\theta$  nên từ đây, ký hiệu  $\pi_\theta$  được sử dụng để chỉ hàm chính sách với tham số  $\theta$ .

Ngoài ra,  $V(s)$  và  $Q(s, a)$  lần lượt là hàm giá trị của một trạng thái  $s$  (state-value function) và hàm giá trị của một hành động  $a$  (action-value function) dưới chính sách  $\pi$ .

Nhiệm vụ RL là đảm bảo đặc tính Markov được gọi là quy trình quyết định Markov MDP. Với bất kỳ mỗi cặp trạng thái  $s$  và hành động  $a$ , xác suất cho mỗi cặp trạng thái tiếp theo  $s'$  và điểm thưởng  $r$  được biểu diễn bởi

$$p(s', r|s, a) = Pr(S_{t+1} = s', R_{t+1} = r|S_t = s, A_t = a), \quad (4.22)$$

Một trong những vấn đề thách thức khi sử dụng RL là sự đánh đổi giữa việc thăm dò (exploration) và việc khai thác (exploitation). Để thu được điểm thưởng cao, tác nhân của RL phải quan tâm đến các hành động đã thử trong quá khứ và tìm ra hành động mang lại điểm thưởng lớn. Nhưng để khám phá, tác nhân cần phải thử các hành động mà chưa được lựa chọn trước đó. Tức là, tác nhân phải khai thác (exploit) những gì đã biết để thu được điểm thưởng, nhưng nó cũng cần phải thăm dò (explore) để tạo ra những hành động tốt hơn trong tương lai. Vấn đề ở đây là nếu không thăm dò cũng không khai thác thì sẽ không thể hoàn thành nhiệm vụ. Tác nhân phải thử nhiều hành động rồi từ từ thấy được hành động tốt. Với mỗi nhiệm vụ ngẫu nhiên,

mỗi hành động phải được thử nghiệm nhiều lần để có được ước lượng đáng tin cậy về điểm thưởng.

Khi kết hợp với các mạng nơ ron sâu (deep neural networks) [106], học tăng cường RL trở thành học sâu tăng cường DRL. Phương pháp on-policy sử dụng để đánh giá hoặc cải tiến chính sách đưa ra các quyết định. Ngược lại, phương pháp off-policy đánh giá hoặc cải tiến một chính sách khác với chính sách sử dụng để tạo ra dữ liệu. Phương pháp học xấp xỉ cho cả hàm chính sách và hàm giá trị được gọi là phương pháp actor-critic và các phiên bản mở rộng của nó [107, 108]. Ở đây, Actor để chỉ việc học chính sách và được sử dụng để lựa chọn hành động (actions) để thực hiện, và Critic là chỉ việc ước lượng hàm giá trị được sử dụng để phán xét hành động tạo bởi Actor. Việc học ở đây luôn luôn là on-policy: mạng critic phải học và phán xét chính sách mà mạng actor đưa ra. Việc bình xét thông qua dạng lỗi TD (temporal-difference). Đây là dạng tín hiệu đầu ra của critic và nó điều hướng việc học trong cả actor và critic. Lỗi TD được mô tả bởi

$$\delta_t = r_{t+1} + \gamma V_t(s_{t+1}) - V(s_t), \quad (4.23)$$

trong đó,  $V_t$  là hàm giá trị được tạo ra bởi critic tại thời điểm  $t$ ,  $\gamma$  is hệ số chiết khấu (discount factor). Lỗi TD này có thể được sử dụng để đánh giá hành động vừa lựa chọn, tức là hành động  $a_t$  ở trạng thái  $s_t$ . Nếu lỗi TD là dương, điều này gợi ý rằng xu hướng lựa chọn hành động  $a_t$  là tốt, trong khi đó, nếu lỗi TD này âm sẽ gợi ý rằng đây là một lựa chọn không tốt.

#### 4.2.2 Mô hình MDP cho bài toán điều hướng tối ưu

Bài toán cho nhiệm vụ lập quỹ đạo chuyển động của robot di động robot di động có thể được mô tả bởi một bài toán tối ưu được định nghĩa bởi  $\langle \mathcal{S}, \mathcal{A}, \mathcal{P}, \mathcal{R} \rangle$ , trong đó,  $\mathcal{S}$ ,  $\mathcal{A}$ ,  $\mathcal{P}$ , và  $\mathcal{R}$  lần lượt là không gian trạng thái (state space), không gian hành động (action space), hàm xác suất chuyển trạng thái (state transition probability function) và hàm thưởng (reward function).

#### Không gian trạng thái (State Space)

Việc lập quỹ đạo cho robot di động phụ thuộc vào vị trí hiện tại của robot di động, trạng thái phân bố của các vật cản xung quanh robot di động, và độ lớn của các kênh truyền trực tiếp và các kênh truyền hay phản xạ qua STAR-IRS. Do đó, không

gian trạng thái được định nghĩa như sau:

$$\mathcal{S}_t = \{\mathbf{q}_t, R_t^c, \mathcal{D}_t; \forall l \in \mathcal{L}\}, \quad (4.24)$$

trong đó,  $\mathcal{D}_t = d_{l,o_i}, \forall o_i \in \mathcal{M}$  là tập các khoảng cách từ robot di động đến các vật cản trong phạm vi phát hiện tại khe thời gian  $t$ .

### Không gian hành động (Action space)

Tại mỗi khe thời gian, hướng di chuyển của robot di động, cấu hình của STAR-IRS, và công suất truyền của AP cần được quyết định. Do đó, không gian hành động được định nghĩa như sau:

$$\mathcal{A}_t = \{\varphi(\mathbf{q}_t), p_t, \phi_{t,n}^r, \phi_{t,n}^t, \beta_{t,n}^r, \beta_{t,n}^t, \forall l \in \mathcal{L}, n \in \mathcal{N}\}, \quad (4.25)$$

trong đó,  $\varphi(\mathbf{q}_t) \in \{f_1, \dots, f_h, \dots, f_H\}$  là hướng xoay của robot di động,  $\beta_{t,n}^r \in \{\beta_1, \dots, \beta_j, \dots, \beta_J\}$  là hệ số độ lớn phản xạ/truyền của mỗi phần tử của STAR-IRS.

### Hàm điểm thưởng (Reward function)

Theo (4.21), hàm mục tiêu của bài toán tối ưu năng lượng tiêu thụ của hệ thống bao gồm năng lượng tiêu thụ bởi AP và robot di động, với các ràng buộc là việc cần đảm bảo yêu cầu về tốc độ truyền thông. Việc thiết kế điểm thưởng (reward) cho việc học cần phải nhất quán với mục tiêu. Do đó, NCS đưa ra một số điểm thưởng như sau:  $r_{E,t} = -W_1 E_R - W_2 E_A, r_{i,t} = -\bar{D}/\bar{D}_0, r_{QoS,t} = \lambda, r_s = 1000$ , trong đó,  $r_{E,t}$  đại diện cho chi phí mỗi bước di chuyển của robot di động,  $r_{i,t}$  là phần thưởng giúp robot di động tìm ra điểm đến, với  $\bar{D}$  là khoảng cách hiện tại giữa robot di động và điểm đích và  $\bar{D}_0$  là khoảng cách tối đa có thể giữa robot di động và điểm đích. Để làm cho robot di động đi hướng đến đích thay vì đi vòng, việc thiết kế điểm thưởng phải tuân thủ yêu cầu  $r_{E,t} + r_{i,t} < 0$ . Khi robot di động đến đích, nó sẽ nhận được điểm thưởng  $r_s$  và dừng lại. Ngoài ra, đối với ràng buộc QoS, robot di động nhận được một điểm thưởng  $r_{QoS} = \lambda$ , trong đó,  $\lambda = -1$  khi robot di động vi phạm yêu cầu về tốc độ truyền thông và ngược lại thì  $\lambda = 0$ .



Như vậy, hàm điểm thưởng được định nghĩa như sau:

$$r_t = \begin{cases} r_{E,t} + r_{i,t} + r_{QoS,t}, & \text{khi chưa đến đích} \\ r_{E,t} + r_{i,t} + r_{QoS,t} + r_s, & \text{khi đến đích} \end{cases} \quad (4.26)$$

### 4.3 Thuật toán học sâu tăng cường DRL (Deep Reinforcement Learning)

Để giải quyết bài toán tối ưu ngẫu nhiên được trình bày trong Mục 4.2.2, các phương pháp dựa trên giá trị của DRL như DDQN [109] thường được sử dụng. Tuy nhiên, các thuật toán này chỉ ước lượng hàm giá trị (hàm Q) và chọn các hành động tối ưu hóa giá trị Q tại mỗi bước. Do chính sách chưa được huấn luyện, các thuật toán dựa trên giá trị đòi hỏi một số lượng lớn mẫu trong quá trình khám phá môi trường. Ngược lại, các phương pháp Actor-critic, điển hình là thuật toán Advantage Actor-critic (A2C) [?], kết hợp kỹ thuật gradient với các hàm giá trị đã được học. Mạng actor định nghĩa một chính sách thực hiện các hành động dựa trên trạng thái, trong khi mạng critic đóng vai trò như một hàm giá trị (tương tự như các thuật toán dựa trên giá trị). Mạng critic đánh giá các hành động của actor và cải thiện chính sách thông qua gradient chính sách. Phương pháp này giúp thuật toán A2C hiệu quả hơn so với các phương pháp truyền thống dựa trên giá trị. Bằng cách học trực tiếp từ môi trường, nó rút ngắn quá trình khám phá và cải thiện hiệu quả huấn luyện. Tuy nhiên, là một thuật toán dựa trên chính sách, A2C liên tục cập nhật chính sách dựa trên dữ liệu mới thu thập từ môi trường. Điều này có thể gây ra các tác động tiêu cực từ các cập nhật lớn cũng như làm quá trình huấn luyện bị ảnh hưởng đáng kể bởi hệ số học và hệ số chiết khấu. Để giải quyết những vấn đề này, thuật toán Proximal Policy Optimization (PPO) [110] cùng với cơ chế cắt bớt là một thuật toán hiệu quả. Độ ổn định của PPO đến từ hàm mục tiêu của nó, sử dụng cơ chế cắt bớt để ngăn chặn các cập nhật chính sách quá lớn có thể dẫn đến bất ổn. Do đó, luận án sử dụng thuật toán PPO để giải quyết bài toán điều hướng tối ưu cho robot di động. PPO là thuật toán dựa trên A2C nên sau đây, luận án trình bày từ thuật toán A2C để làm rõ hơn thuật toán PPO.

#### 4.3.1 Thuật toán A2C (Advantage Actor-Critic)

A2C là một thuật toán học tăng cường on-policy kết hợp kiến trúc actor-critic để giải quyết sự không ổn định của chính sách và giảm thiểu tác động của các tập mẫu

lớn. Thuật toán này dựa vào hai mạng nơ-ron: mạng actor có trách nhiệm lựa chọn hành động và mạng critic cung cấp phản hồi để cải thiện chính sách. Sự kết hợp của hai mạng này giúp cải thiện hiệu suất tổng thể của hệ thống học tăng cường thông qua tối ưu hóa chính sách hành động và ước lượng chính xác giá trị của trạng thái. Gọi  $\mathbf{w}_a$  và  $\mathbf{w}_c$  lần lượt là các tham số trọng số của mạng actor và mạng critic. Gọi  $A(s_t, a_t)$  là hàm tính toán ước lượng lợi thế (the advantage estimation) của cặp trạng thái-hành động (state-action pair)  $(s_t, a_t)$ :

$$A(s_t, a_t) = r_t + \gamma C'(s_{t+1}|\mathbf{w}_c) - C(s_t|\mathbf{w}_c), \quad (4.27)$$

trong đó  $r_t$  là điểm thưởng tức thời nhận được khi thực hiện hành động  $a_t$  ứng với trạng thái  $s_t$ ,  $\gamma$  là hệ số chiết khấu (discount factor),  $C'(s_{t+1}|\mathbf{w}_c)$  là giá trị ước lượng cho trạng thái tiếp theo  $s_{t+1}$  sử dụng mạng Critic  $\mathbf{w}_c$ , và  $C(s_t|\mathbf{w}_c)$  là giá trị ước lượng cho trạng thái hiện tại  $s_t$  sử dụng mạng Critic với bộ tham số  $\mathbf{w}_c$ .

Hàm mất mát  $\mathcal{L}(\mathbf{w}_c)$  tính bằng sai số bình phương trung bình MSE (the mean squared error) giữa giá trị ước lượng lợi thế  $A(s_t, a_t)$  và giá trị mục tiêu bình phương của nó. Mục tiêu là tối thiểu hàm mất mát để nâng cao độ chính xác của giá trị ước lượng lợi thế. Do đó, về mặt toán học, hàm mất mát này được xác định bởi công thức

$$\mathcal{L}(\mathbf{w}_c) = \frac{1}{\bar{B}} \sum_{\bar{b}=1}^{\bar{B}} (A(s_t, a_t))^2, \quad (4.28)$$

trong đó  $\bar{B}$  là kích thước lô dữ liệu (batch) và  $\bar{b}$  là mẫu thứ  $\bar{b}$ -th trong lô dữ liệu.

Hàm mục tiêu được sử dụng để tối ưu hóa chính sách được định nghĩa bởi

$$J(\mathbf{w}_a) = \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t r_t \right]. \quad (4.29)$$

Gradient ascent của hàm mục tiêu của tác nhân  $J(\mathbf{w}_a)$  được thực hiện theo [111]:

$$\nabla_{\mathbf{w}_a} J(\mathbf{w}_a) = \frac{1}{\bar{B}} \sum_{\bar{b}=1}^{\bar{B}} \nabla_{\mathbf{w}_a} \log \pi(a_t|s_t, \mathbf{w}_a) A(s_t, a_t), \quad (4.30)$$

trong đó,  $\nabla_{\mathbf{w}_a}$  là gradient theo các tham số của mạng Actor  $\mathbf{w}_a$ ,  $\pi(a_t|s_t, \mathbf{w}_a)$  là xác suất chính sách lựa chọn hành động  $a_t$  ở trạng thái  $s_t$  với  $\mathbf{w}_a$  cho trước. Các hàm (4.28) và (4.30) sau đó được thực hiện để cập nhật các tham số của mạng Actor và Critic

trong quá trình huấn luyện nhằm tối đa hóa phần thưởng tích lũy kỳ vọng.

$$\mathbf{w}_c = \mathbf{w}_c + \alpha \nabla_{\mathbf{w}_c} \mathcal{L}(\mathbf{w}_c), \quad (4.31)$$

$$\mathbf{w}_a = \mathbf{w}_a + \beta \nabla_{\mathbf{w}_a} J(\mathbf{w}_a), \quad (4.32)$$

trong đó,  $\alpha$  và  $\beta$  lần lượt là tốc độ học (learning rate) của mạng Critic và mạng Actor.

### 4.3.2 Thuật toán PPO (Proximal Policy Optimization)

Do các ràng buộc bổ sung trong tối ưu hóa chính sách, PPO đặc biệt hiệu quả với mục tiêu đạt được sự cân bằng giữa cải thiện chính sách và ổn định trong quá trình huấn luyện. PPO thực hiện điều này bằng cách sử dụng kỹ thuật cắt giảm (clipping) để giới hạn mức độ thay đổi của chính sách trong mỗi bước cập nhật. Điều này giúp cải thiện đáng kể sự ổn định của quá trình học. Cụ thể, việc cập nhật chính sách dựa trên kỹ thuật cắt bớt được biểu diễn bởi:

$$\mathbf{w}_a^{\text{update}} = \mathbf{w}_a^{\text{old}} + \text{clip}(\mathbf{w}_a^{\text{new}}, 1 - \epsilon, 1 + \epsilon) \quad (4.33)$$

trong đó,  $\mathbf{w}_a^{\text{new}}$  là chính sách update mới trước khi thực hiện cắt giảm,  $\epsilon$  là tham số để điều chỉnh mức độ cắt giảm. Hàm  $\text{clip}(\cdot)$  với tham số  $\epsilon$  được sử dụng để giới hạn mức độ thay đổi chính sách, giới hạn này trong khoảng  $[1 - \epsilon, 1 + \epsilon]$ . Như vậy, việc hạn chế sự thay đổi được thực hiện đối với chính sách trong mỗi vòng huấn luyện. Điều này giúp loại bỏ động lực để chính sách hiện tại đi lệch quá xa so với chính sách cũ và đảm bảo rằng chính sách mới vẫn nằm trong một vùng an toàn giúp cho thuật toán vẫn có thể học hiệu quả.

Tại mỗi bước trong tập huấn luyện, AP đóng vai trò như một tác nhân và chọn một hành động  $a_t$  từ trạng thái  $s_t$ . Phần thưởng  $r_t$  mà AP nhận được từ hệ thống được xác định theo (4.26) và trạng thái hiện tại chuyển sang trạng thái mới  $s_{t+1}$ .

Đối với mỗi trạng thái  $s_t$  và hành động  $a_t$ , thuật toán tính đạo hàm của hàm mục tiêu theo trọng số của mạng Actor. Tổng của các đạo hàm này được tính cho tất cả các trạng thái và hành động. Cuối cùng, chính sách mới  $\mathbf{w}_a^{\text{new}}$  được xác định theo

$$\mathbf{w}_a^{\text{new}} = \frac{\partial}{\partial \mathbf{w}} \sum_{s,a} \mathbf{w}_a(s_t, a_t) A(s_t, a_t). \quad (4.34)$$

Hàm mất mát của mạng Actor trong PPO được thiết kế để tránh hiện tượng "over-

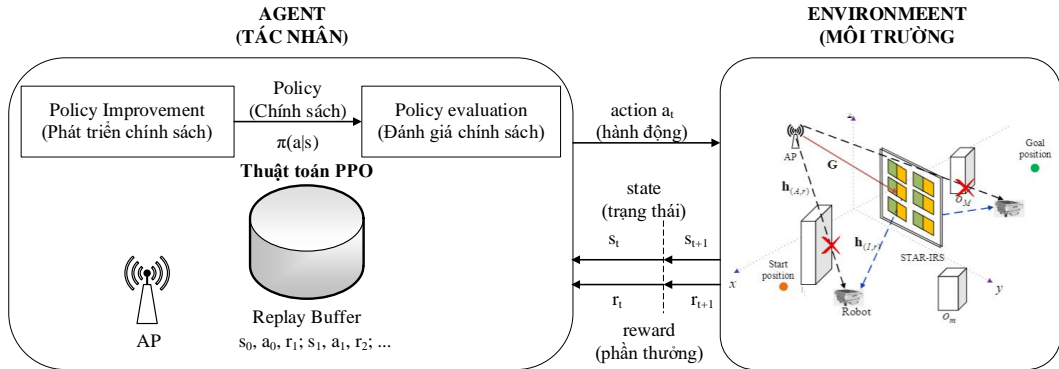
shooting" (cập nhật tham số quá lớn, có thể dẫn đến một chính sách không tối ưu). Đây được gọi là hàm mục tiêu thay thế đã bị cắt bớt (clipped surrogate objective function) và được xác định bởi [110]

$$J(\mathbf{w}_a) = \mathbb{E}_t \left[ A(s_t, a_t) \min \left( p_t(\mathbf{w}_a), \text{clip}(p_t(\mathbf{w}_a), 1 - \epsilon, 1 + \epsilon) \right) \right], \quad (4.35)$$

trong đó,  $p(\mathbf{w}_a) = \frac{\pi(a_t|s_t, \mathbf{w}_a)}{\pi(a_t|s_t, \mathbf{w}_a^{\text{old}})}$  là tỷ lệ xác suất giữa chính sách cũ và chính sách mới,  $\pi(a_t|s_t, \mathbf{w}_a)$  và  $\pi(a_t|s_t, \mathbf{w}_a^{\text{old}})$  lần lượt là chính sách lựa chọn hành động  $a_t$  tại trạng thái  $s_t$  hiện tại và trước đó, hàm  $\text{clip}(p_t(\mathbf{w}_a), 1 - \epsilon, 1 + \epsilon)$  được định nghĩa như sau:

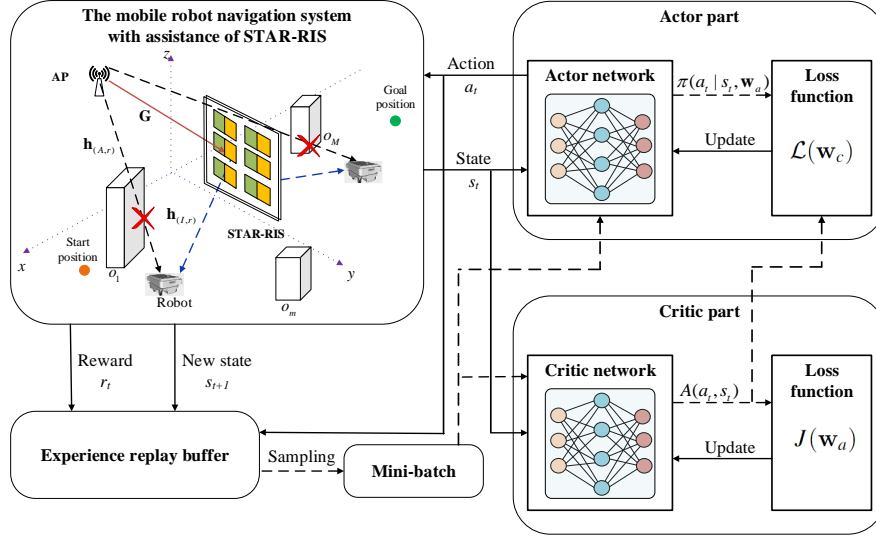
$$\text{clip}(p_t(\mathbf{w}_a), 1 - \epsilon, 1 + \epsilon) = \begin{cases} 1 - \epsilon, & \text{nếu } p(\mathbf{w}_a) < 1 - \epsilon, \\ 1 + \epsilon, & \text{nếu } p(\mathbf{w}_a) > 1 + \epsilon, \\ p(\mathbf{w}_a), & \text{các trường hợp còn lại.} \end{cases} \quad (4.36)$$

Phần tử clip này có nhiệm vụ giữ cho chính sách mới và chính sách cũ không có sự khác biệt quá lớn, và giữ cho tỷ lệ  $p(\mathbf{w}_a)$  luôn nằm trong khoảng  $[1 - \epsilon, 1 + \epsilon]$ . Hình 4.6 mô tả tổng quan thuật toán PPO áp dụng cho bài toán lập quỹ đạo cho robot di động. Chi tiết quy trình thực hiện thuật toán PPO thể hiện trên Hình 4.6.



Hình 4.5: Tổng quan thuật toán PPO bài toán điều hướng tối ưu cho robot di động.

Thuật toán 8 cho thấy các bước thực hiện của thuật toán PPO cho việc kết hợp tối ưu công suất truyền của AP, quỹ đạo chuyển động của robot di động, hệ số truyền và phản xạ của STAR-IRS để tối thiểu năng lượng tiêu thụ của hệ thống trong khi vẫn đảm bảo tốc độ truyền thông tối thiểu, công suất truyền lớn nhất của AP, và robot di động di chuyển được đến đích và không va chạm với vật cản.



Hình 4.6: Quy trình quản lý tài nguyên thuật toán PPO áp dụng cho bài toán điều hướng tối ưu cho robot di động

### 4.3.3 Độ hội tụ và độ phức tạp của thuật toán dựa trên actor-critic

Tại mỗi khe thời gian, thuật toán actor-critic thực hiện hai bước: actor và critic. Bước actor nhằm mục tiêu cập nhật tham số  $\theta$  của chính sách  $\pi_\theta$  bằng cách tối đa hàm trung bình điểm thưởng theo thời gian  $J(\pi_\theta)$  [112]

$$J(\pi_\theta) = \lim_{T \rightarrow \infty} \mathbb{E} \left[ \frac{1}{T} \sum_{t=0}^{T-1} r_{t+1} | \pi_\theta \right] \quad (4.37)$$

bằng phương pháp dốc lạp (gradient ascent), trong đó,  $T$  là số tập huấn luyện (epoch). Do đó, cần phải xấp xỉ gradient của chính sách (policy gradient)  $\nabla_\theta J(\pi_\theta)$ . Tại mỗi bước, một xấp xỉ ngẫu nhiên (stochastic approximation) của  $\nabla_\theta J(\pi_\theta)$  được tạo ra với mẫu  $(s, a, s', r)$ . Tuy nhiên, gradient ngẫu nhiên này sẽ bị lệch do: (1) sai số ước lượng trung bình điểm thưởng, (2) sai số ước lượng các tham số  $w$  của hàm  $V(s)$ , (3) sai số xấp xỉ giá trị  $V(s)$  bằng việc tuyến tính hóa. Gọi giới hạn trên của bình phương sai số xấp xỉ giá trị  $V(s)$  là  $\varepsilon_{approx}$ . Khi đó, nếu  $\varepsilon_{approx}$  càng nhỏ thì thể hiện việc xấp xỉ càng hiệu quả. Bước critic nhằm xấp xỉ ngẫu nhiên cho bài toán đánh giá chính sách với mục tiêu tối thiểu hàm gần với hàm bình phương sai số Bellman [112] với chính sách  $\pi_\theta$ . Tốc độ hội tụ của thuật toán actor-critic được đưa ra ở [112] như sau:

$$\frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E} \|\nabla J(\pi_{\theta_t})\|_2^2 \leq \frac{\rho[J(\pi^*) - J(\pi_{\theta_0} + \sigma^2)]}{\sqrt{T}} + G_\pi^2 \varepsilon_\mu + G_\pi^2 \varepsilon_\omega + G_\pi^2 \varepsilon_{approx},$$

---

**Thuật toán 8** Thuật toán PPO-STAR-IRS điều hướng tối ưu cho robot di động trong môi trường truyền thông có STAR-IRS hỗ trợ

---

**Khởi tạo giá trị ban đầu:** Các trọng số mạng actor  $\mathbf{w}_a$ , các trọng số mạng critic  $\mathbf{w}_c$ , bộ nhớ Replay  $\mathcal{D}$ , số lần lặp tối đa  $T$ , hệ số chiết khấu  $\gamma$  và lô dữ liệu huấn luyện  $\mathcal{B}$ ;

- 1: **for** số lần lặp = 1 đến  $T$  **do**
- 2:   Xóa lô dữ liệu huấn luyện  $\mathcal{B}$
- 3:   **for** mỗi bước **do**
- 4:     Quan sát trạng thái  $s_t = \{\mathbf{q}_t, R_t^c, \mathcal{D}_t\}$
- 5:     Chọn và thực hiện các hành động điều hướng cho robot và điều khiển tham số của STAR-IRS  $a_t = \{f(\mathbf{q}_t), p_t, \phi_{t,n}^{re}, \phi_{t,n}^{tr}, \beta_{t,n}^{re}, \beta_{t,n}^{tr}\}$  sử dụng mạng Actor.
- 6:     Thu được điểm thưởng  $r_t$ . Sau đó, môi trường thay đổi sang trạng thái  $s_{t+1} = \{\mathbf{q}_{t+1}, R_{t+1}^c, \mathcal{D}_{t+1}\}$
- 7:     Lưu  $(s_t, a_t, r_t, s_{t+1})$  vào  $\mathcal{B}$  và tính giá trị của  $A_t$  theo (4.27).
- 8:     **for** mỗi lô dữ liệu nhỏ trong  $\mathcal{B}$  **do**
- 9:        Tính  $\mathbf{w}_a^{\text{new}}$  sử dụng gradient descent theo (4.34)
- 10:        Tính hàm mất mát của mạng Critic theo (4.28)
- 11:        Tính hàm mất mát của mạng Actor theo (4.35)
- 12:        Cập nhật mạng Critic sử dụng gradient descent như sau:
- 13:         $\mathbf{w}_c = \mathbf{w}_c + \alpha \nabla_{\mathbf{w}_c} \mathcal{L}(\mathbf{w}_c)$
- 14:        Cập nhật mạng Actor bằng cách sử dụng kỹ thuật clipping như sau:
- 15:         $\mathbf{w}_a = \mathbf{w}_a^{\text{old}} + \text{clip}(\mathbf{w}_a^{\text{new}} - \mathbf{w}_a^{\text{old}}, -\epsilon, +\epsilon)$
- 16:     **end for**
- 17:   **end for**
- 18: **end for**

---

với những giả sử sau:

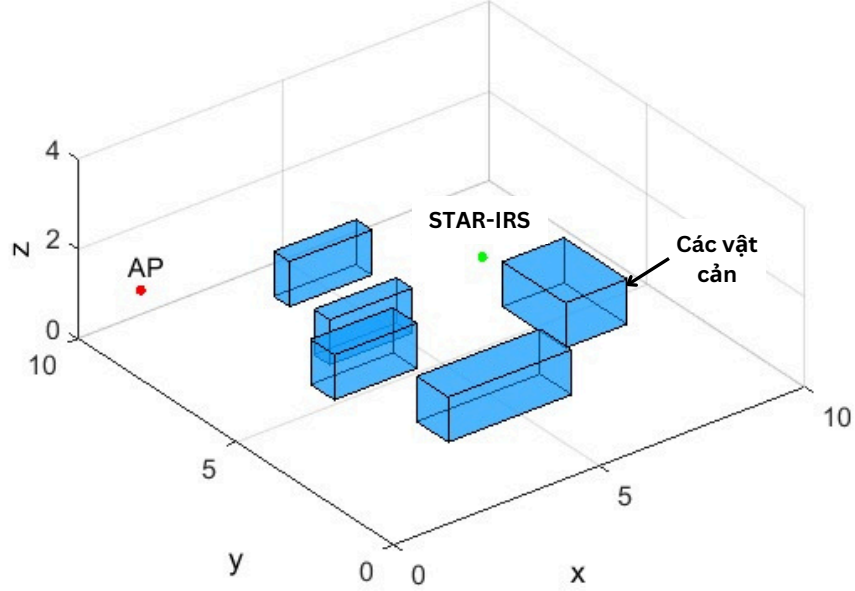
- Tồn tại một giá trị tối ưu của hàm giá trị  $J^* = J(\pi^*)$  là  $\pi^* = \text{argmax}_{\pi} J(\pi)$  với  $J^* \leq +\infty$ .
- Hàm  $\log_{\pi_{\theta}}(a|s)$  bị chặn:  $\|\nabla \log_{\pi_{\theta}}(a|s)\|_2 \leq G_{\pi}, \forall \theta, \forall s, \forall a$ .
- $\rho$ -Lipschitz gradient của  $J(\pi_{\theta})$ :  $\|\nabla_{\theta} J(\pi_{\theta_1}) - \nabla_{\theta} J(\pi_{\theta_2})\|_2 \leq \rho \|\theta_1 - \theta_2\|_2, \forall \theta_1, \forall \theta_2$ .
- Độ lệch chuẩn bị chặn:  $E_{(s,a,s')} \|g_{\theta}(\theta; s, a, s') - \nabla_{\theta} J(\pi_{\theta})\|_2^2 \leq \sigma^2, \forall \theta$

trong đó  $\epsilon$  là sai số ước lượng.

Do tồn tại tốc độ hội tụ nên có thể khẳng định rằng thuật toán sẽ hội tụ đến một điểm tối ưu  $J^*$ .

Độ phức tạp của thuật toán PPO phụ thuộc vào cấu trúc của mạng các mạng nơ ron bao gồm mạng actor và mạng critic. Cấu trúc của hai mạng nơ ron này bao gồm 1 lớp đầu vào, một lớp đầu ra và hai lớp ẩn. Gọi  $U_i^a$  và  $U_i^c$  lần lượt là số lượng nơ ron của mạng actor và mạng critic tại lớp  $i$ . Khi đó, độ phức tạp của thuật toán PPO được tính là  $\mathcal{O}(TK(\sum_{i=2}^4 (U_{i-1}^a U_i^a + U_{i-1}^c U_i^c)))$ , trong đó  $K$  là kích thước của lô dữ liệu training,  $T$  là tổng số lần training.

#### 4.4 Đánh giá hiệu quả mô hình và thuật toán điều hướng cho robot di động



Hình 4.7: Bản đồ môi trường làm việc của robot di động.

Luận án quan tâm đến môi trường trong nhà máy với diện tích hình vuông và kích thước mỗi chiều là 10 m và chiều cao là 5 m như mô tả trên Hình 4.7. Cụ thể, AP và STAR-IRS lần lượt đặt tại vị trí (0, 5, 2) m và (5, 5, 2) m. Vị trí điểm bắt đầu và điểm đích đến của robot di động lần lượt là (1.5, 1.5, 0.5) m và (8, 8, 0.5) m. Số lượng các phần tử của IRS là  $N = 16$ , có 5 vật cản trong môi trường làm việc của robot di động. Kênh LoS và mất mát đường truyền tính theo dB theo [6] như sau:

$$\mathbf{PL}_{LoS} = 31.84 + 21.50 \log_{10}(d_{ru}) + 19 \log_{10}(f_c), \quad (4.38)$$

trong đó,  $d_{ru}$  là khoảng cách 3D tính từ vị trí của ăng ten của robot di động đến AP (hoặc STAR-IRS), và  $f_c = 2$  GHz là tần số sóng mang. Với kênh NLoS, mất mát đường truyền tính theo dB như sau

$$\mathbf{PL}_{NLoS} = \max\{L_{LoS}, 32.4 + 23 \log_{10}(d_{ru}) + 20 \log_{10}(f_c)\}, \quad (4.39)$$

đảm bảo  $\mathbf{PL}_{NLoS} \geq \mathbf{PL}_{LoS}$ . Hệ số Rician của tất cả các kênh liên quan được đặt là 3 dB trong trường hợp các kênh không bị các vật cản che khuất. Thêm nữa, tốc độ truyền thông tại robot di động yêu cầu tối thiểu là  $\bar{\gamma} = 2.5$  bps/Hz, và công suất truyền

tối đa của AP là  $p^{\max} = 10$  dBm.

Trong không gian làm việc của robot di động bao gồm 5 chướng ngại vật với kích thước lần lượt là  $(3 \times 1 \times 1)$ ,  $(2 \times 0.75 \times 1)$ ,  $(2 \times 0.5 \times 1)$ ,  $(1.5 \times 2 \times 1)$ ,  $(2 \times 0.5 \times 1)$ . Vị trí tâm của các chướng ngại vật trên sàn lần lượt là  $(4.75, 3)$ ,  $(3.5, 5.5)$ ,  $(4.25, 6.5)$ ,  $(8, 5)$ ,  $(5, 8.75)$  m. Chiều cao của ăng ten robot di động là  $z_0 = 0.5$ m.

Để thấy được hiệu quả của thuật toán và hệ thống đề xuất sử dụng, hơn thế nữa, để thuận tiện cho việc theo dõi khi so sánh, luận án xem xét đến thuật toán đề xuất và các thuật toán cơ sở sau:

- **PPO-STAR-IRS:** Đây là thuật toán đề xuất dựa trên kỹ thuật học sâu tăng cường DRL được mô tả ở phần 4.3 (Thuật toán 8) được đề xuất sử dụng để giải bài toán điều hướng tối ưu cho hệ thống có STAR-IRS hỗ trợ.
- **A2C-STAR-IRS:** Đây là thuật toán A2C (Advantage Actor Critic) áp dụng cho hệ thống điều hướng cho robot di động có STAR-IRS hỗ trợ, tương tự như thuật toán PPO-STAR-IRS nhưng hàm mục tiêu không có hàm  $\text{clip}(p(\mathbf{w}_a), 1-\epsilon, 1+\epsilon)$ .
- **PPO-IRS:** Thuật toán này tương tự như thuật toán PPO-STAR-IRS nhưng thực hiện cho hệ thống sử dụng bề mặt chỉ phản xạ IRS hỗ trợ.
- **PPO-non-IRS:** Thuật toán này tương tự như thuật toán PPO-STAR-IRS nhưng sử dụng cho hệ thống không có IRS hỗ trợ.

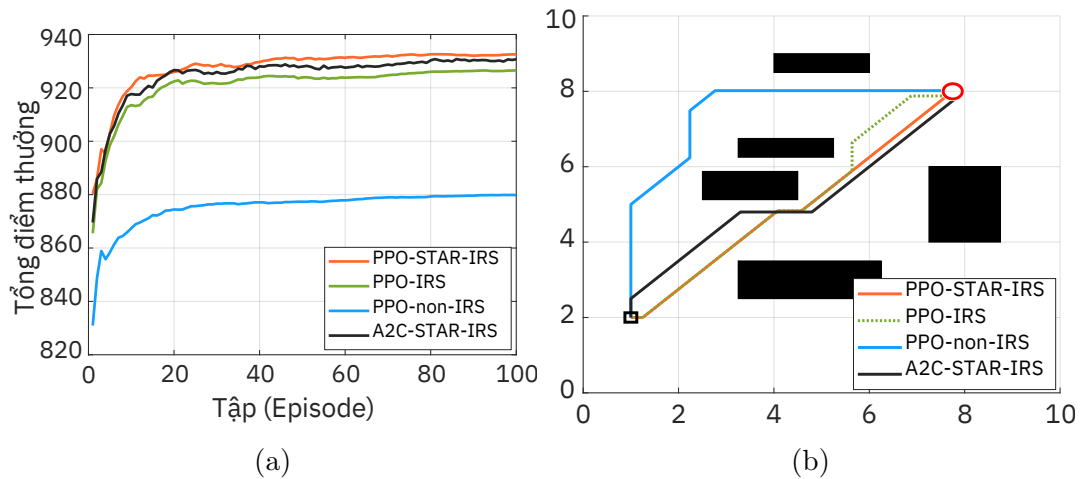
Khi thực hiện mô phỏng, các thuật toán dựa trên PPO và thuật toán dựa trên A2C kể trên đều sử dụng 2 mạng actor và critic, trong đó mỗi mạng đều có 2 lớp ẩn (hidden layer), mỗi lớp gồm có 1024 nút (node). Theo đó, các kết quả mô phỏng sẽ được phân tích ở dưới đây.

Bảng 4.1: Bảng tham số mô phỏng thuật toán điều hướng tối ưu cho robot di động

Tham số	Giá trị	Tham số	Giá trị
$p^{\max}$	20dBm	$N$	16
$d_r$	0.3 m	$\tau$	0.6 s
$v$	0.5 m/s	$\mu$	0.05
$g$	$9.8m/s^2$	$m$	9 kg [41]
$I$	$0.16245kgm^2$	$P_s$	17.8 W
$\mathbf{q}_s$	(1, 2)	$\mathbf{q}_g$	(8, 7.75)

Đầu tiên, luận án xem xét độ hội tụ của các thuật toán. Như thể hiện trên Hình 4.8(a), tất cả các thuật toán đều hội tụ tại mức giá trị điểm thưởng ổn định sau khoảng 50 tập (epoch). Đáng chú ý là thuật toán PPO-STAR-IRS thể hiện tính vượt





Hình 4.8: a) Độ hội tụ của các thuật toán và b) Quỹ đạo tối ưu thu được bởi các thuật toán.

trội hơn ba thuật toán còn lại, gồm PPO-IRS, PPO-non-IRS và A2C-STAR-IRS, về tổng điểm thưởng. Cụ thể, tổng điểm thưởng đạt được bởi PPO-STAR-IRS vượt qua tổng điểm thưởng đạt được bởi A2C-STAR-IRS, PPO-IRS và vượt qua đáng kể so với điểm thưởng thu được bởi PPO-non-IRS. Các kết quả này rõ ràng chứng minh sự hiệu quả của hệ thống và thuật toán đề xuất. Sự cải thiện điểm thưởng của thuật toán PPO-STAR-IRS dẫn đến quỹ đạo di chuyển tốt hơn cho robot. Quỹ đạo tương ứng của bốn thuật toán được thể hiện trong Hình 4.8(b). Trong hình này, các ô đen là các khu vực được che phủ bởi các vật cản. Điểm khởi đầu và điểm đích của robot di động được ký hiệu bằng " $\square$ " và " $\circ$ ", tương ứng. Có thể thấy rằng quỹ đạo thu được bởi thuật toán PPO-STAR-IRS là hiệu quả nhất và cho phép robot di chuyển với quãng đường ngắn nhất. Với thuật toán PPO-non-IRS, robot di động phải đi theo một đường vòng lên phía trước của AP để đạt được yêu cầu về tốc độ truyền thông do có các vật cản giữa vị trí ban đầu và vị trí đích. Điều này dẫn đến robot phải di chuyển theo một quỹ đạo dài hơn khi sử dụng thuật toán PPO-non-IRS. Ngược lại, việc phản xạ và truyền tín hiệu từ STAR-IRS nâng cao chất lượng truyền thông giữa AP và robot di động trong thuật toán PPO-STAR-IRS. Do đó, robot di động di chuyển gần như trực tiếp từ điểm ban đầu đến điểm đích. Với thuật toán PPO-IRS, tín hiệu chỉ phản xạ mà không truyền qua IRS, quỹ đạo từ điểm ban đầu đến IRS tương tự như được thu được bởi thuật toán PPO-STAR-IRS. Tuy nhiên, vì không có tín hiệu truyền qua IRS, robot di động phải di chuyển vòng qua phía trước của IRS để đáp ứng yêu cầu về tốc độ truyền thông. Điều này cũng dẫn đến một quỹ đạo dài hơn khi sử dụng PPO-IRS. Ngoài ra, thuật toán dựa trên PPO sử dụng phương pháp giới hạn (clipping) để điều chỉnh cập

Bảng 4.2: Năng lượng tiêu thụ bởi các thuật toán

Thuật toán	$E_A$ (J)	$E_R$ (KJ)
PPO-STAR-IRS	1.28	12.92
A2C-STAR-IRS	1.31	13.87
PPO-IRS	1.43	14.98
PPO-non-IRS	1.81	25.1

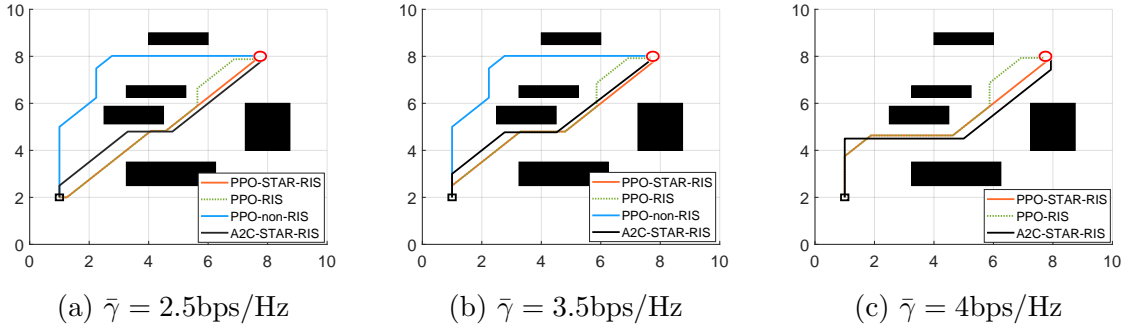
Bảng 4.3: Tổng quãng đường di chuyển của robot với các thuật toán

Thuật toán	$D$ (m)
PPO-STAR-IRS	9
A2C-STAR-IRS	9.5
PPO-IRS	10
PPO-non-IRS	11.5

nhật, ngăn chặn các thay đổi quá mức nhằm cân bằng giữa khám phá và khai thác. Do đó, quỹ đạo thu được từ PPO-STAR-IRS tốt hơn so với A2C-STAR-IRS. Tổng thể, hệ thống được hỗ trợ bởi STAR-IRS chứng minh hiệu quả quỹ đạo vượt trội. Đặc biệt, sử dụng thuật toán dựa trên PPO cho kết quả tốt hơn A2C.

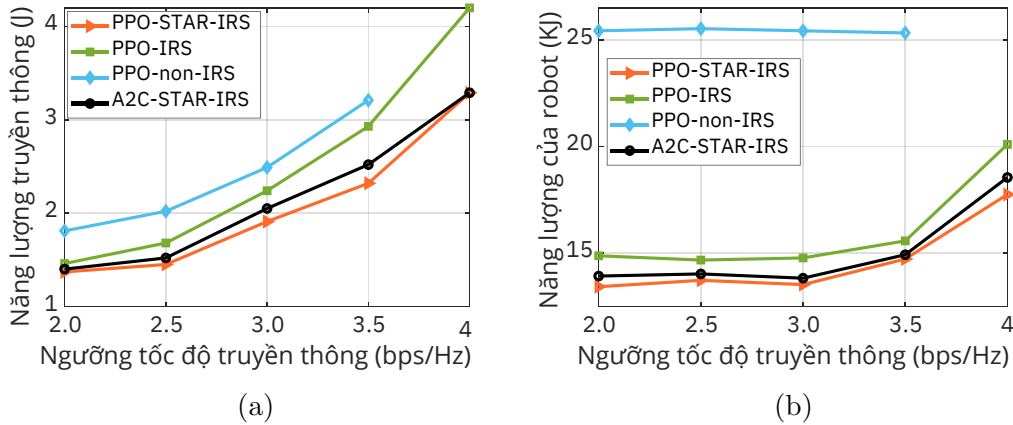
Các số liệu trong Bảng 4.2 và Bảng 4.3 cung cấp thông tin rõ ràng về hiệu quả của hệ thống và thuật toán đề xuất. Cụ thể, mức tiêu thụ năng lượng truyền thông khi sử dụng PPO-STAR-IRS lần lượt giảm 29.3%, 10.5%, và 2.3% so với PPO-non-IRS, PPO-IRS, và A2C-STAR-IRS. Hơn nữa, PPO-STAR-IRS giúp tiết kiệm năng lượng tiêu thụ của robot lên tới 48.5%, 13.8%, và 6.8% khi lần lượt so sánh với PPO-non-IRS, PPO-IRS, và A2C-STAR-IRS. Ngoài ra, tổng quãng đường di chuyển khi sử dụng PPO-STAR-IRS giảm 21.7%, 10%, và 5.3% hi lần lượt so sánh với PPO-non-IRS, PPO-IRS, và A2C-STAR-IRS. Những kết quả này thể hiện rõ ràng rằng hiệu suất của hệ thống hỗ trợ STAR-IRS luôn vượt trội hơn so với các hệ thống IRS thông thường hoặc không có IRS, và thuật toán dựa trên PPO có hiệu suất cao hơn so với thuật toán dựa trên A2C. Điều này sẽ được làm rõ hơn qua các kết quả sau đây.

Tiếp theo, luận án xem xét tác động của yêu cầu về ngưỡng tốc độ truyền thông cho robot di động ( $\bar{\gamma}$ ) lên quỹ đạo tối ưu của robot di động. Như thể hiện trên Hình 4.9, khi  $\bar{\gamma}$  tăng lên, robot gặp nhiều thách thức hơn trong việc tìm quỹ đạo đáp ứng yêu cầu này. Điều này dẫn đến robot di động phải di chuyển một quãng đường dài hơn. Sự khác biệt này thể hiện thể hiện rõ trên các Hình 4.9(a), (b) và (c). Đáng lưu ý là khi  $\bar{\gamma}$  tăng lên đến một mức nhất định (cụ thể ở đây là 4.0 bps/Hz), hệ thống không có IRS hỗ trợ với thuật toán PPO-non-IRS không thể tìm thấy bất kỳ đường đi nào đáp ứng được yêu cầu đó. Ngoài ra, trong tất cả các kịch bản, quỹ đạo thu được bởi



Hình 4.9: Ảnh hưởng của ngưỡng yêu cầu tốc độ truyền thông đến quỹ đạo tối ưu của robot di động.

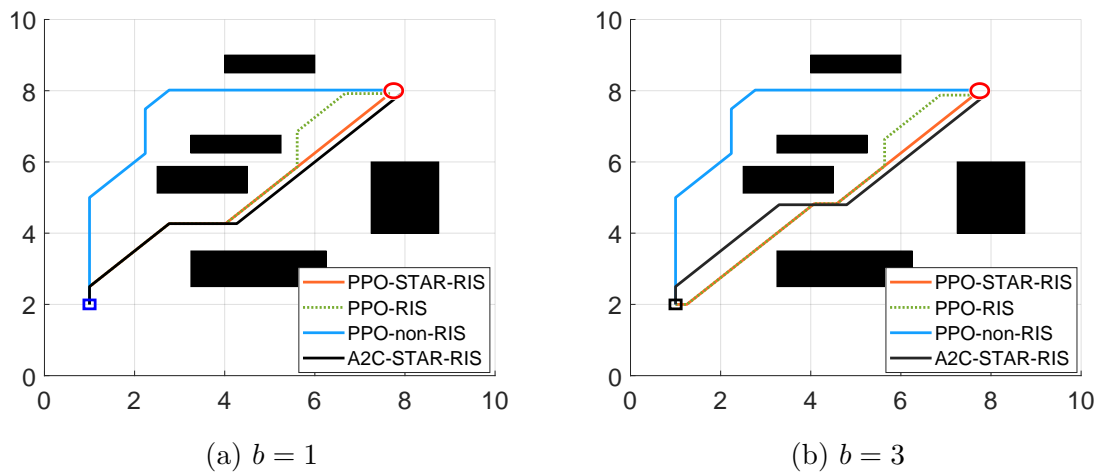
PPO-STAR-IRS luôn là ngắn nhất, và quỹ đạo thu được bởi PPO-non-IRS luôn là dài nhất.



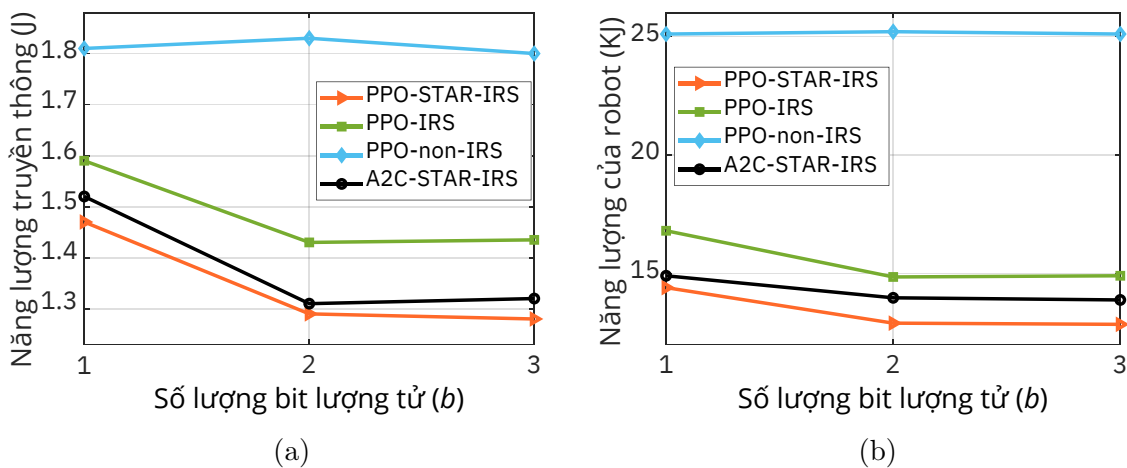
Hình 4.10: Ảnh hưởng của ngưỡng yêu cầu tốc độ truyền thông lên a) năng lượng truyền thông và b) năng lượng tiêu thụ bởi robot di động.

Hình 4.10(a) và (b) lần lượt thể hiện sự ảnh hưởng của  $\bar{\gamma}$  đối với năng lượng truyền thông và năng lượng tiêu thụ của robot di động. Trong Hình 4.10(a), có thể thấy rằng năng lượng truyền thông tăng khi  $\bar{\gamma}$  tăng. Lý do là khi  $\bar{\gamma}$  tăng, AP phải sử dụng công suất cao hơn để đáp ứng yêu cầu này. Tuy nhiên, năng lượng tiêu thụ cho việc truyền thông với thuật toán PPO-STAR-IRS luôn là thấp nhất. Với thuật toán PPO-non-IRS, AP cần sử dụng công suất cao nhất để đáp ứng cùng ngưỡng tốc độ yêu cầu. Ngoài ra, trong Hình 4.10(b), năng lượng tiêu thụ của robot di động gần như không thay đổi khi  $\bar{\gamma}$  tăng đến 3 bps/Hz. Lý do là khi sử dụng công suất cao hơn tại AP, yêu cầu về  $\bar{\gamma}$  đã được đáp ứng. Do đó, quỹ đạo của robot di động thu được bởi mỗi thuật toán là quỹ đạo tối ưu trong mỗi trường hợp. Tuy nhiên, khi  $\bar{\gamma}$  tiếp tục tăng (3.5 bps/Hz hoặc 4.0 bps/Hz) thì mức năng lượng tiêu thụ của robot tăng lên đáng

kể. Điều này có thể lý giải là do khi đó, robot cần di chuyển một quãng đường xa hơn đáng kể để đạt được mức yêu cầu về  $\bar{\gamma}$  dẫn đến năng lượng tiêu thụ tăng. Mặc dù vậy, trong tất cả các trường hợp, năng lượng tiêu thụ của hệ thống có STAR-IRS hỗ trợ luôn là thấp nhất, và năng lượng tiêu thụ của hệ thống không có IRS hỗ trợ luôn là cao nhất. Đặc biệt, thuật toán PPO-STAR-IRS giúp giảm mức tiêu thụ năng lượng truyền thông và năng lượng của robot lên tới 25% và 50% so với thuật toán PPO-non-IRS. Hơn nữa, mức tiêu thụ năng lượng cho liên lạc và robot thu được từ A2C-STAR-IRS cao hơn so với từ PPO-STAR-IRS. Những kết quả này chứng minh hiệu quả của hệ thống và thuật toán đề xuất.



Hình 4.11: Ảnh hưởng của số bit lượng tử ( $b$ ) đến quỹ đạo của robot di động.

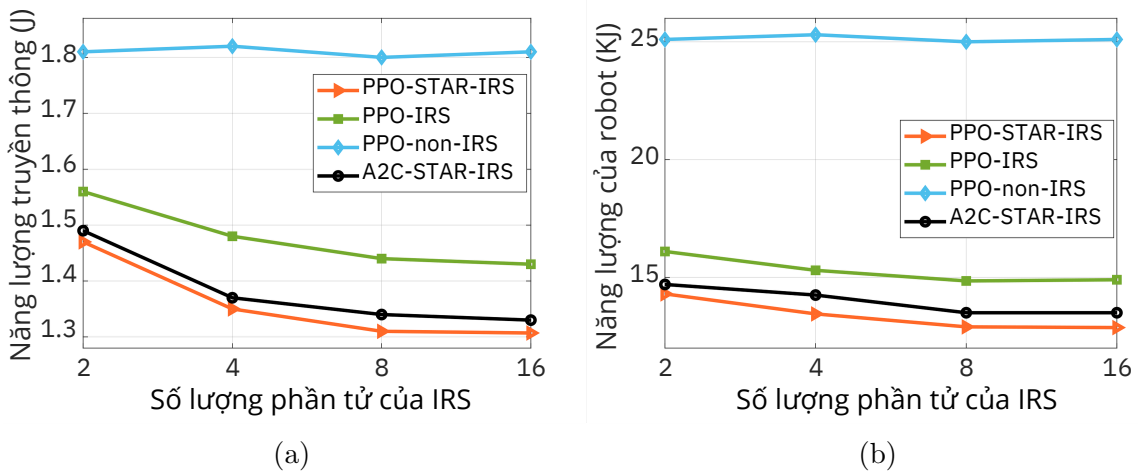


Hình 4.12: Ảnh hưởng của số bit lượng tử ( $b$ ) đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động.

Tiếp theo, luận án thảo luận về tác động của số lượng bit lượng tử độ dịch pha của STAR-IRS ( $b$ ) lên quỹ đạo của robot di động và mức tiêu thụ năng lượng của hệ

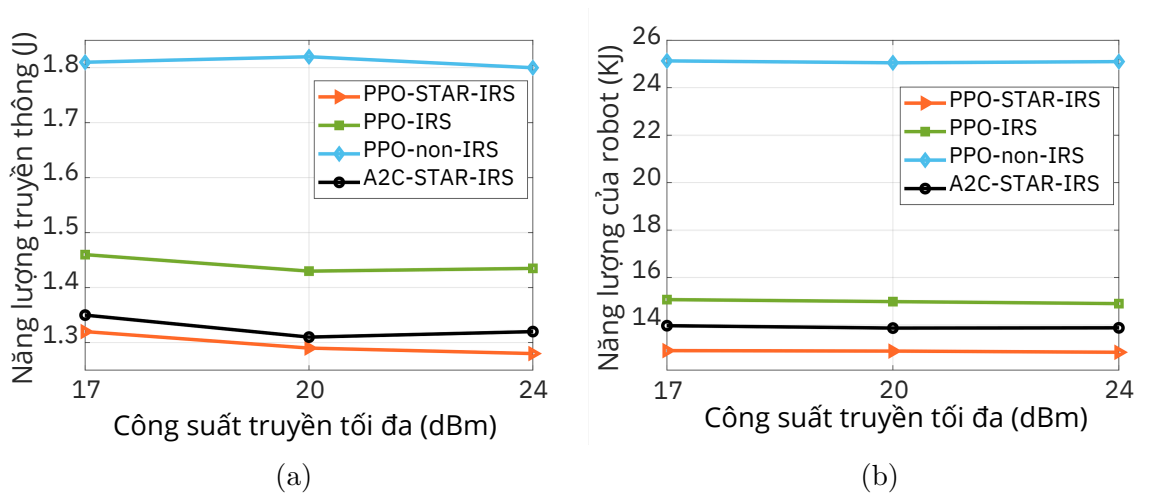
thống. Đầu tiên, như đã thấy Hình 74.11(a) và (b), với 1-bit lượng tử, robot di động cần di chuyển quãng đường dài hơn so với trường hợp 3-bit lượng tử để đáp ứng yêu cầu tốc độ truyền thông mong muốn. Nguyên nhân là khi số lượng bit lượng tử tăng lên, tức là độ phân giải của độ dịch pha tăng, thuật toán có thể tìm ra mức dịch pha gần tối ưu hơn đảm bảo đáp ứng yêu cầu tốc độ truyền thông với quãng đường ngắn hơn. Đây cũng là lý do giải thích cho các kết quả trong Hình 4.12(a) và (b). Cụ thể, cả mức tiêu thụ năng lượng truyền thông và năng lượng tiêu thụ của robot thu được từ PPO-STAR-IRS, A2C-STAR-IRS và PPO-IRS đều giảm dần khi số lượng bit lượng tử tăng lên. Bên cạnh đó, mức tiêu thụ năng lượng của PPO-non-IRS không thay đổi vì nó không phụ thuộc vào số lượng bit lượng tử. Hơn nữa, kết quả thu được với 2 hoặc 3-bit lượng tử không thay đổi do số lượng mức giá trị độ dịch pha với 2 hoặc 3-bit có thể đã đủ lớn để đạt được xấp xỉ kết quả tối ưu. Điều này cho thấy vai trò của độ phân giải khi lượng tử hóa độ dịch pha.

Một điểm đáng lưu ý là kích thước của hệ thống STAR-IRS ( $N$ ) cũng có thể ảnh hưởng đến năng lượng tiêu thụ của hệ thống. Như thấy trên Hình 4.13(a) và (b), khi  $N$  tăng lên, năng lượng tiêu thụ bởi AP và robot di động giảm đi. Lý do là khi số lượng phần tử STAR-IRS tăng lên, SINR tại robot di động cũng tăng lên và tốc độ truyền thông của nó cũng tăng. Do đó, năng lượng tiêu thụ do truyền thông giảm đi, và việc đáp ứng ngưỡng tốc độ truyền thông trở nên dễ dàng hơn khi  $N$  tăng lên. Điều này cho phép robot di động di chuyển một cách linh hoạt hơn với mức năng lượng tiêu thụ thấp hơn.



Hình 4.13: Ảnh hưởng của số phần tử của STAR-IRS đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động.

Cuối cùng, luận án trình bày ảnh hưởng của công suất truyền tối đa ( $p^{\max}$ ) lên



Hình 4.14: Ảnh hưởng của công suất truyền tối đa của AP đến a) năng lượng truyền thông và b) năng lượng tiêu thụ của robot di động.

mức tiêu thụ năng lượng truyền thông và năng lượng của robot trong các thuật toán như Hình 4.14(a) và (b). Hình 4.14(a), khi  $p^{\max}$  tăng, mức tiêu thụ năng lượng truyền thông thu được từ tất cả các thuật toán nhìn chung đều giảm. Nguyên nhân là việc tăng  $p^{\max}$  cho phép robot tìm quỹ đạo với thời gian di chuyển ngắn hơn  $\hat{T}$ , dẫn đến giảm mức tiêu thụ năng lượng truyền thông của AP. Vì lý do này, mức tiêu thụ năng lượng của robot cũng giảm khi  $p^{\max}$  (Hình 4.14(b)). Đặc biệt, với PPO-STAR-IRS, mức tiêu thụ năng lượng của robot giảm từ 12.92 KJ xuống 12.85 KJ khi  $p^{\max}$  tăng từ 17 dBm đến 24 dBm. Cần nhấn mạnh rằng với bất kỳ giá trị nào của  $p^{\max}$ , mức tiêu thụ năng lượng thu được từ PPO-STAR-IRS luôn thấp nhất so với các thuật toán cơ sở.

#### 4.5 Kết luận chương 4

Dựa trên việc phân tích các nghiên cứu về việc hỗ trợ nâng cao chất lượng mạng truyền thông không dây của STAR-IRS, chương 4 đã đề xuất một mô hình hệ thống và thuật toán điều hướng tối ưu cho robot di động trong nhà máy với môi trường truyền thông không dây có STAR-IRS hỗ trợ. Cụ thể, những đóng góp chính của chương 4 như sau:

- Đề xuất mô hình hệ thống điều hướng cho robot di động trong môi trường nhà máy có STAR-IRS hỗ trợ.
- Đề xuất hệ thuật toán PPO-STAR-IRS để giải quyết bài toán lập quỹ đạo nhằm tối thiểu năng lượng tiêu thụ cho robot di động trong môi trường truyền thông

có STAR-IRS hỗ trợ.

Các đóng góp của chương 4 được công bố ở các công trình CT6 và CT7. Công trình CT6 trình bày việc đề xuất mô hình điều hướng cho robot di động trong môi trường truyền thông không dây có STAR-IRS hỗ trợ, từ đó đề xuất thuật toán tối ưu hệ thống nhằm tối thiểu quãng đường di chuyển cho robot di động dựa trên kỹ thuật DRL. Dựa trên các kết quả ban đầu đạt được từ công bố CT6, công trình CT7 phát triển thuật toán tối ưu hệ thống điều hướng cho robot di động trong môi trường truyền thông có STAR-IRS hỗ trợ nhằm tối thiểu năng lượng tiêu thụ cho hệ thống.

[CT6] **N. T. T. Van**, N. M. Tien, N. C. Luong, and L. M. Khoi, "Simultaneously Transmitting And Reflecting (STAR) IRS Enhanced Mobile robot di động Path Planning: A Reinforcement Learning based Approach," in Proceedings of the seventh Vietnam international conference and Exhibition on Control and Automation, VCCA-2024; ISBN 978-604-937-357-2.

[CT7] **N. T. T. Van**, N. C. Luong, H. Le Hung, N. T. Hoa and N. M. Tien, "Minimizing Energy Consumption in Mobile robot di độngics with STAR-RIS in Smart Factories," 2024 Tenth International Conference on Communications and Electronics (ICCE), Danang, Vietnam, 2024, pp. 741-746, doi: 10.1109/ICCE62051.2024.10634674.

## KẾT LUẬN VÀ KIẾN NGHỊ

Robot di động đã và đang được ứng dụng rộng rãi trong đời sống và thu hút được sự chú ý của nhiều nhà nghiên cứu. Theo xu hướng phát triển IoT, robot di động làm việc trong môi trường kết nối, vấn đề truyền thông và năng lượng tiêu thụ của robot là những vấn đề cần được quan tâm. Mặc dù mới dừng lại ở bước đầu nghiên cứu lý thuyết, các kết quả mô phỏng của luận án hứa hẹn mang lại hiệu quả ứng dụng cao.

### 1. Các kết quả đạt được của luận án

Luận án đã đạt được mục tiêu đã đề ra là nghiên cứu đề xuất mô hình mạng và thuật toán tối ưu hệ thống mạng sử dụng trong nhà máy, từ đó đề xuất hệ thống điều hướng cho robot nhằm tối thiểu năng lượng tiêu thụ cho robot di động và hệ thống trong môi trường truyền thông lý tưởng và trong môi trường truyền thông có STAR-IRS hỗ trợ. Ngoài ra, luận án cũng đã mô phỏng và đánh giá các thuật toán và mô hình hệ thống đề xuất.

Các đóng góp mới của luận án:

**1. Đề xuất mô hình mạng DED2D có IRS hỗ trợ và thuật toán tối ưu OTA và N-OTA nhằm nâng cao chất lượng truyền thông hỗ trợ hoạt động của robot di động.**

Mô hình hệ thống này được đề xuất dựa trên các nghiên cứu những lợi ích mà IRS mang lại cho hệ thống truyền thông không dây. Cụ thể, trong hệ thống mạng ở nhà máy với nhiều thiết bị bao gồm robot di động nhận thông tin, các cảm biến không dây thu năng lượng và các thiết bị truyền thông khác, việc sử dụng IRS và các thuật toán tối ưu các tham số hệ số đã không những nâng cao chất lượng truyền thông đảm bảo hoạt động cho robot di động mà còn hỗ trợ hoạt động các thiết bị không dây khác góp phần đảm bảo tính bền vững của hệ thống.

**2. Đề xuất sử dụng thuật toán OSOW và MSOW tối ưu quỹ đạo cho robot di động trong môi trường truyền thông lý tưởng và thuật toán PPO-STAR-IRS tối ưu năng lượng cho robot di động trong môi trường**



***truyền thông có STAR-IRS hỗ trợ.***

Thuật toán OSOW và MSOW dựa trên việc xấp xỉ hàm mục tiêu và các ràng buộc trở thành bài toán tối ưu lồi để giải. Kết quả mô phỏng đã tìm ra được quỹ đạo tối ưu giúp robot di chuyển đến đích với mức năng lượng tiêu thụ tối thiểu trong môi trường truyền thông lý tưởng. Thuật toán PPO-STAR-IRS dựa trên kỹ thuật DRL đã giúp giải quyết bài toán động với quỹ đạo và kênh thay đổi theo thời gian. Kết quả mô phỏng cho thấy hiệu quả của mô hình và thuật toán đề xuất trong việc điều hướng tối ưu cho robot di động.

*Ngoài ra, trong quá trình thực hiện luận án, các nghiên cứu đã được công bố ở các hội nghị khoa học và các tạp chí có uy tín trong nước và quốc tế.*

**2. Hướng phát triển của Luận án**

Một hệ thống điều hướng thông minh cho các robot trong các nhà máy là hướng phát triển của Luận án. Cụ thể, định hướng nghiên cứu trong tương lai của NCS bao gồm:

- Nghiên cứu, phát triển thuật toán điều hướng cho hệ đa robot có STAR-IRS hỗ trợ.
- Nghiên cứu, xây dựng hệ thống thực nghiệm để kiểm chứng các thuật toán đã mô phỏng.

## DANH MỤC CÔNG TRÌNH CỦA TÁC GIẢ

### TẠP CHÍ KHOA HỌC

[CT1] **N. T. T. Van**, H. T. Nguyen, N. C. Luong, N. M. Tien, D. Niyato, and D. I. Kim, “Intelligence Reflecting Surface-Aided Integrated Data and Energy Networking Coexisting D2D Communications,” *IEEE Trans. Wirel. Commun.*, pp. 1–1, 2022.

[CT2] **N. T. T. Van**, N. M. Tien, N. C. Luong, and H. T. K. Duyen, “Energy consumption minimization for autonomous mobile robot: A convex approximation approach,” *Journal of Robotics and Control (JRC)*, vol. 4, no. 3, pp. 403–412, 2023.

### HỘI NGHỊ KHOA HỌC

[CT3] **N. T. T. Van**, N. Manh Tien, H. T. K. Duyen, N. M. Cuong, and D. H. Duân, “A Convex Approximation Method to Optimize Trajectory for AGV Robot Motion Planning,” in *Proceedings of the sixth Vietnam international conference and Exhibition on Control and Automation VCCA-2021*; ISBN 978-604-95-0875-2.

[CT4] **N. T. T. Van**, N. Manh Tien, N. M. Cuong, H. T. K. Duyen, B. T. T. Ha, and B. V. Tuan, “Building SLAM system and Intelligent Navigation for Autonomous Mobile Robot base on ROS,” in *Proceedings of the sixth Vietnam international conference and Exhibition on Control and Automation, VCCA-2021*; ISBN 978-604-95-0875-2.

[CT5] **N. T. T. Van**, N. M. Tien, N. M. Cuong, H. T. K. Duyen, and N. D. Duy, “Constructing an Intelligent Navigation System for Autonomous Mobile Robot Based on Deep Reinforcement Learning”. Cham: Springer International Publishing, 2021, pp. 251–261. [Online]. Available: [https://doi.org/10.1007/978-3-030-76620-7\\_22](https://doi.org/10.1007/978-3-030-76620-7_22).

[CT6] **N. T. T. Van**, N. M. Tien, N. C. Luong, and L. M. Khoi, “Simultaneously Transmitting And Reflecting (STAR) IRS Enhanced Mobile Robot Path Planning: A Reinforcement Learning based Approach,” in *Proceedings of the seventh Vietnam international conference and Exhibition on Control and Automation, VCCA-2024*; ISBN 978-604-937-357-2.

[CT7] **N. T. T. Van**, N. C. Luong, H. Le Hung, N. T. Hoa and N. M.

Tien, "Minimizing Energy Consumption in Mobile Robotics with STAR-RIS in Smart Factories," 2024 Tenth International Conference on Communications and Electronics (ICCE), Danang, Vietnam, 2024, pp. 741-746, doi: 10.1109/ICCE62051.2024.10634674.

## TÀI LIỆU THAM KHẢO

- [1] H. T. K. Duyên, “Điều khiển mặt động thích nghi bám quỹ đạo cho robot tự hành bốn bánh đa hướng,” Luận án Tiến sỹ, Học viện Khoa học Công nghệ, viện Hàn lâm Khoa học và công nghệ Việt Nam, 2020.
- [2] H.-y. Zhang, W.-m. Lin, and A.-x. Chen, “Path planning for the mobile robot: A review,” *Symmetry*, vol. 10, no. 10, p. 450, 2018.
- [3] P. team Netmarks Vietnam, “Shi-v wifi system design proposal,” Tech. Rep., 2019. unpublished.
- [4] M. Noor-A-Rahim, F. Firyaguna, J. John, M. O. Khyam, D. Pesch, E. Armstrong, H. Claussen, and H. V. Poor, “Toward industry 5.0: Intelligent reflecting surface in smart manufacturing,” *IEEE Communications Magazine*, vol. 60, no. 10, pp. 72–78, 2022.
- [5] H. Ren, K. Wang, and C. Pan, “Intelligent reflecting surface-aided 5g-urllc in a factory automation scenario,” *IEEE Transactions on Communications*, vol. 70, no. 1, pp. 707–723, 2022.
- [6] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, “Intelligent reflecting surface enhanced indoor robot path planning: A radio map-based approach,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4732–4747, 2021.
- [7] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, “Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey,” *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, 2020.
- [8] Q. Wu, S. Zhang, B. Zheng, C. You, and R. Zhang, “Intelligent reflecting surface-aided wireless communications: A tutorial,” *IEEE Transactions on Communications*, vol. 69, no. 5, pp. 3313–3351, 2021.

- [9] X. Mu, Y. Liu, L. Guo, J. Lin, and R. Schober, “Simultaneously transmitting and reflecting (star) ris aided wireless communications,” *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3083–3098, 2022.
- [10] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. MIT press, 2018.
- [11] B. Holfeld, D. Wieruch, T. Wirth, L. Thiele, S. A. Ashraf, J. Huschke, I. Aktas, and J. Ansari, “Wireless communication for factory automation: an opportunity for lte and 5g systems,” *IEEE Communications Magazine*, vol. 54, no. 6, pp. 36–43, 2016.
- [12] B. Siciliano, O. Khatib, and T. Kröger, *Springer handbook of robotics*. Springer, 2008, vol. 200.
- [13] G. Hu, W. P. Tay, and Y. Wen, “Cloud robotics: architecture, challenges and applications,” *IEEE Network*, vol. 26, no. 3, pp. 21–28, 2012.
- [14] M. A. Contreras-Cruz, V. Ayala-Ramirez, and U. H. Hernandez-Belmonte, “Mobile robot path planning using artificial bee colony and evolutionary programming,” *Applied Soft Computing*, vol. 30, pp. 319–328, 2015.
- [15] E. W. Dijkstra, “A note on two problems in connexion with graphs,” in *Edsger Wybe Dijkstra: His Life, Work, and Legacy*, 2022, pp. 287–290.
- [16] M. C. Silva, A. C. B. Garcia, and A. Conci, “A multi-agent system for dynamic path planning,” in *2010 Second Brazilian Workshop on Social Simulation*. IEEE, 2010, pp. 47–51.
- [17] P. E. Hart, N. J. Nilsson, and B. Raphael, “A formal basis for the heuristic determination of minimum cost paths,” *IEEE transactions on Systems Science and Cybernetics*, vol. 4, no. 2, pp. 100–107, 1968.
- [18] A.-F. Contreras-González, J.-I. Hernández-Vega, C. Hernández-Santos, and D.-G. Palomares-Gorham, “A method to verify a path planning by a back-propagation artificial neural network.” in *LANMR*, 2016, pp. 98–105.
- [19] A. Stentz, “Optimal and efficient path planning for partially known environments,” in *Intelligent unmanned ground vehicles*. Springer, 1997, pp. 203–220.

- [20] W. Yijing, L. Zhengxuan, Z. Zhiqiang, and L. Zheng, “Local path planning of autonomous vehicles based on a\* algorithm with equal-step sampling,” in *2018 37th Chinese Control Conference (CCC)*. IEEE, 2018, pp. 7828–7833.
- [21] A. Ismail, A. Sheta, and M. Al-Weshah, “A mobile robot path planning using genetic algorithm in static environment,” *Journal of Computer Science*, vol. 4, no. 4, pp. 341–344, 2008.
- [22] P. Martin and A. Del Pobil, “Application of artificial neural networks to the robot path planning problem,” *WIT Transactions on Information and Communication Technologies*, vol. 6, 1970.
- [23] J.-P. Thomesse, “Fieldbus technology in industrial automation,” *Proceedings of the IEEE*, vol. 93, no. 6, pp. 1073–1101, 2005.
- [24] T. Raunholt, I. Rodriguez, P. Mogensen, and M. Larsen, “Towards a 5g mobile edge cloud planner for autonomous mobile robots,” in *2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall)*, 2021, pp. 01–05.
- [25] P. Galambos, “Cloud, fog, and mist computing: Advanced robot applications,” *IEEE Systems, Man, and Cybernetics Magazine*, vol. 6, no. 1, pp. 41–45, 2020.
- [26] N. V. Tinh, P. T. Cat, P. M. Tuan, and T. T. Q. Bui, “Visual control of integrated mobile robot-pan tilt-camera system for tracking a moving target,” in *2014 IEEE International Conference on Robotics and Biomimetics (ROBIO 2014)*, 2014, pp. 1566–1571.
- [27] C. V. Hoàng, “Thiết kế quỹ đạo cho xe tự hành và ứng dụng trên mô hình robot 3 bánh,” *Trường Đại học Bách khoa Hà Nội*, 2020.
- [28] T. V. Dang, “Autonomous mobile robot path planning based on enhanced a\* algorithm integrating with time elastic band,” *MM SCI. J*, vol. 2023, pp. 6717–6722, 2023.
- [29] N. M. Ngọc, N. Toàn, V. V. Quang, and N. N. Vũ, “Một tiếp cận tránh vật cản cho robot amr trong tự động hóa nhà kho,” *TNU Journal of Science and Technology*, vol. 228, no. 02, pp. 62–69, 2023.

- [30] L. A. Nguyen, T. D. Pham, T. D. Ngo, and X. T. Truong, “A proactive trajectory planning algorithm for autonomous mobile robots in dynamic social environments,” in *2020 17th International Conference on Ubiquitous Robots (UR)*, 2020, pp. 309–314.
- [31] L. A. Nguyen, T. D. Ngo, T. D. Pham, and X. T. Truong, “An efficient navigation system for autonomous mobile robots in dynamic social environments,” *International Journal of Robotics and Automation*, vol. 37, no. 1, pp. 97–106, 2022.
- [32] R. Robotin, G. Lazea, and C. Marcu, “Graph search techniques for mobile robot path planning,” in *Engineering the Future*. IntechOpen, 2010.
- [33] A. Yang, J. Pu, C. B. Wong, and P. Moore, “By-pass valve control to improve energy efficiency of pneumatic drive system,” *Control Engineering Practice*, vol. 17, no. 6, pp. 623–628, 2009.
- [34] R. Grebers, M. Gadaleta, A. Paugurs, A. Senfelds, A. Avotins, and M. Pellicciari, “Analysis of the energy consumption of a novel dc power supplied industrial robot,” *Procedia Manufacturing*, vol. 11, pp. 311–318, 2017.
- [35] W. Chen, J. Liu, S. Chen, and L. Zhang, “Energy shaping control for wireless power transfer system in automatic guided vehicles,” *Energies*, vol. 13, no. 11, p. 2959, 2020.
- [36] Y. Mei, Y.-H. Lu, Y. Hu, and C. Lee, “A case study of mobile robot’s energy consumption and conservation techniques,” in *ICAR ’05. Proceedings., 12th International Conference on Advanced Robotics, 2005.*, 2005, pp. 492–497.
- [37] H. Kim and B.-K. Kim, “Minimum-energy translational trajectory planning for battery-powered three-wheeled omni-directional mobile robots,” in *Proc. of the 10th International Conference on Control, Automation, Robotics and Vision*, pp. 1730–1735, 2008.
- [38] Y. Mei, Y.-H. Lu, C. G. Lee, and Y. C. Hu, “Energy-efficient mobile robot exploration,” in *Proc. of the IEEE International Conference on Robotics and Automation*, pp. 505–511, 2006.

- [39] I. Duleba and J. Z. Sasiadek, “Nonholonomic motion planning based on newton algorithm with energy optimization,” *IEEE transactions on control systems technology*, vol. 11, no. 3, pp. 355–363, 2003.
- [40] A. Pal, R. Tiwari, and A. Shukla, *Modified A\* Algorithm for Mobile Robot Path Planning*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2012, pp. 183–193. [Online]. Available: [https://doi.org/10.1007/978-3-642-25507-6\\_16](https://doi.org/10.1007/978-3-642-25507-6_16)
- [41] S. Liu and D. Sun, “Minimizing energy consumption of wheeled mobile robots via optimal motion planning,” *IEEE/ASME Transactions on Mechatronics*, vol. 19, no. 2, pp. 401–411, 2013.
- [42] Y. Liu, Z. Chen, Y. Li, M. Lu, C. Chen, and X. Zhang, “Robot search path planning method based on prioritized deep reinforcement learning,” *International Journal of Control, Automation and Systems*, vol. 20, no. 8, pp. 2669–2680, 2022.
- [43] S. Dogru and L. Marques, “Towards fully autonomous energy efficient coverage path planning for autonomous mobile robots on 3d terrain,” in *2015 European Conference on Mobile Robots (ECMR)*, 2015, pp. 1–6.
- [44] M. F. Jaramillo-Morales, S. Dogru, and L. Marques, “Generation of energy optimal speed profiles for a differential drive mobile robot with payload on straight trajectories,” in *2020 IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*, 2020, pp. 136–141.
- [45] C. H. Kim and B. K. Kim, “Energy-saving 3-step velocity control algorithm for battery-powered wheeled mobile robots,” in *Proceedings of the 2005 IEEE International Conference on Robotics and Automation*, 2005, pp. 2375–2380.
- [46] U. Orozco-Rosas, K. Picos, J. J. Pantrigo, A. S. Montemayor, and A. Cuesta-Infante, “Mobile robot path planning using a qapf learning algorithm for known and unknown environments,” *IEEE Access*, vol. 10, pp. 84 648–84 663, 2022.
- [47] H. Qin, B. Qiao, W. Wu, and Y. Deng, “A path planning algorithm based on deep reinforcement learning for mobile robots in unknown environment,” in *2022 IEEE 5th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC)*, vol. 5, 2022, pp. 1661–1666.



- [48] W. Zhang, W. Wang, H. Zhai, and Q. Li, "A deep reinforcement learning method for mobile robot path planning in unknown environments," in *2021 China Automation Congress (CAC)*, 2021, pp. 5898–5902.
- [49] H. Kabir, M.-L. Tham, and Y. C. Chang, "Internet of robotic things for mobile robots: Concepts, technologies, challenges, applications, and future directions," *Digital Communications and Networks*, vol. 9, no. 6, pp. 1265–1290, 2023. [Online]. Available: <https://www.sciencedirect.com/science/article/pii/S2352864823000883>
- [50] Y. Ren and V. Friderikos, "Path planning optimization based interference awareness for mobile robots in mmwave multi cell networks," *IEEE Transactions on Vehicular Technology*, pp. 1–12, 2024.
- [51] C. C. Ooi and C. Schindelbauer, "Minimal energy path planning for wireless robots," *Mobile Networks and Applications*, vol. 14, no. 3, pp. 309–321, 2009.
- [52] A. Ghaffarkhah and Y. Mostofi, "Optimal motion and communication for persistent information collection using a mobile robot," in *2012 IEEE Globecom Workshops*, 2012, pp. 1532–1537.
- [53] C. Tatino, N. Pappas, and D. Yuan, "Robot trajectory planning with qos constrained irs-assisted millimeter-wave communications," in *ICC 2021 - IEEE International Conference on Communications*, 2021, pp. 1–6.
- [54] Z. Liu, Y. Liu, and X. Chu, "Reconfigurable-intelligent-surface-assisted indoor millimeter-wave communications for mobile robots," *IEEE Internet of Things Journal*, vol. 11, no. 1, pp. 1548–1557, 2024.
- [55] J. Zhao, Y. Zhu, X. Mu, K. Cai, Y. Liu, and L. Hanzo, "Simultaneously transmitting and reflecting reconfigurable intelligent surface (star-ris) assisted uav communications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 10, pp. 3041–3056, 2022.
- [56] S. Zhang, Q. Wu, S. Xu, and G. Y. Li, "Fundamental green tradeoffs: Progresses, challenges, and impacts on 5g networks," *IEEE Communications Surveys Tutorials*, vol. 19, no. 1, pp. 33–56, 2017.

- [57] D. Muirhead, M. A. Imran, and K. Arshad, "A survey of the challenges, opportunities and use of multiple antennas in current and future 5g small cell base stations," *IEEE Access*, vol. 4, pp. 2952–2964, 2016.
- [58] P. Park, S. Coleri Ergen, C. Fischione, C. Lu, and K. H. Johansson, "Wireless network design for control systems: A survey," *IEEE Communications Surveys & Tutorials*, vol. 20, no. 2, pp. 978–1013, 2018.
- [59] C. Liaskos, S. Nie, A. Tsioliariidou, A. Pitsillides, S. Ioannidis, and I. Akyildiz, "Realizing wireless communication through software-defined hypersurface environments," in *2018 IEEE 19th International Symposium on "A World of Wireless, Mobile and Multimedia Networks" (WoWMoM)*, 2018, pp. 14–15.
- [60] Q. Wu and R. Zhang, "Towards smart and reconfigurable environment: Intelligent reflecting surface aided wireless network," *IEEE Communications Magazine*, vol. 58, no. 1, pp. 106–112, 2020.
- [61] C. Cai, H. Yang, X. Yuan, and Y.-C. Liang, "Two-timescale optimization for intelligent reflecting surface aided d2d underlay communication," in *IEEE Global Communications Conference*, 2020, pp. 1–6.
- [62] Q. Wu and R. Zhang, "Intelligent reflecting surface enhanced wireless network via joint active and passive beamforming," *IEEE Transactions on Wireless Communications*, vol. 18, no. 11, pp. 5394–5409, 2019.
- [63] A. H. Phan, H. D. Tuan, H. H. Kha, and D. T. Ngo, "Nonsmooth optimization for efficient beamforming in cognitive radio multicast transmission," *IEEE Transactions on Signal Processing*, vol. 60, no. 6, pp. 2941–2951, 2012.
- [64] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, "Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper gaussian signaling," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2589–2603, 2020.
- [65] S. Gong, X. Lu, D. T. Hoang, D. Niyato, L. Shu, D. I. Kim, and Y.-C. Liang, "Toward smart wireless communications via intelligent reflecting surfaces: A contemporary survey," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 4, pp. 2283–2314, 2020.

- [66] Y. Cao, T. Lv, W. Ni, and Z. Lin, “Sum-rate maximization for multi-reconfigurable intelligent surface-assisted device-to-device communications,” *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7283–7296, 2021.
- [67] S. Zargari, A. Khalili, Q. Wu, M. R. Mili, and D. W. K. Ng, “Max-min fair energy-efficient beamforming design for intelligent reflecting surface-aided swipt systems with non-linear energy harvesting model,” *IEEE Trans. on Vehicular Technology*, 2021.
- [68] Y. Tang, G. Ma, H. Xie, J. Xu, and X. Han, “Joint transmit and reflective beamforming design for irs-assisted multiuser miso swipt systems,” in *IEEE International Conference on Communications (ICC)*, 2020, pp. 1–6.
- [69] C. Pan, H. Ren, K. Wang, M. ElKashlan, A. Nallanathan, J. Wang, and L. Hanzo, “Intelligent reflecting surface aided mimo broadcasting for simultaneous wireless information and power transfer,” *IEEE J. Sel. Areas Commun.*, Feb. 2020.
- [70] Q. Wu and R. Zhang, “Weighted sum power maximization for intelligent reflecting surface aided swipt,” *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 586–590, 2020.
- [71] H. Yu, H. D. Tuan, T. Q. Duong, Y. Fang, and L. Hanzo, “Improper gaussian signaling for integrated data and energy networking,” *IEEE Transactions on Communications*, vol. 68, no. 6, pp. 3922–3934, 2020.
- [72] J. Hu, K. Yang, G. Wen, and L. Hanzo, “Integrated data and energy communication network: A comprehensive survey,” *IEEE Communications Surveys & Tutorials*, vol. 20, no. 4, pp. 3169–3219, 2018.
- [73] C. Xu, L. Song, Z. Han, Q. Zhao, X. Wang, and B. Jiao, “Interference-aware resource allocation for device-to-device communications as an underlay using sequential second price auction,” in *IEEE ICC*, 2012, pp. 445–449.
- [74] A. A. Nasir, H. D. Tuan, D. T. Ngo, T. Q. Duong, and H. V. Poor, “Beamforming design for wireless information and power transfer systems: Receive power-splitting versus transmit time-switching,” *IEEE Trans. Commun.*, vol. 65, no. 2, pp. 876–889, Nov. 2016.

- [75] H. Lee, S.-R. Lee, K.-J. Lee, H.-B. Kong, and I. Lee, “Optimal beamforming designs for wireless information and power transfer in miso interference channels,” *IEEE Trans. on Wireless Communications*, vol. 14, no. 9, pp. 4810–4821, 2015.
- [76] H. Yu, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, “Joint design of reconfigurable intelligent surfaces and transmit beamforming under proper and improper gaussian signaling,” *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 11, pp. 2589–2603, 2020.
- [77] I. Csiszar, “Information geometry and alternating minimization procedures,” *Statistics and decisions*, vol. 1, pp. 205–237.
- [78] H. Yu, H. D. Tuan, T. Q. Duong, Y. Fang, and L. Hanzo, “Improper Gaussian signaling for integrated data and energy networking,” *IEEE Trans. Commun.*, vol. 68, no. 6, pp. 3922–3934, Mar. 2020.
- [79] H. Tuy, *Convex Analysis and Global Optimization (second edition)*. Springer International, 2016.
- [80] D. Peaucelle, D. Henrion, Y. Labit, and K. Taitz, “User’s guide for SeDuMi interface 1.04,” 2002. [Online]. Available: <http://homepages.laas.fr/peaucell/software/sdmguide.pdf>
- [81] Y. Shi, H. D. Tuan, H. Tuy, and S. Su, “Global optimization for optimal power flow over transmission networks,” *Journal of Global Optimization*, vol. 69, no. 3, pp. 745–760, 2017.
- [82] G. Yang, Y. Liao, Y.-C. Liang, O. Tirkkonen, G. Wang, and X. Zhu, “Reconfigurable intelligent surface empowered device-to-device communication underlying cellular networks,” *IEEE Transactions on Communications*, vol. 69, no. 11, pp. 7790–7805, 2021.
- [83] S. Boyd and L. Vandenberghe, *Convex optimization*. Cambridge university press, 2004.
- [84] N. D. Phước, *Tối ưu hóa trong điều khiển và điều khiển tối ưu*. Nhà xuất bản Bách Khoa Hà Nội, 2020.
- [85] M. Grant, S. Boyd, and Y. Ye, “Cvx users’ guide,” *online: http://www.stanford.edu/boyd/software.html*, 2009.

- [86] M. Grant and S. Boyd, “CVX: Matlab software for disciplined convex programming, version 2.1,” <http://cvxr.com/cvx>, Mar. 2014.
- [87] D. Bertsekas, *Convex optimization algorithms*. Athena Scientific, 2015.
- [88] S. K. Malu and J. Majumdar, “Kinematics, localization and control of differential drive mobile robot,” *Global Journal of Research In Engineering*, 2014.
- [89] R. Fierro and F. L. Lewis, “Control of a nonholomic mobile robot: Backstepping kinematics into dynamics,” *Journal of robotic systems*, vol. 14, no. 3, pp. 149–163, 1997.
- [90] H. Tuy, *Convex analysis and global optimization*. Springer, 1998.
- [91] F. H. Ajeil, I. K. Ibraheem, M. A. Sahib, and A. J. Humaidi, “Multi-objective path planning of an autonomous mobile robot using hybrid pso-mfb optimization algorithm,” *Applied Soft Computing*, vol. 89, p. 106076, 2020.
- [92] Y. Mei, Y.-H. Lu, Y. C. Hu, and C. G. Lee, “Deployment of mobile robots with energy and timing constraints,” *IEEE Transactions on robotics*, vol. 22, no. 3, pp. 507–522, 2006.
- [93] L. Hou, L. Zhang, and J. Kim, “Energy modeling and power measurement for mobile robots,” *Energies*, vol. 12, no. 1, p. 27, 2018.
- [94] H. Ding, G. Reißig, D. Groß, and O. Stursberg, “Mixed-integer programming for optimal path planning of robotic manipulators,” in *Proc. of the IEEE International Conference on Automation Science and Engineering*, pp. 133–138, 2011.
- [95] H. P. Williams, *Model building in mathematical programming*. John Wiley & Sons, 2013.
- [96] M. Razaviyayn, “Successive convex approximation: Analysis and applications,” Ph.D. dissertation, University of Minnesota, 2014.
- [97] A. Ben-Tal and A. Nemirovski, *Lectures on Modern Convex Optimization*. Philadelphia: MPS-SIAM Series on Optimi., SIAM, 2001.
- [98] A. Beck, A. Ben-Tal, and L. Tetruashvili, “A sequential parametric convex approximation method with applications to nonconvex truss topology design problems,” *J. Global Optim.*, vol. 47, no. 1, pp. 29–51, May 2010.

- [99] M. Grant and S. Boyd, “Cvx: Matlab software for disciplined convex programming, version 2.1,” 2014.
- [100] S. Boyd and M. Grant, “Graph implementations for nonsmooth convex programs,” in *Recent Advances in Learning and Control*, ser. Lecture Notes in Control and Information Sciences, V. Blondel, S. Boyd, and H. Kimura, Eds. Springer-Verlag Limited, 2008, pp. 95–110, [http://stanford.edu/~boyd/graph\\_dcp.html](http://stanford.edu/~boyd/graph_dcp.html).
- [101] D. Adionel Guimaraes, G. H. Faria Floriano, and L. Silvestre Chaves, “A tutorial on the cvx system for modeling and solving convex optimization problems,” *IEEE Latin America Transactions*, vol. 13, no. 5, pp. 1228–1257, 2015.
- [102] J. Xu, Y. Liu, X. Mu, and O. A. Dobre, “Star-riss: Simultaneous transmitting and reflecting reconfigurable intelligent surfaces,” *IEEE Communications Letters*, vol. 25, no. 9, pp. 3134–3138, 2021.
- [103] Y. Liu, X. Mu, J. Xu, R. Schober, Y. Hao, H. V. Poor, and L. Hanzo, “Star: Simultaneous transmission and reflection for 360° coverage by intelligent surfaces,” *IEEE Wireless Communications*, vol. 28, no. 6, pp. 102–109, 2021.
- [104] B. O. Zhu, K. Chen, N. Jia, L. Sun, J. Zhao, T. Jiang, and Y. Feng, “Dynamic control of electromagnetic wave propagation with the equivalent principle inspired tunable metasurface,” *Scientific reports*, vol. 4, no. 1, p. 4971, 2014.
- [105] X. Gao, X. Mu, W. Yi, and Y. Liu, “Intelligent trajectory design for ris-noma aided multi-robot communications,” *arXiv preprint arXiv:2205.01647*, 2022.
- [106] I. Goodfellow, Y. Bengio, and A. Courville, *Deep learning*. MIT press, 2016.
- [107] V. Konda and J. Tsitsiklis, “Actor-critic algorithms,” in *Advances in Neural Information Processing Systems*, S. Solla, T. Leen, and K. Müller, Eds., vol. 12. MIT Press, 1999. [Online]. Available: [https://proceedings.neurips.cc/paper\\_files/paper/1999/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf](https://proceedings.neurips.cc/paper_files/paper/1999/file/6449f44a102fde848669bdd9eb6b76fa-Paper.pdf)
- [108] J. Schulman, S. Levine, P. Abbeel, M. Jordan, and P. Moritz, “Trust region policy optimization,” in *Proceedings of the 32nd International Conference on Machine Learning*, ser. Proceedings of Machine Learning Research, F. Bach and

- D. Blei, Eds., vol. 37. Lille, France: PMLR, 07–09 Jul 2015, pp. 1889–1897. [Online]. Available: <https://proceedings.mlr.press/v37/schulman15.html>
- [109] Y. Yuan, Z. Li, Z. Liu, Y. Yang, and X. Guan, “Double Deep Q-Network Based Distributed Resource Matching Algorithm for D2D Communication,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 1, pp. 984–993, 2022.
- [110] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, and O. Klimov, “Proximal policy optimization algorithms,” *arXiv preprint arXiv:1707.06347*, 2017.
- [111] Z. Chen, J. Hu, and G. Min, “Learning-based resource allocation in cloud data center using advantage actor-critic,” in *ICC 2019-2019 IEEE International Conference on Communications (ICC)*. IEEE, 2019, pp. 1–6.
- [112] S. Qiu, Z. Yang, J. Ye, and Z. Wang, “On finite-time convergence of actor-critic algorithm,” *IEEE Journal on Selected Areas in Information Theory*, vol. 2, no. 2, pp. 652–664, 2021.
- [113] Z. Sheng, H. D. Tuan, A. A. Nasir, T. Q. Duong, and H. V. Poor, “Power allocation for energy efficiency and secrecy of wireless interference networks,” *IEEE Trans. Wireless Commun.*, vol. 17, no. 6, pp. 3737–3751, Mar. 2018.

## PHỤ LỤC

### PHỤ LỤC A: MỘT SỐ BẤT ĐẲNG THỨC CƠ BẢN

Sử dụng bất đẳng thức Cauchy-Schwarz ta có

$$2xy \leq \frac{\bar{x}\bar{y}}{2} \left( \frac{x}{\bar{x}} + \frac{y}{\bar{y}} \right)^2, \forall (x, y) \in \mathbb{R}_+^2 \text{ and } (\bar{x}, \bar{y}) \in \mathbb{R}_+^2 \quad (\text{A.1})$$

Thay  $y \rightarrow 1/y$  và  $\bar{y} \rightarrow 1/\bar{y}$ , ta có bất đẳng thức

$$2\frac{x}{y} \leq \frac{\bar{x}}{2\bar{y}} \left( \frac{x}{\bar{x}} + \frac{\bar{y}}{y} \right)^2, \forall (x, y) \in \mathbb{R}_+^2 \text{ và } (\bar{x}, \bar{y}) \in \mathbb{R}_+^2 \quad (\text{A.2})$$

Hàm  $f(x) = \ln(1 + 1/x)$  là hàm lồi trong vùng  $x > 0$

$$\ln\left(1 + \frac{1}{x}\right) \geq \ln\left(1 + \frac{1}{\bar{x}}\right) + \frac{1}{1 + \bar{x}} \left(1 - \frac{x}{\bar{x}}\right). \quad (\text{A.3})$$

Hàm lồi  $f(x, y) = \ln(1 + \frac{1}{xy})$  trên  $\mathbb{R}_+^2$

$$\begin{aligned} \ln\left(1 + \frac{1}{xy}\right) &\geq f(\bar{x}, \bar{y}) + \langle \nabla f(\bar{x}, \bar{y}), (x, y) - (\bar{x}, \bar{y}) \rangle \\ &\geq \ln(1 + 1/\bar{x}\bar{y}) + \frac{1/\bar{x}\bar{y}}{\bar{t}(1 + 1/\bar{x}\bar{y})} \left(2 - \frac{x}{\bar{x}} - \frac{y}{\bar{y}}\right). \end{aligned} \quad (\text{A.4})$$

Thay  $x \rightarrow 1/x$  và  $\bar{x} \rightarrow 1/\bar{x}$

$$\ln\left(1 + \frac{x}{y}\right) \geq \ln(1 + \bar{x}/\bar{y}) + \frac{\bar{x}/\bar{y}}{(1 + \bar{x}/\bar{y})} \left(2 - \frac{\bar{x}}{x} - \frac{y}{\bar{y}}\right). \quad (\text{A.5})$$

Từ [113], hàm  $f(x, y, t) = \frac{1}{t} \ln(1 + \frac{1}{xy})$  là hàm lồi trên  $\mathbb{R}_+^3$ . Do đó

$$\begin{aligned} \frac{1}{t} \ln\left(1 + \frac{1}{xy}\right) &\geq f(\bar{x}, \bar{y}, \bar{t}) + \langle \nabla f(\bar{x}, \bar{y}, \bar{t}), (x, y, t) - (\bar{x}, \bar{y}, \bar{t}) \rangle \\ &\geq \frac{2}{\bar{t}} \ln(1 + 1/\bar{x}\bar{y}) + \frac{1/\bar{x}\bar{y}}{\bar{t}(1 + 1/\bar{x}\bar{y})} \left(2 - \frac{x}{\bar{x}} - \frac{y}{\bar{y}}\right) - \frac{\ln(1 + 1/\bar{x}\bar{y})}{\bar{t}^2} t. \end{aligned} \quad (\text{A.6})$$



Thay  $x \rightarrow 1/x$  và  $\bar{x} \rightarrow 1/\bar{x}$

$$\frac{1}{t} \ln\left(1 + \frac{x}{y}\right) \geq \frac{2}{\bar{t}} \ln(1 + \bar{x}/\bar{y}) + \frac{\bar{x}/\bar{y}}{\bar{t}(1 + \bar{x}/\bar{y})} \left(2 - \frac{\bar{x}}{x} - \frac{y}{\bar{y}}\right) - \frac{\ln(1 + \bar{x}/\bar{y})}{\bar{t}^2} t. \quad (\text{A.7})$$

**PHỤ LỤC B: CODE CHƯƠNG 2**

```

clear
clear all
clc
close all
%Declare global variables
global M UI UE N E UR RG RGe RD UG UGe UD D2D DD DG
DGe DR Noise rmin_CU rmin_D2D Pmax_BS Pmax_D2D emin
global rho mode

rho = 0.5;%energy efficient
mode = 1; %0: without IRS; 1:with IRS
%% INITIAL VARIABLES %%
%Fixed values
R_cell = 60; %size of cell
d_D2D = 30; %30; %distance of D2D pairs
B = 10*1e6; % MHZ bandwidth 10*1e6
Noise = 10-(174/10)*B;%AGWN white noise

%Dynamic values
M = 10; % #MIMO antenna
UI = 3; %number of UIs 6
UE = 2; %number of UEs 2
N = 3; % number of D2Ds 6
E = 200; %number of reflected elements
for E=200:100:200
    %E=E+10;
%% GENERATE SCENARIOS %%
%Location of BS, IUs, EUs, D2Ds
%[X_CU, Y_CU, Z_CU, X_CUe, Y_CUe, Z_CUe, X_Dt, Y_Dt, Z_Dt,
X_Dr, Y_Dr, Z_Dr]=Generate_Location_202(R_cell, d_D2D);
filename = strcat('Location_Rev2_1105_300', '.mat');

```

```

%save(filename, 'R_cell', 'd_D2D', 'X_CU', 'Y_CU', 'Z_CU',
'X_CUe', 'Y_CUe', 'Z_CUe', 'X_Dt', 'Y_Dt', 'Z_Dt', 'X_Dr',
'Y_Dr', 'Z_Dr');
load(filename, 'R_cell', 'd_D2D', 'X_CU', 'Y_CU', 'Z_CU',
'X_CUe', 'Y_CUe', 'Z_CUe', 'X_Dt', 'Y_Dt', 'Z_Dt', 'X_Dr',
'Y_Dr', 'Z_Dr');

%Plot_system_model(R_cell, X_CU, Y_CU, Z_CU, X_CUe, Y_CUe,
Z_CUe, X_Dt, Y_Dt, Z_Dt, X_Dr, Y_Dr, Z_Dr);

%Pathloss Model
%[UR, RG, RGe, RD, UG, UGe, UD, D2D, DD, DG,
DGe,DR]=PathlossShadowing_revision3(X_CU, Y_CU, Z_CU, X_CUe,
Y_CUe, Z_CUe, X_Dt, Y_Dt, Z_Dt, X_Dr, Y_Dr, Z_Dr,
shadow_deviation);
filename = strcat('Channel_Rev2_1105_', num2str(E),
'.mat');%0905: Change channel model
%filename = strcat('Channel_Rev_1105_antenna_',
num2str(M), '.mat');
%save(filename, 'UR', 'RG', 'RGe', 'RD', 'UG', 'UGe',
'UD', 'D2D', 'DD', 'DG', 'DGe', 'DR');
load(filename, 'UR', 'RG', 'RGe', 'RD', 'UG', 'UGe', 'UD',
'D2D', 'DD', 'DG', 'DGe', 'DR');

%% MAIN FUNCTION OF ITERATION %%
%network parameters
%pmax_range = 10.^([25:1:27]/10); %20dBm→mW
pmax_range = 10^(24/10);% power of BS
rmin_CU_range = 0.001*log(2);% IU data rate threshold range
%emin = 10^(0/10); % energy threshold
emin_range = 10.^([-20]/10);% energy threshold range
Pmax_D2D = 10^(24/10); %power of D2D transmitter
%rmin_D2D = 0.4*log(2);% D2D data rate threshold

```

```

rmin_D2D_range = [0.1]*log(2);% D2D data rate threshold range

for lap=1:1
%loop parameters
Montein = 0;
NoMonte = length(rmin_D2D_range);
%NoMonte = 20;
run_alg = 1; %1 PGS 2 PGS-time 3 IGS 4 IGS-time
%Sname = 'temp.txt';

RCU_avg = zeros(1, NoMonte);
RD_avg = zeros(1, NoMonte);
while Montein < NoMonte
    Montein = Montein+1;

    Pmax_BS = pmax_range(1);
%    rmin_CU = rmin_CU_range(Montein);
    rmin_CU = rmin_CU_range(1);
    rmin_D2D = rmin_D2D_range(Montein);
    emin = emin_range(1);
    Curloop = 0;
    Noloop = 1;
    RCU_min = zeros(1, Noloop);
    RD_min = zeros(1, Noloop);
    fprintf(' Runing with Pmax %4.1f(mW), PD2D %4.2f ,
    rCUmin %2.4f bps/Hz at %d/%d realizations \n', Pmax_BS,
    Pmax_D2D, rmin_CU/log(2), Montein, NoMonte);

    while(Curloop < Noloop)
        Curloop = Curloop + 1;
        fprintf(' Runing Loop at %d/ %d total \n',
        Curloop, Noloop);
        if(run_alg == 1)

```

```

disp('Run Alg1: without_timeAllocation');
[result, iter, minCU, minE, minECU, minED2D, minD,
rco_norm, rco, beam, P_tot, pDmax, t_i, t_e,
t_ie]=without_timeAllocation();
elseif(run_alg == 2)
disp('Run Alg2: with_timeAllocation');
[result, iter, minCU, minE, minD, rco_norm, rco,
beam, P_tot, pDmax, t_i, t_e, t_d,
t_ied]=with_timeAllocation();
filename = strcat('GA_iterations_Rev2', '.mat');
num2str(Montein)
save(filename,
'result', 'rco', 'beam', 'rco_norm', 'rmin_CU',
'minCU', 'minE', 'emin', 'minD', 'rmin_D2D',
'P_tot', 'pDmax', 't_i', 't_e', 't_d', 't_ied');
end
end
RCU_avg(Montein) = sum(RCU_min)/Noloop;
end
end
end

```

**PHỤ LỤC C: CODE CHƯƠNG 3**

```

clear all
clc
close all

global N T q_sp q_gp Obs R_Obs delta v_max v_min D_cell ...
    d_r d_safe tau_max obj_max...
    max rmin tau_min dsg we M b_col A_col K

%load('parameter.mat');
D_cell = 10;%100m2
T = 51;
N = 6; % number of obstacles in rectangle shape
K = 3; % number of obstacles in circle shape
obj_max = 20;
we = 1;
M = 100;
d_r = 0.5; %m, size of the robot
delta = 0.01; %m, safe distance
v_max = 1; %m/s, maximum velocity of robot
v_min = 0.3; %m/s, maximum velocity of robot
tau_max = 1;
tau_min = 0.2;
%Generate scenarios
%Location of mobile diveces
[q_sp, q_gp, Obs]=Generate_Location(D_cell);
filename = strcat('Location_0211_', '.mat');
save(filename, 'D_cell', 'q_sp', 'q_gp', 'Obs');
%load(filename, 'D_cell', 'q_sp', 'q_gp', 'Obs');
%plot obstacles
%for i=1:N
%    rectangle('Position',Obs(i).q,'FaceColor','k'); hold on;

```

```

%end
%axis([0 10 0 10]); grid on;

dsg = Distance(q_sp,q_gp);
for i=1:10
    Obs(i).safe(1) = Obs(i).q(1) - d_r - delta;
    Obs(i).safe(2) = Obs(i).q(2) - d_r - delta;
    Obs(i).safe(3) = Obs(i).q(3) +2*(d_r+delta);
    Obs(i).safe(4) = Obs(i).q(4) +2*(d_r+delta);
end

    for n=1:N
        b_col(n,1) = Obs(n).safe(1);
        b_col(n,2) = Obs(n).safe(2);
        b_col(n,3) = Obs(n).safe(1)+Obs(n).safe(3);
        b_col(n,4) = Obs(n).safe(2)+Obs(n).safe(4);
    end

Montein = 0;
NoMonte = size(1,1);
while Montein < NoMonte
    Montein = Montein+1;
    Curloop = 0;
    Noloop = 1;
    %RCU_min = zeros(1, Noloop);
    %RD_min = zeros(1, Noloop);
    fprintf('Runing with workspace %d(m2), size of robot
    %4.1f(m), number of obstacles %d\n', D_cell, d_r, N);

    while (Curloop < Noloop)
        Curloop = Curloop+1;
        disp('Run Algorithm ');
        [result , sum_dis, sum_time, iter , robot , cole , v_t,
        omega_t,theta_t,min_q] = Algorithm ();
    end
end

```

```
end
```

```
filename = strcat('result ', '.mat'); %41 , num2str(Montein)  
save(filename, 'result');
```



**PHỤ LỤC D: CODE CHƯƠNG 4**

```

import gym
from gym.utils import seeding
import numpy as np
from gym import spaces
from scipy.linalg import sqrtm
from stable_baselines3 import PPO
from stable_baselines3.common.vec_env import DummyVecEnv
import math

def path_loss_shadowing(X_R, Y_R, Z_R, check_AR, check_IR):
    # global M, E, BS, IRS, KR_AI, mu_AI, sigma_AI, KR_AR,
    mu_AR, sigma_AR, KR_IR, mu_IR, sigma_IR, f_c

    M = 1 # number of antennas of the AP
    E = 16 # number of elements of the IRS
    BS = np.array([0, 2.5, 3])
    IRS = np.array([5, 5, 3])
    KR_AI = 3 # Rician factor 3dB
    mu_AI = np.sqrt(KR_AI / ((KR_AI + 1)))
    sigma_AI = np.sqrt(1 / ((KR_AI + 1)))
    if check_AR:
        KR_AR = 0
    else:
        KR_AR = 3 # Rician factor depend on the location of
        #the robot, it can be 3 or 0
    mu_AR = np.sqrt(KR_AR / ((KR_AR + 1)))
    sigma_AR = np.sqrt(1 / ((KR_AR + 1)))
    if check_IR:
        KR_IR = 0
    else:
        KR_IR = 3 # Rician factor depend on the location of

```

```

#the robot , it can be 3 or 0
mu_IR = np.sqrt(KR_IR / ((KR_IR + 1)))
sigma_IR = np.sqrt(1 / ((KR_IR + 1)))
f_c = 2 * 10**3 # 2000MHz
# Path loss from AP to IRS
AI = {"chan": None, "l": None, "n": None}
Dist = np.sqrt((BS[0] - IRS[0])**2 + (BS[1] - IRS[1])**2
+ (BS[2] - IRS[2])**2)/1000 #Km
path_loss_shasmall = 31.84 + 21.50 * np.log10(Dist)
+ 19 * np.log10(f_c)
channel = np.sqrt((10**((path_loss_shasmall / 10))))
* (np.random.normal(mu_AI, sigma_AI, (1, M))
+ 1j * np.random.normal(mu_AI, sigma_AI, (1, M)))
HR = np.zeros((E, M), dtype=complex)
for e in range(E):
    theta = np.random.uniform(0, np.pi)
    phi = np.random.uniform(0, 2 * np.pi)
    for m in range(M):
        HR[e, m] = np.exp(1j * np.pi * ((e - 1)
* np.sin(np.pi - theta) * np.sin(np.pi + phi)
+ (m - 1) * np.sin(theta) * np.sin(phi)))
AI["chan"] = channel * HR
# Path loss from IRS to the robot
IR = {"chan": None, "l": None, "n": None}
Dist = np.sqrt((BS[0] - IRS[0])**2 + (BS[1] - IRS[1])**2
+ (BS[2] - IRS[2])**2)/1000 #km

if KR_IR == 0: # NLoS case
    path_loss_shasmall = max(path_loss_shasmall, 32.4
+ 23 * np.log10(Dist) + 20 * np.log10(f_c))
else:
    path_loss_shasmall = 31.84 + 21.50 * np.log10(Dist)
+ 19 * np.log10(f_c)

```

```

channel = np.sqrt((10**(path_loss_shasmall / 10)))
*(np.random.normal(mu_IR, sigma_IR, (1, E))
+ 1j * np.random.normal(mu_IR, sigma_IR, (1, E)))

azim = np.arctan((IRS[0] - X_R) / (IRS[1] - Y_R))
elev = np.arcsin((IRS[2] - Z_R) / Dist)
RI = np.zeros((E, E), dtype=complex)
for e in range(E):
    for f in range(E):
        RI[e, f] = np.exp(1j * np.pi * (e - f)
        * np.sin(azim) * np.sin(elev))

# print(channel, RI)
IR["chan"] = channel * sqrtm(RI)
# Path loss from AP to the robot
AR = {"chan": None, "n": None}
Dist = np.sqrt((BS[0] - X_R)**2 + (BS[1] - Y_R)**2
+ (BS[2] - Z_R)**2)/1000 # km
if KR_AR == 0: # NLoS case
    path_loss_shasmall = max(path_loss_shasmall, 32.4
    + 23 * np.log10(Dist) + 20 * np.log10(f_c))
else:
    path_loss_shasmall = 31.84 + 21.50 * np.log10(Dist)
    + 19 * np.log10(f_c)
channel = np.sqrt((10**(path_loss_shasmall / 10)))
* (np.random.normal(mu_AR, sigma_AR, (1, M)) +
1j * np.random.normal(mu_AR, sigma_AR, (1, M)))
AR["chan"] = channel
return AI['chan'], IR['chan'], AR['chan']

```

```

def check_collision(point1, point2, obstacle_centers,
obstacle_sizes):
    obstacle_size = np.array(obstacle_sizes)
    obstacle_center = np.array(obstacle_centers)

```

```

# Calculate coordinate of peak of boxes
obstacle_vertices = np.array([
    [obstacle_center[0] - obstacle_size[0] / 2,
     obstacle_center[1] - obstacle_size[1] / 2],
    [obstacle_center[0] + obstacle_size[0] / 2,
     obstacle_center[1] - obstacle_size[1] / 2],
    [obstacle_center[0] + obstacle_size[0] / 2,
     obstacle_center[1] + obstacle_size[1] / 2],
    [obstacle_center[0] - obstacle_size[0] / 2,
     obstacle_center[1] + obstacle_size[1] / 2]])

# Check if it cut obstacles
for i in range(4):
    j = (i + 1) % 4
    if np.any(
        ((point1[0] - obstacle_vertices[i, 0])
         * (point2[1] - obstacle_vertices[i, 1]) -
         (point1[1] - obstacle_vertices[i, 1])
         * (point2[0] - obstacle_vertices[i, 0]))
         * ((point1[0] - obstacle_vertices[j, 0])
            * (point2[1] - obstacle_vertices[j, 1]) -
            (point1[1] - obstacle_vertices[j, 1])
            * (point2[0] - obstacle_vertices[j, 0]))
         < 0):
        return True # It cut
    return False # It did not cut
def distance(point1, point2):
    return math.sqrt((point1[0] - point2[0])**2 + (point1[1]
    - point2[1])**2)
def collision_detecte(point, obstacle_centers, obstacle_sizes,
min_distance=0.25):
    for center, size in zip(obstacle_centers, obstacle_sizes):
        half_width, half_height = size[0]/2, size[1]/2
        x_min, x_max = center[0] - half_width, center[0]

```

```

+ half_width
y_min, y_max = center[1] - half_height, center[1]
+ half_height
# Check the point is in box or not
if x_min <= point[0] <= x_max and y_min <= point[1]
<= y_max:
    return False # The point is in the box,
    it does not safe

# Calculate the distance from the point to the box
dist = distance(point, center)

# Check the distance is less than the safety distance
or not
if dist < min_distance:
    return False

# The distance is too close, it does not safe
return True #it safe
room_size = (10, 10)
obstacle_sizes = [(3, 1), (2, 0.75), (2, 0.5), (1.5, 2),
(2, 0.5)]
obstacle_centers = [(4.75, 3), (3.5, 5.5), (4.25, 6.5),
(8, 5), (5, 8.75)]
AP = np.array([0, 8])
goal = np.array([8, 8])
IRS = np.array([6, 5])
t1=0.1
t2=0.5
m = 9
v = 0.5
I = 0.16245
P_s = 17.8
L = 30

```

```

g = 9.8
nu = 0.05
noise = 1e-9
rate_min = 170

class RobotEnv(gym.Env):
    # metadata = {'render.modes': ['human']}

    def __init__(self):
        super(RobotEnv, self).__init__()
        self.step_dir = np.array([[ -0.25, 0], [0.25, 0],
        [0, 0.25],
        [0, -0.25]])
        self.num_step = 0
        self.state = np.full(3, None)
        self.state_low = np.array([0, 0, -100])
        self.state_high = np.array([10, 10, 100])
        # build action space (b=3)
        self.action_space = spaces.MultiDiscrete([50, 8, 8, 8,
        8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8, 8])
        self.observation_space = spaces.Box(low=self.state_low,
        high=self.state_high, dtype=np.float32) #build
        state space

    def reset(self):
        self.state[0: 2] = 1.5
        self.state[2] = 0
        self.state = np.array(self.state, dtype=np.float32)
        return self.state

    def seed(self, seed=None):
        self.np_random, seed = seeding.np_random(seed)
        return [seed]

```

```

def step(self, action):
    self.num_step +=1
    done = False
    self.robot = self.state[0:1]

    self.next_state = self.robot + self.step_dir[action[0], :]
    collision_detected = collision_detecte(self.next_state,
    obstacle_centers, obstacle_sizes)
    distance_to_walls = [
        self.next_state[0], room_size[0] - self.next_state[0],
        self.next_state[1], room_size[1] - self.next_state[1]]
    euclidean_dist_to_goal = np.linalg.norm(self.robot - goal)
    if collision_detected == False or min(distance_to_walls)
    <= 0.5:
        self.robot = self.robot +
        self.step_dir[action[0], :]*-1
    else:
        self.robot = self.robot +
        self.step_dir[action[0], :]

    #chanel
    check_IR = check_collision(IRS, self.robot,
    obstacle_centers, obstacle_sizes)
    check_AR = check_collision(AP, self.robot,
    obstacle_centers, obstacle_sizes)
    X_R = self.robot[0]
    Y_R = self.robot[1]
    AI, IR, AR = path_loss_shadowing(X_R, Y_R, 1.5,
    check_AR, check_IR)
    phi1 = np.pi*action[2]/8
    phi2 = np.pi*action[2]/8
    diagonal_matrix = np.zeros((16, 16), dtype=np.complex64)

```

```

diagonal_matrix[:8, :8] = np.exp(1j * phi1)
* np.identity(8)
diagonal_matrix[8:, 8:] = np.exp(1j * phi2)
* np.identity(8)
A = np.linalg.norm(AR +
np.dot(np.dot(np.conjugate(IR), diagonal_matrix), AI))**2
rate = math.log2(1+(A*0.01)/noise) #shuold be 4-5bit/s
# print("rate:", rate, "A:", A)
if rate < rate_min:
    r_rate = (rate_min - rate)/10
else:
    r_rate = 0
direction = math.atan2(self.step_dir[action[0], :][1],
self.step_dir[action[0], :][0]) #direction of robot
omega = abs(direction-self.state[2])/t1
E_R=(m*v**2)/2 + (I*omega**2)/2 + P_s*L*(t1+t2)
+ 2*nu*m*g*v*t2
E_A = 0.01*(t1+t2)
print("EA:", E_A*10, "ER", E_R/100, "khoang cach",
euclidean_dist_to_goal, "r_rate", r_rate, "huong",
direction, "huong cu", self.state[2])
if euclidean_dist_to_goal <= 0.5:
    reward = 10000 -E_A*10 - E_R/100 -
    euclidean_dist_to_goal
    self.num_step = 0
    done = True
else:
    reward = -E_A*10 - E_R/100 - euclidean_dist
    _to_goal - r_rate
if self.num_step == 800:
    done = True
    self.num_step = 0
self.state[0:2] = self.robot

```



```

        self.state[2] = direction
        self.state = np.array(self.state, dtype=np.float32)
        info = {}
        return self.state, reward, done, info
    def render(self, mode='human'):
        pass
env = ENV()

# PPO model
model = PPO("MlpPolicy", env, verbose=1)
model.learn(total_timesteps=500000)
model.save("PPO")
del model
model = PPO.load("PPO")

# Train model
num_episodes = 1000
costs_PPO = []
avg_cost_PPO = []
score_history = []

for t in range(num_episodes):
    done = False
    num_step = 0
    obs = env.reset()
    score = 0
    while not done:
        action, _ = model.predict(obs)
        obs, reward, done, info = env.step(action)
        print(obs)
        score += reward
        # print(obs)
        # print(f"Step: {t+1}, Reward: {reward}")

```