

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ



DƯƠNG NGỌC SƠN

**CÁC THUẬT TOÁN ĐỊNH VỊ KHÔNG DÂY
CHO CÁC DỊCH VỤ DỰA TRÊN VỊ TRÍ**

LUẬN ÁN TIẾN SĨ
NGÀNH KỸ THUẬT VIỄN THÔNG

HÀ NỘI - 2025

ĐẠI HỌC QUỐC GIA HÀ NỘI
TRƯỜNG ĐẠI HỌC CÔNG NGHỆ



DƯƠNG NGỌC SƠN

**CÁC THUẬT TOÁN ĐỊNH VỊ KHÔNG DÂY
CHO CÁC DỊCH VỤ DỰA TRÊN VỊ TRÍ**

LUẬN ÁN TIẾN SĨ
NGÀNH KỸ THUẬT VIỄN THÔNG

CHUYÊN NGÀNH : KỸ THUẬT VIỄN THÔNG
MÃ SỐ : 9520208
CÁN BỘ HƯỚNG DẪN: PGS. TS. DINH THỊ THÁI MAI
 PGS. TS. LÂM SINH CÔNG

HÀ NỘI - 2025

Lời cam đoan

“Tôi xin cam đoan các kết quả trình bày trong luận án này là công trình nghiên cứu của tôi dưới sự hướng dẫn của các cán bộ hướng dẫn. Các số liệu, kết quả trình bày trong luận án là hoàn toàn trung thực và chưa được công bố trong bất kỳ công trình nào trước đây. Các kết quả sử dụng tham khảo đều được trích dẫn đầy đủ theo quy định.”

Hà Nội, ..., ...,

Tác giả

Lời cảm ơn

Luận án này không thể hoàn thành nếu không có sự giúp đỡ của thầy cô, gia đình, bạn bè và các tổ chức.

Tôi xin bày tỏ lòng biết ơn sâu sắc đến PGS. TS. Đinh Thị Thái Mai, PGS. TS. Lâm Sinh Công và cố PGS. TS. Nguyễn Quốc Tuấn. Sự chỉ bảo tận tâm của cô và thầy đã mang lại kiến thức và kỹ năng hết sức quý báu để tôi có thể hoàn thiện luận án một cách tốt nhất.

Tôi xin chân thành gửi lời cảm ơn đến các thầy, cô trong Bộ môn Hệ thống Viễn thông - Khoa Điện tử Viễn thông, Trường Đại học Công Nghệ đã hỗ trợ và giúp đỡ tôi trong suốt quá trình làm Luận án Tiến Sĩ.

Tôi xin biết ơn gia đình và chân thành cảm ơn những người đồng hành, đặc biệt là các em Vũ Thị Lan Nhi, Hoàng Thị Phương Anh, Nguyễn Thu Trang, Ngô Quang Tăng, Tạ Thị Huyền Trang đã hỗ trợ tôi rất nhiều về cả vật chất lẫn tinh thần để tôi có thể học tập đạt kết quả tốt và thực hiện thành công luận án này.

Luận án này được tài trợ bởi tập đoàn Vingroup và hỗ trợ bởi Chương trình học bổng Thạc Sĩ, Tiến Sĩ trong nước của Quỹ Đổi mới và sáng tạo Vingroup (VINIF), Viện Nghiên cứu Dữ liệu lớn, mã số VINIF.2021.TS.094 và VINIF.2022.TS.103.

Mục lục

Danh mục các từ viết tắt	ix
Danh mục hình vẽ	xii
Danh mục bảng	xiii
1 Giới thiệu nội dung nghiên cứu	1
1.1 Động lực nghiên cứu	1
1.1.1 Định vị trong nhà	1
1.1.2 Định vị ngoài trời	3
1.2 Các nghiên cứu liên quan	4
1.2.1 Định vị trong nhà với phép đo RSS	4
1.2.2 Định vị ngoài trời với sóng mi-li-mét (mmWave)	8
1.3 Thách thức và hạn chế	10
1.4 Mục tiêu và đóng góp của luận án	11
1.4.1 Mục tiêu của luận án	11
1.4.2 Đóng góp của luận án	11
1.5 Cấu trúc luận án	13
2 Nền tảng cơ bản	16
2.1 Các công nghệ định vị không dây	16
2.1.1 Wi-Fi [119]	17
2.1.2 Bluetooth năng lượng thấp (BLE) [119]	17
2.1.3 Ultra Wideband (UWB) [119]	18

2.1.4	Sóng mmWave trong hệ thống mạng 5G [57]	18
2.2	Các phép đo của hệ thống định vị không dây	19
2.2.1	Thời gian đến (TOA)	20
2.2.2	Chênh lệch thời gian đến (TDOA)	21
2.2.3	Cường độ tín hiệu (RSS)	22
2.2.4	Góc đến (DOA)	23
2.3	Định vị dựa trên tham số	24
2.3.1	Thuật toán định vị phi tuyến	25
2.3.2	Thuật toán định vị tuyến tính	28
2.4	Định vị không dựa trên tham số	31
2.4.1	K lân cận (KNN)	32
2.4.2	Mạng nơ-ron nhân tạo (ANN)	33
2.4.3	Mạng nơ-ron tích chập (CNN)	34
2.4.4	Mạng nơ-ron hồi quy (RNN)	34
2.4.5	Bộ nhớ Ngắn hạn Dài (LSTM)	35
2.5	Kết luận chương	36
3	Định vị với phép đo RSS trong hệ thống định vị trong nhà sử dụng Bluetooth	37
3.1	Phép đo RSS	38
3.1.1	Cường độ tín hiệu (RSS)	38
3.1.2	Chuyển đổi từ RSS sang khoảng cách.	44
3.2	Định vị trong nhà với cách triển khai đèn hiệu BLE tam giác đều	46
3.2.1	Hệ thống đề xuất	46
3.2.2	Chiến lược đặt đèn hiệu đề xuất	46
3.2.3	Kết quả và thảo luận	48
3.3	Xem xét sự hỗ trợ của học máy trong vấn đề định vị trong nhà	51
3.3.1	Hệ thống đề xuất	53
3.3.2	Kết quả và thảo luận	58
3.4	Kết luận chương	64

4	Định vị với các phép đo TOA, AOA, AOD trong hệ thống mmWave MIMO	66
4.1	Giới thiệu	67
4.2	Định vị dựa trên các tham số kênh MIMO mmWave	71
4.2.1	Mô hình kênh MIMO mmWave	71
4.2.2	Biểu diễn thưa của kênh MIMO mmWave trong miền góc	72
4.2.3	Chuyển đổi từ các tham số kênh truyền sang vị trí mục tiêu . .	73
4.3	Các thuật toán ước lượng thưa họ OMP truyền thống	76
4.3.1	Thuật toán OMP	76
4.3.2	Thuật toán StOMP	76
4.3.3	OMP đồng thời (SOMP)	77
4.4	Phương pháp ước lượng thô không xét đến tương quan giữa các đường dẫn: Cách tiếp cận dựa trên thuật toán StOMP phân tán	78
4.4.1	Phương pháp đề xuất	78
4.4.2	Kết quả và thảo luận	83
4.5	Phương pháp ước lượng thô có xét đến tương quan giữa các đường dẫn	87
4.5.1	Phương pháp đề xuất	87
4.5.2	Kết quả và thảo luận	93
4.6	Phương pháp ước lượng tinh	96
4.6.1	Phương pháp đề xuất	96
4.6.2	Kết quả và thảo luận	103
4.7	Kết luận chương	107
5	Kết luận và công việc trong tương lai	108
5.1	Kết luận	108
5.2	Công việc trong tương lai	110
	Phụ lục	112
	Danh mục công trình khoa học của tác giả liên quan tới luận án	117

Danh mục các từ viết tắt và ký hiệu toán học

Các từ viết tắt

STT	Viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
1	(A)NN	Artificial Neural Network	Mạng nơ-ron (nhân tạo)
2	3GPP	The 3rd Generation Partnership Project	Dự án hợp tác thế hệ thứ ba
3	4G	The Fourth Generation	Mạng di động thế hệ thứ tư
4	5G	the Fifth Generation	Mạng di động thế hệ thứ năm
5	AOA	Angle of Arrival	Góc đến của tín hiệu
6	AOD	Angle of Departure	Góc đi của tín hiệu
7	AP	Access Point	Điểm truy cập
8	BLE	Bluetooth Low Energy	Bluetooth năng lượng thấp
9	BS	Base Station	Trạm gốc
10	CNN	Convolution Neural Network	Mạng nơ-ron tích chập
11	CRLB	Cramér–Rao Lower Bound	Cận dưới Cramér–Rao
12	CSI	Channel State Information	Thông tin trạng thái kênh
13	CWLS	Conventional Weighted Least Square	Bình phương tối thiểu có trọng số thông thường
14	DNN	Deep Neural Network	Mạng học sâu
15	DOA	Direction of Arrival	Hướng đến của tín hiệu

STT	Viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
16	FIM	Fisher Information Matrix	Ma trận thông tin Fisher
17	FP	Fingerprinting	Lấy dấu vân tay
18	GD	Gradient Descent	Suy giảm đạo hàm
19	GNSS	Global Navigation Satellite System	Hệ thống vệ tinh định vị toàn cầu
20	GPS	Global Positioning System	Hệ thống định vị toàn cầu
21	GS	Golden-Section Search	Tìm kiếm sử dụng tỉ lệ vàng
22	IMU	Inertial Measurement Unit	Đơn vị đo quán tính
23	INS	Inertial Navigation system	Hệ thống dẫn đường quán tính
24	IPS	Indoor Positioning System	Hệ thống định vị trong nhà
25	KNN	k Nearest Neighbor	Thuật toán k lân cận
26	LBS	Location-Based Services	Dịch vụ dựa trên vị trí
27	LLS	Linear Least Square	Bình phương tối thiểu tuyến tính
28	LOS	Line of Sight	Đường nhìn thẳng
29	LS	Least Square	Bình phương tối thiểu
30	LSTM	Long Short Term Memory	Bộ nhớ Ngắn hạn dài
31	MIMO	Multiple Input Multiple Output	Nhiều đầu vào nhiều đầu ra
32	MISO	Multiple Input Single Output	Nhiều đầu vào một đầu ra
33	MLE	Maximum Likelihood Estimation/Estimator	Tối đa hàm khả năng
34	MSE	Mean Square Error	Sai số toàn phương trung bình
35	NLOS	Non-Line of Sight	Đường truyền bị che khuất
36	NLS	Non-linear Least Square	Bình phương tối thiểu phi tuyến tính
37	OMP	Orthogonal Matching Pursuit	Thuật toán theo đuổi đối sánh trực giao

STT	Viết tắt	Từ tiếng Anh	Nghĩa tiếng Việt
38	PDR	Pedestrian Dead Reckoning	Kỹ thuật định vị sử dụng IMU
39	PEB	Position Error Bound	Giới hạn lỗi vị trí
40	PLE	Path Loss Exponent	Hệ số suy hao
41	PSR	Probability of Successful Recovery	Xác suất khôi phục thành công (tập hỗ trợ)
42	REB	Rotation Error Bound	Giới hạn sai số góc quay
43	RFID	Radio Frequency Identification	Công nghệ nhận dạng bằng sóng vô tuyến
44	RL	Reinforcement Learning	Học củng cố
45	RSS(I)	Received Signal Strength (Indicator)	Cường độ tín hiệu
46	RTT	Round Trip Time	Thời gian truyền khứ hồi
47	SNR	Signal to Noise Ratio	Tỉ số tín hiệu trên tạp âm
48	SOMP	Simultaneous OMP	Thuật toán OMP đồng thời
49	TDOA	Time Different of Arrival	Khác biệt về thời gian đến của tín hiệu
50	TOA	Time of Arrival	Thời gian đến của tín hiệu
51	UE	User Equipment	Thiết bị người dùng
52	UWB	Ultra Wide Band	Băng siêu rộng
53	V2V	Vehicle to Vehicle	Kết nối phương tiện với phương tiện
54	WKNN	Weighted KNN	Phương pháp KNN có trọng số
55	WLS	Weighted Least Square	Phương pháp LS có trọng số

Ký hiệu toán học

STT	Ký hiệu	Ý nghĩa
1	In nghiêng, chữ thường, hoặc in hoa ví dụ: a , N	Đại lượng vô hướng

STT	Ký hiệu	Ý nghĩa
2	In đậm, chữ thường, ví dụ: a	Véc-tơ, các thành phần của véc-tơ a được ký hiệu là a_i
3	In đậm, chữ hoa, ví dụ: A	Ma trận, các thành phần của ma trận A được ký hiệu là a_{ij} hoặc $[\mathbf{A}]_{ij}$
4	$\mathbb{R}$	Tập số thực
5	$\mathbb{C}$	Tập số phức
6	$\mathbb{E}(\cdot)$	Kỳ vọng
7	$\Re(\cdot)$	Phần thực của số phức
8	$\Im(\cdot)$	Phần ảo của số phức
9	$\angle(\cdot)$	Góc lượng giác hoặc pha của số phức
10	$ \cdot $	Trị tuyệt đối hoặc mô-đun của số phức
11	$\mathbf{0}_K$	Ma trận toàn 0 có kích thước $K \times K$
12	$\mathbf{1}_K$	Ma trận toàn 1 có kích thước $K \times K$
13	$\mathbf{I}_K$	Ma trận đơn vị có kích thước $K \times K$
14	$\mathbb{R}^N$ hoặc $\mathbb{C}^N$	Véc-tơ thực hoặc phức có độ dài bằng N
15	$\mathbb{R}^{M \times N}$ hoặc $\mathbb{C}^{M \times N}$	Ma trận thực hoặc phức có kích thước $M \times N$
16	$(\cdot)^*$	Liên hợp phức (vô hướng)
17	$(\cdot)^\top$	Phép chuyển vị thường (véc-tơ, ma trận)
18	$(\cdot)^H$	Chuyển vị phức (véc-tơ, ma trận)
19	$\otimes$	Tích Kronocker
20	$\odot$	Tích từng phần tử
21	$\langle \cdot, \cdot \rangle$	Tích trong
22	$\text{diag}(\mathbf{a})$	Ma trận đường chéo được tạo bởi các phần tử của véc-tơ a
23	$\ \cdot\ _q$	Chuẩn l_q (l_q -norm), $q = \{0, 1, 2, \text{F}, \dots\}$
24	$[N]$	Tập hợp $\{1, 2, \dots, N\}$

Danh sách hình vẽ

1.1	Các kịch bản nghiên cứu để tạo ra định vị liên mạch.	14
2.1	Từ trái qua phải: Kỹ thuật sử dụng khoảng cách, kỹ thuật sử dụng góc, kỹ thuật sử dụng kết hợp thời gian và góc.	19
2.2	Phương pháp lấy dấu vân tay [60]	31
2.3	Cấu trúc đơn giản của mạng nơ-ron nhân tạo	33
2.4	Mạng tích chập đơn giản (Nguồn: CS-230)	34
2.5	Cấu trúc cơ bản của một mạng Nơ-ron hồi quy (Nguồn: CS-230)	35
2.6	Bên trong một nơ-ron của mạng LSTM	36
3.1	Kết quả thực nghiệm nêu trong [11]. Các đường màu xanh và đỏ thể hiện cho $\Gamma(d)$ với hệ số mất mát lần lượt là 2 (LOS) và 3.5 (NLOS). . .	40
3.2	Cường độ tín hiệu suy giảm theo các hướng khác nhau [60]. Góc 0° và 180° lần lượt ứng với 2 trường hợp máy thu và máy phát không bị che chắn và bị che chắn bởi cơ thể con người ở khoảng cách 2 m.	41
3.3	Phân bố của các giá trị RSS ở cách khoảng cách khác nhau.	42
3.4	Xấp xỉ tuyến tính cho mô hình chuyển đổi từ RSS sang khoảng cách. .	44
3.5	Kiến trúc hệ thống	46
3.6	Chiến lược triển khai đèn hiệu BLE tam giác đều	47
3.7	Vị trí đặt đèn hiệu BLE và đường tham chiếu trong khu vực thí nghiệm.	49
3.8	Phân phối của vị trí ước lượng bởi các kỹ thuật khác nhau trên bản đồ.	50
3.9	Sai số của các phương pháp thông qua biểu đồ hộp.	51
3.10	Phân phối của ước lượng cuối cùng khi thay đổi số lượng đèn hiệu BLE.	52
3.11	Sai số định vị với số lượng đèn hiệu BLE khác nhau.	53

3.12	Kiến trúc hệ thống đề xuất	55
3.13	Cấu tạo mạng nơ-ron đa lớp đề xuất cho bài toán phân loại khu vực. .	56
3.14	Bản đồ khu vực thí nghiệm.	59
3.15	Ma trận nhầm lẫn khi ước lượng tho vị trí máy thu.	60
3.16	Độ chính xác của các phương pháp định vị	61
3.17	Ước lượng tinh không sử dụng mạng nơ-ron	62
3.18	Phân bố của vị trí ước lượng từ sự kết hợp giữa NN và CWLS (NN + CWLS) và sự kết hợp giữa NN và phương pháp WLS đề xuất (NN + PP-WLS).	63
3.19	Lỗi tích của các phương pháp khi sử dụng mạng nơ-ron như một chiến lược chọn lựa đèn hiệu.	64
4.1	Mô hình hình học bao gồm ba đường dẫn, một đường LOS (màu xanh) và hai đường NLOS (màu đỏ).	71
4.2	Định vị với đường dẫn LOS	74
4.3	Định vị với 2 đường dẫn NLOS	75
4.4	Biểu diễn trực quan của a) Kênh trong miền góc không nhiều, b) Kênh được tái tạo lại chỉ sử dụng thuật toán DStOMP, và c) Kênh được tái tạo lại sử dụng kết hợp thuật toán DStOMP và phương pháp Silhouette sửa đổi.	84
4.5	Hiệu năng PSR và MSE của các phương pháp khi thay đổi số sóng mang con N	85
4.6	Hiệu năng PSR và MSE của các phương pháp khi thay đổi số lượng các chùm tia.	87
4.7	Mối quan hệ giữa $\mathcal{X}$ và $\mathcal{Y}$ và các dữ liệu trong chúng	88
4.8	Độ tương quan của các cột trong $\mathbf{\Omega}[1]$ với tín hiệu nhận được. Biên độ của hai đỉnh $x = 62$ và $x = 105$ không phải là giá trị thực vì chúng có tương tác với nhau. Đỉnh tại $x = 62$ bị ảnh hưởng bởi đuôi của đỉnh $x = 105$ và ngược lại.	89
4.9	Mô tả phương pháp đề xuất. Hai tập con $\mathcal{S}_{\theta_1}$ và $\mathcal{S}_{\theta_2}$ biểu diễn hai hàng của kênh trong miền góc.	91

4.10	Biểu diễn PSR là một hàm theo SNR.	93
4.11	PSR của các phương pháp theo SNR với $G_r = 2N_r$	95
4.12	PSR của các phương pháp theo SNR với $G_r = 4N_r$	95
4.13	MSE của các kỹ thuật ước lượng kênh theo SNR khi thay đổi số lượng góc ỏ.	97
4.14	Cách hoạt động của phương pháp tinh chỉnh đề xuất.	100
4.15	Quá trình hội tụ của các tham số ở mức SNR cao	104
4.16	RMSE cho sai số vị trí theo thang dB của thuật toán đề xuất so với thuật toán truyền thống.	105
4.17	RMSE cho sai số về góc quay ăng-ten theo thang dB của thuật toán đề xuất so với thuật toán truyền thống.	106
4.18	a) Thời gian chạy là một hàm của N_b , b) Sai số bình phương trung bình (MSE) và thời gian chạy của thuật toán như một hàm của sai số dừng trong thuật toán tìm kiếm theo tỉ lệ vàng.	106
5.1	Biểu diễn hình học của vấn đề tối ưu vị trí đặt đèn hiệu BLE.	113

Danh sách bảng

2.1	So sánh ưu và nhược điểm của các phép đo phổ biến trong bài toán định vị	20
3.1	RSS trung bình và độ không chắc chắn của nó ở những khoảng cách khác nhau.	43
3.2	Mô hình tính toán khoảng cách cho mỗi khoảng RSSI.	45
3.3	Thống kê chi tiết lỗi của 8 khu vực con khi chỉ sử dụng mạng nơ-ron	59
3.4	So sánh các phương pháp định vị khác nhau về sai số bình phương trung bình (RMSE), trung vị, trung bình, phân vị 70 và 90.	63

Chương 1

Giới thiệu nội dung nghiên cứu

1.1 Động lực nghiên cứu

Với sự phát triển nhanh chóng của Internet, nhiều dịch vụ dựa trên vị trí (LBS) đã xuất hiện trong các lĩnh vực thương mại, giải trí, bảo mật với mục đích tạo ra trải nghiệm mới, thúc đẩy kinh tế,... Rất nhiều trong số dịch vụ trên yêu cầu thiết bị người dùng báo cáo vị trí theo thời gian thực với độ chính xác cao, hoạt động trong các loại môi trường phổ biến và có khả năng chuyển đổi liên mạch giữa môi trường trong nhà và ngoài trời. Một số hệ thống định vị phổ biến, ví dụ như hệ thống định vị toàn cầu (GPS), hoạt động tốt ở môi trường ngoài trời ít vật cản nhưng phải đối mặt với nhiều khó khăn với môi trường trong nhà và môi trường ngoài trời có nhiều vật cản do sự suy giảm tín hiệu và hiệu ứng đa đường. Vì vậy, việc tìm kiếm các giải pháp định vị cho hai loại môi trường này là hết sức cần thiết [117].

1.1.1 Định vị trong nhà

Hạn chế của GPS đã thúc đẩy nghiên cứu về hệ thống định vị trong nhà (IPS). Một số loại hình dịch vụ dựa trên vị trí cho môi trường trong nhà có thể kể đến như:

- **Tăng cường dịch vụ chăm sóc sức khỏe:** Một trong những động lực chính để nghiên cứu định vị trong nhà là nâng cao dịch vụ chăm sóc sức khỏe. Bệnh viện là môi trường phức tạp, nơi vị trí chính xác của nhân viên y tế, bệnh nhân

và thiết bị có thể ảnh hưởng đáng kể đến chất lượng dịch vụ y tế. Ví dụ, trong những tình huống khẩn cấp, việc nhanh chóng xác định vị trí của nhân viên y tế và thiết bị quan trọng có thể cứu được mạng sống của bệnh nhân. Ngoài ra, định vị trong nhà có thể hỗ trợ theo dõi chuyển động của bệnh nhân, đảm bảo họ được chăm sóc kịp thời và giảm khả năng xảy ra sai sót y tế. Bằng cách tích hợp IPS với hồ sơ sức khỏe điện tử, các nhà cung cấp dịch vụ chăm sóc sức khỏe có thể cải thiện hiệu quả quá trình điều trị của bệnh nhân.

- **Tăng cường trải nghiệm bán lẻ:** Ngành bán lẻ được hưởng lợi rất nhiều từ hệ thống định vị trong nhà. Bằng cách quan sát hành vi của khách hàng và mô hình di chuyển của họ, nhà bán lẻ có thể tối ưu hóa cách bố trí cửa hàng và tạo ra trải nghiệm mua sắm mới lạ. Định vị trong nhà cho phép quảng cáo chọn chính xác mục tiêu, gửi khuyến mãi hoặc thông tin sản phẩm trực tiếp đến điện thoại thông minh của khách hàng dựa trên vị trí của họ. Mức độ cá nhân hóa có thể nâng cao sự hài lòng và sự trung thành của khách hàng, cuối cùng là thúc đẩy doanh số bán hàng.
- **Cách mạng hóa quản lý chuỗi cung ứng:** Trong một chuỗi cung ứng, hệ thống định vị trong nhà mang lại những cải tiến đáng kể về tốc độ một sản phẩm tới được tay người tiêu dùng. Kho hàng và trung tâm phân phối là những không gian rộng lớn, nơi việc theo dõi sự di chuyển của hàng hóa là rất quan trọng. IPS có thể cung cấp dữ liệu vị trí theo thời gian thực, cho phép kiểm soát hàng tồn kho tốt hơn. Xe dẫn đường tự động được trang bị khả năng định vị trong nhà có thể tối ưu hóa việc di chuyển và điều hướng trong các môi trường phức tạp, từ đó tìm kiếm hàng hóa nhanh hơn với sai số thấp hơn. Mức độ tự động hóa và độ chính xác cao rất cần thiết để đáp ứng nhu cầu của thương mại hiện đại, đặc biệt trong kỷ nguyên thương mại điện tử.
- **Quản lý tòa nhà thông minh:** Tòa nhà thông minh là một mục tiêu mới giúp định vị trong nhà được thúc đẩy mạnh mẽ. Trong các tòa nhà, IPS có thể góp phần nâng cao hiệu quả sử dụng năng lượng, an ninh và sự thoải mái cho người sử dụng. Định vị trong nhà cũng có thể tăng cường hệ thống an ninh bằng cách

giám sát chuyển động của các cá nhân và phát hiện truy cập trái phép trong thời gian thực. Hơn nữa, người quản lý có thể sử dụng dữ liệu IPS để tối ưu hóa việc sử dụng không gian, nâng cao sự thoải mái cho người sử dụng.

- **Hỗ trợ ứng phó khẩn cấp:** Trong các tình huống khẩn cấp như hỏa hoạn, động đất hoặc các mối đe dọa an ninh, việc biết vị trí chính xác của các cá nhân trong tòa nhà là rất quan trọng để ứng phó và sơ tán hiệu quả. Hệ thống định vị trong nhà có thể cung cấp cho những người cứu thương dữ liệu thời gian thực về vị trí của mọi người, cho phép các hoạt động cứu hộ diễn ra nhanh hơn và hiệu quả hơn.

1.1.2 Định vị ngoài trời

So với các công nghệ định vị mặt đất khác, dải băng rộng của hệ thống 5G mmWave là ứng cử viên tốt nhất để xác định vị trí các đối tượng ở môi trường ngoài trời có độ che chắn cao. Vấn đề định vị và truyền thông được kỳ vọng sẽ có mối quan hệ mật thiết kể từ 5G trở đi. Theo các nghiên cứu gần đây, thông tin vị trí trong hệ thống 5G có các tầm quan trọng nhất định.

- Thứ nhất, thông tin vị trí được sử dụng cho các ứng dụng độc lập với truyền thông di động. Hưởng lợi nhiều nhất là các dịch vụ dựa trên vị trí như các ứng dụng bản đồ, gọi xe thông minh, quảng cáo, tiếp thị, v.v. trong môi trường GPS bị hạn chế như trung tâm thành phố lớn.
- Thứ hai, thông tin vị trí chính xác giúp nhà cung cấp dịch vụ thiết kế một mạng với khả năng phục vụ nhiều người hơn với độ trễ thấp hơn thông qua các chiến lược phân bổ tài nguyên. Trong [95, 96] các tác giả đã chứng minh được rằng một hệ thống liên lạc có thể được hưởng lợi từ thông tin vị trí nếu nó có thể khai thác tính nhất quán của kênh truyền vô tuyến và chuyển động của người dùng. Điều này đạt được nhờ các thiết bị di động báo cáo vị trí hiện tại và hướng di chuyển của chúng.
- Thứ ba, một hướng nghiên cứu tiềm năng mà các thế hệ mạng tiếp theo có thể

đạt được là khả năng định vị cho hệ thống xe thông minh. Thông tin vị trí trong trường hợp này đảm bảo yếu tố an toàn trong giao thông.

Có thể thấy, động lực nghiên cứu các thuật toán định vị cho môi trường hạn chế GPS bắt nguồn từ tiềm năng cải thiện chất lượng cuộc sống, từ tăng cường các dịch vụ chăm sóc sức khỏe, cách mạng hóa trải nghiệm bán lẻ, quản lý tòa nhà thông minh đến thiết kế mạng thông minh hay đảm bảo an toàn trong hệ thống giao thông thông minh. Khi công nghệ tiếp tục phát triển, tầm quan trọng của việc định vị và theo dõi chính xác sẽ ngày càng tăng lên, thúc đẩy sự đổi mới và nghiên cứu hơn nữa trong lĩnh vực này. Vì vậy, luận án này sẽ tập trung nghiên cứu và phát triển các thuật toán định vị sử dụng sóng vô tuyến, phục vụ các loại môi trường mà tín hiệu GPS bị hạn chế như môi trường trong nhà và môi trường ngoài trời với tỉ lệ che chắn cao, ví dụ môi trường đô thị với nhiều nhà cao tầng. Việc nghiên cứu các thuật toán này giúp tạo ra sự liên mạch trong định vị, từ đó nâng cao chất lượng các dịch vụ dựa trên vị trí.

1.2 Các nghiên cứu liên quan

Trong những năm qua, lĩnh vực định vị thu hút sự chú ý đáng kể của các nhà khoa học cũng như các doanh nghiệp. Bằng chứng cho thấy số lượng nghiên cứu và bằng sáng chế tăng dần qua các năm. Cộng đồng nghiên cứu có riêng cho mình một chuỗi hội nghị về định vị và điều hướng trong nhà có tên là *International Conference on Indoor Positioning and Indoor Navigation (IPIN)* và một tạp chí về định vị liên mạch có tên *IEEE Journal of Indoor and Seamless Positioning and Navigation (J-ISPIN)*. Mặc dù có rất nhiều công nghệ và kỹ thuật định vị đã được sử dụng và phát triển, luận án chỉ tập trung vào các nghiên cứu liên quan tới định vị trong nhà sử dụng RSS vì sự phổ biến của nó và định vị ngoài trời sử dụng sóng mmWave vì tiềm năng của nó trong các mạng di động thế hệ tiếp theo.

1.2.1 Định vị trong nhà với phép đo RSS

Cường độ tín hiệu (RSS) là một trong những phép đo được ưa chuộng nhất trong nghiên cứu và phát triển các hệ thống định vị trong nhà. Với phép đo RSS, nhóm tác

giả của nghiên cứu [114] đã phát triển một thuật toán định vị hiệu quả bằng cách sử dụng phương pháp bình phương tối thiểu với ma trận trọng số được tính toán liên quan tới công suất nhận và hệ số lan truyền. Các tác giả trong [115] cho rằng RSS có thể thay đổi theo thời gian, do đó việc sử dụng mô hình suy hao cố định sẽ gây ra lỗi cho quá trình chuyển đổi RSS sang khoảng cách. Họ vì thế đã đề xuất phương pháp hiệu chỉnh RSS theo thời gian thực để thích ứng với sự thay đổi đó. Trong nghiên cứu này, các máy thu được đặt xung quanh một khu vực để đo sự thay đổi về RSS từ một máy phát. Sau đó, thông tin về RSS sẽ được chuyển đến và xử lý tại máy chủ trước khi được truyền xuống nút mục tiêu để giúp nó xác định vị trí. Hơn nữa, các tác giả cũng sử dụng mạng nơ-ron nhân tạo (ANN) và phương pháp tối ưu hóa bầy đàn để huấn luyện mô hình ước lượng khoảng cách thêm chính xác. Một phương pháp đơn giản hơn được khai thác để nâng cao chất lượng cho hệ thống định vị trong nhà là sử dụng bộ lọc Kalman. Các tác giả trong [116] đưa các giá trị RSS đi qua bộ lọc Kalman để ước tính khoảng cách giữa thiết bị mục tiêu và đèn hiệu Bluetooth. Kết quả thực nghiệm cho thấy độ chính xác định vị được cải thiện hơn 78.9% so với phương pháp không sử dụng bộ lọc với sai số trung bình khoảng 0.53 m.

Những nỗ lực trong việc cải thiện độ chính xác của thuật toán dựa trên khoảng cách được trình bày trong [1] và [12]. Trong [1], nhóm tác giả đã đề xuất một phương pháp giúp cải thiện độ chính xác của các phương pháp dựa trên RSS. Trong nghiên cứu này, hiệu ứng hấp thụ do cơ thể con người gây ra được giảm bớt nhờ quá trình ước lượng hệ số suy hao (PLE) tức thời. Sau khi có được PLE tức thời, các giá trị RSS có thể được sử dụng trực tiếp cho cả phương pháp dựa trên và không dựa trên khoảng cách. Trong [12], một thuật toán Trilateration dựa trên mối quan hệ hình học có trọng số được đề xuất. Với việc sử dụng phân tập tần số và lọc Kalman, thuật toán đề xuất đạt được độ chính xác cao với sai số 90% nhỏ hơn 1,82 m đối với trường hợp máy thu di chuyển và sai số trung bình bằng 0,7 m đối với trường hợp máy thu ở trạng thái tĩnh. Trong [8], H. Xia và cộng sự đã tận dụng thông tin bản đồ và bộ lọc hạt (particle filter) để cải thiện độ chính xác. Trong đó, bản đồ được chia thành hai khu vực bao gồm khu vực có thể tiếp cận và không thể tiếp cận. Họ lập luận rằng chỉ những điểm của bộ lọc hạt nằm trong vùng có thể tiếp cận mới được đánh trọng số và

xác định vị trí. Điều tương tự cũng được áp dụng cho các bước di chuyển của người dùng. Tất cả những bước mà bộ lọc tạo ra và đi qua các khu vực không thể tiếp cận đều sẽ được gán trọng số bằng 0. Kết quả cho thấy phương pháp mà họ đề xuất có độ chính xác cao với sai số trung bình khoảng 1.5 m.

RADAR [20] là một trong những khung tham chiếu đầu tiên của phương pháp sử dụng bản đồ vô tuyến (FP). Theo cách tiếp cận này, bản đồ vô tuyến được tạo bằng cách thu thập RSS từ tất cả các máy phát tại các vị trí tham chiếu trên bản đồ. Nhiệm vụ chính của phương pháp này là xác định vị trí của mục tiêu bằng cách so sánh các giá trị RSS đang có với toàn bộ cơ sở dữ liệu sử dụng khoảng cách Euclid. Tuy nhiên, RSS có xu hướng dao động khi có sự thay đổi trong môi trường xảy ra. Do đó, HORUS [21] đã được đề xuất để khắc phục hạn chế này. Thay vì sử dụng trực tiếp các giá trị RSS, phương pháp này sử dụng phân phối của RSS để khớp dữ liệu. So với RADAR, HORUS cho độ chính xác cao hơn nhưng thời gian yêu cầu để ước lượng vị trí lâu hơn vì nó phải thu thập được một lượng phép đo nhất định rồi mới có thể ước lượng được vị trí. Nói cách khác, HORUS có được độ chính xác cao hơn nhờ tận dụng được lợi thế về thông tin tốt hơn. Các nghiên cứu gần đây, ví dụ như [22, 23], đã xem xét việc sử dụng thông tin trạng thái kênh (CSI) vì nó được coi là ổn định hơn và có thể cung cấp thông tin chi tiết hơn so với RSS.

FP kết hợp với thuật toán k -Nearest Neighbor (KNN) cũng là một phương pháp phổ biến trong nghiên cứu định vị trong nhà. Trong nghiên cứu gần đây, Rohan và cộng sự [106] đề xuất một hệ thống kết hợp sử dụng BLE và cảm biến IMU của điện thoại thông minh. Họ đề xuất một phương pháp sử dụng đồng thời KNN và suy luận Bayes để cải thiện độ chính xác của hệ thống. Thuật toán cho hiệu năng tốt hơn 25% so với KNN thông thường và độ chính xác trung bình khoảng 1 m. Để giảm bớt gánh nặng tính toán, Loizos và cộng sự [8] đã đề xuất thuật toán i KNN cho IPS. Trong đó, họ sử dụng hệ thống WiFi để tạo bản đồ vô tuyến và khai thác các vùng lân cận của tín hiệu BLE để giới hạn phạm vi tìm kiếm cho KNN. Phương pháp đề xuất cung cấp một giải pháp ước tính vị trí nhanh với độ chính xác cao.

Học sâu (deep learning) - một kỹ thuật tiên tiến trong học máy, đang được sử dụng ngày càng nhiều trên nhiều lĩnh vực khác nhau. Một trong số chúng chính là định vị

[110, 111, 112, 113]. Đóng vai trò là công cụ ước tính linh hoạt, mạng học sâu [107] có khả năng hiểu được cách mà tín hiệu thay đổi theo cách phức tạp nhờ dữ liệu thu thập được. Có thể nói, độ chính xác của các hệ thống định vị sử dụng mạng học sâu càng trở nên chính xác nếu người ta có càng nhiều dữ liệu về RSS và môi trường.

Với các phương pháp định vị sử dụng khoảng cách, người ta đều phải thực hiện theo hai bước: *i)* ước lượng khoảng cách và sau đó là *ii)* ước lượng vị trí. Trong bước thứ nhất, các giải pháp dạng đóng được ưu tiên hơn (các phương pháp tuyến tính) các giải pháp dựa trên tìm kiếm (các phương pháp phi tuyến tính). Trong [19], nhóm tác giả đã đề xuất một hệ thống định vị theo thời gian thực thông qua RSS từ các thẻ RFID. Đầu tiên họ sử dụng bộ lọc Kalman cho các giá trị RSS chứa nhiễu với mục đích có được các giá trị khoảng cách với độ ổn định và chính xác cao hơn. Trong bước ước tính vị trí, các tác giả đã giới thiệu phương pháp *Heron-Bilateration*, sử dụng hai thẻ RFID và công thức Heron thay vì ba điểm như phương pháp *Trilateration* để xác định vị trí của nút mục tiêu. Kết quả cho thấy hệ thống đề xuất đạt được độ chính xác cao hơn với số nút tham chiếu ít hơn so với phương pháp sử dụng khoảng cách xấp xỉ và *Trilateration*. Các tác giả trong [31] đã đề xuất một giải pháp định vị dựa trên RSS thời gian thực trong điều kiện công suất phát không biết trước và hệ số suy hao (PLE) không phải là một giá trị cố định. Trong nghiên cứu này, PLE được mô hình hóa như một biến ngẫu nhiên Gaussian với giá trị trung bình và phương sai đã biết. Điều này giúp họ giới hạn được độ ngẫu nhiên của PLE. Giới hạn này được coi là phù hợp với môi trường thực tế trong nhà hơn một môi trường lý tưởng. Sau đó, các tác giả đã đề xuất phương pháp WLS hai bước dạng đóng (two-step weighted least square) để giải quyết bài toán. Trong [32], các tác giả đã đề xuất phương pháp định vị dựa trên RSS/TOA bao gồm hai bước: *i)* hiệu chỉnh và *ii)* ước tính vị trí. Trong bước hiệu chỉnh, các tác giả thu thập thông tin từ sự hợp tác giữa các nút cố định và một nút mục tiêu để có được các thông số chính xác hơn về khoảng cách. Tiếp theo, họ đề xuất một giải pháp dạng đóng dựa trên bình phương nhỏ nhất (LS) cho các tham số đã hiệu chỉnh để tìm vị trí nút mục tiêu trong bước thứ hai. Kết quả cho thấy phương pháp đề xuất đạt được độ chính xác gần như tương đương với các phương pháp tối đa hàm khả năng (MLE) với chi phí tính toán thấp hơn.

1.2.2 Định vị ngoài trời với sóng mi-li-mét (mmWave)

Chủ đề nghiên cứu về định vị sử dụng mmWave bắt đầu từ năm 2016. Đến thời điểm hiện tại, chủ đề này vẫn đang tiếp tục nghiên cứu về vấn đề này. Ý tưởng đầu tiên về định vị trong mạng 5G sử dụng một trạm gốc với hai phép đo là thời gian truyền (TOA) và hướng sóng đến (AOA). Nghiên cứu [79] trình bày một khung kết hợp hai phép đo này và bộ lọc Kalman mở rộng để giải quyết một tình huống thực tế, trong đó cả vị trí người dùng và độ lệch đồng bộ đều thay đổi theo thời gian. Độ chính xác đạt được là khá cao, tuy nhiên điều kiện xem xét trong nghiên cứu này là môi trường không có vật cản (LOS) - một điều kiện không thường xuyên có trong môi trường thực tế. Nhóm tác giả trong [80] đề xuất phương pháp định vị và điều hướng cho các phương tiện thông minh sử dụng mmWave trong mạng 5G. Đầu tiên, các tác giả tính toán các giới hạn cơ bản về độ chính xác của vị trí và góc di chuyển thông qua giới hạn Cramér-Rao (CRLB) với ba tình huống: đường truyền thẳng (LOS), đường truyền bị che khuất (NLOS) và đường truyền bị che khuất không thường xuyên. Các giới hạn này chỉ ra rằng thông tin của các đường truyền NLOS giúp ước tính vị trí và hướng của người dùng. Các tác giả cũng đề xuất một kỹ thuật để ước tính vị trí và hướng thông qua ba bước để đạt được giới hạn cơ bản. Bước đầu tiên của thuật toán là khai thác sự thừa của kênh truyền mmWave trong miền góc. Trong bước này, Distributed Compressive Sensing - Simultaneous Orthogonal Matching Pursuit (DCS-SOMP), một thuật toán lấy mẫu nén được áp dụng để ước lượng thô các phép đo TOA, AOA và AOD. Bước thứ hai, Space-Alternating Generalized Expectation-Maximization Algorithm (SAGE), một thuật toán tối đa hóa kỳ vọng, được dùng để tinh chỉnh ước lượng của các thông số kênh. Cuối cùng, bước thứ ba chuyển các thông số ước lượng thành vị trí trong không gian 2 chiều thông qua các mối quan hệ hình học.

Một hướng nghiên cứu tiềm năng cho mạng 5G là định vị cho hệ thống xe thông minh. Mạng 5G giới thiệu một khối chức năng gọi là “road-side units (RSU)”. Chúng có thể được sử dụng như bộ chuyển tiếp để truyền dữ liệu từ phương tiện này sang phương tiện khác hoặc đến các thiết bị khác thuộc lớp mạng cao hơn. Hệ thống định vị sử dụng RSU cho các hệ thống xe thông minh đã được giới thiệu trong [81]. Trong nghiên cứu này, các tác giả đầu tiên đánh giá việc triển khai các RSU để đạt được một

độ chính xác cho trước. Tiếp theo, các tác giả tìm mật độ tối thiểu của các RSU nằm dọc theo các con đường sao cho khả năng môi trường truyền giữa các RSU và phương tiện luôn là LOS. Để xác định vị trí mục tiêu, các tác giả đã sử dụng kết hợp TOA và AOA. Cuối cùng, CRB được sử dụng để đánh giá độ chính xác của hệ thống. Ngoài ra, nhiều tình huống khác nhau đã được đặt ra cho vấn đề định vị xe thông minh. Ví dụ, nghiên cứu [83] đặt ra vấn đề định vị hợp tác để giúp hai phương tiện tránh va chạm tại các giao lộ, nơi các phương tiện đang ở trong môi trường không có GPS. Hai xe phải kết nối với nhau để trao đổi thông tin về vị trí của nhau, từ đó đưa ra quyết định tránh va chạm. Nghiên cứu [84] xem xét vấn đề định vị hợp tác cho các phương tiện thông minh. Vấn đề đặt ra trong bài báo là việc trao đổi thông tin giữa các phương tiện đôi khi gặp khó khăn do lượng thông tin cần xử lý là rất lớn, do đó tiêu tốn một lượng lớn tài nguyên. Vì vậy, mục tiêu của nghiên cứu này là tìm ra phương pháp lập kế hoạch tối ưu sử dụng một lượng phép đo tối thiểu. Các tác giả đề xuất phương pháp dựa trên tính toán phân tán cho phép các phương tiện trao đổi thông tin mà không cần đến bộ điều khiển trung tâm. Vấn đề được trình bày dưới dạng một quá trình Markov có thể quan sát được một phần và được giải quyết bằng cách sử dụng các thuật toán lập lịch với sự trợ giúp của học tăng cường (RL).

Các nghiên cứu mới về định vị trong mạng 5G đến từ 2 nhóm nghiên cứu lớn trên thế giới là nhóm nghiên cứu của giáo sư T.S. Rappaport [85][86][93][94] tại Đại học New York và giáo sư H. Wymeersch tại Đại học Công Nghệ Chalmers [80][87]. Nghiên cứu [85] xem xét khả năng định vị cho mạng kết nối phương tiện với phương tiện (V2V) sử dụng ăng-ten MIMO và sóng mmWave. Nghiên cứu đã chứng minh rằng mạng V2V cho ước lượng AOA, AOD và TOA tốt hơn khi các phương tiện chia sẻ các thông tin về kênh truyền với nhau. Trong khi đó, [86] không những xem xét khả năng định vị trong mạng 5G mà còn trong những mạng di động tương lai. Nhóm tác giả đã chứng minh rằng việc kết hợp dữ liệu (data fusion), học máy và định vị hợp tác cải thiện độ chính xác của vị trí trong mạng 5G và các thế hệ mạng trong tương lai. Nghiên cứu [87] trình bày một phương pháp định vị trong mạng 5G MISO và sóng mmWave với phép đo tức thời (snap-shot measurement). Bài toán được xuất phát từ ý tưởng định vị chỉ sử dụng một phép đo (không qua huấn luyện) và yêu cầu tính toán thời

gian thực. Để giải quyết vấn đề này, nhóm tác giả đề xuất một phương pháp xấp xỉ bài toán tối đa hàm khả năng (MLE) nhiều tham số nhờ khai thác sự thừa của kênh truyền sóng mmWave. Kết quả cho thấy độ chính xác đạt được ngưỡng dưới CRLB với chi phí tính toán trung bình. Các nhóm này cũng đã chứng minh khả năng định vị của sóng mmWave trên thực tế thông qua các thử nghiệm của họ [89, 90, 91]. Trong [91], các tác giả chứng minh mô hình định vị mmWave bằng cách sử dụng dữ liệu đo đạc thực tế. Họ sử dụng kết hợp thông tin AOA và TOA để ước tính vị trí của mục tiêu thông qua cả đường dẫn LOS và NLOS từ một BS duy nhất. Kết quả cho thấy hệ thống đạt độ chính xác trung bình là 5,72 cm ở 28 GHz và 6,29 cm ở 140 GHz. Trong [92], các tác giả cũng trình bày một thử nghiệm thực tế về định vị sử dụng một BS. Nghiên cứu sử dụng một trạm BS 5G mmWave thương mại và bộ công cụ phát triển thiết bị người dùng (UE) được gắn trên một phương tiện thử nghiệm. Các tác giả đã trình bày các mô hình tín hiệu có liên quan, phương pháp xử lý tín hiệu (bao gồm ước tính tham số kênh và ước tính vị trí) và xác thực tính hiệu quả của chúng dựa trên dữ liệu thực được thu thập trong môi trường ngoài trời. Độ chính xác mà họ có thể đạt được là 99% sai số dưới 2,5 mét với ba phép đo AOA, AOD và RTT.

1.3 Thách thức và hạn chế

Có thể thấy, các giải pháp định vị và dẫn đường là chìa khóa để xây dựng và nâng cao trải nghiệm người dùng qua dịch vụ định vị dựa trên vị trí (LBS). Trên thực tế, một số giải pháp định vị trong nhà đã được đề xuất thành công cho các môi trường trong nhà phức tạp và đông đúc, ví dụ: sân bay quốc tế [121], khuôn viên trường đại học [122], bệnh viện [123] hoặc bảo tàng [124]. Tuy nhiên, các công nghệ định vị dường như chưa có đủ tính thực nghiệm để triển khai LBS ở khắp mọi nơi, một phần cho môi trường trong nhà và đặc biệt cho môi trường ngoài trời có tỉ lệ che chắn cao. Các hạn chế của các nghiên cứu về định vị mặt đất hiện tại bao gồm:

- Độ chính xác và khả năng thích ứng với môi trường phức tạp.
- Chi phí cho thiết bị, vận hành và bảo trì chưa được tối ưu.

- Khả năng mở rộng của các hệ thống định vị trong mạng lớn (như thành phố thông minh hoặc hàng triệu thiết bị IoT).
- Các công nghệ mạng tế bào hiện tại chưa đạt được độ chính xác cao cho định vị ngoài trời.

1.4 Mục tiêu và đóng góp của luận án

1.4.1 Mục tiêu của luận án

Mục tiêu của luận án là giải quyết những hạn chế ảnh hưởng trực tiếp tới việc thực thi các thuật toán định vị, phục vụ các dịch vụ dựa trên vị trí như giảm độ phức tạp thuật toán, nâng cao độ chính xác của các thuật toán định vị.

1.4.2 Đóng góp của luận án

Luận án đề xuất các giải pháp định vị hoặc hỗ trợ định vị mang tính thực nghiệm cho hệ thống định vị mặt đất. Các đóng góp chính của luận án được liệt kê như sau:

1. Đề xuất các thuật toán định vị hoặc hỗ trợ định vị cho môi trường trong nhà sử dụng đèn hiệu BLE.
 - *Một chiến lược đặt đèn hiệu BLE hỗ trợ hệ thống định vị trong nhà sử dụng cảm biến điện thoại di động và RSS*: Chiến lược này cho phép triển khai hệ thống định vị trong nhà với số lượng đèn hiệu thấp nhất xong vẫn có khả năng duy trì độ chính xác cao.
 - *Một phương pháp định vị trong nhà sử dụng đèn hiệu BLE giúp cân bằng được ưu và nhược điểm của các phương pháp tiếp cận: sử dụng khoảng cách và không sử dụng khoảng cách (FP)*. Với FP, phương pháp đề xuất chia khu vực định vị thành các ô có diện tích lớn hơn rất nhiều so với phương pháp FP thông thường. Với cách làm này, chúng ta chỉ có được ước lượng thô về vị trí người dùng thông qua mạng nơ-ron nhân tạo. Điểm khác biệt so với các nghiên cứu khác là vị trí thô ước lượng được

mang thêm thông tin về các đèn hiệu gần đó. Các đèn hiệu này được coi là tin cậy và được sử dụng cho quá trình định vị tinh sử dụng một phương pháp WLS đề xuất ở phía sau.

2. Các thuật toán ước lượng kênh hỗ trợ định vị ngoài trời trong dụng hệ thống MIMO mmWave

- **01 phương pháp ước lượng kênh thô dựa trên thuật toán *Stagewise Orthogonal Matching Pursuit (StOMP)* [108] không xem xét sự ảnh hưởng lẫn nhau giữa các đường dẫn.** Trong phần này, luận án xem xét bài toán ước lượng kênh mmWave mà không cần biết trước số lượng đường dẫn. Đầu tiên, luận án đề xuất thuật toán StOMP phân tán, viết tắt là DStOMP (Distributed StOMP), giúp ước lượng kênh trên nhiều sóng mang con thay vì một sóng mang duy nhất. Nó dựa trên thực tế rằng độ chính xác của một hệ thống luôn tốt hơn khi có nhiều phép đo hơn. Tận dụng DStOMP và cấu trúc đặc biệt của kênh, luận án sau đó đề xuất một thuật toán phân cụm dựa trên phương pháp Silhouette có khả năng xác định số lượng đường dẫn.
- **01 phương pháp ước lượng kênh thô dựa trên thuật toán *OMP* có xem xét sự ảnh hưởng lẫn nhau giữa các đường dẫn.** [73] cho rằng các đường dẫn là độc lập với nhau khi số lượng ăng-ten lớn đến vô hạn, do đó có thể ước lượng chúng riêng biệt. Tuy nhiên, điều này không đúng với thực tế vì số lượng ăng-ten là hữu hạn. Vì vậy, trong nghiên cứu này, luận án đề xuất một phương pháp giúp tận dụng cấu trúc Dirichlet của kênh truyền trong miền góc để cố gắng tách sự ảnh hưởng của các đường dẫn tới nhau khi sử dụng thuật toán OMP.
- **01 phương pháp tinh chỉnh các tham số thô có độ phức tạp tính toán thấp song vẫn đạt được độ chính xác cao.** Trong phần này, luận án dựa vào [57] để phát triển một phương pháp tinh chỉnh các tham số hỗ trợ định vị có độ phức tạp tính toán thấp hơn song vẫn đạt được độ chính xác cao. Đầu tiên, thuật toán đề xuất thực hiện tìm kiếm sử dụng

tỉ lệ vàng trên AOA và AOD, sau đó áp dụng LS cho các cột ứng với cặp AOA và AOD vừa tìm được để ước lượng hệ số kênh và xây dựng cơ sở mới theo hướng cơ sở ở bước sau chịu ảnh hưởng của hiệu ứng lệch cơ sở ít hơn bước trước. Sau đó, TOA và độ lợi kênh thực được ước lượng lại bằng cách sử dụng cơ sở vừa tìm thấy ở bước trước thông qua thuật toán DCS-SOMP [64]. Theo cách này, thuật toán đề xuất cho độ chính xác cao tương đương với [57] trong khi đạt được tốc độ tính toán nhanh hơn gấp hai lần.

1.5 Cấu trúc luận án

Chương 1 giới thiệu động lực, các nghiên cứu liên quan, các đóng góp chính và cấu trúc của luận án.

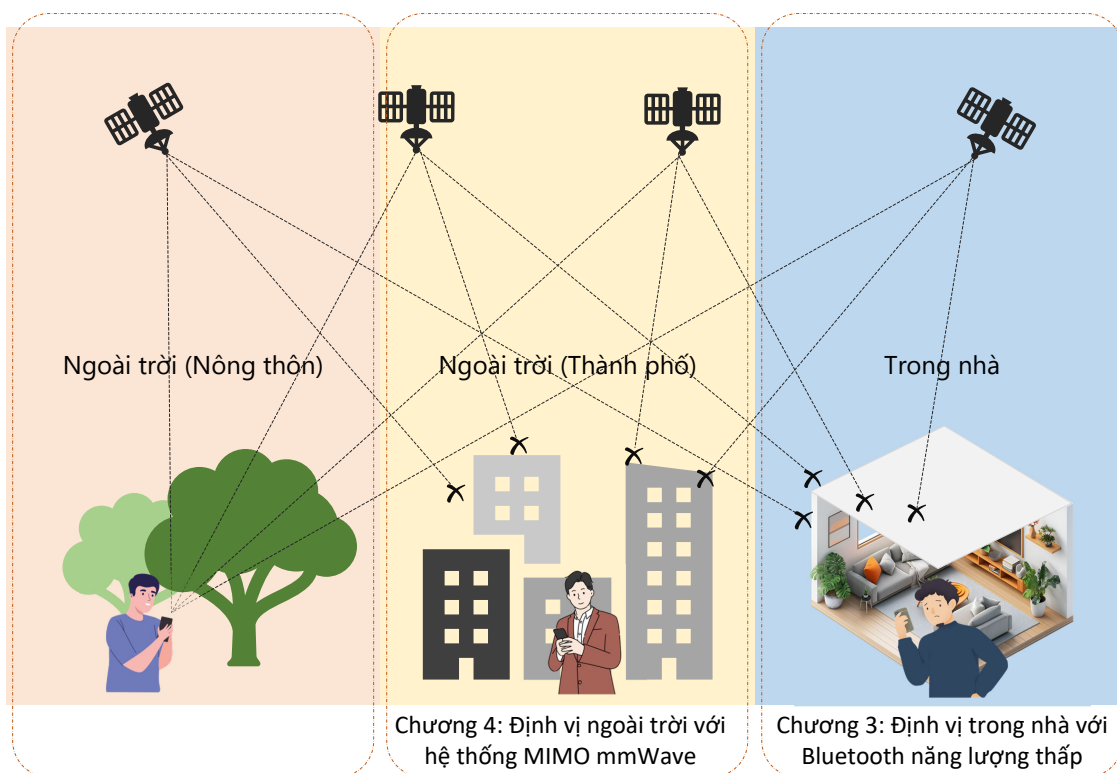
Chương 2 của luận án trình bày các khái niệm và kiến thức cơ bản liên quan tới hệ thống định vị, bao gồm các công nghệ và kỹ thuật định vị phổ biến. Đây là nền tảng để phát triển các thuật toán định vị hoặc hỗ trợ định vị cho các chương sau.

Chương 3 trình bày các thuật toán định vị trong nhà sử dụng đèn hiệu BLE. Trong chương này, luận án đề xuất hai phương pháp định vị bao gồm: một phương pháp quy hoạch đèn hiệu BLE, được trình bày trong nghiên cứu:

- Thai-Mai Dinh-Thi, Ngoc-Son Duong, Quoc-Tuan Nguyen, "Developing a Novel Real-Time Indoor Positioning System Based on BLE Beacons and Smartphone Sensors," *IEEE Sensors Journal*, vol. 21, no. 20, pp. 23055-23068, Oct. 2021, doi: 10.1109/JSEN.2021.3106019.

và một phương pháp lựa chọn đèn hiệu để xác định vị trí, được trình bày trong 02 nghiên cứu:

- Ngoc-Son Duong, Thanh-Phuc Nguyen, Quoc-Tuan Nguyen, Thai-Mai Dinh Thi, "Improving indoor positioning system using weighted linear least square and neural network," *International Journal of Sensor Networks*, vol. 41, no. 2, pp. 67-77, Mar. 2023, doi: 10.1504/IJSNET.2023.129632.



Hình 1.1: Các kịch bản nghiên cứu để tạo ra định vị liên mạch.

- Ngoc-Son Duong, Ngoc-Thuy Nguyen, Phuong-Dung Chu-Thi, Quoc-Tuan Nguyen and Thai-Mai Dinh-Thi, “Cải thiện độ chính xác của hệ thống định vị trong nhà dựa trên phân tích lỗi truyền lan”, *Hội thảo Quốc gia lần thứ XXIV về Điện tử, Truyền thông và Công nghệ Thông tin*, Hanoi, 2021

Kết quả cho thấy chúng giúp có độ chính xác cao đồng thời giải quyết được các vấn đề thực tế như giảm chi phí triển khai hệ thống hoặc cân bằng ưu-nhược điểm của các phương pháp định vị.

Chương 4 trình bày các thuật toán ước lượng các tham số liên quan tới định vị, bao gồm AOA, AOD và TOA cho hệ thống mmWave MIMO. Các thuật toán trong chương này có thể chia thành 2 loại thuật toán là ước lượng thô và ước lượng tinh. Với thuật toán ước lượng thô, luận án đề xuất 2 phương pháp ước lượng kênh dựa trên nền tảng của các thuật toán ước lượng thưa kinh điển như OMP hoặc StOMP. Các cải tiến đều mang tính thực nghiệm cao khi chúng giải quyết được các vấn đề như: Không giả sử biết trước độ thưa của tín hiệu hoặc xét ảnh hưởng của các đường dẫn tới nhau.

Một số công trình của tác giả liên quan tới phần ước lượng thô là:

- Ngoc-Son Duong, Lan-Nhi Vu Thi, Phuong-Dung Chu Thi, Sinh-Cong Lam, Thai-Mai Dinh Thi “A novel Distributed Stagewise Orthogonal Matching Pursuit algorithm for mmWave MIMO channel estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2024, doi: 10.1587/transfun.2024EAL2064.
- Ngoc-Son Duong, Thu-Trang Nguyen, Lan-Nhi Vu Thi, Sinh-Cong Lam, Thai-Mai Dinh-Thi, “An improved OMP-based mmWave channel estimation for MIMO systems," in *Proc. The 3rd International Conference on Advances in Information and Communication Technology*, 2024, doi: 10.1007/978-3-031-80943-9_97

Với thuật toán ước lượng tinh, luận án đề xuất một phương pháp có độ chính xác cao đồng thời có chi phí tính toán thấp hơn so với phương pháp thuần Newton. Một số công trình của tác giả liên quan tới phần ước lượng tinh là:

- Ngoc-Son Duong, Quoc-Tuan Nguyen, Thai-Mai Dinh-Thi, “OMP-Based Channel Estimation With Dynamic Grid for mmWave MIMO Positioning Systems," *IEEE Communications Letters*, vol. 27, no. 10, pp. 2623-2627, Oct. 2023, doi: 10.1109/LCOMM.2023.3303453.
- Ngoc-Son Duong, Quoc-Tuan Nguyen, Khac-Hoang Ngo, Thai-Mai Dinh-Thi, “Sparse Bayesian Learning with Atom Refinement for mmWave MIMO Channel Estimation," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, 2023, doi: 10.1109/SSP53291.2023.10208044.
- Ngoc-Son Duong, Quoc-Tuan Nguyen, Thai-Mai Dinh-Thi, “mmWave Channel Estimation for Location-based Application in 5G MIMO systems," in *Proc. 2022 IEEE Region 10 Conference (TENCON)*, Hong Kong, 2022, pp. 1-6, doi: 10.1109/TENCON55691.2022.9977818.

Chương 5 đưa ra kết luận cho luận án và một số công việc trong tương lai.

Chương 2

Nền tảng cơ bản

2.1 Các công nghệ định vị không dây

Với môi trường ngoài trời có tính mở, GNSS là lựa chọn đầu tiên và có lẽ là tốt nhất để định vị và dẫn đường. Trên thực tế, điện thoại thông minh hiện đại hầu như được trang bị chip GNSS và phần mềm điều hướng (ví dụ Google Maps) để tìm vị trí của điện thoại trên bản đồ và chỉ dẫn người dùng đến điểm mong muốn. Tuy nhiên, khi di chuyển từ không gian mở ngoài trời vào một không gian có tính đóng, độ chính xác của GNSS sẽ giảm đi. Để giải quyết vấn đề này, các công nghệ khác được sử dụng để tạo nên một giải pháp định vị liền mạch (seamless positioning). Các công nghệ này có tính cục bộ, được nhúng thẳng vào thiết bị cần xác định vị trí, ví dụ như hệ cảm biến quán tính, hoặc xây dựng một hệ thống định vị sử dụng sóng vô tuyến mặt đất hoặc dựa trên chính các yếu tố tự nhiên như từ tính của trái đất. Ví dụ, hệ thống tích hợp GNSS/INS có thể cung cấp thông tin vị trí tương đối chính xác, ngay cả trong môi trường không có tín hiệu GPS trong một khoảng thời gian ngắn. Sau khi di chuyển từ ngoài trời vào trong nhà, hệ thống định vị cục bộ này sẽ đo đạc chất lượng tín hiệu GNSS và quyết định chuyển qua định vị sử dụng hệ cảm biến quán tính nếu chất lượng tín hiệu GNSS không đảm bảo. Vì trọng tâm nghiên cứu của luận án này là hệ thống định vị sử dụng các công nghệ không dây, luận án sẽ chỉ thảo luận về các công nghệ không dây phổ biến, góp mặt nhiều nhất trong các nghiên cứu về định vị mặt đất hiện tại.

2.1.1 Wi-Fi [119]

Một trong những ưu điểm chính của công nghệ Wi-Fi là các điểm truy cập Wi-Fi được triển khai ở khắp mọi nơi trong các tòa nhà. Phương pháp sử dụng bản đồ cường độ tín hiệu (FP) được coi là phương pháp hiệu quả nhất để xác định vị trí của một đối tượng trong môi trường trong nhà. Mặc dù có khả năng triển khai cao về mặt lý thuyết, nghiên cứu gần đây đã chỉ ra rằng độ chính xác của công nghệ này tương đối thấp. Sai số định vị đạt được dao động từ 5 m đến 15 m. Để đạt được độ chính xác cao, người ta cần phải lắp đặt các điểm truy cập Wi-Fi với mật độ dày. Điều này là không thực tế vì điểm truy cập Wi-Fi ban đầu được thiết kế để gửi và nhận tín hiệu thay vì mục đích định vị. Ngoài ra, công nghệ này cũng tiêu tốn năng lượng hơn các công nghệ khác như BLE hoặc Zigbee. Các nghiên cứu sử dụng WiFi AP như các trạm gốc phục vụ định vị được thống kê tương đối đầy đủ trong [118].

2.1.2 Bluetooth năng lượng thấp (BLE) [119]

Bluetooth được phát minh bởi Ericsson năm 1994 nhưng nó chỉ được xem xét trong bài toán định vị cho tới khi chuẩn Bluetooth năng lượng thấp đầu tiên được giới thiệu năm 2009 và thực sự thu hút được sự quan tâm của doanh nghiệp và các nhà nghiên cứu khi những giao thức dành riêng cho định vị ra đời năm 2013. Đặc điểm nổi bật nhất của BLE là nó có mức tiêu thụ điện năng rất thấp, cỡ vài microwatt so với 1W của Bluetooth truyền thống. Phạm vi hoạt động của BLE có thể lên đến 100 mét. Nó hoạt động ở hai chế độ là quảng bá và kết nối. Các giao thức hỗ trợ định vị được phát triển chủ yếu dựa trên chế độ quảng bá. Trong chế độ này, thiết bị BLE sẽ gửi các gói tin quảng bá định kỳ đến các thiết bị ở xung quanh khu vực phục vụ của nó. Các thiết bị sau đó có thể trích xuất được thông tin hữu ích từ chính các gói tin quảng bá mà nó nhận được. Một giao thức nổi tiếng do Apple phát triển riêng cho bài toán định vị có tên là iBeacon. Với iBeacon, người ta có thể sử dụng RSS để xác định vị trí người dùng thông qua cả hai phương pháp dựa trên và không dựa trên khoảng cách (FP).

2.1.3 Ultra Wideband (UWB) [119]

Băng siêu rộng (UWB), thuộc chuẩn IEEE 802.15.6, là một công nghệ không dây giúp truyền dữ liệu theo từng cụm (burst) trong thời gian cực ngắn. Nó hoạt động trong dải tần số từ 3.1 GHz đến 10.6 GHz với băng thông rộng (trên 500 MHz). UWB được thiết kế để truyền dữ liệu trong phạm vi ngắn và tiêu thụ điện năng thấp. Những đặc điểm này làm cho UWB phù hợp cho việc định vị và theo dõi đối tượng trong khoảng cách ngắn. Tuy nhiên, bản thân UWB cũng có một số nhược điểm. Thứ nhất, giá thành của các thiết bị UWB không rẻ. Thứ hai, UWB thường hoạt động trong khoảng cách ngắn (từ vài mét đến dưới 20 m). Điều này làm cho nó không phù hợp cho việc định vị ở khoảng cách xa hoặc trong môi trường rộng lớn. Thứ ba, việc lắp đặt và triển khai một hệ thống định vị sử dụng UWB không hề đơn giản vì nó thường yêu cầu đồng bộ chặt chẽ về mặt thời gian giữa các thiết bị trong mạng. Thứ tư, môi trường NLOS ảnh hưởng tới độ chính xác của hệ thống.

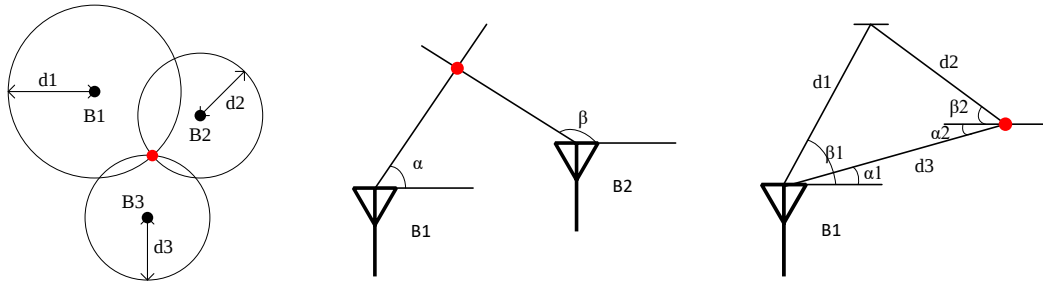
2.1.4 Sóng mmWave trong hệ thống mạng 5G [57]

Công nghệ mmWave được sử dụng trong hệ thống mạng 5G băng tần 2 (FR2) sử dụng sóng điện từ có tần số trong khoảng 30 đến 300 GHz, ứng với các bước sóng trong dải từ 1 mm đến 10 mm. Công nghệ mmWave được sử dụng trong hệ thống mạng 5G với mục đích chính là cung cấp tốc độ dữ liệu và băng thông cao với độ trễ thấp. Sóng mmWave có một số tính chất giúp nó trở nên phù hợp với mục đích định vị như tính tán xạ kém và băng thông rộng. Phần băng thông rộng của dải sóng mmWave trong mạng 5G cho phép khoảng thời gian giữa 2 lần lấy mẫu ngắn, từ đó giúp cải thiện độ chính xác của các phép đo khoảng cách dựa vào thời gian truyền giữa máy thu và máy phát. Kết hợp với hệ thống ăng-ten nhiều phần tử, hệ thống MIMO mmWave có thể tạo ra những búp sóng có độ rộng rất nhỏ, từ đó nâng cao độ chính xác cho các phép đo góc. Tính tán xạ kém của mmWave làm giảm số lượng các thành phần đa đường giữa máy thu và máy phát, do đó kênh truyền chỉ là tổng hợp của đường dẫn LOS và một số ít các đường dẫn NLOS. Một khi xác định được AOA, AOD và TOA của các đường dẫn, hệ thống định vị sử dụng sóng mmWave chỉ cần 1 BS để có thể xác định

được vị trí của người dùng.

2.2 Các phép đo của hệ thống định vị không dây

Vị trí của một nút mục tiêu (máy thu) có thể được xác định bằng cách sử dụng tín hiệu phát ra từ các nút tham chiếu (máy phát) có vị trí biết trước ¹. Trên thực tế, định vị là một trong những vấn đề trọng tâm của nhiều lĩnh vực như ra-đa, sonar, thông tin di động, mạng cảm biến không dây, tương tác người-máy, v.v. Ví dụ, vị trí của một người đang nói có thể được xác định bằng cách sử dụng mảng micro trong các ứng dụng như hội nghị truyền hình hoặc giám sát an ninh. TOA, TDOA, RSS và DOA là các phép đo thường được sử dụng nhất trong bài toán định vị. Trong số các phép này, TOA, TDOA và RSS gián tiếp cung cấp thông tin về khoảng cách giữa máy thu và máy phát, trong khi DOA cung cấp thông tin về góc đến của một tín hiệu. Mặc dù có nhiều cách kết hợp khác nhau nhưng thời gian và góc vẫn là hai phép đo có tính nền tảng trong bài toán định vị. Về mặt trực quan, các thông số khoảng cách và góc được sử dụng như trong Hình 2.1.



Hình 2.1: Từ trái qua phải: Kỹ thuật sử dụng khoảng cách, kỹ thuật sử dụng góc, kỹ thuật sử dụng kết hợp thời gian và góc.

Các phép đo được mô hình hóa thông qua công thức [109]

$$\mathbf{r} = \mathbf{f}(\mathbf{x}) + \mathbf{n}, \quad (2.1)$$

với $\mathbf{r}$ là véc-tơ tín hiệu quan sát được, $\mathbf{x}$ vị trí của mục tiêu cần xác định vị trí, $\mathbf{f}(\mathbf{x})$ là

¹Bài toán coi máy phát là nút mục tiêu và các máy thu là nút tham chiếu là một bài toán tương đương, được biết đến với tên gọi định vị nguồn phát (source localization)

một hàm phi tuyến của $\mathbf{x}$ và $\mathbf{n}$ là nhiễu Gauss. Bảng 2.1 liệt kê các ưu và nhược điểm của các phép đo.

Phép đo	Thông tin	Ưu điểm	Nhược điểm
TOA	Khoảng cách	- Độ chính xác cao	- Hoạt động kém trong môi trường NLOS - Yêu cầu đồng bộ đồng hồ giữa máy thu và máy phát
TDOA	Khoảng cách	- Độ chính xác cao - Chỉ yêu cầu đồng bộ giữa các nút tham chiếu	- Hoạt động kém trong môi trường NLOS
DOA	Góc	- Yêu cầu ít nút tham chiếu - Không yêu cầu đồng bộ	- Yêu cầu mảng ăng-ten thông minh - Hoạt động kém trong môi trường NLOS
RSS	Khoảng cách	- Đơn giản - Không yêu cầu đồng bộ	- Độ chính xác thấp hoặc trung bình

Bảng 2.1: So sánh ưu và nhược điểm của các phép đo phổ biến trong bài toán định vị

2.2.1 Thời gian đến (TOA)

TOA là khoảng thời gian tín hiệu đi từ nút tham chiếu tới nút mục tiêu. Để có được thông tin TOA chính xác, nút mục tiêu và nút tham chiếu phải được đồng bộ chính xác về mốc thời gian. Người ta sẽ không cần đồng bộ nếu sử dụng phép đo TOA hai chiều, tức RTT (Round Trip Time). Quan hệ giữa khoảng cách, ký hiệu là t , và TOA, ký hiệu là t , được biểu diễn qua công thức

$$d = t \times c, \quad (2.2)$$

trong đó c là tốc độ ánh sáng. Bài toán định vị sử dụng phép đo TOA được xây dựng như sau. Gọi $\mathbf{x} = [x, y]^\top$ và $\mathbf{x}_l = [x_l, y_l]^\top, l \in [L]$ lần lượt là tọa độ trong không gian 2 chiều của nút mục tiêu và nút tham chiếu thứ l . Khoảng cách giữa nút mục tiêu và nút tham chiếu thứ l được biểu diễn bằng công thức

$$d_l = \|\mathbf{x} - \mathbf{x}_l\|_2^2 = \sqrt{(x - x_l)^2 + (y - y_l)^2}. \quad (2.3)$$

Trong thực nghiệm, phép đo TOA thường chứa nhiễu. Do đó, khoảng cách quan sát được, $r_{\text{TOA},l}$, thực tế là

$$r_{\text{TOA},l} = d_l + n_{\text{TOA},l} = \sqrt{(x - x_l)^2 + (y - y_l)^2} + n_{\text{TOA},l}, \quad (2.4)$$

trong đó $n_{\text{TOA},l}$ là lỗi khoảng cách do nhiễu trong phép đo TOA nhiễu gây ra. Phương trình (2.4) có thể được biểu diễn bởi [109]

$$\mathbf{r}_{\text{TOA}} = \mathbf{f}_{\text{TOA}}(\mathbf{x}) + \mathbf{n}_{\text{TOA}}, \quad (2.5)$$

trong đó

$$\mathbf{r}_{\text{TOA}} = [r_{\text{TOA},1}, r_{\text{TOA},2}, \dots, r_{\text{TOA},L}]^\top, \quad (2.6)$$

$$\mathbf{n}_{\text{TOA}} = [n_{\text{TOA},1}, n_{\text{TOA},2}, \dots, n_{\text{TOA},L}]^\top \quad (2.7)$$

và

$$\mathbf{f}_{\text{TOA}}(\mathbf{x}) = \mathbf{d} = \begin{bmatrix} \sqrt{(x - x_1)^2 + (y - y_1)^2} \\ \sqrt{(x - x_2)^2 + (y - y_2)^2} \\ \vdots \\ \sqrt{(x - x_L)^2 + (y - y_L)^2} \end{bmatrix}, \quad (2.8)$$

trong đó $\mathbf{f}_{\text{TOA}}(\mathbf{x})$ là một hàm biết trước được đặc trưng bởi $\mathbf{x}$. Bài toán định vị sử dụng phép đo TOA là ước lượng $\mathbf{x}$ khi biết $\mathbf{r}_{\text{TOA}}$.

2.2.2 Chênh lệch thời gian đến (TDOA)

Lấy TOA của một nút tham chiếu đến nút mục tiêu làm tham chiếu, TDOA là phép đo đo sự khác biệt về TOA của những nguồn tham chiếu còn lại với TOA này. Như vậy, để hệ thống hoạt động chính xác, người ra chỉ cần đồng bộ mốc thời gian giữa các điểm tham chiếu với nhau. Điều này làm cho hệ thống sử dụng TDOA có tính thực nghiệm cao hơn TOA vì việc đồng bộ giữa các nút tham chiếu là dễ dàng hơn đồng bộ giữa nút tham chiếu và nút mục tiêu. Tương tự như TOA, nhận TDOA với tốc độ ánh sáng sẽ thu được sự khác biệt về khoảng cách giữa nút đích và và nhiều nút tham chiếu. Bài toán định vị sử dụng phép đo TDOA được xây dựng như sau. Đặt $d_1 = t_1 c$ làm khoảng cách tham chiếu, sự khác biệt về khoảng cách tới nút mục tiêu của nút

tham chiếu thứ l , $l = 2, 3, \dots, L$ và d_1 là

$$d_{l,1} = d_l - d_1. \quad (2.9)$$

Tương tự TOA, ta có

$$r_{\text{TDOA},l} = d_{l,1} + n_{\text{TDOA},l}, \quad l = 2, 3, \dots, L, \quad (2.10)$$

trong đó $n_{\text{TDOA},l}$ là nhiễu của sự khác biệt về khoảng cách do $t_{l,1}$ gây ra. Khi đó, mô hình tín hiệu TDOA dưới dạng véc-tơ có thể biểu diễn bởi [109]

$$\mathbf{r}_{\text{TDOA}} = \mathbf{f}_{\text{TDOA}}(\mathbf{x}) + \mathbf{n}_{\text{TDOA}}, \quad (2.11)$$

trong đó

$$\mathbf{r}_{\text{TDOA}} = [r_{\text{TDOA},2}, r_{\text{TDOA},3}, \dots, r_{\text{TDOA},L}]^\top, \quad (2.12)$$

$$\mathbf{n}_{\text{TOA}} = [n_{\text{TOA},2}, n_{\text{TOA},3}, \dots, n_{\text{TOA},L}]^\top \quad (2.13)$$

và

$$\mathbf{f}_{\text{TOA}}(\mathbf{x}) = \mathbf{d}_1 = \begin{bmatrix} \sqrt{(x-x_2)^2 + (y-y_2)^2} - \sqrt{(x-x_1)^2 + (y-y_1)^2} \\ \sqrt{(x-x_3)^2 + (y-y_3)^2} - \sqrt{(x-x_1)^2 + (y-y_1)^2} \\ \vdots \\ \sqrt{(x-x_L)^2 + (y-y_L)^2} - \sqrt{(x-x_1)^2 + (y-y_1)^2} \end{bmatrix}. \quad (2.14)$$

Bài toán định vị sử dụng phép đo TDOA là ước lượng $\mathbf{x}$ khi biết $\mathbf{r}_{\text{TDOA}}$.

2.2.3 Cường độ tín hiệu (RSS)

RSS là phép đo năng lượng của tín hiệu quan sát được trên nút mục tiêu, phụ thuộc vào khoảng cách tới nút tham chiếu và suy hao theo hàm mũ. Gọi P_t là công suất phát của L nút tham chiếu, RSS tại nút mục tiêu của nút tham chiếu thứ l , $l \in [L]$ được biểu diễn bởi [109]

$$P_{r,l} = K_l P_t d_l^{-\alpha} = K_l P_t \|\mathbf{x} - \mathbf{x}_l\|^{-\alpha}, \quad (2.15)$$

trong đó K_l thể hiện cho tất cả yếu tố ảnh lên công suất nhận được bao gồm độ lợi ăng-ten, chiều cao ăng-ten, góc lệch ăng-ten giữa nút tham chiếu và nút mục tiêu, v.v và α là hệ số suy hao. Tùy vào loại môi trường, hệ số α có thể là các giá trị nằm trong

khoảng từ 2 cho đến 5, với $\alpha = 2$ đại diện cho không gian tự do. Với K_l và α là hai thành phần phụ thuộc vào yếu tố môi trường, người ta đã đề xuất rất nhiều mô hình để giảm thiểu lỗi trong việc chuyển đổi từ RSS sang khoảng cách. Vấn đề này sẽ được trình bày ở các chương sau. Trong nội dung chương này, luận án giả sử P_t , K_l và α đều là các thông tin biết trước. Thông qua đo đạc, người ta thấy rằng RSS suy giảm theo hàm lô-ga tự nhiên

$$\ln(P_{r,l}) = \ln(K_l) + \ln(P_t) - \alpha \ln(d_l) + n_{\text{RSS},l}, \quad (2.16)$$

với $n_{\text{RSS},l}$ là nhiễu. Đặt $r_{\text{RSS},l} = \ln(P_{r,l}) - \ln(K_l) - \ln(P_t)$, mô hình đo đạc sử dụng RSS được đơn giản hóa thành

$$r_{\text{RSS},l} = -\alpha \ln(d_l) + n_{\text{RSS},l}. \quad (2.17)$$

Dạng véc-tơ của mô hình này là

$$\mathbf{r}_{\text{RSS}} = \mathbf{f}_{\text{RSS}}(\mathbf{x}) + \mathbf{n}_{\text{RSS}}, \quad (2.18)$$

trong đó

$$\mathbf{r}_{\text{RSS}} = [r_{\text{RSS},1}, r_{\text{RSS},2}, \dots, r_{\text{RSS},L}]^\top, \quad (2.19)$$

$$\mathbf{n}_{\text{RSS}} = [n_{\text{RSS},1}, n_{\text{RSS},2}, \dots, n_{\text{RSS},L}]^\top, \quad (2.20)$$

và

$$\mathbf{f}_{\text{RSS}}(\mathbf{x}) = \mathbf{p} = -\alpha \begin{bmatrix} \sqrt{(x-x_1)^2 + (y-y_1)^2} \\ \sqrt{(x-x_2)^2 + (y-y_2)^2} \\ \vdots \\ \sqrt{(x-x_L)^2 + (y-y_L)^2} \end{bmatrix}. \quad (2.21)$$

Bài toán định vị sử dụng phép đo RSS là đi ước lượng $\mathbf{x}$ khi biết $\mathbf{r}_{\text{RSS}}$.

2.2.4 Góc đến (DOA)

DOA là góc tới của tín hiệu quan sát tại nút mục tiêu. Từ 2 nút tham chiếu, chúng ta có thể vẽ ra 2 đường thẳng, mỗi đường thẳng đi qua 1 nút tham chiếu với góc tương ứng quan sát được trên nút mục tiêu. Vị trí giao nhau của 2 đường thẳng chính là vị trí ước lượng của nút mục tiêu. Mặc dù hệ thống sử dụng DOA không yêu cầu đồng

bộ về mặt thời gian nhưng các thiết bị trong hệ thống này phải trang bị ăng-ten thông minh để ước lượng góc đến của các nguồn tín hiệu. Gọi ϕ_l là góc đến từ nút tham chiếu thứ $l, l \in [L]$ với tọa độ $\mathbf{x}_l = [x_l, y_l]^\top$, ta có

$$\tan(\phi_l) = \frac{y - y_l}{x - x_l}. \quad (2.22)$$

Phép đo DOA có chứa nhiễu được biểu diễn bởi phương trình

$$r_{\text{DOA},l} = \phi_l + n_{\text{DOA},l} = \frac{y - y_l}{x - x_l} + n_{\text{DOA},l}, \quad (2.23)$$

trong đó $n_{\text{DOA},l}$ là nhiễu. Dạng véc-tơ của (2.23) được biểu diễn bởi [109]

$$\mathbf{r}_{\text{DOA}} = \mathbf{f}_{\text{DOA}}(\mathbf{x}) + \mathbf{n}_{\text{DOA}}, \quad (2.24)$$

trong đó

$$\mathbf{r}_{\text{DOA}} = [r_{\text{DOA},1}, r_{\text{DOA},2}, \dots, r_{\text{DOA},L}]^\top, \quad (2.25)$$

$$\mathbf{n}_{\text{DOA}} = [n_{\text{DOA},1}, n_{\text{DOA},2}, \dots, n_{\text{DOA},L}]^\top \quad (2.26)$$

và

$$\mathbf{f}_{\text{DOA}}(\mathbf{x}) = \boldsymbol{\phi} = \begin{bmatrix} \tan^{-1} \left(\frac{y-y_1}{x-x_1} \right) \\ \tan^{-1} \left(\frac{y-y_2}{x-x_2} \right) \\ \vdots \\ \tan^{-1} \left(\frac{y-y_L}{x-x_L} \right) \end{bmatrix}. \quad (2.27)$$

Bài toán định vị sử dụng phép đo DOA là đi ước lượng $\mathbf{x}$ khi biết $\mathbf{r}_{\text{DOA}}$.

2.3 Định vị dựa trên tham số

Phương pháp định vị có tham số được chia thành hai loại chính là tuyến tính và phi tuyến tính. Phương pháp ước lượng phi tuyến tính sử dụng trực tiếp (2.1) để giải $\mathbf{x}$ bằng cách tối thiểu hóa lỗi [109]

$$\mathbf{e}_{\text{non-linear}} = \mathbf{r} - \mathbf{f}(\tilde{\mathbf{x}}), \quad (2.28)$$

trong đó $\tilde{\mathbf{x}} = [\tilde{x}, \tilde{y}]^\top$ là lời giải tối ưu cho $\mathbf{x}$. Trong khi đó, phương pháp tuyến tính tìm cách biến đổi (2.1) thành một hệ thống tuyến tính

$$\mathbf{b} = \mathbf{A}\mathbf{x} + \mathbf{q}, \quad (2.29)$$

trong đó $\mathbf{b}$ và $\mathbf{A}$ đã biết và $\mathbf{q}$ là nhiễu. Tương tự, phương pháp tuyến tính cố gắng tối thiểu lỗi

$$\mathbf{e}_{\text{linear}} = \mathbf{b} - \mathbf{A}\tilde{\mathbf{x}}. \quad (2.30)$$

2.3.1 Thuật toán định vị phi tuyến

Hai phương pháp phổ biến của loại này là Bình phương tối thiểu phi tuyến và Tối đa hàm khả năng. Lấy phép đo TOA làm đại diện, cách thực thi của các phương pháp được trình bày sau đây có thể áp dụng cho TDOA, RSS hay DOA.

2.3.1.1 Bình phương tối thiểu phi tuyến

Với phép đo TOA: Dựa vào (2.4) và (2.5), hàm mất mát đối với phép đo TOA được định nghĩa bởi [109]

$$\begin{aligned} \mathcal{J}_{\text{NLS, TOA}}(\tilde{\mathbf{x}}) &= \sum_{l=1}^L \left(r_{\text{TOA},l} - \sqrt{(\tilde{x} - x_l)^2 + (\tilde{y} - y_l)^2} \right)^2 \\ &= (\mathbf{r}_{\text{TOA}} - \mathbf{f}_{\text{TOA}}(\tilde{\mathbf{x}}))^{\top} (\mathbf{r}_{\text{TOA}} - \mathbf{f}_{\text{TOA}}(\tilde{\mathbf{x}})). \end{aligned} \quad (2.31)$$

Nhiệm vụ là tìm $\tilde{\mathbf{x}}$ sao cho $\mathcal{J}_{\text{NLS, TOA}}(\tilde{\mathbf{x}})$ nhỏ nhất, hay

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \mathcal{J}_{\text{NLS, TOA}}(\tilde{\mathbf{x}}). \quad (2.32)$$

Với phép đo TDOA: Hàm mất mát được định nghĩa bởi [109]

$$\begin{aligned} \mathcal{J}_{\text{NLS, TDOA}}(\tilde{\mathbf{x}}) &= \sum_{l=2}^L \left(r_{\text{TDOA},l} - \sqrt{(\tilde{x} - x_l)^2 + (\tilde{y} - y_l)^2} + \sqrt{(\tilde{x} - x_1)^2 + (\tilde{y} - y_1)^2} \right)^2 \\ &= (\mathbf{r}_{\text{TOA}} - \mathbf{f}_{\text{TDOA}}(\tilde{\mathbf{x}}))^{\top} (\mathbf{r}_{\text{TDOA}} - \mathbf{f}_{\text{TDOA}}(\tilde{\mathbf{x}})). \end{aligned} \quad (2.33)$$

Ước lượng NLS của $\tilde{\mathbf{x}}$ xác định bởi

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \mathcal{J}_{\text{NLS, TDOA}}(\tilde{\mathbf{x}}). \quad (2.34)$$

Với phép đo RSS: Hàm mất mát được định nghĩa bởi [109]

$$\begin{aligned} \mathcal{J}_{\text{NLS, RSS}}(\tilde{\mathbf{x}}) &= \sum_{l=1}^L \left(r_{\text{RSS},l} + \alpha \ln \left(\sqrt{(\tilde{x} - x_l)^2 + (\tilde{y} - y_l)^2} \right) \right)^2 \\ &= (\mathbf{r}_{\text{RSS}} - \mathbf{f}_{\text{RSS}}(\tilde{\mathbf{x}}))^{\top} (\mathbf{r}_{\text{RSS}} - \mathbf{f}_{\text{RSS}}(\tilde{\mathbf{x}})). \end{aligned} \quad (2.35)$$

Ước lượng NLS của $\tilde{\mathbf{x}}$ xác định bởi

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \mathcal{J}_{\text{NLS, RSS}}(\tilde{\mathbf{x}}). \quad (2.36)$$

Với phép đo DOA: Hàm mất mát được định nghĩa bởi [109]

$$\begin{aligned} \mathcal{J}_{\text{NLS, DOA}}(\tilde{\mathbf{x}}) &= \sum_{l=1}^L \left(r_{\text{DOA},l} - \tan^{-1} \left(\frac{y - y_l}{x - x_l} \right) \right)^2 \\ &= (\mathbf{r}_{\text{DOA}} - \mathbf{f}_{\text{DOA}}(\tilde{\mathbf{x}}))^{\top} (\mathbf{r}_{\text{DOA}} - \mathbf{f}_{\text{DOA}}(\tilde{\mathbf{x}})). \end{aligned} \quad (2.37)$$

Ước lượng NLS của $\tilde{\mathbf{x}}$ xác định bởi

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \mathcal{J}_{\text{NLS, DOA}}(\tilde{\mathbf{x}}). \quad (2.38)$$

Tìm $\hat{\mathbf{x}}$ cho các bài toán trên không phải là một nhiệm vụ đơn giản vì hàm mất mát chứa nhiều điểm cực tiểu cục bộ. Lời giải cho bài toán tối ưu toàn cục có thể tìm được bằng các phương pháp tối ưu cục bộ như suy giảm đạo hàm (GD), phương pháp Newton, phương pháp Gauss-Newton, v.v. với một điểm khởi tạo gần cực tiểu toàn cục. Trong bài toán định vị, điểm khởi tạo thường được chọn thông qua các chiến lược tìm kiếm thô trên toàn bộ tập xác định, ví dụ như chia lưới hoặc tìm kiếm ngẫu nhiên. Quy trình lặp cho phương pháp GD bởi [109]

$$\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} - \alpha \nabla \left(\mathcal{J}_{\text{NLS, TOA}}(\hat{\mathbf{x}}^{(k)}) \right), \quad (2.39)$$

trong α là một số dương giúp điều khiển tốc độ hội tụ cho phương pháp GD. Với phương pháp Newton, quá trình tìm $\hat{\mathbf{x}}$ là [109]

$$\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} - \frac{\nabla \left(\mathcal{J}_{\text{NLS, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)}{\mathbf{H} \left(\mathcal{J}_{\text{NLS, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)}, \quad (2.40)$$

trong đó, $\mathbf{H} \left(\mathcal{J}_{\text{NLS, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)$ và $\nabla \left(\mathcal{J}_{\text{NLS, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)$ lần lượt là ma trận Hessian và véc-tơ gradient tại bước thứ k . Nói chung, tốc độ hội tụ tăng khi α tăng. Tuy nhiên, nếu chúng quá lớn, phương pháp sẽ trở nên không ổn định. Do đó, việc lựa chọn α phải rất khéo léo, sao cho thuật toán vừa hội tụ nhanh song vẫn đảm bảo sự ổn định. Bắt đầu với giá trị khởi tạo $\hat{\mathbf{x}}^{(0)}$, các thuật toán trong (2.40) và (2.39) sẽ phải dừng tại một giá trị thể hiện cho sự hội tụ. Thông thường, thuật toán sẽ dừng khi một trong hai tiêu chí sau được thỏa mãn: *i)* số vòng lặp đạt giá trị tối đa hoặc *ii)* sai số nhỏ hơn một ngưỡng nhất định, tức $\|\hat{\mathbf{x}}^{(k+1)} - \hat{\mathbf{x}}^{(k)}\|_2 \leq \epsilon$, với ϵ là sai số dừng.

2.3.1.2 Tối đa hàm khả năng

Phương pháp ML tối đa hóa hàm mật độ xác suất (PDF) của các phép đo TOA, TDOA, RSS và DOA để tìm ra vị trí máy thu với giả sử biết trước về phân phối lỗi. Xét bài toán ML cho phép đo TOA làm đại diện, giả sử $n_{\text{TOA},l}$ trong (2.4) là nhiễu Gauss với trung bình bằng 0, PDF cho $r_{\text{TOA},l}$ được biểu diễn bởi [109]

$$p(r_{\text{TOA},l}) = \frac{1}{\sqrt{2\pi\sigma_{\text{TOA},l}^2}} \exp\left(-\frac{1}{2\sigma_{\text{TOA},l}^2}(r_{\text{TOA},l} - d_l)^2\right), \quad (2.41)$$

trong đó d_l và $\sigma_{\text{TOA},l}^2$ lần lượt là trung bình và phương sai của phép đo TOA. Tương tự, PDF của $\mathbf{r}_{\text{TOA}}$ có thể viết dưới dạng

$$p(\mathbf{r}_{\text{TOA}}) = \frac{1}{(2\pi)^{\frac{L}{2}} \prod_{l=1}^L \sigma_{\text{TOA},l}} \exp\left(-\frac{1}{2}(\mathbf{r}_{\text{TOA}} - \mathbf{d})^\top \mathbf{C}_{\text{TOA}}^{-1}(\mathbf{r}_{\text{TOA}} - \mathbf{d})\right), \quad (2.42)$$

với $\mathbf{C}_{\text{TOA}}$ là ma trận hiệp phương sai của $\mathbf{r}_{\text{TOA}}$ được định nghĩa bởi

$$\begin{aligned} \mathbf{C}_{\text{TOA}} &= \mathbb{E}((\mathbf{r}_{\text{TOA}} - \mathbf{d})(\mathbf{r}_{\text{TOA}} - \mathbf{d})^\top) \\ &= \mathbb{E}(\mathbf{n}_{\text{TOA}}\mathbf{n}_{\text{TOA}}^\top) \\ &= \text{diag}(\sigma_{\text{TOA},1}^2, \sigma_{\text{TOA},2}^2, \dots, \sigma_{\text{TOA},L}^2). \end{aligned} \quad (2.43)$$

Lô-ga hóa (2.42), ta thu được

$$\ln(p(\mathbf{r}_{\text{TOA}})) = \ln\left(\frac{1}{(2\pi)^{\frac{L}{2}} |\mathbf{C}_{\text{TOA}}|^{\frac{1}{2}}}\right) - \frac{1}{2}(\mathbf{r}_{\text{TOA}} - \mathbf{d})^\top \mathbf{C}_{\text{TOA}}^{-1}(\mathbf{r}_{\text{TOA}} - \mathbf{d}). \quad (2.44)$$

Vì thành phần đầu tiên của (2.44) không phụ thuộc vào $\mathbf{x}$, việc tối đa hóa (2.44) trở thành tối ưu cho cụm biểu thức thứ hai. Hàm mất mát ML của phép đo TOA được định nghĩa bởi

$$\begin{aligned} \mathcal{J}_{\text{ML, TOA}}(\tilde{\mathbf{x}}) &= (\mathbf{r}_{\text{TOA}} - \mathbf{f}_{\text{TOA}}(\tilde{\mathbf{x}}))^\top \mathbf{C}_{\text{TOA}}^{-1}(\mathbf{r}_{\text{TOA}} - \mathbf{f}_{\text{TOA}}(\tilde{\mathbf{x}})) \\ &= \sum_{l=1}^L \frac{\left(r_{\text{TOA},l} - \sqrt{(\tilde{x} - x_l)^2 + (\tilde{y} - y_l)^2}\right)^2}{\sigma_{\text{TOA},l}^2}. \end{aligned} \quad (2.45)$$

Hàm mất mát cho các phép đo TDOA, RSS hay DOA được xây dựng tương tự bằng cách nhân thêm vào hàm mất mát của phương pháp NLS một ma trận hiệp phương sai chứa thống kê về nhiễu của các phép đo tương ứng. Nhiệm vụ là phải đi tìm

$$\hat{\mathbf{x}} = \arg \min_{\tilde{\mathbf{x}}} \mathcal{J}_{\text{ML, TOA}}(\tilde{\mathbf{x}}). \quad (2.46)$$

Từ (2.45), chúng ta có thể thấy MLE là một trường hợp tổng quát của NLS. Khi $\mathbf{C}_{\text{TOA}} = \mathbf{I}_L$ hay $\sigma_{\text{TOA},l}^2, l \in [L]$ của các phép đo là giống nhau, phương pháp ML sẽ trở thành phương pháp NLS. Tuy nhiên, nếu nhiều của các phép đo là khác nhau, $\sigma_{\text{TOA},l}^2$ lớn, tương đương với nhiễu trong $r_{\text{TOA},l}$ lớn, một trọng số $\frac{1}{\sigma_{\text{TOA},l}^2}$ nhỏ sẽ được áp dụng cho đại lượng trên tử số và ngược lại. Áp dụng phương pháp GD cho (2.46), ta có

$$\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} - \gamma \nabla \left(\mathcal{J}_{\text{ML, TOA}}(\hat{\mathbf{x}}^{(k)}) \right). \quad (2.47)$$

Áp dụng phương pháp Newton cho (2.46), ta có

$$\hat{\mathbf{x}}^{(k+1)} = \hat{\mathbf{x}}^{(k)} - \frac{\nabla \left(\mathcal{J}_{\text{ML, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)}{\mathbf{H} \left(\mathcal{J}_{\text{ML, TOA}}(\hat{\mathbf{x}}^{(k)}) \right)}. \quad (2.48)$$

2.3.2 Thuật toán định vị tuyến tính

2.3.2.1 Bình phương tối thiểu tuyến tính (LLS)

Bình phương hai vế của (2.4), ta có

$$r_{\text{TOA},l}^2 = (x - x_l)^2 + (y - y_l)^2 + n_{\text{TOA},l}^2 + 2n_{\text{TOA},l}\sqrt{(x - x_l)^2 + (y - y_l)^2}, \quad (2.49)$$

Thay thế $R = x^2 + y^2$ và $m_{\text{TOA},l} = n_{\text{TOA},l}^2 + 2n_{\text{TOA},l}\sqrt{(x - x_l)^2 + (y - y_l)^2}$ vào (2.49), ta thu được

$$-2x_l x - 2y_l y + R + m_{\text{TOA},l} = r_{\text{TOA},l}^2 - x_l^2 - y_l^2. \quad (2.50)$$

Đặt

$$\mathbf{A} = \begin{bmatrix} -2x_1 & -2y_1 & 1 \\ -2x_2 & -2y_2 & 1 \\ \vdots & \vdots & \vdots \\ -2x_L & -2y_L & 1 \end{bmatrix}, \quad (2.51)$$

$$\boldsymbol{\theta} = [x, y, R]^\top, \quad (2.52)$$

$$\mathbf{b} = \begin{bmatrix} r_{\text{TOA},1}^2 - x_1^2 - y_1^2 \\ r_{\text{TOA},2}^2 - x_2^2 - y_2^2 \\ \vdots \\ r_{\text{TOA},L}^2 - x_L^2 - y_L^2 \end{bmatrix} \quad (2.53)$$

và

$$\mathbf{q} = [m_{\text{TOA},1}, m_{\text{TOA},2}, \dots, m_{\text{TOA},L}]^\top. \quad (2.54)$$

Dạng ma trận của (2.50) là

$$\mathbf{A}\boldsymbol{\theta} + \mathbf{q} = \mathbf{b}. \quad (2.55)$$

Khi đó, $\mathbf{r}_{\text{TOA}}$ ở (2.5) được biến đổi thành $\mathbf{b}$. $\mathbf{A}$ được xây dựng từ vị trí đã biết của các nút tham chiếu và $\boldsymbol{\theta}$ chứa vị trí cần xác định của nút mục tiêu. Khi $m_{\text{TOA},l}$ đủ nhỏ sao cho $\mathbb{E}(\mathbf{q}) \approx 0$, ta có

$$\mathbf{q} \approx \begin{bmatrix} 2n_{\text{TOA},1}\sqrt{(x-x_1)^2 + (y-y_1)^2} \\ 2n_{\text{TOA},2}\sqrt{(x-x_2)^2 + (y-y_2)^2} \\ \vdots \\ 2n_{\text{TOA},L}\sqrt{(x-x_L)^2 + (y-y_L)^2} \end{bmatrix}. \quad (2.56)$$

Ta có xấp xỉ

$$\mathbf{A}\boldsymbol{\theta} \approx \mathbf{b}. \quad (2.57)$$

Hàm mất mát của (2.57) có thể được định nghĩa bởi

$$\begin{aligned} \mathcal{J}_{\text{LLS, TOA}}(\tilde{\boldsymbol{\theta}}) &= (\mathbf{A}\tilde{\boldsymbol{\theta}} - \mathbf{b})^\top (\mathbf{A}\tilde{\boldsymbol{\theta}} - \mathbf{b}) \\ &= \tilde{\boldsymbol{\theta}}^\top \mathbf{A}^\top \mathbf{A} \tilde{\boldsymbol{\theta}} - 2\tilde{\boldsymbol{\theta}}^\top \mathbf{A}^\top \mathbf{b} + \mathbf{b}^\top \mathbf{b}. \end{aligned} \quad (2.58)$$

Có thể thấy, $\mathcal{J}_{\text{LLS, TOA}}(\tilde{\boldsymbol{\theta}})$ là một hàm bậc hai của $\mathcal{J}_{\text{LLS, TOA}}(\tilde{\boldsymbol{\theta}})$ và nó có nghiệm duy nhất. Ước lượng LLS của bài toán này là

$$\hat{\boldsymbol{\theta}} = \arg \min_{\tilde{\boldsymbol{\theta}}} \mathcal{J}_{\text{LLS, TOA}}(\tilde{\boldsymbol{\theta}}) \quad (2.59)$$

Lời giải dạng đóng của bài toán này là

$$\boldsymbol{\theta} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}. \quad (2.60)$$

Cuối cùng, vị trí của nút mục tiêu được trích xuất từ phần tử thứ nhất và thứ hai của véc-tơ $\boldsymbol{\theta}$. Bên cạnh đó, chúng ta có thể thu được một bài toán LLS tương đương bằng cách lấy các phương trình từ thứ hai trở đi trừ cho phương trình thứ nhất. Khi đó, đại

lượng R trong (2.50) sẽ bị loại bỏ và chúng ta thu được một hệ thống tuyến tính với

$$\mathbf{A} = -2 \begin{bmatrix} x_2 - x_1 & y_2 - y_1 \\ x_3 - x_1 & y_3 - y_1 \\ \vdots & \vdots \\ x_L - x_1 & y_L - y_1 \end{bmatrix}, \quad (2.61)$$

$$\mathbf{q} = [m_{\text{TOA},2} - m_{\text{TOA},1}, m_{\text{TOA},3} - m_{\text{TOA},1}, \dots, m_{\text{TOA},L} - m_{\text{TOA},1}]^\top, \quad (2.62)$$

và

$$\mathbf{b} = \begin{bmatrix} r_{\text{TOA},2}^2 - x_2^2 - y_2^2 - (r_{\text{TOA},1}^2 - x_1^2 - y_1^2) \\ r_{\text{TOA},3}^2 - x_3^2 - y_3^2 - (r_{\text{TOA},1}^2 - x_1^2 - y_1^2) \\ \vdots \\ r_{\text{TOA},L}^2 - x_L^2 - y_L^2 - (r_{\text{TOA},1}^2 - x_1^2 - y_1^2) \end{bmatrix}. \quad (2.63)$$

Hệ thống được biểu diễn bởi

$$\mathbf{A}\mathbf{x} + \mathbf{q} = \mathbf{b}. \quad (2.64)$$

Lời giải cho bài toán này là:

$$\hat{\mathbf{x}} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}, \quad (2.65)$$

trong đó $\hat{\mathbf{x}} = [\hat{x}, \hat{y}]^\top$ là vị trí của nút mục tiêu. Các biến đổi tương tự có thể áp dụng cho các phép đo TDOA, RSS hoặc DOA.

2.3.2.2 Thuật toán bình phương tối thiểu tuyến tính có trọng số (WLLS)

Mặc dù phương pháp LLS đơn giản nhưng nó chỉ cho kết quả chính xác khi nhiều giữa các phép đo là độc lập và có phân phối giống hệt nhau (i.i.d). Tuy nhiên, nhiều trong thực tế không hoàn toàn là i.i.d. Do đó, ước lượng LLS có thể là chưa hoàn toàn tối ưu. Độ chính xác định vị có thể được cải thiện bằng cách nhân thêm vào hàm mất mát của LLS một ma trận trọng số, ký hiệu là $\mathbf{W}$. Ta có

$$\begin{aligned} \mathcal{J}_{\text{WLLS, TOA}}(\tilde{\boldsymbol{\theta}}) &= (\mathbf{A}\tilde{\boldsymbol{\theta}} - \mathbf{b})^\top \mathbf{W} (\mathbf{A}\tilde{\boldsymbol{\theta}} - \mathbf{b}) \\ &= \tilde{\boldsymbol{\theta}}^\top \mathbf{A}^\top \mathbf{W} \mathbf{A} \tilde{\boldsymbol{\theta}} - 2\tilde{\boldsymbol{\theta}}^\top \mathbf{A}^\top \mathbf{W} \mathbf{b} + \mathbf{b}^\top \mathbf{W} \mathbf{b}. \end{aligned} \quad (2.66)$$

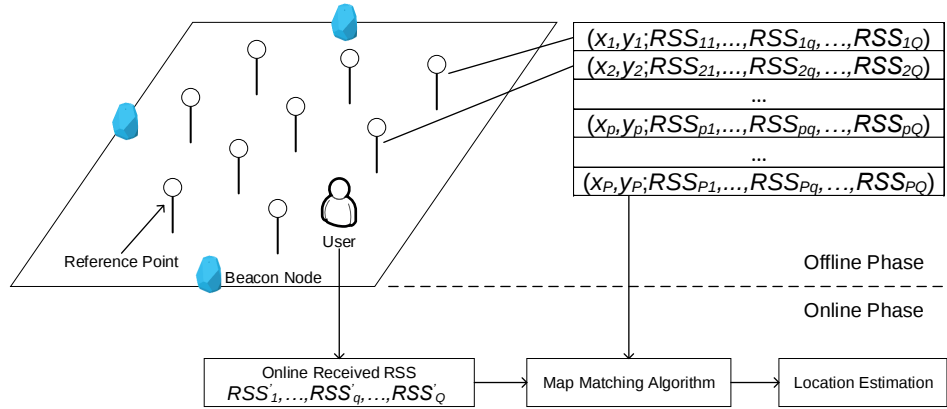
Từ (2.55) và (2.56), có thể thấy $\mathbb{E}(\mathbf{b}) = \mathbf{A}\boldsymbol{\theta}$ có liên quan tới mô hình dữ liệu không thiên lệch (unbias). Do đó, chúng ta có thể xác định $\mathbf{W}$ bằng cách nghịch đảo hiệp

phương sai của $\mathbf{q}$. Ma trận $\mathbf{W}$ tối ưu bằng

$$\begin{aligned}\mathbf{W} &= [\mathbb{E}(\mathbf{q}\mathbf{q}^\top)]^{-1} \\ &= [\text{diag}(\sigma_{\text{TOA},1}^2 d_1^2, \sigma_{\text{TOA},2}^2 d_2^2, \dots, \sigma_{\text{TOA},L}^2 d_L^2)]^{-1} \\ &= \text{diag}\left(\frac{1}{\sigma_{\text{TOA},1}^2 d_1^2}, \frac{1}{\sigma_{\text{TOA},2}^2 d_2^2}, \dots, \frac{1}{\sigma_{\text{TOA},L}^2 d_L^2}\right).\end{aligned}\quad (2.67)$$

2.4 Định vị không dựa trên tham số

Định vị không tham số còn được biết đến với tên gọi là định vị sử dụng bản đồ vô tuyến hay lấy dấu vân tay (FP). So với những phương pháp định vị được tham số hóa như định vị sử dụng khoảng cách, FP có một đặc điểm vượt trội là không phải chuyển đổi giá trị các phép đo (TOA, TDOA hoặc RSS) thành khoảng cách - một việc được cho là gây thêm lỗi khi xác định vị trí nút mục tiêu. Độ ổn định của FP cũng cao hơn các phương pháp dựa trên khoảng cách hay góc vì vị trí được cố định thành những điểm có nhãn và vị trí xác định. Ý tưởng chính của FP là sử dụng dữ liệu lớn để mô tả các vị trí vật lý bằng RSS hoặc CSI trong khu vực quan tâm. Hình 2.2 mô tả cách phương pháp FP hoạt động. Vị trí ước tính được từ FP phải thông qua hai pha: pha



Hình 2.2: Phương pháp lấy dấu vân tay [60]

Ngoại tuyến và pha Trực tuyến. Cụ thể, trong pha Ngoại tuyến, FP chia khu vực quan tâm thành các khu vực con hoặc các điểm tham chiếu. Bước tiếp theo sẽ thực hiện thu thập dữ liệu về RSS hoặc CSI và đánh nhãn cho tất cả các khu vực con hoặc điểm

điểm tham chiếu đó. Dữ liệu sau đó được lưu vào cơ sở dữ liệu dưới dạng các véc-tơ có nhãn vị trí. Toàn bộ cơ sở dữ liệu được gọi là một bản đồ vô tuyến. Giả sử một khu vực quan tâm được triển khai với M nút tham chiếu (còn gọi là đèn hiệu hoặc beacon) và có thể chia thành L khu vực con với tọa độ vật lý là (x_l, y_l) . Với N véc-tơ dữ liệu thu thập cho mỗi khu vực con, cơ sở dữ liệu thu thập được có dạng

$$\{(x_l, y_l) | RSS_{1n}^{(l)}, \dots, RSS_{mn}^{(l)}, \dots, RSS_{Mn}^{(l)}\}, \quad m \in [M], n \in [N], l \in [L]. \quad (2.68)$$

Trong pha Trực tuyến, nút mục tiêu sẽ quan sát dữ liệu mới và sử dụng các thuật toán khớp để so sánh dữ liệu mới này với toàn bộ cơ sở dữ liệu được thu thập trước và đưa ra dự đoán về vị trí. Các véc-tơ dữ liệu mới có dạng

$$\{RSS'_{1n}, \dots, RSS'_{mn}, \dots, RSS'_{Mn}\}, \quad n \in [N]. \quad (2.69)$$

Các thuật toán khớp thường được sử dụng bao gồm: k-lân-cận (KNN), k-lân-cận có trọng số (WKNN), suy luận Bayes, máy véc-tơ hỗ trợ (SVM), mạng nơ-ron nhân tạo (ANN), mạng nơ-ron hồi quy (RNN), v.v.

2.4.1 K lân cận (KNN)

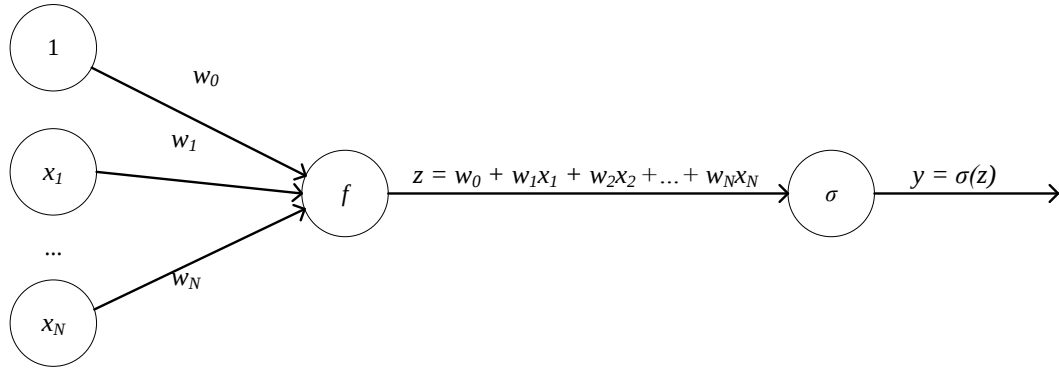
k NN là một trong những thuật toán học máy có giám sát đơn giản nhất, chủ yếu được sử dụng để phân loại dữ liệu. k NN lưu trữ tất cả các dữ liệu với nhãn được thu thập trước và phân loại dữ liệu mới dựa trên một số loại khoảng cách như Euclidean, Cosine, Chi-Square hay Minkowsky. Số k trong k NN là một tham số dùng để phân loại các nhãn dựa trên quá trình bỏ phiếu theo đa số (major voting) hoặc theo trọng số khoảng cách. Tùy thuộc vào loại dữ liệu, người ta sử dụng các loại khoảng cách khác nhau để đo khoảng cách giữa hai điểm $\mathbf{a}$ và $\mathbf{b}$ trong không gian các đặc trưng. Khoảng cách Euclidean là loại được sử dụng rộng rãi nhất. Cho 2 véc-tơ $\mathbf{a} = (x_1, x_2, \dots, x_m)$ và $\mathbf{b} = (y_1, y_2, \dots, y_m)$, khoảng cách Euclidean giữa $\mathbf{a}$ và $\mathbf{b}$ được tính bởi

$$d_{\text{Euclid}}(\mathbf{a}, \mathbf{b}) = \|\mathbf{a} - \mathbf{b}\|_2^2. \quad (2.70)$$

Trong bài toán định vị sử dụng FP, phương pháp kNN tính toán khoảng cách từ véc-tơ RSS mới quan sát được (chưa có nhãn) với tất cả các véc-tơ RSS đã có nhãn trong cơ sở dữ liệu. Sau đó, k khoảng cách gần nhất được sử dụng để xác định nhãn của véc-tơ mới quan sát được.

2.4.2 Mạng nơ-ron nhân tạo (ANN)

Mạng nơ-ron nhân tạo (ANN) là một mô hình toán học được xây dựng dựa trên các mạng nơ-ron sinh học. Giữa lối vào và lối ra chứa các lớp nơ-ron giúp trao đổi thông tin. Một cấu trúc cơ bản của ANN được biểu diễn trong Hình 2.3. Trong đó, lớp đầu



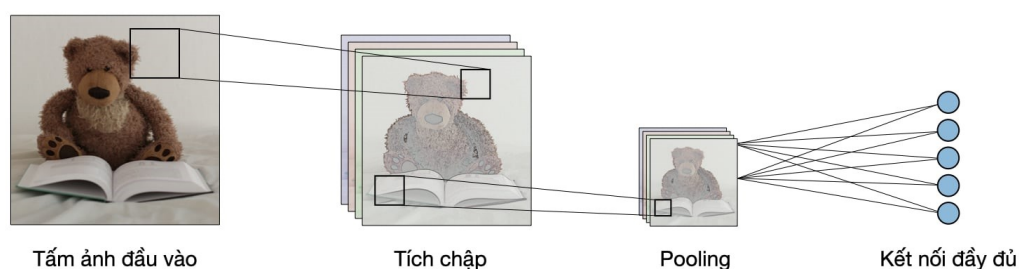
Hình 2.3: Cấu trúc đơn giản của mạng nơ-ron nhân tạo

tiên là các nút chứa dữ liệu đầu vào và tham số độ lệch (bias). Trong bài toán định vị, dữ liệu chính là các giá trị RSS của các khu vực con. Kết nối giữa lớp đầu vào lớp thứ hai thể hiện trọng số ứng với mỗi nút. Lớp thứ hai chứa giá trị tổng hợp từ các nút thuộc lớp thứ nhất và trọng số của nó. Lớp thứ ba là một hàm kích hoạt giúp tính xác suất và phân loại dữ liệu. Mạng ANN làm việc theo hai giai đoạn là học (truyền xuôi) và học củng cố (truyền ngược). Giai đoạn học được thực hiện theo chiều xuôi (theo Hình 2.3 là chu trình từ trái qua phải), giúp tạo ra các kiến thức cho các lớp sau thông qua các trọng số và độ lệch. Sau khi quá trình học diễn ra, người ta sẽ thu được một kết quả ở lối ra. Kết quả này sẽ cho biết được mức độ hiệu quả mà mỗi nút và trọng số đã tạo ra. Sau đó, bước học củng cố sẽ được thực hiện theo chiều ngược lại (theo Hình 2.3 là chu trình từ phải qua trái) theo cách giảm trọng số cho các nút có hiệu quả ước lượng kém (ứng với tỷ lệ lỗi cao) và tăng trọng số cho các nút có hiệu quả ước lượng tốt (ứng với tỷ lệ lỗi thấp). Quá trình học và học củng cố sẽ diễn ra liên tục cho đến khi hàm mất mát đại diện cho toàn mạng đạt cực trị. Quá trình tối ưu hàm mất mát thường liên quan tới các thuật toán suy giảm đạo hàm (gradient descent). Sau khi quá trình học và học củng cố kết thúc, một siêu mặt phẳng (hyper-plane) sẽ được

hình thành và chúng ta có thể sử dụng nó để phân loại các dữ liệu mới.

2.4.3 Mạng nơ-ron tích chập (CNN)

Mạng nơ-ron tích chập (CNN) nhân chập các bộ lọc (kernel) với ma trận đầu vào để tìm ra các đặc trưng của loại dữ liệu nhất định. Một trong số đó là dùng để phân loại. Một mạng CNN đơn giản được minh họa trong Hình 2.4.



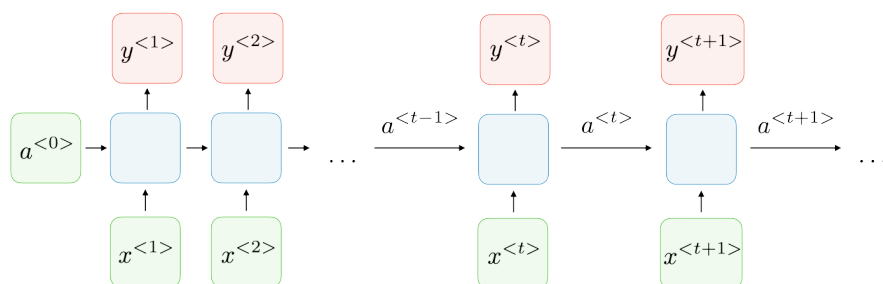
Hình 2.4: Mạng tích chập đơn giản (Nguồn: CS-230)

Mạng tích chập bao gồm: lớp tích chập, lớp tổng hợp (pooling), và mạng nơ-ron nhân tạo cơ bản (ANN). Trong đó, lớp tích chập thực hiện việc trượt một bộ lọc trên ma trận đầu vào. Các tham số trong bước này bao gồm kích thước bộ lọc và tốc độ trượt của bộ lọc (stride). Kết quả đầu ra là một ma trận đặc trưng. Nói cách khác, bộ lọc giúp làm nổi bật các tính chất quan trọng của dữ liệu đầu vào. Lớp tổng hợp thường được sử dụng sau lớp tích chập giúp giảm kích thước ma trận đặc trưng, giảm số lượng tham số cho các lớp phía sau, qua đó tăng tính bất biến, tính ổn định đối với các biến đổi nhỏ của dữ liệu đầu vào. Cuối cùng, mạng ANN với hàm kích hoạt đi kèm giúp tính xác suất và phân loại dữ liệu.

2.4.4 Mạng nơ-ron hồi quy (RNN)

Mạng nơ-ron hồi quy (RNN) là một dạng của mạng nơ-ron trong đó đầu ra từ bước trước được dùng làm đầu vào cho bước hiện tại. Mạng này phù hợp với dữ liệu có tính tuần tự, ví dụ dữ liệu theo thời gian. Cấu trúc cơ bản của một mạng Nơ-ron hồi quy được minh họa bởi Hình 2.5. Cấu trúc RNN khá giống với ANN nhưng có thêm vòng phản hồi qua đó mạng thể ghi nhớ được trạng thái trước đó. Giống như ANN

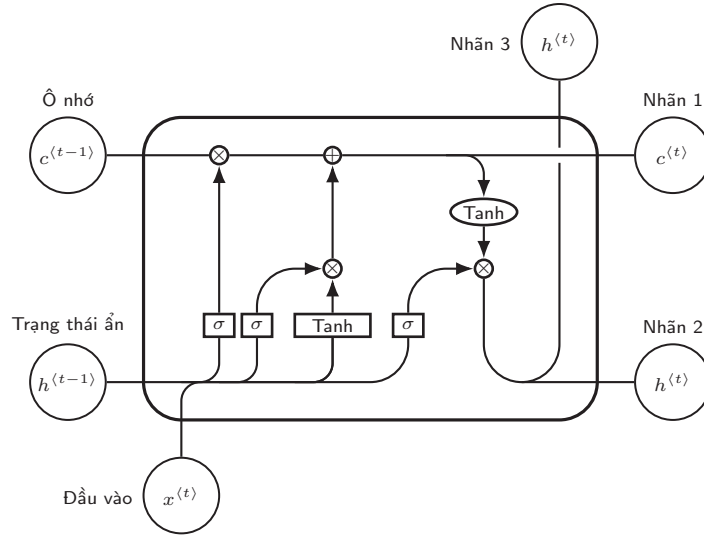
hay CNN, để tối ưu ma trận trọng số của RNN, người ta cần tính được hàm mất mát và sử dụng các phương pháp tối ưu như GD hoặc phương pháp Newton để tìm điểm cực trị của nó. Vấn đề đối với RNN chính là biến mất đạo hàm. Hiện tượng này xảy ra khi đạo hàm của hàm mất mát quá nhỏ hoặc quá lớn trong quá trình học củng cố (truyền ngược). Hiện tượng này không chỉ xảy ra riêng với mạng RNN mà còn ở tất cả các mạng học máy khác. Tuy nhiên, vì RNN có các kết nối cho phép chúng lưu trữ thông tin từ các bước trước đó nên hiện tượng này xảy ra phổ biến hơn. Với bài toán định vị dựa trên RSS, RNN cho phép khai thác mối liên hệ tuần tự của các mẫu theo thời gian đồng thời giảm thiểu sự ảnh hưởng bởi sự không ổn định của RSSI trong thời gian ngắn [68].



Hình 2.5: Cấu trúc cơ bản của một mạng Nơ-ron hồi quy (Nguồn: CS-230)

2.4.5 Bộ nhớ Ngắn hạn Dài (LSTM)

Bộ nhớ Ngắn hạn Dài (LSTM) được thiết kế để giải quyết vấn đề biến mất đạo hàm của RNN. Cấu trúc một nơ-ron của mạng LSTM được minh họa bởi Hình 2.6. Một nơ-ron của mạng LSTM bao gồm: cổng đầu vào hiện tại, cổng quên và cổng đầu ra. Cổng đầu vào quyết định xem liệu thông tin mới có được lưu vào trong bộ nhớ hay không. Nó có đầu vào là dữ liệu hiện tại và trạng thái trước đó. Cổng đầu vào có 2 lớp. Lớp thứ nhất sử dụng hàm **sigmoid** để quyết định xem thông tin nào sẽ được cập nhật. Đầu ra của lớp này là một giá trị nằm trong khoảng 0 đến 1, trong đó 0 ứng với các phần thông tin không quan trọng, sẽ bị loại bỏ và 1 ứng với thông tin quan trọng, sẽ được giữ lại. Tiếp theo, dữ liệu sẽ được đưa qua lớp thứ hai chứa hàm **tanh()**. Hàm **tanh()** ở đây giúp khắc phục hiện tượng biến mất đạo hàm giúp đạo hàm của nó được duy trì được trong khoảng một thời gian dài trước khi tiến về 0. Cổng quên sẽ quyết



Hình 2.6: Bên trong một nơ-ron của mạng LSTM

định thông tin nào của bộ nhớ dài hạn sẽ được giữ lại. Để làm được điều này, giống như cổng đầu vào, người ta đưa đầu vào qua lớp ẩn chứa hàm **sigmoid**. Giá trị sau khi qua hàm **sigmoid** sẽ quyết định bao nhiêu phần trăm của bộ nhớ sẽ được giữ lại. Cổng đầu ra sẽ lấy giá trị đầu vào hiện tại, trạng thái nhớ của lớp ẩn trước đó để tạo ra lớp ẩn mới, được chuyển tiếp sang lớp tiếp theo.

2.5 Kết luận chương

Chương này đã trình bày *i)* các công nghệ không dây bao gồm WiFi, BLE, UWB, mmWave *ii)* các phép đo không dây bao gồm: TOA, TDOA, DOA, RSS và *iii)* các phương pháp định vị không dây cơ bản bao gồm: định vị có tham số và không tham số. Đây là cơ sở cho những đề xuất sẽ trình bày trong hai chương tiếp theo.

Chương 3

Định vị với phép đo RSS trong hệ thống định vị trong nhà sử dụng Bluetooth

Chương 3 xem xét hệ thống định vị trong nhà sử dụng RSS từ đèn hiệu BLE. Luận án bắt đầu bằng việc trình bày các tính chất của phép đo RSS trong thực tế. Từ đó, luận án đề xuất hai phương pháp: *i)* Phương pháp triển khai đèn hiệu sao cho diện tích vùng phục vụ của các đèn hiệu là lớn nhất song vẫn cung cấp độ chính xác định vị cao nhất cho hệ thống định vị kết hợp hệ cảm biến quán tính và phương pháp định vị sử dụng khoảng cách và *ii)* Phương pháp lựa chọn đèn hiệu cho thuật toán định vị LS có trọng số. Trên cơ sở các thuật toán được xây dựng để giảm thiểu chi phí triển khai của hệ thống, luận án đánh giá sai số định vị dựa trên thí nghiệm trong môi trường thực tế. Kết quả cho thấy các phương pháp đều cung cấp một giải pháp định vị có độ chính xác cao hơn so với những hệ thống đơn lẻ, không có sự hỗ trợ của các phương pháp đề xuất.

*Nội dung của Chương 3 liên quan chủ yếu đến tới các công trình nghiên cứu **CT1**, **CT2** và **CT8**.*

3.1 Phép đo RSS

3.1.1 Cường độ tín hiệu (RSS)

RSS là một trong những phép đo có tính ứng dụng nhất cho hệ thống định vị [15]. Theo cách đầy đủ nhất, RSS được mô hình hóa theo mô hình thời gian - không gian $\rho(t, \mathbf{s})$ thông qua công thức [59]

$$\rho(t, \mathbf{s}) = \mathcal{F}(P_{\mathbf{rx}}(t, \mathbf{s})), \quad (3.1)$$

trong đó $\mathbf{s} \subset \mathbb{R}^2$ là một khu vực trong không gian 2 chiều, $P_{\mathbf{rx}}(t, \mathbf{s})$ là cường độ tín hiệu quan sát được trên máy thu và $\mathcal{F}(\cdot)$ là một hàm chuyển đổi $P_{\mathbf{rx}}(t, \mathbf{s})$ sang RSS. Để đơn giản, người ta coi $\mathcal{F}(P_{\mathbf{rx}}(t, \mathbf{s})) = P_{\mathbf{rx}}(t, \mathbf{s})$ với

$$P_{\mathbf{rx}}(t, \mathbf{s}) = \Gamma(d) + \Upsilon(\mathbf{s}) + a(t). \quad (3.2)$$

Trong (3.2):

- $P_{\mathbf{rx}}(t, \mathbf{s})$ được đo theo dBm (log scale). Vì thế, $P_{\mathbf{rx}}(t, \mathbf{s})$ có thể viết thành tổng các thành phần, bao gồm $\Gamma(d)$ - suy hao không tính đến yếu tố không gian, $\Upsilon(\mathbf{s})$ - che khuất (shadowing) và $a(t)$ - pha-đỉnh đa đường. Nếu chỉ xem xét $P_{\mathbf{rx}}(t, \mathbf{s})$ ở một tần số cố định, thành phần liên quan tới tần số có thể loại bỏ.
- $\Gamma(d)$ là một hàm phụ thuộc vào khoảng cách và được biết đến với tên gọi *mô hình suy hao tuyến tính theo từng đoạn (multi-slope piece-wise linear propagation model)* [40].

$$\Gamma(d) = \mathcal{A} - 10 \sum_{h=1}^m \eta_h \log \left(\frac{\langle d \rangle_{\delta_{h-1}}^{\delta_h}}{\delta_{h-1}} \right), \quad (3.3)$$

trong đó $\mathcal{A}$ là cường độ tín hiệu nhận được ở khoảng cách tham chiếu δ_0 . η_h biểu diễn hệ số suy hao của đoạn thứ m trong $\langle d \rangle_a^b = \min \{ \max \{ d, a \}, b \}$, và $\delta_0 < \delta_1 < \delta_2 < \dots < \delta_m = \infty$. $\mathcal{A}$, η_h và δ thường có được thông qua quá trình học trên dữ liệu đo đạc thực tế. Đối với môi trường ngoài trời, $\delta_0 = 1$ km và $m = 1$ vì khoảng cách giữa máy thu và máy phát là tương đối xa và được coi là không có vật cản. Trong khi đó, môi trường trong nhà thường có $\delta_0 = 1$ m và $m \geq 2$ vì

khoảng cách giữa máy thu và máy phát nhỏ và môi trường trong nhà được coi là biến đổi nhanh và có nhiều vật cản.

- $\Upsilon(\mathbf{s})$ mô hình hóa những thay đổi của RSS trong không gian và được biết đến với tên gọi *pha-đỉnh chậm*. Nó thường phụ thuộc vào vị trí và gần như bất biến theo thời gian. Trong thực tế, người ta có thể quan sát được loại pha-đỉnh này bằng cách khảo sát một khoảng cách ở các môi trường khác nhau. Khi đó, mỗi loại môi trường sẽ cho một giá trị $\Gamma(d_0)$ khác nhau nhưng hầu như không thay đổi theo thời gian.
- $a(t)$ mô tả sự biến đổi của RSS theo thời gian và được biết đến với tên gọi *pha-đỉnh nhanh* [7]. Nó được coi là một quá trình dừng (stationary) và có tính ergodic.

Khi có thống kê đầy đủ, công suất tín hiệu trung bình được tính bởi

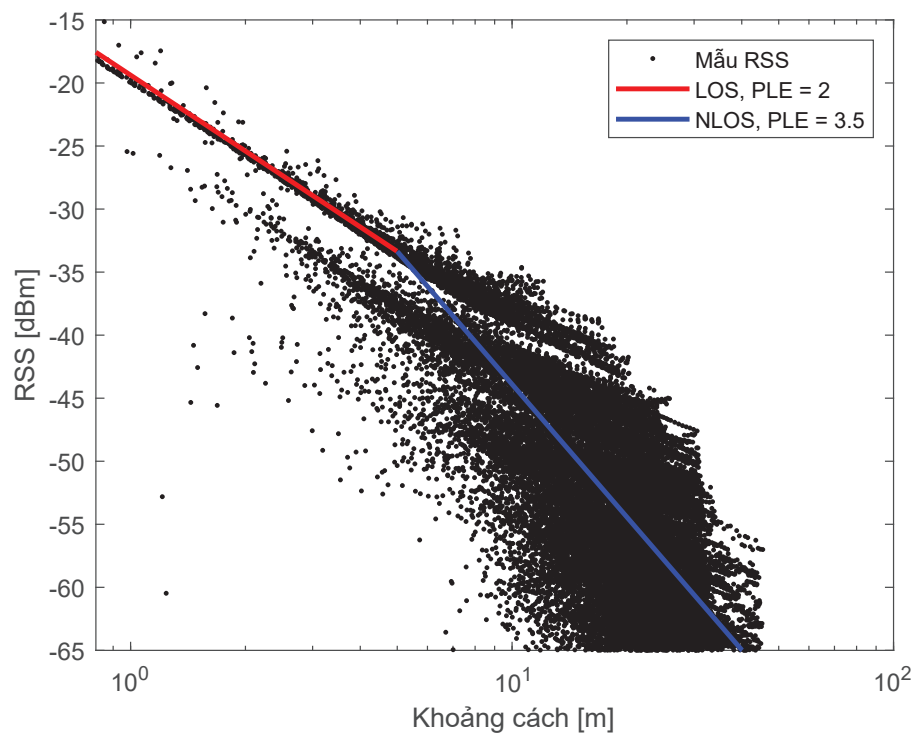
$$\bar{P}_{\text{rx}}(t, \mathbf{s}) = \lim_{T \rightarrow \infty} \frac{1}{T} \int_0^T P_{\text{rx}}(t, \mathbf{s}) dt = \Gamma(d) + \Upsilon(\mathbf{s}). \quad (3.4)$$

Điều này ngụ ý rằng thành phần pha-đỉnh nhanh có thể được loại bỏ thông qua các bộ lọc như bộ lọc trung bình hoặc bộ lọc Kalman. Trong điều kiện không có sự hiện diện của con người, người ta có thể thu được một kết quả đo như trong Hình 3.1. Trong Hình 3.1, các mẫu RSS có xu hướng tạo thành hình sao chổi với phần đầu ứng với khoảng cách gần và phần đuôi ứng với khoảng cách xa hơn. Mô hình ứng với hiện tượng này được gọi là một mô hình không đồng nhất và được biểu diễn bởi [2]

$$\begin{cases} \mathcal{R}_i = \Gamma_i \pm \sigma_{\mathcal{R},i} \\ \sigma_{\mathcal{R},i} = f(|\Gamma_i|) \end{cases}, \quad (3.5)$$

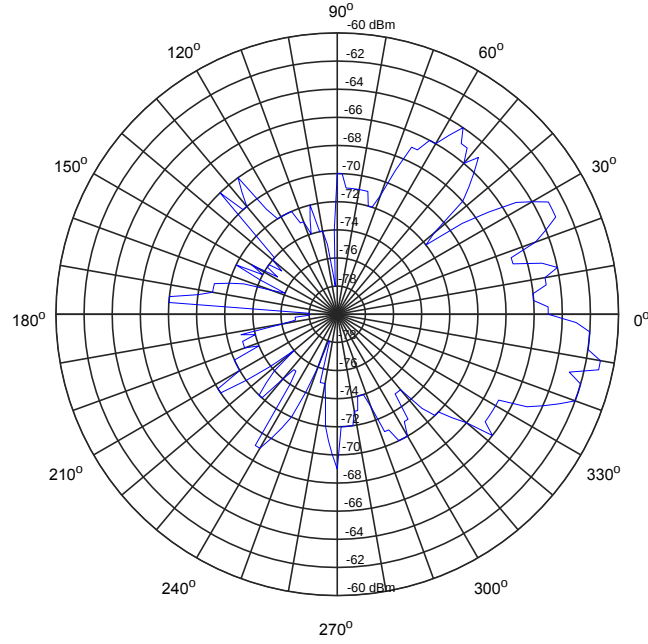
trong đó $\mathcal{R} \sim \mathcal{N}(\Gamma, \sigma_{\mathcal{R}}^2)$ biểu diễn cho giá trị RSS tức thời. Γ và $\sigma_{\mathcal{R}}$ lần lượt là giá trị RSS trung bình và độ lệch chuẩn của nó. Trong khi đó, $f(\cdot)$ là một hàm phụ thuộc vào Γ . Trong điều kiện thực tế, giá trị RSS thậm chí còn biến đổi phức tạp hơn do nhiều yếu tố khác gây ra, đặc biệt là sự hấp thụ của cơ thể con người. Cơ thể con người chứa 70% là nước; do đó, nó gây suy hao mạnh đối với tín hiệu ở tần số 2,4 GHz. Sự suy giảm tín hiệu do cơ thể con người gây ra, được ước tính thông qua [72]

$$PL_H = 30 + 10 \log(2wn), \quad (3.6)$$



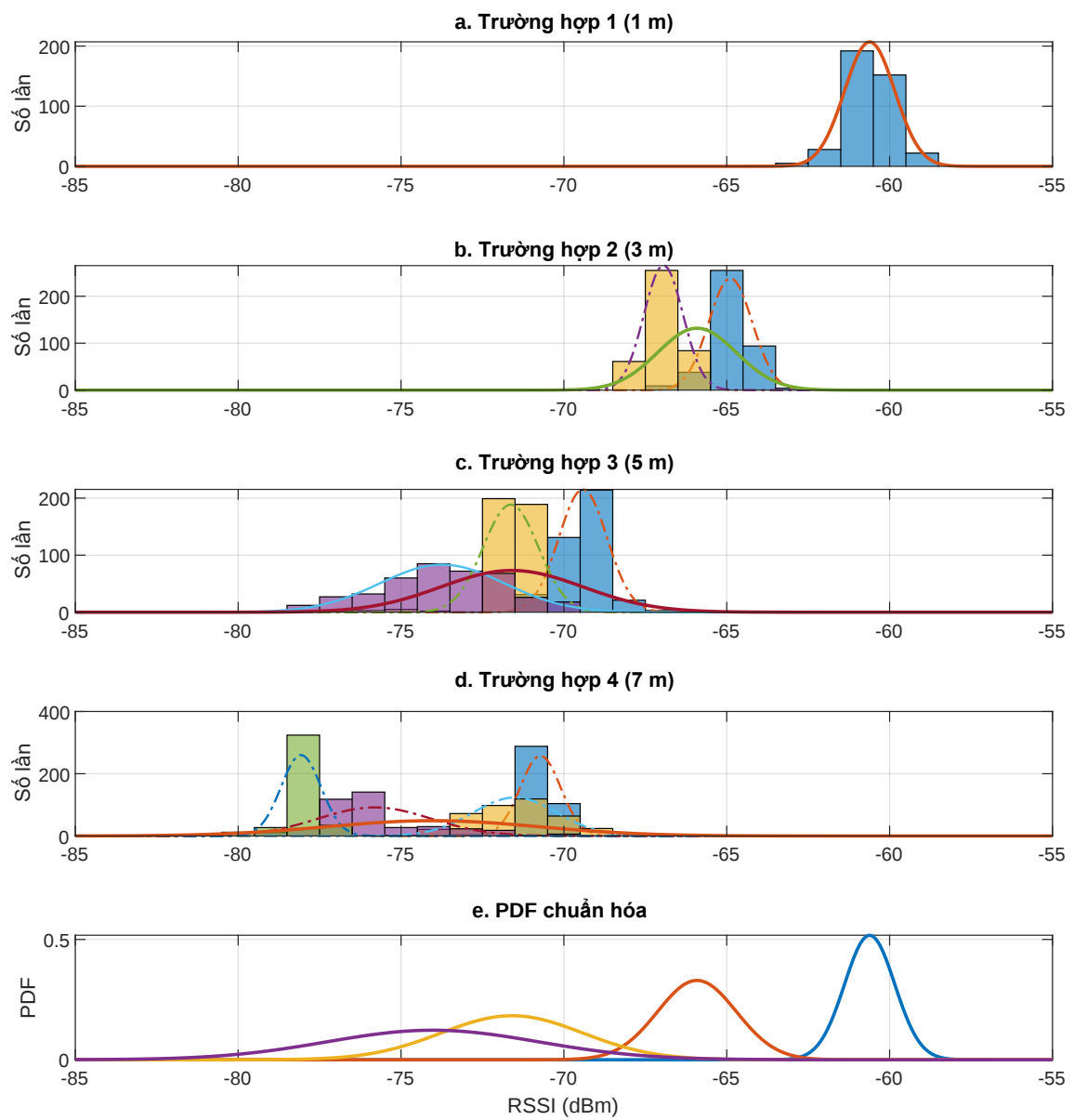
Hình 3.1: Kết quả thực nghiệm nêu trong [11]. Các đường màu xanh và đỏ thể hiện cho $\Gamma(d)$ với hệ số mất mát lần lượt là 2 (LOS) và 3.5 (NLOS).

trong đó w và n là trọng lượng cơ thể và số người hiện diện trong khu vực đo đạc. Các tác giả trong [60] đã thực hiện một thí nghiệm để thấy rõ được sự ảnh hưởng này. Thí nghiệm được thực hiện bằng cách cho người dùng cầm điện thoại thông minh (máy thu) ở trước ngực để đo cường độ tín hiệu theo các hướng khác nhau từ một đèn hiệu BLE (máy phát) được triển khai ở độ cao 1.5 m. Kết quả thí nghiệm được cho trong Hình 3.2.



Hình 3.2: Cường độ tín hiệu suy giảm theo các hướng khác nhau [60]. Góc 0° và 180° lần lượt ứng với 2 trường hợp máy thu và máy phát không bị che chắn và bị che chắn bởi cơ thể con người ở khoảng cách 2 m.

Xem xét sự không chắc chắn của phép đo RSS, bắt đầu với giá trị RSS tại vị trí tham chiếu $\mathcal{A} = -60$ dBm, nếu máy thu quan sát được giá trị này, nó có thể kết luận rằng môi trường giữa nó và máy phát là LOS. Tuy nhiên, nếu máy thu nhận được một giá trị RSS nhỏ hơn, ví dụ -90 dBm, nó không thể kết luận được loại môi trường nào đang hiện hữu giữa nó và máy phát. Nó có thể là môi trường LOS với khoảng cách xa hoặc môi trường NLOS với khoảng cách gần. Vì máy thu không biết nó đang phải đối mặt với loại môi trường nào, nó phải chấp nhận rằng một giá trị RSS có thể đến từ tất cả các tình huống. Kết quả là độ không chắc chắn của một giá trị RSS là cao. Luận án thực hiện khảo sát để tính độ không chắc chắn của RSS ở mọi tình huống có thể xảy ra



Hình 3.3: Phân bố của các giá trị RSS ở các khoảng cách khác nhau.

giữa máy thu và máy phát trong phạm vi quan tâm. Trong mỗi thí nghiệm, 400 giá trị RSS được ghi lại cho các khoảng cách khác nhau (1, 3, 5, 7 m). Dữ liệu sau đó được sử dụng để vẽ Hình 3.3 sử dụng MATLAB. Trong Hình 3.3, biểu đồ thanh được vẽ bằng cách sử dụng hàm `histogram()`. Trong đó, các cột màu xanh lam, vàng, tím và xanh lục lần lượt biểu diễn các điều kiện khác nhau giữa máy thu và máy phát, bao gồm: LOS, bị chặn bởi một bức tường, bị chặn bởi một cột lớn, và bị chặn bởi hai bức tường. Đường nét đứt biểu diễn cho phân phối chuẩn của các dữ liệu tương ứng. Các đường nét liền biểu diễn phân phối sau khi kết hợp các phân phối của dữ liệu ở cùng một khoảng cách nhưng khác loại điều kiện. Luận án sử dụng hàm `fitdist(x, 'Normal')` để tìm giá trị trung bình (μ - μ) và độ lệch chuẩn (σ - σ) trong mỗi trường hợp. Các đường nét đứt và liền được vẽ thông qua μ and σ

$$y = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right). \quad (3.7)$$

Kết quả của μ và σ được tổng hợp trong Bảng 3.1. Trong Bảng. 3.1, giá trị trung

Khoảng cách (m)	Giá trị RSS trung bình (Γ hay μ (dBm))	Độ lệch chuẩn (σ)
1	-60.6125	0.7706
3	-65.9138	1.2096
5	-71.5942	2.181
7	-74.0263	3.2523

Bảng 3.1: RSS trung bình và độ không chắc chắn của nó ở những khoảng cách khác nhau.

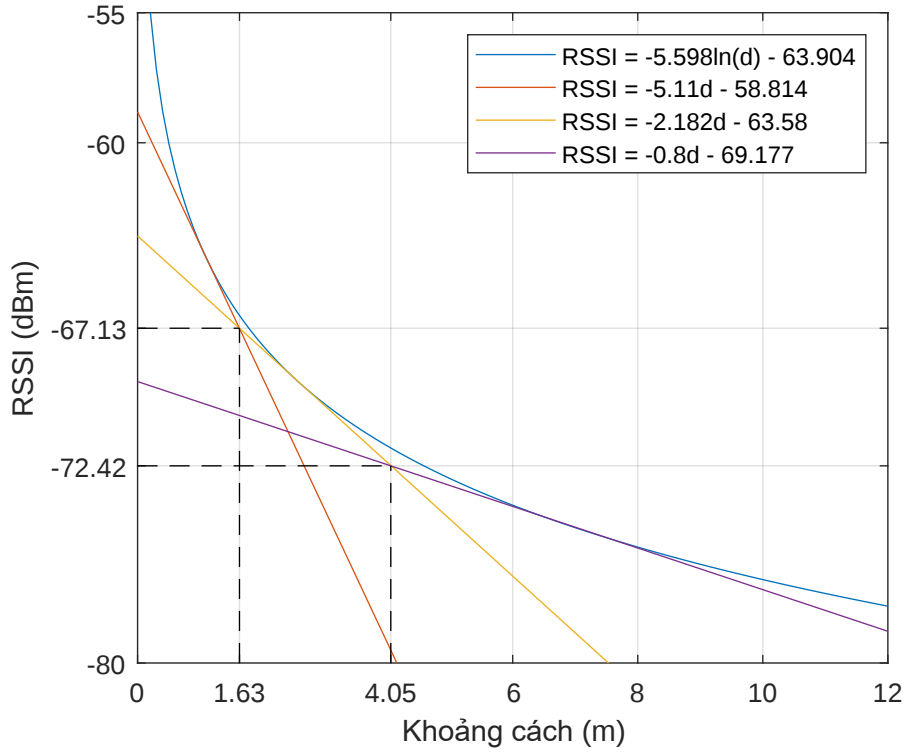
binh (μ -dBm) và độ lệch chuẩn (σ) lần lượt biểu diễn $\Gamma(d)$ và biến thể của $\Upsilon(\mathbf{s})$ trong phương trình 3.2. Sự thay đổi độ lệch chuẩn cho thấy RSS có thể bị ảnh hưởng bởi môi trường giữa máy thu và máy phát (position-dependent). Ở khoảng cách 1 m, một giá trị RSS hầu như chỉ quan sát được từ một đường dẫn (đường dẫn LOS). Kết quả là độ lệch chuẩn nhỏ. Ngược lại, một giá trị RSS có thể quan sát được từ rất nhiều tình huống ở khoảng cách 7 m. RSS trung bình chỉ ra rằng thành phần $a(t)$ có thể được loại bỏ nhờ phương trình (3.4). Sử dụng kết quả có trong Bảng 3.1 và phép hồi quy

tuyến tính, chúng ta thấy rằng có thể sử dụng phương trình

$$\sigma_i = 0.1752|\Gamma_i| - 10.067 \quad (3.8)$$

để tính toán độ tin cậy của một giá trị RSS.

3.1.2 Chuyển đổi từ RSS sang khoảng cách.



Hình 3.4: Xấp xỉ tuyến tính cho mô hình chuyển đổi từ RSS sang khoảng cách.

Môi trường trong nhà được coi là không đồng nhất và là duy nhất. Vì vậy, cách tốt nhất để có một mô hình chuyển đổi tốt từ RSS sang khoảng cách là học trực tiếp từ dữ liệu đo đạc ở chính môi trường đó. Dựa vào dữ liệu đo đạc được và hồi quy lô-ga, luận án tìm được mô hình chuyển đổi từ RSS sang khoảng cách đại diện cho cả môi trường LOS và NLOS là $\Gamma = -5.598 \ln(d) - 63.904$, được thể hiện bằng đường màu xanh lam trong Hình 3.4. Vì RSS dao động mạnh nên biến đổi trên khoảng cách đầu ra thông qua hàm lô-ga càng trở lên khó kiểm soát. Để hạn chế điều này, luận án sử

dùng một xấp xỉ tuyến tính cho từng khoảng RSS khác nhau có dạng

$$\Gamma(d) = \sum_{n=1}^N \alpha_n \langle d \rangle_{\delta_{n-1}}^{\delta_n} + \beta_n, \quad (3.9)$$

trong đó α_n và β_n là hệ số thực của đoạn thứ n và $0 = \delta_0 < \delta_1 < \dots < \delta_N = \infty$. Để tìm một đường xấp xỉ hàm phi tuyến, luận án sử dụng phương pháp Newton [44]. Gọi $f : \mathbb{R} \rightarrow \mathbb{R}$ là một hàm khả vi, chúng ta đi tìm kiếm lời giải $f(d) = 0$, bắt đầu từ khoảng cách tham chiếu d_0 . Tại vòng lặp thứ n , chúng ta thu được d_n và tiếp tục tính toán xấp xỉ cho d_{n+1} thông qua

$$d_{n+1} = d_n - \frac{f(d_n)}{f'(d_n)}. \quad (3.10)$$

Việc này được lặp đi lặp lại cho đến khi n đạt tới giá trị tối đa. Tiếp tuyến tìm được ở các khoảng cách $\{d_0, d_1, \dots, d_N\}$ tạo ra một đường xấp xỉ cho hàm phi tuyến. Với dữ liệu thu thập được, mô hình chuyển đổi từ RSS sang khoảng cách được mô tả bằng phương trình tuyến tính ba đoạn

$$\Gamma(d) = \begin{cases} -5.11d - 58.814, & 0 \leq d < 1.63 \\ -2.182d - 63.58, & 1.63 \leq d < 4.05 \\ -0.8d - 69.177, & 4.05 \leq d < 10 \end{cases} \quad (3.11)$$

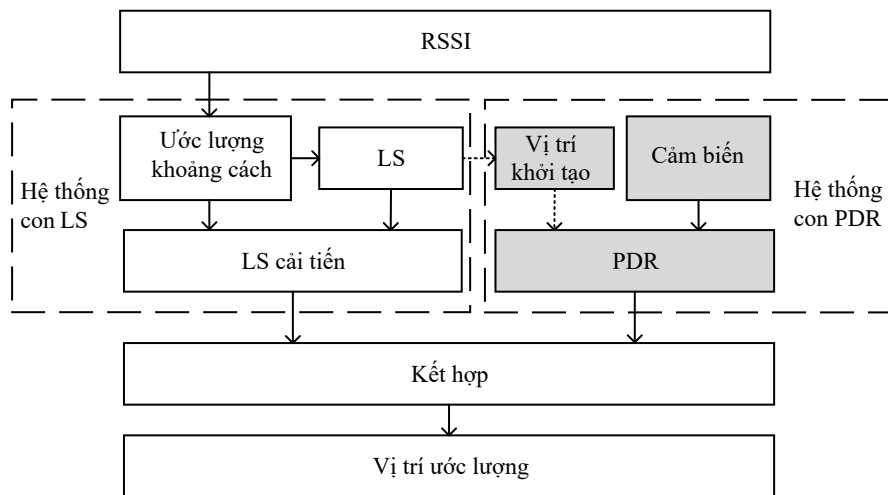
Từ kết quả này, một chiến lược thích ứng sẽ được áp dụng cho RSS đầu vào. Các mô hình này được cho bởi Bảng 3.2.

Bảng 3.2: Mô hình tính toán khoảng cách cho mỗi khoảng RSSI.

Khoảng RSS (dBm)	Mô hình chuyển đổi từ RSS sang khoảng cách
Từ -67.13 đến -60	$\Gamma = -5.11d - 58.814$
Từ -72.42 đến -67.13	$\Gamma = -2.182d - 63.58$
Từ -80 đến -72.42	$\Gamma = -0.8d - 69.177$

3.2 Định vị trong nhà với cách triển khai đèn hiệu BLE tam giác đều

3.2.1 Hệ thống đề xuất



Hình 3.5: Kiến trúc hệ thống

Kiến trúc hệ thống tổng thể được mô tả bởi Hình 3.5. Hệ thống đề xuất bao gồm hai hệ thống con hoạt động song song là hệ thống con PDR và hệ thống con LS. Theo thứ tự từ trên xuống của Hình 3.5, điện thoại thông minh quét tìm tín hiệu BLE có sẵn ở gần, sau đó ước tính khoảng cách để mỗi tín hiệu đó tạo ra vị trí LS. Đồng thời, LS giúp tìm ra vị trí ban đầu cho PDR trong giai đoạn đầu. Trong khi đó, hệ thống con PDR chịu trách nhiệm lấy dữ liệu từ các cảm biến để tính toán sự di chuyển của người dùng. Đầu ra của hệ thống con này là các vị trí PDR liên tục. Cuối cùng, vị trí dựa trên LS và PDR được bộ lọc Kalman hợp nhất để tạo ra các vị trí chính xác hơn.

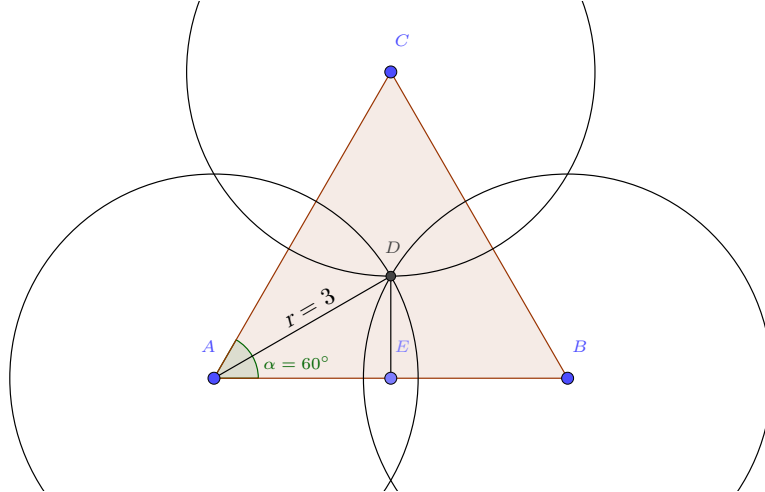
3.2.2 Chiến lược đặt đèn hiệu đề xuất

Một chiến lược đặt đèn hiệu BLE tối ưu đã được trình bày trong [25]. Trong đó, nhóm tác giả đề xuất nên triển khai đèn hiệu theo mô hình tam giác đều. Chiến lược này đảm bảo việc triển khai với số lượng đèn hiệu ít nhất có thể. Tuy nhiên, nó áp dụng

cho phạm vi phát sóng tối đa của đèn hiệu (50 m), và có thể chỉ phù hợp với FP hoặc các kỹ thuật khác. Vì vậy, luận án đề xuất một chiến lược khác cho hệ thống này. Vì LS sử dụng tối thiểu cho 3 đèn hiệu BLE nên trước tiên luận án xem xét đơn vị nhỏ nhất bao gồm ba thiết bị. Gọi Φ_A, Φ_B, Φ_C là các vùng bao phủ đáng tin cậy có bán kính đáng tin cậy là $r = 3$ [6] và có tâm lần lượt là A, B, C . Chiến lược đề xuất yêu cầu hai điều kiện dưới đây phải được thỏa mãn

$$\begin{cases} \max\{\Phi_A \cup \Phi_B \cup \Phi_C\} \\ \Phi_A \cap \Phi_B \cap \Phi_C = D \end{cases}, \quad (3.12)$$

trong đó $D \in \mathbb{R}^2$ là một điểm trong không gian 2 chiều. Điều kiện đầu tiên giảm thiểu số lượng đèn hiệu BLE cần được triển khai. Điều kiện thứ hai là điều kiện phụ của điều kiện đầu tiên, nó đảm bảo vị trí PDR được điều chỉnh với xác suất cao nhất. Để thỏa mãn 2 điều kiện trên, ba điểm A, B và C tạo thành một tam giác đều (xem Phụ lục I). Chiến lược tối ưu được cho trong Hình 3.6. Từ mối quan hệ hình học thể hiện trong Hình 3.6, khoảng cách tối ưu giữa 2 đèn hiệu BLE liền kề bằng $2AE = 2r \cos(30) = 3\sqrt{3}$ m. Lưu ý rằng kết quả của này gần giống với [25], [109] nhưng với mô hình hệ thống khác, cách tiếp cận khác và khoảng cách tối thiểu giữa 2 đèn hiệu BLE cũng khác. Do



Hình 3.6: Chiến lược triển khai đèn hiệu BLE tam giác đều

thực tế là các đèn hiệu BLE được đặt cách đều nhau theo hình dạng đường thẳng, LS sẽ không hoạt động tốt trong môi trường dài và hẹp (ví dụ như hành lang hẹp). Trong trường hợp này, thay vào đó, luận án sử dụng một đèn hiệu BLE duy nhất làm điểm

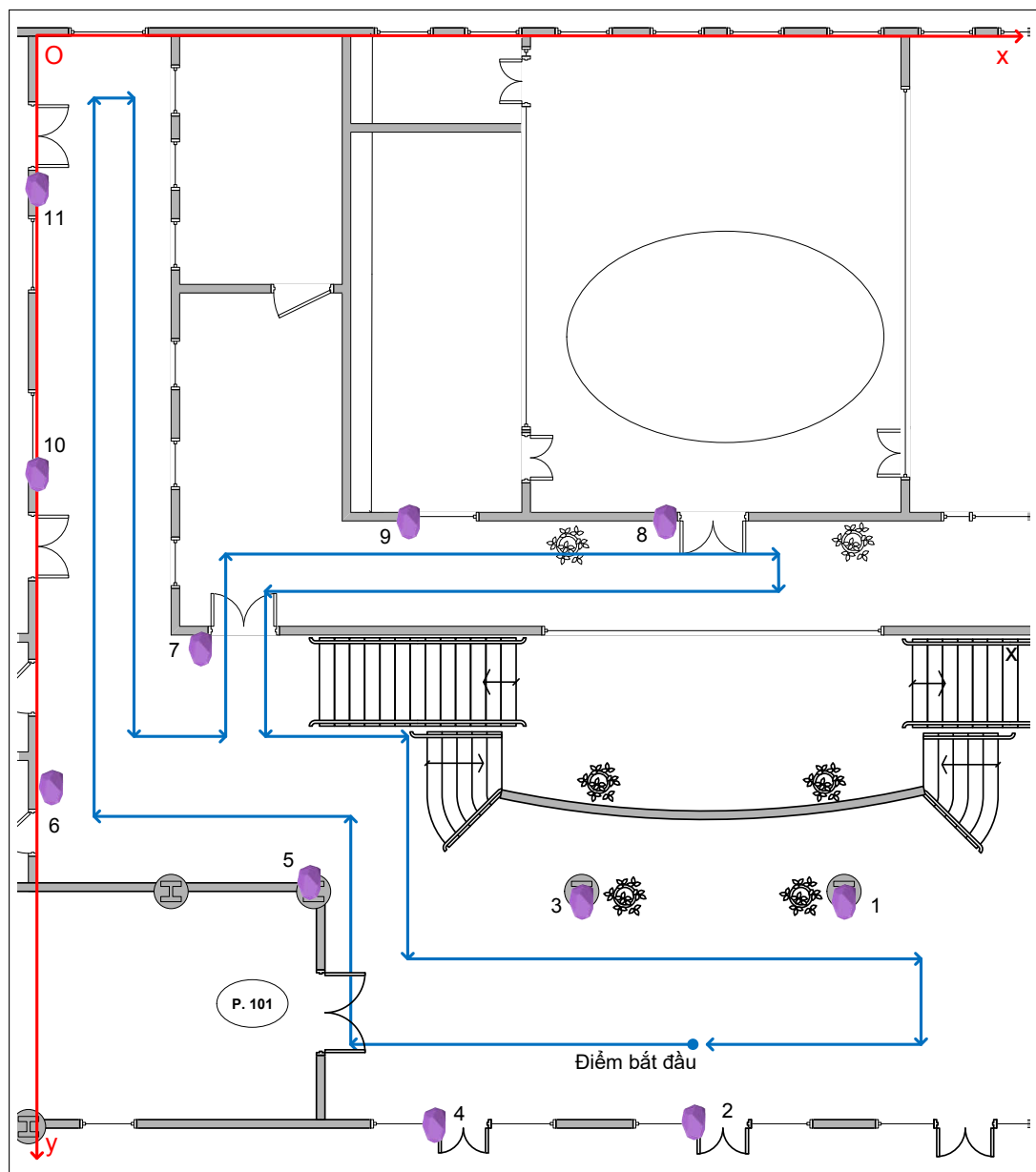
tham chiếu. Vị trí ước lượng của PDR sau đó được di chuyển đến vòng tròn của đèn hiệu BLE gần nhất.

3.2.3 Kết quả và thảo luận

Khu vực thí nghiệm là môi trường trong nhà điển hình với diện tích xấp xỉ 350 m^2 . 11 đèn hiệu BLE được sử dụng để đánh giá độ hiệu quả của hệ thống. Vị trí đèn hiệu BLE tuân theo chiến lược được đề xuất trước đó với khoảng cách tối đa giữa các đèn hiệu BLE liền kề là khoảng 6 đến 8 m. Người dùng di chuyển theo một đường dẫn được vẽ sẵn để đánh giá độ chính xác của hệ thống trong các tình huống khác nhau. Để bắt đầu, thiết bị người dùng sử dụng LS với 3 đèn hiệu BLE gần nhất để ước tính vị trí nó trong mỗi thí nghiệm. Bản đồ, vị trí đèn hiệu BLE và đường đi chính xác được cho trong Hình 3.7. Trong hình này, đường màu xanh lam và các nút màu tím lần lượt biểu thị đường tham chiếu và vị trí của đèn hiệu. Đèn hiệu số 1 - số 7 được đặt trong không gian mở trong khi đèn hiệu số 8 - số 11 được đặt trong không gian đóng.

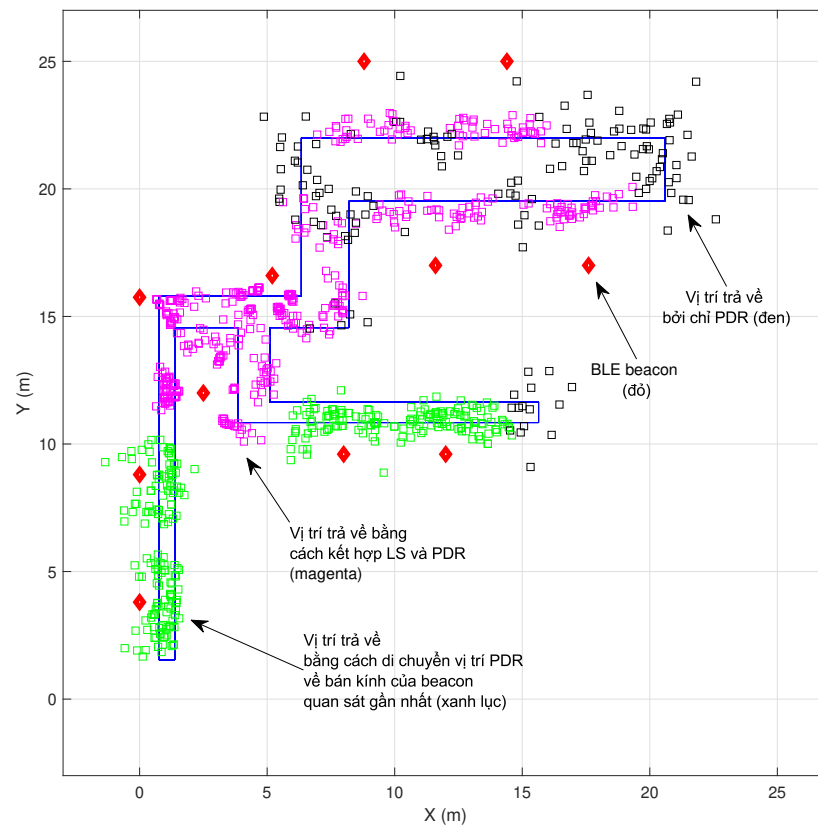
3.2.3.1 Đánh giá độ chính xác của hệ thống

Trong phần này, luận án thực hiện mô phỏng để đánh giá phương pháp triển khai đèn hiệu BLE đề xuất. Mô phỏng này sử dụng các tham số đo thực tế, bao gồm mô hình chuyển đổi từ RSS sang khoảng cách trong phương trình (3.11), mô hình tính toán sự không chắc chắn cho các giá trị RSS trong phương trình (3.8) và thống kê lỗi thực tế của PDR. Vị trí do PDR trả về sẽ được sửa bởi phương pháp LS mỗi khi thiết bị người dùng đọc được một giá trị RSS của bất kỳ đèn hiệu BLE nào lớn hơn -70 dBm . Quá trình sửa bắt đầu bằng việc áp dụng LS cho 3 đèn hiệu BLE có RSS lớn nhất, sau đó di chuyển vị trí ước lượng được bởi LS về vòng tròn có bán kính nhỏ nhất trong số 3 đèn hiệu BLE quan sát được ở phía trên. Vị trí do LS và PDR trả về sau đó được kết hợp lại với nhau thông qua bộ lọc Kalman. Độ lệch so với giá trị thực của PDR và phương sai của nó lần lượt được đặt bằng 1,2 m và 1.44 m. Trong khi đó, phương sai của vị trí do LS trả về được tính thông qua sự lan truyền của mô hình (3.8). Các kết quả được cho bởi Hình 3.8 và Hình 3.9. Như có thể thấy, sự kết hợp giữa PDR và LS xảy ra khi người dùng di chuyển vào khu vực tam giác của 3 đèn hiệu BLE trong



Hình 3.7: Vị trí đặt đèn hiệu BLE và đường tham chiếu trong khu vực thí nghiệm.

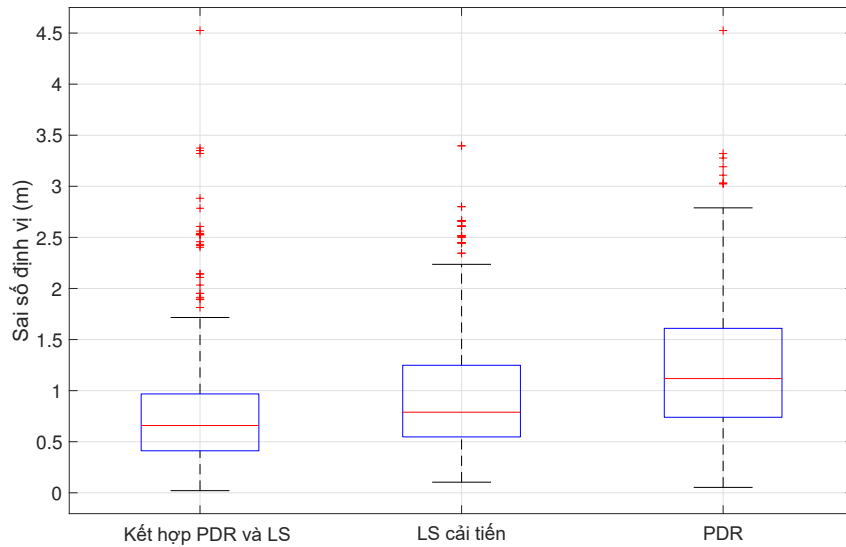
khu vực mở. Cũng trong khu vực này, sự kết hợp sẽ không xảy ra khi mục tiêu không quan sát được đèn hiệu BLE cho RSS lớn hơn -70 dBm. Với khu vực đóng, vị trí có xu hướng tạo thành một vòng tròn xung quanh đèn hiệu BLE gần nhất. Sai số trung bình của PDR, LS đề xuất và vị trí tổng hợp lần lượt là $1,23$ m, $0,98$ m và $0,8$ m. Sự kết hợp của LS và PDR được thực hiện 350 lần trên toàn bộ quãng đường di chuyển, bao gồm 704 bước và giúp cải thiện độ chính xác thêm $18,37\%$.



Hình 3.8: Phân phối của vị trí ước lượng bởi các kỹ thuật khác nhau trên bản đồ.

3.2.3.2 Hiệu năng của hệ thống khi thay đổi số đèn hiệu.

Để đánh giá ảnh hưởng của số lượng đèn hiệu đến độ chính xác của hệ thống, các đèn hiệu dần được loại bỏ khỏi bản đồ với điều kiện luôn tồn tại cụm ba đèn hiệu trong không gian mở. Hiệu năng của của phương pháp đề xuất được biểu diễn trong Hình 3.11 và Hình 3.10. Từ xu hướng trong Hình 3.11, chúng ta có thể thấy rằng độ chính xác định vị tăng lên khi số lượng đèn hiệu tăng lên. Hệ thống có độ chính xác cao nhất

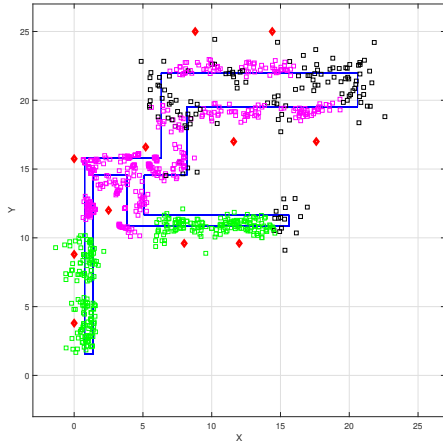


Hình 3.9: Sai số của các phương pháp thông qua biểu đồ hộp.

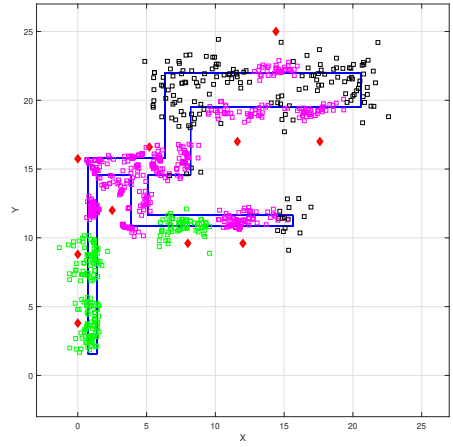
là 0.75 m với 11 đèn hiệu và thấp nhất là 1,1 m với 7 đèn hiệu BLE. Điều này xảy ra vì càng có nhiều đèn hiệu thì khả năng xuất hiện các điểm khắc phục cũng tăng theo. Một kết quả đáng chú ý là sự thay đổi về độ chính xác giữa các trường hợp là không đáng kể khi số lượng đèn hiệu tăng lên. Điều này cho thấy rằng chúng ta có thể có một hệ thống rẽ tiền với độ chính xác gần như tương đương. Ví dụ, trong trường hợp 8 đèn hiệu, số lượng đèn hiệu BLE phục vụ từng khu vực là vừa đủ, tức là 6 đèn hiệu cho không gian mở và 2 đèn hiệu BLE cho hai hành lang.

3.3 Xem xét sự hỗ trợ của học máy trong vấn đề định vị trong nhà

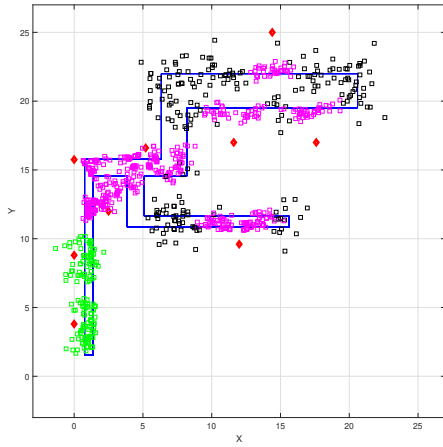
So với FP, các phương pháp định vị được tham số hóa có tính thực nghiệm cao vì sự đơn giản về mặt tính toán cũng như không tiêu tốn quá nhiều thời gian cho việc thu thập dữ liệu. Đánh đổi lại, độ chính xác của chúng chỉ ở mức trung bình trong những trường hợp tổng quát. Trong khi đó, FP là một phương pháp dựa trên dữ liệu lớn. Nó dùng dữ liệu lớn để mô tả sự thay đổi của RSS cho từng vị trí trong môi trường. Chúng ta dễ thấy rằng khi có càng nhiều dữ liệu, độ chính xác của các phương pháp học máy sử dụng để khớp dữ liệu trong FP càng chính xác. Tuy nhiên, dữ liệu luôn



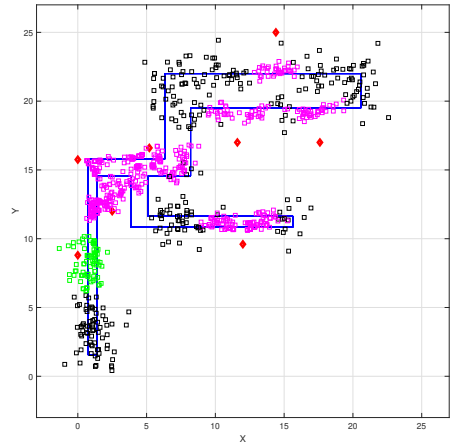
(a) 11 đèn hiệu BLE



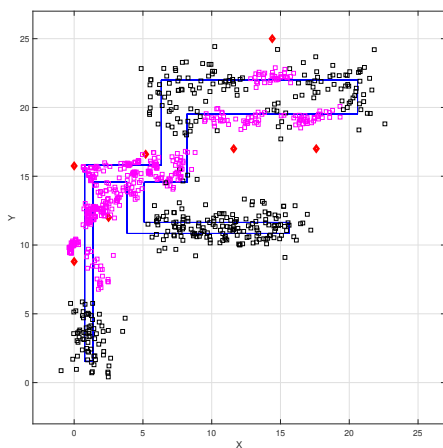
(b) 10 đèn hiệu BLE



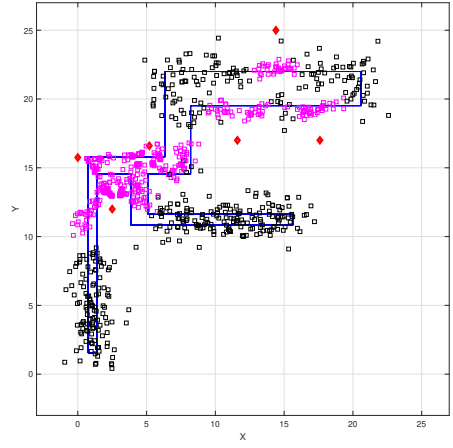
(c) 9 đèn hiệu BLE



(d) 8 đèn hiệu BLE

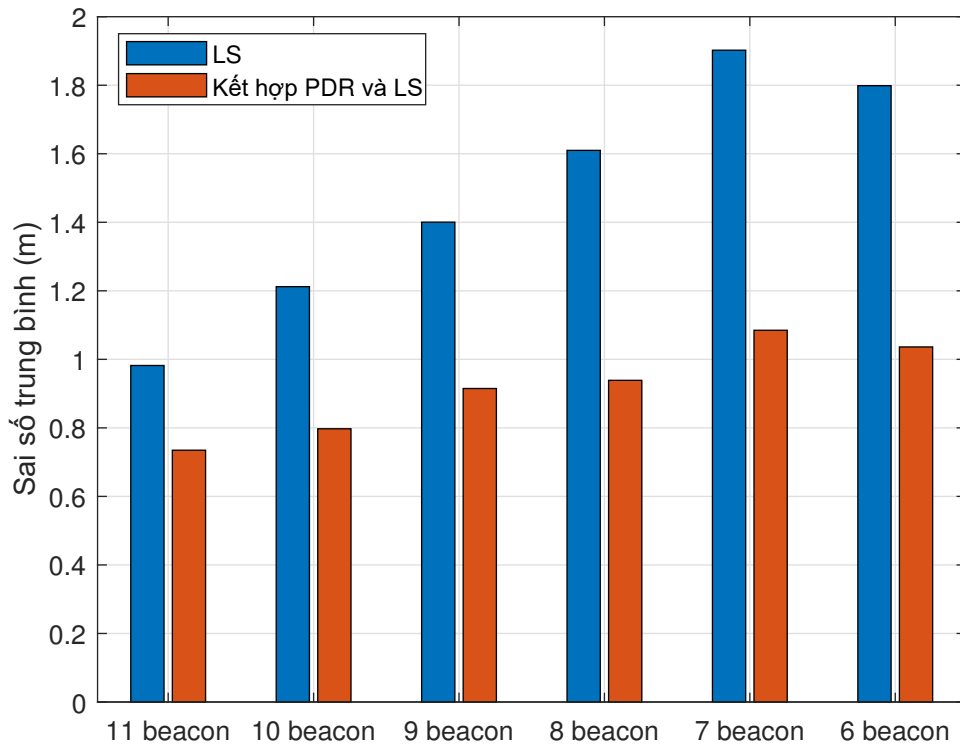


(e) 7 đèn hiệu BLE



(f) 6 đèn hiệu BLE

Hình 3.10: Phân phối của ước lượng cuối cùng khi thay đổi số lượng đèn hiệu BLE.



Hình 3.11: Sai số định vị với số lượng đèn hiệu BLE khác nhau.

đắt giá. Việc thu thập dữ liệu cho một lượng lớn các điểm tham chiếu (hoặc khu vực con) đi ngược lại với tính khả thi. Trong thực tế, người ta chỉ có thể thu thập dữ liệu cho một lượng hữu hạn các khu vực con trong một khu vực lớn, ví dụ các khu vực con có diện tích $10 \times 10 \text{ m}^2$ trong một trung tâm thương mại rộng $100 \times 100 \text{ m}^2$. Kết quả là cách tiếp cận FP có độ phân giải thấp trong ước lượng vị trí. Nó chỉ có thể giúp người ta phân biệt được họ đang thuộc một khu vực thô nào đó và đôi khi ước lượng thô đó không thể đáp ứng cho một số loại hình dịch vụ nhất định. Vì vậy, trong phần này, luận án đề xuất một phương pháp có khả năng cân bằng giữa hai phương pháp định vị tham số và không tham số.

3.3.1 Hệ thống đề xuất

Sơ đồ của hệ thống đề xuất được mô tả bởi Hình 3.12. Hệ thống gồm hai khối chức năng, hoạt động theo kiểu tuần tự, bao gồm khối ước lượng vị trí thô sử dụng mạng

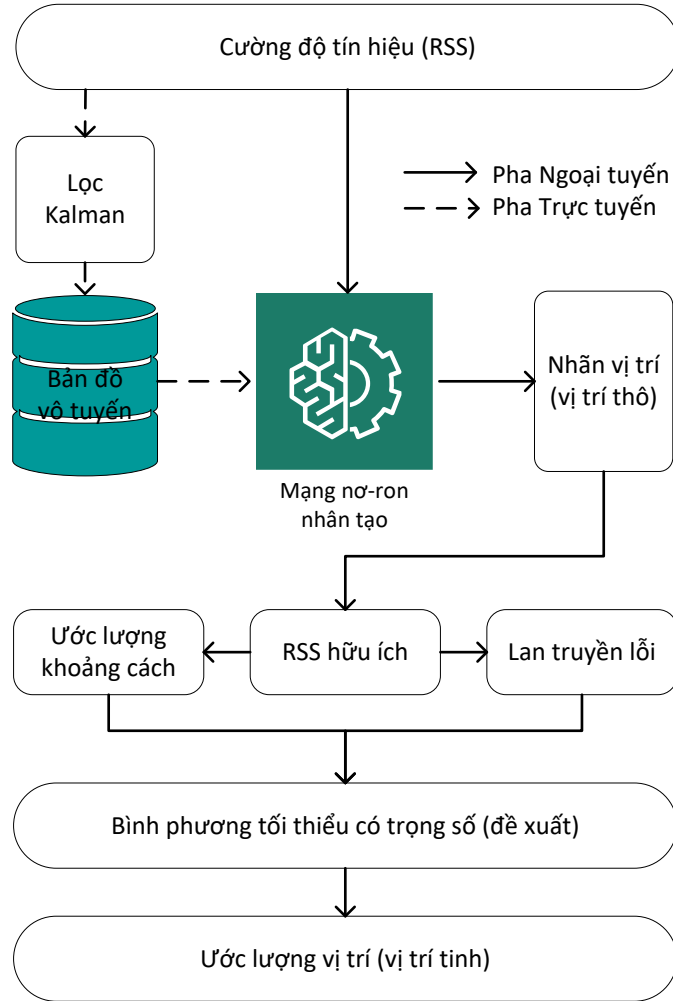
neuron đa lớp (ANN-MLP) và khối ước lượng chính xác vị trí dùng sử dụng phương pháp bình phương nhỏ nhất có trọng số (WLS). Theo ngôn ngữ học máy, hệ thống hoạt động theo 2 giai đoạn, bao gồm giai đoạn thu thập và huấn luyện dữ liệu (training) và bước ước lượng vị trí (testing). Ở giai đoạn thu thập dữ liệu, RSS thô được đưa qua bộ lọc Kalman để loại bỏ hiệu ứng nhiễu. Sau đó, RSS đã lọc của tất cả khu vực con quan sát được từ tất cả các đèn hiệu xung quanh khu vực triển khai sẽ được lưu vào một cơ sở dữ liệu. Sau đó, dữ liệu được huấn luyện bởi mạng neuron nhân tạo nhiều lớp. Trong giai đoạn ước lượng vị trí, RSS thô được sử dụng thay vì RSS được lọc. Đầu ra của giai đoạn này là các nhãn của các khu vực con trên bản đồ. Thông tin về các khu vực con sau đó được sử dụng như một chiến lược lựa chọn đèn hiệu cho quá trình ước lượng vị trí sử dụng LS kế sau. Giá trị RSS từ đèn hiệu đã chọn được gọi là giá trị RSS tin cậy. Những giá trị RSS này cung cấp hai thông tin quan trọng *i)* để tính khoảng cách giữa các đèn hiệu và nút mục tiêu thông qua một hàm ánh xạ và *ii)* để đánh giá độ không chắc chắn của chính mẫu RSS đó. Cuối cùng, mô-đun bình phương nhỏ nhất có trọng số thực hiện phép nhân hệ thống LS thông thường với ma trận có trọng số để đưa ra ước tính vị trí cuối cùng.

3.3.1.1 Hệ thống con sử dụng mạng neuron nhân tạo đa lớp.

Kiến trúc của mạng neuron nhân tạo đa lớp được minh họa bởi Hình 3.13. Chức năng của mạng này là phân loại xem dữ liệu mới quan sát được thuộc vào khu vực con nào, nói cách khác, xác định vị trí thô của người dùng trong không gian hai chiều. Sau khi xác định được một khu vực con, phương pháp LS có trọng số tính toán vị trí của người dùng một cách chi tiết hơn. Mạng neuron đề xuất có hai lớp ẩn. Lớp đầu tiên bao gồm $2N$ neuron, ứng với cặp thông tin về mã định danh và cường độ tín hiệu của tất cả N đèn hiệu. Lớp này được kết nối hoàn toàn với lớp ẩn thứ hai. Hàm kích hoạt mà được sử dụng ở đây là hàm Rectified Linear Unit (ReLU), được biểu diễn bởi công thức

$$\text{ReLU} = \max(0, x) \quad (3.13)$$

Với mong muốn mạng neuron có khả năng phân loại, lớp thứ hai sử dụng M neuron ứng với M khu vực con cần phân loại. Hàm softmax được lựa chọn cho lớp ẩn thứ hai

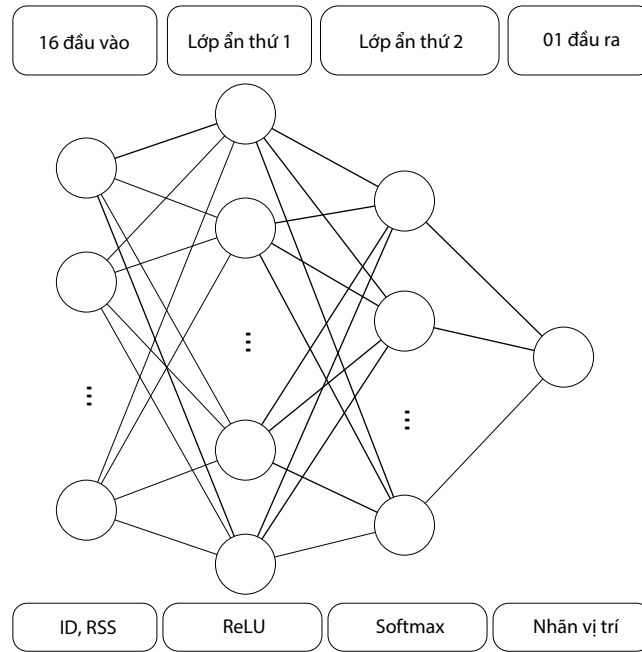


Hình 3.12: Kiến trúc hệ thống đề xuất

này. Đầu ra của hàm softmax đánh giá điểm tin cậy khi phân loại kết quả. Xác suất của đầu ra được cho bởi (nằm trong khoảng từ 0 đến 1)

$$\sigma(z)_i = \frac{e^{z_i}}{\sum_{j=1}^K e^{z_j}} \quad (3.14)$$

Mô hình được huấn luyện theo cả hai hướng là truyền xuôi và truyền ngược. Ngoài ra, do đặc điểm nhiều của phép đo RSS thô, luận án thêm vào một bộ lọc Kalman như một bước xử lý trước để làm mịn dữ liệu và giảm các hiệu ứng sai lệch trước khi huấn luyện mô hình.



Hình 3.13: Cấu tạo mạng nơ-ron đa lớp đề xuất cho bài toán phân loại khu vực.

3.3.1.2 Chiến lược chọn đèn hiệu BLE tin cậy

Sử dụng chiến lược lựa chọn đèn hiệu hiệu quả có thể cải thiện độ chính xác và độ tin cậy của các phương pháp sử dụng khoảng cách, ví dụ như Trilateration [12]. Theo cách thông thường, nút mục tiêu sẽ lắng nghe tất cả các đèn hiệu BLE được triển khai trong môi trường trong nhà và ước tính vị trí thông qua tất cả các khoảng cách có được. Điều này có 2 ảnh hưởng xấu tới hệ thống: *i)* độ phức tạp của thuật toán lớn và *ii)* độ chính xác của vị trí giảm. Chúng ta dễ dàng nhận thấy rằng phương pháp LS có trọng số bao gồm bước nghịch đảo ma trận, do đó, độ phức tạp thuật toán của nó là $\mathcal{O}(n^3)$ trong đó n là số lượng đèn hiệu được phát hiện. Các thủ tục tính toán khác bao gồm phép nhân ma trận có độ phức tạp thuật toán là $\mathcal{O}(n)$. Về lý thuyết, chúng ta có thể thấy đây là một phương pháp có độ phức tạp thuật toán cao, đặc biệt khi n trở nên lớn. Vấn đề thứ hai liên quan đến độ chính xác của hệ thống. Vì RSS bị ảnh hưởng trực tiếp bởi môi trường vật lý, các giá trị RSS quan sát được từ các đèn hiệu BLE không phải đều đáng tin cậy, đặc biệt là những đèn hiệu BLE ở xa mục tiêu. Do đó, chúng ta chỉ nên chọn một số đèn hiệu BLE nhất định để xác định vị trí mục tiêu. Chiến lược đề xuất được thực hiện theo cách sau. Đầu tiên, các mạng ANN-MLP cung

cấp mô hình để ước lượng khu vực thô mà nút đích đang thuộc về. Sau đó, kiến thức về khu vực thô hiện tại sẽ chỉ ra đâu là đèn hiệu BLE ở gần khu vực đó. Tất nhiên, thông tin này phải được cung cấp trước và được lưu trong cơ sở dữ liệu. Ví dụ, khi hệ thống xác định rằng người dùng đang ở khu vực thứ nhất thì chỉ có bốn đèn hiệu BLE gần khu vực đó được sử dụng để ước lượng vị trí (xem Hình 3.14).

3.3.1.3 Hệ thống con sử dụng LS

Giả sử N đèn hiệu BLE được chọn từ chiến lược chọn đèn hiệu được trình bày trong chương trước và một nút mục tiêu nằm gần lượt có tọa độ $[x_i, y_i]^\top$, $i \in \{1, \dots, N\}$ và $[x, y]^\top$ trong không gian 2 chiều. Vị trí của $\mathbf{x}$ được tính bởi

$$\mathbf{x}_{LS} = (\mathbf{A}^\top \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{b}, \quad (3.15)$$

trong đó $\mathbf{A}$ và $\mathbf{b}$ được xây dựng từ tọa độ đã biết của các đèn hiệu và khoảng cách ước lượng được như trong (2.61) và (2.63).

3.3.1.4 Phương pháp LS có trọng số sử dụng phân tích lỗi lan truyền

Phương pháp LS coi độ tin cậy của các khoảng cách đầu là giống nhau. Tuy nhiên, từ Hình 3.1 và Hình 3.3, chúng ta có thể thấy rằng phép đo ứng với RSS cao hơn có độ tin cậy cao hơn. Do đó, bằng cách thêm trọng số vào các phép đo chính xác hơn (tương ứng với khoảng cách ngắn), chúng ta hy vọng sẽ có được độ chính xác cao hơn trong kết quả định vị. Ước tính có trọng số của $\mathbf{x}_{LS}$ có thể được biểu diễn bởi

$$\mathbf{x}_{WLS} = (\mathbf{A}^\top \mathbf{S}^{-1} \mathbf{A})^{-1} \mathbf{A}^\top \mathbf{S}^{-1} \mathbf{b}. \quad (3.16)$$

Ma trận trọng số $\mathbf{S}$ mà luận án chọn chính là ma trận hiệp phương sai của $\mathbf{b}$ - trong đó có lỗi truyền từ RSS đến khoảng cách. Giả sử $\mathbf{b}$ có thể đơn giản hóa bằng $\mathbf{b} = [X_1, X_2, \dots, X_{N-1}]^\top$. Ma trận hiệp phương sai, $\mathbf{S}$, của $\mathbf{b}$ được định nghĩa bởi

$$\mathbf{S} = \begin{bmatrix} \sigma_{X_1}^2 & \sigma_{X_1 X_2} & \cdots & \sigma_{X_1 X_{N-1}} \\ \sigma_{X_2 X_1} & \sigma_{X_2}^2 & \cdots & \sigma_{X_2 X_{N-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{X_{N-1} X_1} & \sigma_{X_{N-1} X_2} & \cdots & \sigma_{X_{N-1}}^2 \end{bmatrix}. \quad (3.17)$$

Giả sử $d_1, d_2, \dots, d_N$ là độc lập với nhau từng đôi một và có phân phối giống hệt nhau. Do đó, $X_1, X_2, \dots, X_{N-1}$ cũng độc lập với nhau từng đôi một. Gọi c là một hằng số, ta có $\sigma_c^2 = 0$, $\sigma_{X+c}^2 = \sigma_X^2$, và $\sigma_{cX}^2 = c^2 \sigma_X^2$. Khi đó, chúng ta có thể dễ dàng tính được $\mathbf{S}$ theo công thức

$$\mathbf{S} = \begin{bmatrix} \sigma_{d_1^2}^2 + \sigma_{d_2^2}^2 & \sigma_{d_1^2}^2 & \cdots & \sigma_{d_1^2}^2 \\ \sigma_{d_1^2}^2 & \sigma_{d_1^2}^2 + \sigma_{d_3^2}^2 & \cdots & \sigma_{d_1^2}^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{d_1^2}^2 & \sigma_{d_1^2}^2 & \cdots & \sigma_{d_1^2}^2 + \sigma_{d_N^2}^2 \end{bmatrix}. \quad (3.18)$$

Các giá trị $\sigma_{d_i^2}^2$ của ma trận hiệp phương sai $\mathbf{S}$ sau đó có thể được tính bởi

$$\sigma_{d_i^2}^2 = \mathbb{E}(d_i^4) - (\mathbb{E}(d_i^2))^2 \quad (3.19)$$

Dựa trên hàm sinh mô-men của biến ngẫu nhiên Gaussian, luận án dễ dàng tính được $\mathbb{E}(d_i^4) = e^{4\mu_{d_i} + 8\sigma_{d_i}^2}$ và $(\mathbb{E}(d_i^2))^2 = e^{4\mu_{d_i} + 4\sigma_{d_i}^2}$. Khi đó, các thành phần không nằm trên đường chéo chính của $\mathbf{S}$ được tính bởi

$$\sigma_{d_1^2}^2 = \exp\left(\frac{\mu_{d_1} + 2\sigma_{d_1}^2}{\mu_{d_1} + \sigma_{d_1}^2}\right) \quad (3.20)$$

Trong khi đó, các phần tử thuộc đường chéo chính của $\mathbf{S}$ có thể được tính bởi (với $j \in \{2, \dots, N\}$)

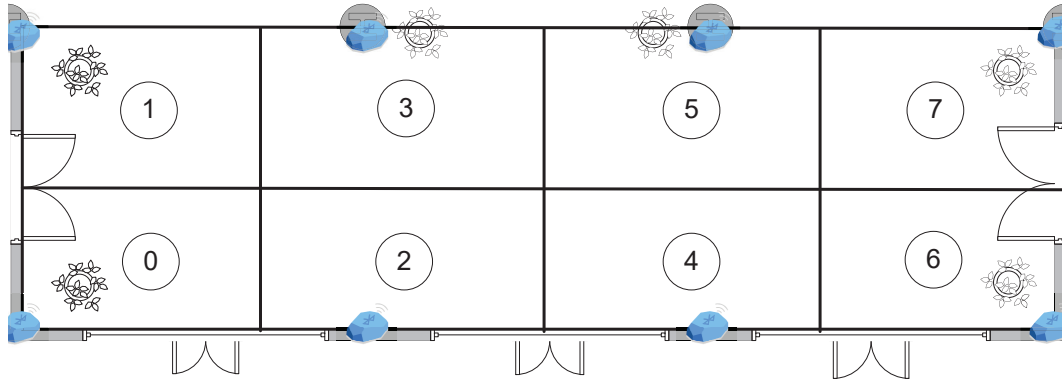
$$\sigma_{d_1^2}^2 + \sigma_{d_j^2}^2 = \exp\left(\frac{\mu_{d_1} + 2\sigma_{d_1}^2}{\mu_{d_1} + \sigma_{d_1}^2}\right) + \exp\left(\frac{\mu_{d_j} + 2\sigma_{d_j}^2}{\mu_{d_j} + \sigma_{d_j}^2}\right). \quad (3.21)$$

Trong phương trình (3.20) và (3.21), $\mu_{d_i}, i \in [N]$ được tính toán thông qua các công thức trong Bảng 3.2, nói cách khác, dựa trên giá trị RSS tức thời ($\mathcal{R}$) quan sát được trên nút mục tiêu và hệ số góc của hàm mà tại đó RSS thuộc về. Trong khi đó, σ_{d_i} được tính qua hệ số góc của hàm lan truyền, thuộc về các mô hình chuyển đổi trong Bảng 3.2 và phương trình (3.8).

3.3.2 Kết quả và thảo luận

Để đánh giá hệ thống đề xuất, luận án chọn sảnh chính tại tầng một tòa nhà G2, Đại học Công Nghệ, ĐHQGHN là khu vực thí nghiệm. Tổng diện tích khu vực là khoảng 100 m^2 , được chia thành một lưới gồm tám khu vực nhỏ hơn có kích thước bằng nhau

và được đánh số từ 1 đến 8. Tám đèn hiệu BLE được bố trí xung quanh khu vực thí nghiệm ở độ cao 1,5 m. Nút mục tiêu là Raspberry Pi 3 (RPi). Sơ đồ khu vực thí nghiệm được thể hiện trong Hình 3.14. RPi đầu tiên được dùng để quét và đọc tín hiệu



Hình 3.14: Bản đồ khu vực thí nghiệm.

từ đèn hiệu BLE. Các dữ liệu sau đó được sử dụng với hai mục đích là huấn luyện và đánh giá độ chính xác. RPi được đặt trên một giá ba chân cố định có chiều cao bằng 1.5 m. Hướng và vị trí thu thập dữ liệu không thay đổi cho tất cả các phép đo. Trong giai đoạn thu thập dữ liệu, 400 mẫu RSSI ở trung tâm của từng khu vực con được ghi lại. Do đó, tổng số mẫu dữ liệu có được là 3200 mẫu cho tám vùng nhỏ. Trong giai đoạn Trực tuyến, luận án sử dụng trực tiếp giá trị RSS thô, với vị trí của RPi là ngẫu nhiên và không di chuyển.

3.3.2.1 Khả năng xác định khu vực con chỉ sử dụng mạng nơ-ron

Bảng 3.3: Thống kê chi tiết lỗi của 8 khu vực con khi chỉ sử dụng mạng nơ-ron

Khu vực con	1	2	3	4	5	6	7	8	TB
Đúng	1196	576	1196	1084	1130	1038	1165	1200	1073
Sai	1	614	4	116	70	161	34	0	125
Độ chính xác	99.92	48.8	99.67	90.33	94.17	86.57	97.16	100	89.6

Xác suất dự đoán chính xác cho 8 khu vực con sử dụng mạng nơ-ron được mô tả bằng ma trận nhầm lẫn trong Hình 3.15. Từ Hình 3.15, chúng ta có thể nhận thấy hầu

1	1196			1				
2	552	576	62					
3		4	1196					
4			116	1084				
5				11	1130	59		
6					137	1038	24	
7							1165	34
8								1200
	1	2	3	4	5	6	7	8

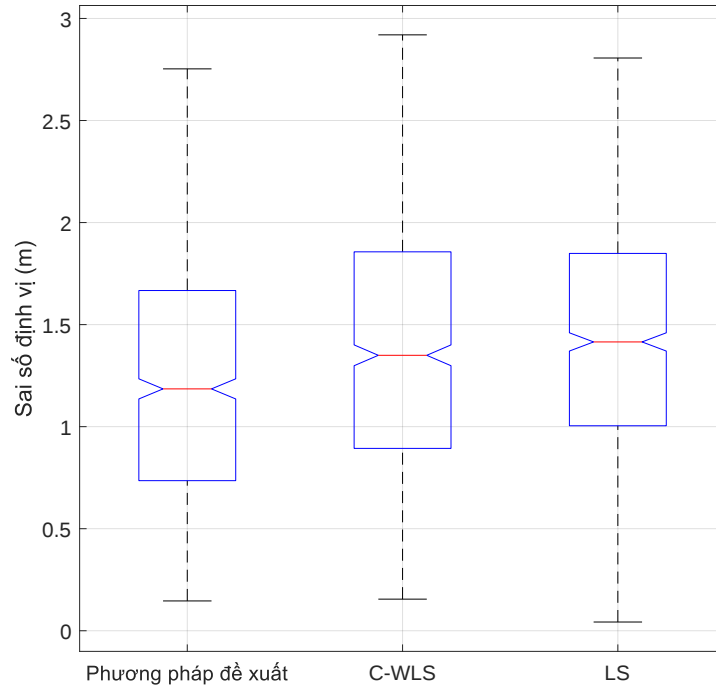
Lớp dự đoán

Hình 3.15: Ma trận nhầm lẫn khi ước lượng tho vị trí máy thu.

hết các khu vực con đều được nhận dạng với tỉ lệ thành công cao (xác suất trên 86%). Ngoại lệ chỉ xảy ra với khu vực con thứ hai khi mà ô này chỉ có xác suất nhận dạng chính xác là 51,5%. Tuy nhiên, khi nhìn vào sự khác biệt giữa nó và khu vực bị nhầm lẫn, chúng ta có thể thấy đây là điều chấp nhận được vì nó chỉ bị nhầm lẫn với các ô ở cạnh (ô thứ nhất và ô thứ ba), tức có sai số không đáng kể. Độ chính xác chi tiết cho từng khu vực con được liệt kê trong Bảng 3.3. Từ Bảng 3.3, chúng ta có thể tính được độ chính xác trung bình của nhãn bằng 89,52%. 10% lỗi có thể được giải thích bởi nhiễu có trong môi trường, bao gồm ảnh hưởng của các thiết bị không dây khác và sự hiện diện của người đi lại quanh khu vực thí nghiệm.

3.3.2.2 Ước lượng tình không sử dụng hỗ trợ của mạng nơ-ron

Trong phần này, độ chính xác của phương pháp đề xuất được so sánh với các phương pháp ước lượng cổ điển không sử dụng mạng nơ-ron. Các phương pháp dùng để so sánh bao gồm LS và LS có trọng số thông thường, viết tắt là CWLS. CWLS nhận hệ

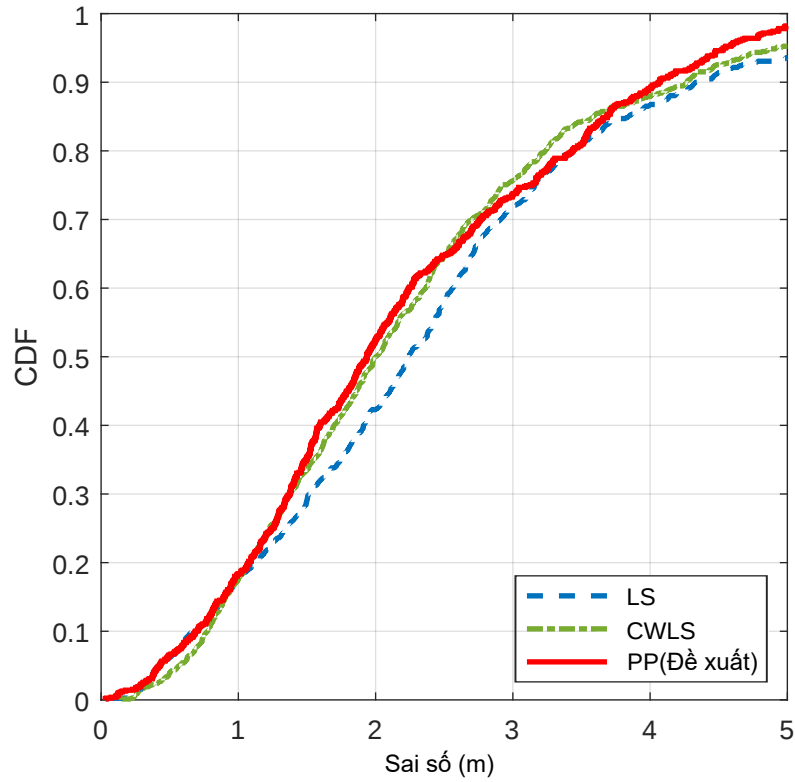


Hình 3.16: Độ chính xác của các phương pháp định vị

thống LS với ma trận trọng số có dạng

$$\mathbf{S} = \begin{bmatrix} d_2 & 0 & \cdots & 0 \\ 0 & d_3 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & d_N \end{bmatrix}. \quad (3.22)$$

Sai số theo mô phỏng và sai số trên thực tế lần lượt được biểu diễn bởi Hình 3.16 và Hình 3.17. Với kết quả mô phỏng được cho bởi Hình 3.16, chúng ta thấy rằng độ chính xác của 3 phương pháp là tương đương nhau với sai số trung bình khoảng 1.35 m. Sai số lần lượt của phương pháp đề xuất, C-WLS và LS là 1.24, 1.39 và 1.44. Như vậy, phương pháp đề xuất làm tốt hơn 14% so với C-WLS và 10% so với C-WLS về mặt lý thuyết. Hình 3.17 so sánh sai số tích lũy của ba phương pháp: LS, CWLS và phương pháp WLS đề xuất khi thử nghiệm trong môi trường thực tế. Nhìn chung, các phương pháp đều cho kết quả gần như giống nhau, tức có sai số dưới 4 m với xác suất 90 %. Trong số 3 phương pháp, LS cho độ chính xác thấp nhất với sai số trung bình khoảng 2.2 m. Điều đáng chú ý là CWLS và phương pháp WLS đề xuất có độ chính xác gần

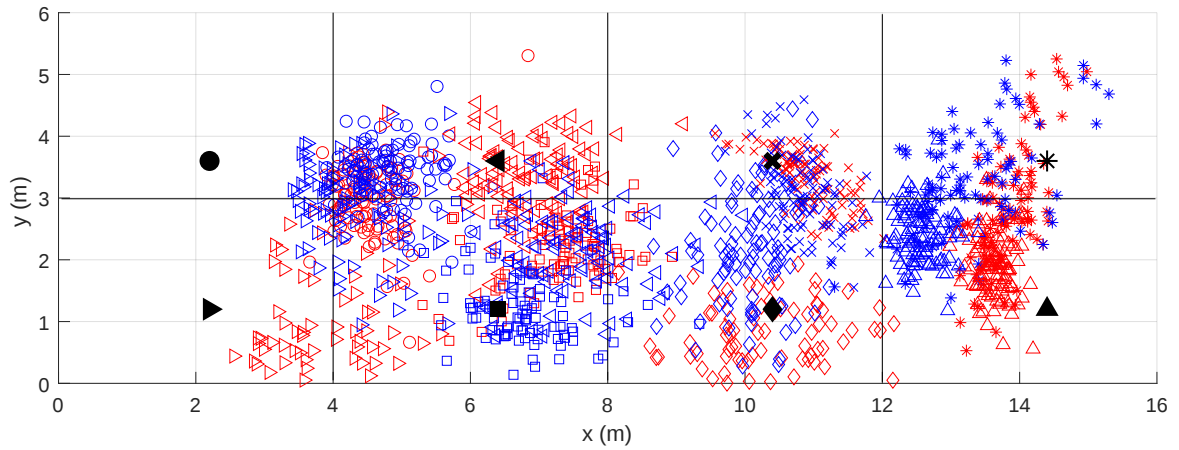


Hình 3.17: Ước lượng tình không sử dụng mạng nơ-ron

như giống nhau với sai số trung bình khoảng 2 m. Điều này xảy ra vì rất khó để có được một mô hình ước lượng khoảng cách từ RSS không lệch (unbiased) (3.11). Điều này cũng đúng với mô hình ở (3.8). Một số nghiên cứu [29, 30] chỉ ra rằng 3 - 4 m là khoảng cách đáng tin cậy. Ngoài phạm vi khoảng cách này, người ta khó có thể ước tính khoảng cách chính xác từ RSS với độ chính xác cao. Nói cách khác, ở khoảng cách gần, người ta dễ dàng có mô hình ước lượng (3.8) và (3.11) không lệch và ngược lại. Nếu không có mạng nơ-ron, các phép đo của tất cả các tín hiệu quan sát được sẽ được đưa vào hệ thống LS. Do đó, ước tính cuối cùng bị ảnh hưởng bởi độ lệch gây ra bởi các phép ước lượng của 3.8 và 3.11. Đây là lý do tại sao độ chính xác của WLS đề xuất không có mạng nơ-ron gần giống như hiệu suất của CWLS.

3.3.2.3 Ước lượng tình sử dụng mạng nơ-ron

Trong phần này này, luận án so sánh ba phương pháp trên với sự tham gia của mạng nơ-ron, nói cách khác, có sự tham gia của chiến lược chọn đèn hiệu BLE. Kết quả được

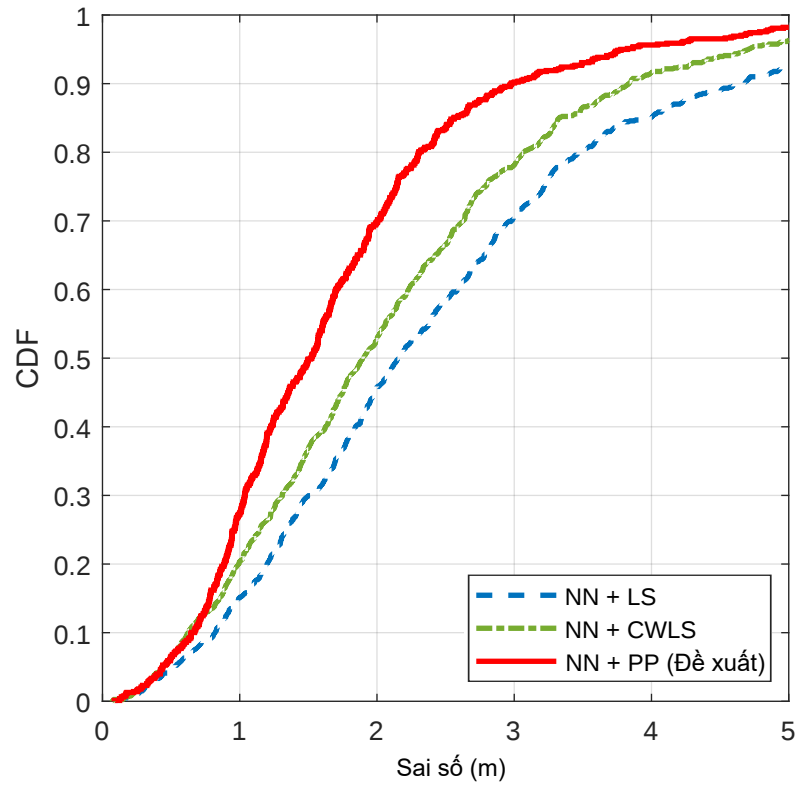


Hình 3.18: Phân bố của vị trí ước lượng từ sự kết hợp giữa NN và CWLS (NN + CWLS) và sự kết hợp giữa NN và phương pháp WLS đề xuất (NN + PP-WLS).

Bảng 3.4: So sánh các phương pháp định vị khác nhau về sai số bình phương trung bình (RMSE), trung vị, trung bình, phân vị 70 và 90.

Thuốc đo	NN + LS	NN + CWLS		NN + PP	
	Sai số (m)	Sai số (m)	Khác biệt (%)	Sai số (m)	Khác biệt (%)
RMSE	2.63	2.27	13.68	1.78	32.32
Trung vị	2.10	1.88	10.47	1.50	28.57
Trung bình	2.30	1.98	13.91	1.57	31.74
Phân vị 75 %	3.07	2.67	13.02	2.07	32.57
Phân vị 90 %	4.13	3.53	14.52	2.67	35.35

biểu diễn trong Hình 3.18, Hình 3.19 và Bảng. Trong Hình 3.18, các khối đặc màu đen biểu diễn cho các vị trí kiểm tra. Các khối rỗng biểu diễn vị trí ước lượng của các vị trí kiểm tra có cùng hình dạng. Các khối màu đỏ và xanh lần lượt biểu diễn vị trí ước lượng bởi sự kết hợp NN + PP-WLS và NN + CWLS. Từ Hình 3.18, chúng ta có thể thấy mô hình vị trí ước tính của CWLS và WLS đề xuất, cả hai đều kết hợp với NN. Có thể thấy, các ước tính hầu như tập trung vào các ô mà chúng thuộc về. Phân bố của các vị trí cũng chỉ ra rằng ước tính của phương pháp đề xuất tốt hơn CWLS. Từ Hình 3.19, chúng ta có thể thấy rằng có sự cải thiện, khoảng 10 - 30%, khi kết hợp mạng nơ-ron và phương pháp dựa trên khoảng cách. Khi tận dụng sức mạnh của mạng



Hình 3.19: Lỗi tích của các phương pháp khi sử dụng mạng nơ-ron như một chiến lược chọn lựa đèn hiệu.

nơ-ron làm chiến lược lựa chọn đèn hiệu, chúng ta có thể thấy rõ sự khác biệt giữa ba phương pháp. Như được thể hiện trong Bảng 3.4, phương pháp WLS đề xuất vượt trội hơn cả LS và CWLS ở 5 chỉ số đánh giá, với mức trung bình cải thiện là 32%. Điều này cũng dễ hiểu vì phương pháp sử dụng kết hợp mạng nơ-ron và WLS đề xuất sẽ loại bỏ các báo hiệu ở xa và đồng thời tận dụng được toàn bộ thông tin, bao gồm cả giá trị trung bình và phương sai của RSS, thay vì chỉ sử dụng khoảng cách trung bình như CWLS.

3.4 Kết luận chương

Chương này bắt đầu với việc tìm hiểu về đặc điểm của phép đo RSS trong môi trường trong nhà. Các kết quả thí nghiệm cho thấy RSS có tính chất không đồng nhất, tức độ tin cậy của phép đo RSS giảm đi khi khoảng cách giữa máy phát và máy thu

tăng lên. Dựa trên đặc điểm này, luận án đề xuất hai phương pháp mới giúp các hệ thống định vị trong nhà đạt độ chính xác cao, bao gồm:

1. Một phương thức triển khai đèn hiệu cho hệ thống định vị trong nhà sử dụng đèn hiệu BLE. Giải pháp này triển khai các đèn hiệu ở dạng tam giác đều. Nó đặc biệt phù hợp với hệ thống định vị sử dụng kết hợp PDR và RSS - nơi mà cả vị trí do PDR và RSS đều không hoàn toàn tin cậy ở mọi thời điểm. Dựa trên tính chất không đồng nhất của RSS, các phép đo RSS được coi là tin cậy ở những khoảng nhất định và được sử dụng để sửa lỗi cho PDR. Chiến lược triển khai đèn hiệu cho phép hệ thống quan sát được các RSS tin cậy nhiều nhất có thể. Kết quả thí nghiệm cho thấy hệ thống định vị trong nhà sử dụng PDR và RSS cùng với chiến lược triển khai này đạt sai số tương đối cao với sai số khoảng 1 m với số lượng đèn hiệu tối thiểu.
2. Một phương pháp kết hợp hài hòa hai cách tiếp cận: dựa trên khoảng cách và không dựa trên khoảng cách. Với kích thước lưới lớn, phương pháp FP đòi hỏi ít nỗ lực thu thập dữ liệu hơn. Để có ước tính chính xác, luận án tận dụng các lưới tham chiếu để chọn ra các đèn hiệu có RSS đáng tin cậy thông qua mạng nơ-ron nhân tạo, phục vụ cho phương pháp dựa trên khoảng cách sau đó. Để đạt được độ chính xác cao hơn, luận án đề xuất một phương pháp bình phương nhỏ có trọng số (WLS). Kết quả thí nghiệm cho thấy độ chính xác của mạng nơ-ron độc lập tương đối cao khi xác định chính xác tới 86% nhãn vị trí. Hơn nữa, với sự trợ giúp của mạng nơ-ron như một chiến lược lựa chọn đèn hiệu, phương pháp WLS đề xuất đạt được độ chính xác cao với sai số trung bình khoảng 1,7 m, cao hơn 15% khi không sử dụng mạng nơ-ron.

Trong chương tiếp theo, luận án sẽ trình bày các thuật toán định vị ngoài trời sử dụng hệ thống MIMO mmWave. Sự phát triển của các thuật toán này là một phần trong việc xây dựng một hệ thống định vị liên mạch giữa môi trường trong nhà và ngoài trời. Ngoài ra, chúng được kỳ vọng không chỉ cải thiện độ chính xác của vị trí mà còn giúp mở rộng ứng dụng của các dịch vụ dựa trên vị trí.

Chương 4

Định vị với các phép đo TOA, AOA, AOD trong hệ thống mmWave MIMO

Chương 4 trình bày các phương pháp ước lượng tham số kênh mmWave cho định vị ngoài trời, bao gồm AOA, AOD và TOA. Với đặc điểm tán xạ kém của sóng mmWave, luận án bắt đầu bằng việc xây dựng bài toán ước lượng kênh thưa. Dựa trên các thuật toán cơ bản ước lượng thưa họ OMP cơ bản, luận án đề xuất các phương pháp ước lượng kênh theo hướng cải thiện độ chính xác hoặc giảm độ phức tạp thuật toán. Cụ thể, luận án đề xuất 01 phương pháp ước lượng tham số kênh không xem xét sự ảnh hưởng giữa các đường dẫn dựa trên thuật toán StOMP và phương pháp Silhouette và 01 phương pháp ước lượng tham số kênh có xem xét sự ảnh hưởng giữa các đường dẫn dựa trên cấu trúc kênh và thuật toán OMP. Nhận thấy các tham số ước lượng ở bước trước đó bị lượng tử hóa bởi việc chia lưới cơ sở, luận án đề xuất 01 phương pháp tinh chỉnh tham số dựa trên hàm mất mát LS, phương pháp suy giảm tuần tự và suy giảm dựa trên đạo hàm. Các thuật toán đề xuất được đánh giá dựa trên *i)* Độ chính xác ước tính, đo bằng Xác suất phục hồi thành công (PSR) và sai số bình phương trung bình (RMSE), *ii)* khả năng chống chịu với nhiễu, *iii)* khả năng chống chịu với các mức liên kết tương hỗ khác nhau và *iv)* độ phức tạp thuật toán hoặc thời gian chạy. Kết quả cho thấy các thuật toán đề xuất đều có sự cải thiện ở ít nhất một trong các yếu

tổ trên.

*Nội dung của Chương 4 liên quan chủ yếu đến tới các công trình nghiên cứu **CT3**, **CT4**, **CT5**, **CT6** và **CT7**.*

4.1 Giới thiệu

Trong vài thập kỷ qua, sự phổ biến của các dịch vụ dựa trên vị trí làm cho vấn đề định vị sử dụng trở thành một chức năng quan trọng. Các hệ thống đầu tiên cung cấp khả năng định vị bao gồm hệ thống định vị toàn cầu (GPS) và Galileo [54]. Với sự tiến bộ trong công nghệ truyền thông không dây, bài toán xác định vị trí mục tiêu trong mạng di động đã thu hút được sự chú ý đáng kể, đặc biệt kể từ khi hệ thống 4G ra đời với những loại hình dịch vụ mới, ví dụ như thực tế tăng cường, trò chơi dựa trên vị trí, chăm sóc sức khỏe và quảng cáo dựa trên vị trí, v.v. Trong thế hệ mạng di động tiếp theo (5G và sau 5G), chúng ta sẽ thấy dịch vụ dựa trên vị trí dần tự khẳng định mình là một phần không thể thiếu trong mạng di động tế bào. Nhận thấy sự quan trọng đó, 3GPP đã xem xét vấn đề định vị trong các tài Release 16 [97] và Release 17 [98]. Theo đó, một số yêu cầu đặt ra với định vị như: Đối với trường hợp thông thường, mức độ sai số cho phép phải nhỏ hơn 50 m với phương ngang và 5 m với phương dọc và độ trễ nhỏ hơn 3 giây. Tuy nhiên, để có thể thương mại hóa dịch vụ định vị hoặc liên quan tới định vị, độ chính xác phải cao hơn. Cụ thể, sai số phải nhỏ hơn 3 m đối với cả phương ngang và phương dọc cho trường hợp định vị trong nhà và sai số phải nhỏ hơn 10 m với phương ngang và 3 m với phương dọc cho trường hợp định vị ngoài trời. Độ trễ cho cả trong nhà và ngoài trời đều phải dưới 1 giây. Các phép đo được khuyến cáo sử dụng bao gồm thời gian truyền (TOA hoặc RTT), góc đi (AOD), góc đến (AOA) và cường độ tín hiệu (RSS). Từ Release 16 trở đi, khả năng ước lượng vị trí các thiết bị người dùng thông qua các phép đo này hoàn toàn có thể biến 5G thành một hệ thống định vị độc lập. Các tính năng định vị 5G được mô tả trong các phiên bản báo cáo kỹ thuật của 3GPP:

- TS 22.071 [75]: Location Services (LCS) Service description - Stage 1: định nghĩa, mô tả logic, các dịch vụ có liên quan tới chức năng định vị, đưa ra cách quản trị

hệ thống, xác định khả năng tương tác với các dịch vụ khác và mô tả các dịch vụ dựa trên vị trí có thể có.

- TS 23.273 [76]: 5G System Location Services (LCS) - Stage 2: định nghĩa kiến trúc hệ thống, xác định các tính năng, quy trình của dịch vụ định vị, quy trình lưu trữ thông tin của người dùng trong hệ thống.
- TS 38.305 [77]: Stage 2 functional specification of UE Positioning in Next-Generation Radio Access Network (NG-RAN): xác định kiến trúc định vị NG-RAN cho UE, các giao thức báo hiệu, các phương pháp định vị UE trong NG-RAN.

Ngoài ra, vì một số nguyên nhân kỹ thuật mà chỉ đến thế hệ mạng thứ 5, vấn đề định vị mới được coi là khả thi. So với hệ thống 5G, các kỹ thuật định vị mà mạng 3G hay 4G có thể sử dụng bao gồm:

- **Ước lượng khoảng cách** thông qua các phép đo như RSS, TOA hoặc TDOA. Với cách này, người ta sử dụng thông tin khoảng cách từ ít nhất 3 trạm gốc để xác định vị trí của một người dùng nằm trong vùng bao phủ của chúng.
- **Ước lượng góc** với sự hỗ trợ của các phương pháp dựa trên ước lượng không gian con như Multiple Signal Classification (MUSIC, tạm dịch: phân loại tín hiệu tổng hợp) [63] hay Estimation of signal Parameters via Rotational Invariance (ESPRIT, tạm dịch: Ước lượng tham số của tín hiệu thông qua phép quay bất biến) [62]. Người ta có thể xác định mục tiêu di động trong mạng với chỉ 2 trạm gốc.

Tuy nhiên, một số thách thức mà định vị trong các hệ thống 3G và 4G (và nói chung trong các hệ thống sử dụng tần số dưới 6 GHz) gặp phải là sự tồn tại của một lượng lớn các thành phần đa đường hoặc yêu cầu nhiều trạm gốc để định vị hoặc sự phức tạp trong việc triển khai.

- **Tồn tại số lượng lớn các thành phần đa đường** Các thành phần đa đường là các bản sao của tín hiệu, được tạo ra trong quá trình truyền sóng trong môi trường vật lý, chẳng hạn như phản xạ, nhiễu xạ, v.v. Tín hiệu quan sát được trong một khoảng thời gian nhất định tại máy thu sẽ là một hỗn hợp gồm các

phiên bản trễ khác nhau của một tín hiệu được gửi đi. Các thành phần tín hiệu này gây ra bất lợi cho các phương pháp sử dụng thời gian truyền, góc đến và RSS. Đối với RSS, đa đường tạo ra sự dao động lớn về công suất nhận được. Theo một cách khác, thời gian đến khác nhau của các thành phần đa đường tạo ra sự sai lệch trong việc xác định khoảng cách từ máy thu đến máy phát. Tương tự, các tín hiệu đa đường đến từ nhiều hướng có thể tạo ra lỗi trong quá trình xác định AoA nếu độ phân giải góc lớn.

- **Định vị sử dụng nhiều trạm gốc** Với hệ thống chỉ sử dụng TOA hoặc TDOA hoặc RSS, người ta cần ít nhất 3 BS để có thể xác định được vị trí mục tiêu. Trong khi đó, người ta cần ít nhất 2 BS để có thể xác định được vị trí mục tiêu nếu hệ thống đó sử dụng chỉ sử dụng AOA.
- **Sự phức tạp của hệ thống định vị và thuật toán trong nó** Hệ thống sử dụng TOA và TDOA yêu cầu đồng bộ chặt chẽ về mặt thời gian giữa máy thu và máy phát hoặc giữa các máy thu với nhau. Trong khi đó, hệ thống sử dụng phép đo góc thường phải sử dụng mảng ăng-ten thông minh với số phần tử lớn để tạo các búp sóng định hướng có độ chia góc nhỏ. Điều này gián tiếp làm cho độ phức tạp của các thuật toán định vị tăng lên.

Tuy nhiên, với đặc điểm tự nhiên của sóng mmWave, sự hỗ trợ của công nghệ MIMO và kỹ thuật tạo chùm sóng có độ phân giải cao, định vị trong các thế hệ mạng tiếp theo trở nên khả thi.

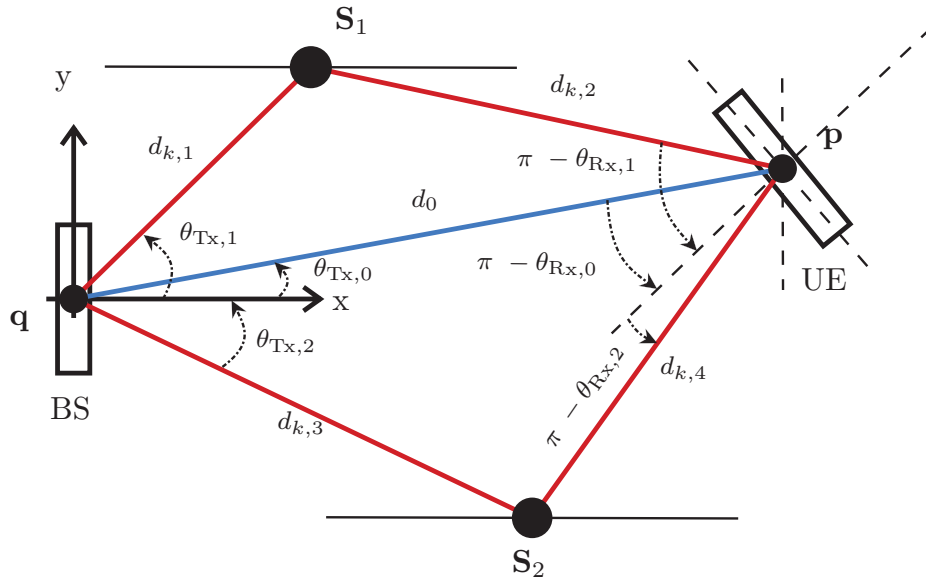
- **Độ phân giải cao** Một công nghệ có phù hợp với định vị hay không phụ thuộc vào một tham số gọi là độ phân giải thô [55]. Độ phân giải thô được định nghĩa bằng quãng đường truyền được của một tín hiệu giữa 2 lần lấy mẫu liên tiếp. Ví dụ, đối với UWB, tín hiệu có băng thông 500 MHz, độ phân giải thô đạt được là khoảng cách mà tín hiệu truyền được trong khoảng thời gian lấy mẫu 2 ms và do đó cho độ phân giải là khoảng 60 cm. Có thể thấy, dải băng thông rộng của sóng mmWave sẽ trở nên phù hợp hơn bất kỳ công nghệ không dây nào trong bài toán định vị.

- Kênh có ít các thành phần đa đường** Tín hiệu sử dụng mmWave tại tần số cao có đặc điểm là suy hao mạnh. Thuộc tính này làm giảm số lượng các thành phần đa đường và do đó kênh trở nên thưa về mặt không gian. [56] chỉ ra rằng số lượng thành phần đa đường nằm trong khoảng 2 - 5 đường. Vì vậy, tín hiệu nhận được là tổng hợp của một lượng nhỏ các thành phần đa đường. Với đặc điểm này, chúng ta có thể phân tích các tham số góc và thời gian của từng đường dẫn để xác định vị trí của một mục tiêu. Thứ nhất, một khi các thành phần đa đường được phân tích, người ta có thể ước tính được vị trí mục tiêu mà chỉ cần sử dụng 1 trạm gốc với đồng thời hai phép đo góc và thời gian truyền. Trong trường hợp này, các thành phần đa đường được sử dụng như một thông tin mang tính xây dựng để cải thiện độ chính xác định vị [57]. Thứ hai, chúng ta có thể sử dụng các thành phần đa đường để định vị không chỉ máy thu mà còn cả các yếu tố liên quan tới môi trường vật lý, ví dụ như vị trí của các điểm/mặt phẳng phản xạ, từ đó góp phần lập lại bản đồ, phục vụ cho các ứng dụng khác.
- Mảng ăng-ten lớn** Một đặc tính khác của hệ thống 5G ảnh hưởng đến khả năng định vị là việc sử dụng ăng-ten có nhiều phần tử. Khi tạo búp sóng, mảng ăng-ten lớn cho mức SNR tốt hơn so với mảng ăng-ten có số lượng phần tử nhỏ hơn theo một hướng nhất định và do đó góp phần tích cực vào việc ước tính các tham số kênh như TOA, AOD và AOA. [73] chỉ ra rằng giới hạn lý thuyết về hiệu suất định vị phụ thuộc trực tiếp vào SNR thu được. Hơn nữa, số lượng lớn ăng-ten cũng tạo búp sóng có độ phân giải góc cao hơn, từ đó góp phần giảm sai số định vị.
- Mạng siêu dày đặc** Khoảng cách giữa các trạm được thu hẹp lại cho phép người dùng có được đường dẫn trực tiếp (LOS) với trạm gốc cao hơn. Xác suất này, về mặt trực quan, làm giảm sai số trong việc ước tính vị trí người dùng. Ngoài ra, các trạm gần nhau cho phép người dùng nhận được thông tin từ nhiều trạm gốc hơn, từ đó cho phép các thuật toán ước lượng chính xác vị trí của nó hơn.

4.2 Định vị dựa trên các tham số kênh MIMO mmWave

4.2.1 Mô hình kênh MIMO mmWave

Luận án xem xét mô hình kênh truyền hoạt động trên N sóng mang con. Kênh giữa trạm gốc (BS) và thiết bị người dùng (UE) là kênh đa đường, có thể được biểu diễn bằng mô hình Saleh-Valenzuela với K đường dẫn. Mô hình hình học của kênh trong không gian 2 chiều được minh họa bởi Hình 4.1. Trong đó, BS và UE được trang bị



Hình 4.1: Mô hình hình học bao gồm ba đường dẫn, một đường LOS (màu xanh) và hai đường NLOS (màu đỏ).

ăng-ten ULA với số phần tử ăng-ten lần lượt là N_t và N_r . Ở dạng ma trận, kênh MIMO được biểu diễn bởi

$$\mathbf{H}(n) = \mathbf{A}_{\text{Rx}}[n]\mathbf{\Gamma}[n]\mathbf{A}_{\text{Tx}}[n]^\top, \quad (4.1)$$

trong đó, $\mathbf{A}_{\text{Tx}}[n] \in \mathbb{C}^{N_t \times K}$ và $\mathbf{A}_{\text{Rx}}[n] \in \mathbb{C}^{N_r \times K}$ lần lượt là véc-tơ điều khiển và véc-tơ đáp ứng. Chúng được định nghĩa bởi

$$\mathbf{A}_{\text{Tx}}[n] = [\mathbf{a}_{\text{Tx},n}(\theta_{\text{Tx},0}), \dots, \mathbf{a}_{\text{Tx},n}(\theta_{\text{Tx},K-1})]^\top, \quad (4.2)$$

và

$$\mathbf{A}_{\text{Rx}}[n] = [\mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},0}), \dots, \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},K-1})]^\top. \quad (4.3)$$

Với mảng ULA, véc-tơ điều khiển $\mathbf{a}_{\text{Tx},n} \in \mathbb{C}^{N_t}$ và véc-tơ đáp ứng $\mathbf{a}_{\text{Rx},n} \in \mathbb{C}^{N_r}$ của đường dẫn thứ k lần lượt được định nghĩa bởi

$$\mathbf{a}_{\text{Tx},n}(\theta_{\text{Tx},k}) = [1, e^{-j\pi \sin(\theta_{\text{Tx},k})}, \dots, e^{-j\pi(N_t-1) \sin(\theta_{\text{Tx},k})}]^\top, \quad (4.4)$$

và

$$\mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},k}) = [1, e^{-j\pi \sin(\theta_{\text{Rx},k})}, \dots, e^{-j\pi(N_r-1) \sin(\theta_{\text{Rx},k})}]^\top. \quad (4.5)$$

Khi viết các véc-tơ điều khiển và đáp ứng như trong (4.4) và (4.5), luận án đã xấp xỉ chiều dài bước sóng của tín hiệu (λ_n) trên sóng mang con thứ n thành λ_c và khoảng cách giữa 2 phần tử ăng-ten liền kề là bằng nhau và bằng $\lambda_c/2$. Ma trận chứa độ lợi kênh và suy hao, $\mathbf{\Gamma}[n]$, được định nghĩa bởi

$$\mathbf{\Gamma}[n] = \begin{bmatrix} \tilde{h}_0 e^{\frac{-j2\pi n \tau_0}{NT_s}} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \tilde{h}_{K-1} e^{\frac{-j2\pi n \tau_{K-1}}{NT_s}} \end{bmatrix}, \quad (4.6)$$

trong đó $\tilde{h}_k = \frac{\sqrt{N_t N_r}}{\rho_k} h_k$ là độ lợi kênh phức. Đồng thời, luận án giả sử kênh là không thay đổi trong thời gian kết hợp. Giả sử BS truyền tín hiệu hoa tiêu $\mathbf{x}$ theo kênh đường xuống thông qua một ma trận tạo búp sóng $\mathbf{F}[n]$. Tín hiệu nhận được là

$$\mathbf{y}^{(g)}[n] = \mathbf{H}[n] \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] + \mathbf{n}^{(g)}[n], \quad (4.7)$$

trong đó, chỉ số (g) biểu thị cho lần truyền thứ g trong chuỗi G lần truyền. Nhiều $\mathbf{n}^{(n)}[n]$ có trung bình bằng 0 và phương sai $\sigma^2 = N_0$ cho 2 chiều thực và phức của tín hiệu. Vấn đề của bài toán định vị trong mạng 5G mmWave là ước lượng các tham số kênh chưa biết, bao gồm AOD, AOA, TOA và độ lợi kênh khi biết $\mathbf{y}^{(g)}[n]$.

4.2.2 Biểu diễn thưa của kênh MIMO mmWave trong miền góc

Chia tập xác định $[-\frac{\pi}{2}, \frac{\pi}{2})$ thành các góc rời rạc với độ chia là G_t , G_r và định nghĩa $\bar{\omega}_T[i], i \in [G_t]$, $\bar{\omega}_R[j], j \in [G_r]$ lần lượt là góc thứ i và góc thứ j . Luận án xây dựng $\mathbf{U}_{\text{Rx}} \in \mathbb{C}^{N_r \times G_r}$ và $\mathbf{U}_{\text{Tx}} \in \mathbb{C}^{N_t \times G_t}$ lần lượt là các đối ứng của $\mathbf{A}_{\text{Rx}}$ và $\mathbf{A}_{\text{Tx}}$. Chúng là các ma trận đơn nhất giúp chuyển đổi tín hiệu sang miền góc

$$\mathbf{U}_{\text{Tx}} = [\mathbf{u}_{\text{Tx}}(\bar{\omega}_{\text{Tx}}[1]), \mathbf{u}_{\text{Tx}}(\bar{\omega}_{\text{Tx}}[2]), \dots, \mathbf{u}_{\text{Tx}}(\bar{\omega}_{\text{Tx}}[G_t])], \quad (4.8)$$

$$\mathbf{U}_{\text{Rx}} = [\mathbf{u}_{\text{Rx}}(\bar{\omega}_{\text{Rx}}[1]), \mathbf{u}_{\text{Rx}}(\bar{\omega}_{\text{Rx}}[2]), \dots, \mathbf{u}_{\text{Rx}}(\bar{\omega}_{\text{Rx}}[G_r])] , \quad (4.9)$$

trong đó $\mathbf{u}_{\text{Tx}}(\bar{\omega}_{\text{Tx}}[i])$ và $\mathbf{u}_{\text{Rx}}(\bar{\omega}_{\text{Rx}}[j])$ lần lượt là cột thứ i và cột thứ j của $\mathbf{U}_{\text{Tx}}$ và $\mathbf{U}_{\text{Rx}}$. Chúng được tạo ra theo cùng một cách tương tự với $\mathbf{a}_{\text{Tx}}$ và $\mathbf{a}_{\text{Rx}}$. Véc-tơ hóa $\boldsymbol{\mu}[n] = \mathbf{H}[n]\mathbf{F}^{(g)}[n]\mathbf{x}^{(g)}[n]$, ta thu được

$$\begin{aligned} \text{vec}(\boldsymbol{\mu}[n]) &= \text{vec} \left(\mathbf{U}_{\text{Rx}} \mathbf{U}_{\text{Rx}}^H \mathbf{H}[n] \mathbf{U}_{\text{Tx}} \mathbf{U}_{\text{Tx}}^H \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] \right) \\ &= \frac{1}{G_t G_r} \left(\mathbf{U}_{\text{Tx}}^H \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] \right)^\top \otimes \mathbf{U}_{\text{Rx}} \text{vec} \left(\mathbf{U}_{\text{Rx}}^H \mathbf{H}[n] \mathbf{U}_{\text{Tx}} \right). \end{aligned} \quad (4.10)$$

Đại lượng $\check{\mathbf{H}}[n] = \mathbf{U}_{\text{Rx}}^H \mathbf{H}[n] \mathbf{U}_{\text{Tx}} \in \mathbb{C}^{G_r \times G_t}$ được gọi là kênh trong miền góc và có thể biểu diễn bởi

$$\check{\mathbf{H}}[n] = \sum_{k=0}^{K-1} \mathbf{U}_{\text{Rx}}^H \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},k}) \tilde{h}_k e^{\frac{-j2\pi n \tau_k}{NT_s}} \mathbf{a}_{\text{Tx},n}^H(\theta_{\text{Tx},k}) \mathbf{U}_{\text{Tx}}. \quad (4.11)$$

Đặt $\boldsymbol{\Omega}[n] = \left(\mathbf{U}_{\text{Tx}}^H \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] \right)^\top \otimes \mathbf{U}_{\text{Rx}}$ và $\check{\mathbf{h}}[n] = \text{vec}(\check{\mathbf{H}}[n])$, (4.7) trở thành

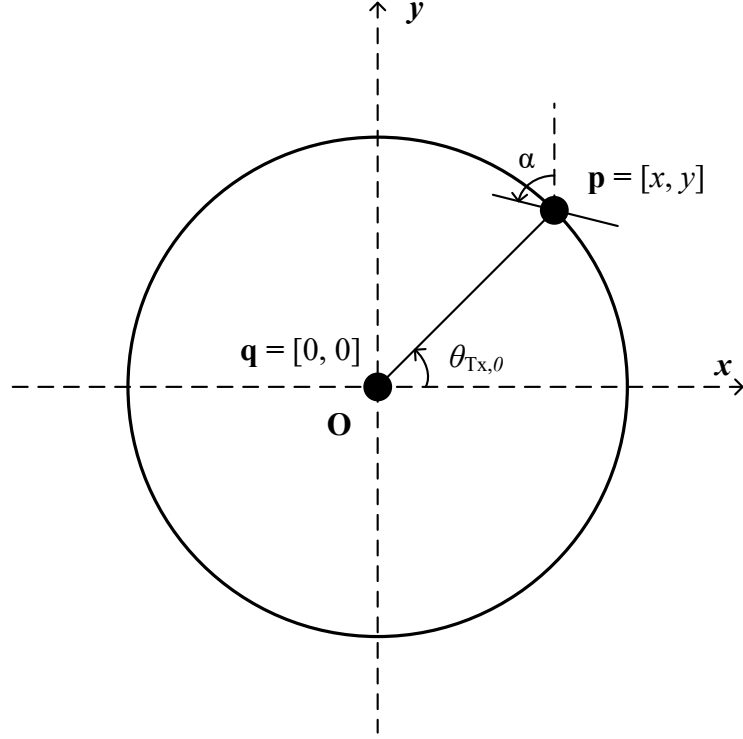
$$\check{\mathbf{y}}[n] = \boldsymbol{\Omega}[n] \check{\mathbf{h}}[n] + \check{\mathbf{n}}[n]. \quad (4.12)$$

4.2.3 Chuyển đổi từ các tham số kênh truyền sang vị trí mục tiêu

Định vị trong hệ thống MIMO mmWave sử dụng các phép đo góc và khoảng cách tương tự những phương pháp truyền thống. So với các hệ định vị trong mạng 3G hay 4G, hệ định vị sử dụng MIMO mmWave trong 5G không sử dụng nhiều BS cùng lúc để xác định vị trí của người dùng thông qua thuật toán LS hay kỹ thuật giao 3 điểm. Về mặt thực quan, người ta có thể xác định vị trí của UE thông qua một BS duy nhất, thậm chí khi không tồn tại đường truyền trực tiếp từ BS tới UE. Cụ thể, số lượng lớn các ăng-ten trên BS có thể tạo ra búp sóng với độ rộng rất nhỏ, một khi hệ thống có thể phân tích các thành phần đa đường, nó sẽ tính toán được vị trí UE thông qua đồng thời phép đo khoảng cách và góc.

4.2.3.1 Định vị với đường dẫn LOS

Xem xét trường hợp tồn tại một đường LOS trong không gian hai chiều, phép đo về thời gian (TOA) giữa UE và BS tạo ra một đường tròn là quỹ tích các điểm mà UE có



Hình 4.2: Định vị với đường dẫn LOS

thể nằm trên đó. Trong khi đó, phép đo về góc (AOD) xác định một hướng duy nhất mà búp sóng trở tới tới UE. Giao điểm giữa góc đi và đường tròn chính là vị trí của UE. Về mặt toán học, các mối quan hệ này được biểu diễn bởi công thức

$$(q_x - p_x)^2 + (q_y - p_y)^2 = d_0^2 \quad (4.13)$$

và

$$\tan(\theta_{Tx,0}) = \frac{p_y - q_y}{p_x - q_x} \quad (4.14)$$

Giải hai phương trình trên ta thu được

$$\mathbf{p} = \mathbf{q} + d_0 \mathbf{u}(\theta_{Tx,0}), \quad (4.15)$$

trong đó $\theta_{Tx,0} = [\cos(\theta_{Tx,0}), \sin(\theta_{Tx,0})]^\top$. Hướng của ăng-ten trên máy thu được ước lượng bởi

$$\alpha = \pi + \theta_{Tx,0} - \theta_{Rx,0}. \quad (4.16)$$

4.2.3.2 Định vị với đường dẫn NLOS

Trong trường hợp không có đường dẫn LOS, người ta vẫn có thể xác định được mục tiêu với ít nhất 2 đường dẫn NLOS. Trong trường hợp này, vị trí của UE được tính bởi

$$\|\mathbf{p} - \mathbf{s}_1\|_2 + \|\mathbf{q} - \mathbf{s}_1\|_2 = d_1 \quad (4.17a)$$

$$\|\mathbf{p} - \mathbf{s}_2\|_2 + \|\mathbf{q} - \mathbf{s}_2\|_2 = d_2 \quad (4.17b)$$

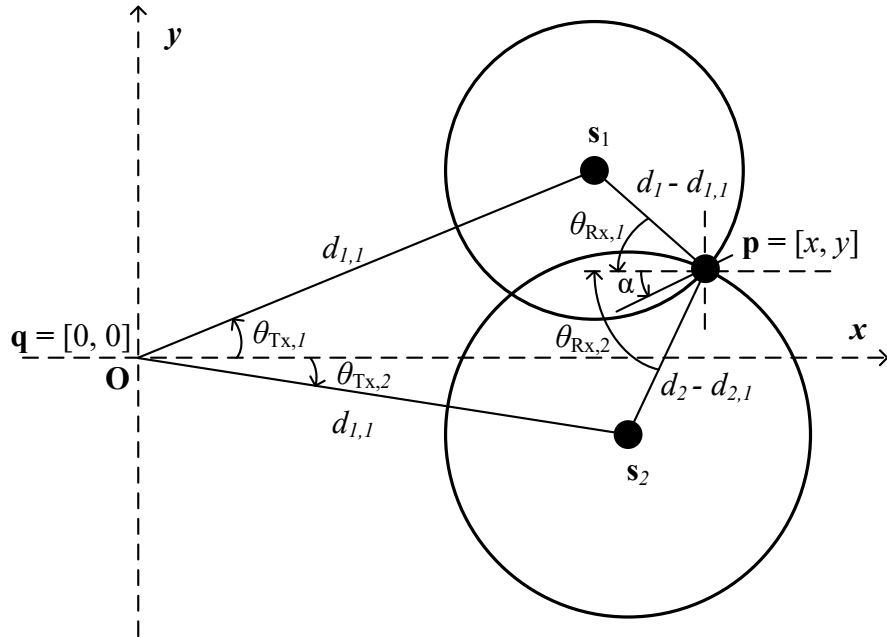
$$\mathbf{s}_1 = \mathbf{q} + d_{1,1} \mathbf{u}(\theta_{\text{Tx},1}) \quad (4.17c)$$

$$\mathbf{s}_2 = \mathbf{q} + d_{2,1} \mathbf{u}(\theta_{\text{Tx},2}) \quad (4.17d)$$

$$\tan(\pi - (\theta_{\text{Rx},1} + \alpha_0)) = \frac{s_{1,y} - p_y}{p_x - s_{1,x}} \quad (4.17e)$$

$$\tan(\pi - (\theta_{\text{Rx},2} + \alpha_0)) = \frac{-s_{2,y} + p_y}{p_x - s_{2,x}} \quad (4.17f)$$

trong đó, $\mathbf{q}$ là biết trước. Tập $\{\theta_{\text{Tx},k}, \theta_{\text{Rx},k}, d_k\}$ chứa các tham số về AOD, AOA và



Hình 4.3: Định vị với 2 đường dẫn NLOS

khoảng cách của đường dẫn NLOS thứ k và có được thông qua quá trình ước lượng kênh. $\mathbf{p}$ là vị trí chưa biết của UE. Tập $\{\mathbf{s}_k, d_{k,1}\}$ chứa tham số chưa biết về vị trí của điểm phản xạ thứ k và khoảng cách từ điểm phản xạ đó đến BS. Xem xét bài toán trong không gian 2 chiều, chúng ta có tất cả 8 tham số cần phải ước lượng.

4.3 Các thuật toán ước lượng thưa họ OMP truyền thống

Để trích xuất các tham số định vị, người ta phải ước tính một vectơ kênh thưa $\mathbf{h}$ trong hệ thống $\mathbf{y} = \mathbf{\Omega}\mathbf{h} + \mathbf{n}$, trong đó $\mathbf{\Omega}$ và $\mathbf{y}$ là những đại lượng đã biết, đại diện cho ma trận cảm nhận và tín hiệu nhận được. Bài toán ước lượng thưa được cho bởi

$$\hat{\mathbf{h}} = \min_{\mathbf{h}} \|\mathbf{h}\|_0 \quad \text{s.t.} \quad \|\mathbf{y} - \mathbf{\Omega}\mathbf{h}\|_2^2 < \epsilon, \quad (4.18)$$

trong đó ϵ là sai số dừng. Đây được coi là bài toán ngược và có thể giải quyết bằng các phương pháp ước lượng thưa họ OMP.

4.3.1 Thuật toán OMP

Thuật toán OMP là một nhánh của các thuật toán khôi phục dựa trên tính *tham lam*. Các thuật toán này có độ phức tạp tính toán thấp nhất trong các loại thuật toán ước lượng thưa và được sử dụng rộng rãi trong lĩnh vực CS. Vì $\mathbf{x}$ chỉ có S phần tử khác không, vec-tơ $\mathbf{y} = \mathbf{A}\mathbf{x}$ sẽ được tính bằng tổ hợp tuyến tính của S cột trong $\mathbf{A}$. Để tái tạo lại được $\mathbf{x}$, chúng ta cần phải xác định được cột nào trong $\mathbf{A}$ tham gia nhiều nhất trong việc cấu thành nên $\mathbf{y}$. Tại mỗi lần lặp, nó chọn cột trong $\mathbf{A}$ có tương quan cao nhất với phần dư tín hiệu. Sau khi tìm ra được cột có đóng góp nhiều nhất, nó trừ phần ảnh hưởng của cột đó đi và tiếp tục tìm kiếm trên phần còn lại. Sau S lần lặp, chúng ta hy vọng rằng có thể tìm được một tập hợp các số là chỉ số của các cột có đóng góp nhiều nhất trong việc tạo thành tín hiệu $\mathbf{y}$. Mã giả của OMP được tóm tắt bằng Thuật toán 1.

4.3.2 Thuật toán StOMP

Thuật toán StOMP là thuật toán được phát triển từ thuật toán OMP. Điểm khác biệt giữa OMP và StOMP là StOMP không cần phải biết trước độ thưa của tín hiệu. Ngoài ra, StOMP là thuật toán dựa trên ngưỡng cứng, nó vì vậy sẽ chọn được nhiều hơn một cột tại mỗi vòng lặp, trái ngược với OMP chỉ lấy một cột. Thuật toán StOMP

Algorithm 1: Thuật toán OMP [69]

Đầu vào: Tín hiệu nhận được: $\mathbf{y}$,

Ma trận cảm nhận,

Độ thưa tín hiệu: S và Sai số dừng: ϵ .

Khởi tạo: Biến đếm $s = 0$, Tập hỗ trợ $I_s = \emptyset$, phần dư $\mathbf{r}_0 = \mathbf{y}$

while $\|\mathbf{r}_s\|_2 \geq \epsilon$ hoặc $s \leq S$ **do**

1. $s = s + 1$;
2. Chọn cột sao cho: $j = \arg \max_j |\langle \mathbf{r}_{s-1}, \boldsymbol{\Omega}_j \rangle|$;
3. Cập nhật chỉ số $I_s = I_{s-1} \cup j$;
4. Cập nhật tập hỗ trợ $\boldsymbol{\Omega}_{I_s} = \{\boldsymbol{\Omega}_i\}_{i \in I_s}$;
5. Ước lượng LS: $\mathbf{h}_s = (\boldsymbol{\Omega}_{I_s}^T \boldsymbol{\Omega}_{I_s})^{-1} (\boldsymbol{\Omega}_{I_s}^T \mathbf{y})$;
6. Cập nhật phần dư: $\mathbf{r}_s = \mathbf{y} - \boldsymbol{\Omega}_{I_s} \mathbf{h}_s$;

end

Đầu ra : $I_s, \mathbf{h}_s$.

thực hiện S vòng lặp, bắt đầu với phần dư được khởi tạo là $\mathbf{r}_0 = \mathbf{y}$, thuật toán duy trì chuỗi ước lượng $I_1, \dots, I_S$, là vị trí của các phần tử khác không trong $\mathbf{h}$. Tại vòng lặp thứ n , StOMP sử dụng bộ lọc phù hợp (matched filter) cho phần dư hiện tại, tạo ra một véc-tơ tương quan giữa phần dư hiện tại và các cột trong $\boldsymbol{\Omega}$, sau đó chọn ra các vị trí có giá trị tương quan cao hơn một ngưỡng được đặt trước, từ đó áp dụng LS cho các cột này và ước lượng được độ lợi kênh. Trong Bước 3, σ_k được tính thông qua $\sigma_k = \|\mathbf{r}_s\|_2 / \sqrt{n}$ với n là số hàng của ma trận $\boldsymbol{\Omega}$ và t_s được chọn dựa theo mức SNR. Ngưỡng này là một phần của bộ lọc phù hợp, giúp xác định tất cả các cột có giá trị tương quan lớn hơn nó. Mã giả của StOMP được tóm tắt trong Thuật toán 2.

4.3.3 OMP đồng thời (SOMP)

Đây là một biến thể của thuật toán OMP, dành riêng cho các bài toán ước lượng thưa trong đó các véc-tơ có chung tập hỗ trợ. Gọi d là chiều dài của tín hiệu nhận được trên một sóng mang con. Khi đó dạng đồng thời của tín hiệu nhận được trên N sóng mang con là $\mathbf{Y} \in \mathbb{C}^{d \times N}$. Với ma trận cảm nhận $\boldsymbol{\Omega}$ có kích thước $d \times M$, bài toán ước lượng

Algorithm 2: Thuật toán StOMP [108]

Đầu vào: Tín hiệu nhận được: $\mathbf{y}$, ma trận cảm nhận $\mathbf{\Omega}$, số vòng lặp S , dung sai: ϵ .

Khởi tạo: bộ đếm vòng lặp $s = 1$, tập hỗ trợ $I_s = \emptyset$, phần dư $\mathbf{r}_0 = \mathbf{y}$

while $\|\mathbf{r}_s\|_2 \leq \epsilon$ hoặc $s = S$ **do**

1. Chuẩn hóa ma trận sensing: $\mathbf{\Omega}_i = \frac{\mathbf{\Omega}_i}{\|\mathbf{\Omega}_i\|_2}$ với $\mathbf{\Omega}_i$ là cột thứ i của $\mathbf{\Omega}$;
2. Tính tương quan: $c_k = \mathbf{\Omega}^H \mathbf{r}_{s-1}$;
3. Chọn cột lớn hơn ngưỡng cho trước: $J_s = \{j : |c_k(j)| > t_s \sigma_s\}$;
4. Cập nhật tập hỗ trợ: $I_s = I_{s-1} \cup J_s$;
5. Áp dụng LS cho các cột vừa tìm được: $\hat{\mathbf{h}} = (\mathbf{\Omega}_{I_s}^H \mathbf{\Omega}_{I_s})^{-1} \mathbf{\Omega}_{I_s}^H \mathbf{y}$;
6. Cập nhật phần dư: $\mathbf{r}_s = \mathbf{r}_{s-1} - \mathbf{\Omega}_{I_s} \hat{\mathbf{h}}$;
7. Tăng số vòng lặp lên 1: $s = s + 1$;

end

Đầu ra: Tập hỗ trợ I_s , kênh ước lượng $\hat{\mathbf{h}}$.

đồng thời trên N sóng mang con là

$$\min_{\mathcal{H}} \|\mathbf{Y} - \mathbf{\Omega} \mathcal{H}\|_F^2 \quad \text{s.t.} \quad \mathcal{H} \text{ có } S \text{ hàng khác không,} \quad (4.19)$$

trong đó, $\mathcal{H} \in \mathbb{C}^{M \times N}$ là ma trận chứa các véc-tơ kênh thừa trên N sóng mang con. Mã giả của SOMP được tóm tắt trong Thuật toán 3

4.4 Phương pháp ước lượng thô không xét đến tương quan giữa các đường dẫn: Cách tiếp cận dựa trên thuật toán StOMP phân tán

4.4.1 Phương pháp đề xuất

4.4.1.1 Thuật toán StOMP phân tán

Ngoại trừ thù chính, cấu trúc Dirichlet còn chứa những thù phụ. Về mặt trực quan, sự mở rộng của các thù phụ tạo thành một đường chéo theo không gian cột và hàng

Algorithm 3: Thuật toán SOMP [24]

Đầu vào: $\mathbf{Y}$, $\mathbf{\Omega}$, S

Khởi tạo: $\mathbf{R}^0 = \mathbf{Y}$ và $S_0 = \emptyset$

while $s < S$ **do**

1. Tăng số vòng lặp lên 1: $s = s + 1$;
2. Chọn cột: $j_s = \arg \min_j \|\mathbf{R}_s^H \boldsymbol{\omega}_j\|_1$, với $\boldsymbol{\omega}_j$ là cột thứ j của $\mathbf{\Omega}$;
3. Cập nhật tập hỗ trợ : $S_s = S_{s-1} \cup \{j_s\}$;
4. Áp dụng LS cho các cột chọn được: $\hat{\mathcal{H}}_s = (\mathbf{\Omega}_{S_s}^H \mathbf{\Omega}_{S_s})^{-1} \mathbf{\Omega}_{S_s}^H \mathbf{Y}$;
5. Trừ phần ảnh hưởng của cột vừa tìm thấy ra khỏi phần dư:
 $\mathbf{R}_s = \mathbf{R}_s - \mathbf{\Omega}_{S_s} \hat{\mathcal{H}}_s$;

end

Đầu ra : $\hat{\mathcal{H}}_S$

của ma trận kênh. Giả sử các đường dẫn không ảnh hưởng tới nhau, luận án gọi vị trí có đỉnh cao nhất là hỗ trợ chính. Luận án tham chiếu phương trình (4.12) tới mô hình lấy mẫu nén phân tán, tức bài toán có ma trận cảm nhận $\mathbf{\Omega}$ có cùng tập hỗ trợ trên các sóng mang con khác nhau. Phương pháp đề xuất hoạt động trong tối đa S bước. Giống với thuật toán OMP, nó trừ đi ảnh hưởng của các cột tìm được khỏi véc-tơ phần dư trong mỗi vòng lặp để tạo ra phần dư mới là xấp xỉ của tín hiệu tương ứng với những phần tử khác không khác. Việc đầu tiên phải làm là chuẩn hóa năng lượng của các cột

$$\bar{\boldsymbol{\omega}}_i[n] = \frac{\boldsymbol{\omega}_i[n]}{\|\boldsymbol{\omega}_i[n]\|_2}. \quad (4.20)$$

Thuật toán bắt đầu với biến đếm $s = 0$, lời giải tạm thời $\bar{\mathbf{h}}^{(0)} = 0$ và phần dư $\mathbf{r}^{(0)} = \bar{\mathbf{y}}$, với $\bar{\mathbf{y}}$ là tín hiệu nhận được, thu được bằng cách xếp chồng tín hiệu nhận được trên các sóng mang con lên nhau. Sau đó, nó tính toán độ tương quan giữa từng cột trong ma trận cảm nhận và phần dư tương ứng. Trong trường hợp phân tán, tương quan của các cột trong ma trận cảm nhận và phần dư được tính bởi

$$c_i^{(s)} = \sum_{n=1}^N |\langle \mathbf{r}^{(s-1)}[n], \bar{\boldsymbol{\omega}}_i[n] \rangle|. \quad (4.21)$$

Quá trình tiếp theo xác định tất cả các vị trí có biên độ vượt quá một ngưỡng cứng

$$\mathcal{J}_i^{(s)} = \{i : |c_i^{(s)}| > t^{(s)}\sigma^{(s)}\}, \quad (4.22)$$

trong đó $\sigma^{(s)}$ là nhiễu cộng trắng và $t^{(s)}$ là ngưỡng cứng. Vì số lượng hàng của ma trận cảm nhận tăng lên N lần và c_i là tích lũy của N sóng mang con nên mức độ nhiễu và ngưỡng cứng được điều chỉnh theo N và lần lượt trở thành $\frac{\|\mathbf{r}^{(s)}\|_2}{N\sqrt{N_t G}}$ và $N\delta^{(s)}$, với $\delta^{(s)}$ là một số phụ thuộc vào mức SNR. Tiếp theo, thuật toán cập nhật các ước tính bằng cách hợp nhất hỗ trợ được xác định trước đó với tập hợp con mới hình thành

$$\mathcal{I}^{(s)} = \mathcal{I}^{(s-1)} \cup \mathcal{J}^{(s)}. \quad (4.23)$$

Sau bước này, nếu $|\mathcal{I}^{(s)}| = |\mathcal{J}^{(s)}|$, thuật toán chỉ cần thực hiện các bước còn lại một lần và dừng ngay sau đó. Gọi $\mathbf{\Omega}_{\mathcal{I}}[n] \in \mathbb{C}^{N_t G \times |\mathcal{I}|}$ là ma trận chứa các cột trong tập hỗ trợ $\mathcal{I}$ trên sóng mang con thứ n , luận án xây dựng ma trận tương ứng $\overline{\mathbf{\Omega}}_{\mathcal{I}} \in \mathbb{C}^{N_t G N \times |\mathcal{I}| N}$ với N sóng mang con như sau

$$\overline{\mathbf{\Omega}}_{\mathcal{I}} = \text{blkdiag}(\mathbf{\Omega}_{\mathcal{I}}[1], \mathbf{\Omega}_{\mathcal{I}}[2], \dots, \mathbf{\Omega}_{\mathcal{I}}[N]). \quad (4.24)$$

Sau đó, độ lợi kênh cho tất cả các sóng mang được ước lượng bởi

$$\overline{\mathbf{h}}_{\mathcal{I}^{(s)}}^{(s)} = \left(\overline{\mathbf{\Omega}}_{\mathcal{I}^{(s)}}^H \overline{\mathbf{\Omega}}_{\mathcal{I}^{(s)}} \right)^{-1} \overline{\mathbf{\Omega}}_{\mathcal{I}^{(s)}}^H \overline{\mathbf{y}}. \quad (4.25)$$

Tiếp theo, ảnh hưởng của các cột vừa tìm được được loại ra khỏi tín hiệu nhận được

$$\mathbf{r}^{(s)} = \overline{\mathbf{y}} - \overline{\mathbf{\Omega}} \overline{\mathbf{h}}^{(s)}. \quad (4.26)$$

Thuật toán sẽ dừng khi một trong hai điều kiện sau được đảm bảo: *i)* số vòng lặp đạt tới ngưỡng tối đa S , *ii)* $\|\mathbf{r}\|_2 < \tau$, với τ là sai số dừng, ví dụ, $\tau = 10^{-6}$. Cuối cùng, $\overline{\mathbf{h}}^{(s)}$ được coi là kết quả cuối cùng.

4.4.1.2 Thuật toán xác định độ thừa của tín hiệu

Kết thúc quá trình trên, chúng ta thu được một tập hợp là các hỗ trợ với số lượng các hỗ trợ thường lớn hơn độ thừa của kênh truyền. Bản chất của các thuật toán họ OMP là đi xác định các hỗ trợ dựa trên hệ số tương quan giữa các cột trong ma trận cảm

nhận và phần dư tín hiệu sau các vòng lặp. Chúng phải đảm bảo sự kiện sau xảy ra với xác suất cao [52].

$$E = \{ \max_{1 \leq i \leq m} |\boldsymbol{\omega}_i^H[n] \mathbf{w}[n]| < \tau \} \quad (4.27)$$

trong đó $\tau := \sigma \sqrt{2(1 + \alpha) \log m}$ với $\alpha > 0$, m là chiều dài của véc-tơ $\boldsymbol{\omega}_i$ và σ là độ lệch chuẩn của nhiễu. Dưới ảnh hưởng của nhiễu σ , thuật toán OMP không thể đảm bảo chọn được đúng hỗ trợ với chỉ một cột trong mỗi vòng lặp. Ngược lại, thuật toán StOMP lựa chọn nhiều cột trong mỗi vòng lặp dựa trên một ngưỡng cứng và đảm bảo chọn được đúng hỗ trợ với xác suất cao hơn OMP. Tuy nhiên, cách này thêm vào bước chọn tập hỗ trợ một sự mơ hồ vì ở mỗi bước nó chọn được nhiều cột thay vì một cột. Để khắc phục hiện tượng này, luận án đề xuất thuật toán Silhouette sửa đổi [66] để ước lượng độ thừa của tín hiệu và giữ lại những hỗ trợ thuộc về thùy chính. Luận án thêm vào phương pháp Silhouette cổ điển những ràng buộc mà chỉ bài toán ước lượng kênh có cấu trúc Dirichlet có. Ràng buộc thứ nhất quan tâm tới bán kính của thùy chính. Khi bán kính của một cụm dữ liệu được xem xét, thuật toán K-means (thuật toán thường đi kèm với phương pháp Silhouette) sẽ loại bỏ những trường hợp mà tại đó khoảng cách từ 1 điểm dữ liệu tới tâm cụm lớn hơn bán kính thùy chính. Sử dụng phương pháp Silhouette với ràng buộc đầu tiên, luận án tìm được số lượng cụm tối ưu và từ đó chọn giá trị lớn nhất trong mỗi cụm làm hỗ trợ chính. Gọi $\mathbf{X} = \{\mathbf{X}_i\}$ là các hỗ trợ ước lượng được sau khi áp dụng ràng buộc thứ nhất. Chúng được mô tả bởi cụm ba tham số như sau $\mathbf{X}_i = [x_i, y_i, z_i]$, với x_i , y_i và z_i lần lượt là chỉ số của AOD, chỉ số của AOA và độ lợi kênh thực của cặp AOD-AOA đó. Để loại bỏ những hỗ trợ không thuộc về thùy chính, thuật toán tìm $\mathbf{X}_j$ thỏa mãn 2 điều kiện

$$\mathbf{X}_j = \begin{cases} \arg \max_{\mathbf{X}_i \in \mathbf{X}, y_i = y_j} z_i, \forall y_i \\ z_i \geq \eta \end{cases}, \quad (4.28)$$

trong đó η là một sai số nhỏ. Điều kiện thứ nhất của (4.28) xem xét sự hợp lý của giá trị AOA và AOD trong môi trường vật lý thực tế. Khi nhiều hỗ trợ có cùng giá trị AOA hoặc AOD, thuật toán sẽ chỉ giữ lại hỗ trợ có độ lợi lớn nhất vì thực tế không có tình huống nào mà nhiều đường dẫn có chung một giá trị AOA hoặc AOD. Ràng buộc thứ hai của (4.28) xem xét độ lợi thực của một giá trị khác không nằm ở một vị

Algorithm 4: Các sử dụng phương pháp Silhouette đề xuất cho thuật toán DStOMP.

Input: Các hỗ trợ với tọa độ của chúng: $\mathbf{X}$

Khởi tạo: Biến đếm (số cụm dùng để thử): $k = 2$

Số lượng cụm tối đa cần kiểm tra: K

Độ rộng thùy chính trong miền góc (Giả sử $G_t = G_r$): r .

for $k = 2 \rightarrow K$ **do**

1. Thực hiện thuật toán phân cụm k-means: $[\mathbf{i}, \mathbf{C}] = \text{kmeans}(\mathbf{X}, k)$;

Kiểm tra bán kính của mỗi cụm

for $j = 1 \rightarrow k$ **do**

2. Trích xuất các hỗ trợ thuộc cụm thứ j -th: $\mathbf{p} = \mathbf{X}(j == \mathbf{i}, :)$;

3. Tính toán khoảng cách từ các điểm dữ liệu tới tâm cụm:

$$\mathbf{d} = \|\mathbf{p} - \mathbf{C}(j, :)\|_2^2;$$

if $\max(\mathbf{d}) > r$ **then**

4. Tăng số lượng cụm lên 1: $k = k + 1$;

5. Phân cụm k-means với giá trị k mới: $[\mathbf{i}, \mathbf{C}] = \text{kmeans}(\mathbf{X}, k)$;

break;

end

end

6. Tính toán giá trị silhouette: $\mathbf{s} = \text{silhouette}(\mathbf{X}, \mathbf{i})$;

7. Tính toán giá trị silhouette trung bình: $\mathbf{c}(k - 1) = \text{mean}(\mathbf{s})$;

end

8. Tìm ra số lượng cụm tối ưu (tương với giá trị silhouette lớn nhất):

$$[:, \hat{k}] = \max(\mathbf{c});$$

9. Tìm các hỗ trợ thỏa mãn các điều kiện (4.28);

10. Thực hiện LS để ước lượng lại độ lợi kênh cho các hỗ trợ tìm được ở bước trước đó;

Đầu ra : Kênh ước lượng được.

trí bất kỳ. Nếu một vị trí có biên độ nhỏ hơn một ngưỡng nhất định, chúng sẽ bị loại trừ. Phương pháp Silhouette với các ràng buộc trên được cho bởi Thuật toán 4.

4.4.2 Kết quả và thảo luận

Xét kênh truyền như trong Hình 4.1 với 2 đường dẫn, trong đó vị trí của BS, UE và một điểm phản xạ lần lượt là $[0, 0]$, $[4, 1]$ và $[2, 2]$. Trong mô phỏng này, luận án chọn tần số trung tâm $f_c = 60$ GHz, băng thông $B = 100$ MHz và tốc độ ánh sáng $c = 3 \times 10^8$ m/s. Số lượng ăng-ten tại máy phát và máy thu lần lượt bằng $N_t = 16$ và $N_r = 16$. Số góc ảo trong miền xác định là $N_b = 32$. Ma trận định dạng chùm tia được định nghĩa là $\mathbf{F} = e^{j\phi}$, trong đó ϕ là một góc ngẫu nhiên. SNR được đặt trong phạm vi $[-10, 20]$ dB và độ lệch chuẩn của nhiễu được tính như sau

$$\sigma = \sqrt{\frac{P_s}{10^{\frac{\text{SNR}}{10}}}}, \quad (4.29)$$

trong đó P_s là năng lượng của tín hiệu. Luận án sử dụng hai chỉ số, bao gồm xác suất xác định chính xác tập hỗ trợ (PSR) và lỗi bình phương trung bình (MSE), để đánh giá hiệu suất của hệ thống. Chúng được định nghĩa bởi

$$\text{PSR} = \frac{\text{số lần thành công}}{\text{tổng số lần thí nghiệm}}, \quad (4.30)$$

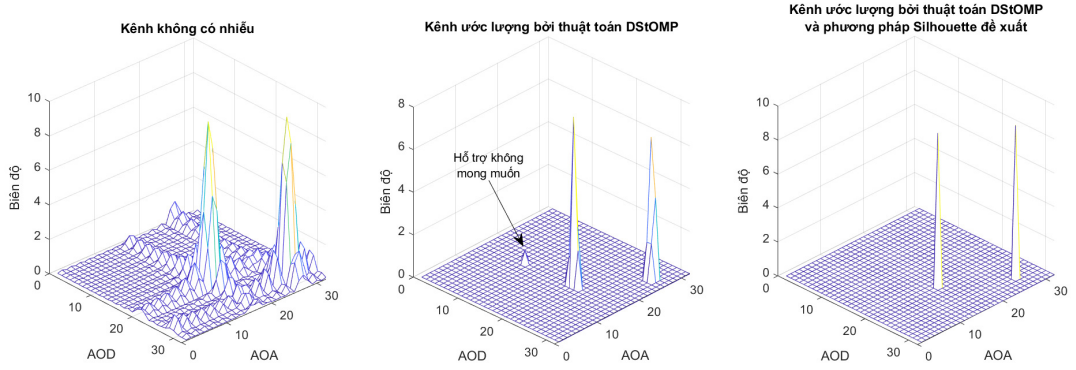
và

$$\text{MSE} = \mathbb{E} \left\{ \frac{\|\mathbf{H} - \hat{\mathbf{H}}\|_F^2}{\|\mathbf{H}\|_F^2} \right\}, \quad (4.31)$$

trong đó $\mathbf{H}$ là kênh gốc và $\hat{\mathbf{H}}$ là kênh được tái tạo lại từ các tham số kênh ước lượng. PSR và MSE có được bằng cách chạy mô phỏng 1000 lần và được đánh giá dựa trên số tín hiệu truyền đi G và số lượng sóng mang sử dụng N . Để đánh giá năng lực của phương pháp đề xuất, luận án so sánh nó với OMP tiêu chuẩn và OMP đồng thời (Simultaneous OMP) [24].

4.4.2.1 Đánh giá kết quả trực quan

Kết quả mô phỏng trực quan của các phương pháp đề xuất được minh họa bởi Hình 4.4. Hình 4.4b biểu diễn khả năng tái tạo lại kênh ban đầu chỉ sử dụng thuật toán

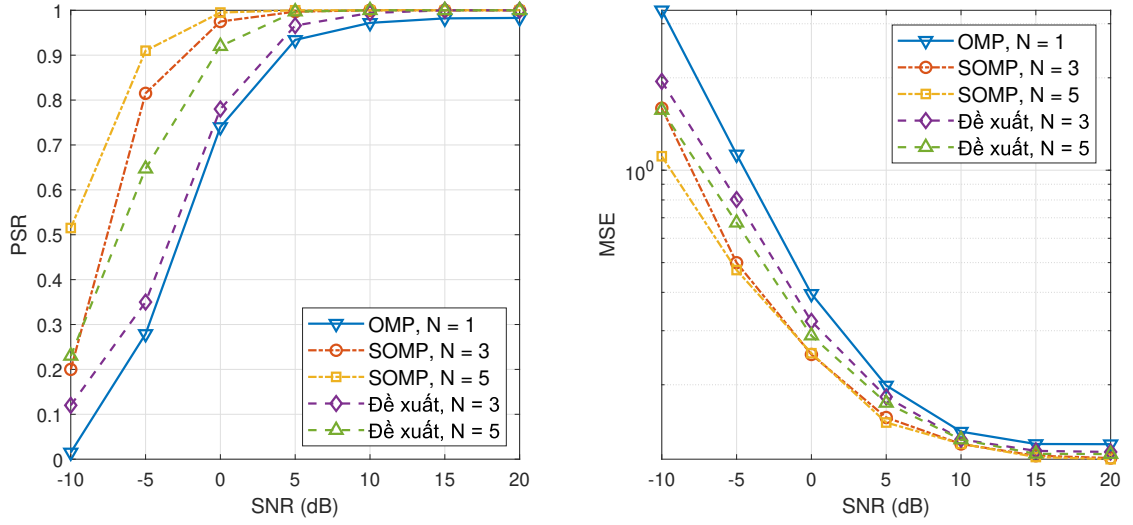


Hình 4.4: Biểu diễn trực quan của a) Kênh trong miền góc không nhiễu, b) Kênh được tái tạo lại chỉ sử dụng thuật toán DStOMP, và c) Kênh được tái tạo lại sử dụng kết hợp thuật toán DStOMP và phương pháp Silhouette sửa đổi.

DStOMP ở mức một mức SNR cao. Trong trường hợp này, những hỗ trợ không mong muốn được thêm vào tập hỗ trợ và do hệ số kênh được ước lượng bởi phương trình (3.15) nên hiệu suất tái tạo kênh tổng thể bị giảm đi nhiều, đặc biệt là về biên độ kênh. Để duy trì biên độ của hỗ trợ chính, LS chỉ nên được thực hiện trên một cột duy nhất của ma trận cảm nhận. Khi phép LS được sử dụng trong StOMP, biên độ của hỗ trợ chính được chia sẻ với các hỗ trợ khác thuộc thùy phụ. Kết quả là độ lợi kênh thực ước lượng được khác xa so với biên độ chính xác. Hình 4.4c biểu thị kênh ước lượng được trong miền góc khi sử dụng cả thuật toán DStOMP và phương pháp Silhouette sửa đổi. Nhờ khả năng xác định số lượng cụm bằng phương pháp Silhouette đề xuất, các hỗ trợ không mong muốn bị loại bỏ và ước tính cuối cùng chỉ dựa trên các hỗ trợ thuộc thùy chính.

4.4.2.2 Đánh giá hiệu năng của phương pháp đề xuất khi N thay đổi

Mô phỏng này đánh giá PSR của bốn phương pháp. Kết quả mô phỏng được biểu diễn trong Hình 4.5. Chúng ta dễ dàng thấy rằng hai phương pháp sử dụng đa sóng mang có PSR cao hơn hai phương pháp sử dụng đơn sóng mang. Điều đáng chú ý là SOMP vượt trội hơn phương pháp đề xuất ở phạm vi SNR thấp (-10 đến 0 dB). Tuy nhiên,



Hình 4.5: Hiệu năng PSR và MSE của các phương pháp khi thay đổi số sóng mang con N

cần nhớ rằng SOMP là phương pháp đòi hỏi kiến thức trước về độ thừa tín hiệu, đây là một lợi thế về mặt thông tin so với phương pháp được đề xuất. Có thể nói, khoảng cách giữa hai phương pháp được tạo ra khi Thuật toán 4 bị sai trong việc ước tính số lượng đường dẫn tồn tại trong kênh. Như thể hiện trong Hình 4.5, chúng ta có thể viết $\lim_{N \rightarrow \infty} \text{PSR} = 1$ vì nhiễu có trong $\mathbf{y}$ là nhiễu trắng và việc tăng N chỉ làm cho công suất nhiễu trong (4.32) ngày càng nhỏ hơn. Ở khoảng SNR từ -10 đến 10, MSE của các phương pháp cho thấy sự khác biệt đơn giản vì PSR của chúng khác nhau. Trong trường hợp này, MSE thấp hơn tương ứng với các giá trị PSR cao hơn. Tuy nhiên, ở khoảng SNR từ 15 đến 20, các phương pháp đa sóng mang có MSE gần như giống nhau. Điều này xảy ra vì một cột của ma trận cảm nhận trên một sóng mang con về cơ bản là một phiên bản quay của một cột khác có cùng chỉ số trên một sóng mang con khác. Sau đó, việc tăng N không làm giảm MSE vì phép chiếu của $\mathbf{y}[n]$ lên không gian con mở rộng bởi tập hỗ trợ có N khác nhau về cơ bản là giống nhau.

4.4.2.3 Đánh giá hiệu năng của phương pháp đề xuất khi G thay đổi

Mô phỏng này đánh giá MSE của hai phương pháp ở các mức SNR khác nhau khi N thay đổi. Kết quả mô phỏng được biểu diễn trong Hình 4.6. Tương tự như đánh giá

trong Chương 4.4.2.2, việc tăng G cũng làm tăng hiệu suất về mặt PSR. Điều này có thể được giải thích thông qua việc xem xét độ tương quan của cột thứ k -th trong $\mathbf{\Omega}$ với tín hiệu nhận được như sau

$$\begin{aligned} c_k = \langle \boldsymbol{\omega}_k, \mathbf{y} \rangle &= \left[(\mathbf{u}_{T,i}^H \mathbf{F} \mathbf{x})^\top \otimes \mathbf{u}_{R,j} \right]^H (\mathbf{U}_T^H \mathbf{F} \mathbf{x})^\top \otimes \mathbf{U}_R \mathbf{y} \\ &= \left[(\mathbf{u}_{T,i}^H \mathbf{F} \mathbf{x})^\top \right]^H (\mathbf{U}_T^H \mathbf{F} \mathbf{x})^\top \otimes \mathbf{u}_{R,j}^H \mathbf{U}_R \mathbf{y} \quad , \\ &= \mathbf{u}_{T,i}^\top \left[(\mathbf{F} \mathbf{x})^\top \right]^H (\mathbf{F} \mathbf{x})^\top \left[(\mathbf{U}_T)^H \right]^\top \otimes \mathbf{u}_{R,j}^H \mathbf{U}_R \mathbf{y} \end{aligned} \quad (4.32)$$

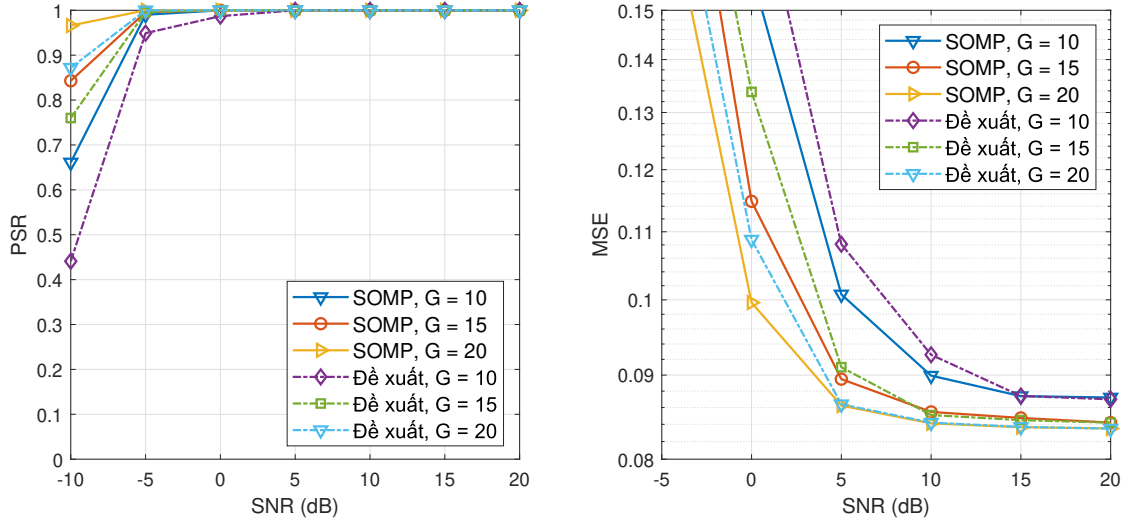
trong đó $\boldsymbol{\omega}_k$ là cột thứ k trong $\mathbf{\Omega}$ và được định nghĩa bởi

$$\boldsymbol{\omega}_k = (\mathbf{u}_{T,i}^H \mathbf{F} \mathbf{x})^\top \otimes \mathbf{u}_{R,j}, \forall i, j \in [N_b] | k = i + N_b^2(j-1), \quad (4.33)$$

Có thể thấy, bản chất của việc tăng G chính là phát tín hiệu theo nhiều góc hơn. G tiến đến vô cùng đồng nghĩa với việc phát tín hiệu ra tất cả các hướng, hay

$$\lim_{G \rightarrow \infty} \left[(\mathbf{F} \mathbf{x})^\top \right]^H (\mathbf{F} \mathbf{x})^\top = \mathbf{I}. \quad (4.34)$$

Khi $\left[(\mathbf{F} \mathbf{x})^\top \right]^H (\mathbf{F} \mathbf{x})^\top = \mathbf{I}$, các giá trị c_k sẽ tạo thành một vec-tơ là phiên bản khuếch đại của kênh ban đầu hay, $|c_k| = \alpha |h_k|$, với $\alpha \in \mathbb{R}_+$ là hệ số khuếch đại. Một khi chùm tia hướng tới chính xác vị trí của UE (bất kể qua đường dẫn LOS hay NLOS), giá trị c_k ứng với góc đó cũng tăng lên, từ đó làm tăng xác suất xác định đúng hỗ trợ. Hiệu ứng này sẽ biến mất khi số lượng chùm tia đủ lớn. Sau đó, việc tăng số lượng chùm tia chỉ làm tăng độ phức tạp của thuật toán mà không cải thiện đáng kể hiệu suất về PSR và MSE. Đáng chú ý là khi G tăng lên thì MSE giảm đi. Điều này xảy ra vì hệ số kênh được ước lượng thông qua LS. Với nhiễu trắng, một hệ thống có G lớn hơn, tức là có nhiều phép đo hơn, sẽ cho ra một ước tính LS với lỗi thấp hơn. Khi G tăng đến vô cực, MSE dự kiến sẽ không giảm đến 0 vì kênh tái tạo lại bởi các tham số kênh bị lượng tử hóa. Trong phạm vi SNR cao và với cùng G , MSE của phương pháp đề xuất và của SOMF là giống hệt nhau vì chúng có cùng cơ chế ước tính hệ số kênh, tức chiếu tín hiệu đã nhận lên không gian con được trải rộng bởi cùng một tập hỗ trợ.



Hình 4.6: Hiệu năng PSR và MSE của các phương pháp khi thay đổi số lượng các chùm tia.

4.5 Phương pháp ước lượng thô có xét đến tương quan giữa các đường dẫn

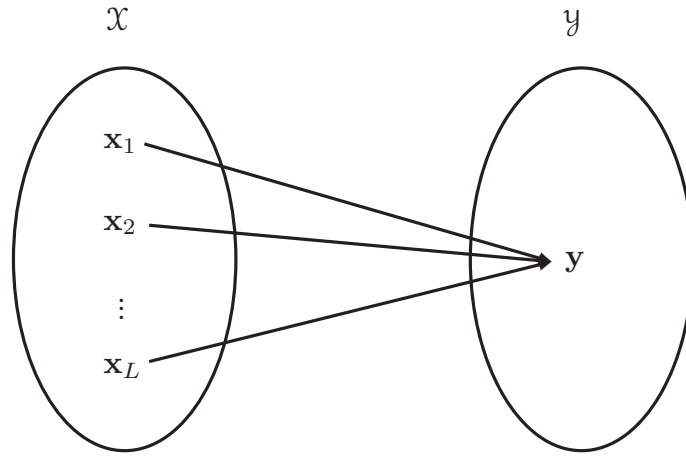
Trong chương này, luận án khai thác lỗ hổng trong giả định của [58] để đề xuất một phương pháp ước lượng kênh mới với độ chính xác cao hơn những phương pháp không xem xét lỗ hổng đó. Các tác giả trong [58] cho rằng khi số lượng ăng-ten đủ lớn thì AOA của các đường dẫn khác nhau phân biệt hoàn toàn với nhau, do đó chúng được ước lượng độc lập với nhau. Tuy nhiên, điều này là không thực tế vì số lượng ăng-ten thường là hữu hạn. Vì vậy, việc xem xét sự ảnh hưởng lẫn nhau giữa các giá trị AOA là điều cần thiết.

4.5.1 Phương pháp đề xuất

Như đã trình bày trước đó, kênh truyền trong miền góc là tổng hợp của các kênh con có độ thưa bằng một. Theo ngôn ngữ thống kê, kênh trong miền góc là dữ liệu có tính chất không đầy đủ (incomplete data). Sự không đầy đủ ngụ ý rằng tồn tại hai không gian mẫu $\mathcal{X}$ và $\mathcal{Y}$ và ánh xạ nhiều - một từ $\mathcal{X}$ sang $\mathcal{Y}$ [45]. Gọi $\mathbf{y}$ là dữ liệu thuộc

không gian $\mathcal{Y}$. Các giá trị có quan hệ trực tiếp với nó được ký hiệu là $\mathbf{x} \in \mathcal{X}$ được gọi là dữ liệu đầy đủ, thường không thể quan sát được trực tiếp mà chỉ có thể quan sát thông qua $\mathbf{y}$. Ví dụ 1 mô tả hiện tượng này.

Ví dụ 1 Giả sử $\mathbf{y} = [y_1, y_2, y_3, y_4]$ và $\mathbf{x} = [x_1, x_2, x_3, x_4, x_5]$ là các véc-tơ có các phần tử là số tự nhiên lớn hơn 0 và giữa chúng tồn tại một mối quan hệ như sau: $y_1 = x_1 + x_2$, $y_2 = x_3$, $y_3 = x_4$, $y_4 = x_5$. Với $y_1 = n$, người ta có rất nhiều sự lựa chọn cho x_1 và x_2 sao cho tổng của chúng bằng n , hay $[x_1, x_2] = \{[0, n], [1, (n-1)], \dots, [(n-1), 1], [n, 0]\}$.



Hình 4.7: Mối quan hệ giữa $\mathcal{X}$ và $\mathcal{Y}$ và các dữ liệu trong chúng

Hệ quả 1 Cho một tập hợp các số biểu diễn một véc-tơ gần thưa¹, ví dụ, $\mathbf{x}_i, i \in [N]$ và các phần tử của nó $\mathbf{x}_i$ thỏa mãn một hàm điều hòa tắt dần, ví dụ, hàm dựa trên nhân Dirichlet, nếu $\mathbf{x} = \sum_{i=1}^N \mathbf{x}_i$, $\mathbf{b} = \sum_{i=1}^N \mathbf{b}_i$ và $\mathbf{A}\mathbf{x}_i = \mathbf{b}_i$ thì

$$\langle \mathbf{a}_k, \mathbf{b} \rangle = \sum_{i=1}^N \langle \mathbf{a}_k, \mathbf{b}_i \rangle. \quad (4.35)$$

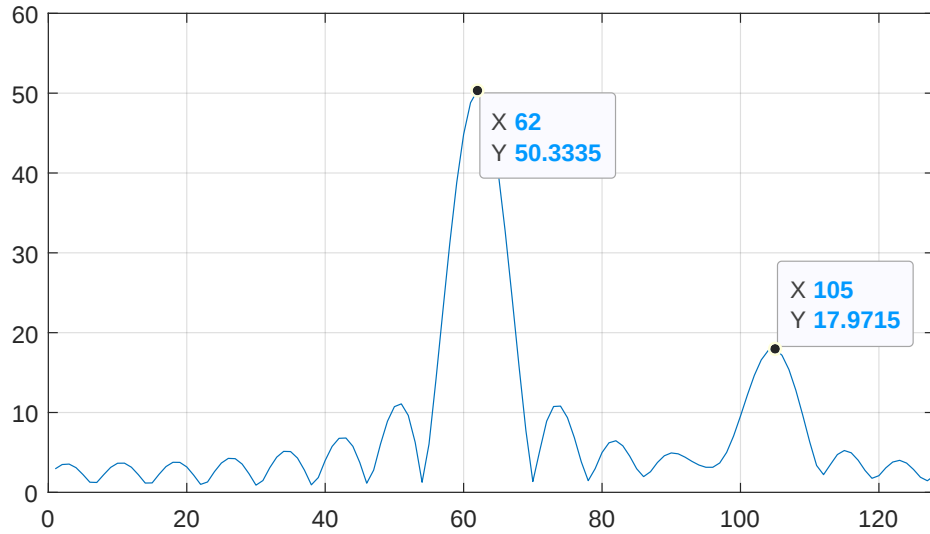
Theo Hệ quả 1, kết quả của OMP truyền thống khi tính toán độ tương quan giữa các cột trong ma trận cảm nhận và phần dư là sự chồng chất của nhiều đường dẫn. Do đó, chúng ta cần phân tách sự đóng góp của mỗi đường dẫn ra từ tín hiệu quan sát

¹theo nghĩa có một số nhỏ phần tử khác không có biên độ không đáng kể tập trung xung quanh một giá trị có biên độ lớn

được. Quá trình phân tách được thực hiện trên các giá trị tương quan như sau. Giả sử các tham số nằm chính xác trên các lưới chia, nhiệm vụ là phải xác định các chỉ số của AOA và tính một tham số biểu diễn cho tính nhất quán về không gian của AOA². Tham số này cho biết các đỉnh ảnh hưởng tới nhau như thế nào trong miền AOA và giới hạn phạm vi tìm kiếm cho bước chọn cột của thuật toán OMP đề xuất sau khi trừ đi ảnh hưởng của nó khỏi tín hiệu nhận được. Đặt $\mathcal{C}_k = |\langle \boldsymbol{\omega}_k, \mathbf{y} \rangle|$, bắt đầu từ $k = 1$, thuật toán thu thập các véc-tơ hệ số $\mathcal{C}_k$ với độ dài G_r . Sau đó, các véc-tơ được cộng dồn (theo chỉ số) để thu được một véc-tơ hệ số khác, gọi là $\mathbf{v}$, có cùng dạng mẫu như mô tả trong Hình 4.8 và được sử dụng để xác định tính nhất quán về không gian giữa các AOA

$$\mathbf{v} = \left[\sum_{j=1}^{G_t} \mathcal{C}_{1,j}, \sum_{j=1}^{G_t} \mathcal{C}_{2,j}, \dots, \sum_{j=1}^{G_t} \mathcal{C}_{G_r,j} \right]. \quad (4.36)$$

Chỉ số của các giá trị AOA là vị trí của L giá trị lớn nhất trong $\mathbf{v}$ thỏa mãn 2 điều



Hình 4.8: Độ tương quan của các cột trong $\boldsymbol{\Omega}[1]$ với tín hiệu nhận được. Biên độ của hai đỉnh $x = 62$ và $x = 105$ không phải là giá trị thực vì chúng có tương tác với nhau. Đỉnh tại $x = 62$ bị ảnh hưởng bởi đuôi của đỉnh $x = 105$ và ngược lại.

kiện: *i)* chúng phải lớn hơn ngưỡng τ và *ii)* là giá trị lớn nhất trong số 3 giá trị liên

²tính nhất quán về mặt không gian là khoảng cách không thay đổi giữa hai góc trong thời gian kết hợp

tiếp. Sau khi có thông tin về AOA từ bước trước, thuật toán tiếp theo sẽ phân tách sự chồng chất giữa các khối G_r giá trị liên tiếp thuộc $\mathcal{C}_k$. Giả sử tồn tại L đường dẫn trong kênh, mỗi đường dẫn được biểu diễn bởi một hàm Dirichlet có biên độ lớn nhất được khuếch đại một lượng $\alpha_i, i \in [L]$. Mỗi giá trị $\mathcal{C}_k, k \in [G_r]$ có thể được biểu diễn thông qua

$$\mathcal{C}_k \approx \left| \sum_{i=1}^L \alpha_i e^{j\pi\phi_{i,k} \frac{N_t-1}{2}} \frac{\sin(\frac{1}{2}N_t\phi_{i,k})}{\sin(\frac{1}{2}\pi\phi_{i,k})} \right|, \quad (4.37)$$

trong đó, $\phi_{i,k}$ là sự khác biệt về góc giữa góc thứ k và góc AOA thứ i . Vì $\phi_{i,k}$ đã biết, không mất tính tổng quát, luận án viết $\phi_{i,k} = e^{j\pi\phi_{i,k} \frac{N_t-1}{2}} \frac{\sin(\frac{1}{2}N_t\phi_{i,k})}{\sin(\frac{1}{2}\pi\phi_{i,k})}$ và đi giải quyết vấn đề có hàm giá như sau

$$\mathcal{J} = \sum_{k=1}^{G_t} \left(\mathcal{C}_k - \left| \sum_{i=1}^L \alpha_i \phi_{i,k} \right| \right)^2. \quad (4.38)$$

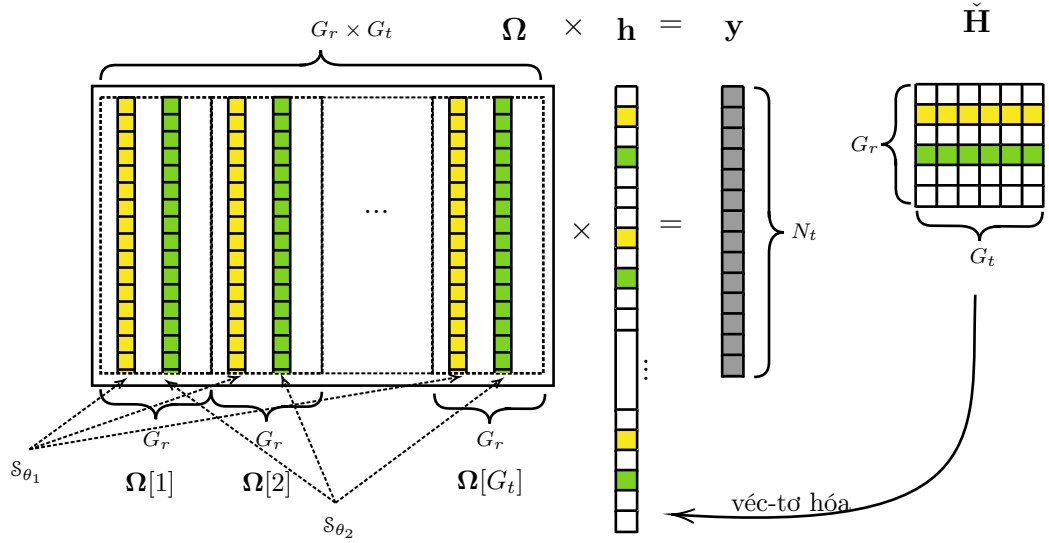
Hàm giá này có thể được giải quyết cho mỗi giá trị α_i sử dụng suy giảm theo đạo hàm (GD)

$$\alpha_i^{(t)} = \alpha_i^{(t-1)} - \eta \frac{\partial \mathcal{J}^{(t-1)}}{\partial \alpha_i}, \quad (4.39)$$

trong đó, $\alpha_i^{(t)}$, η , và $\frac{\partial \mathcal{J}^{(t-1)}}{\partial \alpha_i}$ lần lượt là độ lợi ước lượng của đường dẫn thứ i tại bước thời gian thứ t , hệ số suy giảm, và đạo hàm riêng của hàm giá theo α_i tại bước thời gian trước đó ($t-1$). Các giá trị $\alpha_i^{(0)}$ có thể được khởi tạo bằng $\mathcal{C}_{\theta_i}/N_t$ trong đó $\mathcal{C}_{\theta_i}$ là hệ số của AOA thứ i . Sau khi thực hiện phân tách cho các khối gồm G_t hệ số, chúng ta thu được độ lợi α_i cho L đường dẫn và các giá trị tương quan không bao gồm sự chồng chất, được ký hiệu nghĩa bởi $\tilde{\mathcal{C}}_k$. Sau bước này, thuật toán thực hiện bước chọn cột dựa trên hệ số đã được phân tách và cập nhật tập hỗ trợ. Giống với OMP, bước tiếp theo ước lượng hệ số kênh ứng với các cột đã tìm được ở bước trước đó sử dụng LS. Sau đó, thuật toán trừ ảnh hưởng của cột vừa tìm được và quay lại bước chọn cột.

Nhận xét:

1. Ngoại trừ bước đầu tiên, hiệu năng của việc chọn cột có thể được cải thiện nhờ thông tin của các giá trị AOA ước lượng được. Nếu coi kênh truyền là thừa theo hàng và là tổng hợp của L đường dẫn thì sẽ có L tập con được định nghĩa bởi $\mathcal{S} = [\mathcal{S}_{\theta_1}, \mathcal{S}_{\theta_2}, \dots, \mathcal{S}_{\theta_L}]$ là ứng cử viên góp phần tạo ra tín hiệu nhận được. Trong trường hợp tín hiệu là thừa- L , nếu bước chọn cột thứ k chọn được cột từ tập con



Hình 4.9: Mô tả phương pháp đề xuất. Hai tập con $\mathcal{S}_{\theta_1}$ và $\mathcal{S}_{\theta_2}$ biểu diễn hai hàng của kênh trong miền góc.

$\mathcal{S}_{\theta_i}$ thì bước thứ $k + 1$ sẽ loại trừ tất cả các cột có chung AOA ở tập đó. Cách thực hiện được cho bởi Hình 4.9. Điều này dựa trên thực tế là môi trường vật lý hiếm khi cho phép nhiều đường dẫn có cùng AOA.

2. Phương trình (4.37) chỉ là giá trị gần đúng vì luận án cho rằng chỉ các giá trị AOA tương tác với nhau. Trên thực tế, $\mathcal{C}_k$ cũng chứa ảnh hưởng lẫn nhau theo miền AOD. Tuy nhiên, việc xem xét thêm AOD bài toán sẽ trở nên phức tạp về mặt tính toán.
3. Độ chính xác của bước ước lượng AOA có thể được cải thiện bằng cách thực hiện phân tách $\mathbf{v}$ thông qua phương trình (4.39) cho một số bộ điểm xung quanh L giá trị cao nhất của véc-tơ $\mathbf{v}$. Tập hợp các giá trị AOA tạo ra chênh lệch nhỏ nhất giữa $\mathbf{v}$ và tổng các hệ số đã được tách riêng sẽ đại diện cho giá trị AOA của các góc.

Thuật toán OMP khi xem xét ảnh hưởng giữa các đường dẫn được tóm lược bởi Thuật toán 5.

Algorithm 5: Phương pháp OMP đề xuất cho mô hình các đường dẫn có ảnh hưởng lẫn nhau

Đầu vào: Tín hiệu nhận được: $\mathbf{y}$,

Ma trận cảm nhận: $\mathbf{\Omega} = [\boldsymbol{\omega}_1, \boldsymbol{\omega}_2, \dots, \boldsymbol{\omega}_{G_t G_r}]$,

Độ thưa: L và Sai số dừng: ϵ .

Khởi tạo: Biến đếm $t = 0$, tập hỗ trợ $I_s = \emptyset$, phần dư $\mathbf{r}_0 = \mathbf{y}$

1. Tính độ tương quan giữa các cột của $\mathbf{\Omega}$ và tín hiệu nhận được:

$$\mathcal{C}_k = |\langle \boldsymbol{\omega}_k, \mathbf{y} \rangle|;$$

2. Ước lượng chỉ số của AOA thông qua (4.36) với điều kiện kèm theo:

$$[\theta_1, \theta_2, \dots, \theta_L] = \max(\mathbf{v}, L);$$

3. Tạo ra tập hợp chứa các chỉ số của các hỗ trợ tiềm năng:

$$\mathcal{S} = [\mathcal{S}_{\theta_1}, \mathcal{S}_{\theta_2}, \dots, \mathcal{S}_{\theta_L}];$$

while $\|\mathbf{r}_k\|_2 \geq \epsilon$ *or* $t \leq L$ **do**

4. $t = t + 1$;

5. Tính toán độ tương quan giữa các cột của ma trận cảm nhận trong $\mathbf{\Omega}$ và phần dư: $\mathcal{C}_k = |\langle \boldsymbol{\omega}_k, \mathbf{r} \rangle|$;

6. Thực hiện phân tách các thành phần con theo mỗi cụm G_t giá trị $\mathcal{C}_k$ sử dụng (4.39) và xây dựng $\tilde{\mathcal{C}}_k$;

7. Chọn cột : $k = \arg \max_k \tilde{\mathcal{C}}_k$ với $k \in \mathcal{S}_{\theta_i}$;

8. Cập nhật danh sách hỗ trợ tiềm năng: $\mathcal{S} = \mathcal{S} \setminus \mathcal{S}_{\theta_i}$;

9. Cập nhật tập hỗ trợ $I_s = I_s \cup k$;

10. Ước lượng độ lợi kênh: $\mathbf{h}_s = (\mathbf{\Omega}_{I_s}^H \mathbf{\Omega}_{I_s})^{-1} (\mathbf{\Omega}_{I_s}^H \mathbf{y})$;

11. Trừ đóng góp của hỗ trợ tìm được ra khỏi tín hiệu: $\mathbf{r}_s = \mathbf{y} - \mathbf{\Omega}_{I_s} \mathbf{h}_s$;

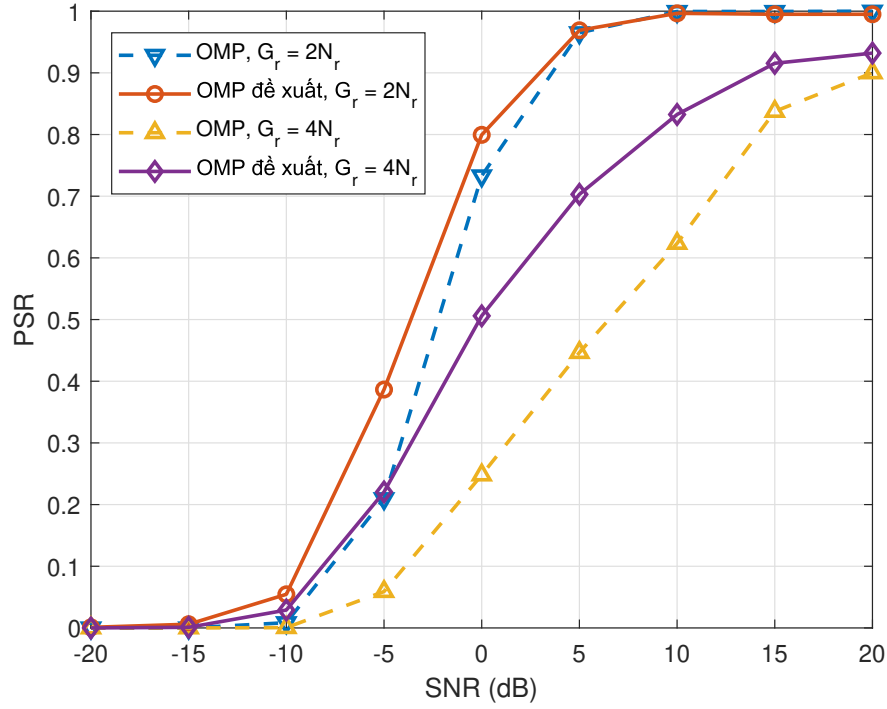
end

Đầu ra : $I_s, \mathbf{h}_s$.

4.5.2 Kết quả và thảo luận

4.5.2.1 PSR dưới điều kiện giá trị liên kết tương hỗ nhỏ

Trong mô phỏng này, luận án đánh giá khả năng xác định đúng tập hỗ trợ khi thay đổi độ phân giải của các góc ảo với mong đợi rằng việc này có thể giảm được lỗi lượng tử hóa gây ra bởi việc chia lưới. Từ Hình 4.10, chúng ta thấy có một xu hướng giảm



Hình 4.10: Biểu diễn PSR là một hàm theo SNR.

chung về PSR đối với cả hai phương pháp OMP khi tăng số lượng góc ảo. Mặc dù tăng số lượng góc ảo có thể làm giảm sai số lượng tử hóa nhưng nó gián tiếp làm tăng giá trị μ - một điều không mong muốn đối với ma trận cảm nhận trong bài toán ước lượng thưa. Hơn nữa, việc tăng số lượng góc ảo còn làm cho kích thước của bài toán trở nên lớn hơn, gây tốn nhiều tài nguyên tính toán hơn. Về xác suất xác định đúng tập hỗ trợ (PSR), phương pháp đề xuất có hiệu năng tốt hơn phương pháp OMP truyền thống. Sự cải thiện này cũng thay đổi tùy vào số lượng góc ảo và mức SNR. Trong trường hợp $G_r = 2N_r$, thuật toán OMP đề xuất có độ chính xác PSR cao hơn OMP truyền thống ở dải SNR thấp (từ -10 đến 5 dB) và cùng độ chính xác PSR với phương pháp này ở dải SNR cao (từ 5 đến 20dB). Về mặt trực quan, sự cải thiện này là không đáng

kể, nhưng nó có ý nghĩa về mặt thực nghiệm vì SNR của kênh mmWave không phải lúc nào cũng cao vì tính chất của mmWave là suy hao mạnh. Sự khác biệt giữa hai phương pháp có thể được quan sát rõ ràng hơn trong trường hợp $G_r = 4N_r$. Tùy thuộc vào mức SNR, thuật toán OMP đề xuất cho mức độ cải thiện khác nhau, từ 10% đến 25% so với OMP truyền thống.

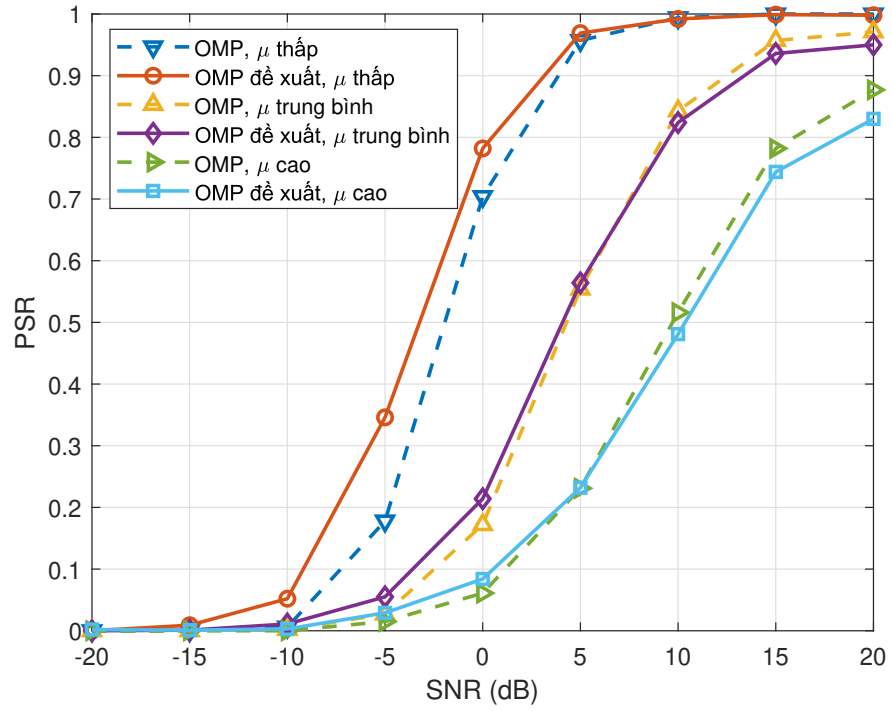
4.5.2.2 PSR dưới điều kiện thay đổi về chỉ số liên kết tương hỗ

Trong mô phỏng này, luận án đánh giá hiệu năng của phương pháp OMP đề xuất dưới ảnh hưởng của giá trị liên kết tương hỗ (μ). Người ta mong muốn rằng giá trị liên kết tương hỗ của ma trận cảm nhận càng nhỏ càng tốt vì hai cột bất kỳ trong ma trận cảm nhận khi đó sẽ tiến đến trực giao, nói cách khác, bài toán có nghiệm duy nhất. Tuy nhiên, hệ thống truyền thông có thể gặp phải tình huống trong đó ma trận cảm nhận có giá trị liên kết tương hỗ cao. Một số bài toán gặp phải tình huống này có thể kể đến như: phân bổ công suất [82], tạo búp sóng với các ràng buộc năng lượng,... Trong thí nghiệm này, giá trị liên kết tương hỗ được tăng lên bằng cách áp đặt các ràng buộc công suất lên ma trận định dạng chùm $\mathbf{F}$ bởi

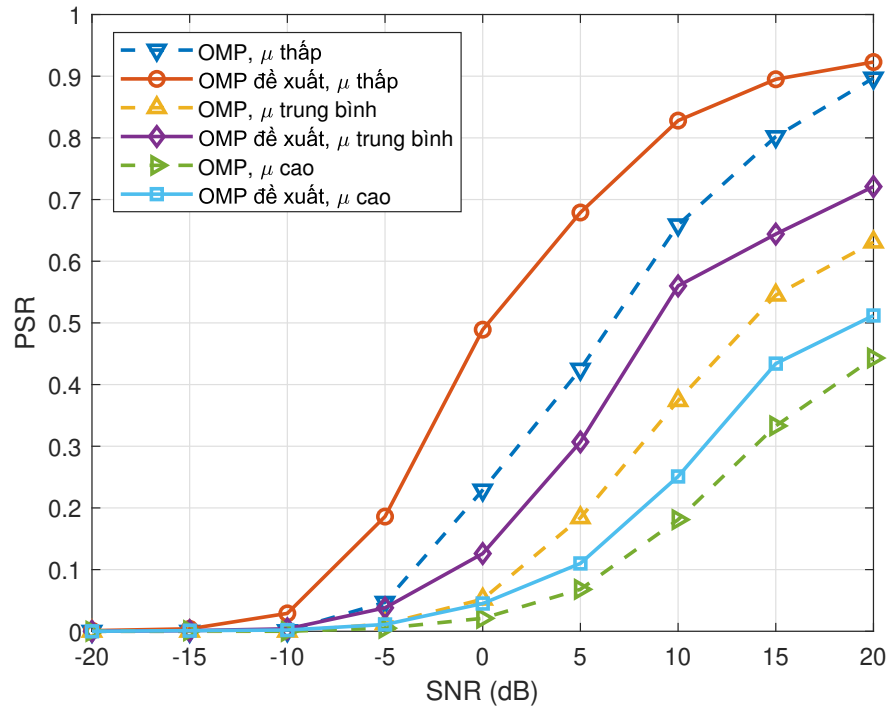
$$\mathbf{F} = \frac{\mathbf{F}}{\alpha \|\mathbf{F}\|_F}, \quad (4.40)$$

trong đó α là các hệ số giúp thay đổi công suất cho $\mathbf{F}$. Việc áp đặt hạn chế công suất cho $\mathbf{F}$ có thể dẫn tới giá trị liên kết tương hỗ cao hơn so với trường hợp bình thường, tức ma trận cảm nhận kém phù hợp hơn với các thuật toán phục hồi tín hiệu thưa như OMP hoặc LASSO. Kết quả mô phỏng trong hai trường hợp, $G_r = 2N_r$ và $G_r = 4N_r$, lần lượt được cho bởi Hình 4.11 và Hình 4.12.

Nhìn chung, độ chính xác về PSR của cả hai phương pháp đều giảm khi giá trị liên kết tương hỗ tăng. Với trường hợp $G_r = 4N_r$, độ chính xác PSR của cả hai phương pháp đều kém hơn trường hợp có $G_r = 2N_r$ vì chúng có chung ràng buộc về năng lượng, nhưng trường hợp thứ hai có số lượng góc ảo lớn hơn, hay μ lớn hơn. Trong kịch bản đầu tiên, sự cải thiện giữa phương pháp OMP đề xuất và OMP truyền thống dần thu hẹp khi μ tăng lên. Sự cải thiện về mặt PSR của phương pháp OMP đề xuất thậm chí còn biến mất ở mức μ cao nhất tại các ngưỡng SNR cao. Trong kịch bản thứ hai, mặc dù xu hướng này vẫn xảy ra, giống như kịch bản thứ nhất, khi SNR tăng lên,



Hình 4.11: PSR của các phương pháp theo SNR với $G_r = 2N_r$.



Hình 4.12: PSR của các phương pháp theo SNR với $G_r = 4N_r$.

chúng ta vẫn có thể thấy sự cải thiện của phương pháp đề xuất. Ở mức 5 dB, mức cải thiện giảm dần từ 22% xuống 10% và cuối cùng là 4% tương ứng với giá trị liên kết tương hỗ thấp, trung bình và cao. Từ các kết quả trên, luận án kết luận rằng phương pháp đề xuất mạnh hơn OMP tiêu chuẩn trong trường hợp ma trận cảm nhận có độ kết hợp lẫn nhau ở mức thấp và trung bình. Cải tiến này đặc biệt hữu ích trong các tình huống có nhiều hạn chế về đường truyền thẳng.

4.5.2.3 Đánh giá MSE của các phương pháp đề xuất

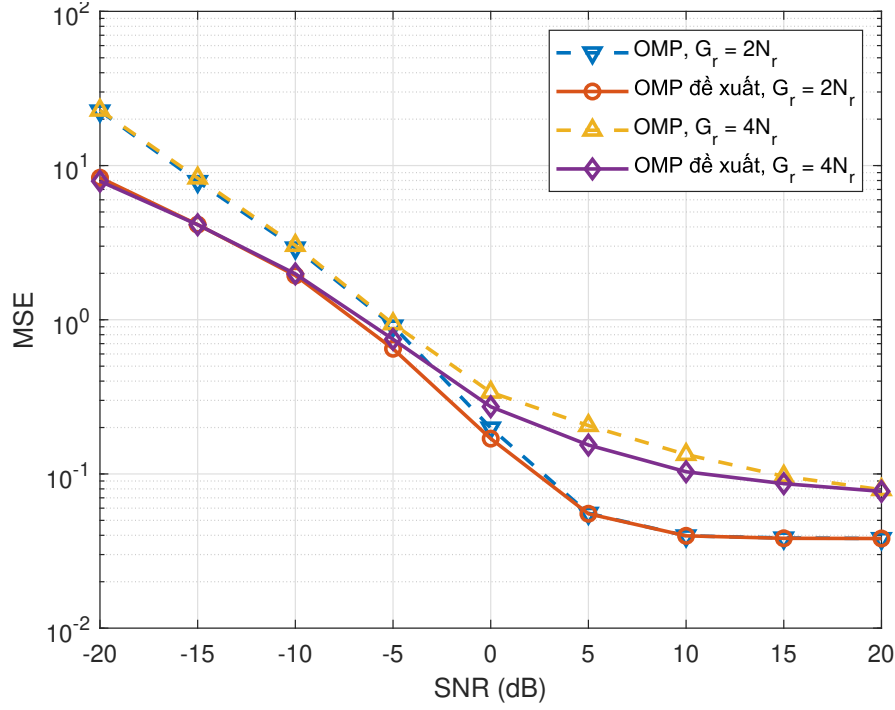
Các kết quả về MSE được biểu diễn trong Hình 4.13. Nhìn chung, việc đánh giá MSE trong hai trường hợp này có ý nghĩa tương tự như việc đánh giá PSR. Như có thể thấy, phương pháp OMP đề xuất vượt trội hơn OMP thông thường trong cả hai trường hợp của G_r . Kết quả này hoàn toàn có thể dự đoán được dựa trên PSR. PSR càng cao, đặc biệt khi nó đạt 100%, thì MSE càng thấp. Ngoài ra, việc tăng số lượng góc ảo không làm giảm MSE của cả hai phương pháp OMP vì thực tế PSR giảm khi số lượng góc ảo tăng. Tương tự như PSR, trong trường hợp $G_r = 2N_t$, phương pháp đề xuất chỉ cải thiện một chút ở mức SNR thấp, trong khi trong trường hợp $G_r = 2N_t$, nó cải thiện trên toàn dải SNR. Ngoài ra, MSE của tất cả các mô phỏng có xu hướng bão hòa khi SNR tăng trong khoảng từ 10 đến 20 dB. Điều này được giải thích bởi các tham số mà OMP ước tính chỉ là giá trị trên lưới; chúng bị hạn chế và không thể giảm ở một mức SNR nhất định.

4.6 Phương pháp ước lượng tình

4.6.1 Phương pháp đề xuất

Từ các hỗ trợ tìm được bằng các thuật toán trình bày trong phần trước, luận án có thể ước lượng các tham số AOD, AOA, TOA, và độ lợi kênh thực thông qua một vài hàm ánh xạ cơ bản [57]. Kênh tái tạo lại được từ các tham số có thể được viết dưới dạng

$$\hat{\mathbf{H}}[n] = \sum_{k=0}^{K-1} |\hat{h}_k| e^{\frac{-j2\pi n \hat{\tau}_k}{NT_s}} f(\bar{\theta}_{\text{Rx},k}, \bar{\theta}_{\text{Tx},k}) \quad (4.41)$$



Hình 4.13: MSE của các kỹ thuật ước lượng kênh theo SNR khi thay đổi số lượng góc ảo.

trong đó, $f(\bar{\theta}_{\text{Rx},k}, \bar{\theta}_{\text{Tx},k})$ được định nghĩa bởi:

$$f(\bar{\theta}_{\text{Rx},k}, \bar{\theta}_{\text{Tx},k}) \triangleq \mathbf{u}_{\text{Rx}}^H(\bar{\theta}_{\text{Rx},k}) \mathbf{a}_{\text{Rx},n}(\bar{\theta}_{\text{Rx},k}) \mathbf{a}_{\text{Tx},n}^H(\bar{\theta}_{\text{Tx},k}) \mathbf{u}_{\text{Tx}}(\bar{\theta}_{\text{Tx},k}) \quad (4.42)$$

ở đây, $\bar{\theta}_{\text{Rx},k}$ and $\bar{\theta}_{\text{Tx},k}$ là các giá trị AOA và AOD bị lượng tử hóa, lần lượt được làm tròn tới các lưới $\mathbf{u}_{\text{Rx}}^H(\bar{\theta}_{\text{Rx},k})$ và $\mathbf{u}_{\text{Tx}}^H(\bar{\theta}_{\text{Tx},k})$. Hiện tượng này được gọi là *lệch cơ sở*. Hình 4.14 (bên trái) mô tả trực quan hiện tượng này. Có thể thấy, các giá trị góc trong thực tế là một giá trị liên tục và có lẽ không có một cách chia nào có thể đảm bảo một lưới đi qua chính xác các giá trị AOA và AOD. Để tìm được các giá trị liên tục này, nhiệm vụ đầu tiên là phải xây dựng một hàm mất mát và thực hiện tối ưu trên nó. Hàm mất mát được định nghĩa bằng cách viết lại phương trình (4.12) theo dạng

$$\mathbf{y}[n] = \sum_{k=0}^{K-1} \underbrace{\Omega[n] \mathbf{h}_k[n] + \mathbf{n}_k[n]}_{\mathbf{y}_k[n]}. \quad (4.43)$$

Xếp chồng tín hiệu nhận được trên các sóng mang con, ta thu được

$$\mathbf{y} = \sum_{k=0}^{K-1} \underbrace{\tilde{\Omega} \mathbf{h}_k + \mathbf{n}_k}_{\mathbf{y}_k}, \quad (4.44)$$

trong đó

$$\check{\mathbf{\Omega}} = \text{blkdiag}\{\mathbf{\Omega}[0], \dots, \mathbf{\Omega}[N-1]\}, \quad (4.45a)$$

$$\check{\mathbf{y}} = [\check{\mathbf{y}}^\top[0], \dots, \check{\mathbf{y}}^\top[N-1]]^\top, \quad (4.45b)$$

$$\check{\mathbf{h}}_k = [\check{\mathbf{h}}_k^\top[0], \dots, \check{\mathbf{h}}_k^\top[N-1]]^\top, \quad (4.45c)$$

$$\check{\mathbf{n}}_k = [\check{\mathbf{n}}_k^\top[0], \dots, \check{\mathbf{n}}_k^\top[N-1]]^\top. \quad (4.45d)$$

Gọi $\boldsymbol{\eta} = [\boldsymbol{\eta}_0, \boldsymbol{\eta}_1, \dots, \boldsymbol{\eta}_{K-1}]^\top$ là tập hợp chứa tất cả các tham số của các đường dẫn và $\hat{\boldsymbol{\eta}}^{(i)}$ là ước lượng của $\boldsymbol{\eta}$ tại bước thứ i . Hàm mất mát theo đường dẫn thứ k có dạng

$$Q(\boldsymbol{\eta}_k | \hat{\boldsymbol{\eta}}^{(i)}) = \|\check{\mathbf{y}} - \sum_{l \neq k, l=0}^{K-1} \check{\mathbf{\Omega}} \hat{\mathbf{h}}_l^{(i)} - \check{\mathbf{\Omega}} \check{\mathbf{h}}_k\|_2^2. \quad (4.46)$$

Ở đây, thuật toán suy giảm tuần tự [70] được sử dụng để tìm được lời giải tối ưu cho bài toán trên. Với thuật toán này, chúng ta sẽ thực hiện tối ưu tham số cho một đường dẫn trong khi giữ cố định những tham số những đường dẫn khác. Trong bước này, nhiệm vụ chính là đi tìm $\boldsymbol{\eta}_k$ sao cho (4.46) đạt cực tiểu, hay

$$\begin{aligned} \hat{\boldsymbol{\eta}}_k^{(i+1)} &= \arg \min_{\boldsymbol{\eta}_k} Q(\boldsymbol{\eta}_k | \hat{\boldsymbol{\eta}}^{(i)}) \\ &= \arg \min_{\boldsymbol{\eta}_k} \|\check{\mathbf{y}} - \sum_{l \neq k, l=0}^K \check{\mathbf{\Omega}} \hat{\mathbf{h}}_l^{(i)} - \check{\mathbf{\Omega}} \check{\mathbf{h}}_k\|_2^2. \end{aligned} \quad (4.47)$$

Với bài toán trong phương trình (4.47), người ta có thể dễ dàng tối ưu nó thông qua phương pháp Newton (xem Phụ lục II) và suy giảm tuần tự. Mặc dù phương pháp Newton sử dụng đến thông tin đạo hàm bậc hai và được kỳ vọng sẽ đem lại sự hội tụ nhanh cho tất cả các tham số nhưng điều này không nhất thiết đúng trong thực tế. Đối với một hàm nhiều biến như (4.47), sự hội tụ nhanh đánh đổi về chi phí tính toán khi mà biểu thức liên quan tới đạo hàm bậc hai có độ dài đáng kể. Ngoài ra, việc tìm ra các hệ số suy giảm cho tất cả các tham số cũng là điều phi thực tế của phương pháp này. Do hàm giá chứa yếu tố ngẫu nhiên (nhiều và ma trận beamforming), các tập hợp hệ số suy giảm mà tại đó sự hội tụ nhanh xảy ra sẽ thay đổi trong những tình huống khác nhau. Nếu người ta sử dụng chung một tập hợp các hệ số suy giảm cho tất cả các tình huống, sự hội tụ đôi khi không xảy ra hoặc xảy ra với số vòng lặp rất lớn. Vì vậy, luận án lựa chọn phương pháp tìm kiếm theo tỉ lệ vàng - một phương pháp tối ưu

số học, để thay thế cho phương pháp Newton. Phương pháp này luôn đảm bảo sự hội tụ xảy ra với số vòng lặp hữu hạn. Sự hội tụ đến giá trị toàn cục được đảm bảo nếu hàm mất mát là hàm lồi và điểm khởi tạo nằm gần giá trị toàn cục. Nếu chỉ sử dụng thuật toán tìm kiếm theo tỉ lệ vàng, quá trình tối ưu các tham số sẽ được thực hiện như trong Thuật toán 6. Với cách thực hiện này, về bản chất chúng ta vẫn phải thực hiện tối ưu trên tất cả các tham số, dẫn đến tốc độ tính toán của phương pháp chưa được cải thiện nhiều. Vì vậy, luận án đề xuất một phương pháp khác, tiêu thụ ít năng lượng tính toán hơn theo lý luận sau đây. Từ phương trình (4.47), chúng ta thấy rằng $\tilde{\mathbf{\Omega}}$ với các cột của nó chính là cơ sở chứa thông tin về AOA và AOD. Nếu không quá quan tâm tới ước lượng của độ lợi kênh và TOA, chúng ta có thể cập nhật các giá trị AOA và AOD sử dụng thuật toán tìm kiếm theo tỉ lệ vàng và ước lượng hệ số kênh cho các cặp AOA-AOD tìm được thông qua LS. Theo cách này, LS đảm bảo (4.47) được tối thiểu trong khi AOA và AOD dần tiếp cận tới các giá trị chính xác của chúng. Tuy nhiên, việc nghịch đảo ma trận lớn $\tilde{\mathbf{\Omega}}$ yêu cầu rất nhiều tài nguyên tính toán. Kết quả là, sử dụng LS để ước tính trực tiếp hệ số kênh $\hat{\mathbf{h}}_k$ là không thực tế. Nhận thấy toàn bộ cơ sở hầu như không thay đổi trong quá trình ước lượng ngoại trừ cơ sở được tạo bởi giá trị AOA và AOD đang được tinh chỉnh nên luận án chỉ sử dụng AOA và AOD tìm được ở mỗi bước để xây dựng cơ sở mới như sau

$$\tilde{\boldsymbol{\omega}}^{(i)}[n] \triangleq \left(\mathbf{u}_{\text{Tx}}^H(\hat{\theta}_{\text{Tx},k}^{(i)}) \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] \right)^T \otimes \mathbf{u}_{\text{Rx}}(\hat{\theta}_{\text{Rx},k}^{(i)}). \quad (4.48)$$

$\mathbf{u}_{\text{Rx}}(\hat{\theta}_{\text{Rx},k}^{(i)})$ được định nghĩa bởi

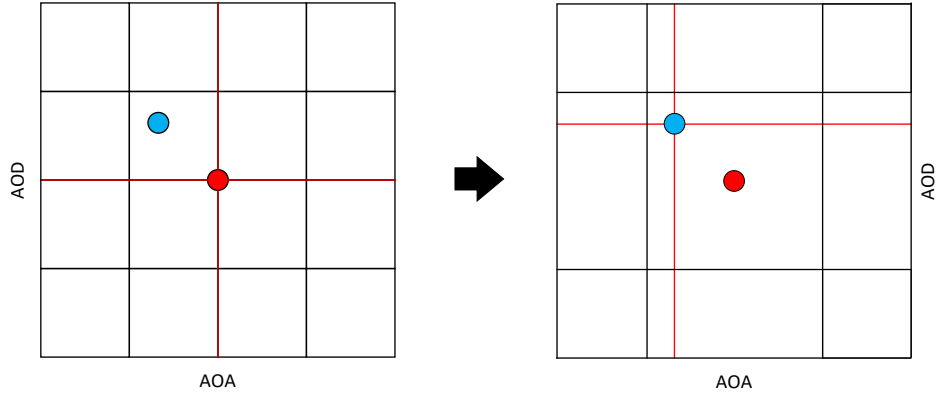
$$\mathbf{u}_{\text{Rx}}(\hat{\theta}_{\text{Rx},k}^{(i)}) = \left[1, e^{-j\pi \sin(\hat{\theta}_{\text{Rx},k}^{(i)})}, \dots, e^{-j\pi(N-1) \sin(\hat{\theta}_{\text{Rx},k}^{(i)})} \right]^T \quad (4.49)$$

$\mathbf{u}_{\text{Tx}}(\hat{\theta}_{\text{Tx},k}^{(i)})$ được định nghĩa tương tự bằng cách thay thế Rx bởi Tx. Theo cách này, thuật toán đề xuất giảm được kích thước bài toán từ $N_r GN \times N_b N_b N$ xuống còn $N_r GN \times N$ và dễ dàng ước lượng hệ số kênh, $\hat{\mathbf{h}}_k^{(i)} = [\hat{h}_k^{(i)}[0], \hat{h}_k^{(i)}[1], \dots, \hat{h}_k^{(i)}[N-1]]^T$, thông qua LS

$$\hat{\mathbf{h}}_k^{(i)} = \left((\boldsymbol{\omega}^{(i)})^H \boldsymbol{\omega}^{(i)} \right)^{-1} (\tilde{\boldsymbol{\omega}}^{(i)})^H \left(\tilde{\mathbf{y}} - \sum_{l \neq k, l=0}^K \tilde{\boldsymbol{\omega}} \hat{\mathbf{h}}_l^{(i)} \right). \quad (4.50)$$

Ở mỗi vòng lặp, thuật toán đề xuất cập nhật lại ma trận cảm nhận bằng cách thay thế cột tìm được ở bước khởi tạo bằng cột có AOA và AOD được hiệu chỉnh và ước lượng

lại kênh trực tiếp bằng thuật toán DCS-SOMP [64]. Với cách này, chỉ số của các hỗ trợ tìm được bởi thuật toán DCS-SOMP ở bước đầu tiên sẽ giống với chỉ số tìm được ở bước cuối cùng. Điểm khác biệt nằm ở chỗ cơ sở ở bước khởi tạo chịu ảnh hưởng của hiệu ứng lệch cơ sở còn cơ sở ở bước tinh chỉnh cuối cùng thì không. Hình 4.14 mô tả cách hoạt động của thuật toán đề xuất. Trong Hình 4.14, chấm tròn màu đỏ và màu xanh lần lượt biểu diễn giá trị lượng tử hóa và giá trị chính xác của AOD và AOA. Các đường màu đỏ biểu diễn các cột là hàm của AOD và AOA ước lượng được. Các đường màu đen là các cột không thay đổi trong quá trình ước lượng tinh.



Hình 4.14: Cách hoạt động của phương pháp tinh chỉnh đề xuất.

Cuối cùng, các thông tin về TOA và độ lợi kênh thực được trích xuất thông qua các hệ số kênh ước lượng được. Vì hệ số kênh là các số phức, có thể viết dưới dạng $z = |z|e^{j\angle(z)}$ nên ước lượng về độ lợi kênh thực được tính bằng giá trị trung bình của các giá trị mô-đun của hệ số kênh phức

$$|\hat{h}_k| = \frac{1}{N} |\hat{\mathbf{h}}_k| \quad (4.51)$$

và ước lượng τ_k thông qua sự khác biệt về pha của các hệ số kênh phức trên các sóng mang con khác nhau

$$\hat{\tau}_k = -\frac{1}{N-1} \left(\angle \left(\hat{\mathbf{h}}_k \right) \right) \times \frac{NT_s}{2\pi} \quad (4.52)$$

Mã giả của phương pháp tinh chỉnh đề xuất được cho bởi Thuật toán 7.

Algorithm 6: Thuật toán tinh chỉnh tham số thô sử dụng tìm kiếm vàng và suy giảm tuần tự cho 2 đường dẫn (1 LOS và 1 NLOS), mỗi đường dẫn gồm 4 tham số

Đầu vào: Chỉ số của đường dẫn: $k = 0 \rightarrow 1$; Chỉ số của tham số: $p = 1 \rightarrow 4$;

Giá trị thô khởi tạo:

$$\hat{\boldsymbol{\eta}}^{(0)} = [\hat{\boldsymbol{\eta}}_0^{(0)}, \hat{\boldsymbol{\eta}}_1^{(0)}]^\top = [\hat{\boldsymbol{\eta}}_{0,p}^{(0)}, \hat{\boldsymbol{\eta}}_{1,p}^{(0)}]^\top = \underbrace{[\hat{\theta}_{\text{Rx},0}^{(0)}, \hat{\theta}_{\text{Tx},0}^{(0)}, \hat{\tau}_0^{(0)}, \hat{h}_0^{(0)}]^\top}_{\text{LOS}}, \underbrace{[\hat{\theta}_{\text{Rx},1}^{(0)}, \hat{\theta}_{\text{Tx},1}^{(0)}, \hat{\tau}_1^{(0)}, \hat{h}_1^{(0)}]^\top}_{\text{NLOS}};$$

Tín hiệu nhận được: $\check{\mathbf{y}}$; Ma trận cảm nhận: $\check{\mathbf{\Omega}}$;

Sai số dừng: ε_p ; Khoảng tìm kiếm: γ_p ;

Tỉ lệ vàng: $\varphi = \frac{\sqrt{5}-1}{2}$; Số vòng lặp tối đa: M ;

Đầu ra : Tham số được tinh chỉnh: $\hat{\boldsymbol{\eta}}^{(M)}$

for $i = 1 \rightarrow M$ **do**

for $k = 0 \rightarrow 1$ **do**

for $p = 1 \rightarrow 4$ **do**

$$\hat{\eta}_{k,p,1}^{(i)} = \hat{\eta}_{k,p}^{(i)} - \gamma_p; \hat{\eta}_{k,p,2}^{(i)} = \hat{\eta}_{k,p}^{(i)} + \gamma_p; \Delta_p = |\hat{\eta}_{k,p,1}^{(i)} - \hat{\eta}_{k,p,2}^{(i)}|;$$

while $\Delta_p > \varepsilon_p$ **do**

$$d_k = \varphi \times (\hat{\eta}_{k,p,2}^{(i)} - \hat{\eta}_{k,p,1}^{(i)});$$

$$x_{k,p,1} = \hat{\eta}_{k,p,1}^{(i)} + d_p; x_{k,p,2} = \hat{\eta}_{k,p,2}^{(i)} - d_p;$$

$$\Delta_p = x_{k,p,1} - x_{k,p,2};$$

if $Q_{k,p}(x_{k,p,1}) > Q_{k,p}(x_{k,p,2})$ **then**

$$| \quad \hat{\eta}_{k,p,2}^{(i)} = x_{k,p,1};$$

else

$$| \quad \hat{\eta}_{k,p,1}^{(i)} = x_{k,p,2};$$

end

end

$$\hat{\eta}_{k,p}^{(i+1)} = 0.5 \times (x_{k,p,1} + x_{k,p,2});$$

end

end

end

Algorithm 7: Thuật toán ước lượng tinh đề xuất

Đầu vào: Tín hiệu nhận được: $\mathbf{y}$; Ma trận cảm nhận: $\tilde{\mathbf{\Omega}}$; Số vòng lặp tối đa:

M , Độ thưa của kênh: K ; Các tham số thô: $\hat{\boldsymbol{\eta}}^{(0)}$; Kênh thô: $\hat{\mathbf{h}}_k^{(0)}$

Đầu ra : $\hat{\boldsymbol{\eta}}^{(M+1)}$

$\hat{\mathbf{H}}^{(0)} = \text{reshape}\left(\hat{\mathbf{h}}^{(0)}, [N_b \times N_b]\right);$

for $k = 0 \rightarrow K - 1$ **do**

$[\mathbf{r}_k, \mathbf{c}_k] = \text{find}(|\hat{\mathbf{H}}^{(0)}| \neq 0);$

end

// Tinh chỉnh sử dụng phương pháp tìm kiếm dựa trên tỉ lệ vàng;

for $i = 1, \dots, M$ **do**

 Cập nhật tuần tự giá trị: $\mathbf{y} - \sum_{l \neq k, l=0}^K \tilde{\mathbf{\Omega}} \hat{\mathbf{h}}_l^{(i)};$

for $k = 0 \rightarrow K - 1$ **do**

$\hat{\theta}_{\text{Tx},k}^{(i)} = \text{GOLDEN SEARCH}(\hat{\theta}_{\text{Tx},k}^{(i-1)}, \tilde{\mathbf{\Omega}});$

$\hat{\theta}_{\text{Rx},k}^{(i)} = \text{GOLDEN SEARCH}(\hat{\theta}_{\text{Rx},k}^{(i-1)}, \tilde{\mathbf{\Omega}});$

 Xây dựng cơ sở mới sử dụng (4.48);

 Đánh giá (4.50) để tối thiểu (4.46);

$|\hat{h}_k^{(i)}| = \frac{1}{N} |\hat{\mathbf{h}}_k^{(i)}|; \hat{\tau}_k^{(i)} = -\frac{1}{N-1} \left(\angle \left(\hat{\mathbf{h}}_k^{(i)} \right) \right) \times \frac{NT_s}{2\pi};$

end

end

// Cập nhật lại ma trận cảm nhận với cơ sở mới;

for $k = 0 \rightarrow K - 1$ **do**

$\mathbf{U}_{\text{Tx}}(:, \mathbf{c}_k) = \mathbf{u}_{\text{Tx}}\left(\hat{\theta}_{\text{Tx},k}^{(M)}\right); \mathbf{U}_{\text{Rx}}(:, \mathbf{r}_k) = \mathbf{u}_{\text{Rx}}\left(\hat{\theta}_{\text{Rx},k}^{(M)}\right);$

end

$\mathbf{\Omega}[n] = \left(\mathbf{U}_{\text{Tx}}^H \mathbf{F}^{(g)}[n] \mathbf{x}^{(g)}[n] \right)^T \otimes \mathbf{U}_{\text{Rx}}; \tilde{\mathbf{\Omega}} = \text{blkdiag}\{\mathbf{\Omega}[0], \dots, \mathbf{\Omega}[N-1]\};$

// Ước lượng lại kênh với ma trận cảm nhận mới;

$\hat{\mathbf{h}}^{(M+1)} = \text{DCS-SOMP}(\mathbf{y}, \tilde{\mathbf{\Omega}}, K);$

for $k = 0, \dots, K - 1$ **do**

$|\hat{h}_k^{(M+1)}| = \frac{1}{N} |\hat{\mathbf{h}}_k^{(M+1)}|; \hat{\tau}_k^{(M+1)} = -\frac{1}{N} \left(\angle \left(\hat{\mathbf{h}}_k^{(M+1)} \right) \right) \times \frac{NT_s}{2\pi};$

end

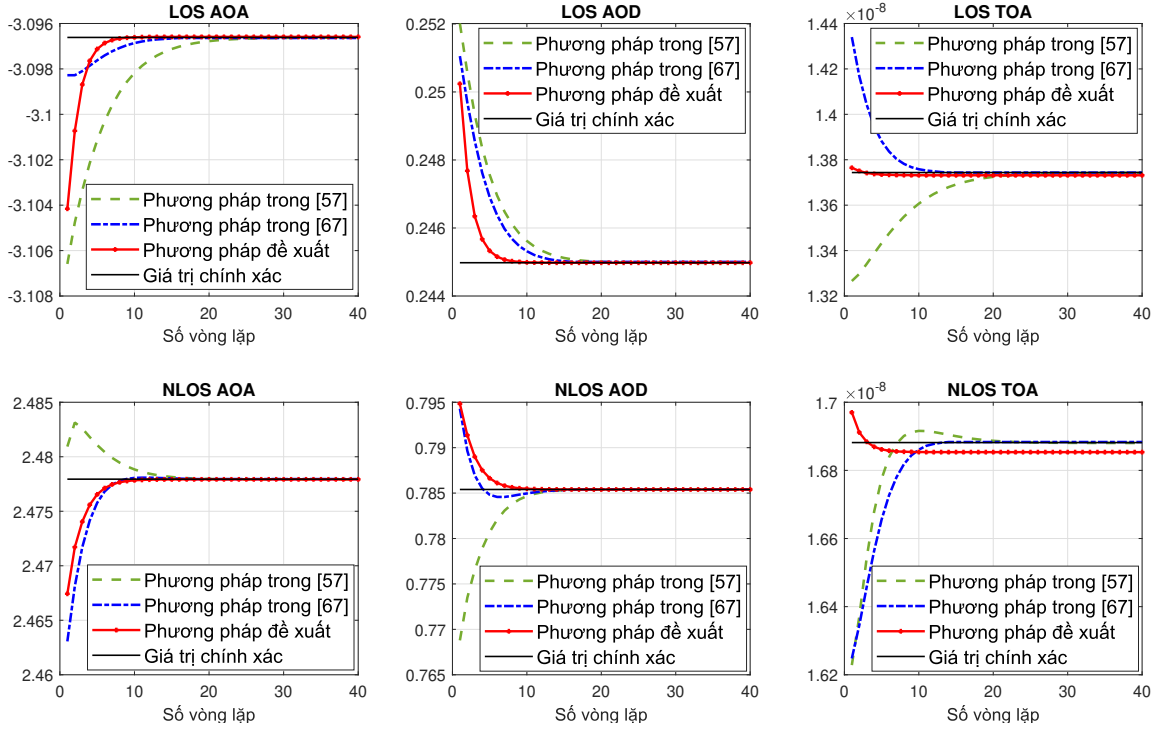
4.6.2 Kết quả và thảo luận

Trong chương này, luận án sẽ sử dụng mô phỏng để đánh giá hiệu suất của phương pháp đề xuất. Luận án xem xét kênh truyền giữa UE và BS trong không gian 2 chiều, bao gồm 1 đường dẫn LOS và 1 đường dẫn NLOS. Hệ thống sử dụng 10 sóng mang con với tần số trung tâm bằng 60 GHz và băng thông bằng 100 MHz. Số lượng ăng-ten ở cả máy thu và máy phát đều là 32. Vị trí của BS được coi là biết trước với tọa độ bằng $\mathbf{q} = [0, 0]^\top$. Vị trí của điểm phản xạ và UE là ẩn số, lần lượt nằm ở tọa độ $\mathbf{p} = [4, 0]^\top$ và $\mathbf{s} = [2, 2]^\top$. Để đơn giản, số liệu thống kê kênh cho đường dẫn LOS và đường dẫn NLOS được coi là giống nhau và độ suy hao chỉ phụ thuộc vào khoảng cách. Phương pháp sử dụng thuật toán Newton cho bước tinh chỉnh được chọn làm phương pháp tiêu chuẩn³. Độ chính xác và độ phức tạp thuật toán là các tiêu chí dùng để đánh giá hiệu suất của phương pháp đề xuất với phương pháp tiêu chuẩn.

4.6.2.1 Đánh giá về độ chính xác và khả năng hội tụ của các phương pháp

Hình 4.15 cho thấy khả năng hội tụ của các phương pháp. Chúng ta dễ dàng thấy rằng tất cả các tham số, bất kể phương pháp nào, đều hội tụ về một giá trị nhất định. Cả ba phương pháp đều cho khả năng hội tụ chính xác đối với tham số AOA và AOD. Trong khi đó, TOA của phương pháp đề xuất không hội tụ về giá trị chính xác của nó. Lý do là TOA và độ lợi kênh chỉ được ước tính dựa trên cơ sở không đầy đủ, tức ước lượng chỉ dựa trên $(\hat{\omega}[n])$. Đây chính là điểm khác biệt, giúp thuật toán đề xuất có độ phức tạp thấp hơn các thuật toán sử dụng cơ sở đầy đủ. Luận án làm được việc này vì kênh truyền trong miền góc được coi là thưa nên cơ sở tương ứng với giá trị khác không của ma trận kênh có thể dùng để xấp xỉ cho toàn bộ hệ thống. Do đó, chỉ có giá trị AOA và AOD là hội tụ tới giá trị chính xác của chúng. Khoảng thời gian hội tụ ổn định của phương pháp đề xuất là sau 6 đến 8 lần lặp. Trong khi đó, các phương pháp chỉ sử dụng tìm kiếm theo tỷ lệ vàng và phương pháp Newton đều hội tụ với tốc độ lần lượt là 10 và 20 lần lặp. Hình 4.16 và 4.17 thể hiện sai số về vị trí và góc của cả ba phương pháp khi SNR thay đổi. Như chúng ta có thể thấy, độ chính xác của cả hai

³Phương pháp này đã được chứng minh trong [57] là có thể đạt được cận dưới CRLB ở các mức SNR cao

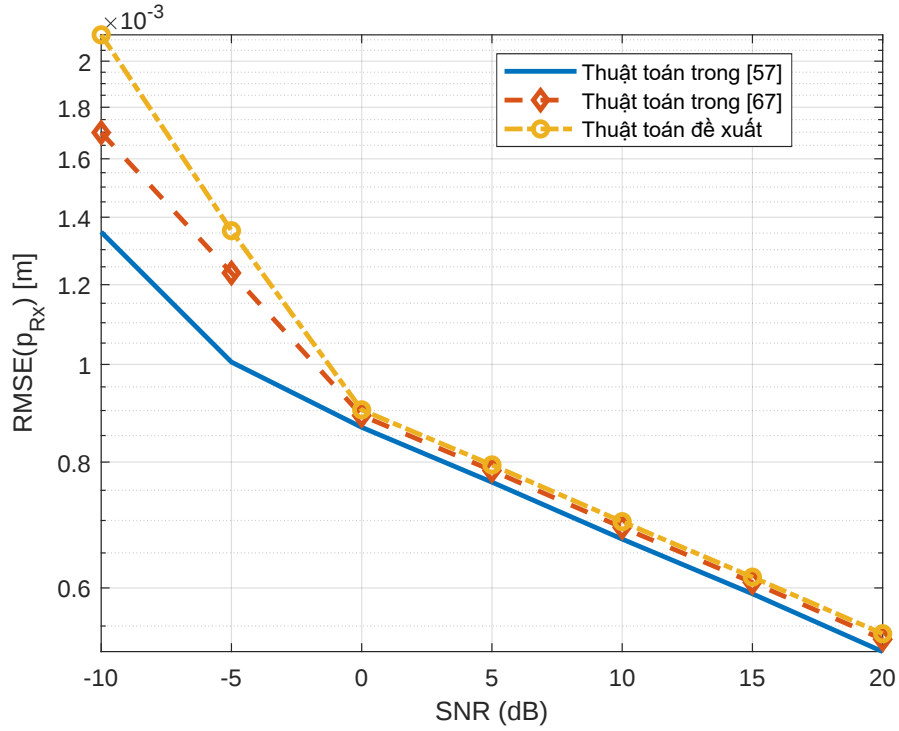


Hình 4.15: Quá trình hội tụ của các tham số ở mức SNR cao

chỉ số đều tăng theo SNR. Hai phương pháp sử dụng tìm kiếm theo tỉ lệ vàng có hiệu suất tương tự nhau và gần tương đương với độ chính xác của phương pháp Newton.

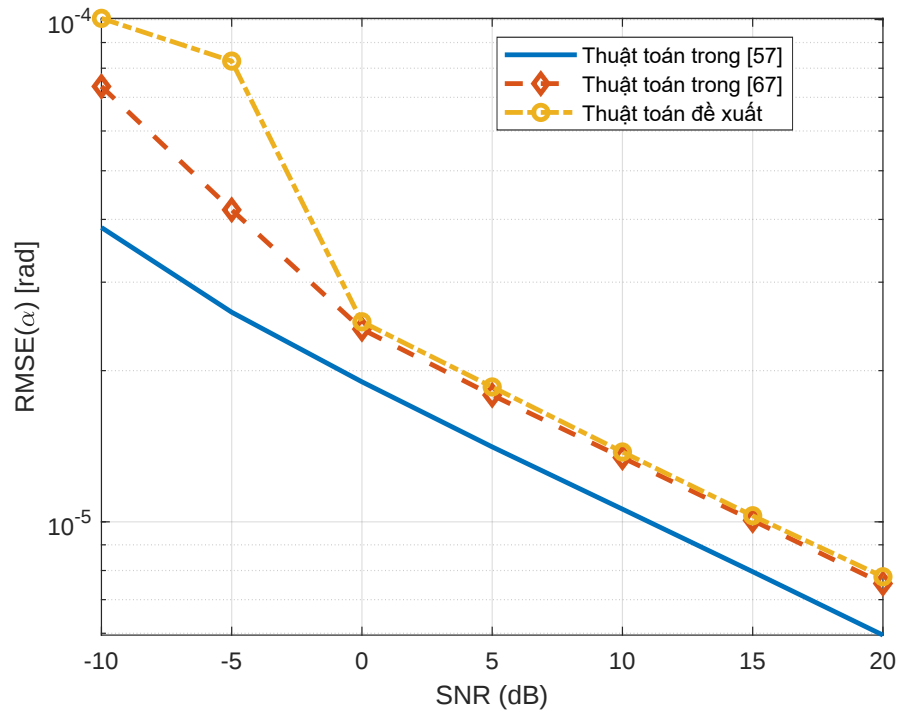
4.6.2.2 Đánh giá độ phức tạp thuật toán

Kết quả mô phỏng cho thấy độ phức tạp của thuật toán DCS-SOMP là không đáng kể so với các thuật toán tinh chỉnh (2s so sánh với 20s). Độ chính xác của các phương pháp hoàn toàn phụ thuộc vào cách mà thuật toán tinh chỉnh thực hiện. Để đạt tới giá trị hội tụ, thuật toán đề xuất cần tới M_1 vòng lặp cho K đường dẫn, trong đó mỗi đường dẫn chứa 2 cặp giá trị AOA và AOD. Trong mỗi vòng lặp, thuật toán thực hiện tối thiểu hàm giá sử dụng tìm kiếm theo tỉ lệ vàng một lần. Với ϵ là sai số dừng của thuật toán GS, thuật toán GS có độ phức tạp là $\mathcal{O}(\log(\epsilon^{-1}))$. Mỗi bước của thuật toán GS thực hiện đánh giá 2 lần giá trị $\tilde{\mathbf{\Omega}}$ với kích thước $N_r G N \times N_b N_b N$. Vì vậy, độ phức tạp thuật toán của bước tinh chỉnh là $\mathcal{O}(2M_1 K (N_r G N \times N_b N_b N) \times \log(\epsilon^{-1}))$. Nếu chỉ sử dụng phương pháp Newton, bước tinh chỉnh sẽ có độ phức tạp là $\mathcal{O}(6M_2 K (N_r G N \times N_b N_b N))$, trong đó M_2 là số vòng lặp mà thuật toán Newton cần đạt được để các tham số cùng hội tụ.

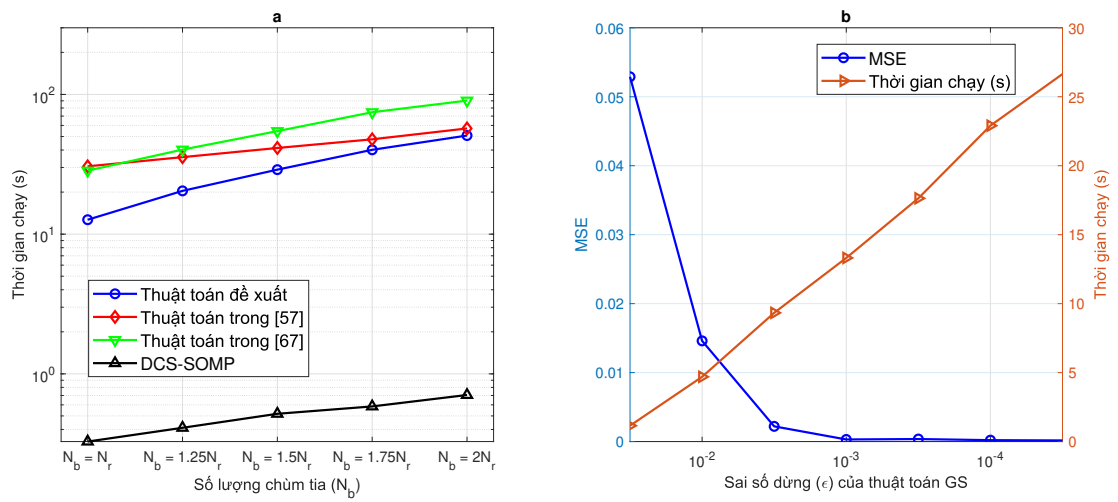


Hình 4.16: RMSE cho sai số vị trí theo thang dB của thuật toán đề xuất so với thuật toán truyền thống.

Để có cái nhìn rõ hơn về độ phức tạp của thuật toán, luận án đo thời gian chạy của các phương thức trong hai trường hợp: *i)* thay đổi N_b trong khi giữ nguyên dung sai của thuật toán GS, và *ii)* thay đổi dung sai từ 10^{-7} tới 10^{-1} trong khi giữ nguyên N_b và thực hiện thí nghiệm ở một mức SNR cao. Kết quả so sánh của các thí nghiệm được cho trong Hình 4.18. Như có thể thấy trong Hình 4.18a, phương pháp đề xuất tiêu thụ ít năng lượng tính toán hơn các phương pháp trình bày trong [57] và [67]. Điều này có thể giải thích bởi phương pháp đề xuất chỉ phải thực hiện thuật toán GS cho hai tham số thay vì bốn tham số như phương pháp Newton trong [57] hay chỉ dùng GS như trong [67]. Bằng chứng là, phương pháp đề xuất chỉ cần $M_2 = 6 \rightarrow 8$ vòng lặp trong khi phương pháp Newton cổ điển cần đến $M_1 = 20$ vòng lặp. Từ Hình 4.18b, chúng ta dễ dàng nhận thấy rằng đánh đổi giữa độ chính xác và thời gian tính toán của phương pháp đề xuất. Nó thể hiện rằng khi ϵ giảm, độ chính xác của phương pháp đề xuất tăng lên nhưng cái giá phải trả là thời gian tính toán. Xét về mặt thực tiễn, phương pháp đề xuất tốt hơn phương pháp Newton vì nó sử dụng các phương pháp



Hình 4.17: RMSE cho sai số về góc quay ăng-ten theo thang dB của thuật toán đề xuất so với thuật toán truyền thống.



Hình 4.18: a) Thời gian chạy là một hàm của N_b , b) Sai số bình phương trung bình (MSE) và thời gian chạy của thuật toán như một hàm của sai số dừng trong thuật toán tìm kiếm theo tỉ lệ vàng.

tối ưu số học và thời gian tính toán nhanh hơn.

4.7 Kết luận chương

Chương này trình bày các phương pháp ước lượng tham số định vị cho hệ thống định vị ngoài trời sử dụng trạm gốc di động 5G sử dụng công nghệ MIMO mmWave. Luận án bắt đầu với đặc điểm và tính chất của kênh truyền mmWave. Tiếp theo, luận án tập trung nghiên cứu các thuật toán ước lượng thưa họ OMP với mục đích ước lượng kênh mmWave có tính thưa trong miền góc. Từ cơ sở của các phương pháp OMP cơ bản, luận án phát triển và đề xuất 03 đóng góp chính là các phương pháp ước lượng các tham số kênh hỗ trợ khả năng định vị cho hệ thống MIMO sử dụng sóng mmWave như sau

1. Luận án mở rộng thuật toán StOMP thành thuật toán StOMP phân tán. Thuật toán này phù hợp với hệ thống sử dụng đa sóng mang và kênh truyền có sự thưa giống nhau trên các sóng mang khác nhau. Việc mở rộng này giúp thuật toán StOMP trở nên trợ với nhiều hơn so với thuật toán StOMP áp dụng cho một sóng mang duy nhất. Từ đó nâng cao được độ chính xác ước lượng thô.
2. Nhận thấy các thuật toán OMP hiện có không xem xét sự ảnh hưởng của các thành phần khác không đến nhau, luận án dựa trên đặc điểm của kênh để đề xuất một thuật toán giải quyết vấn đề này.
3. Nhận thấy các phương pháp ước lượng họ OMP là các phương pháp ước lượng dựa trên các lưới chia được xác định trước, tức giá trị có được từ chúng chỉ là các giá trị bị lượng tử hóa (có tính chất thô), luận án đề xuất một phương pháp ước lượng tinh để giải quyết vấn đề này. Ý tưởng đằng sau phương pháp đề xuất là phương pháp suy giảm tuần tự (coordinate descent) với tìm kiếm Golden-Section và SOMP. Luận án tiến thêm một bước nữa bằng cách đề xuất một phương pháp giảm độ phức tạp của các vấn đề ước tính từ tìm kiếm bốn chiều xuống gần như hai chiều. Kết quả mô phỏng cho thấy phương pháp đề xuất cung cấp hiệu suất gần như tương đương so với thuật toán Newton về độ chính xác với chi phí tính toán thấp hơn.

Chương 5

Kết luận và công việc trong tương lai

5.1 Kết luận

Luận án đã trình bày và phân tích các phương pháp định vị không dây cho cả môi trường trong nhà và môi trường ngoài trời với mục tiêu hướng đến một giải pháp định vị liên mạch giữa hai môi trường này. Sự cần thiết của nghiên cứu xuất phát từ thực tế rằng tín hiệu GPS, vốn là nền tảng của định vị ngoài trời, thường bị suy giảm đáng kể trong môi trường đô thị và môi trường trong nhà. Điều này đặt ra thách thức lớn trong việc cung cấp dịch vụ định vị với độ chính xác cao và liên tục. Luận án đã đề xuất các thuật toán định vị hoặc hỗ trợ định vị cho các hệ thống định vị cục bộ. Từ các thuật toán cơ bản, luận án đã phát triển chúng theo hướng có thể giải quyết các vấn đề còn tồn tại trong các hệ thống định vị. Các thuật toán đề xuất, qua mô phỏng hoặc thí nghiệm thực tế, đều chứng minh được tính hiệu quả của chúng. Các thuật toán này đều có tính thực nghiệm cao, giúp hiện thực hóa định vị liên mạch và là nền tảng để phát triển các dịch vụ dựa trên vị trí như dẫn đường, theo dõi sức khỏe, giao thông thông minh, tiếp thị,...

1. **Chiến lược tối ưu hóa số lượng đèn hiệu:** Luận án đã đề xuất một chiến lược triển khai đèn hiệu BLE dựa trên cấu trúc hình học (tam giác đều) để giảm thiểu số lượng đèn hiệu cần thiết mà vẫn đảm bảo độ phủ sóng tốt trong môi

trường định vị. Kết quả thực nghiệm cho thấy chiến lược này giảm được khoảng 30% số lượng đèn hiệu so với các phương pháp thông thường mà vẫn đạt được độ chính xác định vị dưới 1.5 mét.

2. **Hệ thống định vị kết hợp học máy và thuật toán WLS:** Một hệ thống định vị sử dụng mạng nơ-ron nhân tạo đã được phát triển để chia khu vực định vị thành các ô lớn, từ đó cung cấp các thông tin thô về vị trí. Hệ thống sau đó áp dụng thuật toán WLS cải tiến (PP-WLS) để tinh chỉnh vị trí ước lượng bằng cách khai thác các đèn hiệu gần nhất. Kết quả thử nghiệm cho thấy: *i)* Sai số định vị trung bình giảm 20% so với các phương pháp dựa trên khoảng cách. *ii)* Độ ổn định của hệ thống cao hơn, với sai số định vị trong 90% trường hợp nằm dưới 1.2 mét.
3. **Thuật toán ước lượng kênh thô cho định vị không xem xét sự tương tác giữa các đường dẫn:** Luận án đã phát triển một thuật toán ước lượng kênh thô dựa trên phương pháp StOMP, gọi là DStOMP. Phương pháp này khai thác cấu trúc thưa của kênh truyền mmWave trên nhiều sóng mang con để ước lượng các tham số TOA, AOA và AOD, ngay cả khi số lượng đường dẫn không được biết trước. Các thử nghiệm cho thấy: DStOMP có khả năng xác định chính xác số lượng đường dẫn và các tham số liên quan với độ sai lệch thấp trong các môi trường NLOS. Thuật toán hoạt động hiệu quả với các dữ liệu thu thập từ nhiều sóng mang con, cải thiện khả năng ước lượng tham số so với OMP truyền thống.
4. **Thuật toán ước lượng kênh thô cho định vị có xem xét sự tương tác giữa các đường dẫn:** Phương pháp cải tiến dựa trên thuật toán OMP đã được phát triển để xem xét sự tương tác giữa các đường dẫn trong miền góc của kênh truyền. Bằng cách khai thác cấu trúc Dirichlet của kênh mmWave, thuật toán có thể giảm đáng kể sai số ước lượng so với các phương pháp không xét tương tác đường dẫn.
5. **Thuật toán ước lượng kênh tinh cho định vị:** Phương pháp tinh chỉnh tham số được phát triển dựa trên tìm kiếm theo tỉ lệ vàng, giúp giảm thiểu hiệu ứng

ngoài lưới của tham số góc. Kết hợp với thuật toán LS, phương pháp này cho độ chính xác tương đối cao xong giảm được chi phí tính toán đi đáng kể.

Các thuật toán được đề xuất trong luận án không chỉ giải quyết các vấn đề trong định vị không dây mà còn cung cấp một nền tảng thực nghiệm vững chắc cho các nghiên cứu tiếp theo. Nó hứa hẹn mang lại nhiều ứng dụng thực tiễn trong các lĩnh vực như giao thông thông minh, quản lý đô thị, và dịch vụ dựa trên vị trí. Về mặt khoa học, định vị là một trong những bài toán đa ngành, thúc đẩy nghiên cứu liên quan đến xử lý tín hiệu, mạng truyền thông, và trí tuệ nhân tạo. Về mặt xã hội, cải thiện chất lượng cuộc sống, tối ưu hóa các quy trình, và tăng khả năng kết nối giữa con người và công nghệ.

5.2 Công việc trong tương lai

1. Đối với RSS: Khi các thuật toán định vị đã được nghiên cứu rất kỹ lưỡng, quy hoạch beacon là một trong những công việc còn lại giúp hệ thống định vị đạt độ chính xác cao đồng thời giảm được chi phí triển khai. Do mối quan hệ phức tạp giữa vị trí của các beacon và chất lượng của thuật toán định vị nên việc phân tích hiệu năng chung của hệ định vị dựa trên cấu hình không gian của các beacon cho một thuật toán cụ thể là một vấn đề khó khăn. Vì vậy, các kỹ thuật học máy sẽ trở nên hữu ích trong những tình huống này.
2. Đối với các phương pháp ước lượng kênh thưa:
 - Các bài toán ước lượng kênh họ OMP được coi là những thuật toán tối ưu không lỗi. Mặc dù đơn giản nhưng chúng có độ phức tạp tương đối cao khi số lượng phần tử ở mỗi bên thu và phát tăng lên. Để giải quyết vấn đề này, tác giả cần xem xét thêm các phương pháp ước lượng thưa dành riêng cho các bài toán cho chiều dữ liệu lớn như [51]. Như đã trình bày ở phía trên, tác giả cho rằng bài toán ước lượng kênh thưa có những đặc điểm mà không bài toán nào có, ví dụ cấu trúc kênh đặc biệt. Lợi dụng những đặc điểm đó, tác giả có thể kết hợp hoặc cải tiến các phương pháp có sẵn để đề xuất những phương pháp có độ hiệu quả hơn.

- Các tác giả trong [69] đã chứng minh rằng chiến lược chia lưới không đồng đều sẽ tạo ra một ma trận cảm nhận có giá trị liên kết tương hỗ nhỏ hơn chiến lược chia lưới đồng đều. Vì ma trận cảm nhận có giá trị liên kết tương hỗ nhỏ luôn là đích đến của bài toán ước lượng thưa, tác giả sẽ xem xét vấn đề thiết kế ma trận cảm nhận trong bài toán ước lượng kênh thưa trong tương lai. Các phương pháp thiết kế có thể kể đến như bổ đề Johnson–Lindenstrauss, quá trình Gram-Schmidt, phân tích QR,...
- Mạng thế hệ tương lai hứa hẹn sẽ tích hợp chức năng cảm nhận (radar sensing) và truyền thông (communication) và cùng một hệ thống, gọi là ISAC. ISAC có khả năng cảm nhận môi trường xung quanh một cách chính xác, từ đó cung cấp thông tin không chỉ về vị trí mà còn về trạng thái của môi trường. Điều này có thể hỗ trợ các ứng dụng như quản lý giao thông, giám sát an ninh, và các ứng dụng IoT tiên tiến. Trong tương lai, luận án sẽ xem xét bài toán này để cải thiện chất lượng của dịch vụ định vị.

Phụ lục

Phụ lục I

Xem xét vấn đề đặt 3 đèn hiệu BLE sao cho chúng thỏa mãn điều kiện (3.12), chúng ta mong muốn rằng có thể sử dụng LS để xác định vị trí của mục tiêu, vì vậy việc đặt 3 đèn hiệu BLE trên một đường thẳng là vô lý. Loại trừ trường hợp 3 đèn hiệu BLE nằm trên một đường thẳng, vị trí tương đối mà 3 đèn hiệu BLE có thể tạo ra là: tam giác nhọn, tam giác vuông và tam giác tù. Chúng ta có thể dễ dàng nhận thấy rằng nếu đặt 3 đèn hiệu BLE thành dạng tam giác tù và tam giác vuông, diện tích tạo ra bởi 3 đường tròn có tâm là đỉnh của tam giác nhỏ hơn khi đặt chúng thành dạng tam giác nhọn. Vì vậy, chúng tôi chỉ tìm kiếm giải pháp vị trí tối ưu cho dạng tam giác nhọn. Mô tả trực quan được cho bởi Hình 5.1. Tối đa hóa diện tích bao phủ đáng tin cậy có nghĩa là giảm thiểu tổng diện tích chồng lấp của mỗi 2 đèn hiệu BLE (hình thấu kính, được biểu diễn bởi phần màu xanh trong Hình 5.1). Đặt $\alpha = \angle EAD$, $\beta = \angle DBG$, $\gamma = \angle DCF$ và $AD = BD = CD = r$. Chúng ta dễ dàng nhận thấy rằng các thấu kính đều đối xứng và tổng diện tích của chúng (S_{lens}) bằng

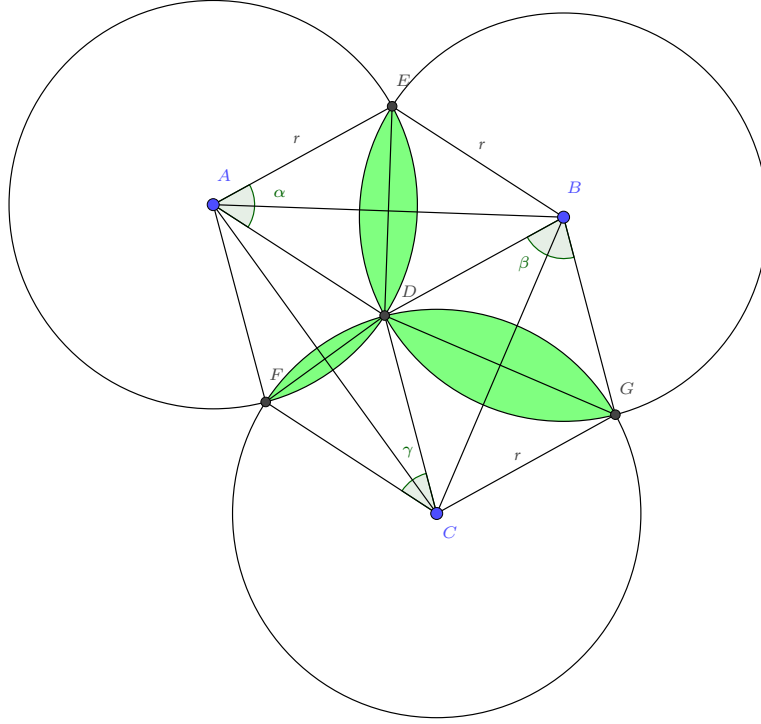
$$S_{\text{lens}} = r^2\alpha - r^2\sin(\alpha) + r^2\beta - r^2\sin(\beta) + r^2\gamma - r^2\sin(\gamma). \quad (5.1)$$

Hơn nữa, $2\angle ADE + 2\angle BDG + 2\angle ADF = 360^\circ$, suy ra

$$2\left(90 - \frac{\alpha}{2}\right) + 2\left(90 - \frac{\beta}{2}\right) + 2\left(90 - \frac{\gamma}{2}\right) = 360^\circ. \quad (5.2)$$

Cuối cùng, $\alpha + \beta + \gamma = 180^\circ$ và (5.1) có thể viết thành

$$S_{\text{lens}} \propto r^2 (180 - (\sin(\alpha) + \sin(\beta) + \sin(\gamma))). \quad (5.3)$$



Hình 5.1: Biểu diễn hình học của vấn đề tối ưu vị trí đặt đèn hiệu BLE.

Nếu chúng ta muốn tối thiểu hóa S_{lens} , chúng ta phải tối đa hóa thành phần $\sin(\alpha) + \sin(\beta) + \sin(\gamma)$. Hơn nữa, vì A , B , và C tạo thành một tam giác nhọn nên α , β , γ đều không lớn hơn 90° . Khi đó, chúng ta có một bài toán tối ưu như sau

$$\max \quad \mathcal{X} = \sin(\alpha) + \sin(\beta) + \sin(\gamma) \quad (5.4a)$$

$$\text{s.t.} \quad \alpha + \beta + \gamma = 180^\circ \quad (5.4b)$$

$$0^\circ < \alpha, \beta, \gamma < 90^\circ \quad (5.4c)$$

Bài toán này tương đương với bài toán tìm giá trị lớn nhất của $\mathcal{X}$, trong đó α, β, γ là các góc trong một tam giác nhọn. Bắt đầu với (5.4b), với $\sin(\gamma) = \sin(\alpha + \beta) = \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha)$, ta có

$$\begin{aligned} \mathcal{X} &= \sin(\alpha) + \sin(\beta) + \sin(\alpha)\cos(\beta) + \sin(\beta)\cos(\alpha) \\ &= \frac{2}{\sqrt{3}} \left(\frac{\sqrt{3}}{2} \sin(\alpha) + \frac{\sqrt{3}}{2} \sin(\beta) \right) + \sqrt{3} \left(\frac{\sin(\alpha)}{\sqrt{3}} \cos(\beta) + \frac{\sin(\beta)}{\sqrt{3}} \cos(\alpha) \right). \end{aligned} \quad (5.5)$$

Lại có $x^2 + y^2 \geq 2xy, \forall x, y \in \mathbb{R}$, ta thu được

$$\begin{aligned}\mathcal{X} &\leq \frac{1}{\sqrt{3}} \left(\sin^2(\alpha) + \frac{3}{4} + \sin^2(\beta) + \frac{3}{4} \right) + \frac{\sqrt{3}}{2} \left(\frac{\sin^2(\alpha)}{3} + \cos^2(\beta) + \frac{\sin^2(\beta)}{3} + \cos^2(\alpha) \right) \\ &= \frac{\sqrt{3}}{2} (\sin^2(\alpha) + \sin^2(\beta) + \cos^2(\alpha) + \cos^2(\beta)) + \frac{\sqrt{3}}{2} \\ &= \frac{3\sqrt{3}}{2}.\end{aligned}\tag{5.6}$$

Vì vậy, giá trị lớn nhất của $\mathcal{X}$ là $\frac{3\sqrt{3}}{2}$. Dấu bằng xảy ra khi $\sin(\alpha) = \sin(\beta) = \sin(\gamma) = \frac{\sqrt{3}}{2}$ và do đó $\alpha = \beta = \gamma = 60^\circ$, và $\angle CAB = 0.5\alpha + 0.5\gamma = 60^\circ$, $\angle CBA = 0.5\alpha + 0.5\gamma = 60^\circ$, $\angle ABC = 0.5\beta + 0.5\alpha = 60^\circ$, và cuối cùng $\triangle ABC$ là tam giác đều (đpcm).

Phụ lục 2: Phương pháp ước lượng tinh sử dụng thuật toán Newton

Đặt $\mathbf{b} = \tilde{\mathbf{y}} - \sum_{l \neq k, l=0}^K \check{\mathbf{\Omega}} \hat{\mathbf{h}}_l^{(m)}$, phương trình (4.46) có dạng một bài toán tối ưu lồi bậc hai $Q = \|\mathbf{b} - \check{\mathbf{\Omega}} \check{\mathbf{h}}_k\|_2^2$, trong đó, $\mathbf{b}$ và $\check{\mathbf{\Omega}}$ là những thành phần biết trước. Nhiệm vụ là đi tìm lời giải tối ưu cho $\check{\mathbf{h}}_k$ - một véc-tơ có độ thưa bằng 1. Tuy nhiên, trong bài toán này, chúng ta không làm việc trực tiếp với $\check{\mathbf{h}}_k$ mà các tham số tạo nên nó, bao gồm $\theta_{\text{Tx}}, \theta_{\text{Rx}}, \tau$ và $\tilde{h}$. Bài toán tối ưu sử dụng phương pháp Newton có thể viết dưới dạng:

$$x_i(t+1) = x_i(t) - \frac{\nabla_i Q(\hat{x}(i, t))}{\nabla_{ii}^2 Q(\hat{x}(i, t))} \tag{5.7}$$

trong đó, $\hat{x}(i, t) = (x_1(t+1), \dots, x_{i-1}(t+1), x_i(t), \dots, x_n(t))$ là giá trị của x được cập nhật tới thời điểm t . Vì tối ưu đồng thời cho tất cả các tham số trong Q cùng một lúc dẫn tới sự tăng nhanh về độ phức tạp thuật toán, chúng tôi sử dụng đạo hàm theo từng biến đơn và sử dụng phương pháp suy giảm tuần tự (coordinate descent) để tìm ra lời giải tối ưu chúng. Ví dụ, nếu τ được cập nhật đầu tiên, chúng tôi sẽ đi tìm lời giải tối ưu cho nó thông qua

$$\tau_k^{(i+1)} = \tau_k^{(i)} - \frac{\partial Q}{\partial \tau_k} \left(\frac{\partial^2 Q}{\partial \tau_k^2} \right)^{-1}. \tag{5.8}$$

trong đó

$$\frac{\partial Q}{\partial \tau_k} = -2(\mathbf{b} - \check{\mathbf{\Omega}} \check{\mathbf{h}}_k)^H \check{\mathbf{\Omega}} \frac{\partial \check{\mathbf{h}}_k}{\partial \tau_k} \tag{5.9}$$

và

$$\frac{\partial^2 Q}{\partial \tau_k^2} = 2(\tilde{\mathbf{\Omega}} \frac{\partial \tilde{\mathbf{h}}_k}{\partial \tau_k})^H (\tilde{\mathbf{\Omega}} \frac{\partial \tilde{\mathbf{h}}_k}{\partial \tau_k}) - 2(\mathbf{b} - \tilde{\mathbf{\Omega}} \tilde{\mathbf{h}}_k)^H \tilde{\mathbf{\Omega}} \frac{\partial^2 \tilde{\mathbf{h}}_k}{\partial \tau_k^2} \quad (5.10)$$

lần lượt là đạo hàm bậc nhất và đạo hàm bậc hai theo τ . Ở đây, $\tilde{\mathbf{h}}_k = \text{vec}(\tilde{\mathbf{H}}_k[n])$, và $\frac{\partial \tilde{\mathbf{h}}_k}{\partial \tau_k}$ được tính bởi:

$$\frac{\partial \tilde{\mathbf{h}}_k}{\partial \tau_k} = \text{vec} \left(\mathbf{U}_{\text{Rx}}^H \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},k}) \tilde{h}_k \frac{\partial \exp\left(\frac{-j2\pi n \tau_k}{NT_s}\right)}{\partial \tau_k} \mathbf{a}_{\text{Tx},n}^H(\theta_{\text{Tx},k}) \mathbf{U}_{\text{Tx}} \right) \quad (5.11)$$

và $\frac{\partial^2 \tilde{\mathbf{h}}_k}{\partial \tau_k^2}$ được tính bởi

$$\frac{\partial^2 \tilde{\mathbf{h}}_k}{\partial \tau_k^2} = \text{vec} \left(\mathbf{U}_{\text{Rx}}^H \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx},k}) \tilde{h}_k \frac{\partial^2 \exp\left(\frac{-j2\pi n \tau_k}{NT_s}\right)}{\partial \tau_k^2} \mathbf{a}_{\text{Tx},n}^H(\theta_{\text{Tx},k}) \mathbf{U}_{\text{Tx}} \right) \quad (5.12)$$

Sau khi cập nhật τ theo (5.8), chúng tôi cập nhật những tham số khác, bao gồm θ_{Tx} , θ_{Rx} , $\tilde{h}$ và giữ nguyên các tham số khác như đầu vào của hàm của một hàm nhiều biến:

$$\theta_{\text{Tx},k}^{(i+1)} = \theta_{\text{Tx},k}^{(i)} - \frac{\partial Q}{\partial \theta_{\text{Tx},k}} \left(\frac{\partial^2 Q}{\partial \theta_{\text{Tx},k}^2} \right)^{-1} \quad (5.13)$$

$$\theta_{\text{Rx},k}^{(i+1)} = \theta_{\text{Rx},k}^{(i)} - \frac{\partial Q}{\partial \theta_{\text{Rx},k}} \left(\frac{\partial^2 Q}{\partial \theta_{\text{Rx},k}^2} \right)^{-1} \quad (5.14)$$

$$\tilde{h}_k^{(i+1)} = \tilde{h}_k^{(i)} - \frac{\partial Q}{\partial \tilde{h}_k} \left(\frac{\partial^2 Q}{\partial \tilde{h}_k^2} \right)^{-1} \quad (5.15)$$

Đạo hàm bậc nhất và bậc hai của các tham số được cho bởi

- Đạo hàm theo TOA:

$$\frac{\partial \exp\left(\frac{-j2\pi n \tau}{NT_s}\right)}{\partial \tau} = -j \frac{2\pi n}{NT_s} \exp\left(\frac{-j2\pi n \tau}{NT_s}\right) \quad (5.16)$$

$$\frac{\partial^2 \exp\left(\frac{-j2\pi n \tau}{NT_s}\right)}{\partial \tau^2} = - \left(\frac{2\pi n}{NT_s} \right)^2 \exp\left(\frac{-j2\pi n \tau}{NT_s}\right) \quad (5.17)$$

- Đạo hàm theo AOA:

$$\frac{\partial \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx}})}{\partial \theta_{\text{Rx}}} = -j\pi n \cos(\theta_{\text{Rx}}) \exp(-j\pi n \sin(\theta_{\text{Rx}})) \quad (5.18)$$

$$\begin{aligned} \frac{\partial^2 \mathbf{a}_{\text{Rx},n}(\theta_{\text{Rx}})}{\partial \theta_{\text{Rx}}^2} &= j\pi n \cos(\theta_{\text{Rx}}) \exp(-j\pi n \sin(\theta_{\text{Rx}})) \\ &\quad \times (j\pi n \cos^2(\theta_{\text{Rx}}) + \sin(\theta_{\text{Rx}})) \end{aligned} \quad (5.19)$$

- Đạo hàm theo AOD được định nghĩa giống với đạo hàm theo AOA bằng cách chuyển Rx thành Tx.
- Đạo hàm theo đội lợi kênh thực:

$$\frac{\partial |\tilde{h}|}{\partial |h|} = 1 \quad (5.20)$$

Danh mục công trình khoa học của tác giả liên quan tới luận án

1. [CT1] Thai-Mai Dinh-Thi, Ngoc-Son Duong, Quoc-Tuan Nguyen, “Developing a Novel Real-Time Indoor Positioning System Based on BLE Beacons and Smart-phone Sensors," *IEEE Sensors Journal*, vol. 21, no. 20, pp. 23055-23068, Oct. 2021, doi: 10.1109/JSEN.2021.3106019. (ISI/SCIE-Q1)
2. [CT2] Ngoc-Son Duong, Thanh-Phuc Nguyen, Quoc-Tuan Nguyen, Thai-Mai Dinh Thi, “Improving indoor positioning system using weighted linear least square and neural network," *International Journal of Sensor Networks*, vol. 41, no. 2, pp. 67-77, Mar. 2023, doi: 10.1504/IJSNET.2023.129632. (ISI/SCIE-Q3)
3. [CT3] Ngoc-Son Duong, Lan-Nhi Vu Thi, Phuong-Dung Chu Thi, Sinh-Cong Lam, Thai-Mai Dinh Thi “A novel Distributed Stagewise Orthogonal Matching Pursuit algorithm for mmWave MIMO channel estimation," *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, 2024, doi: 10.1587/transfun.2024EAL2064. (ISI/SCIE-Q4)
4. [CT4] Ngoc-Son Duong, Quoc-Tuan Nguyen, Thai-Mai Dinh-Thi, “OMP-Based Channel Estimation With Dynamic Grid for mmWave MIMO Positioning Systems," *IEEE Communications Letters*, vol. 27, no. 10, pp. 2623-2627, Oct. 2023, doi: 10.1109/LCOMM.2023.3303453. (ISI/SCIE-Q1)
5. [CT5] Ngoc-Son Duong, Thu-Trang Nguyen, Lan-Nhi Vu Thi, Sinh-Cong Lam, Thai-Mai Dinh-Thi, “An improved OMP-based mmWave channel estimation for

- MIMO systems," *The 3rd International Conference on Advances in Information and Communication Technology*, 2024, doi: 10.1007/978-3-031-80943-9_97
6. [CT6] Ngoc-Son Duong, Quoc-Tuan Nguyen, Khac-Hoang Ngo, Thai-Mai Dinh-Thi, "Sparse Bayesian Learning with Atom Refinement for mmWave MIMO Channel Estimation," in *Proc. IEEE Statistical Signal Processing Workshop (SSP)*, 2023, doi: 10.1109/SSP53291.2023.10208044.
 7. [CT7] Ngoc-Son Duong, Quoc-Tuan Nguyen, Thai-Mai Dinh-Thi, "mmWave Channel Estimation for Location-based Application in 5G MIMO systems," in *Proc. 2022 IEEE Region 10 Conference (TENCON)*, Hong Kong, 2022, pp. 1-6, doi: 10.1109/TENCON55691.2022.9977818.
 8. [CT8] Ngoc-Son Duong, Ngoc-Thuy Nguyen, Phuong-Dung Chu-Thi, Quoc-Tuan Nguyen and Thai-Mai Dinh-Thi, "Cải thiện độ chính xác của hệ thống định vị trong nhà dựa trên phân tích lỗi truyền lan", *Hội thảo Quốc gia lần thứ XXIV về Điện tử, Truyền thông và Công nghệ Thông tin*, Hanoi, 2021

Tài liệu tham khảo

- [1] S. Naghdi and K. O’Keefe, “Trilateration With BLE RSSI Accounting for Pathloss Due to Human Obstacles,” *2019 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Pisa, Italy, 2019, pp. 1-8.
- [2] H. White, “A Heteroskedasticity-Consistent Covariance Matrix Estimator and a Direct Test for Heteroskedasticity”, *Econometrica*, vol. 48, no. 4, p. 817, 1980.
- [3] W. Zhao, S. Han, R. Hu, W. Meng and Z. Jia, “Crowdsourcing and Multisource Fusion-Based Fingerprint Sensing in Smartphone Localization”, *IEEE Sensors Journal*, vol. 18, no. 8, pp. 3236-3247, 2018.
- [4] Chung-Hao Huang, Lun-Hui Lee, C. Ho, Lang-Long Wu and Zu-Hao Lai, “Real-Time RFID Indoor Positioning System Based on Kalman-Filter Drift Removal and Heron-Bilateration Location Estimation”, *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 3, pp. 728-739, 2015.
- [5] Q. Tian, K. Wang and Z. Salcic, “A Low-Cost INS and UWB Fusion Pedestrian Tracking System”, *IEEE Sensors Journal*, vol. 19, no. 10, pp. 3733-3740, 2019.
- [6] Z. Chen, Q. Zhu and Y. Soh, “Smartphone Inertial Sensor-Based Indoor Localization and Tracking With iBeacon Corrections”, *IEEE Transactions on Industrial Informatics*, vol. 12, no. 4, pp. 1540-1549, 2016.
- [7] M. D. Yacoub, “The α - μ distribution: A Physical Fading Model for the Stacy Distribution,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 1, pp. 27–34, 2007.

- [8] H. Xia, J. Zuo, S. Liu and Y. Qiao, “Indoor Localization on Smartphones Using Built-In Sensors and Map Constraints”, *IEEE Transactions on Instrumentation and Measurement*, vol. 68, no. 4, pp. 1189-1198, 2019.
- [9] R. Yadav, B. Bhattarai, H. Gang and J. Pyun, “Trusted K Nearest Bayesian Estimation for Indoor Positioning System”, *IEEE Access*, vol. 7, pp. 51484-51498, 2019.
- [10] A. Belmonte-Hernandez, G. Hernandez-Penaloza, D. Martin Gutierrez and F. Alvarez, “SWiBluX: Multi-Sensor Deep Learning Fingerprint for Precise Real-Time Indoor Tracking”, *IEEE Sensors Journal*, vol. 19, no. 9, pp. 3473-3486, 2019.
- [11] B. -h. Lee, D. Ham, J. Choi, S. -C. Kim and Y. -H. Kim, “Genetic Algorithm for Path Loss Model Selection in Signal Strength-Based Indoor Localization,” *IEEE Sensors Journal*, vol. 21, no. 21, pp. 24285-24296, 1 Nov.1, 2021, doi: 10.1109/JSEN.2021.3110971.
- [12] V. Cantón Paterna, A. Calveras Augé, J. Paradells Aspas and M. Pérez Bulones, “A Bluetooth Low Energy Indoor Positioning System with Channel Diversity, Weighted Trilateration and Kalman Filtering”, *Sensors*, vol. 17, no. 12, p. 2927, 2017.
- [13] R. Liu, C. Yuen, T. Do and U. Tan, “Fusing Similarity-Based Sequence and Dead Reckoning for Indoor Positioning Without Training”, *IEEE Sensors Journal*, vol. 17, no. 13, pp. 4197-4207, 2017.
- [14] S. Pradhan, Y. Bae, J. Pyun, N. Ko and S. Hwang, “Hybrid TOA Trilateration Algorithm Based on Line Intersection and Comparison Approach of Intersection Distances”, *Energies*, vol. 12, no. 9, p. 1668, 2019.
- [15] S. He and K. Shin, “Geomagnetism for Smartphone-Based Indoor Localization”, *ACM Computing Surveys*, vol. 50, no. 6, pp. 1-37, 2018.

- [16] E. C. L. Chan, G. Baciú and S. Mak, “Wireless Tracking Analysis in Location Fingerprinting”, *IEEE International Conference on Wireless and Mobile Computing, Networking and Communications*, Avignon, France, 2008.
- [17] N. Yu, X. Zhan, S. Zhao, Y. Wu and R. Feng, “A Precise Dead Reckoning Algorithm Based on Bluetooth and Multiple Sensors”, *IEEE Internet of Things Journal*, vol. 5, no. 1, pp. 336-351, 2018.
- [18] V. Cantón Paterna, A. Calveras Augé, J. Paradells Aspás and M. Pérez Bulhões, “A Bluetooth Low Energy Indoor Positioning System with Channel Diversity, Weighted Trilateration and Kalman Filtering”, *Sensors*, vol. 17, no. 12, p. 2927, 2017.
- [19] W. Kang and Y. Han, “SmartPDR: Smartphone-Based Pedestrian Dead Reckoning for Indoor Localization”, *IEEE Sensors Journal*, vol. 15, no. 5, pp. 2906-2916, 2015.
- [20] Bahl, P., Padmanabhan, “RADAR: an in-building RF-based user location and tracking system”, *Proceedings IEEE INFOCOM 2000*, pp. 775–784, 2000.
- [21] Youssef, M., Agrawala, A. “The Horus WLAN location determination system”, *Proceedings of the 3rd International Conference on Mobile Systems, Applications, and Services. MobiSys*, 2005.
- [22] Yang, Z., Zhou, Z., Liu, Y. “From RSSI to CSI: indoor localization via channel response”, *ACM Computing Surveys* vol. 46, no. 2, pp. 1–32, 2013.
- [23] Xiao, J., et al. “FIFS: fine-grained indoor fingerprinting system”, *21st International Conference on Computer Communications and Networks (ICCCN)*, pp. 1–7, 2012.
- [24] J. A. Tropp, A. C. Gilbert and M. J. Strauss, “Simultaneous sparse approximation via greedy pursuit,” *Proc. IEEE International Conference on Acoustics, Speech, and Signal Processing*, Philadelphia, PA, USA, pp. v/721-v/724 Vol. 5, 2005.
- [25] J. Rezazadeh, R. Subramanian, K. Sandrasegaran, X. Kong, M. Moradi and F. Khodamoradi, “Novel iBeacon Placement for Indoor Positioning in IoT”, *IEEE Sensors Journal*, vol. 18, no. 24, pp. 10240-10247, 2018.

- [26] S. Subedi, H. Gang, N. Ko, S. Hwang and J. Pyun, “Improving Indoor Fingerprinting Positioning With Affinity Propagation Clustering and Weighted Centroid Fingerprint”, *IEEE Access*, vol. 7, pp. 31738-31750, 2019.
- [27] Weisstein, Eric W. “Least Squares Fitting–Logarithmic.” From MathWorld–A Wolfram Web Resource. <http://mathworld.wolfram.com/LeastSquaresFittingLogarithmic.html>
- [28] Yassin, A., Nasser, Y., Awad, M., Al-Dubai, A., Liu, R., Yuen, C., Aboutanios, E. “Recent advances in indoor localization: a survey on theoretical approaches and applications”, *IEEE Communications Surveys & Tutorials*, vol. 19, no. 2, pp. 1327-1346, 2017.
- [29] A. Zanella, “Best Practice in RSS Measurements and Ranging”, *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2662-2686, 2016.
- [30] D. Feng, C. Wang, C. He, Y. Zhuang and X. Xia, “Kalman-Filter-Based Integration of IMU and UWB for High-Accuracy Indoor Positioning and Navigation”, *IEEE Internet of Things Journal*, vol. 7, no. 4, pp. 3133-3146, 2020.
- [31] B. Hofmann-Wellenhof, James Collins and Herbert Lichtenegger “Global Positioning System: Theory and Practice”, Springer-Verlag GmbH. 2000.
- [32] S. He and K. Shin, “Geomagnetism for Smartphone-Based Indoor Localization”, *ACM Computing Surveys*, vol. 50, no. 6, pp. 1-37, 2018.
- [33] J. Jiao, F. Li, Z. Deng and W. Ma, “A Smartphone Camera-Based Indoor Positioning Algorithm of Crowded Scenarios with the Assistance of Deep CNN”, *Sensors*, vol. 17, no. 4, p. 704, 2017.
- [34] M. Arulampalam, S. Maskell, N. Gordon and T. Clapp, “A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking”, *IEEE Transactions on Signal Processing*, vol. 50, no. 2, pp. 174-188, 2002.

- [35] I. Guvenc, C. Abdallah, R. Jordan, and O. Dedoglu, “Enhancements to RSS-based indoor tracking systems using Kalman filters,” *Proc. Int. Signal Processing Conf.*, 2003.
- [36] Greg Welch and Gary Bishop “An Introduction to the Kalman Filter” Technical Report. University of North Carolina, Chapel Hill, USA, 1995.
- [37] S. Sadowski and P. Spachos, “RSSI-Based Indoor Localization With the Internet of Things”, *IEEE Access*, vol. 6, pp. 30149-30161, 2018.
- [38] L. Kanaris, A. Kokkinis, A. Liotta and S. Stavrou, “Fusing Bluetooth Beacon Data with Wi-Fi Radiomaps for Improved Indoor Localization”, *Sensors*, vol. 17, no. 4, p. 812, 2017.
- [39] T. -M. T. Dinh, N. -S. Duong and K. Sandrasegaran, “Smartphone-Based Indoor Positioning Using BLE iBeacon and Reliable Lightweight Fingerprint Map,” *IEEE Sensors Journal*, vol. 20, no. 17, pp. 10283-10294, 1 Sept.1, 2020, doi: 10.1109/JSEN.2020.2989411.
- [40] A. Goldsmith, *Wireless Communications*. New York, NY, USA: Cambridge University Press, 2005.
- [41] Yan, C. Tiberius, G. Janssen, P. Teunissen and G. Bellusci, “Review of range-based positioning algorithms”, *IEEE Aerospace and Electronic Systems Magazine*, vol. 28, no. 8, pp. 2-27, 2013.
- [42] M. D. Yacoub, “The α - μ distribution: A physical fading model for the stacy distribution,” *IEEE Transactions on Vehicular Technology*, vol. 56, no. 1, pp. 27 –34, 2007.
- [43] Yunye Jin, Hong-Song Toh, Wee-Seng Soh, and Wai-Choong Wong, “A Robust Dead-Reckoning Pedestrian Tracking System with Low Cost Sensors”, *IEEE International Conference on Pervasive Computing and Communications (PerCom)*, Seattle, WA, USA, 2011.

- [44] Weisstein, Eric W. “Newton’s Method.” From MathWorld—A Wolfram Web Resource. [Online]. Available: <https://mathworld.wolfram.com/NewtonsMethod.html>.
- [45] A. P. Dempster, N. M. Laird, and D. B. Rubin, “Maximum likelihood from incomplete data via the EM algorithm,” *Journal of the Royal Statistical Society Series B: Statistical Methodology*, vol. 39, no. 1, pp. 1–22, Sep. 1977. doi:10.1111/j.2517-6161.1977.tb01600.x.
- [46] A. Belloni, V. Chernozhukov, L. Wang, “Square-root LASSO: pivotal recovery of sparse signals via conic programming”, *Biometrika*, vol. 98, no. 4, pp. 791–806, 2011.
- [47] I. F. Gorodnitsky, B. D. Rao, “Sparse signal reconstruction from limited data using focuss: A reweighted minimum norm algorithm”, *IEEE Transactions on Signal Processing*, vol. 45, no. 3, pp. 600–616, 1997.
- [48] E. Candes, J. Romberg, “ l_1 -magic: Recovery of sparse signals via convex programming, Available online at <https://candes.su.domains/software/l1magic/downloads/l1magic.pdf>
- [49] Stephen Boyd, Neal Parikh, Eric Chu, Borja Peleato; Jonathan Eckstein “Distributed Optimization and Statistical Learning via the Alternating Direction Method of Multipliers”, 2011.
- [50] J. Yang and Y. Zhang, “Alternating direction algorithms for ℓ_1 -problems in compressive sensing,” *SIAM Journal on Scientific Computing*, vol. 33, no. 1, pp. 250–278, 2011.
- [51] S.J. Kim, K. Koh, M. Lustig, S. Boyd, and D. Gorinevsky, “An interior-point method for large-scale ℓ_1 -regularized least squares,” *IEEE Journal of Selected Topics in Signal Processing*, vol. 1, no. 4, pp. 606–617, Dec. 2007. doi:10.1109/jstsp.2007.910971

- [52] Z. Ben-Haim, Y. C. Eldar and M. Elad, “Coherence-Based Performance Guarantees for Estimating a Sparse Vector Under Random Noise,” *IEEE Transactions on Signal Processing*, vol. 58, no. 10, pp. 5030-5043, 2010.
- [53] A. Hu and J. He, “Position-Aided Beam Learning for Initial Access in mmWave MIMO Cellular Networks,” *IEEE Systems Journal*, vol. 16, no. 1, pp. 1103-1113, 2022, doi: 10.1109/JSYST.2020.3027757.
- [54] C. J. Hegarty and E. Chatre, “Evolution of the Global Navigation Satellite System (GNSS),” *Proceedings of the IEEE*, vol. 96, no. 12, pp. 1902-1917, Dec. 2008, doi: 10.1109/JPROC.2008.2006090.
- [55] F. Lemic et al., “Localization as a feature of mmWave communication,” *2016 International Wireless Communications and Mobile Computing Conference (IWCMC)*, Paphos, Cyprus, 2016, pp. 1033-1038, doi: 10.1109/IWCMC.2016.7577201.
- [56] S. Geng, J. Kivinen, X. Zhao and P. Vainikainen, “Millimeter-Wave Propagation Channel Characterization for Short-Range Wireless Communications,” *IEEE Transactions on Vehicular Technology*, vol. 58, no. 1, pp. 3-13, Jan. 2009, doi: 10.1109/TVT.2008.924990.
- [57] A. Shahmansoori, G. E. Garcia, G. Destino, G. Seco-Granados and H. Wymeersch, “Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems,” *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822-1835, March 2018, doi: 10.1109/TWC.2017.2785788.
- [58] R. Mendrzik, H. Wymeersch, G. Bauch and Z. Abu-Shaban, “Harnessing NLOS Components for Position and Orientation Estimation in 5G Millimeter Wave MIMO,” *IEEE Transactions on Wireless Communications*, vol. 18, no. 1, pp. 93-107, Jan. 2019, doi: 10.1109/TWC.2018.2877615.
- [59] A. Zanella, “Best Practice in RSS Measurements and Ranging,” *IEEE Communications Surveys & Tutorials*, vol. 18, no. 4, pp. 2662-2686, Fourthquarter 2016, doi: 10.1109/COMST.2016.2553452.

- [60] N. -S. Duong and T. -M. Dinh, "Indoor Localization with lightweight RSS Fingerprint using BLE iBeacon on iOS platform," *2019 19th International Symposium on Communications and Information Technologies (ISCIT)*, Ho Chi Minh City, Vietnam, 2019, pp. 91-95, doi: 10.1109/ISCIT.2019.8905160.
- [61] C. K. Anjinappa, Y. Zhou, Y. Yapici, D. Baron and I. Guvenc, "Channel Estimation in mmWave Hybrid MIMO System via Off-Grid Dirichlet Kernels," *2019 IEEE Global Communications Conference (GLOBECOM)*, Waikoloa, HI, USA, 2019, pp. 1-6, doi: 10.1109/GLOBECOM38437.2019.9013906.
- [62] R. Roy and T. Kailath, "ESPRIT-estimation of signal parameters via rotational invariance techniques," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 37, no. 7, pp. 984-995, July 1989, doi: 10.1109/29.32276.
- [63] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Transactions on Antennas and Propagation*, vol. 34, no. 3, pp. 276-280, March 1986, doi: 10.1109/TAP.1986.1143830.
- [64] M. F. Duarte, S. Sarvotham, D. Baron, M. B. Wakin and R. G. Baraniuk, "Distributed Compressed Sensing of Jointly Sparse Signals," *Conference Record of the Thirty-Ninth Asilomar Conference on Signals, Systems and Computers*, Pacific Grove, CA, USA, 2005, pp. 1537-1541, doi: 10.1109/ACSSC.2005.1600024.
- [65] J. Du, J. Dong and F. Gao, "Simultaneous Channel Estimation and Localization of Terahertz Massive MIMO Systems via Bayesian Tensor Decomposition," *IEEE Communications Letters*, vol. 27, no. 2, pp. 541-545, Feb. 2023, doi: 10.1109/LCOMM.2022.3223336.
- [66] Kaufman L., and P. J. Rousseeuw. Finding Groups in Data: An Introduction to Cluster Analysis. Hoboken, NJ: John Wiley & Sons, Inc., 1990.
- [67] N. -S. Duong, T. -M. Dinh-Thi and Q. -. T. Nguyen, "mmWave Channel Estimation for Location-based Application in 5G MIMO systems," *TENCON 2022 - 2022 IEEE Region 10 Conference (TENCON)*, Hong Kong, 2022, pp. 1-6, doi: 10.1109/TENCON55691.2022.9977818.

- [68] M. T. Hoang, B. Yuen, X. Dong, T. Lu, R. Westendorp and K. Reddy, "Recurrent Neural Networks for Accurate RSSI Indoor Localization," *IEEE Internet of Things Journal*, vol. 6, no. 6, pp. 10639-10651, Dec. 2019, doi: 10.1109/JIOT.2019.2940368.
- [69] J. Lee, G. -T. Gil and Y. H. Lee, "Channel Estimation via Orthogonal Matching Pursuit for Hybrid MIMO Systems in Millimeter Wave Communications," *IEEE Transactions on Communications*, vol. 64, no. 6, pp. 2370-2386, June 2016, doi: 10.1109/TCOMM.2016.2557791.
- [70] S. J. Wright, "Coordinate descent algorithms," *Mathematical Programming*, vol. 151, no. 1, pp. 3-34, Mar. 2015. doi:10.1007/s10107-015-0892-3
- [71] S. E. Trevlakakis et al., "Localization as a Key Enabler of 6G Wireless Systems: A Comprehensive Survey and an Outlook," *IEEE Open Journal of the Communications Society*, vol. 4, pp. 2733-2801, 2023, doi: 10.1109/OJCOMS.2023.3324952.
- [72] M. -C. Gosselin et al., "Estimation Formulas for the Specific Absorption Rate in Humans Exposed to Base-Station Antennas," *IEEE Transactions on Electromagnetic Compatibility*, vol. 53, no. 4, pp. 909-922, Nov. 2011, doi: 10.1109/TEMC.2011.2139216.
- [73] Z. Abu-Shaban, X. Zhou, T. Abhayapala, G. Seco-Granados and H. Wymeersch, "Error Bounds for Uplink and Downlink 3D Localization in 5G Millimeter Wave Systems", *IEEE Transactions on Wireless Communications*, vol.17, no. 8, pp. 4939-4954, 2018.
- [74] E. Leitinger, P. Meissner, C. Rudisser, G. Dumphart and K. Witrisal, "Evaluation of Position-Related Information in Multipath Components for Indoor Positioning", *IEEE Journal on Selected Areas in Communications*, vol. 33, no. 11, pp. 2313-2328, 2015.
- [75] Location Services (LCS); Service Description; Stage 1, Version 17.0.0, 3GPP Standard TS 22.071, Mar. 2022. [Online]. Available:

- <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=584>
- [76] 5G System (5GS) Location Services (LCS); Stage 2, Version 18.1.0, 3GPP Standard TS 23.273, Mar. 2023. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3577>
 - [77] NG Radio Access Network (NG-RAN); Stage 2 Functional Specification of User Equipment (UE) Positioning in NG-RAN, Version 17.4.0, 3GPP Standard TS 38.305, Mar. 2023. [Online]. Available: <https://portal.3gpp.org/desktopmodules/Specifications/SpecificationDetails.aspx?specificationId=3310>
 - [78] T. T. Cai and L. Wang, “Orthogonal Matching Pursuit for Sparse Signal Recovery With Noise,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680-4688, July 2011, doi: 10.1109/TIT.2011.2146090.
 - [79] M. Koivisto et al., “Joint Device Positioning and Clock Synchronization in 5G Ultra-Dense Networks”, *IEEE Transactions on Wireless Communications*, vol. 16, no. 5, pp. 2866-2881, 2017.
 - [80] A. Shahmansoori, G. Garcia, G. Destino, G. Seco-Granados and H. Wymeersch, “Position and Orientation Estimation Through Millimeter-Wave MIMO in 5G Systems”, *IEEE Transactions on Wireless Communications*, vol. 17, no. 3, pp. 1822-1835, 2018.
 - [81] J. del Peral-Rosado, G. Seco-Granados, S. Kim and J. Lopez-Salcedo, “Network Design for Accurate Vehicle Localization”, *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4316-4327, 2019.
 - [82] W. Zhang, M. Dong and T. Kim, “MMV-Based Sequential AoA and AoD Estimation for Millimeter Wave MIMO Channels,” *IEEE Transactions on Communications*, vol. 70, no. 6, pp. 4063-4077, June 2022, doi: 10.1109/TCOMM.2022.3168886.

- [83] M. Fröhle, C. Lindberg, K. Granström and H. Wymeersch, “Multisensor Poisson Multi-Bernoulli Filter for Joint Target–Sensor State Tracking”, *IEEE Transactions on Intelligent Vehicles*, vol. 4, no. 4, pp. 609-621, 2019.
- [84] B. Peng, G. Seco-Granados, E. Steinmetz, M. Frohle and H. Wymeersch, “Decentralized Scheduling for Cooperative Localization With Deep Reinforcement Learning”, *IEEE Transactions on Vehicular Technology*, vol. 68, no. 5, pp. 4295-4305, 2019.
- [85] J. Lota, S. Ju, O. Kanhere, T. S. Rappaport and A. Demosthenous, “MmWave V2V Localization in MU-MIMO Hybrid Beamforming,” *IEEE Open Journal of Vehicular Technology*, doi: 10.1109/OJVT.2022.3170522.
- [86] O. Kanhere and T. S. Rappaport, “Position Location for Futuristic Cellular Communications: 5G and Beyond,” *IEEE Communications Magazine*, vol. 59, no. 1, pp. 70-75, January 2021, doi: 10.1109/MCOM.001.2000150.
- [87] A. Fascista, A. Coluccia, H. Wymeersch and G. Seco-Granados, “Downlink Single-Snapshot Localization and Mapping With a Single-Antenna Receiver,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 7, pp. 4672-4684, July 2021, doi: 10.1109/TWC.2021.3061407.
- [88] A. Kakkavas, H. Wymeersch, G. Seco-Granados, M. H. C. García, R. A. Stirling-Gallacher and J. A. Nossek, “Power Allocation and Parameter Estimation for Multipath-Based 5G Positioning,” *IEEE Transactions on Wireless Communications*, vol. 20, no. 11, pp. 7302-7316, Nov. 2021, doi: 10.1109/TWC.2021.3082581.
- [89] C. Baquero Barneto et al., “Millimeter-wave mobile sensing and environment mapping: Models, algorithms and validation,” *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 3900–3916, 2022. doi:10.1109/tvt.2022.3146003
- [90] Share Technote “5G/NR - Beam Management” [Online] <https://mail.sharetechnote.com/html/5G/5G-Phy-BeamManagement.html>

- [91] O. Kanhere and T. S. Rappaport, “Millimeter Wave Position Location using Multipath Differentiation for 3GPP using Field Measurements,” *GLOBECOM 2020 - 2020 IEEE Global Communications Conference*, Taiwan, 2020, pp. 1-7, doi: 10.1109/GLOBECOM42002.2020.9348110.
- [92] Y. Ge et al., “Experimental Validation of Single BS 5G mmWave Positioning and Mapping for Intelligent Transport,” *IEEE Transactions on Vehicular Technology*, doi: 10.1109/TVT.2024.3418140.
- [93] O. Kanhere and T. S. Rappaport, “Outdoor sub-THz Position Location and Tracking using Field Measurements at 142 GHz,” *ICC 2021 - IEEE International Conference on Communications*, Montreal, QC, Canada, 2021, pp. 1-6, doi: 10.1109/ICC42927.2021.9500482.
- [94] J. Lota, S. Ju, O. Kanhere, T. S. Rappaport and A. Demosthenous, “mmWave V2V Localization in MU-MIMO Hybrid Beamforming,” *IEEE Open Journal of Vehicular Technology*, vol. 3, pp. 210-220, 2022, doi: 10.1109/OJVT.2022.3170522.
- [95] S. Sand, R. Tanbourgi, C. Mensing, and R. Raulefs, “Position aware adaptive communication systems,” *Forty-Third Asilomar Conference on Signals, Systems and Computers*, 2009, pp. 73-77.
- [96] R. Di Taranto, S. Muppirisetty, R. Raulefs, D. Slock, T. Svensson and H. Wymeersch, “Location-Aware Communications for 5G Networks: How location information can improve scalability, latency, and robustness of 5G”, *IEEE Signal Processing Magazine*, vol. 31, no. 6, pp. 102-112, 2014.
- [97] 3GPP TR 38.855 V16.0.0: Study on NR positioning support (Release 16); Technical Report, Mar .2019.
- [98] 3rd Generation Partnership Project (3GPP), Release 17 Description; Summary of Rel-17 Work Items, document TR21.917, Dec. 2022.

- [99] W. U. Bajwa, J. Haupt, A. M. Sayeed, and R. Nowak, “Compressed channel sensing: A new approach to estimating sparse multipath channels,” *Proceedings of the IEEE*, vol. 98, no. 6, pp. 1058–1076, Jun. 2010.
- [100] Y. Han and J. Lee, “Two-stage compressed sensing for millimeter wave channel estimation,” *Proc. IEEE Int. Symp. Inf. Theory (ISIT)*, pp. 860–864, Spain, 2016.
- [101] Y. Wu, Y. Gu, and Z. Wang, “Channel estimation for mmWave MIMO with transmitter hardware impairments,” *IEEE Communications Letters*, vol. 22, no. 2, pp. 320–323, Feb. 2018.
- [102] Li, J., Da Costa, M.F. and Mitra, U. (2022) “Joint localization and orientation estimation in millimeter-wave MIMO OFDM systems via Atomic Norm Minimization,” *IEEE Trans. Signal Process.*, vol. 70, pp. 4252–4264, 2022.
- [103] S. Ji, Y. Xue, and L. Carin, “Bayesian compressive sensing,” *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2346–2356, 2008.
- [104] J. Lee, G.-T. Gil, and Y. H. Lee, “Channel estimation via orthogonal matching pursuit for hybrid MIMO systems in Millimeter Wave Communications,” *IEEE Trans. Commun.*, vol. 64, no. 6, pp. 2370–2386, 2016.
- [105] A. Belloni, V. Chernozhukov, L. Wang, “Square-root LASSO: pivotal recovery of sparse signals via conic programming”, *Biometrika*, vol. 98, no. 4, pp. 791–806, 2011.
- [106] L. Kanaris, A. Kokkinis, A. Liotta and S. Stavrou, “Fusing Bluetooth Beacon Data with Wi-Fi Radiomaps for Improved Indoor Localization”, *Sensors*, vol. 17, no. 4, p. 812, 2017.
- [107] Hornik, K., Stinchcombe, M., White, H. “Multilayer feedforward networks are universal approximators”, *Neural Network*. 2(5), 359–366, 1989.
- [108] D. L. Donoho, Y. Tsaig, I. Drori and J. -L. Starck, “Sparse Solution of Underdetermined Systems of Linear Equations by Stagewise Orthogonal Matching Pursuit,”

- IEEE Transactions on Information Theory*, vol. 58, no. 2, pp. 1094-1121, Feb. 2012, doi: 10.1109/TIT.2011.2173241.
- [109] Y. Chen, J. Yang, W. Trappe and R. Martin, “Impact of Anchor Placement and Anchor Selection on Localization Accuracy”, *Handbook of Position Location*, pp. 435-465, 2019.
 - [110] Wang, X., Wang, X., Mao, S. “CiFi: deep convolutional neural networks for indoor localization with 5 GHz Wi-Fi”, *2017 IEEE International Conference on Communications (ICC)*, pp. 1–6. (2017)
 - [111] Hsieh, C.H., Chen, J.Y., Nien, B.H. “Deep learning-based indoor localization using received signal strength and channel state information”, *IEEE Access*. 7, 33256–33267. (2019)
 - [112] Bencharif, L., Ouameur, M.A., Massicotte, D. “Long short-term memory for indoor localization using WI-FI received signal strength and channel state information”, *2021 IEEE 4th 5G World Forum (5GWF)*, pp. 230–235. (2021).
 - [113] W. Njima, M. Chafii, A. Chorti, R. M. Shubair and H. V. Poor, “Indoor Localization Using Data Augmentation via Selective Generative Adversarial Networks,” *IEEE Access*, vol. 9, pp. 98337-98347, 2021, doi: 10.1109/ACCESS.2021.3095546.
 - [114] Y. Tian, D. Shigaki, W. Wang and C. Ahn, “A weighted least squares method using received signal strength measurements for WLAN indoor positioning system”, *20th International Symposium on Wireless Personal Multimedia Communications (WPMC)*, 2017.
 - [115] G. Li, E. Geng, Z. Ye, Y. Xu, J. Lin and Y. Pang, “Indoor Positioning Algorithm Based on the Improved RSSI Distance Model”, *Sensors*, vol. 18, no. 9, p. 2820, 2018.
 - [116] A. Mackey, P. Spachos and K. Plataniotis, “Enhanced Indoor Navigation System with Beacons and Kalman Filters”, *IEEE Global Conference on Signal and Information Processing (GlobalSIP)*, 2018.

- [117] M. Mallik, A. K. Panja, C. Chowdhury, “Paving the way with machine learning for seamless indoor–outdoor positioning: A survey”, *Information Fusion*, vol. 94, pp. 126-151, 2023. doi: 10.1016/j.inffus.2023.01.023.
- [118] V. Bellavista-Parent, J. Torres-Sospedra and A. Perez-Navarro, “New trends in indoor positioning based on WiFi and machine learning: A systematic review,” *2021 International Conference on Indoor Positioning and Indoor Navigation (IPIN)*, Spain, 2021, pp. 1-8, doi: 10.1109/IPIN51156.2021.9662521.
- [119] F. Zafari, A. Gkelias and K. K. Leung, “A Survey of Indoor Localization Systems and Technologies,” *IEEE Communications Surveys & Tutorials*, vol. 21, no. 3, pp. 2568-2599, thirdquarter 2019, doi: 10.1109/COMST.2019.2911558.
- [120] K. Alinsavath, L. Nugroho, Widyawan, and K. Hamamoto, “The Seamlessness of Outdoor and Indoor Localization Approaches based on a Ubiquitous Computing Environment: A Survey”, *Proceedings of the 2nd ACM International Conference on Information Science and Systems (ICISS '19)*, New York, NY, USA, pp. 316–324. <https://doi.org/10.1145/3322645.3322690>.
- [121] B. Molina, E. Olivares, C. E. Palau and M. Esteve, “A Multimodal Fingerprint-Based Indoor Positioning System for Airports,” *IEEE Access*, vol. 6, pp. 10092-10106, 2018, doi: 10.1109/ACCESS.2018.2798918.
- [122] T. J. Gallagher, B. Li, A. G. Dempster and C. Rizos, “A sector-based campus-wide indoor positioning system,” *2010 International Conference on Indoor Positioning and Indoor Navigation*, Switzerland, 2010, pp. 1-8, doi: 10.1109/IPIN.2010.5648250.
- [123] H. Hakim, R. Renouf and J. Enderle, “Passive RFID Asset Monitoring System in Hospital Environments,” *Proceedings of the IEEE 32nd Annual Northeast Bioengineering Conference*, USA, 2006, pp. 217-218, doi: 10.1109/NEBC.2006.1629830.
- [124] Rubino, I., Xhembulla, J., Martina, A., Bottino, A., and Malnati, G. “MusA: Using Indoor Positioning and Navigation to Enhance Cultural Experiences in a Museum”, *Sensors*, vol. 12, no. 13, pp. 17445-17471, 2013. doi: 10.3390/s131217445