

KỶ YẾU HỘI THẢO KHOA HỌC TRẺ LẦN 4 NĂM 2022 YSC2022

KHƠI THÔNG TRI THỨC - KIẾN TẠO TƯƠNG LAI

LĨNH VỰC ĐIỆN - ĐIỆN TỬ - CÔNG NGHỆ THÔNG TIN



**NHÀ XUẤT BẢN
ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH**

**BỘ CÔNG THƯƠNG
TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH**

**HỘI NGHỊ KHOA HỌC TRẺ LẦN 4
NĂM 2022 (YSC 2022)**

NHÀ XUẤT BẢN ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

HỘI NGHỊ KHOA HỌC TRẺ LẦN 4 NĂM 2022

TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

BAN TỔ CHỨC

- TS. Phan Hồng Hải	Hiệu trưởng	Trưởng ban
- PGS. TS Đàm Sao Mai	Phó Hiệu trưởng	Phó Trưởng ban
- PGS. TS Trịnh Ngọc Nam	Trưởng phòng Quản lý khoa học và Hợp tác quốc tế	Phó Trưởng ban
- GS.TS. Lê Văn Tấn	Trưởng tiểu ban Hóa – Sinh – Thực phẩm – Môi trường	Thành viên
- PGS.TS Huỳnh Trung Hiếu	Trưởng tiểu ban Kỹ thuật - Công nghệ thông tin	Thành viên
- PGS.TS Nguyễn Đức Nam	Trưởng tiểu ban Cơ khí – Xây dựng	Thành viên
- TS. Ngô Ngọc Hưng	Trưởng tiểu ban Khoa học Xã hội và Nhân văn	Thành viên
- TS. Nguyễn Thị Thu Hiền	Trưởng tiểu ban Kinh tế	Thành viên
TS. Lê Ngọc Sơn	Viện trưởng Viện Đào tạo quốc tế và Sau đại học	Thành viên
- TS. Phạm Trần Bích Thuận	Phó trưởng phòng Quản lý khoa học và Hợp tác quốc tế	Thành viên
- Ths. Bùi Đình Tiền	Phó trưởng phòng Quản lý khoa học và Hợp tác quốc tế	Thành viên
- ThS. Phạm Trung Kiên	Trưởng phòng Tổ chức – Hành chính	Thành viên
- ThS. Nguyễn Đình Hiền	Trưởng phòng Tài chính – Kế toán	Thành viên
- ThS. Nguyễn Thị Thương	Giám đốc Trung tâm Thông tin - Truyền thông	Thành viên
- CN. Hồ Văn Thái	Chuyên viên phòng Quản lý khoa học và Hợp tác quốc tế	Thành viên
- KS. Huỳnh Phú Vinh	Bí thư Đoàn Thanh niên Trường	Thành viên
- Ngô Đình Luật	Chủ tịch Hội Sinh viên Trường	Thành viên

BAN CHUYÊN MÔN THẨM DUYỆT BÀI BÁO

- PGS.TS. Huỳnh Trung Hiếu	Trưởng khoa Công nghệ Thông tin	Trưởng ban
- TS. Mai Thăng Long	Trưởng khoa Công nghệ Điện tử	Thành viên
- TS. Lê Ngọc Sơn	Viện Đào tạo quốc tế và Sau đại học	Thành viên
- ThS. Phạm Trung Kiên	Trưởng phòng Tổ chức – Hành chính	Thành viên
- TS. Trần Thanh Ngọc	Trưởng khoa Công nghệ Điện	Thành viên
- TS. Nguyễn Trung Nhân	Trưởng phòng Đào tạo	Thành viên
- PGS.TS. Nguyễn Tấn Lũy	Phó Trưởng khoa Công nghệ Điện tử	Thành viên
- TS. Đặng Thị Phúc	Phó Trưởng khoa Công nghệ Thông tin	Thư ký

BAN BIÊN TẬP

- ThS. Hoàng Phượng Trâm	Phòng Quản lý khoa học và Hợp tác quốc tế	Trưởng ban
- ThS. Nguyễn Phúc Thùy Dương	---	Thành viên
- ThS. Nguyễn Minh Tú Anh	---	Thành viên
- ThS. Bạch Thị Lê	---	Thành viên
- ThS. Nguyễn Thị Lụa	---	Thành viên
- ThS. Nguyễn Diệu Linh	---	Thành viên
- CN. Hồ Văn Thái	---	Thành viên
- CN. Đoàn Thị Hồng Gấm	---	Thành viên

LĨNH VỰC
ĐIỆN – ĐIỆN TỬ
CÔNG NGHỆ THÔNG TIN

DANH MỤC BÀI BÁO

YSC4F.201	11
GOQUICK: CÔNG CỤ HỖ TRỢ QUẢN LÝ NHÀ HÀNG DỰA TRÊN ĐIỆN TOÁN ĐÁM MÂY	
Trần Thị Anh Thi, Đoàn Văn Vĩnh, Nguyễn Thanh Liêm	
YSC4F.202	18
ĐÁNH GIÁ MỨC ĐỘ TIÊU CỰC BÌNH LUẬN TRÊN MẠNG XÃ HỘI	
Phạm Minh Tuấn, Trương Nguyễn Duy Tân, Nguyễn Quang Quốc, Đào Duy Trường, Nguyễn Thị Thanh Hòa, Đặng Thị Phúc	
YSC4F.203	32
THIẾT KẾ HỆ THỐNG ĐIỀU KHIỂN THIẾT BỊ SỬ DỤNG RASPBERRY PI 4 TRÊN NỀN TẢNG IOT HOME ASSISTANT	
Nguyễn Đình Khôi, Phạm Hùng Phương, Nguyễn Tấn Lộc	
YSC4F.204	39
ỨNG DỤNG MÔ HÌNH HỌC MÁY TRONG PHÂN TÍCH CẢM XÚC KHÁCH HÀNG	
Hồ Hoàng Huy Quang, Phạm Thị Thiết, Võ Quang Hoàng Khang	
YSC4F.205	50
HỆ THỐNG PHÁT HIỆN NGƯỜI ĐEO VÀ KHÔNG ĐEO KHẨU TRANG DỰA TRÊN MÔ HÌNH HỌC SÂU	
Nguyễn Hoanh, Nguyễn Công Tước, Phạm Thanh Ngân, Nguyễn Phi Trường, Huỳnh Đức Hải	
YSC4F.206	65
HỆ THỐNG NHẬN DIỆN VÀ ĐO KHOẢNG CÁCH GIỮA NGƯỜI VỚI NGƯỜI SỬ DỤNG HỌC SÂU	
Nguyễn Hoanh, Đỗ Nguyễn Hoàng Long, Huỳnh Trần Minh Thuận, Trần Nguyễn Bá Thọ, Bùi Ngọc Minh Phong	
YSC4F.207	81
MÔ HÌNH TOÁN HỌC Ô TỬ DỌC TRỰC HÌNH E CẤU TRÚC NGUYÊN KHỐI	
Lê Ngọc Hội, Nguyễn Tùng Lâm	
YSC4F.208	93
ỨNG DỤNG CẤU TRÚC ĐIỀU KHIỂN BS ₂ SM MỚI VÀO HỆ TRUYỀN ĐỘNG ĐỘNG CƠ SÁU PHA CHO HỆ THỐNG ĐẨY CỬA TÀU THỦY	
Phạm Thụy Ngọc	
YSC4F.209	106
THIẾT KẾ BỘ VI XỬ LÝ RISC 16 BIT DÙNG FPGA	
Vu Thi Hong Nga, Le Tan Khanh, Bui Quoc Thang	
YSC4F.210	118
THIẾT KẾ VI ĐIỀU KHIỂN 32 BIT TRÊN FPGA	
Vu Thi Hong Nga, Do Truong Vi Ky, Doan Ba Tien	

YSC4F.211.....	128
THE SYSTEM OF MONITORING AND TAKE CARE OF PETS REMOTELY	
Dinh Quang Tuyen, Vu Quoc Lan	
YSC4F.212.....	139
ĐIỀU KHIỂN THÍCH NGHI TRƯỢT BACKSTEPPING CHO MÔ HÌNH TÀU THỦY 3D DÙNG MẠNG NƠ-RON	
Hoàng Thị Tú Uyên	
YSC4F.213.....	153
ĐÁNH GIÁ VỀ NHỮNG CẤU HÌNH VẬN HÀNH, MÔ HÌNH ĐIỀU KHIỂN VÀ CẤU HÌNH NỐI ĐẤT CỦA CÁC DẠNG LƯỚI ĐIỆN MICROGRID	
Lê Duy Phúc, Chung Anh Vũ	
YSC4F.214.....	170
PHÂN TÍCH XÁC SUẤT DỪNG CỦA MẠNG HỢP TÁC HAI CHIỀU CÓ THU HOẠCH NĂNG LƯỢNG VÔ TUYẾN VÀ KỸ THUẬT MÃ HÓA MẠNG SỐ	
Nguyễn Linh Tâm, Đỗ Ngọc Thanh, Đào Thị Thu Thủy	
YSC4F.215.....	178
HỆ THỐNG GIÁM SÁT CHẤT LƯỢNG NGUỒN NƯỚC XẢ THẢI CÔNG NGHIỆP THEO THỜI GIAN THỰC	
Duong Quoc Thang, Nguyen Trong Tuan, Tran Thi Minh Khoa	
YSC4F.216.....	189
TỐI ƯU CHIỀU DÀI CÁP ĐIỆN NHÀ MÁY ĐIỆN GIÓ XA BỜ DỰA TRÊN THUẬT TOÁN LAI SPDPSO VÀ SWEEP HEURISTIC	
Pham Quoc Khanh, Nguyen Huy Khiem	
YSC4F.217.....	199
CẢI TIẾN TÍNH CHẤT ĐIỆN CỦA VẬT LIỆU TGS SỬ DỤNG ỐNG NANOCACBON ĐA VÁCH DẠNG OXI HÓA GIÚP MỞ RỘNG VÙNG HOẠT ĐỘNG CỦA CÁC CẢM BIẾN HỒNG NGOẠI	
Nguyễn Hoài Thương, Hà Văn Đại	
YSC4F.218.....	205
ỨNG DỤNG IOT TRÊN BOARD ESP8266 VÀ GIAO THỨC MQTT CHO MÔ HÌNH QUẢN LÝ, GIÁM SÁT VĂN PHÒNG THÔNG MINH	
Phạm Hồng Thái, Nguyễn Văn Dũng, Nguyễn Anh Vinh	
YSC4F.219.....	216
ỨNG DỤNG GIẢI THUẬT HỌC SÂU VÀ LOGIC MỜ XÂY DỰNG HỆ THỐNG GIÁM SÁT VÀ ĐIỀU KHIỂN CHO NGÔI NHÀ THÔNG MINH TRÊN BOARD RASPBERRY VÀ ARDUINO	
Đặng Thái Hiên, Nguyễn Bá Hoàn, Nguyễn Anh Vinh	
YSC4F.220.....	231
ỨNG DỤNG MÔ HÌNH DEEP SPEECH TRONG NHẬN DẠNG GIỌNG NÓI TIẾNG VIỆT	
Nguyễn Văn Anh Tuấn, Phạm Huỳnh Trường Thịnh, Phạm Trung Sơn, Đặng Thị Phúc	
YSC4F.221.....	240
NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG ĐÁNH GIÁ CẢM XÚC KHÁCH HÀNG THÔNG QUA KHUÔN MẶT	
Nguyễn Văn Duy, Nguyễn Gia Bảo, Trần Thanh Bình	

YSC4F.222.....	250
NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG TRỢ LÝ ẢO CỔ VẤN HỌC TẬP CHO SINH VIÊN IUH	
Nguyễn Văn Duy, Phan Thanh Nhật, Hoàng Ngọc Thụy	
YSC4F.223.....	257
XÂY DỰNG HỆ THỐNG CHUÔNG CỬA THÔNG MINH KHÔNG TIẾP XÚC	
Nguyễn Văn Duy , Phạm Quốc Hoàng	
YSC4F.224.....	265
CHẨN ĐOÁN LỖI XE Ô TÔ THỜI GIAN THỰC	
Nguyễn Thanh Đăng , Trần Minh Quân, Trần Anh Khoa	
YSC4F.225.....	273
XÂY DỰNG ỨNG DỤNG HƯỚNG DẪN DU LỊCH TỰ ĐỘNG TẠI THÀNH PHỐ HỒ CHÍ MINH SỬ DỤNG KỸ THUẬT HỌC SÂU	
Dang Thi Phuc, Nguyen Vu Lam Duy, Tran Thi Minh Khoa, Le Thi Anh Tuyet	
YSC4F.226.....	284
THIẾT KẾ HỆ THỐNG GIÁM SÁT TỪ XA CÔNG SUẤT NĂNG LƯỢNG TÁI TẠO	
Vũ Việt Nam, Nguyễn Lê Vĩnh Phát, Nguyễn Văn Lành	
YSC4F.227.....	295
HOUFLEX: SEARCHING & SMART MANAGEMENT BOARDING HOUSE APPLICATION BASED ON BLOCKCHAIN MECHANISM	
Nguyễn Công Thành Đạt, Mai Kiên Cường, Lưu Tuấn Kha, Trương Công Cường, Tôn Long Phước	
YSC4F.228.....	308
HOẠCH ĐỊNH QUỸ ĐẠO VÀ ĐIỀU KHIỂN BÙ SAI LỆCH ĐỘNG HỌC CHO MOBILE ROBOT	
Trần Hữu Toàn, Vương Hoàng Sang, Cao Văn Thắng	
YSC4F.229.....	327
XÂY DỰNG HỆ THỐNG QUẢN LÝ NHÂN SỰ CHO CÔNG TY CỔ PHẦN BAO BÌ PHƯƠNG NAM	
Nguyễn Văn Chuẩn, Nguyễn Đăng Tuấn Đạt, Nguyễn Thị Hạnh	
YSC4F.230.....	342
ĐỘI NGŨ SHIPPER VÀ SỰ CẦN THIẾT CỦA MỘT ỨNG DỤNG QUẢN LÝ- THEO DÕI TRONG THỜI ĐẠI CÁCH MẠNG 4.0	
Nguyễn Văn Thắng, Nguyễn Văn Hùng	
YSC4F.231.....	351
INTRODUCTION TO MULTI-MARGINAL OPTIMAL TRANSPORT ON SUB-RIEMANNIAN MANIFOLD	
Thanh Sơn Trinh	
YSC4F.232.....	360
TRẠM QUAN TRẮC VÀ DỰ BÁO THỜI TIẾT SỬ DỤNG RASPBERRY PI ỨNG DỤNG MẠNG NƠ RON VÀ GIẢI THUẬT LAN TRUYỀN NGƯỢC	
Lê Thị Cúc , Cao Văn Kiên	

YSC4F.233	371
SỬ DỤNG THUẬT TOÁN HEURISTIC ĐỂ TỐI ƯU BÀI TOÁN GIAO HÀNG	
Lâm Ngọc Long, Võ Hoàng Tuấn, Nguyễn Đình Thanh, Hồ Hoàng Vân Anh, Đoàn Quốc Huy, Lê Phúc Lữ	
YSC4F.234	380
THUẬT TOÁN COOT CHO BÀI TOÁN XÁC ĐỊNH VỊ TRÍ VÀ CÔNG SUẤT NGUỒN ĐIỆN PHÂN TÁN TRÊN LƯỚI ĐIỆN PHÂN PHỐI	
Trịnh Hữu Trường, Võ Tấn Lộc, Dương Thanh Long, Nguyễn Thanh Thuận	
YSC4F.235	388
ẢNH HƯỞNG CỦA NHIỀU KÊNH HỢP TÁC (CCI) TÁC ĐỘNG ĐẾN HIỆU NĂNG HỆ THỐNG MẠNG VÔ TUYẾN NHẬN THỨC SỬ DỤNG NOMA	
Tôn Thất Phùng, Lê Phúc Đức, Trần Hoàng Khải Long	
YSC4F.236	398
NGHIÊN CỨU VÀ THIẾT KẾ BỘ ĐIỀU KHIỂN PI MỜ ĐỂ ĐIỀU KHIỂN TỐC ĐỘ ĐỘNG CƠ KHÔNG ĐỒNG BỘ DỰA TRÊN LABVIEW	
Ngô Thanh Quyền, Nguyễn Văn Thọ, Nguyễn Đức Trung, Phạm Ngọc Duyên, Lê Nguyên Tân	
YSC4F.237	408
PHÂN TÍCH VÀ THIẾT KẾ: FEATURE STORE	
Bui Thi Huyen Trang, Cu Thi Cam Ly, Ho Vo Hoang Duy, Nguyen Quang Bao, Doan Anh Duong, Tran Huu Nhan	
YSC4F.238	415
SỬ DỤNG LÝ THUYẾT TRÒ CHƠI VÀO BÀI TOÁN KHUYẾN MÃI TRONG THƯƠNG MẠI ĐIỆN TỬ	
Lâm Quang Phú, Huỳnh Thị Lưu Ly, Trần Lộc, Trương Đức Tài, Nguyễn Trần Gia Huy, Lê Phúc Lữ	
YSC4F.239	427
SỬ DỤNG THUẬT TOÁN PHÂN CỤM ĐỀ ĐỀ XUẤT VỊ TRÍ TRẠM XE ĐƯA ĐÓN NHÂN VIÊN	
Đào Xuân Hoàng Tuấn, Nguyễn Văn Nam, Nguyễn Đình Nguyên Bắc, Nguyễn Năng Anh, Đồng Mạnh Dũng , Lê Phúc Lữ	
YSC4F.240	439
XÂY DỰNG NỀN TẢNG HỌC TẬP CHẤM CODE TỰ ĐỘNG IUHCODER	
Trần Văn Nhân, Phan Chí Trung, Huỳnh Quốc Bảo, Trần Quốc Trọng, Nguyễn Hải Nam, Nguyễn Thành Tín, Ngô Gia Phát, Nguyễn Huỳnh Anh Thắng, Hoàng Thảo My, Thái Thị Hiền, Nguyễn Hữu Tình, Lê Phúc Lữ	
YSC4F.241	451
IUHEDU – CÔNG CỤ HỖ TRỢ SINH VIÊN HỌC TẬP TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH	
Trương Đức Hoàn, Đoàn Thị Thanh Hồng, Tôn Long Phước	
YSC4F.242	465
MỘT HƯỚNG TIẾP CẬN GIẢI QUYẾT VẤN ĐỀ TẮC NGHẼN TRÊN ỨNG DỤNG WEB BẰNG ĐIỆN TOÁN ĐÁM MÂY	
Đặng Trường An, Nguyễn Việt Tiến, Tôn Long Phước	

YSC4F.243	476
NGHIÊN CỨU MÔ HÌNH HỒI QUY TUYẾN TÍNH ĐỂ ĐÁNH GIÁ SỰ TƯƠNG QUAN ĐIỂM MÔN HỌC	
Nguyễn Xuân Nhựt, Nguyễn Ngọc Dung, Đoàn Văn Thắng	
YSC4F.244	483
GIẢI PHÁP XÁC THỰC NGƯỜI DÙNG THÔNG QUA MÃ OTP SỬ DỤNG SINH TRẮC HỌC GIỌNG NÓI	
Hà Bảo Anh, Đặng Thị Phúc	

YSC4F.201

GOQUICK: CÔNG CỤ HỖ TRỢ QUẢN LÝ NHÀ HÀNG DỰA TRÊN ĐIỆN TOÁN Đám Mây

TRẦN THỊ ANH THI, ĐOÀN VĂN VĨNH, NGUYỄN THANH LIÊM
Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
tranthianhthi@iuh.edu.vn, doanvinh983@gmail.com, thanhliem2n@gmail.com

Tóm tắt. Sự bùng nổ của kỹ thuật, công nghệ đã làm thay đổi cách thức vận hành, hoạt động kinh doanh trong nhiều lĩnh vực, đặc biệt là các lĩnh vực kinh doanh dịch vụ ăn uống. GoQuick giúp chủ cửa hàng nắm bắt được hiệu quả hoạt động kinh doanh của cửa hàng mình một cách nhanh chóng và đơn giản bằng các công cụ quản lý thông minh. Ứng dụng này là app mobile được thiết kế chạy trên Android, IOS với các công nghệ, công cụ hỗ trợ tối ưu nhất như Flutter, Java, Spring Boot, Socket.io, Cloud Computing. Hệ thống quản lý hoạt động kinh doanh tại quán ăn rất quan tâm đến vấn đề bảo mật. Các yêu cầu chuyển giao giữa hai bên Client – Server được bảo mật với chế độ Json Web Token. Định hướng phát triển GoQuick của chúng tôi sẽ tiến đến xây dựng một hệ sinh thái hoàn chỉnh trong lĩnh vực kinh doanh dịch vụ ăn uống.

Từ khoá. Ứng dụng di động, bảo mật thông tin, điện toán đám mây, dịch vụ ăn uống.

GOQUICK: A TOOL TO SUPPORT RESTAURANT MANAGEMENT BASED ON CLOUD COMPUTING

Abstracts. The explosion of techniques and technology has changed the way businesses operate and operate in many fields, especially the food service business. GoQuick helps store owners approach the business performance of their stores quickly and simply with smart management tools. It's a mobile app designed to run on Android and iOS with the most optimal technologies and support tools such as Flutter, Java, Spring Boot, Socket.io, Cloud Computing. The business management system at the restaurant is very concerned with security issues. Request for handover Client - Server secured with Json Web Tokens mode. Our GoQuick development orientation will be to build a complete ecosystem in the food service business.

Keywords. Mobile Application, Security Information, Cloud-Computing, Food Service.

1. GIỚI THIỆU

Trong thời đại kỹ thuật số bùng nổ như hiện nay, việc áp dụng các thiết bị công nghệ hỗ trợ quy trình hoạt động của các hoạt động kinh doanh, quán ăn không còn là việc xa lạ. Các thiết bị di động thông minh như smartphone, máy tính bảng được mọi người sở hữu một cách dễ dàng và phổ biến. Nhận thấy sự phổ biến của các thiết bị di động, xu hướng công nghệ hóa các quy trình truyền thống, nhóm chúng tôi quyết định phát triển ứng dụng giúp đặt món ăn trên thiết bị di động để quản lý các quy trình đặt món, chế biến, phục vụ và thanh toán món ăn.

Quy trình phục vụ tại các quán ăn theo cách truyền thống thường được ghi nhận các món ăn qua tờ đơn, ghim cho bộ phận bếp chế biến và thanh toán dựa trên thông tin món ăn trên tờ đơn. Ứng dụng GoQuick hiện thực quy trình trên qua xử lý phục vụ chọn bàn, gửi thông tin món ăn khách hàng yêu cầu sẽ được ghi nhận và chuyển dữ liệu sang bộ phận chế biến để thực hiện các phản hồi, xử lý thanh toán và xuất hóa đơn sẽ dựa trên thông tin dữ liệu được hệ thống ghi nhận. Việc xử lý và phản hồi qua ứng dụng sẽ tiết kiệm thời gian đặt món, phản hồi món ăn so với phương pháp truyền thống, thông tin thanh toán cũng chính xác và dễ dàng hơn so với tính tay. Ngoài ra với các dữ liệu được lưu trữ trên hệ thống, có thể thống kê dựa trên doanh số bán hàng, góp phần nào giúp người quản lý, chủ quán ăn nắm được tình hình kinh doanh tổng quan của quán ăn theo tháng hoặc một cột mốc nhất định.

Ứng dụng GoQuick được xây dựng hiện thực trên các thiết bị di động sử dụng hệ điều hành Android, được cài đặt dựa trên tệp tin cài đặt được xuất từ hệ thống. Cơ sở dữ liệu, vận hành của Server được xây dựng và quản lý bởi các dịch vụ bên thứ ba của AWS, giúp ứng dụng có thể truy xuất dữ liệu một cách an toàn và ổn định.

2. CÁC CÔNG NGHỆ LIÊN QUAN

2.1. Cloud Computing với AWS

a. Amazon Relational Database Service (Amazon RDS)

Amazon Relational Database Service (Amazon RDS) là một dịch vụ được quản lý giúp bạn dễ dàng thiết lập, vận hành và thay đổi quy cơ sở dữ liệu liên quan hệ trên đám mây. Dịch vụ này cung cấp tính năng hiệu quả về chi phí và có thể thay đổi quy mô, trong khi vẫn quản lý được các tác vụ quản trị cơ sở dữ liệu tiêu tốn nhiều thời gian, giúp bạn thoải mái tập trung vào các ứng dụng và công việc của bạn [1].

b. Amazon Simple Storage Service (S3)

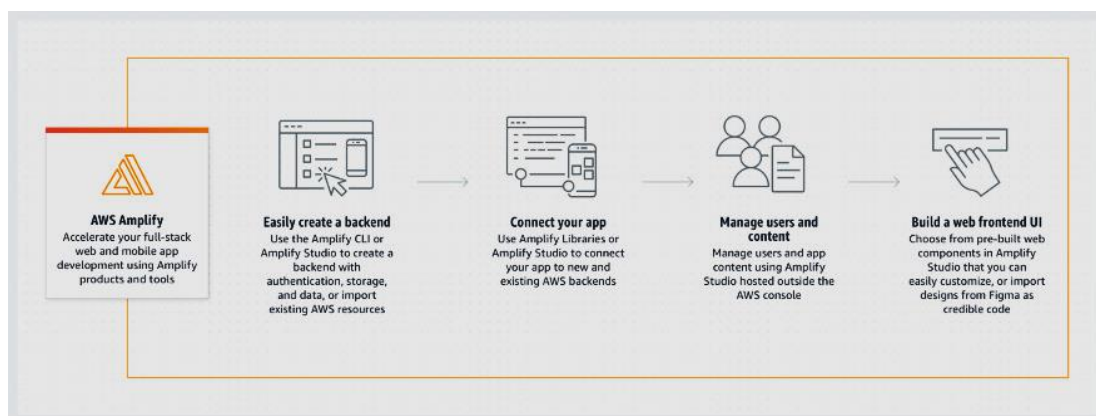
AWS S3 là dịch vụ lưu trữ đối tượng do AWS cung cấp. Đây có lẽ là dịch vụ lưu trữ được sử dụng phổ biến nhất cho người dùng AWS nhờ các tính năng như tính khả dụng, bảo mật cực cao và kết nối đơn giản với các dịch vụ AWS khác. AWS S3 có thể được sử dụng bởi mọi người với tất cả các trường hợp sử dụng như ứng dụng di động/ web, dữ liệu lớn, máy học và nhiều hơn nữa [4, 6].

c. Amazon Elastic Compute Cloud (Amazon EC2)

Đầu tiên, EC2 là viết tắt của Elastic Compute Cloud. Amazon EC2 là một máy ảo cơ bản với các thành phần phần cứng có thể tùy chỉnh và một hệ điều hành. Hệ thống cho phép bạn chạy nhiều máy tính ảo khác nhau và quản lý giống nhau bằng một phần cứng duy nhất. Elastic Compute Cloud là hệ thống dịch vụ chính và được sử dụng nhiều trong hệ sinh thái AWS không lồ. Ví dụ, hệ thống đám mây cung cấp nhiều tính năng, nó tạo điều kiện thuận lợi cho việc tính toán theo yêu cầu và mở rộng khả năng tính toán trong hệ thống đám mây Amazon. EC2 giúp bạn không cần phải mua trước phần cứng. Ngoài ra, không có thêm chi phí khi bảo trì phần cứng đã thuê. Phần cứng ảo tất cả trong một rất dễ sử dụng và cho phép bạn tạo và chạy các ứng dụng với tốc độ cao. Thích ứng với điện toán đám mây trong AWS cho phép bạn khởi chạy nhiều máy chủ ảo. Nó cũng cung cấp khả năng kiểm soát để mở rộng quy mô hoặc thu nhỏ tương ứng với tốc độ của lưu lượng truy cập trang web [5,7].

d. Amplify

AWS Amplify là một tập hợp các công cụ và tính năng được tạo ra có chủ đích, giúp nhà phát triển web và di động frontend xây dựng các ứng dụng fullstack một cách nhanh chóng và dễ dàng trên AWS với khả năng tận dụng linh hoạt các dịch vụ AWS phong phú khi các trường hợp sử dụng của bạn tiến triển (Hình 2.1). Với Amplify, bạn có thể cấu hình backend của ứng dụng web hoặc di động, kết nối ứng dụng của bạn chỉ trong vài phút, xây dựng UI frontend của web một cách trực quan và dễ dàng quản lý nội dung ứng dụng bên ngoài bằng điều khiển AWS. Cung cấp nhanh hơn và dễ dàng thay đổi quy mô mà không cần phải có chuyên môn về đám mây [3].



Hình 2.1. Sơ đồ tính năng của Amplify [3]

2.2. Framework Flutter

Flutter là một nền tảng hỗ trợ phát triển cho các ứng dụng đa nền cho hệ iOS và Android do tập đoàn Google phát triển. Nó được sử dụng vô cùng phổ biến cho nhiệm vụ tạo ra các ứng dụng gốc dành cho Google. Flutter thường bao gồm 2 thành phần chính quan trọng như sau: SDK (Software Development Kit) - Đây là một bộ sưu tập bao gồm các công cụ có thể hỗ trợ cho người dùng có thể phát triển được các ứng dụng nền của mình. Những điều này thường bao gồm các công cụ có trình để biên dịch mã thành mã gốc dành riêng cho hệ điều hành iOS và Android; Framework (UI Library based on widgets)) - Mỗi một tập hợp những thành phần giao diện của người dùng đều có thể thực hiện tái sử dụng vô cùng dễ dàng nên người sử dụng có thể cá nhân hóa tùy theo nhu cầu riêng của bản thân mình.

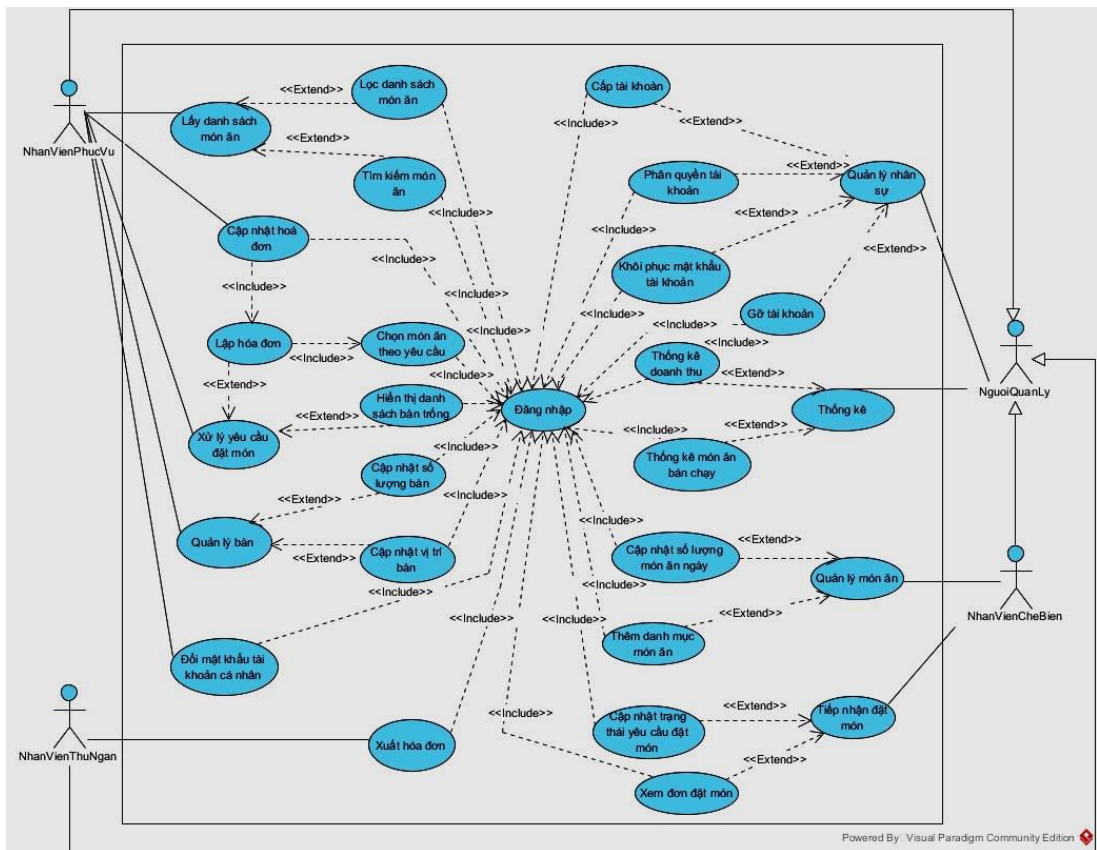
Ngoài ra, Flutter còn được biết đến với các tính năng nổi trội cho việc xây dựng các ứng dụng trên thiết bị di động. Cụ thể cho các tính năng đó là: Tính năng phát triển nhanh chóng các ứng dụng - mọi tính năng hot reload của Flutter giúp cho người dùng có thể sử dụng thử nghiệm nhanh chóng và dễ dàng hơn rất nhiều. Với khả năng xây dựng giao diện cho người dùng cộng thêm các tính năng về sửa lỗi nhanh chóng nên Flutter đang thu hút không ít người sử dụng lựa chọn. Ngoài ra, các trải nghiệm về thực hiện tải lại lần thứ hai đều rất dễ dàng mà không làm mất đi trạng thái ở trên emulator, simulator và device cho cả iOS và Android; UI đẹp mắt và có tính biểu cảm - Flutter thỏa mãn người sử dụng nhờ các widget built-in vô cùng đẹp mắt dựa theo Material Design và Cupertino (iOS-flavor), thì các API sẽ thực hiện chuyển động theo nhiều hướng phong phú, scroll tự nhiên và mượt mà nên có thể tự động nhận thức được các nền tảng cần thiết; Quá trình truy cập với nhiều tính năng và SDK native - giúp các ứng dụng của bạn sẽ trở nên sống động hơn rất nhiều nhờ vào API của platform, SDK của các bên thứ ba và native code. Từ đó, nó sẽ cho phép lập trình viên sử dụng lại được mã Java, Swift và ObjC hiện tại của mình. Nhờ vậy, các truy cập sẽ thực hiện được mọi tính năng mà SDK native dựa trên iOS và Android; Có khả năng phát triển các ứng dụng thống nhất - Nhờ sở hữu các công cụ cũng như thư viện nên người sử dụng có thể dễ dàng đưa ra được ý tưởng của mình vào chính trong cuộc sống trên hệ điều hành iOS và Android. Chính vì thế, nếu bạn là người chưa có nhiều kinh nghiệm cho mục đích phát triển các thiết bị di động thì việc lựa chọn Flutter được xem là một phương pháp dễ dàng và rất nhanh chóng trong việc xây dựng ra một ứng dụng di động tuyệt đẹp.

Cuối cùng, một ưu điểm khác mà Flutter mang lại đó là cộng đồng phát triển lớn mạnh. Flutter đang sở hữu một trong những cộng đồng phát triển vô cùng mạnh mẽ và nó được đánh giá cao nhờ khởi đầu bằng sự chất lượng tuyệt đối: Bạn có thể trao đổi mọi thắc mắc của mình về Flutter trên rất nhiều website nó sở hữu [7, 8].

3. PHÂN TÍCH HỆ THỐNG

3.1. Usecase tổng quát

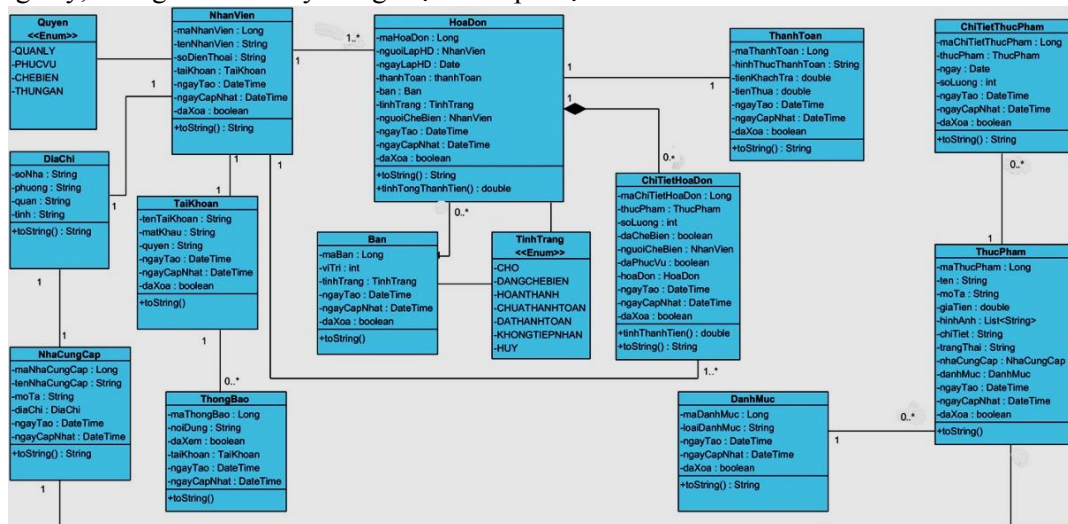
Đối tượng người dùng cho hệ thống GOQUICK gồm có người quản lý, nhân viên phục vụ, nhân viên chế biến và thu ngân. Trong đó, chúng tôi tập trung vào các chức năng chính của hệ thống là hỗ trợ người dùng là nhân viên phục vụ và nhân viên chế biến. Ngoài ra, các chức năng phụ khác cũng được đưa vào để người dùng có thể dễ dàng sử dụng hơn như hỗ trợ các chức năng thanh toán hóa đơn hay thống kê doanh số. Những chức năng này chúng tôi trình bày trong Hình 3.1.



Hình 3.1. Sơ đồ Use Case tổng quát của hệ thống GOQUICK

3.2. Danh sách Sơ đồ lớp

Từ các usecase đã phân tích ở trên thì chúng tôi đã tạo ra được lược đồ lớp của hệ thống GOQUICK. Trong đó, chúng tôi chú trọng đến các lớp như *Hoadon*, *Nhanvien* và *Thucpham*. Trong đó chúng tôi tập trung lớp *Hoadon* nhằm thực hiện các giao dịch của khách hàng đặt món, đồng thời cũng hỗ trợ cho đối tượng nhân viên chế biến dựa vào đó thực hiện các món ăn. Ngoài ra, chúng tôi cũng mở rộng các lớp *Nhanvien* dùng để phân quyền người dùng nhằm đảm bảo tính hợp lý của hệ thống. Chi tiết hơn cho các đối tượng này, chúng tôi trình bày trong lược đồ lớp được thiết kế như Hình 3.2.



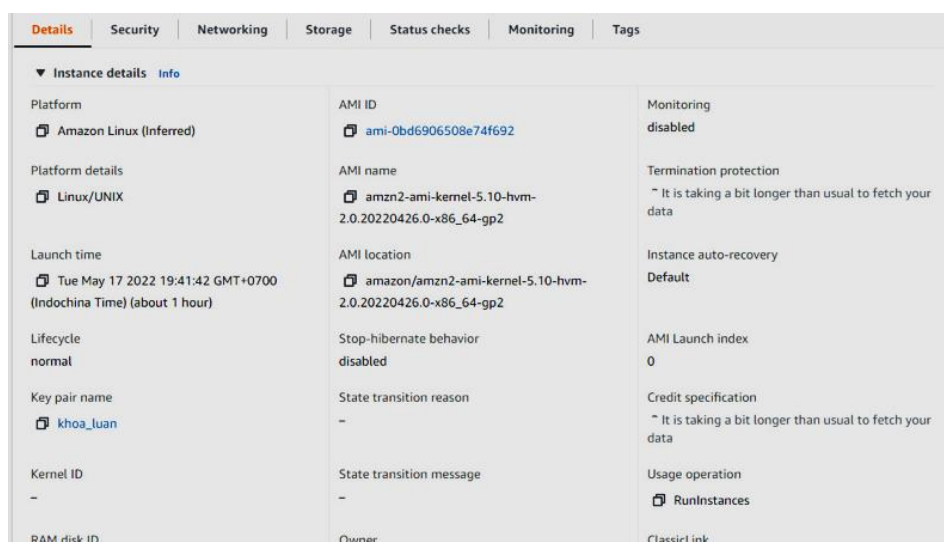
Hình 3.2. Sơ đồ Class Diagram

4. KẾT QUẢ HIỆN THỰC

4.1. Cấu hình phần cứng, phần mềm

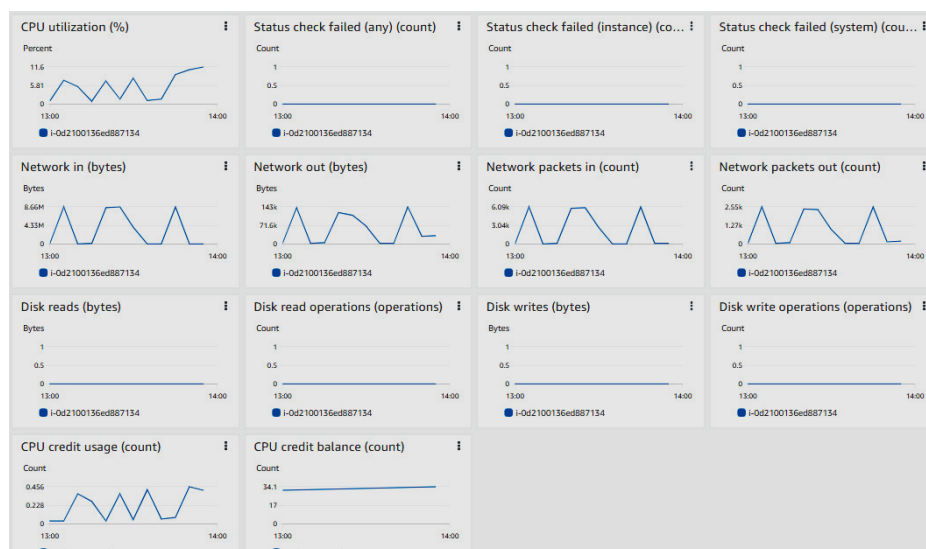
a. Cấu hình Server

Chúng tôi đã tiến hành triển khai hệ thống GOQUICK lên môi trường cloud của AWS. Trong giai đoạn này chúng tôi sử dụng cơ sở hạ tầng của AWS gồm một EC2 instance (hệ điều hành Linux, phiên bản t2.small), được cài sẵn môi trường docker để chạy ứng dụng của chúng tôi (Hình 4.1).



Hình 4.1. Cấu hình máy ảo AWS EC2 cho hệ thống GOQUICK

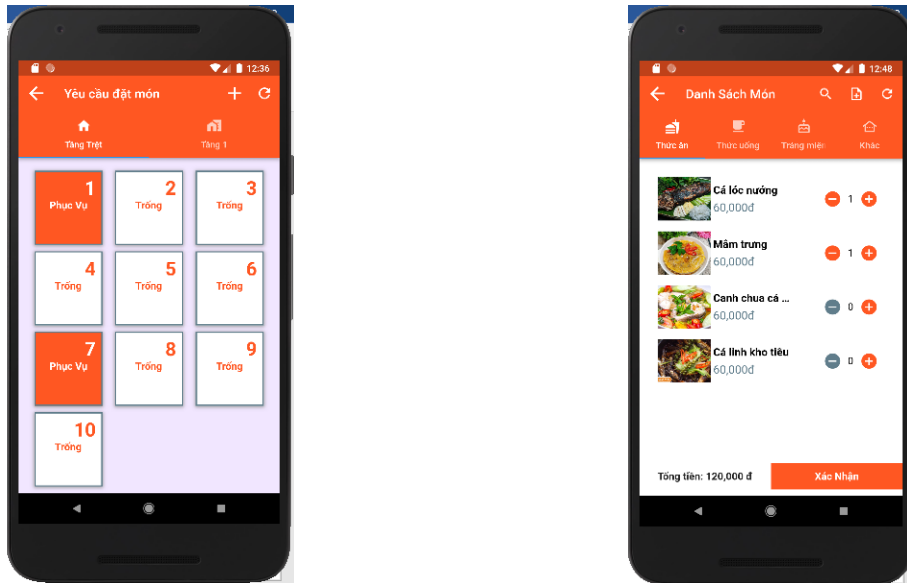
Hệ thống chạy trên EC2 instance hoạt động ổn định. Không có request nào trả về trạng thái failed. Những thông số về tài nguyên sử dụng CPU (CPU credit balance) cũng như dung lượng đường truyền (Network packets) được đảm bảo trong các khoảng thời gian khi hệ thống tăng tốc xử lý. Tuy nhiên đây là phiên bản thử nghiệm nên sẽ dùng các instance có hiệu suất thấp hơn phiên bản triển khai trong thực tế (Hình 4.2).



Hình 4.2 Các thông số Monitoring của hệ thống GOQUICK được triển khai

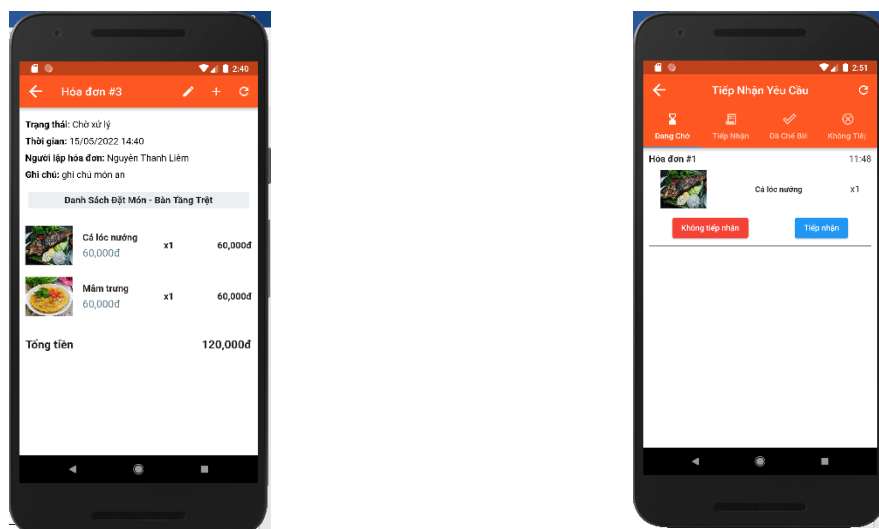
4.2. Giao diện hiện thực màn hình

Để thuận tiện cho việc quản lý các dịch vụ nhà hàng, chúng tôi đã trực quan hóa các chức năng trên hệ thống ứng dụng di động (app) cũng như trên hệ thống Web. Người dùng có thể dễ dàng sử dụng các chức năng đặt món hoặc xử lý các chức năng thanh toán trên ứng dụng này (Hình 4.2.1).



Hình 4.2.1 Giao diện thực hiện chức năng đặt món trên GOQUICK

Ngoài ra, hệ thống cũng hỗ trợ các chức năng cho nhân viên chế biến có thể nhận các món ăn được đặt nhằm chủ động trong chế biến cũng như phản hồi kịp thời đến khách hàng tình trạng của các món ăn với nguyên liệu hiện có. Đồng thời hỗ trợ thanh toán các hóa đơn của khách hàng một cách dễ dàng và chính xác (Hình 4.2.2).



Hình 4.2.2 Giao diện thực hiện chức năng tiếp nhận món ăn trên GOQUICK

5. KẾT LUẬN

Bài báo này chúng tôi đưa giải pháp hỗ trợ quy trình phục vụ tại các quán ăn một cách tiện lợi, hiện đại hơn nhằm nâng cao hiệu suất phục vụ, cũng như quản lý đơn đặt món của quán ăn mà không quá tốn kém về chi phí. Cụ thể, giải pháp của chúng tôi triển khai trên các thiết bị di động Android và IOS, một thiết bị công nghệ phổ biến và phù hợp với mọi người. Để phát triển phần mềm chúng tôi sử dụng Flutter, một ngôn ngữ lập trình dành cho di động mới, hỗ trợ đa nền tảng và có tiềm năng, cộng đồng hỗ trợ to lớn. Để quản lý dữ liệu, các kết nối API giữa phần mềm và cơ sở dữ liệu chúng tôi triển khai Spring Boot JPA, cũng như JWT, Spring Boot Security bảo mật hệ thống cơ sở dữ liệu. Ngoài ra theo xu thế mới với sự phát triển mạnh mẽ của dịch vụ Cloud Computing, chúng tôi triển khai cơ sở dữ liệu với hệ thống RDS. Ngoài ra, sử dụng các dịch vụ EC2, S3, Amplify của nhà cung cấp Amazon để hoàn thiện phần mềm. Tuy nhiên do sự hạn chế về thời gian, cũng như việc khảo sát các hệ thống còn chưa đầy đủ, phần mềm còn hạn chế nhiều tính năng hỗ trợ các nghiệp vụ kinh doanh khác của nhà hàng. Cụ thể những quy trình, nghiệp vụ của các dịch vụ chưa diễn ra thực sự mượt mà với nhiều kịch bản cho tình huống khác nhau, hay còn chưa triển khai phần mềm trên các hệ thống nhà hàng có các nghiệp vụ tương tự. Đây là một hạn chế của phần mềm GOQUICK, trong tương lai chúng tôi sẽ dành nhiều thời gian, cũng như tìm kiếm sự hợp tác để phát triển phần mềm gắn với thực tế, hỗ trợ đa dạng các hoạt động kinh doanh hơn.

TÀI LIỆU THAM KHẢO

- [1] Amazon RDS (Sep 2021), <https://www.techtarget.com/searchaws/definition/Amazon-Relational-Database-Service-RDS>
- [2] Amazon Web Service (Jan 2020), https://aws.amazon.com/vi/rds/?nc2=h_ql_prod_db_rds
- [3] AWS Amplify, <https://aws.amazon.com/vi/amplify/>
- [4] Fast and Cost-Effective Image Manipulation with Serverless Image Handler (NOV 2020), <https://aws.amazon.com/vi/blogs/architecture/fast-and-cost-effective-image-manipulation-with-serverless-image-handler/>
- [5] Introduction to Amazon EC2 (Nov 2015), https://s3-us-west-2.amazonaws.com/uw-s3-cdn/wp-content/uploads/sites/149/2018/12/28193450/Curtis-Bray_Amazon_Introduction-to-Amazon-EC2.pdf
- [6] Introduction to AWS Simple Storage Service (Dec 2021), <https://www.geeksforgeeks.org/introduction-to-aws-simple-storage-service-aws-s3/>
- [7] What is AWS EC2 (Nov 2020), <https://www.jigsawacademy.com/blogs/cloud-computing/what-is-aws-ec2/>
- [8] Why Flutter For Mobile App Development (Feb 2022), <https://doit.software/blog/why-flutter-for-mobile-app-development#screen4 ml>

YSC4F.202

ĐÁNH GIÁ MỨC ĐỘ TIÊU CỰC BÌNH LUẬN TRÊN MẠNG XÃ HỘI

PHẠM MINH TUẤN, TRƯƠNG NGUYỄN DUY TÂN, NGUYỄN QUANG QUỐC,
ĐÀO DUY TRƯỜNG, NGUYỄN THỊ THANH HÒA, ĐẶNG THỊ PHÚC

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
mtp92621@gmail.com, truongnguyenduytan.2016@gmail.com, quangquoc2001@gmail.com,
aotruong123@gmail.com, thanhhoanguyen.iuh@gmail.com, phucdt@iuh.edu.vn

Tóm tắt. Hiện nay, do việc thể hiện quan điểm cá nhân trên mạng xã hội ngày càng dễ dàng hơn nên không thể tránh khỏi những nhóm người có hành vi bình luận tiêu cực trên mạng. Trong dự án nghiên cứu này, chúng tôi sẽ thu thập dữ liệu từ mạng xã hội Facebook và xây dựng mô hình để dự đoán mức độ tiêu cực bình luận. Bộ dữ liệu của chúng tôi thu thập gồm những nội dung bình luận và gán 7 loại nhãn: (1) quảng cáo, (2) thông tin không liên quan, (3) không tiêu cực, (4) tiêu cực nhẹ, (5) thông tin tiêu cực, (6) tiêu cực kỳ thị, (7) tiêu cực từ ngữ thô tục. Với mỗi bình luận được gán nhãn dựa vào kinh nghiệm người gán nhãn và sử dụng các kỹ thuật xử lý ngôn ngữ tự nhiên đưa vào mô hình. Chúng tôi áp dụng kỹ thuật học sâu để xây dựng mô hình phân loại các bình luận này để có độ chính xác cao nhất trên bộ dữ liệu trên 20000 bình luận đã được gán nhãn. Mô hình được triển khai lên các hệ thống mạng xã hội giúp tự động quản lý, chọn lọc bình luận phù hợp, loại bỏ hoặc ghi nhận những cá nhân có bình luận không đúng mục đích.

Từ khóa. mức độ tiêu cực bình luận, học sâu, xử lý ngôn ngữ tự nhiên.

ASSESS THE NEGATIVE LEVEL OF COMMENTS ON SOCIAL NETWORKS

Abstract. Currently, because it is easier to express personal views on social networks, groups of people inevitably have negative comments online. In this research project, we will collect data from the social network Facebook and build a model to predict the level of negative comments. Our data set includes comments and assigns 7 types of labels: (1) advertising, (2) irrelevant information, (3) not negative, (4) mildly negative, (5) negative information, (6) negative stigma, (7) negative profanity. Each comment is labeled based on the tagger's experience and using natural language processing techniques to feed into the model. We apply deep learning techniques to build a classification model of these comments to have the highest accuracy on a dataset of over 20000 labeled comments. The model is deployed to social networking systems to automatically manage and select appropriate comments, and remove or recognize individuals with inappropriate comments.

Keywords. negative comments, deep learning, natural language processing.

1. TỔNG QUAN

Xử lý ngôn ngữ tự nhiên (natural language processing - NLP) là một nhánh của Trí tuệ nhân tạo, tập trung vào việc nghiên cứu sự tương tác giữa máy tính và ngôn ngữ tự nhiên của con người, dưới dạng giọng nói (speech) hoặc văn bản (text). Mục tiêu là giúp máy tính hiểu và thực hiện hiệu quả những nhiệm vụ liên quan đến ngôn ngữ của con người như: tương tác giữa người và máy, cải thiện hiệu quả giao tiếp giữa con người với con người, hoặc đơn giản là nâng cao hiệu quả xử lý văn bản và lời nói...

Ngày nay, tình trạng xuất hiện những bình luận tiêu cực trên Facebook gây ảnh hưởng xấu không nhỏ đến cuộc sống con người, thậm chí là suy sụp tinh thần, tự sát. Cũng chỉ vì những bình luận chứa từ ngữ thô tục, chứa nội dung nhạy cảm; bình luận có xu hướng lôi kéo, dụ dỗ nhằm mục đích lừa đảo; bình luận kỳ thị, chê bai, đánh giá, chỉ trích người khác vô cớ; bình luận chứa từ ngữ không rõ ràng, gây hiểu lầm và vô

tình tạo ra những mâu thuẫn... Ở Việt Nam, Tháng 6.2013, P.U.N., nữ sinh lớp 12 của một trường THPT ở Đà Nẵng đã uống thuốc an thần tự tử. Người nhà N. may mắn phát hiện và đưa đi cấp cứu kịp thời. Nguyên nhân được tiết lộ sau đó, chỉ vì N. bị trang fanpage B.M.T.C.C.H.Đ.T. viết bài vu khống để thỏa mạ, bôi nhọ lên Facebook. Nhiều dân mạng đã a dua chỉ trích, xúc phạm N. thậm tệ. Quá mệt mỏi, N. tìm đến cái chết. Câu chuyện được trích dẫn ở trên cũng không ngoại lệ, bởi có không ít nữ sinh tìm đến cái chết chỉ vì mạng xã hội.

Để giải quyết tình trạng trên, chúng tôi đề xuất xây dựng mô hình đánh giá mức độ tiêu cực bình luận trên mạng xã hội, cụ thể trong bài báo này chúng tôi sẽ khai thác dữ liệu bình luận trên nền tảng Facebook. Kết hợp với ứng dụng bài toán phân loại trong xử lý ngôn ngữ tự nhiên để đưa ra đánh giá mức độ trên bình luận đó.

Vấn nạn bắt nạt trên mạng xã hội là việc bắt nạt gián tiếp thông qua việc sử dụng các công nghệ kỹ thuật số. Việc này hiện đã và đang diễn ra trên các nền tảng mạng xã hội, nhắn tin, chơi game và điện thoại (các thiết bị) di động. Đây là hành vi lặp đi lặp lại nhằm mục đích khiến người bị bắt nạt tức giận, sợ hãi hoặc xấu hổ. Ví dụ: lan truyền những thông tin không đúng sự thật về một ai đó trên mạng xã hội, các tin nhắn hoặc lời đe dọa gây tổn thương về mặt tinh thần qua các nền tảng nhắn tin, mạo danh và thay mặt ai đó để gửi các thông điệp gây ác ý cho người khác.

Bắt nạt trực tiếp và bắt nạt gián tiếp thông qua nền tảng mạng xã hội, nhắn tin,... thường có thể xảy ra song song với nhau. Tuy nhiên việc bắt nạt trực tuyến sẽ để lại dấu ấn kỹ thuật số - một hồ sơ có thể chứng minh là hữu ích và cung cấp bằng chứng giúp ngăn chặn hành vi xâm hại.

Nhóm nghiên cứu sẽ tìm đến những bài đăng trên facebook đề cập về các chủ đề được xã hội quan tâm nhằm lọc ra những bình luận có thái độ tiêu cực. Nhóm nhận thấy với các bình luận tiêu cực trên mạng xã hội sẽ tương đồng với những bình luận bắt nạt, công kích trong nạn bắt nạt trên mạng. Những bình luận đó sẽ là cơ sở để xây dựng nên mô hình phân loại bình luận.

Bài toán phân loại văn bản (Text Classification) là bài toán thuộc nhóm học có giám sát (Supervised learning) trong học máy. Bài toán này yêu cầu dữ liệu cần có nhãn (label). Mô hình sẽ học từ dữ liệu có nhãn đó, sau đó được dùng để dự đoán nhãn cho các dữ liệu mới mà mô hình chưa gặp.

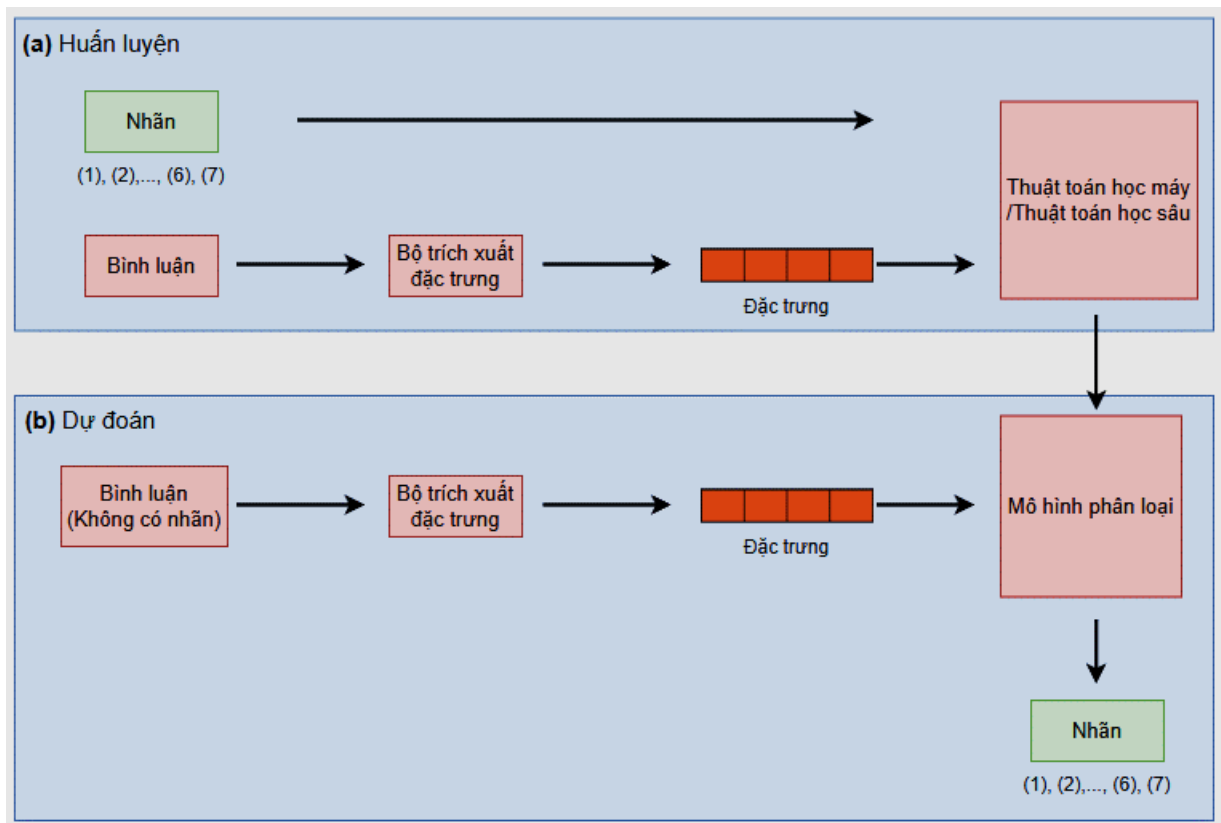
Về các ứng dụng thực tế, bài toán đã được các doanh nghiệp lớn như: Google, Facebook, Tiki, Shopee, Lazada... ứng dụng để giải quyết các vấn đề của riêng họ và đem lại kết quả tốt nên họ vẫn cố gắng duy trì song song phát triển phương pháp hướng đến kết quả tốt nhất. Cụ thể, ứng dụng của google trong việc lọc email spam của gmail, bộ lọc bình luận phù hợp nhất của facebook trên mạng xã hội, hệ thống phân loại thẻ đề xuất tìm kiếm của google, các hệ thống nghiên cứu, tiếp thị sản phẩm của các sàn thương mại điện tử theo lịch sử tìm kiếm của khách hàng...

Nhận thấy tiềm năng của bài toán phân loại văn bản trong xử lý ngôn ngữ tự nhiên. Nhóm sẽ ứng dụng cách làm của bài toán để giải quyết vấn đề mà nhóm hướng tới là lọc ra các bình luận công kích trên mạng xã hội. Với kỳ vọng bằng cách xây dựng được mô hình phân loại này có thể giúp facebook có thể phát hiện những bình luận công kích làm suy giảm nạn bắt nạt trên mạng xã hội.

2. CƠ SỞ LÝ THUYẾT

2.1. Lý thuyết về bài toán phân loại văn bản

Để phân loại một đoạn bình luận trên mạng xã hội cần phải đi qua các bước xử lý cơ bản cho bài toán phân loại văn bản. Cụ thể, sử dụng kỹ thuật học máy để giải quyết bài toán phải đi qua các bước xử lý và xây dựng này.



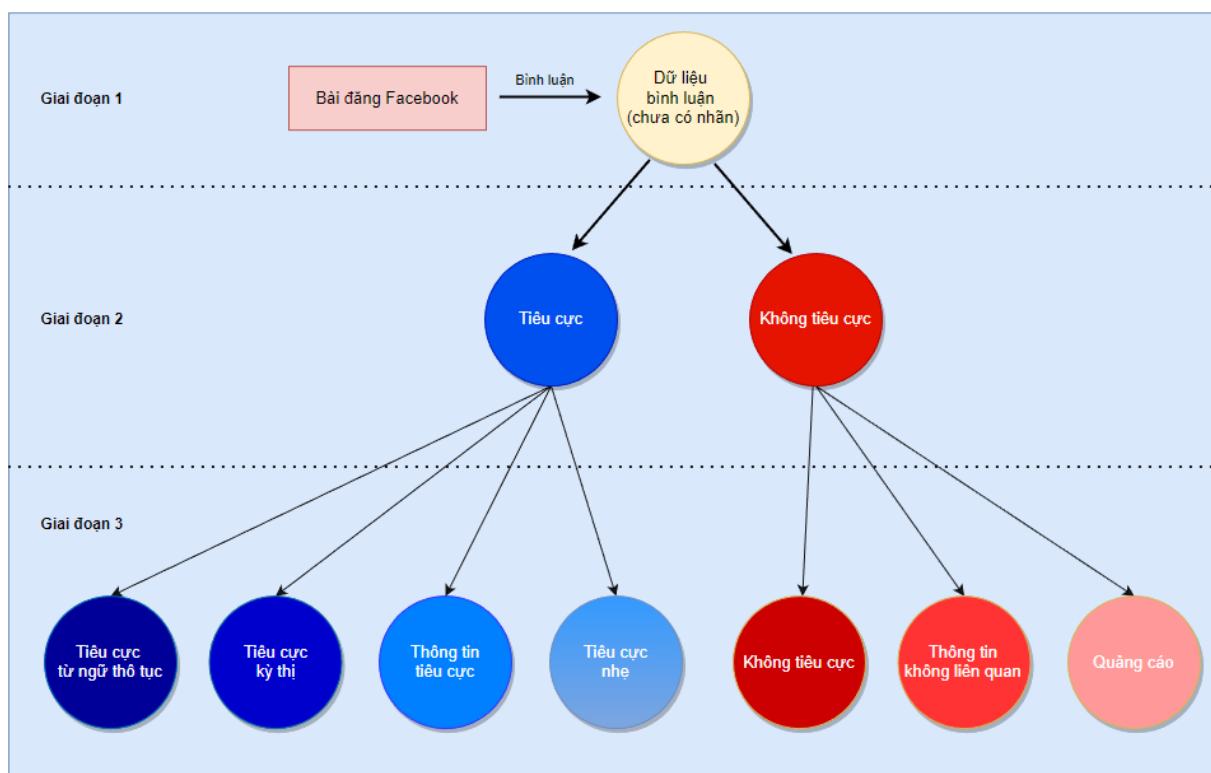
Hình 1: Các bước giải quyết bài toán

Giai đoạn (a): Huấn luyện (training) là giai đoạn mô hình sẽ học các dữ liệu được truyền vào để phục vụ cho mục đích phân loại bình luận. Dữ liệu dùng từ giai đoạn (a) đã được gán nhãn sẵn thành 7 loại mức độ bình luận khác nhau. Tiếp đó dữ liệu đưa qua bộ trích xuất đặc trưng giúp cho máy tính có thể hiểu được các dòng bình luận dưới dạng số. Lúc này những dòng bình luận được đại diện là những vector. Thuật toán máy học sẽ học và tối ưu các tham số để đạt kết quả tốt trên bộ dữ liệu. Nhãn của bình luận sẽ là cơ sở để đánh giá mô hình này tốt hay xấu và dựa vào đó mà tối ưu.

Giai đoạn (b): Dự đoán (prediction) là giai đoạn đưa dữ liệu cần dự đoán đi qua những bước xử lý chuyển đổi số như giai đoạn (a) nhưng sẽ dùng mô hình phân loại đã huấn luyện xong để đưa ra kết quả dự đoán.

2.2. Thu thập và gán nhãn dữ liệu

Trong lĩnh vực học máy, bài toán phân loại văn bản thuộc về một dạng kỹ thuật riêng là học có giám sát. Điều kiện tối thiểu của việc học có giám sát, ngoài vấn đề đòi hỏi cần có dữ liệu mà dữ liệu được thu thập phải được gán nhãn sẵn mới có thể ứng dụng được kỹ thuật học có giám sát. Hiểu được yêu cầu của phương pháp giải quyết bài toán, nhóm chia việc thu thập và gán nhãn dữ liệu thành ba giai đoạn chính.



Hình 2: Các giai đoạn thu thập và gán nhãn dữ liệu

Giai đoạn một, thu thập dữ liệu bằng cách sử dụng selenium để tiến hành lấy dữ liệu từ mạng xã hội Facebook. Lý do chính nhóm quyết định chọn mạng xã hội Facebook, xét về bối cảnh ứng dụng ở Việt Nam theo thống kê mới nhất năm 2022, có 70,4 triệu người dân và chiếm gần 72% tổng số dân Việt Nam đang sử dụng mạng xã hội Facebook. Đồng thời, nhóm áp dụng phương pháp quan sát trong quá trình chọn các bài đăng trên Facebook để lấy dữ liệu. Về mặt lý thuyết, phương pháp quan sát là phương pháp thu thập dữ liệu bằng cách ghi lại có kiểm soát hoặc các hành vi ứng xử của con người. Nhóm dựa theo các tin tức được lên hàng ngày, lấy chủ đề những tin đó làm trọng tâm để tìm đến những bài báo cũng đề cập tin tức tương tự. Cụ thể trong trường hợp này là quan sát về các bình luận của người dùng trong các bài viết khác nhau để chọn ra các bình luận trong các bài viết có xu hướng gây tranh cãi về một sự kiện hay vấn đề đang được cộng đồng mạng quan tâm và tương tác.

Giai đoạn hai, tiến hành gán nhãn dữ liệu được thu thập phục vụ cho kỹ thuật học có giám sát. Về việc phân loại các nhãn trong bộ dữ liệu, cần dựa vào sự tìm hiểu đối với bộ dữ liệu để rút ra các từ khóa, đặc trưng có khả năng xuất hiện trong các chủ đề cụ thể. Cụ thể, nhóm sử dụng các nhận định bản thân về những bình luận tiêu cực để phân loại bình luận đó. Bởi vì phương pháp gán nhãn của nhóm phụ thuộc vào cảm nhận của con người nên nhóm đã thực hiện công đoạn gán nhãn bắt đầu từ hai nhãn đơn giản. Ý nghĩa hai nhãn này đại diện, một - các nhãn tiêu cực (gồm các nhãn: tiêu cực nhẹ, thông tin tiêu cực, tiêu cực kỳ thị, tiêu cực từ ngữ thô tục) và hai - các nhãn không tiêu cực (gồm các nhãn: quảng cáo, thông tin không liên quan, không tiêu cực). Cách làm bắt đầu từ những nhãn cơ bản giúp cho việc đánh giá từ nhận định ít bị nhầm lẫn giữa các nhãn, giảm dữ liệu bị lỗi. Đồng thời từng thành viên trong nhóm cũng tham gia gán nhãn dữ liệu để các dữ liệu đưa vào mô hình được đánh giá khách quan nhất có thể.

Giai đoạn ba, sau khi đã đưa dữ liệu thành hai nhãn cơ bản. Tiếp theo, tiến hành xác định rõ hơn các mức độ tiêu cực của bình luận và chia nhỏ dần theo các mức khác nhau. Các mức cụ thể được gán như: quảng cáo, thông tin không liên quan, không tiêu cực, tiêu cực nhẹ, thông tin tiêu cực, tiêu cực kỳ thị, tiêu cực từ ngữ thô tục. Những nhãn quảng cáo, thông tin không liên quan có vai trò giúp mô hình có thể tự động nhận biết loại những thông tin không cần thiết trong nhãn không tiêu cực. Nhãn không tiêu cực sẽ đại diện các dạng bình luận mặc dù có đề cập đến chủ đề bài đăng nhưng theo hướng tích cực hoặc góp ý. Còn lại những nhãn tiêu cực nhẹ, thông tin tiêu cực, tiêu cực kỳ thị, tiêu cực từ ngữ thô tục là các nhãn mức độ đại diện

cho bình luận công kích. Đặc biệt, các nhãn tiêu cực kỳ thị, tiêu cực từ ngữ thô tục còn đại diện cho các kiểu công kích khác nhau trên mạng xã hội.

Ưu điểm phương pháp gán nhãn là tăng độ chính xác khi gán nhãn dựa trên nhận định và tăng tính khách quan của dữ liệu gán nhãn. Nhược điểm là với một bộ dữ liệu, ta phải phân loại nhiều lần nên phải đầu tư nhiều thời gian vào phân gán nhãn.

2.3. Tiền xử lý dữ liệu

Tiền xử lý dữ liệu là một bước rất quan trọng trong việc giải quyết bất kỳ vấn đề nào trong lĩnh vực học máy. Hầu hết các bộ dữ liệu được sử dụng trong các vấn đề liên quan đến học máy cần được xử lý, làm sạch và biến đổi trước khi một thuật toán học máy có thể được huấn luyện trên những bộ dữ liệu chuẩn bị. Mục tiêu chính, giúp trích được các đặc trưng cần thiết cho mô hình huấn luyện. Bước tiền xử lý có thể được xem là một trong những bước quan trọng để cải thiện mô hình dự đoán và tốn rất nhiều thời gian vì phải quay lại bước này xử lý.

Chi tiết hơn, trong các bước tiền xử lý dữ liệu bao gồm: (1) xử lý ký tự trống, chuyển thành chữ thường, (2) xử lý icon được dùng trên facebook, (3) chuẩn hóa Unicode, (4) chuẩn hóa kiểu gõ, (5) xử lý các ký tự đặc biệt trong một câu, (6) xử lý teencode, (7) xử lý loại bỏ các từ không cần thiết, (8) trích xuất các từ cần thiết đưa vào mô hình.

Đối với các bước xử lý ký tự trống, chuyển thành chữ thường, chuẩn hóa Unicode, chuẩn hóa kiểu gõ, xử lý ký tự đặc biệt là các bước cơ bản mà bất bài toán xử lý ngôn ngữ tự nhiên nào cũng phải thực hiện. Các bước xử lý làm cho các từ trong câu bình luận đồng bộ với nhau. Bởi vì sẽ có các trường hợp mặc dù từ tương đồng ý nghĩa nhưng do khác một số ký tự hoặc định dạng Unicode khiến cho những từ tương đồng thành những từ khác nhau.

Ngoài ra còn có các bước xử lý đặc thù mà nhóm xử lý riêng cho bài toán như: xử lý icon, xử lý teencode, xử lý loại bỏ các từ không cần thiết, trích xuất các từ cần thiết đưa vào mô hình. Do bộ dữ liệu của nhóm là dòng bình luận trên mạng xã hội facebook nên hiển nhiên khi bình luận sẽ thêm các yếu tố icon hoặc teencode. Ý nghĩa các bước xử lý icon, teencode giúp loại bỏ các icon có trong bình luận và chuyển các từ teencode sang chữ thường. Đây là công đoạn tốn nhiều thời gian trong tiền xử lý nhất vì không chỉ xử lý một lần mà cần phải kết hợp trực quan hóa để xử lý rất nhiều lần. Thêm vào đó cũng cần có sự hiểu biết về các ký tự teencode để có thể chuyển đổi chính xác nghĩa của chúng. Cuối cùng, xử lý loại bỏ các từ không cần thiết, trích xuất các từ cần thiết đưa vào mô hình là cần thiết làm cho mô hình huấn luyện khôn ngoan hơn.

2.4. Hướng tiếp cận mô hình

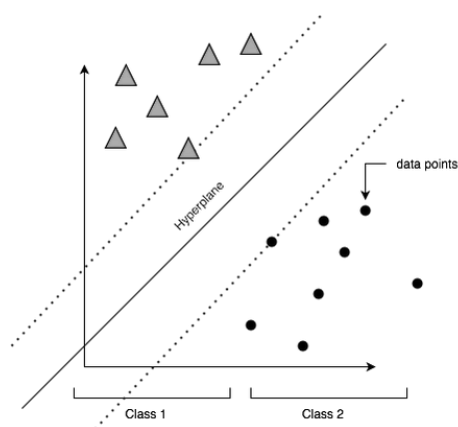
Bằng cách thử qua nhiều loại mô hình khác nhau sẽ đưa ra thông số khác nhau, nhóm sẽ chọn ra mô hình tốt nhất để phân loại bình luận. Cụ thể, các loại mô hình này vừa có cả ở lĩnh vực học máy lẫn học sâu. Tuy các loại mô hình này khác nhau về lĩnh vực nhưng nó chỉ khác ở cách trích xuất đặc trưng từ các dòng bình luận còn các tiền xử lý thì như nhau.

2.5. Mô hình xây dựng

2.5.1. Support Vector Machine

Support Vector Machine (SVM) là một trong những thuật toán classification phổ biến nhất thường được dùng trong việc giải quyết các dạng toán như binary class (hay còn gọi là bài toán 2 lớp). Tuy nhiên, trong bài này chúng tôi sẽ dùng SVM để giải quyết dạng toán multi-class (nhiều lớp).

SVM dựa trên việc tìm kiếm các siêu mặt phẳng (hyperplanes) để làm đường ranh giới phân chia các lớp dữ liệu sao cho khoảng cách từ điểm gần nhất của mỗi lớp tới đường phân chia là lớn nhất, từ đó việc phân chia giữa các lớp sẽ được rạch ròi rõ ràng hơn. Điều này có nghĩa SVM có thể hoạt động trong cả các tập lớp và cho bất kì vector mã hoá của bất kì dữ liệu nào vì mục tiêu của SVM là đo lường sự phân tách của dữ liệu chứ không phải là so khớp các đặc trưng của dữ liệu [8].

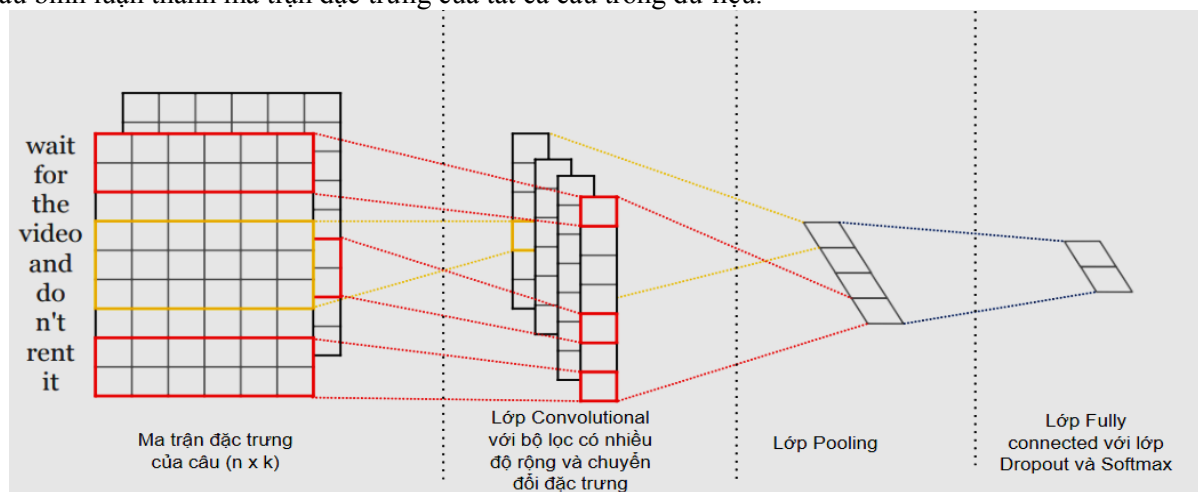


Hình 3: Cách hoạt động SVM đối với 2 lớp

Trong bài toán phân loại văn bản, các từ được chuyển hoá thành các vector, từ đó ta có thể dễ dàng sử dụng thuật toán SVM.

2.5.2. Convolutional Neural Network

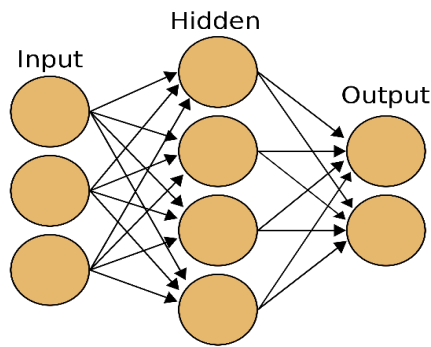
Convolution Neural Network (CNN) là một trong những mô hình thuộc lĩnh vực Học sâu. CNN đơn giản tập trung vào những chữ quan trọng có trong câu. Mặc dù được sử dụng phổ biến trong thị giác máy tính hay xử lý ảnh nhưng CNN vẫn có thể sử dụng trong xử lý ngôn ngữ tự nhiên khi chỉ cần chuyển đổi dữ liệu câu bình luận thành ma trận đặc trưng của tất cả câu trong dữ liệu.



Hình 4: Cấu trúc cơ bản CNN [9]

Cấu trúc CNN bao gồm 3 thành phần chính là lớp convolutional, lớp pooling, lớp fully connected:

- Đối với lớp Convolutional sẽ tính “element-wise” của ma trận con từ ma trận đặc trưng có kích thước được điều chỉnh trước với ma trận kernel có cùng kích thước ma trận con. Tùy vào dạng ma trận kernel ta nhận được các đặc trưng khác nhau. Với mỗi phép tính ta sẽ được một đặc trưng được trích xuất. Sau khi đã tạo ra ma trận đặc trưng mới sẽ đi qua hàm kích hoạt đưa ra đầu ra của lớp Convolutional.
- Lớp Pooling sẽ giúp CNN giảm số lượng tính toán nhưng vẫn có thể trích xuất đặc trưng và giảm số chiều ma trận đặc trưng. Điểm bất lợi mang lại mặc giảm được số lượng tính toán nhưng phải chấp nhận có thể mất đi nhiều đặc trưng quan trọng lúc trích xuất hơn lớp Convolutional.
- Lớp Fully connected sẽ có dạng như một mạng lưới nơ-ron thông thường có đầu vào lúc này được nối lại tạo thành vector chứa các đặc trưng trích xuất từ câu và đầu ra là các lớp cần dự đoán. Thành phần bao gồm 3 phần: lớp đầu vào, lớp ẩn, lớp đầu ra.

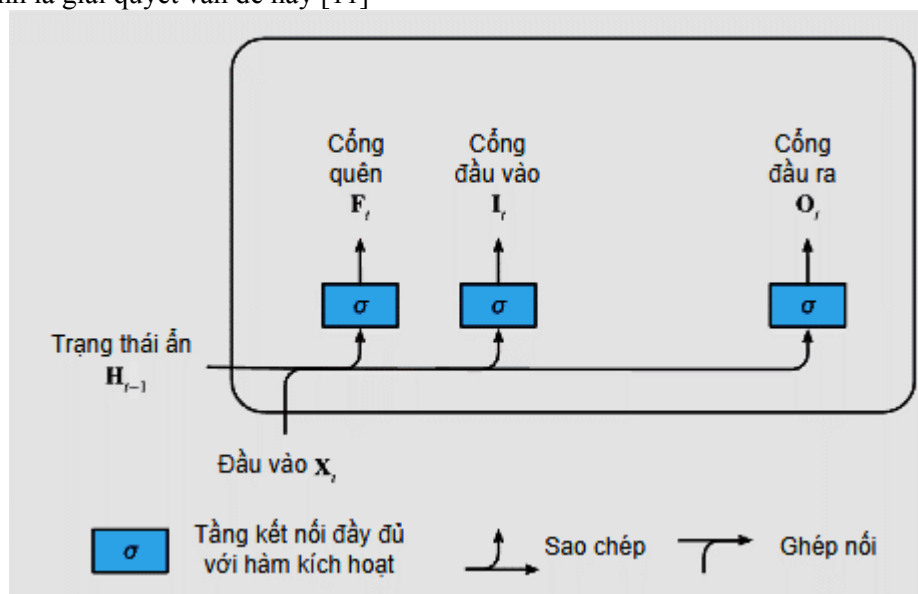


Hình 5: Cấu trúc cơ bản lớp Fully connected

2.5.3. Long short-term memory

Các mạng nơ ron truyền thống và mạng nơ ron tích chập thường không đủ khả năng để xử lý dữ liệu tuần tự, một vấn đề có ý nghĩa trong các bài toán xử lý ngôn ngữ tự nhiên khi sự tuần tự của các câu từ tạo nên ngữ cảnh quyết định ý nghĩa của câu.

Cấu trúc của RNN phù hợp để xây dựng mô hình dữ liệu tuần tự như văn bản hoặc chuỗi thời gian, các mạng nơ ron của RNN sử dụng trạng thái trước đó để lưu và đầu vào hiện tại để cập nhật trạng thái của chính nó. Tuy nhiên nó có một vấn đề mô hình dễ bị biến mất hoặc bùng nổ gradient khi lượng văn bản quá lớn trong khi huấn luyện từ đó gây ảnh hưởng lớn đến kết quả của mô hình. LSTM là một mô hình với nhiệm vụ chính là giải quyết vấn đề này [11]



Hình 6: Cấu trúc LSTM [13]

Cấu trúc LSTM với 4 thành phần:

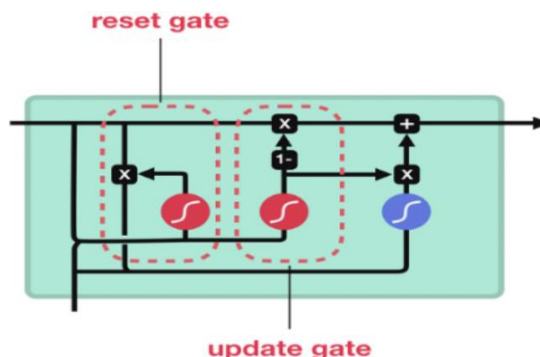
- * Memory unit: lưu trữ tạm thời trạng thái của mạng
- * Input gate: điều chỉnh và kích hoạt đầu vào các cells
- * Output gate: điều chỉnh và kích hoạt đầu ra của các cells
- * Forget gate: đặt lại bộ nhớ của cells

Bằng cách chuyển hoá văn bản thành các số nguyên hoặc thành vector, LSTM có thể huấn luyện từ dữ liệu văn bản. Với bài toán phân loại văn bản, hàm activation là hàm softmax để ra output nhiều lớp, và vì là bài toán phân loại nhiều lớp nên hàm loss được sử dụng categorical_crossentropy

2.5.4. Gated recurrent unit

Gated recurrent unit (GRU) là một phiên bản của LSTM với nguyên tắc hoạt động tương tự LSTM, nhưng có cấu tạo đơn giản hơn. Thông thường LSTM có thể lưu trữ thông tin với dữ liệu dài hơn so với GRU. Tuy nhiên do cấu tạo đơn giản của mình, GRU thường xử lý nhanh hơn LSTM và có thể dễ dàng sử dụng để xây dựng các mạng có cấu trúc phức tạp.

GRU kết hợp Cell state và Hidden state thành 1, do đó nó chỉ có 2 input và 1 output. Ngoài ra, trong GRU chúng ta sử dụng 2 gate là Reset gate và Update gate để quyết định việc lưu trữ và loại bỏ thông tin.



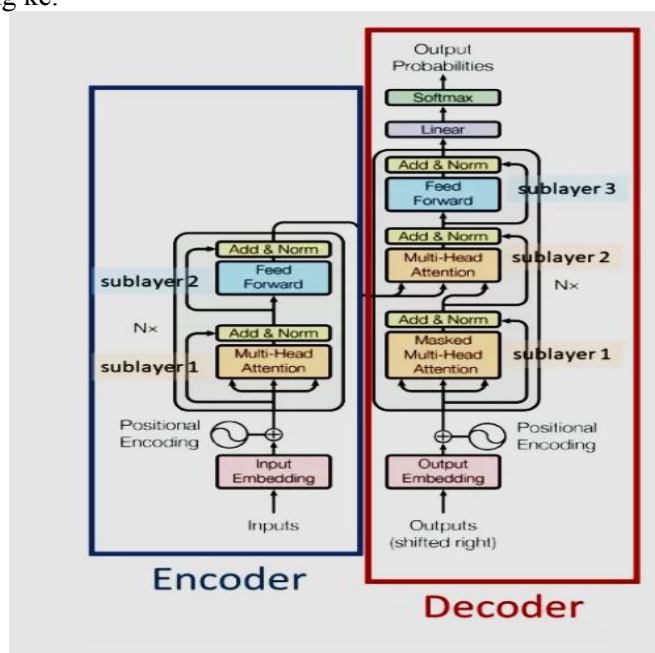
Hình 7: Cấu trúc GRU [14]

2.5.5. Bidirectional-LSTM và Bidirectional-GRU

Bidirectional là kỹ thuật xây dựng mô hình bằng cách kết nối 2 lớp ẩn song song với nhau có cùng đầu ra. Cách làm này giúp lớp đầu ra có thể lấy được thông tin trạng thái trước đó và sắp tới được áp dụng nhiều với cấu trúc RNN, LSTM, GRU.

2.5.6. Transformer

Transformer là loại mô hình học sâu được thiết kế để phục vụ giải quyết nhiều bài toán trong xử lý ngôn ngữ và tiếng nói. Tuy nhiên, khác với RNN, LSTM, GRU... Transformer không xử lý tuần tự các phần tử trong một chuỗi. Điều này có nghĩa Transformer có thể tận dụng khả năng tính toán song song GPU và giảm thời gian xử lý đáng kể.



Hình 8: Cấu trúc Transformer [10]

Đối với cấu trúc mô hình Transformer có thể tổng quát thành 2 thành phần chính: encoder và decoder. Các lớp này có thể xếp chồng lên nhau N lần phụ thuộc vào người xây dựng mô hình.

Lớp Input Embedding có nhiệm vụ nhận vector từ đã được trích xuất đặc trưng bằng các mô hình pretrained word embeddings đã được xây dựng sẵn. Bằng phương pháp trích xuất này ta có thể nhận được vector của từ. Nếu các vector càng giống nhau có nghĩa các từ đó sẽ gần nghĩa nhau hơn.

Tuy nhiên nếu chỉ phụ thuộc vào từ thì vẫn chưa thật sự hợp lý đối với việc xử lý cả câu. Điều này có thể thấy qua việc cũng chính từ đó nhưng nếu đặt ở các vị trí khác nhau sẽ tạo thành ý nghĩa câu khác nhau. Cho nên lớp Positional Encoding được thiết kế để giải quyết vấn đề đó trong xử lý ngôn ngữ tự nhiên.

$$PE_{(pos,2i)} = \sin(pos/10000^{2i/d_{model}})$$

$$PE_{(pos,2i+1)} = \cos(pos/10000^{2i/d_{model}})$$

Trong đó:

- pos là vị trí của từ trong câu.
- PE là giá trị phần tử thứ i trong embedding có độ dài d_{model} .

Lớp Multi-head attention giúp cho mô hình không chỉ sử dụng 1 cách để tìm những từ liên quan đến chính nó mà sẽ dùng nhiều attention chủ yếu tạo ra nhiều cách chú ý từ trong câu.

Lớp Feedforward chỉ là lớp có cấu trúc như một mạng nơ-ron cơ bản bao gồm 3 thành phần chính: Lớp đầu vào, lớp ẩn, lớp đầu ra.

2.6. Đánh giá mô hình

Có rất nhiều lựa chọn về các kỹ thuật đánh giá mô hình khác nhau, tuy nhiên vì để phù hợp với nhu cầu của bài toán, kỹ thuật đánh giá được lựa chọn ở đây là Confusion Matrix. Việc lựa chọn Confusion Matrix cho chúng ta thấy được cụ thể mỗi loại được phân loại như thế nào, lớp nào được phân loại đúng nhiều nhất, và dữ liệu thuộc lớp nào thường bị phân loại nhầm vào lớp khác.

Từ ma trận cơ bản này, ta sẽ có một số thuật ngữ sau:

		Model dự đoán	
		Positive	Negative
Thực tế	Positive	True Positive (Dự đoán Đúng là Positive)	False Negative (Dự đoán Sai là Negative)
	Negative	False Positive (Dự đoán Sai là Positive)	True Negative (Dự đoán Đúng là Negative)

Hình 9: Ma trận Confusion Matrix [12]

- Condition positive (P): Điều kiện chính xác
- Condition Negative (N): Điều kiện không chính xác
- True positive (TP): Số các dự đoán chính xác đúng.
- True negative (TN): Số các dự đoán không chính xác đúng.
- False positive (FP): Số các dự đoán chính xác sai.
- False negative (FN): Số các dự đoán không chính xác sai.

Với các thuật ngữ trên, ta có các chỉ số đánh giá & các khái niệm đi kèm sau: Chỉ số Precision Or Positive Predictive Value (PPV): Tỷ lệ dự đoán đúng là tỷ lệ giữa số sample được tính là True Positive (TP) với tổng số sample được phân loại là Positive (bằng chính TP + FP).

$$Precision = \frac{TruePositive}{TruePositive + FalsePositive}$$

Chỉ số Recall hoặc True Positive Rate (TPR): Độ nhạy - Tỷ lệ chính xác thực tế là tỷ lệ giữa các điểm positive thực được nhận đúng trên tổng điểm positive thực. Như vậy, Recall cao có nghĩa tỉ lệ bỏ sót các sample positive thực thấp.

$$Recall = \frac{TruePositive}{TruePositive + FalseNegative}$$

Như vậy chúng ta đã có 2 khái niệm Precision và Recall và mong muốn 2 giá trị này càng cao càng tốt. Tuy nhiên trong thực tế nếu ta điều chỉnh model để tăng Recall quá mức có thể dẫn đến Precision giảm và ngược lại, cố điều chỉnh mô hình để tăng Precision có thể làm giảm Recall. Nhiệm vụ của chúng ta là phải cân bằng 2 đại lượng này. Thật may có thêm một tham số nữa dung hòa giữa 2 cái và ta có thể căn vào đó để lựa chọn, đó là F-1 Score, là một trung bình hài hòa giữa Precision và Recall.

$$F1 = 2 \frac{Precision * Recall}{Precision + Recall}$$

3. PHÂN TÍCH DỮ LIỆU

3.1. Mô tả bài toán

Ý tưởng xây dựng mô hình bài toán này là thực hiện tìm các từ ngữ, cụm từ đặc trưng trong các loại bình luận. Nghĩa là ánh xạ mỗi từ ngữ, cụm từ vào một trường hợp bình luận nhất định từ đó trích xuất được các đặc trưng cho từng loại bình luận. Các tập từ ngữ, cụm từ trong các đặc trưng cụ thể đó, được sử dụng để làm tăng hiệu suất phân lớp văn bản.

3.2. Phân tích yêu cầu

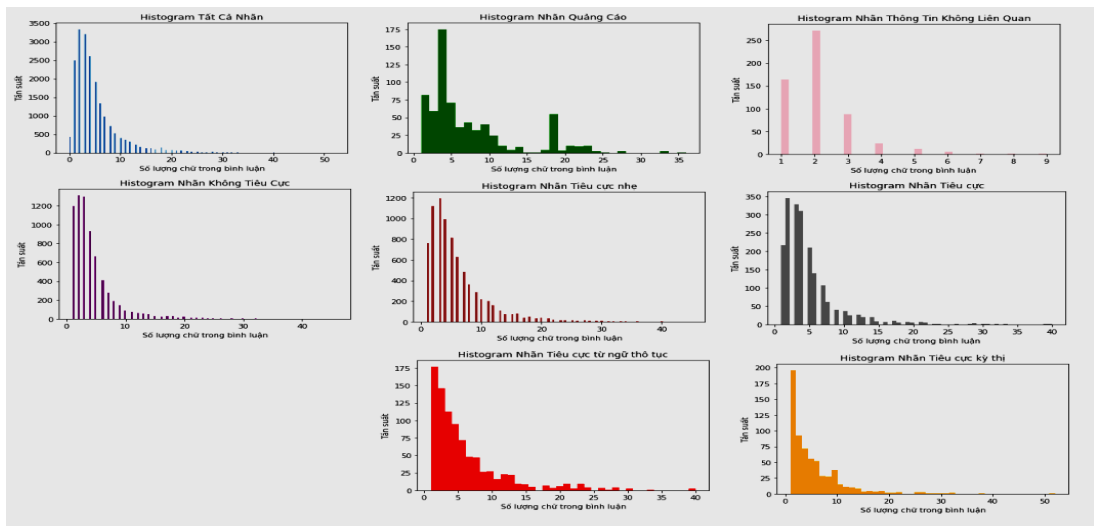
Chuẩn bị môi trường, sử dụng Jupyter Notebook, Google Colab, Kaggle, ... để tiến hành xử lý dữ liệu & train model. Dữ liệu và tiền xử lý dữ liệu, đầu tiên việc chuẩn bị dữ liệu cần được thu thập chính xác và đầy đủ theo mục đích yêu cầu của bài toán (phân loại bình luận tiêu cực) được thu thập từ các trang mạng xã hội (facebook) bằng các công cụ cào dữ liệu (selenium). Để có thể trả lời và đưa ra các đánh giá về mức độ tiêu cực của một bình luận trên mạng xã hội. Ở giai đoạn tiền xử lý dữ liệu, để giải quyết bài toán này đòi hỏi một lượng dữ liệu khá lớn cũng như độ chính xác cao của dữ liệu và tốn nhiều công sức. Tuy nhiên hiện nay cộng đồng xử lý ngôn ngữ tự nhiên tiếng Việt khá mạnh nên có rất nhiều mã nguồn chất lượng. Với độ chính xác khá cao, dễ dàng sử dụng, ở đây bài toán sử dụng công cụ và các bộ thư viện có sẵn để xử lý teencode, icon, emoji, stopwords (vietnamese stopwords) và word_tokenize của thư viện underthesea để tiến hành xây dựng bộ từ điển tiếng Việt với các từ trong câu được tách thành những từ đơn, cụm từ riêng biệt có ý nghĩa. Trích xuất các đặc trưng từ dữ liệu bằng cách tạo các bộ thư viện riêng bằng các từ ngữ phổ biến được tìm thấy & trích xuất từ bộ thư viện từ qua các nhãn khác nhau. Lựa chọn mô hình dự đoán.

3.3. Phân tích dữ liệu

Bảng 1: Phân phối giữa các nhãn dữ liệu

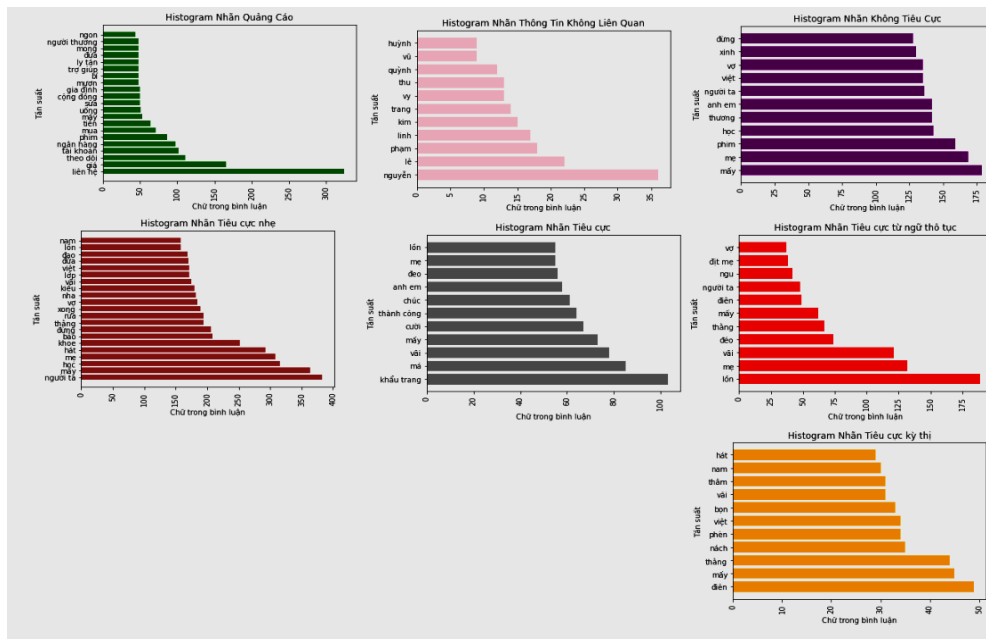
Type of comments	Percentile type of comments	Binary type of comments
Quảng cáo	4 %	Non - Toxic
Thông tin không liên quan	3 %	
Không tiêu cực	36 %	
Tiêu cực nhẹ	40 %	Toxic
Tiêu cực	10 %	
Tiêu cực kì thị	3 %	
Tiêu cực (chứa từ ngữ thô tục)	4 %	

Qua việc thống kê lại bộ dữ liệu, nhóm nhận thấy các nhãn chiếm đa phần là nhãn không tiêu cực, tiêu cực nhẹ. Còn các nhãn còn lại số lượng gần như là tương đồng nhau.



Hình 10: Phân phối độ dài câu bình luận trong tất cả nhãn và của từng nhãn

Từ hình 5 ta nhận thấy được hầu hết phân phối độ dài câu của từng nhãn dữ liệu là gần như nhau. Mặc dù ta thấy số lượng nhãn giữa các lớp là khác nhau



Hình 11: Phân phối các từ được lặp lại nhiều trong từng nhãn dữ liệu

Đối với nhãn quảng cáo, khá dễ phân loại. Có thể thấy các từ được lặp lại nhiều thường liên quan đến thông tin liên lạc, giá cả, giao dịch, sản phẩm.

Nhãn thông tin không liên quan sẽ tập trung những dạng bình luận không mang lại nhiều ý nghĩa về việc muốn đóng góp ý kiến. Cụ thể như các bình luận chỉ tag tên bạn bè, người thân mục đích để giới thiệu xem cùng chứ không có nhiều giá trị. Vì thế đa phần các từ lặp lại nhiều là các tên riêng của người được tag.

Nhãn không tiêu cực là nhãn chiếm tỉ lệ thứ 2 sau nhãn tiêu cực nhẹ. Có do bản chất loại nhãn bình luận còn khá tổng quát nên dễ rơi vào nhãn này. Đồng thời đây cũng là nhãn cong phụ thuộc vào suy nghĩ của người gán nhãn nếu người gán cảm thấy những lời bình luận không có ý nghĩa tiêu cực, chỉ là góp ý cá

nhân hay không có khả năng làm tổn thương người khác sẽ được phân vào nhãn này. Một điểm đặc biệt ở những từ được lặp lại ở nhãn này được xem là phân phối gần đồng nhất.

Nhãn tiêu cực nhẹ có chiếm tỉ lệ cao nhất dữ liệu. Lý do có thể bị ảnh hưởng vì nhóm chọn đa phần những bài báo về những vấn đề được tranh cãi nhiều ngoài xã hội. Nên việc thể hiện chính kiến của người bình luận sẽ có thêm chút cảm xúc để thể hiện sự khẳng định, bảo vệ ý kiến của mình. Riêng nhãn này sẽ dễ nhầm lẫn ở cả 2 nhãn như không tiêu cực và tiêu cực. Vì nó đòi hỏi người gán nhãn đọc thật kỹ câu bình luận liệu này có thật là chỉ thể hiện ý riêng nhưng chỉ theo hướng tiêu cực hay có thêm thông tin để công kích một số nhóm trái ý mình hoặc câu từ bình luận có thể hiện ý kiến một cách mạnh mẽ nhưng lại theo hướng tích cực.

Thông thường, nhãn tiêu cực sẽ được phân loại đối với các bình luận đưa thêm thông tin khác ngoài lẽ về đối tượng để góp phần phản biện cho một ý kiến trái chiều khác. Ngoài ra, còn có dạng bình luận mang châm chọc, đá xéo đối tượng.

Nhãn tiêu cực kì thị, chủ yếu các từ ngữ xuất hiện trong câu ở đây mang ý nghĩa xúc phạm, chế giễu cá nhân, tổ chức, xã hội..., giới tính, ngoại hình, chủng tộc, quan điểm chính trị, thân trạng ... Thường thể hiện một quan điểm tiêu cực về vấn đề nhạy cảm.

Nhãn tiêu cực (có từ ngữ thô tục), chủ yếu các từ ngữ xuất hiện trong câu là các từ có ý nghĩa chửi, mắng, từ lóng, các từ ngữ thô tục. Có ý định ác ý, gây sự nhiều bên để có thêm sự tranh cãi, lăng mạ đối tượng nhắm tới.

4. THỰC NGHIỆM

Đối với bước đưa dữ liệu vào mô hình, nhóm phải kiểm tra thật kỹ liệu mô hình tránh hiện tượng overfitting trên mô hình. Bởi vì hiện tượng này thường xuất hiện khi ta đưa quá nhiều từ được trích xuất từ bộ dữ liệu dẫn đến overfitting ở mô hình. Biết được điều này, nhóm đã điều chỉnh đầu vào bằng cách thống kê các từ thường xuyên lặp lại trên 100 lần để tránh hiện tượng đó. Đồng thời để tăng độ chính xác mô hình mà không bị overfitting, nhóm trích xuất thêm những từ quan trọng chỉ riêng từng nhãn có kết hợp với tập chữ thường xuyên xuất hiện. Tổng cộng số chữ được trích xuất đưa vào mô hình là 200 chữ. Sau đây là bản số liệu mà nhóm đã thử chạy qua các mô và được độ chính xác tốt nhất của từng mô hình.

Bảng 2: Thống kê thông số đánh giá qua từng mô hình đạt được. Với các nhãn được ký hiệu: (1) quăng cáo, (2) thông tin không liên quan, (3) không tiêu cực, (4) tiêu cực nhẹ, (5) thông tin tiêu cực, (6) tiêu cực kỳ thị, (7) tiêu cực từ ngữ thô tục

label	precision	recall	f1-score	accuracy	model
1	0.79	0.68	0.73	0.54	Bi(LSTM)- 2xBi(GRU)
2	0.36	0.14	0.2		
3	0.5	0.58	0.54		
4	0.57	0.64	0.6		
5	0.48	0.23	0.31		
6	0.44	0.23	0.31		
7	0.48	0.36	0.41		
1	0	0	0	0.41	3xTransformer
2	0	0	0		
3	0.37	0.86	0.52		
4	0.53	0.31	0.39		
5	0	0	0		
6	0	0	0		
7	0	0	0		
1	0.88	0.63	0.74	0.51	CNN- Bi(LSTM)- Bi(GRU)
2	0	0	0		
3	0.49	0.5	0.5		
4	0.52	0.71	0.61		
5	0.4	0.14	0.21		

6	0	0	0		
7	0.39	0.33	0.36		
1	0	0	0	0.41	3xGRU
2	0	0	0		
3	0	0	0		
4	0.41	1	0.58		
5	0	0	0		
6	0	0	0		
7	0	0	0		
1	0.85	0.66	0.74	0.54	3xLSTM
2	0.37	0.2	0.26		
3	0.52	0.51	0.51		
4	0.55	0.74	0.63		
5	0.62	0.16	0.26		
6	0.49	0.19	0.28		
7	0.49	0.37	0.42		
1	0.59	0.55	0.57	0.56	SVM
2	0.4	0.51	0.45		
3	0.67	0.53	0.59		
4	0.66	0.58	0.62		
5	0.15	0.69	0.24		
6	0.07	0.57	0.13		
7	0.07	0.3	0.11		

Các mô hình được đề cập ở bảng số liệu hình 7 đều được nhóm chạy cùng một tập dữ liệu chữ trích xuất, số lượng epoch = 10, batch size = 64. Ngoại trừ, SVM là mô hình thuộc lĩnh vực học máy nên không dùng đến các tham số epoch, batch size. Nhìn sơ lược bảng hình 7, có thể thấy được ba mô hình có hiệu suất nổi bật: Bi(LSTM)-2xBi(GRU), 3xLSTM, SVM. Nhưng vào chi tiết, các thông số precision, recall, f1-score về nhãn tiêu cực từ ngữ thô tục lại không cao. Chúng tôi mô hình SVM sẽ không phân loại tốt đối với nhãn tiêu cực từ ngữ thô tục. Nhóm đề xuất có thể sử dụng 1 trong 2 mô hình Bi(LSTM)-2xBi(GRU), 3xLSTM còn lại.

5. KẾT LUẬN

5.1. Kết luận

Tóm lại, từ những dữ liệu thô sơ trên mạng xã hội nhóm vẫn có thể xử lý và giải quyết vấn nạn bắt nạt trên mạng xã hội. Mặc dù kết quả thành công không quá cao nhưng nó có thể là nền tảng để phát triển tốt hơn để góp một phần không nhỏ giải quyết vấn nạn bắt nạt trong thời đại mạng Internet và nhu cầu sử dụng mạng xã hội ngày càng gia tăng không chỉ ở Việt Nam mà còn là các nước trên thế giới.

5.2. Hạn chế

Mặc dù nhóm đã cố gắng xử lý dữ liệu rất nhiều nhưng phần hạn chế lớn nhất nằm ở bộ dữ liệu. Nếu có nguồn nhân lực nhiều hơn ngoài thành viên trong nhóm thì bộ dữ liệu sẽ được gán nhãn với nhiều nhãn định khác nhau giúp cho dữ liệu ngày càng khách quan hơn. Đồng thời nếu có thể, đầu tư thêm thời gian lấy thêm các dữ liệu chiếm ít phần trong bộ dữ liệu.

5.3. Hướng phát triển

Nhóm có thể thử thêm nhiều kỹ thuật xây dựng mô hình khác để có thể cải thiện thêm độ chính xác của mô hình, cụ thể là trong lĩnh vực học sâu. Khi đa phần mô hình phân loại được đều các nhãn với nhau rất có

tiềm năng. Cuối cùng, nhóm cũng hướng tới có thể ứng dụng được mô hình vào một hệ thống lọc trên mạng xã hội một cách tự động hóa.

ACKNOWLEDGMENTS

Chúng em xin bày tỏ lòng biết ơn sâu sắc tới cô Đặng Thị Phúc. Chân thành cảm ơn Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh đã tạo điều kiện cho chúng em học tập và thực hiện đề tài nghiên cứu.

TÀI LIỆU THAM KHẢO

- [1] Deep Learning For NLP: Zero To Transformers & BERT: <https://www.kaggle.com/code/tanulsingh077/deep-learning-for-nlp-zero-to-transformers-bert/notebook>
- [2] fastText Vietnamese word vectors full: <https://www.kaggle.com/datasets/aeryss/fasttext-vietnamese-word-vectors-full>
- [3] vietnamese_stopwords: <https://www.kaggle.com/datasets/linhlpv/vietnamese-stopwords>
- [4] Craw data facebook: https://github.com/thangnch/MiAI_Facebook_Crawl
- [5] Toxic Comment Classification: <https://scholar.smu.edu/datasciencereview/vol3/iss1/13/>
- [6] Teen code dataset: https://xltiengviet.fandom.com/wiki/Danh_s%C3%A1ch_teencode_ph%E1%BB%95_bi%E1%BA%BFn_trong_ti%E1%BA%BFng_Vi%E1%BB%87t
- [7] Xóa biểu cảm: <https://helpex.vn/question/xoa-bieu-tuong-cam-xuc-khoi-chuoi-bang-python-60c31e589f28385abef63e14>
- [8] Support vector machine: Support Vector Machines for Text Categorization 2003. A. Basu, C. Watters, and M. Shepherd
- [9] Convolutional Neural Network: <https://arxiv.org/pdf/1408.5882.pdf>
- [10] Transformer: <https://arxiv.org/pdf/1706.03762.pdf>
- [11] Long short term memory for multi class: <https://towardsdatascience.com/multi-class-text-classification-with-lstm-1590bee1bd17>
- [12] Confusion matrix: <https://machinelearningcoban.com/2017/08/31/evaluation/>
- [13] Long short term memory: https://d2l.ai/vn.com/chapter_recurrent-modern/lstm_vn.html
- [14] Gated recurrent unit: <https://towardsdatascience.com/illustrated-guide-to-lstms-and-gru-s-a-step-by-step-explanation-44e9eb85bf21>

YSC4F.203

THIẾT KẾ HỆ THỐNG ĐIỀU KHIỂN THIẾT BỊ SỬ DỤNG RASPBERRY PI 4 TRÊN NỀN TẢNG IOT HOME ASSISTANT

NGUYỄN ĐÌNH KHÔI, PHẠM HÙNG PHƯƠNG, NGUYỄN TẤN LỘC

Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

18038901.khoi@student.iuh.edu.vn, 18051631.phuong@student.iuh.edu.vn, nguyentanloc@iuh.edu.vn

Tóm tắt. Bài báo trình bày thiết kế hệ thống điều khiển từ xa thiết bị trong nhà áp dụng cho homestay. Hệ thống điều khiển được thiết kế sử dụng bộ điều khiển trung tâm Raspberry Pi dựa trên nền tảng IoT Home Assistant (HASS). Bộ điều khiển trung tâm kết nối, thực hiện lệnh điều khiển, thu thập và gửi dữ liệu của các thiết bị lên Home Assistant để lưu trữ và xử lý. Hệ thống có thể được điều khiển từ xa sử dụng Wi-Fi. Kết quả thử nghiệm đã thành công kết nối và sử dụng hệ thống. Hệ thống thiết kế quy mô nhỏ và hoạt động đơn giản hơn như điều khiển, xem thông tin nhiệt độ, độ ẩm, lịch sử hoạt động thích hợp cho các ứng dụng nhỏ lẻ. Mô hình thiết kế được đề xuất và thực hiện với chi phí chỉ khoảng 170\$. Hệ thống có khả năng kết nối được nhiều thiết bị của các hãng khác nhau nhờ vào hệ sinh thái phong phú của HASS. Cuối cùng, mô hình hệ thống được thiết kế và xây dựng đưa vào sử dụng thực tế để đánh giá hiệu quả của hệ thống.

Từ khóa. IoT, điều khiển từ xa, Home Assistant, Raspberry Pi.

DESIGN OF A REMOTE CONTROL SYSTEM USING RASPBERRY PI 4 BASED ON IOT HOME ASSISTANT PLATFORM

Abstract. This paper presents the design of a remote control system for home appliances applied to homestays. The control system is designed using Raspberry Pi as a central control unit (CCU) based on IoT Home Assistant (HASS) platform. The CCU connects, executes control commands, collects and sends appliance data to Home Assistant for storage and processing. The system can be controlled remotely using Wi-Fi. Test results have successfully connected and used the system. The system is small-scale design and simpler to operate, such as control, view information about temperature, humidity, and operation history, suitable for small applications. It is suitable for small applications. The proposed design model has a cost of only about \$170. The system is capable of connecting many different device brands by using the HASS ecosystem. Finally, the control system model is built and tested to evaluate the effectiveness of the system.

Keywords. IoT, Remote control, Home Assistant, Raspberry Pi.

1. GIỚI THIỆU

Mô hình dịch vụ homestay hiện nay phát triển nhanh chóng và ứng dụng vào du lịch để du khách được trải nghiệm du lịch một cách hoàn hảo, vì ngoài yếu tố phong cảnh, ẩm thực thì nơi ở đóng vai trò hết sức quan trọng cho người du khách. Homestay là loại hình trải nghiệm du lịch hiện đã không còn quá xa lạ với chúng ta, tại đây du khách được nghỉ dưỡng trong những ngôi nhà thực sự ấm áp đầy đủ tiện nghi qua đó dễ dàng tìm hiểu, tận hưởng và khám phá những nét văn hóa, con người, đặc điểm vùng miền một cách thoải mái, dễ dàng [1]. Yếu tố thoải mái, tiện nghi, an toàn là điều cơ bản trong một ngôi nhà homestay và nhu cầu điều khiển công nghệ là tiêu chí mà chủ quản lý homestay đang hướng đến nhằm mang lại trải nghiệm tốt nhất cho du khách. Dịch Covid-19 khiến cho du lịch homestay gặp nhiều khó khăn, việc di chuyển đi lại bị hạn chế nên làm việc tại nhà là lựa chọn an toàn của nhiều người, do vậy nhiều công việc bị ảnh hưởng khi chủ nhà ở những địa phương xa không thể trực tiếp đến được homestay dẫn đến nhu cầu áp dụng IoT [2] để quản lý từ xa.

Nội dung bài báo mô tả hệ thống phần cứng gồm sơ đồ khối hệ thống và sơ đồ kết nối thiết bị, cấu hình phần mềm nền tảng IoT [3] Home Assistant [4][5] để lưu trữ cấu hình, quản lý, điều khiển thiết bị từ xa qua wifi [6], xây dựng giao diện điều khiển gọn gàng, dễ sử dụng và hiển thị đầy đủ thông tin từ cảm biến, tạo chức năng tự động thông báo đơn giản khi có chuyển động. Phương pháp điều khiển gián tiếp qua trung gian được sử dụng với trung tâm là Raspberry Pi 4 [7] và NodeMCU8266 làm nhiệm vụ điều khiển vì ưu điểm điều khiển được nhiều loại thiết bị, dễ quản lý và mở rộng. Sản phẩm thực tế đáp ứng được các mục tiêu đã đề ra, hệ thống hoạt động ổn định trên mô hình homestay.

2. MÔ TẢ HỆ THỐNG

2.1. Thiết kế hệ thống

Kiến trúc thiết kế bao gồm: khối nguồn, khối trung tâm, khối điều khiển, giao diện người dùng và thiết bị đầu cuối.

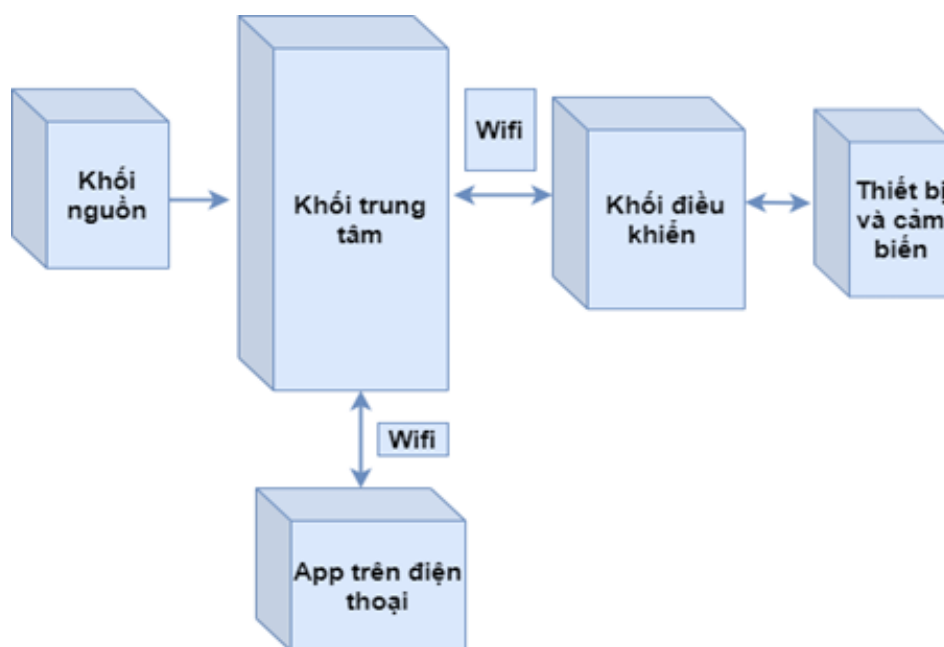
Khối nguồn: cung cấp nguồn cho hệ thống, sử dụng nguồn 5VDC, 12VDC và 220VAC.

Khối trung tâm: là nơi cài đặt HassOS, kết nối với khối điều khiển, truyền nhận, lưu trữ dữ liệu, cấu hình và quản lý các thiết bị trong hệ thống, gửi lệnh điều khiển xuống khối lệnh điều khiển để thực thi

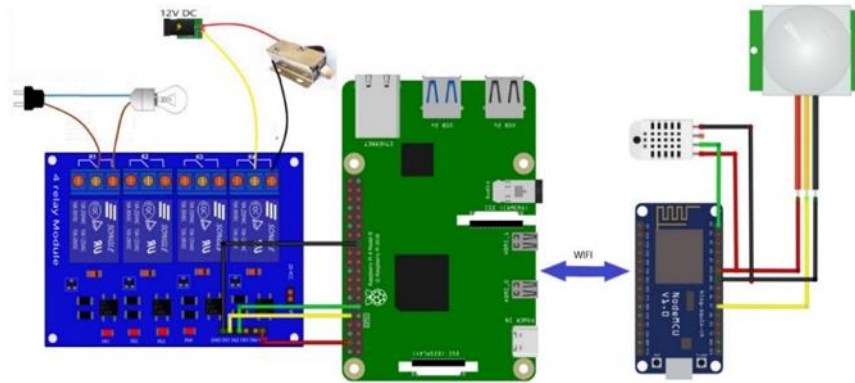
Khối điều khiển: nhận lệnh điều khiển từ khối xử lý trung tâm và thực hiện điều khiển thiết bị, chuyển dữ liệu từ cảm biến về trung tâm.

Ứng dụng điều khiển trên điện thoại: kết nối với Home Assistant và thực hiện điều khiển từ xa.

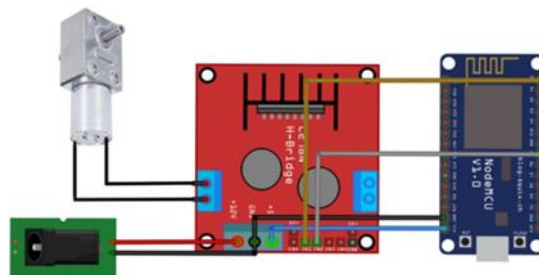
Nhằm hỗ trợ người dùng trong việc quản lý và điều khiển các thiết bị trong ngôi nhà của mình hệ thống cần: tạo được hệ thống phần cứng và kết nối với mạng không dây (wifi), xây dựng giao diện điều khiển cho hệ thống qua nền tảng IoT Home Assistant, thành công điều khiển thiết bị từ xa. Để đáp ứng với mục tiêu đề ra hệ thống sử dụng phần cứng gồm có:



Hình 1: Sơ đồ khối hệ thống

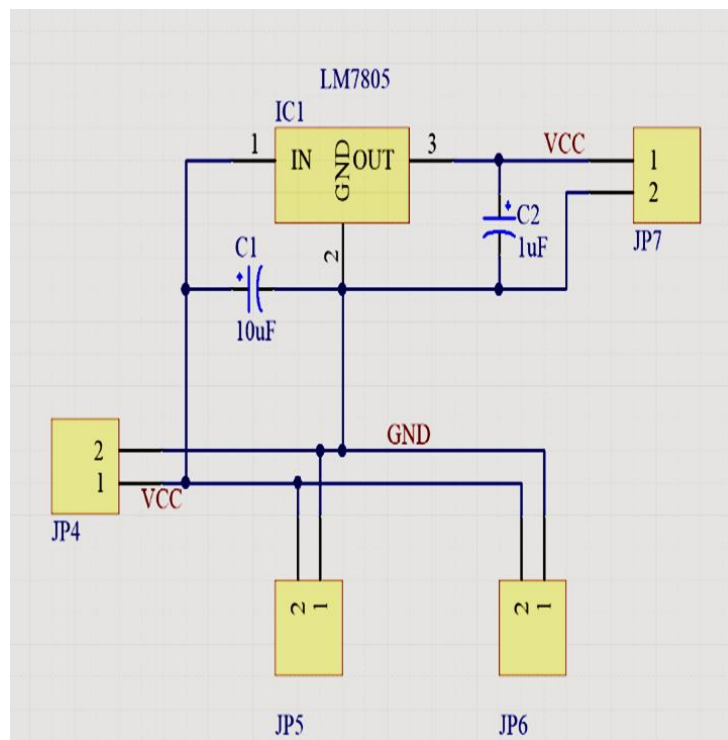


Hình 2: Sơ đồ nối chân điều khiển khóa cửa, đèn, cảm biến



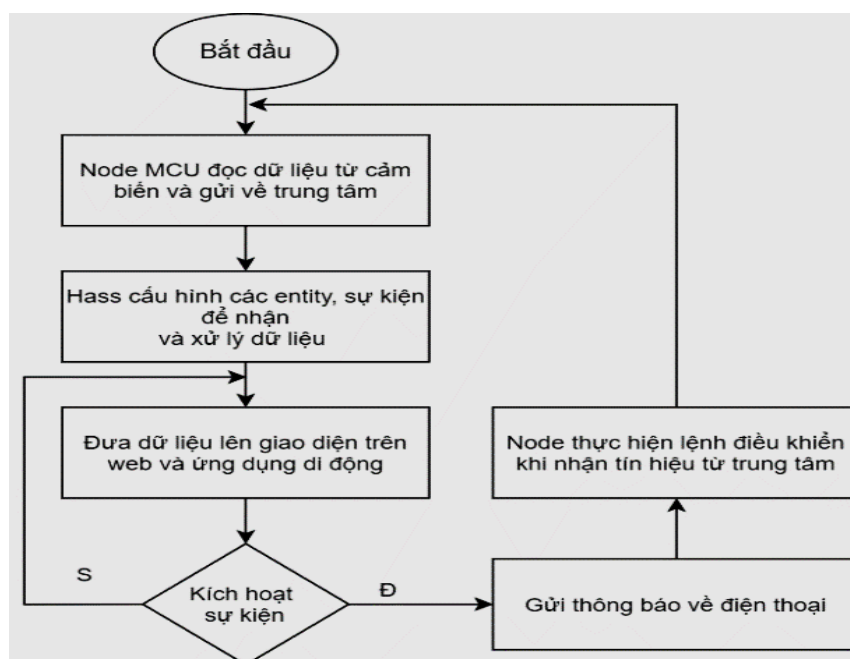
Hình 3: Sơ đồ nối chân điều khiển rèm cửa

Do nhu cầu sử dụng cùng lúc 12V DC và 5V DC nên khối nguồn (Hình 21) sử dụng IC LM7805 có đầu vào tối đa đến 18V, đầu ra ổn định 5V, dòng 1.5A. Có 2 ngõ ra trực tiếp lấy từ đầu vào 12V. Tụ điện 10uF và 1uF để lọc nhiễu ổn định đầu vào và đầu ra.



Hình 4: Sơ đồ nguyên lý mạch nguồn

2.1. Lưu đồ giải thuật



Hình 5: Lưu đồ giải thuật hệ thống

Cài đặt HassOS lên Raspberry và kết nối wifi lần đầu cần khoảng 20 phút. Sau đó đăng ký sử dụng và kết nối dịch vụ cloud để điều khiển ngoại mạng không cần dùng chung wifi, tiếp đến là kết nối NodeMCU đóng vai trò làm gateway vào Hass để truyền nhận dữ liệu từ cảm biến. Thiết lập các entity tạo thực thể để chứa thông tin và dữ liệu cho từng thiết bị sau đó đưa lên giao diện trên web và điện thoại. Nếu sự kiện được kích hoạt sẽ gọi lớp dịch vụ gửi thông báo về điện thoại. Khi nhận được lệnh điều khiển gửi từ trung tâm NodeMCU sẽ thực hiện và cập nhật lại trạng thái thiết bị về trung tâm.

2.2. Cấu hình phần mềm

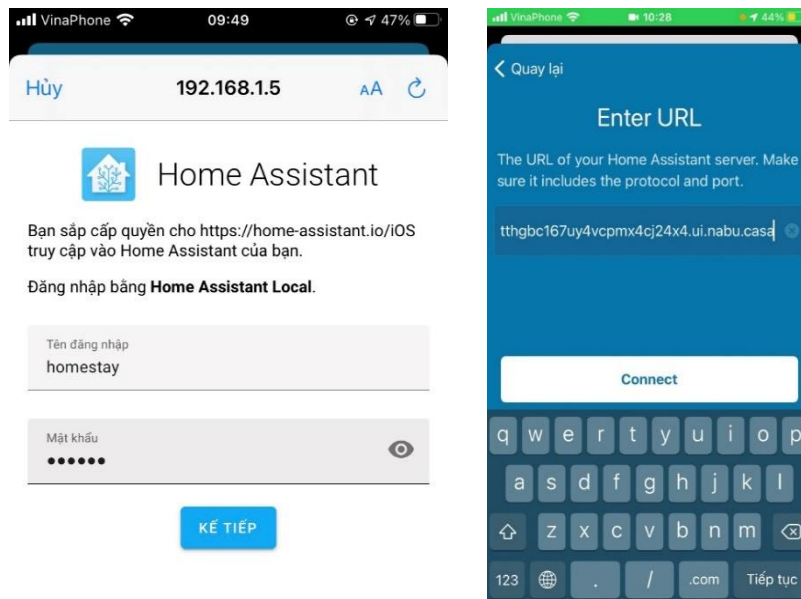
Cài đặt các add-on để thực hiện chỉnh sửa file YAML cấu hình trên Hass và NodeMCU, là chuẩn chính để lưu trữ cấu hình và các tích hợp-integration/component cũng như các platform cú pháp theo dạng key-value [8]. File "configuration.yaml" dùng để chứa cấu hình kết nối thiết bị vào hệ thống bằng cách sử dụng integration và platform [9], dùng integration *switch* cho đèn, khóa cửa và platform "rpi_gpio" để khai báo chân gpio, đặt tên và id mà thiết bị đang dùng trên Raspberry Pi. Dùng *trigger* để tạo thông báo về điện thoại khi có sự thay đổi trạng thái của cảm biến chuyển động từ "off" sang "on". Cấu hình cho NodeMCU cũng theo các bước tương tự và khai báo thêm mạng wifi mà hệ thống đang dùng, dùng integration *binary_sensor* cho cảm biến chuyển động và *sensor* cho cảm biến nhiệt độ, độ ẩm, kết hợp giữa *switch* và *cover* để mở/đóng rèm kéo với thời gian tùy chỉnh. Giao diện thiết kế đơn giản, gọn gàng, gồm các chức năng: điều khiển đơn giản dạng công tắc gạt, xem lịch sử hoạt động từng thiết bị, xem thông tin nhiệt độ, độ ẩm, chuyển động. Hệ thống dùng được trên cả web và ứng dụng điện thoại nên người dùng có thể truy cập ở bất cứ đâu có internet.

2.3. Phương pháp điều khiển

Hệ thống được thiết kế và điều khiển sử dụng phương pháp điều khiển qua trung gian [10], sử dụng Raspberry Pi 4 làm trung tâm, các thiết bị sẽ được kết nối và quản lý đồng bộ nhờ đó chỉ cần gửi lệnh điều khiển từ ứng dụng được cài đặt trên điện thoại đến bộ xử lý trung tâm, trung tâm sẽ phân luồng và chuyển tiếp đến địa chỉ của thiết bị để thực hiện lệnh, tương tự dữ liệu của cảm biến thu thập được cũng đi qua trung tâm để hiển thị lên giao diện người dùng. Ưu điểm lớn của phương pháp này là dễ dàng quản lý, kiểm soát và mở rộng hệ thống và điều khiển được thiết bị không có tích hợp chức năng wifi để kết nối trực tiếp.

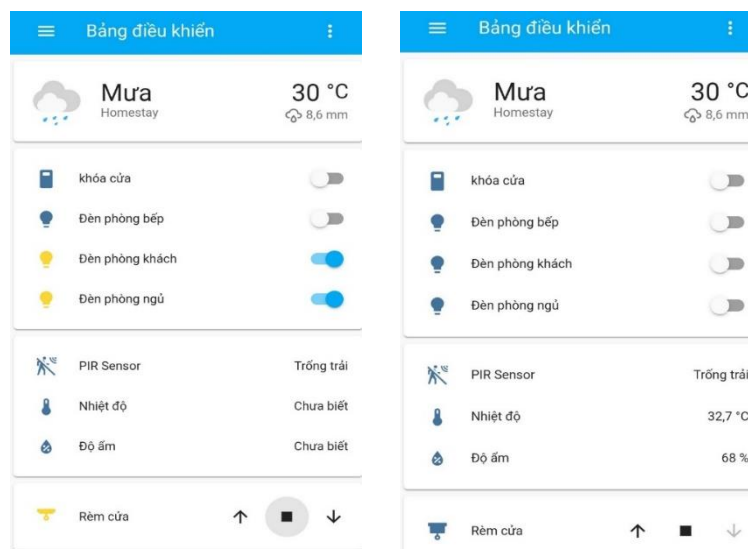
3. KẾT QUẢ VÀ KẾT LUẬN

3.1. Kết quả



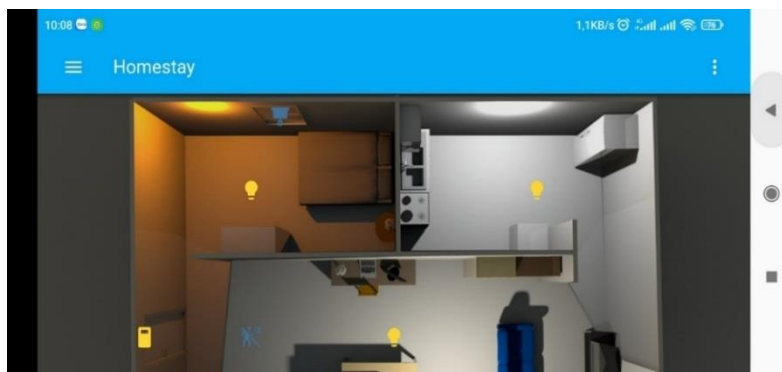
Hình 6: Màn hình đăng nhập và dán link url

Mở ứng dụng sau khi khởi động Hass nếu điện thoại đang dùng chung một mạng wifi với Hass thì sẽ tự động truy cập đến trang đăng nhập hệ thống. Khi dùng một mạng khác để sử dụng sẽ không bắt IP của Pi để vào hệ thống, home assistant cloud giải quyết vấn đề bằng cách kết nối đến Hass và cấp một địa chỉ url. Bằng cách đó chỉ cần dùng đường link được cấp để đăng nhập là có thể truy cập hệ thống.

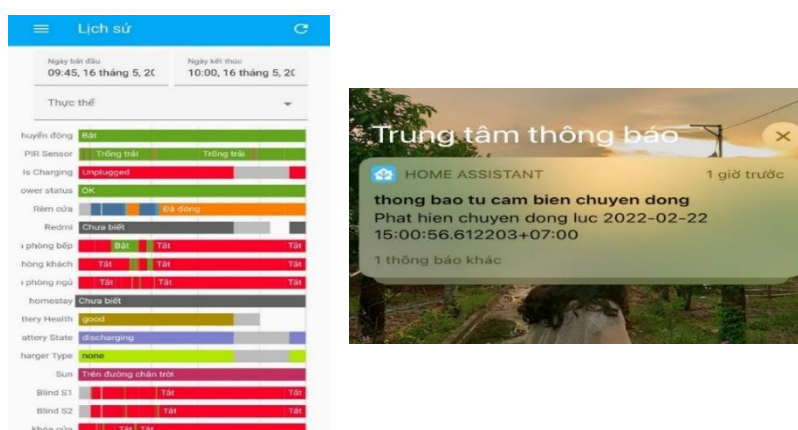


Hình 7: Màn hình điều khiển trên giao diện

Đăng nhập thành công sẽ tiến vào giao diện điều khiển theo dạng công tắc đóng/mở hoặc lựa chọn điều khiển qua mô hình 3D. Trên giao diện hiển thị đầy đủ các thiết bị kết nối đến hệ thống, giao diện chia thành 3 phần rõ ràng gồm khối đèn, rèm cửa, khối hiển thị cảm biến và khối kéo/mở rèm. Thiết bị chưa kết nối vào hệ thống sẽ hiện trạng thái “chưa biết”, sau khi kết nối màn hình sẽ tự động cập nhật lại trạng thái và hiển thị dữ liệu, tương tự với đèn và rèm cửa khi chưa kết nối sẽ ẩn đi các nút nhấn bật/tắt.



Hình 8: Điều khiển trên mô hình 3D



Hình 9: Lịch sử hoạt động và thông báo gửi về điện thoại

3.2. Phân tích, đánh giá kết quả

Phạm vi kết nối lớn nên ở bất cứ đâu có internet đều có thể truy cập đến hệ thống, sử dụng nhiều mạng wifi khác nhau trong phạm vi 1 km đều đăng nhập và sử dụng thành công. Quá trình truyền nhận dữ liệu từ cảm biến có sai số nhỏ, tín hiệu điều khiển truyền xuống các node thực thi ổn định, không bị mất dữ liệu. Dữ liệu bị lỗi hoặc không chính xác trong quá trình truyền nhận do nguồn cấp không đủ, sự ổn định của mạng internet đang sử dụng. Tốc độ gửi thông báo về điện thoại có thời gian trung bình nhỏ hơn 1 giây gần như ngay lập tức sau 10 lần thử trên cảm biến chuyển động. Lệnh điều khiển được truyền trực tiếp về Raspberry Pi không phải qua trung gian bên ngoài và dễ dàng kết nối thiết bị mới, các ngữ cảnh tự động hóa là những ưu điểm của hệ thống. Tuy nhiên quá trình khởi động và kết nối NodeMCU vào Hass còn mất nhiều thời gian chiếm từ 10 đến 15 phút.

3.3. Kết luận

Hệ thống với quy mô nhỏ áp dụng vào ngôi nhà với diện tích vừa và nhỏ. Tính năng cơ bản như điều khiển thiết bị, xem thông số môi trường đáp ứng các nhu cầu cơ bản. Sử dụng các thiết bị đơn giản và nền tảng IoT mã nguồn mở nên có chi phí thấp. Tuy nhiên, chức năng và độ ổn định chưa so được với các hệ thống nhà thông minh hiện đang có trên thị trường.

TÀI LIỆU THAM KHẢO

- [1] N. V. B. Nguyễn Văn Sĩ, "Các yếu tố tác động đến sự hài lòng và lòng trung thành của khách du lịch Homesstay: Nghiên cứu trường hợp Bến Tre," Tạp chí phát triển khoa học & công nghệ, vol. 2, no. 3, pp. 5-17, 2018.

- [2] BT Sinh, “Năng lực và định hướng của Internet vạn vật (IoT) trên thế giới”, *J. Sci. Kỹ thuật. Chính sách Quản lý.*, quyển sách. 6, không. 3, trang 84–92, tháng 10 năm 2017.
- [3] B. Vogel, Y. Dong, B. Emruli, P. Davidsson, and R. Spalazzese, “What Is an Open IoT Platform? Insights from a Systematic Mapping Study,” *Future Internet*, vol. 12, no. 4, p. 73, Apr. 2020, doi: 10.3390/fi12040073
- [4] A. Kalinauskas and S. Dennis, ‘Deploying and testing Home Assistant at HINT’, Dissertation, 2021.
- [5] M. Suvar, G. Daniei Florea, L. Munteanu, "Residential security through the Home Assistant Platform," *MATEC Web Conf.*, Vols. 354, 2022, no. 10th International Symposium on Occupational Health and Safety (SESAM 2021), p. 10, 2021.
- [6] A. Ismail Al-Alawi, "WiFi Technology: Future Market Challenges and Opportunities," *Journal of Computer Science*, vol. 2, no. 1, pp. 13-18, 2006.
- [7] R. A. A. a. O. D. a. A. Marzak, "Internet of Things: a new Interoperable IoT Platform. Application to a Smart Building," *Procedia Computer Science*, vol. 191, no. 1877-0509, pp. 511-517, 2021.
- [8] GEditor, "https://konnected.vn/," 15 6 2020. [Online]. Available: <https://konnected.vn/home-assistant/tin-tuc-va-cap-nhat/home-assistant-tuong-lai-cua-yaml-2020-04-24>.
- [9] GEditor, "https://konnected.vn/," 26 6 2020. [Online]. Available: <https://konnected.vn/home-assistant/home-assistant-thuat-ngu-thong-dung-2020-03-19>
- [10] T. T. Ha, L. N. Quy and V. T. Thanh, "Xây dựng hệ thống điều khiển thiết bị điện ứng dụng công nghệ Internet of Things," in *Tạp chí Khoa học và Công nghệ Đại học Đà Nẵng*, *Tạp chí Khoa học và Công nghệ Đại học Đà Nẵng*, 2017, pp. 26-30.

YSC4F.204

ỨNG DỤNG MÔ HÌNH HỌC MÁY TRONG PHÂN TÍCH CẢM XÚC KHÁCH HÀNG

HỒ HOÀNG HUY QUANG, PHẠM THỊ THIẾT, VÕ QUANG HOÀNG KHANG

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

18067451.quang@student.iuh.edu.vn, phamthithiet@iuh.edu.vn, vqhkhang@iuh.edu.vn

Tóm tắt. Ngày nay với sự phát triển mạnh mẽ của công nghệ thông tin và internet, các website Thương mại điện tử ra đời như một phương tiện hữu ích và tiện lợi giúp khách hàng thực hiện mua hàng trực tuyến cũng như chia sẻ những nhận xét và đánh giá về chất lượng của sản phẩm sau mỗi lần giao dịch. Dựa trên những nhận xét, đánh giá này, các khách hàng tiềm năng khác sẽ an tâm hơn với sản phẩm mình đang quan tâm cũng như các nhà cung cấp sản phẩm sẽ hiểu thêm nhu cầu của khách hàng để từ đó có kế hoạch cải tiến hoặc phát triển sản phẩm tốt hơn. Để thấu hiểu hơn về sản phẩm thông qua các ý kiến bình luận của khách hàng, trong bài báo này, nhóm tác giả nghiên cứu và ứng dụng mô hình PhoBERT và mBERT kết hợp với LSTM để phân tích cảm xúc của khách hàng trong việc đánh giá, nhận xét về quá trình đặt hàng, giao hàng, dịch vụ bán hàng và chất lượng của sản phẩm. Thực nghiệm được đánh giá trên dữ liệu nhận xét, đánh giá về các linh kiện máy tính, điện thoại, đồng hồ thông minh được thu thập trên website thương mại điện tử như Tiki, Shopee, FPT Shop...

Từ khóa. học máy, phân tích cảm xúc, ý kiến khách hàng.

APPLY MACHINE LEARNING MODELS IN SENTIMENT ANALYSIS OF CUSTOMER

Abstract. Nowadays, with the strong evolution of Information technology, many e-commerce websites have established as a useful and convenient way for customers make online purchase as well as share their reviews and evaluate about the quality of products after each transaction. Based on these reviews, potential customers will be more secure with the products they are interested in, by the way, product suppliers will understand more about what customers' needs so that they can make plans to improve or develop them better. For more understanding about product based on reviews from customers, in this paper, we have researched and performed PhoBERT and mBERT model combined with LSTM to analysis customer reviews about: Oder process, product delivery, selling services and product quality. Experimental results on reviews data about computer components, mobile phones, smart watches collected on e-commerce websites such as: Tiki, Shopee, FPT Shop,...

Keywords. machine learning, sentiment analysis, customer reviewers.

1. GIỚI THIỆU

Trong những năm gần đây, với sự phát triển mạnh mẽ của công nghệ thông tin, rất nhiều các website thương mại điện tử ra đời như Tiki, Shopee, FPT Shop, Lazada, Sendo... giúp cho việc mua sắm hàng hóa của khách hàng ngày càng thuận tiện. Bên cạnh tính năng cho phép người dùng lựa chọn sản phẩm, trên mỗi website đều có chức năng cho phép người dùng chia sẻ các ý kiến nhận xét đánh giá, bình luận về chất lượng của sản phẩm sau mỗi lần mua sắm thành công dẫn đến một khối lượng khổng lồ về dữ liệu được sản sinh. Đây là những thông tin dữ liệu có giá trị cho cả khách hàng và doanh nghiệp. Dựa trên những thông tin này, các khách hàng tiềm năng khác sẽ an tâm hơn với sản phẩm mình đang quan tâm cũng như các nhà cung cấp sản phẩm sẽ hiểu thêm nhu cầu của khách hàng để từ đó có kế hoạch cải tiến hoặc phát triển sản

phẩm tốt hơn. Tuy nhiên, rất khó có thể xử lý một lượng lớn thông tin để tìm ra các thông tin có giá trị. Do đó, chúng ta cần xây dựng một hệ thống phân tích đánh giá để phân loại dữ liệu, xem bình luận nào là tích cực, bình luận nào là tiêu cực và dùng các phương pháp phân tích và dự đoán để phân loại cảm nhận của khách hàng. Dựa vào nhu cầu thiết thực trên, các mô hình máy học ra đời với mục đích có thể phân tích tự động các phản hồi từ khách hàng. Các mô hình RNN với biến thể LSMTs hay sự kết hợp giữa CNN với LSTM cũng đã đem lại nhiều kết quả cao dựa trên phân tích từng khía cạnh. Tuy nhiên, việc phân tích từng khía cạnh đôi khi đòi hỏi mô hình máy học tốn rất nhiều thời gian. Để tận dụng được ưu điểm của mô hình máy học trước đó cùng với bộ dữ liệu khổng lồ có sẵn, một mô hình mới ra đời dựa trên việc xem xét ngữ nghĩa theo 2 chiều [9]. Hiện tại, mô hình BERT [3] được xem là 1 trong các mô hình tối ưu nhất cho bài toán xử lý ngôn ngữ tự nhiên. Mô hình BERT ứng dụng kỹ thuật Transformers để biểu diễn từ theo 2 chiều. Hiện nay, mô hình BERT đã được ứng dụng trong phân tích ngôn ngữ tiếng Việt như PhoBERT [7]. Kết quả thực nghiệm trong [7] cho thấy PhoBERT hoạt động tốt và cải thiện nhiều tính năng trong xử lý ngôn ngữ tự nhiên dành cho tiếng Việt. Vì vậy trong bài báo này chúng tôi đề xuất ứng dụng mô hình máy học và xử lý ngôn ngữ tự nhiên để phân tích cảm xúc của khách hàng. Cụ thể, chúng tôi đề xuất ứng dụng và đánh giá hiệu suất giữa mô hình PhoBERT kết hợp với LSTM và mBERT (BERT-Base Multilingual) kết hợp với LSTM để phân tích cảm xúc của khách hàng trong việc đánh giá, nhận xét về quá trình đặt hàng, giao hàng, dịch vụ bán hàng và chất lượng của sản phẩm.

Nội dung tiếp theo của bài báo được phân bố như sau: Phần 2 của bài báo, trình bày các nghiên cứu liên quan. Cơ sở lý thuyết được trình bày ở Phần 3. Mô hình nghiên cứu và ứng dụng sẽ được trình bày ở Phần 4. Phần 5 là kết quả thực nghiệm và đánh giá mô hình. Cuối cùng, kết luận và đề xuất hướng phát triển được trình bày ở Phần 6.

2. CÁC NGHIÊN CỨU LIÊN QUAN

Trong thời đại công nghệ thông tin phát triển, một khối lượng dữ liệu lớn được ghi lại dưới dạng văn bản. Bài toán phân tích cảm xúc là một phương pháp trích xuất thông tin có giá trị từ tập dữ liệu khổng lồ này. Các thông tin nhận xét đánh giá về sản phẩm là các dữ liệu dạng ngôn ngữ tự nhiên cụ thể là dữ liệu văn bản chứa các nhận xét đánh giá của khách hàng về một sản phẩm nào đó. Xử lý ngôn ngữ tự nhiên (natural language processing - NLP) là một nhánh của trí tuệ nhân tạo tập trung vào các ứng dụng trên ngôn ngữ của con người. Một trong những ứng dụng phổ biến nhất của xử lý ngôn ngữ tự nhiên là phân tích cảm xúc của con người với mong muốn xác định, trích xuất và nghiên cứu các trạng thái cảm xúc và thông tin chủ quan một cách có hệ thống thông qua các tài liệu như nhận xét, đánh giá và phản hồi khảo sát,... của khách hàng. Trong những năm gần đây, bài toán phân tích cảm xúc luôn nhận được quan tâm từ các nhà nghiên cứu. Cụ thể, Ahmed và cộng sự [1] đã đưa ra một cái nhìn tổng quan về các khái niệm khác nhau giữa các cấp độ và phương pháp phân tích cảm xúc. Bên cạnh đó, các tác giả đã đề xuất xây dựng bộ dữ liệu đa ngôn ngữ hoặc xây dựng một tập dữ liệu dùng chung được lấy từ ứng dụng của các phương tiện truyền thông xã hội khác nhau như Twitter, Youtube, Facebook,... Saif và đồng sự [9] đề xuất một cách tiếp cận mới là thêm đặc điểm ngữ nghĩa cho các thực thể được trích xuất từ tweet và bổ sung vào tập dữ liệu huấn luyện để phân tích tình cảm. Buche và cộng sự [2], Chen và cộng sự [11], Thanh và cộng sự [12], đã đề xuất một số phương pháp và kỹ thuật xử lý ngôn ngữ tự nhiên trong việc phân tích ý kiến và cảm xúc khách hàng thông qua bình luận trực tuyến. Các mô hình xử lý ngôn ngữ tự nhiên sử dụng một cơ chế tiền xử lý dữ liệu huấn luyện bằng một mô hình được đào tạo từ một lượng lớn các dữ liệu không được gán nhãn. Bên cạnh đó, các mô hình biểu diễn từ phổ biến như Word2vec, Glove hay FastText... không linh hoạt theo ngữ cảnh bởi vì nó sử dụng vector đại diện cho mỗi từ dựa trên một tập ngữ liệu lớn nên không thể hiện được sự đa dạng của ngữ cảnh cụ thể của từ trong từng lĩnh vực hay văn cảnh cụ thể.

Ngoài phương pháp xử lý ngôn ngữ tự nhiên thì bài toán phân tích cảm xúc cũng là một trong những nhiệm vụ phân loại văn bản trong đó dữ liệu đầu vào là các nhận xét đánh giá của người dùng về một sản phẩm cụ thể được gán nhãn tích cực hay tiêu cực. Trong khi đó, máy học là một ứng dụng của Trí tuệ nhân tạo, giúp hệ thống tự động hiểu dữ liệu từ dữ liệu huấn luyện mà không cần lập trình cụ thể. Máy học tập trung vào vấn đề cung cấp hệ thống tự động hiểu dữ liệu và thực hiện các phép phân loại, dự đoán. Hiện tại có nhiều mô hình được đề xuất sử dụng phương pháp tiếp cận máy học kết hợp với phương pháp dựa vào từ vựng để giải quyết bài toán phân tích cảm xúc. Pang và cộng sự [8] đã đề xuất sử dụng 3 phương pháp máy học bao gồm Naïve Bayse, MEM (Maximum Entropy Model) và SVM (support vector machines) để phân

loại xem 1 đánh giá phim là tích cực hay tiêu cực. Các kết quả thực nghiệm của các tác giả cho thấy hiệu suất là tương đối, trong đó thì phương pháp Naïve Bayes có hiệu suất phân loại kém nhất, còn SVM có hiệu suất phân loại tốt nhất. Tuy nhiên, độ chính xác của kết quả phân loại tình cảm không như mong muốn. Để cải tiến độ chính xác của nhiệm vụ phân loại, Zinuddin và cộng sự [15] đã đề xuất sử dụng SVM kết hợp với phương pháp lựa chọn đối tượng Chi-square. Các phương pháp tiếp cận được đề cập ở trên cũng được áp dụng trong các văn bản tiếng Việt như Duyen và các cộng sự [4] đã đề xuất sử dụng 3 phương pháp là Naïve Bayes, MEM và SVM để phân loại các đánh giá về các dịch vụ khách sạn trên trang web Agoda. Vo và cộng sự [13] đã đề xuất một mô hình mới bằng cách tích hợp các ưu điểm của CNN và LSTM. Kết quả thực nghiệm của nhóm tác giả cho thấy mô hình này có kết quả tốt hơn so với CNN, LSTM và SVM trên bộ dữ liệu gồm có 17.500 lượt đánh giá từ các trang mạng xã hội của Việt Nam đã được gán nhãn là tích cực, trung lập, tiêu cực. Vo và cộng sự [14] đã tích hợp quy trình nhúng từ vào quy trình huấn luyện về phân tích tình cảm từ đó đề xuất một mô hình mới bằng cách các vector từ được chuyển qua CNN và sau đó đầu ra được sử dụng làm đầu vào cho mạng LSTM để thực hiện phân loại.

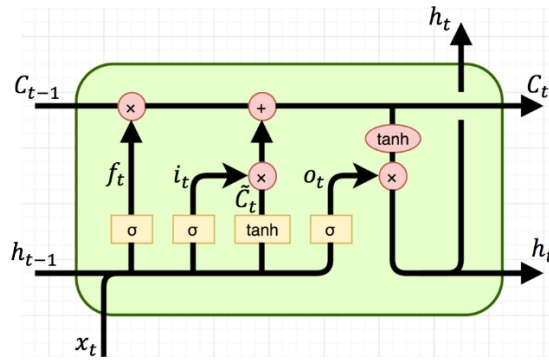
Năm 2018, Google đã công bố một kỹ thuật mới trong xử lý ngôn ngữ tự nhiên được gọi là BERT (Bidirectional Encoder Representations from Transformers) [3]. BERT là một mô hình biểu diễn từ theo 2 chiều sử dụng kỹ thuật Transformers. Đây là một mô hình đã được huấn luyện sẵn trên tập dữ liệu lớn theo cách tự giám sát (self-supervised). Nó chỉ được đào tạo trước trên các văn bản thô, không có con người ghi nhãn chúng theo bất kỳ cách nào với một quy trình tự động để tạo đầu vào và nhãn từ các văn bản đó. Hiện tại, BERT có nhiều phiên bản cải tiến, biến thể khác nhau như RoBERTa (Robustly Optimized BERT Pre-training Approach) [6], DistilBERT [10], mBERT (BERT-Base Multilingual) [3]. Hiện tại, Nguyen và cộng sự [7] đã đề xuất mô hình PhoBERT. Đây là một mô hình đã được huấn luyện sẵn dành riêng cho tiếng Việt. Việc đào tạo biểu diễn từ của PhoBERT dựa trên kiến trúc và cách tiếp cận của RoBERTa. Kết quả thực nghiệm của nhóm tác giả cho thấy PhoBERT hoạt động tốt trong bài toán xử lý ngôn ngữ tự nhiên dành riêng cho tiếng Việt.

Dựa trên các ưu điểm của các công trình trên, để thấu hiểu hơn về sản phẩm thông qua các nhận xét đánh giá của khách hàng, trong bài báo này, chúng tôi đề xuất và ứng dụng mô hình PhoBERT kết hợp với LSTM và mBERT kết hợp với LSTM để phân tích cảm xúc của khách hàng dựa trên dữ liệu đánh giá, nhận xét về quá trình đặt hàng, giao hàng, dịch vụ bán hàng và chất lượng của sản phẩm là các linh kiện máy tính, điện thoại, đồng hồ thông minh được thu thập trên website thương mại điện tử như Tiki, Shopee, FPT Shop...

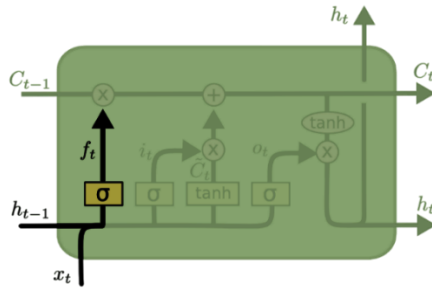
3. CƠ SỞ LÝ THUYẾT

Học sâu (Deep Learning) là một lĩnh vực con của máy học (Machine Learning), mặt khác, học sâu cũng là một lĩnh vực con của Trí tuệ nhân tạo (Artificial Intelligence). Nó cho phép máy học và đưa ra dự đoán dựa trên kinh nghiệm của nó (dữ liệu). Các thuật toán học sâu cố gắng đưa ra các kết luận tương tự như con người bằng cách liên tục phân tích dữ liệu với một cấu trúc logic nhất định. Để đạt được điều này, học sâu sử dụng cấu trúc nhiều lớp (multi-layered) của các thuật toán được gọi là neural networks. Học sâu có 2 kiến trúc lớn là Convolutional Neural Network (CNN) áp dụng cho bài toán có input là ảnh và Recurrent neural network (RNN) áp dụng cho bài toán dữ liệu dạng chuỗi (sequence). Trong xử lý ngôn ngữ tự nhiên, ý tưởng chính của RNN là sử dụng một bộ nhớ để lưu lại thông tin từ những bước tính toán xử lý trước để dựa vào đó có thể đưa ra dự đoán chính xác nhất cho bước dự đoán hiện tại.

LSTM (Long Short Term Memory) là một dạng đặc biệt của RNN, được thiết kế để tránh được vấn đề mất thông tin ở phía trước. Nó có thể nhớ thông tin trong suốt thời gian dài là không cần phải huấn luyện. Hình 1 mô tả kiến trúc hoạt động của LSTM.



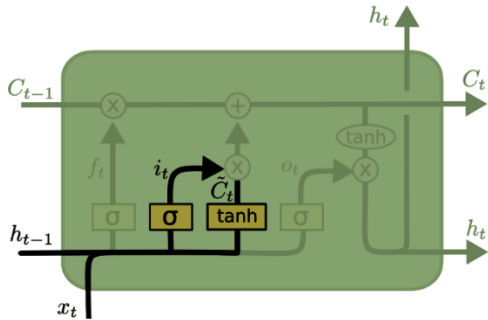
Hình 1: Sơ đồ LSTM cell [16]



$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f)$$

Hình 2: Forget gate [16]

Hình 2 mô tả Forget gate hoạt động trong LSTM. Forget gate có nhiệm vụ quyết định liệu thông tin của Cell state trước có cần được lưu trữ hay không. Thông tin từ Current input và Hidden state được chuyển qua hàm sigmoid với output nằm trong khoảng [0, 1]. Do đó nếu giá trị output gần bằng 1, tức là thông tin cần được giữ lại còn nếu output gần bằng 0 thì phải loại bỏ thông tin.

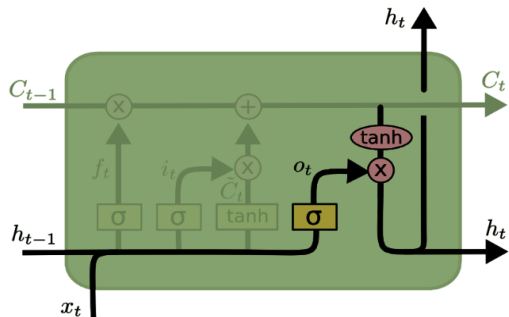


$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i)$$

$$\tilde{C}_t = \tanh(W_C \cdot [h_{t-1}, x_t] + b_C)$$

Hình 3: Input gate [16]

Hình 3 mô tả Input gate. Input gate có nhiệm vụ cập nhật thông tin vào Cell state. Ở đây ta nhân output của sigmoid với output của hàm tanh để quyết định thông tin của thông tin nào được cập nhật.

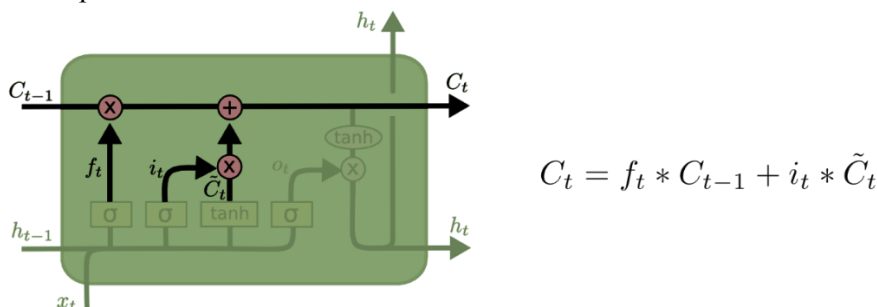


$$o_t = \sigma(W_o [h_{t-1}, x_t] + b_o)$$

$$h_t = o_t * \tanh(C_t)$$

Hình 4: Output gate [16]

Hình 4 mô tả Output gate. Output gate điều chỉnh lượng thông tin có thể ra ngoài y_t và lượng thông tin truyền tới cell tiếp theo.



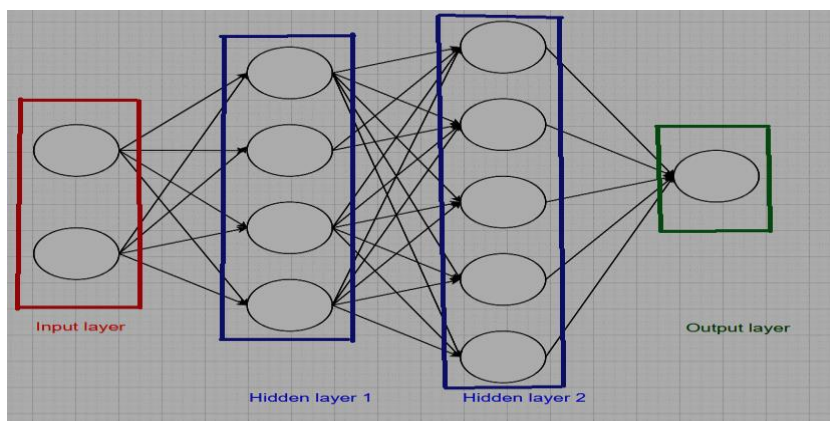
Hình 5: Bộ nhớ trong [16]

Cuối cùng, ta có c_t là bộ nhớ trong của LSTM được mô tả trong hình 5. Bộ nhớ trong tổng hợp thông tin trước đã được lọc qua Forget gate, cộng với trạng thái ẩn đã được lọc bởi Input gate. Bộ nhớ trong sẽ mang thông tin nào quan trọng truyền đi xa hơn và sẽ được dùng khi cần.

Hàm phi tuyến Tanh là hàm kích hoạt trong lớp ẩn. Tác dụng của hàm này là biến đổi những giá trị số thực về khoảng $[-1, 1]$ thông qua công thức $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$

Sigmoid là một hàm phi tuyến với đầu vào là các số thực và cho kết quả nằm trong khoảng $[0, 1]$ và được xem là xác suất trong một số bài toán với $\sigma(x) = \frac{1}{1 + e^{-x}}$

Fully Connected là một lớp được kết nối như một mạng nơ-ron được mô tả trong hình 6, ở đó mỗi node trong layer được kết nối với tất cả các node trong layer trước. Chức năng kết hợp với các đặc trưng lại với nhau sử dụng hàm softmax hoặc sigmoid để phân loại đầu ra.



Hình 6: Fully connected layers [17]

Transformer là một mô hình học sâu được thiết kế để phục vụ giải quyết nhiều bài toán trong xử lý ngôn ngữ và tiếng nói. Khác với RNNs, Transformer không xử lý các phần tử trong một chuỗi một cách tuần tự. Nếu dữ liệu đầu vào là một câu ngôn ngữ tự nhiên, Transformer không cần phải xử lý phần đầu câu trước rồi mới tới phần cuối câu. Do tính năng này, Transformer có thể tận dụng khả năng tính toán song song của GPU và giảm thời gian xử lý đáng kể. Mô hình Transformer được mô tả trong hình 7, ở đó input embedding là 1 vector biểu diễn câu chữ. Transformers có thêm một phần Positional Encoding (PE) [18] để thêm thông tin về vị trí của một từ vì cùng một từ ở vị trí khác nhau của câu có thể mang ý nghĩa khác nhau. PE được tính như sau:

$$PE_{(pos, 2i)} = \sin\left(\frac{pos}{10000^{2i/d_{model}}}\right)$$

$$PE_{(pos, 2i+1)} = \cos\left(\frac{pos}{10000^{2i/d_{model}}}\right)$$

Trong đó, pos là vị trí từ trong câu, d_{model} là số chiều vector embedding và i là vị trí trong vector embedding.

Kiến trúc của BERT bao gồm nhiều lớp bidirectional Transformer encoder. Các phiên bản đều dựa trên việc thay đổi kiến trúc của Transformer tập trung ở 3 tham số gồm L là số lượng các sub-layers trong transformer, H là kích thước của embedding vector (hay còn gọi là hidden size), và A là số lượng head trong multi-head layer. Tên gọi của 2 kiến trúc BERT là $BERT_{BASE}$ ($L = 12, H = 768, A = 12$) và $BERT_{LARGE}$ ($L = 24, H = 1024, A = 16$). Bên cạnh đó, BERT có nhiều phiên bản cải tiến, biến thể như RoBERTa [6,20], ALBERT [5], DistilBERT [10],... BERT là một mô hình đã được huấn luyện sẵn trên tập dữ liệu lớn theo cách tự giám sát (self-supervised). Đầu vào của BERT có thể là một câu đơn hoặc hai câu được gộp lại với nhau. BERT sử dụng 2 tác vụ huấn luyện không giám sát [3, 19] gồm Masked-Language Models (MLM) và Next Sentence Prediction (NSP).

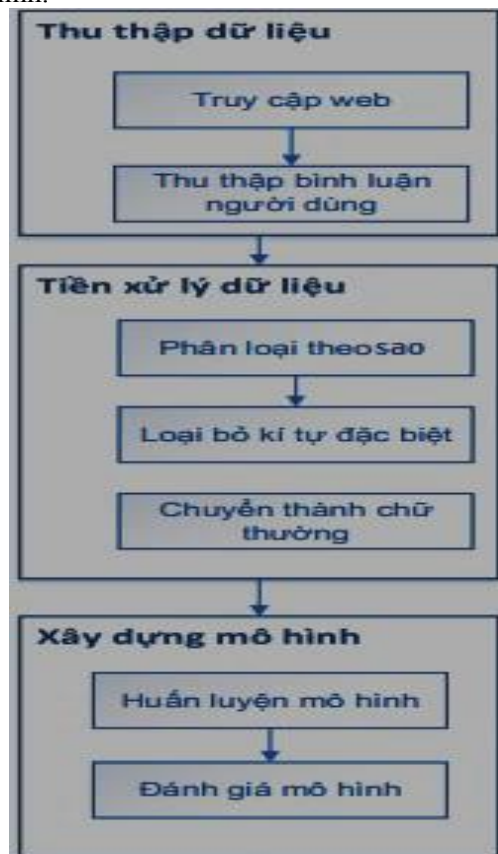
Kiến trúc RoBERTa [6, 20] gần như tương tự với kiến trúc BERT, nhưng quy trình huấn luyện của RoBERTa gồm các bước: loại bỏ Next Sentence Prediction (NSP), huấn luyện nhiều mẫu hơn trong một lần với câu đầu vào dài hơn, huấn luyện với tập dữ liệu lớn hơn.

PhoBERT [7] là một mô hình đã được huấn luyện sẵn dành riêng cho tiếng Việt. Việc huấn luyện của PhoBERT dựa trên kiến trúc và cách tiếp cận giống RoBERTa. Tương tự như BERT, PhoBERT cũng có 2 phiên bản là $PhoBERT_{BASE}$ và $PhoBERT_{LARGE}$, được huấn luyện trên khoảng 20GB dữ liệu bao gồm khoảng 1GB Vietnamese Wikipedia corpus và 19GB còn lại lấy từ Vietnamese news corpus.

mBERT [3] là mô hình BERT đã được huấn luyện sẵn trên tập dữ liệu lớn với 104 ngôn ngữ khác nhau trên Wikipedia. Bộ từ điển của mô hình này khoảng 110000 từ.

4. XÂY DỰNG MÔ HÌNH ỨNG DỤNG

Mô hình tổng quan của ứng dụng được thể hiện trong hình 9 bao gồm các bước như thu thập dữ liệu, tiền xử lý dữ liệu và xây dựng mô hình.

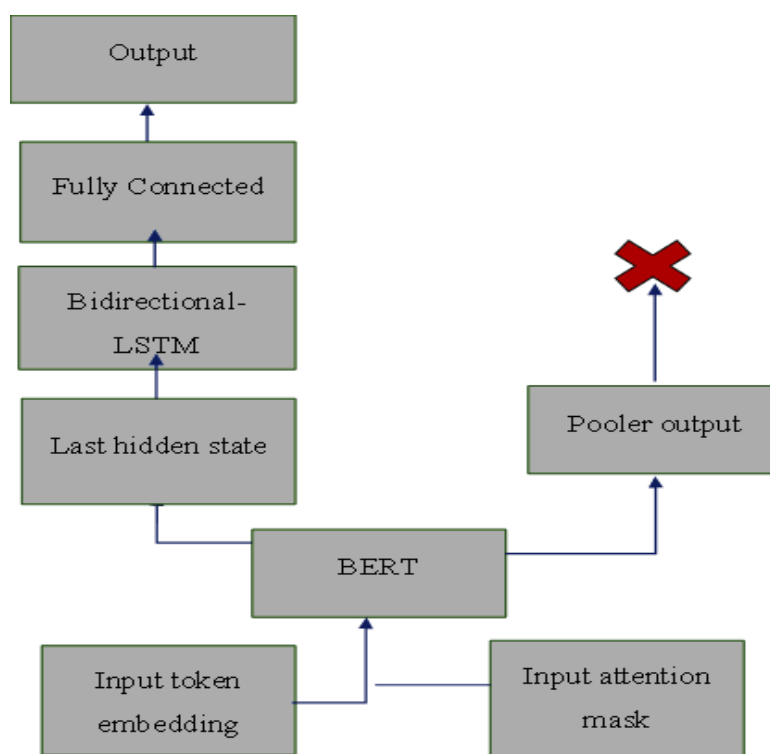


Hình 9. Mô hình tổng quan của ứng dụng

Mô hình đề xuất ở trên được hiện thực bằng ngôn ngữ lập trình Python và Google Colab pro (GPU) để huấn luyện mô hình. Các thực nghiệm được thực hiện trên bộ dữ liệu thực được thu thập trên các website thương mại điện tử như Tiki, Shopee, FPT Shop... Công cụ thu thập dữ liệu là Selenium được sử dụng để thu thập những câu nhận xét, đánh giá về các linh kiện máy tính, điện thoại, đồng hồ thông minh. Sau khi loại bỏ những câu nhận xét không có ý nghĩa, bộ dữ liệu gồm 19703 nhận xét đánh giá. Bộ dữ liệu thu thập được là dữ liệu thô chưa được xử lý nên có thể chứa dữ liệu bị thiếu, sai chính tả, từ viết tắt, ngôn ngữ teen, ký tự đặc biệt,... điều này có thể làm ảnh hưởng đến chất lượng của kết quả phân tích, do đó dữ liệu này cần được xử lý để làm sạch dữ liệu thông qua các bước như:

- Gán nhãn cho mỗi mẫu trong dữ liệu là 1 sao, 2 sao, 3 sao, 4 sao hay 5 sao.
- Chuẩn hoá bộ mã hoá ký tự về dạng các từ có dấu của Unicode dựng sẵn, loại bỏ dấu và các ký tự đặc biệt như (+-*/./,?\`~!@#%&="{};:;')
- Loại bỏ các stopwords tiếng Việt. Các stopwords thường là các từ nối (của, là, có, được, những,...) và các từ đặc trưng của dữ liệu như (tôi, mình,...). Bộ stopwords được tham khảo từ website: <https://github.com/stopwords/vietnamese-stopwords/blob/master/vietnamese-stopwords-dash.txt>;
- Chuẩn hoá các từ viết tắt. Các từ này được tham khảo trên website: <https://gist.github.com/nguyenvanhieuvn/7d9441c10b3c2739499fc5a4d9ea06fb>
- Đưa các từ về dạng chữ viết thường
- Thêm ngẫu nhiên bản sao các câu thuộc lớp có số lượng thấp hơn vào dữ liệu huấn luyện sao cho số lượng mẫu ở các lớp bằng nhau.

Quá trình gán nhãn được dựa trên ý kiến cá nhân và thực hiện theo nguyên tắc sau: các câu không có từ thể hiện tiêu cực như tệ, hư hỏng, xấu,... và có các từ tích cực như rất tốt, tốt, tuyệt, hài lòng... được gán nhãn 5 sao. Các câu có các từ như ổn, cũng ok, không như kỳ vọng lắm, lỗi một chút... được gán nhãn 4 sao. Các câu có các từ như không đẹp, không hài lòng lắm,... được gán nhãn 2 sao. Các câu chỉ có từ thể hiện tiêu cực được gán nhãn 1 sao. Phần còn lại các câu có cả từ tích cực và tiêu cực, các câu bản thân không phân loại được và bình luận đánh giá 3 sao sẽ gán nhãn 3 sao. Bộ dữ liệu sau khi tiền xử lý và gán nhãn được chia làm 2 tập dữ liệu bao gồm tập huấn luyện có 15829 mẫu (~80% mẫu) và tập kiểm thử có 3874 mẫu (~20% mẫu).



Hình 10. Xây dựng mô hình BERT kết hợp với LSTM

Bước xây dựng mô hình do bài báo đề xuất được trình bày trong hình 10. Mô hình được thực hiện như sau: đầu tiên tạo 2 đầu vào cho mô hình gồm “Input token embedding” và “Input attention mask”, sau đó cho 2 đầu vào ở trên vào mô hình BERT đã huấn luyện sẵn của thư viện Huggingface Transformer, và sau đó ta lấy đầu ra “Last hidden state” của mô hình để làm đầu vào của lớp LSTM 2 chiều. Kết quả của LSTM 2 chiều được đưa qua mạng fully connected và ta có đầu ra của mô hình.

5. ĐÁNH GIÁ MÔ HÌNH

Để đánh giá hiệu năng của mô hình đề xuất, chúng ta sử dụng các độ đo gồm độ chính xác Precision (P), độ bao phủ Recall (R), và độ đo trung bình điều hòa của Precision và Recall (F1-score) với cách thức tính cụ thể như sau:

Recall (Độ bao phủ): Tỷ lệ dương thực, chỉ số thể hiện trong tất cả các trường hợp Positive, bao nhiêu trường hợp đã được dự đoán chính xác: $Recall = TP / (TP + FN)$

Precision (Độ chính xác): Tỷ lệ dương được mô hình dự đoán đúng, là số lượng dự đoán được thực hiện chính xác: $Precision = TP / (TP + FP)$

F-score: Trung bình điều hòa của Precision và Recall: $F_1 = 2 * (Recall * Precision) / (Recall + Precision)$

Trong bài toán phân loại nhiều lớp thì:

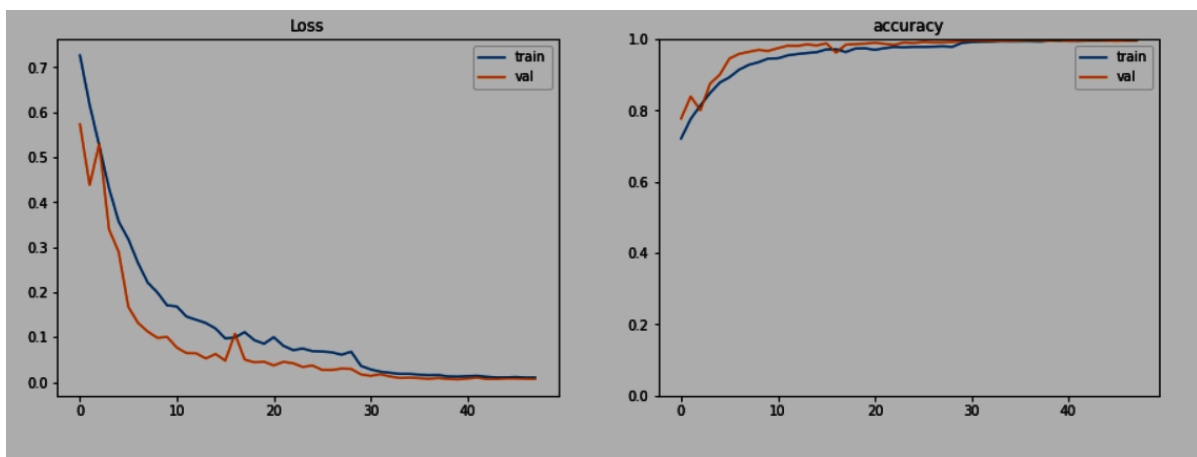
TP (True Positive): Số bình luận được phân lớp đúng.

FN (False Negative): Tổng số bình luận dự đoán sai trên hàng ứng với lớp phân loại.

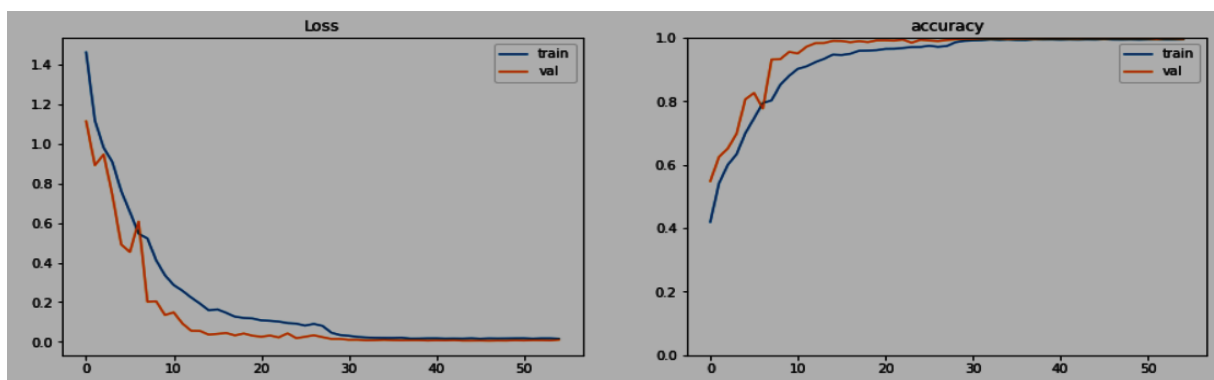
FP (False Positive): Tổng số bình luận dự đoán sai trên cột ứng với lớp phân loại.

Accuracy: độ chính xác trung bình các thuật toán, là tỷ lệ giữa kết quả dự đoán với dữ liệu thực tế.

Các kết quả thực nghiệm trên 2 mô hình là PhoBERT kết hợp với LSTM và mBERT kết hợp với LSTM được thể hiện trong hình 11, hình 12, và bảng 1. Hình 11 và hình 12 cho thấy Train Loss và Validation Loss của hai mô hình đều giảm, trong khi đó Train Accuracy và Validation Accuracy của hai mô hình tăng khá đều chứng tỏ mô hình học tốt không bị overfit.



Hình 11. Loss và Accuracy PhoBERT+ LSTM



Hình 12. Loss và Accuracy mBERT+ LSTM

Bảng 1. Kết quả thực nghiệm

Mô hình	PhoBERT+ LSTM					mBERT+ LSTM				
	1 sao	2 sao	3 sao	4 sao	5 sao	1 sao	2 sao	3 sao	4 sao	5 sao
Precision	90	76	71	83	95	81	62	56	66	90
Recall	89	76	80	83	92	78	63	67	70	82
F1-score	89	76	75	83	93	80	62	61	68	86
Accuracy	86,06 %					74,86 %				
Avg Precision	86 %					76 %				
Avg Recall	86 %					75 %				
Avg F1-score	86 %					75 %				
Thời gian huấn luyện	552s (495 ms/step)					503s (888ms/step)				
Thời gian dự đoán	42s (284ms/step)					36s (294ms/step)				

Dựa vào kết quả thực nghiệm trong bảng 1 chúng ta thấy rằng mô hình PhoBERT kết hợp với LSTM có kết quả phân lớp chính xác hơn so với mô hình mBERT kết hợp với LSTM trong tất cả các lớp. Tuy nhiên, mô hình mBERT kết hợp với LSTM có thời gian huấn luyện và dự đoán nhanh hơn mô hình PhoBERT kết hợp với LSTM nhưng thời gian nhanh hơn không đáng kể.

Kết quả đánh giá của các mô hình trên thực tế cho thấy cả hai mô hình đều phân loại khá tốt trong điều kiện là câu đúng cú pháp và chỉ dùng các từ viết tắt thông thường. Cụ thể, ví dụ như đánh giá của khách hàng Phùng Thị Khánh Ly trên Tiki về sản phẩm Điện Thoại iPhone 13 256GB - Hàng Chính Hãng: “Tiki có hệ thống vận chuyển, shipper riêng nên khi đặt hàng điện tử giá trị cao rất yên tâm. Không lo tráo hàng, nhận hàng phải có otp. Đặt hàng lúc 22h khuya, hôm sau sáng 9h30 đã nhận được. Gói băng keo kỹ, sản phẩm được niêm phong thêm bằng keo của tiki, nếu đã bị gỡ là biết ngay. Nguyên seal, phụ kiện, bảo hành đầy đủ. Màu hồng trên thị trường hiếm với mức hơn các màu còn lại tầm một đến hai triệu nhưng trên tiki tương tự nhau. Đã v trả góp không tốn đồng phí nào. Màu hồng siêu cung luôn á mng, dùng được 2 tuần rồi, check máy chính hãng bảo hành đầy đủ nhe.”. Với nhận xét này thì mô hình PhoBERT kết hợp với LSTM dự đoán 5 sao với tỉ lệ là 0.99971634, và mô hình mBERT kết hợp với LSTM dự đoán 5 sao với tỉ lệ là

0.9982823. Đây là kết quả dự đoán hợp lý của hai mô hình. Tuy nhiên, với đánh giá của khách hàng Khoa Tran trên Tiki về sản phẩm Điện Thoại OnePlus Nord CE 5G (12GB/256G) - Hàng Chính Hãng: “Sản phẩm hơi thất vọng so với kỳ vọng. Nhưng xài bt vẫn OK” thì mô hình PhoBERT kết hợp với LSTM dự đoán 3 sao với tỉ lệ là 0.65818155, trong khi đó mô hình mBERT kết hợp với LSTM thì dự đoán 4 sao với tỉ lệ là 0.949803. Với kết quả dự đoán này, ta thấy mô hình PhoBERT+ LSTM có kết quả chính xác hơn. Bên cạnh đó cũng có những bình luận có hàm ý tương tự nhưng các mô hình cho ra kết quả phân loại khác nhau, cụ thể với đánh giá như “Tốt hơn hàng bán bên Shopee rất nhiều” thì mô hình PhoBERT kết hợp với LSTM dự đoán 4 Sao với tỉ lệ là 0.99722505, còn mô hình mBERT kết hợp với LSTM thì dự đoán 4 sao với tỉ lệ là 0.5237757. Hoặc với đánh giá “Tốt hơn hàng bán bên Shopee nhiều” thì mô hình PhoBERT kết hợp LSTM dự đoán là 5 Sao với tỉ lệ: 0.9996828, mô hình mBERT kết hợp với LSTM dự đoán là 5 sao với tỉ lệ là 0.99982494.

Dựa trên kết quả thực nghiệm ở trên, ta thấy cả hai mô hình trên đều cho ra các kết quả phân lớp tương đối chính xác. Tuy nhiên, mô hình PhoBERT kết hợp với LSTM là mô hình có độ chính xác phân lớp tốt hơn.

6. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Trong bài báo này, chúng tôi đã ứng dụng các mô hình biến thể của BERT gồm PhoBERT và mBERT kết hợp với LSTM vào việc nhận diện cảm xúc khách hàng trên tập dữ liệu nhận xét đánh giá của người tiêu dùng về các sản phẩm linh kiện máy tính, điện thoại, đồng hồ thông minh được thu thập từ các website thương mại điện tử như Tiki, Shopee, FPT Shop,... Đây là một bài toán có tính ứng dụng cao có thể áp dụng trong nhiều vấn đề thực tế, đặc biệt liên quan đến việc phân tích cảm nhận phản hồi của khách hàng về sản phẩm. Kết quả thực nghiệm cho thấy mô hình PhoBERT kết hợp với LSTM có kết quả phân loại tốt hơn so với mBERT kết hợp với LSTM và đặc biệt thực thi tốt trên các câu nhận xét đánh giá có đầy đủ dấu. Các kết quả nghiên cứu trên có thể sử dụng làm cơ sở cho việc xây dựng những hệ thống phân loại đánh giá tự động ở trên các website thương mại điện tử của Việt Nam. Tuy nhiên, công việc gán nhãn cho các nhận xét đánh giá được thực hiện thủ công, do đó dẫn đến tình trạng thiếu khách quan. Bên cạnh đó, bài báo chỉ tập trung phân loại mức độ cảm xúc mà chưa thực sự phân loại các cảm xúc như: vui, buồn, ngạc nhiên, sợ hãi, tức giận hoặc nhiều hơn nữa. Trong tương lai, chúng tôi sẽ tìm hiểu, nghiên cứu phương pháp gán nhãn tự động và phát triển ứng dụng cho phép phát hiện cảm xúc trong các đoạn hội thoại.

TÀI LIỆU THAM KHẢO

- [1]. K. Ahmed, N. E. Tazi and A. H. Hossny, "Sentiment Analysis over Social Networks: An Overview," 2015 IEEE International Conference on Systems, Man, and Cybernetics, pp. 2174-2179, 2015.
- [2]. A. Buche, D. Chandak, A. Zadgaonkar. Opinion mining and analysis: A survey. International Journal on Natural Language Computing (IJNLC), vol.2, no.3, pp. 39-48, 2013.
- [3]. J. Devlin, M. Chang, K. Lee, K. Toutanova. BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. In Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, vol.1, pp. 4171–4186, 2019.
- [4]. N. T. Duyen, N. X. Bach and T. M. Phuong. An empirical study on sentiment analysis for Vietnamese, 2014 International Conference on Advanced Technologies for Communications (ATC 2014), pp. 309-314, 2014.
- [5]. Z. Lan, M. Chen, S. Goodman, K. Gimpel, P. Sharma, R. Soricut. ALBERT: A Lite BERT for Self-supervised Learning of Language Representations. 2020.
- [6]. Y. Liu, M. Ott, N. Goyal, J. Du, M. Joshi, D. Chen, O. Levy, M. Lewis, L. Zettlemoyer, V. Stoyanov. “RoBERTa: A Robustly Optimized BERT Pretraining Approach”, 2019.
- [7]. D. Q. Nguyen, A.T. Nguyen. PhoBERT: Pre-trained language models for Vietnamese. In Findings of the Association for Computational Linguistics: EMNLP 2020, pp. 1037–1042, 2020.

- [8]. B. Pang, L. Lee, S. Vaithyanathan. Sentiment classification using machine learning techniques, Proceedings of the ACL-02 conference on Empirical methods in natural language processing, vol. 10, 2002.
- [9]. H. Saif, Y. He, H. Alani. Semantic Sentiment Analysis of Twitter. In: Proceedings of the 11th international conference on the Semantic Web (ISWC), Boston, MA, 2012.
- [10]. V. Sanh, L. Debut, J. Chaumond, T. Wolf. DistilBERT, a distilled version of BERT: smaller, faster, cheaper and lighter, at the 5th Workshop on Energy Efficient Machine Learning and Cognitive Computing - NeurIPS 2019, 2019.
- [11]. S. Sun, C. Luo, J. Chen. A review of natural language processing techniques for opinion mining systems. Information Fusion, vol. 36, pp.10-25, 2017.
- [12]. H. Thanh, D. Phuc. Analyzing users' interests with the temporal factor based on topic modeling. In Asian conference on intelligent information and database systems, Pa Tong, Thailand: Springer, pp. 105- 115, 2015.
- [13]. Q. -H. Vo, H. -T. Nguyen, B. Le, M. -L. Nguyen. Multi-channel LSTM-CNN model for Vietnamese sentiment analysis. 9th International Conference on Knowledge and Systems Engineering (KSE), pp. 24-29, 2017.
- [14]. K. Vo, T. Nguyen, D. Pham, M. Nguyen, M. Truong, D. Nguyen, T. Quan. Handling negative mentions on social media channels using deep learning Journal of Information and Telecommunication, vol.3, no.3, pp. 271-293, 2019.
- [15]. N. Zainuddin and A. Selamat. Sentiment analysis using Support Vector Machine, 2014 International Conference on Computer, Communications, and Control Technology (I4CT), pp. 333-337, 2014.
- [16]. <https://www.quora.com/What-is-the-difference-between-cell-state-denoted-c-t-and-the-output-of-the-state-denoted-by-h-t-in-a-recurrent-neural-network>
- [17]. <https://cdspninhthuan.edu.vn/fully-connected-layer-la-gi/>
- [18]. Website <https://machinelearningmastery.com/a-gentle-introduction-to-positional-encoding-in-transformer-models-part-1>
- [19]. Website <https://phamdinhhkhanh.github.io/2020/05/23/BERTModel.html>
- [20]. Website: <https://www.geeksforgeeks.org/overview-of-roberta-model>

YSC4F.205

HỆ THỐNG PHÁT HIỆN NGƯỜI ĐEO VÀ KHÔNG ĐEO KHẨU TRANG DỰA TRÊN MÔ HÌNH HỌC SÂU

NGUYỄN HOANH, NGUYỄN CÔNG TUỐC, PHẠM THANH NGÂN,
NGUYỄN PHI TRƯỜNG, HUỖNH ĐỨC HẢI

Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

*nguyenhoanh@iuh.edu.vn, 18054561.tuoc@student.iuh.edu.vn, 18035031.ngan@student.iuh.edu.vn,
18028891.truong@student.iuh.edu.vn, 18032251.hai@student.iuh.edu.vn.*

Tóm tắt. Vai trò của việc đeo khẩu trang ở nơi công cộng trở nên quan trọng, góp phần chính vào việc hạn chế sự lây lan COVID-19. Trong bài báo này, nhóm tác giả xây dựng hệ thống phát hiện người đeo/không đeo khẩu trang dựa trên nền tảng học sâu với ngõ vào là hình ảnh thu được từ camera được lắp đặt ở nơi công cộng. Hệ thống dựa trên mô hình FCOS, là một mô hình học sâu cho phát hiện đối tượng, để phát hiện các trường hợp đeo khẩu trang/không đeo khẩu trang trên hình ảnh thu được. Nhằm cải thiện độ chính xác và tốc độ thực thi, trong bài báo này sử dụng các phương pháp tăng cường dữ liệu để tăng số lượng dữ liệu huấn luyện và sử dụng mạng nơ ron tích chập ResNet. Kết quả huấn luyện và kiểm tra trên bộ dữ liệu cho thấy hệ thống có thể phát hiện được người đeo khẩu trang với tỉ lệ là 95,1% và người không đeo khẩu trang là 91.3%. Ngoài ra, nhóm tác giả còn xây dựng một giao diện trực quan, đầy đủ các thành phần để tạo thành một phần mềm hoàn chỉnh, thân thiện và dễ sử dụng với người dùng. Kết quả thực thi với các ảnh và video thực tế cho thấy hệ thống hoạt động ổn định và đạt được tỉ lệ chính xác cao.

Từ khóa. COVID-19, FCOS, ResNet, học sâu, đeo khẩu trang, không đeo khẩu trang.

FACE MASK DETECTION SYSTEM BASED ON DEEP LEARNING

Abstract. The role of wearing a mask in public places becomes important, it is making a major contribution to limiting the spread of COVID-19. In this article, the authors build a mask-wearing/non-masking detection system based on deep learning with the input of images obtained from cameras installed in public places. The system is based on FCOS model, which is a deep learning model for objects detection, to detect mask wearing/non-masking cases on the acquired images. In order to improve the accuracy and execution speed, in this article, data enhancement methods are used to increase the amount of training data and use ResNet convolutional neural network. The results of training and testing on the dataset show that the system can detect people wearing a mask with a rate of 95.1% and people without a mask at 91.3%. In addition, the author team also built an intuitive interface, full of components to form a complete, user-friendly and easy-to-use software. The performance results with actual photos and videos show that the system operates stably and achieves a high accuracy rate.

Keywords. COVID-19, FCOS, ResNet, deep learning, mask, no mask.

1. GIỚI THIỆU

Với sự phát triển mạnh mẽ của học sâu trong những năm gần đây, các mô hình phát hiện đối tượng dựa trên học sâu được phát triển mạnh mẽ và cho thấy được độ chính xác vượt trội so với các phương pháp truyền thống. Hơn nữa, các mô hình phát hiện đối tượng dựa trên học sâu còn cho thấy được sự hiệu quả với các ứng dụng thực tế khi vẫn cho độ chính xác cao trong nhiều môi trường khác nhau và nhiều loại đối tượng khác nhau về kích thước, tỉ lệ, màu sắc,... Các mô hình phát hiện đối tượng dựa trên học sâu được phát triển để phát hiện nhiều loại đối tượng khác nhau như xe cộ, khuôn mặt, biển số xe, biển báo giao thông,... Các mô hình này thường bao gồm một mạng nơ ron tích chập ở ngõ vào để trích xuất các đặc trưng trên ảnh và một mạng để phát hiện đối tượng ở cuối cùng để phát hiện đối tượng dựa trên các đặc trưng đã trích xuất

trước đó. Một số mô hình phát hiện đối tượng phổ biến như mô hình Faster R-CNN [1], SSD (Single Shot MultiBox Detector) [2], FCOS (Fully Convolutional One-stage Object Detection) [3], FPN (Feature Pyramid Network) [4],...

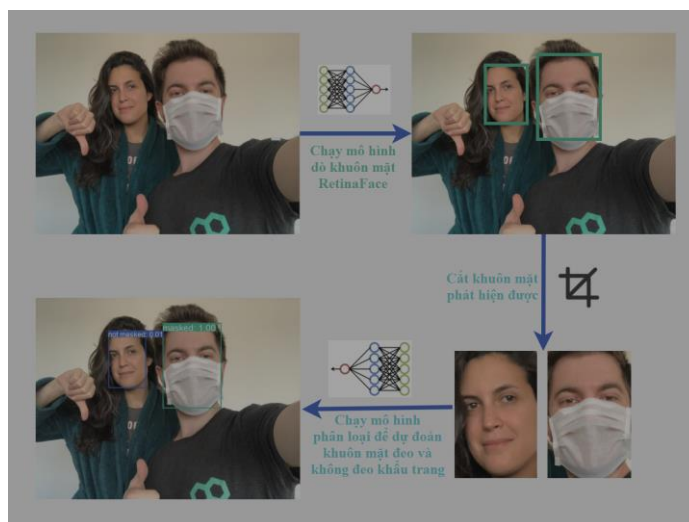
Với các đặc điểm phân tích ở trên, trong bài báo này đề xuất xây dựng một hệ thống phát hiện người đeo/không đeo khẩu trang dựa trên hình ảnh. Hệ thống đề xuất được xây dựng dựa trên mô hình FCOS, là mô hình đã được phát triển cho bài toán phát hiện đối tượng dựa trên hình ảnh. Hệ thống được xây dựng có thể phát hiện tất cả các khuôn mặt đeo/không đeo khẩu trang có trong hình ảnh thu được từ môi trường thực tế. Hệ thống xây dựng được huấn luyện và kiểm tra trên tập dữ liệu Face Mask Detection [5].

2. CÁC CÔNG TRÌNH LIÊN QUAN

Hệ thống phát hiện khuôn mặt đeo khẩu trang là một hệ thống phức tạp, yêu cầu giải quyết cùng lúc hai bài toán. Thứ nhất, hệ thống phải giải quyết bài toán phân lớp, phân biệt được khuôn mặt đeo khẩu trang và khuôn mặt không đeo khẩu trang. Thứ hai, hệ thống phải giải quyết bài toán phát hiện, định vị chính xác vị trí của các khuôn mặt có đeo khẩu trang/không đeo khẩu trang trên ảnh.

Có một số hướng tiếp cận để giải quyết bài toán. Trong [6], nhóm tác giả đề xuất sử dụng một hệ thống gồm hai hệ thống con để giải quyết bài toán phát hiện khuôn mặt đeo, không đeo khẩu trang. Sơ đồ tổng quan của hệ thống này như Hình 1. Hệ thống được đề xuất này gồm có hai hệ thống con. Hệ thống con thứ nhất sử dụng mô hình RetinaFace detector [7] cho bài toán phát hiện khuôn mặt. Hệ thống thứ hai sử dụng mạng MobileNet V1 [8] cho bài toán phân lớp khuôn mặt đeo khẩu trang/không đeo khẩu trang. Kết quả phát hiện khuôn mặt trên ảnh của hệ thống con thứ nhất sẽ là ngõ vào của hệ thống con thứ hai. Ngõ vào của hệ thống con thứ nhất là ảnh chứa các khuôn mặt có đeo khẩu trang/ không đeo khẩu trang. Hướng tiếp cận này có một nhược điểm là hệ thống phức tạp do phải có hai giai đoạn phát hiện và nhận dạng độc lập nhau dẫn đến tốc độ thực thi chậm. Tuy nhiên, hệ thống này cũng có ưu điểm là vì tách rời hai giai đoạn phân lớp khuôn mặt và phát hiện khuôn mặt nên hệ thống có thể phát hiện được các khuôn mặt trong các điều kiện khó khăn [9]. Một hướng tiếp cận khác là sử dụng một mô hình duy nhất cho cả hai bài toán nhận dạng và phát hiện. Với hướng tiếp cận này, mô hình có thể được xây dựng dựa trên các thuật toán xử lý ảnh truyền thống hoặc dựa trên mô hình học sâu. Các mô hình dựa trên học sâu luôn cho độ chính xác cao hơn so với các mô hình sử dụng xử lý ảnh truyền thống. Trong [10] đề xuất xây dựng mô hình phát hiện khuôn mặt thời gian thực dựa trên mô hình YOLO (You Only Look Once) v4 kinh điển. Mô hình được đề xuất bao gồm một số khối cải tiến để giúp hệ thống phát hiện người đeo khẩu trang tốt hơn, đặc biệt là trong điều kiện trời tối. Trong bài báo này, nhóm tác giả đề xuất xây dựng hệ thống phát hiện người đeo/không đeo khẩu trang dựa trên học sâu. Cụ thể hơn, mô hình FCOS được sử dụng để phát hiện khuôn mặt đeo/không đeo khẩu trang. Mô hình được xây dựng có độ chính xác cao và tốc độ thực thi nhanh trong điều kiện thực tế.

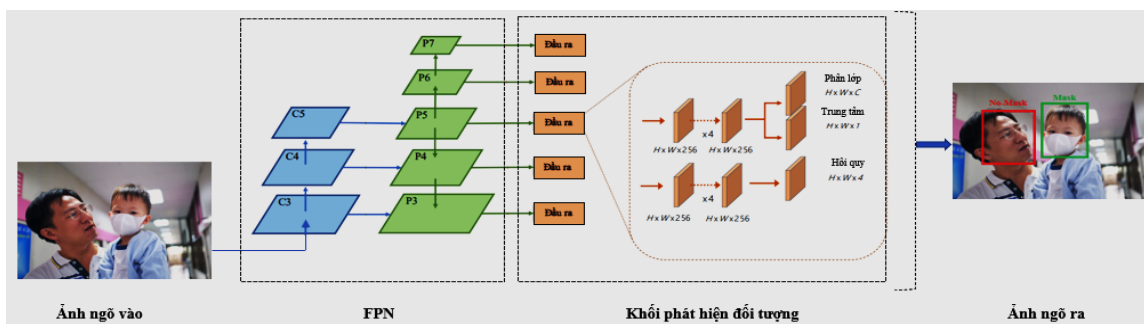
Chi tiết về hệ thống được xây dựng trong bài báo này được trình bày chi tiết trong các phần sau.



Hình 1: Mô hình đề xuất trong [6].

3. MÔ HÌNH ĐỀ XUẤT

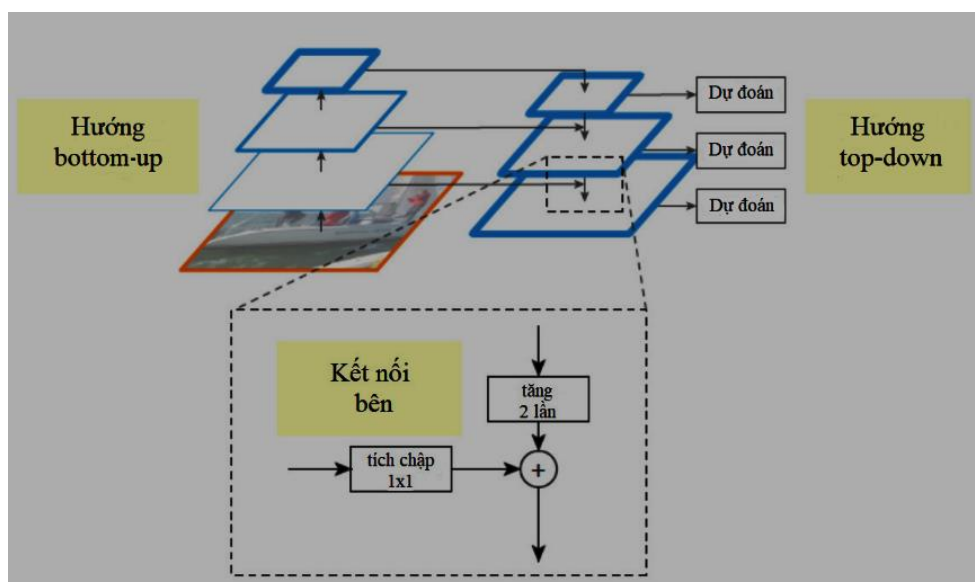
Bài báo này đề xuất sử dụng mô hình FCOS cho hệ thống phát hiện người đeo khẩu trang. Mô hình FCOS là một mô hình đã được đề xuất vào năm 2019 [3] dùng để phát hiện đối tượng trong ảnh. Sự ra đời của mô hình FCOS là một dấu mốc cho sự lột xác hoàn toàn khác so với các mô hình phát hiện đối tượng đã có như RetinaNet, SSD, YOLO v3 và Faster R-CNN dựa vào neo để xác định các hộp. Mô hình FCOS được đề xuất là mô hình không sử dụng neo cũng như đề xuất neo. Bằng cách loại bỏ neo được xác định trước. Mô hình FCOS hoàn toàn tránh được các tính toán phức tạp liên quan đến neo chẳng hạn như tính toán chùng chéo trong quá trình huấn luyện. Quan trọng hơn, nhóm tác giả cũng tránh tất cả các tham số liên quan đến neo. Hình 2 mô tả chi tiết cấu trúc của mô hình FCOS. Như mô tả trong Hình 2, mô hình FCOS gồm có 2 khối chính: Khối FPN và khối phát hiện đối tượng. Đầu tiên, ảnh sẽ được trích xuất các đặc trưng thông qua một mạng nơ-ron tích chập. Các bản đồ đặc trưng ở tầng cuối cùng của khối trích xuất đặc trưng sẽ được sử dụng cho khối tiếp theo. Cuối cùng, dựa trên các đối tượng có kích thước cố định, khối phát hiện đối tượng sẽ cho ra kết quả cuối cùng. Chi tiết các khối của mô hình FCOS được trình bày sau đây. Ngoài các khối chính, một số phương pháp xử lý bên trong mỗi khối cũng sẽ được trình bày chi tiết. Bài báo này đề xuất mô hình phát hiện đối tượng một giai đoạn hoàn toàn tích hợp (FCOS) để giải quyết việc phát hiện đối tượng theo kiểu dự đoán trên mỗi pixel.



Hình 2: Mô hình FCOS [3].

3.1. FPN (Feature Pyramid Networks)

Trong khối FPN [4], dữ liệu ảnh đầu vào được đi qua hai hướng là hướng Bottom-up và hướng Top-down. Hướng Bottom-up của FCOS đề xuất sử dụng mạng nơ-ron tích chập ResNet trong mô hình. Hướng Top-down dùng các lớp tích chập để tái tạo lại hình ảnh có độ phân giải cao với giàu thông tin về ngữ cảnh.



Hình 3. Kiến trúc của FPN [4].

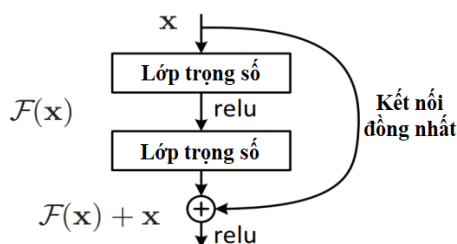
3.1.1. Mạng ResNet (Hướng Bottom-up)

Để trích xuất các đặc trưng từ ảnh ngõ vào, hướng Bottom-up của FCOS đề xuất sử dụng mạng nơ ron tích chập ResNet trong mô hình. Mạng ResNet là một mạng nơ ron tích chập gồm nhiều tầng sử dụng khối Residual. Hình 4 mô tả cấu trúc của một khối Residual. Trọng tâm của khối chính là kết nối đồng nhất giữa các tầng với nhau. Ngõ ra của khối Residual được tính theo công thức sau:

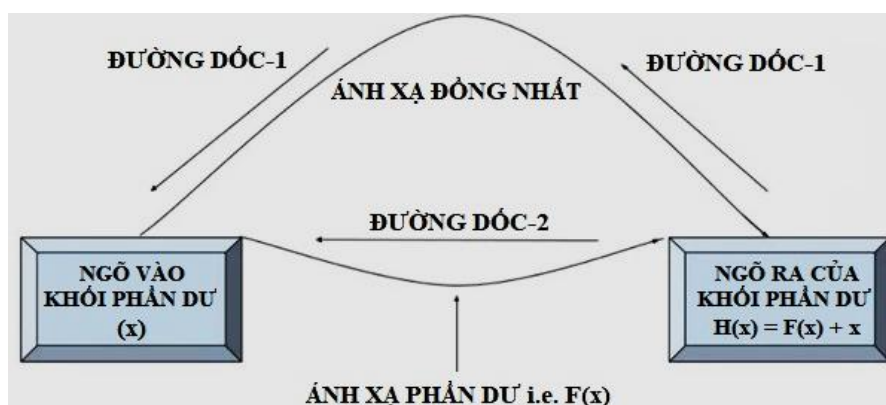
$$y = f(x, \{Wi\}) + x \quad (1)$$

Với x là ngõ vào của khối, Wi là các trọng số ở các tầng.

Với kết nối đồng nhất, quá trình thực thi thuật toán lan truyền ngược để cập nhật giá trị trọng số được mô tả như hình 5. Từ hình 5 ta có thể thấy là quá trình lan truyền ngược sẽ có 2 nhánh là Đường dốc 1 và Đường dốc 2. Đường dốc 1 sẽ là nhánh để cập nhật lại giá trị các trọng số và tính toán lại đạo hàm của sai số dựa vào giá trị trọng số mới. Việc cập nhật này sẽ làm giá trị đạo hàm của sai số giảm dần qua các tầng. Đường dốc 2 là phương pháp hiệu quả để giải quyết vấn đề này. Vì Đường dốc 2 không đi qua tầng có trọng số, do đó giá trị đạo hàm sai số của nó không bị suy giảm, kết quả là đạo hàm của sai số có thể lan truyền ngược đến các tầng ngõ vào để cập nhật giá trị trọng số. Một chú ý là kết nối đồng nhất trong khối Residual là phép toán đơn giản do đó không cần thêm bất cứ thông số nào vào khối, kết quả là mức độ học của khối Residual tương tự lớp tích chập thông thường trong khi có thể tận dụng để xây dựng mạng học sâu nhiều tầng.



Hình 4: Khối Residual để xây dựng mạng ResNet [4].



Hình 5. Thuật toán lan truyền ngược trong khối Residual [11].

3.1.2. Cấu trúc mạng ResNet

Dựa trên các khối Residual, các cấu hình mạng ResNet khác nhau đã được đề xuất, như ResNet-18 [12], ResNet-34 [13], ResNet-50 [14], ResNet-101, ResNet-110, ResNet-152, ResNet-164, ResNet-1202. Các cấu hình này chỉ khác nhau về số lớp. Trong các cấu hình này thì cấu hình ResNet-34 được sử dụng phổ biến do tính hiệu quả cả về độ chính xác và tốc độ thực thi. Trong bài báo này, nhóm tác giả sẽ đi chi tiết cấu hình mạng ResNet-34. Bảng 1 mô tả chi tiết cấu trúc của mạng ResNet-34. Mạng ResNet-34 gồm có 34 tầng. Ảnh ngõ vào sẽ được cho qua 1 tầng tích chập có kích thước bộ lọc là 7×7 theo sau là 1 tầng max pooling có kích thước bộ lọc là 3×3 . Sau đó, sẽ có 4 tầng tích hợp, tầng tích hợp thứ nhất có 3 khối residual,

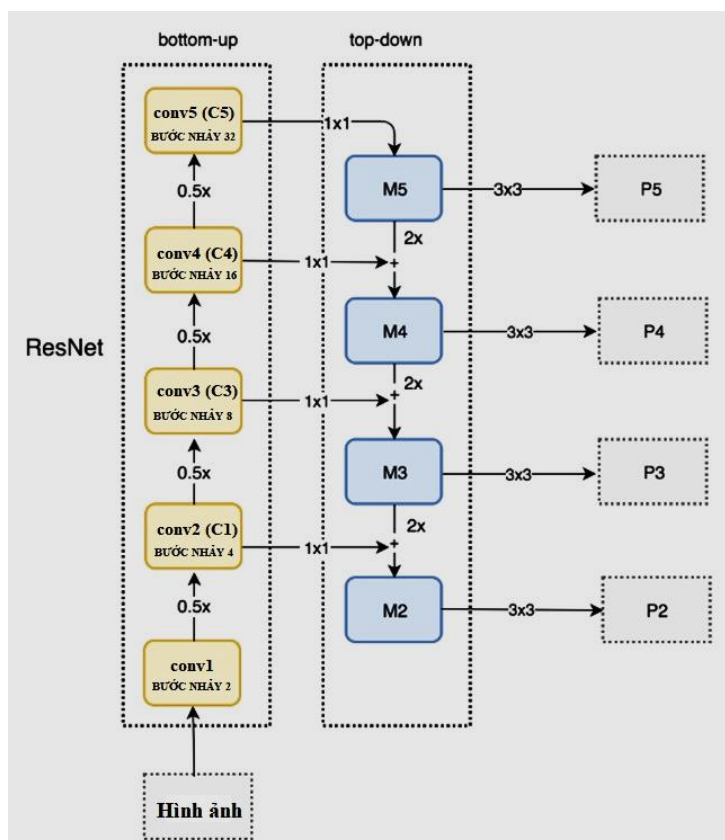
tầng tích hợp thứ 2 có 4 khối residual, tầng tích hợp thứ 3 có 6 khối residual và tầng tích hợp cuối cùng có 3 khối residual. Sau mỗi tầng tích hợp, kích thước của feature map sẽ được giảm đi một nửa nhờ sử dụng tầng tích hợp có stride là 2 bên trong mỗi khối residual. Đồng thời, số kênh của bản đồ đặc trưng cũng được tăng gấp đôi sau mỗi tầng tích hợp.

Bảng 1. Cấu hình chi tiết mạng ResNet-34 [11].

Tên lớp	Kích thước ngõ ra	18 lớp	34 lớp	50 lớp	101 lớp	152 lớp
conv1	112 x 112	7x7, 64, stride 2				
conv2_x	56x56	3x3 max pool, stride 2				
		$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 64 \\ 3 \times 3, & 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 64 \\ 3 \times 3, & 64 \\ 1 \times 1, & 256 \end{bmatrix} \times 3$
conv3_x	28x28	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 128 \\ 3 \times 3, & 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, & 128 \\ 3 \times 3, & 128 \\ 1 \times 1, & 512 \end{bmatrix} \times 8$
conv4_x	14x14	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 256 \\ 3 \times 3, & 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 23$	$\begin{bmatrix} 1 \times 1, & 256 \\ 3 \times 3, & 256 \\ 1 \times 1, & 1024 \end{bmatrix} \times 36$
conv5_x	7x7	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 2$	$\begin{bmatrix} 3 \times 3, & 512 \\ 3 \times 3, & 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, & 512 \\ 3 \times 3, & 512 \\ 1 \times 1, & 2048 \end{bmatrix} \times 3$
	1x1	average pool, 1000-d fc, softmax				
FLOPs		1.8×10^9	3.6×10^9	3.8×10^9	7.6×10^9	11.3×10^9

3.1.3. Khối tái tạo đặc trưng của đối tượng (Hướng Top-down)

Trong [15], hướng Top-down của FPN sử dụng lớp tích chập 1x1 để giảm số kênh của lớp C5 để tạo ra lớp M5. M5 là lớp bản đồ đặc trưng đầu tiên được dùng để dự đoán đối tượng. Theo chiều đi xuống của mũi tên, nhóm tác giả sử dụng thuật toán làm tăng độ phân giải của lớp phía trên lên hai lần. Sau đó tiếp tục sử dụng lớp tích chập 1x1 cho feature map C4 và cộng chúng lại với nhau để thu được bản đồ đặc trưng M4. Để giảm hiệu ứng răng cưa của M4, nhóm tác giả áp dụng lớp tích chập 3x3 để tạo ra bản đồ đặc trưng P4. Lập lại quá trình này để có được P3, P2. Không thực hiện tới P1 vì kích thước bản đồ đặc trưng C1 quá lớn sẽ làm giảm tốc độ xử lý.

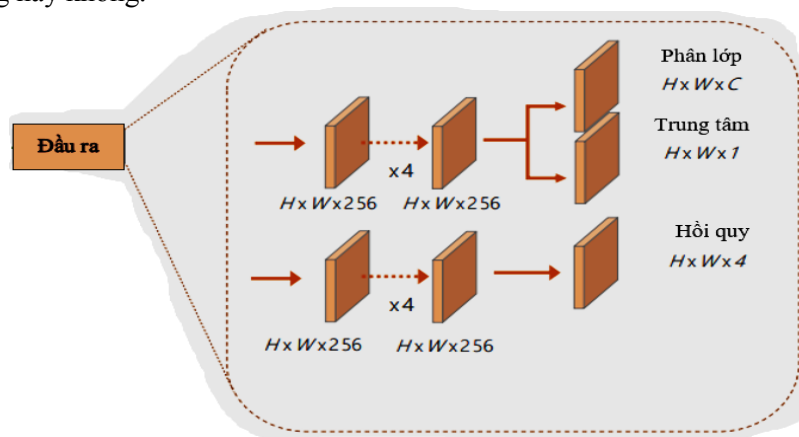


Hình 6: Khối đặc trưng của đối tượng [15].

3.2. KHỐI PHÁT HIỆN ĐỐI TƯỢNG

3.2.1. Khối phân lớp (Classification)

Ở nhánh phân lớp, mô hình sẽ dự đoán xác suất lớp trên mỗi pixel được tính bởi điểm trung tâm. Xác suất lớp có được bằng cách nhân xác suất lớp với điểm trung tâm, thay vì dùng khái niệm neo, thì mô hình sẽ đưa các giá trị trên từng pixel cho qua mạng học, từ đó sẽ cho biết pixel này có chứa hoặc có nằm trên vùng thuộc về đối tượng hay không.



Hình 7. Khối phân lớp [3].

Mục đích của mô hình huấn luyện là tìm ra hàm $f(x)$ và thông qua hàm f tìm được để gán nhãn cho dữ liệu, bước này thường được gọi là học hay huấn luyện.

$$f(x) = y \quad (2)$$

- x: các đặc trưng hay đầu vào của dữ liệu.
- y: nhãn lớp hay đầu ra.

Thông thường để xây dựng mô hình phân lớp cho bài toán này cần sử dụng các thuật toán học giám sát (supervised learning) như k-nearest neighbors [16], Neural Network [17], SVM, Decision tree, Naive Bayes.

3.2.2. Khối trung tâm (Center-ness)

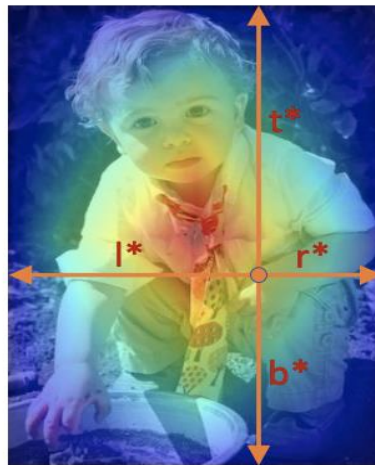
Sau khi sử dụng dự đoán đa cấp trong FCOS, có nhiều hộp giới hạn chất lượng thấp được tạo ra và nằm cách xa tâm của vật thể. Vì vậy, việc đề xuất ra một chiến lược đơn giản giải quyết hiệu quả vấn đề này là vô cùng cần thiết. Cụ thể, bổ sung thêm một nhánh lớp đơn, song song với nhánh phân loại để dự đoán “tâm điểm” của một vị trí cụ thể [3].

Trung tâm mô tả khoảng cách chuẩn hóa từ vị trí trung tâm của vật thể mà vị trí đó chịu trách nhiệm hồi quy cho các thông số l^* , r^* , t^* , b^* và được xác định qua công thức:

$$\text{Trung tâm}^* = \sqrt{\frac{\min(l^*, r^*)}{\max(l^*, r^*)} \times \frac{\min(t^*, b^*)}{\max(t^*, b^*)}} \quad (3)$$

Với l^* , r^* , t^* , b^* là khoảng cách tương ứng từ vị trí trung tâm đến bốn cạnh trái, phải, trên, dưới trên hộp giới hạn và các mục tiêu hồi quy này được xác định bởi công thức (4)

Căn bậc hai được sử dụng để làm chậm quá trình phân rã của trung tâm. Giá trị của trung tâm dao động từ 0 đến 1 khi vị trí tiến lại càng gần tâm của đối tượng và ngược lại và được huấn luyện qua hàm BCE (Binary cross-entropy loss). Trong suốt quá trình thử nghiệm, giá trị của trung tâm được mạng dự đoán nhân với điểm phân loại để điểm cuối cùng trước khi xếp hạng các hộp giới hạn được suy ra, từ đó làm giảm đáng kể số lượng các hộp giới hạn chất lượng thấp được dự đoán nằm ở các vị trí nằm xa tâm vật thể.



Hình 8: Center-ness và bài toán tìm tâm vật thể [3].

3.2.3. Khối hồi quy (Regression)

Theo như [3], sau khi biết pixel có nằm trên vùng thuộc đối tượng, nhánh hồi quy sẽ có nhiệm vụ tính vị trí ngay chính pixel đó đến 4 cạnh của đối tượng như Hình 11 tạo thành 4 khoảng cách (l , t , b , r là khoảng cách từ pixel đến cạnh trên, dưới, trái, phải của hộp giới hạn) và từ 4 khoảng cách ấy sẽ tạo nên hộp vuông cuối cùng.

Hộp ground-truth cho hình ảnh đầu vào được xác định là $\{B_i\}$ có tọa độ $B_i=(x_0^{(i)}, y_0^{(i)}, x_1^{(i)}, y_1^{(i)}, c^{(i)})$ với:

- $x_0^{(i)}, y_0^{(i)}, x_1^{(i)}, y_1^{(i)}$: Tọa độ góc trái phía trên và góc phải phía dưới của hộp giới hạn.
- $c^{(i)}$: Lớp đối tượng.

Nếu vị trí (x, y) được liên kết với 1 hộp giới hạn, các mục tiêu hồi quy đào tạo cho vị trí có thể được tính bằng công thức dưới đây:

$$\begin{aligned}l^* &= x - x_0^{(i)}, t^* = y - y_0^{(i)} \\ r^* &= x_1^{(i)} - x, b^* = y_1^{(i)} - y\end{aligned}\tag{4}$$

Khác với các mô hình phát hiện dựa vào neo, bài báo này sử dụng hồi quy trực tiếp hộp giới hạn mục tiêu tại vị trí. Nói cách khác, nhóm tác giả xem những vị trí như là những mẫu thử huấn luyện thay vì hộp neo trong bộ phát hiện dựa vào neo.



Hình 9: FCOS hoạt động bằng cách dự đoán 1 vector 4D (l^*, t^*, r^*, b^*) [3].

4. KẾT QUẢ THỰC NGHIỆM

Phương pháp đề xuất được thực hiện trên máy tính có cấu hình như sau: CPU Intel Core I7-11375 GPU Nvidia RTX 3060, DDR4 16GB. Ngôn ngữ lập trình được sử dụng là Python với thư viện Pytorch, mmdetection cho các mô hình học sâu và thư viện OpenCv cho các thuật toán xử lý ảnh đơn giản. Chi tiết về quá trình thực nghiệm mô hình được trình bày chi tiết ở các phần sau đây.

4.1. Dataset

Để thực hiện hệ thống phát hiện người đeo/không đeo khẩu trang, bộ dữ liệu Face Mask Detection [5] được sử dụng. Bộ dữ liệu gồm có 853 ảnh với độ phân giải đa dạng được thu thập từ thực tế. Các ảnh trong bộ dữ liệu được chia làm 3 lớp: lớp có đeo khẩu trang, lớp không đeo khẩu trang và lớp đeo khẩu trang không phù hợp. Hình 10 giới thiệu một vài ảnh tượng trưng ứng với mỗi lớp. Mỗi ảnh được dán nhãn và kèm theo trong tập tin dạng xml tương ứng. Các nhãn được gán gồm có thông tin về vị trí hộp vuông của đối tượng và nhãn lớp tương ứng với hộp vuông. Trong bài báo này, để đơn giản nhóm tác giả chỉ sử dụng 2 lớp: lớp có đeo khẩu trang và lớp không đeo khẩu trang. Tất cả các ảnh trong lớp đeo khẩu trang không phù hợp sẽ thuộc về lớp không đeo khẩu trang.



Hình 10: Ảnh từ bộ dữ liệu Face Mask Detection [5].

4.2. Kết quả huấn luyện mô hình

Mô hình FCOS cho bộ dữ liệu đeo/không đeo khẩu trang được huấn luyện với số epochs = 20. Tốc độ học được chọn mặc định như mô hình gốc [1]. Các thông số huấn luyện khác giữ nguyên như mô hình gốc. Kết quả huấn luyện sau 20 epochs được thể hiện như Bảng 2 (có đính kèm trong file train_log.txt). Có thể thấy sau 20 epochs, mô hình có AP ứng với IoU 0.5 là 71.5%.

Bảng 2: Kết quả huấn luyện mô hình sau 20 epochs.

Average Precision (AP)	[IoU= 0.50 : 0.95	area = all	maxDets= 100] = 0.420
Average Precision (AP)	[IoU= 0.50	area = all	maxDets= 100] = 0.715
Average Precision (AP)	[IoU= 0.75	area = all	maxDets= 100] = 0.477
Average Precision (AP)	[IoU= 0.50 : 0.95	area = small	maxDets= 100] = 0.377
Average Precision (AP)	[IoU= 0.50 : 0.95	area = medium	maxDets= 100] = 0.525
Average Precision (AP)	[IoU= 0.50 : 0.95	area = large	maxDets= 100] = 0.808
Average Recall (AR)	[IoU= 0.50 : 0.95	area = all	maxDets= 1] = 0.201
Average Recall (AR)	[IoU= 0.50 : 0.95	area = all	maxDets= 10] = 0.465
Average Recall (AR)	[IoU= 0.50 : 0.95	area = all	maxDets= 100] = 0.503
Average Recall (AR)	[IoU= 0.50 : 0.95	area = small	maxDets= 100] = 0.446
Average Recall (AR)	[IoU= 0.50 : 0.95	area = medium	maxDets= 100] = 0.605
Average Recall (AR)	[IoU= 0.50 : 0.95	area = large	maxDets= 100] = 0.837

Trong đó:

Average Precision (AP):

AP: Độ chính xác trung bình tại IoU=.50: .05 : .95.

AP^{IoU=.50}: Độ chính xác trung bình tại IoU=0.5.

AP^{IoU=.75}: Độ chính xác trung bình tại IoU=0.75.

AP Across Scales:

AP^{small}: Độ chính xác trung bình cho đối tượng nhỏ, với kích thước đối tượng < 32x32 (height x weight) theo kích cỡ pixel.

AP^{medium}: Độ chính xác trung bình cho đối tượng trung bình, với 32x32 < kích thước đối tượng < 96x96.

AP^{large}: Độ chính xác trung bình cho đối tượng lớn, với kích thước đối tượng > 96x96.

Average Recall (AR):

AR^{max=1} : Độ nhạy trung bình với 1 phát hiện trên ảnh.

AR^{max=10} : Độ nhạy trung bình với 10 phát hiện trên ảnh.

AR^{max=100}: Độ nhạy trung bình với 100 phát hiện trên ảnh.

AR Across Scales:

AR^{small}: Độ nhạy trung bình cho đối tượng nhỏ, với kích thước đối tượng < 32x32 (height x weight) theo kích cỡ pixel.

AR^{medium}: Độ nhạy trung bình cho đối tượng trung bình, với 32x32 < kích thước đối tượng < 96x96.

AR^{large}: Độ nhạy trung bình cho đối tượng lớn, với kích thước đối tượng > 96x96.

4.3. Kết quả thực nghiệm mô hình

Mô hình sau khi huấn luyện được đánh giá trên tập dữ liệu testing của bộ dữ liệu. Tập dữ liệu kiểm tra gồm có 213 ảnh được lấy từ 853 ảnh gốc. Trong 213 ảnh kiểm tra có tổng cộng 412 khuôn mặt với 46 khuôn mặt không đeo khẩu trang và 366 khuôn mặt có đeo khẩu trang (bao gồm cả trường hợp đeo khẩu trang không đúng tiêu chuẩn). Kết quả kiểm tra được thống kê như Bảng 3. Mỗi kết quả hộp vuông phát hiện khuôn mặt đeo/không đeo khẩu trang trên tập kiểm tra được so sánh với ground-truth có sẵn trong bộ dữ liệu. Giá trị IoU (Intersection over Union) được tính toán dựa trên hộp vuông được phát hiện dựa trên dự đoán và hộp vuông ground-truth. Nếu IoU của hộp vuông ≥ 0.5 và cùng nhãn với nhãn của hộp vuông ground-truth thì hộp vuông đó được xem là được phát hiện và phân loại chính xác. Với hệ thống phát hiện

đeo/không đeo khẩu trang thì hai kết quả được quan tâm đó là True Negative rate và False Negative rate. True Negative rate sẽ cho biết hệ thống phát hiện chính xác bao nhiêu người không đeo khẩu trang, từ đó giúp người vận hành mô hình có thể có biện pháp ứng phó kịp thời với tình huống. Trong khi đó, False Negative rate sẽ cho biết hệ thống phát hiện sai bao nhiêu trường hợp có đeo khẩu trang (có đeo khẩu trang nhưng hệ thống báo là không đeo khẩu trang). Điều này sẽ giúp người vận hành mô hình đánh giá được mức độ tin cậy của mô hình để đưa ra ứng phó phù hợp. Dựa vào kết quả Bảng 3, hệ thống phát hiện được 42/46 người không đeo khẩu trang với tỉ lệ là 91.3%. Và hệ thống phát hiện sai 38/366 trường hợp có đeo khẩu trang với tỉ lệ là 10,4%.

Bảng 3: Kết quả thử nghiệm mô hình trên tập dữ liệu kiểm tra.

	Số người thực tế có trong tập dữ liệu kiểm tra	Số người phát hiện được và phân lớp chính xác	Số người phát hiện được và phân lớp không chính xác	Số người không phát hiện được
Có đeo khẩu trang	366	306	38	22
Không đeo khẩu trang	46	42	3	1

4.4. Thiết kế giao diện

Để có thể áp dụng mô hình đã huấn luyện vào sử dụng trong thực tế, nhóm tác giả đề xuất xây dựng giao diện để giúp người sử dụng có thể thao tác dễ dàng. Giao diện được xây dựng như Hình 11. Để xây dựng giao diện tích hợp mô hình, thư viện Tkinter được sử dụng. Các thông tin chính trên giao diện được thiết kế bao gồm:

Phần hiển thị thông tin mô hình: gồm tên đề tài; logo.

Phần hiển thị ảnh và kết quả phát hiện: gồm phần hiển thị ảnh gốc (ảnh đầu vào của mô hình FCOS); hiển thị ảnh kết quả của mô hình với người đeo/không đeo khẩu trang được dự đoán trên ảnh

Phần thông tin: Bao gồm tên nhóm tác giả, ngày và giờ hiện tại.

Phần thông tin xử lý: Hiển thị chế độ đang thực hiện hiện tại, khung ảnh đang được xử lý (với video và camera); hiển thị thời gian xử lý trên ảnh/khung ảnh; hiển thị kết quả phát hiện gồm số người có khẩu trang phát hiện được và số người không có khẩu trang phát hiện được.

Phần nút nhấn: Gồm nút nhấn để chọn ảnh, nút nhấn để chọn video, nút nhấn chạy mô hình dùng để chạy mô hình trên từng khung ảnh (do tốc độ xử lý trên video rất nhanh nên nhóm tác giả sử dụng nút nhấn chạy mô hình trên từng khung ảnh để dễ dàng quan sát kết quả), nút chạy camera để lấy ảnh thu được từ camera cho hệ thống xử lý, nút lưu file dùng để lưu lại các thông tin trước đó đã được hiển thị trên giao diện.



Hình 11: Giao diện chính của hệ thống.



Hình 12: Giao diện chọn ảnh/video từ nơi bất kỳ trên máy

4.5. Kết quả chạy hệ thống thực tế

Nhóm tác giả tiến hành chạy thực tế hệ thống với một số ảnh và video thu thập từ Internet cùng với hình ảnh thu được từ camera thực tế. Kết quả chạy hệ thống trên ảnh thể hiện như Hình 13, trên video như Hình 14, hình ảnh thu được trên camera thực tế ở Hình 15, hình ảnh thu được trên camera trong điều kiện môi trường thiếu ánh sáng như Hình 16 và Hình 17 thể hiện kết quả chạy hệ thống với video và bản lưu dữ liệu. Như hiển thị trong Hình 13, hệ thống phát hiện chính xác số người và vị trí người đeo khẩu trang/không đeo khẩu trang trên ảnh. Với video, hệ thống có thể phát hiện chính xác số người và vị trí người đeo khẩu trang/không đeo khẩu trang trên phần lớn khung hình. Còn với hình ảnh thu được trên camera thực tế hệ thống phát hiện được chính xác số lượng người đeo/không đeo khẩu trang trùng khớp với số lượng người mà camera ghi nhận lại được.



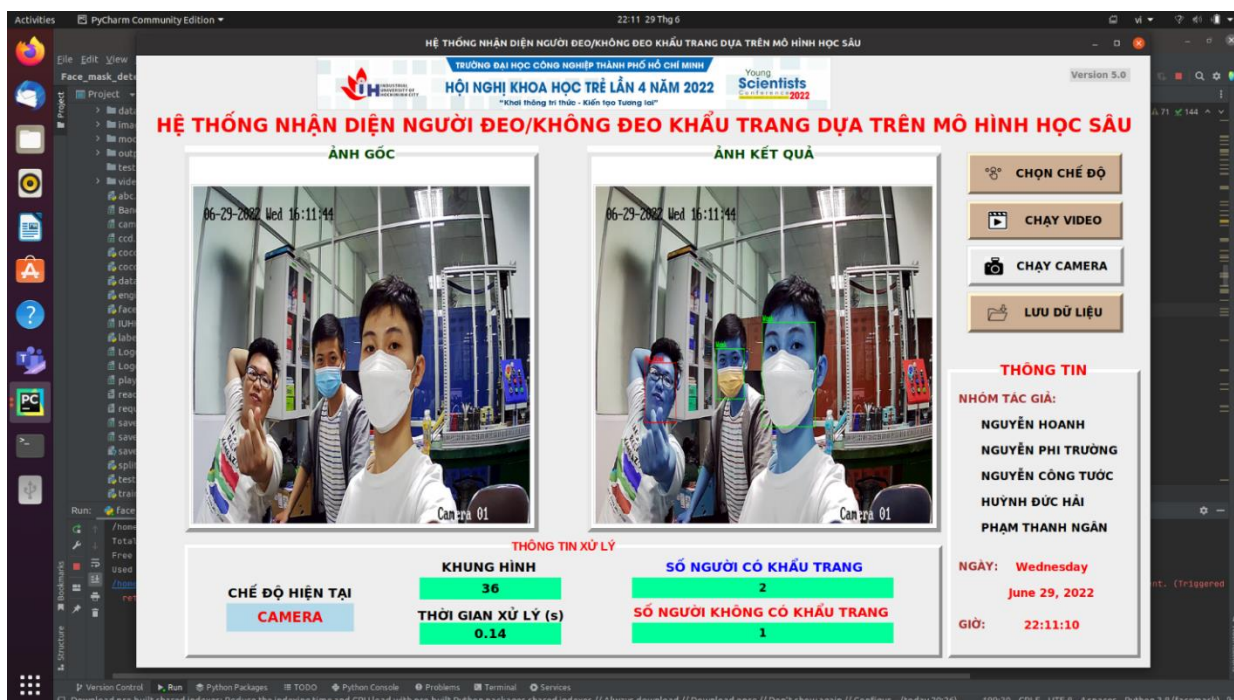
Hình 13: Kết quả chạy hệ thống trên ảnh bất kì đã chọn.

Dựa vào hình 13, khi chọn ảnh có chứa người đeo/không đeo khẩu trang trong thư mục chứa ảnh có sẵn trên máy tính, hệ thống cho ra kết quả chính xác số người đeo và không đeo khẩu trang, nhận diện đúng các đối tượng có trên ảnh với thời gian xử lý nhanh (0,2 giây).



Hình 14: Kết quả chạy hệ thống trên ảnh thu thập từ video.

Hình 14 khi chọn video có chứa đối tượng đeo/không đeo khẩu trang trong thư mục chứa video có sẵn trên máy tính, khi nhấn nút chạy video, hệ thống cho ra kết quả chính xác với tốc độ xử lý nhanh (0,11 giây) và nhận diện đúng các đối tượng trên mỗi khung hình của video.



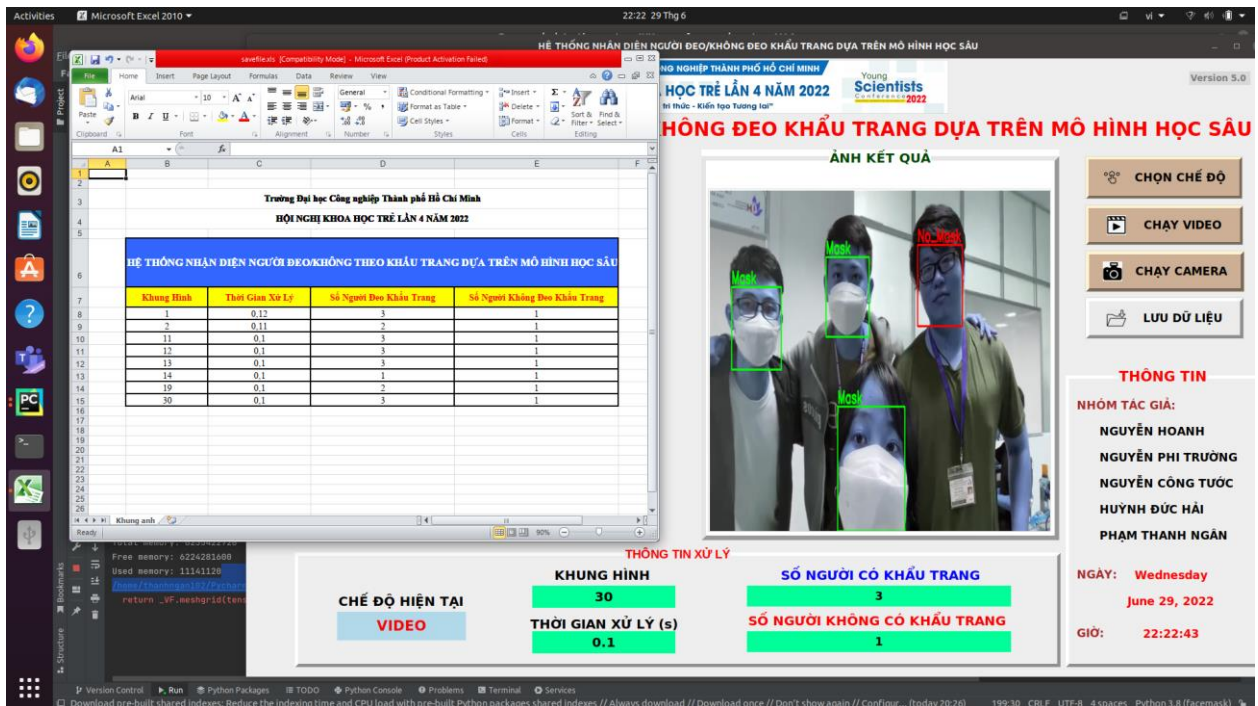
Hình 15: Kết quả chạy hệ thống trên hình ảnh thu thập từ camera thực tế.

Hình 15 khi chọn chế độ chạy camera, hệ thống sẽ lấy khung hình thông qua camera ngay tại thời điểm nhấn nút chạy camera để phân tích, xử lý, phát hiện khuôn mặt đeo và không đeo khẩu trang và hiển thị kết

quả trên màn hình giao diện. Khi nhấn chạy camera lần 2, hệ thống sẽ tiếp tục lấy khung hình thông qua camera tại thời điểm khi nhấn lần 2 để tiếp tục xử lí, phân tích như trên.



Hình 16: Kết quả chạy hệ thống trong điều kiện thiếu ánh sáng. Hình 16 nhóm tác giả vẫn chạy hệ thống thông qua camera như trường hợp ở Hình 15 nhưng trong điều kiện môi trường thiếu ánh sáng và cho ra kết quả nhận diện chính xác số người đeo và không đeo khẩu trang.



Hình 17: Kết quả chạy hệ thống với video và bản lưu dữ liệu.

Như hình trên, nhóm tác giả đã thiết kế nút lưu dữ liệu, cụ thể hơn, sau khi chạy mô hình và cho ra kết quả hiển thị trên màn hình giao diện, nhấn nút lưu dữ liệu và sau đó mở file excel đã được tạo ra sẵn trước đó, các thông tin như khung hình, thời gian xử lý, số người đeo và không đeo khẩu trang tại thời điểm nhấn nút lưu dữ liệu sẽ được hiển thị trên trang tính excel. Khi muốn cập nhật tiếp các kết quả tiếp theo, chỉ nhấn nút lưu dữ liệu và mở lại trang tính excel, các kết quả sẽ được cập nhật lại.

5. KẾT LUẬN

Trong bài báo này nhóm tác giả đề xuất xây dựng hệ thống phát hiện người đeo/không đeo khẩu trang dựa trên mô hình học sâu với dữ liệu thu được từ hình ảnh, video và camera. Hệ thống đề xuất được xây dựng dựa trên mô hình FCOS, là mô hình đã được phát triển với tính chất linh hoạt, đơn giản nhưng vẫn đảm bảo được độ chính xác cao đối với bài toán phát hiện đối tượng.

Hệ thống xây dựng được huấn luyện và kiểm tra trên tập dữ liệu Face Mask Detection [5]. Kết quả huấn luyện và kiểm tra trên bộ dữ liệu cho thấy hệ thống phát hiện người đeo/không đeo khẩu trang đạt được độ chính xác cũng như độ ổn định cao về vấn đề phát hiện cũng như phân loại. Bên cạnh đó, nhóm tác giả còn xây dựng một giao diện với đầy đủ các thành phần để tạo thành một phần mềm hoàn chỉnh, giúp người dùng có thể vận hành hệ thống một cách dễ dàng. Quá trình kiểm tra mô hình với các ảnh và video, camera thu được trên thực tế cho kết quả cao về độ chính xác của mô hình.

TÀI LIỆU THAM KHẢO

- [1] Shaoqing Ren, Kaiming He, Ross Girshick and Jian Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," *Computer Vision and Pattern Recognition*, 2016.
- [2] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott Reed, Cheng-Yang Fu, Alexander C. Berg, "SSD: Single Shot MultiBox Detector," *In Proceedings of the IEEE/CVF international conference on computer vision*, pp. 21-37, 2016.
- [3] Zhi Tian, Chunhua Shen, Hao Chen, and Tong He, "FCOS: Fully Convolutional One-Stage Object Detection," *Computer Vision and Pattern Recognition*, pp. 9627-9636, 2019.
- [4] Tsung-Yi Lin, Piotr Dollár, Ross Girshick, Kaiming He, Bharath Hariharan, Serge Belongie, "Feature Pyramid Networks for Object Detection," *In Proceedings of the IEEE/CVF international conference on computer vision*, pp. 2117-2125, 2017.
- [5] Larxel, "Kaggle," Face Mask Detection, 2020. [Online]. Available: <https://www.kaggle.com/datasets/andrewmvd/face-mask-detection/>.
- [6] Toon Van Craenendonck, Baturay Ofluoglu, "Github," Tutorial Face Mask Detection, 2020. [Online]. Available: <https://github.com/datarootsio/tutorial-face-mask-detection/>.
- [7] Jiankang Deng, Jia Guo, Yuxiang Zhou, Jinke Yu, Irene Kotsia and Stefanos Zafeiriou, "Retinaface: Single-stage dense face localisation in the wild," *Computer Vision and Pattern Recognition*, 2019.
- [8] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, "Mobilenets: Efficient convolutional neural networks for mobile vision applications," *Computer Vision and Pattern Recognition*, 2017.
- [9] Jingwei Liu, Yi Gu, Shumin Han et al, "Feature Rescaling and Fusion for Tiny Object Detection," *IEEE Access*, pp. 62946-62955, 2021.

- [10] Zhihao Cao, Mingfeng Shao, Li Xu, Shaomin Mu, Hongchun Qu, "MaskHunter: real- time object detection of face masks during the COVID-19 pandemic," *IET Image Processing*, pp. 4359-4367, 2020.
- [11] A. Sachan, "Learn Machine Learning, AI & Computer vision," 2020. [Online]. Available: <https://cv-tricks.com/keras/understand-implement-resnets>.
- [12] Satnam Singh, Doris Schicker, "Seven Basic Expression Recognition Using ResNet-18," *Computer Vision and Pattern Recognition*, 2021.
- [13] Li Zhang, Qing Wang, Lei Xie, "Duality temporal-Channel-Frequency attention enhanced speaker representation learning," *Computer Vision and Pattern Recognition*, 2021.
- [14] Ross Wightman, Hugo Touvron, Hervé Jégou, "ResNet strikes back: An improved training procedure in timm," *Computer Vision and Pattern Recognition*, 2021.
- [15] J. Hui, "Understanding Feature Pyramid Networks for object detection (FPN)," 2018. [Online]. Available: <https://jonathan-hui.medium.com/understanding-feature-pyramid-networks-for-object-detection-fpn-45b227b9106c>.
- [16] Pdraig Cunningham, Sarah Jane Delany, "k-Nearest Neighbour Classifiers: 2nd Edition (with Python examples)," *Computer Vision and Pattern Recognition*, 2020.
- [17] Keiron O'Shea, Ryan Nash, "An Introduction to Convolutional Neural Networks," *Computer Vision and Pattern Recognition*, 2015.

YSC4F.206

HỆ THỐNG NHẬN DIỆN VÀ ĐO KHOẢNG CÁCH GIỮA NGƯỜI VỚI NGƯỜI SỬ DỤNG HỌC SÂU

NGUYỄN HOANH, ĐỖ NGUYỄN HOÀNG LONG, HUỲNH TRẦN MINH THUẬN,
TRẦN NGUYỄN BÁ THỌ, BÙI NGỌC MINH PHONG

Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

*nguyenhoanh@iuh.edu.vn, 18040631.long@student.iuh.edu.vn, 18029841.thuan@student.iuh.edu.vn,
18033591.tho@student.iuh.edu.vn, 18031721.phong@student.iuh.edu.vn*

Tóm tắt. Trong thời điểm dịch Covid-19, việc giữ khoảng cách giữa mọi người trong phòng có vai trò quan trọng góp phần làm giảm đi sự lây lan của dịch bệnh, giúp bảo vệ sức khỏe cộng đồng. Trong bài báo này, chúng tôi phát triển hệ thống dựa trên mô hình Faster-RCNN, là mô hình học sâu phát hiện đối tượng là người trên hình ảnh thu được, kết hợp với các thuật toán xử lý ảnh để tính toán số người và khoảng cách giữa người. Nhằm cải thiện độ chính xác và tốc độ xử lý, chúng tôi sử dụng mạng nơ-ron tích chập ResNet kết hợp mạng thu phóng EfficientNet thay cho mạng VGG16 trong mô hình Faster-RCNN ban đầu. Quá trình huấn luyện và kiểm tra trên bộ dữ liệu cho thấy hệ thống có thể phát hiện người đạt độ chính xác với tỉ lệ là 90%, thuật toán đo khoảng cách có sai số so với thực tế thấp khoảng ± 0.3 (m). Ngoài ra, chúng tôi còn xây dựng một giao diện trực quan thân thiện với người sử dụng, cho phép hiển thị và lưu trữ các trường hợp không đúng khoảng cách quy định để tạo thành một phần mềm hoàn chỉnh giúp người dùng nắm bắt thông tin một cách chính xác. Với nghiên cứu này, nhóm chúng tôi hi vọng hệ thống sẽ góp phần bảo vệ sức khỏe cộng đồng thông qua việc phát hiện và giữ đúng khoảng cách quy định.

Từ khóa. COVID-19, Faster-RCNN, EfficientNet, học sâu, tính khoảng cách, tính số người.

COUNTING PEOPLE AND MEASURING SPACE DISTANCES SYSTEM BASEDON DEEP LEARNING

Abstract. During the Covid-19 epidemic, keeping a distance between people in the room plays an important role in reducing the spread of the disease, helping to protect public health. In this paper, we develop a system based on the Faster-RCNN model, which is a deep learning model that detects human objects on the acquired images, combined with image processing algorithms to calculate the number of people. and distance between people. In order to improve the accuracy and processing speed, we use ResNet convolutional neural network combined with EfficientNet zoom network instead of VGG16 network in the original Faster-RCNN model. The training and testing process on the data set shows that the system can detect people with an accuracy of 90%, the distance measurement algorithm has an error as low as ± 0.3 (m). Besides, We also build a user-friendly intuitive interface that allows displaying and storing cases that are not at the specified distance to form a complete software that helps users capture information in a timely manner. exactly. With this study, our team hopes that the system will contribute to protecting public health by detecting and keeping the prescribed distance.

Keywords. COVID-19, Faster-RCNN, EfficientNet, Deep learning, Measuring space distances, Counting people.

1. GIỚI THIỆU

1.1. Lí do nghiên cứu

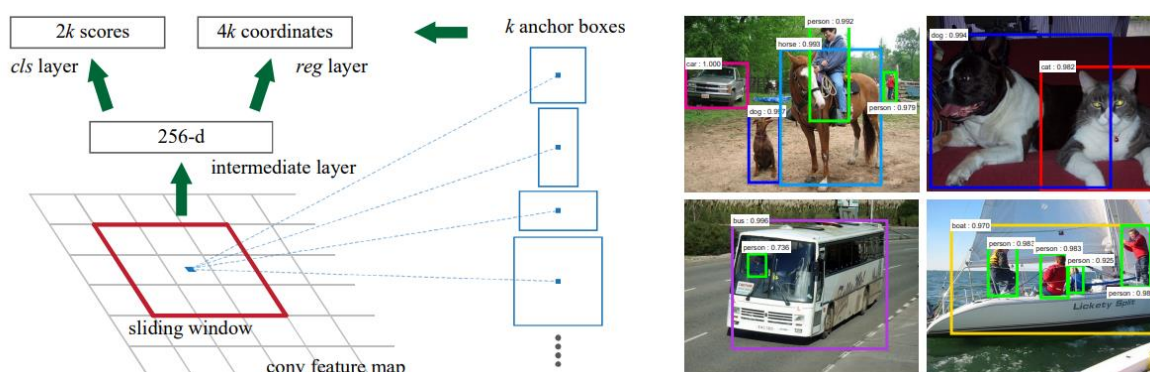
Trong thời điểm dịch COVID-19, việc giữ khoảng cách giữa mọi người trong phòng có vai trò quan trọng góp phần làm giảm đi sự lây lan của dịch bệnh, giúp bảo vệ sức khỏe cộng đồng. Để đảm bảo được các tiêu chí an toàn trong đại dịch, các quốc gia thường quy định việc giữ khoảng cách giữa mọi người trong phòng. Tuy nhiên, vẫn tồn tại các trường hợp không giữ khoảng cách đúng quy định. Việc phát hiện sớm các trường hợp không giữ khoảng cách có thể góp phần không nhỏ vào công tác ngăn ngừa sự lây lan dịch bệnh, giúp bảo vệ sức khỏe cộng đồng. Các phương pháp phát hiện người và đo khoảng cách giữa mọi người nói riêng và phát hiện đối tượng nói chung dựa trên hình ảnh đã được các nhà khoa học phát triển trong thời gian dài. Các phương pháp này có thể chia làm 2 hướng chính: Các phương pháp truyền thống và các phương pháp dựa trên học sâu. Các phương pháp truyền thống sử dụng các đặc trưng cơ bản trên ảnh như màu sắc, hình dạng, kích thước,... cùng với một bộ phân lớp đơn giản như phân lớp tuyến tính - Support Vector Machine (SVM) để phát hiện các đối tượng trên ảnh. Các phương pháp này có một nhược điểm lớn là độ chính xác không cao và dễ bị ảnh hưởng của nhiễu như nhiễu ánh sáng, nhiễu đồ bóng,... Độ chính xác thấp của các phương pháp truyền thống nằm ở sự đa dạng của đối tượng cần phát hiện trong thực tế. Cụ thể hơn, các thuật toán truyền thống thường được phát triển để phát hiện đối tượng trong một số môi trường cụ thể. Với sự đa dạng về môi trường, kích thước, tỉ lệ,... của đối tượng trên ảnh, các thuật toán này thường không chính xác trong các trường hợp thực tế khác biệt.

Với sự phát triển mạnh mẽ của học sâu trong những năm gần đây, các mô hình phát hiện đối tượng dựa trên học sâu được phát triển mạnh mẽ và cho thấy được độ chính xác vượt trội so với các phương pháp truyền thống. Hơn nữa, các mô hình phát hiện đối tượng dựa trên học sâu còn cho thấy được sự hiệu quả với các ứng dụng thực tế khi vẫn cho độ chính xác cao trong nhiều môi trường khác nhau và nhiều loại đối tượng khác nhau về kích thước, tỉ lệ, màu sắc,... Các mô hình phát hiện đối tượng dựa trên học sâu được phát triển để phát hiện nhiều loại đối tượng khác nhau như xe cộ, khuôn mặt, biển số xe, biển báo giao thông,... Các mô hình này thường bao gồm một mạng nơ-ron tích chập ở ngõ vào để trích xuất các đặc trưng trên ảnh và một mạng để phát hiện đối tượng ở cuối cùng để phát hiện đối tượng dựa trên các đặc trưng đã trích xuất trước đó. Một số mô hình phát hiện đối tượng phổ biến như mô hình Faster R-CNN (Regions with CNN features) [1], Single Shot MultiBox Detector (SSD) [2], Fully Convolutional One-stage Object Detection (FCOS), Feature Pyramid Network (FPN),...

1.2. Các công trình liên quan

Trong những năm gần đây, học sâu càng ngày càng được phát triển và ứng dụng nhiều hơn trong cuộc sống. Một số các nghiên cứu như [22],[23] và [24] đã chứng tỏ điều này. Vì vậy, nhóm chúng tôi sử dụng học sâu để giải quyết bài toán nhận diện và tính toán khoảng cách giữa người với người. Hệ thống nhận diện người và đo khoảng cách giữa mọi người là một hệ thống phức tạp, yêu cầu giải quyết cùng lúc ba bài toán. Thứ nhất, hệ thống phải giải quyết bài toán phân lớp, phân biệt được người. Thứ hai, hệ thống phải giải quyết bài toán phát hiện, định vị chính xác vị trí của người trên ảnh. Thứ ba, hệ thống phải giải quyết bài toán đo khoảng cách giữa người. Hướng tiếp cận để giải quyết bài toán là sử dụng một mô hình duy nhất cho cả hai bài toán nhận dạng và phát hiện. Nhóm tác giả trong [5] sử dụng mô hình Faster RCNN có cấu trúc như hình 1 để nhận diện đối tượng là người ở các tư thế khác nhau và phát hiện chính xác vị trí người. Với hướng tiếp cận này, mô hình có thể được xây dựng dựa trên các thuật toán xử lý ảnh truyền thống hoặc dựa trên mô hình học sâu. Các mô hình dựa trên học sâu [3],[4] luôn cho độ chính xác cao hơn so với các mô hình sử dụng xử lý ảnh truyền thống. Với việc sử dụng mạng nơ-ron tích chập ResNet [20] kết hợp mạng thu phóng EfficientNet [11] thay cho mạng VGG16 [8], mô hình đã cải thiện tốc độ xử lý và tỉ lệ chính xác. Trong [9], nhóm tác giả đã xây dựng mô hình phát hiện người theo thời gian thực dựa trên mô hình COCO để phát hiện nhiều người một cách chính xác trong môi trường tự nhiên. Mô hình được đề xuất bao gồm một số khối cải tiến để giúp hệ thống phát hiện người đặc biệt là trong điều kiện trời tối. Về vấn đề tính khoảng cách, các bài nghiên cứu trước [13] sử dụng các thiết bị ngoại vi như cảm biến hoặc camera phụ để tính khoảng cách, trong nghiên cứu này chúng tôi chỉ dùng một camera duy nhất để tính toán khoảng cách dựa trên nhóm tác giả [12]. Dựa vào thuật toán trong [14] dùng phương pháp “Lỗ kim của camera” (pinhole camera) và hệ trục tọa độ, chúng tôi đã tính một cách chính xác khoảng cách các đối tượng là người trong mô hình này. Dựa vào những nghiên cứu và tham khảo trên, chúng tôi xây dựng hệ thống phát hiện người và đo khoảng cách giữa người dựa trên học

sâu với tốc độ xử lý nhanh và kết quả chính xác. Chi tiết về hệ thống được xây dựng trong bài báo này được chúng tôi trình bày chi tiết trong các phần sau đây.

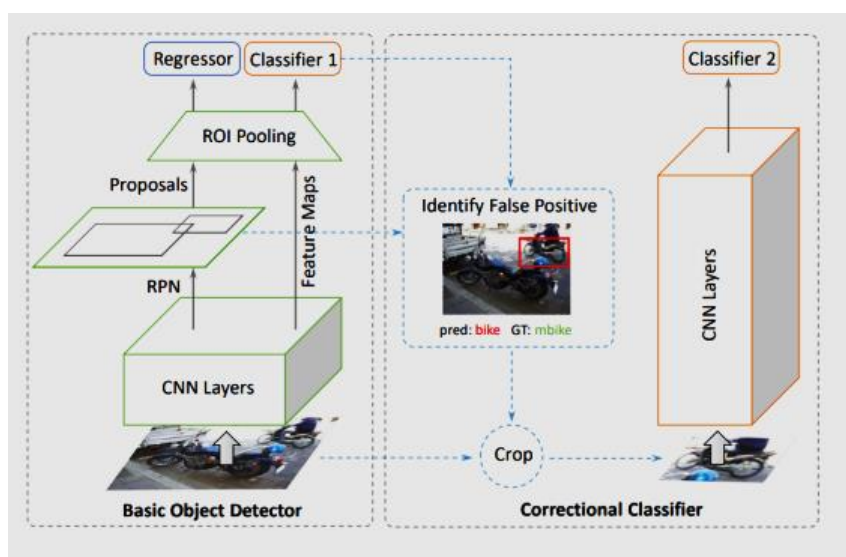


Hình 1. Mô hình được đề xuất trong [1]

1.3. Mô hình đề xuất

Nhóm tác giả sử dụng mô hình Faster R-CNN cho hệ thống phát hiện người. Mô hình Faster R-CNN là một mô hình đã được đề xuất vào năm 2016 [1] dùng để phát hiện đối tượng trong ảnh. Sự ra đời của mô hình Faster R-CNN đã tạo tiền đề cho sự phát triển mạnh mẽ của các mô hình phát hiện đối tượng dựa trên nền tảng học sâu [3], [4]. Nhiều mô hình đã và đang được xây dựng dựa trên nền tảng của mô hình Faster R-CNN để cải thiện khả năng phát hiện đối tượng cũng như ứng dụng để phát hiện các đối tượng trong các điều kiện khó khăn như vật thể nhỏ [10], vật thể có nhiều kích thước khác nhau [4],...

Hình 2 mô tả chi tiết cấu trúc của mô hình Faster R-CNN. Như mô tả trong hình 2, mô hình Faster R-CNN gồm có 4 khối chính: khối trích xuất đặc trưng ảnh; khối phát hiện đối tượng tiềm năng; khối tái tạo đặc trưng đối tượng; và khối phát hiện đối tượng. Đầu tiên, ảnh sẽ được trích xuất các đặc trưng thông qua một mạng nơ-ron tích chập. Các bản đồ đặc trưng ở tầng cuối cùng của khối trích xuất đặc trưng sẽ được sử dụng cho các khối tiếp theo. Khối phát hiện đối tượng tiềm năng là một mô hình phát hiện đối tượng có nhiệm vụ tạo ra các đối tượng tiềm năng (đối tượng có xác suất cao là đối tượng thật). Dựa vào tọa độ của các đối tượng tiềm năng và các đặc trưng của tầng cuối cùng của mạng trích xuất đặc trưng, khối tái tạo đặc trưng đối tượng sẽ trích xuất ra các đối tượng có kích thước cố định. Cuối cùng, dựa trên các đối tượng có kích thước cố định, khối phát hiện đối tượng sẽ cho ra kết quả. Từ đó, đo khoảng cách giữa người và cho kết quả cuối cùng. Chi tiết các khối của mô hình Faster R-CNN được trình bày sau đây. Ngoài các khối chính, một số phương pháp xử lý bên trong mỗi khối cũng sẽ được trình bày chi tiết.

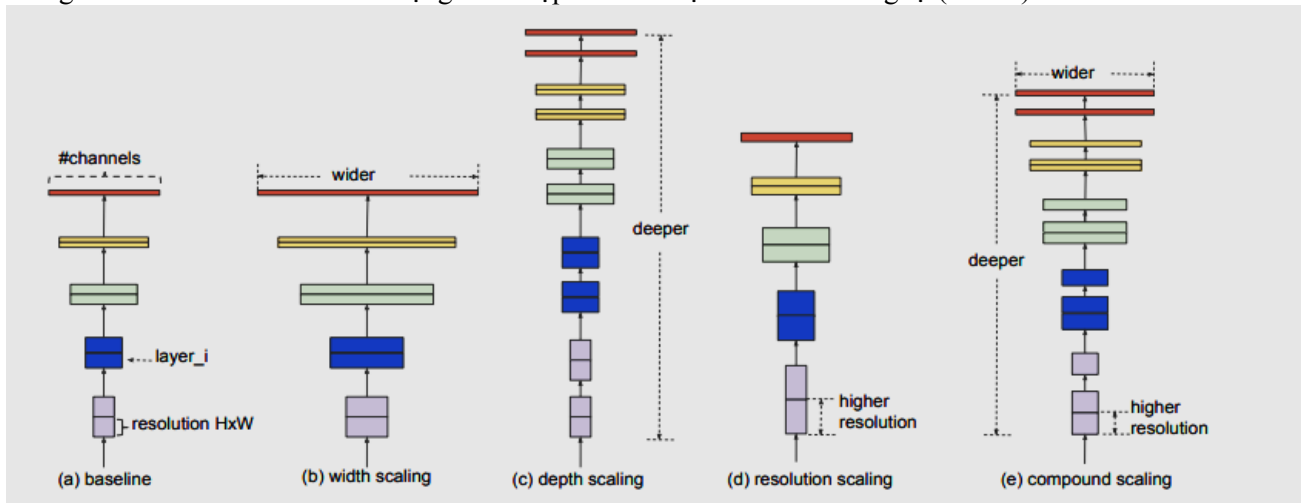


Hình 2. Mô hình Faster R-CNN trong [15]

2. MÔ HÌNH FASTER R-CNN

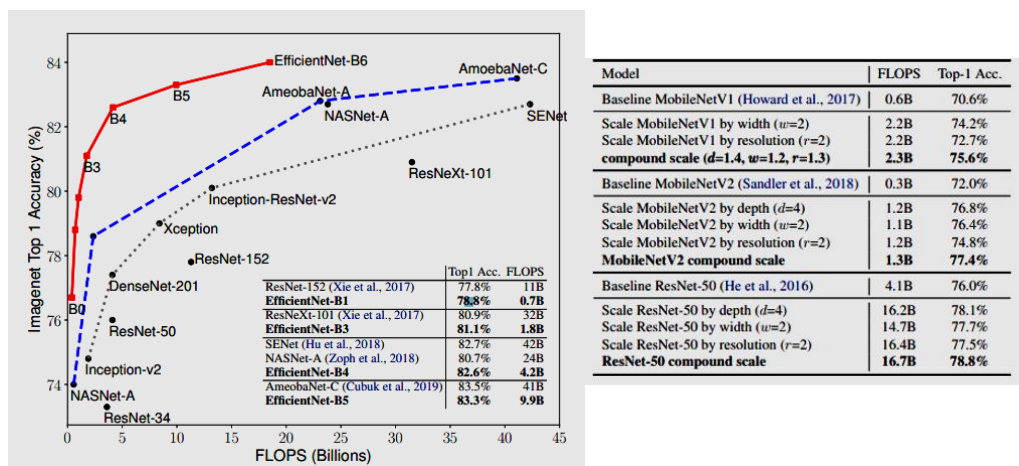
Mạng thu phóng EfficientNet

Để trích xuất các đặc trưng từ ảnh ngõ vào, mô hình Faster R-CNN sử dụng mạng nơ-ron tích chập ResNet50 kết hợp mạng thu phóng EfficientNet trong mô hình. Hình 3 mô tả cách hoạt động của mạng EfficientNet là một mạng nơ-ron phức hợp đạt về độ chính xác với số lượng tham số và FLOPS 1 (tỷ đơn vị) ít hơn nhưng đạt hiệu quả và nhanh hơn các kiến trúc mạng khác. Mạng thu phóng EfficientNets kết hợp sử dụng thu phóng dựa trên độ rộng, độ sâu và độ phân giải của ảnh để đạt được hiệu suất cao mà không ảnh hưởng đến hiệu quả tài nguyên. Trong bài [11], nhóm tác giả đã chứng minh độ chính xác của mạng ResNet50 sau khi sử dụng mạng thu phóng EfficientNet được cải thiện lên 78,8% và ít sử dụng các thông số FLOPS hơn so với các mạng tích chập khác có độ chính xác tương tự (hình 4).



Hình 3. Mạng thu phóng EfficientNet [16]

¹FLOPS (FLoating-point Operations per Second) là một thước đo hiệu suất máy tính.

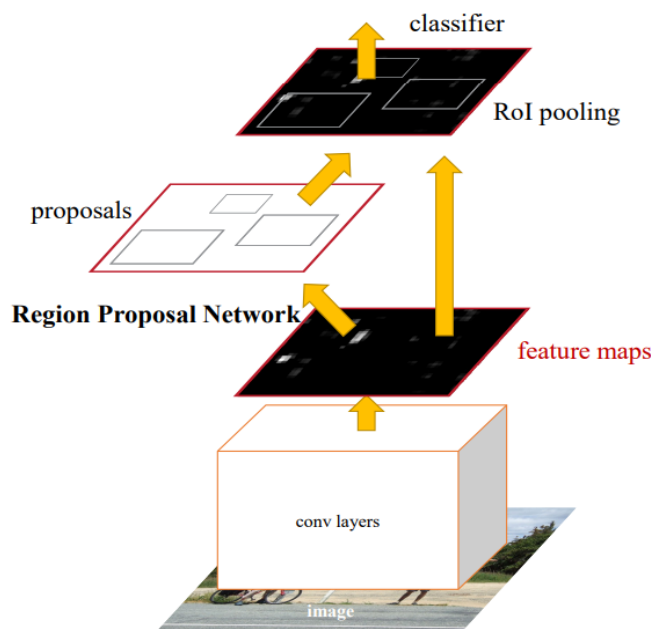


Hình 4. Kết quả quy mô thu phóng của mạng EfficientNet

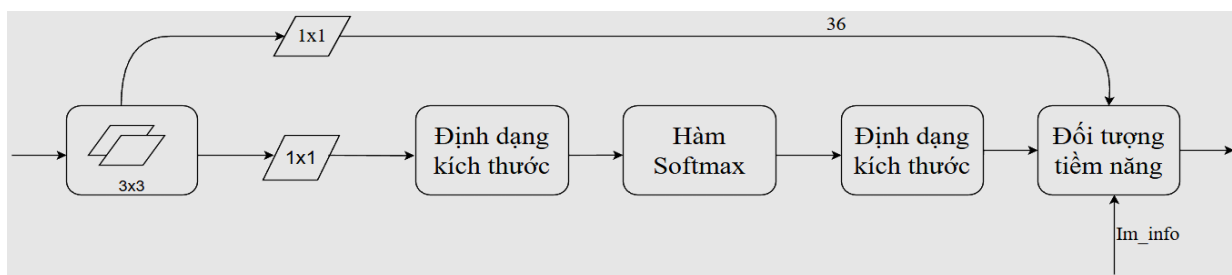
Khối phát hiện đối tượng tiềm năng

Hình 5 mô tả khối phát hiện đối tượng tiềm năng có nhiệm vụ tạo ra các đối tượng tiềm năng (đối tượng có khả năng cao là đối tượng thật) dựa trên các bản đồ đặc trưng ở tầng cuối cùng của khối trích xuất đặc trưng. Khối phát hiện đối tượng tiềm năng trong mô hình Faster R-CNN là một mạng học sâu có cấu trúc như hình 6. Khối gồm có 1 tầng tích chập với ma trận trọng số kích thước 3×3 , theo sau là 2 tầng tích chập

có kích thước ma trận trọng số là 1×1 . Lớp thứ nhất có ngõ ra là $2k$ giá trị với k là số lượng anchor ở mỗi vị trí của bản đồ đặc trưng cho biết kết quả phân lớp của anchor đó là thuộc về lớp đối tượng hay lớp nền. Lớp thứ 2 có ngõ ra là $4k$ giá trị $\{\Delta x, \Delta y, \Delta w, \Delta h\}$ cho biết các giá trị hồi quy dùng để cập nhật vị trí của anchor tương ứng trên ảnh gốc.



Hình 5. Khối phát hiện đối tượng tiềm năng [1]

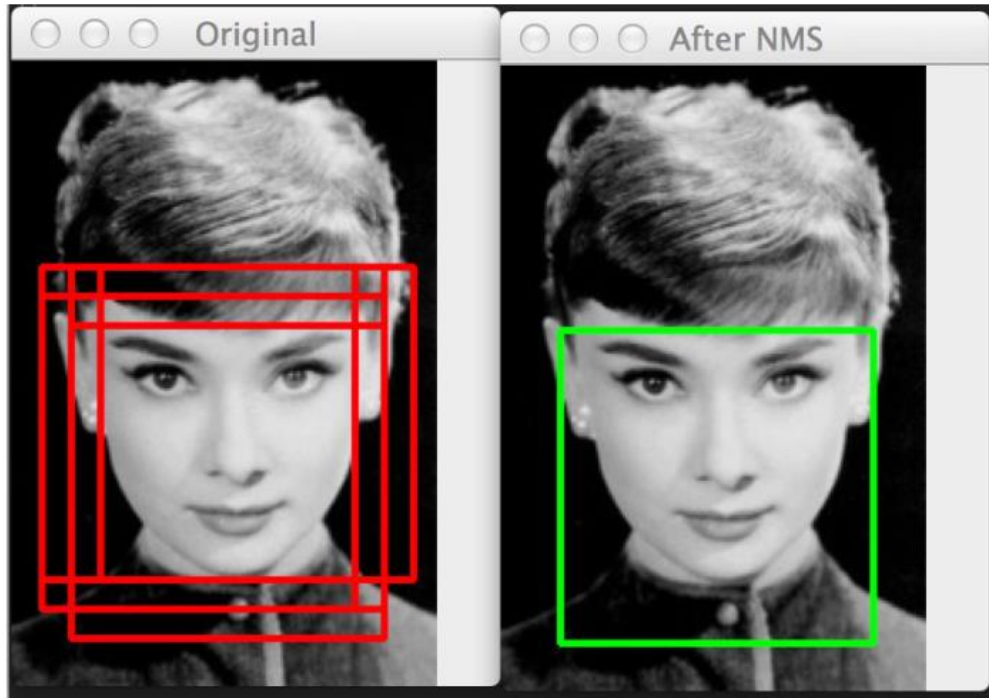


Hình 6. Cấu trúc khối phát hiện đối tượng tiềm năng trong [6]

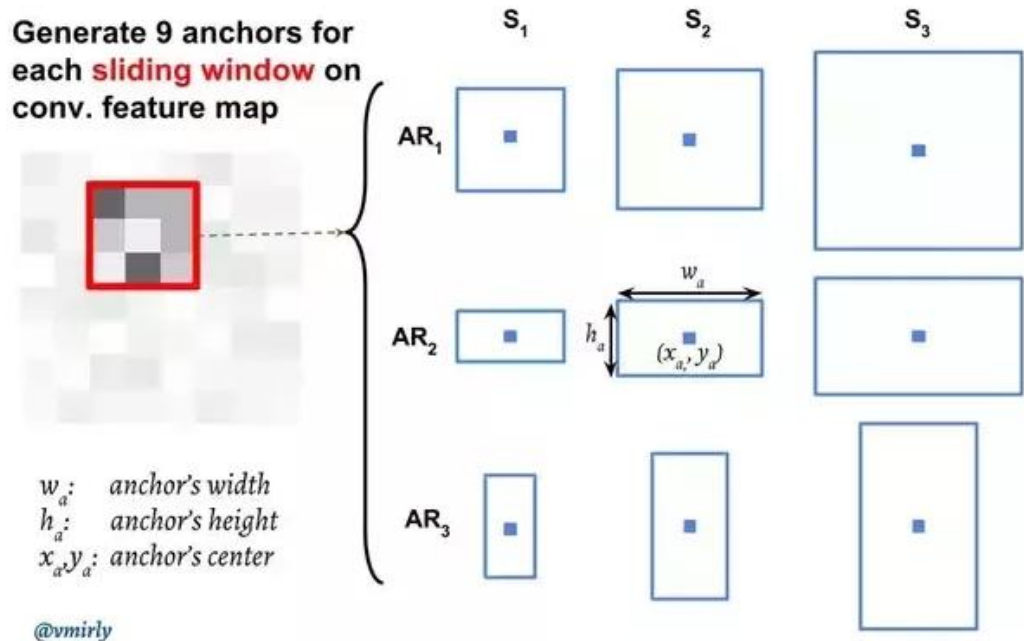
Anchor

Trong bài toán phát hiện đối tượng, ngoài việc phân lớp đối tượng, mô hình còn cần phải xác định vị trí của đối tượng trên ảnh. Vị trí của một đối tượng trên ảnh được định nghĩa bằng một hộp vuông có 4 thông số $\{x,y,h,w\}$ trong đó $\{x,y\}$ cho biết tọa độ của hộp vuông và $\{h,w\}$ cho biết chiều cao và chiều dài của hộp vuông đó (hình 7 và 8). Do đặc trưng của đối tượng trên ảnh là có kích thước và tỉ lệ đa dạng nên các hộp vuông thể hiện vị trí đối tượng này cũng có kích thước và tỉ lệ đa dạng. Để xác định các hộp vuông này, mô hình Faster R-CNN đưa ra khái niệm anchor, được xem là các hộp vuông tiền định nghĩa có kích thước và tỉ lệ nhất định tại một vị trí trên bản đồ đặc trưng. Để giải quyết vấn đề về sự đa dạng của tỉ lệ và kích thước của đối tượng, anchor trong mô hình Faster R-CNN cũng được tiền định nghĩa với tỉ lệ và kích thước khác nhau. Cụ thể, tại mỗi vị trí (pixel) của bản đồ đặc trưng ở tầng cuối cùng, mô hình Faster R-CNN định nghĩa 9 anchor với 3 kích thước $\{128 \times 128, 256 \times 256, 512 \times 512\}$ và 3 tỉ lệ $\{1:1, 1:2, 2:1\}$. Mỗi anchor sẽ có kích thước và tỉ lệ cố định trong tổ hợp 9 anchor này. Hình 8 thể hiện 9 anchor này tại 1 vị trí trên bản đồ đặc trưng. Trong quá trình huấn luyện mạng và thực thi mạng, các anchor này sẽ là ngõ vào của khối phát hiện đối tượng tiềm năng. Khối phát hiện đối tượng tiềm năng có nhiệm vụ là phân lớp các anchor này thành các anchor có khả năng là đối tượng và các các anchor là nền và thực hiện hồi quy tọa độ của các anchor có khả năng là đối tượng để mô tả tốt hơn vị trí của hộp vuông tương ứng. Kết thúc của khối phát

Hiện đối tượng tiềm năng sẽ là tọa độ của các hộp vuông có khả năng là đối tượng thật. Các tọa độ này sẽ được sử dụng bởi khối tái tạo đặc trưng đối tượng để tạo ra các đối tượng có kích thước cố định. Các đối tượng có kích thước cố định sẽ dùng làm ngõ vào của khối cuối cùng để tìm đối tượng thật trên ảnh.



Hình 7: Hộp vuông thể hiện vị trí đối tượng trên ảnh trong [17].

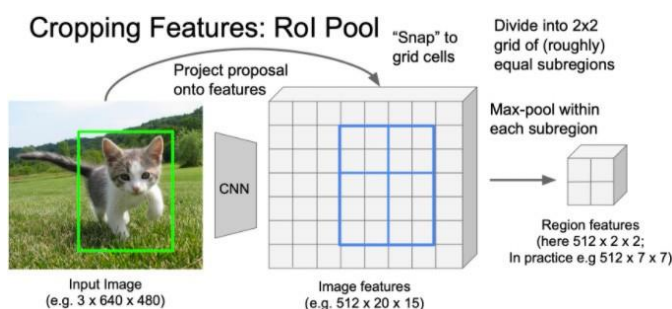


Hình 8. 9 Anchor tại mỗi vị trí của bản đồ đặc trưng [17]

Khối tái tạo đặc trưng đối tượng

Khối tái tạo đặc trưng đối tượng có nhiệm vụ là tạo ra các đối tượng tiềm năng có kích thước cố định (ví dụ, $7 \times 7 \times 256$) dựa vào bản đồ đặc trưng ở lớp cuối cùng của khối trích xuất đặc trưng và tọa độ của các đối tượng tiềm năng ở ngõ ra của khối phát hiện đối tượng tiềm năng. Mô hình Faster R-CNN đề xuất sử dụng

khối tái tạo đặc trưng đối tượng để từ đó tạo ra tất cả các đối tượng tiềm năng có cùng kích thước. Việc tạo ra các đối tượng tiềm năng cùng kích thước là quan trọng vì khi đó khối phát hiện đối tượng cuối cùng chỉ cần dùng chung một mô hình mạng để phát hiện đối tượng cho tất cả các đối tượng tiềm năng. Cơ chế làm việc của khối tái tạo đặc trưng đối tượng được minh họa như ví dụ hình 9. Đầu tiên, khối tái tạo đặc trưng đối tượng chia đối tượng tiềm năng bất kỳ có kích thước $h \times w$ thành đối tượng tiềm năng có kích thước cố định $H \times W$ với xấp xỉ $h/H \times w/W$. Sau đó, phương pháp trích xuất giá trị lớn nhất được sử dụng để trích xuất giá trị lớn nhất trên đối tượng tiềm năng cũ sang đối tượng tiềm năng mới. Kết quả của khối tái tạo đặc trưng đối tượng là n đối tượng tiềm năng có cùng kích thước.

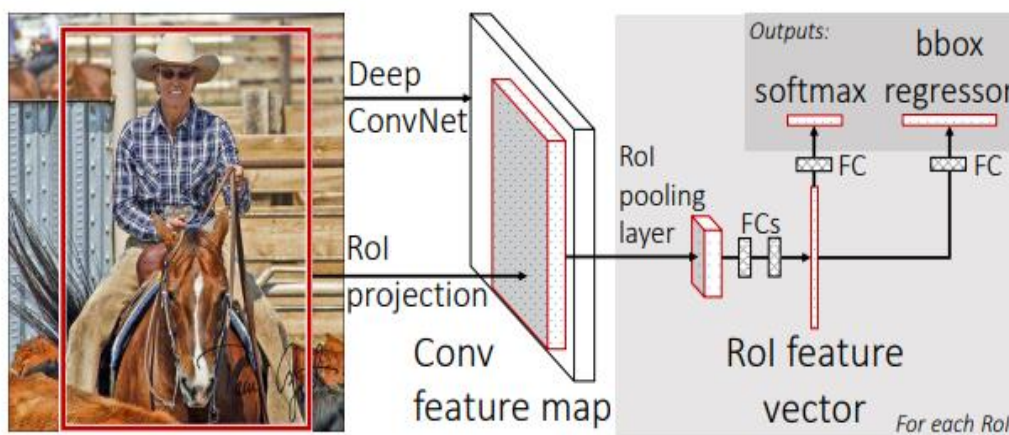


Hình 9. Cơ chế làm việc của khối tái tạo đặc trưng đối tượng.

(Nguồn: CS231n: Convolutional Neural Networks for Visual Recognition: Stanford - Spring 2021).

Khối phát hiện đối tượng

Khối cuối cùng trong mô hình Faster R-CNN là khối phát hiện đối tượng. Khối phát hiện đối tượng sẽ có ngõ vào là các đối tượng tiềm năng có cùng kích thước được tạo ra bởi khối tái tạo đặc trưng đối tượng. Hình 10 mô tả cấu trúc của khối phát hiện đối tượng được đề xuất trong mô hình Faster R-CNN. Như mô tả trong hình 10, trước tiên khối biến đổi các đối tượng tiềm năng ngõ vào thành các vector cột. Tiếp theo đó sẽ là 2 tầng kết nối fully connected (FC) với 4096 nơ-ron và hàm kích hoạt ReLU ở mỗi lớp được sử dụng để học các đặc trưng của đối tượng tiềm năng. Cuối cùng sẽ là 2 tầng kết nối: tầng thứ nhất có $N+1$ nơ-ron với N là số lớp của bộ dữ liệu, giá trị của mỗi nơ-ron ở tầng này chính là số điểm của mỗi lớp ứng với từng đối tượng tiềm năng; tầng thứ 2 có $4N$ nơ-ron biểu thị 4 giá trị hồi quy ứng với mỗi lớp để cập nhật lại tọa độ của hộp vuông biểu thị vị trí đối tượng trong ảnh. Với cấu trúc trên, khối phát hiện đối tượng sẽ phân lớp mỗi đối tượng tiềm năng về một lớp tương ứng, đồng thời cập nhật tọa độ hộp vuông của đối tượng tiềm năng. Như vậy, sau lớp cuối cùng ta sẽ có được tọa độ của các đối tượng thật trên ảnh và lớp tương ứng với đối tượng đó. Vì các đối tượng tiềm năng có thể trùng nhau trên cùng vị trí một vật thể, để xử lý việc các kết quả trùng nhau này thì thuật toán Non maximum suppression (NMS) [19] được sử dụng.

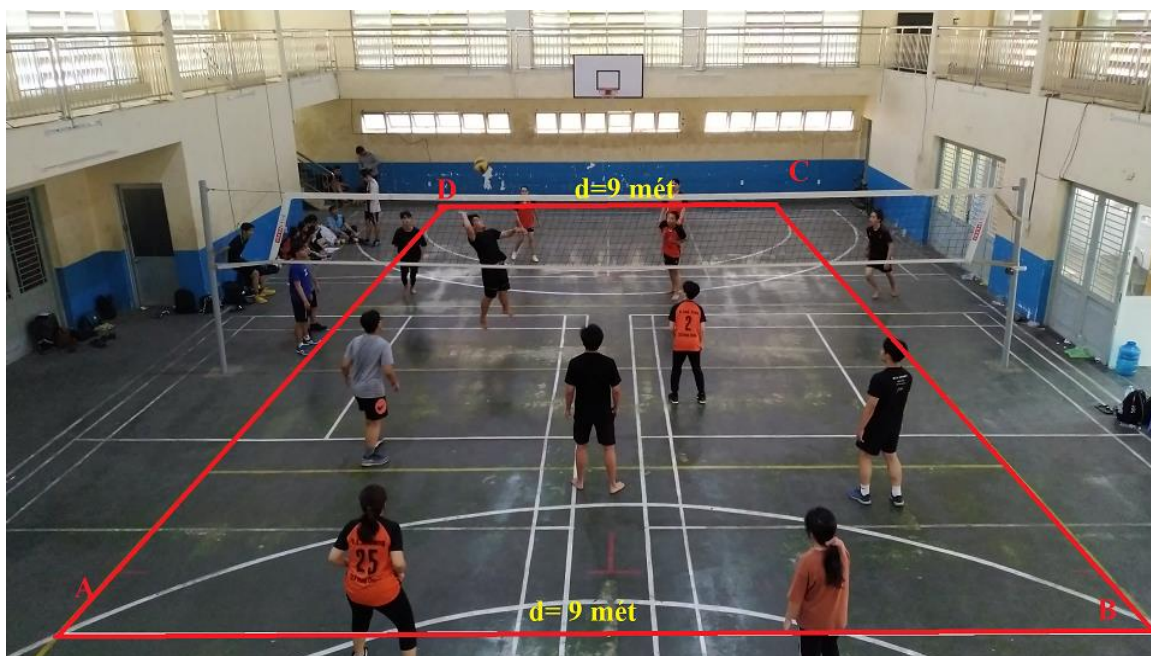


Hình 10. Cấu trúc của khối phát hiện đối tượng dùng trong mô hình Faster R-CNN[18]

3. THUẬT TOÁN TÍNH KHOẢNG CÁCH

Khởi tạo giá trị cơ sở

Trong thực tế, độ dài đoạn $AB = DC = d = 9$ (m) được đo như hình 11. Nhưng trên ảnh thu được thì ta thấy đoạn $AB > DC$. Dùng thuật toán tính độ dài pixel 2 đoạn AB , CD và tất cả các đoạn thẳng nằm giữa (AB , CD) được giới hạn 2 đầu bởi AD và BC ta thu được độ dài pixel của từng đoạn với cùng độ dài d .



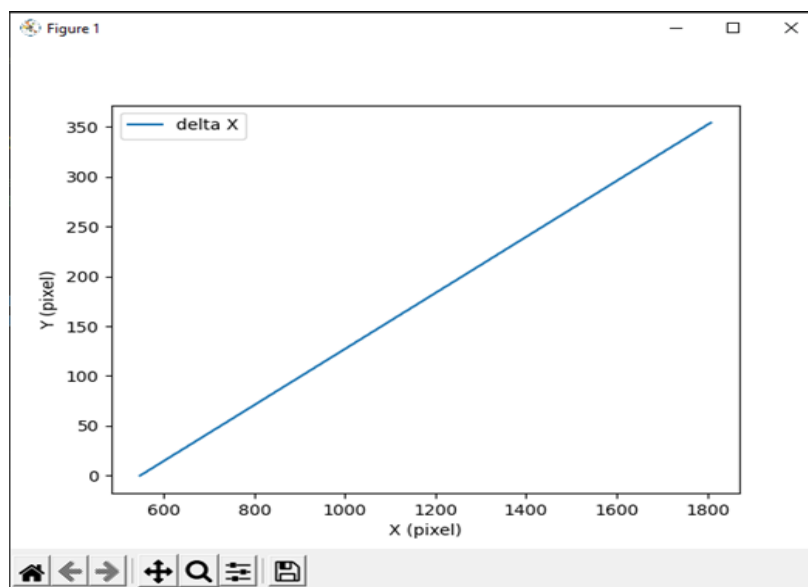
Hình 11. Khoảng cách đo thực tế.

Trên thực tế, ta biết được tất cả các đoạn vàng trên (Hình 12) đều bằng $d=9$ (mét) và ảnh thu được trên camera thì độ dài (pixel) của chúng được tăng khi càng về phía camera - tăng theo trục Y .



Hình 12. Khoảng cách pixel đo được bằng thuật toán trên các đoạn vàng.

Qua tính toán, ta có độ dài các đoạn vàng (Hình 12) tăng tuyến tính theo (Hình 13) với X là số pixel của đoạn vàng (Hình 12).



Hình 13. Biểu đồ tương quan độ dài pixel các đoạn vàng (Hình 26) theo trục Y.

Ta chọn đoạn đã được đo thực tế đặt ở gần camera nhất (đoạn AB trong hình 11) làm gốc thì ta có:

$$P = \frac{n}{d} \quad (\text{pixel/m}) \quad (1)$$

Trong đó: n là số pixel ứng với đoạn AB.
 d là độ dài thực tế của AB.

Với bất kì đoạn vàng khác đoạn AB ta luôn có:

$$P1 = \frac{n1}{d} \quad (\text{pixel/m}) \quad (2)$$

Trong đó: n1 là số pixel ứng với đoạn vàng (Hình 12).

$$P1 = \frac{n1}{n} \cdot P \quad (3)$$

Trong đó: $(n1)/n = k$ thể hiện độ phóng đại kích thước k (lần) của P1 so với P.

Ta tiến hành làm như trên với tất cả các đoạn vàng để thu được một tập hợp:

$$S = \left\{ k(i) = \frac{n(i)}{n} \mid i \in N, k \in R \right\} \quad (4)$$

Tập hợp S chứa tất cả các tỉ lệ phóng đại kích thước so với đoạn ban đầu.

Tính khoảng cách thực tế với 2 điểm

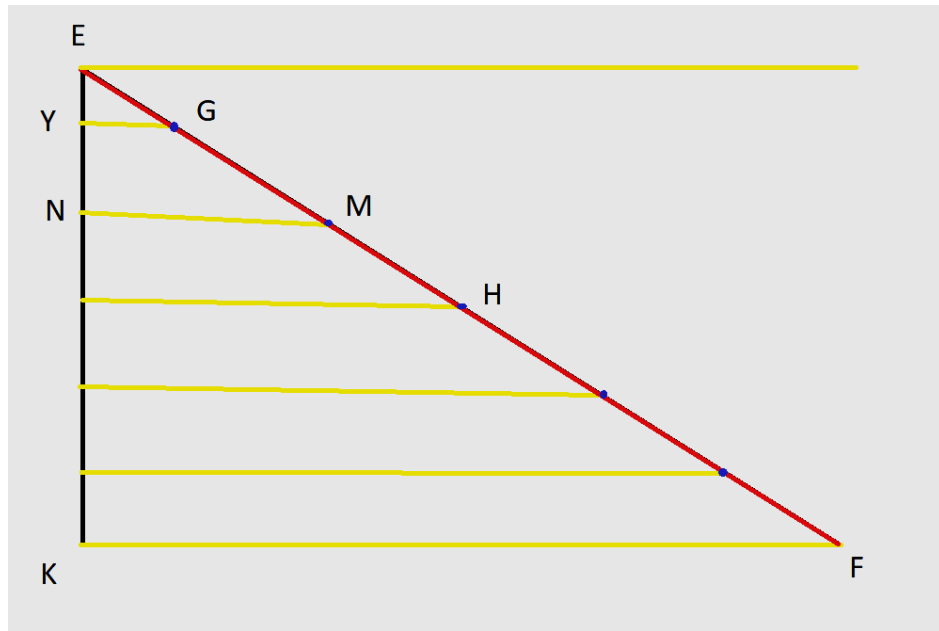
Ta có tọa độ 2 điểm (E và F) là tâm của hộp vuông của người và mô đun của đoạn EF (pixel) > 1 thì sẽ được cắt bởi n đường vàng (Hình 12) có n - 1 tỉ lệ phóng đại kích thước k ($k \in S$). Dùng định lí Thales để chứng minh và tính toán độ dài đoạn EF đã được phóng đại so với đoạn gốc.

Công thức tính mô đun 2 điểm:

$$|EF| = \sqrt{(X_2 - X_1)^2 + (Y_2 - Y_1)^2} \quad (5)$$

Trong đó: E(X1,Y1), F(X2,Y2)

Cơ sở tính toán thuật toán và chứng minh:



Hình 14. Tam giác chứng minh hệ quả Thales với đoạn EF là khoảng cách cần tìm

Xét tam giác EKF (Hình 14). Ta có $YG \parallel KF$:

$$\frac{EY}{EK} = \frac{EG}{EF} \quad (6)$$

Biến đổi tương đương:

$$EG = \frac{EY}{EK} \cdot EF \quad (7)$$

Xét tam giác EKF Ta có $MN \parallel KF$:

$$\frac{EN}{EK} = \frac{EM}{EF} \quad (8)$$

Lấy (8) - (6) ta được:

$$\frac{EN}{EK} - \frac{EY}{EK} = \frac{EM}{EF} - \frac{EG}{EK}$$

Biến đổi tương đương:

$$\frac{YN}{EK} = \frac{GM}{EF} \quad (9)$$

Biến đổi tương đương:

$$GM = \frac{YN}{EK} \cdot EF \quad (10)$$

Chứng minh tương tự cho đoạn MH và các đoạn tiếp theo.

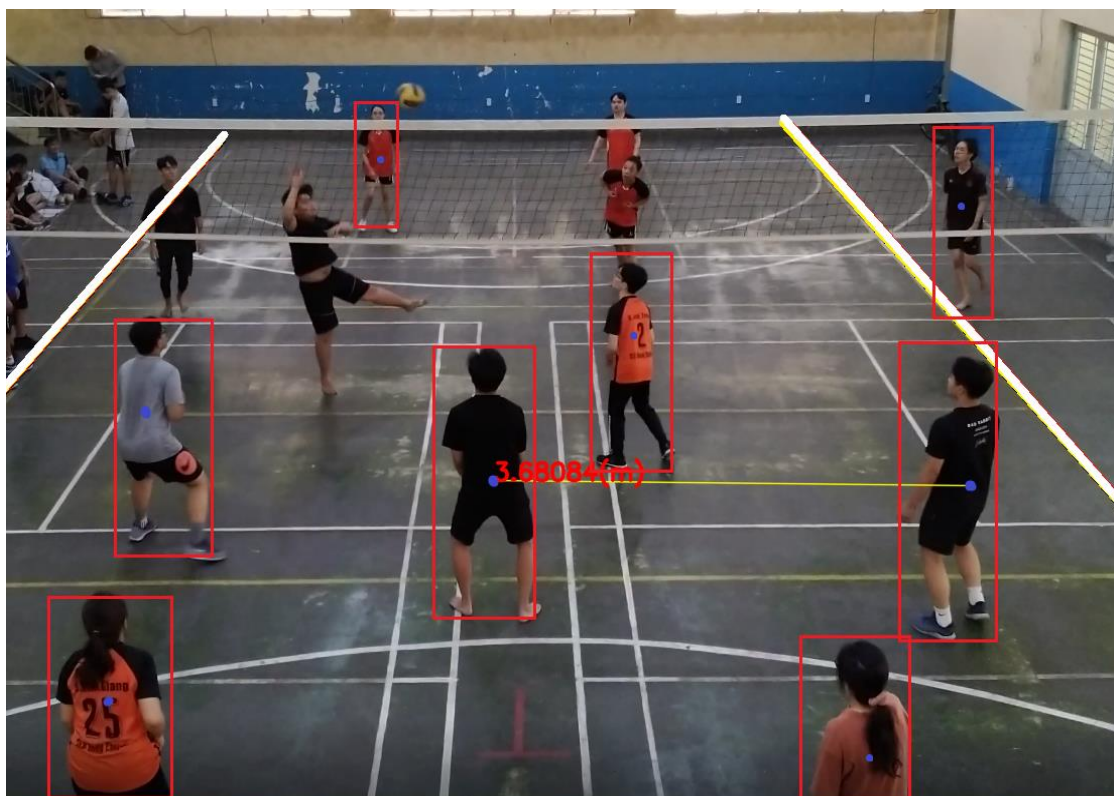
Khoảng cách thực tế EF (mét) theo tỉ lệ tương ứng là:

Từ (7) và (9) suy ra:

$$EF(m) = \left(\frac{EG}{k_0} + \frac{GM}{k_1} + \dots \right) \cdot \frac{1}{P} = \left(\frac{EY}{k_0} + \frac{YN}{k_1} + \dots \right) \cdot \frac{EF(\text{pixel})}{EK} \cdot \frac{1}{P} \quad (11)$$

Kết quả thực tế như ở hình 15 cho thấy thuật toán có tính khả thi.

4. HUẤN LUYỆN VÀ KIỂM TRA MÔ HÌNH NHẬN DIỆN NGƯỜI

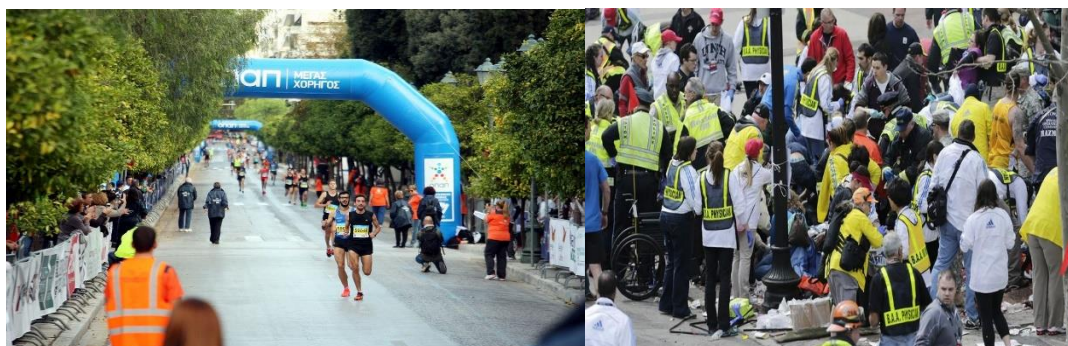


Hình 15: Kết quả khi chạy kiểm tra thuật toán

Phương pháp đề xuất được thực hiện trên máy tính có cấu hình như sau: CPU Intel Core I3-10105F, GPU Nvidia RTX 3060, RAM DDR4 16GB. Ngôn ngữ lập trình được sử dụng là Python với thư viện Pytorch cho các mô hình học sâu và thư viện OPENCV cho các thuật toán xử lý ảnh đơn giản. Chi tiết về quá trình thực nghiệm mô hình được trình bày chi tiết ở các phần sau đây.

Tập dữ liệu huấn luyện

Để thực hiện hệ thống phát hiện người, bộ dữ liệu có tên là “Human Detection” được sử dụng. Bộ dữ liệu gồm có 14175 ảnh với độ phân giải đa dạng được thu thập từ thực tế. Các ảnh trong bộ dữ liệu được chia làm 3 lớp: lớp là người, lớp không là người và lớp nền. Hình 16 giới thiệu một vài ảnh tương trưng với mỗi lớp. Mỗi ảnh được dán nhãn và kèm theo trong file xml tương ứng. Các nhãn được gán gồm có thông tin về vị trí hộp vuông của đối tượng và nhãn lớp tương ứng với hộp vuông.



Hình 16. Ảnh từ bộ dữ liệu Human Detection.

Vì bộ dữ liệu Human Detection không phân chia ảnh cho huấn luyện và kiểm tra, chúng tôi tiến hành chia ngẫu nhiên bộ dữ liệu thành 2 phần: huấn luyện và kiểm tra với tỉ lệ là 5:1. Chương trình để phân chia bộ

dữ liệu được viết trong file `split_dataset.py`.

Kết quả huấn luyện mô hình

Mô hình Faster R-CNN cho bộ dữ liệu nhận dạng người được huấn luyện với số epochs¹ = 20. Tốc độ học được chọn mặc định như mô hình gốc [1]. Các thông số huấn luyện khác giữ nguyên như mô hình gốc. Kết quả huấn luyện sau 20 epochs được thể hiện như hình 17 (có đính kèm trong file `train_log.txt`). Có thể thấy sau 20 epochs, mô hình có độ chính xác trung bình ứng với Intersection over Union² (IoU) 0.5 là 71.5%.

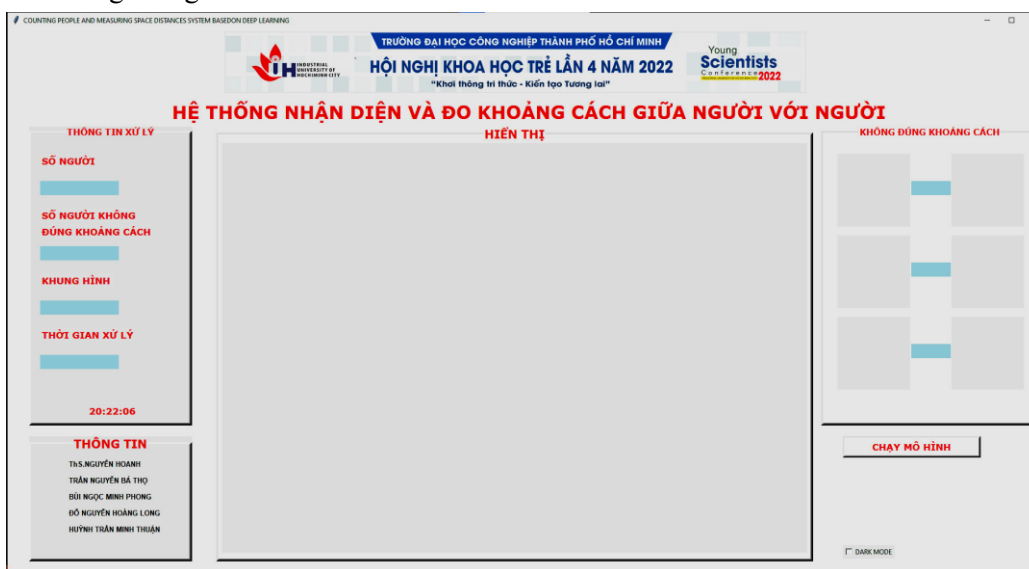
Average Precision	(AP) @[IoU=0.50:0.95 area= all maxDets=100]	= 0.420
Average Precision	(AP) @[IoU=0.50 area= all maxDets=100]	= 0.715
Average Precision	(AP) @[IoU=0.75 area= all maxDets=100]	= 0.477
Average Precision	(AP) @[IoU=0.50:0.95 area= small maxDets=100]	= 0.377
Average Precision	(AP) @[IoU=0.50:0.95 area=medium maxDets=100]	= 0.525
Average Precision	(AP) @[IoU=0.50:0.95 area= large maxDets=100]	= 0.808
Average Recall	(AR) @[IoU=0.50:0.95 area= all maxDets= 1]	= 0.201
Average Recall	(AR) @[IoU=0.50:0.95 area= all maxDets= 10]	= 0.465
Average Recall	(AR) @[IoU=0.50:0.95 area= all maxDets=100]	= 0.503
Average Recall	(AR) @[IoU=0.50:0.95 area= small maxDets=100]	= 0.446
Average Recall	(AR) @[IoU=0.50:0.95 area=medium maxDets=100]	= 0.605
Average Recall	(AR) @[IoU=0.50:0.95 area= large maxDets=100]	= 0.837

Hình 17. Kết quả huấn luyện mô hình sau 20 epochs.

5. THIẾT KẾ GIAO DIỆN

Để có thể áp dụng mô hình đã huấn luyện vào sử dụng trong thực tế, chúng tôi đề xuất xây dựng giao diện để giúp người sử dụng có thể thao tác dễ dàng. Giao diện được xây dựng như hình 18. Để xây dựng giao diện tích hợp mô hình, thư viện Tkinter được sử dụng. Các thông tin chính trên giao diện được thiết kế gồm:

- **Phần hiển thị thông tin phần mềm:** Gồm tên đề tài; thông tin tác giả; logo.
- **Phần hiển thị ảnh và kết quả phát hiện:** hiển thị ảnh kết quả của mô hình với người được dự đoán trên ảnh; hiển thị frame ảnh đang được xử lý (với video); hiển thị thời gian xử lý trên ảnh/frame ảnh; hiển thị kết quả phát hiện gồm số người phát hiện được, hiển thị một số trường hợp không đúng khoảng cách; hiển thị khoảng cách giữa người.
- **Phần nút nhấn:** nhấn nút để khởi động hệ thống với mặc định là camera được chọn.
- **Checkbox:** đổi giữa lightmode và darkmode.



Hình 18: Giao diện chính của hệ thống.

¹ Epochs: Một lần đưa tất cả ảnh vào mạng nơ-ron.

² Intersection over Union: Là tỉ số dự đoán hộp vuông của đối tượng với đối tượng thật.



Hình 19: Giao diện chọn video từ nơi bất kỳ trên máy.

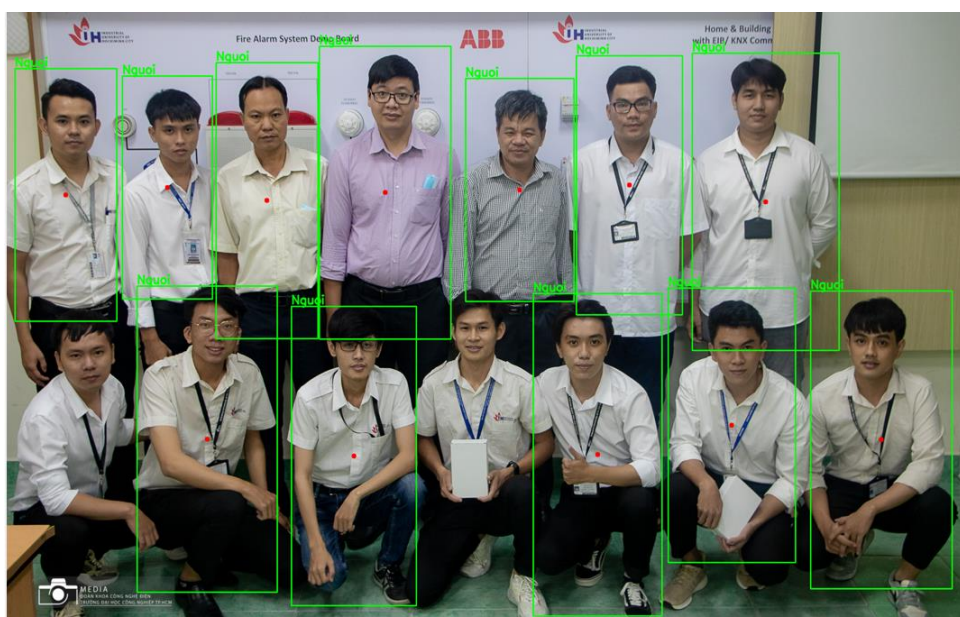
6. KẾT QUẢ CHẠY HỆ THỐNG THỰC TẾ

Kết quả chạy hệ thống nhận diện người

Ta thấy rằng, độ chính xác và hiệu quả của bộ dữ liệu được xây dựng “Human Dectection” cao và có thể phát hiện chính xác đối tượng là người trong từng khung ảnh mà camera thu được. Chúng tôi tiến hành kiểm tra mô hình thông qua file *test.py* bằng những video thực tế và được thống kê lại như trong bảng 1.

Bảng 1: Kết quả thực nghiệm mô hình trên file *test.py*.

Số người trên khung hình	Số người nhận diện được	Số người chưa nhận diện được	Tỉ lệ chính xác	Thời gian xử lí 1 khung ảnh
14	12	2	86%	0.02s



Hình 20. Kết quả nhận diện người trên file *test.py*

Nguồn: Ảnh thu thập trên Internet

Kết quả chạy hệ thống nhận diện người và đo khoảng cách

Nhóm tác giả tiến hành chạy thực tế hệ thống với video thu thập từ camera với các góc độ khác nhau. Hình 21 và hình 22 cho thấy hệ thống hoạt động ổn định và có độ chính xác cao. Phần lớn người trong khung hình đều được hệ thống nhận dạng một cách chính xác, tốc độ xử lý của hệ thống đáp ứng tốc độ thực tế của camera.



Hình 21. Kết quả chạy hệ thống trên video quay thực tế.



Hình 22. Kết quả chạy hệ thống trên video quay thực tế.

Bảng 2 cho thấy hệ thống có tính chính xác, khả năng tính toán của hệ thống sẽ tối ưu nhất khi đối tượng ở trong khoảng 0(m) đến 5(m) và khung ảnh không bị nhiễu. Ngược lại, hệ thống sẽ khó phát hiện đối tượng ở xa camera và sai số tính toán sẽ càng cao. Ở các trường hợp khác, các đối tượng trong khung ảnh xếp chồng lên nhau thì hệ thống không thể xác định chính xác khoảng cách giữa họ. Kết luận, đối tượng khi càng xa camera thì thuật toán sẽ có sai số càng lớn, vậy để mô hình hoạt động trong khoảng 0(m) đến 10(m) tính từ camera sẽ đem lại sự tối ưu.

Bảng 2: Kết quả thử nghiệm tính khoảng cách.

Khoảng cách tính từ camera (m)	0 (m) - 5 (m)	5 (m) - 10 (m)	10 (m) - 15 (m)	15 (m) - 20 (m)
Khoảng cách tính được(m)	3,102	3,276	3,373	3,495
Khoảng cách thực tế (m)	3	3	3	3
Sai số (m)	0,102	0,276	0,373	0,495
Sai số (%)	3,4%	9,2%	12,43%	16,67%

7. KẾT LUẬN

Nghiên cứu đã đạt được một số kết quả nhất định, chúng tôi đã xây dựng hệ thống nhận diện người và đo khoảng cách giữa các đối tượng là người dựa trên mô hình Faster RCNN và học sâu, đem lại nhiều ý nghĩa thực tiễn trong công cuộc phòng chống dịch bệnh do tiếp xúc gần nói chung và đại dịch COVID nói riêng. Hiện tại, mô hình đã hoạt động ổn định trong điều kiện thực tế, tỉ lệ nhận diện người chính xác trên từng khung ảnh đạt trên 86%. Ngoài ra, sai số tính toán khoảng cách so với thực tế ± 0.3 (mét) do việc đo và cấu hình trong phần mềm. Tốc độ xử lý của hệ thống đã được cải thiện hơn để đáp ứng được nhu cầu xử lý thời gian thực, kết hợp với một giao diện trực quan và dễ sử dụng góp phần giúp đội ngũ quản lý nắm bắt thông tin một cách kịp thời và chính xác. Nghiên cứu đã góp phần bảo vệ sức khỏe cộng đồng ở những nơi công cộng.

Bên cạnh đó, mô hình còn nhiều hạn chế về khả năng nhận diện người chưa được thông minh, phát hiện sai. Cụ thể, mô hình chưa phát hiện đầy đủ đối tượng là người trong các trường hợp khung ảnh bị nhiễu, bị thiếu sáng, bị quá nhỏ, ... Thuật toán đo khoảng cách còn nhiều sai sót khi đối tượng càng xa camera thì sai số khoảng cách so với thực tế càng lớn. Nếu hai đối tượng trong khung ảnh chồng lên nhau thì mô hình sẽ không xác định được khoảng cách. Thuật toán đo khoảng cách chỉ hoạt động chính xác trong không gian đã được đo đạc từ trước.

Để khắc phục những hạn chế nêu trên chúng tôi đề xuất nghiên cứu xây dựng một hệ thống sử dụng mạng nơ-ron ResNet152 để tăng tốc độ xử lý và tính chính xác của mô hình Faster RCNN. Về thuật toán đo khoảng cách, để tăng tính chính xác chúng tôi đề xuất sử dụng phương pháp sử dụng 2 camera kết hợp các cảm biến.

TÀI LIỆU THAM KHẢO

- [1] Ren, S., He, K., Girshick, R., et al.: "Faster r-cnn: towards real-time object detection with region proposal networks", Adv. Neural. Inf. Process. Syst., 2016, pp. 91–99, arXiv:1506.01497.
- [2] Liu, W., Anguelov, D., Erhan, D., et al.: "SSD: single shot MultiBox detector". European Conf. on Computer Vision, Amsterdam, The Netherlands, 2016, pp. 21–37.
- [3] Xinyi Zhou; Wei Gong; WenLong Fu; Fengtong Du. " Application of deep learning in object detection." IEEE Xplore: 29 June 2017, Accession Number: 16997426.
- [4] G Chandan; Ayush Jain; Harsh Jain; Mohana, " Real Time Object Detection and Tracking Using Deep Learning and OpenCV." , IEEE Xplore: 03 January 2019 : 18357982
- [5] Shih-Chung Hsu; Yu-Wen Wang; Chung-Lin Huang , "Human Object Identification for Human-Robot Interaction by Using Fast R-CNN", IEEE Xplore: 05 April 2018: 17669499.
- [6] Bin Liu; Wencang Zhao; Qiaoqiao Sun "Study of object detection based on Faster R-CNN" IEEE Xplore: 01 January 2018, Accession Number: 17484408.
- [7] M Lokanath, K Sai Kumar and E Sanath Keerthi "Accurate object classification and detection by faster-RCNN" ,Citation M Lokanath et al 2017 IOP Conf. Ser.: Mater. Sci. Eng. 263 052028.
- [8] Zakariya Qawaqneh, Arafat Abu Mallouh, Buket D. Barkana. " Deep Convolutional Neural Network for Age Estimation based on VGG-Face Model." Computer Vision and Pattern Recognition (cs.CV), arXiv:1709.01664 [cs.CV] , <https://doi.org/10.48550/arXiv.1709.01664>.
- [9] George Papandreou, Tyler Zhu, Nori Kanazawa, Alexander Toshev, Jonathan Tompson, Chris Bregler, Kevin Murphy , " Towards Accurate Multi-Person Pose Estimation in the Wild." Proceedings of the IEEE Conference on

- Computer Vision and Pattern Recognition (CVPR), 2017, pp. 4903-4911.
- [10] Liu, Jingwei, Yi Gu, Shumin Han, Zhibin Zhang, Jiafeng Guo, and Xueqi Cheng. "Feature Rescaling and Fusion for Tiny Object Detection." *IEEE Access* 9 (2021): 62946-62955.
- [11] Mingxing Tan, Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks". arXiv:1905.11946v5 [cs.LG] 11 Sep 2020 : <https://github.com/tensorflow/tpu/tree/master/models/official/efficientnet>.
- [12] Friendly, F., Harizahayu, H., Sembiring, R. W., & Br Sembiring, B. S. (2021). "Measurement of Image Distance Using Only Camera On Object Detection OpenCV". *INFOKUM*, 10(1), 56-63. Retrieved from <http://infor.seaninstitute.org/index.php/infokum/article/view/217>.
- [13] Ik-Hwan Kim; Do-Eun Kim; You-Sung Cha; Kwang-hee Lee; Tae-Yong Kuc "An embodiment of stereo vision system for mobile robot for real-time measuring distance and object tracking", *Xplore*: 26 December 2007 : 9853364.
- [14] de La Bourdonnaye, A., Doskočil, R., Křivánek, V., & Štefek, A. (2022). Practical Experience with Distance Measurement Based on Single Visual Camera. *Advances in Military Technology*, 7(2), 49–56. Retrieved from <https://www.aimt.cz/index.php/aimt/article/view/1583>
- [15] Bowen Cheng, Yunchao Wei, Honghui Shi, Rogerio Feris, Jinjun Xiong, Thomas Huang. "Revisiting RCNN: On Awakening the Classification Power of Faster RCNN", *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 453-468
- [16] Zhang, D., Liu, Z., & Shi, X. (2020) "Transfer Learning on EfficientNet for Remote Sensing image Classification" 2020 5th International Conference on Mechanical, Control and Computer Engineering (ICMCCE).
- [17] Mike Liao "Detecting Objects in (almost) Real-time: FasterRCNN Explained with Code": <https://towardsdatascience.com/fasterrcnn-explained-part-1-with-code-599c16568cff>
- [18] Ross Girshick, "Fast R-CNN" *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, 2015, pp. 1440-1448
- [19] Neubeck; L. Van Gool, "Efficient Non-Maximum Suppression" *IEEE Xplore*: 18 September 2006 : 9210072.
- [20] Mukti, I. Z., & Biswas, D. (2019), "Transfer Learning Based Plant Diseases Detection Using ResNet50". 2019 4th International Conference on Electrical Information and Communication Technology (EICT).
- [21] Sheldon Mascarenhas, Mukul Agarwal "A comparison between VGG16, VGG19 and ResNet50 architecture frameworks for Image Classification", *IEEE Xplore*: 03 February 2022.
- [22] Zhang, Y., Song, C., & Zhang, D. (2020). "Deep Learning-based Object Detection Improvement for Tomato Disease". *IEEE Access*, 1–1.
- [23] Xiao, B., & Kang, S.-C. (2021). "Development of an Image Data Set of Construction Machines for Deep Learning Object Detection". *Journal of Computing in Civil Engineering*, 35(2), 05020005.
- [24] Limeng Qiao, Yuxuan Zhao, Zhiyuan Li, Xi Qiu, Jianan Wu, Chi Zhang; "Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)", 2021, pp. 8681-8690

YSC4F.207

MÔ HÌNH TOÁN HỌC Ồ TỪ DỌC TRỤC HÌNH E CẤU TRÚC NGUYÊN KHỐI

LÊ NGỌC HỘI^{1*}, NGUYỄN TÙNG LÂM²

¹Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

²Trường Đại học Bách khoa Hà Nội

lengochoi@iuh.edu.vn, lam.nguyentung@hust.edu.vn

Tóm tắt. Dòng điện xoáy được tạo ra trong cơ cấu chấp hành ồ đỡ từ dọc trục cấu tạo nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator. Do đó, dòng điện xoáy ảnh hưởng lớn đến động lực học và thiết kế điều khiển cơ cấu chấp hành ồ từ dọc trục cấu tạo nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator. Trong các công trình nghiên cứu của các tác giả trước đây đã đưa ra được mô hình toán học của ồ từ có cấu trúc nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator tương đối hoàn chỉnh. Tuy nhiên, các tác giả chỉ đưa ra trong trường hợp ồ từ dọc trục hình C. Đối với ồ từ dọc trục hình E được ứng dụng nhiều trong thực tế, nhưng các tác giả chưa đưa ra được mô hình toán học cả tuyến tính lẫn phi tuyến. Do đó, để so sánh với mô hình toán học hình C và hoàn thiện mô hình hóa ồ từ cấu trúc nguyên khối, trong bài báo này nhóm tác giả đã đưa ra mô hình toán học ở dạng tuyến tính và phi tuyến trên miền thời gian của ồ từ dọc trục hình E cấu trúc nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator. Kết quả mô phỏng bằng phần mềm Matlab cho thấy từ trở và từ thông trong công thức gần đúng bám sát công thức chính xác với sai số rất nhỏ. Từ đó cho thấy mô hình toán học tìm được là đáng tin cậy.

Từ khóa. ồ từ dọc trục, cấu tạo nguyên khối, từ trở hiệu dụng, mô hình toán học, đạo hàm cấp phân số.

MATHEMATICAL MODELLING OF NON-LAMINATED MAGNETIC BEARINGS WITH E-TYPE ACTUATOR GEOMETRY

Abstract. Eddy currents are generated in a non-laminated geometry axial magnetic bearing actuator by a time-varying magnetic field, which has a great influence on the dynamics and control design. In previous research, a mathematical model of a non-laminated magnetic bearing drive that operates when supplied by an alternating current to the stator windings is considered. However, the authors only focused on C-type geometry axial magnetic bearings, while both the linear and nonlinear mathematical models for E-type geometry axial magnetic bearings widely used in practice have not been yet studied. Therefore, to compare with the C-type geometry model and complete the mathematical model of non-laminated geometries. In this paper, both linear and nonlinear models in the time domain of the axial magnetic bearing of the E-type geometry supplied by an alternating current to the stator windings are introduced and demonstrated by evaluation by simulation. The simulation results using Matlab/Simulink environment show that the reluctance and flux in the approximate formula follow the exact formula with very small errors. From this, it shows that the found mathematical model is reliable.

Keywords. Axial magnetic bearing, non-laminated geometry, effective reluctance, mathematical modelling, fractional-order.

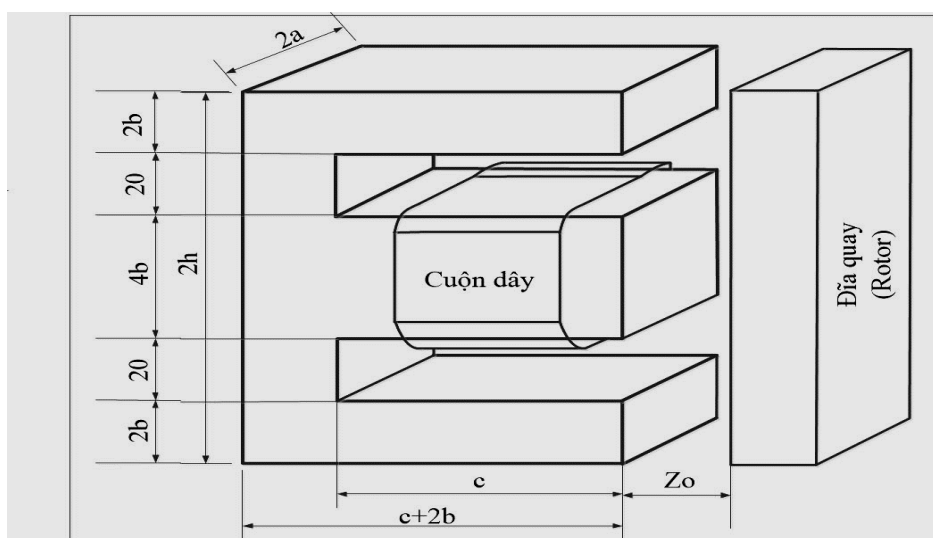
1. GIỚI THIỆU

Ồ đỡ từ đã được nghiên cứu cho nhiều ứng dụng công nghiệp và khoa học, bao gồm máy móc có cơ cấu chuyển động [1], [2], băng tải kim loại [3], quy trình phủ kim loại [4], quang khắc [5], và hệ thống servo công cụ [6]. Mạch từ (stator và rotor) của ồ đỡ từ thường được ghép bằng các lá thép kỹ thuật để giảm tổng

hao dòng xoáy khi có từ thông biến thiên trong trong vật liệu sắt từ. Tuy nhiên, đối với ổ từ dọc trục trong một số ứng dụng các cơ cấu chấp hành thường cấu tạo nguyên khối. Đặc biệt, do yêu cầu về độ bền cơ học nên đĩa quay của ổ từ dọc trục thường được cấu tạo nguyên khối. Ngoài ra, do mối quan tâm về chi phí nên các cơ cấu chấp hành của ổ đỡ từ được cấu tạo nguyên khối. Trong các ứng dụng của ổ đỡ từ cấu tạo nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây, dòng xoáy sẽ ảnh hưởng sâu sắc đến hoạt động của cơ cấu chấp hành và phải được xem xét trong mô hình hóa hệ thống và thiết kế bộ điều khiển. Công trình đầu tiên về mô hình toán học ổ đỡ từ cấu tạo nguyên khối đã được thực hiện bởi Zmood [7], người đã trình bày mô hình toán học cho ổ từ hình C cấu tạo nguyên khối có tỷ lệ chiều rộng rất lớn so với chiều cao. Công trình thứ hai về động lực học ổ đỡ cấu tạo nguyên khối đã được trình bày bởi Feeley và Ahlstrom [15], tác giả cũng đã đưa ra được động lực học trên miền thời gian. Một công bố tiếp theo về mô hình hóa của cơ cấu chấp hành hình C cấu tạo nguyên khối đã được trình bày bởi Feeley [9]. Mặc dù các mô hình toán học được đưa ra trong [7], [9] và [15] khá đơn giản. Tuy nhiên, độ chính xác của các công bố trên là khá kém. Lý do chính là trong cả ba bài báo, các tác giả đã giả định rằng mật độ từ thông trong mặt cắt ngang của khe hở không khí giống như mật độ từ thông trong tiết diện phần sắt từ. Đây là một giả định chỉ đúng cho phân tích tĩnh nhưng không đúng cho trường hợp cấp dòng xoay chiều vào 2 đầu cuộn dây vì sinh ra dòng điện xoáy. Kucera and Ahrens [11], các tác giả đã trình bày các kết quả phân tích mô tả mối quan hệ dòng điện-lực tương ứng đối với stator hình trụ và hình chữ C. Tuy nhiên, tác giả cho rằng mật độ từ thông khe hở không khí không phụ thuộc vào tần số của trường điều hòa. Hơn nữa, tham số d (độ dày của các phần chia nhỏ trong cơ cấu chấp hành) trong mô hình toán học phân tích này phải được xác định từ kết quả thực nghiệm, không tương ứng với bất kỳ hình dạng và vật liệu của cơ cấu chấp hành. Do đó, kết quả này không phù hợp với việc tối ưu hóa thiết kế cơ cấu chấp hành ổ từ dọc trục. Mặc dù các tác giả [14] và [21] đã đưa ra mô hình toán học ổ từ dọc trục hình C cấu tạo nguyên khối tương đối hoàn chỉnh. Tuy nhiên, đối với ổ từ dọc trục hình E cấu trúc nguyên khối các tác giả trước đây chưa đưa ra được mô hình toán học dạng tuyến tính và phi tuyến trên miền thời gian. Hơn nữa đối với ổ từ hình E cấu trúc nguyên khối ứng dụng nhiều trong thực tế là rất cần thiết. Trong bài báo này nhóm tác giả đã trình bày một phương pháp để đưa ra động lực học tuyến tính và phi tuyến của ổ từ dọc trục hình E cấu trúc nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator. Bài báo này được tổ chức như sau: Phần 2 mô tả ngắn gọn về từ trở hiệu dụng và từ thông ổ từ dọc trục hình E cấu trúc nguyên khối. Phần 3. Trình bày về mô hình toán học dạng tuyến tính và phi tuyến trên miền thời gian của ổ từ dọc trục hình E cấu trúc nguyên khối. Kết luận được rút ra trong phần 4.

2. TỪ TRỞ HIỆU DỤNG VÀ TỪ THÔNG Ổ TỪ DỌC TRỤC HÌNH E

Xét ổ từ dọc trục hình E cấu tạo nguyên khối được có kích thước được thể như hình 1 và thông số như bảng 1.

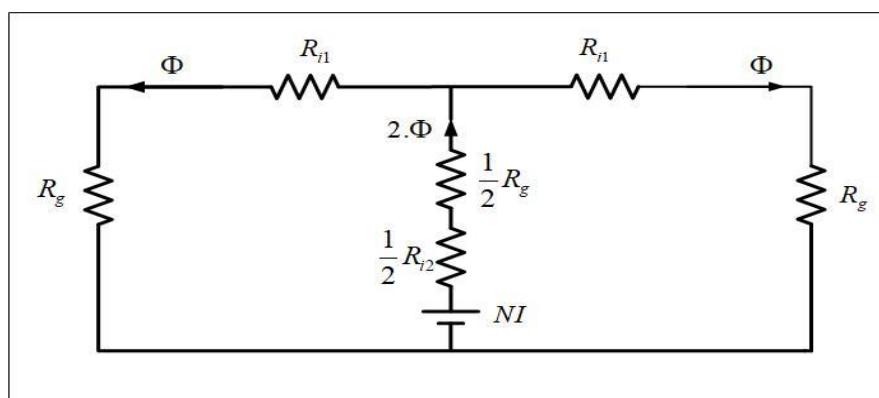


Hình 1: Ổ từ dọc trục hình E cấu trúc nguyên khối

Bảng 1: Thông số cơ cấu chấp hành ổ từ dọc trục hình C

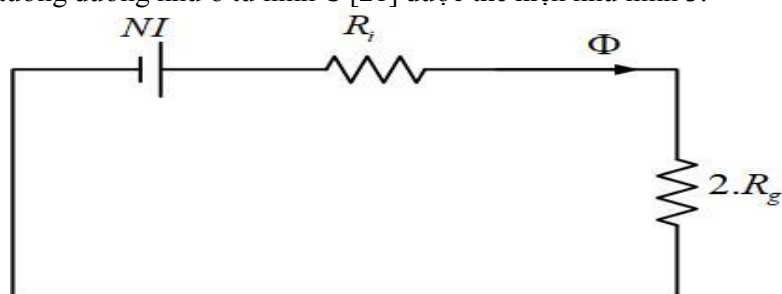
STT	Ký hiệu	Mô tả thông số	Giá trị
1	σ	Độ dẫn điện của sắt	2.5×10^6 S/m
2	a	1/2 Chiều rộng của cực từ	27.5 mm
3	b	1/2 Chiều cao của cực từ	10 mm
4	c	Chiều rộng cơ cấu chấp hành	60mm
5	h	1/2 chiều dài cơ cấu chấp hành	60mm
6	N	Số cuộn dây	280
7	z_o	Khoảng cách khe hở không khí danh định	0.2mm
8	μ_r	Độ từ thẩm tương đối	5000
9	μ_o	Độ từ thẩm chân không	$4\pi \times 10^{-7}$ T.m/A
10	I	Dòng điện	3.5A

Từ sơ đồ cấu trúc của ổ từ dọc trục hình E cấu tạo nguyên khối hoạt động khi cấp dòng điện xoay chiều như hình 1 suy ra được mạch từ tương đương được thể hiện trên hình 2.



Hình 2: Mạch từ tương đương của ổ từ hình E cấu trúc nguyên khối

Từ kích thước của ổ từ dọc trục hình E cấu tạo nguyên khối hoạt động khi cấp dòng điện xoay chiều như hình 1 và sơ đồ mạch từ tương đương của ổ từ hình E như hình 2 có thể tách mạch từ của ổ từ dọc trục hình E thành 2 mạch từ tương đương như ổ từ hình C [21] được thể hiện như hình 3.



Hình 3: Mạch từ tương đương của ổ đỡ từ cấu tạo nguyên khối

Trong đó:

$R_i = R_{i1} + R_{i2}$: Từ trở hiệu dụng phần tử sắt trong hình 3.

R_g : Từ trở hiệu dụng phần tử không khí trong hình 3.

NI: Sức từ động.

Φ : Từ thông chạy trong hình 3.

Theo [14] và [21] từ trở hiệu dụng của mạch từ hình C được chia thành 2 phần chính được thể hiện trên hình 3 bao gồm: Từ trở hiệu dụng phần tử không khí và từ trở hiệu dụng phần tử sắt từ.

Đối với tổng từ trở hiệu dụng chính xác của phần tử không khí theo [14] xác định như sau:

$$R_i = R_i^0 \frac{1}{\frac{\tanh(\alpha b)}{\alpha b} + \sum_{n=1}^{\infty} \frac{8}{(2n-1)^2 \pi^2} \frac{\alpha^2 \tanh(\beta_{2n} a)}{\beta_{2n}^2 \beta_{2n} a}} \quad (1)$$

Theo [14] xấp xỉ của (1) ta được từ trở hiệu dụng gần đúng:

$$R_i(s) = R_i^0 + \frac{l_i}{4(a+b)} \sqrt{\frac{\sigma}{\mu_r \mu_0}} \sqrt{s} \quad (2)$$

Đối với tổng trở hiệu dụng chính xác của phân tử không khí theo [14] xác định như sau:

$$R_g = \frac{z_0}{\mu_0} \frac{1}{\frac{4a \tanh(\alpha_1 b)}{\alpha_1} + \sum_{n=1}^m \left(\frac{4\sqrt{2}\alpha_1 b^{1/2}}{(2n-1)\pi\beta_{1n}^{3/2}} \right)^2 \tanh(\beta_{1n} a)} \quad (3)$$

Theo [14] và [21] xấp xỉ của (3) ta được từ trở hiệu dụng gần đúng:

$$\tilde{R}_g = R_g^0 + k_1 \sqrt{s} \quad (4)$$

Trong đó:

$$k_1 = \left(\left(\begin{aligned} & -\frac{243215552}{253125} e^{\frac{2\pi a}{b}} + \frac{243215552}{253125} e^{\frac{5\pi a}{b}} - \frac{146015552}{253125} e^{\frac{6\pi a}{b}} + \frac{48815552}{253125} e^{\frac{7\pi a}{b}} + \\ & \frac{93440}{81} e^{\frac{3\pi a}{b}} - \frac{93440}{81} e^{\frac{4\pi a}{b}} + \frac{146015552}{253125} e^{\frac{\pi a}{b}} - \frac{48815552}{253125} \end{aligned} \right) x \sqrt{\frac{\sigma}{\mu_r \mu_0}} \right. \\ \left. + \frac{b^2}{6\pi^5 a^2 (-e^{\frac{7\pi a}{b}} + e^{\frac{6\pi a}{b}} - e^{\frac{5\pi a}{b}} - e^{\frac{2\pi a}{b}} + e^{\frac{\pi a}{b}} - 1)} + \frac{b}{6a} \right) \quad (5)$$

$s=j\omega$ là toán tử laplace

Từ trở tổng trong ổ từ hình C cấu trúc nguyên khối có mạch từ được xác định theo hình 3:

$$R_C(s) = R_i(s) + 2R_g(s) = \left(\frac{l_i}{\mu_r \mu_0 A} + \frac{2z_0}{\mu_0 A} \right) + k\sqrt{s} = R_g^0 + k\sqrt{s} \quad (6)$$

Từ trở tổng trong ổ từ hình E cấu trúc nguyên khối được có mạch từ xác định theo hình 2:

$$R_E(s) = \frac{1}{2}R_{i1}(s) + \frac{1}{2}R_{i2}(s) + R_g(s) = \frac{1}{2}R_i(s) + R_g(s) = \left(\frac{l_i}{2\mu_r \mu_0 A} + \frac{z_0}{\mu_0 A} \right) + \frac{1}{2}k\sqrt{s} = \frac{1}{2}R_g^0 + \frac{1}{2}k\sqrt{s} \quad (7)$$

Trong đó:

$$k = 2k_1 + \frac{l_i}{4(a+b)} \sqrt{\frac{\sigma}{\mu_r \mu_0}} \quad (8)$$

k_1 : được xác định lần lượt theo (5)

$$R_i^0 = \frac{l_i}{\mu_r \mu_0 A}, \quad R_g^0 = \frac{z_0}{\mu_0 A}, \quad l_i = 2c + 2h. \quad (9)$$

Ta có từ thông trong ổ từ hình C được xác định theo công thức sau:

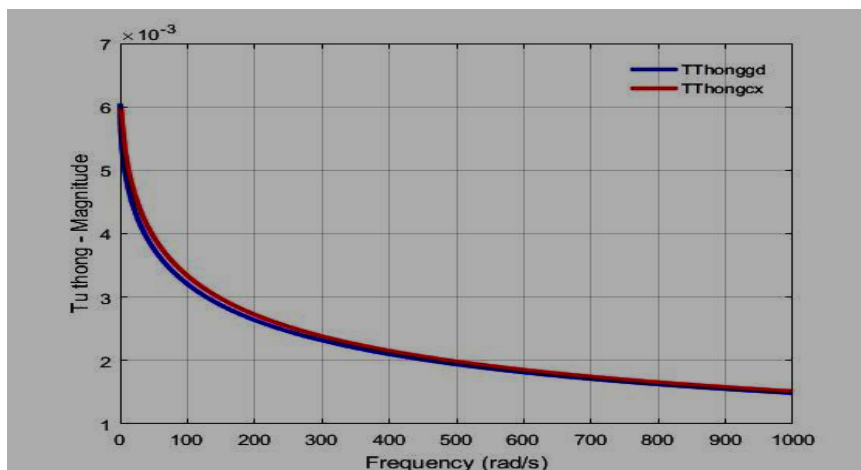
$$\Phi(\omega) = \frac{N.I}{2R_g(s) + R_i(s)} \quad (10)$$

Suy ra từ thông trong ổ từ hình E được xác định từ từ thông ổ từ hình C như sau:

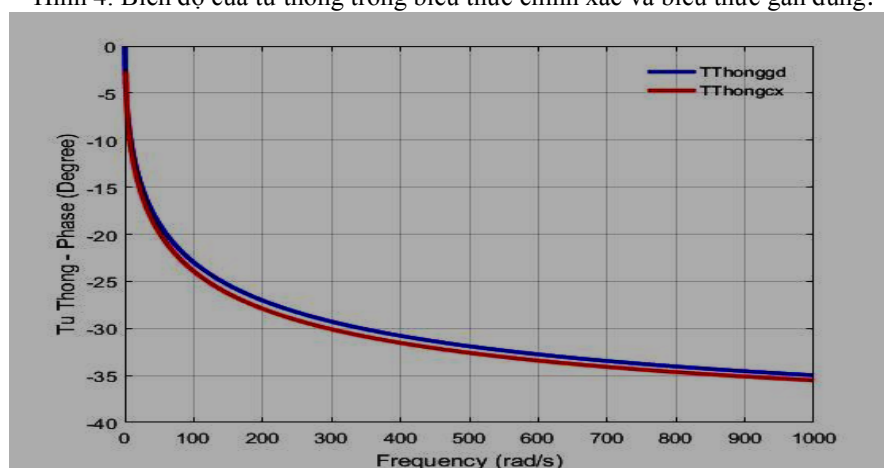
$$\Phi_1(\omega) = 2\Phi = \frac{2N.I}{2R_g(s) + R_i(s)} \quad (11)$$

Từ (6), (7), (10) và (11) tác giả đưa ra nhận xét: đối với từ trở của hình E bằng 1/2 từ trở của hình C, đối với từ thông của ổ từ dọc trục hình E lớn gấp 2 lần từ thông của ổ từ dọc trục hình C, đối với lực điện từ của ổ từ dọc trục hình E lớn gấp 2 lần lực điện từ của ổ từ dọc trục hình C.

Kết quả mô phỏng matlab (10) đối với công thức chính xác khi từ trở hiệu dụng $R_i(s)$ và $R_g(s)$ được xác định theo công thức chính xác như [14] và công thức gần đúng khi từ trở hiệu dụng $R_i(s)$ và $R_g(s)$ được xác định theo công thức gần đúng như [14] và [21] được thể hiện trên hình 4-7. Kết quả mô phỏng cho thấy đường từ thông trong công thức gần đúng tiến sát với đường từ thông trong công thức chính xác.



Hình 4: Biên độ của từ thông trong biểu thức chính xác và biểu thức gần đúng.



Hình 5: Pha của từ thông trong biểu thức chính xác và biểu thức gần đúng.

Ta có công thức tính sai số giữa từ thông chính xác và từ thông gần đúng được tính như sau:

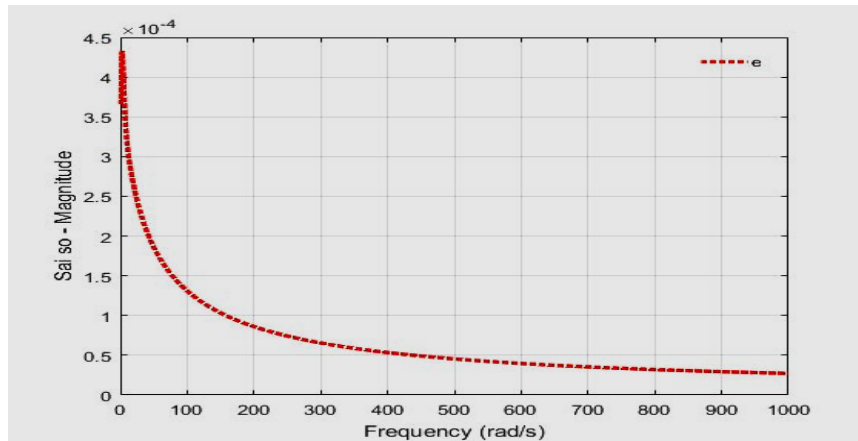
$$e = \Phi_{1cx}(\omega) - \Phi_{1gd}(\omega) \quad (12)$$

Trong đó:

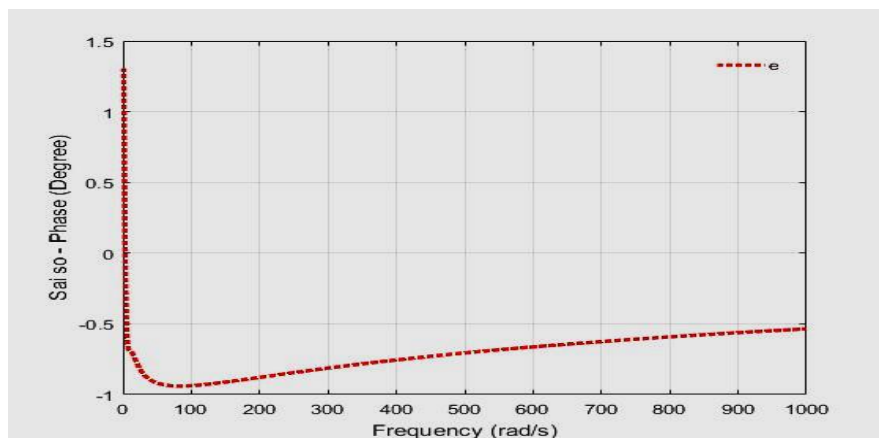
$\Phi_{1cx}(\omega)$: Từ thông ổ từ dọc trục hình E được tính theo công thức chính xác

$\Phi_{1gd}(\omega)$: Từ thông ổ từ dọc trục hình E được tính theo công thức gần đúng.

Từ đó ta tiến hành mô phỏng sai số ta được kết quả như hình 6-7.



Hình 6: Biên độ sai số giữa từ thông trong biểu thức chính xác và biểu thức gần đúng.



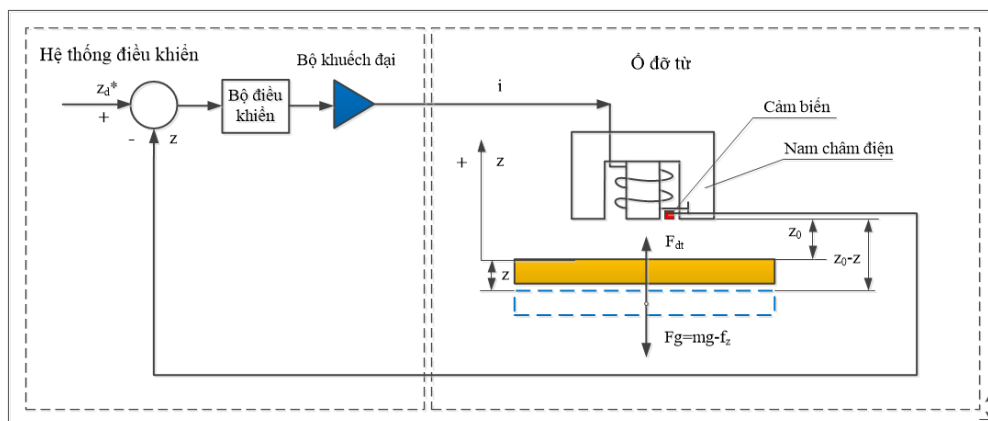
Hình 7: Pha sai số giữa từ thông trong biểu thức chính xác và biểu thức gần đúng.

Từ hình 6-7 tác giả nhận thấy sai số giữa từ thông được tính theo công thức chính xác và từ thông được tính theo công thức gần đúng là khá nhỏ và giá trị sai số luôn mang giá trị dương, tức là giá trị từ thông trong công thức gần đúng bao giờ cũng nhỏ hơn giá trị từ thông trong công thức chính xác. Giá trị sai số cực đại $e_{max}=4.4 \times 10^{-4}$ Wb, giá trị nhỏ nhất $e_{min}=0.3 \times 10^{-4}$ Wb. Điều đó nói lên rằng công thức gần đúng của từ thông đã tiến sát với công thức chính xác của từ thông.

3. MÔ HÌNH TOÁN HỌC

3.1. Mô hình tuyến tính

3.1.1. Mô hình toán học một cực từ



Hình 8: Cấu trúc điều khiển và hoạt động ổ từ cấu trúc nguyên khối 1 cực từ.

Cấu trúc hệ thống điều khiển được thể hiện như hình 8 gồm: một rotor được treo tự do tại một khoảng cách mong muốn z_0 so với cơ cấu điện từ, một cảm biến vị trí đo độ sai lệch z giữa vị trí mong muốn z_0 với vị trí thực của rotor và đưa tín hiệu này đến bộ điều khiển. Nhiệm vụ của bộ điều khiển là cung cấp một tín hiệu điều khiển thích hợp nhằm ổn định hệ thống, duy trì vị trí của rotor tại vị trí mong muốn, tức là làm thỏa mãn sự cân bằng giữa lực hấp dẫn F_g được tạo ra với trọng lực rotor mg tại điểm làm việc và các ngoại lực f_z tác động theo phương thẳng đứng. Bộ điều khiển sẽ xử lý thông tin thu được và gửi tín hiệu điều khiển đến bộ khuếch đại công suất để tạo thành dòng điện xoay chiều đưa đến cuộn dây của cơ cấu điện từ, dòng điện này sẽ tạo ra được lực điện từ F_{dt} giúp cân bằng lực tác động lên rotor theo phương trọng trường. Về cơ bản, luật điều khiển sẽ hoạt động theo cách thức: khi rotor dịch chuyển xuống so với vị trí mong muốn, bộ điều khiển sẽ đưa ra một tín hiệu làm tăng dòng điện điều khiển, từ đó gia tăng độ lớn lực điện từ để kéo rotor trở lại vị trí danh định của nó và ngược lại.

Theo [14], [16]-[18] và lực điện từ của 1 cực từ hình E lớn gấp 2 lần hình C, do đó lực điện từ trên miền tần số của 1 cực từ cấu tạo nguyên khối khi cấp dòng điện xoay chiều vào 2 đầu cuộn dây stator là:

$$F(s) = K_i \cdot \frac{2R^0}{R^0 + k\sqrt{s}} \cdot I(s) + K_z \cdot \frac{2R^0}{R^0 + k\sqrt{s}} \cdot Z(s) \quad (13)$$

Biến đổi tương đương (13) trở thành (14):

$$F(s) + \frac{k\sqrt{s}}{R^0} F(s) = 2K_i \cdot I(s) + 2K_z \cdot Z(s) \quad (14)$$

Biến đổi laplace ngược (14) trở thành (15):

$$F(t) + \frac{k}{R^0} \frac{d^{1/2} F(t)}{dt^{1/2}} = 2K_i \cdot i + 2K_z \cdot z \quad (15)$$

Theo định luật Newton II, phương trình động lực mô tả 1 cực từ:

$$m \frac{d^2 z}{dt^2} = F(t) - mg + f_z \quad (16)$$

Từ (16) thế F(t) vào (15) trở thành:

$$m \frac{d^2 z}{dt^2} + \frac{km}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - \frac{k}{R^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} = 2K_i \cdot i + 2K_z \cdot z - mg + f_z \quad (17)$$

Biến đổi tương đương (17) tương đương:

$$\frac{d^2 z}{dt^2} = \frac{2K_i}{m} \cdot i + \frac{2K_z}{m} \cdot z - g - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + \frac{f_z}{m} \quad (18)$$

Trong đó:

f_z là nhiễu bên ngoài.

m là khối lượng của đĩa quay (rotor).

$l_i = 2h + 2c$ là chiều dài đường từ trường của sắt từ của ổ từ hình C.

k được xác định như (8) là hệ số dòng xoáy.

$$R^0 = \frac{1}{\mu_0 A} \left(2z_0 + \frac{l_i}{\mu_r} \right) \text{ là tổng từ trở tĩnh của ổ từ hình C.}$$

$$K_i = \frac{2N^2 i_0}{\mu_0 A (R^0)^2} \text{ là hệ số tỉ lệ với dòng điện.}$$

$$K_z = \frac{1}{(R^0)^3} \left(\frac{2Ni_0}{\mu_0 A} \right)^2 \text{ là hệ số tỉ lệ với độ dịch chuyển.}$$

Xét trường hợp 1: Không có nhiễu bên ngoài khi đó (18) trở thành:

$$\frac{d^2 z}{dt^2} = \frac{2K_i}{m} \cdot i + \frac{2K_z}{m} \cdot z - g - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} \quad (19)$$

Xét trường hợp 2: Nếu nhiễu là hằng số khi đó (18) trở thành:

$$\frac{d^2 z}{dt^2} = \frac{2K_i}{m} \cdot i + \frac{2K_z}{m} \cdot z - g - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} + \frac{1}{m} f_z \quad (20)$$

Từ (18) biến đổi đối trong đương thành:

$$\frac{d^2z}{dt^2} = \frac{2K_i}{m} \cdot i + \frac{2K_z}{m} \cdot z - g - \frac{k}{R^0} \frac{d^{5/2}z}{dt^{5/2}} + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + \frac{1}{m} f_z = f(z) + g(z) \cdot i + d(z) \quad (21)$$

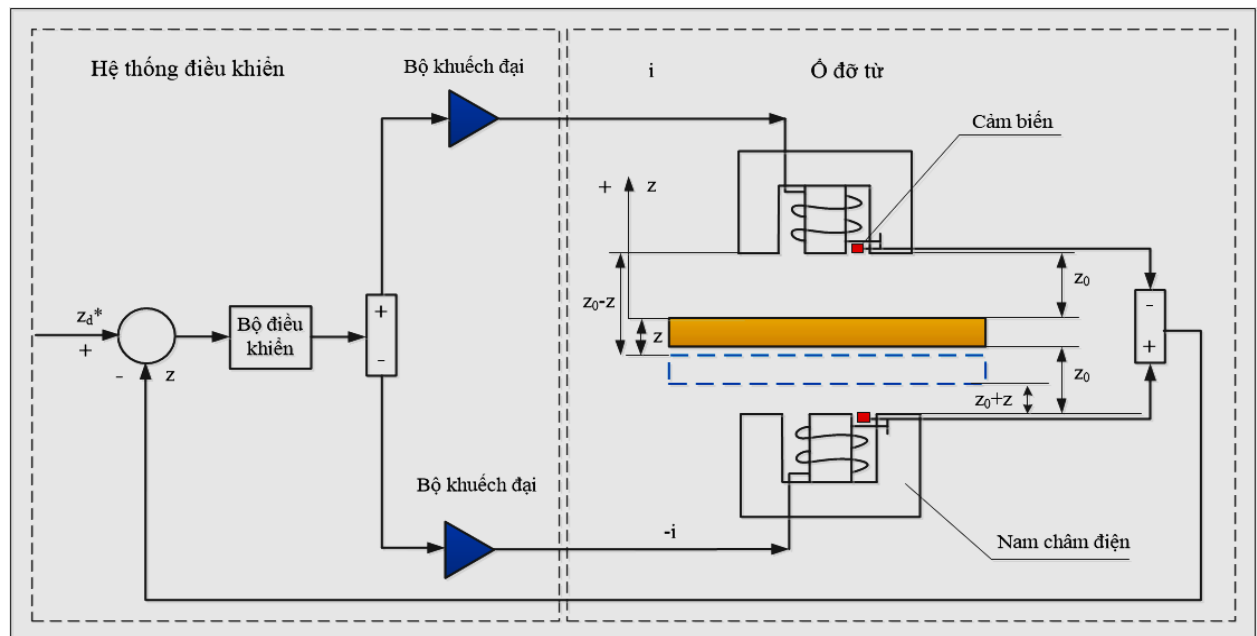
Trong đó:

$$f(z) = \frac{2K_z}{m} z - g, \quad g(z) = \frac{2K_i}{m} \quad d(z) = \frac{1}{m} f_z + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} - \frac{k}{R^0} \frac{d^{5/2}z}{dt^{5/2}}$$

Vậy (21) là động lực học tuyến tính tổng quát trên miền thời gian của 1 cực từ cấu tạo nguyên khối hoạt động khi cấp dòng xoay chiều vào 2 đầu cuộn dây stator. Trong đó, k được xác định như (8) là hệ số biểu thị sự ảnh hưởng của dòng xoáy. Khi k = 0 thì mô hình toán học trở thành trường hợp cơ cấu chấp hành của ổ từ được cấu tạo xếp lớp. Cũng trong phương trình (21) ta thấy đây là động lực học có chứa đạo hàm cấp phân số 5/2 của đối tượng bên trong và đạo hàm cấp phân số 1/2 của nhiễu bên ngoài, đây là dạng mô hình toán học tương đối mới. Kết quả mô hình toán học hình E (21) cho thấy mô hình toán học hình E cũng tương tự như hình C, nhưng hệ số tỉ lệ với dòng điện và hệ số tỉ lệ với độ dịch chuyển của ổ từ nguyên khối hình E gấp đôi trong mô hình toán học của ổ từ nguyên khối hình C.

3.1.2. Mô hình toán học một cặp cực từ

Xét nguyên lý điều khiển và hoạt động của ổ từ cấu trúc nguyên khối một cực từ thể hiện như hình 9:



Hình 9: Cấu trúc điều khiển và hoạt động ổ từ cấu trúc nguyên khối 1 cặp cực từ.

Cấu trúc hệ thống điều khiển 1 cặp cực từ được thể hiện như hình 9 gồm: một rotor (đĩa quay) được treo tự do tại một khoảng cách mong muốn z_0 so bởi 2 nam châm điện từ hình E, hai cảm biến vị trí đo độ sai lệch z giữa vị trí mong muốn z_0 với vị trí thực của rotor và đưa tín hiệu này đến bộ điều khiển. Nhiệm vụ của bộ điều khiển là cung cấp một tín hiệu điều khiển thích hợp nhằm ổn định hệ thống, duy trì vị trí của rotor tại vị trí mong muốn, tức là làm thỏa mãn sự cân bằng giữa lực hấp dẫn F_g được tạo ra với trọng lực rotor mg tại điểm làm việc và các ngoại lực f_z tác động theo phương thẳng đứng. Bộ điều khiển sẽ xử lý thông tin thu được và gửi tín hiệu điều khiển đến bộ khuếch đại công suất để tạo thành dòng điện đưa đến cuộn dây của cơ cấu điện từ, dòng điện trên 2 nam châm điện từ được điều khiển theo cơ chế đẩy hút sẽ tạo ra được lực điện từ $F_{dt}=F_1-F_2$ giúp cân bằng lực tác động lên rotor theo phương trọng trường. Về cơ bản, luật điều khiển sẽ hoạt động theo cách thức: khi rotor dịch chuyển xuống so với vị trí mong muốn, bộ điều khiển sẽ đưa ra một tín hiệu làm tăng dòng điện điều khiển một nam châm điện xa rotor hơn và giảm dòng điện trên nam châm điện còn lại, từ đó gia tăng độ lớn lực điện từ để kéo rotor trở lại vị trí danh định của nó và ngược lại. Theo (15) phương trình lực từ trên miền thời gian đối với cực từ phía trên:

$$F_1(t) + \frac{k}{R^0} \frac{d^{1/2} F_1(t)}{dt^{1/2}} = 2K_i \cdot i + 2K_z \cdot z \quad (22)$$

Đối với cực từ phía dưới do ổ đỡ từ hoạt động theo nguyên tắc đẩy kéo, cực từ phía có dòng điện mang dấu dương, cực từ phía dưới có dòng điện mang dấu âm và chiều chuyển động xuống phía dưới ngược chiều với chiều dương của trục tọa độ nên lực điện từ trên miền tần số:

$$F_2(s) = -K_i \cdot \frac{2R^0}{R^0 + k\sqrt{s}} \cdot I(s) - K_z \cdot \frac{2R^0}{R^0 + k\sqrt{s}} \cdot Z(s) \quad (23)$$

Tương tự như trên phương trình lực từ trên miền thời gian đối với cực từ phía dưới:

$$F_2(t) + \frac{k}{R^0} \frac{d^{1/2} F_2(t)}{dt^{1/2}} = -2K_i \cdot i - 2K_z \cdot z \quad (24)$$

Theo định luật Newton II, phương trình động lực mô tả cực từ:

$$m \frac{d^2 z}{dt^2} = -C \frac{dz}{dt} + F_1(t) - F_2(t) - mg + f_z \quad (25)$$

Từ (22), (24) và (25) ta suy ra phương trình:

$$m \frac{d^2 z}{dt^2} = 4K_i \cdot i + 4K_z \cdot z - C \frac{dz}{dt} - \frac{Ck}{R^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{km}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - mg + \frac{k}{R^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + f_z \quad (26)$$

Biến đổi tương đương (26) trở thành:

$$\frac{d^2 z}{dt^2} = \frac{4K_i}{m} i + \frac{4K_z}{m} z - \frac{C}{m} \frac{dz}{dt} - \frac{Ck}{mR^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - g + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + \frac{f_z}{m} \quad (27)$$

Xét trường hợp 1: Không có nhiễu bên ngoài khi đó (27) trở thành:

$$\frac{d^2 z}{dt^2} = \frac{4K_i}{m} i + \frac{4K_z}{m} z - \frac{C}{m} \frac{dz}{dt} - \frac{Ck}{mR^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - g \quad (28)$$

Xét trường hợp 2: Nếu nhiễu là hằng số khi đó (27) trở thành:

$$\frac{d^2 z}{dt^2} = \frac{4K_i}{m} i + \frac{4K_z}{m} z - \frac{C}{m} \frac{dz}{dt} - \frac{Ck}{mR^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - g + \frac{f_z}{m} \quad (29)$$

Từ (27) ta biến đổi tương đương thành:

$$\frac{d^2 z}{dt^2} = \frac{4K_i}{m} i + \frac{4K_z}{m} z - \frac{C}{m} \frac{dz}{dt} - \frac{Ck}{mR^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} - g + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + \frac{f_z}{m} = f(z) + g(z) \cdot i + d(z) \quad (30)$$

Trong đó:

C: Hệ số giảm chấn của đĩa quay

$$f(z) = \frac{4K_z}{m} z - g, \quad g(z) = \frac{4K_i}{m} i, \quad d(z) = -\frac{C}{m} \frac{dz}{dt} - \frac{Ck}{mR^0} \frac{d^{3/2} z}{dt^{3/2}} - \frac{k}{R^0} \frac{d^{5/2} z}{dt^{5/2}} + \frac{k}{mR^0} \frac{d^{1/2}(f_z)}{dt^{1/2}} + \frac{f_z}{m}$$

Vậy (30) là động lực học tuyến tính tổng quát trên miền thời gian của ổ từ dọc trục (1 cặp cực từ) cấu tạo nguyên khối hoạt động khi cấp dòng điện xoay chiều vào hai đầu cuộn dây stator.

Trong đó, k được xác định như (8) là hệ số biểu thị sự ảnh hưởng của dòng xoáy. Khi k = 0 thì mô hình toán học trở thành trường hợp cơ cấu chấp hành của ổ từ được cấu tạo xếp lớp. Cũng trong phương trình (35) ta thấy đây là động lực học có chứa đạo hàm cấp phân số 3/2, 5/2 của đối tượng bên trong và đạo hàm cấp phân số 1/2 của nhiễu bên ngoài, đây là dạng động lực học tương đối mới. So với trường hợp 1 cực từ mô hình toán học có chứa thêm thành phần đạo hàm cấp phân số 3/2 của đối tượng, hệ số tỉ lệ với dòng điện và hệ số tỉ lệ với độ dịch chuyển gấp 2 lần. Kết quả mô hình toán học hình E (30) cho thấy mô hình toán học hình E cũng tương tự như hình C, hệ số tỉ lệ với dòng điện và hệ số tỉ lệ với độ dịch chuyển của ổ từ nguyên khối hình E gấp đôi trong mô hình toán học của ổ từ nguyên khối hình C.

3.2. Mô hình phi tuyến

Do [14], [16]-[18] chỉ đưa ra mô hình toán học dạng tuyến tính trên miền tần số, mô hình toán học dạng phi tuyến tác giả chưa đưa ra. Đối với [21] chỉ mới đưa ra được động lực học dạng tuyến tính và phi tuyến trên miền thời gian đối với ổ từ hình C. Trong bài báo này tác giả trình bày phương pháp xác định mô hình toán học của ổ từ cấu tạo nguyên khối hình E ở dạng phi tuyến trên miền tần số cũng như trên miền thời gian.

Xét trường hợp một cực từ thể hiện như hình 9:

Lực từ được tính theo công thức:

$$F(\omega) = \frac{\Phi_1^2}{2\mu_0 A} = \frac{4N^2 I^2}{2\mu_0 A (R(s))^2} = \frac{2\mu_0 \mu_r^2 AN^2 I^2}{\left((l_i + 2\mu_r(z_0 - z)) + \mu_0 \mu_r A k \sqrt{s} \right)^2} \quad (31)$$

Nhận thấy trong phương trình (31), k là hệ số biểu thị ảnh hưởng của dòng xoáy. Khi k=0 thì mô hình toán học trở thành trường hợp cơ cấu chấp hành của ổ từ được cấu tạo xếp lớp.

Đặt:

$$B = 2\mu_0 \mu_r^2 AN^2 I^2, \quad E = l_i + 2\mu_r(z_0 - z) \quad (32)$$

$$D = k\mu_0 \mu_r A, \quad A = 4ab \quad (33)$$

Khi đó (31) trở thành:

$$F(s) = \frac{B}{(E + D\sqrt{s})^2} = \frac{B}{E^2 + 2ED\sqrt{s} + D^2 s} \quad (34)$$

$$\Leftrightarrow E^2 F(s) + 2ED\sqrt{s} F(s) + D^2 s F(s) = B \quad (35)$$

Biến đổi laplace ngược 2 vế của phương trình (35) tương đương (coi B, D, E là hằng số):

$$\Leftrightarrow E^2 F(t) + 2ED \frac{d^{1/2} F(t)}{dt^{1/2}} + D^2 \frac{dF(t)}{dt} = B\delta(t) \quad (36)$$

Nhận thấy phương trình (36) là phương trình có chứa đạo hàm cấp phân số của lực theo thời gian. Đây là dạng mô hình toán học mới mà chưa có công trình nào đã công bố về ổ đỡ từ. Trong đó hệ số B tỉ lệ thuận với bình phương của dòng điện, E phụ thuộc vào khoảng cách dịch chuyển so với vị trí cân bằng, D là hằng số.

Xét trường hợp một cặp cực từ thể hiện như hình 9:

Từ mô hình động lực học của một cực từ như (36) suy ra mô hình động lực của một cặp cực từ được viết như sau:

$$\frac{di_1}{dt} = -\frac{2(z_0 - z)Ri_1}{N^2 \mu_0 A} + \frac{i_1 v}{z_0 - z} + \frac{2(z_0 - z)u_1}{N^2 \mu_0 A} \quad (37)$$

$$\frac{di_2}{dt} = -\frac{2(z_0 + z)Ri_2}{N^2 \mu_0 A} + \frac{i_2 v}{z_0 + z} + \frac{2(z_0 + z)u_2}{N^2 \mu_0 A} \quad (38)$$

$$v = \frac{dz}{dt} \quad (39)$$

$$\frac{d^2 z}{dt^2} = \frac{1}{M} (F_1 - F_2 - mg + f_z) \quad (40)$$

$$\frac{d^{1/2} F_1}{dt^{1/2}} = -\frac{D}{2B_1} \frac{dF_1}{dt} - \frac{B_1}{2D} F_1 + A_1 \quad (41)$$

$$\frac{d^{1/2} F_2}{dt^{1/2}} = -\frac{D}{2B_2} \frac{dF_2}{dt} - \frac{B_2}{2D} F_2 + A_2 \quad (42)$$

Trong đó:

$$A_1 = \frac{\mu_r N^2 I_1^2 \delta}{k(l_i + 2\mu_r(z_0 - z))}, \quad B_1 = (l_i + 2\mu_r(z_0 - z))$$

$$A_2 = \frac{\mu_r N^2 I_2^2 \delta}{k(l_i + 2\mu_r(z_0 + z))}, \quad B_2 = (l_i + 2\mu_r(z_0 + z))$$

Như vậy (37)-(42) là các phương trình ở dạng miền thời gian. Đây chính là động lực học của 1 cặp cực từ hình E cấu tạo nguyên khối trên miền thời gian, trong các phương trình toán học cũng tương tự như hình C, nhưng khác nhau về hệ số. Từ đó tạo tiền đề cho các phương pháp điều khiển phi tuyến sẽ được triển khai tương lai, từ đó hoàn thiện hơn về phương pháp điều khiển ổ từ cấu tạo nguyên khối.

4. KẾT LUẬN

Bài báo này nhóm tác giả đã trình bày một phương pháp để đưa ra động lực học tuyến tính và phi tuyến của ổ từ dọc trục hình E cấu trúc nguyên khối hoạt động khi cấp dòng xoay chiều vào hai đầu cuộn dây stator. Tạo thành cơ sở toán học hoàn thiện để phân tích và thiết kế hệ thống điều khiển ổ từ dọc trục cấu tạo nguyên khối các hình dạng khác nhau. Từ đó tạo tiền đề cho các nghiên cứu thiết kế hệ thống điều khiển ổ đỡ từ cấu tạo nguyên khối để nâng cao chất lượng hoạt động của ổ từ cấu tạo nguyên khối nói chung. Các thiết kế hệ thống điều khiển đang và sẽ được nghiên cứu dựa trên các cách tiếp cận sau: a) phương pháp điều khiển trên miền tần số và trên miền thời gian (mô hình tuyến tính), b) phương pháp điều khiển phi tuyến như: Backstepping, điều khiển trượt..., điều khiển hiện đại như Mạng Nơ Ron, điều khiển thích nghi... Kết quả của các phương pháp điều khiển này sẽ được công bố trong các bài báo trong tương lai.

REFERENCES

- [1] R. Fittro and C. Knospe, “Rotor compliance minimization via mu-control of active magnetic bearings,” *IEEE Trans. Control Syst. Technol.*, vol. 10, no. 2, pp. 238–249, Mar. 2002.
- [2] M. D. Noh, S.-R. Cho, J.-H. Kyung, S.-K. Ro, and J.-K. Park, “Design and implementation of a fault-tolerant magnetic bearing system for turbomolecular vacuum pump,” *IEEE/ASME Trans. Mechatronics*, vol. 10, no. 6, pp. 626–631, Dec. 2005.
- [3] H. Hayashiya, D. Iizuka, H. Ohsaki, and E. Masada, “A novel combined lift and propulsion system for a steel plate conveyance by electromagnets,” *IEEE Trans. Magn.*, vol. 34, no. 4, pp. 2093–2095, Jul. 1998.
- [4] D. L. Trumper, M. Weng, and R. Ritter, “Magnetic suspension and vibration control of beams for non-contact processing,” in *Proc. 1999 IEEE Conf. Control Appl.*, Kona, HI, vol. 1, Aug., pp. 551–557.
- [5] P. Subrahmanyam and D. Trumper, “Active vibration isolation design for a photolithographic stepper,” presented at the 6th Int. Symp. Magn. Bearings, Cambridge, MA, Aug. 1998.
- [6] H. Gutierrez and P. Ro, “Sliding-mode control of a nonlinear-input system: Application to a magnetically levitated fast-tool servo,” *IEEE Trans. Ind. Electron.*, vol. 45, no. 6, pp. 921–927, Dec. 1998.
- [7] R. B. Zmood, D. K. Anand, and J. A. Kirk, “The influence of eddy currents on magnetic actuator performance,” *Proc. IEEE*, vol. 75, no. 2, pp. 259–260, Feb. 1987.
- [8] R.L. Stoll, *The Analysis of Eddy Currents*. London, U.K.: Oxford Univ. Press, 1974.
- [9] J. J. Feeley, “A simple dynamic model for eddy currents in a magnetic actuators,” *IEEE Trans. Magn.*, vol. 32, no. 2, pp. 453–458, Mar. 1996.
- [10] L. Zhu, C. Knospe, and E. Maslen, “An analytical model of a nonlaminated cylindrical magnetic actuator including eddy currents,” *IEEE Trans. Magn.*, vol. 41, no. 4, pp. 1248–1258, Apr. 2005.
- [11] Kucera and M. Ahrens, “A model for axial magnetic bearings including eddy currents,” in *Proc. 3rd Int. Symp. Magn. Suspension Technol.*, Tallahassee, FL, Dec.1995, pp. 421–436.
- [12] L. Zhu, C. Knospe, and E. Maslen, “Frequency domain modeling of nonlaminated C-shaped magnetic actuators,” in *Proc. 9th Int. Symp. Magn. Bearings*, Lexington, KY, Aug. 2004, pp. 1–6.
- [13] C. Bonnet and J. R. Partington, “Coprime factorizations and stability of fractional differential systems,” *Syst. Control Lett.*, vol. 41, no. 3, pp. 167–174, Oct. 2000.
- [14] Lei Zhu, “Non-laminated Magnetic Actuators: Modeling and Performance Limitations”, A Dissertation Presented to the Faculty of the School of Engineering and Applied Science University of Virginia.2005.

- [15] Joseph J. Feeley and Daniel J. Ahlstrom, “A New Eddy Current Model for Magnetic Bearing Control System Design¹”, 4th NASA Symposium on VLSI Design 1992.
- [16] L. Zhu, C. Knospe, and E. Maslen, “Frequency domain modeling of nonlaminated C-shaped magnetic actuators,” in *Proc. 9th Int. Symp. Magn. Bearings*, Lexington, KY, Aug. 2004, pp. 1–6.
- [17] Lei Zhu and Carl R. Knospe, *Senior Member, IEEE*, “Modeling of Nonlaminated Electromagnetic Suspension Systems”, *IEEE/ASME TRANSACTIONS ON MECHATRONICS*, VOL. 15, NO. 1, FEBRUARY 2010
- [18] Carl R. Knospe, *Senior Member, IEEE*, and Lei Zhu, “Performance Limitations of Non-Laminated Magnetic Suspension Systems” *IEEE TRANSACTIONS ON CONTROL SYSTEMS TECHNOLOGY*, VOL. 19, NO. 2, MARCH 2011.
- [19] Zackary W. Whitlow “*Modeling and Control of Non-laminated Active Magnetic Thrust Bearings*” In Partial Fulfillment of the requirements for the Degree Master of Science (Mechanical and Aerospace Engineering), December 2014
- [20] Henry Spece, Roger Fittro and Carl Knospe * “*Optimization of Axial Magnetic Bearing Actuators for Dynamic Performance*” Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA 22904, USA, September 2018.
- [21] Lê Ngọc Hội^{1,2}, Phạm Hữu Luân², Nguyễn Quang Địch¹, Nguyễn Tùng Lâm¹, “MÔ HÌNH HÓA Ổ TỪ CÓ CẤU TRÚC NGUYÊN KHỔP”, *Measurement, Control, and Automation*, Vol. 2 No. 2 No. 2 (2021): July-December 2021.

YSC4F.208

ỨNG DỤNG CẤU TRÚC ĐIỀU KHIỂN BS_SM MỚI VÀO HỆ TRUYỀN ĐỘNG ĐỘNG CƠ SÁU PHA CHO HỆ THỐNG ĐẨY CỦA TÀU THỦY

PHAM THUY NGOC

Industrial University of Ho Chi Minh city

ngocpham1020@gmail.com

Tóm tắt. Trong bài báo này, cấu trúc điều khiển chế Backstepping_Sliding mới (BS_SM) kết hợp bộ quan sát tốc độ thích nghi sử dụng mạng nơ-ron (VM_SC_MRAS) được đề xuất để điều khiển vector không cảm biến hệ truyền động động cơ sáu pha (SPIM) cho hệ thống đẩy của tàu thủy. Thiết kế của bộ điều khiển dựa trên cấu trúc BS và SM mới cải thiện hiệu suất và độ bền vững của bộ điều khiển. Bộ quan sát tốc độ thích nghi VM_SC_MRAS được đề xuất để ước tính tốc độ của chân vịt tàu. Sự kết hợp của bộ điều khiển BS-SM với bộ quan sát tốc độ thích nghi VM_SC_MRAS có thể bù đắp cho sự không chắc chắn gây ra bởi thay đổi thông số máy, lỗi đo và nhiễu tải, cải thiện hiệu suất động và tăng cường sự bền vững của hệ thống truyền động SPIM, hệ thống điều chỉnh chính xác tốc độ của tàu theo tốc độ tham chiếu, các đáp ứng của dòng điện và mô-men của động cơ nhanh. Các kết quả mô phỏng được thực hiện cho thấy hệ thống đề xuất cho hệ thống động lực của tàu là khả thi, tàu hoạt động chính xác, ổn định theo phương pháp điều khiển được đề xuất.

Từ khóa. Điều khiển Backstepping, điều khiển Sliding mode, hệ truyền động động cơ không đồng bộ sáu pha, điều khiển vector FOC, hệ thống đẩy của tàu thủy.

APPLICATION OF NEW BS_SM CONTROL STRUCTURE INTO SIX-PHASES INDUCTION MOTOR DRIVES FOR THE PROPULSION SYSTEM OF SHIPS

Abstract. In this paper, a (BS_SM) new Backstepping_Sliding mode control structure combined with a (VM_SC_MRAS) improved stator current MRAS based on adaptive speed observer using a neuron network and sliding mode are proposed for sensorless vector control for the propulsion system of the ship. The design of the controller is based on new BS and SM structure to improve its performance and robustness. VM_SC_MRAS improved adaptive speed observer is proposed to estimate the speed of the propeller. The combination of BS-SM controller with VM_SC_MRAS adaptive speed observer can compensate for the uncertainties caused by the machine parameter variations, measurement errors, and load disturbances. It also improves dynamic performance and enhances the robustness of the SPIM drive system, perfects the tuning of the speed reference values, and quickens the response of the motor current and torque, high accuracy of speed regulation. The simulation results lead to the conclusion that the proposed system for the propulsion system of ship is feasible. The simulation results on a test ship propeller showed that the proposed control approach operates satisfactorily.

Keywords. backstepping control, SM control, six phase induction motor drives, FOC vector control, the propulsion system of Ship.

1. INTRODUCTION

In recent years, due to the greenhouse effect and declining fossil energy reserves, and the increasing price of gasoline. there are increasing challenges to sustainable development of transport. For these reasons, research on electric propulsion systems is growing to replace main propulsion systems as well as auxiliary systems in ships in particular and another means of transportation in general. Other means include electric

vehicles and those in aeronautics. [142-154] to develop solutions for energy and environmental issues. A lot of attention is paid to the development of electric propulsion systems in marine vessels (fishing vessels) such as size, weight, noise, vibration, precision control, quality and performance. The drive systems are the main aspects of any marine propulsion system, and they are especially important for fishing vessels. In ship applications, multi-phase motor drives are becoming more and more popular than traditional three-phase machines [155-156], due to their outstanding advantages for application in systems. The propulsion systems feature fault tolerance [159-161], high torque [155], low torque [156], reduced vibration and noise [157,158], etc. Among multi-phase motors, asymmetric SPIM is one of the most popular types of multi-phase motors. However, the conventional FOC strategies using PI control cannot provide satisfactory quality for high performance drives. In order to overcome these, recently nonlinear control methods have been developed to replace PID controllers [8], [9]. The nonlinear techniques have been proposed for SPIM drive systems such as [10-17]: Regulatory linear control [10], slip model control theory (SM) [11], Backstepping control (BS) [12], Fuzzy Logic (FL) control [13] and neural network control (NN) [14], prediction control [15], passive control [16], Hamiltonian control [17], etc. Among the nonlinear control methods, BS receives a great deal of attention due to the systematic design and effective control of nonlinear responses [18]. The greatest advantage of this method is the flexibility, the ability to retain useful nonlinear elements and pursue good stability and tracking goals. However, detailed and accurate information about SPIM parameters is required when designing a traditional BS controller. So, it is difficult to use the independent BS control method to meet the requirements of high quality drive systems. In fact, BS method has been studied to combine with other control methods, such as combined with sliding mode control [19], neural network [20], fuzzy logic system (FLS) [21] to get an effective control system. In this paper, the author proposes a combined control structure where a BS controller is applied in external speed closed-loop control. The SPIM model parameters are updated for the controller to minimize the effect of changing the parameters on the efficiency of the controller. The controller design based on BS technology has also been improved, adding the tracking error integration component to improve its sustainability. The SM control technique is proposed for internal loop current control to improve the quality of control and ensure the stability, the drive system's accuracy speed response, and improve robustness against the sensitivity of changes in machine parameters and load disturbances.

On the other side, the implementation of vector control techniques requires accurate information about the speed and rotor flux. The current, voltage and speed sensors are often used to provide speed information and estimate the rotor flux. However, the recent trend eliminates the mechanical sensors used to measure speed because they are sensitive, cause noise, increase cost, size, weight and reduce reliability [22]. There has been a lot of attention from researchers to reduce the number of sensors. In this paper, the authors have focused on the accurate rotor flux estimation, and estimated the speed and resistance values of R_s and R_r by the improved stator current based on reference model adaptive system (SC_MRAS). The system makes use of a neural network to improve the performance of observer and controller for the high performance SPIM drives. In detail, in the proposed NN_SC_MRAS observer, the reference model uses the stator current components free of pure integration problems and insensitive to motor parameter variations. In this SC_MRAS scheme is, first: Adaptive Model that uses a two layer linear neural network, which is trained online by a linear LS algorithm. This algorithm requires less computation effort and overcomes some drawbacks caused by its inherent nonlinearity as in previously published literature [19]. This significantly improves the performance of the proposed observer. Second: the adaptive model based on NN is implemented in the prediction mode. This improvement ensures the proposed observer operates with better accuracy and stability. Third: the rotor flux, which is needed for the stator current estimation of the adaptive model. It is identified by the Voltage Model (VM) with the stator resistance value estimated online to enhance the performance of the proposed NN_SC_MRAS observer. In addition, using VM will avoid the instability in the regenerating mode. In this proposed system, the rotor resistance value also has been estimated based on its variation proportional to that one of stator resistance. Then the estimated resistance values were updated for the current observer to estimate the current more exactly. Finally, the modified Euler integration has been used in the adaptive model to solve the instability problems due to the discretization of the rotor equations of the machine to enhance the performance of observer. The effectiveness of these proposed schemes is verified by simulation using MATLAB/ Simulink.

The paper is organized into five sections, in section 2, the basic theory of the model of the SPIM and the SPIM drive are presented. Section 3 introduces BS_ SM controller, NN_ SC MRAS observer also is introduced in this part. Simulation and the discussion are presented in Section 4. Finally, the conclusion is provided in Section 5.

2. MODEL OF SPIM AND ELECTRIC PROPULSION SYSTEM

2.1. Model of SPIM

The drive system under study consists of an SPIM fed by a six-phase Voltage Source Inverter (VSI) and a DC link. A detailed scheme of the drive is provided in Fig.1.

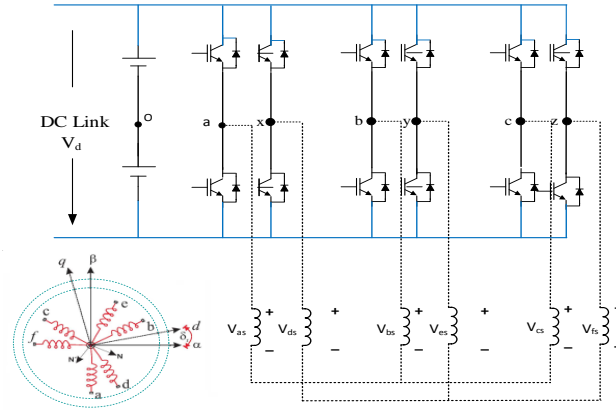


Figure. 1. A SPIM drive general diagram

By applying the Vector Space Decomposition (VSD) technique introduced in [32], the original six-dimensional space of the machine is transformed into three two-dimensional orthogonal subspaces in the stationary reference frame (D-Q), (x - y) and (z1 -z2). This transformation is obtained by means of 6x6 transformation matrix (Eq.1). In order to develop SPIM model for control purposes, some basic assumptions should be made. Hence, the windings are assumed to be sinusoidally distributed, the magnetic saturation, the mutual leakage inductances, and the core losses are neglected.

$$T_6 = \frac{1}{3} \begin{bmatrix} 1 & -\frac{1}{2} & -\frac{1}{2} & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & 0 \\ 0 & \frac{\sqrt{3}}{2} & -\frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & -\frac{1}{2} & -\frac{1}{2} & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ 0 & -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & \frac{1}{2} & \frac{1}{2} & -1 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix} \quad (1)$$

The electrical matrix equations in the stationary reference frame for the stator and the rotor may be written as

$$\begin{aligned} [V_s] &= [R_s][I_s] + p([L_{ss}][I_s] + [L_{sr}][I_r]) \\ [0] &= [R_r][I_r] + p([L_{rr}][I_r] + [L_{rs}][I_s]) \end{aligned} \quad (2)$$

where: [V], [I], [R], [L] and [Lm] are voltage, current, resistance, self and mutual inductance vectors, respectively. P is differential operator. Subscript r and s related to the rotor and stator resistance respectively. Since the rotor is squirrel cage, [Vr] is equal to zero. The electromechanical energy conversion only takes place in the DQ subsystem:

$$T_2 = \begin{bmatrix} \cos(\delta_r) & -\sin(\delta_r) \\ \sin(\delta_r) & \cos(\delta_r) \end{bmatrix} \quad (3)$$

where δ_r is the rotor angular position referred to the stator as shown in Fig. 1.

As these equations imply, the electromechanical conversion only takes place in the D-Q subspace and the other subspaces just produce losses. Therefore, the control is based on determining the applied voltage in the $\alpha\beta$ reference frame. With this transformation, the 6PIM control technique is similar to the classical three phase IM FOC. The control for the motor in the stationary reference frame is difficult, even for a three phase IM. So the transformation of SPIM model in a dq rotating reference frame to obtain currents with dc components [23] của 1 is necessary. A transformation matrix must be used to represent the stationary reference fame (α - β) in the dynamic reference (d - q). This matrix is given:

$$\begin{bmatrix} V_{s\alpha} \\ V_{s\beta} \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} R_s + PL_s & 0 & PL_m & 0 \\ 0 & R_s + PL_s & 0 & PL_m \\ PL_m & \omega_r L_m & R_r + PL_r & \omega_r L_r \\ -\omega_r L_m & PL_m & -\omega_r L_m & R_r + P \end{bmatrix} \quad (4)$$

2.2. Model of electric propulsion system

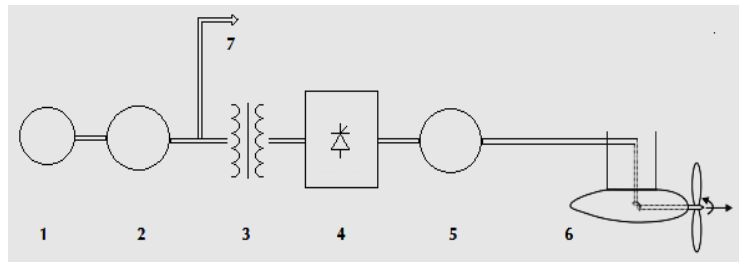


Fig. 2 The diagram of proposed drive in the ship

1. Primary engine (Diesel, electrical or Gas engine, etc.)
2. Synchronous generator
3. Power transformer
4. Six-phase drive system
5. Motor (Multi_phase induction motor)
6. Propeller
7. Other loading needs (pumps, winches, lighting, etc.)

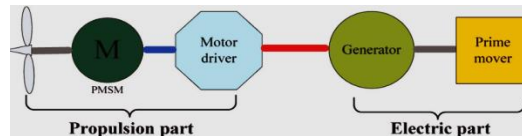


Fig. 3. Ship propulsion system using SPIM [24]

The principle of the FOC is to align the d axis of the rotating frame (d-q) with the desired flux. Therefore, the flux will be controlled by the direct component of the stator current (i_{sd}) and the torque by the

quadratic component (i_{sq}). In this case we obtain: $\psi_{rq} = 0; \psi_{rd} = \psi_{rd}$. Using Eqs. The equation of the SPIM drive in the rotary reference coordinates can be written as follows:

$$\left\{ \begin{array}{l} L_s \frac{di_{sd}}{dt} = -ai_{sd} + L_s \omega_e i_{sq} + bR_r \psi_{rd} + cu_{sd} \\ L_s \frac{di_{sq}}{dt} = -ai_{sq} + L_s \omega_e i_{sd} + b_r \omega_e \psi_{rd} + cu_{sq} \\ \frac{d\omega_r}{dt} = \frac{3}{2} P \frac{\delta \sigma L_s}{J} (\psi_{rd} i_{sq}) - \frac{T_L}{J} - B\omega_r \\ \frac{d\psi_{rd}}{dt} = \frac{L_m}{\tau_r} i_{sd} - \frac{1}{\tau_r} \psi_{rd} \end{array} \right. \quad (5)$$

where

$$\sigma = 1 - \frac{L_m^2}{L_s L_r}; \delta = \frac{L_m}{\sigma L_s L_r}; a = \frac{L_m^2 R_r + L_r^2 R_s}{\sigma L_r^2};$$

$$b = \frac{L_m^2 R_r}{\sigma L_r^2}; c = \frac{1}{\sigma}; \tau_r = \frac{L_r}{R_r}$$

The electromagnetic torque and the slip frequency can be expressed in dq reference frame:

$$T_e = \frac{3}{2} n_p \frac{L_m}{L_r} \psi_{rd} i_{sq} \quad (6)$$

$$\omega_{sl} = \frac{L_m}{L_r} \psi_{rd} i_{sq} \quad (7)$$

We also have:

$$T_e = T_L + B\omega_r + J \frac{d\omega_r}{dt} \quad (8)$$

To propel the ship, SPIM supplies torque directly to the propeller. Then the propeller creates thrust to drive the ship. Therefore, SPIM, propeller and hull together form the propulsion system in the electric propulsion system ship. SPIM's load is the hull resistance that the propeller thrust needs to overcome. The hydrodynamic properties of the hull propeller can be expressed as in [24]

$$k_1 m \frac{dv}{dt} = P(1 - T_d) - R_f \quad (9)$$

$$P = k_p \rho n^2 D_p^4$$

$$R_f = k_2 v^2$$

where k_1 is the water coefficient, m is the hull weight, v is the speed of the ship, P is the propeller thrust, T_d is the coefficient of thrust reduction, R_f is the hull resistance, k_2 is the power factor. In the hull, ρ is the sea density, D_p is the diameter of the propeller, k_p is the thrust coefficient and n is the propeller rotation speed. The load torque of propeller is:

$$T_p = k_m \rho n^2 D_p^5 = kn^2 \quad (10)$$

where, k_m is the torque factor. T_p load torque is equal to T_L (SPIM external load), k is the composite coefficient of hull resistance.

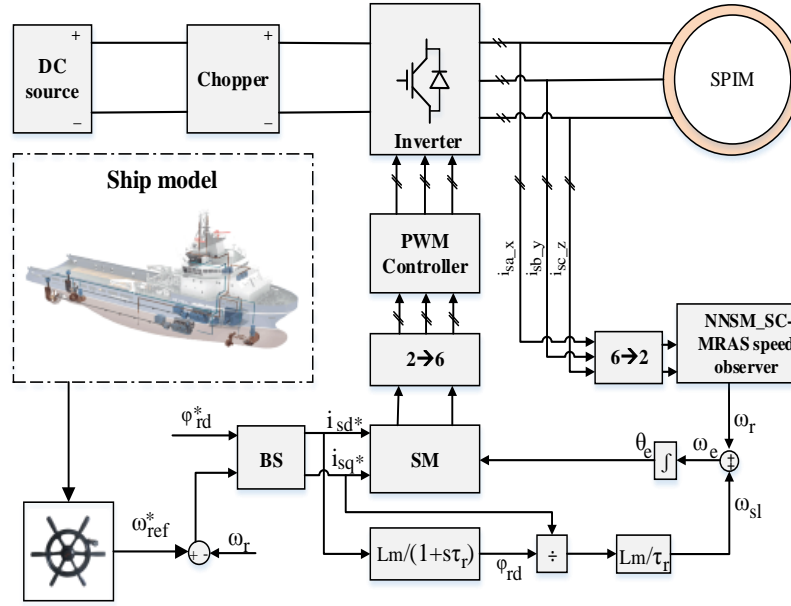


Fig. 4 FOC drive using BS_PCH combined

3. THE BS_SM CONTROLLER COMBINATES NN_SC MRAS OBSERVER FOR SENSORLESS VECTOR CONTROL OF SPIM DRIVES

3.1. The proposed BS controller for outer speed control and rotor flux loops

The purpose of this study is to design a simple control law but for high dynamic system to establish performance, eliminate load disturbance and the effect of motor parameter variations. The influence from the change of parameters and the load disturbance can be significantly reduced by adding a tracking error integration when designing the BS speed controller and updating the rotor resistance for BS control. BS techniques are a systematic and recursive method for synthesizing nonlinear control rules. The stability and performance of the subsystems is studied by Lyapunov theory. Therefore, at each step of the design, a virtual command is created to ensure the convergence of subsystems.

$$\begin{aligned} \varepsilon_{\omega} &= (\omega_r^* - \omega_r) + k'_{\omega} \int_0^t (\omega_r^* - \omega_r) dt \\ \varepsilon_{\psi} &= (\psi_{rd}^* - \psi_{rd}) + k'_{\psi} \int_0^t (\psi_{rd}^* - \psi_{rd}) dt \end{aligned} \quad (11)$$

The error dynamical equations are

$$\begin{aligned} \dot{\varepsilon}_{\omega} &= \dot{\omega}_r^* - \frac{3}{2} P \frac{\delta \sigma L_s}{J} \psi_{rd} i_{sq} + \frac{T_l}{J} + B \omega_r + k'_{\omega} (\omega_r^* - \omega_r) \\ \dot{\varepsilon}_{\psi} &= \dot{\psi}_{rd}^* + \frac{L_m}{\tau_r} i_{sd} + \frac{1}{\tau_r} \psi_{rd} + k'_{\psi} (\psi_{rd}^* - \psi_{rd}) \end{aligned} \quad (12)$$

To obtain the virtual controller of speed and rotor flux loop, the following Lyapunov function candidate is considered:

$$V_{(\omega, \psi)} = \frac{1}{2} (\varepsilon_{\omega}^2 + \varepsilon_{\psi}^2) \quad (13)$$

Differentiating V :

$$\begin{aligned} \frac{dV_{(\omega,\psi)}}{dt} &= \varepsilon_{\omega} \frac{d\varepsilon_{\omega}}{dt} + \varepsilon_{\psi} \frac{d\varepsilon_{\psi}}{dt} \\ &= \varepsilon_{\omega} \left[\frac{d\omega_r^*}{dt} - k_t \psi_{rd} i_{sq}^* + \frac{T_l}{J} + B\omega_r + k_{\omega}' (\omega_r^* - \omega_r) \right] \\ &+ \varepsilon_{\psi} \left[\frac{d\psi_{rd}^*}{dt} + \frac{L_m}{\tau_r} i_{sd}^* + \frac{1}{\tau_r} \psi_{rd} + k_{\psi}' (\psi_{rd}^* - \psi_{rd}) \right] \end{aligned} \quad (14)$$

where: $k_t = \frac{3}{2} n_p \frac{\delta\sigma L_s}{J}$, k_{ω} , k_{ψ} are positive design constants that determine the closed-loop dynamics. To $V' < 0$, the stabilizing virtual controls are chosen as

$$\begin{aligned} i_{sq}^* &= \frac{1}{k_t \psi_{rd}} \left[k_{\omega} \varepsilon_{\omega} + \frac{d\omega_r^*}{dt} + \frac{T_l}{J} + B\omega_r + k_{\omega}' (\omega_r^* - \omega_r) \right] \\ i_{sd}^* &= \frac{\tau_r}{L_m} \left[k_{\psi} \varepsilon_{\psi} + \frac{d\psi_{rd}^*}{dt} + \frac{1}{\tau_r} \psi_{rd} + k_{\psi}' (\psi_{rd}^* - \psi_{rd}) \right] \end{aligned} \quad (15)$$

We obtain:

$$\frac{dV_{(\omega,\psi)}}{dt} = -k_{\omega} \varepsilon_{\omega}^2 - k_{\psi} \varepsilon_{\psi}^2 < 0 \quad (16)$$

3.2. The inner current loop controllers using SM

In this paper, the inner current control objective is to make the measured stator currents reach the desired currents value i^* . SMC is adopted for the currents loop of SPIM drives, an adaptation law for this SMC scheme is derived based on Lyapunov theory to ensure stability and fast error dynamics.

Defining the current tracking errors as:

$$\varepsilon_{i_{sd}} = i_{sd}^* - i_{sd}; \varepsilon_{i_{sq}} = i_{sq}^* - i_{sq} \quad (17)$$

Consider s_1 and s_2 are the two sliding surfaces are defined for i_{sd} and i_{sq} , respectively. These sliding surfaces can be described as

$$\begin{aligned} s_1 &= \varepsilon_{i_{sd}} + k_1 \int_0^t \varepsilon_{i_{sd}} dt \\ s_2 &= \varepsilon_{i_{sq}} + k_2 \int_0^t \varepsilon_{i_{sq}} dt \end{aligned} \quad (18)$$

where k_1 and k_2 are the undetermined coefficient. The time derivative of Lyapunov function yields:

$$\begin{aligned} \dot{s}_1 &= \dot{\varepsilon}_{i_{sd}} + k_1 \varepsilon_{i_{sd}} \\ \dot{s}_2 &= \dot{\varepsilon}_{i_{sq}} + k_2 \varepsilon_{i_{sq}} \end{aligned} \quad (19)$$

Substituting Equation (5) to Equation (19) and combining to the sliding mode exponential approach law, we get the following equation:

$$s_2 = \varepsilon_{i_{sq}} + k_2 \varepsilon_{i_{sq}} = \dot{i}_{sq}^* + \frac{1}{L_s} [a i_{sq} - L_s \omega_s i_{sd} - b_r \omega_r \psi_{rd} - c u_{sd}] + k_2 \varepsilon_{i_{sq}} = u_{s_2} \quad (20)$$

$$s_1 = \varepsilon_{i_{sd}} + k_1 \varepsilon_{i_{sd}} = \dot{i}_{sd}^* + \frac{1}{L_s} [a i_{sd} - L_s \omega_s i_{sq} - b R_r \psi_{rd} - c u_{sd}] + k_1 \varepsilon_{i_{sd}} = u_{s_1}$$

The second order sliding controller which was proposed:

$$\begin{cases} u_{s1} &= -[\alpha_1 \text{sign}(s_1) + \beta_1 \text{sign}(ps_1)] \\ u_{s2} &= -[\alpha_2 \text{sign}(s_2) + \beta_2 \text{sign}(ps_2)] \end{cases} \quad (21)$$

where, $\alpha_1, \alpha_2, \beta_1, \beta_2$ are positive constants with: $\begin{cases} \alpha_1 > \beta_1 > 0 \\ \alpha_2 > \beta_2 > 0 \end{cases}$. The SM control law can be found using

Lyapunov theory and defining the Lyapunov function candidate:

$$V = \frac{1}{2}(s_1^2 + s_2^2) \quad (22)$$

The time derivative of Lyapunov function can be calculated as:

$$\dot{V} = s_1 \dot{s}_1 + s_2 \dot{s}_2 \quad (23)$$

According to Lyapunov theory, if the function \dot{V} is negative definite, this will ensure that the state trajectory will be driven and attracted toward the sliding surface s . Once reached, it will remain sliding on it until the origin is reached asymptotically.

$$\begin{aligned} u_{sd} &= \frac{L_s}{c} \left[\alpha_1 \text{sign}(s_1) + \beta_1 \text{sign}(ps_1) + pi_{sd}^* \right. \\ &\quad \left. + \frac{1}{L_s} (ai_{sd} - L_s \omega_s i_{sq} - bR_r \psi_{rd}) + k_1 \varepsilon_{isd} \right] \\ u_{sq} &= \frac{L_s}{c} \left[\alpha_2 \text{sign}(s_2) + \beta_2 \text{sign}(ps_2) + pi_{sq}^* \right. \\ &\quad \left. + \frac{1}{L_s} (ai_{sq} - L_s \omega_s i_{sd} - b\omega_r \psi_{rd}) + k_2 \varepsilon_{isq} \right] \end{aligned} \quad (24)$$

This ensures that:

$$\begin{aligned} \dot{V} &= -[\alpha_1 \text{sign}(s_1) + \beta_1 \text{sign}(ps_1)][s_1] \\ &\quad -[\alpha_2 \text{sign}(s_2) + \beta_2 \text{sign}(ps_2)][s_2] < 0 \end{aligned} \quad (25)$$

3.3. Structure of the VM_SC_MRAS Speed Observer

In this scheme, the measured stator current components are also used as the reference model of the MRAS observer to avoid the use of a pure integrator and reduce influence of motor parameter variation as in [25]. The adaptive model is a two-layer linear NN to estimate the stator current has been trained online by means of a least-squares algorithm. In this VM_SC_MRAS observer proposed, the adaptive model based on the ADALINE has been improved. A linear least-square algorithm, which is more suitable than a nonlinear one, like the BPN, is used to reduce the computation effort and overcome some drawbacks, caused by its inherent nonlinearity. Furthermore, the employment of the adaptive model in prediction mode leads to a quicker convergence of the algorithm, a higher bandwidth of the speed control loop, a better behavior at zero-speed, and lower speed estimation errors both in transient and steady state conditions.

Rearranging (26), the matrix equation is obtained in prediction mode; see (28). This matrix equation can be solved by any least square technique.(Eq. 32)

$$\begin{aligned} \hat{i}_{sD}(k) &= w_1 \hat{i}_{sD}(k-1) + w_2 u_{sD}(k-1) + w_3 \hat{\psi}_{rd}(k-1) + w_4 \hat{\psi}_{rq}(k-1) + w_5 \hat{i}_{sD}(k-2) - w_6 u_{sD}(k-2) \\ \hat{i}_{sQ}(k) &= w_1 \hat{i}_{sQ}(k-1) + w_2 u_{sQ}(k-1) + w_3 \hat{\psi}_{rq}(k-1) - w_4 \hat{\psi}_{rd}(k-1) + w_5 \hat{i}_{sQ}(k-2) - w_6 u_{sQ}(k-2) \\ &\quad - w_7 \hat{\psi}_{rD}(k-2) - w_8 \hat{\psi}_{rQ}(k-2) \\ &\quad - w_7 \hat{\psi}_{rQ}(k-2) + w_8 \hat{\psi}_{rD}(k-2) \end{aligned} \quad (26)$$

$$w_1 = 1 - \frac{3TR_s}{2\sigma L_s} - \frac{3TL_m^2}{2\sigma L_r L_s T_r}; w_2 = \frac{3T}{2\sigma L_s}; w_3 = -\frac{3TL_m}{2\sigma L_r L_s T_r}; w_4 = \frac{3TL_m}{2\sigma L_r L_s} \omega_r; w_5 = \frac{3TR_s}{2\sigma L_s} + \frac{TL_m^2}{2\sigma L_r L_s T_r}; w_6 = \frac{T}{2\sigma L_s} \quad (27)$$

$$Ax = b \quad (28)$$

Matrix equation (26) can be written as (28), This is a classical matrix equation of the type, where A is called a “data matrix”, b is called an “observation vector,” and ω_r is the scalar unknown. In this application a classical LS algorithm in a recursive form has been employed. This algorithm is described in detail in [26]. Fig. 5 shows the block diagram of the VM_SC_MRAS speed observer. In this paper, the authors focus on analysing OLS algorithm.

3.3.1. Rotor Speed Estimation Algorithm

$Ax \sim b$ is the linear regression problem under hand. All LS problems have been generalized by using a parameterized formulation (generalized LS) of an error function whose minimization yields the corresponding solution. This error is given by:

$$E_{LS} = \frac{(Ax - b)^T (Ax - b)}{1 - \xi + \xi X^T X} \quad (29)$$

where T represents the transpose and ξ is equal to 0.5 for TLS, 1 for DLS and 0 for OLS.

Using OLS algorithm, this error is given by:

$$E_{OLS} = (Ax - b)^T (Ax - b) \quad (30)$$

where:

$$(Ax - b) = \varepsilon = \begin{bmatrix} \varepsilon_{isD} \\ \varepsilon_{isQ} \end{bmatrix} = \begin{bmatrix} i_{sD}(k) - \hat{i}_{sD}(k) \\ i_{sQ}(k) - \hat{i}_{sQ}(k) \end{bmatrix} \quad (31)$$

This error can be minimized with a gradient descent method:

$$\omega_r(k+1) = \omega_r(k) - \eta \gamma(k) a(k) \quad (32)$$

Where

$$\gamma(k) = a(k)^T a(k) - b(k) \quad (33)$$

with η is the learning rate, a(k) is the row of A fed at instant k, and b(k) is the corresponding observation:

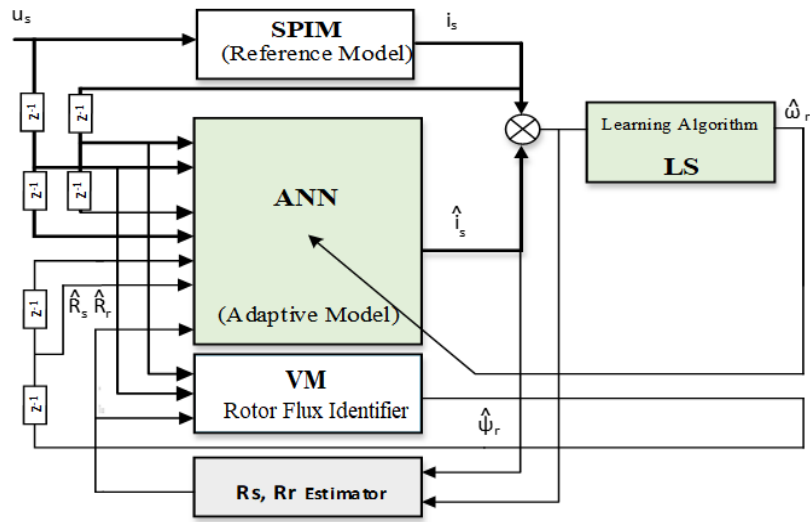


Fig 5 VM_SC_MRAS speed observer

3.3.2. Rotor Flux identifier and stability analysis

In this proposed control scheme, the rotor flux values were generated base on equation VM.

$$\begin{aligned} \frac{d\hat{\psi}_{rd}}{dt} &= \frac{L_r}{L_m}(u_{sD} - R_s i_{sD} - \sigma L_s \frac{di_{sD}}{dt}); \\ \frac{d\hat{\psi}_{rq}}{dt} &= \frac{L_r}{L_m}(u_{sQ} - R_s i_{sQ} - \sigma L_s \frac{di_{sQ}}{dt}); \end{aligned} \quad (34)$$

4. SIMULINK AND DISCUSSION

To confirm the quality of the SPIM sensorless control system for the ship propulsion system, the simulations were performed in this section. The simulation model was established by Matlab/Simulink tool. The simulations were performed based on the surveys conducted for the ship propulsion system in [24], [27], [28].

SPIM parameters: 1HP, 220V, 50 Hz, 4 pole, 1450 rpm. $R_s = 10.1\Omega$, $R_r = 9.8546\Omega$, $L_s = 0.833457$ H, $L_r = 0.830811$ H, $L_m = 0.783106$ H, $J_i = 0.0088$ kg.m². R_s is nominal value of stator resistance

In the free running ship velocity survey conducted based on [24, Figure 4-6], to evaluate the controller quality and estimate of the BS_SM method that combines VM_SCMRAS in the low and zero speed range, the low-speed reference step from initial value of 100 rpm was down to 5 rpm. As observed in Figure 6a, it is easy to see that the actual speed of the SPIM closely follows the reference speed, the estimated error and the speed tracking error are very small and nearly zero in both the established and transient mode as Fig.6b. The instant torque response according to load torque requirement 6c. The motor stator current shown in Figure 6d shows a good sine wave in a six-phase static reference frame. The $i_{s\alpha\beta}$ current response in the corresponding stator reference frame also indicates that the $i_{s\alpha}$ current is well controlled and kept stable throughout the survey. This provides the rotor flux value following the reference flux. The quality of the sensorless control system is confirmed at a very low speed range of 5 rpm. In this speed range the observer still works very accurately, the estimated speed is very close to the set speed, the estimated error does not change and is close to zero. The controller also controls the speed and torque well within this speed range, the actual speed of the motor closely follows the reference speed.

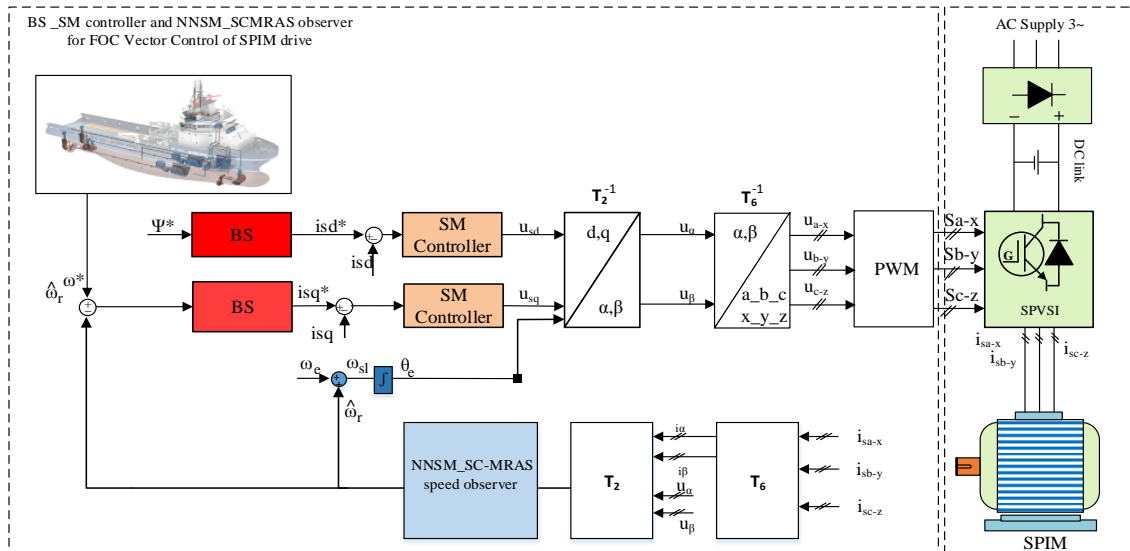


Fig. 4 Vector control of SPIM drive using RBF based on PI_SM control structure

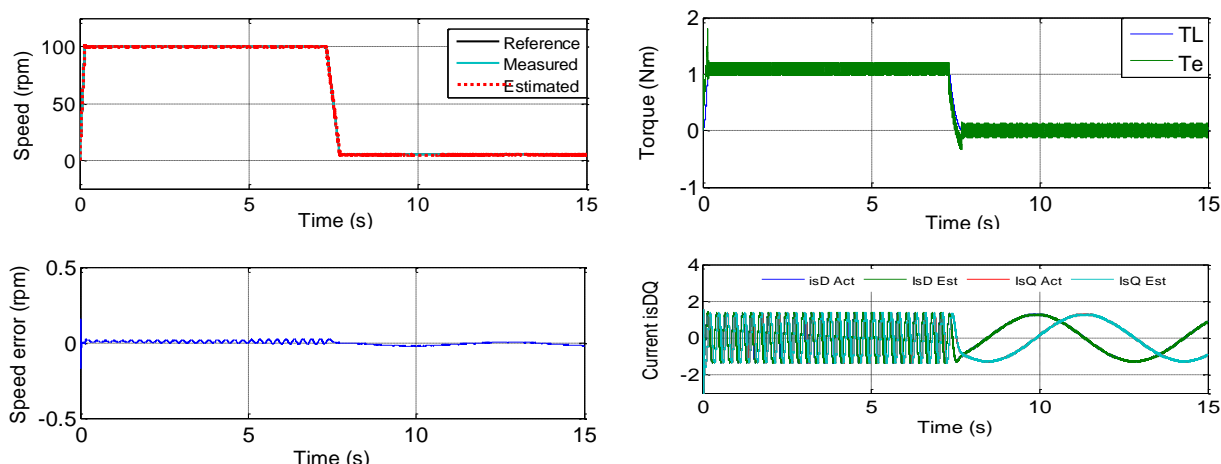


Fig. 6 Torque, stator current, speed response, estimated error in case of speed change : 100 rpm -5 rpm, [24, Figure 4-6]

Another case with the speed reference being accelerated from the initial 50 rpm to 100 rpm also was surveyed according to [27, Figure 9]. Figure 7 shows the torque, stator current, speed response in case the speed changes in the speed range: 50 rpm-100 rpm. Observing the simulation results, one can see that, when the reference speed is changed, BS_SM and VM_SCMRAS worked effectively. The speed control and estimation quality are very good. The electromagnetic torque of SPIM and the load torque of the propeller, the actual I_{sd} current and the estimated current in Figure 6 show that the current and the torque are respond fast, or instantaneous.

From the results obtained in Figure 6 and Figure 7, we can see that the BS_SM controller works effectively, controlling speed and torque well. The propeller speed can accurately track the reference speed, instantaneous current and torque response in both in transient and steady state conditions. The survey results also show that the VM_SCMRAS observation works well even when the propeller rotates at a very low speed (5 rpm). Estimated errors and small speed tracking errors approach zero in the two survey cases.

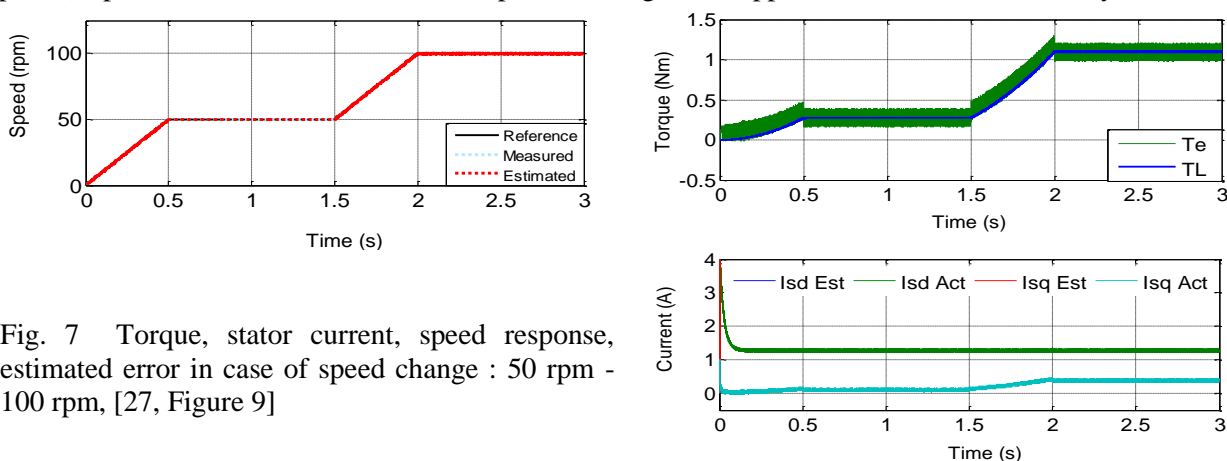


Fig. 7 Torque, stator current, speed response, estimated error in case of speed change : 50 rpm - 100 rpm, [27, Figure 9]

5. CONCLUSION

In the paper, the proposed sensorless strategy has given the ship's electric propulsion system a smaller structure and higher reliability, and a more accurate operation system. The electric propulsion systems in ships require a wide range of working speeds, especially at low speed ranges. The VM_SCMRAS adaptive speed observer and the BS_SM controller meets the requirements of ship propulsion systems because they can accurately estimate and control SPIM's thrust even at zero and low speed regions. From the research results it has been confirmed that sensorless control using BS_SM and the VM_SCMRAS speed estimator provides accurate control and estimation signals for propulsion systems. In fact, the performance of the ship is greatly influenced by the quality of the propulsion motor drive control system. The application of modern control techniques to improve the efficiency of this ship propulsion have high practical significance.

REFERENCES

- [1] C. C. Chan et al., “Electric vehicles charge forward,” *IEEE Power Energy Mag*, Vol. 2, pp. 24–33, 2004.
- [2] V. Erginer, M. H. Sarul, “High Performance and Reliable Torque Control of Permanent Magnet Synchronous Motors in Electric Vehicle Applications”, *Elektronika ir Elektrotechnika*, Vol. 19, pp. 41-46, 2013.
- [3] I. Husain, “Electric and hybrid vehicles: design fundamentals”, CRC press, 2011.
- [4] C. C. Chan, “The state of the art of electric and hybrid vehicles,” *Proc. IEEE*, Vol. 90, pp. 247–275., 2002.
- [5] A. Haddoun^{1,2}, M.E.H. Benbouzid¹, D. Diallo³, R. Abdessemed⁴, J. Ghouili⁵ and K. Srairi “Comparative Analysis of Estimation Techniques of SFOC Induction Motor for Electric Vehicles”, *Proceedings of the 2008 International Conference on Electrical Machines*, Oct, 2018.
- [6] S. Woosuk, S. Jincheol, and J. Yu-seok, "Energy-Efficient and Robust Control for High-Performance Induction Motor Drive With an Application in Electric Vehicles," *IEEE Transactions on Vehicular Technology*, Vol. 61, pp. 3394-3405, 2012.
- [7] Farid Khoucha, Khoudir Marouani, Abdelaziz Kheloui, Mohamed Benbouzid. “A Sensorless Direct Torque Control Scheme Suitable for Electric Vehicles,” *Electromotion*, Vol. 16, pp.89-97, 2009.
- [8] J. W. Finch and D. Giaouris, “Controlled AC electrical drives,” *IEEE Trans. Ind. Electron*, Vol. 55, pp. 481–491., Feb. 2008.
- [9] W. Leonhard, *Control of Electrical Drives*: Springer, 2001.
- [10] Hamou Ait Abbasa, Mohammed Belkheirib, Boubakeur Zegninia, *Feedback Linearization Control of An Induction Machine Augmented by Single Hidden Layer Neural Networks*, *International Journal of Control*, Vol. 89, pp.140-155, 2015.
- [11] Yassine Kali, Magno Ayala 2, Jorge Rodas 2, Maarouf Saad 1, Jesus Doval-Gandoy 3, Raul Gregor 2 and Khalid Benjelloun 4, “Current Control of a Six-Phase Induction Machine Drive Based on Discrete-Time Sliding Mode with Time Delay Estimation” *Energies*, Vol. 12, 2019.
- [12] Mehazzem F, Nemmour AL, Reama A, Benalla H. “Nonlinear integral backstepping control for induction motors”. In: *Proceedings of 2011 International Aegean Conference on Electrical Machines and Power Electronics and Electromotion, Joint Conference, Istanbul, Turkey, 2011*.
- [13] M.R Jovanovic, B. Bamieh, “Architecture Induced by Distributed Backstepping Design”, *IEEE Transactions on Automatic Control*, Vol. 52, pp. 108 – 113, Jan 2007.
- [14] Yassine Kali, Magno Ayala 2, Jorge Rodas 2, Maarouf Saad 1, Jesus Doval-Gandoy 3, Raul Gregor 2 and Khalid Benjelloun 4, “Current Control of a Six-Phase Induction Machine Drive Based on Discrete-Time Sliding Mode with Time Delay Estimation” *Energies*, 12(1), 2019.
- [15] Nasir Uddin M. Huang Z. R.; Siddique Hossain A. B. M., “Development and Implementation of a Simplified Self-Tuned Neuro–FuzzyBased IM Drive”. *IEEE Transactions on Industry Applications*, Vol. 50 (1), 51-59, 2014.
- [16] I. Gonzalez-Prieto, M.J. Duran, J.J. Aciego, C. Martin, F. Barrero, “Model Predictive Control of Six-phase Induction Motor Drives Using Virtual Voltage Vectors” *IEEE Transactions on Industrial Electronics*, Vol. 65(1), 2018.
- [17] Zhiping Cheng, Liucheng Jiao, Hamiltoniaian Modeling and Passivity-based Control of Permanent Magnet Linear Synchronous Motor, *Journal Of Computers*, Vol. 26, pp. 501-508, February 2013.

- [18] Abderrahmen Zaafouri, Chiheb Ben Regaya, Hechmi Ben Azza, Abdelkader Châari, “zDSP-based adaptive backstepping using the tracking errors for high-performance sensorless speed control of induction motor drive”, *ISA Transactions*, Vol. 60, pp. 333-347, 2016.
- [19] Q. Su, W. Quan, G. Cai, et al., Improved adaptive backstepping sliding mode control for generator steam valves of non-linear power systems, *IET Control Theory Appl.*, 2017, Vol. 11 , pp. 1414–1419.
- [20] 20Z. Liu, B. Chen, C. Lin, Adaptive neural backstepping for a class of switched nonlinear system without strict-feedback form, *IEEE Trans. Syst. Man Cybern. Syst.*, 2017, Vol. 47 , pp.1315–1320.
- [21] H. Li, L. Wang, H. Du, et al., Adaptive fuzzy backstepping tracking control for strict-feedback systems with input delay, *IEEE Trans. Fuzzy Syst.*, 2017, Vol. 25 , pp. 642–652.
- [22] Flah A, Novak M, Lassaad S. An Improved Reactive Power MRAS Speed Estimator With Optimization for a Hybrid Electric Vehicles Application. *J. Dyn. Syst. Meas. Control.*, 2018, Vol. 140, pp.061 016.
- [23] E. Levi, Multiphase electric machines for variable-speed applications, *IEEE Transactions on Industrial Electronics.* , 2008, Vol 55, pp. 1893 – 1909.
- [24] Zhixiong Li, Xinping Yan and Zhongxiao Peng, “Ship electric propulsion with a sensorless permanent magnet synchronous motor: A simulation study”, *Journal Engineering for the Maritime Environment*, Vol. 226, No.4, 2015.
- [25] Cirrincione G, Cirrincione M, Hérault J, Van Huffel S. The MCA EXIN neuron for the minor component analysis: Fundamentals and comparisons. *IEEE Trans. Neural Netw.*, 2002, Vol. 13, pp. 160–187.
- [26] Ngoc Thuy Pham, Khuong Huu Nguyen, “Sensorless Control for High Performance SPIM Drives Based on the Improved Rotor Flux Identifier Using Sliding Mode” *International Journal of Intelligent Engineering and Systems* (ISSN: 2185-3118) , Aug 2019, Volume 12, No.4, pp. 291-305.
- [27] [26] [173]1.Jun-Jie Ren, Yan-Cheng Liu, Ning Wang, Si-Yuan Liu, “Sensorless control of ship propulsion interior permanent magnet synchronous motor based on a new sliding mode observer”, *ISA Transactions*, Vol ,No. , 2014.
- [28] F. Temen, S. Siala and P. Noy “Multiphase Induction Motor Sensorless Control For Electric Ship Propulsion” *Second International Conference on Power Electronics, Machines and Drives (PEMD 2004)*, pp. 556-561.

YSC4F.209

THIẾT KẾ BỘ VI XỬ LÝ RISC 16 BIT DÙNG FPGA

VU THI HONG NGA¹, LE TAN KHANH², BUI QUOC THANG³

Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

¹vuthihongnga@iuh.edu.vn, ²letankhanhkk2000@gmail.com, ³Thangbuik14.090@gmail.com

Tóm tắt. Nội dung bài báo này nói về quá trình thiết kế bộ vi xử lý Reduced Instructions Set Computer (RISC) 16 bit dùng Field Programmable Gate Array (FPGA) [10]. Chúng tôi dùng ngôn ngữ mô tả phần cứng Verilog để mô tả đầy đủ phần cứng của vi mạch, việc này giúp tiết kiệm tài nguyên và thời gian thiết kế vi mạch. Cấu trúc phần cứng của bộ vi xử lý bao gồm bộ đếm chương trình (PC), bộ nhớ lệnh, tệp thanh ghi, đơn vị số học và logic (ALU), bộ nhớ dữ liệu, đơn vị điều khiển [1]. Tập lệnh của RISC gồm 15 lệnh số học và logic 16 bit đã được thiết kế, mô phỏng và kiểm chứng [5]. Mỗi lệnh được thực thi trong 1 chu kỳ máy, tốc độ biên dịch cao với tần số $f = 50\text{MHz}$. Các khối chức năng thiết kế sử dụng mô hình hành vi, được mô phỏng và kiểm chứng bằng phần mềm ModelSim [13], sau đó thực thi trên FPGA qua kit ALTERA FPGA CycloneIV EP4CE6E22C8N. Bộ vi xử lý sau khi thiết kế được kết nối với các ngoại vi để kiểm tra chức năng và hiệu suất làm việc.

Từ khóa. Bộ vi xử lý RISC, FPGA, Verilog HDL, ModelSim, ALU, PC, bộ nhớ lệnh, CycloneIV EP4CE6E22C8N.

DESIGN OF 16 BIT RISC PROCESSOR USING FPGA

Abstract. The content of this article talks about the design process of Reduced Instructions Set Computer (RISC) processors using FPGA. Using Verilog HDL, hardware description language to describe the chip's hardware, this saves time and resources in IC design. The hardware structures of a processor includes a program counter (PC), instruction memory, file register, Arithmetic-Logic Unit (ALU), data memory, and control unit. The processor instruction set of 15 16-bit logic and arithmetic instructions that has been simulated and verified. Each instruction is executed in 1 machine cycle, high compilation speed with frequency $f = 50\text{MHz}$. Function blocks are designed using behavioral models, simulated and verified by ModelSim software, and then implemented on FPGAs through the ALTERA FPGA Cyclone IV EP4CE6E22C8N kit. The designed processor connected to peripherals such as single led, 7-segment led, push button, switch, DC motors to check the function and working performance of the IC.

Keywords. RISC processor, FPGA, Verilog HDL, ModelSim, ALU, PC, instruction memory, CycloneIV EP4CE6E22C8N.

1. GIỚI THIỆU

Ngày nay, lĩnh vực thiết kế chip ngày càng phát triển và được nhiều bạn trẻ quan tâm. Tuy nhiên các bộ vi xử lý hiện nay vẫn còn nhiều hạn chế trong việc khắc phục tốc độ xử lý lệnh và hiệu suất làm việc. Vì vậy việc thiết kế thêm các bộ vi xử lý có tốc độ biên dịch cao, công suất làm việc lớn và có khả năng giao tiếp dễ dàng với các thiết bị ngoại vi khác là rất cần thiết. Bộ vi xử lý RISC với các tập lệnh rút gọn sẽ đáp ứng tốt các yêu cầu nâng cấp, cải tiến cũng như hoàn thiện hơn các chức năng của các bộ vi xử lý thông thường. Bộ vi xử lý RISC thiết kế trên công nghệ FPGA giúp tiết kiệm tốt đa chi phí sản xuất phần cứng, tốc độ làm việc cao với tần số $f = 50\text{MHz}$, có thể thay đổi phần cứng linh hoạt theo yêu cầu của người dùng [4][5]. Bộ vi xử lý RISC 16 bit sau khi được mô tả phần cứng dùng ngôn ngữ Verilog HDL, sẽ được mô phỏng và kiểm tra chức năng hoạt động trên ModelSim, sau đó sẽ nạp xuống kit ALTERA FPGA CycloneIV EP4CE6E22C8N, đồng thời được kết nối với các ngoại vi như led đơn, led 7 đoạn, nút nhấn, công tắc, động cơ DC để kiểm tra chức năng và hiệu suất làm việc của IC [6].

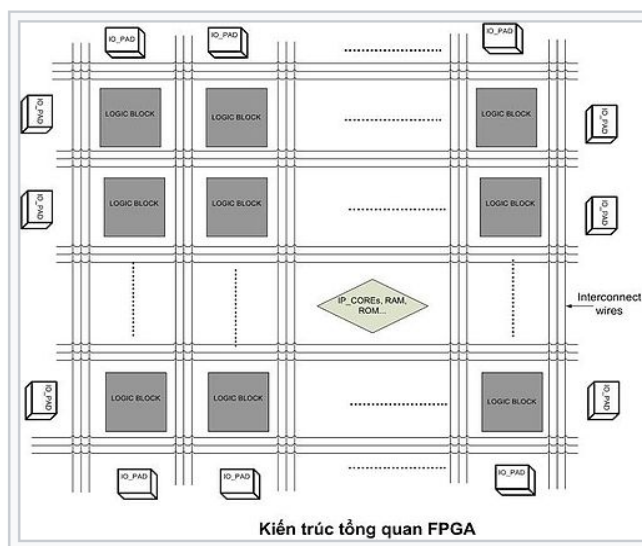
2. CƠ SỞ LÝ THUYẾT

2.1. Giới thiệu Field Programmable Gate Array (FPGA)

FPGA là một thiết bị bán dẫn hoạt động dựa trên ma trận các Configurable Logic Blocks, phần lớn chức năng điện bên trong thiết bị có thể được thay đổi bởi kỹ sư thiết kế [11].

FPGA được xác định bởi mảng các mạch con kỹ thuật số kết nối với nhau, thực hiện những chức năng chung đồng thời cung cấp mức độ linh hoạt cao. FPGA là thiết bị logic khả trình, FPGA không tự làm bất cứ điều gì nhưng nó có thể cấu hình để trở thành bất kỳ mạch kỹ thuật số nào được yêu cầu.

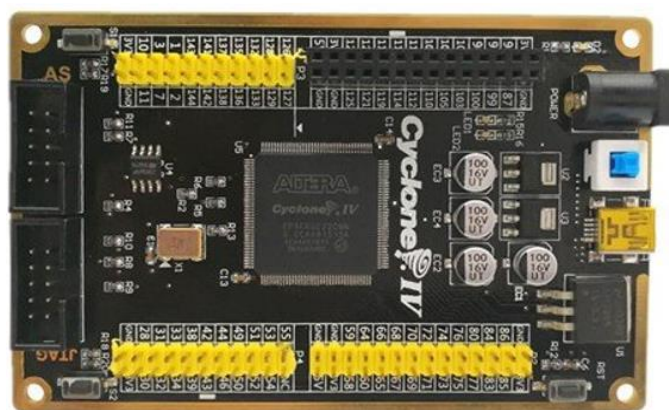
Để lập trình cho FPGA, người ta thường dùng các ngôn ngữ mô tả phần cứng như VHDL, Verilog, ...



Hình 2: Cấu trúc tổng quan của FPGA

2.2. Kit ALTERA CycloneIV EP4CE

Kit FPGA Altera Cyclone IV EP4CE [14] có tốc độ xử lý cao, tần số làm việc $f = 50$ MHz, dễ sử dụng để thiết kế các mạch số vừa và nhỏ và kết nối với các thiết bị bên ngoài.



Hình 3: Kit FPGA Altera Cyclone IV EP4CE

Thông số kỹ thuật Kit FPGA Altera Cyclone IV EP4CE:

- Nguồn: 5VDC cấp qua cổng USB hoặc giắc tròn DC.
- Chip chính: ALTERA FPGA CycloneIV thế hệ IV EP4CE6E22C8N.
- Chip cấu hình là EPCS4N hỗ trợ JTAG/AS mode (self upgraded).

- Hỗ trợ SOPC, NIOSII development (khi dùng NIOS cần thêm module SRAM).
- Thạch anh 50MHz độ chính xác cao cấp xung cho toàn hệ thống.
- Tích hợp IC ổn áp.
- 1 nút reset, cũng có thể lập trình nút này như nút nhấn (input key).
- 1 công tắc nguồn dạng nhấn giữ.
- 1 led báo nguồn.
- 2 hàng rào đôi là các chân IO của chip được đánh dấu và đặt tên rõ ràng.
- Có sẵn 2 đế nạp JTAG và AS.

2.3. Các phần mềm lập trình

Phần mềm Quartus II

Quartus II là công cụ phát triển phần mềm của Altera. Phần mềm được tích hợp và đóng gói đầy đủ giúp cho việc mô tả và tổng hợp các mạch số [11]. Các dòng FPGA của Altera gồm nhiều loại như: Cyclone, MAX, APEX, FLEX ...

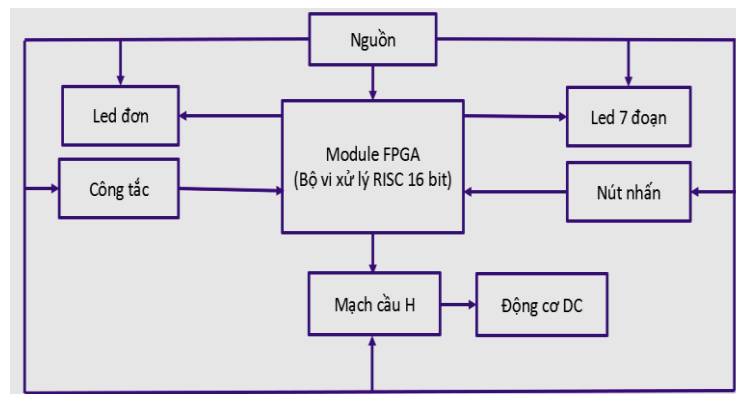
Quartus II có thể phân tích và tổng hợp các thiết kế HDL, cho phép nhà phát triển biên dịch thiết kế của họ, thực hiện phân tích thời gian, kiểm tra sơ đồ RTL, mô phỏng phản ứng của thiết kế với các kích thích khác nhau và định cấu hình thiết bị mục tiêu với lập trình viên. Quartus II bao gồm triển khai VHDL và Verilog HDL để mô tả phần cứng, chỉnh sửa trực quan các mạch logic, kiểm tra lỗi và mô phỏng.

Phần mềm ModelSim

ModelSim là chương trình mô phỏng và debug cho các thiết kế ASIC và FPGA cực kỳ mạnh mẽ và hiệu quả [13]. ModelSim thường sử dụng các ngôn ngữ như Verilog HDL, VHDL, System C. ModelSim hiện nay có các phiên bản là PE, LE, SE, XE. ModelSim PE và LE dùng cho các thiết kế nhỏ và trung bình, trong khi bản SE dùng cho các thiết kế FPGA lớn hơn.

3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1. Sơ đồ khối toàn hệ thống



Hình 4: Sơ đồ khối toàn hệ thống

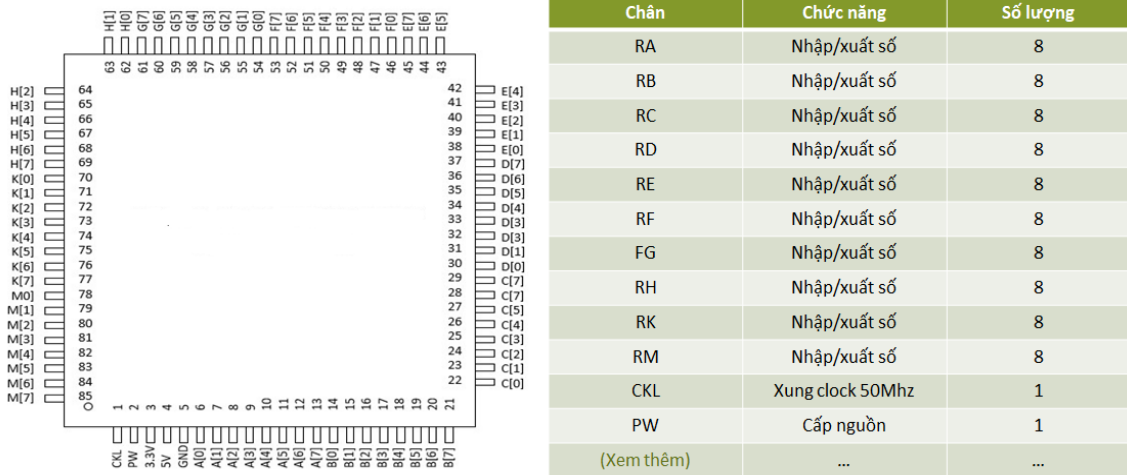
Chức năng các khối thành phần:

- Module FPGA: Tạo bộ vi xử lý RISC 16-bit [3].
- Khối led đơn: mô hình phần cứng bao gồm 8 led đơn.
- Khối led 7 đoạn: gồm 4 led 7 đoạn anode chung.
- Khối nút nhấn: mô hình phần cứng bao gồm 4 nút nhấn.
- Khối công tắc: mô hình phần cứng bao gồm 19 công tắc.
- Mạch cầu H: gồm 1 module L298N.
- Động cơ DC: động cơ DC 12V.

3.2. Thiết kế bộ vi xử lý RISC 16 bit

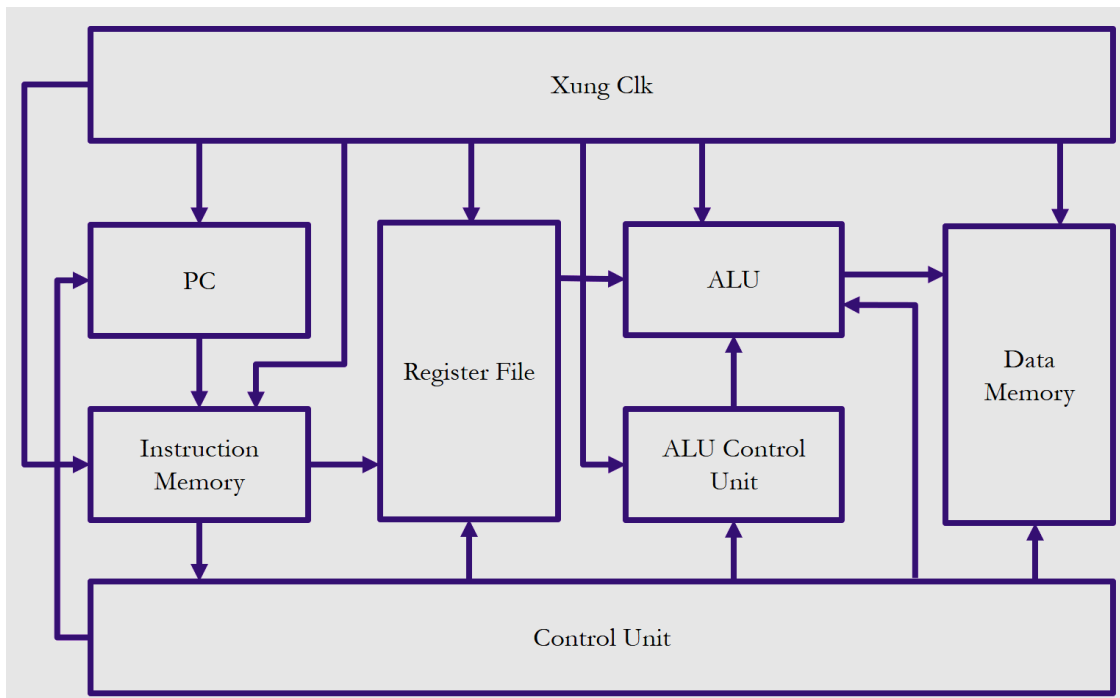
Tạo bộ vi xử lý RISC 16 bit nhúng trên FPGA, Cấu trúc phần cứng của bộ vi xử lý bao gồm bộ đếm chương trình (PC), bộ nhớ lệnh, tệp thanh ghi, đơn vị số học và logic (ALU), bộ nhớ dữ liệu, đơn vị điều khiển [6]. Tần số hoạt động là 50MHz, bộ nhớ RAM có dung lượng là 270 KB.

Sơ đồ chân bộ vi xử lý RISC 16 bit



Hình 5: Sơ đồ chân bộ vi xử lý RISC 16 bit

Sơ đồ khối của bộ vi xử lý risc 16 bit



Hình 6: Sơ đồ khối của bộ vi xử lý RISC 16 bit

Các khối chức năng của bộ vi xử lý RISC 16 bit [7]:

- Program Counter (PC): Bộ đếm chương trình.
- Instruction Memory: bộ nhớ lệnh.

- Register File: Là tập thanh ghi của bộ vi xử lý RISC 16 bit.
- ALU: Là đơn vị số học và logic, được điều khiển bởi ALU control.
- ALU Control Unit: Có chức năng điều khiển ALU.
- Data Memory: Là bộ nhớ dữ liệu của bộ vi xử lý RISC 16 bit
- Control Unit: Đơn vị điều khiển tạo ra các tín hiệu dùng để điều khiển các bộ phận còn lại của bộ vi xử lý sau khi nhận opcode trong mã lệnh của bộ nhớ lệnh.
- Xung Clk: cung cấp xung chuẩn cho FPGA hoạt động với $f = 50$ MHz.

Tập lệnh bộ vi xử lý RISC 16 bit

Tập lệnh của bộ vi xử lý RISC 16-bit gồm 15 lệnh, thuộc trong 3 nhóm lệnh chính [8][9]

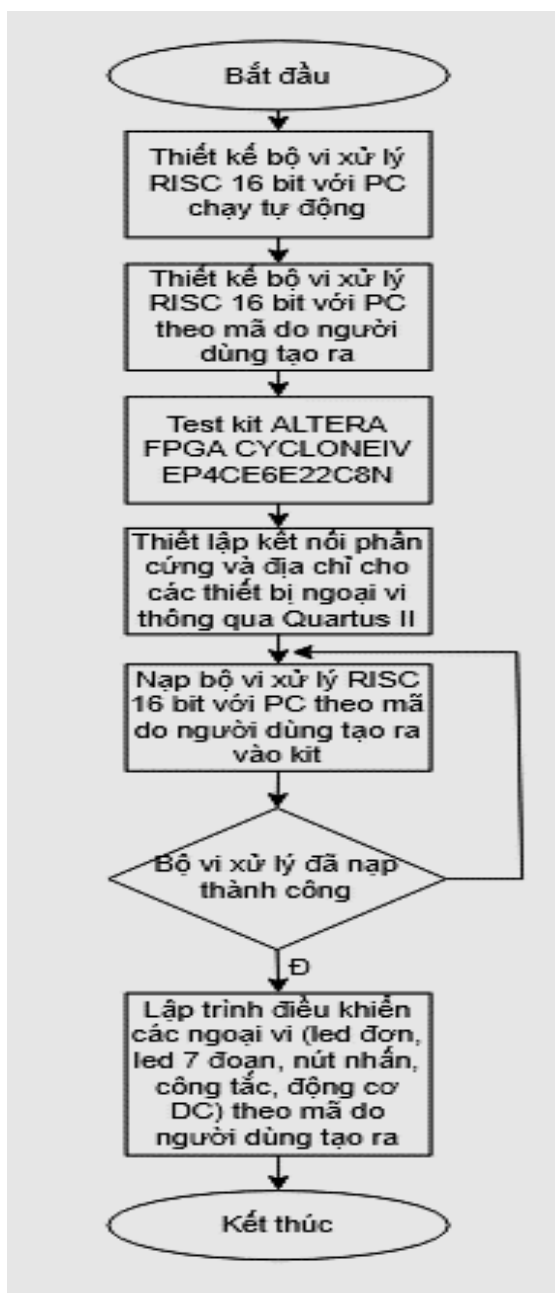
- Nhóm lệnh truy cập bộ nhớ: lệnh load word là lệnh nạp dữ liệu từ bộ nhớ vào thanh ghi, lệnh store word là lệnh lưu dữ liệu từ thanh ghi vào data memory.
- Nhóm lệnh số học và logic: lệnh cộng, lệnh trừ, lệnh dịch trái, lệnh dịch phải, lệnh NOT logic, lệnh AND logic, lệnh OR logic, lệnh so sánh hơn.
- Nhóm lệnh nhảy rẽ nhánh: lệnh rẽ nhánh khi bằng (BEQ), lệnh rẽ nhánh khi không bằng (BEN), lệnh nhảy (J).

Chức năng của các lệnh được mô phỏng và kiểm tra chức năng hoạt động bằng phần mềm Model Sim

Bảng 1: Tập lệnh của bộ vi xử lý RISC 16-bit

	Instruction	Opcode	Operation
1	Load	0000	$R0 \leftarrow \text{Mem}(R2 + 0)$
2	Load	0000	$R1 \leftarrow \text{Mem}(R2 + 0)$
3	ADD	0010	Add $R2 \leftarrow R0 + R1$
4	Store	0001	$\text{Mem}(R1 + 0) \leftarrow R2$
5	SUB	0011	$R2 \leftarrow R0 - R1$
6	INVERT	0100	$R2 \leftarrow !R0$
7	Logical Shift Left	0101	$R2 \leftarrow R0 \ll R1$
8	Logical Shift Right	0110	$R2 \leftarrow R0 \gg R1$
9	AND	0111	$R2 \leftarrow R0 \text{ AND } R1$
10	OR	1000	$R2 \leftarrow R0 \text{ OR } R1$
11	Set on Less Than	1001	$R2 \leftarrow 1 \text{ if } R0 < R1$
12	ADD	0010	$R0 \leftarrow R0 + R0$
13	BEQ	1011	Branch to jump if $R0=R1$
14	BNE	1100	Branch to jump if $R0 \neq R1$
15	J	1101	jump to the beginning address

3.3. Lưu đồ giải thuật toàn hệ thống



Hình 7: Lưu đồ giải thuật toàn hệ thống

3.4. Nguyên lí hoạt động toàn hệ thống

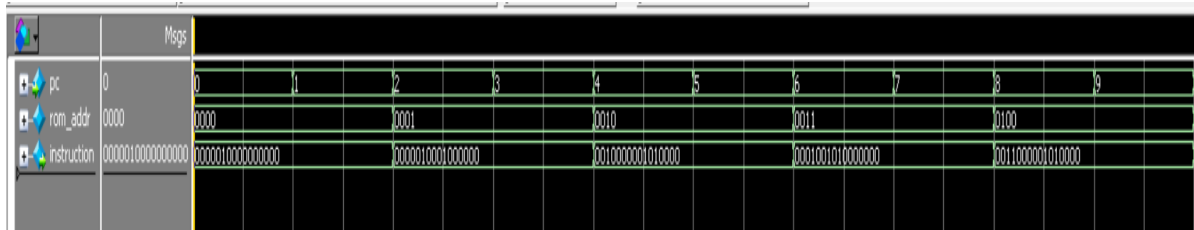
Thiết kế bộ vi xử lý RISC 16-bit bằng ngôn ngữ mô tả phần cứng Verilog trên phần mềm Quartus II, với các khối thành phần là: khối Program Counter, khối Instruction Memory, khối Register File, khối Arithmetic Logic Unit, khối Data Memory, khối Control Unit. Thiết kế tập lệnh do người lập trình tạo ra và được mô phỏng, kiểm chứng bằng phần mềm ModelSim. Bộ vi xử lý RISC 16 bit sau khi thiết kế phần cứng sẽ được nạp xuống kit ALTERA FPGA CycloneIV EP4CE6E22C8N, sau đó được lập trình kết nối các thiết bị ngoại vi là led đơn, led 7 đoạn, nút nhấn, công tắc, động cơ DC để kiểm tra tính ổn định, tốc độ biên dịch và hiệu suất làm việc của chip.

4. KẾT QUẢ ĐẠT ĐƯỢC

4.1. Kết quả mô phỏng các khối chức năng của bộ vi xử lý RISC 16 bit

Khối Instruction Memory

Kết quả mô phỏng khối Instruction Memory trên ModelSim, với mỗi lần PC tăng 2 giá trị sẽ cho ra mã lệnh tương ứng với địa chỉ.

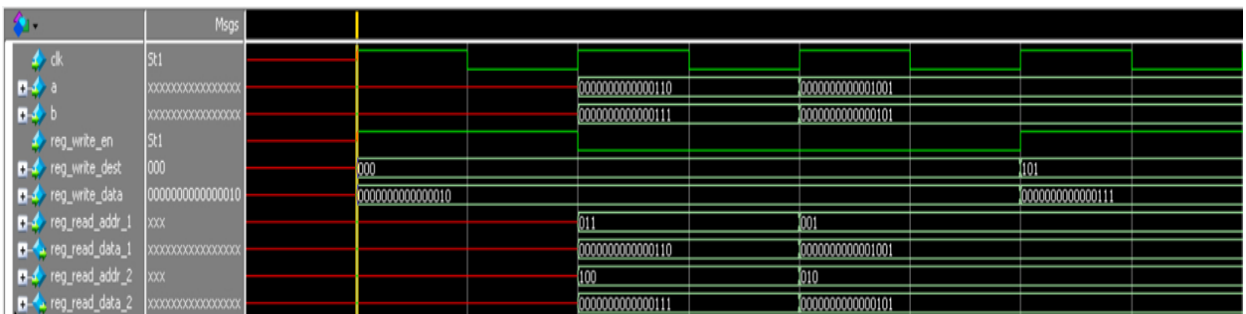


Hình 8: Kết quả mô phỏng khối Instruction Memory trên ModelSim

Khối Register File

Kết quả mô phỏng khối Register File trên ModelSim, thực hiện với dữ liệu sau:

- Ghi dữ liệu 2 vào thanh ghi R0.
- Nạp dữ liệu a = 6, b = 7 vào thanh ghi R3, R4. Rồi đọc dữ liệu trong 2 thanh ghi R3, R4.
- Nạp dữ liệu a = 9, b = 5 vào thanh ghi R1, R2. Rồi đọc dữ liệu trong 2 thanh ghi R1, R2.
- Ghi dữ liệu 7 vào thanh ghi R5.



Hình 9: Kết quả mô phỏng khối Register File trên ModelSim

Khối Data Memory (bộ nhớ dữ liệu)

Kết quả mô phỏng khối Data Memory trên ModelSim, thực hiện với các dữ liệu sau :

- Lần lượt ghi dữ liệu 3,10,12 vào thanh ghi R0, R1, R2.
- Sau đó đọc dữ liệu trong thanh ghi R0, R1, R2.



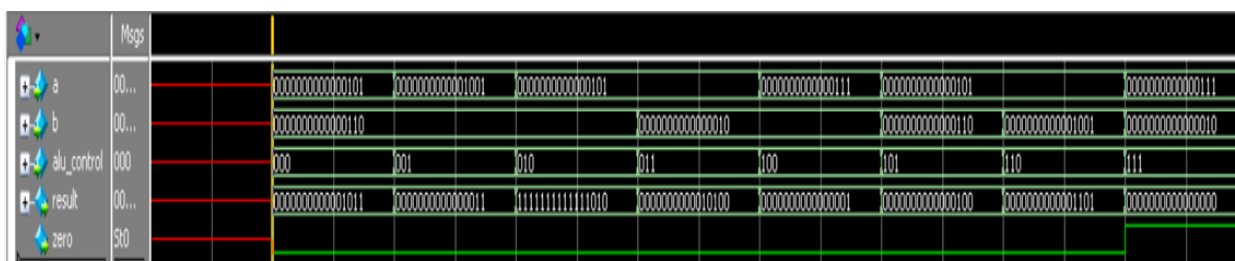
Hình 10. Kết quả mô phỏng khối Data Memory trên ModelSim

Khối Arithmetic Logic Unit

Kết quả mô phỏng khối ALU trên ModelSim, đưa lần lượt các tín hiệu alu_control để thực hiện các phép tính sau:

Bảng 2: Mô tả phép tính ALU thông qua tín hiệu alu_control

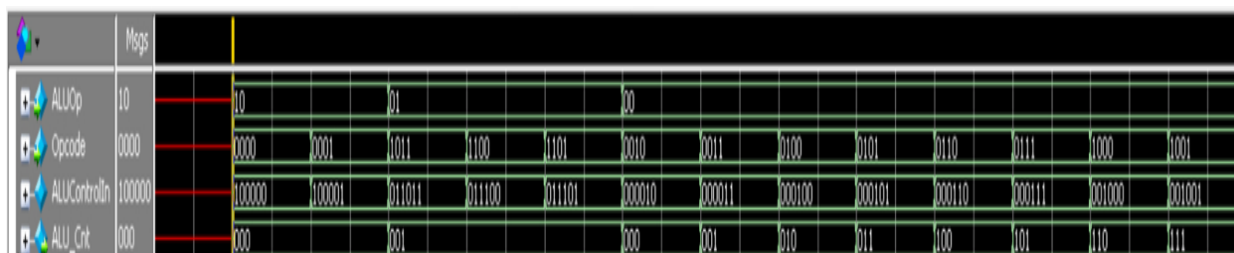
alu_control	Phép tính ALU	Mô tả
000	ADD	Cộng
001	SUB	Trừ
010	INVERT	Đảo
011	LSL	Dịch trái
100	LSR	Dịch phải
101	AND	AND logic
110	OR	OR logic
111	SLT	So sánh hơn



Hình 11: Kết quả mô phỏng khối ALU trên ModelSim

Khối ALU Control Unit

Kết quả mô phỏng khối ALU Control Unit trên ModelSim, đưa tất cả các đầu vào có thể của ALU và Opcode, kiểm tra đầu ra của mô phỏng với đầu ra chuẩn.



Hình 12. Kết quả mô phỏng khối ALU Control Unit trên ModelSim

Bảng 3: Mô tả đầu ra chuẩn của ALU Control

ALU Control				
ALUOp	Opcode	ALUcnt	Phép tính ALU	Kiểu lệnh
10	xxxx	000	ADD	LW, SW
01	xxxx	001	SUB	BEQ, BNE
00	0010	000	ADD	ADD
00	0011	001	SUB	SUB

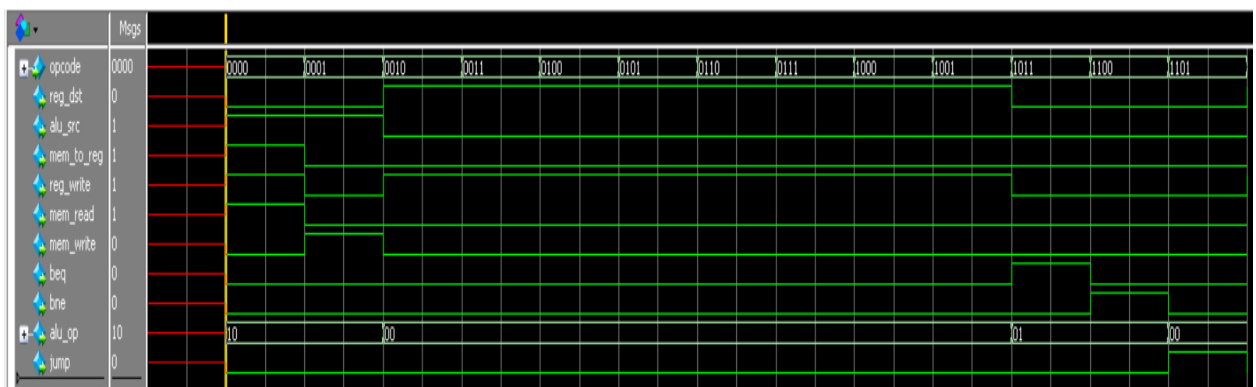
00	0100	010	INVERT	INVERT
00	0101	011	LSL	LSL
00	0110	100	LSR	LSR
00	0111	101	AND	AND
00	1000	110	OR	OR
00	1001	111	SLT	SLT

Khối Control Unit

Kết quả mô phỏng khối Control Unit trên ModelSim, đưa các đầu vào có thể của Opcode và so sánh đầu ra của mô phỏng với đầu ra chuẩn.

Bảng 4: Mô tả các đầu ra chuẩn của đơn vị điều khiển

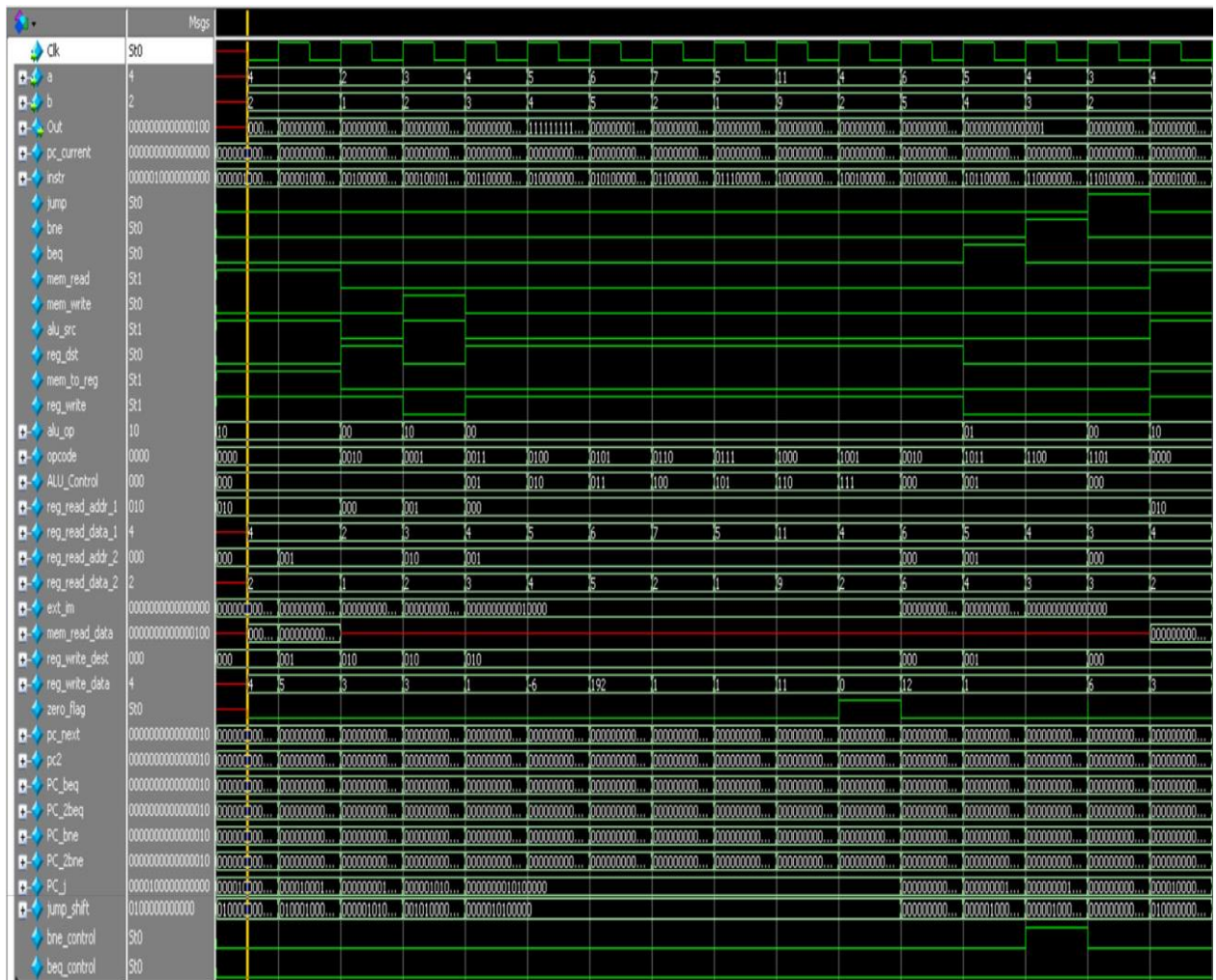
Đầu vào Opcode	Đầu ra chuẩn									
	reg_dst	alu_src	mem_to_reg	reg_write	mem_read	mem_write	Beq	Bne	alu_op	Jump
0000	0	1	1	1	1	0	0	0	10	0
0001	0	1	0	0	0	1	0	0	10	0
0010	1	0	0	1	0	0	0	0	00	1
0011	1	0	0	1	0	0	0	0	00	0
0100	1	0	0	1	0	0	0	0	00	0
0101	1	0	0	1	0	0	0	0	00	0
0110	1	0	0	1	0	0	0	0	00	0
0111	1	0	0	1	0	0	0	0	00	0
1000	1	0	0	1	0	0	0	0	00	0
1001	1	0	0	1	0	0	0	0	00	0
1011	0	0	0	0	0	0	1	0	01	0
1100	0	0	0	0	0	0	0	1	01	0
1101	0	0	0	0	0	0	0	0	00	1



Hình 13: Kết quả mô phỏng khối Control Unit trên ModelSim

4.2. Kết quả mô phỏng bộ vi xử lý RISC 16 bit

Kết quả mô phỏng bộ vi xử lý RISC 16 bit trên ModelSim [12][13].



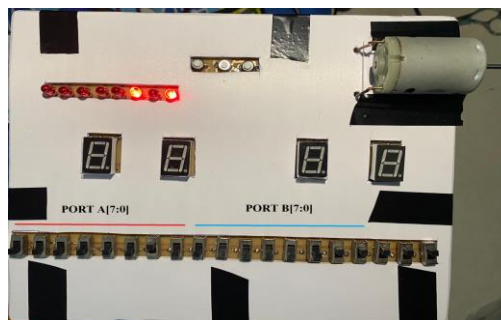
Hình 14: Kết quả mô phỏng bộ vi xử lý RISC 16 bit trên ModelSim

4.3. Kết quả kiểm tra kết nối bộ vi xử lý RISC 16 bit với các ngoại vi

Kết quả kiểm tra lệnh số học ADD:

Với PC = 16'b0000000000000100 ứng với địa chỉ là 4'b0010 sẽ cho ra lệnh ADD 2 thanh ghi.

Cho đầu vào là công tắc PORT A = 2, B = 3. Kết quả hiển thị trên led đơn là nhị phân sáng số 5 là đúng theo lệnh ADD.

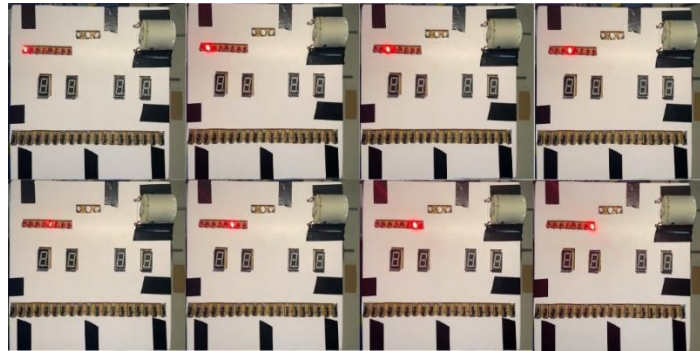


Hình 15: Kết quả thực hiện lệnh ADD trên board phần cứng

Kết quả kiểm tra lệnh Logical Shift Right

Với PC = 16'b0000000000001110 ứng với địa chỉ 4'b0111 sẽ cho ra lệnh dịch phải.

Kết quả hiển thị trên 8 led đơn, sáng đuôi 1 led đơn từ trái sang phải.



Hình 16: Kết quả thực hiện lệnh Logical Shift Right trên board phần cứng

Kết quả kiểm tra lệnh logic AND

Với PC = 16'b0000000000010000 ứng với địa chỉ là 4'b1000 sẽ cho ra lệnh AND 2 thanh ghi. Cho đầu vào là công tắc PORT A = 15, B = 13. Kết quả hiển thị trên led đơn là nhị phân sáng số 13 là đúng theo lệnh AND.

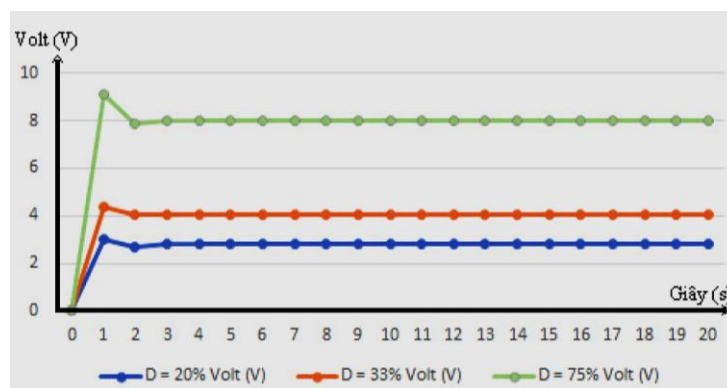


Hình 17: Kết quả thực hiện lệnh AND trên board phần cứng

Kết quả kiểm tra lệnh LOAD

Với PC = 16'b0000000000000000 ứng với địa chỉ 4'b0000 sẽ cho ra lệnh LOAD dùng để tạo xung PWM để điều khiển tốc độ động cơ DC.

Thực hiện đo điện áp ngõ ra với D = 20%, 33.3%, 75%. Để kiểm tra tính chính xác, nhóm đã thực hiện lập bảng giá trị điện áp ngõ ra trung bình qua 100 lần lấy mẫu. Kết quả quá trình điều khiển tốc độ động cơ DC với 3 tốc độ khác nhau tương ứng với chu kỳ nhiệm vụ D = 20%, 33.3%, 75%.



Hình 18: Đồ thị biểu diễn điện áp ngõ ra DC với D = 20%, 33.3%, 75%

5. KẾT LUẬN

Bài báo đã trình bày cách thiết kế bộ vi xử lý RISC 16-bit trên FPGA. Phần cứng sử dụng là kit ALTERA CycloneIV EP4CE6E22C8N và sử dụng phần mềm Quartus II, kết hợp với ModelSim. Kết quả mà bài báo đạt được là thiết kế được một bộ vi xử lý RISC 16-bit trên FPGA với các khối chức năng: bộ đếm chương trình, bộ nhớ lệnh, tệp thanh ghi, đơn vị số học và logic, bộ nhớ dữ liệu, đơn vị điều khiển. Tập lệnh của bộ vi xử lý gồm 15 lệnh số học và logic 16 bit đã được thiết kế, mô phỏng và kiểm chứng. Mỗi lệnh được thực thi trong 1 chu kỳ máy, tốc độ biên dịch cao với tần số $f = 50\text{MHz}$. Bộ vi xử lý RISC 16-bit sau khi thiết kế có tốc độ biên dịch nhanh và có thể dễ dàng kết nối với các ngoại vi để lập trình điều khiển.

TÀI LIỆU THAM KHẢO

- [1] Analysis of 16-Bit and 32-Bit RISC Processors, Animesh Kulshreshtha, Anmol Moudgil, Abhishek Chaurasia, Bharat Bhushan, 2021.
- [2] Design and Verification of 16 bit RISC Processor Using Vedic Mathematics, Ankita Yadav, Varsha Bendre, 2021.
- [3] Design And Comparison Of Risc Processors Using Different ALU Architectures, Sunita M.s, 2021.
- [4] Implementation of 16 bit RISC Processor by FSM, Manoj Barfa, Deepak Sharma, 2022.
- [5] Reduced Instruction Set Computer, Sarah L. Harris, David Harris, 2022.
- [6] MIPS Instruction Set, Prof. James L. Frankel, Harvard University, 2016.
- [7] The Design of a RISC Architecture and its Implementation with an FPGA, Niklaus Wirth, 2015.
- [8] Design and Analysis of 16-bit RISC Processor, Shraddha M. Bhagat, Sheetal U. Bhandari, 2018.
- [9] Design, Implementation and Testing of 16 bit RISC Processor, V. R. Gaikwad, 2013.
- [10] Design and Implementation of 16-Bit RISC Processor On FPGA, Mr. Rajkumar D. Komati, Aditya Kolekar, Kunal Kasbekar, Ms. Avanti Godbole, 2020.
- [11] https://en.wikipedia.org/wiki/Field-programmable_gate_array
- [12] <https://www.fpga4student.com/>
- [13] <https://www.microsemi.com/>
- [14] <https://hshop.vn/>

YSC4F.210

THIẾT KẾ VI ĐIỀU KHIỂN 32 BIT TRÊN FPGA

VU THI HONG NGA¹, DO TRUONG VI KY², DOAN BA TIEN³

Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

¹vuthihongnga@iuh.edu.vn, ²dotruongviky@gmail.com, ³doanbatien100@gmail.com

Tóm tắt. Trong bài báo này, chúng tôi tập trung vào thiết kế và thực thi vi điều khiển nhúng 32 bit bằng cách sử dụng ngôn ngữ mô tả phần cứng Verilog (VHDL). VHDL giúp chúng tôi mô tả đầy đủ phần cứng của vi mạch (IC), tối thiểu tài nguyên vật lý và giảm thời gian thiết kế [1]. Cấu trúc phần cứng của bộ vi điều khiển gồm module PIO, Interrupt, Timer, ADC, UART, PWM. Tất cả các modules của vi điều khiển nhúng 32 bit được mã hóa trong Verilog để dễ dàng mô tả và thực thi phần cứng. Các khối được thiết kế sử dụng mô hình hành vi và được mô phỏng và kiểm chứng bằng phần mềm ModelSim, sau đó thực thi trên Kit DE0-Nano Development and Education với chip Cyclone IV EP4CE22F17C6N của Altera và phần mềm Nios II.

Từ khóa. Vi điều khiển, FPGA, Verilog, HDL, ModelSim, PWM, UART, DE0-Nano, Cyclone IV EP4CE22F17C6N, Phần mềm Nios II.

DESIGNING 32 BIT MICRO CONTROLLER ON FPGA

Abstract. In this paper, we focus on 32 bits embedded micro controller design and implementation by using Verilog Hardware Description Language (VHDL). The VHDL helps us to fully describe the Integrated Circuit (IC) hardware, minimize physical resources and reduce design time. The hardware structure of the micro controller includes the PIO module, Interrupt, Timer, ADC, UART, PWM. All modules in the design codes in Verilog for easy hardware characterization, verification, simulation and implementation. The blocks design using behavioral modeling, then they are simulated and verified using ModelSim software. Finally, they execute on the DE0-Nano Education and Development Kit with Altera's Cyclone IV EP4CE22F17C6N and software Nios II.

Keywords. Microcontroller, FPGA, Verilog, HDL, ModelSim, PWM, UART, DE0-Nano, Cyclone IV EP4CE22F17C6N, Nios II software.

1. GIỚI THIỆU

Hiện nay ngành công nghệ điện tử phát triển ngày càng mạnh mẽ, đặc biệt là công nghệ Field Programmable Gate Array (FPGA). Những thiết bị phức tạp hầu hết hiện nay đều được giải quyết trên cơ sở hệ vi xử lý. Bên cạnh đó, nhờ tính linh động cao trong quá trình thiết kế, cho phép FPGA giải quyết những bài toán phức tạp nhờ mật độ tích hợp cổng logic cao. FPGA còn có thể được ứng dụng trong những bài toán đòi hỏi khối lượng tính toán lớn. Đặc biệt, với khả năng tái lập trình, người dùng có thể thay đổi các thiết kế phần cứng theo nhiều mục đích khác nhau [6].

Dựa trên những ưu điểm mạnh mẽ và tính ứng dụng thực tiễn cao của FPGA, chúng tôi đã chọn nội dung nghiên cứu “Thiết kế vi điều khiển 32 bit trên FPGA”. Vi điều khiển nhúng 32 bit được mô tả phần cứng bằng ngôn ngữ verilog, sau đó được nạp xuống chip Cyclone IV EP4CE22F17C6N. Vi điều khiển được thiết kế có đầy đủ các module PIO, Interrupt, Timer, ADC, UART, PWM. Ưu điểm của vi điều khiển thiết kế trên FPGA là tốc độ biên dịch cao, dễ dàng kết nối với các ngoại vi, thay đổi các cấu hình phần cứng do có thể tái lập trình được nhiều lần, dễ bảo trì, sử dụng thuận tiện cho việc nghiên cứu và chế tạo, phục vụ cho dự án thiết kế số vừa và nhỏ.

2. CƠ SỞ LÝ THUYẾT

2.1. Giới thiệu Field Programmable Gate Array

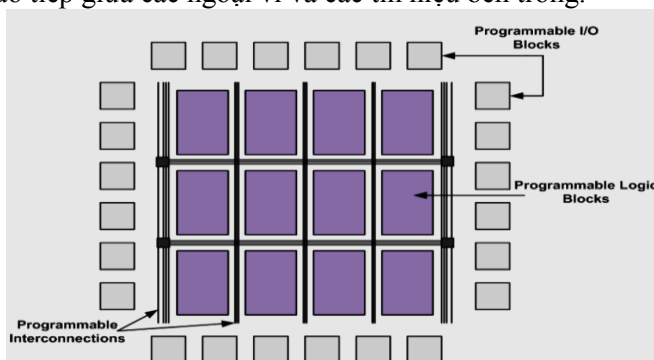
Field Programmable Gate Array (FPGA) là một thiết bị bán dẫn hoạt động dựa trên ma trận các khối logic có thể tái cấu hình (Configurable Logic Blocks), theo đó phần lớn chức năng điện bên trong thiết bị có thể được thay đổi bởi kỹ sư thiết kế [8].

FPGA được xác định bởi mảng các mạch con kỹ thuật số kết nối với nhau, thực hiện những chức năng chung đồng thời cung cấp mức độ linh hoạt cao. FPGA thuộc về một loại thiết bị được gọi là logic khả trình. FPGA không tự làm bất cứ điều gì, thay vào đó nó có thể cấu hình để trở thành bất kỳ mạch kỹ thuật số nào được yêu cầu.

Thiết kế hay lập trình cho FPGA được thực hiện chủ yếu bằng các ngôn ngữ mô tả phần cứng như VHDL, Verilog, AHDL.

Cấu trúc cơ bản của FPGA bao gồm 3 thành phần chính:

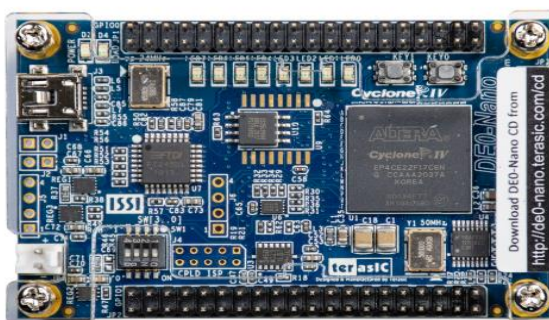
- Khối logic có thể tái cấu hình (CLB) thực hiện các chức năng logic.
- Các kết nối bên trong (Programmable Interconnect) có thể lập trình để kết nối các đầu vào và đầu ra của các CLB và các khối I/O bên trong.
- Các khối I/O cung cấp giao tiếp giữa các ngoại vi và các tín hiệu bên trong.



Hình 1: Mô hình cấu trúc cơ bản của FPGA

2.2. Kit De0 Nano Development and Education

Kit De0-Nano Development and Education là bo mạch sử dụng chip ALTERA FPGA Cyclone IV EP4CE22F17C6N [5]. Kit De0-Nano có kích thước nhỏ gọn phù hợp cho các thiết kế số vừa và nhỏ, thường được sử dụng cho mục đích giáo dục, phục vụ các nghiên cứu ứng dụng trên FPGA.



Hình 2: Kit De0 Nano Development And Education

Thông số kỹ thuật của Kit De0 Nano Development and Education:

- Chip chính: Altera FPGA Cyclone IV EP4CE22F17C6N. Có 22,320 cổng logic, 594Kbit bộ nhớ nhúng, 153 chân I/O FPGA tối đa.

- Nguồn: Cổng USB loại mini-AB (5V). Hai chân DC 5V của đầu cắm GPIO (5V). Đầu cắm nguồn bên ngoài 2 chân (3.6-5.7V).
- Bộ nhớ: 32MB SDRAM, 2Kb I2C EEPROM.
- Hỗ trợ SOPC, NIOSII development (khi dùng NIOS cần thêm module SRAM).
- Thạch anh 50MHz độ chính xác cao cung cấp xung cho toàn hệ thống.
- Bộ chuyển đổi ADC: Bộ chuyển đổi ADC128S022, 8 kênh, 12 bit.
- Tích hợp sẵn 8 led đơn, 2 nút nhấn và 4 công tắc và các cổng GPIO.
- Kit De0 Nano Development and Education có 22,320 Logic elements.

2.3. Các phần mềm sử dụng

2.3.1. Phần mềm Quartus II

Quartus II là công cụ phát triển phần mềm của Altera cung cấp môi trường thiết kế hệ thống trên chip có thể lập trình (SOPC) toàn diện. Là một phần mềm thiết kế thiết bị logic khả trình.

Quartus II có thể phân tích và tổng hợp các thiết kế HDL, cho phép nhà phát triển biên dịch thiết kế của họ, thực hiện phân tích thời gian, kiểm tra sơ đồ RTL, mô phỏng phản ứng của thiết kế với các kích thích khác nhau và định cấu hình thiết bị mục tiêu với lập trình viên. Quartus II bao gồm triển khai VHDL và Verilog để mô tả phần cứng, chỉnh sửa trực quan các mạch logic và mô phỏng dạng sóng [8].

2.3.2. Phần mềm ModelSim

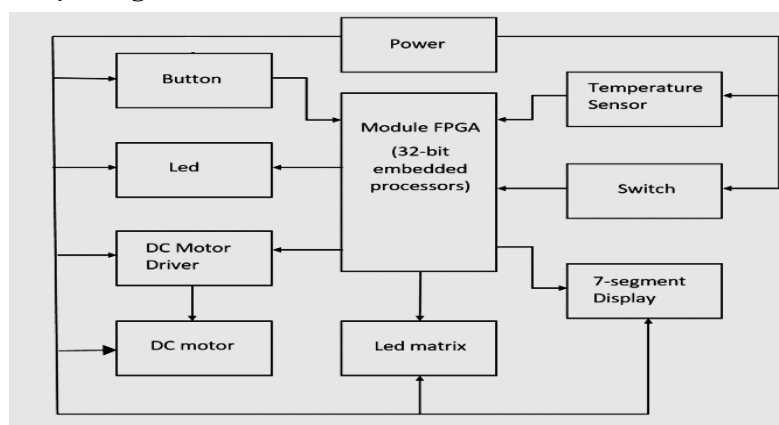
ModelSim là một chương trình mô phỏng và gỡ lỗi thiết kế FPGA cực kỳ mạnh mẽ và hiệu quả. Khi thiết kế FPGA, chúng ta có thể sử dụng ModelSim để mô phỏng, kiểm tra, chạy thử chương trình. ModelSim cho phép debug các ngôn ngữ như Verilog HDL, VHDL, System C. Phần mềm này có các phiên bản như là PE, LE, SE, XE. ModelSim PE và LE sử dụng cho các cấu trúc thiết kế nhỏ và trung bình, còn phiên bản SE dùng cho các thiết kế FPGA lớn hơn [11].

2.3.3. Công cụ xây dựng phần mềm Nios II

Bộ công cụ phần mềm Nios II EDS chuyên dùng để viết phần mềm chạy trên bộ vi xử lý Nios II. Nios II Software Build Tools for Eclipse cung cấp nhiều thuộc tính và công cụ mã nguồn mở nhằm tạo ra chương trình phần mềm Nios II. Các tính năng bao gồm như sau: hỗ trợ thư viện Newlib C, trình biên dịch cho C và C++, trình chỉnh sửa gói hỗ trợ bo mạch Nios II, bộ lập trình Intel Quartus Prime, trình điều hướng nguồn, trình chỉnh sửa và trình gỡ lỗi [2][4].

3. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

3.1. Sơ đồ khối toàn hệ thống



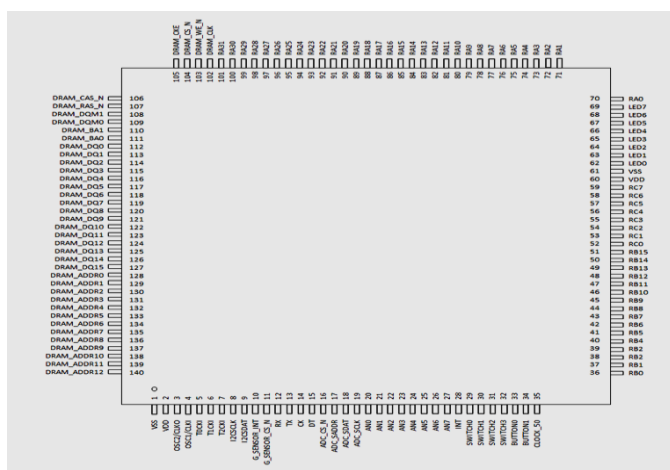
Hình 3: Sơ đồ khối toàn hệ thống

Chức năng các khối thành phần [7]:

- Module FPGA: Sử dụng chip Altera FPGA Cyclone IV EP4CE22F17C6N, đây được coi là bộ điều khiển của hệ thống, nhận dữ liệu từ nút nhấn và từ cảm biến.
- Khối switch: Dùng để điều khiển các thiết bị ngoại vi.
- Khối temperature sensor: Cảm biến nhiệt độ LM35 đo nhiệt độ từ môi trường, sau đó tín hiệu analog được đưa vào hệ thống xử lý.
- Khối button: Dùng để điều khiển các thiết bị ngoại vi.
- Khối DC motor: Động cơ DC.
- Khối DC Motor Driver: Nhận tín hiệu điều khiển từ bộ PWM trên FPGA để điều chỉnh tốc độ động cơ.

3.2. Sơ đồ chân vi điều khiển nhúng 32 bit

Tạo chip nhúng vi điều khiển 32 bit trên FPGA, chip có các module chức năng là: PIO, ADC, Timer, UART, PWM. Tần số hoạt động là 50MHz, bộ nhớ RAM được thiết kế là 64K bytes. Các PIO có số input/output linh hoạt theo yêu cầu của người sử dụng. Ngoài ra ta có thể tạo thêm các khối chức năng khác bằng cách thay đổi cấu hình phần cứng thì thiết lập trên Qsys.



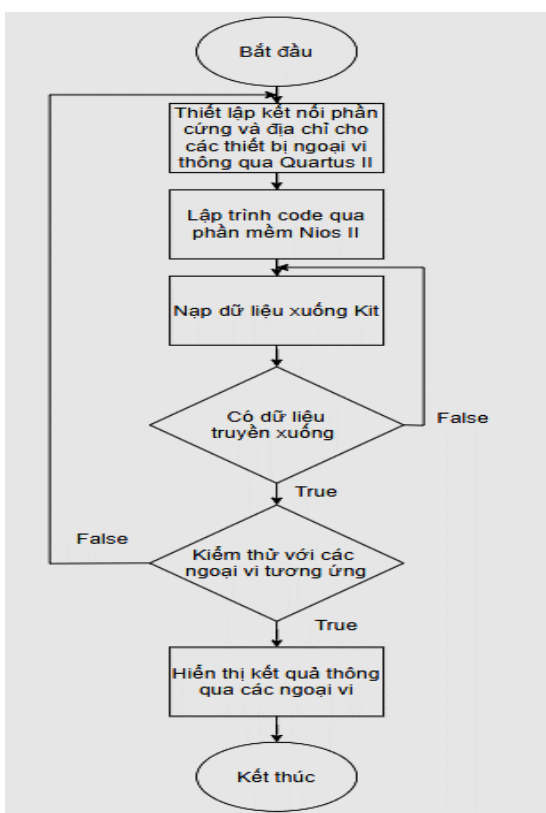
Hình 4: Sơ đồ chân vi điều khiển nhúng 32 bit

Các khối chức năng của chip nhúng vi điều khiển 32 bit:

- Module PIO có 5 PORT: PORTA: 32 chân, PORTB: 16 chân, PORTC, PORTE: 8 chân, PORTF: 11 chân.
- Module PWM: Bao gồm 2 bộ điều khiển độ rộng xung 32 bit.
- Module ADC: gồm 8 kênh tương tự, độ phân giải 12 bit (tương đương giá trị từ 0-4095) và thời gian lấy mẫu 50Kps – 200Kps.
- Module UART: truyền và nhận dữ liệu nối tiếp với tần số 50MHz.
- Module Timer: gồm 3 bộ định thời 32 bit (Timer0, Timer1, Timer2), dùng để định thời hoặc đếm.
- Module ngắt: Bao gồm 2 lớp ngắt là lớp ngắt cơ bản và lớp ngắt ngoại vi. Các lớp ngắt cơ bản bao gồm ngắt tràn Timer0, ngắt ngoài... Lớp ngắt ngoại vi bao gồm các ngắt ngoại vi như ngắt tràn Timer1, Timer2, ngắt hoàn thành việc chuyển đổi ADC...

3.3. Nguyên lí hoạt động toàn hệ thống

Lưu đồ hoạt động của hệ thống



Hình 5: Lưu đồ hoạt động của hệ thống

Nguyên lí hoạt động của hệ thống

Tạo chip nhúng vi điều khiển 32 bit bằng công cụ Qsys trên Quartus II. Chip nhúng có các module chức năng: PIO, ADC, Timer, UART, PWM. Tần số hoạt động là 50MHz, bộ nhớ RAM được thiết kế là 64K bytes. Các ngoại vi như LED, LED 7 đoạn, nút nhấn, led ma trận, động cơ DC sẽ nhận tín hiệu từ vi điều khiển để lập trình ứng dụng nhằm kiểm tra chức năng của các module.

Các thiết bị ngoại vi được kết nối với vi điều khiển thông qua 5 PORTA,B,C,D,E. Lập trình điều khiển nút nhấn, công tắc, led để kiểm tra chức năng của module PIO, ngắt, TIMER. Vi điều khiển xuất xung từ module PWM để điều khiển tốc độ động cơ DC. Vi điều khiển lấy thông số nhiệt độ từ môi trường thông qua cảm biến nhiệt độ LM35 và tiến hành chuyển đổi tín hiệu analog sang tín hiệu số thông qua khối ADC và hiển thị kết quả đo nhiệt độ trên led 7 đoạn.

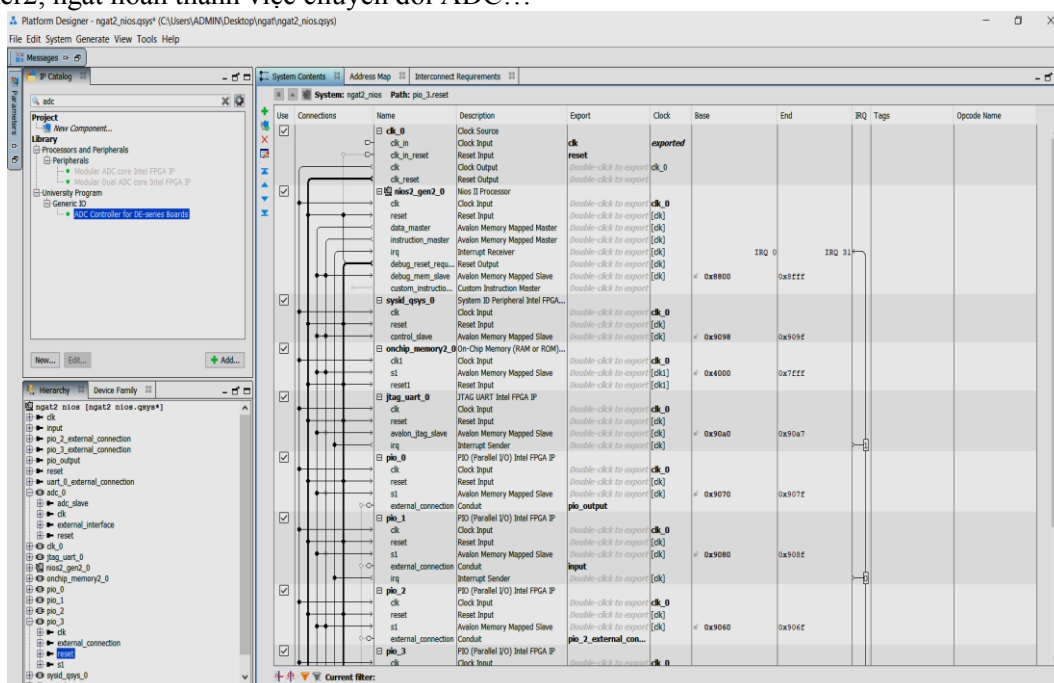
4. KẾT QUẢ ĐẠT ĐƯỢC

4.1. Lập trình phần cứng trên Qsys

Các khối chức năng của chip nhúng vi điều khiển 32 bit [10]:

- CPU NiosII là: Nios2_gen2_0
- Module PIO có 5 PORT: PORTA: 32 chân, PORTB: 16 chân, PORTC, PORTE: 8 chân, PORTF: 11 chân.
- Module PWM: Bao gồm 2 bộ điều khiển độ rộng xung 32 bit.
- Module ADC: gồm 8 kênh tương tự, độ phân giải 12 bit (tương đương giá trị từ 0-4095) và thời gian lấy mẫu 50Kps – 200Kps.
- Module UART: truyền và nhận dữ liệu nối tiếp với tần số 50MHz.
- Module Timer: gồm 3 bộ định thời 32 bit (Timer0, Timer1, Timer2), dùng để định thời hoặc đếm.

- Module ngắt: Bao gồm 2 lớp ngắt là lớp ngắt cơ bản và lớp ngắt ngoại vi. Các lớp ngắt cơ bản bao gồm ngắt tràn Timer0, ngắt ngoại... Lớp ngắt ngoại vi bao gồm các ngắt ngoại vi như ngắt tràn Timer1, Timer2, ngắt hoàn thành việc chuyển đổi ADC...



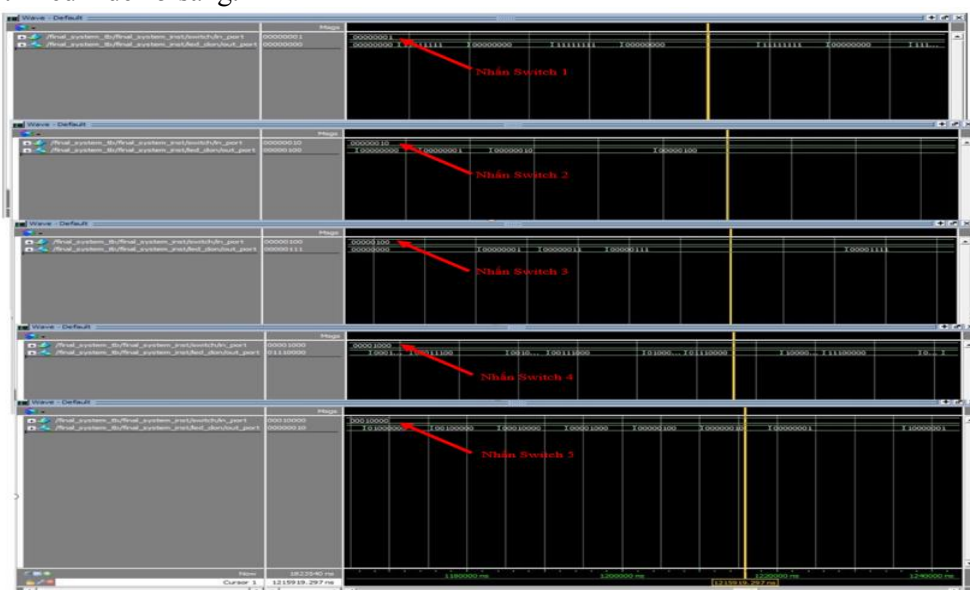
Hình 6: Kết quả thiết kế chip nhúng 32 trên Qsys

Sau khi lập trình và biên dịch đã tạo ra chip nhúng vi điều khiển 32 bit bằng công cụ Qsys trên Quartus II. Gồm có các module chức năng: PIO, ADC, Timer, UART, PWM, số logic elements của Kit DE0-Nano Development and Education đã dùng là 1926 logic cells trên tổng số 3303 logic cells trên phần cứng.

4.1. Kiểm tra chức năng của vi điều khiển nhúng 32 bit

4.1.1. Kết quả kiểm tra chức năng module PIO

Kết quả thực hiện kết nối module PORTE với 8 LED đơn, PORTA với 5 công tắc 1 đến 5: Khi nhấn công tắc 1 đến 5 thì led 1 đến 5 sáng.



Hình 7: Kết quả kiểm tra chức năng module PIO

Kết quả thực hiện kết nối module PIO với các LED matrix, công tắc. Khi gạt công tắc 6 thì led matrix hiển thị hình trái tim sáng tắt, sau 3s thì đổi màu xanh, đỏ, cam.



Hình 8: Kết quả kiểm tra chức năng module PIO (tt)

4.1.2. Kết quả kiểm tra chức năng module TIMER

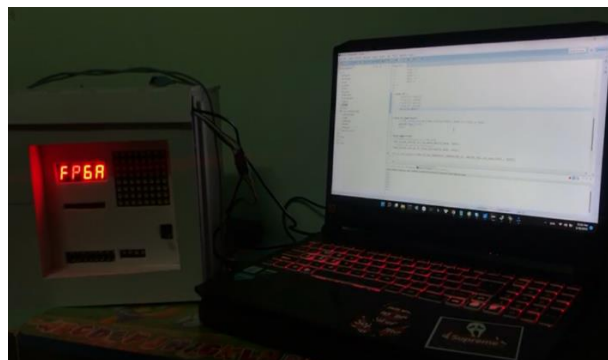
Kết quả thực hiện kết nối TIMER0, PIO, Led bảy đoạn. Lập trình điều khiển led bảy đoạn hiển thị các trị đếm số từ 0-99, cứ sau 5s sẽ thực hiện chương trình ngắt sẽ hiển thị chữ FPGA lên LED 7 đoạn.



Hình 9: Kết quả kiểm tra chức năng TIMER

4.1.3. Kết quả kiểm tra chức năng module INTERRUPT

Kết quả kiểm tra chức năng ngắt ngoài. Chương trình chính hiển thị chữ FPGA trên led bảy đoạn, khi nhấn nút 1 thì chuyển sang chương trình ngắt, hiển thị “IUH” trong 1s và quay lại chương trình chính.



Hình 10: Kết quả kiểm tra chức năng module INTERRUPT

4.1.4. Kết quả kiểm tra chức năng module ADC

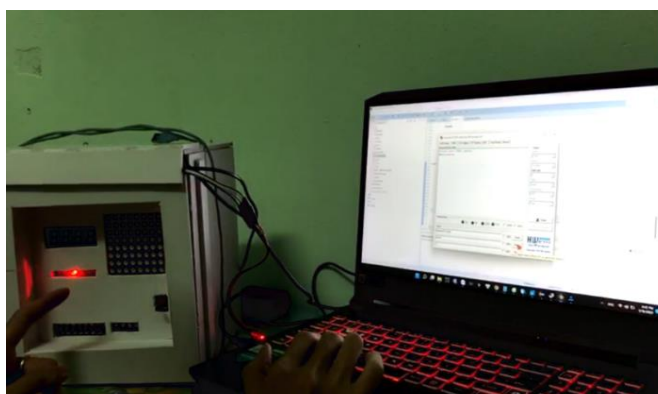
Kết quả kiểm tra chức năng module ADC. Sử dụng cảm biến nhiệt độ LM35 để đọc nhiệt độ môi trường và tiến hành chuyển đổi tín hiệu analog sang tín hiệu số thông qua module ADC và hiện thị nhiệt độ đo được ra led bảy đoạn.



Hình 11: Kết quả kiểm tra chức năng module ADC

4.1.5. Kết quả kiểm tra chức năng module UART

Kết quả kiểm tra chức năng module UART, truyền và nhận dữ liệu từ kit đến PC và từ PC xuống kit. Truyền chữ hello từ kit sang máy tính, nhận chuỗi kí tự từ máy tính sang kit nếu đúng thì led sáng đuôi nếu sai thì led chớp tắt.



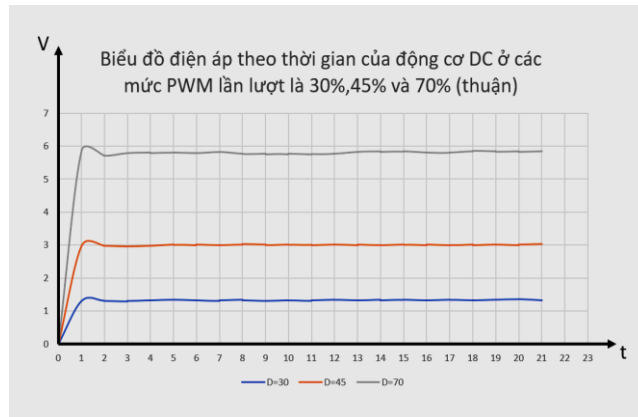
Hình 12: Kết quả kiểm tra chức năng module UART

4.1.6. Kết quả kiểm tra chức năng module PWM

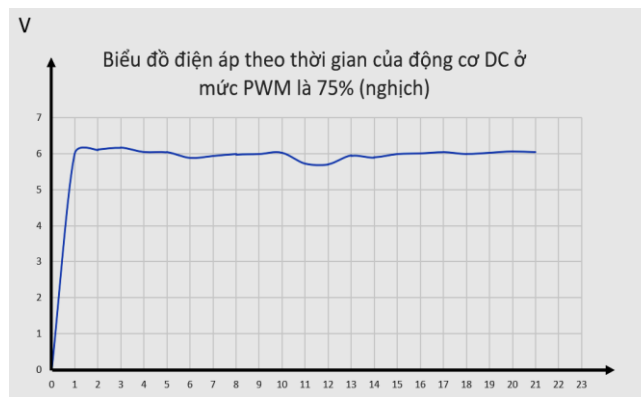
Module PWM tạo hai xung PWM có tần số $f = 5 \text{ KHz}$ và có chu kỳ nhiệm vụ (D) thay đổi để điều khiển tốc độ, đảo chiều quay động cơ [9].

Điều chỉnh tăng giảm độ rộng xung thông qua nút nhấn với độ rộng thay đổi (D=30%, 45%, 70%, ở các tốc độ này ta có thể thấy được tốc độ thay đổi rõ ràng), đảo chiều quay của motor (D = 75%).

Vì điều khiển tạo xung PWM để điều khiển tốc độ động cơ DC. Để kiểm tra tính chính xác, nhóm tiến hành đo và lập bảng giá trị điện áp ngõ ra cầu H để quan sát. Sau nhiều lần đo, nhóm đã thu được kết quả sau:



Hình 13: Đồ thị biểu diễn điện áp ngõ ra DC theo chiều thuận



Hình 14: Đồ thị biểu diễn điện áp ngõ ra DC theo chiều nghịch

Khi tiến hành đo đảo chiều, nếu giữ nguyên vị trí que đo điện áp thì kết quả thu được sẽ là âm. Vì thế khi đảo chiều que đo, sẽ thu được kết quả dương.



Hình 15: Ảnh test thực tế chức năng PWM

5. KẾT LUẬN

Bài báo đã trình bày về cách thiết kế SoC trên FPGA, phần cứng sử dụng là Kit DE0-Nano Development and Education với chip Cyclone IV EP4CE22F17C6N của Altera và phần mềm Nios II là chương trình viết bằng ngôn ngữ C. Kết quả mà bài báo đạt được cuối cùng đó là thiết kế được vi điều khiển 32 bit dựa trên kit FPGA và kết nối với các thiết bị ngoại vi để kiểm tra chức năng đã thiết kế cho vi điều khiển. Dựa vào kết quả thực nghiệm kiểm tra các chức năng đã tạo ra. Ưu điểm của chip nhúng thiết kế trên FPGA là tốc

độ xử lý cao, cấu hình phần cứng có thể thay đổi linh hoạt, có thể kết nối dễ dàng với các ngoại vi khác và có thể sử dụng để lập trình ứng dụng như các vi điều khiển khác trên thị trường.

TÀI LIỆU THAM KHẢO

- [1] Pong P. Chu, “Embedded SoPC Design with NIOS II Processor and Verilog Examples”, Wiley, Hoboken, New Jersey, 2012.
- [2] Han-Way Huang, Nannan He, “Experience of Teaching Embedded System Design using FPGAs”, Minnesota State University, Mankato, ASEE-NMWSC2013-0040, 2013.
- [4] Norman F. Schneidewind, “Digital Logic and Microprocessor Design”, 2012.
- [5] User Manual, “DE0-Nano”, Terasic technologies, 2018
- [6] TS Vũ Đức Lung, Ths Lâm Đức Khải, Ths Phan Đình Duy, “Ngôn ngữ lập trình phần cứng Verilog”, ĐHQG TPHCM, 2013
- [7] Dr. David, J Greaves, “System on Chip Design and Modelling”, Computer Science Tripos Easter Term, 2011
- [8] https://en.wikipedia.org/wiki/Field-programmable_gate_array
- [9] Implementation of a PID control PWM Module on Altera DE0 Kit Using FPGA, 2016.
- [10] FPGA architecture and verification of built in self-test (bist) for 32-bit adder/subtractor using de0-nano fpga and an analog discovery 2 hardware, 2021.
- [11] <https://www.microsemi.com/>, 2020
- [12] <https://hshop.vn/>, 2021

YSC4F.211

THE SYSTEM OF MONITORING AND TAKE CARE OF PETS REMOTELY

DINH QUANG TUYEN¹, VU QUOC LAN²

Industrial University of Ho Chi Minh City

¹ dinhquangtuyen@iuh.edu.vn, ² 02.vuquoclan.29@gmail.com

Abstract. Nowadays, owning one or more pets is commonplace in Vietnam. The question is how can we care for pets when we have to go on a business trip or when we leave the city to return to our hometown because of epidemic or holidays. We can not bring pets with us. We also can not find a place to send them. For that reasons, we built a monitoring system to take care of pets remotely. Our system can automatic supply food and water for pets via the Internet. The system uses Raspberry Pi as a personal server. In addition, Home Assistant helps to integrate all control functions to a webpage instead of a setup application, which is convenient for remote control via other mobile devices. Compared with existing systems on the market, ours takes advantage of the capabilities of Raspberry Pi, a popular device. In addition to feeding/drinking for pets, the system also has the ability to schedule feeding/drinking, along with monitoring feeding/drinks by camera.

Keywords. Raspberry Pi, personal server, Home Assistant, Pet care system.

HỆ THỐNG GIÁM SÁT VÀ CHĂM SÓC THÚ CUNG TỪ XA

Tóm tắt. Ngày nay, việc sở hữu riêng cho mình một hoặc nhiều thú cưng cũng không còn xa lạ với nhiều người. Vấn đề đặt ra là chúng ta sẽ xử lý thế nào trong một số trường hợp như khi phải đi công tác xa, lúc phải rời thành phố về quê do dịch bệnh hay trong những ngày lễ tết mà không thể mang theo và cũng không thể tìm được chỗ gởi. Từ những lý do đó, chúng tôi xây dựng một hệ thống chăm sóc thú cưng từ xa. Hệ thống của chúng tôi có thể tự động cung cấp thức ăn và nước uống cho thú cưng từ xa qua Internet. Hệ thống sử dụng Raspberry Pi như một máy chủ cá nhân. Thêm vào đó, Home Assistant giúp cho việc kết hợp tất cả các chức năng điều khiển lên một trang web thay vì phải xây dựng các ứng dụng, điều này thuận lợi cho việc điều khiển từ xa thông qua các thiết bị di động khác nhau. So với các hệ thống hiện có trên thị trường, hệ thống tận dụng khả năng của Raspberry Pi, một thiết bị phổ biến, có sẵn. Ngoài việc cho thú cưng ăn/ uống, hệ thống còn cung cấp thêm khả năng lên lịch cho ăn/uống, cùng với việc theo dõi cho ăn/uống qua camera.

Từ khóa. Raspberry Pi, máy chủ cá nhân, Home Assistant, hệ thống chăm sóc thú cưng.

1. INTRODUCTION

With the strong development of IoT, using a compact embedded computer like Raspberry PI [2] as a personal server aims to take full advantage of the capabilities of an embedded computer. It also aims to bring users another view of the real capabilities of a computer as small as an ATM card. The initial goal was to research and design a system in which the Raspberry PI plays the role of a server configured with the necessary services (Home Assistant [6], DDNS [7]). The focus is to support the ability to manage, monitor and control the devices over the Internet. It is also important to connect the Raspberry PI to the supporting devices in the system (Current switch module 4 Relay 5V [3], Speed reducer motor code 310 [4], mini water pump motor code 370 [5], webcam, food container, water pipe and feed tray, led light bulb 5V). It is critical to ensure that the devices work well with a voltage of 5V provided from the GPIO of the Raspberry PI.

"The system of monitoring and taking care of pets remotely". Besides the pet feeding/drinking function [1], the system also has the ability to schedule the pet feeding/drinking function, along with monitoring

feeding/drinks function through a camera. All operations are controlled on a web interface. With DDNS, the system can not only be used within the same local area network (LAN), but also outside the Internet, we just need to access the correct domain name to be able to access and control the device without depending on any cloud server.

In addition, we can see the practical significance that the system can carry out for us. This system works well in difficult situations such as the current epidemic, and has a poteneial to meet the needs of many people if they are away from home but still want to care for their pets. This system is built in an intuitive, easy to use, accessible way. Through that, many people will reduce the difficult thought when they want to raise pets but do not have time to take care of them.

2. DESIGN SYSTEM

2.1. The components

2.1.1. Raspberry Pi 4

Acting as a server configured with necessary services to support the ability to manage, monitor and control remote devices over the Internet.



Fig. 1. Raspberry PI 4

Advantages:

- Small size but still provides full connectivity through ports such as USB 2.0, USB 3.0 (on Raspberry Pi 4), HDMI, Ethernet.
- There is WIFI and Bluetooth.
- Low power consumption helps them to work regularly for many hours and still save monthly expenses.
- Application in many different projects and products with low budget.

Defect:

- For configuration, knowledge of Linux operating system is required.
- There is no ability to upgrade RAM, some current Raspberry PI versions have the highest RAM version of only 8GB
- The operating system is installed on a MicroSD memory card with a limited lifespan.

2.1.2. Current switch module 4 Relay 5V

The Relay module receives control signals from the Raspberry PI and executes the opening / closing of the devices. Relay module is a common component used in many projects and products that need stability. The main part of the Module is the Relays. These relays are simply K-locks that are electrically switched on and off.



Fig. 2. Module 4 Relay 5V

Specifications:

- Operating voltage: 5V
- The voltage of the relay is 0V, so it can be used for both 5V and 3V3 signals.
- Maximum switching load voltage is 250V AC - 10A or 30V DC - 10A
- There is an indicator light on each relay.
- Dimensions: 76 x 56mm

2.1.3. Speed reduction motor code 310

The speed reducer motor code 310 is used to do the work of pushing food from the food container to the food tray. Unlike many common motors, the speed reducer motor will give a low rotation speed, when the motor current is interrupted, it will stop working due to the nature of the reducer motor, which has a low inertia force.



Fig. 3. Motor giảm tốc mã 310

Specifications:

- Rated voltage: 5V DC
- No-load current: 10mA
- No-load speed: 10rpm (rpm)
- Suitable voltage: 3-12V (6-24 revolutions)
- Weight: 80 grams

2.1.4. Mini water pump motor code 370

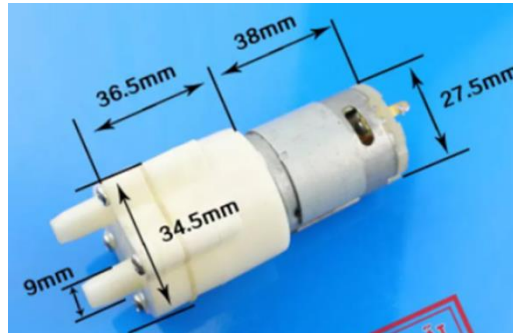


Fig. 4. Mini water pump motor code 370

The mini water pump motor code 370 takes on the task of taking water from the water tank and then bringing the water to the drinking water tray. Unlike other submersible motors commonly found in aquariums, this pump motor has a design of 1 pump and 1 suction that can be reversed by reversing the input voltage.

Specifications:

- Voltage: DC 5V
- Flow: 1.5L per minute.
- Pumping height (without rubber hose) ~20cm
- Motor length: 38mm
- Weight: 60 grams

2.1.5. Webcam



Fig. 5. Webcam

The webcam used had a sufficient resolution to use “720x480” with joints to adjust the angle. The webcam is located on the front and is placed over the food container. With this placement along with the ability to customize the viewing angle, the Webcam will give a wider and more comprehensive view to help fully monitor both feeding and drinking.

2.1.6. Food container



Fig. 6. Food container

The original design of the food container was to divide the dry foods contained in the container by manually turning the knob to simultaneously rotate the rubber wheel in the container. To make the design smarter, we're going to make some changes. Turn the container's manual crown inward and remove the crown. Then connect the reducer motor to the rubber wheel. At this time, the reducer motor will be controlled to rotate the rubber wheel.

2.1.7. Water pipes and Food troughs

Water pipe: Used for installation in the model is a transparent flexible hose with an inner diameter of 9mm to match the size of the pump/suction head of the pump motor. This type of tube is often used in aquariums because of its safe, non-degradable and durable characteristics.



Fig. 7. Water pipe

Feed tray: The trough can be flexibly moved back and forth to match the position of the food tray. The trough is large enough to cover the mouth of the food container to prevent food from falling in the other direction.



Fig. 8. Feed tray

2.1.8. 5V led bulb

A 5V LED bulb with low power consumption and low heat generation will be used as a lighting system for the model to make it easier to observe the feeding and drinking through the camera. The bulb will be attached to the lid of the food container. In this position, the bulb will still ensure good lighting without causing glare when the pet eats/drinks.



Fig. 9. 5V led bulb

2.1.9. Home Assistant (HASS):

Home Assistant is a bridge for smart home devices using different IoT technologies, providing solutions to store data on-premises to help ensure safety, not using the cloud platform (Cloud). Home Assistant provides components with integration capabilities. Instead of each company having to install a different app to be able to add and use it, Home Assistant helps to gather everything into one place, which is both easy to manage and easy to create flexible contexts for devices. All applications work seamlessly together. Controlling everything from an internal "server" not only ensures the safety of the home, and personal data, but also ensures consistency every time the network connection has problems. Users need to have Python programming and system knowledge to install and customize Home Assistant. In this topic, we use Home Assistant to integrate all controls on a web interface that any browser on any smart device's OS can access via the Internet.

2.1.10. Dynamic DNS (DDNS)

DDNS is an acronym for Dynamic Domain Name System, which is used to map the network modem IP address to the domain name every time a system IP change occurs. DDNS was created as an alternative to static IP addresses. It makes it possible for users to directly access the remote system and not need to use the static IP address generated by the ISP. In this topic, we use DDNS to provide a domain name for mobile devices to access to control the system remotely through a web page on the system's server.

2.2. Operating principles

The model will have a surveillance camera on the front for users to observe the amount of food and water in the trays placed in front of the model. A light bulb is placed high above the camera to create a good lighting angle for the surveillance system. The feeding rotary motor is located at the rear of the model, the motor's head will connect to the rubber wheel that is neatly placed in the food container. This ensures that food does not fall out when the motor has stopped rotating, the pump motor responsible for drinking is also located near the feeding motor to ensure the aesthetics of the wiring into the circuit box.

2.3. Overall model of the system

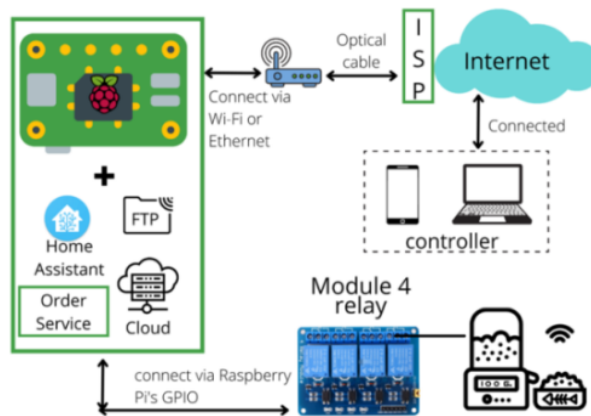


Fig. 10. Overall model of the system.

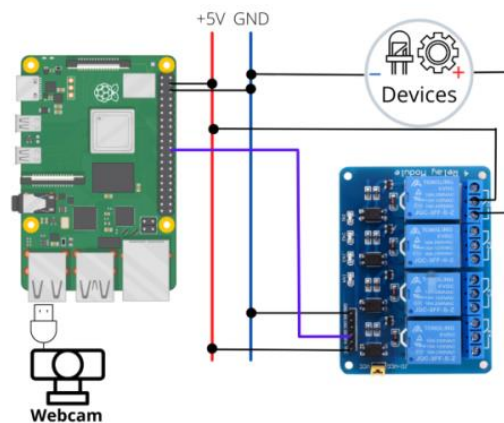


Fig. 11. System hardware connection diagram.

Table 5. Pin connection diagram between Module 4 Relay 5V and Raspberry PI 4

pins on Module 4 Relay 5V	GPIO pins on Raspberry PI 4
VCC	5V
GND	Ground
IN1	GPIO25
IN2	GPIO08
IN3	GPIO07
IN4	GPIO01

Raspberry Pi 4 acts as a personal server installed with appropriate services to provide the necessary features for the system, receiving signals from client computers and transmitting them to the corresponding devices you want to control.

The Raspberry Pi 4 server connects to the Internet via a wired connection (via Ethernet port) or a wireless connection (WIFI). The clients that are allowed to access (have an account) will easily control, manage and monitor the system remotely via the Internet.

The Relay module in the system acts as an electrical fence to increase safety and stability when controlling devices connected to Raspberry Pi 4 via GPIO pins. Except for the Webcam used for monitoring, connected to the Raspberry Pi via USB, other devices such as motors and lights are connected according to the pinout as follows Table 6.

2.4. Block Diagram

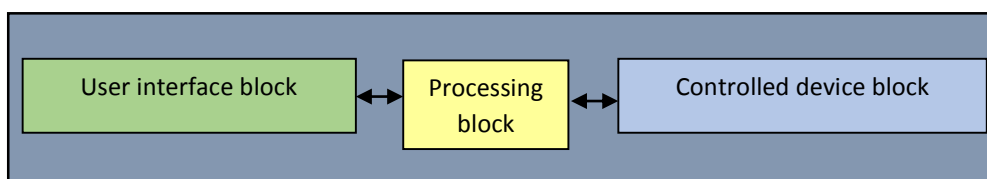


Fig. 12. The block diagram of the system

- Processing block: As an intermediate bridge connecting to the user interface block to receive control signals, the received signal continues to be transmitted by the Raspberry PI's GPIO control pin and the corresponding IN pins in the Module. Relay will receive and execute the opening/closing of devices, the open/ close status of devices will be sent to the control interface in the user interface block to notify users so that they can tell. subsequent decisions. In addition, the signal supply pins must be connected to the corresponding controlled devices according to the predetermined pinout.

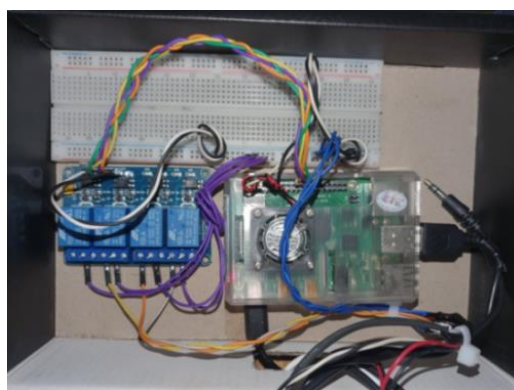


Fig. 13. Processing block

- Controlled device block: The devices that the processing block connects to, used to execute control signals from the user.



Fig. 14. Controlled device block

- User interface block: The user will be provided with a control interface which is a customizable web page with the ability to configure it provided by the Home Assistant service. From the control interface the user performs control operations to the desired devices. The on/off status of the device is updated to the console for the user to monitor and make the next decision.

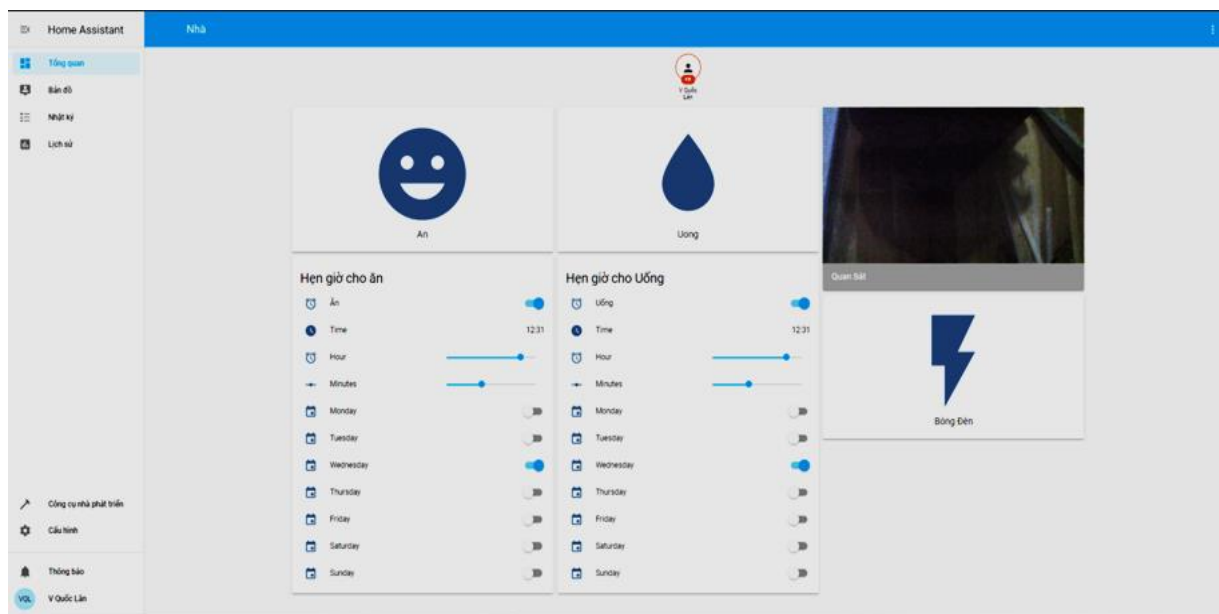


Fig. 15. User interface block

2.5. Algorithm Flowchart

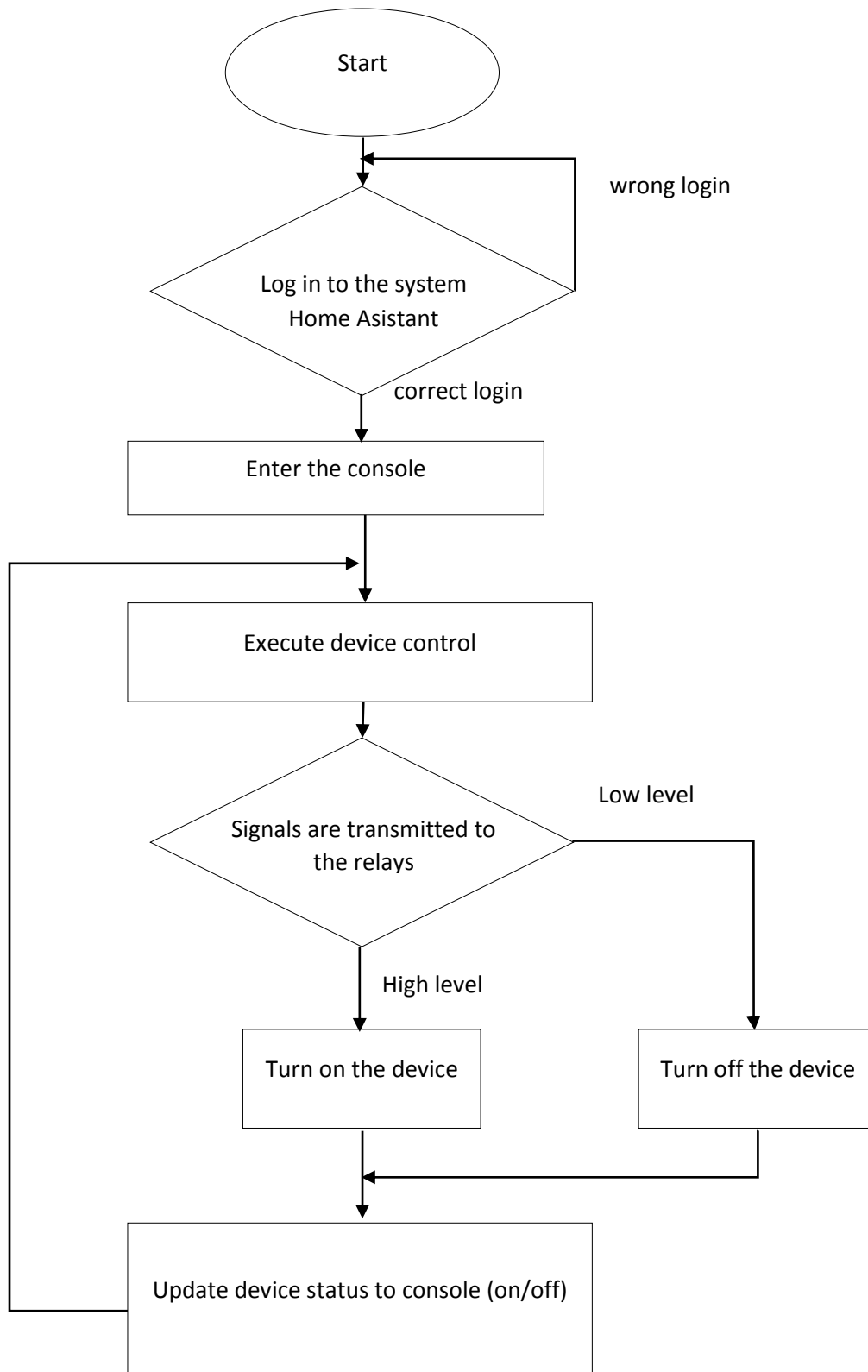


Fig.16. Algorithmic flowchart of the remote feeding and drinking control and management system.

After powering up and accessing the system's domain name with a browser, users will see a request to log in from the Home Assistant system. After entering the correct account, the user will be taken to the main interface page of the system where they can manage, monitor and control the system. The console has a screen displaying images taken directly from the surveillance camera.

3. RESULTS AND DISCUSSION

- Accessibility: With DDNS, we have set up the Home Assistant to be accessible anywhere via the Internet. Experiments on many different smart devices accessing the system via the Internet give stable access and control results.
- Latency: using Raspberry Pi 4 as a personal server, when performing control signals will be transmitted and executed almost immediately. Unlike systems controlled via a 3rd Server service, there will be higher latency. However, here using Raspberry Pi 4 with 1GB RAM version, sometimes the access and processing process is somewhat delayed but not significant.
- Operational ability: Devices such as motors and leds are connected through the Relay module and the IN pins of the Relay module will be connected to the GPIO pins, so the control to switch on and off the current is more stable and safer.
- Limitations: Using Raspberry Pi 4 with low RAM version, it sometimes flickers.

4. CONCLUSIONS

After performing the experimental run, the system is almost entirely guaranteed to meet the initial requirements. The system operates stably, access and control operations have low latency. The model meets almost all criteria such as: beautiful, neat, full of features and safe for users.

Some development directions need attention:

- Complete system optimization to be able to operate more smoothly and stably.
- Expand to connect more smart devices to the system and through the Home Assistant service to be able to control, monitor and manage most devices in the family.

REFERENCES

- [1] Yixing Chen, Maher Elshakankiri, "Implementation of an IoT based Pet Care System", 2020 Fifth International Conference on Fog and Mobile Edge Computing (FMEC).
- [2] raspberrypi.vn. [Online]. Available: raspberrypi.vn. [Accessed: June. 24, 2022].
- [3] dientudat.com/module-4-relay. [Online]. Available: dientudat.com/module-4-relay. [Accessed: June. 24, 2022].
- [4] dosangtao.vn/motor-giam-toc-3-12v-24rpm-ma-310-motor-giam-toc-5v-10rpm. [Online]. Available: raspberrypi.vn. [Accessed: June. 24, 2022].
- [5] dosangtao.vn/may-bom-nuoc-mini-12v-370. [Online]. Available: dosangtao.vn/may-bom-nuoc-mini-12v-370. [Accessed: June. 24, 2022].
- [6] "Tìm hiểu Home Assistant," [Online]. Available: <https://fptshop.com.vn/tin-tuc/danh-gia/dieu-khien-nha-thong-minh-home-assistant-131464>. [Accessed: June. 24, 2022].
- [7] "Tìm hiểu DDNS" [Online]. Available: <https://fptcloud.com/ddns-la-gi/>. [Accessed: June. 24, 2022].

YSC4F.212

ĐIỀU KHIỂN THÍCH NGHI TRƯỢT BACKSTEPPING CHO MÔ HÌNH TÀU THỦY 3D DÙNG MẠNG NƠ-RON

HOÀNG THỊ TÚ UYÊN

Bộ môn Tự Động hóa, Khoa Công nghệ Điện

Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

hoangthituuyen@iuh.edu.vn

Tóm tắt. Bài báo đề xuất sử dụng thuật toán trượt Backstepping thích nghi trên cơ sở mạng nơ-ron nhân tạo cho bài toán bám quỹ đạo của tàu thủy ba bậc tự do có thông số không xác định. Bộ điều khiển trượt Backstepping thích nghi (ASMB) đảm bảo rằng tàu thủy bám quỹ mong muốn và loại bỏ nhiễu. Những thông số và hàm chưa biết của mô hình tàu được xấp xỉ bằng mạng nơ-ron hướng tâm. Tính ổn định của hệ kín ASMB được chứng minh bằng lý thuyết Lyapunov. Kết quả mô phỏng trên máy tính cho thấy đặc tính bám quỹ đạo của thuật toán đề xuất tốt hơn so với bộ điều khiển thích nghi backstepping.

Từ khóa. Kỹ thuật backstepping, điều khiển trượt, mạng nơ-ron, tàu thủy ba bậc tự do.

ADAPTIVE SLIDING MODE BACKSTEPPING CONTROL FOR 3 DOF SURFACE SHIP BASED ON NEURAL NETWORKS

Abstract. This paper proposes an adaptive sliding mode backstepping control algorithm for the tracking problem of uncertain ship model based on neural networks. The adaptive sliding mode backstepping controller (ASMB) ensures that the surface ship tracks desired trajectory and resists disturbance. The unknown functions of ship-model are approximated by radial basis function networks. The stability analysis of closed ASMB system will be given by the Lyapunov theorem. Simulation results on PC show that trajectory tracking properties of the proposed algorithm are better than the adaptive backstepping controller.

Keywords. Backstepping technique, Sliding mode control, Neural networks, 3 DOF surface ship.

1. GIỚI THIỆU

Phương tiện hàng hải có vai trò to lớn trong các lĩnh vực vận tải, khảo sát, giám sát, nghiên cứu khôi phục môi trường biển và nhiều ứng dụng trong quân sự. Việc điều khiển để tàu bám quỹ đạo khi trên biển sẽ giảm được sức lao động, đồng thời việc bám quỹ đạo chính xác có ý nghĩa to lớn trong khảo sát môi trường biển và trong quân sự. Để nâng cao chất lượng hệ thống điều khiển lái tàu thủy thì việc nghiên cứu các phương pháp điều khiển hiện đại ứng dụng cho các hệ thống lái của tàu đóng vai trò quan trọng. Trên thế giới, việc nghiên cứu hệ thống lái tự động đã được tiến hành trong một thời gian dài. Hiệp hội quốc tế về tự động hóa (IFAC) nhận định hệ thống điều khiển lái tàu là bài toán điều khiển điển hình khó giải quyết [9]. Lý do chính là bởi điều khiển lái tàu thủy có nhiều thách thức xuất phát từ thực tế:

- 1) Môi trường làm việc của phương tiện hàng hải là động, phức tạp và không có cấu trúc, điều này dẫn đến nhiều loạn không dự báo được đối với hệ thống điều khiển, ví dụ như dòng chảy đại dương, sóng và gió.
- 2) Mô hình động lực học của tàu là mô hình phi tuyến bất định nên việc thiết kế bộ điều khiển gặp không ít khó khăn.

Do tầm quan trọng của phương tiện hàng hải mà bài toán điều khiển và nhận dạng các thành phần bất định của phương tiện này nhận được nhiều quan tâm, chú ý. Nhiều bài báo đã sử dụng kỹ thuật thiết kế tựa mô hình và nhiều giải pháp khác cho bài toán điều khiển chuyển động của phương tiện hàng hải [11][18]. Trong đó, bộ điều khiển thích nghi tựa mô hình truyền thống được thiết kế để giải quyết bài toán điều khiển

phương tiện hàng hải với động lực học của phương tiện là tuyến tính, thông số của mô hình không biến đổi theo thời gian. Tuy nhiên, do phương tiện hàng hải làm việc trong môi trường đại dương có nhiễu lớn, không biết trước và không xác định như dòng chảy đại dương, sóng và gió nên việc tuyến tính hóa động lực học của phương tiện sẽ dẫn đến sự xuất hiện của sai số mô hình và thành phần động lực học không mô hình. Những yếu tố không xác định này có thể kích thích động lực học không mô hình tần số cao, phá vỡ hàm của bộ điều khiển thích nghi và gây lên sự bất ổn định của hệ thống vòng lặp kín [4]. Vì vậy việc xử lý động lực học không xác định là điều quan trọng, tạo ra thách thức trong điều khiển chuyển động của phương tiện hàng hải.

Điều khiển nơ-ron bắt nguồn từ khả năng học và điều khiển của con người, từ đó tạo thành khả năng thực hiện nhiều nhiệm vụ phức tạp trong môi trường không xác định của hệ thống. Với cấu trúc phân bố song song, khả năng thích nghi và xấp xỉ vạn năng của mạng nơ-ron mà điều khiển nơ-ron thích nghi rất phù hợp cho điều khiển hệ thống phi tuyến không xác định, phức tạp trong những ứng dụng trong công nghiệp [13] [15][16], cũng như được sử dụng để giải quyết thách thức trong điều khiển chuyển động của phương tiện hàng hải [12, 14, 17]. Đặc điểm của điều khiển nơ-ron thích nghi gồm: (i) thiết kế và phân tích dựa trên lý thuyết ổn định Lyapunov; (ii) sự ổn định và đặc tính của hệ thống điều khiển vòng kín được xác định một cách rõ ràng; (iii) trọng số của mạng nơ-ron được huấn luyện trực tuyến dựa trên hàm Lyapunov của hệ kín, không sử dụng kỹ thuật tối ưu hóa như những phương pháp huấn luyện trước kia. Nhờ đó mà điều khiển nơ-ron thích nghi khắc phục được sự bất lợi của bộ điều khiển nơ-ron dựa trên tối ưu hóa, là rất khó phân tích tính ổn định và đánh giá chất lượng của hệ thống kín [3].

Những năm gần đây, tổng hợp bộ điều khiển phi tuyến dựa trên hàm điều khiển Lyapunov được ứng dụng trong nhiều lĩnh vực. Việc xác định hàm điều khiển Lyapunov luôn là một thách thức. Năm 1990, Petar V. Kokotovic là người đầu tiên đề xuất kỹ thuật backstepping để xác định hàm điều khiển Lyapunov. Kỹ thuật này có khả năng áp dụng cho lớp đối tượng phi tuyến dạng truyền ngược [6]. Bộ điều khiển tổng hợp dựa trên kỹ thuật backstepping mới chỉ đảm bảo tính ổn định cho hệ kín, chưa có khả năng kháng nhiễu và không tổng hợp được bộ điều khiển cho các đối tượng có mô hình bất định. Việc kết hợp các kỹ thuật điều khiển hiện đại với nhau sẽ giúp ta giải quyết được những nhược điểm trên. Ví dụ khả năng kháng nhiễu của hệ thống sẽ tốt hơn khi kết hợp kỹ thuật backstepping với điều khiển trượt.

Bài báo đề xuất xây dựng bộ điều khiển lái cho tàu thủy ba bậc tự do dựa trên sự kết hợp kỹ thuật backstepping với điều khiển trượt, những thành phần bất định trong mô hình động học của tàu thủy sẽ được xấp xỉ bằng mạng nơ-ron hướng tâm (RBF).

2. NHỮNG VẤN ĐỀ TOÁN HỌC CƠ BẢN

2.1. Mô hình động lực học của tàu thủy

Khi tàu thủy chuyển động trên đại dương giống như tàu chuyển động trên mặt phẳng nằm ngang, tiếp tuyến với bề mặt trái đất. Khi đó chuyển động của tàu thủy thường được mô tả bởi các thành phần chuyển động tiến, chuyển động dạt và chuyển động quay hướng, các chuyển động lên xuống, chuyển động quay lắc và chuyển động quay lật bị bỏ qua. Do đó từ mô hình chuyển động sáu bậc tự do của phương tiện hàng hải, phương trình chuyển động của tàu thủy chỉ còn ba bậc tự do như sau [2][8]:

$$\begin{cases} \dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\eta})\mathbf{v} \\ \mathbf{M}\dot{\mathbf{v}} + \mathbf{C}(\mathbf{v})\mathbf{v} + \mathbf{D}(\mathbf{v})\mathbf{v} + \mathbf{g}(\boldsymbol{\eta}) = \boldsymbol{\tau} + \Delta(\boldsymbol{\eta}, \mathbf{v}) \end{cases} \quad (1)$$

trong đó $\boldsymbol{\eta} = [x, y, \psi]^T \in R^3$ là véc-tơ gồm vị trí (x, y) và hướng (ψ) trong khung tọa độ gắn trái đất e -frame, $\mathbf{v} = [u, v, r]^T \in R^3$ là véc-tơ vận tốc trong hệ tọa độ gắn thân tàu, gồm vận tốc chuyển động tiến (u) , vận tốc chuyển động dạt (v) và vận tốc quay hướng (r) , \mathbf{M} là ma trận quán tính của tàu, $\mathbf{C}(\mathbf{v})$ là ma trận Coriolis và lực hướng tâm, $\mathbf{D}(\mathbf{v})$ là ma trận suy giảm thủy động lực học, $\mathbf{J}(\boldsymbol{\eta})$ là ma trận quay xung quanh trục z, $\mathbf{g}(\boldsymbol{\eta})$ là véc-tơ lực đẩy và lực trọng trường, $\Delta(\boldsymbol{\eta}, \mathbf{v})$ gồm các véc-tơ lực và momen nhiễu từ môi trường và các thành phần không xác định của mô hình tàu, $\boldsymbol{\tau} \in R^3$ là véc-tơ lực và momen. Các ma trận thông số mô hình tàu được biểu diễn như sau:

$$\mathbf{J}(\boldsymbol{\eta}) = \begin{bmatrix} \cos \psi & -\sin \psi & 0 \\ \sin \psi & \cos \psi & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad (2)$$

$$\mathbf{M} = \begin{bmatrix} m - X_u & 0 & 0 \\ 0 & m - Y_v & mx_g - Y_r \\ 0 & mx_g - N_v & I_z - N_r \end{bmatrix} \quad (3)$$

$$\mathbf{C}(\mathbf{v}) = \begin{bmatrix} 0 & 0 & c_{13}(\mathbf{v}) \\ 0 & 0 & c_{23}(\mathbf{v}) \\ -c_{13}(\mathbf{v}) & -c_{23}(\mathbf{v}) & 0 \end{bmatrix} \quad (4)$$

$$\mathbf{D}(\mathbf{v}) = \begin{bmatrix} d_{11}(\mathbf{v}) & 0 & 0 \\ 0 & d_{22}(\mathbf{v}) & d_{23}(\mathbf{v}) \\ 0 & d_{32}(\mathbf{v}) & d_{33}(\mathbf{v}) \end{bmatrix} \quad (5)$$

$$\mathbf{g}(\boldsymbol{\eta}) = \begin{bmatrix} g_1(\boldsymbol{\eta}) \\ g_2(\boldsymbol{\eta}) \\ g_3(\boldsymbol{\eta}) \end{bmatrix} \quad (6)$$

$$\Delta(\boldsymbol{\eta}, \mathbf{v}) = \begin{bmatrix} \Delta_1(\boldsymbol{\eta}, \mathbf{v}) \\ \Delta_2(\boldsymbol{\eta}, \mathbf{v}) \\ \Delta_3(\boldsymbol{\eta}, \mathbf{v}) \end{bmatrix} \quad (7)$$

trong đó: $\mathbf{J}(\boldsymbol{\eta})^T \mathbf{J}(\boldsymbol{\eta}) = \mathbf{I}$; $c_{13}(\mathbf{v})$, $c_{23}(\mathbf{v})$, $d_{11}(\mathbf{v})$, $d_{22}(\mathbf{v})$, $d_{23}(\mathbf{v})$, $d_{32}(\mathbf{v})$, $d_{33}(\mathbf{v})$, $g_1(\boldsymbol{\eta})$, $g_2(\boldsymbol{\eta})$, $g_3(\boldsymbol{\eta})$, $\Delta_1(\boldsymbol{\eta}, \mathbf{v})$, $\Delta_2(\boldsymbol{\eta}, \mathbf{v})$, $\Delta_3(\boldsymbol{\eta}, \mathbf{v})$ là những hàm chưa biết và phụ thuộc vào véc-tơ vận tốc $\mathbf{v} = [u, v, r]$. Vì vậy $\mathbf{C}(\mathbf{v})$, $\mathbf{D}(\mathbf{v})$, $\mathbf{g}(\boldsymbol{\eta})$ và $\Delta(\boldsymbol{\eta}, \mathbf{v})$ là những ma trận, véc-tơ hàm bất định. \mathbf{M} là ma trận quán tính của tàu và thỏa mãn tính chất $\mathbf{M} = \mathbf{M}^T > 0$ (đối xứng xác định dương đã biết [8]).

Mục tiêu của bài toán là điều khiển tàu bám theo quỹ đạo đặt mong muốn $\boldsymbol{\eta}_d$, với mô hình tham chiếu có dạng:

$$\boldsymbol{\eta}_d(\xi) = [x_d(\xi), y_d(\xi), \psi_d(\xi)]^T \quad (8)$$

Giả thiết 1: Quỹ đạo $\boldsymbol{\eta}_d(\xi)$ và đạo hàm riêng thứ nhất, thứ hai bị chặn đồng dạng trên \mathbb{R}^3 . Quỹ đạo tham chiếu là chuyển động lặp lại.

2.2. Mạng nơ-ron nhân tạo RBF

Để nhận dạng những thành phần bất định trong mô hình, bài báo sử dụng mạng nơ-ron RBF để tận dụng những ưu điểm của mạng này: i) Mạng nơ-ron RBF có khả năng xấp xỉ hàm vạn năng và được sử dụng rộng rãi trong nhận dạng và điều khiển thích nghi; ii) Mạng nơ-ron RBF có thể được biểu diễn theo mẫu hồi quy tham số tuyến tính, là tích của véc-tơ trọng số nơ-ron và véc-tơ hồi quy. Trong đó, thành phần của véc-tơ hồi quy là những hàm phi tuyến đầu vào của mạng RBF; iii) Mạng nơ-ron RBF với tín hiệu vào là những hàm có chu kỳ thì có véc-tơ hồi quy thỏa mãn điều kiện kích thích bền PE[10] [19]. Với điều kiện PE là điều kiện quan trọng để hội tụ thông số và nhận dạng chính xác động học hệ thống trong lý thuyết nhận dạng hệ thống và điều khiển thích nghi.

Mạng RBF là mạng 2 lớp nhiều đầu vào, nhiều đầu ra (MIMO). Lớp vào gồm các nơ-ron hướng tâm, đầu ra phụ thuộc vào véc-tơ tín hiệu đầu vào được tính bằng công thức [13]:

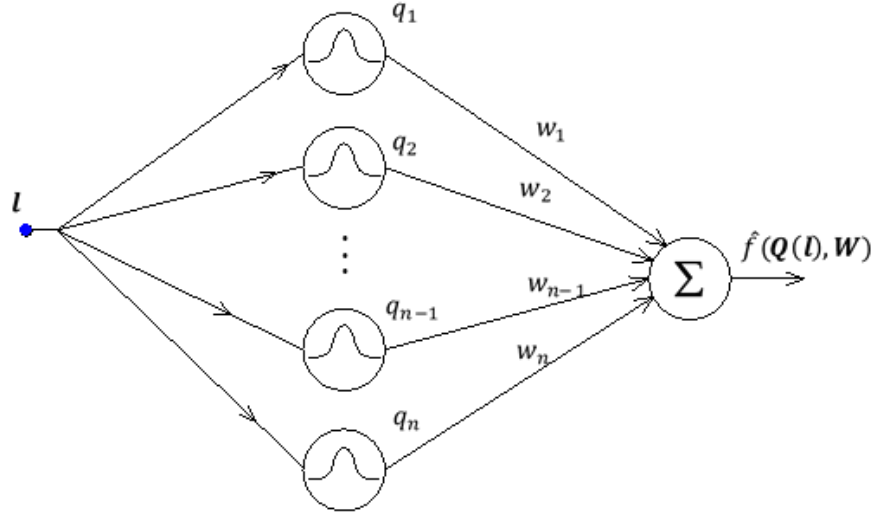
$$q_i(\mathbf{l}) = \exp \left[\frac{-(\mathbf{l} - \boldsymbol{\mu}_i)^T (\mathbf{l} - \boldsymbol{\mu}_i)}{\zeta_i^2} \right], i = 1, 2, \dots, n \quad (9)$$

trong đó: $\boldsymbol{\mu}_i = [\mu_{i1}, \mu_{i2}, \dots, \mu_{im}]^T$ là trọng tâm của hàm hướng tâm và ζ_i là độ tản. Véc-tơ đầu ra của lớp vào được biểu diễn bằng $\mathbf{Q}(\mathbf{l}) = [q_1(\mathbf{l}), \dots, q_n(\mathbf{l})]^T$.

Hàm đầu ra của mạng được xác định như sau:

$$\hat{f}(\mathbf{l}) = \sum_{i=1}^n w_i q_i(\mathbf{l}) = \mathbf{w}^T \mathbf{Q}(\mathbf{l}) \quad (10)$$

trong đó $\mathbf{l} \in \Omega_l \subset R^m$ là véc-tơ đầu vào, $\mathbf{W} = [w_1, w_2, \dots, w_n]^T \in R^n$ là véc-tơ trọng số.



Hình 19. Xấp xỉ hàm bất định bằng mạng RBF

Mạng nơ-ron có khả năng xấp xỉ một hàm phi tuyến bất kỳ với độ chính xác tùy ý. Điều đó có nghĩa là, đối với hàm liên tục trơn bất kỳ $f(\mathbf{l}): \Omega_l \rightarrow R$, trong đó $\Omega_l \subset R^m$ là một tập compact, hàm $f(\mathbf{l})$ được xấp xỉ bằng mạng nơ-ron (10) (với số nơ-ron lớp vào n là đủ lớn), luôn tồn tại một véc-tơ trọng số \mathbf{W}^* sao cho ứng với mỗi $\epsilon^* > 0$ nhỏ tùy ý, ta có:

$$\max_{\mathbf{l} \in \Omega_l} |f(\mathbf{l}) - \mathbf{W}^{*T} \mathbf{Q}(\mathbf{l})| < \epsilon^* \quad (11)$$

hay

$$f(\mathbf{l}) = \mathbf{W}^{*T} \mathbf{Q}(\mathbf{l}) + \epsilon(\mathbf{l}), \forall \mathbf{l} \in \Omega_l \quad (12)$$

trong đó $|\epsilon(\mathbf{l})| < \epsilon^*$.

Kết hợp phương trình (10) và (12), ta có:

$$\Delta f = \hat{f}(\mathbf{l}) - f(\mathbf{l}) = \widetilde{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) - \epsilon, \quad \forall \mathbf{l} \in \Omega_l \quad (13)$$

trong đó $\widetilde{\mathbf{W}} = \widehat{\mathbf{W}} - \mathbf{W}^*$, $\widehat{\mathbf{W}}$ là ước lượng của \mathbf{W}^*

3. THUẬT TOÁN TRƯỢT BACKSTEPPING THÍCH NGHI CHO TÀU THỦY DỰA TRÊN MẠNG NƠ-RON RBF

3.1. Tổng hợp bộ điều khiển

Biến đổi hệ phương trình (1) thành hệ truyền ngược chặt:

$$\begin{cases} \dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\eta}) \boldsymbol{\nu} \\ \dot{\boldsymbol{\nu}} = -\mathbf{M}^{-1} \mathbf{C}(\boldsymbol{\nu}) \boldsymbol{\nu} - \mathbf{M}^{-1} \mathbf{D}(\boldsymbol{\nu}) \boldsymbol{\nu} - \mathbf{M}^{-1} \mathbf{g}(\boldsymbol{\eta}) + \mathbf{M}^{-1} \Delta(\boldsymbol{\eta}, \boldsymbol{\nu}) + \mathbf{M}^{-1} \boldsymbol{\tau} \end{cases} \quad (14)$$

Áp dụng phương pháp trượt backstepping cho đối tượng có dạng truyền ngược chặt trong hệ (14) như sau:

Bước 1: Xét hệ con thứ nhất: $\dot{\boldsymbol{\eta}} = \mathbf{J}(\boldsymbol{\eta}) \boldsymbol{v}$

Định nghĩa các véc-tơ sai lệch:

$$\mathbf{z}_1 = \boldsymbol{\eta} - \boldsymbol{\eta}_d \quad (15)$$

$$\mathbf{z}_2 = \boldsymbol{v} - \boldsymbol{\alpha}_1 \quad (16)$$

trong đó $\boldsymbol{\alpha}_1$ là tín hiệu điều khiển ảo của hệ con thứ nhất. Đạo hàm \mathbf{z}_1 theo thời gian ta có:

$$\dot{\mathbf{z}}_1 = \dot{\boldsymbol{\eta}} - \dot{\boldsymbol{\eta}}_d \quad (17)$$

$$\dot{\mathbf{z}}_1 = \mathbf{J}(\boldsymbol{\eta}) \boldsymbol{v} - \dot{\boldsymbol{\eta}}_d \quad (18)$$

Chọn tín hiệu điều khiển ảo:

$$\boldsymbol{\alpha}_1 = -\mathbf{J}^T(\boldsymbol{\eta})(c_1 \mathbf{z}_1 - \dot{\boldsymbol{\eta}}_d) \quad (19)$$

trong đó c_1 là hằng số thiết kế dương

$$\dot{\mathbf{z}}_1 = \mathbf{J}(\boldsymbol{\eta})(\mathbf{z}_2 + \boldsymbol{\alpha}_1) - \dot{\boldsymbol{\eta}}_d = -c_1 \mathbf{z}_1 + \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_2 \quad (20)$$

Đạo hàm \mathbf{z}_2 theo thời gian ta có:

$$\dot{\mathbf{z}}_2 = \dot{\boldsymbol{v}} - \dot{\boldsymbol{\alpha}}_1 = \mathbf{M}^{-1}[\boldsymbol{\tau} - (\mathbf{C}(\boldsymbol{v})\boldsymbol{v} + \mathbf{D}(\boldsymbol{v})\boldsymbol{v} + \mathbf{g}(\boldsymbol{\eta})) + \Delta(\boldsymbol{\eta}, \boldsymbol{v})] - \dot{\boldsymbol{\alpha}}_1 \quad (21)$$

$$\dot{\mathbf{z}}_2 = \mathbf{M}^{-1}[\boldsymbol{\tau} - (\mathbf{C}(\boldsymbol{v})\boldsymbol{v} + \mathbf{D}(\boldsymbol{v})\boldsymbol{v} + \mathbf{g}(\boldsymbol{\eta})) + \Delta(\boldsymbol{\eta}, \boldsymbol{v}) - \mathbf{M}\dot{\boldsymbol{\alpha}}_1] \quad (22)$$

Chọn hàm Lyapunov cho hệ (17) như sau:

$$V_1 = \frac{1}{2} \mathbf{z}_1^T \mathbf{z}_1 \quad (23)$$

Đạo hàm V_1 theo thời gian ta có:

$$\dot{V}_1 = \mathbf{z}_1^T \dot{\mathbf{z}}_1 = -c_1 \mathbf{z}_1^T \mathbf{z}_1 + \mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_2 \quad (24)$$

Bước 2:

Đặt biến mới \mathbf{s} như sau:

$$\mathbf{s} = \boldsymbol{\lambda}_1 \mathbf{z}_1 + \mathbf{M} \mathbf{z}_2 \quad (25)$$

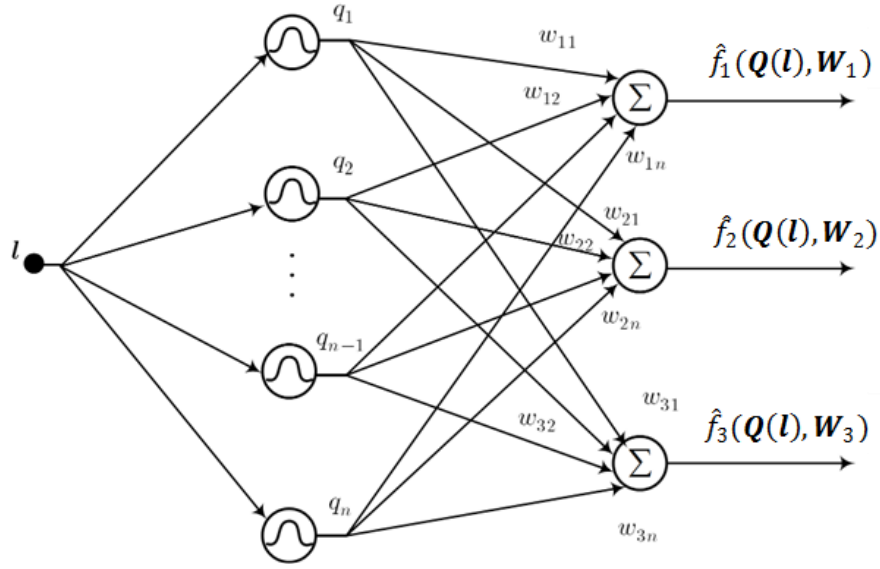
với $\boldsymbol{\lambda}_1$ là ma trận hằng số thiết kế xác định dương được chọn, biến \mathbf{s} được tài liệu [5] gọi là mặt trượt.

$$\dot{\mathbf{s}} = \boldsymbol{\lambda}_1 \dot{\mathbf{z}}_1 + \mathbf{M} \dot{\mathbf{z}}_2 \quad (26)$$

$$\dot{\mathbf{s}} = \boldsymbol{\lambda}_1 \dot{\mathbf{z}}_1 + \mathbf{M} \mathbf{M}^{-1}[\boldsymbol{\tau} - (\mathbf{C}(\boldsymbol{v})\boldsymbol{v} + \mathbf{D}(\boldsymbol{v})\boldsymbol{v} + \mathbf{g}(\boldsymbol{\eta})) + \Delta(\boldsymbol{\eta}, \boldsymbol{v}) - \mathbf{M}\dot{\boldsymbol{\alpha}}_1] \quad (27)$$

Chọn luật điều khiển:

$$\boldsymbol{\tau} = -c_2 \mathbf{h}(\mathbf{s}) - \frac{\mathbf{s} \mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_2}{\mathbf{s}^T \mathbf{s} + \delta} - \boldsymbol{\lambda}_1 \dot{\mathbf{z}}_1 + \mathbf{M}\dot{\boldsymbol{\alpha}}_1 + (\mathbf{C}(\boldsymbol{v})\boldsymbol{v} + \mathbf{D}(\boldsymbol{v})\boldsymbol{v} + \mathbf{g}(\boldsymbol{\eta})) - \Delta(\boldsymbol{\eta}, \boldsymbol{v}) \quad (28)$$



Hình 20. Cấu trúc mạng nơ-ron xấp xỉ véc-tơ hàm bất định

Luật điều khiển (28) chỉ thực hiện được khi biết rõ các ma trận thông số mô hình $C(\mathbf{v})$, $D(\mathbf{v})$ và $\mathbf{g}(\boldsymbol{\eta})$ và biên chặn trên của nhiễu môi trường $\Delta(\boldsymbol{\eta}, \mathbf{v})$. Tuy nhiên như đã thấy trong các tài liệu [2][8] việc xác định chính xác các thông số của ma trận $C(\mathbf{v})$, $D(\mathbf{v})$ và $\mathbf{g}(\boldsymbol{\eta})$ hay biên chặn trên của nhiễu tác động của môi trường (như sóng, gió, dòng chảy đại dương...), sai số đo của các cảm biến là rất khó khăn và tốn nhiều công sức. Bài báo gom các thành phần bất định thành một véc-tơ hàm bất định, véc-tơ hàm bất định này được xấp xỉ bằng mạng nơ-ron hàm hướng tâm.

$$\mathbf{F}(\mathbf{l}) = -(\mathbf{C}(\mathbf{v})\mathbf{v} + \mathbf{D}(\mathbf{v})\mathbf{v} + \mathbf{g}(\boldsymbol{\eta}) - \Delta(\boldsymbol{\eta}, \mathbf{v})), \quad \mathbf{F}(\mathbf{l}) \in \mathbf{R}^3 \quad (29)$$

với:

$$\mathbf{F}(\mathbf{l}) = [f_1(\mathbf{l}), f_2(\mathbf{l}), f_3(\mathbf{l})]^T, \quad \mathbf{l} = [\boldsymbol{\eta}^T, \mathbf{v}^T]^T \quad (30)$$

Véc-tơ hàm bất định (29) được xấp xỉ bằng mạng nơ-ron RBF với đầu ra là $\hat{\mathbf{F}}(\mathbf{l})$. Cấu trúc mạng nơ-ron xấp xỉ véc-tơ hàm bất định của hệ lái tàu biển diễn trong Hình 20. Véc-tơ đầu vào của mạng nơ-ron \mathbf{l} theo Giả thiết 1 và công thức (1) là những tín hiệu lặp lại (có chu kỳ) nên mạng nơ-ron RBF thỏa mãn điều kiện PE và đầu ra của mạng theo (12) có dạng như sau:

$$f_i(\mathbf{l}) = \mathbf{W}_i^{*T} \mathbf{Q}(\mathbf{l}) + \epsilon_i(\mathbf{l}) \quad i = 1, 2, 3 \quad (31)$$

trong đó $\mathbf{W}_i^* \in \mathbf{R}^n$ là các véc-tơ trọng số cần xác định, $|\epsilon_i(\mathbf{l})| \leq \epsilon_i^*$ là sai số xấp xỉ với hằng số $\epsilon_i^* > 0$. Do \mathbf{W}_i^* chưa biết nên ta lấy $\widehat{\mathbf{W}}_i$ để ước lượng \mathbf{W}_i^* . Khi đó $\mathbf{F}(\mathbf{l})$ được ước lượng qua $\hat{\mathbf{F}}(\mathbf{l})$. Như vậy luật điều khiển (28) có dạng sau:

$$\boldsymbol{\tau} = -\mathbf{c}_2 \mathbf{h}(\mathbf{s}) - \mathbf{c}_3 \mathbf{s} - \frac{\mathbf{s} \mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_2}{\mathbf{s}^T \mathbf{s} + \delta} - \lambda_1 \dot{\mathbf{z}}_1 + \mathbf{M} \dot{\boldsymbol{\alpha}}_1 + \hat{\mathbf{F}}(\mathbf{l}) \quad (32)$$

Hàm $\hat{\mathbf{F}}(\mathbf{l}) = [\hat{f}_1(\mathbf{l}), \hat{f}_2(\mathbf{l}), \hat{f}_3(\mathbf{l})]$ là đầu ra của mạng RBF Hình 20. Cấu trúc mạng nơ-ron xấp xỉ véc-tơ hàm bất định có dạng như sau:

$$\hat{f}_i(\mathbf{l}) = \widehat{\mathbf{W}}_i^T \mathbf{Q}(\mathbf{l}) \quad i = 1, 2, 3 \quad (33)$$

Ma trận trọng số $\widehat{\mathbf{W}}_i$ được huấn luyện để $\widehat{\mathbf{W}}_i \rightarrow \mathbf{W}_i^*$ theo luật chỉnh định thích nghi sau:

$$\dot{\widehat{\mathbf{W}}}_i = \dot{\widehat{\mathbf{W}}}_i = \Gamma_i [\mathbf{Q}(\mathbf{l}) z_{2i} - \sigma \widehat{\mathbf{W}}_i], \quad i = 1, 2, 3 \quad (34)$$

trong đó $\widehat{\mathbf{W}}_i = \widehat{\mathbf{W}}_i - \mathbf{W}_i^*$, Γ_i là ma trận chỉnh định thích nghi, đối xứng xác định dương $\Gamma_i \in \mathbf{R}^{n \times n}$ và σ là thông số thiết kế dương tùy chọn.

Như vậy biểu thức (22) được viết lại như sau:

$$\dot{\mathbf{z}}_2 = \mathbf{M}^{-1}[\boldsymbol{\tau} + \mathbf{F}(\mathbf{l}) - \mathbf{M}\dot{\boldsymbol{\alpha}}_1] \quad (35)$$

Biểu thức (27) được viết lại như sau:

$$\dot{\mathbf{s}} = \lambda_1 \dot{\mathbf{z}}_1 + \mathbf{M}\mathbf{M}^{-1}[\boldsymbol{\tau} + \mathbf{F}(\mathbf{l}) - \mathbf{M}\dot{\boldsymbol{\alpha}}_1] \quad (36)$$

$$\dot{\mathbf{s}} = \lambda_1 \dot{\mathbf{z}}_1 - \frac{\mathbf{s}\mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_1}{\mathbf{s}^T \mathbf{s} + \delta} - \lambda_1 \dot{\mathbf{z}}_1 + \mathbf{M}\dot{\boldsymbol{\alpha}}_1 - c_2 \mathbf{h}(\mathbf{s}) - c_3 \mathbf{s} - \widehat{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) + \mathbf{W}^{*T} \mathbf{Q}(\mathbf{l}) + \boldsymbol{\epsilon}(\mathbf{l}) - \mathbf{M}\dot{\boldsymbol{\alpha}}_1 \quad (37)$$

trong đó $\widehat{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) = [\widehat{\mathbf{W}}_1^T \mathbf{Q}(\mathbf{l}), \widehat{\mathbf{W}}_2^T \mathbf{Q}(\mathbf{l}), \widehat{\mathbf{W}}_3^T \mathbf{Q}(\mathbf{l})]^T$,

$$\mathbf{W}^{*T} \mathbf{Q}(\mathbf{l}) = [\mathbf{W}_1^{*T} \mathbf{Q}(\mathbf{l}), \mathbf{W}_2^{*T} \mathbf{Q}(\mathbf{l}), \mathbf{W}_3^{*T} \mathbf{Q}(\mathbf{l})]^T$$

$$\dot{\mathbf{s}} = -\frac{\mathbf{s}\mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_1}{\mathbf{s}^T \mathbf{s} + \delta} - c_2 \mathbf{h}(\mathbf{s}) - c_3 \mathbf{s} - \widetilde{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) + \boldsymbol{\epsilon}(\mathbf{l}) \quad (38)$$

trong đó $\widetilde{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) = [\widetilde{\mathbf{W}}_1^T \mathbf{Q}(\mathbf{l}), \widetilde{\mathbf{W}}_2^T \mathbf{Q}(\mathbf{l}), \widetilde{\mathbf{W}}_3^T \mathbf{Q}(\mathbf{l})]^T$

3.2. Chứng minh tính ổn định của hệ và kết luận

Chứng minh:

Chọn hàm Lyapunov của hệ kín như sau:

$$V_2 = V_1 + \frac{1}{2} \mathbf{s}^T \mathbf{s} + \frac{1}{2} \sum_{i=1}^3 \widetilde{\mathbf{W}}_i^T \boldsymbol{\Gamma}_i^{-1} \widetilde{\mathbf{W}}_i \quad (39)$$

Đạo hàm V_2 theo thời gian:

$$\dot{V}_2 = \dot{V}_1 + \mathbf{s}^T \dot{\mathbf{s}} + \sum_{i=1}^3 \widetilde{\mathbf{W}}_i^T \boldsymbol{\Gamma}_i^{-1} \dot{\widetilde{\mathbf{W}}}_i \quad (40)$$

Thế biểu thức (24) và (38) vào biểu thức (40)

$$\begin{aligned} \dot{V}_2 = & -c_1 \mathbf{z}_1^T \mathbf{z}_1 + \mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_2 + \mathbf{s}^T \left[-\frac{\mathbf{s}\mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_1}{\mathbf{s}^T \mathbf{s} + \delta} - c_2 \mathbf{h}(\mathbf{s}) - c_3 \mathbf{s} - \widetilde{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) + \boldsymbol{\epsilon}(\mathbf{z}) \right] \\ & + \sum_{i=1}^3 \widetilde{\mathbf{W}}_i^T \boldsymbol{\Gamma}_i^{-1} \dot{\widetilde{\mathbf{W}}}_i \end{aligned} \quad (41)$$

trong đó δ là hằng số dương được chọn có giá trị rất nhỏ nên $\frac{\mathbf{s}^T \mathbf{s}}{\mathbf{s}^T \mathbf{s} + \delta} \approx 1$ do đó:

$$\mathbf{s}^T \left(-\frac{\mathbf{s}\mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_1}{\mathbf{s}^T \mathbf{s} + \delta} \right) \approx -\mathbf{z}_1^T \mathbf{J}(\boldsymbol{\eta}) \mathbf{z}_1 \quad (42)$$

Mặt khác: $\mathbf{s}^T \widetilde{\mathbf{W}}^T \mathbf{Q}(\mathbf{l}) = \sum_{i=1}^3 \widetilde{\mathbf{W}}_i^T \mathbf{Q}(\mathbf{l}) \mathbf{s}_i$

Vì vậy \dot{V}_2 được viết lại như sau:

$$\dot{V}_2 = -c_1 \mathbf{z}_1^T \mathbf{z}_1 - c_2 \mathbf{s}^T \mathbf{h}(\mathbf{s}) - c_3 \mathbf{s}^T \mathbf{s} + \mathbf{s}^T \boldsymbol{\epsilon}(\mathbf{z}) + \sum_{i=1}^3 \widetilde{\mathbf{W}}_i^T (\boldsymbol{\Gamma}_i^{-1} \dot{\widetilde{\mathbf{W}}}_i - \mathbf{Q}(\mathbf{l}) \mathbf{s}_i) \quad (43)$$

Thế biểu thức (34) vào biểu thức (43) ta có:

$$\dot{V}_2 = -c_1 \mathbf{z}_1^T \mathbf{z}_1 - c_2 \mathbf{s}^T \mathbf{h}(\mathbf{s}) - c_3 \mathbf{s}^T \mathbf{s} + \mathbf{s}^T \boldsymbol{\epsilon}(\mathbf{z}) - \sum_{i=1}^3 \sigma \bar{\mathbf{W}}_i^T \bar{\mathbf{W}}_i \quad (44)$$

Sử dụng bất đẳng thức Young ta có:

$$\mathbf{s}^T \boldsymbol{\epsilon} \leq \mathbf{s}^T \boldsymbol{\epsilon}^* \leq \mathbf{s}^T \mathbf{s} + \frac{\boldsymbol{\epsilon}^{*T} \boldsymbol{\epsilon}^*}{4} \quad (45)$$

Ta lại có:

$$2\bar{\mathbf{W}}_i^T \bar{\mathbf{W}}_i \geq \|\bar{\mathbf{W}}_i\|^2 - \|\mathbf{W}_i^*\|^2 \quad (46)$$

Từ đó suy ra:

$$\begin{aligned} \dot{V}_2 \leq & -c_1 \mathbf{z}_1^T \mathbf{z}_1 - c_2 \mathbf{s}^T \mathbf{h}(\mathbf{s}) - (c_3 - 1) \mathbf{s}^T \mathbf{s} - \frac{\sigma}{2\lambda_{\max}(\boldsymbol{\Gamma}_i^{-1})} \sum_{i=1}^3 \bar{\mathbf{W}}_i^T \boldsymbol{\Gamma}_i^{-1} \bar{\mathbf{W}}_i + \left(\frac{\boldsymbol{\epsilon}^{*T} \boldsymbol{\epsilon}^*}{4}\right. \\ & \left. + \frac{\sigma}{2} \sum_{i=1}^3 \|\mathbf{W}_i^*\|^2\right) \end{aligned} \quad (47)$$

trong đó $\lambda_{\max}(\boldsymbol{\Gamma}_i^{-1})$ là giá trị riêng lớn nhất của ma trận $\boldsymbol{\Gamma}_i^{-1}$

Chọn:

$$\frac{\sigma}{2\lambda_{\max}(\boldsymbol{\Gamma}^{-1})} = k_0 \quad (48)$$

$$c_1 = c_2 = k_0 \quad (49)$$

$$c_3 = k_0 + 1$$

$$\rho = \frac{\boldsymbol{\epsilon}^{*T} \boldsymbol{\epsilon}^*}{4} + \frac{\sigma}{2} \sum_{i=1}^3 \|\mathbf{W}_i^*\|^2 \quad (50)$$

Khi đó biểu thức (47) được viết lại như sau:

$$\dot{V}_2 \leq -2k_0 V_2 + \rho \quad (51)$$

Từ (51) ta có:

$$0 \leq V_2(t) \leq V_2(t_0) \exp(-2k_0(t - t_0)) + \frac{\rho}{2k_0} \quad (52)$$

Biểu thức (51) và (52) cho thấy hệ kín sẽ ổn định ISS với miền hấp dẫn có bán kính ρ , các biến trạng thái $\mathbf{z}_1, \mathbf{s}, \bar{\mathbf{W}}$ tiến về lân cận của không gian sau một khoảng thời gian hữu hạn $t \geq T_1$.

Kết luận: Hệ thống kín bao gồm hệ lái tàu (1) có mô hình tham chiếu (8) với Giả thiết 1, được điều khiển bằng luật điều khiển (32) và luật cập nhật trọng số (34) đảm bảo:

1. Hệ thống sẽ ổn định ISS.
2. Sai số bám đầu ra $\boldsymbol{\eta} - \boldsymbol{\eta}_d$ hội tụ tới lân cận điểm không trong khoảng thời gian hữu hạn $t \geq T_1$.

4. KẾT QUẢ MÔ PHỎNG KIỂM CHỨNG THUẬT TOÁN TRÊN PHẦN MỀM MATLAB

Để kiểm chứng tính đúng đắn của bộ điều khiển trượt backstepping thích nghi, xấp xỉ bất định bằng mạng nơ-ron nhân tạo, bài báo thực hiện mô phỏng số trên phần mềm Matlab/Simulink với các tham số được tham khảo trong các tài liệu [8] [7] [1] như sau:

$$\begin{aligned} c_{13}(\mathbf{v}) &= -24.6612v - 1.0948r, & c_{23}(\mathbf{v}) &= 25.8v, & d_{11}(\mathbf{v}) &= 0.7225 + 1.3274|u| + 5.8664u^2, \\ d_{22}(\mathbf{v}) &= 0.8612 + 36.2823|v| + 8.05|r|, & d_{23}(\mathbf{v}) &= -0.1079 + 0.845|v| + 3.450|r|, & d_{32}(\mathbf{v}) &= \\ & -0.1052 - 5.0437|v| - 0.13|r|, & d_{33}(\mathbf{v}) &= 1.9 - 0.08|v| + 0.75|r|, & g_1(\boldsymbol{\eta}) &= g_2(\boldsymbol{\eta}) = g_3(\boldsymbol{\eta}) = 0. \end{aligned}$$

Thông số ma trận quán tính như sau: $\mathbf{M} = [25.800 \ 0 \ 0; 0 \ 24.6612 \ 1.0948; 0 \ 1.0948 \ 2.7600]$

Đây là các tham số của tàu CyberShip II (CS2) [8], được xây dựng trong phòng thí nghiệm Marine Cybernetics Đại học khoa học và công nghệ Na Uy. CS2 là một bản sao thu nhỏ tỉ lệ 1:70 của một tàu dịch vụ (tàu cung cấp các nhu yếu phẩm cho giàn khoan dầu).

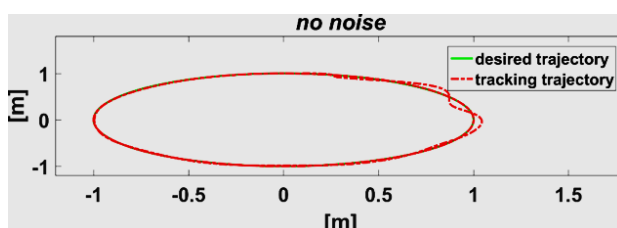
Trong mô phỏng, bài báo sử dụng mạng nơ-ron hai lớp hướng tâm, lớp đầu gồm 1728 nút mạng với trọng tâm mạng μ_i được phân bố đều trên $[-4.4, 4.4] \times [-4.4, 4.4] \times [-4.4, 4.4]$, độ tản của hàm RBF được chọn bằng nhau cho tất cả các nút ở lớp đầu với $\zeta_i = 1.2$ $i = 1, \dots, 1728$. Lớp ra là ba nơ-ron tuyến tính. Các thông số dương thiết kế cho hệ thống được lựa chọn như sau:

$c_1 = c_2 = k_0 = 0.1$, $c_3 = k_0 + 1$, $\Gamma_1 = 10I$, $\Gamma_2 = 10I$, $\Gamma_3 = 8I$, với I là ma trận đơn vị, $I \in R^{1728 \times 1728}$. và $\sigma = 2 * k_0 * \max(\text{eig}(\Gamma^{-1}))$, $\lambda_1 = 300$.

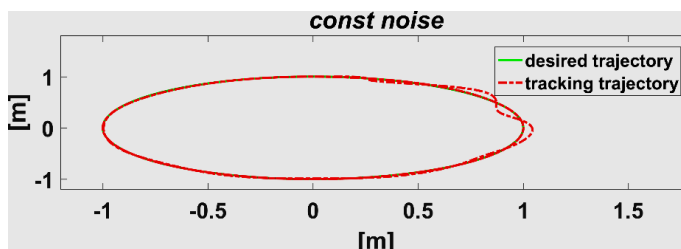
Các trạng thái khởi tạo ban đầu: $\eta_d(0) = [0, 1, 0]^T$ và $\eta(0) = [0.1, 1, 0.0873]^T$

Để xét khả năng chống nhiễu, hệ thống sẽ được mô phỏng với nhiều loại nhiễu hàm được tham khảo trong các tài liệu [1][17], nhiễu hằng và nhiễu trắng như sau:

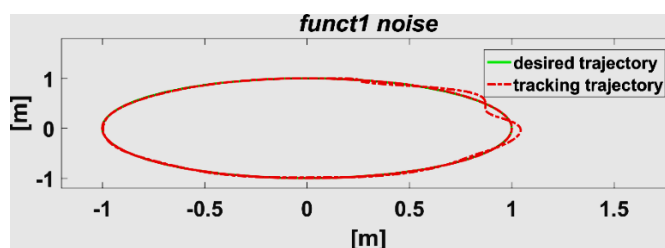
- Hệ thống không chịu tác động của nhiễu.
- Hệ thống chịu tác động của nhiễu hằng: với hàm step có biên độ bằng 4 [N][Nm].
- Hệ thống chịu tác động của nhiễu hàm 1: $\Delta(\eta, v) = [1, 0.01|u|^2 + 0.05, -0.1|r|^3 + \sin(v)]^T$ như tài liệu tham khảo [1].
- Hệ thống chịu tác động của nhiễu hàm 2: $\Delta(\eta, v) = J^T(\eta)f_e(t)$ với $f_e(t) = [3 + 0.8 \sin(0.7t) + 0.2 \sin(0.5t) + 0.2 \sin(0.9t), 0, 0]^T$ như tài liệu tham khảo [17].
- Hệ thống chịu tác động của nhiễu trắng với biên độ bằng 4[N][Nm].



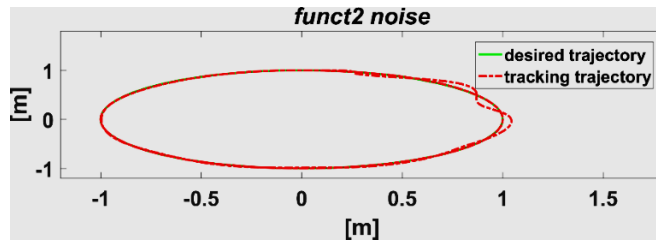
Hình 21. Quỹ đạo bám của tàu thủy khi không có nhiễu tác động



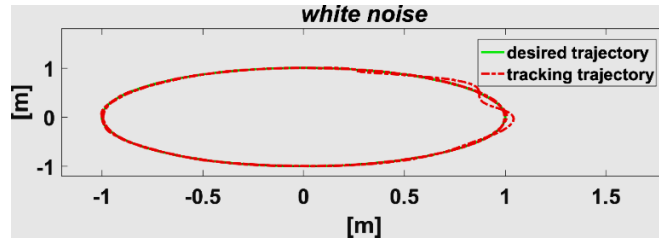
Hình 22. Quỹ đạo bám của tàu thủy khi có nhiễu hằng tác động



Hình 23. Quỹ đạo bám của tàu thủy khi có nhiễu hàm 1 tác động

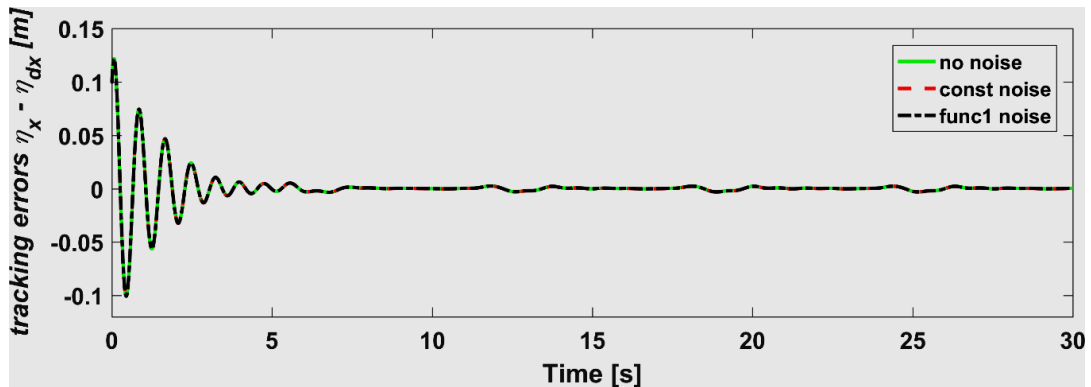


Hình 24. Quỹ đạo bám của tàu thủy khi có nhiễu hàm 2 tác động

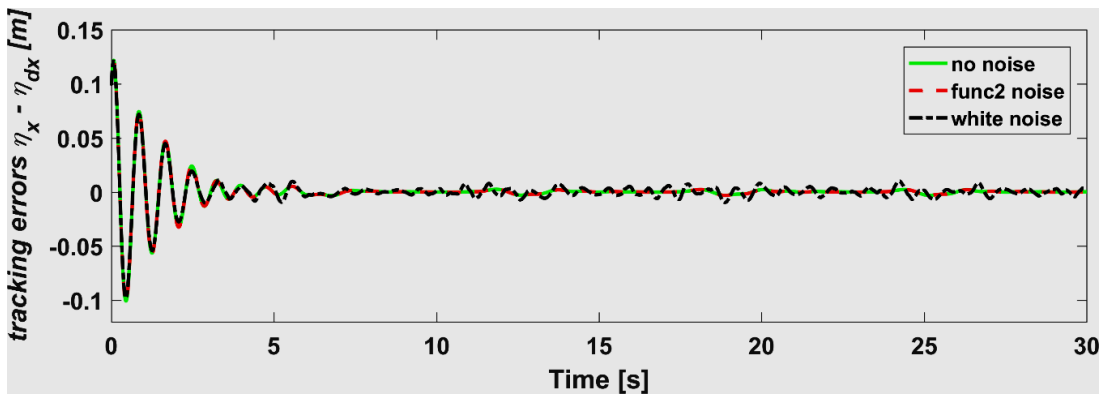


Hình 25. Quỹ đạo bám của tàu thủy khi có nhiễu trắng tác động

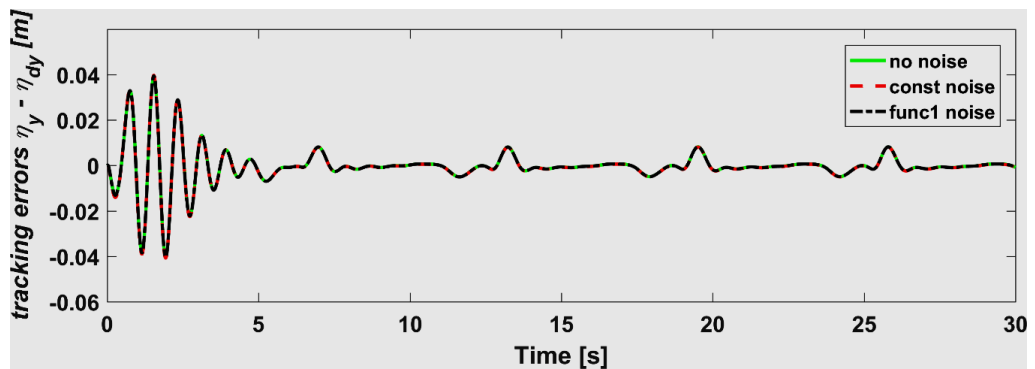
Hình 21 - Hình 25 cho thấy dưới sự tác động của các nhiễu môi trường khác nhau, hệ thống vẫn ổn định, bám được quỹ đạo mong muốn, chỉ dao động nhỏ xung quanh quỹ đạo đặt. Để xét sự dao động này, bài báo phân tích sai số bám trong từng trường hợp như sau:



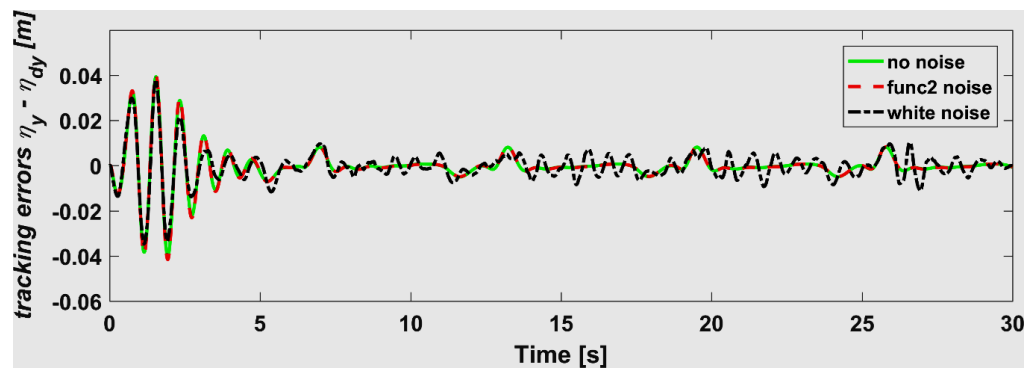
Hình 26. Sai số bám theo trục x: không có nhiễu “-“, nhiễu hằng “- -“, nhiễu hàm 1 “-.”



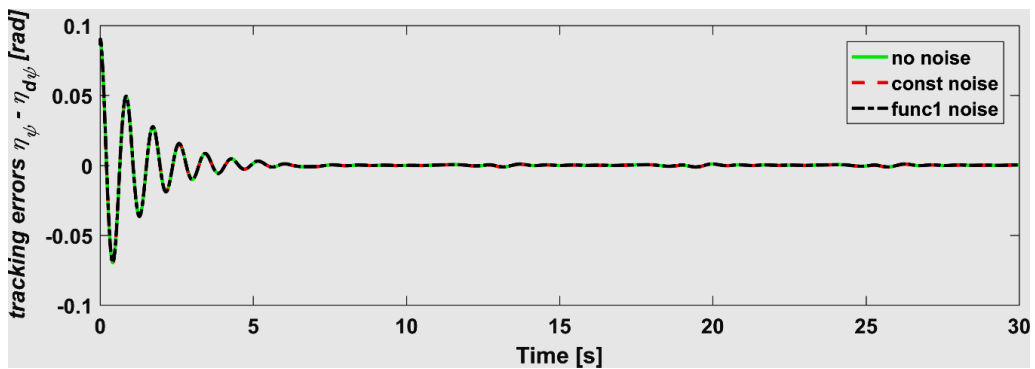
Hình 27. Sai số bám theo trục x: không có nhiễu “-“, nhiễu hàm 2 “- -“, nhiễu trắng “-.”



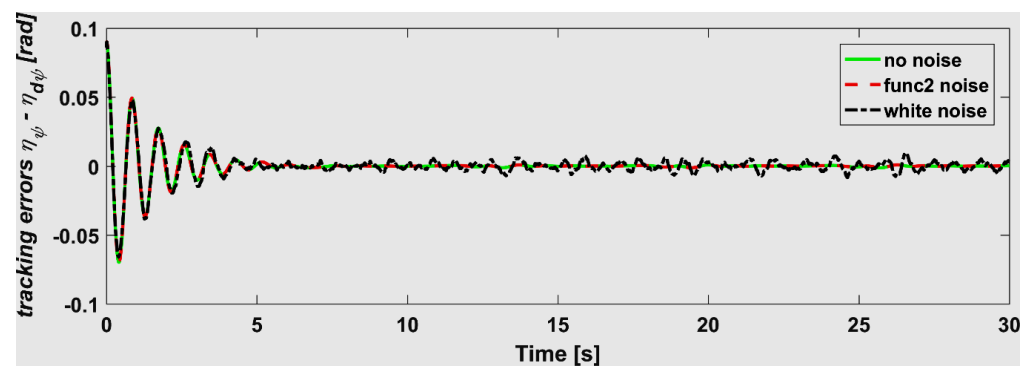
Hình 28. Sai số bám theo trục y : không có nhiễu “-“, nhiễu hằng “-“, nhiễu hàm 1 “-.”



Hình 29. Sai số bám theo trục y : không có nhiễu “-“, nhiễu hàm 2 “-“, nhiễu trắng “-.”

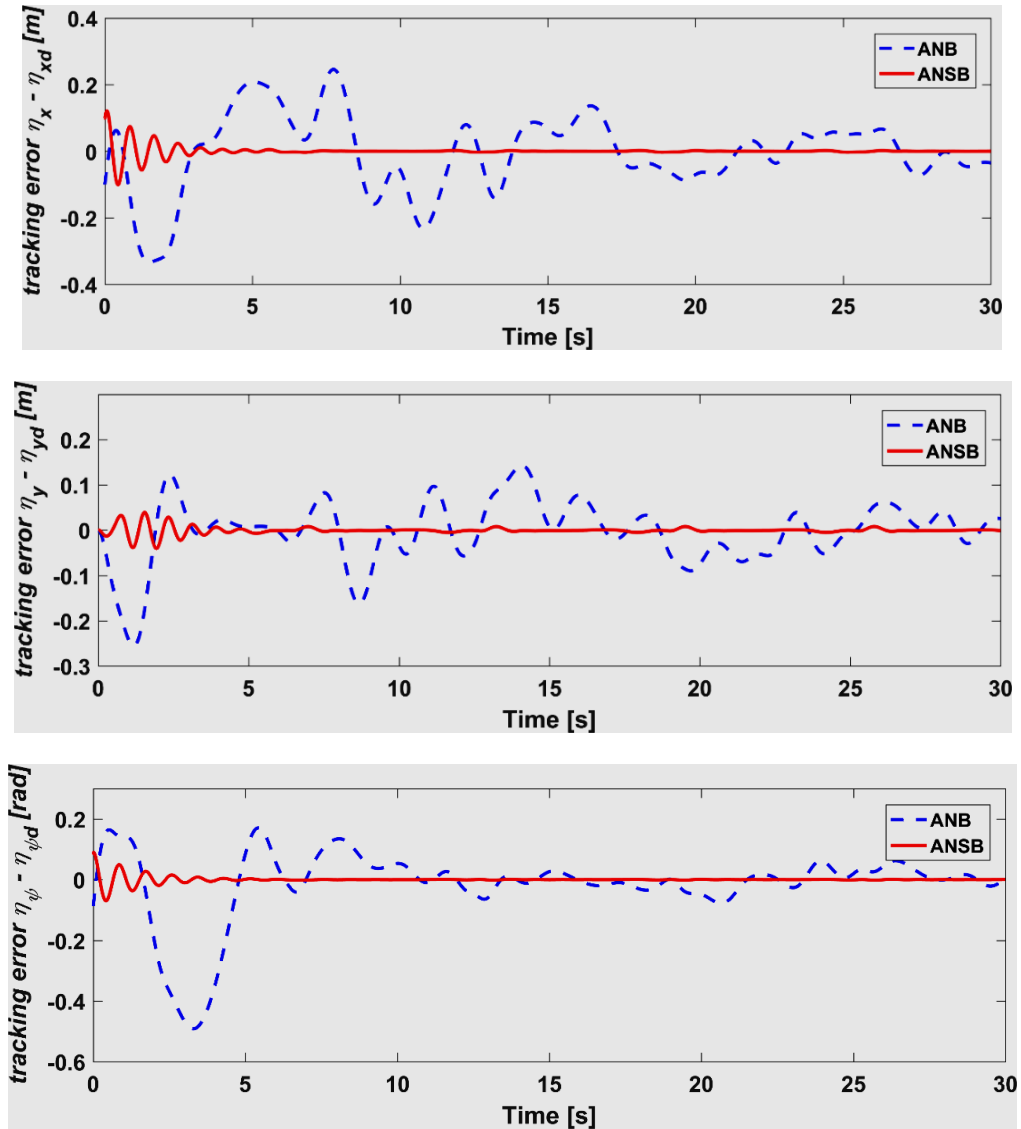


Hình 30. Sai số bám theo ψ : không có nhiễu “-“, nhiễu hằng “-“, nhiễu hàm 1 “-.”



Hình 31. Sai số bám theo ψ : không có nhiễu “-“, nhiễu hàm 2 “-“, nhiễu trắng “-.”

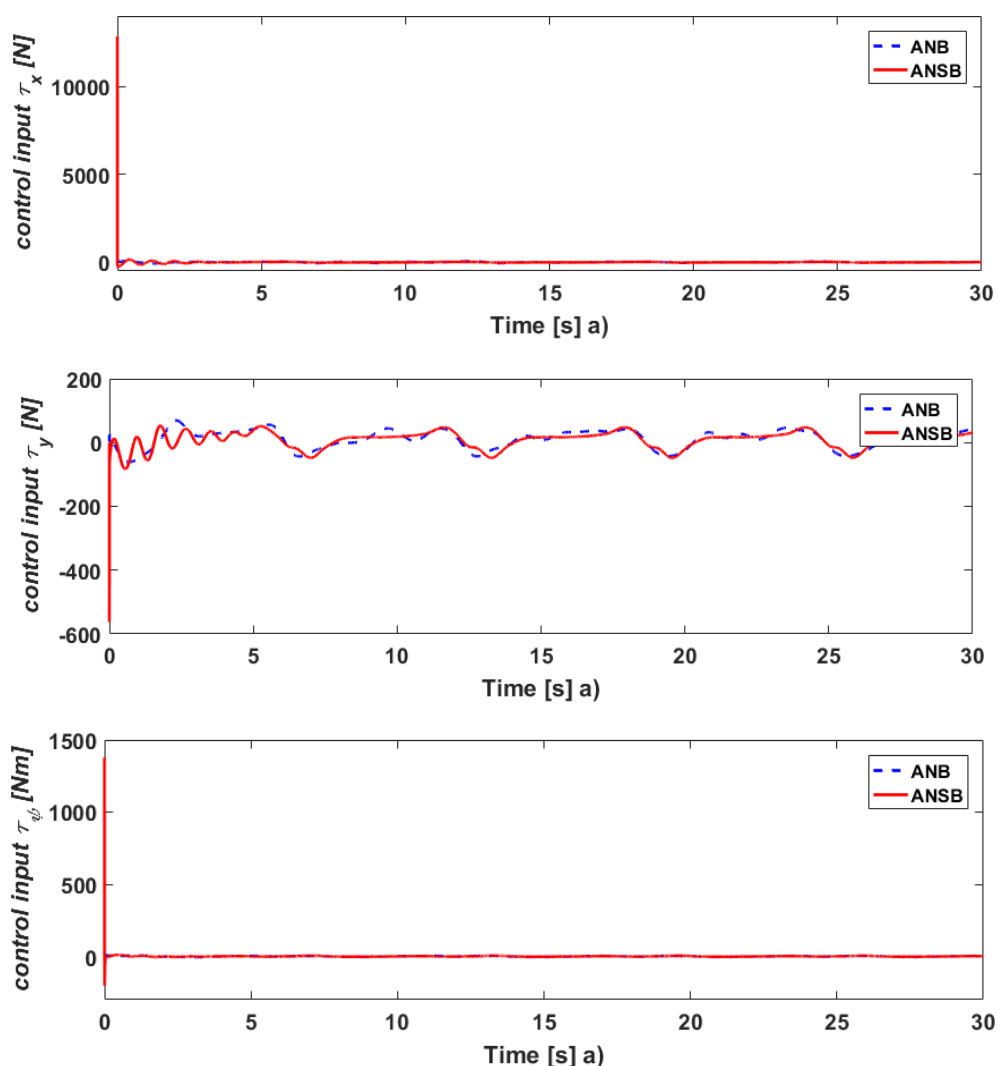
Để thể hiện ưu điểm của thuật toán mà bài báo đề xuất, trong mô phỏng bài báo thực hiện so sánh bộ điều khiển đã đề xuất (ANSB) với bộ điều khiển backstepping thích nghi (ANB), xấp xỉ các thành phần bất định bằng mạng nơ-ron mà tài liệu [1] đề xuất với tác động của nhiễu hàm 1: $\Delta(\boldsymbol{\eta}, \boldsymbol{v}) = [1, 0.01|u|^2 + 0.05, -0.1|r|^3 + \sin(v)]^T$. Kết quả mô phỏng thể hiện như sau:



Hình 32. So sánh sai số bám của hai bộ điều khiển ANB (“--”) và ANSB (“—”)

Hình 32 cho thấy sai số bám của bộ điều khiển ANSB tiến về lân cận không nhanh hơn bộ điều khiển ANB và có giá trị nhỏ hơn sai số bám của bộ điều khiển ANB.

Hình 33 so sánh tín hiệu điều khiển $\boldsymbol{\tau}$ của hai bộ điều khiển ANB và ANSB. Kết quả mô phỏng từ Hình 32, Hình 33 cho thấy: 1) Các bộ điều khiển đều bám quỹ đạo đặt. 2) Sai số bám của bộ điều khiển bài báo đề xuất (ANSB) nhỏ hơn phương pháp ANB. 3) Tín hiệu điều khiển của bộ ANSB ban đầu dao động mạnh hơn ANB do bộ điều khiển này có kết hợp với SMC 4) Thời gian quá độ của phương pháp ANSB là 10(s), sau đó tín hiệu điều khiển ổn định, tương tự tín hiệu điều khiển của bộ ANB.



Hình 33. So sánh tín hiệu điều khiển τ của hai bộ điều khiển ANB (“- -”) và ANSB (“—”)

5. KẾT LUẬN

Bài báo đã trình bày phương pháp tổng hợp bộ điều khiển trượt backstepping nơ-ron thích nghi cho mô hình tàu thủy ba bậc tự do có thành phần bất định giống như trong các tài liệu [1][8] đã đề xuất. Các thành phần bất định được đưa về dạng véc-tơ hàm bất định và được xấp xỉ bằng mạng nơ-ron hướng tâm, bộ điều khiển nơ-ron thích nghi được thiết kế dựa trên điều khiển trượt kết hợp với kỹ thuật backstepping. Bài báo đã chứng minh tính ổn định của hệ kín. Phần mô phỏng cho thấy tín hiệu đầu ra bám quỹ đạo tham chiếu, sai số bám nhỏ hơn bộ điều khiển nơ-ron thích nghi dựa trên kỹ thuật backstepping.

Bài báo sử dụng mạng nơ-ron để xấp xỉ véc-tơ hàm bất định mà không sử dụng đơn thuần nguyên lý điều khiển trượt vốn đã có tính bền vững với thành phần bất định hàm vì các lý do sau: i) Để sử dụng điều khiển trượt thì cần phải biết biên chặn của véc-tơ bất định hàm. Như vậy phải tiến hành nghiên cứu, thực nghiệm để xác định biên chặn của véc-tơ bất định hàm bao gồm cả các thành phần bất định $C(\mathbf{v})$, $D(\mathbf{v})$, $\mathbf{g}(\boldsymbol{\eta})$, $\Delta(\boldsymbol{\eta}, \mathbf{v})$. Công việc này khó khăn và tốn kém do phải thực hiện trên môi trường đại dương nhiều biến động, tác động ở các vùng miền khác nhau. Sử dụng mạng nơ-ron với khả năng xấp xỉ vạn năng để xấp xỉ véc-tơ hàm bất định trong mô hình tàu sẽ bỏ được công việc khó khăn và tốn kém nói trên. ii) Phương pháp điều khiển trượt tạo ra hiện tượng rung, độ lớn của rung đồng biến với biên chặn của thành phần bất định. Để giảm rung người ta giảm biên chặn bằng cách sử dụng mạng nơ-ron để ước lượng các thành phần bất định. Khi đó thành phần bất định được thay bằng sai lệch ước lượng cũng là thành phần bất định nhưng có biên chặn nhỏ hơn. Đây là ưu điểm của thuật toán mà bài báo đề xuất.

TÀI LIỆU THAM KHẢO

- [1] Dai S.-L., Wang C., and Luo F. (2012): *Identification and learning control of ocean surface ship using neural networks*, IEEE Transactions on Industrial Informatics, vol. 8, pp. 801-810.
- [2] Fossen T. (2002): *Marine control systems: guidance, navigation and control of ships, rigs and underwater vehicles* vol. 28, 2002.
- [3] Ge S. S., Hang C. C., Lee T. H., and Zhang T. (2013): *Stable adaptive neural network control* vol. 13: Springer Science & Business Media, 2013.
- [4] Ioannou P. A. and Sun J. (1996): *Robust adaptive control* vol. 1: PTR Prentice-Hall Upper Saddle River, NJ, 1996.
- [5] Koshkouei A. J., Zinober A. S., and Burnham K. J. (2004): *Adaptive sliding mode backstepping control of nonlinear systems with unmatched uncertainty*, Asian Journal of control, vol. 6, pp. 447-453.
- [6] Krstic M., Kanellakopoulos I., and Kokotovic P. V. (1995): *Nonlinear and adaptive control design* vol. 222: Wiley New York, 1995.
- [7] Skjetne R. (2005): *The maneuvering problem*, 1, NTNU, PhD-thesis.
- [8] Skjetne R., Fossen T. I., and Kokotović P. V. (2005): *Adaptive maneuvering, with experiments, for a model ship in a marine control laboratory*, Automatica, vol. 41, pp. 289-298.
- [9] Smith R., "IFAC Benchmark Problems," ed: McGraw Hill, 1994.
- [10] Wang C. and Hill D. J. (2006): *Learning from neural control*, IEEE Transactions on Neural Networks, vol. 17, pp. 130-146.
- [11] Arie T., Itoh M., Senoh A., Takahashi N., Fujii S., and Mizuno N. (1986): *An adaptive steering system for a ship*, IEEE Control Systems Magazine, vol. 6, pp. 3-8.
- [12] Burns R. S. (1995): *The use of artificial neural networks for the intelligent optimal control of surface ships*, IEEE Journal of Oceanic Engineering, vol. 20, pp. 65-72.
- [13] Ge S. S. and Wang C. (2002): *Direct adaptive NN control of a class of nonlinear systems*, IEEE Transactions on Neural Networks, vol. 13, pp. 214-221.
- [14] Leonessa A., VanZwieten T., and Morel Y. (2006): "Neural network model reference adaptive control of marine vehicles," in *Current trends in nonlinear systems and control*, ed: Springer, 2006, pp. 421-440.
- [15] Orłowska-Kowalska T. and Kaminski M. (2011): *FPGA implementation of the multilayer neural network for the speed estimation of the two-mass drive system*, IEEE transactions on Industrial Informatics, vol. 7, pp. 436-445.
- [16] Orłowska-Kowalska T. and Szabat K. (2007): *Neural-network application for mechanical variables estimation of a two-mass drive system*, IEEE Transactions on Industrial Electronics, vol. 54, pp. 1352-1364.
- [17] Tee K. P. and Ge S. S. (2006): *Control of fully actuated ocean surface vessels using a class of feedforward approximators*, IEEE Transactions on Control Systems Technology, vol. 14, pp. 750-756.
- [18] Van Amerongen J. (1984): *Adaptive steering of ships—A model reference approach*, Automatica, vol. 20, pp. 3-14.
- [19] Wang C. and Hill D. (Year): *Persistence of excitation, RBF approximation and periodic orbits*, in *2005 International Conference on Control and Automation*, 2005, pp. 547-552.

YSC4F.213

ĐÁNH GIÁ VỀ NHỮNG CẤU HÌNH VẬN HÀNH, MÔ HÌNH ĐIỀU KHIỂN VÀ CẤU HÌNH NỐI ĐẤT CỦA CÁC DẠNG LƯỚI ĐIỆN MICROGRID

LÊ DUY PHÚC , CHUNG ANH VŨ

*Khoa Công nghệ điện, Viện Kỹ thuật, Trường Đại học Công nghệ Thành phố Hồ Chí Minh
leduyphuc@iuh.edu.vn, vu2020hocmai@gmail.com*

Tóm tắt. Lưới điện Microgrid (MG) là một thành phần trong hệ thống điện hiện đại ngày nay với khả năng vận hành linh hoạt nhằm duy trì độ tin cậy cung cấp điện cao. Bài báo này cung cấp những tổng quan về cấu hình của lưới điện MG, các mô hình điều khiển và cấu hình nối đất tương ứng trước khi thực hiện việc phân tích, đánh giá và so sánh nhằm chỉ ra các ưu nhược điểm của từng loại cấu hình MG. Cụ thể, các cấu hình phổ biến như MG điện xoay chiều (ACMG), MG điện một chiều (DCMG) và MG dạng hỗn hợp (HMG) sẽ được khảo sát. Tiếp theo, các mô hình điều khiển tương ứng với ACMG, DCMG và HMG sẽ được trình bày cụ thể nhằm làm nổi bật các nguyên lý vận hành và phản ứng của từng cấu hình MG khi hoạt động nối lưới/tách lưới. Cuối cùng, những cấu hình nối đất tương ứng với từng dạng cấu hình vận hành của MG cũng được trình bày nhằm mục đích so sánh các đặc điểm vận hành với nhau.

Từ khóa. Microgrid, cấu hình vận hành, mô hình điều khiển, hệ thống nối đất.

REVIEWS ON THE OPERATIONAL CONFIGURATIONS, CONTROL STRATEGIES AND GROUNDING SYSTEMS FOR THE DIFFERENT TYPES OF MICROGRID

Abstract. Nowadays, a microgrid (MG) is an important part of the modern electrical system due to its operational flexibility and high reliability. This paper reviews the possible operational configurations, control strategies, and grounding systems of MG before giving an analysis and comparative results. Specifically, the alternating-current MG (ACMG), direct-current (DCMG), and hybrid MG (HMG) are surveyed. Next, control strategies for these MG types will be introduced to emphasize the operational principles and behaviors corresponding to each MG configuration when in on-grid/off-grid mode. Finally, the grounding systems are also surveyed to indicate the comparison characteristics of each type of MG.

Keywords. Microgrid, operating configuration, control model, grounding system

1. GIỚI THIỆU

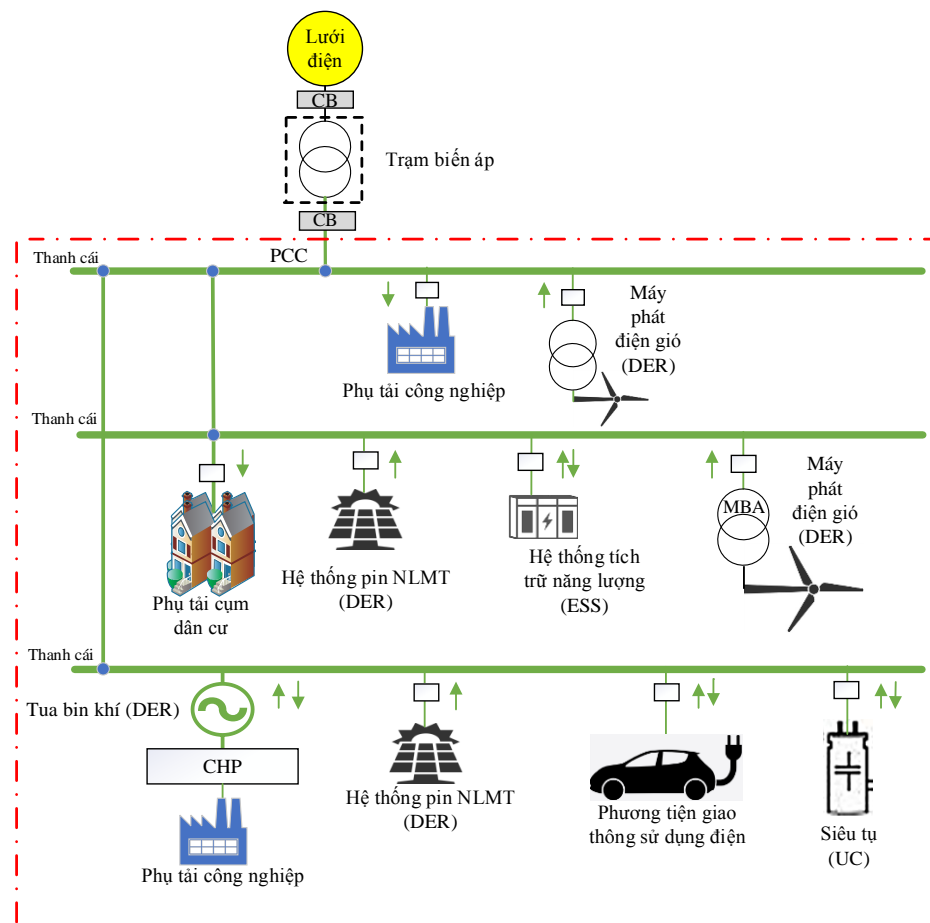
Ngày nay, việc khai thác quá mức các nguồn nhiên liệu hoá thạch dẫn đến sự cạn kiệt. Chính vì lý do đó, để phần nào giảm thiểu sự phụ thuộc vào nguồn nhiên liệu này, thế giới đã và đang liên tục nghiên cứu để ứng dụng các nguồn năng lượng sạch. Qua đó, giúp cải thiện sự thiếu hụt năng lượng trong tương lai và giúp giảm thiểu ô nhiễm môi trường. Tuy nhiên, sự gia tăng đột biến của các nguồn năng lượng sạch, dưới dạng nguồn năng lượng phân tán (Distributed Energy Resource – DER) đã có những tác động tiêu cực nhất định đến cấu hình và khả năng vận hành của lưới điện chính. Các mô hình lưới điện Microgrid (MG) đã được đề xuất để giảm thiểu những tác động này và giúp cho việc vận hành lưới điện phù hợp hơn [1]. Đối với những lưới điện quy mô lớn, MG có thể được tận dụng như một phần tử điều phối công suất và hoạt động để đáp ứng các nhu cầu của lưới điện truyền tải cũng như lưới điện phân phối. MG còn có khả năng đáp ứng các yêu cầu huy động nguồn trong những trường hợp sự cố xuất hiện trên lưới điện phân phối. Đặc biệt hơn hết, MG còn có tác dụng cải thiện đáng kể độ tin cậy cung cấp điện, giảm tổn thất điện năng và duy trì ổn định điện áp cho các phụ tải có liên kết với chúng. Bài báo tập trung giới thiệu và phân tích các cấu hình vận hành, mô hình điều khiển và cấu hình nối đất tương ứng với các dạng lưới điện MG. Từ đó,

giúp các kỹ sư vận hành và kỹ sư thiết kế hệ thống điện có cái nhìn tổng quan nhất về cấu hình lưới điện, công nghệ điều khiển và hệ thống nối đất liên quan đến MG hiện nay.

2. CÁC CẤU HÌNH ĐIỂN HÌNH CỦA LƯỚI ĐIỆN MG

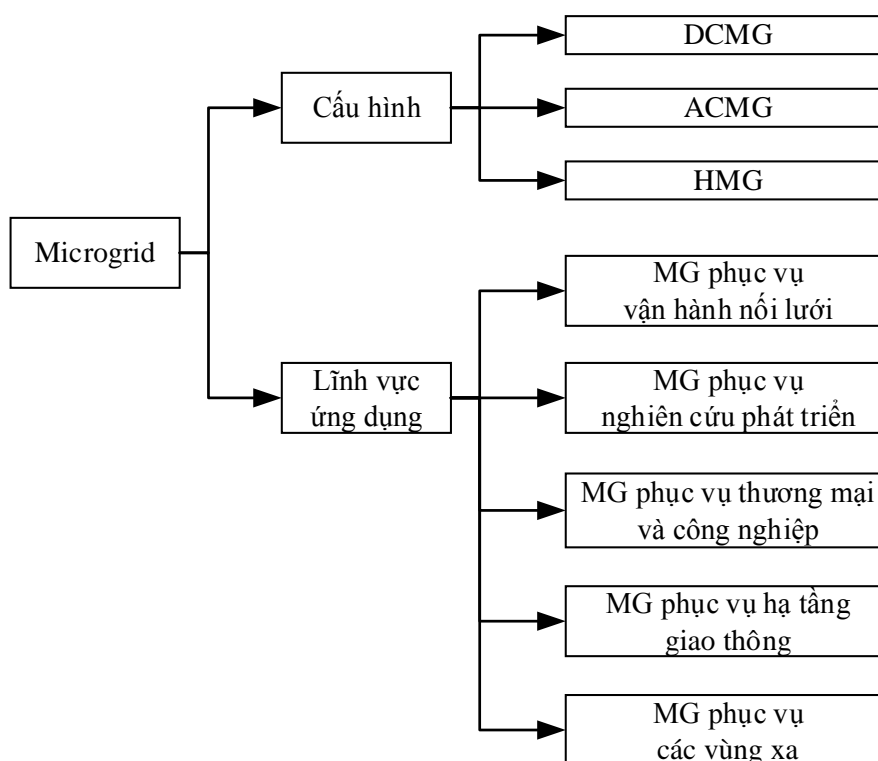
2.1. Định nghĩa, phân loại và tiêu chuẩn liên quan đến MG

Theo [1][2], lưới điện MG được định nghĩa là một hệ thống điện có quy mô vừa và nhỏ, bao gồm các nguồn DER và các phụ tải với những đặc tính vận hành linh hoạt nhằm đáp ứng được những nhu cầu sử dụng khác, như biểu diễn tại **Hình 1**. Bên cạnh đó, lưới điện MG có khả năng vận hành độc lập mà hoàn toàn không phụ thuộc vào nguồn lưới điện chính [3].



Hình 1: Cấu hình lưới điện MG điển hình có tích hợp các DER và phụ tải [4]

Lưới điện MG có thể được xem là một thành phần quan trọng khi phát triển lưới điện thông minh trong tương lai. Nếu dựa vào cấu hình và đặc tính vận hành, lưới điện MG có thể được chia thành 03 dạng chính như sau: i) lưới điện MG sử dụng điện xoay chiều – ACMG; ii) lưới điện MG sử dụng điện một chiều – DCMG; iii) lưới điện MG hỗn hợp – HMG. Ngoài ra, các MG có thể được phân loại dựa trên lĩnh vực ứng dụng, chẳng hạn như, 1) MG vận hành nối lưới, 2) MG dành cho nghiên cứu phát triển, 3) MG phục vụ thương mại và công nghiệp, 4) MG phục vụ hạ tầng giao thông và 5) MG dành cho đảo độc lập về mặt địa lý, như **Hình 2**.

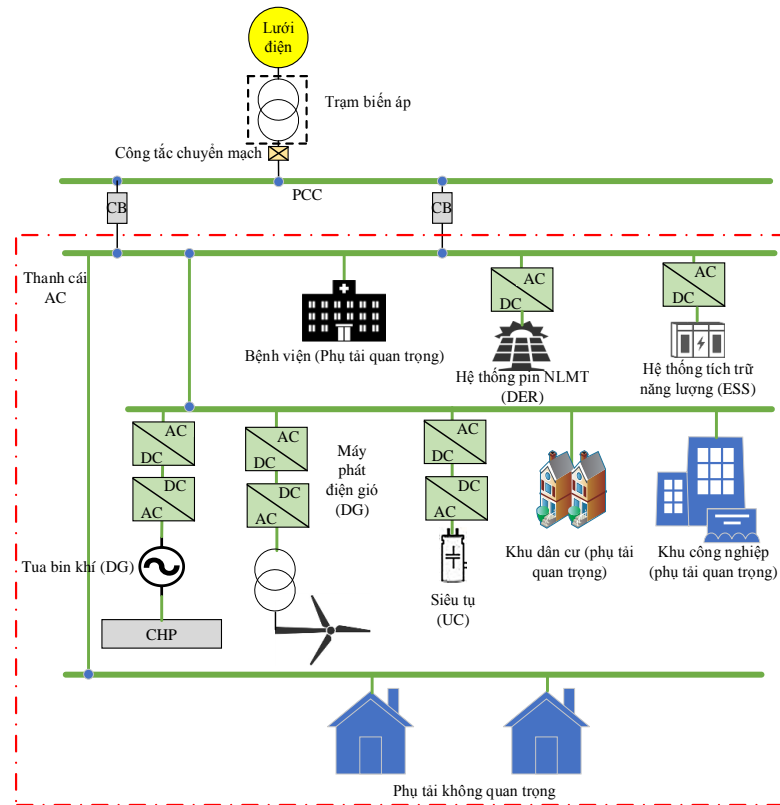


Hình 2: Phân loại lưới điện MG theo cấu hình và lĩnh vực ứng dụng

Hầu hết các thiết kế MG ngày nay đều tuân theo các tiêu chuẩn về kết nối giữa lưới điện truyền thống và DER. Sự thâm nhập cao của DER trong lưới điện phân phối đã làm giảm sự phụ thuộc và khả năng kiểm soát của nguồn điện truyền thống. Qua đó, dẫn đến việc thay đổi các đặc điểm vận hành của lưới điện này. Khi được kết nối với lưới điện chính, các hệ thống MG có quy mô và công suất nhỏ không gây khó khăn trong việc kết nối, điều khiển và vận hành. Tuy nhiên, việc tích hợp số lượng lớn các MG vào lưới điện chính là vấn đề cần quan tâm của các kỹ sư vận hành hệ thống điện. Khó khăn đặt ra chính là duy trì sự ổn định và độ tin cậy cung cấp điện. Vì thế, Hiệp hội Kỹ sư Điện và Điện tử (Institute of Electrical and Electronic Engineers – IEEE) đã xuất bản Tiêu chuẩn IEEE 1547-2018 (bản Tiêu chuẩn đầu tiên được xuất bản là IEEE 1547-2003) để đưa ra những ràng buộc về mặt kết nối, tương tác của các DER với các hệ thống điện chính.

Các yêu cầu quy định trong Tiêu chuẩn IEEE 1547-2018 cũng được áp dụng cho các hệ thống lưới điện MG, vì một MG có thể được xem là một cụm các DER. Ngoài ra, Tiêu chuẩn này còn giúp người kỹ sư tích hợp hiệu quả các DER vào hệ thống lưới điện chính và thảo luận bảy vấn đề về các chế độ vận hành khác nhau, gồm: 1) Khả năng tạo công suất phản kháng; 2) Điều chỉnh điện áp và công suất phản kháng; 3) Hiệu suất hoạt động; 4) Yêu cầu về nhiễu động điện áp; 5) Các yêu cầu về mật độ đóng/cắt; 6) Chế độ vận hành độc lập và bảo vệ hiệu quả; 7) Chất lượng điện năng.

2.2. Cấu hình lưới điện MG dùng điện xoay chiều – ACMG



Hình 3: Cấu hình của lưới điện ACMG cùng với nguồn DER và các phụ tải

Lưới điện MG sử dụng điện xoay chiều (ACMG) được nâng cấp từ hệ thống điện truyền thống. Đây được xem là cấu hình phổ biến nhất, thường được sử dụng để phục vụ cho việc nghiên cứu và triển khai lưới điện MG trong tương lai. Bằng cách sử dụng cơ sở hạ tầng lưới điện xoay chiều sẵn có (như hệ thống phân phối, máy biến áp, hệ thống bảo vệ, v.v...), cấu hình của ACMG dễ dàng được thiết kế và triển khai. Lưới điện ACMG lần đầu tiên được nghiên cứu và phát triển bởi Hiệp hội CERTS vào năm 1998 và được giới thiệu là hệ thống lưới điện chỉ có một cụm máy phát và hệ thống lưu trữ kích cỡ nhỏ, với khả năng tự cách ly khỏi lưới điện chính một cách nhanh chóng mà không gây gián đoạn cung cấp điện cho các phụ tải [5]. Một ví dụ để minh họa cấu hình của lưới điện ACMG được trình bày tại **Hình 3**. Cụ thể, ACMG cung cấp điện cho ba thanh cái, hai trong số đó chứa các phụ tải quan trọng, hệ thống lưu trữ năng lượng và hệ thống năng lượng mặt trời. Thanh cái còn lại chứa các nhóm phụ tải không quan trọng. Lưới điện ACMG có thể thay đổi trạng thái liên kết với lưới điện chính thông qua điểm nối chung (Point of Common Coupling – PCC) và công tắc tĩnh [6]. Điều này cho thấy khả năng linh hoạt trong quá trình vận hành của ACMG. Khi hoạt động ở chế độ nối lưới, các nguồn điện bên trong ACMG buộc phải hoạt động với điện áp và tần số tương ứng với lưới điện. Do đó các thông số kỹ thuật của phụ tải, máy phát và thiết bị lưu trữ năng lượng phải được đảm bảo yêu cầu vận hành nối vào lưới điện chính. Đây là ưu điểm khi triển khai cấu hình ACMG. Mặt khác, một trong những nhược điểm chính là yêu cầu số lượng lớn các thiết bị điện tử công suất và kết nối phức tạp (bộ biến tần và bộ chuyển đổi nối tiếp) để đồng bộ hoá các DER với lưới điện chính. Hiệu suất và độ tin cậy của ACMG có thể bị giảm nếu như không được thiết kế và vận hành phù hợp. Tóm lại, để đạt được hiệu suất và độ tin cậy cao, ACMG cần có nhiều bước chuyển đổi hơn so với DCMG [7].

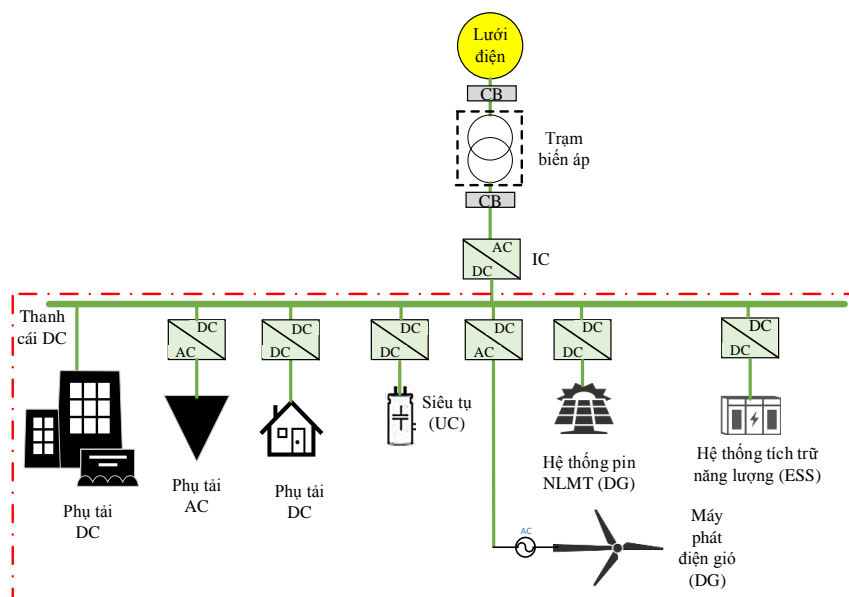
2.3. Cấu hình lưới điện MG dùng điện một chiều – DCMG

Trong lưới điện MG sử dụng điện một chiều, DCMG, các nguồn năng lượng đều ở dạng điện một chiều. Theo đó, việc tích hợp các bộ chuyển đổi DC/DC là cần thiết để đảm bảo hiệu quả vận hành và chất lượng điện năng khi vận hành độc lập với lưới điện phân phối [8]. Một ví dụ về cấu hình của lưới điện DCMG

được minh họa trong **Hình 4** [9]. Bộ chuyển đổi AC/DC đóng vai trò là thiết bị liên kết DCMG với lưới điện chính tại điểm PCC (tại vị trí thanh cái DC). Việc điều chỉnh điện áp của thanh cái DC thông thường được thực hiện bởi bộ chuyển đổi AC/DC nên chất lượng điện áp được đảm bảo dù cho điện áp của lưới điện chính đạt chất lượng thấp.

So với lưới điện ACMG thì DCMG có cấu hình đơn giản hơn, chi phí đầu tư thấp hơn và có lợi thế về mặt tổng thể. Việc đơn giản hóa cấu hình và vận hành là các yêu cầu tiên quyết khi xây dựng DCMG. Ngoài ra, DCMG có tổn thất thấp hơn trong đường dây phân phối nếu so sánh với ACMG bởi vì không có tổn hao về cảm kháng. Hơn nữa, lưới điện DCMG có khả năng vượt qua sự cố dễ dàng, không bị ảnh hưởng bởi lưới điện chính khi xảy ra mất điện hoặc sụt áp do năng lượng dự trữ trong các tụ điện DC hoặc do việc điều chỉnh điện áp tại vị trí liên kết IC (Interconnecting Link).

Mặt khác, lưới điện DCMG có một số hạn chế như cần phải xây dựng đường dây phân phối DC và khó tương thích với hệ thống lưới điện chính. Ngoài ra, các hệ thống bảo vệ của DCMG còn chưa hoàn chỉnh, chủ yếu liên quan đến tiêu chuẩn, hướng dẫn cũng như kinh nghiệm thực tế còn hạn chế. Một nhược điểm khác của DCMG là không thể kết nối trực tiếp với phụ tải AC. Vì vậy, cần có bộ chuyển đổi bổ sung. Tuy nhiên, nhược điểm chính của cấu hình lưới điện DCMG này là phải xử lý toàn bộ dòng công suất trên lưới điện phân phối thông qua một bộ chuyển đổi công suất tại vị trí IC. Vì vậy việc này có thể làm giảm độ tin cậy của lưới DCMG [10].

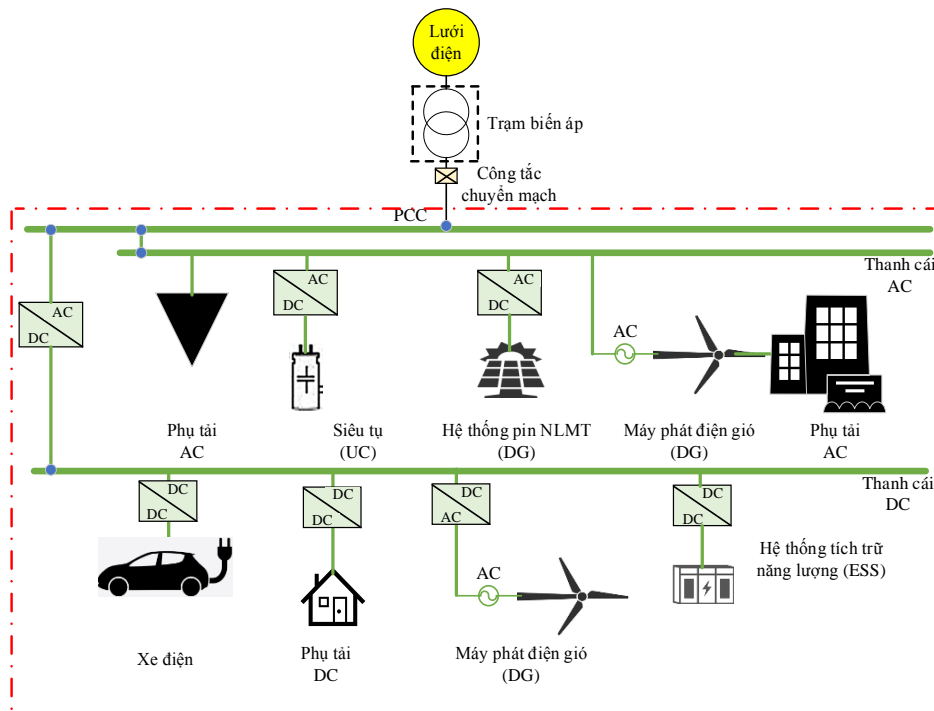


Hình 4: Cấu hình lưới điện DCMG cùng với nguồn DER và các phụ tải [9]

2.4. Cấu hình lưới điện MG hỗn hợp – HMG

Cấu hình hỗn hợp HMG là cấu hình lưới kết hợp cả ưu điểm của cấu hình ACMG và DCMG. Cụ thể, sự tích hợp của hai cấu hình ACMG và DCMG với nhau trong cùng một lưới điện hỗ trợ việc kết nối trực tiếp và vận hành linh hoạt với bất kỳ dạng nguồn năng lượng DER hoặc phụ tải nào. Một cấu hình HMG điển hình được thể hiện tại **Hình 5**, trong đó lưới điện ACMG và DCMG được liên kết thông qua bộ chuyển đổi công suất hai chiều AC-DC [10]. Phụ tải AC được kết nối với thanh cái AC trong khi phụ tải DC được kết nối với thanh cái DC thông qua bộ chuyển đổi nguồn. Các thành phần như DG và ESS cũng được kết nối với thanh cái AC và DC. Có thể thấy rằng, thanh cái AC được phân nhiệm vụ kết nối với những thiết bị hiện hữu trong lưới điện ACMG và thanh cái DC sử dụng cho những thành phần phụ tải DC mới. Theo đó, giảm số lượng bộ chuyển đổi và việc vận hành tương đối đơn giản hơn [10].

Tóm lại, một hệ thống lưới điện HMG có thể được phát triển dựa trên lưới điện chính hiện hữu. Theo đó, các thiết bị kết nối là bộ chuyển đổi nguồn chính AC/DC và một hệ thống mạng truyền dẫn thông tin. Đây cũng chính là nguyên nhân dẫn đến vấn đề chi phí đầu tư có thể cao hơn so với cấu hình ACMG.



Hình 5: Cấu hình lưới điện HMG cùng với nguồn DER và các phụ tải

2.5. Một số nhận xét và đánh giá

Các cấu hình phổ biến nhất của lưới điện MG đã được trình bày trong phần này, bao gồm ba loại chính: ACMG, DCMG và HMG. Trong đó, ACMG là cấu hình được sử dụng nhiều nhất bởi vì có ưu điểm về khả năng kết nối trực tiếp, liên tục với các phụ tải và lưới điện chính ở chế độ vận hành nối lưới thường xuyên. Ngoài ra, cấu hình này tạo điều kiện thuận lợi cho việc triển khai việc áp dụng lưới điện MG trên các mạng lưới phân phối điện đã có sẵn hạ tầng. Đây được xem là yếu tố chính khiến ACMG trở thành cấu hình được sử dụng nhiều nhất tính đến thời điểm hiện tại.

Nếu các phụ tải ở trong lưới điện MG yêu cầu nguồn điện chất lượng cao thì cấu hình DCMG là sự lựa chọn phù hợp bởi vì điện áp của thanh cái DC được tạo ra bằng thiết bị điện tử không phụ thuộc vào tần số/điện áp/các dao động trên lưới điện chính. Bên cạnh đó, cấu hình DCMG cung cấp khả năng quản lý năng lượng được lưu trữ trong các thiết bị ESS một cách dễ dàng. Tuy nhiên, bộ chuyển đổi điện năng chính để liên kết lưới điện MG với lưới điện chính là một điểm yếu của cấu hình này. Theo đó, có thể ảnh hưởng đáng kể đến độ tin cậy cung cấp điện nếu như nâng cao tính đáp ứng về chất lượng điện năng. Độ tin cậy cũng phụ thuộc vào các quá trình chuyển đổi năng lượng và các bộ chuyển đổi điện tử công suất phức tạp thì dễ bị hỏng hóc hơn các bộ chuyển đổi đơn giản. Theo nhận định này thì lưới điện MG có cấu hình hỗn hợp có số lượng bộ chuyển đổi thấp nhất (và có các bộ chuyển đổi đơn giản hơn). Trên thực tế, vẫn còn nhiều vấn đề tồn tại để thực nghiệm cấu hình lưới điện hỗn hợp này. **Bảng 1** trình bày tóm tắt về ưu và nhược điểm của ACMG và DCMG.

Bảng 1: So sánh ưu và nhược điểm của ACMG và DCMG

Hiệu suất	ACMG	DCMG
Hiệu suất chuyển đổi	Kém hiệu quả do có nhiều giai đoạn chuyển hoá năng lượng	Hiệu quả do có ít giai đoạn chuyển hoá năng lượng
Bộ điều khiển đồng bộ	Khó, trạng thái quá độ của ACMG phức tạp hơn DCMG	Đơn giản hơn ACMG vì cấu hình liên kết khá đơn giản

Vấn đề về chất lượng nguồn điện	Điện áp, tần số, góc pha và hệ số công suất tương đối phức tạp và cần được quan tâm	Trong DCMG, chỉ duy nhất vấn đề về chất lượng điện cần được quan tâm, do đó sẽ dễ dàng hơn trong việc cung cấp điện cho khách hàng
Tiêu chuẩn hoá	Các tiêu chuẩn cho ACMG đã hoàn thiện vì có cùng tiêu chuẩn với lưới điện chính	Tiêu chuẩn của DCMG vẫn còn đang xây dựng và triển khai
Tích hợp nguồn năng lượng thay thế	Khó, vì cần nhiều mức chuyển đổi hơn để tích hợp với DCDG và dòng AC cần đồng bộ khi tích hợp với ACDG có tần số cao	Tích hợp tốt với PV, pin nhiên liệu, tua bin gió, v.v...
Giao diện kết nối với lưới điện chính	Có thể kết nối với lưới điện chính thông qua điểm PCC	Cần sử dụng bộ chuyển đổi công suất 2 chiều để kết nối với lưới điện chính
Bảo vệ	Có thể dựa trên công nghệ bảo vệ từ lưới điện chính	Vẫn đang nghiên cứu
Độ ổn định	Độ ổn định thấp, tần số cần phù hợp khi đồng bộ hoá, bộ điều khiển phức tạp và có nhiều mức chuyển đổi năng lượng.	Độ ổn định cao, chỉ điện áp cần điều khiển, ít công đoạn chuyển hóa năng lượng, làm tăng tính mạnh mẽ của hệ thống
Độ tin cậy	Thấp hơn DCMG vì có nhiều thiết bị được sử dụng trong ACMG có thể dẫn đến tỉ lệ hỏng hóc cao trên toàn hệ thống	Cao
Chi phí	Chi phí tương đối thấp so với DCMG có cùng quy mô, khi có sẵn các thiết bị hiện có như dây điện, máy phát, v.v....	Chi phí tương đối thấp so với ACMG có cùng quy mô khi xây dựng lưới điện MG hoàn toàn mới.
Cấu hình liên kết	Tương đối phức tạp do sử dụng số lượng lớn bộ chuyển đổi	Cấu hình liên kết tương đối đơn giản

3. CÁC MÔ HÌNH ĐIỀU KHIỂN CỦA LƯỚI ĐIỆN MG TƯƠNG ỨNG VỚI TỪNG CẤU HÌNH VẬN HÀNH

Để lưới điện MG hoạt động một cách tối ưu và đạt hiệu suất cao, sự có mặt của các bộ điều khiển cùng với các yêu cầu về mặt vận hành tương ứng với từng dạng cấu hình là rất cần thiết. Trong phần này, tác giả sẽ trình bày các vấn đề về kỹ thuật điều khiển của từng cấu hình lưới điện MG đã trình bày tại Mục 2.

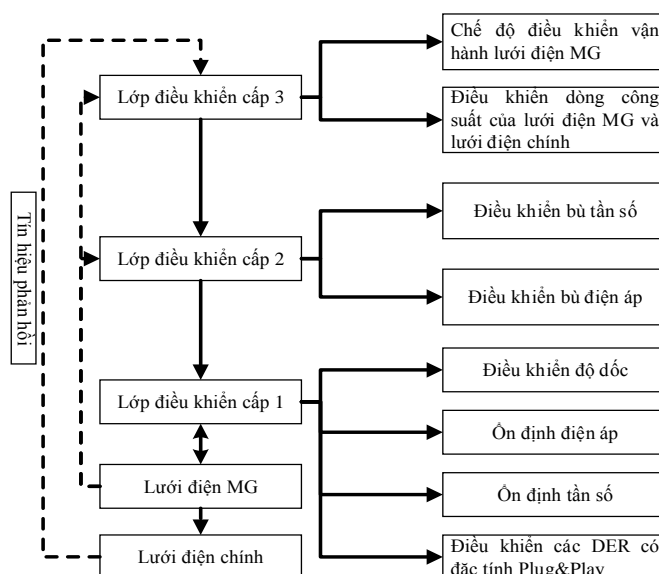
3.1. Các mô hình điều khiển dành cho lưới điện ACMG

Ở cấu hình ACMG, các kỹ thuật điều khiển cần đảm bảo yêu cầu về tính nhất quán, hoạt động ổn định và hiệu quả về mặt kinh tế. Các vấn đề về việc điều khiển ổn định khi áp dụng cấu hình ACMG bao gồm:

- Kiểm soát điện áp và tần số;
- Phối hợp giữa các DER và trao đổi công suất;
- Đồng bộ hoá lưới điện MG;
- Kiểm soát dòng điện giữa nhiều lưới điện ACMG;
- Không gián đoạn khi thay đổi chế độ vận hành từ nối lưới sang tách lưới và ngược lại;
- Hiệu quả về mặt kinh tế;
- Giảm tổn thất trong quá trình sản xuất và truyền tải điện năng;
- Giảm thiểu mất điện.

3.1.1. Cơ chế điều khiển phân cấp

Để đáp ứng các vấn đề điều khiển đã đề cập ở trên, ba cấp độ điều khiển (cấp 1, cấp 2 và cấp 3) được áp dụng trong cấu hình lưới điện ACMG, hay còn được gọi là kiểm soát phân cấp. Mô hình điều khiển phân cấp của lưới điện ACMG được thể hiện trong Hình 6.



Hình 6: Phân cấp điều khiển các phần tử trong và ngoài lưới điện ACMG

Điều khiển cấp 1

Các chức năng chính của điều khiển cấp 1 là:

1. Duy trì điện áp và tần số trong giới hạn cho phép;
2. Chia sẻ công suất tác dụng và công suất phản kháng khi các DER hoạt động song song;
3. Đáp ứng khả năng vận hành Plug&Play – khả năng tự động phát hiện và thay đổi để thích nghi với từng cấu hình vận hành của ACMG.

Điều khiển cấp 1 sẽ hoạt động nhanh hơn so với điều khiển cấp 2 và cấp 3. Trong quá trình chuyển đổi chế độ từ nối lưới sang vận hành độc lập, lớp điều khiển cấp 1 sẽ phải duy trì điện áp và tần số ổn định. Nếu không, việc phát điện không đủ đáp ứng nhu cầu phụ tải sẽ gây ra hiện tượng dao động điện áp và tần số, gây mất độ ổn định trong lưới điện ACMG. Lớp điều khiển cấp 1 thực thi tính toán bằng nhiều vòng lặp nhằm xác định giá trị điện áp/dòng điện phù hợp để điều chỉnh các DER trong ACMG. Cụ thể, các DER có thể thay đổi cơ chế điều khiển điện áp sang cơ chế điều khiển nguồn dòng và ngược lại. Theo [11][12] các phương pháp điều khiển độ dốc (droop-control) và dựa trên trở kháng ảo thường được sử dụng để điều khiển việc chia sẻ công suất giữa các DER với lưới điện ACMG.

Điều khiển cấp 2

Mặc dù, lớp điều khiển cấp 1 đã được phân công nhiệm vụ duy trì điện áp và tần số ổn định nhưng vẫn có sự sai số về giá trị điện áp/tần số ở trạng thái xác lập. Vì vậy, lớp điều khiển cấp 2 sẽ được thiết lập để điều chỉnh bù vào phần chênh lệch tần số và điện áp. Lớp điều khiển này hoạt động chậm hơn so với lớp điều khiển cấp 1 [13][14].

Điều khiển cấp 3

Sau khi độ lệch tần số và điện áp ở trạng thái xác lập đã được điều chỉnh bởi lớp điều khiển cấp 2, lớp điều khiển cấp 3 sẽ thực hiện việc kiểm soát dòng công suất trao đổi giữa lưới điện chính và lưới điện ACMG. Cụ thể, khi vận hành ở chế độ nối lưới, lớp điều khiển cấp 3 sẽ tính toán để điều chỉnh các tham số vận hành của các DER trong lưới điện ACMG. Vì đây là lớp điều khiển cuối cùng và quan trọng nên buộc phải vận hành sau lớp điều khiển cấp 1 và 2 [15][16].

3.1.2. Cơ chế điều khiển dựa trên giao tiếp truyền thông

Các nguồn DER trong lưới điện ACMG cần được phối hợp một cách hợp lý để có thể cung cấp công suất đầy đủ và liên tục cho các phụ tải. Theo đó, khi áp dụng cơ chế điều khiển dựa vào giao tiếp truyền thông, các DER, ESS và phụ tải sẽ được kết nối vào cùng một hạ tầng mạng truyền thông để trao đổi thông tin và phối hợp vận hành [17]. Nhờ vậy, các yêu cầu điều khiển điều chỉnh điện áp và chia sẻ công suất dễ dàng được thực thi.

1. Mô hình điều khiển tập trung

Trong mô hình điều khiển tập trung, một bộ điều khiển trung tâm được sử dụng để thực thi các lệnh điều khiển đến DER và các cơ cấu chấp hành trong lưới điện ACMG. Dựa trên các dữ liệu phụ tải trong quá khứ và trạng thái vận hành của ACMG, bộ điều khiển trung tâm có thể dự báo phụ tải và lập lịch vận hành lưới điện trong tương lai [18][19]. Việc triển khai mô hình điều khiển tập trung chỉ phù hợp với lưới điện ACMG quy mô nhỏ do những giới hạn về khả năng kết nối kênh truyền và khả năng quản lý dữ liệu của các phần tử bên có mặt trong lưới điện.

2. Mô hình điều khiển phân tán

Ở mô hình điều khiển phân tán, mỗi bộ chuyển đổi công suất đều được tích hợp thêm một bộ điều khiển riêng biệt mà không tập trung tại bất kỳ trung tâm điều khiển nào [18][19]. Đối với mô hình điều khiển này, điểm đấu nối PCC được sử dụng làm cơ sở tham chiếu cho các DER thực hiện tính toán tự động và tiến hành phân phối công suất theo từng yêu cầu vận hành cụ thể.

3. Mô hình điều khiển theo cơ chế chủ - tớ

Trong phương pháp điều khiển này, một số DER trong lưới điện ACMG được phân chia nhiệm vụ điều chỉnh điện áp (DER chính) và xác định dòng điện cần bơm vào lưới để gửi đến các DER còn lại (DER phụ) [20][21]. Các DER phụ sẽ liên tục theo dõi giá trị tính toán từ DER chính để có thể tự phân phối công suất. Bởi vì các DER phụ đã giao tiếp trực tiếp với các DER chính nên việc kiểm tra tính tuân thủ và sự đồng bộ của các DER khi vận hành là không cần thiết. Tóm lại, mô hình điều khiển này mang lại sự hiệu quả đối với ứng dụng điều phối dòng công suất tối ưu theo mục tiêu vận hành cụ thể. Mặt khác, trong trường hợp các DER chính bị lỗi, các DER phụ sẽ đổi vai trò thành DER chính. Theo đó, mô hình điều khiển này cho thấy ưu điểm về khả năng vận hành dự phòng trong những trường hợp lỗi kênh truyền. Những phân tích và so sánh giữa ba mô hình điều khiển này được trình bày tóm tắt trong **Bảng 2**.

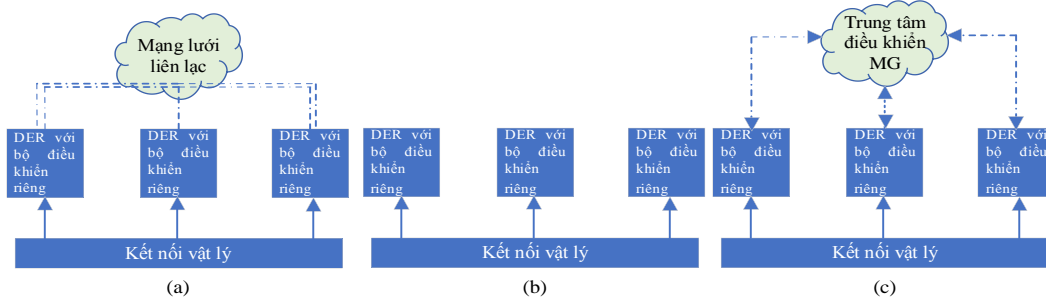
Bảng 2: So sánh các mô hình điều khiển lưới điện ACMG dựa trên giao tiếp truyền thông

Mô hình điều khiển	Ưu điểm	Nhược điểm
Điều khiển tập trung	<ul style="list-style-type: none"> - Đáp ứng tốt về tính nhất quán khi thực thi phân phối công suất trong cả hai trạng thái quá độ và xác lập; - Hoạt động hiệu quả khi MG vận hành điều khiển điện áp và tần số. 	<ul style="list-style-type: none"> - Bắt buộc phải có băng thông cao; - Chi phí đầu tư cao; - Mạng truyền dẫn thông tin liên lạc phải được trang bị theo cơ chế “dự phòng”; - Cả hệ thống phải đồng nhất.
Điều khiển phân tán	<ul style="list-style-type: none"> - Đáp ứng tốt về yêu cầu điều khiển điện áp và công suất không đổi; - Vận hành bình thường trong trường hợp mất kết nối tại một điểm; - Yêu cầu dung lượng băng thông ở mức trung bình. 	<ul style="list-style-type: none"> - Tính thống nhất của hệ thống thấp.
Điều khiển theo cơ chế chủ - tớ	<ul style="list-style-type: none"> - Đáp ứng tốt về yêu cầu điều khiển công suất trong trạng thái xác lập. 	<ul style="list-style-type: none"> - Cần có dung lượng băng thông lớn để giao tiếp.

3.1.3. Các mô hình điều khiển dành cho lưới điện DCMG

Lưới điện DCMG bao gồm các bộ chuyển đổi công suất tương ứng với từng DER và hệ thống lưu trữ năng lượng ESS. Mỗi bộ chuyển đổi công suất đều có khả năng đo lường điện áp, dòng điện và mạch vi xử lý

điều khiển bên trong. Để lưới điện DCMG có thể hoạt động ổn định, các mô hình điều khiển khác nhau như điều khiển tập trung, phi tập trung và điều khiển phân tán có thể được dùng nếu xét từ khía cạnh điều khiển theo giao tiếp truyền thông giữa các thiết bị. Hình 7 biểu diễn ba cấu hình điều khiển lưới điện DCMG thường gặp.



Hình 7: Các mô hình điều khiển cơ bản của DCMG: (a) điều khiển phân tán, (b) điều khiển phi tập trung, (c) điều khiển tập trung

1. Mô hình điều khiển phi tập trung

Điều khiển phi tập trung được thực hiện dựa vào dữ liệu tại các nút nguồn/tải (Data Bus Signal – DBS), bộ điều khiển cục bộ (Local Controller – LC) và tín hiệu truyền trên dây dẫn (Power Line Signal – PLS) để có thể đáp ứng được yêu cầu vận hành. Trong điều khiển phi tập trung, tín hiệu giá trị điện áp tại thanh cái DC tại điểm PCC là cơ sở để phối hợp vận hành giữa các DER, ESS và lưới điện chính. Do đó, độ tin cậy và tính hiệu quả của mô hình điều khiển này phụ thuộc vào việc đo lường các giá trị điện áp của thanh cái DC. Bên cạnh đó, vì đặc tính độc lập của mô hình điều khiển phi tập trung, lưới điện DCMG vẫn có thể hoạt động ngay cả khi sự cố xảy ra ở bộ điều khiển cục bộ.

2. Mô hình điều khiển tập trung

Ở mô hình điều khiển tập trung, nguồn DER và phụ tải trong lưới điện DCMG được tích hợp vào trong một hệ thống điều khiển trung tâm (Control Center – CC) bằng mạng truyền thông kỹ thuật số. Các nguồn điện này được điều khiển bởi CC thông qua việc giao tiếp với các hệ thống điều khiển chính/phụ [22][23]. Về mặt hệ thống, mô hình này cho thấy khả năng quản lý các phần tử bên trong DCMG hiệu quả, nhưng yêu cầu hệ thống phải đồng nhất và phương án dự phòng trong những trường hợp lỗi kết nối.

3. Mô hình điều khiển phân tán

Mô hình điều khiển phân tán được hiểu là mô hình không có thiết bị điều khiển trung tâm và tất cả các LC giao tiếp với nhau thông qua giao tiếp kỹ thuật số chuyên dụng. Mô hình điều khiển phân tán không đòi hỏi sự đồng nhất về mặt thiết bị như mô hình điều khiển tập trung. Ưu điểm của mô hình này nằm ở sự hiệu quả khi quản lý năng lượng và điều chỉnh điện áp tại thanh cái DC nhờ vào khả năng kiểm soát các nguồn điện DER trong lưới điện DCMG. Tuy nhiên, tính ổn định và độ phức tạp cao là những điểm yếu của mô hình điều khiển này [24]. Bảng 3 sẽ trình bày tóm tắt những thuộc tính cơ bản của các mô hình điều khiển đã nêu.

Bảng 3: So sánh đặc tính cơ bản của những mô hình điều khiển trong lưới điện DCMG [25]

Mô hình	Điều khiển phi tập trung	Điều khiển phân tán	Điều khiển tập trung
Kênh truyền	Có	Không	Có
Công nghệ giao tiếp	Công nghệ kỹ thuật số	Dây truyền tín hiệu	Công nghệ kỹ thuật số
Quyết định điều khiển	Tại một thiết bị thông qua điện áp DC tại điểm PCC	Tại thiết bị, dựa vào điện áp đo lường được	Tại CC
Độ tin cậy của chức năng	Chịu đựng được các trường hợp mất kết nối	Phụ thuộc vào độ tin cậy của giao thức truyền thông	Một điểm mất kết nối sẽ gây ảnh hưởng đến toàn bộ quá trình điều khiển

Chi phí	Thấp	Thấp	Tốn kém vì cần dự phòng đường dây giao tiếp
---------	------	------	---

3.1.4. Các mô hình điều khiển dành cho lưới điện HMG

Cấu trúc lưới điện MG hỗn hợp HMG là cấu hình tối ưu nhất tính đến thời điểm hiện tại vì nó kết hợp các ưu điểm của cả ACMG và DCMG [26-28].

Về cơ bản, việc điều khiển các phần tử bên trong lưới điện HMG dựa trên những yêu cầu dành cho cả lưới điện ACMG và DCMG. Theo đó, mô hình điều khiển dành cho hệ thống HMG tương đối phức tạp. Trong Hình 8, mô hình điều khiển dành cho lưới điện HMG được phân thành ba cấp, gồm: điều khiển cấp 1, điều khiển cấp 2 và điều khiển cấp 3.

1. Điều khiển cấp 1

Trong lớp điều khiển cấp 1, việc điều khiển điện áp/dòng điện của những bộ chuyển đổi công suất được thực hiện tại các DER và ESS. Tùy thuộc vào số lượng DER tham gia điều khiển điện áp, phương pháp điều khiển có thể là:

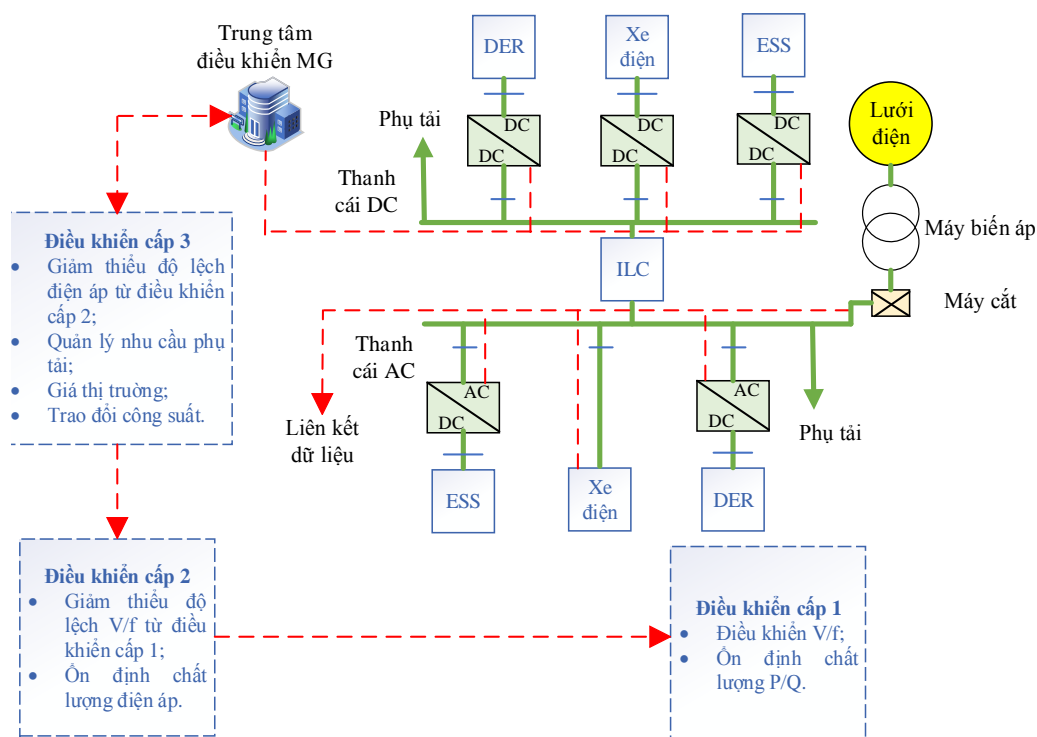
- Sử dụng một bộ chuyển đổi công suất của một DER để vận hành ở chế độ V/f; trong khi đó, các DER còn lại hoạt động ở chế độ P/Q; hoặc
- Sử dụng hai/ba bộ chuyển đổi công suất của DER để vận hành ở chế độ V/f và các DER còn lại hoạt động ở chế độ P/Q. Tuy nhiên, cần sự đồng bộ về mặt vận hành của các bộ chuyển đổi công suất DER khi hoạt động nối lưới.

2. Điều khiển cấp 2

Điều khiển cấp 2 đảm nhận nhiệm vụ điều khiển để bù phần điện áp còn thiếu/thừa tại DCMG và phần điện áp/tần số thiếu/thừa tại ACMG. Việc triển khai điều khiển cấp 2 đòi hỏi phải có cơ sở hạ tầng truyền thông phức tạp, dẫn đến chi phí đầu tư cho hạ tầng cao.

3. Điều khiển thứ 3

Lớp điều khiển này đảm nhận việc điều khiển dòng công suất thực và phản kháng giữa lưới điện HMG để đảm bảo điện áp và tần số tương thích với lưới điện chính.



Hình 8: Sơ đồ điều khiển phân cấp dành cho lưới điện HMG (với ILC là bộ chuyển đổi công suất hai chiều giữa lưới điện ACMG và DCMG – Interlinking converter)

4. CÁC CẤU HÌNH NỐI ĐẤT TƯƠNG ỨNG VỚI TỪNG CẤU TRÚC VẬN HÀNH CỦA LƯỚI ĐIỆN MG

4.1. Cấu hình nối đất sử dụng cho lưới điện ACMG

Trung tính cách ly trong lưới điện ACMG ba pha ba dây

Hệ thống điện ba pha ba dây với trung tính cách ly như được biểu diễn ở Hình 9. Cấu hình nối đất này có ưu điểm là không bị ảnh hưởng về mặt vận hành khi sự cố một pha chạm đất (LG) xảy ra, bởi vì đã được thiết kế để chịu đựng hiện tượng điện áp vọt [29]. Theo đó, giúp hạn chế hậu quả khi sự cố LG xảy ra. Tuy nhiên, hệ thống vẫn cần một dây nối đất ở phía DC để đảm bảo cách ly nguồn điện AC với các thiết bị DC. Mặt khác, việc nối đất vẫn có thể thực hiện thông qua một trở kháng cao (thay vì trực tiếp) trong những ứng dụng buộc phải nối đất. Tuy nhiên, việc phát hiện sự cố LG trong lưới điện trung tính cách ly không đơn giản, vốn có thể dẫn đến việc chuyển hoá thành sự cố LL (2 pha chạm nhau) nếu duy trì trong thời gian dài.

Bên cạnh đó, phương án bảo vệ dựa trên điện áp thường được sử dụng trong cấu hình nối đất này có độ tin cậy không cao [30]. Điều này có thể dẫn đến thời gian khôi phục sự cố dài hơn sau khi sự cố LG xảy ra trong hệ thống ACMG ba pha có trung tính cách ly. Nhìn chung, nếu việc bảo vệ lưới điện ACMG là quan trọng thì không nên áp dụng cấu hình trung tính cách ly. Mặt khác, nếu lưới điện chính đã vận hành sẵn theo cấu hình nối đất này thì hệ thống nối đất của lưới điện ACMG buộc phải đồng nhất.

Trung tính nối đất trực tiếp đơn điểm trong lưới điện ACMG ba pha ba dây

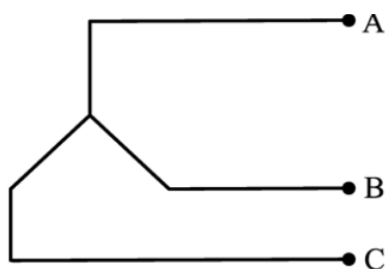
Cấu hình trung tính nối đất trực tiếp đơn điểm được thể hiện trong Hình 10 [31][32]. Cấu hình này có ưu điểm trong việc hỗ trợ thiết bị bảo vệ phát hiện được sự cố xảy ra trong lưới điện ACMG. Mặt khác, việc áp dụng cấu hình nối đất này phụ thuộc vào hệ thống nối đất của lưới điện chính để đảm bảo về mặt đồng bộ và tính nhất quán khi thiết kế, cài đặt cho hệ thống bảo vệ.

Trung tính nối đất trực tiếp đa điểm trong lưới điện ACMG ba pha bốn dây

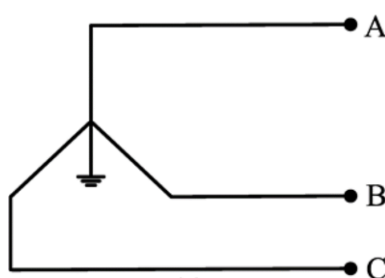
Cấu hình trung tính nối đất trực tiếp đa điểm dành cho hệ thống ACMG ba pha bốn dây có thể tham khảo tại Hình 11. Cấu hình này có nhiều ưu điểm như: (i) giảm sự ảnh hưởng của điện áp quá độ; (ii) tối ưu về số lượng thiết bị bảo vệ quá điện áp; (iii) trở kháng thứ tự không thấp hơn so với trường hợp nối đất đơn điểm; (iv) độ an toàn cao; (v) khả năng cung cấp điện cho tải một pha. Thông tin so sánh các cấu hình nối đất được trình bày chi tiết tại Bảng 4. Để đánh giá sự hiệu quả của những cấu hình nối đất dành cho lưới điện ACMG, ta cần dựa vào các tiêu chí sau: khả năng đảm bảo độ tin cậy hoạt động cho các rô-le bảo vệ, khả năng cách điện của hệ thống và tính đảm bảo an toàn cho con người. Một số giải pháp được liệt kê như sau:

1. Nối đất tại cả hai phía trung áp và hạ áp của lưới điện ACMG và các máy biến áp của nguồn DER;
2. Nối đất theo hệ thống nối đất của lưới điện chính.

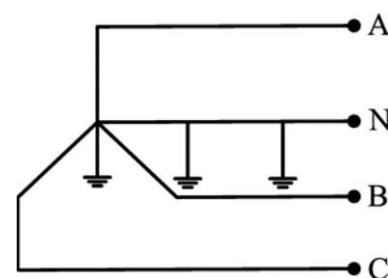
Nhìn chung, cấu hình trung tính nối đất trực tiếp đa điểm được xem là sự lựa chọn phù hợp nhất cho lưới điện ACMG ba pha bốn dây.



Hình 9: Cấu hình lưới điện ACMG ba pha ba dây có trung tính cách ly



Hình 10: Cấu hình lưới điện ACMG ba pha ba dây có trung tính nối đất trực tiếp đơn điểm



Hình 11: Cấu hình lưới điện ACMG ba pha bốn dây có trung tính nối đất trực tiếp đa điểm

Bảng 4: Bảng so sánh các phương án nối đất khác nhau liên quan đến lưới điện ACMG [33]

Cấu hình nối đất	Không nối đất	Trung tính cách ly	Nối đất qua điện trở cao	Nối đất qua trở kháng cao	Nối đất qua điện trở thấp	Nối đất qua trở kháng thấp	Trung tính nối đất trực tiếp
Ứng dụng	Không khuyến khích do hiện tượng quá độ điện áp	Máy phát trung áp	Từ 600V trở xuống, ưu tiên cung cấp điện liên tục	Không được sử dụng do quá áp	Dùng cho hệ thống phân phối trung áp, chỉ dùng cho tải 3 pha	Điện áp định mức từ dưới 600V hoặc hơn 15kV	Điện áp định mức từ dưới 600V hoặc hơn 15kV
Tỉ lệ các thông số thành phần đối xứng	Không áp dụng	Không áp dụng	$\frac{R_0}{X_1} > 100$ $\frac{R_0}{X_1} \leq 1$	$\frac{X_0}{X_1} > 10$ $\frac{R_0}{X_0} < 2$	$0 < \frac{X_0}{X_1} < 10$ $\frac{R_0}{X_0} \geq 2$	$3 \leq \frac{X_0}{X_1} \leq 10$ $0 \leq \frac{R_0}{X_1} \leq 1$	$0 \leq \frac{X_0}{X_1} \leq 1$ $0 \leq \frac{R_0}{X_1} \leq 0,1$
Tỉ lệ dòng sự cố LG	$\frac{I_{LG}}{I_{3\phi}} < 1\%$	$\frac{I_{LG}}{I_{3\phi}} < 1\%$	$\frac{I_{LG}}{I_{3\phi}} < 1\%$	$5\% \leq \frac{I_{LG}}{I_{3\phi}} \leq 25\%$	$\frac{I_{LG}}{I_{3\phi}} \leq 25\%$	$25\% \leq \frac{I_{LG}}{I_{3\phi}} \leq 100\%$	$\frac{I_{LG}}{I_{3\phi}} > 95\%$
Bảo vệ khi xảy ra sự cố LG	Rơ-le bảo vệ quá áp (59N)	Không áp dụng	Rơ-le bảo vệ quá áp (59N)	Không áp dụng	Rơ-le bảo vệ quá dòng chạm đất có thời gian (51G)	Rơ-le bảo vệ quá dòng chạm đất có thời gian (51G)	Rơ-le bảo vệ quá dòng chạm đất có thời gian (51G)
Yêu cầu điện áp cách điện	V_{LL}	V_{LL}	V_{LL}	V_{LL}	V_{LN}	V_{LN}	V_{LN}
Điện áp quá độ khi xảy ra sự cố LG	$V_{LG} > 3$	$V_{LG} \leq 2,73$	$V_{LG} \leq 2,73$	$V_{LG} \leq 2,73$	$V_{LG} \leq 2,5$	$V_{LG} \leq 2,3$	$V_{LG} \leq 1,5$
Thiết bị chống sét	Dây trung tính không được nối đất	Dây trung tính không được nối đất	Dây trung tính không được nối đất	Dây trung tính không được nối đất	Dây trung tính không được nối đất	Dây trung tính được nối đất nếu dòng sự cố LG là 60% hoặc hơn	Dây trung tính được nối đất
Chi phí bảo trì	Cao	Thấp	Thấp	Hợp lý	Hợp lý	Hợp lý	Hợp lý
Khôi phục cung cấp điện khi xảy ra sự cố LG	Có thể, nhưng không khuyến khích	Tốt	Tốt	Không thể	Không thể	Không thể	Không thể
Tính an toàn	Thấp	Phù hợp	Phù hợp	Phù hợp	Phù hợp	Phù hợp	Thấp
Khả năng chịu nhiệt	Thấp	Thấp	Thấp	Thấp	Cao	Cao	Cao
Gây nhiễu cho các đường	Trung bình	Thấp	Thấp	Thấp	Cao	Cao	Cao

dây liên lạc							
Biên độ của dòng điện cảm ứng	Không áp dụng	Không đáng kể	Không đáng kể, nhỏ	Đánh giá cao	Đánh giá cao	Đánh giá cao	Đánh giá cao
Phối hợp bảo vệ	Khó	Rất tốt	Rất tốt	Tốt	Tốt	Tốt	Tốt
Độ nhạy	Trung bình	Trung bình	Trung bình	Tốt	Tốt	Tốt	Tốt

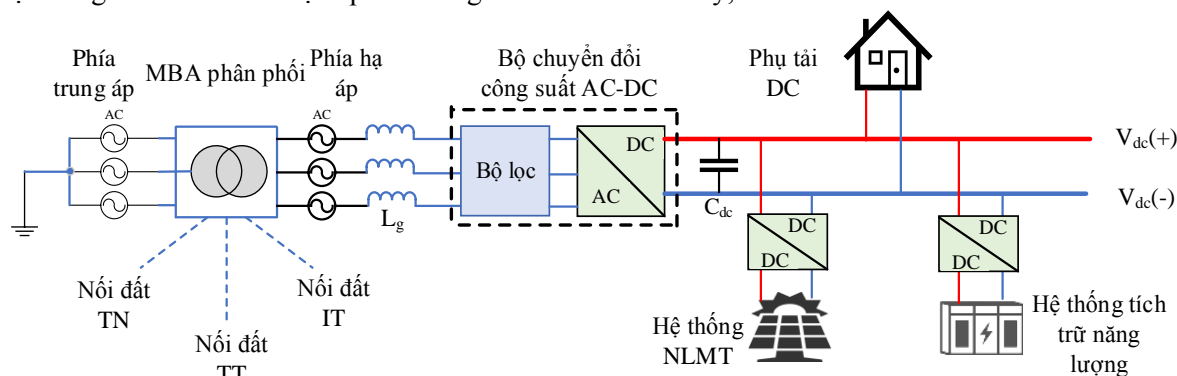
4.2. Cấu hình nối đất sử dụng cho lưới điện DCMG

Để thiết kế hệ thống nối đất cho lưới điện DCMG cần thỏa mãn hai yêu cầu sau: 1) giảm thiểu dòng điện DC rò, và 2) đảm bảo an toàn tối đa cho người sử dụng/thiết bị. Việc lựa chọn cấu hình nối đất thích hợp cho hệ thống DCMG phức tạp hơn nhiều so với hệ thống ACMG. Một số bài báo [34-36] đã thảo luận về các phương pháp nối đất khả thi cho hệ thống điện một chiều. Một số bài báo khác đã nêu tầm quan trọng của cấu hình nối đất lưới điện AC trên hệ thống DC [37][38]. Trên thực tế, ta có thể thấy nhiều phương án nối đất cho lưới điện DCMG như sau:

- Không kết nối với đất;
- Nối đất tại điểm trung tâm của nguồn;
- Nối đất qua điện trở cao;

Ngoài ra, để hệ thống nối đất có thể vận hành tốt, ta cần phải xem xét đến khuyến cáo/hướng dẫn của nhà sản xuất đối với từng loại nguồn để đảm bảo cách bố trí nối đất thích hợp nhất.

Tương tự như lưới điện ACMG, cấu hình nối đất của lưới điện DCMG chắc chắn ảnh hưởng đến việc phát hiện các sự cố nối đất. Hệ thống DCMG liên kết với lưới điện xoay chiều thông qua các bộ chuyển đổi công suất AC/DC. Các bộ chuyển đổi công suất này phải đảm bảo yêu cầu về chống nhiễu điện từ và sóng hài. Vì vậy, chúng thường được tích hợp những bộ lọc tần số cao/thấp ở phía AC và/hoặc DC. Tụ liên kết DC được dùng để điều chỉnh điện áp DC trong cấu hình nối đất này, như Hình 12.



Hình 12: Sơ đồ hệ thống DCMG với hệ thống nối đất phía lưới điện xoay chiều

4.3. Cấu hình nối đất sử dụng cho lưới điện HMG

Lưới điện HMG là sự kết hợp giữa ACMG và DCMG nên hệ thống nối đất của dạng lưới điện này là sự tổng hợp của hai cấu hình nối đất ACMG và DCMG. Theo đó, các yêu cầu các vấn đề về kỹ thuật sẽ được phân chia theo hệ thống nối đất dành cho khu vực AC và hệ thống nối đất dành cho khu vực DC. Đối với khu vực AC, yêu cầu khả năng đảm bảo các đặc tính hoạt động phù hợp của rơ-le bảo vệ, khả năng cách điện của hệ thống và tính đảm bảo an toàn cho con người. Đối với khu vực DC, yêu cầu giảm thiểu dòng điện DC rò và an toàn tối đa cho người sử dụng/thiết bị. Tóm lại, việc xây dựng hệ thống nối đất cho cấu hình lưới điện HMG có thể dựa trên các cấu hình nối đất của lưới điện dạng ACMG và DCMG. Để đảm bảo các yêu cầu về kỹ thuật và tính an toàn cho hệ thống, ta cần thực hiện thiết kế cấu hình nối đất và đề xuất những phương án phù hợp theo yêu cầu vận hành MG.

5. KẾT LUẬN

Hệ thống lưới điện MG được xem là một giải pháp phù hợp hướng tới việc tích hợp và sử dụng các nguồn từ DER vào lưới điện chính. Trong bài báo này, các cấu hình vận hành MG khác nhau đã được trình bày, bao gồm ba loại chính: ACMG, DCMG và HMG. Cấu trúc ACMG là cấu hình phổ biến nhất do có thể kết nối trực tiếp với các phụ tải đến lưới điện chính ở chế độ nối lưới. Có thể thấy rằng, độ tin cậy vận hành của cấu hình ACMG cũng phụ thuộc khá nhiều vào các quá trình chuyển đổi năng lượng xảy ra trong các bộ chuyển đổi công suất. Theo nhận định này, cấu hình lưới điện HMG có số lượng bộ chuyển đổi thấp nhất vì các DER và phụ tải có thể được kết nối với bộ nghịch lưu/chỉnh lưu tương thích với đặc tính của chúng. Ngoài ra, để có được nguồn điện chất lượng cao cho các phụ tải thì cấu hình lưới điện DCMG là phù hợp nhất bởi vì điện áp ở thanh cái được tạo ra bằng thiết bị điện tử công suất, do đó hoàn toàn độc lập với tần số và điện áp của lưới điện chính. Bên cạnh đó, cấu hình lưới điện DCMG cung cấp khả năng quản lý dễ dàng các nguồn năng lượng của ESS. Bài báo cũng đã trình bày chi tiết các mô hình điều khiển cho cả ba loại lưới điện MG bao gồm ACMG, DCMG và HMG. Trong đó, điều khiển theo mô hình phân cấp giúp lưới điện MG đảm bảo độ tin cậy bằng cách duy trì điện áp, tần số ổn định và điều chỉnh công suất hợp lý. Cuối cùng, các cấu hình nối đất cũng đã được trình bày cụ thể trong bài báo này. Theo đó, để đảm bảo tính an toàn trong khi vận hành và cho người dùng, cần có phương án nối đất phù hợp với mỗi cấu hình lưới điện MG.

TÀI LIỆU THAM KHẢO

- [1] T. S. Usstun, C. Ozansoy, and A. Zayegh, Recent developments in microgrid and example cases around the world – A review, *Renew. Sustain. Energy Rev.*, vol. 15, no. 8, pp. 4030-4041, Oct. 2011.
- [2] A. Haytham and Vijay K. Sood, A Comprehensive Review on Microgrid Architectures for Distributed Generation, *IEEE Electrical Power and Energy Conference (EPEC)*, 2019.
- [3] G. Platt, A. Berry, and D. Cornforth, What role for microgrid?, in *Smart grid: integrating renewable, distributed & efficient energy*, Elsevier/Academic Press, pp. 185-207, 2012.
- [4] N. Hatziargyriou, Microgrid: architectures and control, *John Wiley and Sons Ltd. IEEE Press*, 2014.
- [5] Consultant report: Integration of distributed energy resources, *The CERTS MicroGrid concept. in California Energy Commission*, 2013.
- [6] I. Patrao, E. Figueres, G. Garcerá, and R. González-Medina, Microgrid architectures for low voltage distributed generation, *Renew. Sustain. Energy Rev.*, vol. 43, pp. 415–424, Mar. 2015
- [7] J. J. Justo, F. Mwasilu, J. Lee, and J.-W. Jung, AC-microgrids versus DC-microgrids with distributed energy resources: A review, *Renew. Sustain. Energy Rev.*, vol. 24, pp. 387–405, Aug. 2013.
- [8] H. Gabbar, Smart Energy Grid Engineering. *Elsevier Inc.*, 2016.
- [9] E. Planas, J. Andreu, J. I. Gárate, I. Martínez de Alegría, and E. Ibarra, AC and DC technology in microgrids: A review, *Renew. Sustain. Energy Rev.*, vol. 43, pp. 726–749, Mar. 2015.
- [10] E. Unamuno and J. A. Barrena, Hybrid ac/dc microgrids—Part I: Review and classification of topologies, *Renew. Sustain. Energy Rev.*, vol. 52, pp. 1251–1259, Dec. 2015.
- [11] A. Bidram, V. Nasirian, A. Davoudi, and F. L. Lewis, Control and modeling of microgrids, in *Cooperative Synchronization in Distributed Microgrid Control*. Cham, Switzerland: Springer, pp. 7–43, Feb. 2017.
- [12] F. Katiraei and M. R. Iravani, Power management strategies for a microgrid with multiple distributed generation units, *IEEE Trans. Power Syst.*, vol. 21, no. 4, pp. 1821–1831, Nov. 2006.
- [13] A. Bidram and A. Davoudi, Hierarchical structure of microgrids control system, *IEEE Trans. Smart Grid*, vol. 3, no. 4, pp. 1963–1976, Dec. 2012.

- [14] M. Savaghebi, A. Jalilian, J. C. Vasquez, and J. M. Guerrero, Secondary control scheme for voltage unbalance compensation in an islanded droop-controlled microgrid, *IEEE Trans. Smart Grid*, vol. 3, no. 2, pp. 797–807, Jun. 2012.
- [15] E. Barklund, N. Pogaku, M. Prodanovic, C. Hernandez-Aramburo, and T. C. Green, Energy management in autonomous microgrid using stability-constrained droop control of inverters, *IEEE Trans. Power Electron.*, vol. 23, no. 5, pp. 2346–2352, Sep. 2008.
- [16] A. Pantoja and N. Quijano, A population dynamics approach for the dispatch of distributed generators, *IEEE Trans. Ind. Electron.*, vol. 58, no. 10, pp. 4559–4567, Oct. 2011.
- [17] T. L. Vandoorn, J. D. M. De Kooning, B. Meersman, and L. Vandeveldel, Review of primary control strategies for islanded microgrids with power-electronic interfaces, *Renew. Sustain. Energy Rev.*, vol. 19, pp. 613–628, Mar. 2013.
- [18] T. F. Wu, K. Siri, and J. Banda, The central-limit control and impact of cable resistance in current distribution for parallel-connected DC-DC converters, in *Proc. Power Electron. Specialist Conf. (PESC)*, vol. 1, Taipei, Taiwan, pp. 694–702, 1994.
- [19] J. Banda and K. Siri, Improved central-limit control for paralleloperation of DC-DC power converters, in *Proc. IEEE Power Electron. Specialist Conf. (PESC)*, vol. 2, Atlanta, GA, USA, pp. 1104–1110, Aug. 1995.
- [20] T. Caldognetto and P. Tenti, Microgrids operation based on master–slave cooperative control, *IEEE J. Emerg. Sel. Topics Power Electron.*, vol. 2, no. 4, pp. 1081–1088, Dec. 2014.
- [21] J. Tan, H. Lin, J. Zhang, and J. Ying, A novel load sharing control technique for paralleled inverters, in *Proc. IEEE 34th Annu. Conf. Power Electron. Specialist (PESC)*, vol. 3, Acapulco, Mexico, pp. 1432–1437, Aug. 2003.
- [22] F. Valenciaga and P. F. Puleston, Supervisor control for a stand-alone hybrid generation system using wind and photovoltaic energy, *IEEE Trans. Energy Convers.*, vol. 20, no. 2, pp. 398–405, Jun. 2005.
- [23] F. Valenciaga and P. F. Puleston, High-order sliding control for a wind energy conversion system based on a permanent magnet synchronous generator, *IEEE Trans. Energy Convers.*, vol. 23, no. 3, pp. 860–867, Sep. 2008
- [24] T. Dragičević, X. Lu, J. C. Vasquez, and J. M. Guerrero, Dc microgrids—Part I: A review of control strategies and stabilization techniques, *IEEE Trans. Power Electron.*, vol. 31, no. 7, pp. 4876–4891, Sep. 2015
- [25] M. Ahmed, A. Vahidnia, L. Meegahapola, and M. Datta, Impact of multiple motor loads on dynamic performance and stability of microgrids, in *Proc. IEEE Int. Conf. Ind. Tech. (ICIT)*, Melbourne, VIC, Australia, pp. 1704–1709, Feb. 2019.
- [26] N. Eghtedarpour and E. Farjah, Power control and management in a hybrid AC/DC microgrid, *IEEE Trans. Smart Grid*, vol. 5, no. 3, pp. 1494–1505, May 2014.
- [27] E. Planas, J. Andreu, J. I. Gárate, I. M. de Alegría, and E. Ibarra, AC and DC technology in microgrids: A review, *Renew. Sustain. Energy Rev.*, vol. 43, pp. 726–749, Mar. 2015.
- [28] Patrao, E. Figueres, G. Garcerán, and R. González-Medina, Microgrid architectures for low voltage distributed generation, *Renew. Sustain Energy Rev.*, vol. 43, pp. 415–424, Mar. 2015.
- [29] Application Guide for IEEE STD 1547(TM), IEEE standard for interconnecting distributed resources with electric power systems, *IEEE STD 1547.2-2008*, pp.1-217, Apr. 2009

- [30] S. Teimourzadeh, F. Aminifar, M. Davarpanah and J. M. Guerrero, Macroprotections for Microgrids: Toward a new protection paradigm subsequent to distributed energy resource integration, *IEEE Ind. Electr. Magazine*, vol. 10, no. 3, pp. 6-18, Sept. 2016.
- [31] D. M. Bui, S. L. Chen, K. Y. Lien, Y. R. Chang, Y. D. Lee, and J. L. Jiang, Investigation on transient behaviors of a uni-grounded low voltage AC microgrid and evaluation on its available fault protection methods: Review and proposals, Elsevier, *Renewable and Sustainable Energy Reviews*. vol. 75, pp. 1417–1452, 2017.
- [32] D. M. Bui, S. L. Chen, K. Y. Lien, and J. L. Jiang, A generalized fault protection structure for uni-grounded low-voltage AC microgrids, *2015 IEEE Innovative Smart Grid Technologies - Asia (ISGT ASIA)*, Bangkok, pp. 1-6, 2015.
- [33] J. Mohammadi, F. Badrkhani Ajaei and G. Stevens, Grounding the AC Microgrid, in *IEEE Transactions on Industry Applications*, vol. 55, no. 1, pp. 98-105, Jan.-Feb. 2019, doi: 10.1109/TIA.2018.2864106.
- [34] D. Paul, DC traction power system grounding, *IEEE Trans. Ind. Appl.*, vol. 38, no. 3, pp. 818–824, May 2002.
- [35] M. Noritake, T. Lino, A. Fukui, K. Hirose, and M. Yamasaki, A study of the safety of the DC 400 V distribution system, in *Proc. IEEE 31st Int. Telecommun. Energy Conf. (INTELEC)*, pp. 1–8, Oct. 2009.
- [36] K. Hirose et al., Grounding concept considerations and recommendations for 400VDC distribution system, in *Proc. IEEE 33rd Int. Telecommun. Energy Conf. (INTELEC)*, pp. 1–8, Oct. 2011.
- [37] K. Xing, F. C. Lee, J. S. Lai, T. Gurjit, and D. Borojevic, Adjustable speed drive neutral voltage shift and grounding issues in a DC distributed system, in *Proc. IEEE-IAS Annu. Meet.*, pp. 517–524, Oct. 1997.
- [38] M. E. Baran and N. R. Mahajan, DC distribution for industrial systems: Opportunities and challenges, *IEEE Trans. Ind. Appl.*, vol. 39, no. 6, pp. 1596–1601, Nov. 2003.

YSC4F.214

PHÂN TÍCH XÁC SUẤT DỪNG CỦA MẠNG HỢP TÁC HAI CHIỀU CÓ THU HOẠCH NĂNG LƯỢNG VÔ TUYẾN VÀ KỸ THUẬT MÃ HÓA MẠNG SỐ

NGUYỄN LINH TÂM¹, ĐỖ NGỌC THANH², ĐÀO THỊ THU THÚY³

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

¹ tamlinh1201@gmail.com, ² thanh dabac2000@gmail.com, ³ daothithuthuy@iuh.edu.vn

Tóm tắt. Những năm gần đây, việc thu thập năng lượng thông qua tần số vô tuyến là một phương pháp đang được quan tâm nghiên cứu. Bài báo này nghiên cứu một mô hình mạng hợp tác hai chiều có khả năng truyền năng lượng và tín hiệu đồng thời gồm hai nguồn và một thiết bị chuyển tiếp giải mã-chuyển tiếp. Mô hình sử dụng phương pháp thu thập năng lượng phân chia theo công suất và kỹ thuật mã hóa mạng số với kênh truyền Rayleigh fading và có xem xét tồn tại đường truyền trực tiếp giữa hai nguồn. Xác suất dừng của hai nguồn được phân tích để đánh giá hiệu năng của hệ thống và các biểu thức toán học được kiểm chứng bằng phương pháp mô phỏng Monte-Carlo sử dụng phần mềm Matlab. Đồng thời các kết quả của mô hình được đề xuất cũng được so sánh với mô hình không có đường truyền trực tiếp giữa hai nguồn.

Từ khóa. Thu thập năng lượng, mã hóa mạng số, giải mã và chuyển tiếp, xác suất dừng, đường truyền trực tiếp.

OUTAGE PROBABILITY ANALYSIS OF TWO-WAY COOPERATIVE NETWORK WITH WIRELESS ENERGY HARVESTING AND DIGITAL NETWORK ENCODING

Abstract. In recent years, energy harvesting (EH) through radio frequency (RF) is a method of interest to research. This paper studies a two-way cooperation network capable of transmitting energy and signal simultaneously consists of two sources and a relay decode-forward (DF). The model uses power-splitting protocol (PS) to energy harvesting and digital network encoding techniques with Rayleigh fading and considers the existence of a direct path between the two sources. The system outage probabilities of two sources are analyzed to evaluate the performance of the system and the mathematical expressions are verified by the Monte-Carlo simulation method using Matlab software. Besides, the results of proposed model are also compared with the model without direct links between the two sources.

Keywords. Energy harvesting, digital network encoding, decoding and forwarding, the outage probability, direct link.

1. GIỚI THIỆU

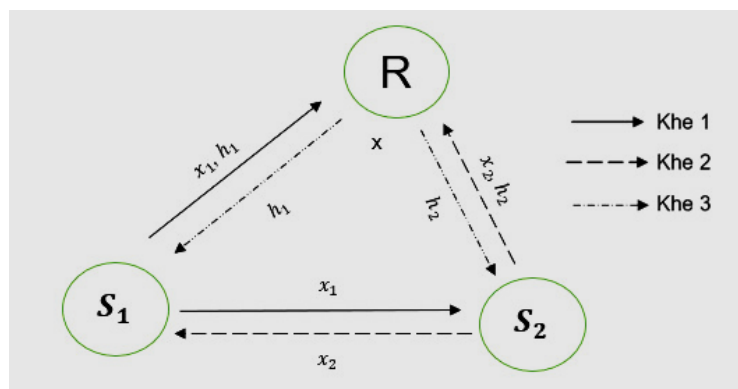
Mạng hợp tác hai chiều hay mạng vô tuyến chuyển tiếp hai chiều là một giải pháp giúp cải thiện thông lượng điểm truy cập mạng, cải thiện chất lượng phủ sóng tín hiệu. Hiện nay, thu thập năng lượng vô tuyến đang được nhiều nhà khoa học quan tâm nghiên cứu trong tình hình bùng nổ các thiết bị sử dụng mạng truyền thông vô tuyến. Việc thu thập năng lượng vô tuyến cho phép các nút chuyển tiếp có thể thu thập và sử dụng được năng lượng từ nguồn phát vô tuyến để cải thiện tốc độ và chất lượng tín hiệu truyền dẫn, tăng cường năng lượng cho các nút chuyển tiếp trong hệ thống mạng chuyển tiếp. Có 2 giao thức thu thập năng lượng: giao thức phân chia theo thời gian (TS) và giao thức phân chia theo công suất (PS) [1][2]. Bên cạnh đó kỹ thuật mã hóa mạng số (DNC) cũng được sử dụng để cải thiện hiệu năng của mạng hợp tác hai chiều [3].

Trong bài báo này, chúng tôi xem xét và phân tích mô hình mạng hợp tác hai chiều có một thiết bị chuyển tiếp có khả năng thu hoạch năng lượng. Giao thức phân chia theo công suất được sử dụng để thiết bị chuyển

tiếp thu thập năng lượng từ tín hiệu vô tuyến từ hai nguồn. Sau đó thiết bị chuyển tiếp dùng kỹ thuật DNC để tạo một tín hiệu mới và sử dụng năng lượng thu được để chuyển tiếp tín hiệu. Bài báo có xem xét xác suất dừng của hệ thống trong trường hợp có tồn tại đường truyền trực tiếp và không tồn tại đường truyền trực tiếp giữa hai nguồn.

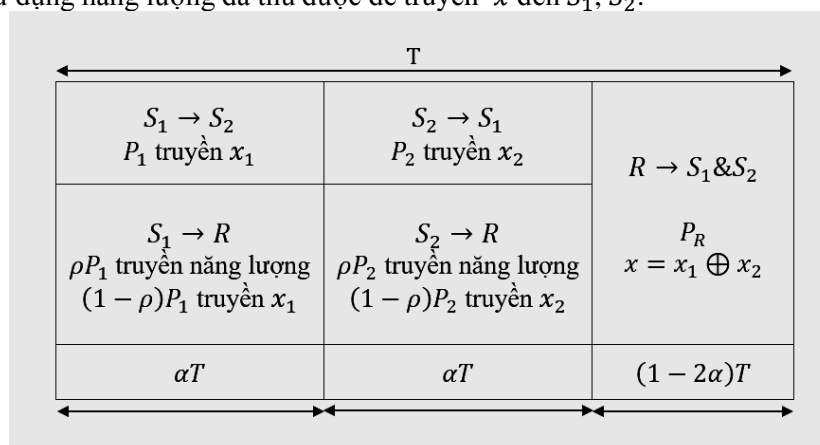
Trong các phần tiếp theo, chúng tôi trình bày sơ đồ hệ thống mạng chuyển tiếp hai chiều có đường trực tiếp giữa hai nguồn, tính toán quá trình truyền nhận, phân tích hiệu năng của từng nguồn. Sau đó đưa ra mô phỏng và các nhận xét của chúng tôi về mô hình. Và cuối cùng là phần kết luận.

2. MÔ HÌNH NGHIÊN CỨU



Hình 1: Sơ đồ hệ thống

Sơ đồ hệ thống bao gồm hai nguồn S_1, S_2 và một thiết bị chuyển tiếp R . Hai nguồn S_1, S_2 truyền tín hiệu x_1, x_2 cho nhau thông qua R và đường trực tiếp. Hệ thống hoạt động ở chế độ bán song công. Ngoài ra thiết bị chuyển tiếp có khả năng thu thập năng lượng từ hai nguồn sử dụng phương pháp phân chia theo công suất. Quá trình truyền nhận tín hiệu gồm ba khe thời gian như hình 2. Ở hai khe thời gian đầu, nút nguồn S_1, S_2 truyền tín hiệu đến relay đồng thời truyền trực tiếp đến người dùng còn lại. Ở khe thời gian thứ ba thiết bị chuyển tiếp sử dụng kỹ thuật mã hóa mạng số kết hợp hai tín hiệu thu được thành tín hiệu mới $x = (x_1 \oplus x_2)$ và R sử dụng năng lượng đã thu được để truyền x đến S_1, S_2 .



Hình 2: Sơ đồ khối truyền nhận tín hiệu của mô hình

Trong đó T là tổng thời lượng của một khối truyền đơn và $0 < \alpha < 0.5$ là hệ số phân chia khe thời gian. Các quá trình chuyển tiếp tín hiệu và năng lượng trong hai khe thời gian đầu bằng nhau và bằng αT , thời gian của giai đoạn cuối bằng $(1 - 2\alpha)T$. ρ là hệ số phân chia công suất ($0 < \rho < 1$). P_1, P_2 là công suất truyền của hai nguồn.

Ở bài báo này chúng tôi sử dụng kênh truyền Rayleigh [3][4], các hệ số kênh truyền và các khoảng cách giữa (S_1 và R), (S_2 và R), (S_1 và S_2) được kí hiệu h_1, h_2, h và d_1, d_2, d tương ứng. Độ lợi kênh truyền $\varphi_1 = |h_1|^2, \varphi_2 = |h_2|^2, \varphi = |h|^2$ là các biến ngẫu nhiên có hàm phân phối tích lũy (CDF) và hàm mật độ

xác suất (PDF): $F_x(x) = 1 - \exp\left(\frac{-x}{\lambda_k}\right)$, $f(x) = \frac{1}{\lambda_k} \exp\left(\frac{-x}{\lambda_k}\right)$ với $\lambda_k \in \{\lambda_1, \lambda_2, \lambda\}$, $\lambda_k = d_k^{-\beta}$, và β là hệ số mũ suy hao đường truyền [3].

2.1. Quá trình truyền nhận và tỉ lệ tín hiệu trên nhiễu

Tại khe thời gian đầu: S_1 truyền tín hiệu x_1 và năng lượng đến thiết bị chuyển tiếp và đồng thời truyền tín hiệu x_1 truyền trực tiếp đến S_2 . Tín hiệu nhận được ở thiết bị chuyển tiếp.

$$y_{1,R} = \sqrt{P_1} h_1 x_1 + \eta_R. \quad (1)$$

η_R là nhiễu trắng có phân bố Gaussian (AWGN) với trị trung bình bằng 0 và phương sai N_0

Tỉ lệ tín hiệu trên nhiễu (SNR) tại thiết bị chuyển tiếp để giải mã tín hiệu x_1 :

$$\gamma_{1,R} = \frac{(1-\rho)|h_1|^2 P_1}{N_0} = (1-\rho)|h_1|^2 \psi = (1-\rho)\phi_1 \psi. \quad (2)$$

Năng lượng mà thiết bị chuyển tiếp thu thập từ tín hiệu RF của S_1 .

$$E_{h1} = \eta \rho \alpha P_1 |h_1|^2 T. \quad (3)$$

Xét kênh truyền trực tiếp, tín hiệu nhận được và SNR tại S_2 :

$$y_{1,2} = \sqrt{P_1} h x_1 + \eta_2. \quad (4)$$

$$\gamma_{1,2} = \frac{P_1 |h|^2}{N_0}. \quad (5)$$

Với η_R, η_1, η_2 là nhiễu trắng có phân bố Gaussian (AWGN) với trị trung bình bằng 0 và phương sai N_0

Tương tự tại khe thời gian thứ hai: S_2 truyền tín hiệu x_2 và năng lượng đến relay và đồng thời truyền tín hiệu x_2 trực tiếp đến S_1 . Tín hiệu nhận được và SNR ở thiết bị chuyển tiếp.

$$\gamma_{2,R} = h_2 x_2 + \eta_R, \quad (6)$$

$$\gamma_{2,R} = \frac{(1-\rho)|h_2|^2 P_1}{N_0} = (1-\rho)|h_2|^2 \psi = (1-\rho)\phi_2 \psi. \quad (7)$$

Năng lượng mà thiết bị chuyển tiếp thu thập từ tín hiệu RF của S_2

$$E_{h2} = \eta \rho \alpha P_2 |h_2|^2 T. \quad (8)$$

Xét kênh truyền trực tiếp, tín hiệu nhận được và SNR tại S_1

$$\gamma_{2,1} = \sqrt{P_2} h x_2 + \eta_1. \quad (9)$$

$$\gamma_{2,1} = \frac{P_2 |h|^2}{N_0}. \quad (10)$$

Tổng 2 năng lượng thu được tại nút chuyển tiếp R do S_1 và S_2 cung cấp:

$$E_h = E_{h_1} + E_{h_2} = \eta\rho\alpha(P_1|h_1|^2 + P_2|h_2|^2)T. \quad (11)$$

Chọn $P_1 = P_2 = P$, công suất phát trung bình của nút chuyển tiếp trong khe thời gian thứ ba là:

$$P_R = \frac{E_h}{(1-2\alpha)T} = \frac{\eta\rho\alpha}{1-2\alpha} P(|h_1|^2 + |h_2|^2). \quad (12)$$

Tại khe thời gian thứ ba: Thiết bị chuyển tiếp sử dụng năng lượng nhận được từ S_1, S_2 trong khe 1 và khe 2 truyền tín hiệu x_1, x_2 để giải mã và chuyển tiếp $x = (x_1 \oplus x_2)$ trở về hai nguồn. Tín hiệu nhận được ở S_1 và S_2 tương ứng là:

$$y_{R,1} = \sqrt{P_R}h_1(x_1 \oplus x_2) + \eta_1. \quad (13)$$

$$y_{R,2} = \sqrt{P_R}h_2(x_1 \oplus x_2) + \eta_2. \quad (14)$$

Tỉ lệ tín hiệu trên nhiễu thu để giải mã tín hiệu x tại hai nguồn:

$$\gamma_{R,1} = \frac{P_R|h_1|^2}{N_0}. \quad (15)$$

$$\gamma_{R,2} = \frac{P_R|h_2|^2}{N_0}. \quad (16)$$

Sử dụng kỹ thuật lựa chọn kết hợp (SC) tại nguồn thu khi có 2 tín hiệu đến từ đường trực tiếp và qua thiết bị chuyển tiếp. Tỉ lệ tín hiệu trên nhiễu kết hợp để giải mã các tín hiệu tại S_1, S_2 là:

$$\gamma_{S_1} = \max\left\{\gamma_{2,1}, \min\left\{\gamma_{2,R}, \gamma_{R,1}\right\}\right\}, \quad (17)$$

$$\gamma_{S_2} = \max\left\{\gamma_{1,2}, \min\left\{\gamma_{1,R}, \gamma_{R,2}\right\}\right\}, \quad (18)$$

2.2. Biểu thức xác suất dừng của hai nguồn

Trong phần này tìm biểu thức xác suất dừng của từng nguồn. Xác suất dừng xảy ra tại nguồn thu khi cả hai đường truyền trực tiếp và truyền qua thiết bị chuyển tiếp đều không thành công. Hay tỉ lệ tín hiệu trên nhiễu kết hợp tại nguồn thu nhỏ hơn giá trị ngưỡng γ_t . Cụ thể xác suất dừng tại S_1 được biểu diễn như sau:

$$OP_{S_1} = \Pr[\gamma_{S_1} < \gamma_t] = \left\{ \Pr[\gamma_{2,R} < \gamma_t] + \Pr[\gamma_{2,R} \geq \gamma_t, \gamma_{R,1} < \gamma_t] \right\} \Pr[\gamma_{2,1} < \gamma_t], \quad (19)$$

Sau nhiều phép tính phức tạp ta được biểu thức xác suất dừng tại nguồn S_1 như sau:

$$\begin{aligned} OP_{S_1} &= \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda_2}\right) + \exp\left(-\frac{\tilde{\gamma}_t}{\lambda_1\psi}\right) \left[1 - \exp\left(-\frac{1-\rho}{\lambda_1\kappa}\right) \right] + \frac{\lambda_2}{\lambda_1 - \lambda_2} \right. \\ &\times \frac{1}{\lambda_1} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_2\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_2\kappa\psi}\right) - \frac{\lambda_1}{\lambda_1 - \lambda_2} \exp\left[-\left(\frac{1}{\lambda_2} - \frac{1}{\lambda_1}\right) \frac{\tilde{\gamma}_t}{\psi}\right] \\ &\times \frac{1}{\lambda_1} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_1\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_1\kappa\psi}\right) \left. \right\} \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda}\right) \right\}. \end{aligned} \quad (20)$$

với $\kappa = \frac{\eta\rho\alpha}{1-2\alpha}$, $\tilde{\gamma}_t = \frac{\gamma_t}{1-\rho}$ và $\Gamma(\alpha, x)$ là hàm gamma không đầy đủ [5].

Tương tự biểu thức xác suất dừng tại S_2 như sau:

$$\begin{aligned}
 OP_{S_2} &= \left\{ \underbrace{\Pr[\gamma_{1,R} < \gamma_t]}_{H_4} + \underbrace{\Pr[\gamma_{1,R} \geq \gamma_t, \gamma_{R,2} < \gamma_t]}_{H_5} \right\} \underbrace{\Pr[\gamma_{1,2} < \gamma_t]}_{H_6}, \\
 OP_{S_2} &= \left\{ \Pr[\gamma_{1,R} < \gamma_t] + \Pr[\gamma_{1,R} \geq \gamma_t, \gamma_{R,2} < \gamma_t] \right\} \Pr[\gamma_{1,2} < \gamma_t] \\
 &= \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda_1}\right) + \exp\left(-\frac{\tilde{\gamma}_t}{\lambda_1\psi}\right) \left[1 - \exp\left(-\frac{1}{\lambda_2} \frac{1-\rho}{\kappa}\right) \right] + \frac{\lambda_1}{\lambda_2 - \lambda_1} \right. \\
 &\quad \times \frac{1}{\lambda_2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_1\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_1\kappa\psi}\right) - \frac{\lambda_2}{\lambda_2 - \lambda_1} \exp\left[-\left(\frac{1}{\lambda_1} - \frac{1}{\lambda_2}\right) \frac{\tilde{\gamma}_t}{\psi}\right] \\
 &\quad \left. \times \frac{1}{\lambda_2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_2\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_2\kappa\psi}\right) \right\} \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda}\right) \right\}.
 \end{aligned} \tag{21}$$

Xét hai trường hợp $\lambda_1 \neq \lambda_2$ và $\lambda_1 = \lambda_2$ để ta xem xét xác suất dừng tại S_1, S_2 :

- Xét trường hợp $\lambda_1 \neq \lambda_2$ thì xác suất dừng tại S_1 được tính (20) và xác suất dừng tại S_2 được tính (21).
- Xét trường hợp $\lambda_1 = \lambda_2$ thì:

Từ (20) xác suất dừng tại S_1 được tính như sau:

$$\begin{aligned}
 OP_{S_1} &= \{H_1 + H_2\} H_3 \\
 &= \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda_2}\right) + \left[\exp\left(-\frac{\tilde{\gamma}_t}{\lambda_1\psi}\right) \left[1 - \exp\left(-\frac{1}{\lambda_1} \frac{1-\rho}{\kappa}\right) \right] \right. \right. \\
 &\quad \left. \left. + \frac{1}{\lambda_1} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_2\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_2\kappa\psi}\right) \right. \right. \\
 &\quad \left. \left. - \frac{1}{\lambda_1} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_1\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_1\kappa\psi}\right) \right\} 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda}\right)
 \end{aligned} \tag{22}$$

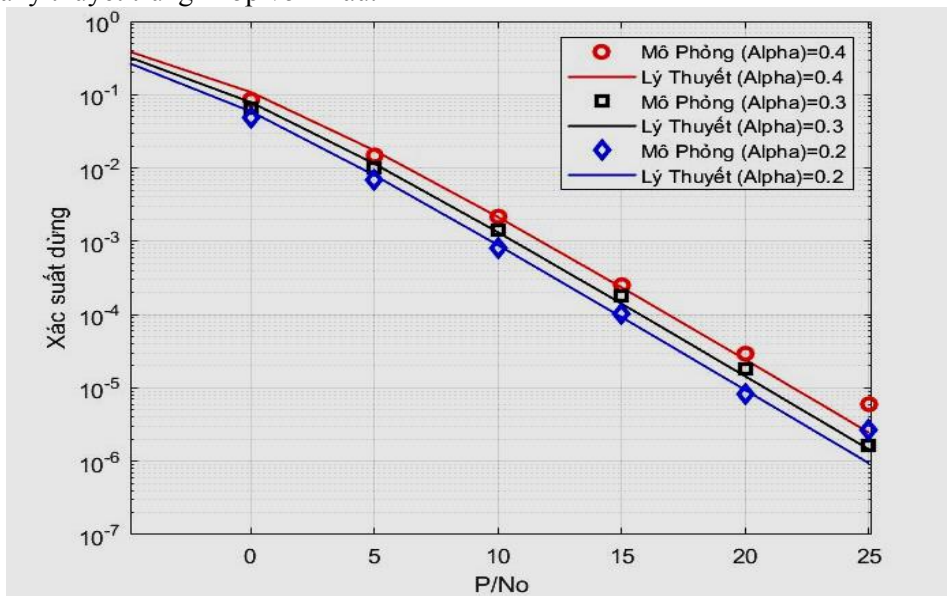
Từ (21) xác suất dừng tại S_2 được tính như sau:

$$\begin{aligned}
 OP_{S_2} &= \{H_4 + H_5\} H_6 \\
 &= \left\{ 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda_1}\right) + \left[\exp\left(-\frac{\tilde{\gamma}_t}{\lambda_1\psi}\right) \left[1 - \exp\left(-\frac{1}{\lambda_2} \frac{1-\rho}{\kappa}\right) \right] \right. \right. \\
 &\quad \left. \left. + \frac{1}{\lambda_2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_1\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_1\kappa\psi}\right) \right. \right. \\
 &\quad \left. \left. - \frac{1}{\lambda_2} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \left(\frac{-\gamma_t}{\lambda_2\kappa\psi}\right)^{k+1} \Gamma\left(-k-1, \frac{\kappa}{1-\rho} \frac{\gamma_t}{\lambda_2\kappa\psi}\right) \right\} 1 - \exp\left(-\frac{\tilde{\gamma}_t}{\psi\lambda}\right)
 \end{aligned} \tag{23}$$

3. MÔ PHỎNG VÀ NHẬN XÉT

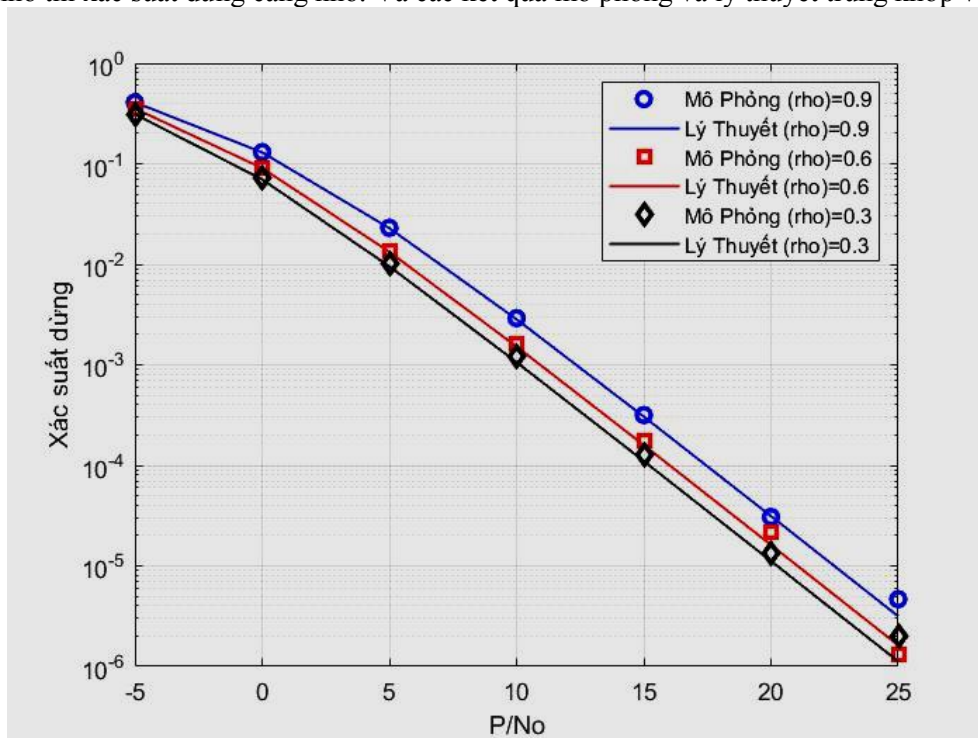
Trong phần này, các thông số được chọn như sau: tọa độ $S_1(0, 0)$, tọa độ $S_2(1, 0)$, tọa độ $R(0.3, 0.5)$ tương ứng $d_1 = 0.583$, $d_2 = 0.86$, $d = 1$ và $\beta = 3$. Do tính đối xứng của mô hình nên các hình biểu diễn xác

suất dừng của nguồn S_1 . Các đường liền nét là biểu thức phân tích, các kí hiệu là biểu thức mô phỏng.
 Hình 3: biểu diễn xác suất dừng của nguồn S_1 theo tỉ lệ công suất phát trên nhiễu P/N_0 (dB) với ba giá trị hệ số phân chia khe thời gian α . Quan sát hình ta thấy P/N_0 càng tăng thì xác suất dừng của hệ thống càng giảm. Bên cạnh đó α nhỏ thì xác suất dừng nhỏ dẫn đến hệ thống hoạt động tốt hơn. Cuối cùng các kết quả mô phỏng và lý thuyết trùng khớp với nhau.



Hình 3: Xác suất dừng khi $R=1$ bps/Hz, $\rho=0.3$, $\alpha=0.2$, $\alpha=0.3$, $\alpha=0.4$

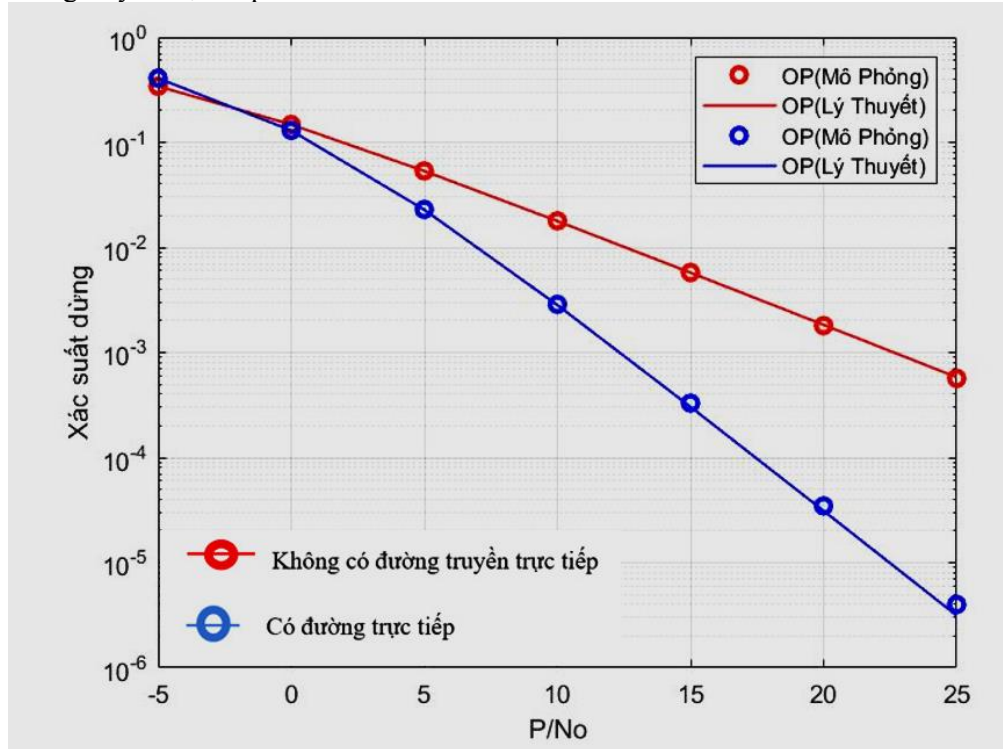
Hình 4: Biểu diễn xác suất dừng của nguồn S_1 theo tỉ lệ công suất phát trên nhiễu P/N_0 (dB) với ba giá trị hệ số phân chia công suất ρ . Quan sát hình ta thấy P/N_0 càng tăng thì xác suất dừng của hệ thống càng giảm như hình 3. Hệ số phân chia công suất thay đổi cũng ảnh hưởng đến xác suất dừng, hệ số phân chia công suất càng nhỏ thì xác suất dừng càng nhỏ. Và các kết quả mô phỏng và lý thuyết trùng khớp với nhau.



Hình 4: Xác suất dừng khi ρ thay đổi ($\rho=0.3$, $\rho=0.6$, $\rho=0.9$), $R=1$ bps/Hz, $\alpha=0.4$

Hình 5: Biểu diễn xác suất dừng của nguồn S_1 theo tỉ lệ công suất phát trên nhiễu P/N_0 (dB) trong hai trường

hợp: thứ nhất là đường màu đỏ không có đường trực tiếp và thứ hai là đường màu xanh có đường truyền trực tiếp. Quan sát hình ta thấy khi giá trị P/N_0 (dB) nhỏ ở đoạn $[-5, -2]$ xác suất dừng của S_1 trong hai trường hợp không chênh lệch nhiều. Tuy nhiên khi giá trị P/N_0 (dB) tăng ở đoạn $[-2, 25]$ thì xác suất dừng trong trường hợp có đường truyền trực tiếp có xu hướng giảm nhanh hơn so với trường hợp không có đường truyền trực tiếp. Điều này cho thấy hiệu năng của hệ thống có đường truyền trực tiếp tốt hơn so với hệ thống không có đường truyền trực tiếp.



Hình 5: So sánh xác suất dừng (OP) khi có đường trực tiếp và không có đường trực tiếp với $\rho=0.4$, $R=0.3$ bps/Hz, $\alpha=0.4$

4. KẾT LUẬN

Bài báo này nghiên cứu một mô hình mạng hợp tác hai chiều có khả năng truyền năng lượng và tín hiệu đồng thời. Mô hình sử dụng phương pháp PS và kỹ thuật DNC với kênh truyền Rayleigh fading và có xem xét tồn tại đường truyền trực tiếp giữa hai nguồn. Các biểu thức toán học tường minh của xác suất dừng của hai nguồn đã được xác định. Và xác suất dừng của hai nguồn được khảo sát theo P/N_0 (dB) với các hệ số phân chia khe thời gian và hệ số phân chia công suất thay đổi. Trong bài báo cũng so sánh xác suất dừng trong trường hợp có và không có đường truyền trực tiếp giữa hai nguồn, kết quả cho thấy khi có đường truyền trực tiếp xác suất dừng của hệ thống nhỏ hơn. Cuối cùng phương pháp mô phỏng Monte-Carlo sử dụng phần mềm Matlab đã chứng minh được tính đúng đắn của biểu thức lý thuyết.

TÀI LIỆU THAM KHẢO

- [1] H.N.Huy, N.T.Hải, Đ.T.T.Thùy, N.T.Luân, and P.T.Tín, "Thu thập năng lượng trong mạng đa truy cập phi trực giao khi có đường trực tiếp," Hội nghị Khoa học trẻ lần 3 năm 2021 (YSC2021) – IUH. ISBN: 978-604-920-123-3, p. tr. 5-10, 2021.
- [2] H.N.Huy, N.T.Hải, Đ.T.T.Thùy, "Thu hoạch năng lượng trong mạng hợp tác hai chiều," Khóa Luận Tốt Nghiệp. Trường Đại Học Công Nghiệp Thành phố Hồ Chí Minh, 2021.
- [3] T.-T. T. Dao and P. N. Son, "Uplink non-orthogonal multiple access protocol in two-way relaying networks: realistic operation and performance analysis," in 2020 7th NAFOSTED Conference on Information and Computer

Science (NICS), 2020, pp. 399-404: IEEE.

[4] M. V. Tan N. Nguyen , Phuong T. Tran, “Wireless energy harvesting meets receiver diversity: A successful approach for two-way half-duplex relay networks over block Rayleigh fading channel. Computer Networks journal,” Computer Networks 172 (2020) 107176. p. tr. 5-12, 2020.

[5] I. S. G. a. I. M. Ryzhik, *Table of integrals, series, and products*, Seventh ed. Elsevier, 2007.

YSC4F.215

HỆ THỐNG GIÁM SÁT CHẤT LƯỢNG NGUỒN NƯỚC XẢ THẢI CÔNG NGHIỆP THEO THỜI GIAN THỰC

DUONG QUOC THANG ¹, NGUYEN TRONG TUAN ², TRAN THI MINH KHOA ^{3*}

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

¹ duongquochang2629@gmail.com, ² nguyentrongtuanfpt@gmail.com, ³ ttmkhoa@iuh.edu.vn

Tóm tắt. Ở nhiều khu công nghiệp, cụm công nghiệp tại Việt Nam, nước thải công nghiệp vẫn là nguyên nhân chính gây ô nhiễm nguồn nước chính tại những con sông. Việc tuân thủ các quy định tại các doanh nghiệp còn mang tính hình thức, việc vận hành hệ thống xử lý nước thải tập trung còn mang tính đối phó, chưa thực sự hiệu quả. Bài báo này trình bày chi tiết một cách tiếp cận việc xây dựng hệ thống giám sát chất lượng nguồn nước thải công nghiệp. Hệ thống được phát triển sử dụng các kỹ thuật vạn vật kết nối kết hợp với các dịch vụ điện toán đám mây. Các chỉ số đo chất lượng nguồn nước thải công nghiệp được liên tục gửi về máy chủ thông qua kết nối mạng không dây. Dữ liệu thu thập sẽ được phân tích tại máy chủ, kết quả phân tích về chất lượng nguồn nước thải, quá trình xả thải có phù hợp với quy định hay không sẽ được thông báo cho người quản trị.

Từ khóa. Mạng lưới vạn vật kết nối, nền tảng đám mây, nước thải công nghiệp, hệ thống giám sát nguồn nước, Microsoft Azure.

AN APPROACH FOR REAL TIME INDUSTRIAL WASTEWATER MONITORING SYSTEM

Abstract. In many parts of Vietnam, industrial wastewater remains the main cause of water pollution in rivers. Compliance with regulations in enterprises is still formal, the operation of centralized wastewater treatment systems is still coping and not really effective. This paper presents a detailed approach of recent works carried out in the field of smart wastewater quality monitoring. A simpler solution for wastewater quality monitoring based on Internet of Things technology combining with cloud services in order to develop a real time wastewater quality monitoring system. The system is used for transferring wastewater indicators to the cloud server over the Internet. The collected data are then analyzed. The system also provides a warning to a remote user about the quality of the wastewater.

Keywords. Internet of things, cloud platform, industrial wastewater, smart water monitoring.

1. GIỚI THIỆU

Ô nhiễm nước là sự thay đổi theo chiều tiêu cực của các tính chất vật lý – hóa học – sinh học của nước, với sự xuất hiện các chất lạ ở thể lỏng, rắn làm cho nguồn nước trở nên độc hại với con người và sinh vật. Nước bị ô nhiễm ở các khu vực nước ngọt (sông, suối, ao, hồ...) và các vùng ven biển, vùng biển khép kín do lượng muối khoáng và hàm lượng các chất hữu cơ quá dư thừa làm cho các quần thể sinh vật trong nước không thể đồng hóa được. Kết quả làm cho hàm lượng oxy trong nước giảm đột ngột, các khí độc tăng lên, tăng độ đục của nước dẫn đến việc gây ra suy thoái thủy vực. Ô nhiễm nước có nguyên nhân từ các loại chất thải công nghiệp được thải ra lưu vực các con sông mà chưa qua xử lý đúng mức, các loại phân bón hóa học và thuốc trừ sâu ngấm vào mạch nước ngầm, ao hồ, nước thải sinh hoạt từ các khu dân cư ven sông gây ô nhiễm trầm trọng, ảnh hưởng đến sức khỏe của người dân, sinh vật và vi sinh vật sống trong khu vực.

Các thông số của nguồn nước bao gồm: Độ dẫn điện EC (Electric Conduction), nồng độ chất rắn hòa tan TDS (Total Dissolved Solids), nhiệt độ, độ đục, độ PH, nồng độ oxy hòa tan trong nước DO (Dissolved Oxy), tổng lượng chất rắn lơ lửng TSS (Total Suspended Solids) trong nước... là các thông số quan trọng để chúng ta xác định được mức độ ô nhiễm nguồn nước. Vì vậy, một hệ thống cảnh báo mức độ ô nhiễm nguồn nước có hiệu quả nên dựa trên việc đo đạc, thu thập và phân tích thường xuyên các thông số này tại các vị trí thiết yếu của nguồn nước như: sông, suối, ao hồ, nơi xả thải của khu công nghiệp. Xây dựng hệ thống giám sát tự động chất lượng nguồn nước đã nhận được sự quan tâm của các nhà nghiên cứu cũng như những ai quan tâm đến môi trường.

Trên thế giới cũng như tại Việt Nam cũng có khá nhiều các dự án xây dựng các hệ thống liên quan đến giám sát chất lượng nguồn nước kết hợp với công nghệ thông tin. Trong bài báo này, nhóm tìm hiểu, thiết kế và xây dựng mô hình hệ thống giám sát chất lượng nguồn nước xả thải công nghiệp theo thời gian thực, sử dụng các công nghệ mới như vạn vật kết nối và điện toán đám mây. Chức năng của hệ thống bao gồm theo dõi, thông báo và dự đoán cũng như cảnh báo về chất lượng nguồn nước xả thải để người dùng có những tác động phù hợp.

2. CÁC NGHIÊN CỨU LIÊN QUAN

Trong nghiên cứu [1], tác giả ứng dụng IoT trong việc xây dựng một hệ thống quan trắc và cảnh báo mức độ ô nhiễm nguồn nước trong ao hồ, bao gồm phao nổi có cảm biến để đo đạc và mạch để thu thập truyền các giá trị nồng độ pH, nhiệt độ, oxy hòa tan và độ dẫn điện. Các giá trị này được gửi lên Internet và được hiển thị trong ứng dụng máy tính dùng Labview. Tuy nhiên, bài báo không đề cập đến việc sử dụng cơ sở dữ liệu để lưu trữ thông tin hoặc dùng các ứng dụng di động để truy cập. So sánh với các nghiên cứu đã được đề cập, bài báo hướng đến việc nghiên cứu và chế tạo các node cảm biến tích hợp và cơ động cho phép hệ thống dễ dàng thay đổi vị trí các khu vực khác nhau sao cho khoảng cách của các node cảm biến với node trung tâm tầm 7-8 Km. Giống các nghiên cứu trên hệ thống dùng các ứng dụng dành cho thiết bị di động và máy tính để hiển thị thông tin, lưu trữ kết quả đo trên Server. Điểm khác biệt là cấu trúc các node cảm biến không dây truyền thông qua chuẩn LoRa và việc lấy dữ liệu để thực hiện phân tích cụ thể hơn cho từng trường hợp xảy ra trong môi trường, đồng thời tối ưu hóa tài nguyên và trao đổi thông tin với các tổ chức, cá nhân khác nhau, hệ thống kết nối với nhiều loại cảm biến xác định được nhiều tham số để tăng cường độ chính xác trong việc phân tích dữ liệu đưa ra mức cảnh báo ô nhiễm môi trường nước. Tất cả điều này là rất quan trọng vì nó dễ dàng tiếp cận với tất cả các đối tượng khác nhau, giảm thiểu rủi ro nếu có các bất thường về nguồn nước, đặc biệt là nước sinh hoạt.

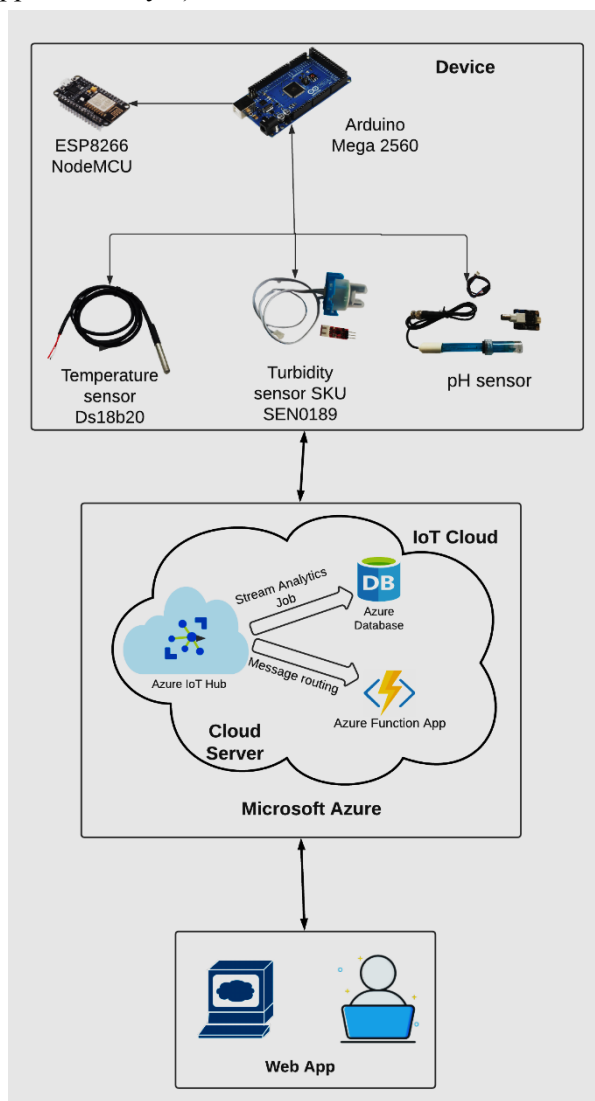
Nghiên cứu của nhóm tác giả [2] đề xuất giải pháp xây dựng hệ thống giám sát chất lượng nguồn nước thông minh dựa trên kỹ thuật vạn vật kết nối. Hệ thống được sử dụng trong việc thu thập dữ liệu nguồn nước thông qua hệ thống con thu thập dữ liệu từ các cảm biến, quá trình truyền các dữ liệu từ thiết bị đến cơ sở dữ liệu thông qua mạng Internet, dữ liệu được lưu trữ để kiểm tra, phân tích và gửi cảnh báo đến người dùng về sự thay đổi các thông số về chất lượng nguồn nước so với thông số chuẩn. Hệ thống cũng cung cấp trang web để người dùng có thể theo dõi về nguồn nước.

Trong thực tế đã có các hệ thống giám sát tự động chất lượng nguồn nước qua GPRS sử dụng để giám sát và xác định loại ô nhiễm nước, nồng độ các chất gây ô nhiễm nước khác nhau và xu hướng thay đổi của chúng. Đây là hệ thống quan trọng trong quản lý tài nguyên nước và bảo vệ môi trường nước. Phát hiện kịp thời với việc giám sát hiệu quả ô nhiễm nguồn nước để lên kế hoạch dự phòng thích hợp và có ảnh hưởng tốt đến sức khỏe và an toàn cho con người. Hệ thống giám sát tự động chất lượng nước dựa trên nền tảng truyền nhận dữ liệu không dây GPRS (General Packet Radio Service) và hệ thống được kết hợp với màn hình hiển thị các thông số chất lượng nước. Modem truyền thông GPRS kết nối với màn hình chất lượng nước thông qua Serial Port (RS232/RS485/RS442). Nó có thể truyền tín hiệu điều khiển từ hệ thống và nhận tín hiệu thu thập được từ màn hình hiển thị các thông số gửi về cho trung tâm giám sát [3][4].

3. ỨNG DỤNG IOT-CLOUD TRONG HỆ THỐNG GIÁM SÁT NGUỒN NƯỚC XẢ THẢI CÔNG NGHIỆP THEO THỜI GIAN THỰC

Trong phần này, nhóm đề xuất kiến trúc xây dựng hệ thống giám sát nguồn nước xả thải công nghiệp tại Việt Nam ứng dụng kỹ thuật vạn vật kết nối (IoT) và các dịch vụ điện toán đám mây (cloud services) thu thập các thông tin về chất lượng nước thải công nghiệp bao gồm đo các thông tin chỉ số về độ đục, độ pH,

nhiệt độ, lưu lượng nước. Các chỉ số này sẽ được gửi về máy chủ thông qua mạng internet và lưu trong cơ sở dữ liệu đám mây để phục vụ việc phân tích, dự đoán chất lượng nguồn nước. Nếu có các bất thường xảy ra đối với nguồn nước hoặc các chỉ số vượt quá mức quy định thì hệ thống sẽ gửi cảnh báo qua tin nhắn điện thoại đến người dùng và người quản trị. Bên cạnh đó, người quản trị cũng có thể theo dõi tình trạng nguồn nước xả thải thông qua trang web hiển thị các thông tin của hệ thống theo thời gian thực. Hình 1 mô tả kiến trúc hệ thống, gồm ba tầng nghiệp vụ: tầng thiết bị (Device Layer), tầng đám mây (Cloud layer) và tầng ứng dụng (Application layer).



Hình 1: Hệ thống IoT-Cloud giám sát chất lượng nguồn nước xả thải

- **Tầng thiết bị:** gồm các cảm biến thu thập dữ liệu như pH, nhiệt độ, ... gửi về tầng đám mây thông qua giao thức truyền thông MQTT (Message Queue Telemetry Transport)
- **Tầng đám mây:** nhận dữ liệu từ tầng thiết bị thông qua công nghệ IoT Hub (một phần của Azure IoT) để thiết lập giao tiếp giữa thiết bị – đám mây – thiết bị.
- **Tầng ứng dụng:** sử dụng framework.net Core và EF Core (Entity Framework Core) để xử lý dữ liệu Azure Database hiện thực ra Web hoặc App. Các thiết bị cảm biến sẽ được lắp đặt ở cửa xả thải để đo, thu thập dữ liệu và gửi về database thông qua Service của Azure là IoT Hub Sau khi đã thu thập dữ liệu thì Server sẽ xử lý các thông số từ database và gửi đến client. Client sẽ hiển thị các thông số thu thập và xử lý được kèm theo thời gian thực để người dùng theo dõi.

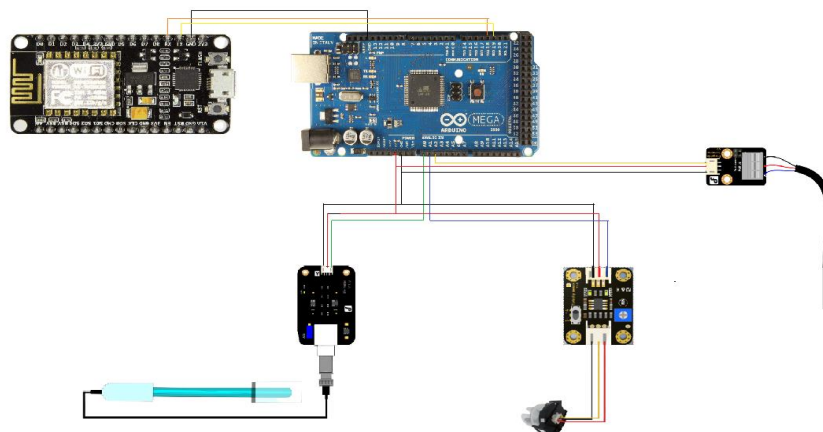
3.1. Thiết kế phần cứng

Bảng 1 trình bày chi tiết các thiết bị và thông số kỹ thuật của các thiết bị cần thiết để xây dựng hệ thống cảm biến và giám sát nguồn nước.

Thiết bị	Thông số kỹ thuật
ESP NodeMCU	Chip: ESP8266EX WiFi: 2.4 GHz hỗ trợ chuẩn 802.11 b/g/n Điện áp hoạt động: 3.3V Điện áp vào: 5V thông qua cổng USB Số chân I/O: 11 (tất cả các chân I/O đều có Interrupt/PWM/I2C/One-wire, trừ chân D0) Số chân Analog Input: 1 (điện áp vào tối đa 3.3V) Bộ nhớ Flash: 4MB Giao tiếp: Cable Micro USB Hỗ trợ bảo mật: WPA/WPA2 Tích hợp giao thức TCP/IP Lập trình trên các ngôn ngữ: C/C++, Micropython, NodeMCU - Lua
Arduino	Vi điều khiển ATmega328P Điện áp hoạt động 5V Điện áp vào khuyến dùng 7-12V Điện áp vào giới hạn 6-20V Digital I/O pin 14 (trong đó 6 pin có khả năng băm xung) PWM Digital I/O Pins 6 Analog Input Pins 6 Cường độ dòng điện trên mỗi I/O pin 20 mA Cường độ dòng điện trên mỗi 3.3V pin 50 mA Flash Memory 32 KB (ATmega328P) 0.5 KB được sử dụng bởi bootloader SRAM 2 KB (ATmega328P) EEPROM 1 KB (ATmega328P) Tốc độ 16 MHz Chiều dài 68.6 mm Chiều rộng 53.4 mm Trọng lượng 25 g
Thiết bị đo nồng độ pH	Điện áp hoạt động: 5V Dải đo chỉ số pH: 0 - 14 pH Dữ liệu đầu ra: Analog Nhiệt độ làm việc: 0 - 60 °C Độ chính xác: ± 0.1pH (25 °C) Thời gian đáp ứng: ≤ 1 phút Sử dụng điện cực công nghiệp với kết nối chuẩn BNC Tích hợp Led chỉ thị nguồn Kết nối đầu ra PH2.0 3 chân Kích thước: 43mm * 32mm
Thiết bị đo độ đục	Đầu vào: 3.3-5V (với mạch chống ngược) Đầu ra: tương tự mặc định (có thể thiết lập đầu ra cao và thấp) Dải đo độ đục: 0-1000 NTU Analog Out : 0-4.5V
Thiết bị đo nhiệt độ	Điện áp sử dụng: 3~5.5VDC Dòng tiêu thụ: 1~1.5mA

	<p>Chuẩn giao tiếp: Digital TTL 1-Wire (chỉ 1 chân Data duy nhất, lưu ý cần nối chân Data của cảm biến lên mức cao VCC qua điện trở kéo 4k7 Ohm hoặc 10k Ohm trước khi kết nối với Vi điều khiển). Khoảng nhiệt độ đo được: -55~125°C Độ chính xác (sai số): ±0.5°C Độ phân giải: 9~12 bit (có thể thiết đặt) Thời gian phản hồi < 750ms Thiết kế dạng dây chống nước dài 1m với vỏ thép bảo vệ chắc chắn.</p>
--	---

Sơ đồ kết nối thiết bị được trình bày trong Hình 2.

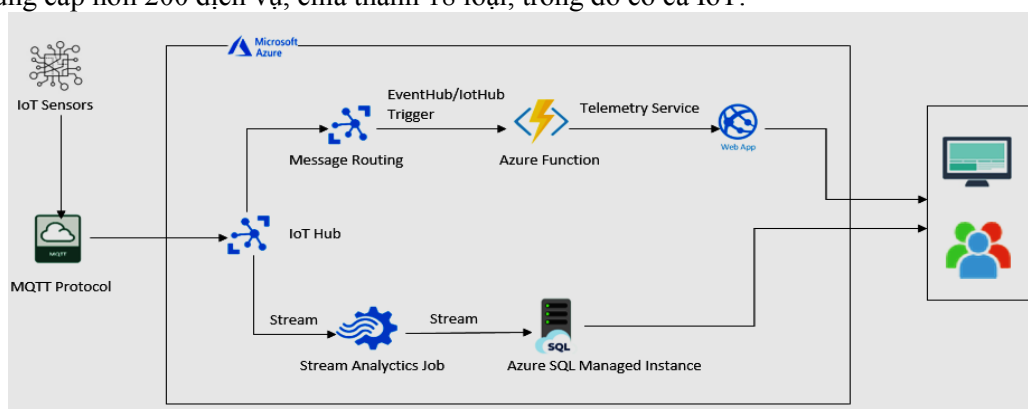


Hình 2. Sơ đồ kết nối thiết bị

Các thiết bị cảm biến sẽ kết nối và gửi dữ liệu về Arduino thông qua các chân Digital, nhiệm vụ của Arduino lúc này là gom lại tất cả dữ liệu mà cảm biến đo được và chuyển qua ESP để có thể gửi về Cloud. Việc giao tiếp giữa Arduino và ESP lúc này là thông qua giao thức UART [12] (chuẩn giao tiếp nối tiếp với sự hỗ trợ của phần cứng – hardware)

3.2. Thiết kế phần mềm

Tầng đám mây trong kiến trúc IoT áp dụng hệ thống này sử dụng Microsoft Azure. Azure [9] là nền tảng điện toán đám mây (cloud computing platform) và một cổng trực tuyến (online portal) cho phép người dùng truy cập và quản lý các dịch vụ (service) và tài nguyên (resource) đám mây do Microsoft cung cấp, Azure cung cấp hơn 200 dịch vụ, chia thành 18 loại, trong đó có cả IoT.



Hình 3. Sơ đồ mô tả quy trình hoạt động của hệ thống

Dữ liệu từ các cảm biến được truyền về máy chủ đám mây (Cloud Server) và sử dụng các dịch vụ của Microsoft Azure bao gồm: Azure IoT Hub, Stream Analytics Job, Azure Function App, Azure Database (Azure SQL Managed Instance), Azure Web App.

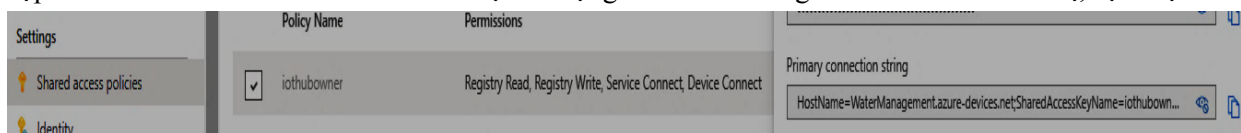
Azure IoT Hub [6] là một dịch vụ được quản lý, được lưu trữ trên đám mây, hoạt động như một trung tâm

thông báo trung tâm để giao tiếp hai chiều giữa ứng dụng IoT của bạn và các thiết bị mà nó quản lý. IoT Hub có chức năng định tuyến bằng message giúp thiết lập sử dụng gửi message để điều khiển thiết bị kết nối. Ngoài ra, IoT Hub còn có thể tích hợp với các dịch vụ khác của Azure để xây dựng các giải pháp end-to-end hoàn chỉnh, IoT Hub Của Azure có thể xây dựng trên nhiều nền tảng được hỗ trợ bao gồm nhiều bản phân phối Linux, Windows và hệ điều hành thời gian thực. Khi các dữ liệu được thu thập sẽ gửi về IoT Cloud bằng những message (trong đó mỗi message bao gồm một chuỗi JSON chứa các thông số mà thiết bị đo được) thông qua cổng giao thức MQTT. Để kết nối thiết bị với IoT Hub cần phải có một Connection Strings (được IoT Hub tự sinh ra khi đăng ký thiết bị trên IoT Hub).

```
HostName=WaterManagement.azure-devices.net;DeviceId=MyIoTDevice;SharedAccessKey=18N0BWVOQ6NzEOhsIIY96TkQE8J9QDgVBy38XBJAZO4=
```

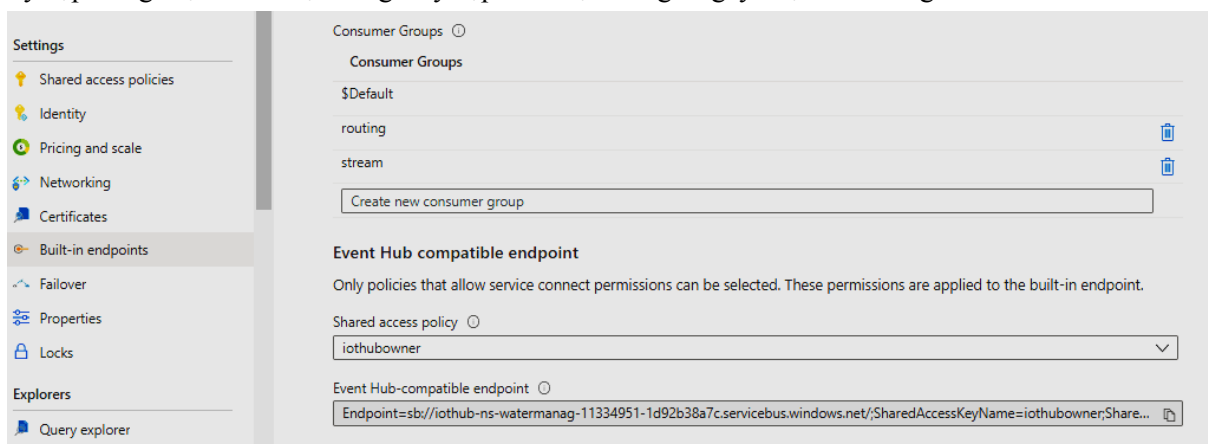
Ngoài Connection String của IoT Device thì còn thứ quan trọng nữa trong IoT Hub giúp các dịch vụ khác trên Azure có thể truy cập và truy xuất dữ liệu từ IoT Hub đó là Shared access policies key và Build-in-endpoints:

+ Shared access policies key: về cơ bản, nó chứa quyền truy cập trung tâm IoT. Các chính sách truy cập chứa khóa và chuỗi kết nối có thể được sử dụng để kết nối trung tâm IoT từ các thiết bị, dịch vụ.



Hình 4. Mục Shared access policies của IoT Hub

+ Build-in-endpoints: giống như một định tuyến cho phép các dịch vụ khác đọc các Device – to – cloud messages thông qua Event Hub-compatible endpoint. Ngoài Event Hub-compatible endpoint thì cái quan trọng thứ hai trong Build-in-endpoints là Consumer group [11] – Một nhóm người dùng mà người dùng có thể thêm hoặc xóa đi, một cài đặt tựa như các hệ thống nhắn tin khác để các dịch vụ khác có thể truy cập lấy dữ liệu dựa trên các Consumer cụ thể. Ví dụ như có hai dịch vụ cùng truy cập IoT Hub là Stream Analytics và Azure Function thì ta sẽ thêm 2 consumer vào Consumer group, Hình 5, để mỗi dịch vụ có thể truy cập riêng biệt tránh việc cùng truy cập vào một luồng sẽ gây chậm trễ và nghẽn.

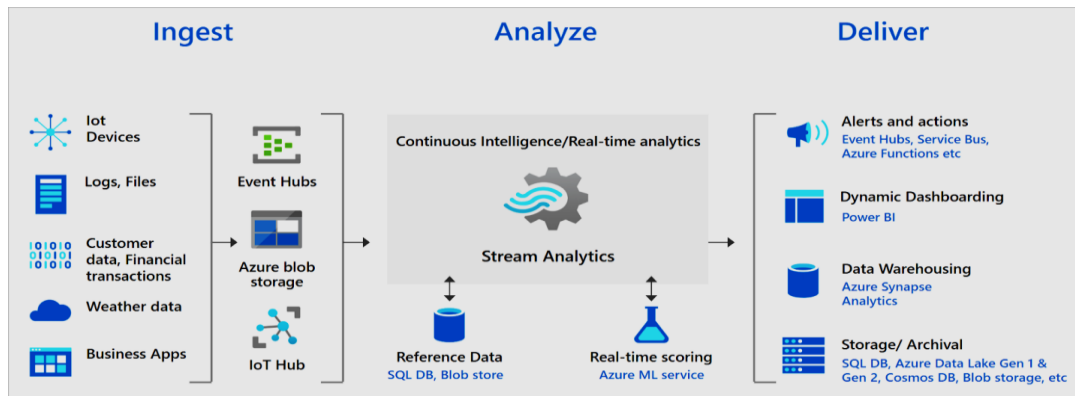


Hình 5. Mục Build-in-endpoints của IoT Hub

Các message sau khi được gửi lên cloud thì IoT Hub (được coi là cổng giúp thiết bị là Azure Cloud liên kết với nhau) sẽ lưu trữ tạm thời lại toàn bộ message vì sau một ngày thì IoT Hub sẽ reset lại bộ nhớ - điểm này khiến cho bộ nhớ lẫn tốc độ xử lý của Cloud không bị quá tải. Chính vì việc IoT Hub sẽ reset lại bộ nhớ nên ta sẽ dùng đến dịch vụ Stream Analytics để giúp lưu trữ dữ liệu vào Database.

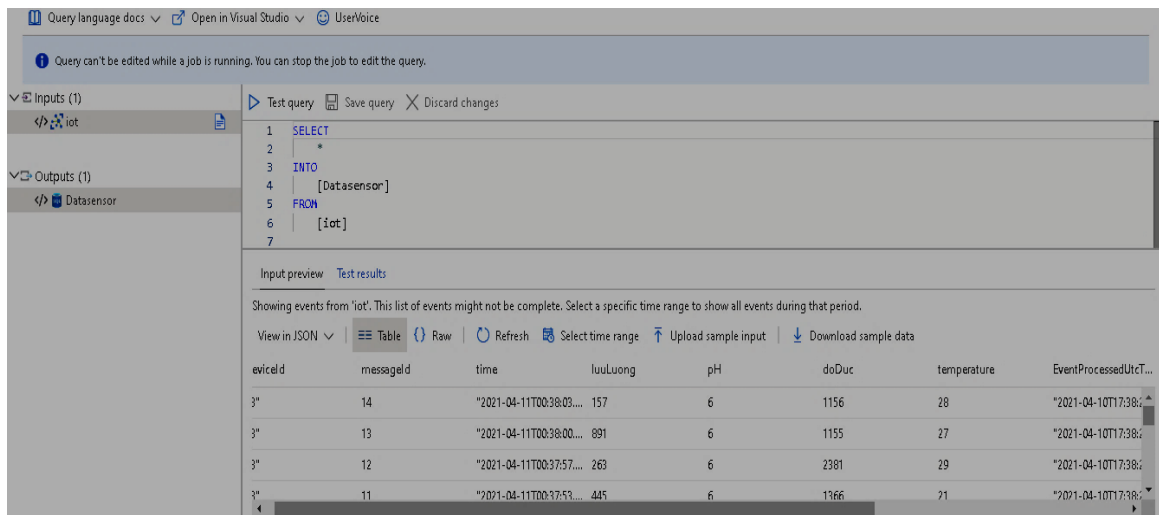
Stream Analytics [7], Hình 6, là một công cụ phân tích thời gian thực và xử lý sự kiện phức tạp, công việc Azure Stream Analytics bao gồm đầu vào, truy vấn và đầu ra (cụ thể trong hệ thống ở đây đầu vào là IoT Hub và đầu ra là Azure Database). Dựa trên ngôn ngữ truy vấn SQL, có thể được sử dụng để dễ dàng lọc, sắp xếp, tổng hợp và kết hợp dữ liệu phát trực tuyến trong một khoảng thời gian (có thể dùng dịch vụ này

để xử lý dữ liệu trên Cloud trước khi lưu trữ).



Hình 6. Cơ chế hoạt động của Stream Analytics

Stream Analytics sẽ phân tích chuỗi JSON nằm trong từng message trong IoT Hub và thêm vào Database, mỗi chuỗi JSON trong một message sẽ tương ứng với một record trong một bảng của Database. Tất cả những điều đó thông qua câu lệnh truy vấn SQL với input là từ IoT Hub và output là Database, Hình 7.

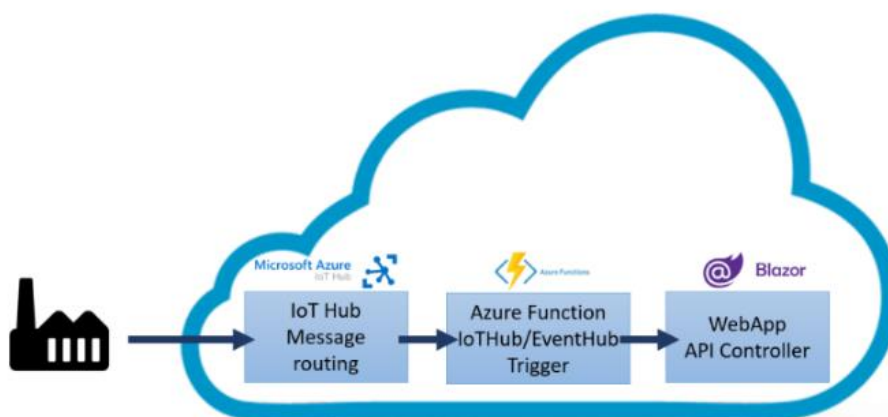


Hình 7. Giao diện xử lý của Stream Analytics

Để Stream Analytics có thể truy cập được hai đầu vào và ra (cụ thể ở đây là IoT Hub và Azure Database) thì cần quan tâm hai vấn đề khi tạo input và output:

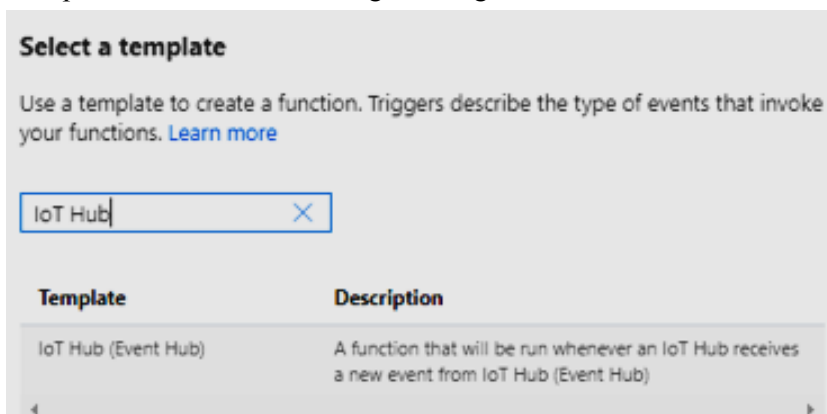
- + Inputs (IoT Hub): cần quan tâm Shared access policies key
- + Outputs (Azure Database): cần quan tâm Connection String

Cơ sở dữ liệu Azure SQL [10][11] là một nền tảng được quản lý hoàn toàn dưới dạng công cụ cơ sở dữ liệu dịch vụ (PaaS) xử lý hầu hết các chức năng quản lý cơ sở dữ liệu như nâng cấp, và lỗi, sao lưu và giám sát mà không cần sự tham gia của người dùng. Cơ sở dữ liệu Azure SQL luôn chạy trên phiên bản ổn định mới nhất của công cụ cơ sở dữ liệu SQL Server và hệ điều hành được và lỗi với tính khả dụng 99,99%. Các khả năng PaaS được tích hợp trong Cơ sở dữ liệu Azure SQL cho phép bạn tập trung vào các hoạt động quản trị và tối ưu hóa cơ sở dữ liệu theo miền cụ thể, rất quan trọng cho doanh nghiệp của bạn. Với Cơ sở dữ liệu Azure SQL, có thể tạo lớp lưu trữ dữ liệu hiệu suất cao và sẵn có cho các ứng dụng và giải pháp trong Azure. Cơ sở dữ liệu SQL có thể là lựa chọn phù hợp cho nhiều ứng dụng đám mây hiện đại vì nó cho phép xử lý cả dữ liệu quan hệ và cấu trúc phi quan hệ, chẳng hạn như đồ thị, JSON, không gian và XML. Tất cả dữ liệu được visualize real time về web, ứng dụng nhờ vào kỹ thuật Telemetry (một kỹ thuật điều khiển, truyền nhận dữ liệu từ xa để giám sát).

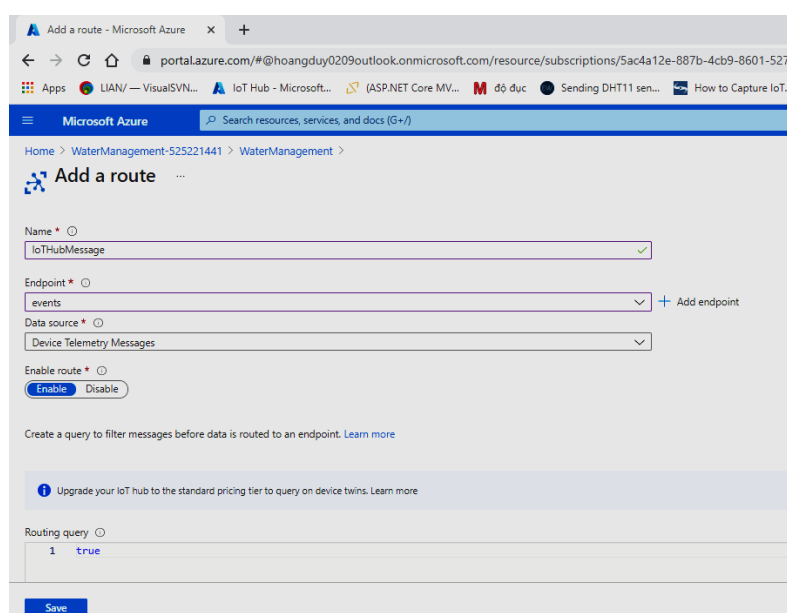


Hình 8. Kiến trúc áp dụng Telemetry để truyền dữ liệu IoT Hub ra web [8]

Sử dụng Azure Function App làm cầu nối để có thể truyền nhận từ IoT Hub ra Azure Web App, khi tạo một Function thì sẽ cho người dùng lựa chọn template phù hợp với nhu cầu sử dụng (cụ thể áp dụng cho hệ thống này thì là template IoT Hub). Tuy nhiên để Function App có thể truy cập lấy dữ liệu từ IoT Hub thì cần phải cấu hình bên phía IoT Hub đó là Message routing.



Hình 9a. Chọn template phù hợp



Hình 9b. Tạo một message routing

Sau khi Azure Function đã có thể truy cập và truy vấn IoT Hub thì viết code cho function để function truy vấn dữ liệu từ IoT Hub và chuyển về web:

```
using System;
using System.Net.Http;
using System.Text;

public static void Run(string myIoTHubMessage, ILogger log)
{
    log.LogInformation($"IoT Hub trigger function processed a message: {myIoTHubMessage}");
    var url = "https://khoaluantotnghiep.azurewebsites.net/api/values";
    var stringContent = new StringContent(myIoTHubMessage, Encoding.UTF8, "application/json");

    using var httpClient = new HttpClient();

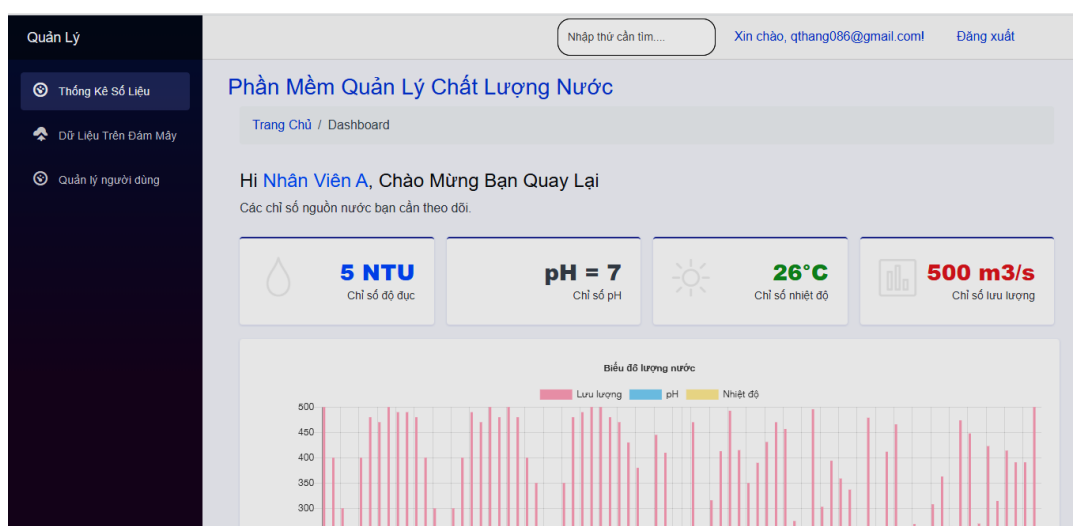
    var response = httpClient.PostAsync(url, stringContent).Result;
    Console.WriteLine($"Response {response.StatusCode}");
}
```

Biến *myIoTHubMessage* là biến mặc định của IoT Hub dùng để lưu trữ các message được gửi từ thiết bị về IoT Hub để cho function truy cập.

4. KẾT QUẢ VÀ HƯỚNG PHÁT TRIỂN

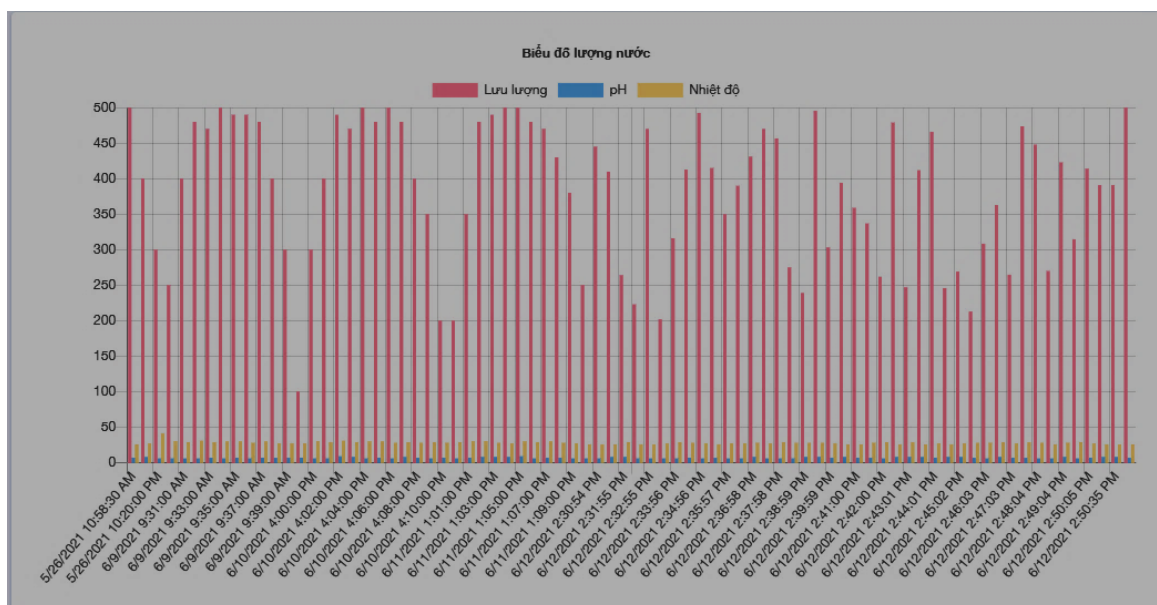
4.1. Kết quả

Tất cả các dữ liệu thu thập được hiển thị trên giao diện chính của trang web như Hình 10. Trong đó, hàng đầu tiên sẽ cập nhật các chỉ số của các thiết bị đo được theo thời gian thực mỗi khi dữ liệu được gửi về máy chủ.



Hình 10. Giao diện chính của trang web

Ngoài việc cập nhật các chỉ số theo thời gian thực thì, các thông số sẽ còn được biểu diễn thông qua biểu đồ ở Hình 11 để người quản lý có đánh giá một cách tổng quát. Khi có chỉ số nào vượt quá mức cho phép thì trang web sẽ thông báo cho người dùng hoặc người quản trị.



Hình 11. Biểu đồ chỉ số đo nguồn nước theo thời gian thực

4.2. Kết luận và hướng phát triển

Hệ thống hoàn thành các chức năng chính trong giám sát và phân tích chất lượng nguồn nước xả thải, đáp ứng các yêu cầu của người dùng. Hệ thống vẫn đang còn phát triển ứng dụng để có thể hoàn thiện sản phẩm một cách tốt nhất, thời gian xử lý và trả kết quả nhanh nhất và nhược điểm của hệ thống chưa đo được hết tất cả chỉ số chất lượng nước công nghiệp.

Hướng phát triển: Phát triển hệ thống có thể đo tất cả chỉ số; Thời gian khi đo và xuất ra được chất lượng nguồn nước nhanh; Phát triển thêm một vài chức năng như dự đoán được chất lượng nước; Phát triển ứng dụng quản lý chất lượng trên điện thoại; Phát triển thêm từ hệ thống giám sát nước thải công nghiệp có thể sang nước thải sinh hoạt hoặc một hệ thống không khí, ...; Dùng các thiết bị có độ bền và chất lượng cao hơn để sản phẩm có thể có giá trị cao.

PHỤ LỤC

Giá trị C của các thông số làm cơ sở để tính nồng độ tối đa cho phép trong khí thải

TT	Thông số	Đơn vị	Giá trị C		
			A	B ₁	B ₂
1	Bụi tổng	mg/Nm ³	400	200	100
2	Lưu huỳnh đioxit, SO ₂	mg/Nm ³	1.500	500	500
3	Nitơ oxit NO _x (tính theo NO ₂)	mg/Nm ³	1.000	850	750
4	Cadimi và hợp chất (tính theo Cd)	mg/Nm ³	20	5	1
5	Chì và hợp chất (tính theo Pb)	mg/Nm ³	10	5	2
6	Tổng chất hữu cơ dễ bay hơi, VOC	mg/Nm ³		20	20
7	Benzo(a)pyren	mg/Nm ³	-	0,1	0,1
8	Amoniac và các hợp chất amoni (tính theo NH ₃)	mg/Nm ³	76	50	30
9	Axit clohydric, HCl	mg/Nm ³	200	50	20

10	Flo, HF, hoặc các hợp chất vô cơ của Flo (tính theo HF)	mg/Nm ³	50	20	10
11	Hydro sunfua, H ₂ S	mg/Nm ³	7,5	7,5	5
<i>Hàm lượng oxy tham chiếu trong khí thải công nghiệp sản xuất thép là 7%</i>					

Cột A Bảng 1 quy định giá trị C của các thông số ô nhiễm trong nước thải công nghiệp sản xuất thép khi xả vào nguồn nước được dùng cho mục đích cấp nước sinh hoạt.

TÀI LIỆU THAM KHẢO

- [1] Vu Van Thanh, “Nghiên cứu ứng dụng IoT trong quan trắc và cảnh báo mức độ ô nhiễm nguồn nước”, *Tạp chí điện tử khoa học và công nghệ*, Vol. 17, No. 1.1, Trang 74-79, 2019.
- [2] Geetha, S., Gouthami, S. Internet of things enabled real time water quality monitoring system. *Smart Water* 2, 1 (2016).
- [3] Giới thiệu hệ thống giám sát tự động chất lượng nước qua GPRS (bkaii.com.vn)
- [4] <https://asin.com.vn/he-thong-giam-sat-chat-luong-nuoc-tu-dong-online-asin-viet-nam.html>
- [5] <https://www.env.go.jp/air/tech/ine/asia/vietnam/files/law/QCVN%2005-2013.pdf>
- [6] <https://docs.microsoft.com/vi-vn/azure/iot-hub/>
- [7] <https://docs.microsoft.com/en-us/azure/stream-analytics/stream-analytics-introduction>
- [8] <https://sandervandeveldde.wordpress.com/2020/11/04/sending-iot-hub-telemetry-to-a-blazor-web-app/>
- [9] https://vinsep.com/kien-thuc/microsoft-kien-thuc/azure/tim-hieu-ve-microsoft-azure-cloud-computing/#Microsoft_Azure_la_gi
- [10] <https://docs.microsoft.com/en-us/azure/azure-sql/database/sql-database-paas-overview>
- [11] <https://sandervandeveldde.wordpress.com/2017/06/21/the-only-thing-standing-between-iot-hub-and-functions-are-consumers-groups/>
- [12] <https://dientutuonglai.com/giao-tiep-uart-la-gi.html>

YSC4F.216

TỐI ƯU CHIỀU DÀI CÁP ĐIỆN NHÀ MÁY ĐIỆN GIÓ XA BỜ DỰA TRÊN THUẬT TOÁN LAI SPDPSO VÀ SWEEP HEURISTIC

PHAM QUOC KHANH¹, NGUYEN HUY KHIEM²

Faculty of Electrical Engineering Technology, Industrial University of Ho Chi Minh City

¹ phamquockhanh@iuh.edu.vn, ² nguyenhuykhiem@iuh.edu.vn

Tóm tắt. Trong các nhà máy điện gió xa bờ (OWF), các tuabin máy phát điện gió (WT) được thiết kế cách xa nhau để giảm tác động của “wake effect”. Điều này làm gia tăng chiều dài dây cáp truyền tải điện dẫn đến tăng chi phí lắp đặt và mua sắm ban đầu của dự án điện gió vì các đường dây điện này rất đắt đỏ. Bài toán tối ưu đường dây điện trong OWF (OWFCRP) có vai trò quan trọng trong nâng cao tính cạnh tranh của năng lượng gió trong thị trường điện. Có nhiều nghiên cứu được đề xuất trong những năm gần đây để giải quyết vấn đề này và đã đem lại kết quả. Với sự phát triển của các OWF thì cần thêm nhiều hơn các nghiên cứu về OWFCRP để áp dụng trong tình hình phát triển năng lượng tái tạo của Việt Nam hiện nay. Bài báo đề xuất phương pháp lai giữa thuật toán PSO rời rạc kiểu xác suất và thuật toán Sweep để giải quyết bài toán OWFCRP. Sử dụng cùng một OWF đã vận hành để đánh giá hiệu quả của thuật toán tối ưu, kết quả tối ưu được bằng thuật toán lai được đề xuất cho thấy chi phí của đường dây gom điện cho hệ thống được giảm đi so với chi phí của phương pháp đã thi công.

Từ khóa. Nhà máy điện gió xa bờ, bài toán tối ưu đường dây điện trong nhà máy điện gió xa bờ, thuật toán PSO rời rạc kiểu xác suất, thuật toán Sweep.

OFFSHORE WIND FARM CABLE ROUTING PROBLEM BASED ON HYBRID ALGORITHM SPDPSO AND SWEEP HEURISTIC

Abstract. In offshore wind power plants (OWF), the wind turbines (WT) are spaced apart to reduce the “wake effect”. This increases the length of power transmission cables leading to increased installation and initial procurement costs of wind power projects because these power lines are costly. Offshore Wind Farm Cable Routing Problem of OWF (OWFCRP) plays an essential role in improving the competitiveness of wind energy in the electricity market. Several studies have been proposed in recent years to address this problem and have yielded results. However, with the development of OWFs, more studies on OWFCRP are needed to apply in the current situation of renewable energy development in Vietnam. The article proposes a hybrid method between the Selective Probabilistic Discrete Particle Swarm Optimization and the Sweep algorithm to solve the OWFCRP problem. Using the same actual OWF to evaluate the efficiency of the optimization algorithm, the optimization results obtained by the proposed hybrid algorithm show that the cost of the power collection line for the system is reduced compared to the cost of the system. of the applied method.

Keywords. Offshore Wind Power Plant, Offshore Wind Farm Cable Routing Problem, Selective Probabilistic Discrete Particle Swarm Optimization, Sweep algorithm.

1. GIỚI THIỆU

Nhu cầu sử dụng điện ngày càng tăng cao bởi quá trình công nghiệp hóa - điện khí hóa đất nước. Mỗi năm hệ thống điện cả nước tăng thêm hơn 7500 MW tham gia hòa lưới điện [1]. Việc tăng công suất phát điện bằng các nguồn năng lượng truyền thống cũng gặp nhiều khó khăn. Các nguồn thủy điện lớn có thể phát

điện đã tận dụng hết trong khi các nguồn thủy điện nhỏ vấp phải sự phản đối ngày càng cao bởi phá rừng, hủy hoại môi trường sinh thái và tăng sức tàn phá của lũ lụt mỗi mùa mưa đến. Các nhà máy nhiệt điện than gây ô nhiễm môi trường không khí và giá than tăng cao do khó khăn trong khai thác và bất ổn chính trị [2]. Các nguồn năng lượng hạt nhân lại chưa được cho phép phát triển bởi nhiều yếu tố kinh tế chính trị hiện nay [3].

Các nhà máy phát điện sử dụng năng lượng tái tạo như điện gió và điện mặt trời là một trong các giải pháp tiềm năng để góp phần đảm bảo nhu cầu điện năng cho đất nước và đảm bảo thực hiện các cam kết bảo vệ môi trường của chính phủ về phát thải khí nhà kính. Hiện nay nhiều nhà máy điện gió đã được xây dựng và vận hành trên cả nước và một số dự án đã được đưa vào kế hoạch xây dựng trong các năm tiếp theo. Một trong những trở ngại lớn trong việc phát triển năng lượng gió là chi phí xây dựng ban đầu lớn so với các nguồn năng lượng khác. Do đó vấn đề tối ưu chi phí xây dựng các trang trại gió là một vấn đề cần phải nghiên cứu để giảm chi phí xây dựng, từ đó nâng cao tính cạnh tranh của điện gió với các nhà máy điện truyền thống khác.

Trong các chi phí đầu tư ban đầu thì chi phí cho hệ thống dây thu gom điện từ các trụ gió đến trạm biến áp chiếm từ 4 đến 5 % chi phí đầu tư ban đầu của cả hệ thống điện gió [4]. Nguyên nhân của việc này là các trụ gió phải được xây dựng cách xa nhau từ 5 đến 7 lần đường kính quét của cánh tubin gió nhằm loại bỏ tác động của hiện tượng “Wake effect” [5]. Hiện tượng này có nguyên nhân là sự thay đổi tốc độ gió khi xuyên qua cánh quạt của tubin gió đứng trước tác động vào tubin gió đứng phía sau hướng gió. Khoảng cách giữa các trụ gió tăng lên sẽ làm chi phí cho hệ thống dây dẫn điện tăng theo. Nên vấn đề này rất được quan tâm nghiên cứu trong những năm gần đây và đặc biệt là trong các hệ thống điện gió xa bờ [5].



Hình 34 Hình ảnh minh họa cho hiện tượng “Wake effect” trong [5]

Bài toán định tuyến cáp điện trang trại gió (Wind Farm Cable Routing Problem – WFCRP) là bài toán kết nối tất cả các tubin trong một trang trại với một điểm thu gom, được gọi là trạm biến áp ngoài khơi. Cụ thể hơn, đối với các trang trại điện gió ngoài khơi, chi phí cơ sở hạ tầng điện lên tới 4% –5% tổng chi phí vốn đầu tư, vì vậy điều quan trọng là phải thiết kế các giải pháp càng gần càng tốt với mức tối ưu. Nghiên cứu trong [5] là một ví dụ cụ thể trong định tuyến cáp trang trại gió.

Với mục tiêu cực tiểu chi phí đầu tư trang trại gió, nhiều nghiên cứu được đề xuất trong thời gian qua để giải các bài toán này [6][7][8]. WFCRP đã được nhiều tác giả nghiên cứu, sử dụng cả kỹ thuật chính xác và kỹ thuật heuristic.

Các kỹ thuật heuristic [6][7][8] lựa chọn các cấu trúc liên kết các trang trại gió dựa trên thuật toán thử sai có lựa chọn. Việc lựa chọn các ứng viên tiềm năng ban đầu thực hiện bằng cách lựa chọn ngẫu nhiên. Điều này cho phép loại bỏ các cực tiểu địa phương trong quá trình thực thi. Tuy nhiên, các trụ gió thường chỉ liên kết với các trụ khác gần đó và các đường dây liên kết giữa chúng không được chéo lên nhau. Điều này làm cho việc sử dụng các thuật toán heuristic thường không được hiệu quả cao khi có nhiều ràng buộc trong quá trình thực hiện.

Phương pháp Sweep [9] là sự phát triển tiếp theo của phương pháp heuristic với việc đảm bảo không có sự chéo nhau giữa các đường dây kết nối các trụ gió với nhau. Phương pháp này được thực hiện bằng cách sắp xếp các trụ gió dựa trên góc tạo bởi các trụ gió và trạm kết nối xa bờ. Phương pháp này có nhược điểm chính là không thể lựa chọn được các kết nối tối ưu giữa các trụ gió trong một nhóm do chúng được xác định cứng dựa trên thứ tự góc kết nối tăng dần (hoặc giảm dần).

Bài báo đề xuất phương pháp một phương pháp mới để giải quyết bài toán tối ưu hệ thống dây dẫn hệ thống

điện gió xa bờ sử dụng thuật toán liên kết giữa phương pháp Sweep và thuật toán tối ưu hóa bầy đàn rời rạc kiểu xác suất (Selective Probabilistic Discrete Particle Swarm Optimization – SPDPSO). Việc lựa chọn nhóm dựa trên thuật toán Sweep cho phép không xuất hiện tình trạng chéo dây giữa các trụ gió và giảm không gian tìm kiếm các ứng viên phù hợp. Thuật toán SPDPSO cho phép tận dụng sức mạnh của các thuật toán tối ưu có chọn lựa một cách hiệu quả. Kết quả thực hiện dựa trên thuật toán được đề xuất cho thấy hiệu quả tốt hơn về chi phí đầu tư định khi so sánh với cấu trúc hệ thống dây dẫn đã được thi công trong thực tế.

2. BÀI TOÁN ĐỊNH TUYẾN CÁP ĐIỆN TRANG TRẠI GIÓ

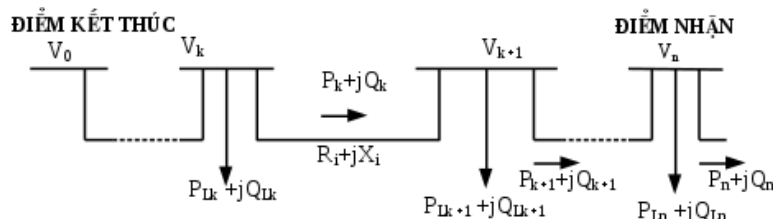
2.1. Tính toán phân bố công suất trên các cáp điện của trang trại gió

Đối với hệ thống cáp điện kết nối các trụ gió với nhau và đưa tới đầu gom thì vấn đề xác định công suất chạy trên các đường dây kết nối là rất quan trọng. Đây là bước quan trọng để xác định loại dây thích hợp. Do dây cáp điện có chi phí rất cao nên nếu chọn loại dây lớn hơn yêu cầu thì không hiệu quả về kinh tế, còn ngược lại chọn loại dây quá nhỏ thì không an toàn trong vận hành.

Với số lượng các trụ gió trong một trang trại gió thường rất lớn và cấu trúc phổ biến của chúng là đầu hình tia, do đó có nhiều phương pháp khác nhau để xác định công suất chạy trên các đường dây được đề xuất tương tự như xác định công suất truyền trên các đường dây của lưới điện phân phối hình tia.

Trong các phương pháp tính toán phân bố công suất được đề xuất, phương pháp tính toán kiểu Backward/Forward [10][11] được sử dụng phổ biến bởi tính đơn giản, yêu cầu ít các phép tính và đơn giản trong giải thuật.

Nguyên lý hoạt động của phương pháp Backward/Forward cải tiến được sử dụng thông qua việc tìm ra tất cả các điện áp nút của LDPP. Từ các điện áp này có thể tính trực tiếp dòng điện, tổn thất và các trạng thái ổn định khác cho LDPP hình tia thay thế cho phương pháp truyền thống nhằm giảm thời gian tính toán. Backward là tính dòng điện hoặc dòng công suất với điện áp được cập nhật và được tính bắt đầu từ nút cuối cùng về nút gốc. Forward là tính sụt áp với dòng điện hay dòng công suất cập nhật. Điện áp nút được cập nhật trong Forward bắt đầu từ nút gốc đến nút cuối cùng, được mô tả như Hình 35.



Hình 35. Sơ đồ đơn tuyến LDPP

Tính toán tổn thất công suất gần đúng trên các đường dây từ nút cuối lưới tính lên. Tiến hành đọc dữ liệu đường dây và tải. Trong cấu hình đường dây quy ước, nút nguồn đặt là nút số 2 và nút tải đặt là nút số 1. Giả sử nút (k+1) là nút cuối lưới thì nút này sẽ là nút loại 1 và có một liên kết đến nút thứ (k). Sau khi xác định xong, tiến hành tính tổn thất công suất trên đường dây liên kết với nút cuối lưới (k+1). Tổn thất công suất tác dụng và công suất phản kháng của đường dây (k, k+1) được xác định lần lượt theo các phương trình (1) và (2) như sau:

$$P_{loss(k,k+1)} = R_{(k,k+1)} \cdot \frac{P_{L,k+1}^2 + Q_{L,k+1}^2}{V_m^2} \quad (1)$$

$$Q_{loss(k,k+1)} = X_{(k,k+1)} \cdot \frac{P_{L,k+1}^2 + Q_{L,k+1}^2}{V_m^2} \quad (2)$$

Với $P_{loss(k,k+1)}$ là tổn thất công suất tác dụng trên đường dây đoạn từ nút k tới nút (k+1); $Q_{loss(k,k+1)}$ là tổn thất công suất phản kháng trên đường dây đoạn từ nút k tới nút (k+1). R, X lần lượt là tổng trở và tổng dẫn của đường dây, $P_{L,k+1}$, $Q_{L,k+1}$ lần lượt là công suất tác dụng và công suất phản kháng của phụ tải tại nút k+1,

V_m là điện áp nguồn lưới điện.

Sau khi tính xong tổn thất công suất và tổn thất điện áp trên đường dây thì ta tiến hành cộng dồn tải và tổn thất công suất trên nhánh vừa tính vào nút đầu của nhánh đó và tiến hành xóa bỏ dữ liệu đường dây và cắt đường dây ra khỏi lưới. Làm như vậy cho đến khi gặp nút nguồn (loại 2) thì dừng lại, tức là tổng số liên kết giữa các nút = 0.

$$P'_{L,k} = P_{L,k} + P_{L,k+1} + P_{loss(k,k+1)} \quad (3)$$

$$Q'_{L,k} = Q_{L,k} + Q_{L,k+1} + Q_{loss(k,k+1)} \quad (4)$$

Như vậy, ta có thể xác định được gần đúng công suất chạy trên các đường dây bằng cách sử dụng công suất chạy vào nút cuối đường dây được tính ra trong các công thức (3) và (4). Lưu ý là do tính chất ngược nhau của tải và máy phát, nên dòng công suất chạy ngược từ trụ gió về đầu máy biến áp tăng áp và ngược với lưới điện phân phối thông thường khi dòng điện chạy theo hướng từ trạm biến áp về các tải ở phía xa. Tuy nhiên, hoàn toàn có thể sử dụng phương pháp này để tính toán công suất chạy trên các đường dây.

a. Lựa chọn dây cáp điện của trang trại gió

Sau khi đã xác định được công suất chạy vào các đường dây gom công suất điện gió, giá trị công suất này được sử dụng để xác định kích cỡ dây mang được dòng điện với công suất tương ứng. Để có thông tin về giá bán dây và công suất truyền tải của dây, Houria Smail trong nghiên cứu [12] đã đề xuất thông số về cáp điện như được thể hiện trong Bảng 7

Bảng 7. Thông tin về một số loại cáp điện dành cho trang trại điện gió xa bờ

Loại dây	Cost (MEuro/km)	code	Pmax (MW)	R (Ω /km)	X (Ω /km)
3x120	0.31	1	18.2	0.14	0.11
3x150	0.36	2	20.3	0.11	0.11
3x185	0.42	3	22.8	0.09	0.11
3x240	0.49	4	26.2	0.07	0.1
3x300	0.52	5	30.4	0.06	0.1
3x400	0.6	6	33.9	0.04	0.1
3x500	0.69	7	38	0.03	0.09
3x630	0.64	8	40.4	0.03	0.09
3x800	0.87	9	44.3	0.02	0.09

Dựa vào công suất chạy trên đường dây, khi so sánh với công suất lớn nhất cho phép Pmax (MW) của mỗi loại dây tương ứng trong Bảng 7 sẽ chọn được loại kích thước tương ứng cho mỗi đường dây.

b. Chi phí mua sắm dây cáp điện của trang trại gió

Tổng chi phí mua sắm dây dẫn được xác định bằng tổng giữa tích chiều dài dây dẫn nhân với chi phí dây dẫn thể hiện trong cột 2 bảng Bảng 7. Nếu gọi chiều dài đường dây nối giữa hai trụ gió i và j là L_{ij} và giá bán dây điện của đường dây nối giữa hai trụ gió i và j là P_{ij} , tổng chi phí cho hệ thống thu gom điện gió sẽ được tính dựa trên phương trình (5) với $A_{ij} = 0$ nếu không có dây dẫn giữa hai trụ gió i và j và bằng 1 nếu ngược lại.

$$C_{total} = \sum P_{ij} * L_{ij} * A_{ij} \quad (5)$$

Như vậy, với cấu hình kết nối các trụ gió với nhau thông qua hệ thống cáp điện thì chiều dài và kích thước cáp điện sẽ khác nhau tương ứng. Điều này làm cho chi phí mua sắm dây điện cũng theo đó mà thay đổi theo. Các thuật toán tối ưu khác nhau được đề xuất để tìm được kết nối các đường dây sao cho chi phí mua dây là nhỏ nhất. Điều này là cần thiết nhằm tối ưu chi phí xây dựng trang trại điện gió.

c. Thuật toán tối ưu spd-pso

Thuật toán tối ưu bầy đàn-chuỗi rời rạc (SPD-PSO) [13] được đề xuất để giải quyết vấn đề bố trí tải sử dụng công thức ma trận nhị phân đại diện cho các phần tử trong giải pháp bầy đàn khi áp dụng khái niệm sử dụng vận tốc phần tử như kiểm soát xác suất trong việc cập nhật các yếu tố giải pháp trong các phần tử. Ngoài ra, để giữ gìn bản chất thuật toán PSO cơ bản ban đầu ở vị trí mới (giải pháp) của phần bị ảnh hưởng bởi vị trí hiện tại của nó, cũng như các vị trí tốt nhất của phần tử và toàn quần thể, một lựa chọn cho cơ chế trong thuật toán đề xuất được đưa ra để cho phép các yếu tố giải pháp cập nhật của một phần tử có nguồn gốc từ giải pháp hiện tại, phần tử tốt nhất hoặc giải pháp tốt nhất toàn cục. Do đó, dựa trên các đặc điểm thiết kế của đề xuất Thuật toán PSO rời rạc, phương pháp tối ưu hóa mới được đặt tên Tối ưu hóa xác suất hạt rời rạc chọn lọc (từ đó được ký hiệu là SPD-PSO).

Vị trí phần tử X và vận tốc V trong đề xuất thuật toán SPD-PSO là ma trận và các phương trình cập nhật vận tốc và phần tử đã được sửa đổi như thể hiện trong giải mã của thuật toán SPD-PSO mới được nêu trong công thức (6) và (7). Phần còn lại của phần này sẽ trình bày chi tiết các cơ chế được sử dụng để theo dõi và cập nhật vị trí của các phần tử trong SPD-PSO, cũng như công thức của ma trận vận tốc phần tử và tính toán các giá trị mục tiêu của biện pháp trong tổ hợp bố trí, trong đó mục tiêu tối ưu hóa là để giảm thiểu giá trị mục tiêu của các giải pháp được đề cử.

Tóm loại điểm khác nhau giữa PSO và SPD-PSO là công thức cập nhật vận tốc và vị trí mới của các cá thể để tạo ra quần thể mới:

$$V_k^{t+1} = I_{Q1 \geq Q2} \cdot (P_{best,k} - X_k^t) + I_{Q1 < Q2} \cdot (G_{best} - X_k^t) \quad (6)$$

$$X_k^{t+1} = X_k^t + I_{Q3} \cdot (V_k^{t+1}) \quad (7)$$

d. Thuật toán Sweep trong tối ưu chi phí mua sắm dây cáp điện của trang trại gió

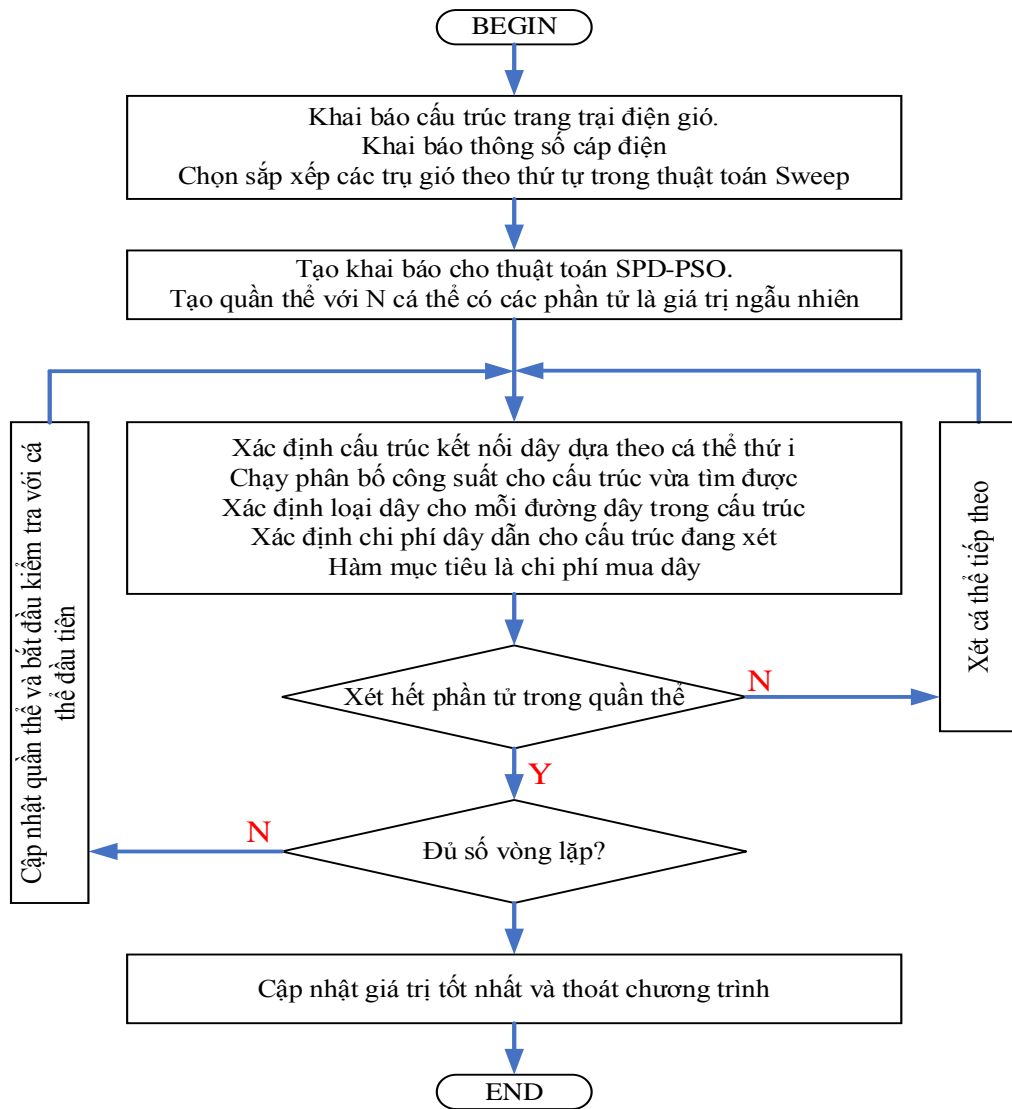
Trong nhiều trường hợp thực tế, công suất tối đa của cáp luôn bị giới hạn ở một mức nhất định. Những yêu cầu riêng về công suất truyền tải lớn luôn kèm chi phí cao hơn bởi các đơn đặt hàng riêng lẻ luôn chịu một chi phí cao hơn sản xuất hàng loạt. Điều này sẽ khống chế số lượng các trụ gió gắn vào các đường dây. Trên thực tế, những ràng buộc này đòi hỏi phải phân chia các tuabin theo cách cân bằng, để các tuabin trong mỗi nhóm tương ứng với một phát tuyến con của trạm biến áp và do đó có tổng công suất có thể được truyền bởi một dây cáp duy nhất. Ngoài ra, vấn đề chống chéo dây dẫn giữa các trụ gió cũng được xem xét vì điều này giúp giảm không gian tìm kiếm cấu trúc phù hợp.

Thuật toán Sweep được Gillet and Miller [9] đề xuất lần đầu năm 1974 và được Fischetti và Pisinger [6] áp dụng hiệu quả trong tối ưu định tuyến hệ thống cáp điện gió xa bờ. Thuật toán Heuristic có ba tham số:

- Thông số trụ gió được chọn làm trụ gió ban đầu để quét.
- Hướng quét các trụ gió (thuận chiều hay ngược chiều kim đồng hồ, từ trên xuống dưới, từ trái sang phải).
- Số trụ gió trong một nhóm. Các trụ điện gió trong một nhóm sẽ liên kết với nhau và kết nối đến trạm biến áp gom điện. Số lượng nhỏ nhất và lớn nhất của các trụ gió trong mỗi nhóm phụ thuộc vào khả năng mang tải của các cáp điện. Cáp điện có khả năng mang được công suất của ít trụ gió nhất được chọn làm số trụ gió nhỏ nhất và ngược lại. Đây là các thông số đã biết ở đầu vào thuật toán. Việc xác định số trụ gió trong một nhóm được thực hiện thủ công theo phương pháp thử sai.

Áp dụng thuật toán SPD-PSO kết hợp thuật toán Sweep trong tối ưu chi phí mua sắm dây cáp điện của trang trại gió.

Trong bài báo này, một thuật toán tối ưu định tuyến hệ thống thu gom công suất điện gió được xây dựng bằng cách phối hợp hoạt động giữa hai thuật toán SPD-PSO kết hợp với thuật toán Sweep. Thuật toán SPD-PSO sẽ xác định chính xác số lượng các trụ điện gió trong một nhóm trụ gió. Thuật toán Sweep có nhiệm vụ đưa các trụ gió vào các nhóm và xác định các kết nối bên trong mỗi trụ với tiêu chí là đường dây từ một trụ đến các trụ còn lại là ngắn nhất. Lưu đồ thực hiện của thuật toán tối ưu lai được đề xuất như trong Hình 36 bên dưới.

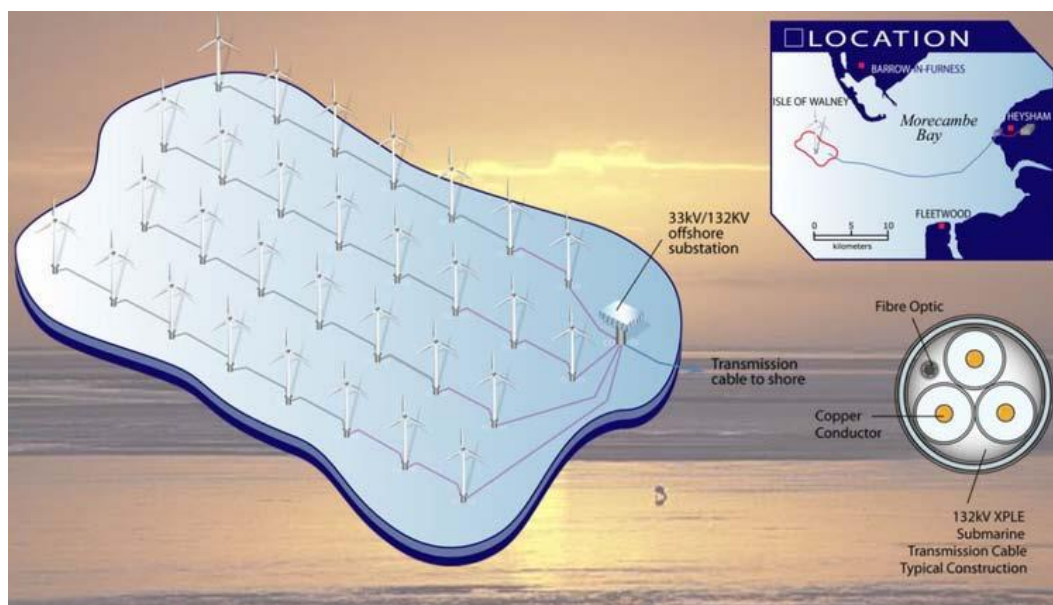


Hình 36. Lưu đồ thuật toán tối ưu định tuyến cáp ngầm trang trại gió được đề xuất

4. KẾT QUẢ MÔ PHỎNG

a. Mô hình trang trại gió “Barrow Offshore Wind Farm”

Để đánh giá hiệu quả của thuật toán tối ưu đường dây kết nối các trang trại gió thu được, trang trại gió xa bờ “Barrow Offshore Wind Farm” được sử dụng để so sánh các kết quả thu được. Các thông số của trang trại gió được đề xuất trong các nghiên cứu [12][14]. Đây là một trang trại điện gió ngoài khơi nằm ở phía Tây nước Anh. Khu vực đặt “Barrow Offshore Wind Farm” là một khu vực có diện tích khoảng 10 km². Trang trại điện gió này có tổng cộng 30 tuabin gió Vestas-V9 (Hình 37) được phân bố thành bốn hàng, hai với bảy và hai với tám tuabin gió. Khoảng cách giữa các tuabin gió là khoảng 500 m theo hướng gió bình thường và giữa các hàng là khoảng 750 m theo hướng gió chính. Công suất định mức của mỗi tuabin gió là 3 MW (tổng công suất lắp đặt là 90 MW). “Barrow Offshore Wind Farm” có một trạm biến áp ngoài khơi 132/33 kV với một máy biến áp công suất 120 MVA. Nguồn điện do “Barrow Offshore Wind Farm” tạo ra được xuất ra điểm đầu nối lưới điện thông qua cáp truyền tải dưới biển dài 27.5 km.



Hình 37 Vị trí địa lý và cấu trúc trang trại gió xa bờ “Barrow Offshore Wind Farm” [12]

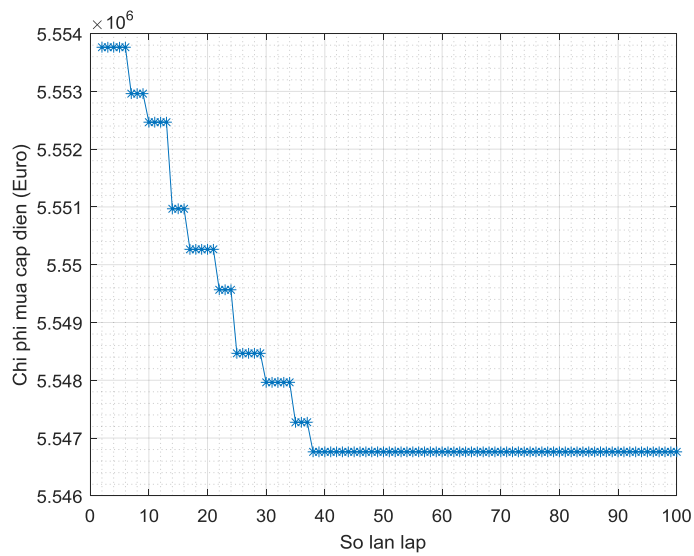
b. Kết quả mô phỏng tối ưu định tuyến cáp điện trang trại gió

Thuật toán tối ưu SPD-PSO lai thuật toán Sweep được đề xuất áp dụng để đánh giá hiệu quả. Với các thông số của chương trình được cho trong Bảng 8, kết quả tốt nhất qua mỗi lần lặp khi thực thi thuật toán đề xuất được thể hiện như trong Hình 38.

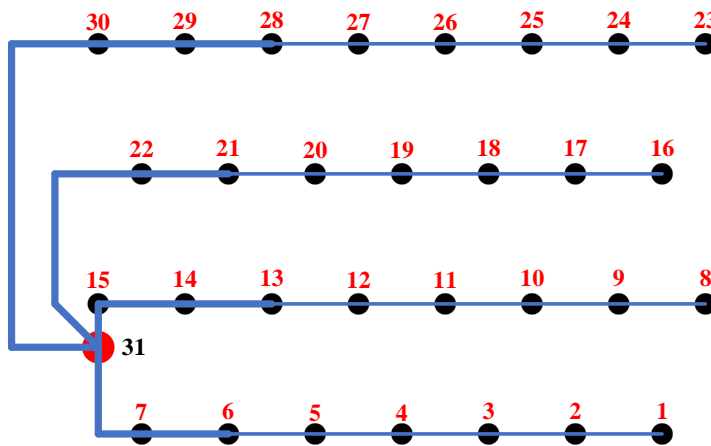
Bảng 8. Thông số Thuật toán tối ưu SPD-PSO lai thuật toán Sweep

Thông số	Giá trị
Số biến số trong một cá thể	6
Số cá thể trong quần thể	10
Số vòng lặp tối đa	100
Hệ số nhận thức tự thân Φ	0.7
Hệ số nhận thức xã hội δ	0.3
Số trụ điện gió nhỏ nhất trong một nhóm	4
Số trụ điện gió lớn nhất trong một nhóm	10

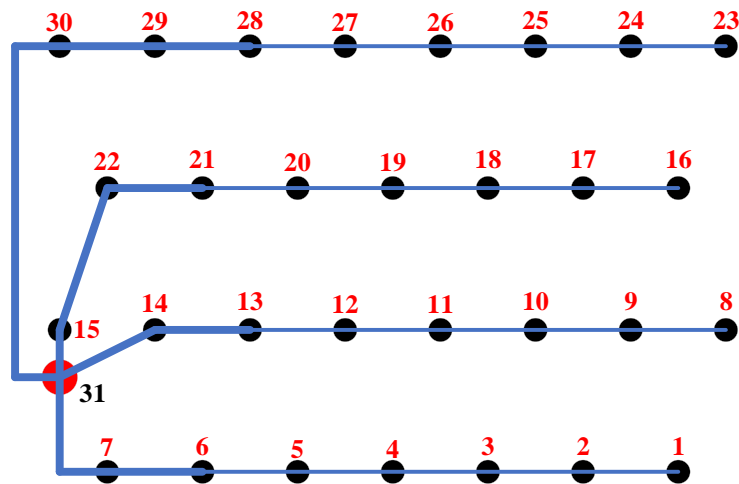
Kết quả kết nối khi được tối ưu chi phí đường dây được thể hiện như trong Hình 40 và kết nối thực tế được thể hiện như trong Hình 39. Nhận thấy rằng khi được tối ưu thì cấu trúc đường dây đề xuất không còn giống cấu trúc hiện hữu. Kết quả chi phí đường dây được xác định dựa trên cấu trúc đường dây và chi phí của mỗi đường dây được cho trong Bảng 7. Có thể nhận thấy rằng chi phí đường dây hiện hữu là 5711613 (Euro) và chi phí cho cấu trúc tối ưu mới là 5546761 (Euro). Như vậy bằng cách tối ưu kết nối đường dây điện thì đã giảm chi phí đi dây 2.89 % chi phí đầu tư đường dây ban đầu.



Hình 38 Kết quả chi phí đường dây gom điện trang trại gió qua mỗi lần lặp



Hình 39 Cấu trúc đường dây kết nối trang trại gió hiện tại



Hình 40 Cấu trúc đường dây kết nối trang trại gió tối ưu từ giải thuật

5. KẾT LUẬN

Bài báo đã đề xuất giải thuật tối ưu lai để giải quyết bài toán tối ưu đường dây dẫn điện giữa các trụ điện gió. Phương pháp tối ưu được đề xuất cho phép xác định cấu trúc đường dây mới có chi phí cấp điện là thấp hơn so với cấu trúc đã được lắp đặt trước đây. Điều này cho phép tiết kiệm hơn nữa chi phí đầu tư hệ thống điện gió nếu được áp dụng ngay trong quá trình thiết kế hệ thống cáp dẫn điện. Ưu điểm chính của phương pháp đề xuất là không gian tìm kiếm giải pháp được giảm đáng kể: số lượng nhóm trụ gió và số lượng trụ gió trong nhóm là những biến duy nhất được xem xét trong bài toán tối ưu hóa dựa trên thuật toán tối ưu hoá bầy đàn rời rạc kiểu xác suất. Không gian tìm kiếm được thu gọn dựa trên thuật toán Sweep.

TÀI LIỆU THAM KHẢO

- [1] “Tập đoàn Điện lực Việt Nam: Thích ứng an toàn, linh hoạt, hiệu quả.” <https://www.qdnd.vn/kinh-te/tin-tuc/tap-doan-dien-luc-viet-nam-san-sang-thich-ung-an-toan-linh-hoat-va-hieu-qua-683434> (accessed Jun. 06, 2022).
- [2] “Tình hình cung cấp than cho sản xuất điện còn nhiều khó khăn, EVN kêu gọi tiết kiệm điện.” <https://www.evn.com.vn/d6/news/Tinh-hinh-cung-cap-than-cho-san-xuat-dien-con-nhieu-kho-khan-EVN-keu-goi-tiet-kiem-dien-6-12-30271.aspx> (accessed Jun. 06, 2022).
- [3] “Nội dung họp báo Chính phủ chuyên đề về Dự án điện hạt nhân.” <https://baochinhphu.vn/noi-dung-hop-bao-chinh-phu-chuyen-de-ve-du-an-dien-hat-nhan-102212659.htm> (accessed Jun. 06, 2022).
- [4] D. Cazzaro, M. Fischetti, and M. Fischetti, “Heuristic algorithms for the Wind Farm Cable Routing problem,” *Appl. Energy*, vol. 278, no. May, p. 115617, Nov. 2020, doi: 10.1016/j.apenergy.2020.115617.
- [5] M. Fischetti and D. Pisinger, “Mathematical Optimization and Algorithms for Offshore Wind Farm Design: An Overview,” *Bus. Inf. Syst. Eng.*, vol. 61, no. 4, pp. 469–485, Aug. 2019, doi: 10.1007/s12599-018-0538-0.
- [6] M. Fischetti and D. Pisinger, “Optimizing wind farm cable routing considering power losses,” *Eur. J. Oper. Res.*, vol. 270, no. 3, pp. 917–930, Nov. 2018, doi: 10.1016/j.ejor.2017.07.061.
- [7] P. Hou, W. Hu, and Z. Chen, “Optimisation for offshore wind farm cable connection layout using adaptive particle swarm optimisation minimum spanning tree method,” *IET Renew. Power Gener.*, vol. 10, no. 5, pp. 694–702, May 2016, doi: 10.1049/iet-rpg.2015.0340.
- [8] A. Cerveira, A. de Sousa, E. J. S. Pires, and J. Baptista, “Optimal Cable Design of Wind Farms: The Infrastructure and Losses Cost Minimization Case,” *IEEE Trans. Power Syst.*, vol. 31, no. 6, pp. 4319–4329, Nov. 2016, doi: 10.1109/TPWRS.2016.2521700.
- [9] B. E. Gillett and L. R. Miller, “A Heuristic Algorithm for the Vehicle-Dispatch Problem,” *Oper. Res.*, vol. 22, no. 2, pp. 340–349, Apr. 1974, doi: 10.1287/opre.22.2.340.
- [10] S. Ouali and A. Cherkaoui, “An Improved Backward/Forward Sweep Power Flow Method Based on a New Network Information Organization for Radial Distribution Systems,” *J. Electr. Comput. Eng.*, vol. 2020, pp. 1–11, Jan. 2020, doi: 10.1155/2020/5643410.
- [11] G. W. Chang, S. Y. Chu, and H. L. Wang, “An Improved Backward/Forward Sweep Load Flow Algorithm for Radial Distribution Systems,” *IEEE Trans. Power Syst.*, vol. 22, no. 2, pp. 882–884, May 2007, doi: 10.1109/TPWRS.2007.894848.
- [12] H. Smail, R. Alkama, and A. Medjdoub, “Optimal design of the electric connection of a wind farm,” *Energy*, vol. 165, pp. 972–983, Dec. 2018, doi: 10.1016/j.energy.2018.10.015.

- [13] Ming Shu Seah, Whye Loon Tung, and T. Banks, “A Novel Discrete Particle Swarm Optimization approach to large-scale survey planning,” in 2015 11th International Conference on Natural Computation (ICNC), Aug. 2015, vol. 2016-Janua, pp. 261–268. doi: 10.1109/ICNC.2015.7378001.
- [14] H. D. Betancourt, “Stochastic Optimization of Offshore Wind Power Plants Operation For Maximizing Energy Generation focusing on the electric power system optimization and cost minimization,” 2016.

YSC4F.217

CẢI TIẾN TÍNH CHẤT ĐIỆN CỦA VẬT LIỆU TGS SỬ DỤNG ỐNG NANOCACBON ĐA VÁCH DẠNG OXI HÓA GIÚP MỞ RỘNG VÙNG HOẠT ĐỘNG CỦA CÁC CẢM BIẾN HỒNG NGOẠI

NGUYỄN HOÀI THƯƠNG, HÀ VĂN ĐẠI

*Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
nguyenthuongfee@iuh.edu.vn*

Tóm tắt. Trong nghiên cứu này, chúng tôi chọn cải tiến vật liệu TGS, một loại vật liệu sắt điện cổ điển được ứng dụng rộng rãi trong các cảm biến nhiệt hồng ngoại nhưng khả năng ứng dụng đã trở nên bão hòa do pha sắt điện - vùng quan trọng cho các ứng dụng - tồn tại ở nhiệt độ khá thấp ($< 49\text{ }^{\circ}\text{C}$). Bằng cách kết hợp với ống nanocarbon đa vách dạng oxit hóa cho thấy, nhiệt độ chuyển pha TGS có thể lên tới $52.3\text{ }^{\circ}\text{C}$ ở điều kiện khô ráo và $57.7\text{ }^{\circ}\text{C}$ dưới độ ẩm tương đối 90%. Ngoài ra, vùng tần số xuất hiện hiệu ứng điện cực – một trong những hiệu ứng có thể làm giảm hiệu suất hoạt động của vật liệu ở tần số thấp - cũng được khảo sát.

Từ khóa. Vật liệu sắt điện, sự chuyển pha, hiệu ứng kích thước, hiệu ứng điện cực.

IMPROVING ELECTRICAL PROPERTIES OF TGS USING OXIDIZED MULTIWALLED CARBON NANOTUBES FOR EXPANDING THE OPERATING REGION OF IR SENSORS

Abstract. In this study, we chose improving TGS - a classical ferroelectric that had been widely applied in infrared sensors, but its applicability has become saturated because the ferroelectric phase – a characteristic region essential for applications - exists at fairly low temperatures ($<49\text{ }^{\circ}\text{C}$). By combining with oxidized multiwalled carbon nanotubes (OMWCNT), the TGS phase transition temperature shifted to $52.3\text{ }^{\circ}\text{C}$ in dry conditions, and $57.7\text{ }^{\circ}\text{C}$ under a relative humidity of 90%. In addition, the frequency region for the existence of the electrode effect - one of the effects that can reduce the performance of a material at low frequencies - was also investigated.

Keywords. Ferroelectrics, phase transition, size effects, electrode effects.

1. GIỚI THIỆU

Sự phát triển nhanh chóng của công nghệ điện – điện tử luôn song hành cùng quá trình khám phá liên tục các vật liệu mới nhằm đáp ứng nhu cầu ngày càng cao của thực tiễn. Trong bức tranh tổng thể về thành công của các thiết bị công nghệ, vật liệu sắt điện không thể không nhắc đến với các ứng dụng phổ biến trong transistor [1], máy phát nano [2], cảm biến [3], bộ nhớ [4], tích trữ năng lượng [5], màn hình tinh thể lỏng [6], v.v. ...

Triglycine sulfate (TGS) cũng là một trong những vật liệu sắt điện được biết đến rộng rãi trong giới nghiên cứu và ứng dụng thực tiễn như máy dò hồng ngoại, hình ảnh nhiệt và các ứng dụng cảm biến hồng ngoại [7]. Cảm biến nhiệt điện dựa trên TGS là nhạy cảm đồng đều với các bức xạ trong dải bước sóng từ tia cực tím đến tia hồng ngoại xa, và không yêu cầu làm mát cho hoạt động so với máy dò lượng tử, nơi yêu cầu làm mát ở nhiệt độ thấp.

Theo thời gian, các vật liệu điện – điện tử nói chung và vật liệu sắt điện nói riêng ngày càng trở nên thông minh hơn, kích thước nhỏ gọn hơn nhưng lại tích hợp nhiều tính năng hơn. Để làm được điều này, một phương pháp đơn giản và hiệu quả là cải tiến tính chất điện của các vật liệu truyền thống sẵn có nhờ kết hợp với một hoặc nhiều vật liệu khác có cấu trúc nano.

Ống nano cacbon đa vách (MWCNT) được biết đến với các ứng dụng nổi bật trong chấn nhiễu điện từ, các bộ chuyển hóa năng lượng, tụ điện, cảm biến [8][9]. Ưu điểm nổi bật của ống nano cacbon nằm ở khả năng dẫn điện tốt, độ bền nhiệt – điện cao, có độ rộng bề mặt lớn nhờ kích thước nano cùng với mật độ thấp giúp tương tác điện dễ dàng, từ đó giúp cải thiện đáng kể tính chất điện của các vật liệu chứa nó. Việc sử dụng ống nano cacbon đa vách kết hợp với nhiều loại vật liệu sắt điện khác nhau không phải là vấn đề mới. Hàng loạt nhà nghiên cứu trong các công trình [10],[11] đã sử dụng ống nano cacbon đa vách để nâng cao khả năng chuyển hóa năng lượng cơ học thành tín hiệu điện, cải thiện hiệu suất chuyển hóa quang điện, tính chất sắt điện cũng như độ bền nhiệt

Mặc dù ý tưởng sử dụng ống nano cacbon không còn xa lạ nhưng còn nhiều vấn đề chưa được giải quyết triệt để. Thứ nhất, việc sử dụng ống nano cacbon thường được sử dụng dạng thô sau khi tổng hợp để đưa thẳng vào vật liệu, trong khi ống nano cacbon sau khi bị oxy hóa (OMWCNT) có sự biến đổi về thành phần cấu trúc và sự ảnh hưởng của nó lên tính chất điện của vật liệu sẽ thay đổi. Tuy vậy, điều này lại ít được đề cập đến. Thứ hai, sự chuyển pha của vật liệu sắt điện – vấn đề quan trọng trong việc ứng dụng sắt điện làm cảm biến nhiệt ít được xem xét khi nghiên cứu ảnh hưởng của ống nano cacbon. Nói cách khác, việc sử dụng ống nano cacbon đa vách ở dạng thông thường hoặc sau khi bị oxy hóa để cải tiến tính chất điện của vật liệu sắt điện truyền thống rất đáng quan tâm và cần phải làm rõ hơn nữa. Một mặt để làm rõ bản chất khoa học về mặt cơ bản, mặt khác để mở rộng tiềm năng ứng dụng của các vật liệu sẵn có. Xuất phát từ tính cấp thiết đó, trong nghiên cứu này, ống nano cacbon đa vách được dùng để chế tạo vật liệu nanocomposite với TGS, đồng thời nghiên cứu vùng tồn tại của pha sắt điện TGS – vùng quan trọng cho các ứng dụng thực tiễn. Ngoài ra, ảnh hưởng của điện cực – hiệu ứng có thể làm nhiễu tín hiệu điện đi ra từ vật liệu cũng được khảo sát. Bên cạnh đó, các thí nghiệm về sự ảnh hưởng của độ ẩm lên các tham số điện cũng được thực hiện.

2. PHƯƠNG PHÁP CHẾ TẠO VẬT LIỆU VÀ THIẾT BỊ ĐO

2.1. Quy trình chế tạo vật liệu và thiết bị đo

Vật liệu nanocomposite sắt điện trên nền tảng OMWCNT và TGS được chế tạo bằng phương pháp hóa học thông qua tổ các máy chế tạo vật liệu chuyên dụng như máy phát siêu âm (Qsonica Q70), máy khuấy từ gia nhiệt (IKA C-MAG HS7), kính hiển vi điện tử quét (A FE-SEM S4800 Hitachi, Japan), máy quang phổ Bruker Tensor 37 (USA).

Trước hết, MWCNT cần phải được oxy hóa bằng cách đưa vào dung dịch HNO_3 (300 mL), sau đó đun sôi và khuấy bằng máy sấy từ gia nhiệt (IKA C-MAG HS7) ở nhiệt độ $110\text{ }^\circ\text{C}$ trong 5h. Sau khi khuấy xong, dung dịch được để nguội ngoài không khí. Hỗn hợp thu được sẽ được rửa sạch và ly tâm trong nước ion ở $\text{pH} = 7$, sau đó sấy khô ở nhiệt độ $110\text{ }^\circ\text{C}$ trong 12h. Kết quả nhận được sẽ là ống nanocarbon đa vách đã bị oxy hóa và sẵn sàng để tổng hợp vật liệu nanocomposite với TGS. Tiếp theo, một lượng xác định OMWCNT được đưa vào dung dịch bão hòa TGS ở $20\text{ }^\circ\text{C}$ theo tỉ lệ khối lượng OMWCNT/TGS bằng 5%. Để tăng sự phân tán của ống nanocarbon vào vật liệu, máy phát siêu âm Qsonica Q70 được đưa vào hỗn hợp nhận được cùng với hệ thống khuấy tự nhiên của IKA C-MAG HS7. Công đoạn này được duy trì trong 6h trong bình kín để hạn chế sự thất thoát hơi nước trước khi mở ra bốc hơi tự nhiên. Cuối cùng, một hỗn hợp sệt được hình thành và đưa đi gia nhiệt ở $110\text{ }^\circ\text{C}$ trong 4h ta sẽ thu được nanocomposite OMWCNT/TGS. Vật liệu nanocomposite MWCNT/TGS cũng được tổng hợp bằng quy trình hoàn toàn tương tự, chỉ khác ở giai đoạn đầu là ống nanocarbon đa vách không cần oxy hóa trước khi bắt đầu tổng hợp vật liệu. Vật liệu chế tạo xong sẽ được kiểm tra hình thái và nhóm chức năng hóa học bằng kính hiển vi điện tử quét (A FE-SEM S4800 Hitachi, Japan), máy quang phổ Bruker Tensor 37 (USA). Nếu kết quả kiểm tra đáng tin cậy, vật liệu thu được sẽ tiếp tục dùng để kiểm tra sự chuyển pha và hiệu ứng điện cực trên hai tổ hợp máy LCR-821 và máy phân tích trở kháng Solartron 1260A (UK) kết nối với modul mở rộng INTERFACE 1269A (Hình 1)



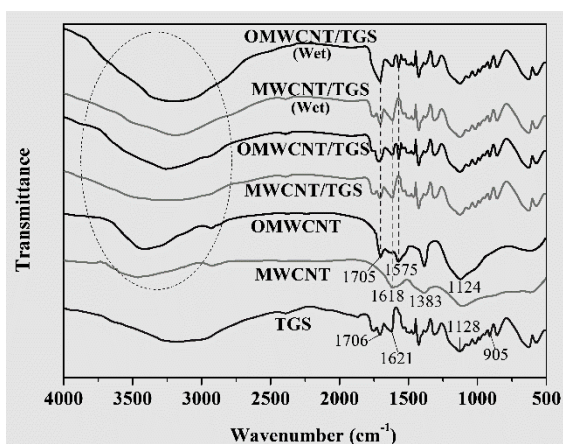
Hình 1. Máy phân tích trở kháng Solartron SI 1260 (1) và modul mở rộng Solartron 1296 (2)

3. KẾT QUẢ THỰC NGHIỆM

3.1. Cấu trúc vật liệu

Trước hết, phổ hồng ngoại cho thấy độ tin cậy rất cao của các vật liệu ban đầu MWCNT và TGS với các đỉnh hấp thụ đặc trưng C-H (2930 cm^{-1}), COOH (1618 cm^{-1}), C=C (1383 and 1124 cm^{-1}) trong MWCNT, và C=O (1706 cm^{-1} , COO- (1621 cm^{-1}), SO $^{4-}$ ($1128 - 909\text{ cm}^{-1}$) của TGS. Bên cạnh đó, đỉnh hấp thụ khá sâu thu được ở 3436 cm^{-1} (MWCNT) hoặc 3250 cm^{-1} (TGS) có liên quan trực tiếp đến nhóm OH. Trong vật liệu TGS, đỉnh hấp thụ này còn gây ra bởi liên kết -H và C-H. Sau khi bị oxi hóa, một vài bất thường đã được ghi nhận trong phổ hồng ngoại đối với OMWCNT, bao gồm sự xuất hiện 2 đỉnh mới tại 1705 (OH) và 1575 cm^{-1} (C=O). Điều này cũng được báo cáo trong các nghiên cứu trước đây. Đặc biệt, quá trình oxi hóa làm cho đỉnh hấp thụ tại 3436 cm^{-1} trở nên sâu và mở rộng hơn. Nguyên nhân có thể là do sự gia tăng đáng kể các nhóm OH trên bề mặt ống nanocarbon đa vách sau khi bị oxi hóa. Đối với các mẫu nanocomposite MWCNT/TGS và OMWCNT/TGS đã bị loại bỏ nước, phổ hồng ngoại hầu như chỉ thể hiện các đỉnh hấp thụ của TGS. Có thể hiểu là do TGS chiếm hàm lượng lớn trong vật liệu. Tuy nhiên, các đỉnh đặc trưng tại 1705 , 1575 cm^{-1} (OMWCNT) hoặc 1618 , 1383 cm^{-1} (MWCNT) cũng được tìm thấy. Hầu hết những sự thay đổi lớn đều ghi nhận sau khi lưu trữ mẫu trong điều kiện ẩm ướt với sự mở rộng dải $3500 - 2800\text{ cm}^{-1}$. Như đã trình bày ở trên, điều này không quá khó hiểu vì số lượng nhóm OH trong OMWCNT sau khi bị oxi hóa có khả năng giữ nước rất tốt. Thêm vào đó, ảnh hưởng của độ ẩm dẫn đến sự hấp thụ mạnh hơn tại 1705 cm^{-1} .

Kết quả khảo cấu trúc cho thấy có sự tồn tại của các nhóm OH trên bề mặt của ống nanocarbon sau khi bị oxi hóa rất nhạy với độ ẩm của môi trường. Những thông tin trên rất hữu ích trong việc giải thích các tính chất điện trong các phép đo tiếp theo trong nghiên cứu này. Đồng thời giúp định hướng trong quá trình cải tiến vật liệu để phù hợp hơn với các ứng dụng thực tế. Cần lưu ý thêm rằng, lớp OH có đặc tính cách điện tốt nên sẽ giúp ngăn sự ngắn mạch trong vật liệu dưới tác dụng của điện áp.



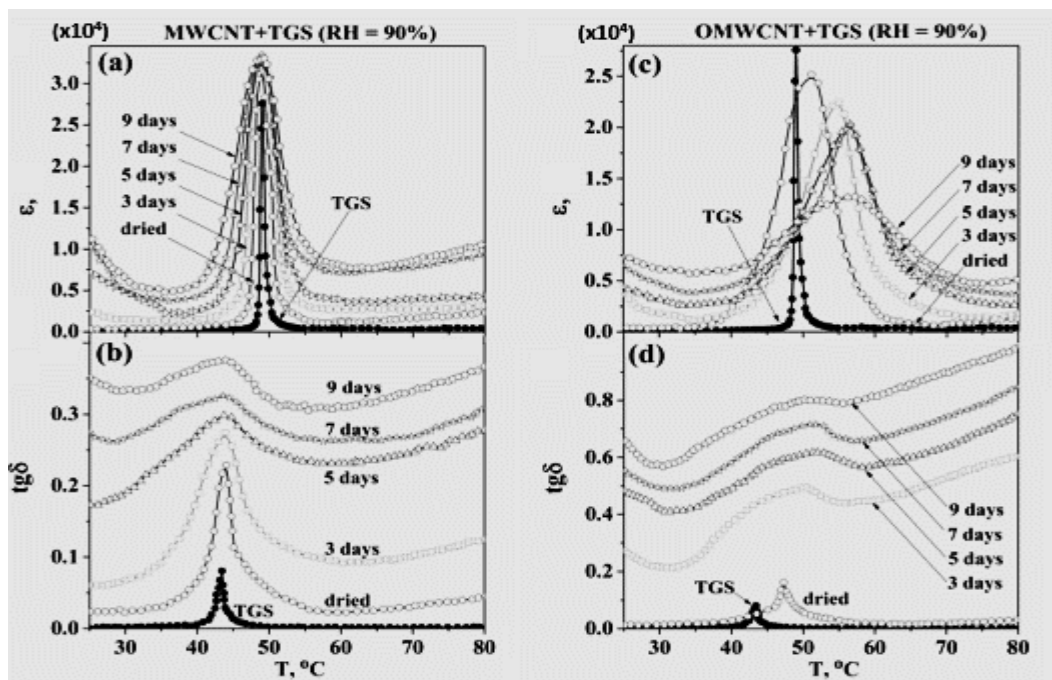
Hình 2. Kết quả phân tích nhóm chức của các thành phần vật liệu và nanocomposite MWCNT/TGS và OMWCNT ở điều kiện khô ráo và độ ẩm tương đối 90%

3.2. Khảo sát sự chuyển pha

Kết quả nghiên cứu sự chuyển pha ở điều kiện khô và ẩm ướt được trình bày ở hình 3. Kết quả cho thấy, sự phụ thuộc của hằng số điện môi vào nhiệt độ thu được trong thực nghiệm khá phù hợp với định luật Curie-Weiss [12]:

$$\varepsilon = \frac{C}{|T - T_c|} \quad (1)$$

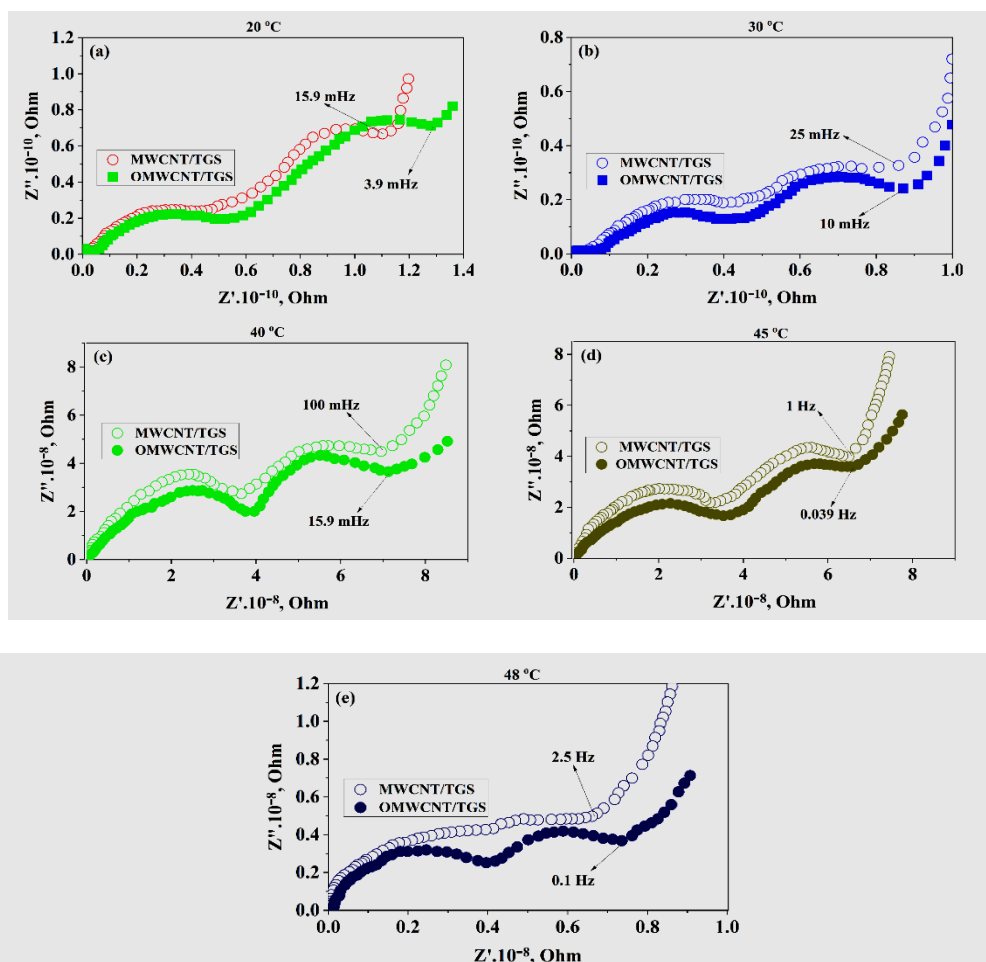
Trong đó, ε là hằng số điện môi, C là hằng số phụ thuộc vào bản chất từng loại vật liệu sắt điện, T_c là nhiệt độ chuyển pha và T là nhiệt độ môi trường. Việc tuân theo định luật này chứng tỏ ống nanocarbon đa vách dạng thường và oxi hóa không làm thay đổi bản chất sắt điện TGS, ngược lại còn giúp cải thiện hai tham số quan trọng là nhiệt độ chuyển pha. Mặc dù hằng số điện môi trong trường hợp sử dụng OMWCNT có giảm nhưng vẫn khá cao và không ảnh hưởng nhiều về mặt ứng dụng. Dưới ảnh hưởng của độ ẩm tương đối tương đối cao (RH=90%), sự chuyển pha vật liệu có sự ảnh hưởng rất rõ rệt, đặc biệt đối với sự tham gia của ống nanocarbon đa vách bị oxi hóa (Hình 3). Có thể thấy, nhiệt độ chuyển pha của nanocomposite OMWCNT/TGS được cải thiện đáng kể (lên đến 55.2 °C) sau 3 ngày lưu trữ. Việc kéo dài thời gian lưu trữ không làm tăng đáng kể nhiệt độ chuyển pha nhưng lại có xu hướng kéo hằng số điện môi xuống thấp hơn. Mặc dù vậy, hằng số điện môi dần đạt giá trị ổn định khoảng 10.000 nếu kéo dài thêm. Giá trị này vẫn chấp nhận được trong điều kiện ứng dụng thực tế. Trong điều kiện có độ ẩm cao như Việt Nam, kết quả nghiên cứu này có ý nghĩa quan trọng. Đối với cả hai vật liệu MWCNT/TGS và OMWCNT/TGS, sự ảnh hưởng của độ ẩm gây ra sự tổn hao rất cao, đây là nhược điểm cần cải thiện.



Hình 3. Kết quả khảo sát sự chuyển pha của vật liệu MWCNT/TGS (a, b) và OMWCNT/TGS (c,d) trong điều kiện khô và ẩm ướt trong các thời gian lưu trữ khác nhau. Kết quả của TGS tinh khiết cũng được đưa vào để so sánh.

3.3. Khảo sát hiệu ứng điện cực

Để thấy được sự xuất hiện của hiệu ứng điện cực, thông thường, chúng ta cần phân tích biểu đồ Nyquist $Z''(Z')$ [13], trong đó Z' là phần thực hay điện trở thuần và Z'' - phần ảo của trở kháng (Hình 4).



Hình 4. Biểu đồ Nyquist $Z''(Z')$ đối với composite MWCNT/TGS và OMWCNT/TGS ở nhiệt độ 20 °C (a), 30 °C (b), 40 °C (c), 45 °C (d) và 48 °C (e)

Kết quả trên cho thấy, đối với các composite được khảo sát, ngoài hai cung tròn đặc trưng cho TGS còn có sự xuất hiện của phần đuôi (Hình 4). Phần này tương ứng với tần số thấp nhất trong dải khảo sát, nghĩa là thời gian tích thoát rất cao. Hay nói cách khác, phần bất thường này trên biểu đồ Nyquist gây ra do các điện tích không gian bị kẹt lại ở lớp phân cách giữa điện cực và composite dẫn đến dịch chuyển chậm hơn sự thay đổi của tần số. Đây chính là bản chất của hiệu ứng điện cực. Có thể nhận thấy, tần số xuất hiện của hiệu ứng điện cực tăng khi tăng nhiệt độ, nghĩa là thời gian tích thoát của các điện tích giảm. Điều này phản ánh bản chất bán dẫn của vật liệu. Đặc biệt hơn, đối với ống nanocarbon đa vách bị oxi hóa, hiệu ứng điện cực xuất hiện yếu hơn và ở tần số thấp hơn so với ở dạng thường. Đây là điều bình thường bởi OMWCNT có khả năng dẫn điện kém hơn. Hiệu ứng điện cực tồn tại nhưng trong khoảng tần số khá thấp, chỉ nhỏ hơn 2.5 Hz trong vùng phân cực của vật liệu. Đối với các cảm biến hồng ngoại sử dụng vật liệu TGS, các tần số hiển nhiên lớn hơn rất nhiều, do đó sự ảnh hưởng của hiệu ứng điện cực là không đáng kể. Tuy nhiên, đối với từng loại điện cực và công nghệ chế tạo điện cực khác nhau, vùng ảnh hưởng này sẽ khác nhau. Trong các khảo sát trên, các điện cực được chế tạo khá thô sơ và chất lượng không cao. Nếu trong điều kiện chế tạo thực tế bằng công nghệ tiên tiến hơn, ảnh hưởng của hiệu ứng điện cực này có thể sẽ yếu hơn. Đây là kết quả triển vọng.

KẾT LUẬN

Kết quả nghiên cứu thu được đã góp một phần nhỏ trong nỗ lực chung của các nhà nghiên cứu bằng việc phát triển thành công hai loại vật liệu mới trên nền tảng vật liệu sắt điện TGS truyền thống với ống nano cacbon đa vách dạng thông thường và dạng oxi hóa. Bằng phương pháp đo đặc thực nghiệm kết hợp với

phân tích lý thuyết cho thấy, ống nano cacbon đa vách bị oxi hóa đáp ứng nhu cầu tốt hơn trong việc mở rộng vùng hoạt động của vật liệu TGS từ 49 °C lên 52.3 °C ở điều kiện khô ráo hoặc lên đến 57.7 °C trong điều kiện ẩm ướt. Đây là kết quả rất khả quan trong việc mở rộng vùng hoạt động của các cảm biến hồng ngoại. Trong điều kiện khảo sát với hàm lượng ống nano cacbon trong vật liệu không vượt quá 5% và được tán nhỏ bằng sóng siêu âm, độ ẩm tương đối 90% vẫn không gây ra ngắn mạch trong vật liệu. Ngoài ra, vùng tần số lớn hơn 2.5 Hz sẽ giúp vật liệu tránh được ảnh hưởng của hiệu ứng điện cực (trong điều kiện biên độ điện áp AC đặt vào nhỏ hơn 10 V/cm). Mặc dù vậy, hệ số tổn hao vẫn còn cao, đặc biệt khi sử dụng vật liệu trong điều kiện độ ẩm cao phần nào sẽ làm giảm hiệu suất thiết bị trong ứng dụng.

TÀI LIỆU THAM KHẢO

- [1] G. Tan *et al.*, Diode-like rectification characteristics of BiFeO₃-based $Zn_{1-x}Ni_xFe_2O_4$ bilayered films for application of ferroelectric field effect transistors, *J. Alloys Compd.* Vol. 851, pp.156-818, 2021.
- [2] Y.-P. Jiang *et al.*, Layer-dependent solvent vapor annealing on stacked ferroelectric P(VDF-TrFE) copolymers for highly efficient nanogenerator applications, *Polymer.* Vol. 204, pp. 122-822, 2020.
- [3] Z. L. Wang *et al.*, Progress in nanogenerators for portable electronics, *Mater. Today.* Vol. 15, pp.532-543, 2012.
- [4] Y. Chen *et al.*, Ferroelectric domain dynamics and stability in graphene Oxide-P(VDF-TrFE) multilayer films for ultra-high-density memory application, *Carbon.* Vol.144, pp. 15-23, 2019.
- [5] S. Ishaq *et al.*, Enhancement of dielectric and ferroelectric properties in flexible polymer for energy storage applications, *Ceramic International.* Vol. 46, no. 15, pp. 24649-24660, 2020.
- [6] A. Debnath *et al.*, Effect of non-mesogenic chiral terphenylate on the formulation of room temperature ferroelectric liquid crystal mixtures suitable for display applications, *J. Mol. Liq.* Vol. 292, pp. 111317, 2019.
- [7] Ashok K. Batra *et al.*, Growth and characterization of doped DTGS crystals for infrared sensing devices, *Material letters.* Vol. 57, pp. 3943-3948, 2003.
- [8] S. Salaeh *et al.*, Highly enhanced electrical and mechanical properties of methyl methacrylate modified natural rubber filled with multiwalled carbon nanotubes, *Polym. Test.* Vol. 85, pp. 106417, 2020.
- [9] N. George *et al.*, Nanosilica decorated multiwalled carbon nanotubes (CS hybrids) in natural rubber latex, *Polymer.* Vol. 161, pp. 170-180, 2019.
- [10] A. Cacciotti *et al.*, Design and development of advanced BaTiO₃/MWCNTs/ PVDF multi-layered systems for microwave applications, *Compos. Struct.* Vol. 224, pp. 111075, 2019.
- [11] A. Pal *et al.*, Enhancement in energy storage and piezoelectric performance of three phase (PZT/MWCNT/PVDF) composite, *Mater. Chem. Phys.* Vol. 244, pp. 122639, 2020.
- [12] M. Trainer, Ferroelectrics and the Curie-Weiss law, *Eur. J. Phys.* Vol. 21, pp. 459-464, 2000.
- [13] Bing-Ang Mei *et al.*, Physical Interpretations of Nyquist Plots for EDLC Electrodes and Devices, *J. Phys. Chem. C.* Vol. 122 (1), pp. 194-206, 2018.

YSC4F.218

ỨNG DỤNG IOT TRÊN BOARD ESP8266 VÀ GIAO THỨC MQTT CHO MÔ HÌNH QUẢN LÝ, GIÁM SÁT VĂN PHÒNG THÔNG MINH

PHẠM HỒNG THÁI, NGUYỄN VĂN DŨNG, NGUYỄN ANH VINH

Khoa công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

Tóm tắt. Trong các văn phòng ngày nay, ở các công ty, doanh nghiệp, vấn đề đưa một văn phòng làm việc bình thường trở thành một văn phòng thông minh đang là việc vô cùng cần thiết và không thể thiếu. Hiểu được nhu cầu đó, nhóm tác giả đã ứng dụng công nghệ điều khiển và giám sát thông minh vào văn phòng làm việc nhằm giúp nâng cao chất lượng môi trường làm việc. Trong phạm vi bài báo, nhóm tác giả sẽ xây dựng hệ thống giám sát của văn phòng cho phép thu thập tín hiệu từ các cảm biến, sử dụng vân tay để kiểm soát nhân viên ra vào, khi có nhân viên bấm vân tay, cửa sẽ được mở ra và cập nhật số người đang có trong văn phòng, tự động bật tắt đèn dựa theo điều kiện của ánh sáng hiện tại, chuông báo động sẽ kêu khi trong văn phòng phát hiện có khói, điều khiển các thiết bị trong phòng làm việc như máy lạnh, đèn, cửa cuốn, thông qua máy tính và điện thoại di động kết nối wifi. Nhằm đáp ứng mục tiêu của nghiên cứu đã đề ra, nhóm tác giả sử dụng các công nghệ hiện đại như các giao thức theo thời gian thực, vạn vật kết nối, xây dựng mô hình văn phòng làm việc thông minh với chi phí hợp lý và dễ dàng lắp đặt cho văn phòng trong thực tế.

Từ khóa. Văn phòng thông minh, Internet of Things, điều khiển tự động, cảm biến vân tay.

APPLICATION IOT ON BOARD ESP8266 AND MQTT PROTOCOL FOR SMART OFFICE MODEL MANAGEMENT AND SUPERVISION

Abstract. In today's offices, in corporate companies, the problem of turning an ordinary office into a smart office is extremely necessary and indispensable. Understanding that need, the authors have applied intelligent control and monitoring technology to the office to help improve the quality of the working environment and increase labor productivity. Within the scope of the article, the authors will build an office monitoring system that uses fingerprints to control employees entering and exiting, when an employee presses their fingerprints in, the door will be opened and the number of people updated. are in the office, automatically turn lights on and off based on current lighting conditions, an alarm will sound when smoke is detected in the office, control equipment in the office such as air conditioners, lights, rolling doors, through computers and mobile phones connected to wifi. In order to meet the objectives of the proposed research, the authors use modern technologies such as real-time protocols, connected things, building a working office model. Smart work at a reasonable cost and easy to install.

Keywords. Smart Office, Internet of Things, Automatic Control, Fingerprint.

1. GIỚI THIỆU

Hiện nay, công nghệ IoT đã trở nên phổ biến và được áp dụng vào rất nhiều lĩnh vực như nông nghiệp, công nghiệp, công nghệ... Một số ứng dụng của IoT trong cuộc sống hiện tại có thể kể đến như: quản lý đô thị thông minh, hệ thống băng điều khiển thông minh, bóng đèn thông minh, đồ gia dụng thông minh, nhà kho thông minh...

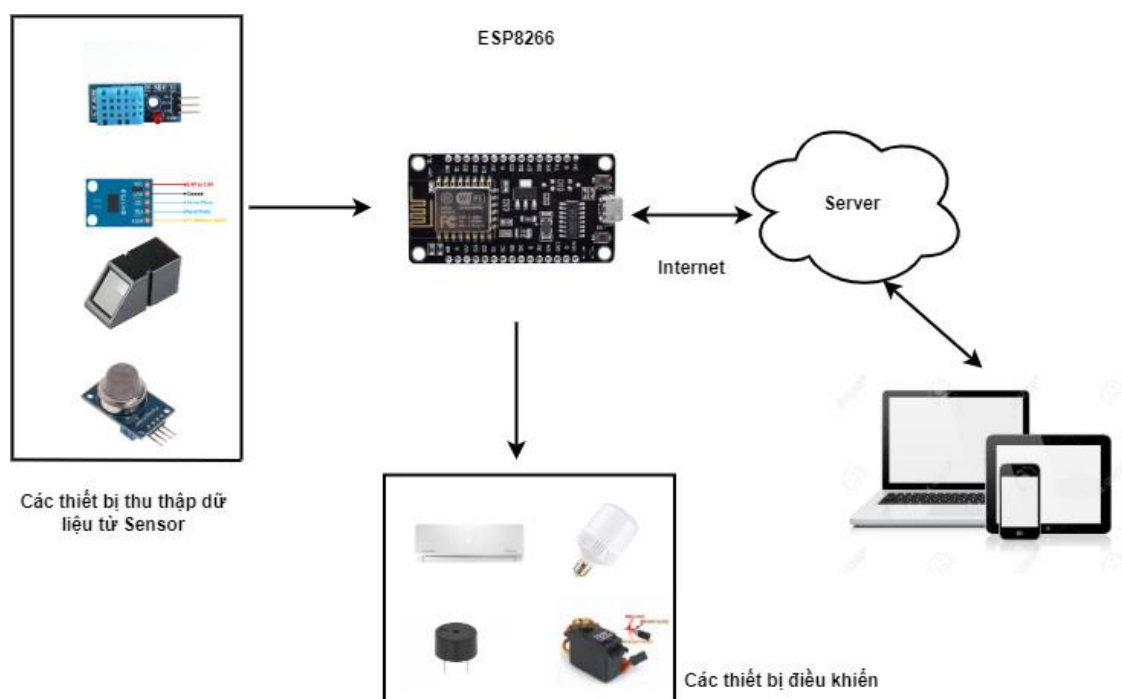
Trong tương lai, với sự phát triển của mạng 5G hứa hẹn sẽ là thời kì bùng nổ của IoT trên toàn cầu, ước tính trong vòng mười năm từ 2020 đến 2030, các thiết bị IoT sẽ tăng từ 75 tỷ lên 100 tỷ. Công nghệ phát triển rất nhanh và được ứng dụng rất nhiều trong cuộc sống, các văn phòng cũng ngày càng xuất hiện nhiều

cho nên việc ứng dụng công nghệ vào văn phòng làm việc để biến một văn phòng làm việc bình thường thành một văn phòng thông minh là việc vô cùng cần thiết và gần như không thể thiếu của các công ty, doanh nghiệp. Công việc văn phòng đã trở thành một công việc phổ biến. Đặc biệt, văn phòng ngày nay không chỉ là một nơi để làm việc mà nó còn chứa đựng giá trị, bộ mặt thương hiệu của doanh nghiệp. Với sự phát triển của Internet, Smartphone hay các thiết bị cảm biến, IoT đang trở thành xu thế mới của thế giới, các ứng dụng của IoT vào hệ thống thông minh đang được khẳng định một cách rõ rệt và được ứng dụng vô cùng rộng. Từ đó các mô hình làm việc tại văn phòng làm việc ngày càng được nâng cấp và trang bị các chức năng hiện đại để nâng cao chất lượng làm việc và mang lại năng suất lao động tốt nhất.

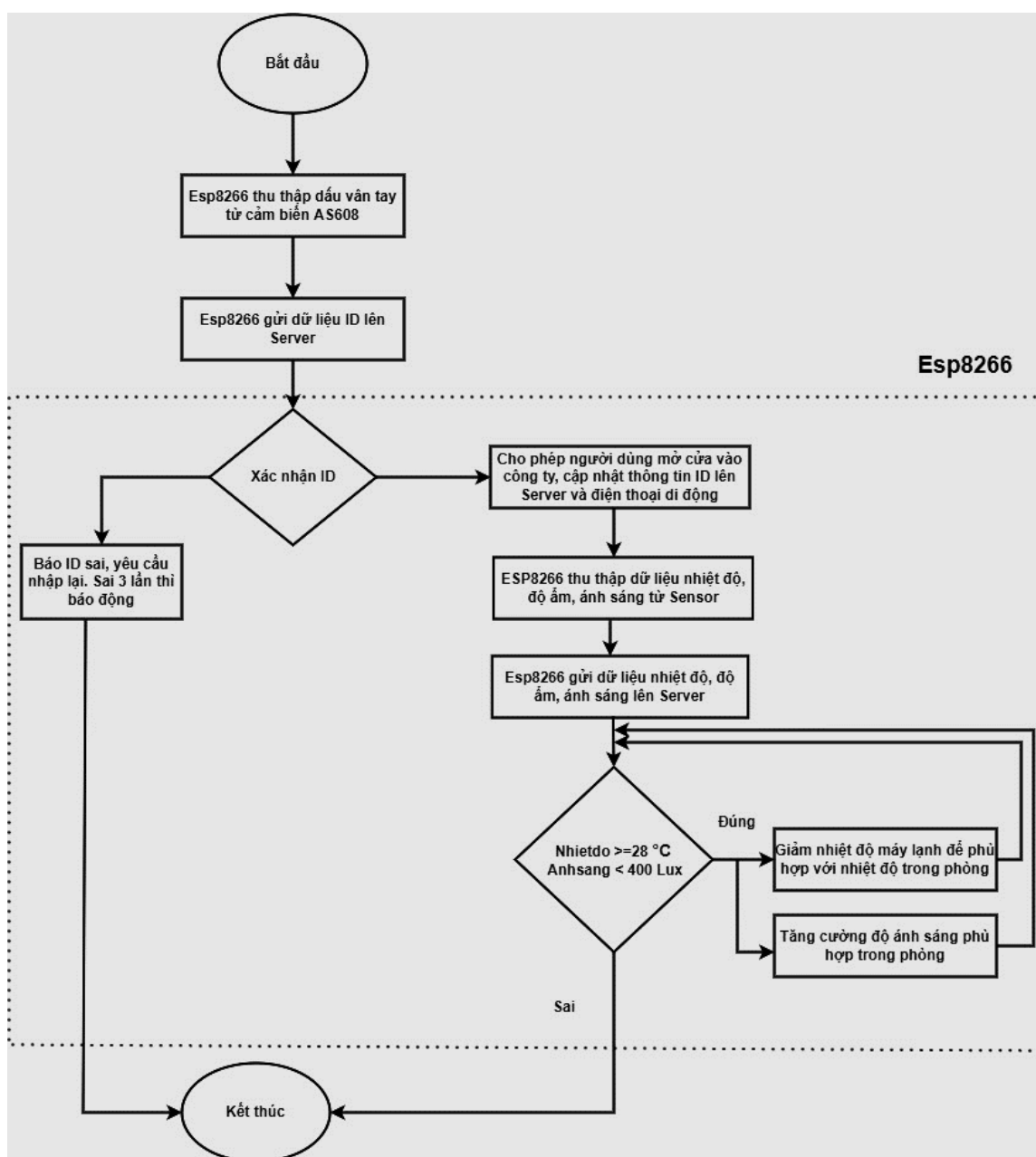
Khi công nghệ và đời sống con người ngày càng phát triển, các ứng dụng IoT ngày càng được quan tâm và nghiên cứu, việc xây dựng mô hình văn phòng làm việc thông minh dựa trên ứng dụng của IoT mang lại rất nhiều lợi ích cho doanh nghiệp, giúp cho doanh nghiệp vận hành một cách trơn tru, hiện đại qua đó thúc đẩy doanh số để doanh nghiệp phát triển. Gần đây, nhóm tác giả của bài báo [8][9] đã ứng dụng IoT vào trong đời sống, xây dựng các hệ thống giám sát, điều khiển dựa trên nền tảng IoT với các giao thức kết nối khác nhau cùng sự phát triển mạnh mẽ của công nghệ, các thiết bị thông minh giúp hỗ trợ phát triển công việc. Trong bài báo này, chúng tôi xây dựng một hệ thống văn phòng làm việc thông minh với Board ESP8266 và giao thức MQTT, hệ thống có khả năng điều khiển các thiết bị trong văn phòng như đèn, máy lạnh, cửa cuốn, chuông báo... giúp người dùng có thể dễ dàng kiểm soát và nâng cao chất lượng làm việc. Ngoài ra bài báo còn giúp cho sinh viên có thêm kiến thức về IoT, quá trình trao đổi thông tin và các giao thức mạng như giao thức mạng wifi IoT publish/subscribe MQTT (Message Queuing Telemetry Transport).

2. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

Hệ thống sử dụng vi điều khiển ESP8266 Node MCU để làm trung tâm kết nối giữa các cảm biến và phần mềm, thiết bị điều khiển. Vi điều khiển ESP8266 sẽ được kết nối tới 1 Server trung tâm, gọi là MQTT Broker Server, vi điều khiển ESP8266 sẽ thu thập dữ liệu điều khiển từ các cảm biến và tiến hành đưa lên Server bằng giao thức MQTT Publishing theo các topic đã đăng ký từ trước, đồng thời, các thiết bị, phần mềm điều khiển cũng có thể gửi các tín hiệu về ESP8266 theo giao thức MQTT Subscribe để tiến hành điều khiển các thiết bị cảm biến.



Hình 2.1: Sơ đồ tổng quát của hệ thống



Hình 2.2: Lưu đồ giải thuật của hệ thống

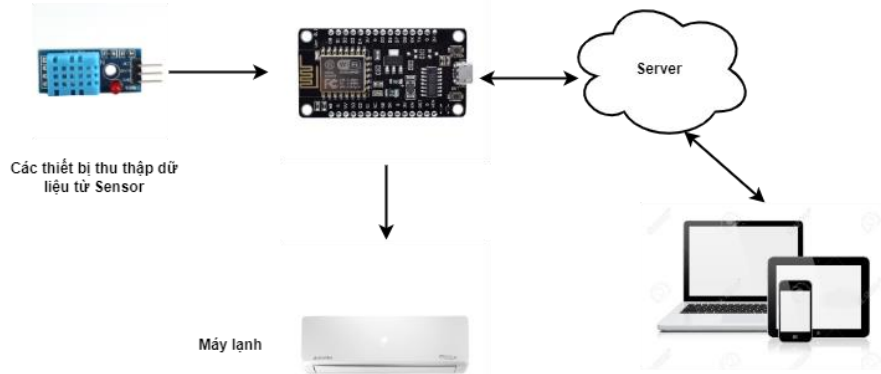
3. THIẾT KẾ PHẦN CỨNG

Danh sách các phần cứng sử dụng và chức năng của từng phần.

Thiết bị phần cứng	Chức năng
ESP8266 NodeMCU	Kết nối, điều khiển các thiết bị cảm biến thông qua mạng WiFi
Động cơ Motor DC	Điều khiển để hiện trạng thái thay cho máy lạnh
DHT11	Thu thập dữ liệu nhiệt độ và độ ẩm
BH1750	Thu thập dữ liệu ánh sáng
RC Servo MG996	Điều khiển đóng mở cửa cuốn
OLED SH1106G	Màn hình hiển thị thông tin nhiệt độ, độ ẩm, ánh sáng và vân tay cửa ra vào

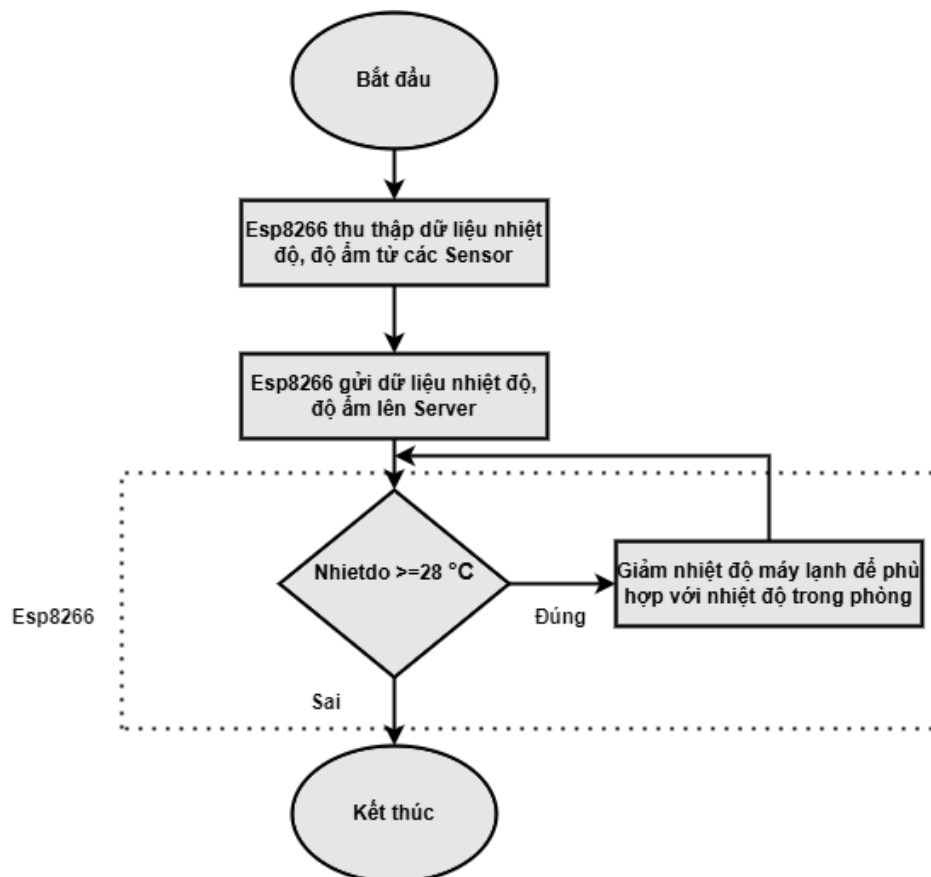
MQ-135	Phát hiện khói trong không gian phòng
AS608	Nhận dạng vân tay kiểm soát ra vào

Bảng 3.1: Danh sách các thiết bị phần cứng



Hình 3.1: Sơ đồ kết nối phần cứng điều khiển máy lạnh

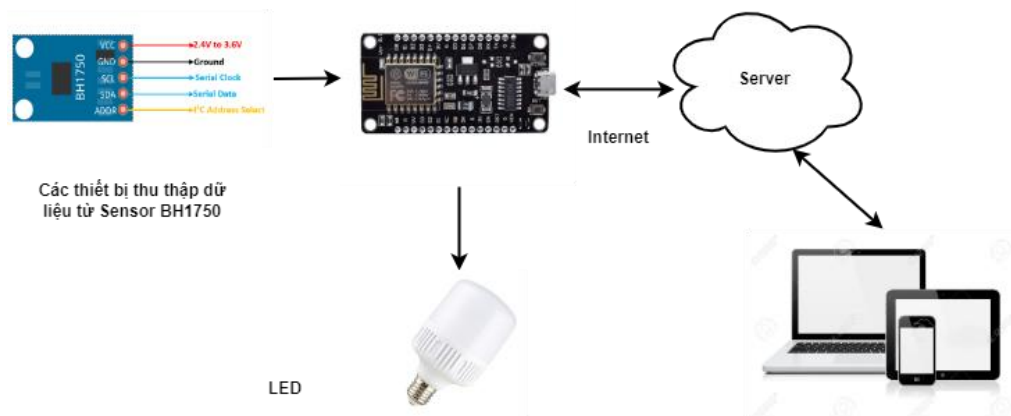
3.1. Kết nối phần cứng cho module điều khiển máy lạnh



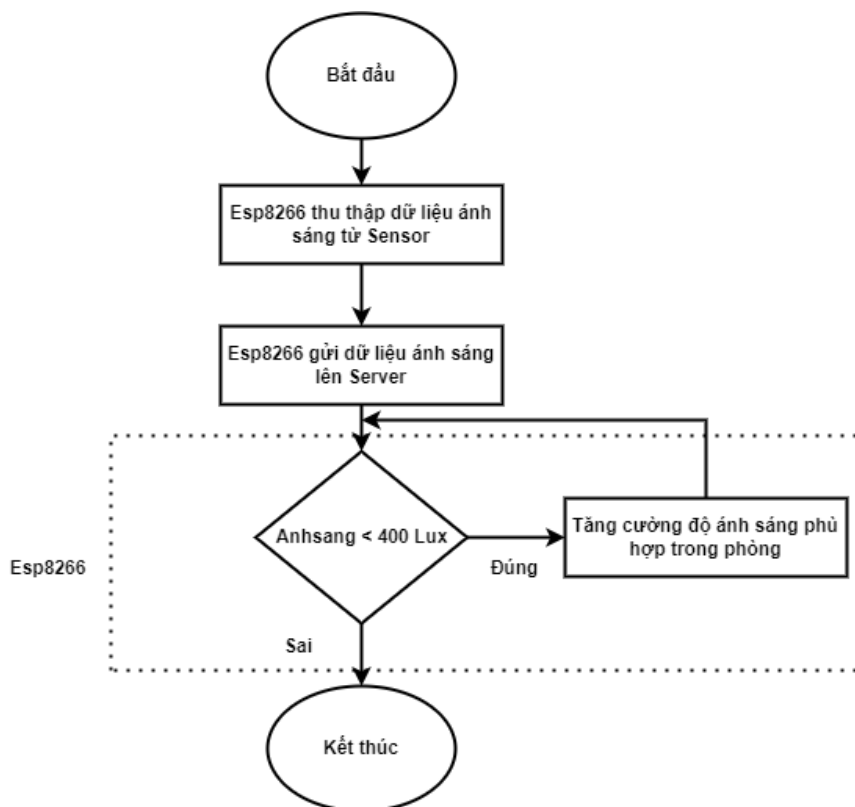
Hình 3.2: Lưu đồ giải thuật tự động kiểm soát nhiệt độ

Hình 3.2 trình bày lưu đồ giải thuật tự động kiểm soát nhiệt độ của hệ thống, nguyên lý hoạt động của chức năng giám sát nhiệt độ và độ ẩm trong phòng được cảm biến nhiệt độ DHT11 đo, truyền tín hiệu qua ESP8266 NodeMCU. Tại đây, thông qua kết nối WiFi truyền dữ liệu nhiệt độ, độ ẩm lên Server MQTT để

lưu trữ. Dữ liệu thông qua Server xử lý và hiển thị nhiệt độ, độ ẩm trên giao diện web, App Android. Dữ liệu được thể hiện bằng biểu đồ dạng đường hiển thị theo thời gian thực đo được đồng thời bật tắt máy lạnh để điều chỉnh nhiệt độ trong phòng cho phù hợp.



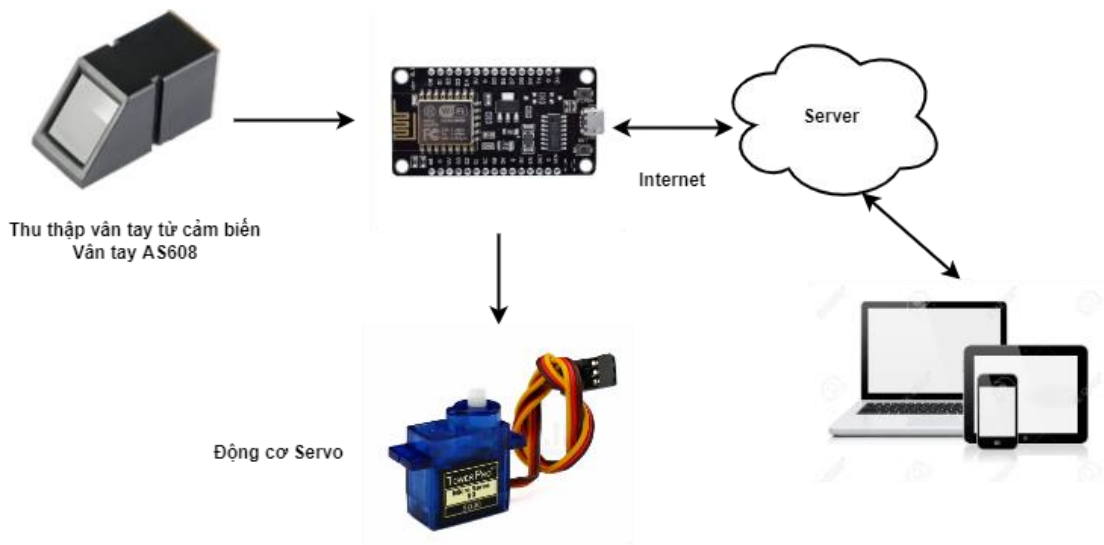
Hình 3.3: Sơ đồ điều khiển đèn dựa trên cảm biến ánh sáng



Hình 3.4: Lưu đồ giải thuật điều khiển đèn dựa trên cường độ ánh sáng thu thập được

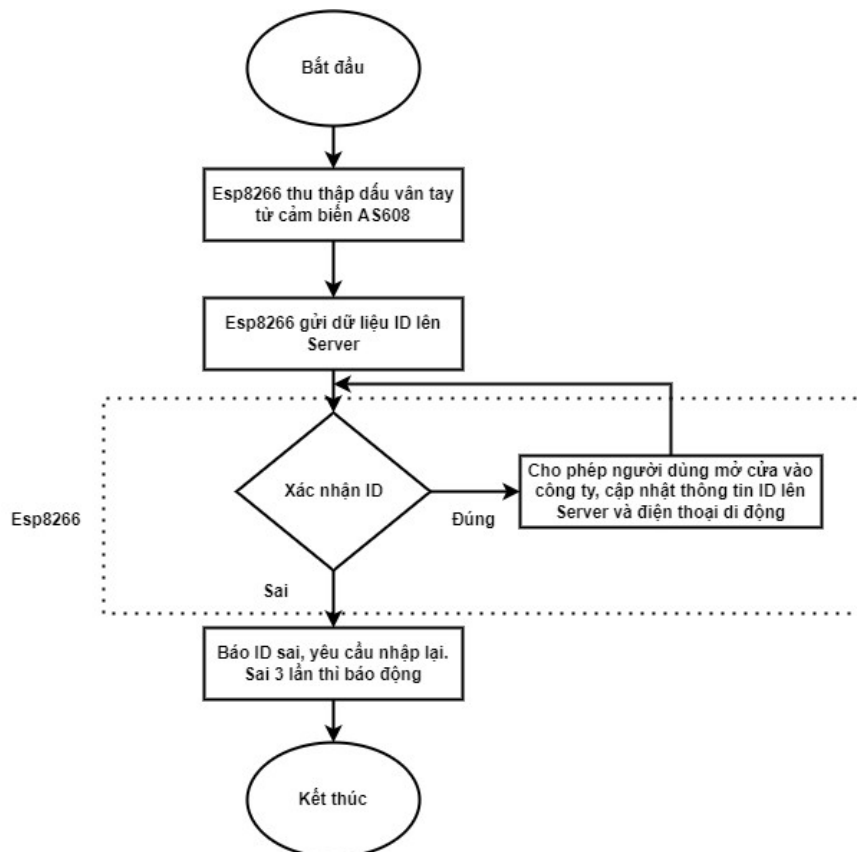
Hình 3.3 là sơ đồ kết nối giữa cảm biến BH1750, LED và ESP8266. Xây dựng chức năng giám sát đồng thời cung cấp dữ liệu ánh sáng trong phòng và tự động điều khiển bật tắt đèn.

Hình 3.4 thể hiện lưu đồ giải thuật điều khiển đèn dựa trên cường độ ánh sáng thu thập được Nguyên lý hoạt động của chức năng giám sát nhiệt độ và độ ẩm trong phòng được cảm biến nhiệt độ DHT11 đo, truyền tín hiệu qua ESP8266 NodeMCU. Tại đây, thông qua kết nối WiFi truyền dữ liệu nhiệt độ, độ ẩm lên Server MQTT để lưu trữ. Dữ liệu thông qua Server xử lý và hiển thị dữ liệu ánh sáng trên giao diện web, App Android. Dữ liệu được thể hiện bằng biểu đồ dạng đường hiển thị theo thời gian thực đo được đồng thời bật tắt đèn lạnh để điều chỉnh cường độ ánh sáng trong phòng cho phù hợp.



Hình 3.5: Sơ đồ hệ thống kiểm soát ra vào

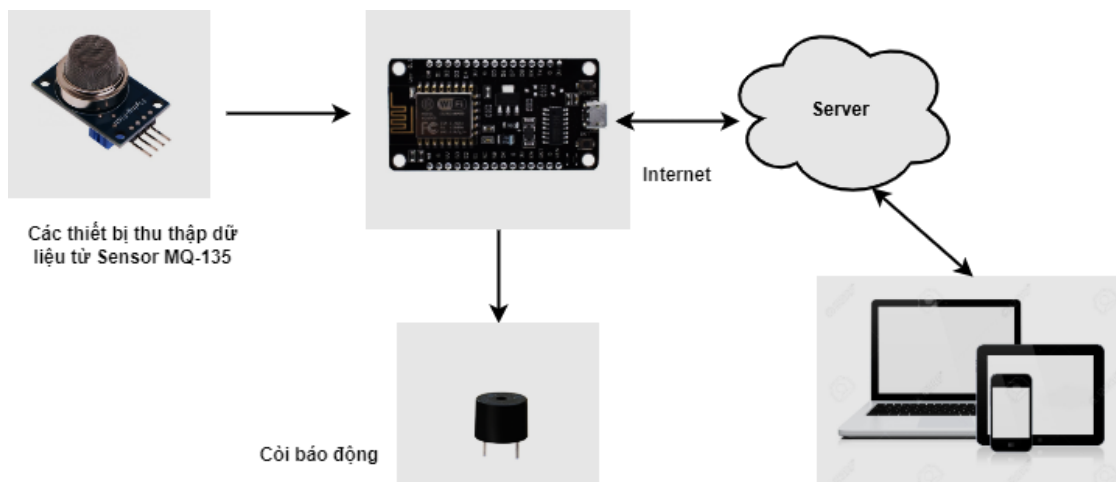
Hình 3.5 là sơ đồ hệ thống kiểm soát ra vào của văn phòng, các thiết bị cảm biến và động cơ sẽ được nối với ESP8266. Server sẽ là nơi lưu trữ số lượng và người ra vào trong văn phòng, các thông tin sẽ được cập nhật liên tục theo thời gian thực



Hình 3.6: Lưu đồ giải thuật hệ thống kiểm soát ra vào

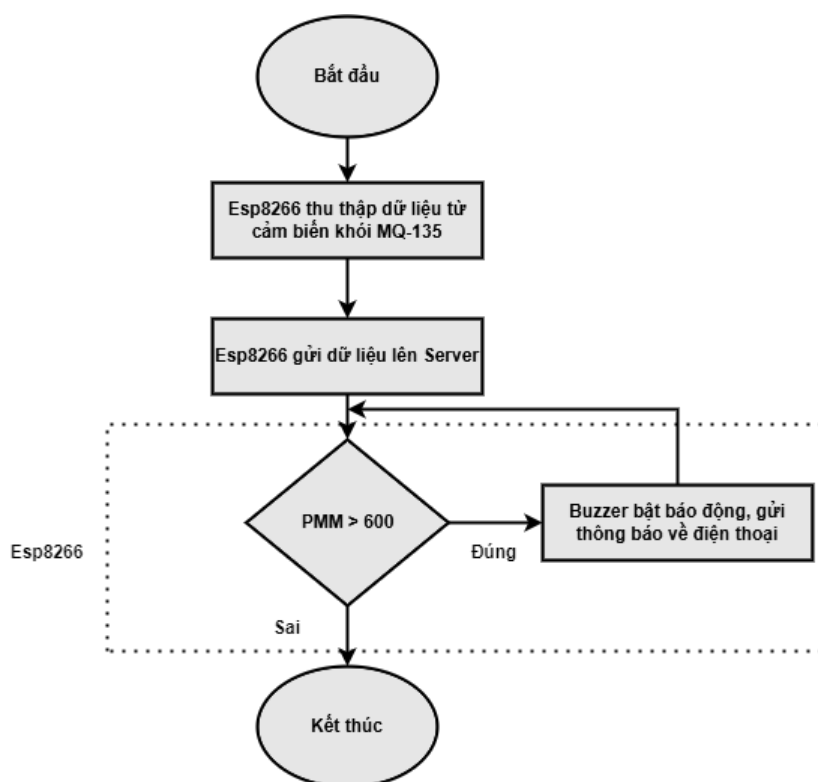
Hình 3.6 mô tả cách hệ thống kiểm soát ra vào hoạt động, nguyên lý hoạt động của chức năng thu thập dấu vân tay được cảm biến vân tay AS608 truyền tín hiệu qua ESP8266 NodeMCU. Tại đây, thông qua kết nối WiFi truyền dữ liệu dấu vân tay (ID) lên Server MQTT để lưu trữ. Dữ liệu thông qua Server xử lý nếu xác thực dấu vân tay trùng khớp thì hiển thị dữ liệu thông tin trên App Android đồng thời mở cửa (Servo). Mỗi

vân tay sẽ được gán với 1 ID, khi người dùng thực hiện mở cửa bằng vân tay, cảm biến sẽ gửi tín hiệu về và so sánh với vân tay đã đăng ký trên hệ thống, nếu đúng vân tay, cửa sẽ được mở và cho phép người dùng vào văn phòng, ngược lại, nếu vân tay không chính xác, màn hình sẽ hiện thông báo và yêu cầu người dùng giữ lại vân tay của mình, sau 3 lần không chính xác, hệ thống sẽ phát hiện chuông báo và cảnh báo có người xâm nhập trái phép.



Hình 3.7: Sơ đồ báo động cháy trong văn phòng

Hình 3.7 là sơ đồ kết nối giữa cảm biến MQ-135, Buzzer và ESP8266. Xây dựng chức năng giám sát đồng thời cung cấp dữ liệu không khí trong phòng và tự động bật còi báo động Buzzer. Chất lượng không khí sẽ được cảm biến MQ-135 thu thập và cập nhật liên tục lên server và hệ thống, khi phát hiện khói, cảm biến sẽ gửi tín hiệu về vi điều khiển và ngay lập tức chuông báo sẽ được kích hoạt để báo động cho mọi người trong văn phòng



Hình 3.8: Lưu đồ giải thuật hệ thống báo cháy trong văn phòng

Nguyên lý hoạt động của hệ thống phát hiện khói, cảm biến Sensor MQ-135 được kết nối với ESP8266. Tại đây, thông qua kết nối WiFi truyền dữ liệu nồng độ khí Gas lên Server MQTT để lưu trữ. Khi nồng độ khí Gas vượt qua mức cho phép thì nó sẽ tự động bật còi cảnh báo và gửi thông báo đến App Android.

4. THIẾT KẾ PHẦN MỀM

Để thu thập dữ liệu và điều khiển hệ thống, nhóm tác giả đã tiến hành viết chương trình trên App Android và trang web sử dụng Node-RED, sử dụng giao thức MQTT kết hợp cùng Server HiveMQ để tiến hành publish/subscribe điều khiển các thiết bị, ngoài ra nhóm tác giả còn sử dụng Firebase để lưu trữ dữ liệu vân tay và thu thập dữ liệu trong App Android.

a. App Android

App bao gồm giao diện chính khi khởi động gồm các phòng ban, hệ thống kiểm soát ra vào và bảng biểu đồ thu thập dữ liệu từ các cảm biến như nhiệt độ, độ ẩm, ánh sáng.

Trong một phòng sẽ bao gồm thông tin về nhiệt độ, độ ẩm, ánh sáng thông qua việc publish các tín hiệu theo các topic cùng với các nút nhấn để điều khiển các thiết bị như ánh sáng, đèn, điều hòa thông qua việc subscribe các topic và hiển thị số người trong văn phòng thông qua server Firebase.

Giao diện Attendance sẽ hiển thị thông tin của người đã mở cửa ra vào hệ thống thông qua vân tay, các vân tay sẽ được đăng ký ID và khi người dùng truy cập bằng vân tay, ID sẽ được nhận dạng và cập nhật trên hệ thống, đồng thời cửa mở ra và cho phép người dùng vào văn phòng.

b. Node-RED

Node-RED là một công cụ lập trình để kết nối các thiết bị phần cứng, API và dịch vụ trực tuyến, cung cấp một trình chỉnh sửa dựa trên trình duyệt giúp dễ dàng kết nối các luồng với nhau bằng cách sử dụng các nút trong bảng điều khiển được cung cấp sẵn.

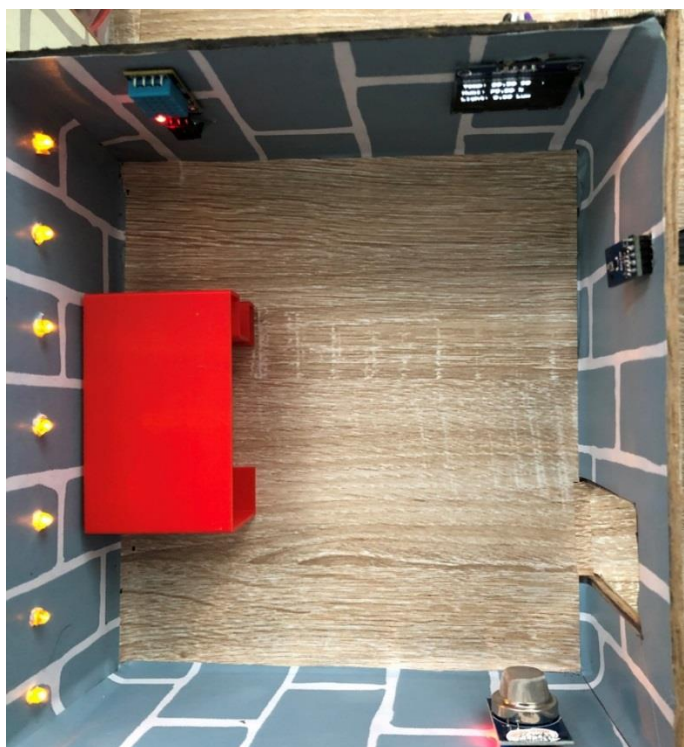
5. KẾT QUẢ THỰC HIỆN

c. Mô hình văn phòng làm việc

Thông qua mô hình ta có thể theo dõi được các chỉ số như nhiệt độ, ánh sáng, độ ẩm qua màn hình OLED, điều khiển được điều hòa, bật tắt đèn led thông qua ứng dụng được viết trên hệ điều hành Android và giao diện điều khiển trên Web, ngoài ra hệ thống còn thu thập được khói trong văn phòng để cảnh báo nếu có sự cố cháy xảy ra, hệ thống sử dụng vân tay sinh trắc học để mở cửa cuốn tại cửa chính của văn phòng.



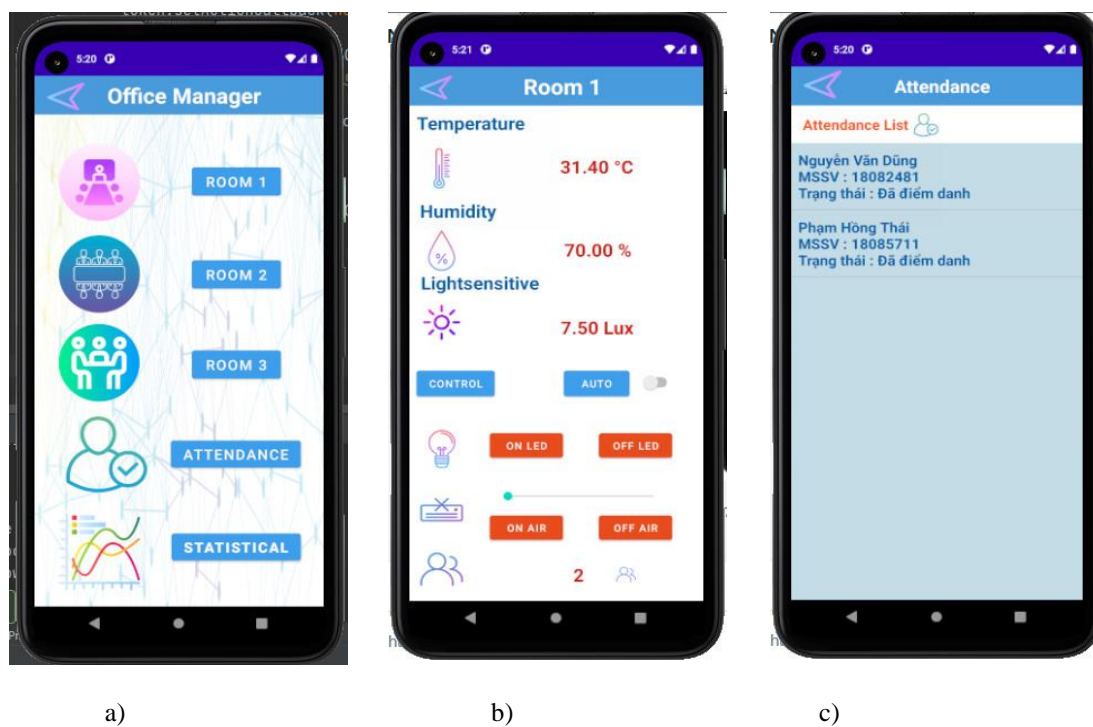
Hình 5.1: Hình ảnh mô hình



Hình 5.2: Hình minh họa một phòng trong hệ thống

d. Chương trình quản lý

Giao diện quản lý trên App Android sẽ bao gồm 3 phần chính là phần trang chủ, giao diện từng phòng ban và giao diện điểm danh.

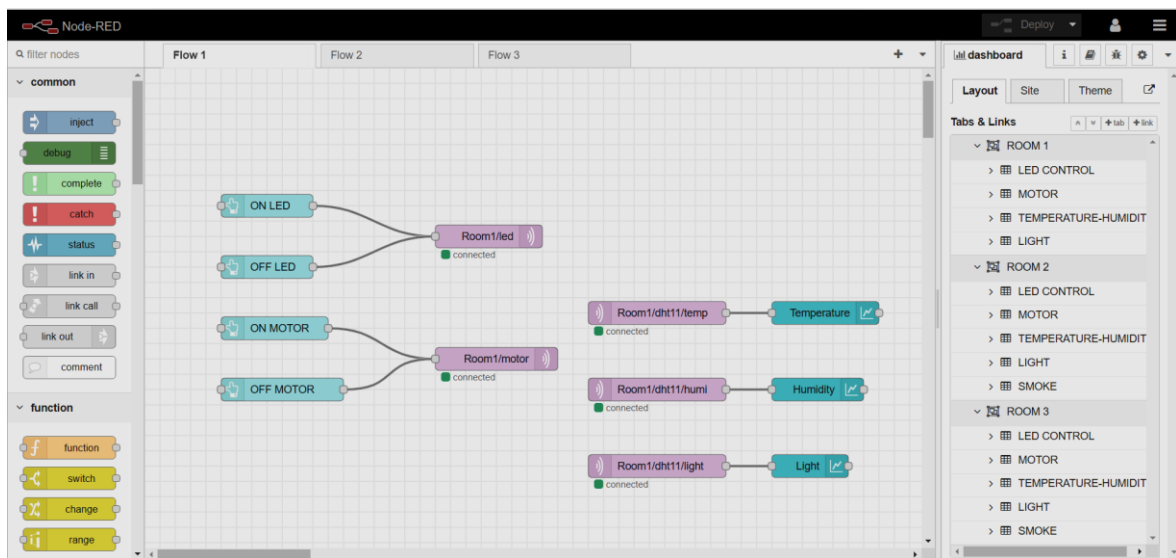


Hình 5.3: Giao diện hệ thống điều khiển trên điện thoại

Hình 5.3: a) là giao diện chính của App khi khởi động bao gồm các phòng ban, hệ thống kiểm soát ra vào và bảng biểu đồ thu thập dữ liệu từ các cảm biến như nhiệt độ, độ ẩm, ánh sáng; b) là giao diện trong một phòng sẽ bao gồm thông tin về nhiệt độ, độ ẩm, ánh sáng thông qua việc publish các tín hiệu theo các topic cùng với các nút nhấn để điều khiển các thiết bị như ánh sáng, đèn, điều hòa thông qua việc subscribe các topic và hiển thị số người trong văn phòng thông qua server Firebase; c) là giao diện Attendance hiển thị thông tin của người đã mở cửa ra vào hệ thống thông qua vân tay, các vân tay sẽ được đăng ký ID và khi người dùng truy cập bằng vân tay, ID sẽ được nhận dạng và cập nhật trên hệ thống, đồng thời cửa mở ra và cho phép người dùng vào văn phòng.

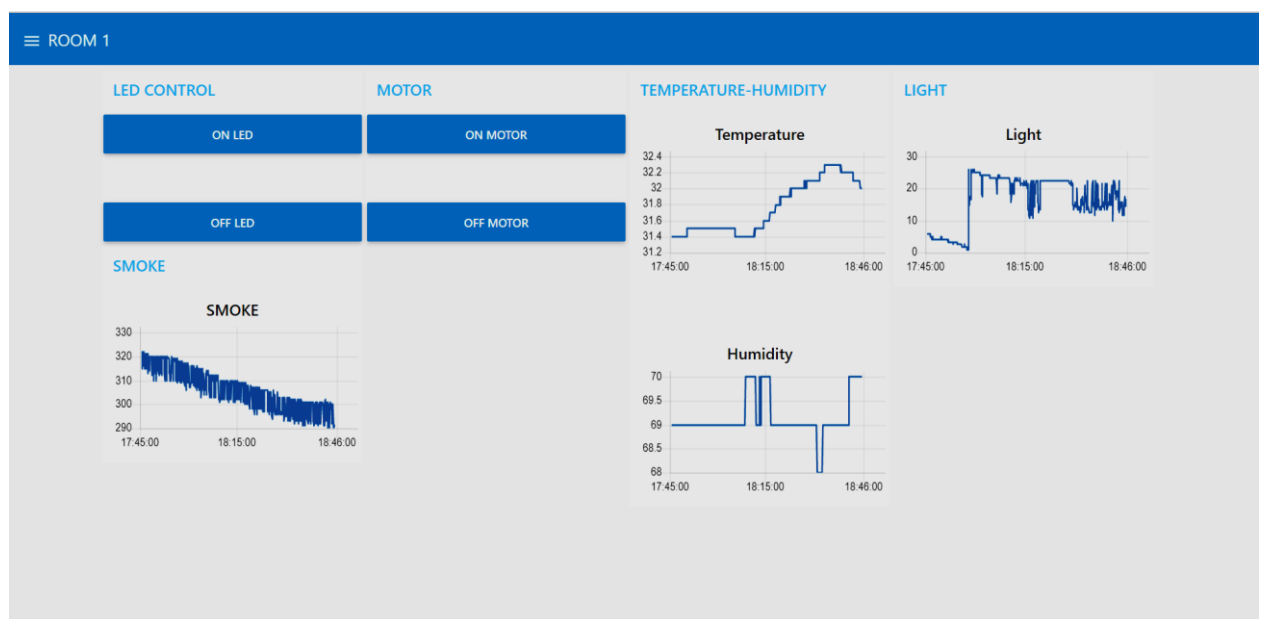
Ngoài ra chúng tôi còn xây dựng giao diện thu thập dữ liệu và điều khiển trên trang web thông qua Node-RED.

Để khởi động Node-RED, ta truy cập vào localhost để tiến hành kết nối các thiết bị với server broker và tiến hành truyền nhận dữ liệu theo publish/subscribe.



Hình 5.4: Giao diện Node-RED

Để điều khiển, ta truy cập vào địa chỉ localhost/admin



Hình 5.5: Giao diện điều khiển và thu thập dữ liệu

6. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Từ việc nghiên cứu lý thuyết và tiến hành xây dựng mô hình Smart Office để giám sát các chỉ số về nhiệt độ, độ ẩm, ánh sáng và đồng thời điều khiển các thiết bị trong văn phòng làm việc như hệ thống đèn chiếu sáng, điều hòa, chuông báo động, cửa cuốn tích hợp nhận dạng vân tay. Hệ thống có thể giám sát các chỉ số, điều khiển các thiết bị trong văn phòng, điều khiển cửa cuốn tự động thông qua việc nhận dạng vân tay. Hướng đến xây dựng một ứng dụng hoàn chỉnh hơn trên hệ điều hành Android, phát triển nhiều tính năng cho hệ thống để biến hệ thống thành hệ thống tự động hoàn toàn, có thể tự điều chỉnh, bật tắt các thiết bị dựa vào dữ liệu thu thập được. Thêm vào đó là xây dựng hệ thống bảo mật cho người sử dụng để tránh trường hợp bị mất dữ liệu cá nhân. Mô hình có thể nâng cấp để ứng dụng cho các công ty có số lượng nhân viên lớn, hệ thống có thể tích hợp mở rộng lớn, cập nhật liên tục theo thời gian thực.

TÀI LIỆU THAM KHẢO

- [1] B. INTELLIGENCE, "The Internet of Everything," BI INTELLIGENCE, 2015.
- [2] P. M. T. (TuanPM), Internet Of Thing Cho Người Mới Bắt Đầu, IoT Maker VietNam, 2017.
- [3] Dipa Soni , Ashwin Makwana, "A Survey On MQTT : A Protocol Of Internet Of Things (IoT)," Charusat University , Gujarat , India, 2017.
- [4] I. Inc, "All Datasheet," InterBird Inc, [Online]. Available: <https://www.alldatasheet.com/>.
- [5] J. Cardle, Android App Development in Android Studio, Manchester Academic.
- [6] Paul Blundell , Diego Torres Milano, Learning Android Application Testing, Birmingham - Mumbai: Packt, 2015.
- [7] Khalid Alghamdi , Ali Alqazzaz , Anyi Liu , Hua Ming, IoTVerif : An Automated Tool to Verify SSL/TLS Certificate Validation in Android MQTT Client Applications, USA: Oakland University, March 19-21 , 2018.
- [8] Trần Thị Minh Hoa, Nguyễn Cao Anh Minh, Nguyễn Thị Hậu, "ỨNG DỤNG IOT TRONG HỆ THỐNG QUẢN LÝ VÀ GIÁM SÁT NGÔI NHÀ THÔNG MINH THEO THỜI GIAN THỰC," Tạp chí Khoa học và Công nghệ, 2021.
- [9] Võ Minh Phụng, Dương Thị Thanh Hiền, Võ Tiến Phúc, "NGHIÊN CỨU XÂY DỰNG MÔ HÌNH ĐIỀU KHIỂN NHÀ THÔNG MINH SỬ DỤNG KẾT NỐI BLUETOOTH, GSM," Tạp chí Khoa học Đại học Đà Lạt, 2018.
- [10] P. M. Tuấn, Internet Of Things Cho Người Mới Bắt Đầu, IoT Maker, 2017.
- [11] A. Datasheet, "Electronic Components Datasheet Search," [Online]. Available: <https://www.alldatasheet.com/>.
- [12] J.Cardle, Android App Development for Android Studio, Manchester Academic.
- [13] Christophe Le Gal, Jerome Martin, Augustin Lux, James L.Crowley , "Smart Office: Design of an Intelligent Environment".

YSC4F.219

ỨNG DỤNG GIẢI THUẬT HỌC SÂU VÀ LOGIC MỜ XÂY DỰNG HỆ THỐNG GIÁM SÁT VÀ ĐIỀU KHIỂN CHO NGÔI NHÀ THÔNG MINH TRÊN BOARD RASPBERRY VÀ ARDUINO

ĐẶNG THÁI HIÊN, NGUYỄN BÁ HOÀN, NGUYỄN ANH VINH

*Khoa công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh,
nguyenanhvinh@iuh.edu.vn*

Tóm tắt. Trong bài báo này, chúng tôi nghiên cứu xây dựng mô hình ngôi nhà thông minh sử dụng giải thuật logic mờ và học sâu. Hệ thống sử dụng Arduino thu thập dữ liệu từ các cảm biến, Raspberry Pi thực hiện xử lý, tín hiệu ngõ ra điều khiển tự động thiết bị như đèn, rèm cửa, máy lạnh,... sao cho ngôi nhà luôn trong trạng thái lý tưởng và tiết kiệm điện năng. Ngoài ra chúng tôi còn phát triển chức năng nhận diện khuôn mặt dùng kỹ thuật học sâu để giám sát an ninh vào/ra cho ngôi nhà. Kết quả bài báo có thể là giải pháp tiết kiệm năng lượng, bảo mật cho ngôi nhà thông trong thực tế.

Từ khóa. ngôi nhà thông minh, logic mờ, học sâu, nhận diện khuôn mặt, Raspberry Pi, Arduino.

BUILDING A FUZZY LOGIC AND DEEP LEARNING BASED MONITORING AND CONTROL SYSTEM ON RASPBERRY AND ARDUINO BOARDS FOR SMART HOME

Abstract. In this paper, we investigate and build a smart house model using Fuzzy logic and deep learning. Arduino board collects signals from sensors while Raspberry pi board performs processing of the received signals. The output signals are used to control devices such as lights, curtains, air conditioners, ... The aim is to maintain an ideal and save energy state for the house. In addition, we also develop the facial recognition function based on deep learning for the security of the house entrance. The obtained results can be a power-saving and security solution for smart homes.

Keywords. Smart home, fuzzy logic, deep learning, face recognition, Raspberry pi, Arduino.

1. GIỚI THIỆU

Với sự phát triển không ngừng của khoa học hiện đại, con người đã ngày càng nâng cao đời sống của mình hơn và luôn mơ ước tới một cuộc sống hiện đại và tiện nghi nhất. Chính từ những nhu cầu đó, con người đã có rất nhiều sáng tạo phục vụ cho cuộc sống. Và ý tưởng “ngôi nhà thông minh” cũng xuất phát từ nhu cầu thực tiễn như vậy.

Nhà thông minh là ngôi nhà được lắp đặt các thiết bị điện, điện tử có thể được điều khiển bằng các phương pháp tự động hoặc bán tự động. Nhiều công trình nghiên cứu thiết kế và xây dựng ngôi nhà thông minh được thực hiện, chẳng hạn [1-4]. Đa số bài báo tập trung sử dụng board vi điều khiển có chức năng khai thác khả năng xử lý của hệ thống lưu trữ và/hoặc xử lý từ xa thông qua mạng Internet [1-3], GSM, Bluetooth [4]. Rõ ràng các giải pháp này luôn đối mặt với rủi ro về bảo mật thông tin. Khác với các bài báo [1-4], tài liệu [5] gợi ý cách điều khiển máy lạnh theo phương pháp logic mờ. Trong đó nhiệt độ, độ ẩm ngõ ra nằm trong khoảng nhất định.

Trong bài báo này, chúng tôi đề xuất việc xây dựng một hệ thống giám sát và điều khiển cho ngôi nhà thông minh sử dụng kết hợp board Raspberry pi và board Arduino. Việc thu thập tín hiệu cảm biến được thực hiện nhờ vào board Arduino. Trong khi board Raspberry pi đảm nhiệm việc xử lý tín hiệu bằng thuật toán logic mờ để tạo ra các tín hiệu điều khiển cho các thiết bị như máy lạnh, đèn, rèm cửa, chuông báo động, ... Các yếu tố nhiệt độ, độ ẩm, ánh sáng trong ngôi nhà được duy trì trong khoảng mát mẻ, có lợi cho sức khỏe

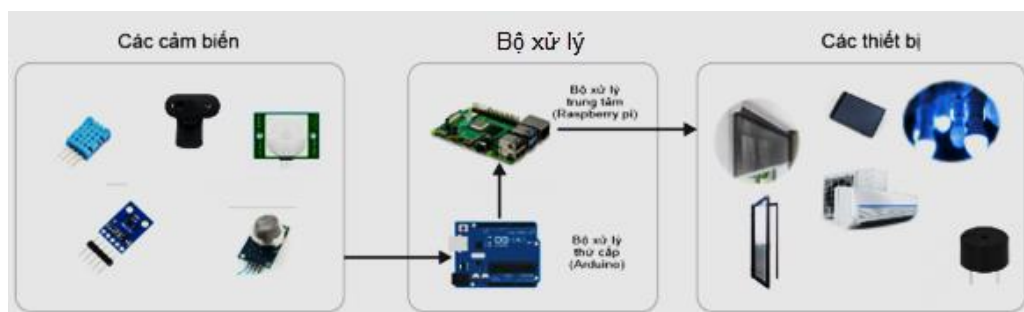
người dùng. Bên cạnh đó, chức năng nhận diện khuôn mặt dùng kỹ thuật học sâu để giám sát an ninh vào/ra cho ngôi nhà cũng được phát triển.

Nội dung còn lại của bài báo được bố trí như sau. Tiếp theo phần giới thiệu là phần xây dựng hệ thống. Các kết quả thực nghiệm được trình bày kế đó. Sau cùng, kết luận và một số hướng phát triển cho bài báo được đưa ra.

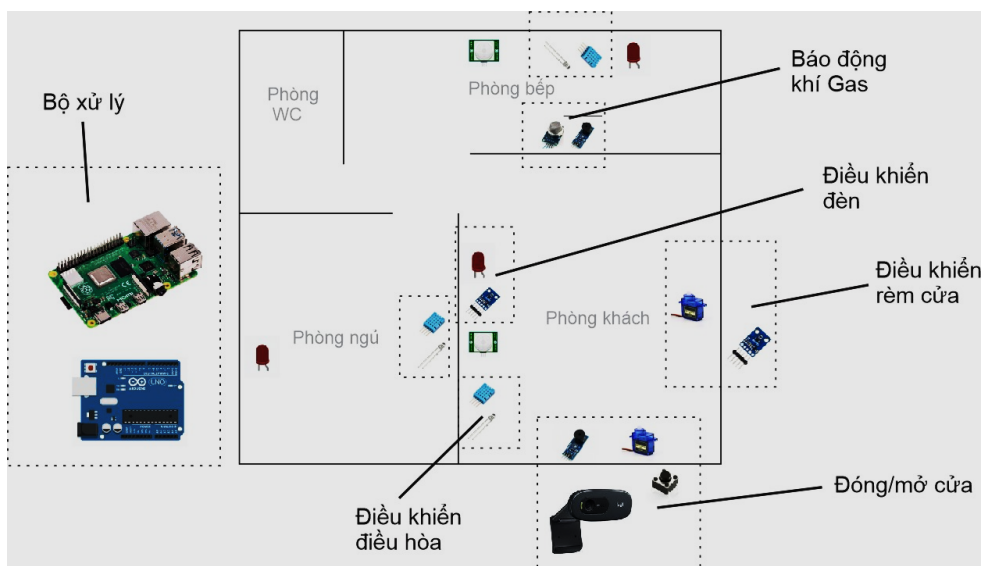
2. PHÂN TÍCH VÀ THIẾT KẾ HỆ THỐNG

2.1. Tổng quan hệ thống

Hệ thống giám sát và điều khiển ngôi nhà thông minh được minh họa ở Hình 2.1. Hệ thống gồm ba phần: các cảm biến, bộ xử lý và các thiết bị được kết nối như Hình 2.1.a. Trong đó, chúng tôi sử dụng board Raspberry Pi 4 và board Arduino Uno R3 làm bộ xử lý. Board Arduino đóng vai trò xử lý thứ cấp thực hiện nhận tín hiệu vào từ các cảm biến ánh sáng, nhiệt độ, độ ẩm và từ webcam chuyển sang dạng tín hiệu phù hợp để chuyển tới Board Raspberry Pi. Tới lượt mình, board Raspberry Pi thực hiện các tác vụ xử lý trung tâm các tín hiệu nhận được. Tín hiệu ngõ ra được dùng để điều khiển các thiết bị.



a. Sơ đồ kết nối hệ thống.



b. Sơ đồ bố trí phần cứng

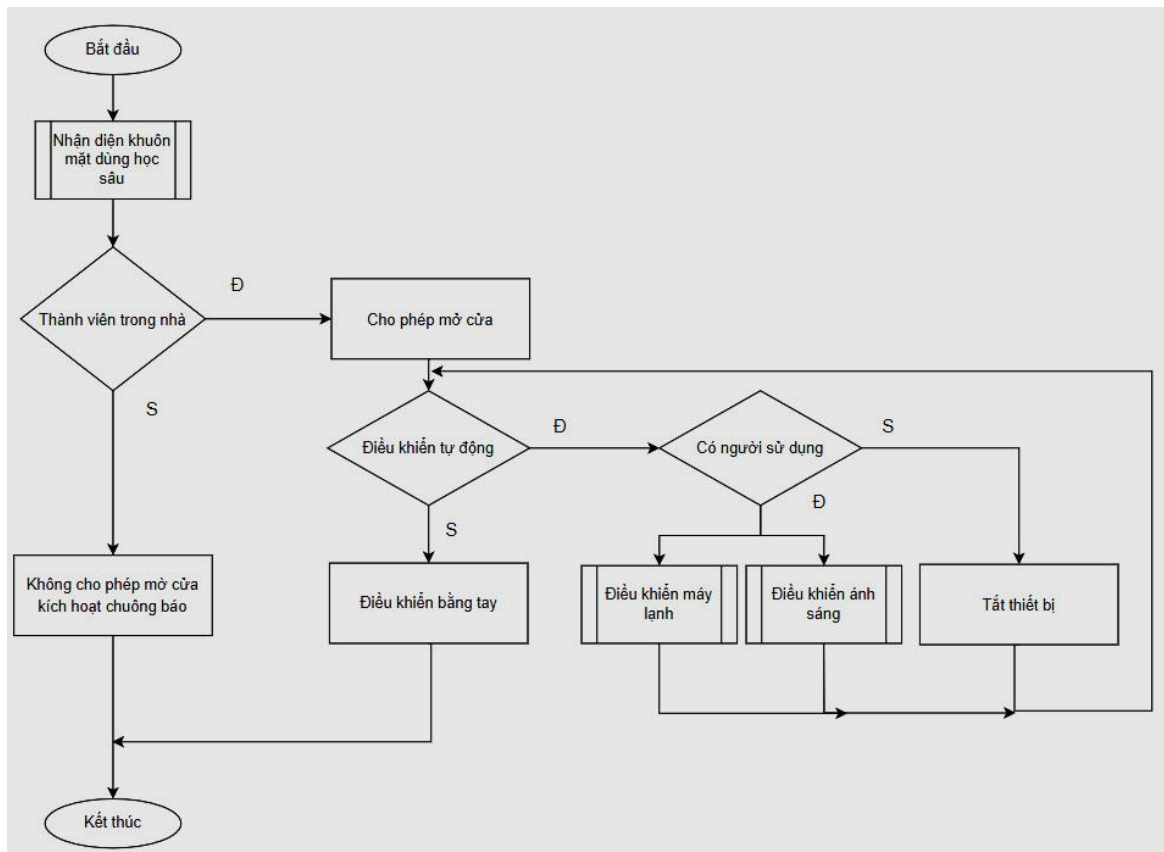
Hình 2.1. Hệ thống giám sát và điều khiển ngôi nhà thông minh

Chúng tôi xây dựng một mô hình ngôi nhà gồm phòng khách, phòng ngủ, phòng bếp và phòng WC và bố trí các thiết bị phần cứng vào các phòng như Hình 2.1.b. Bên dưới là các chức năng của hệ thống:

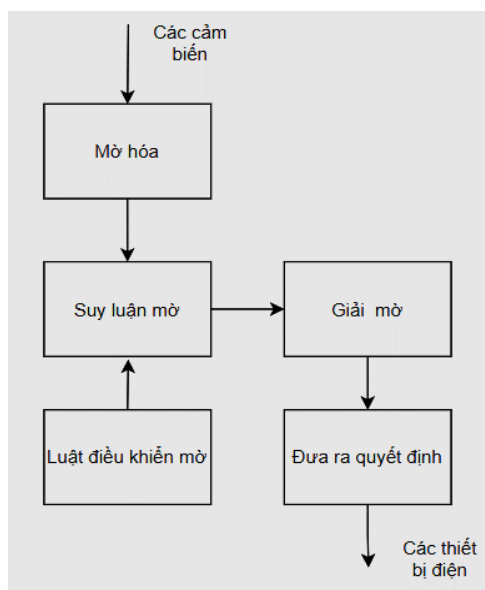
- Giám sát vào/ra: webcam thu nhận hình ảnh hiển thị trên LCD và gửi về Raspberry Pi 4 để nhận dạng khuôn mặt. Button có vai trò kích hoạt sự kiện. Servo đóng mở cửa dưới sự điều khiển của Raspberry Pi 4. Buzzer có tác dụng như chuông báo.
- Điều khiển điều hòa: Cảm biến DHT11 đọc giá trị nhiệt độ, độ ẩm từ môi trường gửi về cho Arduino

- Uno R3. LED phát hồng ngoại sẽ điều khiển máy lạnh như 1 cái remote thông qua Raspberry Pi 4.
- Điều khiển rèm cửa: Cảm biến BH1750 đo đặc giá trị ánh sáng ở bên ngoài gửi về Arduino Uno R3. Động cơ servo điều khiển rèm.
- Điều khiển đèn: Cảm biến BH1750 đo đặc giá trị ánh sáng trong phòng gửi về Arduino Uno R3. Đèn LED thấp sáng phòng.
- Báo động khí Gas: Cảm biến khí gas MQ2 đọc giá trị từ môi trường rồi gửi về cho Arduino Uno R3. Buzzer báo động khi khí Gas quá cao.
- Cảm biến PIR: Phát hiện người dùng để bật/tắt thiết bị khi không có người sử dụng.

Hình 2.2 cho thấy cách thức hệ thống hoạt động. Đầu tiên người dùng nhận diện khuôn mặt dùng học sâu, nếu là thành viên trong gia đình, hệ thống sẽ cho phép mở cửa để người sử dụng có thể tiến vào trong nhà. Còn nếu không phải là thành viên trong nhà, thì hệ thống sẽ kích hoạt chuông báo. Kế tiếp, người sử dụng có thể lựa chọn điều khiển các thiết bị trong nhà bằng tay hoặc lựa chọn hệ thống điều khiển tự động. Đối với điều khiển tự động, cảm biến PIR được dùng để phát hiện có người sử dụng hay không. Nếu không có người sử dụng hệ thống sẽ tắt các thiết bị để tiết kiệm và tránh gây lãng phí năng lượng. Nếu có người sử dụng hệ thống sẽ điều khiển các thiết bị điện theo giải thuật mờ như Hình 2.3 [5].



Hình 2.2: Cách hoạt động của hệ thống

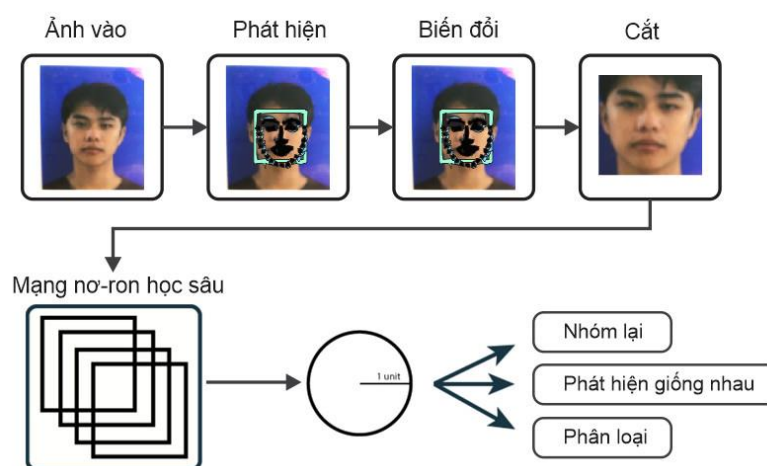


Hình 2.3: Giải thuật logic mờ [5]

Thuật toán mờ này bao gồm các bước: làm mờ, suy luận mờ (ra quyết định với cơ sở quy tắc mờ), giải mờ và đầu ra rõ nét để kiểm soát nhiệt độ, độ ẩm, ánh sáng của ngôi nhà thông minh. Webcam thực hiện thu hình ảnh, bộ xử lý trung tâm xử lý, nhận diện khuôn mặt để kiểm soát vào/ra. Hệ thống sử dụng Raspberry Pi 4 làm bộ điều khiển trung tâm, Arduino Uno R3 là bộ điều khiển thứ cấp tiếp nhận tín hiệu từ cảm biến, DHT11 làm cảm biến nhiệt độ, độ ẩm, BH1750 làm cảm biến ánh sáng, ... Tiếp theo, chúng tôi sẽ trình bày việc thiết kế các mô-đun nhận diện giám sát vào/ra, mô-đun điều khiển máy điều hòa, và mô-đun điều khiển mờ đèn.

2.2. Mô-đun nhận diện khuôn mặt giám sát vào/ra

Hệ thống mở khóa điện thoại dựa trên nhận dạng khuôn mặt được Apple tung ra trên iPhone của mình vào năm 2017 đã làm mưa làm gió trên thế giới và trở thành một công nghệ đột phá kể từ đó trong ngành công nghiệp điện thoại thông minh. Tương tự, các thiết bị an ninh, giám sát và tự động hóa gia đình gần đây đã trở thành một phần không thể thiếu trong lối sống của nhiều người. Hệ thống mở khóa cửa sử dụng hệ thống nhận dạng khuôn mặt làm mật khẩu để mở/đóng cửa ngôi nhà thông minh.



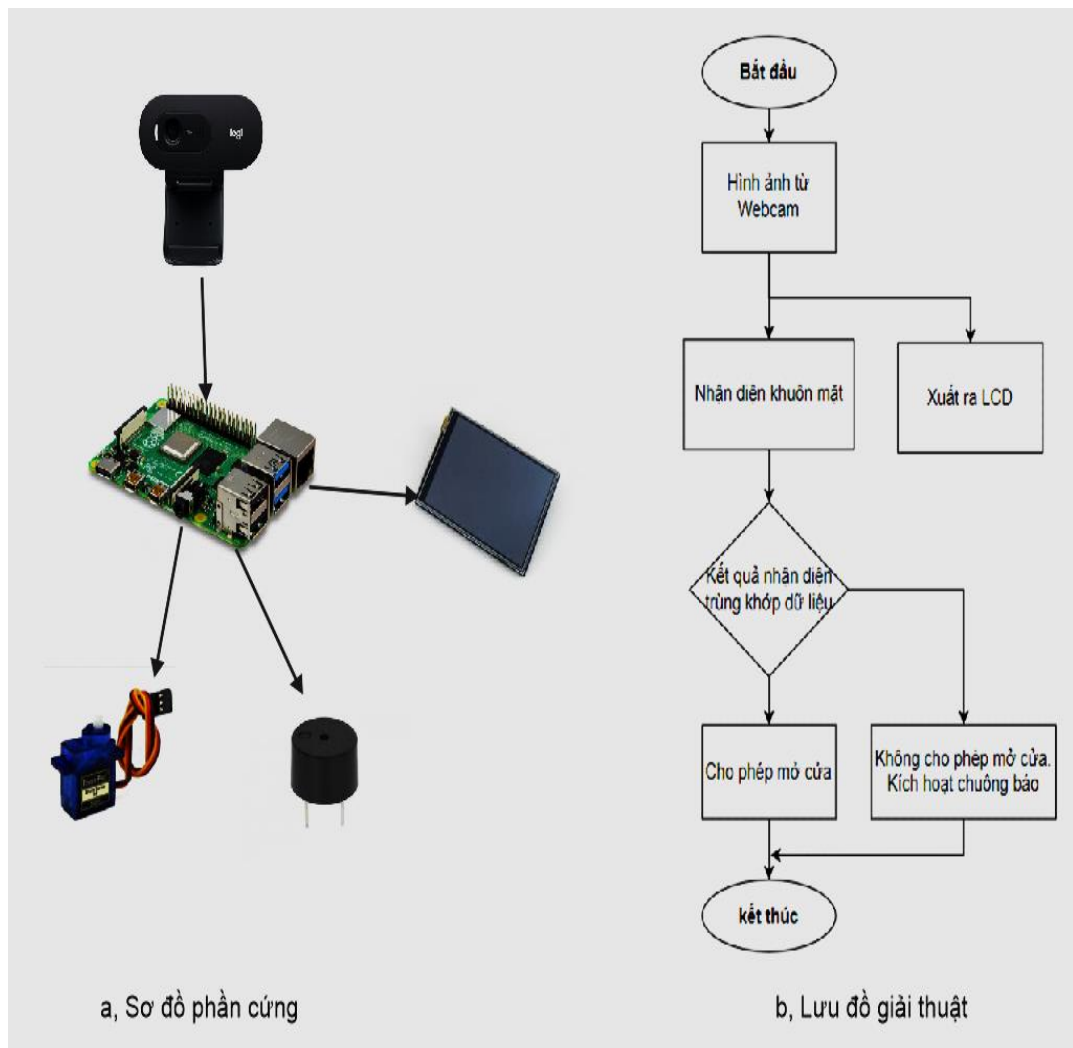
Hình 2.4: Kỹ thuật học sâu để nhận dạng khuôn mặt [7]

Trình tự các bước nhận dạng khuôn mặt theo phương pháp học sâu được thể hiện trong Hình 2.4. Sau khi được lưu lại bằng webcam, ảnh vào sẽ được tiền xử lý để cắt lấy ảnh khuôn mặt. Nếu khuôn mặt được quan sát khớp với các khuôn mặt chính được định cấu hình trong ứng dụng, nó sẽ gửi một thông báo tới cửa để

mở khóa. Webcam liên tục quan sát nguồn cấp dữ liệu video trực tiếp và phát hiện khuôn mặt người trong đó. Nếu khuôn mặt của con người khớp với khuôn mặt của chìa khóa đã được định cấu hình thì nó sẽ mở khóa cửa. Ứng dụng hoàn chỉnh được phát triển bằng ngôn ngữ lập trình Python [6]. Các thư viện được sử dụng gồm cho mô-đun nhận dạng khuôn mặt gồm:

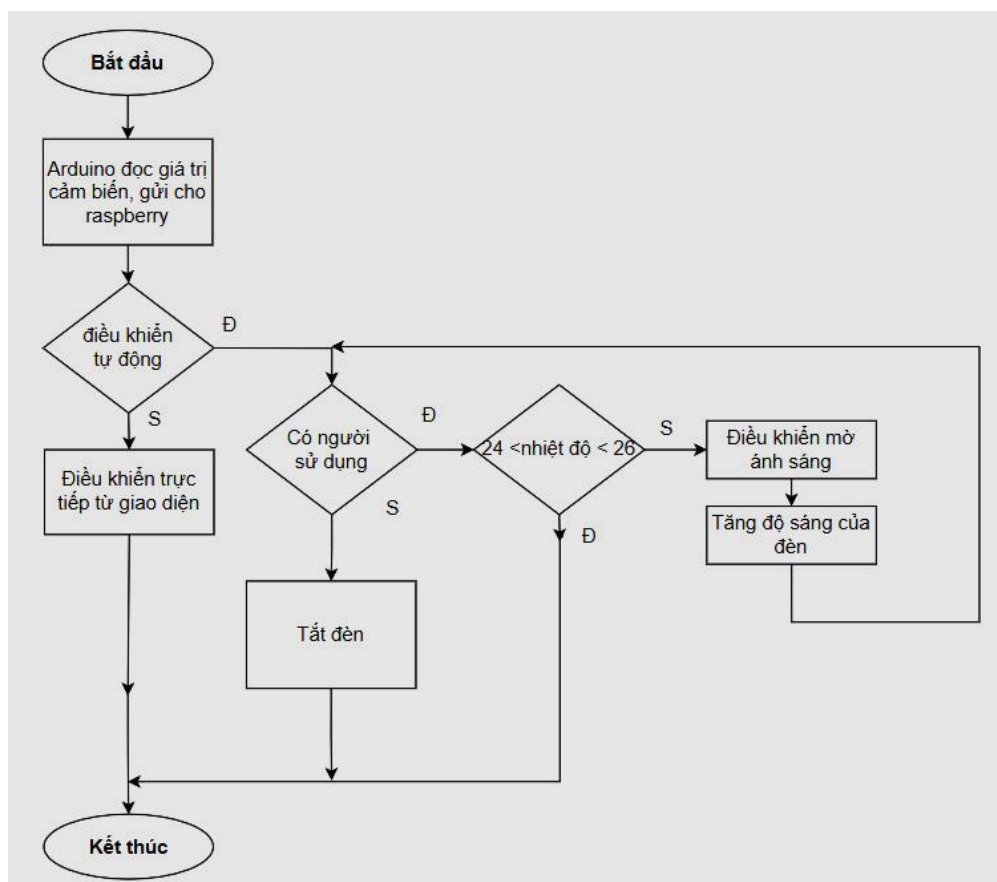
- dlib - thư viện học máy và học sâu
- face_recognition - thư viện nhận dạng khuôn mặt.
- RPi.GPIO - Thư viện điều khiển các đầu ra chân GPIO của Raspberry Pi
- PyQt5 - Các liên kết Python cho thư viện Qt C ++ để phát triển giao diện người dùng và các công cụ hiển thị.
- OpenCV - Thư viện Computer Vision để đọc hình ảnh.

Sơ đồ phần cứng hệ thống mở cửa dựa trên nhận diện khuôn mặt thể hiện ở Hình 2.5. Ở đây nút nhấn có vài trò kích hoạt sự kiện, khi nút nhấn được nhấn webcam sẽ thu nhận hình ảnh và phát hiện khuôn mặt người trong đó. Nếu khuôn mặt con người khớp với khuôn mặt đã được cấu hình, hệ thống sẽ hoạt động với mở cửa cho người nhà thì Raspberry pi sẽ điều khiển động cơ servo mở cửa. Còn nếu không khớp với các khuôn mặt đã được cấu hình hệ thống sẽ hoạt động với chế độ đón khách, Raspberry pi sẽ điều khiển buzzer kêu lên như một cái chuông thông báo. Người dùng có thể quan sát được ai đang ở ngoài dựa trên giao diện điều khiển trên màn hình LCD và có thể lựa chọn mở cửa hoặc không.



Hình 2.5: Mô-đun nhận dạng khuôn mặt giám sát vào/ra

2.3. Mô-đun điều khiển mờ máy điều hòa

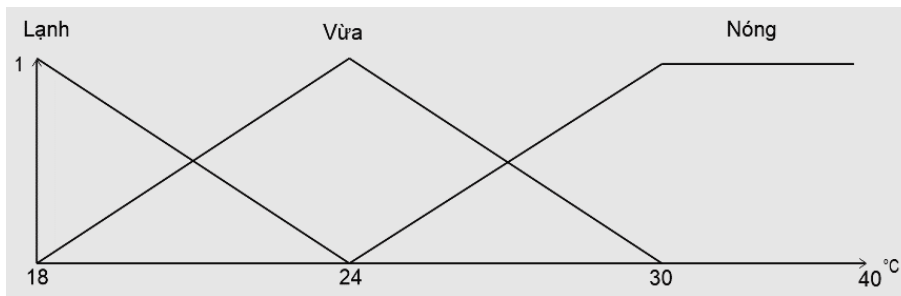


Hình 2.6: Hệ thống điều khiển máy điều hòa.

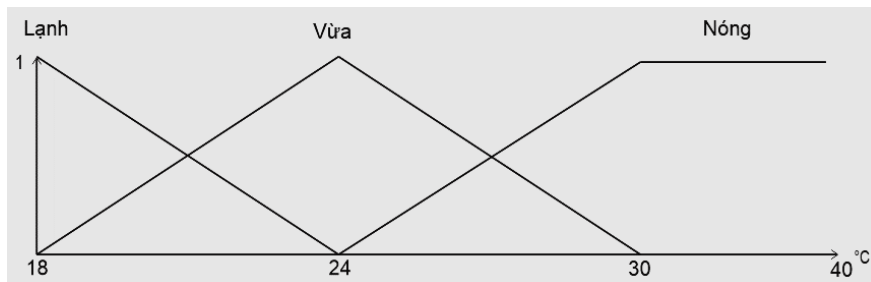
Bài báo đề xuất thiết kế hệ thống điều khiển nhiệt độ và độ ẩm trong phòng của ngôi nhà thông minh sử dụng logic mờ dựa trên Raspberry làm điều và điều khiển từ xa phát tia hồng ngoại để điều khiển nhiệt độ và cài đặt chế độ trong điều hòa để điều khiển nhiệt độ và độ ẩm của trong phòng. Nhiệt độ phù hợp với cơ thể giao động trong khoảng 24 – 26°C. Nhiệt độ quá thấp có thể sẽ gây ra một số bệnh như ngạt mũi, viêm họng, ...

Hình 2.6 cho thấy sơ đồ kết nối giữa cảm biến PIR, DHT11, led hồng ngoại, Arduino Uno R3 và Raspberry Pi 4. Xây dựng chức năng giám sát nhiệt độ, độ ẩm trong ngôi nhà và điều khiển máy lạnh theo hai chế độ. Cảm biến DHT11 đo nhiệt độ, độ ẩm trong nhà đồng thời cung cấp dữ liệu độ ẩm xây dựng chế độ điều khiển tự động máy lạnh theo độ ẩm trong nhà đã thu thập qua cảm biến DHT11, với chế độ tự động điều khiển ở máy lạnh thêm điều kiện có chuyển động của con người mới thực hiện mở máy lạnh, dữ liệu theo dõi chuyển động người thu thập thông qua cảm biến chuyển động hồng ngoại HC-SR501 và cập nhật trạng thái máy lạnh thông qua led hồng ngoại. Led hồng ngoại thu thập tín hiệu của máy lạnh để xử lý và phát ra tín hiệu điều khiển máy lạnh. Ngoài ra, người dùng có thể tự mình điều khiển theo ý thích của bản thân.

Đầu tiên để thiết kế bộ điều khiển mờ cho máy điều hòa, trước hết cần phải mờ hóa các giá trị ngõ vào/ra của máy điều hòa không khí. Với giá trị ngõ vào sẽ được lấy từ cảm biến nhiệt độ và độ ẩm (DHT11). Thông thường các loại máy điều hòa sẽ điều khiển nhiệt độ trong khoảng từ 18°C đến 30°C. Vì vậy giá trị ngõ ra sẽ nằm trong phạm vi này. Lựa chọn biến ngôn ngữ và biểu diễn dưới dạng hàm liên thuộc Hình 2.7 và Hình 2.8.

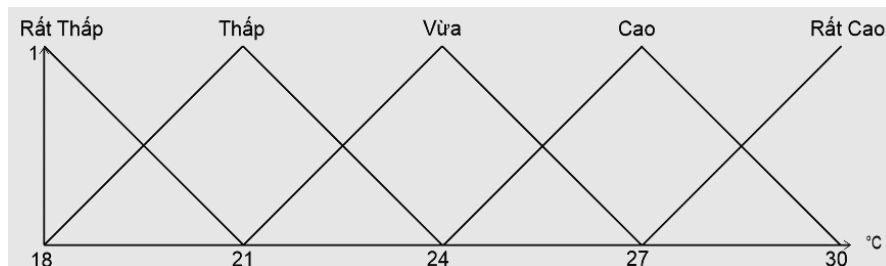


Hình 2.7: Nhiệt độ phòng từ cảm biến



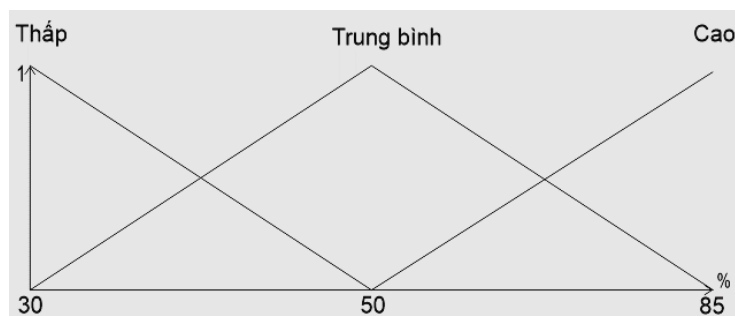
Hình 2.8: Nhiệt độ mục tiêu điều khiển

Để điều khiển được máy điều hòa không khí, đối với phần điều khiển nhiệt độ cần có 2 giá trị ngõ vào đó là nhiệt độ phòng và nhiệt độ mục tiêu. Giá trị ngõ ra sẽ được chia ra thành 5 tập mờ như Hình 2.9.

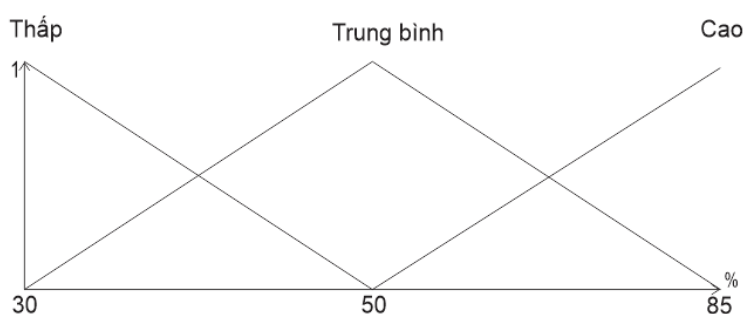


Hình 2.9: Nhiệt độ mục tiêu điều khiển

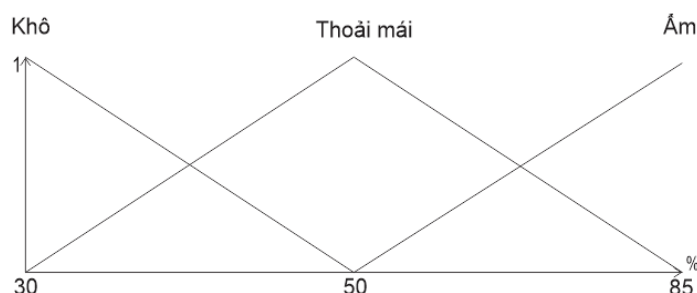
Tương tự như nhiệt độ, đối với độ ẩm các giá trị đầu vào/ra cũng sẽ được mờ hóa như Hình 2.10, Hình 2.11 và Hình 2.12.



Hình 2.10: Độ ẩm phòng từ cảm biến



Hình 2.11: Độ ẩm mục tiêu điều khiển



Hình 2.12: Giá trị ngõ ra của điều khiển độ ẩm

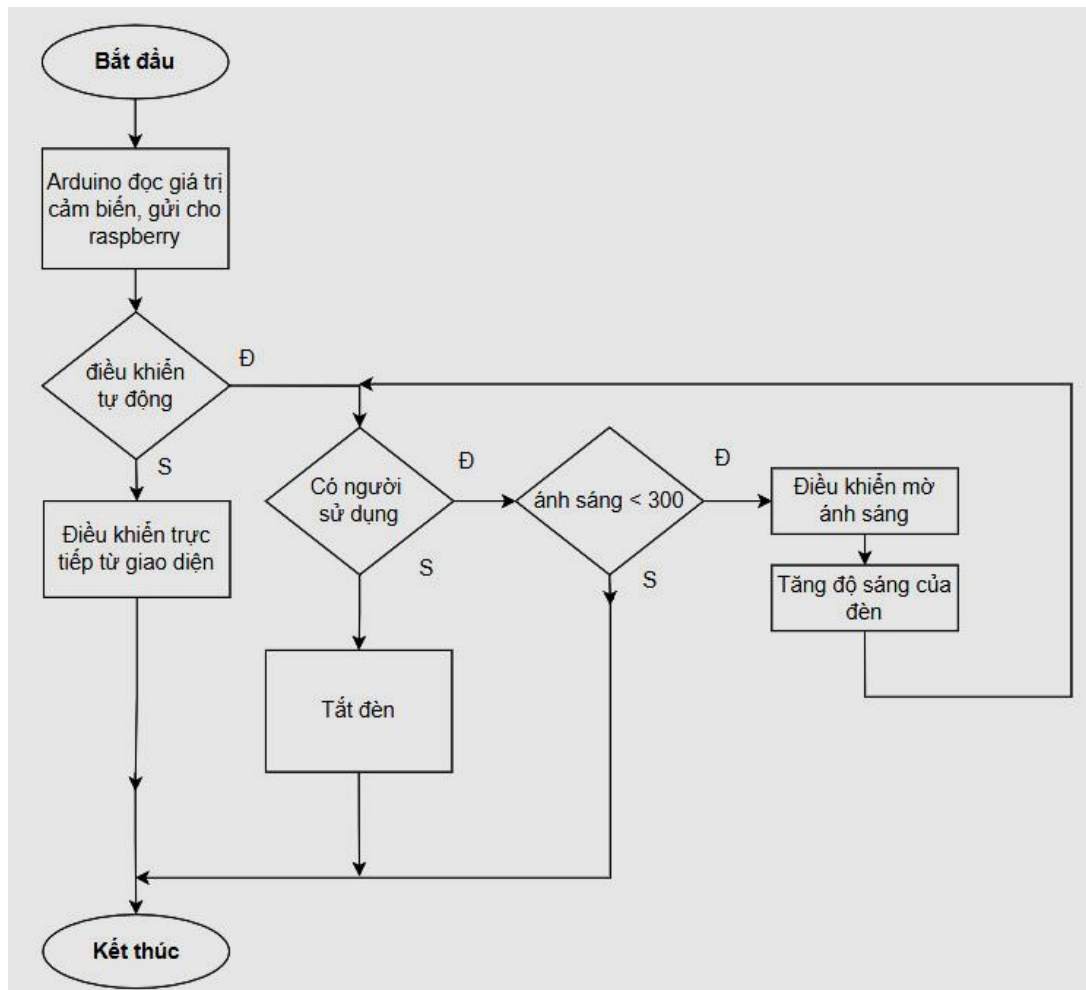
Bảng 2.1: Luật điều khiển cho bộ điều khiển mờ máy lạnh

$T_{\text{Cảm biến}} \backslash T_{\text{Mục tiêu}}$	Lạnh	Vừa	Nóng
Lạnh	Thấp	Cao	Rất cao
Vừa	Thấp	Vừa	Cao
Nóng	Rất thấp	Thấp	Cao

Sau khi đã mờ hóa các giá trị vào/ra thì tiếp theo sẽ chọn luật điều khiển cho bộ điều khiển mờ, luật điều khiển cho bộ điều khiển mờ điều hòa không khí biểu diễn ở Bảng 2.1. Chúng tôi đã sử dụng thư viện Skfuzzy của SciPy để lập trình bộ điều khiển mờ bằng ngôn ngữ lập trình Python và chạy trên Raspberry pi 4.

2.4. Điều khiển mờ kiểm soát ánh sáng

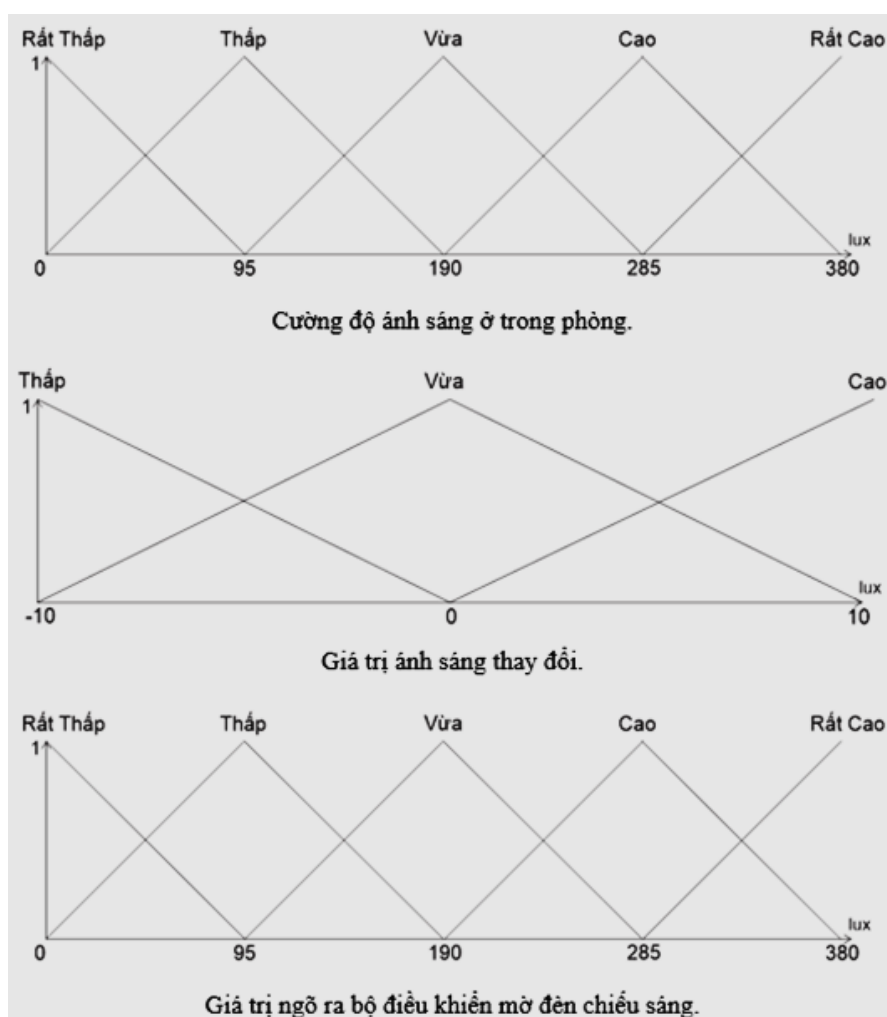
Ngoài các giải pháp kỹ thuật đang được ứng dụng như sử dụng đèn có hiệu suất cao hơn, việc tận dụng nguồn ánh sáng tự nhiên cần được xem xét tận dụng do tính sẵn có và liên tục trong suốt thời gian làm việc trong ngày. Sự phối hợp điều khiển giữa ánh sáng tự nhiên và ánh sáng nhân tạo cần xây dựng bộ điều khiển đáp ứng được sự thay đổi liên tục ánh sáng tự nhiên trong ngày. Độ tính chất linh hoạt trong điều khiển dựa trên hệ thống cơ sở tri thức, bộ điều khiển fuzzy được đề xuất nhằm đáp ứng lại những thay đổi của ánh sáng tự nhiên, yếu tố luôn thay đổi theo thời gian trong ngày. Bên cạnh đó, bộ điều khiển fuzzy còn có khả năng chống lại tín hiệu nhiễu của môi trường do hiện tượng mây che hay bóng râm thoáng qua do hàm ngõ ra là không tuyến tính. Theo tiêu chuẩn ánh sáng, độ lợi của phòng khách nên có giá trị khoảng 300 lux trở lên. Ngoài ra có các yếu tố khác như độ đồng đều, chỉ số hoàn màu,...



Hình 2.13: Hệ thống điều khiển kiểm soát ánh sáng

Sơ đồ kết nối giữa cảm biến PIR, BH1750, led hồng ngoại, Arduino Uno R3 và Raspberry Pi 4 được thể hiện trong Hình 2.13. Hệ thống kiểm soát ánh sáng dùng 2 cảm biến BH1750 để thu thập ánh sáng ở trong phòng và ở ngoài. Cảm biến chuyển động thu thập thông qua cảm biến chuyển động hồng ngoại HC-SR501 và cập nhật trạng thái để điều khiển rèm cửa và ánh sáng của đèn trong phòng. Bộ điều khiển mờ sẽ xử lý dữ liệu ánh sáng nhận được từ cảm biến để đưa ra tín hiệu điều khiển thích hợp.

Độ rọi (có đơn vị là lux) không phải là phép đo trực tiếp của năng lượng ánh sáng, mà nó là sự cảm nhận của mắt người. Vì thế, hệ số chuyển đổi sẽ thay đổi theo thành phần bước sóng hay nhiệt độ màu của ánh sáng. Để điều khiển hệ thống chiếu sáng cho ngôi nhà thông minh, cần phải cân bằng giữa ánh sáng tự nhiên và ánh sáng nhân tạo. Các giá trị ngõ vào/ra sẽ được mờ hóa như Hình 2.14 cho bộ điều khiển mờ hệ thống đèn chiếu sáng. Luật điều khiển của bộ điều khiển mờ cho hệ thống ánh sáng được biểu diễn ở Bảng 2.2, phụ thuộc vào 2 giá trị đầu vào và đưa ra giá trị điều khiển.



Hình 2.14: Giá trị ngõ ra của điều khiển độ ẩm

Bảng 2.2: Luật điều khiển cho bộ điều khiển mờ hệ thống đèn chiếu sáng

Ánh sáng trong phòng / Giá trị chênh lệch	Rất thấp	Thấp	Vừa	Cao	Rất cao
Âm	Rất cao	Rất cao	Cao	Vừa	Thấp
Cân bằng	Cao	Cao	Vừa	Thấp	Rất thấp
Dương	Cao	Vừa	Thấp	Rất thấp	Rất thấp

3. KẾT QUẢ

3.1. Mô hình ngôi nhà thông minh và giao diện phần mềm điều khiển

Hình 3.1 là mô hình ngôi nhà thông minh mà nhóm đã xây dựng và hoàn thiện theo ý tưởng mô hình ở Hình 2.1. Giao diện phần mềm điều khiển tương ứng thể hiện ở Hình 3.2.



a. Mô hình thực tế



b. Minh họa hệ thống hoạt động

Hình 3.1: Mô hình ngôi nhà thông minh thực tế

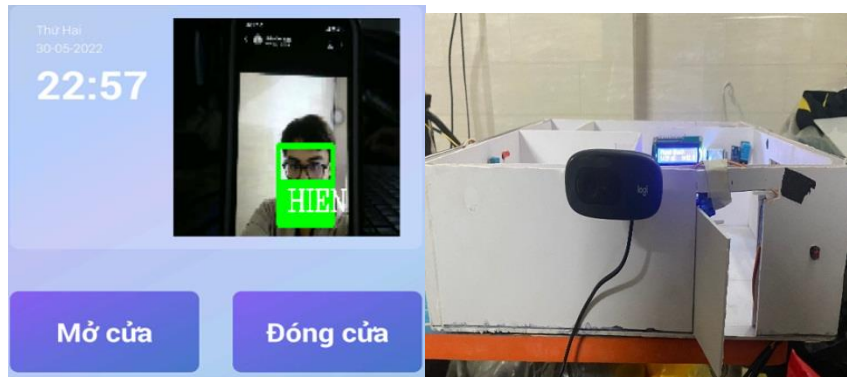
Giao diện chính cho phép người dùng chọn phòng cần điều khiển thiết bị. Người dùng có thể điều khiển các chức năng cho từng phòng tương ứng với các mô-đun điều khiển với 2 chế độ “tự động” và “tự chỉnh” chúng tôi đã trình bày ở Phần 2.



Hình 3.2: Giao diện giám sát và điều khiển ngôi nhà thông minh

3.2. Thực nghiệm nhận dạng khuôn mặt giám sát vào/ra

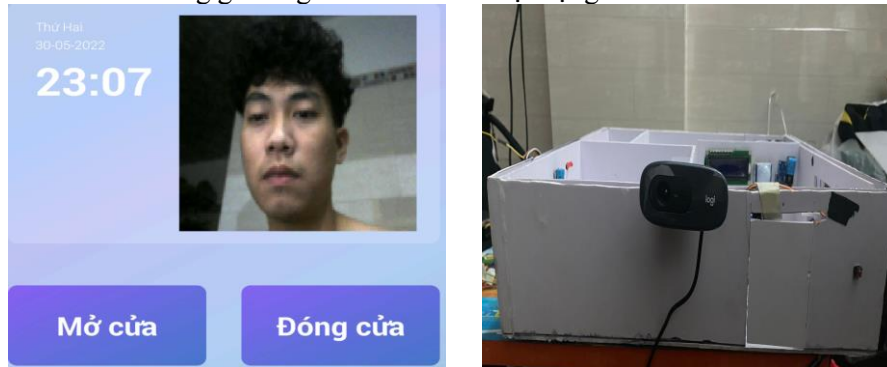
Kết quả nhận dạng đối với trường hợp người vào là người trong gia đình được thể hiện ở Hình 3.3. Giao diện phần mềm ở Hình 3.3.a cho thấy khuôn mặt của một người tên “HIEN” được nhận dạng. Cửa ra vào được mở ngay sao đó như Hình 3.3.b.



a. Hình ảnh khuôn mặt và tên tương ứng b. Cửa mở

Hình 3.3: Kết quả nhận dạng người trong gia đình

Kết quả nhận dạng đối với trường hợp người vào là người ngoài gia đình được thể hiện ở Hình 3.4. Trong trường hợp này khuôn mặt của người này cũng được đóng khung và thể hiện trên giao diện phần mềm như ở Hình 3.4.a nhưng cửa vẫn đóng như quan sát ở Hình 3.4.b. Các kết quả trên xác nhận tính mô-đun nhận dạng khuôn mặt thành viên trong gia đình giám sát vào/ra hoạt động chính xác.

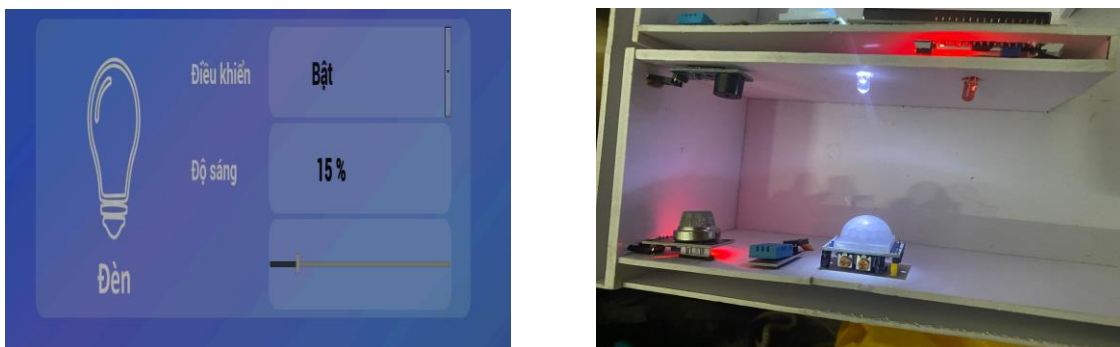


a. Hình ảnh khuôn mặt b. Cửa đóng

Hình 3.4: Kết quả nhận dạng người ngoài gia đình

3.3. Thử nghiệm chức năng tự chỉnh

Kết quả điều khiển bóng đèn² thông qua giao diện với độ sáng 15% được thể hiện ở Hình 3.5. Giao diện điều khiển và đèn thay đổi độ sáng được cho thấy lần lượt ở Hình 3.5.a và Hình 3.5.b.

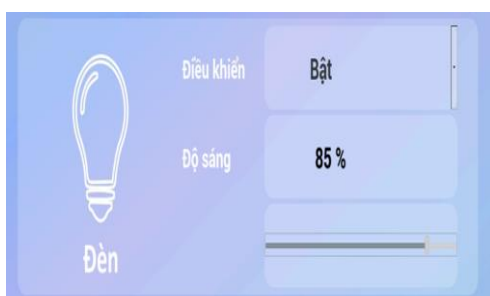


a. Giao diện điều khiển đèn b. Kết quả điều khiển đèn

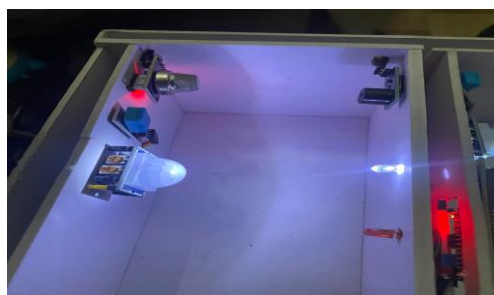
Hình 3.5: Kết quả điều khiển đèn bằng giao diện với độ sáng 15%

² Bài báo nhằm ứng dụng giải thuật mờ để điều khiển thiết bị trong đó có đèn LED đơn được sử dụng trong mô hình ngôi nhà thông minh là hoàn toàn phù hợp và được sử dụng phổ biến trong các nghiên cứu liên quan ví dụ [1, 2]. Để kiểm tra tính khả thi trong điều khiển thực tế, bóng đèn kết hợp với mạch điện tử công suất được dùng thay vì LED.

Tương tự, đèn sáng hơn khi được điều chỉnh ở độ sáng 85% cho thấy ở Hình 3.6. Kết quả cho thấy chức năng tự điều khiển thiết bị trực tiếp bằng giao diện hoạt động hiệu quả.



a. Giao diện điều khiển đèn



b. Kết quả điều khiển đèn

Hình 3.6: Kết quả điều khiển đèn bằng giao diện với độ sáng 85%

3.4. Thực nghiệm điều khiển mờ cho máy điều hòa

Bảng 3.1: Kết quả điều khiển nhiệt độ

Nhiệt độ mục tiêu	Nhiệt độ cảm biến	Nhiệt độ điều khiển	Chế độ điều khiển
25	18	27.748	COOL
25	19	27.239	COOL
25	20	26.767	COOL
25	21	26.385	COOL
25	22	25.877	COOL
25	23	25.281	COOL
25	24	25.148	COOL
25	25	25.148	COOL
25	26	25.148	COOL
25	27	24.307	COOL
25	28	23.73	COOL
25	29	23.248	COOL
25	30	22.748	COOL
25	31	22.748	COOL

Trước tiên chúng tôi kiểm tra điều khiển mờ cho thông số nhiệt độ. Đặt nhiệt độ mục tiêu cố định ở mức 25°C, độ ẩm không thay đổi ở mức 50%. Bằng cách thay đổi giá trị nhiệt độ thu được từ cảm biến, hệ thống xử lý, đưa ra kết quả điều khiển máy lạnh là nhiệt độ điều khiển. Bảng 3.1 cho thấy kết quả điều khiển. Ta có thể thấy nhiệt độ thay đổi để đạt được nhiệt độ mục tiêu và duy trì ổn định tại giá trị này.

Tiếp theo chúng tôi thực nghiệm về chế độ kiểm soát độ ẩm. Tiến hành tương tự, chúng tôi đặt độ ẩm mục tiêu cố định ở mức 50%, độ ẩm thu được từ cảm biến thay đổi. Chạy chương trình điều khiển và thu được kết quả ở Bảng 3.2. Kết quả cho thấy khi độ ẩm quá cao, hệ thống sẽ điều khiển máy điều hòa hoạt động ở chế độ DRY MODE để làm khô, giảm bớt lại độ ẩm trong không khí. Ngược lại, khi độ ẩm thấp, hệ thống điều khiển máy lạnh hoạt động ở chế độ COOL MODE để tăng độ ẩm trong không khí. Điều này mang đến sự thoải mái, dễ chịu cho người sử dụng.

Bảng 3.2: Kết quả điều khiển chế độ hoạt động của máy lạnh

Độ ẩm mục tiêu	Độ ẩm mục tiêu	Chế độ điều khiển
50	40	COOL
50	45	COOL
50	50	COOL
50	55	COOL
50	60	DRY
50	65	DRY
50	70	DRY
50	75	DRY
50	80	DRY

3.5. Thử nghiệm điều khiển mờ độ sáng đèn

Chúng tôi duy trì mức chênh lệch ánh sáng là bằng 0 (lux). Bằng cách thay đổi giá trị độ rọi thu được từ cảm biến, sau khi thực hiện xử lý bằng bộ điều khiển logic mờ, chúng tôi thu được tín hiệu điều khiển đèn như ở Bảng 3.3. Kết quả cho thấy ánh sáng trong phòng luôn đảm bảo ở mức vừa phải nhưng vẫn đảm bảo tiết kiệm năng lượng.

Bảng 3.3: Kết quả điều khiển đèn

Độ rọi trong phòng (lux)	Mức chênh lệch độ rọi (lux)	Tín hiệu điều khiển đèn (%)
80	0	78
100	0	76
120	0	71
140	0	66
160	0	60
180	0	50
200	0	44
220	0	37
240	0	33
260	0	28
280	0	23
300	0	21
320	0	18
340	0	15
360	0	9
380	0	0

3.6. Thử nghiệm điều khiển mờ cho rèm cửa

Hệ thống điều khiển rèm cửa được thực nghiệm và kiểm tra giống như thực nghiệm điều khiển đèn. Chúng tôi duy trì mức chênh lệch ánh sáng bằng 0 (lux). Thay đổi giá trị độ rọi thu được từ cảm biến, sau khi thực hiện xử lý bằng bộ điều khiển logic mờ, tín hiệu điều khiển đèn thể hiện ở Bảng 3.4. Kết quả cho thấy rèm cửa được đóng/mở để ánh sáng bên ngoài chiếu vào trong nhà không quá cao. Mục đích là tận dụng được ánh sáng tự nhiên cho căn phòng. Rèm cửa đóng hoàn toàn khi tín hiệu đạt 0 %.

Bảng 3.4: Kết quả điều khiển rèm cửa

Độ rọi bên ngoài (lux)	Mức chênh lệch độ rọi(lux)	Tín hiệu điều khiển rèm (%)
100	0	85
150	0	85
200	0	85
250	0	85
300	0	77
350	0	73
400	0	71
450	0	69
500	0	68
550	0	59
600	0	53
650	0	50
700	0	44
750	0	35
800	0	34
850	0	32

4. KẾT LUẬN

Chúng tôi đã trình bày một giải pháp xây dựng hệ thống giám sát và điều khiển cho ngôi nhà thông minh. Các thiết bị điện như đèn, điều hòa, rèm cửa,... được điều khiển thông minh theo giải thuật logic mờ. Bên cạnh đó chế độ điều khiển thiết bị trực tiếp thông qua giao diện cũng được phát triển kèm theo. Các hoạt động vào/ra ngôi nhà được giám sát bằng mô-đun nhận dạng khuôn mặt thành viên gia đình dựa trên phương pháp học sâu. Kết quả thực nghiệm cho thấy các mô-đun nhận dạng khuôn mặt và điều khiển thiết bị hoạt động tốt. Hướng phát triển tiếp theo có thể là nâng cao tốc độ xử lý của mô-đun nhận dạng khuôn mặt dùng phương pháp học sâu. Kế đó là tích hợp chức năng điều khiển thiết bị bằng giọng nói vào hệ thống.

TÀI LIỆU THAM KHẢO

- [1] T. T. M. Khoa, N. C. A. Minh, và N. T. Hậu, Ứng dụng IOT trong hệ thống quản lý và giám sát ngôi nhà thông minh theo thời gian thực, *Tạp chí Khoa học và Công nghệ, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh*, số 50, 2021.
- [2] N. V. Thăng, P. T. Minh, N. C. Toàn, và N. T. Đức, Ứng dụng Internet of things xây dựng ngôi nhà thông minh, *Tạp chí Khoa học Công nghệ Hàng hải*, số 47, 2016.
- [3] N. V. Tiến và H. X. Bình, Nghiên cứu ứng dụng trí thông minh nhân tạo và Arduino điều khiển các thiết bị điện trong ngôi nhà thông minh, *Tạp chí Khoa học Công nghệ Hàng hải*, số 59, 2019.
- [4] V. M. Phụng, D. T. T. Hiên, và V. T. Phúc, Nghiên cứu xây dựng mô hình điều khiển nhà thông minh sử dụng kết nối Bluetooth, GSM, *Tạp chí Khoa học Đại học Đà Lạt*, tập 8, số 3, 2018.
- [5] D. Norris, *Home Automation with Raspberry Pi*, *Mc Graw Hill Education*, 2019.
- [6] T. T. B. Nguyễn và H. Q. Phạm, Lập trình hệ thống nhúng với Raspberry, *Nhà xuất bản Thanh Niên*, 2019.
- [7] B. C. Michał Koziarski, Recognition with Deep Neural Networks in Presence of Noise - Dealing with and Taking Advantage of Distortions, *Integrated Computer-Aided Engineering*, vol. 24, no. 4, pp. 337-349, 2017.

YSC4F.220

ỨNG DỤNG MÔ HÌNH DEEP SPEECH TRONG NHẬN DẠNG GIỌNG NÓI TIẾNG VIỆT

NGUYỄN VĂN ANH TUẤN, PHẠM HUỲNH TRƯỜNG THỊNH, PHẠM TRUNG SƠN,
ĐẶNG THỊ PHÚC

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
nvatuan3@gmail.com, thanh7174@gmail.com, sonphukl@gmail.com, phucdt@iuh.edu.vn

Tóm tắt. Nghiên cứu về nhận dạng giọng nói đã có từ rất lâu, nhưng có rất ít nghiên cứu trong việc ứng dụng học sâu vào nhận dạng giọng nói ngôn ngữ tiếng Việt. Trong bài nghiên cứu này, chúng tôi giải quyết bài toán nhận dạng giọng nói tiếng Việt bằng mô hình Deep Speech được giới thiệu lần đầu vào năm 2015. Kết quả thực nghiệm trên bộ dữ liệu Vivos có được kết quả cơ bản để nghiên cứu sâu hơn vào những kiến trúc mô hình và bộ dữ liệu khác.

Từ khóa. Nhận dạng giọng nói, học sâu, Deep Speech, nhận dạng giọng nói tiếng Việt.

VIETNAMESE VOICE RECOGNITION USING DEEP LEARNING MODEL

Abstract. Research on speech recognition has existed for a long time, but there has been very little research work on the application of deep learning to the Vietnamese language. In this study, we solve the Vietnamese language speech recognition problem using a Deep Speech model first published in 2015. Experimental results on the Vivos dataset provide basic results for further research into other model architectures and datasets.

Keywords. Speech recognition, deep learning, Deep Speech, Vietnamese speech recognition.

1. GIỚI THIỆU BÀI TOÁN

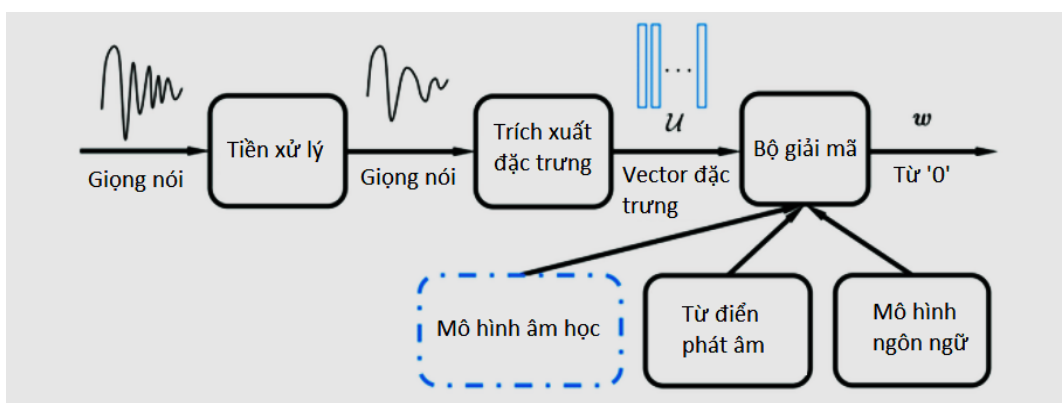
Trí tuệ nhân tạo (AI) đóng một vai trò quan trọng trong sự phát triển của xã hội vì những tiềm năng và ứng dụng to lớn của nó mang lại trong những năm gần đây. Một số ứng dụng AI nổi tiếng có thể kể đến như Siri (Apple), Google Assistant (Google), Cortana (Microsoft) và Watson (IBM). Tất cả các ứng dụng kể trên đều có liên quan đến giọng nói của con người hay còn được biết đến đó là “Trợ lý ảo”. Trợ lý ảo là ứng dụng có thể giao tiếp, hỗ trợ con người thông qua văn bản và giọng nói. Trong các ứng dụng gần đây, trợ lý ảo đã được triển khai trong nhiều lĩnh vực như ô tô tự lái, tổng đài ảo, báo nói, dịch máy và giáo dục. Quá trình của những ứng dụng này là nhận yêu cầu từ người dùng, xử lý yêu cầu và phản hồi các yêu cầu. Đối với các ứng dụng dựa trên giọng nói, tự động nhận dạng giọng nói (Automatic Speech Recognition - ASR) đóng vai trò là bước đầu tiên của quá trình này. ASR có một vai trò quan trọng vì nó ảnh hưởng đến chất lượng của những ứng dụng liên quan. Do đó việc xây dựng mô hình nhận dạng giọng nói mang tính cấp thiết trong chủ đề này.

Tiếng Việt là ngôn ngữ chính thức duy nhất của nước Việt Nam. Hệ thống chữ viết Việt Nam được xây dựng vào thế kỉ XVII, hệ thống chữ viết này dựa trên các kí tự la mã có dấu phụ, dễ nhớ dễ viết hơn chữ viết trước đây dựa trên các kí tự Trung Quốc được gọi là chữ Nôm - ngày nay còn được gọi là chữ Quốc ngữ [1] [2]. Có thể nói, tiếng Việt rất phong phú về giọng nói, mỗi một vùng miền là một giọng nói đặc trưng khác nhau. Gần đây, đã có nhiều kĩ thuật học máy khác nhau có thể xử lí nhận dạng giọng nói cho tiếng Việt. Tuy nhiên, nghiên cứu đối với lĩnh vực này có một số trở ngại do đặc điểm của giọng nói ngôn ngữ Việt Nam. Do đó để giải quyết những trở ngại trên, trong bài báo này chúng tôi đưa ra đánh giá và so sánh giữa các mô hình nhận dạng giọng nói sử dụng kĩ thuật học sâu (deep learning) đang đạt kết quả tốt

trên một số ngôn ngữ. Từ đó chọn ra mô hình, kiến trúc phù hợp nhất với tiếng Việt về tốc độ và độ chính xác khi nhận dạng.

$$W^* = \arg \max_{W \in V^*} P(W|X) \quad (1)$$

Bài toán nhận dạng giọng nói có thể được giải quyết bằng cách phân rã phương trình (1) [3] thành mô hình âm thanh, mô hình ngôn ngữ và từ điển phát âm sử dụng định lý Bayes. Phương pháp này đã có từ lâu và thiên hướng nặng về giả định của bài toán luôn là độc lập có điều kiện và tối ưu riêng lẻ giữa các thành phần mô hình khác nhau. Có thể hiểu hệ thống này hoạt động bằng cách nhận đầu vào là dạng âm thanh đã được xử lý, sau đó dùng mô hình âm thanh là GMM-HMM để tìm phân phối xác suất sắp xếp giữa các khung âm thanh và các ký hiệu phiên âm chính xác, sau đó đưa vào trong từ điển phát âm và mô hình ngôn ngữ để làm cho kết quả chính xác hơn. Hình dưới mô tả hệ thống nhận dạng giọng nói kết hợp.



Hình 1: Hệ thống nhận dạng giọng nói kết hợp (Author)

Với sự ra đời của học sâu [4], kiến trúc DNN (mạng sâu) có thể thay thế cho GMM truyền thống để dự đoán phân phối xác suất các ký hiệu phiên âm. Một điểm trừ của hệ thống này, cũng giống hệ thống trước đó là việc tối ưu các thành phần của mô hình một cách riêng lẻ, không kết hợp được với nhau (huấn luyện mô hình ngôn ngữ không phụ thuộc vào âm thanh, huấn luyện mô hình âm thanh mà không phụ thuộc vào mô hình ngôn ngữ để lấy ngữ cảnh). Mà bởi vì tối ưu riêng lẻ, nên mỗi mô hình sẽ có một trạng thái hội tụ khác nhau, mà khi kết hợp với nhau sẽ không đưa ra kết quả tối ưu toàn cục được. Một ví dụ mà khiến hệ thống nhận dạng giọng nói kết hợp này kém hiệu quả là âm thanh được thu thập ở môi trường nhiễu nhiều (tiếng còi, tiếng xe, tiếng máy...) và sự đa dạng người nói (vùng miền, tuổi tác, cách nói, cao độ, ...). Vì thế, cần phải có một kiến trúc giải quyết được những điểm trừ trên, phải tối ưu một cách đồng thời với nhau các thành phần, kiến trúc đơn giản hơn và việc huấn luyện cũng sẽ đơn giản hơn, trực tiếp giải $P(W|X)$ mà không phải phân rã ra thành nhiều thành phần.

Với kiến trúc nhận dạng giọng nói đầu cuối (end-to-end speech recognition) (Li J. , 2021), những vấn đề trên hầu như đều có thể giải quyết. Cặp âm thanh (đầu vào) - phiên dịch (đầu ra) là thứ duy nhất cần đưa vào mô hình. Ngoài ra, việc sử dụng một kiến trúc đầu cuối giúp chuyển trực tiếp từ âm thanh sang đoạn phiên dịch, mà không cần thông qua phiên âm (có thể bỏ được từ điển phát âm ở hệ thống cũ). Hệ thống mới tối ưu trực tiếp các thành phần bằng back-propagation, giúp cho mô hình nhận dạng giọng nói đạt hiệu quả tốt hơn.

Kiến trúc mà nhóm sử dụng trong đề tài này là kiến trúc Deep Speech [6], được trình bày đầu tiên vào năm 2015, đến từ Baidu Research - Lab trí tuệ nhân tạo tại Thung lũng Silicon. Ở nghiên cứu gốc, các tác giả cho rằng hệ thống nhận dạng giọng nói đầu cuối Deep Speech có thể khắc phục các nhược điểm của hệ thống kết hợp cũ, bao gồm: có thể khắc phục được độ nhiễu trong dữ liệu âm thanh, sự vang của âm thanh và mức độ đa dạng người nói khác nhau bằng cách học trực tiếp sự thay đổi từ dữ liệu. Bên cạnh đó, hệ thống có kiến trúc đơn giản với BiRNN là mô hình chủ đạo nắm bắt thông tin theo thời gian của dữ liệu, kết hợp với CTC [7] - hàm lỗi giúp thay thế cho HMM truyền thống trong việc sắp xếp các khung âm thanh với đoạn phiên dịch, đạt hiệu quả tốt trong bài toán nhận dạng giọng nói, vì thế giúp Deep Speech đạt State-of-the-art trên hệ thống nhận dạng giọng nói thương mại vào năm 2015.

2. DỮ LIỆU

2.1. Thu thập dữ liệu

Để đảm bảo độ đa dạng dữ liệu chúng tôi đã thu thập tất cả 6 bộ dữ liệu khác nhau. Số lượng người nói được ước tính là 100 với giới tính được cân bằng và nhiều độ tuổi khác nhau. Tuy nhiên, phần lớn bản ghi được thu thập từ người có độ tuổi từ 20 - 30 do đó mô hình sẽ nhận diện tốt nhất với những lời nói to rõ. Bảng dưới đây là thống kê về những dữ liệu đã được thu thập.

Bảng 1: Thông tin dữ liệu

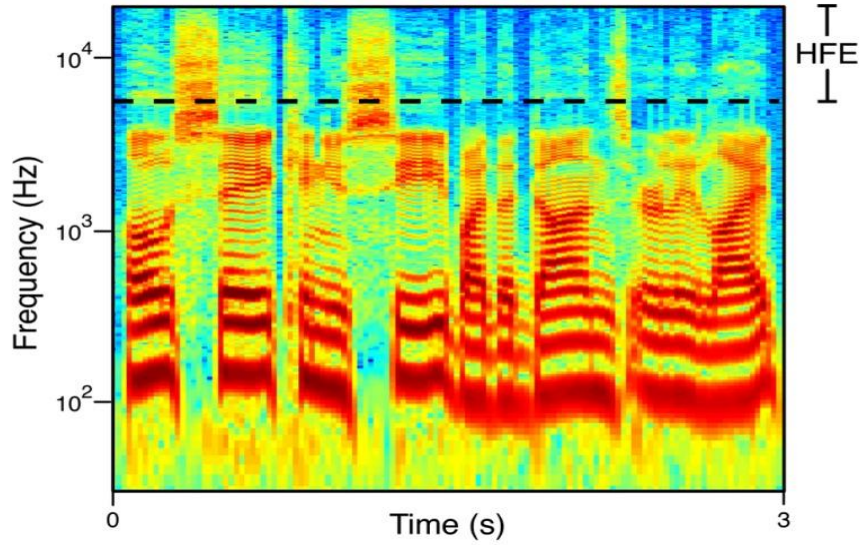
Tên dữ liệu	Số lượng bản ghi	Tổng thời gian (Giờ)	Thời gian trung bình (Giây)	Thời gian dài nhất (Giây)	Thời gian ngắn nhất (Giây)
VIVOS Corpus	12,360	15.67	8	29.4	1.42
VLSP2020 100h	56,427	100	6.47	80.64	0.32
FPT Open Speech Dataset (FOSD)	25,912	30	4.15	48.3	0.4
CV Corpus 8.0	15,305	16.69	3.92	23.2	0.5
Thu âm đọc sách	1,236	2.74	7.98	29.4	1.42
Cào từ Youtube	13,400	17.42	4.68	90.34	0.74
Tất cả data	124,640	182.52	5.86	50.21	0.79

Hầu hết dữ liệu được sử dụng là các bộ dữ liệu chuẩn được công khai trên internet. Tuy nhiên qua lấy mẫu cho thấy dữ liệu chưa đa dạng về miền nội dung nên nhóm đã thu thập thêm 2 bộ dữ liệu qua việc cào âm thanh từ Youtube và tự thu âm qua việc đọc sách. Các âm thanh lời nói cùng transcript được thu thập thông qua API mã nguồn mở của Youtube nên đảm bảo được độ chính xác của dữ liệu. Nhằm không để dữ liệu bị nhiễu, các bản ghi tự thu âm qua việc đọc sách được chúng tôi ghi lại trong môi trường hoàn toàn yên tĩnh và không có tạp âm.

2.2. Xử lý dữ liệu

Resample: Âm thanh lời nói trong dữ liệu có tỉ lệ mẫu (sample rate) cao làm cho việc tải dữ liệu vào mô hình và huấn luyện rất tốn kém thời gian và bộ nhớ máy tính. Việc chuyển đổi tỉ lệ mẫu cho tất cả các mẫu dữ liệu về 16000 cho thấy được hiệu quả tốt. Vì đây là tỷ lệ mẫu mà âm thanh lời nói khi nghe vẫn không đổi về mặt cảm nhận nhưng độ lớn của bản ghi đã được giảm xuống. Đồng thời, việc này cũng giúp giảm bớt các tính toán trong quá trình huấn luyện do kích thước ma trận đã được giảm xuống.

Spectrogram: là một cách biểu diễn âm thanh theo thời gian, tần số và độ lớn trên một biểu đồ. Cùng một từ được nói giữa những người nói khác nhau vẫn có spectrogram gần giống nhau. Do đó, để tổng quát được lời nói đưa vào bởi nhiều người nói khác nhau spectrogram tỏ ra rất hiệu quả. Dữ liệu sau khi resample sẽ được chuyển sang spectrogram để đưa vào mô hình.



Hình 2: Spectrogram của một đoạn âm thanh

3. MÔ HÌNH DEEP SPEECH

3.1. Kiến trúc Deep Speech

Thành phần chính của hệ thống Deep Speech [6] là mạng hồi quy chuyên tiếp (RNN), được huấn luyện để “ăn vào” dạng spectrogram của giọng nói và sinh ra các đoạn phiên dịch tương ứng. Hệ thống sẽ sử dụng spectrogram làm các feature chính cho mô hình, được chuyển hóa trực tiếp từ các file âm thanh đầu vào. Mục tiêu của RNN là chuyển đầu vào (spectrogram) thành phân phối xác suất trên bộ ký tự đầu ra, nếu gọi x là đoạn giọng nói, gọi y là đoạn phiên dịch thực tế thì $\hat{y} = P(c_t | x)$ là đoạn phiên dịch dự đoán, trong đó $c_t = \{a, \text{ă}, \text{â}, \text{á}, \text{à}, \dots, z, :, _ \}$, mà ký tự cuối “_” là ký tự rỗng (blank) dùng trong CTC.

Mô hình Deep Speech được kết hợp bởi 5 lớp ẩn. Với đầu vào x , các node ẩn tại lớp l được xác định bởi $h_t^{(l)}$ với kí quy định là $h_t^{(0)}$ là đầu vào. Ba lớp đầu tiên không hồi quy (chỉ là lớp ẩn thông thường). Với lớp đầu tiên, tại mỗi thời gian t , đầu ra phụ thuộc vào khung spectrogram x_t cùng với một ngữ cảnh C khung mỗi bên (thường là 5, 6, 7). Lớp không hồi quy còn lại thao tác trên dữ liệu độc lập với mỗi khung thời gian. Vì vậy, với mỗi thời gian t , 3 lớp đầu tiên được tính toán bằng công thức sau:

$$h_t^{(l)} = g(W^{(l)}h_t^{(l-1)} + b^{(l)}) \quad (2)$$

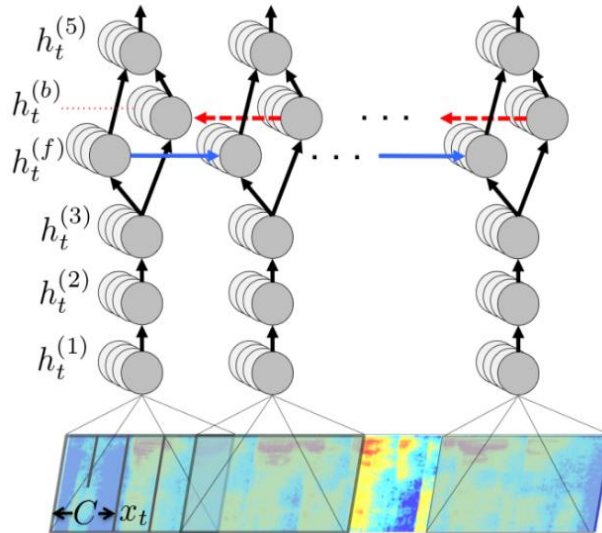
$g(z) = \min\{\max\{0, z\}, 20\}$ là hàm kích hoạt ReLU được cắt bớt và $W^{(l)}$, $b^{(l)}$ là ma trận trọng số và tham số thiên vị (bias) cho lớp thứ l . Lớp thứ tư là lớp hồi quy chuyên tiếp hai chiều (Bi-RNN). Lớp này chứa hai tập node ẩn: một tập với chiều thuận $h_t^{(f)}$, một tập với chiều ngược lại $h_t^{(b)}$, công thức như thế này:

$$\begin{aligned} h_t^{(f)} &= g(W^{(4)}h_t^{(3)} + W_r^{(f)}h_{t-1}^{(f)} + b^{(4)}) \\ h_t^{(b)} &= g(W^{(4)}h_t^{(3)} + W_r^{(b)}h_{t+1}^{(b)} + b^{(4)}) \end{aligned} \quad (3)$$

Tầng thứ năm nhận cả hai tập node ẩn theo chiều thuận và theo chiều ngược lại của RNN làm đầu vào $h_t^{(5)} = g(W^{(5)}h_t^{(4)} + b^{(5)})$ mà $h_t^{(4)} = h_t^{(f)} + h_t^{(b)}$. Đầu ra của lớp này được đưa qua hàm softmax và tạo thành một phân phối xác suất trên từng ký tự cho từng khung thời gian t và ký tự k trong tập từ vựng:

$$h_{t,k}^{(6)} = \hat{y}_{t,k} = P(c_t = k | x) = \frac{\exp(W_k^{(6)}h_t^{(5)} + b_k^{(6)})}{\sum_j \exp(W_j^{(6)}h_t^{(5)} + b_j^{(6)})} \quad (4)$$

Sau khi đã tính được $P(c_t|x)$, có thể tính lỗi với hàm mất mát CTC $\mathcal{L}(\hat{y}, y)$ và tính đạo hàm bằng back-propagation để truyền ngược lại các lớp trước, dùng Stochastic Gradient Descent (SGD) để tối ưu đồng thời các lớp, vì thế kiến trúc này mới gọi là kiến trúc đầu cuối, tối ưu từ đầu đến cuối. Hình dưới đây mô tả kiến trúc của Deep Speech.



Hình 3: Kiến trúc Deep Speech

3.2. Connectionist Temporal Classification (Hàm mục tiêu CTC)

Để mô hình hóa giọng nói hiệu quả, ta phải đi mô hình hóa cấu trúc theo thời gian của giọng nói. Đối với hệ thống kết hợp trước đây, mô hình HMM là bắt buộc trong việc căn chỉnh giữa đoạn giọng nói và đoạn phiên dịch (căn lề - alignment là một tác vụ nhằm điều chỉnh sự giống nhau của đầu vào và đầu ra theo từng khung thời gian). Có nhiều sự thay đổi trong kiến trúc nhận dạng giọng nói kết hợp, như là HMM, GMM-HMM [8], DNN-HMM [9], HMM-RNN [10], ... nhưng mô hình HMM vẫn là thành phần chính để thực hiện việc căn chỉnh.

HMM không phải kiến trúc học sâu, việc sử dụng trong hệ thống nhận dạng đầu cuối sẽ gặp khó khăn, việc sử dụng HMM để căn chỉnh lề và xuất ra các phiên âm có thể không thực hiện được vì lượng dữ liệu ngày càng lớn. Vì thế, người ta đã sáng tạo ra một phương pháp mới trong việc căn chỉnh lề và huấn luyện những bài toán có chuỗi đầu vào và đầu ra chưa được căn chỉnh là Connectionist Temporal Classification (CTC). Đưa trước một đoạn đầu vào ngữ âm $X = [x_1, x_2, \dots, x_T]$ là đoạn thuộc tính âm thanh với đoạn phiên dịch tương ứng là $Y = [y_1, y_2, \dots, y_U]$. Một căn chỉnh chính xác giữa X và Y chưa biết trước, và ta biết có sự khác nhau giữa độ dài của X và Y , thường là $T \ll U$ (độ dài đoạn âm thanh luôn dài hơn đoạn phiên dịch). Một cách căn chỉnh giữa X và Y được mô tả ở hình dưới đây.

Thuộc tính âm thanh đầu vào	x_1	x_2	x_3	x_4	x_5	x_6
Căn chỉnh thông thường	c	c	a	a	a	t
Đầu ra Y	c		a			t

Hình 4: Căn chỉnh thông thường của âm thanh

Nhưng cách căn chỉnh lề ở trên có hai vấn đề: Thứ nhất, trong nhận dạng giọng nói, đoạn đầu vào có thể có nhiều đoạn im lặng (không có giọng nói) mà có thể không căn chỉnh với đoạn chữ nào ở đầu ra. Thứ hai, ta có thể sẽ không bao giờ xuất ra được một từ nào mà có những ký tự liên tiếp giống nhau (ví dụ 2 chữ “o” trong “xoong”) bởi vì theo phương pháp trên, nó sẽ bị gộp lại thành một.

Phương pháp CTC làm giảm bớt vấn đề trên bằng cách đưa thêm một ký tự “blank” trong bộ từ vựng và coi nó như một ký hiệu phân biệt rỗng. Ký tự này sẽ được bỏ sau khi các ký tự liền kề giống nhau được gộp lại, nên ký tự này có thể coi như căn chỉnh lề cho các đoạn âm thanh yên lặng. Tuy nhiên, nó cho phép trực tiếp căn chỉnh giữa đầu vào và đầu ra mà không xuất ra phân loại trong bộ từ vựng. Hình dưới đây mô tả một ví dụ căn chỉnh lề giữa đoạn đầu vào và đoạn phiên dịch “hello world”.

Thuộc tính âm thanh đầu vào X:	x_1	x_2	x_3	x_4	x_5	x_6	x_7	x_8	x_9	x_{10}	x_{11}	x_{12}	x_{13}	x_{14}	x_{15}	x_{16}	x_{17}
Căn chỉnh được dự đoán:	h	e	l	l	ϵ	l	o	ϵ	ϵ	_	w	o	o	ϵ	r	l	d
Hợp nhất các dự đoán lặp lại:	h	e	l	ϵ	l	o	ϵ	_	w	o	ϵ	r	l	d			
Xóa ký tự trắng:	h	e	l		l	o			_	w	o		r	l	d		
CTC dự đoán đầu ra Y:	h	e	l	l	o	_	w	o	r	l	d						

Hình 5: Căn chỉnh âm thanh bằng CTC

Với cách đại diện này, ta có thể trực tiếp căn chỉnh 1:1 giữa hai độ dài chuỗi đầu vào và đầu ra khác nhau. Ký tự blank là ϵ , hình dưới đây đều mô tả từ “hello” với hai căn chỉnh khác nhau

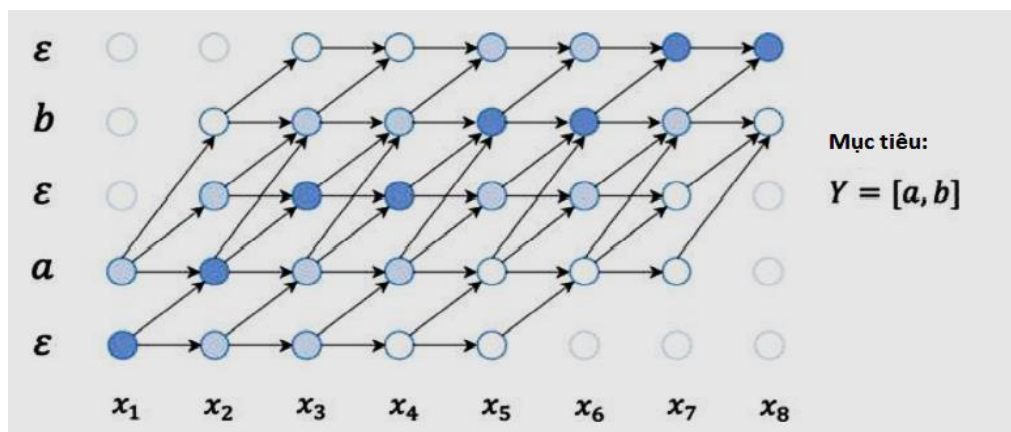
$$\begin{aligned} [h, e, l, \epsilon, l, o, \epsilon] &\rightarrow \text{“hello”} \\ [h, \epsilon, e, l, \epsilon, l, o] &\rightarrow \text{“hello”} \end{aligned}$$

Hình 6: Các cách căn chỉnh của cùng từ "hello"

Bởi vì ta có thể thêm bao nhiêu ký tự blank trước hoặc sau ký tự thực tế đều được, nên ta có thể tưởng tượng đầu ra của chúng ta đều có blank trước mỗi ký tự.

$$Y = [\epsilon, y_1, \epsilon, y_2, \dots, \epsilon, y_u] \tag{5}$$

Nhiều đoạn căn chỉnh khác nhau đều có thể đưa ra cùng một kết quả đúng, và vì thế, tất cả đoạn căn chỉnh đúng phải được xem xét. Những đoạn căn chỉnh này được dùng để tính toán xác suất xảy ra của một đoạn căn chỉnh.



Hình 7: Căn chỉnh trong CTC

Sau khi đưa vào một đoạn âm thanh đầu vào, Deep Speech sẽ xuất ra một phân phối xác suất trên bộ ký tự, và trên tất cả các cách căn chỉnh khác nhau. Phân phối xác suất $P(Y|X)$, được tính toán bằng cách tổng tất cả các cách căn chỉnh khác nhau giữa đầu vào và đầu ra, được thể hiện ở hình dưới.

Theo toán học, ta có thể định nghĩa xác suất có điều kiện của một đoạn căn chỉnh α_t , là tích của mỗi trạng thái trong chuỗi.

$$P(\alpha|X) = \prod_t^T P(\alpha_t|X) \quad (6)$$

Tất cả các đường dẫn căn chỉnh khác nhau phải được xem xét độc lập, nên ta có thể tổng tất cả xác suất của tất cả cách căn chỉnh, đưa trước xác suất có điều kiện cho một đoạn đầu vào (X, Y) .

$$P(Y|X) = \sum_{A \in A_{X,Y}} \prod_{t=1}^T P(\alpha_t|X) \quad (7)$$

trong đó $A_{X,Y}$ là tập những căn chỉnh hợp lệ. Thuật toán quy hoạch động được sử dụng để tính toán hàm lỗi CTC. Kết hợp tất cả lại, ta có hàm lỗi CTC được sử dụng để huấn luyện mô hình Deep Speech.

$$L_{CTC}(X, Y) = -\log \sum_{\alpha \in A_{X,Y}} \prod_{t=1}^T P(\alpha_t|X) \quad (8)$$

4. THỰC NGHIỆM VÀ ĐÁNH GIÁ

4.1. Phương pháp đánh giá: Word Error Rate

Trong bài toán này, nhóm sử dụng Word Error Rate (WER) [11] để đánh giá kết quả mô hình. WER là một trong những thước đo phổ biến về hiệu suất của một hệ thống nhận dạng giọng nói hoặc hệ thống dịch máy (machine translation system). WER có nguồn gốc từ khoảng cách Levenshtein, hoạt động ở cấp độ từ thay vì cấp độ âm vị. WER giúp so sánh hiệu quả của các hệ thống khác nhau hoặc cũng có thể để đánh giá một hệ thống để cải tiến hệ thống đó. WER được tính bằng công thức sau:

$$WER = \frac{S + D + I}{N} = \frac{S + D + I}{S + D + C} \quad (9)$$

Trong đó:

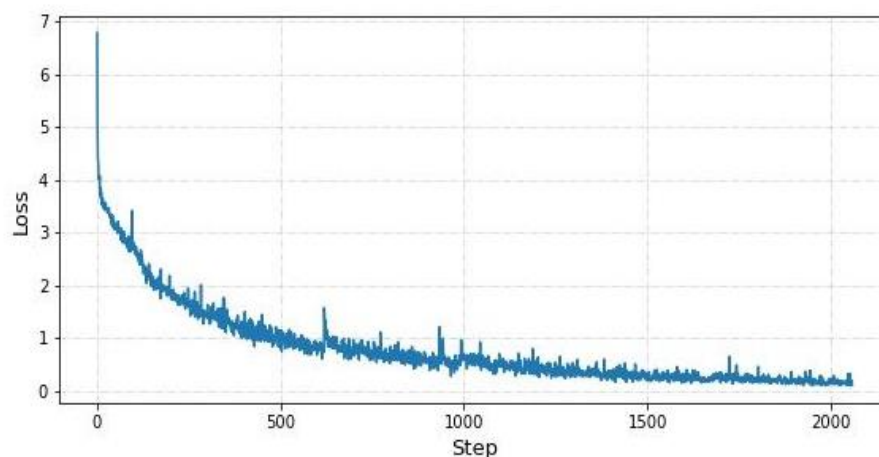
- S là số lần thay thế (substitution): xảy ra khi một từ được thay thế, ví dụ "noose" → "moose".
- D là số lần xóa (deletion): xảy ra khi một từ bị loại bỏ khỏi một đoạn văn bản, ví dụ "turn it around" → "turn around".
- I là số lần chèn (insertion): xảy ra khi một từ được chèn thêm vào mà chưa được nói.
- C là số lượng từ đúng.
- N là số lượng từ đã nói ($N = S + D + C$).

Về cơ bản có thể hiểu WER là tổng số lỗi chia cho tổng số từ.

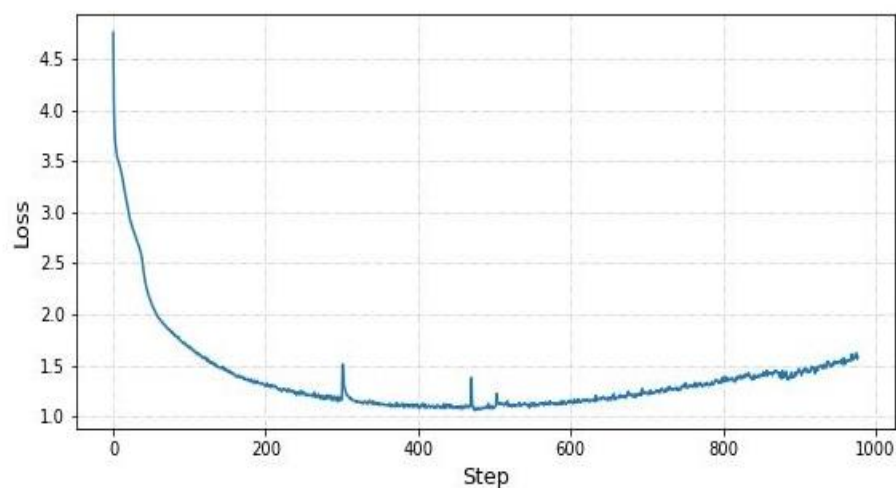
4.2. Phương pháp thực nghiệm

Nhóm đã sử dụng Kaggle làm nền tảng chính để huấn luyện mô hình Deep Speech, do hạn chế về thời gian nên mỗi phiên của Kaggle chỉ chạy được 12h là tối đa, chúng tôi tận dụng triệt để khoảng thời gian này để huấn luyện và lưu lại model để huấn luyện tiếp trong phiên sau.

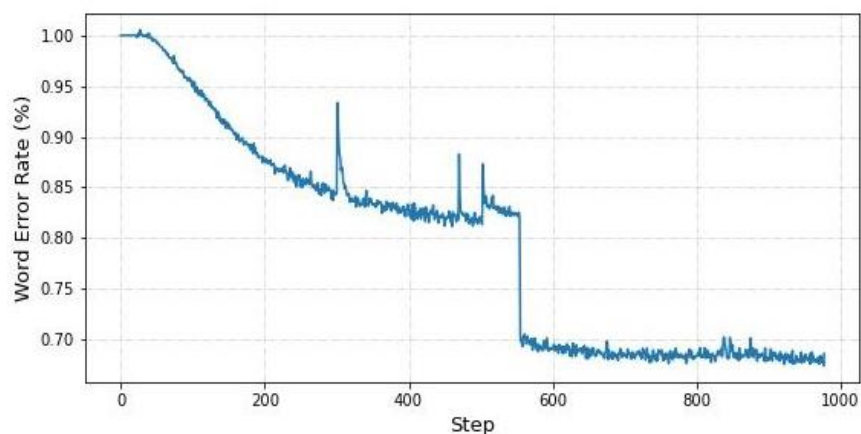
Tổng số phiên là 9, thời gian huấn luyện là 110 giờ và 58 phút, chạy trên 978 epoch tất cả, batch size là 4. Do hạn chế về lượng GPU có thể sử dụng, nên nhóm chỉ sử dụng một bộ dữ liệu VIVOS để phục vụ cho việc huấn luyện và kiểm thử.



Hình 8: Độ lỗi theo thời gian của tập huấn luyện



Hình 9: Độ lỗi theo thời gian của tập kiểm thử



Hình 10: Word Error Rate theo thời gian của tập kiểm thử

Word Error Rate trên tập test được giảm dần theo thời gian, trong quá trình huấn luyện, chúng tôi ngắt giữa chừng để điều chỉnh tham số thì được một đợt giảm WER mạnh, được mô tả trong hình. Dưới đây là một số mẫu câu được dự đoán từ mô hình.

Thực tế: nhỏ đưa tay ra cào hết bùn khô trên cổ tớ
Dự đoán: nhỏ đưa tay ra cào hết bông cỏ trên cổ tớ
Thực tế: và khi thấy tôi vẫn kiên định trả lời trót quớt
Dự đoán: và khi thấy tôi vẫn kiên định trả lời chết rhuộc
Thực tế: hằng năm chúng ta thải ra hàng tỉ thiết bị điện tử
Dự đoán: khăn năn chúng xê phễ ra hàn ty tiết bi điền tử
Thực tế: hoặc cho vào vại muối để lợn ăn dần
Dự đoán: hoặc cho vào vại mui đa lợn ăn dần
Thực tế: phương châm sống của tôi là người không vì mình trời tru đất diệt
Dự đoán: phương thân sống của tôi lđan người không qyề mạnh trợ chủ đắc duiệt
Thực tế: cho em góp một viên gạch
Dự đoán: cho em róc một viên gặt
Thực tế: cái vị chát chát và đắng ngắt của cà phê nguyên chất
Dự đoán: cây vùi chếc giác bà đóng ngắt của cà phê uyên chất
Word Error Rate đo được cuối cùng trên tập dữ liệu kiểm thử của Vivos là 68%.

5. KẾT LUẬN VÀ HƯỚNG NGHIÊN CỨU TƯƠNG LAI

Trong bài nghiên cứu này, Deep Speech đã được áp dụng trong nhận dạng giọng nói tiếng Việt và thực nghiệm trên bộ Vivos. Độ lỗi từ đạt được trên tập kiểm thử là 68% trên mô hình Deep Speech 30 triệu tham số. Kết quả này làm nền cho các nghiên cứu tương lai của nhóm, bao gồm: thu thập thêm dữ liệu, tăng lượng dữ liệu sử dụng cho phần thực nghiệm kết quả, thay đổi kiến trúc mô hình Deep Speech thành các kiến trúc khác và so sánh với các kiến trúc khác nhau để đưa ra được kiến trúc mô hình tối ưu cho nhận dạng giọng nói tiếng Việt.

TÀI LIỆU THAM KHẢO

- [1] A. G. Haudricourt, "The origin of the peculiarities of the Vietnamese alphabet," 2010.
- [2] M. Alves, "Linguistic Research on the Origins of the Vietnamese Language: An Overview," *Journal of Vietnamese Studies*, tập 1, pp. 104-130, 02-2006.
- [3] J. L. J. W. Uday Kamath, *Deep Learning for NLP and Speech Recognition*, Springer Cham, 2019.
- [4] J. a. W. J. a. E. J. P. a. F. J. a. M. R. a. Y. N. B. Novoa, "DNN-HMM based Automatic Speech Recognition for HRI Scenarios," trong *2018 13th ACM/IEEE International Conference on Human-Robot Interaction (HRI)*, 2018, pp. 150-159.
- [5] J. Li, "Recent Advances in End-to-End Automatic Speech Recognition," trong *arXiv*, 2021.
- [6] C. C. J. C. B. C. G. D. E. E. R. P. S. S. S. A. C. A. Y. N. Awni Hannun, "Deep Speech: Scaling up end-to-end speech recognition," in *arXiv*, 2014.
- [7] A. a. F. S. a. G. F. a. S. J. Graves, "Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks," *ICML 2006 - Proceedings of the 23rd International Conference on Machine Learning*, tập 2006, pp. 369-376, 01-2006.
- [8] L. D. Dong Yu, *Automatic Speech Recognition*, Springer London, 2015.
- [9] L. a. Z. Y. a. J. D. a. Z. Y. a. W. F. a. G. I. a. V. E. a. S. H. Li, "Hybrid Deep Neural Network--Hidden Markov Model (DNN-HMM) Based Speech Emotion Recognition," in *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, 2013, pp. 312-317.
- [10] A. a. J. N. a. M. A.-r. Graves, "Hybrid speech recognition with Deep Bidirectional LSTM," trong *2013 IEEE Workshop on Automatic Speech Recognition and Understanding*, 2013, pp. 273-278.
- [11] "Wikipedia," [Trực tuyến]. Available: https://en.wikipedia.org/wiki/Word_error_rate. [Đã truy cập 1 5 2022].

YSC4F.221

NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG ĐÁNH GIÁ CẢM XÚC KHÁCH HÀNG THÔNG QUA KHUÔN MẶT

NGUYỄN VĂN DUY, NGUYỄN GIA BẢO, TRẦN THANH BÌNH

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

nguyenvanduy@iuh.edu.vn, nguyengiabao.arnold@gmail.com, binhtranthanh031200@gmail.com

Tóm tắt. Biểu hiện trong giao tiếp hàng ngày của mỗi người là một phần rất quan trọng trong việc truyền tải thông tin, cảm xúc, cảm nhận hay cụ thể là sự thành công, hiệu quả của buổi giao tiếp đó hoặc đơn giản chỉ là cảm xúc của khách hàng đối với dịch vụ của nhà hàng, khách sạn... Trong bài báo này, nhóm tác giả phân tích các phương pháp nhận dạng cảm xúc khác nhau và kết quả của các biểu hiện ấy. Chúng tôi đã chọn phương pháp học sâu sử dụng mạng nơ-ron tích chập kết hợp với thuật toán “Haar Cascade” để phát hiện cảm xúc khuôn mặt và phân loại. Tập dữ liệu phổ biến về nhận dạng mặt người FER-2013 được dùng trong quá trình thực nghiệm. Nhóm tác giả đã nghiên cứu và thiết kế “Hệ thống đánh giá tự động thông qua cảm xúc trên khuôn mặt” sẽ cho biết cảm xúc của đối tượng đang được nhận diện, thống kê về ứng dụng di động. Từ đó có thể kế thừa và phát triển tùy theo lĩnh vực, dự án yêu cầu.

Từ khóa. Học sâu, nhận dạng cảm xúc, cảm xúc khuôn mặt, mạng nơ-ron tích chập.

RESEARCH AND DESIGN AUTOMATIC ASSESSMENT SYSTEM THROUGH EMOTION ON FACES

Abstract. Individual expression in daily communication is a very important part in conveying information, feelings or specifically the success and effectiveness of that communication or simply facial expression. In this paper, the authors analyze different methods of emotion recognition and the results of our research in facial expressions. We have chosen a deep learning method using convolutional neural network combined with “Haar Cascade” algorithm for facial emotion detection and classification. The popular dataset on human face recognition FER-2013 is used in the experimental process. The authors have researched and designed the "Automatic evaluation system through facial emotions" that will analyze and give the results of the emotions of the object being recognized, and the of the mobile application of this system. From there, it is possible to inherit and develop according to the required fields and projects.

Keywords. Deep learning, emotion recognition, facial emotion, convolutional neural networks.

1. GIỚI THIỆU

Mehrabian, một người nổi tiếng nhà tâm lý học người Mỹ, cho rằng trong giao tiếp hàng ngày của con người, thông tin được truyền tải bằng ngôn ngữ và giọng nói lần lượt chiếm 7% và 38% tổng lượng thông tin, trong khi thông tin được truyền tải bằng nét mặt chiếm 55%, ngoài ra còn có 6 biểu hiện cơ bản của con người quyết định tính hiệu quả và sự thành công của buổi giao tiếp đó. Vì vậy, biểu hiện trong giao tiếp hàng ngày của mỗi chúng ta là một phần rất quan trọng trong việc truyền tải thông tin. Đó là quá trình phức tạp và quan trọng, thể hiện cảm xúc, cảm nhận, tâm tư mà ta muốn gửi đến đối phương trong các buổi giao tiếp. Xử lý dữ liệu cảm xúc của con người đã và đang trở thành xu thế ở Việt Nam nói riêng và trên toàn thế giới nói chung với các ứng dụng như: rô bốt an ninh, rô bốt phục vụ, trò chơi đồ họa tương tác,... Vì lý do này, rất nhiều nghiên cứu đã được thực hiện với sự tiến bộ của khoa học công nghệ, chúng ta hoàn toàn có thể biết được chính xác cảm xúc của người sử dụng dựa trên thời gian thực, để kịp thời điều chỉnh sản phẩm sao cho phù hợp với thực tế khách quan.

Một trong những đề tài phổ biến nhất thời điểm hiện tại (2022) trong lĩnh vực thị giác máy tính đó là “Nhận dạng cảm xúc trên nét mặt”. Có rất nhiều ứng dụng dùng cảm xúc trên khuôn mặt của con người làm dữ liệu để tính toán, xử lý từ đó đưa ra các quyết định như: tương tác giữa người và máy, giám sát an ninh,... Ngoài ra, hệ thống còn được ứng dụng trong khối ngành khoa học hành vi để cung cấp các loại thông tin xã hội (tuổi tác, nguồn gốc, giới tính,...) và được cấp phép sử dụng trong khoa học về y tế nhằm theo dõi mức độ của các cơn đau, điều trị tâm thần, chậm phát triển, lo lắng và trầm cảm, ...

2. TỔNG QUAN

2.1. Mô tả quá trình

Nhóm tác giả sẽ tìm hiểu các lý thuyết như cảm xúc khuôn mặt, Convolution Neural Network trong nhận diện cảm xúc, từ đó thiết kế hệ thống và xây dựng theo đúng thiết kế. Bước cuối cùng là thực nghiệm và đánh giá kết quả đạt được.

2.2. Cơ sở lý thuyết

2.2.1. Nhận diện cảm xúc trên khuôn mặt

Đây là một bài toán phân lớp tương đối tiêu chuẩn, đã được nghiên cứu trong một thời gian khá dài. Một hệ thống nhận diện cảm xúc khuôn mặt thường được triển khai gồm 3 bước:

Bước 1: Nhận ảnh và tiền xử lý: Ảnh khuôn mặt được lấy từ nguồn dữ liệu tĩnh (chẳng hạn như từ file, database), hoặc động (từ livestream, webcam, camera,...), nguồn dữ liệu này có thể trải qua một số bước tiền xử lý nhằm tăng chất lượng hình ảnh để giúp việc phát hiện cảm xúc trở nên hiệu quả hơn.

Bước 2. Trích xuất các đặc trưng: Bước rất quan trọng, đặc biệt với các phương pháp truyền thống, các đặc trưng khuôn mặt được tính toán dựa trên các thuật toán có sẵn, kết quả thường là một vector đặc trưng làm đầu vào cho bước sau.

Bước 3. Phân lớp và nhận diện cảm xúc: Đây là một bài toán phân lớp điển hình, rất nhiều các thuật toán có thể áp dụng trong bước này như KNN, SVM, LDA, HMM,...

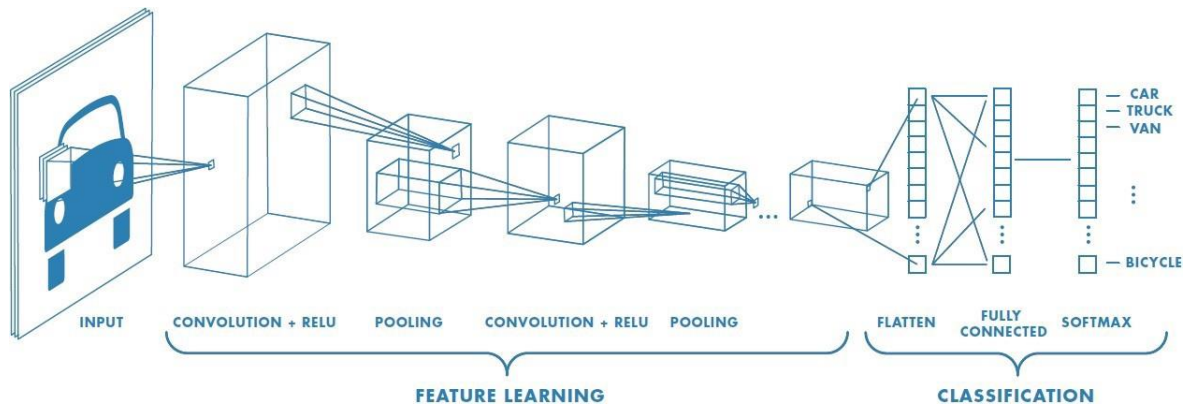
Tuy được ứng dụng rộng rãi, nhưng bài toán phát hiện cảm xúc khuôn mặt vẫn là một thách thức lớn vì độ chính xác của những hệ thống hiện nay vẫn còn khá thấp.

2.2.2. Convolution Neural Network (CNN)

Mô hình mạng nơ ron tích chập (CNN) là một trong những mô hình của của học sâu đem lại kết quả có độ chính xác cao, dùng để nhận dạng và phân loại hình ảnh, với nhận dạng đối tượng. CNN phân loại hình ảnh bằng cách lấy hình ảnh đầu vào, xử lý và phân loại theo các danh mục nhất định. Máy tính xử lý hình ảnh đầu vào như một ma trận các pixel. Và nó phụ thuộc vào độ phân giải của hình ảnh dựa trên độ phân giải của hình ảnh mà máy tính nhận thấy $H \times W \times D$ (H: chiều cao, W: chiều rộng, D: độ dày).

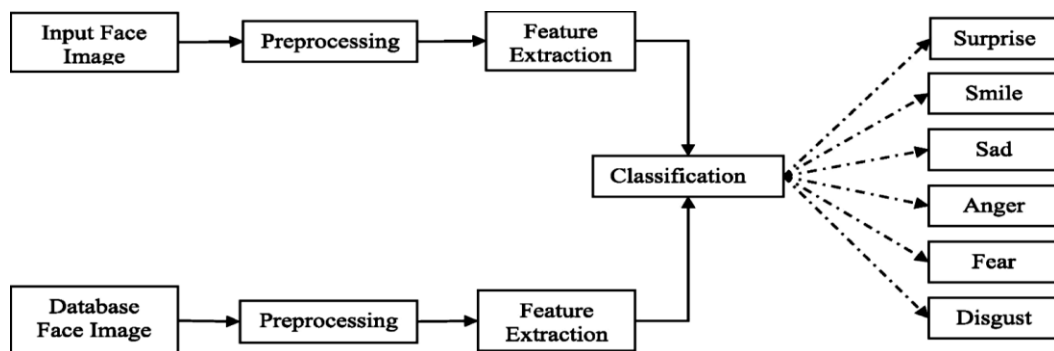
Kiến trúc cơ bản của CNN sử dụng các lớp: convolution layer + nonlinear layer, pooling layer, fully connected layer. Các lớp này liên kết với nhau theo một thứ tự nhất định.

Thông thường, một ảnh sẽ được lan truyền qua tầng convolution layer + nonlinear layer đầu tiên, sau đó các giá trị tính toán được sẽ lan truyền qua pooling layer, bộ ba convolution layer + nonlinear layer + pooling layer, có thể được lặp lại nhiều lần trong mạng. Và sau đó được lan truyền qua lớp kết nối đầy và softmax để tính xác suất ảnh đó chứa vật thể gì. Hình sau cho thấy luồng mạng nơ ron tích chập hoàn chỉnh để xử lý hình ảnh đầu vào và phân loại các đối tượng dựa trên các giá trị [5].

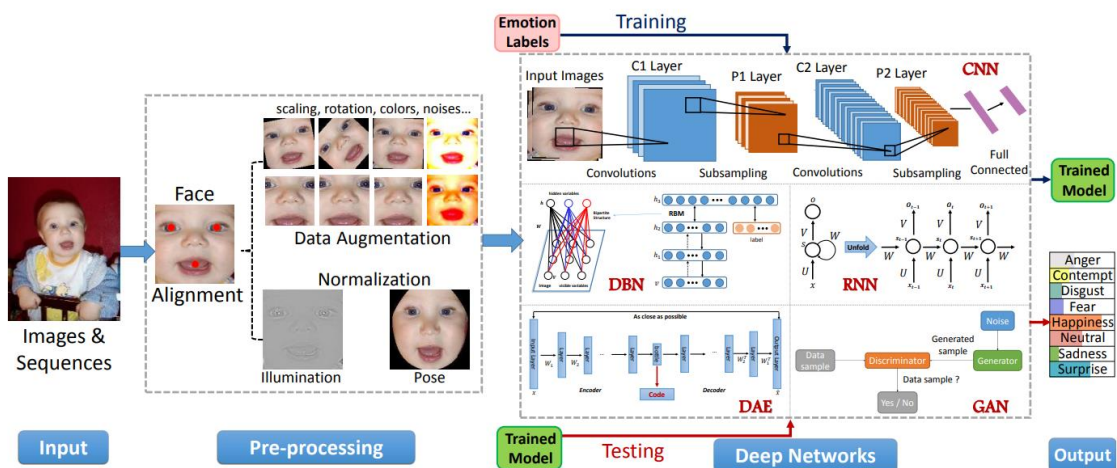


Hình 1: Xử lý hình ảnh đầu vào và phân loại các đối tượng dựa trên giá trị của mô hình mạng nơ ron tích chập

2.2.3. Phương pháp truyền thống và hiện đại



Hình 2: Một kiến trúc của hệ thống nhận dạng cảm xúc khuôn mặt bằng phương pháp truyền thống

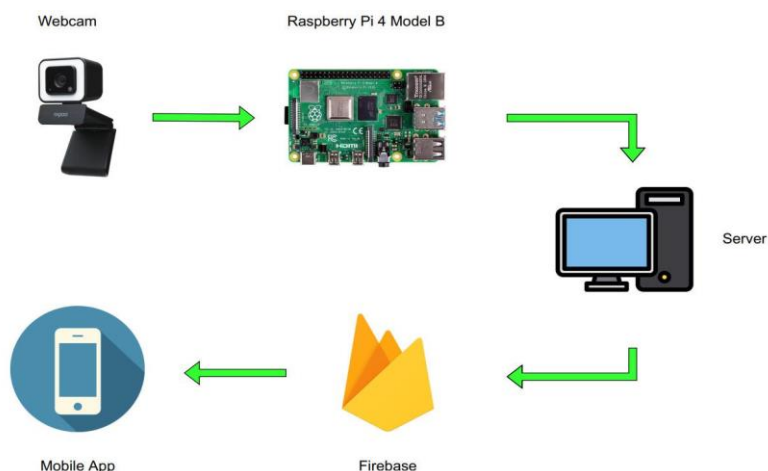


Hình 3: Kiến trúc của hệ thống nhận dạng cảm xúc khuôn mặt bằng phương pháp học sâu [5]

So với phương pháp truyền thống có 3 giai đoạn: tiền xử lý, trích xuất đặc trưng và phân loại thì phương pháp hiện đại chỉ có 2 giai đoạn: tiền xử lý và phân loại dùng học sâu (Deep learning). Trong những năm gần đây, học sâu có độ chính xác hơn phương pháp truyền thống vì nó không phải qua bước trích xuất các đặc trưng một cách tường minh, nó sẽ thực hiện song song với phân loại.

3. THIẾT KẾ VÀ PHÂN TÍCH HỆ THỐNG

3.1. Thiết kế hệ thống



Hình 4: Tổng quan hệ thống

Mô tả sơ lược về hệ thống:

Raspberry Pi 4 Model B đóng vai trò như trung tâm tiếp nhận dữ liệu: nhận hình ảnh thu thập từ Webcam Rapoo sau đó thực hiện các bước tiền xử lý ảnh, sau đó gửi cho server.

Server là bộ xử lý: phát hiện ra khuôn mặt, đưa ra kết quả cảm xúc: vui, buồn, ngạc nhiên, sợ hãi, tức giận, ghê tởm và trung tính theo thời gian thực từ ảnh nhận được từ Raspberry Pi 4 Model B. Sau đó, Server sẽ gửi kết quả là các chuỗi : “vui”, “buồn”,... đến cơ sở dữ liệu Firebase.

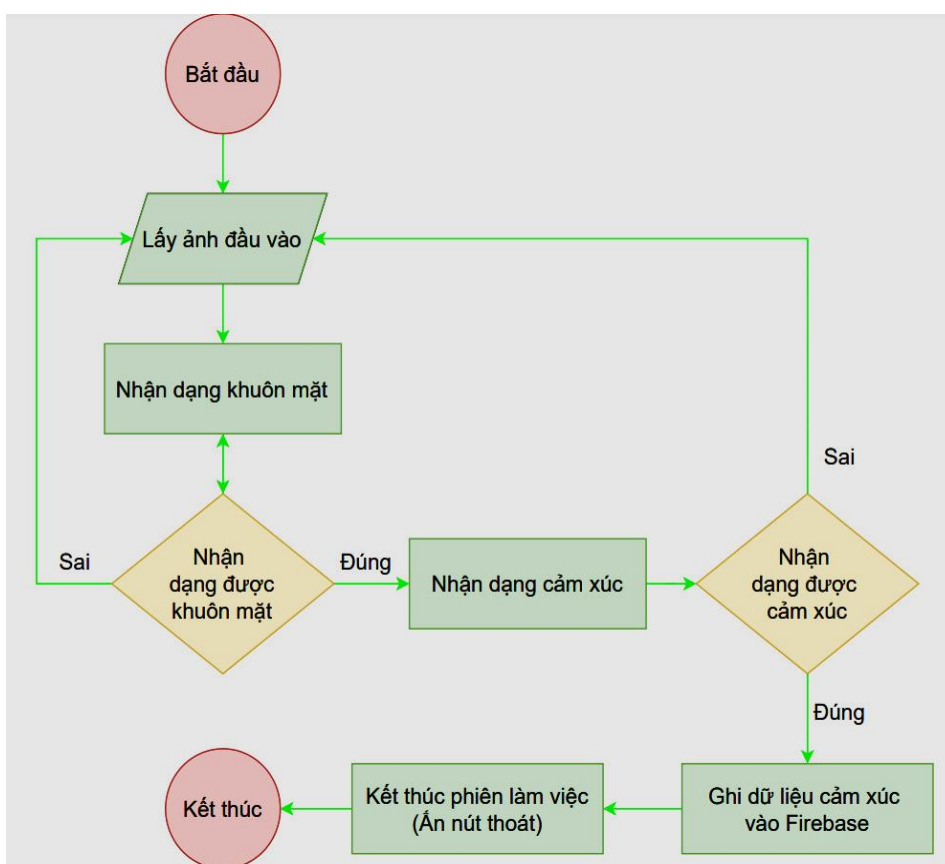
Cơ sở dữ liệu Firebase sử dụng dịch vụ Firebase Realtime Database để nhận và gửi dữ liệu từ Server đến Mobile App.

Bảng 1: Cấu hình hệ thống thử nghiệm

STT	THÀNH PHẦN	CẤU HÌNH
1	Camera	Webcam Rapoo C270L FHD 1080p
2	Raspberry Pi	RAM: 2GB Model: Pi 4 Model B
3	Server	CPU: I3 OS: Window 10 Ram: 8GB HDD: 60GB
4	Cơ sở dữ liệu	Firebase: Firebase Realtime Database
5	Điện thoại	OS: Android 7 CPU: MediaTek Helio P60 8 nhân RAM: 4GB ROM: 64GB Wifi: Wi-Fi 802.11 a/b/g/n, Wi-Fi hotspot Mạng di động: Hỗ trợ 4G

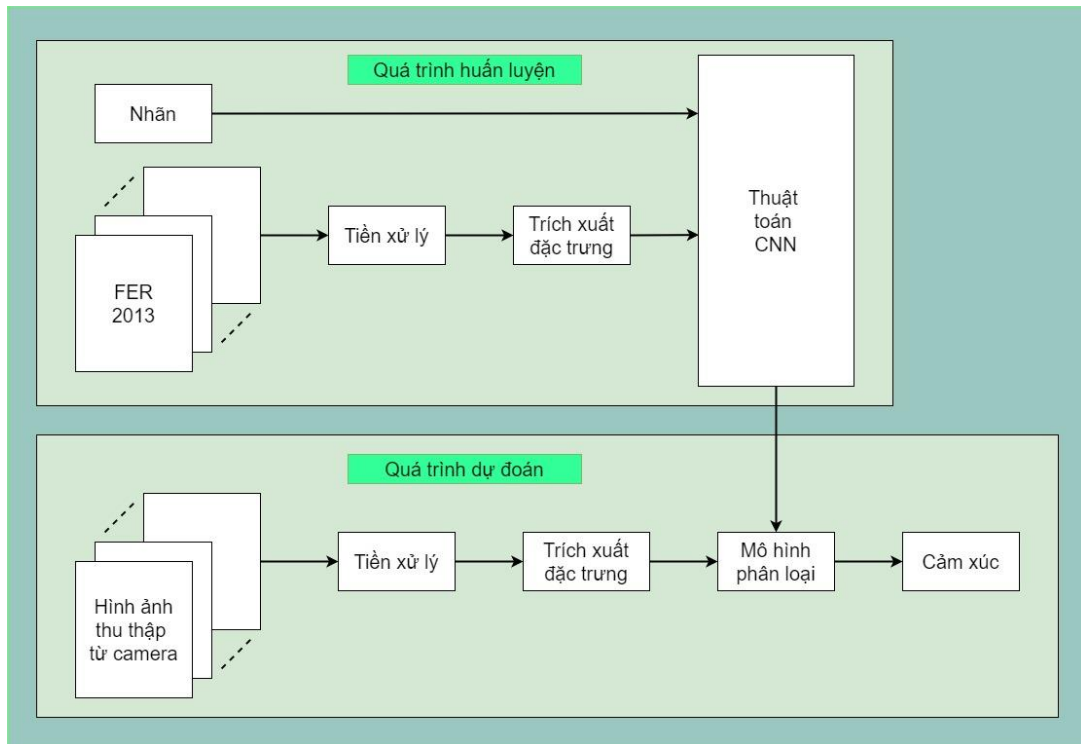


Hình 5: Kết nối phần cứng hệ thống



Hình 6: Lưu đồ giải thuật hệ thống

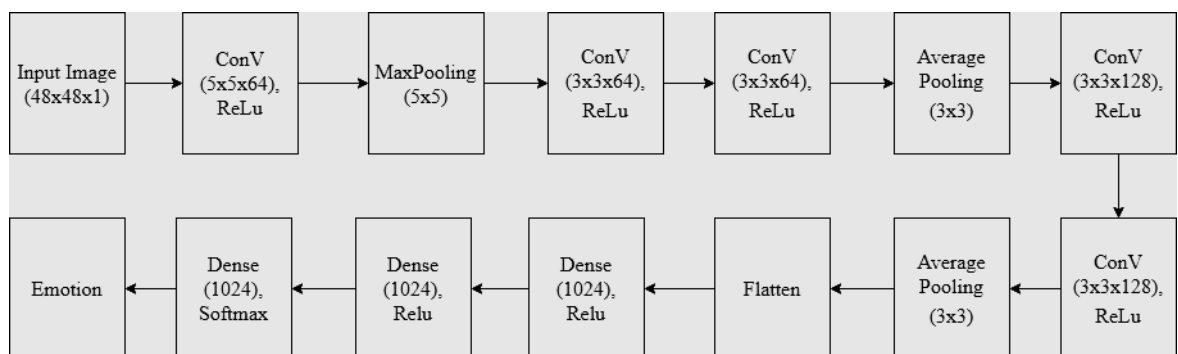
3.2. Xây dựng mô hình nhận diện cảm xúc trên khuôn mặt



Hình 7: Quy trình xây dựng mô hình nhận dạng cảm xúc khuôn mặt

Đầu vào của quá trình huấn luyện là dataset FER2013, một bộ dataset phổ biến với 35,887 grayscale ảnh khuôn mặt có kích thước 48×48 pixels. Bộ data gồm 7 loại: trung tính, tức giận, ghê tởm, sợ hãi, vui, buồn và ngạc nhiên. Quá trình này gồm 3 bước: tiền xử lý, trích xuất đặc trưng và huấn luyện sử dụng các thuật toán CNN. Đầu ra của quá trình huấn luyện là mô hình được xây dựng và các tham số tối ưu tương ứng cho mô hình.

Quá trình dự đoán đầu vào là hình ảnh thu thập được từ camera qua các giai đoạn tiền xử lý, trích xuất đặc trưng khuôn mặt, sau đó qua mô hình phân loại. Đầu ra là kết quả cảm xúc khuôn mặt.



Hình 8: Mô hình huấn luyện dựa trên CNN [1]

Dữ liệu đầu vào các ảnh cần huấn luyện sẽ được lấy từ database FER-2013, các hình ảnh đã được resize về kích thước 48x48x1.



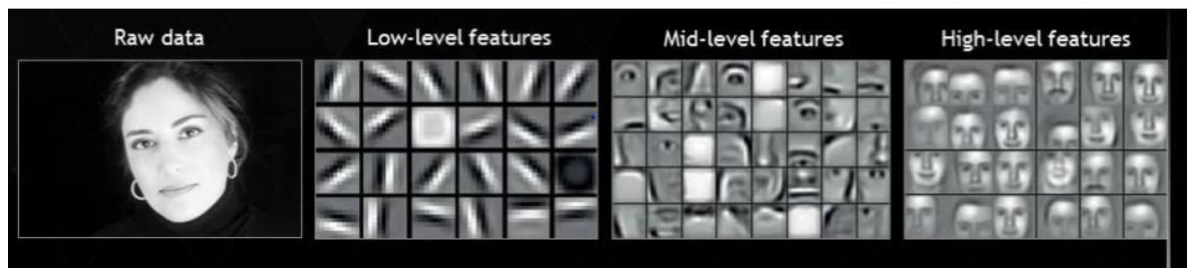
Hình 9: Dữ liệu đầu vào

Sau đó gán nhãn cho từng ảnh theo bảng sau:

Bảng 2: Dữ liệu nhận dạng khuôn mặt

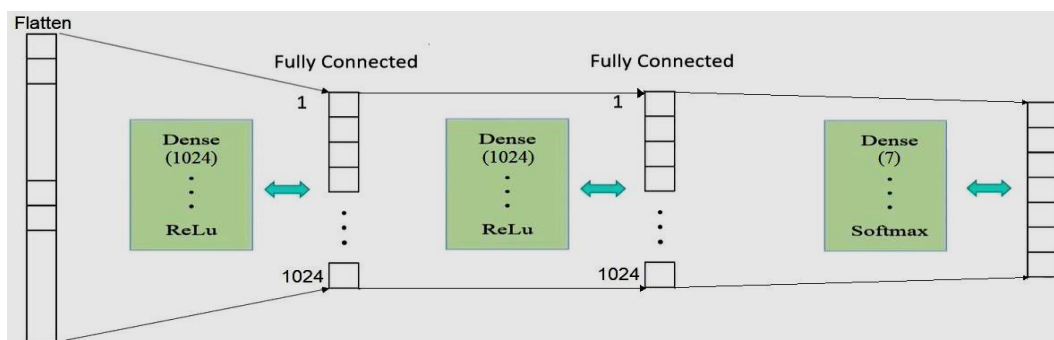
Định danh cá nhân	Ảnh đầu vào	Dữ liệu đầu ra cuối cùng của CNN
#0	Anger	[[1.0.0.0.0.0.0]]
#1	Disgust	[[0.1.0.0.0.0.0]]
#2	Fear	[[0.0.1.0.0.0.0]]
#3	Happiness	[[0.0.0.1.0.0.0]]
#4	Sadness	[[0.0.0.0.1.0.0]]
#5	Surprise	[[0.0.0.0.0.1.0]]
#6	Neutral	[[0.0.0.0.0.0.1]]

Các ảnh sau khi qua các lớp tích chập (Convolution Layer) sẽ lọc ra các đặc trưng của khuôn mặt cần nhận diện.



Hình 10: Một số đặc trưng

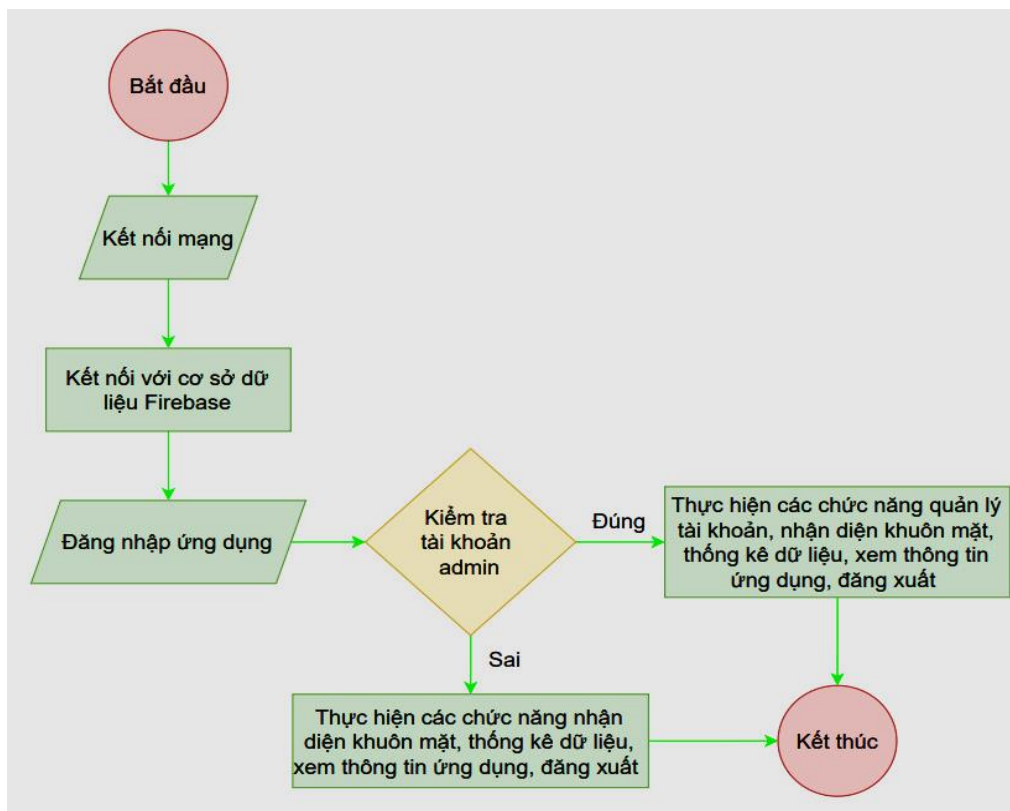
Các ảnh sau khi qua lớp gộp (Pooling Layer) sẽ giữ lại được các đặc trưng chung nhất của khuôn mặt. Cuối cùng đi qua lớp kết nối đầy đủ (Fully Connected Layer) với hàm softmax cho ra xác suất kết quả.



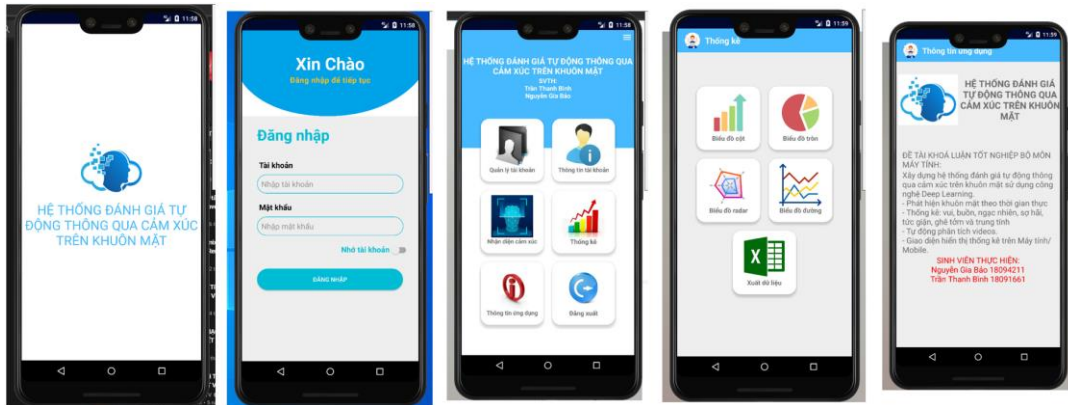
Hình 11: Lớp kết nối đầy đủ

Dữ liệu cảm xúc khuôn mặt sẽ được thống kê và đưa lên firebase để lưu trữ và cung cấp cho các của ứng dụng người dùng trên nền tảng điện thoại.

3.3. Thiết kế ứng dụng điện thoại



Hình 12: Lưu đồ giải thuật ứng dụng điện thoại



Hình 13 Giao diện ứng dụng điện thoại

4. THỰC NGHIỆM VÀ ĐÁNH GIÁ KẾT QUẢ

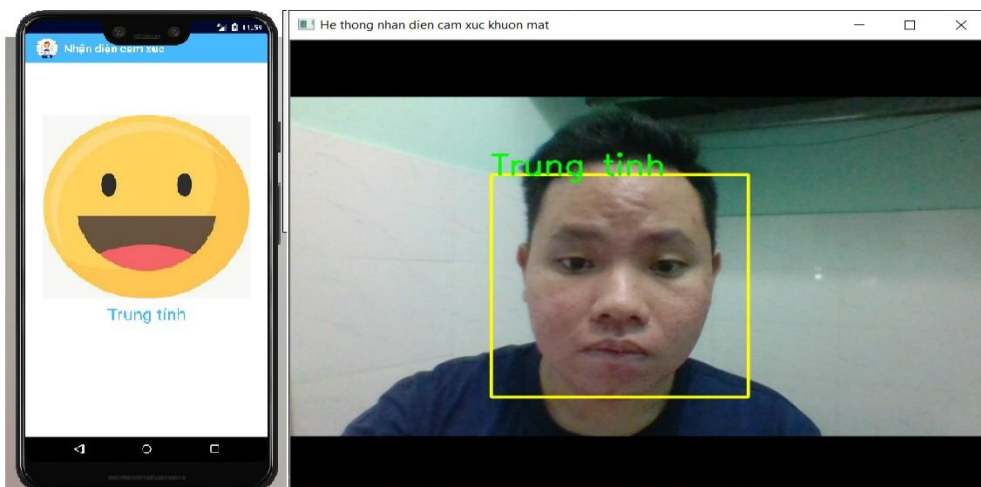
4.1. Thử nghiệm

Bảng 3: Kết quả thử nghiệm

Tác giả	Thử nghiệm	Raw	Raw + Face Landmarks	Raw + Face Landmarks + Hog	ResNet50
animeHorseman	SVM	x	46.6%	56.0%	x
animeHorseman	CNN	72.5%	73.4%	73.3%	x
Thử nghiệm của nhóm tác giả	SVM	x	67.2%	66.8%	x
Thử nghiệm của nhóm tác giả	CNN	x	52.7%	52.2%	91.7%

Bên cạnh các thử nghiệm của các tác giả "animeHorseman" và "ducmanhktkd" với các mô hình Raw, Face landmarks, Hogs,... đã cho ra các kết quả nhận diện chính xác khoảng từ 46% đến 74% (như bảng trên) là chưa tối ưu và hiệu quả.

Nhóm tác giả đã thử nghiệm với ResNet50 - một mạng thần kinh phức tạp tạo độ sâu 50 lớp, dùng để giải quyết vấn đề giảm độ dốc, sâu hơn, tăng tính chính xác, hiệu quả của model - và cho ra mức độ chính xác khoảng 91,7%.



Hình 14: Kết quả đạt được

4.2. Kết quả đạt được

- Nhận diện được cảm xúc trên khuôn mặt theo thời gian thực.
- Ứng dụng trên smartphone thống kê được dữ liệu dưới nhiều dạng biểu đồ và tệp Excel, có chức năng đăng nhập cho người dùng và chức năng “Quản lý tài khoản” (admin).

4.3. Khó khăn

- Chưa so sánh dữ liệu theo ngày, tháng hoặc chia theo từng tài khoản đăng nhập (có thể khắc phục được).
- Trong một số trường hợp thiếu sáng, khuất một phần khuôn mặt, ... sẽ không nhận diện được hoặc nhận diện sai cảm xúc.

4.4. Hướng phát triển

“Hệ thống đánh giá tự động thông qua cảm xúc trên khuôn mặt” có thể tiếp tục phát triển hoặc kế thừa vào các sản phẩm tích hợp, các hệ thống lớn hơn: Hệ thống phân tích và dự đoán cảm xúc sử dụng trong y tế để cảnh báo đến người thân và bác sĩ khi bệnh nhân có vấn đề bất thường để xử lý nhanh chóng, giúp cứu sống nhiều bệnh nhân; Hệ thống đánh giá dựa trên cảm xúc khách hàng ở các hệ thống siêu thị tiện lợi, trung tâm thương mại,... nhằm cải thiện và nâng cao chất lượng dịch vụ, phục vụ, chăm sóc khách hàng trong khối ngành thương mại, ...

5. KẾT LUẬN

Bài viết tập trung vào nghiên cứu, thiết kế và phân tích hệ thống đánh giá tự động thông qua cảm xúc trên khuôn mặt. Xây dựng thành công hệ thống dựa trên thiết kế ban đầu và chạy thử nghiệm. Mặc dù có một số khó khăn do hình ảnh bị khuất hoặc thiếu sáng dẫn đến nhận diện sai nhưng ứng dụng trên điện thoại nói riêng và cả hệ thống nói chung đã hoạt động ổn định, cho ra các kết quả, thống kê như mục tiêu đặt ra ban đầu. Ngoài ra, có thể cập nhật thêm các tính năng so sánh theo ngày, tháng, năm hoặc theo từng tài khoản đăng nhập để mô tả dữ liệu một cách trực quan hơn trong tương lai.

TÀI LIỆU THAM KHẢO

- [1] N. T. Duyên, T. X. Nam and N. T. Tùng, "MỘT MÔ HÌNH HỌC SÂU CHO PHÁT HIỆN CẢM XÚC KHUÔN MẶT" Trường ĐH Thủy lợi, 2019.
- [2] Nguyễn Văn Duy "Bài giảng: Lập trình cho thiết bị di động" Trường ĐH Công Nghiệp TPHCM, 2021.
- [3] I. Revina and W. S. Emmanuel, "A Survey on Human Face Expression Recognition Techniques," Journal of King Saud University, 2018.
- [4] D. Smirnov, "Emotion Recognition Using Facial Feature Extraction," Rowan University, 2019.
- [5] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," IEEE, 2018.
- [6] S. Albawi and T. A. Mohammed, "Understanding of a Convolutional Neural Network," Turkey, 2017.
- [7] amineHorseman. (2019, 3 27). Facial expression recognition using CNN in Tensorflow. Retrieved from <https://github.com/amineHorseman/facial-expression-recognition-using-cnn>
- [8] Developers, G. Connect to Firebase | Android Developers. Retrieved from <https://developer.android.com/studio/write/firebase>
- [9] Forogh, P. (2022). Python Firebase with Firebase Admin SDK. Retrieved from <https://www.udemy.com/course/python-firebase-with-firebase-admin-sdk/#instructor-1>

YSC4F.222

NGHIÊN CỨU VÀ XÂY DỰNG HỆ THỐNG TRỢ LÝ ẢO CỐ VẤN HỌC TẬP CHO SINH VIÊN IUH

NGUYỄN VĂN DUY, PHAN THANH NHẬT, HOÀNG NGỌC THỤY

Khoa Công nghệ điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
nguyenvanduy@iuh.edu.vn, pphanthanhnhatt@gmail.com, vim10042000@gmail.com

Tóm tắt. Ngày nay, ứng dụng AI để xây dựng các hệ thống trả lời, tư vấn tự động đang dần thay thế con người trong nhiều lĩnh vực khác nhau. Với mong muốn số hóa vấn đề về cố vấn học tập và giải đáp thắc mắc về học vụ cho sinh viên. Nhóm tác giả lựa chọn “Nghiên cứu và xây dựng hệ thống trợ lý ảo cố vấn học tập cho sinh viên IUH”. Hệ thống không chỉ tương tác với sinh viên mọi lúc mọi nơi mà còn tương tác có ngữ cảnh, đọc và hiểu được nhiều định dạng câu hỏi khác nhau: text, video, voice. Đồng thời, hệ thống cũng tự học các câu hỏi mới để trả lời chính xác. Hệ thống sử dụng công nghệ dialogflow của google (API.AI) để training và các thuật toán về xử lý ngôn ngữ tự nhiên cho Tiếng Việt để xử lý dữ liệu từ phía sinh viên. Hệ thống được xây dựng dưới dạng ứng dụng trên điện thoại thông minh, chạy được cả trên IOS và Android.

Từ khóa. Trí tuệ nhân tạo, chatbot hỗ trợ sinh viên, trợ lý ảo thông minh.

RESEARCHING AND BUILD A VIRTUAL ASSISTANT SYSTEM OF ACADEMIC ADVISORS FOR IUH STUDENTS

Abstract. Today, applying AI to build automated answering and consulting systems is gradually replacing humans in many different fields. With the desire to digitize the issue of academic advisors and answer academic questions for students. The authors selected "Research and build a virtual assistant system of academic advisors for IUH students". The system not only interacts with students anytime, anywhere, but also interacts with context, reads and understands many different question formats: text, video, voice. At the same time, the system also learns new questions to answer correctly. The system uses google's dialogflow technology (API.AI) for training and natural language processing algorithms for Vietnamese to process data from students. The system is built as an application on smartphones, running on both iOS and Android.

Keywords. AI, chatbot to support students, smart virtual assistant.

1. GIỚI THIỆU

Ở nước ta, việc giải đáp thắc mắc cho sinh viên liên quan đến học tập của bộ phận công tác sinh viên qua tin nhắn trực tuyến đang được ưa chuộng. Tuy nhiên, việc này còn thực hiện một cách thủ công và gặp nhiều khó khăn như: tốn rất nhiều thời gian và chi phí trả lương cho nhân sự chỉ để trả lời cho những câu đơn giản và giống nhau. Chính vì vậy, nhu cầu cấp thiết là cần một hệ thống điều khiển thông minh, tự động mang lại hiệu quả cao hơn và công nghệ chatbot là một sự lựa chọn hoàn hảo.

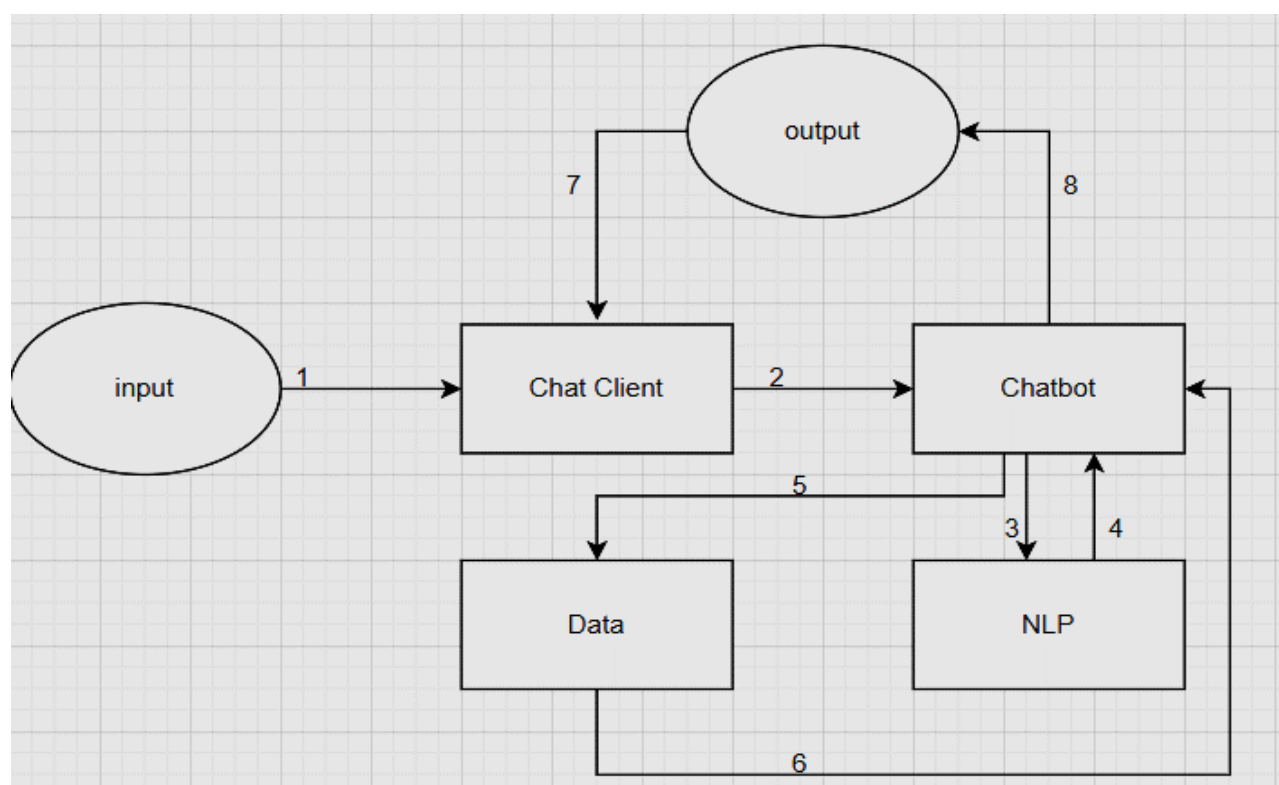
Hiện nay, thói quen sử dụng tin nhắn tức thời, chatbot và mạng xã hội đã trở thành một phần cuộc sống. Sinh viên sử dụng công nghệ không chỉ để giải trí mà còn thực hiện hầu hết các tác vụ như trao đổi thông tin đặc biệt là tài liệu học tập. Khi sinh viên có câu hỏi về buổi học quan trọng, các sự kiện sắp diễn ra, thắc mắc về học vụ, học phần, học phí thì không cần liên lạc với giáo viên và chờ đợi vài tiếng để phản hồi. Sinh viên có thể nhắn tin trực tiếp tới chatbot. Mọi thông tin có thể giải đáp ngay trong thời gian thực với một câu hỏi tới chatbot.

Trên thị trường, có nhiều ứng dụng chatbot nổi tiếng như: Zalo Chat, Facebook Messenger hay Google Chatbot đang được nhiều đơn vị sử dụng trong nhiều lĩnh vực: giáo dục, y tế, thương mại điện tử... Nhưng tất cả là các sản phẩm thương mại, để cụ thể hóa cho một đối tượng yêu cầu người dùng phải cấu hình rất phức tạp. Nhóm tác giả đưa ra ý tưởng với đối tượng cụ thể là sinh viên IUH và tập dữ liệu để huấn luyện là từ Niên giám của IUH.

2. THIẾT KẾ HỆ THỐNG

2.1. Mô hình tư vấn học vụ

Hệ thống chatbot với dữ liệu được training với bản hội thoại có sẵn để xây dựng giúp chatbot có thể trả lời câu hỏi thắc mắc của học sinh, sinh viên. Dựa vào câu hỏi của học sinh, sinh viên chatbot sẽ trả lời đúng với những câu hỏi của học sinh, sinh viên muốn tìm.



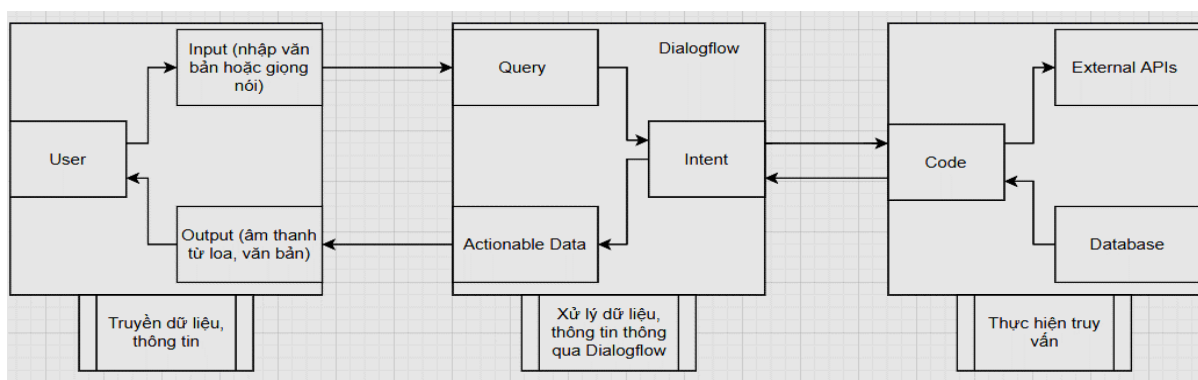
Hình 1: Mô hình tư vấn học vụ

Cơ chế hoạt động của chatbot

1. Người dùng có câu hỏi dạng văn bản cần trả lời
2. Người dùng nhập đoạn câu hỏi trên chat Client
3. Chatbot gửi đoạn câu hỏi về máy học
4. Sử dụng NLP trích xuất các thông tin cần thiết của người dùng và gửi về cho Chatbot
5. Yêu cầu dữ liệu câu trả lời từ các thông tin cần thiết đã được xử lý
6. Dữ liệu câu trả lời được trả về cho Chatbot
7. Chatbot gửi dữ liệu câu trả lời đến Chat Client
8. Chat Client hiển thị câu trả lời cho người dùng

2.2. Thiết kế sơ đồ hệ thống

Với mục tiêu xây dựng mô hình tương tác giữa người với máy thông qua hệ thống trợ lý ảo tự động để điều khiển các thiết bị điện, hệ thống được chia thành các phần chính như sau: Truyền dữ liệu, thông tin; Xử lý dữ liệu, thông tin thông qua Dialogflow [3].



Hình 2: Sơ đồ hệ thống trợ lý ảo tự động

Khối truyền dữ liệu, thông tin: Giúp người dùng có thể nhận và ghi được những dữ liệu và thông tin mà họ mong muốn. Trong đó: Khối User là người sử dụng có thể thực hiện viết ra hoặc chọn ra những từ khóa phù hợp và cụ thể để chatbot có thể nhận được thông tin trả lời một cách chính xác nhất. Khối Input giúp người sử dụng có thể nhập dữ liệu từ bàn phím hoặc giọng nói sau đó được tổng hợp lại và được gửi đến khối truy vấn dữ liệu trong Dialogflow. Khối output sẽ trả lời những thắc mắc thông qua loa, qua màn hình hoặc các tín hiệu từ các thiết bị.

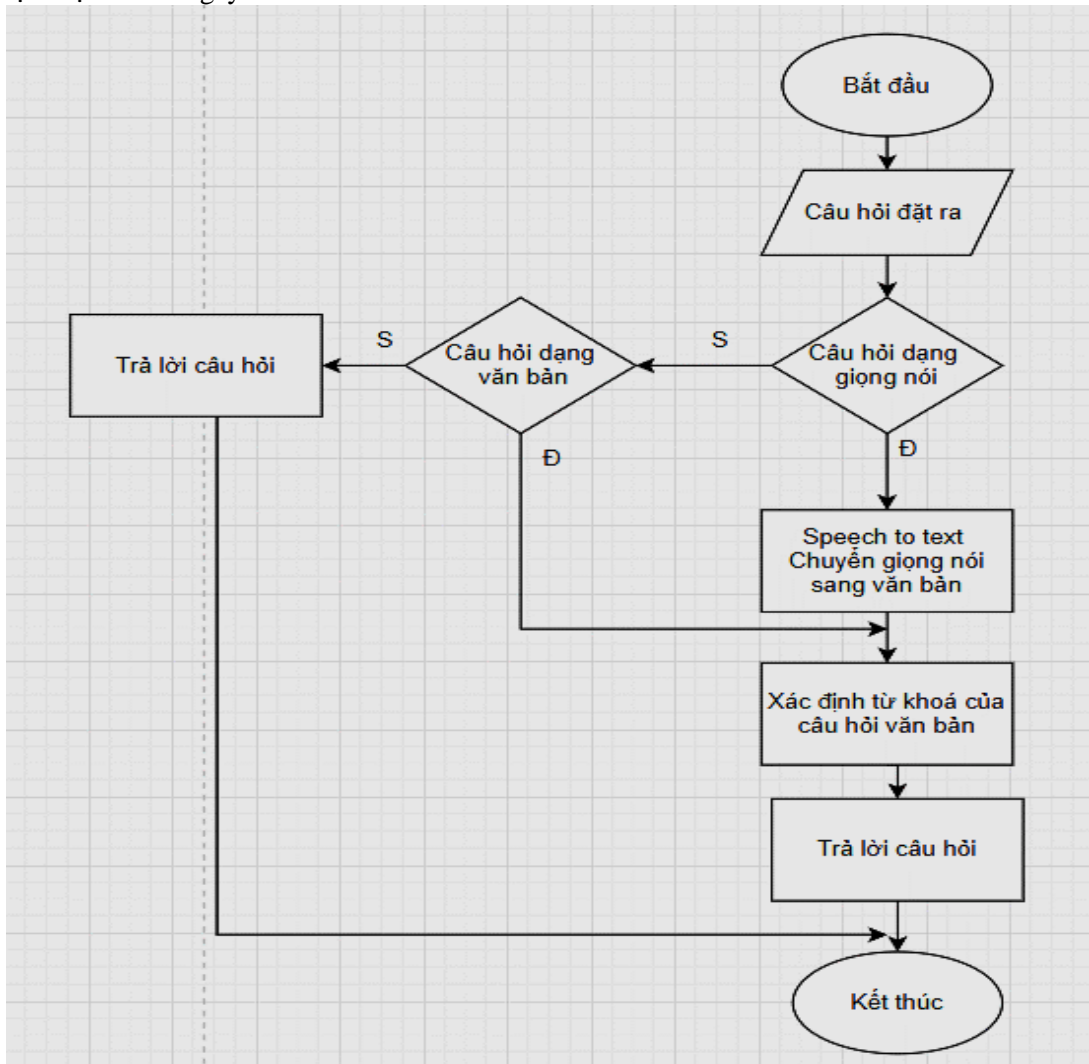
Khối xử lý dữ liệu, thông tin thông qua Dialogflow: Từ thông tin ngôn ngữ tự nhiên của người sử dụng sẽ xử lý trước khi đưa vào khối truyền dữ liệu để chuyển đổi ngôn ngữ tự nhiên thành ngôn ngữ máy. Trong đó: Khối truy vấn (Query) xử lý các yêu cầu dưới dạng ngôn ngữ tự nhiên thành ngôn ngữ máy có thể thao tác được và sử dụng các kỹ thuật học máy để giải quyết các vấn đề. Các yêu cầu truy vấn sau đó được trả về dữ liệu có cấu trúc ở định dạng JSON với một hành động xác định và các tham số cho hành động đó. Khối thao tác dữ liệu (Actionable Data) nhận các tham số, dữ liệu từ khối ý định và xử lý thành dạng dữ liệu có thể thao tác được, sau đó chuyển dữ liệu đến các thiết bị đầu ra.

Khối truy vấn gồm: Khối cơ sở dữ liệu (firebase) được sử dụng vào là firebase nhằm mục đích nhận dữ liệu và đẩy các file hình ảnh vào màn hình app để thực hiện trả lời các câu hỏi của người dùng. Khối code đồng bộ các dữ liệu trong khối ý định với cơ sở dữ liệu là firebase, sau đó đẩy dữ liệu từ Dialogflow lên màn hình nhằm đưa ra câu chào khởi đầu và những câu truy vấn được tích hợp sẵn trước thay vì đặt câu hỏi. Khi kích hoạt xong sẽ đóng khối vai trò và lấy dữ liệu từ API. Khối API khi các khối ý định cụ thể của người dùng được phân tích bởi Dialogflow, đồng bộ dữ liệu trong khối code và gửi đến API, lúc này API sẽ được sử dụng để tìm phản hồi trong các cơ sở dữ liệu bên ngoài. Khối API sẽ chuyển các dữ liệu chứa tham số đó tới máy chủ nơi chúng được xử lý. Tiếp theo, máy chủ sẽ tùy xuất dữ liệu đó, phân tích nó, thực hiện các hành động cần thiết và gửi lại dữ liệu cho hệ thống chúng ta. Sau đó, ứng dụng sẽ diễn giải dữ liệu này và hiển thị cho chúng ta thông tin chúng ta muốn.

2.3. Thiết kế ứng dụng

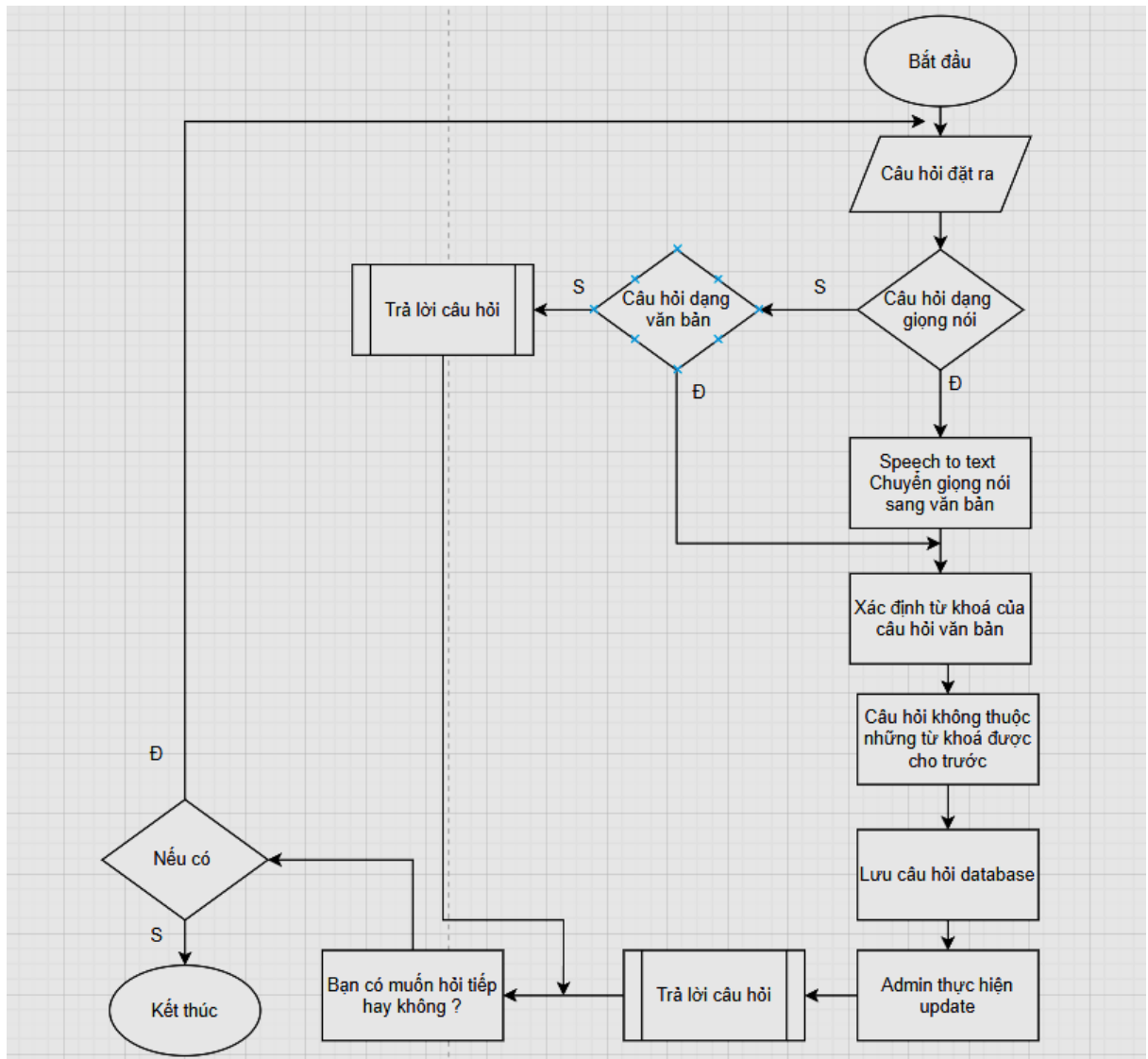
Hệ thống được xây dựng trên framework React Native với ngôn ngữ lập trình là javascript, Framework React Native được sử dụng do sự linh động và chạy trên đa nền tảng.

Khởi thực hiện trả lời ngay vào câu hỏi:



Lưu đồ 1: Hệ thống chatbot với câu hỏi tích hợp sẵn

+ Câu hỏi không có sẵn, rủi ro:



Lưu đồ 2: Lưu đồ giải thuật với câu hỏi chưa được tích hợp

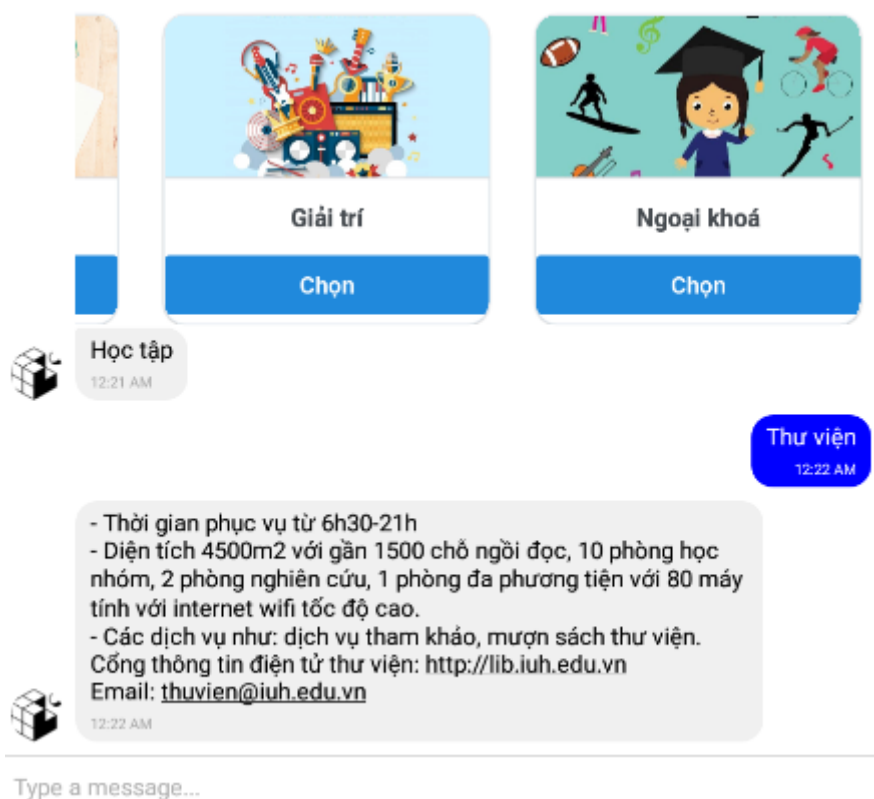
3. KẾT QUẢ VÀ ĐÁNH GIÁ THỰC NGHIỆM

3.1. Thực nghiệm

Bảng 1: Dữ liệu thực nghiệm

STT	Câu hỏi	Số lượng câu hỏi	Từ khóa	Trả lời đúng (%)
1	Thông tin về trường, review trường.	10	Review, thông tin trường	90%
2	Câu hỏi về điểm, đóng học phí	10	kế toán	80%
3	Địa chỉ, cơ sở của trường	10	Địa chỉ trường, cơ sở trường	90%

Nhiều câu hỏi của sinh viên, hệ thống chưa được thêm dữ liệu nên khi huấn luyện không trả lời được đúng với câu hỏi của sinh viên. Khả năng tương tác của chatbot phản hồi nhanh cũng như ghi nhận được thời gian tương tác với chatbot.



Hình 3: Kết quả đạt được

3.2. Kết quả đạt được

Các câu hỏi, câu trả lời từ khóa đều được lưu trữ dữ liệu trên dialogflow để dàng thay đổi được các câu hỏi, câu trả lời, từ khóa theo ý muốn của người sử dụng.

Ứng dụng dễ dàng cho sinh viên sử dụng, cũng như có thể trả lời đúng các câu hỏi thắc mắc của sinh viên, lưu được thời gian trò chuyện của sinh viên.

Khó khăn:

- Chưa có thể chuyển thành giọng nói được, chỉ có thể viết bằng văn bản.
- Dữ liệu thực nghiệm còn ít nên độ chính xác của hệ thống chưa đảm bảo.
- Trong một số trường hợp các câu hỏi, từ khóa không được lưu trong dữ liệu dialogflow thì chatbot không hiểu được và không ghi nhận lại thông tin để cập nhập lên hệ thống.

Hướng phát triển:

Nhóm tác giả tiếp tục nghiên cứu “HỆ THỐNG XÂY DỰNG TRỢ LÝ ẢO CỔ VẤN HỌC TẬP CHO SINH VIÊN IUH” về khả năng nhận diện dữ liệu đầu vào dạng text, dạng media. Đồng thời cải tiến các thuật toán cho xử lý nhanh hơn, thực nghiệm trên thực tế để hoàn thiện.

4. KẾT LUẬN

Chatbot có thể ứng dụng cho nhiều trường khác nhau, không chỉ ứng dụng trong học tập chatbot còn có thể ứng dụng trong doanh nghiệp, y tế và mọi lĩnh vực khác nhau. Trong bài báo này, nhóm tác giả đưa ra mô hình ứng dụng cụ thể. Sản phẩm đang được nhóm tác giả tiếp tục cải thiện và hoàn thiện hơn trong thời gian tới.

TÀI LIỆU THAM KHẢO

- [1] Bashmakov, P. (2018). *Advanced Natural Language Processing Tools*. Retrieved from <https://www.intellectsoft.net/blog/advanced-natural-language-processing-tools-for-bot-makers/>
- [2] ddhuy. (n.d.). *Khái quát Text to speech*. Retrieved from viblo.asia: <https://viblo.asia/p/khai-quat-ve-text-to-speech-ORNZqdMGK0n>
- [3] Dialogflow. (2018). *Dialogflow*. Retrieved from <https://cloud.google.com/dialogflow/docs/>
- [4] Google.com. (n.d.). "Speech-to-Text,". Retrieved from Google: <https://cloud.google.com/speech-to-text>.
- [5] Hoang, N. V. (2020, 9 20). *Giới thiệu Google Speech to Text API dịch một audio file sang text trên nền tảng Nodejs*. Retrieved from viblo.asia: <https://viblo.asia/p/gioi-thieu-google-speech-to-text-api-dich-mot-audio-file-sang-text-tren-nen-tang-nodejs-jvElayBDlkw>
- [6] Nam, D. V. (2017). *Dialogflow – Các khái niệm*. Retrieved from [dialogflowvn.wordpress.com:https://dialogflowvn.wordpress.com/2017/11/18/dialogflow-cac-khai-niem/](https://dialogflowvn.wordpress.com/2017/11/18/dialogflow-cac-khai-niem/)
- [7] Tiến, N. T. (2019). *Nghiên cứu và xây dựng chatbot hỗ trợ người dùng trong ngân hàng*. Retrieved from tailieu.vn: <https://tailieu.vn/doc/luan-van-thac-si-cong-nghe-thong-tin-nghien-cuu-va-xay-dung-chatbot-ho-tro-nguoi-dung-trong-ngan-ha-2279448.html>

YSC4F.223

XÂY DỰNG HỆ THỐNG CHUÔNG CỬA THÔNG MINH KHÔNG TIẾP XÚC

NGUYỄN VĂN DUY, PHẠM QUỐC HOÀNG

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

nguyenvanduy@iuh.edu.vn, phamquochuang630@gmail.com

Tóm tắt. Chuông cửa là nơi chứa đựng nhiều vi khuẩn có thể lây nhiễm phổ biến nhất trong nhà, bệnh viện, nhà máy và đặc biệt là những ngôi nhà cũ. Vì vậy, yêu cầu về các sản phẩm chuông cửa thông minh không tiếp xúc là xu hướng ngày càng tăng, không chỉ tại Việt Nam mà cả trên thế giới. Để đáp ứng nhu cầu đó, nhóm tác giả trình bày “*Xây dựng hệ thống chuông cửa thông minh không tiếp xúc*” trong bài báo này. Một hệ thống chuông cửa cần kết hợp các yếu tố: giám sát, nhận diện và cảnh báo sớm thời gian thực. Để thay thế giải pháp tiếp xúc vào chuông cửa bằng giải pháp không tiếp xúc, nhóm tác giả sử dụng các thuật toán AI trong nhận diện khuôn mặt, xử lý thân nhiệt và âm thanh để xác nhận có phải người quen hay không. Thông qua nhận diện khuôn mặt, hệ thống sẽ đưa ra lời cảnh báo hoặc thông báo cho chủ nhà. Hệ thống được xây dựng gồm các khối phần cứng: Camera, Raspberry Pi, Cảm biến thân nhiệt, Cảm biến siêu âm. Kết hợp với các thuật toán AI để xử lý dữ liệu vào ra theo thời gian thực. Ngôn ngữ lập trình hệ thống: Python, Dart.

Từ khoá. chuông cửa không tiếp xúc, chuông cửa thông minh, Python, Dart, AI

BUILD CONTACTLESS SMART DOORBELL SYSTEM

Abstract. Doorbells are home to many of the most common infectious bacteria found in homes, hospitals, factories, and older homes. Therefore, the demand for contactless smart doorbell products is an increasing trend not only in Vietnam but also in the world. To meet that need, the authors present "Building contactless smart doorbell system" in this paper. A doorbell system needs to combine elements: real-time monitoring, identification and early warning. To replace the contact solution on the doorbell with a contactless solution, the authors use AI algorithms in face recognition, body temperature and sound processing to confirm whether it is an acquaintance or not. Through facial recognition, the system will issue an alert or notify the homeowner. The system is built with hardware blocks: Camera, Raspberry Pi, Body temperature sensor, Ultrasonic sensor. Combined with AI algorithms to process input and output data in real time. System programming languages: Python, Dart.

Keywords. contactless doorbell, smart doorbell, Python, Dart, AI

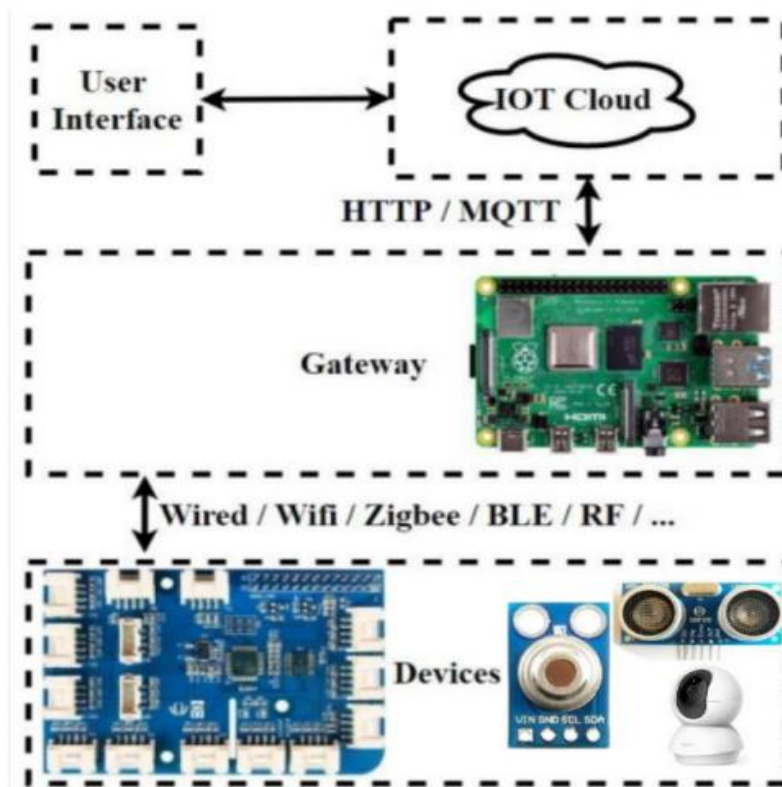
1. ĐẶT VẤN ĐỀ

Hiện nay, chuông cửa được sử dụng rộng rãi không chỉ tại nhà riêng mà cả các phòng làm việc, nơi công cộng. Đây là vị trí có nhiều tiếp xúc với nhiều đối tượng khác nhau, dẫn đến sẽ là nơi chứa đựng nhiều vi khuẩn có thể lây nhiễm phổ biến nhất trong nhà, bệnh viện, nhà máy và đặc biệt ở những ngôi nhà cũ. Vì vậy, yêu cầu về các sản phẩm chuông cửa thông minh không tiếp xúc là xu hướng ngày càng tăng không chỉ tại Việt Nam mà cả trên thế giới. Để đáp ứng nhu cầu đó, nhóm tác giả trình bày “*Xây dựng hệ thống chuông cửa thông minh không tiếp xúc*” trong bài báo này. Một hệ thống chuông cửa cần kết hợp các yếu tố: giám sát, nhận diện và cảnh báo sớm thời gian thực. Để thay thế giải pháp tiếp xúc vào chuông cửa bằng giải pháp không tiếp xúc, nhóm tác giả sử dụng các thuật toán AI trong nhận diện khuôn mặt, xử lý thân nhiệt và âm thanh để xác nhận có phải người quen hay không. Thông qua nhận diện khuôn mặt, hệ thống sẽ đưa ra lời cảnh báo hoặc thông báo cho chủ nhà. Hệ thống được xây dựng gồm các khối phần cứng:

Camera, Raspberry Pi, Cảm biến thân nhiệt, Cảm biến siêu âm. Sản phẩm sẽ giúp an toàn cho người sử dụng và chủ nhà.

2. THIẾT KẾ HỆ THỐNG

2.1. Tổng quan hệ thống



Hình 1: Tổng quan mô hình

Hệ thống sử dụng module phần cứng là Raspberry đóng vai trò là khối xử lý trung tâm để tiếp nhận dữ liệu đầu vào: nhận tín hiệu từ cảm biến, từ camera. Sau đó, khối này sẽ gửi dữ liệu nhận được từ camera đến khối xử lý ảnh.

Khối xử lý ảnh ta phát hiện khuôn mặt và nhận diện được người đó có phải khách quen không theo thời gian thực. Thông qua đó ta gửi về thông tin phân tích được về Firebase.

Cơ sở dữ liệu Firebase thông qua dịch vụ Firebase Realtime Database để nhận và gửi dữ liệu đến Mobile App.

2.2. Nhận diện người cần nhấn chuông cửa

Hệ thống nhận diện người nhấn chuông cửa bằng các bước sau:

- + Phát hiện xem có người đứng trước cửa: dùng cảm biến thân nhiệt để đo nhiệt độ xem đồ vật đứng trước cửa có phải con người không.
- + Phát hiện xem người đứng trước cửa có ý định bấm chuông cửa: dùng cảm biến siêu âm xác định khoảng cách giữa người với chuông cửa để xác định người đến có ý định bấm chuông cửa.
- + Do thân nhiệt của con người dao động trong khoảng 36,44 độ C vào 6 giờ sáng và tới 36,84 độ C vào lúc 6 giờ tối. Lúc đó, ta có thể xác định vật đứng trước cửa là con người và thông qua đó ta có thể kích hoạt cảm biến siêu âm để xác định người đó có phải khách đến chơi bằng cách đo khoảng cách từ người đó đến cửa (khoảng 50 cm).

Bằng cách lặp lại quá trình trên ta có thể xác nhận có vị khách đến chơi nhà và kích hoạt hệ thống chuông cửa hoặc cảnh báo.

2.3. Nhận diện khách quen

Hệ thống nhận diện khách quen thông qua việc xử lý ảnh bằng các bước như sau:

- + Phát hiện khuôn mặt: căn cứ theo vị trí khuôn mặt trong bức ảnh.
- + Căn chỉnh khuôn mặt: thuật toán sẽ tìm các điểm mốc rồi dựa vào điểm này để chỉnh sửa ảnh (xoay, tịnh tiến, thu nhỏ) sao cho ảnh khuôn mặt nhìn chính diện nhất có thể.
- + Biểu diễn: biểu diễn khuôn mặt thành các số đo dùng để xác minh khách quen.
- + Xác minh: Dựa vào số đo và so sánh với số đo được lưu trữ để xác định khách quen.

Camera thu được một bức ảnh có kích thước 480x640x3, nó sẽ được xác định vị trí của khuôn mặt trong bức ảnh và kích thước của khuôn mặt rồi ta cắt bức ảnh khuôn mặt ra rồi chỉnh sửa nó về kích thước 96x96x3.

Khuôn mặt đó sẽ tiếp tục được xử lý lại để ta có được các điểm mốc đặc biệt trong khuôn mặt (góc mắt, đỉnh mũi hay khoé miệng).

Từ những điểm mốc đặc biệt đó ta phải biểu diễn bức ảnh thành các tính chất để giảm sự biến thiên trong bức ảnh của một người trong khi đó thì phải tăng sự biến thiên đối với bức ảnh của hai người khác nhau (quá trình này cực kỳ quan trọng trong quá trình xác minh khuôn mặt). Từ đó ta có thể kết luận hai khuôn mặt này cùng một người.

2.4. Phát hiện khuôn mặt.

Trong ngôn ngữ lập trình Python các bức ảnh được biểu diễn ở dạng numpy array 3 chiều theo thứ tự chiều dọc, chiều ngang và chiều sâu. Ví dụ một mảng numpy array như sau:

$$I = [[1, 0, 0]], [0, 1, 2]]$$

Trong mảng này, các số ở giữa mỗi cặp ngoặc vuông bên trong là các phần tử của một hàng, dấu phẩy của các cặp ngoặc vuông phân tách các hàng với nhau. Việc đánh số trong các hàng và cột được bắt đầu bằng số 0 và trả về số 1. Để lấy phần tử ở giá trị một chiều nào đó, ta dùng dấu “.” ở chiều khác.

Để xử lý dữ liệu nhanh nhất và có độ chính xác cao ta có thể sử dụng thuật toán Viola-Jones được cài đặt trong OpenCV.

2.5. Căn chỉnh khuôn mặt

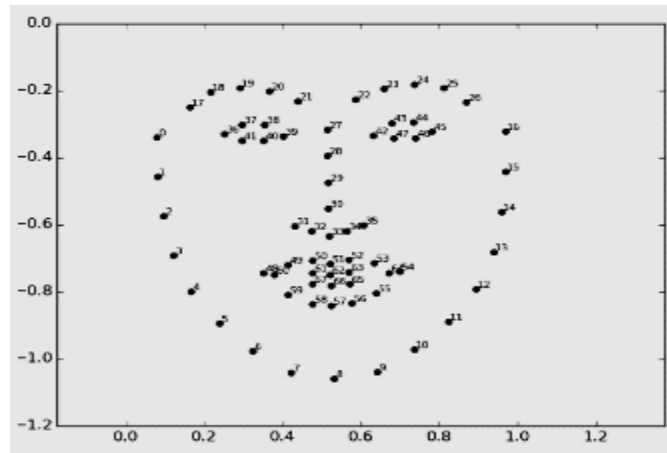
Để căn chỉnh khuôn mặt ta phải tìm được 68 điểm landmarks trên khuôn mặt bằng cách sử dụng mã nguồn mở Dlib trong thư viện xử lý ảnh. Sau khi phát hiện các điểm landmarks, ta có thể dựa vào các điểm này để biến đổi khuôn mặt sao cho góc nhìn được chính diện nhất. Ở đây ta sẽ sử dụng phép biến đổi affine. Cụ thể là tìm một ma trận H, ma trận này biến đổi các điểm có tọa độ $(x_i, y_i, 1)$ trong bức ảnh gốc thành các điểm có tọa độ (x'_i, y'_i) [3] trong bức ảnh mới sao cho:

$$\begin{bmatrix} x'_i \\ y'_i \end{bmatrix} = H \begin{bmatrix} x_i \\ y_i \\ 1 \end{bmatrix} \quad (1)$$

Để tìm được H, ta cần dựa vào một tính chất của phép biến đổi affine như sau: với hai hình tam giác bất kỳ, ta luôn tìm được một phép biến đổi affine để biến đổi tam giác này thành tam giác kia [5]. Dựa vào tính chất này, ta sẽ chọn ra ba điểm, tìm ma trận H sao cho với ma trận H, ba điểm landmarks được chọn trong bức ảnh gốc được biến đổi thành ba điểm landmarks chuẩn tương ứng. Hình 1 bên dưới minh họa 68 điểm landmarks chuẩn mà bài báo này sử dụng [1].

Gọi A, B, C là ba điểm landmarks được chọn trong bức ảnh gốc và A', B', C' là ba điểm tương ứng của landmarks chuẩn, từ phương trình (1) ta có:

$$\begin{bmatrix} x_{A'} \\ y_{A'} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{bmatrix} \begin{bmatrix} x_A \\ y_A \\ 1 \end{bmatrix} \quad (2)$$



Hình 2: Các điểm landmarks chuẩn bị sử dụng

$$\begin{bmatrix} x_{B'} \\ y_{B'} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{bmatrix} \begin{bmatrix} x_B \\ y_B \\ 1 \end{bmatrix} \quad (3)$$

$$\begin{bmatrix} x_{C'} \\ y_{C'} \end{bmatrix} = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{bmatrix} \begin{bmatrix} x_A \\ y_A \\ 1 \end{bmatrix} \quad (4)$$

Giải hệ phương trình ta tìm ra H:

$$\begin{bmatrix} H_{11} \\ H_{12} \\ H_{13} \\ H_{21} \\ H_{22} \\ H_{23} \end{bmatrix} = \begin{bmatrix} x_A & y_A & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_A & y_A & 1 \\ x_B & y_B & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_B & y_B & 1 \\ x_C & y_C & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & x_C & y_C & 1 \end{bmatrix}^{-1} \begin{bmatrix} x_{A'} \\ y_{A'} \\ x_{B'} \\ y_{B'} \\ x_{C'} \\ y_{C'} \end{bmatrix} \quad (5)$$

Ma trận biến đổi cuối cùng:

$$H = \begin{bmatrix} H_{11} & H_{12} & H_{13} \\ H_{21} & H_{22} & H_{23} \end{bmatrix} \quad (6)$$

Các điểm landmarks số 39, 42, 33 được sử dụng để biến đổi khuôn mặt.

2.6. Biểu diễn khuôn mặt

Để so sánh hai khuôn mặt bài báo đã sử dụng mạng neurou đã được luyện sẵn của thư viện OpenFace. Mạng neuron có input là một bức ảnh màu RGB với kích thước 96x96x3 và output là một vector chứa 128 số đo (được gọi là embeddings). Khoảng cách embedding giữa hai bức ảnh sẽ lớn nếu chúng thuộc về hai người khác nhau và ngược lại.

Gọi x là một bức ảnh đầu vào và $f(x)$ là vector chứa các embeddings của x . Mạng này được huấn luyện sao cho:

$$\|f(x_i^a) - f(x_i^p)\|_2^2 + \alpha < \|f(x_i^a) - f(x_i^n)\|_2^2 \quad (7)$$

Trong đó x_i^a là một bức ảnh đối chiếu, x_i^p là một bức ảnh khác của cùng một người với x_i^a và x_i^n là ảnh khác của một người khác khác với x_i^a và x_i^p . Từ bất đẳng thức (7) ta thấy rằng khoảng cách của các embeddings hai khuôn mặt của hai người khác nhau luôn lớn hơn khoảng cách của các embeddings của hai khuôn mặt cùng một người ít nhất là α (với mạng neuron của OpenFace, $\alpha = 0,2$).

2.7. Xác minh khuôn mặt

Như vậy khoảng cách Ruclidean của hai khuôn mặt của hai người luôn lớn hơn α , nên khi lập trình cần tính khoảng cách giữa các khuôn mặt thu được từ camera, nếu nhỏ hơn 0,2 thì kết luận khuôn mặt của cùng một người và ngược lại. Tuy nhiên để bảo đảm tính an toàn ta nên chọn ngưỡng thấp hơn 0,2, trong bài này tôi chọn ngưỡng 0,18.

2.8. Đọc ảnh từ sever

Để đọc ảnh từ server trả về, trước hết cần phải kết nối với server. Server trong bài báo này trả về ảnh dạng JPEG; đây là chuẩn nén ảnh phổ biến trên internet. Nếu server chỉ trả về một bức ảnh thì ta chỉ cần giải mã chuỗi dữ liệu nhận được trong biến stream là đã được một bức ảnh. Tuy nhiên, trong bài báo này, server liên tục trả ảnh về nên cần phải có cách để đọc từng chuỗi mã hóa ảnh trong stream ra. Để tìm được chuỗi mã hóa ảnh, ta cần tìm các chuỗi '\xff\xd8' và '\xff\xd9'; đây là các chuỗi mở đầu và kết thúc được quy ước trong định dạng ảnh JPEG. Sau khi đọc được chuỗi mã hóa, ta sẽ xóa các giá trị từ '\xd9' trở về trước. Quá trình đọc ảnh tiếp tục được lặp lại để đọc các ảnh tiếp theo.

3. GIẢI PHÁP TRUYỀN THÔNG

3.1. Truyền hình ảnh qua mạng LAN

Mạng LAN (Local Area Network) là tên của một loại mạng cục bộ. Các thiết bị kết nối trong mạng này có thể chia sẻ thông tin, dữ liệu với nhau mà không bị can thiệp, tấn công từ các thiết bị ngoài mạng. Việc truyền hình ảnh qua mạng LAN giúp các thiết bị trong mạng có thể xem được hình ảnh từ camera ở cửa gửi về. Để truyền được hình ảnh qua mạng LAN trên máy tính Raspberry Pi (sử dụng hệ điều hành Debian) chúng tôi sử dụng một dịch vụ (service) đi kèm, đó là motion.

- Để cài đặt dịch vụ, chúng ta nhập vào terminal câu lệnh: `sudo apt-get install motion`.
- Để bắt đầu sử dụng dịch vụ chúng ta gõ vào terminal câu lệnh: `sudo service motion start`.
- Để dừng dịch vụ sử dụng câu lệnh: `sudo service motion stop`.

Nếu việc truyền hình ảnh không như ý muốn, chúng ta có thể thay đổi các thông số trong tập tin cấu hình của dịch vụ bằng câu lệnh: `sudo nano/etc/motion/motion.conf`, lúc đó tập cấu hình được mở ra và chúng ta chỉnh sửa lại dựa trên chú thích trước mỗi thông số, hoặc tìm hiểu để biết thêm cách chỉnh sửa các thông số cho phù hợp.

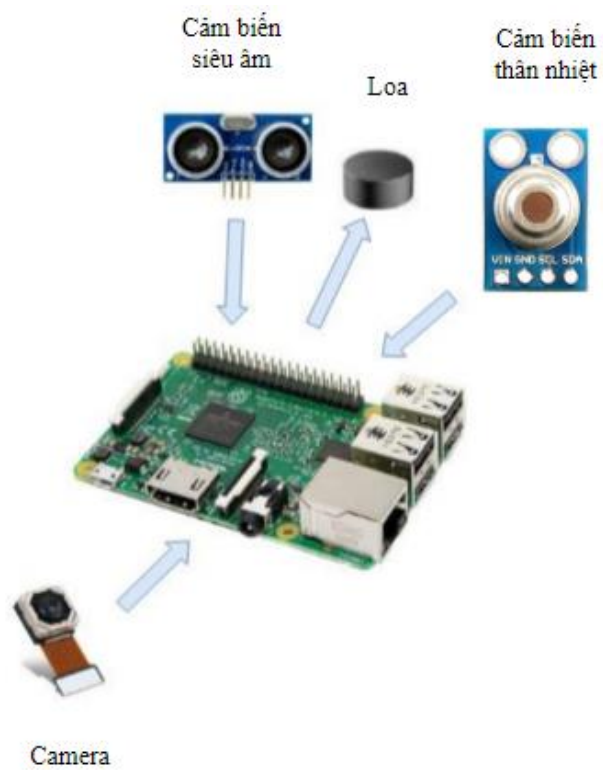
3.2. Giao tiếp với thiết bị khác qua giao thức MQTT

MQTT là một giao thức truyền tin hiệu quả, đáng tin cậy dựa trên nền tảng giao thức TCP/IP được sử dụng phổ biến hiện nay. MQTT thích hợp với việc truyền tải những dữ liệu nhỏ, yêu cầu tốc độ nhanh và sự đồng bộ dựa trên phương pháp đăng ký/xuất bản (subscribe/publish) lên một kênh bất kỳ.

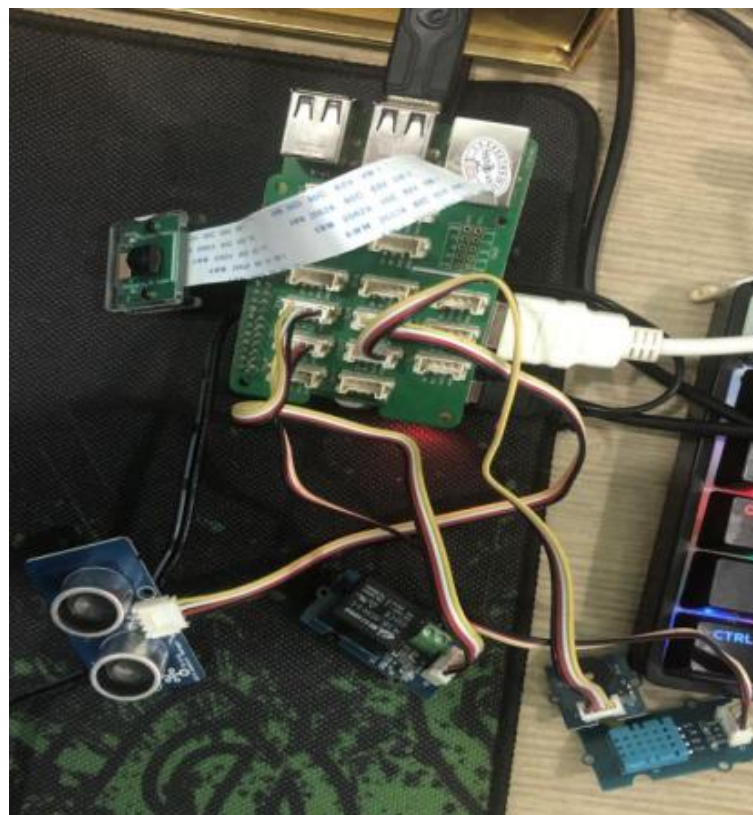
MQTT có hỗ trợ ngôn ngữ python. Để cài đặt gói giao thức MQTT cho python, sử dụng câu lệnh `pip install paho-mqtt`. Để thêm thư viện MQTT vào chương trình python, sử dụng `import paho.mqtt.client as mqtt`. Để đăng ký nội dung lên một kênh (topic), sử dụng `client.publish("topic", data, qos=0, retain=False)`. Để đăng ký nội dung từ một kênh bất kỳ, sử dụng câu lệnh `client.subscribe("topic", qos = 0)`. Chú ý, yêu cầu kênh đó phải tồn tại, có nghĩa là chúng ta phải chỉ rõ địa chỉ máy chủ nơi mà kênh được thiết lập và tồn tại, sử dụng `client.connect` (địa chỉ máy chủ, cổng, thời gian giữ tin).

4. MÔ HÌNH THỰC NGHIỆM

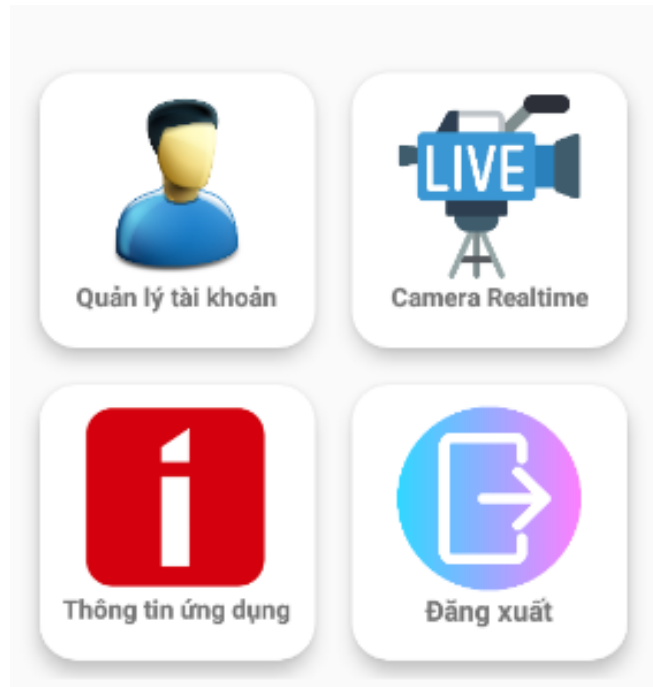
Từ những phương án trên chúng tôi đã xây dựng một mô hình thực tế và đã thử nghiệm các phương án nhiều lần và đạt được kết quả mong muốn. Chương trình đã nhận diện được khuôn mặt của khách quen và tự động nhấn chuông mà không cần chạm vào.



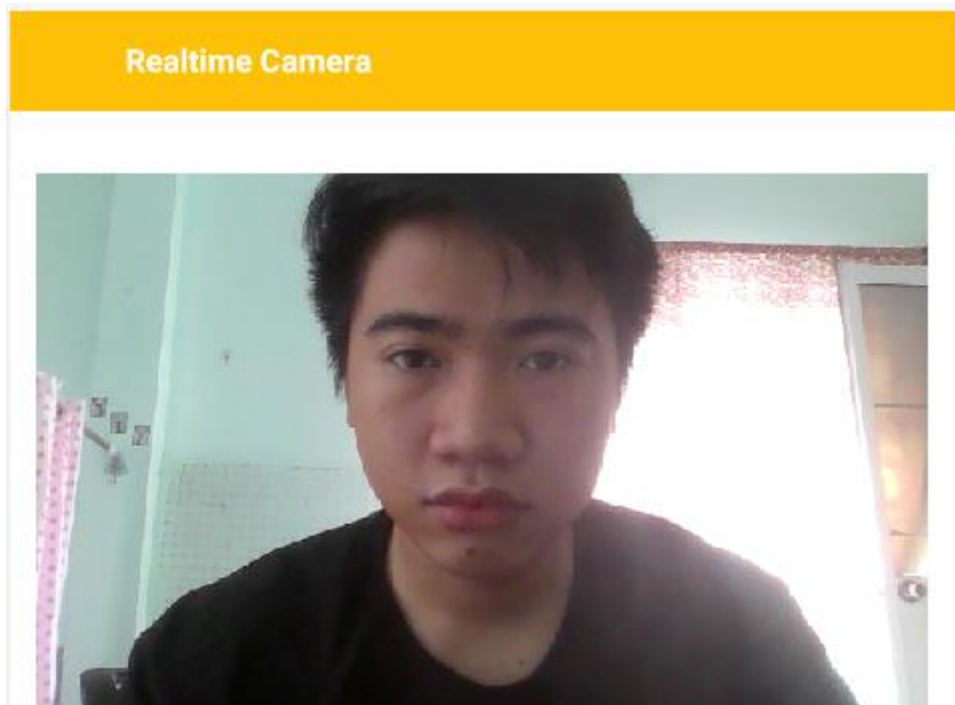
Hình 3: Sơ đồ kết nối phần cứng



Hình 4: Phần cứng của hệ thống



Hình 5: Phần mềm hệ thống



Hình 6: Kết quả phần mềm

5. KẾT LUẬN

Ưu điểm: bài báo đã đưa ra được phương án xây dựng hệ thống: “Chuông cửa không tiếp xúc”. Đây là một ý tưởng không hề mới xong có thể phát triển thành sản phẩm và có tính ứng dụng thực tế cao. Một số mô hình, công nghệ và thuật toán cũng đã được trình bày trong bài báo.

Hướng phát triển: Trong thời gian tới, nhóm tác giả sẽ tiếp tục nghiên cứu tinh chỉnh phần cứng và thực nghiệm trên tập dữ liệu lớn hơn để lựa chọn và cải thiện hiệu suất của sản phẩm. Đồng thời, nhóm tác giả cũng sẽ so sánh với các sản phẩm hiện có trên thị trường để hoàn thiện sản phẩm.

TÀI LIỆU THAM KHẢO

- [1] Amos, B., Ludwiczuk, B. & Satyanarayanan, M., 2016, OpenFace: A general-purpose face recognition library with mobile application, CMU-CS-16-118, CMU School of Computer Science
- [2] Gonzalez, R.C. & E., W.R., 2007, Digital Image Processing. 3rd ed. Pearson
- [3] Le Tien Dung, Trinh Ngoc Van, Nguyen Van Kien, Nguyen Duc Tuong. “Thiết kế hệ thống điều khiển cửa tự động sử dụng công nghệ xử lý giọng nói và công nghệ xử lý ảnh”. *Tạp Chí Khoa học Và Công nghệ - Đại học Đà Nẵng*, vol 3, số p.h 112.1, Tháng Ba 2017, tr 22-27, <https://jst-ud.vn/jst-ud/article/view/2545>.
- [4] https://docs.opencv.org/4.x/d3/df2/tutorial_py_basic_ops.html
- [5] <https://mindovermetal.org/ung-dung-cua-cam-bien-sieu-am-1639476077/>

YSC4F.224

CHẨN ĐOÁN LỖI XE Ô TÔ THỜI GIAN THỰC

NGUYỄN THANH ĐĂNG, TRẦN MINH QUÂN, TRẦN ANH KHOA

Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

nguyenthahdang@iuh.edu.vn, 18057771.quan@student.iuh.edu.vn, 18072151.khoa@student.iuh.edu.vn

Tóm tắt. Có rất nhiều thiết bị chẩn đoán lỗi xe ô tô hiện nay như máy GSCAN, TECHSTREAM, MAXISYS, FCAR C8-W, ZENITH Z5, AUTEL MS909. Các dòng máy này hỗ trợ rất tốt cho công tác bảo trì và sửa chữa ô tô, nhưng giá thành rất cao. Bài báo này đưa ra phương pháp đọc lỗi DTC (Diagnostic Trouble Code) trên xe ô tô, dễ thực hiện với chi phí thấp có tính năng quét lỗi trên nhiều hãng xe, từ đó liên kết đến các hãng xe và truy vấn vào các mã lỗi P,U,B,C giúp cho kỹ thuật viên chẩn đoán và đưa ra phương pháp sửa chữa chính xác. Ngoài ra, bài báo còn xây dựng cơ sở dữ liệu trực tiếp lên thiết bị. Để kiểm chứng độ hiệu quả của bài báo, sẽ thực hiện trên phần cứng Raspberry giao tiếp qua chuẩn OBD-II (On Board Diagnostics II).

Từ khóa. OBDII, Mã lỗi P,U,B,C, DTC.

REAL-TIME CAR FAULT DIAGNOSIS

Abstract. There's an extensive diversity of diagnostic equipment for car faults such as GSCAN, TECHSTREAM, MAXISYS, FCAR C8-W, ZENITH Z5, AUTEL MS909. These machines are very subsidiary for car maintenance and retrieve work, but the cost is formidable. The article gives a quick, easy to enforce with depleted cost reading errors DTC (Diagnostic Trouble Code) disposal on cars with the scanning errors feature on many car assemblers, thereby linking to car assemblers and querying for fault codes P, U, B, C help technicians determine and provide retrieve methods properly. In addition, the article also builds a database directly on the device. To verify the article's capability, it will be done on Raspberry hardware that communicates via OBD-II (On Board Diagnostics II) standard.

Keywords. OBDII, trouble code P,U,B,C, DTC.

1. GIỚI THIỆU

Hiện nay đã có rất nhiều bài báo nghiên cứu về việc truyền nhận dữ liệu qua OBD-II trên xe hơi như [1] giám sát mức tiêu thụ nhiên liệu ở tốc độ 5km/h với mật độ giao thông thấp thông qua Mode1 của OBD-II, [2] giám sát thông số xe vượt quá giá trị vận hành bình thường qua Mode1, [3] dự đoán hệ số khí thải dựa trên giám sát thông số NO_x của xe hơi qua Mode1, [4] dự đoán tuổi thọ động cơ, nước làm mát dựa trên thông số cảm biến tốc độ, cảm biến áp suất không khí, cảm biến mức nhiên liệu thông qua Mode1, [5] giám sát lưu lượng khí thải xe hơi thông qua Mode1, giám sát thông số ABS/ESB dựa vào quá trình phân tích lực phanh trên hệ thống con lăn [6], thực hiện Mode3 chẩn đoán lỗi thông qua giao tiếp với mạng CAN (Control Area Network) [7] [8] ước tính mức tiêu thụ nhiên liệu thông qua việc giám sát cảm biến tốc độ động cơ, dữ liệu đọc OBD-II cũng đã được lưu trên đám mây tiện lợi cho việc theo dõi và sửa chữa [9]. Các tác giả đã đọc được các thông số quan trọng trong giám sát xe hơi thông qua Mode1, Mode3. Đây là hai Mode đọc các thông số cảm biến và mã lỗi DTC, ngoài ra việc xóa các lỗi đọc được thông qua Mode4 chưa được đề cập đến, việc xác định thông tin về xe ô tô thông qua Mode9 cũng chưa được quan tâm đến, đây là thông tin rất quan trọng để hỗ trợ xác định nhà sản xuất và sửa chữa. Từ những hạn chế trên, bài báo này mạnh dạn sử dụng raspberry để đọc các Mode quan trọng như Mode1, Mode3, Mode4, Mode9, dữ liệu đọc các Mode này được lưu trực tiếp vào bộ nhớ của raspberry, qua đó liên kết đến các tài liệu sửa chữa của các hãng ô tô.

2. OBD-II

Bảng 1 cho thấy các PID (Proportional Integral Derivative) OBD-II tiêu chuẩn như được định nghĩa bởi SAE J1979. Phản hồi dự kiến cho mỗi PID được đưa ra, cùng với thông tin về cách chuyển phản hồi thành dữ liệu có ý nghĩa.

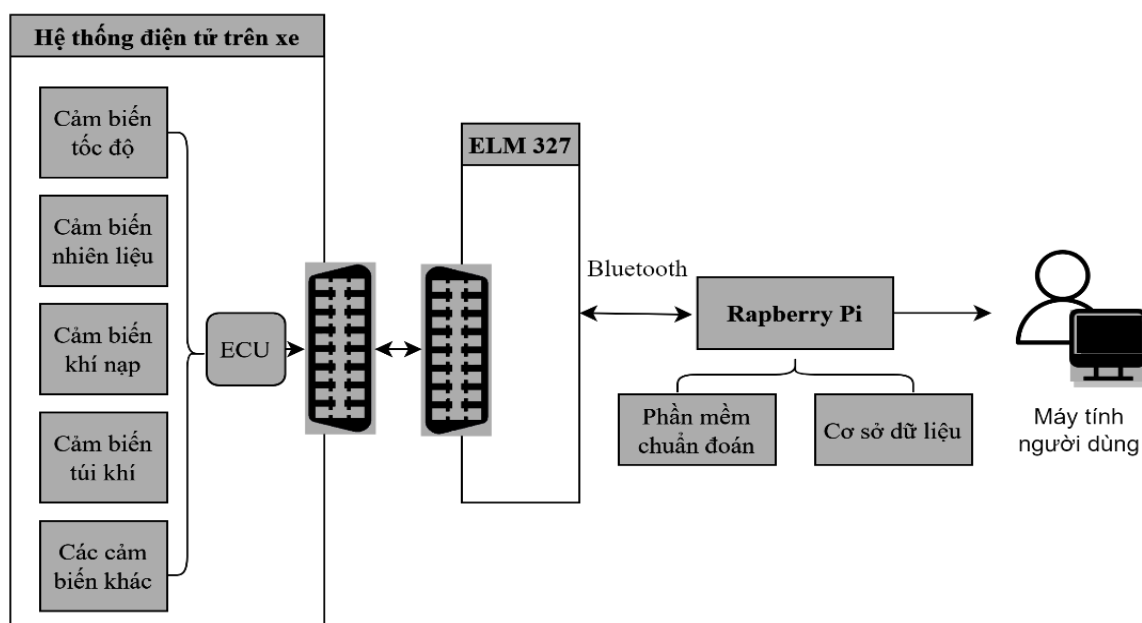
Bảng 9: Các Mode OBD-II

Mode	Mô tả
01	Hiển thị dữ liệu hiện tại - Dữ liệu trực tiếp
02	Hiển thị dữ liệu cố định
03	Hiển thị mã lỗi
04	Xóa/Xóa mã sự cố chẩn đoán và các giá trị được lưu trữ
05	Kết quả kiểm tra, giám sát cảm biến oxy
06	Kết quả kiểm tra, giám sát thành phần/hệ thống khác (Kết quả kiểm tra, giám sát cảm biến oxy chỉ CAN)
07	Hiển thị mã sự cố chẩn đoán đang chờ xử lý (được phát hiện trong chu kỳ lái xe hiện tại hoặc cuối cùng)
08	Kiểm soát hoạt động của thành phần/hệ thống trên bo mạch
09	Yêu cầu thông tin xe
0A	Mã sự cố chẩn đoán (DTC) (DTC đã xóa)

3. XÂY DỰNG HỆ THỐNG

3.1. Sơ đồ khối hệ thống

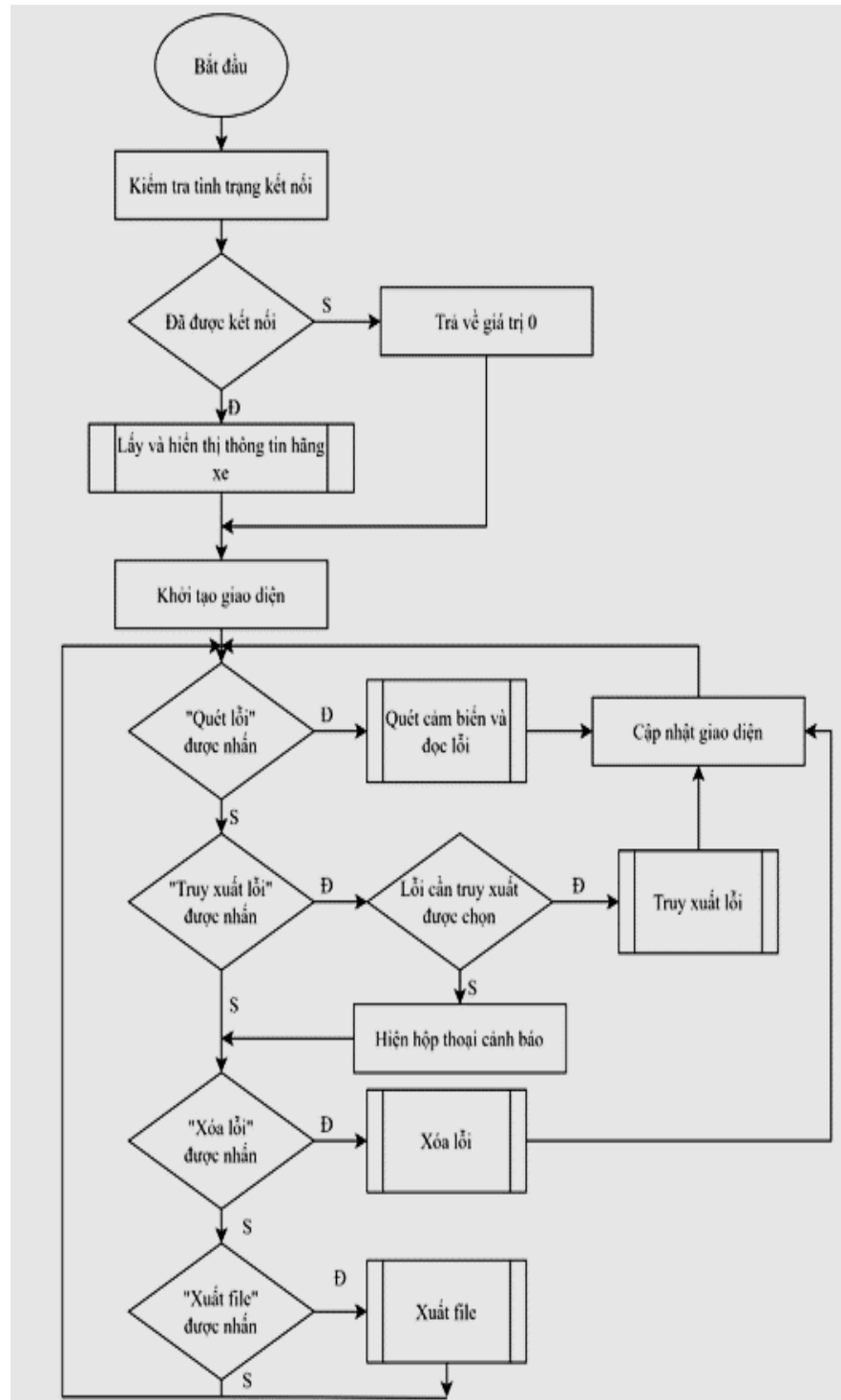
Xe hơi chứa các cảm biến bao gồm: cảm biến tốc độ, cảm biến nhiên liệu, cảm biến khí nạp, cảm biến túi khí... Các cảm biến này được điều khiển bởi ECU (Electronic Control Unit), thông qua ECU dữ liệu cảm biến thu thập được sẽ truyền lên cổng OBD-II thông qua mạng CAN. Các dữ liệu được truyền bao gồm các thông số cảm biến, mã lỗi DTC và thông tin xe hơi.



Hình 3: Sơ đồ khối hệ thống

Thiết bị đọc ELM327 được cắm trực tiếp vào cổng OBD-II xe hơi và kết nối Raspberry thông qua bluetooth. ELM327 như một trung gian truyền nhận dữ liệu giữa ECU trên xe hơi với Raspberry. Người dùng dễ dàng sử dụng phần mềm chẩn đoán hệ thống từ xa.

3.2. Lưu đồ hoạt động



Hình 4: Lưu đồ giải thuật

Khi bắt đầu, chương trình kiểm tra tình trạng kết nối với xe và lấy thông tin hãng xe lưu vào bộ nhớ tạm. Nếu không được kết nối, chương trình sẽ hiện cảnh báo cho đến khi thiết bị được kết nối. Khi kết nối thành công, chương trình lấy được mã VIN từ xe, phần mềm sẽ truy vấn tên hãng xe tương ứng và đọc các giá trị cảm biến được gửi từ ECU sau đó hiển thị lên giao diện giám sát.

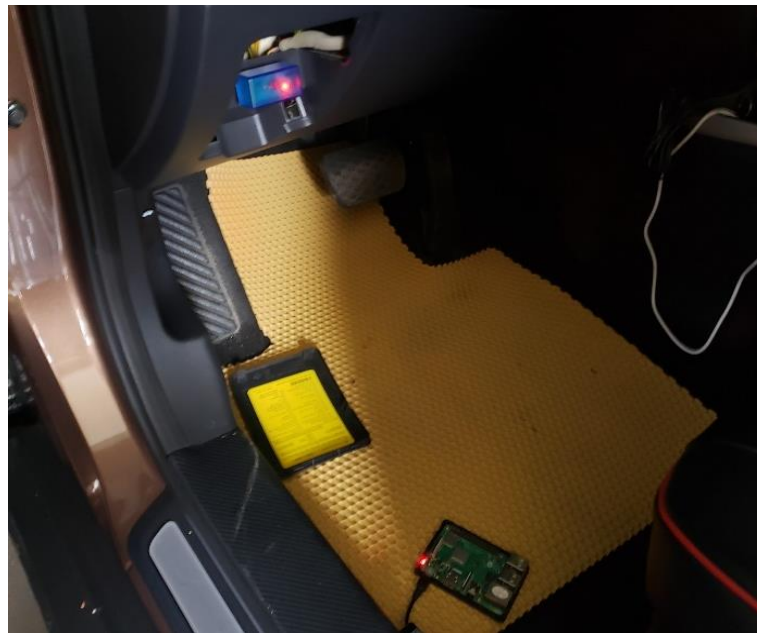
Khi thực hiện quét lỗi, chương trình sẽ đọc về các lỗi từ ECU và đưa lên giao diện giám sát theo định dạng danh sách.

Khi chọn chức năng xóa lỗi, hệ thống sẽ hiện lên thông tin lỗi đã được xóa. Nếu xác nhận, hệ thống tiến hành xoá lỗi trên xe.

Các mã lỗi đọc được từ ECU được lưu trữ dưới dạng file excel và lưu vào bộ nhớ của Raspberry Pi.

4. KẾT QUẢ THỰC NGHIỆM

Sau khi kết nối phần cứng với xe theo hình 3 và khởi động phần mềm sẽ quét thấy mã VIN (Vehicle Identification Numbers) bằng Mode9 và giá trị từ các cảm biến bằng Mode1 trên xe ô tô và được xử lý để lấy thông tin hãng xe và thông số các cảm biến hiển thị lên giao diện mô tả ở hình 4.



Hình 3: Kết nối phần cứng

Thông tin		
Hãng xe:	SsangYong	In file lỗi
Thông tin cảm biến		
Tải động cơ: 31.76%	Nhiệt độ nước làm mát: 52.0°C	Áp suất ống nạp: 35.0kPa
Tốc độ vòng tua: 894.0RPM	Tốc độ xe: 0.0Km/h	Nhiệt độ khí nạp: 33.0°C
Vị trí bướm ga: 13.33%	Thời gian chạy động cơ: 105.0s	Khoảng cách đã chạy: 0.0Km
Áp suất khí nạp: 99.0kPa	Van báo quá tải: 23.14%	Nhiệt độ môi trường: 33.0°C

Hình 4: Đọc thông tin hãng xe và thông số cảm biến qua Mode9, Mode1

Chuẩn đoán lỗi	
Mã lỗi	Mô tả
C1260	Lỗi cảm biến góc lái
P0134	Cảm biến Oxy dây 1 cảm biến 1 không hoạt động
P0461	Cảm biến bình nhiên liệu hiệu suất thấp
P0131	Nguồn yếu cảm biến Oxy dây 1 cảm biến 1

Quét lỗi Xóa lỗi Truy xuất file sửa

Tên sinh viên thực hiện đề tài

Trần Minh Quân - 18057771 Trần Anh Khoa - 18072151

Hình 5: Đọc lỗi DTC qua Mode3

Truy cập đến file sửa lỗi như hình 6 bằng phương thức truy vấn đến cơ sở dữ liệu với thông số mã lỗi lấy từ Mode 3 và mã VIN từ Mode9 mô tả ở hình 4 và hình 5.

2012 SsangYong L4-2.0L Bản quyền © 2013, ALLDATA10.52SS Trang 1

TẤT CẢ các mã sự cố chẩn đoán (DTC): Biểu đồ mã P Thông tin chung

Kiểm soát động cơ (SULEV)
P0461 Cảm biến mức nhiên liệu \ 'A ' \ Dải mạch / Hiệu suất - Thông tin chung

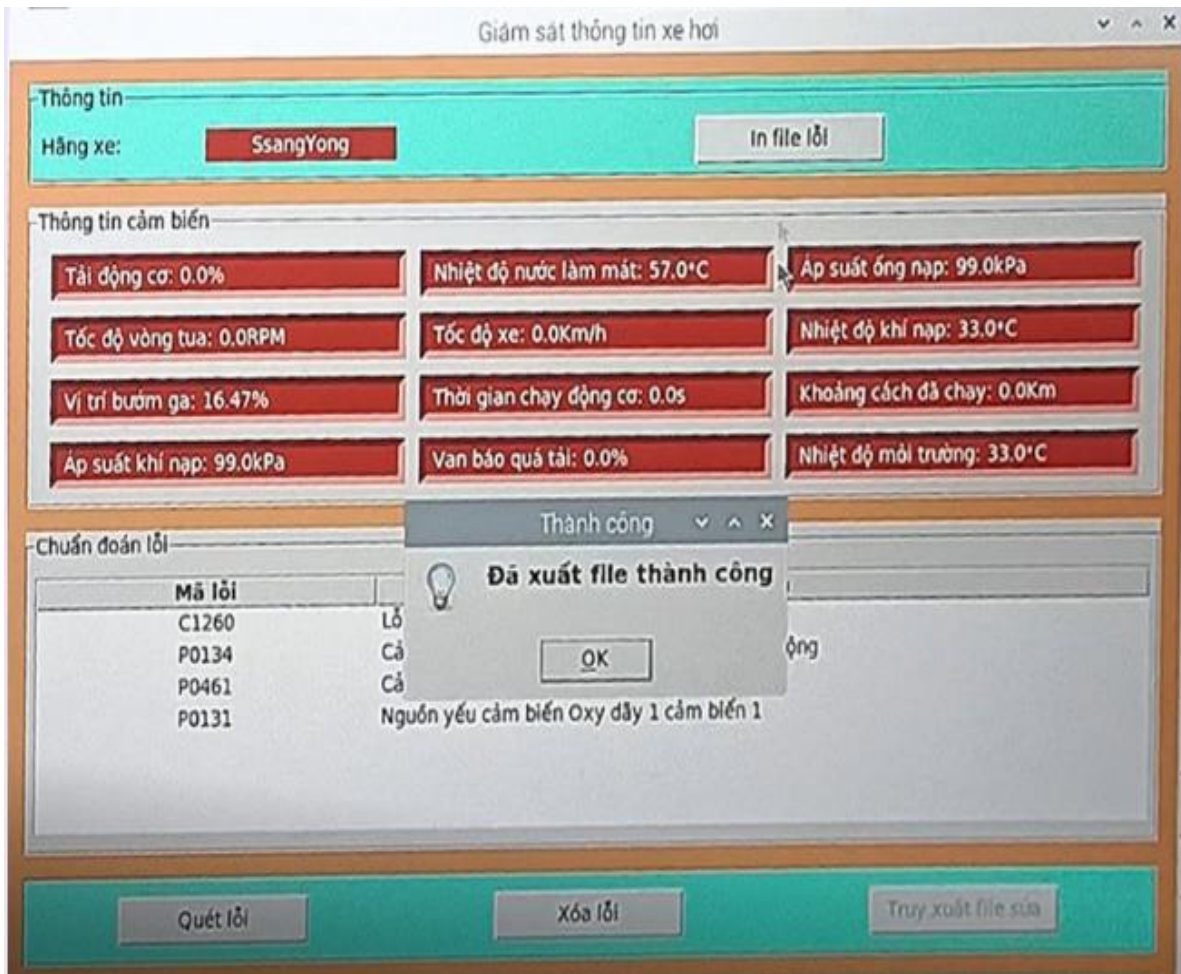
Vị trí thành phần



Mô tả chung

Cảm biến mức nhiên liệu (FLS) đo mức nhiên liệu trong thùng nhiên liệu và thông tin của nó được sử dụng để cung cấp cho PCM thông tin liên quan đến điều kiện giám sát của hệ thống thanh lọc hộp EVAP để phát hiện rò rỉ. Nếu cảm biến mức nhiên liệu được theo dõi để phát hiện bị kẹt ở mức cao hay thấp hoặc chênh lệch giữa điện áp của cảm biến mức nhiên liệu và độ lệch điện áp của cảm biến mức nhiên liệu vượt quá giá trị giới hạn trong khi lái xe, điều này được PCM hiểu là lỗi.

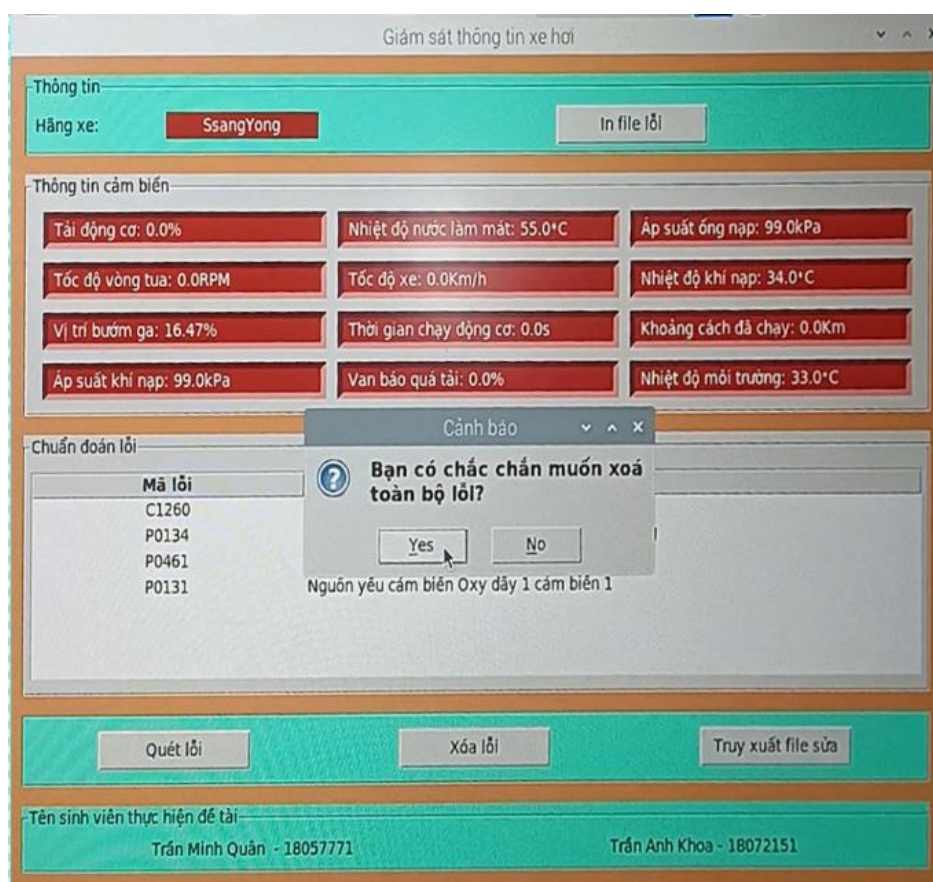
Hình 6: Truy cập đến file sửa lỗi



Hình 7: Giao diện phần mềm xuất file
 Danh sách mã lỗi được lưu thành công như hình 7. File danh sách lỗi được xây dựng như hình 8.

B	C	D	E	F	G
Mã lỗi	Mô tả				
C1260	Lỗi cảm biến góc lái				
P0134	Cảm biến Oxy dây 1 cảm biến 1 không hoạt động				
P0461	Cảm biến bình nhiên liệu hiệu suất thấp				
P0131	Nguồn yếu cảm biến Oxy dây 1 cảm biến 1				

Hình 8: Danh sách mã lỗi được xuất



Hình 9: Giao diện phần mềm xóa lỗi

Các thông số lỗi được lưu trữ ở bộ nhớ đệm trên xe hơi nên sau khi sửa chữa hoàn tất và quét lại vẫn sẽ hiện ra những lỗi cũ do đó việc xóa lỗi qua Mode4 là cần thiết minh họa như hình 9.

5. KẾT LUẬN

Bài báo sử dụng Raspberry Pi để đọc các Mode quan trọng, Mode1 đọc thông số các cảm biến trên xe ô tô, Mode3 đọc các mã lỗi trả về từ xe, Mode4 xóa mã lỗi trên bộ nhớ đệm xe hơi, Mode9 đọc thông tin xe hơi lưu vào Raspberry Pi.

Với độ phổ biến của máy tính nhúng Raspberry hiện tại với giá thành rẻ có thể dễ dàng làm ra một hệ thống đơn giản nhưng vẫn đảm bảo được các chức năng đọc các Mode thông qua ECU, qua đó liên kết đến các tài liệu sửa chữa của các hãng ô tô. Kết quả thực nghiệm đã được thực hiện trên dòng xe SsangYong, Kia, Hyundai cho thấy hệ thống đọc được các mã lỗi DTC và thông số cảm biến trên xe ô tô một cách chính xác. Từ đó cho thấy tính khả thi của bài báo trong ứng dụng thực tế và với giá thành rẻ hơn nhiều [10] so với các thiết bị chuyên dụng [11].

TÀI LIỆU THAM KHẢO

- [1] A. P. M. S. Dimitrios Rimpasa, "OBD-II sensor diagnostics for monitoring vehicle operation and consumption," 18 October 2019.
- [2] R. C. B. A. L.-M. B. Radu-Petru Fotescu, "System for monitoring the parameters of vehicle," 31 December 2020.
- [3] L. W. Z. J. S. Q. Z. X. Z. Y. a. X. Z. Jiguang Wang, "Research on Actual Road Emission Prediction Model of Heavy-Duty Diesel Vehicles Based on OBD Remote Method and Artificial Neural Network," 2021.

- [4] A. Srinivasan, "IoT Cloud Based Real Time Automobile Monitoring System," 18 October 2018.
- [5] S. Wierzbicki, "Evaluation of the effectiveness of on-board diagnostic systems in controlling exhaust gas emissions from motor vehicles," pp. 75-79, 2019.
- [6] P. S. a. K. D. Andrzej Gajek, "Algorithms for diagnostics of the hydraulic pressure modulators of ABS / ESP systems in stand," 30 July 2018.
- [7] H. M. S. S. J. H. K. K. Tae Un Kang, "Automated Reverse Engineering and Attack for CAN Using OBD-II," 15 April 2019.
- [8] H. A. M. A. a. B. A. Y. A. Tamer Abukhalil, "Fuel Consumption Using OBD-II and Support Vector Machine Model," 25 Jan 2020.
- [9] A. M. Sohail, K. S. Khattak, A. Iqbal, Z. H. Khan and A. Ahmad, "Cloud-based Detection of Road Bottlenecks Using OBD-II Telematics," 30 Nov. 2019.
- [10] "rasberrypi.com," [Online]. Available: <https://www.raspberrypi.com/products/raspberry-pi-3-model-b/?resellerType=industry>.

YSC4F.225

XÂY DỰNG ỨNG DỤNG HƯỚNG DẪN DU LỊCH TỰ ĐỘNG TẠI THÀNH PHỐ HỒ CHÍ MINH SỬ DỤNG KỸ THUẬT HỌC SÂU

DANG THI PHUC, NGUYEN VU LAM DUY, TRAN THI MINH KHOA, LE THI ANH TUYET

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

*phucdt@iuh.edu.vn, 18075581.duy@student.iuh.edu.vn, ttmkhoa@iuh.edu.vn,
lethianhtuyet@iuh.edu.vn*

Tóm tắt. Ngày nay với nhu cầu du lịch tăng cao, nhiều khách du lịch có nhu cầu tự khám phá, trải nghiệm mà không cần phải theo lịch trình. Tuy nhiên khi du lịch tự túc chắc chắn sẽ gặp nhiều khó khăn, đặc biệt là khi du lịch đến nước ngoài. Để hỗ trợ cho nhu cầu du lịch tự túc này, đặc biệt là đối với khách nước ngoài khi tới Thành phố Hồ Chí Minh. Chúng tôi xây dựng ứng dụng hỗ trợ khách du lịch tự túc khi đến Thành phố Hồ Chí Minh, có thể tự động khám phá tất cả địa điểm của Thành phố Hồ Chí Minh mà không cần đến hướng dẫn viên du lịch. Ứng dụng cho phép khách du lịch, thông qua camera của điện thoại biết được thông tin chi tiết như lịch sử, sự kiện, địa điểm... Nơi mà họ đang đặt chân đến. Chúng tôi sử dụng các mô hình học sâu để phân loại các địa điểm, dữ liệu được thu thập đa dạng giúp làm tăng độ chính xác. Chúng tôi huấn luyện trên các mô hình VGG16, ResNet50 và MobileNet, kết quả thu được độ chính xác tương ứng là 77.3%, 84.51% và 82.42%. Kết quả cho thấy độ chính xác của mô hình ResNet50 và MobileNet chênh lệch không đáng kể. Chúng tôi lựa chọn mô hình MobileNet để triển khai lên hệ thống bởi kiến trúc mạng nhẹ, phù hợp với thiết bị di động. Với kết quả này chúng tôi sẽ phát triển hệ thống cho cả nước để phục vụ cho khách du lịch tự túc.

Từ khóa. Học sâu, mạng nơ-ron tích chập, VGG16, ResNet50, mobileNet, ứng dụng phục vụ khách du lịch tự túc.

BUILDING AN AUTOMATIC TOURIST GUIDE APPLICATION IN HO CHI MINH CITY USING DEEP LEARNING TECHNIQUES

Abstract. Nowadays, with increasing tourism demand, many tourists have the need to discover and experience on their own without being interfered with by schedules. However, self-sufficient travel will certainly face many difficulties, especially when traveling to foreign countries. To support this demand for self-sufficient travel, especially for foreign visitors when coming to Ho Chi Minh City, we built an application to support self-sufficient tourists when coming to Saigon to discover all of them automatically. places of Saigon without a tour guide. The application allows tourists through the phone's camera to know detailed information such as history, events, locations. where they are going. We use deep learning models to classify the locations, the data collected varies greatly to increase the accuracy. We train on the models VGG16, ResNet50 and MobileNet, the results are obtained with the accuracy of 77.3%, 84.51% and 82.42%, respectively. The results show that the accuracy of ResNet50 and MobileNet models is not significantly different. We choose the MobileNet model to deploy on the system because of its lightweight network architecture, suitable for mobile devices. With this result, we will develop a system for the whole country to serve self-sufficient tourists.

Keywords. Deep Learning, Convolutional Neural Networks, VGG16, ResNet50, MobileNet, self-sufficient travel application.

1. GIỚI THIỆU

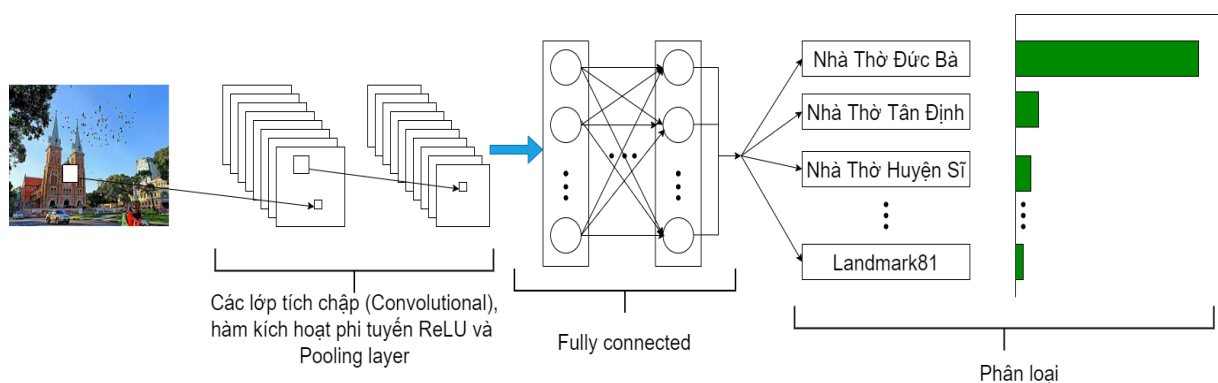
Trong những năm gần đây, sự tiến bộ vượt bậc của ngành công nghệ thông tin, đồng thời nó cũng là sự phát triển của cuộc cách mạng công nghiệp 4.0, hiện nay nó đang tác động mạnh mẽ đến các lĩnh vực trong đời sống con người. Thị trường du lịch toàn cầu cũng đang chứng kiến sự thay đổi, với sự lên ngôi của xu hướng du lịch kết hợp với ứng dụng công nghệ và thị trường Việt Nam cũng đang thay đổi để bắt kịp xu hướng hiện nay.

Nhìn thấy sự tiện ích đó, nhiều khách du lịch có nhu cầu thích tham quan, khám phá, trải nghiệm một mình mà không cần đến sự quản lý, lịch trình của những tour du lịch. Đồng thời nếu không có sự hướng dẫn của các hướng dẫn viên du lịch thì những chuyến đi tự túc sẽ gặp rất nhiều khó khăn đặc biệt là du khách nước ngoài, khi họ không biết thông tin nơi mình đã đến. Điều đó dẫn đến những trải nghiệm của họ không còn hấp dẫn nữa. Nhìn thấy sự khó khăn đấy, chúng tôi đã tìm hiểu, nghiên cứu và đưa ra quyết định xây dựng một ứng dụng hỗ trợ khách du lịch tự túc khi đến Thành phố Hồ Chí Minh, để cho khách du lịch nước ngoài có thể tự động khám phá các thông tin, địa điểm của Thành phố Hồ Chí Minh mà không cần đến hướng dẫn viên du lịch.

Với sự kết hợp giữa công nghệ và du lịch mà điều đó đã giúp thị trường du lịch của Việt Nam có sự đa dạng và mới mẻ hơn trước bạn bè quốc tế. Hệ thống ứng dụng du lịch thông minh bao gồm các thông tin địa điểm du lịch ở Thành phố Hồ Chí Minh từ hệ thống ứng dụng du lịch thông minh mà khách du lịch biết được những thông tin nơi mình đến. Nhằm nâng cao và đem đến trải nghiệm cho những khách du lịch tự túc khi đến với Việt Nam, trong bài báo, chúng tôi xây dựng ứng dụng hướng dẫn du lịch tự động tại Thành phố Hồ Chí Minh sử dụng kỹ thuật học sâu.

2. CÁC THUẬT TOÁN LIÊN QUAN

Mạng Neural tích chập (Convolutional Neural Network - CNN) [2] cho phép mô hình trích xuất được các đặc trưng của hình ảnh đầu vào, được đưa vào mô hình huấn luyện. Nó không giống với các mô hình đã phát triển lâu đời khác, CNN thực hiện bằng cách lấy các giá trị pixel thô của hình ảnh đầu vào và đem các giá trị đó đi qua từng lớp khác nhau trong mô hình, để huấn luyện và đi qua các tầng cuối để tính toán các đặc điểm, từ đó đưa ra xác suất để tiến hành phân loại. Kiến trúc CNN được mô tả ở hình 1.



Hình 1: Mô hình Convolutional Neural Network

Kiến trúc mạng CNN được xây dựng từ rất nhiều tầng cơ bản khác nhau như convolution layer, nonlinear layer, fully connected layer. Các lớp này được liên kết với nhau theo một số thứ tự nhất định. Thông thường thì một ảnh sẽ được truyền qua các tầng convolution layer, nonlinear layer đầu tiên, sau đó mô hình sẽ tính toán các giá trị pixel đầu vào và tới các tầng pooling layer. 3 lớp convolution layer, nonlinear layer và pooling layer có thể lặp lại nhiều lần trong network và cuối cùng đi qua các lớp cuối như fully connected layer và softmax để tính toán đưa ra xác suất thực hiện quá trình phân loại.

Lớp tích chập (Convolution Layer) [3] là một khối cốt lõi của CNN và nó là nơi diễn ra các quá trình tính toán phức tạp. Để thực hiện quá trình này nó yêu cầu một vài thành phần như dữ liệu đầu vào, bộ lọc (filter) và các giá trị đặc trưng. Chúng sẽ sử dụng các bộ lọc (filter) có kích thước nhỏ hơn so với ảnh thường và

được sử dụng với kích thước 3x3 hoặc 5x5, áp dụng vào một vùng của ảnh và tiến hành nhân tích chập giữa bộ filter và giá trị đặc trưng của các điểm ảnh đó.

Lớp Pooling [3] được dùng sau các lớp convolution nó tiến hành quét qua toàn bộ dữ liệu ảnh. Khác với lớp convolution, pooling layer không thực hiện các phép toán tích chập mà nó tiến hành lấy mẫu, khi cửa sổ trượt trên các giá trị ảnh, chỉ có các giá trị được xem là đại diện cho đặc trưng của ảnh thì được giữ lại và các phương pháp được sử dụng phổ biến như Max Pooling (giá trị lớn nhất), Average Pooling (giá trị trung bình). Max pooling được dùng khi bộ lọc di chuyển qua các giá trị đầu vào, nó sẽ tính toán và lựa chọn các giá trị có pixel lớn nhất để in ra mảng đầu ra, tuy nhiên phương pháp này được sử dụng thường xuyên hơn so với phương pháp average pooling. Average pooling được tính khi bộ lọc di chuyển qua các giá trị đầu vào, nó sẽ tính các giá trị trung bình trong trường tiếp nhận để gửi ra mảng đầu ra. Trong khi lớp max pooling khi sử dụng thì có nhiều thông tin bị mất, tuy nhiên average pooling giúp giảm thiểu được sự phức tạp và nâng cao tính hiệu quả, hạn chế được rủi ro khi bị overfitting. Lớp Pooling có vai trò giúp giảm kích thước ảnh xuống làm giảm được các công việc tính toán của mô hình, giúp mô hình có thể huấn luyện nhanh hơn và tăng hiệu quả tính toán. Sau khi đi qua pooling layer kích thước ảnh sẽ được thu nhỏ lại nhưng vẫn giữ được các giá trị đặc trưng cho hình ảnh.

Hàm Phi Tuyến: Sau khi đi qua các convolutional layer và max pooling layer chúng ta áp dụng các đơn vị tuyến tính chỉnh lưu (ReLU). Hàm ReLU tương đối giống với tế bào thần kinh của chúng ta một kích thước đủ lớn để tạo ra các phi tuyến tính cho các giá trị $x > 0$ và trả về 0 khi nó không đáp ứng điều kiện. Phương pháp này đã có hiệu quả khi giải quyết các độ dốc [4].

Lớp Fully Connected: Sau khi ảnh được truyền qua nhiều convolutional layer và pooling thì model đã học được tương đối các đặc điểm của ảnh thì tensor của output, layer cuối cùng sẽ được làm phẳng thành vector và đưa vào một lớp được kết nối như một mạng Neural. Với layer được kết hợp với các tính năng với nhau để tạo ra một mô hình. Cuối cùng sử dụng softmax để phân loại ảnh đầu ra [5].

Nhiều mô hình CNN nổi tiếng được nghiên cứu và xây dựng trên các tập dữ liệu lớn ImageNet và cho ra hiệu suất rất tốt như AlexNet, VGGNet, ResNet, MobileNet. Trong bài toán hướng dẫn du lịch tự động chúng tôi dùng các kiến trúc mạng này để trích xuất các đặc trưng từ các frame ảnh.

3. MÔ HÌNH CHO BÀI TOÁN HƯỚNG DẪN DU LỊCH TỰ ĐỘNG

Để giải quyết cho bài toán hướng dẫn du lịch tự động, trong bài báo chúng tôi đã đề xuất ra 3 mô hình: mô hình VGG16, ResNet50 và MobileNet để đánh giá và so sánh từ đó chưa ra được hướng phát triển để xây dựng cho bài toán mang tính hiệu quả nhất cho thực tế.

3.1. Mô hình MobileNet

Các mô hình CNN được đề xuất như VGG16, ResNet50 để huấn luyện thì đưa ra kết quả có độ chính xác cao nhưng chúng đều có một hạn chế chung đó là chúng thường ít phù hợp với các ứng dụng trên các thiết bị di động hay các hệ thống nhúng khác. Nếu muốn triển khai trên các mô hình có hệ thống ứng dụng real time thì chúng cần một cấu hình rất mạnh mẽ đối với các hệ thống có các phần cứng hỗ trợ như: Raspberry Pi, Nano hay các ứng dụng chạy trên các thiết bị di động. Do đó chúng cần một mô hình nhẹ hơn và có độ chính xác cao. MobileNet [6] là một mô hình kiến trúc mạng tích chập (CNN – convolutional neural network) được sử dụng cho việc phân loại hình ảnh và tìm vật thể trong ảnh hoặc video (hình 2). Mô hình này được thiết kế tối ưu về tốc độ tính toán các lớp tích chập và giảm bớt số lượng tham số trong mô hình. Điều này khiến cho kiến trúc MobileNet phù hợp với các thiết bị di động hay những chiếc máy tính nhúng với cấu hình thấp như Raspberry Pi hay Jetson Nano, Google TPU Dev Board... MobileNet sử dụng các tập hợp có thể phân tách theo chiều sâu để xây dựng neural network. Nó làm giảm đáng kể số lượng các tham số. Mô hình MobileNet dựa trên sự phân tách theo chiều sâu. Chúng tôi giới thiệu hai tham số giúp cân bằng hiệu quả giữa độ trễ và độ chính xác. Kiến trúc mô hình MobileNet bao gồm:

- Spatially Separable Convolutions (SSC): So với mạng nơ-ron tích chập truyền thống, SSC giúp giải quyết các bài toán theo thời gian thực khả thi hơn rất nhiều. SSC sẽ chia một convolution thành hai bước. SSC dùng chủ yếu các tham số về chiều của không gian của ảnh và kernel, nghĩa là với chiều rộng (width), chiều dài (height) và số lượng kênh (depth).
- Depthwise Separable Convolutions (DSC): Kiến trúc của MobileNet dựa trên Depthwise Separable Convolutions (DSC) chia một kernel thành hai kernel riêng biệt là Depthwise Convolution và Pointwise

Convolution để thực hiện hai phép tích chập riêng biệt. Điều này sẽ giúp giảm thiểu số lượng phép tính đáng kể, cụ thể sẽ được trình bày dưới đây.

Đối với phép tính của tích chập truyền thống hàm tính toán có công thức:

$$f_s(K, M, N, F) = D_K \cdot D_K \cdot M \cdot N \cdot D_F \cdot D_F \quad (1)$$

Trong đó hàm tính toán trên phụ thuộc vào M, số lượng kênh đầu ra N, với kernel kích thước $D_K \times D_K$ và feature map kích thước $D_F \times D_F$. Đối với Depthwise convolution, hàm chi phí tính toán được tính theo công thức:

$$f_{DC}(K, M, F) = D_K \times D_K \cdot M \cdot D_F \cdot D_F \quad (2)$$

Như vậy, số lượng phép tính trong Depthwise Convolution đã giảm đi N lần so với Standard Convolution. Số lượng phép toán tại Depthwise Convolution không phụ thuộc vào số lượng kênh của đầu vào. Với thành phần Pointwise Convolution, số lượng các phép toán cần có sau khi đi qua sẽ là:

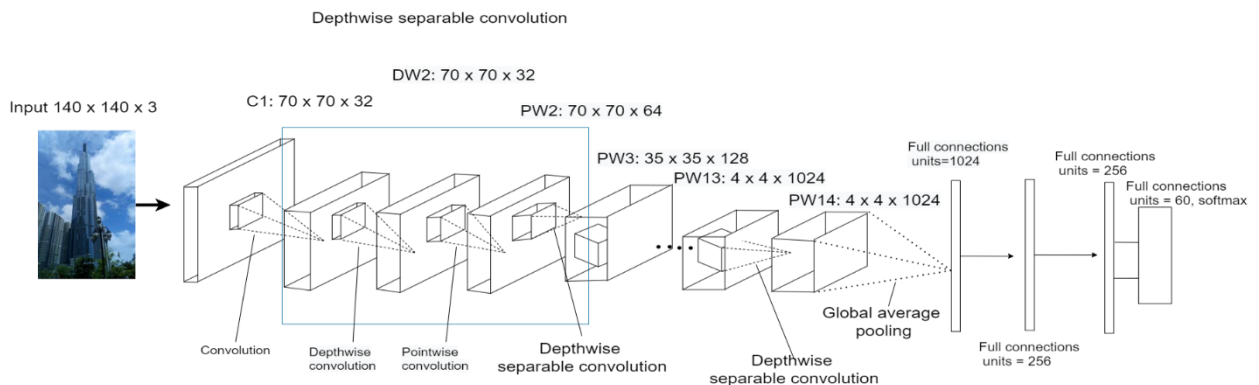
$$f_{PC}(M, N, F) = M \cdot N \cdot D_F \cdot D_F \quad (3)$$

Kết hợp với các thành phần Pointwise Convolution từ công thức (2) đến (3) hàm tính toán của Depthwise Separable Convolution là:

$$f_{DSC}(K, M, N, F) = D_K \cdot D_K \cdot M \cdot D_F \cdot D_F + M \cdot N \cdot D_F \cdot D_F \quad (4)$$

Từ công thức (4) so sánh với Standard convolution (1) ta được:

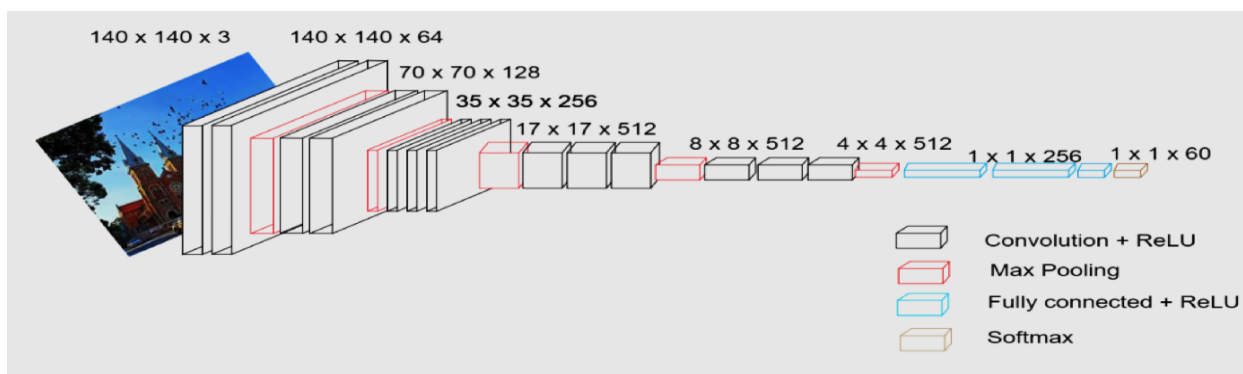
$$R_{SC-DSC}(K, N) = \frac{R_{DSC}(K, M, N, F)}{R_{SC}(K, M, N, F)} = \frac{1}{N} + \frac{1}{D_K^2} \quad (5)$$



Hình 2: Kiến trúc mô hình MobileNet[6]

3.2. Mô hình VGG16

Mô hình VGG16 [7] được xây dựng với các bộ lọc tích chập là một kiến trúc gồm 16 lớp. VGG16 bao gồm nhiều convolution layer, pooling layer... Trong đó có 13 lớp phức hợp và 3 lớp kết nối đầy đủ và lần đầu tiên VGG16 xuất hiện về các khối tích chập Block, đây là những kiến trúc gồm một tập hợp các lớp CNN được lặp lại giống nhau. VGG16 cũng là kiến trúc đầu tiên được thay đổi thứ tự của các Block khi xếp nhiều lớp CNN và Max Pooling thay vì xen kẽ chỉ một lớp CNN và Max Pooling. Mạng VGG16 có kích thước tham số khoảng 134 triệu bao gồm 13 lớp tích chập (Convolution Layer), 5 lớp Max Pooling với kích thước kernel 2x2 và 3 lớp Fully Connected, cuối cùng là một hàm softmax để tính toán xác suất cho 60 lớp. Mô hình VGG16 được mô tả ở hình 3:



Hình 3: Kiến trúc mô hình VGG16[7]

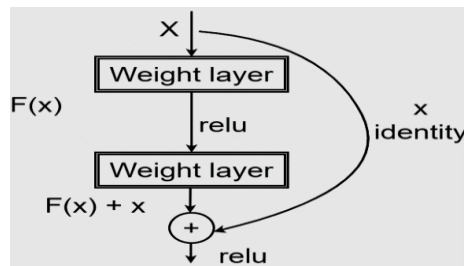
VGG16 bao gồm 13 lớp convolution, 5 lớp max-pooling và 3 lớp fully connected. Do đó số lớp có tham số có thể điều chỉnh là 16 bao gồm 13 lớp Convolution và 3 lớp fully connected. Số lượng bộ lọc của khối đầu tiên trong kiến trúc mạng là 64, tiếp theo thì sẽ được nhân đôi lên có khi đạt đến con số 512 và mô hình này được hoàn thiện với hai lớp ẩn là lớp fully connected hidden và output layer. Hai lớp được kết nối đầy đủ và có cùng số neural là 256. Lớp đầu ra bao gồm 60 lớp neural tương ứng với số loại trong tập dữ liệu. Điểm mạnh của mô hình VGG16 là mạng có kiến trúc khá sâu từ đó nó giúp ích cho việc cải thiện được độ chính xác của mô hình tốt hơn, nó hầu như kế thừa các điểm mạnh của AlexNet và cải tiến thêm những gì mà AlexNet chưa thể đạt được. VGG16 là kiến trúc đầu tiên thay đổi thứ tự của Block khi xếp nhiều layers CNN và Max Pooling thay vì xen kẽ chỉ một layer CNN và Max Pooling. Các layers CNN sâu hơn thì có thể trích lọc các đặc trưng tốt hơn so với chỉ một CNN. VGG16 chỉ sử dụng các kích thước bộ lọc, có kích thước nhỏ 3x3 thay vì nhiều kích thước như AlexNet. Vì kích thước bộ lọc nhỏ sẽ giúp giảm số lượng các tham số cho mô hình và mang lại hiệu quả tính toán hơn. Nhược điểm của mô hình VGG16 là mô hình tuy có độ chính xác trong các tập dữ liệu rất cao, nhưng nó không phù hợp với các ứng dụng trên mobile hay các hệ thống nhúng có khả năng tính toán thấp. VGG16 có số lượng tham số nhiều nên khó ứng dụng vào ứng dụng realtime. Ngoài ra, VGG16 có cấu trúc trích xuất thông tin theo một đường thẳng cho nên nó dẫn đến việc các thông tin được thực hiện từ trên xuống dễ bị đánh mất.

3.3. Mô hình ResNet50

Mô hình ResNet50 [8] là kiến trúc được sử dụng phổ biến nhất trong thời điểm hiện tại. ResNet50 cũng là kiến trúc sớm được áp dụng batch normalization. Những kiến trúc gần đây thường cải tiến độ chính xác nhờ gia tăng chiều sâu của mạng CNN. Nhưng thực nghiệm cho việc gia tăng chiều sâu đến một ngưỡng nào đó thì độ chính xác của mô hình sẽ bão hòa, thậm chí còn làm mô hình kém đi độ chính xác hơn. Khi đi qua quá nhiều tầng thì thông tin gốc bị mất đi, chính vì thế ResNet50 đã giải quyết bằng cách sử dụng kết nối tắt (Skip Connection) (hình 4). Các kết nối tắt này hoạt động là giảm bớt vấn đề gradient biến mất bằng cách thiết lập một kết nối tắt thay thế cho gradient đi qua. Ngoài ra, chúng cho phép mô hình học một chức năng nhận dạng. Điều này đảm bảo rằng các lớp cao hơn của mô hình không hoạt động kém hơn các lớp thấp hơn. Các khối dư giúp các lớp học các chức năng nhận dạng dễ dàng. Kết quả là, ResNet50 cải thiện hiệu quả của mạng nơ-ron sâu với nhiều lớp nơ-ron hơn trong khi giảm thiểu tỷ lệ lỗi. Nói cách khác, các kết nối tắt bỏ qua đầu ra từ các lớp trước và đầu ra của các lớp xếp chồng lên nhau, giúp cho việc đào tạo các mạng sâu hơn nhiều so với trước đây có thể. Các kết nối tắt giúp giữ thông tin không bị mất bằng cách kết nối từ các lớp trước đó, tới các lớp phía sau và bỏ qua một vài lớp trung gian.

Kỹ thuật kết nối tắt (Skip Connection) thay đổi cách học của mạng học sâu, giả sử chúng ta gọi hàm mục tiêu của cần học là $H(x)$, trước đây ta sẽ dùng các mạng để học từ $x \rightarrow H(x)$. Bây giờ với kỹ thuật skip connection, $H(x) = F(x) + x$, với $F(x)$ là những đặc trưng mạng học được “bổ sung” vào x thay vì trực tiếp tác động vào x . Ta gọi đây là Residual Learning, $F(x)$ là phần “residue” và x là “identity”. Đầu tiên, việc thay đổi mạng học thành những Residual Block ở trên khiến cho mạng học sâu của ta “dễ học hơn”, khi $H(x)$ là một hàm kết hợp giữa 1 identity function $I(x) = x$ và 1 phần “residue” $F(x)$. Giả sử, trong quá trình tối ưu, việc tối ưu hàm $F(x)$ không cải thiện cho hàm mục tiêu $H(x)$ (thậm chí là làm tệ đi), bài toán tối ưu đơn giản sẽ zero out $F(x)$ và vẫn lưu giữ lại $H(x) = 0 + I(x)$, giúp cho mô hình học được các identity mapping và tránh bị hiện tượng degrade như đã nói. Thứ 2 là vấn đề về độ phức tạp của mô hình, việc thêm nhiều

lớp vào mạng học sâu chắc chắn sẽ khiến cho khả năng học mô hình tăng lên, tuy nhiên không có gì đảm bảo được rằng, ta có thể dễ dàng tìm kiếm được optimal solution hơn so với mô hình nhỏ hơn. Skip connection trong ResNet50 có 2 dạng là Identity mapping và Convolutional block. Identity mapping trong trường hợp này residual mapping đơn giản là việc cộng trực tiếp x vào đầu ra của các stacked block $F(x)$. Đây là một cách sử dụng khá phổ biến trong thiết kế mạng ResNet50, nếu như input activation có cùng số chiều với output activation. Convolutional block một trường hợp khác là thay vì cộng trực tiếp các giá trị input activation. Trường hợp này có thể được thực hiện trong trường hợp input activation và output activation có số chiều khác nhau. Lúc này đầu ra được xác định như sau $y = F(x ; W_i) + \text{Conv}(x)$. Việc sử dụng các kết nối tắt chính là ưu điểm của mô hình giữ được các thông tin của hình ảnh khi trong quá trình thực hiện huấn luyện mà không bị mất, bằng cách kết nối từ layer trước đó tới layer phía sau và bỏ qua một vài layers trung gian. Tuy nhiên, nhược điểm của ResNet50 chính là số lượng tham số lớn từ đó không được ứng dụng vào ứng dụng có phần cứng yếu hoặc các hệ thống nhúng khác.

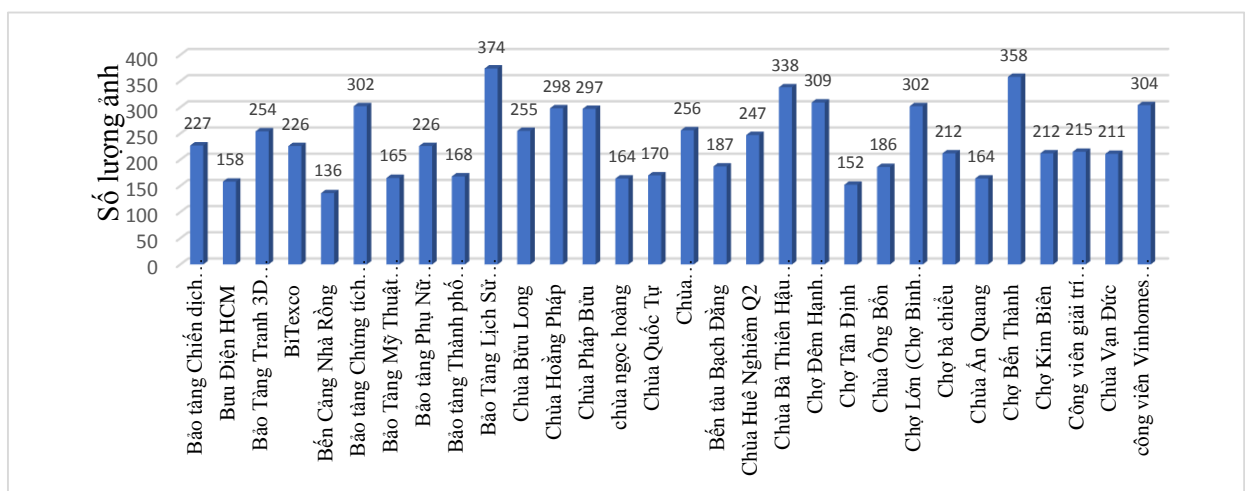


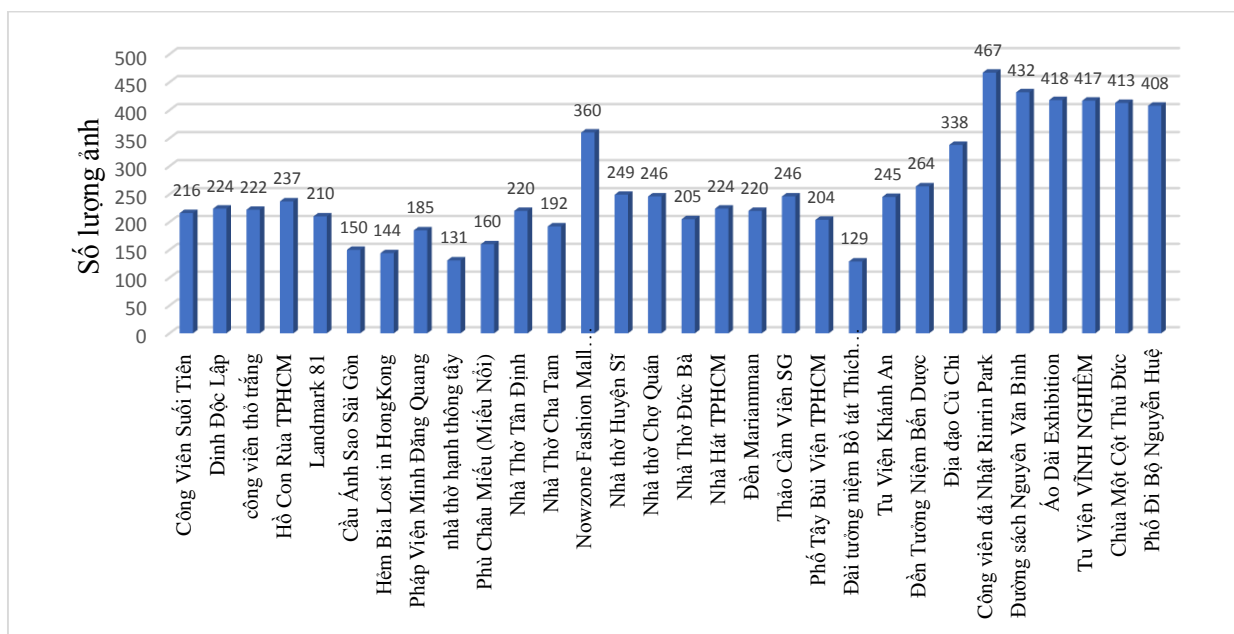
Hình 4: Mô hình kết nối tắt (Skip Connection) trong RetsNet [8]

4. KẾT QUẢ THỰC NGHIỆM

4.1. Bộ dữ liệu

Bộ dữ liệu chúng tôi sử dụng là bộ dữ liệu hình ảnh được thu thập một cách thủ công từ các công việc như cắt các khung ảnh từ những video ở nhiều nguồn khác nhau, thu thập từ hình ảnh thực tế, từ đó chúng tôi tiến hành gán nhãn từng hình ảnh cho từng lớp khác nhau. Tập dữ liệu bao gồm với hơn 14846 hình ảnh của các địa danh khác. Để chuẩn bị cho quá trình huấn luyện cho mô hình chúng tôi tiến hành chia ra các tập dữ liệu để đào tạo làm 3 phần, trong đó 12024 ảnh của tập huấn luyện, 1485 của tập kiểm thử và 1337 của tập hình ảnh test với 60 lớp tương ứng với 60 địa danh khác nhau trong khu vực TP. HCM. Phân bố dữ liệu của các lớp được mô tả ở hình 5.

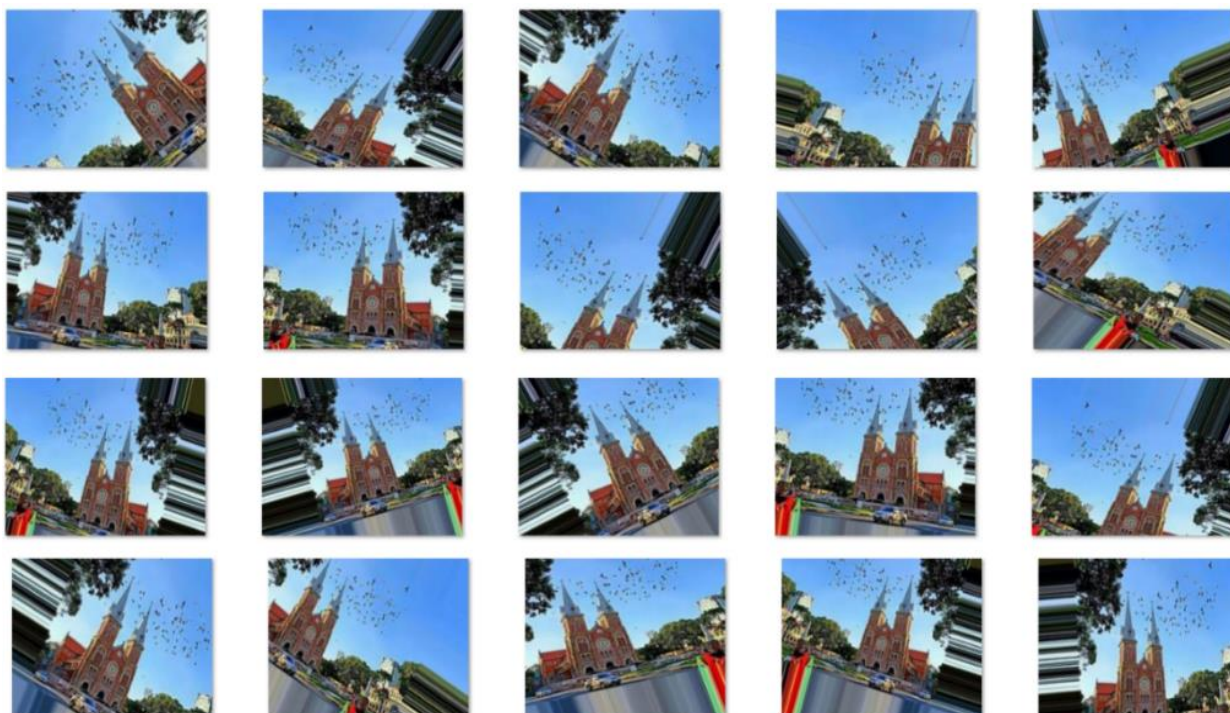




Hình 5: Biểu đồ phân bố dữ liệu thu thập của các địa điểm du lịch ở TP.HCM

4.2. Tăng cường dữ liệu

Phương pháp tăng cường dữ liệu (data augmentation) là phương pháp được dùng phổ biến nhất để giảm được việc dữ liệu quá khớp hoặc trong quá trình huấn luyện mà mô hình có kết quả không được khả quan. Do đó chúng tôi áp dụng công việc tăng cường dữ liệu trên các tập dữ liệu gốc bằng cách lật ngang, xoay hình ảnh, chia kích thước và thay đổi độ sáng và không gian màu RGB ngẫu nhiên. Mục tiêu là làm thay đổi các tỷ lệ của từng chi tiết các đặc điểm ảnh khác nhau. Dữ liệu tạo mới được mô tả như hình 6.



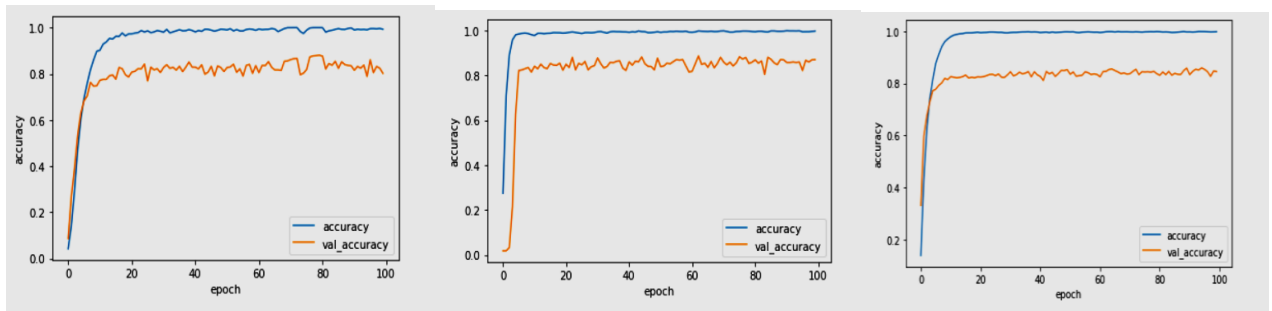
Hình 6: Tăng cường dữ liệu

4.3. Thiết kế mô hình

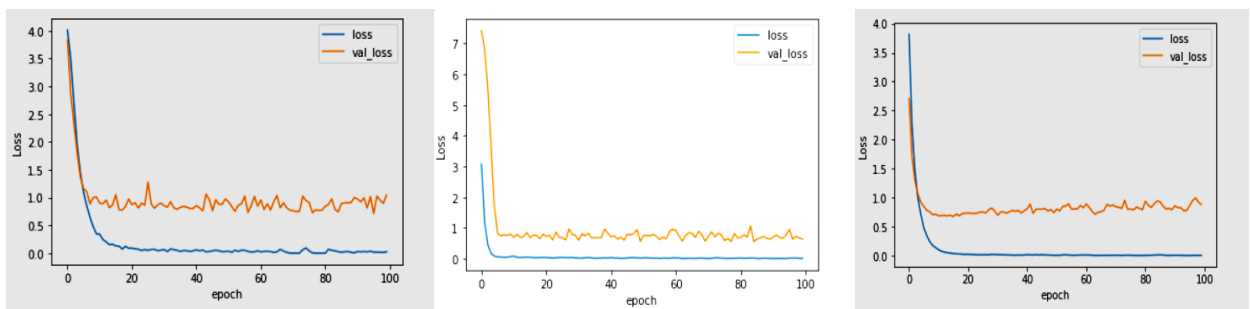
Để huấn luyện mô hình cho bài toán hướng dẫn du lịch tự động, chúng tôi sử dụng công cụ Google Colab để tận dụng các tính năng của GPU. Đối với các mô hình đã đề xuất chúng tôi thiết kế các giá trị siêu tham số mô hình như: Giá trị learning rate là 0.0001, batch size là 64 và epoch là 100. Các giá trị learning rate được điều chỉnh phù hợp trong quá trình huấn luyện để cho mô hình đạt được độ chính xác cao nhất.

4.4. Kết quả so sánh

Kết quả huấn luyện các mô hình trên Google Colab cho chúng tôi thấy độ chính xác của 3 mô hình trên tập dữ liệu của đề tài đều đạt trên 80%.



Hình 7: Độ chính xác của các mô hình tương ứng; a) VGG16, b) ResNet50 và c) MobileNet



Hình 8: Hàm mất mát của các mô hình tương ứng; a) VGG16, b) ResNet50 và c) MobileNet

Bảng 1: Bảng so sánh các mô hình trên tập ImageNet

	VGG16	ResNet50	MobileNet
Size (MB)	528	98	16
Top-1 Accuracy	71.30%	74.90%	70.40%
Top-5 Accuracy	30.10%	92.10%	89.50%
Parameter	138.4M	25.6M	4.3M
Depth	16	107	55
Time (ms) per inference step (CPU)	69.5	58.2	22.6
Time (ms) per inference step (GPU)	4.2	4.6	3.4

Các mô hình VGG16, ResNet50, MobileNet được huấn luyện trên tập Image trong các cuộc thi thường niên thi cho ra được các giá trị và mức độ chính xác theo từng mô hình, từ đó chúng ta so sánh và xác định mô hình ResNet có độ chính xác cao hơn VGG16, MobileNet.

Bảng 2: Bảng so sánh trên tập test

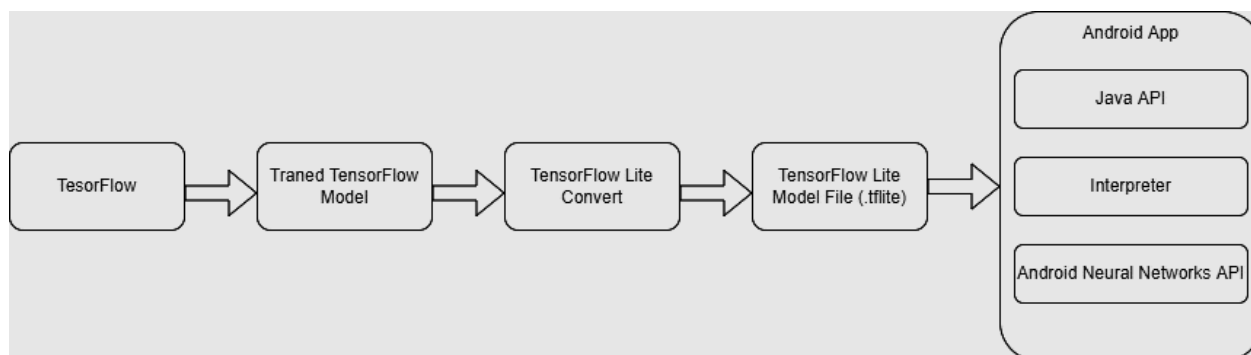
Mô hình	Accuracy	Precision	Recall:	ROC AUC	F1 score
VGG16	77.30%	80.79%	76.94%	98.40%	76.94%
RetsNet50	84.51%	86.17%	84.36%	99.23%	84.41%
MobileNet	82.42%	82.63%	81.38%	99%	81.38%

Chúng tôi tiến hành đánh giá độ chính xác trên tập dữ liệu test gồm 1485 hình ảnh. Kết quả được mô tả ở bảng 2. Kết quả cho thấy mô hình cho thấy, dựa trên các giá trị đánh giá Accuracy, precision, recall, ROC AUC, F1 score thì mô hình ResNet50 đạt kết quả cao hơn so với các mô hình khác.

4.5. Triển khai mô hình lên ứng dụng

Sau khi huấn luyện các mô hình và đưa ra những kết quả so sánh, từ đó chúng tôi chọn ra mô hình có độ chính xác cao. Vì đặc thù của các thiết bị như app mobile hay các hệ thống nhúng khác đều hạn chế về phần cứng thế nên khi triển khai mô hình AI lên thì cần phải đáp ứng được với nhu cầu đó. Để tiến hành đưa một mô hình AI lên App thì chúng ta cần xử lý để giảm kích thước mô hình xuống, để tiết kiệm được độ trễ xử lý của hệ thống và không gian bộ nhớ... Tensorflow lite là phiên bản nhẹ hơn của Tensorflow cũng do Google phát triển. Tensorflow Lite được thiết kế để chạy các mô hình trên thiết bị di động.

Như chúng ta cũng đã biết một mô hình học sâu có hàng chục triệu tham số để tính toán và cần một lượng tài nguyên lớn và GPU để hỗ trợ cho công việc huấn luyện, từ đó sẽ dẫn đến cho mô hình chúng ta trở nên nặng hơn, khi triển khai mô hình vào thì nó dẫn đến hiệu suất của ứng dụng giảm đi và giảm khả năng trải nghiệm cho người dùng. Ngược lại nếu một mô hình có quá ít tham số thì khả năng học không hiệu quả và cũng dẫn đến mô hình không được đưa vào sử dụng. Cho nên việc nén mô hình từ TensorFlow sang TensorFlow Lite là rất cần thiết để giúp cho mô hình trở nên gọn nhẹ mà vẫn giữ được độ chính xác trong các thiết bị nhúng. Các bước triển khai được mô tả ở hình 9.

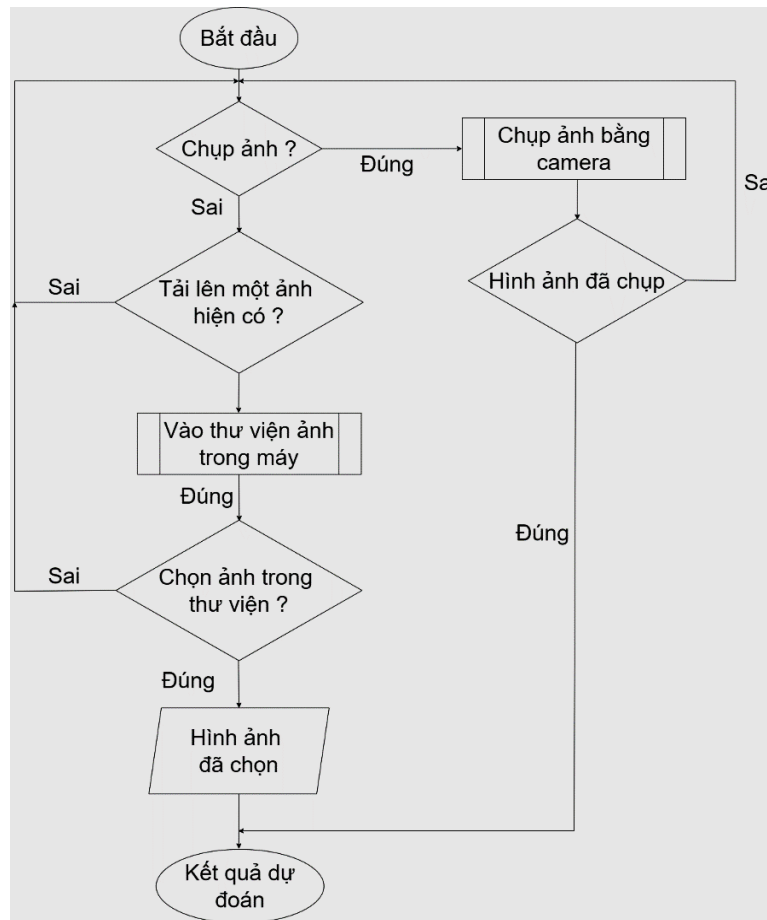


Hình 9: Sơ đồ triển khai lên ứng dụng

Ứng dụng Travel Guide được mô tả ở hình 10. Trong ứng dụng Travel Guide người dùng cần đưa các hình ảnh vào ứng dụng để hệ thống có thể xử lý và đưa ra kết quả. Đối với sơ đồ được thực hiện bằng camera thì sơ đồ được mô tả như sau:

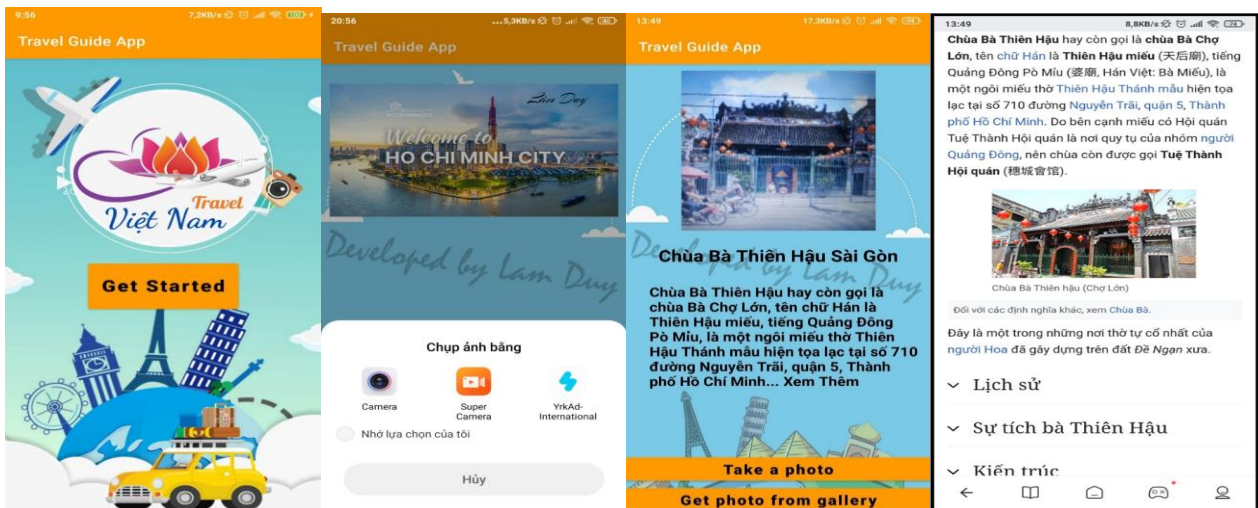
Sau khi người dùng chọn chức năng chụp ảnh thì ứng dụng sẽ hiện lên giao diện để cho người dùng có thể chọn được ứng dụng chụp ảnh, tiếp theo người dùng cần thực hiện các thao tác chụp ảnh. Nếu mà người dùng chụp ảnh bị mờ hoặc không không đúng ý thì có thể chọn chụp lại, nếu ảnh vừa ý thì người dùng nhấn dấu tích và để cho hệ thống xử lý.

Sau khi đợi hệ thống xử lý xong và trả về kết quả, ứng dụng Travel Guide sẽ hiển thị các thông tin về địa điểm đó. Nếu người dùng muốn xem chi tiết thì có thể nhấn “Xem Thêm” thì ứng dụng sẽ đưa người dùng đến trang thông tin chính thống của địa điểm đó.



Hình 10: Triển khai mô hình lên ứng dụng

Kết quả ứng dụng được mô tả ở hình 11.



5. KẾT LUẬN

Trong bài báo, chúng tôi xây dựng một ứng dụng hướng dẫn du lịch tự động ở TP.HCM bằng kỹ thuật Deep Learning. Mô hình được huấn luyện trên tập dữ liệu được thu thập thủ công bao gồm 14846 hình ảnh với 60 lớp khác nhau được thu thập trên các địa điểm du lịch tại TP.HCM. Chúng tôi tiến hành huấn luyện trên

các mô hình học sâu như VGG16, ResNet50 và MobileNet. Các mô hình được kiểm chứng trên tập test gồm 1485 hình ảnh và được so sánh đánh giá độ chính xác cũng như hiệu quả của mô hình. Kết quả cho thấy 2 mô hình ResNet50 và MobileNet có độ chính xác khá cao và chênh lệch không đáng kể. Tuy nhiên thời gian huấn luyện của ResNet50 lâu hơn so với MobileNet và số lượng tham số nhiều dẫn đến dung lượng khá lớn. Với hướng phát triển của đề tài là triển khai trên các hệ thống ứng dụng mobile có phần cứng còn bị hạn chế. Dựa trên kết quả của các mô hình và đem đi so sánh chúng tôi lựa chọn mô hình MobileNet vì nó đáp ứng được độ chính xác và thời gian huấn luyện cũng như dung lượng mô hình nhẹ đem lại hiệu quả tốt nhất cho người dùng. Trong bài báo còn một số hạn chế là dữ liệu còn khá ít không đáp ứng được hết nhu cầu của người sử dụng và mô hình vẫn chưa đạt được độ chính xác cao nhất. Trong tương lai chúng tôi sẽ xây dựng bộ dữ liệu đa dạng hơn, nhiều địa điểm hơn, xây dựng thêm các chức năng mới cho ứng dụng và tăng cường phần cứng để huấn luyện mô hình tốt hơn.

TÀI LIỆU THAM KHẢO

- [1] TF. Sultana, A. Sufian and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), 2018, pp. 122-129, doi: 10.1109/ICRCICN.2018.8718718.
- [2] S. Albawi, T. A. Mohammed and S. Al-Zawi, "Understanding of a convolutional neural network," 2017 International Conference on Engineering and Technology (ICET), 2017, pp. 1-6, doi: 10.1109/ICEngTechnol.2017.8308186.
- [3] M.D. Zeiler, R. Fergus, *Visualizing and Understanding Convolutional Networks*, In: Fleet, D., Pajdla, T., Schiele, B., Tuytelaars, T. (eds) Computer Vision – ECCV 2014. ECCV 2014. Lecture Notes in Computer Science, vol 8689. Springer, Cham. https://doi.org/10.1007/978-3-319-10590-1_53.
- [4] N. Jmour, S. Zayen and A. Abdelkrim, "Convolutional neural networks for image classification," 2018 International Conference on Advanced Systems and Electric Technologies (IC_ASET), 2018, pp. 397-402, doi: 10.1109/ASET.2018.8379889.
- [5] A. Krizhevsky, Ilya Sutskever, Geoffrey E. Hinton, "ImageNet classification with deep convolutional neural networks." *Communications of the ACM* 60 (2012): 84 - 90.
- [6] S. Bouguezzi, H. Faiedh and C. Souani, "Slim MobileNet: An Enhanced Deep Convolutional Neural Network," 2021 18th International Multi-Conference on Systems, Signals & Devices (SSD), 2021, pp. 12-16, doi: 10.1109/SSD52085.2021.9429519.
- [7] M. A. E. Muhammed, A. A. Ahmed and T. A. Khalid, "Benchmark analysis of popular ImageNet classification deep CNN architectures," 2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon), 2017, pp. 902-907, doi: 10.1109/SmartTechCon.2017.8358502.
- [8] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.

YSC4F.226

THIẾT KẾ HỆ THỐNG GIÁM SÁT TỪ XA CÔNG SUẤT NĂNG LƯỢNG TÁI TẠO

VŨ VIẾT NAM, NGUYỄN LÊ VĨNH PHÁT, NGUYỄN VĂN LÀNH

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

18076191.nam@student.iuh.edu.vn, 18076431.phat@student.iuh.edu.vn, nguyenvanlanh@iuh.edu.vn

Tóm tắt. Vai trò năng lượng cho sự thúc đẩy kinh tế đã kích thích sử dụng các nguồn năng lượng tái tạo như năng lượng mặt trời, năng lượng gió và thủy điện ngày càng tăng. Tuy nhiên, xây dựng các hệ thống năng lượng tái tạo bị ràng buộc vào vị trí địa lý để có được công suất lớn. Vì thế, mà công việc thu thập nhiều dữ liệu hệ thống để giám sát và bảo trì trở nên khó khăn. Do đó, bài báo này đề xuất thiết kế và xây dựng một hệ thống giám sát từ xa các thông số dữ liệu như cường độ nắng và gió, các thông số điện áp, cường độ dòng điện và công suất của nhà máy điện mặt trời và gió. Để thực hiện nhiệm vụ này, hệ thống được thiết kế như sau. Thứ nhất, nhiều mạch cảm biến để thu thập các dữ liệu được xây dựng. Thứ hai, các dữ liệu này được truyền lên trung tâm dựa vào công nghệ IoT làm nền tảng. Thứ ba, dữ liệu có được sẽ mang đi phân tích và xử lý để cung cấp cho người dùng những thông tin thiết yếu cho việc giám sát và bảo trì. Thứ tư, xây dựng giao diện người dùng ở các thiết bị như máy tính, smartphone. Cuối cùng, mô hình và kết quả thực nghiệm được trình diễn để xem xét khả năng đáp ứng của hệ thống đề xuất.

Từ khóa. Hệ thống giám sát từ xa, năng lượng tái tạo, Internet of Things, IoT.

A REMOTE MONITORING SYSTEM DESIGN FOR RENEWABLE ENERGY CAPACITY

Abstract. The energy demand for economic development and growth in each country has promoted and increased the use of renewable energy sources such as solar energy, wind power, hydropower, ... to replace depleted fossil energy source, leading to rising costs, environmental pollution and global climate change. However, building renewable energy systems depends on geographic location to get large capacity. This makes it more difficult to collect system data for monitoring and maintenance. Therefore, this paper proposes to design and build a system that can remotely monitor data parameters such as intensity of sunlight and wind, parameters of voltage, amperage and capacity of solar and wind power plants. To perform this task, the system is designed according to the following steps. Firstly, build a sensor system to collect data for monitoring purposes. Secondly, this data is transmitted to the center for processing based on Internet of things (IoT) technology and data transfer protocols. Thirdly, the obtained data will be analyzed and aggregated to provide users with the necessary information for monitoring and maintenance. Fourthly, build a user interface that can be used easily on devices such as computers, smartphones anywhere. Finally, the experimental model and system experimental results are presented to test the effectiveness of the renewable energy capacity remote monitoring system.

Keyword. Remote monitoring system, Renewable energy, Internet of things, IoT.

1. GIỚI THIỆU

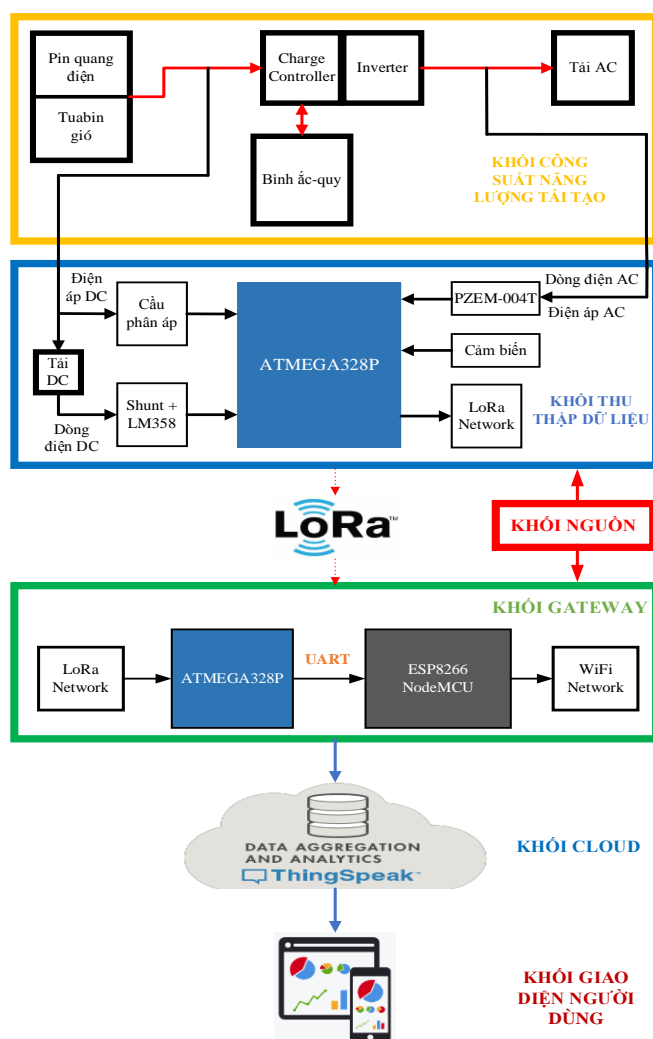
Thế giới đang trải qua quá nhiều biến đổi trên thị trường năng lượng do các nguồn năng lượng hoá thạch ngày dần cạn kiệt, do đó việc phát triển năng lượng tái tạo là giải pháp thay cho các nguồn năng lượng hoá thạch là điều tất yếu. Để khai thác năng lượng tái tạo sao cho hiệu quả, năng suất cao, cũng như bảo trì thì việc giám sát những thông số công suất của các trạm năng lượng tái tạo là cực kỳ quan trọng. Các thông số quan trọng đó bao gồm điện áp, dòng điện, công suất và có thể thêm vài thông số phụ như tần số, hệ số

năng lượng, cường độ nắng và gió của từng trạm năng lượng tái tạo. Cùng với sự bùng nổ mạnh mẽ của công nghệ Internet of Things (IoT) được ứng dụng trong nhiều lĩnh vực, trong đó không thể thiếu là lĩnh vực năng lượng. Nghiên cứu về lĩnh vực này, có một số bài báo đã công bố. Bài báo [1][2] trình bày thiết kế hệ thống giám sát hòa lưới cho mạng năng lượng mặt trời. Bài báo [3][6] nghiên cứu về hệ thống thu thập dữ liệu và cảm biến tự cung cấp nguồn cho mạng năng lượng tái tạo. Bài báo [7][16] nghiên cứu về hệ thống IoT và ứng dụng trong năng lượng mặt trời. Tuy nhiên các bài báo này chưa thu thập đầy đủ các thông số để giám sát và chưa xử lý dữ liệu sang thông tin quan trọng cho giám sát và bảo trì hệ thống. Do đó trong bài báo này, nhóm thiết kế mô hình hệ thống giám sát từ xa các trạm năng lượng tái tạo thu thập đầy đủ hơn và cải thiện tính trực quan của dữ liệu qua phân tích các biểu đồ. Ý tưởng thiết kế được xem xét các khía cạnh là thực thi dễ dàng, ít tốn kém chi phí, thu thập dữ liệu. Mô hình sẽ là một hệ thống IoT hoàn thiện gồm 3 thành phần cơ bản: phần cứng, phần mềm và truyền thông kèm nhiều tính năng mở.

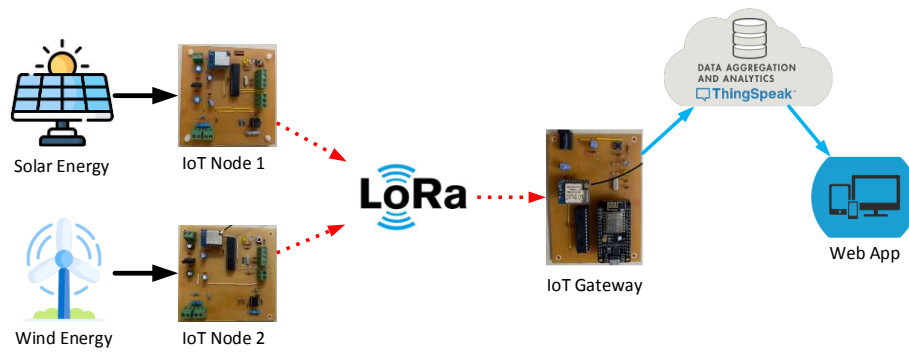
2. THIẾT KẾ HỆ THỐNG

Mô hình hệ thống thiết kế được thể hiện trong Hình 1 và hình 2 bao gồm các khối sau: khối công suất năng lượng tái tạo, khối thu thập dữ liệu, khối Gateway, khối cloud và giao diện người dùng. Dưới đây sẽ thể hiện chi tiết các khối. Ý tưởng thiết kế nền tảng mô hình dựa trên một hệ thống IoT bao gồm 4 thành phần chính: IoT Network, IoT Node, IoT Gateway và IoT Server.

2.1. Thiết kế hệ thống IoT

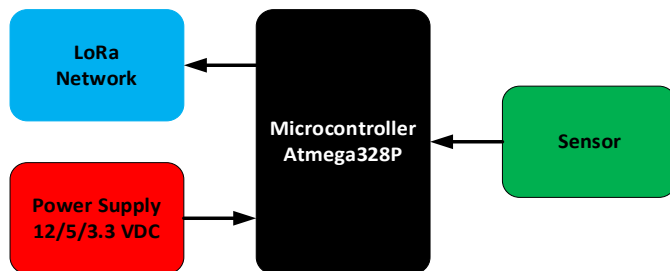


Hình 1: Sơ đồ tổng quát hệ thống giám sát

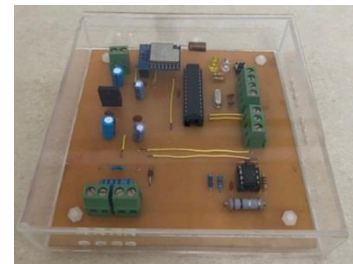


Hình 2: Sơ đồ khối hệ thống theo mô hình IoT

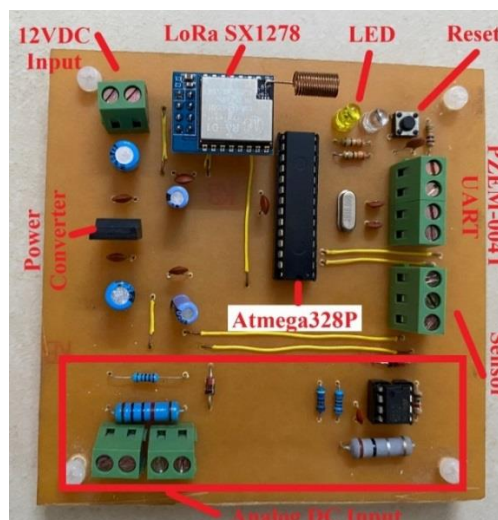
- Thiết kế IoT Networks: Trước khi thiết kế một hệ thống IoT, việc chọn lựa mạng không dây liên kết các thiết bị với nhau là công việc rất quan trọng trong thiết kế. Nó không chỉ về khả năng truyền, nhận dữ liệu mà còn xét đến các yếu tố quan trọng như khoảng cách truyền, năng lượng tiêu thụ, tần số. Hiện nay, có rất nhiều công nghệ mạng không dây sử dụng trong IoT như Bluetooth, Wi-Fi, Zigbee, ... Xét trong hệ thống của nhóm, thì việc kết nối các trạm năng lượng tái tạo ở khoảng cách xa thì mạng LoRa được lựa chọn trong thiết kế này vì những ưu điểm nổi bật về phạm vi phủ sóng, tiêu thụ năng lượng.
- Thiết kế IoT Node: Chức năng là thu thập các thông tin về điện áp, dòng điện, công suất DC và AC thông qua các giá trị tương tự (analog) ở cảm biến và truyền thông nối tiếp UART. Sau khi thu thập, IoT Node sẽ xử lý dữ liệu và truyền thông tới IoT Gateway bằng mạng không dây LoRa.



a) Sơ đồ khối của IoT Node



b) Phần cứng của IoT Node



c) Thành phần bên trong phần cứng của IoT Node

Hình 3: Sơ đồ khối (a) và phần cứng (b, c) của IoT Node

- IoT Node gồm các phần cứng: bộ xử lý trung tâm, phần thu thập dữ liệu, phần truyền thông và phần nguồn.
- Bộ xử lý trung tâm được lựa chọn là vi điều khiển Atmega328P tiêu thụ dòng điện nhỏ. Vi điều khiển này chứa 32KB Flash, 2KB SRAM, 1KB EEPROM, tốc độ xung nhịp 16MHz.
 - Phần truyền thông gồm mạng LoRa không dây và có dây UART. Phần cứng LoRa sử dụng module LoRa SX1278 được kết nối qua giao thức SPI với vi điều khiển. Truyền thông nối tiếp UART với module đo điện áp AC PZEM-004T kết nối với vi điều khiển.
 - Phần thu thập dữ liệu sẽ tiến hành việc đọc thông tin qua các giá trị tương tự analog qua cầu phân áp, bộ khuếch đại và giao tiếp UART với module PZEM-004T
 - Phần nguồn được thiết kế lấy từ bộ năng lượng lưu trữ của các trạm năng lượng tái tạo với các mức điện áp 12/5/3.3VDC phân phối điện áp hoạt động cho các phần của IoT Node.

Thiết kế mạch thu thập điện áp và dòng điện DC:

- Thông số điện áp DC được thu thập bởi vi điều khiển Atmega328P như Hình 3 và Hình 4(a). Đối với các vi điều khiển Atmega328P thì việc đọc tín hiệu tại các chân analog với mức điện áp từ 0 – 5VDC. Vì thế, sử dụng cầu phân áp để đọc các giá trị điện áp cao hơn 5VDC để vi điều khiển có thể đọc được.

Điện áp ra của cầu phân áp được tính theo công thức sau:

$$V_{out} = \frac{V_{cc} \times R_5}{R_4 + R_5} \quad (1)$$

Sau đó, điện áp ra này được xây dựng chương trình con tính toán.

Thông số dòng điện DC được thu thập bởi sơ đồ mạch hình 4(b). Sử dụng tải DC mắc nối tiếp với điện trở Shunt. Tiếp đến, đi qua bộ khuếch đại thuật toán opamp LM358 đưa vào vi điều khiển đọc.

Công thức tính toán điện áp ra tại chân out 1 của IC op-amp LM358 mạch khuếch đại không đảo:

$$V_{out} = \left(1 + \frac{R_8 + R_9}{R_{10}}\right) V_{in} = \left(1 + \frac{R_8 + R_9}{R_{10}}\right) V_{shunt} \quad (2)$$

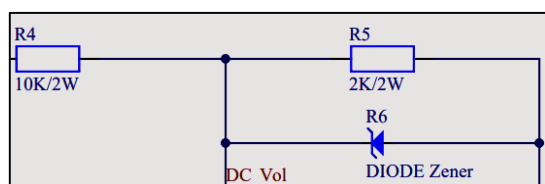
$$\Rightarrow V_{shunt} = V_{out} \div \left(1 + \frac{R_8 + R_9}{R_{10}}\right) \Rightarrow I_{shunt} = \frac{V_{shunt}}{R_{shunt}} = \frac{V_{shunt}}{0.1} (A) = I_{Load} \quad (3)$$

Thu thập công suất DC: Để thu thập được công suất DC, thì sử dụng công thức để tính toán công suất

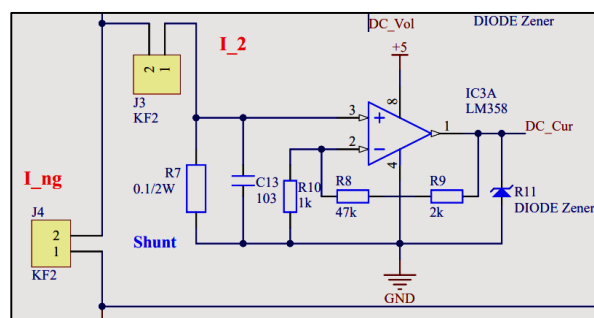
$$P = U \times I \quad (4)$$

Trong đó: U là điện áp DC và I dòng điện DC vừa thu thập được.

Thu thập dữ liệu AC: Sử dụng module PZEM-004T ở Hình 3 để thu thập 6 thông số bao gồm: điện áp, dòng điện, công suất hoạt động, tần số, hệ số công suất và năng lượng hoạt động. Thông qua giao tiếp UART với vi điều khiển Atmega328P.



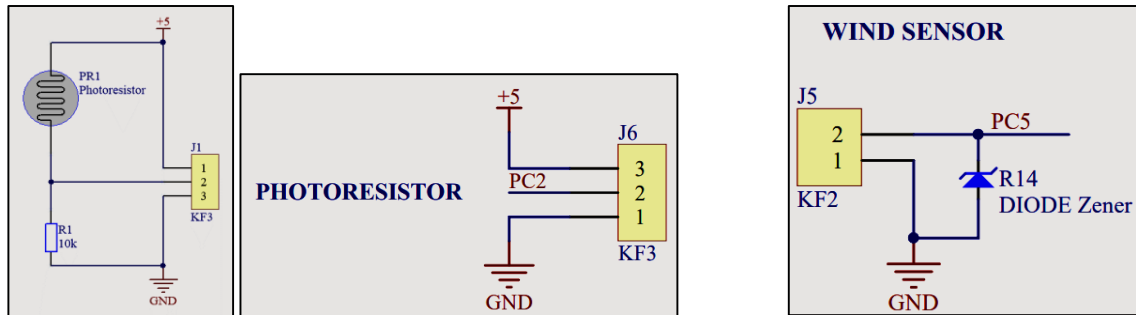
a) Sơ đồ mạch thu thập điện áp DC



b) Sơ đồ mạch thu thập dòng điện DC

Hình 4: Sơ đồ thiết kế mạch thu thập điện áp và dòng điện DC

Thiết kế mạch thu thập dữ liệu nắng và gió: Cường độ nắng và gió được thu thập bằng cách đọc giá trị tương tự từ mạch cảm biến phát hiện nắng như trong Hình 5(a) và tuabin gió phụ như trong Hình 5(b). Tiếp đến các giá trị này so sánh với ngưỡng analog được thiết lập để nhận biết có hay không có nắng/ gió.

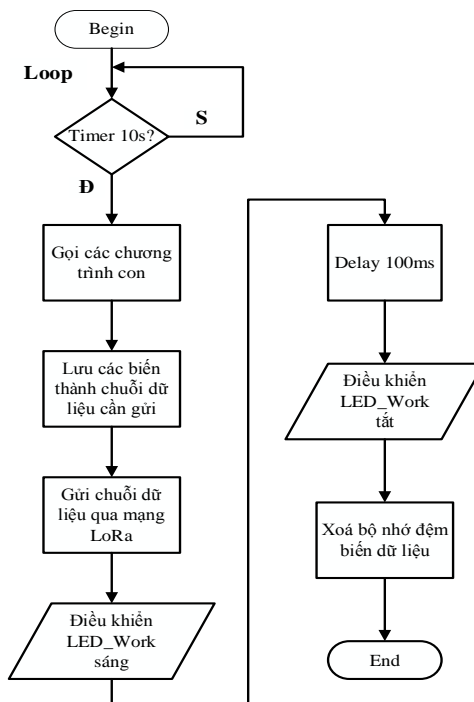


a) Sơ đồ mạch phát hiện nắng và đầu nối với vi điều khiển Atmega328P

b) Sơ đồ đầu nối từ tuabin gió phụ với vi điều khiển Atmega328P

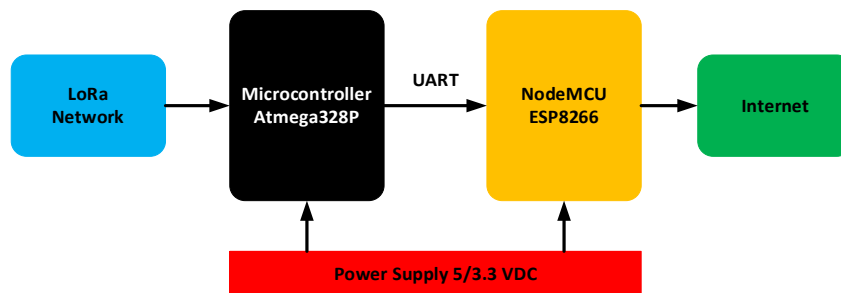
Hình 5: Sơ đồ mạch thu thập dữ liệu nắng và gió

Lưu đồ giải thuật gửi dữ liệu qua truyền thông LoRa của IoT Node được thể hiện qua Hình 6 như sau:



Hình 6: Giải thuật cho IoT Node

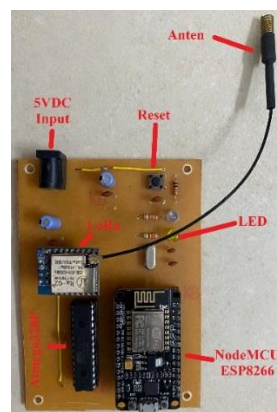
Thiết kế IoT Gateway: Sau khi các IoT Node gửi dữ liệu thì module IoT Gateway sẽ nhận dữ liệu này qua mạng LoRa không dây, xử lý các dữ liệu này và sau đó gửi lên Server ThingSpeak.



a) Sơ đồ khối của IoT Gateway



b) Phần cứng của IoT Gateway



c) Thành phần bên trong phần cứng của IoT Gateway

Hình 7: Sơ đồ khối (a) và phần cứng (b, c) của IoT Gateway

IoT Gateway có các phần cứng: Bộ nhận dữ liệu, bộ gửi dữ liệu, phần truyền thông và phần nguồn.

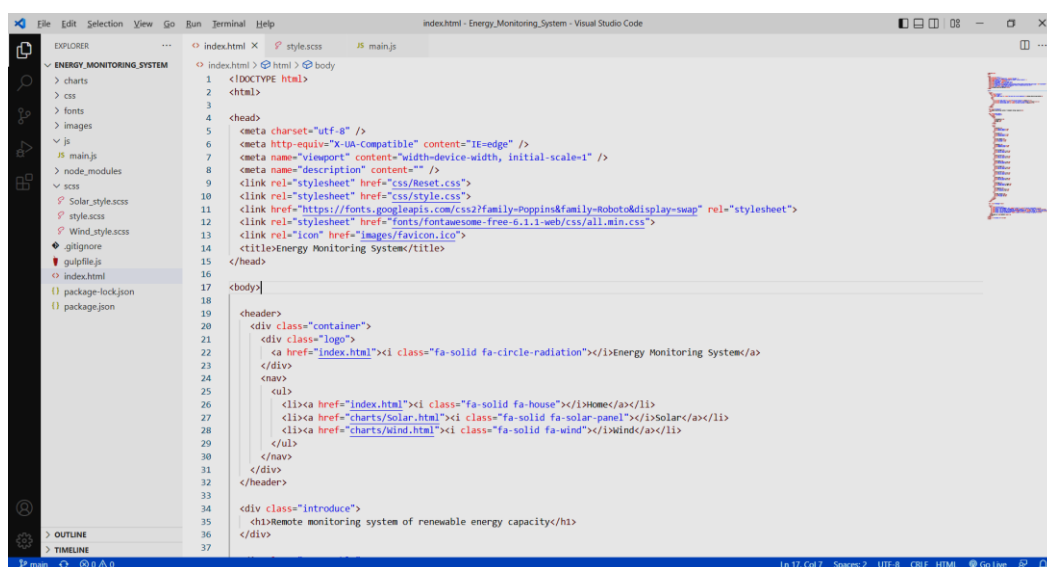
- Bộ nhận dữ liệu được lựa chọn là vi điều khiển Atmega328P.
- Bộ gửi dữ liệu được lựa chọn là module Wi-Fi ESP8266 CP2102 NodeMCU LUA.
- Phần truyền thông là các mạng không dây LoRa, truyền thông nối tiếp UART và mạng Internet Wi-Fi.
- Phần nguồn được thiết kế lấy từ Adapter 220VAC – 5VDC phân phối điện áp hoạt động tới các phần IoT Gateway.

Lựa chọn IoT Server: Ngày nay, có rất nhiều nền tảng đám mây lưu trữ dữ liệu đóng góp cho các công nghệ IoT. Ở đây, nhóm sử dụng một ứng dụng cung cấp mã nguồn mở cho các dự án IoT đó là nền tảng miễn phí Thingspeak. Nó cung cấp các API tích lũy và truy xuất dữ liệu từ thiết bị bằng giao thức HTTP thông qua đường truyền Internet mà nhóm đang sử dụng.

2.2. Thiết kế giao diện người dùng

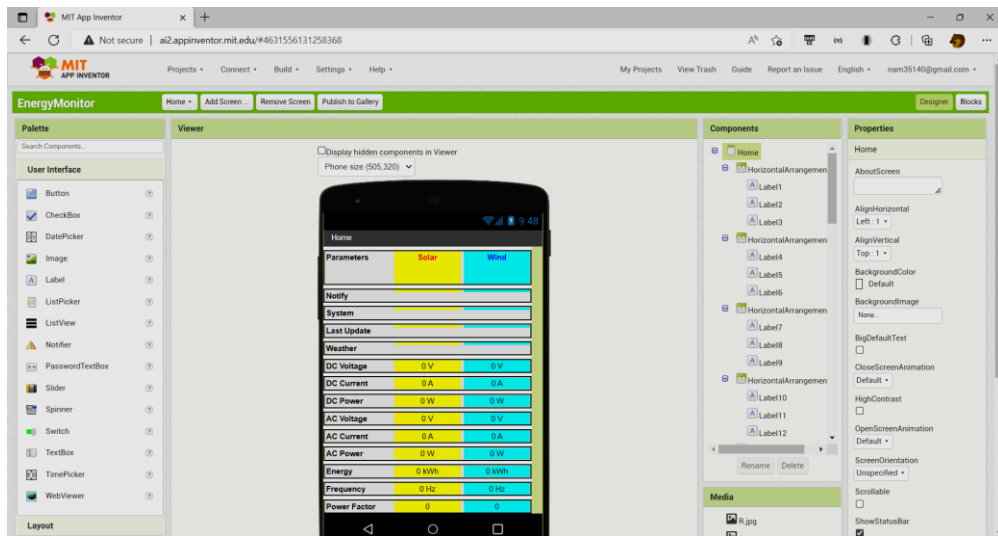
Xây dựng giao diện người dùng cho hệ thống giám sát từ xa bao gồm 2 giao diện là Web và App được minh họa ở các Hình 8 và hình 9.

Thiết kế giao diện Web: Giao diện Web được thiết kế đơn giản, cơ bản cho người dùng bằng Front-end: HTML, CSS, JavaScript được minh họa như Hình 8.



Hình 8: Sử dụng phần mềm Visual Studio Code để code giao diện Web

Thiết kế giao diện App Android: Sử dụng ứng dụng MIT App Inventor để tạo ra app mobile trên hệ điều hành Android được minh họa như Hình 9.

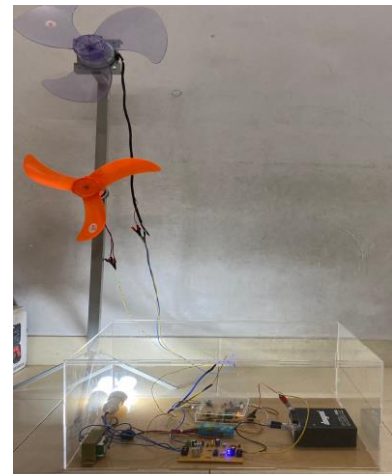


Hình 9: Ứng dụng Web của MIT App Inventor để tạo app Android

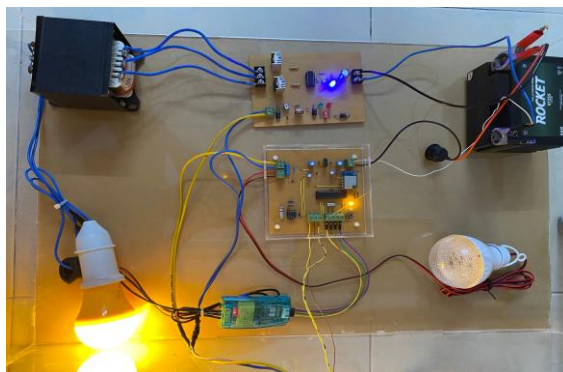
2.3. Mô hình thực nghiệm



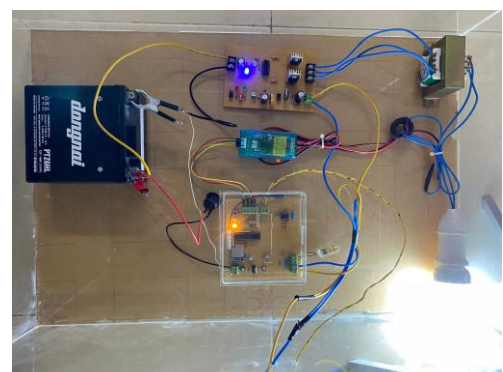
a) Mô hình năng lượng mặt trời



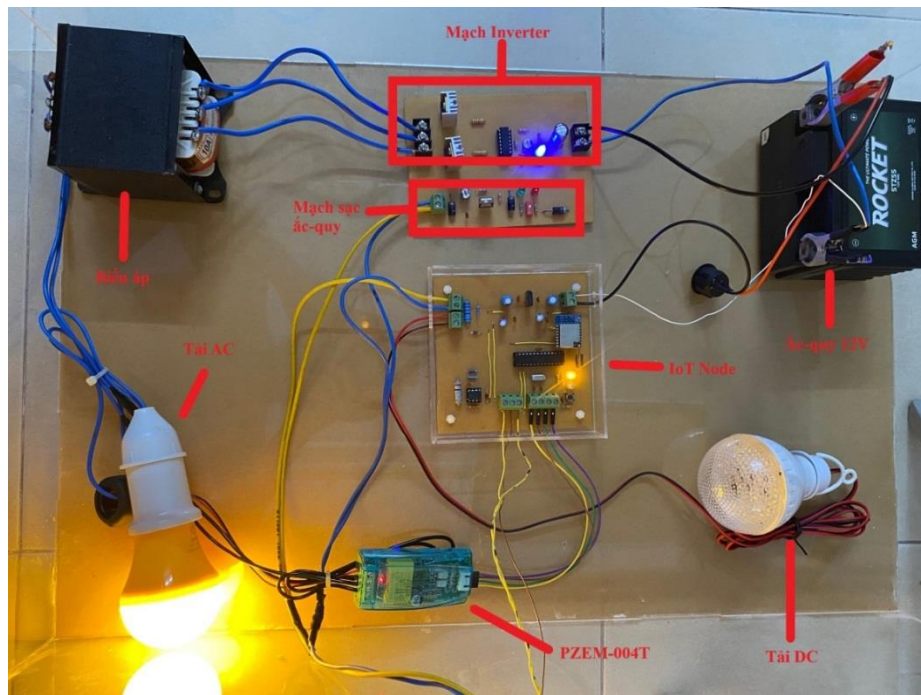
b) Mô hình năng lượng gió



c) Bên trong mô hình thu thập năng lượng mặt trời



d) Bên trong mô hình thu thập năng lượng gió



e) Các thành phần của mô hình thu thập

Hình 10: Mô hình thực nghiệm

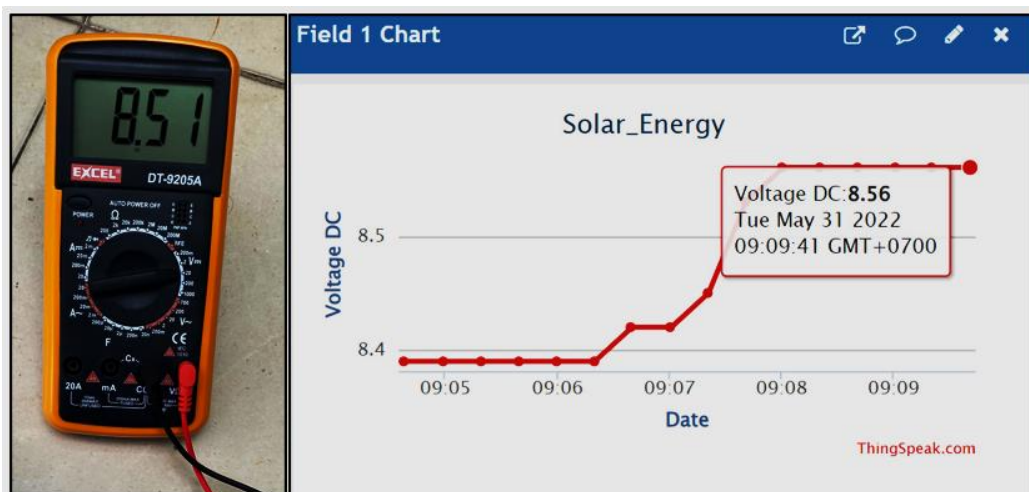
Mô hình thực nghiệm hoàn chỉnh được thể hiện trong Hình 10. Mô hình bao gồm hệ thống năng lượng mặt trời trong Hình 10(a) và hệ thống năng lượng gió trong Hình 10(b), phần cứng của các khối công suất năng lượng tái tạo, thu thập dữ liệu, khối IoT, khối giao diện người dùng, khối lưu trữ năng lượng và khối nguồn trong các Hình 10(c)-(e).

3. KẾT QUẢ THỰC NGHIỆM

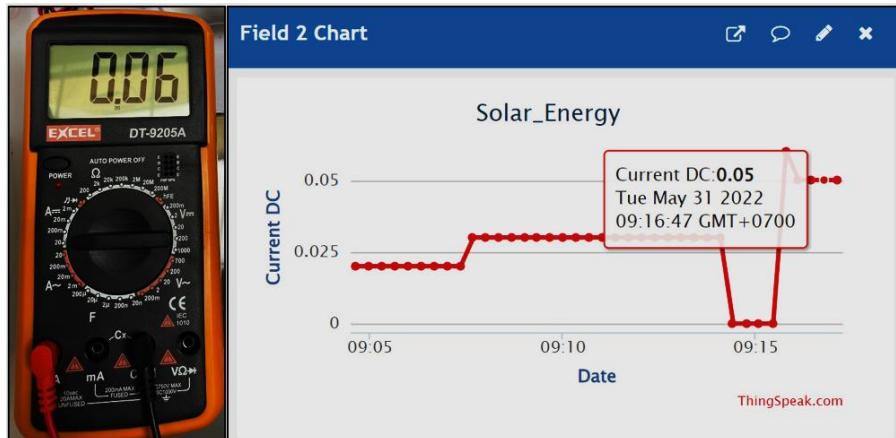
Sau khi thiết kế các bo mạch hoàn chỉnh, thực nghiệm kiểm tra các chức năng cụ thể là thu thập dữ liệu, truyền nhận dữ liệu qua mạng không dây LoRa được thực hiện để đánh giá độ chính xác của hệ thống vừa thiết kế.

3.1. Thực nghiệm khả năng thu thập dữ liệu

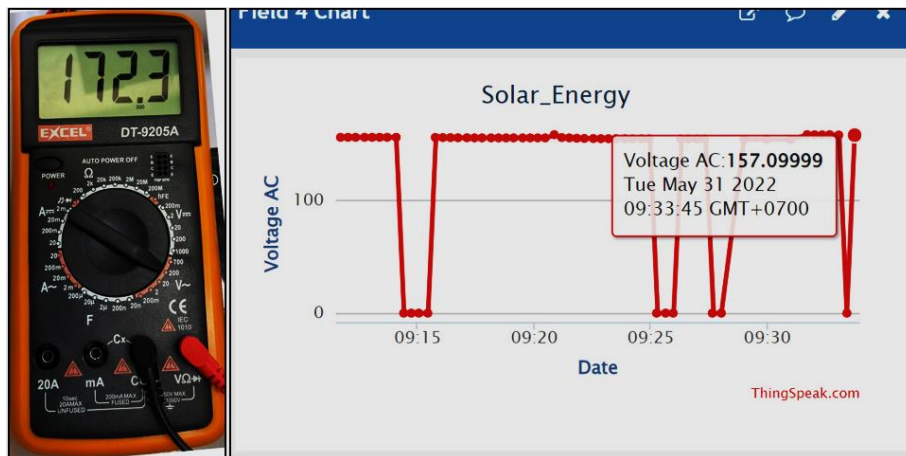
Kết quả thực nghiệm thu thập dữ liệu của mô hình thiết kế được thể hiện trong các Hình (11 – 14).



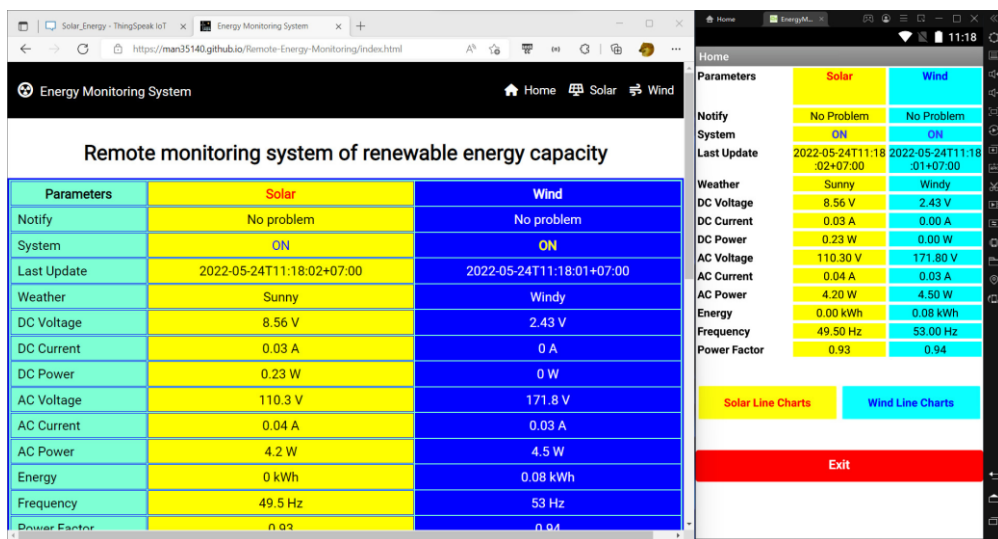
Hình 11: Thu thập dữ liệu điện áp DC



Hình 12: Thu thập dữ liệu dòng điện DC



Hình 13: Thu thập dữ liệu điện áp AC



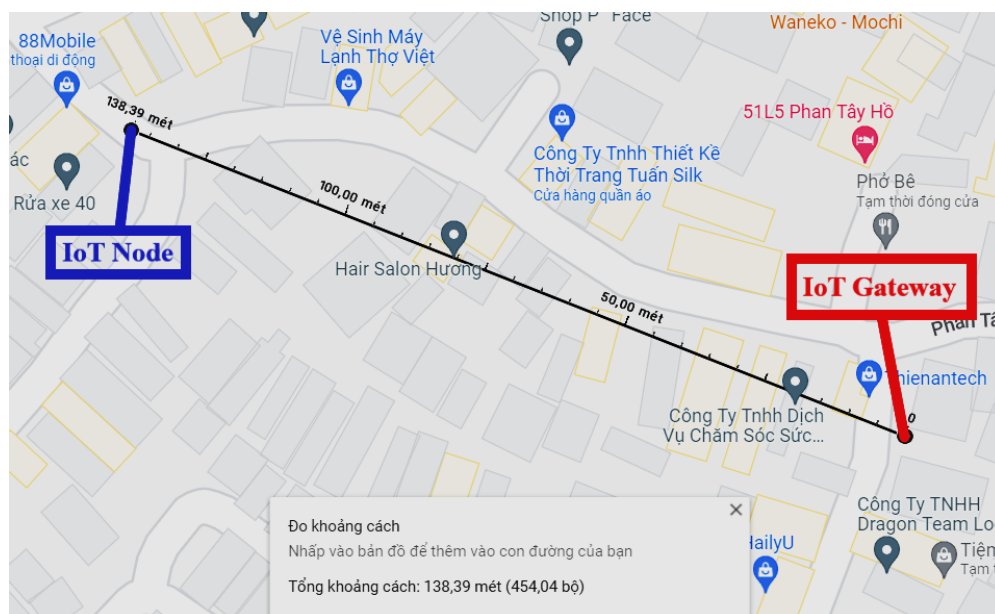
Hình 14: Giao diện giám sát Web và App các thông số của hệ thống

Hình 11 thể hiện kết quả thu thập dữ liệu điện áp DC. Hình 12 thể hiện kết quả thu thập dữ liệu dòng điện DC. Hình 13 thể hiện kết quả thu thập dữ liệu điện áp AC. Hình 14 thể hiện kết quả giao diện giám sát Web

và App các thông số hệ thống. Các dữ liệu về điện áp DC, dòng điện DC và điện áp AC được thu thập mỗi 20s tại phần cứng IoT Node bằng các mạch đọc giá trị analog. Sau đó, gửi các dữ liệu này đến module IoT Gateway qua mạng LoRa và tiếp tục gửi lên đám mây để lưu trữ. Ở dưới mỗi hình, bên trái là đồng hồ vạn năng số đo điện áp, dòng điện từ trạm năng lượng mặt trời để so sánh với khả năng thu thập dữ liệu của hệ thống. Kết quả cho thấy, sai số không đáng kể.

3.2. Thực nghiệm khoảng cách truyền nhận qua mạng LoRa

Tiến hành việc di dời module IoT Node để xác minh phạm vi tối đa mà mạng không dây LoRa có thể truyền dữ liệu. Khi module IoT Gateway nhận được tín hiệu từ module IoT Node truyền đến thì LED lam sẽ sáng lên. Kết quả xác minh cho thấy, khoảng cách truyền dữ liệu giữa 2 module ở trong khu vực thiết kế, có nhiều chướng ngại vật như các nhà lầu, cây cối là khoảng 138,39 mét. Hình 15 thể hiện kết quả thực nghiệm truyền nhận qua mạng LoRa.



Hình 15: Khoảng cách truyền và nhận giữa IoT Gateway và IoT Node

4. KẾT LUẬN

Bài báo trình bày ý tưởng thiết kế một hệ thống giám sát từ xa công suất năng lượng tái tạo bằng cách ứng dụng công nghệ IoT. Với nền tảng được kiến nghị bao gồm 4 thành phần là module IoT Node sử dụng chip Atmega328P, module IoT Gateway sử dụng chip Atmega328P và module NodeMCU ESP8266, IoT Network sử dụng mạng LoRa và Internet, điện toán đám mây miễn phí ThingSpeak. Thông qua các kết quả kiểm chứng, các mô hình hệ thống chạy khá tốt và ổn định, thu thập dữ liệu khá chính xác, sai số thấp, truyền nhận dữ liệu khá nhanh. Giao diện người dùng hiển thị các thông số, các thông báo khá tốt. Trong tương lai có thể cải tiến và xây dựng hệ thống hoàn thiện hơn nữa, đặc biệt là giao diện người dùng có thêm nhiều tính năng.

TÀI LIỆU THAM KHẢO

- [1] C. Lin, G. Jie, Z. Wu, and W. Rui, “Design of networked monitoring system of PV grid-connected power plant,” Proceedings of 2011 International Conference on Electronic and Mechanical Engineering and Information Technology, EMEIT 2011, 2011, vol. 3, pp. 1169-1172.
- [2] S. Tiwari and R. N. Patel, “Real Time Monitoring of Solar Power Plant and Automatic Load Control,” Engineering and Systems (SCES), 2015 IEEE Students Conference on, 2015, pp. 1-6.

- [3] A. Soetedjo, Y. I. Nakhoda, and D. Suryadi, “Development of data acquisition system for hybrid power plant,” 2013 13th International Conference on Quality in Research, QiR 2013 - In Conjunction with the 2nd International Conference on Civic Space, ICCS 2013, 2013, pp. 197-201.
- [4] D. Delimustafic, J. Islambegovic, A. Aksamovic, and S. Masic, “Model of a hybrid renewable energy system: Control, supervision and energy distribution,” 2011 IEEE International Symposium on Industrial Electronics, 2011, pp. 1081-1086.
- [5] P. Guerriero, F. Di Napoli, G. Vallone, V. Dalessandro, and S. Daliento, “Monitoring and diagnostics of PV plants by a wireless self-powered sensor for individual panels,” IEEE Journal of Photovoltaics, vol. 6, no. 1, pp. 286-294, 2016.
- [6] P. Guerriero, G. Vallone, M. Primato, F. Di Napoli, L. Di Nardo, V. D’Alessandro, and S. Daliento, “A wireless sensor network for the monitoring of large PV plants,” 2014 International Symposium on Power Electronics, Electrical Drives, Automation and Motion, 2014, pp. 960-965.
- [7] A. Tsagaris and D. G. Triantafyllidis, “Data monitoring system for supervising the performance assessment of a photovoltaic park,” CINTI 2012 - 13th IEEE International Symposium on Computational Intelligence and Informatics, Proceedings, 2012, pp. 385-389.
- [8] Naser Hossein Motlagh, Mahsa Mohammadrezaei, Julian Hunt, Behnam Zakeri, *Internet of Things (IoT) and the Energy Sector*, 2020.
- [9] Jayaharsha Kandimalla, Dr.D.Ravi Kishore, *Web Based Monitoring of Solar Power Plant Using Open Source IOT Platform Thingspeak and Arduino*, 2017.
- [10] Z. Atika, W. Z. Leow, I. Iszaidy, Y.M. Irwan, I. Safwati, M. Irwanto, N.M.Wafi, S. X. Saw, *Development A Portable Solar Energy Measurement System*, 2021.
- [11] Youssef Cheddadi, Hafsa Cheddadi, Fatima Cheddadi, Fatima Errahimi, Najia Es-sbai, *Design and implementation of an intelligent low-cost IoT solution for energy monitoring of photovoltaic stations*, 2020.
- [12] W Priharti, A F K Rosmawati, I P D Wibawa, *IoT based photovoltaic monitoring system application*, 2019.
- [13] Manish Katyarmal, Suyash Walkunde, Arvind Sakhare, Mrs.U.S.Rawandale, *Solar power monitoring system using IoT*, 2018.
- [14] Nguyen Phuoc Hoang Khang, Nguyen Chi Nhan, *Solar power monitoring system based on IoT technology*, 2021.
- [15] N. T. M. Hiền, *Các nguồn năng lượng tái tạo tối sự phát triển bền vững và giảm thiểu biến đổi khí hậu*, 2022.
- [16] Pham Quang Tri, Nguyen Ngoc Son, Cao Van Kien, *Design IoT Lab Platform*, Applied in University Curricula, 2020.

YSC4F.227

HOUFLEX: SEARCHING & SMART MANAGEMENT BOARDING HOUSE APPLICATION BASED ON BLOCKCHAIN MECHANISM

NGUYỄN CÔNG THÀNH ĐẠT, MAI KIÊN CƯỜNG, LƯU TUẤN KHA, TRƯƠNG CÔNG CƯỜNG,
TÔN LONG PHƯỚC

Khoa Công nghệ thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
datma113112111@gmail.com, luutkha@gmail.com, truongcongcuon0302@gmail.com,
maikiencuongiuh@gmail.com, tonlongphuoc@iuh.edu.vn

Tóm tắt. Tìm kiếm phòng để thuê hoặc cho thuê ở một đô thị lớn là một trong những xử lý rất phổ biến trong cuộc sống. Hơn nữa, trước xu thế chuyển đổi số hiện nay, việc xử lý các nghiệp vụ này với cơ chế thông minh vẫn là một thử thách. Vì thế, chúng tôi đề xuất ứng dụng có tên gọi là Houflex. Houflex là một ứng dụng hỗ trợ tìm kiếm phòng và quản lý phòng thông minh. Để đáp ứng vấn đề quản lý phòng, chúng tôi sử dụng cơ chế Blockchain cho việc tạo hợp đồng thông minh (smart contract). Cơ chế này cần sự đồng thuận giữa 2 bên và xác thực thông qua các khối (block). Khi các điều khoản được đồng thuận, tiền sẽ được chuyển tự động mà không cần qua trung gian. Qua đó, hệ thống sẽ tự động quản lý các hợp đồng thuê phòng một cách rõ ràng và minh bạch. Chúng tôi tập trung phát triển 2 phân hệ là ứng dụng Web và ứng dụng trên thiết bị di động. Houflex được triển khai đồng bộ với nhau nhằm tăng cường sự tiện ích và dễ dàng hơn cho người sử dụng. Công nghệ mà chúng tôi lựa chọn cho việc phát triển Houflex như sau: Phía back-end, sử dụng ngôn ngữ lập trình Java với Springboot framework kết hợp với NodeJS. Phía front-end, chúng tôi sử dụng ngôn ngữ Javascript với React framework cho web và React-Native cho di động. Ứng dụng sẽ được triển khai trên nền tảng cloud với các dịch vụ của AWS và Docker. Hiện tại, ứng dụng sẽ được triển khai cho các bạn sinh viên trong trường IUH với phạm vi tìm kiếm ở khu vực TP. HCM. Trong tương lai, chúng tôi sẽ mở rộng tìm kiếm cho toàn quốc và hỗ trợ tích hợp thêm nhiều tiện ích cho việc tìm kiếm phòng và quản lý hợp đồng thông minh.

Từ khóa. Ứng dụng Web, Ứng dụng di động, Hợp đồng thông minh, Chuỗi khối, Điện toán đám mây.

HOUFLEX: SEARCHING & SMART MANAGEMENT BOARDING HOUSE APPLICATION BASED ON BLOCKCHAIN MECHANISM

Abstract. Looking for rooms to rent or to rent out rooms in a large metropolis is one of the very common life handles. Moreover, in the face of the current digital transformation trend, handling these transactions with intelligent mechanisms is still a challenge. We recommend an app called Houflex. Houflex is an application that supports room search and smart room management. To meet the room management problem, we use Blockchain mechanism for smart contract creation. This mechanism needs consensus between two parties and validation through blocks. Once the terms are agreed, the funds will be transferred automatically without the need for an intermediary. Thereby, the system will automatically manage the room rental contracts in a clear and transparent way. We focus on developing 2 modules: Web application and Mobile application. Houflex is deployed synchronously with each other to increase convenience and ease for users. The technology that we choose for Houflex development is as follows: Back-end, using Java programming language with Springboot framework combined with NodeJS. On the Front-end, we use Javascript language with React framework for web and React-Native for mobile. The application will be deployed in the cloud with AWS and Docker services. Currently, the application will be deployed for students in IUH with the search scope in the HCMC area. In the future, we will expand the search to the whole country and support the integration of more utilities for room search and smart contract management.

Key words. Web Application, Mobile Application, Smart Contract, Blockchain, Cloud Computing

1. GIỚI THIỆU TỔNG QUAN

Hiện nay, nhu cầu tìm kiếm thuê và cho thuê phòng trọ là một trong những nhu cầu phổ biến, quen thuộc và có phân cấp thiết đối với nhiều người. Đặc biệt là đối với sinh viên, hay những người mới từ các tỉnh thành khác đến sinh sống và làm việc ở Tp. Hồ Chí Minh. Việc tìm được một căn phòng hay nhà phù hợp, trong khi thiếu thông tin về đường sá, vị trí, văn hoá và giá cả còn hạn chế là một thách thức lớn. Song song đó, việc tìm phòng, thương lượng, ký hợp đồng cũng tiêu tốn rất nhiều thời gian. Mặt khác, đối với những người có nhu cầu cho thuê ở Tp. Hồ Chí Minh, việc quản lý một lúc nhiều căn hộ, nhiều phòng trọ khác nhau, nhiều hợp đồng, nhiều thỏa thuận các chỉ số dịch vụ đi kèm với dịch vụ cho thuê như tiền điện, tiền nước, ... thật sự gặp nhiều rủi ro và tốn kém. Cuối cùng, việc thanh toán hoá đơn trong các giao dịch hàng tháng cũng còn nhiều hạn chế. Do đó, việc quản lý một căn hộ với các dịch vụ kèm theo là một khối lượng công việc cần được hỗ trợ tự động từ các công cụ phần mềm hiện nay. Cũng tương tự như người thuê, thời gian đối với những chủ nhà bận rộn thì càng quý báu hơn. Vì thế mà việc rút ngắn thời gian cho việc ký kết hợp đồng, ghi lại các chỉ số dịch vụ và gửi, nhận hoá đơn cũng như là các nghiệp vụ vô cùng cấp thiết. Trên thực tế, nhiều chủ nhà phải thuê riêng một nhân sự quản lý cho mỗi căn hộ cho thuê của mình, để đảm bảo việc vận hành của các dịch vụ được ổn định. Hiện nay, những dịch vụ như cho thuê nhà trên hệ thống Airbnb cũng được phát triển mạnh ở các nước phát triển cũng như tại Việt nam. Những mô hình thành công này cũng là nguồn cảm hứng cho các hệ thống dịch vụ tìm kiếm cho thuê phòng tại Tp. Hồ Chí Minh sớm thành hiện thực.

Từ những khảo sát trên, chúng tôi đã đề xuất ứng dụng Houflex, với 2 nền tảng là website và ứng dụng di động nhằm hỗ trợ cho bài toán trên. Houflex giúp hỗ trợ cho người dùng tìm kiếm một phòng trọ trong một khu vực nhất định với các tiêu chí phù hợp với nhu cầu. Chúng tôi trực quan hóa các phòng trọ và các thông tin đi kèm lên trên bản đồ cho người dùng có thể tương tác trực tiếp với hệ thống. Mọi thông tin về phòng như: diện tích phòng, giá cả, vị trí, xếp loại, chúng tôi sẽ hiển thị tất cả kết quả lên trên một tấm bản đồ để người dùng có thể dễ dàng hình dung và biết được chính xác vị trí mà mình muốn thuê ở đâu, giá cả thế nào... Hơn nữa, chúng tôi còn dùng đồ họa để phân loại các phòng trọ theo màu sắc, đánh giá trực quan hóa các thông tin.

Với hai hệ thống trên, chúng tôi đề xuất hướng giải quyết cho các khớp lệnh trong các giao dịch thỏa thuận giữa bên cho thuê và khách hàng thuê bằng công nghệ Blockchain. Những thỏa thuận này chúng tôi đưa vào các hợp đồng thỏa thuận thông minh (smart contract). Hợp đồng thông minh này sẽ dựa trên các điều khoản mà hai bên đã được đọc và ký kết. Thông qua điều khoản đó, nếu người thuê chấp nhận thì sẽ đi đến ký hợp đồng thuê phòng. Khi mọi thứ đã thanh toán thành công, hợp đồng sẽ được thành lập. Nó sẽ dựa vào các thỏa thuận mà 2 bên đã đồng ý mà thực thi các ràng buộc đã được tự động hóa. Những thỏa thuận trong hợp đồng được thực thi trong từng tháng sẽ được lưu thành các chuỗi khối (block) nhằm đảm bảo tính minh bạch cho các giao dịch này. Hệ thống Houflex cũng dựa vào các block này nhằm đảm bảo khớp lệnh các giao dịch chuyên khoản cho các bên liên quan. Hệ thống chúng tôi cũng đề xuất ví điện tử thông qua dịch vụ VNPAY nhằm giúp giao dịch trở nên thuận tiện và hiệu quả.

2. CÁC CÔNG NGHỆ LIÊN QUAN

2.1. Những framework hỗ trợ phát triển ứng dụng

2.1.1. Spring-boot Framework

Spring Boot là một framework của ngôn ngữ lập trình java [1]. Nó giúp cho việc hỗ trợ cấu hình các quá trình phát triển của một ứng dụng sử dụng spring. Điều này giúp cho lập trình viên tập trung vào được các logic nghiệp vụ của một dự án phần mềm mà không phải lo nghĩ tới việc cấu hình các dependency mà dự án cần. Đây là một framework phổ biến và được sử dụng nhiều nhất đối với các lập trình viên sử dụng ngôn ngữ Java. Hơn nữa, với cộng đồng đóng góp to lớn, mọi câu hỏi thắc mắc về framework này gần như đã có đầy đủ câu trả lời cũng như tài liệu.

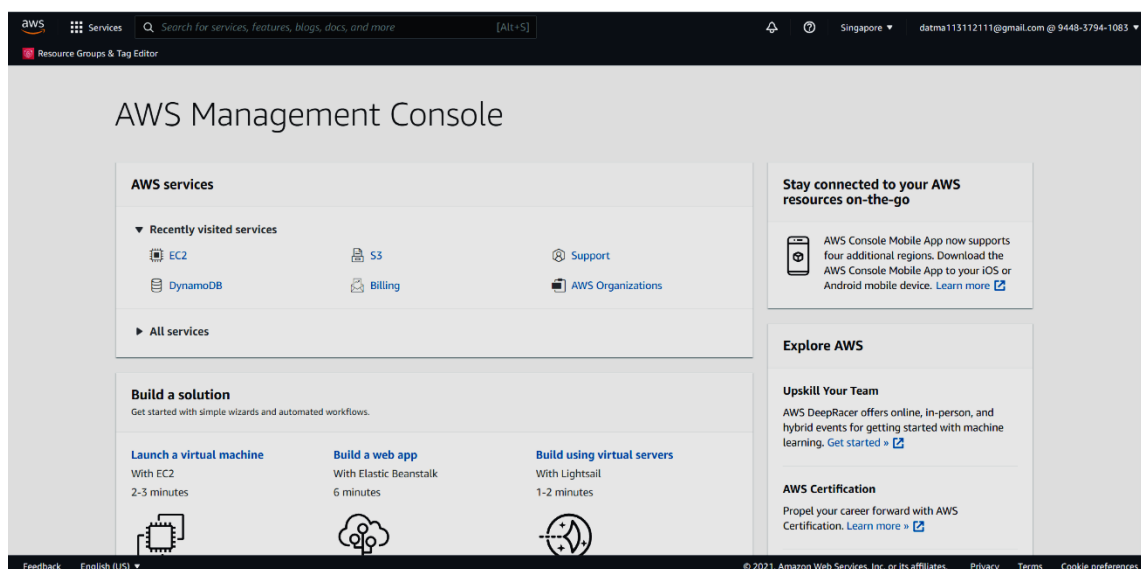
2.1.2. ReactJS - React-Native

ReactJS là một thư viện của ngôn ngữ lập trình Javascript [2]. Thư viện này hỗ trợ cho việc tạo giao diện người dùng cùng các xử lý tương tác của người dùng với giao diện. ReactJS hoạt động tương tự như một framework nên mọi người thường xếp ReactJS vào framework. Nó giúp cho việc xây dựng các xử lý logic và vấn đề hiển thị dữ liệu của một trang web trở nên dễ dàng hơn bao giờ hết. Đối với các xử lý của JavaScript về DOM (document object model) có thể vô cùng khó khăn thì đối với ReactJS, nó trở nên cực kỳ đơn giản. React Native là một framework do công ty công nghệ nổi tiếng Facebook phát triển nhằm mục đích giải quyết bài toán hiệu năng của Hybrid Application và bài toán chi phí khi mà phải viết nhiều loại ngôn ngữ riêng cho từng nền tảng di động [3].

2.1.3. Cloud computing

2.1.3.1. EC2

Amazon Elastic Compute Cloud (Amazon EC2) là dịch vụ web cung cấp các nền tảng máy tính rộng và sâu nhất với đa dạng các tiện ích về: hệ điều hành, lưu trữ, cấu hình máy... Có thể nói, EC2 là một máy ảo đám mây toàn diện nhất mà Amazon cung cấp cho người dùng có kết nối mạng ethernet lên đến 400 Gbp/s [4].



Hình 2.1.1 Màn hình console manager của AWS

2.1.3.2. S3

Amazon S3 là một trong những dịch vụ của AWS cung cấp cho người dùng khả năng lưu trữ và truy xuất dữ liệu trên nền tảng cloud. Điều này nghĩa là tất cả các file sẽ được lưu trữ trên hệ thống của AWS và người dùng ở bất cứ đâu, nơi nào cũng đều có thể truy xuất được đến nó. Không những vậy, Amazon S3 còn cung cấp khả năng bảo mật tốt và hiệu suất sử dụng cực kỳ hiệu quả. Chính vì thế mà đây là dịch vụ rất được ưa chuộng cho các nhà phát triển (developer) sử dụng. Về chi phí cho việc sử dụng dịch vụ này thì cực kỳ thấp, đồng thời khả năng sao lưu dự phòng và bảo mật của dữ liệu cũng được đánh giá cao cho dịch vụ này [5].

2.1.3.3. CloudFront

Amazon CloudFront (Amazon CF) là một trong những dịch vụ thay đổi domain của AWS. Nói chính xác hơn, Amazon CF giúp cho một trang web có thể truyền tải tới người truy cập một cách nhanh nhất dựa trên các data center gọi là các vùng biên (edge location). Amazon CF thường tích hợp chung với các dịch vụ lưu trữ, gần gũi và phổ biến nhất có lẽ là Amazon S3 bucket. Amazon CF sẽ truy vấn đến nguồn bucket mà người dùng đã khai báo để lấy thông tin và trả về một trang domain mới cho người truy cập trang web [6].

2.1.3.4. CodeCommit

AWS CodeCommit được hiểu là dịch vụ kiểm soát nguồn được quản lý, bảo mật để lưu trữ các kho Git riêng tư và với quy mô rất linh hoạt. Các đội ngũ công tác một cách bảo mật để cải thiện mã ngay cả khi ở trạng thái truyền và lưu trữ thì AWS CodeCommit là dịch vụ đáng để sử dụng [5]. Với dịch vụ này người dùng sẽ không cần phải tự quản lý hệ thống kiểm soát nguồn riêng hay việc thay đổi quy mô cơ sở hạ tầng. Nó giúp lưu trữ mọi thứ cả về mã cho đến giá trị nhị phân. Bởi các tính năng đó mà AWS Codecommit là sự lựa chọn được đánh giá cao, phối hợp với các công cụ khác dựa trên Git mà chúng ta đang sử dụng [7].

2.1.3.5. CodePipeline

AWS CodePipeline là dịch vụ phân phối liên tục giúp người dùng tự động hóa các Pipeline phát hành dành cho các bản cập nhật ứng dụng và cơ sở hạ tầng nhanh chóng, ổn định. Việc xây dựng, thử nghiệm và phát hành sẽ được CodePipeline tự động hóa một cách tốt nhất dựa trên mô hình phát hành đã được người dùng tạo ra trước đó. Điều này giúp người dùng phân phối một cách nhanh chóng và ổn định giữa các tính năng và bản cập nhật. Ngoài ra, AWS CodePipeline có thể tích hợp các dịch vụ của Github. Đây là công cụ CI/CD tích hợp cùng với AWS CodeCommit [8].

2.1.3.6. Elastic Load Balancing

Elastic Load Balancing (ELB) là dịch vụ của AWS sử dụng để cấu hình cho việc phân bổ số lượng truy cập của người dùng tới địa chỉ của trang web của người dùng. ELB sẽ tự động tạo ra nhiều vùng để có thể tự động phân chia khi người dùng truy cập vào website đang sẵn sàng hoạt động (Available Zone) [9].

2.1.3.7. Blockchain

Blockchain là một công nghệ mới về lưu trữ dữ liệu. Đây là công nghệ mã hóa tất cả dữ liệu thành các khối (gọi là block). Các khối này liên kết với nhau để tạo thành một chuỗi dài. Mỗi khi có một thông tin hay một giao dịch mới xảy ra, thông tin cũ không mất đi. Thay vào đó là thông tin mới sẽ được lưu vào một khối mới và lần lượt được nối vào khối cũ để tạo thành một chuỗi mới. Ưu điểm hơn, Blockchain có thể triển khai trên nhiều máy chủ một cách tự động và sao lưu lại trên đó kết nối đồng thời với hệ thống chính. Vì thế mà mọi người đều có thể xem và kiểm tra giao dịch của mình. Cũng chính vì lý do đó, Blockchain ngăn chặn được các tác vụ sửa đổi, giả mạo nhằm đảm bảo tính minh bạch và chính xác nhất trong các giao dịch điện tử. Vì các lý do trên mà Blockchain được ứng dụng cực kỳ rộng rãi trong nhiều lĩnh vực, ngành nghề khác nhau [10].

Trong hệ thống Houflex, chúng tôi sử dụng Blockchain để lưu lại các lịch sử giao dịch, các hợp đồng. Dựa vào lịch sử giao dịch đó, hệ thống có thể truy ra được các vết của các khối block trước đó. Điều này có thể đảm bảo được tính toàn vẹn của dữ liệu. Khi xuất ra các lịch sử giao dịch, hệ thống sẽ xem xét các khối block có hợp lệ hay không? Có nghĩa là hệ thống sẽ kiểm tra mã hash của toàn bộ khối. Nếu một chuỗi khối (Blockchain) hợp lệ - tất cả mã hash trước của khối là hợp lệ thì có thể xác định, các giao dịch đó là hợp lệ. Đồng thời, hợp đồng giao dịch đó cũng là hợp lệ và việc khớp lệnh cho các giao dịch này được thực thi trên các ví điện tử [11].

3. HOUFLEX: ỨNG DỤNG HỖ TRỢ TÌM KIẾM PHÒNG VÀ QUẢN LÝ PHÒNG TRỢ DỰA TRÊN NỀN TẢNG BLOCKCHAIN

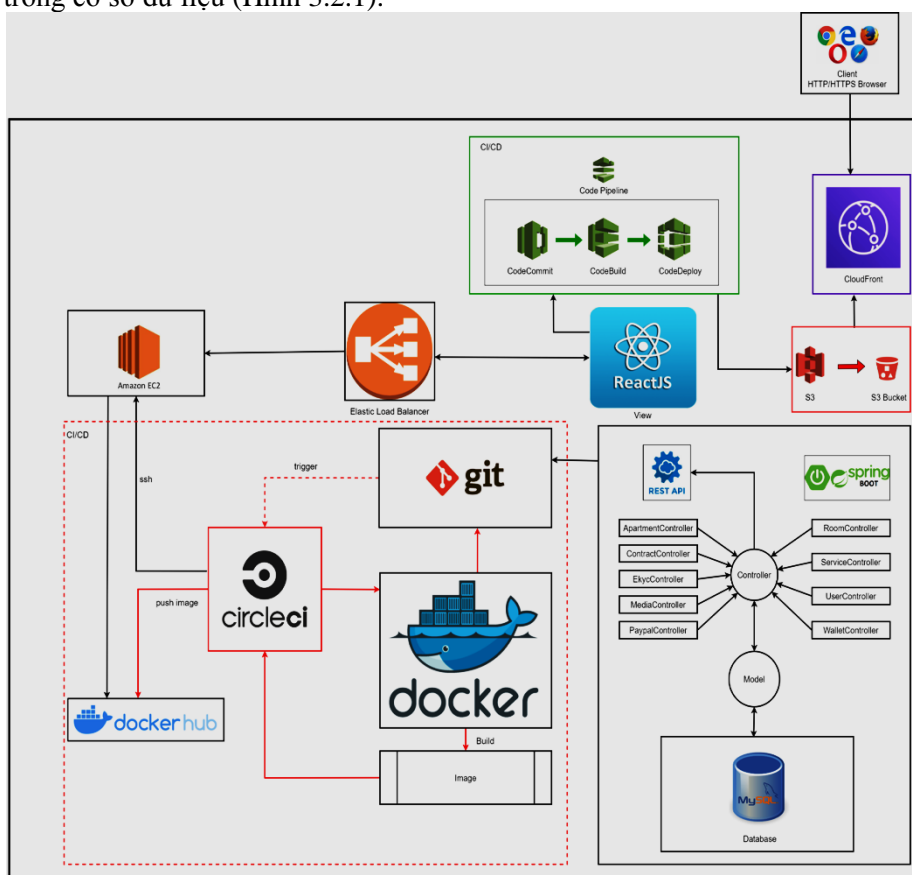
3.1. Giới thiệu

Để triển khai các giải pháp mà chúng tôi nêu ra trong phần giới thiệu, chúng tôi đề xuất xây dựng ứng dụng Houflex. Ứng dụng được triển khai dựa vào các công nghệ như Spring Boot, ReactJS/React-Native. Đồng thời, để ứng dụng Houflex người dùng (người thuê và cho thuê), chúng tôi cũng dùng các dịch vụ như AWS và CI/CD. Hệ thống Houflex gồm ứng dụng trên điện thoại (app) và hệ thống website. Trong đó, React Native tạo ứng dụng trên điện thoại hệ điều hành Android/iOS, phục vụ cho người dùng. ReactJS tạo website với các chức năng tương tự đồng bộ với app. Spring Boot chịu trách nhiệm tạo ra các Restful-API, xử lý các request. MySQL dùng lưu trữ dữ liệu của toàn hệ thống như các thông tin phòng, thông tin người dùng, hợp đồng... Các dịch vụ điện toán đám mây sử dụng trong AWS gồm: Elastic Load Balancing (ELB); EC2; S3; CodePipeline; CodeCommit. Cụ thể các dịch vụ này đảm nhiệm các công việc như sau: EC2 chịu

trách nhiệm tạo máy ảo lưu trữ mã nguồn cho server. Load Balancing cân bằng tải các request. CodePipeline quản lý các quá trình CI/CD của website. CodeCommit để lưu trữ mã nguồn của web. S3 dùng để lưu trữ các dữ liệu dưới dạng file/folder.

3.2. Kiến trúc phần mềm

Ứng dụng Houflex áp dụng kiến trúc MVC. MVC là một mẫu kiến trúc phần mềm để tạo lập giao diện người dùng trên máy tính và được áp dụng rộng rãi trên các kiến trúc của một hệ thống website. Kiến trúc MVC sẽ được chia làm 3 phần đúng theo tên viết tắt của nó là: View – Model – Controller. **View**: sử dụng công nghệ ReactJS cho phía web, React-Native cho phía app để thực thi giao diện người dùng. Sử dụng các phương thức HTTP/HTTPS để kết nối với Controller của server. **Controller**: Lấy dữ liệu từ Model để xây dựng các RESTful API để View có thể truy cập được và lấy dữ liệu hiển thị ra giao diện. **Model**: Chịu trách nhiệm lưu trữ các dữ liệu từ phía View vào cơ sở dữ liệu. Hơn nữa, Model tạo các entity để tham chiếu tới các đối tượng trong cơ sở dữ liệu (Hình 3.2.1).



Hình 3.2.1 Kiến trúc của Houflex

3.3. Phân tích hệ thống Houflex

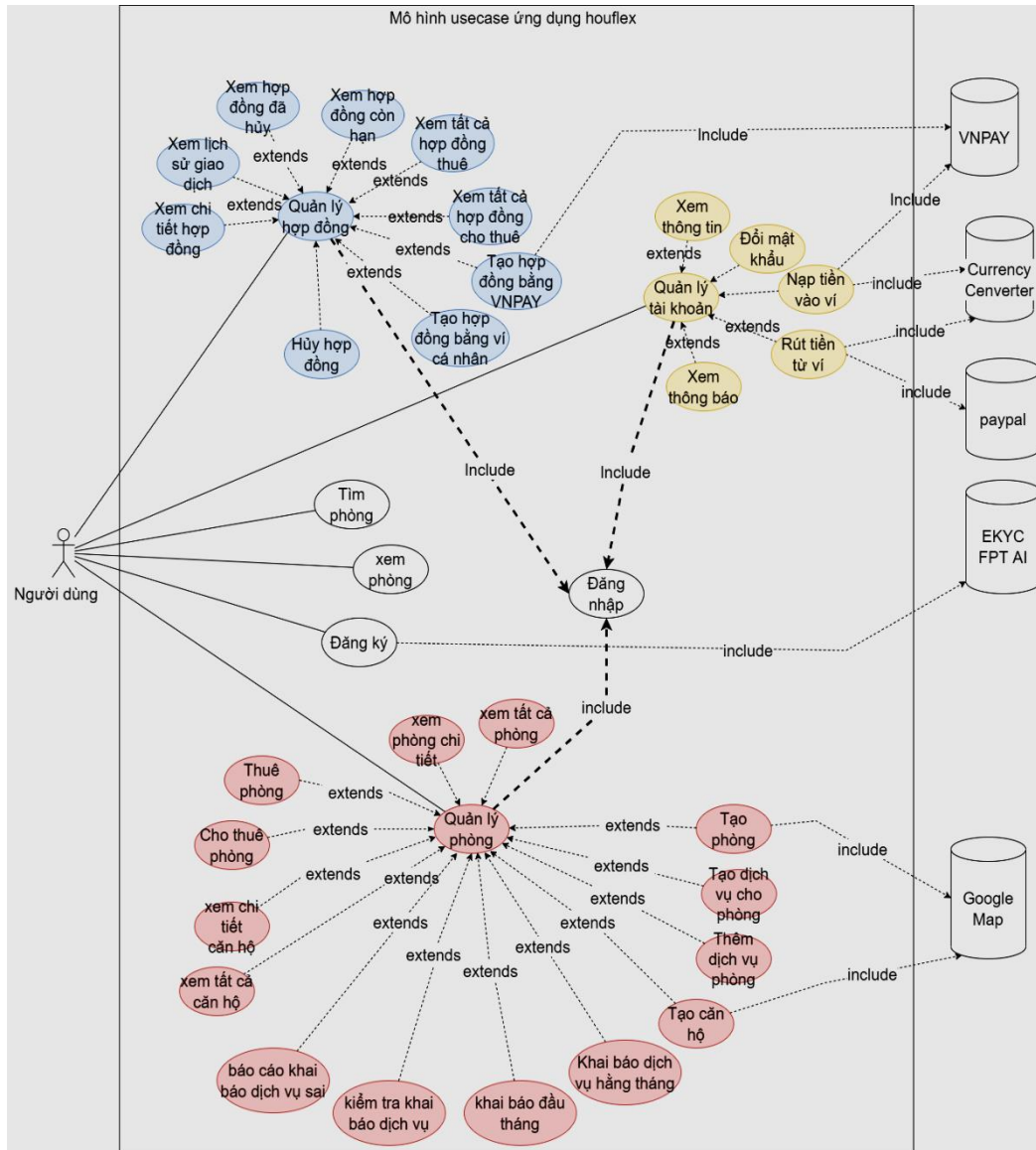
Đối tượng người dùng cho hệ thống Houflex gồm có người thuê và chủ nhà. Trong đó, chúng tôi tập trung vào các chức năng chính của hệ thống là hỗ trợ người muốn cho thuê cũng như người đi thuê phòng và tự động hóa tuần tự của các giao dịch trong hợp đồng đã được lập bởi người dùng. Ngoài ra, các chức năng phụ khác cũng được đưa vào để người dùng có thể dễ dàng sử dụng hơn như hỗ trợ các ví điện tử, thực hiện các giao dịch thanh toán thông qua ví điện tử VNPAY. Những chức năng này chúng tôi trình bày trong Bảng 3.3.1.

Bảng 3.3.1 Danh sách các chức năng trong Houflex

Đối tượng	Tên chức năng	Mô tả
Người dùng chung	Quản lý hợp đồng	Quản lý các hợp đồng của chính mình gồm các chức năng như: xem lịch sử giao dịch hợp đồng, tạo hợp đồng mới, hủy hợp đồng...
	Quản lý tài khoản	Gồm các chức năng: xem thông tin cá nhân, đổi mật khẩu, nạp tiền, rút tiền
	Tìm kiếm phòng	Tìm kiếm danh sách phòng theo tiêu chuẩn của chính mình.
	Xem phòng	Xem thông tin chi tiết của căn phòng đã tìm kiếm được.
	Đăng ký	Đăng ký tài khoản ứng dụng Houflex.
	Đăng nhập	Đăng nhập tài khoản vào ứng dụng.
	Quản lý thông báo	Cho phép nhận các thông báo từ hệ thống, hợp đồng liên quan, ...
Chủ nhà	Tạo căn hộ, phòng	Chủ nhà tạo các căn hộ, phòng lên hệ thống để các người dùng khác có thể thuê.
	Xem danh sách căn hộ, phòng	Xem lại danh sách chi tiết các căn hộ, phòng của chính mình đã đăng lên.
	Quản lý dịch vụ của phòng	Cho phép thêm, xóa hoặc sửa các dịch vụ của phòng khi phòng chưa được người khác thuê.
	Khai báo chỉ số dịch vụ	Chủ nhà phải khai báo chỉ số dịch vụ lần đầu khi hợp đồng bắt đầu có hiệu lực và có thể nhập chỉ số tháng đầu nếu như người dùng không khai báo.
	Kiểm tra khai báo dịch vụ	Xem lại các khai báo hàng tháng của người thuê.
	Yêu cầu khai báo lại chỉ số dịch vụ hàng tháng.	Cho phép yêu cầu người thuê khai báo lại chỉ số dịch vụ hàng tháng nếu như người thuê nhập không đúng với thực tế.
Người đi thuê	Thuê phòng	Cho phép tạo hợp đồng với các phòng hợp lệ.
	Khai báo chỉ số hàng tháng	Yêu cầu khai báo chỉ số hàng tháng, nếu có chậm trễ thì sẽ bị phạt theo quy định của hợp đồng.
Hệ thống	Yêu cầu nhập chỉ số dịch vụ cho phòng.	Gửi thông báo yêu cầu nhập chỉ số dịch vụ cho phòng tới người thuê hoặc chủ nhà.
	Tự động thanh toán	Tự động thanh toán tiền phòng hàng tháng.
	Thông báo hệ thống	Gửi các thông báo hệ thống tới người dùng.
	Thông báo cảnh báo	Gửi các thông báo cảnh báo tới người dùng như: không đủ số dư, quá hạn thanh toán, ...

Từ những chức năng trên, chúng tôi tập trung hiện thực các chức năng chính trong hệ thống App cũng như Web HouFlex như: Khởi tạo phòng, thuê phòng và các chức năng tự động của hệ thống. Sau đó chúng tôi mở rộng các chức năng tương ứng như trong Hình 3.3.1. Trong mô hình use-case chúng tôi trình bày, người dùng chính là actor chính mà chúng tôi hướng đến (chủ nhà, người đi thuê). Trong đó, những use-case chính và phức tạp về nghiệp vụ đòi hỏi nhiều xử lý cho hệ thống đó là: use-case số UC09-Lập hợp đồng cho thuê và use-case UC10-Thanh toán hợp đồng. Ngoài ra, chúng tôi cũng mở rộng các use-case cho các chức năng của hệ thống nhằm tạo điều kiện thuận lợi cho người dùng. Tuy nhiên, những chức năng của các dịch vụ khác chúng tôi chưa đề cập ở đây. Chúng tôi xây dựng các chức năng trên tích hợp cùng các bên thứ 3 như:

Google Map API, Vnpay, Paypal, Currency Converter và EKYC FPT AI. Đồng thời, dựa trên những chức năng đã nêu, chúng tôi tiến hành thiết kế mô hình quan hệ cho ứng dụng.



Hình 3.3.1 Mô hình Use-case của Houflex

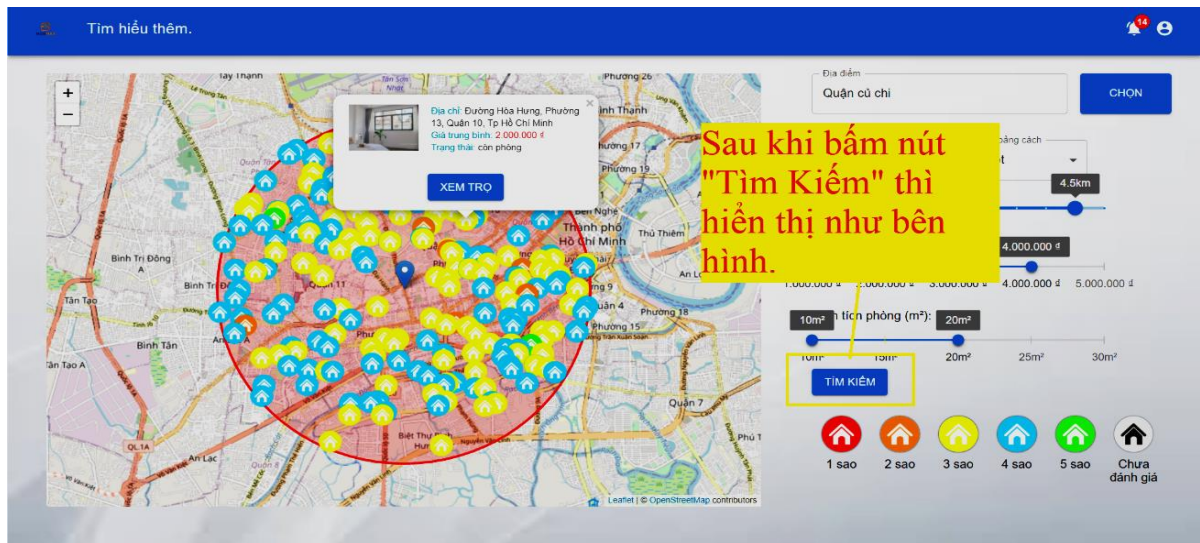
3.4. Giao diện đồ họa

3.4.1. Hệ thống Website của Houflex

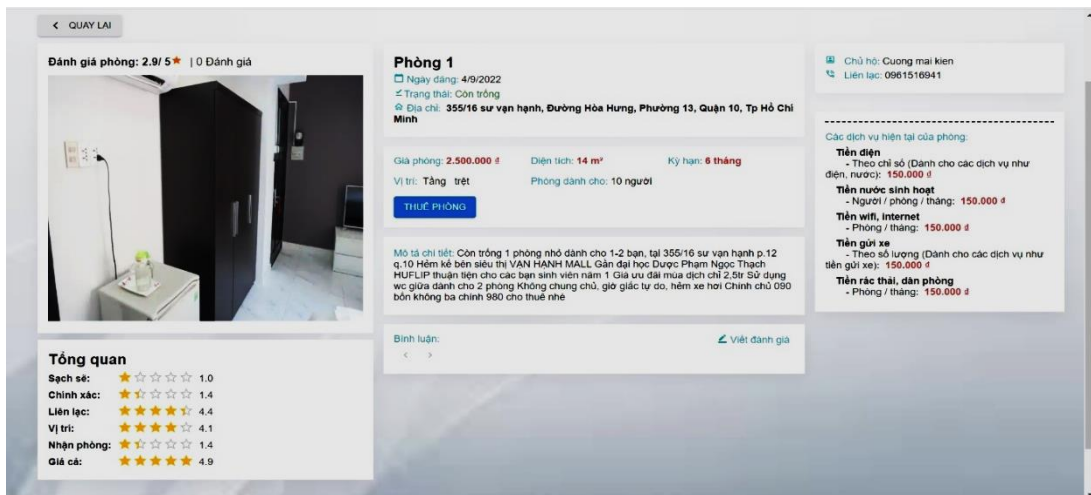
Để thuận tiện cho việc tìm phòng, chúng tôi đã trực quan hóa bản đồ và hiển thị các thông tin phòng trọ lên giao diện để người dùng dễ dàng sử dụng (Hình 3.4.1.1).

Các phòng trọ sẽ được hiển thị theo sắp xếp từ đánh giá người dùng (sao cho các chọn lựa). Các sao này sẽ được gửi về từ phản hồi của người đã thuê phòng trọ trước đó. Sẽ có các mức độ từ 1-5 sao tương ứng cho các mức độ hài lòng của người dùng. Hệ thống sẽ tính toán trung bình sao của phòng trọ và hiển thị lên với các màu tương ứng. Khi bấm vào tìm kiếm, người dùng sẽ có thể xem chi tiết của một phòng trọ với các thông tin liên quan được hệ thống hiển thị một cách đầy đủ (Hình 3.4.1.2).

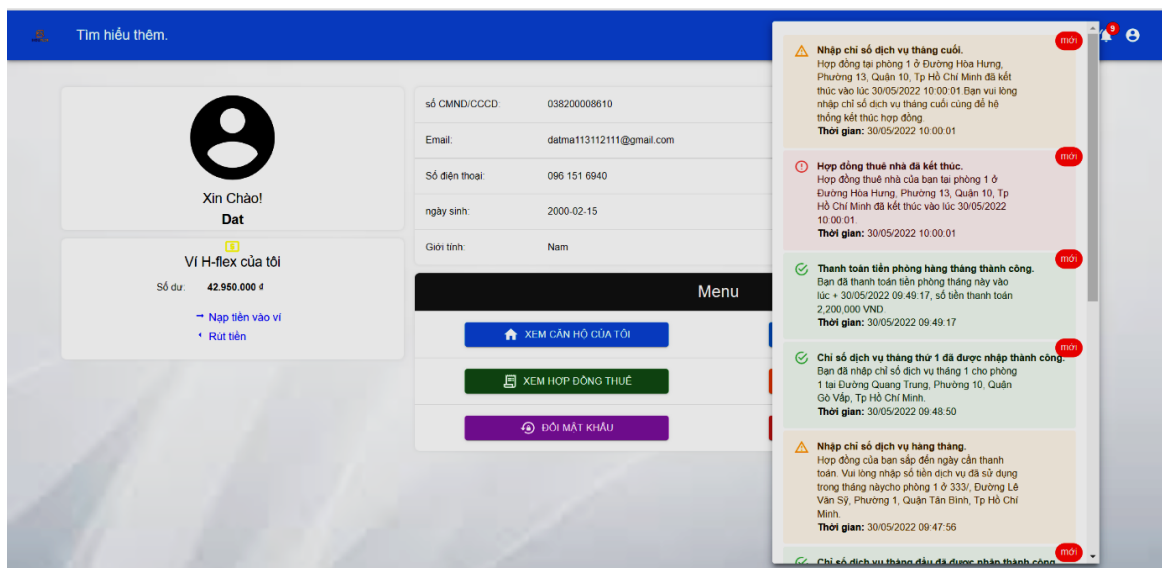
Khi thuê phòng, người dùng sẽ phải ký một hợp đồng với người thuê và thực hiện các thanh toán liên quan. Khi đã xong các thanh toán, hợp đồng sẽ được lập thành công. Người dùng cũng nhận được các thông báo tương ứng (Hình 3.4.1.3).



Hình 3.4.1.1 Giao diện trang tìm phòng Houflex



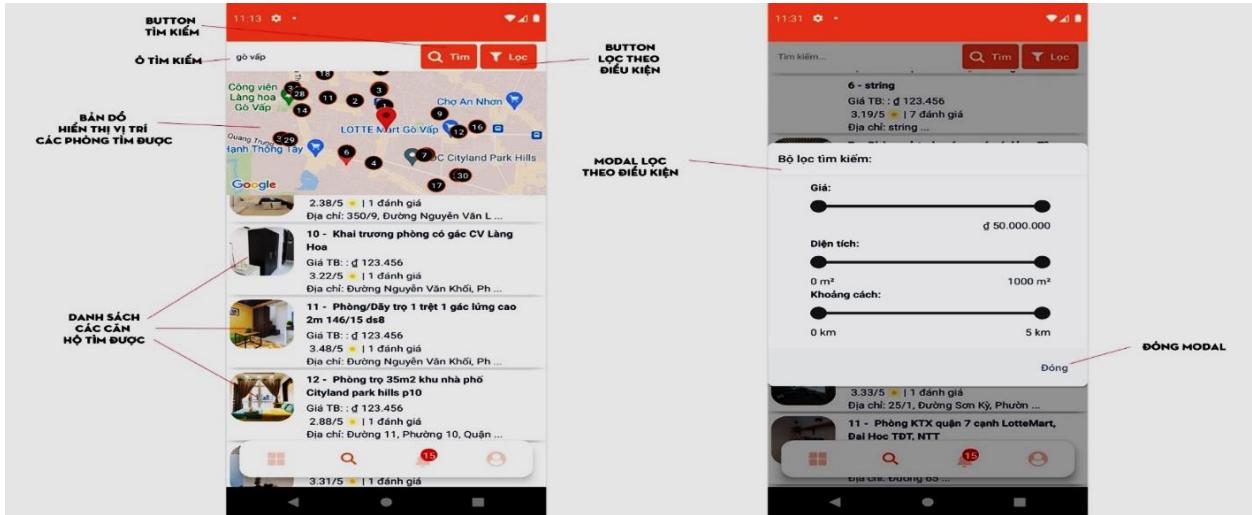
Hình 3.4.1.2 Giao diện thông tin phòng được tìm thấy trên hệ thống Houflex



Hình 3.4.1.3 Trang thông tin và thông báo các giao dịch thành công trên hệ thống Houflex

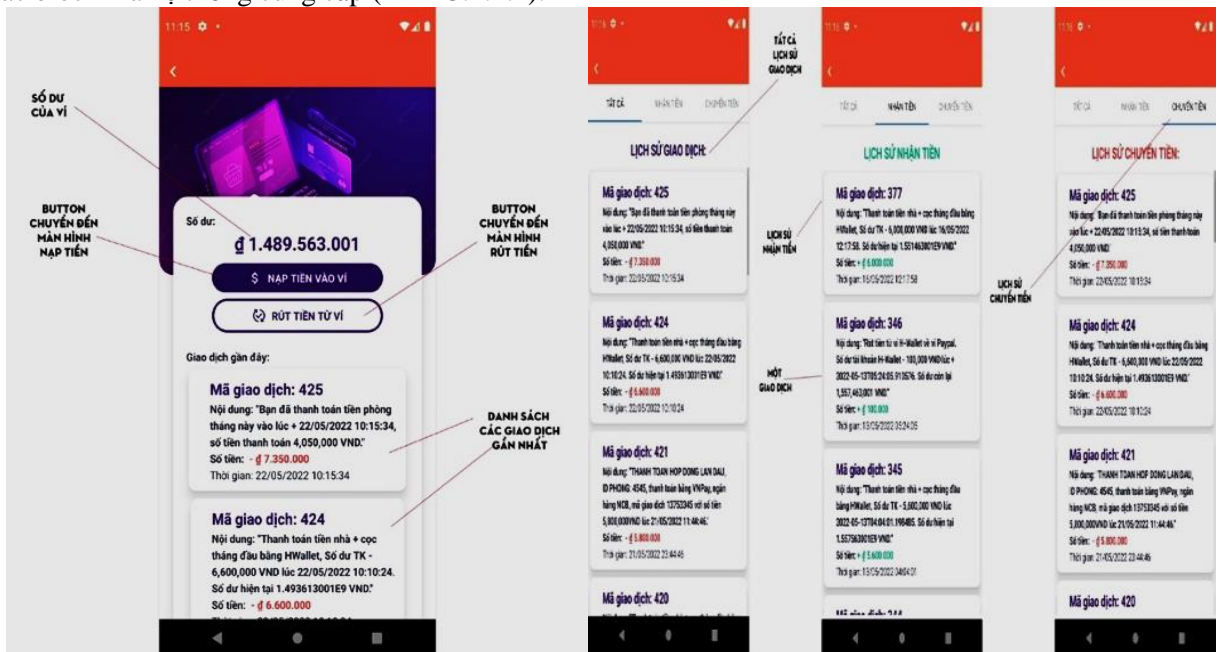
3.4.2. Phía App trên thiết bị di động của hệ thống Houflex

Chúng tôi giới thiệu sơ lược về giao diện của ứng dụng Houflex chạy trên thiết bị di động Android. Sau khi người dùng đã vào trang tìm kiếm phòng và căn hộ theo nhu cầu, giao diện sẽ như sau:



Hình 3.4.2.1 Màn hình tìm phòng chạy trên ứng dụng di động Houflex

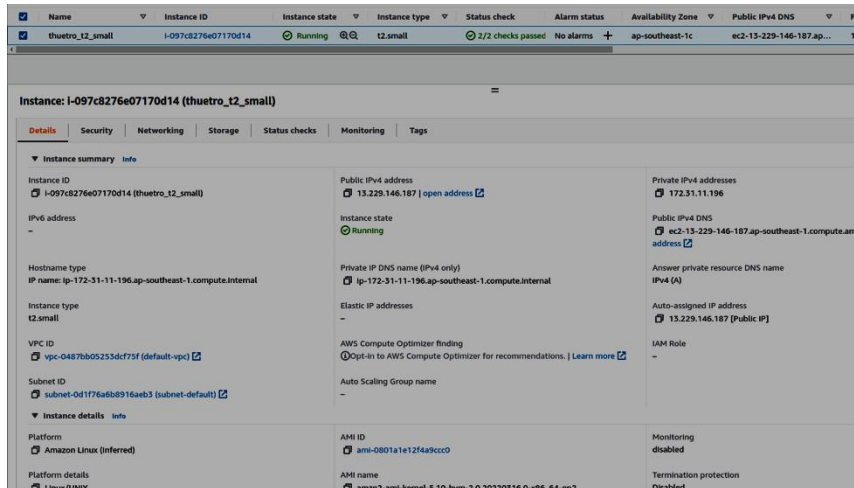
Ở màn hình tìm kiếm, người dùng sẽ nhập địa chỉ muốn thuê trọ vào ô tìm kiếm, bấm nút **Lọc** để thêm các tiêu chí tìm kiếm như giá phòng, diện tích phòng, khoảng cách đến căn hộ. Sau đó người dùng chọn chức năng tìm kiếm, kết quả tìm kiếm phù hợp sẽ được hiển thị trực tiếp lên trên bản đồ hoặc hiển thị dưới dạng danh sách. Màn hình lịch sử sẽ hiển thị các lịch sử liên quan đến giao dịch tiền như là thuê phòng thành công, thanh toán tiền phòng thành công, yêu cầu khai báo chỉ số cho các dịch vụ tại tháng đầu tiên. Khi bấm vào một giao dịch sẽ hiển thị chi tiết về giao dịch đó như tổng tiền của các dịch vụ đã dùng trong tháng đó. Tất cả các giao dịch và khớp lệnh cho các giao dịch này sẽ được thực hiện một cách tự động dựa vào các block mà hệ thống cung cấp (Hình 3.4.2.2).



Hình 3.4.2.2 Màn hình ví giao dịch và lịch sử giao dịch với sự giám sát của các Block

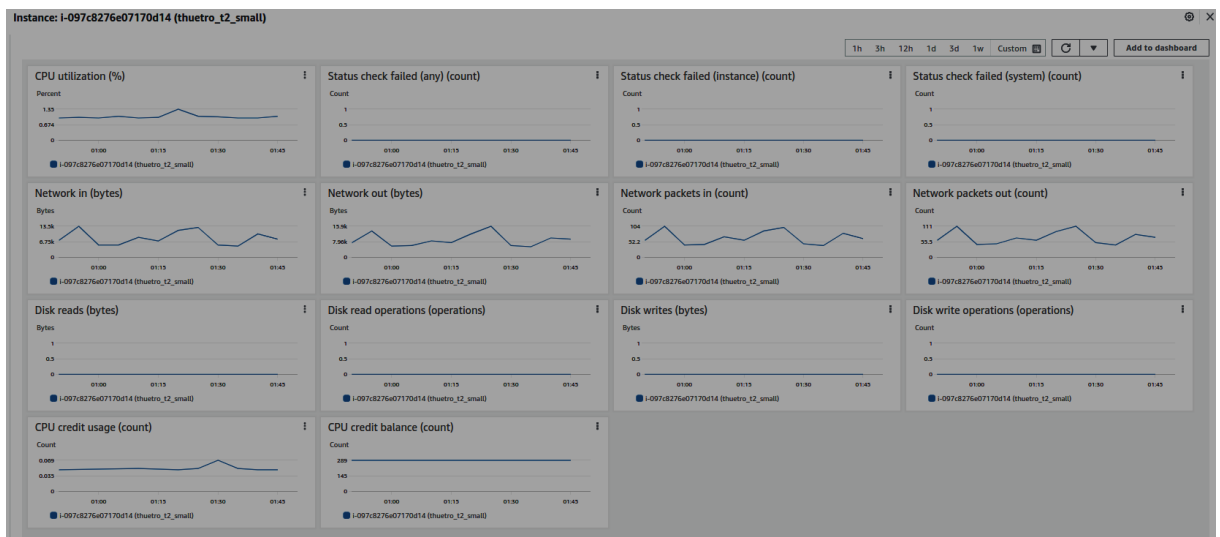
4. KIỂM NGHIỆM THỰC TẾ

Chúng tôi đã tiến hành triển khai hệ thống Houflex lên môi trường cloud của AWS. Trong giai đoạn này chúng tôi sử dụng cơ sở hạ tầng của AWS gồm một EC2 instance (hệ điều hành Linux, phiên bản t2.small), được cài sẵn môi trường docker để chạy ứng dụng của chúng tôi (Hình 4.1).



Hình 4.1 Cấu hình chi tiết của EC2 instance

Hệ thống chạy trên EC2 instance hoạt động ổn định. Không có request nào trả về trạng thái failed. Những thông số về tài nguyên sử dụng CPU (CPU credit balance) cũng như dung lượng đường truyền (Network packets) được đảm bảo trong các khoảng thời gian khi hệ thống tăng tốc xử lý (Bảng 4.1 mô tả ý nghĩa của các thông số). Tuy nhiên đây là phiên bản thử nghiệm nên sẽ dùng các instance có hiệu suất thấp hơn phiên bản triển khai trong thực tế (Hình 4.2).



Hình 4.2 Các thông số Monitoring của hệ thống Houflex được triển khai

Bảng 4.1 Ý nghĩa các chỉ số trong monitoring

Network in (bytes)	Số byte mà ứng dụng sẽ nhận được trên tất cả các giao diện mạng. Số liệu này xác định khối lượng lưu lượng mạng đến cho một ứng dụng duy nhất. Đơn vị: Byte
Network out (bytes)	Số byte được gửi bởi ứng dụng trên tất cả các giao diện mạng. Số liệu này xác định khối lượng lưu lượng mạng đi từ một ứng dụng duy nhất. Đơn vị: Byte
Network packets in (count)	Số lượng gói tin mà ứng dụng nhận được trên tất cả các giao diện mạng. Số liệu này xác định khối lượng lưu lượng đến về số lượng gói trên một ứng dụng. Đơn vị: count
Disk reads (bytes)	Các byte được đọc từ tất cả các khối lượng lưu trữ ứng dụng có sẵn cho ứng dụng. Số liệu này được sử dụng để xác định khối lượng dữ liệu mà ứng dụng đọc từ đĩa cứng. Điều này có thể được sử dụng để xác định tốc độ của ứng dụng. Nếu không có khối lượng lưu trữ nào, thì giá trị là 0 hoặc số liệu không được báo cáo. Đơn vị: Byte
Disk, read operations (operations)	Các thao tác đọc đã hoàn thành từ tất cả các khối lượng lưu trữ ứng dụng có sẵn cho ứng dụng trong một khoảng thời gian cụ thể. Nếu không có khối lượng lưu trữ nào, thì giá trị là 0 hoặc số liệu không được báo cáo. Đơn vị: count
Disk writes (bytes)	Các byte được ghi vào tất cả các khối lượng lưu trữ có sẵn cho ứng dụng đó. Số liệu này được sử dụng để xác định khối lượng dữ liệu mà ứng dụng ghi vào đĩa cứng. Điều này có thể được sử dụng để xác định tốc độ của ứng dụng. Nếu không có khối lượng lưu trữ nào, thì giá trị là 0 hoặc số liệu không được báo cáo. Đơn vị: Byte
Disk, write operations (operations)	Các thao tác ghi đã hoàn thành vào tất cả khối lượng lưu trữ có sẵn cho ứng dụng trong một khoảng thời gian cụ thể. Nếu không có khối lượng lưu trữ nào, thì giá trị là 0 hoặc số liệu không được báo cáo. Đơn vị: count
CPU credit usage (count)	CPU Credit quy định số lượng CPU bùng nổ của một phiên bản. Có thể sử dụng CPU Credit này để tăng hiệu suất của CPU trong suốt thời gian Burst. Giả sử đang vận hành ứng dụng ở 100% hiệu suất CPU trong 5 phút, sẽ tiêu tốn 5 (tức là $5 * 1,0$) CPU Credit. Tương tự, nếu chạy một ứng dụng ở 50% hiệu suất CPU trong 5 phút, sẽ tiêu tốn 2,5 (tức là $5 * 0,5$) Tín dụng CPU.
CPU credit balance (count)	CPU Credit Balance là số lượng CPU Credit còn lại trong tài khoản.

```

[INFO] [dev]:[2022-05-30 11:00:36.907]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/user/info/1 - ...
[INFO] [dev]:[2022-05-30 11:00:36.912]:[k.a.s.UserService]:[1326]:[27.64.18.220]:GET /api/v1/user/info/1 {"userId":1} - get another user
[INFO] [dev]:[2022-05-30 11:00:36.954]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/apartment/rooms/4366 - ...
[INFO] [dev]:[2022-05-30 11:00:36.957]:[k.a.s.ApartmentService]:[1158]:[27.64.18.220]:GET /api/v1/apartment/rooms/4366 {"apartmentId":4366,"
apartment id: 4366
[INFO] [dev]:[2022-05-30 11:00:36.962]:[k.a.s.ApartmentService]:[1160]:[27.64.18.220]:GET /api/v1/apartment/rooms/4366 {"apartmentId":4366,"
[INFO] [dev]:[2022-05-30 11:01:30.936]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/rating/room/4358 - ...
[INFO] [dev]:[2022-05-30 11:01:30.940]:[k.a.s.RatingService]:[156]:[27.64.18.220]:GET /api/v1/rating/room/4358 {"roomId":4358,"pageable"
[INFO] [dev]:[2022-05-30 11:01:30.943]:[k.a.s.RatingService]:[61]:[27.64.18.220]:GET /api/v1/rating/room/4358 {"roomId":4358,"pageable"
[INFO] [dev]:[2022-05-30 11:01:31.010]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/rating/room/4358/overview - ...
[INFO] [dev]:[2022-05-30 11:01:46.533]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/contract/detail/4358 - ...
[INFO] [dev]:[2022-05-30 11:02:03.876]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/contract/detail/4358 - ...
room id : 4358 , renterId : 3
[INFO] [dev]:[2022-05-30 11:02:05.292]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:POST /api/v1/user/email/gen-otp - ...
[INFO] [dev]:[2022-05-30 11:02:05.300]:[k.a.s.UserService]:[95]:[27.64.18.220]:POST /api/v1/user/email/gen-otp {"email":"luutkha@gmail
[INFO] [dev]:[2022-05-30 11:02:05.304]:[k.a.s.UserService]:[1106]:[27.64.18.220]:POST /api/v1/user/email/gen-otp {"email":"luutkha@gmail
[INFO] [dev]:[2022-05-30 11:02:05.307]:[k.a.s.UserService]:[1161]:[27.64.18.220]:POST /api/v1/user/email/gen-otp {"email":"luutkha@gmail
[INFO] [dev]:[2022-05-30 11:02:05.917]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/contract/detail/4358 - ...
room id : 4358 , renterId : 3
[INFO] [dev]:[2022-05-30 11:02:18.196]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/user/info - ...
[INFO] [dev]:[2022-05-30 11:02:18.198]:[k.a.s.UserService]:[263]:[27.64.18.220]:GET /api/v1/user/info {} - get user info id: 3
[INFO] [dev]:[2022-05-30 11:02:19.082]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/notification/list - ...
[INFO] [dev]:[2022-05-30 11:02:19.084]:[k.a.s.NotificationService]:[126]:[27.64.18.220]:GET /api/v1/notification/list - get all notification
[INFO] [dev]:[2022-05-30 11:02:19.089]:[k.a.s.NotificationService]:[128]:[27.64.18.220]:GET /api/v1/notification/list - total elements: 37
[INFO] [dev]:[2022-05-30 11:02:27.755]:[k.a.i.LoggingInterceptor]:[123]:[27.64.18.220]:GET /api/v1/apartment/search - ...
    
```

Hình 4.3 Màn hình logs của hệ thống Houflex khi triển khai trên AWS

Chúng tôi ghi log các thông tin như thời gian, api được gọi, ip, nội dung của request để phục vụ cho việc giám sát hệ thống cũng như việc truy vết khi có nhu cầu. (Hình 4.3).

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Trong bài báo này, chúng tôi đã trình bày hai giải pháp cho việc xây dựng ứng dụng Houflex đó là ứng dụng di động và ứng dụng Web nhằm giải quyết bài toán xây dựng hệ thống tìm kiếm và quản lý hệ thống thuê phòng trên nền tảng Blockchain. Cả 2 giải pháp chúng tôi đều sử dụng các công nghệ phù hợp để phát triển ứng dụng một cách hiệu quả và đảm bảo các tiêu chí cho một hệ thống thông minh. Chúng tôi sử dụng React-Native cho việc phát triển ứng dụng di động và ReactJS cho ứng dụng Web. Houflex sử dụng Spring-Boot để tạo ra các Restful-API để tương tác với phía Client. Sử dụng được một số dịch vụ dựa trên nền tảng đám mây của AWS để phục vụ cho việc triển khai hệ thống như: CodePipeline – CodeBuild – CodeDeploy: để tự động thực hiện CI/CD cho website. CodeCommit để lưu trữ mã nguồn của Website. Phía server sử dụng EC2 – Load Balancing để cân bằng tải cho các API. Áp dụng các công cụ CI/CD cho phía server. Docker tự động biến mã nguồn thành các Image và đóng gói thành các Container nhằm phục vụ cho việc tự động triển khai hệ thống. Circle CI quản lý các container của docker. Áp dụng được Blockchain để giải quyết bài toán hợp đồng thông minh. Hệ thống sẽ dựa trên các điều khoản đã được hệ thống lập trình ra sẵn để thiết lập hợp đồng thuê giữa người thuê và chủ nhà. Sau khi hợp đồng được thiết lập thành công, hợp đồng sẽ tự động triển khai đúng như những gì điều khoản đã đề ra. Mỗi hành động mà hợp đồng triển khai sẽ được lưu trữ vào các Block. Hệ thống sẽ kiểm tra tính hợp lệ của các block nhằm xác minh quá trình thực hiện chính xác những gì đã cam kết. Ứng dụng còn có ví điện tử sử dụng riêng cho hệ thống phục vụ cho việc thanh toán hợp đồng. Số dư trong ví cũng có thể chuyển khoản về tài khoản của người dùng qua ví Paypal. Hệ thống cũng ứng dụng Blockchain để truy vết được các giao dịch đã được thực hiện. Khi một giao dịch được thực hiện giữa người dùng, hệ thống sẽ lưu trữ đồng thời dữ liệu vào cả cơ sở dữ liệu lẫn chuỗi khối. Điều này nhằm bảo vệ tính xác thực và chính xác của các giao dịch và bảo vệ được quyền lợi của người dùng. Trong tương lai, ứng dụng Houflex sẽ được triển khai không chỉ trong phạm vi Thành phố Hồ Chí Minh mà còn có thể phát triển ra trên cả nước. Hơn nữa, các chức năng về quản lý các hợp đồng thông minh sẽ được hoàn thiện hơn, minh bạch, rõ ràng hơn nữa cùng kết hợp với các công cụ tìm kiếm phòng tiện nghi và dễ dàng. Về quy mô của ứng dụng, Houflex cũng có thể mở rộng ra không chỉ phục vụ cho việc thuê phòng trọ mà còn có thể cho thuê các loại nhà, khách sạn, thậm chí có thể cho thuê các loại dịch vụ phục vụ cho việc phát triển du lịch.

TÀI LIỆU THAM KHẢO

- [1]. Sim, W., Choi, W.: Spring boot-based web application development for providing information on security vulnerabilities and patches for open source software. *Journal of Korea Society of Digital Industry and Information Management* 17(4) (2021) 77–83
- [2]. Rawat, P., Mahajan, A.N.: Reactjs: A modern web development framework. *International Journal of Innovative Science and Research Technology* 5(11) (2020)
- [3]. ReactJS Library Version 16.8 (Release February 16, 2019), [online], <https://reactjs.org/>, Accessed 20/05/2022
- [4]. Amazon Web Service – EC2 (2022), [online], Secure and resizable Cloud Compute, <https://aws.amazon.com/ec2>, Accessed 20/05/2022
- [5]. Amazon Web Service – S3 (2022), [online], Cloud Object Storage, <https://aws.amazon.com/s3/>, Accessed 20/05/2022
- [6]. Amazon Web Service – CloudFront (2022), [online], Content Delivery Network, <https://aws.amazon.com/cloudfront/>, Accessed 20/05/2022

- [7]. Amazon Web Service – CodeCommit (2022), [online] Managed Source Control Service, <https://aws.amazon.com/codecommit/>, Accessed 20/05/2022
- [8]. Amazon Web Service – CodePipeline (2022), [online] Continuous Integration and Continuous Delivery, <https://aws.amazon.com/codepipeline/>, truy cập ngày 20/05/2022
- [9]. Amazon Web Service – ELB (2022), [online], Elastic Load Balancing <https://aws.amazon.com/elasticloadbalancing>, Accessed 20/05/2022
- [10]. Thạc sĩ Nguyễn Nhi Quang (2021), “Ứng dụng Blockchain trong giao dịch L/C tại các ngân hàng thương mại Việt Nam”, Tạp chí Công nghệ và Ngân hàng số, ngày 11/08/2021
- [11]. Alharby, M., Van Moorsel, A.: Blockchain-based smart contracts: A systematic mapping study. arXiv preprint arXiv:1710.06372 (2017)

YSC4F.228

HOẠCH ĐỊNH QUỸ ĐẠO VÀ ĐIỀU KHIỂN BÙ SAI LỆCH ĐỘNG HỌC CHO MOBILE ROBOT

TRẦN HỮU TOÀN, VƯƠNG HOÀNG SANG, CAO VĂN THẮNG

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

tranhuutoan@iuh.edu.vn, svuonghoangsang@gmail.com, caothang15102000@gmail.com

Tóm tắt. Bài báo này trình bày thuật toán điều khiển bù sai lệch động học cho Robot di động ứng dụng trong môi trường tác vụ thực tế bám theo quỹ đạo được hoạch định trước. Cụ thể, việc phát triển thuật toán bù động học cho Robot di động nhằm mang đến khả năng xử lý chính xác hơn và bám vị trí. Để giải quyết vấn đề, bài báo trình bày một phương thức điều khiển dựa trên phương pháp đo sai lệch giữa 2 bánh xe của robot và tính toán để bù cho bánh xe bị lệch kèm theo điều kiện ràng buộc về vận tốc ổn định để tránh quán tính. Quỹ đạo được hoạch định cho trước. Để kiểm chứng mức độ hiệu quả của thuật toán đề xuất, mô hình Robot di động di chuyển bằng bánh xích được xây dựng để kiểm tra thực nghiệm. Kết quả nghiên cứu ban đầu cho thấy thuật toán bù sai lệch động học và hoạch định quỹ đạo được đề xuất cải thiện đáng kể chất lượng điều khiển bám quỹ đạo. Với quỹ đạo đường thẳng và quỹ đạo chữ L, thuật toán điều khiển bù động học cải thiện trung bình 31,2% đến 72,5% sai lệch so với điều khiển PID thông thường.

Từ khóa. Bộ điều khiển PID, robot di động, điều khiển quỹ đạo robot di động, động học robot di động.

PLANNING TRAJECTORY AND CONTROL OF COMPENSATED KINEMATIC VARIATION FOR MOBILE ROBOT

Abstract. This paper presents the kinematic deviation compensation control algorithm for mobile robot applied in the real task environment following the pre-planned trajectory. Specifically, the development of a kinematic compensation algorithm for mobile robots to provide more precise handling and position tracking. To solve the problem, the paper presents a control method based on the method of measuring the difference between the two wheels of the robot and calculating it to compensate for the wheel deviation with the constraint of the stable velocity in order to solve the problem. avoid inertia. The trajectory is planned based on inverse kinematics. To test the effectiveness of the proposed algorithm, the crawler mobile robot model is built to test. experiment. The initial research results show that the proposed trajectory planning and kinematic offset compensation algorithm significantly improves the quality of trajectory tracking. With straight-line trajectory and L-shaped trajectory, the kinematic compensation control algorithm on average improves 31.2% to 72.5% deviation compared with conventional PID control.

Keywords. PID controller, mobile robot trajectory control, mobile robots.

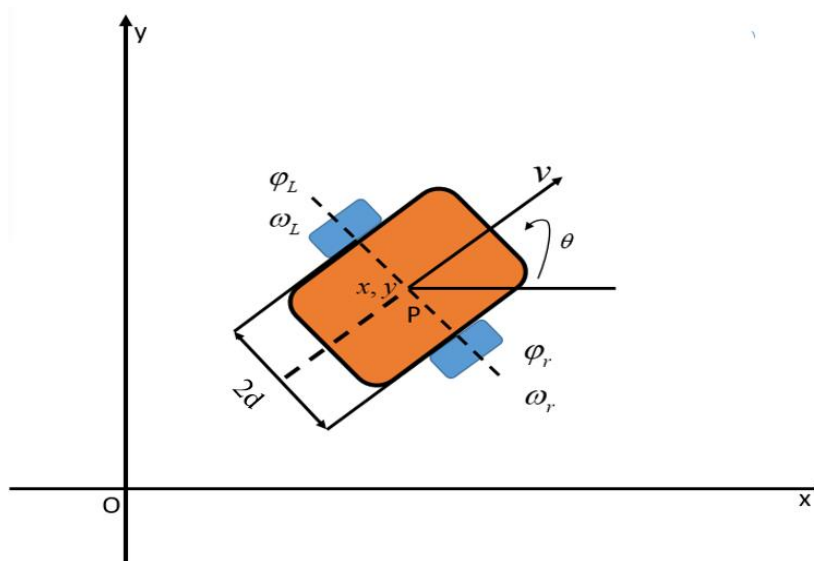
1. GIỚI THIỆU

Mobile Robot đang được ứng dụng rộng rãi trong các ngành nghề: vũ trụ, kho hàng, khai thác... do tính đơn giản và dễ điều khiển, giúp con người làm những việc đòi hỏi sự chính xác do đó Mobile Robot thường được dùng các thuật toán có bộ điều khiển chính xác và các cảm biến để định vị vị trí một cách chính xác. Để điều khiển Mobile Robot bám quỹ đạo tốt thường sử dụng các bộ điều khiển như: bộ điều khiển mờ (Fuzzy Logic) và điều khiển cổ điển (PI, PID), mạng thần kinh nhân tạo (nơron) và đã có một số bài nghiên cứu về điều khiển Mobile Robot như: [10] “Thiết kế bộ điều khiển chuyển động cho robot tự hành di chuyển theo đường hoạch định dựa trên logic mờ” của Phan Tú Châu hay [13] “Ứng dụng mạng nơron điều khiển mobile robot di chuyển tránh vật cản” của Lê Thanh Hoàn, [1] “Control of Wheeled Mobile Robots: An

Experimental Overview” của các tác giả Alessandro De Luca, Giuseppe Oriolo, and Marilena Vendittelli, [8] “Modelling and trajectory tracking of wheeled mobile robots” của N.Leena và K.K.Saju. Các bài báo trên đưa ra phương pháp điều khiển bám quỹ đạo dựa trên các thuật toán hiện đại như: Fuzzy Logic hay mạng thần kinh nhân tạo thì ở đây bài nghiên cứu này sẽ trình bày thêm phương pháp khắc phục hạn chế của bộ điều khiển cổ điển (PI,PID) thông thường, hạn chế ở đây khi điều khiển PID độc lập cho hai động cơ thì chịu các yếu tố bất định làm cho hai động cơ khi hoạt động cùng vận tốc cùng thông số nhưng sẽ không giống nhau hoàn toàn, để khắc phục được sự sai lệch đó bài nghiên cứu đưa ra phương pháp sử dụng thuật toán bù sai lệch động học để cân chỉnh vận tốc của hai động cơ. Bài báo này sẽ trình bày kết quả thực nghiệm và kiểm chứng chất lượng điều khiển giữa bộ điều khiển PID có sử dụng thuật toán bù sai lệch động học và bộ điều khiển PID thông thường. Bên cạnh đó sẽ đánh giá kết quả và thảo luận các hướng phát triển cho kết quả nghiên cứu.

2. CẤU HÌNH VÀ ĐỘNG HỌC ROBOT

2.1. Cấu hình Robot



Hình 41 Động học Mobile Robot

Cấu trúc của Robot di động được mô tả bằng 3 biến:

$$q = [x, y, \theta]^T$$

Trong đó:

ω_r : vận tốc góc của bánh phải

ω_l : vận tốc góc của bánh trái

ϕ_r : góc quay của bánh lái bên phải

ϕ_l : góc quay của bánh lái bên trái

v : vận tốc

P : trọng tâm của chiếc xe

d : khoảng cách từ tâm xe tới bánh xe

θ : góc định hướng của Mobile Robot

R : bán kính bánh xe

x, y : tọa độ của góc P

2.2. Động học thuận Mobile RoBot

Vận tốc v có thể tìm thấy theo phương trình:

$$v = \frac{R(\omega_r + \omega_l)}{2} \quad (1)$$

Ở đây vận tốc góc sẽ được tính bằng đạo hàm góc quay bánh lái bánh xe:

$$\omega_r = \frac{d\varphi_r}{dt} \quad (2)$$

$$\omega_l = \frac{d\varphi_l}{dt} \quad (3)$$

Vị trí và Hướng của Mobile Robot được xác định bởi 1 bộ phương trình vi phân:

$$\dot{x} = \frac{(R \cos \theta)(\omega_r + \omega_l)}{2} \quad (4)$$

$$\dot{y} = \frac{(R \sin \theta)(\omega_r + \omega_l)}{2} \quad (5)$$

$$\dot{\theta} = \frac{R(\omega_r - \omega_l)}{2d} \quad (6)$$

Rút gọn:

$$\dot{x} = v \cos \theta \quad (7)$$

$$\dot{y} = v \sin \theta \quad (8)$$

Khi đó ta sẽ có dạng ma trận:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} v \\ \dot{\theta} \end{bmatrix} \quad (9)$$

Mô hình động học của vận tốc xe v và vận tốc góc w được biểu diễn dưới dạng ma trận:

$$\begin{bmatrix} v \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \frac{R}{2} & \frac{R}{2} \\ \frac{R}{2d} & \frac{-R}{2d} \end{bmatrix} \begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} \quad (10)$$

Từ công thức (9) và (10) suy ra:

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \cos \theta & 0 \\ \sin \theta & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{R}{2} & \frac{R}{2} \\ \frac{R}{2d} & \frac{-R}{2d} \end{bmatrix} \begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{\theta} \end{bmatrix} = \begin{bmatrix} \frac{R}{2} \cos \theta & \frac{R}{2} \cos \theta \\ \frac{R}{2} \sin \theta & \frac{R}{2} \sin \theta \\ \frac{R}{2d} & \frac{-R}{2d} \end{bmatrix} \begin{bmatrix} \omega_r \\ \omega_l \end{bmatrix} \quad (11)$$

Trong trường hợp này, coi chuyển động của Mobile Robot như một hệ thống cơ khí nơi tồn tại 3 ràng buộc:

$$\begin{aligned} \dot{x} \sin \theta - \dot{y} \cos \theta &= 0 \\ \dot{x} \cos \theta + \dot{y} \sin \theta &= R\omega_r - b\dot{\theta} \\ \dot{x} \cos \theta + \dot{y} \sin \theta &= R\omega_r + b\dot{\theta} \end{aligned} \quad (12)$$

Các ràng buộc viết dưới dạng ma trận:

$$A = \begin{bmatrix} \sin \theta & -\cos \theta & 0 & 0 & 0 \\ \cos \theta & \sin \theta & b & R & 0 \\ \cos \theta & \sin \theta & -b & 0 & -R \end{bmatrix} \quad (13)$$

2.3. Động học ngược Mobile Robot dựa trên thực nghiệm

Khi động cơ quay 1 góc α thì bánh xe robot phải quay được 1 góc β với tỷ lệ truyền k ở đây xe có tỷ lệ truyền của bộ giảm tốc 30:1 và tỷ lệ truyền của bánh xích 14:11 ta có:

$$\beta = \alpha \cdot k = \alpha \cdot \left(\frac{1}{30} \cdot \frac{11}{14} \right) \quad (14)$$

Để tính được số vòng cần quay của động cơ để robot có thể đi được một quãng đường là: $L_{xe}(cm)$

Giả sử biến đầu vào: v và θ ta suy ra được công thức tính ω_r và ω_l dựa trên động học của mobile robot ta suy ngược ra sẽ có công thức:

$$\begin{cases} v = \frac{R}{2}(\omega_r + \omega_l) \Rightarrow \frac{2v}{R} = \omega_r + \omega_l \\ \theta = \frac{R(\omega_r - \omega_l)}{2d} \Rightarrow \frac{\theta \cdot (2d)}{R} = \omega_r - \omega_l \end{cases} \Rightarrow \begin{cases} \omega_l = \frac{2v - \theta \cdot (2d)}{2R} \\ \omega_r = \frac{2v + \theta \cdot (2d)}{2R} \end{cases} \quad (15)$$

Ở đây công thức tính vận tốc góc của bánh xe robot qua tỷ lệ truyền $\omega = \omega_0 \cdot k$ từ đây suy ngược lại tính φ_r và φ_l từ công thức $\omega = \frac{d\varphi}{dt}$ ta có:

$$\begin{aligned} \omega_r = \omega_{0r} \cdot k &\Rightarrow \omega_r = \frac{\varphi_{0r}}{t_1} \cdot k \Rightarrow \varphi_{0r} = \frac{\omega_r \cdot t_1}{k} \\ \omega_l = \omega_{0l} \cdot k &\Rightarrow \omega_l = \frac{\varphi_{0l}}{t_1} \cdot k \Rightarrow \varphi_{0l} = \frac{\omega_l \cdot t_2}{k} \end{aligned} \quad (16)$$

Trong đó:

t_1 : thời gian lấy mẫu của động cơ phải

t_2 : thời gian lấy mẫu của động cơ trái

φ_{0r} : góc quay bánh lái bên phải qua tỷ lệ truyền (rad)

φ_{0l} : góc quay bánh lái bên trái qua tỷ lệ truyền (rad)

Dựa trên công thức thực nghiệm tính được quãng đường khi robot đi được dựa trên số xung n :

$$L_{xe} = n \cdot h_{eso_k} \quad (17)$$

Ở đây, $heso_kc$ có được nhờ thực nghiệm đo được từ thực tế và suy luận ra kết hợp với công thức để bánh xe đi được quãng đường $x(cm)$ thì bánh xe quay một góc $\varphi = \frac{180x}{\pi R}$ (độ) Mà $x = Lxe$ suy ra công thức để biết quãng đường hiện tại:

$$Lxe = x = \frac{\varphi\pi R}{180} \quad (18)$$

Sau khi có được quãng đường ta suy ngược lại số xung cần cấp cho động cơ:

$$n = \frac{x}{heso_kc} \quad (19)$$

đổi số xung cần cấp sang độ ta có:

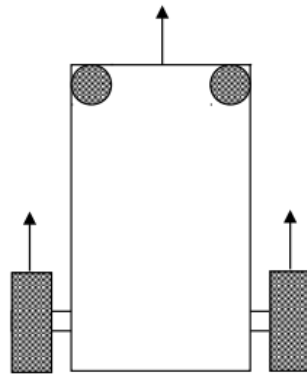
$$toadogoc = n \cdot \frac{360}{PPR} \quad (20)$$

Ở đây PPR là độ phân giải encoder của động cơ từ đây đổi sang số vòng cần phải quay của động cơ:

$$sovongquay = \frac{toadogoc}{360} \quad (21)$$

3. HOẠCH ĐỊNH ĐƯỜNG ĐI CHO MOBILE ROBOT

3.1. Di chuyển thẳng



Hình 42 Mobile Robot đi thẳng

Động cơ được điều khiển bởi vi điều khiển. Vi điều khiển sẽ tạo xung tại các cổng song song và cung cấp trực tiếp cho mạch điều khiển động cơ.

Giả sử ta muốn robot đi được một quãng đường là x (cm) \Rightarrow Bánh xe robot (bánh xe phát động) cần phải quay được một góc là $\varphi = \frac{x}{R}$ (rad) hay $\varphi = \frac{180x}{\pi R}$ (độ) dựa trên động học ngược, thông qua thực nghiệm

tính toán được các công thức liên hệ (18): $L_xe = x = \frac{\varphi\pi R}{180}$ (cm)

Mà $Lxe = x = n \cdot heso_kc$ suy ra:

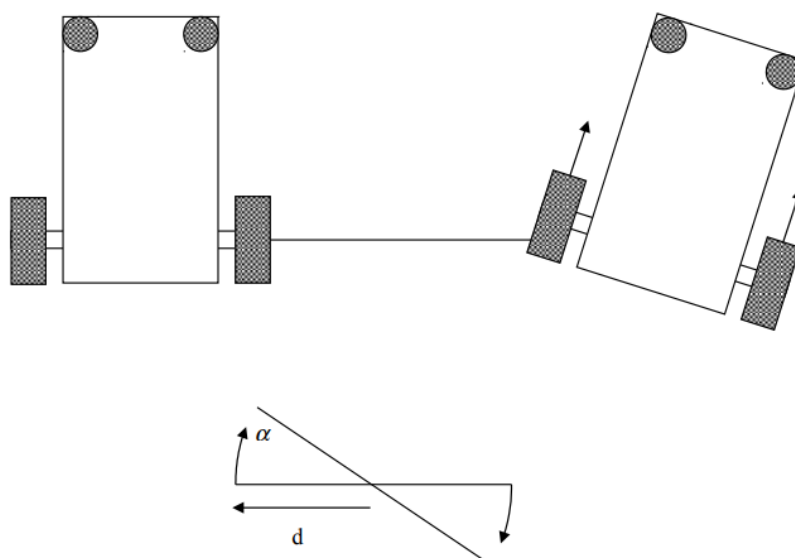
$$n = \frac{x}{heso_kc} \Rightarrow n = \frac{\varphi\pi R}{heso_kc.180} \quad (22)$$

Như vậy, ở bài toán thực tế khi cho đầu vào là v và θ ta tính ngược lại được ω_l và ω_r dựa trên công thức (15):

$$\begin{cases} v = \frac{R}{2}(\omega_r + \omega_l) \Rightarrow \frac{2v}{R} = \omega_r + \omega_l \\ \theta = \frac{R(\omega_r - \omega_l)}{2d} \Rightarrow \frac{\theta \cdot (2d)}{R} = \omega_r - \omega_l \end{cases} \Rightarrow \begin{cases} \omega_l = \frac{2v - \theta \cdot (2d)}{2R} \\ \omega_r = \frac{2v + \theta \cdot (2d)}{2R} \end{cases}$$

Từ đó ta tính được φ và suy ra được số xung cần cấp để xe đi thẳng theo quỹ đường mong muốn.

3.2. Di chuyển xoay



Hình 43 Mobile Robot xoay tại tâm 2 trục bánh xe

Gọi góc quay là α (độ) và giả sử robot quay tại tâm giữa 2 trục bánh xe.

Suy ra: bánh trái cần đi một quãng đường và bánh phải cần đi một quãng đường tương đương với x (cm) là quãng đường:

$$x = \frac{d\alpha\pi}{180} \quad (23)$$

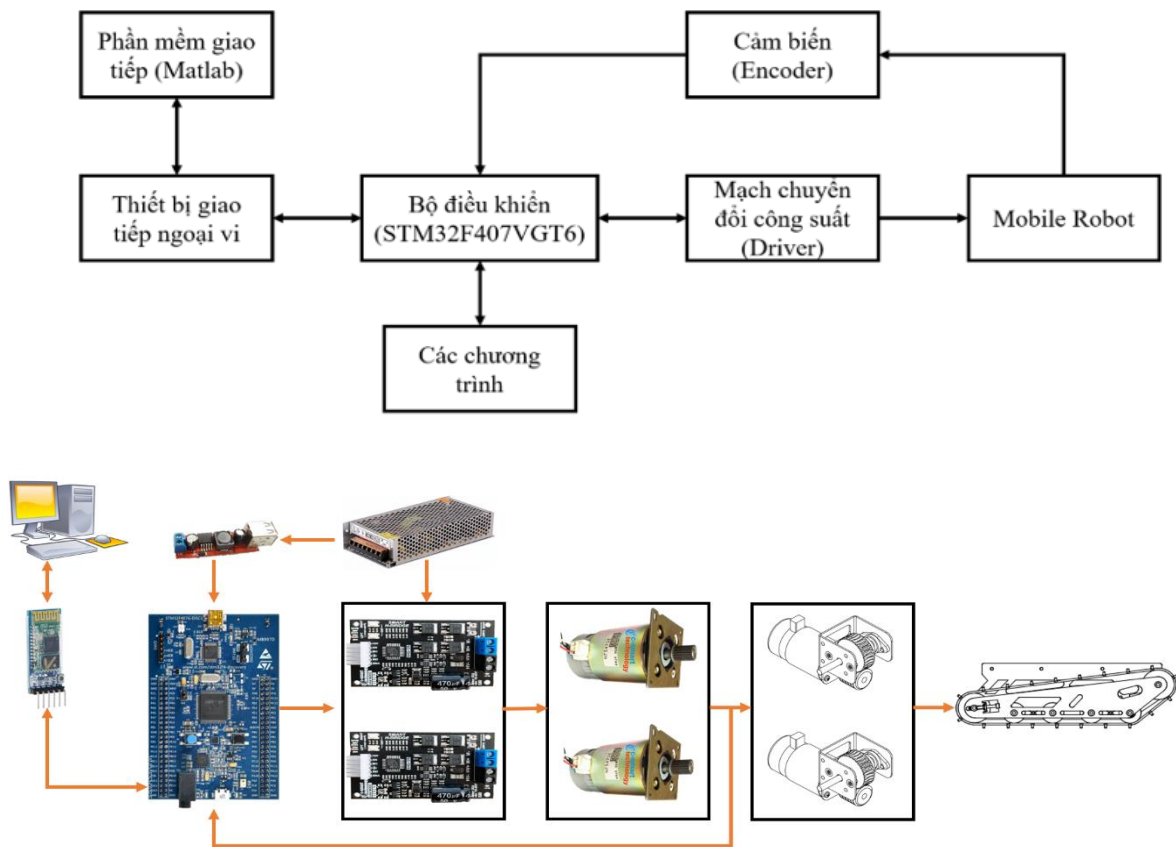
Liên hệ với công thức tính số xung cấp khi đi thẳng ta có:

$$n = \frac{x}{hesokc} \Rightarrow n = \frac{d\alpha\pi}{180 \cdot hesokc} \quad (24)$$

Như vậy, n là số xung cần cấp cho mỗi động cơ khi quay 1 góc α (độ) tại tâm giữa 2 trục bánh xe.

4. PHẦN CỨNG ROBOT

Sơ lược về sơ đồ phần cứng Mobile Robot gồm các thành phần: Phần mềm matlab để giao tiếp nhận dữ liệu từ vi điều khiển trung tâm (STM32F407VGT6); thiết bị ngoại vi (BLUETOOTH HC_05) là trung gian bắt tay giao tiếp giữa phần mềm matlab và vi điều khiển trung tâm (STM32F407VGT6); mạch chuyên đổi công suất (Driver) điều khiển động cơ DC servo được tích hợp chung với phần Mobile Robot; cảm biến đảm nhận vai trò đọc dữ liệu từ động cơ DC servo đưa về cho vi điều khiển trung tâm xử lý; các chương trình là các thuật toán được tính toán sẵn.



Hình 44 Sơ đồ khối phần cứng Mobile Robot

5. THIẾT KẾ BỘ ĐIỀU KHIỂN CHO ROBOT

Bộ điều khiển vi tích phân tỉ lệ (PID - Proportional Integral Derivative) là một bộ cơ chế phản hồi vòng điều khiển tổng quát (bộ điều khiển). Bộ điều khiển này được ứng dụng phổ biến trong những hệ thống điều khiển sản xuất công nghiệp. Trong các bộ điều khiển phản hồi, bộ điều khiển PID được ứng dụng nhiều nhất. Với bộ điều khiển vi tích phân tỉ lệ sẽ tính toán một giá trị "sai số", giá trị đó là hiệu số giữa giá trị đặt mong muốn và giá đo thông số biến đổi. Từ đó bộ điều khiển sẽ làm giảm thiểu tối đa giá trị sai số thông qua quá trình điều chỉnh giá trị tín hiệu điều khiển đầu vào. Bộ điều khiển PID sẽ là bộ điều khiển tốt nhất trong trường hợp người dùng không có những kiến thức cơ bản về hệ thống điều khiển.

Giải thuật của bộ điều khiển PID bao gồm ba giá trị thông số riêng biệt, vì vậy còn được gọi là điều khiển ba khâu, đó là: tỉ lệ, tích phân và đạo hàm, được viết tắt là P, I, và D. Giá trị tỉ lệ mô tả tác động của sai số của hiện tại, giá trị tích phân mô tả tác động của tổng các sai số của quá khứ, và giá trị vi phân mô tả tác động của tốc độ biến đổi giá trị sai số đó. Tổng chập của ba giá trị này được sử dụng để điều chỉnh quá trình hoạt động thông qua một thông số điều khiển như vận tốc động cơ hay vị trí của bánh xe. Từ đó, những giá trị này có thể biểu hiện rõ mối quan hệ về thời gian: I lệ thuộc vào tích lũy các sai số quá khứ, P lệ thuộc vào sai số hiện tại và D sẽ dự đoán những giá trị về sai số tương lai.

Bộ điều khiển PID sẽ điều chỉnh sai lệch $e(t)$ của hệ thống về 0 sao cho quá trình quá độ thỏa mãn với các yêu cầu cơ bản nhất về chất lượng:

Nếu sai lệch $e(t)$ chưa bằng 0 thì với giá trị $u_I(t)$, PID sẽ tạo tín hiệu điều chỉnh.

Nếu sai lệch tính $e(t)$ càng lớn thì với giá trị $u_P(t)$, tín hiệu điều chỉnh $u(t)$ sẽ càng lớn.

Bộ điều khiển PID được thể hiện với mô hình vào-ra như sau:

$$u(t) = K_R \left[e(t) + \frac{1}{T_C} \int_0^t e(\tau) d\tau + T_V \frac{de(t)}{dt} \right] \quad (25)$$

Trong đó:

K_R : hệ số tỷ lệ

T_C : hằng số thời gian tích phân

T_V : hằng số thời gian vi phân

Theo một số sách lý thuyết điều khiển tự động ta có hàm truyền rời rạc trên miền ảnh z:

$$\frac{U(z)}{E(z)} = G_{PID}(z) = K_p + \frac{K_i T}{2} \cdot \frac{z+1}{z-1} + \frac{K_d}{T} \cdot \frac{z-1}{z} \quad (26)$$

Ta được tín hiệu điều khiển PID cho hệ rời rạc:

$$u(k) = \frac{\alpha e(k) + \beta e(k-1) + \gamma e(k-2) + \Delta u(k-1)}{\Delta} \quad (27)$$

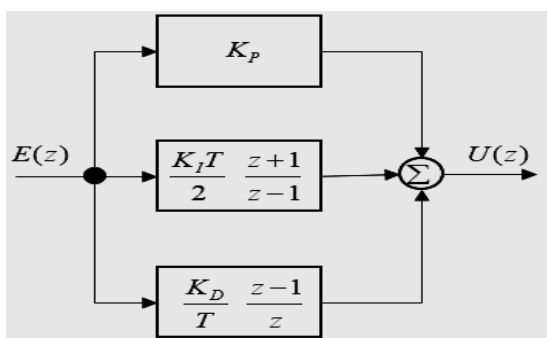
$$\alpha = 2TK_p + K_i T^2 + 2K_d$$

$$\beta = T^2 K_i - 4K_d - 2TK_p \quad (28)$$

$$\gamma = 2K_d$$

$$\Delta = 2T$$

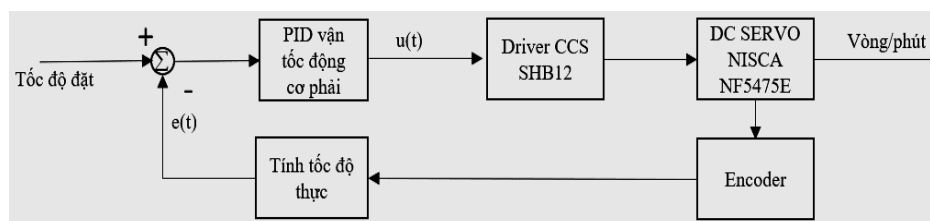
Với T: là thời gian chu kỳ lấy mẫu



Hình 45 PID rời rạc

5.1. Thiết kế bộ điều khiển PID vận tốc dùng vị trí kiểm soát

Bộ PID vận tốc cho 1 bánh lái:



Hình 46 PID rời rạc vận tốc cho 1 bánh lái

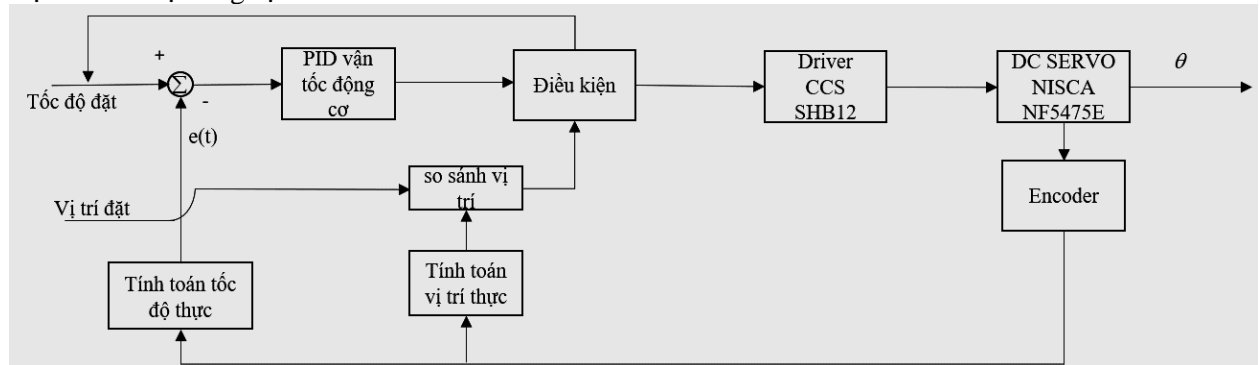
Tính toán giá trị vận tốc đo được để so sánh với giá trị vận tốc đặt:

Tính giá trị vận tốc dựa trên giá trị Encoder đọc về ta đổi ra số vòng/phút bằng công thức:

$$Tocdothuc = \frac{(Encoder_xung \cdot (60 \cdot (1000 / T)))}{PPR} \quad (29)$$

Encoder_xung : số xung hiện tại đọc từ Encoder.

PPR : độ phân giải encoder
 T : chu kỳ lấy mẫu
 Bộ PID tốc độ dùng vị trí kiểm soát:



Hình 47 PID rời rạc vận tốc dùng vị trí kiểm soát

Tính toán giá trị vị trí hiện tại:
 Chuyển giá trị xung sang giá trị độ

$$vitrihiientai = Encoder_xung \cdot \frac{360^0}{PPR} \quad (30)$$

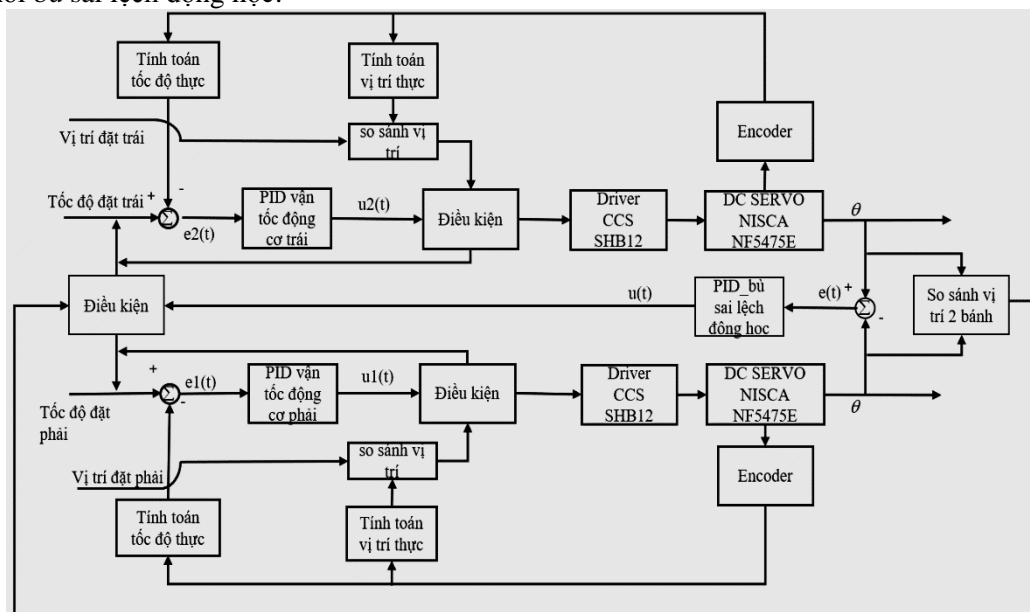
$Encoder_xung$: số xung hiện tại đọc từ Encoder.

PPR : độ phân giải encoder

Giải thích khối điều kiện của 2 bên bánh: ở đây khối điều kiện chia ra 3 giai đoạn, 2 giai đoạn đầu là tăng tốc và giảm tốc và giai đoạn cuối kiểm tra vị trí chính xác chưa, giai đoạn tăng tốc và giảm tốc tùy vào người thiết kế mong muốn. Ở phần thiết kế này, chia ra làm 2 đoạn là 90% quãng đường đầu sẽ là tăng tốc và 10% đoạn đường còn lại sẽ là giảm tốc vì khi động cơ quay sẽ có quán tính với tốc độ càng nhanh quán tính càng lớn do đó cần giảm tốc lại để giảm đi quán tính và làm giảm đi sai số khi gần đạt vị trí mong muốn.

5.2. Thiết kế bộ PID bù sai lệch động học

Sơ đồ khối bù sai lệch động học:



Hình 48 PID rời rạc bù sai lệch động học

Khối thuật toán PID bù động học: thuật toán PID tính sai số giữa góc quay bánh bên trái và góc quay bánh bên phải, bên nào lớn hơn thì thuật toán sẽ lấy góc quay bên lớn hơn trừ đi góc quay nhỏ hơn tính ra sai số và đưa ra tín hiệu điều khiển từ tín hiệu điều khiển đó kết hợp với khối thuật toán điều kiện để bù sai lệch động học cho xe.

Khối điều kiện: trong khối điều kiện chia ra làm 2 điều kiện là nếu góc quay bánh trái lớn hơn góc quay bánh phải thì tốc độ đặt ban đầu của bánh trái trừ đi tín hiệu điều khiển tính từ khối thuật toán PID bù động học và tốc độ đặt ban đầu bánh phải sẽ cộng cho tín hiệu điều khiển, ngược lại góc quay bánh phải lớn hơn góc quay bánh trái thì tốc độ đặt ban đầu của bánh phải trừ đi cho tín hiệu điều khiển của khối thuật toán PID bù động học và tốc độ đặt ban đầu bánh trái cộng cho tín hiệu điều khiển. Khi động cơ chạy độc lập 2 khối PID thì tốc độ bám vẫn có sai lệch dù là tốc độ đặt bằng nhau nhưng khả năng bám sẽ khác nhau do đó thông qua khối PID bù động học này tìm ra khoảng sai lệch của vị trí góc quay bánh phải và góc quay bánh trái để cân chỉnh vận tốc của 2 bên động cơ tìm ra điểm cân bằng.

- Mọi quan hệ giữa góc quay 2 bánh xe và động học Mobile Robot:

Để tính được tọa độ x, y ta suy luận từ góc quay 2 bánh xe bằng công thức sau:

Ta có công thức tính quãng đường đi của xe dựa trên thực nghiệm ở công thức số (17), bằng cách cho xe chạy 3 lần ở 3 vị trí khác nhau và dựa trên số xung xe chạy được ở 3 lần đó và quãng đường đi được suy ra được hệ số giữa số xung và quãng đường và chọn ra hệ số gần đúng nhất ta có công thức:

$$Lxe = n \cdot heso_kc$$

Mà n là số xung cần cấp cho bánh xe mà giá trị đặt ban đầu là số vòng suy ra góc quay ta áp dụng công thức (30):

$$vitrihientai = Encoder_xung \cdot \frac{360^0}{PPR}$$

Ở đây $Encoder_xung = n$, góc quay bánh xe = $vitrihientai$ như vậy, từ tọa độ góc quay ta tính ngược lại số xung và tính được quãng đường Lxe có được quãng đường, ta dùng công thức ở phần động học để tính ra được φ_r và φ_l từ công thức: $\varphi = \frac{180x}{\pi R}$

Có φ_r và φ_l ta suy ra được ω_r và ω_l dựa trên công thức (2) và (3):

$$\omega_r = \frac{d\varphi_r}{dt}$$

$$\omega_l = \frac{d\varphi_l}{dt}$$

Có ω_r và ω_l ta có được 2 tọa độ x, y và góc hướng θ dựa trên công thức (4), (5) và (6):

$$\dot{x} = \frac{(R \cos \theta)(\omega_r + \omega_l)}{2}$$

$$\dot{y} = \frac{(R \sin \theta)(\omega_r + \omega_l)}{2}$$

$$\dot{\theta} = \frac{R(\omega_r - \omega_l)}{2d}$$

Như vậy mối liên hệ giữa góc quay 2 bánh xe tỷ lệ thuận với tọa độ x, y và góc hướng θ , có nghĩa sai số ở góc quay 2 bánh xe càng lớn thì tọa độ x, y và góc hướng θ càng lớn và ngược lại sai số góc quay giữa 2 bánh xe càng nhỏ thì tọa độ x, y và góc hướng θ càng nhỏ khi đi thẳng. Do đó, khi bù sai lệch động học sẽ sử dụng vị trí 2 góc quay bánh xe để tính sai số ta có công thức:

Nếu $\theta_r > \theta_l$:

$$e_r = \theta_r - \theta_l \quad (31)$$

Nếu $\theta_r < \theta_l$:

$$e_r = \theta_l - \theta_r \quad (32)$$

Sau đó tính toán PID rời rạc cho ra giá trị tín hiệu điều khiển là $u(t)$ từ tín hiệu điều khiển này ta cộng trừ vào tốc độ đặt ban đầu của 2 động cơ ta có công thức:

Nếu $\theta_r > \theta_l$:

$$\begin{aligned} \text{toctodat_phai} &= \text{toctodat_phai} + u(t) \\ \text{toctodat_trai} &= \text{toctodat_trai} - u(t) \end{aligned} \quad (33)$$

Nếu $\theta_r < \theta_l$:

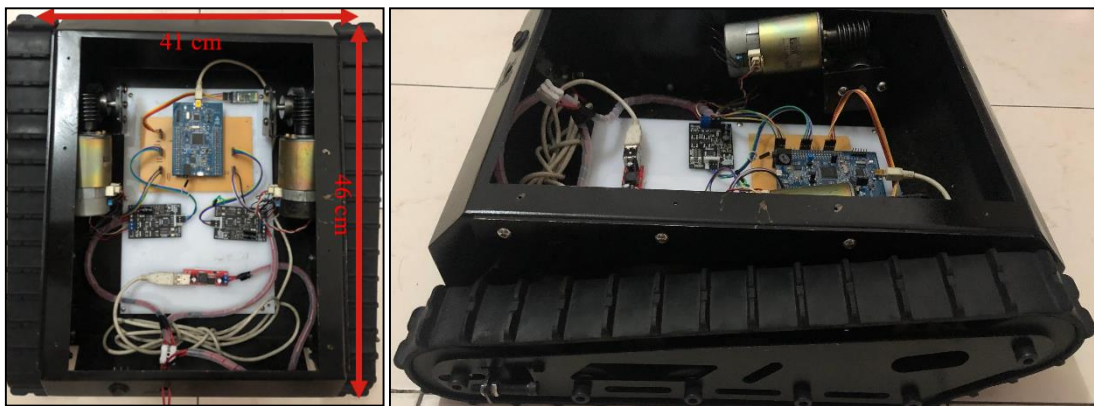
$$\begin{aligned} \text{toctodat_phai} &= \text{toctodat_phai} + u(t) \\ \text{toctodat_trai} &= \text{toctodat_trai} - u(t) \end{aligned} \quad (34)$$

Vì khi chạy 2 động cơ quay với một vận tốc đặt như nhau nhưng có nhiều yếu tố khác tác động (phần cơ khí mỗi bên sai lệch,...) làm cho bản chất 2 động cơ không thể chạy cùng với một vận tốc như nhau, khi đó dựa trên phần thuật toán bù sai lệch động học này để tìm ra khoảng cân bằng giữa 2 động cơ và giảm sai số khi tính toán động học. Vì sai lệch giữa 2 góc quay sẽ tỷ lệ thuận với ω_r và ω_l từ đó ảnh hưởng đến sai lệch khi tính toán tọa độ x, y và góc hướng θ .

6. THI CÔNG VÀ KẾT QUẢ THỰC NGHIỆM

6.1. Mô hình thực tế

Thực hiện được kết quả nghiên cứu dựa trên mô hình Mobile Robot bánh xích đã được thiết kế với cấu hình sau:

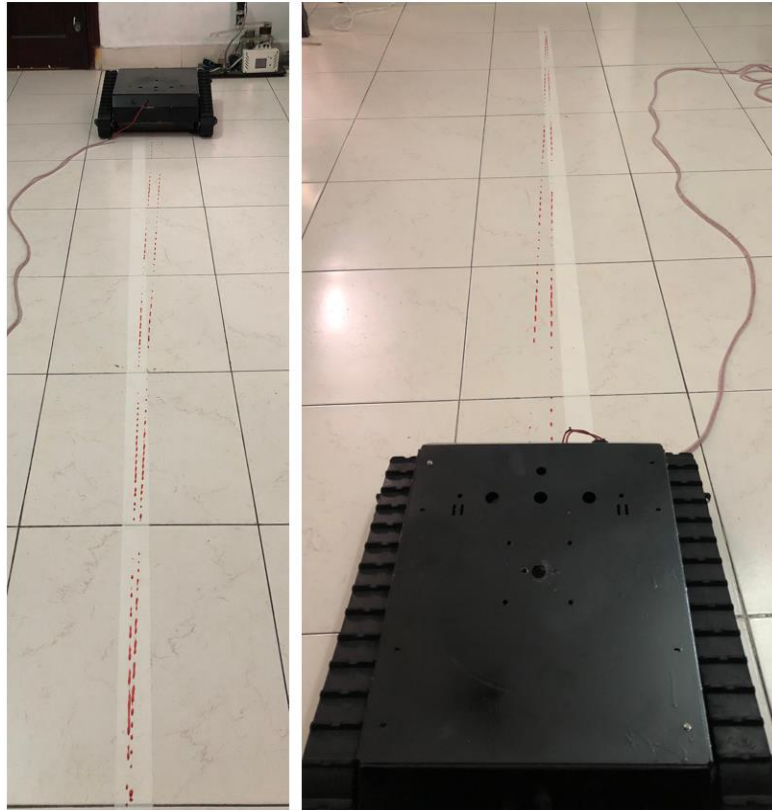


Hình 49 Cấu hình Mobile Robot thực tế
 Bảng 1. Các thông số cấu hình Mobile Robot

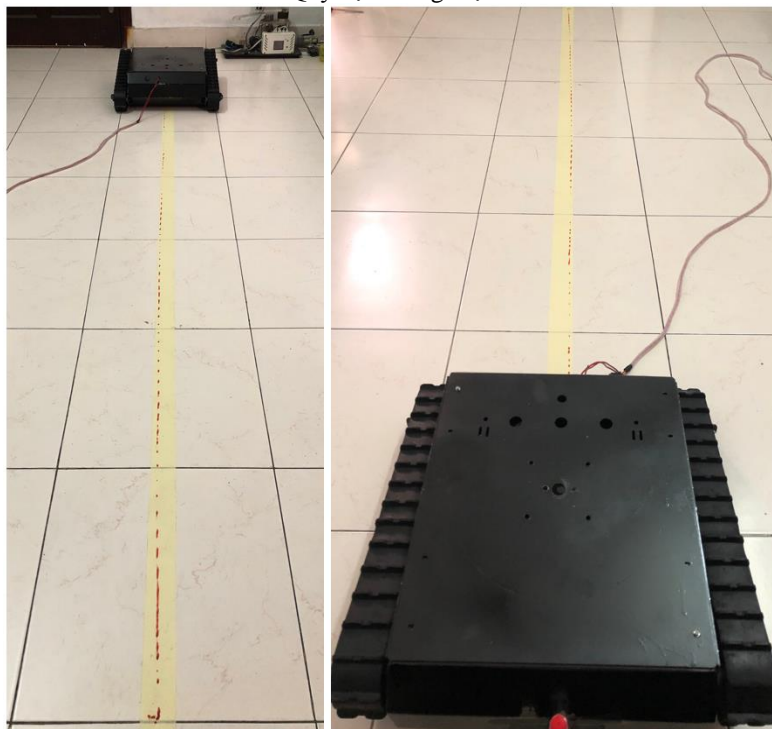
Mô tả	Số liệu	Đơn vị
Chiều dài xe	0.46	m
Chiều rộng xe	0.41	m
Bán kính bánh xe dẫn động	0.0325	m
Chu vi đai bánh xích	1.05	m
Tỷ lệ truyền bộ giảm tốc	30:1	
Tỷ lệ truyền bánh xích	14:11	

6.2. Kết quả thực nghiệm

6.2.1. Quỹ đạo thẳng thực tế

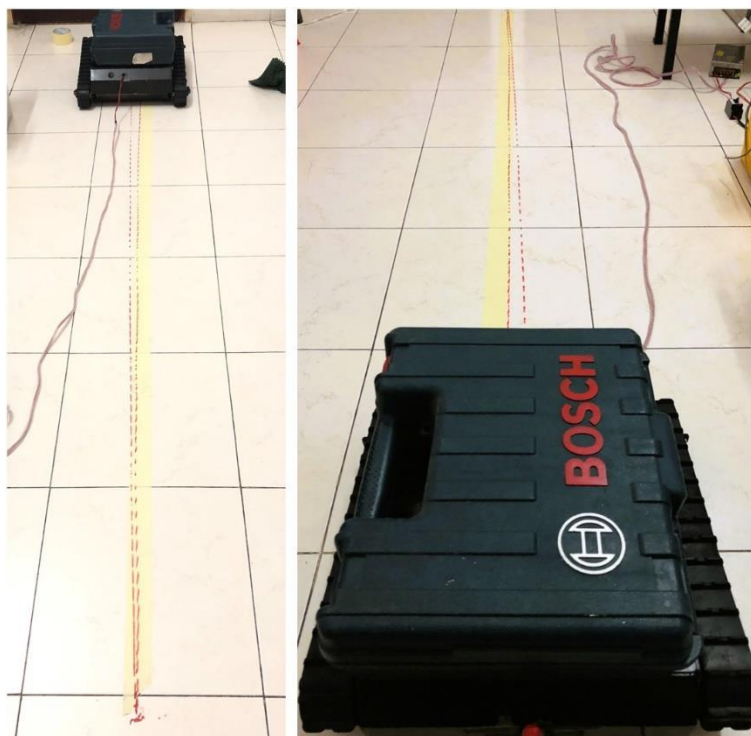


Hình 50 Quỹ đạo thẳng thực tế - PID

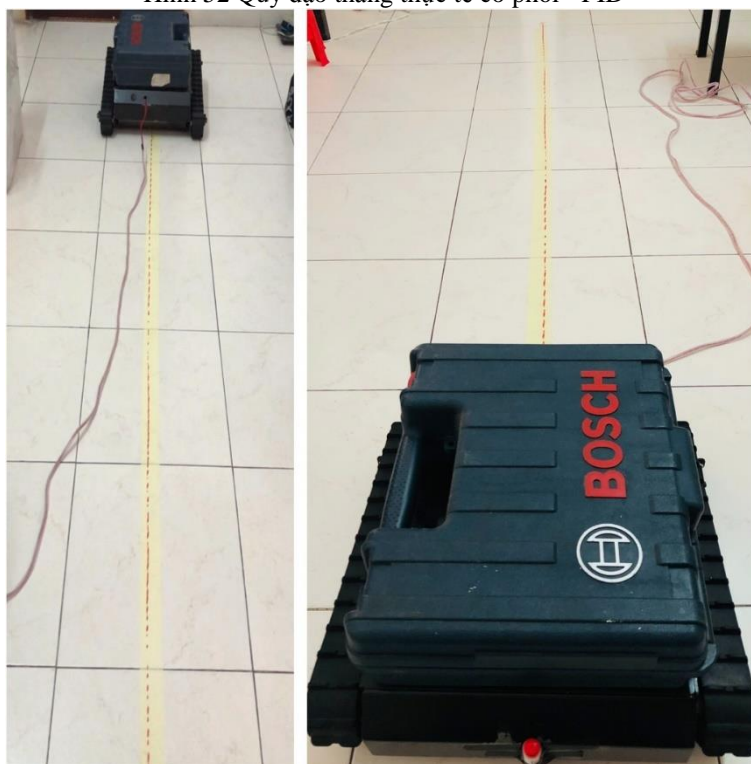


Hình 51 Quỹ đạo thẳng thực tế - PID bù sai lệch động học

6.2.2. Quỹ đạo thẳng có phôi thực tế

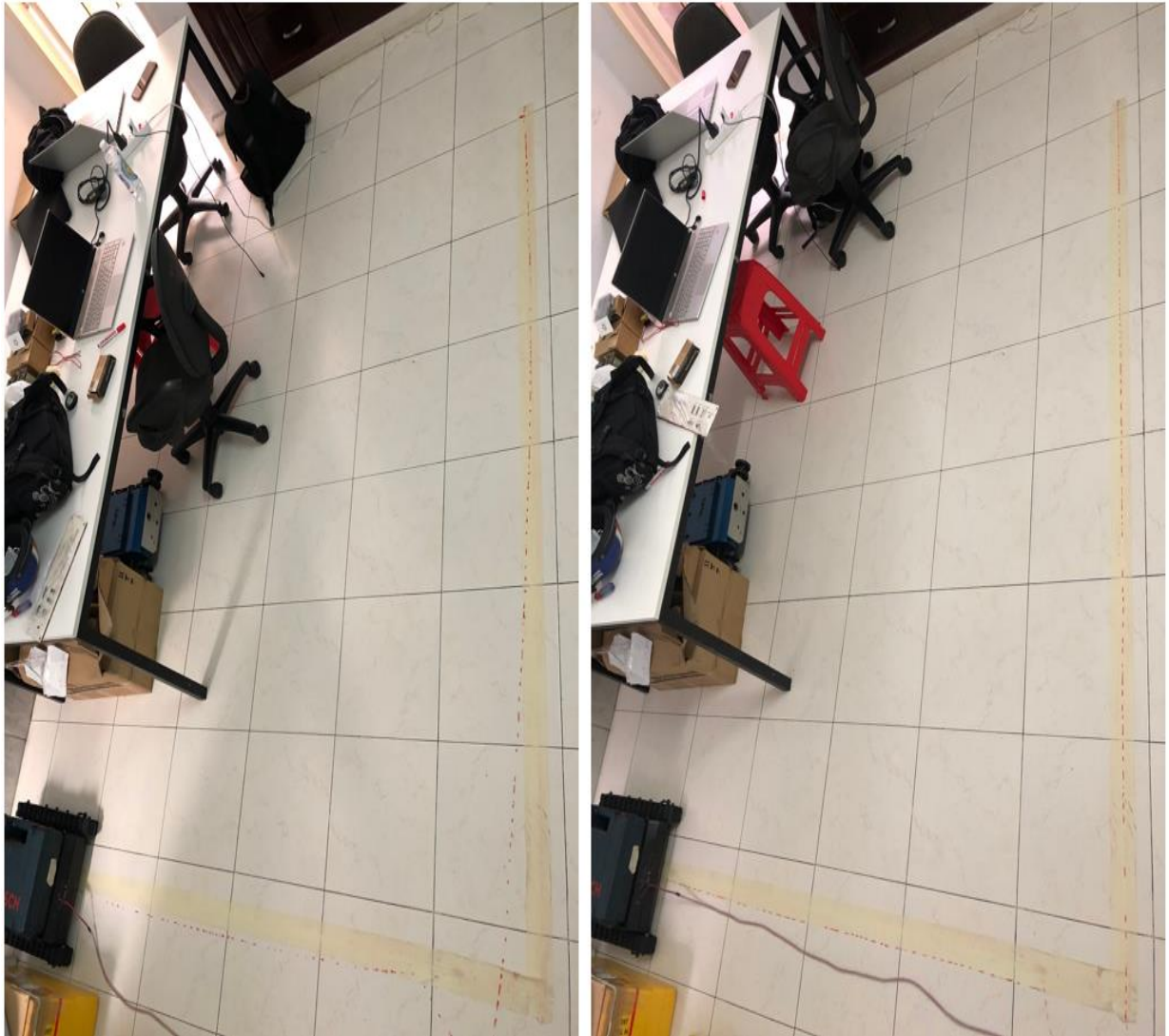


Hình 52 Quỹ đạo thẳng thực tế có phôi - PID



Hình 53 Quỹ đạo thẳng thực tế có phôi - PID bù sai lệch động học

6.2.3. Quỹ đạo L có phiê thực tế

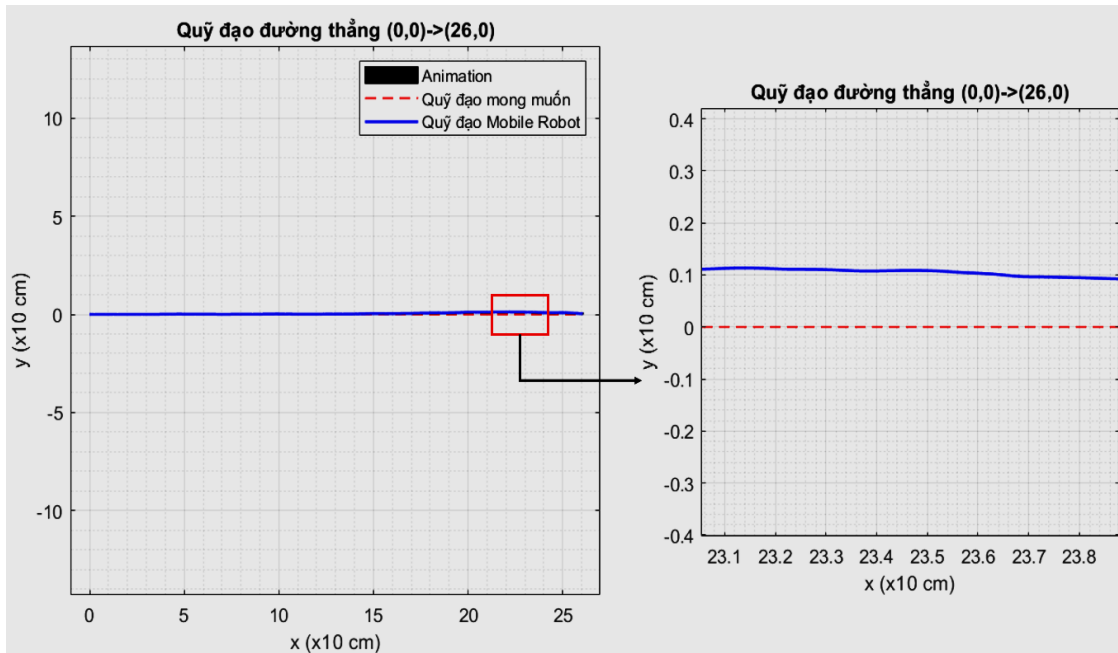


Hình 54 Quỹ đạo L thực tế so sánh giữa PID và PID bù sai lệch động học

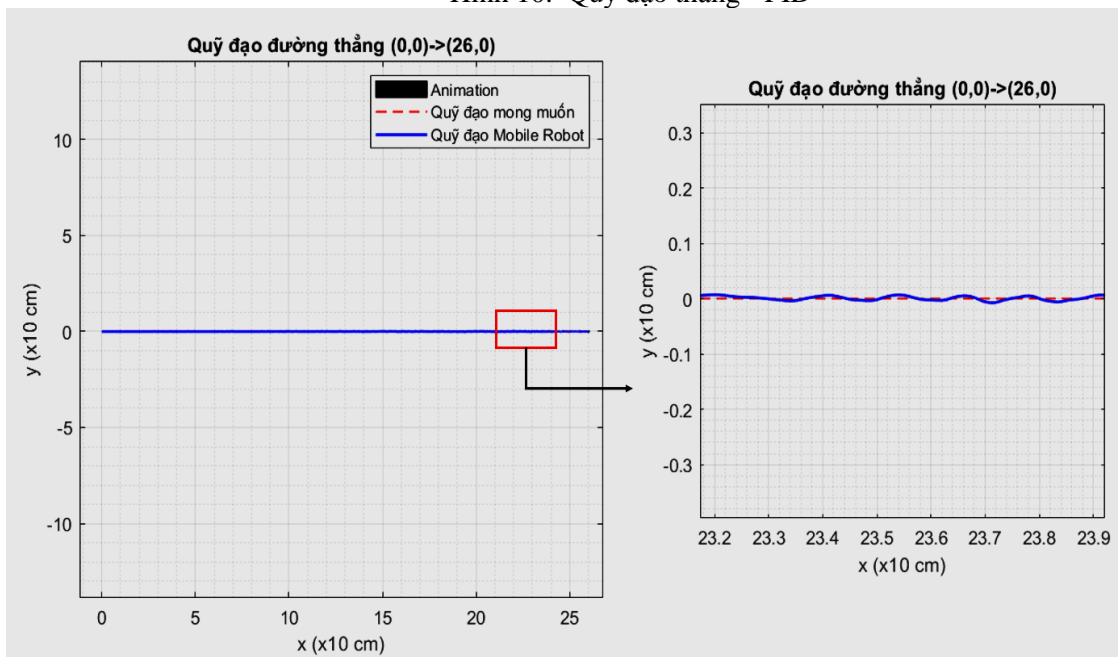
6.3. Đánh giá kết quả thực nghiệm

Để kiểm chứng hiệu quả của giải thuật điều khiển đề xuất, bài báo trình bày việc xây dựng một mô hình Mobile Robot di chuyển bằng 2 bánh xích được đánh giá quỹ đạo hoạch định qua phần mềm Matlab/m.file. Với tín hiệu đặt cho Mobile Robot là các quỹ đạo cho trước (đường thẳng, chữ L trái, chữ L phải...), bài báo tiến hành so sánh kết quả điều khiển lần lượt của 2 bộ điều khiển PID và PID có sử dụng thuật toán bù sai lệch động học.

6.3.1. Quỹ đạo di chuyển đường thẳng



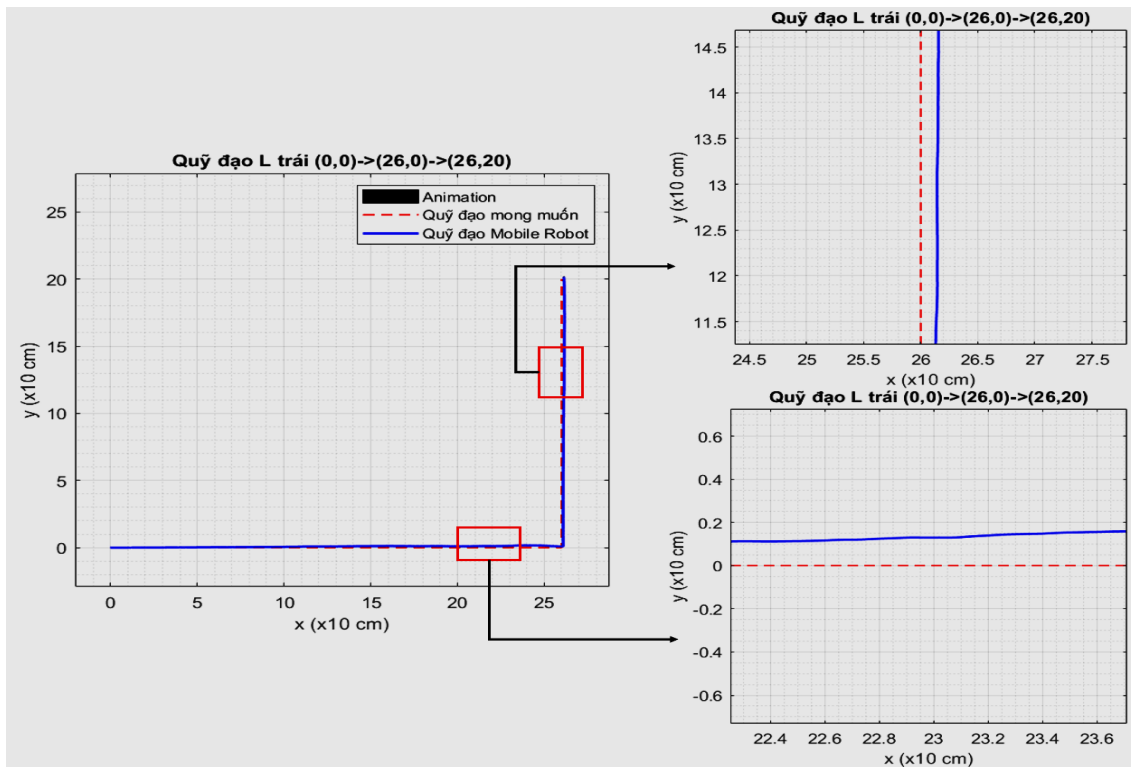
Hình 10. Quỹ đạo thẳng - PID



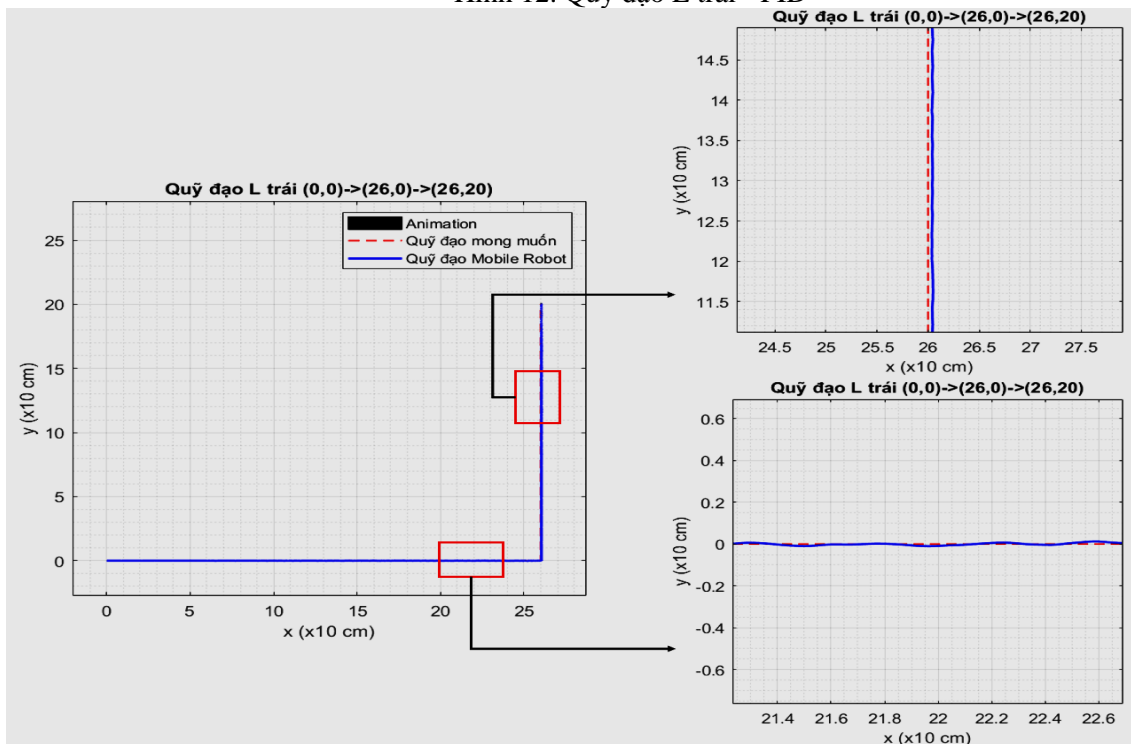
Hình 11. Quỹ đạo thẳng - PID bù sai lệch động học

Nhận xét: đánh giá quỹ đạo thẳng giữa 2 kết quả giữa 2 thuật toán PID và PID có bù sai lệch động học thì PID có bù sai lệch động học cho ra kết quả bám tốt hơn so với PID thường khi đi đường thẳng.

6.3.2. Quỹ đạo di chuyển L trái



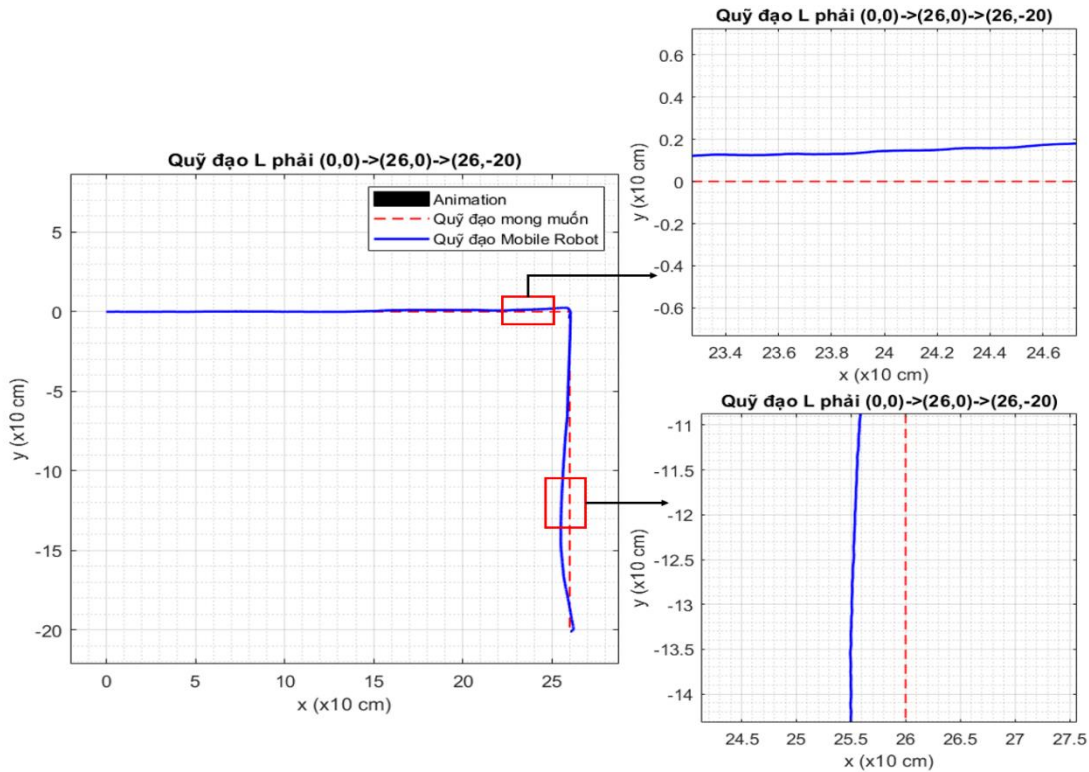
Hình 12. Quỹ đạo L trái - PID



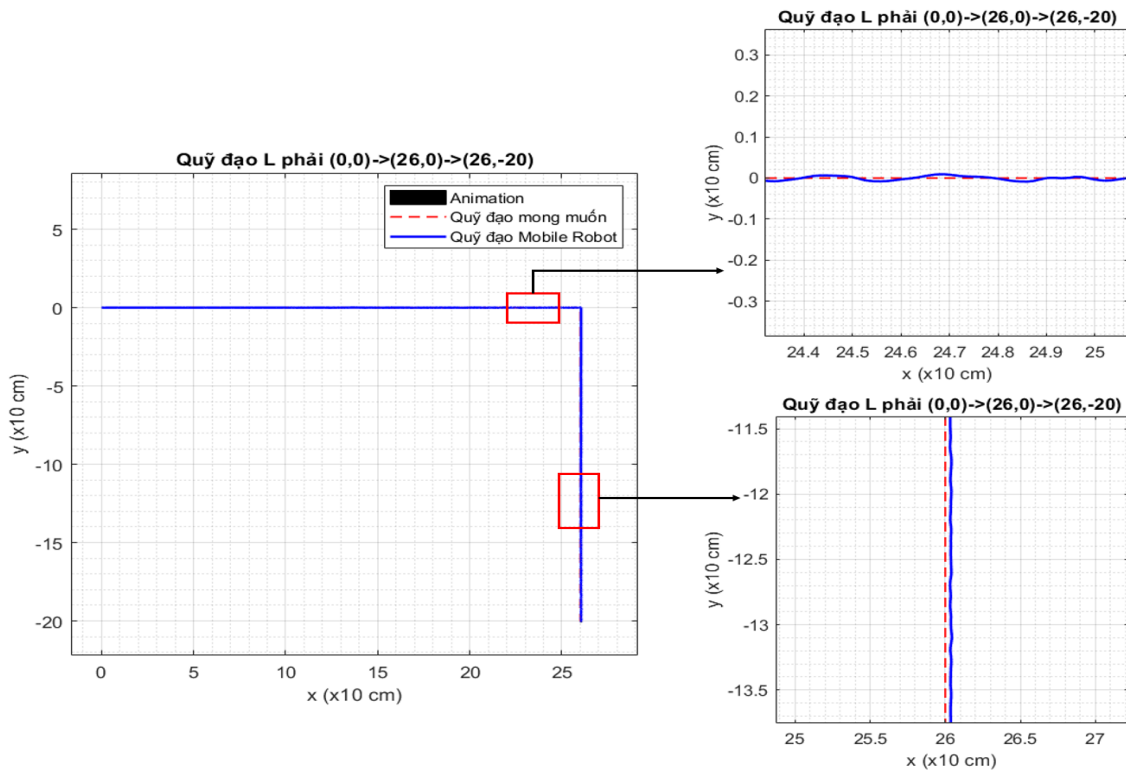
Hình 13. Quỹ đạo L trái - PID bù sai lệch động học

Nhận xét: đánh giá quỹ đạo L trái cho ra kết quả tương tự như quỹ đạo thẳng thuật toán PID có bù sai lệch động học cho ra kết quả bám tốt hơn so với PID thường.

6.3.3. Quỹ đạo di chuyển L phải



Hình 14. Quỹ đạo L phải - PID



Hình 15. Quỹ đạo L phải - PID bù sai lệch động học

Nhận xét: đánh giá quỹ đạo L phải, thuật toán PID có bù sai lệch động học cho ra kết quả bám tốt hơn và ít sai lệch hơn so với PID thường.

Từ kết quả trên, có thể cho ta thấy trường hợp sử dụng PID có sử dụng thuật toán bù sai lệch động học cho 2 bên bánh Mobile Robot cho ra kết quả tốt hơn khi sử dụng bộ điều khiển PID không có thuật toán bù sai lệch động học. Kết quả cho thấy khi sử dụng bộ điều khiển PID có sử dụng thuật toán bù sai lệch động học có biên độ dao động nhỏ hơn và sai số của góc hướng θ ít hơn bộ điều khiển PID.

Để kiểm tra hiệu quả bộ điều khiển PID có sử dụng thuật toán bù sai lệch động học, ta so sánh góc hướng θ mong muốn và góc hướng θ thực tế (vì khi góc hướng sai lệch so với góc hướng mong muốn thì dẫn đến quỹ đạo thay đổi theo). Do kết quả khó có thể đánh giá qua quan sát trực tiếp nên ta thực hiện so sánh định lượng, sử dụng công thức sai số bình phương trung bình (MSE) như sau:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \tilde{y}_i)^2 \quad (35)$$

Bảng 2. Sai số bình phương trung bình (MSE) của các kết quả thực nghiệm

Thí nghiệm		Quỹ đạo		
		Thẳng θ	Chữ L trái θ	Chữ L phải θ
Lần 1	PID	$1,1503.10^{-5}$	0,0017	0,0014
	PID (bù động học)	$1,0399.10^{-7}$	$9,3969.10^{-4}$	0,0011
Lần 2	PID	$2,3828.10^{-5}$	0,0015	0,0015
	PID (bù động học)	$1,1040.10^{-7}$	$9,5971.10^{-4}$	0,0011

Dựa vào sai số định lượng được tính theo công thức sai số bình phương trung bình, có thể đánh giá rằng trường hợp điều khiển PID có sử dụng thuật toán bù sai lệch động học cho ra kết quả tốt hơn.

7. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Qua quá trình thực hiện xây dựng phương pháp hoạch định quỹ đạo và đưa ra thuật toán bù sai lệch động học để điều khiển cho Robot di động ứng dụng trong thực tế. Kết quả cho thấy chất lượng của bộ điều khiển PID có sử dụng thuật toán bù sai lệch động học tốt hơn hẳn bộ điều khiển PID thông thường thông qua việc đánh giá tọa độ. Việc sử dụng thuật toán mềm cho bộ PID để bù sai lệch động học cho thấy chất lượng điều khiển tốt hơn khi Robot di chuyển theo đường thẳng, sai số góc hướng di chuyển nhỏ hơn so với khi sử dụng bộ điều khiển PID không có bù sai lệch động học, các đánh giá trên được thực hiện trong môi trường thực nghiệm lý tưởng (ma sát thấp giữa bánh xích và sàn để tránh hiện tượng sai lệch về mặt cơ khí). Do điều kiện về thời gian, cơ sở vật chất nên bài báo còn nhiều hạn chế. Tuy nhiên, từ các kết quả thực nghiệm và đánh giá đã cho thấy hiệu quả của việc xây dựng thuật toán mềm cho bộ điều khiển PID để bù sai lệch động học so với bộ điều khiển thông thường khi Robot di chuyển thẳng. Bài nghiên cứu này đưa ra hướng giải pháp dựa trên thuật toán mềm để triệt tiêu sai số thấp nhất nhưng chưa giải quyết được vấn đề cơ khí để bù sai lệch quá trình đó nên tích hợp thêm một số cảm biến khoảng cách, góc quay để kiểm soát tốt hơn. Việc phát triển từ cơ sở lý thuyết, động học Mobile Robot thành mô hình thực tế rất hữu ích và có khả năng ứng dụng trong công nghiệp, các chuyên ngành về Mobile Robot.

TÀI LIỆU THAM KHẢO

- [1] G. O. a. M. V. Alessandro De Luca, "Control of Wheeled Mobile Robots: An Experimental Overview", Dipartimento di Informatica e Sistemistica.
- [2] Y. Ali Gholipour, "Dynamic tracking control of Nonholonomic Mobile Robot with Model Reference Adaptation for", pp.3188-3122,2003.
- [3] Egerstedt, "Control of mobile robots", 2013.

- [4] C. B. H. Maaref, "Sensor-based navigation of a mobile robot in an indoor environment", *Robotics and Autonomous System*, Vol 38 (2002), pp 1-18.
- [5] A. M. , M. P. G.-C. J. L.Martínez, "Approximating Kinematics for Tracked Mobile Robots".
- [6] G. Mester, "Motion Control of Wheeled Mobile Robots", Serbia: Department of Informatics, Polytechnical Engineering College, 2006.
- [7] G. Mester, "Obstacle Avoidance of Mobile", *International Symposium on Intelligent*, 2007.
- [8] K. N.Leena, "Modelling and trajectory tracking of wheeled mobile robots", India: Federal Institute of Science and Technology, India.
- [9] A. A. A. W. P. a. M. Y. M. Nwe, "Software implementation of obstacle detection and for wheeled mobile robot avoidance", *Engineering and Technology*, 42, 572–577.
- [10] P. C. Tú, "Thiết kế bộ điều khiển chuyển động cho robot tự hành di chuyển theo đường hoạch định dựa trên logic mờ", Đại Học Sư Phạm Kỹ Thuật Thành phố Hồ Chí Minh
- [11] n. h. p. Đinh văn thịnh, "Nghiên cứu điều khiển mobile robot vận chuyển hàng hóa theo quỹ đạo", Trường Đại học Sư phạm Kỹ thuật Thành phố Hồ Chí Minh, 2016.
- [12] v. m. Peri, "Fuzzy logic controller for an autonomous", India: Jawaharlal Nehru Technological University, India, 2002.
- [13] L. T. Hoàn, "Ứng dụng mạng nơron điều khiển Mobile Robot", Trường Đại học Kỹ thuật Công nghệ Thành phố Hồ Chí Minh, 2012.
- [14] V. M. Đ. Tăng Quốc Nam, "Nghiên cứu thiết kế hệ thống di động cho robot tự hành có khả năng vận động theo quỹ đạo cho trước", Đề tài NCKH cấp Học viện Kỹ thuật Quân sự, 2005.

YSC4F.229

XÂY DỰNG HỆ THỐNG QUẢN LÝ NHÂN SỰ CHO CÔNG TY CỔ PHẦN BAO BÌ PHƯƠNG NAM

NGUYỄN VĂN CHUẨN, NGUYỄN ĐĂNG TUẤN ĐẠT, NGUYỄN THỊ HẠNH
Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
vanchuan21123@gmail.com, tuandatmtd@gmail.com, ngthihanh@iuh.edu.vn

Tóm tắt. Phần mềm nhân sự đang là giải pháp hiệu quả trong quản trị nguồn nhân lực của doanh nghiệp. Đặc biệt trong các doanh nghiệp có quy mô lớn với số lượng nhân sự lớn được phân bố ở nhiều bộ phận khác nhau và hoạt động sản xuất diễn ra liên tục thì phần mềm quản trị nhân sự lại đóng vai trò quan trọng. Nắm bắt nhu cầu thực tế, chúng tôi quyết định xây dựng hệ thống quản lý nhân sự trên nền tảng của Odoo với mục đích giúp doanh nghiệp dễ dàng quản lý nhân sự một cách tự động, dễ dàng lưu trữ và tiết kiệm chi phí. Hệ thống cho phép doanh nghiệp quản lý từ khâu tuyển dụng, quản lý nhân viên, quản lý hợp đồng, chấm công ngày nghỉ và tính lương. Hệ thống cũng cung cấp cho lãnh đạo các cấp trong doanh nghiệp các báo cáo quản trị linh động, mạnh mẽ, chuyên sâu và hỗ trợ lập kế hoạch nguồn lực hằng năm. Hệ thống này được áp dụng cho công ty cổ phần Bao bì Phương Nam.

Từ khóa. nhân sự, quản lý nhân sự, Odoo

BUILDING HUMAN RESOURCE MANAGEMENT SYSTEM FOR PHUONG NAM PACKAGING CORPORATION

Abstract. Human resource software is an effective solution in human resource management of enterprises. Especially in large-scale enterprises with a large number of employees distributed in many different departments and continuous production activities, human resource management software plays an important role. Grasping the actual needs, we decided to build a human resource management system on the Odoo platform with the aim of helping businesses easily manage personnel automatically, easily store and save costs. Our system allows enterprises to manage from recruitment, employee management, contract, time-off, payroll. Our system also provides leaders at all levels in the business with flexible, powerful, in-depth management reports and supports annual resource planning. This system is applied to Phuong Nam Packaging Corporation.

Keywords. Human Resources, Human Resource Management, Odoo

1. GIỚI THIỆU

Ngày nay, Việt Nam đã bước vào giai đoạn hội nhập quốc tế sâu sắc và toàn diện hơn bao giờ hết. Trong bối cảnh đó con người luôn là nguồn lực quan trọng hàng đầu, quyết định sự thành bại của mỗi tổ chức, doanh nghiệp. Bởi vì một dự án có kế hoạch tốt đến đâu, công nghệ hiện đại đến đâu nhưng nếu không có một nguồn lực đủ số lượng và chất lượng để triển khai thì không thể thành công được. Yếu tố nhân sự của công ty do đó luôn được coi trọng hàng đầu như nhiệm vụ sống còn trong thời đại cạnh tranh hiện nay, yêu cầu các công ty phải quản lý và thu hút được nguồn lực nhân sự,... để giữ cho tổ chức mình có thể hoạt động tốt. Tuy nhiên, thực trạng hiện nay việc quản lý nhân sự không được chú trọng trong các doanh nghiệp cụ thể là công ty cổ phần bao bì Phương Nam đang lưu trữ thông tin nhân sự trên giấy tờ, không có các quy trình chuẩn để làm việc gây nên việc quản lý nhân sự trong công ty không được hiệu quả, ảnh hưởng đến năng suất làm việc chung. Từ những điều đã phân tích ở trên, chúng tôi đã quyết định xây dựng một hệ thống quản lý nhân sự với tên “XÂY DỰNG HỆ THỐNG QUẢN LÝ NHÂN SỰ CHO CÔNG TY CỔ

PHẦN BAO BÌ PHƯƠNG NAM” nhằm giải quyết các vấn đề về quản lý dữ liệu và thực hiện các quy trình cơ bản của bộ phận nhân sự trong công ty.

2. CƠ SỞ LÝ THUYẾT

2.1. Quản lý nhân sự trong doanh nghiệp

Quản lý nhân sự hiểu một cách đơn giản là công tác quản lý, khai thác, sử dụng lao động - nhân sự - của một công ty, doanh nghiệp, tổ chức, ... một cách hợp lý và hiệu quả. Chức năng quản lý nhân sự liên quan đến các công việc như thu hút, tuyển dụng, đào tạo, tính lương, đánh giá, khen thưởng nhân viên, ... Không chỉ vậy quản lý nhân sự còn liên quan đến các công tác kiểm tra, điều chỉnh nhằm đáp ứng nhu cầu và sử dụng, duy trì, bảo vệ, phát triển tiềm năng của người lao động một cách hiệu quả nhất.



Hình 1. Quản lý nhân sự trong doanh nghiệp [1]

Đối với doanh nghiệp, quản lý nhân sự đóng vai trò rất quan trọng giúp doanh nghiệp có sự đồng bộ, ổn định trong việc thực thi các quy định, nhiệm vụ, đảm bảo các bộ phận đều hoạt động một cách trơn tru. Một doanh nghiệp muốn hoạt động hiệu quả thì trước hết mỗi nhân viên cần hoạt động hiệu quả. Đảm bảo quyền lợi của người lao động, xây dựng môi trường làm việc thoải mái sẽ kích thích sự nhiệt tình, sáng tạo và tình yêu nghề của mỗi người lao động. Năng suất mỗi người đều tăng lên một chút cũng đủ để giúp doanh nghiệp thu được lợi nhuận khủng. Quan trọng hơn là bộ máy nhân sự càng ngày càng vững chắc, đoàn kết, ổn định.

2.2. Tổng quan về Enterprise Resource Planning (ERP) và phần mềm Odoo

Hoạch định tài nguyên doanh nghiệp (ERP) là một hệ thống dùng để hoạch định tài nguyên trong một tổ chức, một doanh nghiệp. Một hệ thống ERP điển hình bao hàm tất cả những chức năng cơ bản của một tổ chức. Một phần mềm ERP, nó tích hợp những chức năng chung của một tổ chức vào trong một hệ thống duy nhất. Thay vì phải sử dụng phần mềm kế toán, phần mềm nhân sự-tiền lương, quản trị sản xuất... song song, độc lập lẫn nhau thì ERP gom tất cả vào chung 1 gói phần mềm duy nhất mà giữa các chức năng đó có sự liên thông với nhau.

Odoo là một phần mềm ERP mã nguồn mở (open-source), nghĩa là khả năng tùy chỉnh và phát triển các tính năng của phần mềm này là vô hạn. Cụ thể hơn, ngoài các phân hệ hay module cơ bản của Odoo như POS (Point of Sale), CRM (Customer relationship management), quản lý kho, quản lý nhân sự, ... thì Odoo cho phép bạn hoàn toàn chỉnh sửa hay thêm bớt những tính năng hoặc tạo ra các phân hệ mới mà bạn mong muốn. Ngoài ra, Odoo còn cung cấp các tính năng bảo mật cho các cộng đồng về công nghệ kinh doanh và phát triển phần mềm trên toàn thế giới.

Hệ thống quản lý nhân sự cho công ty cổ phần bao bì Phương Nam được xây dựng trên Odoo với các chức năng chính bao gồm quản lý tuyển dụng, quản lý nhân viên, quản lý ngày nghỉ và quản lý bảng lương.



Hình 2. Chức năng của phần mềm Odoo [2]

2.3. Tổng quan về ngôn ngữ Python

Python là ngôn ngữ lập trình hướng đối tượng, cấp cao, mạnh mẽ, được tạo ra bởi Guido van Rossum. Nó dễ dàng để tìm hiểu và đang nổi lên như một trong những ngôn ngữ lập trình nhập môn tốt nhất cho người lần đầu tiếp xúc với ngôn ngữ lập trình. Python hoàn toàn tạo kiểu động và sử dụng cơ chế cấp phát bộ nhớ tự động. Python có cấu trúc dữ liệu cấp cao mạnh mẽ và cách tiếp cận đơn giản nhưng hiệu quả đối với lập trình hướng đối tượng. Cú pháp lệnh của Python là điểm cộng vô cùng lớn vì sự rõ ràng, dễ hiểu và cách gõ linh động làm cho nó nhanh chóng trở thành một ngôn ngữ lý tưởng để viết script và phát triển ứng dụng trong nhiều lĩnh vực, ở hầu hết các nền tảng. Bản thân Odoo được xây dựng dựa trên ngôn ngữ lập trình Python.

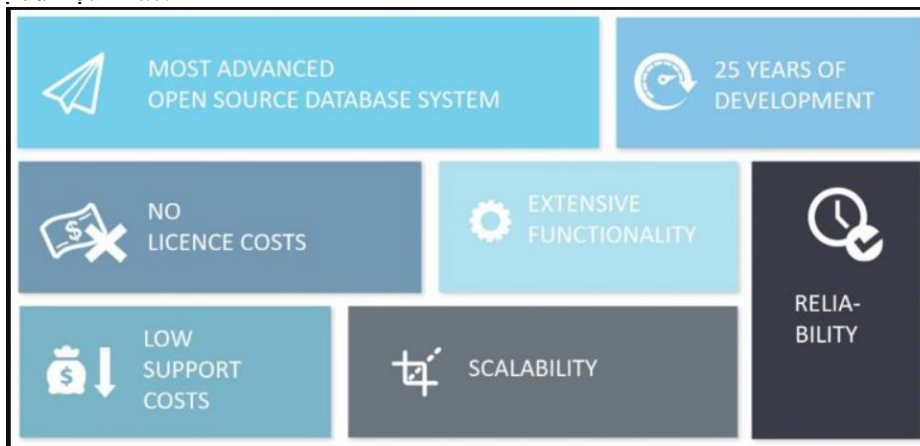


Hình 3. Chức năng của ngôn ngữ Python [3]

2.4. Tổng quan về PostgreSQL

PostgreSQL là một hệ thống quản trị cơ sở dữ liệu quan hệ đối tượng (object-relational database management system) có mục đích chung, hệ thống cơ sở dữ liệu mã nguồn mở tiên tiến nhất hiện nay, được phát triển dựa trên POSTGRES 4.2 tại phòng khoa học máy tính Berkeley, Đại học California. PostgreSQL được thiết kế để chạy trên các nền tảng tương tự UNIX. Tuy nhiên, PostgreSQL sau đó cũng được điều chỉnh linh động để có thể chạy được trên nhiều nền tảng khác nhau như Mac OS X, Solaris và Windows.

PostgreSQL là một phần mềm mã nguồn mở miễn phí. Mã nguồn của phần mềm khả dụng theo license của PostgreSQL, một license nguồn mở tự do. Theo đó, bạn sẽ được tự do sử dụng, sửa đổi và phân phối PostgreSQL dưới mọi hình thức. PostgreSQL không yêu cầu quá nhiều công tác bảo trì bởi có tính ổn định cao. Do đó, nếu bạn phát triển các ứng dụng dựa trên PostgreSQL, chi phí sở hữu sẽ thấp hơn so với các hệ thống quản trị dữ liệu khác.



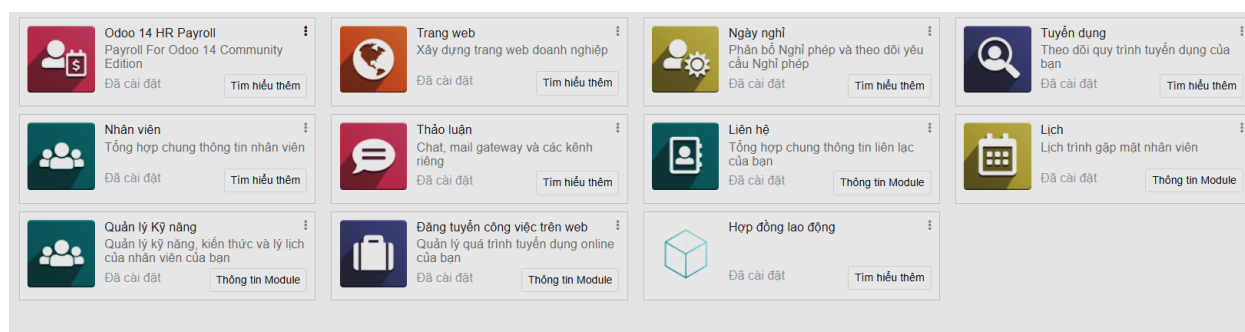
Hình 4. Lợi ích của PostgreSQL [4]

3. TRIỂN KHAI HỆ THỐNG QUẢN LÝ NHÂN SỰ

3.1. Kiến trúc hệ thống đề xuất

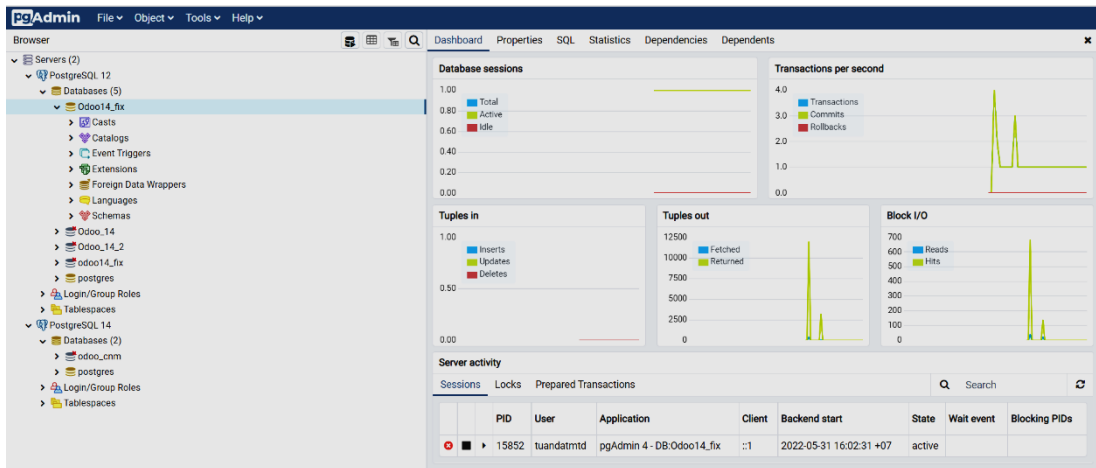
Để triển khai hệ thống quản lý nhân sự thì chúng tôi đề xuất cài đặt các phân hệ trên Odoo 14 Community như hình 5, cụ thể các phân hệ này áp dụng với các quy trình như sau:

- Phân hệ quản lý tuyển dụng: Tuyển dụng, Đăng tuyển công việc lên web, Trang web, Lịch.
- Phân hệ quản lý thông tin nhân viên: Nhân viên, Hợp đồng lao động, Liên hệ, Quản lý kỹ năng, Thảo luận.
- Phân hệ quản lý ngày nghỉ: Ngày nghỉ.
- Phân hệ quản lý bảng lương: Odoo 14 HR Payroll.



Hình 5. Các phân hệ sử dụng trong hệ thống

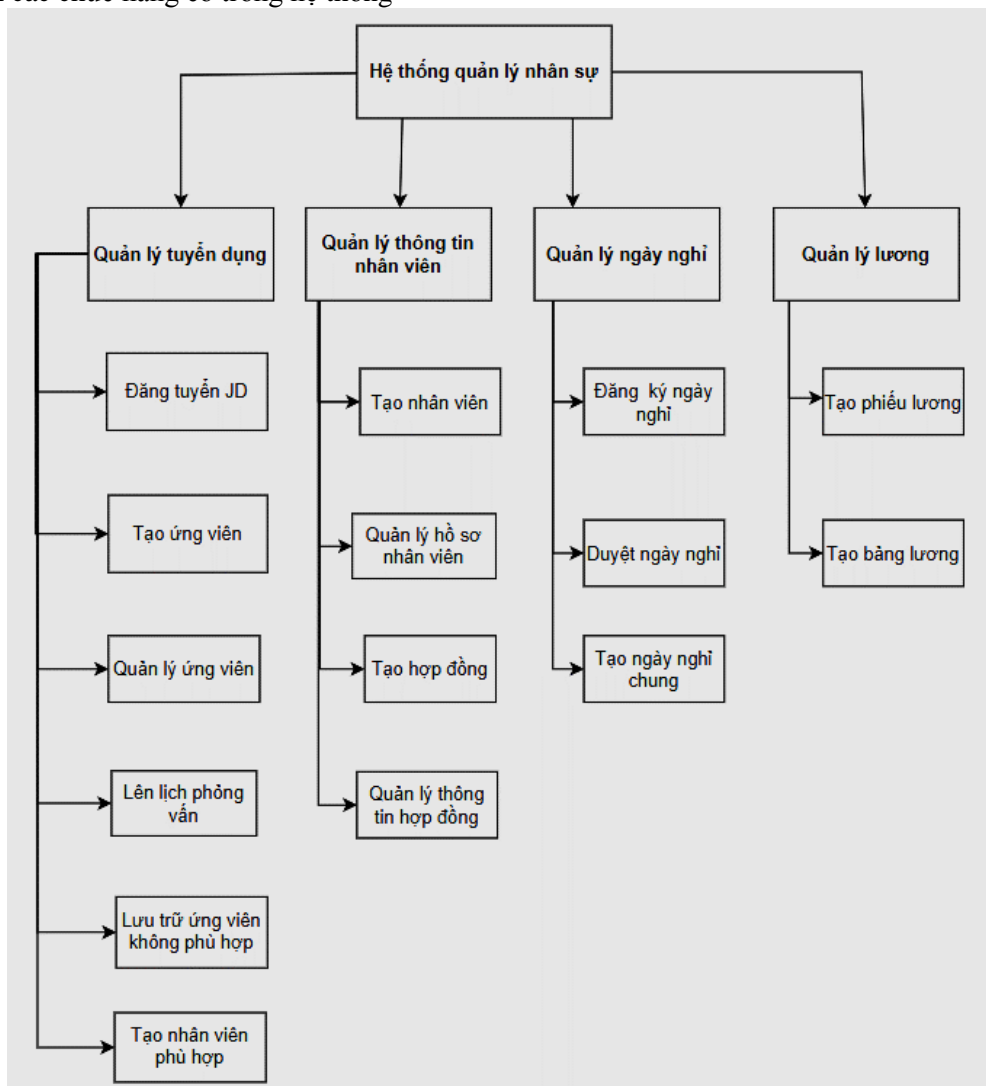
Bên cạnh các phân hệ để sử dụng, chúng tôi tiến hành cài đặt hệ thống quản trị cơ sở dữ liệu PostgreSQL 12 là một dạng cơ sở dữ liệu quan hệ - đối tượng để lưu trữ toàn bộ dữ liệu của hệ thống được thể hiện như hình 6.



Hình 6. Hệ thống cơ sở dữ liệu PostgreSQL 12

3.2. Phân tích và thiết kế hệ thống

Tổng quan các chức năng có trong hệ thống



Hình 7. Sơ đồ phân cấp các chức năng

Về hệ thống quản lý nhân sự có các chức năng chính bao gồm: quản lý quy trình tuyển dụng, quản lý thông tin nhân viên – hợp đồng, quản lý quy trình đăng ký ngày nghỉ của nhân viên và tính toán lương một cách tự động. Các quy trình này có sự gắn kết với nhau để quản trị hệ thống một cách có tổ chức cụ thể như sau: ứng viên thông qua website tuyển dụng để nộp đơn ứng tuyển và được bộ phận tuyển dụng thực hiện quy trình tuyển dụng của công ty; trải qua quá trình tuyển dụng thành công ứng viên sẽ trở thành nhân viên cho công ty và được tạo hợp đồng; sau khi trở thành nhân viên của công ty nhân viên có thể thực hiện chức năng đăng ký các ngày nghỉ của mình và trưởng phòng (quản lý cấp trên) duyệt yêu cầu nghỉ đó; dựa vào thông tin của hợp đồng và số lượng ngày nghỉ của nhân viên, nhân viên nhân sự sẽ tiến hành tạo các phiếu lương và bảng lương một cách tự động để hoàn tất các chức năng hiện có trong hệ thống.

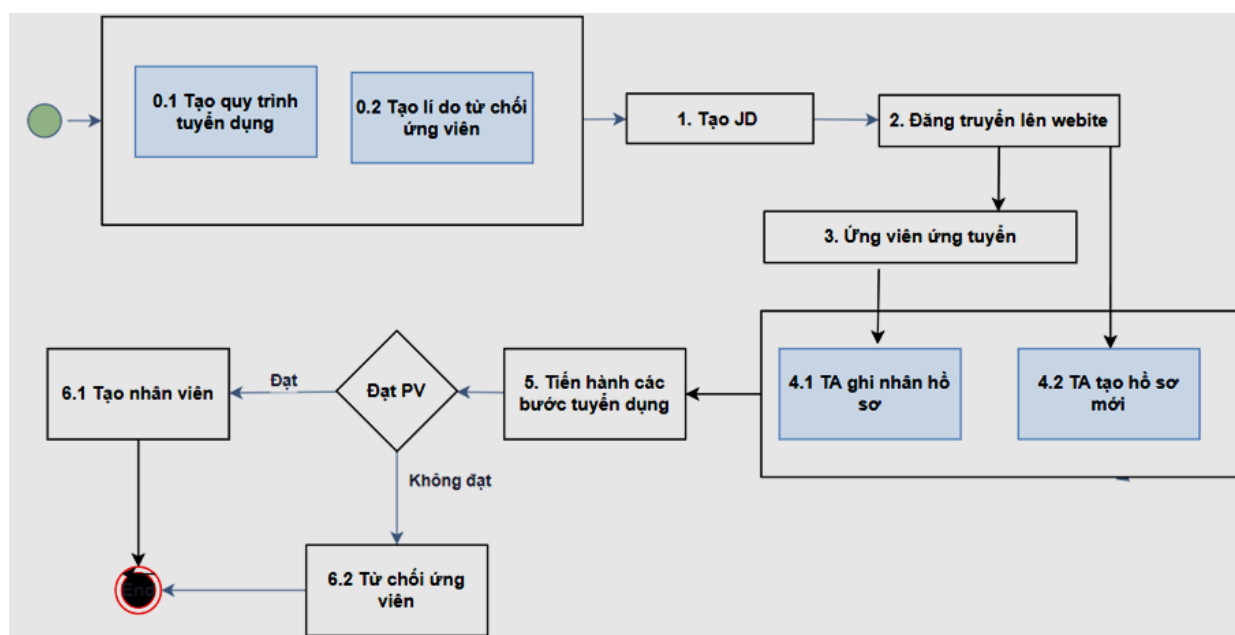
Và trong hệ thống quản lý nhân sự này có các đối tượng làm việc trên hệ thống như là ứng viên, nhân viên công ty, trưởng phòng, nhân viên tuyển dụng và nhân viên nhân sự. Mỗi đối tượng sẽ có các chức năng cụ thể được miêu tả chi tiết ở bảng 3.2.1.

Bảng 3.2.1 Các chức năng trong hệ thống

Đối tượng	Chức năng	Mô tả
Ứng viên	Xem mô tả công việc	Xem các bản mô tả công việc phù hợp cho vị trí cần ứng tuyển
	Nộp đơn ứng tuyển	Cho phép ứng viên nộp hồ sơ cá nhân vào hệ thống để tham gia vào quá trình tuyển dụng
Nhân viên công ty	Xem và đăng ký ngày nghỉ	Cho phép nhân viên trong công ty xem số lượng và đăng ký được ngày nghỉ cho bản thân.
Trưởng phòng	Duyệt ngày nghỉ	Duyệt các yêu cầu đăng ký ngày nghỉ
Nhân viên tuyển dụng	Đăng tuyển JD	Cho phép nhân viên tuyển dụng đăng những vị trí cần tuyển lên website tuyển dụng
	Tạo ứng viên	Nhân viên tạo ứng viên mới
	Quản lý ứng viên	Thay đổi các thông tin, trạng thái của ứng viên
	Lên lịch phỏng vấn	Lên lịch phỏng vấn ứng viên để theo dõi
	Lưu trữ ứng viên	Các ứng viên không phù hợp được lưu trữ vừa để chấm dứt quy trình tuyển dụng vừa để tiến hành theo dõi và phân tích trong các giai đoạn sau
	Tạo ứng viên phù hợp	Tạo nhân viên mới cho các ứng viên phù hợp
	Xem báo cáo tuyển dụng	Xem các báo cáo tuyển dụng theo các tiêu chí cụ thể
Nhân viên nhân sự	Tạo nhân viên	Tạo nhân viên mới vào hệ thống
	Quản lý hồ sơ nhân viên	Quản lý các thông tin nhân viên trong hệ thống
	Tạo hợp đồng	Tạo hợp đồng mới cho nhân viên

Quản lý hồ sơ hợp đồng	Theo dõi thông tin hợp đồng của nhân viên
Tạo ngày nghỉ chung	Tạo các ngày nghỉ chung của công ty (lễ, tết...) để ngày nghỉ được tính vào lương của toàn thể nhân viên
Tạo phiếu lương	Tạo các phiếu lương tự động được liên kết các thông tin từ hợp đồng lao động và số lượng ngày nghỉ
Tạo bảng lương	Tạo bảng lương cho công ty theo từng giai đoạn

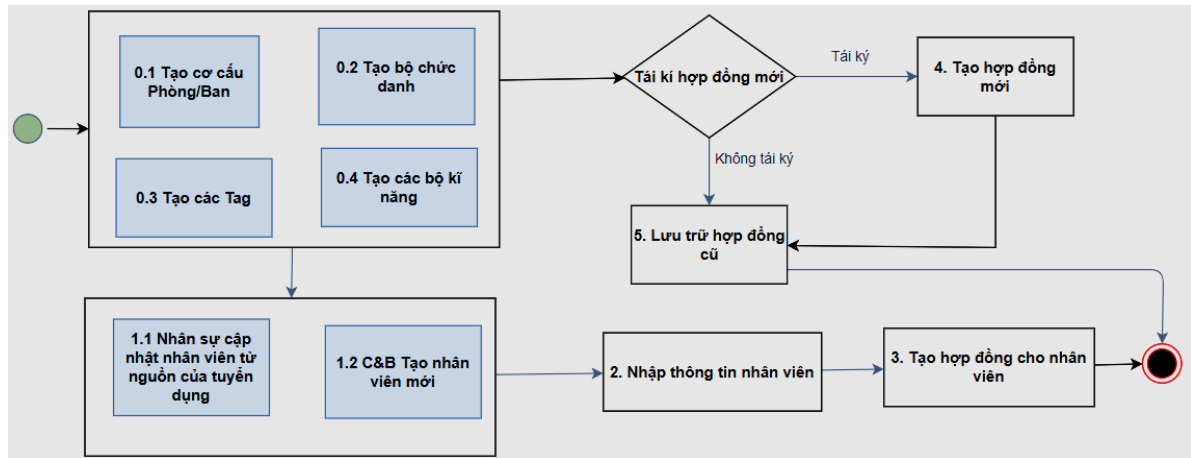
3.2.1. Chức năng quản lý tuyển dụng



Hình 8. Sơ đồ quy trình quản lý tuyển dụng

Ban đầu trước khi sử dụng hệ thống, chúng tôi cần thiết lập các quy trình tuyển dụng trên hệ thống, xây dựng các bộ lý do từ chối ứng viên để có thể sử dụng hệ thống đáp ứng được các quy trình hiện có của công ty. Tiếp đến xuất phát từ nhu cầu tuyển dụng của công ty, bộ phận tuyển dụng tạo các bản mô tả công việc (JD) phù hợp với các vị trí cần ứng tuyển và tiến hành thiết kế các bản mô tả công việc lên website tuyển dụng để cho các ứng viên bên ngoài công ty có thể xem và nộp đơn ứng tuyển. Sau khi có hồ sơ từ việc ứng viên nộp đơn ứng tuyển (hoặc bộ phận tuyển dụng có thể tạo trực tiếp), tiến hành các bước tuyển dụng theo các nghiệp vụ của công ty. Trong trường hợp ứng viên thỏa các yêu cầu của công ty thì tạo nhân viên cho ứng viên đó, nếu như không thỏa mãn yêu cầu thì lưu trữ ứng viên để phân tích và cải thiện quy trình tuyển dụng của công ty.

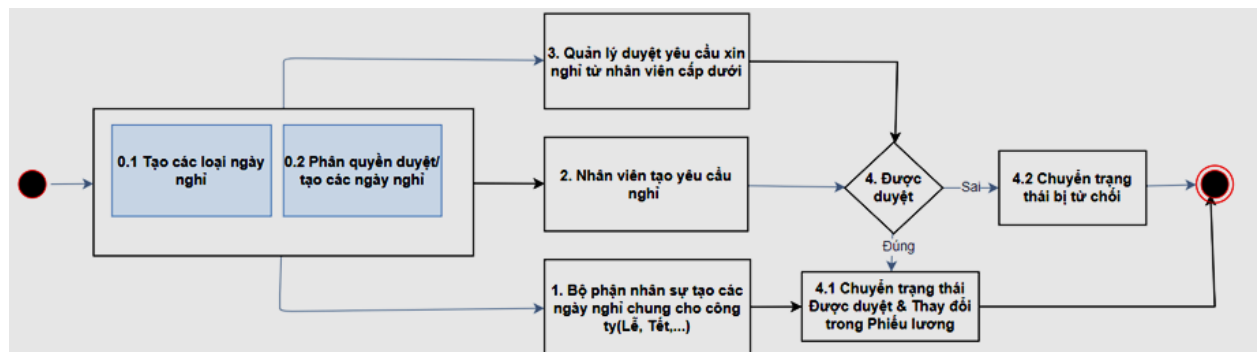
3.2.2. Chức năng quản lý nhân viên



Hình 9. Sơ đồ quy trình quản lý nhân viên

Trước khi có thể sử dụng được phân hệ nhân viên, chúng tôi phải vào định nghĩa hệ thống đáp ứng được hiện trạng thực tế của công ty đang sử dụng như: các phòng ban, vị trí công việc... Sau đó các nhân viên mới được tạo thông qua hai cách là tuyển từ ứng viên đạt yêu cầu hoặc nhân viên nhân sự vào tạo trực tiếp trên phân hệ nhân viên và nhập đầy đủ thông tin nhân viên, tạo hợp đồng theo mức thỏa thuận giữa hai bên. Các thông tin liên quan đến lương và phụ cấp nằm trong hợp đồng sẽ được dùng để tính tự động trong bảng lương.

3.2.3. Chức năng quản lý ngày nghỉ

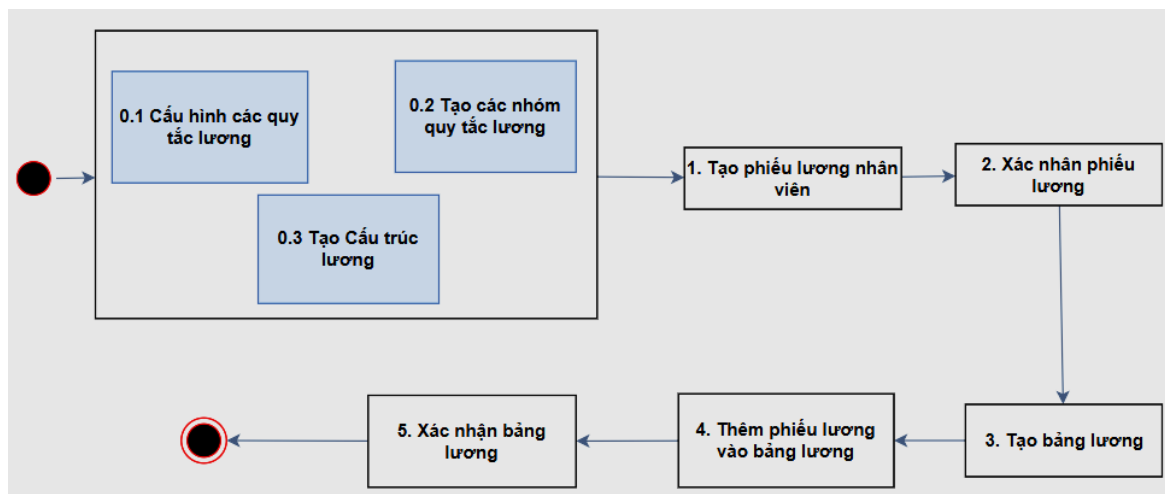


Hình 10. Sơ đồ quy trình quản lý ngày nghỉ

Để có thể thực hiện được chức năng quản lý ngày nghỉ, yêu cầu đặt ra chúng tôi phải cài đặt, định nghĩa các loại ngày nghỉ phù hợp, đáp ứng được các loại ngày nghỉ của công ty như: tạo các loại nghỉ phép, thiết lập trạng thái duyệt của ngày nghỉ đó, phân bổ cấp phát cho loại ngày nghỉ... Sau đó khi có nhu cầu muốn đăng ký ngày nghỉ, nhân viên lên trực tiếp hệ thống của nhân sự với tài khoản đã được cấp, tiến hành đăng ký ngày nghỉ với đầy đủ thông tin như loại nghỉ, thời gian, lý do. Tiếp đến, người quản lý trực tiếp của nhân viên đăng nhập vào hệ thống để duyệt thông tin nghỉ của nhân viên cấp dưới, những thông tin đăng ký và duyệt ngày nghỉ được lưu trữ trong hệ thống để tính toán thời gian nghỉ và thông tin lương trong phiếu lương.

Đến kỳ nghỉ chung của công ty (lễ, tết, ...), nhân viên bộ phận nhân sự sẽ thực hiện tạo ngày nghỉ chung cho toàn công ty trên hệ thống. Những ngày nghỉ được tạo cho toàn công ty sẽ được thể hiện lên giao diện của từng nhân viên, giúp cho nhân viên nhận biết được ngày nghỉ của mình, thông tin này cũng được cập nhật lên phiếu lương của từng nhân viên.

3.2.4. Chức năng quản lý bảng lương



Hình 11. Sơ đồ quy trình quản lý bảng lương

Để có thể tính toán được lương của nhân viên phù hợp theo Luật Lao động tại Việt Nam và các bộ luật liên quan, chúng tôi cần cấu hình từ ban đầu các quy tắc lương như: bảo hiểm xã hội, bảo hiểm y tế, bảo hiểm tai nạn, thuế thu nhập cá nhân, giảm trừ người phụ thuộc, tăng ca, lương trước thuế, lương sau thuế,... từ các quy tắc lương đó chúng tôi tiến hành gắn chúng với thông tin lương trong hợp đồng lao động và số lượng ngày nghỉ của nhân viên để có thể tạo các phiếu lương, bảng lương một cách tự động. Ở hình 3.8 minh họa cách định nghĩa cách tính thuế thu nhập cá nhân trên hệ thống. Sau đó cứ đến mỗi kỳ lương, nhân viên nhân sự sẽ vào hệ thống tiến hành tạo phiếu lương cho nhân viên và bảng lương cho công ty. Chức năng này được sử dụng tự động căn cứ vào mức lương của hợp đồng lao động, số lượng ngày nghỉ và theo chính sách tính lương tại Việt Nam.

```
Tính toán  
Loại tiên  
Mã Python  
Python Code  
luong_gross = categories.BASIC + categories.ALW + categories.OT  
nguo_i_phu_thuoc = employee.children  
  
if contract.wage >=29800000:  
    bhxh= 29800000 * 0.08  
else:  
    bhxh= contract.wage * 0.08  
  
if contract.wage >=29800000:  
    bhyt= 29800000 * 0.015  
else:  
    bhyt = contract.wage * 0.015  
  
if contract.wage >=85172000:  
    bhtn = 85172000 * 0.01  
else:  
    bhtn = contract.wage * 0.01  
  
bh=bhxh+bhyt+bhtn  
tinh_thue = luong_gross-bh-11000000-(nguo_i_phu_thuoc*4400000)  
if 0<tinh_thue<=5000000:  
    result = tinh_thue*0.05  
elif 5000000<tinh_thue<=10000000:  
    result = tinh_thue*0.1-250000  
elif 10000000<tinh_thue<=18000000:  
    result = tinh_thue*0.15-750000  
elif 18000000<tinh_thue<=32000000:  
    result = tinh_thue*0.2-1650000  
elif 32000000<tinh_thue<=52000000:  
    result =tinh_thue*0.25-3250000  
elif 52000000<tinh_thue<=80000000:  
    result = tinh_thue*0.3-5850000  
elif tinh_thue>80000000:  
    result = tinh_thue*0.35-9850000  
else: result=0
```

Hình 12. Định nghĩa cách tính thuế thu nhập cá nhân trên hệ thống

3.2.5. Hiện thực hệ thống

Nộp đơn ứng tuyển

MỤC ĐÍCH CÔNG VIỆC

Vị trí công việc này sẽ đóng góp gì vào Mục Tiêu & Chiến lược chung của tổ chức và phòng ban?
Điều gì đo lường mức độ thành công của vị trí công việc này tại tổ chức?

Vị trí này sẽ thực hiện việc kiểm tra nghiệp vụ của kế toán, thực hiện báo cáo tài chính, báo cáo thuế định kỳ.

Tham mưu về chuyên môn, hướng dẫn nghiệp vụ kế toán liên quan.

KIẾN THỨC

- Tốt nghiệp đại học/cao đẳng chuyên ngành kế toán, kiểm toán, tài chính tại các trường chuyên về kinh tế, tài chính kế toán.

KỸ NĂNG

- Sử dụng thành thạo các phần mềm kế toán Misa, MS office
- Kỹ năng quản lý, sắp xếp công việc.
- Kỹ năng phân tích, tổng hợp tốt.
- Khả năng lập kế hoạch, làm việc nhóm, làm việc dưới môi trường áp lực
- Có kỹ năng làm việc với số liệu, tổng hợp và phân tích báo cáo.

THÁI ĐỘ

- Khát khao và nhiệt huyết được cống hiến & thể hiện năng lực bản thân.
- Là người cẩn thận, trung thực, quyết đoán.

KINH NGHIỆM

- Có ít nhất 2 năm kinh nghiệm làm kế toán kho trong lĩnh vực thương mại nông sản có trên 200 mã hàng.
- Tối thiểu 1 năm kinh nghiệm vị trí trợ lý giám đốc, trợ lý kế hoạch
- Am hiểu về chuẩn mực kế toán, các qui định của pháp luật, thuế trong lĩnh vực kế toán.

Ứng tuyển ngay

Hình 13. Giao diện mô tả công việc chi tiết

Chức năng nộp đơn ứng tuyển là chức năng dành cho những ứng viên tìm kiếm công việc thông qua website tuyển dụng của công ty. Chức năng này giúp cho ứng viên có thể nộp sơ yếu lý lịch và các thông tin cơ bản đến bộ phận tuyển dụng, từ đó bộ phận tuyển dụng ghi nhận được hồ sơ và tiếp tục công việc tuyển dụng của mình. Và để có thể tiếp nhận hồ sơ từ bên ngoài, bộ phận tuyển dụng phải thiết kế giao diện cho các vị trí công việc cụ thể, biểu mẫu dùng cho ứng viên nhập thông tin và sau đó đăng tuyển lên website như các hình 3.9, 3.10.

Đơn ứng tuyển
Giám Sát Kho Và Chế Biến

Your Name * Nguyễn Văn Tiến

Your Email * vantien1234@gmail.com

Your Phone Number * 0983337262

Short Introduction 3 năm kinh nghiệm

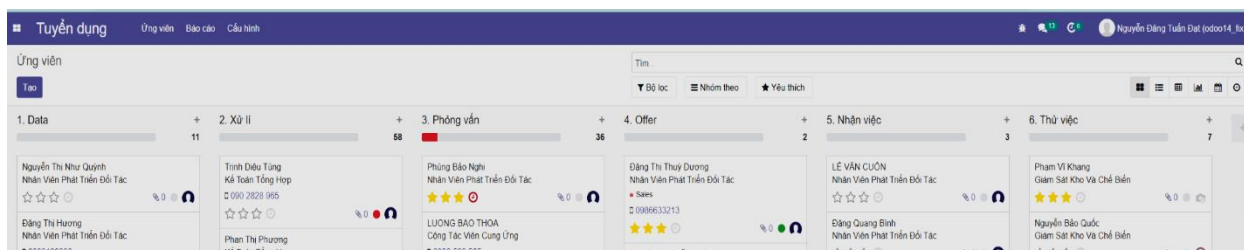
Resume Choose File Cv_Nguyenvantien.pdf

Submit

Hình 14. Giao diện form ứng tuyển

Quản lý tuyển dụng

Chức năng này dành cho nhóm nhân viên tuyển dụng của công ty phục vụ các chức năng cơ bản về tuyển dụng như: sàng lọc thông tin, lên lịch phỏng vấn, theo dõi quy trình quản lý tuyển dụng, tạo nhân viên cho ứng viên phù hợp, từ chối các ứng viên không phù hợp, hỗ trợ xem các báo cáo liên quan đến quá trình tuyển dụng từ đó đưa ra các phân tích, cải tiến quy trình tuyển dụng được diễn ra tốt hơn. Hình 3.11 là hình ảnh các giai đoạn của tuyển dụng đang được thực hiện.



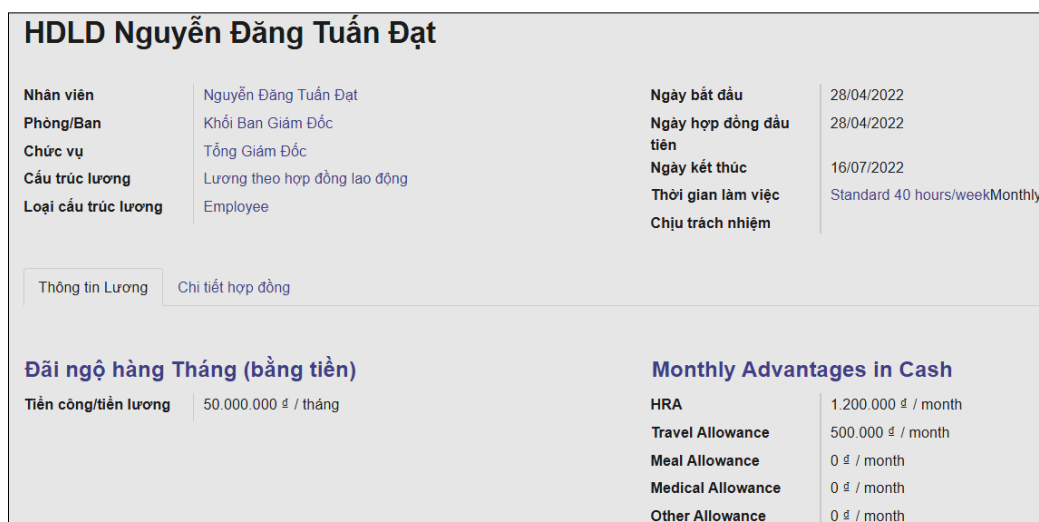
Hình 15. Giao diện quản lý tuyển dụng

Quản lý thông tin nhân viên, hợp đồng

Chức năng này phục vụ chủ yếu cho nhóm nhân viên nhân sự, nhằm giải quyết các vấn đề: theo dõi các thông tin của nhân viên, quản lý các thông tin về hợp đồng, hỗ trợ việc cảnh báo nhân viên hợp đồng sắp hết hạn hay trùng lặp hợp đồng ... Hình 14 và hình 5 là hình minh họa thông tin của nhân viên và chi tiết hợp đồng.



Hình 16 Giao diện quản lý thông tin nhân viên

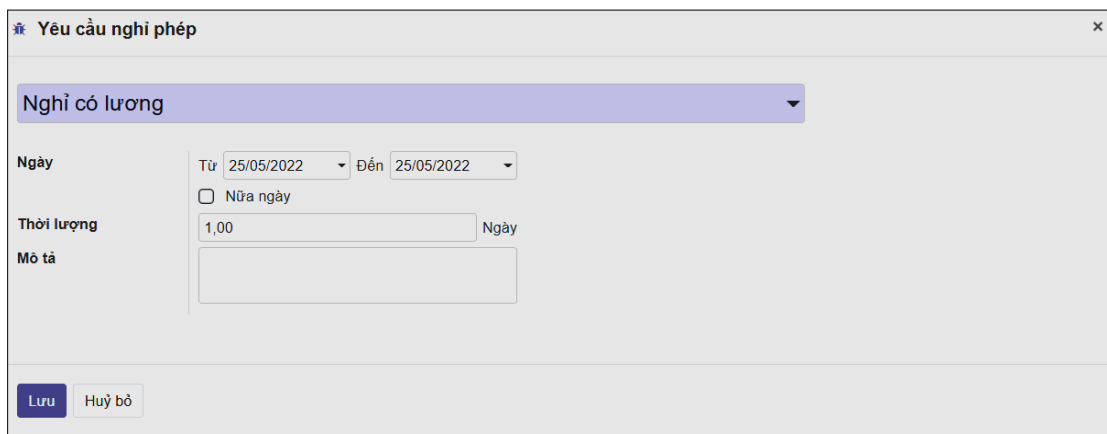


Hình 17. Giao diện quản lý thông tin hợp đồng

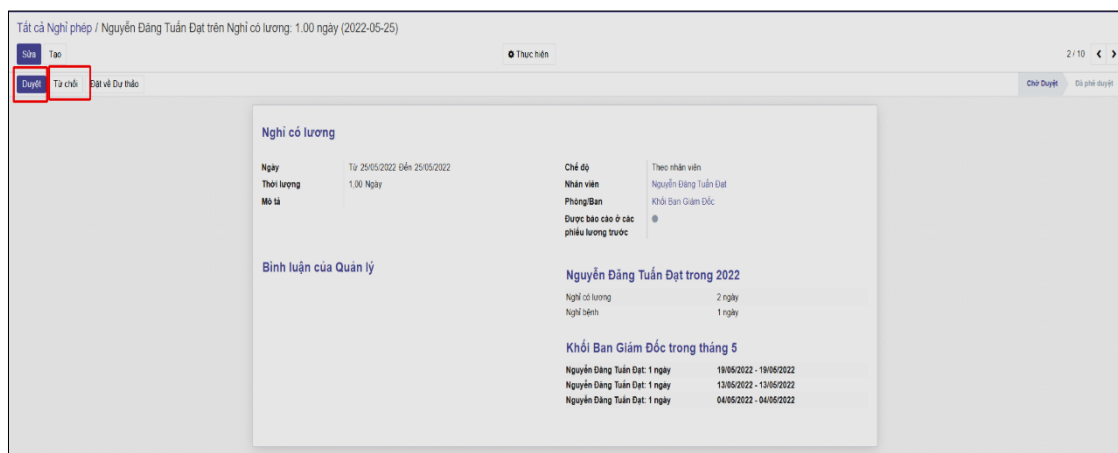
Quản lý ngày nghỉ nhân viên

Chức năng này đáp ứng cho toàn thể nhân viên hiện có của công ty, hệ thống cung cấp cho nhân viên nơi để đăng ký ngày nghỉ và trưởng phòng (quản lý cấp trên) vào duyệt yêu cầu xin nghỉ, trong trường hợp yêu

cầu nghỉ được duyệt, các thông tin này sẽ được liên kết đến bảng lương để tính toán. Nhân viên nhân sự vào hệ thống để tạo các ngày nghỉ chung của công ty (lễ, tết ...) để hệ thống ghi nhận và tính lương một cách chính xác. Xem chi tiết các đơn đăng ký nghỉ và duyệt nghỉ tại các hình 3.14, hình 3.15.



Hình 18. Giao diện đăng ký ngày nghỉ nhân viên



Hình 19. Giao diện duyệt yêu cầu nghỉ

Quản lý bảng lương

Chức năng được thực hiện để hỗ trợ cho nhân viên nhân sự trong việc tính lương cho nhân viên. Căn cứ vào mức lương, phụ cấp trong hợp đồng và số lượng ngày nghỉ của nhân viên (đã đề cập ở trên) mà hệ thống có thể tạo phiếu lương nhân viên và bảng lương cho công ty một cách tự động không yêu cầu nhân viên nhân sự thao tác tính toán nào trong việc quản lý bảng lương này.

16 Phiếu lương							
Nguyễn Đăng Tuấn Đạt							
Chu kỳ	01/05/2022 - 31/05/2022			Hợp đồng	HDLD Nguyễn Đăng Tuấn Đạt		
Tham chiếu	SLIP/065			Cấu trúc lương	Lương theo hợp đồng lao động		
Tên phiếu lương	Phiếu lương của Nguyễn Đăng Tuấn Đạt cho tháng 5-2022			Giấy bảo có	<input type="checkbox"/>		
Ngày làm việc & Dữ liệu nhập ngoài Tính toán Lương Chi tiết theo nhóm quy tắc lương Thông tin kế toán							
Tên	Mã	Danh mục	Số lượng	Tỷ lệ (%)	Quy tắc	Tổng tiền	Tổng
Lương Cơ bản	BASIC	Cơ sở	1,00	100,00	Lương Cơ bản	50.000.000,00	50.000.000,00
House Rent Allowance	HRA	Phụ cấp	1,00	100,00	House Rent Allowance	1.200.000,00	1.200.000,00
Travel Allowance	Travel	Phụ cấp	1,00	100,00	Travel Allowance	500.000,00	500.000,00
Meal Allowance	Meal	Phụ cấp	1,00	100,00	Meal Allowance	0,00	0,00
Medical Allowance	Medical	Phụ cấp	1,00	100,00	Medical Allowance	0,00	0,00
Other Allowance	Other	Phụ cấp	1,00	100,00	Other Allowance	0,00	0,00
Tăng ca ngày thường ca ngày	OTO	OT	1,00	100,00	Tăng ca ngày thường ca ngày	0,00	0,00
Tăng ca ngày nghỉ ca ngày	OTW	OT	1,00	100,00	Tăng ca ngày nghỉ ca ngày	0,00	0,00
Tăng ca ngày lễ ca ngày	OTS	OT	1,00	100,00	Tăng ca ngày lễ ca ngày	0,00	0,00
Bảo hiểm xã hội	BHXH	SHUI	1,00	100,00	Bảo hiểm xã hội	2.384.000,00	2.384.000,00
Bảo hiểm y tế	BHYT	SHUI	1,00	100,00	Bảo hiểm y tế	447.000,00	447.000,00
Bảo hiểm tai nạn	BHTN	SHUI	1,00	100,00	Bảo hiểm tai nạn	500.000,00	500.000,00
Giảm trừ người phụ thuộc	DPD	Dependant	1,00	100,00	Giảm trừ người phụ thuộc	8.800.000,00	8.800.000,00
Thuế thu nhập cá nhân	TNCN	Tax	1,00	100,00	Thuế thu nhập cá nhân	4.063.800,00	4.063.800,00
Tổng	GROSS	Tổng	1,00	100,00	Tổng	51.700.000,00	51.700.000,00
Lương Thực Lĩnh	NET	Lợi nhuận ròng	1,00	100,00	Lương Thực Lĩnh	44.305.200,00	44.305.200,00

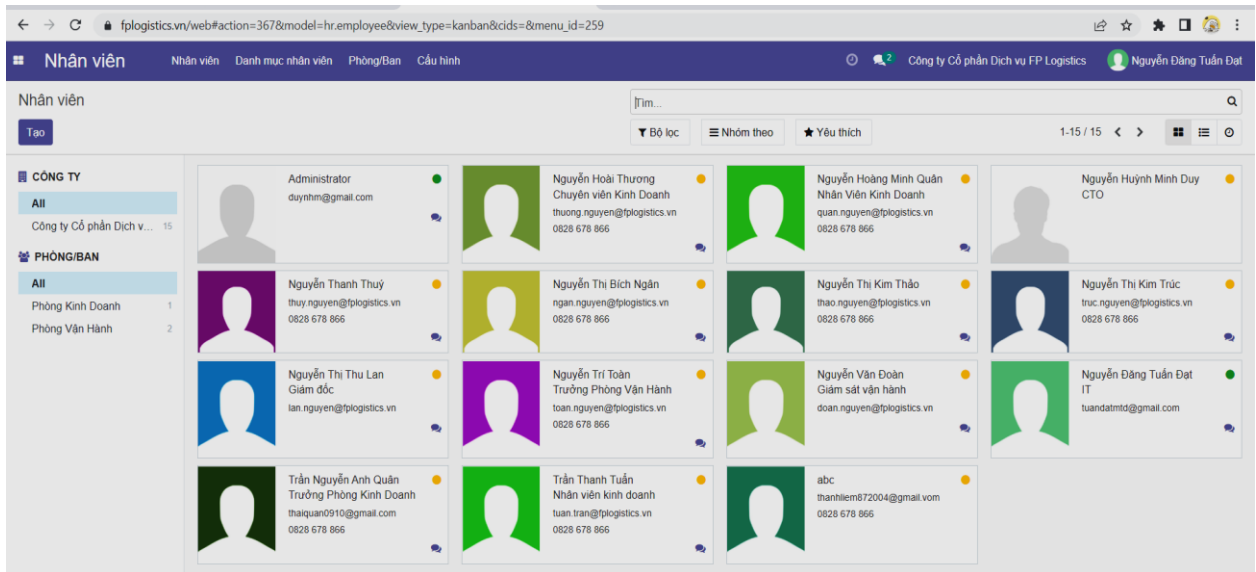
Hình 20. Phiếu lương chi tiết của nhân viên

Bảng lương tháng 5-2022					
Chu kỳ	01/05/2022 - 31/05/2022			Giấy bảo có	<input type="checkbox"/>
Phiếu lương					
Tham chiếu	Người lao động	Tên phiếu lương	Từ ngày	Đến ngày	Trạng thái
SLIP/067	Trần Ngọc Minh Thông	Phiếu lương của Trần Ngọc Minh Thông cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/066	Phạm Vĩ Khang	Phiếu lương của Phạm Vĩ Khang cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/065	Nguyễn Đăng Tuấn Đạt	Phiếu lương của Nguyễn Đăng Tuấn Đạt cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/064	Nguyễn Văn Chuẩn	Phiếu lương của Nguyễn Văn Chuẩn cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/063	Nguyễn Trường Thành	Phiếu lương của Nguyễn Trường Thành cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/062	Nguyễn Bảo Quốc	Phiếu lương của Nguyễn Bảo Quốc cho tháng 5-2022	01/05/2022	31/05/2022	Draft
SLIP/061	Bùi Thị Tuyết Dung	Phiếu lương của Bùi Thị Tuyết Dung cho tháng 5-2022	01/05/2022	31/05/2022	Draft

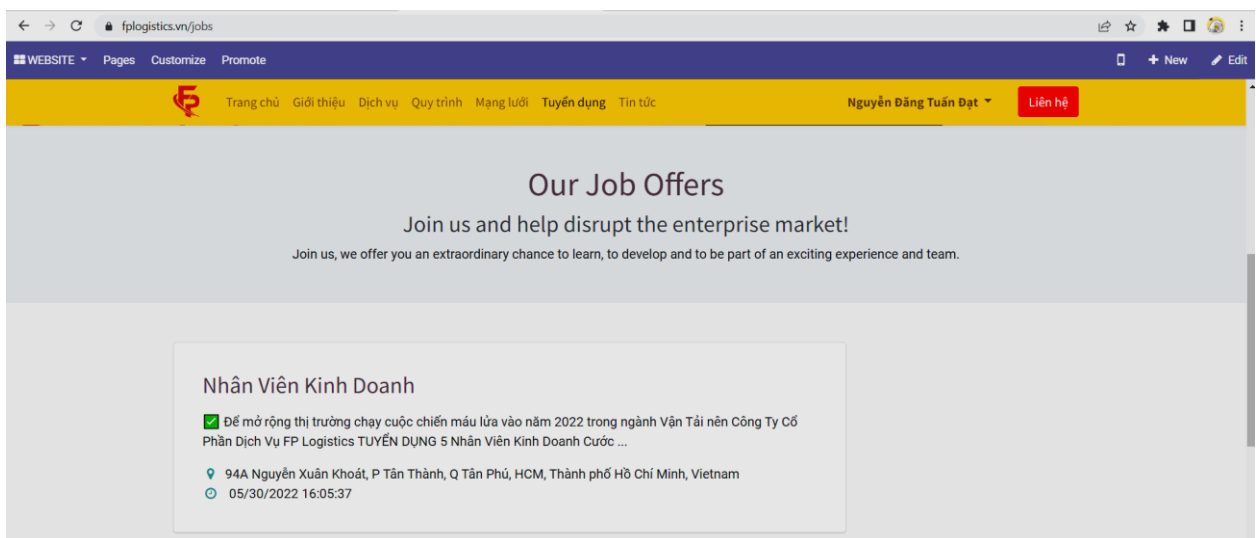
Hình 21. Bảng lương chi tiết cho công ty

4. KIỂM THỬ VÀ ỨNG DỤNG THỰC TẾ

Với sản phẩm hệ thống quản lý nhân sự lần này, chúng tôi đã triển khai cho một số công ty vừa và nhỏ để tối ưu hóa hiệu quả quản lý nhân sự, cụ thể là: công ty cổ phần dịch vụ FP Logistics, công ty TNHH ItapHoa, ... được minh họa ở hình 4.1 và hình 4.2.



Hình 29. Giao diện quản lý nhân viên của công ty FP Logistics



Hình 30. Giao diện website tuyển dụng công ty FP Logistics

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Từ những phân tích và đánh giá hiện trạng quản lý nhân sự của công ty cổ phần Bao Bì Phương Nam, chúng tôi đã đề xuất giải pháp cải thiện quy trình quản lý nhân sự bằng hệ thống Odoo. Trong bài báo này, chúng tôi đã trình bày giải pháp bao gồm chức năng quản lý tuyển dụng, chức năng quản lý nhân viên, chức năng quản lý hợp đồng, chức năng quản lý ngày nghỉ, chức năng quản lý bảng lương nhằm hỗ trợ các nghiệp vụ cho các nhóm nhân viên: nhân viên công ty, nhân viên tuyển dụng, nhân viên nhân sự, trưởng phòng và ứng viên tại công ty cổ phần Bao Bì Phương Nam và một số công ty khác. Ưu điểm của sản phẩm lần này là đã thay thế được việc lưu trữ dữ liệu trên giấy tờ thành lưu trữ trên cơ sở dữ liệu, việc sử dụng mã nguồn mở có sẵn của Odoo giúp tiết kiệm được chi phí, dễ dàng tùy chỉnh phù hợp với từng công ty cụ thể tại Việt Nam. Chúng tôi đã xây dựng hệ thống quản lý nhân sự dựa trên phần mềm ERP mã nguồn mở Odoo phiên bản thứ 14 và ngôn ngữ Python. Hệ thống này cũng đã cài đặt và triển khai thử nghiệm đến người dùng của công ty cổ phần bao bì Phương Nam, trong tương lai chúng tôi sẽ nâng cấp và triển khai hệ thống này lên nền tảng điện toán đám mây, mở rộng quy mô hoạt động đến nhiều doanh nghiệp và tích hợp website tuyển

dụng vào các dịch vụ của các doanh nghiệp để cho việc quản lý nhân sự và kết quả kinh doanh được thực hiện một cách hiệu quả nhất.

TÀI LIỆU THAM KHẢO

- [1] "Vai trò quan trọng của quản lý nhân sự trong doanh nghiệp," [Online]. Available: <https://fososoft.com/vai-tro-cua-quan-ly-nhan-su-trong-doanh-nghiep/>. [Accessed 29 5 2022].
- [2] Concepts in Enterprise Resource Planning, 3rd Editions By Ellen Monk and Bret Wagner, Thomson Course Technology 2008.
- [3] "Python là gì? Tại sao nên chọn Python?," [Online]. Available: <https://quantrimang.com/python-la-gi-tai-sao-nen-chon-python-140518>. [Accessed 29 5 2022].
- [4] "ADVANTAGES OF POSTGRESQL," [Online]. Available: <https://www.cybertec-postgresql.com/en/postgresql-overview/advantages-of-postgresql/>. [Accessed 30 5 2022].
- [5] "Tại sao doanh nghiệp cần sử dụng phần mềm ERP?," [Online]. Available: <http://phanmemungdungvn.com/tai-sao-doanh-nghiep-can-su-dung-phan-mem-erp/>. [Accessed 26 5 2022].
- [6] "Triển khai ERP: Các bước thực hiện và chi phí thực tế," [Online]. Available: <https://magenest.com/vi/chi-phi-va-qy-trinh-trien-khai-erp/>. [Accessed 27 5 2022].
- [7] Jake VanderPlas. Python Data Science Handbook. O REILLY, 2016.
- [8] Igor Milovanović, Dimitry Foures, Giuseppe Vettigli. Python Data Visualization Cookbook, Second Edition. BIRMINGHAM – MUMBAI, 2015.

YSC4F.230

ĐỘI NGŨ SHIPPER VÀ SỰ CẦN THIẾT CỦA MỘT ỨNG DỤNG QUẢN LÝ- THEO DÕI TRONG THỜI ĐẠI CÁCH MẠNG 4.0

NGUYỄN VĂN THẮNG, NGUYỄN VĂN HÙNG
Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
nguyenvanthang@iuh.edu.vn, nvhhungnv@gmail.com

Tóm tắt. Hiện nay, khi thế giới đang dần sống trong thời đại 4.0 (công nghiệp 4.0), các cửa hàng ngày càng mọc nhiều lên. Từ đó dẫn đến người tiêu dùng có nhu cầu mua hàng qua mạng thông qua các trang mạng xã hội hoặc các ứng dụng đặt hàng trực tuyến. Để đáp ứng được nhu cầu đó các cửa hàng bán online hoặc các cá nhân có nhu cầu trao đổi hàng hóa trong phạm vi nhất định thì cần phải thông qua những nhân viên giao hàng hay còn được gọi là Shipper. Nhận thấy tầm quan trọng của vấn đề nói trên, chúng tôi nghiên cứu “xây dựng ứng dụng quản lý Shipper” nhằm tạo ra một hệ thống mạng lưới kết nối các Shipper với tất cả các cửa hàng hoặc các cá nhân có nhu cầu mua bán qua mạng với khả năng linh động hơn trong việc mua hàng và nhận hàng. Ứng dụng client chạy trên hệ điều hành Android sẽ do Shipper sử dụng. Ứng dụng sẽ gửi thông tin vị trí nhân viên cũng như các thông tin liên quan về cho server. Nhân viên có thể xem được thông tin thời gian giao hàng, nơi nhận, nơi giao, khoảng cách từ nơi giao đến nơi nhận, tiền thu, thông tin khách hàng,... Với việc tích hợp Google Maps api và realtime thời gian thực giữa Shipper, người quản lý có thể dễ dàng kiểm tra được các vị trí hiện tại của các Shipper và tình trạng online/offline theo thời gian thực và cũng như tình trạng của các đơn hàng của từng Shipper.

Từ khóa. Google Maps api, realtime, hệ thống mạng lưới kết nối.

SHIPPER TEAM AND THE NEED OF A MANAGEMENT- TRACKING APP IN THE 4.0 REVOLUTIONARY AGE

Abstract. Nowadays, when the world is gradually living in the 4.0 era (industry 4.0), more and more shops are springing up. That leads to consumers wanting to buy goods online through social networking sites or online ordering applications. In order to meet that demand, online stores or individuals wishing to exchange goods within a certain range need to go through delivery workers, also known as Shippers. Realizing the importance of the above problem, we researched "building a Shipper management application" to create a network system connecting Shippers with all stores or individuals who have a need to buy. sell online with greater flexibility in purchasing and receiving. The client application running on the Android operating system will be used by the Shipper. The application will send employee location information as well as related information to the server. Employees can view information about delivery time, place of receipt, place of delivery, distance from the place of delivery to the place of receipt, money collected, customer information, etc. With the integration of Google Maps api and real-time real-time between Shipper, the manager can easily check the current location of the Shippers and the online/offline status in real time and also the status of the orders of each shipper.

Keywords. Network connection system, Google Map api, Realtime.

1. GIỚI THIỆU

Đối với những người đang theo mảng kinh doanh và bán hàng qua internet thì ngoài việc quan tâm đến chất lượng sản phẩm thì khâu quan trọng cần để tâm đến đó chính là dịch vụ vận chuyển hàng hóa tới tay khách hàng.

Dịch vụ bán hàng online ngày càng phát triển đồng nghĩa với việc các nhà bán hàng phải tăng cường dịch vụ giao hàng tận nơi. Tuy vậy, hiện nay để tìm được một người giao hàng uy tín là điều không hề đơn giản hoặc có khi gặp nhiều trở ngại trong việc quản lý đội ngũ giao hàng này. Thuê một shipper, chuyển 1 gói hàng cho họ vận chuyển thì dễ. Nhưng số lượng shipper nhiều lên, hàng hóa cũng tăng và làm sao để việc quản lý shipper và quản lý bán hàng được thuận tiện, logic là vấn đề khó. Giao hàng không hiệu quả cũng ảnh hưởng không ít tới doanh thu của cửa hàng. Nếu bạn đang và sẽ là nhà bán hàng thì quản lý shipper là vấn đề bạn cần quan tâm nhất bây giờ.

Quản lý chưa bao giờ là một việc đơn giản, nhất là công việc quản lý giao hàng. Trên thực tế sẽ có khá nhiều những vấn đề phát sinh với đơn hàng trong “Chặng đường” đi từ người bán tới người mua. Việc mua bán hàng trực tuyến hiện nay quá sôi động và nhộn nhịp vì tính tiện dụng mà nó mang lại.

Nếu không có một phần mềm quản lý Shipper thì mọi người sẽ thường gặp một số tình trạng như sau:

Thứ nhất ‘Khách hàng từ chối không nhận đơn vì không hài lòng với shipper’, đây là một trong những sự cố thường xuyên xảy ra ở các shop bán hàng. Đã có rất nhiều trường hợp khi hàng vận chuyển đến tận nơi rồi mà khách hàng lại không nhận bởi thấy phí ship cao do shipper tự ý tăng giá hay là gọi điện thoại cho khách hàng không được thì shipper mang đơn về lại. Điều này gây ảnh hưởng nhiều đến uy tín và doanh thu của cửa hàng.

Thứ hai ‘Không nhận được thông tin và lộ trình vận chuyển đơn hàng của shipper’, với một nhân viên làm việc tại văn phòng thì việc quản lý rất dễ nhờ camera, máy chấm công và quản lý trực tiếp. Nhưng với nhân viên giao hàng thì khó. Đa số thời gian làm việc của họ là ở bên ngoài, nhiệm vụ của shipper là buổi sáng rời khỏi kho với một xe hàng hóa, chiều về hàng hóa được thay bằng tiền mặt và nộp lại. Nhà quản lý không biết được thời gian ở ngoài shipper làm gì, đi đâu và thái độ làm việc như thế nào. Thái độ cư xử với khách hàng là rất quan trọng nếu không kiểm soát chặt chẽ thì cửa hàng sẽ dễ bị ảnh hưởng uy tín, thương hiệu. Từ những khó khăn mà các nhà bán hàng thường gặp phải, thì một phần mềm quản lý Shipper ra đời là rất cần thiết trong thời đại mới, thời đại cách mạng 4.0.

2. CÁC CÔNG NGHỆ SỬ DỤNG

2.1. NodeJS

2.1.1. Khái niệm NodeJS

NodeJS là một nền tảng (platform) phía Server side được xây dựng, vận hành trên V8 JavaScript runtime của Chrome giúp xây dựng và phát triển các ứng dụng mạng một cách nhanh chóng và có khả năng mở rộng dễ dàng.

Node.js ra mắt vào năm 2009, bởi Ryan Dahl và nhận sự bảo trợ từ công ty Joyent (Hoa Kỳ). Phần core bên dưới của Node.js hầu hết được viết bằng C++. Chính vì thế mà tốc độ xử lý và hiệu năng khá cao. Bên cạnh đó, Node.js còn tạo ra ứng dụng có tốc độ xử lý nhanh với thời gian thực (real-time). Node.js cũng được áp dụng để tạo các sản phẩm có lượng truy cập lớn, mở rộng nhanh và đổi mới công nghệ... hoặc tạo dự án startup tiết kiệm thời gian nhất.

2.1.2. Đặc điểm của NodeJS

Nodejs có nhiều đặc điểm nổi bật, vượt trội. Nắm được các đặc điểm này sẽ giúp bạn hiểu sâu hơn về Nodejs, đồng thời đưa ra quyết định sử dụng nền tảng này chính xác.

- **Bất đồng bộ và phát sinh sự kiện (Non-blocking and Event Driven):** Tất cả các APIs của thư viện Node.js đều bất đồng bộ (non-blocking), NodeJS không cần đợi một API trả về dữ liệu. Server chuyển sang một API khác sau khi gọi nó và có cơ chế riêng để gửi thông báo và nhận phản hồi về các hoạt động của Node.js mà API đã gọi.
- **Tốc độ nhanh:** Phần core phía dưới được viết gần như toàn bộ bằng C++ kết hợp Chrome V8 Engine nên tốc độ xử lý công việc của Node.js cực nhanh, nhưng vẫn đảm bảo được tính chính xác.
- **Đơn giản – Hiệu năng cao:** Node.js sử dụng một mô hình luồng đơn luồng (single thread) và các sự kiện lặp (event-loop). Cơ chế sự kiện cho phép phía Server trả về phản hồi theo non-blocking, đồng thời tăng hiệu quả sử dụng. Các luồng đơn cung cấp dịch vụ cho nhiều request hơn hẳn Server truyền thống.
- **Không lưu bộ nhớ đệm (non buffer):** Nền tảng Node.js không có vùng nhớ đệm, tức không cung cấp khả năng lưu trữ dữ liệu buffer



2.2. Angular

2.2.1. Khái niệm Angular

Angular là một JavaScript framework dùng để viết giao diện web (Front-end). Đây là một sản phẩm được viết bởi Misko Hevery và một người bạn của anh là Adam Abrons. Sau đó, chính thức “ra mắt” vào 20/10/2010. Hiện tại, sản phẩm này đang được Google duy trì. Hiểu đơn giản, Angular là một khung làm việc của JavaScript MVC phía máy khách (client) nhằm phát triển ứng dụng web động.

AngularJS là từ dùng để nói về Angular 1 (ra đời năm 2009), được viết bằng JavaScript. Angular là từ gọi chung cho Angular 2 trở lên (ra đời năm 2016), được viết bằng TypeScript – phiên bản nâng cao của JavaScript.

Angular được thay đổi rất nhiều từ AngularJS. Angular đã thiết kế lại từ đầu nên có nhiều khái niệm đã thay đổi từ AngularJS. Kiến trúc của Angular và AngularJS hoàn toàn khác nhau. Hiện tại AngularJS cũng không còn được Google hỗ trợ nâng cấp nữa.

2.2.2. Những tính năng nổi bật của Angular

- Cơ chế Two-Way Data Binding

Đây là tính năng được developer đánh giá là ấn tượng nhất của Angular. Data binding tự động, tức thời. Bất cứ thay đổi nào trên view, dù là nhỏ nhất, cũng đều được tự động cập nhật thuộc tính “model” vào component class và ngược lại.

Thêm vào đó, Angular cũng hỗ trợ property binding. Developer có thể ràng buộc thuộc tính HTML với thuộc tính của component class, data sẽ tự động xuất hiện bên trong view thông qua việc điều khiển DOM.

- Hỗ trợ cơ chế Routing (điều hướng) mạnh mẽ

Angular có cơ chế routing tải trang một cách bất đồng bộ trên cùng một trang cho phép chúng ta tạo SPA. Đa số các ứng dụng Web không chỉ có 1 view hay một page duy nhất, mà sẽ cung cấp nhiều view khác nhau tương ứng với với các chức năng chính. Ví dụ: các trang liên hệ, giới thiệu chung, nội dung, báo giá,... trong một website. Hiện thị đúng view vào đúng thời điểm là mục đích của điều hướng (routing).

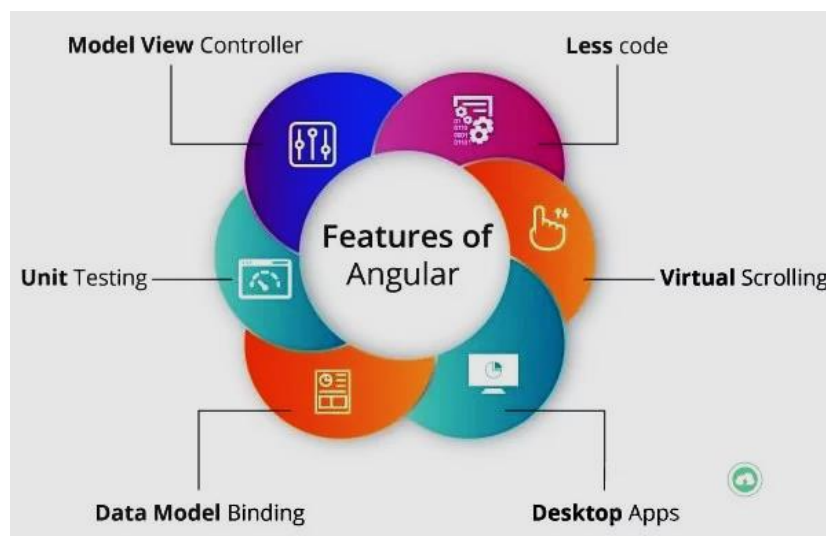
Angular giúp định nghĩa các đường dẫn (route) cho mỗi page view của ứng dụng. Developer sẽ kích hoạt route dựa trên tương tác của người dùng (user).

Mở rộng HTML: Nhờ Angular, developer có thể sử dụng cấu trúc lập trình giống như điều kiện IF, vòng lặp FOR, những biến địa phương “local variables”, ... để render các control.

Thiết kế module hoá: Angular tiếp cận theo hướng thiết kế module hoá. Người dùng phải tạo các Angular Module để tổ chức tốt hơn và quản lý source code.

Hỗ trợ làm việc với hệ thống Backend: Hỗ trợ làm việc với backend server, thực thi bất cứ logic nào và nhận dữ liệu về – Angular được xây dựng như thế!

Cộng đồng hỗ trợ rộng lớn: Nhiều tài liệu phong phú, đa dạng từ cơ bản đến nâng cao các API của Angular, cũng như có hẳn một Tutorial Basic được xây dựng nên bởi Angular team.
Mã nguồn mở và được Google hỗ trợ, liên tục cập nhật.



2.3. React Native

2.3.1. Khái niệm React Native

Được phát triển bởi Facebook, React Native là một framework hướng đến phát triển ứng dụng di động đa nền tảng. Với sự trợ giúp của React Native, lập trình viên (developer) có thể sử dụng JavaScript để tạo ra mobile apps (ứng dụng di động) hỗ trợ cho cả nền tảng Android và iOS. Instagram, Facebook, Skype,... là những ứng dụng nổi bật sử dụng React Native.



2.3.2. Ưu điểm

Có thể tái sử dụng code: React Native cho phép các developer có thể tái sử dụng code trong khi phát triển các ứng dụng đa nền tảng. Đặc biệt, developer có thể tái sử dụng hầu như 80-90% các đoạn code thay vì phải viết và tạo các ứng dụng riêng biệt cho các nền tảng khác nhau.

Ưu điểm này giúp người dùng:

- Tiết kiệm thời gian và giảm chi phí phát triển của một ứng dụng.

- Tận dụng nguồn nhân lực tốt hơn
- Duy trì ít code hơn, ít bugs hơn
- Các tính năng trong cả 2 platforms cũng tương tự nhau.

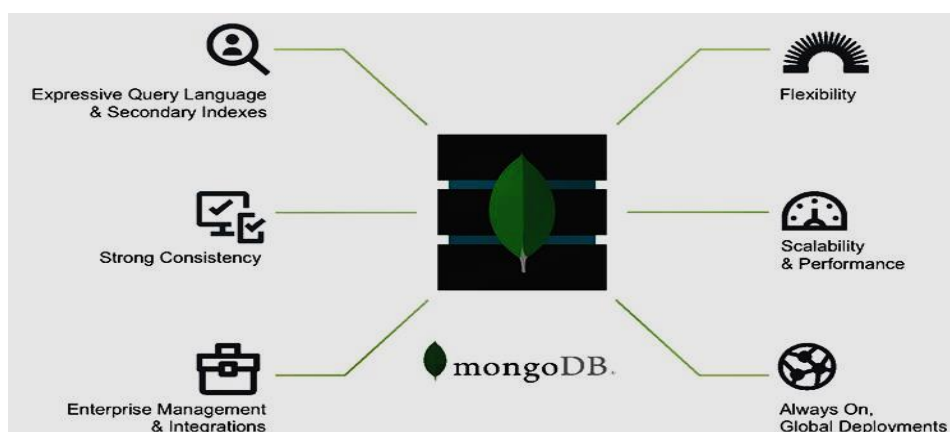
2.3.3. Nhược điểm

- Yêu cầu Native code.
- Hiệu năng kém hơn so với Native App.
- Bảo mật chưa thật sự tốt do dùng JavaScript. Do sử dụng JavaScript, người dùng cũng sẽ bị ảnh hưởng bởi những đặc điểm của JavaScript: dễ làm dễ sai, dẫn đến khó duy trì về sau.
- Quản lý bộ nhớ.
- Tùy biến chưa thật sự tốt ở một số module.
- Không thích hợp cho các app cần năng lực tính toán cao (hash, crypto, etc).

2.4. MongoDB (NO SQL)

NoSQL là 1 dạng CSDL mã nguồn mở và được viết tắt bởi: None-Relational SQL hay có nơi thường gọi là Not-Only SQL. NoSQL được phát triển trên Javascript Framework với kiểu dữ liệu là JSON và dạng dữ liệu theo kiểu key và value. NoSQL ra đời như là 1 mảnh vá cho những khuyết điểm và thiếu sót cũng như hạn chế của mô hình dữ liệu quan hệ RDBMS (Relational Database Management System - Hệ quản trị cơ sở dữ liệu quan hệ) về tốc độ, tính năng, khả năng mở rộng,... Với NoSQL bạn có thể mở rộng dữ liệu mà không lo tới những việc như tạo khóa ngoại, khóa chính, kiểm tra ràng buộc .v.v ... NoSQL bỏ qua tính toàn vẹn của dữ liệu và transaction để đổi lấy hiệu suất nhanh và khả năng mở rộng. NoSQL được sử dụng ở rất nhiều công ty, tập đoàn lớn, ví dụ như FaceBook sử dụng Cassandra do FaceBook phát triển, Google phát triển và sử dụng BigTable,...

Vậy MongoDB là một hệ quản trị cơ sở dữ liệu mã nguồn mở, là CSDL thuộc NoSql và được hàng triệu người sử dụng. MongoDB là một database hướng tài liệu (document), các dữ liệu được lưu trữ trong document kiểu JSON thay vì dạng bảng như CSDL quan hệ nên truy vấn sẽ rất nhanh. Với CSDL quan hệ chúng ta có khái niệm bảng, các cơ sở dữ liệu quan hệ (như MySQL hay SQL Server...) sử dụng các bảng để lưu dữ liệu thì với MongoDB chúng ta sẽ dùng khái niệm là collection thay vì bảng. So với RDBMS thì trong MongoDB collection ứng với table, còn document sẽ ứng với row, MongoDB sẽ dùng các document thay cho row trong RDBMS. Các collection trong MongoDB được cấu trúc rất linh hoạt, cho phép các dữ liệu lưu trữ không cần tuân theo một cấu trúc nhất định. Thông tin liên quan được lưu trữ cùng nhau để truy cập truy vấn nhanh thông qua ngôn ngữ truy vấn MongoDB



2.5. Google Map API

Ứng dụng Google Map là một dịch vụ bản đồ trực tuyến trên web và app miễn phí do Google phát hành và quản lý, ứng dụng hỗ trợ nhiều dịch vụ, tính năng cho người dùng nhưng phổ biến nhất là dò đường, chỉ đường, tìm vị trí; hiển thị những tuyến đường tối ưu cho từng phương tiện tham gia giao thông, ngoài ra còn hướng dẫn cách bắt xe và chuyển tuyến xe dành cho người tham gia các phương tiện lưu thông công

cộng (xe bus), hiển thị những địa điểm xung quanh vị trí người dùng hoặc vị trí chỉ định như ATM, trạm xăng, bệnh viện,...

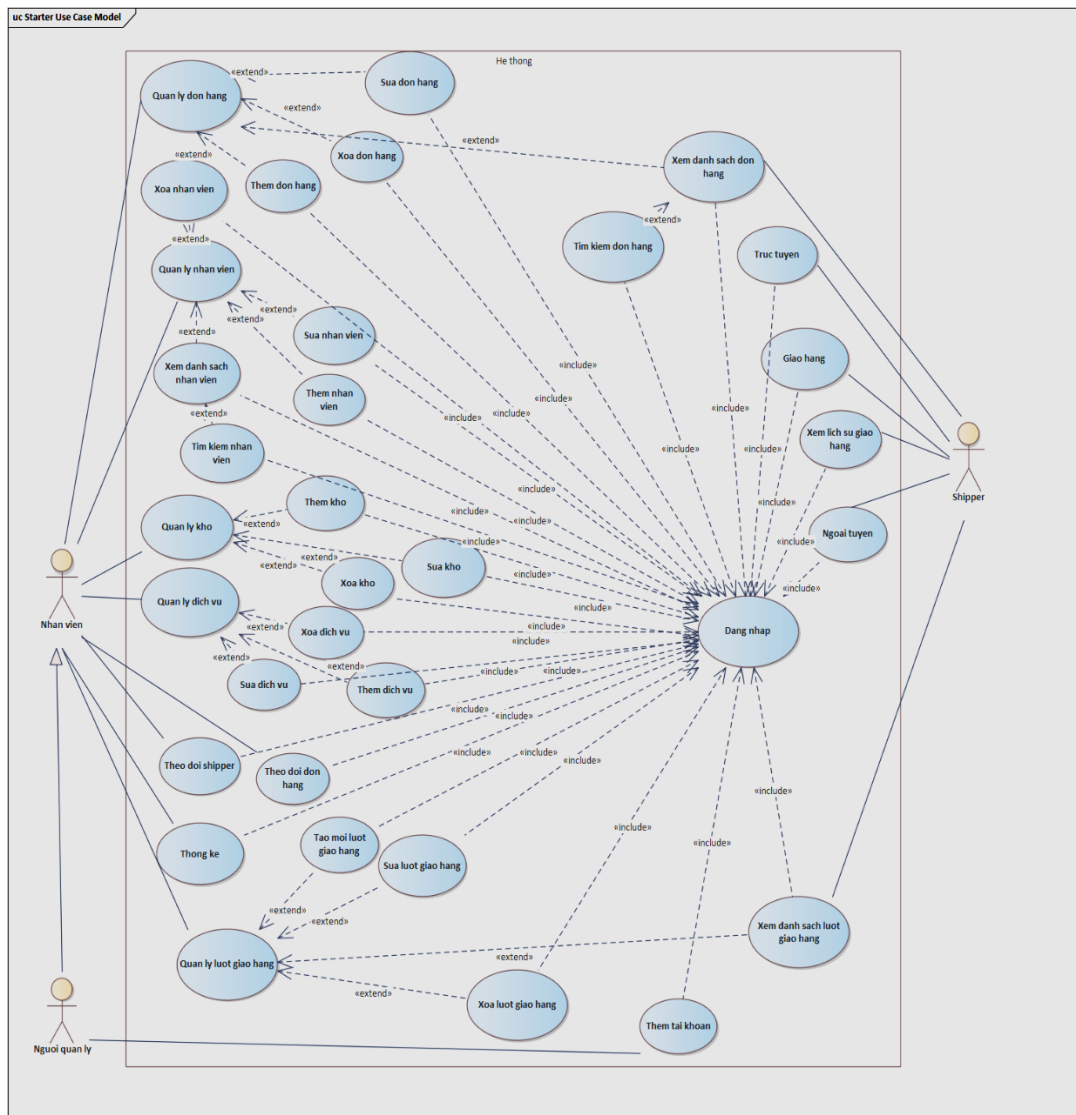
Vậy Google Map API là gì? Hiểu đơn giản đây là một phương pháp cho phép một website B có thể sử dụng dịch vụ hoặc hiển thị nội dung của một trang web khác, ở đây là website A – Google Map (thông qua Map API), dịch vụ bản đồ của website A (Map) sẽ được nhúng vào website B (Website cá nhân), tại trang web B có thể sử dụng những dịch vụ mà Google Map cung cấp thông qua Google Map API như: di chuyển, zoom, đánh dấu trên bản đồ,...

Hiện nay, các ứng dụng xây dựng trên nền tảng Google Maps như Grab thường sử dụng Google Map API để nhúng bản đồ vào trang web hoặc ứng dụng thông qua ngôn ngữ Javascripts, chính vì vậy mà việc sử dụng API từ Google cũng khá dễ dàng. Đồng thời Map API cũng đã được nâng cấp lên phiên bản mới, không chỉ hỗ trợ cho máy tính và website truyền thống mà còn cả thiết bị di động, giúp ứng dụng hoạt động nhanh hơn và hiệu quả hơn.

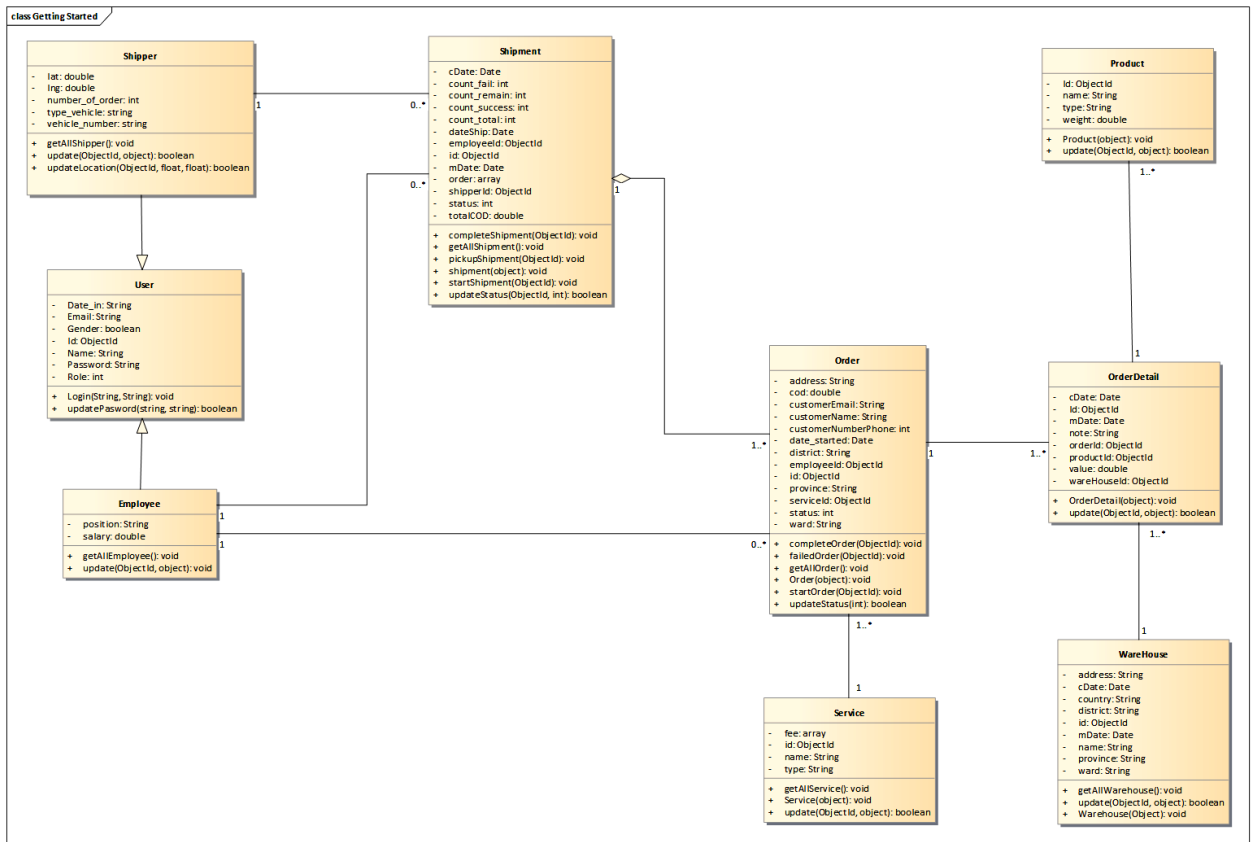
3. PHÂN TÍCH – TRIỂN KHAI “ỨNG DỤNG QUẢN LÝ SHIPPER”

3.1. Phân tích thiết kế hệ thống Ứng dụng quản lý shipper.

Sơ đồ Use case

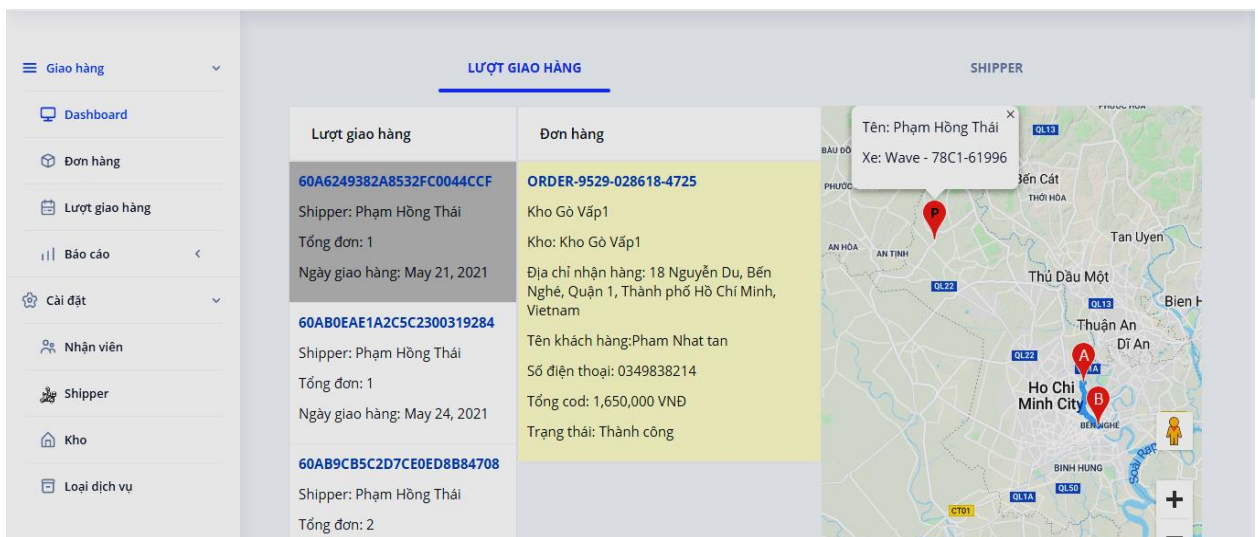


Class diagram

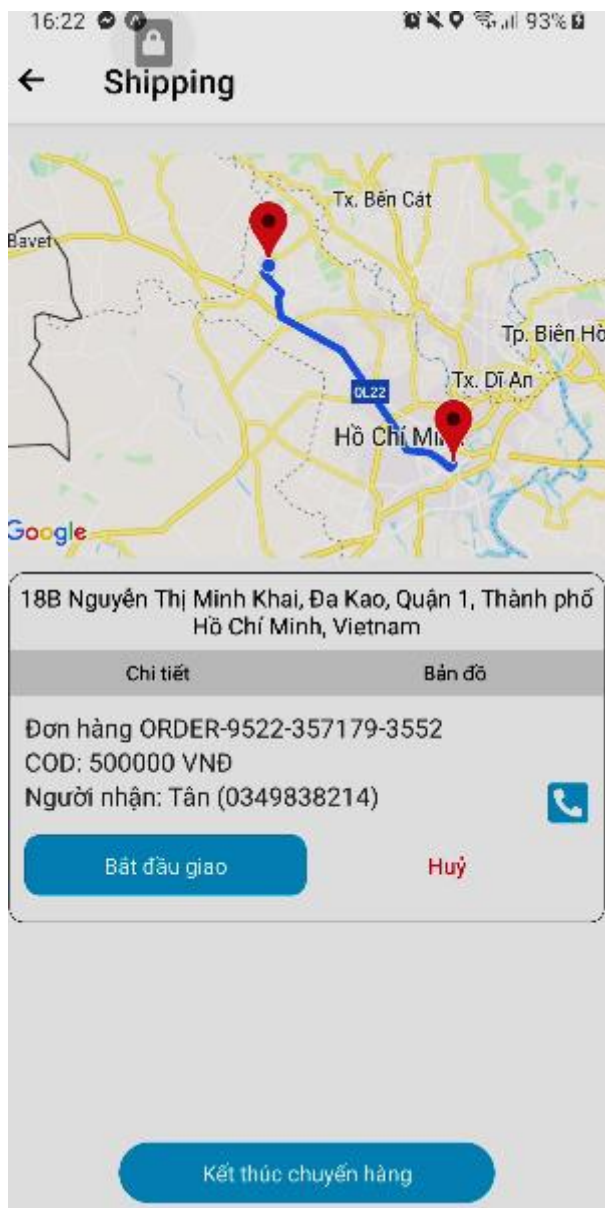


3.2. Giao diện đồ họa (Các màn hình đại diện)

Trang tổng quan – Theo dõi shipper, lượt giao hàng



Màn hình giao hàng trên app Shipper



4. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

4.1. Kết quả đạt được

- Xây dựng thành công website, app mobile quản lý shipper.
- Shipper có thể đăng nhập, thay đổi trạng thái làm việc (Online, Offline), giao hàng (hiển thị vị trí hiện tại, chỉ đường trên bản đồ), xem lịch sử giao hàng, xem thông tin cá nhân.
- Về phần quản lý: Nhân viên có thể quản lý các danh sách (Nhân viên, kho, shipper, loại dịch vụ, đơn hàng, lượt giao hàng), có thể thêm và xóa (nhân viên, kho, shipper, loại dịch vụ), cập nhật thông tin (nhân viên, kho, shipper, loại dịch vụ), có thể thêm và cập nhật (thay đổi shipper, ngày giao hàng) cho lượt giao hàng, có thể thêm và hủy đơn hàng.
- Ứng dụng đã tự động hóa tất cả quy trình: kết nối với nhiều đơn vị vận chuyển, cập nhật trạng thái đơn hàng, thông kê chi phí giao hàng từ các đơn hàng do shipper vận chuyển.

- Ứng dụng quản lý shipper giúp cho quản lý chặt chẽ từng khâu trong quy trình giao hàng giúp nhà bán hàng kiểm soát trạng thái đơn hàng dễ dàng. Từ đó có giải pháp đẩy nhanh tiến độ, mở rộng phạm vi, mở rộng đối tượng phục vụ.
- Nhân viên giao hàng được hỗ trợ tối đa nhờ hệ thống định vị GPS, giúp tìm ra địa điểm nhanh chóng, chính xác.
- Bảo mật thông tin khách hàng, thông tin đơn hàng và cả thông tin shipper.

4.2. Hướng phát triển cho ứng dụng:

- Phát triển thêm App mobile chạy trên nền tảng iOS.
- Phát triển thêm tính năng Web Service API để hỗ trợ kết nối đến các dữ liệu từ công ty khác.
- Ứng dụng sẽ có thêm các chức năng như: Cập nhật liên tục quãng đường đi của shipper, nhắn tin – gọi điện trao đổi giữa shipper và khách hàng.
- Ứng dụng phải được bổ sung thêm chức năng quản lý các hình ảnh trong quá trình giao hàng tới khách hàng.
- Tọa độ - Vị trí của shipper phải được cập nhật thường xuyên.

TÀI LIỆU THAM KHẢO

- [1] Những vướng mắc của chủ shop với shipper và Làm thế nào để quản lý shipper hiệu quả chính xác <https://www.sapo.vn/blog/khuc-mac-shipper-bai-toan-kho-da-co-loi-giai-dap>.
- [2] Những khó khăn khi quản lý nhân viên giao hàng và hướng giải quyết <https://www.webtretho.com/f/y-tuong-kinh-doanh-khoi-nghiep/nhung-kho-khan-khi-quan-ly-nhan-vien-giao-hang-va-huong-giai-quyet>.
- [3] Quản lý giao hàng 4.0, không sổ sách, không thất thoát, không chờ đợi ! <https://blog.abit.vn/quan-ly-giao-hang/>.
- [4] Dottie Schindlinger, Brian Stafford (2019), “Quản lý chuyên nghiệp”. NXB Lao động.
- [5] Nguyễn Văn Vị (2002), Phân tích và thiết kế hệ thống thông tin quản lý, NXB Thống Kê.

YSC4F.231

INTRODUCTION TO MULTI-MARGINAL OPTIMAL TRANSPORT ON SUB-RIEMANNIAN MANIFOLD

THANH SON TRINH

Faculty of Information Technology, Industrial University of Ho Chi Minh City

trinhthanhson@iuh.edu.vn

Abstract. In this paper, we introduce the multi-marginal optimal transport on sub-Riemannian manifold to minimize the total cost, which is defined to be the sum of the squared sub-Riemannian distances. As the works of Kim and Pass on Riemannian manifold [13]; and Pass, Pinamonti and Vedovato on Heisenberg group [16], we give a Kantorovich dual formula and prove the existence of a solution to our problem under structure of sub-Riemannian manifold.

Keywords. Multi-marginal, Kantorovich duality, Sub-Riemannian manifold.

1. INTRODUCTION

In 2010s, the theory of classical optimal transport problems was developed by many authors [4,5,6,7,9,19,20,21]. It has many applications in economics, image processing, PDEs, probability and statistics, and logarithmic Sobolev inequalities. For more details, readers can see [3,7,8,15,17,18,19]. The primal optimal transport problem was introduced by Kantorovich in 1940s [9,10]. This problem is presented in the form

$$\inf_{\gamma \in \Pi(\mu_1, \mu_2)} \int_{X_1 \times X_2} c(x_1, x_2) d\gamma(x_1, x_2),$$

where μ_1 and μ_2 are Borel probability measures on Polish (separable and complete) metric spaces X_1 and X_2 , respectively; $c: X_1 \times X_2 \rightarrow (-\infty, +\infty]$ is a cost function, and we denote by $\Pi(\mu_1, \mu_2)$ the set of all Borel probability measures γ on $X_1 \times X_2$ with the first and the second marginals are μ_1 and μ_2 , respectively, this means that

$$\gamma(A_1 \times X_2) = \mu_1(A_1), \gamma(X_1 \times A_2) = \mu_2(A_2),$$

for every Borel subsets A_i of X_i , $i = 1, 2$.

In 2015, Kim and Pass have generalized the primal optimal transport problem by introducing and investigating the multi-marginal optimal transport problem on Riemannian manifolds in paper [13]. Let μ_1, \dots, μ_k be k Borel probability measures on a manifold M , this problem is written in the following form

$$\inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k),$$

where $\Pi(\mu_1, \dots, \mu_k)$ is the set of all Borel probability measures γ on $M^k = M \times \dots \times M$ with marginals μ_1, \dots, μ_k , this means that

$$\gamma\left(M^{i-1} \times A \times M^{k-i}\right) = \mu_i(A),$$

for every Borel subset A of M and every $i = 1, \dots, k$; and C is a cost function defined by

$$c(x_1, \dots, x_k) = \inf_{y \in M} \sum_{i=1}^k d_R^2(y, x_i),$$

with d_R is the Riemannian metric.

After that, the multi-marginal optimal transport problem was considered more by Pass, Pinamonti and Vedovato in [16]. In this paper, they studied this problem on Heisenberg structure.

In this paper, we introduce multi-marginal optimal transport problem over structure of sub-Riemannian manifolds which generalize both problems in [13] and [16], since the sub-Riemannian structure cover both structures of Riemannian manifold and Heisenberg group. Given k Borel probability measures μ_1, \dots, μ_k on a sub-Riemannian manifold (M, Δ, g) which is defined in Section 2, our problem is formulated as

$$(M) \quad \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k),$$

with C is defined by

$$c(x_1, \dots, x_k) = \inf_{y \in M} \sum_{i=1}^k \lambda_i d_{SR}^2(x_i, y),$$

where d_{SR} is the sub-Riemannian distance and $\lambda_1, \dots, \lambda_k > 0$ such that $\sum_{i=1}^k \lambda_i = 1$.

Our first main result is the existence result for our problem.

Theorem 1.1 *If (M, d_{SR}) is a complete metric space then problem (M) has solutions.*

Next, we give the Kantorovich dual problem for our problem. The dual problem to (M) is defined by

$$(D) \quad \max \left\{ \sum_{i=1}^k \int_M f_i(x_i) d\mu_i(x_i) : f_i \in L^1(\mu_i), \sum_{i=1}^k f_i(x_i) \leq c(x_1, \dots, x_k) \right\}$$

In our second main result, we show a link between our problem and its dual problem. In the following result, a C -conjugate solution (f_1, \dots, f_k) to (D) will be defined in Section 3.

Theorem 1.2 *There exists a C -conjugate solution (f_1, \dots, f_k) to (D). Moreover, if γ is an optimal for problem (M) then we have*

$$\sum_{i=1}^k f_i(x_i) = c(x_1, \dots, x_k) \quad \gamma - a.e$$

Our second main result is a generalized version of [15, Theorem 2.2] in sub-Riemannian structure.

Our paper is organized as follows. In Section 2, we review notations and structure of sub-Riemannian geometry. In Section 3, we recall our multi-marginal optimal transport problem and its Kantorovich dual problem. Furthermore, we will give the proof for our main results, Theorem 1.1 and Theorem 1.2 in Section 3. Finally, we give our conclusion in Section 4.

2. SUB-RIEMANNIAN GEOMETRY

Throughout in this paper, M denotes a smooth connect manifold of dimension $n \geq 2$.

A smooth distribution Δ of rank $m \leq n$ is a rank m subbundle of the tangent bundle TM i.e, for every $x \in M$ there exists a neighborhood V_x of x in M and m smooth vector fields (X_x^1, \dots, X_x^m) linearly independent on V_x such that

$$\Delta(y) = span\{X_x^1(y), \dots, X_x^m(y)\}, \text{ for every } y \in V_x.$$

In this case, we say that m smooth vector fields (X_x^1, \dots, X_x^m) is a local frame in V_x for the distribution Δ . We call that a distribution Δ on M admits a global frame if there exists m smooth vector fields X_1, \dots, X_m on M such that

$$\Delta(x) = span\{X_1(x), \dots, X_m(x)\}, \text{ for every } x \in M.$$

We denote by for every $\mathfrak{S}^\infty(M)$ the space of smooth vector fields on M . The Lie algebra of vector fields generated by $\mathfrak{S}^\infty(M)$, denoted by $Lie(\mathfrak{S}^\infty(M))$, is the smallest vector subspace S of $\mathfrak{S}^\infty(M)$ satisfying.

$$Lie\{X_x^1, \dots, X_x^m\}(y) = T_y M, \text{ for every } y \in V_x.$$

A set of r smooth vector fields X_1, \dots, X_r is called a generating frame for Δ on M if

$$\Delta(x) = span\{X_1(x), \dots, X_r(x)\}, \text{ for every } x \in M.$$

According [18, Proposition 1.3], if Δ is a distribution of rank $m \leq n$ on M then there exists a generation frame $\{X_1, \dots, X_r\}$ for Δ .

A sub-Riemannian manifold is given by a triple (M, Δ, g) , where Δ is a totally nonholonomic on M and g is a Riemannian metric on M .

Let λ be the Lesbegue measure on $[0,1]$. A map $f : [0,1] \rightarrow \mathbb{R}^n$ is *absolutely continuous* if for each $\varepsilon > 0$, there exists $\delta > 0$ such that for every family of disjoints subintervals $\{(a_k, b_k)\}_{k \in \mathbb{N}}$ in $[0,1]$ satisfying $\sum_{k \in \mathbb{N}} b_k - a_k < \delta$, we have $\sum_{k \in \mathbb{N}} |f(b_k) - f(a_k)| < \varepsilon$. If $g : [0,1] \rightarrow \mathbb{R}^n$ is absolutely continuous then it is differentiable λ -a.e on $[0,1]$ and its derivative $f'(t)$ is integrable with respect to λ on $[0,1]$. A function $f : [0,1] \rightarrow \mathbb{R}^n$ is *absolutely continuous with square integrable derivative* if it is absolutely continuous and $f \in L^2([0,1], \mathbb{R}^n)$. A map $f : [0,1] \rightarrow M$ is *absolutely continuous* (absolutely continuous with square integrable derivative) if it is absolutely continuous (absolutely continuous with square integrable derivative) in charts.

A continuous path $\gamma : [0,1] \rightarrow M$ is an *horizontal path* with respect to Δ if it is absolutely continuous with square integrable derivative and satisfies $\frac{d\gamma}{dt} \in \Delta(\gamma(t))$ a.e. $t \in [0,1]$. If $\gamma : [0,1] \rightarrow M$ is horizontal then from [1, Lemma 3.11] we know that the map $\varphi : [0,1] \rightarrow \mathbb{R}$ defined $\varphi(t) := \sqrt{g\left(\frac{d\gamma(t)}{dt}, \frac{d\gamma(t)}{dt}\right)}$ by is measurable and hence integrable. Therefore, we can define its length by

$$L(\gamma) = \int_0^1 \sqrt{g\left(\frac{d\gamma(t)}{dt}, \frac{d\gamma(t)}{dt}\right)} dt$$

Since λ is absolutely continuous with square integrable derivative, its length $L(\gamma)$ is finite. As M is connected and Δ is totally nonholonomic distribution on M , by Chow-Rashevsky Theorem [18, Theorem 1.14], for every $x, y \in M$ there exists a horizontal path $\gamma : [0,1] \rightarrow M$ such that $\gamma(0) = x$ and $\gamma(1) = y$. For every $x, y \in M$ the *sub-Riemannian distance* is defined by

$$d_{SR}(x, y) := \inf \left\{ L(\gamma) \mid \gamma \text{ is an horizontal, } \gamma(0) = x, \gamma(1) = y \right\}.$$

Then the sub-Riemannian distance is finite and continuous on $M \times M$ (see [18, Proposition 1.16]).

For every $x, y \in M$, an horizontal path $\gamma : [0,1] \rightarrow M$ is called a sub-Riemannian minimizing geodesic between x and y if $d_{SR}(x, y) = L(\gamma)$. We say that the metric space (M, d_{SR}) is a geodesic if for every $x, y \in M$ there exists a sub-Riemannian minimizing geodesic between x and y . If the metric space (M, d_{SR}) is complete the applying [18, Theorem 2.4] we get that the infimum in (2.1) is attained. This means (M, d_{SR}) is a geodesic space. Furthermore, the topology defined by the distance d_{SR} coincides with the original topology of M (see [3, Theorem 3.31] or [18, Proposition 1.16]). Therefore, (M, d_{SR}) is a separable space.

In this paper, we always assume that Δ is distribution of rank $m \leq n$ and it admits a global frame $\{X_1, \dots, X_m\}$. We also assume that (M, d_{SR}) is a complete metric space.

3. MAIN RESULTS

In this section, we introduce the multi-marginal optimal transport over sub-Riemannian manifold structure. Then, we will prove our main results Theorem 1.1 and Theorem 1.2.

For every integer $k \geq 2$, we consider k Borel probability measures μ_1, \dots, μ_k on M . We denote by

$\Pi(\mu_1, \dots, \mu_k)$ the set of all Borel probability measures on $M^k = M \times \dots \times M$ with marginals μ_1, \dots, μ_k

. We recall that $\gamma \in \Pi(\mu_1, \dots, \mu_k)$ if and only if γ is a Borel probability measure on M^k and satisfies

$$\gamma(M^{i-1} \times A \times M^{k-i}) = \mu_i(A),$$

for every Borel subset A of M and every $i = 1, \dots, k$.

Next, we recall that our multi-marginal optimal transport problem by considering the following problem.

$$(M) \quad \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k),$$

where the cost function $c : M^k \rightarrow R$ is defined by

$$c(x_1, \dots, x_k) = \inf_{y \in M} \sum_{i=1}^k \lambda_i d_{SR}^2(x_i, y),$$

with $\lambda_1, \dots, \lambda_k$ are k positive real numbers such that $\sum_{i=1}^k \lambda_i = 1$.

Now, we give the proof of our first main result.

Proof of Theorem 1.1. We choose a sequence of $\alpha^n \in \Pi(\mu_1, \dots, \mu_k)$ such that

$$\lim_{n \rightarrow +\infty} \int_{M^k} c(x_1, \dots, x_k) d\alpha^n(x_1, \dots, x_k) = \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k)$$

As we presented in Section 2, (M, d_{SR}) is separable space and thus, it is a Polish metric space. Hence, applying [9, Lemma 12], one gets that the set $\Pi(\mu_1, \dots, \mu_k)$ is compact. Hence, passing to a subsequence we can assume that $\lim_{n \rightarrow +\infty} \alpha^n = \alpha$ under the weak-topology for some $\alpha \in \Pi(\mu_1, \dots, \mu_k)$. Then, we get that

$$\inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k) \leq \int_{M^k} c(x_1, \dots, x_k) d\alpha(x_1, \dots, x_k)$$

Furthermore, by [5, Lemma 3.2] we obtain that the cost function c is lower-semi continuous and bounded from below. This implies that

$$\begin{aligned} \inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k) &= \lim_{n \rightarrow +\infty} \int_{M^k} c(x_1, \dots, x_k) d\alpha^n(x_1, \dots, x_k) \\ &= \liminf_{n \rightarrow +\infty} \int_{M^k} c(x_1, \dots, x_k) d\alpha^n(x_1, \dots, x_k) \\ &\geq \int_{M^k} c(x_1, \dots, x_k) d\alpha(x_1, \dots, x_k). \end{aligned}$$

Therefore,

$$\inf_{\gamma \in \Pi(\mu_1, \dots, \mu_k)} \int_{M^k} c(x_1, \dots, x_k) d\gamma(x_1, \dots, x_k) = \int_{M^k} c(x_1, \dots, x_k) d\alpha(x_1, \dots, x_k)$$

Hence, the problem (M) has solutions. □

Next, we give a Kantorovich dual problem for our problem (M) . We recall that the dual problem to (M) is to maximize

$$(D) \quad \sum_{i=1}^k \int_M f_i(x_i) d\mu_i(x_i)$$

among all m -tuples (f_1, \dots, f_k) of functions $f_i \in L^1(\mu_i)$ which satisfied

$$\sum_{i=1}^k f_i(x_i) \leq c(x_1, \dots, x_k),$$

for every $(x_1, \dots, x_k) \in M^k$.

We say that m -tuples of functions (f_1, \dots, f_k) is \mathcal{C} -conjugate if for every $i \in \{1, \dots, k\}$ we have

$$f_i(x_i) = \inf_{x_j \in M, j \neq i} \left\{ c(x_1, \dots, x_k) - \sum_{j \neq i} f_j(x_j) \right\}.$$

The following well-known result can be found in [12].

The problem (D) has a solution.

To conclude this section, we give the proof of Theorem 1.2.

Proof of Theorem 1.2. In this proof, we use the ideas in [15, Theorem 2.2].

By Theorems 1.1 and 3.1, let γ and (g_1, \dots, g_k) be solutions of problems (M) and (D), respectively.

Then, for every $i \in \{1, \dots, k\}$ and $x_i \in M_i$ we define f_i inductively by

$$f_i(x_i) := \inf_{x_j \in M_j, j \neq i} \left\{ c(x_1, \dots, x_k) - \sum_{j=1}^{i-1} f_j(x_j) - \sum_{j=i+1}^k g_j(x_j) \right\}.$$

Observe that

$$f_k(x_k) = \inf_{x_j \in M_j, j=1, k-1} \left\{ c(x_1, \dots, x_k) - \sum_{j=1}^{k-1} f_j(x_j) \right\}.$$

Thus, for every $i=1, \dots, k$ one has

$$f_i(x_i) \leq \inf_{x_j \in M, j \neq i} \left\{ c(x_1, \dots, x_k) - \sum_{j \neq i} f_j(x_j) \right\}.$$

Moreover, for every $i=1, \dots, k$ and $x_i \in M$ we have that

$$g_i(x_i) \leq c(x_1, \dots, x_k) - \sum_{j=1}^{i-1} u_j(x_j) - \sum_{j=i+1}^k g_j(x_j).$$

Hence, $g_i \leq f_i$ on M for every $i=1, \dots, k$. This implies that

$$f_i(x_i) \geq \inf_{x_j \in M, j \neq i} \left\{ c(x_1, \dots, x_k) - \sum_{j \neq i} f_j(x_j) \right\}.$$

Therefore,
$$f_i(x_i) \geq \inf_{x_j \in M, j \neq i} \left\{ c(x_1, \dots, x_k) - \sum_{j \neq i} f_j(x_j) \right\}.$$

This means that (f_1, \dots, f_k) is a \mathcal{C} -conjugate.

Next, by the definitions of f_i and g_i for $i=1, \dots, k$, we obtain that

$$\begin{aligned} \int_{M^k} c(x_1, \dots, x_k) d\gamma &= \sum_{i=1}^k \int_M g_i(x_i) d\mu_i \\ &\leq \sum_{i=1}^k \int_M f_i(x_i) d\mu_i \\ &\leq \int_{M^k} c(x_1, \dots, x_k) d\gamma. \end{aligned}$$

So that we must have

$$\sum_{i=1}^k f_i(x_i) = c(x_1, \dots, x_k) \quad \gamma - a.e.$$

Hence, we get the result. □

4. CONCLUSION

In this work, we introduce and investigate the multi-marginal optimal transport on sub-Riemannian manifold. The contributions of our paper are two-fold.

- 1) We are the first one providing the multi-marginal optimal transport over a sub-Riemannian manifold that cover both problems in [13] and [16].
- 2) By considering the properties of the cost function

$$c(x_1, \dots, x_k) = \inf_{y \in M} \sum_{i=1}^k \lambda_i d_{SR}^2(x_i, y)$$

we can prove that our problem has solutions. Moreover, we also introduce a Kantorovich dual problem for our problem and show the link between them.

Comparing with the cost function which is studied in [15], in this paper, we only focus on the cost function which is defined based on the squared sub-Riemannian distances. We can extend this cost function by considering the following cost function

$$c_p(x_1, \dots, x_k) = \inf_{y \in M} \sum_{i=1}^k \lambda_i d_{SR}^p(x_i, y),$$

for every integer number $p \geq 1$. Our main results Theorem 1.1 and Theorem 1.2 are also true under the cost function c_p .

The theory of multi-marginal optimal transport plays an important role to study the Wasserstein barycenter problem, for details, readers can see [2], [5], [9], and [14]. Therefore, we can apply our main results in this paper for Wasserstein barycenter over sub-Riemannian structure. This problem can be considered in the future.

REFERENCES

- [1] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *Introduction to geodesics in sub-Riemannian geometry*, Geometry, analysis and dynamics on sub-Riemannian manifolds. Vol. II, EMS Ser. Lect. Math., Eur. Math. Soc., Zurich, 2016, 1 – 83.
- [2] Martial Agueh and Guillaume Carlier, *Barycenters in the Wasserstein space*, SIAM J. Math. Anal. 43 (2011), no. 2, 904–924
- [3] Andrei Agrachev, Davide Barilari, and Ugo Boscain, *A Comprehensive Introduction to Sub-Riemannian Geometry*, Cambridge Studies Advanced Mathematics, vol. 181, Cambridge University Press, 2019.
- [4] Luigi Ambrosio, Nicola Gigli, and Giuseppe Savaré, *Gradient flows in metric spaces and in the space of probability measures*, Lectures in Mathematics ETH Zurich, Birkhauser Verlag, Basel, 2005.
- [5] Nhan-Phu Chung and Minh-Nhat Phung, *Barycenters in the Hellinger-Kantorovich Space*, Appl. Math. Optim. 84.2 (2021), 1791–1820.
- [6] Nhan-Phu Chung and Thanh-Son Trinh, *Unbalanced optimal total variation transport problems and generalized Wasserstein barycenters*, Proceedings of the Royal Society of Edinburgh. Section A: Mathematics. 152.3 (2022), 674–700.
- [7] Nhan-Phu Chung and Thanh-Son Trinh, *Weak optimal entropy transport problems*, arXiv: 2101.04986.
- [8] Alfred Galichon, *Optimal transport methods in economics*, Princeton University Press, Princeton, NJ, 2016.
- [9] Thibaut Le Gouic and Jean-Michel Loubes, *Existence and consistency of Wasserstein barycenters*, Probab. Theory Related Fields 168 (2017), no. 3 – 4, 901–917.
- [10] L. Kantorovich, *On the translocation of masses*, C. R. (Doklady) Acad. Sci. URSS (N.S)37(1942), 199–201.
- [11] L. Kantorovich, *On a problem of Monge*, Zap. Nauchn. Sem. S.-Peterburg. Otdel. Mat. Inst. Steklov. (POMI) 312(2004), no. 4.
- [12] Hans G. Kellerer, *Duality theorems for marginal problems*, Z. Wahrsch. Verw. Gebiete 67 (1984), 399–432.
- [13] Young-Heon Kim and Brendan Pass, *Multi-marginal optimal transport on Riemannian manifolds*, Amer. J. Math. 137 (2015), no. 4, 1045–1060.
- [14] Young-Heon Kim and Brendan Pass, *Wasserstein barycenters over Riemannian manifolds*, Adv. Math. 307 (2017), 640–683.
- [15] Brendan Pass, *Uniqueness and Monge solutions in the multimarginal optimal transportation problem*, SIAM J. Math. 43(2011), no. 6, 2758–2775.
- [16] Brendan Pass, Andrea Pinamonti, and Mattia Vedovato, *Multi-marginal optimal transport on Heisenberg group*, arXiv: 2003.10727.
- [17] Gabriel Peyré and Marco Cuturi, *Computational Optimal Transport*, Foundations and Trends in Machine Learning 11 (2019), no. 5-6, 355-607.
- [18] Ludovic Rifford, *Sub-Riemannian geometry and optimal transport*, SpringerBriefs in Mathematics, Springer, Cham, 2014.
- [19] Filippo Santambrogio, *Optimal transport for applied mathematicians*, Progress in Nonlinear Differential Equations and their Applications, vol. 87, Birkhauser/Springer, Cham, 2015. Calculus of variations, PDEs, and modeling.

- [20] Cédric Villani, *Topics in optimal transportation*, Graduate Studies in Mathematics, vol. 58, American Mathematical Society, Providence, RI, 2003.
- [21] Cédric Villani, *Optimal transport*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 338, Springer-Verlag, Berlin, 2009. Old and new.

YSC4F.232

TRẠM QUAN TRẮC VÀ DỰ BÁO THỜI TIẾT SỬ DỤNG RASPBERRY PI ỨNG DỤNG MẠNG NƠI RỒN VÀ GIẢI THUẬT LAN TRUYỀN NGƯỢC

LÊ THỊ CÚC, CAO VĂN KIÊN

Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

lethicucdmt102@gmail.com, caovankien@iuh.edu.vn

Tóm tắt. Trạm quan trắc và dự báo thời tiết có thể thường xuyên được tìm thấy ở hầu hết các phòng nghiên cứu của nhiều trường đại học, vì nó tương đối dễ xây dựng, mô hình hóa và điều khiển về mặt lý thuyết. Hệ thống bao gồm một bộ vi xử lý Raspberry Pi, các cảm biến nhiệt độ, độ ẩm, cảm biến áp suất không khí. Bộ vi xử lý có chức năng tiếp nhận các tín hiệu chuyển đổi điều khiển hoạt động xử lý cùng với cơ cấu chấp hành để đưa ra tín hiệu. Phần cốt lõi của hệ thống là thuật toán mạng neural huấn luyện bằng giải thuật Backpropagation dùng các dữ liệu được lưu trữ trên cloud (thingspeak). Sử dụng phần mềm Matlab để lấy dữ liệu đã thu thập được từ cloud bao gồm dữ liệu cảm biến nhiệt độ - độ ẩm, cảm biến áp suất làm đầu vào và xử lý để đưa ra được dữ liệu đầu ra về thời tiết. Sau mỗi lần hệ thống hoạt động sẽ xác định được những thông số tín hiệu đầu ra tốt nhất để dự đoán thời tiết ít nhất sau 2 tiếng.

Từ khóa. Trạm thời tiết, Raspberry Pi, Backpropagation algorithms, dữ liệu trung gian

WEATHER OBSERVATIONS AND FORECASTS USING RASPBERRY PI NEURAL NETWORK APPLICATION AND DO LAN AIGORITHM

Abstract. Observatories and weather forecasters can often be found in most research laboratories of many universities, as it is relatively easy to build, model, and theoretically control. The system includes a Raspberry Pi microprocessor, temperature, humidity, and air pressure sensors. The microprocessor has the function of receiving the conversion signals that control the processing operation together with the actuator to output the signal. The core part of the system is a neural network algorithm trained by the Backpropagation algorithm using data stored in the cloud (thingspeak). Use Matlab software to get data collected from the cloud including temperature sensor data - humidity, pressure sensor as input and process to give output data about weather. After each operation, the system will determine the best output signal parameters to predict the weather at least after 2 hours.

Keywords. weather station, Raspberry Pi, Backpropagation algorithms, data midata.

1. GIỚI THIỆU

Hiện nay trên thực tế có rất nhiều hệ thống về mạng lưới thời tiết để ứng dụng vào cuộc sống của con người như dự đoán sản lượng mặt trời, dự đoán lượng mưa, dự đoán về sự trì hoãn của hãng hàng không [1-4]... Với sự phát triển mạnh mẽ của IoT (Internet of Things) và giám sát dữ liệu đã được mở rộng thông qua web và các thiết bị được kết nối Internet [5]. Từ đó tạo được sự thuận tiện và hiện đại trong cuộc sống của con người.

Tại Việt Nam cũng như trên thế giới đang áp dụng nhiều phương pháp quan trắc mới ứng dụng công nghệ tự động hóa cho ta chuỗi số liệu liên tục với độ chính xác cao. León và cộng sự [6] sử dụng Arduino làm trạm thu thập dữ liệu thời tiết. Varghese và cộng sự [7] sử dụng Raspberry Pi làm trạm dự báo thời tiết.

Bên cạnh đó, ngày càng nhiều các phương pháp, mô hình phân tích dự báo thời tiết được công bố, đa dạng trong hướng tiếp cận đã tạo ra thách thức lớn cho con người trong vấn đề tìm hiểu và áp dụng mô hình hiện

có một cách hiệu quả [8-10]. Mỗi hướng tiếp cận có nhiều mô hình khác nhau. Tiếp cận dựa vào dữ liệu tự thu thập, từ những dữ liệu các trạm thời tiết có sẵn. Tiếp cận dựa trên máy học: thuật toán học có giám sát (supervised learning), thuật toán học không giám sát (unsupervised learning). Và hướng tiếp cận kết hợp là kết hợp hai hoặc nhiều mô hình. Với sự đa dạng mô hình thuộc các hướng tiếp cận trên, đòi hỏi người sử dụng phải hiểu rõ mô hình mình chọn để làm và đưa ra được mô hình đó có thể áp dụng chính xác và giải quyết bài toán đặt ra.

Vấn đề đặt ra là làm cách nào để cho hệ thống có thể hoạt động đạt được hiệu quả tốt nhất, đồng thời đáp ứng được những yêu cầu đưa ra. Bên cạnh đó, sự hoạt động ổn định của một hệ thống giữ vai trò rất quan trọng trong mọi lĩnh vực. Vì chúng gắn liền với sinh hoạt hàng ngày của cuộc sống con người. Vì vậy, đề tài này không những là một thực tại khách quan mà nó còn đóng vai trò đặc biệt quan trọng thực sự ở hiện tại cũng như trong tương lai sau này. Do đó, việc dự báo thời tiết một cách chính xác và khách quan là một nhu cầu cần thiết đây chính là lý do mà nhóm nghiên cứu quyết định chọn đề tài này.

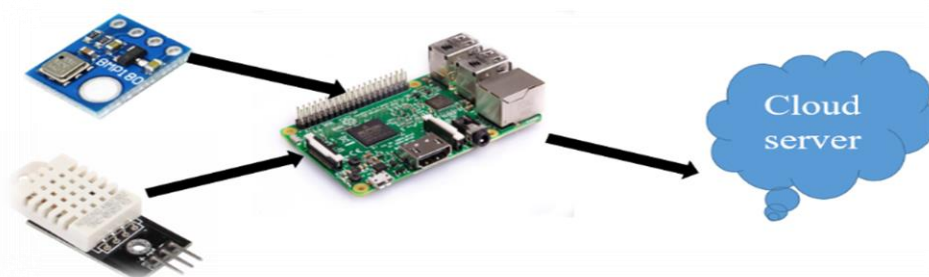
Bài báo sử dụng phần cứng Raspberry Pi để thu thập các dữ liệu nhiệt độ, độ ẩm và áp suất không khí, sau đó ứng dụng mạng nơ ron để dự báo giá trị nhiệt độ, độ ẩm không khí sau 2 tiếng áp dụng giải thuật Backpropagation.

Cấu trúc bài báo gồm 5 mục. Mục 2 giới thiệu về mô hình phần cứng và cơ sở lý thuyết mạng nơ ron, giải thuật Backpropagation. Mục 3 trình bày về quy trình xây dựng hệ thống dự báo thời tiết. Mục 4 nói về thực nghiệm và cuối cùng mục 5 đưa ra các kết luận.

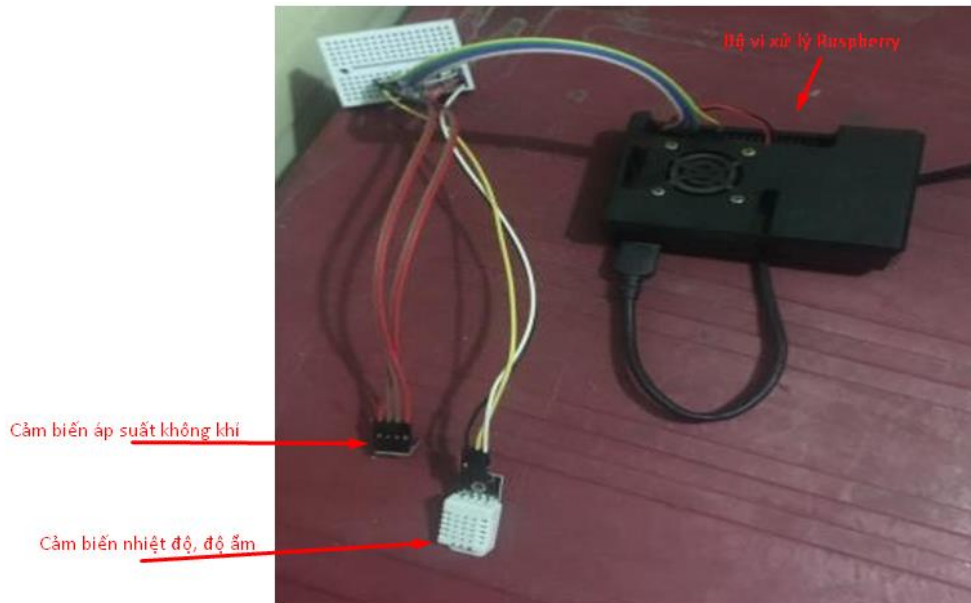
2. MÔ HÌNH PHẦN CỨNG VÀ CƠ SỞ

2.1. Xây dựng mô hình phần cứng và thu thập dữ liệu

Mô hình phần cứng (H.1) gồm bộ vi điều khiển trung tâm (Raspberry pi3) thực hiện làm bộ xử lý chính để nhận và truyền tín hiệu đi. Các cảm biến đảm nhiệm chức năng đo lường các giá trị môi trường và chuyển thành các tín hiệu điện áp chuyển về cho bộ xử lý. Để có được hiệu suất làm việc cao về cả phạm vi đo nhiệt độ lẫn phạm vi đo độ ẩm, mức sai số của các thông số cũng nhỏ nhất thì ở mô hình phần cứng là DHT22, BMP180 và bộ vi xử lý Raspberry PI 3 được sử dụng để lưu trữ dữ liệu được thu thập vào đám mây. Quá trình thu thập dữ liệu sẽ được thực hiện và lưu trữ trên đám mây thingspeak. Để phù hợp với bài toán đặt ra thì em phải xử lý tốc độ lấy mẫu rất nhiều lần và đã chọn ra tốc độ lấy mẫu là cứ sau 15 phút sẽ gửi lên thingspeak 1 mẫu. Những lúc mất tín hiệu truy cập wifi thì dữ liệu sẽ không được gửi lên và chờ khi có tín hiệu gửi lên.



H.1. Mô hình phần cứng trên lý thuyết

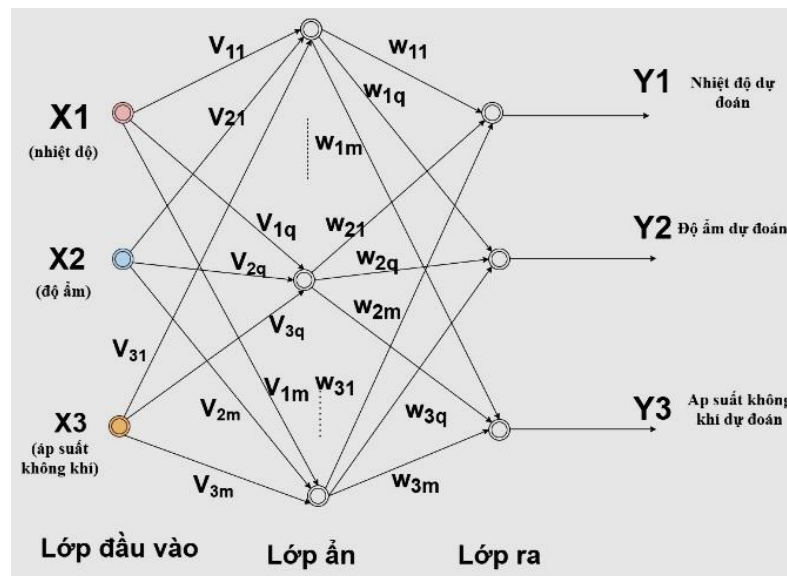


H.2. Mô hình phần cứng trên thực tế

Để thu thập các thông số dữ liệu cho hệ thống cần thiết kế một mô hình phần cứng (H.2) thực hiện các chức năng thu thập dữ liệu môi trường và gửi đi lưu trữ. Mô hình phần cứng gồm bộ vi điều khiển trung tâm (Raspberry pi3) thực hiện làm bộ xử lý chính để nhận và truyền tín hiệu đi. Các cảm biến đảm nhiệm chức năng đo lường các giá trị môi trường và chuyển thành các tín hiệu điện áp chuyển về cho bộ xử lý.

2.2. Cấu trúc thuật toán mạng neural huấn luyện bằng giải thuật Backpropagation

Một mạng neural (H.3) bao gồm một lớp vào, một lớp ra và một hoặc nhiều lớp ẩn nằm giữa lớp vào và lớp ra.



H.3 Sơ đồ cấu trúc mạng neural

Đầu vào: x_j ($j = 1, \dots, 3$ cho bài toán X1: nhiệt độ; X2: độ ẩm; X3: áp suất không khí)

Lớp vào: chọn số nơ ron ($q = 1, \dots, m$ trong bài toán dự báo thời tiết chọn 8 nơ ron) trong đó q là nơ ron thứ q của lớp vào; v_{jq} là trọng số đầu vào thứ j nối với nơ ron thứ q (tổng số có $qj = 3 \times 8 = 24$ trọng số trong bài toán dự báo thời tiết)

$$Net_q = \sum_{j=1}^m v_{qj} x_j \quad (1)$$

Lớp ẩn: đầu ra số của lớp vào chính là đầu vào của lớp ẩn, là một hàm của đầu vào tương tự

$$z_q = f_q (net_q) \quad (2)$$

Hàm tác động đầu ra của mỗi nơ ron của tất cả các lớp: chọn hàm tansig $f_q (.)$:

$$f_q (net_q) = \frac{2}{1 + e^{-2net_q}} - 1 \quad (3)$$

Chọn n=M nơ ron cho lớp ẩn. Số này được chọn theo thực nghiệm bài toán từ những kết quả tối ưu và giá trị lỗi trong quá trình huấn luyện.

Lớp ra: Tổng đầu ra của nơ ron thứ i (net_i) và đầu ra Y_i có thể xác định theo:

$$net_i = \sum_{q=1}^l w_{iq} z_q = \sum_{q=1}^l w_{iq} f \left(\sum_{j=1}^m v_{qj} x_j \right) \quad (4)$$

$$y_i = f (net_i) = f \left(\sum_{q=1}^l w_{iq} z_q \right) = f \left(\sum_{q=1}^l w_{iq} f \left(\sum_{j=1}^m v_{qj} x_j \right) \right) \quad (5)$$

Trong đó i là số nơ ron thứ I trong n nơ-ron đầu ra được chọn tùy thuộc vào cách mã hóa đặc trưng đầu ra. Trong bài toán đã nêu số nơ-ron lớp ra là n=3.

2.3. Luật học bằng giải thuật Backpropagation

Thuật toán Backpropagation tìm kiếm một vector các trọng số giúp cực tiểu hóa lỗi tổng thể của hệ thống đối lập với học tập.

Giải thuật Backpropagation bao gồm hai giai đoạn:

Giai đoạn học: xác định các tham số của W_{ij} của mạng.

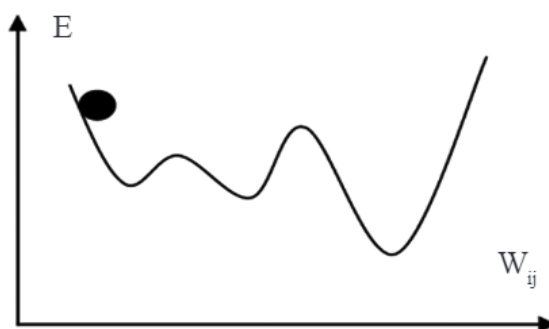
Giai đoạn thực hiện bằng giải thuật Backpropagation lỗi (giai đoạn đủ điều kiện để chạy) chạy để kiểm tra và thể hiện khả năng suy diễn của mạng đối với những mẫu lạ (Chưa được học). Để mạng hoạt động được cần xác định giá trị tất cả các tham số w_{ij} của các lớp. Phương pháp chung để xác định chúng là phương pháp “Thử - Sai – Chỉnh”.

Thử có nghĩa là thử ngẫu nhiên các giá trị trọng số ban đầu W_{ij} ngoài ra cần thêm giá trị tốc độ học “ η ” cùng với tốc độ sai số đầu ra E (trong công thức 6) tối thiểu để chạy được toán học.

Sai có nghĩa là theo họ có giám sát: là hiệu của đầu ra mong muốn d_i với đầu ra thực tế Y_i của từng nơ ron ở mỗi lớp.

Chỉnh có nghĩa là điều chỉnh trọng số với giá trị “Thử” ban đầu trong một lượng ΔW_{ij} (trong công thức 7).

$$E = \frac{1}{2} \sum_{i=1}^n (d_i - y_i)^2 \quad (6)$$



H.4. Tối thiểu cục bộ

Sử dụng phương pháp hạ Gradient (H.4): tức là tiến trình cập nhật trọng số theo hướng gradient, trọng số liên kết giữa các lớp (ví dụ giữa lớp ẩn và lớp ra) được cập nhật bởi:

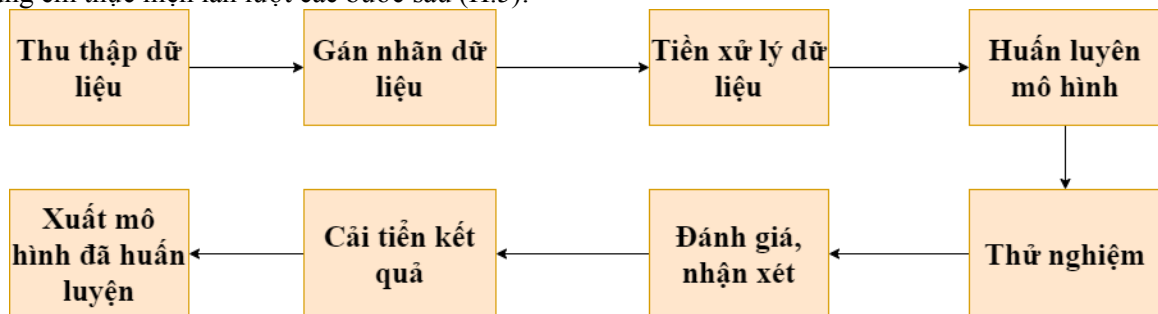
$$\Delta w_{iq} = -\eta \left[\frac{\partial E}{\partial y_i} \right] \left[\frac{\partial y_i}{\partial w_{iq}} \right] \left[\frac{\partial net_i}{\partial w_{iq}} \right] \quad (7)$$

Trong đó: η là hằng số học. Nó chỉ tốc độ cập nhật trọng số từng bước lặp của thuật toán. Kỹ thuật hạ có thể được áp dụng để tính trọng số W_{ij} từ lớp đầu vào đến lớp đầu ra gọi là tính trọng số theo kiểu truyền thẳng. Ngược lại, có thể tính trọng số từ lớp ra ngược về lớp đầu (làm sai số có xu hướng giảm dần). Thuật toán ngược từ ra tới vào gọi là thuật toán lan truyền ngược.

3. XÂY DỰNG HỆ THỐNG DỰ BÁO THỜI TIẾT

3.1. Phân tích bài toán

Để xây dựng hoàn chỉnh một bộ dữ liệu cho nhiệm vụ phân tích dự báo thời tiết tại khu vực em đang sống, chúng em thực hiện lần lượt các bước sau (H.5):



H.5 Sơ đồ khối tổng quát của bài toán dự báo thời tiết

Bước 1: Thu thập dữ liệu

Dữ liệu sẽ được lấy từ thingspseak về phần mềm matlab để xử lý. Để giải quyết bài toán tốt thì ngoài mô hình xử lý tốt thì phải có bộ dữ liệu phù hợp nên dữ liệu là yếu tố quan trọng góp phần trực tiếp vào kết quả của bài toán. Cần có bộ dữ liệu đủ tính đa dạng, đủ nhiều để mô hình có thể học được những tính năng cần thiết.

Bước 2: Gán nhãn dữ liệu

Dữ liệu sau khi thu thập được sẽ được tiến hành gán nhãn trên từng dòng dữ liệu. Việc gán nhãn ảnh hưởng rất lớn đến chất lượng mô hình vì mô hình sẽ học và dự đoán dựa theo nhãn của dữ liệu. Vì vậy, nếu gán nhãn sai gây ra trường hợp nhiễu dữ liệu đồng nghĩa với việc mô hình sẽ học và đưa ra các dự đoán sai.

Bước 3: Tiền xử lý dữ liệu

Để có kết quả tốt, dữ liệu trước khi đưa vào mô hình sẽ xử lý để có dữ liệu sạch, bước này tuy đơn giản nhưng lại là bước quan trọng để bớt nguy cơ nhiễu của dữ liệu, từ đó đến Bước 4 huấn luyện mô hình thì mô hình sẽ học các đặc trưng tốt hơn.

Bước 4: Huấn luyện mô hình

Ở bước này, trước khi đưa vào huấn luyện ta tiến hành tiền xử lý dữ liệu và tách dữ liệu theo dấu thời gian dự đoán rồi mới đưa dữ liệu vào kiến trúc mô hình đã xây dựng cùng tinh chỉnh các tham số để giải quyết bài toán tốt nhất, đây cũng là bước quan trọng nhất để mô hình có thể đưa ra các kết quả tốt nhất. Trong quá trình huấn luyện, mô hình sẽ học các đặc trưng từ bộ dữ liệu và cuối cùng tạo ra mô hình dự báo thời tiết.

Bước 5: Thử nghiệm

Sau khi có được mô hình đã huấn luyện, ta tiến hành thử nghiệm để biết mô hình có học đúng cách hay không. Chúng ta đưa cơ sở dữ liệu xem mô hình có trả về kết quả thời tiết đúng không. Sau đó ta có thể tinh chỉnh lại các tham số trong mô hình để có kết quả tối ưu nhất.

Bước 6: Đánh giá, nhận xét

Dựa vào kết quả thử nghiệm ở Bước 4, đưa ra những nhận xét về kết quả dự đoán của mô hình. Đồng thời đưa ra các nhận xét, đánh giá, so sánh về độ hiệu quả của các dữ liệu đưa vào theo dấu thời gian trong việc giải quyết bài toán và giải thích nguyên nhân những kết quả này. Cuối cùng rút ra đánh giá về độ phù hợp của mô hình trong bài toán này.

Bước 7: Cải tiến kết quả

Từ các đánh giá ở Bước 5, biết được nguyên nhân sai sót của các dự đoán sai. Các dự báo chưa chính xác sẽ tiến hành sửa chữa, loại bỏ những nhãn được gán sai, bổ sung những trường hợp vì ít dữ liệu mà dẫn đến dự đoán sai. Vì vậy, bước này không chỉ loại bỏ đi các dữ liệu nhiễu mà còn tăng khả năng học của mô hình để mô hình đạt độ chính xác cao hơn.

Bước 8: Xuất mô hình đã huấn luyện

Sau khi đã có kết quả tốt, xuất ra mô hình để sử dụng cho các mục đích của bài toán đặt ra. Mô hình sẽ được sử dụng để dự đoán thời tiết từ các thông số của dữ liệu đầu vào.

3.2. Dữ liệu cho thực nghiệm

Trong quá trình thu thập dữ liệu cứ sau mỗi 15 phút sẽ nhận được một mẫu, có nghĩa là 4 lần mỗi giờ. Vì trong vòng 60 phút quá trình thu thập nhiệt độ, độ ẩm, áp suất không khí không có thay đổi nhiều. Bộ dữ liệu với 3 tham số chính: nhiệt độ, độ ẩm, áp suất không khí với số bản ghi là 10000.

Sau khi theo dõi bảng dữ liệu thì sẽ có 4 đầu thời gian ($4/4 = 1$ giờ). Dữ liệu này sẽ được sử dụng để dự đoán nhiệt độ - độ ẩm, áp suất không khí sau 8 đầu thời gian ($8/4 = 2$ giờ). Từ đó ta có thể xây dựng được mô hình (hình 4.15) được xây dựng trên mạng neural gồm 3 lớp: 1 lớp đầu vào, 1 lớp ẩn và 1 lớp ra.[1]

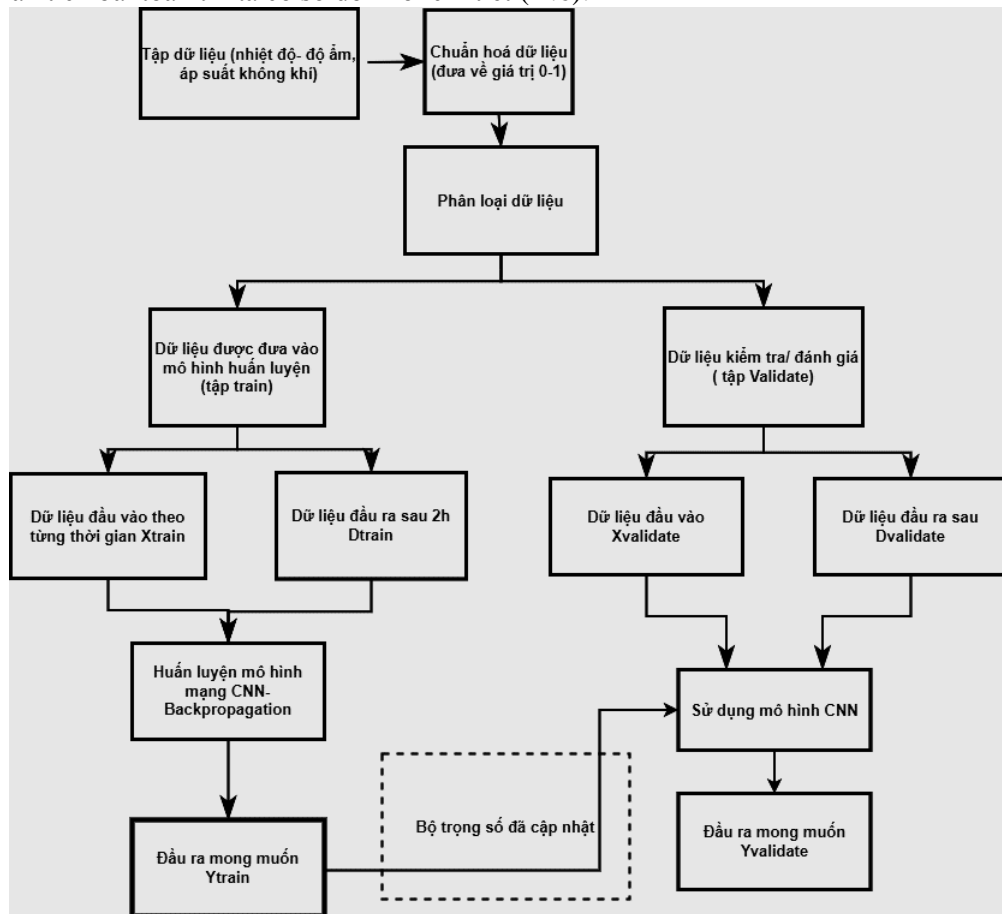
Tiếp theo đó thì bộ dữ liệu sẽ có phạm vi đo/chức năng khác nhau cho nên sẽ có các giá trị thông số dữ liệu sẽ khác nhau, nên đã thực hiện chuẩn hóa để giới hạn các giá trị của đối tượng trong cùng một phạm vi $[0, 1]$ trước khi đưa vào mô hình đào tạo.

Để huấn luyện mô hình Backpropagation trên tập dữ liệu được, ta tiến hành chia tập dữ liệu thành hai nhóm chính:

Nhóm 1: Dữ liệu đưa vào huấn luyện (training) bao gồm có dữ liệu đầu vào theo từng thời gian X_{train} và dữ liệu đầu ra theo 2h huấn luyện D_{train} .

Nhóm 2: Dữ liệu kiểm tra/đánh giá (validate) bao gồm có dữ liệu đầu vào theo từng thời gian $X_{validate}$ và dữ liệu đầu ra theo 2h huấn luyện $D_{validate}$.

Sau khi phân tích bài toán thì ta có sơ đồ khối chi tiết (H.6):



H.6 Sơ đồ khối quá trình phân tích bài toán chi tiết

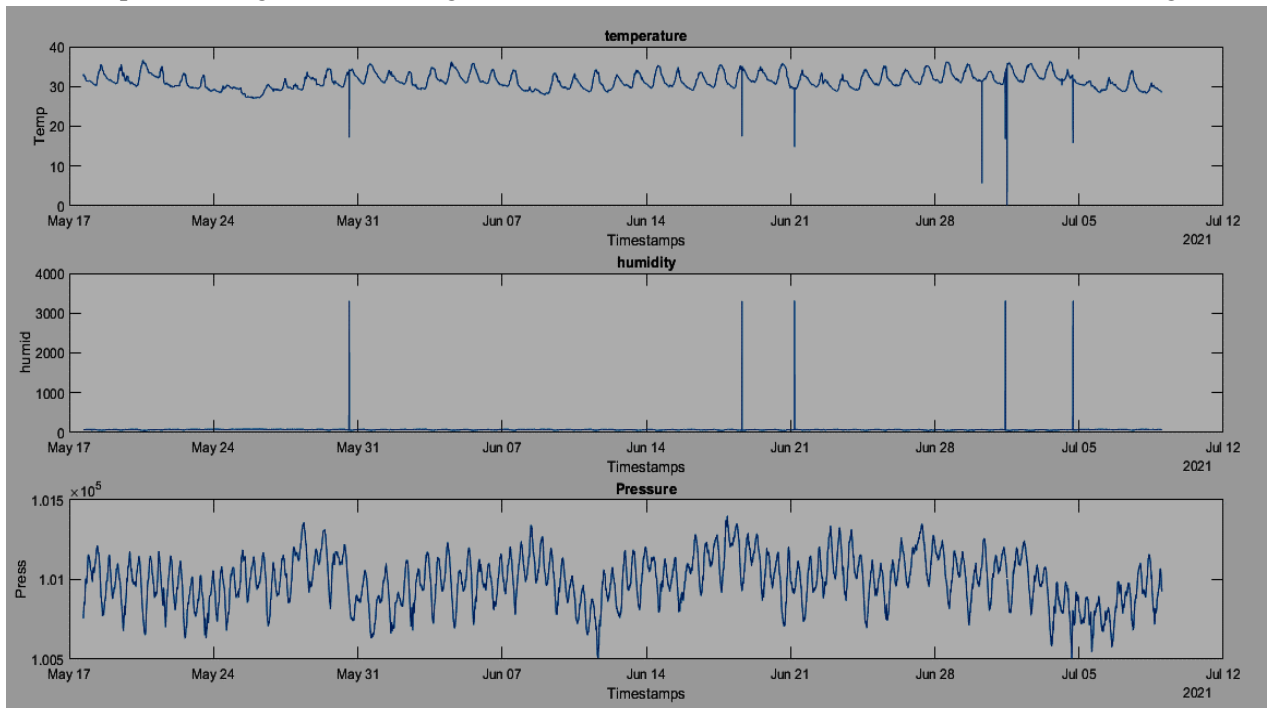
4. KẾT QUẢ THỰC NGHIỆM

Bước 1: Phân tích kiến trúc mô hình

Tính tới thời điểm hiện tại bộ dữ liệu đã được thu thập khá nhỏ, chỉ được 10000 dòng dữ liệu nên phải được chia tập huấn luyện (train) đủ lớn để mô hình có thể bao quát được dữ liệu, phần còn lại chia đều cho tập đánh giá (validation) và tập kiểm tra (test) đủ để đánh giá tốt mô hình, thường thì tỉ lệ của tập đánh giá và tập kiểm tra bằng 1/3 hay 1/4 so với tập huấn luyện. Vì vậy, bộ dữ liệu đã được chia thành các tập huấn luyện tập kiểm thử và tập đánh giá với tỉ lệ lần lượt cho các tập train/ test là 60/20.

Bước 2: Gán nhãn dữ liệu

Dữ liệu đánh giá các đầu ra của thời tiết sau khi thu thập được sẽ được tiến hành gán nhãn nhiệt độ, độ ẩm, áp suất không khí một cách thủ công trên từng dòng dữ liệu. Sau khi gán nhãn, kết quả cuối cùng thu được bộ dữ liệu Database có ma trận 10000×3 . Cụ thể nhãn nhiệt độ có 10000 hàng cột 1, độ ẩm có 10000 hàng cột 2 và áp suất không khí 10000 hàng cột 3. Để cho dễ nhìn thì hình H.7 đã thể hiện dữ liệu ở dạng đồ thị

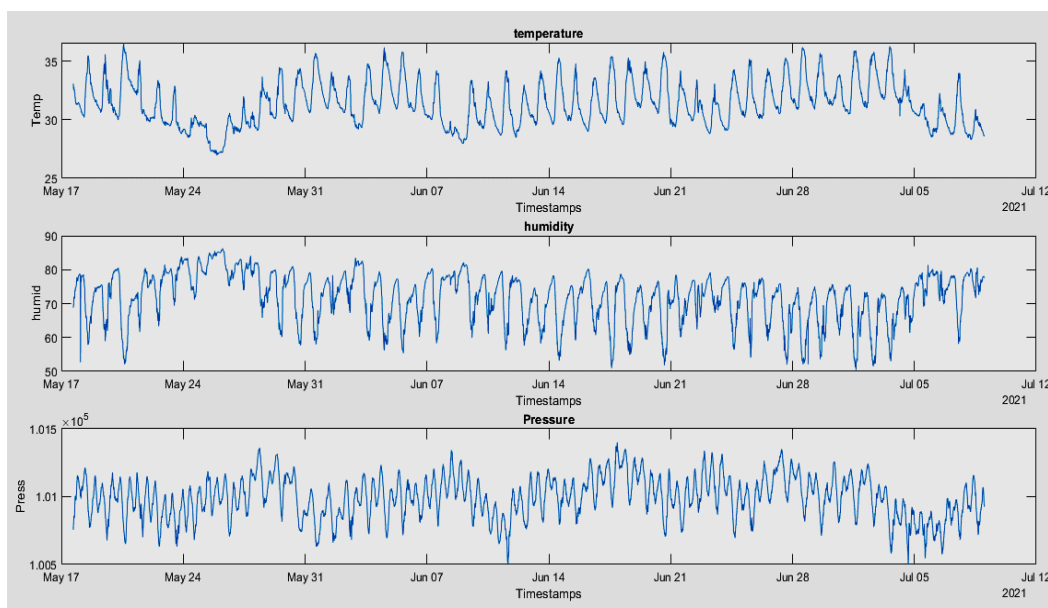


H.7 Minh họa gán nhãn ở dạng đồ thị

Bước 3: Tiền xử lý dữ liệu

- Chuẩn hóa dữ liệu (H.8):
- Xóa các dữ liệu không đúng, ngắt quãng
- Sửa các thông số rời rạc

- Xem kỹ số liệu khi đưa vào đúng/không đúng



H.8 Chuẩn hóa dữ liệu

Từ H.7 ta có thể phân biệt với dữ liệu trước khi chưa qua làm sạch H.8 qua bảng B.1 như sau:

B.1 Phân biệt quá trình chuẩn hóa dữ liệu và chưa chuẩn hóa dữ liệu

Đặc điểm	Chưa chuẩn hóa dữ liệu (hình 4.18)	Đã chuẩn hóa dữ liệu (hình 4.19)
Số lượng dòng	8000x4	7993x4
Hình ảnh dữ liệu quan sát	Dữ liệu rời rạc Dữ liệu bị đứt quãng Khó quan sát	Dữ Liệu tục Dễ quan sát

Sau khi qua quá trình chuẩn hóa trên được sử dụng để tạo bộ dữ liệu huấn luyện mô hình Backpropagation. Có ba mẫu dữ liệu, tiến hành tạo trên mỗi mẫu 7101 dòng làm tập huấn luyện (training data) và 1982 mẫu tương ứng với dữ liệu ở 120 phút làm tập validate. Sau khi đã có mô hình huấn luyện, ta tiến hành thử nghiệm các siêu tham số để kiểm tra kết quả dự đoán theo dấu thời gian.

Các tham số huấn luyện mô hình

Learning rate (tốc độ học_eta) là một siêu tham số kiểm soát mức độ thay đổi mô hình để đáp ứng với sai số ước tính mỗi khi trọng số của mô hình được cập nhật. Cụ thể, learning rate là một siêu tham số có thể cấu hình được sử dụng trong việc đào tạo mạng nơ-ron có giá trị dương nhỏ, thường nằm trong khoảng từ 0 đến 1. Việc chọn learning rate là một thách thức vì giá trị quá nhỏ có thể dẫn đến quá trình luyện tập lâu dài, trong khi giá trị quá lớn có thể dẫn đến việc học một bộ trọng lượng tối ưu mới quá nhanh hoặc quá trình luyện tập không ổn định.[6]

Epoch là số lần mà chúng ta đưa tất cả dữ liệu huấn luyện vào mô hình. Tại mỗi epoch, tập dữ liệu được chia thành các batch (size nhỏ hơn) với kích thước bằng batch-size để đưa vào mô hình (đặc biệt trong trường hợp dữ liệu quá lớn, không thể đưa hết cả tập dữ liệu vào huấn luyện được).

Batch size là số lượng dữ liệu huấn luyện được sử dụng trong một iteration (chỉ số lần mà tham số của mô hình được cập nhật).

B.2 Số liệu thực nghiệm

Thời gian dự đoán	No ron ẩn	Tốc độ học	Max_epoch	Sai số E (nhiệt độ)	Sai số E (độ ẩm)	Sai số E (áp suất không khí)
120 phút	27	0.03	1000	4.57	0.0078	2.2
	35	0.35	10000	5.5	0.0075	2.5
	41	0.5	1000	NaN	NaN	NaN
	41	0.001	100000	2.4	1.18	1.98
	41	0.03	1000	3.0877	0.0079	3.4

	41	0.0001	10000	4.8	0.00996	1.5
	41	0.0009	100000	1.7	0.065	3.01
	41	0.03	10000	7.6	0.06	6.47

Sau khi đã thử nghiệm mọi trường hợp ta chọn ra trường hợp đúng nhất với các tiêu chí sau:

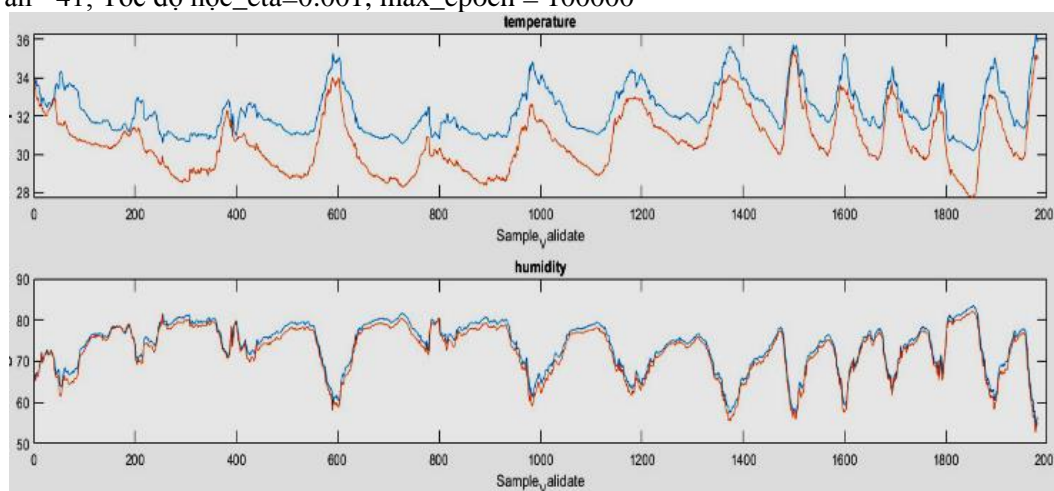
- Tỷ lệ sai số nhỏ
- Kết quả đầu ra và kết quả đầu ra mong muốn phải bằng hoặc xấp xỉ nhau

Nhận xét: Qua quá trình thử nghiệm thì sẽ cho ta thấy số nơ ron ẩn là quá trình được thử đi thử lại rất nhiều lần. Nếu số nơ-ron ít quá thì sẽ không đủ phức tạp để huấn luyện hiệu quả. Nếu số nơ-ron ẩn nhiều quá thì quá trình nó dài hơn và không có hiệu quả. Đối với tần số học khi thay đổi tốc độ học thì nó sẽ ảnh hưởng/tác động đến thời gian học (Max_epoch). Khi tốc độ học quá nhanh thì dù thời gian học có nhiều đến đâu cũng không thể cho ra kết quả đúng được. Khi tốc độ chậm quá thì làm cho hiệu suất học không hiệu quả và tốn thời gian hội tụ. Ngõ ra của quá trình huấn luyện và sai số sẽ được tính ở mỗi epoch trọng số và ngưỡng mới được tính dựa trên tốc độ học hiện tại. Ngõ ra và sai số mới sau đó mới được tính.

Sau khi phân tích hệ thống và tìm hiểu giải thuật, ta đã xây dựng được chương trình dự báo thời tiết sau 2h chạy trên matlab:

Kết quả huấn luyện sau 120 phút (H.9) ta có các tham số sau:

Nơ ron ẩn =41; Tốc độ học_eta=0.001; max_epoch = 100000



H.9 Kết quả huấn luyện sau 120 phút (2h)

Trong đó:

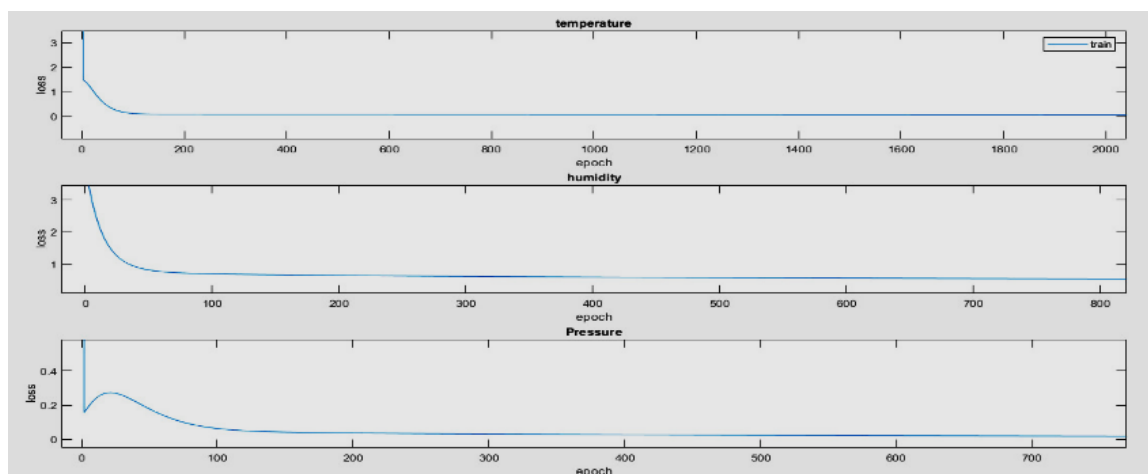
+ Đường đứt màu cam (train) là đồ thị nhiệt độ -độ ẩm được dự báo.

+ Đường nét liền màu xanh (validate) là đồ thị nhiệt độ - độ ẩm ban đầu.

Nhận xét: có thể thấy rằng kết quả của validate và train nó tương đồng với nhau, tiến tới với nhau. Và giá trị sai số trung bình nó hội tụ về 0. Cho thấy kết quả đầu ra sai lệch so với kết quả đầu ra mong muốn không đáng kể.

Đánh giá kết quả

Quá trình huấn luyện cho mô hình Backpropagation –CNN cho bài toán phân tích dữ liệu để dự báo thời tiết được biểu diễn ở H.10



H.10 Biểu diễn độ lỗi trong quá trình huấn luyện dự đoán 120 phút của mô hình Backpropagation – CNN

Nhận xét: Giá trị độ lỗi của mô hình trên tập huấn luyện giảm nhanh từ epoch 0 đến hết quá trình huấn luyện.

- Đối với nhiệt độ giảm mạnh từ 2.6 xuống 0.4 và có xu hướng tiếp tục giảm thêm hơn nữa. Mô hình giảm mạnh từ epoch đầu tiên đến epoch thứ 2. Mô hình đạt độ lỗi nhỏ nhất tại epoch thứ 100000 với độ lỗi là 0.0564
- Đối với độ ẩm tăng nhẹ từ 2.1 xuống 6.9 và giảm từ epoch thứ 8000 với độ lỗi là 0.432 có xu hướng tiếp tục giảm thêm hơn nữa. Mô hình đạt độ lỗi nhỏ nhất tại epoch thứ 100000 với độ lỗi là 0.2951
- Đối với áp suất không khí giảm mạnh từ 2.2 xuống 0.0024 và có xu hướng tiếp tục giảm thêm hơn nữa. Mô hình giảm mạnh từ epoch đầu tiên đến epoch thứ 2. Mô hình đạt độ lỗi nhỏ nhất tại epoch thứ 100000 với độ lỗi là 0.0023

5. KẾT LUẬN

Từ kết quả thực nghiệm có thể thấy, trên tập dữ liệu về đánh giá, mô hình CNN -Backpropagation cho kết quả đánh giá trên sự sai số. Tất cả đều ở mức độ 70-80% ở mức độ tương đối. Chúng ta có thể tham khảo một số công trình sử dụng LSTM để có thể thấy được sự hiệu quả của mô hình. Chẳng hạn như, theo nhóm tác giả “Prabhanshu Attri, Yashika Sharma, Kristi Takach, Falak Shah” [11] Tác giả đã sử dụng mô hình LSTM để giải quyết bài toán trên bộ dữ liệu “Trạm Thời tiết, Viện Hóa sinh Max Planck ở Jena, Đức” lấy về. Bộ dữ liệu này chỉ có 5 nhãn. Có thể thấy, mô hình của em đã đạt được độ chính xác tương đối dù phân trên 3 nhãn và dữ liệu huấn luyện nhỏ hơn những bộ dữ liệu khác. Thử nghiệm này cho thấy, mô hình mạng Nơ ron kết hợp giải thuật Backpropagation có thể được dùng cho phân tích dự báo thời tiết trong lĩnh vực khí tượng học và khí hậu học.

TÀI LIỆU THAM KHẢO

- [1] BABOO, S.Santhosh, SHEREEF and I.Kadar, "International journal of environmental science and development," An efficient weather forecasting system using artificial neural network, vol. 1.4, p. 321, 2010.
- [2] Lagouvardos, K., Kotroni, V., Bezes, A., Koletsis, I., Kopania, T., Lykoudis, S., ... & Vougioukas, S. (2017). The automatic weather stations NOANN network of the National Observatory of Athens: operation and database. *Geoscience Data Journal*, 4(1), 4-16.
- [3] Kreuzer, D., Munz, M., & Schlüter, S. (2020). Short-term temperature forecasts using a convolutional neural network—An application to different weather stations in Germany. *Machine Learning with Applications*, 2, 100007.

- [4] Yonekura, K., Hattori, H., & Suzuki, T. (2018, December). Short-term local weather forecast using dense weather station by deep neural network. In 2018 IEEE international conference on big data (big data) (pp. 1683-1690). IEEE.
- [5] Shafique, K., Khawaja, B. A., Sabir, F., Qazi, S., & Mustaqim, M. (2020). Internet of things (IoT) for next-generation smart systems: A review of current challenges, future trends and prospects for emerging 5G-IoT scenarios. *Ieee Access*, 8, 23022-23040.
- [6] León, J. R. C., Martínez-González, R. F., Medina, A. M., & Peralta-Pelaez, L. A. (2017). Raspberry pi and arduino uno working together as a basic meteorological station. arXiv preprint arXiv:1711.09750.
- [7] Kapoor, P., & Barbhuiya, F. A. (2019, October). Cloud Based Weather Station using IoT Devices. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON) (pp. 2357-2362). IEEE.
- [8] Al-Kahlout, M. M., Ghaly, A. M. A., Mudawah, D. Z., & Abu-Naser, S. S. (2020). Neural Network Approach to Predict Forest Fires using Meteorological Data. *International Journal of Academic Engineering Research (IJAER)*, 4(9).
- [9] Petković, B., Petković, D., Kuzman, B., Milovančević, M., Wakil, K., Ho, L. S., & Jermittiparsert, K. (2020). Neuro-fuzzy estimation of reference crop evapotranspiration by neuro fuzzy logic based on weather conditions. *Computers and Electronics in Agriculture*, 173, 105358.
- [10] Sivaneasan, B., Yu, C. Y., & Goh, K. P. (2017). Solar forecasting using ANN with fuzzy logic pre-processing. *Energy procedia*, 143, 727-732.
- [11] Prabhanshu Attri, Yashika Sharma, Kristi Takach, Falak Shah, “Timeseries forecasting for weather prediction”, 2020/07/20, https://keras.io/examples/timeseries/timeseries_weather_forecasting/

YSC4F.233

SỬ DỤNG THUẬT TOÁN HEURISTIC ĐỂ TỐI ƯU BÀI TOÁN GIAO HÀNG

LÂM NGỌC LONG, VÕ HOÀNG TUẤN, NGUYỄN ĐÌNH THANH, HỒ HOÀNG VÂN ANH,
ĐOÀN QUỐC HUY, LÊ PHÚC LỮ

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
lamngoclong1004@gmail.com, tuanvohoang137@gmail.com, thanhnd2462245@gmail.com
vananh7702@gmail.com, doanquochuy1411@gmail.com, lplu@iuh.edu.vn

Tóm tắt. Trong bối cảnh hiện nay, mỗi khi các shipper nhận được các đơn hàng thì họ phải chọn ra tuyến đường phù hợp để giao hàng theo yêu cầu. Ứng dụng được sử dụng thường xuyên là dựa trên nền tảng Openstreetmap API để tìm đường đi giữa hai địa điểm. Tuy nhiên, công nghệ này chưa giải quyết triệt để bài toán tìm đường đi qua nhiều địa điểm sao cho tổng quãng đường là ngắn nhất và hạn chế được việc lặp lại đường đã đi càng nhiều càng tốt. Vì vậy, nhóm đề xuất sử dụng thuật toán heuristic là 2-OPT (2-Optional Practical Training) để tìm ra một phương án tối ưu mang tính địa phương trong thời gian thực. Bài nghiên cứu sẽ được thực hiện dựa trên các tuyến đường của Thành phố Hồ Chí Minh và trực quan hóa dữ liệu lên Openstreetmap sử dụng các API được hỗ trợ sẵn.

Từ khóa. shipper, 2-opt, google map API, giao hàng nhanh, bài toán người giao hàng, thuật toán heuristic.

USING THE HEURISTIC ALGORITHM TO OPTIMIZE THE DELIVERY PROBLEM

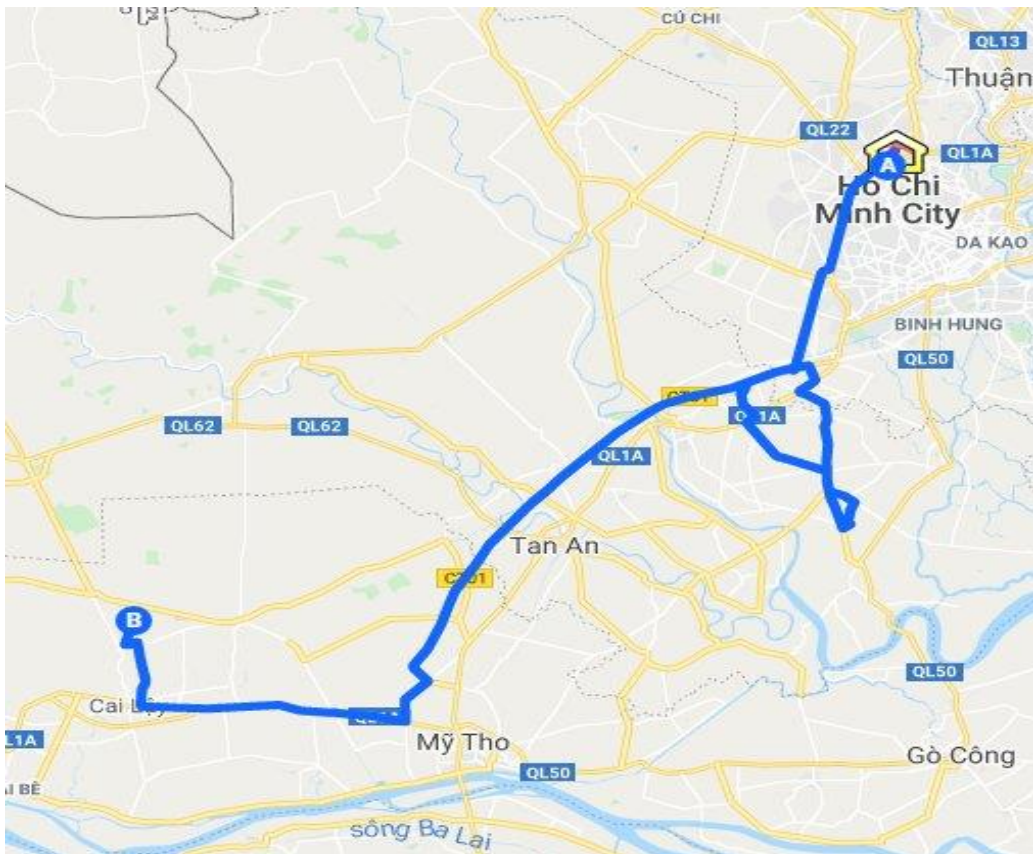
Abstract. Nowadays, whenever a shipper receives some orders, he must choose the appropriate routes to deliver as required. The popular used by the delivery shippers is based on the Openstreetmap API. However, it has not fully solved the problem of finding the way among multiple destinations minimizing the total travelling distance also limiting the repetition of the route traveled as much as possible. So our team studies the use of one heuristic algorithm, which is 2-OPT (2-Optional Practical Training) to find out the local optimal in real time. This research will focus on the routes of Ho Chi Minh City and visualized the data on Openstreetmap using the supported API.

Keywords. shipper, 2-opt, google map API, fast delivery, travelling salesman problem, heuristic algorithm.

1. ĐẶT VẤN ĐỀ

1.1. Giới thiệu

Tại [1], bản đồ trực tuyến đang được sử dụng ngày càng rộng rãi trong nhiều lĩnh vực, đặc biệt là trong việc tìm đường và đánh dấu các địa danh lên trên bản đồ. Nó không chỉ được sử dụng nhiều trong các giao diện website mà còn được dùng nhiều trong các ứng dụng điện thoại. Người dùng không chỉ dễ dàng tìm kiếm địa chỉ và định vị được vị trí mà còn có thể tìm kiếm được đường đi nhanh chóng từ nơi họ đứng đến địa điểm cần tìm một cách nhanh chóng, chính xác.



Hình 1.1 - Ứng dụng bản đồ trực tuyến dùng tìm đường

Theo [2], Google Map chính là 1 trong số nhiều ứng dụng được sử dụng phổ biến nhất nhưng bất tiện của Google Map hiện tại là chỉ cung cấp hai địa điểm đi và đến, không thể mở rộng lộ trình quá 10 địa điểm, nếu muốn phải thao tác thủ công nhiều lần, đó là chưa kể khi người dùng muốn được gợi ý lộ trình ngắn nhất. Việc mở rộng lộ trình qua nhiều địa điểm được ứng dụng rất nhiều trong thực tế chẳng hạn như: Busmap, giao hàng, du lịch,... những mảng liên quan đến việc di chuyển qua nhiều địa điểm; việc cải thiện thêm tính năng này sẽ tối ưu hơn rất nhiều, rút ngắn thời gian di chuyển và chi phí phát sinh. Trong thực tế sự phát triển mạnh mẽ của thương mại điện tử, mua sắm online dẫn đến ngành giao hàng cũng phải đẩy mạnh để đuổi kịp nhu cầu của khách hàng, để tăng hiệu quả giao hàng nhanh chóng đòi hỏi các công ty vận chuyển phải tích hợp thêm công nghệ mới trong việc giao hàng. Các shipper thường phải giao hàng cho khách ở các địa điểm khác nhau tại 1 thành phố có thể nhận nhiều đơn giao hàng và các đơn hàng này cũng cần được giao cho khách tại các đích đến khác nhau. Vì vậy mà việc giao hàng trở nên mất thời gian khi các địa điểm nhận hàng có khi trùng lặp lại đoạn đường đi của shipper, vừa mất thời gian vừa mất thêm chi phí vận chuyển. Nhu cầu tìm kiếm đường đi phù hợp được ưu tiên hàng đầu, cần phải đảm bảo đơn hàng giao đến các địa điểm phù hợp với thứ tự hợp lý sao cho đoạn đường đi ngắn nhất, nhanh nhất có thể [4].

Trên hành trình giao hàng, công nghệ cũng giúp rút ngắn thời gian di chuyển cho shipper và tiết kiệm chi phí vận chuyển cho người mua qua những giải pháp về AI (trí thông minh nhân tạo), Machine Learning (học máy) hay bằng Google Map API được nhiều đơn vị vận chuyển áp dụng như Viettel Post, Giao Hàng Nhanh, Giao hàng tiết kiệm,... đã thiết lập nên những cung đường tối ưu và ngắn nhất, giúp người giao hàng nhanh chóng giao hàng tới tay người nhận ở nhiều điểm khác nhau. Sự phát triển ngành giao hàng đã dẫn tới ngày càng có nhiều shipper giao hàng, để tăng hiệu quả giao hàng nhanh chóng, đội ngũ chuyên dụng đang được phát triển ngày càng mạnh. Trong thực tế các shipper này thường phải giao hàng cho khách tại nhiều địa điểm khác nhau, tại một thành phố có thể nhận nhiều đơn giao hàng và các đơn hàng này cũng cần phải được giao cho khách tại các đích đến khác nhau.

Do đó, việc các shipper tìm kiếm đường đi phù hợp gặp nhiều khó khăn do cần phải đảm bảo giao được hàng khi đi qua các địa điểm cần giao hàng phù hợp, cùng với thứ tự hợp lý để đảm bảo giao hàng nhanh

nhất có thể. Với sự phát triển của khoa học, công nghệ thông tin, việc tính toán các đường đi giao hàng được tin học hóa, nhằm rút ngắn thời gian và đảm bảo độ chính xác trong cung đường di chuyển. Vấn đề tính toán tự động các phương án đường đi giao hàng đã được nghiên cứu bởi nhiều tác giả. Ở Việt Nam, ngành giao hàng ngày càng phát triển, đặt ra yêu cầu phải có khả năng nghiên cứu, xây dựng các phần mềm hỗ trợ tính toán đường đi cho các shipper giao hàng. Tuy nhiên, thực tế lĩnh vực này chưa được quan tâm nhiều... Để góp phần giải quyết vấn đề trên nhóm nghiên cứu đề xuất sử dụng thuật toán heuristic là 2-opt để tối ưu bài toán giao hàng [4].

1.2. Phát biểu bài toán

Bài toán được phát biểu như sau: “*Hãy tìm một hành trình cho một người giao hàng đi qua n điểm khác nhau, mỗi điểm đi qua một lần sao cho tổng quãng đường cần đi là ngắn nhất. Giả sử rằng có con đường nối trực tiếp từ giữa hai điểm bất kỳ và các thông tin về khoảng cách đó được cho trước*”.

Các địa điểm trên có thể coi là đại lý, cửa hàng mà người giao hàng cần tiếp cận. Đây còn được gọi là bài toán người giao hàng – Traveling salesman problem (TSP) rất nổi tiếng trong lĩnh vực Khoa học máy tính. Cho đến thời điểm hiện tại, việc giải quyết triệt để bài toán này trong thời gian đa thức vẫn còn là vấn đề mở (vấn đề P-NP).

Về việc trực quan các dữ liệu, hiện nay với hàng loạt các bản đồ trực tuyến ra đời mặc dù vị trí chưa được ổn định nhưng với sự tiện lợi và đáp ứng các yêu cầu thì nhóm nghiên cứu vẫn chọn OpenStreetMap API để trực quan hóa thuật toán vì dữ liệu ở đây là mở và vẫn đáp ứng được các yêu cầu nghiên cứu. Cấu trúc và nguyên lý hoạt động của giải pháp Heuristic được đề xuất giới thiệu trong phần 2. Trong khi đó, những kết quả mô phỏng và thảo luận được trình bày trong phần 3. Và cuối cùng, những kết luận của bài viết được trình bày trong phần 4.

2. THUẬT TOÁN HEURISTIC

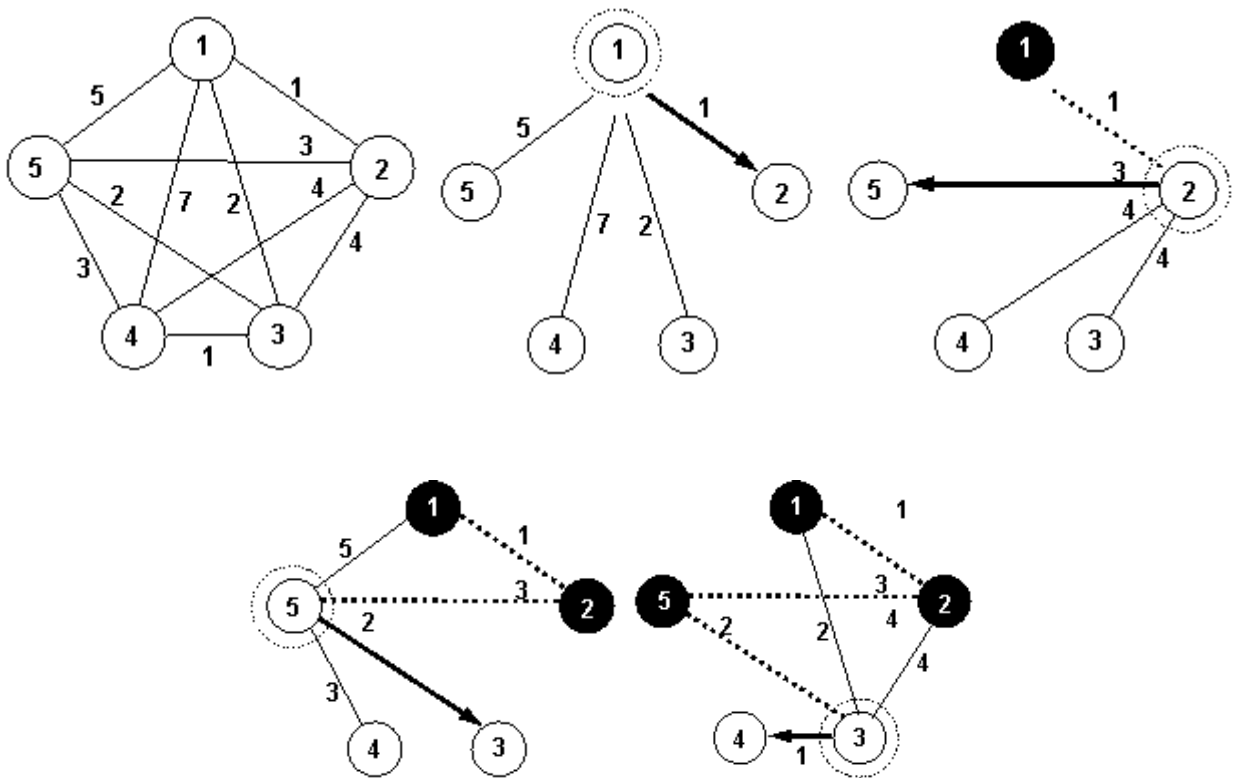
2.1. Bài toán hành trình ngắn nhất - ứng dụng nguyên lý Greedy

Tất nhiên ta có thể giải bài toán nêu trên, bằng cách liệt kê tất cả con đường có thể đi, tính chiều dài của mỗi con đường đó rồi tìm con đường có chiều dài ngắn nhất. Tuy nhiên, cách giải này lại có độ phức tạp $O(n!)$, cụ thể là một hành trình là một hoán vị của n điểm, do đó, tổng số hành trình là số lượng hoán vị của một tập n phần tử là $n!$. Do đó, khi số đại lý tăng thì số con đường phải xét sẽ tăng lên rất nhanh.

Một cách giải đơn giản hơn nhiều và thường cho kết quả tương đối tốt là dùng một thuật giải Heuristic ứng dụng nguyên lý Greedy [5]. Tư tưởng của thuật giải như sau:

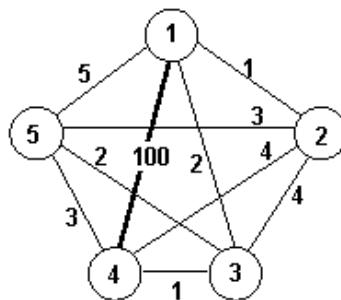
- Từ điểm khởi đầu, ta liệt kê tất cả quãng đường từ điểm xuất phát cho đến n đại lý rồi chọn đi theo con đường ngắn nhất.
- Khi đã đi đến một đại lý, chọn đi đến đại lý kế tiếp cũng theo nguyên tắc trên. Nghĩa là liệt kê tất cả con đường từ đại lý ta đang đứng đến những đại lý chưa đi đến. Chọn con đường ngắn nhất. Lặp lại quá trình này cho đến lúc không còn đại lý nào để đi.

Bạn có thể quan sát hình sau để thấy được quá trình chọn lựa. Theo nguyên lý Greedy, ta lấy tiêu chuẩn hành trình ngắn nhất của bài toán làm tiêu chuẩn cho chọn lựa cục bộ. Ta hy vọng rằng, khi đi trên n đoạn đường ngắn nhất thì cuối cùng ta sẽ có một hành trình ngắn nhất. Điều này không phải lúc nào cũng đúng. Với điều kiện trong hình tiếp theo thì thuật giải cho chúng ta một hành trình có chiều dài là 14 trong khi hành trình tối ưu là 13. Kết quả của thuật giải Heuristic trong trường hợp này chỉ lệch 1 đơn vị so với kết quả tối ưu. Trong khi đó, độ phức tạp của thuật giải Heuristic này chỉ là $O(n^2)$ [6].



Hình 2.1 - Giải bài toán sử dụng nguyên lý Greedy

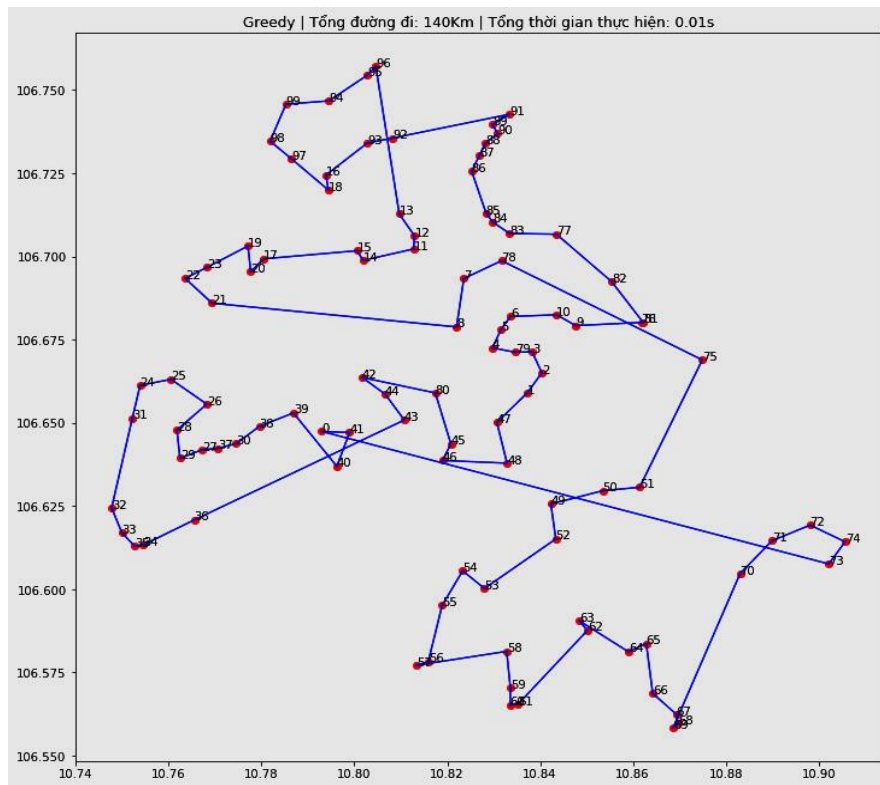
Tất nhiên, thuật giải theo kiểu Heuristic đôi lúc lại đưa ra kết quả không tốt, thậm chí rất tệ như trường hợp ở hình sau [6].



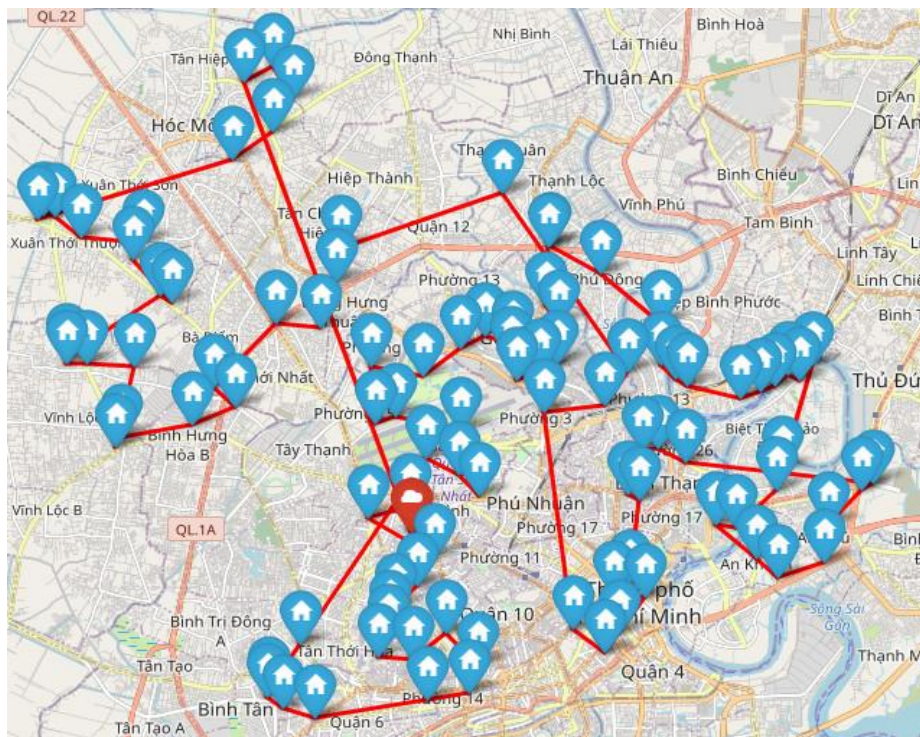
Hình 2.2 – Trường hợp thuật toán không hiệu quả

2.2. Thuật toán Greedy algorithm

Cùng thử nghiệm với 100 điểm như thuật toán 2-opt ta thu được kết quả khi áp dụng thuật toán Greedy algorithm để giải bài toán giao hàng:



Hình 2.3 - Mặt phẳng tọa độ Oxy khi biểu diễn thuật toán Greedy algorithm ở 100 địa điểm khác nhau tại Thành phố Hồ Chí Minh



Hình 2.4 – Bản đồ hình chim bay khi thực hiện thuật toán Greedy algorithm ở 100 địa điểm khác nhau tại Thành phố Hồ Chí Minh

3. THUẬT TOÁN TỐI ƯU 2-OPT

3.1. Giới thiệu

Theo [7], thuật toán 2-opt là một thuật toán tối ưu đơn giản, dựa vào phương pháp tìm kiếm cục bộ (Local Search), được Croes giới thiệu từ năm 1958 và được phát triển và ứng dụng hiệu quả trong nhiều lĩnh vực khác nhau. TSP là một trong những vấn đề được nghiên cứu nhiều nhất trong tổ hợp tối ưu hóa. Trong bài toán traveling salesman problem (TSP), chúng ta được đưa ra n thành phố với khoảng cách theo cặp của chúng. Nhiệm vụ là tìm một chuyến đi ngắn nhất đi qua mỗi thành phố chính xác một lần [8].

Công thức cơ bản của thuật toán 2-opt được thể hiện như sau:

Cho đồ thị n đỉnh đầy đủ cả có trọng số $G=(V\text{-tập đỉnh}, E\text{-tập cạnh})$ có hoặc vô hướng. Ta cần tìm chu trình Halmilton (đi qua tất cả các đỉnh, mỗi đỉnh một lần):

$$v_1 \rightarrow v_2 \rightarrow \dots \rightarrow v_n \rightarrow v_1 \text{ với } v_i \in V, v_i \neq v_j \text{ và } i, j \in \{1, 2, \dots, n\}. \text{ có tổng trọng số}$$

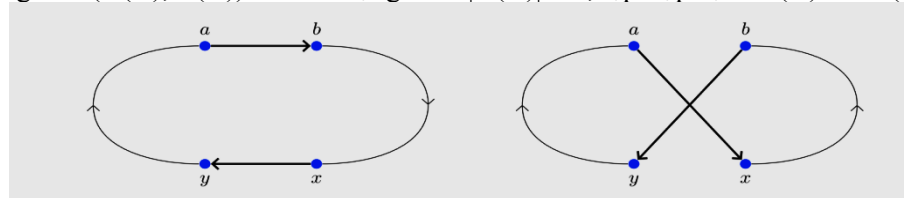
$$\sum_{i=1}^n c_{i,i+1} \text{ với } c_{i,j} \text{ là trọng số của cạnh } (v_i, v_j), \text{ quy ước } i+1 \equiv 1 \text{ là nhỏ nhất [8].}$$

Tuy nhiên, phương pháp 2-opt cũng có hạn chế như các phương pháp khác, do khả năng bài toán local-search bị tắc tại các giá trị tối ưu cục bộ. Để giải quyết vấn đề này, trong phạm vi đề tài, thuật toán 2-opt được sử dụng kết hợp với phương pháp “Bây”, tức là dựa trên một tập hợp các phương án ban đầu tối ưu hóa riêng rẽ các phương án này để được kết quả sau cùng tối ưu.

3.2. Ý tưởng bài toán 2-OPT

Bắt đầu với một điểm tùy ý trong chuyến đi, phương pháp heuristic 2-Opt liên tục thay thế hai cạnh của đường đi bằng hai cạnh khác, miễn là điều này mang lại một chuyến đường đi ngắn hơn. 2-Opt dừng khi không thể cải thiện thêm đường đi trong chuyến đi. Một chuyến đi mà phương pháp heuristic 2-Opt không thể cải thiện được thêm nữa gọi là 2-optimal. Thuật toán 2-Opt liên tục thay thế hai cạnh từ chuyến đi bằng hai cạnh khác để chuyến đi đưa ra kết quả ngắn hơn. Cụ thể hơn như sau:

Cho đồ thị vô hướng $G = (V(G), E(G))$ với số lượng đỉnh $|V(G)| = n$, tập hợp cạnh $E(G)$ chứa $(nC2)$ cạnh.



Hình 3.1 - Thao tác biến đổi chính của thuật toán

Đồ thị bên trái (left) đã tạo ra 2 cạnh mới (a,b) và (x,y) để thay thế cho 2 cạnh (a,x) và (b,y), từ đó tạo ra chuyến đi mới với chi phí tốt hơn có thể trước đó [9].

Cho một chuyến đi T và hai cạnh $\{a, b\}$ và $\{x, y\}$ trong T , có hai khả năng thay hai cạnh này bằng hai cạnh khác. Hoặc chúng ta có thể chọn cặp $\{a, x\}$ và $\{b, y\}$ hoặc chúng ta có thể chọn cặp $\{a, y\}$ và $\{b, x\}$. Chính xác thì một trong hai cặp này sẽ cho kết quả là một chuyến đi khác một lần nữa. Nếu không biết các cạnh khác của T , chúng ta không thể đưa ra quyết định khả năng nào trong hai khả năng mà chúng ta phải chọn. Do đó, chúng ta sẽ giả định rằng chuyến đi T là một chu kỳ định hướng, tức là các cạnh của T có hướng sao cho mỗi đỉnh có đúng một cạnh tới và một cạnh đi [9].

Sử dụng quy ước này, chỉ có một khả năng để trao đổi một cặp cạnh sao cho tập hợp cạnh mới lại là một chuyến đi: các cạnh có hướng (a, b) và (x, y) phải được thay thế bằng các cạnh (a, x) và (b, y).

Lưu ý rằng để có được một chu trình có định hướng một lần nữa, một phải đảo ngược hướng của đoạn giữa b và x. Nếu (a, b) và (x, y) là hai cạnh trong một chuyến đi T và chúng ta có $c(a, x) + c(b, y) < c(a, b) + c$

(x, y) thì chúng ta nói rằng việc thay thế các cạnh (a, b) và (x, y) trong T bằng các cạnh (a, x) và (b, y) là một thay đổi cải tiến. Do đó, 2-Opt có thể được xây dựng như sau:

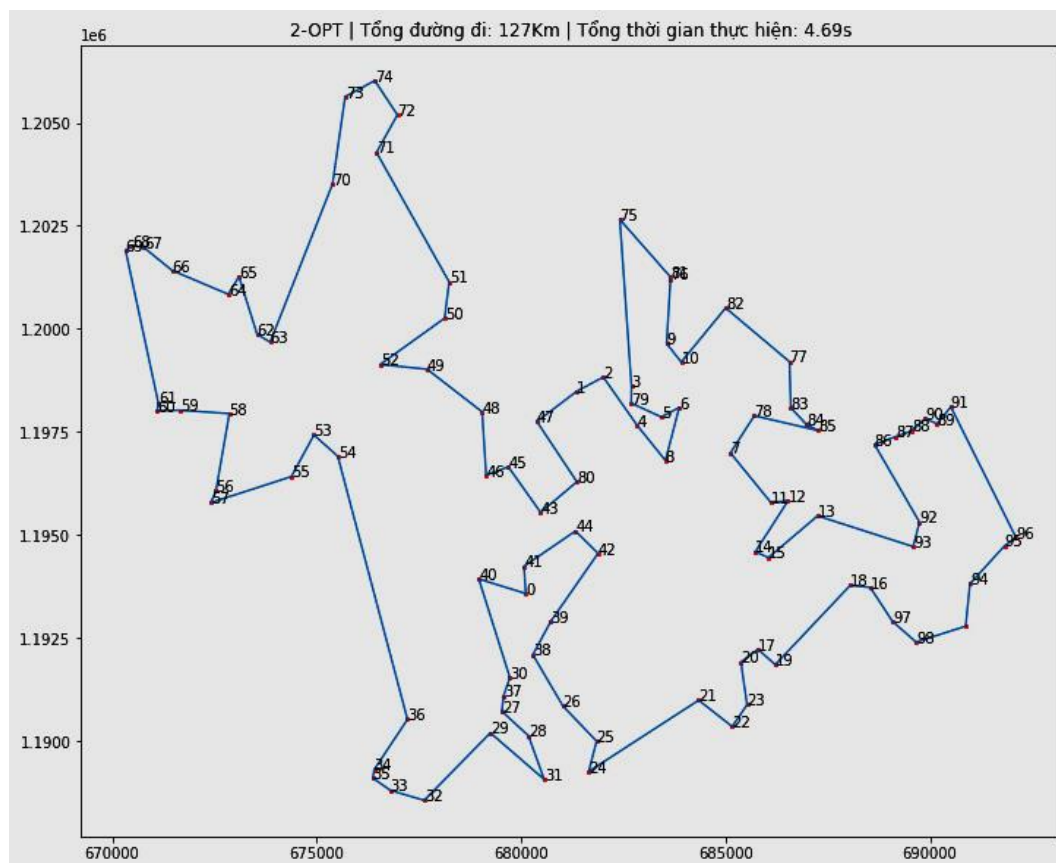
- Bắt đầu với 1 điểm bất kỳ trong T
- Trong khi tồn tại 2 cạnh có thể cải thiện trong T, thực hiện thay đổi 2 cạnh
- Đưa ra chuyến đi T

Mã giả: #Hàm thay đổi chuyến đi:

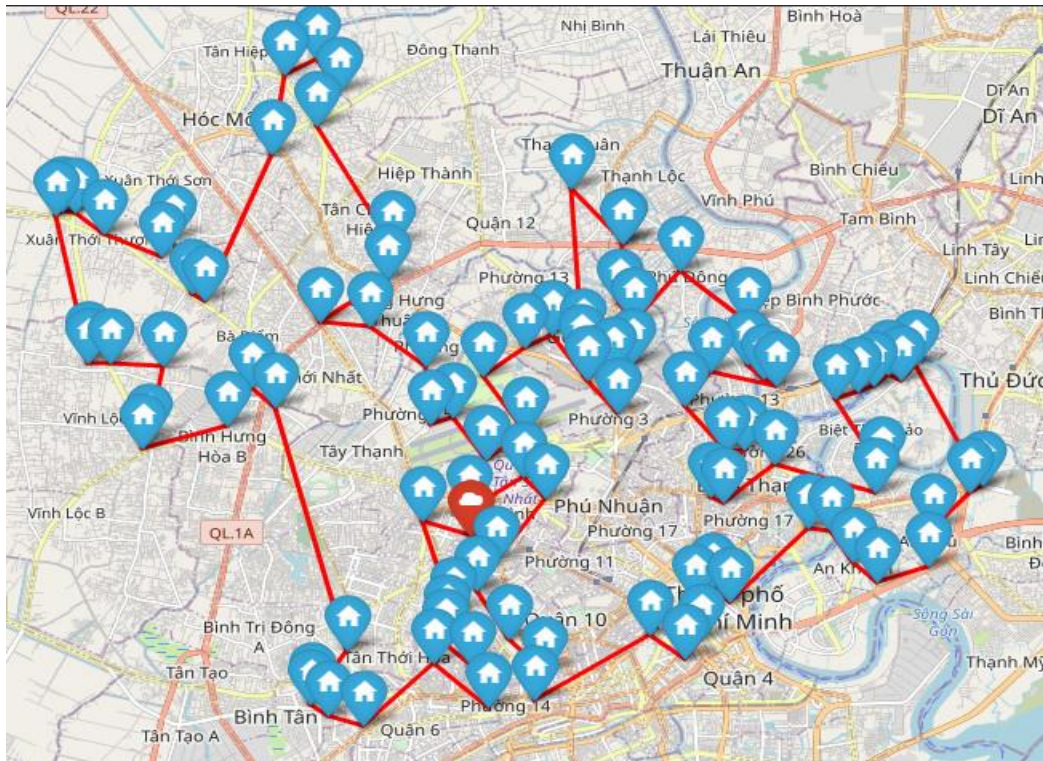
```
def 2optSwap (route, i, k)
    1. take route[0] to route[i-1] and add them in order to new_route
    2. take route[i] to route[k] and add them in reverse order to new_route
    3. take route[k+1] to end and add them in order to new_route
    return new_route;
```

Chú ý rằng thao tác trên sẽ được lặp lại nhiều lần đến khi nào thu được một quãng đường với độ dài chấp nhận được thì thôi. Do đó, trên thực tế, thuật toán sẽ chọn một tham số đầu vào là số lần refinement và sẽ cân nhắc giữa 2 yếu tố: độ tốt của lời giải và thời gian thực thi thuật toán.

3.3. Mô tả



Hình 3.2 - Mặt phẳng tọa độ Oxy khi biểu diễn thuật toán 2-opt ở 100 địa điểm khác nhau tại TP. Hồ Chí Minh



Hình 3.3 - Bản đồ hình chim bay khi biểu diễn thuật toán 2-opt ở 100 địa điểm khác nhau tại Thành phố Hồ Chí Minh

Bên dưới, ta có bảng so sánh đặc điểm đặc trưng giữa 2 thuật toán:

Bảng so sánh tính tối ưu giữa 2-OPT và Greedy Algorithm dựa trên kết quả kiểm thử

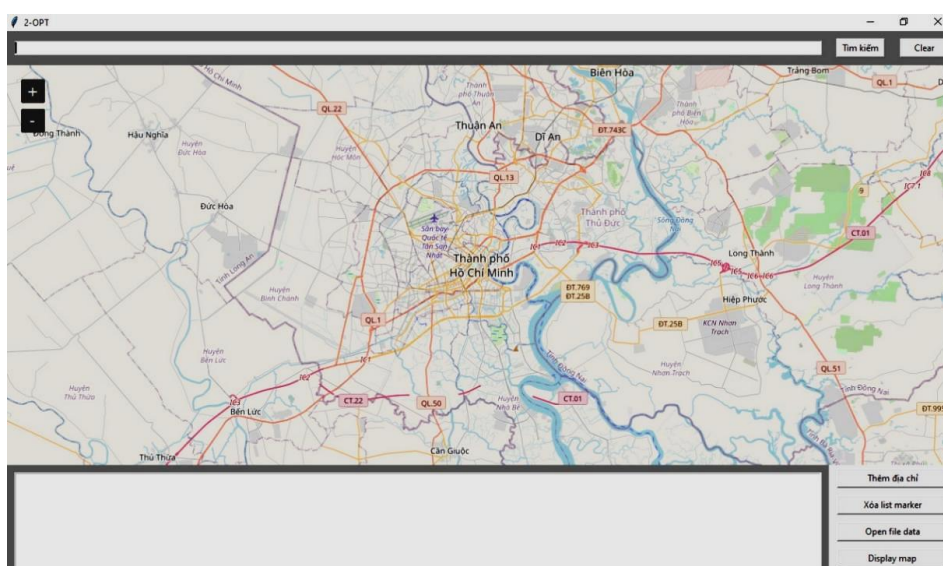
Yếu tố	2-OPT Algorithm	Greedy Algorithm
Mô tả thuật toán	Dựa trên một phương án có sẵn, tiến hành tìm các vị trí chưa tối ưu và tiến hành thay đổi đường đi để thu được cách tốt hơn	Tìm kiếm những điểm xung quanh gần nhất để di chuyển
Tổng quãng đường	Ngắn hơn	Dài hơn
Thời gian thực hiện chương trình	Dài hơn	Ngắn hơn
Độ phức tạp thuật toán	$O(k.n^2)$ với k là số lần refinement	$O(n^2)$

4. KẾT LUẬN

Qua quá trình nghiên cứu đề tài “Sử dụng thuật toán heuristic để tối ưu bài toán giao hàng”, nhóm tác giả đã đạt được một số kết quả sau (code sản phẩm nghiên cứu):

Github: <https://github.com/antphb/HNKHT>

Link sản phẩm: https://terabox.com/s/18KuyYenHWS6_wEYFvaw0Q



Hình 4.1. Giao diện sản phẩm của nhóm đã phát triển được

Đồng thời so sánh hai thuật toán đó là thuật toán 2-OPT và thuật toán tham lam, từ kết quả so sánh nhóm tác giả chọn thuật toán 2-OPT để giải quyết bài toán tìm đường đi ngắn nhất dành cho người giao hàng. Kết quả mô phỏng cho thấy khi sử dụng thuật toán heuristic là 2-opt để tối ưu bài toán thì quãng đường thực hiện và thời gian sẽ ngắn hơn. Cuối cùng, nhóm tác giả đã nghiên cứu kiến thức cơ bản về ứng dụng và đã tạo ra một ứng dụng chạy trên nền tảng phi trung lập - đó là ứng dụng tìm đường đi ngắn nhất. Một thư viện quan trọng mà nhóm có sử dụng là GeoPy 2.2.0 [10]. Kết quả khích lệ trong nghiên cứu này là cơ sở cho các nghiên cứu sâu hơn trong bài toán giao hàng - một bài toán chưa được tối ưu toàn cục.

TÀI LIỆU THAM KHẢO

- [1] Ngọc Bách, Bản đồ trực tuyến: nhu cầu tất yếu của xã hội, Báo điện tử đại biểu nhân dân, 2011.
- [2] Google map, <https://www.google.com/maps/@9.779349,105.6189045,11z?hl=vi-VN>.
- [3] Nishant Singh, Traveling Salesman Problem (TSP) Implementation, website: www.geeksforgeeks.org, 2022.
- [4] Website northwestern.edu, *Heuristic algorithms*. [Online]. Available: https://optimization.mccormick.northwestern.edu/index.php/Heuristic_algorithms
- [5] Website programiz.com, *GreedyAlgorithm*. [Online]. Available: <https://www.programiz.com/dsa/greedy-algorithm>
- [6] Website tek4.vn, Thuật toán tham lam. [Online]. Available: <https://tek4.vn/khoa-hoc/cau-truc-du-lieu-va-giai-thuat/thuat-toan-tham-lam-greedy-algorithms>
- [7] Website wikipedia.org, *2-opt*. [Online]. Available: <https://en.wikipedia.org/wiki/2-opt>
- [8] Pedrohfsd, Optimization with 2-OPT - Part 1, 2017.
- [9] Account rrz0 of stackoverflow, 2-opt algorithm to solve the Travelling Salesman Problem, 2018. [Online]. Available: <https://stackoverflow.com/questions/53275314/2-opt-algorithm-to-solve-the-travelling-salesman-problem>
- [10] GeoPy 2.2.0 documentation, [Online]. Available : <https://geopy.readthedocs.io/en/stable/>

YSC4F.234

THUẬT TOÁN COOT CHO BÀI TOÁN XÁC ĐỊNH VỊ TRÍ VÀ CÔNG SUẤT NGUỒN ĐIỆN PHÂN TÁN TRÊN LƯỚI ĐIỆN PHÂN PHỐI

TRỊNH HỮU TRƯỜNG, VÕ TẤN LỘC, DƯƠNG THANH LONG, NGUYỄN THANH THUẬN

Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
trinhhuutruong@iuh.edu.vn, votanloc@iuh.edu.vn, duongthanhlolong@iuh.edu.vn,
nguyenthanhthuan@iuh.edu.vn

Tóm tắt. Bài báo này sử dụng thuật toán tối ưu Coot để tìm vị trí và công suất nguồn điện phân tán trên hệ thống điện phân phối. Mục tiêu của bài toán xem xét là giảm tổn thất công suất có xem xét đến ràng buộc điện áp. Thuật toán tối ưu Coot là thuật toán tối ưu mới lấy ý tưởng về cách thức di chuyển của bầy chim Coot trên mặt nước. Kết quả kiểm tra thuật toán được thực hiện trên lưới điện phân phối 33 nút. Hiệu quả của thuật toán Coot cho bài toán xem xét được so sánh với thuật toán tối ưu bầy đàn. Kết quả tính toán cho thấy thuật toán Coot là một trong những thuật toán tiềm năng cho bài toán này.

Từ khóa. thuật toán tối ưu Coot, nguồn điện phân tán, thuật toán tối ưu bầy đàn, tổn thất công suất.

COOT ALGORITHM FOR DETERMINATION OF LOCATION AND CAPACITY OF DISTRIBUTED GENERATIONS ON THE DISTRIBUTION ELECTRIC NETWORK

Abstract. This paper presents the Coot optimization algorithm to determine position and capacity of the distributed generations on the distribution electric network. The considered objective is to reduce the power loss considering the voltage constraint. Coot algorithm is a recent algorithm with the idea of how Coot birds move on the water surface. The validated results of the algorithm are implemented on the 33-node distribution electric network. The efficiency of the Coot algorithm for the problem of location and capacity determination of distributed generations is compared to particle swarm optimization algorithm. The calculation results show that the Coot optimization algorithm is one of potential algorithms for this problem.

Keywords. Coot optimization algorithm, distributed power source, power loss, particle swarm optimization.

1. GIỚI THIỆU

Nguồn điện phân tán DG (Distributed Generation) là các nguồn điện nhỏ được kết nối với lưới điện phân phối hoặc gắn các phụ tải [1]. DG được chia làm hai nhóm khác nhau dựa vào nguồn năng lượng sơ cấp. Nhóm thứ nhất là nhóm năng lượng xanh gồm máy phát điện mặt trời, gió và thủy điện nhỏ, nhóm này không phát thải khí nhà kính. Với nhóm này việc lắp DG có lợi ích về môi trường. Nhóm thứ hai phát thải khí nhà kính gồm máy phát diesel, khí ...[2]. Lắp đặt DG trong lưới phân phối có nhiều lợi ích về kỹ thuật như giảm tổn thất công suất, cải thiện điện áp [3] việc này kéo theo những lợi ích về kinh tế. Tuy nhiên, nếu lắp đặt DG không phù hợp sẽ gây ảnh hưởng về tổn thất và chi phí lớn hơn [4]. Do đó, việc xác định vị trí và công suất DG là việc rất quan trọng.

Để giải bài toán tối ưu lắp đặt DG, các thuật toán tối ưu được phát triển gần đây thường được sử dụng. Trong đó có loại thuật toán meta-heuristics, loại thuật toán này là thuật toán tối ưu gần đúng với thời gian thích hợp và áp dụng với các vấn đề phức tạp [5]. Loại thuật toán này có thể được chia thành bốn nhóm [6]. Nhóm đầu tiên là nhóm thuật toán tiến hóa (Evolutionary Algorithms) được phát triển tương đối lâu, bao gồm thuật toán GA (Genetic Algorithm) [7] [8], thuật toán DE (Differential Evolution) [9], thuật toán SFS (Stochastic Fractal Search) [10]. Nhóm thứ hai là nhóm thuật toán dựa trên bầy đàn (Swarm Based Algorithms) được phát triển dựa vào cách thức di chuyển của các con vật khác nhau ví dụ như thuật toán

PSO (Particle Swarm Optimization) [11], thuật toán BA (Bat Algorithm) [12], thuật toán COA (Coyote Algorithm) [13], thuật toán ACSA (Adaptive Cuckoo Search) [14], thuật toán SSA (Salp Swarm Algorithm) [15], thuật toán TLO-GWO (Teaching Learning Optimization - Grey Wolf Optimizer) [16], thuật toán ALO (Ant Lion Optimization Algorithm) [4] [17], thuật toán FA (Firefly Algorithm) [18] [19], thuật toán CSA (Crow Search Algorithm) [20], thuật toán HBMO (Honey Bee Mating Optimization) [21], thuật toán WOA (Whale Optimization Algorithm) [22]. Nhóm thứ ba là nhóm thuật toán dựa trên hành vi của con người (Human-Based Algorithms), thuật toán MTLBO (Modified Teaching–Learning Based Optimization) [1] là một trong số đó. Nhóm cuối cùng là nhóm thuật toán dựa theo các hiện tượng vật lý chẳng hạn như thuật toán IWD (Intelligent Water Drop) [23]. Các thuật toán như trên giải quyết nhiều vấn đề trong lưới phân phối như cấu hình lại mạng phân phối [24], dò tìm công suất cực đại trong hệ thống pin năng lượng mặt trời [25]. Có thể thấy rằng, nhóm các thuật toán meta-heuristics đang phát triển và được ứng dụng rộng rãi vào các bài toán trong lưới phân phối.

Thuật toán tối ưu Coot (Coot algorithm) là một thuật toán mới được đề xuất năm 2021. Thuật toán mô phỏng cách thức di chuyển tìm thức ăn của đàn chim Coot. Trong đó có di chuyển không đồng đều và về sau là di chuyển đồng đều. Di chuyển của đàn chim còn phụ thuộc vào các con chim đầu đàn. Thuật toán đã giải quyết được các vấn đề trong lĩnh vực cơ khí và xây dựng [6]. Trong nghiên cứu này thuật toán Coot được sử dụng để tối ưu lắp đặt nguồn điện phân tán với mục tiêu để giảm thiểu tổn thất công suất và nâng cao điện áp áp của lưới điện.

2. MÔ HÌNH BÀI TOÁN XÁC ĐỊNH VỊ TRÍ VÀ CÔNG SUẤT DG

Trong bài báo hàm mục tiêu bài toán tối ưu lắp đặt nguồn điện phân tán là tối thiểu tổn thất công suất sau khi lắp nguồn phân tán:

$$\text{Min}(\Delta P) \text{ với } \Delta P = \sum_{k=1}^H \frac{P_k^2 + Q_k^2}{U_k^2} R_k \quad (1)$$

Trong đó, H là số đường dây của lưới điện. $P_k + jQ_k$ là công suất biểu kiến trên đường dây k có điện trở R_k . U_k là điện áp cuối đường dây k.

Điều kiện ràng buộc của bài toán xem xét là giới hạn điện áp các nút:

$$V_{\min} \leq V_j \leq V_{\max} \quad (2)$$

Trong đó, V_j là biên độ điện áp nút j. $[V_{\min}, V_{\max}]$ là giới hạn điện áp cho phép.

Hàm mục tiêu của bài toán (Fobj) được xác định bởi biểu thức (3):

$$\text{Fobj} = \Delta P + Z(\max(V_{\min} - V, 0) + \max(V - V_{\max}, 0)) \quad (3)$$

Với, Z là hệ số phạt.

3. ỨNG DỤNG THUẬT TOÁN COOT ĐỂ XÁC ĐỊNH VỊ TRÍ VÀ CÔNG SUẤT DG

Thuật toán này mô phỏng di chuyển thường xuyên và không đều của loài chim Coot. Đàn chim di chuyển theo những con chim đầu đàn để hướng về mục tiêu (thức ăn). Có bốn di chuyển của chim Coot trên mặt nước bao gồm di chuyển ngẫu nhiên, di chuyển dây chuyền, điều chỉnh vị trí theo các con chim đầu đàn và di chuyển theo các con chim đầu đàn hướng tới các khu vực tối ưu.

Bước 1: Khởi tạo

Khởi tạo ngẫu nhiên dựa theo biểu thức (4):

$$C(i) = \text{rand}(1, b)(u - d) + d \quad (4)$$

Trong đó, $C(i)$ là vị trí của cá thể. b là số lượng biến, với bài toán có m nguồn phân tán $b = 2m$. d là giới hạn dưới của không gian tìm kiếm. Tương tự, u là giới hạn trên của các biến.

Mỗi biến có giới hạn khác nhau được xác định theo biểu thức (5):

$$u = [u_1, u_2, \dots, u_b], d = [d_1, d_2, \dots, d_b] \quad (5)$$

Trong đó, d_1 đến d_m là giới hạn dưới của nút nên giới hạn này bằng 2 (do nút 1 là nút select). d_{m+1} đến d_{2m} là giới hạn dưới của công suất nguồn phân tán và bằng 0. u_1 đến u_m là giới hạn trên của nút. u_{m+1} đến u_{2m} là giới hạn trên của công suất nguồn phân tán.

Trong quần thể vừa tạo, chọn ngẫu nhiên N_L cá thể đầu đàn.

Bước 2: Tính toán hàm mục tiêu

Giá trị hàm mục tiêu của mỗi cá thể được tính toán theo biểu thức (3).

Bước 3: Tính toán vị trí mới

Mỗi cá thể được cập nhật theo ba cách bao gồm di chuyển ngẫu nhiên, di chuyển dây chuyền và di chuyển theo các cá thể đầu đàn với xác suất cập nhật của mỗi cách lần lượt là 50%, 25% và 25%.

Di chuyển ngẫu nhiên: Di chuyển này được thực hiện ngẫu nhiên được mô tả ở biểu thức (6):

$$Q = rand(1, d)(u - d) + d \quad (6)$$

Di chuyển này của cá thể khám phá các phần khác nhau của không gian tìm kiếm. Nếu bị mắc kẹt trong tối ưu cục bộ, di chuyển này của cá thể sẽ giúp thuật toán thoát khỏi tối ưu cục bộ. Vị trí mới theo biểu thức (7):

$$C(i) = C(i) + AR_2Q - C(i) \quad (7)$$

Trong đó, R_2 là số ngẫu nhiên có giá trị trong khoảng $[0,1]$. A được tính theo biểu thức (8):

$$A = 1 - \frac{it}{Iter} \quad (8)$$

Với, it là số lần lặp hiện tại. $Iter$ là số lần lặp tối đa.

Di chuyển dây chuyền: Vị trí trung bình của hai cá thể được xác định theo di chuyển dây chuyền được mô tả toán học như sau [26]:

$$C(i) = 0.5(C(i - 1) + C(i)) \quad (9)$$

Trong đó, $C(i - 1)$ là vị trí cá thể trước nó.

Di chuyển theo các cá thể đầu đàn: Các cá thể di chuyển theo các cá thể đầu đàn. Mỗi cá thể phải chọn cho mình một cá thể đầu đàn để điều chỉnh vị trí của mình. Vị trí trung bình của các cá thể đầu đàn được tính toán và các cá thể Coot còn lại cập nhật vị trí theo vị trí trung bình. Biểu thức (10) mô tả phương pháp lựa chọn cá thể đầu đàn:

$$k = 1 + mod(i, N_L) \quad (10)$$

Trong đó, N_L là số lượng cá thể đầu đàn. k là số chỉ số của cá thể đầu đàn. mod là hàm trả về số dư của phép chia i và N_L .

$C(i)$ được cập nhật vị trí theo cá thể đầu đàn k theo biểu thức (11).

$$C(i) = L(k) + 2R_1cos(2R\pi) \times (L(k) - C(i)) \quad (11)$$

Trong đó, $L(k)$ là cá thể đầu đàn được chọn. R_1 là số ngẫu nhiên trong khoảng $[0,1]$. R là số ngẫu nhiên trong khoảng $[-1,1]$.

Cá thể đầu đàn trong quần thể được cập nhật vị trí theo cách sau: Các cá thể di chuyển về mục tiêu (khu vực tối ưu), vậy nên các cá thể đầu đàn cập nhật vị trí của chúng đến mục tiêu. Biểu thức (12) được sử dụng để cập nhật các cá thể đầu đàn:

$$L(i) = \begin{cases} BR_3 \cos(2R\pi) (G_{best} - L(i)) + G_{best} & R_4 < 0.5 \\ BR_3 \cos(2R\pi) (G_{best} - L(i)) - G_{best} & R_4 \geq 0.5 \end{cases} \quad (12)$$

Trong đó, G_{best} là vị trí tốt nhất được tìm thấy. R_3, R_4 là số ngẫu nhiên trong khoảng $[0,1]$. R là số ngẫu nhiên trong khoảng $[-1,1]$. B được tính theo biểu thức (13):

$$B = 2 - \frac{it}{Iter} \quad (13)$$

Bước 4: Kiểm tra giới hạn các biến điều khiển

Giới hạn nút và giới hạn công suất nguồn phân tán được điều chỉnh về các giới hạn cho phép bằng cách sử dụng các biểu thức (14) đến (16):

$$T_p = C(i) > u \quad (14)$$

$$T_m = C(i) < d \quad (15)$$

$$C(i) = C(i)(\sim(T_p + T_m)) + uT_p + dT_m \quad (16)$$

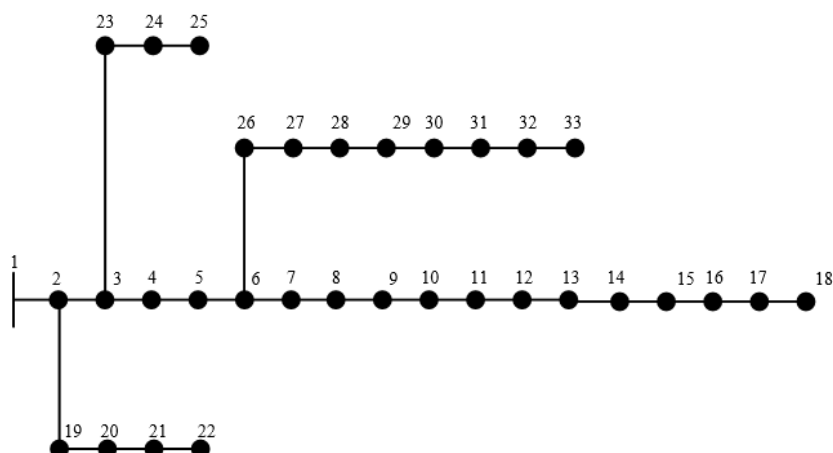
Bước 5: Kiểm tra điều kiện dừng tìm kiếm

Nếu $it < Iter$, tăng số vòng lặp $it = it + 1$ và quay lại bước 2 để tiếp tục thực hiện, ngược lại thì dừng thuật toán và xuất kết quả tìm kiếm. Khi đó G_{best} là kết quả của bài toán.

4. KẾT QUẢ TÍNH TOÁN

Trong mục này, thuật toán tối ưu Coot được sử dụng để tối ưu lắp đặt DG cho lưới điện 33 nút. Sơ đồ lưới điện được trình bày trong Hình 1. Hiệu quả của thuật toán Coot được so sánh với thuật toán PSO. Hai thuật toán được lập trình trong môi trường Matlab và cấu hình máy tính là CPU 1.6 GHz và RAM 12 GB. Số nguồn phân tán trong lưới được giới hạn bằng ba. Hệ số phạt được đặt bằng 1000, quy mô dân số là 30 còn

số vòng lặp là 500. Để so sánh hai thuật toán, cả Coot và PSO được chạy 50 lần độc lập.

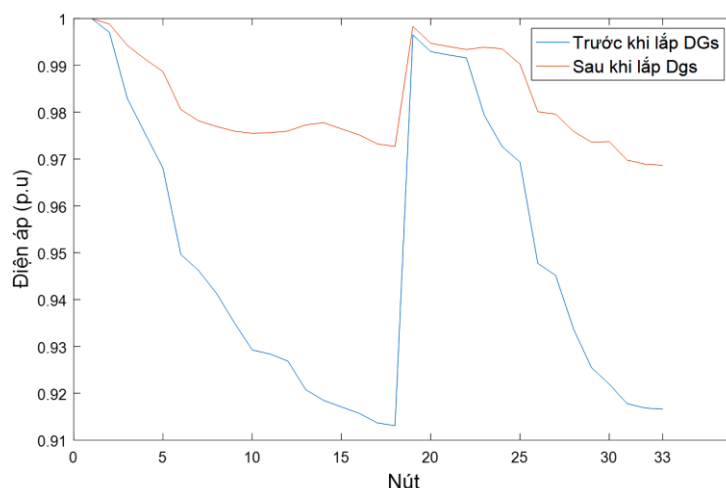


Hình 1. Sơ đồ lưới phân phối 33 nút

Tổng thất công suất ở lưới 33 nút ban đầu là 202.6863 kW và biên độ điện áp nhỏ nhất trong hệ đơn vị tương đối là 0.9131. Kết quả tối ưu vị trí và công suất của DG được trình bày ở Bảng 1. Đặt ba DG có công suất 1.07144, 0.753967 và 1.09944 MW tại các nút 30, 14 và 24, tổng thất công suất giảm từ 202.6863 kW xuống còn 71.45991 kW. Biên độ điện áp nhỏ nhất theo hệ đơn vị tương đối tăng từ 0.9131 lên 0.96865. Theo Hình 2, điện áp của hệ thống được cải thiện đáng kể. Điện áp tại các nút sau khi lắp đặt DG trong hệ đơn vị tương đối tại các nút đều tăng lên so với trước khi lắp đặt DG.

Bảng 1. Kết quả lắp đặt DG tối ưu cho mạng 33 nút.

Mục	Chưa lắp DG	Coot	PSO
Công suất DG (MW) [nút lắp đặt]	-	1.07144 (30) 0.753967 (14) 1.09944 (24)	0.754 (14) 1.0994 (24) 1.0714 (30)
Tổng công suất DG	-	2.9248	2.9248
ΔP (kW)	202.6863	71.45991	71.4599
Vmin (p.u)	0.9131	0.96865	0.9687



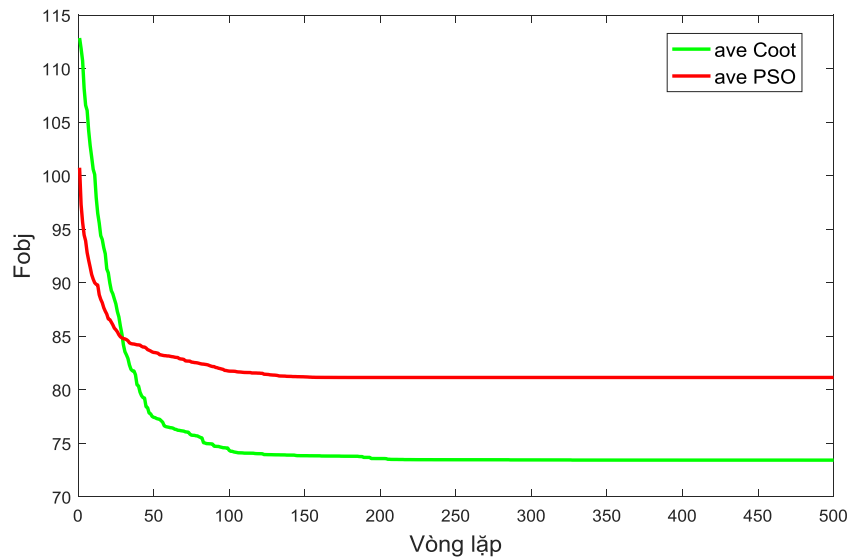
Hình 2. Điện áp trước và sau khi lắp DG

Kết quả so sánh giữa phương pháp dựa trên thuật toán Coot và PSO ở Bảng 2 cho thấy giá trị tối ưu thu được của hai phương pháp là giống nhau. Tuy nhiên giá trị lớn nhất (Max), trung bình (Ave) và độ lệch chuẩn (Std) hàm mục tiêu của Coot tốt hơn so với PSO. Về mặt thời gian tính toán, thời gian chạy của Coot là dài hơn so với PSO.

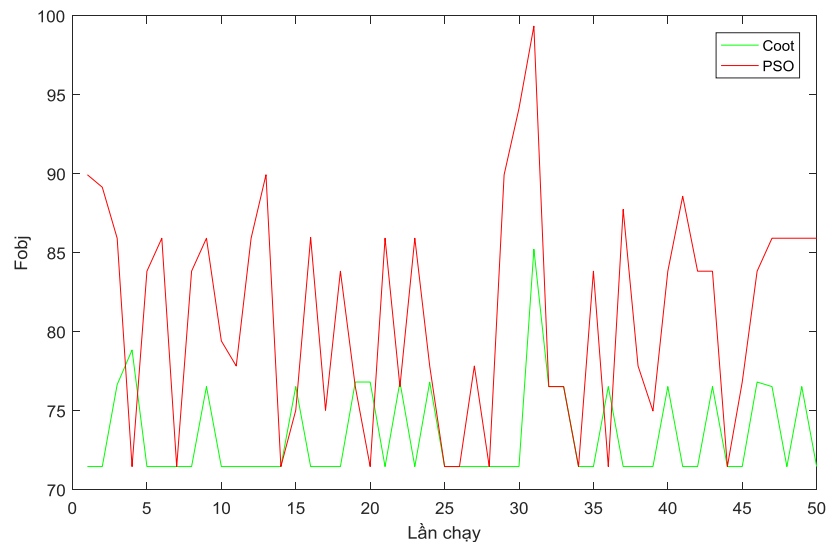
Bảng 2. So sánh kết quả giữa Coot và PS33

Mục	Max	Min	Ave	Std	Time (s)
Coot	85.2143	71.45991	73.4325	3.0295	3384.708021
PSO	99.3375	71.4599	81.1529	7.0046	2237.157258

Giá trị trung bình trong 50 lần chạy của thuật toán Coot và PSO được trình bày trong Hình 3. Kết quả cho thấy trong mỗi vòng lặp, giá trị hội tụ của thuật toán Coot tốt hơn so với PSO. Giá trị tối ưu mỗi lần chạy được thể hiện ở Hình 4. Giá trị tối ưu của PSO giao động mạnh hơn và thường cao hơn so với Coot. Kết quả này cho thấy hiệu quả cao hơn của thuật toán Coot so với PSO cho bài toán tối ưu lắp đặt DG.



Hình 3. Giá trị trung bình của Coot và PSO trong mỗi vòng lặp



Hình 4. Giá trị tối ưu trong mỗi lần chạy

Kết quả so sánh với các phương pháp trước đây được trình bày ở Bảng 3. Tổng thất công suất thu được bởi thuật toán Coot nhỏ hơn HSA, FWA và ACSA. So với SSA, kết quả thu được bởi thuật toán Coot chỉ cao hơn 0.00991 so với SSA.

Bảng 3. So sánh hiệu quả giữa thuật toán Coot với một số thuật toán khác.

Mục	Coot	HSA [27]	FWA [28]	SSA [15]	ACSA [14]
Công suất DG (MW) (nút lắp đặt)	1.07144 (30)	0.1070 (18)	0.5897 (14)	0.7536 (33)	0.7798 (14)
	0.753967 (14)	0.5724(17)	0.1895 (18)	1.1004 (23)	1.1251 (24)
	1.09944 (24)	1.0426 (33)	1.0146 (32)	1.0706 (29)	1.3496 (30)
Tổng công suất DG	2.9248	1.7256	1.7936	2.9246	3.2545
ΔP (kW)	71.45991	96.76	88.68	71.45	74.26
Vmin (p.u)	0.96865	0.9670	0.9680	0.9686	0.9778

5. KẾT LUẬN

Bài báo đã đề xuất phương pháp tối ưu lắp đặt DG dựa trên thuật toán Coot. Hàm mục tiêu xem xét là giảm tổn thất công suất và cải thiện điện áp của lưới điện. Thuật toán được áp dụng cho lưới phân phối 33 nút. Kết quả đối sánh với thuật toán PSO cho thấy phương pháp đề xuất Coot đạt hiệu quả cao hơn so với PSO. Kết quả đối sánh với các nghiên cứu khác cũng cho thấy hiệu quả của phương pháp dựa trên thuật toán Coot cho bài toán tìm kiếm vị trí lắp đặt và công suất tối ưu của DG trên lưới điện phân phối.

REFERENCES

- [1] J. A. Martín García and A. J. Gil Mena, “Optimal distributed generation location and size using a modified teaching-learning based optimization algorithm,” *Int. J. Electr. Power Energy Syst.*, vol. 50, no. 1, pp. 65–75, 2013, doi: 10.1016/j.ijepes.2013.02.023.
- [2] M. Pesaran H.A., M. Nazari-Heris, B. Mohammadi-Ivatloo, and H. Seyedi, “A hybrid genetic particle swarm optimization for distributed generation allocation in power distribution networks,” *Energy*, vol. 209, p. 118218, 2020, doi: 10.1016/j.energy.2020.118218.
- [3] S. Abdi and K. Afshar, “Application of IPSO-Monte Carlo for optimal distributed generation allocation and sizing,” *Int. J. Electr. Power Energy Syst.*, vol. 44, no. 1, pp. 786–797, 2013, doi: 10.1016/j.ijepes.2012.08.006.
- [4] E. S. Ali, S. M. Abd Elazim, and A. Y. Abdelaziz, “Ant Lion Optimization Algorithm for optimal location and sizing of renewable distributed generations,” *Renew. Energy*, vol. 101, pp. 1311–1324, 2017, doi: 10.1016/j.renene.2016.09.023.
- [5] A. Kaveh and A. Dadras, “A novel meta-heuristic optimization algorithm: Thermal exchange optimization,” *Adv. Eng. Softw.*, vol. 110, pp. 69–84, 2017, doi: 10.1016/j.advengsoft.2017.03.014.
- [6] I. Naruei and F. Keynia, “A new optimization method based on COOT bird natural life model,” *Expert Syst. Appl.*, vol. 183, Nov. 2021, doi: 10.1016/j.eswa.2021.115352.
- [7] A. Abdelkader, A. Rabeh, D. Mohamed Ali, and J. Mohamed, “Multi-objective genetic algorithm based sizing optimization of a stand-alone wind/PV power supply system with enhanced battery/supercapacitor hybrid energy storage,” *Energy*, vol. 163, pp. 351–363, 2018, doi: 10.1016/j.energy.2018.08.135.
- [8] A. Silvestri, A. Berizzi, and S. Buonanno, “Distributed generation planning using genetic algorithms,” *Int. Conf. Electr. Power Eng. PowerTech Budapest 1999*, p. 257, 1999, doi:

- 10.1109/PTC.1999.826689.
- [9] L. D. Arya, A. Koshti, and S. C. Choube, “Distributed generation planning using differential evolution accounting voltage stability consideration,” *Int. J. Electr. Power Energy Syst.*, vol. 42, no. 1, pp. 196–207, 2012, doi: 10.1016/j.ijepes.2012.04.011.
- [10] T. T. Tran, K. H. Truong, and D. N. Vo, “Stochastic fractal search algorithm for reconfiguration of distribution networks with distributed generations,” *Ain Shams Eng. J.*, vol. 11, no. 2, pp. 389–407, 2020, doi: 10.1016/j.asej.2019.08.015.
- [11] M. H. Moradi and M. Abedini, “A combination of genetic algorithm and particle swarm optimization for optimal DG location and sizing in distribution systems,” *Int. J. Electr. Power Energy Syst.*, vol. 34, no. 1, pp. 66–74, 2012, doi: 10.1016/j.ijepes.2011.08.023.
- [12] T. Yuvaraj, K. R. Devabalaji, and K. Ravi, “Optimal allocation of DG in the radial distribution network using bat optimization algorithm,” *Lect. Notes Electr. Eng.*, vol. 436, pp. 563–569, 2017, doi: 10.1007/978-981-10-4394-9_55.
- [13] T. N. Ton, T. T. Nguyen, A. V. Truong, and T. P. Vu, “Optimal Location and Size of Distributed Generators in an Electric Distribution System based on a Novel Metaheuristic Algorithm,” *Eng. Technol. Appl. Sci. Res.*, vol. 10, no. 1, pp. 5325–5329, 2020, doi: 10.48084/etasr.3372.
- [14] T. T. Nguyen, A. V. Truong, and T. A. Phung, “A novel method based on adaptive cuckoo search for optimal network reconfiguration and distributed generation allocation in distribution network,” *Int. J. Electr. Power Energy Syst.*, vol. 78, pp. 801–815, 2016, doi: 10.1016/j.ijepes.2015.12.030.
- [15] K. S. Sambaiah and T. Jayabarathi, “Optimal reconfiguration and renewable distributed generation allocation in electric distribution systems,” *Int. J. Ambient Energy*, vol. 42, no. 9, pp. 1018–1031, 2021, doi: 10.1080/01430750.2019.1583604.
- [16] S. Arabi Nowdeh *et al.*, “Fuzzy multi-objective placement of renewable energy sources in distribution system with objective of loss reduction and reliability improvement using a novel hybrid method,” *Appl. Soft Comput. J.*, vol. 77, pp. 761–779, 2019, doi: 10.1016/j.asoc.2019.02.003.
- [17] M. J. Hadidian-Moghaddam, S. Arabi-Nowdeh, M. Bigdeli, and D. Azizian, “A multi-objective optimal sizing and siting of distributed generation using ant lion optimization technique,” *Ain Shams Eng. J.*, vol. 9, no. 4, pp. 2101–2109, 2018, doi: 10.1016/j.asej.2017.03.001.
- [18] M. M. Othman, W. El-khattam, Y. G. Hegazy, and A. Y. Abdelaziz, “Electrical Power and Energy Systems Optimal placement and sizing of voltage controlled distributed generators in unbalanced distribution networks using supervised firefly algorithm,” *Int. J. Electr. POWER ENERGY Syst.*, vol. 82, pp. 105–113, 2016, doi: 10.1016/j.ijepes.2016.03.010.
- [19] N. Khuan, S. R. A. Rahim, M. H. Hussain, A. Azmil, and S. A. Azmi, “Integration of distributed generation and compensating capacitor in radial distribution system via firefly algorithm,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 1, pp. 67–73, 2019, doi: 10.11591/ijeecs.v16.i1.pp67-73.

- [20] M. Abdelbadea, T. A. Boghdady, and D. K. Ibrahim, “Enhancing active radial distribution networks by optimal sizing and placement of DGs using modified crow search algorithm,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 16, no. 3, pp. 1179–1188, 2019, doi: 10.11591/ijeecs.v16.i3.pp1179-1188.
- [21] N. Ghadimi, “Using HBMO Algorithm to Optimal Sizing & Sitting of Distributed Generation in Power System,” *Bull. Electr. Eng. Informatics*, vol. 3, no. 1, pp. 1–8, 2014, doi: 10.12928/eei.v3i1.179.
- [22] M. N. Morshidi, I. Musirin, S. R. A. Rahim, M. R. Adzman, and M. H. Hussain, “Whale optimization algorithm based technique for distributed generation installation in distribution system,” *Bull. Electr. Eng. Informatics*, vol. 7, no. 3, pp. 442–449, 2018, doi: 10.11591/eei.v7i3.1276.
- [23] D. Rama Prabha, T. Jayabarathi, R. Umamageswari, and S. Saranya, “Optimal location and sizing of distributed generation unit using intelligent water drop algorithm,” *Sustain. Energy Technol. Assessments*, vol. 11, pp. 106–113, 2015, doi: 10.1016/j.seta.2015.07.003.
- [24] T. T. Nguyen and A. V. Truong, “Distribution network reconfiguration for power loss minimization and voltage profile improvement using cuckoo search algorithm,” *Int. J. Electr. Power Energy Syst.*, vol. 68, pp. 233–242, 2015, doi: 10.1016/j.ijepes.2014.12.075.
- [25] J. Ahmed and Z. Salam, “A Maximum Power Point Tracking (MPPT) for PV system using Cuckoo Search with partial shading capability,” *Appl. Energy*, vol. 119, pp. 118–130, 2014, doi: 10.1016/j.apenergy.2013.12.062.
- [26] S. Mirjalili, A. H. Gandomi, S. Zahra, and S. Saremi, “Salp Swarm Algorithm : A bio-inspired optimizer for engineering design problems,” *Adv. Eng. Softw.*, vol. 0, pp. 1–29, 2017, doi: 10.1016/j.advengsoft.2017.07.002.
- [27] R. S. Rao, K. Ravindra, K. Satish, and S. V. L. Narasimham, “Power loss minimization in distribution system using network reconfiguration in the presence of distributed generation,” *IEEE Trans. Power Syst.*, vol. 28, no. 1, pp. 317–325, 2013, doi: 10.1109/TPWRS.2012.2197227.
- [28] A. Mohamed Imran, M. Kowsalya, and D. P. Kothari, “A novel integration technique for optimal network reconfiguration and distributed generation placement in power distribution networks,” *Int. J. Electr. Power Energy Syst.*, vol. 63, pp. 461–472, 2014, doi: 10.1016/j.ijepes.2014.06.011.

YSC4F.235

ẢNH HƯỞNG CỦA NHIỀU KÊNH HỢP TÁC (CCI) TÁC ĐỘNG ĐẾN HIỆU NĂNG HỆ THỐNG MẠNG VÔ TUYẾN NHẬN THỨC SỬ DỤNG NOMA

TÔN THẮT PHÙNG, LÊ PHÚC ĐỨC, TRẦN HOÀNG KHẢI LONG

*Khoa Công nghệ Điện tử, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
tonthatphung@gmail.com, Lpthien180@gmail.com, khailong368@gmail.com*

Tóm tắt. Hiện nay vấn đề về giới hạn của phổ tần là một điều đáng quan tâm và trước nhu cầu ngày càng tăng đối với các dịch vụ truyền thông không dây, vì vậy việc tìm kiếm một chiến lược thực tế để sử dụng hiệu quả phổ vô tuyến đã làm phát sinh và gọi đó là nhiễu đồng kênh (CCI) do tái sử dụng tần số trong các mạng không dây, trong đó một giải pháp khả thi (NOMA) được ủng hộ và được đáp ứng nhu cầu người dùng hiện tại. Trong bài báo cáo này nghiên cứu, đánh giá chỉ số hiệu suất hệ thống của NOMA được nghiên cứu trên kênh mờ dần $\kappa - \mu$ fading với sự can thiệp của nhiễu đồng kênh (CCI) và đồng thời của ipSIC và so sánh chúng với hiệu suất của pSIC của cùng một người dùng. Hàm mật độ xác suất (PDF) và hàm phân bố tích lũy (CDF) của giới hạn trên của SINR được phân tích. Xác định các kênh truyền mờ dần là $\kappa - \mu$ fading. Đối với hai điểm đến, xây dựng các công thức xác suất ngừng hoạt động một cách chính xác và đồng thời chạy các mô phỏng Monte-Carlo để chứng minh các biểu thức toán lý thuyết chứng minh là phù hợp.

Từ khóa. kênh mờ dần $\kappa - \mu$, nhiễu đồng kênh, hủy nhiễu liên tiếp không hoàn hảo, hệ thống đa truy cập phi trực giao.

EFFECT OF CO-CHANNEL INTERFERENCE (CCI) IMPACT ON NETWORK WITH COGNITIVE RADIO FUNCTION USING NOMA

Abstract. Currently the issue of spectrum limitation is a matter of concern and in view of the increasing demand for wireless communication services, so the search for a practical strategy to efficiently use the radio spectrum has given rise to so-called co-channel interference (CCI) due to frequency reuse in wireless networks, where a possible solution (NOMA) is advocated and is met with current user needs. In this report will study, evaluate the system performance index of NOMA studied on $\kappa - \mu$ fading channel with the intervention of co-channel interference (CCI) and at the same time of ipSIC and compare them with the performance of pSIC of the same user. The probability density function (PDF) and cumulative distribution function (CDF) of the upper limit of the SINR are analyzed. Define channels as $\kappa - \mu$ fading. For the two destinations, construct the probability formulas to stop working correctly and simultaneously run Monte-Carlo simulations to demonstrate that the theoretical mathematical expressions prove to be appropriate.

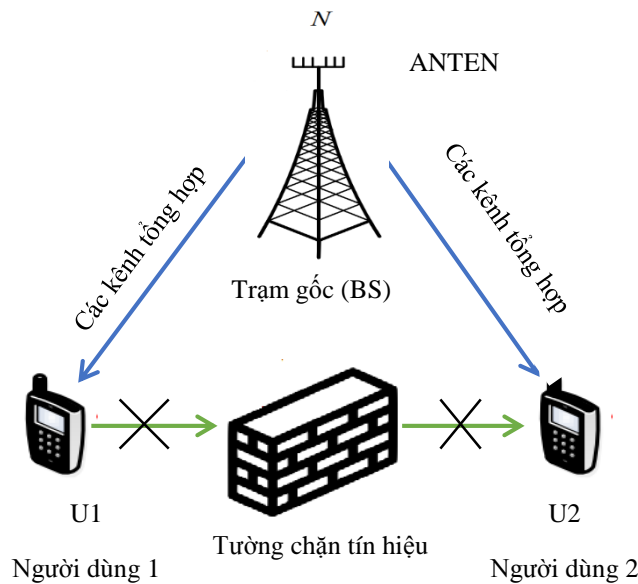
Keywords. $\kappa - \mu$ fading, co-channel interference, imperfect successive interference cancellation, non-orthogonal multiple access.

1. GIỚI THIỆU

Mạng vô tuyến nhận thức (CRN) là một hướng nghiên cứu mới trong những năm gần đây, đây giải pháp để tăng khả năng sử dụng phổ bằng sử dụng các loại phổ không được dùng hoặc chưa được khai thác, dùng để định vị các cấu hình thông số khác nhau khi đang di chuyển (như là dải tần, dạng sóng, công suất truyền). Từ đó có thể can thiệp trực tiếp vào quá trình truyền, truy cập thông tin theo ngữ cảnh của môi trường vô tuyến[9]. Đa truy cập phi trực giao (NOMA) cho phép nhiều người dùng cùng chia sẻ các kênh truyền một cách nhanh chóng và hiệu quả thông qua lớp phủ tuyến tính đơn giản hoặc ghép kênh. Công nghệ NOMA cải thiện hiệu quả phổ, giảm tối thiểu trì hoãn kênh truyền[11]-[14]. NOMA được chia thành hai loại:

NOMA miền công suất và miền mã[12], [14], [15]. Nhiều đồng kênh (CCI) xảy ra khi tín hiệu ở đầu phát tín hiệu xuất hiện sự giao thoa giữa hai hoặc nhiều sóng tín hiệu khi phát cùng một tần số. Không thể sử dụng bộ lọc để loại bỏ nhiễu vì các máy thu sử dụng cùng một tần số, nhiều đồng kênh chỉ có thể được giảm thiểu bằng cách thiết kế mạng di động (cellular) phù hợp với sử dụng. Trong cùng một nhóm tần số sử dụng không ảnh hưởng nhau, vì khoảng cách giữa các ô này phải đủ lớn. Tái sử dụng tần số là việc phân bổ cùng một tần số vô tuyến cho các vị trí địa lý khác nhau trong mạng mà không làm ảnh hưởng đến chất lượng của liên kết trên giao diện vô tuyến do nhiễu kênh lân cận và đồng kênh[6]. Trong quá trình truyền tín hiệu sẽ xảy ra hiện tượng fading giữa các kênh truyền tổng quát, đó là hiện tượng rất phổ biến trong môi trường truyền không dây đa đường đã tác động đến cường độ và xoay pha tín hiệu không giống nhau giữa mọi thời điểm và ở mọi tần số khác biệt. Các tín hiệu RF truyền trong môi trường gây ra các hiện tượng phản xạ, tán xạ hay nhiễu xạ. Do có độ trễ pha khác nhau giữa các tín hiệu nên sẽ có suy hao nhất định, dẫn đến biên độ giữa các tín hiệu khác nhau nên xuất hiện các hàm phân bố xác suất (PDF) khác nhau[10]. Trong bài báo này, sẽ tập trung phân tích các kênh truyền tổng quát (kênh Rayleigh, kênh Nakagami-m, kênh Rician) để điều chỉnh tham số (κ, μ) để thu được thông lượng kênh ổn định. Đặc biệt, nghiên cứu về các hàm PDF, CDF, xác suất dừng để cho hệ thống ổn định tránh bị pha tạp nhiễu trong các kênh. Phần còn lại của bài báo này được tổ chức như sau. Trong Phần 2, chúng tôi trình bày về mô hình hệ thống đường truyền xuống (downlink). Phân tích xác suất dừng, hiệu suất được trình bày trong Phần 3. Mô phỏng Simulink sau đó được trình bày trong Phần 4 để xác nhận rằng các phân tích của chúng tôi so với hiểu biết thực tế. Cuối cùng chúng tôi kết thúc bài báo trong Phần 5.

2. PHÂN TÍCH HỆ THỐNG ĐƯỜNG TRUYỀN XUỐNG



Hình 5: Mô hình hệ thống đường truyền xuống của kênh tổng hợp

Trong hình 1, ta giả sử tại mô hình mạng sơ cấp ở trên đã tốt, ổn định do đó trong hình 1 được xét giá các giá trị theo mô hình mạng thứ cấp được trình bày như sau: hệ thống được phát tín hiệu từ trạm gốc N (BS) đến các User 1 và User 2 gồm các kênh tổng hợp của tín hiệu bị pha tạp lẫn nhau gây ảnh hưởng đến tín hiệu khi truyền đến và làm cho tín hiệu không còn thuần khiết như ban đầu lúc anten gửi đi. Tường chặn tín hiệu có tác dụng làm ngăn cản sự truyền qua lại giữa các tín hiệu để những loại sóng, tín hiệu của User 1 tác động đến User 2 và ngược lại gây ra các hiện tượng nhiễu đồng kênh giữa hai đầu thu. Trong bài báo này, tất cả các kênh truyền được giả định theo mô hình kênh rayleigh, kênh nakagami_m, kênh rician. Xét sự truyền dữ liệu giữa trạm gốc BS và người dùng US, tín hiệu nhận được tại US được mô tả như sau:

$$y_n = \sqrt{P}h_n x_{BS} + \omega_n, \quad (1)$$

Với P là công suất truyền tải năng lượng của BS, x_{BS} là dữ liệu được gửi bởi BS, h_n là hệ số kênh truyền giữa BS và US và \square_n là nhiễu cộng tại US. Nhiễu \square_n được xem như là một biến ngẫu nhiên có phân phối Gauss giá trị trung bình bằng 0 và phương sai là N_0 .

Với mô hình hệ thống thiết kế cho thấy một trạm BS được trang bị với N ăng ten để phục vụ cho U_1 và U_2 trong NOMA, giả định rằng BS đang được trang bị n ăng ten và trong đó $n = [1, 2, \dots, N]$, từ (1) thì công thức được mô tả cho tại U_1 và U_2 :

$$y_{n,1} = g_{n,1}(\sqrt{Pb_1}s_1 + \sqrt{Pb_2}s_2) + \omega_{n,1}, \quad (2)$$

$$y_{n,2} = g_{n,2}(\sqrt{Pb_1}s_1 + \sqrt{Pb_2}s_2) + \omega_{n,2}, \quad (3)$$

Với $g_{n,i}$ được định nghĩa như độ lợi kênh của các liên kết $BS - U_1$ và $BS - U_2$, và $s = \sqrt{Pb_1}s_1 + \sqrt{Pb_2}s_2$ được giả định như là thông tin được truyền tải chồng chập từ trạm gốc BS cho U_1 và U_2 , trong đó s_1 với s_2 là các tín hiệu U_1 và U_2 lần lượt, và b_1 và b_2 là các hệ số phân bổ công suất phụ thuộc vào $b_2 > b_1$ và $b_2 + b_1 = 1$. Hơn thế nữa, U_1 và U_2 nhận được tín hiệu chồng chập và tín hiệu nhiễu vòng đồng thời.

Trong đó $\omega_{n,i} \sim CN(0, N_0)$ với $i = 1, 2$ là nhiễu được xem như một biến ngẫu nhiên có phân phối Gauss (AWGN) giá trị trung bình bằng 0 và phương sai là N_0 .

Tại U_2 có thể phát hiện s_2 bằng cách điều chỉnh ở s_1 ở (2),(3) như một giao thoa đồng kênh. Thì tín hiệu (SINR) ở U_2 được biến đổi như sau:

$$\gamma_{U_2}^n = \frac{\rho b_2 |g_{n,2}|^2}{1 + \rho b_1 |g_{n,2}|^2}, \quad (4)$$

Để đơn giản hóa biểu diễn và tính toán thì $\gamma_{n,i} = \rho |g_{n,i}|^2, i \in \{1, 2\}$, và gọi $\rho = \frac{P}{N_0}$ là SNR được truyền

tại NOMA users. Đối với việc xác nhận tín hiệu U_2 tức tín hiệu s_2 thì đầu tiên U_1 được thực hiện SIC để giải mã với mục đích là loại bỏ tín hiệu của U_2 , sau đó xử lý để giải mã tín hiệu của chính nó mà không bị can thiệp bởi ảnh hưởng nhiễu.

Đối với điều này, SINR tức thời tại U_1 để phát hiện s_2 được đưa ra công thức là:

$$\gamma_{U_2 \rightarrow U_1}^n = \frac{\rho b_2 |g_{n,1}|^2}{1 + \rho b_1 |g_{n,1}|^2} = \frac{b_2 \gamma_{n,1}}{1 + b_1 \gamma_{n,1}} \quad (5)$$

Sau khi thực hiện xong qui trình SIC ở trường hợp trên và đồng thời dẫn đến một lượng nhiễu dư dương và nhiễu ước lượng kênh mà không xuất hiện nhiễu của s_1 với SIC không hoàn hảo (ipSIC) và SIC hoàn hảo (pSIC), tỷ lệ tín hiệu trên nhiễu tức thời (SNR) ở U_1 để phát hiện tín hiệu s_1 với công thức là:

$$\gamma_{U_1}^{n,pSIC} = b_1 \gamma_{n,1}, \quad (6)$$

$$\gamma_{U_1}^{n,ipSIC} = \frac{b_1 \gamma_{n,1}}{1 + \rho |h_1|^2} \quad (7)$$

Trong đó $|h_1|^2 \sim CN(0, \lambda_1)$ được coi là độ lợi kênh truyền giữa BS và US, trong trường hợp $0 \leq \lambda_1 \leq 1$ là mức nhiễu còn lại do SIC không hoàn hảo gây ra và $CN \sim (a, b)$ là phân phối chuẩn phức tạp với trung bình a và phương sai b .

3. PHÂN TÍCH XÁC SUẤT DỪNG

3.1. Phân tích kênh truyền

Đặt $\gamma_n = \gamma_{n_1} = \gamma_{n_2}$, hàm mật độ xác suất (PDF) của γ_n được đưa ra bởi công thức [3],[4]sau:

$$p_{\gamma_n}(x|\kappa, \mu, \rho) = G(\kappa, \mu, \rho) x^{(\mu-1)/2} e^{-\lambda_g x} I_{\mu-1}(2\sqrt{\kappa\mu\lambda_g x}), x > 0 \quad (8)$$

Trong đó $\lambda_g = \frac{(1+\kappa)\mu}{\rho}$ và $G(\kappa, \mu, \rho) = \frac{\lambda_g^{(\mu+1)/2}}{[\kappa\mu]^{(\mu-1)/2} e^{\kappa\mu}}$, $\kappa > 0$ lần lượt là tổng năng lượng của các thành phần chi phối và tổng công suất của tỷ lệ sóng phân tán. $\mu > 0$ mô tả số lượng các cụm đa đường và $I_\nu(x)$ biểu thị ν -th của hàm Bessel. Tương tự, các PDF của γ_i , $i \in \{1, 2\}$ trong κ - μ các kênh mờ dần có thể được thể hiện bằng $P\gamma_i(x|k_i, \mu_i, \rho)$. Đối với các giá trị tùy ý κ và μ , hàm phân phối tích lũy (CDF) của γ_n có thể nhận được từ [2], (38):

$$F_{\gamma_n}(x|\kappa, \mu, \rho) = 1 - Q_\mu(\sqrt{2\kappa\mu}, \sqrt{2\lambda_g x}), x > 0 \quad (9)$$

Với hàm Marcum Q tổng quát $Q_m(\alpha, \beta) = \frac{1}{\alpha^{m-1}} \int_\beta^\infty x^m \exp(-\frac{x^2 + \alpha^2}{2}) I_{m-1}(\alpha x) dx$

Đối với các giá trị nguyên của μ và đồng thời ta thay biến để biến đổi công thức (9) theo dạng công thức hàm marcum Q [1], (4.33) và $F_{\gamma_n}(x|\kappa, \mu, \rho)$ có thể được viết lại và biểu diễn một cách đơn giản như:

$$F_{\gamma_n}(x|\kappa, \mu, \rho) = e^{-\kappa\mu - \lambda_g x} \sum_{r=\mu}^{\infty} \left(\frac{\lambda_g x}{\kappa\mu}\right)^{r/2} I_r(2\sqrt{\kappa\mu\lambda_g x}), \quad (10)$$

Sử dụng công thức $I_\nu(z) = \sum_{k=0}^{\infty} \frac{1}{k! \Gamma(\nu + k + 1)} \left(\frac{z}{2}\right)^{\nu+2k}$ được định nghĩa là hàm Bessel và thay thế biến q [16], (8.445) trong công thức $I_\nu(z)$ ta thu được dạng biến đổi:

$$F_{\gamma_n}(x|\kappa, \mu, \rho) = e^{-\kappa\mu - \lambda_g x} \sum_{r=\mu}^{\infty} \sum_{n=0}^{\infty} \frac{\kappa^n \mu^n \lambda_g^{n+r} x^{n+r}}{n! \Gamma(n+r+1)}, \quad (11)$$

3.2. Xác suất dừng tại hai máy thu

Trường hợp 1: $0 < \lambda \leq 1$, triệt nhiễu tuần tự không hoàn hảo (ipSIC) đối với xác suất dừng của User 1.

$$\begin{aligned} P_1^{ipSIC} &= \Pr(\max_{n \in N} \{\gamma_{U_2 \rightarrow U_1}^{\sim n}\} < z_2 \cup \max_{n \in N} \{\gamma_{U_1}^{\sim n, ipSIC}\} < z_1) \\ P_1^{ipSIC} &= \prod_{n=1}^N \left[1 - \Pr\left(\gamma_{U_2 \rightarrow U_1}^{\sim n} \geq z_2, \gamma_{U_1}^{\sim n, ipSIC} \geq z_1\right) \right] \\ P_1^{ipSIC} &= \left[1 - \Pr\left(\gamma_1 \geq \tilde{\eta}_2, \gamma_1 \geq \tilde{\eta}_1 (1 + \rho |h_1|^2)\right) \right]^N \end{aligned} \quad (12)$$

Trong đó $z_i = 2^{2R_i} - 1$ là hệ số nhiễu cho trước với $i=1, 2$. R_i là tốc độ dữ liệu U_i . $\tilde{\eta}_2 = \frac{z_2}{b_2 - z_2 b_1}$ và $\tilde{\eta}_1 = \frac{z_1}{b_1}$

là hệ số suy hao đường truyền. Giả sử $\tilde{\eta}_1 (1 + \rho |h_1|^2) \gg \tilde{\eta}_2$, P_1^{ipSIC} sẽ được tính theo cách sau

$$P_1^{ipSIC} = \left[1 - \Pr \left(\gamma_{n,1} \geq \tilde{\eta}_1 (1 + \rho |h_T|^2) \right) \right]^N$$

$$P_1^{ipSIC} = \left\{ 1 - \int_0^\infty f_{|h_T|^2}(x) \left[1 - F_{\gamma_{n,1}} \left(\tilde{\eta}_1 (1 + \rho x) \right) \right] dx \right\}^{N'} \quad (13)$$

Theo công thức (11), (13)

$$P_1^{ipSIC} = \left(1 - \lambda_T \int_0^\infty e^{-\frac{x}{\lambda_T}} dx + \frac{e^{-\kappa\mu - \lambda_g \tilde{\eta}_1}}{\lambda_T} \sum_{r=\mu}^\infty \sum_{n=0}^\infty \frac{\kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_1^{\sim n+r}}{n! \Gamma(n+r+1)} \int_0^\infty e^{-x(\lambda_g \tilde{\eta}_1 \rho)} e^{-x(\frac{x}{\lambda_T})} (1 + \rho x)^{n+r} dx \right)^N$$

$$P_1^{ipSIC} = \left(1 - \lambda_T \int_0^\infty e^{-\frac{x}{\lambda_T}} dx + \frac{e^{-\kappa\mu - \lambda_g \tilde{\eta}_1}}{\lambda_T} \sum_{r=\mu}^\infty \sum_{n=0}^\infty \frac{\kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_1^{\sim n+r}}{n! \Gamma(n+r+1)} \int_0^\infty e^{-x(\frac{1}{\lambda_T} + \lambda_g \tilde{\eta}_1 \rho)} (1 + \rho x)^{n+r} dx \right)^N \quad (14)$$

Áp dụng định luật nhị thức Newton, ta có:

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k = \sum_{k=0}^n \binom{n}{k} x^k y^{n-k} \quad (15)$$

Từ công thức (14), (15) suy ra:

$$P_1^{ipSIC} = \frac{e^{-\kappa\mu - \lambda_g \tilde{\eta}_1}}{\lambda_T} \sum_{r=\mu}^\infty \sum_{n=0}^\infty \sum_{q=0}^{\infty} \binom{n+r}{q} \frac{\kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_1^{\sim n+r} \rho^q}{n! \Gamma(n+r+1)} \int_0^\infty e^{-x(\frac{1}{\lambda_T} + \lambda_g \tilde{\eta}_1 \rho)} x^q dx \quad (16)$$

Dựa vào [16] (3.351.5) Công thức xác suất dừng dạng đóng của User 1:

$$P_1^{ipSIC} = \left[e^{-\kappa\mu - \lambda_g \tilde{\eta}_1} \sum_{q=0}^{\infty} \sum_{r=\mu}^{\infty} \sum_{n=0}^{\infty} \binom{n+r}{q} \frac{q! \kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_1^{\sim n+r} \lambda_T^q \rho^q}{n! \Gamma(n+r+1) (1 + \lambda_T \lambda_g \tilde{\eta}_1 \rho)^{(q+1)}} \right]^N \quad (17)$$

Trường hợp 2: $\lambda = 0$, triệt nhiễu tuần tự hoàn hảo (pSIC) đối với xác suất dừng của User 2.

$$P_1^{pSIC} = \Pr \left(\max_{n \in N} \{ \gamma_{U_2 \rightarrow U_1}^{\sim n} \} < z_2 \cup \max_{n \in N} \{ \gamma_{U_1}^{\sim n, ipSIC} \} < z_1 \right)$$

$$P_1^{pSIC} = \prod_{n=1}^N \left[1 - \Pr \left(\gamma_{n,1}^{\sim n} \geq \tilde{\eta}_2, \gamma_{n,1} \geq \tilde{\eta}_1 \right) \right]$$

$$P_1^{pSIC} = \left[1 - \Pr(\gamma_{n,1} \geq \tilde{\eta}_{\max}) \right]^N$$

$$P_1^{pSIC} = \left[F_{\gamma_{n,1}} \left(\tilde{\eta}_{\max} \right) \right]^N \quad (18)$$

với $\tilde{\eta}_{\max} = \left[\tilde{\eta}_1, \tilde{\eta}_2 \right]^+$, $[a, b]^+ = \max(a, b)$

Từ đó áp dụng các công thức (11), (18), (9) ta xây dựng:

$$\begin{aligned}
 P_1^{pSIC} &= \left[Q_\mu \left(\sqrt{2\kappa\mu}, \sqrt{2\lambda_g \tilde{\eta}_{\max}} \right) \right]^N \\
 P_1^{pSIC} &= \left[e^{-\kappa\mu - \lambda_g \tilde{\eta}_{\max}} \sum_{r=\mu}^{\infty} \sum_{n=0}^{\infty} \frac{\kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_{\max}^{n+r}}{n! \Gamma(n+r+1)} \right]^N,
 \end{aligned} \tag{19}$$

Từ công thức tính P_1^{pSIC} có thể suy ra công thức xác suất dừng của P_2 :

$$\begin{aligned}
 P_2^{pSIC} &= \Pr(\max_{n \in N} \{\gamma_{U_2}^{\sim n}\} < z_2) \\
 P_2^{pSIC} &= \prod_{n=1}^N \left[1 - \Pr(\gamma_{U_2}^{\sim n} \geq z_2) \right] \\
 P_2^{pSIC} &= \left[1 - \Pr(\gamma_{n,2} \geq \tilde{\eta}_2) \right]^N, \\
 P_2^{pSIC} &= \left[F_{\gamma_{n,2}}(\tilde{\eta}_2) \right]^N
 \end{aligned} \tag{20}$$

Công thức xác suất dừng dạng đóng của U_2 :

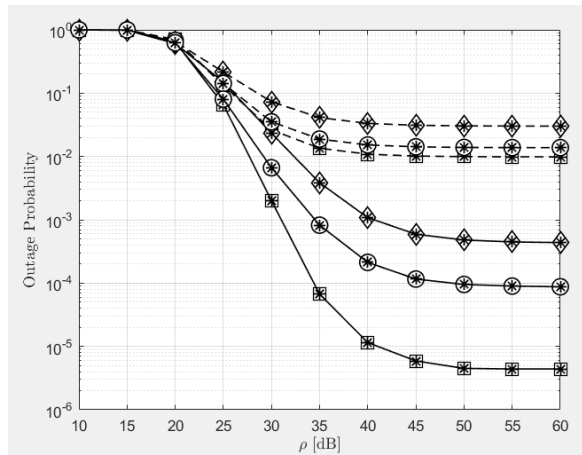
$$\begin{aligned}
 P_2^{pSIC} &= \left[Q_\mu \left(\sqrt{2\kappa\mu}, \sqrt{2\lambda_g \tilde{\eta}_2} \right) \right]^N \\
 P_2^{pSIC} &= \left[e^{-\kappa\mu - \lambda_g \tilde{\eta}_2} \sum_{r=\mu}^{\infty} \sum_{n=0}^{\infty} \frac{\kappa^n \mu^n \lambda_g^{n+r} \tilde{\eta}_2^{n+r}}{n! \Gamma(n+r+1)} \right]^N,
 \end{aligned} \tag{21}$$

4. KẾT QUẢ VÀ THẢO LUẬN

Trong phần này, các kết quả phân tích thu được trình bày dưới dạng đồ thị sẽ cung cấp thông tin chi tiết về thông lượng tại các kênh truyền tổng hợp.

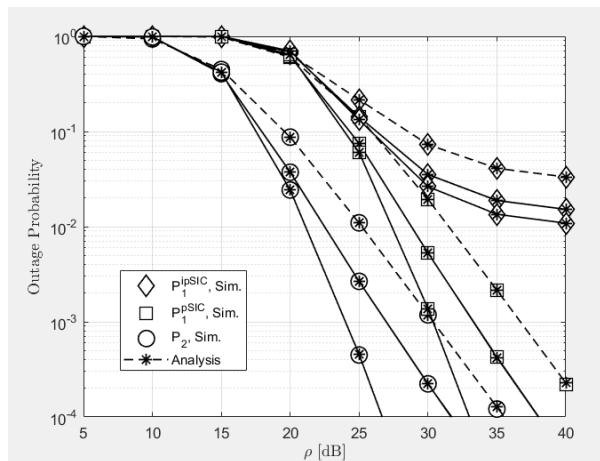
Bảng 1: Các thông số được sử dụng mô phỏng

Thông số của hệ thống	Giá trị tương ứng
Phép thử trong mô phỏng Monte Carlo	10^6
Số Ăng – ten của trạm BS	$N = 2$
Hệ số phân bổ công suất	$(b_1, b_2) = \{0.1 ; 0.9\}$
Tốc độ dữ liệu tại U_1	2 bits/s/Hz
Tốc độ dữ liệu tại U_2	1 bits/s/Hz
Mức ảnh hưởng của tín hiệu nhiễu	$\lambda_1 = 0.001$



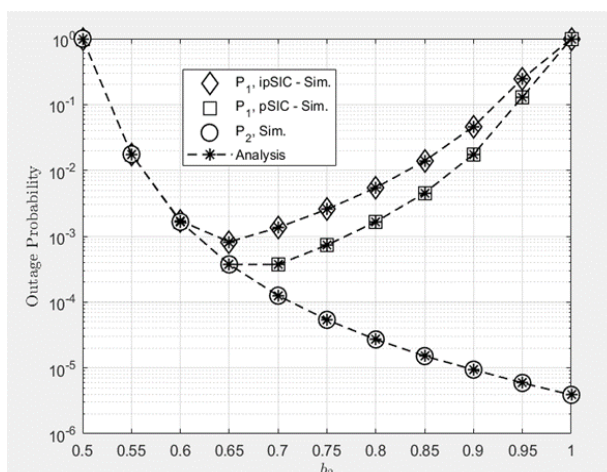
Hình 2. Xác suất của U_1 cho ipSIC với ρ [dB] cho các kênh mờ dần khác nhau tại 2 giá trị λ_1

Hình 2 quan sát được các đường cong OP tức các đường bị cong do bị ảnh hưởng bởi nhiễu tiếp cận được sự ổn định ở các giá trị SNR lớn vì đây là dấu hiệu của lỗi do SIC đưa ra nói chung tại máy thu. Tuy nhiên, các đường cong OP sẽ có hiệu suất cũng như có các giá trị khác nhau tùy thuộc vào giá trị λ_1 . Lý do là ipSIC sẽ hoạt động dưới các trường hợp có mức tín hiệu nhiễu lớn. Và hiệu suất của OP sẽ đạt được mức ổn định nhất dưới kênh truyền của Nakagami – m dù khi ta tăng mức công suất thì giá trị vẫn không thay đổi.

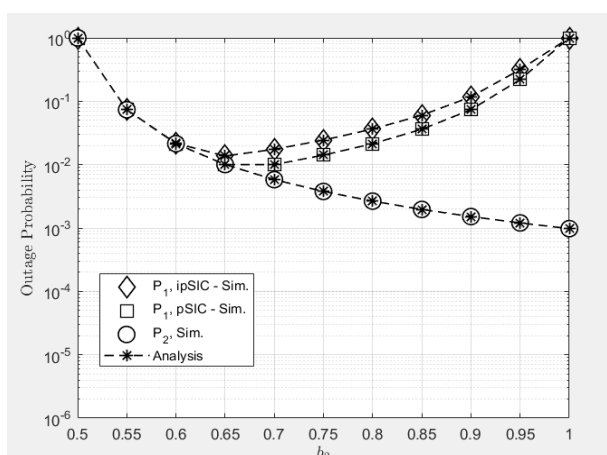


Hình 3. Xác suất dừng so với ρ [dB] cho kênh fading khác nhau với $\lambda_1 = 0.001$

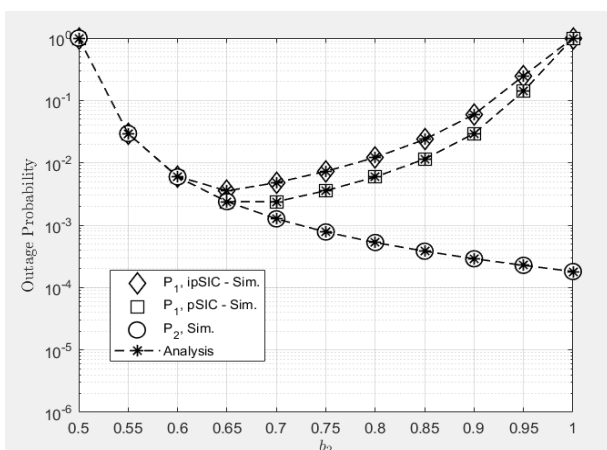
4. Từ hình 3, mô tả xác suất dừng OP cho U_1 ở cả 2 trường hợp với các điều kiện pSIC và ipSIC đồng thời cũng mô tả xác suất tại U_2 so với tín hiệu SNR truyền cho các kênh mờ khác nhau với $\lambda_1 = 0.001$ ta có thể quan sát thấy được các giá trị khác nhau của các thông số phân phối $\kappa - \mu$ tác động đến OP và hiệu suất cũng như độ ổn định của OP sẽ đạt tốt nhất trong trường hợp $(\kappa - \mu) = (2,0)$ tức kênh Nakagami- m cho cả hai U_1 và U_2 . Đồng thời ta có thể quan sát OP cho U_2 không đạt được hiệu suất ổn định vì lý do SIC không được sử dụng ở bộ thu do vậy tại U_2 vẫn bị ảnh hưởng bởi nhiễu với các giá trị SNR nhỏ. Nhưng khác đường cong OP cho U_1 sẽ đạt được hiệu suất gần được ổn định ở các giá trị SNR lớn khi pSIC được sử dụng ở máy thu thay vì U_2 .



Hình 4. Xác suất dừng so với b_2 với giá trị cho kênh Nakagami - m , với $\lambda_1 = 0.01$ và $R_1 = R_2 = 0.5$



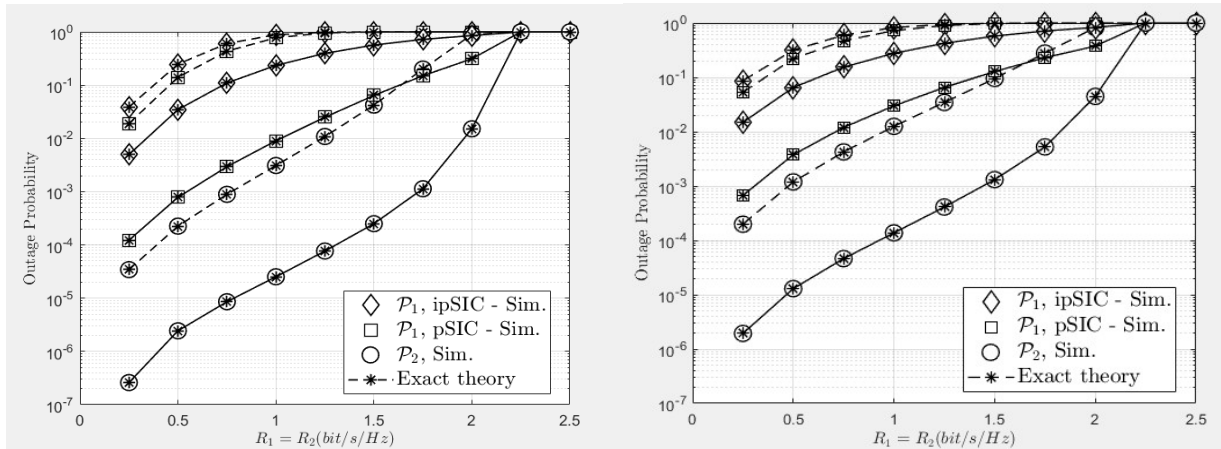
Hình 5. Xác suất dừng so với b_2 với giá trị cho kênh Rayleigh khác nhau, với $\lambda_1 = 0.01$ và $R_1 = R_2 = 0.5$



Hình 6. Xác suất dừng so với b_2 với giá trị cho kênh Rician, với $\lambda_1 = 0.01$ và $R_1 = R_2 = 0.5$

Từ 3 hình 4,5,6 có thể quan sát thấy được các giá trị khác nhau của các thông số phân phối $\kappa - \mu$ tác động đến OP và hiệu suất cũng như độ ổn định của OP sẽ đạt tốt nhất trong trường hợp $(\kappa - \mu) = (2, 0)$ tức kênh Nakagami- m cho cả hai U_1 và U_2 và tương tự với 2 kênh còn lại sẽ hoạt động kém hiệu quả hơn cũng như sự ổn định. Tuy nhiên ở các kênh mờ dần khác nhau từ các hệ số $\kappa - \mu$ ở giai đoạn $b_2 = 0.65$ thì xảy ra trường hợp SIC cho cả hai điều kiện $pSIC$ và $ipSIC$ tại U_1 đang hoạt động một cách kém hiệu quả hơn cho với tại

xác suất U_2 khi không có SIC, điều này cũng chứng minh rằng mối quan hệ giữa hệ số phân bổ năng lượng cả b_2 và b_1 sẽ có ảnh hưởng lên hiệu suất của OP.



Hình 7. Xác suất dừng với tốc độ dữ liệu cho 2 kênh fading khác nhau

Với hình 7 mô tả xác suất dừng OP cho U_1 với các điều kiện pSIC và ipSIC và đồng thời cũng mô tả xác suất dừng tại U_2 với tốc độ dữ liệu được thay đổi qua các giai đoạn khác nhau với giá trị của $R_1 = R_2$ được thay đổi cũng như trong khi $\rho[dB]$ được thay đổi, cho $b_2 = 0.95$ và $b_1 = 0.05$ với $\lambda_l = 0.01$, với đường nét đứt biểu thị cho $\rho[dB] = 15$ và đường nét liền biểu thị cho $\rho[dB] = 25$. Từ hình 4.9, ta có thể quan sát thấy được nếu các giá trị $R_1 = R_2$ hay $\rho[dB]$ được thay đổi theo các giai đoạn thì nó sẽ tác động đến xác suất dừng OP. Hiệu suất OP tốt nhất đạt được ở $R_1 = R_2 = 0.3$ (bit/s/Hz) cho cả U_1 và U_2 . Ngoài ra, đối với U_1 lẫn cả U_2 thì OP chậm đỉnh ở mức các giá trị cao của $R_1 = R_2$.

KẾT LUẬN

Bài báo đã nghiên cứu được mạng giao tiếp NOMA với hai người dùng U_1 và U_2 song đó vẫn chưa đi sâu về mạng vô tuyến nhận thức xét trên mô hình hệ thống cũng như chưa đánh giá được các hiệu năng như tỷ lệ lỗi bit hay tỷ lệ lỗi ký tự. Chúng tôi cũng đã chọn công thức toán học cho OP trong trường hợp máy thu có pSIC. Các biểu thức này được mô phỏng bằng Monte – Carlo bằng cách sử dụng các thông số $\kappa - \mu$ tương ứng với các kênh phân phối fading và cụ thể là các kênh như Nakagami – m, Rician hay Rayleigh. Về bài báo nghiên cứu đã mô phỏng thành công tuy nhiên không có nhiều cải tiến khi so sánh với các bài báo khác tuy nhiên có thể kể đến tại mô phỏng khảo sát tốc độ truyền dữ liệu ảnh hưởng đến hệ thống đã có thể khảo sát bổ sung thêm kênh phân phối fading bằng việc sử dụng thông số $\kappa - \mu$ nhằm để làm rõ hơn về ảnh hưởng của tốc độ truyền qua nhiều kênh fading. Các mô phỏng bổ sung đã được thực hiện bằng cách sử dụng phân phối fading chung của κ và μ là các yếu tố mấu chốt để sử dụng chính. Như có thể thấy từ các mô phỏng, thì có thể kết luận rằng các mô hình SIC không hoàn hảo sẽ có phần tốt hơn các mô hình SIC hoàn hảo.

TÀI LIỆU THAM KHẢO

- [1] M. K. Simon and M.-S. Alouini, “Digital Communication over Fading Channels”, John Wiley & Sons, vol. 95, 2005
- [2] N. Y. Ermolova and O. Tirkkonen, “Outage Probability Analysis in Generalized Fading Channels with Co-Channel Interference and Background Noise: $\eta - \mu/\eta - \mu$, $\eta - \mu/\kappa - \mu$ and $\kappa - \mu/\eta - \mu$ Scenarios,” IEEE Trans. Wireless Commun., vol. 13, no. 1, pp. 291–297, January 2014, doi: 10.1109/TWC.2013.110813.130624.
- [3] M. D. Yacoub, “The $\kappa - \mu$ distribution and the $\eta - \mu$ distribution,” IEEE Anten. Prop. Mag., vol. 49, no. 1, pp. 68-81, Feb. 2007, doi: 10.1109/MAP.2007.370983.

- [4] W. Li and M. Chen, "Outage capacity of dual-hop decode-and-forward relaying system over generalized fading channels," 2010 2nd International Conference on Future Computer and Communication, 2010, pp. V3-827-V3-831, doi: 10.1109/ICFCC.2010.5497387.
- [5] Phạm Quốc Cường, "Kỹ thuật đa truy cập phi trực giao NOMA trong mạng 5G", Đà Nẵng, 9-2018, pp 6-8
- [6] Nguyễn Quốc Điền, "Đánh giá hiệu năng của mạng truyền thông cộng tác dưới tác động của suy hao phân cứng và nhiễu đồng kênh", TP.HCM, 2015, pp 10 -13
- [7] Nguyễn Văn Chính, "Về truyền thông kết hợp trong môi trường vô tuyến nhận thức: cải thiện và đánh giá hiệu năng mạng thứ cấp", Hà Nội, 2017, pp 17-19
- [8] Chi-Bao Le, Hong-Nhu Nguyen, Ngoc-Long Nguyen, Miroslav Voznak, Nhan Duc Nguyen, "Considering the $\kappa - \mu$ fading channels adopted in multiple antennas downlink non-orthogonal multiple access", Vol. 11, No. 1, February 2022, pp. 336~345, ISSN: 2302-9285, DOI: 10.11591/eei.v11i1.3453.
- [9] Adrian Popescu, "Cognitive Radio Networks", Blekinge Institute of Technology, Dept. of Communications and Computer Systems, Karlskrona, Sweden. pp11-12
- [10] Võ Văn Thắng, "Đánh giá hiệu năng mạng vô tuyến ngẫu nhiên đa người dùng trên các kênh truyền fading tổng quát", TP.HCM, 2017, pp 17-21
- [11] Ding, Fan, and Poor, "Impact of user pairing on 5G nonorthogonal multiple-access downlink transmissions," IEEE Trans. Veh. Technol, vol. 65, no. 8, pp. 6010-6023, 2015
- [12] L. Dai, B. Wang, Y. Yuan, S. Han, I. Chih-lin, and Z. Wang, "Non-orthogonal multiple access for 5G: solutions, challenges, opportunities, and future research trends," in IEEE Communications Magazine, vol. 53, no. 9, pp. 74-81, September 2015, doi: 10.1109/MCOM.2015.7263349
- [13] Y. Liu, Z. Qin, M. ElKashlan, Z. Ding, A. Nallanathan, and L. Hanzo, "Nonorthogonal Multiple Access for 5G and Beyond," in Proceedings of the IEEE, vol. 105, no. 12, pp. 2347-2381, Dec. 2017, doi: 10.1109/JPROC.2017.2768666.
- [14] H. Sadia, M. Zeeshan, and S. A. Sheikh, "Performance analysis of downlink power domain NOMA under fading channels," 2018 ELEKTRO, 2018, pp. 1-6, doi: 10.1109/ELEKTRO.2018.8398247
- [15] S. M. R. Islam, N. Avazov, O. A. Dobre, and K. Kwak, "Power-Domain Non-Orthogonal Multiple Access (NOMA) in 5G Systems: Potentials and Challenges," in IEEE Communications Surveys & Tutorials, vol. 19, no. 2, pp. 721-742, Secondquarter 2017, doi: 10.1109/COMST.2016.2621116.
- [16] I. S. Gradshteyn and I. M. Ryzhik, "Tables of Integrals, Series and Products, 6th ed. New York: Academic Press", 2000.

YSC4F.236

NGHIÊN CỨU VÀ THIẾT KẾ BỘ ĐIỀU KHIỂN PI MỜ ĐỂ ĐIỀU KHIỂN TỐC ĐỘ ĐỘNG CƠ KHÔNG ĐỒNG BỘ DỰA TRÊN LABVIEW

NGÔ THANH QUYỀN, NGUYỄN VĂN THỌ, NGUYỄN ĐỨC TRUNG, PHẠM NGỌC DUYÊN,
LÊ NGUYỄN TÂN

*Khoa Công nghệ Điện, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh²
ngothanhquyen@iuh.edu.vn , nguyenvantho@iuh.edu.vn , 18045671.trung@student.iuh.edu.vn ,
18052921.duyen@student.iuh.edu.vn, 18045931.tan@student.iuh.edu.vn*

Tóm tắt. Do điều khiển thông minh, điều khiển mờ, mạng thần kinh nhân tạo hay giải thuật di truyền đã và đang được áp dụng rộng rãi trong việc điều khiển các đối tượng phi tuyến. Nên trong bài báo này, sự kết hợp giữa bộ điều khiển PI truyền thống và bộ điều khiển mờ, còn gọi là PI mờ (Hybrid of PI and Fuzzy) được đề xuất cho việc điều khiển tốc độ trên đối tượng phi tuyến là động cơ không đồng bộ 3 pha. Hiệu quả của sự kết hợp này sẽ được so sánh trực tiếp với bộ điều khiển PI hoặc bộ điều khiển mờ đơn lẻ. Trong bài báo này, việc điều khiển, giám sát và thu thập dữ liệu sẽ được thực hiện trên máy tính thông qua phần mềm LabView và card điều khiển NI myRIO 1900. Với bộ điều khiển được đề xuất, thời gian quá độ và độ vọt lố (overshoot) được tối ưu hơn so với hai phương pháp còn lại trong trường hợp khi có tải thay đổi.

Từ khóa. bộ điều khiển logic, PI mờ

RESEARCH AND DESIGN OF DIFFERENT PID CONTROLLER FOR SPEED CONTROL OF LABVIEW BASED

Abstract. Intelligent control, limited as fuzzy control, artificial neural network or moving algorithm and is widely applied in the control of non-object routes. Therefore, in this paper, the combination of controlled and fuzzy PID transmission system, also called fuzzy PID (Hybrid of PID and Fuzzy) is proposed for speed control on object route is a engine. not 3 phase engine. The efficiency of this combination will be compared directly with the PID controller or single fuzzy controller. In this paper, the controller, monitoring and decimal data will be implemented on the computer through LabView software and NI MYRIO 1900 card controller. With the controller output, the transient time and Overshoot is more optimized than the other two methods in the case of variable loads.

Keywords. PID, PI, Fuzzy, Hybrid of PID and Fuzzy, and Labview, overshoot..

1. GIỚI THIỆU

Với quá trình công nghiệp hóa hiện đại hóa thì các quy trình điều khiển tự động đang ngày phát triển đa dạng và mạnh mẽ, trong đó điều khiển tốc độ động cơ là vấn đề cốt lõi và được xem xét đầu tiên trong hầu hết các hệ thống điều khiển tự động. Các bộ điều khiển công nghiệp phổ biến như bộ điều khiển P, bộ điều khiển PI, bộ điều khiển PID truyền thống thường được sử dụng để điều khiển tốc độ động cơ, chẳng hạn như bộ điều khiển tích phân tỉ lệ, gọi tắt là bộ điều khiển PI (Proportional– Integral controller) được ứng dụng rất phổ biến trong việc vận hành hệ thống các quá trình công nghiệp, do khả năng điều khiển hiệu quả, cấu trúc đơn giản và phạm vi ứng dụng tương đối rộng. Tuy nhiên, đối với một số hệ thống, việc hiệu chỉnh bộ điều khiển PI bằng phương pháp này mất khá nhiều thời gian để thực nghiệm, phân tích và chỉnh định tham số phù hợp. Nhiều sản phẩm phần cứng khác nhau, phương thức khác nhau được sử dụng cho thuật toán điều khiển tốc độ động cơ. Các bộ điều khiển này không đáp ứng tốt được các yêu cầu khi có tải hoặc khi tốc độ thay đổi. Do đó, một số cấu trúc điều khiển nâng cao đã được sử dụng để cải thiện hiệu quả

của kiểm soát PI trong các điều kiện hoạt động phức tạp như tải hoặc tốc độ thay đổi lớn. Sự kết hợp giữa bộ điều khiển PI và bộ điều khiển mờ tạo ra bộ điều khiển Hybrid of PI and Fuzzy hay còn gọi là PI mờ mang lại hiệu quả tốt hơn so với các phương pháp điều khiển PI truyền thống với overshoot (độ vọt lố) thấp hơn và khả năng kiểm soát tốc độ tốt hơn khi có tải thay đổi.

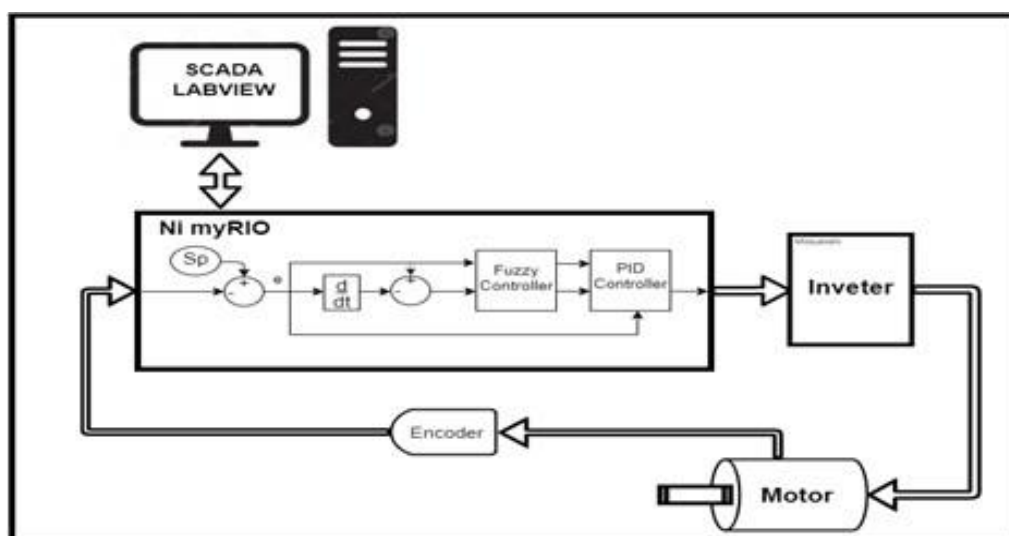
Về nguyên tắc, có hai kiểu tinh chỉnh bộ điều khiển PI đã được nghiên cứu và triển khai ứng dụng. Đó là, tinh chỉnh ngay trong quá trình hoạt động của hệ thống và tinh chỉnh trước khi vận hành hệ thống hoặc chỉ cho hệ thống vận hành một thời gian nhất định. Trong đó, việc áp dụng giải thuật di truyền được xem là một phương pháp hiệu quả để tinh chỉnh trước khi vận hành bộ điều khiển. Tuy nhiên, việc tìm kiếm các giá trị tối ưu theo thời gian thực của các thông số bộ điều khiển trong miền xác định của chúng đòi hỏi rất nhiều thời gian, cũng như có thể phải cần đến mô hình toán học của đối tượng hoặc điều khiển PI mờ.

Trong bài báo này, nhóm tác giả đề xuất giải thuật điều khiển thích nghi bằng việc kết hợp giữa bộ điều khiển PI truyền thống và bộ điều khiển mờ thông qua phần mềm Labview có tích hợp sẵn card nhúng NI myRIO 1900 để điều khiển tốc độ động cơ không đồng bộ 3 pha trong điều kiện khi có tải thay đổi. Kết quả điều khiển của phương pháp đề xuất được phân tích và so sánh về thời gian quá độ, độ vọt lố (overshoot) và sai số tĩnh so với hai phương pháp điều khiển PI và điều khiển mờ. Bộ cục của bài báo được trình bày như sau: phần 2 giới thiệu về hệ thống và phương tiện nghiên cứu, phần 3 trình bày giải thuật điều khiển cho hệ thống, phần 4 là kết quả thực nghiệm và so sánh, phần 5 là đưa ra các kết luận.

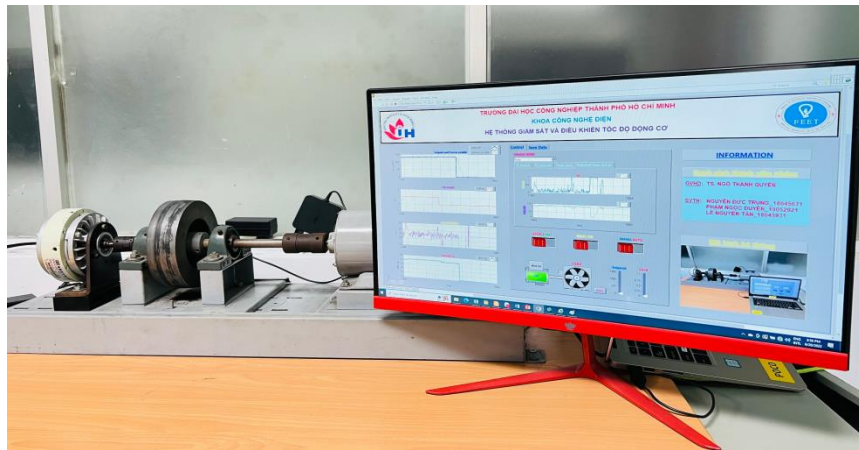
2. MÔ TẢ HỆ THỐNG

Đối tượng nghiên cứu trong bài báo này là một động cơ không đồng bộ 3 pha (0.75kW), được điều khiển tốc độ thông qua một biến tần Mitsubishi (0.75kW). Tín hiệu điều khiển từ máy tính được xuất ra thông qua card điều khiển NI (National Instruments) myRIO 1900 dưới dạng điện áp 0 - 10V và đưa đến ngõ vào của biến tần. Biến tần nhận tín hiệu điện áp ở ngõ vào (0 - 10V) và biến đổi thành điện áp có tần số thay đổi (0 - 60 Hz) để điều khiển tốc độ động cơ không đồng bộ. Tốc độ quay của trục động cơ được hồi tiếp về card điều khiển NI myRIO thông qua bộ đọc tốc độ Encoder (2048 xung/vòng). Chương trình điều khiển và giao diện SCADA thực hiện việc thu thập, giám sát dữ liệu được thực hiện trên máy tính dựa trên phần mềm LabView. Sơ đồ khối điều khiển tốc độ động cơ không đồng bộ và mô hình thực tế được trình bày tương ứng như trên Hình 1 và Hình 2.

Phần mềm LabView (Laboratory Virtual Instrumentation Engineering Workbench) là một phần mềm máy tính được phát triển bởi National Instruments. LabView dùng trong hầu hết các phòng thí nghiệm, lĩnh vực khoa học kỹ thuật như tự động hóa, điều khiển, điện tử, cơ điện tử, hàng không, hóa sinh, điện tử y sinh. Ở nghiên cứu này, nhóm tác giả sử dụng phần mềm LabView phục vụ cho việc nghiên cứu.



Hình 1: Sơ đồ khối của hệ thống PI mờ



Hình 2: Mô hình thực tế của hệ thống

3. THIẾT KẾ BỘ ĐIỀU KHIỂN

3.1. Bộ điều khiển PI

Do tính chất phi tuyến của động cơ không đồng bộ mà các bộ điều khiển PI thông thường không thể kiểm soát tốt khi điều kiện tải thay đổi. Khi động cơ có tải và không tải, các thông số PI phải được thay đổi theo thời gian thực để hệ thống đảm bảo chất lượng điều khiển ở tốc độ cao và thấp. Tín hiệu ngõ ra của bộ điều khiển PI được xác định như sau:

$$u(t) = K_p e(t) + K_I \int_0^t e(t) dt \quad (1)$$

$$e(t) = SP - PV(t) \quad (2)$$

Trong đó: K_p và K_i lần lượt là độ lợi tỉ lệ và độ lợi tích phân của bộ điều khiển, $e(t)$ là giá trị sai số của hệ thống, SP là giá trị đặt (setpoint), $PV(t)$ là giá trị tốc độ động cơ theo thời gian (process variable). Nguyên hàm sai số được tính bằng công thức:

$$ce(t) = e(t) + e(t - 1) \quad (3)$$

Ở đây $e(t-1)$ là giá trị sai số trước đó.

Nhiệm vụ của người thiết kế bộ điều khiển PI là chọn lựa bộ giá trị $\{K_p, K_i\}$ thỏa mãn các yêu cầu về chất lượng điều khiển. Các thông số độ lợi K_p và K_i ảnh hưởng đến bảng thông tin này được xem là cơ sở tri thức để xác định luật chỉnh định mờ bộ điều khiển PI.

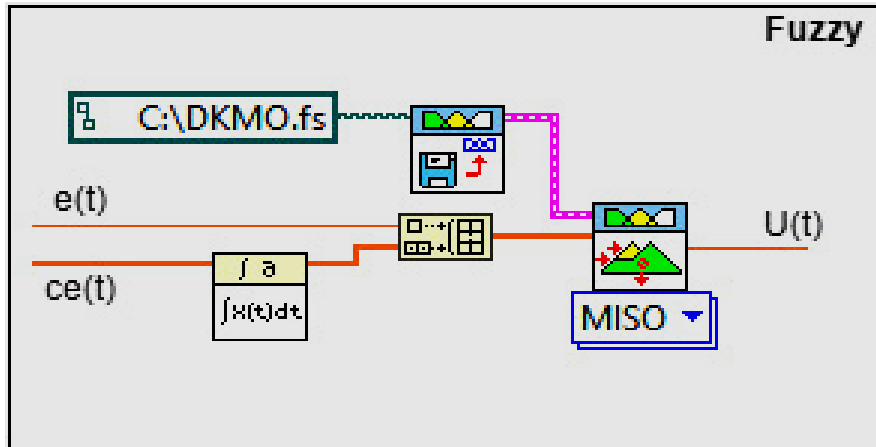
Bảng 1: Ảnh hưởng của việc tăng các thông số của bộ điều khiển PID

Đáp ứng vòng kín	Thời gian lên	Vọt ló	Thời gian xác lập	Sai số xác lập
K_p	Giảm	Tăng	Thay đổi nhỏ	Giảm
K_i	Giảm	Tăng	Tăng	Thay đổi nhỏ
K_d	Thay đổi nhỏ	Giảm	Giảm	Thay đổi nhỏ

3.2. Bộ điều khiển Fuzzy (Bộ điều khiển mờ)

3.2.1. Bộ điều khiển Fuzzy ngõ ra điện áp điều khiển đối tượng

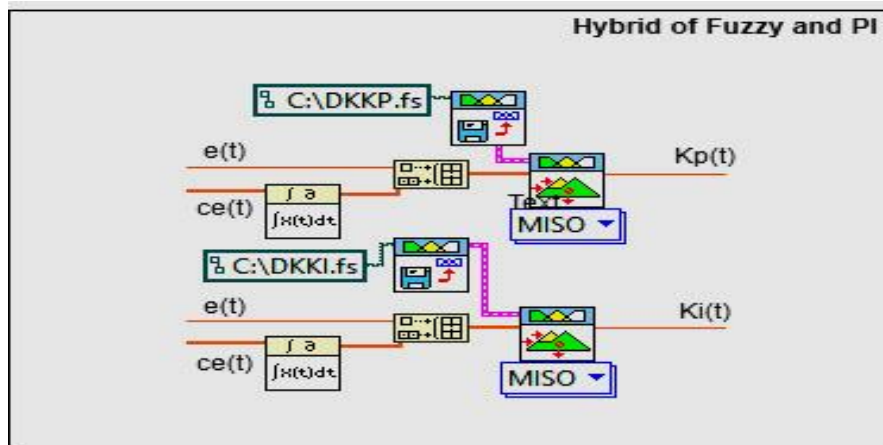
Hiện nay hệ thống điều khiển Fuzzy(mờ) được sử dụng rộng rãi mọi lĩnh vực nghiên cứu và thực tế. Đối với lĩnh vực nghiên cứu được các sinh viên áp dụng lí thuyết của điều khiển mờ trong điều khiển hệ thống.



Hình 3: Cấu trúc bộ điều khiển mờ điện áp ngõ ra

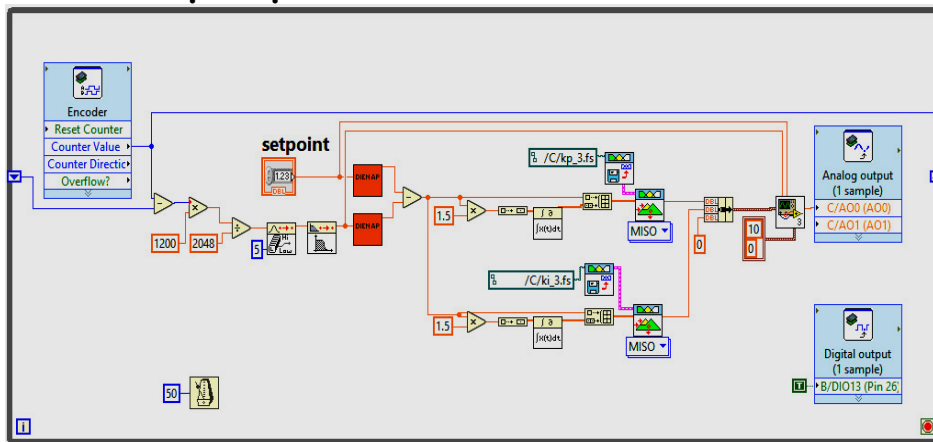
3.2.2. Bộ điều khiển Fuzzy tinh chỉnh thông số cho PI

Bộ điều khiển này được đưa vào cấu trúc điều khiển PI mờ ở Hình 1. Với bộ điều khiển này có đầu vào là $\{e(t):ce(t)\}$ và ngõ ra cung cấp các giá trị thông số K_p, K_i cho bộ điều khiển PI.



Hình 4: Bộ điều khiển mờ tinh chỉnh các thông số cho PI

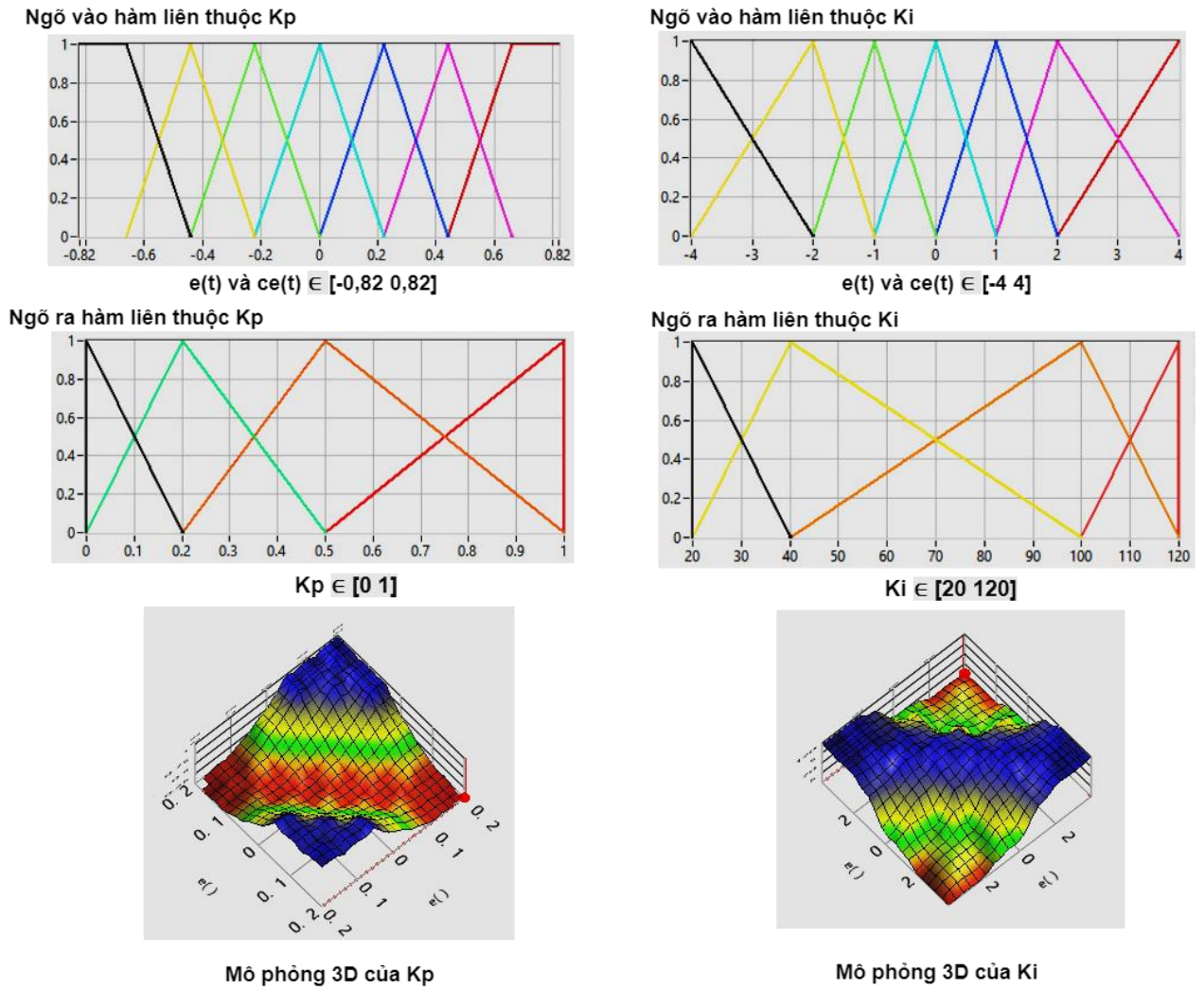
3.2.3. Cách tính và xác định luật mờ



Hình 5: Bộ điều khiển PI mờ trên LabView

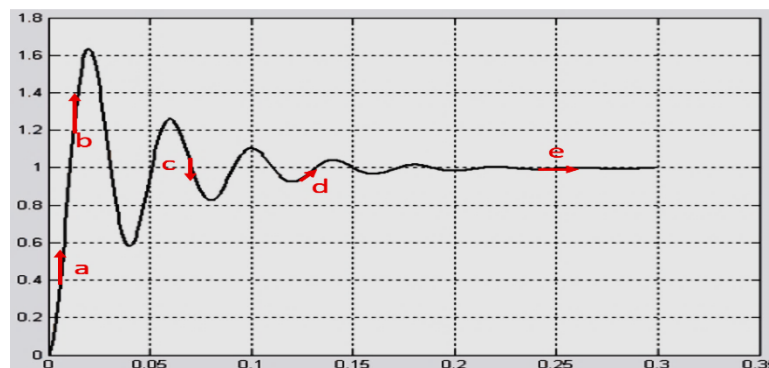
Xác định các biến ngôn ngữ: Ngõ vào $e(t)$ là giá trị sai số và $ce(t)$ là giá trị nguyên hàm sai số, Ngõ ra K_p và K_i . Xác định số lượng tập mờ cho mỗi biến đầu vào là 7 $\{e(t), ce(t)\} = \{RRN, RN, N, TB, L, RL, RRL\}$ và

đầu ra là 4 $K_p = \{TB, L, RL, RRL\}$, $K_i = \{RRN, RN, N, TB\}$. Trong đó RRN là tín hiệu rất rất nhỏ tăng dần đến tín hiệu RRL là tín hiệu rất rất lớn.



Hình 6: Hàm liên thuộc của các thông số K_p và K_i

Xây dựng hàm liên thuộc: dựa vào tính chất của hệ thống PID thì thành phần thông số K_p ảnh hưởng tới quá trình quá độ của hệ thống và thành phần thông số K_i ảnh hưởng đến quá trình xác lập (sai số xác lập) nên bộ điều khiển Fuzzy tinh chỉnh thông số PI được xây dựng thành bộ luật như sau:



Hình 7: Đáp ứng của bộ điều khiển

Tại a thì $e \gg 0$; $ce \gg 0$ nên K_p lớn và K_i nhỏ, Tại b thì $e \ll 0$; $ce \ll 0$ nên K_p lớn và K_i nhỏ, Tại c thì $e > 0$; $ce \{<0; =0; >0\}$ nên K_p nhỏ và K_i lớn, Tại d thì $e > 0$ nhỏ ; $ce > 0$ nhỏ nên K_p giảm dần và K_i tăng dần, Tại e thì $e = 0$; $ce = 0$ nên $K_p = 0$ và K_i lớn

$$\text{If } e = RRL \text{ and } ce = RRL \text{ Then } K_p = RRL \text{ and } K_i = TB \quad (4)$$

Xây dựng các hàm liên thuộc vào ra để điều khiển K_p . Giới hạn ngõ vào/ra bộ điều khiển mờ K_p với $e(t)ce(t) = [-0.82;0,82]$; và $K_p = [0;1]$.

Bảng 2: Bảng luật hợp thành bộ điều khiển thông số K_p

$e(t) \backslash ce(t)$	RRN	RN	N	TB	L	RL	RRL
RRN	RRL	RRL	RRL	RL	L	TB	TB
RN	RRL	RRL	RL	L	TB	TB	TB
N	RRL	RL	L	TB	TB	TB	L
TB	RL	L	TB	TB	TB	L	RL
L	L	TB	TB	TB	L	RL	RRL
RL	TB	TB	TB	L	RL	RRL	RRL
RRL	TB	TB	L	RL	RRL	RRL	RRL

Xây dựng hàm liên thuộc vào ra để điều khiển K_i . Giới hạn ngõ vào/ra bộ điều khiển mờ K_i với $e(t)ce(t) = [-4;4]$; và $K_i = [20;120]$.

Bảng 3: Bảng luật hợp thành bộ điều khiển PI mờ của K_i

$e(t) \backslash ce(t)$	RRN	RN	N	TB	L	RL	RRL
RRN	RRN	RRN	RN	RN	N	TB	TB
RN	RRN	RN	RN	N	TB	TB	TB
N	RN	RN	N	TB	TB	TB	N
TB	RN	N	TB	TB	TB	N	RN
L	N	TB	TB	TB	N	RN	RN
RL	TB	TB	TB	N	RN	RN	RRN
RRL	TB	TB	N	RN	RN	RRN	RRN

Cơ chế suy diễn mờ được chọn là MAX-MIN và giải mờ theo phương pháp trọng tâm, được xác định bởi biểu thức (4):

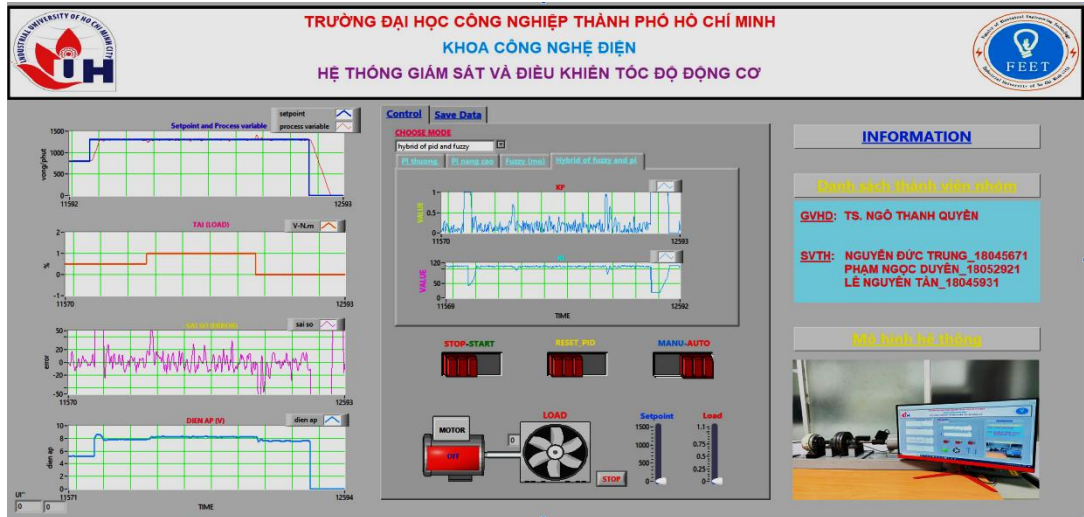
$$\{K_p, K_i\} = \frac{\sum_{i=1}^R b_i \mu_i}{\sum_{i=1}^R \mu_i} \quad (5)$$

Với $\{K_p, K_i\}$ là thông số bộ điều khiển PI sẽ đạt được; b_i và μ_i lần lượt là hoành độ điểm trung bình và giá trị của hàm liên thuộc ngõ ra xác định bởi luật thứ i trong R luật tác động tại thời điểm xem xét.

3.2. Giao diện scada trên labview

SCADA (viết tắt tiếng Anh: Supervisory Control And Data Acquisition) là hệ thống điều khiển giám sát

và thu thập dữ liệu. Nhằm hỗ trợ con người trong quá trình giám sát và điều khiển từ xa. Giao diện gồm một biểu đồ để quan sát từ đó đưa ra phương pháp điều chỉnh thông số cho hệ thống điều khiển. Bảng điều khiển và điều chỉnh các thông số để vận hành hệ thống. Và dùng để lưu trữ dữ liệu thuận tiện cho công tác nghiên cứu đánh giá.



Hình 8: Giao diện Scada của hệ thống

4. KẾT QUẢ THỰC NGHIỆM

4.1. Công thức tính toán các thông số

Ở mô hình này để đánh giá chất lượng các bộ điều khiển thì chúng ta dựa trên 3 tiêu chí sau:

Độ vọt lố: (Percent of Overshoot – POT) là đại lượng đánh giá mức độ vọt lố của hệ thống khi đáp ứng của hệ thống vượt quá giá trị xác lập của nó. Độ vọt lố được tính bằng công thức:

$$POT = \frac{c_{MAX} - c_{xl}}{c_{xl}} \times 100\% \quad (6)$$

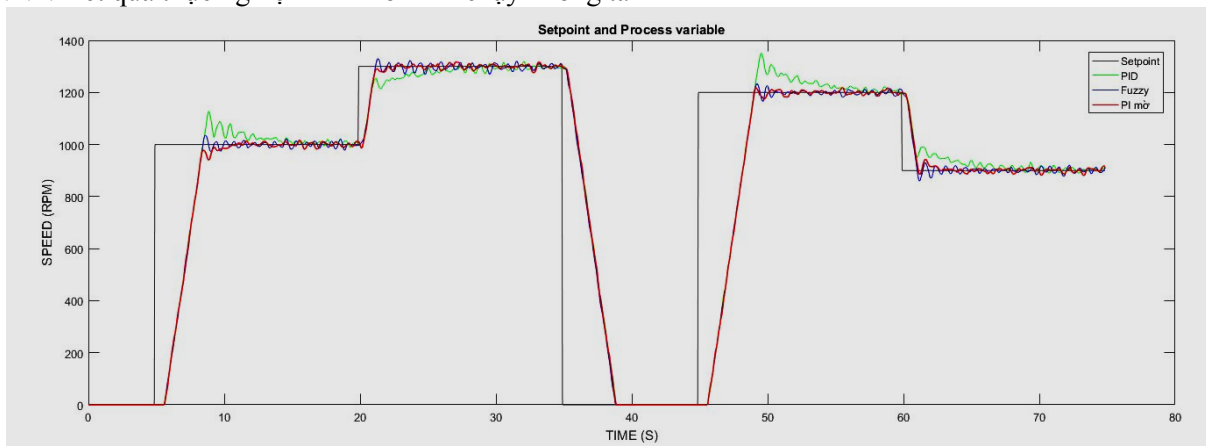
Trong đó: POT là độ vọt lố (%), c_{MAX} là giá trị cực đại của c(t), c_{xl} là giá trị xác lập của c(t). Thời gian xác lập: (T_{xl}) của hệ thống, được xác định khi độ sai số của hệ thống bằng không, ổn định cho đến khi có tác động mới và giá trị xác lập của nó không vượt quá ±5% (0.05). Sai số xác lập: (Error – E_{xl}) là sai số của hệ thống khi thời gian tiến đến vô cùng và được xác định bằng phương pháp MSE (Mean square error).

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - y_i^{\sim})^2 \quad (7)$$

Trong đó: MSE là sai số bình phương trung bình, y_i là giá trị đặt, y_i[~] là giá trị tốc độ trả về theo thời gian.

4.2. Kết quả thực nghiệm mô hình khi chạy không tải và có tải

4.2.1. Kết quả thực nghiệm khi mô hình chạy không tải



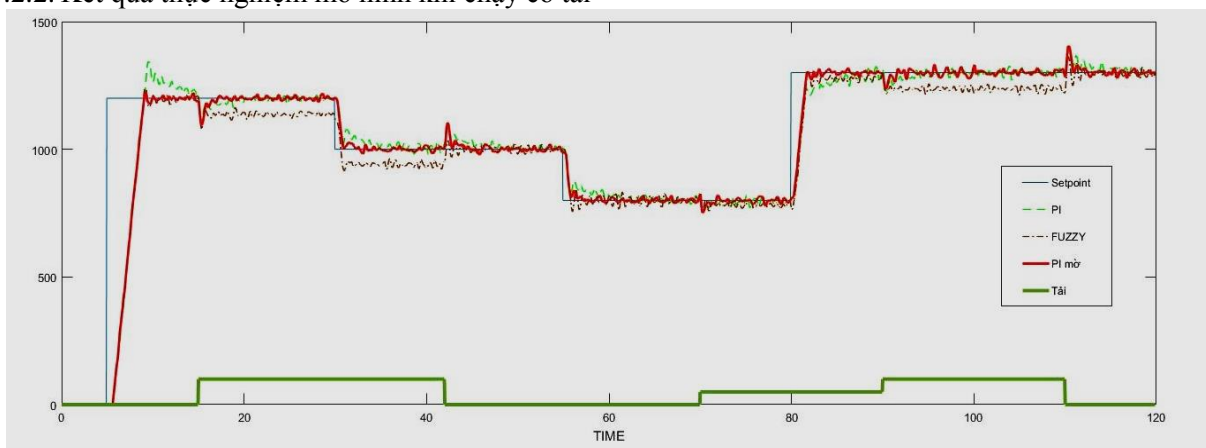
Hình 9: Kết quả thực nghiệm khi mô hình chạy không tải

Bảng 4: So sánh 3 phương pháp ở chế độ khởi động và không tải

	POT (%)	Txl(s)	exl
PID	12.47%	10.05	228.44
FUZZY	3.46%	5	81.48
PI MỜ(Hybrid)	0.64%	5.5	68.85

Dựa vào hình 5.1 và tính toán kết quả ở bảng 5.1 thì khi hệ thống chạy không tải bộ điều khiển PI mờ cho độ vọt lố tốt nhất với $POT = 0.64\%$, khi xác lập thì sai số xác lập được tính theo phương pháp MSE thì bộ điều khiển PI mờ cũng đưa ra kết quả tốt. Thời gian quá độ của hệ thống Fuzzy và PI mờ khá tương đồng nhau với thời gian quá độ khoảng 5 giây, điều này cho thấy hệ thống PI mờ thừa hưởng được những ưu điểm của bộ điều khiển Fuzzy.

4.2.2. Kết quả thực nghiệm mô hình khi chạy có tải

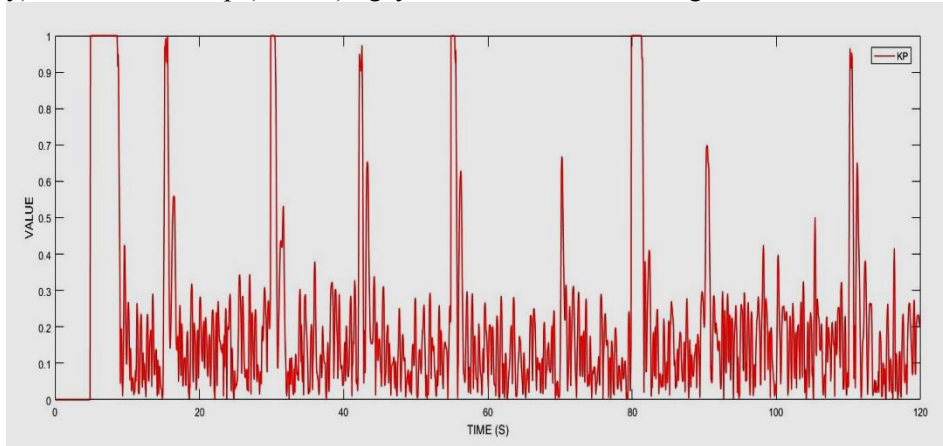


Hình 10: Kết quả thực nghiệm khi mô hình chạy có tải

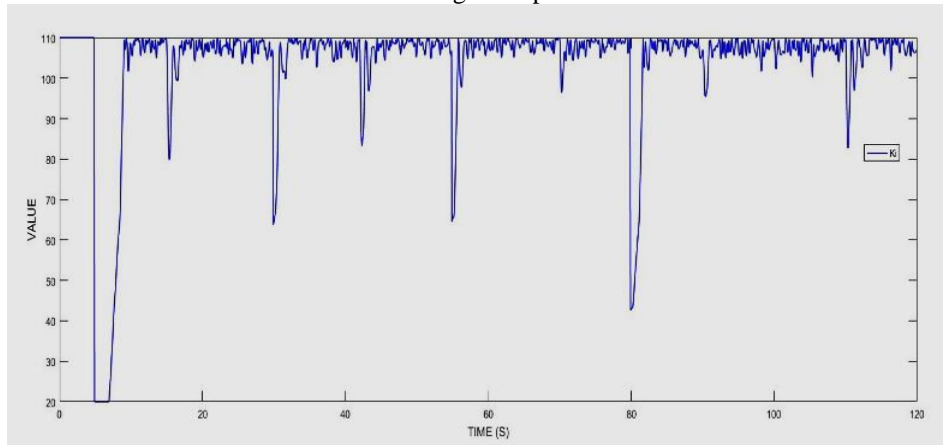
Bảng 5: So sánh 3 phương pháp thời điểm chạy có tải

	POT (%)	$T_{xl}(s)$	e_{xl}
PID	11.96%	9.95	66.3
FUZZY	3.35%	4,85	4055.58
PI MỜ(Hybrid)	2.06%	4.85	40.97

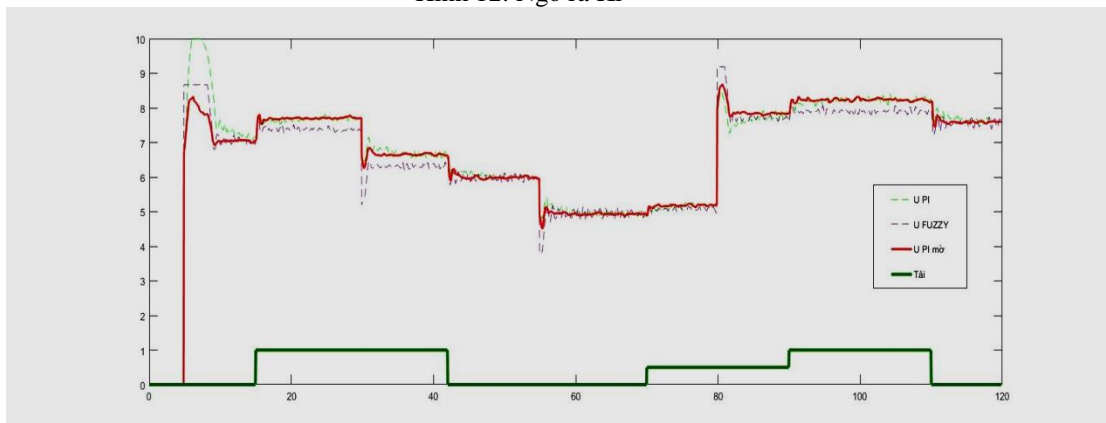
Đối với trường hợp này thì phương pháp điều khiển Fuzzy không còn đáp ứng tốt khi tải thay đổi (sai số xác lập quá lớn). Hệ thống điều khiển PI mờ đáp ứng tốt tất cả các tiêu chí về độ vọt lố (2.06%), thời gian quá độ (4.85 giây) lẫn sai số xác lập (40.97) ngay cả khi có tải hoặc không tải.



Hình 11: Ngõ ra Kp



Hình 12: Ngõ ra Ki



Hình 13: Ngõ ra điện áp của 3 bộ điều khiển

Theo hình 11, hình 12 và hình 13 ta thấy khi sử dụng phương pháp PI mờ khi sử dụng động cơ, với khâu

mờ tinh chỉnh K_p , K_i theo thời gian thì cho ra điện áp cấp cho biến tần điều khiển động cơ ổn định và mượt hơn 2 phương pháp PI và Fuzzy.

5. KẾT LUẬN

Đối với hệ thống điều khiển tốc độ động cơ theo phương pháp PI mờ, dùng luật mờ để chỉnh định hệ số K_p và hệ số K_i của bộ điều khiển PI thì hệ thống cho kết quả tối ưu, hiệu quả hơn bộ điều khiển PI truyền thống và bộ điều khiển Fuzzy. Độ vọt lố của phương pháp này không đáng kể nhỏ hơn so với 2 bộ điều khiển PI và bộ điều khiển Fuzzy. Thời gian xác lập của hệ thống PI mờ được cải thiện tốt hơn bộ điều khiển PI và độ ổn định của phương pháp PI mờ rất cao khi đạt trạng thái xác lập.

Dựa vào những kết quả và đánh giá ở trên thì hệ thống sử dụng phương pháp PI mờ được thừa hưởng những hiệu quả tối ưu từ 2 phương pháp trên, nên việc cho ra đáp ứng là rất tốt với các thông số như độ vọt lố thấp, thời gian xác lập khá nhanh, sai số tĩnh tối ưu và thích nghi tốt khi hệ thống thay đổi tốc độ hoặc cấp tải cho hệ thống. Đối với phương pháp PI: Việc điều khiển cho đáp ứng khá tốt, sai số xác lập tương đối nhưng khi thay đổi điều kiện (tốc độ hoặc tải) thì hệ thống vẫn cho đáp ứng nhưng không tốt với các thông số như độ vọt lố cao hơn, thời gian xác lập lâu và sai số tĩnh lớn hơn so với 2 phương pháp còn lại. Đối với phương pháp Fuzzy: Khi bắt đầu khởi động và điều khiển hệ thống cho ra thời gian xác lập tốt nhất so với 2 phương pháp PI và PI mờ, các thông số như độ vọt lố, thời gian xác lập thì bộ điều khiển PI mờ đạt kết quả tốt nhất so với 2 phương pháp còn lại.

Nhìn chung thì các phương pháp điều khiển đều cho các kết quả đáp ứng tốt cho hệ thống. Còn hạn chế về sai số và độ vọt lố, 2 bộ điều khiển PI, Fuzzy đều đáp ứng không tốt được khi cấp tải. Mặc dù phương pháp PI mờ không phải là phương pháp tốt nhất hiện nay nhưng nó lại khắc phục được những hạn chế của 2 hệ thống còn lại và cho ra kết thực nghiệm tốt nhất trong lần nghiên cứu này. Để khắc phục những hạn chế cũng như tối ưu hơn cho các phương pháp trên cần phải có nhiều thời gian theo dõi, phân tích và chỉnh định các tham số để cho ra được hệ thống tốt hơn và ứng dụng được với đời sống.

TÀI LIỆU THAM KHẢO

- [1] Hệ thống điều khiển thông minh, Huỳnh Thái Hoàng, Nhà xuất bản Đại học Quốc Gia Thành phố Hồ Chí Minh, 2014.
- [2] Nguyễn Chí Ngôn. Tối ưu hóa bộ điều khiển PID bằng giải thuật di truyền. Tạp chí Khoa học Đại học Cần Thơ, số 2008:9, trang 241-248, 2008.
- [3] Tống Ngọc Minh. Luận văn Sử dụng LabVIEW điều khiển bám đuổi mức nước. Đại học Công nghiệp Thành phố Hồ Chí Minh, 2015.
- [4] Lý thuyết điều khiển tự động, Nguyễn Thị Phương Hà, Huỳnh Thái Hoàng, Nhà xuất bản Đại học Quốc gia Thành phố Hồ Chí Minh, 2016.
- [5] Nguyễn Văn Thiện. Nghiên cứu thiết kế bộ điều khiển pid mờ, Số hóa bởi Trung tâm Học liệu – Đại học Thái Nguyên, 2010.
- [6] Đặng Đức Công, Đặng Thái Giáp. Thiết kế hệ thống điều khiển tốc độ động cơ DC sử dụng bộ điều khiển PID và PID mờ (Fuzzy-PID). Đại học Bách khoa Hà Nội, 2013.
- [7] Salami, M. and G. Cain. An adaptive PID controller based on genetic algorithm processor. IEEE Conf. Publ. No. 414, 12-14 Sep 1995, pp. 88-93, 2012.

YSC4F.237

PHÂN TÍCH VÀ THIẾT KẾ: FEATURE STORE

BUI THI HUYEN TRANG¹, CU THI CAM LY¹, HO VO HOANG DUY¹, NGUYEN QUANG BAO¹,
ĐOAN ANH DUONG¹, TRAN HUU NHAN²

¹ Industrial University of Ho Chi Minh City,

²One Mount Group

20122031.trang@student.iuh.edu.vn, 19521291.ly@student.iuh.edu.vn,

20087541.duy@student.iuh.edu.vn, 20083601.bao@student.iuh.edu.vn,

20082631.duong@student.iuh.edu.vn, nhantran.est@gmail.com

Tóm tắt. Trong nghiên cứu này, chúng tôi mô tả hệ thống Feature Store giúp thúc đẩy, rút gọn quy trình xử lý dữ liệu cho Data Scientists. Các mô hình Machine Learning cần một bộ dữ liệu sạch để có thể đưa ra những dự đoán và kết quả chính xác hơn. Mục đích nghiên cứu là từ bộ dữ liệu gốc (raw data) thông qua các thao tác xử lý dữ liệu (feature pipelines) để tạo ra bộ dữ liệu sạch cho người dùng có thể tùy ý sử dụng. Mặc dù Pipeline có thể giúp các Data Scientist thử nghiệm với các mô hình Machine Learning, nhưng sẽ có xu hướng trở nên phức tạp theo thời gian và khó quản lý. Điều này thúc đẩy việc sử dụng Feature Store được tiêu chuẩn hóa cho quy trình Data Cleansing, Data Preprocessing và Features Engineering, giúp giảm chi phí xử lý dữ liệu cho các mô hình dự đoán bằng cách chia sẻ dữ liệu ở trạng thái “feature” giữa các machine learning model. Chúng tôi đã thực hiện việc phân tích dữ liệu và thiết lập các yêu cầu kỹ thuật. Khi Feature Store được xây dựng với nhiều Feature hơn, các Feature được sử dụng lại trên các dự án Machine Learning, nó giúp việc xây dựng các mô hình mới một cách dễ dàng và nhanh chóng.

Từ khoá. Data Cleansing, Data Preprocessing, Features Engineering, Feature Pipelines, Feature Store.

ANALYZE AND DESIGN: FEATURE STORE

Abstract. In this study, we describe the Feature Store system that helps to accelerate and shorten the data processing process for Data Scientists. Machine Learning models need a clean data set to be able to make more accurate predictions and outcomes. The research purpose is from the original dataset (raw data) through data processing operations (feature pipelines) to create a clean dataset for the user that can be used optionally. Though Pipelines can help Data Scientists to experiment with Machine Learning models, they tend to become complex over time and difficult to manage. This promotes the use of a standardized Feature Store for the data cleansing process (data cleansing), the procedure before data processing (data preprocessing), and features engineering, therefore, help data processing costs for predictive models by sharing "feature" data among machine learning models. We performed the data analysis and established the technical requirements. When a FS with more features undergoes a construction, such new features, which boost the productivity of building new models, are applied on the ML project.

Keywords. Data Cleansing, Data Preprocessing, Features Engineering, Feature Pipelines, Feature Store.

1. GIỚI THIỆU

Ngày nay dữ liệu được coi là vàng trong kỷ nguyên số, “data makes the world go around” (dữ liệu khiến thế giới chuyển động). Mọi hệ thống dù thông minh đến đâu thì cũng cần dữ liệu để có thể vận hành, tuy nhiên các thuật toán hiện nay đều không làm việc được với dữ liệu thô. Vì thế các bước xử lý dữ liệu thô là vô cùng quan trọng để các thuật toán học máy có thể làm việc hiệu quả với các dữ liệu được cung cấp. Với mục đích có thể tạo được một công cụ hỗ trợ Machine Learning trong bước xử lý dữ liệu, chúng tôi đã thực

hiện đề tài này nhằm tạo một Feature store hoàn chỉnh, thực hiện được cơ bản các bước xử lý thông tin thô để tạo một cơ sở dữ liệu cơ bản có thể sử dụng được. Để thực hiện được mục tiêu đó, chúng tôi đã dùng tìm hiểu và dùng các kỹ thuật như split data một cách random, kỹ thuật upload dữ liệu lên bigquery, kỹ thuật data cleaning, data processing, feature engineering. Đã có một nghiên cứu nghiên quan trước đó giới thiệu về một số Feature Store được thực hiện bởi Viện nghiên cứu về Điện tử và Viễn thông Hàn Quốc nhưng việc mô tả chi tiết về phương pháp thực hiện và kết quả nghiên cứu chưa được giới thiệu cụ thể, vì vậy nhóm chúng tôi đã tiến hành nghiên cứu này với mong muốn làm rõ quy trình của một Feature Store đồng thời bàn luận về những lợi ích của Feature Store đem lại [1]. Kết quả đạt được sau quá trình nghiên cứu chính là thiết kế được một Feature Store.

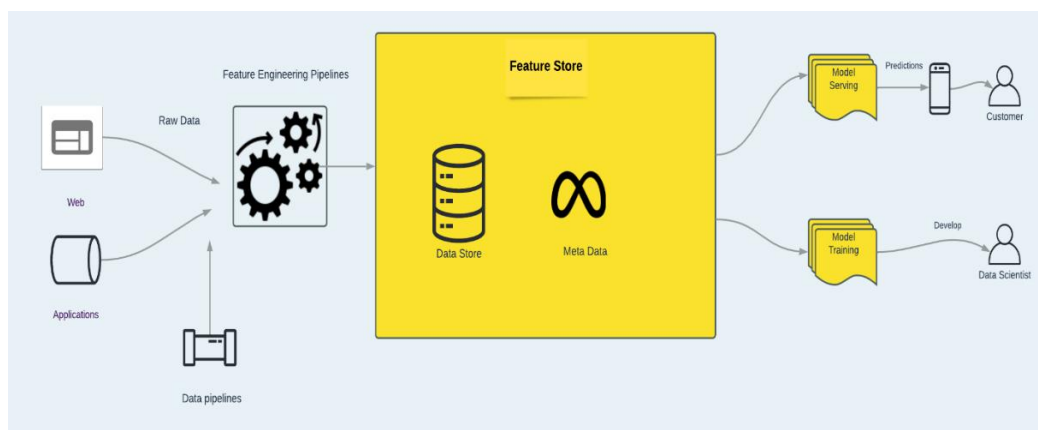
2. CÁC NGHIÊN CỨU LIÊN QUAN

Hiện nay, các dự án Máy học (ML) trong ngành phải lấy từ nhiều nguồn dữ liệu khác nhau thường làm phức tạp quá trình phân tích dữ liệu. Xử lý dữ liệu lộn xộn, chuyển đổi sang các định dạng có thể sử dụng, trích xuất tính năng và kỹ thuật mất ~70% thời gian phát triển. Việc đảm bảo tính nhất quán thường là một thách thức do sự chênh lệch giữa cơ sở hạ tầng phát triển và sản xuất. Ngoài ra, khó có thể tái sử dụng các tính năng này trên các dự án và nhóm. Đến thời điểm này có nhiều nghiên cứu liên quan cùng chủ đề như của Sukant Kumar và cộng sự “Offline and Online Feature Store for faster and consistent Machine Learning Modelling in Wellness Domain” vào ngày 7 tháng 12 năm 2021 trong nghiên cứu này họ sử dụng Snowflake và DBT cho Feature Store tính năng ngoại tuyến và sử dụng Amazon DynamoDB cho Feature Store tính năng trực tuyến và cung cấp các feature trong thời gian thực và hiệu quả mang lại là giảm ít nhất 50% thời gian xử lý dữ liệu của các nhà phát triển. Kết quả khi sử dụng Feature Store là thông thường phải quét qua 30.9GB dữ liệu mất khoảng 323 giây nhưng quét 559.8MB bằng Feature Store thì mất khoảng 4.2 giây giảm được 95% thời gian chạy cho một dự án. Do đó, dự án này dự kiến hoàn thành trong hai quý nhưng chỉ trong một quý đã hoàn thành [1].

3. PHƯƠNG PHÁP

Trong mô hình này, từ dữ liệu ban đầu chúng tôi tiến hành tách dữ liệu thành 2 bộ dữ liệu khác nhau 1 bộ 90% của bộ dữ liệu gốc và 1 bộ 10% của bộ dữ liệu gốc. Trong quá trình tách dữ liệu một cách chúng tôi sử dụng sklearn.model_selection.train_test_split() của thư viện Scikit learn để tách dữ liệu. Chúng tôi sử dụng 90% dữ liệu để xây dựng mô hình feature store và 10% dữ liệu còn lại dùng để kiểm tra và đánh giá mô hình.

Data cleaning thực hiện sửa, xóa dữ liệu không chính xác, dữ liệu bị hỏng, không đúng định dạng, dữ liệu bị trùng lặp, hoặc không đầy đủ trong tập dữ liệu. Nếu bộ dữ liệu không tốt sẽ làm giảm hiệu suất của mô hình trong quá trình xử lý và phân tích. Bộ dữ liệu Superstore khá sạch vì vậy không cần phải qua nhiều công đoạn Data cleaning. Với bộ dữ liệu hiện tại, chúng tôi đã thực hiện làm sạch dữ liệu (data cleaning) loại bỏ các dữ liệu bị thiếu góp phần làm tăng độ chính xác của dữ liệu, đẩy nhanh tốc độ và hiệu quả của quá trình phân tích dữ liệu.



Hình 1. Quy trình Feature Store

Data processing tiến hành xử lý dữ liệu từ dữ liệu thô ban đầu thành dữ liệu có thể sử dụng được. Quá trình này được thực hiện bởi các nhà khoa học dữ liệu. Dữ liệu thô ban đầu được thu thập, sàng lọc, sắp xếp, xử lý, phân tích, lưu trữ và sau đó được trình bày ở định dạng có thể đọc được [2]. Với phần processing thì trong bảng “raw data” chúng tôi thực hiện thao tác thêm một cột có tên là “is_processed” và cột này có giá trị là true/false nếu dữ liệu mới ghi vào bảng “raw data” thì mặc định is_processed = false nếu dữ liệu đã được processed bằng code python thì cột is_processed sẽ cập nhật lại là is_processed = true.

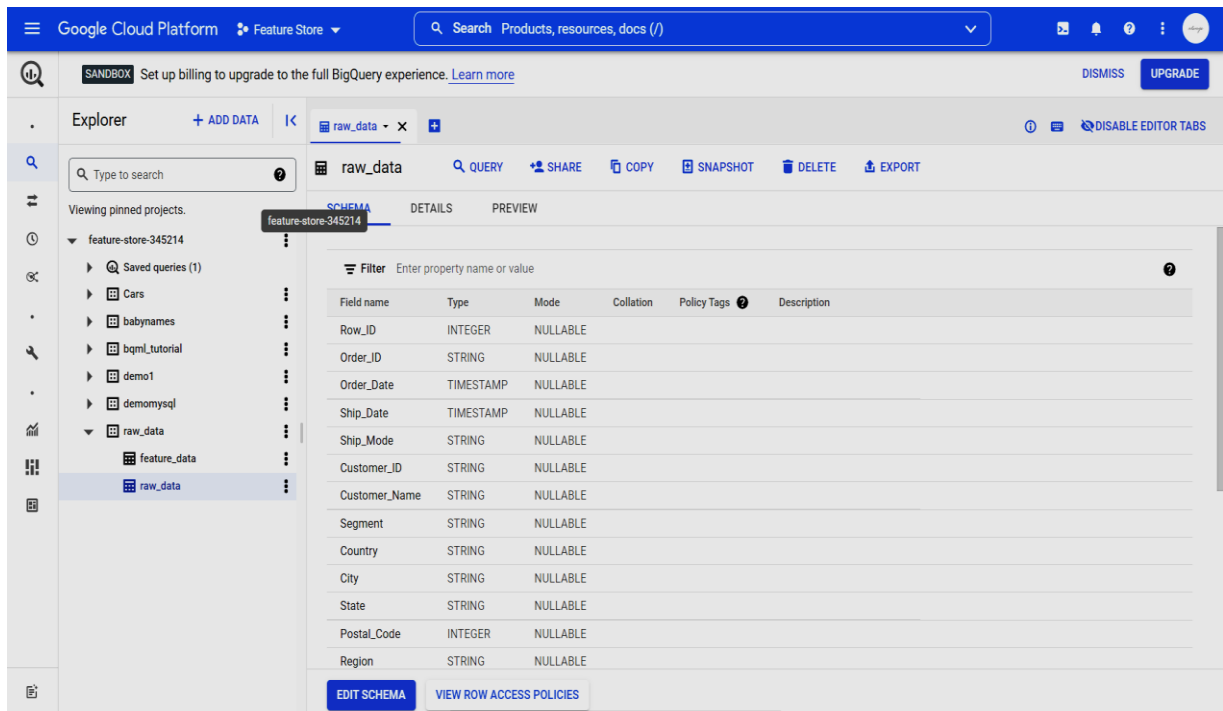
Feature engineering là quá trình chuẩn hóa bộ dữ liệu gốc (raw data) thành các tập các thuộc tính có ý nghĩa, giúp biểu diễn tập dữ liệu ban đầu một cách tốt hơn. Giúp giải quyết các bài toán một cách dễ dàng. Trong máy học việc có một bộ dữ liệu đã được chuẩn hóa giúp cải thiện đáng kể độ chính xác của mô hình. Feature engineering sẽ biểu diễn tập dữ liệu ban đầu một cách tốt nhất sao cho nó có thể tương thích nhất với mô hình hiện tại của chúng ta. Khi cần giải quyết một bài toán nào đó việc tập dữ liệu thô ban đầu chứa rất nhiều những thông tin và các thể có vài cột thông tin không quan trọng và không liên quan đến bài toán hiện tại của chúng ta. Nhưng nếu đưa hết chúng vào mô hình để training thì mô hình sẽ chạy rất chậm và cho ra độ chính xác không cao. Thay vì thế chúng ta sử dụng kỹ thuật feature engineering để có thể đưa ra các thuộc tính phù hợp và cần thiết nhất của bài toán, ngoài ra có thể thêm một vài thuộc tính cần thiết khác để bổ sung vào tập dữ liệu giúp giải quyết bài toán tốt hơn. Kết quả và độ chính xác của mô hình bị ảnh hưởng trực tiếp bởi tập thuộc tính của bộ dữ liệu, do đó ta cần xác định tốt cấu trúc của tập thuộc tính sao cho diễn đạt một cách hiệu quả nhất bản chất của bộ dữ liệu ban đầu. Việc có được một bộ dữ liệu tinh gọn sẽ làm đơn giản hóa độ phức tạp tính toán của mô hình nhờ thế mà tính toán sẽ được tối ưu mô hình sẽ chạy nhanh hơn và cho ra độ chính xác cao hơn. [3, 4, 5, 6]

BigQuery là một kho dữ liệu trên nền tảng đám mây của Google (Google Cloud Platform) cho phép lưu trữ, chạy các truy vấn trên các tập dữ liệu lớn. Google BigQuery có khả năng mở rộng và dễ sử dụng, cho phép các nhà phát triển khai thác và xử lý dữ liệu mạnh mẽ. Và Google BigQuery là một trong những nền tảng chạy các truy vấn SQL cực nhanh, giúp dễ dàng kết hợp tập dữ liệu công khai, dữ liệu thương mại với tập dữ liệu của chúng ta. BigQuery còn cung cấp một nền tảng linh hoạt cho machine learning và trí tuệ nhân tạo. BigQuery tích hợp với CloudML Engine và TensorFlow để đào tạo các mô hình mạnh mẽ trên tập dữ liệu có cấu trúc.

Trong nghiên cứu này chúng tôi sử dụng 90% dữ liệu của bộ dữ liệu gốc (Superstore) lưu vào BigQuery với tên bảng là raw_data. 10% dữ liệu còn lại chúng tôi thực hiện upload dữ liệu lên. BigQuery bằng phương thức load_table_from_file. Sau khi dữ liệu đã được xử lý (processed) toàn bộ dữ liệu vừa xử lý sẽ được thêm vào một bảng mới trên BigQuery với tên bảng là feature_data và lưu trữ trên Bigquery.

Trên nền tảng điện toán đám mây của Google (Google Cloud) chúng tôi sử dụng kho dữ liệu BigQuery để truy vấn tập dữ liệu Superstore. Chúng tôi tiến hành tạo 2 bảng ‘raw_data’ để lưu trữ dữ liệu thô ban đầu và bảng ‘feature_data’ để lưu trữ dữ liệu sau khi xử lý. Sau đó, trong bảng ‘raw_data’ chúng tôi thêm một cột ‘is_processed’ cột này có giá trị là True/ False. Nếu dữ liệu mới được ghi vào bảng ‘raw_data’ thì ‘is_processed’ có giá trị là False. Nếu dữ liệu đã được xử lý bằng code Python thì cột ‘is_processed’ sẽ được cập nhật giá trị thành True.

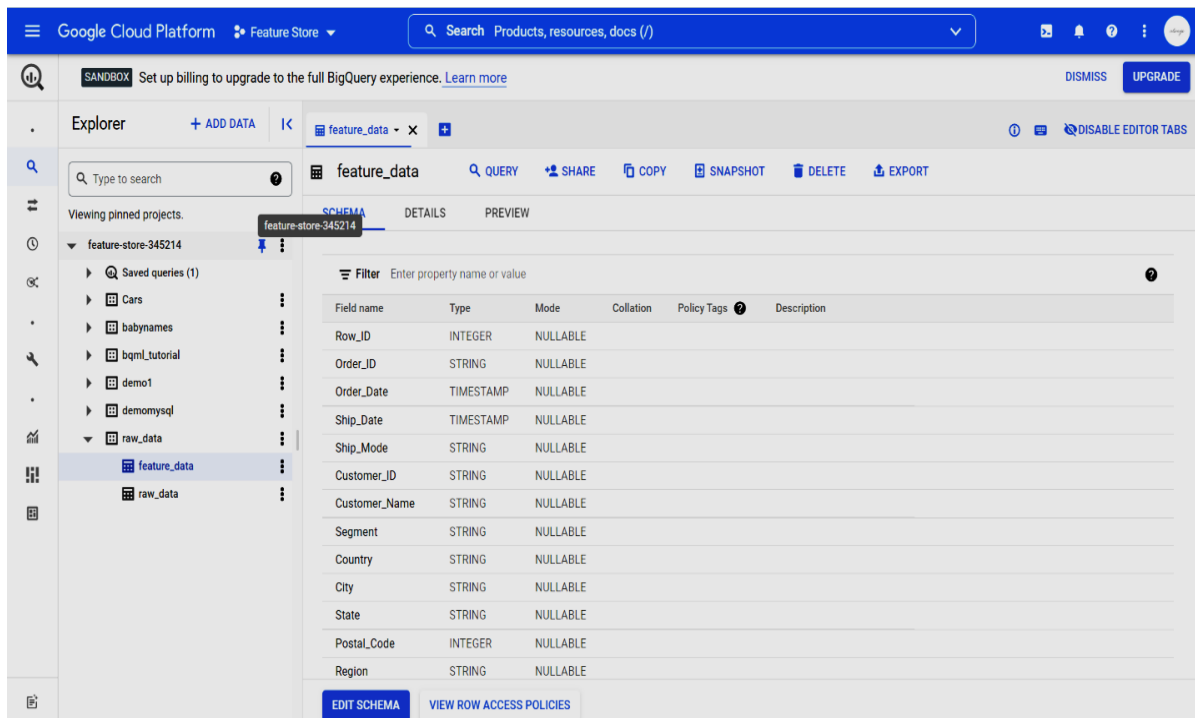
Trong nghiên cứu này từ dữ liệu ban đầu (‘raw_data’) chúng tôi sử dụng kỹ thuật xây dựng đặc trưng mới (feature construction) để phân tách ngày, tháng, năm của thời gian đặt hàng (‘Order Date’) và thời gian giao hàng (‘Ship Date’) thành các cột riêng biệt lần lượt là: ‘Order Date Day’, ‘Order Date Month’, ‘Order Date Year’, ‘Ship Day’, ‘Ship Month’, ‘Ship Year’. Và chúng tôi tiến hành tạo một cột mới cột này mô tả thời gian từ khi đặt hàng đến lúc giao hàng ‘Time To Deliver’ cột này được xác định bằng cách lấy ngày giao hàng trừ cho ngày đặt hàng (Ship Day – Order Date Day). Sau khi có ‘Time To Deliver’ chúng tôi loại bỏ tất cả các giá trị âm và xác định phương thức giao hàng ‘Shipping Method’ bằng các thông số của cột ‘Time To Deliver’. Nếu ‘Time To Deliver’ nhỏ hơn hoặc bằng 2 thì phương thức giao hàng sẽ là “giao hàng hỏa tốc” (“Express delivery”), nếu ‘Time To Deliver’ lớn hơn 2 và nhỏ hơn bằng 4 thì phương thức giao hàng là “giao hàng nhanh” (“Fast delivery”), ‘Time To Deliver’ lớn hơn 4 phương thức giao hàng sẽ là “giao hàng tiết kiệm” (“Economical delivery”). Việc sử dụng kỹ thuật này giúp chúng ta có thể dễ dàng quan sát, đánh giá thời gian khách hàng có thể nhận được hàng của mình kể từ khi đặt hàng. Từ đó biểu diễn tập dữ liệu ban đầu một cách rõ ràng, chính xác hơn.



The screenshot shows the Google Cloud Platform Feature Store interface. The 'raw_data' table schema is displayed with the following fields:

Field name	Type	Mode	Collation	Policy Tags	Description
Row_ID	INTEGER	NULLABLE			
Order_ID	STRING	NULLABLE			
Order_Date	TIMESTAMP	NULLABLE			
Ship_Date	TIMESTAMP	NULLABLE			
Ship_Mode	STRING	NULLABLE			
Customer_ID	STRING	NULLABLE			
Customer_Name	STRING	NULLABLE			
Segment	STRING	NULLABLE			
Country	STRING	NULLABLE			
City	STRING	NULLABLE			
State	STRING	NULLABLE			
Postal_Code	INTEGER	NULLABLE			
Region	STRING	NULLABLE			

Hình 2. Bảng 'raw_data'



The screenshot shows the Google Cloud Platform Feature Store interface. The 'feature_data' table schema is displayed with the following fields:

Field name	Type	Mode	Collation	Policy Tags	Description
Row_ID	INTEGER	NULLABLE			
Order_ID	STRING	NULLABLE			
Order_Date	TIMESTAMP	NULLABLE			
Ship_Date	TIMESTAMP	NULLABLE			
Ship_Mode	STRING	NULLABLE			
Customer_ID	STRING	NULLABLE			
Customer_Name	STRING	NULLABLE			
Segment	STRING	NULLABLE			
Country	STRING	NULLABLE			
City	STRING	NULLABLE			
State	STRING	NULLABLE			
Postal_Code	INTEGER	NULLABLE			
Region	STRING	NULLABLE			

Hình 3. Bảng 'feature_data'

```
def process(data):
    data['Order_Date'] = pd.to_datetime(data['Order_Date'])
    data['Ship_Date'] = pd.to_datetime(data['Ship_Date'])

    data['Order_Date_Month'] = data['Order_Date'].dt.month
    data['Order_Date_Day'] = data['Order_Date'].dt.day
    data['Order_Date_Year'] = data['Order_Date'].dt.year
    data['Ship_Month'] = data['Ship_Date'].dt.month
    data['Ship_Day'] = data['Ship_Date'].dt.day
    data['Ship_Year'] = data['Ship_Date'].dt.year
    data['Time_To_Deliver'] = data['Ship_Day'] - data['Order_Date_Day']
    data.drop(data.loc[data['Time_To_Deliver'] < 0].index, inplace=True)
    data.loc[data['Time_To_Deliver'] <= 2, 'Shipping_Method'] = 'Express delivery'
    data.loc[(data['Time_To_Deliver'] > 2) & (data['Time_To_Deliver'] <= 4), 'Shipping_Method'] = 'Fast delivery'
    data.loc[data['Time_To_Deliver'] > 4, 'Shipping_Method'] = 'Economical delivery'
```

Hình 4. Mô tả kĩ thuật xây dựng đặc trưng mới (feature construction)

Bảng dưới đây chúng tôi trích xuất từ 6 dòng dữ liệu đầu trên trong bảng ‘raw_data’. Trong thực tế sau khi thực thi bảng sẽ có đến 881 dòng và có đầy đủ hết các phương thức ship hàng. Chúng tôi tạo một cột mới cột này mô tả thời gian từ khi đặt hàng đến lúc giao hàng ‘Time To Deliver’ cột này được xác định bằng cách lấy ngày giao hàng trừ cho ngày đặt hàng (Ship Day – Order Date Day). Bảng này giúp chúng tôi xác định được phương thức giao hàng và thời gian từ khi order đến khi nhận được hàng.

Bảng 1. Kết quả phân tách cột ‘Order Date’ và ‘Ship Date’

Order Date	Ship Date	Order Date Month	Order Date Day	Order Date Year	Ship Month	Ship Day	Ship Year	Time To Deliver	Shipping Method
2015-07-20 00:00:00+00:00	2015-07-20 00:00:00+00:00	7	20	2015	7	20	2015	0	Express delivery
2015-04-07 00:00:00+00:00	2015-04-07 00:00:00+00:00	4	7	2015	4	7	2015	0	Express delivery
2015-08-23 00:00:00+00:00	2015-08-23 00:00:00+00:00	8	23	2015	8	23	2015	0	Express delivery
2015-09-14 00:00:00+00:00	2015-09-14 00:00:00+00:00	9	14	2015	9	14	2015	0	Express delivery
2015-11-26 00:00:00+00:00	2015-11-26 00:00:00+00:00	11	26	2015	11	26	2015	0	Express delivery
2015-11-26 00:00:00+00:00	2015-11-26 00:00:00+00:00	11	26	2015	11	26	2015	0	Express delivery

4. KẾT QUẢ

Sau khi qua các bước xử lý bộ dữ liệu ban đầu sẽ trở nên sạch và chi tiết hơn. Dữ liệu được xử lý cũng cho ra kết quả chính xác và rõ ràng hơn. Từ đó các nhà khoa học dữ liệu có thể sử dụng ngay bộ dữ liệu để tiến hành xây dựng các mô hình theo mong muốn giúp tiết kiệm thời gian và công sức. Họ không cần phải lãng phí thời gian và sức lực của mình vào các quy trình xử lý dữ liệu dài dòng và phức tạp. Sau khi dữ liệu qua các quy trình xử lý chúng tôi có một bộ dữ liệu trực quan và các thông tin được thể hiện rõ ràng hơn.

Từ bộ dữ liệu thô ban đầu chúng tôi tải dữ liệu lên bảng ‘raw_data’ (BigQuery), sau khi tiến hành các quy trình xử lý dữ liệu, đầu tiên là Data cleaning (làm sạch dữ liệu) chúng tôi tiến hành sửa, xóa các dữ liệu không chính xác. Tiếp đến là Data processing (xử lý dữ liệu) trong phần này chúng tôi thêm cột “is_processed” có giá trị là False khi dữ liệu mới được đưa vào bảng và True sau khi dữ liệu đã được xử lý. Feature engineering phân tách các cột ngày, tháng, năm đặt hàng và giao hàng để từ đó rút ra được phương

thức giao hàng và đánh giá xem đó là giao hàng nhanh hay chậm. Để cho ra bộ dữ liệu sạch và trực quan, chúng tôi tiến hành cập nhật toàn bộ dữ liệu đã qua các bước xử lý vào bảng ‘feature_data’ (BigQuery) cột “is processed” trong bảng ‘feature_data’ luôn có giá trị là True và bộ dữ liệu này có thể được sử dụng ngay. Bộ dữ liệu ban đầu bao gồm 9994 dòng và 21 cột thể hiện doanh số bán hàng. Kết quả sau khi được xử lý bộ dữ liệu mới sẽ được cập nhật vào bảng ‘feature_data’ trên BigQuery, bộ dữ liệu sau khi xử lý có 30 cột và cột ‘is_processed’ sẽ được cập nhật thành True. Bộ dữ liệu đã qua xử lý sẽ và cho kết quả một cách rõ ràng và trực quan hơn. Từ đó giúp cho việc sử dụng bộ dữ liệu để xây dựng mô hình học máy một cách dễ dàng và chính xác hơn.

Bảng 2: Dữ liệu sau khi được cập nhật trên bảng ‘feature_data’

DATA PROCESSED !!																													
Row ID	Order ID	Order Date	Ship Date	Ship State	Customer ID	Customer Name	Segment	Country	City	State	Postal Code	Region	Product ID	Category	Sub Category	Product Name	Sales	Quantity	Discount	Profit	Is processed	Order Date Month	Order Date Day	Order Date Year	Ship Month	Ship Day	Ship Year	Time To Deliver	Shipping Method
4395	CA-2015-142665	2015-07-20 00:00:00-00:00:00	2015-07-20 00:00:00-00:00:00	Same Day	578-20245	Scott Thomas	Consumer	United States	Springfield	Ohio	45503	East	OFF-EP-10002393	Office Supplies	Storage	FollowUp Your Sites Storage Cubes	25.994	1	0.2	-5.1968	True	7	20	2015	7	20	2015	0	Express delivery
8223	CA-2015-185906	2015-06-07 00:00:00-00:00:00	2015-06-07 00:00:00-00:00:00	Same Day	734-21485	Valerie Mackinn	Consumer	United States	Marietta	Georgia	30062	South	OFF-PA-10004841	Office Supplies	Paper	Xerox 226	58.32	9	0.0	27.9926	True	4	7	2015	4	7	2015	0	Express delivery
2329	CA-2015-110639	2015-08-23 00:00:00-00:00:00	2015-08-23 00:00:00-00:00:00	Same Day	836-19405	Rick Hanson	Consumer	United States	New York City	New York	10009	East	OFF-PA-10004930	Office Supplies	Paper	Personal Creatives Ink Jet Carbs and Labels	45.92	4	0.0	23.9008	True	8	15	2015	8	15	2015	0	Express delivery
3127	CA-2015-186766	2015-08-14 00:00:00-00:00:00	2015-08-14 00:00:00-00:00:00	Same Day	DNA-13015	Diana Martin	Consumer	United States	New York City	New York	10009	East	TEC-MS-10009379	Technology	Machine	Apple T4110ZDD Micro-Cut Stapler	2799.96	4	0.0	1371.9804	True	9	14	2015	9	14	2015	0	Express delivery
3818	CA-2015-124710	2015-11-26 00:00:00-00:00:00	2015-11-26 00:00:00-00:00:00	Same Day	LS-17245	Lynn Smith	Consumer	United States	Utica	New York	12501	East	OFF-BC-10001033	Office Supplies	Binders	PowerPoint Dine Binders by Wilson Jones	17.088	4	0.2	5.7672	True	11	26	2015	11	26	2015	0	Express delivery
3819	CA-2015-124730	2015-11-26 00:00:00-00:00:00	2015-11-26 00:00:00-00:00:00	Same Day	LS-17245	Lynn Smith	Consumer	United States	Utica	New York	12501	East	OFF-LA-10001128	Office Supplies	Labels	Avery Address Shipping Labels for Typewriters, 4" x 2"	93.15	9	0.0	44.712	True	11	26	2015	11	26	2015	0	Express delivery
767	US-2015-120740	2015-06-13 00:00:00-00:00:00	2015-06-13 00:00:00-00:00:00	Same Day	PS-18970	Paul Stevenson	Home Office	United States	Los Angeles	California	90049	West	OFF-AP-10002040	Office Supplies	Appliances	Stapler P9000V10-GET 9 Outer Stapler	106.96	2	0.0	31.0384	True	4	15	2015	4	15	2015	0	Express delivery
9275	US-2015-134054	2015-10-10 00:00:00-00:00:00	2015-10-10 00:00:00-00:00:00	Same Day	FC-14035	Fred Chang	Corporate	United States	Altamonte Springs	New Mexico	87105	West	OFF-AP-10000937	Office Supplies	Art	Delta Touchdowns Core-Lock Colored Pencils, 48-Color Set	255.85	7	0.0	112.574	True	10	10	2015	10	10	2015	0	Express delivery
3822	CA-2015-142769	2015-08-13 00:00:00-00:00:00	2015-08-13 00:00:00-00:00:00	Same Day	BP-18980	Roni Pflanzig	Consumer	United States	Seattle	Washington	98103	West	FUR-PC-10001914	Furniture	Furniture	C-Line Office Keepers Polypropylene Roller With Velcro Backings	14.19	3	0.0	5.5341	True	9	13	2015	9	13	2015	0	Express delivery
7677	CA-2015-142911	2015-10-08 00:00:00-00:00:00	2015-10-08 00:00:00-00:00:00	Same Day	JK-11215	Jack Geria	Consumer	United States	Charlotte	North Carolina	28205	South	TEC-AC-10002331	Technology	Accessories	Manali 74 Minutes CD-R, 10-Pack	23.472	3	0.2	4.9978	True	10	8	2015	10	8	2015	0	Express delivery
7611	CA-2015-103958	2015-07-23 00:00:00-00:00:00	2015-07-23 00:00:00-00:00:00	First Class	AG-10270	Alphadiah Greiv	Consumer	United States	New York City	New York	10011	East	TEC-AC-10003314	Technology	Accessories	Cas Logic 2-4GHz Wireless Keyboard	99.98	2	0.0	7.9994	True	7	23	2015	7	24	2015	1	Express delivery
3131	CA-2015-136099	2015-11-24 00:00:00-00:00:00	2015-11-24 00:00:00-00:00:00	First Class	CC-12200	Chad Cunningham	Home Office	United States	Los Angeles	California	90049	West	TEC-AC-10001142	Technology	Accessories	First Data PD010 DVD Pad	39.5	5	0.0	39.5	True	12	16	2015	12	17	2015	1	Express delivery
4962	CA-2015-156817	2015-03-27 00:00:00-00:00:00	2015-03-27 00:00:00-00:00:00	First Class	AB-10015	Aaron Bergs	Consumer	United States	Seattle	Washington	98103	West	FUR-CE-10000477	Furniture	Furniture	Global Push Button Manager's Chair, Judge	48.712	1	0.2	5.4801	True	3	7	2015	3	8	2015	1	Express delivery
117	CA-2015-112229	2015-06-25 00:00:00-00:00:00	2015-06-25 00:00:00-00:00:00	Second Class	RK-19940	Ryan Crews	Consumer	United States	Columbus	Ohio	43229	East	OFF-PC-10002504	Office Supplies	Envelopes	Typical Top-Opening Steel & Solid Envelopes, Plain White	130.464	6	0.2	44.0316	True	8	25	2015	8	27	2015	2	Express delivery
9871	CA-2015-114199	2015-11-01 00:00:00-00:00:00	2015-11-01 00:00:00-00:00:00	First Class	EA-13955	Eria Aishbrook	Corporate	United States	Massia	Ohio	45049	East	OFF-BC-10001659	Office Supplies	Binders	OSC Standard Thru-8-Ring Covers	23.428	3	0.7	-17.9424	True	11	1	2015	11	3	2015	2	Express delivery
3077	CA-2015-134372	2015-09-20 00:00:00-00:00:00	2015-09-20 00:00:00-00:00:00	Second Class	SV-20385	Seth Varnau	Consumer	United States	Houston	Texas	77070	Central	OFF-PC-10004943	Office Supplies	Storage	Personal Folder Holder, Ebony	44.84	5	0.2	5.605	True	4	20	2015	4	22	2015	2	Express delivery
8282	CA-2015-139973	2015-05-20 00:00:00-00:00:00	2015-05-20 00:00:00-00:00:00	Second Class	MC-17390	Mary Collamore	Corporate	United States	San Marcos	Texas	78666	Central	OFF-PA-10002389	Office Supplies	Paper	Xerox 213	10.368	2	0.2	3.6328	True	5	20	2015	5	22	2015	2	Express delivery
125	US-2015-132030	2015-11-26 00:00:00-00:00:00	2015-11-26 00:00:00-00:00:00	Second Class	AD-10180	Alan Dinnageau	Home Office	United States	Houston	Texas	77041	Central	FUR-CE-10004063	Furniture	Chairs	Global Deluxe High-Back Manager's Chair	600.558	3	0.3	48.5794	True	12	26	2015	12	28	2015	2	Express delivery
4948	CA-2015-184469	2015-06-25 00:00:00-00:00:00	2015-06-25 00:00:00-00:00:00	Second Class	GK-14020	Grace Kelly	Corporate	United States	Salem	Oregon	97301	West	OFF-AP-10003473	Office Supplies	Art	Avery 95-Line SilverBall Pen Triple Fluorescent Highlighters, 4-Pack	71.652	11	0.2	17.9008	True	6	25	2015	6	27	2015	2	Express delivery
4890	CA-2015-113681	2015-11-20 00:00:00-00:00:00	2015-11-20 00:00:00-00:00:00	First Class	CL-11995	Carol Tripp	Consumer	United States	Tucson	Arizona	85705	West	FUR-PC-10001876	Furniture	Furniture	Computer Room Manager, 14"	51.948	2	0.2	10.3836	True	12	20	2015	12	22	2015	2	Express delivery
289	CA-2015-111451	2015-11-28 00:00:00-00:00:00	2015-11-28 00:00:00-00:00:00	First Class	KL-16515	Kelly Lamplin	Consumer	United States	Colorado Springs	Colorado	80906	West	OFF-PA-10004824	Office Supplies	Fasteners	1/8" Hole Coated Wire Paper Clips in Organizer Box, 100/Box	27.552	3	0.2	9.2688	True	12	26	2015	12	28	2015	2	Express delivery
2459	CA-2015-184210	2015-11-18 00:00:00-00:00:00	2015-11-18 00:00:00-00:00:00	Second Class	FR-19240	Phere Weiser	Consumer	United States	Louisville	Colorado	80027	West	OFF-PA-10002239	Office Supplies	Paper	Geopostix 25cm Carbs, Blank, White, 8 1/2" x 11"	35.908	4	0.2	12.5228	True	11	18	2015	11	20	2015	2	Express delivery
9020	CA-2015-100706	2015-11-14 00:00:00-00:00:00	2015-11-14 00:00:00-00:00:00	Second Class	LE-16810	Laural Ellman	Consumer	United States	Springfield	Virginia	22153	South	TEC-AC-10001314	Technology	Accessories	Cas Logic 2-4GHz Wireless Keyboard	99.98	2	0.0	7.9994	True	12	16	2015	12	18	2015	2	Express delivery
8296	CA-2015-111914	2015-05-07 00:00:00-00:00:00	2015-05-07 00:00:00-00:00:00	First Class	GD-14390	Geitman Durich	Corporate	United States	Altamonte Springs	Virginia	22204	South	OFF-BC-10004164	Office Supplies	Binders	Staple Duty Pro Binders	11.88	2	0.0	5.346	True	5	5	2015	5	7	2015	2	Express delivery

Chúng tôi sử dụng các phương pháp trích xuất đặc trưng, các kĩ thuật tính năng và từ dữ liệu thô ban đầu ('raw_data') sau khi qua các bước xử lý thì bảng dữ liệu 'feature_data' sẽ là một bảng dữ liệu rõ ràng hơn. Bảng dữ liệu sau khi được cập nhật lên 'feature_data' mô tả dữ liệu một cách trực quan hơn bằng này phân tách ngày tháng năm giao hàng và nhận hàng của khách hàng để từ đó có thể tính toán chính xác thời gian mà khách hàng có thể nhận được đơn hàng mình đã đặt là bao lâu đơn hàng có thể trong ngày hoặc hơn. Từ thời gian nhận được đơn hàng của khách hàng có thể đưa ra phương thức giao hàng của cửa hàng dành cho khách hàng là “giao hàng hỏa tốc” (“Express delivery”), “giao hàng nhanh” (“Fast delivery”), “giao hàng tiết kiệm” (“Economical delivery”). Từ đó có thể giúp các nhà khoa học dữ liệu muốn sử dụng bộ dữ liệu này để xây dựng các mô hình máy học tiết kiệm thời gian và công sức để xử lý dữ liệu. Dữ liệu sau khi xử lý sẽ có nhiều cột hơn bộ dữ liệu thô ban đầu giúp dữ liệu được biểu diễn trực quan và chi tiết hơn. Trong phương pháp này còn một số hạn chế, khi lượng dữ liệu quá nhiều thì thời gian chờ xử lý sẽ tăng lên đáng kể.

5. KẾT LUẬN

Machine Learning là mô hình dự đoán cực kỳ mạnh mẽ, tuy nhiên, việc xây dựng hệ thống học máy rất khó, đòi hỏi các nền tảng và công cụ chuyên biệt. Trong thực tế, sự kết hợp giữa nền tảng dữ liệu và AI chưa nhận được nhiều sự quan tâm [7]. Trong bài báo này, sự cần thiết và tầm quan trọng của việc giới thiệu Feature Store cho môi trường hợp tác giữa công việc Data Engineering và Data Science đã được đề cập và xu hướng công nghệ của Feature Store đã được nghiên cứu bằng cách giới thiệu các tính năng của mô hình. Feature Store cung cấp các giải pháp giúp người dùng có thể xử lý dữ liệu thô một cách dễ dàng.

Những thách thức trong tương lai bao gồm các yêu cầu thiết lập môi trường hợp tác giữa công việc Data Engineering và công việc Data Science nhằm góp phần giảm chi phí phát triển mô hình AI, cải thiện hiệu suất mô hình AI, tăng độ tin cậy cũng như cải thiện sự thuận tiện khi vận hành mô hình.

LỜI CẢM ƠN

Để hoàn thành thuận lợi bài nghiên cứu “Phân tích và thiết kế Feature Store”, nhóm chúng tôi xin chân thành cảm ơn Thạc sĩ Nguyễn Hữu Tình, anh Trần Hữu Nhân đã dành nhiều thời gian và tâm huyết để chỉ dẫn, góp ý và động viên nhóm hoàn thành bài nghiên cứu. Bên cạnh đó, nhóm chúng tôi cũng xin cảm ơn Ban tổ chức Hội nghị Khoa học Trẻ lần thứ 4 năm 2022 Trường Đại học Công nghiệp Thành phố Hồ Chí Minh đã đọc bài nghiên cứu và cho nhóm chúng tôi những lời nhận xét quý báu.

TÀI LIỆU THAM KHẢO

- [1] S. Kumar, R. Velaga, S. Mohanty and P. Saripalli, “Offline and Online Feature Store for faster and consistent Machine Learning Modelling in Wellness Domain,” 2021. [Online]. Available: https://easychair-www.easychair.org/publications/preprint_download/NWKd. [Accessed 06/26/2022].
- [2] A. Agrawal, “Step-by-step Data Preprocessing & EDA,” 2018. [Online]. Available: <https://www.kaggle.com/code/agrawaladitya/step-by-step-data-preprocessing-eda/notebook?fbclid=IwAR3mEP-RKmP9cIHi8UPJNxX4GYEzLhz4Vg6go6-xgorrGAveKd2gQJtyQvU>. [Accessed 06/26/2022].
- [3] A. Zheng and A. Casari, Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists. California: O’Reilly Media, 2018.
- [4] E. P. K. Reddy, “Step by Step process of Feature Engineering for Machine Learning Algorithms in Data Science,” 2021. [Online]. Available: <https://www.analyticsvidhya.com/blog/2021/03/step-by-step-process-of-feature-engineering-for-machine-learning-algorithms-in-data-science/>. [Accessed 06/26/2022].
- [5] H. Patel, “What is Feature Engineering — Importance, Tools and Techniques for Machine Learning,” 2021. [Online]. Available: <https://towardsdatascience.com/what-is-feature-engineering-importance-too=ls-and-techniques-for-machine-learning-2080b0269f10#:~:text=Fe=ature>. [Accessed 06/26/2022].
- [6] J. Kumar M J, Feature Store for Machine Learning. Birmingham: Packt, 2022.
- [7] M. Kuhn and K. Johnson, Feature Engineering and Selection: A Practical Approach for Predictive Models. Boca Raton, Florida; London; New York: CRC Press, 2020.

YSC4F.238

SỬ DỤNG LÝ THUYẾT TRÒ CHƠI VÀO BÀI TOÁN KHUYẾN MÃI TRONG THƯƠNG MẠI ĐIỆN TỬ

LÂM QUANG PHÚ, HUỖNH THỊ LƯU LY, TRẦN LỘC, TRƯƠNG ĐỨC TÀI,
NGUYỄN TRẦN GIA HUY, LÊ PHÚC LŨ

*Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
lamquangphu2176@gmail.com, luuly1163@gmail.com, tranloc120603@gmail.com,
taitruong256@gmail.com, huycmdt@gmail.com, lephuclu@iuh.edu.vn*

Tóm tắt. Khuyến mãi, tặng quà để tăng nhu cầu mua hàng là vấn đề quan trọng trong thương mại điện tử, một trong các cách khuyến mãi quen thuộc là quay số/rút thăm trúng thưởng hoặc được công ty thiết kế dưới hình thức trò chơi. Các chương trình đó vừa giúp tri ân khách hàng vừa giúp công ty xây dựng hình ảnh vừa có thể quảng bá sản phẩm. Việc thiết kế một trò chơi hợp lý để cân đối giữa tỷ lệ thắng - thua, tỷ lệ trúng thưởng với quỹ tiền thưởng, phần quà, đảm bảo đôi bên cùng hài lòng là điều thiết thực và đòi hỏi các tính toán xác suất, kiểm chứng mô hình hợp lý.

Do đó, trong bài nghiên cứu này, nhóm sẽ nghiên cứu việc tính toán kỳ vọng, phân bố xác suất của các đối tượng trong trò chơi khuyến mãi, đưa ra phương án thưởng với tỷ lệ hợp lý. Bên cạnh đó, tìm hiểu việc giả lập một mô hình xác suất để kiểm chứng các tính toán bằng các quá trình ngẫu nhiên. Dựa trên các thư viện rất mạnh của Python cũng như Visual Basic trong MS Excel, nhóm sẽ thiết kế một chương trình hoàn chỉnh để hỗ trợ người quản lý chương trình khuyến mãi có thể sử dụng nhằm đảm bảo mọi thứ sẽ diễn ra đúng kế hoạch.

Từ khóa. Lý thuyết trò chơi, xác suất, tính toán kỳ vọng, phân bố xác suất.

USING GAME THEORY TO SOLVE THE PROMOTION PROBLEM IN E-COMMERCE

Abstract. Promotional campaigns and gifts for customers increasing shopping demand is an important operation in the e-commerce scope. Among lots of promotion methods, we can point out some familiar ones, such as lotteries, lucky draws, or online games. These promotions both help in showing gratitude to customers and helping the companies build their trademarks and promote their products. Designing a reasonable game to balance the win-loss ratio, the winning ratio of players, limited funds, gifts, and ensuring mutual satisfaction are practical and require probability calculations, and valid model verification.

In this work, our team will study the calculation of expectations, the probability distribution of the objects in the promotion game and offer a reward plan with a reasonable rate. In addition, research to simulate a probabilistic model to verify calculations using stochastic processes. Based on tools like the strong library in Python or Visual Basic in MS Excel, we are going to design a program that supports the campaign manager in controlling everything that will work for the right purpose.

Keywords. game theory, probability, calculation of expectations, probability distribution, lotteries, lucky lucky.

1. ĐẶT VẤN ĐỀ

Khuyến mãi, tặng quà của công ty là một trong những yếu tố góp phần thúc đẩy doanh nghiệp tiếp tục mở rộng và phát triển. Các công ty và nhà tiếp thị từ lâu đã dựa vào sức mạnh của việc tặng quà để đảm bảo mối quan hệ tốt hơn với khách hàng tiềm năng và khách hàng hiện tại, để thể hiện sự đánh giá cao đối với

doanh nghiệp của họ. Cùng với sự phát triển của công nghệ và kỹ thuật, việc khuyến mãi cũng dần chuyển đổi cơ chế thông qua các trò chơi trực tuyến như rút thăm / quay số... [1].

Một trò chơi khuyến mãi thành công cần phải đảm bảo hai yếu tố là quảng bá phát triển và tỷ lệ thắng thua hợp lý, tỷ lệ trúng thưởng so với quỹ thưởng để đảm bảo đôi bên cùng có lợi. Việc tính toán tỷ lệ thắng thua không đúng có thể khiến các doanh nghiệp phải tốn chi phí tương đương 1/5 thu nhập hàng năm và chi phí này sẽ càng gia tăng nếu tỷ lệ càng sai lệch [2].

Qua báo cáo của một số nghiên cứu về tỷ lệ thắng thua và giá trị kì vọng trong trò chơi đổi thưởng, trong trường hợp này là trò chơi quay số với số 6. Thì tỷ lệ để người chơi có thể quay ra được ô đặc biệt là 10% và các ô còn lại là 18%. Còn trong trò rút thăm may mắn tỷ lệ để khách hàng có thể trúng thưởng là 0.6% [2].

Những trò này có những hạn chế là không thể thay đổi được xác suất xảy ra, không có thống kê cụ thể rõ ràng về xác suất của từng trường hợp. Khiến các doanh nghiệp khó khăn trong việc đặt ra giải thưởng tương ứng hoặc đảm bảo tỷ lệ thắng thua hợp lý để các doanh nghiệp không bị thua lỗ.

Và chúng tôi nhận ra rằng thách thức trong trường hợp này chính là việc điều chỉnh tỷ lệ thắng thua hợp lý. Vì vậy trong bài báo cáo này chúng tôi sẽ đề xuất một chương trình có thể giúp người dùng tính toán xác suất các trường hợp xảy ra thông qua hai trò chơi là: “Bầu Cua” và “Rút bài 52 lá”, kèm theo đó người dùng cũng có thể điều chỉnh tỷ lệ theo mong muốn. Chương trình sẽ dựa trên các thông số người chơi điều chỉnh và lựa chọn để tính tỷ lệ xác suất phù hợp, thống kê, phân bố cơ cấu giải thưởng hợp lý để người chơi đều sẽ được quà và doanh nghiệp cũng không bị thua lỗ.

Bên cạnh 2 mô hình trò chơi đã nêu trên đề tài của nhóm còn kèm theo một chương trình coupon code giúp mã hóa, sinh code khuyến mãi cho từng khách hàng tham gia. Hướng đến giải quyết vấn đề lỗi mã khuyến mãi trong cuộc sống như (trùng lặp, sai lệch, dễ đoán...).

Các phần tiếp theo của nghiên cứu bao gồm việc trình bày cơ sở lý thuyết của mô hình nghiên cứu, phương pháp nghiên cứu, Kết quả nghiên cứu, và cuối cùng là thảo luận các hàm ý tính toán liên quan. Cuối cùng, nghiên cứu cũng trình bày một số hạn chế để lưu ý cho các tác giả khác thực hiện các nghiên cứu liên quan về sau này.

2. PHƯƠNG PHÁP NGHIÊN CỨU VÀ CÁC CÔNG NGHỆ LIÊN QUAN

2.1. Ngôn ngữ lập trình phát triển ứng dụng

2.1.1. Ngôn ngữ lập trình Python

Python là một ngôn ngữ lập trình mã nguồn mở rất mạnh mẽ, hỗ trợ một loạt các thư viện bổ sung. Một thư viện cụ thể tuyệt vời để phân tích và tạo nên các ứng dụng desktop là PyQt5.

Python là một ngôn ngữ lập trình động, có mục đích chung, được thông dịch, cấp cao được sử dụng rộng rãi. Triết lý thiết kế của nó nhấn mạnh đến khả năng đọc mã và cú pháp của nó cho phép các lập trình viên diễn đạt các khái niệm bằng ít dòng mã hơn có thể bằng các ngôn ngữ như C++ hoặc Java. [3]

Ngôn ngữ này cung cấp các cấu trúc nhằm cho phép viết các chương trình rõ ràng ở cả quy mô nhỏ và quy mô lớn. Python hỗ trợ nhiều mô hình lập trình, bao gồm các kiểu lập trình hoặc thủ tục hướng đối tượng, mệnh lệnh và chức năng. Nó có hệ thống kiểu động và quản lý bộ nhớ tự động và có một thư viện tiêu chuẩn lớn và toàn diện. [4]

2.1.2. Ngôn ngữ lập trình C++

Là một ngôn ngữ lập trình đa năng, ngôn ngữ này mở rộng đáng kể theo thời gian và C++ hiện đại ngày nay có các tính năng hướng đối tượng, chung sức và chức năng ngoài tiện ích cho bộ nhớ cấp thấp. Thao tác đơn giản hầu như rất nhanh trong việc xử lý dữ liệu và kiểm thử [5]. Vì vậy nhóm chúng tôi quyết định dùng ngôn ngữ này cho phần kiểm thử, tính đúng đắn trong công thức xác suất mà nhóm xây dựng trên ứng dụng.

2.1.3. Modulo PyQt 5

PyQt5 là một thư viện trong ngôn ngữ lập trình Python, được liên kết và phát triển từ khung Qt của C++ có thể được sử dụng để tạo giao diện người dùng đồ họa trên máy tính để bàn. PyQt5 cung cấp cho người dùng

tất cả các chức năng phức tạp của Qt C++ đồng thời cho phép người dùng triển khai nhanh chóng bằng Python [6].

2.2. Mô hình xác suất của trò chơi bầu cua

2.2.1. Mô hình lý thuyết : Cho n con xúc xắc đặc biệt, mỗi con có các mặt:

“BẦU”, “CUA”, “TÔM”, “CÁ”, “NAI”, “GÀ”

tùy ý, người dùng có thể thay đổi không nhất thiết là 1 việc điều chỉnh theo mong muốn này sẽ thay đổi tỉ lệ xuất hiện ở các mặt. Ở đây quy ước: các mặt trên có index là 0,1,2,3,4,5 tương ứng trong các mảng dữ liệu. Tính xác suất để khi gieo đồng thời n con xúc xắc đó, ta nhận được một số mặt của các loại trên với số lượng mong muốn.

2.2.2 Phân tích

Gọi là số lượng các loại mặt xuất hiện trong trò chơi này, ở đây ta có giá trị số này bằng 6, tương ứng với 6 loại mặt. Xét dữ liệu List<int [NUMBER OF TYPE]> dices thể hiện thông tin của các con xúc xắc, bao gồm tất cả n phần tử là dices [0], dices [1] ...dices [$n - 1$]. Mỗi phần tử như thế lại là 1 mảng chứa số lượng mặt của con xúc xắc tương ứng.

Chẳng hạn con xúc xắc thứ 1 có 3 mặt là “BẦU”, 1 mặt là “CUA”, 1 mặt là “TÔM”, 1 mặt là “CÁ”, 4 mặt là “NAI” và 0 mặt là “GÀ” thì ta có dices [0] = (3,1,1,1,4,0).

Kết quả mong muốn cũng có dạng một danh sách như sau:

List <KeyValuePair<int, int>> result

trong đó, số lượng phần tử của list này sẽ không vượt quá NUMBER_OF_TYPE (không nhất thiết bằng đúng số này mà có thể nhỏ hơn), cho biết số lượng tương ứng của các mặt muốn có được.

Chẳng hạn, ta muốn tính xác suất để khi tung các con xúc xắc, ta thu được đúng 2 mặt “BẦU”, 1 mặt “CUA” và không có mặt nào là “GÀ”, ta có:

result = (< 0,2 >, < 1,1 >, < 5,0 >)

Ràng buộc dữ liệu:

+ Mỗi key xuất hiện không quá 1 lần.

+ Tổng các value của tất cả các cặp phải không vượt quá n là số lượng xúc xắc.

⇒ Các dữ liệu vi phạm ràng buộc này đều cho ra xác suất bằng 0.

2.2.3 Cách tính xác suất

Gọi A là biến cố thỏa mãn đề bài, không gian mẫu của phép thử tung xúc xắc là Ω [7].

Khi đó, ta cần tính:

$$P(A) = \frac{n(A)}{n(\Omega)}$$

Bước 1: Tính $n(\Omega)$

Tính tổng số lượng các mặt của các con xúc xắc (tính tổng của 6 phần tử), đặt là $a_0, a_1, a_2 \dots a_{n-1}$. Rõ ràng các số này đều phải là số lớn hơn 0. Khi đó: $n(\Omega) = a_0, a_1, a_2 \dots a_{n-1}$.

Bước 2: Tính $n(A)$

Bước này là quan trọng nhất. Đầu tiên, ta kiểm tra trong các giá trị values của mặt, có giá trị nào là 0 hay không. Nếu có, nghĩa là các mặt tương ứng không được xuất hiện, ta duyệt mảng dữ liệu các con xúc xắc và loại ra các mặt vi phạm được một bộ xúc xắc mới.

Nếu có một con xúc xắc nào còn lại 0 mặt, ta trả về kết quả là 0. Tiếp theo, ta sẽ trải qua các bước sinh các thông tin [7].

Bước 3: Tính xác suất $P(A)$

Khi đã có 2 giá trị $n(A)$, $n(\Omega)$, ta chia $n(A)$ có số lượng hoán vị của các giá trị giống nhau (chẳng hạn cần 3 mặt “BẦU” thì sau đó, ta chia $\frac{n(A)}{3!}$) và cứ thế [7].

Cuối cùng, xác suất cần tính chính là : $P(A) = \frac{n(A)}{n(\Omega)}$

2.3. Mô hình xác suất của trò chơi rút bài

Trò chơi rút bài của nhóm dựa trên trò chơi Mậu binh [8] là một trò chơi dân gian rất được ưa chuộng. Trong Mậu binh, có thể chơi 2 - 4 người, mỗi người chơi có 13 lá bài (sử dụng bài Tây). Tùy vào tính toán chiến thuật của mỗi người chơi, các lá bài này sẽ được sắp xếp thành các bộ khác nhau, tính tối ưu của các bộ là không đổi [8].

Luật chơi: sẽ dựa trên luật trò chơi Mậu Binh. Người chơi sẽ rút một số lượng lá bài (3, 4 hoặc 5 lá) để tạo thành một bộ bài được rút. Dựa vào kết quả xác suất đã được tính trước để xác định phần thưởng cho người chơi rút được bộ bài đó. Các kết quả rút bài có thể đạt được như sau:

Hình 1 – Giải thích luật chơi 5 lá

Tên	Ý nghĩa	Ví dụ	Điều kiện quyết định cao thấp
Mậu Thầu (en: High Card)	Không có liên kết nào với các lá bài khác	A♠	Xét lá bài cao nhất, sau đó xét các lá bài cao tiếp theo
Đôi (en: Pair)	Hai lá bài đồng số	10♣ 10♥	Xét đôi cao thấp, tiếp đến là các lá bài lẻ cao nhất
Thú (en: Two Pairs)	Hai đôi	J♦ J♣ 8♣ 8♠	Xét đôi cao thấp, tiếp đến là đôi thấp hơn và sau cùng là (các) lá bài lẻ
Xám chi/Xám cô (en: Three of a Kind)	Ba lá bài đồng số	Q♣ Q♥ Q♠	Xét cao thấp của bộ ba
Sảnh (en: Straight)	Năm lá bài trong một chuỗi số (nhưng không đồng chất)	J♣ 10♦ 9♥ 8♣ 7♥	Xét cao thấp của lá bài cao nhất của chuỗi
Thùng (en: Flush)	Năm lá bài cùng màu, đồng chất (nhưng không cùng một chuỗi số)	7♣ Q♣ 10♣ K♣ A♣	Xét lá bài cao nhất, sau đó các lá bài cao tiếp theo
Cù lũ (en: Full House)	Một bộ ba và một bộ đôi	K♣ K♠ K♥ 9♥ 9♣	Xét cao thấp của bộ ba
Tứ quý (en: Four of a Kind)	Bốn lá đồng số	A♥ A♦ A♠ A♣	Xét cao thấp lá bài của tứ quý
Thùng phá sảnh (en: Straight Flush)	Năm lá bài cùng màu, đồng chất, cùng một chuỗi số	Q♣ J♣ 10♣ 9♣ 8♣	Xét lá bài cao nhất

2.3.1. Công thức tính xác suất

Xác suất được tính theo công thức tổng quát sau: $P(A) = \frac{n(A)}{n(\Omega)}$. [7]

Trong đó: A là biến cố.
 P(A) là xác suất của biến cố A.
 n(A) là số phần tử của biến cố A.
 n(Ω) là không gian mẫu.

Số phần tử của biến cố A được tính như sau:

Gọi m là số lượng lá bài cần lấy của mỗi trường hợp với m = 5 ta thu được bảng tính các biến cố trong các trường hợp như sau:

Chọn ngẫu nhiên 5 trong 52 lá bài: $C_{52}^5 \Rightarrow n(\Omega) = C_{52}^5$.

Bảng 1- Tính các biến cố và xác suất trong rút bài 5 lá

CÁC TỔ HỢP	GIẢI THÍCH CÁCH TÍNH	n(A)	n(Ω) =
Thùng phá sảnh	Có 10 cách tạo sảnh từ A2345 đến 10JQKA trong 1 bộ màu: C_{10}^1 . Bộ bài có 4 bộ màu.	$N(1) = C_{10}^1 \cdot 4 = 40$	$P(1) = \frac{n(1)}{n(\Omega)} = \frac{1}{64974}$
Tứ quý	Có 13 cách chọn 4 lá cùng số: C_{13}^1 . Chọn cả 4 lá trong 4 lá cùng số: C_4^4 . Vị trí còn lại có $52-4=48$ cách chọn.	$n(2) = C_{13}^1 \cdot C_4^4 \cdot 48 = 624$	$P(2) = \frac{n(2)}{n(\Omega)} = \frac{1}{4165}$

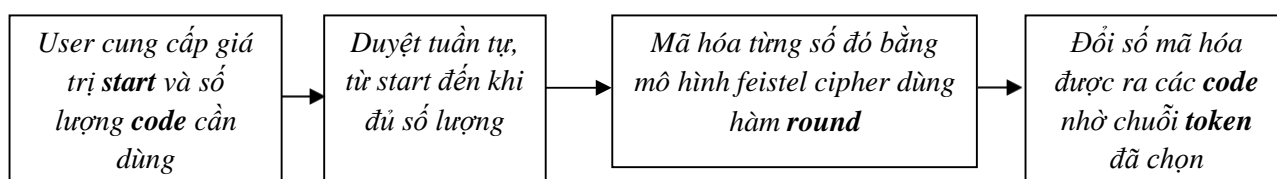
Cù lũ	Có 13 cách chọn 3 cây giống nhau C_{13}^1 . Chọn cả 3 lá trong 4 lá cùng số: C_4^3 . Có 12 cách chọn 1 đôi trong các lá còn lại: C_{12}^1 . Chọn cả 2 lá trong 4 lá cùng số: C_4^2 .	$n(3) = C_{13}^1 \cdot C_4^3 \cdot C_{12}^1 \cdot C_4^2 = 3744$	$P(3) = \frac{n(3)}{n(\Omega)} = \frac{6}{4165}$
Thùng	Chọn 5 lá trong 13 lá cùng màu: C_{13}^5 . Bộ bài có 4 màu. Mỗi màu có 10 trường hợp 5 lá ấy là 1 dãy số. Loại trường hợp là thùng phá sảnh.	$n(4) = C_{13}^5 \cdot 4 - n(1) = 5108$	$P(4) = \frac{n(4)}{n(\Omega)} = \frac{1277}{649740}$
Sảnh	Có 10 cách tạo sảnh từ A2345 đến 10JQKA: C_{10}^1 . Mỗi lá bài có 4 cách chọn do bộ bài có 4 màu. Sảnh abcde có $4.4.4.4.4 = 4^5$ cách chọn. Loại trường hợp là thùng phá sảnh.	$n(5) = C_{10}^1 \cdot 4^5 - n(1) = 10200$	$P(5) = \frac{n(5)}{n(\Omega)} = \frac{5}{1274}$
Sám chi	Chọn 1 bộ số trong 13 bộ số: C_{13}^1 Chọn 3 lá cùng số trong 4 lá cùng số: C_4^3 Chọn 2 lá ở 12 bộ số còn lại và 2 lá đó phải ở mỗi bộ khác nhau: $C_{12}^2 \cdot (C_4^1)^2$	$n(6) = C_{13}^1 \cdot C_4^3 \cdot C_{12}^2 \cdot (C_4^1)^2 = 54912$	$P(6) = \frac{n(6)}{n(\Omega)} = \frac{88}{4165}$
Hai đôi	Chọn 2 bộ khác số trong 13 bộ số: C_{13}^2 Chọn 2 trong 4 lá ở mỗi bộ số đã chọn: $C_4^2 \cdot C_4^2$ Vị trí còn lại có $52-8=44$ cách chọn	$n(7) = C_{13}^2 \cdot C_4^2 \cdot C_4^2 \cdot 44 = 123552$	$P(7) = \frac{n(7)}{n(\Omega)} = \frac{198}{4165}$
Một đôi	Chọn 1 bộ số trong 13 bộ số: C_{13}^1 Chọn 2 lá trong 4 lá ở bộ số đã chọn: C_4^2 Chọn 3 lá trong 12 bộ số còn lại và 3 lá đó ở mỗi bộ số khác nhau: $C_{12}^3 \cdot 4^3$	$n(8) = C_{13}^1 \cdot C_4^2 \cdot C_{12}^3 \cdot 4^3 = 1098240$	$P(8) = \frac{n(8)}{n(\Omega)} = \frac{352}{833}$
Mậu thầu	Chọn 5 bộ số trong 13 bộ số: C_{13}^5 . Mỗi bộ số chọn 1 lá trong 4 lá: $(C_4^1)^5$. Loại các trường hợp thùng và sảnh.	$n(9) = C_{13}^5 \cdot (C_4^1)^5 - n(4) - n(5) - n(1) = 1302540$	$P(9) = \frac{n(9)}{n(\Omega)} = \frac{1277}{2548}$

2.4. Mô hình sinh coupon code

Coupon là vé/chứng từ / phiếu giảm giá/ mã giảm giá được nhà sản xuất/ nhà bán lẻ tung ra trong các chương trình khuyến mãi. Mã Coupon (coupon code) cũng là một dạng coupon nhưng thay vì sử dụng phiếu thì người dùng sẽ sử dụng một đoạn mã để nhập vào lúc thanh toán online trên trang web. Khi người mua hàng sử dụng coupon thì họ được giảm một mức giá trong hoá đơn mua hàng so với việc mua thông thường [9].

2.4.1. Thuật toán sinh Coupon Code

Yêu cầu: mã hóa các số trong một dãy cho trước, chẳng hạn từ $1 \rightarrow 10^6$ thành các chuỗi ký tự gồm chữ cái, chữ số $A \rightarrow Z, a \rightarrow z, 0 \rightarrow 9$ một cách ngẫu nhiên, ở đây quy ước gọi là code. Bên cạnh đó, từ các code đã được mã hóa, có thể giải mã ngược lại thành số ban đầu. Mô hình chính của bước mã hóa là:



Hình 2 - Mô tả quy trình thuật toán sinh coupon code

Quy trình cơ bản để mã hóa một coupon code như sau:

- Bước 1: Người dùng nhập vào một token là một chữ cái hoặc chữ số đôi một khác nhau
- Bước 2: Mã hóa bằng mô hình Feistel Cipher dùng hàm round
- Bước 3: Tạo code có thể sử dụng được từ số đã mã hóa
- Bước 4: Ứng dụng sẽ giải mã code để lấy lại số ban đầu

Về nguyên tắc thì hai số đầu vào phân biệt thì hai code đầu ra cũng phân biệt, tuy nhiên để đảm bảo hơn cho người dùng, ta có thể thực hiện việc kiểm tra các code sinh ra được. Để làm được điều này nhóm đã sử dụng một số cách phổ biến như:

Sử dụng nhiều hash set, với code mới sinh ra, kiểm tra 2 ký tự đầu vào của nó và đổi ra thứ tự tương ứng vào hash set. Lưu vào hash set với thứ tự tương ứng(cơ chế hash set sẽ xóa phần tử bị trùng), cuối cùng ta chỉ cần tính tổng phần tử của hash set này lại để so sánh với số lượng dữ liệu đưa vào.

2.4.2. Hạn chế sự tuần hoàn của mã hóa

Việc tuần hoàn code đã mã hóa ở đây là khó tránh khỏi, vì tất cả các bước không có sử dụng hàm random. Cần nhấn mạnh rằng việc tuần hoàn này là chỉ của một **nhóm ký tự**, chẳng hạn 2 ký tự đầu trong code, chứ không phải toàn bộ code.

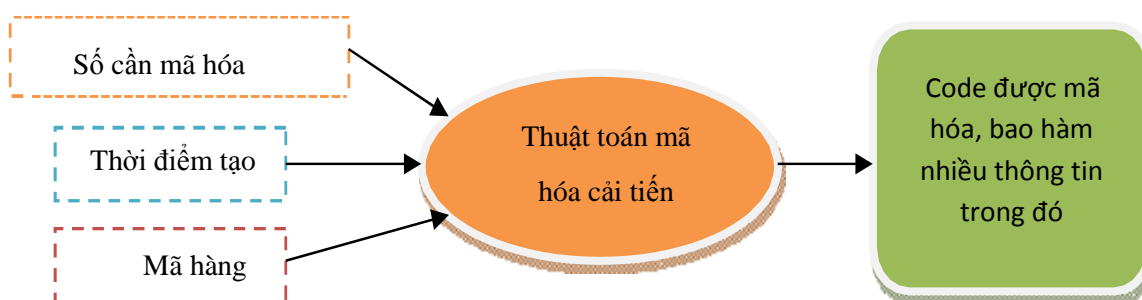
Có thể khó thấy việc lặp lại này ở một khoảng ngắn, nhưng nó có thể lặp lại với chu kỳ 100 hoặc 1000, tùy đặc trưng của hàm round. Tất nhiên, với độ dài đó cũng đã an toàn nhưng nếu muốn tốt hơn, ta có thể bổ sung một bước xáo trộn để loại bỏ hoàn toàn tính chu kỳ đó. Để tránh điều này, ta có thể bổ sung thêm một bước nhỏ là **shuffle** các số trong đoạn đó.

Tức là các số trong khoảng muốn gen sẽ được xáo trộn lại và đưa vào hàm mã hóa. Thứ tự xáo trộn sẽ được lưu vào database.

Khi đó, thay vì sử dụng i để mã hóa, ta sẽ dùng $shuffle[i]$. Các số được mã hóa ra ở bước 3 cũng sẽ bị xáo trộn theo, tránh được chuyện bị lặp theo chu kỳ.

2.4.3. Mã hóa các yếu tố khác

Bên cạnh các số ban đầu, ta còn có thể bổ sung thêm tham số thời gian, mã hàng, ... vào thuật toán mã hóa.



Hình 3 - Quy trình cải tiến của thuật toán mã hóa

Tất nhiên một khi đã mã hóa được thì với các cách đặc thù, có thể giải mã và lấy lại được các thông tin đó để kiểm tra, đối chiếu dễ dàng. Điều quan trọng là user sử dụng các thông tin đó trong việc bảo mật thể nào.

3. TÍNH TOÁN VÀ THỰC NGHIỆM

3.1. Quá trình thực hiện chương trình khuyến mãi

Khi một công ty muốn tạo một chương trình khuyến mãi đến khách hàng. Ta sẽ dựa vào các yêu cầu về số lượng giải, tổng giá trị chương trình... và kết hợp với kết quả xác suất tính được để tạo ra một chương trình khuyến mãi phù hợp nhất.

Việc thu thập dữ liệu để tạo chương trình khuyến mãi để phù hợp doanh thu của công ty hoặc so với quỹ giải thưởng cần phải tính toán chính xác. Tổng giải thưởng có thể được tính từ nhiều nguồn khác nhau như: số lượng sản phẩm bán được, lợi nhuận của công ty... hay được trích một phần vốn, quỹ của công ty để góp phần vào chương trình khuyến mãi. Tuy có thể kết hợp từ nhiều nguồn tiền khác nhau nhưng nhìn chung nó đều xuất phát từ số lượng sản phẩm bán được. Qua đó vẫn đảm công bằng chương trình khuyến mãi cho số lượng khách hàng đó.

Khi xác định được các dữ liệu đầu vào, ta tiến hành thiết kế chương trình khuyến mãi theo các hình thức nêu trên:

+ Rút thăm, mã khuyến mãi (coupon code): số lượng giải và giá trị từng giải thưởng được xác định ngay từ đầu. Quỹ tiền khuyến mãi sẽ ổn định, ít rủi ro cho công ty.

+ Rút bài, bầu cua: có thể điều chỉnh tỷ lệ, cơ cấu và giá trị giải thưởng.

Nhưng phải cần tính toán cẩn thận để tạo ra chương trình khuyến mãi thực tế được chính xác.

3.2. Tình huống thực tế qua chương trình bầu cua

Với dữ liệu chuẩn thông thường là 3 con xúc xắc có 6 mặt khác nhau, ta cần tính xác suất để có 2 con “BẦU” thì dữ liệu và các bước trên cụ thể như sau:

$$dices = \{(1,1,1,1,1,1), (1,1,1,1,1,1), (1,1,1,1,1,1)\} \text{ và } result = \langle 0, 2 \rangle.$$

Không gian mẫu tính được là $n(\Omega) = 6.6.6 = 216$.

Dãy số ở bước 2.1 là $\{0,0\}$. Các bộ tính thứ tự ở bước 2.2 là:

$$\{(0,1), (1,0), (0,2), (2,0), (1,2), (2,1)\}.$$

Xét bộ index $(0,1)$, nghĩa là con xúc xắc thứ 0 và 1 đều phải cho ra kết quả là “BẦU”, con xúc xắc còn lại không phải bầu nên có 5 cách chọn. Do đó, số lượng cần tìm ở đây là $1.1.5 = 5$. Duyệt qua hết 6 bộ số, ta được

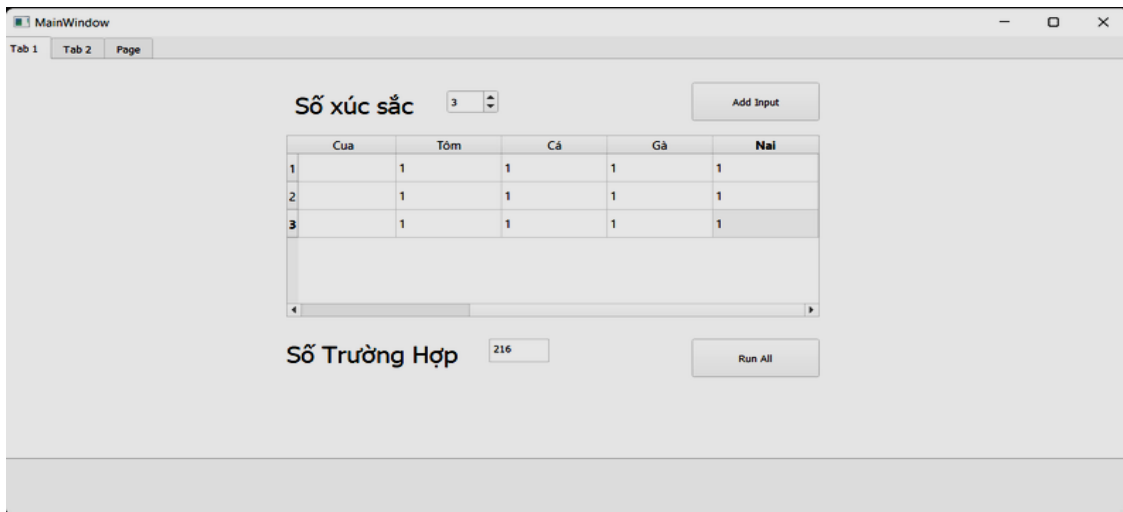
$$n(A) = 6.5 = 30. \text{ Tuy nhiên, do ở đây ta xét 2 con “BẦU” giống nhau nên } n(A) \leftarrow = \frac{n(A)}{3!}.$$

$$\text{Xác suất cho biến cố A này là } \frac{15}{216} = 0.06944444(4).$$

Ta có thể thấy việc tính toán thông thường bằng tay để kiểm tra và xác định là rất lâu, thông qua ứng dụng của chúng tôi người dùng chỉ cần nhập dữ liệu đầu vào gồm:

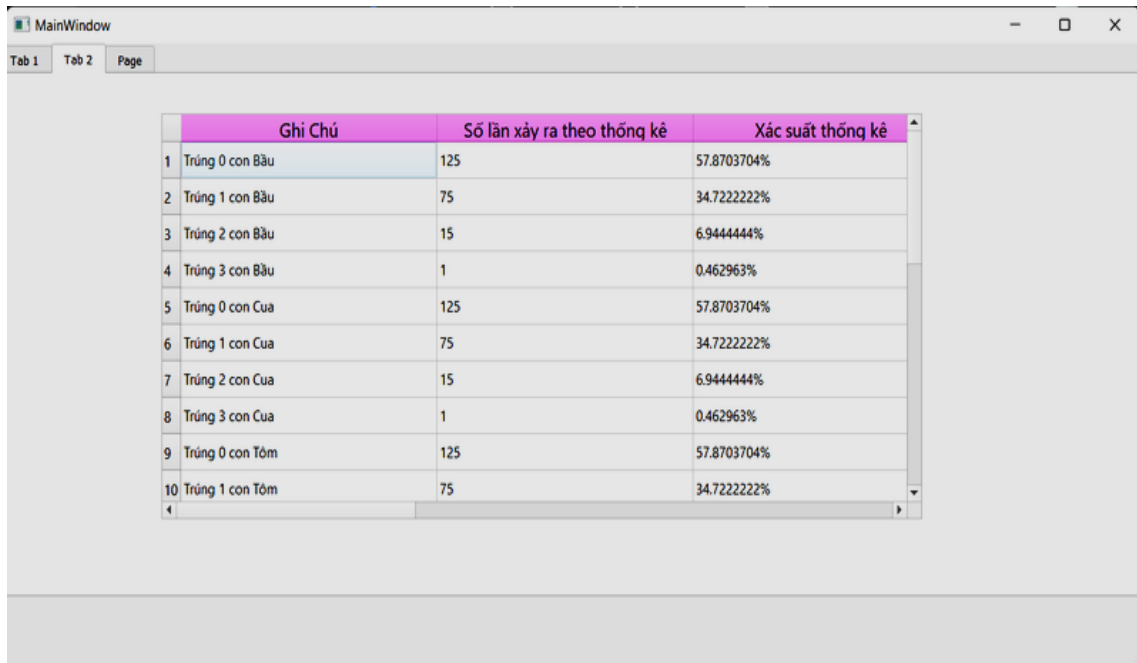
+Số lượng xúc xắc

+Tỉ lệ mong muốn xuất hiện ở các mặt



Hình 4 – Nhập dữ liệu đầu vào của người dùng

Chương trình sẽ xuất ra số trường hợp xảy ra và thống kê xác suất chi tiết cụ thể của từng trường hợp



Hình 4 – Hình 5 - Kết quả trường hợp thử nghiệm dices = {(1,1,1,1,1), (1,1,1,1,1), (1,1,1,1,1)}

Chương trình còn sẽ tính số lượng khách hàng trúng được và sẽ đề nghị quỹ thưởng phù hợp cho từng mục so với tổng quỹ thưởng ban đầu người dùng nhập vào.

3.3. Tình huống thực tế qua trò chơi rút bài

Trong dịp cuối năm, một chi nhánh công ty T muốn tạo một chương trình tri ân đến khách trong năm vừa qua. Với lượng sản phẩm bán được là 100.000.000 thùng mì tôm. Với giá mỗi thùng mì là 100.000 VNĐ, lợi nhuận thu được sau khi đã trừ hết các chi phí là 25% doanh thu. Và chi nhánh công ty này muốn trích ra 20% lợi nhuận để tổ chức khuyến mãi cho 100.000.000 khách hàng đã mua mì (giá sử dụng trung bình mỗi người mua một thùng mì). Yêu cầu đặt ra là tạo một chương trình khuyến mãi với xác suất được tính cho trò chơi rút thăm 5 lá bài.

Đầu tiên, ta xác định quỹ giải thưởng dựa vào các số liệu trên:

Doanh thu = (Số lượng sản phẩm) * (Giá sản phẩm) = 100.000.000 * 100.000 = 10 nghìn tỷ VNĐ.

Lợi nhuận = (Doanh thu) * (Tỷ lệ lợi nhuận) = 10 nghìn tỷ * 25% = 2500 tỷ VNĐ.

Quỹ khuyến mãi = (Lợi nhuận) * (phần trăm trích làm khuyến mãi) = 2500 tỷ * 20% = 500 tỷ VNĐ.

Dựa vào kết quả xác suất:

	SỐ BIẾN CỐ (Ω=2598960)	XÁC SUẤT
Thùng phá sảnh	40	0.00154%
Tứ quý	624	0.024%
Cù lũ	3744	0.144%
Thùng	5108	0.197%
Sảnh	10200	0.393%
Xám chi	54912	2.1129%
Thú	123552	4.754%
Đôi	1098240	42.257%
Mậu thầu	1302540	50.118%

Hình 6 - Kết quả xác suất rút 5 lá bài

Ta có thể thiết kế chương trình rút tham trúng thưởng cho các mục: thùng phá sảnh, tứ quý, cù lũ, thùng, sảnh, xám chi, thú. Đôi và mậu thầu ta sẽ cho vào danh sách không trúng thưởng.

Áp dụng định luật Zipf, tần số suất hiện của một mục trúng thưởng sẽ tỷ lệ nghịch với giá trị giải thưởng [10].

Tổng quỹ thưởng của chương trình khuyến mãi này là 500 tỷ ta sẽ chia đều cho 7 khả năng trúng thưởng trên = 500 tỷ / 7 = 72 tỷ mỗi phần.

Và số lượng khách hàng có khả năng trúng thưởng là: $(40+624+3744+5108+10200+54912+123552) * 100.000.000 / 2598960 = 7\ 625\ 357$ khách hàng.

Tỷ lệ trong danh sách mục trúng thưởng là:

(Thùng phá sảnh : Tứ quý : Cù lũ : Thùng : Sảnh : Xám chi : Thú)
 (1 : 15 : 93 : 127 : 255 : 1372 : 3088)

Suy ra, số lượng khách hàng có khả năng trúng từng danh mục đó là (đơn vị khách hàng):

- + Thùng phá sảnh: 1540
- + Tứ quý: 23100
- + Cù lũ: 143220
- + Thùng: 195580
- + Sảnh: 392700
- + Xám chi: 2112880
- + Thú: 4755520

Vậy với mỗi mục giá trị tiền là 72 tỷ thì ta có giá trị từng mục trúng thưởng là:

- + Thùng phá sảnh: 46.7 triệu VNĐ
- + Tứ quý: 3,119 triệu VNĐ
- + Cù lũ: 500 ngàn VNĐ
- + Thùng: 368 ngàn VNĐ
- + Sảnh: 183346 ngàn VNĐ
- + Xám chi: 34 ngàn VNĐ
- + Thú: 15 ngàn VNĐ

Với giá trị giải thưởng mỗi phần như trên, chi nhánh công ty đó có thể tạo ra giải thưởng bằng tiền hoặc phần quà tương ứng.

3.4. Kết quả mô hình coupon code

Coupon Code không dựa trên các kết quả xác suất, để phân bổ cơ cấu giải thưởng, số lượng giải thưởng như trò chơi: “Rút bài” hay “Bầu cua”. Vì vậy tùy vào yêu cầu của công ty, nhà đầu tư trò chơi mà chương trình sẽ tạo ra một lượng phiếu trúng thưởng và giải thưởng cố định sẵn. Với cách thực hiện chương trình khuyến mãi như vậy sẽ giúp công ty, nhà đầu tư trò chơi tính chính xác quỹ thưởng và không có sai số trong việc triển khai chương trình.

Bảng 2 - Thời gian trung bình kiểm tra trùng với độ dài coupon là 6.

Dữ liệu (số coupon)	Thời gian (s)
1 triệu	0,166
2 triệu	0,415
3 triệu	0,706
4 triệu	0,896
5 triệu	1,220
6 triệu	1,575
7 triệu	1,746
8 triệu	1,950
9 triệu	2,474
10 triệu	2,657
11 triệu	2,876
12 triệu	3,527
13 triệu	3,811
14 triệu	4,361
15 triệu	4,679

3.5. Kiểm chứng tính đúng đắn của công thức

Để kiểm chứng công thức, chúng tôi đã viết 1 chương trình code bằng Python để tính tất cả các trường hợp có thể xảy ra. Ý tưởng là vét cạn các trường hợp bằng 5 vòng for lồng nhau, sau đó lưu các giá trị vào một vector và kiểm tra xem có thỏa mãn điều kiện hay không. Độ phức tạp cao nhưng thời gian thực thi cũng tốn vài giây, và do chỉ dùng một lần để kiểm chứng nên chấp nhận được.

```
ll omega()
{
    return (52*51*50*49*48)/(1*2*3*4*5);
}

ll thungphasanh()
{
    ll thungphasanh=0;

    for (ll i=0; i<56-4; i++)
        for (ll j=i+1; j<56-3; j++)
            for (ll k=j+1; k<56-2; k++)
                for (ll l=k+1; l<56-1; l++)
                    for (ll m=l+1; m<56; m++)
                        {
                            vector<ll> v;
                            v.push_back(i);
                            v.push_back(j);
                            v.push_back(k);
                            v.push_back(l);
                            v.push_back(m);
                            ll dem[4];
                            for (ll n=0; n<4; n++) dem[n]=0;
                            for (ll n=0; n<5; n++)
                                {
                                    ++dem[v[n]%4];
                                    if (dem[v[n]%4]==5)
                                        {
                                            bool check=true;
                                            if (v[4]/4-v[0]/4==4)
                                                {
                                                    for (ll n=1; n<5; n++) if (v[n]/4-v[n-1]/4!=1) check=false;
                                                }
                                            else check=false;
                                            if (check==true) ++thungphasanh;
                                        }
                                }
                        }
    return thungphasanh;
}
```

Hình 7 - Code kiểm tra tính đúng đắn của công thức

Kết quả hoàn toàn trùng khớp với những gì chúng tôi đã tính toán ở trên, và kết quả của chương trình như sau:

```
PS D:\YSC> cd "d:\YSC\" ; if ($?) { g++ kiemtra.cpp -o kiemtra } ; if ($?) { .\kiemtra }
Thung pha sanh: 40
Tu quy: 624
Cu lu: 3744
Thung: 5108
Sanh: 10200
Xam chi: 54912
Thu: 123552
Doi: 1098240
Mau thau: 1302540
```

Hình 8 - Kết quả code kiểm tra tính đúng đắn của công thức

4. KẾT LUẬN

Các trò chơi và mô hình khuyến mãi trên thị trường hiện nay là nhu cầu tất yếu và đòi hỏi thực tế từ nền kinh tế thị trường. Tuy nhiên các trò chơi và mô hình này xuất hiện khuyết điểm lớn là không ổn định và không đảm bảo tối ưu về tỷ lệ thắng và thua. Mặc dù các mô hình hiện nay có thể điều chỉnh tỷ lệ theo mong muốn nhưng vẫn chưa có thể thống kê lại các giá trị kì vọng đối với từng trường hợp xảy cũng như việc đề xuất tỷ lệ đổi thưởng hợp cho người dùng khi so với tổng quỹ thưởng ban đầu. Bên cạnh đó một số mô hình sinh code khuyến mãi vẫn chưa tối ưu vẫn có thể xảy ra tình trạng trùng lặp, không thể suy ngược lại mã hóa ban đầu hay mất quá lâu trong việc sinh code.

Vì vậy kết quả trong bài nghiên cứu này là một trong những giải pháp để giải quyết các tình trạng nêu trên. Dựa trên những kết quả mà chương trình của nhóm đưa ra người dùng có thể dễ dàng phân bố cơ cấu giá trị giải thưởng đồng thời đảm bảo tỷ lệ thắng thua hợp lý. Chương trình của nhóm không chỉ đơn thuần là tính toán mà còn giải quyết các vấn đề mã hóa mã khuyến mãi trong các chương trình rút thăm, tri ân của các công ty, dịch vụ... Kết quả khích lệ của bài báo cáo này là cơ sở cho các nghiên cứu sâu hơn cũng như cung cấp giải pháp để giải quyết các bài toán tri ân khuyến mãi khách hàng, của thị trường, việc mã hóa mã khuyến mãi của lĩnh vực thương mại điện tử và các chủ đề nghiên cứu về giá trị kỳ vọng trong nền kinh tế thị trường.

TÀI LIỆU THAM KHẢO

- [1] Robert Morello, *The Effects of Advertising and Promotion*, 01/2022. [Online]. Available: <https://smallbusiness.chron.com/effects-advertising-promotion-59158.html>.
- [2] Website wikipedia.org, *Redemption game*. [Online]. Available: https://en.wikipedia.org/wiki/Redemption_game.
- [3] User cfeld2stanfordedu of website stanford.edu, *Getting started with Python*. [Online]. Available: <https://opensource.stanford.edu/getting-started-python>.
- [4] Website wikipedia.org, *Python programming language*. [Online]. Available: [https://en.wikipedia.org/wiki/Python_\(programming_language\)](https://en.wikipedia.org/wiki/Python_(programming_language)).
- [5] Website wikipedia.org, *C++*. [Online]. Available: <https://en.wikipedia.org/wiki/C%2B%2B>.
- [6] Bilal Naeem, *Module PyQt5*. [Online]. Available: <https://www.educative.io/answers/what-is-pyqt>.
- [7] Alan Hájek, *Interpretations of Probability*, 2019. [Online]. Available: <https://plato.stanford.edu/entries/probability-interpret/>.
- [8] Website wikipedia.org, *Mậu binh*. [Online]. Available: https://vi.wikipedia.org/wiki/M%E1%BA%ADu_binh.
- [9] Website wikipedia.org. [Online]. https://en.wikipedia.org/wiki/Zipf%27s_law.
- [10] Website vinahost.vn, *[Coupon là gì] – Voucher là gì | Coupon khác gì Voucher*. [Online]. Available: <https://vinahost.vn/coupon-la-gi/>.

YSC4F.239

SỬ DỤNG THUẬT TOÁN PHÂN CỤM ĐỂ ĐỀ XUẤT VỊ TRÍ TRẠM XE ĐƯA ĐÓN NHÂN VIÊN

ĐÀO XUÂN HOÀNG TUẤN, NGUYỄN VĂN NAM, NGUYỄN ĐÌNH NGUYỄN BẮC,
NGUYỄN NĂNG ANH, ĐỒNG MẠNH DŨNG, LÊ PHÚC LŨ

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
hoangtuan.salmon@gmail.com, nguyenvannam14056969@gmail.com, nguyenbac.iuh@gmail.com,
nanganha10@gmail.com, dongdung0407@gmail.com, lephucclu@iuh.edu.vn

Tóm tắt. Việc tổ chức đưa đón nhân viên đi làm mỗi buổi sáng là một hoạt động thiết yếu của nhiều công ty lớn. Với lượng xe có hạn và thông tin khảo sát về chỗ ở của các nhân viên, người quản lý sẽ cân nhắc việc bố trí các trạm xe buýt sao cho hợp lý, giảm thiểu tổng quãng đường di chuyển của nhân viên. Nhằm giải quyết bài toán đó, trong nghiên cứu này, nhóm chúng tôi tìm hiểu sử dụng thuật toán phân cụm K-means để chia nhỏ thành các khu vực mà nhân viên ở gần nhau, dựa trên dữ liệu đầu vào là danh sách địa chỉ nhân viên, sau đó chúng tôi mã hóa địa chỉ bằng ArcGIS API và tính toán ma trận khoảng cách theo đường chim bay dựa trên kinh độ - vĩ độ quy đổi sang tọa độ cực. Trên cơ sở có các cụm, nhóm cũng tìm hiểu triển khai thuật toán Convex Hull tìm bao lồi cho các khu vực và trực quan hóa lên bản đồ thành các miền riêng biệt cho dễ quan sát.

Từ khóa. trạm xe đưa đón, thuật toán phân cụm K-means, thuật toán Convex Hull, ArcGIS API.

USING CLUSTERING ALGORITHMS TO RECOMMEND THE LOCATIONS OF SHUTTLE BUS

Abstract. Nowadays, the picking up of the staff by shuttle bus every morning is a significant activity of every big company. With a limited number of vehicles and the information from the staff's residence survey, the manager can consider some fixed bus stations that minimize the staff's moving distances. To solve this problem, in this research, our group used K-means clustering algorithm to subdivide by area of employees near each other, based on the input data is a list of employee addresses, we then encode the address using the ArcGIS API and calculate distance matrix as crow flies based on longitude - latitude converted to polar coordinates. Based on the clusters, our group also study the convex hull algorithm to find out the boundary of every group of staff and visualize it on the map for viewing purpose.

Keyword. shuttle bus station, clustering algorithm K-means, Convex Hull algorithm, ArcGIS API.

1. ĐẶT VẤN ĐỀ

Hiện nay việc đi làm hàng ngày có thể là một thách thức đối với mỗi người và mất nhiều thời gian hơn bình thường vì các phương tiện giao thông công cộng còn nhiều mặt hạn chế, các vấn đề quá tải, các quy định an toàn, dịch bệnh, kẹt xe, hình ảnh xe buýt đông đúc và ngột ngạt, tập nập trên con phố lúc 8 giờ sáng ... đối với những nhân viên ở xa nơi làm việc, thì việc đi làm sẽ là rất khó khăn, điều đó làm giảm năng suất và hiệu suất làm việc của nhân viên. Trên thực tế, theo một cuộc khảo sát của Robert Half [1], gần một phần tư số người lao động đã phải rời bỏ việc làm do đường đi làm không tốt.

Vì vậy trong các công ty, tập đoàn, doanh nghiệp lớn việc đưa đón nhân viên mỗi bữa sáng là một hoạt động thiết yếu, ví dụ thực tế như Samsung Việt Nam, họ đã có một hệ thống xe buýt với gần 1,000 xe và nhiệm vụ của nó là đưa đón các nhân viên đi làm hàng ngày ở trong phạm vi 60 km xung quanh nhà máy [2]. Một số công ty lớn trong ngành vận tải nhân viên như Arjun Travels, Ambassador Tours & Travels,

Safetrax, Vee Vee Bus Services, Wings Travels, Transdev ... Tại Hoa Kỳ, Transdev đã vận chuyển 6.000 người ở San Francisco đến nơi làm việc của họ ở Thung lũng Silicon, cho các công ty như Apple, Google, Microsoft và Genentech [3]. Với số lượng xe có hạn thì doanh nghiệp cần phải giải bài toán cách để quản lý và bố trí các trạm xe buýt hợp lý nhằm giảm thiểu quãng đường di chuyển của nhân viên để đến với nơi làm việc. Nếu bài toán này được giải quyết hiệu quả thì sẽ đem lại một lợi ích to lớn đối với doanh nghiệp. Về mặt kinh tế sẽ tiết kiệm cho doanh nghiệp một chi phí như là giảm chi phí nhiên liệu, chi phí xe đưa đón, chi phí tuyển tài xế, nhân viên sẽ tiết kiệm được thời gian khi đến nơi làm việc do quãng đường đã được giảm thiểu đáng kể và sẽ không phát sinh thêm chi phí đi lại vì đã có doanh nghiệp đưa đón. Về phía công ty, nhân viên đi làm đầy đủ và đúng giờ sẽ là tiền đề để đảm bảo chất lượng công việc cho doanh nghiệp. Về mặt xã hội, nhu cầu cho nhân viên đi làm bằng xe riêng giảm, giảm áp lực lên hạ tầng giao thông, giảm tắc đường. Về mặt môi trường, khi mà số lượng xe giảm thiểu, quãng đường đi được rút gọn thì sẽ góp phần vào việc giảm đáng kể lượng khí thải carbon dioxide bảo vệ môi trường.

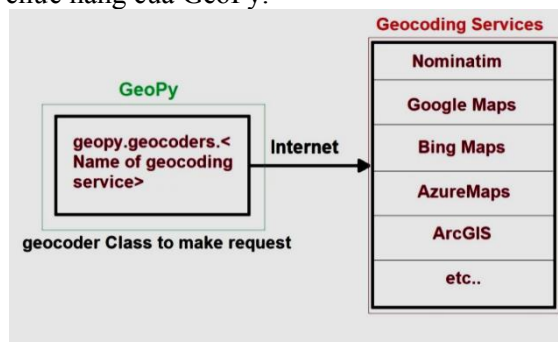
Chính vì vậy, nhóm chúng tôi đã nghiên cứu thuật toán phân cụm để giải quyết vấn đề này bằng cách thu thập thông tin địa chỉ của nhân viên để quy đổi sang kinh độ và vĩ độ, trên cơ sở các thông tin đó chúng tôi sử dụng thuật toán phân cụm để trực quan hoá vị trí lên bản đồ và xác định các điểm tọa độ phù hợp với số cụm qua đó đề xuất phương án đặt vị trí đặt xe tốt nhất. Dựa trên những thống kê nghiên cứu và vận dụng các thuật toán để đưa ra giải pháp cho bài toán đặt trạm xe buýt đáp ứng với nhu cầu và phù hợp với chiến lược của doanh nghiệp.

2. MÃ HÓA ĐỊA CHỈ

Như chúng tôi đã nói ở trên, dữ liệu địa chỉ đầu vào là danh sách địa chỉ nhân viên trong đó chứa các thông tin như: số nhà, tên đường, quận/huyện, tỉnh thành và thành phố và từ dữ liệu thô này, chúng tôi sẽ xử lý trả về kinh độ, vĩ độ. Sau một thời gian nghiên cứu, chọn lọc, chúng tôi đã đưa ra hai phương án tối ưu nhất, phù hợp: Dịch vụ mã hóa địa lý Geocode API (Application Programming Interface – API) và Selenium truy vấn lấy kinh độ, vĩ độ.

2.1. Dịch vụ mã hóa địa lý Geocode API

Để lấy được kinh độ vĩ độ, thì GeoPy một thư viện của ngôn ngữ Python được xem như là một ứng dụng khách, cầu nối, liên kết với một số dịch vụ web mã hóa địa lý phổ biến để truy xuất ra kinh độ, vĩ độ. Hình dưới đây cho biết ý tưởng về chức năng của GeoPy:



Hình 1. Ảnh ví dụ các bên thứ ba liên kết với GeoPy.

Như trong hình trên mã hóa địa lý được cung cấp bởi một trong số nhiều dịch vụ khác nhau chúng ta có thể lựa chọn tùy vào nhu cầu hiện tại của khách hàng. Mỗi dịch vụ định vị địa lý, chẳng hạn như Google Maps, Bing Maps hoặc Nominatim, ArcGIS có một lớp riêng trong `geopy.geocoders` tóm tắt API của dịch vụ. Mỗi bộ mã hóa địa lý xác định ít nhất một phương thức mã hóa địa lý, để phân giải một vị trí từ một chuỗi từ đó truy xuất thành một cặp tọa độ và có thể xác định một phương pháp, ngược lại, phương pháp này phân giải một cặp tọa độ thành một địa chỉ [4].

Ở bài báo này, chúng tôi sử dụng thư viện GeoPy kết hợp với dịch vụ ArcGIS với ưu điểm là nhanh chóng trong việc truy xuất kinh độ, vĩ độ và họ cung cấp 1 triệu yêu cầu truy vấn miễn phí mỗi tháng. Nhưng có nhược điểm là độ chính xác thấp và độ lệch trung bình chúng tôi nghiên cứu khoảng 50m cho một địa chỉ. Cách thức lấy kinh độ, vĩ độ:

Đầu tiên, chúng tôi gọi ra dịch vụ mã hóa địa lý ArcGIS tạo ra một phiên bản của lớp ArcGIS. Sau khi tạo lớp, chúng tôi áp dụng phương pháp mã hóa địa lý để truy xuất ra vị trí từ dữ liệu địa chỉ đầu vào. Sử dụng phương thức ‘latitude’ và ‘longitude’ để lấy kinh độ vĩ độ.

2.2. Selenium truy vấn lấy kinh độ, vĩ độ

Phương pháp này được chúng tôi nghĩ ra thông qua nhiều lần tra địa chỉ thực tế trên Google Chrome. Chúng tôi nhận thấy rằng mỗi lần tra cứu một địa chỉ thực tế nào đó thì kinh độ, vĩ độ sẽ được cập nhật thông qua đường dẫn. Và mục tiêu chúng tôi là lấy đường dẫn đó về và sử dụng phương thức Regex trích xuất ra kinh độ, vĩ độ. Ưu điểm của phương pháp này cho kết quả tương đối chính xác, và không tốn kém chi phí nhưng nhược điểm là việc truy vấn lâu.

Trước tiên, chúng tôi đã tìm địa chỉ “Trường Đại học Công nghiệp Thành phố Hồ Chí Minh”:

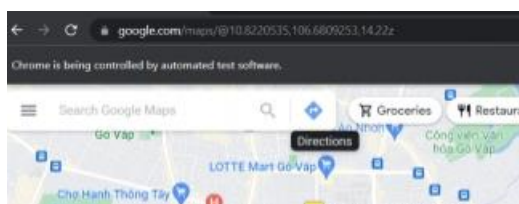


Hình 2. Ảnh ví dụ về lấy kinh độ, vĩ độ.

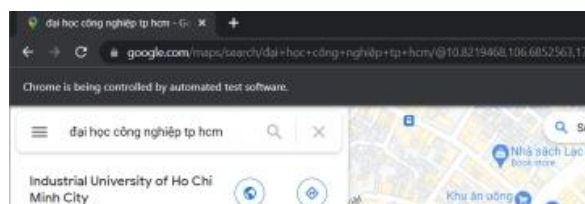
Và để thực hiện ý tưởng này, chúng tôi đã sử dụng thư viện Selenium của Python cho phép tương tác từng bước với trang web của Google Maps. Sau đó, chúng tôi cài đặt ChromeDriver để kiểm tra tự động các ứng dụng web trên trình duyệt Google Chrome, cung cấp khả năng điều hướng đến trang web và nhập dữ liệu địa chỉ đầu vào để tìm kiếm. Nhưng để truy cập tới trang Google Maps thì chúng tôi phải khởi tạo phiên bản WebDriver của Google Chrome truyền vào đường dẫn lưu trữ ChromeDriver.

Sau đó chúng tôi tận dụng phương thức driver.get để điều hướng đến trang web Google Maps thông qua đường dẫn <https://www.google.com/maps/>. WebDriver sẽ đợi đến khi trang được tải đầy đủ trước khi trả lại quyền kiểm soát cho thử nghiệm hoặc tập lệnh của chúng tôi.

Tiếp đến chúng tôi tạo ra một hàm tìm kiếm tự động trên Google Maps bằng cách xác định tên class trên khung tìm kiếm trong trang Google Maps truyền vào phương thức find_element_by_name để định vị, vị trí nhập địa chỉ. Tiếp theo, chúng tôi gửi chuỗi địa chỉ cần tìm vào, điều này tương tự như nhập các phím bằng bàn phím. Và sử dụng phương thức send_key (Keys.RETURN), tương tự như nút tìm kiếm trên trang. Sau đó đường dẫn mới trên trang Google Maps sẽ được cập nhật mới.



Hình 3. Đường dẫn trước khi tìm kiếm.



Hình 4. Đường dẫn sau khi tìm kiếm.

Rõ ràng sau khi tìm kiếm chúng tôi đã có được đường dẫn và việc sau đó là chúng tôi sẽ lấy đường dẫn đó về bằng phương thức driver.current_url. Việc cuối cùng, chúng tôi sử dụng RegEx để trích xuất ra kinh độ vĩ.

3. TÍNH KHOẢNG CÁCH THEO LÝ THUYẾT

Để đo khoảng cách vị trí giữa các nhân viên chúng tôi sử dụng công thức Haversine là công thức dùng trong xác định khoảng cách đường tròn lớn giữa hai điểm trên một mặt cầu dựa trên kinh độ và vĩ độ của hai điểm đó[5]. Đường hasrsine có thể được biểu thị dưới dạng hàm lượng giác như sau:

$$Haversine(\theta) = \text{Sin}^2\left(\frac{\theta}{2}\right) \quad (1)$$

Đường hasrsine của góc ở tâm (là d / r) được tính theo công thức sau:

$$\left(\frac{d}{r}\right) = \text{havesine}(\Phi_2 - \Phi_1) + \cos(\Phi_1) \cos(\Phi_2) \text{haversine}(\lambda_2 - \lambda_1) \quad (2)$$

Trong đó r là bán kính trái đất (6371 km), d là khoảng cách giữa hai điểm, Φ_1, Φ_2 là vĩ độ của hai điểm, và λ_1, λ_2 là kinh độ của hai điểm tương ứng [6].

Giải d bằng cách sử dụng hàm sin nghịch đảo, chúng ta nhận được:

$$d = 2 \cdot r \cdot \arcsin\left(\sqrt{\sin^2\left(\frac{\Phi_2 - \Phi_1}{2}\right) + \cos(\Phi_1) \cos(\Phi_2) \sin^2\left(\frac{\lambda_2 - \lambda_1}{2}\right)}\right) \quad (3)$$

Kết hợp với thư viện `math` của python để xây dựng chương trình tính khoảng cách và biểu diễn thành ma trận nhiều vị trí khác nhau của các nhân viên với dữ liệu vào ban đầu là số vị trí “ n ” và các dòng tiếp theo lần lượt là “`lat A lon A`”, “`lat B lon B`”, ..., “`lat n lon n`”

Mô tả: Cách tính ma trận khoảng cách

```
1. import math
2. def Haversine_formula(lat1, long1, lat2, long2):
3.     lat1 = math.radians(lat1)
4.     # Hoặc lat1 = lat1 * math.pi/180
5.     lat2 = math.radians(lat2)
6.     # Hoặc lat2 = lat2 * math.pi/180
7.     long1 = math.radians(long1)
8.     # Hoặc long1 = long1 * math.pi * 180
9.     long2 = math.radians(long2)
10.    # Hoặc long2 = long2 * math.pi * 180
11.    temp = math.asin(math.sqrt(math.sin((lat2-lat1)/2)**2 + math.cos(lat1) *
12.        math.cos(lat2) * math.sin((long2-long1)/2)**2))
13.    d = 2 * r * temp
14.    return round(d, 3)
15. r = 6371 # bán kính trái đất theo dữ liệu từ google
16. n = int(input())
17. lati = []
18. longi = []
19. for i in range(n):
20.     x, y = map(float, input().split())
21.     lati.append(x)
22.     longi.append(y)
23. for i in range(len(lati)):
24.     for j in range(len(longi)):
25.         print(Haversine_formula(lati[i], longi[i], lati[j], longi[j]), end
26.             = " ")
27.     print()
```

4. THUẬT TOÁN PHÂN CỤM K-MEANS CLUSTERING

4.1. Giới thiệu về thuật toán K-means

K-means được ra đời bởi nhà toán học người Ba Lan là Hugo Steinhaus vào năm 1956. Thuật toán phân cụm được giới thiệu trong bài báo bằng tiếng Pháp “*Sur la Divison de Cops Matériels en Parties*”. Năm 1967, James MacQueen đưa ra thuật ngữ K-means trong bài báo cáo “*Some Methods for Classification and Analysis of Multivariate Observations*”.

Thuật toán phân cụm K-means là phương pháp phân các dữ liệu điểm đầu vào thành các cụm khác nhau. Thuật toán K-means có nhiều ứng dụng, trong đó trí tuệ nhân tạo và học máy là sử dụng nhiều nhất.

Phương pháp của thuật toán K-means là tạo và cập nhật liên tục điểm trung tâm và phân các điểm dữ liệu vào các nhóm khác nhau. Ban đầu, tạo ra trung điểm ngẫu nhiên và gán mỗi điểm từ tập dữ liệu vào tâm

gần nó nhất. Sau đó thuật toán sẽ cập nhật điểm trung tâm và lặp lại các bước cho tới khi được kết quả gần như chính xác. [7]

4.2 Độ phức tạp của thuật toán K-means

4.2.1 Mô tả thuật toán

Cho một tập hợp điểm (x_1, x_2, \dots, x_n) , phân cụm K-means giúp chia n điểm thành k ($\leq n$) cụm với tập hợp cụm $S = \{S_1, S_2, \dots, S_k\}$. Được tuân theo quy tắc:

$$\underset{S}{\operatorname{argsin}} \sum_{i=1}^k \sum_{x \in S_i} \|x - \mu_i\|^2 = \underset{S}{\operatorname{argsin}} \sum_{i=1}^k |S_i| \operatorname{Var} S_i \quad (4)$$

Với μ_i là giá trị trung bình của các điểm trong S_i . Điều này tương đương với việc giảm thiểu độ lệch bình phương theo cặp của các điểm trong cùng một cụm:

$$\underset{S}{\operatorname{argsin}} \sum_{i=1}^k \frac{1}{|S_i|} \sum_{x, y \in S_i} \|x - y\|^2 \quad (5)$$

Từ đó, ta có thể suy ra công thức tương đương:

$$|S_i| \sum_{x \in S_i} \|x - \mu_i\|^2 = \sum_{x \neq y \in S_i} \|x - y\|^2 \quad (6)$$

Vì tổng phương sai không đổi nên với công thức trên có thể tối đa hóa tổng bình phương độ lệch giữa các điểm trong các cụm khác nhau.[7]

4.2.2. Thuật toán tiêu chuẩn

Với tập k điểm trung tâm ban đầu m_1, m_2, \dots, m_k thuật toán được thực hiện liên tục và lặp lại hai bước:

Bước chỉ định: Gán mỗi điểm cho cụm có giá trị trung bình gần nhất với khoảng cách Euclide bình phương nhỏ nhất. Ta có:

$$S_i = \{x_p : \|x_p - m_i\|^2 \leq \|x - m_j\|^2 \forall j, 1 \leq j \leq k\} \quad (7)$$

Mỗi điểm x_p được chỉ định cho chính xác một tập S .

Bước cập nhật: Tính toán lại giá trị trung bình (trọng tâm) cho từng cụm:

$$m_i = \frac{1}{|S_i|} \sum_{x_j \in S_i} x_j \quad (8)$$

Thuật toán được dừng lại khi giá trị của các phép gán không còn thay đổi. [7]

4.3. Ưu điểm và nhược điểm của thuật toán K-means

Thuật toán này rất phù hợp khi số cụm K cho trước luôn ổn định, có khả năng mở rộng và dễ dàng sửa đổi với các dữ liệu phân cụm mới, trải qua một vài vòng lặp hữu hạn thì vẫn giữ được tính hội tụ của một cụm, các cụm cũng sẽ được tách biệt rõ ràng, không dễ xảy ra hiện tượng một đối tượng xuất hiện chung với các cụm dữ liệu khác. Đặc biệt hơn là chúng tạo ra cụm chặt chẽ hơn so với các thuật toán phân cụm khác nếu các cụm có dạng hình cầu. Tuy nhiên vẫn có một số hạn chế như cần phải xác định số cụm là bao nhiêu trước khi triển khai phân cụm, không thể đảm bảo đạt được tối ưu toàn cục, từ đó kết quả in ra sẽ phụ thuộc nhiều vào việc chọn số K cụm để triển khai đồng thời khó khăn cho việc xử lý dữ liệu ngoại lệ, và chỉ khi tính được trọng tâm của cụm mới ứng dụng được thuật toán này [8].

4.4 Các bước thực hiện thuật toán

Bước 1:

Chúng tôi nhận đầu vào là một tệp excel (có dạng .xlsx). Sau đó sử dụng thư viện pandas để đọc file lưu vào trong biến dạng DataFrame, và trong file đó ít nhất phải chứa cột “Address” chứa địa chỉ thực. Mỗi phần tử của danh sách là một danh sách khác chứa các giá trị mục cho các đối tượng địa lý.

G	H	I	J
Họ đệm	Tên	Giới tính	Address
1	Triệu Quốc An	Nam	1, Hoàng Cầu, Ô Chợ Dừa, Đống Đa, Hanoi, Vietnam,
1	Nguyễn Đức Anh	Nam	3, Thành Công, Khu tập thể Bắc Thành Công, Thành Công, Ba Đình District, Hanoi, Vietnam,
1	Dương Thế Bảo	Nam	Ngõ 3 Thái Hà, Trung Liệt, Hanoi, Hoàn Kiếm District, Hanoi, Vietnam,

Hình 5. Danh sách chứa các đối tượng địa lý đầu vào.

Bước 2:

Khởi tạo K trọng tâm bất kì.

Bước 3:

Sau đó chúng tôi tính ma trận khoảng cách. Gán các điểm vào cụm có trọng tâm gần nó nhất.

Bước 4:

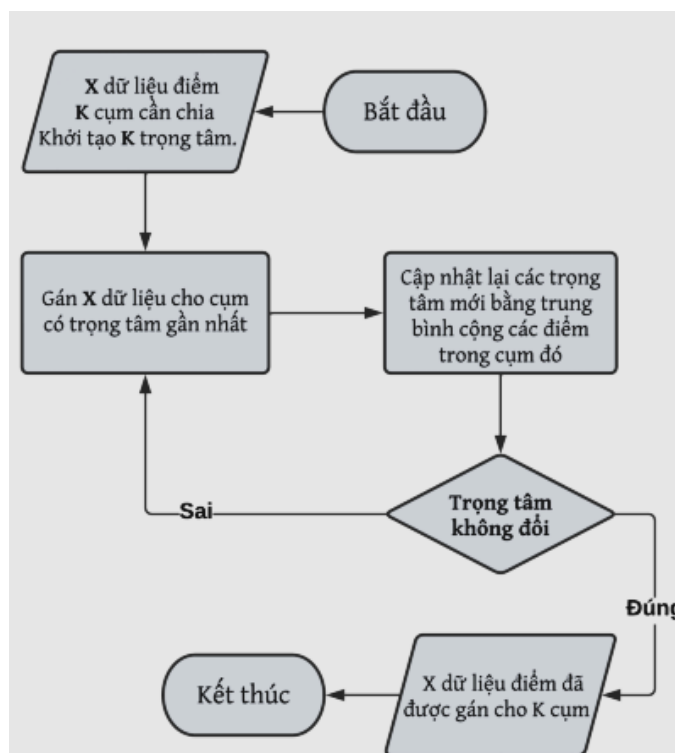
Nếu việc gán dữ liệu vào cụm ở bước 02 không thay đổi so với vòng lặp trước đó thì ta dừng và đưa ra kết quả.

Bước 5:

Cập nhật các trọng tâm mới bằng cách lấy trung bình cộng của các điểm được gán ở mỗi cụm trước đó.

Bước 6:

Trở lại bước 02.



Hình 6. Sơ đồ tóm tắt các bước thực hiện của thuật toán.

5. SỬ DỤNG THUẬT TOÁN VÀO BÀI TOÁN PHÂN CỤM XE ĐƯA RƯỚC

5.1. Cách thức cài đặt

Như đã biết thì thuật toán K-means được dùng để gom dữ liệu đã cho thành các cụm dữ liệu, sao cho dữ liệu đó đến với tâm cụm gần nó nhất. Dựa trên những đặc tính của thuật toán K-means, nhóm nghiên cứu đã áp dụng vào việc phân cụm nơi ở nhân viên thành các cụm các khu vực sao cho đường đi từ nơi ở của nhân viên đến vị trí tâm cụm đó là ngắn nhất. Theo dữ liệu thu thập được là vị trí chỗ ở của các nhân viên cần đưa đón.

Dữ liệu đầu vào theo đúng các trường như nhau:

<Số nhà>, <Tên đường>, <Phường/Xã>, <Quận/Huyện>, <Thành/Thị xã/Huyện>, <Quốc Gia>.

Ví dụ

12, Nguyễn Văn Bảo, Phường 04, Quận Gò Vấp, Thành phố Hồ Chí Minh, Việt Nam.

Qua thông tin địa chỉ thu thập được, bằng kỹ thuật dữ liệu nhóm nghiên cứu chuyển đổi những dòng địa chỉ đó trở thành tọa độ tương ứng.

Ví dụ với địa chỉ “12, Nguyễn Văn Bảo, Phường 04, Quận Gò Vấp, Thành phố Hồ Chí Minh, Việt Nam”.

Thì ta sẽ thu được cặp kinh độ, vĩ độ là 10.8222053, 106.6874994.

Dựa vào các địa chỉ, chúng tôi đã viết chương trình để chuyển đổi địa chỉ thực thành tọa độ địa lí tương ứng mà máy tính có thể tính toán và xử lý được.

Sau khi chuyển hoá hết toàn bộ thông tin vị trí, đây là lúc mà thuật toán K-means được đưa vào áp dụng, với dữ liệu đầu vào là các điểm, sau đó tính toán khoảng cách địa lý giữa các điểm với nhau. Cuối cùng là phân các điểm đó thành K cụm tương ứng (Với số K phụ thuộc vào yêu cầu của công ty nhưng vẫn đảm bảo số lượng cụm phải nhỏ nhất).

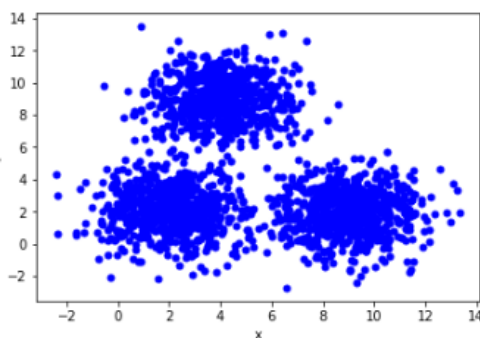


Hình 7. Kết quả của nhóm đạt được sau khi phân cụm.

5.2. Mô phỏng

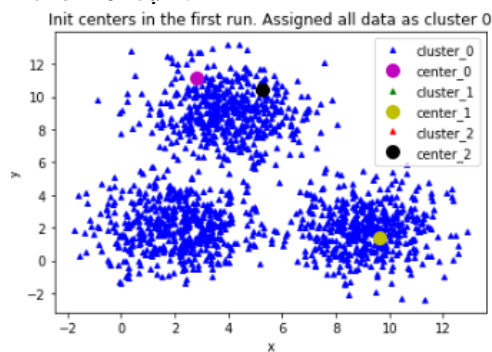
Thuật toán K-means là thuật toán tuân theo một quy trình lặp đi lặp lại trong đó, nó cố gắng giảm thiểu khoảng cách của các điểm dữ liệu so với các điểm trọng tâm [9] nhờ đó mà độ chính xác giữa các điểm được tối đa hóa. Để chứng minh cho điều này chúng tôi đã thực hiện một mô phỏng về quá trình thực hiện của thuật toán K-means.

Đầu tiên, chúng tôi khởi tạo 500 điểm và số trung tâm là 3:



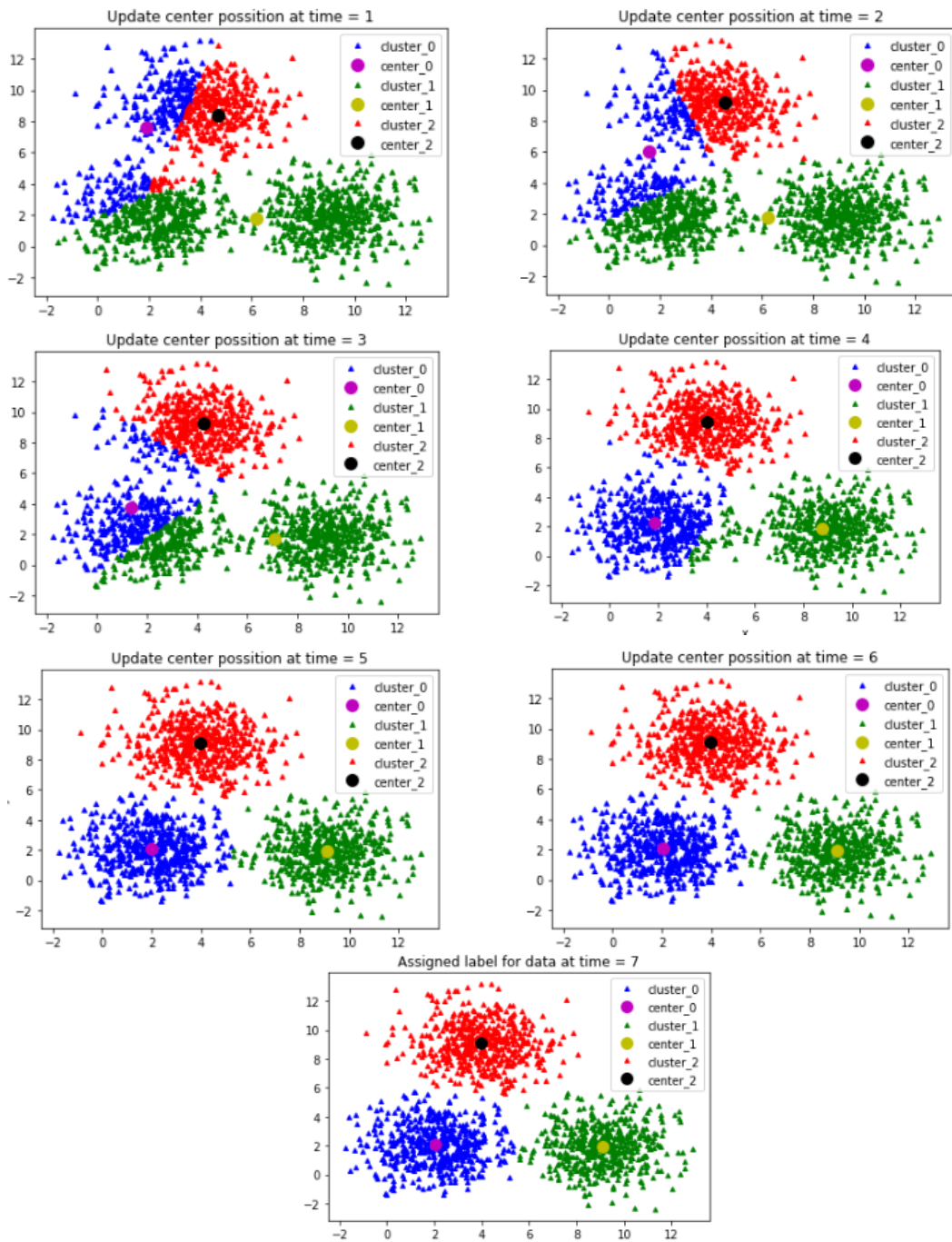
Hình 8. Các điểm được trực quan hóa.

Sau đó chúng tôi sử dụng thuật toán K-means bắt đầu khởi tạo 3 điểm trung tâm, do các điểm trung tâm không thể xác định một cách chính xác. Vì thế chúng tôi sẽ khởi tạo 3 điểm trung tâm một cách ngẫu nhiên và xác định chúng làm trung tâm cho mỗi cụm.



Hình 9. Các điểm trung tâm được khởi tạo ngẫu nhiên.

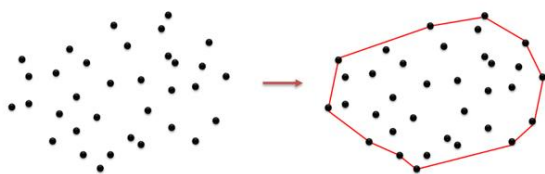
Khi khởi tạo xong 3 điểm trung tâm, chúng tôi bắt đầu cập nhật tâm cho từng cụm liên tục và chỉ dừng nếu như tọa độ của tâm không thay đổi:



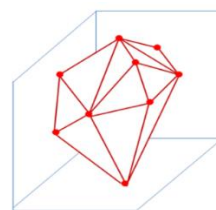
Hình 10. Mô tả quá trình cập nhật 3 cụm.

6. THUẬT TOÁN CONVEX HULL

Trong hình học tính toán (computational geometry), bao lồi (Convex Hull) là phần lồi của một tập hợp các điểm được xem là tập lồi nhỏ nhất sao cho diện tích nhỏ nhất khi ở trong hình học phẳng và thể tích nhỏ nhất trong hình học không gian mà chứa mọi điểm trong tập hợp đó [10].



Hình 11. Bao lồi trong hình học phẳng.



Hình 12. Bao lồi trong hình học không gian.

Hiện nay thuật toán để tìm bao lồi được áp dụng cho nhiều lĩnh vực khác nhau như toán học, thống kê, tối ưu hóa tổ hợp, kinh tế, mô hình hóa hình học, vật lý lượng tử, khoa học vật liệu, khoa học dữ liệu,...

Khi nhìn bằng mắt vào một tập hợp các điểm, con người chúng ta có thể dễ dàng để nhận biết được đâu là vỏ lồi và nó đi qua những điểm nào, nhưng để máy tính làm được việc này thì không đơn giản như vậy, và chúng ta cần viết những đoạn mã để xử lý công việc này. Để tìm tập hợp các điểm thuộc phần lồi chúng ta có thể sử dụng thuật toán sau.

Đầu tiên chúng ta cần tạo một class (python) để lưu tọa độ kinh độ vĩ độ của các địa chỉ

Mô tả: Tạo class lưu kinh độ, vĩ độ.

```
1. class Point:
2.     def __init__(self, x, y):
3.         self.x = x
4.         self.y = y
```

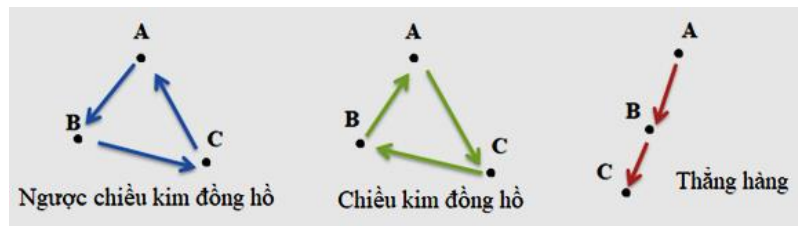
Sau đó tạo ra một hàm để tìm ra một điểm bất kì nằm trên bao lồi ví dụ tọa độ trên cùng bên trái, hàm dưới đây sẽ trả về index của tọa độ ngoài cùng bên trái và nếu có nhiều tọa độ như vậy (x bằng nhau) thì sẽ lấy tọa độ có vĩ độ (y) lớn nhất.

Mô tả: Hàm trả về tọa độ trên cùng bên trái của tập tọa độ.

```
1. def Left_index(points):
2.     #Tìm điểm xa nhất bên trái
3.     minn = 0
4.     for i in range(1,len(points)):
5.         if points[i].x < points[minn].x:
6.             minn = i
7.         elif points[i].x == points[minn].x:
8.             if points[i].y > points[minn].y:
9.                 minn = i
10.    return minn
```

Tiếp đó chúng ta viết hàm xác định vị trí của 3 điểm p, q, r
Định hướng của bộ 3 điểm có thứ tự trong mặt phẳng có thể là:

- Ngược chiều kim đồng hồ
- Chiều kim đồng hồ
- Thẳng hàng



Hình 13. Sơ đồ cho thấy các định hướng của 3 điểm A, B, C.

Ý tưởng để tính toán định hướng là tính độ dốc của từng đoạn thẳng AB, BC, ví dụ tọa độ các điểm là $A(x_1, y_1)$, $B(x_2, y_2)$, $C(x_3, y_3)$

$$\text{Độ dốc của đoạn thẳng AB là : } d_1 = (y_2 - y_1) / (x_2 - x_1) \quad (9)$$

$$\text{Độ dốc của đoạn thẳng BC là : } d_2 = (y_3 - y_2) / (x_3 - x_2) \quad (10)$$

Nếu $d_1 > d_2$ hướng theo kim đồng hồ (rẽ phải)

Từ những giá trị d_1, d_2 ở trên chúng ta có thể kết luận rằng, hướng phụ thuộc vào dấu hiệu của biểu thức dưới:

$$(y_2 - y_1) * (x_3 - x_2) - (y_3 - y_2) * (x_2 - x_1) \quad (11)$$

Nếu biểu thức dương thì A, B, C theo chiều kim đồng hồ, nếu âm thì ngược chiều kim đồng hồ, còn nếu bằng không thì A, B, C thẳng hàng[11].

Dưới đây là code thực hiện ý tưởng trên:

Mô tả: Hàm kiểm tra định hướng của ba điểm.

```

1. def orientation(p, q, r):
2.     val = (q.y - p.y) * (r.x - q.x) - (q.x - p.x) * (r.y - q.y)
3.     if val == 0:
4.         return 0
5.     elif val > 0:
6.         return 1
7.     else:
8.         return 2
    
```

Hàm sẽ trả về các giá trị sau

- 0 □ p, q và r thẳng hàng
- 1 □ Theo chiều kim đồng hồ
- 2 □ Nếu Ngược chiều kim đồng hồ

Cuối cùng là thực hiện thuật toán Convex Hull [12].

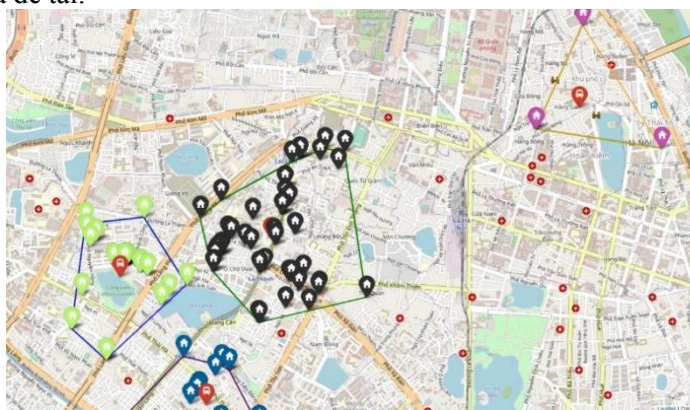
Mô tả: Thuật toán Convex Hull.

```

1. def convexHull(points, n):
2.     if n < 3:
3.         return
4.     l = Left_index(points)
5.     hull = []
6.     p = l
7.     q = 0
8.     while(True):
9.         hull.append(p)
10.        q = (p + 1) % n
11.        for i in range(n):
12.            if(orientation(points[p],
13.                            points[i], points[q]) == 2):
    
```

```
14.                                     q = i
15.             p = q
16.             if(p == 1):
17.                 break
18.     ans = []
19.     for each in hull:
20.         ans.append([points[each].x, points[each].y])
21.     return ans
```

Khi tìm được điểm chắc chắn nằm trong bao lồi thì xuất phát từ điểm này ta sẽ đi tìm các điểm khác nằm trong bao lồi cho đến khi quay trở lại điểm ta chọn lúc đầu. Độ phức tạp của thuật toán là $O(n^2)$ đủ nhanh để đáp ứng nhu cầu của đề tài.



Hình 14. Kết quả của nhóm đạt được sau khi bao lồi.

7. TỔNG KẾT

Mô hình sử dụng thuật toán phân cụm để đề xuất vị trí trạm đưa đón nhân viên với độ chính xác cao trong thực tế đã phân nào giúp giảm tình trạng kẹt xe ở những vùng đô thị đông đúc vì số lượng nhân viên trong một công ty rất lớn, đồng thời giúp giảm chi phí đi lại của mỗi nhân viên so với việc đi bằng phương tiện cá nhân. Song với đó, lượng khí thải từ các phương tiện cá nhân thải ra môi trường được giảm xuống. Đồng thời, sử dụng mô hình này sẽ tính toán được vị trí lắp đặt trạm tối ưu nhất, giúp giảm tình trạng ãn công ty muộn, từ đó thúc đẩy năng suất công việc tăng cao. Hiện nay với sự phát triển nhanh chóng của khoa học – công nghệ, thông qua quá trình nghiên cứu người dùng cũng sẽ dễ dàng tiếp cận cho việc lưu thông của mình theo ý muốn.

Kết quả nghiên cứu trên sẽ là bước đệm tiếp theo cho việc đi sâu hơn và phát triển trong lĩnh vực giao thông vận tải hiện nay. Nhằm cải tiến chất lượng dịch vụ, chúng tôi dự định sẽ tạo ra một sản phẩm web, app dựa trên bài nghiên cứu này để người dùng có trải nghiệm tốt hơn với một giao diện đẹp mắt, sáng tạo và các dịch vụ hỗ trợ người dùng có hiệu quả.

Mở rộng đối tượng sử dụng là những người đi làm, học sinh, sinh viên, ... và sau đó chúng tôi sẽ cải tiến và thêm một số tính năng như: Dự đoán số cụm cần thiết với dữ liệu đầu vào là danh sách các địa chỉ, tính khoảng cách dựa trên đường đi thực tế, tối ưu thời gian biến đổi địa chỉ về kinh độ, vĩ độ giúp ứng dụng hoạt động hiệu quả hơn, tối thiểu số lần gọi API để tiết kiệm chi phí, có tính năng chỉ đường, tính khoảng cách giữa xe buýt và trạm xe theo thời gian thực để người dùng chủ động thời gian trong việc chờ xe buýt, tài xế xe buýt xác định được con đường đi đến trạm một cách dễ dàng và nhanh nhất. Đó cũng chính là những dự kiến nghiên cứu của nhóm chúng tôi trong tương lai.

TÀI LIỆU THAM KHẢO

[1] Robert Half, Nearly One-Quarter Of Workers Have Left A Job Due To A Bad Commute (2018). [Online].

Available: <https://press.roberthalf.com/2018-09-24-Nearly-One-Quarter-Of-Workers-Have-Left-A-Job-Due-To-A-Bad-Commute-According-To-Robert-Half-Survey>.

- [2] Phạm Trung Hiền, Trần Sơn Bách, Lê Minh Sơn, Nguyễn Thị Hạnh, Ngô Xuân Quảng, Phạm Thanh Trà, Samsung Việt Nam chăm sóc nhân viên thế nào? (2020). [Online]. Available:
<https://www.vietnamplus.vn/samsung/tong-hop-cac-hoat-dong-phuc-loi-cho-nhan-vien-Samsung-VN.html>.
- [3] Transdev, Corporate and industrial transport. [Online]. Available: <https://www.transdev.com.au/our-solutions/our-worldwide-solutions/corporate-institutional-and-industrial-site-transport/>.
- [4] GeoPy, Welcome to GeoPy's documentation! (2016). [Online]. Available:
<https://geopy.readthedocs.io/en/latest/#nominatim>.
- [5] Wikipedia, Haversine formula [Online]. Available: https://en.wikipedia.org/wiki/Haversine_formula.
- [6] User prakhar7 of GeeksforGeeks, Haversine formula to find distance between two points on a sphere (2022). [Online]. Available:
<https://www.geeksforgeeks.org/haversine-formula-to-find-distance-between-two-points-on-a-sphere/>.
- [7] Wikipedia, k-means clustering. [Online]. Available: https://en.wikipedia.org/wiki/K-means_clustering.
- [8] Bùi Anh Kiệt, Thuật toán K-means và ứng dụng trong thực tế, Báo cáo môn khai phá dữ liệu và kho dữ liệu, ĐH Công nghệ thông tin – ĐHQG TP HCM, 2012.
- [9] Hans-Hermann Bock, Clustering Methods: A History of k-Means Algorithms, Springer Link, 2007.
- [10] Wikipedia, Convex-Hull [Online]. Available: http://wcipeg.com/wiki/Convex_hull.
- [11] GeeksforGeeks, Orientation of 3 ordered points (2022). [Online]. Available:
<https://www.geeksforgeeks.org/orientation-3-ordered-points/>.
- [12] Wikipedia, Bao lỗi. [Online]. Available: https://vi.wikipedia.org/wiki/Bao_l%E1%BB%93i.

YSC4F.240

XÂY DỰNG NỀN TẢNG HỌC TẬP CHẤM CODE TỰ ĐỘNG IUHCODER

TRẦN VĂN NHÂN, PHAN CHÍ TRUNG, HUỖNH QUỐC BẢO, TRẦN QUỐC TRỌNG, NGUYỄN HẢI NAM, NGUYỄN THÀNH TÍN, NGÔ GIA PHÁT, NGUYỄN HUỖNH ANH THẮNG, HOÀNG THẢO MY, THÁI THỊ HIỀN, NGUYỄN HỮU TÌNH, LÊ PHÚC LỮ

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

tranvannhan1911@gmail.com, chitrong08082001@gmail.com, quocbao642002@gmail.com, quoctrong.13022002@gmail.com, hainamdev@gmail.com, ngthanhtin68@gmail.com, ngogiaphat0802@gmail.com, anhthang764@gmail.com, hoangthaomy48@gmail.com, hienthai1234thptdh@gmail.com, nhtinh.aw@gmail.com, lephuclu@iuh.edu.vn

Tóm tắt. Nhằm phát triển cộng đồng học lập trình, tạo môi trường học tập cho các sinh viên CNTT và các học sinh THPT có định hướng theo ngành CNTT, góp phần thu hút học sinh gia nhập Trường Đại học Công nghiệp Thành phố Hồ Chí Minh, chúng tôi đã tiến hành nghiên cứu, tìm hiểu để hướng đến triển khai nền tảng học tập IUHCoder. IUHCoder là hệ thống chấm code tự động được phát triển dựa trên mã nguồn mở DMOJ. Nhóm định hướng sản phẩm sẽ cung cấp giải pháp để giáo viên tổ chức kiểm tra, đánh giá năng lực lập trình của sinh viên, cũng như chia sẻ kiến thức giữa các học sinh sinh viên nhằm tạo cầu nối giữa các học sinh THPT và Trường Đại học Công nghiệp Thành phố Hồ Chí Minh. Hệ thống sẽ được chúng tôi áp dụng các công nghệ mới như kỹ thuật bảo mật hệ thống sandbox trong quá trình chấm code, web socket trong việc kết nối thời gian thực giữa website và hệ thống chấm code. Cuối cùng triển khai nền tảng trên cloud computing mà ở đây là AWS.

Từ khóa. giải pháp dạy và học lập trình, chấm code trực tuyến, mã nguồn mở, sandbox, web sockets, điện toán đám mây

BUILDING A LEARNING PLATFORM ONLINE JUDGE IUHCODER

Abstract. In order to develop the programming learning community, create a learning environment for IT students and IT-oriented high school students, and contribute to attracting students to join the Industrial University of Ho Chi Minh City, we have conducted research, learning to work towards implementing an IUHCoder learning platform. IUHCoder is an automatic code grading system developed based on the open-source DMOJ. The product-oriented group will provide a solution for teachers to organize testing and assessment of students' programming ability and share knowledge among students to create a bridge between high school students and the Industrial University of Ho Chi Minh City. We will apply the system with new technologies such as sandbox security techniques in the code scoring process, and web sockets in the real-time connection between the website and the code scoring system. Deploying the platform on cloud computing which is AWS.

Keywords. solutions for teaching and learning programming, online judge, open-source, sandbox, web sockets, cloud computing.

1. ĐẶT VẤN ĐỀ

Trong thời đại công nghệ hiện nay, việc giảng dạy các môn học về lập trình là điều không thể thiếu và không chỉ riêng ở bậc Đại học mà còn xuất hiện ở cả THCS - THPT, thậm chí là Tiểu học. Các ngôn ngữ lập trình thường được sử dụng là: C/C++, Python,... Với các bài tập có chung một đặc điểm là: cho trước thông tin đầu vào (input), học sinh/sinh viên phải tìm cách sử dụng các câu lệnh một cách thích hợp để giải quyết bài toán và in ra kết quả (output) cho từng bộ dữ liệu kiểm tra (test case). Với số lượng học viên trong

các lớp khá đông, số lượng bài tập nhiều, việc xử lý chấm từng bài, từng test case một đòi hỏi nhiều thời gian và ít nhiều có khả năng sai sót, thiếu bao quát hết các trường hợp. Những điều đó đặt ra một nhu cầu khá thiết thực là xây dựng một nền tảng chấm code tự động, trong đó, giáo viên sẽ chuẩn bị sẵn các đề bài cùng bộ test, học viên khi làm bài chỉ cần gửi đoạn code của mình lên hệ thống và được biên dịch, chạy ra kết quả và so khớp với các output cho trước. Các công việc này sẽ giúp tự động hóa các quy trình làm việc thủ công, việc chấm điểm có thể diễn ra ngay lập tức khi học viên nộp bài.

Thêm vào đó, thực trạng hiện nay cho thấy khó để có thể tìm kiếm được các nền tảng nào vừa chấm code tự động vừa có thể quản lý lớp học, các kỳ thi, các bài tập, bài nộp của học sinh, sinh viên một cách dễ dàng. Giáo viên cho dù có công cụ để chấm test tự động nhưng vẫn gặp không ít gặp khó khăn trong việc thống kê điểm số, đánh giá kết quả của lớp hoặc tạo ra các kỳ thi học thuật riêng. Về phía học sinh, sinh viên, đối tượng chính sử dụng nền tảng này cũng đang còn thụ động trong việc tiếp cận lập trình và gặp khó khăn trong quá trình học tập, tìm kiếm các nguồn tài liệu phù hợp với mục đích và năng lực của mình.

Khảo sát trực tiếp từ một số giảng viên khoa Công nghệ thông tin của trường Đại học Công nghiệp TP.HCM, nhóm chúng tôi cũng thấy được các yêu cầu này từ phía các thầy cô, cụ thể là với các môn như: Nhập môn lập trình, Kỹ thuật lập trình, Cấu trúc dữ liệu và giải thuật,... Đồng thời, cũng là các thành viên của Câu lạc bộ lập trình H3.2PL, nhóm thấy được tầm quan trọng của việc xây dựng nền tảng chấm code tự động trong công tác tuyển chọn thành viên đầu vào, tổ chức các kỳ thi hàng tuần, hàng tháng để rèn luyện cho mục tiêu lập trình thi đấu, Olympic Tin học. Và mở rộng hơn nữa, khoa CNTT của nhà trường đang ngày một phát triển và xây dựng hình ảnh vững chắc, uy tín, song song đó là với yêu cầu sáng tạo ra thêm những phương pháp mới, những kênh thông tin thích hợp để góp phần thu hút, chiêu mộ được các học sinh THPT khắp nơi trên cả nước.

2. CÁC CÔNG NGHỆ LIÊN QUAN

2.1. Ngôn ngữ lập trình và các framework phát triển ứng dụng

2.1.1. Ngôn ngữ lập trình Python

Python là một ngôn ngữ lập trình bậc cao cho các mục đích lập trình đa năng, do Guido van Rossum tạo ra và lần đầu ra mắt vào năm 1991. Python được thiết kế với ưu điểm mạnh là dễ đọc, dễ học và dễ nhớ. Python là ngôn ngữ có hình thức rất sáng sủa, cấu trúc rõ ràng, thuận tiện cho người mới học lập trình và là ngôn ngữ lập trình dễ học; được dùng rộng rãi trong phát triển trí tuệ nhân tạo. Cấu trúc của Python còn cho phép người sử dụng viết mã lệnh với số lần gõ phím tối thiểu [1].

2.1.2. Ngôn ngữ lập trình Javascript

JavaScript, theo phiên bản hiện hành, là một ngôn ngữ lập trình thông dịch được phát triển từ các ý niệm nguyên mẫu. Ngôn ngữ này được dùng rộng rãi cho các trang web (phía người dùng) cũng như phía máy chủ (với Nodejs). Nó vốn được phát triển bởi Brendan Eich tại Hãng truyền thông Netscape với cái tên đầu tiên Mocha, rồi sau đó đổi tên thành LiveScript, và cuối cùng thành JavaScript. Giống Java, JavaScript có cú pháp tương tự C, nhưng nó gần với Self hơn Java. .js là phần mở rộng thường được dùng cho tập tin mã nguồn JavaScript [2]. Nếu như HTML được ví như khung xương của con người, CSS được ví như da thịt thì JavaScript được ví như các cơ bắp, thần kinh để giúp con người có thể hoạt động, thao tác,... Tương tự như vậy trên một website, HTML có tác dụng khai báo các thành phần được sử dụng, CSS sử dụng để trang trí các thành phần trên trang Web, còn JavaScript có tác dụng xử lý các tương tác của con người lên Website như cuộn chuột, nhấn chuột,... Ngoài ra nó còn có các chức năng như xử lý, hiển thị dữ liệu và nhiều chức năng khác.

2.1.3. Framework Django

Django là một mã nguồn mở miễn phí dựa trên Python Framework, theo kiến trúc mô hình (model) - mẫu (template) - khung nhìn (view) (MTV). Nó được duy trì bởi Django Software Foundation (DSF), một tổ chức độc lập được thành lập dưới dạng phi lợi nhuận [3]. Web framework là một bộ các thành phần giúp các nhà phát triển phần mềm phát triển website nhanh hơn và dễ dàng hơn [4]. Khi xây dựng một website cần các chức năng như đăng nhập, đăng ký, đăng xuất, bảng điều khiển website,... May mắn là django đã làm mọi thứ, ta chỉ việc sử dụng lại hay phát triển dựa trên các thành phần đã được xây dựng trước đó. Từ

đó nhà phát triển không phải xây dựng lại từ những thứ cơ bản như biểu mẫu (form), lớp truy xuất cơ sở dữ liệu (Data Access Object), trang quản trị (admin),... Giúp rút ngắn thời gian, công sức để phát triển một website.

2.2. Mã nguồn mở

“Open-source software is software with source code that anyone can inspect, modify, and enhance” [5]. Mã nguồn mở (Open Source) là khái niệm để chỉ những phần mềm, ứng dụng có mã nguồn công khai (mở). Cho phép tất cả mọi người xem, thay đổi, chỉnh sửa, hay sử dụng mã nguồn này để phát triển các phần mềm khác [6]. Các phần mềm mã nguồn mở thường là những phần mềm hoàn chỉnh, được trang bị các tính năng cơ bản. Người dùng có thể dựa trên phần mềm này để viết thêm giao diện, tính năng và tích hợp vào để tạo được phần mềm tùy biến theo ý thích. Việc sử dụng mã nguồn mở giúp tạo website nhanh hơn vì mã nguồn mở thường đã là một hệ thống hoạt động hoàn chỉnh. Ngoài ra nó cũng làm giảm chi phí phát triển bởi mã nguồn mở là hoàn toàn miễn phí, làm giảm thời gian, công sức phát triển phần mềm. Một mã nguồn mở cũng được cho là bảo mật và an ninh hơn so với các website code tay thông thường. Lý do là vì có nhiều người đọc được mã nguồn, do đó sẽ sớm phát hiện các lỗ hổng, và tung ra các bản vá tức thì, Cho nên các website mã nguồn mở được cập nhật bảo mật thường xuyên.

Để tạo nên một ứng dụng hoặc website phải mất nhiều thời gian, công sức và nhiều tài nguyên. Giá trị tạo ra sản phẩm lại không đảm bảo được tính ổn định, dễ gặp nhiều lỗi và khó nâng cấp cho sau này. Nhận thấy được những lợi ích của việc sử dụng mã nguồn mở, chúng tôi đã nghiên cứu và lựa chọn mã nguồn mở DMOJ làm cơ sở để phát triển hệ thống sau này. DMOJ là một nền tảng chấm code tự động trực tuyến (Online Judge) hiện đại được viết trên framework Django. Nếu xét về mức độ hoàn thiện và hỗ trợ phát triển thì DMOJ xứng đáng là số 1. Mã nguồn này hiện vẫn được phát triển và cập nhật đều đặn. DMOJ có các tính năng cơ bản như bài tập, cuộc thi, hệ thống xếp hạng, hỗ trợ chấm nhiều ngôn ngữ khác nhau, hệ thống chạy trên môi trường sandbox,...

2.3. Cloud computing

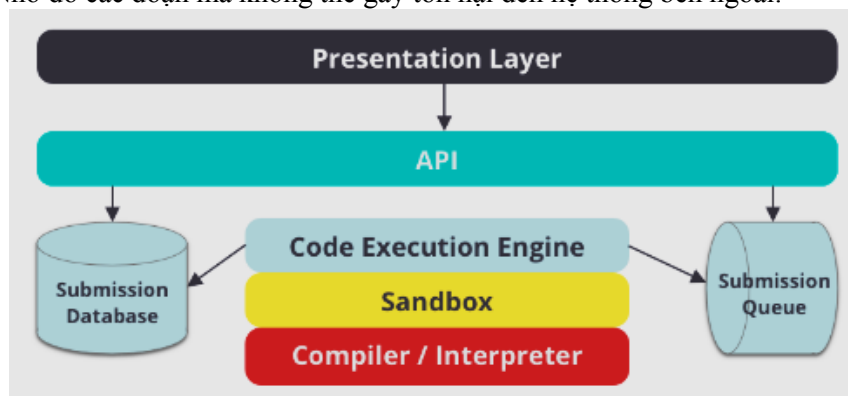
Trước khi khái niệm Cloud Computing xuất hiện, cơ sở hạ tầng ở thời điểm đó khi cần một dịch vụ gì người dùng cũng đều phải tự làm và tự đầu tư hay đối với các hệ thống nhỏ thì phải thuê các hosting bởi các nhà cung cấp trên internet. Lấy ví dụ người dùng muốn tự lập một website bán hàng thì người dùng phải mua một web server riêng cho website đó. Hoặc người dùng muốn lưu trữ hay sao lưu thì phải tự sắm cho mình một vài chiếc ổ cứng,... Chúng ta có thể dễ dàng nhận thấy được mức chi phí đầu tư ban đầu là khá lớn, chưa kể đến việc nó còn rườm rà và phức tạp. Nó chỉ phù hợp với những người có kiến thức chuyên sâu về CNTT. Và còn cả trường hợp người dùng không sử dụng hết tài nguyên đó gây ra thất thoát và lãng phí. Lúc này Cloud Computing ra đời để giải quyết vấn đề này. Điện toán đám mây (Cloud Computing) là việc ảo hóa các tài nguyên tính toán và các ứng dụng, thay vì việc sử dụng một hoặc nhiều máy chủ thật. Hiểu một cách đơn giản Cloud Computing là việc cung cấp dịch vụ điện toán hoàn toàn trên Internet. Hay đúng hơn là việc cung cấp tài nguyên phù hợp nhu cầu người dùng hoàn toàn qua Internet. Các dịch vụ ở đây có thể bao gồm máy chủ, lưu trữ hay phần mềm,...

Hiện nay đã có rất nhiều nhà cung cấp các dịch vụ đám mây như Google Cloud, Microsoft Azure, Amazon Web Services (AWS). AWS là nền tảng đám mây toàn diện và được sử dụng rộng rãi nhất, cung cấp trên 200 dịch vụ đầy đủ tính năng từ các trung tâm dữ liệu trên toàn thế giới. AWS sẽ cung cấp cho người dùng một nền tảng cloud với cơ sở hạ tầng đáng tin cậy có thể mở rộng và chi phí thấp, có thể phục vụ mọi nhu cầu của nhiều cá nhân, doanh nghiệp mà không hề giới hạn. AWS thường bao gồm nhiều nhóm dịch vụ đa dạng khác nhau như điện toán (compute), lưu trữ (storage), cơ sở dữ liệu (database), mạng lưới (network) [7]. Để áp dụng Cloud Computing vào hệ thống chúng tôi đã tạo một máy chủ ảo (VPS) trên dịch vụ AWS EC2. Việc sử dụng AWS EC2 giúp chúng tôi giảm thiểu việc phải đầu tư vào phần cứng cũng như thiết lập hệ thống, mạng,... Bên cạnh đó dễ dàng mở rộng, thu hẹp (scale) hệ thống tùy thuộc vào nhu cầu.

2.4. Sandbox

Sandbox là một kỹ thuật quan trọng trong lĩnh vực bảo mật có tác dụng cô lập các ứng dụng, ngăn chặn các phần mềm độc hại để chúng không thể làm hỏng hệ thống máy tính, hay cài các mã độc vào hệ thống [8]. Về cơ bản, Sandbox là một môi trường dùng để chạy phần mềm và môi trường đó được nằm trong sự kiểm soát chặt chẽ. Sandbox giúp hạn chế chức năng của một đoạn mã, cấp quyền cho một đoạn mã nào đó chỉ

được thực hiện một số chức năng nhất định, từ đó không thể thực hiện những can thiệp khác có thể làm nguy hại cho máy tính người dùng. Ví dụ như các đoạn mã nguy hiểm như xóa các tệp hệ thống, tắt máy,... Các hệ thống chấm code tự động trực tuyến (Online Judge) thường gặp phải các mối đe dọa bảo mật do gửi mã độc hại, điều này có thể gây ra lỗi máy chủ, sự cố, mất mát dữ liệu,... Vì thế IUHCoder sử dụng mô hình Sandbox như một cơ chế bảo mật. Khi thiết kế hệ thống Online Judge, để đảm bảo hoạt động an toàn của các đoạn mã, người dùng phải được phép gửi mã chỉ có thể truy cập các tài nguyên hạn chế do Sandbox cung cấp. Do vậy, mô hình Sandbox được xây dựng, tránh việc người dùng gửi mã độc làm hỏng hệ thống. Sandbox security được tách biệt với phía web, tương tác thông qua một giao thức tương tác và các thủ tục tiêu chuẩn. Nhiệm vụ chính của Sandbox là đánh giá yêu cầu từ phía web, biên dịch và chạy các chương trình do người dùng gửi, sau đó xuất kết quả chạy theo tệp input đầu vào tiêu chuẩn đặt trước, sau đó đóng gói và gửi trở lại phía website. IUHCoder sử dụng Sandbox bởi nó giúp hệ thống có khả năng mở rộng tốt, thuận tiện triển khai Sandbox security trên một máy chủ khác để cải thiện hiệu suất hệ thống. Việc mở rộng chỉ yêu cầu sửa đổi một lượng nhỏ mã để cải thiện tính ổn định của mạng. Thực thi song song và đồng thời các tính năng mang lại hiệu suất cao, độ tin cậy và ổn định cao. Khi biên dịch mã mà người dùng gửi tới, hệ thống sẽ biên dịch thông qua một môi trường sandbox như hình bên dưới. Nhờ đó các đoạn mã không thể gây tổn hại đến hệ thống bên ngoài.



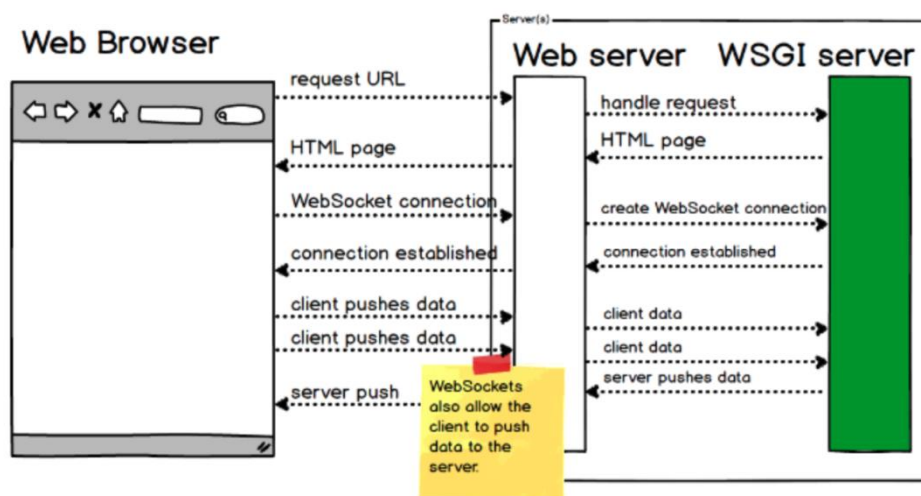
Hình 1. Mô hình sandbox trong hệ thống Online Judge

2.5. WebSocket

WebSocket là một giao thức được sử dụng rộng rãi cho việc phát triển ứng dụng thời gian thực (realtime). WebSocket cho phép tạo kết nối hai chiều giữa client và server. Nó cung cấp một cách thức bền bỉ, độ trễ thấp để hỗ trợ giao tiếp giữa client và server. Sử dụng WebSocket, người dùng có thể tạo ra một ứng dụng real-time đúng nghĩa như nhắn tin, soạn văn bản online cùng nhau, giao dịch chứng khoán hay game online nhiều người chơi cùng lúc [9], [10]. WebSocket API: Cho phép ứng dụng của người dùng kiểm soát giao thức WebSocket và phản hồi lại với những sự kiện được thông báo bởi server. Nếu server có dữ liệu gửi về client hoặc khi tài nguyên mà ứng dụng quan tâm thay đổi trạng thái, nó lập tức gửi thông báo về phía client. Một khi được tạo lập các message WebSocket có thể được truyền đến và đi thông qua các method của WebSocket. Bản chất đồng bộ của WebSocket có nghĩa là một khi kết nối WebSocket được mở, ứng dụng sẽ lắng nghe những sự kiện. Để bắt đầu lắng nghe những sự kiện, người dùng có thể thêm hàm gọi lại (callback) vào đối tượng WebSocket hoặc sử dụng phương thức để thêm các sự kiện lắng nghe.

Bởi vì WebSocket cung cấp giao thức 2 chiều mạnh mẽ nên có độ trễ rất thấp và dễ sửa chữa. Thông tin trả về từ WebSocket là vô cùng nhanh chóng nên được sử dụng trong nhiều trường hợp cần thời gian thực như chat, hiển thị biểu đồ,... Trong những trường hợp này, ta không thể dùng HTTP để kết nối vì nếu gửi lệnh AJAX liên tiếp tới server để lấy dữ liệu mới và cập nhật chúng lên màn hình, như thế sẽ tốn nhiều tài nguyên, traffic và thời gian trả lại không chính xác. Khi sử dụng giao thức TCP để kết nối bằng WebSocket, hệ thống không cần quá nhiều kết nối như phương pháp Comet streaming. API của WebSocket cũng khá dễ sử dụng trực tiếp. Khả năng tương thích giữa các nền tảng cao (web, máy tính để bàn, thiết bị di động). WebSocket giúp giảm kích thước của HTTP header từ 500-1000 lần. Giao thức WebSocket Secure sử dụng mã hóa SSL và TLS tiêu chuẩn để thiết lập kết nối an toàn giữa máy khách và máy chủ. Khách hàng và máy chủ có thể trao đổi một cách an toàn các tin nhắn WebSocket được mã hóa.

Chúng tôi đã áp dụng WebSocket vào hệ thống chấm bài thi, cập nhật trực tiếp trạng thái đúng sai của từng testcase, hiển thị kết quả cuối cùng và các lỗi biên dịch hay cảnh báo theo thời gian thực (realtime).



Hình 2. Cấu trúc cơ bản của một hệ thống sử dụng WebSocket

3. IUHCODER: NỀN TẢNG HỌC TẬP CHẤM CODE TỰ ĐỘNG

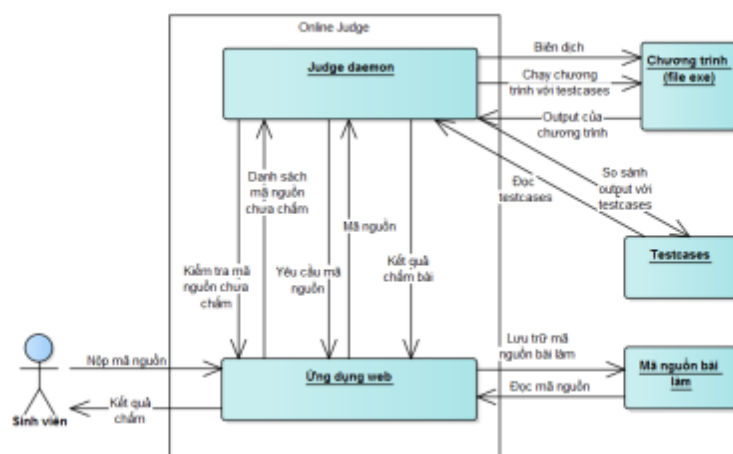
3.1. Hệ thống Online Judge

Online Judge là hệ thống chấm bài trực tuyến để kiểm tra các chương trình trong các cuộc thi về lập trình [11]. Học viên chỉ cần nộp file code lên và hệ thống sẽ chấm bài một cách tự động thông qua các đầu vào (input) và so khớp kết quả (output) của chương trình đó với kết quả đúng của đề bài (expect output). Sau đó kết quả sẽ được thông báo tới học viên. Thông thường một hệ thống Online Judge bao gồm các chức năng chính như:

- Bài tập: Tại đây học viên có thể làm bài tập để tra dồi vốn kiến thức, kỹ năng lập trình của mình.
- Cuộc thi: Cho phép học viên tham gia các kỳ thi nhằm tra dồi kiến thức, tranh đua thứ hạng với các học viên khác.
- Hệ thống xếp hạng: Đây cũng là chức năng tiêu biểu của một hệ thống Online Judge, điều này làm tăng sự thú vị khi lập trình. Mỗi học viên sẽ có một chỉ số xếp hạng, chỉ số này sẽ được thay đổi sau khi tham gia các cuộc thi xếp hạng.

Biểu đồ ngữ cảnh hệ thống, quy trình cơ bản để chấm một bài code như sau:

- Bước 1: Người dùng chọn bài tập sau đó nộp file code hoặc dán code vào form và bấm nộp.
- Bước 2: Ứng dụng web gửi mã nguồn tới hệ thống chấm code (judge).
- Bước 3: Hệ thống chấm code sẽ biên dịch chương trình và chạy với từng input trong bộ test case.
- Bước 4: Hệ thống chấm code thực hiện so sánh kết quả đã được xuất ra ở bước 3 với kết quả đúng (expect output) trong bộ test case.
- Bước 5: Hệ thống chấm điểm dựa trên số test case đúng.
- Bước 6: Trả kết quả về ứng dụng web và hiển thị cho người dùng.



Hình 3. Biểu đồ ngữ cảnh cơ bản của một hệ thống online judge

Khi áp dụng hệ thống online judge, nó sẽ giúp học viên dễ dàng học lập trình với các bài toán mà hệ thống đưa ra, giải quyết vấn đề và nhận được kết quả ngay lập tức. Đối với giảng viên, điều này sẽ làm giảm đi thời gian chấm từng bài code của học viên (công việc tốn nhiều công sức). Thay vào đó hệ thống sẽ tự động chấm tất cả các bài nộp và trả về kết quả cuối cùng dựa trên bộ test case.

3.2. Hệ thống Learning Management System (LMS)

Hiện nay, có rất nhiều hệ thống quản lý lớp học online cho học sinh - sinh viên đặc biệt là sinh viên CNTT, một trong số đó có thể kể đến là LMS, CMS. Hiện tại, trường Đại học Công Nghiệp Thành phố Hồ Chí Minh cũng đang sử dụng hệ thống quản lý lớp học LMS.

LMS là một phần mềm quản lý liên quan đến việc học tập, bao gồm lập tài liệu, theo dõi, báo cáo, tự động hóa và phân phối các khóa học, chương trình đào tạo,... [12],[13],[14] Áp dụng hệ thống LMS vào IUHCoder nhằm giúp theo dõi chi tiết tất cả hoạt động của học sinh - sinh viên, qua đó sẽ thống kê và xuất báo cáo tự động hoàn toàn cho giảng viên. Từ đó giảng viên ghi nhận được toàn bộ tiến độ cải thiện kỹ năng, học tập của từng học sinh - sinh viên của mình. Toàn bộ giáo trình cũng như là các bài kiểm tra, khảo sát đều sẽ được lưu trữ trực tuyến, trên hệ thống quản lý lớp học này. Sinh viên sẽ có thể nhận thông báo từ lớp học bất kỳ lúc nào và ở đâu, trên các lớp học LMS sẽ có các diễn đàn để trao đổi các chủ đề học tập những điều này mang lại một tâm trạng thoải mái khi tiếp thu việc học và thoải mái về mặt thời gian. Mọi hoạt động của việc học tập đều được lưu trữ lại trên hệ thống; việc cập nhật, truy cập nội dung cũng chỉ trong vài cú click chuột, giúp cho sinh viên có thể thao tác dễ dàng trên hệ thống.

Kết hợp hệ thống LMS vào hệ thống Online Judge sẽ làm hệ thống trở nên mạnh mẽ, đa năng hơn đối với giáo viên cũng như học viên. Giáo viên có thể quản lý lớp học, các tài nguyên của lớp học như bài tập, tài liệu, giáo trình, các câu hỏi lý thuyết ôn tập,... Tạo các kỳ thi, xuất kết quả các cuộc thi một cách tự động, điểm số cũng được đánh giá tự động dựa trên hệ thống Online Judge.

3.3. Phân tích hệ thống IUHCoder

Hệ thống IUHCoder là sự kết hợp giữa hệ thống Online Judge và hệ thống LMS, các actor chính được xác định bao gồm: học viên, giảng viên, quản trị tổ chức, quản trị hệ thống. Các chức năng của từng actor được liệt kê trong bảng 1 dưới đây.

Bảng 1. Danh sách các chức năng trong hệ thống IUHCoder

Actor	Tên chức năng	Mô tả
Học viên	Xem thông tin cá nhân	Cho phép học viên xem thông tin cá nhân của người khác hoặc của bản thân.

Đăng nhập	Cho phép học viên đăng nhập bằng tài khoản đã đăng ký
Đăng ký	Cho phép học viên đăng ký tài khoản
Quên mật khẩu	Cho phép học viên quên mật khẩu
Đổi mật khẩu	Cho phép học viên đổi mật khẩu của tài khoản đang đăng nhập
Sửa thông tin cá nhân	Cho phép học viên sửa thông tin cá nhân của mình
Xem các bài tập	Cho phép học viên xem danh sách bài tập trong hệ thống IUHCoder
Giải các bài tập	Cho phép học viên làm các bài tập. Bao gồm việc đọc đề, nộp bài code, xem các bài nộp của bản thân, báo cáo vấn đề của bài tập.
Tìm kiếm, lọc bài tập	Cho phép học viên tìm kiếm, lọc bài tập
Bình luận	Cho phép học viên bình luận vào các bài tập, cuộc thi, tài liệu
Xem cuộc thi	Cho phép học viên xem danh sách các cuộc thi đã qua, đang diễn ra, cuộc thi sắp tới
Tham gia các cuộc thi	Cho phép học viên tham gia các cuộc thi. Bao gồm việc đăng ký tham gia, làm các đề bài trong cuộc thi, tham gia ảo cuộc thi, bình luận cuộc thi.
Xem thống kê cuộc thi	Cho phép học viên xem các thống kê của cuộc thi
Xem bảng xếp hạng cuộc thi	Cho phép học viên xem bảng xếp hạng của cuộc thi
Xem các bài nộp	Cho phép học viên xem các bài nộp trên hệ thống
Xem bảng xếp hạng người dùng	Cho phép học viên xem bảng xếp hạng người dùng
Xem danh sách tổ chức	Cho phép học viên xem danh sách các tổ chức có trong hệ thống
Đọc tài liệu	Cho phép học viên đọc tài liệu, làm các câu hỏi ôn tập ở mỗi tài liệu, bình luận vào tài liệu
Xem bài viết trên diễn đàn	Cho phép học viên xem các bài viết trên diễn đàn
Tham gia diễn đàn	Cho phép học viên tham gia diễn đàn, thêm bài viết vào diễn đàn, đánh giá bài viết, bình luận

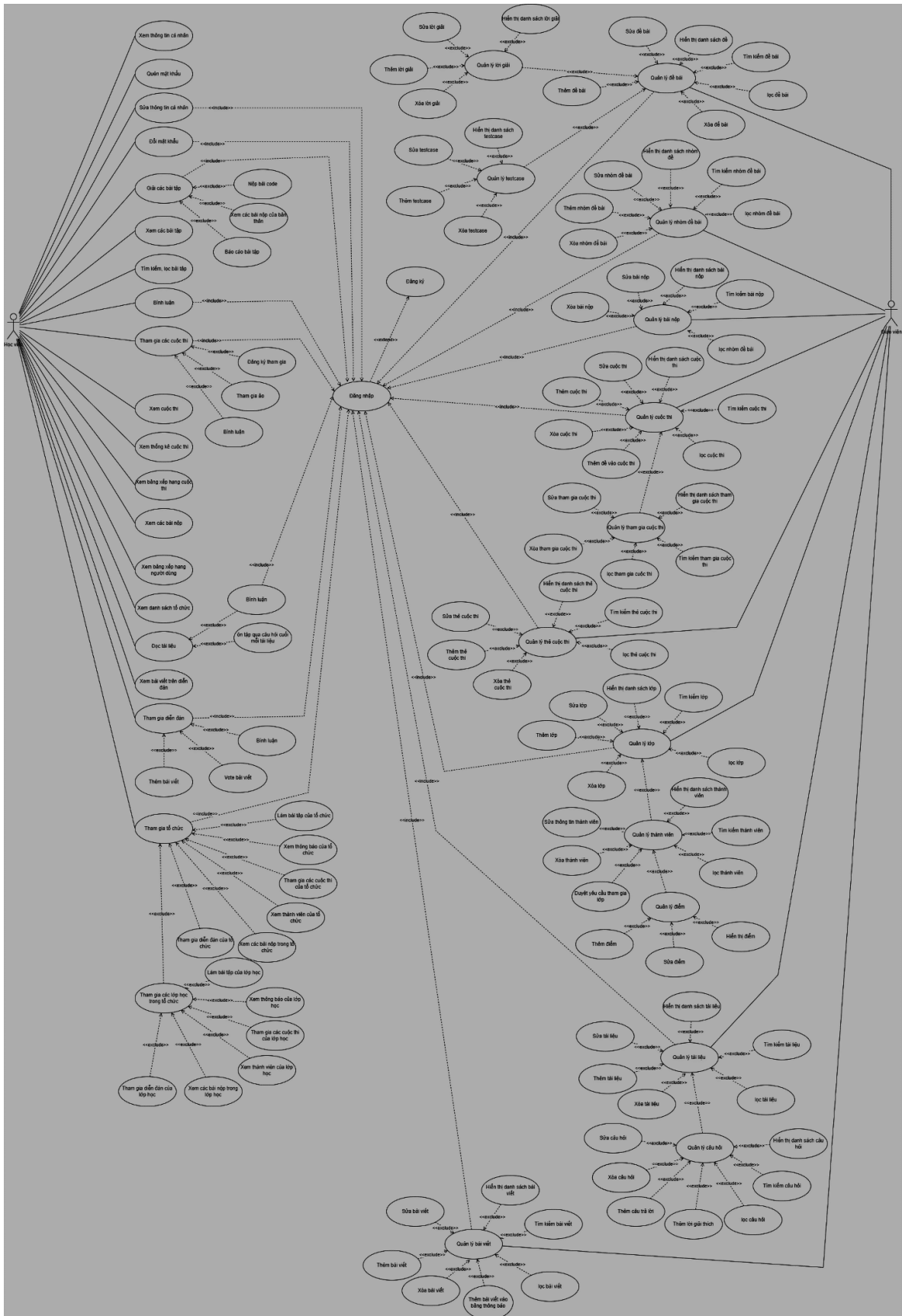
	Tham gia tổ chức	Cho phép học viên tham gia các tổ chức, làm bài tập của tổ chức, xem thông báo, tham gia các cuộc thi, xem danh sách thành viên, xem các bài nộp, tham gia diễn đàn, tham gia các lớp học của tổ chức.
	Tham gia các lớp học trong tổ chức	Cho phép học viên tham gia lớp học trong tổ chức, làm bài tập, xem thông báo, tham gia các cuộc thi, xem danh sách thành viên, xem các bài nộp, tham gia diễn đàn của lớp học.
Giáo viên	Quản lý đề bài	Cho phép giáo viên quản lý đề bài, hiển thị, thêm, sửa, tìm kiếm, lọc, xóa đề bài, thêm các lời giải cho đề bài, thêm các testcase cho đề bài.
	Quản lý lời giải	Cho phép giáo viên quản lý lời giải, hiển thị, thêm, sửa, xóa lời giải.
	Quản lý testcase	Cho phép giáo viên quản lý testcase của đề bài, hiển thị, thêm, sửa, xóa testcase.
	Quản lý nhóm đề bài	Cho phép giáo viên quản lý nhóm đề bài, hiển thị, thêm, sửa, tìm kiếm, lọc nhóm đề bài, thêm các đề bài vào nhóm đề bài.
	Quản lý bài nộp	Cho phép giáo viên quản lý bài nộp, hiển thị, sửa, tìm kiếm, xóa bài nộp.
	Quản lý cuộc thi	Cho phép giáo viên quản lý cuộc thi, hiển thị, thêm, sửa, tìm kiếm, lọc, xóa cuộc thi, thêm đề bài vào cuộc thi, thêm thẻ cuộc thi.
	Quản lý tham gia cuộc thi	Cho phép giáo viên quản lý tham gia cuộc thi, hiển thị, sửa, xóa, tìm kiếm, lọc tham gia cuộc thi
	Quản lý thẻ cuộc thi	Cho phép giáo viên quản lý thẻ cuộc thi, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc thẻ cuộc thi.
	Quản lý lớp	Cho phép giáo viên quản lý lớp, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc lớp, quản lý thành viên của lớp.
	Quản lý thành viên	Cho phép giáo viên quản lý thành viên, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc thành viên, duyệt yêu cầu tham gia, quản lý điểm của thành viên.
	Quản lý điểm	Cho phép giáo viên quản lý điểm, hiển thị, thêm, sửa, xóa điểm.
	Quản lý tài liệu	Cho phép giáo viên quản lý tài liệu, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc tài liệu, quản lý câu hỏi của tài liệu.

	Quản lý câu hỏi	Cho phép giáo viên quản lý câu hỏi, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc câu hỏi, thêm các đáp án vào câu hỏi.
	Quản lý bài viết diễn đàn	Cho phép giáo viên quản lý diễn đàn, hiển thị, thêm, sửa, xóa, tìm kiếm, lọc bài viết, thêm bài viết vào bảng thông báo.
Quản trị tổ chức	Quản lý tổ chức	Cho phép quản trị tổ chức quản lý tổ chức, sửa các thông tin liên quan của tổ chức.
	Quản lý phân quyền	Cho phép quản trị tổ chức tạo nhóm quyền, phân quyền cho các thành viên trong tổ chức.
Quản trị hệ thống	Quản lý tổ chức	Cho phép người quản trị quản lý các tổ chức, hiển thị, thêm, xóa, sửa, tìm kiếm, lọc các tổ chức, chỉ định người quản trị tổ chức.
	Quản lý hệ thống	Cho phép người quản trị quản lý hệ thống, thêm các thanh điều hướng, cấu hình hệ thống, chuyên hướng,...
	Quản lý phân quyền	Cho phép người quản trị tạo các nhóm quyền, phân quyền cho các người dùng trong hệ thống.
	Quản lý máy chấm	Cho phép người quản trị quản lý các máy chấm, hiển thị, thêm, sửa, xóa máy chấm.

4. THỰC NGHIỆM THỰC TẾ

Chúng tôi đã tiến hành triển khai hệ thống IUHCoder trên thực tế. Hiện tại hệ thống đã có hơn 170 người dùng, 11 cuộc thi đã được tổ chức. Chúng tôi đã tiến hành thực nghiệm dựa trên lớp DHKHDL17A, với hơn 60 thành viên. Chúng tôi đã tổ chức 2 cuộc thi dành cho các bạn sinh viên lớp DHKHDL17A nhằm đánh giá kỹ hơn về hệ thống. Một cuộc thi đã được tổ chức nhằm đánh giá năng lực lập trình theo thể thức OLP, tức sinh viên sẽ làm bài cá nhân, cuộc thi diễn ra trong 3 tiếng, có 55 sinh viên tham gia với hơn 1600 bài được nộp tới hệ thống. Một cuộc thi khác được tổ chức theo thể thức ICPC, tức sinh viên sẽ làm bài theo đội 4 người, gồm 17 đội tham gia với 320 bài được nộp tới hệ thống.

Qua thực nghiệm, hệ thống IUHCoder đang vận hành theo đúng kỳ vọng, không diễn ra tình trạng tắc nghẽn, quá tải, các bài nộp được chấm và trả về kết quả gần như ngay lập tức. Các chức năng của hệ thống đáp ứng được các yêu cầu mà chúng tôi đã đề ra. Cũng dựa trên thực nghiệm, chúng tôi đã ghi chép lại các bất tiện, khó khăn trong quá trình sử dụng dựa trên các báo cáo và góp ý của các thành viên lớp DHKHDL17A.



Hình 4. Sơ đồ Usecase của IUHCoder với Actor Học viên và Giáo viên

Xếp hạng	Tên truy cập	Điểm	1	2	3	4	5	6	7	8	9	10
1	Đào_Xuân_Hoàng_Tuấn	972	100	100	100	100	100	72	100	100	100	100
2	Nguyễn_Văn_Nam	960	100	100	100	100	100	60	100	100	100	100
3	Bảng_Hữu_Tri	910	100	100	100	100	70	60	100	80	100	100
4	Trương_Đức_Tê	863	100	0	100	100	100	63	100	100	100	100
5	Hoàng_Tiến_Anh	715	100	0	100	10	100	5	100	100	100	100
6	Hoàng_Ngọc_Tên	703	100	8	100	40	30	25	100	100	100	100
7	Cao_Phan_Nhánh_Duy	685	100	30	100	40	30	55	100	100	30	100
8	Lê_Thanh_Di	670	100		100	10	60		100	100	100	100
9	Phạm_Chị_Nguyên	660	100		100	20	60		100	80	100	100
10	Vũ_Hoàng_Phúc	650	100	0	100	100	100		100	50	60	40
11	Nguyễn_Nhã_Anh	640	100	80	100	100	10	60	100	50		40
12	Nguyễn_Dĩnh_Nguyên_Bắc	580	100	0	100	20	100		100	100		60
13	Dương_Văn_Tài_K17D5	560	100		100	30	0		100	100	100	30
13	Lê_Thị_Minh_Trang	560	100		100	100	30		100	100	30	
15	Tăng_Quang_Nhật_Nam	535	100		100	100	0	5	100	100		30
16	Nguyễn_Tiến_Gia_Huy	530	100		100	0	30		100	100		100
17	Nguyễn_Văn_Siền	500	100		100	100			100	100		

Hình 5. Bảng xếp hạng cuộc thi Đánh giá năng lực lập trình - DHKHD17A

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Chúng tôi đã cơ bản giải quyết được 2 bài toán trong việc giảng dạy lập trình. Đầu tiên là hỗ trợ giáo viên trong quá trình chấm code, giúp học viên có nguồn tài nguyên học tập, chấm code và đưa ra kết quả ngay lập tức. Thứ hai là giúp giáo viên có thể quản lý lớp học, quản lý các thành viên, các bài thi, bài tập và tài liệu trong lớp học. Chúng tôi đã tiến hành thực nghiệm trên thực tế để đánh giá các chức năng, hệ thống đầy đủ và chi tiết hơn. Chúng tôi đã áp dụng được các công nghệ vào hệ thống IUHCoder như mã nguồn mở, cloud computing, sandbox, websocket,... Tuy nhiên hệ thống IUHCoder vẫn còn những hạn chế về mặt chức năng, nhiều chức năng còn chưa được hoàn thiện, chưa hợp logic và các lỗi chưa được sửa. Về mặt hệ thống, để tiết kiệm chi phí chúng tôi chỉ sử dụng EC2 để triển khai môi trường phát triển. Với môi trường sản xuất chúng tôi đã tận dụng lợi thế của mình để triển khai hệ thống trên một máy chủ được đặt tại câu lạc bộ H3.2. Vì một vài khó khăn về mạng lưới internet, chúng tôi đã sử dụng Cloudflared Tunnel để có thể dễ dàng đưa hệ thống đến với người dùng. Tuy nhiên việc này vẫn gặp nhiều bất cập như máy chủ bị tắt do mất điện, lỗi mạng,... Trong tương lai hệ thống sẽ hoàn toàn được triển khai trên Cloud Computing, các chức năng sẽ được hoàn thiện hơn, giải quyết được các vấn đề mà nhóm đang gặp phải. Bước đầu chúng tôi sẽ triển khai các chức năng mới, xây dựng các bộ giáo án, tài liệu ở mức độ cơ bản dành cho sinh viên khóa K18 Trường Đại học Công nghiệp Thành phố Hồ Chí Minh ở các bộ môn như nhập môn lập trình, kỹ thuật lập trình, cấu trúc dữ liệu và giải thuật,... Đồng thời triển khai dạy các kiến thức liên quan đến cuộc thi DrCT, quảng bá hệ thống với các trường Trung học cơ sở, Trung học phổ thông và các trường Đại học, Cao đẳng khác.

Khi còn là ý tưởng, IUHCoder đã đặt ra hướng đi và sứ mệnh cho dự án là thay đổi cách dạy và học của sinh viên - giảng viên CNTT trường Đại học Công Nghiệp Thành phố Hồ Chí Minh nói riêng và các trường Đại học khác trên cả nước nói chung. Bắt đầu từ học kì I 2022, IUHCoder sẽ hoàn thiện và đưa vào giảng dạy các giáo trình, tài liệu mà đội ngũ đã hoàn thành cho các môn học như: nhập môn lập trình, kỹ thuật lập trình, cấu trúc dữ liệu và giải thuật,... Chưa dừng lại ở đó, IUHCoder mong muốn sẽ phát triển mô hình dạy và học này không chỉ ở Đại học Công Nghiệp Thành phố Hồ Chí Minh mà chúng tôi mong muốn mô hình giảng dạy này sẽ nhân rộng ra cho nhiều trường khác không chỉ là Đại học. Bên cạnh việc phát triển giảng dạy các môn học ở trường, đội ngũ chúng tôi đã và đang xây dựng một môi trường học thuật uy tín và chất lượng trên nền tảng IUHCoder, và trong tương lai chúng tôi sẽ phát triển nhiều hơn nữa các cuộc thi này với quy mô và tầm cỡ lớn hơn tạo nên một nền thương hiệu học thuật IUHCoder. Với các cuộc thi khá mới mẻ ở Việt Nam như DrCT, IUHCoder mong muốn sẽ là nền tảng thúc đẩy các cuộc thi này phát triển và

lớn mạnh với các bộ giáo án - đề thi thử mà chúng tôi sẽ chuẩn bị, bên cạnh việc tổ chức các lớp dạy và học ngay trên nền tảng này.

TÀI LIỆU THAM KHẢO

- [1] Wikipedia, Python (ngôn ngữ lập trình). [Online]. Available: [https://vi.wikipedia.org/wiki/Python_\(ngôn_ngữ_lập_trình\)](https://vi.wikipedia.org/wiki/Python_(ngôn_ngữ_lập_trình)).
- [2] Wikipedia, JavaScript. [Online]. Available: <https://vi.wikipedia.org/wiki/JavaScript>.
- [3] Wikipedia, Django (web framework). [Online]. Available: [https://en.wikipedia.org/wiki/Django_\(web_framework\)](https://en.wikipedia.org/wiki/Django_(web_framework)).
- [4] Website tutorial.djangogirls.org, What is Django? [Online]. Available: <https://tutorial.djangogirls.org/en/django/>.
- [5] Website opensource.com, What is open source? [Online]. Available: <https://opensource.com/resources/what-open-source>.
- [6] Website vpsmaychua.com, Open Source là gì? Những kiến thức cơ bản về Mã nguồn mở (2021). [Online]. Available: <https://vpsmaychua.com/open-source-la-gi-nhung-kien-thuc-co-ban-ve-ma-nguon-mo-2610>.
- [7] Website aws.amazon.com, Điện toán đám mây với AWS. [Online]. Available: <https://aws.amazon.com/vi/what-is-aws/>.
- [8] Tìm hiểu Sandbox là gì? Một kỹ thuật bảo vệ sự an toàn cho máy tính thiết bị. [Online]. Available: <https://cuongquach.com/tim-hieu-sanbox-la-gi.html>.
- [9] Introduction to WebSockets. [Online]. Available: <https://www.linode.com/docs/guides/introduction-to-websockets/#websocket-events>.
- [10] WebSockets. [Online]. Available: <http://stevlacey.github.io/websocket-examples/>.
- [11] Online judge. [Online], Available: https://code.fandom.com/wiki/Online_judge.
- [12] Wikipedia, Learning management system. [Online]. Available: https://en.wikipedia.org/wiki/Learning_management_system.
- [13] Hệ thống quản lý học tập (LMS) là gì? [Online]. Available: <https://thinkingschool.vn/quan-ly-dao-tao/he-thong-quan-ly-hoc-tap-lms-la-gi/>.
- [14] lms là gì? Vai trò và chức năng của phần mềm Lms. [Online]. Available: <https://jpweb.vn/lms-la-gi-vai-tro-va-chuc-nang-cua-phan-mem-lms/>.

YSC4F.241

IUHEDU – CÔNG CỤ HỖ TRỢ SINH VIÊN HỌC TẬP TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

TRƯƠNG ĐỨC HOÀN, ĐOÀN THỊ THANH HỒNG, TÔN LONG PHƯỚC
Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
hoantruong6814@gmail.com, hongblack2000@gmail.com, tonlongphuoc@iuh.edu.vn

Tóm tắt. Trường Đại học Công nghiệp Thành phố Hồ Chí Minh (IUH) là một trong 10 trường đại học hàng đầu Việt Nam trong bảng xếp hạng QS năm 2021. Trường đào tạo sinh viên đa ngành nghề, đa bậc học. Trường có cơ sở hạ tầng đạt tiêu chuẩn và đội ngũ giảng viên có trình độ cao. Hiện có hơn 30.000 sinh viên đang theo học và nghiên cứu tại các khoa và viện nghiên cứu. Tuy nhiên, hệ thống phần mềm hỗ trợ sinh viên trong quá trình học tập có một số vấn đề. Thứ nhất, nhà trường hiện đang sử dụng hệ thống Web để quản lý và hỗ trợ sinh viên trong công việc học tập và nghiên cứu. Hệ thống này thường gặp sự cố tại một vài thời điểm nhất định trong năm học. Cụ thể như thời điểm đăng ký học phần đầu mỗi học kỳ, hệ thống thường bị quá tải gây tắc nghẽn hệ thống. Thứ hai, hệ thống phần mềm chưa được triển khai trên các thiết bị di động trong khi thiết bị này ngày càng được phổ biến với sinh viên. Để giải quyết các vấn đề này, chúng tôi xây dựng một hệ thống phần mềm có tên là IUHEdu. Hệ thống này có hai phân hệ là Web ứng dụng và Ứng dụng trên điện thoại di động. Hệ thống giúp cho công việc hỗ trợ sinh viên trong quá trình học và nghiên cứu tại trường thuận tiện hơn. Chúng tôi đề xuất sử dụng các dịch vụ của AWS nhằm giải quyết vấn đề cân bằng tải cho các giao tác tại một thời điểm tắc nghẽn. Chúng tôi đã sử dụng công nghệ như ngôn ngữ lập trình Java và Javascript, React Native, ReactJS, Spring Boot và Cloud Computing, Github Action để xây dựng Front-end và Back-end. Hiện chúng tôi đã triển khai hệ thống này cho sinh viên ngành Kỹ thuật phần mềm thuộc Khoa Công nghệ Thông tin. Trong tương lai, chúng tôi cũng sẽ phát triển công cụ này bằng cách tối ưu hóa các chức năng và gắn kết với các hệ thống khác đã có của nhà trường.

Từ khóa. ứng dụng web, ứng dụng trên di động, dịch vụ AWS, điện toán đám mây

IUHEdu- A TOOL SUPPORT FOR STUDENTS AT INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY

Abstract. Industrial University of Ho Chi Minh City (IUH) is one of the top 10 universities in Vietnam in the QS rankings in 2021. The university trains students in a wide range of disciplines and levels of study. The university has a standard infrastructure and highly qualified teaching staff. Currently, there are more than 30,000 students studying and doing research at faculties and research institutes. However, the software system that supports students in the learning process has some problems. Firstly, the IUH is currently using the Web system to manage and support students in learning and research. This system often crashes at certain times during the school year. Second, the software system has not been implemented on mobile devices while these devices are increasingly popular with students. To solve these problems, we build a software system called IUHEdu. It has two modules: Web application and Mobile application. The system makes the work of supporting students in the process of studying and researching at the university more convenient. We recommend using AWS services to solve the load balancing problem of congested transactions at a time. We used technology such as Java and JavaScript programming languages, React Native framework, ReactJS, Spring Boot and Cloud Computing, Github Action to build Front-end and Back-end. Currently, we have implemented this system for students of Software Engineering in the Faculty of Information Technology. In the future, we will also develop this tool by optimizing the way it functions and integrates with other existing systems of the university.

Keywords. web Application, mobile application, AWS, cloud Computing

1. GIỚI THIỆU TỔNG QUAN

Giáo dục – đào tạo giữ vai trò hết sức quan trọng đối với sự tồn tại và phát triển của mỗi quốc gia cũng như toàn nhân loại. Trong bối cảnh hiện nay, sự phát triển mạnh mẽ của khoa học công nghệ, cùng với quá trình toàn cầu hóa và hội nhập quốc tế sâu rộng, đặc biệt là xu thế tất yếu của nền kinh tế tri thức đã tác động mạnh đến mọi lĩnh vực của đời sống xã hội, trong đó có giáo dục. Trong các trường Đại học, nhiệm vụ quan trọng hàng đầu của sinh viên chính là nhiệm vụ học tập, bằng các hoạt động học tập, người học tự hình thành và phát triển nhân cách của mình. Trường Đại học có trách nhiệm tạo mọi điều kiện thuận lợi nhất về việc quản lý sinh viên toàn trường, hỗ trợ, hướng dẫn, giúp đỡ, tạo các sân chơi giải trí... để người học có thể hoàn thành tốt nhất việc học của mình. Để hỗ trợ việc học tập của sinh viên và việc quản lý sinh viên toàn trường phải cần một chương trình chuyên nghiệp và hiện đại. Như chúng ta đã biết, ngày xưa, có rất nhiều trường Đại học lập ra chương trình hỗ trợ sinh viên học tập còn nhiều hạn chế, bất cập, hiệu quả còn thấp và chưa đầy đủ các chức năng dẫn đến sinh viên rất khó khăn trong việc học của sinh viên. Nhu cầu phát triển chất lượng cao đặt ra yêu cầu cho các trường đại học không ngừng đổi mới, cải cách lại chương trình hỗ trợ học tập cho sinh viên để nâng cao hơn nữa chất lượng và hiệu quả giáo dục. Muốn “cải tiến” lại một chương trình thì có rất nhiều yếu tố cần để có một chương trình hỗ trợ sinh viên tốt nhất. Điều quan trọng ở đây khi quản lý sinh viên phải cần lập ra danh sách sinh viên trong lớp và trong lớp học phần, quản lý việc đăng ký học phần, xếp lịch học/ lịch thi, lưu điểm số... chính vì điều đó mà phải cần đến sự tỉ mỉ, rõ ràng từng chi tiết. Nếu làm tất cả những việc đó trên các bộ hồ sơ, giấy tờ thì thông tin của sinh viên có thể không được đồng bộ, không chính xác. Vì thấu hiểu điều đó mà chúng tôi đã xây dựng một công cụ phần mềm để giúp đỡ trong việc quản lý việc học tập của sinh viên. Những giải pháp này có thể tránh được tình trạng mất mát các dữ liệu cũng như đồng bộ hóa các nghiệp vụ đào tạo. Trong thời điểm hiện tại, đa phần sinh viên đều sở hữu cho mình một chiếc điện thoại thông minh. Các thiết bị này có thể chạy được các ứng dụng trên nền tảng Internet một cách dễ dàng. Hiện tại, nhà trường cũng đã xây dựng hệ thống website có chế độ responsive tương thích với mọi kích thước màn hình hỗ trợ cho các thiết bị di động có thể sử dụng tốt. Tuy nhiên, những hạn chế của các ứng dụng web responsive có thể kể ra như tải trang chậm, xử lý đồng bộ dữ liệu không hiệu quả, khả năng chia sẻ tài nguyên còn hạn chế và chiếm tài nguyên băng thông đường truyền lớn do phải tải lượng lớn các phần không cần thiết các tài nguyên của trang Web, dẫn đến hiệu suất của phần mềm không được tốt. Từ đó, chúng tôi đề xuất một hướng giải quyết các vấn đề trên bằng cách xây dựng một ứng dụng hỗ trợ sinh viên thực hiện các nghiệp vụ khi học tập tại trường. Hiện việc triển khai các ứng dụng chạy trên các thiết bị di động cũng có nhiều công cụ hỗ trợ. Chúng tôi cũng quan tâm đến các vấn đề như giải quyết cân bằng tải của hệ thống, đồng thời phải đảm bảo được tính an toàn hiệu quả của hệ thống. Từ đó, chúng tôi chọn lựa các công nghệ cũng như các dịch vụ điện toán đám mây nhằm tổ chức xây dựng ứng dụng, lưu trữ dữ liệu cũng như cấu hình máy chủ một cách hiệu quả và an toàn. Hệ thống này sẽ giúp cho Nhà trường giải quyết những bất cập hiện nay như quá tải trong các kỳ đăng ký học phần của sinh viên, hay tra cứu thông tin học tập thông qua một hệ thống web tốn kém tài nguyên và rủi ro.

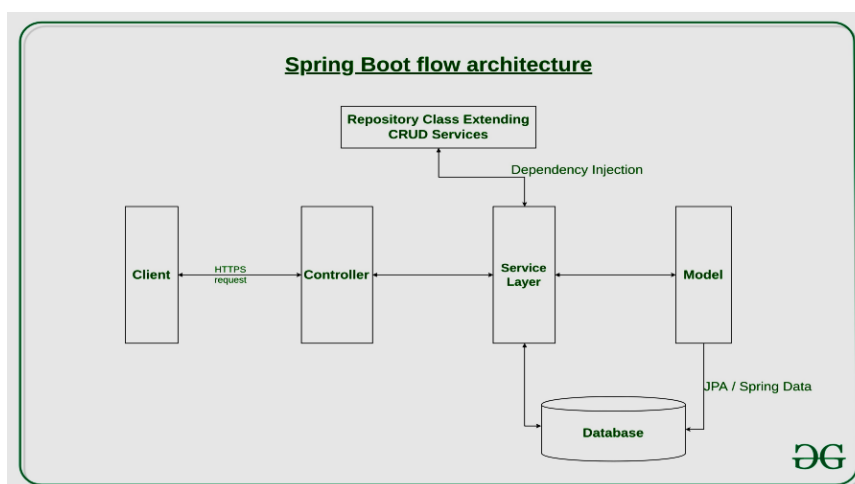
Bài báo của chúng tôi chia làm 5 phần: *Phần 1 Giới thiệu tổng quan về hệ thống; Phần 2 Các công nghệ liên quan; Phần 3 Công cụ IUHEdu; Phần 4 Kiểm nghiệm thực tế và Phần 5 Kết luận và hướng phát triển.* Trong đó *Phần 3* là phần đóng góp chính của bài báo.

2. CÁC CÔNG NGHỆ LIÊN QUAN

2.1 Framework Spring Boot

Spring Boot là khuôn mẫu (framework) phát triển dựa trên ngôn ngữ Java trong hệ sinh thái Spring framework [1, 2]. Nó giúp cho các lập trình viên đơn giản hóa quá trình lập trình một ứng dụng với Spring. Trong Spring Boot, người lập trình chỉ tập trung vào việc phát triển các nghiệp vụ cho ứng dụng mà không quan tâm nhiều về kiến trúc và mã nguồn của ứng dụng. Một kiến trúc webserver xây dựng dựa trên spring boot gồm các thành phần cơ bản như sau: Repository Class Extending, Service Layer, Model, Controller (*Hình 2.1.1*). Repository Class Extending là các thư viện hoặc framework được phát triển và duy trì bởi cộng đồng các lập trình viên Java để hỗ trợ quá trình xây dựng hệ thống một cách nhanh chóng hoặc cũng

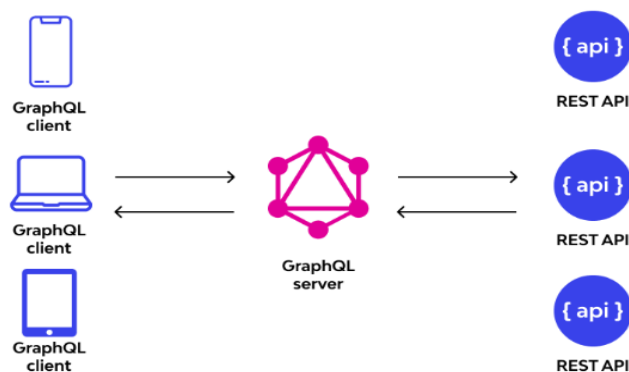
có thể là các lớp chứa các phương thức xử lý, truy xuất dữ liệu (CRUD) do lập trình viên viết ra; Service Layer là lớp dịch vụ chứa các phương thức (hàm) xử lý các yêu cầu như chỉnh sửa, lớp Service thường sử dụng các phương thức của các lớp trong Repository Class Extending (bằng cách gọi tới một hay kết hợp nhiều phương thức với nhau) để truy xuất, thêm, xóa dữ liệu trong cơ sở dữ liệu. Nó là nơi tập trung các xử lý nghiệp vụ của hệ thống Back-end; Model là thành phần được đính kèm theo các truy vấn của Services Layer xuống cơ sở dữ liệu. Model đóng nhiều vai trò quan trọng như phát sinh hệ cơ sở dữ liệu theo liên kết các lớp của lập trình viên, chuyển đổi kết quả truy xuất từ cơ sở dữ liệu thành các thực thể lớp; Controller là thành phần giao tiếp trực tiếp với phía client, tồn tại dưới dạng các API. Model có trách nhiệm xử lý yêu cầu của client, xác minh các yêu cầu đó. Model truy cập đến phương thức phù hợp trong Service Layer để đạt được kết quả mong muốn, trả kết quả về cho phía client.



Hình 2.1.1 Kiến trúc của luồng hoạt động Spring Boot[2]

2.2 GraphQL

GraphQL là một loại ngôn ngữ truy vấn và công nghệ áp dụng cho phía server. Nó được sử dụng rộng rãi cho các giao diện lập trình ứng dụng của API để đảm bảo rằng máy khách nhận được dữ liệu cần thiết. QL trong tên gọi là viết tắt của ngôn ngữ truy vấn (Query Languages). Được tạo bởi Facebook vào năm 2012 trong khi đang chế tạo ứng dụng FB, GraphQL hiện có rất nhiều mục đích và chức năng. GraphQL ra đời phục vụ cho việc API tìm kiếm dữ liệu mạnh mẽ; Giao diện API dễ học và làm việc; Một ngôn ngữ truy vấn mã nguồn mở làm cho các API nhanh chóng và linh hoạt; Một phương pháp hiệu quả để hợp lý hóa việc quản lý bảo mật cho API; Một cách để cung cấp cho các chuyên gia khả năng phát triển API với các phương pháp mong muốn; Một ngôn ngữ truy vấn hòa hợp với một môi trường phát triển tích hợp nhất định một cách dễ dàng và hiệu quả; Cơ sở để phát triển các API sẽ hoạt động theo cách mà khách hàng/ người dùng cuối mong muốn hoặc mong đợi; Một sự thay thế lý tưởng cho REST [3, 4].



2.2.2 Sơ đồ hoạt động của GraphQL [4]

Ưu điểm của GraphQL [3]

- GraphQL schema tự động tạo ra một nguồn đáng tin cậy trong ứng dụng GraphQL.
- Chỉ với một request, client có thể có được nhiều thông tin họ cần làm giảm thiểu số lượng requests.
- Hỗ trợ tối đa trong việc kiểm soát và xử lý data type từ đó hạn chế sự sai lệch trong giao tiếp giữa server và client.
- GraphQL giúp ứng dụng của bạn có thể phát triển thêm API mà không làm ảnh hưởng lên các truy vấn đã có.
- GraphQL không yêu cầu một kiến trúc ứng dụng cụ thể mà có thể hoạt động như một Rest API và có thể làm việc với các công cụ API hiện có.
- Tài liệu để học về GraphQL có sẵn và rất chi tiết, dễ dàng học, tiếp thu. Bạn có thể học tập tương tự nếu đã từng học về Rest API.

Nhược điểm của GraphQL [3]

- Rất nhiều extension mã nguồn mở của GraphQL không tương thích và không thể hoạt động với Rest API.
- Nhiều truy vấn sẽ bị GraphQL chuyển lên server khiến server sẽ chịu thêm nhiều công việc và làm cho server trở nên phức tạp hơn.
- Công việc để triển khai ứng dụng GraphQL và server sẽ có thể nhiều hơn việc phát triển Rest API.
- Bộ nhớ đệm phức tạp hơn Rest rất nhiều
- Người viết API sẽ có thể phải đảm nhiệm cả việc bảo trì cho GraphQL schema.
- Tùy thuộc theo cách triển khai, GraphQL sẽ có thể yêu cầu những chiến lược quản lý API khác nhau và cách thức khác nhiều so với Rest API.

2.3 React Native

Được phát triển bởi Facebook, React Native là một framework hướng đến phát triển ứng dụng di động đa nền tảng. Với sự trợ giúp của React Native, lập trình viên (developer) có thể sử dụng JavaScript để tạo ra mobile apps (ứng dụng di động) hỗ trợ cho cả nền tảng Android và iOS. Instagram, Facebook, Skype,... là những ứng dụng nổi bật sử dụng React Native. React Native hoạt động được nhờ tích hợp hai thread là Main Thread và JS Thread cho ứng dụng mobile. Trong đó: Main Thread: cập nhật giao diện người dùng (UI) và xử lý tương tác người dùng; JS Thread: thực thi và xử lý code Javascript. Hai luồng Main Thread và JS Thread này hoạt động độc lập với nhau. Hai Thread sẽ tương tác với nhau nhờ một Bridge (cầu nối). Chiếc cầu này sẽ chuyển đổi dữ liệu qua lại giữa các Thread [5]. Ngoài ra, React nói chung và React Native nói riêng là công nghệ được yêu thích hiện nay. Phong cách xây dựng mã nguồn trong React Native là phân chia các khối chức năng giúp dễ dàng trong việc quản lý, bảo trì, sửa lỗi, nâng cấp ứng dụng. Tương tự React Native, ReactJs là công nghệ sử dụng việc xây dựng ứng dụng web dựa trên JavaScript được duy trì và phát triển bởi Facebook. ReactJs được nhận xét là dễ học, dễ hiểu, dễ làm quen nhưng hiệu quả lại cao, phù hợp với nhiều ứng dụng web, kể cả những ứng dụng có quy mô cực lớn (như Facebook). Tương tự React Native, ReactJs cũng có cộng đồng lớn mạnh, nhiều thư viện và framework hỗ trợ [5].

2.4 ReactJs

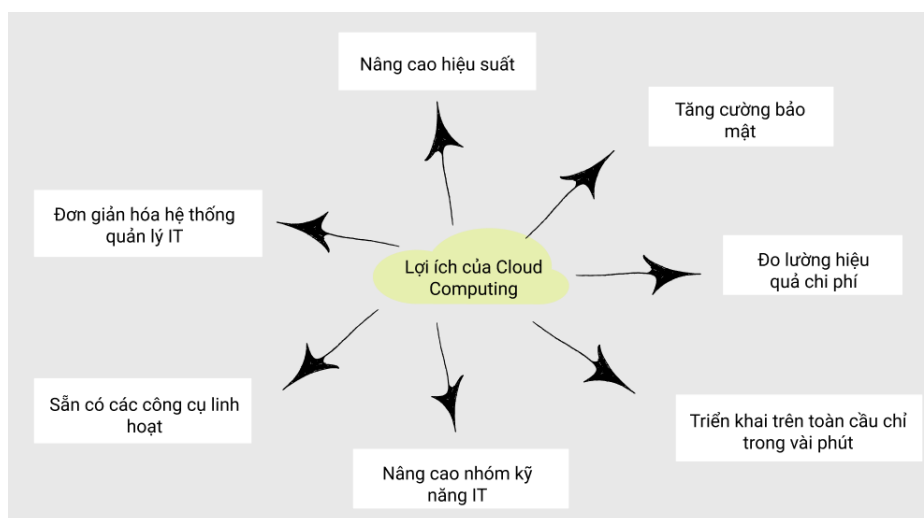
ReactJS là một mã nguồn mở (opensource) được phát triển bởi Facebook, ra mắt vào năm 2013, bản thân nó là một thư viện Javascript được dùng để xây dựng các tương tác với các thành phần trên website. Một trong những điểm nổi bật nhất của ReactJS đó là việc render dữ liệu không chỉ thực hiện được trên tầng Server mà còn ở dưới Client nữa [6].

2.5 Cloud Computing với AWS

Điện toán đám mây (Cloud Computing) là việc phân phối các tài nguyên công nghệ thông tin theo nhu cầu qua Internet với chính sách thanh toán theo mức sử dụng [7]. Thay vì mua, sở hữu và bảo trì các trung tâm dữ liệu và máy chủ vật lý, bạn có thể tiếp cận các dịch vụ công nghệ, như năng lượng điện toán, lưu trữ và cơ sở dữ liệu, khi cần thiết, từ nhà cung cấp dịch vụ đám mây như Amazon Web Services (AWS).

Amazon Elastic Compute Cloud (Amazon EC2) là một dịch vụ web cung cấp khả năng tính toán an toàn, có thể thay đổi kích thước trên đám mây. Nó được thiết kế để làm cho điện toán đám mây quy mô web trở nên dễ dàng hơn cho các nhà phát triển. Giao diện dịch vụ web đơn giản của Amazon EC2 cho

phép bạn lấy và định cấu hình dung lượng với mức tối thiểu. Nó cung cấp cho bạn quyền kiểm soát hoàn toàn các tài nguyên máy tính của bạn và cho phép bạn chạy trên môi trường máy tính đã được chứng minh của Amazon [8].



Hình 2.5.1 Cloud Computings [7]

Tính đàn hồi. Amazon EC2 cho phép bạn tăng hoặc giảm nhu cầu tài nguyên trong vòng vài phút, không cần phải là hàng giờ hoặc nhiều ngày. Bạn có thể thực hiện một, hàng trăm, thậm chí hàng ngàn trường hợp máy chủ cùng lúc. Dịch vụ hosting đám mây linh hoạt. Bạn có thể lựa chọn nhiều loại instance, các hệ điều hành và các gói phần mềm. Amazon EC2 cho phép bạn chọn cấu hình bộ nhớ, CPU, bộ nhớ instance, và kích thước phân vùng khởi động với mục tiêu tối ưu cho sự lựa chọn của hệ điều hành và ứng dụng. Bảo mật đám mây tại AWS là ưu tiên cao nhất. Amazon EC2 cung cấp một môi trường đáng tin cậy cao. Dịch vụ này chạy trong cơ sở hạ tầng mạng đã được kiểm chứng của Amazon và các trung tâm dữ liệu đầy đủ tiêu chuẩn cao.

Ưu điểm:

- Quen thuộc với người dùng AWS và dễ sử dụng.
- An toàn cao vì các chính sách có thể được định cấu hình để quản lý quyền và kiểm soát quyền truy cập vào hình ảnh của người dùng sử dụng vai trò và người dùng AWS IAM mà không cần phải quản lý thông tin đăng nhập trực tiếp trên các phiên bản EC2.
- Không có phí trả trước hoặc cam kết. Bạn chỉ phải trả cho lượng dữ liệu bạn lưu trữ trong kho lưu trữ của mình và dữ liệu được chuyển lên Internet.
- Tích hợp chặt chẽ với Amazon ECS và Docker CLI, cho phép bạn đơn giản hóa quy trình phát triển và sản xuất của mình.

Nhược điểm:

- Thiếu cái nhìn sâu sắc về việc sử dụng số đăng ký.
- Khó sử dụng với ứng dụng khách docker vì nó yêu cầu tạo mã thông báo tạm thời.
- Có thể tốn kém nếu các vùng chứa đang được triển khai không có trong AWS.

Amazon Simple Storage Service (Amazon S3) là dịch vụ lưu trữ đối tượng với giao diện dịch vụ web đơn giản để lưu trữ và truy xuất bất kỳ lượng dữ liệu nào từ bất kỳ đâu trên web. S3 là dịch vụ lưu trữ dạng object storage. S3 chúng ta có thể lưu trữ file, ảnh, video hoặc rất nhiều thứ khác. S3 có nhiều kiểu lưu trữ, hỗ trợ những kiểu truy cập khác nhau. Bạn có thể cấu hình S3 intelligent – tiering cho phép tự động chuyển các kiểu lưu trữ. Lưu trữ dữ liệu trong Amazon S3 và bảo mật dữ liệu khỏi bị truy cập trái phép bằng các tính năng mã hóa và công cụ quản lý truy cập. S3 là dịch vụ lưu trữ đối tượng duy nhất cho phép chặn truy cập công khai vào tất cả các đối tượng ở mức bucket hoặc tài khoản với S3 Block Public Access [9].

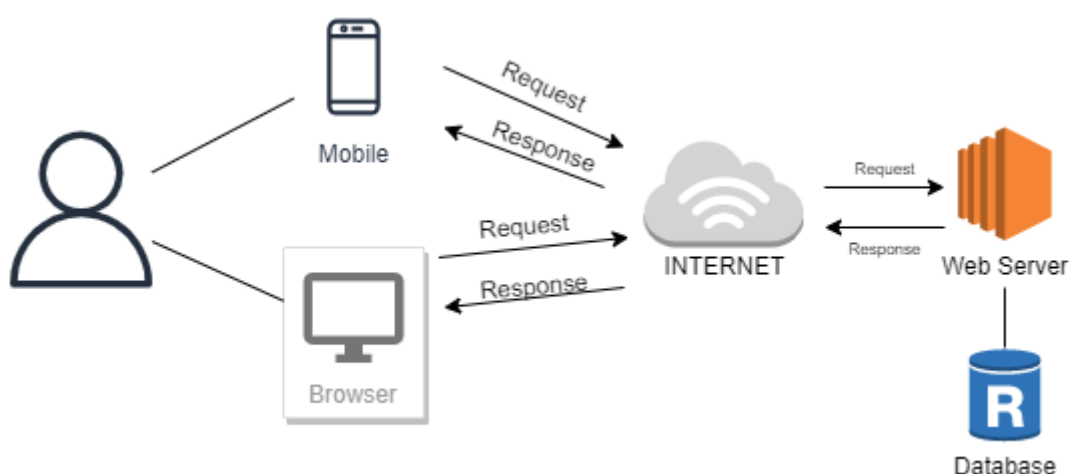
Amazon Relational Database Service (Amazon RDS) giúp dễ dàng thiết lập, vận hành và mở rộng quy mô cơ sở dữ liệu quan hệ trên đám mây. Nó cung cấp khả năng tiết kiệm chi phí và có thể thay đổi kích thước trong khi tự động hóa và các tác vụ quản trị tốn thời gian như cung cấp phân cứng, thiết lập cơ sở dữ liệu, vá lỗi và sao lưu. Nó giúp bạn tập trung vào các ứng dụng của mình để bạn có thể cung cấp cho chúng

hiệu suất nhanh, tính sẵn sàng cao, bảo mật, và khả năng tương thích mà chúng cần [9]. Thao tác tăng giảm tài nguyên dễ dàng chỉ với một vài cú nhấp chuột hoặc với một lệnh gọi API mà không phải ngừng hoạt động. Một vài CSDL cho phép chạy một hoặc nhiều read replicas giảm truy cập đọc dữ liệu tới database chính. Bảo vệ rất tốt database trước các cuộc tấn công.

3. IUHEdu: CÔNG CỤ HỖ TRỢ HỌC TẬP CHO SINH VIÊN TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

3.1 Giới thiệu

Để đảm bảo hiện thực yêu cầu bài toán đưa ra, chúng tôi sử dụng các công nghệ như ReactJs, React Native, Spring Boot, MySql. Để triển khai ứng dụng đến người dùng (giáo vụ, giảng viên, sinh viên) chúng tôi cũng sử dụng các dịch vụ như Amazon Web Service (AWS), CI/CD. Hệ thống của chúng tôi bao gồm ứng dụng di động (mobile app) và hệ thống website (web app) như Hình 3.1.1 đã mô tả kiến trúc của hệ thống.

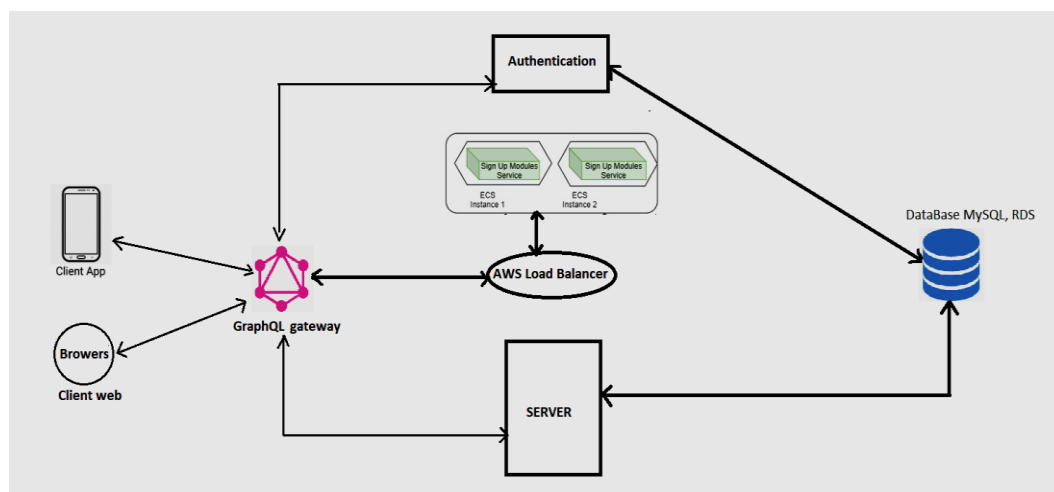


Hình 3.1.1 Sơ đồ hoạt động hệ thống

Trong đó, chúng tôi sử dụng React Native để xây dựng ứng dụng cho thiết bị di động Android/IOS, phục vụ các chức năng của sinh viên như đăng ký học phần, xem chương trình khung, xem điểm, xem lịch học... ReactJs được sử dụng để xây dựng hệ thống website phục vụ cho giáo vụ, giảng viên và người quản trị hệ thống truy cập để quản lý công tác tuyển sinh, học vụ của sinh và quản lý các hạng mục khác liên quan. Spring Boot dùng để tạo ra các API, xử lý các request trả về response để phục vụ cho các ứng dụng website và mobile. MySql được sử dụng để lưu trữ toàn bộ dữ liệu của hệ thống, như dữ liệu sinh viên, học phần, lớp học phần, ... Các dịch vụ điện toán đám mây sử dụng như Elastic Compute Cloud (EC2), Relational Database Service (RDS), S3, Auto Scaling, Load Balancer. Các nhiệm vụ mà các service này đảm nhiệm như EC2 là một máy ảo Linux dùng để cài đặt server, RDS là server database dùng để lưu trữ dữ liệu của toàn hệ thống, S3 dùng để lưu trữ file tài liệu và hình ảnh cho toàn hệ thống. AWS Auto Scaling là tính năng tự động nhân rộng để đảm bảo rằng các phiên bản Amazon EC2 đủ để chạy các ứng dụng cài đặt. Chúng ta có thể tạo một nhóm AWS Auto Scaling trong các phiên bản EC2. Load Balancer (Cân bằng tải) là công việc giúp phân phối hiệu quả các traffic từ ngoài truy cập đến các server dịch vụ. Giúp việc mở rộng khả năng chịu tải bằng phương pháp mở rộng theo chiều ngang trở nên hiệu quả.

3.2 Kiến trúc hệ thống

Chúng tôi sử dụng kiến trúc Client – Server để xây dựng ứng dụng, bởi vì mô hình này hoạt động theo cách là khi máy client gửi yêu cầu đến server, lúc này server sẽ xử lý yêu cầu và trả kết quả về cho client. Client sẽ không chia sẻ bất kỳ tài nguyên nào của họ. Mô hình Client Server có một số ưu điểm chính để lựa chọn như tính tập trung, tính bảo mật, có khả năng mở rộng và khả năng truy cập (Hình 3.2.1).



Hình 3.2.1. Cấu trúc của hệ thống IUHEdu.

3.3 Phân tích hệ thống IUHEdu của trường

Đối tượng người dùng hệ thống IUHEdu bao gồm quản trị viên và sinh viên của trường. Hệ thống tập trung vào khó khăn của trường đang gặp phải là hệ thống bị quá tải khi đến thời gian đăng ký học phần. Thêm vào đó còn có các chức năng của quản trị viên cũng được thêm vào hệ thống để hỗ trợ cho phân đăng ký học phần cũng như là quản lý sinh viên. Các chức năng chính của quản trị viên và sinh viên được mô tả ở bảng sau (Bảng 3.3.1).

Bảng 3.3.1 Danh sách các chức năng trong hệ thống IUHEdu

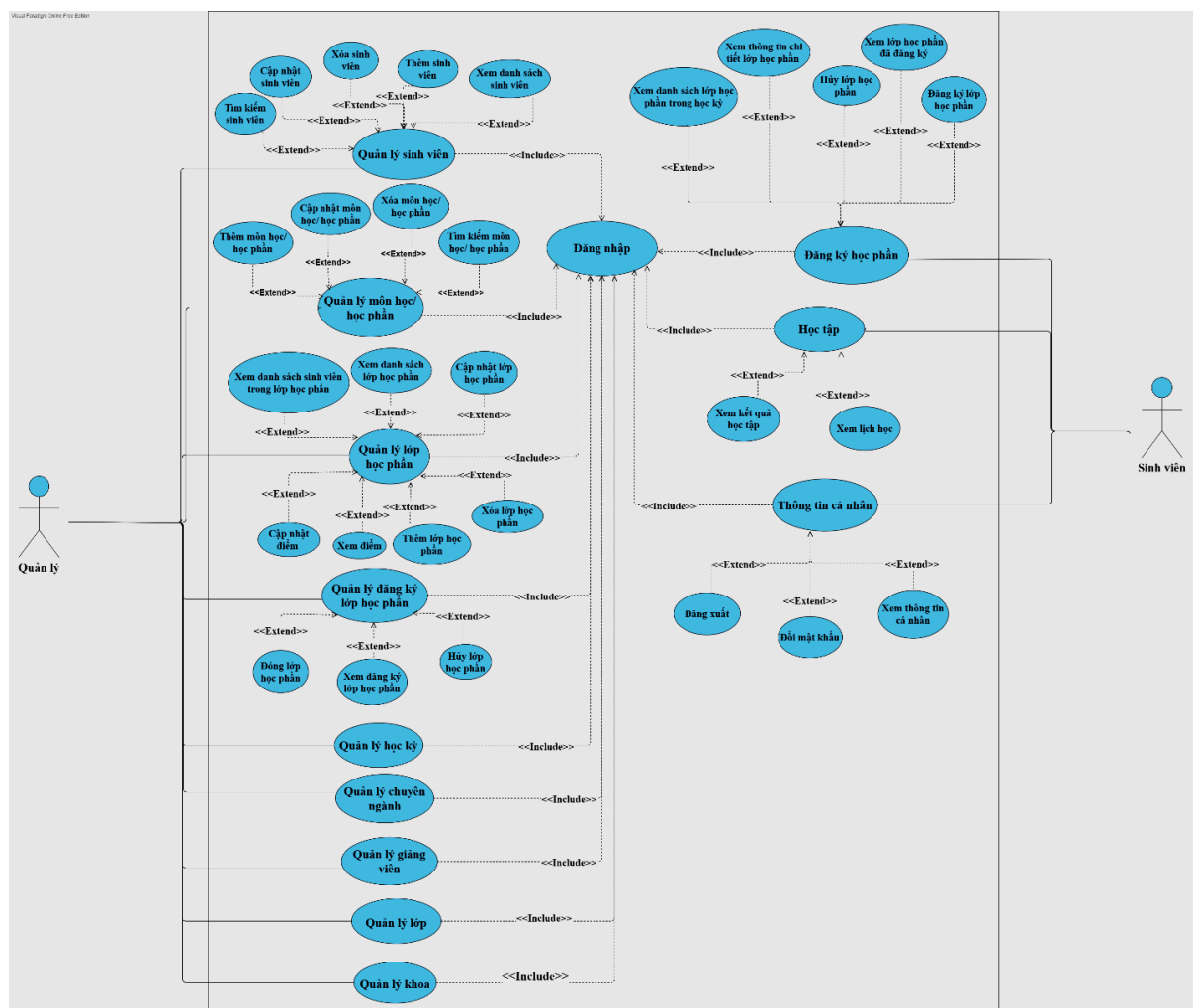
Đối tượng	Tên chức năng	Mô tả
Sinh viên	Xem thông tin cá nhân	Xem thông tin cá nhân của sinh viên đã đăng nhập vào ứng dụng
	Đăng nhập tài khoản	Đăng nhập vào hệ thống với tài khoản được tạo
	Xem lịch học	Xem lịch học của các môn học đã đăng ký trong kỳ đó, xem lịch theo tuần, xem lịch theo ngày.
	Đăng ký học phần	Gồm các chức năng như xem danh sách học phần đã đăng ký, đăng ký lớp học phần, hủy đăng ký lớp học phần, xem chi tiết lịch học lớp học phần, kiểm tra các lớp học phần có bị trùng nhau hay không.
	Quản lý thông báo	Cho phép nhận thông báo lớp học phần thay đổi, xem lịch sử thông báo.
	Xem chương trình khung	Xem thông tin các môn học theo học kỳ của sinh viên.
	Xem kết quả học tập	Xem kết quả các môn đã học của các học kỳ, xem chi tiết điểm của từng môn học.
Quản trị viên	Quản lý khoa	Cho phép thêm, xóa, sửa thông tin của khoa, xem danh sách khoa.
	Quản lý chuyên ngành	Cho phép thêm, xóa, sửa thông tin của chuyên ngành, xem danh sách chuyên ngành.
	Quản lý sinh viên	Cho phép thêm (Thêm bằng file excel hoặc thêm 1 sinh viên), xóa, sửa thông tin của sinh viên, xem danh sách sinh viên.
	Quản lý môn học	Cho phép thêm, xóa, sửa thông tin của môn học, xem danh sách môn học.
	Quản lý học phần	Cho phép thêm, xóa, sửa thông tin của học phần, xem danh sách học phần.

	Quản lý lớp học phần	Cho phép thêm, xóa, sửa thông tin của lớp học phần, xem danh sách lớp học phần.
	Quản lý giảng viên	Cho phép thêm, xóa, sửa thông tin của lớp học phần, xem danh sách giảng viên theo từng chuyên ngành.
	Quản lý học kỳ	Cho phép thêm, xóa, sửa thông tin của lớp học phần, xem danh sách học kỳ.

Từ những chức năng trên, chúng tôi tập trung vào các chức năng chính như: Đăng ký học phần, quản lý lớp học phần. Sau đó chúng tôi mở rộng các chức năng tương ứng như trong Hình 3.3.1.

Trong mô hình use-case chúng tôi trình bày, gồm có hai thành phần actor chính mô tả cho hai đối tượng người dùng mà chúng tôi hướng đến (sinh viên, nhân viên quản trị trường học). Trong đó, những use-case chính và phức tạp về nghiệp vụ đòi hỏi nhiều xử lý cho hệ thống đó là: use-case quản lý đăng ký học phần và use-case đăng ký học phần. Ngoài ra, chúng tôi cũng mở rộng các use-case cho các chức năng của hệ thống nhằm tạo điều kiện thuận lợi cho người dùng là sinh viên và nhân viên quản trị.

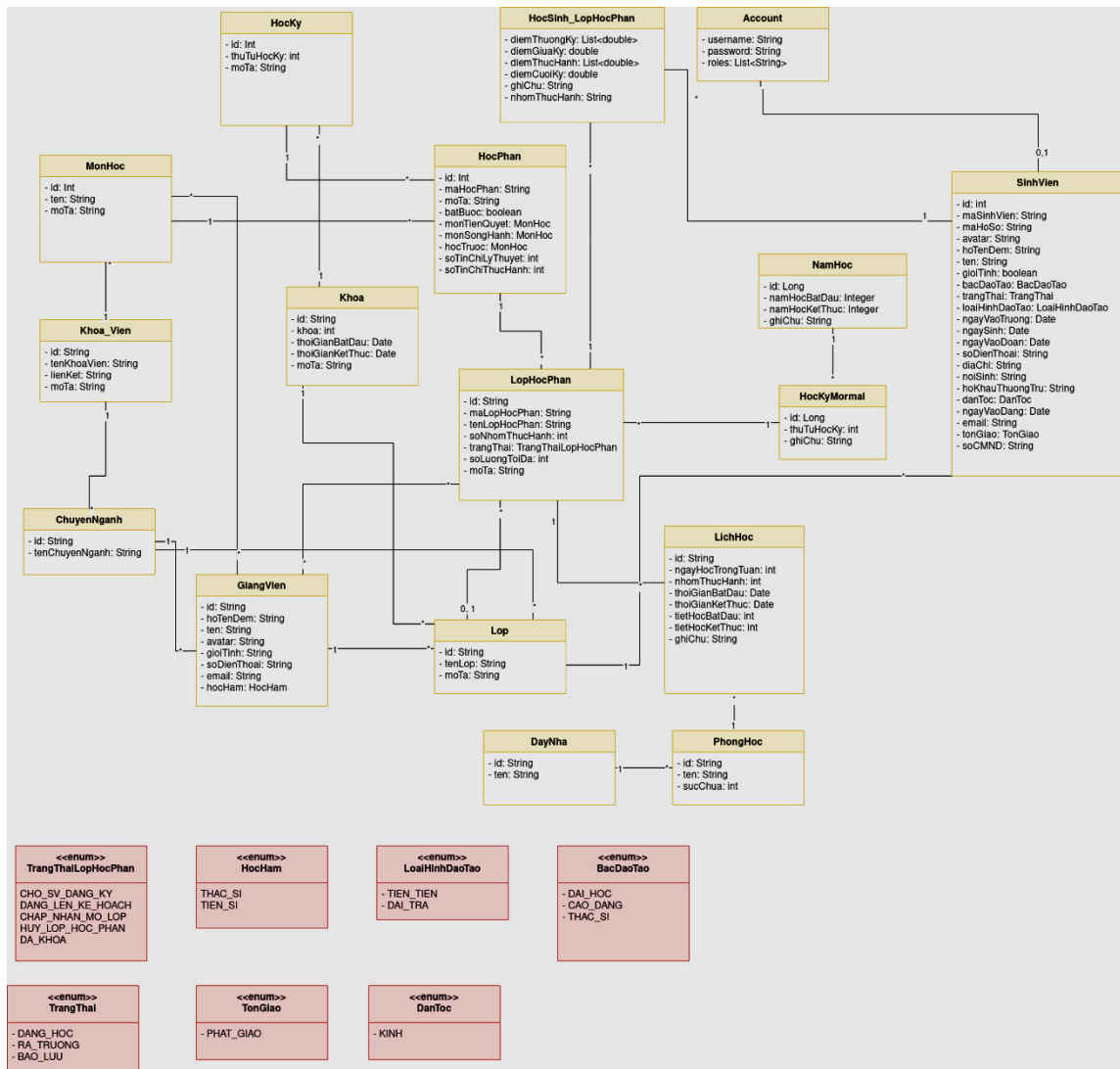
Từ các usecase đã phân tích ở trên thì chúng tôi đã tạo ra được lược đồ cơ sở dữ liệu của IUHEdu. Trong đó, chúng tôi chú trọng đến bảng sinh viên và bảng lớp học phần. Hai bảng này giúp có khá nhiều dữ liệu cho chức năng chính là đăng ký học phần và quản lý học phần. Ngoài ra, còn các bảng dữ liệu khác hỗ trợ cho chức năng chính đồng thời làm rõ các mối quan hệ trong lúc xử lý dữ liệu. Lược đồ dữ liệu chúng tôi sẽ triển khai dựa vào lược đồ Lớp được thiết kế như Hình 3.3.1



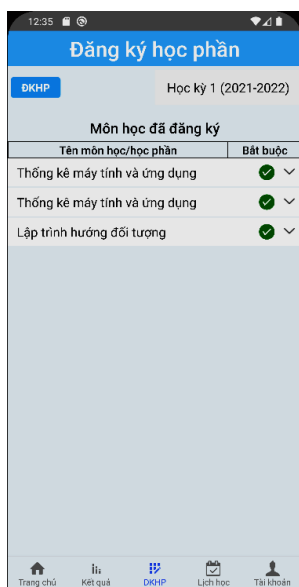
Hình 3.3.1 Sơ đồ use case của hệ thống IUHEdu

3.4 Giao diện đồ họa

Hệ thống IUHedu triển khai trên hai nền tảng là Web app và Mobile app. Trong đó, hệ thống app cho phép người dùng download và cài đặt trên các thiết bị chạy trên nền Android/iOS. Công cụ này sẽ có giao diện người dùng thân thiện và đảm bảo tính ổn định trên các dòng điện thoại khác nhau. Chúng tôi cũng xin giới thiệu vài giao diện cho các chức năng trên mobile và Web. Sinh viên đăng nhập vào hệ thống bằng tài khoản mà nhà trường cung cấp là mã số sinh viên và mật khẩu. Sau khi đăng nhập thành công sẽ hiện giao diện trang chủ. Ở thanh menu chọn chức năng đăng ký học phần là một chức năng chính của hệ thống. Chức năng này cho sinh viên xem được chi tiết lớp, kiểm tra lịch trùng, kiểm tra môn đã đăng ký hay những lớp học phần được mở trong học kỳ đó. Để thực hiện chức năng, sinh viên cần thực hiện các bước sau: **Bước một**, sinh viên phải chọn học kỳ cần đăng ký học phần, danh sách học kỳ hiển thị dưới dạng danh sách (dropdown) và nhấn nút “ĐKHP” (Hình 3.4.3)

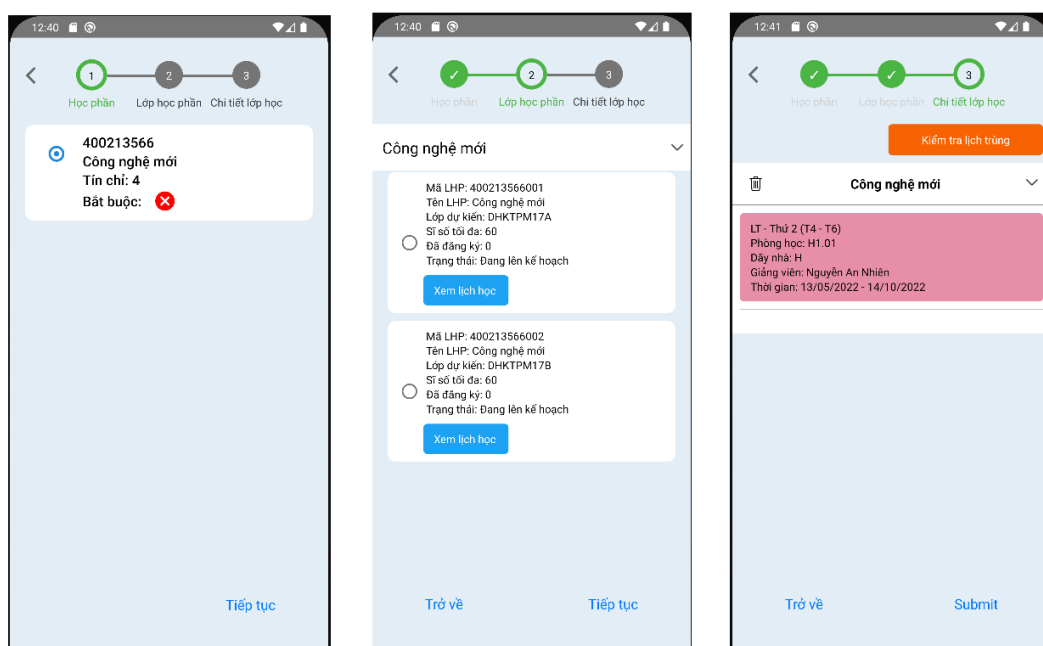


Hình 3.4.2 Sơ đồ Class diagram của hệ thống IUHedu



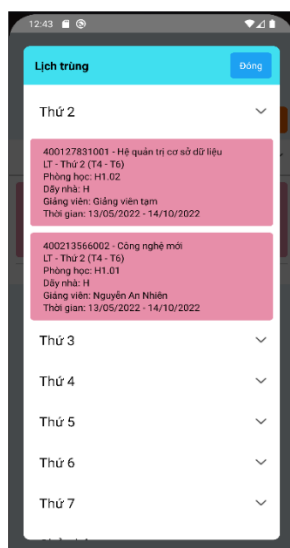
Hình 3.3.2. Giao diện xem các môn đã đăng ký học phần

Bước hai, sinh viên phải chọn các học phần muốn đăng ký và nhấn nút “Tiếp tục”. Sau khi tiếp tục, ứng dụng sẽ tìm nạp và hiển thị danh sách những lớp học phần tương ứng với học phần đã chọn. **Bước ba**, sinh viên chọn các lớp muốn đăng ký học và nhấn nút “Tiếp tục”. Ứng dụng sẽ tìm nạp và hiển thị chi tiết (thời khóa biểu, nhóm thực hành) của lớp học đó. Nếu đó là lớp có học phần thực hành, sinh viên phải chọn nhóm thực hành cho mình rồi mới có thể chọn nút đăng ký. **Bước bốn**, sinh viên nhấn nút “Submit”, ứng dụng sẽ thông báo rõ ràng đăng ký thành công hay thất bại, nếu thất bại thì sẽ nêu rõ lý do (Hình 3.4.3).



Hình 3.4.3 Giao diện chọn học phần

Người dùng có thể nhấn nút kiểm tra lịch trùng để kiểm tra xem lớp học phần và nhóm thực hành đã chọn có trùng với những môn học đã đăng ký trước đó hay không. Ứng dụng sẽ hiển thị danh sách các học phần bị trùng lịch học (Hình 3.4.4)



Hình 3.4.4 Giao diện xem lịch trùng

Sau khi đăng ký học phần thì có thể xem được tính năng xem lịch học theo tuần. Sinh viên có thể tìm một ngày bất kỳ để xem lịch học hoặc có thể xem lịch học ở tuần hiện tại, kế tiếp, tuần quá khứ (Hình 3.4.5). Để xem được kết quả của môn học, sinh viên chọn tính năng kết quả (Hình 3.4.6), xem tổng quan điểm trung bình các môn học trong các học kỳ và có thể nhấn vào môn học để có thể xem chi tiết điểm của môn học đó.



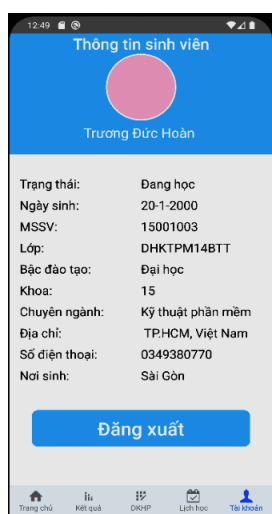
Hình 3.4.5 Giao diện lịch học



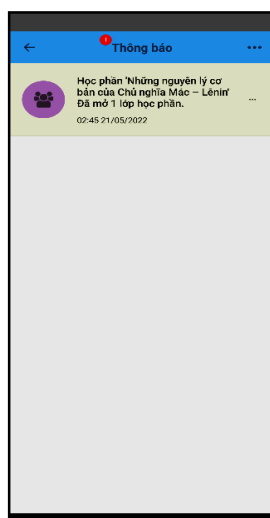
Hình 3.4.6 Giao diện kết quả học tập

Chức năng tài khoản giúp cho sinh viên có thể xem được chi tiết thông tin sinh viên và thực hiện chức năng đăng xuất (Hình 3.4.7). Ngoài những tính năng ở trên thì chúng tôi còn có chức năng thông báo (Hình 3.4.8) và chương trình khung (Hình 4.1). Chức năng thông báo dùng để thông báo các thay đổi về thông tin

lớp học phần cho sinh viên. Chức năng chương trình khung giúp sinh viên biết được học kỳ này gồm có những môn nào, bao nhiêu tín chỉ.



Hình 3.4.7 Giao diện thông tin sinh viên

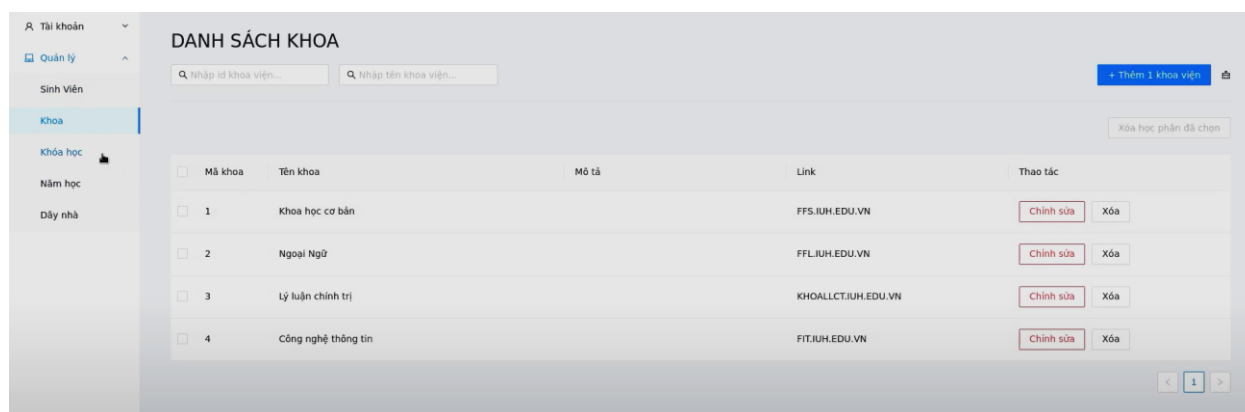


Hình 3.4.8 Giao diện thông báo



Hình 3.4.9 Giao diện chương trình khung

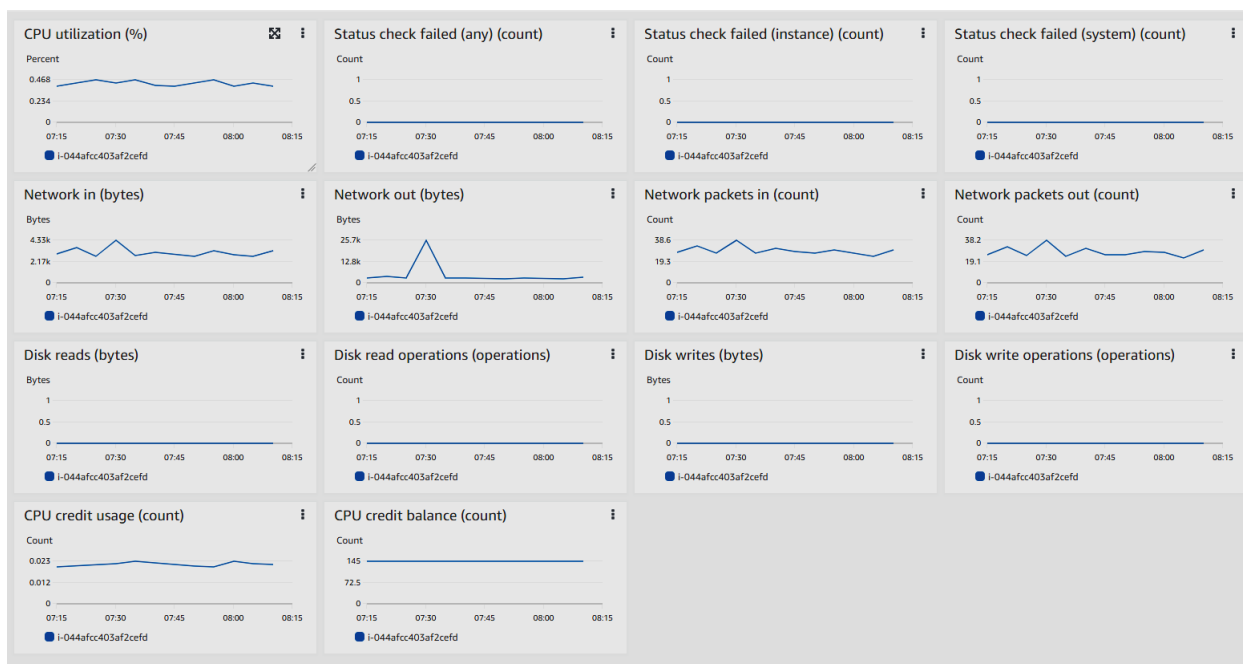
Ngoài ra, các tính năng xử lý trên Web cho nhân viên đào tạo (Quản trị viên) chúng tôi cũng đã triển khai. Nhân viên đào tạo có thể đăng nhập vào hệ thống và thực hiện các chức năng như: thêm các chương trình đào tạo, mở các học phần trong các học kỳ,... Bên cạnh đó, hệ thống cũng sẽ cho phép thực hiện các chức năng với ràng buộc về thời gian cũng như các ràng buộc nghiệp vụ khác trong công tác đào tạo tại trường (Hình 3.4.10).



Hình 3.4.10 Giao diện quản lý của hệ thống IUHEdu dành cho nhân viên đào tạo

4. KIỂM NGHIỆM THỰC TẾ

Chúng tôi sử dụng dịch vụ ECS Cluster của AWS để triển khai hệ thống, ECS Cluster, ECS là một service quản lý container có tính scale cao và nhanh. Dễ dàng run, stop, hay quản lý container ở trong một cluster. Chính vì tính scale cao của dịch vụ ECS Cluster đã giúp cho hệ thống hoạt động một cách ổn định và được đảm bảo trong khoảng thời gian khi hệ thống tăng tốc xử lý đồng thời cũng có thể theo dõi, giám sát các tài nguyên bên trong dịch vụ ECS Cluster thông qua chức năng monitoring như Hình 4.1. Trong hình này, chúng tôi cũng đã thử nghiệm cho các thời điểm xử lý cùng lúc nhiều giao dịch (đăng ký học phần) tại một thời điểm và ghi nhận các thông số về sử dụng về tài nguyên theo cấu hình ECS.



Hình 4.1 ECS Cluster Monitoring trên AWS cho hệ thống IUHedu

Do hệ thống phần mềm đang trong phiên bản thử nghiệm nên cấu hình máy chủ cloud sẽ sử dụng một cấu hình không quá cao để thử nghiệm. Trong tương lai, khi đưa sản phẩm vào thực tế sẽ nâng cấp cấu hình dịch vụ lên. Đồng thời phiên bản mobile của hệ thống cũng sẽ được triển khai lên trên Google Play và AppStore với tên là IUHedu để thuận tiện phục vụ cho việc sử dụng của sinh viên.

5. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài báo này chúng tôi đã đưa ra giải pháp cho bài toán xây dựng công cụ *IUHedu* nhằm hỗ trợ cho sinh viên tại Trường Đại học Công nghiệp Thành phố Hồ Chí Minh thực hiện các xử lý học vụ tại trường. Trong đó, giải pháp đầu tiên chúng tôi đã triển khai ứng dụng chạy trên thiết bị di động có tên là *IUHedu*. Ứng dụng này chạy trên cả hai nền tảng thông dụng là Android và iOS. Giải pháp thứ hai, chúng tôi cũng triển khai cho công cụ trên và hệ thống website trên nền tảng điện toán đám mây AWS. Đồng thời cũng kiểm chứng được tính ổn định của hệ thống khi triển khai trên nền tảng này. Trong cả hai giải pháp trên, chúng tôi đều áp dụng công nghệ phù hợp nhằm tối ưu tài nguyên cũng như đảm bảo được tính hiệu quả của hệ thống. Cụ thể là phần ứng dụng trên di động chúng tôi triển khai trên công nghệ như React Native tạo ứng dụng, ReactJS tạo website cho quản trị, Spring Boot tạo các hàm API cho GraphQL gateway. Phần triển khai trên điện toán đám mây chúng tôi dùng Elastic Container Service (ECS); RDS; Auto Scaling; Load Balancer của AWS. Tuy nhiên, công cụ vẫn còn một số hạn chế là quá trình chúng tôi thử nghiệm chỉ với số lượng truy xuất vừa và nhỏ. Hệ thống cần thêm việc thử nghiệm DOS server, tạo ra lượng truy cập ảo (có thể lên tới 100.000 request) để tiến hành monitoring, maintain nhằm đảm bảo tính mở rộng của hệ thống. Ngoài ra, các chức năng trên hệ thống web dành cho nhà quản trị vẫn còn hạn chế, chưa đáp ứng hết các chức năng hỗ trợ. Phần ứng dụng trên thiết bị di động cũng chưa được triển khai trên các store để hỗ trợ cho sinh viên sử dụng. Trong tương lai, chúng tôi sẽ tìm hiểu và khắc phục những tính năng còn thiếu sót và đưa ra các bài toán mới, cách giải quyết mới để giải quyết tình trạng trên.

TÀI LIỆU THAM KHẢO

1. Sim, W., Choi, W.: Spring boot-based web application development for providing information on security vulnerabilities and patches for open source software. Journal of Korea Society of Digital Industry and Information Management 17(4) (2021) 77–83

2. Spring boot Architecture. [online] <https://dzone.com/articles/spring-boot-architecture-and-workflow> [Accessed 22/04/2022]
3. Seabra, M., Nazário, M.F., Pinto, G.: Rest or graphql? a performance comparative study. In: Proceedings of the XIII Brazilian Symposium on Software Components, Architectures, and Reuse. (2019) 123–132
4. Doglio, F., Doglio, Corrigan: REST API Development with Node. js. Volume 331. Springer (2018)
5. “Tìm hiểu về React Native”. [online]. Địa chỉ: <https://wiki.tino.org/react-native-la-gi/> [Accessed 17/04/2022]
6. Rawat, P., Mahajan, A.N.: Reactjs: A modern web development framework. International Journal of Innovative Science and Research Technology 5(11) (2020)
7. “Điện toán đám mây là gì?”. [online]. Địa chỉ: <https://aws.amazon.com/vi/what-is-cloud-computing/> [Accessed 14/04/2022]
8. Chapin, J., Roberts, M.: Programming AWS Lambda: build and deploy serverless applications with Java. " O'Reilly Media, Inc." (2020)
9. Mukherjee, S.: Benefits of AWS in modern cloud. Available at SSRN 3415956 (2019)

YSC4F.242

MỘT HƯỚNG TIẾP CẬN GIẢI QUYẾT VẤN ĐỀ TẮC NGHẼN TRÊN ỨNG DỤNG WEB BẰNG ĐIỆN TOÁN Đám Mây

ĐẶNG TRƯỜNG AN, NGUYỄN VIỆT TIẾN, TÔN LONG PHƯỚC,

Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh

andang.100202@gmail.com, tiennguyenk71.work@gmail.com, tonlongphuoc@iuh.edu.vn

Tóm tắt. Tắc nghẽn trong quá trình xử lý trên các ứng dụng Web hiện nay có nhiều nguyên nhân, trong đó nguyên nhân chính vẫn là tại một thời điểm hệ thống có quá nhiều truy xuất. Nhằm giảm thiểu số lượng tắc nghẽn trong các hệ thống này, chúng tôi đề xuất hướng tiếp cận sử dụng các dịch vụ trong điện toán đám mây cho ứng dụng. Chúng tôi lấy một trường hợp để minh họa cho hướng đề xuất này là hệ thống ứng dụng Web và thiết bị di động trong công tác hỗ trợ sinh viên học tập tại Trường Đại học Công nghiệp Thành phố Hồ Chí Minh. Hệ thống sẽ có hai phân hệ là Web và ứng dụng di động. Chúng tôi đã dùng các công nghệ như Ngôn ngữ lập trình Java, Javascript, NPM, NodeJS, Express và các dịch vụ AWS và Heroku để xây dựng hệ thống trên. Kiến trúc phần mềm chúng tôi dùng kỹ thuật Microservice để tối ưu kiến trúc và hiệu suất các xử lý trong hệ thống. Trong đó vấn đề cốt lõi cho việc giảm thiểu tắc nghẽn chúng tôi dùng dịch vụ AutoScaling, Load Balancer của AWS. Chúng tôi cũng đã thử nghiệm hệ thống này cho các sinh viên thuộc chuyên ngành Kỹ thuật phần mềm thuộc Khoa Công nghệ Thông tin. Các giao dịch thường xảy ra tắc nghẽn đã được kiểm chứng và mang lại hiệu quả cao. Trong tương lai, chúng tôi sẽ triển khai cho sinh viên toàn trường và mở rộng cho các trường khác.

Từ khoá. tắc nghẽn xử lý, ứng dụng Web, ứng dụng di động, AWS, Heroku, điện toán đám mây

AN APPROACH TO THE PROBLEM OF CONGESTION IN WEB APPLICATIONS BY CLOUD COMPUTING

Abstract. Congestion in processing on Web applications today has many causes, of which the main cause is still because the system has too many accesses at a time. In order to reduce the number of bottlenecks in these systems, we propose an approach to using services in cloud computing for the application. We take a case to illustrate this proposed direction, which is the Web application system and mobile device in supporting students studying at the Industrial University of Ho Chi Minh City (IUH). The system will have two modules: Web and Mobile applications. We used technologies like Java Programming Language, JavaScript, NPM, NodeJS, Express and AWS services to build the above system. Our software architecture uses Micro-service technology to optimize the architecture and performance of the processes in the system. In which the core problem for reducing congestion we use Autoscaling service, Load Balancer of AWS and Heroku. We also tested the system for students in the Software Engineering major in the Faculty of Information Technology. Transactions that are often congested are proven and highly effective. In the future, we will roll out to students across the IUH and expand to other university.

Keywords. congestion, web Application, mobile Application, AWS, Heroku, Cloud Computing

1. GIỚI THIỆU TỔNG QUAN

Hiện nay với việc phát triển của công nghiệp 4.0, nhu cầu sử dụng và phát triển trên môi trường kỹ thuật số có trên mọi lĩnh vực, trong đó có lĩnh vực giáo dục. Cụ thể tại Trường Đại học Công nghiệp Thành phố Hồ Chí Minh, số lượng sinh viên nghiên cứu và học tập tại trường ngày càng tăng, kéo theo đó là nhu cầu quản lý học tập ngày càng cao. Chúng tôi đã xây dựng chương trình hỗ trợ học tập sinh viên IUH với hai nền tảng đó là Website và Android. Đây là một chương trình có tính ứng dụng cao, bám sát được chương

trình phát triển của nhà trường. Chương trình được tạo ra để hỗ trợ nhân viên quản lý chương trình đào tạo tại trường (Môn học phần, lớp học phần, sinh viên, xếp lịch học...) quan trọng hơn là đơn giản hoá việc mở lớp học phần, đăng ký học phần và giảm thiểu tắc nghẽn dữ liệu. Còn về phía sinh viên, chương trình có thể hỗ trợ việc học tập và nghiên cứu của sinh viên tại trường (Đăng ký học phần, xem điểm, xem lịch học, xem chương trình khung, xem công nợ...) trở nên dễ dàng và thuận tiện hơn. Đáng chú ý nhất là việc đăng ký học phần thường dẫn đến quá tải hệ thống khi quá nhiều truy xuất vào hệ thống vào một thời điểm. Điều này đã dẫn đến việc các sinh viên chậm đăng ký học phần và tệ hơn là không thể đăng ký học phần.

Một trong những nguyên nhân chính dẫn đến việc quá tải này là do có quá nhiều lượt truy xuất tới hệ thống trong một thời điểm dẫn tới quá tải hệ thống và nguy hiểm hơn là sập hệ thống. Giải pháp đơn giản nhất cho vấn đề này chính là nâng cấp cấu hình của máy chủ lên để có thể xử lý được nhiều yêu cầu trong một thời điểm hơn. Tuy nhiên giải pháp này sẽ có một số câu hỏi được đặt ra như: Nâng cấp như thế nào, có rủi ro không? Nâng cấp trong bao lâu? Chi phí nâng cấp như thế nào? Hiện tượng tắc nghẽn chỉ xảy ra khi tới đợt đăng ký học phần thì có đáng để nâng cấp không?

Để giải quyết các vấn đề đó, chúng tôi đã đưa ra các giải pháp như dùng điện toán đám mây (cloud computing) nhằm tổ chức lưu trữ và xử lý các nghiệp vụ; xây dựng hệ thống website và ứng dụng trên di động (app) cho sinh viên dễ dàng xử lý và giảm tải cho máy chủ. Điện toán đám mây là một giải pháp có tính xu hướng toàn cầu cho các cá nhân, tổ chức, doanh nghiệp vừa và nhỏ. Hiện nay, nhu cầu tạo ra các hệ thống dịch vụ trực tuyến phục vụ cho nhu cầu cá nhân, kinh doanh ngày càng nhiều. Những hệ thống này khi triển khai trên điện toán đám mây sẽ có chi phí rẻ, cấu hình đơn giản nhưng vẫn an toàn, bảo mật tốt, khả năng nâng cấp bảo trì hệ thống dễ dàng. Hơn hết, trong các ưu điểm trên là khả năng cân bằng tải, đó là ưu điểm tuyệt vời của điện toán đám mây [1]. Trong đó, Scale là dịch vụ có khả năng tự động điều chỉnh số lượng máy chủ để đáp ứng nhu cầu truy cập của khách hàng (tăng khi yêu cầu nhiều và giảm khi yêu cầu ít). Do đó, nhà phát triển hệ thống không còn phải đầu tư quá nhiều chi phí chỉ để giải quyết nhu cầu cho việc sử dụng tài nguyên của hệ thống trong một quãng thời gian ngắn trong năm. Đó cũng là lý do chúng tôi chọn giải pháp này cho bài toán xây dựng hệ thống hỗ trợ đào tạo cho sinh viên.

Để triển khai giải pháp trên, chúng tôi đã triển khai hệ thống web với kiến trúc micro-services. Micro-service là một kỹ thuật phát triển phần mềm, một biến thể của kiến trúc hướng dịch vụ (SOA), trong đó cấu trúc một ứng dụng như một tập hợp các dịch vụ được ghép với nhau [2]. Trong kiến trúc micro-service, các dịch vụ được xử lý tốt và các giao thức rất nhẹ. Lợi ích của việc phân tách một ứng dụng thành các dịch vụ nhỏ hơn là nó cải thiện tính mô đun. Điều này làm cho ứng dụng trở nên gọn nhẹ, dễ hiểu, dễ phát triển và trở nên linh hoạt hơn. Cụ thể, chúng tôi chia hệ thống làm hai phần chính: Front-end và Back-end. Front-end của chúng tôi ở đây chính là ứng dụng di động cho sinh viên, ứng dụng website cho quản trị viên và sinh viên của nhà trường. Front-end sẽ đảm nhận trách nhiệm nhận dữ liệu, hiển thị dữ liệu cũng như các tương tác cho người dùng. Back-end chúng tôi xây dựng hai API cho hai nền tảng là Web và Android, hệ thống máy sẽ xử lý, lấy dữ liệu từ hệ cơ sở dữ liệu MySQL và trả về dưới định dạng JSON. Ngoài ra, khả năng phát triển song song và tương đối độc lập giữa Front-end và Back-end cũng là ưu điểm cho việc bảo trì và nâng cấp.

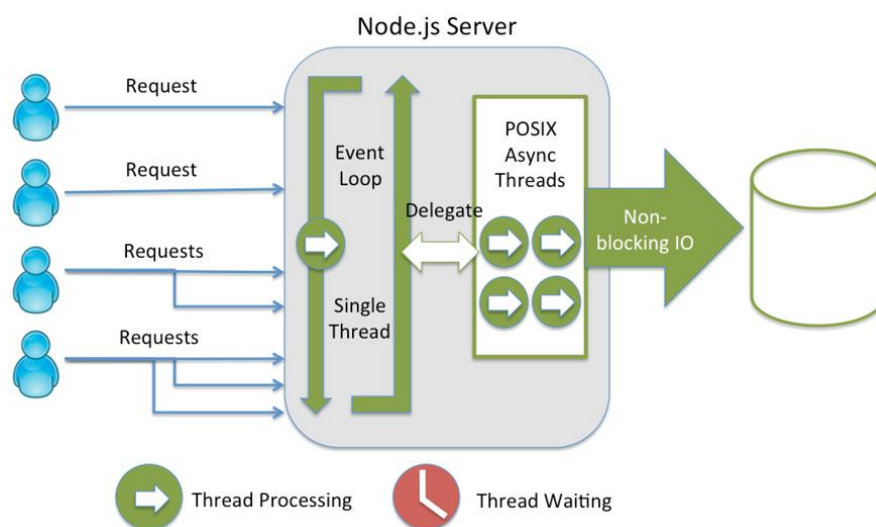
Còn về phần triển khai hệ thống thì quá trình sẽ tốn rất nhiều công sức, người lập trình viên phải cài đặt môi trường thích hợp để chạy chương trình. Khi đó có nhiều vấn đề phát sinh thêm như có thể phiên bản không tương thích, xung đột với các phần mềm khác... Để tránh các vấn đề đó và làm cho việc triển khai hệ thống một cách dễ dàng hơn, chúng tôi đã lựa chọn sử dụng Heroku. Heroku là nền tảng đám mây cho phép xây dựng, triển khai, quản lý và mở rộng ứng dụng (PaaS – Platform as a service). Heroku chạy các ứng dụng trong dynos – nó là một máy ảo mà có thể tăng giảm sức mạnh dựa vào độ lớn của ứng dụng [3, 4]. Nó rất linh hoạt và dễ sử dụng, cung cấp cho một con đường đơn giản nhất để đưa sản phẩm tiếp cận người dùng. Nó giúp các nhà phát triển tập trung vào phát triển sản phẩm mà không cần quan tâm đến việc vận hành máy chủ hay phần cứng... Chúng tôi sử dụng Heroku để deploy web và cơ sở dữ liệu. Đối với web, sử dụng github để lưu trữ và quản lý code, sau đó từ github kết nối tới Heroku để deploy. Còn với cơ sở dữ liệu, chúng tôi sử dụng dịch vụ JawsDB MySQL để làm nơi lưu trữ cơ sở dữ liệu.

Bài báo của chúng tôi chia làm 4 phần: *Phần 1 Giới thiệu tổng quan về hệ thống; Phần 2 Các công nghệ liên quan; Phần 3 Công cụ ATIUH; Phần 4 Kết luận và hướng phát triển.* Trong đó *Phần 3* là phần đóng góp chính của bài báo.

2. CÁC CÔNG NGHỆ LIÊN QUAN

2.1 Node.js

Node JS được phát triển từ Javascript V8 Engine từ năm 2009 bởi Ryan Dahl với mục đích đơn giản hóa việc xây dựng cho các ứng dụng web [5]. Nó cung cấp đầy đủ các thư viện phong phú ở dạng Javascript Module khác nhau giúp đơn giản hóa việc lập trình và giảm thời gian ở mức thấp nhất. Nó có thể hoạt động trên nhiều nền tảng khác nhau: MAC OS, Microsoft Windows, Linux. NodeJS dần được xem là một trong những công cụ hoàn hảo dành cho các ứng dụng chuyên sâu về dữ liệu theo khoảng thời gian thực. Theo khảo sát của Stack Overflow về các framework, nền tảng được sử dụng nhiều nhất năm 2019, NodeJS đã giành vị trí số 1 với số lượng người dùng lên đến gần 50%. Một hệ thống Website được xây dựng từ Node JS có một khả năng hoạt động rất ổn định. Đặc biệt, tốc độ xử lý tương đối cao nhờ cơ chế bất đồng bộ (non-blocking), giúp Website dễ dàng xử lý hàng ngàn kết nối cùng một lúc mà không gặp bất cứ khó khăn nào.



Hình 2.1.1. Kiến trúc luồng hoạt động của NodeJS [5]

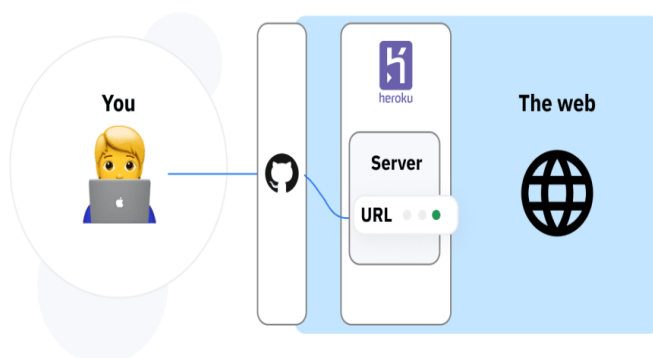
2.2 MySQL

MySQL là một hệ thống quản trị cơ sở dữ liệu mã nguồn mở (Relational Database Management System, viết tắt là RDBMS) dùng để tạo và quản lý các cơ sở dữ liệu theo hình thức quản lý các mối liên hệ giữa chúng. Độ phổ biến của MySQL là cực kỳ lớn, nó được sử dụng bởi các ứng dụng Web rất lớn như Facebook, Twitter, Youtube,... cho mục đích lưu trữ dữ liệu. Đồng thời nó tương thích với nhiều hạ tầng máy tính như Windows, Linux, MacOS, Ubuntu [6]. Máy tính hay hệ máy tính được cài đặt phần mềm MySQL dành cho server để giúp bạn lưu trữ dữ liệu trên đó, để máy khách có thể truy cập và quản lý. Còn máy khách (client) không hẳn là phải cài phần mềm MySQL mà là một phần mềm có các dòng lệnh truy vấn lên MySQL server. MySQL hoạt động theo mô hình client-server. Mỗi khi máy khách (client) cần truy cập dữ liệu, chúng sẽ kết nối tới máy chủ. Một máy client sẽ liên lạc với máy server trong một mạng nhất định. Mỗi client có thể gửi một request từ giao diện người dùng (Graphical user interface – GUI) trên màn hình, và server sẽ trả về kết quả như mong muốn, miễn là cả hai hiểu nhau.

2.3 Heroku

Heroku là nền tảng đám mây cho phép các xây dựng, triển khai, quản lý và mở rộng ứng dụng (PaaS – Platform as a service) [3, 4]. Heroku chạy các ứng dụng trong dynos – nó là một máy ảo mà có thể tăng giảm sức mạnh dựa vào độ lớn của ứng dụng. Nó rất linh hoạt và dễ sử dụng, cung cấp cho một con đường đơn giản nhất để đưa sản phẩm tiếp cận người dùng. Nó giúp các nhà phát triển tập trung vào phát triển sản phẩm mà không cần quan tâm đến việc vận hành máy chủ hay phần cứng... Heroku cho phép triển khai, chạy và quản lý các ứng dụng được viết bằng Ruby, Node.js, Java, Python, Clojure, Scala, Go và PHP.

Heroku cung cấp Database miễn phí như: ClearDB MySQL, New Relic, cung cấp các SSL miễn phí, liên kết với Github để có thể sử dụng một cách dễ dàng nhất. Ngoài ra Heroku được cung cấp 1000 giờ miễn phí khi cài đặt tài khoản thanh toán.

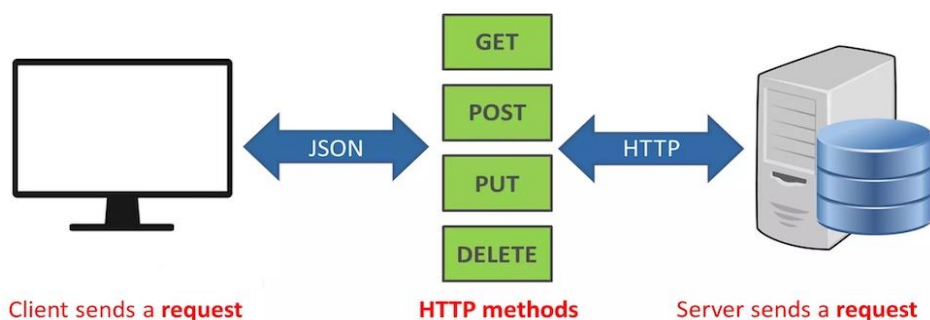


Hình 2.3.1. Mô hình deploy ứng dụng lên Heroku [4]

2.4 Restfull API

RESTful API là một trong những kiểu thiết kế API được sử dụng phổ biến hiện nay dùng cho các ứng dụng web để quản lý các resource [7]. Quy định cách sử dụng HTTP method là chức năng quan trọng nhất của REST. Một số phương thức thường được sử dụng là:

- Phương thức GET sẽ lấy một Resource hoặc danh sách các Resource từ URL cung cấp.
- Phương thức POST sẽ tạo mới một Resource từ URL cung cấp.
- Phương thức PUT sẽ cập nhật thông tin cho Resource từ URL cung cấp.
- Phương thức DELETE sẽ xóa thông tin Resource từ URL cung cấp.

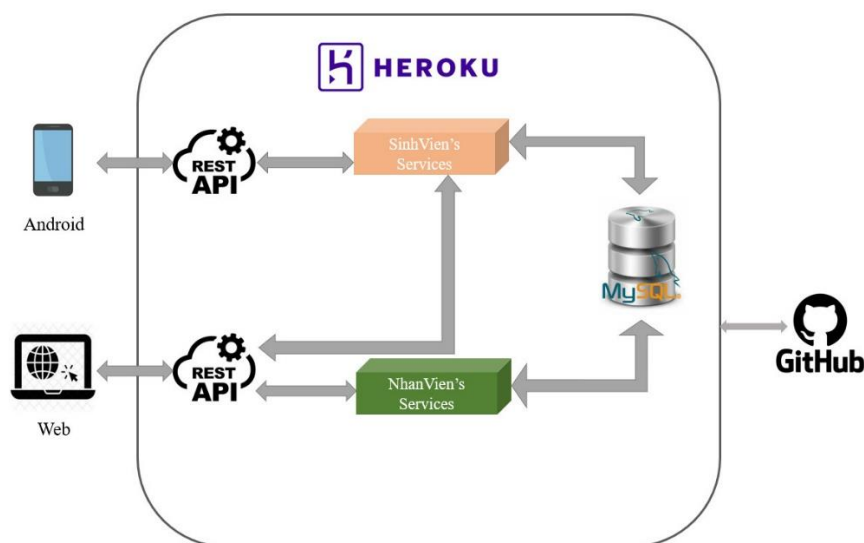


Hình2.4.1. Rest API – Giao tiếp client và server [7]

3. ATIUH: CÔNG CỤ HỖ TRỢ HỌC TẬP TẠI TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

3.1 Kiến trúc phần mềm

Micro-service là một kỹ thuật phát triển phần mềm, một biến thể của kiến trúc hướng dịch vụ (SOA), trong đó cấu trúc một ứng dụng như một tập hợp các dịch vụ được ghép với nhau. Trong kiến trúc micro-service, các dịch vụ được xử lý tốt và các giao thức rất nhẹ. Đó cũng là lý do chúng tôi đề xuất cho kiến trúc của chương trình theo mô hình Microservice. Trong kiến trúc này chúng tôi tách các service thành SinhVien Service và NhanVien Service. Đồng thời, chúng tôi cũng tổ chức dữ liệu trên Heroku và sử dụng Load Balancer, AutoScaling của Heroku để triển khai hệ thống trên nền tảng đám mây (Hình 3.1.1).



Hình 3.1.1 Kiến trúc của hệ thống hỗ trợ học tập ATIUH

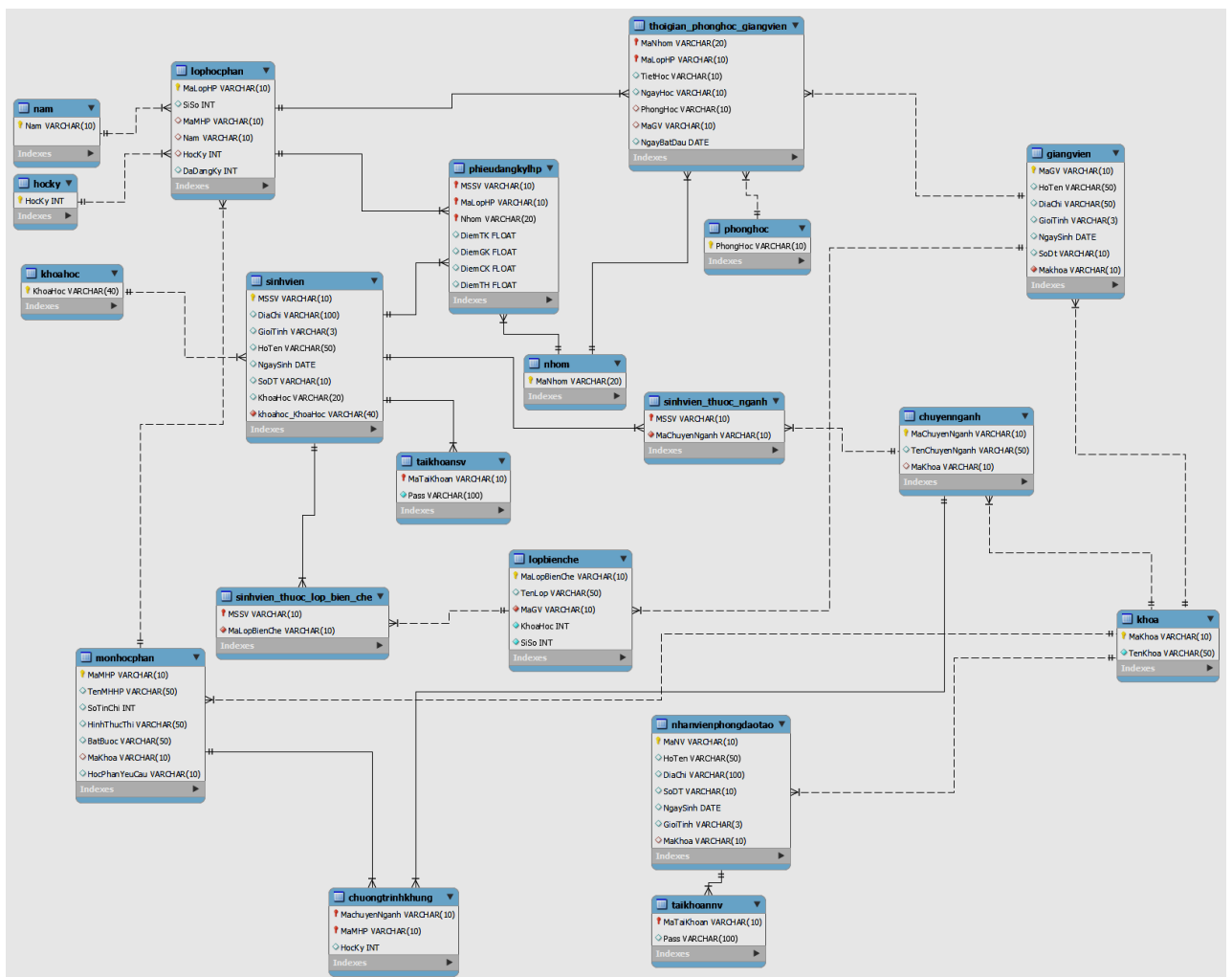
3.2 Phân tích hệ thống

Đối tượng người dùng cho hệ thống gồm có sinh viên và quản trị viên của trường. Các chức năng của từng đối tượng được mô tả trong Bảng 3.3-1. Trong đó, chúng tôi tập trung vào các chức năng chính của hệ thống mà hiện nay nhà trường đang gặp phải khó khăn như đăng ký học phần. Ngoài ra, các chức năng của nhân viên phòng đào tạo cũng được đưa vào nhằm hỗ trợ cho hệ thống đồng bộ các nghiệp vụ đào tạo và các thao tác của sinh viên.

Bảng 3.2.1 Bảng chức năng của công cụ trong hệ thống hỗ trợ học tập ATIUH

Đối tượng	Tên chức năng	Mô tả
Sinh viên	Đăng nhập	Đăng nhập vào hệ thống với tài khoản được tạo
	Cập nhật mật khẩu	Đổi mật khẩu của tài khoản sau khi đăng nhập vào hệ thống
	Xem thông tin cá nhân.	Xem thông tin cá nhân của sinh viên thuộc tài khoản đã đăng nhập
	Xem chương trình khung	Xem các môn học phần cần học theo từng học kỳ
	Xem công nợ	Xem học phí của từng môn học đã đăng ký và tổng tiền phải thanh toán của tài khoản sinh viên đã đăng nhập
	Xem kết quả học tập	Xem điểm của các môn học phần mà sinh viên đã học
	Đăng ký học phần	Gồm các chức năng như xem danh sách học phần đã đăng ký, đăng ký lớp học phần, hủy đăng ký lớp học phần
Quản trị viên	Quản lý sinh viên	cho phép thêm, xóa, sửa thông tin của sinh viên, xem danh sách sinh viên, , đặt lại mật khẩu cho sinh viên
	Quản lý giảng viên	Cho phép thêm, xóa, sửa thông tin của giảng viên, xem danh sách giảng viên
	Quản lý nhân viên	Cho phép thêm, xóa, sửa thông tin của nhân viên, xem danh sách nhân viên, đặt lại mật khẩu cho nhân viên
	Quản lý khoa	Cho phép thêm, xóa, sửa thông tin của khoa, xem danh sách khoa
	Quản lý chuyên ngành	Cho phép thêm, xóa, sửa thông tin của chuyên ngành, xem danh sách các chuyên ngành
	Quản lý lớp học phần	Cho phép thêm, xóa, sửa thông tin của lớp học phần, xem danh sách lớp học phần

Quản lý môn học phần	Cho phép thêm, xóa, sửa thông tin của môn học phần, xem danh sách môn học phần
Xếp chương trình khung	Cho phép xếp các môn đã tạo vào một chuyên ngành để tạo chương trình khung cho ngành
Chia chuyên ngành cho sinh viên	Cho phép sắp xếp sinh viên vào chuyên ngành
Xếp lịch học	Cho phép sắp xếp lịch học cho các lớp học phần đã tạo trước đó
Nhập điểm	Cho phép nhập điểm cho các sinh viên trong các lớp có sinh viên đăng ký



Hình 3.2.55: Lược đồ cơ sở dữ liệu của hệ thống hỗ trợ học tập ATIUH

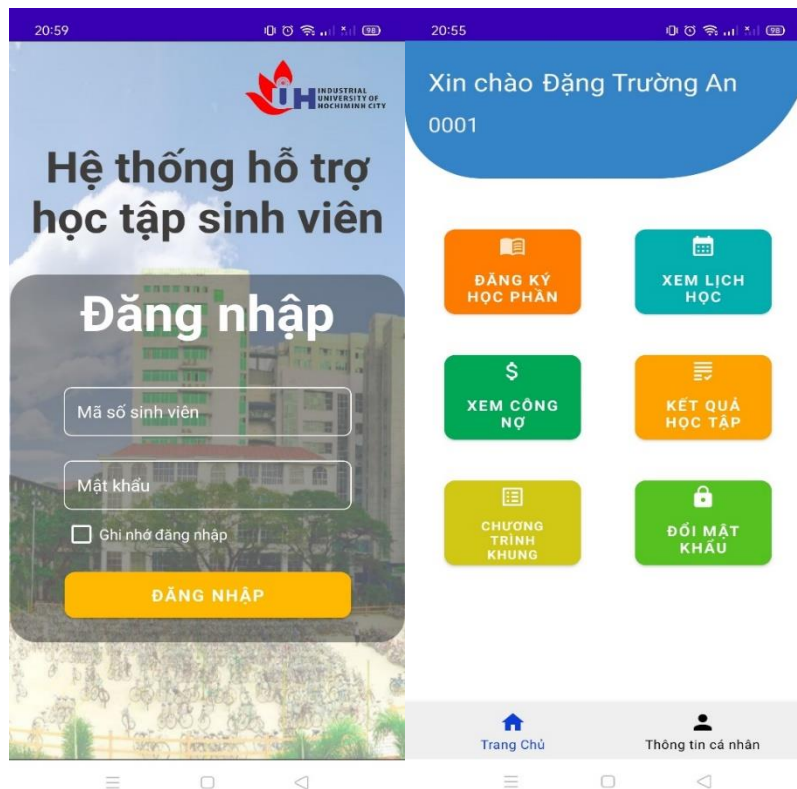
Với danh sách các chức năng trên, chúng tôi xây dựng mô hình dữ liệu thể hiện toàn bộ cấu trúc cơ sở dữ liệu của ATIUH (Hình 3.2.2).

3.3 Giao diện đồ họa

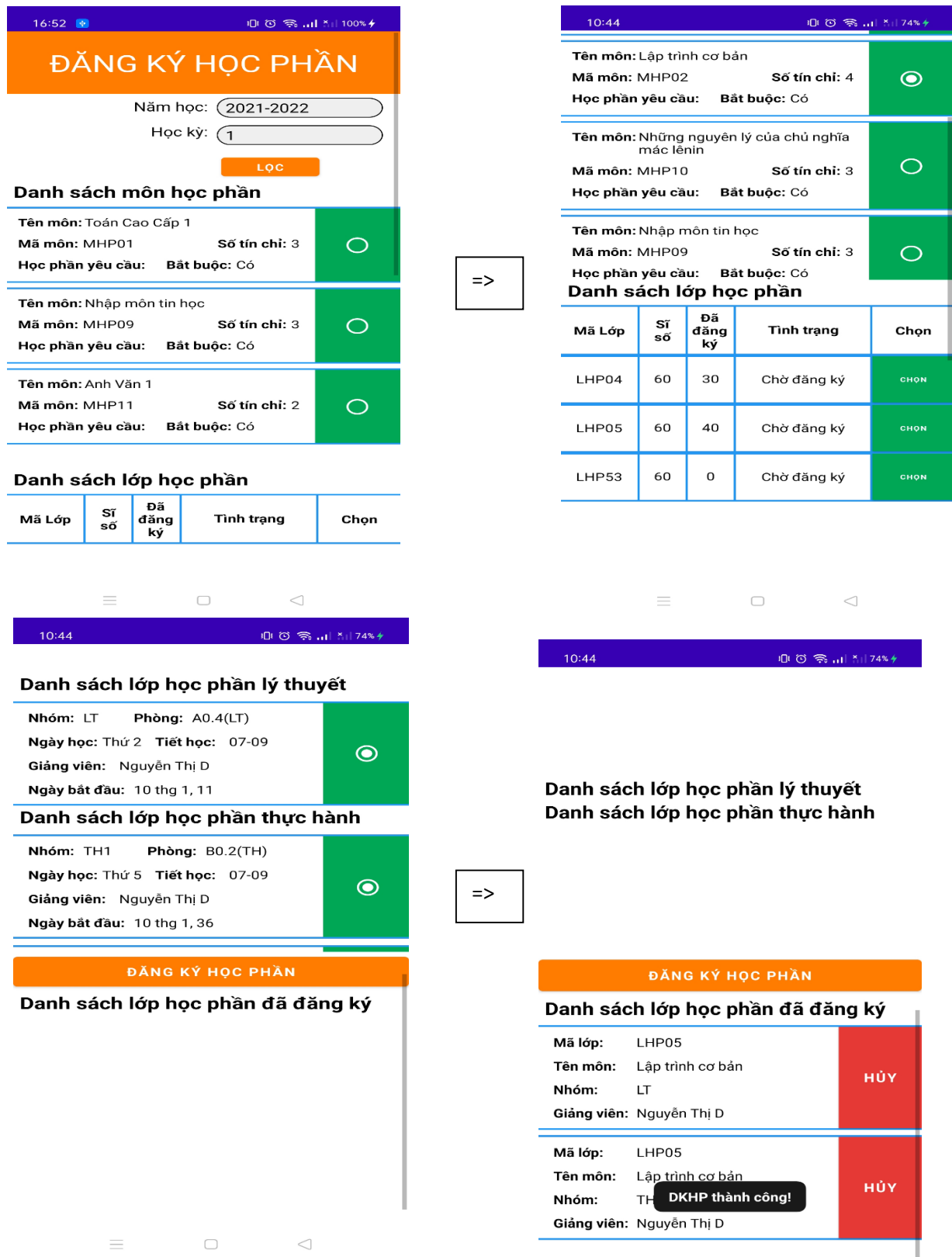
Công cụ ATIUH triển khai trên hai nền tảng là Web và App. Trong đó hệ thống App cho phép người dùng download và cài đặt trên các thiết bị chạy trên nền Android. Công cụ này sẽ có giao diện người dùng thân thiết và đảm bảo tính ổn định trên các dòng điện thoại khác nhau.

Chúng tôi cũng xin giới thiệu vài giao diện cho các chức năng trên ứng dụng trên. Đầu tiên là giao diện đăng nhập của người dùng là sinh viên vào hệ thống. Người dùng đăng nhập vào hệ thống với MSSV và

Mật khẩu do người quản lý cung cấp. Sau khi người dùng bấm đăng nhập, ứng dụng sẽ hiển thị thông báo rõ ràng đăng nhập thành công hoặc thất bại và ghi rõ lý do thất bại. Ô ghi nhớ đăng nhập để lưu trạng thái đăng nhập. Giao diện trang chủ của ứng dụng sẽ hiển thị tên sinh viên và mã số sinh viên ở trên cùng màn hình. Ở giữa màn hình là các nút nhấn để thực hiện các chức năng mở các giao diện tương ứng (Hình 3.3.1).



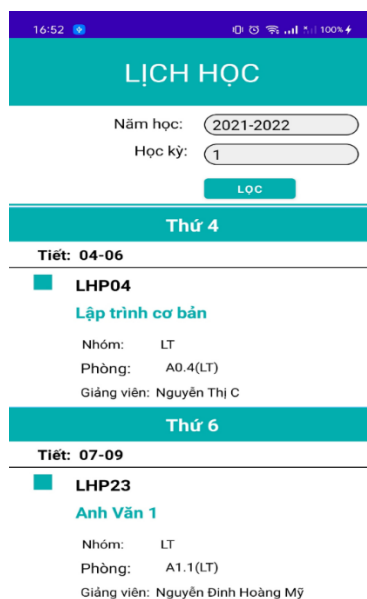
Hình 3.3.1. Giao diện chính của ATIUH trên di động



Hình 3.3.2: Giao diện đăng ký học phần ứng dụng ATIUH

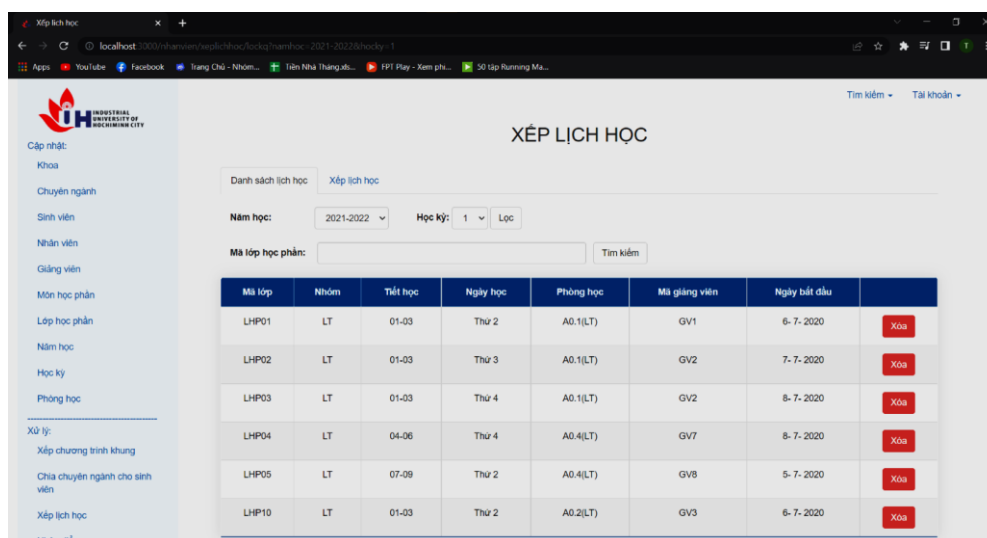
Giao diện đăng ký học phần. Giao diện này sẽ có tùy chọn năm học, tùy chọn học kỳ, nút nhấn lọc, nút nhấn đăng ký học phần, danh sách môn học phần, danh sách lớp học phần, danh sách lớp học phần lý thuyết,

danh sách lớp học phần thực hành, danh sách lớp học phần đã đăng ký. Sau khi người dùng chọn năm học và học kỳ muốn đăng ký rồi bấm lọc, ứng dụng sẽ hiển thị danh sách môn học phần và danh sách các lớp học phần đã đăng ký nếu có. Sau khi người dùng chọn một môn học phần trong danh sách môn học phần, ứng dụng sẽ hiển thị danh sách các lớp học phần tương ứng với môn học phần đã chọn nếu có. Sau khi người dùng chọn một lớp học phần, ứng dụng sẽ hiển thị danh sách lớp học phần lý thuyết và lớp học phần thực hành tương ứng với lớp học phần đã chọn nếu có. Cuối cùng, khi người dùng chọn lớp học phần lý thuyết và lớp học phần, ứng dụng sẽ hiển thị nút nhấn đăng ký học phần. Lúc này người dùng nhấn đăng ký học phần, ứng dụng sẽ hiển thị thông báo đăng ký học phần thành công hoặc thất bại và ghi rõ lý do thất bại (Hình 3.3.2)



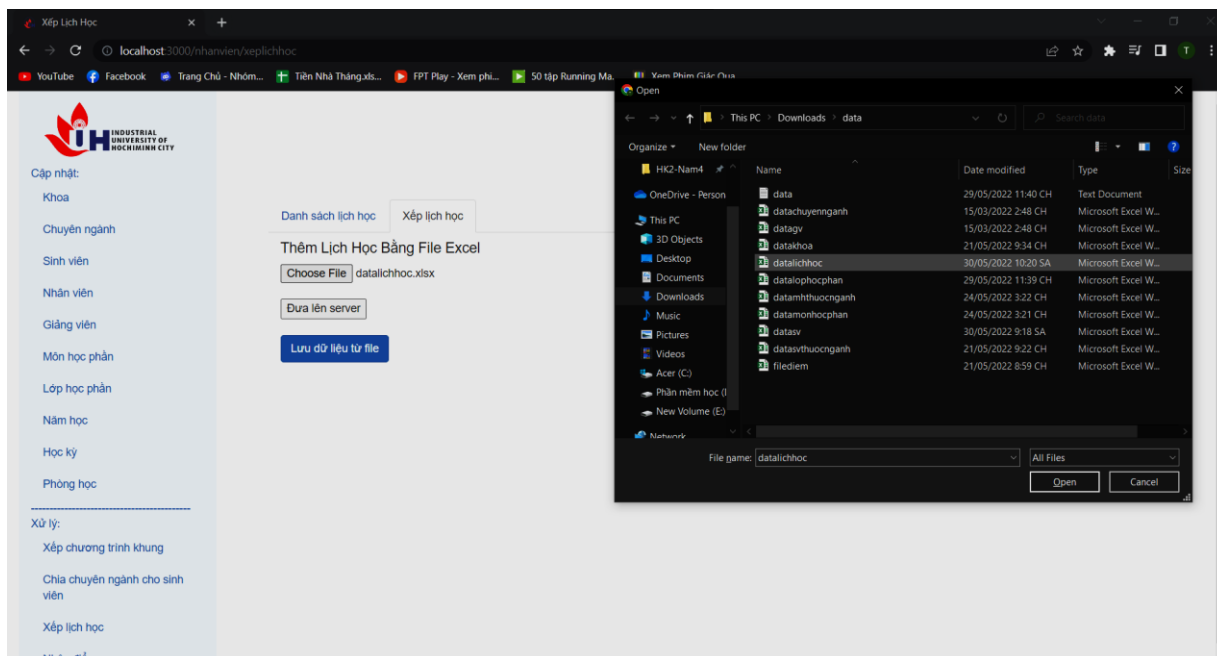
Hình 3.3.3: Giao diện xem lịch học của ứng dụng ATIUH

Giao diện xem lịch học sẽ có tùy chọn năm học, tùy chọn học kỳ, nút nhấn lọc, lịch học. Sau khi người dùng chọn năm học và học kỳ rồi bấm lọc, ứng dụng sẽ hiển thị lịch học nếu có (Hình 3.3.3). Các giao diện quản lý về phần lớp học phần. Ở các giao diện này nhân viên có thể lọc lớp học phần theo môn học phần hoặc theo từ khoá trong phần tìm kiếm. Hệ thống sẽ hiển thị đầy đủ thông tin lớp học phần. Cho phép nhân viên thêm từng lớp, thêm bằng tập tin Excel, thông tin của lớp học phần mới vào hệ thống.



Hình 3.3.4: Giao diện xếp lịch học (Web)

Giao diện xếp lịch học giúp nhân viên Phòng Đào tạo có thể xem lịch học có trong hệ thống theo năm và học kỳ đã lọc, hoặc có thể tìm kiếm thông tin lịch học của một lớp học phần thông qua phần tìm kiếm theo mã lớp (Hình 3.3.4).



Hình 3.3.5. Xếp lịch học bằng cách upload file excel (Web)

Nhân viên Phòng Đào tạo có thể xếp lịch học một cách dễ dàng và nhanh chóng thông qua file excel. Sau đó, nhân viên có thể cập nhật các thông tin đó bằng cách upload file lên server và lưu dữ liệu từ file (Hình 3.3.5).

4. KẾT LUẬN VÀ HƯỚNG PHÁT TRIỂN

Bài báo này chúng tôi đã giới thiệu được hệ thống hỗ trợ sinh viên học tập tại Trường Đại học Công nghiệp Thành phố Hồ Chí Minh (có tên gọi là *ATIUH*). Đưa ra được hai giải pháp để giải quyết bài toán đã đề ra đó là tắc nghẽn hệ thống. Giải pháp đầu tiên chúng tôi đã triển khai được ứng dụng chạy trên thiết bị di động Android. Ứng dụng này đã được triển khai thử nghiệm cho sinh viên chuyên ngành Công nghệ Phần mềm tại trường. Giải pháp thứ hai là áp dụng được điện toán đám mây vào xây dựng được ứng dụng Web. Hệ thống Web này được triển khai trên nền tảng dịch vụ điện toán đám mây Heroku. Đồng thời cũng đã kiểm chứng được tính ổn định của hệ thống khi triển khai trên cả hai nền tảng Website và Android. Hiện thực hoá thành công các chức năng cần thiết cho sinh viên (Đăng ký học phần, xem lịch học, xem điểm,...) cũng như nhân viên Phòng Đào tạo (quản lý khoa, quản lý sinh viên, xếp lịch học, chia chuyên ngành,...). Chúng tôi cũng tìm hiểu được thêm nhiều công nghệ lập trình như NodeJS, RestFull API, Java và các dịch vụ AWS, Heroku nhằm xử lý các chức năng của hệ thống. Đồng thời, để giải quyết vấn đề cân bằng tải cho các giao dịch trong hệ thống trong các giai đoạn đăng ký học phần bằng các dịch vụ Heroku. Chúng tôi đã triển khai được chương trình ra môi trường thực tế bằng công cụ Heroku với các dịch vụ đi kèm như Load Balancer và AutoScaling, đáp ứng được phần nào nhu cầu đã đề ra trước đó về quá tải. Tuy nhiên, các chức năng nghiệp vụ trong chương trình còn hạn chế như khả năng quản lý công nợ hay thanh toán trực tuyến cho hệ thống thì vẫn còn chưa được giải quyết. Trong tương lai, chúng tôi sẽ cố gắng giải quyết các hạn chế trên.

REFERENCES

1. Srivastava, P., Khan, R. A review paper on cloud computing. International Journal of Advanced Research in Computer Science and Software Engineering 8(6) (2018) 17–20

2. Mohsin, A., Janjua, N.K. A review and future directions of soa-based software architecture modeling approaches for system of systems. *Service Oriented Computing and Applications* 12(3) (2018) 183–200
3. Lee, B.H., Dewi, E.K., Wajdi, M.F. Data security in cloud computing using aes under Heroku cloud. In *Proceeding of the 27th wireless and optical communication conference (WOCC)*, IEEE (2018) 1–5
4. <https://dev.to/hookdeck/deploying-a-server-to-heroku-1e4b> accessed 10/04/2022
<https://nodejs.org/en/docs/> accessed 10/04/2022
5. Ongo, G., Kusuma, G.P. Hybrid database system of mysql and mongodb in web application development. In *Proceeding of the 2018 International Conference on Information Management and Technology (ICIMTech)*, IEEE (2018) 256–260
6. Doglio, F., Doglio, Corrigan REST API Development with Node. js. Volume 331. Springer (2018)

YSC4F.243

NGHIÊN CỨU MÔ HÌNH HỒI QUY TUYẾN TÍNH ĐỂ ĐÁNH GIÁ SỰ TƯƠNG QUAN ĐIỂM MÔN HỌC

NGUYỄN XUÂN NHỰT¹, NGUYỄN NGỌC DUNG², ĐOÀN VĂN THẮNG³

¹Khoa Công nghệ Thông tin – Trường Cao đẳng Công thương Thành phố Hồ Chí Minh,

²Khoa Công nghệ Thông tin – Trường Đại học Công nghiệp Thành phố Hồ Chí Minh;

nhutnhutnhut@gmail.com, nguyennngocdung@iuh.edu.vn, doanvanthang@iuh.edu.vn

Tóm tắt. Trong bài báo này, chúng tôi sử dụng phương pháp hồi quy tuyến tính trong lý thuyết thống kê để đánh giá sự tương quan giữa các điểm thành phần của môn học nhằm tạo điều kiện cho các giảng viên và nhà quản lý dễ dàng kiểm soát và đánh giá chất lượng dạy và học trong sinh viên.

Từ khóa. hồi quy tuyến tính, lý thuyết thống kê, gglots

STUDYING LINEAR REGRESSION MODEL TO EVALUATE THE CORRELATION OF COMPONENT SCORES IN COURSES

Abstract. In this paper, we use the linear regression method in statistical theory to evaluate the correlation between the component points of the subject in order to create conditions for lecturers and managers to easily control and evaluate the quality of teaching and learning among students.

Keywords. Linear regression, statistical theory, gglots

1. GIỚI THIỆU

Thống kê là tập hợp các phương pháp dùng để thu thập, trình bày và mô tả đặc tính của dữ liệu nhằm mục đích rút ra các quy luật chi phối các hiện tượng đang quan sát và đưa ra một quyết định.

Cơ sở của lý thuyết thống kê là lý thuyết xác suất và thống kê toán. Hiện nay thống kê được ứng dụng rộng rãi trong nhiều lĩnh vực như: thống kê dân số, thống kê xã hội, thống kê trong kinh doanh, thống kê trong y học, thống kê trong giáo dục,...

Thống kê có hai chức năng cơ bản sau:

- Thống kê mô tả (Descriptive Statistics) là các phương pháp sử dụng để tóm tắt hoặc mô tả một tập hợp dữ liệu. Đó là các phương pháp: thu thập dữ liệu, sắp xếp dữ liệu, trình bày tóm tắt dữ liệu, phân tích dữ liệu...

- Thống kê suy luận (Inferential Statistics) là các phương pháp mô hình hóa trên các dữ liệu quan sát để giải thích những biến thiên có tính ngẫu nhiên và không chắc chắn nhằm rút ra các suy diễn về quá trình, hay các đặc trưng của các đối tượng nghiên cứu. Đó là các phương pháp: ước lượng thống kê, kiểm định giả thuyết thống kê, phân tích phương sai, hồi quy tương quan, dự báo thống kê.

Phân tích hồi quy là nghiên cứu mối liên hệ phụ thuộc của 1 biến (gọi là biến phụ thuộc) vào một hay nhiều biến khác (gọi là các biến độc lập), với ý tưởng ước lượng hoặc/và dự đoán giá trị trung bình (tổng thể) của các biến phụ thuộc trên cơ sở các biến độc lập (dựa trên mẫu).

Trong bài báo này, chúng tôi áp dụng phân tích hồi quy trong lý thuyết thống kê để phân tích phổ điểm nhằm giúp cho giảng viên có một cái nhìn tổng quát khi kết thúc một học kỳ học tập hoặc kết hợp với việc khảo sát chất lượng dạy-học của môn học.

2. CÁC NGHIÊN CỨU LIÊN QUAN

Khái niệm hồi quy tuyến tính lần đầu tiên được đề xuất bởi Sir Francis Galton vào năm 1894. Hồi quy tuyến tính là một kiểm định thống kê áp dụng cho một tập dữ liệu để xác định và định lượng mối quan hệ giữa

các biến được xem xét. Các kiểm định thống kê đơn biến như Chisquare, kiểm định chính xác của Fisher, kiểm định t và phân tích phương sai (ANOVA) không cho các tính toán làm ảnh hưởng đồng biến/ yếu tố gây nhiễu khác trong quá trình phân tích (Chang 2004). Tuy nhiên, tương quan và hồi quy là các phép thử cho phép các nghiên cứu kiểm soát tác động của các yếu tố gây nhiễu trong sự hiểu biết về mối quan hệ giữa hai biến (Chang 2003).

Một số công trình gần đây phân tích về mặt lý thuyết cho các mô hình tuyến tính [Belkin và cộng sự, 2019, Hastie và cộng sự, 2019, Bartlett và cộng sự, 2019, Mu-thukumar và cộng sự, 2019, Bibas et cộng sự, 2019, Mitra, 2019, Mei và Montanari, 2019, Liang và Rakhlin, 2018, Liang và cộng sự, 2019, Xu và Hsu, 2019, Dereziński và cộng sự, 2019, Lampinen và Ganguli, 2018, Deng và cộng sự., 2019]. Ở cấp độ cao, các nghiên cứu này phân tích rủi ro thử nghiệm của các công cụ ước lượng trong hồi quy tuyến tính được phân tích ở mức quá cao với các giá trị giả định khác nhau trên các biến số.

Trong khi thực hiện nhiều thử nghiệm hồi quy tuyến tính cho biến khác nhau, phải tồn tại một số yêu cầu như hệ số cố định và nhiễu đồng biến. Breusch và A. R. Pagan [11] giải thích rằng khi các yêu cầu trên không được thỏa mãn thì việc giải thích kết quả hồi quy tuyến tính đa biến có thể không đúng.

Breusch và Pagan đưa ra hai phép hồi quy bình phương nhỏ nhất thay vì thực hiện kiểm tra tỷ lệ khả năng như một điều kiện thông thường. Tuy nhiên, nó có thể được tính bằng hai hồi quy bình phương nhỏ nhất. Loại bỏ việc sử dụng các phép tính lặp đi lặp lại là cần thiết để có được các ước tính khả năng tối đa của các hằng số trong toàn bộ mô hình được đưa ra trong bài báo.

3. MỘT SỐ KHÁI NIỆM LIÊN QUAN

3.1 Hồi quy tuyến tính đơn giản

Cho hai biến ngẫu nhiên X và Y, được quan trắc thực nghiệm bởi hai mẫu cỡ n: $X: X_1, X_2, \dots, X_n; Y: Y_1, Y_2, \dots, Y_n$.

Y có quan hệ tuyến tính với X, nếu $Y_i = \alpha + \beta X_i + \varepsilon_i, i=1,2,\dots,n$

Với: ε_i là biến ngẫu nhiên theo luật phân phối chuẩn $N(0; \sigma^2)$

α : gọi là chặn (*intercept*)

β : gọi là độ dốc (*slop hay gradient*)

Các hệ số này được ước tính từ dữ liệu. Phương pháp ước tính là phương pháp bình phương bé nhất (*least squares method*). Phương pháp này tìm α, β để $\sum_{i=1}^n [y_i - (\alpha + \beta x_i)]^2$ đạt giá trị bé nhất

Khi đó

$$\hat{\beta} = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sum_{i=1}^n (x_i - \bar{x})^2}, \hat{\alpha} = \bar{y} - \hat{\beta} \bar{x}$$

Chú ý: $\hat{\alpha}, \hat{\beta}$ là các ước lượng xấp xỉ của α, β . Từ $\hat{\alpha}, \hat{\beta}$ ta có $\hat{y}_i = \hat{\alpha} + \hat{\beta} x_i$, khi đó đại lượng $(y_i - \hat{y}_i)$ gọi là phần dư (*residual*). Phương sai của phần dư có thể ước lượng bằng.

$$s^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n - 2}$$

Cho ví dụ: xét dữ liệu nghiên cứu độ Cholesterol trong máu của 18 đối tượng nam như sau (BMI: tỷ số trọng lượng (kg) với chiều cao bình phương (cm²). Ước tính hệ số tương quan giữa độ tuổi và Cholesterol

id	age	bmi	chol
1	46	25.4	3.5
2	20	20.6	1.9
3	52	26.2	4.0
4	30	22.6	2.6
5	57	25.4	4.5
6	25	23.1	3.0
7	28	22.7	2.9
8	36	24.9	3.8
9	22	19.8	2.1
10	43	25.3	3.8
11	57	23.2	4.1

12	33	21.8	3.0
13	22	20.9	2.5
14	63	26.7	4.6
15	40	26.4	3.2
16	48	21.2	4.2
17	28	21.2	2.3
18	49	22.8	4.0

Để phân tích hồi quy tuyến tính đơn giản cho hai đại lượng age, chol ta có:

```
> lm(chol~age)
```

Call:

```
lm(formula = chol ~ age)
```

Coefficients:

```
(Intercept)      age
    1.08922      0.05779
```

Trong kết quả này, mô tả chol là một hàm số của age, với $\hat{\alpha} = 1.0892; \hat{\beta} = 0.05779$, nghĩa là ta có phương trình tuyến tính $\hat{y}_i = 1.08922 + 0.05779 * \hat{x}_i$

3.2 Hồi quy tuyến tính đa biến

Trong trường hợp tổng quát, một biến ngẫu nhiên Y có thể phụ thuộc tuyến tính với nhiều biến ngẫu nhiên X_1, X_2, \dots, X_k . Khi đó có thể biểu diễn dưới dạng $Y = \alpha + X\beta + \varepsilon$

Với

$$Y = \begin{bmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{bmatrix}, X = \begin{bmatrix} x_{11} & x_{21} & \dots & x_{k1} \\ x_{12} & x_{22} & \dots & x_{k2} \\ \dots & \dots & \dots & \dots \end{bmatrix}, \beta = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \dots \\ \beta_n \end{bmatrix}, \alpha = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \dots \\ \alpha_n \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \dots \\ \varepsilon_n \end{bmatrix}$$

Để tìm các giá trị xấp xỉ $\hat{\alpha}, \hat{\beta}$ cũng dựa vào phương pháp bình phương bé nhất tương tự như trường hợp hồi quy tuyến tính đơn giản (*một biến*).

Tính toán các giá trị $\hat{\alpha}, \hat{\beta}$ cũng sử dụng câu lệnh trong ngôn ngữ R tương tự hồi quy tuyến tính đơn giản: `lm()`, `summary()`

Ví dụ, tính toán $\hat{\alpha}, \hat{\beta}$ ứng với trường hợp tìm quan hệ tuyến tính giữa chol với age và bmi

```
> mreg<-lm(chol~age+bmi)
> summary(mreg)

Call:
lm(formula = chol ~ age + bmi)

Residuals:
    Min       1Q   Median       3Q      Max
-0.3762 -0.2259 -0.0534  0.1698  0.5679

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  0.455458   0.918230   0.496   0.627
age          0.054052   0.007591   7.120 3.5e-06 ***
bmi          0.033364   0.046866   0.712  0.487
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.3074 on 15 degrees of freedom
Multiple R-squared:  0.8815,    Adjusted R-squared:  0.8657
F-statistic: 55.77 on 2 and 15 DF,  p-value: 1.132e-07
```

Với kết quả tính toán ta có:

$$Chol = 0.455458 + 0.054052 * age + 0.033364 * bmi$$

4. PHÂN TÍCH ĐÁNH GIÁ SỰ TƯƠNG QUAN GIỮA CÁC ĐIỂM THÀNH PHẦN CỦA MÔN HỌC

Để phân tích phổ điểm, trong bài báo này chúng tôi xin giới thiệu cấu trúc cơ sở dữ liệu như sau.

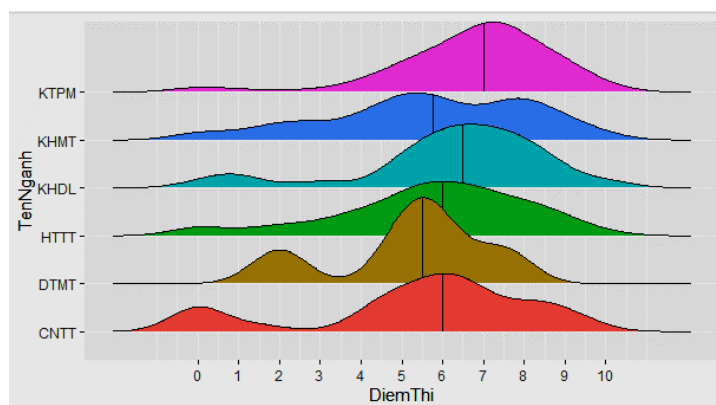
Bảng 1: Cấu trúc cơ sở dữ liệu minh họa để phân tích.

TT	MaSV	Lop	TenMon	Khoa	Nganh	DQT	DiemThi	DiemKT
1	2113011213	DT13SMN01	HeCSDL	CNTT	HTTT	1.8	8.5	8.5
2	2113011224	DT13SMN01	HeCSDL	CNTT	KHMT	2	8.5	8.5
3	2114012741	DT14SSH01	HeCSDL	CNTT	KTPM	1.6	8	8
4	2112011007	DT12CTT01	HeCSDL	CNTT	HTTT	8	8.5	8.5
5	2112011025	DT12CTT01	HeCSDL	CNTT	KHMT	9	9.5	9.5
...

4.1 Phân bố tần suất điểm thi

Biểu đồ phân bố thường dùng là biểu đồ phân bố tần suất, bao gồm trục x chỉ các loại điểm số và trục y là tần suất (số lần xuất hiện giá trị của x) hoặc dùng xác suất (tính phần trăm trên kích thước mẫu). Với biểu đồ này, ta so sánh được điểm số nào xuất hiện nhiều nhất hoặc thấp nhất, dữ liệu tập trung hay phân tán để đánh giá chất lượng. Để tiện theo dõi trực quan sự biến động của điểm số ta có thể quan sát hình dáng của biểu đồ theo đường cong nối các đỉnh của các cột.

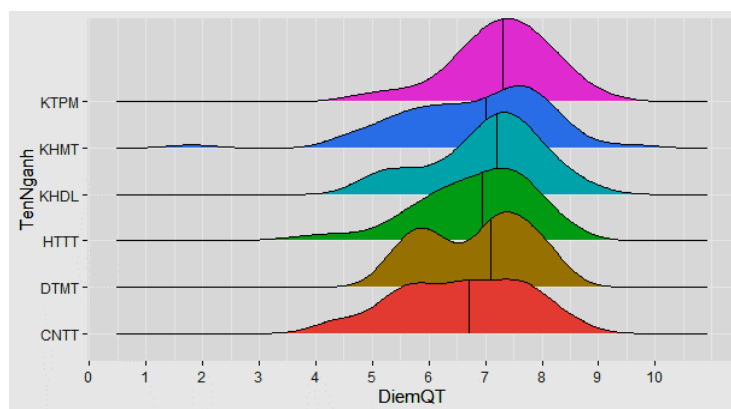
4.1.1 Điểm thi



Hình 1: Phân bố điểm thi.

Đánh giá: Điểm thi được phân bố rộng từ 0 ->10, có sự phân hóa tương đối cao các đối tượng. Điểm thi tập trung từ điểm 7-> điểm 10, trong đó điểm 7.0 có tần suất cao nhất.

4.1.2 Điểm quá trình



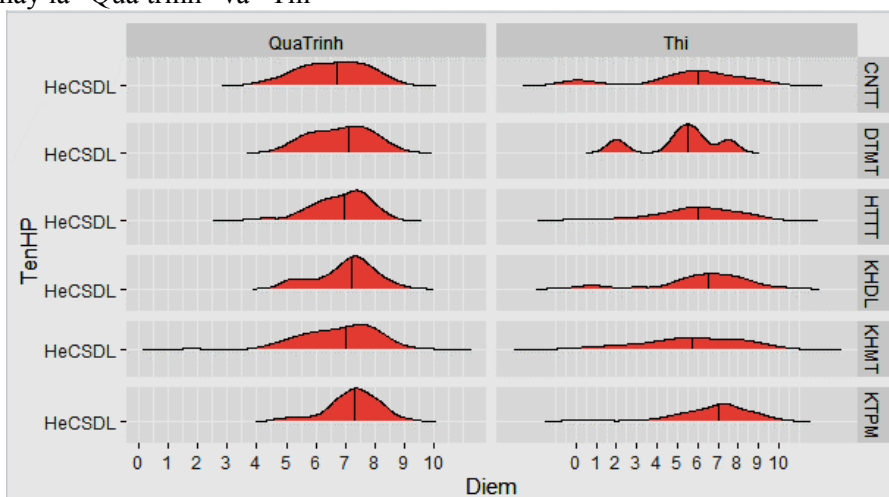
Hình 2: Phân bố điểm quá trình.

Đánh giá: Điểm quá trình được phân bố rộng từ 0 ->10, có sự phân hóa tương đối cao các đối tượng. Điểm quá trình tập trung từ điểm 7-> điểm 10, trong đó điểm 7.0 có tần suất cao nhất.

4.1.3 Kết hợp điểm thi và điểm quá trình

Đây là loại biểu đồ kết hợp hai phân bố trên cùng một biểu đồ. Cách trình bày này tạo điều kiện cho việc so sánh sự phân bố điểm thi và điểm quá trình.

Tuy nhiên, cách biểu diễn này phải thay đổi về cơ sở dữ liệu, phải thêm một cột chung là loại điểm có giá trị của cột này là “Quá trình” và “Thi”



Hình 3: Phổ điểm kết hợp.

Đánh giá:

Phổ điểm thi có phân bố rộng hơn so với điểm quá trình, phân hóa đối tượng sinh viên tốt hơn, số lượng sinh viên có điểm quá trình lớn hơn điểm thi chiếm số lượng rất lớn, đặc biệt điểm quá trình cao hơn điểm thi cao một cách bất thường.

4.2 Phân tích phổ điểm theo hồi quy tuyến tính

Tương quan giữa điểm quá trình và điểm thi toàn mẫu

Đây là loại biểu đồ điểm (Scatter) biểu diễn mối tương quan giữa 2 đại lượng, điểm quá trình và điểm thi, mỗi điểm trên hệ tọa độ là một đối tượng sinh viên. Cũng trên biểu đồ này sẽ biểu diễn xu hướng tăng giảm theo sự phụ thuộc của 2 đại lượng (gọi là sự tương quan giữa 2 đại lượng). Mối tương quan này được đánh giá theo hệ số r (Hệ số tương quan Pearson), theo kinh nghiệm giá trị r biểu diễn sự tương quan như sau:

$$-1 \leq r \leq 1$$

$r > 1$: Khi x tăng, y tăng

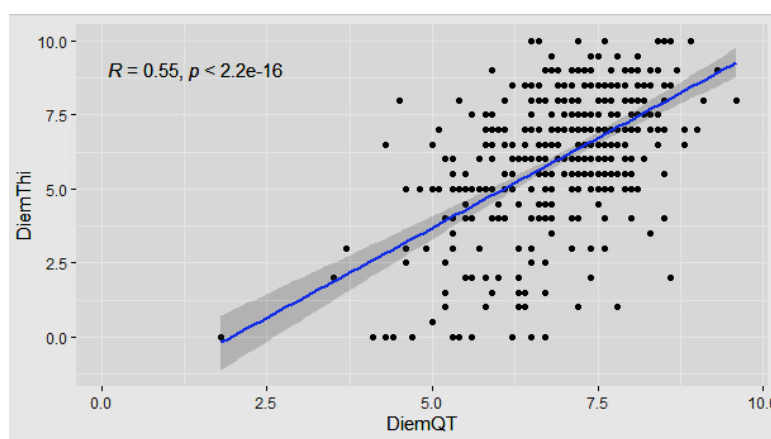
$r > -1$: Khi x tăng, y giảm

$r > 0.8$: Tương quan mạnh

$r = 0.4 - 0.8$: Tương quan trung bình

$r < 0.4$: Tương quan yếu

$r > 0$: Không có mối liên hệ nào



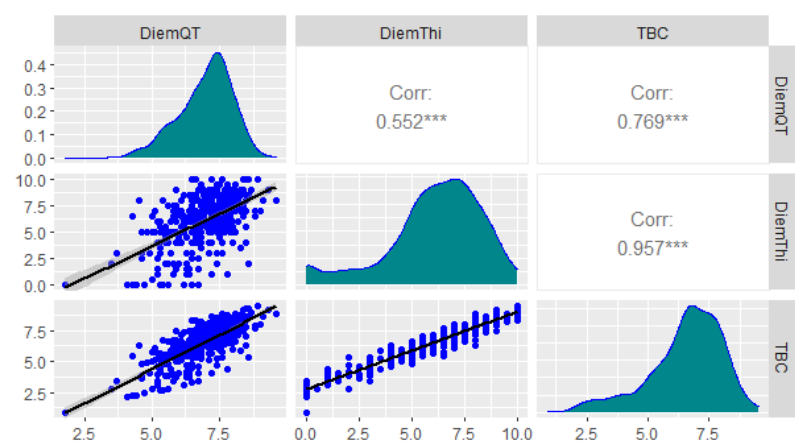
Hình 4: Hệ số tương quan giữa điểm thi và điểm quá trình.

Đánh giá: Hệ số tương quan giữa điểm quá trình và điểm thi rất thấp, chứng tỏ rằng chưa có mối liên hệ nào giữa điểm quá trình và điểm thi, chưa khẳng định được điểm quá trình cao thì điểm thi cao và ngược lại.

Mối quan hệ giữa Điểm quá trình, Điểm thi và Điểm học phần

Đây là đa biểu đồ, phối hợp nhiều loại biểu đồ trên cùng một biểu đồ, tạo điều kiện đánh giá đa chiều hơn về phân phối tần suất, sự tương quan, hệ số tương quan.

Biểu đồ này nhằm đánh giá tổng hợp từ các loại biểu đồ đã phân tích ở trên. Nhìn vào biểu đồ để đánh giá về phân phối tần suất của điểm quá trình, điểm thi và điểm học phần và mối tương quan giữa các loại điểm này.



Hình 5: Hệ số tương quan giữa điểm thi, điểm quá trình, và trung bình chung.

Đánh giá: Hệ số tương quan giữa điểm thi và điểm học phần rất cao, điều này cũng hiển nhiên vì điểm thi có trọng số cao trong điểm học phần.

5. KẾT LUẬN

Kết quả học tập của sinh viên được giảng viên chấm điểm thông qua các cột điểm số: điểm quá trình, điểm thi và điểm tổng kết môn học. Qua phân tích nhận thấy rằng việc đánh giá ở điểm quá trình là quá cao và dữ liệu tập trung (phổ điểm nhọn, không phân hóa được chất lượng sinh viên). Trong khi đó điểm thi lại thấp hơn nhiều so với điểm quá trình, phổ điểm thi rộng hơn, phân hóa chất lượng sinh viên tốt hơn so với điểm quá trình. Một vấn đề nữa trong phân tích phổ điểm cho thấy chưa có mối liên hệ nào giữa điểm quá trình và điểm thi, bởi lẽ thông thường sinh viên có điểm quá trình tốt thì điểm thi sẽ tốt.

TÀI LIỆU THAM KHẢO

- [1] Babak Shahbaba (2012), *Biostatistics with R*, Springer.
- [2] Bartlett, P. L., Long, P. M., Lugosi, G., and Tsigler, A. (2019). Be-nign overfitting in linear regression. arXiv preprint arXiv:1906.11300.
- [3] Belkin, M., Hsu, D., and Xu, J. (2019). Two models of double de-scent for weak features. arXiv preprint arXiv:1903.07571.
- [4] Bibas, K., Fogel, Y., and Feder, M. (2019). A new look at an old problem: A universal learning approach to linear regression. arXiv preprint arXiv:1905.04708.
- [5] Breusch, T. S., & Pagan, A. R. (1979). A simple test for heteroscedasticity and random coefficient variation. *Econometrica: Journal of the Econometric Society* , 1287 - 1294
- [6] Dương Thiệu Tống (2000), *Thống kê ứng dụng trong nghiên cứu khoa học giáo dục*. Nhà xuất bản đại học Quốc gia Hà Nội.
- [7] Dereziński, M., Liang, F., and Mahoney, M. W. (2019). Exact expressions for double descent and implicit regularization via surrogate random design.
- [8] G.Jay Kens (2010), *Introduction to Probability and Statistics Using R*, First Edition.
- [9] Hadley Wickham (2015), *Ggplot2: Elegant Graphics for Data Analysis*. Springer pub-lisher.
- [10] Hoàng Trọng & Chu Nguyễn Mộng Ngọc (2005), *Phân tích dữ liệu nghiên cứu với SPSS*. Nhà xuất bản Thống kê.
- [11] Mei, S. and Montanari, A. (2019). The generalization error of random features regression: Precise asymptotics and double descent curve. arXiv pre-print arXiv:1908.05355.
- [12] Nguyễn Đức Thuần, Đoàn Văn Thắng (2016), *Thống kê máy tính*, Nhà xuất bản Thông tin truyền thông.
- [13] Nguyễn Văn Tuấn (2018), *Phân tích dữ liệu với R nghiên cứu với SPSS*. Nhà xuất bản tổng hợp TP Hồ Chí Minh.
- [14] Nguyễn Văn Tuấn, *Phân tích số liệu và biểu đồ bằng R*.

YSC4F.244

GIẢI PHÁP XÁC THỰC NGƯỜI DÙNG THÔNG QUA MÃ OTP SỬ DỤNG SINH TRẮC HỌC GIỌNG NÓI

HÀ BẢO ANH, ĐẶNG THỊ PHÚC

*Khoa Công nghệ Thông tin, Trường Đại học Công nghiệp Thành phố Hồ Chí Minh
baoanhcr7123456@gmail.com, phucdt@iuh.edu.vn*

Tóm tắt: Sinh trắc học giọng nói là một công nghệ dựa vào việc nhận dạng các mẫu giọng để xác minh danh tính của một người. Việc sử dụng mật khẩu một lần (One-Time Password – OTP) một loại xác thực đã quen thuộc với tất cả chúng ta, đặc biệt là những người hay giao dịch, thanh toán online hoặc xác thực bảo mật hai lớp, nhưng đa phần người dùng phải nhập liệu bằng tay. Nghiên cứu này giới thiệu giải pháp xác thực người dùng thông qua mã OTP sử dụng sinh trắc học giọng nói của chính người dùng đó với mục tiêu nâng cao mức độ bảo mật, bảo vệ người dùng khỏi nhiều mối nguy hại tiềm ẩn, giảm thiểu hành vi giả mạo trong giao dịch ngoài ra có thể ứng dụng ở nhiều lĩnh vực khác trong kỷ nguyên chuyển đổi số.

Từ khóa: Sinh trắc học giọng nói, xác thực người dùng, xác thực người nói, mật khẩu một lần

USER AUTHENTICATION USING A ONE-TIME PASSWORD AND VOICE BIOMETRICS

Abstract: Voice biometrics is a technology that relies on recognizing voice patterns to verify a person's identity. The use of One Time Password (OTP), a type of authentication, is familiar to all of us, especially those who often transact, pay online, or authenticate with two-factor authentication, but the most user must enter data manually. This study introduces a solution to authenticate users through OTP code by using that user's voice biometrics to improve security, protect users from possible risks, and at the same time reduce fraud in transactions. This research can be applied to many other fields in the age of digital transformation.

Keywords: Voice biometrics, user authentication, speaker verification, one-time password

1. GIỚI THIỆU

Trong thời đại phát triển mạnh về khoa học và công nghệ hiện đây vấn đề bảo mật thông tin trong các giao dịch, thanh toán online hoặc xác thực người dùng luôn được quan tâm tìm ra các giải pháp công nghệ nâng cao mức độ bảo mật như bảo mật vân tay, móng mắt hoặc bảo mật bằng nhận diện khuôn mặt ngoài ra một giải pháp có khả năng xác minh danh tính người dùng thông qua đặc điểm giọng nói cá biệt của mỗi người là sinh trắc học giọng nói. Sinh trắc học giọng nói là một công nghệ sử dụng giọng nói của một người nó dựa vào các đặc điểm riêng biệt về tone giọng, cao độ, cường độ, đặc trưng chất giọng làm đặc điểm sinh học để xác thực họ, còn được gọi là xác minh giọng nói hoặc nhận dạng người nói. Ở nghiên cứu này của chúng tôi chia ra thành hai bài toán xử lý, một là bài toán xác thực người nói chúng tôi sử dụng mô hình pretrained “X-vectors”[1] để lấy ra các vector embedding giọng nói thông qua đó huấn luyện các mô hình học máy nhận dạng người nói, hai là bài toán phân loại âm thanh ở bài toán này chúng tôi áp dụng kỹ thuật trích xuất đặc trưng Mel frequency cepstral coefficients (MFCC) thực hiện biến đổi âm thanh đầu vào về thang đo tần số Mel sau đó sử dụng kiến trúc mạng Convolutional Neural Networks (CNN) AlexNet [2] để phân lớp các có nhãn là từ 1 đến 5 ứng với đó là các số sẽ có trong mật khẩu một lần (OTP) sau đó tiến hành so khớp với mã OTP được sinh ra nếu trùng khớp thì quá trình xác thực người dùng thông qua mã OTP sử dụng sinh trắc học giọng nói đã hoàn thành ngược lại nếu mã OTP người nói cung cấp khác so với hệ thống gửi ra thì quá trình xác thực không hoàn thành.

2. CÁC NGHIÊN CỨU LIÊN QUAN

Sinh trắc học giọng nói (Voice Biometric) hiện nay đã được sử dụng tương đối phổ biến trên thế giới, vì những ưu điểm mà công nghệ này mang lại. Nó thường được sử dụng trong các ứng dụng Internet Banking sử dụng song song với bảo mật vân tay để cung cấp thêm nhiều lựa chọn cho khách hàng ngoài ra cũng được dùng trong tổng đài Call center, đây là ứng dụng đột phá trong ngành Call center mà chỉ những phần mềm Call center tiên tiến mới đạt được. Khi tích hợp Voice Biometric, khách hàng gọi đến tổng đài mà không cần phải cung cấp câu hỏi bảo mật, mã PIN, hoặc số chứng minh thư cho tổng đài viên, thay vào đó mọi thứ được diễn ra nhanh chóng bởi sự đảm nhận của Robot. Giúp tăng vượt trội về khả năng bảo mật. Các hệ thống nhận dạng người nói dựa trên mạng nơ-ron sâu (DNN) trước đây thường sử dụng các phương pháp pooling đã thành công cho các tác vụ nhận dạng hình ảnh, chẳng hạn như average pooling [3, 4] hoặc các lớp được kết nối đầy đủ (fully connected layers) [5, 6] để cô đọng thông tin ở cấp khung thành các biểu diễn ở mức phát âm. Một số công trình gần đây đã đề xuất mang lại các phương pháp học sâu tương tự để nhận dạng người nói [7, 8, 9, 10, 11].

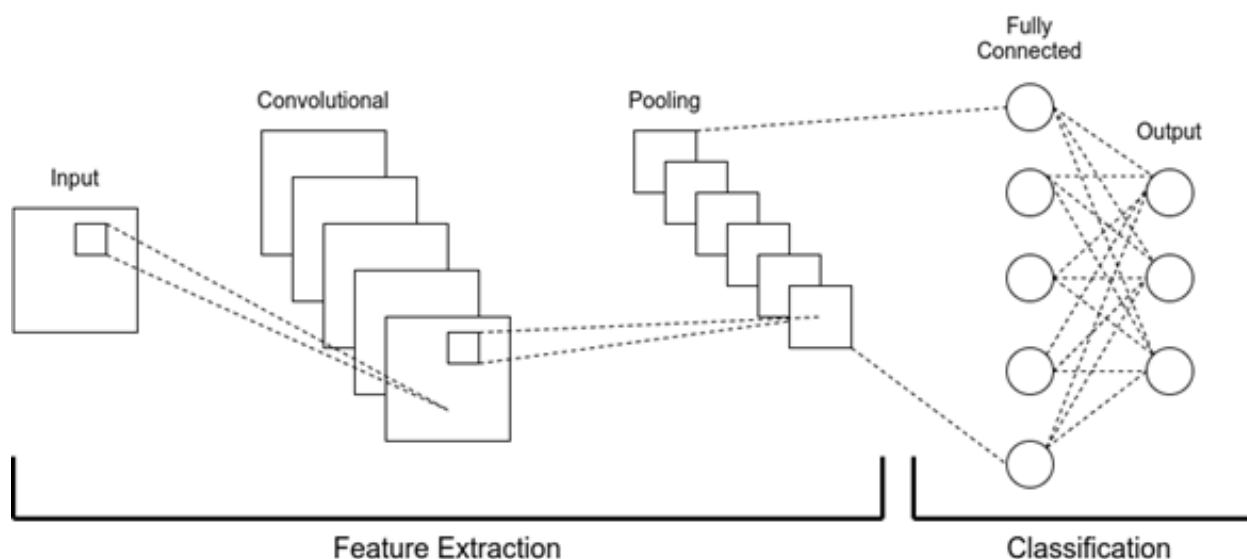
3. PHƯƠNG PHÁP TIẾP CẬN

3.1. Dữ liệu

Trong nghiên cứu này chúng tôi sử dụng bộ dữ liệu người nói được trích một phần từ tập dữ liệu Vietnamese Language and Speech Processing (VLSP) 2021 trong Hội thảo quốc tế lần thứ 8 được tổ chức tại Trường Đại học Công nghệ Thông tin, ĐHQG Tp. Hồ Chí Minh [12]. Dữ liệu trích rút gồm 748 người nói với 13988 tập tin âm thanh, ngoài ra chúng tôi còn thu thập giọng nói từ 24 sinh viên khóa 15 ngành Khoa học dữ liệu Trường Đại học Công nghiệp Thành phố Hồ Chí Minh. Dữ liệu thu thập gồm 24 lớp, mỗi lớp bao gồm 5 tập tin. Mỗi tập tin âm thanh kéo dài 1 phút, chỉ chứa giọng đọc của một người về một số cố định. Đối với mỗi tập tin như vậy, chúng tôi tách ra thành từng tập tin nhỏ (là một đoạn âm thanh chứa một từ). Vì dữ liệu do nhiều người khác nhau thu thập, ở các môi trường khác nhau nên chúng ta cần loại bỏ bớt dữ liệu nhiễu (noise), chuyển đổi tần số lấy mẫu (sample rate) về 16kHz và chuyển đổi số kênh (channel) về 1 kênh (mono). Các bước tiền xử lý dữ liệu âm thanh trên được chúng tôi sử dụng phần mềm WavePad để thực hiện. Dữ liệu sau khi qua bước tiền xử lý thu được 7000 tập tin âm thanh đã được gán nhãn đầy đủ.

3.2. Mô hình phân loại số

Mục tiêu của mô hình là phân loại ra 5 lớp (1, 2, 3, 4, 5) từ tập dữ liệu tự thu thập. Thông thường khi phân tích một tín hiệu bất kỳ, người ta hay sử dụng biến đổi Fourier để biểu diễn tín hiệu trong miền phổ. Tuy nhiên, tai nghe con người lại không có sự nhạy cảm như nhau với mọi dải tần số, nó cảm nhận kém hơn ở những tần số cao hơn 1000hz. Do đó việc mô hình hóa tính chất này của tai người làm tăng khả năng nhận dạng của hệ thống. Một trong những cách phổ biến nhất là sử dụng thang đo Mel, đây cũng là thang đo nhận thức về cao độ (pitch). Mỗi mel là một đơn vị của cao độ được xác định sao cho với các đoạn tín hiệu tiếng nói tương đương về mặt cao độ (những đường biểu diễn pitch của chúng giống nhau), thì cũng được phân tách bằng số lượng đơn vị mel như nhau. Ở bài toán này chúng tôi ánh xạ trực y tần số (frequency) lên thang mel để tạo thành mel spectrogram (dạng hình ảnh). Sau đó sử dụng mô hình mạng Neural Network AlexNet để phân lớp. Kiến trúc cơ bản của mạng CNN bao gồm hai phần chính. Một là, phần phân tích và nhận dạng các đặc điểm khác nhau của một ảnh trong quá trình xử lý được gọi là trích xuất đặc trưng (Feature Extraction). Hai là, lớp được kết nối đầy đủ (fully connected - FC) sử dụng đầu ra của quá trình tích chập và dự đoán hình ảnh dựa trên các đặc trưng được trích xuất ở giai đoạn trước được gọi là phân loại (Classification). Hình 1 dưới đây mô tả kiến trúc cho mô hình phân loại số.

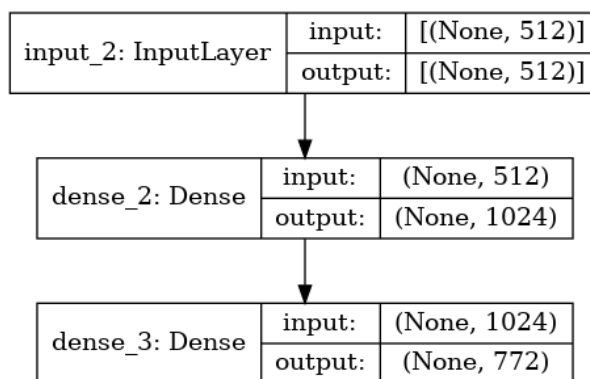


Hình 1. Kiến trúc mô hình phân loại số

Kiến trúc CNN được cấu thành từ ba layer đó là convolutional layer, pooling layer và fully connected layer (FC). Các layer này được xếp chồng lên nhau thành một kiến trúc CNN. Ngoài ba layer này còn có hai tham số quan trọng đó là dropout và hàm kích hoạt (activation function). Mô hình AlexNet đã tham gia cuộc thi ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC 2012) được gọi với tên phổ biến là AlexNet theo tên của tác giả đầu tiên là Alex Krizhevsky. (A. Krizhevsky, I. Sutskever, G.E. Hinton, 2012). Mạng AlexNet có khoảng 56 triệu tham số và 650.000 neurons bao gồm 5 lớp Convolution, 3 Fully connected Layers ngoài ra vì đầu vào của dữ liệu có kích thước không cố định nên chúng tôi đã sử dụng thêm 1 lớp Adaptive average pooling cho bất kì kích thước đầu vào nào cũng sẽ thích ứng và còn lại chỉ cần xác định đầu ra. Cuối cùng là 1 hàm softmax cho ra xác suất của ảnh thuộc về 5 lớp của bộ dữ liệu.

3.3. Mô hình xác thực người nói

Mục tiêu phân loại được 772 lớp, trong đó có 24 lớp là dữ liệu tự thu thập và số còn lại là dữ liệu VLSP 2021. Toàn bộ dữ liệu sau khi trích xuất đặc trưng được chia ra thành 3 tập huấn luyện, phát triển, thử nghiệm theo tỷ lệ 6:2:2. Kiến trúc mô hình phân loại người nói sử dụng lớp đầu vào là 1 vector embedding có kích thước (512,). Sau đó dữ liệu đi qua lớp FC với 1024 nodes, sử dụng hàm kích hoạt là relu. Sau đó là 1 lớp FC tiếp với đầu ra 772 lớp và hàm kích hoạt là softmax để đưa ra xác suất. Hình 2 dưới đây là kiến trúc mạng và mã hiện thực bằng tensorflow của mô hình xác thực người nói.



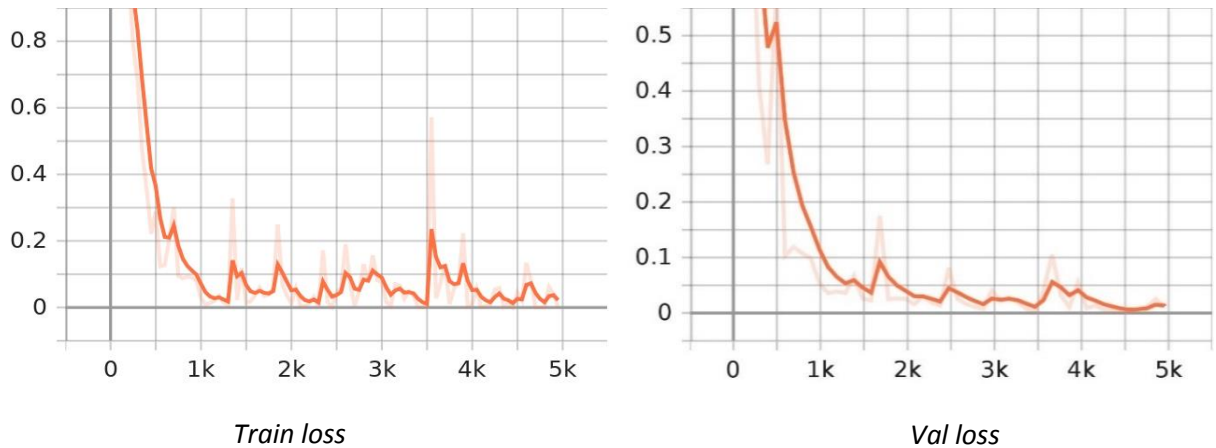
Hình 2. Kiến trúc mạng của mô hình xác thực người nói

Ngoài dùng mô hình Deep Learning trên, chúng tôi còn sử dụng thêm các mô hình machine learning và deep learning khác. Bao gồm Extreme Learning Machine, Logistic Regression, Support Vector Machine, K-Nearest Neighbors, Decision Tree, Random Forest, Perceptron. Các mô hình này để thực nghiệm cho việc xác thực người nói. Kết quả sẽ được trình bày trong phần đánh giá.

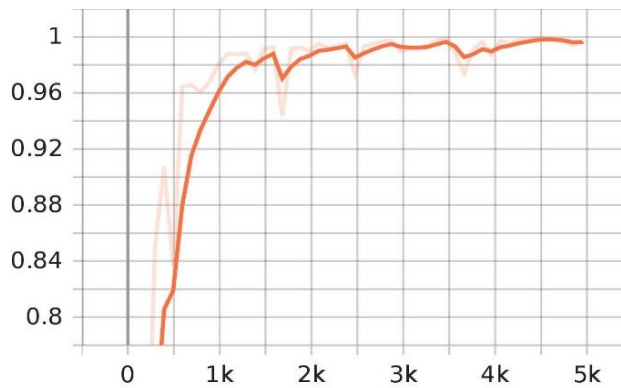
4. ĐÁNH GIÁ

4.1. Mô hình phân loại số

Đánh giá hiệu quả mô hình trên tập validation và tập test.



Hình 3. Biểu đồ mất mát trên dữ liệu huấn luyện và kiểm thử

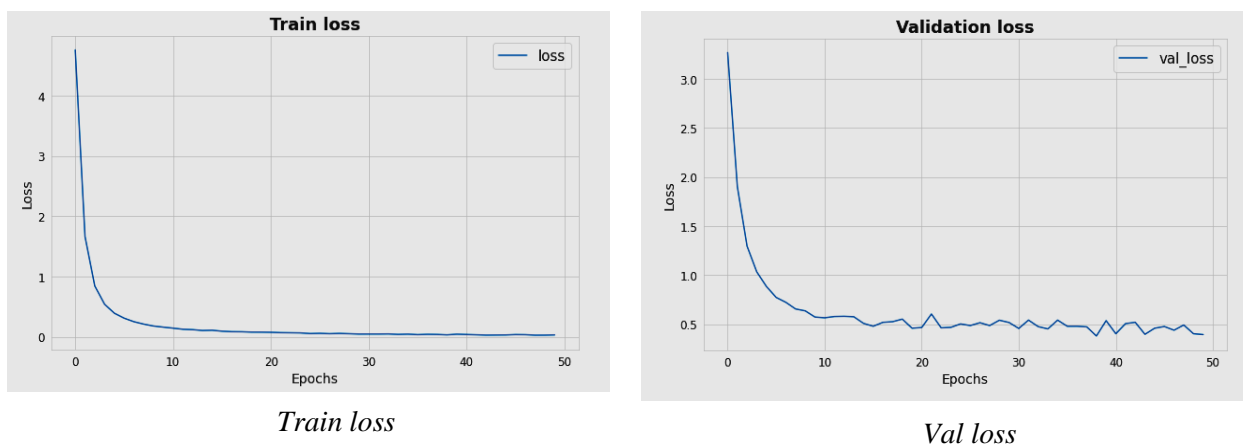


Hình 4. Biểu đồ độ chính xác trên tập kiểm thử

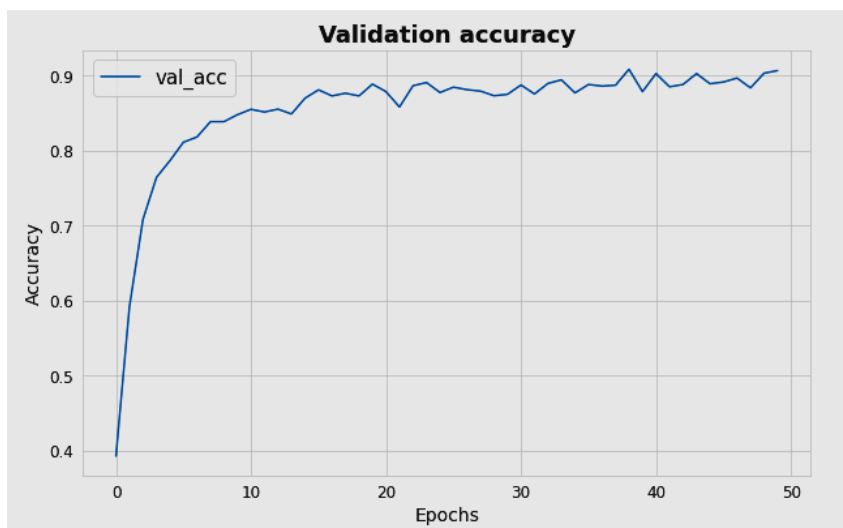
	Accuracy	Loss
Validation	0.996	0.111
Test	0.995	0.0199

4.2. Mô hình xác thực người nói

Đánh giá hiệu quả mô hình trên tập validation và tập test.



Hình 5. Biểu đồ mất mát trên dữ liệu huấn luyện và kiểm thử



Hình 6. Biểu đồ độ chính xác trên tập kiểm thử

4.3. Một số mô hình xác thực người nói khác

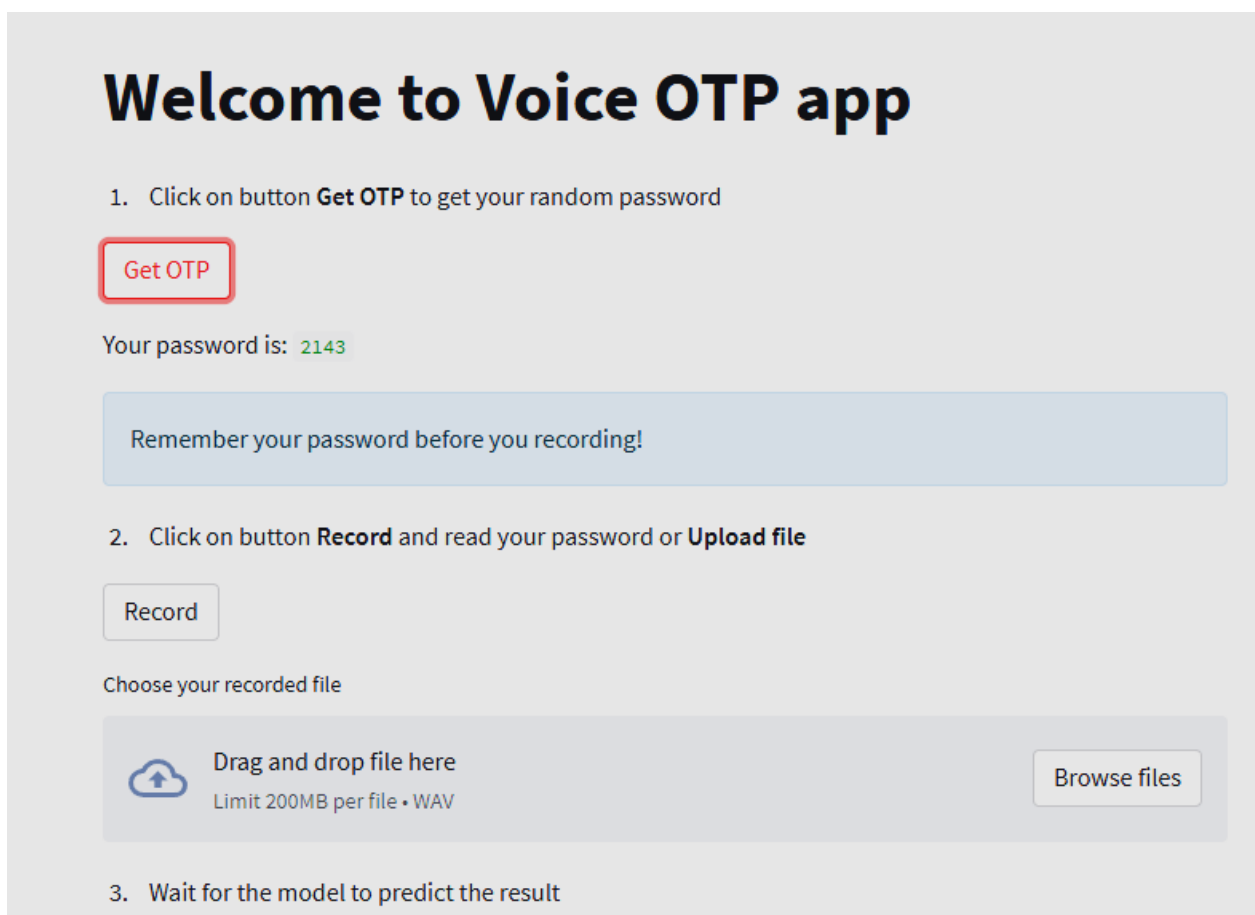
Bảng 1. Đánh giá thời gian thực hiện và độ chính xác trên tập train và tập test

		Accuracy	Loss	
Validation		90%	0.4798	
Test		89%	0.4694	
STT	Model	Time cost (ms)	Accuracy train	Accuracy test
1	Extreme Learning Machine	2959.84	67.22%	61.68%
2	Logistic Regession	140415.39	96.91%	89.92%
3	Support Vector Machine	80158.69	57.74%	60.86%
4	K-Nearest Neighbors	11.48	94.48%	91.79%
5	Decision tree	440980.14	100%	43.27%
6	Random Forest	1061295.29	100%	84.04%
7	Perceptron	398853.21	98.94%	88.76%

Trong các mô hình huấn luyện trên ta nhận thấy mô hình K-Nearest Neighbors có thời gian huấn luyện ngắn nhất và cũng cho ra được độ chính xác khá cao, ngoài ra xuất hiện hiện tượng overfitting ở các mô hình khác. Đây là hiện tượng mô hình quá khớp với dữ liệu huấn luyện nó sẽ đúng trên tập huấn luyện nhưng kết quả trên tập kiểm thử thì sẽ tệ, ở mô hình Decision tree có độ chính xác trên tập huấn luyện là 100% nhưng kết quả độ chính xác trên tập kiểm thử chỉ 43%.

4.4. Kết quả hiện thực mô hình bằng web

Người dùng sẽ có thể lấy mã OTP bằng nút ‘Get OTP’ sau đó sẽ thu âm giọng nói bằng nút ‘Record’ sau khi tiến hành thu âm xong người dùng sẽ submit lên và hệ thống sẽ trả lại kết quả.



Welcome to Voice OTP app

- Click on button **Get OTP** to get your random password

Get OTP

Your password is: 2143

Remember your password before you recording!

- Click on button **Record** and read your password or **Upload file**

Record

Choose your recorded file

Drag and drop file here
Limit 200MB per file • WAV

Browse files

- Wait for the model to predict the result

Welcome to Voice OTP app

1. Click on button **Get OTP** to get your random password

Get OTP

2. Click on button **Record** and read your password or **Upload file**

Record

Choose your recorded file

Drag and drop file here
Limit 200MB per file • WAV

Browse files

3. Wait for the model to predict the result

Recording...

Press q button on your keyboard to stop recording!

Hệ thống sẽ xác thực người nói là ai sau đó sẽ so khớp mã OTP người nói cung cấp nếu đúng với mã OTP của hệ thống sinh ra thì quá trình xác thực hoàn thành ngược lại nếu mã OTP người nói cung cấp khác so với hệ thống gửi ra thì quá trình xác thực không thành công.

Welcome to Voice OTP app

1. Click on button **Get OTP** to get your random password

Get OTP

2. Click on button **Record** and read your password or **Upload file**

Record

Choose your recorded file

Drag and drop file here
Limit 200MB per file • WAV

Browse files

3. Wait for the model to predict the result

Press q button on your keyboard to stop recording!

Segmented successfully!

Speaker: Phạm Thành Trung

OTP predict: 2143

You passed!

5. KẾT LUẬN – ĐỀ NGHỊ

5.1. Kết luận

Sau các giai đoạn thu thập, tiền xử lý dữ liệu âm thanh kết quả nhận được ở mô hình phân loại số đạt độ chính xác 99.6% trên tập huấn luyện và 99.5 trên tập kiểm thử, ở mô hình xác thực người nói đạt độ chính xác 90% trên tập huấn luyện và 89% trên tập kiểm thử, ngoài ra một số mô hình như K-Nearest Neighbors, Logistic Regression, Perceptron lần lượt đạt kết quả rất khả quan 94.48%, 96.91%, 98.94% trên tập huấn luyện và 91.79%, 89.92%, 88.76% trên tập kiểm thử.

Với kết quả đạt được chúng tôi đã xây dựng được mô hình phân loại số và phân loại người nói, bước đầu đã hoàn thành mục tiêu ban đầu đề ra có thể xác thực người dùng thông qua mã OTP sử dụng sinh trắc học giọng nói của người dùng đó với mục tiêu nâng cao mức độ bảo mật, bảo vệ người dùng khỏi nhiều mối nguy hại tiềm ẩn, giảm thiểu hành vi giả mạo trong giao dịch tài chính bên cạnh đó có thể ứng dụng vào các lĩnh vực khác trong đời sống công nghệ chuyển đổi số ngày nay.

5.2. Kiến nghị

Hạn chế

Bên cạnh các kết quả đạt được thì nghiên cứu của chúng tôi còn nhiều mặt hạn chế như dữ liệu huấn luyện còn hạn chế, các mô hình huấn luyện còn xuất hiện hiện tượng overfitting dẫn đến việc ứng dụng vào thực tiễn vẫn chưa thực hiện được.

Hướng nghiên cứu tương lai

Huấn luyện mô hình với số lượng dữ liệu nhiều hơn, cải thiện tốc độ xử lý của hệ thống, nâng cấp các giai đoạn bài toán bằng các mô hình học sâu hiện đại đánh giá và so sánh kết quả để tìm ra mô hình tối ưu nhất cho bài toán.

Nhóm nghiên cứu cũng đề ra việc ứng dụng bài toán để giải quyết cho vấn đề trong lĩnh vực giáo dục, cụ thể là việc học online của học sinh, sinh viên. Cụ thể giáo viên/giảng viên thường khó xác định được học sinh/sinh viên có đang ngồi học hay không cách giải quyết hiện tại là có thể yêu cầu người học bật camera trong suốt quá trình học sẽ gặp khó khăn nếu điều kiện học sinh/sinh viên không đáp ứng được thiết bị camera do đó nhóm nghiên cứu đề xuất cách quản lý người học bằng xác thực giọng nói. Ở một thời điểm nhất định học sinh/sinh viên sẽ phải thực hiện xác thực bản thân đang tham gia lớp học thông qua việc đọc ngẫu nhiên mã OTP để hệ thống xác thực. Hy vọng hướng nghiên cứu này có thể cải thiện chất lượng buổi học online của giáo viên/giảng viên đồng thời xây dựng ý thức học tập tốt cho học sinh/sinh viên.

TÀI LIỆU THAM KHẢO

- [1] D. Snyder, D. Garcia-Romero, G. Sell, D. Povey, and S. Khudanpur, “X-vectors: Robust DNN embeddings for speaker recognition,” in Proc. IEEE Int. Conf. Acoust., Speech, Signal Process., 2018, pp. 5329–5333.
- [2] A. Krizhevsky, I. Sutskever, G.E. Hinton. (2012). ImageNet Classification with Deep Convolutional Neural Networks. *Advances in Neural Information Processing Systems*, 25, 1097-1105.
- [3] C. Li, X. Ma, B. Jiang, X. Li, X. Zhang, X. Liu, Y. Cao, A. Kannan, and Z. Zhu, “Deep speaker: an end-to-end neural speaker embedding system,” arXiv preprint arXiv:1705.02304, 2017
- [4] Wan, Q. Wang, A. Papir, and I.L. Moreno, “Generalized end-to-end loss for speaker verification,” in 2018 IEEE International Conference on Acoustics, Speech and Signal
- [5] Y. Lukic, C. Vogt, O. Durr, and T. Stadelmann, “Speaker identification and clustering using convolutional neural networks,” in IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP). IEEE, 2016, pp. 1–6.

- [6] I. Lopez-Moreno, J. Gonzalez-Dominguez, O. Plchot, D. Martinez, J. Gonzalez-Rodriguez, and P. Moreno, “Automatic language identification using deep neural networks,” in Acoustics, Speech and Signal Processing (ICASSP), 2014 IEEE International Conference on. IEEE, 2014, pp. 5337–5341.
- [7] D. Snyder, D. Garcia-Romero, D. Povey, and S. Khudanpur, “Deep neural network embeddings for textindependent speaker verification,” Proc. Interspeech 2017, pp. 999–1003, 2017.
- [8] S. Shon, H. Tang, and J. Glass, “Frame-level speaker embeddings for text-independent speaker recognition and analysis of end-to-end model,” arXiv preprint arXiv:1809.04437, 2018.
- [9] K. Okabe, T. Koshinaka, and K. Shinoda, “Attentive statistics pooling for deep speaker embedding,” arXiv preprint arXiv:1803.10963, 2018.
- [10] W. Cai, Z. Cai, X. Zhang, X. Wang, and M. Li, “A novel learnable dictionary encoding layer for end-to-end language identification,” arXiv preprint arXiv:1804.00385, 2018.
- [11] W. Cai, J. Chen, and M. Li, “Exploring the encoding layer and loss function in end-to-end speaker and language recognition system,” arXiv preprint arXiv:1804.05160, 2018.
- [12] "VLSP 2021 - Vietnamese Speaker Verification" p.<https://vlsp.org.vn/vlsp2021/eval/vsv>.

HỘI NGHỊ KHOA HỌC TRẺ LẦN 4 NĂM 2022 (YSC 2022)
LĨNH VỰC ĐIỆN – ĐIỆN TỬ - CÔNG NGHỆ THÔNG TIN
TRƯỜNG ĐẠI HỌC CÔNG NGHIỆP THÀNH PHỐ HỒ CHÍ MINH

NHÀ XUẤT BẢN
ĐẠI HỌC CÔNG NGHIỆP TP HỒ CHÍ MINH
12 Nguyễn Văn Bảo – P. 4 – Q. Gò Vấp – TPHCM
ĐT: (028) 3894 0390 – 816 ; **Fax:** (028) 3994 0650
Email: nhaxuatban@iuh.edu.vn

Chịu trách nhiệm xuất bản: TRƯỜNG NGỌC THỜI

Biên tập: LÊ THỊ TIÊU NHI
Sửa bản in: ĐOÀN THANH ĐIỀN
Trình bày bìa: VĂN SANG

Đối tác liên kết: Trường ĐH Công nghiệp Thành phố Hồ Chí Minh

ISBN: 978-604-920-155-4

In 51 cuốn khổ 20 × 28 cm theo Quyết định xuất bản số: 21/QĐ-NXBĐHCN ngày 30/09/2022 với xác nhận đăng ký xuất bản số 3268-2022/CXBIPH/2-19/ĐHCNTPHCM ngày 19/09/2022. In tại Xưởng in NXB Đại học Công nghiệp TPHCM, nộp lưu chiểu tháng 11/2022.

Young
Scientists
Conference **2022**

INDUSTRIAL UNIVERSITY OF HO CHI MINH CITY

ISBN: 978-604-920-155-4



9 786049 201554

SÁCH KHÔNG BÁN