http://icmic-conf.org/

ICMIC 2023

The 2nd International Conference on Maritime IT Convergence

"Enhancing Convergence of ICT in the Artificial Intelligence Era"

PROCEEDINGS

Date August 23 (Wed) ~ 25 (Fri), 2023

Venue Shilla Stay, Jeju Island, Korea

Technically Co-Sponsored by IEICE Communications Society

Organized by KICS The Korean Institute of Communications and Information Sciences

Patrons LG U+, KIOST, JY SYSTEM (주)제이와이시스템, MET-SOLUTIONS Marine Equipment & Technology Solutions, ICT CRC ICT Convergence Research Center, AURI 한국산학연협회 Korea Association of University, Research Institute and Industry, kt

General Information

The International Conference on Maritime IT Convergence (ICMIC) aims to promote convergence activities of maritime, terrestrial and aerial communications as well as related advanced wireless communications. Recognizing that maritime communications will usually involve ship-to-ship and ship-to-shore communication, maritime ICT technologies will bring some mature but still evolving terrestrial wired/wireless communication technologies to its own use for future smart maritime communications. More specifically, the conference will focus on addressing challenges of maritime communications with ICT convergence or advanced future communications over various industrial sectors, academia and practice engineers. The conference will include keynote sessions, invited special sessions and technical paper sessions. You are invited to submit papers in all areas of maritime communications, applications for ICT convergence, technologies, or advanced future communications.

Proceedings of International Conference on Maritime IT Convergence (eISSN: 2983-3132)

Greetings

Message from the General Co-Chair of ICMIC 2023

First and foremost, I extend my sincere gratitude to all of you for participating in ICMIC 2023. This is the 2nd international conference held this year, following the successful event last year. It holds great significance as it focuses on maritime IT convergence and aims to introduce and share the latest technologies in the related IT field.

During the conference, a total of about 80 papers will be presented, accompanied by keynote speeches delivered by three distinguished foreign experts in both on-site and virtual sessions.

ICMIC 2023 is a dynamic international conference organized collaboratively by the Busan, Ulsan, and Gyeongnam Branches of The Korean Institute of Communications and Information Sciences (KICS), the largest academic institute in the field of ICT in Korea.

The successful hosting of ICMIC 2023 was made possible by the active cooperation of KIOST (Korea Institute of Ocean Science & Technology), LG Uplus, JY SYSTEM, MET-SOLUTIONS, and AURI (Korea Association of University, Research Institute, and Industry).

Furthermore, we are committed to elevating ICMIC to an even higher level, and we hope that all participants in this event will continue to collaborate and contribute to its development.

Over the course of three days, from August 23rd (Wed) to August 25th (Fri), we gather in Jeju, the beautiful city of dreams. I wish you all fruitful discussions, achievements, and opportunities to rejuvenate.

Warm regards,

General Co-chair of ICMIC 2023

Prof. Dong Myung Lee

Message from the Organizing Committee Chairs of ICMIC 2023

Dear members of The Korean Institute of Communications and Information Sciences (KICS) and participants of the ICMIC 2023.

ICMIC 2023 will be held in August, in the Jeju Island. The event will be on/off hybrid. This event is undoubtedly the most promising conference in the field of convergence of maritime, terrestrial and aerial communications. The keynote speeches by Prof. Chao Zhang from School of Aerospace Engineering, Tsinghua University, Beijing, China, Prof. Murat Yuksel from University of Central Florida, USA, and Prof. Rafael Pérez-Jiménez from University of Las Palmas de Gran Canaria, Spain are prepared along with various oral and poster presentations.

The Shilla Stay in Jeju Island seems to be a perfect place for participants to comfortably enjoy offline events and relaxation. It is expected that many professionals in related fields will be able to exchange ideas and collaborate through the event.

Lastly, we would like to thank the organizing committees and staffs for the generous support. We would like to ask many of you to participate online/offline, and wish everyone good health and well-being.

Sincerely,

ICMIC Organizing Committee Chairs of ICMIC 2023

Committees

International Advisory Committee

- Zabih Ghassemlooy (Northumbria University, UK)
- Een-Kee Hong (Kyung Hee University, Korea)
- Do Hyung Kang (KIOST, Korea)
- Dong Jin Kang (Korean Society of Oceanography, Korea)
- Anand Srivastava (IIIT Delhi, India)
- Murat Uysal (Ozyegin University, Turkey)

International Steering Committee

- Trio Adiono (Institute of Technology at Bandung, Indonesia)
- Chang-jun Ahn (Chiba University, Japan)
- Robert Bestak (Czech Technical University, Czech Republic)
- Manav R. Bhatnagar (IIT Delhi, India)
- Dong Wook Cho (Chungbuk Provincial University, Korea)
- Francisco Jose Lopez Hernandez(Universidad Polite cn icade Madrid,Spain)
- Sevia M. Idrus (Universiti Teknologi Malaysia, Malaysia)
- Yeon Man Jeong (Gangneung-Wonju National University, Korea)
- Dong Seong Kim (Kumoh National Institute of Technology, Korea)
- Woo Yong Lee (ETRI, Korea)
- Prasant Kumar Sahu (IIT Bhubaneswar, India)
- Himal A. Suraweera (University of Peradeniya, Sri Lanka)

General Co-Chairs

- Yeon Ho Chung (Pukyong National University, Korea)
- Dong Myung Lee (Tongmyong University, Korea)

Technical Program Committee Chairs

- Seungjae Baek (KIOST, Korea)
- Yonggang Kim (Kongju National University, Korea)

Technical Program Committee

- Motassem Al-Tarazi (University of Nebraska–Lincoln, USA)
- Sudhanshu Arya (Pukyong National University, Korea)
- Willy Anugrah Cahyadi (Telkom University, Indonesia)
- Indranath Chatterjee (Tongmyong University, Korea)
- Yu-Wen Chen (New York City College of Technology, New York, USA)
- Hyuckjin Choi (Kyushu University, Japan)
- Chanjun Chun (Chosun University, Korea)
- Vo Le Cuong (Hanoi University of Science and Technology, Vietnam)
- Minseok Han (Korea Naval Academy, Korea)
- Minkyu Je (Korea Advanced Institute of Science and Technology, Korea)
- Hyoyoung Jung (KOPTI, Korea)
- Beomjun Kim (Yonam Institute of Technology, Korea)
- Gyung Bae Kim (Seowon University, Korea)
- Kyungsook Kim (National Institute of Advanced Industrial Science and Technology, Japan)
- Ryangsoo Kim (ETRI, Korea)
- Suk Chan Kim (Pusan National University, Korea)
- Sunyong Kim (Dongseo University, Korea)
- Yohan Kim (Dongseo University, Korea)
- Yonghun Kim (KATECH, Korea)
- Jin Hwan Koh (Gyeongsang National University, Korea)
- Wang Sang Lee (Gyeongsang National University, Korea)
- Sung Hoon Lim (Hallym University, Korea)
- Jenn-Wei Lin (Fu Jen Catholic University, New Taipei City, Taiwan)
- Suman Malik (IIT Kanpur, India)
- Sifat Rezwan (TU Dresden, Germany)
- Rehenuma Tasnim Rodoshi (George Mason University, USA)
- Prakriti Saxena (INRIA, France)
- Yujae Song (Kumoh National Institute of Technology, Korea)
- Arif Ullah (Chosun University, Pakistan)
- Min Sang Yoon (Fujitsu Network Communications, Texas, USA)

Finance Chair

- Taewoon Kim (Pusan National University, Korea)

Publicity Chair

- Wooyeol Choi (Chosun University, Korea)

Registration Chair

- Hyuntae Cho (Tongmyong University, Korea)

Publication Chair

- Dongwan Kim (Dong-A University, Korea)

Web Chair

- Incheol Shin (Pukyoung National University)

Table of Contents

Oral Sessions

Oral Session 1A : KIOST Special Session I

Chair : Yujae Song (Kumoh National Institute of Technology, Korea)

Oral Session 1B : KIOST Special Session II

Chair : Taewoon Kim (Pusan National University, Korea)

Oral Session 2A : Maritime Internet of Things

Chair : Yujae Song (Kumoh National Institute of Technology, Korea)

Table of Contents

Oral Session 2B : Maritime Artificial Intelligence I

Chair : Trio Adiono (Bandung Institute of Technology, Indonesia)

Oral Session 3A : Advanced Maritime Control and Management

Chair : Yonggang Kim (Kongju National University, Korea)

Oral Session 3B : Maritime Artificial Intelligence II

Chair : Taewoon Kim (Pusan National University, Korea)

Oral Session 4A : Maritime Wireless Communications

Chair : Yonggang Kim (Kongju National University, Korea)

Table of Contents

Oral Session 4B : Object Recognition and Pattern Awareness

Chair : Arif Ullah (Chosun University, Korea)

Oral Session 5A : 6G, 5G, and LTE-Advanced Networks I

Chair : Woo Yong Lee (ETRI, Korea)

Oral Session 5B : ICT Convergence Technologies I

Chair : Dong-Seong Kim (Kumoh National Institute of Technology, Korea)

Oral Session 6A : 6G, 5G, and LTE-Advanced Networks II

Chair : Siti Hasunah Mohammad (Universiti Teknologi Malaysia, Malaysia)

Oral Session 6B : ICT Convergence Technologies II

Chair : Dong-Seong Kim (Kumoh National Institute of Technology, Korea)

Oral Session 7A : Information and Communication Technologies I

Chair : Arif Ullah (Chosun University, Korea)

Oral Session 7B : Big Data and Smart Computing

Chair : Hyuntae Cho (Tongmyong University, Korea)

Oral Session 8A : Information and Communication Technologies II

Chair : Sevia Mahdaliza Idrus (Universiti Teknologi Malaysia, Malaysia)

Oral Session 8B : Systems, Applications, and Data Science

Chair : Raghavendra Ganiga (Chosun University, Korea)

Poster Session

Poster Session

Chair : Yonghun Kim (KATECH, Korea)

Oral Sessions

Underwater Optical Beam Orientation Angle and Divergence Angle Control based on Deep Reinforcement Learning

Huicheol Shin
Korea Institute of Ocean Science and Technology
South Korea
shc0305@kiost.ac.kr

Yujae Song
Yeungnam University
South Korea
yjsong@yu.ac.kr

Seungjae Baek
Korea Institute of Ocean Science and Technology
South Korea
baeksj@kiost.ac.kr

***Abstract*—In this work, we first analyze a communication performance in underwater optical wireless communication (UOWC) between a transmit node fixed to the seabed (e.g., underwater sensor) and a receive node moving on sea surface (i.e., ship, unmanned surface vehicle). Based on the performance analysis, we then propose a two-phase deep reinforcement learning based algorithm which jointly optimize beam orientation angle and beam divergence angle at the transmit node to establish a reliable UOWC link. The proposed algorithm maximizes instantaneous signal-to-noise ratio by minimizing the beam divergence angle while maintaining the angle difference (i.e., inclination angle) between the selected beam orientation and the actual orientation between the transmit node and receive node below 2°.**

***Index Terms*—Deep reinforcement learning, underwater optical wireless communication, beam divergence angle, beam orientation angle**

I. Introduction

Underwater optical wireless communication (UOWC) has gained significant attention in recent years due to its potential for high-speed data transmission, wide bandwidth, and low latency in underwater environment. However, UOWC faces several performance degradation challenges caused by various factors such as absorption, scattering, attenuation, fading, and misalignment. These challenges arise due to the harsh underwater environment, including waves, turbidity, and turbulence [1]. To address these challenges and improve the reliability of UOWC system, researchers have been exploring various solutions [2], [3]. In this paper, we propose a two-phase deep reinforcement learning-based algorithm for jointly controlling the beam orientation angle and beam divergence angle in UOWC system. The primary objective of this algorithm is to reduce the probability of disconnection between the transmit node and receive node by mitigating misalignment issues.

II. System model

In this work, we focus on a laser-based UOWC link that comprises a transmit node, which serves as an underwater sensor, and a receive node, which can be a ship or an unmanned surface vehicle floating above sea level. As illustrated in Fig. 1, transmit node is fixed in position on the seabed and is responsible for measuring various underwater environmental

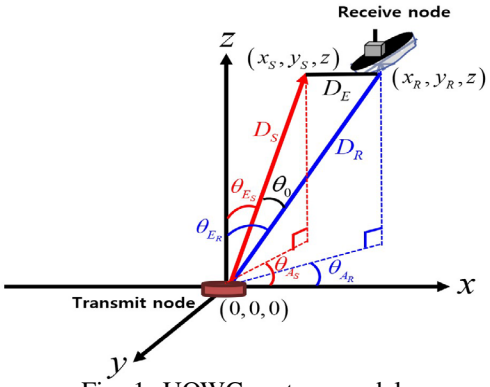


Fig. 1: UOWC system model

data. Whereas, the receive node is mobile and its target location (i.e., a location where the underwater sensor is installed vertically) is determined based on global positioning system (GPS) information. Despite of determined location, the receive node can irregularly move or shake due to various external factors such as wave and turbulence in the ocean.

In the considered UOWC environment, a beam orientation angle and an actual orientation angle can be expressed as azimuth angle (i.e., θ_{A_S} and θ_{A_R}) and elevation angle (i.e., θ_{E_S} and θ_{E_R}), and the correlation with three-dimensional location coordinates is as follow:

$$\theta_A = \tan^{-1}(\frac{y}{x}), \tag{1}$$

$$\theta_E = \tan^{-1}(\frac{\sqrt{x^2+y^2}}{z}), \tag{2}$$

If we know the location of the sea level where the optical beam reaches and the location of the receive node, the linear distance (i.e., D_S and D_R) and horizontal distance difference D_E can be calculated based on the sea level coordinates as follows:

$$D_S = \sqrt{x_S^2 + y_S^2 + z^2}, \tag{3}$$

$$D_R = \sqrt{x_R^2 + y_R^2 + z^2}, \tag{4}$$

$$D_E = \sqrt{(x_S - x_R)^2 + (y_S - y_R)^2}. \tag{5}$$

In addition, the inclination angle θ_0 between the beam

orientation angle and the actual orientation angle is determined by the linear distance and the horizontal distance difference as follows:

$$\theta_0 = \cos^{-1}\left(\frac{D_R^2 + D_S^2 + D_E^2}{2D_R D_S}\right). \quad (6)$$

In an UOWC environment, the signal-to-noise-ratio (SNR) of communication link can be calculated as shown in the following equation:

$$SNR = \frac{I_P}{I_N}, \quad (7)$$

where I_N is the noise current that defined as the summation of the thermal noise, shot noise, and dark current noise [4] and I_P is the received optical current that calculated by the performance of photodiode and received power, of which the received power can be obtained as follows:

$$P_R = P_T \eta_T \eta_R P_L G_L, \quad (8)$$

where P_T is the transmission power, η_T and η_R are transmission and receive efficiency, respectively, P_L is the propagation loss, and G_L is the geometric loss. The propagation loss of an optical signal in the underwater environment can be calculated as follows:

$$P_L = \exp\left\{c(\lambda)\frac{D_S}{cos(\theta_0)}\right\}, \quad (9)$$

where $c(\lambda)$ is the attenuation coefficient. The geometric loss of underwater optical signal can be calculated as follows:

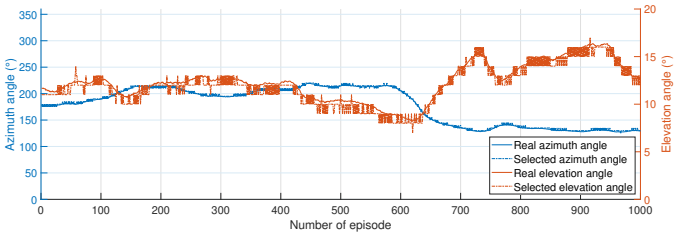
$$G_L = \begin{cases} \frac{\alpha cos(\theta)}{2\pi D_S^2[1-cos(\theta_0)]}, \theta_0 \leqslant \theta \\ 0, \qquad\qquad \text{otherwise}, \end{cases} \quad (10)$$

where α is aperture area of optical modem and θ is beam divergence angle. According to the above equation, G_L can be obtained only when the beam divergence angle of transmit node is greater than inclination angle. Accordingly, in order to succeed in data transmission, the receive node must be located within the beam range. Therefore, the transmit node needs to minimize the inclination angle by adjusting the beam orientation angle according to the location of the receive node and maintain the beam divergence angle larger than the inclination angle. To this end, we proposed a two-phase deep reinforcement learning algorithm for determining the beam orientation angle and beam divergence angle of the transmit node.

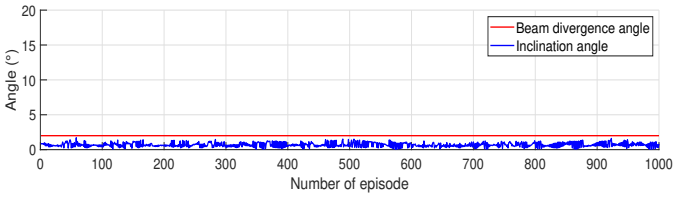
III. Simulation Result

To analyze the performance of the proposed algorithm, we performed the simulation by utilizing GPS data from the unmanned surface vehicle (USV) floating at sea level as the location data of receive node.

Fig. 2 (a) shows the learning result of the outer agent that controls the beam orientation angle by learning the location of the receive node. As a result, the average angle difference between the azimuth angle and elevation angle are $0.94\,^\circ$ and $0.36\,^\circ$. Accordingly, the angle difference (i.e., inclination angle) is maintained below $2\,^\circ$. In Fig. 2 (b), the inner agent maintains the beam divergence angle larger than the inclination angle to prevent the link disconnection, and as a result, a reliable UOWC link without link disconnection is established.



(a) Outer agent simulation result



(b) Inner agent simulation result

Fig. 2: Performance of proposed algorithm.

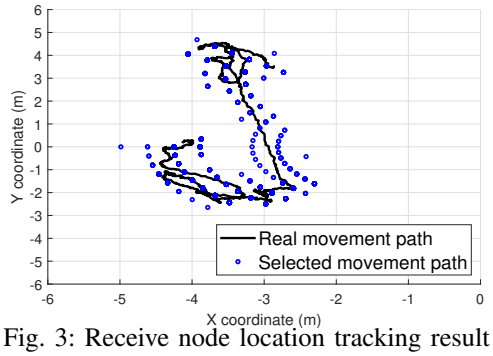


Fig. 3: Receive node location tracking result

Finally, Fig. 3 shows the movement path of the receive node and the movement path of the optical beam according to the orientation angle selected by the proposed algorithm.

Acknowledgment

This research was a part of the project titled 'Development of Polar Region Communication Technology and Equipment for Internet of Extreme Things (IoET)', funded by Ministry of Science and ICT (MSIT).

References

[1] N. Saeed, A. Celik, T.Y. Al-Naffouri and M. Alouini, "Underwater optical wireless communications, networking, and localization: A survey," *Ad Hoc Networks*, vol.94, pp.1-34, 2019.

[2] A. Celik et al., "End-to-end performance analysis of underwater optical wireless relaying and routing techniques under location uncertainty," *IEEE Trans. Wirel. Commun.*, vol.19, pp.1167-1181, 2020.

[3] H. Shin, Y. Kim, S. Baek and Y. Song, "Distributed learning for dynamic channel access in underwater sensor networks," *Entorpy 2020*, vol.22, pp.1-16, 2020.

[4] M.A.A. Ali, "Comparison of modulation techniques for underwater optical wireless communication employing APD receivers," *Research journal of applied sciences, engineering and technology*, vol.10, no.6, pp.707-715, 2015.

Underwater Environmental Radiation Measurement System

Jungmin Seo
Maritime ICT & Mobility Research Department
Korea Institute of Ocean Science and Technology(KIOST)
Busan, Korea
jmseo@kiost.ac.kr

Soo Mee Kim
Maritime ICT & Mobility Research Department
Korea Institute of Ocean Science and Technology(KIOST)
Busan, Korea
smeekim@kiost.ac.kr

***Abstract*—In this study, we compared the environmental radiation measurements on-land and underwater using our developed waterproof environmental gamma-ray detection system. The system incorporates an aluminum case and O-rings for the waterproof design of the NaI(Tl) scintillation detector from Ortec. Additionally, the communication method has been improved by changing to Ethernet-based USB communication, enhancing the communication distance capability. The experiments involved bundling three radioactive sources (Cs-137, Ba-133, and Co-60) and comparing the measured values at natural state (background) and distances of 25, 50, 75, and 100 cm both on land and underwater. The results revealed that all three radiation sources could be simultaneously measured. As the distance increased, the measured values decreased due to attenuation by water. Moreover, the measured values in the underwater environment were consistently lower compared to the values measured in the land environment at the same distance because of the higher attenuation by water. Based on these updated findings, this study contributes to a better understanding of radiation measurements in different environments, highlighting the importance of considering distance-dependent attenuation effects in both land and underwater radiation measurements.**

Keywords—underwater environmental radiation monitoring, waterproof case design, NaI(Tl) scintillation detector

I. INTRODUCTION

Radiation measurement plays a crucial role in various fields, including environmental monitoring, nuclear power plants, and radiation therapy. The accurate assessment of radiation levels in different environments, such as land and underwater, is essential for understanding the potential risks and developing appropriate safety measures[1]. In this study, we aim to compare radiation measurements between land and underwater environments using a waterproof environmental radiation detection system.

To provide a comprehensive analysis, we conducted experiments using three radiation sources: Cs-137, Ba-133, and Co-60. The measurements were performed at various distances (25, 50, 75, and 100 cm) from the radiation sources. Also, we measured the background spectra without any radiation sources. It was confirmed that the damping effect occurred by comparing the values measured in the land and underwater environments.

II. METHOD

The waterproof radiation detection system used in this study consisted of the following components shown in Fig. 1,

- ORTEC 905-4 NaI(Tl) Scintillation Detector [2]: The scintillation detector served as the primary sensor for radiation measurement. It was connected to a data acquisition system on the ground for signal processing and data recording.
- Waterproof Case and O-ring Seal: To ensure the system's integrity underwater, a custom-designed waterproof case with an O-ring seal was used to protect the detector from water ingress.USB over IP and POE Splitter: These devices enabled power supply and Ethernet communication for the system, allowing remote control and data transmission.

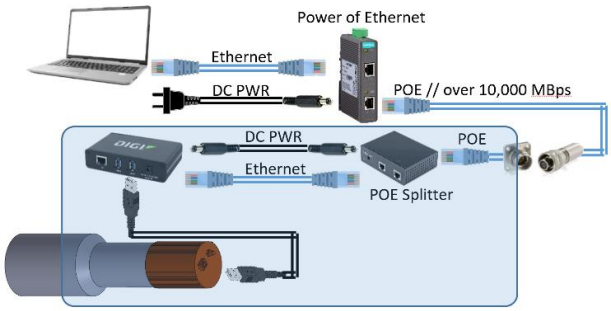


Fig. 1 Diagram of underwater radiation measurement system

Fig.2 shows the composition and arrangement of each element in the waterproof case.

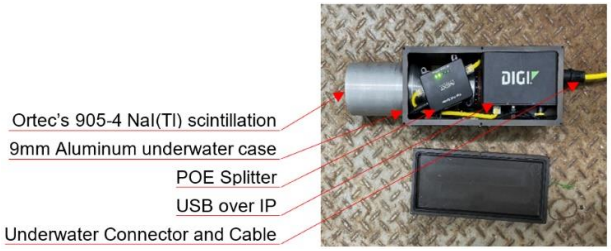
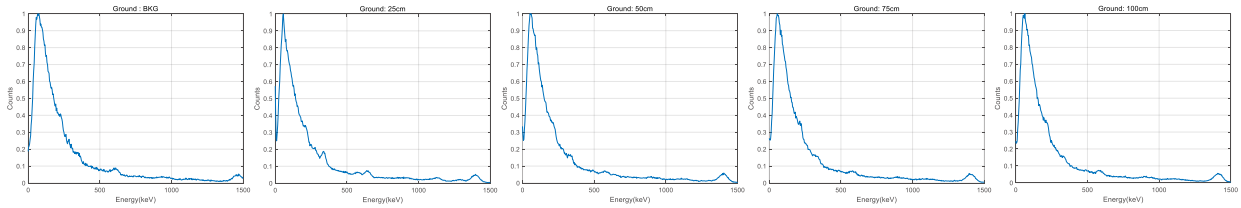


Fig. 2 System Configuration

III. RESULTS

The experimental setup involved measuring the radiation sources in both land and underwater environments. The radioactive sources (Cs-137, Ba-133, and Co-60) were bundled together and placed at a fixed distance from the detector. The measurements were performed at the background (natural state) without three radioactive sources, and then measured at the distances of 25, 50, 75, and 100 cm from three radioactive sources. The measurement duration for each distance was set to 10 minutes.

- Ground (V_{bias} = 830 V) : 10 min simultaneous acquisition with three radioactive source



- Underwater (V_{bias} = 830 V) : 10 min simultaneous acquisition with three radioactive source

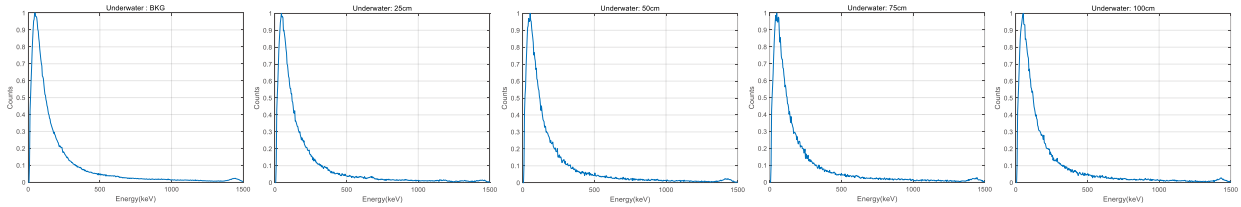


Fig. 4 Measured ground and underwater radiation spectra at different distances

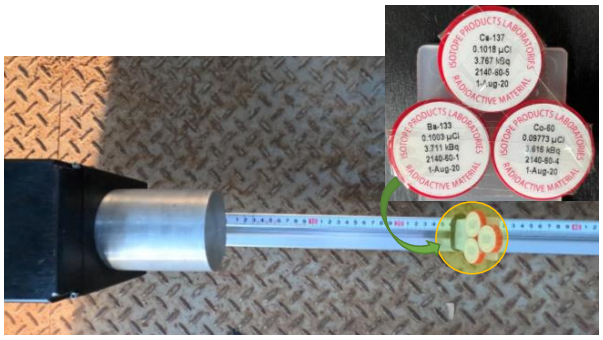


Fig. 3 Experimental setup for environmental radiation measurement from three radioactive sources at a distance of 25 cm

For example, Fig. 3 shows the experimental setup of three radioactive sources assembled and fixed at 25 cm away from the underwater radiation detector. The ground environment was measured in an empty water tank, and the underwater measurement was conducted after filling water up to a depth of 1 m in the same water tank. To maintain the same measurement conditions, the detector was placed in a waterproof case when measuring in land environment. The voltage bias (V_bias) is set to 830V for all measurements.

The acquired data were analyzed to compare the measured radiation levels measured in land and underwater environments. Fig. 4 compared the radiation measurements in land and underwater environments. The results indicate that all three radioactive sources could be simultaneously measured using the underwater environmental radiation detection system. As the distance increased, the measured radiation levels decreased in land and underwater environments. However, it was observed that the attenuation effect was more prominent in the underwater environment. At the same distance, the radiation levels measured underwater were consistently lower compared to the land measurements. This discrepancy can be attributed to the different densities of air and water, which results in higher attenuation of radiation underwater.

IV. Conclusion

The findings of this study demonstrate the importance of considering the environmental conditions and distance-dependent attenuation effects when performing radiation measurements. Computerized analysis is needed to determine the attenuation of the radiation dose in the environment. The results clearly indicate that the underwater environment poses additional challenges in accurately assessing radiation levels. The higher attenuation observed underwater suggests that special consideration is required when evaluating radiation risks in underwater settings. The developed waterproof radiation detection system proved to be effective in both land and underwater environments, providing reliable measurements under challenging conditions. In conclusion, this study contributes to a better understanding of radiation measurements in different environments, highlighting the significance of distance-dependent attenuation effects.

Further research includes the study of attenuation according to density and the development of equipment capable of measuring radiation dose at various depths.

Acknowledgment

This research was a part of the project titled ‘Development of Smart Processing Technology for Sea Foods’, funded by the Ministry of Oceans and Fisheries, Korea.

References

[1] PART, N. G. Radiation protection and safety of radiation sources International Basic Safety Standards. 2011.

[2] https://www.ortec-online.com/products/radiation-detectors/scintillation-detectors/scintillation-detector-types/905-series

A Study on the Analysis of Micro-Doppler Signature for Human Gait Parameters Estimation

Min Kim
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
skymin@kiost.ac.kr

Seungjae Baek
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
baeksj@kiost.ac.kr

Eugin Hyun
Division of Automotive Technology
DGIST
Daegu, South Korea
braham@dgist.ac.kr

Youngseok Jin
Division of Automotive Technology
DGIST
Daegu, South Korea
ysjin@dgist.ac.kr

Jieun Bae
Division of Automotive Technology
DGIST
Daegu, South Korea
jebae@dgist.ac.kr

Inoh Choi
Electronics and Communications Engineering
Korea Maritime & Ocean University
Busan, South Korea
inohchoi@kmou.ac.kr

***Abstract*—The compact IoT radar sensors can be used to analyze indoor walking human. For this purpose, micro-Doppler signature is usually extracted in 2D time-frequency domain and used. However, the complex computations of these method can be burdensome for compact radar sensor. To efficiently analyze human gait, we herein propose a new scheme to estimate gait parameters by extracting phase components of received radar echo. In experiments using 5 GHz CW radar hardware, we observed that our proposed scheme is not only efficient but also capable of performing accurate and robust gait parameters estimation even in the presence of unexpected motion.**

***Index Terms*—Radar sensor, gait analysis, micro-Doppler signature**

I. Introduction

Human gait refers to the pattern of human movement during locomotion, and gait analysis involves the systematic study on the human motion. Therefore, the parameters including step width, cadence (gait rate), and velocity are considered in gait analysis, because they offer insights into both intrinsic and extrinsic characteristics of human movement. Recently, in order to classify users and recognize their specific behavior, radar signal processing techniques have been employed for indoor human gait analysis [1], [2]. However, given the limited computational resources of Internet of Things (IoT) compact radar sensors, these complex computations of existing gait analysis method can be burdensome for these sensors.

In the proposed method, the phase components of received radar echoes are extracted to efficiently analyze human gait without generating the spectrogram which is joint time-frequency representation.

II. Echo Signal Model

In this section, we introduce an echo signal model to analysis gait pattern of a walking human. We have assumed that the walking human consists of P scattering points, as shown in Fig. 1. The radar echo signal, $s(t)$, returned by walking human can be represented as:

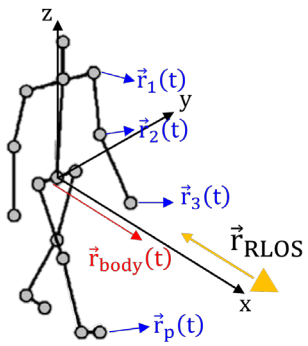


Fig. 1. Geometry of the walking human for radar echo signal.

$$s(t) = \sum_{p=1}^{P} \sigma_p exp\left(j\frac{4\pi}{\lambda} \left\{ \vec{r}_{body}^{T} + \vec{r}_{p}^{T}(t) \right\} \cdot \vec{r}_{RLOS} + j\phi \right), \tag{1}$$

where σ_p is the complex amplitude for pth scattering point; $\vec{r}_{RLOS}$ is a RLOS vector; $\vec{r}_{body}$ is a vector for whole body movement; $\vec{r}_p(t)$ is a vector for pth scattering point; λ is the wavelength of radar signal; ϕ is a component related to stationary clutter; and superscript T is the transpose operator. To analyze micro-Doppler signature (MDS) for walking human's limbs, the phase terms of received radar echoes should be extracted. To this end, the radar echo in (1) can be derived as follows based on scenarios, in which limbs of walking human move at a constant cycle:

$$s(t) = \sigma_{tot} \cdot exp\left(j\frac{4\pi}{\lambda} \left\{ x_{body}(t) + w_d(t) \otimes x_{tot}(t) \right\} + j\phi \right), \tag{2}$$

where σ_{tot} is the amplitude of $s(t)$; $x_{body}(t) \approx vt + at^2/2$ is the hole body movement along RLOS; $x_{tot}(t)$ is the mixed

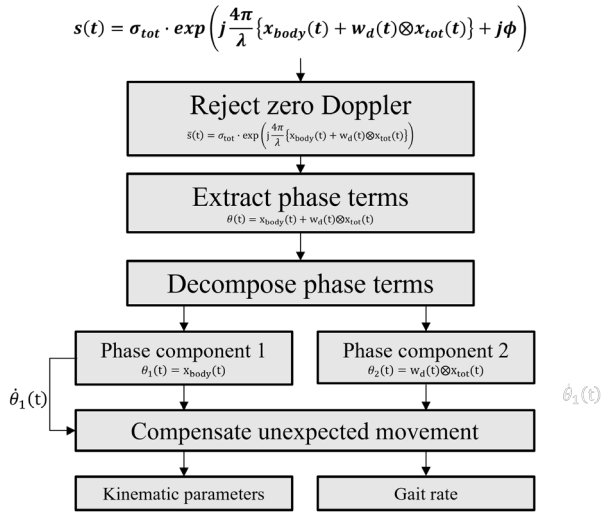


Fig. 2. Flowchart of the proposed scheme.

micro motions for scatterers of limbs; and

$$w_d(t) = \sum_{k=0}^{N} \delta\left(t - \frac{k}{f_g}\right). \tag{3}$$

where $\delta(\cdot)$ is the delta function; and f_g is the human gait rate.

III. Proposed Scheme

A complete flowchart of the proposed method is shown in Fig. 2. In the proposed method, the phase components of received radar echoes are extracted to analyze MDS. To address the multiple scatterers problem [3], we model the mixed signals reflected by multiple scattering points of the walking human's limbs based on locomotion scenarios. Subsequently, the zero-Doppler components are filtered out to minimize interference by the stationary clutters. And then, phase terms are decomposed into components related to whole body movements excluding the limbs' motion and micro motions of the limbs. If there are unexpected movements, the phase components are distorted. To compensate these phase distortions, the gradient of phase component for whole body movements is used. Finally, we can estimate the kinematic and gait parameters of a walking human.

IV. Experimental Results

To verify the performance of the proposed scheme in the presence of a unexpected motion via measurements, we used a 5 GHz CW radar sensor. Fig. 3 shows the extracted and decomposed phase components for limbs' micro motion. Because there are unexpected movements during locomotion, the phase components are distorted as shown in Fig. 3a. As a result of applying the compensation algorithm, distortion of phase term can be minimized as shown in Fig. 3b.

To estimate gait rate, spectrum analysis was conducted by applying FFT to compensated phase component. Consequently, the proposed method is robust to phase distortion

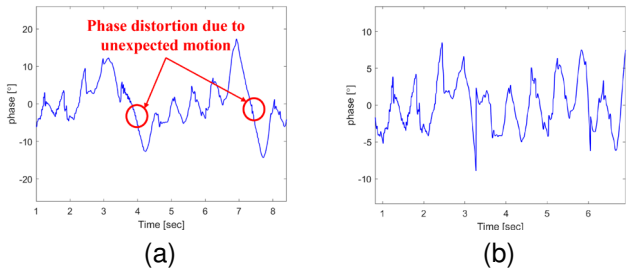


Fig. 3. Phase components for micro motion of walking human's limbs (a) before, and (b) after compensation.

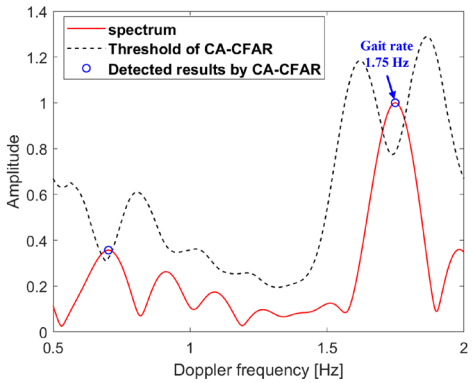


Fig. 4. Representation of the spectrum for compensated phase components in Fig. 3b.

caused by the critical unexpected movements and can efficiently estimate gait rate of walking human in 1D frequency domain.

V. Conclusion

In this paper, a new scheme based on phase extraction and compensation was proposed to efficiently estimate gait parameters in walking human involving unexpected movements. Based on this result, we intend to expand the research to real-time IoT sensor for user classification and gait analysis.

Acknowledgment

This research was supported in part of the project titled 'Development of Polar Region Communication Technology and Equipment for Internet of Things (IoET)', funded by Ministry of Science and ICT (MSIT)

References

[1] Bo Wang, Hao Zhang, and Yong-Xin Guo, "Radar-Based Soft Fall Detection Using Pattern Contour Vector",IEEE Internet of Tings Journal, Vol. 10, No. 3, pp.2519–2527, 2023.

[2] S. Rani, A. Chowdhury, T. Chakravarty, and A. Pal, "Exploiting Unique State Transitions to Capture Micro-Doppler Signatures of Human Actions Using CW Radar, Vol. 21, No. 24, pp. 27878–27886, 2021.

[3] Nan Su, F. Dai, H. Liu, and Bo Zhang, "Micro-Doppler frequency estimation and association via the inverse-free complex SBL TVAR method", IET Radar, Sonar & Navigation, Vol. 15, No. 10, pp. 1247–1265, 2021.

Illuminance-invariant RSW Block Detection via Image Enhancement and Mask R-CNN

Yong-Soo Ha
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
ysha@kiost.ac.kr

Myoung Hak Oh
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
omyhak@kiost.ac.kr

Sang Ki Jung
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
jeongsk313@kiost.ac.kr

Hae Yong Park
Maritime ICT & Mobility Research
KIOST
Busan, South Korea
hypark@kiost.ac.kr

Minh-Vuong Pham
Department of Ocean Engineering
Pukyong national university
Busan, South Korea
vuongpm@dau.edu.vn

Yun-Tae Kim
Department of Ocean Engineering
Pukyong national university
Busan, South Korea
yuntkim@pknu.ac.kr

***Abstract*— Mask R-CNN, a representative instance segmentation technique, was used to evaluate the detection performance of a laboratory reinforced soil retaining wall (RSW) block for changes in illuminance. To enhance images with changes in illumination, gamma correction (GC), histogram equalization (HE), and low-light image enhancement via illuminance map (LIME) were set as comparison groups, and the performance of image enhancement of the proposed RGB equalization was evaluated. It became. Each Mask R-CNN model was trained based on the original image dataset and the dataset to which each image enhancement method was applied. The performance of block detection for individual trained models was evaluated. The image improved through the RE method showed excellent LOE, PSNR, and SSIM compared to other methods, and in the case of the RE model, it showed excellent image enhancement performance and excellent block detection performance.**

Keywords—mask R-CNN, illuminance, image enhancement, detection, RGB equalization

I. INTRODUCTION

Mask R-CNN [1], a representative instance segmentation method, performs fully convolutional networks (FCN) based on the bounding box and classification information estimated by faster R-CNN. The method performs instance segmentation by estimating the mask without loss of location information. Mask R-CNN had been used in various fields to accurately detect objects with pixel unit. In order to continuously monitor the structure and analyze its behavior, a camera must be installed and monitored at a specific location. However, segmentation-based detection studies including Mask R-CNN have not analyzed the detection performance of changes in illuminance that occur over time for the same camera viewpoint. Therefore, in order to apply RSW monitoring to the field, it is necessary to quantitatively evaluate the detection performance for individual illuminance because the illuminance varies depending on the brightness and weather.

In this study, the performance of RSW block detection according to the application of the image enhancement method is evaluated for fixed camera images affected by illumination according to changes in time and weather in the field. We employed several image enhancement algorithms such as Gamma correction, Histogram equalization, low-light image enhancement via illuminance map [2], RGB equalization (proposed) to enhance the quality of datasets. Also, the Mask R-CNN was adapted to each enhanced model to identify and localize RSW blocks with different illuminances and compare the performance of detection regardless of illuminance.

II. METHOD

Figure 1 describes the workflow used in this study. Image enhancement is performed by introducing individual image processing techniques for block images according to various illuminances, and performance of image enhancement is evaluated. Then, for the original dataset and each corrected dataset, a block detection model based on Mask R-CNN is built and trained, and the performance of detection of each model is evaluated.

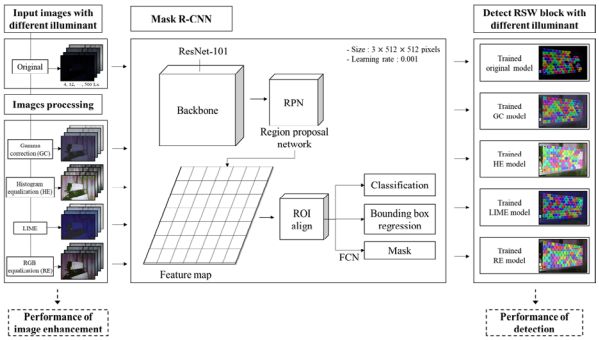


Fig. 1. Flowchart of image enhancement and detection of RSW block with different illuminanc

A. Image enhancement

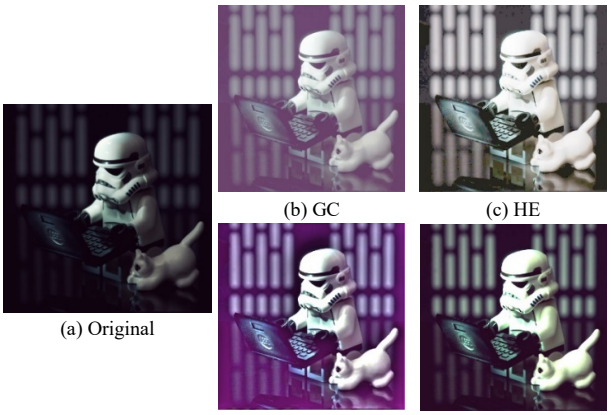
The methods such as GC, HE, and LIME focused on improving the quality of images. RE equalizes the average value of a specific or all RGB to target intensity and aims to make each image pair similar. The intensity corrected through RE based on the RGB equalization function (Ref) for the input image R is as shown in Eq. 1. In this study, the block intensity of 200 was targeted and improved.

$$\mathrm{Ref}(m, n) = \sum_{i=1}^{m} \sum_{j=0}^{n} k(i, j)$$

$$R_{(x, y)\_enhanced} = R_{(x, y)\_input} \times \frac{\mathrm{Ref}(m, n)}{m \times n} \times f_e,$$

Where, m and n are the height and width of image R and specific area, f_e is the equalization factor, k(i, j) is intensity corresponding to i, j.

Figure 2 provides several enhanced images using GC, HE, and LIME. Guo et al. The image used in 2013 was used, and the GC method uses $\gamma = 0.2$ and the RE method uses $f_e = 100$ to show the corrected image. Compared to the original image, the image quality has been improved, and the pixel intensity characteristics are different depending on the individual method.



(a) Original (b) GC (c) HE (d) LIME (e) RE

Fig. 2. The comparison of enhancement using different methods

Performance of block detection

Accuracy, precision, recall, F1 score are normally used to evaluate the performance of instance segmentation methods. Pixels of output images were divided into true positive (TP) (blocks correctly identified); false positive (FP) (background pixels incorrectly identified as block pixels); true negative (TN) (background pixels correctly identified); and false negative (FN) (block pixels identified as background pixels) results. Precision refers to the proportion of correct positive predictions among all positive predictions. Recall considers the completeness of positive predictions (i.e., the number of crack pixels that were correctly classified). F1 Score is the harmonic mean of precision and recall, which we used as an indicator of overall accuracy.

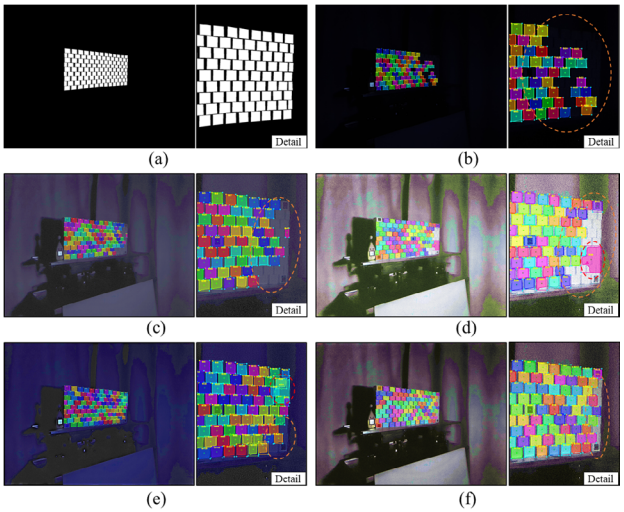
III. Results and Disscussion

Table 2 show performance comparison with different enhancement methods. Each performance metric was obtained by comparing the original image and the enhanced image at individual illumination levels, and SSIM_500 obtained results by comparing the enhanced images with the enhanced image at 500 Lux illumination. Although some illuminance and enhancement methods show better values, in general, the RE method shows the best performance in all performance metrics.

Figure 3 shows the ground truth of the laboratory RSW at 4 Lux and the detection results with each trained model. The original model detected many blocks even though it was difficult to visually identify them. However, significantly fewer blocks were detected compared to other models. In the HE and the LIME model, an error occurred in which multiple blocks were detected as one block. In the RE model, all blocks except 7 out of 121 blocks were detected, showing the best detection performance.

Table 1 Performance comparison with different enhancement methods

Metrics	Enhancement methods			
	GC	HE	LIME	RE
LOE (↓)	223.05	239.97	590.50	**176.92**
PSNR (↑)	27.53	27.93	27.59	**28.02**
SSIM (↑)	0.40	0.32	0.32	**0.48**
SSIM_500 (↑)	0.52	0.47	0.27	**0.56**



(a) (b) (c) (d) (e) (f)

Fig. 3. Block detection results and detailed results with different Mask R-CNN model at 4 Lux: (a) ground truth, (b) trained original model, (c) trained GC model, (d) trained HE model, (e) trained LIME model, (f) trained RE model.

IV. Conclusion

In this study, the change in illumination that occurs in the field was considered, and an image enhancement model was applied to correct it. The image enhancement performance of the proposed RE model was the best, and the RE trained model using this showed the best RSW block measurement performance. Through the development process, we plan to develop and apply a technology that can detect RSW blocks with perfect performance even at low illumination, taking into account changes in field illumination.

Acknowledgment

This research was a part of the project titled 'Development of smart maintenance monitoring techniques to prepare for disaster and deterioration of port infra structures', funded by the Ministry of Oceans and Fisheries, Korea.

References

[1] K. He, G. Gkioxari, P. Dollár and R. Girshick, “Mask r-cnn,” In Proceedings of the IEEE international conference on computer vision, pp. 2961-2969, 2017.

[2] X. Guo, Y. Li, and H. Ling, “LIME: Low-light image enhancement via illumination map estimation”. IEEE Transactions on image processing, 26(2), pp. 982-993, 2017.

A Study on Rule and Marine Environment Learning-based Unmanned Surface Vehicle Swarm Control

Jeong, SangKi
Maritime ICT & Mobility Research Department
Korea Institute of Ocean Science & Technology
BUSAN, Republic of Korea
jeongsk313@kiost.ac.kr

Oh, Myounghak
Maritime ICT & Mobility Research Department
Korea Institute of Ocean Science & Technology
BUSAN, Republic of Korea
omyhak@kiost.ac.kr

Park, Haeyong
Maritime ICT & Mobility Research Department
Korea Institute of Ocean Science & Technology
BUSAN, Republic of Korea
hypark@kiost.ac.kr

***Abstract*—**The article describes a study on the development of a system for controlling multiple small unmanned surface vehicles (USVs) in a swarm for marine research purposes at sea. The study aimed to overcome limitations in acquiring broadband data in a wide range and diverse missions at sea by using swarm control techniques. The study used a long short-term memory (LSTM) model to learn disturbance information and predict maritime disturbances. The predicted ocean currents were used to generate a swarm USV control system for USV formations. The study conducted a comparative analysis of the designed USV model results and those generated by the simulator, and the effectiveness of the USV mathematical model and behavior control rules were verified. The system could potentially contribute to the exploration of marine data and resources

Keywords—Unmanned surface vehicles (USV), Recurrent Neural Network (RNN), Long Short-Term Memory models (LSTM), Artificial neural networks PID(ANN-PID)

I. INTRODUCTION

To overcome these limitations, this dissertation focuses on developing a system for intensive mission performance in a wide area through swarm control of multiple small USVs. The study employs the long short-term memory (LSTM) algorithm of recurrent neural networks (RNN) to learn the disturbances that affect movement control.[1][2]Using the learned model, a method to predict future disturbances and improve the performance of the artificial neural network-PID (ANN-PID) control algorithm using the predicted disturbance data are examined. The study aims to develop a swarm control system that can effectively manage multiple small USVs for various marine research and exploration purposes.[3]

II. BEHAVIORAL CONTROL OF USV SWARM

To control the behavior of a swarm of USVs, it is necessary to understand the position and distance traveled by each USV. A small USV equipped with two rear thrusters and a front lateral thruster was used in this study.[4]

A. Behavioral control of USV swarm

To control the behavior of the USV swarm, each USV uses GPS. As shown in Fig. 4 of the swarm USVs, the searchable radial distance of each USV is represented by d_r. Here, i is the number of USVs i ranging from 1 to n. j represents the surrounding USV.[3]

Fig. 1 shows the behavior of the swarm USV according to the measured distance d of the adjacent USV. As shown in Fig. 2, the distance that maintains the maximum search radius with the adjacent USV is called d_{max}. Furthermore, Fig. 2shows the minimum search radius distance to avoid collision with an adjacent USV as d_{min}.

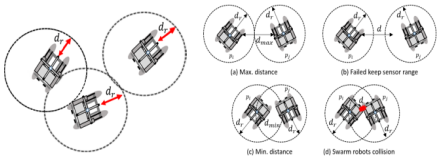


Fig. 1. Behavior of swarm USVs by distance.

TABLE 1. THE BEHAVIOR OF SWARM ROBOT BY MEASURED

	The behavior of swarm USVs
$d = d_{max}$	keep the maximum distance (Fig 2(a))
$d > d_{max}$	Move to adjacent USV(Fig 2(b))
$d = d_{min}$	Maintaining the minimum distance between USVs(Fig 2(c))
$d < d_{min}$	Avoid collisions with adjacent USVs(Fig 2(d))

B. Behavioral rules for USV swarm

For each robot to search the space using the adjacent robot while the search radius in the unknown space, the area is divided into Zone of Repulsion (ZoR), Zone of Orientation (ZoO), and Zone of Attraction (ZoA), as shown in Fig. 4 Within the searchable radius of each robot, an action rule suitable for each area is planned.

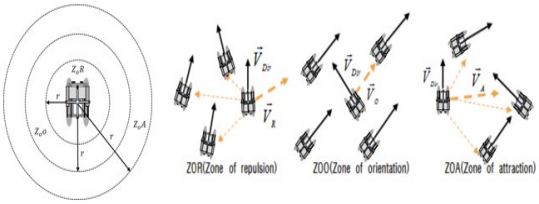


Fig. 2. Property of behavior rule for swarm USVs

To control the robot's behavior for each area in the ZoR area, the swarm robot maintains a minimum distance to avoid collision with an adjacent robot. In the ZoO area, the robot moves in the same direction as the average direction of the moving swarm robots to maintain a constant distance from an adjacent robot. In addition, because the swarm robots are dispersed in the ZoA area, they move in the direction where adjacent robots exist. The behavior rule for controlling the behavior of the swarm robot in each of these areas is expressed in Equations (1-3) and can be implemented by calculating the direction vector in each area. In the N_Zor area, the direction vector of the USV is expressed using the position of each robot as in Equation (1).

$$\vec{v}_R = \begin{cases} \vec{v}_{Dir} & , N_{ZoR} = 0 \\ -\frac{1}{N_{ZoR}} \sum_{j=1}^{N_{ZoR}} \frac{(p_j - p_i)}{|p_j - p_i|} & , N_{ZoR} > 0 \end{cases} \quad (1)$$

Here, p_i denotes the position (x_i, y_i) of the USV i at sea, and p_j denotes the position of the surrounding USV (x_j, y_j). $\vec{v}_{Dir}$ represents the moving direction of the USV i. N_{ZoR}, N_{ZoO}, N_{ZoA} are the number of USVs in each domain. To determine the direction of USV i in the domain N_{ZoO}, it is expressed using the direction vector of the surrounding USV j, as in Equation (x).

$$\vec{v}_O = \begin{cases} \vec{v}_{Dir} & , N_{ZoO} = 0 \\ \frac{1}{N_{ZoO}} \sum_{j=1}^{N_{ZoO}} \frac{\vec{v}_j}{|v_j|} & , N_{ZoO} > 0 \end{cases} \quad (2)$$

In the domain of N_{ZoA}, the direction vector is expressed using the position of each USV, as in Equation (4).

$$\vec{v}_A = \begin{cases} \vec{v}_{Dir} & , N_{ZoA} = 0 \\ \frac{1}{N_{ZoA}} \sum_{j=1}^{N_{ZoA}} \frac{(p_j - p_i)}{|p_j - p_i|} & , N_{ZoA} > 0 \end{cases} \quad (3)$$

In addition, to prevent overlapping actions for the N_{ZoR}, N_{ZoO}, N_{ZoA} domains of each USV, priority is assigned according to the existence of USVs in each domain. To determine the priority of each robot's action, if USV exists in the area of N_{ZoR} as in Equation (4), other actions are restricted, and the appropriate action is considered in the area N_{ZoR}. If N_{ZoO} and N_{ZoA} exist simultaneously, the final direction vector can be obtained as the average of $\vec{v}_O$, and $\vec{v}_A$,.

$$\vec{v}_c = \begin{cases} \vec{v}_R, & if \quad N_{ZoR} > 0 \\ \vec{v}_O, & if \quad N_{ZoR}, N_{ZoA} = 0 \wedge N_{ZoO} > 0 \\ \vec{v}_A, & if \quad N_{ZoR}, N_{ZoO} = 0 \wedge N_{ZoA} > 0 \\ \frac{1}{2}\left(\vec{v}_O + \vec{v}_A\right), & if \quad N_{ZoR} = 0 \wedge N_{ZoO}, N_{ZoA} > 0 \end{cases} \quad (4)$$

If the behavior rules of USV swarm are planned for each domain of ZoR, ZoO, and ZoA, and overlapping actions are prevented for each domain, the swarm robot avoids collisions between adjacent USVs to cooperatively navigate the space and an optimal search radius can be obtained and maintained while moving. In addition, if there is a robot that is separated from the swarm, it can move to an adjacent USV and maintain its formation. Therefore, when moving to the target point, the swarm behavior can be controlled by avoiding collisions with adjacent USVs from the initial position and moving while maintaining its position and range with adjacent USVs.

III. Learning Ocean Currents of USV Swarm

To control the movement of multiple small USVs that form a swarm, it is necessary to analyze and predict the control system as well as many external forces of the ocean as control factors. [5] Some of the factors are the ocean current, wind, and waves. In this study, learning and prediction were made based on the measured data of ocean currents, which have the greatest influence among these factors.[6]

A. Analysis of Ocean Currents

Ocean currents are environmental factors that significantly affect objects moving in the sea according to the shape of the vessel as the vessel moves. This increases the control error during missions because a constant external force on the vessel is applied. The current interferes with the movement of multiple USVs forming the swarm, making it difficult to successfully complete the mission. If these external disturbance factors are analyzed and used as control factors, the stability and effectiveness of a multiple USV control system forming a swarm will increase. To measure these ocean currents, a sensor was installed in the USV to measure the velocity and $Current$ angle of the external force. To measure the flow velocity, GPS can be used to measure ground speed, yaw angle, and position (latitude and longitude). The flow velocity sensor can measure the speed (water velocity) at which the hull moves in the fluid with the flow velocity. Therefore, the speed of ocean currents is measured by sensors mounted on each of the multiple USVs moving above sea level. The measured data can be expressed by Equations 5 and 6.

$$V_G = V_w + V_c \quad (5)$$

By decomposing the vector of magnitude and direction into x and y directions, the component of the three velocities can be expressed as Equation 6 by the relationship between the magnitude of the vector and the Earth's fixed coordinate system.

$$\begin{aligned} V_G &= [|V_G|\cos\psi, |V_G|\sin\psi] \\ V_w &= [|V_w|\cos\psi_w, |V_w|\sin\psi_w] \\ V_c &= [|V_c|\cos\psi_c, |V_c|\sin\psi_c] \end{aligned} \quad (6)$$

B. Learning Ocean Currents

The model learned the measured time-series ocean current data using LSTM, predicted the next ocean current speed, and applied it as a disturbance to each USV control algorithm.

The measured data are normalized and converted into data having a size between 0 and 1. In Equation 7, three gates are represented by f_t, I_t, O_t and two outputs are represented by h_t, C_t. Furthermore, the parameters of LSTM are W_i, W_f, W_o, W_c. Finally, $\widetilde{C_t}$ is a new output of the input gate, and the output of I_t is determined according to the degree of reflection. Empirically, the function softsign was used as the activation function, and the Adam Optimizer was applied.

$$\begin{aligned} f_t &= sigmoid(W_f \cdot [h_{t-1}, x_t] + b_f \\ I_t &= sigmoid(W_I \cdot [h_{t-1}, x_t] + b_I \\ O_t &= sigmoid(W_O \cdot [h_{t-1}, x_t] + b_O \\ h_t &= O_t * \text{softsign}(C_t) \\ C_t &= f_t * C_{t-1} + I_i * \widetilde{C_t} \\ \widetilde{C_t} &= \text{softsign}\,(W_c \cdot [h_{t-1}, x_i] + b_c \end{aligned} \quad (7)$$

The data from the ocean current measuring sensor operated in the sea were used in the LSTM model constructed in this study and applied to predict future disturbance.

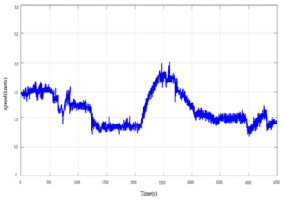


Fig. 4 Ocean current measurement data.

Fig. 4 is the actual measurement data from the ocean. Data measurements were taken for 75 minutes. The graph in Fig. 10 was derived by learning the actual data. In the graph of

Fig. 5, green indicates 1800 data out of 4500 actually measured data, and these data were used for training and testing. In addition, blue is actual data and red is data predicted by learning the previous 1800 data points.

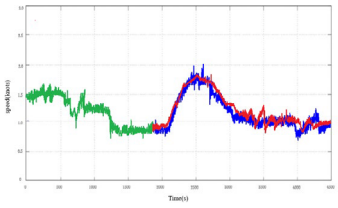


Fig. 5 Ocean current prediction data result.

From the predicted results (red) in Fig. 5, there is an error in the section where the measured data changes rapidly; however, the error between the actual data and the predicted data showed an average error of 0.12 knots.

IV. Control Algorithm of USV Swarm

To control the movement of multiple small USVs that form a swarm, it is necessary to analyze and predict the control system as well as many external forces of the ocean as control factors. Some of the factors are the ocean current, wind, and waves. In this study, learning and prediction were made based on the measured data of ocean currents, which have the greatest influence among these factors.

A. USU Control Algorithm

In this study, because the velocity of ocean currents is learned and used as a controlling factor of small USVs forming a swarm, the control algorithm must design a controller that can respond appropriately to external information. Therefore, the controller applied to the small USV was modeled by ANN-PID.

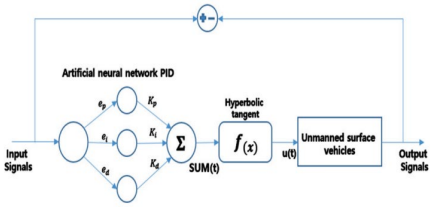


Fig. 6. Artificial neural network PID.

The three inputs in Fig. 6 are the error value, error integral value, and error differential value used as the basis of the PID control algorithm, which is input into a nonlinear hyperbolic tangent function.

When the error is adequately reduced and the control performance of a certain range for the stability of the system is satisfied, the algorithm is stopped and the calculated gain value is maintained.

B. USV Running Algorithm

The navigation algorithm of a small USV is following a target point or driving toward a specific direction. This method is called line of sight (LOS), where GNSS-based navigation is used. USV driving in this study was also designed based on the LOS algorithm.[7]

$$\psi_p = tan^{-1}\left[\frac{Y_k - Y(t)}{X_k - X(t)}\right]$$

$$\rho^2(t)=[X_k - X(t)]^2 + [Y_k - Y(t)]^2 < \rho_c^2 \quad (8)$$

In Equation 8, $[X(t), Y(t)]$ is the position of the USV, and $[X_k, Y_k]$ is the target position. In addition, after reaching the target position, ρ_c denotes a position radius for determining whether the next designated position has been reached.

V. Control System of USV Swarm with the Learning Ocean Current Model

The small USV used in this study determines the direction of travel by calculating the designated position and its heading angle using a propellant placed laterally in the front. [8]When a certain range of the designated position is reached, the lateral propellant is controlled by calculating the heading angle again to move to the next position. Meanwhile, the aforementioned ANN-PID control algorithm is applied for precise control. [9][10] The influence of external force is learned and predicted during movement, and this is applied to the control system to design a control system that can move to the exact position.

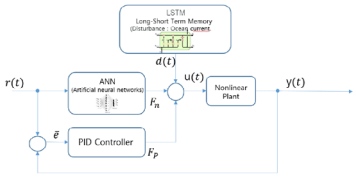


Fig. 7. ANN-PID Control Algorithm Using learning Ocean Current Model

The article describes a control algorithm for improving the performance of an unmanned surface vehicle (USV) swarm by predicting ocean currents and applying the predicted data to an artificial neural network (ANN)-proportional integral derivative (PID) control algorithm. The control algorithm consists of two parts: the first part is an ANN that predicts the magnitude of the next disturbance using measured ocean current data, while the second part is the ANN-PID control algorithm that adjusts the control gain based on feedback weights. The USV control algorithm combines these two learning models by adding the predicted ocean current components to the weight-adjusted output value of the ANN-PID control algorithm to determine the next output value delivered to the system.

VI. Validation of the Swarm Control Algorithm

The verification was conducted in two ways.

1) By analyzing the similarity between the actual driving path of the USV swarm at the sea and the USV swarm path generated by the simulator, including the behavior control rules of the USV swarm and USV designed in this study, the mathematical model of the designed USV and behavior control rules are verified for accuracy.

2) To check the validity of the ocean current learning algorithm and ANN-PID control algorithm, the position error was obtained by comparing the movement path of the USV swarm with and without the disturbance learning algorithm, and the ANN-PID control algorithm applied to the simulator designed in this study.

The position, movement speed, and direction of each USV were used in a simulation vessel design based on driving data during real sea operation. Fig. 15 shows the results created by applying the real driving information to the simulator, including the designed USV model and behavior control rules of the USV swarm.

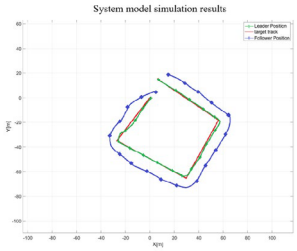


Fig. 8. System model simulation results(verification_1)

The article discusses a comparison between real sea test data and a designed simulator's results, which revealed some discrepancies due to environmental conditions like wind, tides, and waves. Despite this, the movement trajectory of the real sea test and simulator were found to be similar. The article concludes that the USV model and behavior control rules designed to have a position average error of less than 10% can be considered an adequate simulation of a USV swarm's operation at sea. Fig. 9 shows a graph comparing the disturbance learning algorithm and the ANN-PID control algorithm when a disturbance (Ocean Currents) with a specific direction is applied to the operating environment.

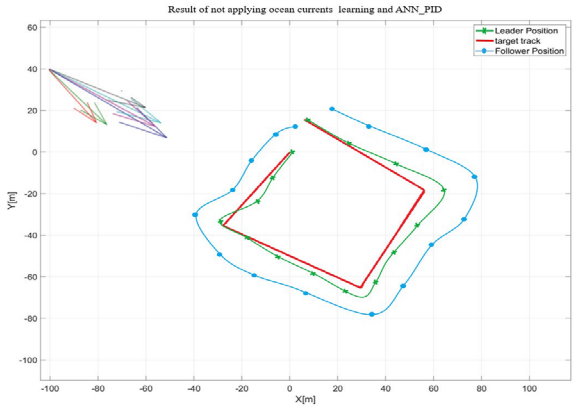


Fig. 9(a). Not applying ocean currents learning and ANN_PID(verification_2)

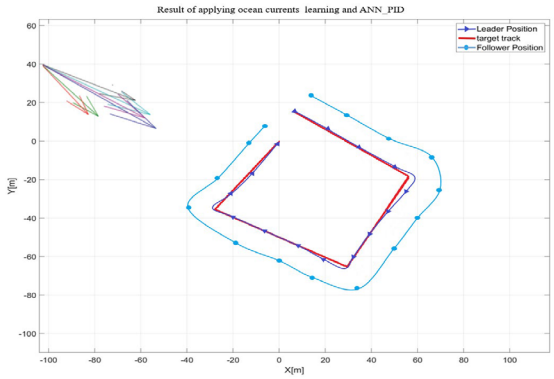


Fig. 9(b). applying ocean currents learning and ANN_PID(verification_2)

As a result of swarm operation by applying the disturbance learning algorithm and ANN-PID control algorithm, the large retention performance between USVs was increased by 12.4% and decreased position average error by 14.8% compared to the result of not applying the disturbance learning algorithm.

VII. CONCLUSIONS

The study proposes an algorithm that uses artificial neural networks (ANN) and proportional-integral-derivative (PID) control to improve the control performance of unmanned surface vehicle (USV) swarms in ocean environments. The algorithm learns and predicts ocean currents as an external force and applies it to the controller output. The study uses a 3-degree-of-freedom USV model with a fluid force to verify the algorithm and compare the control performance with and without the algorithm. The results show that applying the predicted data of external force increased the formation maintaining performance by 12.4% and decreased the position error during movement by 14.8%. The study plans to add other external factors like waves and wind for effective swarm control in the future.

ACKNOWLEDGMENT

“ This research was a part of the project titled 'Development of smart maintenance monitoring techniques to prepare for disaster and deterioration of port infra structures', funded by the Ministry of Oceans and Fisheries, Korea.". Also, hhis research was supported in part of the project titled `Development of Polar Region Communication Technology and Equipment for Internet of Extreme Things (IoET)', funded by Ministry of Science and ICT (MSIT)

REFERENCES

[1] Dae-Hyeong Ji, Hyeung-Sik Cho, Sang-Ki Jeong, Ji-Youn Oh, Seo-Kang Kim and Sam-Sang You, A study on heading and attitude estimation of underwater track vehicle, Advances in Technology Innovation 4(2) (2019), 84–93.

[2] Asgeir J. Sorensen, A server of dynamic positioning control systems, Annual Reviews in Control 35(1) (2011), 123–136.

[3] I. S. Jacobs and C. P. Bean, “Fine particles, thin films and exchange anisotropy,” in Magnetism, vol. III, G. T. Rado and H. Suhl, Eds. New York: Academic, 1963, pp. 271–350.

[4] T.I. Fossen, Guidance and Control of Ocean Vehicles, John Wiley & Sons Ltd, 1994. S.K Jeong, H.S Choi, J.I Kang, J.Y Oh, and S.K. Kim, Design and control of navigation system for hybrid underwater glider, Journal of Intelligent & Fuzzy Systems 36 (2019), 1057–1072.

[5] Jeong, Sang-Ki, Ji Dea-Hyeong, Oh, Ji-Youn, Seo Jung-Min, Choi Hyeung-Sik, Disturbance learning controller design for unmanned surface vehicle using LSTM technique of recurrent neural network, journal of intelligent and fuzzy system,2021;40.4;8001-8011

[6] Y. Yorozu, M. Hirano, K. Oka, and Y. Tagawa, “Electron spectroscopy studies on magneto-optical media and plastic substrate interface,” IEEE Transl. J. Magn. Japan, vol. 2, pp. 740–741, August 1987 [Digests 9th Annual Conf. Magnetics Japan, p. 301, 1982].

[7] T.I. Fossen, Handbook of Marine Craft Hydrodynamics and Motion Control; John Wiley & Sons: Hoboken, NJ, USA, 2011.

[8] Sepp Hochreiter, Bengio Yoshua, Frasconi Paolo and Jurgen ¨ Schmidhuber, Gradient flow in recurrent nets: the difficulty of learning long-term dependencies. In Kremer and Kolen, editors, A Field Guide to Dynamical Recurrent Neural Net works. IEEE Press, 2001.

[9] S.K Jeong, H.S Choi, J.I Kang, J.Y Oh and S.K. Kim, Design and control of navigation system for hybrid underwater glider, Journal of Intelligent & Fuzzy Systems 36 (2019), 1057–1072.

[10] Shin DH, Bae SB, Beak WK, et al. Way-Point Tracking of AUV using Fuzzy PD Controller. *Journal of KIIT* 2013; 11.5: 1-

A Study on Swarm USV's Equipped with Fault Coping Algorithm

1st Jihyeong Lee
Maritime ICT·Mobility
Research Department
KIOST
Busan, Republic of Korea
wlgud0818@kiost.ac.kr

2nd Sang Ki Jeong
Maritime ICT·Mobility
Research Department
KIOST
Busan, Republic of Korea
jeongsk313@kiost.ac.kr

3rd Saehun Baeg
Vessel Operation & Observation Team
KIOST
Busan, Republic of Korea
saehun@kiost.ac.kr

***Abstract*— In this study, to overcome the limitations of a single unmanned surface vehicle(USV), a study was conducted on the swarm control algorithm of the USV. Among the various swarm control methods, a leader-follower swarm control method was selected and studied, and in order to solve the fatal problem of the leader-follower swarm control method, a study on the fault-coping algorithm was conducted together.**

Keywords—Swarm control algorithm, Unmanned Surface Vehicle, Leader-Follower swarm control method, Fault-Coping Algorithm

I. INTRODUCTION

The multi-entity system using swarm control has a wider range of mission applications such as search and environmental investigation than single-entity systems. In addition, swarm control algorithms are receiving a lot of attention recently because they have various advantages such as improved mission performance, increased efficiency, and energy saving by using multiple objects at the same time[1-3]. In studies on existing swarm control, methods are classified into behavior-based swarm control, virtual structure swarm control, and leader-follower swarm control[4]. Among them, the leader-follower swarm control is a control method in which, when the leader unmanned mobile vehicle moves along a preset trajectory, the following unmanned mobile vehicles move while maintaining a certain distance and heading angle from the leader unmanned mobile vehicle. This method has the advantage of being easy to understand mathematically of the swarm control algorithm, easy to implement, and theoretically, possible to increase the number of followers without limit. However, there is a disadvantage that there is a fatal problem for the entire swarm formation in the event of a failure of the leader unmanned mobile vehicle[5-6]. Therefore, in this study, we propose a new algorithm to overcome the disadvantages of the leader-follower swarm control method, which has advantages do.

II. LEADER-FOLLOWER BASED SWARM MODELING

The USV(Unmanned Surface Vehicle) used in this study uses a method of moving by independently rotating two thrusters, port side, and starboard side. At this time, assuming that the thruster does not cause slip or cavitation in the water, the USV has the non-holonomic constraint condition as shown in Equation (1).

$$\begin{bmatrix}\dot{X}_c(t)\\ \dot{Y}_c(t)\\ \dot{e}_c(t)\end{bmatrix} = \begin{bmatrix}\cos(\theta_c(t)) & 0\\ \sin(\theta_c(t)) & 0\\ 0 & 1\end{bmatrix}\begin{bmatrix}v_c(t)\\ w_c(t)\end{bmatrix} \quad (1)$$

Here, $X_c(t), Y_c(t)$ indicates the position of the USV on the X-axis and Y-axis, θ_c means the heading angle of USV. v_c, w_c means the linear velocity and angular velocity of the USV[8].

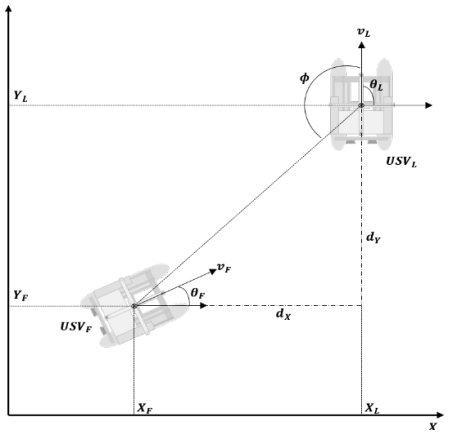


Fig. 1. Kinematic model of USV.

In Fig.1, the leader USV is expressed using the subscript 'L', and the follower USV is expressed using the subscript 'F'. X_L, Y_L are the current position of the leader USV, and θ_L is the heading angle of the leader USV. Likewise, X_F, Y_F are the position of the follower USV, and θ_F is heading angle of the follower USV. v_L, v_F are the straight speed of each USV, and w_L, w_F represents the angular velocity of each USV.

The role of the follower USV is to form a formation with the leader USV while maintaining a certain distance and angle from the mission trajectory calculated by the leader USV. That is, a position having a specific distance and a specific angle designated by the land operation console is created from the position of the leader USV and the follower USV follows the trajectory. Equation (2) is an expressing the position of the formation followed by the follower USV from the leader USV.

$$\begin{aligned} X_F^d(t) &= X_L(t) + d_{ref}\cos(\phi_{ref} + \theta_L(t)) \\ Y_F^d(t) &= Y_L(t) + d_{ref}\sin(\phi_{ref} + \theta_L(t)) \\ \theta_F^d(t) &= \theta_L \end{aligned} \quad (2)$$

Here, $X_F^d(t), Y_F^d(t)$ represents the position that the follower USV should follow, and d_{ref} is the distance of the trajectory that the leader USV and the follower USV should follow. In this case, since the distance cannot have a (-) value, it is always expressed as a positive number. ϕ_{ref} means the angle to be maintained from the leader USV, and this value is expressed as a constant, not a function.

By differentiating Equation (2), the tracking USV coordinate change can be obtained, which can be expressed as Equation (3).

$$\dot{X}_F^d(t) = v_L(t)\cos(\theta_L(t)) - d_{ref}\sin(\phi_{ref} + \theta_L(t))w_L(t)$$
$$\dot{Y}_F^d(t) = v_L(t)\sin(\theta_L(t)) + d_{ref}\cos(\phi_{ref} + \theta_L(t))w_L(t) \quad (3)$$

III. FAULT-COPING ALGORITHM

In this study, we propose a fault-coping algorithm to overcome one of the biggest disadvantages of the leader-follower swarm control method, which is that it is impossible to maintain the entire formation in the event of a failure of the leader USV. The failure situation considered in this study is a situation in which the thruster does not rotate normally due to marine floating matter being caught in the thruster, which is the most frequent failure situation in the offshore platform.

Fig. 2 shows the block diagram of the algorithm proposed in this study.

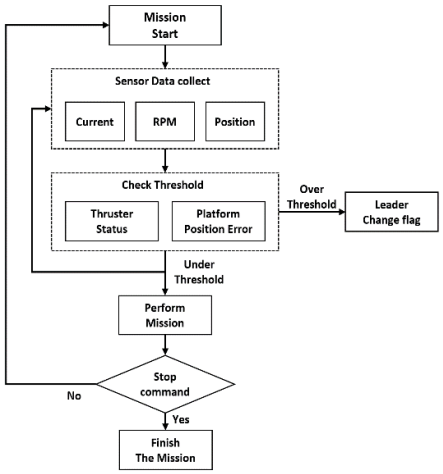


Fig. 2. Block diagram of fault-coping algorithm

The algorithm works in the following order. When the mission starts, data such as the total current amount of the platform, RPM value of the thruster, and location information of the USV are acquired using sensors mounted on the USV. Then, by inputting the acquired data into the Check Threshold block, it is judged whether the state of the data is higher or lower than the threshold value. If the data status is similar to or lower than the threshold value, it is judged to be normal and the mission is continued. At this time, the mission is repeatedly performed unless a mission stop command is input from the operator. However, if the data state is greater than the threshold value, it is judged as a failure state, and the Leader Change flag is raised. When the Leader Change flag goes up, the USV that has a malfunction stops, and its authority and mission are transferred to the USV that has the closest distance among USVs that have not currently malfunctioned. In the algorithm, the Check Threshold block compares the current and voltage applied to the thruster, and the RPM value output from the thruster with the arbitrarily set Threshold value, and the Thruster Status block, which determines whether or not there is a failure, and the current location information of the platform, heading angle, azimuth information, etc. It is designed by combining the Platform Position Error block that determines whether there is a failure by comparing it with the set threshold value.

IV. RESULT OF ACTUAL SEA AREA TEST

The performance test of the fault-coping algorithm was conducted in the actual sea area. At the time of the experiment, the sea state 3 with a wave height of 0.5~1m, a swell height of 0.6m, and a swell period of 8 seconds. The fault-coping algorithm normally identifies the sudden change in current amount and RPM value that occur in a fault situation where marine floating objects are caught in the thruster, and the failure is judged in the Check Threshold block, and the leader change flag is raised normally, so that the follower USV took over the role of the leader USV and carried out its mission. Among the tracking results of the fault-coping algorithm shown in Fig. 3.

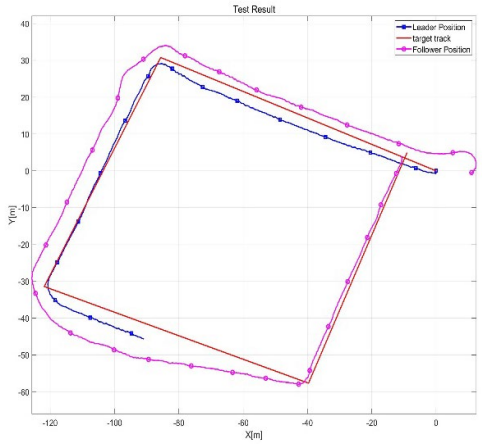


Fig. 3. Tracking result of fault-coping algorithm test

V. CONCLUSION

This research conducted a study on the fault-coping algorithm to overcome the disadvantage of a leader-follower swarm control method. For the study, we have developed USV to perform actual missions, and we have developed a swarm control algorithm using a fault-coping algorithm. In addition, the performance of the algorithm proposed in this study was verified through the test in the actual sea area.

ACKNOWLEDGMENT

This research was supported in part of the Korea Institute of Ocean Science and Technology under Grant PEA0132.

REFERENCES

[1] D.Cruz, J. Mcclintock, B. Perteet, A.A.O.Orqueda, Y.Cao and R.Fierro, "Decentralized cooperative control", Proceeding of IEEE Control System Magazine, 2007, pp. 58-78.

[2] J.T.Feddema, C.Lewis, and D.A. Schoenwald, "Decentralized control of cooperative robotic vehicles: theory and application", Proceeding of IEEE Trans. on Robotics and Automation, 2002, vol.44(5), pp 1343-1349

[3] Y.Gu, B.Seanor, G.Campa, M.R.Napolitano, L.Rowe, S.Gururajan and S.Wan, "Design and flight testing evaluation of formation control laws", Proceeding of IEEE Trans. On Control Systems Technology, 2006, vol.14, pp. 1105-1112

[4] K.R.Lee, "Design of Decentralized Behavior-Based Network for the Formation Control and the Obstacle Avoidance of Multiple Robot", Proceeding of Ajou University, 2018

[5] S.B.Park, J.H. Park, "A study on the formation control algorithms of Multi-Obstaces collision avoidance for autonomous Navigation of swarm marine unmanned moving vehicles", Proceeding of Journal of the KNST, 2020, 3(1), pp. 56-61.

[6] M.H.Tak, J.S.Kim, Y.H.Joo, S.H.Ji, "Formation Control for the Obstacle Avoidance of Swarm Robots based Leader-Follower Robots", Proceeding of the Korean Institute of Electrical Engineers, 2013, 10, pp. 171-172

Study on The Atmospheric Measurement Monitoring System of a Smart Buoy Robot Equipped With a SAW-based Fine Dust Sensor

Dong-Woo Man
Maritime ICT & Mobility Research Department
KIOST
Busan, South Korea
mandongw@kiost.ac.kr

Hyun-Sik Kim
Department of Mechanical and Robotics Engineering
Tongmyong University
Busan, South Korea
hyunskim@tu.ac.kr

Sang Ki Jeong
Maritime ICT & Mobility Research Department
KIOST
Busan, South Korea
jeangsk313@kiost.ac.kr

Changjoo Shin
Maritime ICT & Mobility Research Department
KIOST
Busan, South Korea
cjshin@kiost.ac.kr

O-Soon Kwon
Maritime ICT & Mobility Research Department
KIOST
Busan, South Korea
oskwon@kiost.ac.kr

Se-Yong Park
Department of Advanced Materials Engineering
Tech University of Korea
Siheung-si, Gyeonggi-do, South Korea
psy881024@tukorea.ac.kr

***Abstract*— As interest in air pollution increases, research on how to measure fine dust in advance is needed. It is known that fine dust is mostly introduced through the West Sea. Therefore, measuring at sea is more accurate and has the advantage of being able to measure fine dust in advance. Accordingly, this paper proposes a method for measuring fine dust at sea using a SBR(Smart Buoy Robot). The proposed system has advantages such as power consumption, size, and accuracy by using a SAW(Surface Acoustic Wave)-based fine dust sensor.**

Keywords—Smart Buoy Robot, Fine dust, Sensor, Robot, Surface Acoustic Wave

I. INTRODUCTION

Recently, interest in air pollution is increasing both domestically and internationally. Accordingly, various studies related to fine dust sensors are being conducted in Korea. Currently, it is understood that fine dust is mainly introduced through the West Sea in Korea. Therefore, for a more accurate study of the inflow and distribution of fine dust, it is better to measure it at sea[1].

Atmospheric measurement results require cumulative data for analysis. Therefore, it is necessary to continuously measure fine dust, but it is difficult to install a system that can continuously measure at sea due to problems such as cost and power supply.

Therefore, in this paper, we propose a marine fine dust measurement system. It is a method of collecting and measuring fine dust sensors by attaching them to buoys that are common in the sea. Existing buoys are fixed in one place through anchoring, so there is a problem in that they do not have a certain direction due to wind and waves[2][3].

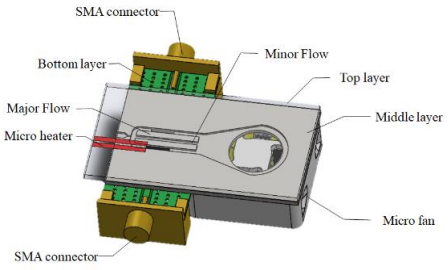
To solve this problem, this paper proposes a measurement method using SBR(Smart Buoy Robot). Since SBR is capable of self-propulsion, it has the advantage of being able to collect in the desired direction through position control[4]. The MEMS-based FBAR (Film Bulk Acoustic Resonator) fine dust sensor has high sensitivity, but has the disadvantage that the process of making the sensor is very complicated[5]. However, based on SAW (Surface Acoustic Wave) Resonator, it is relatively easy to process and suitable for mass production. Therefore, in this paper, the installation of a SAW-based fine dust sensor was considered [6].

This research was supported in part of the Korea Institute of Ocean Science and Technology under Grant PEA0132.

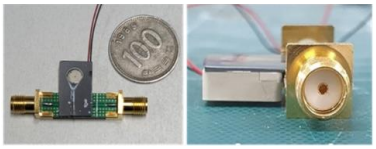
II. SYSTEM CONFIRUGATION

A. Fine dust sensor base on SAW

The conceptual diagram of the components constituting the SAW-based fine dust sensor used in this study is shown in same as Figure 1 (a). Figure 1(b) is the fabricated fine dust sensor. The manufactured fine dust sensor was produced in a small size of 20 × 22 × 5mm(w × l × h).



(a)



(b)

Fig. 1. (a) The schematic diagram of Fine dust sensor based on SAW, (b) prototype Finde dust sensor

The configuration of the fine dust sensor is composed of an uppermost layer including a micro-heater first and a middle

layer including a virtual impactor and having a flow path for controlling the flow of the entire fluid. The third is the bottom layer where the PCB with the SAW resonator is mounted and the SMA connector is mated. Finally, there are 4 micro-fans to determine the total flow[6].

The amount of fine dust according to the inflow air flow rate was measured through a small Grimm type optical dust sensor (HI-100, HUMAi, Korea). As shown in Figure 6, it can be seen that the system is stable by showing the result that the fine dust measurement value increases linearly as the flow rate increases.

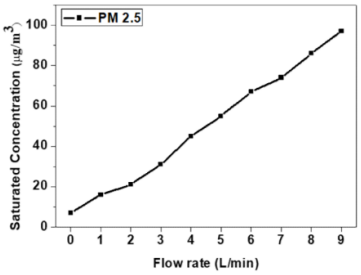


Fig. 2. Change in the amount of fine dust measured in the chamber according to the air flow rate

B. SBR with fine dust sensor

The SBR used in the experiment is composed as shown in Figure 2 below. The current prototype of SBR is composed of non-standard buoys to facilitate the charging of electric ships and marine robots. The cross-sectional area was made in an X shape to make it easy for ships to anchor.

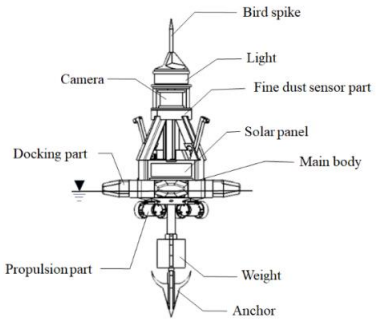


Fig. 3. Fine dust sensor installed SBR

A light that informs the location with light, which is the main purpose of SBR, is attached to the top, followed by a camera for control and situation identification, and a SAW-based fine dust sensor.

III. CONTROLLER DESIGN

SBR controls the bow direction using AHRS (Attitude Heading Measurement System) and controls the propeller output value using the PD controller using the current position error value at the target position.

The PA (Position Accuracy) discriminant is used to control the position of the SBR. When the target position (RP: Reference Position) x_r and y_r are determined, the PA discriminant related to the position control performance index is defined as in Equation 1 as follow:

$$r_i = \sqrt{(x_i - x_r)^2 + (y_i - y_r)^2} \quad (1)$$

In this equation, PA_m means the maximum position error value of the buoy robot. x_i and y_i are the position values of the SBR at the ith sampling point, and r_i is the radius of the SBR at the i-th sampling point.

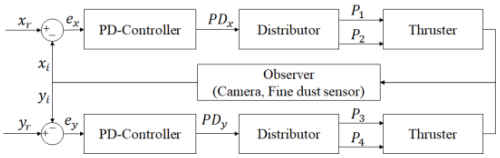


Fig. 4. Block diagram position control system

Figure 3 shows the block diagram of the PD controller for position control of the SBR. In Figure 3, e_x is the x-coordinate error, and e_y is the y-coordinate error. PD-Controller derives PD_x, PD_y with x, y errors, and inputs control values to each motor p_1, p_2, p_3, p_4, through Distributor.

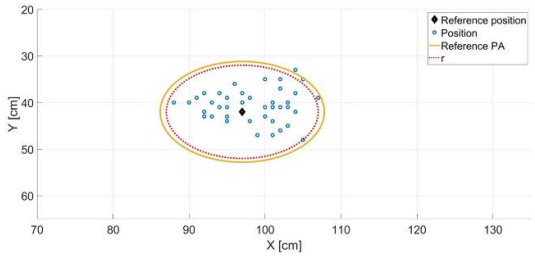


Fig. 5. Performance of position control used a camera instead GPS

Looking at the result in Figure 4, the target position is (97,42) in camera coordinates and the reference PA has a radius of 7.1cm. The position control was performed inside the circle related to the discriminant, and considering that the small circles are concentrated near the target point, it can be seen that the position control is well performed.

IV. CONCLUSION

In this study, a method for measuring fine dust on the sea is proposed. The proposed system has strengths in the limited power problem due to the size of the marine robot by using the SAW type fine dust sensor, which has advantages such as cost reduction, reduced power consumption, and ultra-small size compared to the existing optical fine dust sensor. In addition, it is easy to collect marine fine dust by using an SBR capable of position control. Since only experiments on individual systems have been carried out so far, we plan to

conduct fine dust measurement experiments in the actual sea in the future.

ACKNOWLEDGMENT

This research was supported in part of the Korea Institute of Ocean Science and Technology under Grant PEA0132.

REFERENCES

[1] Sean Shi-An, Lim, and Harini Manivannan. "A Comparative Study on the Characteristics of the Inflow Path of Fine Dust and the Trend of Climate Change in Korea and Singapore Using Satellite Data" IRC-SET 2022: Proceedings of the 8th IRC Conference on Science, Engineering and Technology, August 2022, Singapore. Singapore: Springer Nature Singapore, 2023.

[2] H. G. Park, H. S. Kim, “Mechanism Development and Position Control of Smart Buoy Robot”, Journal of Ocean Engineering and Technology, Vol. 35, No. 4, pp. 305-312, 2021.

[3] H. G. Park, “Study on Development and Position Control of Smart Buoy Robot”, Master thesis, 2021.

[4]

[5] D. W. Man, C. J. Shin, J. M. Seo and O. S. Kwon, “A Study on the Detection and Position Control of Thruster Failure of Smart Buoy Robot”, Journal of the Korea Academia-Industrial cooperation Society, Vol. 24, No. 5, pp. 326-332, 2023.

[6] I. Paprotny, F. Doering, and R. M. White, “MEMS Particulate Matter (PM) monitor for cellular deployment”, SENSORS, 2010 IEEE, Waikoloa, HI, USA, pp. 2435-2440, 2010

[7] S. Y. Park, “A Study on design and fabrication of PM sensor and gas sensor based on computational fluid dunamycis”, Ph.D thesis, 2022.

[8] S. Y. Park, M. J. Kim and H. C. Lee, “”Computational Fluid Dynamics Simulation, Microelectromechanical System Fabrication, and Radio-Frequency Evaluation of the PM2.5 Fine Dust Sensor Based on the Surface Acoustic Wave Resonator”, Journal of Nanoelectronics and Optoelectronics, Vol. 16, pp. 1-9, 2021.

Automatic Lifting Cable Control System for Large Circular Steel Pipes based on Verticality Monitoring System

Sungmin Koo
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
smkoo@kiost.ac.kr

Haeyong Park
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
hypark@kiost.ac.kr

Myounghak Oh
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
omyhak@kiost.ac.kr

Seungjae Baek
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
baeksj@kiost.ac.kr

***Abstract*—In this paper, we propose an automatic lifting cable control system for large circular steel pipes based on the verticality monitoring system. Specifically, the system includes a monitoring component that provides real-time data on the pipe's alignment and an actuator control module that maintains perpendicularity by utilizing the data acquired from the monitoring system. The system is implemented with a tilt sensor, a hydraulic actuator, and an embedded controller. Experimental results show that the proposed system can efficiently measure the roll/pitch information in real-time. Furthermore, we show that the system properly operates and secures a verticality less than 0.2° in the suction bucket-based model experiment.**

Keywords—suction bucket, lifting cable, automatic control system, verticality, monitoring

I. INTRODUCTION

A suction bucket foundation is a construction method that first locates an empty upturned bucket to a target site, then installs the foundation by suctioning the air and water inside the bucket [1]. By applying a suction pressure to the bucket greater than the penetration resistance, the bucket foundation penetrates into the ground due to the difference in water pressure between the inside and the outside. The suction bucket has traditionally been used as a mooring anchor [2, 3]. But, many studies have been conducted for using offshore structures such as the foundations of offshore wind turbines [4-17] and temporary structures for the construction of offshore bridges [18] because of the advantages of fast construction speed, simplicity, and less noise compared to the impact piling method at depths of more than 20 meters.

Securing the verticality of the bucket is crucial as tilting the pile can affect the stability and performance of the structure [13]. However, the bucket's inclination can occur due to the unevenness of the subsea ground. In general, the repeated intrusion-pulling method is used to secure verticality in the current suction bucket foundation construction process. However, this relies heavily on the experience and skill of field workers, and it also takes a relatively long time. Various studies have proposed techniques to secure the verticality of the suction bucket such as adjusting the suction pressure [15, 16], modulating the length of the lifting cable [13, 17], and utilizing a tilt sensor-based monitoring system for real-time observations during construction [19]. However, these techniques still depend substantially on field workers' expertise and are time-intensive.

In this paper, we propose an automatic lifting cable control system based on a real-time monitoring system to secure the verticality of large circular steel pipes. The proposed system incorporates a real-time monitoring component that provides real-time data on the pipe's alignment and an actuator control module that automatically adjusts the lifting cable's length to maintain perpendicularity. The real-time monitoring system is implemented with a tilt sensor, a Linux-based embedded controller, and a Windows-based monitoring program. Specifically, the monitoring system consists of a sensor unit that collects raw data, a controller unit that processes the data and wirelessly transmits the information, and a display unit that shows the vertical information of the circular steel pipe. The actuator control module comprises a lifting cable, a hydraulic actuator, a serial communication part, and a controller with an algorithm for adjusting the length of the actuator. We demonstrated that the proposed system could show vertical information in real-time and secure a verticality of 0.2° in model experiments.

The remainder of this paper is organized as follows. Section 2 presents the detailed design and implementation of our proposed system. The experimental results are presented in Section 3. Finally, Section 4 concludes this paper.

II. AUTOMATIC LIFTING CABLE CONTROL SYSTEM

A. Verticality 3D Monitoring System.

The verticality monitoring system includes a sensor unit, a controller unit, and a display unit as shown in Fig. 1. The sensor unit should consist of a sensor that can measure a slope of at least two axes. For data reliability, it can present a different level of information by fusing the data collected by same-type or heterogeneous sensors. Therefore, the sensor unit is designed to be equipped with various types of sensors as required, such as UART inputs like RS-232/422/485, and analog/digital inputs. The controller unit receives data from various sensors, preprocess the collected data, and transmits the verticality information of the circular steel pipe using

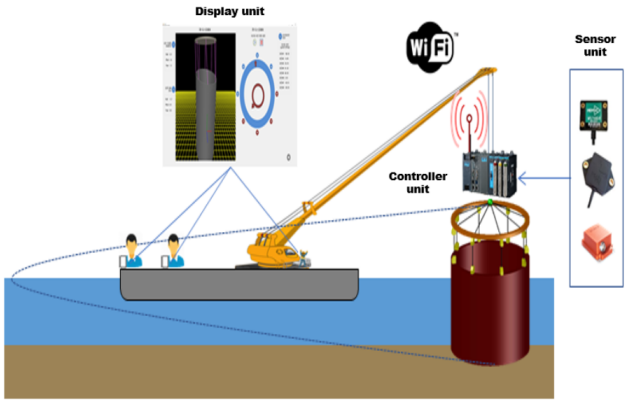


Fig. 1. Verticality monitoring system overview

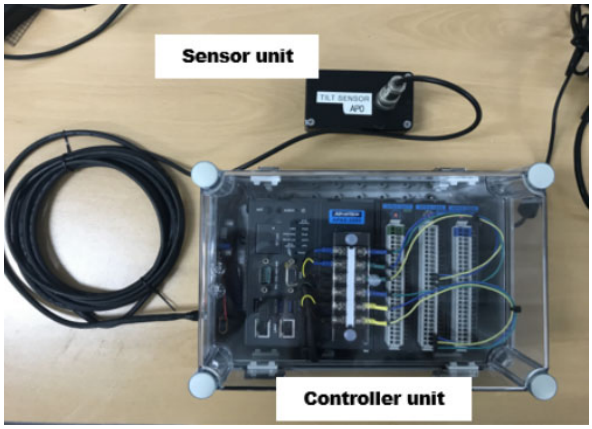


Fig. 2. Sensor and Controller parts for verticality 3D monitoring system

wired/wireless communication. The display unit receives verticality data of the circular steel pipe from the controller unit and presents it in an intuitive 3D display. Transmitting and receiving data between the controller and display units can use wired/wireless communication. But, the network condition is generally inadequate in a marine construction environment. Accordingly, the data is transmitted through the wireless internal network.

- The sensor part is implemented with the MTLT105D-R developed ACEINNA, which can measure two-axis tilt. The MTLT105D-R sensor measures the two axes of pitch and roll at a frequency of 50Hz and transmits it via RS-232 communication. The sensor is separately watertight to prevent damage due to exposure to seawater in a marine environment.

- The controller part is implemented with the APAX-5580 from ADVANTECH, which can collect multiple sensor inputs. The APAX-5580 has various ports for collecting UART/analog/digital input signals, and additional ports can easily be added utilizing expansion slots. The data collected by the sensors is calibrated in the controller unit using a Kalmann filter and then sent to the display unit. Because The TCP protocol has a retransmission function when sending and receiving data, we adopt the TCP protocol, which easily ensures data availability. The implemented sensor and controller parts are shown in Fig. 2.

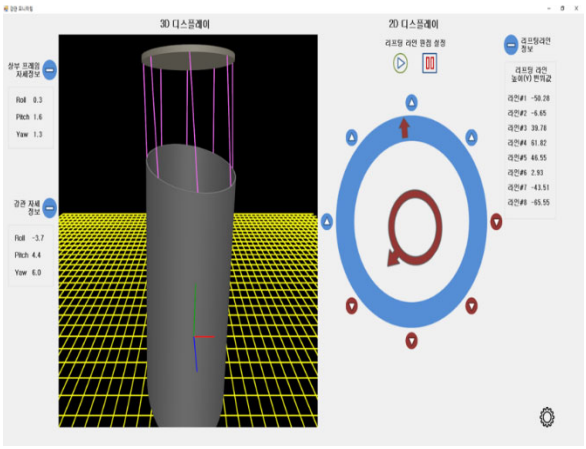


Fig. 3. Main window of the display unit

Fig. 4. Setting window of the display unit

- The monitoring program displayed on the display part consists of the upper frame and lower steel pipe verticality information, 3D display screen, 2D display screen, lifting line information, and setting window button as shown in Fig. 3. The 3D display screen shows real-time 3D information about the state of the upper frame and the verticality of the lower steel pipe sent from the controller unit. The display can be zoomed in/out and rotated. The 2D display screen is a two-dimensional flat screen showing the circular steel pipe from above. The verticality of the circular steel pipe is displayed at the center. If the verticality exceeds the threshold, the user can intuitively understand the direction and degree of tilt of the steel pipe through the changing color of the indicators. The lifting line information is designed to display the height of each lifting line relative to the initial reference value. This facilitates accurate measurements and analysis by enabling the numerical calculation of the displacement for the entire lifting line. The setting window can be accessed by clicking the button at the bottom right of the main screen. In the setting window as shown in Fig. 4, the user can configure the shape of the upper frame, the number of lifting lines, the tilt threshold of the steel pipe, and the connection to the controller.

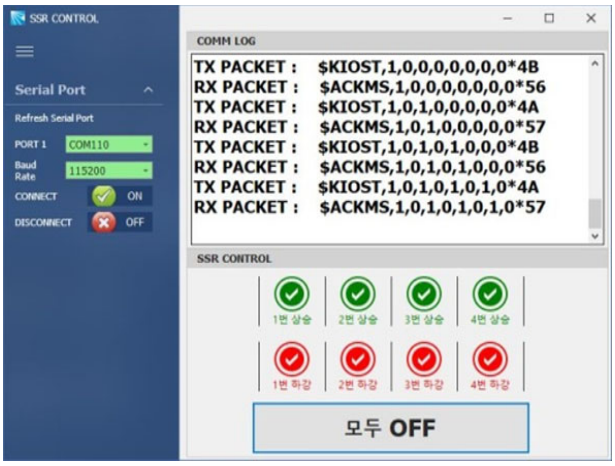


Fig. 5. Actuator control module GUI

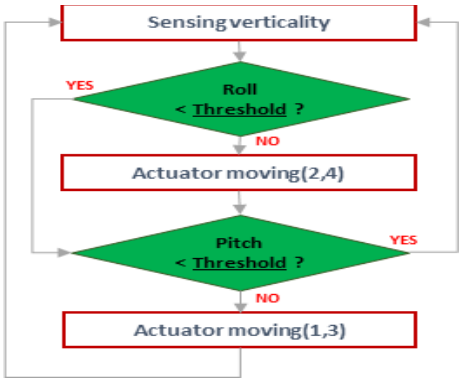


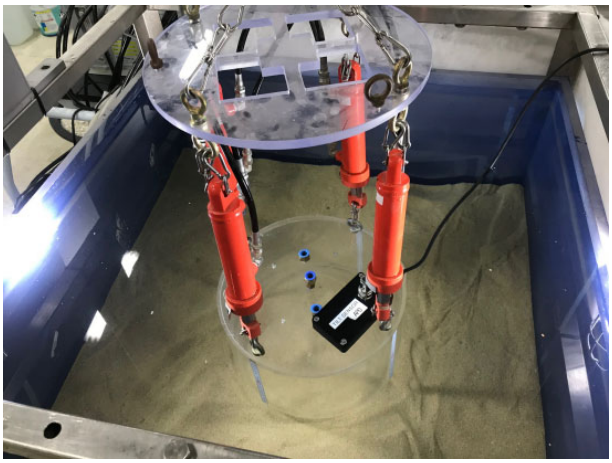
Fig. 6. Actuator control algorithm

B. Actuator Control Module

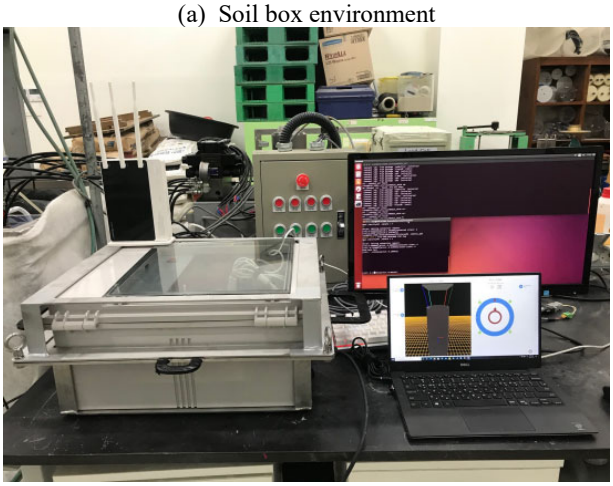
The actuator control module to adjust the length of the lifting cable consists of a lifting cable, a hydraulic actuator, a relay circuit to flip the actuator switch, a controller to control the relay circuit, and a serial communication circuit to receive data from the vertical monitoring system. The actuator control module can be controlled manually or automatically. Each actuator can be manually controlled through the up/down buttons in the GUI program as shown in the Fig. 5. In addition, the actuator can be automatically controlled based on the data collected from the verticality monitoring system to secure verticality by connecting the verticality monitoring system and the actuator control module with the serial interface. The program for automatically controlling the actuator was implemented with a simple algorithm as shown in Fig. 6. The program automatically adjusts the length of the actuator until it is within the tilt threshold based on the verticality information collected by the vertical monitoring system.

III. Experimental Result

The suction bucket foundation method has no verticality standard of its own. But, it must be constructed within the allowable verticality provided in the specification or letter of construction design [15]. In order to effectively apply the proposed system to real suction bucket foundation construction, the time delay of verticality monitoring and the precision of the verticality are crucial.



(a) Soil box environment



(b) Automatic lifting cable control system

Fig. 6. Model experiment environment to demonstrate automatic lifting cable control system

A. Time delay of verticality monitoring

The time delay for data display is given by Equation (1), the time delay for T_{cal} is given by Equation (2), the time delay for T_{trans} is given by Equation (3), and the time delay for T_{disp} is given by Equation (4).

$$\mathrm{T} = T_{cal} + T_{trans} + T_{disp} \quad (1)$$

Where, T_{cal} denotes data calculate, T_{trans} denotes data transfer T_{disp} denotes data display

$$T_{cal} = T_{sens} + T_{hw} + T_{c-app} \quad (2)$$

Where, T_{sens} denotes sensor, T_{hw} denotes h/w line T_{c-app} denotes controller application

$$T_{trans} = T_{c-os} + T_m + T_{s-os} \quad (3)$$

Where, T_{c-os} denotes controller OS protocol stack, T_m denotes h/w medium T_{s-os} denotes server OS protocol stack

$$T_{disp} = T_{g-app} + T_{mon} \quad (4)$$

Where, T_{g-app} denotes GUI application, T_{mon} denotes monitor

T_{sens} is the time delay at which the sensor measures the verticality information. The sensor applied to our system operates at 50Hz, so it can read data every 20ms. T_{hw} means the time delay for the sensed data to transmit the controller via the hardware line, which takes less than 1 psec. T_{c-os} and T_{s-os} represent the time delay caused by transmitting and receiving data in the operating system. T_{c-app} and T_{g-app} indicate the time delay resulting from data processing by the controller and PC. The CPUs of the APAX 5580 operates at 1.7 GHz, and the PC can process more than 2000 instructions in about 1 μsec because the clock speed of the current CPUs used in PCs exceeds 2.0 GHz. T_m indicates the time delay for the controller's data to be transmitted to the 3D GUI program. The transmission delay is about 11.8ms which measure by *ping* command in a wi-fi environment. T_{mon} is the time delay for the monitor to display the data, which takes about 16ms because most monitors currently operate more than 60Hz. Accordingly, the proposed verticality monitoring system takes about 48.8ms from data acquisition to the verticality information display and can operate in real-time.

B. Model Experiment

Fig. 7 is an indoor model experiment of suction bucket penetration to demonstrate an automatic lifting cable control system. The model suction bucket measured a diameter of 30cm and a length of 20cm. The soil box measured 80cm in width, length, and height. Standard sand was placed inside the soil box up to a height of 40cm, and then the soil box was filled with water to a depth of 30cm. Actuators were installed at four points on the top plate, and a two-axis tilt sensor was installed to check the tilt. To enable effective penetration, we first forced the model suction bucket into the ground for about 10 cm. Subsequently, we conduct suction penetration with the automatic lifting cable control system automatic lifting cable control system. We confirmed that the verticality of the bucket was secured by operating the piston of the actuator according to the tilt state. The proposed system calibrated the verticality from 0.6° before suction penetration to 0.2° after penetration.

IV. Conclusion

In this paper, we propose an automatic lifting cable control system based on a real-time verticality monitoring system to secure the verticality of large circular steel pipes. The proposed system consists of a verticality monitoring system and an actuator control module. The verticality monitoring system minimizes the verticality error by setting the origin of the verticality when the sensor is horizontal. Furthermore, it can be utilized in various fields that require precise verticality such as bridge construction in addition to controlling the verticality of steel pipes as it offers various adjustable options according to the construction type. The actuator control module can secure the verticality quickly compared to existing techniques because it automatically adjusts the length of the lifting cable according to the tilt state. The proposed system also operates in real-time, and it has been demonstrated to secure precise verticality of less than 0.2° with model experiments. Therefore, it is effectively applicable to actual construction sites. In the future, the proposed system is expected to demonstrate its reliability through the actual suction bucket foundation construction in a real-sea environment.

Acknowledgment

This work was supported by research grants from Korea Institute of Ocean Science and Technology(PEA0132).

References

[1] Myunghak Oh, Osoon Kwon, Keun-Soo Kim and Insung Jang, "Economic Feasibility of Bucket Foundation for Offshore Wind Farm", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 13, No 4 pp. 1908-1914, 2012.

[2] Hogervorst, J.R, "Field trials with large diameter suction piles", Proc. Offshore Technology Conference, Houston, Texas, USA, May, 1980.

[3] Senpere, Denis, and Gerard A. Auvergne. "Suction Anchor Piles - A Proven Alternative to Driving or Drilling." Proc. Offshore Technology Conference, Houston, Texas, USA, May 1982.

[4] Heejung Youn, Insung Jang, Myounghak Oh, Osoon Kwon, and Sungjun Jung, "Trend in suction bucket foundation for offshore wind turbine", KGS Fall National Conference, KGC, Gyeonggi, Korea, pp.494-403, September 2010.

[5] Heejung Youn, Insung Jang, Myounghak Oh, Osoon Kwon, and Sungjun Jung, "Trend in suction bucket foundation for offshore wind turbine", KGS Fall National Conference, KGC, Gyeonggi, Korea, pp.494-403, September 2010.

[6] Ryu, M. S., Lee J. S., Kwag, D. J., and Seo, Y. H., "Verification of tripod suction pile applicability through dynamic characteristic analysis of offshore wind turbine at each installation stage" *Journal of wind energy*, Vol. 10, No. 3, pp. 12-21, 2019.

[7] Kim, B., Kim, Y., Jin, B., Bae, K., and Youn, H., "Numerical analysis on tilting control of suction pile for offshore wind power" *Journal of the Korean Geo-Environmental Society*, Vol. 17, No. 9, pp. 5-12, 2016.

[8] Houlsby, G. T. and Byrne, B. W. (2000). "Suction caisson foundations for offshore wind turbines and anemometer masts" *Wind engineering*, Vol. 24, No. 4, pp. 249-255.

[9] Houlsby, G. T., Ibsen, L. B. and Byrne, B.W., "Suction caissons for wind turbines", *Proc., Int. Symp. on Frontiers in offshore geotechnics*, Taylor&Francis group, London, pp. 75-94, 2005.

[10] LeBlanc, C., Houlsby, G. T. and Byrne, B.W., "Response of stiff piles in sand to long term cyclic loading", *Geotechnique*, Vol. 60, No. 2, pp. 79-90, 2010.

[11] Kim, Y. S. and Jang, Y. S., "Analysis of load capacity and deformation behavior of suction pile installed in sand", J*ournal of the Korean Geotechnical Society*, Vol. 27, No. 11, pp. 27-37, 2011.

[12] Lee, J. H., Kim, D. W., Chung, M. K., Kwak, K. S. and Jung, Y. H., "Numerical analysis of the suction pile behavior with different lateral loading locations", Journal of the Korean Geotechnical Society, Vol. 27, No. 4, pp. 67-76, 2011.

[13] Park, Haeyong, Oh, Myounghak, Kim, Sungwon, Han, Taekhee, and Baek, Seungjae, "Model Tests on Verticality Correction using Individual Length Control Method for Lifting Cable during Suction Pile Penetration", *J. Korean Soc, Hazard Mitig,* Vol. 19, No.5, pp.217-223, 2019.

[14] Korea Ocean Research & Development Institute (KORDI), "Suction foundation structure technique manual", KORDI, Korea, 2010.

[15] Kwag, D., Oh, M., and Kwon, O, "Field installation tests of monopod suction pile and tripod suction buckets", Proceedings of the ASME 2013 32nd International Conference on Ocean, Offshore and Arctic Engineering OMAE2013, Nantes, France, pp. 1-8, June 2013.

[16] Kim, Y.S., Bae, K.T., Lee, J.P., Joung, J.W., and Choo, Y.W, "Model tests for tiling control of suction bucket foundation for offshore wind turbine", *J. Korean Soc. Hazard Mitig*, Vol. 17, No. 2, pp. 207-218, 2017.

[17] Haeyoung Park, Myounghak Oh, and Hyoun Kang, "Effect of lifting Cables tension on verticality correction of suction pile during ground penetration", *Journal of Coastal Disaster Prevention*, Vol. 7, No. 2, pp. 63-72, 2020.

[18] Ministry of Land, Infrastructure, and Transport (MOLIT), "Development of erection method for marine bridge foundations", MOLIT, Korea, 2015.

[19] Koo, Sungmin, Park, Haeyong, Oh, Myounghak, and Baek, Seungjae, " Verticality 3D Monitoring System for the Large Circular Steel Pipe", *Journal of the Korea Academia-Industrial cooperation Society*, Vol. 21, No.11, pp.870-877, 2020.

Sea-Water-Battery for maritime Application

Juhyun Kim
Maritime ICT·Mobility
Research Department
KIOST
Busan, Korea
jhk@kiost.ac.kr

Hyun Kang
Maritime ICT·Mobility
Research Department
KIOST
Busan, Korea
hkang@kiost.ac.kr

Seungjae Baek
Maritime ICT·Mobility
Research Department
KIOST
Busan, Korea
baeksj@kiost.ac.kr

***Abstract*— This research paper explores the utilization of marine specialized batteries in various marine devices to overcome the challenges posed by corrosion and damage caused by seawater. Two types of marine-specific batteries are investigated: seawater batteries and lithium-ion batteries. The specifications of these battery systems are presented, highlighting their capacities, dimensions, and weights. Real-sea experiments are conducted to evaluate the performance and functionality of marine devices equipped with these specialized batteries. The findings demonstrate the successful integration of the maritime specialized batteries in different marine applications, including light buoys, leisure vessels, and unmanned surface vehicles. The results reveal extended battery life, improved power output, and enhanced system stability, showcasing the potential of these batteries in addressing the limitations of existing battery technologies in the marine environment.**

Keywords—marine specialized battery, problems with existing battery, maritime application, stable power supply

I. INTRODUCTION

Lithium-based and lead-acid batteries are commonly employed in a wide range of battery-powered marine devices, including light buoys, marine unmanned vehicles, and electric propulsion ships. Unfortunately, the current batteries used in these applications face significant challenges such as corrosion and damage caused by seawater. As a result, they suffer from limitations such as shortened usage time, vulnerability to flooding, and the potential for environmental pollution due to hazardous material leakage [1, 2, 4]. In an effort to alleviate this issue, this paper introduces a maritime specialized battery. We demonstrate that the integration of this maritime specialized battery in various marine devices leads to improved performance and functionality [1, 4].

II. MARINE SPECIALIZED BATTERIES

Within the demanding marine environments where marine devices are employed, extensive research is currently underway to develop maritime specialized batteries capable of withstanding the effects of seawater exposure. This research aims to address the challenges posed by corrosive conditions. Two distinct types of marine-specific batteries have emerged: seawater battery-based systems and lithium-ion battery-based systems [5].

A. Seawater battery

Seawater batteries (SWBs) leverage the vast availability of seawater as a virtually limitless resource for the storage and generation of electrical energy. These batteries store electrical energy by harnessing the chemical reactions involving water and the sodium ions dissolved within the seawater. By utilizing the unique properties of seawater, SWBs offer a sustainable and efficient method of storing and producing electrical energy. [6].

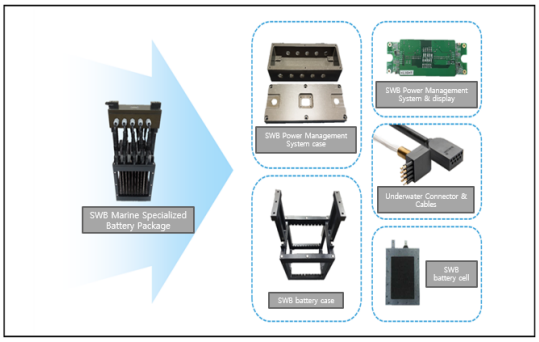


Fig. 1. Sewater battery as a marine-specific power source

Cell	8Wh (2.8V, 3Ah)	160*181*30(mm)	0.95kg
Unit Module	144Wh (12V, 12Ah)	444*167*205(mm)	18kg

Fig. 2. Seawater Battery specifications

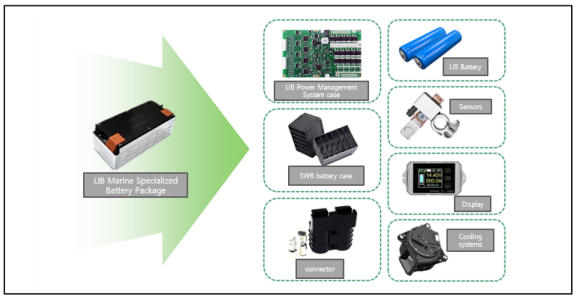


Fig. 3. Marine-specific power supplies based on lithium-ion batteries

Cell	11Wh (3.7V, 3Ah)	φ18*70(mm)	50g
Unit Module	1.2KWh (12V, 40Ah)	360*120*230(mm)	13kg

Fig. 4. Li-ion Battery Marine Power Supply Specifications

- The specifications of the seawater battery system designed for maritime applications are illustrated in Figure 2. The individual cell of the seawater battery possesses a capacity of 8Wh (2.8V, 3Ah) with dimensions measuring 160×181×230mm and weighing 0.95kg. Furthermore, the unit module of the seawater battery offers a capacity of 144Wh (12V, 12Ah), featuring dimensions of 444×167×205mm and weighing 18kg.

B. *Lithium-ion battery*

Lithium-ion batteries are extensively employed due to their ability to be repeatedly used between 500 to 2000 cycles and their compact size and weight relative to batteries with similar capacities [3]. Notably, these batteries do not contain environmentally regulated substances such as cadmium, lead, or mercury. Moreover, they are free from the memory effect, which diminishes their rechargeable capacity over time. Additionally, lithium-ion batteries possess the advantage of delivering higher power output compared to conventional battery technologies [5].

- The specifications of the marine-specific lithium-ion battery-based system are presented in Figure 4. The Li-ion battery cell within this system offers a capacity of 11Wh (3.7V, 3Ah) with dimensions measuring 18×70mm and a weight of 50g. Furthermore, the unit module of this battery system boasts a capacity of 1.2kWh (12V, 40Ah) and features dimensions of 360×120×230mm, weighing 13kg.

III. Marine Application Testing Equipped with Marine Specialized Batteries

In recent times, there has been a notable surge in research focused on diverse marine devices. This paper particularly highlights the introduction of marine-specific batteries in various marine applications, such as light buoys, vessels, and unmanned surface vehicles.

A. *Light buoys*

A light buoy serves as a navigational marker to guide ships along specific routes or waterways. However, existing light buoys often face the issue of frequent lights going out due to limited battery life. To overcome this challenge, we implemented marine specialized batteries [1, 4]. Four seawater batteries were installed, supplying 3.9V and 1A of power. Additionally, a remote control device was added to manage the lights from a distance.

Through real-sea experiments, we found that the light buoy equipped with marine specialized batteries operated autonomously for four days without requiring any external power input. These results demonstrate the effectiveness of using marine specialized batteries in solving the battery life issue and improving the performance of light buoys in marine environments.

B. *Marine leisure vessels*

Existing leisure ships are commonly powered by internal combustion engines or batteries to drive their motors, leading to issues such as marine environmental pollution, noise, lengthy charging periods, durability problems, and fire hazards in the event of accidents. However, by implementing marine specialized batteries with hydrogen fuel cells as the primary energy source in a hybrid configuration, we have successfully mitigated the problem of marine pollution while resolving issues related to corrosion and failure caused by seawater [1, 3, 4].

During a real-sea test, a hybrid battery-powered ship demonstrated impressive performance. With five individuals on board, the hybrid system supplied a consistent 53.7V and 56.3A of power, achieving a maximum power output of 2.5kW, a top speed of 4 knots, and a cruising time of 4 hours. These results highlight the efficacy of the hybrid battery system in addressing the challenges associated with traditional power sources, paving the way for cleaner, more efficient, and reliable leisure ships in marine environments.

Fig. 5. Four seawater-based marine specialized batteries for light buoy

Fig. 6. Light buoy

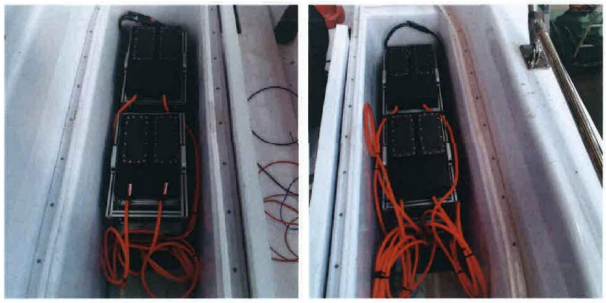


Fig. 7. Seawater-based marine specialized batteries for marine leisure vessel

Fig. 8. Marine leisure vessels

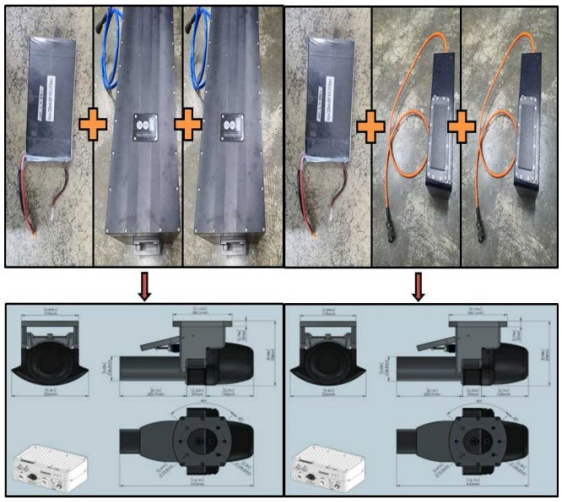


Fig. 9. Seawater-based marine batteries (left) Lithium-ion-based marine batteries (right) for unmanned water boat

Fig. 10. Unmanned water boat

C. Unmanned surface vehicle

Unmanned surface vehicles (USVs) are marine unmanned vehicles designed for various maritime applications, including navigation, surveillance, and environmental data collection. The USV proposed in this paper is equipped with a hybrid system consisting of both lithium and seawater batteries, providing a voltage output of 21V to the entire system. In order to evaluate the performance of the unmanned watercraft, real-sea experiments were conducted, comparing parameters such as hull transmission, motor functionality, current flow method, battery operating time, attitude values (roll, pitch, and yaw), and system stability tests.

The experimental results revealed that the targeted attitude values (roll, pitch, and yaw) were accurately measured even without the movement of the hull. Furthermore, the operation time of the communication equipment was quantitatively measured to be approximately 2 hours with a stable power supply, yielding a measured voltage of 20.26V and a continuous discharge current of 1.5A. These findings, demonstrate the successful integration of the hybrid battery system in the USV, showcasing its capability to meet the power demands and maintain system stability during real-sea operations.

IV. Conclusion

This paper introduces and evaluates the application of maritime specialized batteries in marine devices. The integration of seawater batteries in light buoys improved their performance, enabling extended operation without power input. The implementation of hybrid battery systems in leisure vessels mitigated marine pollution and addressed durability and safety concerns. The hybrid system in unmanned surface vehicles showcased accurate attitude measurement and stable power supply. Overall, the use of maritime specialized batteries offers enhanced performance, extended battery life, improved power output, and system stability, paving the way for cleaner and more efficient marine operations. Future work will focus on further optimization and exploration of these battery systems to enhance their capabilities and expand their applications in the maritime industry.

Acknowledgment

This research was supported in part of the project titled `Development and Demonstration of Seawater Secondary Battery Module over 1kWh for Marine Equipment', funded by Ministry of Trade, Industry and Energy

References

[1] S. Koo, S. Jeong, J. Seo, Y. Song, O. Kwon, Y. C. Kim, S. Baek, "Sea-water Battery for Maritime Applications", Global Oceans 2020: Singapore – U.S. Gulf Coast, pp. 1-4, Oct. 2020.

[2] Frölicher, T.L., Fischer, E.M., and Gruber, N., "Marine heatwaves under global warming," Nature 560, 360–364, Aug. 2018

[3]

[4] J.-K. Kim, E. Lee, H. Kim, C. Johnson, J. Cho, Y. Kim, "Rechargeable seawater battery and its electrochemical mechanism", ChemElectroChem. Wiley, Vol. 2(3). pp. 328-332, Nov. 2014.

[5] Jinmao Chen, Wanli Xu, Xudong Wang, Shasha Yang, Chunhua Xiong, "Progress and Applications of Seawater-Activated Batteries" Sustainability Vol 15, no. 1635 p. 1635

[6] Sungmin Koo, Seungjae Baek, Hyoun Kang, Ah Hyun Park, "Hybrid Battery System based Fish Biologger" Vol. 22, no.11, pp. 870-876(7 pages), 2021

[7] Kim, Youngjin et al., "High-Capacity Anode Materials for Sodium-Ion Batteries", Chemistry A European Journal 20(38), 11980-11992. (2014)

[8] Hyoun Kang, Mee Kyung Kim, Hyun Soo Yoon, Sungmin Koo, "Research on the development of Unmanned Surface Vehicles with Power Supply Specialized for Maritime Environment" Journal of the Korea Academia-Industrial cooperation Society Vol. 23, No. 10 pp. 712-718, 2022.

Advancing Real-Time Monitoring and Prediction: An Indoor Air Pollutant System Leveraging Internet-of-Things and Deep Learning Models

Jamal Yousaf
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
Busan, Republic of Korea
j.yousaf@donga.ac.kr

Anna Camille Sanchez
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
Busan, Republic of Korea
acsanchez.ce@gmail.com

Seung-Beom Lee
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
Busan, Republic of Korea
dda1109@naver.com

Hyeon-Ju Jang
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
Busan, Republic of Korea
pinkhj1218@naver.com

Kyung-Hee Park
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
ABLEECO
Busan, Republic of Korea
blueshout@naver.com

Dong-Keon Kim
Department of ICT Integrated Ocean Smart Cities Engineering
Dong-A University
Busan, Republic of Korea
dkkzone@dau.ac.kr

***Abstract*— This paper presents the development of an intelligent monitoring system aimed at tracking and predicting multiple indoor air pollutants in real-time. Leveraging advanced sensors, the internet-of-things (IoT), and a cloud-based platform, the system enables continuous monitoring of controllable air quality parameters, including temperature, humidity, carbon dioxide (CO_2), carbon monoxide (CO), and particulate matter ($PM_{2.5}$, PM_{10}) concentrations. The design principles, implementation strategies, and preliminary findings of our ongoing research are outlined. Despite being in the early stages, initial experiments showcase the system's potential to deliver timely feedback on indoor air quality, empowering proactive measures for ensuring healthier and more comfortable living and working environments. This work not only contributes to the field of indoor air quality monitoring but also paves the way for future advancements in intelligent, data-driven approaches for air pollutant control systems.**

Keywords—indoor air quality, real-time, monitoring, Internet-of-Things

I. Introduction

Indoor environments, such as homes, offices, schools, and other confined spaces, have become the primary settings where people spend the majority of their time, particularly in light of the Covid-19 pandemic. Regrettably, indoor air quality is frequently compromised by various factors, including inadequate ventilation, the presence of pollutants, and the accumulation of allergens. Subpar indoor air quality can give rise to an array of health issues, ranging from respiratory problems to allergies and even long-term chronic conditions. Consequently, ensuring optimal indoor air quality is of utmost importance, as it directly impacts human health, comfort, and productivity. The objective of this study is to develop an Internet-of-Things (IoT)-based indoor air monitoring system that enhances indoor environmental quality through continuous monitoring, prediction, and control, leveraging cloud computing and deep learning techniques. The use of various sensors based on the Internet of Things (IoT) and machine learning has become increasingly prevalent in monitoring and predicting indoor air quality [1]. Real-time measurements of indoor air pollutants have been taken using IoT applications, with recommendations made for improving ventilation and reducing indoor activities [2].

II. Related Literature and Design Principles

Previous studies on indoor air quality assessment have often suffered from limitations such as a random selection of pollutants, short-term data collection, and data loss due to Wi-Fi communication issues. Furthermore, these studies have often lacked data accessibility and an indoor air quality (IAQ) index, failing to capture the dynamic and real-time nature of indoor air quality variations. Consequently, occupants lack timely information to proactively address air quality concerns and improve the quality of the air they breathe.

To address these shortcomings, this study focuses specifically on controllable pollutants, namely carbon dioxide (CO_2), carbon monoxide (CO), PM_{10}, $PM_{2.5}$, as well as temperature and humidity. Unlike previous research, this study aims to collect data over a more extended period, spanning at least one full year, to obtain a comprehensive understanding of indoor air quality and develop an accurate IAQ index. Additionally, the proposed monitoring device incorporates redundant data saving mechanisms to prevent loss of collected information, ensuring data integrity and availability for users.

Moreover, this ongoing work aims to achieve automated responses for improved indoor air quality management. The integration of the intelligent indoor air monitoring device with existing smart home systems will enable advanced functionalities in the future. For instance, in response to poor air quality, the device will have the potential to trigger the ventilation system to increase airflow or activate air purifiers to effectively mitigate pollutants. By continuously monitoring indoor air quality parameters, occupants will have the ability to optimize ventilation systems, adjust heating and cooling settings, and ultimately reduce energy consumption, while ensuring a healthy and productive indoor environment. Although the development of these automated responses is part of future work, the proposed system sets the foundation for achieving these goals, with the potential to significantly enhance indoor air quality management in the future.

III. Device Design and Functionality

In this section, key design principles and functionality of the proposed indoor air monitoring device are discussed.

A. Hardware

The proposed device is equipped with a diverse array of sensors capable of measuring various crucial indoor air quality parameters, encompassing temperature, humidity, carbon monoxide (CO) and carbon dioxide (CO_2) levels, as well as $PM_{2.5}$ and PM_{10} concentrations. Through wireless communication, the device seamlessly transmits the collected data to a cloud-based platform, where cutting-edge data analytics techniques are employed to process and analyze the information in real time. This enables continuous monitoring and real-time analysis of indoor air quality. To ensure data integrity, the device incorporates local storage capabilities, safeguarding against potential network disruptions. The following components are listed in Table 1.

TABLE I. COMPONENTS OF THE DEVICE

Component	Function	Parameters measured
Raspberry Pi 4B (8GB)	Microcontroller	
MH-Z19B	Carbon dioxide sensor	CO_2
CUBIC PM2008M	Fine dust sensor	PM2.5 PM10
DFROBOT SEN0377	Carbon monoxide sensor	CO
PIM458 Enviro+	Air quality sensor	Temperature Humidity

B. Implementation Strategy: Code Development

The software architecture of the device is implemented using Python, a versatile programming language. The coding process entails the development of essential code modules and functions to fulfill the software requirements. For most of the sensors used, readily available Python libraries and modules are leveraged, simplifying the integration process. Additionally, specific functions are developed to calculate the Indoor Air Quality (IAQ) index and classify the IAQ status based on the Air Korea Guidelines. These functions ensure accurate and reliable acquisition of sensor data, accounting for potential error sources and handling data anomalies or outliers. Upon data collection, the device securely transmits the gathered information to both the cloud and an external SD card for further analysis and storage. The initial configuration of the device is visually depicted in Figure 1, providing an overview of its open design.

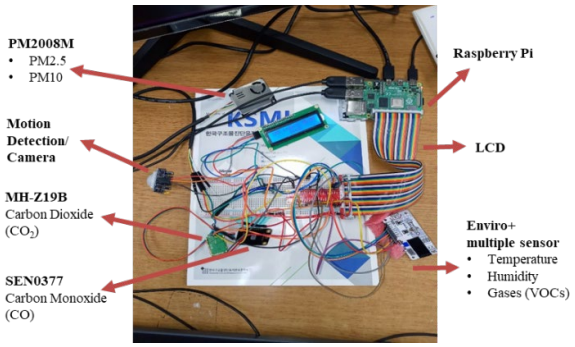


Fig. 1. Device configuration

C. Integration and Testing

While the device is still in the early stages of development, initial demonstrate its capability for real-time monitoring. The device successfully captures variations in temperature, humidity, CO and CO_2 levels, $PM_{2.5}$ and PM_{10} concentrations, providing real-time insights into the indoor environment's air quality. A screenshot of the cloud interface showing the received data from the Raspberry Pi is shown in Figure 2.

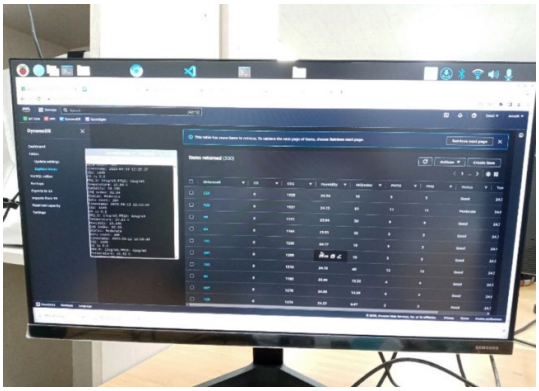


Fig. 2. Data transfer from Raspberry Pi Terminal to AWS cloud platform

IV. DISCUSSION AND FUTURE WORK

This work-in-progress paper focuses on developing a device for the monitoring stage of indoor air quality assessment. The device successfully monitors air quality parameters, determines the air quality index, and transmits data to the cloud. The next steps include calibration and validation of the sensors, data analysis using deep learning for prediction, and integration into field experiments. The proposed intelligent indoor air monitoring device integrates advanced sensors, IoT capabilities, and deep learning techniques to provide real-time information on various air quality parameters. Its goal is to revolutionize indoor air quality monitoring, enhance control and management, and contribute to healthier living and working environments.

ACKNOWLEDGMENT

This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Korea Ministry of Education (No.2016R1A6A1A03012812).

REFERENCES

[1] Mumtaz, R., Zaidi, S. M. H., Shakir, M. Z., Shafi, U., Malik, M. M., Haque, A., ... & Zaidi, S. A. R. (2021). Internet of things (IoT) based indoor air quality sensing and predictive analytic—A COVID-19 perspective. Electronics, 10(2), 184.

[2] Esfahani, S., Rollins, P., Specht, J. P., Cole, M., & Gardner, J. W. (2020, October). Smart city battery operated IoT based indoor air quality monitoring system. In 2020 IEEE SENSORS (pp. 1-4). IEEE.

Building protection during a typhoon: an early warning window break detection system for occupants' safety in ocean front coastal shoreline cities

Christian Matthew Mabborang
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
matthewmabborang@rocketmail.com

Regidestyoko Wasistha Harseno
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
regidestyoko@gmail.com

Nur Indah Mukharromah
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
nurindahmukharromah@gmail.com

Bin na Choi
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
bn2374@naver.com

Won gi Kwon
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
dnjsrl9565@naver.com

Dong keon Kim
Department of ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
dkkzone@dau.ac.kr

***Abstract*— This paper presents a novel early warning system designed to detect window breakage and ensure the safety of occupants in coastal cities during typhoons. The system comprises two main processes: the warning process and the unalert process. In the warning process, an audible alarm and LED indicator are activated when the stress and natural frequencies detected surpass 70% of the window glass fracture stress and natural frequencies, respectively. In contrast to existing window break detection systems, the proposed system prioritizes early warnings, thereby mitigating the potential risks of personal injury and property damage resulting from shattered windows. Experimental validation of the proposed system is conducted using two different window sizes.**

Keywords*— *early warning, window break, detection system, safety, typhoon, coastal cities.

I. Introduction

In recent years, the frequency of typhoons, powerful tropical cyclones characterized by strong winds, has increased due to climate change [1]. These typhoons predominantly impact the northwest Pacific Ocean region, including East Asian nations like South Korea [2]. The strong winds result in significant damage, not only to load-bearing structural elements but also to envelope structures such as windows. The destruction of windows not only leads to economic losses but also disrupts and endangers people's lives [3]. Research [4] indicates a growing number of incidents in South Korea where exterior materials, including window, have fallen off buildings due to strong winds. Aging apartment, lack of structural monitoring, and insufficient sealant are some of the contributing factors to the detachment window.

Over the five-year period, South Korea experienced 24 typhoons, averaging 4.8 typhoons per year [5], exceeding the normal occurrence rate of 3.1 per year. The wind speed intensity resulting from typhoons has progressively increased, from 144.67 km/h in 2019 to 200 km/h in 2022 [6]. Noteworthy typhoon-induced incidents include a Seogwipo resident being hospitalized in 2014 due to Typhoon Nakri shattering a window [7], the windows of a Haeundae building in Busan being broken by Typhoon Kong-rey in 2018 [8], and a large glass window in Busan being shattered by Typhoon Tapah in 2019, resulting in it falling on the sidewalk below [9]. In 2020, broken store windows were reported in Seogwipo and Jeju City during Typhoon Bavi [10], and Typhoon Maysak caused at least two deaths in Busan due to broken windows [11]. The rising frequency of high-intensity typhoons in the Pacific Ocean [12] and the escalating cost of typhoon-related damages in South Korea, which increased from 0 KRW in 2017 to 222.54 billion KRW in 2020 [13], highlight the need for risk assessment of structures vulnerable to strong winds.

A window break detection system has been developed that is capable of providing an alarm signal upon detecting a broken window [14–15]. However, to further enhance the safety of occupants, there is a need for an early warning detector system. In this study, a novel early warning window break detection system is proposed to mitigate the risk of injuries or property damage caused by broken windows during typhoons.

II. Early Warning Window Break Detection System

Every material has a natural frequency, which is the frequency at which it vibrates when excited. Our proposed system utilizes a piezoelectric sensor to detect stress and monitor natural frequencies resulting from wind loads. This cost-effective system maintains the aesthetic integrity of the window by putting the sensor in each window edge where the vibrations are most effectively transmitted, as shown in Fig. 1. To anticipate window breakage, stress and natural frequency thresholds are established. Increasing stress and shifting natural frequencies indicate glass weakening and an elevated risk of breakage.

Table I presents the equipment list and budget. The electronic circuit setup presented in Fig. 2 includes an Arduino as the controller and a piezoelectric, which will generate a voltage when it is physically deformed by a vibration, as the sensor. The Arduino receives and processes data from both sensors using its built-in programming capabilities. This processed data can be utilized to trigger different actions or responses, such as sending notifications through sound and light. Additionally, the proposed system's framework and flowchart are illustrated in Figs. 1 and 2, respectively.

TABLE I. EQUIPMENT LIST AND BUDGET

No.	Equipment	Quantity	Remarks	Total Price
1	Arduino Uno R3	1	Mini-PC	KRW 28,000
2	Ceramic Piezoelectric Buzzer 27mm [FQ-076]	4	Sensor	KRW 720
3	5 Pi High Brightness LED Translucent [5R3HT-10]	3	Alarm system	KRW 210
4	IMTI2D2001AP	1	Alarm system	KRW 290
5	Socket jumper cable 40P (color) (F/F) 10cm [CH254]	40	Connector	KRW 700
6	Socket jumper cable 40P (color) (M/F) 10cm [CH254]	40	Connector	KRW 700
7	BC337 Transistor	5	Connector	KRW 300
8	1/2W 1% Axial Resistor 102F (1KΩ)	20	Connector	KRW 100
9	MCP3008-I/P	1	Connector	KRW 5,700
10	Breadboard 400-pin Half Size Breadboard	4	Connector	KRW 2,800
Grand Total Price				**KRW 39,520**

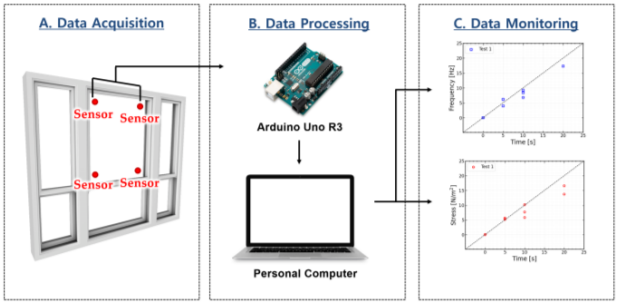


Fig. 1. Proposed system framework

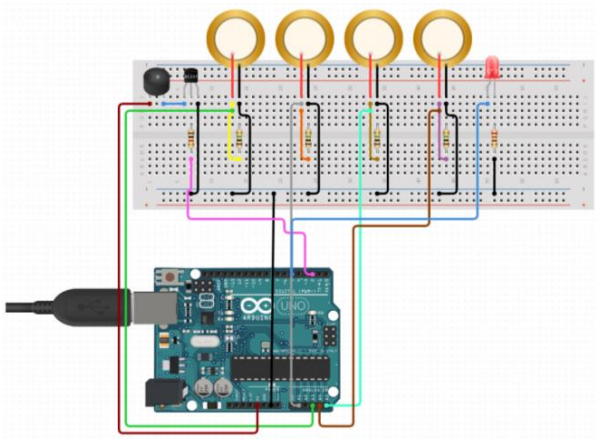


Fig. 2. Electronic circuit setup

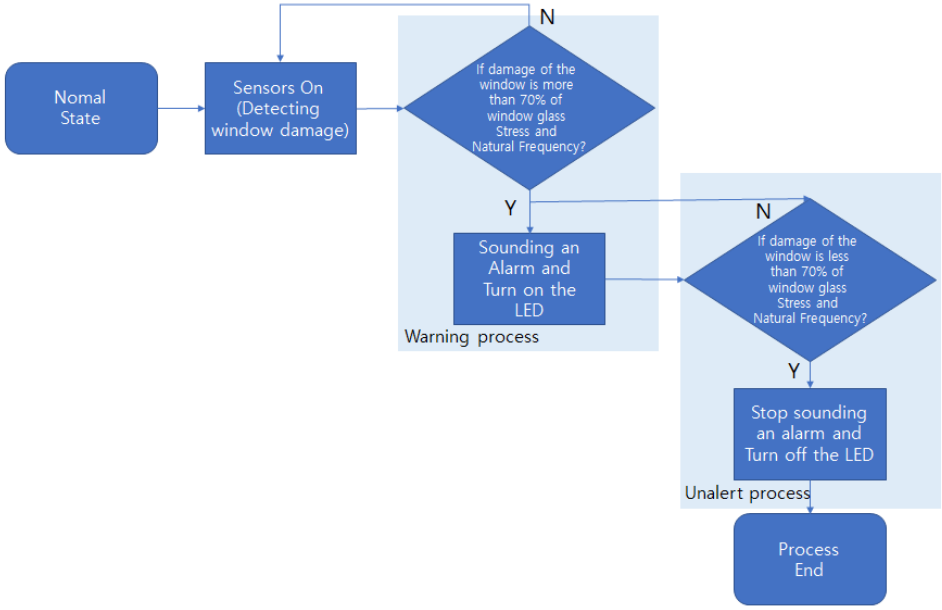


Fig. 3. Proposed system flowchart

III. EXPERIMENTAL PROGRAM

A simplified building model is employed, utilizing window glass panels of two different sizes: 1 m × 1 m and 0.5 m × 0.5 m, both having a thickness of 5 mm. The simplified model takes the form of a foam cube, wherein the window glass is positioned on one of the faces, as depicted in Fig. 4. The properties of the window glass are detailed in Table 2.

The stress threshold utilized for the proposed window design is determined based on a comprehensive full-scale glass window experiment conducted by [16] under the support

of the Korea National Disaster Management Research Institute. This experiment covered various thicknesses, reinforcements, and installation conditions. However, for the purposes of this paper, only a specific type of window with a thickness of 5 mm is considered. The fracture stress obtained from the experiment, specifically for the 5 mm-thick window glass, amounted to 4263 Pa. Additionally, the natural frequency threshold is determined by calculating the natural frequency of the glass as presented in (1) based on [17] in accordance with its characteristic properties. Both the stress and natural frequency thresholds are established at 70% of the fracture load and the calculated natural frequency, respectively.

$$f_{natural} = (1/2\pi)\,(K/M)^{0.5} \quad (1)$$

where $f_{natural}$ is the natural frequenzy (Hz), K is the stiffness of the structure (N/m), and M is the mass of the structure (kg).

The experiment will be conducted in a consistent location for both window glass sizes, utilizing the proposed system. Prior to conducting the experiment, the defined threshold will be allocated to the proposed model. The reading from the sensor will be calibrated to the windspeed which is provided by the open data source from [18].

TABLE II. WINDOW GLASS PROPERTIES

Properties	General window glass
Density (kg/m^3)	2500
Hardness	6 – 7
Young's modulus (N/mm^2)	7.35×10^4
Coefficient of linear expansion (strain/ºC)	9.54
Thermal conductivity (W/mºC)	0.60

Fig. 4. Simplified building model

IV. RESULTS AND DISCUSSIONS

The piezoelectric sensor was used to collect data from two different window sizes: 0.5 m × 0.5 m and 1 m × 1 m, as depicted in Figs. 5 and 6, respectively. The data collection process for the and 0.5 m × 0.5 m cm window size lasted approximately 45 minutes. The collected data exhibited discernible patterns at specific times, indicating variations in the load experienced by the sensor due to changes in wind speed during the data acquisition period. Higher sensor readings corresponded to increased stress from external loads, while lower readings indicated reduced stress. These findings demonstrate the sensor's capability to accurately detect changes induced by external forces.

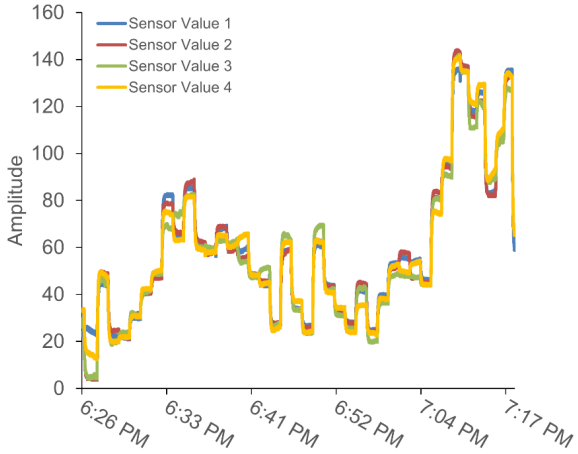


Fig. 5. Data reading from all sensors in case of 0.5 m × 0.5 m window size

In Fig. 6, corresponding to the 1 m × 1 m window size case, the sensor readings exhibited a similar pattern to the previous case. The data collection duration for this case was approximately 15 minutes. The observed data pattern displayed a distinct trend at specific time intervals, reflecting the pressure experienced by the sensors. The purpose of employing different window sizes was to investigate whether variations in data readings would occur. However, it was found that the sensor readings from the different window sizes were nearly identical. These results suggest that, in real-world deployment scenarios, the window size does not significantly impact the sensor's readings.

Moreover, in both cases, the sensors positioned at each corner of the window exhibited similar readings, indicating consistent data across the different sensor locations. The placement of sensors in each corner was primarily intended to detect variations in stress levels that could arise during typhoon events.

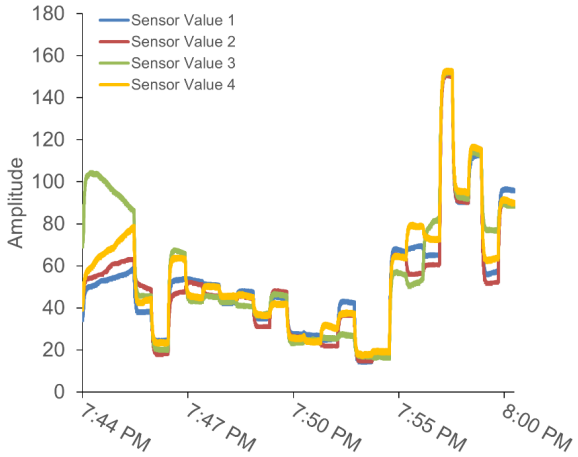


Fig. 6. Data reading from all sensors in case of 1 m × 1 m window size

In future studies, we aim to calibrate our sensor readings using wind speed data obtained from [18]. The primary objective of calibration is to enhance the comprehensiveness

and meaningfulness of the collected data. By establishing a direct correlation between our sensor readings and the stress levels over time, the calibrated data will provide more accurate and valuable insights. This calibration process will enable us to refine our understanding of the relationship between sensor readings and the corresponding wind speeds, ultimately improving the reliability and applicability of our sensor data in assessing stress variations.

V. Conclusions

In conclusion, this paper presents a novel early warning system designed to detect window breakage during typhoons, with a focus on ensuring the safety of occupants in coastal cities. The system incorporates a warning process and an unalert process, where an audible alarm and LED indicator are activated when stress and natural frequencies detected surpass predefined thresholds. By prioritizing early warnings, the proposed system mitigates the potential risks of personal injury and property damage resulting from shattered windows. Experimental validation of the system using different window sizes confirmed its effectiveness in detecting changes in stress levels and natural frequencies induced by external forces. Moreover, the study demonstrated that the location of the sensors within the window had consistent readings, highlighting the system's reliability in detecting stress variations.

For future studies, our research aims to calibrate the sensor readings using wind speed data from [18]. This calibration process will establish a direct correlation between the sensor readings and the stress levels over time, enabling more comprehensive and meaningful data analysis. By refining our understanding of the relationship between sensor readings and wind speeds, we can improve the reliability and applicability of our sensor data in assessing stress variations. The calibrated data will provide more accurate insights and contribute to a more comprehensive risk assessment of structures vulnerable to strong winds. Overall, the proposed early warning system offers a promising approach to enhancing the safety of occupants during typhoons and can serve as a foundation for further research in the field of structural monitoring and risk mitigation.

Acknowledgment

This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Korea Ministry of Education(No.2016R1A6A1A03012812).

References

[1] Ministry of the Interior and Safety (2021). *Disaster & safety management in Korea*. Ministry of the Interior and Safety.

[2] Choi, J. W., Cha, Y., Kim, H. D., & Kang, S. D. (2016). Climatological features of Korea-landfalling tropical cyclones. *Advances in Meteorology, 2016*, 1-15.

[3] Sim, V., & Jung, W. (2019). Wind fragility of glass windows with tape and film reinforcement. *Journal of the Korean Society of Hazard Mitigation, 19*(6), 33-39.

[4] Lee, D. Y., Cho, B. H., Kim, H. J., Kim, G. Y., Jung, D. U., & Shin, S. H. (2022). Cases of falling off exterior materials of domestic buildings due to strong winds and guideline for prevention of dropouts. *Architectural Institute of Korea Conference, 42*(2), 534-535.

[5] Korea Meteorological Administration (2022). *Year-by-year typhoon occurrence and typhoon status*. Korea Meteorological Administration.

[6] World Data Info (2023). *Most violent typhoons in South Korea since 2018*. World Data Info.

[7] Korean Times (2014, August 3). Heavy rain from Typhoon Nakri soaks southern s. Korea. *Korea Times*. https://www.koreatimes.co.kr/www/nation/2023/02/113_162186.html?utm_source=CU.

[8] Korea Herald (2018, October 7). Typhoon Kong-rey leaves 2 dead, 1 missing. *Korea Herald*. https://m.koreaherald.com/amp/view.php?ud=20181006000080.

[9] Korea Times (2019, September 22). Typhoon Tapah approaching Korea's southern region. *Korea Times*. https://www.koreatimes.co.kr/www/nation/2020/10/113_275929.html

[10] Korea Joongang Daily (2020, August 26). Typhoon Bavi uproots trees, will graze Seoul tonight. *Korea Joongang Daily*. https://koreajoongangdaily.joins.com/2020/08/26/national/socialAffairs/Typhoon-Bavi-Korea-Summer-2020/20200826181500391.html

[11] Reuters (2020, September 3). At least two people killed as Typhoon Maysak hits waterlogged Koreas. *Reuters*. https://www.reuters.com/article/us-asia-storm-southkorea-northkorea-idUSKBN25U043.

[12] Typhoon Committee (2015). *Typhoon committee operational manual 2015*. World Meteorological Organization.

[13] Ministry of the Interior and Safety (2022). *Cost of damage caused by typhoons in South Korea from 2011 to 2020 (in billion South Korea won)*. Ministry of the Interior and Safety.

[14] Venkat, K. (2010). Glass-Breakage Detector Uses One Microcontroller. *EDN, 55*(17), 39-45.

[15] Bhosale, M., Marwade, S., Arke, V., & Shirsath, A. (2017). A Review of Glass Break Detector System. *International Journal for Research & Development in Technology, 7*(3).

[16] Disaster Research Laboratory (2012). *Windows aerodynamic performance demonstration*. National Disaster Management Research Institute.

[17] Al-Ansari, L. S., Ali, M., Al-Hajjar, H., & Husam, J. A. (2018). Calculating the natural frequency of cantilever tapered beam using classical Rayleigh, modified Rayleigh and finite element methods. *International Journal of Engineering & Technology, 7*(4), 4866-4872.

[18] Korea Meteorological Administration (2015). *Longitudinal Meteorological Observation (ASOS) - Data*. Korea Meteorological Administration.

Improving Stadium Crowd Monitoring with a Smart Fusion of Vision-Based and Piezoelectric Sensor

Ana Claudinne E. Olivas
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
olivas.ace@gmail.com

Erdina Tyagita Utami
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
erdina.utami@si.itera.ac.id

Rizky Pitajeng
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
rizky.pitajeng@hotmail.com

Jong Bin Seong
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
whdqls759@naver.com

Seong Yeop Woo
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
yubi7456@naver.com

Dongwan Kim, Ph.D.
ICT Integrated Ocean Smart City Engineering
Dong-A University
Busan, South Korea
dongwankim@dau.ac.kr

***Abstract*— Crowd crush disasters are of significant concern for public safety. Stadiums, known for hosting large-scale events with massive crowds, are susceptible to crowd disasters. This study proposes a crowd detection system that addresses the accuracy limitations of vision-based technologies, such as CCTV. To improve accuracy, piezoelectric sensors are integrated alongside CCTV monitoring. These sensors accurately detect human walking pressure, while CCTV provides visual information. By analyzing the temporal characteristics of both data sources, data fusion enhances crowd monitoring. This approach offers a novel method to effectively monitor crowd dynamics, enabling timely intervention and prevention of crowd crush situations. The system contributes to public safety by overcoming accuracy drawbacks and ensuring comprehensive crowd monitoring.**

Keywords*— *crowd monitoring, CCTV, piezoelectric sensor, data fusion.

I. INTRODUCTION

Crowd crush is a phenomenon that occurs when the density of people in an area is high enough to put pressure on individuals that can cause the crowd to collapse on itself, causing each person to experience injuries or difficulty of breathing, among others. The crowd crush incident, sourced from multiple references [1,2], has been categorized as one of the most fatal events in history, considering its occurrence in various locations. Among all the locations, stadiums are classified in the top 3 most frequent locations where the crowd crush incidents have taken place after open areas and indoor building, respectively. With the increasing number of incidents each century, crowd crush remains an important area of study.

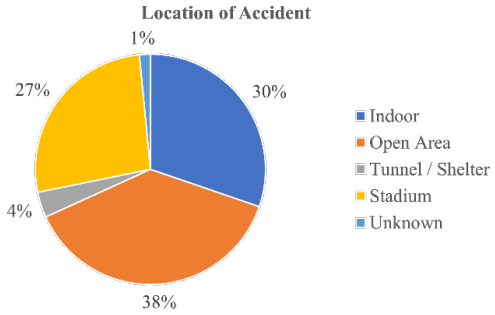


Fig. 1. Locations of deadliest crowd crush events in history

Several crowd detection systems [3,4] which have been studied include the use of vision-based technology (CCTV, drone, satellite), wireless based technology (RFID tag, mobile phone, Wi-Fi, Bluetooth), and social-media and web data mining based. Among them, vision-based systems using camera and CCTV have been widely used and developed as it is one of the oldest technologies in this domain. However, the use of cameras alone has some shortcomings in case of calculating the number of crowds. The system might encounter the loss of vision due to the brightness of the environment. Furthermore, the loss of vision can also be caused by the device position, overlapping object, and occlusion, hence the accuracy might decrease with the increase of crowd density [3,4,5]. To overcome this issue, the simultaneous use of other sensors alongside the use of camera is becoming more common. Some of them used the combination of camera and Wi-Fi signal strength measurement to increase the accuracy [6,7]. However, the use of Wi-Fi technology has also its drawback due to some uncertainties whether the Wi-Fi is switched on or if one person brings any or more than one mobile devices.

This study attempts to address the issue of occlusion by using piezoelectric sensor in combination of computer vision. By using this sensor, the crowd's pressure data in a certain area can be added into the computer vision data to verify the number of people in the area and increase accuracy.

II. METHODOLOGY

A. Crowd Monitoring System Design

We initiated an outline of technical aspect to monitor the crowd is depicted in Fig 1. In this study, a camera and a piezoelectric sensor were used as crowd monitoring devices. The end goal is to come up with a monitoring system that signals when the density in an area of interest reaches certain thresholds based on the Guide of Safety at Sport Grounds (2008), the criteria of crowd evacuation stability [8]:

- Stable: $D \leqslant 4$ NP/meter2
- Critical stable: 4 NP/meter2 $\leqslant D < 7$ NP/meter2
- Unstable, 7 NP/meter2 $< D$

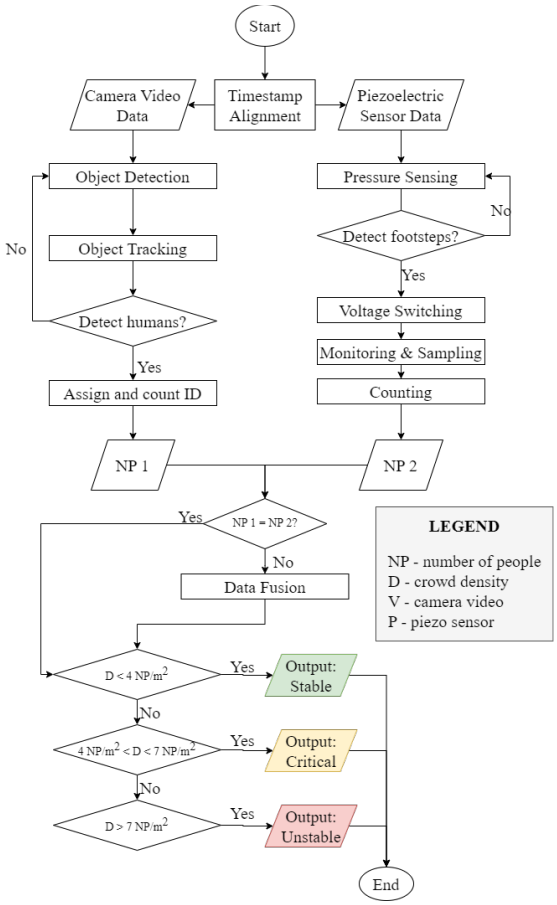


Fig. 2. Proposed reasearch scheme

To simulate the floor cover of a stadium gate, we have employed an acrylic panel with a 5 mm thickness and the dimension of 300 mm x 300 mm. In order to simulate the pressure of human footstep, we worked with 3 dolls, each with a mass of 35 grams. All the equipment would be linked into a personal computer and a microSD would be used to save the signal. Fig 3 is the setting arrangement for the small-scale crowd monitoring system.

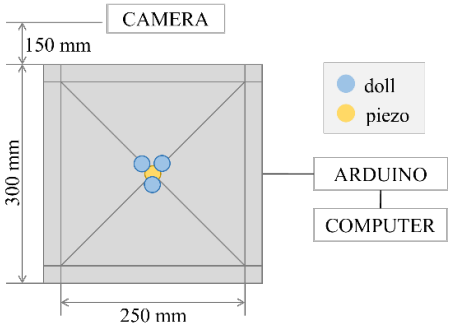


Fig. 3. Propose setting arrangement for small-scale crowd monitoring system

B. Image Processing

YOLOv4 was selected to implement the image processing phase of this study. YOLOv4 is a deep learning model used for object detection, one that is widely used for crowd counting. Figure 4 shows the YOLOv4 workflow with three main parts: backbone, neck, and head. The backbone is where features from the input image are extracted. At the neck, the extracted features are combined for a more thorough interpretation of the input image. The head takes the neck output for predicting the classes of the objects detected in bounding boxes.

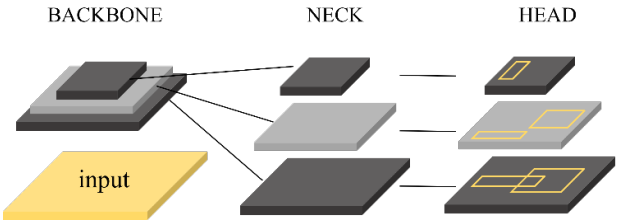


Fig. 4. YOLOv4 architecture one-stage detector

C. Vibration Signal Using Piezoelectric Sensor and Microcontroller

Our method, as depicted in Figure 3, utilizes a single piezoelectric sensor to detect footsteps and compensate for the camera's inaccuracies in identifying individuals at the gate. The sensor's high accuracy in detecting human walking pressure, exceeding 90% [9,10], makes it an ideal choice for our approach. We also utilized the aforementioned sensor to detect human footprints. In contrast to previous studies, an oscilloscope and signal generator was occupied [11] however, we used an Arduino-based microcontroller (ATmega328p) connected to a personal computer to display the process collecting the signal. The sampled signal was converted into range of voltage representing footprint generated by dolls stomping.

III. Expected Result

A. RGB Camera Data

YOLOv4 has been used to data processing for video camera. as can be seen in Fig. 5, YOLOv4 recognizes people and other objects from an RGB camera. If the distance between populations is less than 7NP/meter², people are well recognized, but if it is greater than that, the recognition number will fail. To do this, the area of the location of interest is fed into the code to compute the crowd density per unit of time.

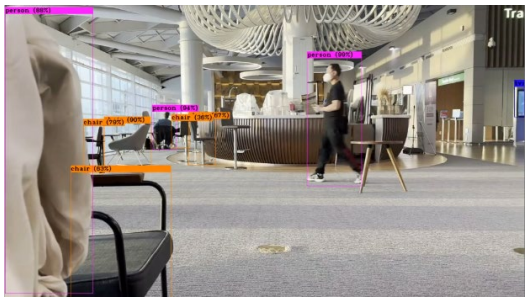


Fig. 5. Object detection using YOLOv4

B. Piezoelectric Sensor Data

The result of the experiment explained in section II shows a notable distinction in the voltage values generated by 1, 2, and 3 dolls. This finding strongly suggests the potential use of piezoelectric sensors as crowd detectors.

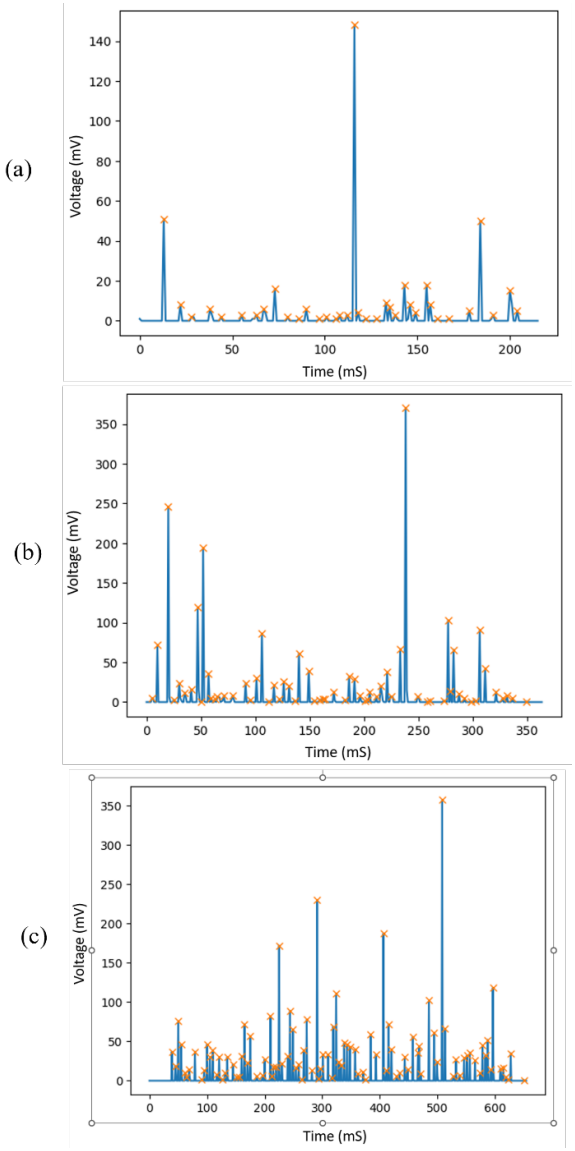


Fig. 6. Signal Processing from piezoelectric sensor for (a) 1 doll, (b) 2 dolls, and (c) 3 dolls

In the scenario where one doll is placed on the plate, a voltage of below 20mV is generated. When two dolls are present, the voltage mostly remains below 50mV, and with three dolls, it similarly stays below 70mV. This observation suggests that the average voltage can potentially provide information about the number of people. The collected signal data, as depicted in Fig. 5, undergoes analysis to extract relevant features and effectively handle outlier. Peak voltage value serves as one of the significant features that can be used to identify the number of people generating the signal. If the distance between the population is greater than 7NP/meter², the number of people measured by the piezoelectric sensor and the number of people recognized by the camera is compared, and the difference is determined that the camera does not recognize the person and finds the person.

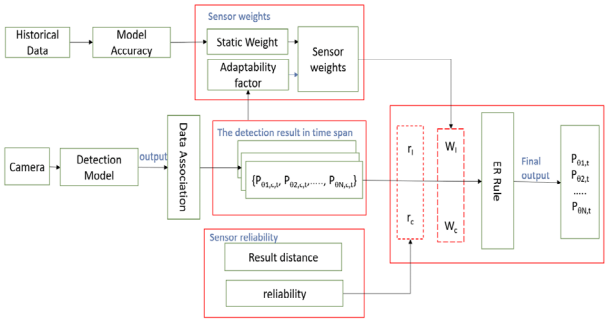


Fig. 7. Fusion Algorithm for vision based technology, and piezoelectric device.

IV. Conclusion and Future Work

This study proposes a crowd monitoring system involving computer vision and piezoelectric sensors, wherein the camera detects people in a crowd and the piezoelectric sensors detect the pressure of human footsteps, which is then associated with a number of people. This was done to address the issue of occlusion, or the blocking of people by other people as a limitation of the camera field of view and location. Currently, the system is able to detect and track people from a camera video.

There are several potential avenues for future research in the field. Some aspects that upcoming research must be considered, include:

1. Applying the Kalman Filter algorithm [12] for data processing. The technique offers a recursive approach to estimate the state of a dynamic system, even in the presence of uncertainty. This widely-used and valuable estimation method is particularly well-suited for implementation in linear systems due to its simplicity. However, when dealing with nonlinear systems, the recommended choice is to implement the extended version of Kalman filtering known as the Modification Kalman Filter.

2. Previous research has already explored wiring techniques in tile floors such as the development of a Smart Carpet system [13]. Building upon these existing findings, our future work aims to enhance and implement innovative wiring solutions that enable seamless integration of technology within floor surfaces.

Acknowledgment

This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Korea Ministry of Education (No.2016R1A6A1A03012812).

References

[1] Wikipedia (2023, April). “List of fatal crowd crushes” available : https://en.wikipedia.org/wiki/List_of_fatal_crowd_crushes.

[2] Word Crowd Disasater Web App (2023, April).” World Crowd Disaster” Available : https://yorku.maps.arcgis.com/apps/webappviewer/index.

[3] D. Sharma, A.P. Bhondekar, A.K Shukla, C.A. Ghanshyam,. “Review on Technological Advancements in Crowd Management.” *Journal of*

Ambient Intelligence and Humanized Computing, **2016**, 9(3), 485-495, doi: 10.1007/s12652-016-0432-x.

[4] U. Singh, J.F. Determe, F. Horlin, and P.D. Doncker. “Crowd Monitoring: State-of-the-Art and Future Directions.” IETE Technical Review (Institution of Electronics and Telecommunication Engineers, India), **2020**, 38(6), 578-594, doi: 10.1080/02564602.2020.1803152

[5] B.Yogameena, C. Nagananthini. “Computer Vision Based Crowd Disater Avoidance System: A Survey.” International Journal of Disaster Risk Reduction. 22, pp. 95-129.

[6] D. Duives, T. van Oijen, S. Hoogendoorn, “Enhancing Crowd Monitoring System Functionality Through Data Fusion: Estimating Flow Rate From W-Fi Traces and Automated Counting System Data,” 2020 Sensors, Switzerland, pp. 1-25.

[7] T. Miyaki, T. Yamasaki, K. Aizawa. “Multi-sensor fusion tracking using visual information and Wi-Fi location estimation.” In Proceedings of the 2007 1st ACM/IEEE International Conference on Distributed Smart Cameras, ICDSC, Vienna, Austria, 25–28 September 2007; pp. 275–282.

[8] Department for Culture, Media, and Sport, “ Guide to Safety at Sports Grounds”, 5th ed, United Kingdom: TSO, 2008, pp. 205-206.

[9] S. Akiyama, M. Yoshida, Y. Moriyama, H. Suwa and K. Yasumoto, "Estimation of Walking Direction with Vibration Sensor based on Piezoelectric Device," 2020 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Austin, TX, USA, 2020, pp. 1-6.

[10] Y. Kashimoto, M. Fujimoto, H. Suwa, Y. Arakawa and K. Yasumoto, "Floor vibration type estimation with piezo sensor toward indoor positioning system," 2016 International Conference on Indoor Positioning and Indoor Navigation (IPIN), Alcala de Henares, Spain, 2016, pp. 1-6.

[11] Y. Yu, X. Qin, S. Hussain, W. Hou, and T. Weis, “Pedestrian Counting Based on Piezoelectric Vibration Sensor,” Applied Sciences, vol. 12, no. 4, p. 1920, Feb. 2022, doi: 10.3390/app12041920.

[12] A. Alofi, A.A. Alghamdi, R. Alahmadi. N. Aljuaid. and Hemalatha, M. “A Review of Data Fusion Techniques. International Journal of Computer Applications”, 2017. 167, pp. 37-41.

[13] D.Savio and T.Ludwig.” Smart-Carpet: A Foot Step Tracking Interface. “21st International Conference on Advanced Information Networking and Applications Workshops(AINAW’07), Niagara Falls, ON, Canada,2007, pp. 754-760, doi: 10.1109/AINAW.2007.338.

Multiple Access in HAP-Enabled Maritime IoT Networks: A Deep Reinforcement Learning Approach

Thanh Phung Truong, Nam-Phuong Tran, Cuong Manh Ho, Thi Thu Hien Pham, Anh-Tien Tran, and Sungrae Cho

School of Computer Science and Engineering, Chung-Ang University, Seoul 06974, South Korea.
Email: {tptruong, tnphuong, hmcuong, ptthien, attran}@uclab.re.kr, srcho@cau.ac.kr.

***Abstract*—This study considers a maritime Internet of Things (IoT) network, where a high-altitude platform (HAP) is deployed to serve the IoT devices (IoTDs) on sailing ships. Here, the HAP allocates its communication resource to the ships using frequency division multiple access (FDMA), and IoTDs on each ship share the same communication resource allocated for their ship using non-orthogonal multiple access (NOMA). Accordingly, we formulate an optimization problem to maximize the system sum rate by optimizing the resource allocation at HAP and transmit power at IoTDs. Then, we propose a deep reinforcement learning (DRL) framework that applies the deep deterministic policy gradient algorithm to decide the variables. The simulation results illustrate the proposed algorithm's convergence and performance under various environmental parameters.**

***Index Terms*—maritime IoT networks, high-altitude platform, multiple access, deep reinforcement learning**

I. Introduction

Recently, maritime activities such as maritime tourism, oceanic mineral research, and offshore fishing have grown significantly, creating a demand for IoT services [1]. In response to this observation, aerial radio access networks emerge as a potential option for supplementing terrestrial communication systems, which support global access, including maritime communications [2]. Therefore, the research on maritime IoT networks with the assistance of aerial networks has increased recently [3]–[5]. In particular, the work in [3] considered a combined satellite-terrestrial communication infrastructure to assist intelligent vessel traffic services. The authors introduced a two-phase data-driven machine learning framework aimed at improving the quality of vessel trajectory records obtained from automatic identification system networks. The authors in [4] also investigated a maritime IoT system with the combined satellite-terrestrial automatic identification system base stations, where they proposed a long short-term memory-based prediction framework that predicts the spatiotemporal vessel trajectories. In [5], the authors considered a maritime IoT network with the help of an unmanned aerial vehicle (UAV), where unmanned surface vehicles (USVs) offload their tasks to the UAV using NOMA. They focused on a problem that minimizes energy consumption by optimizing resource allocation for computing, USVs offloading strategy, transmit power, and UAV trajectory with the latency requirement. Multiple access techniques such as NOMA are crucial in the next-generation network. However, to the best of our knowledge, the consideration of multiple access techniques in maritime IoT networks is still limited. Therefore, in this work, we study a multiple access scheme that combines FDMA and NOMA in a maritime IoT network with the assistance of a HAP. The main contribution of our work are summarized as follows:

- We investigate a HAP-enabled maritime IoT network scenario, where a multiple access scheme that combines FDMA and NOMA is introduced. Here, we formulate a maximization problem of system sum rate where the communication resource allocation at HAP and the transmit power at IoTDs are considered optimization variables.
- To address the issue, we present a deep reinforcement learning framework that uses the deep deterministic policy gradient algorithm to determine optimization variables, where problem constraints are satisfied using the softmax function.
- To analyze the system performance, we simulate a scenario with the variation in the environmental parameters. The outcomes of the simulation reveal how the proposed algorithm achieves convergence and demonstrates system performance across a range of parameters.

The rest of this work are arranged as follows. We formulate the problem in Section II, where the proposed framework is introduced in Section III. The simulation is presented in Section IV. Section V concludes this work.

II. Problem Statement

We consider the uplink communication in a HAP-enabled maritime IoT network system shown in Fig. 1, where the HAP serves N ships sailing on the sea. At each ship $n, n \in \mathcal{N} \triangleq \{1, 2, \ldots, N\}$, K_n IoTDs connect to the HAP, which form a virtual cluster. Accordingly, the HAP simultaneously

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-RS-2023-00258639) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

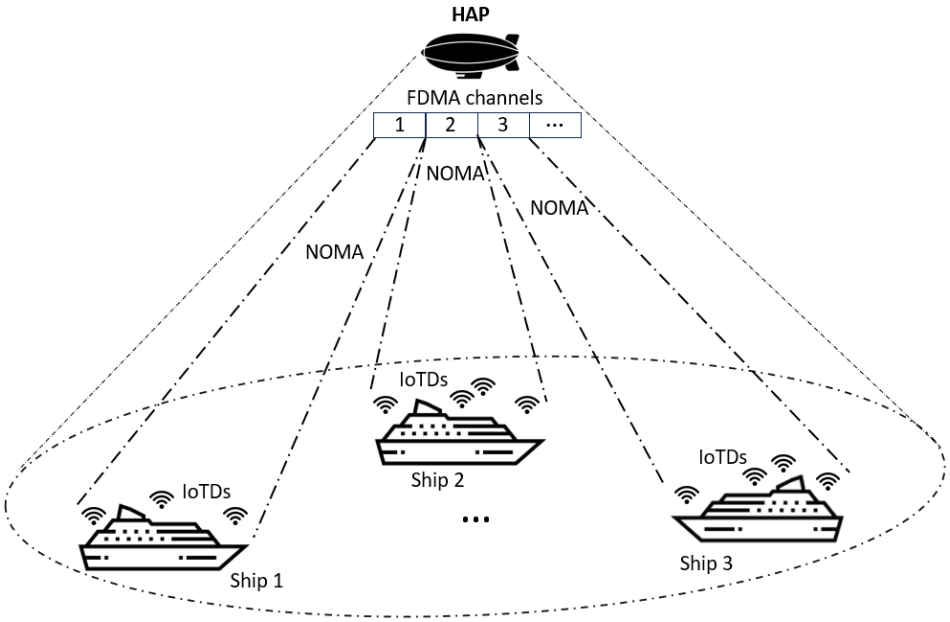


Fig. 1: HAP-Enabled Maritime IoT Network System.

serve N clusters, each has K IoTDs. To enhance the uplink communication performance, we deploy a multiple access that combines the FDMA and NOMA to the system, where the HAP divides its communication resource into N sub-channels, each for one cluster. In each cluster n, K_n IoTDs share the same communication resource intended for their ship, which is enhanced by applying the NOMA technique. We consider the HAP to have perfect channel state information (CSI), allowing it to progressively decode each message after receiving them from IoTDs using the successive interference cancellation (SIC) technique [6]. Without sacrificing generality, we choose the decoding order following the channel gains between IoTDs and HAP in descending order [7]. As the air-to-ocean communication, the channel between $k-$th IoTD in cluster n and HAP is considered as the Line-of-Sight (LoS) propagation, where the channel gain can be expressed as [8]

$$h_{(n,k)} = \frac{h_0}{d^{\alpha}_{(n,k)}}, \tag{1}$$

where h_0 denotes the channel gain of LoS propagation at the reference distance $d_0 = 1$m and α denotes the path loss exponent.

Accordingly, considering IoTDs in cluster n are ordered based on the descending order of channel gain, i.e., $h_{(n,1)} \geq h_{(n,2)} \geq \cdots \geq h_{(n,K_n)}$, the transmission rate of $k-$th IoTD in $n-$th cluster is calculated as

$$r_{(n,k)} = f_n B \log_2 \left(1 + \frac{p_{(n,k)} h_{(n,k)}}{\sum_{i=k+1}^{K_n} p_{(n,i)} h_{(n,i)} + \sigma^2} \right), \tag{2}$$

where B denotes the HAP communication bandwidth, f_n is the fraction of the HAP bandwidth allocated to cluster n, $p_{(n,k)}$ is the corresponding transmit power, and σ^2 is the power of the additive white Gaussian noise.

This study aims to maximize the sum rate of all IoTDs in all cluster by optimizing the resource allocation at HAP and the transmit power of IoTDs. By denoting $\mathcal{F} = \{f_n, n \in \mathcal{N}\}, \mathcal{P} = \{p_{(n,k)}, n \in \mathcal{N}, k \in \{1, 2, \ldots, K_n\}\}$, we formulate the system sum rate maximization problem as

$$(P): \max_{\mathcal{F},\mathcal{P}} \sum_{n\in\mathcal{N}} \sum_{k=1}^{K_n} r_{(n,k)}, \tag{3a}$$

$$\text{s.t.} \sum_{n\in\mathcal{N}} f_n = 1, \tag{3b}$$

$$p_{(n,k)} \leq P_{max,(u,k)}, \tag{3c}$$

where the constraint in (3b) indicates that the sum of allocated resource cannot exceed the resource capacity at HAP, and constraint (3c) is to ensure the maximum transmit power of each IoTD. Due to the non-convex constraint in (3b) and the massive number of real-time observations from the dynamic environment, it is demanding to address the problem using traditional methods. Therefore, we introduce a DRL framework for solving the optimization problem.

III. Proposed DRL Framework

We model the problem as a reinforcement learning (RL) model [9], with HAP acting as the RL agent and the entire system acting as the environment. Here, we define the state space, action space, and reward function at each time slot t as follows.

- State space: The state space indicates the environment observation affecting the model's reward. Therefore, in this work, we define the state space including all IoTDs' channel gain information:

$$s[t] = \{h_{(n,k)}[t], n \in \mathcal{N}, k \in \{1, 2, \ldots, K_n\}\}. \tag{4}$$

Algorithm 1 Proposed DRL-based Algorithm

Initialize environment and neural networks.
for $e \in \mathbf{E}$ **do**
  Obtain the first state $s[t]$.
  while in episode **do**
    Interactive process:
      Determine action $a[t]$ using main actor network.
      Calculate $p'_{(n,k)}$ and f'_n.
      Perform $p'_{(n,k)}$, $f'_n \rightarrow$ obtain $s[t+1]$, $r[t]$.
      Store $\{s[t], a[t], r[t], s[t+1]\}$ into buffer for training.
      Update state $s[t] \leftarrow s[t+1]$.
    Training process:
      Get random sample batch form buffer.
      Update parameters θ^{μ} and θ^{Q} by (7) and (8).
      Update parameters of target networks by (9).
  end while
end for
return the trained actor networks θ^{μ^*}.

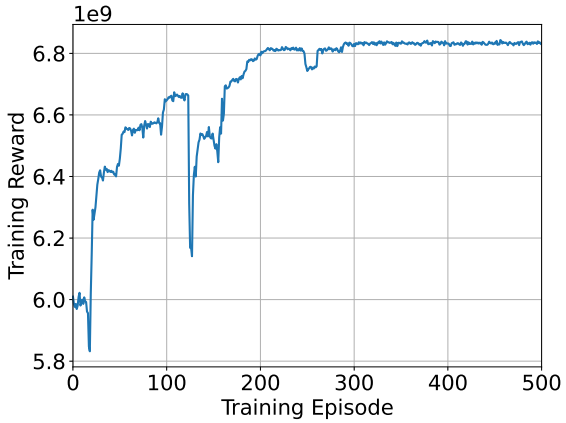


Fig. 2: Training convergence.

- Action space: It contains all optimization variables the agent has to decide at each time slot, which includes resource allocation and transmit power variables:

$$a[t] = \{\mathcal{F}[t], \mathcal{P}[t]\}. \tag{5}$$

- Reward function: The reward at each time slot is defined by the system sum rate, which is calculated as

$$r[t] = \sum_{k=1}^{K_n} r_{(n,k)}. \tag{6}$$

To decide the action, we apply the deep deterministic policy gradient algorithm to train the agent, which includes four neural networks: main and target actor networks $\left(\mu(s|\theta^{\mu}) \text{ and } \mu'(s|\theta^{\mu'})\right)$, and main and target critic networks $\left(Q(s,a|\theta^{Q}) \text{ and } Q'(s,a|\theta^{Q'})\right)$, where $\theta^{\mu}, \theta^{\mu'}, \theta^{Q}, \theta^{Q'}$ are the corresponding networks parameters [10]. The agent observes the environment's state and decides the action using the main actor network, which parameters are trained by the policy gradient:

$$\nabla_{\theta^{\mu}} J = \frac{1}{B} \sum_{i=1}^{B} \left(\nabla_a Q(s,a|\theta^{Q})|_{s=s_i, a=\mu(s_i)} \nabla_{\theta^{\mu}} \mu(s_i|\theta^{\mu}) \right), \tag{7}$$

where B and s_i denote the size and the state in sample i of the training sample batch, respectively. The the main critic network parameters are updated by minimizing the loss function given as

$$L = \frac{1}{B} \sum_{i=1}^{B} \left(y_i - Q(s_i, a_i|\theta^{Q}) \right)^2, \tag{8}$$

where $y_i = r_i + \gamma Q'\left(s'_i, \mu'(s'_i|\theta^{\mu'})|\theta^{Q'}\right)$, $\gamma \in [0,1]$ denotes the discount factor, a_i, s'_i, and r_i are action, next state, and reward in sample i of the sample batch. The parameters of target networks are calculated according to a soft update:

$$\begin{aligned} \theta^{\mu'} &\leftarrow \tau\theta^{\mu} + (1-\tau)\theta^{\mu'}, \\ \theta^{Q'} &\leftarrow \tau\theta^{Q} + (1-\tau)\theta^{Q'}, \end{aligned} \tag{9}$$

where τ denotes the soft update coefficient. In addition, the training process explores the samples via additional noise to the actor policy as follows:

$$a[t] = \mu(s[t]|\theta^{\mu}) + \mathcal{OU}[t], \tag{10}$$

where $\mathcal{OU}(t)$ follows the Ornstein-Uhlenbeck process. The action determined from the neural network can be scale to a specific value range by activation function. Due to constraint (3c), we scale the range of action to the interval $[0,1]$, the decided transmit power is then calculated as

$$p'_{(n,k)} = p_{(n,k)} P_{max,(u,k)}, \tag{11}$$

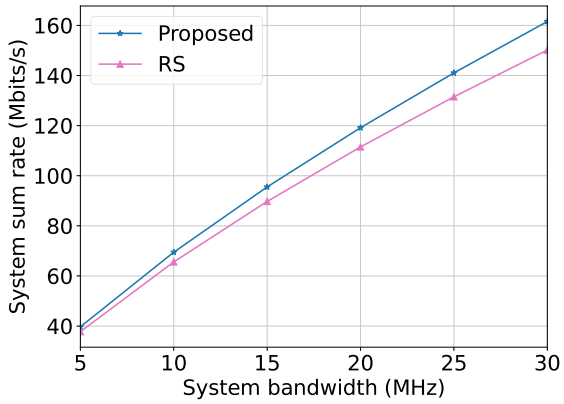
where $p_{(n,k)} \in [0,1]$ is determined from the main actor network. Furthermore, to satisfy the constraint (3b), we apply the softmax function to the resource allocation variables, the decided resource allocation can be expressed as

$$f'_n = \text{softmax}(\mathbf{f})_n = \frac{e^{f_n}}{\sum_{i \in \mathcal{N}} e^{f_i}}, \tag{12}$$

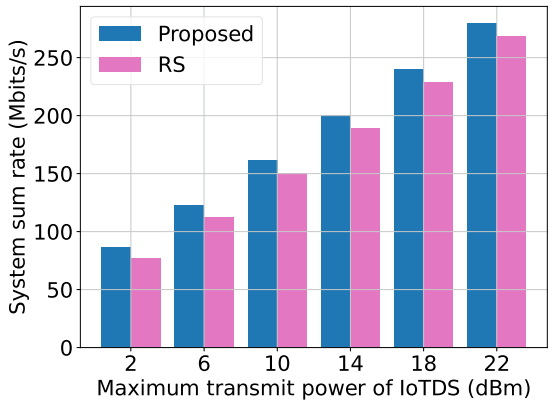
where $\mathbf{f} = \{f_n, n \in \mathcal{N}\}$. Consequently, Algorithm 1 depicts the proposed algorithm, which runs over the E episodes. In each interactive step, the agent observe the environment state $s[t]$, determines the action by using the main actor network as in (10). Then, the transmit power and resource allocation variables are calculated according to (11) and (12) to satisfy the problem constraints. Accordingly, the agent performs decided variables and get the reward and the next state. The environment state is then updated for the next interaction. Besides, a tuple of experience is put into a buffer after each interaction for preparing samples for the training process, where the neural networks' parameters are updated. Finally, the trained main actor network is obtain to interact with the environment.

IV. Performance Evaluation

In this section, we simulate a scenario including 5 ships, each containing 40 IoTDs, to evaluate the proposed framework performance, the HAP serves a range of 500 (meters) radius at an altitude of 20 (Km). The other parameters are set as:



(a) Communication bandwidth.



(b) Maximum transmit power.

Fig. 3: Performance evaluation results.

$h_0 = -30$ (dBm), $\alpha = 2$, $\sigma^2 = -174$ (dBm/Hz). The communication bandwidth varies from 5 to 30 MHz, the maximum transmit power of each IoTD is in the range between 2 (dBm) and 22 (dBm). Furthermore, we define a random scheme to compare the performance of the proposed algorithm. The random scheme is defined as

- Random scheme (RS): Instead of using the actor network, this scheme chooses the action randomly.

First, we analyze the convergence of the proposed algorithm, where the neural networks are built with two hidden layers, with the number of nodes in the first and second layers being 512 and 256, respectively. The other parameters are set as: actor learning rate $= 0.001$, critic learning rate $= 0.01$, $\gamma = 0.99$, $\tau = 0.1$, $B = 16$, buffer size $= 50000$. We train the model for 500 episodes, each has 100 steps, the training result is shown in Fig. 2. The model converges after about 300 episodes, where its reward is stable after this point and hits a value of 6.82×10^9.

Next, we evaluate the system's performance with the variation of environmental parameters, the results are illustrated in Fig. 3. As shown in Fig. 3a, the system sum rate increases with the rise in bandwidth, where it gains about four times when expanding bandwidth from 5 (MHz) to 30 (MHz). In Fig. 3b, we illustrate the change in system sum rate when increasing the maximum transmit power of IoTDs from 2 (dBm) to 22 (dBm). As a result, the system sum rate rises when the transmit power increases, rising by about 39 (Mbits/s) as the power gains by 4 (dBm). Besides, the proposed scheme's performance consistently outperforms the RS in both results.

V. Conclusion

This study considered a maritime IoT network scenario, where a HAP hovers in the sky and serves IoTDs in sailing ships. Here, we investigated multiple access that combines the FDMA and NOMA techniques to assist uplink communication. In such a system, we formulated a problem of maximizing the system sum rate by optimizing the communication resource allocation at HAP and the transmit power of IoTDs. To resolve the problem, we proposed a DRL framework that applied the deep deterministic policy gradient algorithm. The simulation results analyzed the proposed algorithm's convergence and performance with different environmental parameters.

References

[1] T. Wei, W. Feng, Y. Chen, C.-X. Wang, N. Ge, and J. Lu, "Hybrid satellite-terrestrial communication networks for the maritime internet of things: Key technologies, opportunities, and challenges," *IEEE Internet of Things Journal*, vol. 8, no. 11, pp. 8910–8934, 2021.

[2] N.-N. Dao, Q.-V. Pham, N. H. Tu, T. T. Thanh, V. N. Q. Bao, D. S. Lakew, and S. Cho, "Survey on aerial radio access networks: Toward a comprehensive 6g access infrastructure," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 2, pp. 1193–1225, 2021.

[3] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, and M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6g-enabled maritime iot systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5374–5385, 2021.

[4] R. W. Liu, M. Liang, J. Nie, W. Y. B. Lim, Y. Zhang, and M. Guizani, "Deep learning-powered vessel trajectory prediction for improving smart traffic services in maritime internet of things," *IEEE Transactions on Network Science and Engineering*, vol. 9, no. 5, pp. 3080–3094, 2022.

[5] L. P. Qian, H. Zhang, Q. Wang, Y. Wu, and B. Lin, "Joint multi-domain resource allocation and trajectory optimization in uav-assisted maritime iot networks," *IEEE Internet of Things Journal*, vol. 10, no. 1, pp. 539–552, 2023.

[6] Y. Liu, Z. Ding, M. Elkashlan, and H. V. Poor, "Cooperative non-orthogonal multiple access with simultaneous wireless information and power transfer," *IEEE Journal on Selected Areas in Communications*, vol. 34, no. 4, pp. 938–953, 2016.

[7] V. D. Tuong, T. P. Truong, T.-V. Nguyen, W. Noh, and S. Cho, "Partial computation offloading in noma-assisted mobile-edge computing systems using deep reinforcement learning," *IEEE Internet of Things Journal*, vol. 8, no. 17, pp. 13 196–13 208, 2021.

[8] T.-H. Nguyen, T. P. Truong, N.-N. Dao, W. Na, H. Park, and L. Park, "Deep reinforcement learning-based partial task offloading in high altitude platform-aided vehicular networks," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 1341–1346.

[9] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*. The MIT Press, 2018.

[10] T. P. Lillicrap, J. J. Hunt, A. Pritzel, N. Heess, T. Erez, Y. Tassa, D. Silver, and D. Wierstra, "Continuous control with deep reinforcement learning." in *ICLR (Poster)*, 2016.

CNN and RNN Based Deep Learning Methods for Fish's Appetite Detection

Agape D'sky[‡], Infall Syafalni[‡§], Nana Sutisna[‡§], Hadi Hariyanto[¶], Trio Adiono[‡§], and Yeon Ho Chung [‖]
[‡]School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia
[§]University Center of Excellence on Microelectronics, Bandung Institute of Technology, Indonesia
[¶]Department of Information Systems, Telkom University, Indonesia
[‖] Pukyong National University, Korea
e-mail: agapedsky@students.itb.ac.id

***Abstract*—In this paper, several deep learning models are created to assess fish's appetite. Several works that had been done before are also explained within. A 2D-CNN model and a 3D-CNN-RNN model were built for this experiment. Both models can accurately determine the fish's appetite based on the given labels, $i.e.$ with 99.34% and 97.83% accuracy respectively. In terms of functionality attainment, the 2D-CNN model performs better than the other model.**

I. Introduction

Sustainable Development Goals (SGD) are one of the biggest concerns for human beings, and one of these goals is to realize zero hunger [1]. Humans need food as one of their primary needs, which is essential to achieving living sustainability. However, these recent decades have been showing us that without an effort to control the food production and consumption rate, the food itself will eventually run out, thus leading to a Malthusian Catastrophe [2]. This control can be manifested in lots of different disciplines, and maritime is one of them (regarding fish and other seafood).

In 2021, the global fish consumption rate increased by 1% compared to 2018-2020. Furthermore, it has been estimated that in 2030, the consumption rate will increase by 11% [3]. However, the food production rate does not compete with it. The food production rate grows in a more fluctuating manner with many external factors influencing it.

The two main contributors to the fish production rate are wild capture and aquaculture. Among these two, aquaculture is the one that can be exploited thoroughly without causing environmental problems if done the right way. Still, there often are flaws in its operation, mostly caused by the conventional method being used, thus lessening its efficiency.

In aquaculture, the feed usually takes about 60%-70% to the total budget [3], while 60% of the given feed actually gets wasted. Facing this problem, lots of ideas using technology have been developed and implemented to decrease the production cost without much altering the growth of the fish.

This paper discusses the development of the implementation of Artificial Intelligence methods to solve fish feeding problems using deep learning analysis.

II. Preliminaries

A. Fundamental Concepts

1) Fish's Appetite Indicator: Lots of aspects can be assessed to determine the fish's appetite, prior to and during the feeding. This was documented based on observations and several interviews at lotic aquaculture $tilapia$ ponds in West Java, Indonesia.

Prior to the feeding, the fish's hunger can actually be seen through their gathering behavior. In an intensive aquaculture system, fish will gather around the feeding spot at the usual feeding time. Also, the water quality often represents the fish's appetite quite well. For example, in a low level of dissolved oxygen (DO), fish usually tends not to eat. During the feeding, the absence of rapid movement from the fish (including jumps and splashes) reflects that the process should be ended.

2) Conventional Aquaculture: In $tilapia$ ponds located in West Java, most fish farmers use sole intuition to determine everything, $i.e.$ the feeding dose, time, water quality parameters, etc, thus increasing the feed waste upon feeding. They use their experiences to determine the fish's appetite and water quality using bare hands.

Recent emerging technologies are definitely able to fix these conventional ways of aquaculture. Automation can deliver accurate measurements to control feeding doses. Furthermore, the utilization of Artificial Intelligence (AI) and Internet of Things (IoT) are suitable to decide feeding time and decisions.

B. Previous Works

Previous works on this topic mainly talk over feeding control using modified actuators and AI.

Chang, et al. (2005) [4] built a nonabsolute time-based eel-feeding controller using IR sensors. The main idea of this system is that the eel's body can reflect IR waves. In an intensive aquaculture pond, the flocking behavior of the eels will present during the usual feeding time, thus alarming the sensor regarding the appetite of the eels. This method, according to the claims, performs well in an intensive recirculating system with environmental parameters adjusted. However, the usage of photoelectric sensors only provides feedback regarding the fish's flocking behavior, which might not be sufficient to assess fish appetite without control over environmental parameters.

Zhao et al. (2017) [5] developed a computer vision to assess fish's appetite based on their spontaneous behavior upon stimulus. The system was built specifically for Recirculating Aquaculture Systems (RAS), where every environmental parameter is controlled so that the water will not get turbid. A camera is placed above the pond to detect the fish's movement and location. The spontaneous movement of the fish upon stimulus will be indicated as a high appetite condition. The limitation of this system lies in its reliance on the color feature.

Subakti et al. (2017) [6] used an accelerometer to measure the water surface's motion during the feeding time. The water surface's motion tends to be fluctuating during feeding due to the flocking and splashing behavior of the fish. This method is effective for conventional ponds since the water's turbidity will not bother the sensor's reliability. Nevertheless, in an open aquaculture system, the water surface's motion sometimes does not represent the flocking behavior of the fish, *e.g.* rain, harsh wind, *etc.*

Hu et al. (2022) [7] also developed a computer vision-based feeding system to detect the water surface motion to assess fish's appetite. This system can be applied to larger ponds. The result of this research shows the reliability of deep learning applications in aquaculture, especially using videos as its input. That being said, the author is interested in the usage of different types of deep learning methods for aquaculture. Additionally, the experiment takes place in a smaller-scale aquaculture system to verify the method's versatility.

III. Proposed Method

The proposed method utilizes a few deep learning models using images and videos (sequences of images) as its input. Videos are taken at the feeding spots to describe the fish's appetite by assessing the fish's movement and gathering behavior upon feeding. The setup illustration is given in Fig. 1. For the experiment, a medium-sized aquarium (50 cm x 30 cm x 25 cm) is used, containing 60 fish (*nile tilapia*). A set of filter is installed inside the aquarium to assure the water quality so that the fish can survive.

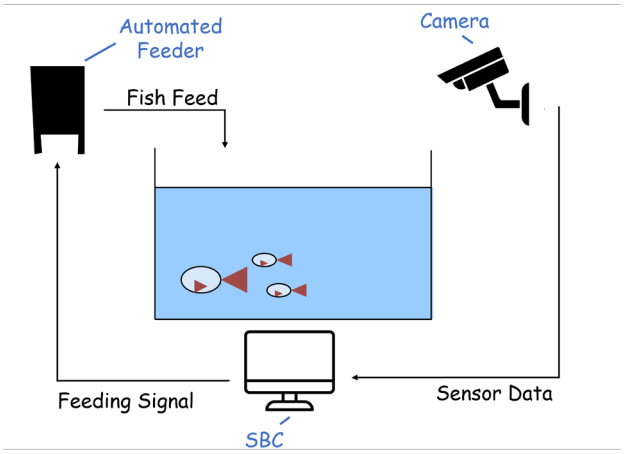


Fig. 1. Proposed Method Setup

A. Data Acquisition

The data are taken manually using a smartphone. Videos are taken and parsed into several images. The videos are also parsed into shorter videos (creating a different kind of dataset), with each containing 20 frames. The parsed images and videos then are cropped and rescaled into a specific dimension.

The data is divided into 2 classes (binary: high and low appetite), and several procedures are held in order to obtain the desired data. The high appetite data are taken during the feeding process right after the fish are starved for more than 8 hours. The low appetite data are shot subsequent to the feeding. The low appetite data are also taken when no significant activity occurs inside the aquarium to introduce more variance.

B. Deep Learning Method

There are 2 different models that are used in this experiment, *i.e.* a standard 2-dimensional Convolutional Neural Network (CNN) model (see Fig. 2) and a 3-dimensional Convolutional-Recurrent Neural Network (CNN-RNN) model (see Fig. 3). As the name suggests, each model uses a different type of data for its input (images and videos). The two models are compared to each other to determine the best model in terms of performance in the experiment.

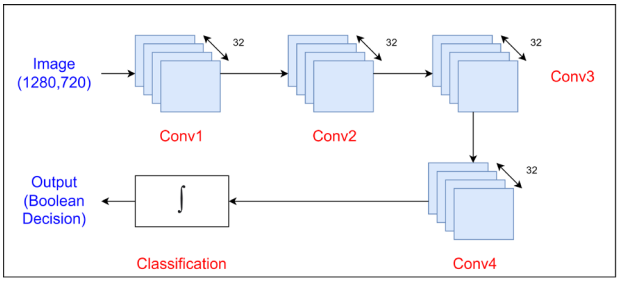


Fig. 2. 2-Dimensional CNN Architecture

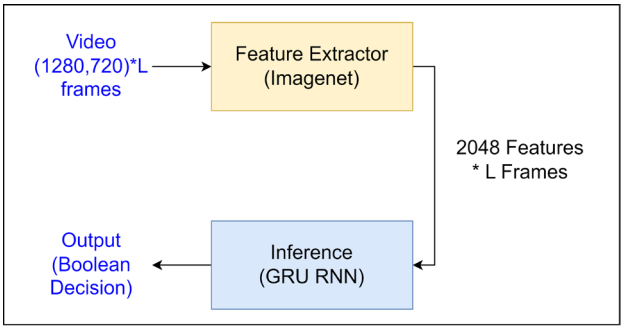


Fig. 3. 3-Dimensional CNN-RNN Architecture

A convolutional neural network layer can be described as a layer that processes a value inside a pixel and the values adjacent to that pixel with a set of values inside an operator (often called a kernel), thus gaining a specific feature from that pixel based on the kernel values, as shown in Fig. 4. At the end of this operation, an activation function is applied to the output

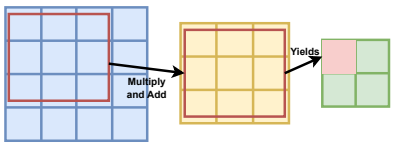


Fig. 4. A 2-Dimensional CNN Illustration

to introduce nonlinearity. This algorithm can be expanded in higher dimension problems (3, 4, and so on), simply by adding another degree of freedom to its input and kernel.

The model that uses video as its input utilizes a 2-dimensional algorithm to extract certain features from the image. For feature extraction, a pre-trained model from Imagenet is used. Obtained features are then fed to inference layers. These layers utilize a gated recurrent unit (GRU) [8], a sub-algorithm of the recurrent neural network (RNN) model. RNN itself is a network similar to an artificial neural network that passes the previous calculation to be considered on the next computation.

Throughout the model development, several parameters are tuned to give the best performance while still considering the computational resource of the model. These parameters include layer sizes, number of filters, and activation functions. In a certain layer, layer size represents the amount of information that is fed into the process, the number of filters represents the quantity of extracted features as the output of the layer, and the activation function helps to introduce nonlinearity to the system. On the other hand, the training parameters, such as loss functions and learning rate, are also adjusted empirically by means of trial and error.

IV. Experimental Results

The first model (2-D CNN model) detailed architecture is shown in Table 1. This model consists of 4 convolution layers, followed by pooling layers. This model takes 1280*720 pixel colored images as inputs and folds them into a batch of 6 images. Training is done using 1082 images, with 419 labeled as high appetite and the other 663 labeled as low appetite. Finally, the model outputs an array of 2 floating points, describing the inference result based on the value (one respects to hungry condition, and the other one to full condition).

The data that is used in this experiment shows distinctive differences for each condition (See Fig. 5). If the fish are hungry, the color of the fish becomes a bit sharper because the fish collectively tend to go upward the tank. When they are full, the color becomes more subtle. Also, if the fish are overfed, the remaining pellets can also be an indicator of the fish's appetite (Fig. 6).

Testing data are separated from the training data, containing 39+113 images from the same aquarium, showing the low appetite and high appetite data respectively. It is used to test the reliability of the model.

The second model (3-D CNN-RNN model), on the other hand, uses the architecture described in Table 2. As implied in Section III.A, this model's input videos are the same data used for the 2D-CNN model, but packed as sequences of frame per data. This way, the value of the information that is fed to the model persists, while the number of data decreases a lot (only 165 video data remain). The masking method is also used so that videos with less than 20 frames can still be included in the training and testing procedure.

Fig. 7 and 8 show the training result. In terms of training loss, the 2D-CNN model converges a lot faster than the 3D-CNN-RNN model. This most likely happens because there are a lot more extracted features gained from the InceptionV3,

TABLE I
Detailed 2-D CNN Architecture

Function	Input Dimension	Output Dimension
Input Layer	1280*720*3	1280*720*3
Rescaling	1280*720*3	1280*720*3
$Conv2D_1$	1280*720*3	1276*716*32
Max $Pooling_1$	1276*716*32	425*238*32
$Conv2D_2$	425*238*32	423*236*32
Max $Pooling_2$	423*236*32	141*78*32
$Conv2D_3$	141*78*32	139*76*32
Max $Pooling_3$	139*76*32	46*25*32
$Conv2D_4$	46*25*32	44*23*32
Global Average Pooling	44*23*32	32
Output Layer	32	2

Architecture used: 4 layer of convolutions with average pooling and softmax classification function. Using 20 epoches on training

TABLE II
Detailed 3-D CNN-RNN Architecture

Function	Input Dimension	Output Dimension
Resize and Rescale	20*1280*720*3	20*200*2000*3
Imagenet InceptionV3	20*200*2000*3	20*2048
GRU_1	20*2048	20*16
GRU_2	20*16	8
Dropout	8	8
Dense	8	2

Architecture used: Feature extractor using Imagenet InceptionV3 model, followed by 2 layers of GRU and a classification layer. Using 20 epoches on training

Fig. 5. Sample Training Data (left) Actively Feeding (right) Full

Fig. 6. Sample Training Data (Highly Overfed Condition)

TABLE III
TESTING RESULT

Model	2D-CNN	3D-CNN-RNN
Parameters	30,242	99.882
Training Data Count	1082	165
Testing Data Count	152	46
Loss	0.0146	0.1068
Accuracy	0.9934	0.9783

thus dramatically increasing the model's complexity. As for the training accuracy, there seems not to be any significant difference between the two.

The final testing result of each model as well as other essential information is summarized in Table 3. Similar to the training result, the testing result shows that the two models do not differ significantly in terms of loss and accuracy. The number of parameters used in the 3D-CNN-RNN model is way higher than the 2D-CNN model since the InceptionV3 layer is pre-trained, and thus, the latter performs well in terms of attainment of the desired functionality. It is still not evident enough, however, to say that the 2D-CNN model is better than the 3D-CNN model in a more general condition. It is notable that the 3D-CNN-RNN model data count is way lower than the 2D-CNN model. This means that with fewer number of inference, the 3D-CNN-RNN model is still able to catch up with the other model.

V. CONCLUSION

From the progress mentioned, there are several conclusions to be taken:

1) The proposed 2D-CNN and 3D-CNN-RNN model can successfully determine the fish's appetite based on the given label with 99.34% and 97.83% testing accuracy respectively
2) In terms of functionality, the 2D-CNN model excels better than the 3D-CNN-RNN model since it achieves the training target ($i.e.$ high accuracy) with much less parameters and the same value of information.

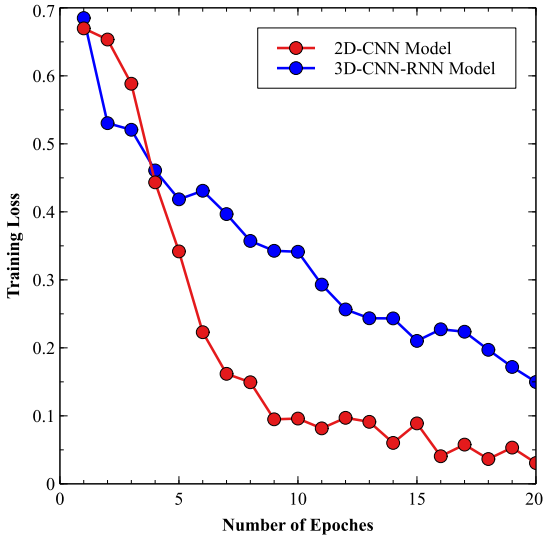


Fig. 7. Models' Training Loss

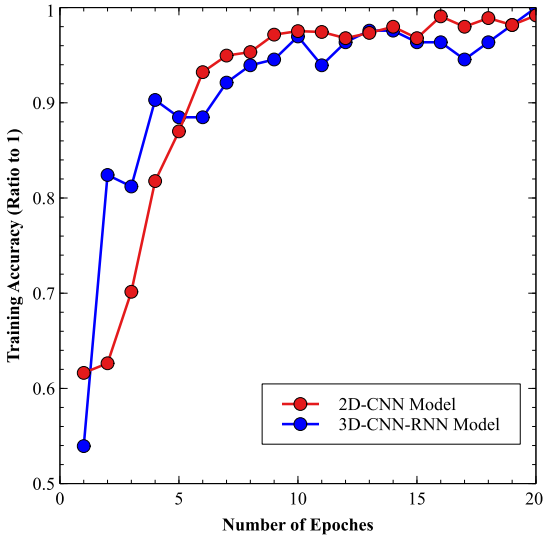


Fig. 8. Models' Training Accuracy

REFERENCES

[1] United Nations, The 2030 Agenda and the Sustainable Development Goals: An opportunity for Latin America and the Caribbean (LC/G.2681-P/Rev.3), Santiago, 2018.

[2] Ramya, R., Babu, C. C., Akshay, P. (2020). *Population Backfire: Hindsight to Malthusian Catastrophe*. Shanlax International Journal of Arts, Science and Humanities, 8(1), 121–126. https://doi.org/10.34293/sijash.v8i1.3248

[3] Organization for Economic Co-operation and Development. and Food and Agriculture Organization of the United Nations., *OECD-FAO Agricultural Outlook 2021-2030*. 2021.

[4] Chang, C. M., Fang, W., Jao, R. C., Shyu, C. Z., Liao, I. C. (2005). Development of an intelligent feeding controller for indoor intensive culturing of eel. Aquacultural Engineering, 32(2), 343–353. https://doi.org/10.1016/j.aquaeng.2004.07.004

[5] Zhao, J., Bao, W. J., Zhang, F. D., Ye, Z. Y., Liu, Y., Shen, M. W., Zhu, S. M. (2017). *Assessing appetite of the swimming fish based on spontaneous collective behaviors in a recirculating aquaculture system*. Aquacultural Engineering, 78, 196–204. https://doi.org/10.1016/j.aquaeng.2017.07.008

[6] Subakti, A., Khotimah, Z. F., Darozat, F. M. (2017). *Preliminary study of acceleration based sensor to record nile tilapia (Oreochromis niloticus)*

feeding behavior at water surface. Journal of Physics: Conference Series, 795(1). https://doi.org/10.1088/1742-6596/795/1/012060

[7] Hu, W. C., Chen, L. B., Huang, B. K., Lin, H. M. (2022). *A Computer Vision-Based Intelligent Fish Feeding System Using Deep Learning Techniques for Aquaculture*. IEEE Sensors Journal, 22(7), 7185–7194. https://doi.org/10.1109/JSEN.2022.3151777

[8] Cho, K., van Merrienboer, B., Gulcehre, C., Bahdanau, D., Bougares, F., Schwenk, H., Bengio, Y. (2014). *Learning Phrase Representations using RNN Encoder-Decoder for Statistical Machine Translation*. http://arxiv.org/abs/1406.1078

[9] Li, D., Wang, Z., Wu, S., Miao, Z., Du, L., Duan, Y. (2020). *Automatic recognition methods of fish feeding behavior in aquaculture: A review*. In Aquaculture (Vol. 528). Elsevier B.V. https://doi.org/10.1016/j.aquaculture.2020.735508

Machine Learning based Newly Built Ship Fuel Consumption Estimation

Kihun Shin
School of Computer Science and Engineering
Pusan National University
Busan, Korea
skh2929209@pusan.ac.kr

Dongkeun Jeon
School of Computer Science and Engineering
Pusan National University
Busan, Korea
happyjdk@pusan.ac.kr

Yunju Baek
School of Computer Science and Engineering
Pusan National University
Busan, Korea
Yunju@pusan.ac.kr

***Abstract*— This paper proposes a machine learning-based approach to predict ship fuel consumption based on various operational factors, which is essential for shipping companies to optimize their operations, reduce costs, and minimize their environmental impact. The ship fuel consumption estimation research paper presents a novel approach to accurately predict fuel consumption in the maritime industry. By utilizing data obtained from actual sea trials, the study develops a practical and effective model tailored to individual vessels. The paper explores various methodologies, such as machine learning algorithms and traditional analytical models, and compares their performance using evaluation metrics. By comparing the performance of different machine learning models, the study demonstrates the superiority of the Extra Tree model, achieving MSE values of 0.7986 for HFO_PORT and 0.1264 for MDO+LSMGO on the dataset. These results can be utilized to design fuel-efficient routes and optimize ship operations, leading to significant cost savings and emissions reductions.**

Keywords—Ship Fuel Consumption, Machine Learning, Data Analysis, Maritime Industry

I. INTRODUCTION

The maritime industry is responsible for transporting approximately 80% of the world's goods, making it a critical component of global commerce However, shipping is also a significant source of greenhouse gas emissions, accounting for around 3% of global CO_2 emissions [1,2]. In addition, Ship fuel consumption accounts for about three-quarters of the cost used in operations. Accurate prediction of ship fuel consumption is crucial for optimizing ship operations and reducing costs and emissions. In recent years, Machine Learning (ML) has emerged as a powerful tool for analyzing large datasets and identifying patterns and relationships between different parameters [3].

The topic of environmental protection and energy management related to fuel is related to regulation. Recently, a system that enables the development of sensor technology and multi-source data collection has been installed, enabling fuel consumption analysis and prediction through data analysis [4,5]. The traditional approach for predicting fuel consumption has been based on analytical models, which require detailed knowledge of the ship's characteristics and operational conditions. However, these models often fail to capture the complex relationships between various factors that influence fuel consumption.

ML techniques can potentially provide a more accurate and reliable prediction by automatically learning from data. With the right training data and suitable model selection, ML-based prediction methods often yield more accurate predictions than data analysis. This is particularly the case with large or complex datasets [6]. And ML algorithms are designed to handle high-dimensional data and can automatically capture complex patterns and interactions between variables [7]. Statistical methods often require manual parameter tuning and can be labor-intensive to scale. However, ML models can be trained automatically, which makes them scalable for large dataset.

Therefore, we propose a ML-based approach for ship fuel consumption prediction. Using data extracted from actual sea trials, we identify the correlations between various variables and employ dimensionality reduction techniques to create a highly accurate model. Subsequently, we compare the fuel consumption prediction accuracy across multiple models. Proposed system is studies actual multidimensional test data of Shipbuilding and analyzes data related to fuel consumption. The rest of this paper is organized as follows: Section II provides an overview of dataset and describes the data preprocessing process and fuel co; Section IV presents the results and discussion; and Section V concludes the paper.

II. PROPOSED ESTIMATION METHOD

A. Dataset

The newly built vessel will be tested at sea for several days before being awarded. It is tested by driving at sea, such as speed using GPS, engine unmanned tests, steering tests, adjustment performance tests, and main engine start tests. Through sea trials, data such as operating time, fuel consumption, external environment, vessel condition, and engine performance is collected in monitoring and control systems such as Alarm Monitoring System (AMS), Integrated Automation System (IAR), and Voyage Data Recorder (VDR) during operation. To predict fuel consumption during voyage planning or while sailing, only input variables that can be adjusted on the actual vessel or that have an impact should be selected.

The dataset is measured data obtained through sea trials after the ship was built. The ship underwent several days of operational tests at sea to check whether the main engine and related systems meet various domestic and international regulations under actual operating conditions. This data comes from that process. It contains information about external environmental factors, fuel conditions, and the status of the ship, all stored on a time basis.

To establish criteria for this, a variable selection method was adopted through expert advice and statistical validation. The selection method is as follows.

Fig. 1. Example of 'WIND_RELATIVE DIRECTION' data.

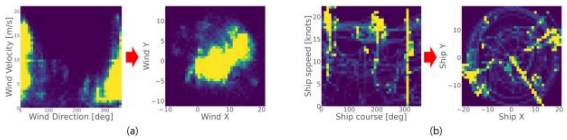


Fig. 2. Data transformation using radians vector based on the correlation between difference two data sets.

1) Based on the expertise of maritime navigation professionals, key variables related to fuel consumption were selected. Then, considering the variable data on ship energy efficiency measurements published by the United States Maritime Administration, designated variables were chosen.

2) To analyze fuel consumption, data preprocessing was performed by removing missing and outlier values and modifying data that did not conform to statistical techniques. Data selection was then performed through statistical validation, using Pearson correlation coefficients to confirm related indicators of fuel consumption data and scatter plots to confirm the relationship between each data.

3) The data includes operating time, information on various fuels, sea information as an external environmental factor, vessel condition and situation information, vessel speed, direction, current, trim, list, and engine status, as shown in Table 1. The data was collected during sea trials from '2021/06/22 08:13:22' to '2021/06/26 13:04:08' by the vessel's sensors at a frequency of 1Hz. The total number of data is 363,048 x 29 with 29 labeling categories.

B. *Preprosessing*

Missing values in data analysis can distort the characteristics of the actual data and reduce the accuracy of prediction results. Therefore, data is transformed to cleanse the values. To detect data anomalies, a time series plot was first created to visually inspect the data, and the mean, standard deviation, minimum, and maximum values of the entire dataset were examined to determine the presence of outliers.

In Fig. 1, an investigation was conducted on values that could not possibly exist. 'WIND_RELATIVE_DIRECTION' and 'SHIP_COURSE' are represented as values between 0 and 360, which indicate direction. Values that do not adhere to this range are considered outliers. Such data can degrade the performance of the model when used as input. Therefore, 'WIND_TRUE_SPEED' and 'SPEED_BY_GPS' were transformed using radian vectors.Additionally, 'WIND_TRUE_SPEED' and 'WIND_RELATIVE_DIRECTION' were converted into an easily interpretable form as radian vector values 'Wx' and 'Wy'. Similarly, 'SPEED_BY_GPS' and 'SHIP_COURSE' were also applied and transformed into 'Sx' and 'Sy', respectively. Through Fig. 2, we can see that the data has been transformed into a format that is easier for the model to understand. In (a), we can see that the data is skewed to one side. A transformation has been applied to evenly distribute the data between the two features. After the transformation, we can confirm from the graph that the distribution has become more uniform. This allows for more accurate model training. Similarly, in (b), the distribution of values between the two features has been changed through a transformation. This also has the same effect.

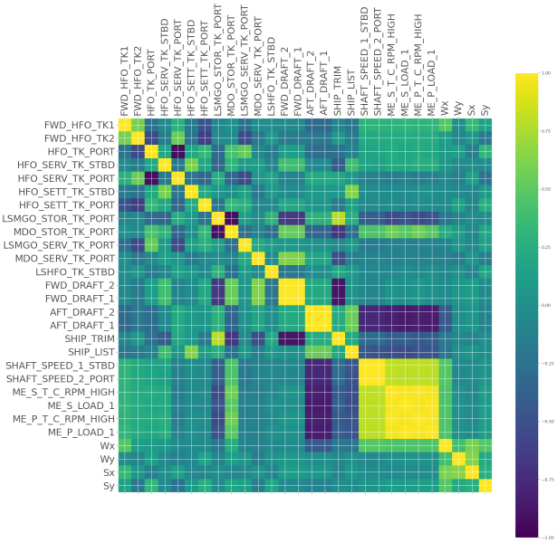


Fig. 3. Heatmap representing the correlation between all features.

C. *Feature Selection*

The goal is to derive the final data by selecting variables to be used in the prediction model through the analysis and investigation of fuel consumption based on the correlation with the actual ship fuel consumption values, using statistical graphs. A total of 3 stages of analysis were conducted, including Correlation heatmap, Histogram, and Skewness, through which the data was refined.

In The closer the correlation coefficient is to 1, the stronger the positive linear relationship between the two features, meaning they exhibit similar trends. Conversely, the closer it is to -1, the stronger the negative linear relationship, indicating they show opposing trends and that the two features cannot be combined. This was utilized to merge the features. In Fig. 3, it was observed that 'HFO_TK_PORT' and 'HFO_SERV_TK_PORT' have an inverse relationship, and 'LSMGO_STOR_TK_PORT' and 'MDO_STOR_TK_PORT' also exhibit an inverse relationship. Furthermore, 'FWD_DRAFT_1' and 'AFT_DRAFT_2' were found to be similar, and subsequent engine-related values showed strong correlation and similarity in the data.

Fig. 4 show that in fuel values such as 'FWD_HFO_TK1', 'FWD_HFO_TK2', 'HFO_SETT_TK_STBD', 'HFO_SERV_TK_STBD', 'MDO_SERV_TK_PORT', and 'LSHFO_TK_STBD', it was observed that they have very small fluctuations. Additionally, values such as 'DRAFT', 'SHART_SPEED', 'RPM_HIGH', and 'LOAD' appeared to have similar quantities and shapes. Therefore, they were replaced with their respective mean values ('mean_FWD_Draft', 'mean_AFT_Draft', 'mean_ShaftSpeed', 'mean_RPM_High', 'mean_Load'). Subsequently, through the analysis of skewness, the data with the highest skewness value was identified, and it was confirmed that there would be no issues when applying it to the model.

In order to extract the most similar data to fuel consumption among various fuels, a study on fuels was conducted. The fuels used in this experiment are HFO (Heavy Fuel Oil), MDO (Marine Diesel Oil), LSMGO (Low Sulfur Marine Gas Oil), and LSHFO (Low Sulfur HFO).

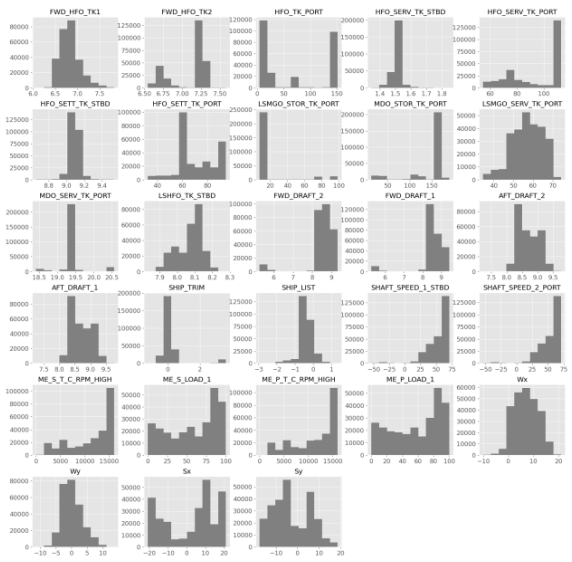


Fig. 4. Histogram representing the distribution of all features.

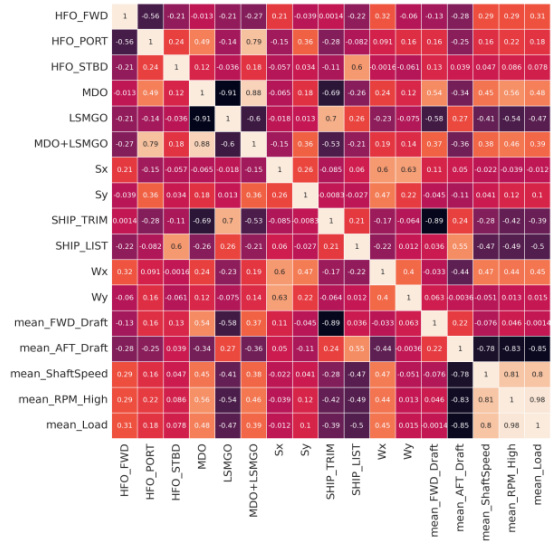


Fig. 5. Pearson correlation coefficient for each feature influenceed by external environment.

D. *Fuel Consumption Predict*

To predict consumption, suitable regression models among machine learning models were used. For the regression machine learning model data, the train:test dataset ratio was set to 8:2 in the final dataset, and the model was trained after normalizing the simple average data. Among fuel values, 'HFO_PORT' and 'MDO+LSMGO' had the most significant relation with the training label and provided the best results; thus, they were each trained and designated as the answer values in the regression model. Negative MSE was used as an evaluation metric to analyze various regression models. For the regression analysis machine learning model for consumption prediction, five models [Linear Regressor (LR), LASSO, ElasticNet (EN), K Neighbors Regressor (KNN), Classification and Regression Tree (CART)] were first examined. Subsequently, four ensemble models [Ada Boost Regressor (AB), GBM (Gradient Boosting Regressor), RF (Random Forest Regressor), ET (Extra Tree Regressor)] were used, and after parameter tuning, the results were analyzed. RF and ET models were performed with an estimator set to 10.

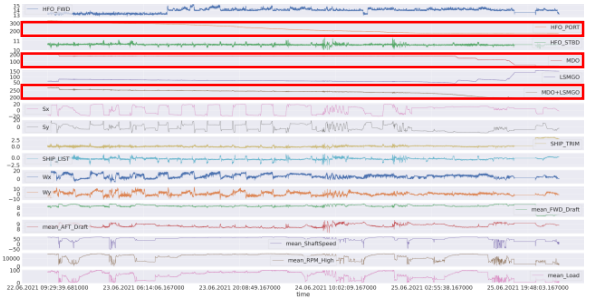


Fig. 6. Tme series graph representing the input features.

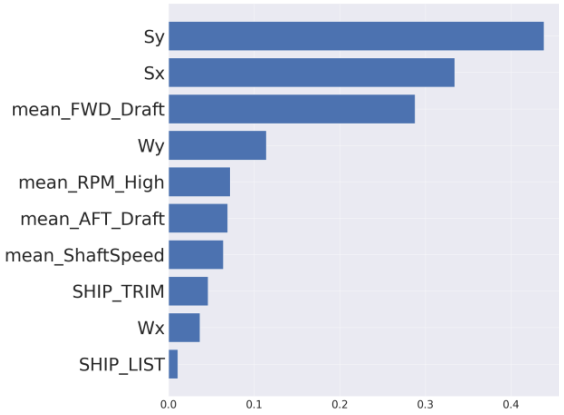


Fig. 7. Rank of the label determining the fuel consumption to be predicted.

III. Result and Discussion

A. *Results of fuel consumption analysis*

Based on the results of three analyses, the labels and features related to fuel consumption were found out. Looking at the Pearson correlation coefficient according to the fuel, external environment, and internal state of the ship, a larger absolute value implies a greater influence. Based on this, we tried to find the factor that has the most significant impact on the fuel and the fuel that is most influenced. First, by examining the Pearson correlation coefficient values in Fig. 5, it was determined that 'MDO' and 'LSMGO' values were the fuels most influenced by external environmental factors and the ship's status and engine values.

As confirmed in Fig. 5, the fuels most influenced are 'HFO_PORT', 'MDO', and 'MDO+LSMGO'. To verify this, we represented it in a graph over time. In Fig. 6, 'HFO_PORT', 'MDO', and 'MDO+LSMGO' values were found to be linear and decreasing, confirming that they are necessary labels for predicting fuel consumption values. And those features were blanked the red box.

The ranking is represented by the sum of the absolute values of the Pearson correlation coefficient. From the top, they can be seen as the factors that have the most influence. Fig. 7 shows the ranking of the most important features when determining fuel consumption for the prediction model. The vector values representing the ship's speed and direction, 'Sy' and 'Sx', are the most critical features, followed by the ship's status values 'mean_FWD_Draft' and 'SHIP_TRIM'. External factors 'Wy' and 'Wx' were found to have relatively lower impacts.

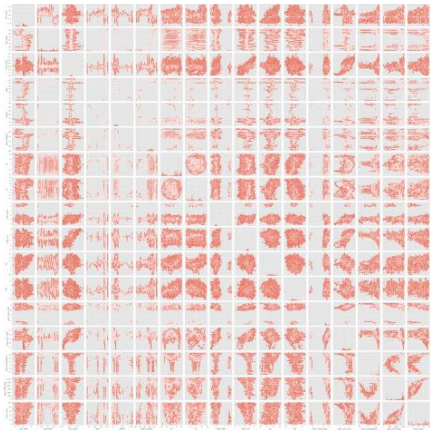


Fig. 8. Scatter plot for verifying linear and non-linear relationships.

Lastly, the linear and non-linear relationships were identified through Fig. 8. Irregularly distributed data means that there is no relationship between the fuel and the elements. To confirm the previous assumption, we drew a scatter plot and were able to support it. As a result, it was determined that the most suitable label for fuel consumption prediction is 'MDO+LSMGO'. We identified the factor that has the most significant impact on fuel, and based on this, we conducted an experiment to create a machine learning-based fuel consumption prediction model.

B. Results of fuel consumption prediction

After training the model selected for the comparison of measurement results, we calculated the error between each feature's predictions. We provided the external and ship's status as input and outputted the predicted fuel quantity, representing the error with the ground truth. Table 1 shows the MSE values for the different ML models tested. The Extra Tree model achieved the best performance, with the lowest MSE. This result indicates that the Extra Tree model was able to capture the complex relationships between the various factors and accurately predict fuel consumption.

TABLE I. RESULTS OF FUEL CONSUMPTION PREDICTION FOR EACH MODEL.

Algorithm	HFO_PORT MSE	MDO+LSMGO MSE
Linear	1864.6246	164.0108
LASSO	1905.5674	175.4442
ElasticNet	2169.2923	210.5297
K-Neighbor	16.2241	1.6343
Classification and Regrssion Tree	3.7491	0.3441
Ada Boosting	1720.2329	122.0007
Gradient Bossting	447.6449	36.7359
Random Forest	2.6006	0.2500
Extra Tree	0.7986	0.1264

IV. CONCLUSION

In this paper, We found out the features that most significantly impact fuel consumption by utilizing data generated during sea trials for testing constructed ships and proposed a machine learning-based approach to predict ship fuel consumption using various operational factors. Among various factors, the ones that have the most significant impact are 'Sx' and 'Sy', which are 'SPEED_BY_GPS' and 'SHIP_COURSE'. This result was derived through the Pearson correlation coefficient. We compared the performance of different ML models and demonstrated the effectiveness of the Extra Tree model, which achieved an MSE value of 0.7986, 0.1264 on our dataset. The most appropriate features for the final fuel consumption prediction are 'HFO_PORT' and 'MDO+LSMGO'. It also found that the factors that contribute the most to fuel consumption are the speed of the ship and the direction of the ship. Our findings can be used by shipping companies to optimize their operations for significant cost savings and reduced emissions. In conclusion, this research significantly advanced the understanding of ship fuel consumption by identifying key operational factors and their impact.

In future research, it would be worthwhile to conduct further analysis on outliers and additional data, and to incorporate deep learning models for performance comparison and evaluation. We plan to study methods for reducing the size of the machine learning and deep learning models that have been developed for implementation in the embedded systems on the ship. Lastly, through research on real-time inference models for consumption prediction, we will investigate a system that can use fuel more efficiently.

ACKNOWLEDGMENT

This work was supported by Korea Institute for Advancement of Technology(KIAT) grant funded by the Korea Government(MOTIE) (P0017006, The Competency Development Program for Industry Specialist at 2023).

REFERENCES

[1] UNCTAD, "Review of Maritime Transport 2021," United Nations Conference on Trade and Development, 2021.

[2] International Maritime Organization, "GHG emissions from ships," 2021. [Online]. Available: https://www.imo.org/en/OurWork/Environment/PollutionPrevention/AirPollution/Pages/GHG-Emissions-from-Ships.aspx. [Accessed: 28-Apr-2023].

[3] Q. Wang, X. Wang, Q. Zeng, and X. Zhu, "Support vector regression for prediction of container ship fuel consumption," Applied Sciences, vol. 7, no. 6, pp. 564, 2017.

[4] Y. Zhu, Y. Zuo, and T. Li, "Modeling of ship fuel consumption based on multisource and heterogeneous data: Case study of passenger ship," MDPI, vol. 9, no. 3, pp. 273, Mar. 2021. [Online]. Available: https://www.mdpi.com/2077-1312/9/3/273. [Accessed: 28-Apr-2023].

[5] N. Bialystocki and D. Konovessis, "On the estimation of ship's fuel consumption and speed curve: A statistical approach," in Journal of Ocean Engineering and Science, vol. 1, no. 2, pp. 157-166, 2016.

[6] S. Wang, B. Ji, J. Zhao, W. Liu, and T. Xu, "Predicting ship fuel consumption based on LASSO regression," in Transportation Research Part D: Transport and Environment, vol. 65, pp. 817-824, 2018.

[7] C. Ji, "Predicting fuel consumptions and exhaust gas emissions for LNG carriers via Machine Learning with hyperparameter optimization," in Proc. TOS21, 2021. [Online]. Available: https://onepetro.org/SNAMETOS/proceedings-abstract/TOS21/2-TOS21/463111. [Accessed: 28-Apr-2023].

Machine Learning for the Prediction of Water Quality for Smart Aquaculture

1st Dae-Won Kim
Digital Convergence Research Laboratory
Electronics and Telecommunications Research Institute
Yuseong-gu, Daejeon, 34129, KOREA
dwk@etri.re.kr

2nd Sumin Park
Digital Convergence Research Laboratory
Electronics and Telecommunications Research Institute
Yuseong-gu, Daejeon, 34129, KOREA
s2uuum@etri.re.kr

3rd Jae-Young Jung
Digital Convergence Research Laboratory
Electronics and Telecommunications Research Institute
Yuseong-gu, Daejeon, 34129, KOREA
jjy72@etri.re.kr

***Abstract*—Water quality in fish farms is crucial for the survival and growth of fish. This work presents the results of applying several machine learning methods to predict water quality in a fish farm. We used time series collected from sensors installed in an eel farm and applied six different machine learning methods (ARIMA, XGBoost, LightGBM, LSTM, SCINet, and DLinear) to predict water quality. As a result, the DLinear model showed the best prediction performance. We also investigated the feature importance estimated during the model training.**

***Index Terms*—machine learning, water quality prediction, time series**

I. Introduction

Smart aquaculture is an important technology that has the potential to make aquaculture more sustainable, efficient, and productive [1]. In a smart aquaculture system, monitoring water quality can help reduce pollution and waste, which also improves the quality of aquaculture products by ensuring that fish are reared in a healthy and stress-free environment [2]–[4]. To this end, the smart aquaculture system continuously monitors a number of water quality variables including dissolved oxygen, water temperature, acidity, oxygen in air, oxidation-reduction potential, and more.

In this manuscript, we present the results of predicting water quality using machine learning applied to time series collected from sensors installed in an eel farm. The time series and the pre-processing applied to the collected data are explained in the sections II and III respectively. Section IV presents the six machine learning methods that we applied in this work and their application results. A summary and future work are given in section V.

II. Data

We collected the time series using the AI[1] Hub website[2], which provides data for AI training from either national or international institutions and companies. The data collected includes values measured by seven sensors installed in an eel farm, each of which is dissolved oxygen (DO), water temperature, acidity (pH), CO_2, oxidation-reduction potential (ORP), O_2 in air, and light intensity. Table I shows some examples of the data collected. The data was collected from 15 July to 28 December 2021. The total number of rows is approximately 200,000 with a time interval of one minute.

[1]Artificial Intelligence
[2]https://www.aihub.or.kr

III. Data Preprocessing

As the performance of machine learning models depends on a training set, we first examined the data to build a robust training set. Fig. 1 shows an example of the collected time series. The top panel is the raw data and the bottom panel is the data after filling in the missing data point using the nearest neighbor interpolation method. If the duration of the missing data region is longer than one hour (e.g. the green and red boxes in the figure), we did not use it for model training.

We normalized the data by scaling each feature to be assigned to the range [0, 1]. Figure 2 shows the distribution of the data after normalization.

After preprocessing the data, we examined the Pearson correlation coefficients (r) between the seven features listed in Table I. The results are shown in Figure 3. The figure shows a weak correlation between DO and ORP ($r = 0.24$), which is a known characteristic of water in a tank. The correlation coefficient between DO and O_2 is -0.34 (i.e. anti-correlation), which is not an expected result because these two variables are generally correlated. This could be due to a malfunction in the sensor or an error in the way the data is being stored. Further investigation, however, is beyond the scope of this work as the eel farm where the data were measured is not accessible.

IV. Results of Machine Learning Application

To build a training set using the preprocessed data mentioned in the previous section, we applied a three-minute moving average to reduce data fluctuations that could be caused by various reasons, such as sensor failures. We then

TABLE I
EXAMPLES OF THE COLLECTED DATA.

Date	DO (ppm)	temperature ($^{\circ}C$)	pH	CO_2 (ppm)	ORP (mV)	O_2 in air (ppm)	light intensity (cd)
2021-07-15 19:03:00	8.8	30.3	5.28	160.0	14.6	5.4	4.0
2021-07-15 19:04:00	8.8	30.3	5.27	160.1	14.6	5.3	3.9
2021-07-15 19:05:00	8.8	30.4	5.27	160.3	14.5	5.3	4.0
...	...	...	...	...	...	...	
2021-12-28 17:55:00	5.5	28.7	4.96	231.8	31.9	5.5	4.0

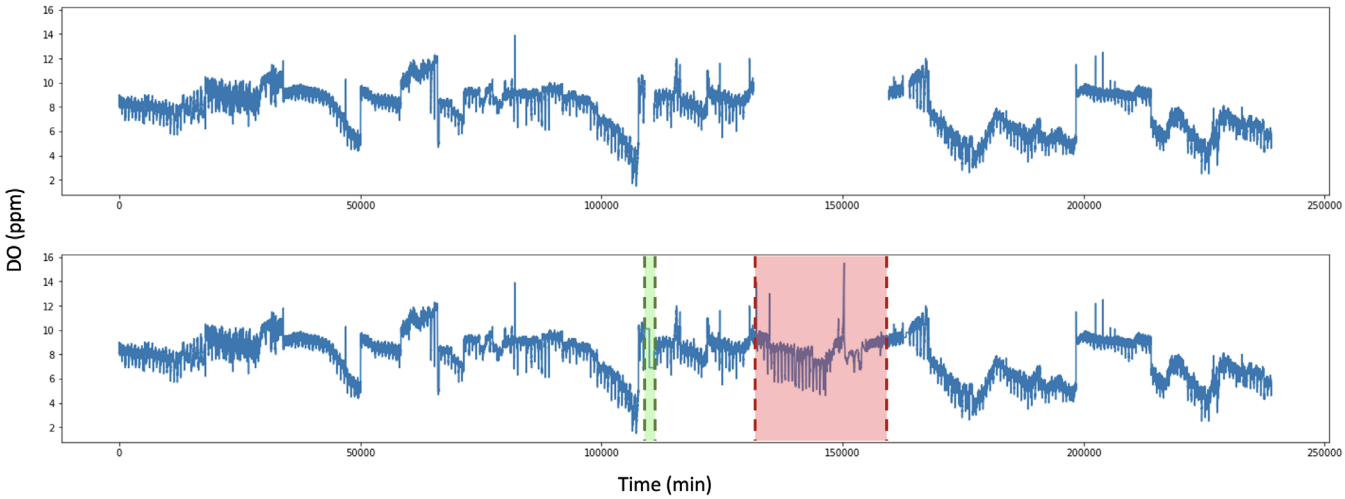


Fig. 1. Top: time series with missing data. Bottom: time series after filling in the missing data.

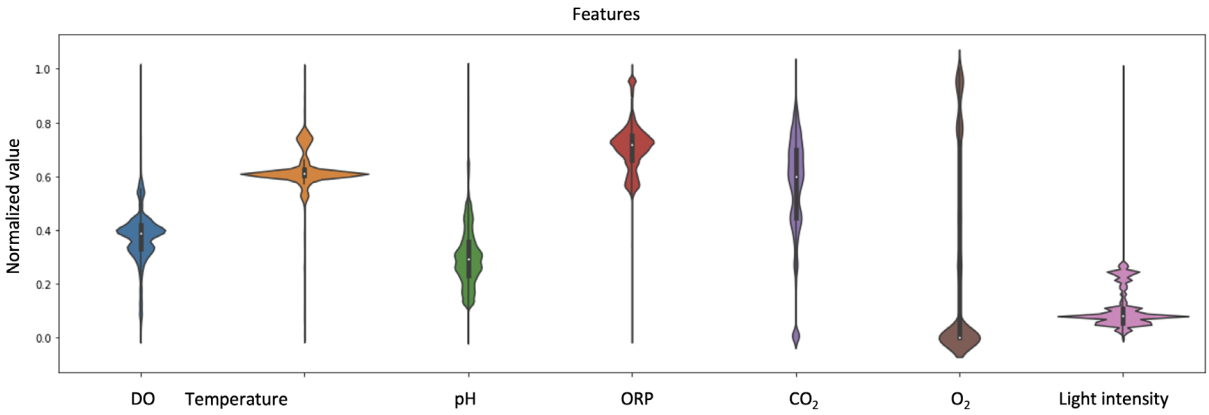


Fig. 2. Distribution of data after normalization.

split the data chronologically into ∼60%, ∼25%, and ∼15% subsets and used these as the training, validation, and test sets respectively.

We then trained six machine learning models using the training and validation sets: ARIMA [5], XGBoost [6], LightGBM [7], LSTM [8], SCINet [9], and DLinear [10]. We used six features (i.e. temperature, pH, CO_2, ORP, O_2 in air and light intensity) to predict DO. The duration of the input data was 60 minutes and the output was five minutes. We tested other combinations of time duration for input and output data (i.e. 60, 90, 120, and 150 minutes for input data and 5, 10, and 15 minutes for output data), and we found that a combination of 60 minutes and five minutes gave the best prediction results.

To evaluate trained models, we used mean squared error (MSE), mean absolute percentage error (MAPE), and root mean squared error (RMSE). Table II shows the MSE, MAPE, and RMSE measured using the test set. As the table shows, ARIMA has a much worse prediction performance than other methods. This may indicate that the time series is not stationary, or that the ARIMA model is not properly optimized.

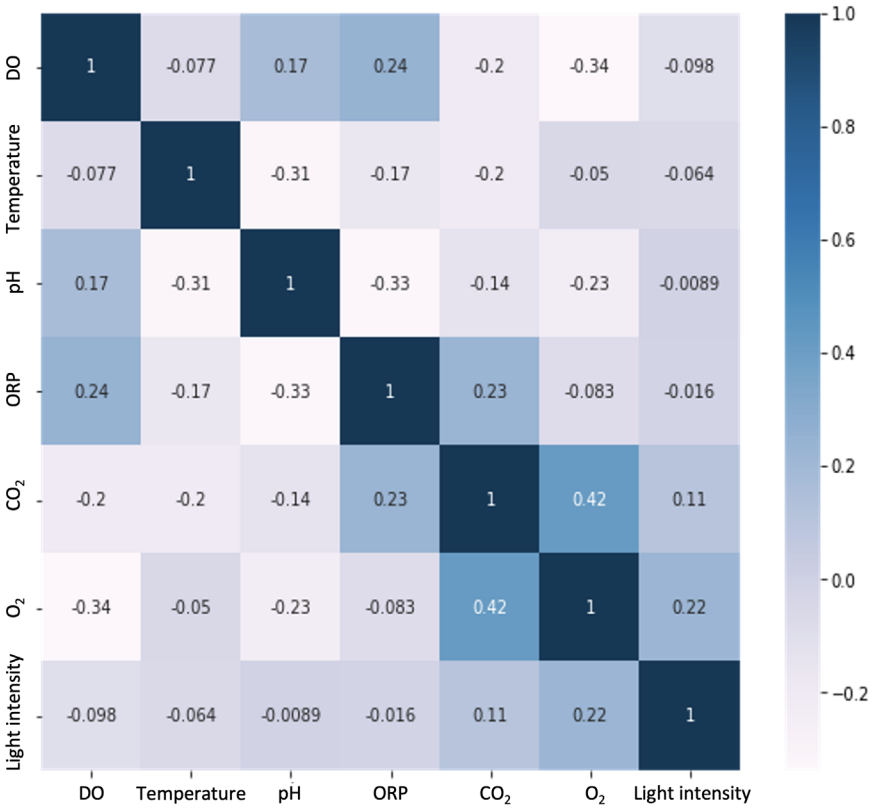


Fig. 3. Pearson correlation between the seven features.

DLinear which handles data with trend and seasonality patterns shows the best performance. This is an expected result, as DLinear is known to often outperform Transformer [11], one of the most widely used deep learning models for time series forecasting. LSTM performs slightly worse, but almost as well as DLinear.

In Fig. 4, we show the feature importance of the six input features estimated using XGBoost. Each column shows the importance of the feature minutes before the prediction time. The color bar shows a normalized importance score. Higher values indicate more important features. The following information can be inferred from the figure:

- Temperature 40 minutes prior has the strongest influence on the prediction of DO at the present time.
- Temperature, pH, and CO_2 about 55∼60 minutes before, temperature about one∼three minutes before, and O_2 about two∼seven minutes before show the moderate influence.
- ORP or light intensity does not show a strong influence on DO prediction.

V. Summary and Future Work

We have introduced water quality prediction using several machine learning methods and time series collected from sensors installed in an eel farm. The data provide seven features (i.e. DO, temperature, pH, CO_2, ORP, O_2 in air and light intensity). Using six features as input data (i.e. temperature, pH, CO_2, ORP, O_2 in air, and light intensity), we trained six machine learning models to predict DO. Among the trained models, we found that the DLinear model showed the best prediction performance. We also investigated the feature importance of the six input features and found that temperature, pH, CO_2, and O_2 affect the prediction results.

Based on the results mentioned in the previous sections, we will investigate other machine learning models to improve the prediction results. In addition to testing other models, extensive preprocessing will be considered, such as enhanced smoothing other than simple moving averages. We also plan to install additional sensors to measure water quality (e.g. turbidity). In this work, we have investigated the importance of six features, but we have not investigated why such an effect occurs or whether the result is physically meaningful. In future work, we will also investigate these in more detail.

The data used in this work was collected using the AI Hub website and is publicly available data, which is useful for preliminary investigations for data preprocessing and machine learning model selection. However, to apply machine learning to water quality prediction in real-world cases, we are building test beds for flatfish farming in several locations. The test beds are equipped with several sensors that measure water quality such as DO, pH, CO_2, and so on. Using the data collected from the test beds, we will develop a water quality prediction

TABLE II
PERFORMANCE OF THE SEVEN MACHINE LEARNING MODELS IN PREDICTING WATER QUALITY.

Model	MSE	MAPE	RMSE
ARIMA	6.936	2.205	0.404
XGBoost	0.119	0.166	0.345
LightGBM	0.113	0.149	0.337
LSTM	0.008	0.071	0.09
SCINet	0.104	0.154	0.323
DLinear	**0.007**	**0.059**	**0.081**

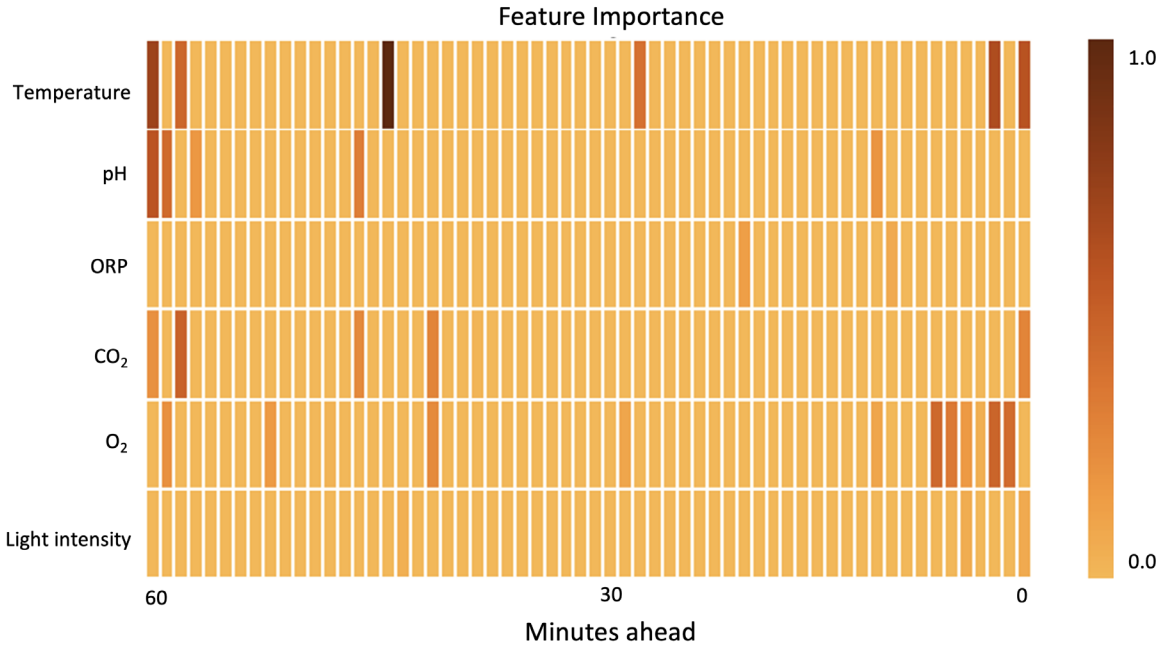


Fig. 4. Feature importance.

system based on machine learning and time series analysis.

VI. ACKNOWLEDGEMENT

This research was supported by Korea Institute of Marine Science & Technology Promotion(KIMST) funded by the Ministry of Oceans and Fisheries, Korea(20220596).

REFERENCES

[1] Vo, T.T.E.; Ko, H.; Huh, J.-H.; Kim, Y. Overview of Smart Aquaculture System: Focusing on Applications of Machine Learning and Computer Vision. Electronics 2021, 10, 2882. https://doi.org/10.3390/electronics10222882

[2] P Arulampalam, F M Yusoff, M Shariff, A T Law, P S Srinivasa Rao. 1998. Water quality and bacterial populations in a tropical marine cage culture farm. Aquaculture Research 29, 617-624. https://doi.org/10.1046/j.1365-2109.1998.00248.x.

[3] P Anusuya Devi, P Padmavathy, S Aanand and K Aruljothi. 2017. Review on water quality parameters in freshwater cage fish culture. International Journal of Applied Research 3(5), 114-120.

[4] Hu Z, Li R, Xia X, Yu C, Fan X, Zhao Y. A method overview in smart aquaculture. Environ Monit Assess. 2020 Jul 8;192(8):493. doi:10.1007/s10661-020-08409-9. PMID: 32642861.

[5] Harvey, A.C. (1990). ARIMA Models. In: Eatwell, J., Milgate, M., Newman, P. (eds) Time Series and Statistics. The New Palgrave. Palgrave Macmillan, London. https://doi.org/10.1007/978-1-349-20865-4_2

[6] Tianqi Chen and Carlos Guestrin. XGBoost: A Scalable Tree Boosting System. SIGKDD Conference on Knowledge Discovery and Data Mining, 2016 https://arxiv.org/abs/1603.02754

[7] Ke, Guolin and Meng, Qi and Finley, Thomas and Wang, Taifeng and Chen, Wei and Ma, Weidong and Ye, Qiwei and Liu, Tie-Yan. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. Advances in Neural Information Processing Systems, 2017

[8] Sepp Hochreiter and Jurgen Schmidhuber. Long Short-Term Memory. In Neural Computation, volume 9, 1997. http://dx.doi.org/10.1162/neco.1997.9.8.1735

[9] Minhao Liu, Ailing Zeng, Muxi Chen, Zhijian Xu, Qiuxia Lai, Lingna Ma, Qiang Xu. SCINet: Time Series Modeling and Forecasting with Sample Convolution and Interaction. Neural Information Processing Systems, 2022. https://arxiv.org/abs/2106.09305

[10] Ailing Zeng, Muxi Chen, Lei Zhang, Qiang Xu. Are Transformers Effective for Time Series Forecasting? Association for the Advancement of Artificial Intelligence, 2023 https://arxiv.org/abs//2205.13504

[11] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, Illia Polosukhin., Attention Is All You Need. Advances in Neural Information Processing Systems, 2017. https://arxiv.org/abs/1706.03762

Water Level Prediction at Han River Bridge using wavelet-GRU Model

Mark Joseph Calubad
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
mjcalubad@donga.ac.kr

Handikajati Kusuma Marjadi
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
handikajati.kusuma@gmail.com

Hubdar Hussain
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
hubdarkhan@gmail.com

So Young Eom
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
2273454@donga.ac.kr

Eun Seok Jang
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
dnstjr5627@naver.com

Junbong Jang
ICT Integrated Safe Ocean Smart Cities Engineering
Dong-A University
Busan, South Korea
jjang@dau.ac.kr

***Abstract*—Accurate prediction of water level fluctuations in rivers is essential for effective planning and mitigation of potential flood-related risks. In this study, four different Recurrent neural networks (RNN) models were developed to predict the Han River's water level, based on collected data from 2020 to 2022 which was used to train and test the proposed models. Two input combinations were investigated to analyze and improve the accuracy of the models. For each model + input combination, the predicted water level values after t hours (t = 0, 1, 2) were correlated to the corresponding measured water level values and used coefficient of determination, R^2, to evaluate accuracy of prediction of each combination of model and input. The results showed that at t = 0 and t = 1, wavelet-GRU performed best with R^2 values of 0.99091 and 0.98743, respectively, while wavelet-LSTM performed best at t = 2 with a value of 0.98867.**

Keywords— Recurrent neural networks (RNN), wavelet-GRU, wavelet-LSTM, Water level monitoring

I. INTRODUCTION

Floods result in substantial social and economic damage and the loss of human lives, making them a major harm factor. Densely populated areas situated adjacent to and downstream of major rivers face a potential threat due to their proximity to flooding hazards. For planning and maintaining water resource systems as well as lowering the danger of floods, accurate river water level estimating is crucial [1]. In addition, the nature of the river flow process is considered nonlinear and affected by a variety of elements, including the mantle that covers the surface of the river basin, the way that rain falls, the topography of the riverbed, and climatic conditions. To lessen or minimize the effects of floods, a large number of predictive strategies that require a tremendous quantity of data for forecasting have been proposed [2].

Currently, there are two common methods for predicting the river water level. The first approach uses mathematical simulations of the water's flow's hydrodynamic process. The mathematical models utilized in the first approach are based on hydraulics and hydrology principles, making it a widely-used technique in the past [3]. These models typically demand a lot of input data, like rainfall predictions and topography data, which might not always be available or might be challenging to get. The parameters may vary per region and executing the model is time consuming which makes it hard to use for prediction.

The second methodology for predicting river water level involves the data-driven models and the statistical relationship between input and output data forms the basis of these approaches [1]. The Artificial Neural Network (ANN) model is one of the popular data-driven models. This statistical procedure does not require assumptions regarding the amount of data needed for the model and the model's capacity to handle both linear and nonlinear systems. Due to the simplicity of its model structure, ANN model has become widely used in many fields of research and engineering as computer science has advanced over the past two decades. To reduce mistakes and increase the prediction accuracy of neural network models, researchers have used a variety of strategies, such as employing them alone or in conjunction with process-based models. These characteristics make the data-driven approaches possible alternatives to conventional hydrological water level-predicting strategies.

In recent years, there has been significant research on deep neural network (DNN) data-based intelligent models with multiple hidden layers, addressing some of the limitations of traditional artificial neural network (ANN) models. However, the temporal aspect of modelling river water level is not considered in ANN and DNN. To tackle the issue of including temporal effect in modelling, long short-term memory (LSTM) and gated recurrent unit (GRU) models, which are based on recurrent neural networks (RNN), have emerged as highly effective in processing time series data [4]. LSTM and GRU both have their own advantages and disadvantages. LSTM is great with larger dataset but is slower compared to GRU. GRU, on the other hand, can be as accurate as LSTM with a faster processing time and less memory consumption. However, GRU struggles with larger datasets [5].

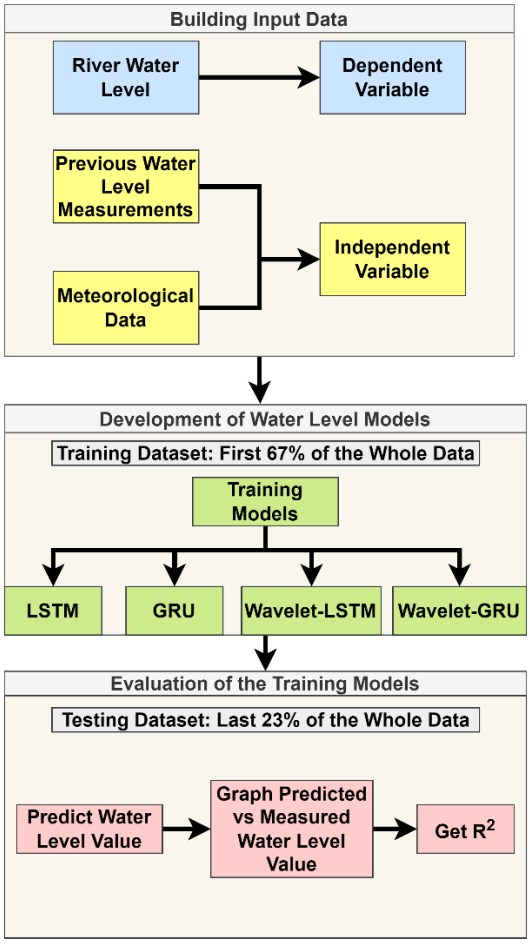


Fig. 1. General Methodology

Wavelet decomposition is a powerful method for processing and analyzing data, particularly in hydrologic time-series analysis [6]. It enables the examination of multi-temporal scale structures, identification of deterministic components, noise removal, and complexity quantification. Wavelet decomposition technique can effectively separate the low-frequency trend information from time series based on the fluctuation frequency, enabling the subsequent analysis of high-frequency fluctuations over time. Using wavelet decomposition has shown its effectiveness in improving both LSTM and GRU models for different applications [7]. Hence, we employed four models LSTM, GRU, wavelet-GRU, and wavelet-LSTM, which is the widely used wavelet decomposition-RNN modelling ensemble technique for water level modelling [8], and a key objective of this study is to integrate the advantageous features of wavelet decomposition along with GRU, which has been used in other applications involving time series data [9], as the modelling technique and compare its capability in modelling water level against wavelet-LSTM technique.

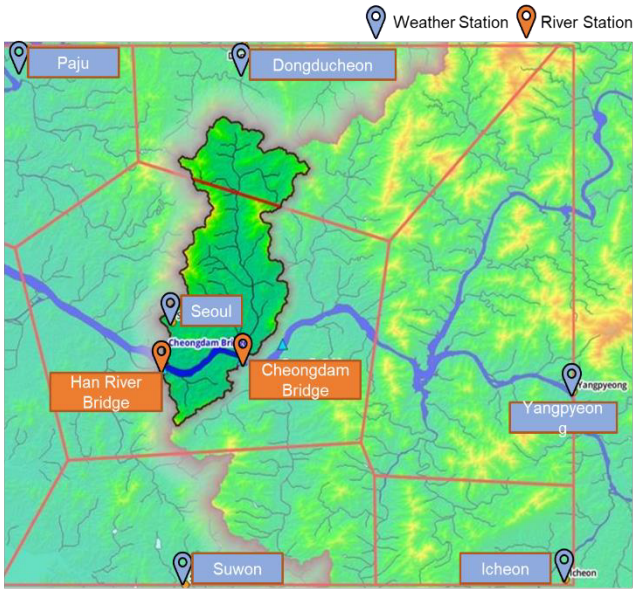


Fig. 2. Study Area

II. Methodology

The adopted methodology consists of three main tasks: building input data, developing prediction models, and evaluating performance as shown in Fig. 1. For the required data, the Han River Bridge located on Han River, Seoul was selected as the focus of our study. Initially, Hydrological analysis was done by using QGIS software to get the catchment area and Thiessen polygon as shown in Fig. 2. The datasets of Han River water level and weather parameters of three closet stations (Seoul, Suwon, and Dongducheon) were obtained from the websites of Han River Flood Control Office (website: https://www.hrfco.go.kr/eng/eng_main.do) for water level measurements and the Korea Meteorological Administration (website: https://data.kma.go.kr/resources/html/en/aowdp.html) for weather data, respectively. The water level of the Han River bridge was used as a dependent variable and the Weather parameters including (temperature, relative humidity, pressure, and precipitation) from relevant weather stations and the previous water level measurements were set up as the independent variables. The dataset was divided into two sets: a training dataset and a testing dataset.

Around 67% of the primary dataset's data samples were used to train the models. The training dataset then serves as the input in the following modelling techniques: LSTM, GRU, wavelet-LSTM, wavelet-GRU. The best setting for each model was then

National Research Foundation (NRF) funded by the Korea Ministry of Education (No.2016R1A6A1A03012812).

used to predict the water level values at certain time periods. In the last step, the coefficients of determination (R^2) were calculated by relating the measured water level values to the predicted water level values based on 23% of the whole data as testing dataset for the evaluation of developed models.

III. Results

The viability of the different data-driven methods; Long-short-term-memory (LSTM), gated recurrent unit (GRU), wavelet-LSTM, and wavelet-GRU was investigated for water level prediction t ($t = 0,1,2$) hours after. Table 1 sums up the test performances of the LSTM, GRU, wavelet-LSTM, and wavelet-GRU in predicting water level at different times of prediction. As it is clear from Table 1, R^2 decreases in general as any model tries to predict water level measurement at a later time period. The wavelet-GRU model with water level only as an input had the highest R^2 value of 0.99091 when trying to predict values at $t = 0$. The other model + input combination also showed great fit at the same time of prediction. This is to be expected since the water level value being predicted is near the input values in terms of temporal position.

TABLE 1 Fit of Predicted vs Measured Water Level Values at Different Times of Prediction

Methodology	**Input**	**R^2 of predicted vs measured values after time t in hours**		
		$t = 0$	$t = 1$	$t = 2$
LSTM	Water Level Only	0.98651	0.94828	0.91312
	Water Level + Weather Data	0.98661	0.94528	0.91764
wavelet-LSTM	Water Level Only	0.99084	0.95763	0.92359
	Water Level + Weather Data	0.98886	0.95973	**0.98867**
GRU	Water Level Only	0.98667	0.94612	0.91392
	Water Level + Weather Data	0.98708	0.94683	0.91649
wavelet-GRU	Water Level Only	**0.99091**	0.96110	0.92415
	Water Level + Weather Data	0.98806	**0.98743**	0.98748

At $t = 1$, the model + input combination that gave the best fit of predicted versus measured values is wavelet-GRU with water level and weather data as input. The said model + input combination resulted to an R^2 value of 0.98743. At this point, the effect of adding weather data showed a gradual increase in the fit of predicted values. The effect of decomposing the input values also showed slight increase in effect when predicting water level values as shown by comparing the results of LSTM and GRU models with wavelet-LSTM and wavelet-GRU model counterparts.

More evidence of the effect of wavelet decomposition and adding weather data in the input is shown in the results at $t = 2$. The combination of wavelet-LSTM as modelling technique with water level and weather data as input had predicted values with an R^2 value of 0.98867. Predicting water level values after 2 hours generally showed a good fit. When using only wavelet-LSTM or wavelet-GRU as modelling techniques or by adding weather parameters in the input data the increase in the fit of predicted values is only minimal. However, when performing wavelet-LST or wavelet-GRU and adding weather parameters as input data in predicting the water level measurements as seen from the jump. All these observations are shown in figure 3.

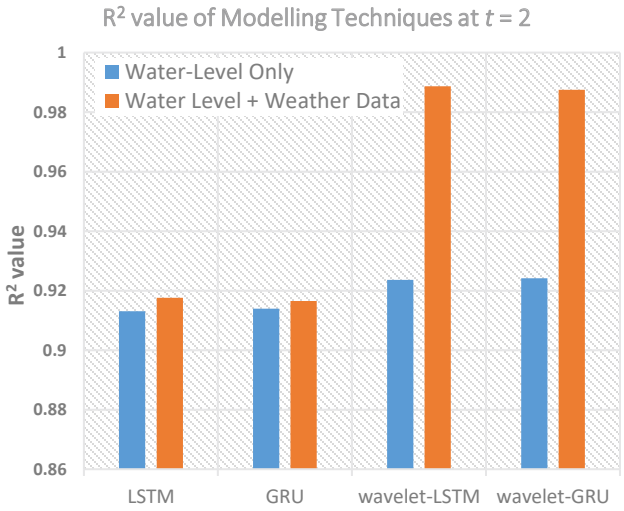


Fig. 3. R^2 values of Modelling Techniques at t = 2 hours

It should be noted that one of the reasons for modelling with wavelet decomposition to have a greater capacity to predict values compared to the general LSTM or GRU modelling is the capacity to model the general trend and the noise of the time series data separately. Also having different input parameters in the time series data, such as weather data which also affects the water level, that have their own general trend and noise when decomposed also helps in finding more correlation which helps in making a more accurate predicted value.

One trend that was observed when comparing the predicted values with the corresponding measured values is that the models tend to underpredict as shown in the fitted graph and histogram shown Figures 4 and 5, respectively. One of the

reasons can be the limited hyperparameters that were considered for tuning. Another reason might be the effect of seasonality in water level measurements.

Wavelet-GRU modelling compared to the widely used wavelet-LSTM modelling showed greater or comparable results in general when predicting water level measurements at different time periods. This goes to show the capability of wavelet-GRU modelling technique in predicting water level measurements at the Han River bridge monitoring station.

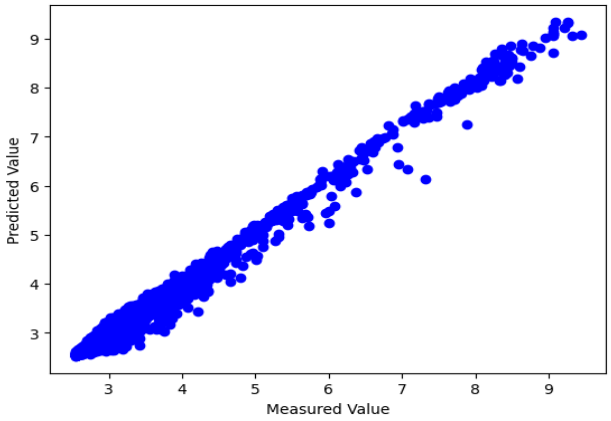


Fig. 4. Predicted vs Measured Water Level Value of wavelet-LSTM model using Water Level and Weather Data Input

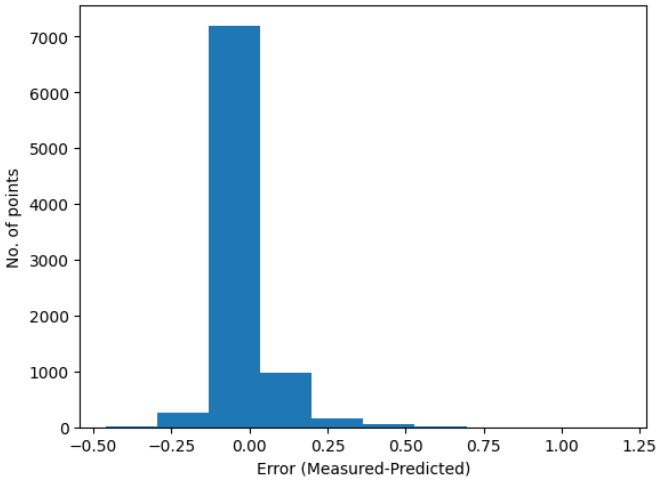


Fig. 5. Histogram of wavelet-LSTM model using Water Level and Weather Data Input Error

IV. Conclusions

This study presents four recurrent neural networks long short-term memory (LSTM), gated recurrent unit (GRU), Wavelet-LSTM, and Wavelet-GRU models to predict the Han River water level at Han River bridge, Seoul.

All models were explored for two input combination using only water level and water level with weather parameters. All models predicted the water level with satisfactory accuracy and achieved R^2 values greater than 0.91 which shows good fit of the predicted values when compared to the measured water level values. The wavelet-GRU modelling provided the best results with highest value of coefficient of determination of R^2 = 0.9909080779 using only water level input parameter at predicting water level value at $t = 0$.

The trends and observations of the predicted water level values correlated with the measured water level values were also presented and discussed. Wavelet-GRU modelling showed promise and consistency in predicting water levels at different time periods. Possible future studies that can be explored are the following: the effectivity of the wavelet-GRU in predicting water level at a further time, checking the effect of seasonality in the prediction of water level, consider other hyperparameters when building the model to mitigate underprediction, and checking the validity of wavelet-GRU in predicting water level values at different river monitoring stations.

Acknowledgment

This work was supported by the Basic Science Research Program through the National Research Foundation (NRF) funded by the Korea Ministry of Education (No. 2016R1A6A1A03012812).

References

[1] Mosavi, A., Ozturk, P., & Chau, K. (2018). Flood prediction using machine learning models: Literature review. Water, 10(11), 1536. https://doi.org/10.3390/w10111536

[2] Olivia Muslim, T., Najah Ahmed, A., Malek, M. A., El-Shafie, A., & EL-Shafie, A. (2018). Investigating the impact of wind on sea level rise using multilayer perceptron neural network (MLP-NN) at coastal area, Sabah. International Journal of Civil Engineering and Technology, 9(12), 646–656.

[3] Ehteram, M., Othman, F. B., Yaseen, Z. M., Afan, H. A., Allawi, M. F., Malek, M. B. A., Ahmed, A. N., Shahid, S., Singh, V. P., & El-Shafie, A. (2018). Improving the Muskingum flood routing method using a hybrid of particle swarm optimization and bat algorithm. Water (Switzerland), 10(6), 1–21. https://doi.org/10.3390/w10060807

[4] Bowes, B.D., Sadler, J.M., Morsy, M.M., Behl, M., Goodall, J.L., 2019. Forecasting groundwater table in a flood prone coastal city with long short-term memory and recurrent neural networks. Water 11. https://doi.org/10.3390/w11051098

[5] Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling (arXiv:1412.3555). arXiv. http://arxiv.org/abs/1412.3555

[6] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. Neural computation, 31(7), 1235-1270.

[7] J. Tzeng, Y. -R. Lai, M. -L. Lin, Y. -H. Lin, & Y. -C. Shih. (2020). Improve the LSTM and GRU model for small training data by wavelet transformation. 2020 International Joint Conference on Neural Networks (IJCNN), 1–6. https://doi.org/10.1109/IJCNN48605.2020.9206840

[8] Liang, Z., Liu, Y., Hu, H., Li, H., Ma, Y., & Khan, M. Y. A. (2021). Combined Wavelet Transform With Long Short-Term Memory Neural Network for Water Table Depth Prediction in Baoding City, North China Plain. *Frontiers in Environmental Science*, *9*, 780434. https://doi.org/10.3389/fenvs.2021.780434

[9] Biazon, V., & Bianchi, R. (2020). Gated Recurrent Unit Networks and Discrete Wavelet Transforms Applied to Forecasting and Trading in the Stock Market. Anais Do Encontro Nacional de Inteligência Artificial e Computacional (ENIAC 2020), 650–661. https://doi.org/10.5753/eniac.2020.12167

Implementation of Internet of Things for Predictive Electromechanical Valve Monitoring System to Enhanced Real Time Remote Observation

Z Mokhtar
School of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
zulkarnain85@graduate.utm.my

S M Idrus
School of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
sevia@utm.my

S Mad Saad
School of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
shaharil@utm.my

S I Azmi
School of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
asrul@utm.my

M Z Mohamed
Instrument & Control Section
Engineering Department
Sungai Udang, Melaka, Malaysia
mohdzurix@petronas.com

M R Salim
School of Electrical Engineering
Universiti Teknologi Malaysia
Skudai, Johor, Malaysia
mrashidi@utm.my

***Abstract*—Electromechanical Valve has been used wherever it is required for fluid flow and has to be controlled automatically. They are being utilised more and more in a wide range of plants and machinery. A valve may be chosen to perfectly suit the application at hand thanks to the range of possible designs. With respect to this advancement of technology regarding electromechanical valves, this paper will portray the most recent commonly produced solutions and technologies on electromechanical valve monitoring systems to determine the conditions of the electromechanical valve that can avoid any coil failures, identify any degradation or issues with the valve and its parts and fix them before the valve fails. Electromechanical valve monitoring is crucial to ensuring that the machinery is operating as efficiently as possible. This study reviews and presents proposed experimental works of electromechanical valve monitoring systems. It is widely used in many major and high revenue industries, from agriculture to the dynamic & technology advancement automotive manufacturing, and it plays a crucial role in keeping businesses running smoothly. An effective EMV monitoring solution is intended to enable large plantation sites, real-time analysis, and the provision of online monitoring systems based on IOT deployment, especially for the strategic area at the oil and gas industry**

Keywords—Electromechanical valve, Electrical coil, Monitoring system, , Rotork

I. INTRODUCTION

This Electromechanical valves are electromechanical devices that are controlled by electricity and are used to regulate the flow or direction of liquids or gas. Electrical energy is converted into mechanical motion or mechanical energy via an electromechanical valve [9]. An electromechanical valve is made up of two parts: a solenoid (a copper-insulated coil) and a valve body. The plunger (armature or core), seal, spring, orifice, shade ring, intake port, and outlet port are the other important components of the electromechanical valve that work together to function as a whole. In engineering practice, an electromechanical valve is an electric component that uses electric current to produce the strength of the magnetic field that can regulate fluid. These components are used to turn off, unload, dose, distribute, or mix fluids. This is performed by delivering an electric current via a coiled wire, creating a magnetic field, and therefore altering the position of a plunger. The location of the plunger governs the passage of the working fluid (s) [1] EMVs are essential parts among many technologies. Their appeal stems mostly from their basic and durable design, as well as their low cost. Electromechanical valves are utilized in a variety of products, including vehicles, household systems, and industrial machines. A liquid electromechanical valve regulates the flow of a liquid medium, whereas a pneumatic electromechanical valve controls the movement of air in a circuit. Electromechanical valves are frequently used for a purpose.

II. WORKING PRINCIPAL

In simpler terms, an electromechanical device regulate the flow of fluid by changing the position of the valve when an electric current is passed through a solenoid coil. A magnetic field is created by passing an electric current through the coil and pulls the orifice upward, opening the plunger. This is the fundamental principle that allows solenoid valves to open and close [15].

To understand the technical flow of the electromechanical valve model more, Figure 2 depicts simplified schematic views of the valve's closed and open positions. The valve is kept closed in the absence of solenoid excitation by the return spring (Fig. 1(a)). When a DC voltage is applied to the solenoid, the resulting magnetic force displaces the moving part against the return spring, causing the valve to open (Fig. 1(b)) and a flow cross section to develop through an orifice.

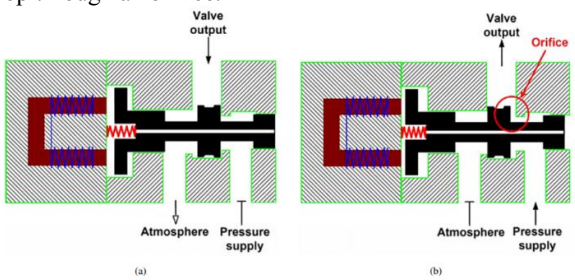


Figure 1. (a) & (b) [15] : Working Principle of Electromechanical Valve

III. Results and Discussion

Experimental works for electromechanical valve parameter characterizations can provide invaluable insights into the optimal operation of components. This can include detailed cross-validation studies to map out behaviors that are often invisible in ordinary tests or simulations. Taking the time to research these parameters and to lay down the boundaries around them can help engineers make more reliable decisions about how their electromechanical valves will react in real application, allowing them to maximize performance and reliability. Knowing these parameters can also allow for cost reduction efforts when developing systems containing electromechanical valve components, as well as improved maintenance practices due to a better understanding of expected performance.

A. Measured Performance Criteria

The coil resistance can play a significant role in understanding the aging condition of electromagnetic coil or coil failure in Electromechanical Valve. When the Electromechanical Valves do fail, a significant change in the measured current and resistance is expected. From the data findings, a suitable threshold can be established as follow which can used to predict the coil failure of Electromechanical Valve:-

- Control Condition: The current and resistance of the coil was observed to be the baseline.
- Shorting condition: The current value is increased while the resistance is decreased.
- Overheating condition: The current value is decreased while the resistance is increased.

B. Control Condition

The In the normal control condition, the Electromechanical Valve (EMV) coil resistance and temperature was observed to be the baseline of the electromechanical valve as per figure below:-

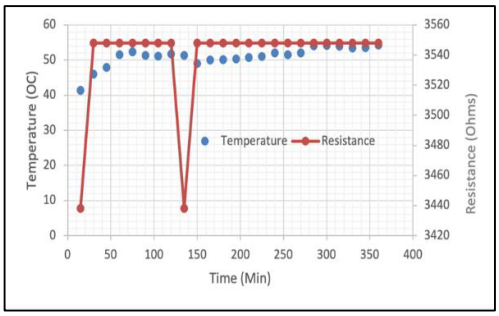


Figure 2. EMV Control Condition Parameter Behavior

C. Shorting Condition

The Electromechanical Valve (EMV) current value increased from its control condition 0.031 A to 0.055 A when there was water in the SOV. The changes of resistance value is vice versa as per table below:-

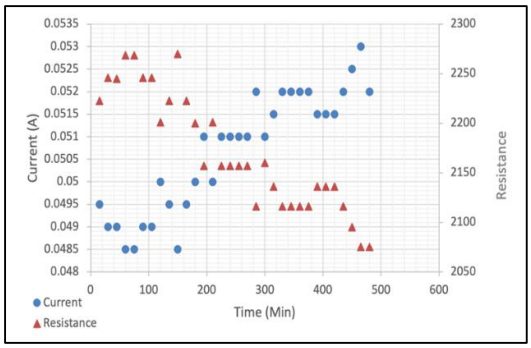


Figure 3. EMV Shorting Condition Parameter Behavior

D. Overheating Condition

The current value was decreased from 0.031A (control condition) to 0.029A when the coil at high surrounding ambient temperature. The changes of resistance value is vice versa as per table below:-

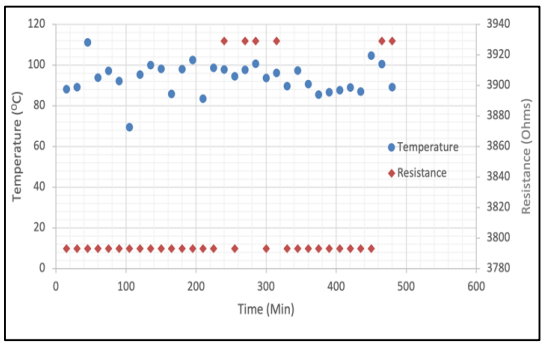


Figure 4. EMV Overheating Condition Parameter Behavior

E. Predictive Electromechanical Valve (EMV) Remote Monitoring System

In depth research of the EMV monitoring system shows that it has been implemented in many applications with a functional method. Further approach in the EMV monitoring system relies on the working principles and conditions of the EMV. Hence, it is significant to understand the type of monitoring system used to monitor the device. Based on the research, most monitoring solutions are using a traditional method to determine the condition and behavior of the EMV.

The proposed predictive EMV Remote Monitoring System configuration as per block diagram below for holistic illustration: -

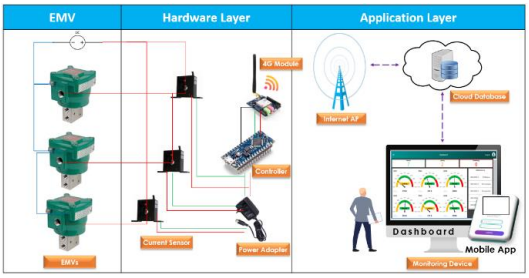


Figure 5. Proposed Configuration for Predictive EMV Remote Monitoring System

A. *Predictive Electromechanical Valve (EMV) Remote Monitoring System*

IoT platform for predictive monitoring of electromechanical valve offers convenience for organizations and businesses. This technology platform delivers dashboard and app-based notifications that allow users to remotely access their devices no matter the location. Users receive real-time updates on their valves' performance, helping them anticipate possible malfunctions before they occur. The dashboard also contains customizable analytics on the temperature, resistance, current, and healthiness so industry can obtain additional insights into their systems. Predictive analytics from IoT technology can improve productivity and reduce operational costs by preventing unexpected outages caused by unwanted malfunctions. Figure xx shown below are the proposed Predictive Monitoring for EMV.

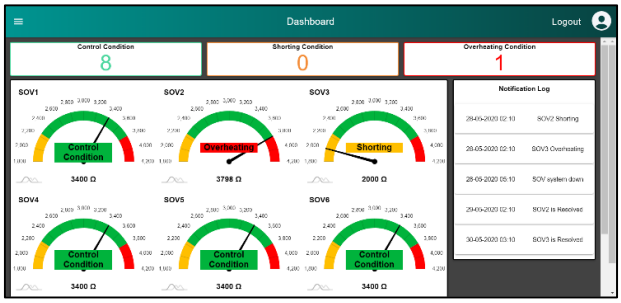


Figure 6. Proposed IoT Platform for Predictive Monitoring of EMV

V. CONCLUSION

The reason behind this paper is to emphasize the methods of monitoring the conditions of an Electromechanical Valve (EMV). Researchers that are engaged in electromechanical valve monitoring systems may find this review paper to be very beneficial. Reliability Testing Procedures, DCS/Safety SVM, Fault Detection Method, Diagnostic System, and Dynamic Model are a few techniques for keeping an eye on the electromechanical valve's condition.

In conclusion, electromechanical valves are an essential component for many industries. With the ability to precisely control flow of a wide variety of fluids, electromechanical valves are versatile and can provide a great deal of value creation across many applications. Beyond that, electromechanical valves offer safety advantages as they can be programmed to operate in predictive manners, perhaps taking over when manual operation would pose too much of a risk. The reliability and accuracy they bring forth makes electromechanical valves a long-term benefit to industry worldwide.

ACKNOWLEDGMENT

The This work was supported by Universiti Teknologi Malaysia (UTM) Institutional grant vote 08G49 and UTM-Industry Research Contract 4C489 funded by iSmartUrus Sdn Bhd. We wish to acknowledge the support and collaborative work received from Instrument & Control Section, Engineering Department, MRCSB, PETRONAS.

REFERENCES

[1] N. J. Jameson, M. H. Azarian, and M. Pecht, "Fault diagnostic opportunities for solenoid operated valves using physics-of-failure analysis," in 2014 International Conference on Prognostics and Health Management, Cheney, WA, USA, Jun. 2014, pp. 1–6. doi: 10.1109/ICPHM.2014.7036385.

[2] X. Qian and C. Zhan, "Laboratory Reliability Testing Design of Solenoid Operated Directional-Control Valves for Industrial Application," p. 6.

[3] S.-H. Jo, B. Seo, H. Oh, B. D. Youn, and D. Lee, "Model-Based Fault Detection Method for Coil Burnout in Solenoid Valves Subjected to Dynamic Thermal Loading," IEEE Access, vol. 8, pp. 70387–70400, 2020, doi: 10.1109/ACCESS.2020.2986537.

[4] L. Yan, Lu Zhang, Z. Jiao, and C. Gerada, "Design and modeling of a rectangular direct-acting solenoid valve," in 2016 IEEE Chinese Guidance, Navigation and Control Conference (CGNCC), Nanjing, China, Aug. 2016, pp. 2336–2340. doi: 10.1109/CGNCC.2016.7829156.

[5] D. Dong and X. Li, "Simulation and experimental research on the response of a novel high-pressure pneumatic pilot-operated solenoid valve," p. 5.

[6] S. Li, P. Guo, W. Jiang, H. Ding, and D. Yu, "Research on response characteristics and parameters optimization of high-speed solenoid valve," in 2015 34th Chinese Control Conference (CCC), Hangzhou, China, Jul. 2015, pp. 2327–2332. doi: 10.1109/ChiCC.2015.7259997.

[7] I.-Y. Lee, "Switching Response Improvement of a High Speed On/Off Solenoid Valve by Using a 3 Power Source Type Valve Driving Circuit," in 2006 IEEE International Conference on Industrial Technology, Mumbai, India, 2006, pp. 1823–1828. doi: 10.1109/ICIT.2006.372511.

[8] F. Meng, G. Tao, and P. P. Luo, "Dynamic analysis of proportional solenoid for automatic transmission applications," in 2014 International Conference on Mechatronics and Control (ICMC), Jinzhou, China, Jul. 2014, pp. 1120–1124. doi: 10.1109/ICMC.2014.7231727.

[9] S. V. Angadi and R. L. Jackson, "A critical review on the solenoid valve reliability, performance and remaining useful life including its industrial applications," Eng. Fail. Anal., vol. 136, p. 106231, Jun. 2022, doi: 10.1016/j.engfailanal.2022.106231.

[10] D. L. Lavanya, S. J. Kanwar, A. K. Nanda, M. Ravikiran, and P. L. Bora, "Design, Development and Performance Evaluation of Automated Relay Controller for Solenoid Valve Operation During Static Test of Rocket Motors," in 2018 International Conference on Power, Energy, Control and Transmission Systems (ICPECTS), Chennai, Feb. 2018, pp. 275–280. doi: 10.1109/ICPECTS.2018.8521634.

[11] hc, "Solenoid Valve Applications: 13 Most Common Uses," T&X Solenoid, Nov. 19, 2019. https://startersolenoid.net/solenoid-valve-applications/ (accessed Jul. 20, 2022).

[12] M. K. Zavarehi, P. D. Lawrence, and F. Sassani, "Nonlinear modeling and validation of solenoid-controlled pilot-operated servovalves," vol. 4, no. 3, p. 11, 1999.

[13] "13 pub026-002-00_0312.pdf."

[14] A. D. Sullivan, "Solenoid operated valve diagnostic system," CA2138724A1, Jun. 23, 1995 Accessed: Jul. 21, 2022. [Online]. Available: https://patents.google.com/patent/CA2138724A1/en

[15] M. Taghizadeh, A. Ghaffari, and F. Najafi, "Modeling and identification of a solenoid valve for PWM control applications," Comptes Rendus Mécanique, vol. 337, no. 3, pp. 131–140, Mar. 2009, doi: 10.1016/j.crme.2009.03.009.

[16] S. V. Angadi et al., "Reliability and life study of hydraulic solenoid valve. Part 2: Experimental study," Engineering Failure Analysis, vol. 16, no. 3, pp. 944–963, Apr. 2009, doi: 10.1016/j.engfailanal.2008.08.012.

[17] Rotork Fluid Systems, "Understanding the Smart Valve Monitor," no. 03/12, 2012

A Cyclic Run-Through Model for processing BPSec Elements in BPv7 Bundles

Cheol Hea Koo
Mission Operation and Infrastructure
Service Division
Korea Aerospace Research Institute
Daejeon, Republic of Korea
[ORCID: 0000-0002-7180-1476]

***Abstract*—BPSec in a bundle of Bundle Protocol version 7 provides a variety of ways to ensure data integrity and confidentiality during bundle transactions. Due to the complexity of the organization of BPSec elements within a bundle, processing bundles possessing BPSec elements can be complicated, and caution must be taken as the scope of block security expands. This paper presents a potential analysis model of complex BPSec elements in a BPv7 bundle to concisely understand the configuration and relationship between the security block and its target block at a DTN destination node. This will enable dedicated software to handle those BPSec elements efficiently and more deterministic way.**

Keywords—BPv7, bundle protocol, BPSec, BIB, BCB, DTN

I. INTRODUCTION

BPSec (Bundle Protocol Security)[1-2] in a BPv7 bundle[3] is a security extension for the Delay-Tolerant Networking (DTN) technology, which is designed to provide communication in situations where end-to-end network connectivity is not always possible, such as in remote or challenging environments. BPSec provides security services for DTN bundles, which are the basic unit of data transmission in DTN networks. These security services include data integrity, confidentiality, and authentication, which are essential for secure communication in DTN networks including space/air, vehicular and maritime networks. BPSec adds security blocks to DTN bundles, which contain information about the security services used to protect the bundle. These blocks can be used to verify the authenticity of the sender, ensure the integrity of the data, and provide confidentiality of the bundle contents.

While applying integrity and confidentiality mechanisms to bundles in the BPv7 standard using BPSec, the process of constructing the security block along with its target block can become complex and vulnerable to malicious manipulation. Therefore, it is important to quickly and safely understand the structure of the received security blocks in a bundle to address any potential software issues. The purpose of this paper is to propose a practical and intuitively implementable cyclic run-through model for investigating and processing the complex relationship between BPSec integrity and confidentiality security blocks and their respective target blocks. This model is designed to provide a clear understanding of how to handle these elements at a DTN node to which the BPSec elements are finally destined.

II. NOTABLE RELATIONSHIP BETWEEN SECURITY BLOCK AND TARGET BLOCK

To further understand BPSec elements, it is important to note that they consist of a security block, including the BIB (bundle integrity block) and BCB (bundle confidentiality block), and their respective target block, which may include an extension block, primary block, or payload block.

BIB and BCB have different roles and responsibilities with regard to providing and handling the necessary functions of bundle's security capabilities in DTN networks.

Care must be taken when composing and combining BPSec elements. Similarly, equal caution must be applied to understanding the structure of received BPSec elements, as their complexity may cause software to misunderstand their precedence or waste processing time unnecessarily, and to react a falsely configured BPSec elements as a privileged.

For efficient processing of BPSec elements in BPv7 bundles, a concise and simple understanding model is required.

A. BIB and target block

Some notable points that should be considered when a receiving DTN node encounters BIB blocks and its target block can be summarized as follows:

- BIB can have multiple security target blocks with the same security context and configuration. This mechanism is called "*Target Multiplicity*." However, there is no predetermined position between BIB with target multiplicity and its target block except primary block (Block #0) and payload block (Block #1). In other words, BIB block(s) can appear at any position in a bundle, so there is no guarantee of completeness in processing BIB block(s) with target multiplicity until all primary, extension, and payload blocks have been identified and checked to determine whether they are the target block(s) of the BIB.

B. BCB and target block

Some notable points that should be considered when a receiving DTN node encounters BCB blocks and its target block can be summarized as follows:

- Every target block of BCB should have the highest priority for handling over other BIB or extension blocks, as a target block of BCB needs to be decrypted first before any other block can be processed. Otherwise, the bundle block processing engine may not be able to recognize the contents of the *block-type-specific data* of the encrypted target block, leading to assertion errors or unnecessary removal of parts or all of the received bundle.
- Primary block (Block #0) cannot be a target block of BCB.

C. BIB and BCB

There is no requirement for a specific position of locating BIB or BCB in a BPv7 bundle. For example, BIB (#3) can appear before BCB (#2). The forced positional rule is only applied to the primary block and payload block; the primary block (Block #0) must appear first in a bundle, and the payload block (Block #1) must be located at the end of the bundle.

III. EFFICIENT UNDERSTANDING BPSEC ELEMENTS IN A BUNDLE

Although BPv7 and BPSec documents in [1-4] provide rules and precautions for configuring BCB and BIB, it is important to locate a short and concise path for effective processing BPSec elements in a BPv7 bundle, especially for embedded software that processes bundles with relatively limited computing power. This is particularly critical when a DTN engine encounters high-bandwidth DTN transactions and bundles that are heavily protected by BPSec, which can affect most components of a bundle.

An efficient way to investigate BPSec elements in a bundle can be broken down into several steps. Investigating BPSec elements in a bundle involves establishing a clear structural relationship between the security block and its target block and actually means creating a *'building-block'* that can be referenced when finding the shortest path. This can reduce the software processing burden and complexity.

A. Populating building-block of BPSec elements

The actual processing of the BPSec elements in a security bundle cannot start until all bundle blocks are received and recognized. This is because the logical and positional relationship between the security blocks and their target blocks can only be known when all bundle blocks are identified. By populating the building blocks of BPSec elements, the following relationships can be identified, and an appropriate relationship model can be recorded:

- Block number (#) and block type.
- If it is a security block, a list of its target block(s).
- These building blocks should be arranged in the following sequence: Primary block - BCBs (if any) - BIBs (if any) - extension blocks (if any) - payload block. This ensures that out-of-order BPSec elements in a bundle are formalized in the expected order and can be treated *as deterministic way*[4].

I suggest that the populated building-blocks for received BPSec elements should resemble the structure shown in Figure 1.

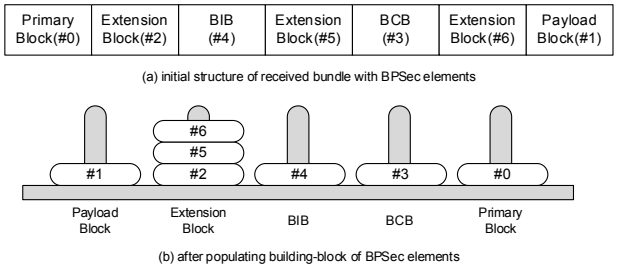


Fig. 1. Populating building-block for BPSec elements

B. Cyclic run-through the building-block

After populating the building blocks for an entire received bundle, which may consist of multiple security blocks, it is suggested to perform a cyclic run-through (CRT) search for processing the building blocks. This helps to minimize the path and effort required for bundle security processing and simplify logical consideration during making decision as a result of the security processing. As shown in (b) of Fig. 1, the CRT search starts from the primary block at the rightmost side and proceeds to the payload block at the leftmost side.

It is desirable to process the BCB first because an encrypted bundle block needs to be decrypted into plain text so a bundle processing engine can recognize the structure and contents of the block-type-specific data in any extension or payload block.

Fig. 2 illustrates a proposed CRT search model, which is designed for use in limited processing environments, such as embedded computing. It should be noted that there may be more flexible ways to perform this search under different computing conditions and approaching ways.

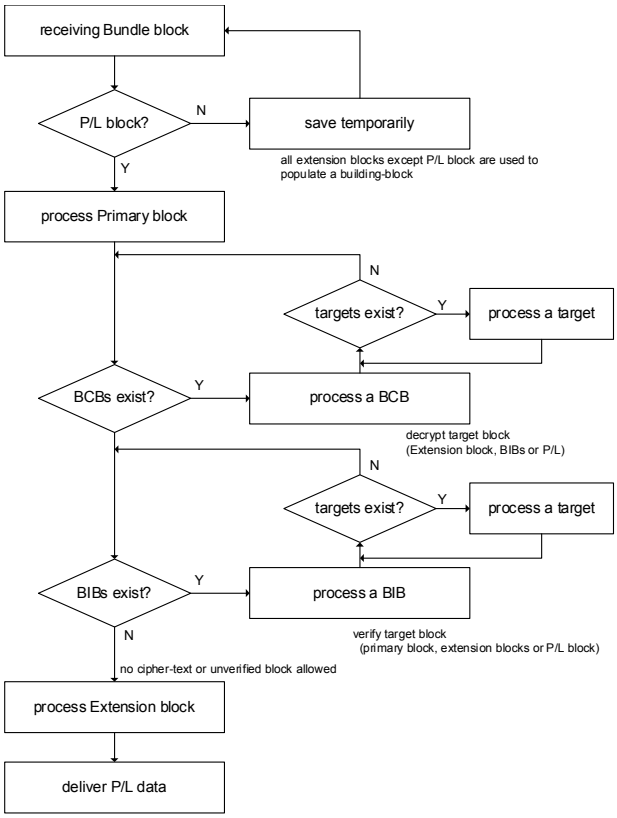


Fig. 2. Proposed sequence diagram of CRT model

IV. CONCLUSION

This CRT model for understanding and processing BPSec elements in a BPv7 bundle has been tested using the KARI's in-house development software of DTN, i3DTN. Full compliance has been checked against examples of bundle security use cases shown in [2]. Further work is needed to perform interoperability tests with other DTN implementations that support BPSec functionality.

REFERENCES

[1] E. Birrane and K. McKeever, "Bundle Protocol Security (BPSec)," RFC-9172, January 2022.

[2] E. Birrane, A. White, and S. Heiner, "Default Security Contexts for Bundle Protocol Security (BPSec)," RFC-9173, January 2022.

[3] S. Burleigh, K. Fall, and E. Birrane, "Bundle Protocol Version 7," RFC-9171, January 2021.

[4] E. Birrane, S. Heiner, and K. McKeever, Securing delay-tolerant networks with BPSec, NJ: Wiley & Sons, 2023.

Secret File Sharing Methodology for Multi-Users

Jae-Mu Choi
Dept. of Information Security
Tongmyong Unversity
Busan, Republic of Korea
tree4843@naver.com

Sung-Hwa Han
Dept. of Information Security
Tongmyong Unversity
Busan, Republic of Korea
shhan@tu.ac.kr

***Abstract*—Confidential information is shared not only in the enterprise environment but also in the personal environment, and the file sharing method is used the most. A number of stakeholders use the file encryption method most often to maintain confidentiality in order to share important information, but there is a disadvantage that confidentiality is not maintained if the encryption key is exposed. In this study, we propose a secret file sharing methodology that allows multiple users to share important information while maintaining confidentiality. The proposed methodology has the advantage of minimizing the leakage of confidentiality when the encryption key is leaked. In addition, in order to verify the effectiveness of the proposed methodology, it was empirically implemented and unit functions were checked. However, since this study is at the level of a shared model, additional research is needed to develop it into a service or solution level.**

Keywords—Secret files, Important Documents, Key management, Cryptography, Access Control

I. INTRODUCTION

In an enterprise environment, critical information is used and shared with stakeholders. This share should be kept confidential by allowing only limited access [1]. Various methods are used for this purpose, but each has its own advantages and disadvantages. If the symmetric key method is used, confidentiality is not maintained when the key is exposed. In addition, if an asymmetric key algorithm is used, only limited stakeholders can access it, but a PKI for information sharing must be established.

In this study, we propose a secret file sharing methodology that uses symmetric keys but allows access only to multiple stakeholders. In addition, it is empirically implemented to verify the effectiveness of the proposed methodology.

II. RELATED WORKS

A. File Sharing Requirements

In an enterprise environment, a web service or SNS is used as a platform for information exchange. A basic information exchange method is file transfer. However, this method is bad for maintaining confidentiality. In the process of transferring the file, unauthorized persons can access it. If the file to be delivered is lost, malicious attackers can exploit it.

In particular, a file may need to be shared by multiple users. Only authorized users should be able to access the file, and a method to block access by unauthorized users is needed.

B. Encryption Alogorithm for Protect Confidentiality

The most effective way to ensure the confidentiality of information recorded in files is encryption [2]. When the encryption algorithm is applied to the information to be protected, the information is converted to cipher text, which is unknown until decryption. The encryption method is divided into a symmetric key method and an asymmetric key method, and the strengths and weaknesses of each are shown in Table 1 [3].

TABLE I. ENCRYPTION ALGORITHM FEATUER

Type	Strengths	Weaknesses
Symmetric Key Algorithm	· high speed · Simple cipher application structure	· Key management burden
Asymmetric Key Algorithm	· High security strength · High application power	· Low performance · PKI required for cipher apply

For such an enterprise environment, a symmetric key or asymmetric key method can be used to share a secret file with multiple users. If the symmetric key method is used, construction can be easy, but there is a disadvantage that confidentiality cannot be maintained if the encryption key is exposed. Conversely, in order to use the asymmetric key method, a PKI must first be established, which has the disadvantage of high cost compared to functionality.

C. Secret File Sharing Requirements

In order for multiple users to share the secret file, the following requirements must be satisfied.

- For a simple structure, a symmetric key algorithm is used. This ensures fast encryption.
- Only authorized users should be able to access the secret file. Access by unauthorized users is blocked.
- Each user uses an individual crypto key, and the crypto key is not shared. Even if one crypto key is exposed, the secret file is protected.

III. SECRET FILE SHARING METHODOLOGY

A. Cripto Key Manage Model

Figure 1 shows the method for multiple users to share the secret file proposed in this study.

The secret file is encrypted with the document crypto key. Only one key is enforced per secret file. This document crypto key is again encrypted with the user crypto key. By applying this structure, each user crypto key can be used to access the secret file. If an unauthorized user accesses the secret file, plain text cannot be verified because it is encrypted.

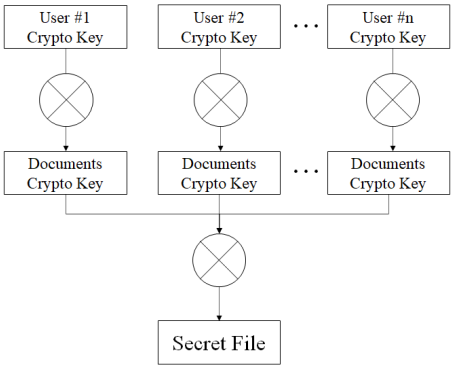


Fig. 1. Secret file sharing structure for multi-users

If the user crypto key is exposed, the confidentiality of the secret file can be maintained by removing the enforcement of the key.

B. *Implements*

The effectiveness of the secret file sharing methodology proposed in this study needs to be verified. For this purpose, I actually implemented it and checked the following unit functions.

- Operation of the secret file encryption/decryption function using the document crypto key
- Operation of the document crypto key encryption function using the user crypto key
- Check to allow authorized users to access the secret file (Open)
- Check secret file access rejection of unauthorized persons

As a result of checking each unit function, it was confirmed that the entire function operated normally. Therefore, the secret file sharing methodology proposed in this study satisfies the four requirements.

IV. Conclustion

The frequency of using files containing important information continues to increase not only in enterprise environments but also in personal environments. In addition, the scope of stakeholders continues to change. For this environment, this study proposed a secret file sharing methodology that only a large number of users can access, and verified its effectiveness through empirical implementation.

However, since this study only proposed a methodology that allows multiple stakeholders to share a secret file using a symmetric key encryption algorithm, additional research is needed to expand to service ice or solutions by applying business processes.

Acknowledgment

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information & communications Technology Promotion)(2018-0-018740301001).

References

[1] F. V. Malmborg and P. A. Strachan, “Advocacy Coalitions and Paths to Policy Change for Promoting Energy Efficiency in European Industry,” Energies, vol.16, no.9, pp.3785, 2023.

[2] K. Wang and J. Guo, “Secure Search over Encrypted Enterprise Data in the Cloud,” International Journal of Network Security, vol.25, no.1, pp.103-112, 2023.

[3] I. Hussain, A. Anees and T. A. Al-Maadeed, “A hybrid encryption system for communication and financial transactions using RSA and a novel symmetric key algorithm, “ Bulletin of Electrical Engineering and Informatics, vol.12, no.2, pp.1148-1158, 2023.

[4] J. Liu, T. Zhao, J. Cao and P. Li, “Interval type-2 fuzzy neural networks with asymmetric MFs based on the twice optimization algorithm for nonlinear system identification,” Information Sciences, vol.629, pp.123-143, 2023

Enhancing Maritime Safety in Mooring Operations through VR Technology

1st Pinar Bilgin
Centre of Excellence in Maritime Safety
Singapore Polytechnic
Singapore, Singapore
pinar_bilgin@sp.edu.sg

2nd Vincent Lee
Centre of Excellence in Maritime Safety
Singapore Polytechnic
Singapore, Singapore
vincent_lee@sp.edu.sg

3rd Tan Rijian
Immersive Experience Technology Center
Singapore Polytechnic
Singapore, Singapore
tan_rijian@sp.edu.sg

4th Goh Zhe An
Immersive Experience Technology Center
Singapore Polytechnic
Singapore, Singapore
goh_zhe_an@sp.edu.sg

5th Chit Htwe
Singapore Maritime Academy
Singapore Polytechnic
Singapore, Singapore
chit_htwe@sp.edu.sg

6th Nor Mohamed Bin Nawawi
Vopak Terminals Singapore Pte Ltd
Singapore, Singapore
nor.mohamed@vopak.com

7th Hui Chen Tee
Vopak Terminals Singapore Pte Ltd
Singapore, Singapore
hui.chen.tee@vopak.com

***Abstract*—This article addresses the growing need for improved maritime training in the pursuit of sustainable operations and cost reduction within the maritime industry. As a fundamental pillar of the maritime industry, mooring training is of utmost importance due to its significance in both ensuring the safety of operations and its overall importance in maritime activities. By integrating immersive technologies such as VR, AR and MR, we can enhance mooring training, fostering a safer and more efficient learning environment that accelerates knowledge acquisition. In this article, we present a novel VR scenario that combines practical training and assessment for mooring operations. Within a secure and controlled environment, our VR scenario provides comprehensive instruction on mooring equipment, ropes, and essential shore-side mooring zones. Through interactive simulations featuring bollards, capstans, and other vital elements, we provide a comprehensive and immersive training experience. Our approach aims to enhance safety, efficiency, and knowledge acquisition in mooring operations.**

***Index Terms*—virtual reality, maritime safety, mooring training, mooring operations**

I. Introduction

Mooring operations are a vital part of maritime industries, ensuring the safe and efficient berthing and anchoring of vessels. Although the maritime industry has a fairly good safety record, maritime accidents could lead to devastating consequences, due to human factors or equipment errors [1], [2]. Although various automation systems and high level equipment are being developed to make maritime operations safer and more sustainable, mooring operations will require people for the foreseeable future [2], [3]. As such, mooring training is critical for various personnel involved in maritime operations.

Traditionally, this training is conducted through a combination of classroom instruction, on-the-job training, and simulation-based training. In classroom training, trainees are taught the basics of mooring operations, safety procedures, and the types of equipment used, as well as different types of mooring lines, winches, and the forces acting on a moored ship. They are also taught about the dangers of mooring and safety precautions they should take during mooring operations. On-the-job training is the practical part of the training where the trainees get hands-on experience under the supervision of experienced crew members. They learn how to handle the equipment and how to communicate effectively during mooring operations. Although on-the-job training is extremely useful, it is critical to eliminate all potential risks with training activities for new trainees. Simulation-based training is a relatively recent training method that can address the risk factor of on-the-job training by first introducing the trainees to the environment via virtual reality, where they can practice the mooring techniques they learned in theory in a safe and controlled setting [4]. This type of training can be very effective as it allows the trainees to gain experience without the risk of injury or damage to the equipment. We can use simulators to replicate different scenarios, weather conditions, and emergencies that may not be feasible or too dangerous to experience in real life.

Emerging immersive technologies, such as virtual reality (VR), augmented reality (AR) and mixed reality (MR) have created new possibilities for maritime simulators and simulation-based experiences for training and operations [4]–[9]. VR, AR and MR have been considered by training organizations as a technology that can significantly improve

seafarer's performance and competency [4]. These training methods are also relatively cheaper, more immersive than theoretical training without the added risk, more compact and accessible in comparison to traditional methods [5], [6], [8], [9].

II. Methodology

A. VR System and Tools

In the developed VR-based mooring operations training program, we are using a high-powered computer that can handle VR applications, the HTC Vive Pro head-mounted display, 4 HTC Vive base stations for motion capture within a room, and Unity software for VR scenario creation. With this setup, the following interaction methods are supported.

- Click: Users can utilize HTC controllers to interact with the environment.
- Eye-Gaze: Users can trigger responses or access more information by focusing their gaze on specific objects or areas.
- Under-the-eye: Users are provided with detailed instructions or prompts related to certain objects or actions within the environment.
- Movement inside a zone: Users are able to physically move within designated zones within the environment. The zone setup is 3 meters by 3 meters.

B. Environment Design

Our VR design includes two modes; Practice and Assessment. The Practice mode is designed as a sort of tutorial in which the users can get used to the VR environment and practice the movements allowed within the design. The Assessment mode is where the user will be doing the full mooring operation without the limits put in during Practice mode. Further details for the two modes are provided below.

1) Practice mode: In this mode, the user will perform three exercises;

- Teleporting (movement within the confines of the environment)
- Pulling of ropes onto bollards
- Use of capstan

a) Teleporting: Teleportation is segmented into 3 zones. The user can walk around and interact with objects within each zone. The user will use the Eye-Gaze feature to teleport from zone to zone. To complete this exercise, the user must teleport to all 3 target locations. After the exercise, a pop-up window will be shown to either Retry or Exit this exercise.

b) Pulling of ropes onto bollards: In this exercise, the user will learn how to use the controllers (left, right) to interact with the ropes and bollard. A monkey fist will be thrown by the ship crew, the user can interact with the monkey fist via pulling the heaving line, mooring the rope, and securing the mooring rope on the bollards. To complete this exercise, the user must secure 2 mooring ropes over the bollard. An excerpt from the scene is shown in Figure 1. After the exercise, a pop-up window will be shown to either Retry or Exit this exercise.

c) Use of capstan: In this exercise, the user will learn how to operate the capstan. The user will connect the messenger rope to one end of the heaving line, and loop the other end over the capstan and start drawing the messenger line. The user will trigger the buttons on the ground to start/stop capstan. To complete this exercise, the user must secure 2 mooring ropes over the hook of the capstan. An excerpt of this scene is shown in Figure 2. After the exercise, a pop-up window will be shown to either Retry or Exit this exercise.

2) Assessment mode: In this mode, the user will be doing the full mooring procedure. The mooring configuration is 3-2-2. Mooring is done at the bow of the ship. The setting for this mode is fine weather during the day with no rain and wind.

For this mode, the user will begin in the office VR environment. A pop-up will inform them that a ship will be arriving for mooring, and they will have to put on their Personal Protective Equipment. Once done, a pop-up will inform the user that they have successfully equipped themselves and are ready to exit the office and proceed to the jetty. At the jetty, the ship will arrive and the user will follow the mooring configuration 3-2-2.

a) Mooring configuration at the spring line bollard is 2: In the first zone, monkey fist is thrown down by the crew. The user must use both controllers to hold the monkey fist and pull the heaving line. The user must pull the heaving line until the first spring line is within reach. Once the spring line is within reach, the user will grab the first spring line rope eye and place it over the bollard using the controller. Then, the user will secure the heaving line onto the railing nearby by holding the heaving line with the controller and walking to the railing. After hanging the heaving line on to the railing, the user will untie the heaving line from the first spring line by clicking on the knot between the heaving line and the mooring rope. Second spring line will be released. The user then can start to pull the heaving line until the second spring line is within reach. Then, the user can grab the second spring line rope eye and place it over the bollard. The user will then untie the heaving line from the second spring line and take the heaving line from the railing. After this, the user will move to the next target destination (breast line platform/breasting dolphin) using teleport.

b) Mooring configuration at the breast line is 2: Moving on to the second zone, the breast line platform/breasting dolphin, the user will bring the heaving line near one end of the messenger line and release the grip to attack the heaving line to the messenger line. The user will then pick up the other end of the messenger line using the controllers and loop over the capstan. Once done, the ship crew will begin drawing the heaving line until the messenger line reaches the ship deck. The ship crew will release the messenger line together with the first breast line. The user will trigger the buttons on the ground to start the capstan to draw the messenger line. Once the first breast line is within reach, the user will trigger the buttons on the ground or move away from the capstan front to stop the capstan front rotating. The user can then grab the first breast line and place it over to the capstan hook. The user will

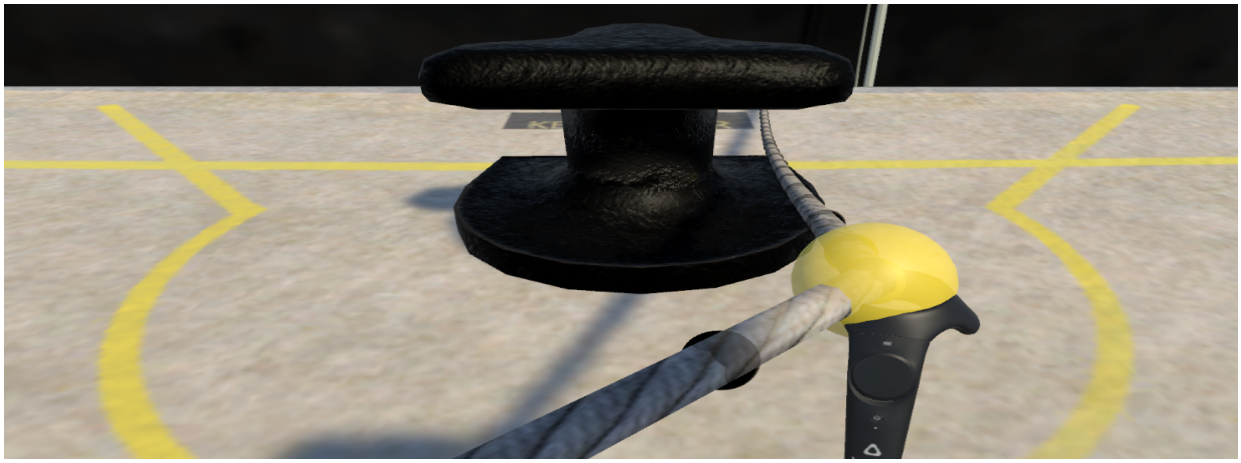


Fig. 1. VR Practice mode - Bollard exercise

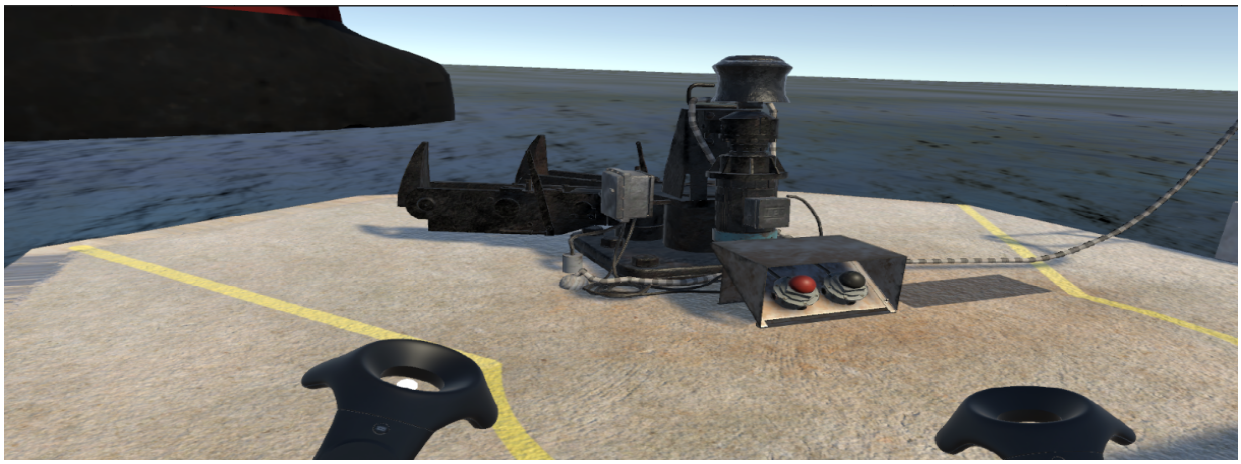


Fig. 2. VR Practice mode - Capstan exercise

Fig. 3. VR Assessment mode - Screenshot from assessment while running capstan

untie the messenger line from the first breast line and re-loop the messenger line on the capstan. Once done, the ship crew will begin drawing the messenger line until it reaches the ship deck. The ship crew will release the messenger line tied to the second breast line. The user will trigger the buttons on the ground to start the capstan to draw the messenger line. Once the second breast line is within reach, the user will trigger the buttons on the ground or move away from the capstan to stop the capstan and grab the second breast line and place it over to the capstan hook. The user will untie the messenger line from the second breast line. After this, the user will pick up the messenger line and move to the next target destination (stern line platform/stern mooring dolphin) using teleport.

c) Mooring configuration at the stern line is 3: Finally, when the user arrives at the stern line platform/stern mooring dolphin, the ship crew will start to release the messenger line tied to the first stern line. Still holding on to the other end of the messenger line, the user will loop over the capstan. Then, the user will trigger the buttons on the ground to start the capstan and capstan will start to draw the messenger line. Once the first stern line is within reach, the user will trigger the buttons on the ground or move away from the capstan to stop the capstan. The user will grab the first stern line and place it over to the first capstan hook and untie the messenger line from the first stern line. The user will re-loop the messenger line on the capstan. Once done, the ship crew will then begin drawing the messenger line until it reaches the ship deck. The ship crew will release the messenger line tied to the second stern line. The user will pick up the other end of the messenger line using the controllers and loop over the capstan. The user will start the capstan to draw the messenger line. Once the second stern line is within reach, the user will stop the capstan and grab the second mooring rope and place it over to the first capstan hook. The user will untie the messenger line from the second stern line and re-loop the messenger line on the capstan. Once done, the ship crew will then begin drawing the messenger line until it reaches the ship deck. The ship crew will release the messenger line tied to the third stern line. The user will pick up the other end of the messenger line and loop over the capstan. The user then starts the capstan to draw the messenger line. Once the third stern line is within reach, the user will stop the capstan and grab the third stern line and place it over to another capstan hook. The user will untie the messenger line from the third mooring rope. All lines must be heaved up and tensioned by the ship and the user must stand behind the snap back zone during the process of tensioning. After this, the mooring procedure will be complete, and a pop-up window will be shown to either Retry or Exit this Assessment.

C. Conclusion and Future Direction

By embracing the power of VR technology, our Mooring VR application offers a transformative solution to enhance maritime safety in mooring operations. Through an immersive and realistic training experience, the trainees can effectively learn and familiarize themselves with the intricacies of the mooring process in a safe, controlled, and interchangeable environment. This will ultimately reduce the occurrence of accidents due to human factors and improve overall safety. With the deployment of this innovative training solution, we aim to significantly improve the way mooring operators are prepared for their critical roles in the maritime industry. In our ongoing efforts for system enhancement, we plan to incorporate additional diverse mooring configurations and environmental conditions to enhance the functionality and practicality of the system. We also intend to integrate mixed reality capabilities, leveraging physical elements such as real-ropes and haptic technologies, such as robotic hands or haptic gloves, to enrich the training experience. These technologies have the potential to enhance realism and tactile feedback, offering trainees a more immersive and engaging environment for practicing mooring operations.

ACKNOWLEDGMENT

Singapore Polytechnic (SP) would like to thank Vopak Terminals Singapore Pte Ltd for their collaboration in the development of this virtual reality (VR) scenario for mooring operations.

REFERENCES

[1] C. Hetherington, R. Flin, and K. Mearns, "Safety in shipping: The human element," Journal of Safety Research, vol. 37, no. 4, pp. 401-411, 2006.
[2] B. Sluiskes, "Safety in Mooring," Terra et Aqua, no. 143, pp. 14-19, 2016.
[3] A. Jha, D. Subedi, P. O. Løvsland, I. Tyapin, L. R. Cenkeramaddi, B. Lozano, and G. Hovland, "Autonomous mooring towards autonomous maritime navigation and offshore operations," in 2020 15th IEEE Conference on Industrial Electronics and Applications (ICIEA), Nov. 2020, pp. 1171-1175.
[4] R. Stone, "Virtual reality for interactive training: an industrial practitioner's viewpoint," International Journal of Human-Computer Studies, vol. 55, no. 4, pp. 699-711, 2001.
[5] K. Hjelmervik, S. Nazir, and A. Myhrvold, "Simulator training for maritime complex tasks: an experimental study," WMU Journal of Maritime Affairs, vol. 17, no. 1, pp. 17-30, 2018.
[6] G. Makransky and S. Klingenberg, "Virtual reality enhances safety training in the maritime industry: An organizational training experiment with a non-WEIRD sample," Journal of Computer Assisted Learning, vol. 38, no. 4, pp. 1127-1140, 2022.
[7] E. Markopoulos, P. Markopoulos, N. Laivuori, C. Moridis, and M. Luimula, "Finger tracking and hand recognition technologies in virtual reality maritime safety training applications," in 2020 11th IEEE International Conference on Cognitive Infocommunications (CogInfoCom), Sept. 2020, pp. 000251-000258.
[8] S. C. Mallam, S. Nazir, and S. K. Renganayagalu, "Rethinking maritime education, training, and operations in the digital era: Applications for emerging immersive technologies," Journal of Marine Science and Engineering, vol. 7, no. 12, p. 428, 2019.
[9] H. Shen, J. Zhang, B. Yang, and B. Jia, "Development of an educational virtual reality training system for marine engineers," Computer Applications in Engineering Education, vol. 27, no. 3, pp. 580-602, 2019.

Response Assessment from Fuzz Testing on Cyber Physical Systems using Artificial Intelligence

Incheol Shin
Department of Computer Engineering
Pukyong National University
Busan, South Korea
icshin@pknu.ac.kr

***Abstract*—The fuzz testing has been considered as an indispensable test-generation tool to identify system vulnerabilities in software security. The testing method utilizes automated directed randomness to travel a variety of execution paths in the systems. The generation of fuzzing data in the fuzz testing has been studied in the wide range of literature, but the identifying abnormal states of the systems after the exposing their defects is analogous for the test to determine whether the testing results are successful or not. In this work, we have been designed a fuzzing response assessment architecture for the vulnerability identification testing on CPS(Cyber Physical Systems) of the smart ships.**

Keywords—fuzz testing, cyber physical systems, security vulnerability assessments.

I. INTRODUCTION

The cyber-physical system has been refered the systems where software and hardware components are seamlessly integrated to conduct well defined tasks, generally. Increasingly, the forementioned operations are highly automated and distributed amongst multiple agents. The fuzz testing technique has been explored to identify the defects of the system by generating and inputting mutated data. The testing disciplines are fundamental security requirements for the CPS to maintain sound status before the deployment in the fields[1].

The generating the fuzzing data has been developed as an autonomous identification of security vulnerabilities in the system. However, the process of the determining the security vulnerabilities from the target systems after injecting the mutated fuzzing data is still analogous for testers to perform. That is, the responses or feedbacks from the system after the processing of the autonomously generated fuzzing data should be analyzed by the testers in order to determine if the fuzzing data cause the fault or crash of the systems. Consequently, the purpose of the fuzz testing is to autonomously cause the abnormal states of the target system, but the final determination of the result from the fuzzing data is still left to human resources.

This work proposes a design to relieve the problem of analogous determining the system states or responses by clustering input data to the system and output data from the system before performing the fuzz testing. Our system has been combined with supervised A.I.(Artificial Intelligence) algorithms to classify the input data by the output data. In addition to the classification, our system also adopts the unsupervised A.I. algorithm to find the similarities between the mutated fuzzing data and input data. By exploring those two autonomous algorithms, we would be able to predict the system responses from the fuzzing data, then the identification of defects in the system could be revealed as a vulnerability to be fixed.

II. DESIGN OF THE RESPONSE ASSESSMENT BASED FUZZ TESTING ARCHITECTURE

A. Architecture for Response Assessment based Fuzz Testing on Cyber Physical Systems

The Fig. 1. shows the design of response assessment based fuzz testing architecture to overcome the limits on the current testing structure.

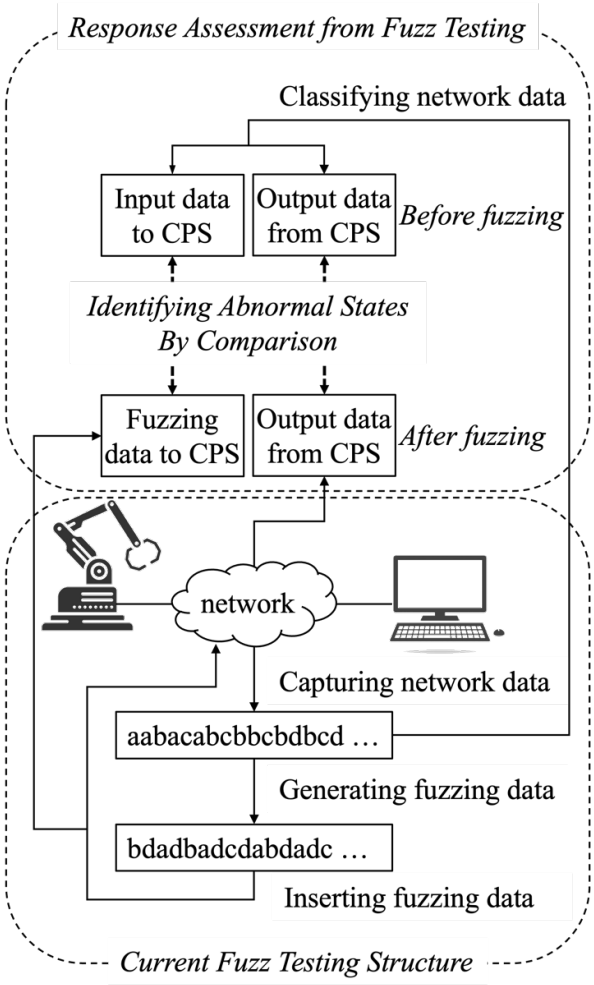


Fig. 1. Response Assessment for Fuzz Testing

An existing fuzz testing architecture is composed with 3 parts, capturing data, generating fuzzing data and inserting the fuzzing data to the target systems. As forementioned in the introduction, since the final determination of the successful fuzz testing is still left to testers, our testing model adopts the

response assessment parts to overcome the problem of analogous process beyond the autonomous analysis.

B. *Advantages of the Response Assessment based Fuzz Testing*

This design explores two A.I. models, supervised and unsupervised algorithms, to build fully automated fuzz testing to identify vulnerabilities of the target systems. The supervised algorithm to learn the classification of input data by output data before injecting the fuzzing data, and the clustering algorithm, unsupervised learning model, groups the mutated fuzzing data with input data by the similarity between them in order to predict the states of the systems after processing the fuzzing data. Consequently, the fuzz testing is fully automated from the generating fuzzing data to identify the system states.

III. Conclusion

Due to the nature of the cyber physical systems, the identification of vulnerabilities on the system before the deployment is one of the crucial processes in the developments. Even though the purpose of the fuzz testing is automated vulnerability assessment of the target system, the final process of determining the response from the system by processing the fuzzing data has been analogous for testers to analyze. However, this work proposes a novel method to augment the existing fuzz testing for fully automated security tool.

References

[1] S. Sheikhi, E. Kim, P. S. Duggirala and S. Bak, "Coverage-Guided Fuzz Testing for Cyber-Physical Systems," 2022 ACM/IEEE 13th International Conference on Cyber-Physical Systems (ICCPS), Milano, Italy, 2022, pp. 24-33, doi: 10.1109/ICCPS54341.2022.00009.

[2] G Baheti, Radhakisan, and Helen Gill. "Cyber-physical systems." The impact of control technology 12.1 (2011): 161-166.

[3] J Wolf, Wayne. "Cyber-physical systems." Computer 42.03 (2009): 88-89.

[4] Alguliyev, Rasim, Yadigar Imamverdiyev, and Lyudmila Sukhostat. "Cyber-physical systems and their security issues." Computers in Industry 100 (2018): 212-223.

Ship Classification Based on AIS Data in Changhua Wind Farm Channel

I-Lun Huang
Maritime Development and Training Center
National Taiwan Ocean University
Keelung, Taiwan, ROC
allen1622@gmail.com

Man-Chun Lee
Department of Merchant Marine
National Taiwan Ocean University
Keelung, Taiwan, ROC
mclee@ntou.edu.tw

June-Chen Huang
Department of Merchant Marine
National Taiwan Ocean University
Keelung, Taiwan, ROC
jchuang@ntou.edu.tw

***Abstract*—The AIS information classifies ships as fishing boats, passenger ships, cargo ships, or oil tankers. Container ships, bulk carriers, general cargo ships, or vehicle carriers are all cargo ships in the AIS code. However, the type of ship is closely related to the maneuverability of the ships, and the maneuverability, acceleration, course stability, steering performance, and stopping distance of different ship types are significantly different. Therefore, in the application of maritime supervision, it becomes an important task to correctly classify and identify these unknown types of ships. This paper develops a ship-type classification model by using Machine learning algorithms based on AIS data. Applied to the Changhua wind farm channel, the experimental test shows that extracting the static and dynamic behavior characteristics of ships can improve the performance of the classifier, and the accuracy rate reaches 97%. The proposed method can solve the problem of insufficient types of ships in AIS information and can be used as characteristic data for subsequent ship navigation anomaly detection. At the same time, the test results also detected abnormal types of ships, showing cases of fishing boats disguised as bulk ships.**

Keywords—Ship-type classification, AIS data, Machine learning, Random forest

I. INTRODUCTION

With the development of the global economy, the number of ships and the volume of transportation is increasing rapidly. The risk of maritime traffic is increasing significantly, especially in busy and complex waters. Serious marine traffic accidents often cause loss of life and irreparable economic loss. Therefore, improving the efficiency of maritime traffic management and supervision has become an important research topic.

In recent years, ship AIS spatiotemporal data have been widely used in the exploration of maritime traffic characteristics, analysis of ship behavior and anomaly detection, maritime safety, maritime situational awareness, maritime rescue, marine pollution, ship supervision, and other fields [1]. The status of ships at sea can be easily identified based on instant AIS information. It can also extract ship navigation features from historical data, construct maritime traffic models or ship behavior models, and be used as important information for maritime supervision and abnormal behavior detection [2]. However, unusual and erroneous information is often hidden in the bulk of AIS records. In particular, some vessels cheat detection by deliberately turning off transponders, falsifying position data, or intentionally transmitting falsified identification data, such as ship type, to conduct illegal operations or even illegal exploration [3]. It is often difficult to find out the true type or identity of certain ships in historical databases, which not only presents challenges for maritime surveillance but also interferes with further data mining and subsequent analysis. In addition, the existing AIS information only classifies ships into fishing boats (ship code 30), passenger ships (code 60), cargo ships (code 70), or oil tankers (code 80), etc. Although container ships, bulk ships, general cargo ships, or vehicle carriers are all cargo ships in the AIS code. However, it is obvious that the type of ship is closely related to the maneuverability of the ship. Generally speaking, different types of ships have different maneuverability, such as acceleration, course stability, steering performance, and stopping distance. For example, the maneuverability of container ships is generally better than that of bulk ships. Therefore, in the subsequent application of maritime supervision, the crew or Vessel Traffic Service (VTS) operators must grasp in advance to predict the behavior of ships in a confined and busy area [4]. Therefore, correctly classifying and identifying these ships becomes an important task.

Traditional maritime navigation supervision primarily relied on charts and radars and was judged by the operator's long-duty experience. The development and innovation of artificial intelligence technologies, such as machine learning and deep learning methods, provide advanced tools for AIS data mining and information extraction. The use of artificial intelligence algorithms to extract potential maritime traffic features from AIS data has become an important research topic for the development of maritime surveillance systems. At present, some studies have achieved certain results by extracting ship features and combining machine learning methods for ship classification. As a preliminary study of maritime surveillance, ship type recognition can be used to identify ship types, provide more detailed ship classification information, and play a vital role in promoting the application of marine surveillance.

The purpose of this research is to apply AIS data to develop ship classification technology, which is a basic project for abnormal detection of ship behavior in the smart navigation safety maritime surveillance system.

II. RELATED WORK

The existing ship type recognition research is mainly divided into two types according to the adopted data types: one is based on image data, such as ship photos taken in ports, aerial or satellite images, and the other is based on information such as speed and heading of AIS data. The image-based ship-type recognition method [5, 6, 7] mainly uses unsupervised learning convolutional neural networks (CNN) to identify the basic features of ship images. However, optical sensors cannot be used at night or in bad weather conditions, and the image-based ship-type recognition system loses its function.

One of the features extracting procedures of ship classification methods based on AIS data is to convert the AIS spatiotemporal data of ship trajectories into graphic data and use graphic recognition algorithms to train the ship type

recognition system [8, 9, 10]. Yang et al. (2022) [10] used AIS data to generate ship trajectory images, including static, standard navigation, and maneuvering states. Use convolutional neural network (CNN) to identify 8 types of ships (that are fishing boats, tugboats, sailing boats, leisure boats, passenger ships, cargo ships, crude oil/oil product ships, and others) from ship track images and the corresponding ship type codes in AIS data are 30, 31, 36, 37, 60, 70, 80, and 90), and obtaining an accuracy rate of 87.5%.

Another feature extraction procedure for ship classification methods based on AIS data is to directly extract the geometric features of ship shape and the motion features of ship behavior from AIS static and dynamic data without converting them into graphic data. And then use the features as learning data for the ship classification model [1, 2, 4, 11, 12, 13]. Wang et al. (2021) [12] used the original AIS static message to contain five fields, A, B, C, D, and draft, as well as the ship type code, to extract the length, width, draft, and geometric features of the ship, such as length and width Static features such as ratio, perimeter, area, etc., where A, B, C, and D are the distances from the reporting reference point O to the bow, stern, port, and starboard of the ship, respectively. Five types of ships, passenger ships, tugboats, oil tankers, fishing boats, and cargo ships, were identified by applying the random forest method with an accuracy rate of 86.14%. The results show that static trajectories and ship shape characteristics may be similar between certain types of ships, such as cargo ships and cruise ships. Only extracting static features is not enough to construct a classification model capable of distinguishing five types of ships. The static and dynamic information of AIS data should be combined to improve the performance of ship-type classification. Yan et al. (2022) [1] used space-borne AIS data, which has the advantages of wide coverage, long tracking time, and rich ship types, to extract static features and dynamic behavior features and analyze and extract them. In addition to the length, width, and draft of the ship, static features refer to the geometric features proposed by Lang et al. (2018) [13], such as naïve perimeter, naïve area, aspect ratio, and shape complex. Dynamic behavior features include ship position, voyage distance, ship speed, and other feature quantities. Based on these two types of features, the random forest model machine learning algorithm is used for five types of ships, including cargo ships, oil tankers, fishing boats, passenger ships, and tugboats, with an accuracy rate of 92.7%. The model test results show that the accuracy rate is obviously improved by combining geometric features and dynamic behavior features.

Compared with the method of converting AIS ship trajectory spatiotemporal data into graphic data, it is simpler, faster, and more efficient to directly extract geometric features and motion behavior features from AIS data and is suitable for real-time aviation safety supervision systems. Baeg & Hammond (2023) [4] proposed a ship classification method for AIS data in Danish waters to solve the problem of missing and tampering ship type information in AIS data. Static and dynamic features (ship geometry, motion features, and time features) were extracted from AIS data, the features of the Danish water's geographical region were integrated, and the ink features of the sketch recognition design were proposed to represent the ship trajectory type. There are a total of 39 features. Various classification algorithms, such as random forests and decision trees, are used for performance comparison. The results and discussion show that Random Forest outperforms other classifiers in classifying AIS data. The classification accuracy of the four ship types can reach 84.05%. In particular, the accuracies of fishing boats and passenger boats were 0.951 and 0.946, respectively, confirming very high results. However, the practical features that distinguish cargo ships from oil tankers need to be further explored.

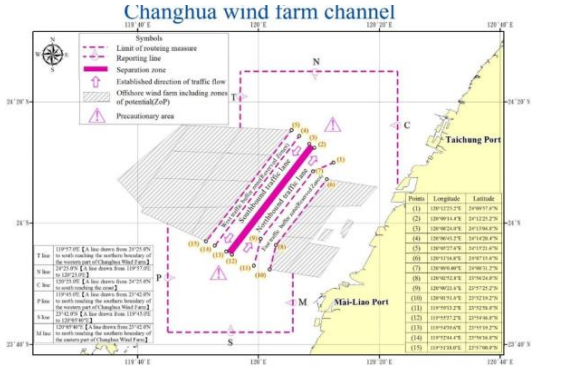
The above review and analysis show that the current research is mainly to solve the problem of missing and tampering ship type information in AIS data. The ship-type classification method aims to confirm the classification of AIS ship-type codes. However, the cargo ships that also belong to the AIS ship type code 70 include bulk carriers, container ships, and general cargo ships, and the ship type code 80 includes crude oil tankers and oil product tankers. The differences in maneuverability of various types of ships significantly affect the navigation behavior of ships. Before the development of the ship navigation safety supervision system, there was an urgent need to develop ship-type identification methods for bulk ships, container ships, general cargo ships, crude oil tankers, or oil product tankers.

III. Data Analysis

3.1 Changhua Wind Farm Channel

The Changhua wind farm channel was officially implemented in October 2021, as shown in Fig. 1. Traffic separation system is adopted, according to International Regulations for Preventing Collisions at Sea (COLREGs), which consists of a southbound traffic lane, a northbound traffic lane, a separation zone, and buffer zones on both sides. The width of the northbound and southbound channels is 2 NM.

Changhua Offshore is a vital water area connecting Kaohsiung Port with Taipei and Keelung Ports. It is also an international shipping route connecting Northeast Asia to Southeast Asia. Many ships pass through this area, and the traffic is very heavy. Fig. 2 shows the maritime traffic density map before and after the implementation of the Changhua wind farm channel. Before the implementation of the channel, the ship tracks spread widely, and the southbound and northbound tracks overlapped. After the channel implementation, the ship tracks concentrated in the southbound and northbound channels, respectively. On average, there are at least 150 ships passing through this channel every day, and most of the ships can sail in compliance with the regulations of the channel.



Maritime and Port Bureau, MOTC.

Fig. 1 The Changhua wind farm channel

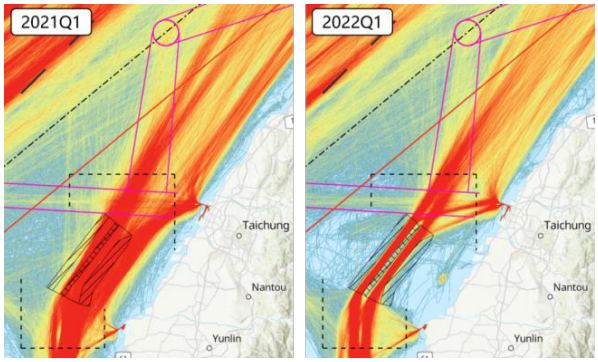


Fig. 2 Traffic density before and after wind farm channel execution

3.2 Data Processing

The AIS data used in this study are provided by the Maritime and Port Bureau, MOTC. The amount of AIS historical records is nearly 10 billion for the period year 2022, which is equivalent to 2TB of massive data. And we extract the AIS trajectory of the cargo ship (Ship and Cargo Type code: 70) passing through the channel. The data quality of AIS significantly affects the performance of ship type recognition and the classification model. The data processing was conducted according to the AIS numerical standard of the International Telecommunication Union (ITU) [14]. The AIS data with apparent abnormalities, such as longitude exceeding 180° , latitude exceeding 90° , etc., was cleared. The static and dynamic data of ship platform shape and navigation parameters are extracted, which include MMSI that can identify a single ship, navigation parameters (SOG, COG, THD, Longitude, Latitude), ship geometric parameters (A, B, C, D, Draught), and the "Record Time" that can create AIS time-space sequences.

In order to effectively analyze the navigation characteristics of ships in the channel, the data representing the track line are normalized. Refer to [15] to create a set of analysis gate lines perpendicular to the southbound traffic lane in the head section of the channel. The distance between two adjacent crossing-line is taken as a constant of 100m, and there are 50 crossing-lines in total. Through spatial analysis calculations, each trajectory passing through the channel intersects with the crossing lines. The ship's lateral position, speed, and heading data, provided by the AIS information, are recorded at each crossing-line. So that the temporal and spatial distribution of the training data is consistent, as shown in Fig. 3. Finally, our research obtained a total of 9,503 cargo ship trajectories (Bulk carrier 4340, container 3895, general cargo 1268), 47,515 static feature parameters, 475,150 intersection points of trajectories and crossing-lines, and 1,425,450 dynamic feature data.

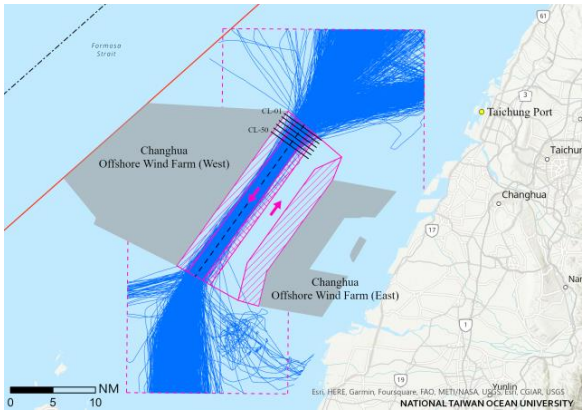


Fig. 3 The trajectories of cargo ships (type 70) in the southbound traffic lane with crossing-lines.

IV. Classification Model

4.1 Feature Extraction

The data fields A, B, C, and D in AIS represent the distances from the antenna or reference point O to the bow, stern, port side, and starboard side, respectively. Then, the features of ship length and width can be calculated by Eq. (1):

$$\begin{cases} Length = A + B \\ Width = C + D \end{cases} \tag{1}$$

In addition, ship Draught is also included, as well as the geometric characteristics of the ship shape, including the Perimeter, Area, Aspect Ratio, Shape Complexity, and position ratio of the bridge over the ship length, named Bridge Position Ratio. Because most antenna positions or reference O points are located on the bridge, this value can be obtained from field A. This parameter is essential because, due to the influence of the shipbuilding industry, large or ultra-large container ships adopt the form of double bridges, and the antenna or reference O point will be set on the front bridge, which is different from traditional container ships. And generally, container ships still have space for stacking containers behind the bridge. Still, the main loading positions of bulk carriers and general cargo ships are in front of the bridge, so the value A of bulk carriers and general cargo ships will be closer to the ship's length. Furthermore, general cargo ships are among the smaller vessels, so that the draught will be shallower than bulk carriers and container ships. And in the case of the same length, the cargo-carrying characteristics of a container ship need to conform to the depth of most ports, so the draught of a container ship will be smaller than that of a bulk carrier. The relevant definitions are given in Eq. (2).

$$\begin{cases} Perimeter = 2 \times (Length + Width) \\ Area = Length \times Width \\ Aspect\ Ratio = Width/Length \\ Shape\ Complexity = (Length + Width)^2/(Length + Width) \\ Bridge\ Position\ Ratio = A/Length \end{cases} \tag{2}$$

To improve the accuracy of the ship-type classification model, the dynamic voyage features in AIS Data are also extracted in this study. In processing dynamic information data, through the crossing-line method in section 3.2, the navigation information originally recorded on the AIS is converted to a specific crossing-line, and the coordinates data are converted into the lateral distance from the center line of the lane on each crossing-line. The navigational parameters collected on each crossing line represent the information of ships passing through a specific channel section along the channel direction.

In summary, 11 features were extracted in this study, including ship geometry and trajectory behavior features. To further understand the relationship between each feature of the three ship-type, container, bulk carrier, and general cargo ship. In terms of static features, the data distribution shows that the length, width, area, and perimeter of the three ship types have apparent differences compared with other static parameters. In contrast, in terms of behavior features, the COG and lateral position is similar in addition to the difference in speed distribution. We speculate that the reason for such data distribution is that the ship is limited by the channel during navigation. Hence, the distribution trend of COG and lateral

position is consistent, and only the speed is different due to the influence of different ship types. Finally, we eliminated the COG and lateral positions of the trajectory behavior features and only modeled with nine feature data.

4.2 Classification Model

This research uses the classification algorithm provided by Python scikit-learn Class to conduct a series of experiments to evaluate the extracted features for ship-type classification. After testing, the standard classification metric, such as accuracy and F1 score, of the random forest algorithm is obviously better than other methods in this problem, so the random forest algorithm is adopted.

Because in the tabbed data set, the number of trajectories of the three ship types is unevenly distributed. To make the weights of the supervised learning classifier consistent for all kinds of ship types, we randomly select 1000 trajectories for each class, a total of 3000 records as training data, and randomly choose 300 works from the rest of the set as test data. We iteratively performed 100 tests, and the accuracy distribution is shown in Fig. 4. In all tests, the scores are higher than 90% nearly 95 times, indicating that the classifier's performance is excellent. In the highest scoring case, both the accuracy and F1 score are 97.3%, and the confusion matrix of the three types is also shown in Fig. 4.

Among the 8 misclassified ships in Fig. 4, the initial labels of the two ships were bulk carrier and general cargo ship. However, after the model's classification, they are classified as container ships. After reconfirming through the ship database on the Internet (www.vesselfinder.com), it was found that the initial labeling of the bulk carrier was actually a container ship. Significantly, On the other ship, its label is a general cargo ship, and the AIS ship-type code is also 70 (cargo). But the ship database on the Internet shows it is a fishing boat. As a result, the ship classifier in this study, in addition to checking the correctness of the original label, can also detect ships with abnormal ship-type codes.

V. CONCLUSIONS

This paper develops a ship-type classification model by using Machine learning algorithms based on AIS data. The accuracy of the ship classification model can be improved by analyzing the AIS data of the Changhua wind farm channel and comprehensively extracting 9 features, including static features and dynamic features (considering the speed, course, and lateral distance). After testing, the standard classification indicators of the random forest algorithm, such as accuracy and F1 score, are significantly better than other methods on this problem, so the random forest algorithm is used.

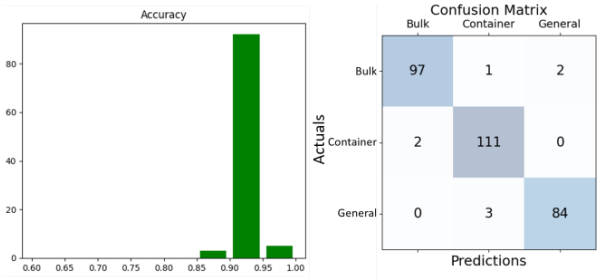


Fig. 4 Distribution of Accuracy of 100 test results and Confusion matrix of the best score.

Applied to the Changhua wind farm channel, extracting the static and dynamic behavior characteristics of ships can improve the performance of the classifier, and the accuracy rate reaches 97%. The cargo ship (container ship, bulk ship, and general cargo ship) classification method proposed in this study can solve the problem of insufficient ship types in AIS information. And it can be used as characteristic data for subsequent ship navigation anomaly detection. At the same time, the test results also detected abnormal types of ships, showing cases of fishing boats disguised as bulk ships.

ACKNOWLEDGMENT

The support of this research by the Maritime and Port Bureau, MOTC, TAIWAN through the grant MPB1160602C005 and National Science and Technology Council, TAIWAN through the grant MOST 111-2410-H-019-017 are acknowledged.

REFERENCES

[1] Yan, Z., Song, X., Zhong, H., Yang, L., & Wang, Y. (2022). Ship Classification and Anomaly Detection Based on Spaceborne AIS Data Considering Behavior Characteristics. *Sensors*, *22*(20), 7713.

[2] Sheng, K., Liu, Z., Zhou, D., He, A., & Feng, C. (2018). Research on ship classification based on trajectory features. *The Journal of Navigation*, *71*(1), 100-116.

[3] McCauley, D. J., Woods, P., Sullivan, B., Bergman, B., Jablonicky, C., Roan, A., Hirshfield, M., Boerder, K. & Worm, B. (2016). Ending hide and seek at sea. *Science*, *351*(6278), 1148-1150.

[4] Baeg, S., & Hammond, T. (2023). Ship Type Classification Based on The Ship Navigating Trajectory and Machine Learning. In Joint Proceedings of the ACM IUI Workshops.

[5] Wang, Y., Wang, C., & Zhang, H. (2018). Ship classification in high-resolution SAR images using deep learning of small datasets. *Sensors*, *18*(9), 2929.

[6] Li, J., Qu, C., & Shao, J. (2017, November). Ship detection in SAR images is based on an improved, faster R-CNN. In *2017 SAR in Big Data Era: Models, Methods, and Applications (BIGSARDATA)* (pp. 1-6). IEEE.

[7] Xu, Y., & Lang, H. (2020). Ship classification in SAR images with geometric transfer metric learning. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(8), 6799-6813.

[8] Luo, P., Gao, J., Wang, G., & Han, Y. (2021). Research on Ship Classification Method Based on AIS Data. In Computer Supported Cooperative Work and Social Computing: 15th CCF Conference, Chinese CSCW 2020, Shenzhen, China, November 7–9, 2020, Revised Selected Papers 15 (pp. 222-236). Springer Singapore.

[9] Li, T., Xu, H., & Zeng, W. (2021, September). Ship Classification Method for Massive AIS Trajectories Based on GNN. In *Journal of Physics: Conference Series* (Vol. 2025, No. 1, p. 012024). IOP Publishing.

[10] Yang, T., Wang, X., & Liu, Z. (2022). Ship Type Recognition Based on Ship Navigating Trajectory and Convolutional Neural Network. *Journal of Marine Science and Engineering*, *10*(1), 84.

[11] Sánchez Pedroche, D., Amigo, D., García, J., & Molina, J. M. (2020). Architecture for trajectory-based fishing ship classification with AIS data. *Sensors*, *20*(13), 3782.

[12] Wang, Y., Yang, L., Song, X., & Li, X. (2021, November). Ship classification based on random forest using static information from AIS data. In *Journal of Physics: Conference Series* (Vol. 2113, No. 1, p. 012072). IOP Publishing.

[13] Lang, H., Wu, S., & Xu, Y. (2018). Ship classification in SAR images improved by AIS knowledge transfer. *IEEE Geoscience and Remote Sensing Letters*, *15*(3), 439-443.

[14] Series, M. (2014). Technical characteristics for an automatic identification system using time-division multiple access in the VHF maritime mobile band. Recommendation ITU: Geneva, Switzerland, 1371-1375.

[15] Huang, J. C., Nieh, C. Y., & Kuo, H. C. (2019). Risk assessment of ships maneuvering in an approaching channel based on AIS data. Ocean Engineering, 173, 399-414.

Deep learning-based retaining wall management system using YOLOv7

Soon Ryan Kwon
Dept. of Electronic Engineering,
Tongmyong University Busan, Korea
srkwon@tu.ac.kr

Woo Seung Lim
Dept. of Electronic Engineering,
Tongmyong University Busan, Korea
dntpdl0530@naver.com

Min Seok Kim
Dept. of Electronic Engineering,
Tongmyong University Busan, Korea
snake1227@naver.com

Enkhtuvshin Myagmar
Dept. of Electric, Electronic and
Information Communication Engineering
Busan, Korea
tuvshoo@naver.com

Abstract—The retaining wall safety management system monitors the condition of retaining wall structures and detects damages early to ensure safety. This paper aims to design and implement a deep learning-based retaining wall management system using YOLOv7, an AI-based automated inspection system utilizing surveillance cameras, to prevent human accidents and property loss.

Keywords—safe management system, deep learning, YOLO, AI, retaining wall

I. INTRODUCTION

The retaining wall safety management system monitors the condition of retaining wall structures and detects damages early to ensure safety. The system aims to secure the stability of retaining walls by taking appropriate measures.

Several domestic and foreign cases of retaining wall collapse are as follows:

- In 2015, a retaining wall collapsed at an apartment site in Gwangju, spilling about 1,000 tons of earth and concrete into the parking lot, damaging many parked cars[1].
- At night in 2018, an accident occurred in the Sangdo Kindergarten building in Dongjak-gu, Seoul, when the retaining wall that supported the foundation of the kindergarten building collapsed and the three-story kindergarten building was bent. Had it occurred during the day, it could have been a disaster affecting more than 120 children[2].
- In 2019, at least 27 people were killed and more than 70 injured when three retaining walls collapsed due to heavy rain in the western Indian state of Maharashtra[3].

Various studies have been conducted to effectively cope with above types of retaining wall collapse accidents in advance. In domestic patent cases, [4] and [5] were proposed. [6] was presented as an overseas research case.

Existing sensor-based retaining wall safety management systems have several limitations and problems. Due to sensor network limitations, it can be difficult to sufficiently monitor the condition of retaining wall structures. Additionally, false alarms may occur due to noise, measurement errors, and environmental changes in the process of collecting, processing, and analyzing large amounts of sensor data.

To overcome these limitations and improve the performance and reliability of the retaining wall safety management system, it is necessary to utilize innovative technologies such as real-time monitoring based on deep learning image recognition.

This paper aims to design and implement a deep learning-based retaining wall management system using YOLOv7, an AI(Artificial Intelligence)-based approach, to automatically inspect and ensure the safety of retaining walls, thereby preventing casualties.

II. SYSTEM ARCHITECTURE

The architectural diagram of the deep learning-based retaining wall management system using YOLOv7 is shown in Figure 1.

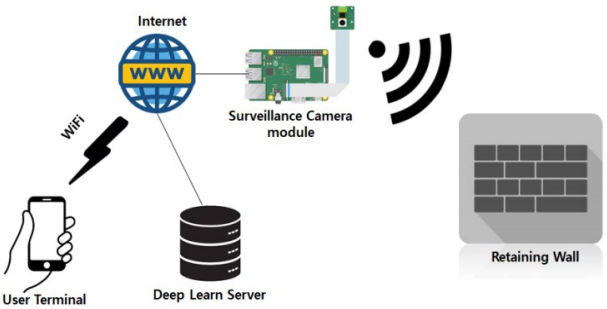


Fig. 1. Structural diagram of the deep learning-based retaining wall management system.

The system consists of the following components: a surveillance camera module that captures real-time footage of the retaining wall and sends it to the deep learning server, a deep learning server that analyzes the received wall footage for events and notifies the user's terminal when an event occurs, and a user terminal that receives crack alarms and remotely monitors the status of the retaining wall. The surveillance camera module is based on Raspberry Pi 4.

III. SYSTEM DESIGN

A. Function Definitions

1) Surveillance Camera Module

- Real-time video recording and transmission to the deep learning server

2) Deep Learning Server

- Deep learning training (retaining wall cracks, non-cracked retaining wall)
- Detection of deep learning event objects in the received video

- Push notification to the user terminal in the detection of event objects (retaining wall cracks)
- Displaying the corresponding video on the server in the detection of event objects (retaining wall cracks)

3) User Terminal

- Receiving push notification upon event occurrence
- Monitoring the corresponding video upon event occurrence

B. Deep Learning Model

The development environment for the deep learning model in the server system was implemented using a PC with GPU. The YOLOv7 object detection algorithm was used for the training model, which predicted the positions of objects and created bounding boxes. Image training was conducted using Darknet. The detailed process for implementing the model is as follows:

- *Acquiring the Dataset for Image Training* : to acquire images for deep learning training, open-source AI images were obtained. Approximately 1,500 retaining wall images were selected.
- *Image Labeling Process* : the labeling program, LabelImg, was used to attach labels to the images. Bounding boxes were created, and the areas for supervised learning of retaining wall crack images were specified, as shown in Figure 2.

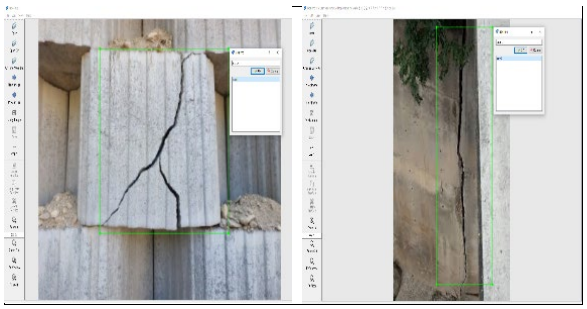


Fig. 2. Image labeling for retaining wall (video samples).

IV. System Implementation and Test Results

A. Surveillance Camera Module

As shown in Figure 3, the surveillance camera module was implemented by attaching a Pi camera to Raspberry Pi 4.

Fig. 3. Surveillance camera module based on the Raspberry Pi 4.

The operating system used was Raspbian. HTTP(HyperText Transfer Protocol) was used as the communication method between the Raspberry Pi and the server. Images of retaining wall cracks used pre-recorded video footage instead of actual footage from the site, as shown in Figure 4.

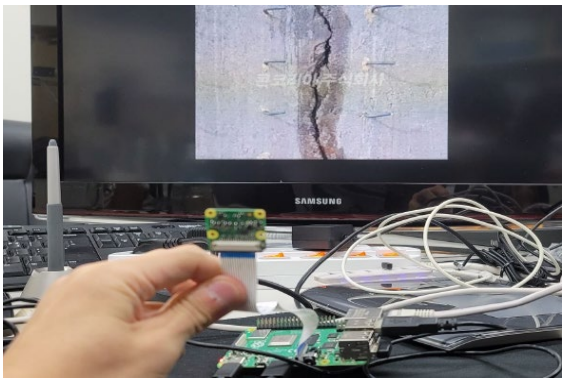
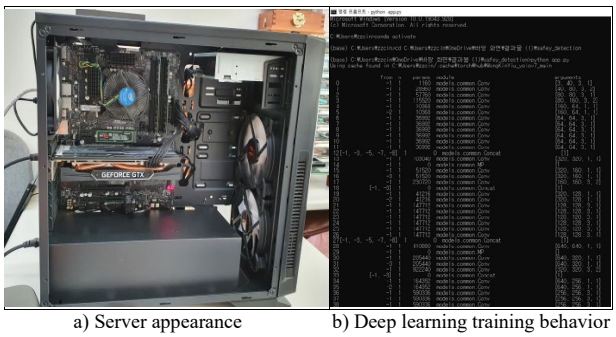


Fig. 4. Pre-recorded video screen shots via camera module

B. Deep Learning Server

Deep learning server appearance and deep learning training behavior are shown in Figure 5.

- Processor: Intel® Xeon® E-2136 CPU @ 3.30GHz, 3.31GHz
- RAM: 16GB
- OS: Window 10 Pro



a) Server appearance b) Deep learning training behavior

Fig. 5. Deep learning server appearance and deep learning training behavior

The test results of deep learning-based retaining wall crack video analysis are shown in Figure 6.

a) Original image b) Crack detection image

Fig. 6. Test results of deep learning-based retaining wall crack video analysis.

C. User Terminal

The user terminal was implemented based on the Android OS, providing push notifications for risk events and functionality for video monitoring.

The initial screen display of the user terminal is shown in Figure 7.

Fig. 7. Initial screen display of the user terminal.

The user terminal receives crack alerts and displays the monitoring screen, as shown in Figure 8.

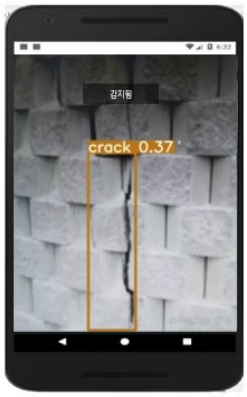


Fig. 8. Receipt of crack alert and monitoring screen on the user terminal.

D. Test Results

As reviewed above, it was confirmed that the functions defined and designed in the server and user terminal worked normally during the functionality test.

Additionally, although not explicitly defined as a function, simulation results using tools for the loss value trend of crack image recognition through deep learning training over 1,200 iterations revealed a crack recognition error rate of approximately 0.2%. Based on the test results mentioned above, the evaluation of the implemented system's crack recognition accuracy indicated an average accuracy level of approximately 99.8%.

V. Conclusion

In this paper, a deep learning-based retaining wall management system using YOLOv7 was designed and implemented. The system captured retaining wall crack sites as video information through surveillance cameras, analyzed the cracks using deep learning techniques, notified the occurrence of cracks to the user terminal, and enabled monitoring of the crack sites. The paper presented the system's architectural diagram, performed experiments on the implemented functionalities, and evaluated the system's accuracy.

Future tasks include enhancing and adding functionalities, as well as improving performance, to develop a competitive commercial service system that can be applied in real-world environments.

Acknowledgment

This work was supported by a grant from Tongmyong University Innovated University Research Park (I-URP) funded by Busan Metropolitan City, Republic of Korea. (IURP2301).

References

[1] Authorities search for possible casualties in Gwangju accident : http://www.koreaherald.com/view.php?ud=20150205000713&ACE_SEARCH=1

[2] Preschool building collapse in Seoul narrowly misses disaster: http://www.koreatimes.co.kr/www/nation/2023/07/113_255222.html

[3] Three retaining walls collapse after heavy rains in western India 27 deaths (total): http://www.yna.co.kr/view/AKR20190702111451104

[4] Patent: Daehan GeoENC Co., Ltd. - Retaining wall and stone management system using LoRa network system: Patent Registration Number: Korea/1020180172975

[5] Patent: Korea Structural Safety Corporation - Inspection System for Safty of Retaining Walls and Inspection Method Utilizing the System: Patent Registration Number: Korea/1020210054277

[6] "Asset Management for Retaining Walls", Thechnical Report, Department of Civil & Environmental Engineering, University of Michigan, March 2020.

Learning-based phase shift of UAV-mounted RIS in maritime communications

Donghyeon Kim
AI & Robotics Tech Lab (ART Lab)
Seongnam, Republic of Korea
donghyeon@artlab.ai

Yonggang Kim*
Division of Computer Science and Engineering
Kongju National University
Cheonan, Republic of Korea
ygkim@kongju.ac.kr

***Abstract*—In maritime communications, UAVs could be deployed as relay nodes to support communications between source nodes and destination nodes. For efficient relaying, reconfigurable intelligent surface (RIS) could be attached to the UAV. The reflecting elements of RIS reflects the signals from source node to the desired direction to improve transmission performance. However, unstable channel state caused by weather, waves, or fluctuations of the nodes may still degrade the transmission performance. In this study, we propose a method that utilizes learning technique to compensate the channel estimation uncertainty for efficient phase shift or RIS.**

I. Introduction

In maritime communications, a ship or a buoy on the sea may transmit data to the center on the ground. However, as the distance from a source node to a destination node becomes longer, it is hard to achieve high data rates without deploying relay nodes. For efficient and cost-effective coverage extension in relay system, reconfigurable intelligent surfaces (RISs) has recently attracted a lot of attention for relay transmission system. A RIS is a programmable structure that controls incoming wireless signals by changing the electric and magnetic properties of reflecting elements at the surface. The phases of incoming signals are shifted as desired direction, and thus, the signals are beamformed and constructively combined at the destination with the appropriate phase shift techniques. Because of easiness in deployment and relatively low cost compared to the conventional relay nodes equipped with signal processing functions, scenario of utilizing RIS for data transmission has been actively studied.

Phan *et al.* studied cooperative relay transmission with multiple RISs over Nakagami-m fading channels [1]. The direct link between a source node and a destination node is cooperatively considered with the multiple paths passing through multiple RISs that reflects the signals from the source node to the destination node. The authors analyze the performance of considered relay system over Nakagam-m fading channels, and showed that the performance could be efficiently improved with RISs. Li *et al.* studied RIS assisted UAV communications where a UAV and a source node on the ground communicate with each other through RIS for relaying [2]. Owing to the mobility of UAV, the phase of RIS should be reactively adjusted following the trajectory of UAV. The authors jointly design the UAV trajectory and the phase shift of RIS. Zhai *et al.* [3] studied a UAV-mounted RIS assisted mobile edge computing where a UAV-mounted RIS assists communication between the ground users and an MEC server. The authors proposed a method that increases the energy efficiency of the system with consideration of UAV trajectory, RIS passive beamforming, and MEC resource allocation.

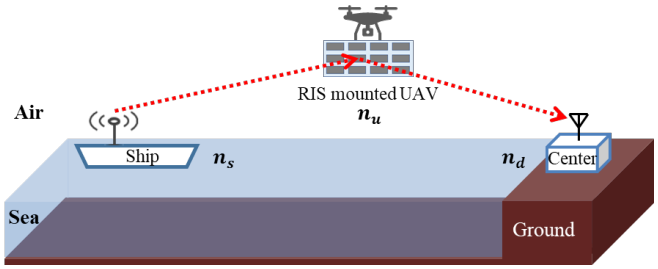


Fig. 1. RIS-assisted FD relay transmission.

The UAV and RIS are efficiently deployed to enhance the transmission efficiency and network coverage. However, a ship or a buoy may have large fluctuation of locations in maritime environments such as waves or weather, and thus, channel could be unstable and may hard to adjust phase of RIS for relay transmissions. In this study, we utilizes learning technique for phase shift to compensate the unstable channel conditions in maritime communications.

II. Learning-Based Phase Shift of RIS

We consider a network scenario where a source node n_s transmits data to a destination node n_d through UAV-mounted RIS n_u as shown in Fig. 1. The source node could be a buoy or a ship on the sea surface in which tries to transmit data to the destination such as monitoring center on the ground. After the data transmissions from the source node n_s, the destination node n_d may transmit acknowledgement whether the message is successfully received. However, when the distance from the source node to the destination node, data transmission with high data rates could be hard to be performed. To efficiently enhance the network coverage, UAV-mounted RIS could be deployed in the middle of the transmission path to relay data transmissions. Reflecting elements in the RIS make phase shift of incoming signals so that the signals are propagated toward the desired direction. Let the RIS consist of R reflecting elements, i.e., $\mathcal{R} = \{r_1, r_2, \cdots, r_R\}$. Each

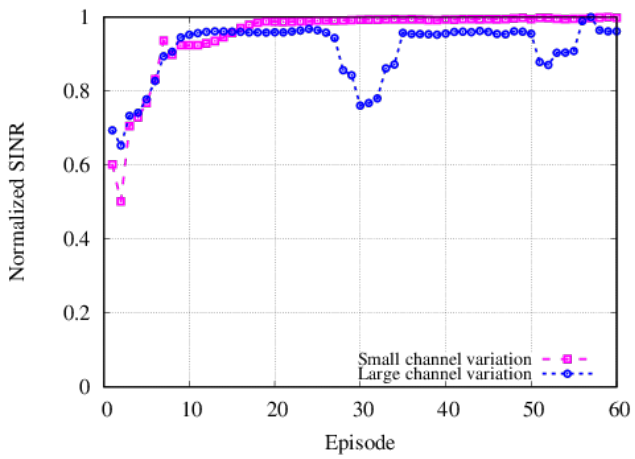


Fig. 2. Normalized SINR of relay transmissions.

element is represented as $r_l = |r_l|e^{j\varphi_l}$, $r_l \in \mathcal{R}$ where $|r_l| \in [0, 1)$ for the passive RIS without amplifiers. Then, the feature of RIS is represented as follows [4]:

$$\Theta = \text{diag}(|r_1|e^{j\varphi_1}, |r_2|e^{j\varphi_2}, \cdots, |r_R|e^{j\varphi_R}). \quad (1)$$

Θ represents the phase shift status of RIS. The phase shift status of RIS affects the performance of relay transmissions using UAV-mounted RIS in the considered maritime communications scenario.

The relay transmission channel from n_s to n_d through n_u could be unstable because of unstable environments such as fluctuation of locations, transmission distance, or whether. The destination node may transmit the acknowledgement messages to the relay UAV node and the source node. The acknowledgement message may contain information about whether the siganl-to-interference-plus-noise ratio (SINR) increases. At the UAV-mounted RIS which has control function for RIS phase shift, aggregate the SINR level change information for the duration T and adjusts the phases of reflecting elements. The proposed method exploits deep Q-network (DQN) for phase shift or RIS. The state and action are modeled as phases of reflecting elements and phase adjustment step, respectively. According to the phases of reflecting elements, phase adjustment step is chosen among the feasible actions to improve the reward calculated by averaging the SINR levels over time T.

III. Performance Evaluation

We evaluate the performance of the DQN-based phase shift in RIS-assisted relay transmissions. For simple evaluation, we assume a network environment where exists single pair of data transmission supported by eight elements of RIS, i.e., $R = 8$. The phase of reflecting elements are shifted as π/Δ, $\Delta = 360$. The transmission channel has variation owing to the maritime environment, and the phase uncertainties are assumed to be 1 degree for small channel variation and 2 degree for large channel variation at most.

Figure 2 shows the normalized SINR with regard to the training episode. When the channel variation is small, the SINR converges after 20 episodes, and shows no degradation after the convergence. However, when the channel variation is large, there is drastic performance degradation points during learning periods. Although the learning-based RIS-assisted relay transmission could improve the throughput performance, channel estimation uncertainty in maritime communications could severely degrade the relay performance.

IV. Conclusion and Future Work

In this study, we studied the relay transmissions with UAV-mounted RIS in maritime communications. UAV-mounted RIS could extend the network coverage. However, the channel fluctuation or estimation uncertainty could be severe owing to the characteristics of maritime environments. The proposed method utilizes learning techniques to compensate the channel variations. The evaluation results show that the learning-based phase shift of RIS improves the throughput performance in maritime relay communications. For future work, we model the channel estimation uncertainty in maritime communications where source nodes are densely deployed. With the transmission information including SINR change over time, the channel uncertainty could be compensated and predicted for better throughput performance. With the channel uncertainty modeling, we try to design more sophisticated method that efficiently compensate the maritime communication channel for efficient relay system using UAV-mounted RIS.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2022-00166739).

References

[1] V.-D. Phan, B. C. Nguyen, T. M. Hoang, T. N. Nguyen, P. T. Tran, B. V. Minh, and M. Voznak, "Performance of cooperative communication system with multiple reconfigurable intelligent surfaces over Nakagami-m fading channels," *IEEE Access*, vol. 10, pp. 9806–9816, 2022.

[2] S. Li, B. Duo, X. Yuan, Y.-C. Liang, and M. Di Renzo, "Reconfigurable intelligent surface assisted UAV communication: Joint trajectory design and passive beamforming," *IEEE Wireless Communications Letters*, vol. 9, no. 5, pp. 716–720, 2020.

[3] Z. Zhai, X. Dai, B. Duo, X. Wang, and X. Yuan, "Energy-efficient UAV-mounted RIS assisted mobile edge computing," *IEEE Wireless Communications Letters*, vol. 11, no. 12, pp. 2507–2511, 2022.

[4] S. Atapattu, R. Fan, P. Dharmawansa, G. Wang, J. Evans, and T. A. Tsiftsis, "Reconfigurable intelligent surface assisted two–way communications: Performance analysis and optimization," *IEEE Transactions on Communications*, vol. 68, no. 10, pp. 6552–6567, 2020.

Beamforming Gain Field Experiments for Long Range Wireless Communication in LoS-MIMO Channel

Woo Yong Lee and Keunyoung Kim
Mobile Communication Research Division,
Telecommunication & Media Research Laboratory
Electronics and Telecommunications Research Institute (ETRI)
Daejeon, Republic of Korea
{wylee, kykim12}@etri.re.kr

Minki Jee, Namho Kim, Wonseog Ko, and Kim Younggyun

Wiznova, Inc.
Seongnam, Republic of Korea

{gs_minki, ykim}@ wiznova.com

***Abstract*— In an outdoor Line of Sight (LOS) mobile communication environment where power consumption, size and weight of equipment are limited, such as in extreme cold regions, communication equipment conditions for unmanned exploration vehicles require high-speed communication of 10Mbps or more over a long distance of 50km or more. In this paper, we focus on the beamforming gain and examine the possibility of long-distance high-speed communication. In order to overcome signal attenuation due to long-distance transmission, we apply beamforming as a multiple input/output (MIMO) communication technology, and communication distance of 50km and 10Mbps at 5.8GHz carrier frequency (to avoid collision with crevasses detection radar frequency). And we verify the operation of a wireless communication system that simultaneously provides transmission rates. In this paper, for a field experiment of a wireless communication system for unmanned exploration vehicles in extreme cold regions, four 12dBi omni-directional antennas and a transmit power of 20dBm were used at an outdoor transmission distance of 52km, and we measured the RF received field strength -86 ~ -88 dBm and data rates of 10.8 ~ 11 Mbps were obtained.**

Keywords*— Beamforming, line of sight MIMO, long range communication, spatial diversity gain

I. INTRODUCTION

Existing low-power long-distance communication (LPWA) services have built their own low-power IoT network using unlicensed carrier frequencies rather than using mobile communication networks. In the case of SigFox, which uses unlicensed bands, commercialization is continuously expanding. Looking at the IoT (Internet of Things) network service using SigFox, up to 140 messages can be sent and received per device per day, and the message size is generally around 12 bytes [1]. In the case of LoRa WAN, technology development is being led by the LoRa Alliance, a global alliance, and mobile communication operators such as Switzerland's Swisscom and Dutch operator KPN, as well as major hardware and software companies, have participated in the alliance [2].

Developed to overcome the limitations of technology using existing unlicensed frequency bands, LTE-MTC considers key requirements such as long-distance low-power communication, low power consumption, low construction cost, low-cost equipment, large-scale terminal access, and stable coverage. Reference [3] reports the increase in IoT products of Cat-NB1 and Cat-M1 products, as well as the release of chipsets and modules compliant with 3GPP Rel-14 Cat-NB2. However, end devices of NB-IoT consume additional power due to synchronous communication and QoS processing, and OFDM/FDMA access schemes consume more power [4].

Beamforming technology can be used as a method to increase spatial diversity gain, and a recent study is hierarchical beamforming to reduce antenna training time and overhead [5]. Under the assumption that the same phase shift is applied to each analog subarray to solve the hierarchical beam search complex problem, a cross-correlation-based beam search technique is proposed in the subarray [6], and a subspace projection-based AoA (Angle of Arrival) technique suggested [7]. In the subarray-based method, the phase ambiguity of AoA detection is resolved through additional training symbols, post-processing in the frequency domain, noise subspace projection, and iterative update [8].

In this paper, we analyze the spatial diversity of the MXN line-of-sight multiple input/output communication channel in an interference channel environment without scattering of electromagnetic waves, and simulate the 2X2 LOS-MIMO channel as an example of the possibility of achieving maximum diversity gain in communication over 50 km. and conducted a field test on the related beamforming technology.

II. SPATIAL DIVERSITY GAIN ANALYSIS IN LOS-MXN MIMO CHANNELS

In an interference channel environment with almost no scattering of electromagnetic waves, when N transmitters attempt simultaneous transmission to M receivers, the transmitters cause interference to each receiver. The relationship causing interference between the transmission x and the reception y is represented by the area of the reception signal $\mathbf{y} \in C^{2\times1}$ as shown in Equation (1).

$$\mathbf{y} = \mathbf{H}\mathbf{x} + \mathbf{N_0}. \quad (1)$$

Here, $\mathbf{H} \in \mathbf{C}^{2\times2}$: channel vector, $\mathbf{x} \in C^{2\times1}$: transmission signal vector, $\mathbf{N_0} \in C^{2\times1}$: AWGN (Additive White Gaussian Noise).

In a line-of-sight wireless communication system, the channel environment to transmit data through multiple transmit/receive antennas can be set as shown in Figure 1.

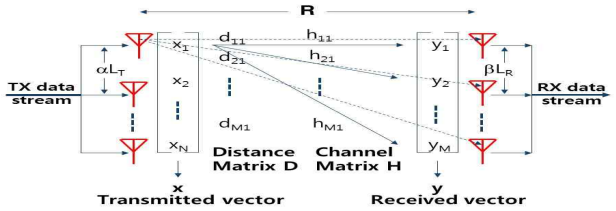


Figure 1. Parameter relationship between transmitting and receiving antennas in a line-of-sight wireless communication system that adjusts the antenna spacing.

Where, R is the effective straight-line range between transmitter and receiver, L_t: transmitting antenna separation

distance, L_r: receiving antenna separation distance, $\alpha>1$: Tx antenna separation distance control factor, and $\beta>1$: receiving antenna separation distance control factor. The (m, n)-th complex channel gain h_{mn} in the channel matrix **H** can be showed by Equation (2).

$$h_{mn} = \frac{\lambda}{4\pi d_{m,n}} e^{-j\frac{2\pi}{\lambda}d_{m,n}} \quad (2)$$

The communication distance d_{mn} from the m-th receive antenna to the n-th transmit antenna is expressed as the following (3) when λ is the wavelength of the carrier radio frequency.

$$d_{mn} \cong R\left\{1+\frac{1}{2}\left\{\frac{(m-1)\alpha L_t+(n-1)\beta L_r}{R}\right\}^2\right\} \quad (3)$$

When the channel capacity ***C*** of the line-of-sight communication system is the transmission power E_x and the received noise N_0 is, it can be expressed as the following equation [9].

$$C = \log_2 \det\left(I_M + \frac{E_x}{NN_0} HH^H\right) \quad (4)$$

Here, in the case of MXN LoS MIMO channels, HH^H can be expressed as a product of a channel matrix H and a Hermitian H^H matrix. At this time, the spatial diversity gain SDG is defined as follows when SNR $\rightarrow \infty$ [10, 11].

$$SDG = \lim_{SNR\to\infty} m\left(1-\frac{\log\det\left(I_M+\frac{SNR}{n}HH^H\right)}{\min(m,n)\log SNR}\right)n\left(1-\frac{\log\det\left(I_n+\frac{SNR}{n}HH^H\right)}{\min(m,n)\log SNR}\right) \quad (5)$$

For example, considering 2X2 LoS-MIMO channels, SDG from Equation (5) can be approximated by Equation (6) [10].

$$SDG \cong \left[2-\log 4\left\{1-\left(\cos 2\pi\frac{\alpha^2L^2}{\lambda R}\right)^2\right\}\right]^2 \quad (6)$$

III. 2×2 LOS-MIMO Channel Capacity Simulation

We set the transmission distance up to 50 km, the carrier frequency 5 GHz (λ=6 cm), and the antenna separation distance as $(\alpha L)^2$= 250. At this time, the distance between the transmitting and receiving antennas is a virtual distance modified by an adjustment factor. In this 2x2 LOS-MIMO communication system environment, spatial multiplexing gain for communication distance was simulated. The results of the simulation of spatial diversity gain for long-distance communication systems are shown in Figure 2.

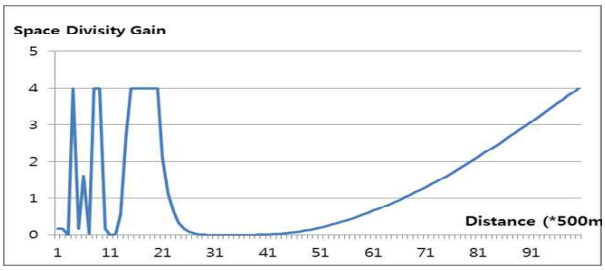


Figure 2. Space diversity gain simulation results for distance in a line-of-sight wireless communication system up to 50 km.

IV. Field Experiments on Long-Range 4x4 LOS-MIMO Beamforming Scheme

The field experiment environment is set with a LoS transmission distance of 50 km, a carrier frequency of 5.8 GHz, and the distance between the transmit and receive antennas is set as phase factors that adjust the delay length of the transmit/receive signals input/output to each antenna. The experimental site is located 52km between Deokjeok-do and Seokmo-do, and is shown in Figure 3 below.

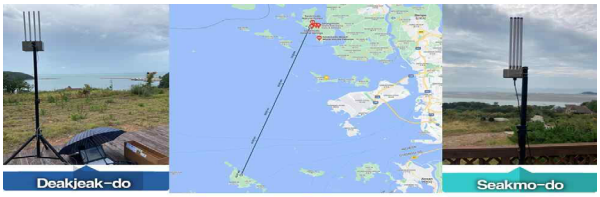


Figure 3. Experiment site between Deokjeok-do and Seokmo-do to measure 52km communication performance.

Four 12dBi omni-directional antennas were applied to each of the transmit and receive antennas. Figure 4 shows the results of a 52km long-distance transmission field test. The vertical axis represents the transmission rate and the horizontal axis represents the time. At this time, in Figure 4, it was confirmed that the transmission rate was 10.8 to 11 Mbps at the RF receiving field strength of -86 to -88 dBm, achieving more than 10 Mbps.

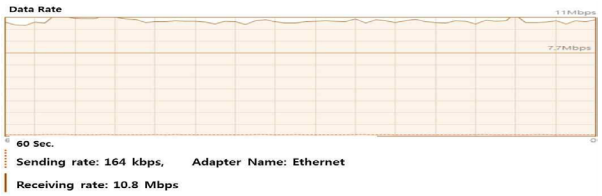


Figure 4. Field test performance results at a transmission distance of 52 km.

Acknowledgment

This paper is a research conducted with the support of the Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the government (Ministry of Science and ICT) in 2023. [No. 2021-0626, Development of Polar Region Communication Technology and Equipment for Internet of Extreme Things (IoET)].

References

[1] C. Gomez, J. C. Veras, R. Vidal, L. Casals and J. Paradells, "A Sigfox Energy Consumption Model," Sensors (Basel, Switzerland) - MDPI, v. 19, pp. 681-694, 2019.

[2] T. Bouguera, J. Diouris, J. Chaillout, R. Jaouadi, and G. Andrieux, "Energy Consumption Model for Sensor Nodes Based on LoRa and LoRaWAN." Sensors (Basel, Switzerland), vol. 18, no. 7, 2018..

[3] S.-M. Oh and J. Shin, "An efficient small data transmission scheme in the 3GPP NB-IoT system," IEEE Communications Letters, vol. 21, no. 3, pp. 660–663, 2017.

[4] K. Mekki, E. Bajic, F. Chaxel, and F. Meyer, "A comparative study of LPWAN technologies for large-scale IoT deployment," ICT Express, 2018.

[5] S. Noh, J. Song, and Y. Sung, "Fast Beam Search and Refinement for Millimeter-Wave Massive MIMO Based on Two-Level Phased Arrays," IEEE Trans. Wireless Commun., vol. 19, no. 10, pp. 6737–6751, Oct. 2020.

[6] J. Zhang, X. Huang, V. Dyadyuk, and Y. Guo, "Massive hybrid antenna array for millimeter-wave cellular communications," IEEE Wireless Commun., vol. 22, no. 1, pp. 79–87, Feb. 2015.

[7] F. Shu et al., "Low-complexity and high-resolution DOA estimation for hybrid analog and digital massive MIMO receive array," IEEE Trans. Commun., vol. 66, no. 6, pp. 2487–2501, Jun. 2018.

[8] C. Qin, J. A. Zhang, X. Huang, and Y. J. Guo, "Virtual-subarray-based angle-of-arrival estimation in analog antenna arrays," IEEE Wireless Commun. Lett., vol. 9, no. 2, pp. 194–197, Feb. 2020.

[9] E. Telatar, "Capacity of multi-antenna Gaussian channels," Eur. Trans. Telecommun., vol. 10, no. 6, pp. 585-595, Nov. Dec. 1999.

[10] W. Y. Lee, K. Kim, Y. Kim, and D. Kwon, "A Study on Spatial Multiplexing Gain in LOS-2X2 MIMO Intelligent Reconfiguration Channel Environment," in Proc. 12th Int. Conf. on ICT Convergence (ICTC), pp. 590-592, Oct. 2021.

[11] L. Zheng and D. N. C. Tse, "Diversity and Multiplexing: A Fundamental Tradeoff in Multiple-Antenna Channels," IEEE Trans. Inform. Theory, vol. 49, no. 5, pp. 1073–1096, May 2003.

Development of an Wartime-Level Target Tracking System using a Kalman Filter

Hongki Kim
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
khgui1111@navy.ac.kr

Taeho Kim
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
xogh0627@navy.ac.kr

Youngjun Lee
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
yeongjun0105@naver.com

Sangje Oh
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
osjae0430@navy.ac.kr

***Abstract*— In this paper, we introduce a target tracking method at the exhibition level that employs a standard Kalman filter in a tracking radar system to predict the trajectory of target loss. The proposed system aims to enhance tracking probability by reducing tracking signal noise and narrowing the interval between radar measurement values through computer vision-level analysis. Furthermore, the accuracy of the results can be enhanced by predicting and generating the target disappearance segment using the predicted value from the Kalman filter and linearizing the target's continuous position.**

Keywords—Wartime-level, target tracking system, kalman filter, predicting and generating the target disappearance section

I. INTRODUCTION

Modern naval warfare has evolved to prioritize long-range missile-centered engagements, prompting navies worldwide to develop strategies and tactics focused on acquiring and engaging targets beyond the enemy's detection range. The Republic of Korea Navy, for instance, has equipped its combat ships, including the Chungmugong Yi Sun-shin destroyer (DDH-2), with anti-aircraft and ship radars for long-range attacks. To ensure improved anti-aircraft detection capabilities, the DDH-2 class ships are equipped with three radars: AN/SPS-49, MW-08, and STIR-240, surpassing the capabilities of their predecessors. However, with these radars being over 20 years old, they face challenges in effectively countering the latest weapon systems. As part of the DDH-2 upgrade plans, the installation of a domestic combat system optimized for the MW-08 radar is being considered.

Tracking algorithms employed by these radars process information received through antennas, predict the target's next location, and provide coordinates for guided weapons interception. However, the MW-08 radar operates mechanically, allowing only discrete tracking, and as the interval between measurements increases, the probability of false detection rises.

Recognizing this issue, this paper aims to address the problem by reducing tracking signal noise and increasing tracking probability through computer vision-level analysis [1]. Additionally, the predicted value of the Kalman filter is utilized to generate an expected target path, enabling continuous tracking even in the event of target loss. The proposed system is illustrated in the flowchart above.

The primary objective is to enhance the effectiveness and combat capabilities of combat ships by supplementing the existing tracking methods employed by older ships. In the current tracking system, the exhibitor's screen presented to radar operators showcases measurements processed using internal noise cancellation technology. However, once developed, this system lacks frequent upgradeability due to its complex operational requirements, and tracking is interrupted when direction cannot be determined due to weather conditions or wave interference.

To overcome these challenges, this paper proposes a method to analyze the screen itself based on the Kalman filter [2]-[3]. This approach involves noise removal, linearization of the target's continuous position, and further enhancement of result accuracy. By utilizing the predictive capabilities of the Kalman filter, the lost portion of the target's trajectory can be compensated. Additionally, error resonance is mitigated by eliminating the error resonance value. This method aims to improve the accuracy of older radars and increase the precision of guided weapons.

II. THEORETICAL BACKGROUND

A. Research on tracking algorithms using existing extended Kalman filters and particle filters

Step 1 : Selection of standard Kalman filters

Traditional tracking radars analyze measurements by inputting them into built-in filters. Among them, Extended Kalman Filter (EKF) is suitable for non-linear model tracking similar to motion on the inertial coordinate system of an object. This is why EKF is also used in Aegis radar tracking systems that use phase array antennas or mechanical radar-based Korean combat systems. However, since this filter operates nonlinearly, the amount of computation increases exponentially in the operation of a high-dimensional vector. Anti-aircraft targets that require confirmation of three-dimensional orientation are

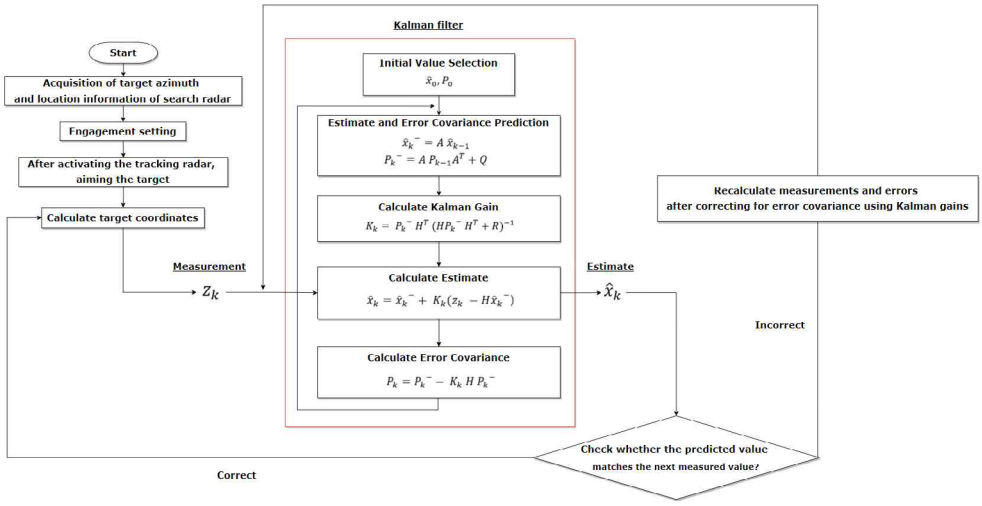


Fig. 1. Flowchart of proposed wartime-level target tracking system using a kalman filter

particularly computational. Using EKF in this system, which applies filters every frame, makes real-time calculations difficult, resulting in tracking one step behind, and multiple targets cannot be tracked at the same time.

Particle Filter, one of the filters commonly used with EKF, was also considered [4]. However, the computational complexity of this filter was also great, and it was excluded because it was difficult to solve the problem of cursing at the dimension when analyzing the location over time. Since the standard Kalman filter operates linearly, there is a probability of error compared to other filters in the inertial coordinate system. Nevertheless, this filter was selected based on the fact that the computational speed was fast, and that most of the target motion did not deviate significantly from the linear. In addition, it was judged that there was no major problem in the application of the linear filter because the nonlinear filters already built in the existing system estimate the motion of the target close to the linear. The pre-processing process of the image and the change of existing parameters performed to reduce the tracking error of the Kalman filter will be described later [4].

Step 2: Image preprocessing and changing filter parameters

The process of preprocessing an image so that the Kalman filter is easy to calculate is essential in the above system [1]. Since filters are applied once per frame, multiple frames per second fall short of the real-time tracking specification [3]. In addition, it is easy to find a tracking target only when the contrast between the background and the target in the video is clear. In the sample image, the target was set to green on a black background to increase the contrast. Filter parameters are largely related to the preprocessing of images and are greatly affected by the set values due to the nature of a linearly operated system. Motion model is consist of Constant Velocity and Constant Acceleration.

As one of them, you can choose a state closer to the motion of the object. The sample image was set to Constant Velocity because it was close to constant velocity motion. The location of the object can be specified when the filter is substituted with 'initial location', and the code recognizes the target on its own and automatically specifies it through the command.

B. Compare code parameters with Kalman filter variables

The value that can be determined in advance in the Kalman filter algorithm is a system model consisting of A, H, Q, and R.

Table 1. Compare code parameters with Kalman filter variables

A	System Matrix	n*n matrix
H	Output Matrix	m*n maxtrix
Q	System noise covariance matrix	n*n diagnal matrix
R	Measurement noise covariance matrix	m*m diagonal matrix
x_k	state variable	n*1 column vector
v_k	Measurement	m*1 column vector

Table 1. shows four variables constituting the system model and two initial values. Variables are values that the user determines in advance when configuring the Kalman filter, and are calculated with the initial value to predict the estimated value and error covariance. Since this code also operates as a Kalman filter, we will look at the corresponding part in Table 1.

The 'Motion Model' refers to the state of motion of the system, so it corresponds to A. 'Initial Location' corresponds to because it specifies the initial position of the target. 'Initial Estimate Error' is an H that defines the relationship between and . 'Motion Noise' and 'Measurement Noise' are represented by Q and R, respectively. 'Segmentation threshold' is a parameter related to the preprocessing of an image, and the sensitivity of the filter can be set through this value, but the value adjustment is unnecessary because the contrast of the image itself is clear.

III. System Architecture and Methodology

A. Implementation of single target tracking in video

An experiment was constructed to prove the performance of the system presented in this study [2]. First, a circuit and a radar screen for acquiring image samples were implemented. For security reasons, it was not possible to directly obtain images detected by the radar during the operation, so an ultrasonic radar was produced using a microcomputer. An ultrasonic sensor was attached on top of the servomotor to detect the 180° orientation through mechanical rotation. Next, the information detected by the radar was implemented as a display screen through processing. Through Fig. 2, the information detected by the radar is displayed in a green circle as shown in Fig. 3. It was set to a diameter suitable for recognition by the Kalman filter, and two tracks per second were recorded for high accuracy.

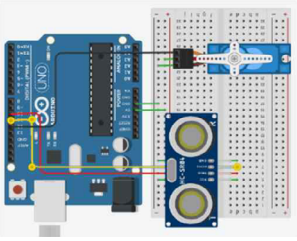


Fig. 2. Ultrasonic radar realized with MC circuit

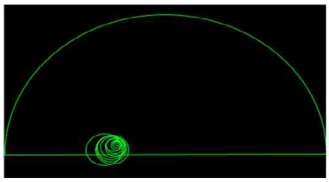


Fig. 3. Radar screen displayed in processing

B. Predict and track the target's movement paths

Based on the images obtained above, target tracking was performed. We designed a code that tracks a single target in the screen using MATLAB, and we want to find out how well it tracks consecutive targets.

Step 1: Synthesis of target movement frame

Fig. 4 (a) is the result of combining all the frames by analyzing the screen inside the code, and Fig. 4 (b) shows the path in red by tracking the target by each frame. This algorithm is performed automatically, and it can be seen that it appears as a smooth straight line in contrast to Fig. 4 (a). Next, an experiment was conducted to detect the predicted path when the target currently being tracked was lost.

Step 2: The movement path of the target tracked by the Kalman filter

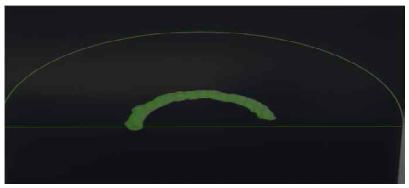
The process of preprocessing an image so that the Kalman filter is easy to calculate is essential in the above system. Since filters are applied once per frame, multiple frames per second fall short of the real-time tracking specification. In addition, it is easy to find a tracking target only when the contrast between the background and the target in the video is clear. In the sample image, the target was set to green on a black background to increase the contrast.

Step 3: Synthesis of target movement frame

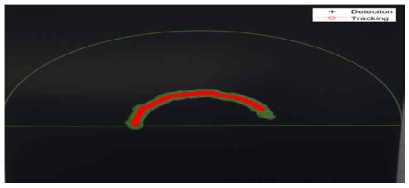
Fig. 4 (c). shows the state where the value input to the radar during target tracking is 0. Existing radars are delivered as they are on this screen. Fig. 4 (d). shows prediction of the vanishing target of the Kalman filter. However, using the program presented in this paper, it can be confirmed that a red line is drawn within the path where the target is lost.

Step 4: The movement path of the target tracked by the Kalman filter

This is the mechanism by which the Kalman filter predicts in-between paths. Through high precision, the predicted path can be displayed as a reference material, and based on the path, tracking continues without interruption.



(a) Synthesis of target movement frame



(b) Movement path of the target tracked by the Kalman filter

(c) Synthesis of target movement frame

(d) Movement path of the target tracked by the Kalman filter

Fig. 4. Result of predict and track the target's movement paths

IV. Results and Discussion

In this part, Matlab based code simulates a target tracking method at the exhibition level that can predict the loss path using a standard Kalman filter in a tracking radar system. This code demonstrates how a Kalman filter can be used to estimate the state of a tracked object based on noisy observations, providing a more accurate prediction of the object's path.

Step 1: System Initialization and Model Definition

The code starts by initializing the initial state of the tracked object and the initial covariance matrix. The system model is defined using the state transition matrix (A) and the observation matrix (H). The state transition matrix describes how the state of the system evolves over time, while the observation matrix relates the measurements to the state variables.

Step 2: Process and Observation Noise Covariance

The covariance matrices for process noise (Q) and observation noise (R) are defined. These matrices represent the uncertainty or noise in the system and observations, respectively.

Step 3: True Path and Observation Generation

A set of true states for the tracked object is generated using the generate true states function. This function uses the system model, initial state, and process noise covariance to generate the true states. Observations are generated by adding observation noise to the true states using the generate observations function. This function uses the observation matrix, true states, and observation noise covariance to generate the observations.

Step 4: Kalman Filter Initialization

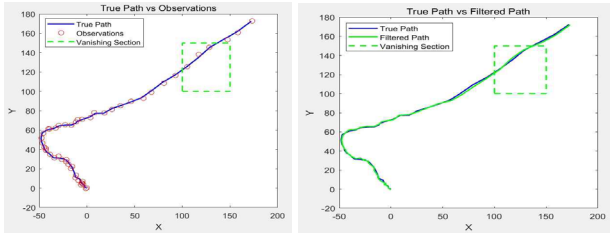
The filtered states and covariances are initialized with zeros. The filtered states will contain the estimated states of the tracked object at each time step, and the filtered covariances will contain the corresponding covariance matrices.

Step 5: Kalman Filter Update

The main loop applies the Kalman filter to estimate the states of the tracked object based on the observations. At each time step, the filter goes through two steps: prediction and update.

a. Prediction Step: The prediction step estimates the next state based on the previous filtered state. The predicted state is obtained by multiplying the state transition matrix (A) with the previous filtered state. The predicted covariance is obtained by propagating the previous filtered covariance using the state transition matrix and adding the process noise covariance (Q).

b. Update Step: The update step incorporates the new observation into the prediction to improve the state estimate. The innovation is the difference between the observation and the predicted state transformed by the observation matrix. The innovation covariance is obtained by propagating the predicted covariance using the observation matrix and adding the observation noise covariance. The Kalman gain is calculated by multiplying the predicted covariance with the transpose of the observation matrix and scaling it by the inverse of the innovation covariance. Finally, the filtered state is updated by adding the product of the Kalman gain and the innovation to the predicted state, and the filtered covariance is updated using the Kalman gain.



(a) True Path vs Observations (b) True Path vs Filtered Path

Fig. 5. Result of proposed wartime-level target tracking system

V. CONCLUSION

The system proposed in this paper utilizes radar measurements to generate images and employs Kalman filters for noise cancellation and prediction paths. The use of standard Kalman filters ensures low computational load and enables real-time target processing. This simple yet powerful program offers a potential upgrade for the aging combat systems currently in use by the Korean Navy, and it can serve as an additional system when equipped with a new combat system. The improved tracking probability provided by this system holds significant benefits for the performance of combat ships in modern battlefields where anti-aircraft detection capabilities are crucial.

However, as the development of this system is still in its early stages, several tasks need to be addressed. Firstly, since re-analysis is conducted based on radar-generated images, inherent time delays occur. In situations where guided weapons approach the self-ship at high speeds, there is a possibility that the Kalman filter may encounter a hit before the analysis is completed. Secondly, the system is currently unable to detect multiple targets within the code progression. To devise a multi-object tracking system, the movement paths of individual targets need to be separated, and unless specified directly by the operator, machine learning techniques must be employed, further complicating the implementation process.

REFERENCES

[1] A. Yilmaz, O. Javed and M. Shah, "Object Tracking: A survey," ACM Computing Surveys, Vol.38, Issue. 4, pp. 13-es, 2006.

[2] S. V. Kothiya and K. B. Mistree, "Notice of Removal: A review on real time object tracking in video sequences," 2015 International Conference on Electrical, Electronics, Signals, Communication and Optimization (EESCO), Visakhapatnam, India, 2015, pp. 1-4, 2015.

[3] P. R. Gunjal, B. R. Gunjal, H. A. Shinde, S. M. Vanam and S. S. Aher, "Moving Object Tracking Using Kalman Filter," 2018 International Conference On Advances in Communication and Computing Technology (ICACCT), Sangamner, India, pp. 544-547, 2018.

[4] S. Mahfouz, F. Mourad-Chehade, P. Honeine, J. Farah, and H. Snoussi, "Target tracking using machine learning and Kalman filter in wireless sensor networks," IEEE Sensors Journal, vol.14, no.10, pp.3715-3725, 2014.

Investigation of Naval Application Strategies for Underwater Laser Communication

Donghyun Kwak
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
wisdom0328@navy.ac.kr

Seungil Jeon
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
117jeon @navy.ac.kr

Jung-rak Choi
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
wjdgkr465900@navy.ac.kr

Minseok Han
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
mshan1024@navy.ac.kr

***Abstract*— This paper investigates the naval application of laser communication as a potential replacement for traditional sound wave communication in underwater environments. A laser transmitter/receiver system utilizing Arduino was configured, and a water tank experiment was conducted to verify the feasibility of communication in diverse underwater conditions. Building upon these findings, our analysis emphasizes applications in naval tactical communication, remote sensing, and underwater drone control. Furthermore, we propose an improvement plan to address current technical limitations and enhance performance in this field.**

Keywords—Underwater laser communication, changing underwater environment, applying MATLAB & Arduino to naval operations

I. INTRODUCTION

In recent years, there has been a growing demand for underwater marine exploration, underwater environment monitoring, navigation, and naval operations, leading to active research in the field of underwater communication technology. Conventional underwater acoustic communication technologies have made progress in wireless communication, but they still face limitations such as low bandwidth, transmission bandwidth constraints, high latency, and vulnerability to security breaches. As a result, laser optical communication technology has emerged as a promising alternative in underwater communication.

Underwater laser communication offers numerous advantages, including fast transmission speed, high bandwidth, low latency, and strong security [1]. These advantages make underwater laser communication highly valuable for enhancing the efficiency and stability of naval operations. This study aims to verify the feasibility of laser communication in diverse underwater environments through experiments and focuses on its potential applications in various naval environments and missions. To comprehend the benefits of underwater laser communication for the navy, we compare and analyze it against conventional acoustic wave communication technology, exploring its utilization in operational fields such as naval tactical communication, remote sensing, and underwater drone control. This paper presents the future prospects of underwater laser communication technology and discusses necessary improvements and requirements to overcome current technological limitations and enhance overall performance. The objective is to maximize the effective implementation of underwater laser communication in naval operations, ultimately improving power and information delivery capabilities.

In conclusion, this paper highlights the benefits derived from the adoption of underwater laser communication and explores its potential as an innovative technology capable of replacing conventional sonic communication technology. Through this research, valuable insights are provided regarding the design and utilization of efficient underwater laser communication systems in areas such as naval tactical communication, remote sensing, and underwater drone control. The findings of this study contribute to enhancing the efficiency and safety of naval operations.

II. THEORETICAL BACKGROUND

Underwater laser communication is a communication method that transmits information using a laser in water, and has many advantages over underwater acoustic communication that is mainly used for underwater communication.

When radar sizes are the same, underwater laser communication transmits data over ultra-high frequencies higher than underwater acoustic communication ensuring fast data transmission speed. In Table 1, underwater acoustic communication shows a data transmission speed of Kbps, while underwater laser communication shows a data transmission speed of Gbps.

Unlike other communication, underwater laser communication is capable of real-time communication. This is a very important factor in marine research and military use [2]. In addition, since it is below the surface of the water, it is not affected by weather changes, and stable communication is possible, such as not affecting other underwater equipment.

Table 1. Comparison of underwater wireless communication technologies

Parameter	Acoustic	RF	Optical
Attenuation	Distance and frequency dependent (0.1–4 dB/km)	Frequency and conductivity dependent (3.5–5 dB/m)	0.39 dB/m (ocean) 11 dB/m (turbid)
Speed	1500 ms^{-1}	2.3×10^8 ms^{-1}	2.3×10^8 ms^{-1}
Data Rate	kbps	Mbps	Gbps
Latency	High	Moderate	Low
Distance	more than 100 km	≤10 m	10–150 m (500 m potential)
Bandwidth	1 kHz–100 kHz	MHz	150 MHz
Frequency Band	10–15 kHz	30–300 MHz	5×10^{14} Hz
Transmission Power	10 W	mW–W	mW–W

III. System Architecture and Methodology

A. Hardware design

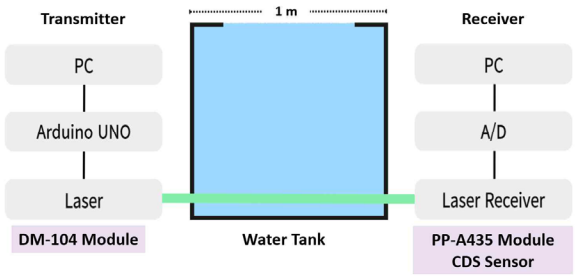
Experiments on underwater laser communication were conducted separately for each communication process. In the first experiment, when data is transmitted through the laser, the desired message is transformed into data and transmitted, and it is received and confirmed to be converted into the correct message. In the second experiment, it was confirmed whether the communication works in the underwater situation, In the third experiment, the limits of laser communication in underwater conditions were measured and confirmed.

To implement the above experiment, the laser was emitted using the DM-104 Arduino laser module. This module transmits data by emitting a 650nm red laser with a voltage of 5V. As the receiving device, PP-A435 module and CDS light sensor module were used. The PP-A435 module is a laser receiving sensor, and the receiving unit receives the laser emitted from the laser module and analyzes the data contained in it. The CDS illuminance sensor module was used as a measurement tool to check the limits of laser communication in various underwater situations by measuring the intensity of light in underwater communication situations.

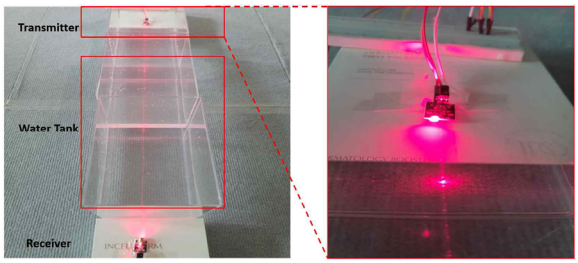
B. Software design

The DM-104 module converts the 'Red Sub Found' message converted into binary code into a laser signal and transmits it at regular intervals. Characters are converted to 1s and 0s, and output values of 1s and 0s are delivered at intervals of 10 microseconds. With the input signal, the module emits laser and transmits data. In order to increase communication efficiency, Arduino's SPI parallel communication, which transmits several bits at the same time, is applied rather than serial communication, which transmits several bits one by one. Since parallel communication transmits bits simultaneously, it greatly reduces transmission time and increases transmission speed, enabling efficient communication.

The PP-A435 module receives binary messages sent to the laser. The laser signal is received through the receiver, and the received binary data is converted into text and decoded into a message. The time interval of the receiver is set to 10 microseconds like the transmitter to minimize the occurrence of errors. The receiver receives the transmitted binary code and rearranges it back into a string. Since the string of transmitted data is transmitted at once, the decoded string is displayed as 'Red Sub Found' in the form of 'Received Message: ' in the result window.



(a) System architecture



(b) Experimental Setup

Fig. 1. System architecture and experimental setup

IV. RESUTLS AND DISCUSSIONS

```
Command Window
  turbidity: 5     suitable lasser wave length: 532nm
  turbidity: 20    suitable lasser wave length: 633nm
  turbidity: 55    suitable lasser wave length: 785nm
  turbidity: 180   suitable lasser wave length: 1064nm
```

Fig. 2. Optimal wavelength for turbidity

The experiment was conducted in an empty water tank, a water tank with water, a situation seawater, and a situation in which impurities were added to seawater.

According to the results, laser communication was possible in the situation of empty water tanks, water tanks only, and seawater. However, it was confirmed that communication was impossible in the situation where impurities were added, and the number of communication failures increased as the concentration of impurities increased. More accurately, the CDS level in the empty tank is about 6 lux, and the CDS level increased in the case of seawater, and when impurities were added. The increase in CDS levels means that the substances in the water tank interfered with the laser's progress and made communications difficult.

In particular, when comparing dust and plastic debris, in can be seen that each level is different in impurities. In Table 3, communication was more difficult than plastic debris in the dust situation. Table 3 shows that the communication state is different even though the same amount of impurities are added. The CDS of the dust was 270 lux at 7g and the plastic debris were 131 lux. At this time, communication was not possible in the soil and it was possible in plastic debris. When confirming in eyesight, the dust situation was no reached by scattering in the middle of the laser process due to the turbidity of the water, However, in plastic debris, turbidity was not as bad as dust, and it was confirmed that the laser reached the receiver. In the experiment through this, it was confirmed that turbidity was the factor that most influenced laser communication.

It proposes a method of varying the wavelength if the laser as a method for communication even in murky situations. There is already a method of calculating turbidity by varying the laser wavelength. Accordingly, it is possible to obtain a laser wavelength suitable for each turbidity. A code that presents the optimal wavelength of the laser for turbidity was written through the Matlab code, and the execution results are shown in Fig. 2.

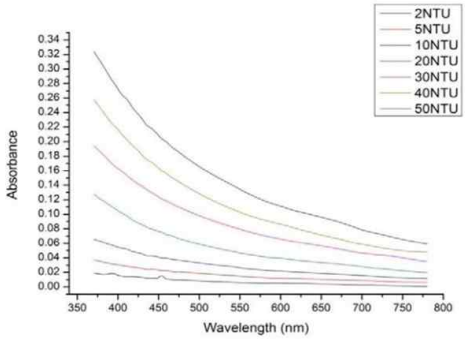


Fig. 3. Relation between turbidity and laser wavelength [3]

Table 2. Laser communication according to the environment

Experiment	1st	2nd	3rd
empty tank	o	o	o
basic water	o	o	o
sea water (120g of salt)	o	o	o
sea water (240g of salt)	o	o	o
sea water (120g of salt + dirt 3.5g)	x	o	x
sea water (120g of salt + dirt 7g)	x	x	x
sea water (120g of salt + plastic debris 3.5g)	o	x	o
sea water (120g of salt + plastic debris 7g)	x	o	x

Table 3. Comparison of measured values of CDS ambient light sensor according to environment

Experiment	1st	2nd	3rd	average
general situation	698	702	699	699.67
empty tank	6	5	7	6
basic water	9	12	11	10.67
sea water (120g of salt)	23	26	24	24.3
sea water (240g of salt)	30	29	32	30.3
sea water (120g of salt + dirt 3.5g)	180	175	177	177.3
sea water (120g of salt + dirt 7g)	270	269	273	270.76
sea water (120g of salt + plastic debris 3.5g)	88	91	92	90.3
sea water (120g of salt + plastic debris 7g)	134	129	132	131.67

As shown in Fig. 2, the wavelength of the laser increases as the turbidity increases. This phenomenon can be explained in Fig. 3. Fig. 3 shows a graph extracted from a paper explaining the relationship between laser and turbidity. In this graph, the laser shows a phenomenon in which the absorption rate decreases at a high turbidity as the wavelength increases. As the turbidity increases, the wavelength of the laser increases, and the wavelength value of the paper and the laser wavelength value of each turbidity presented in this study almost match, showing that the code presented and the result value are reliable [4].

V. Utilization plan for navy application

Based on the results of the experiment, we will find out what advantages laser underwater communication and in what direction it is applied to the Navy.

Submarine communication is capable of communication between submarine-warship, submarine-submarine, and submarine-aircraft, and has high bandwidth and reliability, enabling stable communication between the inside and outside of the submarine. In particular, even in deep-water areas, stable communication is possible compared to sound wave communication, and it has the advantage of being able to transmit a large number of sensor data quickly. In fact, the U.S. Navy had succeeded in laser communication between aircraft and submarines. In this experiment, the results showed that the laser of blue light was the most suitable in the deep sea, and the data transmission success rate was more than 80% using a laser of about 532nm wavelength [5]. Such submarine communication technology has also been actively introduced by the Korean Navy, and the importance of submarines is currently drawing attention, and it is one of the technologies that can be developed with great power.

In marine operations, communication between underwater drones and robots is expected to play an important role as unmanned technologies such as drones and robots are further developed in the future. Multiple operations using this unmanned technology are more efficient than a single operation, and laser communication between drones and robots more quickly and accurately, and transmit/receive a large number of data.

This paper deals with underwater sensor data communication and communication objects such as submarines and underwater drones(AUVs), and proposes underwater reconnaissance and threat detection methods using them. Underwater sensor data communication can be used as an alarm system by detecting collisions, sinks, and other hazards underwater by utilizing submarine sensor installation and laser communication. These technologies will greatly help prepare for submarine threats from North Korea and neighboring countries and take control of the sea.

VI. Conclusion

In this paper, we investigated the "Naval Application of Laser Underwater Communication Utilizing MATLAB & Arduino". We examined the potential effectiveness of laser communication underwater compared to traditional methods and explored its possible applications in the navy.

Throughout our research, we successfully developed a laser-based underwater communication system using MATLAB and Arduino. Our findings revealed that turbidity had the most significant impact on communication performance. To address this issue, we experimented with varying the laser wavelength and supported our approach with evidence from relevant studies. Based on this information, we developed MATLAB code that recommended different wavelengths based on turbidity levels, offering a potential solution.

The laser communication system showcased higher data transmission rates compared to current acoustic communication methods, providing increased bandwidth and reduced latency underwater. Moreover, laser communication demonstrated remarkable resistance to interference, enhancing communication stability in underwater environments. This laser-based underwater communication system holds promise for diverse naval applications. For instance, it could greatly enhance the efficiency and precision of submarine communication, seabed sensor networks, and unmanned underwater exploration equipment. This advancement would significantly improve the speed of information collection and transmission, delivering substantial benefits in future operational scenarios.

In conclusion, our study has demonstrated that laser-based underwater communication outperforms conventional acoustic communication methods, offering significant value in naval operations. As this technology continues to evolve, we anticipate further advancements in naval communication systems, ultimately contributing to comprehensive maritime security and operational support.

References

[1] S. Kim, “A Study on Sloped Pulse Position Modulation in Underwater Laser Wireless Communication,” Master’s Thesis, Korea Maritime & Ocean University, 2018.

[2] D. Kwon, “Underwater Wireless Optical Communication,” EIRIC, vol.45, no.5, pp.49-56, 2018.

[3] Y. Wen, Y. Hu, and X. Wang, “Application of a colorimer for Turbidity measurement,” Journal of Physics: Conference Series 679, pp.1-4, 2015.

[4] Y. Abueejela, A. S. Guinauim, and A. Algitta, “Measuring Turbidity of Water Based Photo Voltaic Cell and Laser Light,” JEEEIT Transactions,.vol.1, no.1, pp.49-54, 2020.

[5] Y. Kim and J. Choi, “Blue Green Lasers and Their Application to Underwater Communication,” Journal of korean institute of communications and information sciences, vol.16, no.12, pp.1211-1218, 1991.

Car Brand & Model and Body Part Detection Algorithms based on CNN and Transfer Learning for Repair Cost Estimation

Man Zhang
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
mandyz0216@gmail.com

Min Htet Thar
Research Team
EasyGeo Co.
Busan, Republic of Korea
htatmin332@gmail.com

Dong Myung Lee
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
dmlee@tu.ac.kr

***Abstract*—In this paper, the car brand & model and body part detection algorithms based on convolutional neural network (CNN) and transfer learning for repair cost estimation is proposed. The algorithm consists of five models: car brand & model detection module, car body parts detection module, car damage classification module, damaged car part calculation module and repair cost estimation module. The study is expected to help auto insurance companies simplify their cost estimation process, identify car brands and models, and classify car damages.**

Keywords—Car Damage, Object Detection, Car Brand & Model Detection, Image Classification, CNN, Transfer Learning

I. INTRODUCTION

In recent years, with the development of computer vision and artificial intelligence technology, this technology has been applied to insurance companies [1]. In this paper, an algorithm to estimate the cost of car repair by using convolutional neural networks (CNN) and transfer learning is proposed in this paper. This algorithm consists of five modules: car brand & model detection module, car body parts detection module, car damage classification module, damaged car parts calculation module and repair cost estimation module.

II. PROPOSED ALGORITHM

The basic design ideas of proposed algorithm mainly uses CNN, transfer learning, car brand and model dataset and vehicle dataset. Datasets are built by sites such as Kaggle, COCO dataset, and others such as the Sandford Cars dataset [2-3]. Through the training model, the car brand and model was identified and classified, the body parts were identified, and the damaged parts were classified by percentage calculation. Through the comprehensive analysis of a number of index results, the maintenance cost of damaged car is estimated.

The proposed algorithm consists of five modules such as car brand & model detection module, body parts detection module, car damage classification module, calculation module for damaged area of the vehicle part, and repair cost estimation module [2]. The car brand & model detection module are used to detect the automobile itself and determine the brand and model of the automobile. The car body parts detection module is used to detect all parts of the car, circle each part, such as bumper, fender, door, glass, head-lamp and tail-lamp, and calculate the area of each part.

Car damage classification module checks the damaged parts of the body, such as scratches, dents, damage, etc., and draw the damaged area and calculate the area of the damaged area, and calculates the percentage of the damaged area in this part. The repair cost estimation module is based on the repair and replacement costs of each part of the car brand and model, and the final result is used to estimate the cost required for car damage maintenance.

III. EXPERIMENTS AND ANALYSIS

The transfer learning method based on DenseNet161 in the proposed algorithm has the highest accuracy of 93.85%. In addition, in reference [4], the highest accuracy was obtained by using ResNet50, which was 76.68%. The accuracy of DenseNet161 adopted in this paper is higher than that of Reference [4]. Therefore, DenseNet161 method used in this paper is adopted for the car brand and model detection module.

IV. CONCLUSION

In this paper, an algorithm for damage detection and repair cost estimation based on car accidents is presented. The accuracy of the four methods used for detecting car brand and type reached 17.3%, 91.1%, 88.3%, and 93.85%, respectively. This study is expected to have an important utilization value in the field of calculating repair costs that occur in the process of settling automobile claims if it goes through the process of supplementing technology for commercialization in the future.

ACKNOWLEDGMENT

This work was supported by a grant from Tongmyong University Innovated University Research Park (I-URP) funded by Busan Metropolitan City, Republic of Korea. (IURP2301)

REFERENCES

[1] A. Shirode, T. Rathod, P. Wanjari, A. Halbe, "Car Damage Detection and Assessment Using CNN," In Proc. 2022 IEEE Delhi Section Conference (DELCON), IEEE, pp.1-5, 11-13 Feb. 2022.

[2] Min Htet Thar, Dong Myung Lee, "Vehicle Damage Volume and Traffic Accident Types Recognition Algorithm based on Mask Regional-CNN," In Proc. 2022 KICS Winter Conference (ISSN: 2383-8302), KICS, vol. 77, pp.0436-0437, 09-11 2022.

[3] Najmeddine Dhieb, Hakim Ghazzai, Hichem Besbes, and Yehia Massoud, "A very deep transfer learning modelfor vehicle damage detection and localization," In IEEE 31st International Conference on Microelectronics (ICM), pp.158–161, 2019.

[4] https://patricia-schutter.medium.com/car-image-recognition-with-convolutional-neural-network-applications-e791c98c9d72

Indoor Positioning Algorithm Using a Decision Tree Model in UWB Environment

Ho Chul Lee
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
calmtot@gmail.com

Woo Hyun Kim
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
kimwoohyun219@gmail.com

Dong Myung Lee
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
dmlee@tu.ac.kr

***Abstract*— In this paper, an indoor positioning algorithm that improves positioning accuracy and prediction performance by using a decision tree model in an ultra-wideband (UWB) environment is proposed. It then uses the decision tree model to accurately predict the coordinates of the UWB tag in the next direction. In addition, a total of 4 UWB anchor nodes and 1 tag node were placed in an indoor 6m × 6m × 2.5m space, and then the 3D positioning coordinate group was measured using the trilateration technique.**

Keywords*— *WPAN, UWB, Decision Tree, TWR, Indoor Positioning,

I. INTRODUCTION

Recently, research on positioning technology using ultra-wideband (UWB), belonging to the wireless personal area network (WPAN) technology category, has received a lot of attention [1-2]. However, the level of positioning technology using UWB is still in its infancy, and there are many tasks to be further studied. In this paper, an indoor positioning algorithm that can improve positioning precision and positioning prediction performances using a decision tree model in a UWB environment is proposed.

II. PROPOSED INDOOR POSITIONING ALGORITHM

The positioning process using UWB sensors and decision trees consists of the following four steps: 1) training of the decision tree model; 2) testing of the decision tree model; 3) real-time running of the decision tree model; and 4) updating of the dataset.

The proposed algorithm uses two decision tree models, the distance decision tree model and the direction decision tree model, which are overlaid in order to predict the next coordinate of the UWB tag [3-4]. The proposed algorithm selects the distance range class using the distance decision tree model and then selects the direction of the tag using the direction decision tree model. Proper selection of the distance range and direction is critical for accurate prediction of the next coordinate of the UWB tag.

III. EXPERIMENTS AND RESULT ANALYSIS

A total of four UWB anchor nodes are placed at the vertices of a 6m × 6m × 2.5m cube indoors using a tripod, and the tag node for positioning is placed inside the building cube designated as the vertex of the anchor node. The estimated position coordinates of the tag relative to the coordinates (10cm × 10cm) of one fixed tag can be displayed in the order of time in the 3D space. It can be seen that the estimated position coordinates of the tag are measured in various ways due to the noise caused by the multipath generation of the UWB radio signal as shown in Fig. 1. The experiment method is as follows: First, after setting 4 anchors and 1 tag in the UWB measuring equipment, measure the distance and calculate the location coordinates using the trilateration method. Then, after placing tags at 10 cm intervals at each coordinate point, the location coordinates of the tag and the distance between each anchor are measured for 30 seconds at 100ms intervals.

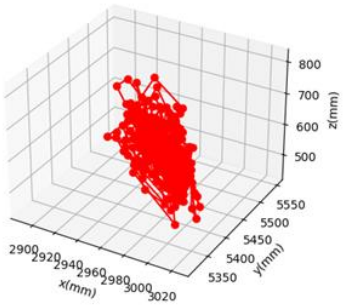


Fig. 1. Cluster map of measurement coordinates in 3-dimensional coordinates (3.0m × 5.4m × 0.6m).

IV. CONCLUSION

In this paper, an indoor positioning algorithm using a decision tree in a UWB environment is proposed and its performance was tested. In the future, we plan to generate simulated learning data and conduct simulations, followed by testing with actual UWB-equipped devices to validate the proposed algorithm.

ACKNOWLEDGMENT

This research was supported by the MISP (Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information & communications Technology Promotion) (2018-0-018740301001).

This work was supported by a grant from Tongmyong University Innovated University Research Park (I-URP) funded by Busan Metropolitan City, Republic of Korea. (IURP2301)

REFERENCES

[1] B. J. Jang, "Principles and Trends of UWB Positioning Technology," The Journal of Korean Institute of Electromagnetic Engineering and Science, vol. 33, no. 1, pp. 1-11, Jan. 31. 2022.

[2] S. J. Lee, J. W. Park, Y. B. Ko, "Random Forest Based Precise Localization Algorithm for UWB-Embedded Car Keys," The Journal of Korean Institute of Communications and Information Sciences, vol. 46, no.12, pp. 2310-2318, 2021.

[3] B. Kamiński, M. Jakubczyk, P. Szufel, "A framework for sensitivity analysis of decision trees," Central European Journal of Operations Research, vol. 26, pp. 135–159, May 24. 2017.

[4] A. Musa, G. D. Nugraha, H. Han, D. Choi, S. Seo, J. Kim, "A decision tree-based NLOS detection method for the UWB indoor location tracking accuracy improvement," International Journal of Communication Systems, Jun. 2019.

Path Estimation Algorithm based on W-RTMDet HPE-WP in Indoor Real-time Obstacle Environment

Tae Wan Kim
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
wanykt@gmail.com

Woo Hyun Kim
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
kimwoohyun219@gmail.com

Dong Myung Lee
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
dmlee@tu.ac.kr

Abstract—In this paper, a path estimation algorithm based on weight-real-time object detectors HPE-weight position (W-RTMDet HPE-WP) is proposed. In addition, the error distance of the movement trajectory of proposed algorithm is compared with RTMDet B-Box and RTMDet PSHP according to RTMDet, Resnet50, RSN50, and HRNet w32 pose model. In conclusion, the error distance of the movement trajectory and the error distance of the movement coordinates in our proposed algorithm were found to be significantly superior to both RTMDet B-Box and RTMDet PSHP.

Keywords-Path Estimation, Human Detection, Localization, W-RTMDet, HPE-WP

I. Introduction

In the path estimation method using HPE (Human Pose Estimation), the lower body information is very important among the joint information of the posture estimation of the moving person. In particular, since humans walk upright, ankle information is essential for accurate path estimation [1-3]. However, it is difficult to obtain accurate joint information in an obstacle environment using the method using HPE [4-6]. In an indoor real-time obstacle environment based on video, the accuracy of human object detection and pose estimation significantly decreases due to obstacles. In particular, it is difficult to obtain ankles information due to obstacles.

II. Algorithm Design

The proposed HPE-weight position (W-RTMDet HPE-WP) algorithm consists of 4 steps such as preparation step, W-RTMDet step, HPE step and localization step as shown in Fig. 1.

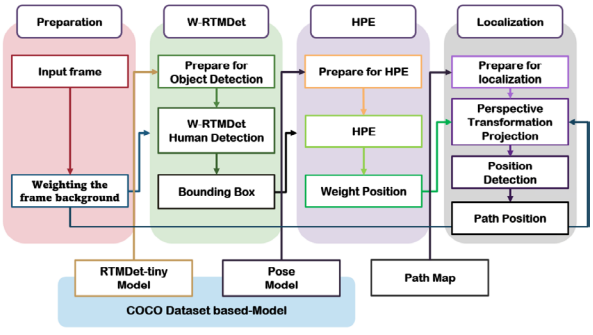


Fig. 1. System architecture of W-RTMDet HPE-WP for realtime indoor localization.

III. Experiments And Results

The experiment was done by moving the human object from 5m to 0m from the camera's position in a virtual obstacle environment. The algorithm proposed in the experiment was compared and analyzed with RTMDet B-Box and RTMDet PSHP [6] for each Pose Model, Resnet50, RSN50, and HRNet w32.

The results of the experiment, regarding the performance of the error distance for the movement trajectory are presented. It was confirmed that the error distance of moving trajectory in the proposed algorithm was decreased in all pose models compared to RTMDet B-Box and RTMDet PSHP. It was decreased by an average of 37.34cm compared to RTMDet B-Box, and by an average of 4.92cm compared to RTMDet PSHP.

IV. Conclusion

In this paper, the error distance of moving trajectory of proposed algorithm is confirmed to be far superior compared to RTMDet B-Box and RTMDet PSHP. The localization methods using pose models in real-time indoor obstacle environments exist errors according to the distance. However, our proposed algorithm achieved satisfactory results in obstacle environments, and it would be helpful to improve accuracy if we can select a pose model suitable for changes in distance. In conclusion, the error distance of moving trajectory and the error distance of moving coordinates of our proposed algorithm are confirmed to be far superior compared to RTMDet B-Box and RTMDet PSHP.

In the future, we plan to have experiments in real-time indoor moving environments based on proposed algorithm. In addition, we intend to apply our proposed algorithm to services that estimate the path of people in a CCTV environment.

Acknowledgment

This research was supported by the MISP (Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information & communications Technology Promotion) (2018-0-018740301001).

References

[1] Shaoqing Ren, Kaiming He, Ross Girshick, Jian Sun, “Faster r-cnn: Towards real-time object detection with region proposal networks,” Advances in NIPS 2015, pp. 91-99, Dec. 2015.

[2] Bin Xiao, Haiping Wu, Yichen Wei, “Simple baselines for human pose estimation and tracking, In Proceedings of the European conference on computer vision (ECCV), pp.466-481, Aug. 21 2018.

[3] Evan Cheshire, Cibele Halasz, and Jose Krause Perin, “Player tracking and analysis of basketball plays,” Proceedings of European Conference of Computer Vision, 2015.

[4] C. Zheng, et.al., “Deep Learning-Based Human Pose Estimation: A Survey,” arXiv preprint arXiv:2012.13392, 2020.

[5] T. W. Kim and D. M. Lee, “A Design of Learning based Human Positioning Algorithm using Images,” Proceedings of ICAIIC 2021, Jeju Island, Republic of Korea, pp.247-250, Apr. 13-16 2021.

[6] (Chengqi Lyu, et al., “RTMDet: An Empirical Study of Designing Real-Time Object Detectors,” arXiv preprint arXiv:2212.07784, Dec. 16 2022.

FPA Detection from GRF Data for Accurate Gait Modification

Ansary Shafew
Department of Electronics Engineering
Dong-A University
Busan, South Korea
shafewansary@gmail.com

Akter Morsheda
Department of Electronics Engineering
Dong-A University
Busan, South Korea
morshedaakteranni@gmail.com

Dongwan Kim
Department of Electronics Engineering
Dong-A University
Busan, South Korea
dongwankim@dau.ac.kr

***Abstract*— Knee osteoarthritis is a prevalent condition that can result from joint load, which refers to the force applied to a weight-bearing knee joint. Several factors contribute to this condition, including injuries, high-impact activities, and obesity. Gait modification is a commonly used technique to reduce knee load and potentially slow the progression of knee osteoarthritis. This modification changes the foot progression angle (FPA), which is the angle between the direction of walking and the long axis of the foot. The present study aims to develop a system to monitor and provide real-time feedback on gait modification in a home environment. The system incorporates an inertial measurement unit (IMU) sensor to detect normal, toe-in, and toe-out walking postures by measuring the ground reaction force (GRF) when attached to a person. The dissimilarities between the acquired data suggests that this system can leverage few-shot learning (FSL) to accurately classify the walking posture, enabling efficient and effective monitoring of gait modification in real time which is the novelty of this work in progress research.**

Keywords*— *foot progression angle, ground reaction force, gait cycle

I. Introduction

A frequent orthopedic condition that puts a lot of strain on a person's knees is excessive medial knee loading [1]. Chronic exposure to this condition may cause serious health issues such knee osteoarthritis, which can limit mobility [2]. Multiple factors, such as high-impact activities, injuries, and excessive weight gain, can cause knee loading [3]. One effective method to remedy this issue is to modify the gait by adjusting the foot progression angle (FPA) using camera feeds and pressure pads [5]. But the operation process, high cost and non-portability of the equipment, limits its application outside clinical settings, preventing real-time feedback. Therefore, to overcome these limitations Nan Jiang [6] et al proposed a method called the EarWalk system with an embedded inertial measurement unit (IMU) sensor in earbuds. This method can identify a person's FPA during walking with a 95% accuracy in laboratory settings. However, the system's performance is limited by the controlled environment of the testing setup, as a person's gait and FPA can vary unpredictably in real-world situations. The purpose of this research is to enhance the existing method of identifying FPA and simplify its implementation by incorporating deep learning automated feature extraction with minimal pre-processing. The resulting method is expected to offer higher accuracy and precision. The work can be divided into three steps: testing ground reaction force (GRF) data using an IMU sensor, creating a database of normal, toe-in, and toe-out walking postures from 40-50 individuals, and designing a few-shot learning (FSL) model to train and test the data.

In the initial GRF data acquisition phase, promising results were obtained using an IMU sensor. Graphs obtained from different walking postures were compared, and the dissimilarities were observed, indicating the potential for using a supervised FSL model. Additionally, resampling the data into smaller samples could increase the data points and a 2D convolutional neural network (CNN) model can also be used to potentially improve precision. This is a novel approach for real time gait monitoring and modification. The general working process is described followed by the experimental procedure section. The paper concludes by presenting the results and describing future works.

II. General Working Principle

Toe-in and toe-out walking postures involve pointing the foot inwards or outwards during walking, causing the inner knee to receive induced GRF as the foot contacts the ground. This leads to counterclockwise rotation of the shinbone around the knee joint due to GRF, increasing compression and stress on the inner knee joint. These postures have been found effective in reducing medial knee loading by altering the direction of applied GRF closer to the knee center through changing the foot-ground contact region [4]. By reducing the rotation caused by the GRF and minimizing the distance between the GRF and shinbone, this shift lessens the stress on the inner knee. This impact is demonstrated in Figure 1.

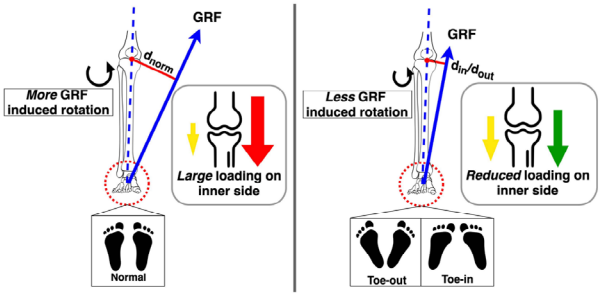


Fig. 1. GRF changes based on toe-in and toe-out where d_{norm} is the normal walking GRF which reduces to d_{in} or d_{out}, based on toe-in or toe-out [6]

The primary aim of this study was to identify the variance in the induced GRF that is transmitted from the foot to the body. Previous research indicates that IMU sensors can be installed at various points in the human body to detect GRF. The original EarWalk system [6] was designed to detect GRF from the earbud's IMU sensor. However, in this study, a commercial IMU sensor, namely the BNO055, was selected for data collection. The waist region was selected as the ideal location for mounting the IMU sensor, as it can capture the GRF forces produced by different walking patterns and output them as distinct signal values. Figure 2. depicts the overall experimental process. The data was collected from one individual by mounting the experimental apparatus firmly around the waist. The apparatus consists of an Arduino Uno, the Adafruit BNO055 IMU sensor, a Bluetooth HC-06 module and a 9V battery to power the device. The data was sent from the Arduino to a laptop where it was collected using a serial monitor 'Putty' and exported as a *.csv* file for easier evaluation. The sampling rate of the Bluetooth device was set

to 10hz, and communication throughput was 9600. The Arduino code for the IMU sensor found on the Adafruit website, was modified to output sixty seconds of data and then stop. The process can then be restarted by pressing the ‘r’ key on the laptop.

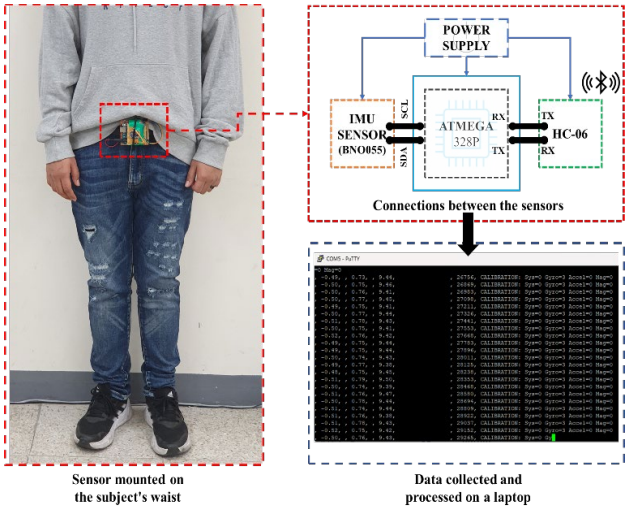


Fig. 2. Overall experimental process

The data was later processed and normalized using MatLab and results were observed. The processing was a simple time vs acceleration graph. During the processing, the sample data was equally scaled to 20 seconds to see the visible differences between the different FPA. Even though visible differences can be seen, on a large scale for many individuals the FPA is vastly different as gait is a type of biometrics for an individual person [6]. However, there are still differences that can be identified by a deep learning model that humans simply cannot comprehend.

III. Experimental Procedure And Result

The experimental procedure entailed collecting data for each of the FPA patterns for one minute to assess differences between them. The participant was instructed to walk in a straight line in laboratory condition for one minute each in normal, toe-in, and toe-out configurations to avoid data overflow while a smartphone video capture was used as a reference. The height, weight and age of the participant were 172 cm, 80 kg, and 29 years respectively.

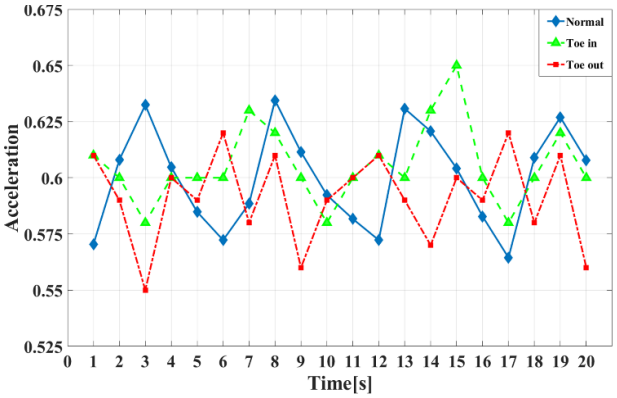


Fig. 3. Difference between normal, toe-in and toe-out GRF patterns

The system was mounted securely on the participant's waist, and they were instructed to walk at a slow pace of 1 m/s. Figure 3. depicts the induced GRF of an individual's walking or gait cycle for each of the three tested FPA patterns. The study found that each FPA resulted in unique recurring GRF patterns as shown in Figure 3. This suggests the potential to train and test with an FSL model. This method could later be used to detect FPA patterns in real-time.

IV. Conclusion

The results of the initial data acquisition phase for gait analysis using an IMU sensor have demonstrated promising outcomes. The graphs obtained from different walking postures showed enough dissimilarities which indicates the potential for automated feature extraction through deep learning. Additionally, resampling the data into smaller samples can result in more data points for the deep learning model, thereby potentially improving precision. In future work, a database comprising normal, toe-in, and toe-out walking postures from 40-50 individuals will be created. The collected data will then be utilized to train and test an FSL model that can accurately identify the postures. The dataset parameters can be increased though resampling and also be used for automated feature extraction system on a 2D CNN model, which will be augmented with necessary tools and modifications to enhance its simplicity and accuracy. Finally, the model will be tested in real-life scenarios to provide live feedback of FPA from individual subjects. These findings hold significant implications for developing a robust and reliable system for gait analysis in clinical settings.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government. (NRF-2021R1F1A1062443)

References

[1] S. Amin, N. Luepongsak, C. A. McGibbon, M. P. LaValley, D. E. Krebs, and D. T. Felson, "Knee adduction moment and development of chronic knee pain in elders," Arthritis Care and Research, vol. 51, no. 3, pp. 371-376, 2004. doi: 10.1002/art.20396. PMID: 15188321.

[2] J. Jack II, L. Miller, and J. Block, "Quality of Life in Patients with Knee Osteoarthritis: A Commentary on Nonsurgical and Surgical Treatments," The open orthopaedics journal, vol. 7, pp. 619-623, 2013. doi: 10.2174/1874325001307010619

[3] T. A. Donelon, T. Dos’Santos, G. Pitchers, et al., "Biomechanical Determinants of Knee Joint Loads Associated with Increased Anterior Cruciate Ligament Loading During Cutting: A Systematic Review and Technical Framework," Sports Med - Open, vol. 6, p. 53, 2020. doi: 10.1186/s40798-020-00276-5

[4] A. Chang, D. Hurwitz, D. Dunlop, J. Song, S. Cahue, K. Hayes, and L. Sharma, "The relationship between toe-out angle during gait and progression of medial tibiofemoral osteoarthritis," Annals of the rheumatic diseases, vol. 66, pp. 1271-1275, 2007. doi: 10.1136/ard.2006.062927

[5] J. Charlton, J. Eng, L. Li, and M. Hunt, "Learning Gait Modifications for Musculoskeletal Rehabilitation: Applying Motor Learning Principles to Improve Research and Clinical Implementation," Physical Therapy, vol. 101, 2020. doi 10.1093/ptj/pzaa207

[6] N. Jiang, T. Sim, and J. Han, "EarWalk: towards walking posture identification using earables," pp. 35-40, 2022. [Online]. Available: 10.1145/3508396.3512883.

On the Performance of Covert Communications under Nakagami distribution with Friendly Jamming

Nguyen Trong Huan and Tran Trung Duy
Posts and Telecommunications Institute of Technology
Ho Chi Minh City, Vietnam,
e-mail: {huannt,duytt}@ptit.edu.vn,
huannt@ptithcm.edu.vn

Lam-Thanh Tu
Faculty of Electrical and Electronics Engineering,
Ton Duc Thang University
Ho Chi Minh City, Vietnam,
e-mail: tulamthanh@tdtu.edu.vn

Sang Nguyen Quang
Science and Technology Application for Sustainable Development Research Group,
Ho Chi Minh City University of Transport
Ho Chi Minh City, Vietnam
e-mail: sang.nguyen@ut.edu.vn

Yeon Ho Chung
Department of Information and Communications Engineering
Pukyong National University
Busan, South Korea
yhchung@pknu.ac.kr

***Abstract*— In this paper, we address the performance of covert communications with the help of a friendly jammer under the Nakagami-m distribution. Particularly, we study the false alarm probability (FAP) and the miss detection probability (MDP) of the warden. Our findings show that FAP and MDP have contrary behaviors with respect to the transmit power of the jammer.**

***Keywords*–Covert communications, false alarm probability, miss detection probability**

I. Introduction

With the exponential growth of wireless-connected devices, information security has become one of the critical issues in wireless networks along with the long-term challenging, namely, energy efficiency and spectral efficiency. However, the conventional cryptography-based approach is only effective provided that the end devices are capable to handle large and sophisticated computations. Nonetheless, most of the end devices in the Internet of Things (IoTs) are low-cost and have fewer computation capabilities. Besides the cryptography-based approach, there is another low-cost approach that is lying in the physical layer called physical layer security (PLS). The main drawback of the PLS approach is that eavesdroppers still receive information from legitimate links. They, as a consequence, can employ several advanced techniques to wiretap the secure information such as transmit and/or receiver diversity schemes and exploit channel state information (CSI) of the main link to either wiretap or send malicious information to the intended receiver. Fortunately, a pioneer work of Bash and others in [1] has shed light on information security. Particularly, by hiding the secure information detected by the wardens, one can guarantee information security and achieve so-called *covert capacity*. Such communications are named covert communications. Compared with PLS, it also employs the physical layer, however, the approach is totally different. In covert communications, the transmitted signals are secretly transmitted, i.e., hiding in the background noise, so that the wardens do not realize whether the transmitter has transmitting information or not while in the PLS, the eavesdroppers always know that the transmitter is transmitting. Thus, it can theoretically provide better protection [2].

Although covert communications have much potential compared with the cryptography approach and PLS, the number of works in the literature was limited [3]–[6]. More precisely, the authors in [3] studied the probability of detection error and average covert rate of the dual-hop networks where the relay secretly transmits its own information. Two relay selection schemes were proposed in [4] to improve the covert rate in the IoTs networks. The performance of the delay-tolerant covert communications with fixed and random transmit power was addressed in [5]. They showed that the amount of secrecy information successfully transmitted logarithmically increased under the finite block length. The performance of the multiple antennae covert communications was studied in [6] where the locations of both the interferer and wardens are modeled according to tools from stochastic geometry. The covert throughput was maximized under the impact of both small-scale fading and large-scale path-loss. Their findings illustrated that the centralized antenna systems provided better performance compared with its counterpart, i.e., distributed antennae systems.

Nevertheless, all of the above-mentioned work do not consider the Nakagami fading as well as the impact of the friendly jamming to improve the hidden probability. In the present work, different from the above-mentioned work, we studied covert communications under the impact of Nakagami fading and with the help of a friendly jammer. Particularly, we derive the false alarm probability (FAP) and miss detection probability (MDP) of the warden.

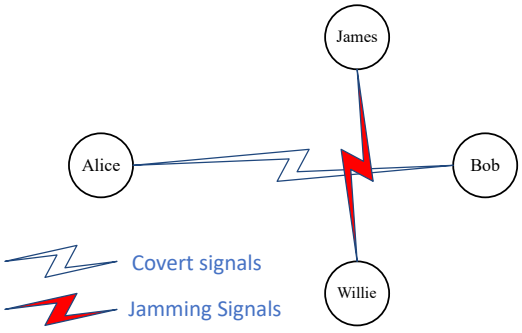


Fig. 1. System model

II. System Model

Let us consider covert communications comprised of a source node Alice denoted by A, a legitimate destination Bob denoted by B, a warden William denoted by W, and a friendly jammer denoted by J. Alice secretly transmits information to Bob and Willie detects whether Alice is transmitting or not. Jame, on the other hand, continuously creates artificial noise (AN) toward Willie to help Alice transmit information to Bob. In this work, we consider the friend jamming scenario, which means that both Alice and Bob know the signals transmitted by Jame. Willie, yet, does not know this signal. The whole transmission is taken place in a time slot and all nodes are assumed to equip with a single antenna. We further assume that the channel coefficient follows by Nakagami-m distribution and the fading changes independently between time slots. The received signals at Willie under the n time slot is formulated as follows:

$$y_W[n] = \begin{cases} \sqrt{P_J} h_{J,W} x_J[n] + n_W[n] & H_0 \\ \sqrt{P_A} h_{A,W} x_A[n] + \sqrt{P_J} h_{J,W} x_J[n] + n_W[n] & H_1 \end{cases}, \quad (1)$$

where P_u, x_u, $u \in \{A, J\}$ are the transmit power and transmitted signals of Alice and Bob; n_W is the additive Gaussian white noise (AWGN) at Willie; $h_{u,W}$ is the channel coefficient from node u to Willie and is followed by Nakagami-m distribution with shape and scale parameters m_u and δ_u. Here, we assume that x_u, $u \in \{A, J\}$ has unit power, i.e., $\mathbb{E}\left\{|x_u|^2\right\} = 1, \forall u$, $\mathbb{E}\{.\}$ is the expectation operator. H_0 is the hypothesis that Alice does not transmit information while H_1, on the contrary, is the hypothesis that Alice transmits information. It is noted that James periodically changes its jamming signals so that Willie is not able to recognize its pattern so that can subtract these signals.

The false alarm probability and miss detection probability are defined as

$$\begin{aligned} \text{FAP} &= \Pr\{D_1 | H_0\} = \Pr\left\{ P_J |h_{J,W}|^2 + \sigma_W^2 \geq \gamma \middle| H_0 \right\} \\ \text{MDP} &= \Pr\{D_0 | H_1\} \\ &= \Pr\left\{ P_A |h_{A,W}|^2 + P_J |h_{J,W}|^2 + \sigma_W^2 < \gamma \middle| H_1 \right\}. \end{aligned} \quad (2)$$

Here $\Pr\{.\}$ is the probability operator, D_0 and D_1 are the events that Willie make a favorable decision on H_0 and H_1, γ is the detection threshold, $\sigma_W^2 = -174 + \text{NF} + 10\log_{10}\text{BW}$ [dBm] is the noise variance of Willie.

III. Numerical Results

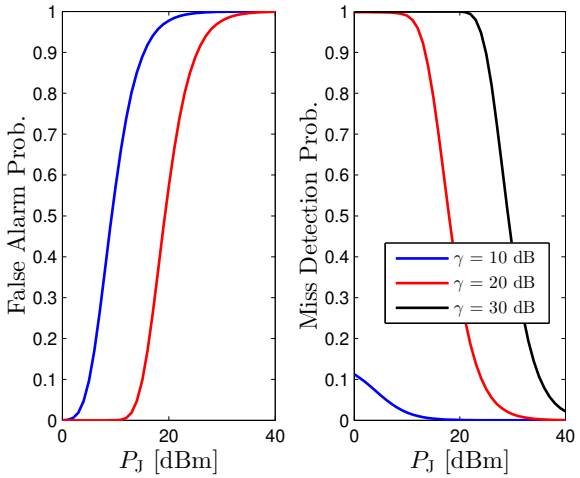


Fig. 2. FAP and MDP vs. P_{rmJ} with various values of γ.

Fig. 2 illustrates the performance of the FAP and MDP with respect to the transmit power of the jammer with various values of γ. We observe that FAP and MDP have a contrary behavior with respect to the P_{rmJ}. Particularly, MDP is a monotonic decreasing function while FAP is an increasing function regarding P_{rmJ}.

References

[1] B. A. Bash, D. Goeckel, and D. Towsley, "Limits of reliable communication with low probability of detection on AWGN channels," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 9, pp. 1921–1930, Sept. 2013.

[2] X. Chen *et al.*, "Covert Communications: A Comprehensive Survey," *IEEE Communications Surveys & Tutorials*, vol. 25, no. 2, pp. 1173-1198, Secondquarter 2023, doi: 10.1109/COMST.2023.3263921.

[3] Y. Su, H. Sun, Z. Zhang, Z. Lian, Z. Xie and Y. Wang, "Covert Communication With Relay Selection," *IEEE Wireless Communications Letters*, vol. 10, no. 2, pp. 421-425, Feb. 2021, doi: 10.1109/LWC.2020.3033786.

[4] C. Gao, B. Yang, X. Jiang, H. Inamura and M. Fukushi, "Covert Communication in Relay-Assisted IoT Systems," *IEEE Internet of Things Journal*, vol. 8, no. 8, pp. 6313-6323, 15 April15, 2021, doi: 10.1109/JIOT.2021.3051694.

[5] S. Yan, B. He, X. Zhou, Y. Cong and A. L. Swindlehurst, "Delay-Intolerant Covert Communications With Either Fixed or Random Transmit Power," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 1, pp. 129-140, Jan. 2019, doi: 10.1109/TIFS.2018.2846257.

[6] T. -X. Zheng, H. -M. Wang, D. W. K. Ng and J. Yuan, "Multi-Antenna Covert Communications in Random Wireless Networks," *IEEE Transactions on Wireless Communications*, vol. 18, no. 3, pp. 1974-1987, March 2019, doi: 10.1109/TWC.2019.2900915.

A Review on Resource Scheduling and Association Scheme in Cell-Free Networks

Yunseong Lee
School of Computer Science and Engineering
Chung-Ang University
Seoul, Republic of Korea
yslee@uclab.re.kr

Sungrae Cho
School of Computer Science and Engineering
Chung-Ang University
Seoul, Republic of Korea
srcho@cau.ac.kr

***Abstract*—In order to guarantee the quality of service in 6G, cell-free networks are attracting attention. Cell-free is one of the massive MIMO technologies that perform cooperative transmission by distributing multiple APs It guarantees high SINR and connectivity compared to tradistional networks. This paper introduces the state-of-the-art schemes of cell-free and user-centric cell-free networks.**

***Index Terms*—Cell-free, User-centric Cell-free, Clustering, mMIMO**

I. Introduction

Cell-free is a massive multiple-input multiple-output (mMIMO) technology in which all access points (APs) provide services to users at the same frequency and time. The APs are connected to one central processing unit (CPU) The user is guaranteed a higher signal than the traditional network. In addition, even if the user moves, connectivity is maintained thanks to a number of connected APs, that is, handovers do not occur. However, as the AP performs channel measurement and data transmission for all users, there is a problem of excessive traffic overhead in the fronthaul. In this paper, we study the resource and user association scheme in cell-free.

II. Resource and Association in Cell-Free Network

Ngo et al. [1] proposed a cell-free mMIMO technique with a pilot assignment and power control algorithm to improve throughput. The scheme maximizes the smallest user data rates compared with traditional small cell networks. Zhang et al. [2] suggested a resource allocation with a heuristic method as a neighborhood field optimization to increase energy efficiency and throughput. To decide the optimal solution to the complex hypothesis task, the heuristic algorithm is combined with an adaptive boosting method. Vu et al. [3] designed a framework for uplink training and resource allocation with federated learning. In order to optimize transmit power and processing frequency, the schemes aim to minimize channel training time and execution time of learning. Le et al. [4] proposed a user clustering and beamforming scheme to optimize the transmit power of APs. The clustering scheme based on k-means decides the center uesrs for clusters and allocates transmit power through the inner approximation framework. However, since exchanging channel information between all APs and CPU for transmission to UE in front haul occurs, it causes increasing computational and power consumption overhead.

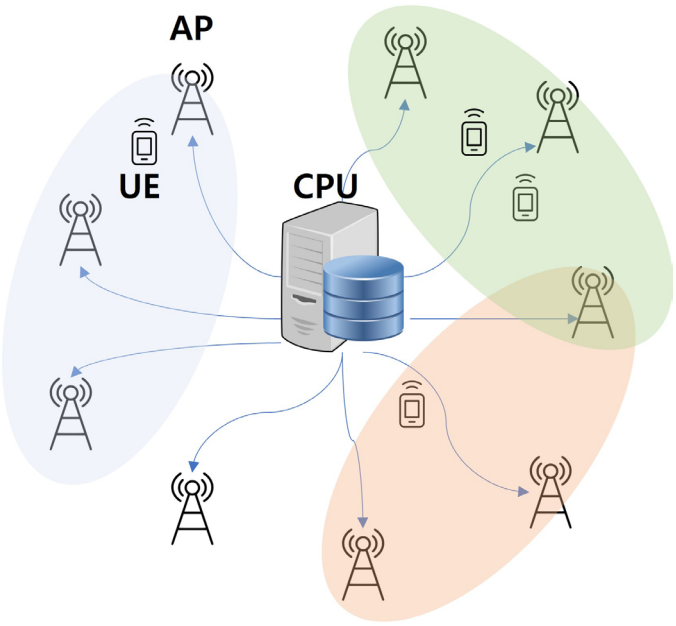


Fig. 1. User-centric cell-free

In the user-centric cell-free, a user receives packets from a set of APs, not all APs. In user-centric, the user connects only a few APs, not all APs. This reduces fronthaul traffic while ensuring high signal. However, there are challenges such as optimal AP clustering and itra-handover between CPUs. In [5], the authors extended the user-centric cell-free including power allocation to maximize the lower bound for sum rate and the minimum spectral efficiency lower bound. Xiao et al. [6] analyzed the CPU handover and AP handover in a hybrid AP selection method that combines conventional coordinated multipoint with the joint transmission for pure user-centric cooperation. The scheme shows that the handover rates are affected by the size of the AP within the CPU and the size of the AP set. Zaher et al. [7] suggested a soft handover approach using pilot association and clustering control. The scheme minimizes the handover of a master AP that assigns pilot sequences instead of other APs in a cluster.

Ammar et al. [8] proposed distributed user scheduling and beamforming schemes with low computational complexity. In order to optimize resource allocation, the Hungarian algorithm and fractional programming are applied. Demirhan et al. [9] designed a beamforming scheme for wireless front haul between CPU and APs to improve end-to-end data rate. To reduce computing complexity for max-min fair optimization, the scheme applies a suboptimal iterative solution.

Unlike the common cell-free, the user-centric cell-free is not completely handover free. However, most studies of user-centric cell-free networks do not consider the handover issue. Some studies [10], [11] analyzed the handover rate depending on the cluster size of APs. These schemes show that the handover rate is affected by cluster size. Kibinda et al. [12] designed the handover pattern in user-centric networks and proposed a handover scheme based on the UE path. The scheme predicts UE mobility patterns and decides whether or not to conduct the handover.

Although existing works try to improve performance for cell-free, there are several challenges such intra-handover problem, dynamic power control, and clustering interference. Therefore, we are supposed to study adaptive resource allocation scheme for cell-free network.

III. Conclusion

In this paper, we introduce the various network scheme for cell-free and user-centric. In future work, we suggest a joint optimization scheme for clustering for user-centric and resource allocation.

Acknowledgment

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-RS-2023-00258639) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation).

References

[1] H. Q. Ngo, A. Ashikhmin, H. Yang, E. G. Larsson, and T. L. Marzetta, "Cell-free massive mimo versus small cells," *IEEE Transactions on Wireless Communications*, vol. 16, no. 3, pp. 1834–1850, 2017.

[2] X. Zhang, H. Qi, X. Zhang, and L. Han, "Energy-efficient resource allocation and data transmission of cell-free internet of things," *IEEE Internet of Things Journal*, vol. 8, no. 20, pp. 15 107–15 116, 2020.

[3] T. T. Vu, D. T. Ngo, N. H. Tran, H. Q. Ngo, M. N. Dao, and R. H. Middleton, "Cell-free massive mimo for wireless federated learning," *IEEE Transactions on Wireless Communications*, vol. 19, no. 10, pp. 6377–6392, 2020.

[4] Q. N. Le, V.-D. Nguyen, O. A. Dobre, N.-P. Nguyen, R. Zhao, and S. Chatzinotas, "Learning-assisted user clustering in cell-free massive mimo-noma networks," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 12, pp. 12 872–12 887, 2021.

[5] S. Buzzi, C. D'Andrea, A. Zappone, and C. D'Elia, "User-centric 5g cellular networks: Resource allocation and comparison with the cell-free massive mimo approach," *IEEE Transactions on Wireless Communications*, vol. 19, no. 2, pp. 1250–1264, 2019.

[6] Y. Xiao, P. Mähönen, and L. Simić, "Mobility performance analysis of scalable cell-free massive mimo," *arXiv preprint arXiv:2202.01488*, 2022.

[7] M. Zaher, E. Björnson, and M. Petrova, "Soft handover procedures in mmwave cell-free massive mimo networks," *arXiv preprint arXiv:2209.02548*, 2022.

[8] H. A. Ammar, R. Adve, S. Shahbazpanahi, G. Boudreau, and K. V. Srinivas, "Distributed resource allocation optimization for user-centric cell-free mimo networks," *IEEE Transactions on Wireless Communications*, vol. 21, no. 5, pp. 3099–3115, 2021.

[9] U. Demirhan and A. Alkhateeb, "Enabling cell-free massive mimo systems with wireless millimeter wave fronthaul," *IEEE Transactions on Wireless Communications*, 2022.

[10] W. Bao and B. Liang, "Optimizing cluster size through handoff analysis in user-centric cooperative wireless networks," *IEEE Transactions on Wireless Communications*, vol. 17, no. 2, pp. 766–778, 2017.

[11] E. Demarchou, C. Psomas, and I. Krikidis, "Intelligent user-centric handover scheme in ultra-dense cellular networks," in *GLOBECOM 2017-2017 IEEE Global Communications Conference*. IEEE, 2017, pp. 1–6.

[12] N. M. Kibinda and X. Ge, "User-centric cooperative transmissions-enabled handover for ultra-dense networks," *IEEE Transactions on Vehicular Technology*, vol. 71, no. 4, pp. 4184–4197, 2022.

Spectrum Efficient Precoding Scheme for Full-Duplex Communications with Imperfect CSI

Teklu Merhawit Berhane and Wooyeol Choi*
Department of Computer Engineering, Chosun University
Gwangju, Republic of Korea
{mberhane63, wyc}@chosun.ac.kr

***Abstract*—The utilization of full-duplex (FD) in combination with multiuser (MU) MIMO exhibits a strong potential as an optimal technological solution for wireless communications beyond 5G. In FD-MU-MIMO systems, prevalent factors that limit performance include self-interference (SI), co-channel interference (CCI), multiuser interferences (MUI), and noise. Furthermore, the system model incorporates the consideration of imperfect channel state information within the context of the Rician fading channel. In this study, we propose a signal-leakage-plus-noise ratio precoding scheme (SLNR-PS), which is not restricted by antenna constraints and is found to exhibit superior spectral efficiency gain as compared to both zero-forcing and block-diagonalization precoding methodologies as validated using simulation results. The results indicate that, even in the presence of high SI, channel error, and Rician factor, the FD-MU-MIMO system provides significant SE improvements over HD MU-MIMO systems based on SLNR-PS.**

***Index Terms*—Full-duplex, Rician fading channel, imperfect CSI, spectral efficiency**

I. Introduction

Next-generation wireless communication systems will require reliable and spectrally efficient transmission techniques to support emerging high-data-rate applications. Full-duplex (FD) systems have received much attention from academia in 5G and beyond technologies due to their ability to double spectral efficiency (SE) when compared to half-duplex (HD) for mmWave, short-range, and different applications [1]. In FD transmissions, self-interference (SI) is typically much greater than the co-channel interference (CCI) of other users, preventing double sum-rate gains. As a result, it is crucial to effectively manage SI in order for FD systems to achieve spectral efficiency that is significantly higher than HD systems. Several SI mitigation techniques exist, including antenna cancellation combined with RF interference cancellation [2], analog cancellation [3], and digital baseband interference suppression [4]. Insightful current studies illustrate SI cancellation in FD devices illustrated that residual SI power at the noise floor could be considerably mitigated by SI cancellation [5]–[8]. It is clear from this result that FD-based systems can be applied to practical situations. FD-based MU-MIMO (FD-MU-MIMO) system is a system in which downlink (DL) and uplink (UL) channel users receive and transmit data streams, respectively, from/to base station (BS) at the same resources. This system severely suffers due to interferences such as SI, CCI, and multi-user interferences (MUI) in both channels. To mitigate these interferences, relying on sub-optimal and lower complex transmission schemes is necessary. Most related works have focused on SE maximization for FD-MU MIMO systems with zero-forcing/block-diagonalization (ZF/BD) precoder designs, and a few studies based on signal-leakage-to-noise-ratio precoding scheme (SLNR-PS) when the transmitter is available a channel state information (CSI) [9]–[11].

In addition, the SE of FD-MU-MIMO systems for Gaussian distribution and independent and identically distributed (i.i.d.) Rayleigh's fading environment was widely analyzed. Nevertheless, in a real-world circumstance, there is a line-of-sight (LOS) link between the transmitter and the receiver [12], [13]. For instance, in short-range and mmWave communications, for example, there is an LOS link, and in such fading conditions, the channel is rather described by the Rician fading model. Mathematically, the random channel matrix in a MIMO channel with a Rician-faded channel is a complex Gaussian matrix with a nonzero mean matrix, unlike in an i.i.d. Rayleigh-faded MIMO channel where the channel matrix is of zero mean. Clearly, the Rayleigh-faded model can be viewed as a special case of the Rician fading model by setting the mean is set to zero. To get a more realistic assessment of the capacity increasing potential of MIMO channels, it is important to evaluate this capacity in a Rician fading environment. This fading channel model comprises a direct path (LOS) and a non-line-of-sight path (NLOS). To include this scenario, we use a flat Rician flat fading channel.
Motivated by those works, in this work, the effects of imperfect channel state information (ICSI) on sum rates of DL and UL channel users, taking into account CCI at the DL and SI at the UL channel for single stream and multiple antennas per user over Rician fading channels is examined.

- By using the proposed SLNR-PS, the sum rates of UL and DL channel users are achieved. For FD-assisted systems, the sum rate of both channels results in a new total SE.
- In this study, the SEs of all the FD modes of interest are compared with those of the conventional half-duplex system. A half-duplex MU-MIMO mode base station's total spectral efficiency can be calculated by dividing the

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) under a grant (No. NRF-2021R1I1A3050535).
*Corresponding author

downlink channel sum rate by the uplink channel sum rate without taking SI into account.

II. FD-based System Model and Proposed Techniques

Our work considers the FD mode of base-station equipped with total antennas $N_T = N_t + N_r$, in which N_t and N_r designate the transmitting and receiving antennas at the BS sequentially. The numbers of antennas for each user can be represented as $N_{d,j}, j = 1, 2, \cdots, D_u$ and $N_{u,j}, j = 1, 2, \cdots, U_u$ for DL and UL channels, respectively. The Rician non-frequency selective fading channel components can be written as $\mathbf{H}_{m,j} = \sqrt{\frac{\kappa}{\kappa+1}}\bar{\mathbf{H}}_{m,j} + \sqrt{\frac{1}{\kappa+1}}\tilde{\mathbf{H}}_{m,j},\ j = 1, 2, \cdots, U$ [14], where m and U are changeable according to the UL and DL. As well, κ stand for the Rician factor, $\bar{\mathbf{H}}_{m,j}$ refers the specular matrix and $\tilde{\mathbf{H}}_{m,j}$ designates the Rayleigh entries with zero-mean unit-variance, which is $\mathcal{CN}(0, \mathbf{I}_{N_{m,j}})$. Moreover, it is assumed that the BS and channel users have inaccurate information about the state of the channel. Based on the imperfect channel state information (ICSI), the channel model is formulated as $\mathbf{H}_{m,j} = \widetilde{\mathbf{H}}_{m,j} + \mathbf{\Delta}_{m,j}$, where $\widetilde{\mathbf{H}}_{r,i}$ and $\mathbf{\Delta}_{r,i}$ represent the nominal and uncertainty channels, respectively. The elements of $\mathbf{\Delta}_{r,i}$ represented as zero-mean i.i.d complex Gaussian random variables with variance $\sigma^2_{\mathbf{\Delta}}$ [15].

A. Downlink Transmission

The signal model for the DL transmission can be described as follows. The received vector at the j^{th} user formulated as

$$\begin{aligned}\mathbf{y}_{d,j} = \mathbf{H}_{d,j}\mathbf{z}_{d,j}\mathbf{x}_{d,j} + \mathbf{H}_{d,j}\sum_{r\neq j}^{U_d}\mathbf{z}_{d,r}\mathbf{x}_{d,r} + \mathbf{H}_{cci,j}\mathbf{x}_{cci,j} \\ + \mathbf{n}_{d,j},\ j = 1, \cdots, U_d,\end{aligned} \tag{1}$$

where $\mathbf{z}_{d,j}$ is precoding vector and $\mathbf{x}_{d,j}$ is the intended signal. Then, we normalize the vectors $\mathbf{x}_{d,j}$ and $\mathbf{z}_{d,j}$ as $\mathrm{E}[|\mathbf{x}_{d,j}|^2] = 1$ and $||\mathbf{z}_{d,j}||^2 = 1$, respectively. The four elements in the sequence shown on the right side of (1) are the desired signal, MUI, CCI, and noise.

We examine the strategic transmission scheme for the DL system model. For this system, a signal-leakage-to-noise-ratio precoding scheme (SLNR-PS) is proposed to suppress CCI and maximize the sum rate despite co-channel interference for a given receiver structure by only using the CCI plus-noise covariance matrix at the base station. Then, the $N_{d,j} \times N_{d,j}$ CCI-plus-noise covariance matrices at user j can be expressed as

$$\mathbf{J}_j = \mathbf{H}_{cci,j}\mathbf{W}_{cci,j}\mathbf{H}^H_{cci,j} + \sigma^2_n\mathbf{I}_{N_{d,j}}, \tag{2}$$

where $\mathbf{W}_{cci,j} = \mathrm{E}[\mathbf{x}_{cci,j}\mathbf{x}^H_{cci,j}]$ stand for the covariance matrices of the CCI and $\mathrm{Tr}(\mathbf{W}_{cci,j}) = \mathbf{P}_{cci,j}$. The whitening matrix can be defined by $\mathbf{T}_j = \mathbf{J}_j^{-1/2}$. Then, (1) can be multiplied by $\mathbf{T}_j$ as well, considering the ICSI, and it becomes

$$\begin{aligned}\tilde{\mathbf{y}}_{d,j} = \mathbf{T}_j\widetilde{\mathbf{H}}_{d,j}\mathbf{z}_{d,j}\mathbf{x}_{d,j} + \mathbf{T}_j\widetilde{\mathbf{H}}_{d,j}\sum_{l\neq i}^{U_d}\mathbf{z}_{d,r}\mathbf{x}_{d,r} \\ + \sum_{j=1}^{U_d}\mathbf{T}_j\mathbf{\Delta}_{d,j}\mathbf{z}_{d,j}\mathbf{x}_{d,j} + \tilde{\mathbf{n}}_{d,j},\end{aligned} \tag{3}$$

where $\tilde{\mathbf{n}}_{d,j} = \mathbf{T}_j(\mathbf{H}_{cci,j}\mathbf{x}_{cci,j} + \mathbf{n}_{d,j})$ with covariance $\mathbf{I}_{N_{d,j}}$ because of the whitening matrix.

Using the figure of merit of SLNR, the optimization problem for user j can be described as shown in (4).

According to SLNR as the optimization metric in [16], [17], the optimization problem for user j by making use of the Rayleigh-Ritz quotient can be formulated as

$$\begin{aligned}\frac{\mathbf{z}^H_{d,j}(\widetilde{\mathbf{H}}^H_{d,j}{\mathbf{T}_j}^H\mathbf{T}_j\widetilde{\mathbf{H}}_{d,j} + \sigma^2_{\mathbf{\Delta}}{\mathbf{T}_j}^H\mathbf{T}_j)\mathbf{z}_{d,j}}{\mathbf{z}^H_{d,j}(N_{d,j}\sigma^2_{\tilde{n}} + \delta\sigma^2_{\mathbf{\Delta}}{\mathbf{T}_j}^H\mathbf{T}_j + \bar{\mathbf{H}}^H_{d,j}\mathbf{T}^H_j\mathbf{T}_j\bar{\mathbf{H}}_{d,j})\mathbf{z}_{d,j}} \\ \leq \lambda_{max}(\widetilde{\mathbf{H}}^H_{d,j}{\mathbf{T}_j}^H\mathbf{T}_j\widetilde{\mathbf{H}}_{d,j} + \sigma^2_{\mathbf{\Delta}}{\mathbf{T}_j}^H\mathbf{T}_j, \\ N_{d,j}\sigma^2_{\tilde{n}} + \delta\sigma^2_{\mathbf{\Delta}}\mathbf{T}^H_j\mathbf{T}_j + \bar{\mathbf{H}}^H_{d,j}\mathbf{T}^H_j\mathbf{T}_j\bar{\mathbf{H}}_{d,j})\end{aligned}, \tag{5}$$

where $\bar{\mathbf{H}}_{d,j} = [\widetilde{\mathbf{H}}^H_{d,1}\cdots\widetilde{\mathbf{H}}^H_{d,j-1}\widetilde{\mathbf{H}}^H_{d,j+1}\cdots\widetilde{\mathbf{H}}^H_{d,U_d}]^H$ and $\delta = U_d - 1$.

Based on general eigenvalue decomposition (GEVD) in [18], the objective function (5) can be simplified as follows. Thus, the optimal precoder for j^{th} user is given by

$$\begin{aligned}\mathbf{z}^o_{d,j} \propto max.\ eigenvector\ ((N_{d,j}\sigma^2_{\tilde{n}} + \delta\sigma^2_{\mathbf{\Delta}}\mathbf{T}^H_j\mathbf{T}_j + \\ \bar{\mathbf{H}}^H_{d,j}\mathbf{T}^H_j\mathbf{T}_j\bar{\mathbf{H}}_{d,j})^{-1}\widetilde{\mathbf{H}}^H_{d,j}\mathbf{T}^H_j\mathbf{J}_j\widetilde{\mathbf{H}}_{d,j} + \sigma^2_{\mathbf{\Delta}}\mathbf{T}^H_j\mathbf{T}_j).\end{aligned} \tag{6}$$

B. Uplink Transmission

In this subsection, the system model and proposed design for UL channel transmission will be presented as follows. The received signal vector $\mathbf{y}_u \in \mathcal{C}^{N_r\times 1}$ can be written as

$$\mathbf{y}_u = \mathbf{H}_{u,j}\mathbf{v}_{u,j}\mathbf{x}_{u,j} + \sum_{r\neq j}^{U_u}\mathbf{H}_{u,r}\mathbf{v}_{u,r}\mathbf{x}_{u,r} + \mathbf{H}_{si,j}\mathbf{x}_{si,j} + \mathbf{n}_u, \tag{7}$$

where $\mathbf{H}_{u,j} \in \mathcal{C}^{N_r\times N_{u,j}}$ is the channel matrices for user j, $\mathbf{n}_u \in \mathcal{C}^{N_r\times 1}$ refers to AWGN, $\mathbf{x}_{si,j} \in \mathcal{C}^{N_t\times 1}$ is the SI transmitted vector, and $\mathbf{H}_{SI,i}$ represents the SI channel matrix. The SI channel generated as $\mathbf{H}_{si,j} \sim \mathcal{CN}(\sqrt{\frac{\kappa_1\sigma^2_{si}}{\kappa_1+1}}\check{\mathbf{H}}_{si,j}, \frac{\sigma^2_{si}}{\kappa_1+1}\mathbf{I}_{N_rN_t})$, where the Rician factor $\kappa_1 = 1$ without loss of generality that is used to characterize the residual SI according to [1]. Additionally, σ^2_{si} is added to parameterize the efficiency of SI suppression mechanisms. The desired signal, multiuser interference, SI, and noise of the j^{th} user are denoted at the right-hand side of (7), respectively. The normalization of the precoder and intended signal vectors of the j^{th} user can be expressed as $||\mathbf{v}_{u,j}||^2 = 1$ and $\mathrm{E}[|\mathbf{x}_{u,j}|^2] = 1$, sequentially.

We next suggest the SLNR strategy, which minimizes SI while also maximizing the sum rate for a specific receiver in the presence of SI, employing the SI-plus-noise covariance

$$\mathbf{z}_{d,j}^{o} = \arg \max_{\mathbf{z}_{d,j} \in \mathcal{C}^{N_t \times 1}} \frac{||\mathbf{T}_j \widetilde{\mathbf{H}}_{d,j} \mathbf{z}_{d,j}||^2 + \sigma_{\mathbf{\Delta}}^2 ||\mathbf{T}_j \mathbf{z}_{d,j}||^2}{N_{d,j}\sigma_{\tilde{n}}^2 + (U_d - 1)\sigma_{\mathbf{\Delta}}^2 ||\mathbf{T}_j \mathbf{z}_{d,j}||^2 + \sum_{r \neq j}^{U_d} ||\mathbf{T}_j \widetilde{\mathbf{H}}_{d,r} \mathbf{z}_{d,j}||^2} \quad \text{subject to} \quad ||\mathbf{z}_{d,j}||^2 = 1, \qquad j = 1, \cdots, U_d. \tag{4}$$

matrix at the transmitter. The $N_r \times N_r$ SI-plus-noise covariance matrix of (7) is therefore represented as

$$\mathbf{A}_j = \mathbf{H}_{si,j} \mathbf{Q}_{si,j} \mathbf{H}_{si,j}^H + \sigma_n^2 \mathbf{I}_{N_r}, \tag{8}$$

where $\mathbf{Q}_{si,j} = \mathrm{E}[\mathbf{x}_{si,j} \mathbf{x}_{si,j}^H]$ is the covariance matrix of SI and $\mathrm{Tr}(\mathbf{Q}_{si,j}) = P_{si,j}$. Let us define a whitening matrix $\mathbf{L}_j = \mathbf{A}_j^{-1/2}$ and multiply (7) by $\mathbf{L}_j$, and we obtain the following expression which also shows the effect of ICSI.

$$\begin{aligned} \tilde{\mathbf{y}}_{u,j} =& \mathbf{L}_j \widetilde{\mathbf{H}}_{u,j} \mathbf{v}_{u,j} \mathbf{x}_{u,j} + \mathbf{L}_j \sum_{r \neq j}^{U_u} \widetilde{\mathbf{H}}_{u,r} \mathbf{v}_{u,r} \mathbf{x}_{u,r} \\ &+ \sum_{j=1}^{U_u} \mathbf{L}_j \mathbf{\Delta}_{u,j} \mathbf{v}_{u,j} \mathbf{x}_{u,j} + \tilde{\mathbf{n}}_{U,i} \end{aligned}, \tag{9}$$

where $\tilde{\mathbf{n}}_{u,j} = \mathbf{L}_j(\mathbf{H}_{si,j}\mathbf{x}_{si,j} + \mathbf{n}_{u,j})$ and this has covariance of identity. The components of $\mathbf{\Delta}_{u,j}$ distributed as circularly symmetric complex Gaussian i.i.d. with $\sigma_{\mathbf{\Delta}}^2$ variance and zero mean.

Using the leakage concept, the optimization problem is given in (10).

The precoder design based on the Rayleigh-Ritz quotient can be expressed by

$$\begin{aligned} &\frac{\mathbf{v}_{u,j}^H (\widetilde{\mathbf{H}}_{u,j}^H {\mathbf{L}_j}^H \mathbf{L}_j \widetilde{\mathbf{H}}_{u,j} + \sigma_{\mathbf{\Delta}}^2 {\mathbf{L}_j}^H \mathbf{L}_j) \mathbf{v}_{u,j}}{\mathbf{v}_{u,j}^H (\sigma_{\tilde{n}}^2 + (U_u - 1)\sigma_{\mathbf{\Delta}}^2 {\mathbf{L}_j}^H \mathbf{L}_j + \bar{\mathbf{H}}_{u,j}^H \mathbf{L}_j^H \mathbf{L}_j \bar{\mathbf{H}}_{u,j}) \mathbf{v}_{u,j}} \\ &\leq \gamma_{max} (\widehat{\mathbf{H}}_{u,j}^H {\mathbf{L}_j}^H \mathbf{L}_j \widehat{\mathbf{H}}_{u,j} + \sigma_{\mathbf{\Delta}}^2 {\mathbf{L}_j}^H \mathbf{L}_j, \\ &\sigma_{\tilde{n}}^2 + (U_u - 1)\sigma_{\mathbf{\Delta}}^2 \mathbf{L}_j^H \mathbf{L}_j + \bar{\mathbf{H}}_{u,j}^H \mathbf{L}_j^H \mathbf{L}_j \bar{\mathbf{H}}_{u,j}) \end{aligned} \tag{11}$$

which $\bar{\mathbf{H}}_{u,j} = [\widetilde{\mathbf{H}}_{u,1}^H \cdots \widetilde{\mathbf{H}}_{u,j-1}^H \widetilde{\mathbf{H}}_{u,j+1}^H \cdots \widetilde{\mathbf{H}}_{u,U_u}^H]^H$.

The optimal precoder using GEVD is given by

$$\begin{aligned} \mathbf{v}_{u,j}^{o} \propto\ & max.\ eigenvector\ ((\sigma_{\tilde{n}}^2 + (U_u - 1)\sigma_{\mathbf{\Delta}}^2 \mathbf{L}_j^H \mathbf{L}_j \\ &+ \bar{\mathbf{H}}_{u,j}^H \mathbf{L}_j^H \mathbf{L}_j \bar{\mathbf{H}}_{u,j})^{-1} \widetilde{\mathbf{H}}_{u,j}^H \mathbf{L}_j^H \mathbf{L}_j \widetilde{\mathbf{H}}_{u,j} + \sigma_{\mathbf{\Delta}}^2 \mathbf{L}_j^H {\mathbf{L}_j}^H). \end{aligned} \tag{12}$$

Accordingly, the total spectral efficiency of FD-MU-MIMO mode is calculated, and given by

$$\mathrm{R}_{FD} = \mathrm{R}_d + \mathrm{R}_u, \tag{13}$$

where R_d designates the sum rate of the DL channel obtained using an optimized precoder in (6). As well, R_u represents the sum rate of the uplink channel by substituting the optimal precoder in (12). The complete spectral effectiveness of half-duplex communication corresponds to the combination of the DL total rate and the attainable UL total rate without any SI.

III. Simulation Results and Discussion

We simulate FD-multiuser MIMO under Rician flat fading conditions with channels suffering from channel error. In this study, the maximum transmitted power at every user is assumed to be equal for all users in which $P_{ul,j} = P_{ul}$. As well, the noise power is expressed as $\sigma_{n,j}^2 = \sigma_{\tilde{n}}^2 = 1$ for $j = 1, 2, \cdots, U$, where U stand for downlink or uplink channel users. It is assumed that the number of users is two per channel. The total number of antennas at the half-duplex system is taken $N_T = N_t + N_r$. We assume that the power transmitted at the base station P_{tr} = 27 dBm and at uplink channel users P_{ul} = 23 dBm. We acquired the results by averaging 1000 realizations per channel.

Figure 1 presents the achievable SE of the full-duplex and half-duplex system by varying imperfect CSI scenarios for different Rician factors. Scenario 1: we vary the Rician factor $\kappa = 1$ and $\kappa = 5$ and set the number of antennas at the BS $N_t = N_r = 6$ and at both channel users $N_{d,j} = N_{u,j} = 3$. Scenario 2: we increase the error variance from $\sigma_{\Delta}^2 = 0.01$ to $\sigma_{\Delta}^2 = 0.5$. For both scenarios, this figure confirms the significant improvement in SE of full-duplex over half-duplex for a given transmission power. However, the total SE of the HD and FD systems decreases significantly as the Rician coefficient increases. This is because the Rician fade effect resembles an AWGN channel. In addition, a strong LOS link reduces channel capacity, i.e. MIMO communication merits a richly distributed environment. Although the SE of the FD-MU-MIMO mode system decreases as the Rician coefficient increases, even now it still provides good performance compared to the HD MU-MIMO mode of communication system.

Figure 2 shows the performance comparisons of the SLNR, BD, and ZF precoding schemes under the imperfect CSI condition when the number of antennas at the base station varies. We observe that as N_t increases from $N_t = 4$ to $N_t = 8$, the SEs for SLNR, BD, and ZF precoding schemes significantly improve. According to the figure, SLNR-PS provides a significant performance improvement over both methods. This implies that the proposed SLNR-based precoding not only attains good SE gain by compromising the interference and noise but also can overcome the dimensional constraints on ZF and BD precoding schemes and make use of all the channel degrees of freedom. Thus, we can conclude that the SLNR-PS provides an effective solution to achieve the optimum SE with a low-computational complexity and it performs well.

IV. Conclusion

In this paper, an SLNR scheme has been proposed to mitigate co-channel interference in the context of downlink

$$\mathbf{v}_{u,j}^{o} = \arg \max_{\mathbf{v}_{u,j} \in \mathcal{C}^{N_{u,j} \times 1}} \frac{||\mathbf{L}_j \widetilde{\mathbf{H}}_{u,j} \mathbf{v}_{u,j}||^2 + \sigma_{\mathbf{\Delta}}^2 ||\mathbf{L}_j \mathbf{v}_{u,j}||^2}{\sigma_{\tilde{n}}^2 + (U_u - 1)\sigma_{\mathbf{\Delta}}^2 ||\mathbf{L}_j \mathbf{v}_{u,j}||^2 + \sum_{r \neq j}^{U_u} ||\mathbf{L}_j \widetilde{\mathbf{H}}_{u,j} \mathbf{v}_{u,j}||^2}. \quad (10)$$
$$\text{subject to} \quad ||\mathbf{v}_{u,j}||^2 = 1, \qquad j = 1, \cdots, U_u.$$

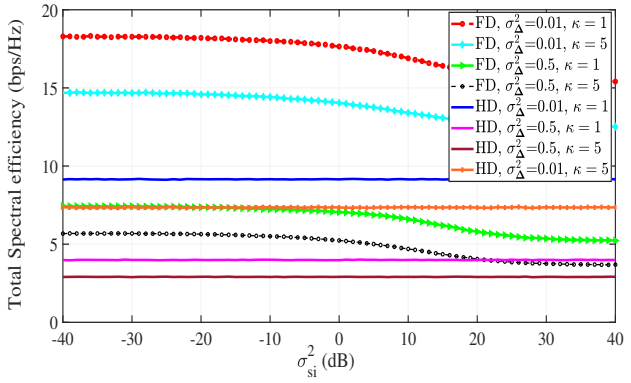


Fig. 1. Total spectral efficiency versus SI σ^2_{si} for FD and HD mode of operation.

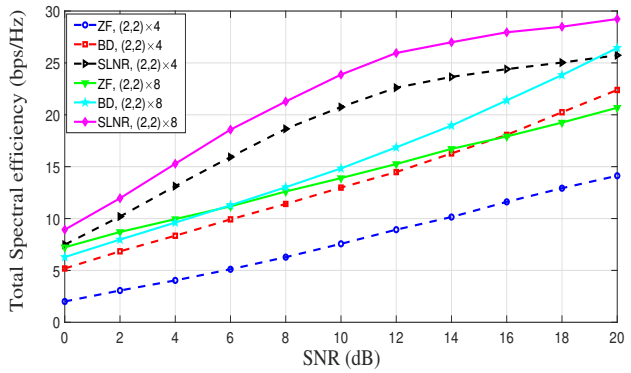


Fig. 2. Total spectral efficiency versus SNR of different precoding schemes for $\sigma^2_{si} = -80$ dB, Rician factor $\kappa = 1$ and $\sigma^2_{\Delta} = 0.01$.

transmission while also maximizing the sum rate. This is achieved through the utilization of only the CCI plus noise covariance matrix at the base station, which is used in conjunction with a specified receiver structure. Furthermore, we present another methodology for uplink communication that takes into consideration SI. To enable interference cancellation, a model incorporating a whitening filter at the receiver and a precoder design based on the covariance of noise and SI is employed. Consequently, the issue of enhancing precoding coefficients is depicted as the maximization of the SLNR-PS for all users concurrently. The objective function is optimized through the utilization of the general eigenvalue decomposition approach. Upon equating the optimized beamformer to the sum rates of both the DL and UL systems, the total sum spectral efficiency of the overall system can be attained. The present investigation involves the assessment of the efficacy of the FD and HD systems within the context of a small-cell system through the implementation of the proposed precoding schemes.

REFERENCES

[1] M. Duarte, C. Dick, and A. Sabharwal, "Experiment-driven characterization of full-duplex wireless systems," *IEEE Trans. Wireless Commun.,* vol. 11, no. 12, pp. 4296-4307, Dec. 2012.

[2] J. I. Choi, M. Jain, K. Srinivasan, P. Levis and S. Katti, "Achieving single channel, full duplex wireless communication," *MobiCom 2010, ACM*, Chicago, Illinois, USA, pp. 1-12, Sept. 2010.

[3] S. Chen, M. A. Beach and J. P. McGeehan, "Division-free duplex for wireless applications," *Electronics Letters,* vol. 34, no. 2, pp. 147-148, Jan. 1998.

[4] B. Day, A. Margetts, D. Bliss, and P. Schniter, "Full-duplex bidirectional MIMO: Achievable rates under limited dynamic range," *IEEE Trans. Signal Process.,* vol. 60, no. 7, pp. 3702-3713, Jul. 2012.

[5] L. Irio and R. Oliveira, "Distribution of the residual self-interference power in in-band full-duplex wireless systems," *IEEE Access,* vol. 7, pp. 57516-57526, 2019.

[6] V. Singh, A. Gadre, and S. Kumar, "Full duplex radios: Are we there yet?" *in Proc. 19th ACM Workshop Hot Topics Netw.*, pp. 375–386, Nov. 2020

[7] M. Dibaei and A. Ghaffari, "Full-duplex medium access control protocols in wireless networks: a survey," *Springer Wireless Networks*, pp. 1–19, 2020.

[8] K. Komatsu, Y. Miyaji and H. Uehara, "Theoretical Analysis of In-Band Full-Duplex Radios With Parallel Hammerstein Self-Interference Cancellers," *IEEE Transactions on Wireless Communications,* vol. 20, no. 10, pp. 6772-6786, Oct. 2021.

[9] D. Nguyen, L.-N. Tran, P. Pirinen, and M. Latva-aho, "On the spectral efficiency of full-duplex small cell wireless systems," *IEEE Trans. Wireless Commun.,* vol. 13, no. 9, pp. 4896-4910, Sep. 2014.

[10] S. Huberman and T. Le-Ngoc, "Full-duplex MIMO precoding for sum-rate maximization with sequential convex programming," *IEEE Transactions on Vehicular Technology,* vol. 64, no. 11, pp. 5103-5112, 2015.

[11] M. B. Teklu, W.-X. Meng, C. Li, and L. Chen, "Spectral efficiency evaluation of full-duplex mode of communications based on SLNR approach," *Int. J. Commun. Syst.,* vol. 32, no. 15, p. e4006, Oct. 2019.

[12] M.-J. Kim and Y.-C. Ko, "Sequential transceiver design based on block diagonalization and interference alignment for full duplex multiuser MIMO system," in *Int. Conf. Info. Commun. Technol. Conv. (IEEE ICTC 2014),* pp. 516-517, Jun. 2014.

[13] J. Mirza, G. Zheng, S. Saleem and K. Wong, "Optimization of Uplink CSI Training for Full- Duplex Multiuser MIMO Systems," in *IEEE Communications Letters,* vol. 23, no. 12, pp. 2325-2329, Dec. 2019.

[14] Maaref A. and Aissa S. Capacity of MIMO rician fading channels with transmitter and receiver channel state information. *IEEE Trans. Wireless Commun.* 2008; 7(5): 1687-1698.

[15] H. Vaezy, M. M. Naghsh, M. J. Omidi, and E. H. M. Alian, "Efficient transmit covariance design in MIMO interference channel," *IEEE Trans. Veh. Technol.,* vol. 67, no. 7, pp. 5793-5805, Jul. 2018.

[16] M. Sadek, A. Tarighat, and A. Sayed, "A leakage-based precoding scheme for downlink multi-user MIMO channels," *IEEE Trans. Wireless Commun.,* vol. 6, no. 5, pp. 1711-1721, May. 2007.

[17] M. Sadek, A. Tarighat, and A. Sayed, "Active antenna selection in multiuser MIMO communications," *IEEE Trans. Sig. Proc.* vol. 55, no. 4, pp.1498-1510, Apr. 2007.

[18] M. Schubert and H. Boche, "Solution of the multiuser downlink beamforming problem with individual SINR constraints," *IEEE Trans. Vehic. Tech.,* vol.53, no.1, pp.18-28, Jan. 2004

Linearly Scaling Achievable Rates through Cell-Free Massive MIMO Systems

Keunyoung Kim
Mobile Communication Research Division
Electronics and Telecommunications Research Institute
Daejeon, Korea
kykim12@etri.re.kr

Woo Yong Lee
Mobile Communication Research Division
Electronics and Telecommunications Research Institute
Daejeon, Korea
wylee@etri.re.kr

***Abstract*—This paper demonstrates that the spatial DoF can be effectively converted into linearly scaling achievable rates by deploying cell-free massive multiple input multiple out systems when the antennas are densely and evenly distributed over the service area and interference suppressing schemes are adopted.**

***Index Terms*—degrees of freedom, capacity, cell-free massive MIMO**

I. Introduction

Multiple input Multiple out (MIMO) systems provide the spatial dimension which can be leveraged as the dimensions of time and frequency. It is well-established that the capacity of point-to-point MIMO systems scales linearly with the minimum number of antennas between the transmitter and receiver at high signal to noise ratio (SNR) [1]. That is, in additive white Gaussian noise channel with M inputs and N outputs, the capacity can be shown to be $\min(M, N)\log(\text{SNR})$ at high SNR. In multi-user (MU) MIMO systems, which encompass MIMO multiple access channels (MAC) and MIMO broadcast channels (BC), the capacity exhibits a linear growth pattern as the number of antennas or users is augmented.

The factor in the capacity formula that undergoes linear multiplication is referred to as the spatial degrees of freedom (DoF) [2]. The linear increase in capacity with the spatial DoF suggests that the achievable rate for each user may remain constant even as the number of users increases. However, in most common wireless communication environments, the achievable rate tends to decrease as the number of users increases. This is primarily because each user does not uniformly experience high SNR regardless of their locations, and the presence of interference typically leads to a degradation in achievable rates.

This paper demonstrates that the spatial DoF can be effectively converted into linearly scaling achievable rates by deploying distributed antennas evenly and effectively managing interference. Such a translation can be realized in cell-free (CF) MIMO systems.

This work was supported by the Korea Institute of Marine Science & Technology Promotion (KIMST) funded by the Korea government (MSIT) in 2023 (No. 2021-0626, Development of Polar Region Communication Technology and Equipment for Internet of Extreme Things (IoET))

II. Degrees of Freedom and CF massive MIMO

The definition of the spatial DoF is given as [2]:

$$\eta \triangleq \lim_{\rho\to\infty} \frac{C_\Sigma(\rho)}{\log(\rho)}, \tag{1}$$

where $C_\Sigma(\rho)$ is the sum capacity at SNR ρ. In MU-MIMO systems equipped with M base station antennas and N single antenna users, $\eta = \min(M, N)$.

CF MIMO systems are equipped with a large number of access points (APs), and each AP can be uniformly deployed over the service area. This allows CF MIMO systems to achieve a large number of degrees of freedom (DoFs), which is equal to the number of users. Additionally, the close-in distance between an AP and a user can provide macro diversity gain, which further improves the system performance [3].

The key ingredient to achieve linearly scaling achievable rate stems from the DoF gain linearly scaling up as the number of users. However, in most wireless communication system, the achievable rate decreases as the number of users increases. This is due to the fact that the SNR value largely varies depending on locations and the interference from other users is treated as noise.

The key to achieving linearly scaling achievable rates is the DoF gain, which scales linearly with the number of users. However, in most wireless communication systems, the achievable rate decreases as the number of users increases. This is because the SNR varies significantly depending on the user's location, and the interference from other users is treated as noise. In CF massive MIMO systems, APs can be densely deployed, which results in a short distance between BSs and users. This leads to a uniformly high SNR for all users. Additionally, coherent signal processing is used to manage interference. As a result of the uniformly high SNR and interference management, linearly scaling achievable rates are possible in real wireless systems. This means that the rate of each user does not decrease even if the number of users increases or users are located at the edge of each AP's coverage area.

The linearly scaling achievable rate was demonstrated in a simulation. The simulation considered 40 or 400 APs deployed over a 120 m by 50 m square area on the ceiling. Fig. 1 shows

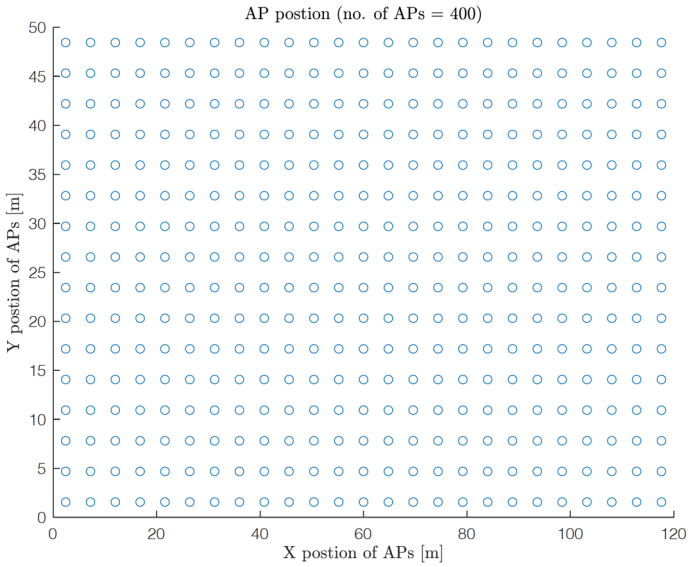


Fig. 1. Location of APs when the number of APs is 400.

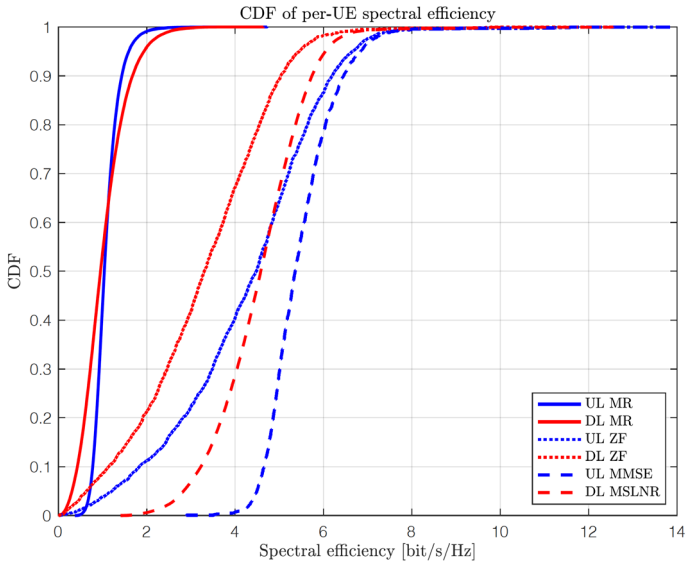


Fig. 2. Spectral efficiency of each user when the number of APs is 40.

the explicit locations of 400 APs. Users are randomly located over the square.

Beamforming schemes are chosen to maximize the received power, minimize the received interference, or to maximize the signal to interference and noise ratio. The simulation parameters and setup are largely based on the indoor hotspot model in the ITU-R guideline document for the evaluation of 5G technologies [4].

Figs. 2 and 3 show the spectral efficiency of each user when the numbers of APs are 40 and 400, respectively. UL MR, DL MR, UL ZF, DL ZF, UL MMSE, and DL MSLNR represent uplink maximum ratio combining, downlink maximum ratio transmission, uplink zero-forcing, downlink zero-forcing, uplink signal to interference and noise maximization, and downlink signal to leakage and noise maximization schemes respectively.

III. Discussions

The very steep slope of the lines in Fig. 3 verifies the spectral efficiency is almost the same regardless of user's location. When the number of users is the same as that of APs, the performances of ZF is less than those of MMSE or MSLNR. This is because null space dimension is used up for cancelling interference, thus no dimension is not available for maximizing signal power. When the number of APs is much more that that of users, the performance gap between ZF and MMSE is negligible. The performance of MR is much less that those of ZF and MMSE/MSLNR, which indicates that MR does not effectively suppress interference in CF massive MIMO systems. When a massive number of antennas are co-located in a base station, MR processing is expected to provide interference management as a consequence of law of large numbers. However, this positive effect does not seem to be observable in CF massive MIMO systems. The results of varying number of users are skipped in this paper.

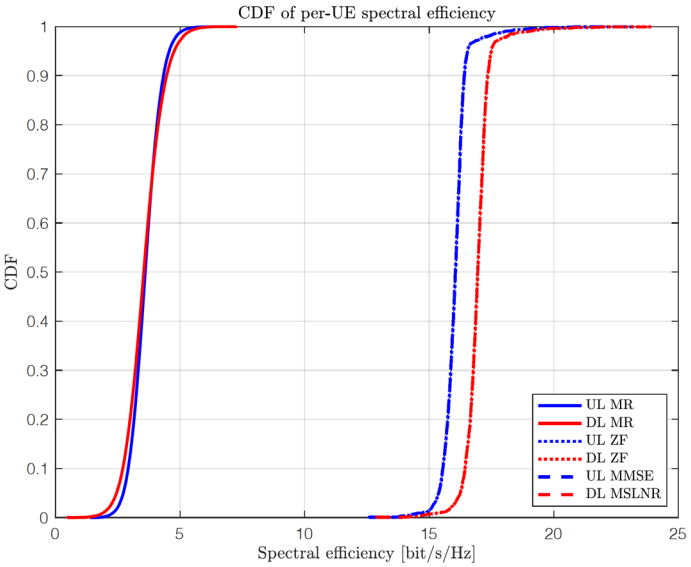


Fig. 3. Spectral efficiency of each user when the number of APs is 400.

With different numbers of users and APs, the linearly scaling achievable rates is shown to be possible in [3].

References

[1] E. Telatar, "apacity of multi-antenna Gaussian channels," European Trans. on Telecomm. ETT, vol. 10, pp. 585?596, November 1999.

[2] S. A. Jafar and M. J. Fakhereddin, "Degrees of freedom for the MIMO interference channel," IEEE Trans. Inf. Theory, vol. 53, no. 7, pp. 2637-2642, Jul. 2007.

[3] K. Kim, J. Myung, Y. -J. Ko and W. Y. Lee, "Beamforming and Power Optimization for User Fairness in Cell-Free MIMO Systems," in IEEE Access, vol. 11, pp. 51032-51046, 2023, doi: 10.1109/ACCESS.2023.3278795.

[4] Report ITU-R M.2412-0, ?Guidelines for evaluation of radio interface technologies for IMT-2020,? ITU-R, Report, Oct. 2017.

Multi-Feature Concatenation for Speech Dependent Automatic Speaker Identification in Maritime Autonomous Vehicles

Judith Nkechinyere Njoku *, Cosmas Ifeanyi Nwakanma†, Jae-Min Lee *, and Dong-Seong Kim *
*IT Convergence Engineering, † ICT Convergence Research Center, Kumoh National Institute of Technology, Korea.

***Abstract*—Automatic speaker identification is a crucial task that can help ensure safety in the use of maritime autonomous vehicles (MAVs). Most of the best solutions to automatic speaker identification tasks have been performed using speech. One challenge is that speech features that work perfectly for speech tasks may not be suitable for automatic speaker identification tasks. As a result, the selection of the best features to employ is critical. To address this issue, this work explores multiple speech features based on different configurations and trains a K-nearest neighbor algorithm using these features. The best performance was obtained from a concatenation of gammatone cepstral coefficients, Mel frequency cepstral coefficients, and pitch. This configuration achieved a validation accuracy of 0.98 and 98.76% and a test accuracy of 0.76 and 84.54% based on the Jaccard similarity index and F1 score metrics, respectively.**

***Index Terms*—automatic speaker identification, maritime autonomous vehicles, features, KNN, concatenation**

I. Introduction

In recent years, there has been a tremendous rise in the development and deployment of autonomous vehicles in diverse domains, including the maritime sector [1], [2]. Maritime autonomous vehicles (MAVs) offer a lot of benefits, including a reduction in operational costs, improved safety, and enhanced efficiency [2], [3]. With this advancement comes a major concern: Security. One security measure that has been explored is the use of automatic speaker identification (ASI) systems, which can identify users based on specific features in their voices [4], [5]. Although ASI systems are majorly based on the users' speech, the task of speaker identification should not be confused with speech recognition. While speech recognition systems identify the words said by the user, ASI systems identify the user itself. Research shows that certain speech features which might be suitable for speech recognition tasks, may not thrive well when used for speaker identification tasks. As a result, it is important that the best features for ASI tasks be carefully selected [6].

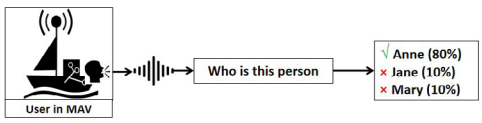


Fig. 1. Typical speaker identification in maritime autonomous vehicles

ASI is very important in MAV, as it helps to ensure that only authorized users can control the vehicle or, where necessary, access it [7], as illustrated in Fig. 1. Machine learning (ML) has been extensively applied to speech recognition and ASI tasks, with impressive results. To develop a robust ASI ML model, there are so many things to consider; including the feature extraction method. Feature extraction plays a key role in the design of the ASI system, as due to the variety of speeches made by different genders, certain features might prove redundant compared to others [8]. Moreover, considering that every individual's vocal tract has a different shape which helps to identify them when listening, it is important for an ideal ASI system to be able to perform its task based on some unique features.

Mel frequency cepstral coefficients (MFCC) are features that are frequently used for speech-related tasks. MFCC has been successfully applied for ASI tasks as presented in [9]. The same features were also employed in [10], along with timbre features, to identify speakers when the speech is in whispers. [11] also employed MFCC features for text-independent ASI.

Motivation - All these works have shown the robustness of MFCC features in ASI systems. However, gammatone cepstral coefficients (GTCC) [12] have been shown to be more robust to noise than MFCC [13] and have thus been preferred for ASI [13]. One other feature commonly applied for ASI is *pitch* [14]. The pitch has been used to improve the performance of ASI tasks in [15], where hybrid features comprising pitch and MFCC were used to develop a more robust ASI system.

Contribution - To overcome these above-listed limitations, and enhance the robustness of speaker identification in MAVs, this paper investigates the effectiveness of multi-feature concatenation. The main objective is to experiment with a hybrid of other prospective speech features and select the best-performing ones for ASI. An investigation and experimentation of the possible concatenation of multiple features were conducted for the development of a robust ASI system. The performances of several configurations of multiple feature concatenations were compared based on the Jaccard index, F1 score, and processing time.

II. Methodology

A. Problem Formulation

Consider a maritime autonomous vehicle $\mathcal{MAV}$ with a set of authorized users $\mathcal{N} = \{1, \ldots, N\}$ and a set of unauthorized users $\mathcal{M} = \{1, \ldots, M\}$, as illustrated in Fig. 2.

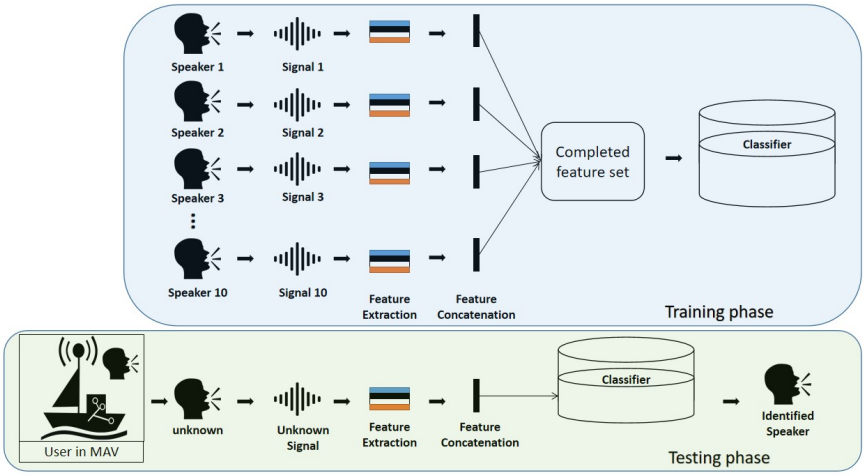


Fig. 2. Proposed system model for automatic speaker identification in maritime autonomous vehicles

It is assumed that users can issue speech control prompts to $\mathcal{MAV}$, and $\mathcal{MAV}$ can, in turn, recognize those prompts and take action. To ensure safety, it is important that $\mathcal{MAV}$ takes instructions from only users in the $\mathcal{N}$ set. Given that all users from sets $\mathcal{N}$ and $\mathcal{M}$ issue speech commands with speech signal x, where x has a frequency that falls between $\mathcal{H} = \{h, \ldots, H\}$, and $\mathcal{F} = \{f, \ldots, F\}$ features can be generated to represent x, we term the users *speakers*, and formulate the closed set ASI problem. This problem requires an algorithm A that can be embedded into $\mathcal{MAV}$ after being trained on the best features from the set $\mathcal{F}$, obtained from a signal x from speakers in the set $\mathcal{N}$, to accurately recognize speakers only from this set with the minimum error and in the shortest possible time.

B. Feature Extraction

Feature extraction aims to convert the speech signals to a format that clearly shows the distinctive attributes of the speaker's voice. As illustrated in Fig. 2, a typical ASI system is composed of two modes: the model training mode and the model testing mode or recognition mode. These two modes both contain a feature extraction unit. In this unit, the feature extractor transforms the speech signals into feature vectors which are numerical samples that will be fed to the training model. In this paper, five different features were employed.

1) Mel Frequency Cepstral Coefficients (MFCC): MFCCs are critical features for ASI tasks. They are extracted through a cepstral analysis of the speech signal. The cepstral analysis of a speech signal separates the source components into excitation source and vocal tract source.

2) Gammatone Cepstral Coefficients (GTCC): The gammatone (GT) filter banks are an improvement over mel scale triangular filters. The GT filter is a linear filter that is represented by an impulse response $g(t)$ given by:

$$g(t) = at^{n-1}e^{-2\pi bt}\cos(2\pi ft + \phi), \tag{1}$$

where ϕ represents the phase of the carrier in $radians$ and f is denotes the center frequency in Hz, n is the order of the filter, a is the amplitude, t is the time, and b is the bandwidth of the filter in Hz.

3) Pitch: The pitch can be generated by first applying a pre-emphasis filter to the speech signal, which would enhance the high-frequency content of the signal. When Pitch features are employed, short-term energy (StE) and zero crossing rate (ZCR) features are used to help differentiate speech from silence and unvoiced from voiced speech respectively.

C. Automatic Speaker Identification

1) Dataset: The dataset employed was obtained from the *"Common voice"* dataset from Mozilla and consists of 10 speakers with a uniform distribution between males and females. Each speaker utters short sentences at a signal frequency of $48kHz$. This dataset was suitable for this work, as it registers different races, sexes, and accents. As shown in Table I, a total of $28,844$ speech samples were used in the experiment.

2) System Model: This work conducted experiments on different sets of fused features for speaker identification. As illustrated in Fig. 2, the experiments were conducted in two phases: the model training phase and the model testing phase.

The performance of five different single and concatenated feature configurations was compared based on the accuracy and time of validation and testing. The following paragraphs detail the different hybrid feature configurations.

i. GTCC features: The GTCC features were extracted using a Hamming window of length $0.03 \times samplerate$ and an overlap length of $0.025 \times samplerate$.
ii. MFCC features: The MFCC features were extracted using a Hamming window of length $0.03 \times samplerate$ and an overlap length of $0.025 \times samplerate$.
iii. GTCC-Pitch features: This hybrid feature consists of the GTCC, pitch, short-time energy, and ZCR features. The MFCC features were extracted using a Hamming window of length $0.03 \times samplerate$ and an overlap length of $0.025 \times samplerate$. The StE features were extracted

TABLE I
STATISTICS OF DATASET

Dataset / Speaker index	1	2	3	4	5	6	7	8	9	10	Total
Training data	3094	1424	4339	2128	2661	2904	2290	968	1857	1456	**23,121**
Test data	756	441	1029	525	732	673	461	226	481	399	**5,723**
Total dataset	3850	1835	5368	2653	3393	3577	2751	1194	2338	1855	**28,844**

with a threshold of 0.005, while the ZCR features were extracted with a threshold of 0.2.

iv. MFCC-Pitch features: This hybrid feature consists of the MFCC, pitch, StE, and ZCR features. The MFCC features were extracted using a Hamming window of length $0.03 \times samplerate$ and an overlap length of $0.025 \times samplerate$. The energy features were extracted with a threshold of 0.005, while the ZCR features were extracted with a threshold of 0.2.

v. GTCC-MFCC-Pitch features: This hybrid feature was made up of GTCC, MFCC, and Pitch features along with the ZCR and energy features.

Since the ZCR and StE features are used to decide when to use the pitch feature, it was only employed for hybrid features involving pitch.

3) Feature Concatenation:

i. Normalization: To prevent bias in the classifier, all features were normalized by deducting the mean and dividing by the standard deviation, thus ensuring that they are all on the same scale.

ii. Concatenation: All feature vectors were concatenated, and the output is a matrix, which serves as an input to the KNN classifier. In this matrix, each of the columns matches a particular feature.

D. K-Nearest Neighbor

The K-nearest neighbor (KNN) algorithm is a nonparametric classifier that uses a supervised learning approach to perform its classification tasks. KNN is used to identify the nearest neighbors to a given point based on distance metrics. This study adopted the Euclidean distance as denoted below.

$$d(x,y) = \sqrt{\sum_{i=1}^{n} (y_i - x_i)^2} \tag{2}$$

We also used $k = 5$ neighbors and a *squared inverse* distance weight.

III. Model Training & Evaluation

A. Simulation details

All systems were simulated in MATLAB. The dataset was imported using a helper function that helps to download and organize the data. The dataset was split into two parts. 80% was used for training, while the remaining 20% was used for testing. All models were trained using a k-fold cross-validation approach, where $k = 5$. This divides the dataset into folds and ensures that each fold is used to test the models at intervals during the model-training process. A total of $23,121$ data points were used for training the models. Optimizing validation and test accuracy, the number of neighbors was set to 5, and the squared-inverse weighted Euclidean distance metric was employed. In the testing phase, the trained KNN classifier was tested using test speech signals from each of the speakers. Like in the training phase, the features were extracted and normalized before the KNN model is tested. The speaker for each frame was then predicted.

B. Results

1) Metrics: The results of the models were evaluated based on three metrics: F1 score, Jaccard index, and processing time. The F1 score measures the accuracy of the validation and test sets, while the Jaccard similarity index is usually used to estimate the similarity between two sets of data, in this case, the predicted and the true label (speaker). The F1 score is calculated using the statutory formula defined:

$$F_1 = \frac{2 \times Precision \times Recall}{Precision + Recall} \tag{3}$$

The Jaccard index score is calculated using the following formula:

$$J(y,\hat{y}) \frac{|y \cap \hat{y}|}{|y| + |\hat{y}| - |y \cap \hat{y}|} = \frac{\hat{y}}{y + y - \hat{y}}, \tag{4}$$

where y is the total number of samples, and $\hat{y}$ is the number of correctly predicted samples.

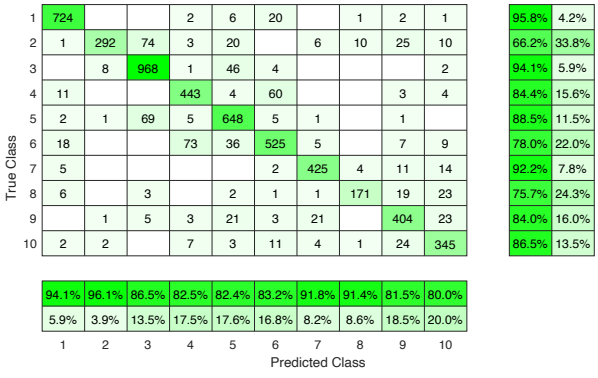


Fig. 3. Test confusion matrix of best performing concatenated features

2) ASI Result: Table II and Fig. 3 represent the results from the experiments. The best-performed concatenated configuration among all the implemented models was the GTCC-MFCC-Pitch feature concatenation, which scored the highest F1 score and Jaccard similarity index of 98.76% and 0.98 on the validation set and 84.54% and 0.76 on the test set, respectively, as illustrated in Table II.

TABLE II
F1 SCORE, JACCARD INDEX, AND COMPUTATION TIME RESULTS OF ALL FEATURE CONCATENATION CONFIGURATIONS

Metrics & Features	F1 Score (%)		Jaccard Index		Computation time (s)	
	Validation	Testing	Validation	Testing	Validation	Testing
GTCC [13], [16]	96.99	69.29	0.94	0.56	**0.15**	**0.251**
MFCC [9]	97.66	81.5	0.95	0.71	1.56	0.26
GTCC + Pitch	96.82	71.23	0.94	0.58	0.18	0.26
MFCC+ Pitch [15]	97.73	82.23	0.96	0.73	0.44	0.29
MFCC + GTCC + Pitch	**98.76**	**84.54**	**0.98**	**0.76**	0.29	0.50

Table II shows that the best computation time was obtained when only GTCC features were used, for both validation and testing. However, the F1 score and Jaccard index for this feature are too low. Although the best-performing configuration of MFCC, GTCC, and Pitch have a high computation time, the accuracy is quite high when compared with the other approaches.

In Fig. 3, the identification results for test identification of speakers are represented. It indicates that all accuracy was above 97%, with the best being 99.7% for speaker 1, and the least being 97.3% for speaker 4.

IV. CONCLUSION

This study was aimed at comparing the performance of different speech features including MFCC, GTCC, pitch, energy, and zero crossing rate, when used independently and when fused together, for the task of automatic speaker identification for MAVs. The configuration with MFCC, GTCC, Pitch, short-term energy, and zero crossing rate performed best. It however incurs a high computation time when compared to the other configurations. It is expected that the experiments conducted will contribute towards further study on speech feature extraction for ASI tasks within and outside the MAVs domain. Future works will focus on developing a real-time ASI model with considerably low computation time and appreciable accuracy on these selected features.

ACKNOWLEDGMENTS

This research was supported by Priority Research Centers Program through the NRF funded by the MEST (2018R1A6A1A03024003), MSIT Korea (1711175292/2022-IT-RD-0084-01), and by MSIT under the Innovative Human Resource Development for Local Intellectualization support program (IITP-2023-2020-0-01612) supervised by the IITP.

REFERENCES

[1] J. N. Njoku, G. O. Anyanwu, I. S. Igboanusi, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "State-of-the-art object detectors for vehicle, pedestrian, and traffic sign detection for smart parking systems," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 1585–1590.

[2] M. Barbier, E. Bensana, and X. Pucel, "A generic and modular architecture for maritime autonomous vehicles," in *2018 IEEE/OES Autonomous Underwater Vehicle Workshop (AUV)*, 2018, pp. 1–6.

[3] G. Zhang, S. Liu, X. Zhang, and W. Zhang, "Event-triggered cooperative formation control for autonomous surface vehicles under the maritime search operation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 11, pp. 21 392–21 404, 2022.

[4] L. Tan, K. Yu, L. Lin, X. Cheng, G. Srivastava, J. C.-W. Lin, and W. Wei, "Speech emotion recognition enhanced traffic efficiency solution for autonomous vehicles in a 5g-enabled space–air–ground integrated intelligent transportation system," *IEEE Transactions on Intelligent Transportation Systems*, vol. 23, no. 3, pp. 2830–2842, 2022.

[5] A. M. Pfalzgraf, C. Sullivan, and D. S. Sánchez, "Autonomous vehicle speaker verification system," in *Bradley University Electrical and Computer Engineering Senior Capstone projects*, 2014. [Online]. Available: http://ee.bradley.edu/projects/proj2014/avsvs/AVSVS_Final_Paper.pdf

[6] J. N. Njoku, C. I. Nwakanma, and D.-S. Kim, "Evaluation of spectrograms for keyword spotting in control of autonomous vehicles for the metaverse," in *Proceedings of Symposium of the Korean Institutes of Communications and Information Sciences (KICS)*, vol. 78, 2022, pp. 1777–1778. [Online]. Available: https://journal-home.s3.ap-northeast-2.amazonaws.com/site/2022s/abs/0086.pdf

[7] R. S. Lakshmi, S. Veena, and B. R. Reddy, "Development of continuous automatic speech recognition system for controlling of mavs through natural speech," in *2017 International Conference on Recent Advances in Electronics and Communication Technology (ICRAECT)*, 2017, pp. 144–148.

[8] S. S. Tirumala, S. R. Shahamiri, A. S. Garhwal, and R. Wang, "Speaker identification features extraction methods: A systematic review," *Expert Systems with Applications*, vol. 90, pp. 250–271, 2017. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0957417417305535

[9] F.-Y. Leu and G.-L. Lin, "An mfcc-based speaker identification system," in *2017 IEEE 31st International Conference on Advanced Information Networking and Applications (AINA)*, 2017, pp. 1055–1062.

[10] V. M. Sardar and S. D. Shirbahadurkar, "Speaker identification of whispering speech: an investigation on selected timbrel features and knn distance measures," *International Journal of Speech Technology*, vol. 21, pp. 545–553, 2021.

[11] M. Soleymanpour and H. Marvi, "Text-independent speaker identification based on selection of the most similar feature vectors," *International Journal of Speech Technology*, vol. 20, pp. 99–108, 2017.

[12] X. Valero and F. Alias, "Gammatone cepstral coefficients: Biologically inspired features for non-speech audio classification," *IEEE Transactions on Multimedia*, vol. 14, no. 6, pp. 1684–1689, 2012.

[13] X. Zhao and D. Wang, "Analyzing noise robustness of mfcc and gfcc features in speaker identification," in *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 7204–7208.

[14] J. Guglani and A. Mishra, "Automatic speech recognition system with pitch dependent features for punjabi language on kaldi toolkit," *Applied Acoustics*, vol. 167, p. 107386, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0003682X20304904

[15] M. A. Nasr, M. Abd-Elnaby, A. S. El-Fishawy, S. El-Rabaie, and F. E. Abd El-Samie, "Speaker identification based on normalized pitch frequency and mel frequency cepstral coefficients," *International Journal of Speech Technology*, vol. 21, pp. 941–951, 2018.

[16] B. Ayoub, K. Jamal, and Z. Arsalane, "Gammatone frequency cepstral coefficients for speaker identification over voip networks," in *2016 International Conference on Information Technology for Organizations Development (IT4OD)*, 2016, pp. 1–5.

Leveraging Deep Learning for Anomaly Detection in AIS for Secured Maritime Navigation

Urslla Uchechi Izuazu, Vivian Ukamaka Ihekoronye, Dong-Seong Kim, Jae Min Lee
Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea
(uursla8, vivian, dskim, ljmpaul)@kumoh.ac.kr

***Abstract*—Automatic Identification System (AIS) was introduced in the maritime domain to enhance sea traffic safety by broadcasting messages containing vessel identification, position, and speed, so as to help prevent collisions and improve situational awareness. However, the dynamic nature of AIS exposes it to various cyber-attacks, requiring protection against unauthorized access. This work leveraged the multi-layer perceptron (MLP) to design an intrusion detection model for anomaly detection in the Maritime Transport System (MTS) network. Our model was evaluated on the CIC-IDS2017 cyber-security dataset. The simulation results show the proposed model's superiority over state-of-the-art models, in terms of high accuracy and timeliness in detecting attacks.**

***Index Terms*—AIS, Anomaly, Intrusion Detection, Maritime transport system.**

I. Introduction

The Automatic Identification System (AIS), an Internet of Things (IoT) technology, is a vital component of the Maritime Transportation System (MTS). It plays a crucial role in enhancing safety and security in maritime operations by providing essential communication and tracking capabilities. AIS was introduced to the maritime domain with the objective of improving situational awareness and ensuring the safety of sea traffic [1]. On-board ships embedded with AIS transceivers broadcast information obtained from the global navigation satellite system (GNSS) to authorities and vessels in their region periodically. However, AIS suffers from the fundamental drawback of being inherently unreliable due to its open wireless medium architecture which exposes it to different malicious interference (zero-day attacks) in the network [2], necessitating the need for intelligent security solutions capable of detecting new cyber threats. This requirement arises from the fact that emerging cyber-attacks targeted at the MTS have exposed the inadequacies of existing intrusion detection systems, which rely on predictable rules and fixed attack patterns, rendering them ineffective in identifying and addressing novel threats within the network [3].

Machine learning (ML) and deep learning (DL) IDMs have been proposed for MTS networks, but not specifically for AIS. For instance, in [4], a computational-efficient 1-dimensional convolutional neural network (1-CNN) was proposed, achieving an accuracy of 86%. Also, authors in [5], designed an ML-based model leveraging adaptive incremental passive-aggressive ML for IoT-enabled MTS networks, and their model achieved an accuracy of 99%. However, the above solution lacks target systems, and they fail to consider the temporal dynamics and environmental factors peculiar to the MTS networks. Therefore, this paper presents a robust DL-based anomaly detection in AIS tracks to protect the vessels from attacks. Specifically, this study makes the following contributions:

1) The design of a DL-based framework, to be embedded in the AIS of each vessel to enhance the security of MTS via anomaly detection.
2) Evaluated the impact of different activation functions on the convergence speed of the model.
3) The use of Principal Component Analysis (PCA) to eliminate irrelevant features and improve detection accuracy.

II. System Methodology

A. Conceptualized Detection Architecture and Model Structure

As detailed in Fig. 1, the MTS network seamlessly connects sea and land areas, allowing for smooth navigation and communication. AIS transponders installed on vessels are essential to this network, enabling real-time data exchange between vessels, coastal stations, and satellites. Our innovative self-defense model, a five-layer sequential neural network model can be integrated into the AIS system on each vessel, actively monitoring and analyzing network traffic data in real-time. This robust system is designed to identify and resist any anomalies that deviate from the established normal network patterns, ensuring the security of legitimate vessels from malicious ones, as well as the integrity of the MTS network. By employing advanced monitoring and analysis techniques, our model safeguards the network against potential threats and enhances the overall reliability of maritime operations.

B. Dataset Description

The CIC-IDS2017 dataset comprising 36 features and 1,061,342 instances with 9 attack types and normal traffic was utilized. It was divided into an 80:20 train-test split. Redundant features were removed using the PCA technique, thus, reducing the curse of dimensionality and increasing the performance of the model.

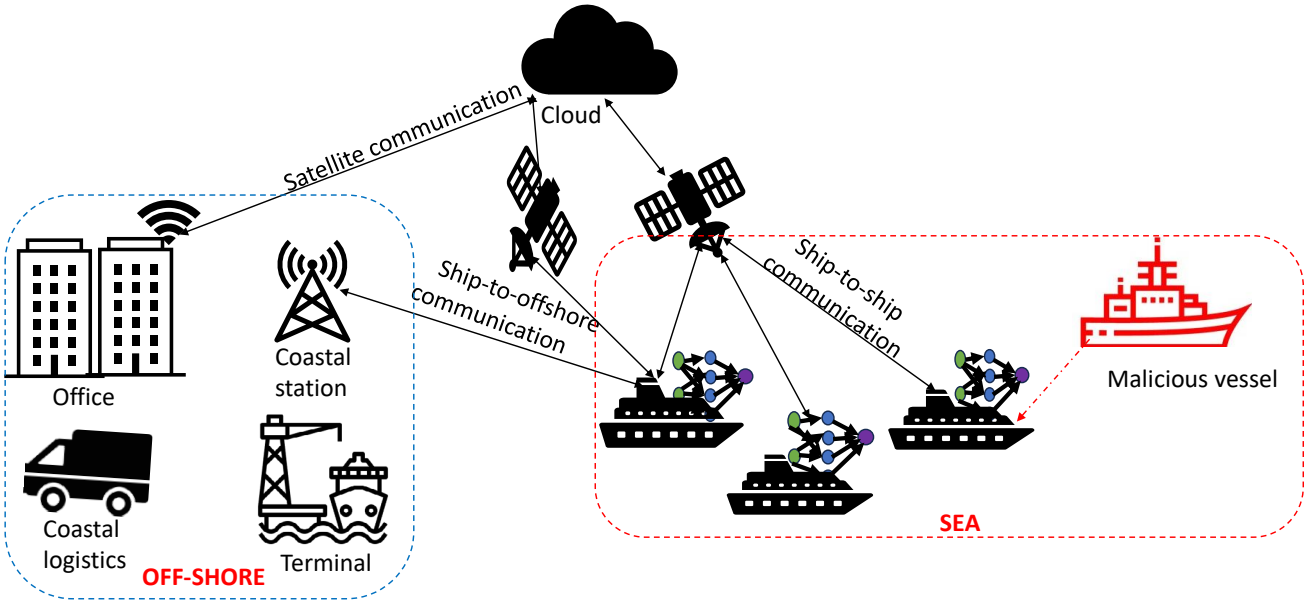


Fig. 1. Proposed Model Architecture

C. Experiment

The simulation was done on Windows 10 operating system, Python with Tensorflow 2.9, Intel Core i5-7400 processor, 8GB RAM, and Tesla K80 GPU. For the simulation parameter, a 5-layer architecture was used with Adam optimizer for 50 epochs. A Categorical cross-entropy loss function and a batch size of 30. The activation functions employed were ReLU and softmax.

III. Performance Evaluation/ Results Discussion

Firstly, we explored the effectiveness of different activation functions to determine the model's optimal performance while keeping other parameters constant. Table I indicates that Relu is the best candidate for our model. The proposed MLP model was evaluated in terms of its accuracy in detecting attacks.

Fig. 2 illustrates the accuracy of the model on both the train and validation set, demonstrating a remarkable accuracy of 99.2% in the prediction of complex attacks in the network, within a training time of 187sec, The proposed MLP model demonstrates efficiency in securing the MTS from malicious intrusions.

In comparison with recent related studies, as indicated in Table II, our proposed model exhibits the highest performance accuracy of 99.2%, with precision, recall, and F1-score of 99.1%, 99.2%, 99.2% respectively, outperforming the performances of existing models.

IV. Conclusion

This study proposed a DL framework for anomaly detection in the Automatic Identification System(AIS) track, Our pro-

TABLE I
Assessing the Impact of Activation Function on the MLP

Activation function	Batch Size	Epoch	Accuracy	Loss	Precision	Recall	Time
Relu	30	50	**99.2**	**0.026**	**99.1**	**99.2**	**187**
Leaky Relu	30	50	98.7	0.035	98.9	98.9	200
Tanh	30	50	98.8	0.032	98.8	98.9	205

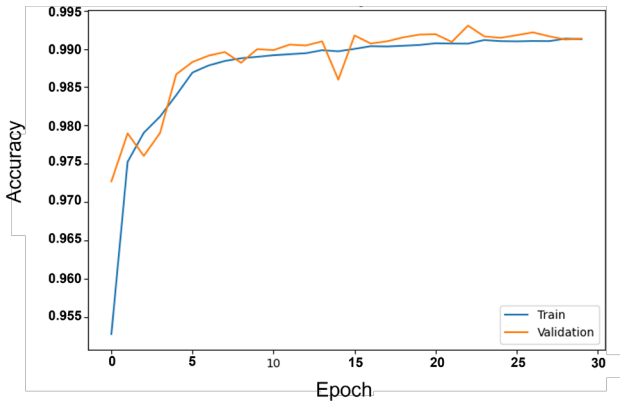


Fig. 2. Accuracy of proposed model

TABLE II
Comparison of Performance with other Studies from Literature

Ref	Model	Accuracy (%)	Precision (%)	Recall (%)	F1-score (%)
[5]	AL-PAML	98.7	99.0	99.0	99.0
[4]	LCNN	86.0	88.0	86.0	85.7
Our model	**MLP**	**99.2**	**99.1**	**99.2**	**99.2**

posed model shows optimal accuracy and efficiency in detecting attacks and enhancing MTS security when compared with current state-of-the-art. Future work aims at improving model timeliness, also we hope to develop a framework leveraging blockchain technology to guarantee robust security.

ACKNOWLEDGMENT

This research work was funded by MIST (Ministry of Science and ICT), Korea. Under the Innovation Human Resource Development for local intellectualization support program(IITP-2023-2020-0-01612) supervised by the IITP (Institute of Information and Communications Technology Planning & Evaluation.

REFERENCES

[1] B. J. Tetreault, "Use of the automatic identification system (ais) for maritime domain awareness (mda)," in *Proceedings of Oceans 2005 Mts/Ieee*. IEEE, 2005, pp. 1590–1594.

[2] A. Goudossis and S. K. Katsikas, "Towards a secure automatic identification system (ais)," *Journal of Marine Science and Technology*, vol. 24, pp. 410–423, 2019.

[3] V. U. Ihekoronye, S. O. Ajakwe, D. Kim, and J. M. Lee, "Hierarchical intrusion detection system for secured military drone network: A perspicacious approach," in *MILCOM 2022 - 2022 IEEE Military Communications Conference (MILCOM)*, 2022, pp. 336–341.

[4] X. Su, G. Zhang, M. Zhang, B. Ye, and H. Xing, "Intrusion Detection for Marine Meteorological Sensor Network," in *2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking*. IEEE, 2022, pp. 190–196.

[5] E. Gyamfi, J. A. Ansere, M. Kamal, M. Tariq, and A. Jurcut, "An adaptive network security system for iot-enabled maritime transportation," *IEEE Transactions on Intelligent Transportation Systems*, 2022.

A Parked-Vehicle-assisted Computation Offloading Scheme in Multi-access Edge Computing

Xuan-Qui Pham and Dong-Seong Kim
ICT Convergence Research Center, Kumoh National Institute of Technology, Korea
Email: {pxuanqui, dskim}@kumoh.ac.kr

***Abstract*—In this paper, we propose a parked-vehicle-assisted computation offloading scheme, which employs parked vehicles as opportunistic computing resources to expand the capacity of conventional multi-access edge computing paradigms for offloading tasks from mobile devices. The objective is to maximize the total offloading utility in terms of the benefit of latency reduction through offloading and the cost of using computing and networking resources. The preliminary results verify the superiority of our proposal over conventional schemes.**

***Index Terms*—computation offloading, edge computing, vehicular network.**

I. Introduction

Multi-access edge computing (MEC) has emerged as a paradigm that facilitates the offloading of resource-intensive tasks (e.g., face recognition, video analytics, and online gaming) from mobile devices (MDs) to MEC servers at the network edge (e.g., base stations and WiFi access points). Additionally, recent research introduced the utilization of parked vehicles (PVs) to assist MEC servers for computation offloading, which can relieve the resource congestion of MEC servers during the peak time [1]. This is promoted by the ubiquity of PVs with underutilized computing resources and the development of vehicle-to-everything communication technologies. However, the majority of research has primarily concentrated on full offloading schemes, where MDs either fully process or fully offload their tasks. Meanwhile, partial offloading schemes, which parallelize each task on both local and remote resources, have often been overlooked [1].

In this paper, we propose a PV-assisted offloading scheme to maximize the total utility of all MDs in terms of service latency and the cost of using computing and networking resources. Specifically, we derive a closed-form expression of the optimal offloading ratio and resource allocation, and solve the task assignment problem by using a binary version of the whale optimization algorithm (BWOA) [2]. The efficiency of our proposal is verified by preliminary evaluation results.

II. System Model and Problem Formulation

Fig. 1 presents the system model consisting of a MEC-enabled base station (BS), and a set $\mathcal{N}$ of MDs and a set $\mathcal{M}$ of PVs within the communication coverage of the BS. Each

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the Innovative Human Resource Development for Local Intellectualization support program (IITP-2023-2020-0-01612) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation). It was also supported by Priority Research Centers Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education, Science and Technology (2018R1A6A1A03024003).

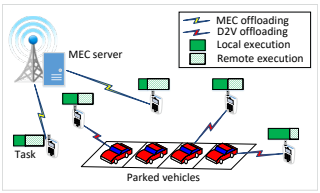


Fig. 1: System model.

MD has a computation task characterized as (B_i, C_i), where B_i and C_i denote input data size (in bits) and computation density (in CPU cycles per bit). In our model, the BS serves as a resource controller to monitor and select reliable PVs for computation offloading. Denoting λ_i as the offloading ratio, each MD i can locally execute $(1-\lambda_i)B_i$ bits of its task and offload the rest to the MEC server ($j=0$) or a PV $j \in \mathcal{M}$.

Denoting f_i^l as the local capacity of MD i, the partial local computing latency is expressed by

$$t_{ij}^{loc} = \frac{(1-\lambda_i)B_i C_i}{f_i^l}. \tag{1}$$

where f_i^l represents the computing capacity of MD i.

For computation offloading, we consider the same communication model as in [1]. Let R_i^{MEC} and R_{ij}^{D2V} denote the transmission rates from MD i to the MEC server and to nearby PVs $j \in \mathcal{M}$, respectively. The partial offloading latency of the MEC offloading mode is expressed by

$$t_{i0}^{off} = \frac{\lambda_i B_i}{R_i^{MEC}} + \frac{\lambda_i B_i C_i}{f_{i0}}, \tag{2}$$

where f_{i0} denotes the computing resources allocated to MD i by the MEC server. The partial offloading latency of the device-to-vehicle (D2V) offloading mode is expressed by

$$t_{ij}^{off} = \frac{\lambda_i B_i}{R_{ij}^{D2V}} + \frac{\lambda_i B_i C_i}{F_j}, \tag{3}$$

where F_j denotes the computing capacity of the PV j.

The overall latency is $t_{ij} = max\{t_{ij}^{loc}, t_{ij}^{off}\}, j \in \mathcal{M} \cup \{0\}$.

Let $\mathbf{a} = \{a_{ij} | i \in \mathcal{N}, j \in \mathcal{M} \cup \{0\}\}$ be the task assignment profile, where $a_{ij} = 1$ indicates MD i offloads its task to computing node j and $a_{ij} = 0$ otherwise. The utility is then defined based on the benefit of latency reduction and the cost paid for communication and computing resources as follows:

$$W_i = \sum_{j \in \mathcal{M} \cup \{0\}} a_{ij} W_{ij} = \sum_{j \in \mathcal{M} \cup \{0\}} a_{ij} [G_t(t_i^L - t_{ij}) - P^{trans} \lambda_i B_i - P_j^{proc} f_{ij}], \tag{4}$$

where G_t denotes the unit gain of latency reduction, t_i^L denotes the latency of fully local computation, and P^{trans} and P_j^{proc} respectively represent the unit price of data transmission and

computation resources when offloading to computing node j. The problem formulation can be expressed as follows:

$$\max_{\boldsymbol{\lambda},\boldsymbol{a},\boldsymbol{f}} \quad \sum_{i\in\mathcal{N}} W_i \tag{5}$$

$$\text{s.t.} \sum_{j\in\mathcal{M}\cup\{0\}} a_{ij} \le 1, \quad \forall i \in \mathcal{N}, \tag{5a}$$

$$\sum_{i\in\mathcal{N}} a_{ij} \le 1, \quad \forall j \in \mathcal{M}, \tag{5b}$$

$$\sum_{i\in\mathcal{N}} a_{i0} f_{i0} \le F_0, \tag{5c}$$

$$t_{ij}^{loc} = t_{ij}^{off}, \qquad \forall i \in \mathcal{N}, j \in \mathcal{M}\cup\{0\}, \tag{5d}$$

$$0 \le \lambda_i \le 1, a_{ij} \in \{0,1\}, f_{i0} \ge 0, \quad \forall i \in \mathcal{N}, j \in \mathcal{M}\cup\{0\}. \tag{5e}$$

where (5a) and (5b) represents the offloading relationships. (5c) represents the MEC resource allocation constraint. (5d) indicates the conditions for minimizing latency in the partial offloading. And (5e) imposes boundaries of the variables.

III. Proposed Approach

To solve the above-mentioned problem, we decompose it into two subproblems as follows:

A. Optimization of offloading ratio and resource allocation

Given (1), (2), and (5d), the computing resources of the MEC server allocated to MD i can be expressed by

$$f_{i0} = \frac{\lambda_i B_i C_i}{(1-\lambda_i)B_i C_i / f_i^l - \lambda_i B_i / R_{i0}}. \tag{6}$$

Substituting (6) into W_{ij} in (4) and then analyzing the function $W_{i0}(\lambda_i)$ by differentiation, we can obtain the optimal offloading ratio of MD i in the MEC offloading mode:

$$\lambda_i^* = \begin{cases} 0, & \text{if } \frac{G_t B_i C_i}{f_i^l} - P^{trans} B_i \le 0 \quad \text{or} \quad \overline{A} \le 0 \\ \overline{A}, & \text{otherwise,} \end{cases} \tag{7}$$

where $\overline{A} = \frac{B_i C_i R_{i0} - B_i C_i R_{i0} f_i^l \sqrt{\frac{P_0^{proc}}{G_t B_i C_i - P^{trans} B_i f_i^l}}}{B_i C_i R_{i0} + B_i f_i^l}$.

For the D2V offloading mode, the optimal ratio is

$$\lambda_i^* = \frac{B_i C_i R_{ij} F_j}{B_i C_i R_{ij} f_i^l + (B_i C_i R_{ij} + B_i f_i^l) F_j}. \tag{8}$$

B. Optimization of task assignment

Given the obtained offloading ratio and resource allocation, the problem (5) can be rewritten as follows:

$$\max_{\boldsymbol{a}} \quad \sum_{i\in\mathcal{N}} W_i \tag{9}$$

$$\text{s.t.} \quad (5a), (5b), (5c), a_{ij} \in \{0,1\}.$$

To solve this, we adopt the whale optimization algorithm (WOA), which is inspired by the prey hunting strategy of humpback whales. Each whale represents a candidate solution or search agent, whose position is iteratively updated according to three main operations: encircling prey, bubble-net attacking method (exploitation phase), and search for prey (exploration phase) by using adaptive parameters. The original WOA is only appropriate for unconstrained continuous optimization problems. Therefore, we adopt the penalty method as the constraint-handling technique. The fitness function is

$$Fitness(\boldsymbol{a}) = -\sum_{i\in\mathcal{N}} W_i(\boldsymbol{a}) + P(\boldsymbol{a}). \tag{10}$$

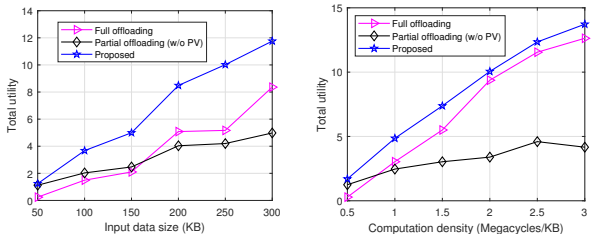
Algorithm 1 Pseudocode of the BWOA-based scheme

1: **Initialization**:
- Initialize the population of random search agents (i.e., task assignment solutions) and the parameters of the BWOA
- Based on the obtained offloading ratio and resource allocation according to Section III-A, compute the fitness value of each search agent and determine the best solution

2: **repeat**
3: **for** each search agent **do**
4: Update parameters in the BWOA
5: Update the position of the current search agent
6: **end for**
7: Compute the fitness of each search agent
8: Update $\boldsymbol{a}^*$ of the best search agent
9: **until** reaching the maximum number of iterations T_{max}
10: **return** the best search agent

where $P(\boldsymbol{a})$ is the penalty function. The minus sign in front of the objective function is to convert to a minimization problem. On the other hand, we use the sigmoid function in [2] as the transfer function to develop a BWOA. The pseudocode of the proposed BWOA-based scheme is shown in Algorithm 1.

IV. Preliminary Evaluation Results



(a) (b)

Fig. 2: Total utility versus: (a) Data size, (b) Density.

Using similar simulation settings as in [1], we consider a single cell coverage of 200 m × 200 m, in which a MEC-enabled BS ($F_0 = 10$ GHz) and 10 PVs offer the computation offloading service for 20 MDs. We compare the performance of our proposed scheme with two baselines: *full offloading* scheme and *partial offloading without PV assistance* scheme. Fig. 2 presents the system utility with respect to varying input data size and processing density. We can see that the partial offloading can improve the system utility compared to the full offloading, even in some cases using only the MEC server to offload. And our proposal takes a joint optimization of both offloading modes to achieve the best total utility.

V. Conclusion

This paper introduced an efficient partial computation offloading scheme optimizing both the MEC and D2V offloading modes. The initial findings showcased the efficiency of the proposed scheme. In-depth performance analysis with more baselines and under different scenarios is left for future work.

References

[1] X.-Q. Pham *et al.*, "Partial computation offloading in parked vehicle-assisted multi-access edge computing: A game-theoretic approach," *IEEE Trans. Veh. Technol.*, vol. 71, no. 9, pp. 10 220–10 225, 2022.

[2] V. Chahar and D. Kumar, "Binary whale optimization algorithm and its application to unit commitment problem," *Neural Computing and Applications*, vol. 32, 04 2020.

Optimizing Transmission Repetition in Industrial IoT Networks Using Particle Swarm Optimization

Won Jae Ryu
ICT Convergence Research Center
Kumoh national Institute of Technology
Gumi, Republic of Korea
wj0828@kumoh.ac.kr

Dong-Seong Kim*
Dept. of IT Convergence Engineering
Kumoh national Institute of Technology
Gumi, Republic of Korea
dskim@kumoh.ac.kr

***Abstract*—This paper discusses the use of Particle Swarm Optimization (PSO) for transmission repetition optimization using Irregular Repetition Slotted ALOHA (IRSA) in wireless networks. This study deals with a method to minimize collisions and improve transmission efficiency in Industrial IoT networks. In Industrial IoT, the power of devices can cause battery drain, so many transmissions can increase battery drain. In this paper, we describe how to reduce transmission collisions and reduce power consumption by using PSO. The simulation results show that the proposed method can reduce power consumption.**

***Index Terms*—Uplink, Irregular Repetition Slotted ALOHA, Optimization, Particle Swarm Optimization, Industrial IoT**

I. Introduction

The Internet of Things (IoT) is a technology that allows physical objects to be connected to the Internet and exchange data. These objects are typically composed of sensors, cameras, and actuators, and their main function is to communicate with each other, and collect, process, and analyze data. In industrial IoT systems, wireless communication technology plays a very important role. Wireless communication is easier to install and maintain than traditional wired networks, and it is suitable for mobile devices [1]. In addition, IoT devices typically operate on low power, so low-power technology plays a very important role in wireless communication. Therefore, the development of wireless communication and low-power technology in IoT is very important. In a wireless IoT network, if a grant-free uplink scheme is applied, when many IoT devices communicate with each other, transmission collisions can occur, which can cause problems in data transmission. If Irregular Repetition Slotted ALOHA (IRSA) is used instead of the commonly considered Slotted ALOHA (SA) scheme, higher transmission efficiency can be confirmed. IRSA is one of the effective Medium Access Control (MAC) protocols in wireless communication, and it minimizes transmission collisions using an irregular repetition structure [2]. This collision recovery is possible using Successive Interference Cancellation (SIC) technology. In addition, IRSA can minimize transmission collisions using an irregular repetition pattern, and due to these advantages, IRSA can enable more stable transmission than SA in grant-free uplinks. However, repeated transmission can affect battery consumption in IoT networks, where low-power systems are typically introduced. Therefore, in this paper, we use Particle Swarm Optimization (PSO) to optimize the repetition pattern for IRSA's irregular repeated transmission, and we aim to simultaneously achieve collision recovery and energy consumption minimization [3] [4]. The composition of this paper is as follows. In Chapter 2, we explain IRSA and the application of PSO. In Chapter 3, we discuss the simulation results, and in Chapter 4, we discuss the conclusion and future research.

II. System Model and Proposed Scheme

This chapter discusses how to optimize the IRSA retransmission-related coefficients using PSO. The optimization is performed in the direction of minimizing the number of retransmissions while eliminating collisions. First, what IRSA is going be explained. IRSA is a random access protocol based on slotted ALOHA. When transmitting data, IRSA intentionally causes collisions by transmitting data through an irregular repetition pattern. In the event of a collision, the Successive Interference Cancellation (SIC) recovers the collided data. The following is the formula for retransmission between timeslots and users.

$$\Lambda(x) \triangleq \sum_{l} \Lambda_l x^l, \qquad \Psi(x) \triangleq \sum_{l} \Psi_l x^l. \tag{1}$$

In Equation (1), is the probability distribution of the transmission pattern that the node will transmit. is the probability distribution of the number of transmissions received per timeslot, that is, the probability distribution of collisions.

$$q > \lambda(1 - e^{-qG\lambda}), \qquad \forall q \in (0, 1]. \tag{2}$$

In Equation (2), q represents the probability of decoding. That is, the first q in the above formula is the decoding probability at i+1, and the q used in the calculation of the exponent multiplier is the decoding probability at i. In this case, q_{i+1} must be greater than q_i. So, by repeating the above process, the decoding probability can be increased. If q at $i+1$ is less than or equal to q at i, it means that the SIC of IRSA is not working properly, and the probability distribution of retransmission in (1) is incorrect. In order to set this probability distribution within the correct range and at the same time reduce the number of retransmissions to

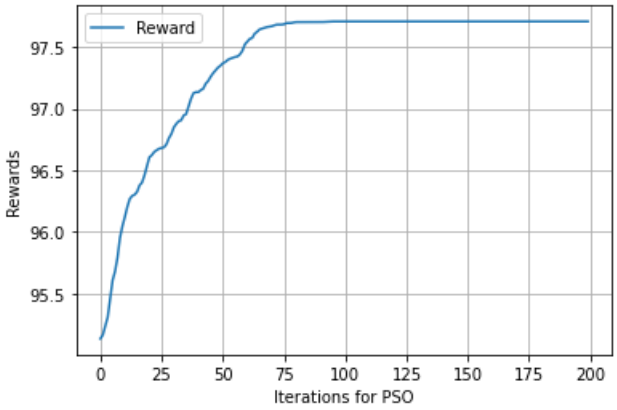


Fig. 1. Rewards

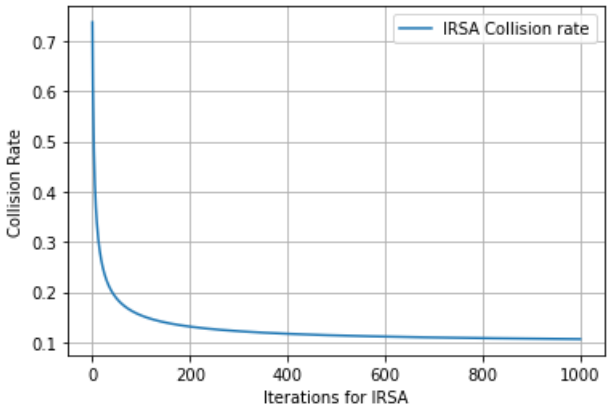


Fig. 2. Collision Rates

reduce the power consumption of the node, PSO was used. In other words, optimization was performed through PSO. The pseudo-code for the fitness function used for optimization is as follows.

Algorithm 1: Fitness Function

Input: the number of timeslots n
Output: score s

```
coefs = getDist(Λ(x))
for q in 0.1:0.01:1 do
    result = λ(q, coefs)
    if q ≥ result then
        s= -10
if s ≠ 10 then
    for j in 1:n do
        s =s + n - coefs[j]*(j+1)
```

In PSO, the particles are initialized randomly and then move through the search space according to the velocity and position update equations. The optimization process is repeated until the fitness function value converges.

III. Simulation

This chapter discusses the results of optimizing the IRSA to minimize the number of iterations and resolve collisions using PSO. The simulation parameters were set to 50 particles and 200 iterations. The number of timeslots in a frame was set to 10, and the number of transmission nodes was set to 7. As you can see in Figure 1, the fitness score increases as the number of iterations increases through PSO. This shows that the number of unnecessary iterations can be reduced, which can save energy. Figure 2 shows the collision rate as the number of SIC iterations increases. We can see that the collision rate decreases as the number of iterations increases. This means that the optimized distribution by PSO can reduce the collision rate, just like the fixed distribution that previous research has suggested.

IV. Conclusion

In Industrial IoT systems, battery consumption is a major problem. To solve this problem and the collision during transmission, this paper proposes a method to use Particle Swarm Optimization (PSO) to optimize the transmission repetition using the Irregular Repetition Slotted ALOHA (IRSA) protocol. IRSA is one of the effective MAC protocols in wireless communication, and it minimizes transmission collisions using an irregular repetition structure. This IRSA protocol was improved to avoid transmission collisions and reduce power consumption by using PSO. To this end, the proposed method was simulated to confirm that it can reduce power consumption. These research results are expected to be helpful for various research to reduce power consumption in IoT networks.

Acknowledgment

This work was supported in part by the Ministry of Science and ICT (MSIT), Korea, under the Grand Information Technology Research Center Support Program, supervised by the Institute for Information and Communications Technology Planning and Evaluation (IITP), under Grant IITP-2023-2020-0-01612, in part by the Priority Research Centers Program under Grant 2018R1A6A1A03024003, and in part by the Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education, Science and Technology, under Grant 2022R1I1A1A01069334.

References

[1] Farooq, M. Umar, et al. "A review on internet of things (IoT)." International journal of computer applications 113.1 (2015): 1-7.
[2] Paolini, Enrico. "Finite length analysis of irregular repetition slotted ALOHA (IRSA) access protocols." 2015 IEEE International Conference on Communication Workshop (ICCW). IEEE, 2015.
[3] Ryu, Won Jae, Gandeva Bayu Satrya, and Soo Young Shin. "Bio-Inspired Scheduling for Factory Automation in the TD-LTE System." IETE Technical Review 39.5 (2022): 1189-1207.
[4] Ryu, Won Jae, and Soo Young Shin. "Performance evaluation of a power allocation algorithm based on dynamic blocklength estimation for URLLC in the multicarrier downlink NOMA systems." Turkish Journal of Electrical Engineering & Computer Sciences 29.1 (2021): 310-320.

Orbital Angular Momentum Enabled Device-To-Device Communications

Jiarui Pang
Department of Artificial Intelligence Convergence
Pukyong National University
Busan, South Korea
jiaruipang@pukyong.ac.kr

Sudhanshu Arya
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, NJ, USA
sarya@stevens.edu

Yeon Ho Chung
Department of Artificial Intelligence Convergence
Pukyong National University
Busan, South Korea
yhchung@pknu.ac.kr

***Abstract*— Internet of Things (IoT) refers to the incorporation of billions of devices connected through the internet. To enable connectivity among IoT devices, Radio Frequency (RF)-based technology is most predominantly deployed in various environments. However, optical wireless communication (OWC) has shown promising and significant improvements in overall performance. Orbital angular momentum (OAM), which is the revolutionary property of light, describes the twist or helical shape of the electromagnetic field. When combined with OWC, OAM is believed to enhance data transmission significantly. This paper focuses on OAM-based OWC in IoT and the objective is to analyze the performance of the system in terms of data transmission rate and bit error rates (BER) for various spatial separations between IoT devices as the transmitter and the receiver. Moreover, we make an important observation that the achievable data rate varies non-linearly with the BER requirements of the OAM-enabled IoT system at a given link range.**

Keywords— OAM, OWC, Internet of Things

I. INTRODUCTION

The definition of Internet of Things (IoT) encompasses a huge expansive domain that establishes the foundation for integrating the physical devices and digital layers. In recent years, billions of IoT devices have been deployed and this tremendous trend is anticipated to continue in the foreseeable future. A crucial component facilitating the traffic of information in IoT networks is wireless communication, which plays an indispensable role for the connectivity between primary entities. Among prevalent wireless solutions in contemporary IoT domains, Radio Frequency (RF) based technologies, such as Wi-fi, Bluetooth, and Zigbee, are the most common solutions, and they are utilizing low power operating in communication links affected by loss and noises [1]. Due to their performance and capacity for distance coverage, stability, power consumption and penetration ability, RF-based wireless communications have evolved into a well-established method of transmitting and facilitating the data traffic between IoT devices.

Although RF-based technologies provide an easily deployable and cost-effective solution, it also faces some severe challenges, such as interference, spectrum congestion and latency. RF signals are strictly regulated therefore, they pose a challenge due to the interferences of other devices operating in the same frequency band [2]. Additionally, as the number of users grows, congestion is likely to occur in the absence of sophisticated advanced infrastructure and management.

To further address the mentioned concerns, additional measurements have been implemented using optical wireless communication (OWC). OWC has been proposed as an alternative to complement and coexist with RF-based technologies, therefore enabling promising results for wireless communication in a broad range of IoT domains [2]. By operating in another domain of spectrum, OWC is possible to effectively mitigate the above issues in terms of spectrum licensing [2], larger capacity offered by the modulated optical beam, and higher scalability when it comes to a large number of connected devices. OWC utilizes visible or invisible lights for data transmission between transceivers and the optical components in OWC establish a better figure of design due to its size, weight, and power-cost design in comparison to RF components [2]. In recent years, orbital angular momentum (OAM) has been undertaken to leverage the capacity of OWC. The property of OAM represents a structure light, allowing it to be obtained as an optical beam with a tailored spatial amplitude [3]. In addition, the implementation of multiple channel modes and multiplexing with OAM results in sophistical improvements in minimizing crosstalk and leveraging the data transmission quality [3].

OAM integrated OWC is envisioned to bolster wireless communication in IoT networks by offering higher capacity, lower latency, and increased efficiency compared to traditional RF-based solutions. The future OAM-based OWC in IoT should provide optimal data rate in transmission channel with low bit error rates (BER). This paper delves into the deployment of OAM-based OWC for device-to-device communications in IoT by elucidating the system model of OAM-based OWC. It presents an analysis of data transmission rate in the system channel and conclusions are drawn.

II. SYSTEM MODEL

The system model of OAM based OWC is represented in Figure 1. The parameters used are shown in Table 1. Moreover, the channel impulse response for OAM-OWC in IoT is represented as follow:

$$h_{p=}^{l}\beta_y\left(\frac{\lambda}{4\pi d_y}\right)\exp\{-j[kd_y + l\frac{\pi}{2}\frac{k}{2F}(r^2 - r_y^2)]\}\exp\left[-(\frac{r^2-r_y^2}{w^2(z)})\right](\frac{r_y}{r})^{|l|}\frac{[(\nabla-1)^p p^{p+|l|}]}{(\nabla_y-1)^p p_y^{p+|l|}} \quad (1)$$

where l is the different channel response input of OAM, ρ is the transmission power, r is related to the radius of the OAM propagation circle shown in Figure 1. And the transmit power is shown as

$$\rho = \frac{2r^2}{w^2(z)} \quad (2)$$

Figure 1 illustrates how OAM beam's propagation occurs along the channel between the transmitter (T_x) and the receiver (R_x)

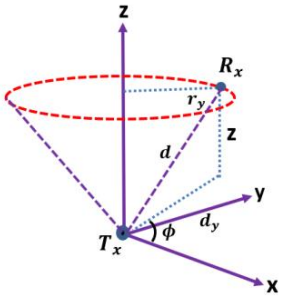


Figure 1. OAM-based OWC transmission

Table I: SYSTEM PARAMETERS

Transmitter-Receiver Distance (d)	0.1 m to 1 m
Radius of OAM beam propagation	r_y
Transmit Power (P_{tx})	10^{-4}
OAM Mode Index (l)	2
Bandwidth	10^4

III. Simulation

As a proof-of-concept, we present the achievable data rate relative to the link range for different reliability constraints, i.e., the maximum bit-error-rate (BER) a system can tolerate. The simulation parameters are illustrated in Table I. As expected, as the link range increases, the achievable data rate, for the given reliability constraint, reduces. We make another important observation that the achievable data rate an OAM-enabled IoT system can support for a given link range varies non-linearly with the BER requirements of the system.

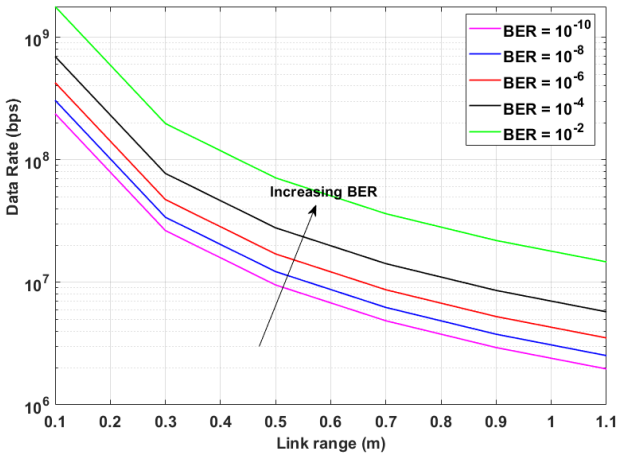


Figure 2. Achievable data rate relative to the link range for different BER requirements.

IV. Conclusion

In this work, the OAM-based OWC for data transmission in IoT networks has been considered in terms of data rate and BER. Moreover, the impact of varying distances between the devices has been examined. Further research will be undertaken to obtain a more comprehensive understanding of the system model's characteristics and performance by incorporating other essential parameters that have not been addressed in this work.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2023R1A2C2006860).

References

[1] Ala Al-Fuqaha, Mohsen Guizani, Mehdi Mohammadi, Mohammed Aledhari and Moussa Ayyash, "Internet of Things: A Survey on Enabling Technologies, Protocols, and Applications", *IEEE Communication Surveys & Tutorials, Vol. 17, NO.4, Fourth Quarter 2015*

[2] Abdulkadir Celik, Imene Romdhane, Georges Kaddoum and Ahmed M.Eltawil, "A Top-Down Survey on Optical Wireless Communications for the Internet of Things", *IEEE Communications Surveys & Tutorials, Vol.25, No.1, First Quarter 2023*

[3] Ming Li, "Orbital-Angular-Momentum Multiplexing Optical Wireless Communications With Adaptive Modes Adjustment in Internet-of-Things Networks.", *IEEE Internet of Things Journal, Vol. 6, No.4, August 2019*

Multiuser Orbital Angular Momentum using Identical Modes

Aye Yadanar Win
Department of Artificial Intelligence Convergence
Pukyong National University
Busan, South Korea
yadanar_aw@pukyong.ac.kr

Sudhanshu Arya
School of Systems and Enterprises
Stevens Institute of Technology
Hoboken, NJ, USA
sarya@stevens.edu

Yeon Ho Chung
Department of Artificial Intelligence Convergence
Pukyong National University
Busan, South Korea
yhchung@pknu.ac.kr

Tu Lam Thanh
Department of Electronics and Telecommunications Engineering,
Ton Duc Thang University,
Vietnam
tulamthanh@tdtu.edu.vn

***Abstract*—In this paper, we propose a novel approach for multiuser communication using the same orbital angular momentum (OAM) mode. Unlike previous studies that utilize different OAM modes for individual users, our approach allows multiple users to share the same OAM mode, thus increasing capacity. Instead of using different OAM modes for each user, we propose utilizing the Gouy phase to enable multiple users to transmit and receive using the same OAM mode. We analyze the feasibility of this approach through simulations and our results demonstrate the potential of multiuser OAM communication using the same OAM mode for increasing the capacity of optical communication systems.**

Keywords—Orbital Angular Momentum (OAM), multiuser, same OAM mode, Gouy phase.

I. INTRODUCTION

Optical communication systems have become increasingly important in recent years, driven by the growing demand for high-speed data transmission[1]. One promising approach to increase the capacity of optical communication systems is to use OAM, which involves transmitting data using optical beams with different OAM modes[2]. However, multiuser OAM communication presents some challenges. One of these challenges is the limited number of available OAM modes for practical OAM transmission systems. Another significant challenge is channel estimation for different OAM modes. The number of channels corresponding to all OAM modes is substantial, leading to a heavy overhead for channel estimation. Recovering the transmit signals corresponding to different OAM modes becomes challenging when the number of channels is relatively large[2].

In this paper, we propose a novel approach to multiuser OAM communication using the same OAM mode and the Gouy phase. Instead of using different OAM modes for each user, we utilize the Gouy phase to enable multiple users to transmit and receive using the same OAM mode. The feasibility of our proposed technique is demonstrated through simulations, and the results show that it is a promising solution for future multi-user OAM communication systems.

II. SYSTEM MODEL

As shown in Fig. 1, we consider the Gaussian wave for OAM generation. It is important to note that, for all the users, we use the Gaussian wave expression that corresponds to the same OAM mode index value, i.e., OAM Mode, $l = 1$. However, for each user in multiuser OAM transmission systems, we have different Gouy phase values. Using these different phase values for each user, we can identify them at the receiver. It means all users have the same OAM mode values but different Gouy phase values. We use Gouy phase value to detect the users (or their information/data/message/ signal). By varying these Gouy phase values, we can generate pulses that are distinguishable for different users. The same OAM mode is used for all users, which simplifies the implemen- tation of the system.

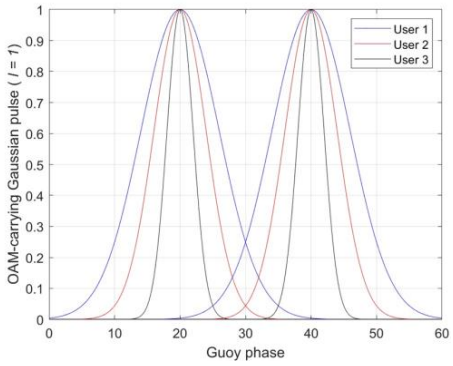


Figure 1. Multiuser OAM Communication by exploiting the Gouy Phase Characteristics of the OAM-carrying beam ($l = 1$).

III. PERFORMANCE ANALYSIS

When multiple users transmit in the same OAM mode, the signals can interfere with each other [3], leading to a reduction in system performance. We focus on calculating the cross-correlation between users to assess the orthogonality of the OAM modes used by different users. Cross-correlation between users is given by [4]

$$\emptyset_{xy}(k) = \sum_{j=\max(0,k)}^{\min(M-1+k,N-1)} x_j y_{j-k} \quad ,$$

$$k = -(M+1), \dots, 0, \dots, (N-1) \qquad (1)$$

where k is the lag, and $\emptyset_{xy}(k)$ represents the cross-correlation of x_j and y_j at lag k. The summation is taken over all indices of j where both x_j and y_{j-k} are defined. The range of k is from

-(M-1) to (N-1), which ensures that all possible lags are considered. In the expression (1), the multiplication of x_j and y_{j-k} at each time index j and subsequent summation over all j, gives the measure of similarity between the two signals at lag k.

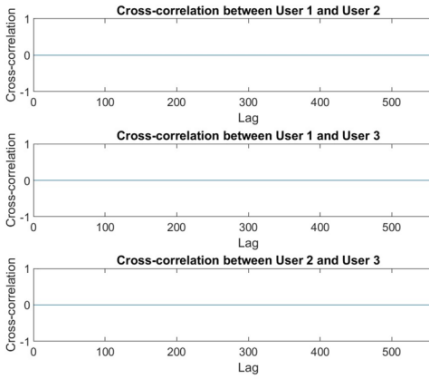


Figure 2. Cross-correlation between different users.

Simulation result of the cross-correlation relative to the lag between users are presented to evaluate the system performance. Figure 2 shows that the cross-correlation between the users is minimized when the OAM mode used by each user is orthogonal to the OAM modes used by all other users.

IV. CONCLUSIONS

In this paper, we have presented a multiuser OAM communication system that uses the same OAM mode for all users, which can increase the capacity of optical communication systems. We have shown that by varying the Guoy phase values, we can generate pulses that are well separated and distinguishable for different users. Our result shows the potential of multiuser OAM communication using the same OAM mode for increasing the capacity of optical communication systems. Future work could focus on improving the performance of the system by reducing noise and increasing the number of users in the system.

ACKNOWLEDGMENT

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2023R1A2C2006860).

REFERENCES

[1] A. Sawant, I. Lee, B. C. Jung and E. Choi, "Ultimate Capacity Analysis of Orbital Angular Momentum Channels," in IEEE Wireless Communications, vol. 28, no. 1, pp. 90-96, February 2021, doi: 10.1109/MWC.001.2000258.

[2] R. Chen, H. Zhou, M. Moretti, X. Wang and J. Li, "Orbital Angular Momentum Waves: Generation, Detection, and Emerging Applications," in IEEE Communications Surveys & Tutorials, vol. 22, no. 2, pp. 840-868, Secondquarter 2020, doi: 10.1109/COMST.2019.2952453.

[3] W. Son, H. Lee and B. C. Jung, "Achievable Rate of Multi-User Mode-Division Multiplexing Using Orbital Angular Momentum," 2019 IEEE 90th Vehicular Technology Conference (VTC2019-Fall), Honolulu, HI, USA, 2019, pp. 1-2, doi: 10.1109/VTCFall.2019.8891483.

[4] Bourke, Paul, " Cross Correlation", Auto Correlation—2D Pattern Identification (1996).

Joint Demodulation and Mode Detection with Longitudinal View of Vortex Beam for OAM-FSO Communications

Young Jae Moon*, Sudhanshu Arya‡, Yeon Ho Chung*, *Senior Member, IEEE*, Trio Adiono†
*Department of Information and Communications Engineering, Pukyong National University, Korea
†School of Electrical Engineering and Informatics, Bandung Institute of Technology, Indonesia
‡School of Systems and Enterprises, Virginia Polytechnic Institute and State University, United States of America

***Abstract*—A joint demodulation and mode comparison between distinct signals using the longitudinal view of the vortex beam is presented. This proposed technique is confirmed by identifying different beams in a longitudinal two-dimensional image made from the product of two different OAM Lagurre-Gaussian beams. As future work, it will incorporate convolutional neural networks to explore the utilization of OAM demultiplexing.**

***Index Terms*—Free-space optical communication, Longitudinal View, Neural network, Optical Vortices**

I. Introduction

One of the essential properties of electromagnetic signals is orbital angular momentum (OAM), which allows a high user capacity and ensures high modulation capabilities. This vortex beams (VB) feature enhances OAM-based communication in free-space optical (FSO) settings by enabling Tbit/s speed rate in transmission environment [1].

When Laguerre-Gaussian (LG) beams are used, the intensity beam pattern displays a 2D image that utilizes an assortment of OAM modes [2]. Demultiplexing and differentiating OAM mode signals, however, are problematic issues that restrict the detection range, speed, and distinguishing accuracy when the product of two different Lagurre-Gausssian (pLG) beams is exposed to atmospheric interference. To overcome these problems, researchers have proposed various approaches in order to achieve high procedural efficiency of OAM demodulation and mode comparison in an FSO environment [3], [4].

In this paper, a joint demodulation and mode separation method based on the longitudinal view of the vortex beam with perceptron is proposed for the FSO-OAM communication. Section 2 presents the background of the LG beam. The results are presented in Section 3. In Section 4, conclusions are drawn.

II. Lauguerre-Gaussian Beams

A kind of radiated beam with high chromaticity which uses an approximate Gaussian amplitude envelope is known as a Laguerre-Gaussian (LG) beam. The ring-shaped asymmetrical modes of the beam use Laguerre polynomials in dimensions in order to solve the paraxial Helmholtz equation, resulting in a Gaussian intensity profile in the horizontal plane.

When the Laguerre-Gaussian beam moves across the free-space optical environment with a cylindrical axis, the vortex beam will create an annulated intensity distribution along the x-directional polarization on the $+z$ vertical plane. However, the longitudinal perspective is considered when the vortex is structured with finite-difference time-domain (FDTD) method, which covers a wide frequency range with a single simulation run and naturally treats nonlinear material properties [5]. The x component of scattering magnetic field $H_{s,x}^{(n+1/2)}$ is located at XY plane $z = (k_s - 1/2)\Delta z$, when we assume the total field/scattering field (TF/SF) interface is perpendicular to the z axis. We denote (i, j, k), and t as the grid position indexes along the x, y, z respective and the time instant, respectively. Therefore, the x electric field component of the LG beam is:

$$E_{LG}^{(n)}(i+\frac{1}{2}, j, k_s) = E_{LG,x}[(i+\frac{1}{2})\Delta x, j\Delta y, k_s\Delta z, n\Delta t] \quad (1)$$

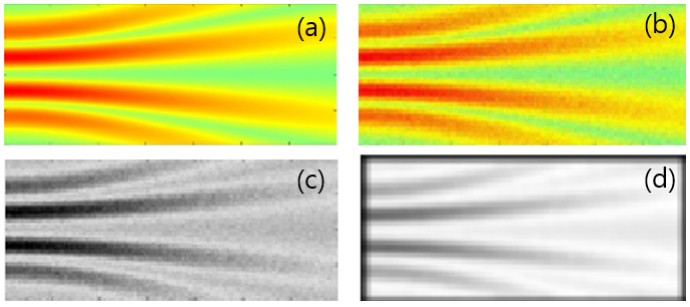


Fig. 1. (a) A received two-dimension image with Gaussian noise and motion filter added. (b) Resized image of received ones. (c) Resized Image filtered by a graying filter. (d) Convolutional filter image with grayed filter and Gaussian filter. These processes are made with the longitudinal image of the OAM beams.

III. OAM Mode Detection and Analysis

Now, we describe the detection process and its analysis for the simulation using longitudinal intensity profiles. As we follow the identical simulation setting from [4], the overall procedures will subsequently occur with the processes from Fig. 1 (a) to Fig. 1 (d).

An ideal offset for each sample of data is considered with twenty, sixty, eighty, and a hundred OAM data samples, using the stable variable model to differentiate the number of OAM data samples d and the offset number o. Note that the offset means the mode difference between two beams which offers

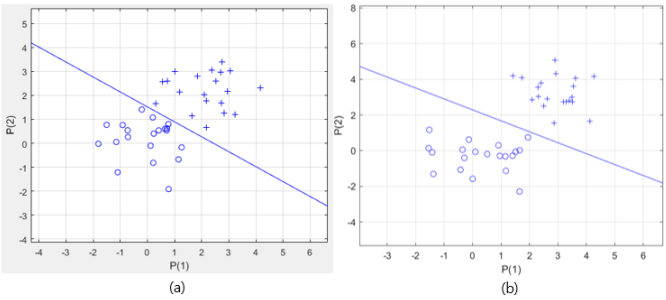


Fig. 2. Simulation results with 20 data samples. (a) o = 2. (b) o = 3.

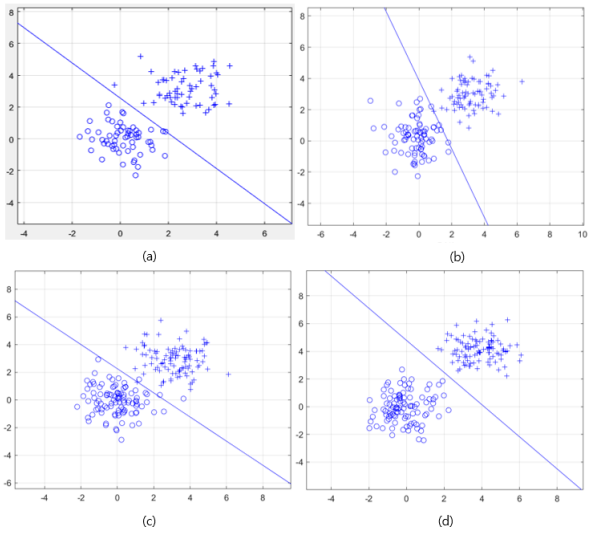


Fig. 3. (a) Simulation results for d = 60 and o = 3. (b) d = 80 and o = 3. (c) d =100, o = 3. (d) d = 100, o = 4.

a consistent delimitation. In Fig. 2, the two distinct optical vertices are separated when o = 2 with the twenty samples. However, the data of the two different beams are incomparable due to a lack of training data.

As we increase the number of samples to sixty, the offset of two cannot guarantee a stable decision boundary set. The offset number needs to be increased to three, as shown in Fig. 3 (a). Fig. 3 (b) also demonstrates that o = 3 is an inadequate number to establish a stable separation when d is eighty. When the difference of the mode is three with a hundred samples, the data of each beam intrude the decision boundary, as shown in Fig. 3 (c). When the minimal offset is set to four, the decision boundary is clearly set, as shown in Fig. 3 (d), which is found to be an adequate offset number.

With the results presented, we can assume that the ideal offset number changes as we increase the data set number; so, it is important to investigate the aspect of the minimal adequate number required for the size of the data samples.

IV. Conclusion

Our proposed joint demultiplexing has been verified, which uses a machine learning method perceptron to separate the individual OAM signals based on the intensity patterns of each of the respective vortex beams. With the technique separating the different modes with the longitudinal 2D pLG image, we can find the adequate minimal offset with the number of data samples we use. In theory, it is well anticipated that our work will be compatible with OAM demultiplexing solutions.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korean government(MSIT) (2023R1A2C2006860).

References

[1] L. Allen, S. M. Barnett, and M. J. Padgett, Optical Angular Momentum. Boca Raton, FL, USA: CRC Press, 2016, pp. 109–186.
[2] Bassene, Pascal et al., "Nonlinear Conversion of Orbital Angular Momentum States of Light." Frontiers in Optics, 2021, paper JTh5A.85.
[3] J. Li, M. Zhang, D. Wang, S. Wu, and Y. Zhan, "Joint atmospheric turbulence detection and adaptive demodulation technique using the CNN for the OAM-FSO communication," Opt. Express, vol. 26, no. 8, pp. 10494–10508, 2018.
[4] Young Jae Moon, Sudhanshu Arya and Yeon Ho Chung, "Demodulation and Mode Separation Analysis using Perceptron for the OAM-FSO Communication", KICS Summer Conference, Jeju, Korea (to be published)
[5] Zhili Lin, Xudong Chen, Weibin Qiu and Jixiong Pu, "Propagation Characteristics of High-Power Vortex Lauguerre-Gaussian Laser Beams in Plasma", Applied Science, April 2018

Orbital Angular Momentum: Recent Technologies in Optical Wireless Communications

Siti Hasunah Mohammad
Lightwave Communication Research Group
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
shasunah2@graduate.utm.my

Yeon Ho Chung
Department of Information and Communications Engineering
Pukyong National University
Busan, Korea
yhchung@pknu.ac.kr

Sevia Mahdaliza Idrus
Department of Communication Engineering
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
sevia@fke.utm.my

Nadiatulhuda Zulkifli
Department of Communication Engineering
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
nadia@fke.utm.my

***Abstract*—The simultaneous transmission of multiple orthogonal beams using structured light has attracted a lot of attention for its potential to boost communications capacity. One of a kind is the orbital angular momentum beam, which has been used in several applications worldwide owing to its special qualities of a helical phase structure and doughnut intensity profile, which led to high transmission capacity and spectral efficiency. This paper reviews the fundamentals of orbital angular momentum and its current progress. Then, the research challenges with regard to the use of optical angular momentum in optical wireless communications are discussed.**

Keywords—orbital angular momentum, optical communications, optical wireless communications

I. INTRODUCTION

The term "orbital angular momentum" (OAM) refers to a class of structured light beams, unique light fields with customized spatial structure and variable amplitude, phase, and polarization distribution [1]. OAM has been extensively explored in optical communications in terms of OAM modulation, multiplexing, and multicasting due to its unique characteristics of a helical phase structure and doughnut intensity profile [2]. These OAM properties may provide a helpful degree of freedom in the development of high-capacity and high-spectral-efficiency optical communication systems.

In this paper, we first introduce the concepts of OAM in optical communications. Then, the recent research progress of OAM technologies in optical wireless communications along with the research challenges is discussed.

II. CONCEPTS OF OAM IN OPTICAL COMMUNICATIONS

Optical communications or more broadly electromagnetic communications involve different physical dimensions or degrees of freedom of electromagnetic waves (EM) to transmit data. While OAM optical communication is also a type of communication using EM waves to transfer data information, unlike conventional EM waves, the energy of OAM waves is focused within a circle [3].

In OAM, the light is coiled like a screw surrounding the propagation axis as shown in Fig. 1. The twisting feature causes the light waves to cancel each other out at the axis. The helical modes are identified by a positive or negative integer value of l. If $l = 0$, the mode is not helical, and the wavefronts are multiple disconnected surfaces. If $l = \pm 1$, the wavefront is formed as a single helical surface, with a step length equal to the wavelength λ. If $|\, l \,| \geqslant 2$, the wavefront is composed of $|\, l \,|$ distinct but intertwined helices, with the step length of each helix surface equal to $|\, l \,|\lambda$.

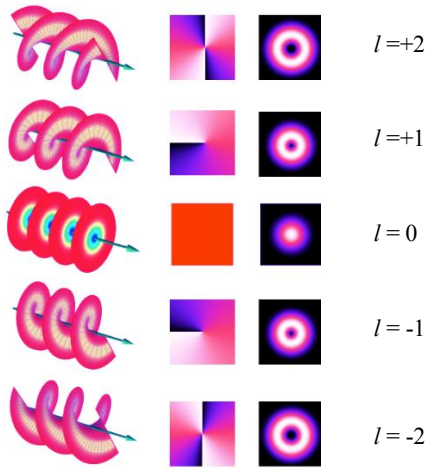


Fig. 1. OAM beam helical structures, phase fronts, and corresponding intensity distributions.

In theory, OAMs can exist in an endless number of states in free space, increasing the data transmission rate in the medium. This feature also allows for multiplexing and a limitless channel capacity. Since beams with distinct OAM states are orthogonal, data from multiple users can be transmitted at the same time. At the receiver, each orthogonal channel can be completely filtered and decoded. OAM supports several users concurrently without multiuser disruption. Multilevel modulation can be used for OAM states, resulting in a huge constellation size and good spectrum efficiency.

III. Research on OAM in Optical Wireless Communications and its Challenges

Modulating, multiplexing, and multicasting data information are common methods of optical communications that use distinct physical dimensions of photons, and so does OAM optical communication. OAM modulation refers to data information encoded by various OAM beams, in which data information is directly encoded by time-varying OAM states. A few works on OAM modulation for free-space optical system has been demonstrated since 2004, involving spatial light modulators [4], high-speed OAM modulation [5] and spatial array modulation [6]. For OAM multiplexing, when carrying distinct OAM values, multi-channel data information modulated by other physical dimensions like as amplitude and phase can be distinguished from one another, i.e., different OAM beams are employed as independent carriers to transport different data information. OAM multiplexing technique is applied in a few studies regarding ultra-high spectral efficiency [7] and real-world long-distance application scenario [8]. On the other hand, OAM multicasting duplicates data information into multiple copies (carrying different OAM values) for multiple end users (i.e., one-to-many communications) [9].

In optical wireless field, the terahertz OAM (THz-OAM) beam has been used to accommodate the ever-increasing wireless capacity by combining THz radiation and OAM technologies [10]. THz technology can provide high speed wireless communication with large available bandwidth, while OAM technology can increase the capacity of a communication system by multiplexing several orthogonal modes. The convergent THz-OAM technology has significant potential for next-generation communication systems, but some limitations must be solved before this technology can be realized, such as deployment cost due to expensive THz devices and large OAM generation and detection components size.

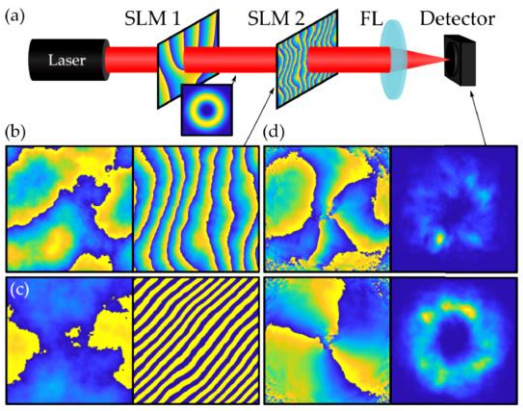


Fig. 2 (a) A typical configuration for generating a mode before aberrating it through turbulence. (FL: Fourier Lens; SLM: Spatial Light Modulator). Holograms for (b) strong (top, SLM) and (c) weak (bottom, DMD) turbulence with and without gratings. (d) Phase and intensity of l =3 beams through the setup for strong (top panel) and weak (bottom panel) turbulence strengths [11].

Furthermore, atmospheric turbulence is a key obstacle for optical wireless communications based on OAM, as it affects inter-model crosstalk between distinct OAM beams, leading to OAM detection difficulties [11], as shown in Fig. 2. Spatial modes of OAM beams are susceptible to distortion and are adversely disturbed by air turbulence. Scattering of one mode into another adds noise to a quantum connection, lowering security. Extenuating this challenge remains an open topic that is being thoroughly researched, both in the laboratory and in the field.

IV. Conclusion

The ideas of OAM have been briefly described in this paper, along with its unique features, modulation, multiplexing, and multicasting strategies. As long as the research challenges are overcome, the application of OAM in optical wireless communications holds great potential for next generation communications.

Acknowledgment

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (2023R1A2C2006860).

References

[1] S. Ramachandran and P. Kristensen, "Optical vortices in fiber," Nanophotonics, vol. 2, no. 5–6, pp. 455–474, 2013, doi: 10.1515/nanoph-2013-0047.

[2] J. Wang, S. Chen, and J. Liu, "Orbital angular momentum communications based on standard multi-mode fiber (invited paper)," APL Photonics, vol. 6, no. 6, pp. 060804-1–15, 2021, doi: 10.1063/5.0049022.

[3] L. Wang, X. Ge, R. Zi, and C. X. Wang, "Capacity Analysis of Orbital Angular Momentum Wireless Channels," IEEE Access, vol. 5, pp. 23069–23077, 2017, doi: 10.1109/ACCESS.2017.2763679.

[4] G. Gibson et al., "Free-space information transfer using light beams carrying orbital angular momentum," Opt. Express, vol. 12, no. 22, pp. 5448–5456, Nov. 2004, doi: 10.1364/OPEX.12.005448.

[5] J. Wang, J. Liu, S. Li, Y. Zhao, J. Du, and L. Zhu, "Orbital angular momentum and beyond in free-space optical communications," Nanophotonics, vol. 11, no. 4, pp. 645–680, 2022, doi: 10.1515/nanoph-2021-0527.

[6] L. Zhu, A. Wang, M. Deng, B. Lu, and X. Guo, "Experimental demonstration of multiple dimensional coding decoding for image transfer with controllable vortex arrays," Sci. Rep., vol. 11, no. 1, pp. 1–9, 2021, doi: 10.1038/s41598-021-91553-0.

[7] A. J. Willner et al., "Experimental demonstration of 20 Gbit/s data encoding and 2 ns channel hopping using orbital angular momentum modes," Opt. Lett., vol. 40, no. 24, pp. 5810–5813, Dec. 2015, doi: 10.1364/OL.40.005810.

[8] L. B. Du et al., "Long-reach wavelength-routed TWDM PON: Technology and deployment," J. Light. Technol., vol. 37, no. 3, pp. 688–697, 2019, doi: 10.1109/JLT.2018.2850343.

[9] A. E. Willner and C. Liu, "Perspective on using multiple orbital-angular-momentum beams for enhanced capacity in free-space optical communication links," Nanophotonics, vol. 10, no. 1, pp. 225–233, 2020, doi: 10.1515/nanoph-2020-0435.

[10] H. Yang, S. Zheng, W. He, X. Yu, and X. Zhang, "Terahertz orbital angular momentum: Generation, detection and communication," China Commun., vol. 18, no. 5, pp. 131–152, 2021, doi: 10.23919/JCC.2021.05.009.

[11] M. A. Cox, N. Mphuthi, I. Nape, N. Mashaba, L. Cheng, and A. Forbes, "Structured Light in Turbulence," IEEE J. Sel. Top. Quantum Electron., vol. 27, no. 2, pp. 1–21, 2020, doi: 10.1109/jstqe.2020.3023790.

SHAP-Based Intrusion Detection Framework for Zero-Trust IoT Maritime Security

Ebuka Chinaechetam Nkoro *, Judith Nkechinyere Njoku†, Cosmas Ifeanyi Nwakanma†,
Jae-Min Lee *, and Dong-Seong Kim *
*IT Convergence Engineering, † ICT Convergence Research Center, Kumoh National Institute of Technology, Korea.

***Abstract*—By integrating the Explainable AI (XAI) into Marine cyber defense systems, the lack of trustworthiness, and low interpretability of complex black box Intrusion Detection Systems (IDS) models can be satisfied. XAI is a key focus, especially in achieving a Zero trust cybersecurity strategy in Marine communication networks. In this article we develop an IDS framework that detects different kinds of network attacks, using a publicly available 2023 Edge-IIoTset dataset, achieving an average accuracy of 95% and leveraging an explainability algorithm (SHAP), for better explainability of the proposed model's prediction output thereby satisfying the systematic justification of interpretable and trustworthy IDS models.**

***Index Terms*—Cyber Security, Explainable AI, Maritime Security, encryption, Zero Trust.**

I. INTRODUCTION

Maritime network traffic management has become an essential task for the Maritime industry due to connected Internet of things (IoT) sensors that collect and transmit various kinds of benign and anomalous network traffic [1]. To mitigate Marine cyberattacks, adopting a zero-trust model would enable security experts to classify, interpret, and gain a better understanding of artificial intelligence (AI) informed intrusion detection systems as illustrated in Fig. 1.

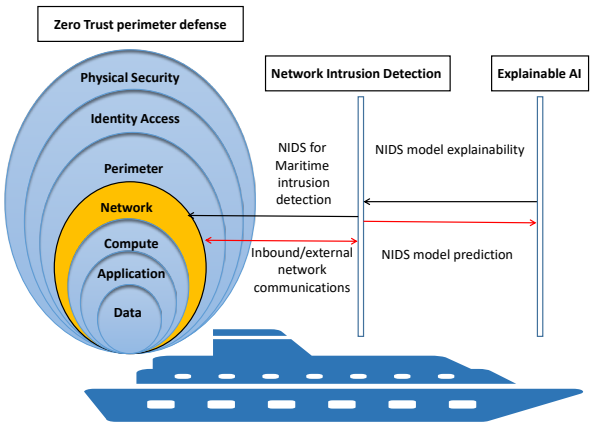


Fig. 1. Sequence diagram of the Zero trust perimeter defense strategy for Marine networks, with insight-driven feedback using an Explainable AI Network Intrusion Detection (NIDS) model.

A *Zero Trust* cybersecurity strategy adopts the *"trust no one, verify everything"* principle, to authenticate network users, encrypt data communications, and mitigate threats just-in-time [2]. Within this strategy, various perimeters of security measures are layered to prevent or slow the advance of an attacker who may aim at tampering with the confidentiality, integrity, and availability of a system. The Marine industry, suffers various kinds of security breaches, in the form of Denial of Service attacks (DDoS), fake distress calls, and malware attacks [3] To mitigate against Shore-based cyber-attacks, intrusion detection systems (IDS) have been employed by various researchers as a layer of the *zero-trust* model, to differentiate malicious alerts, traffic types, and vulnerability scanning [4] from benign communications.

In Ship communications, an IDS can effectively perceive, access, and relay generated traffic data of vessel communications and forbid flagged anomalous activities using AI algorithms [5]. Different threat types that exist in Marine networks may be categorized as intentional or internal and can be targeted towards Marine logistics and navigation, using sophisticated malware types [6]. Thus, there has been a need to adopt progressive solutions which provide increased defense against growing cyber attacks in Marine communications.

As Marine IDS rely on AI algorithms, an Explainable AI (XAI) approach provides a better understanding of the "why, what, and how" of classification algorithms, and their results since many of them are too complex for human understanding (opaque) [7]. Additionally, most predictions are partial in nature, consider only the accuracy, and do not judge models based on increased transparency. Marine IDS in the same regard also need to be explainable, trustworthy, and resilient against threats, to attain better organizational goals.

Motivation - Previous research on cyber resilience for Marine networks using AI algorithms, has not considered explainable methods which can provide transparency for predicting Marine threats and mitigation. By ignoring model explainability, a lack of better validation of classification results may be missing and thus can be linked to the growing rate of false alarms and attacks in Marine networks. Also, debugging the robustness and knowledge discovery of AI-proposed models seems to be missing by not leveraging different explainability methods.

Contribution - This paper's contributions are two-fold:

- This paper adopts a fine-tuned Light Gradient Boost decision tree classifier, for efficient and adaptable IoT Marine network traffic classification.
- An exploration of the SHapley Additive exPlanations (SHAP) explainability method is employed for the proposed IDS classification algorithm. The visual plots pro-

vided yields more insights towards the prediction output of the AI model, for Maritime cyber resilience.

II. RELATED WORK

Previous research has been made in the field of IDS for secure Marine network operations. A Federated learning approach for Marine IDS was investigated by authors in [5]. Their federated learning technique was designed to save computing and storage overhead, with an accuracy of 87%, 500 rounds of training, and using the old NSL-KDD dataset. Another work in [8] proposed an intrusion detection for hydrographic station networks with a hybrid CNN and BiLSTM method, using the NSL KDD dataset, while obtaining an F1-score of 87.35%. Although their approach was effective in identifying deep features, the low accuracy of their results can not be ignored, putting into consideration efficiency requirements and the need for low false positive rates required for a zero trust model in Marine networks. Authors in [9], proposed a CVAE-GAN-based approach to process the imbalanced NSK-KDD dataset for IDS in Marine networks. Within their approach, a data generation module was initiated to improve minority class samples, using the OPTICS denoising algorithm. The authors' classification accuracy of their proposed data augmentation method yielded a micro-average ROC accuracy curve of 95% with *5 classes* of network traffic.

These approaches did not consider model explainability regarding proposed models, as there has been a great demand for the explainability and trustworthiness of these models. Within the context of Zero Trust IDS for Marine security at the network layer, AI-based decisions should be continuously questioned and evaluated to improve their ability, integrity, and benevolence- meaning the extent to which the AI decisions are not self-harmful [10].

III. METHODOLOGY

Marine communications protocols (using TCP/IPV6), have been associated with many cyber threats as they interact with the outside world [11]. Fig. 2 explains the overall system model of the proposed intrusion detection framework. This paper suggests that attackers can breach the confidentiality, availability, and integrity of organizational data, and various communication nodes in marine networks. Since previous research using AI methods has been able to successfully predict normal and benign attacks, current cyber-risk prevention methods now demand more for XAI-based IDS. This section highlights a stage-by-stage and efficient ML-based IDS with a visual explainability method- SHAP for AI-based Marine IDS.

Implementation Environment: The IDS framework was carried out using the sci-kit learn library, which provides tree-based APIs. The experiment was run on Jupyter Notebook with an Intel(R) Core(TM) i3-7100 CPU @ 3.90GHz, and RAM 8.00 GB.

A. Dataset Description and Pre-processing

The publicly available Edge-IIoTset dataset [12] satisfies modern TCP/IPV4/6 IoT and IIoT applications. This dataset contains network traffic of DoS/DDoS, Information gathering, Man in middle attacks, Injection attacks, and Malware attacks, suitable for the proposed IDS model. A good understanding of network traffic features informs our selection of 12 important features. Data preprocessing, and cleaning is critically performed to remove outliners, missing values and NAN values. The relevant training features are employed for the multi-class classification task. Full code and dataset prepossessing code is available in the provided GitHub link: https://github.com/nkoro/SHAP_Marine_IDS

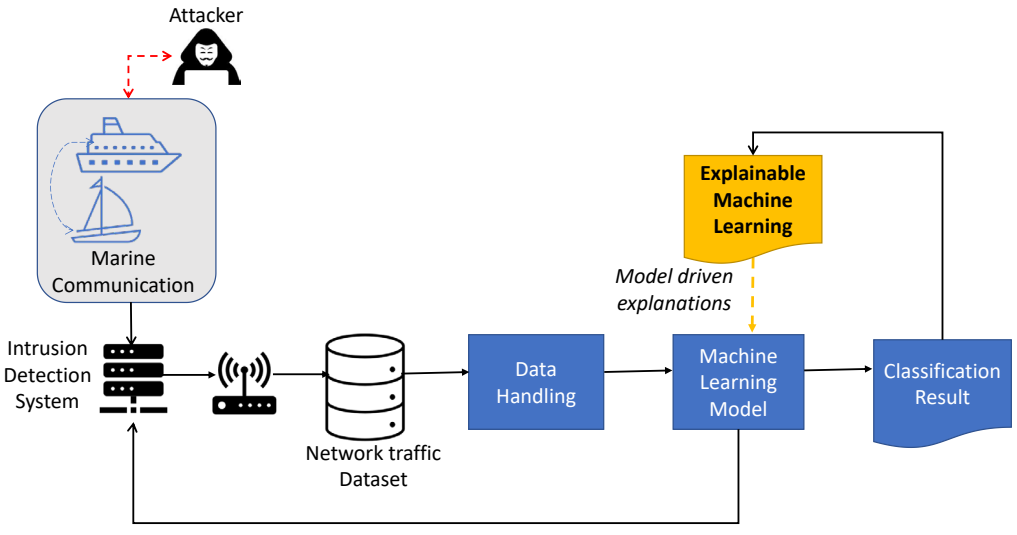


Fig. 2. System model, showing an explainable machine learning approach for network intrusion detection in Marine networks.

B. Light Gradient Boost Classifier

Light gradient boosting (LGBoost) is a widely used decision tree algorithm incorporated for classification tasks due to its precise, fast, and efficient features. The optimal parameters selected are represented in Table I. In general, the LBGoost algorithm maximizes parallel learning and voting for speeding up the learning process and grows leaves by dividing the leaf nodes with an optimization technique of an arbitrary differentiable loss function. A summarized loss function equation of the XGBoost model is shown below:

$$F_m(x) = F_{m-1}(x) + \gamma_m h_m(x) \tag{1}$$

$$\gamma_m = argmin_\gamma \sum_{i=1}^{n} L(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)). \tag{2}$$

TABLE I
OPTIMAL PARAMETER VALUES OF THE PROPOSED MODEL

Parameters	**Value**
Total Features	12
n-estimators	100
random state	42
bagging frequency	5
bagging fraction	0.75
cross-validation	10 k-fold

IV. PERFORMANCE EVALUATION

A. Evaluation Metrics

In terms of F1-Score, average accruacy in Table II precision, recall, and computation time, the confusion matrix in Fig. 3 and full github code results, summarizes the performance of the Lgb classifier. The F1-score of the multiclass classification results (*Backdoor, DDoSHTTP, DDoSICMP, DDoSTCP, DDoSUDP, Normal, PortScanning, Ransomware, SQL injection, Uploading Vulnerabilityscanner, XSS*) yielded were 99%, 87%, 86%, 100%, 94%, 100%, 95%, 92%, 93%, 99%, 91% with a computation time of 71 seconds, and 2.4 ms per sample.

Following the requirements of robust, efficient, and timely detection of intrusions for Marine systems, table II presents a comparison of previous approaches using a recent traffic dataset that reflects modern-day attack scenarios as compared to the NSL KDD dataset which still suffers from adaptability issues in modern, real-world scenarios. Furthermore, the decentralized model training method using Federated Learning (FL), as explored in [5], is computationally expensive and does not yield increased accuracy against anomaly attack prediction as compared to deep learning methods in [8], [9], using the same NSL-KDD dataset. The proposed approach in this paper adopts contemporary measures against cyber attacks in terms of increased accuracy, timely detection, and increased number of anomaly samples/classes using a more recent dataset.

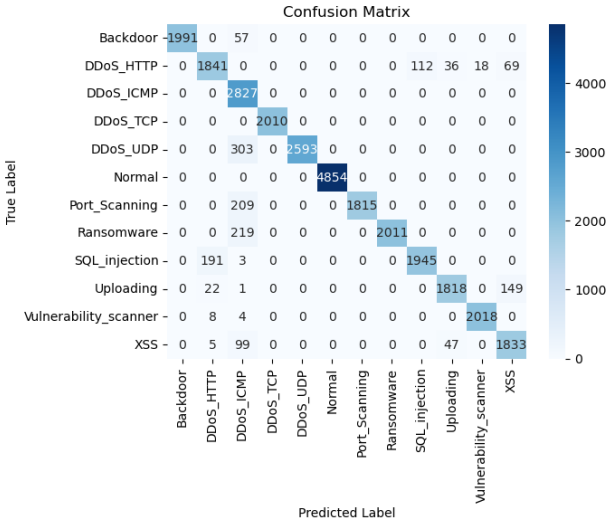


Fig. 3. Showing the confusion matrix of the multi-class traffic type classification.

TABLE II
AVERAGE ACCURACY IN TERMS OF ACCURACY, ATTACK CLASSES, AND USING THE 2022 EDGE-IIOTSET DATASET

Method	**No. of classes**	**Avg. Accuracy**	Dataset
Federated Learning [5]	5	87	NSL-KDD (2009)
CNN/BiLSTM [8]	5	87.35	NSL-KDD
CVAE-GAN [9]	5	95	NSL-KDD
Proposed method	12	95	Edge-IIoTset (2022)

B. Model Explainability with SHAP

The SHAP explainability method is a cooperative game theory method introduced by authors in [13], which basically provides the marginal value of contributions made by a feature or subset of features within a model's prediction. Since most models with large data are black-box in nature, the SHAP method addresses the issue of clarity and reason for model predictions. SHAP values can be calculated by the equation provided below:

$$\emptyset_i(f,x) = \sum_{z' \subseteq x'} \frac{|z'!|(M - |z'| - 1)!}{M!}[f_x(z' - f_x(z')i)] \tag{3}$$

where M is the total number of features, $\emptyset_i$ denotes the shapely value for feature i, in model (f), with input data points (x). z', $f_x(z')i)$ represents the model output with and without target feature. This paper adopts the SHAP *importance plot* to visualize the test data.

As shown in Fig. 4, the various network traffic features are ranked based on their average impact on the model, thus providing various insights like model debugging, feature selection, feature engineering, clearer explanations, and the individual feature importance for model prediction. For instance, if the *tcp.srcport* is randomized to have no prediction power in the model, this action will force a low accuracy and overall model performance. Another adopted plot in Fig. 5 shows a sample data prediction and its most important features that contributed to the prediction output. In this visualized plot,

the feature values highlighted in red within the range $>$ -0.03 are weighted important features that contributed to the single prediction of the multiclass network traffic type output. The tree black box model is better understood in this plot, thus we can rely on and transparently attribute training features for better understanding and improved model classification.

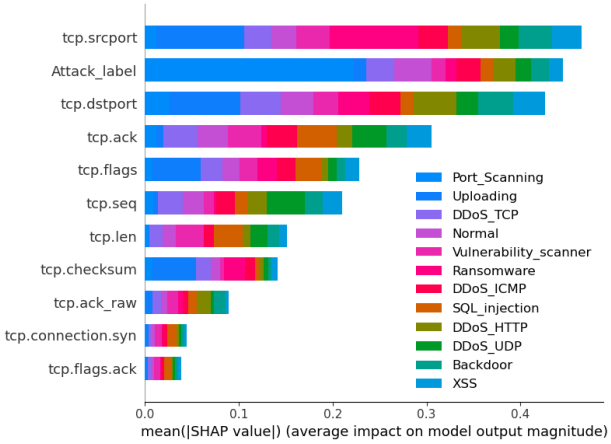


Fig. 4. SHAP feature importance on model output, with test data

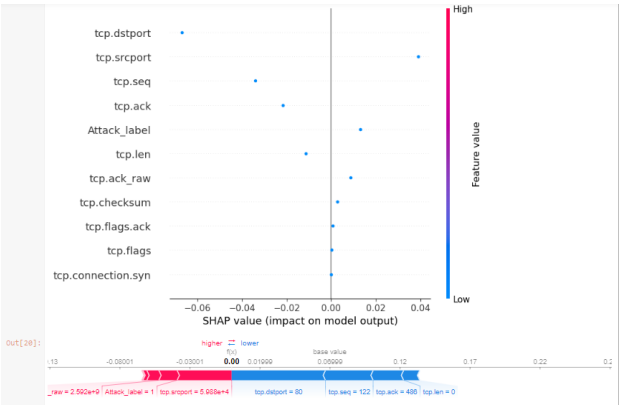


Fig. 5. SHAP contribution for one sample data prediction

V. CONCLUSION

The development of explainability approaches for black box AI models in the context of Zero-trust Marine security has been driven by the relevance of accuracy, transparency, and interpretability in IDS. This paper developed and evaluated a SHAP-based method that evaluates the explainability of an LGBoost algorithm for intrusion detection in marine networks. Future directions that would be worth exploring include reducing the time complexity of SHAP processes to enable faster and more efficient model interpretability. Investigating the explainability of deep learning and federated learning models should also be considered.

ACKNOWLEDGMENT

This research was supported by Priority Research Centers Program through the NRF funded by the MEST (2018R1A6A1A03024003), and by MSIT under the Innovative Human Resource Development for Local Intellectualization support program (IITP-2023-2020-0-01612) supervised by the IITP.

REFERENCES

[1] R. W. Liu, J. Nie, S. Garg, Z. Xiong, Y. Zhang, and M. S. Hossain, "Data-driven trajectory quality improvement for promoting intelligent vessel traffic services in 6g-enabled maritime iot systems," *IEEE Internet of Things Journal*, vol. 8, no. 7, pp. 5374–5385, 2021.

[2] M. Shore, S. Zeadally, and A. Keshariya, "Zero trust: The what, how, why, and when," *Computer*, vol. 54, no. 11, pp. 26–35, 2021. [Online]. Available: https://doi.org/10.1109/MC.2021.3090018

[3] G. Potamos, A. Peratikou, and S. Stavrou, "Towards a maritime cyber range training environment," in *2021 IEEE International Conference on Cyber Security and Resilience (CSR)*, 2021, pp. 180–185.

[4] E. C. Nkoro, C. I. Nwakanma, J.-M. Lee, and D.-S. Kim, "Industrial network attack vulnerability detection and analysis using shodan eye scanning technology," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, 2022, pp. 886–889.

[5] W. Liu, X. Xu, L. Wu, L. Qi, A. Jolfaei, W. Ding, and M. R. Khosravi, "Intrusion detection for maritime transportation systems with batch federated aggregation," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2503–2514, 2023.

[6] I. Ashraf, Y. Park, S. Hur, S. W. Kim, R. Alroobaea, Y. B. Zikria, and S. Nosheen, "A survey on cyber security threats in iot-enabled maritime industry," *IEEE Transactions on Intelligent Transportation Systems*, vol. 24, no. 2, pp. 2677–2690, 2023.

[7] C. I. Nwakanma, L. A. C. Ahakonye, J. N. Njoku, J. C. Odirichukwu, S. A. Okolie, C. Uzondu, C. C. Ndubuisi Nweke, and D.-S. Kim, "Explainable artificial intelligence (xai) for intrusion detection and mitigation in intelligent connected vehicles: A review," *Applied Sciences*, vol. 13, no. 3, 2023. [Online]. Available: https://www.mdpi.com/2076-3417/13/3/1252

[8] T. Hou, H. Xing, X. Liang, X. Su, and Z. Wang, "A marine hydrographic station networks intrusion detection method based on lcvae and cnn-bilstm," *Journal of Marine Science and Engineering*, vol. 11, no. 1, 2023. [Online]. Available: https://www.mdpi.com/2077-1312/11/1/221

[9] X. Su, T. Tian, L. Cai, B. Ye, and H. Xing, "A cvae-gan-based approach to process imbalanced datasets for intrusion detection in marine meteorological sensor networks," in *2022 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*, 2022, pp. 197–203.

[10] P. Laplante and J. Voas, "Zero-trust artificial intelligence?" *Computer*, vol. 55, no. 2, pp. 10–12, 2022. [Online]. Available: https://ieeexplore.ieee.org/document/9714079

[11] K. Tran, S. Keene, E. Fretheim, and M. Tsikerdekis, "Marine network protocols and security risks," *Journal of Cybersecurity and Privacy*, vol. 1, no. 2, 2021. [Online]. Available: https://www.mdpi.com/2624-800X/1/2/13

[12] M. A. Ferrag, O. Friha, D. Hamouda, L. Maglaras, and H. Janicke, "Edge-IIoTset: A New Comprehensive Realistic Cyber Security Dataset of IoT and IIoT Applications for Centralized and Federated Learning," *IEEE Access*, vol. 10, pp. 40 281–40 306, 2022.

[13] L. S. Shapley and M. Shubik, "The Assignment Game I: The core," *International Journal of game theory*, vol. 1, no. 1, pp. 111–130, 1971.

Historical Data Aware Sensor Data Prediction for Internet of Things Longevity

Made Adi Paramartha Putra*†, Gabriel Avelino Sampedro †‡, Jae-Min Lee†, and Dong-Seong Kim†
*Informatic Enginering, STMIK Primakara, Denpasar, Indonesia
†Department of IT Convergence Engineering, Kumoh National Institute of Technology, Gumi, South Korea.
‡Faculty of Information and Communication Studies, UP Open University, Laguna, Philippines.
{mdparamartha95*†, garsampedro †‡, ljmpaul†, dskim†}@kumoh.ac.kr

***Abstract*—This paper presents a historical data-aware approach aimed at improving the longevity of IoT sensor data by incorporating deep learning techniques for sensor data prediction. In contrast to previous studies that focused primarily on enhancing prediction accuracy using deep learning models, this paper highlights the significance of both prediction accuracy and energy efficiency. An LSTM model is developed, utilizing various variations of historical data look-back as input for generating sensor predictions. The simulation results demonstrate that utilizing a total of 8 historical data points yields superior model performance with a coefficient of determination of 95.66 compared to other variations. Additionally, an experimental test is conducted to assess the energy efficiency of utilizing different look-back data.**

***Index Terms*—Data prediction, deep learning, historical data, internet of things**

I. Introduction

The Internet of Things (IoT) facilitates the connection between microcontrollers and sensors, enabling them to monitor environmental conditions and communicate over networks [1]. The adoption of IoT technology has led to over 50 billion connected devices, with a projected increase to 200 billion by 2030 [2]. The energy consumption of IoT devices is directly influenced by the number of operations they perform, including data sensing, logging, transmission, and transients.

Among the IoT device operations, data transmission consumes the most energy, making it vital to minimize the number of transmissions for device longevity. Retransmission often occurs during high network loads or collisions, leading to reduced device availability [3]. Artificial Intelligence (AI), including Machine Learning (ML) and Deep Learning (DL), has been utilized to enhance performance across various domains, including data prediction for IoT sensors. DL-based data prediction has shown promising results in estimating missing sensor data caused by transmission collisions [4].

Shu et al. proposed an energy-efficient dual prediction scheme using the LMS algorithm and LSTM network [5]. Guzel et al. utilized an adaptive-network-based fuzzy inference system and LSTM network to predict sensor data [6]. Cheng et al. introduced a bidirectional LSTM-based data prediction approach leveraging values from other sensor nodes [7]. While these studies focused on improving prediction accuracy, the overall energy savings of sensor nodes were not fully considered. Therefore, this study incorporates simulation and implementation results to evaluate the prediction performance of sensor data, addressing the gap in considering both accuracy and energy efficiency. The key contributions of this paper are summarized as follows:

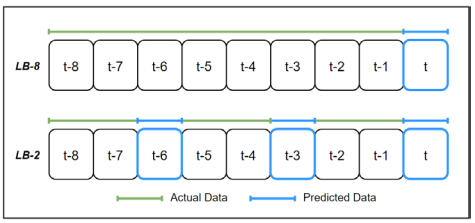


Fig. 1. Sensor data prediction employing various look-back mechanisms.

- Assess the performance of the deep learning (DL) model by employing different methods of incorporating historical data, aiming to enhance prediction outcomes across multiple performance metrics.
- Compute and compare the overall energy efficiency attained by adopting sensor data prediction techniques.

II. Proposed System

An extended version of the gated recurrent unit (GRU) network, known as the long short-term memory (LSTM) network, is well-suited for time-series data forecasting or regression tasks. The LSTM network consists of three distinct gates: the input gate, forget gate, and output gate. These gates employ sigmoid and tanh activation functions to extract relevant features from the input data. The operations performed by the gates within the LSTM network can be expressed as follows:

$$i_t = \sigma_g\left(W_i\, x_t + U_i\, h_{t-1} + b_i\right) \quad , \tag{1}$$

$$f_t = \sigma_g\left(W_f\, x_t + U_f\, h_{t-1} + b_f\right) \quad , \tag{2}$$

$$o_t = \sigma_g\left(W_o\, x_t + U_o\, h_{t-1} + b_o\right) \quad , \tag{3}$$

The input gate equation at time t is denoted by equation (1), while equations (2) and (3) represent the forget gate and output gate equations, respectively.

The input data used in the LSTM model is of utmost importance in achieving accurate prediction results. Utilizing preprocessed data, as opposed to raw data, can significantly enhance prediction accuracy. Fig. 1 illustrates the data prediction technique utilizing historical data. For instance, $LB-2$ and

TABLE I
COMPARATIVE ANALYSIS OF PERFORMANCE ACROSS DIFFERENT VALUES OF HISTORICAL DATA LOOK-BACK.

	LSTM								1D-CNN
	LB-1	LB-2	LB-4	LB-8	LB-16	LB-32	LB-64	LB-128	
RMSE	0.3573	0.3096	**0.2740**	*0.2790*	0.3217	0.4286	0.4726	0.5777	0.3620
MAE	0.3034	0.2618	*0.2260*	**0.2252**	0.2563	0.3631	0.3922	0.4549	0.3068
MAPE	1.2923	1.1166	**0.9620**	*0.9563*	1.0841	1.5595	1.6722	1.9158	1.3058
R2	0.9261	0.9445	0.9566	***0.9549***	0.9401	0.8397	0.8706	0.8064	0.9241

$LB-8$ require two and eight actual data points, respectively, to generate the prediction. The length of historical data, or the quantity of past data, is another critical factor in improving the data estimation process. However, selecting an inappropriate number of past data points can result in underfitting or overfitting of the model. Hence, this paper evaluates different amounts of historical data as inputs for the DL model. The variations in the quantity of past data are conducted using a geometric sequence approach, following the given formula:

$$LB = ar^{(n-1)} \quad , \tag{4}$$

The formula given in equation (4) calculates the number of historical data (LB) considered in the evaluation process. In this equation, a represents the initial term, which is set to 1, and r represents the common ratio, which is set to 2. Using this formula, the number of historical data is adjusted and determined as $LB = 1, 2, 4, 8, 16, 32, 64, 128$.

III. RESULTS AND DISCUSSION

The comparison of overall performance among different values of historical data look-back and the 1D-CNN model is presented in Table I. A total of eight variations of look-back values were assessed, and in terms of RMSE, the LSTM network with $LB-4$ achieved the lowest error value of 0.2740, closely followed by $LB-8$ with 0.2790. In relation to the MAE performance metric, the minimum error was observed with $LB-8$ and $LB-4$, with values of 0.2252 and 0.2260, respectively. It is important to note that a simple increase in the number of look-back data does not guarantee improved prediction results. Insufficient values can lead to underfitting, while excessively large values can result in overfitting. Therefore, careful consideration of the selection of look-back data in the LSTM network is advisable to maximize the model's performance. Moreover Fig. 2 illustrates the energy efficiency using different look-back data. The findings indicate that the utilization of data prediction resulted in an overall efficiency improvement of around 4685 rounds or 20%. Reducing the number of historical data to LB-1 or LB-2 has the potential to enhance energy efficiency.

IV. CONCLUSION

This study presents the utilization of historical data-aware for sensor prediction to prolong the lifespan of IoT devices. Historical data variation of the LSTM network is evaluated to obtain the best input configuration. The results demonstrate that data prediction can enhance the network's lifetime by up to 20% with LB-4 and 11% with LB-8 while minimizing prediction errors. The selection of historical data significantly impacts the accuracy of predictions, underscoring the importance of careful consideration in sensor data prediction.

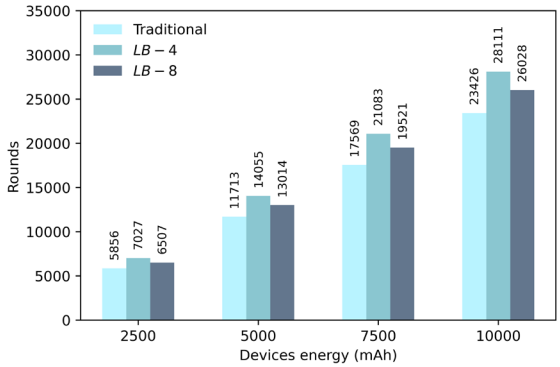


Fig. 2. Total rounds comparison between traditional transmission and network with data prediction.

ACKNOWLEDGMENT

This work was supported by Priority Research Centers Program through the NRF funded by the MEST(2018R1A6A1A03024003) and by the MSIT, Korea, under the Innovative Human Resource Development for Local Intellectualization support program (IITP-2023-2020-0-01612) supervised by the IITP and by NRF-2022R1I1A3071844.

REFERENCES

[1] H. Tran-Dang, N. Krommenacker, P. Charpentier, and D.-S. Kim, "Toward the internet of things for physical internet: Perspectives and challenges," *IEEE Internet of Things Journal*, vol. 7, no. 6, pp. 4711–4736, 2020.

[2] Y. B. Zikria, R. Ali, M. K. Afzal, and S. W. Kim, "Next-generation internet of things (iot): Opportunities, challenges, and solutions," *Sensors*, vol. 21, no. 4, 2021.

[3] L. Xu *et al.*, "A survey of clustering techniques in wsns and consideration of the challenges of applying such to 5g iot scenarios," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1229–1249, 2017.

[4] M. A. P. Putra, A. P. Hermawan, D.-S. Kim, and J.-M. Lee, "Energy efficient-based sensor data prediction using deep concatenate mlp," in *IEEE Int. Conf. on Emerging Tech. and Factory Auto. (ETFA)*, 2021.

[5] T. Shu, J. Chen, V. K. Bhargava, and C. W. de Silva, "An energy-efficient dual prediction scheme using lms filter and lstm in wireless sensor networks for environment monitoring," *IEEE Internet of Things Journal*, vol. 6, no. 4, pp. 6736–6747, 2019.

[6] M. Guzel, I. Kok, D. Akay, and S. Ozdemir, "Anfis and deep learning based missing sensor data prediction in iot," *Concurrency and Computation: Practice and Experience*, vol. 32, no. 2, p. e5400, 2020.

[7] H. Cheng, Z. Xie, L. Wu, Z. Yu, and R. Li, "Data prediction model in wireless sensor networks based on bidirectional LSTM," *EURASIP J. on Wireless Commun. and Networking*, vol. 2019, no. 1, p. 203, 2019.

Generative Adversarial Network (GAN)-enhanced Digital Twin for Edge Computing

Hoa Tran-Dang
Dept. of IT convergence engineering
Kumoh National Institute of Technology
Gumi-si, South of Korea
{hoa.tran-dang}@kumoh.ac.kr

Dong-Seong Kim
Dept. of IT convergence engineering
Kumoh National Institute of Technology
Gumi-si, South of Korea
dskim@kumoh.ac.kr

***Abstract*—Driven by 6G networking, Edge Intelligence (EI) makes the most of the widespread edge resources to gain Artificial Intelligence (AI) insight. Future time-critical and data-intensive applications need distributed AI (DAI) and analytics solutions on the Edge computing platforms to enable EI from small devices to whole industrial factories. To deal with critical challenges of DAI implementation such as communication reliability, resource constrains and heterogeneity of edge devices, and dynamic nature of edge computing environment, the integration of digital twin (DT) technology is important to form an efficient framework. With this framework, efficient DT models are developed for edge devices, edge servers, and edge networks to predict the states of physical entities using probabilistic graphical models (PGMs) and machine learning (ML) algorithms, especially Generative Adversarial Network (GANs). This paper introduces the perspective for develop a framework to create the DT models for edge computing using GAN.**

***Index Terms*—Edge Computing, Digital Twin, Distributed AI, Edge Intelligence, Industrial IoT, 6G Networking, Generative Adversarial Network (GANs), Probabilistic Graphical Models (PGMs).**

I. Introduction

In recent years, the rise of edge computing has revolutionized the way we process data and perform complex computations. Enabled by 6G networking technology, edge computing has emerged as a promising ubiquitous and pervasive computing solution for IoT-connected devices to fulfill the requirements of 5G mobile networks. Fundamentally, edge computing supports to provide the low-latency computing services by offloading computation tasks to the nearby edge servers [1]. As edge devices become more powerful and ubiquitous, they are increasingly being used to run artificial intelligence (AI) algorithms at the edge of the network, allowing for faster and more efficient processing of data [2]. As a result, edge intelligence paradigms has been emerged to provide various intelligent service and applications [3].

Distributed AI refers to the use of multiple, interconnected devices and systems to perform AI tasks. Rather than relying on a single centralized processing unit, distributed AI systems are designed to distribute the processing load across a network of edge devices, cloud servers, and other connected systems. This approach can offer numerous advantages over traditional centralized AI, including improved scalability, fault tolerance, and reduced latency. Additionally, distributed AI can enable the creation of intelligent, autonomous systems that can operate in real-time, even in environments with limited connectivity or computational resources. As such, distributed AI has the potential to revolutionize a wide range of industries, from manufacturing and logistics to healthcare and transportation.

However, designing a DAI solution is still challenging in the large-scale edge computing system mainly due to the resource constrains and heterogeneity of edge devices and edge servers, the network size, and dynamic nature of edge computing. Edge devices typically have limited processing power, memory, and energy resources, which makes it challenging to implement complex AI algorithms on them. Therefore, the AI models must be optimized to work efficiently on edge devices. In addition, in a distributed AI system, edge devices need to communicate with a centralized server or other edge devices. However, communication latency can be a significant bottleneck, especially when dealing with real-time applications. Furthermore, the reliability of the network connecting edge devices can also be a challenge. The network may be unreliable, or edge devices may lose connectivity, leading to data loss or delays.

Recently, both the academia and industry have shown great interest in developing and applying digital twin (DT) technology for intelligent resource allocation and network management in the edge computing systems [4]. By using DTs to simulate edge devices and their behavior, it is possible to create distributed AI systems that can be trained and optimized in real-time, without the need for a central processing unit. In this context, DTs can serve as a bridge between the physical and digital worlds, allowing for the creation of intelligent, autonomous systems that can operate independently at the edge. This approach has numerous potential applications, ranging from autonomous vehicles and smart factories to healthcare and energy management. Overall, the use of digital twins for distributed AI in edge computing represents a promising avenue for advancing the field of AI and unlocking new possibilities for intelligent systems in a wide range of industries. In these regards, this paper introduces an framework integrating the DTs and edge computing to support the efficient and robust implementation of DAI solutions in industrial IoT applications.

II. DTs for Edge Computing

A DT is a digital representation of a physical product, and process with required accuracy and fidelity to simulate, and predict the physical performance, thus supporting the optimized decisions as shown in Fig. 1.

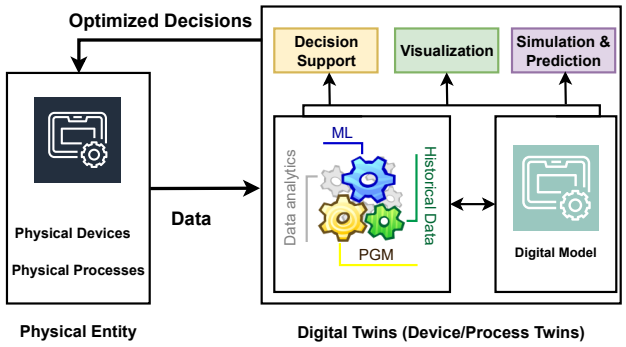


Fig. 1. DT modeling and features

When combined with edge computing, DTs can offer even greater benefits. There are existing related works proposing the optimal offloading in the MEC (Mobile Edge Computing) systems with the assistance of DT. For example, the DT is used to provide the estimated states of edge servers and training data to the centralized base station to derive the optimal offloading solution that minimize the offloading latency while maintaining the acceptable long-term migration cost. In [5], the DT supports UEs to select high-quality MEC servers, thus the offloading is efficient in term of energy consumption and latency. Similarly, the optimal server selection is derived by the assistance of DT in [6] through a set of iterative optimization processes. The paper [7] proposes a digital twin-empowered mobile edge computing (MEC) framework for intelligent vehicular lane-changing. The framework involves using DTs to model the road conditions and vehicle dynamics, which are used to predict the best lane-changing decisions for the vehicle. The MEC is used to enhance the performance of the digital twin-based lane-changing system by offloading some of the computational tasks to a nearby server. The proposed framework is evaluated using real-world vehicular data, and the results show that the digital twin-empowered MEC system achieves higher accuracy and efficiency compared to traditional approaches. The authors conclude that the use of digital twins in combination with MEC can significantly enhance the functionality and performance of intelligent vehicular systems, opening up new possibilities for the design and implementation of future intelligent transportation systems.

III. AI-enhanced Predictive DTs

A. DT Modeling

Despite their potential, the development of DTs is still very challenging, since it requires the collaboration of experts in multiple fields and the use of robust and affordable computational tools, able to integrate multiple physics as well as diverse solving technologies. The kind of resources that are only available to original equipment manufacturers (OEMs) or Tier-1 manufacturers [8].

However, predictive DT models can be built by machine learning (ML) and data analytics approach fed the historical data of physical entity, and probabilistic graphical models (PGMs) to accurately estimate the future states and performance of physical entities [9], [10].

As illustrated in Fig. 2, a PGM is constructed as a formal mathematical representation of a digital twin and its associated physical asset [11]. The declarative and general nature of the proposed DT model make it rigorous yet flexible, enabling its application at scale in a diverse range of application areas.

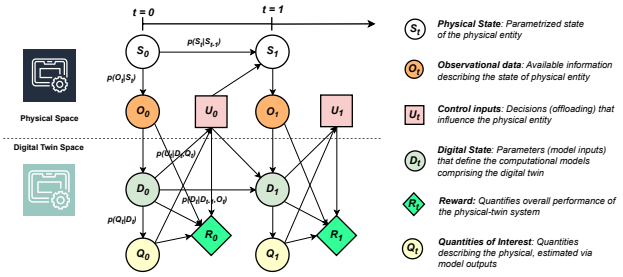


Fig. 2. An example of PGM.

B. DT Models enhanced by AI

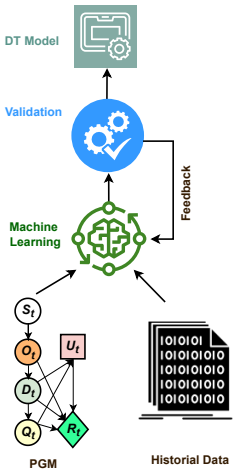


Fig. 3. DT modeling by ML and PGM

When fed with the historical data of physical entities, machine learning algorithms can be developed to create the efficient DT models for physical twins [12] (Fig. 3).

IV. Generative Adversarial Network (GAN)-based DTs

GANs are a type of deep neural network that are used to generate new data that is similar to a given set of training data. To create digital twins using GAN, you need to follow the following steps:

- Data Collection: Collect high-quality data from the physical twin in the form of images, videos, and sensor data.
- Data Preprocessing: Clean, preprocess, and normalize the collected data.
- GAN Training: Use the preprocessed data to train a GAN model. In this step, the generator network of the GAN learns to generate new data that is similar to the training data, while the discriminator network learns to distinguish between real and fake data.
- Digital Twin Generation: Once the GAN model is trained, use the generator network to generate new data that represents the digital twin.
- Validation and Refinement: Validate the generated digital twin data against the physical twin data and refine the model accordingly.
- Deployment: Deploy the digital twin model to simulate the behavior of the physical twin and use it for various applications, such as predictive maintenance, performance optimization, and fault detection.

V. Summary

To deal with critical challenges of DAI implementation such as communication reliability, resource constrains and heterogeneity of edge devices, and dynamic nature of edge computing environment, the integration of digital twin (DT) technology is important to form an efficient framework. With this framework, efficient DT models are developed for edge devices, edge servers, and edge networks to predict the states of physical entities using probabilistic graphical models (PGMs) and machine learning (ML) algorithms. In particularly, to implement DAI solutions efficiently, the developed DT models must predict and estimate the states of physical entities accurately enabled by the power of AI. This document introduces the perspective for develop a framework to implement AI algorithms on AI chips to create the predictive DT models.

VI. Framework of AI-enhanced Predictive DT Implementation on Chips

The Snapdragon Neural Processing Engine (SNPE) is a Qualcomm Snapdragon software accelerated runtime for the execution of deep neural networks. With SNPE, users can:

- Execute an arbitrarily deep neural network
- Execute the network on the Snapdragon CPU, the Adreno GPU or the Hexagon DSP.
- Convert Caffe, Caffe2, ONNXTM and TensorFlowTM models to a SNPE Deep Learning Container (DLC) file
- Quantize DLC files to 8 bit fixed point for running on the Hexagon DSP
- Debug and analyze the performance of the network with SNPE tools

Model training is performed on a popular deep learning framework (Caffe, Caffe2, ONNX and TensorFlow models are supported by SNPE as illustrated in Fig. 4.

After training is complete the trained model is converted into a DLC file that can be loaded into the SNPE runtime.

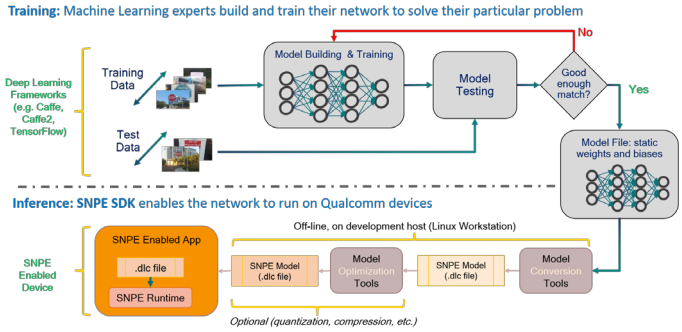


Fig. 4. DT modeling by ML and PGM

This DLC file can then be used to perform forward inference passes using one of the Snapdragon accelerated compute cores.

The basic SNPE workflow consists of only a few steps:

- Convert the network model to a DLC file that can be loaded by SNPE.
- Optionally quantize the DLC file for running on the Hexagon DSP.
- Prepare input data for the model.
- Load and execute the model using SNPE runtime.

Acknowledgments

This work was supported by Priority Research Centers Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology" (2018R1A6A1A03024003), and by the MSIT (Ministry of Science and ICT), Korea, under the Innovative Human Resource Development for Local Intellectualization support program (IITP-2023-2020-0-01612) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)" and the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT (RS-2023-00249687).

References

[1] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Communications Surveys and Tutorials*, vol. 19, no. 3, pp. 1628–1656, 2017.

[2] C. Chaccour and W. Saad, "Edge intelligence in 6g systems," in *Computer Communications and Networks*. Springer International Publishing, pp. 233–249.

[3] K. B. Letaief, Y. Shi, J. Lu, and J. Lu, "Edge artificial intelligence for 6g: Vision, enabling technologies, and applications," *IEEE Journal on Selected Areas in Communications*, vol. 40, no. 1, pp. 5–36.

[4] W. Sun, H. Zhang, R. Wang, and Y. Zhang, "Reducing offloading latency for digital twin edge networks in 6g," vol. 69, no. 10, pp. 12 240–12 251.

[5] T. Liu, L. Tang, W. Wang, Q. Chen, and X. Zeng, "Digital-twin-assisted task offloading based on edge collaboration in the digital twin edge network," *IEEE Internet of Things Journal*, vol. 9, no. 2, pp. 1427–1444.

[6] T. Do-Duy, D. Van Huynh, O. A. Dobre, B. Canberk, and T. Q. Duong, "Digital twin-aided intelligent offloading with edge selection in mobile edge computing," *IEEE Wireless Communications Letters*, vol. 11, no. 4, pp. 806–810.

[7] B. Fan, Y. Wu, Z. He, Y. Chen, T. Q. Quek, and C.-Z. Xu, "Digital twin empowered mobile edge computing for intelligent vehicular lane-changing," *IEEE Network*, vol. 35, no. 6, pp. 194–201.

[8] Y. Wu, K. Zhang, and Y. Zhang, "Digital twin networks: A survey," *IEEE Internet of Things Journal*, vol. 8, no. 18, pp. 13 789–13 804.

[9] M. G. Kapteyn, D. J. Knezevic, and K. E. Willcox, "Toward predictive digital twins via component-based reduced-order models and interpretable machine learning."

[10] M. G. Kapteyn and K. E. Willcox, "From physics-based models to predictive digital twins via interpretable machine learning," *CoRR*, vol. abs/2004.11356. [Online]. Available: https://arxiv.org/abs/2004.11356

[11] M. G. Kapteyn, J. V. R. Pretorius, and K. E. Willcox, "A probabilistic graphical model foundation for enabling predictive digital twins at scale," *Nature Computational Science*, vol. 1, no. 5, pp. 337–347.

[12] D. Koller and N. Friedman, *Probabilistic Graphical Models: Principles and Techniques*, ser. Adaptive Computation and Machine Learning series. MIT Press. [Online]. Available: https://books.google.co.kr/books?id=dOruCwAAQBAJ

Examining the Effect of PEMFC Gasket Materials Under Aging Techniques and Tensile Testing

SooHyun Park

Dept. Mechanical Engineering
Kumoh National Institute of Technology

Gumi, Republic of Korea
psh2344342@naver.com

JangWook Hur

Dept. Mechanical Engineering
Kumoh National Institute of Technology

Gumi, Republic of Korea
hhjw88@kumoh.ac.kr

1

Abstract—**Proton-exchange membrane fuel cells (PEMFC) gasket material degradation is a concern in electric vehicles because it can lead to leaks of hydrogen fuel, which is highly flammable. This can be caused by exposure to high temperatures and pressures, as well as exposure to hydrogen fuel itself. Results showed that the LSR material outperformed the EPDM material, with lower average RMSE and MAE values of 0.3% and 0.25%, respectively, compared to the EPDM values of 0.5% and 0.75%, respectively**

Keywords—accelerates aging, FEM, PEMFC, polynomial regression, machine learning

I. INTRODUCTION (*HEADING 1*)

Hydrogen fuel cells, of which proton-exchange membrane fuel cells (PEMFC) are a type, have been researched as a potential clean energy source because they produce only water vapour as a byproduct. The gasket material for hydrogen fuel cells is a crucial component in ensuring the safe and efficient operation of the fuel cell system. Overall, these gasket materials/sealants have excellent chemical and physical properties that make them suitable for hydrogen fuel cell applications. EPDM and silicon rubber are known for their flexibility and durability, fluoroelastomers for their chemical and high-temperature resistance, and PTFE for their chemical and high-temperature resistance. Hydrogen fuel cells operate at high temperatures and pressures and are exposed to harsh chemicals and gases. Therefore, selecting a gasket material that can withstand these conditions without degrading over time is essential. Interestingly, we have designed a degradation framework considering tensile testing and accelerated aging techniques in selecting the best material for the hydrogen fuel cell.

II. MATERIAL SELECTION FRAMEWORK

A. Selecting a Template (Heading 2)

Selecting the suitable gasket material is crucial in ensuring the reliable and efficient performance of a Proton Exchange Membrane Fuel Cell (PEMFC). Following these steps, you can create a PEMFC gasket material selection framework to help you select the best gasket material for your specific fuel cell application. During FEA simulations, contact pressure and von Mises stress are calculated based on the component's material properties, loads, and geometric characteristics. The results of these simulations 277 can be used to make informed design decisions and improve the performance and reliability of the component. The Figure 2 explains the breakdown process for the gasket materials and decision-making paradigm.

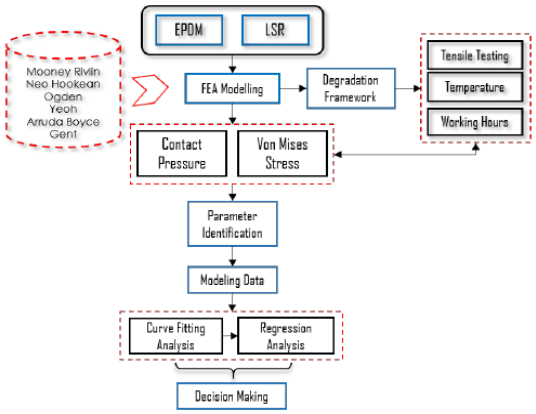
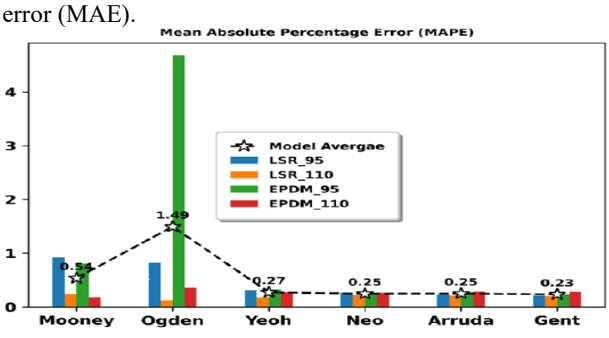


Figure 1 PEMFC Material Selection Flowchart

III. NON-LINEAR REGRESSION PERFORMANCE METRICS

Regression metrics are used to evaluate the performance of a regression model, which predicts a continuous numerical output variable based on one or more input variables. Mean Absolute Percentage Error (MAPE) measures the average percentage difference between the predicted and actual values. R-Squared (R^2): This measures the proportion of variance in the output variable that the model explains. It ranges from 0 to 1, with higher values indicating a better fit. Root Mean Squared Error (RMSE) is the square root of the MSE and provides the mean error in the same units as the output variable. The average absolute difference between the predicted and real values is measured by mean absolute error (MAE).



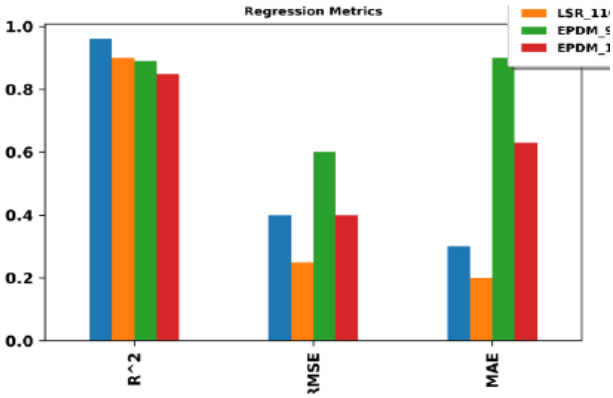


Figure 2 Regression metrics for EPDM and LSR PEMFC materials considering aging conditions

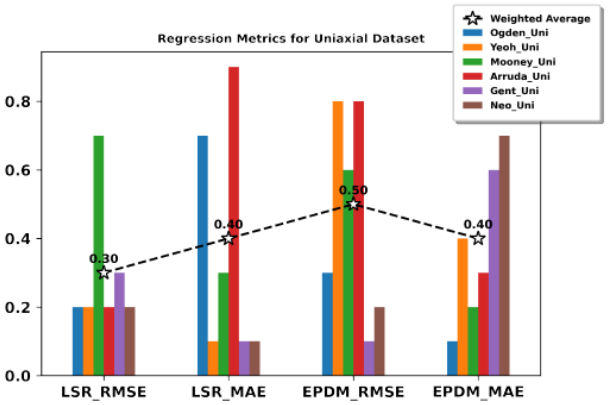


Figure 3 Regression metrics for EPDM and LSR PEMFC materials considering tensile conditions

Figure 2 shows the regression assessment plot considering the aging conditions deployed to the PEMFC gasket material. Subsequently, we tried to check to performance metrics of the EPDM and LSR material considering uniaxial tensile conditions. Figure 3 gives insight to the regression metrics analysis under uniaxial tensile conditions for the gasket materials. The LSR material had the least RMSE value of 0.30% compared to the EPDM material with RMSE value of 0.50%. However, under the MAE assessment, both PEMFC gasket material under consideration had the same average value of 0.50% considering the tensile conditions. Overall, the LSR gasket material is a suitable gasket material in terms of the decision framework analysis considering both the aging and tensile conditions which is enough paradigm in choosing the right material.

IV. CONCLUSIONS

This study showed that EPDM and LSR have suitable mechanical properties for PEMFC gasket applications. However, the FEA analysis revealed that LSR is more resistant to von Mises stress than EPDM, making it the better choice for higher-pressure applications. The LSR material had the lowest RMSE value of 0.30%, indicating higher accuracy compared to the EPDM material, which had an RMSE value of 0.50%. Additionally, the LSR gasket material had a significantly lower average MAE value of 0.25%, compared to the EPDM material with an average MAE value of 0.75%. These findings suggest that LSR is a better material for PEMFC gaskets under aging conditions, and could potentially improve the performance and durability of fuel cells. Further research could investigate the long-term effects of using LSR gasket materials in fuel cells to validate these results.

ACKNOWLEDGMENT

THIS STUDY WAS CONDUCTED AS A RESULT OF RESEARCH BY THE MINISTRY OF COMMERCE, INDUSTRY AND ENERGY AND THE KOREA INSTITUTE OF INDUSTRIAL TECHNOLOGY EVALUATION AND MANAGEMENT (K_G012000998304).

REFERENCES

[1] Singla, MK; Nijhawan and P.Oberoi; A.S Hydrogen fuel cell technology for cleaner future a review. Environ Sci Pollut Res 28, 15607-15626(2021)

[2] John Meurig T.; Peter P. E.; Peter J. D.; Gari P. O.; Decarbonising energy: The developing international activity in hydrogen technologies and fuel cells, Journal of Energy Chemistry, Volume 51, 2020

[3] Asif J.; Sikander R.; Tanveer I.; Hafiza Aroosa A. K.; Haris M. K.; Babar Azeem, M.Z. Mustafa, Abdulkader S. H.; Current status and future perspectives of proton exchange membranes for hydrogen fuel cells, Chemosphere, Volume 303, Part 3, 2022

Machine Learning-Based Energy-Efficient Task Offloading for Mobile Edge Computing in UAV-Assisted Maritime Networks

Van Dat Tuong
School of Computer Science and Engineering
Chung-Ang University
Seoul, Republic of Korea
vdtuong@uclab.re.kr

Nguyen Cong Luong
Faculty of Computer Science
PHENIKAA University
Hanoi, Vietnam
luong.nguyencong@phenikaauni.edu.vn

Sungrae Cho
School of Computer Science and Engineering
Chung-Ang University
Seoul, Republic of Korea
srcho@cau.ac.kr

***Abstract*—Maritime activities have been rapidly conducted for economic development. To support maritime devices used in these activities to reduce consuming energy, mobile edge computing (MEC) is combined with unmanned aerial vehicles (UAVs) to provide powerful computing capabilities with UAV servers. A two-layer UAV-assisted maritime network is considered that includes a centralized Macro-UAV (M-UAV), a group of distributed Small-UAVs (S-UAVs), and multiple maritime ship devices. This study aims to minimize energy consumption for all devices in the network. Because of the special characteristics of the maritime environment with unstable channel conditions, it is challenging to solve the formulated problem using conventional optimization approaches. Therefore, we solve the energy-efficient task offloading problem using a machine learning approach.**

***Index Terms*—Energy-efficient task offloading, machine learning, maritime network, mobile edge computing (MEC), unmanned aerial vehicle (UAV) trajectory design**

I. Introduction

In recent years, the importance of the vast ocean on economic development based on its great benefits of natural resources and potential for trade, tourism, telecommunication, and military services has been well aware by many countries [1]. Thus, numerous ocean activities have been invested in and conducted globally using advanced equipment and technology. Besides traditional activities, e.g., fishery and transportation, lots of new activities, e.g., environmental monitoring [2], seabed resource explorations, maritime tourism and rescue [3], and military activities have been implemented. Furthermore, a large number of smart devices, such as mobile devices, remote robots, high-resolution cameras, specialized equipment, etc., have been used in these activities. The presence of communication networks in the oceanic environment, known as maritime networks, is promoted in order to optimize the energy and timing costs for processing the tasks and getting the results. Therefore, it is necessary to design an efficient solution to reduce the costs of task completion in maritime networks.

Although the devices are often installed with limited computational capability, they have to run computation-intensive tasks resulting from the rapid development of the Internet of Things (IoT) applications and services. To tackle this challenge, the task offloading technique with mobile edge computing (MEC) is considered as an efficient solution. In MEC, powerful computing resources are provided to support task offloading for maritime devices. Compare with cloud computing technology, MEC employs computing servers at the network edge, which is significantly close to the user terminals; hence, it can reduce consuming energy and transmission delay for the users [4]. Considering energy as a vital element of IoT devices [5], this paper focuses on optimizing energy consumption for maritime devices.

However, deploying MEC in a maritime network is difficult because the maritime environment is lack of common communication infrastructures such as terrestrial base stations. To this end, unmanned aerial vehicles (UAV) can be used to combine with MEC in maritime networks [6]. Specifically, MEC servers are deployed in UAVs to support task offloading for maritime users. Based on its flexibility and high altitude, a UAV can establish line-of-sight links with maritime devices. Moreover, UAVs often operate at a lower altitude than satellites that are closer to the maritime terminals for fast response.

There are many differences between communications in the maritime environment and the terrestrial environment. First, the coverage size of maritime networks is often larger than that of terrestrial networks. But the density of user terminals in terrestrial networks can be significantly higher than that in maritime networks. This results in long communication links in the maritime networks. Second, the accessibility in terrestrial networks is significantly better than that in

maritime networks. Current maritime networks mostly use high-frequency (HF) and very high-frequency (VHF) radio channels. Third, considering the geographical factor, the two environments have big different climatic conditions. Affected by frequent extreme weather conditions such as high humidity and sea clutter, radio channels in maritime networks are more unstable than those in terrestrial networks. These differences between maritime networks and terrestrial networks create big challenges to apply conventional optimization approaches for solving a specific problem of maritime networks. Meanwhile, machine learning approaches that can flexibly adapt to the unstable maritime environment are considered a potential approach to solving problems of maritime networks [7].

Motivated by the above discussions, this paper studies the energy-efficient task offloading problem in MEC for UAV-assisted maritime networks. The aim is to optimize the energy consumption of all maritime devices.

II. System Model

This section introduces a two-layer UAV-assisted maritime network and the task offloading operations.

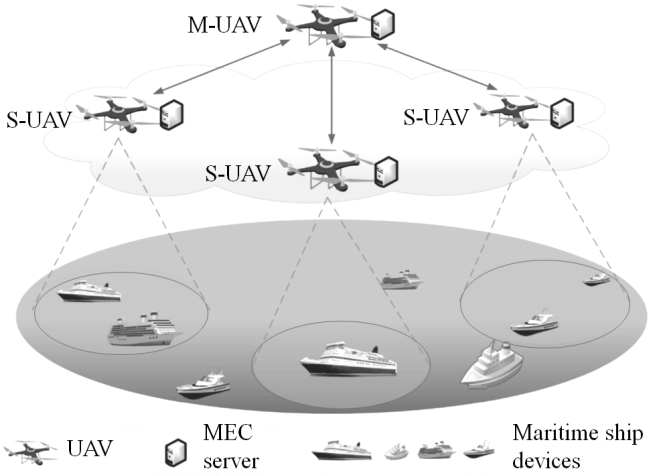


Fig. 1. Task offloading in UAV-assisted maritime networks.

We consider a two-layer UAV-assisted maritime network as illustrated in Fig. 1. The network comprises a Macro-UAV (M-UAV), a group of distributed Small-UAV (S-UAVs), and multiple maritime ship devices, which have to complete computation-intensive tasks. In the first layer, the tasks can be offloaded from maritime devices to S-UAV servers. Notably, the computational capability of each S-UAV server is limited. Therefore, when each S-UAV is overloaded with offloading tasks, the tasks are offloaded further to M-UAV as depicted in the second layer. The M-UAV server is assumed to be empowered with infinite computing capability so that it can process any offloaded task. Because the energy consumed for task processing is significantly larger than that for result transmission, we only consider energy consumption for task processing and task uploading. The aim is to minimize the energy consumption for task completion by optimizing the trajectory of the UAVs.

Acknowledgment

This research was supported by the MSIT (Ministry of Science and ICT), Korea, under the ITRC (Information Technology Research Center) support program (IITP-2023-RS-2023-00258639) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

References

[1] Y. Huo, X. Dong, and S. Beatty, "Cellular communications in ocean waves for maritime internet of things," *IEEE Internet of Things Journal*, vol. 7, no. 10, pp. 9965–9979, 2020.

[2] V. Fontana, J. M. D. Blasco, A. Cavallini, N. Lorusso, A. Scremin, and A. Romeo, "Artificial intelligence technologies for maritime surveillance applications," in *2020 21st IEEE International Conference on Mobile Data Management (MDM)*. IEEE, 2020, pp. 299–303.

[3] L. Liu, Q. Gu, L. Li, and X. Lai, "Research on maritime search and rescue recognition based on agent technology," in *2020 International Conference on Artificial Intelligence and Electromechanical Automation (AIEA)*. IEEE, 2020, pp. 201–205.

[4] B. Cao, L. Zhang, Y. Li, D. Feng, and W. Cao, "Intelligent offloading in multi-access edge computing: A state-of-the-art review and framework," *IEEE Communications Magazine*, vol. 57, no. 3, pp. 56–62, 2019.

[5] V. D. Tuong, N.-N. Dao, W. Noh, and S. Cho, "Dynamic time division duplexing for green internet of things," in *2022 International Conference on Information Networking (ICOIN)*. IEEE, 2022, pp. 356–358.

[6] H. Wang, Y. Wang, Y. Ma, and B. Lin, "Resource allocation for ofdm-based maritime edge computing networks," in *2020 13th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. IEEE, 2020, pp. 983–988.

[7] N. V. Tam, N. Q. Hieu, N. T. T. Van, N. C. Luong, D. Niyato, and D. I. Kim, "Adaptive task offloading in coded edge computing: A deep reinforcement learning approach," *IEEE Communications Letters*, vol. 25, no. 12, pp. 3878–3882, 2021.

Low power consumption high precision Battery Management System

*WanHae Jeon
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
jeonwahae@sunmoon.ac.kr

Hikasa Akio
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
3445djwgodjw@sunmoon.ac.kr

MinSo Bang
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
alsth1004@sunmoon.ac.kr

Yokoi Dakayuki
dept. of advenced automotive eng.
Univ. of Sunmoon
Yuseong-gu, Daejeon city
hyojee.choi@jwpre.com

Khurshid Hussain
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
khurshid12@sunmoon.ac.kr

Innyeal Oh*
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
innyealoh@sunmoon.ac.kr

***Abstract*—An indirect measurement method with an isolated structure was proposed for safe battery voltage measurement. The battery voltage of all battery cells was monitored using a single indirect measurement path to solve the problem of device imbalance and temperature difference. In a process of sequentially recognizing a battery cell voltage through a single path, a battery monitoring voltage is sequentially generated together with a known voltage, REF_V, and a battery voltage to be accurately measured at any temperature. As a result of the design, the temperature compensated battery voltage was obtained within 5.12 usec based on 16 battery cells, The proposed single path multi-cell technology reduced power consumption to 1.6 mW, which is 1/3 of the results of previous studies. This paper automatically outputs a battery code value for battery cell recognition along with REF_V and battery cell voltage, so Battery system operations can be performed without the control of CPU.**

Keywords—monitoring, BMS, low power, precision, isolation, indirect method, temperature

INTRODUCTION

As environmental problems such as warming and air pollution have recently emerged, countries around the world are expanding the supply of electric vehicles as part of their eco-friendly policies. However, as electric vehicles are distributed, electric vehicle fires continue to occur. This electric vehicle fire is mainly caused by excessive battery overload caused by BMS(Battery Management System) malfunction [1]. The existing BMS measured the voltage of the battery by physically connecting each battery cell and controller. However, if the battery voltage operating up to 1, 000V or higher incorrectly affects the control board operating at low voltage, the control board is destroyed and the battery cannot be managed, leading to a battery explosion. Thus, to solve these problems, an indirect measurement technology was proposed and designed to physically isolate the battery cell and the low voltage control BMS board [2]. As battery usage expands, BMSs developed from Infineon, TI (Texas Instrument), and LT (Linear Technology) are being implemented as semiconductors [3],[4]. TI and LT companies have also implemented wireless functions by adding them to BMS circuits. However, TI and LT's wireless functions are limited to transmitting battery cell voltage information obtained through BMS operations to the outside. In terms of delivering battery information to the outside, it is convenient for users (or drivers) to easily obtain battery information, but risk factors that can be destroyed by affecting BMS semiconductor circuits operating at low voltages of thousands of volts have not been eliminated. In the paper [2] that proposed a safe BMS technology by separating a high voltage battery of thousands of volts and a BMS circuit of low voltage operation, there was a difference in temperature characteristics for each LED in using multiple LEDs for each battery cell. As a result, there is a problem of causing an error in voltage recognition for each battery cell. This paper proposed that it could accurately monitor several battery cell voltages through a single path and compensate the temperature at the same time. Each battery cell was assigned a code to recognize the battery cell voltage by sending it along with a code to determine which battery cell voltage was measured when measuring voltage over a single path.

HIGH PRECISION BMS DESIGN

The existing isolated BMS monitoring has a monitoring circuit having a parallel structure for each battery cell, and thus has an error in a monitoring voltage for each battery cell due to a device process differences. In particular, a difference in temperature characteristics of each device according to temperature increases an error in a battery cell voltage monitoring value with respect to temperature. In the technology [2] operating on this basis, it operates with a voltage error of 5 mV. However, accurate voltage recognition for each battery cell through BMS monitoring operations is essential for stable battery operation. In this paper, for the accurate operation of monitoring the battery cell voltage according to the device error of the parallel structure, a BMS monitoring circuit that recognizes all battery cell voltages through one single path is implemented. Figure 1 shows the single-path multi-battery cell BMS monitoring structure that supports high-precision battery cell voltage measurement. As illustrated in Fig. 1, the isolated battery cell voltage monitoring is composed of Switch Array Block, Code Generation(Gen.) Block, and Clock Gen. Block. Clock Gen. Block is a circuit that generates a clock signal required to implement a BMS monitoring circuit that recognizes all battery cell voltages with one serial circuit, and Code Gen. Block is a battery cell that controls the operation of the Switch Array Block. And the Switch Array Block is a circuit that sequentially monitors the voltage of N battery cells through a single-path isolation circuit based on the Seq-M-1 signal obtained from Code Gen. Block. The unit cell battery voltage and the REF_V obtained are transmitted to an electronic control unit (ECU) having an isolated structure, and thus the battery is managed based on the isolated ECU to support stable battery operation.

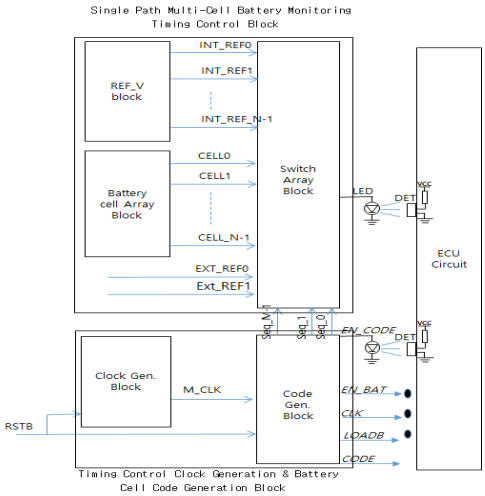


Figure 1. Proposed isolated structure for battery voltage monitoring

A. Timing Control and Battery Code generation circuit design

A circuit that generates a control clock and a battery code is required to obtain a signal required for the operation of the Switch Array Block in Figure 1. For this performance, the Clock Gen. Block and Code Gen. Block in Figure 1 were designed, and Figure 2 shows them in detail.

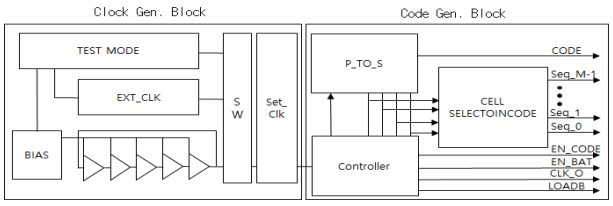
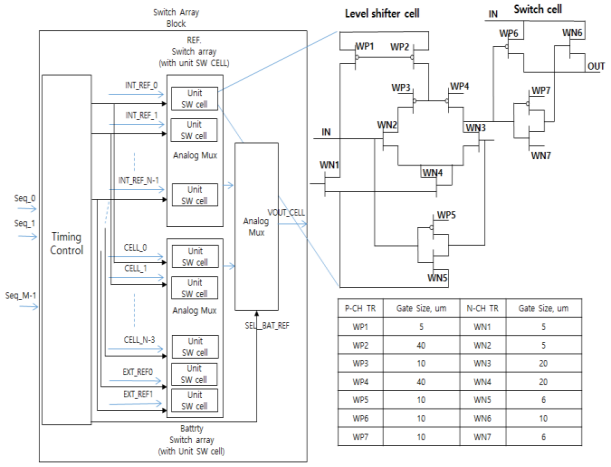


Figure 2. Control Clock Generation and Battery Code Generation Circuit

The Clock Gen. Block shown in Figure 2 is selected between EXT_CLK from the outside and an internal clock generated internally. Through TEST MODE, EXT_CLK and internal clock can be selected to verify Code Gen. Block, and the voltage and current of BIAS of the internal clock can be monitored to see if there is any problem with the operation of the internal clock. Code Gen. Block in Figure 2 receives the final clock selected from Clock Gen. Block to generate a signal to sequentially obtain the battery cell and REF_V voltage output to control the operation of Switch Array Block in Figure 1, and recognizes the EN_BAT signal, EN_CODE signal, and the battery cell selection signal, Seq_M-1 CODE to get the selected battery voltage.

B. Switch Array Block Design

Figure 3 shows the Switch Array Block operating based on the Seq_M-1 signal obtained from Code Gen. Block. A signal generated from Code Gen. Block is input to Switch Array Block, and it is sequentially input to the Unit analog Switch(SW) cell together with each battery cell voltage and internal REF_V (voltage for temperature compensation) to be selected. In order to be used as a reference signal, REF_V, for temperature compensation, 16 reference signals from 3.0V to 4.3V were internally made with 0.08V unit step. It was output sequentially through the REF Switch array shown in figure 3. In addition, for the battery unit cell voltages to be monitored sequentially, the Battery Switch array shown in Fig. 3 is configured to be output. In measuring all battery cell voltages, the battery cell voltages were measured together with internal (or external) REF_V to perform a temperature compensation operation. If 16 battery cell voltages are measured until the final temperature compensated battery cell voltage is obtained, all battery voltages are recognized within 5.12usec time.



P-CH TR	Gate Size, um	N-CH TR	Gate Size, um
WP1	5	WN1	5
WP2	40	WN2	5
WP3	10	WN3	20
WP4	40	WN4	20
WP5	10	WN5	6
WP6	10	WN6	10
WP7	10	WN7	6

Figure 3. Design Switch Array Block Circuit Structure

C. Single Path Multi-Cell Temperature Compensation

The temperature compensation is performed by sequentially determining the battery cell voltage compared with a known REF_V value. The temperature of the battery cell voltage may be compensated based on not only the internal REF_V, but also the external REF_V. For the internal REF_V value, the value from 3.0V to 4.3V was divided into 16 parts so that the REF_V value can be sequentially obtained. The step value of one internal REF_V is a value of 0.08V, and the measured battery cell voltage is set as the battery cell voltage value with the REF_V proximity value. The battery voltage value was not only defined with a value close to that per step, but also corrected through the linear formula of the internal REF_V to obtain an accurate battery voltage value. The existing temperature compensation technology sets a lookup table (LUT) that defines a battery voltage value by temperature based on a temperature sensor value attached to a battery, and when measured, a battery cell voltage with respect to a corresponding temperature is output. The exact battery cell voltage cannot be known for a temperature not set in the LUT, and the LUT-based existing technology cannot recognize the exact battery cell voltage because the temperature causes inaccurate temperature compensation. Figure 4 shows the temperature compensation operation algorithm proposed based on the designed single path multi-battery monitoring circuit. Through the temperature compensation algorithm in Fig. 3, Temperature compensated battery cell voltage is set after receiving the temperature compensation from the 00th battery cell to the Nth battery cell (CELL_N-1), and then receives the temperature compensation internal or external REF_V sequentially from INT_REF00 to INT_REF_N-1.

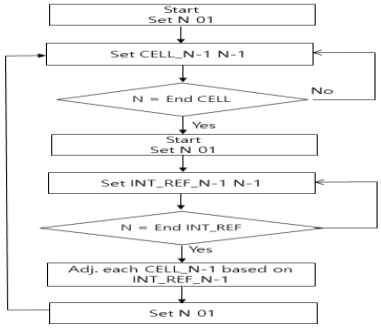


Figure 4. Temperature compensation algorithm for monitering voltage.

THE RESULT

This paper presents a single-path multi-cell-based BMS monitoring circuit. Figure 5 shows the timing chart of the signal in which the designed Switch Array Block, Clock Gen. Block, and Code Gen. Block in Figure 1 operate. In order to sequentially output the voltage of the battery multi-cell to one path, the P_to_S[N-1:0] signal, which is a code signal for selecting the battery cell, and the VOUT_CELL signal, which is a battery voltage, are displayed. The VOUT_CELL voltage is a signal for outputting the REF_V voltage and the voltages of the battery cells, and it can be seen in Fig. 5 that the voltage is output with the code signal P_to_S[N-1:0] together with the VOUT_CELL so as to know which REF_V and which battery cell voltage. In detail, the EN_CODE signal receives a code signal in a low state for EN_CODE, and a battery cell voltage or REF_V in a low state for EN_BAT so that the ECU may receive a voltage value separately with a code when receiving a multi-cell voltage and REF_V.

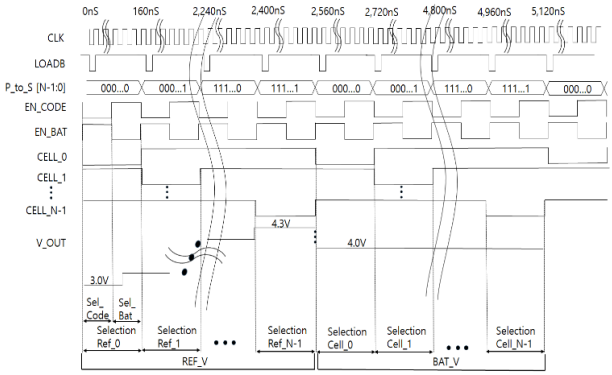


Figure 5. Monitering voltage Operation Timing Chart

In order to stably transmit the code signal, a LOADB signal was transmitted before transmitting the code value together with the clock signal. The LOADB signal was transmitted after the half-clock signal after the EN_CODE signal was generated as Low, and the code signal was transmitted after the half-clock signal after the LOADB signal was generated. In the verification, the half clock speed was set to 20 nsec. After the code transmission, the EN_BAT signal becomes low after two clock signals after the code value is transmitted to take the battery cell voltage or the REF_V stably. As shown in Fig. 5, it shows the REF_V after the VOUT_CELL output signals (BAT_V) was sequentially output. It can be found that the REF_V voltage is output from 3.0V to 4.3V in 0.08V unit. As a result of the design, it was possible to eliminate errors due to differences in characteristics of devices with one LED. Assuming that the battery pack is composed of 16 battery cells, it was verified that each of 16 temperature compensation REF_V and all battery cell voltages were recognized within 5.12. Internally, the Clock Gen. and Code Gen. circuits were designed to enable BMS operation without a CPU control. In the existing technology [2], 16 LEDs of about 300uW per unit were used, and the total power consumption was 4.8mW (= 0.3mW x 16 cells). However, in this paper , by using one LED of 300uW, the entire power consumption operated at a total of 1.6mW consumption, including 1.3mW, as a single path multi-cell monitoring core function. This can reduce power consumption by 1/3 of the results of previous studies, resulting in reduced battery consumption in the BMS operation. The chip layout showing the Clock Gen. Block, Code Gen. Block, Switch Array Block is shown in Figure 6.

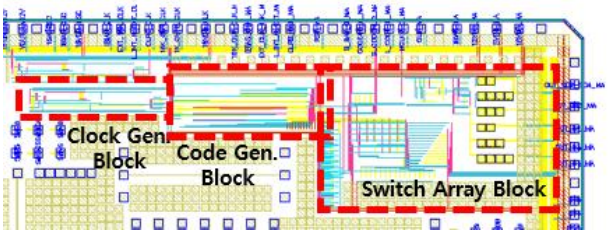


Figure 6. BMS Layout for monitering voltage.

CONCLUSION

This paper is designed with an isolated single-path multi-cell structure so that the BMS semiconductor circuit operates without a problem even in a high voltage battery environment that occurs up to 1,000 volts in electric vehicles. The battery cell monitoring operation, which performs temperature compensation based on a 16-cell battery pack, can be implemented in 5.12 usec time, and all battery cell voltages can be monitored in a single path. In this paper, we were able to overcome the difference in device characteristics by performing a single path multi-cell BMS operation. The single-path multi-cell BMS operation could reduce power consumption from 4. 8mW to 1. 6mW, which is one-third of the existing design [2], hence reducing battery consumption in the BMS operation. This paper automatically generates REF_V and battery cell voltage, which are temperature compensation voltages, together with a code value so that a battery voltage of a multi-cell could be sequentially known, so that an accurate BMS operation is possible without the ECU control.

ACKNOWLEDGMENT

The chip fabrication was supported by the IC Design Education Center(IDEC), Korea.

REFERENCES

[1] Hyung-Sik Kim, Soo-Ho Lee, Tae-Dong Kim, A-Young Choi, "Experimental Study on Fire Characteristics of Adjacent Electric Vehacles, " Journal of the Korea Academia-Industrial cooperation Society, Vol. 23, No. 4, pp. 343-350, Apr. 2022

[2] Jin Uk Yeon, Ji Whan Noh, Innyeal Oh, "Temperature-Compensated Overcharge Protection Measurement Technology, " Advances in Science, Technology and Engineering Systems Journal Vol. 8, No. 2, pp. 24-29, Mar. 2023

[3] Infineon, "TLE9012DQU, Li-Ion battery monitoring and balancing IC," Infineon web-site, 2022.02

[4] Texas Instruments, "BQ79616-Q1, 16-S automotive precision battery monitor, balancer and integrated protector with ASIL-D compliance," Texas Instruments web-site, 2022.09

Isolated cell balancing circuit

* Hikasa Akio
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
3445djwgodjw@sunmoon.ac.kr

Wanhae Jeon
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
jeonwahae@sunmoon.ac.kr

Daehun Kim
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
kdeh1218@sunmoon.ac.kr

Yokoi minori
dept. of advenced automotive eng.
Univ. of Sunmoon
Yuseong-gu, Daejeon city
a5years@sunmoon.ac.kr

Rimhea Kim
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
kimrimhoe123@sunmoon.ac.kr

Innyeal Oh*
dept. of advenced automotive eng.
Univ. of Sunmoon
Asan city, Chungnam provence
innyealoh@sunmoon.ac.kr

***Abstract*—Batteries are being thermal runaway or fire as electric vehicles operate on battery voltage lately. The cause of the thermal runaway fire is caused by voltage differences battery cells when charging and discharging batteries. Therefore, it is important to accurate balancing operation the battery with BMS. The balancing operations adjust the difference between high voltage and low voltage among battery cells. In this paper, the balancing circuit was designed based on a separate structure as a circuit for stably maintaining the balance of each unit battery in the battery pack structure constituting the high voltage. Next, the balancing circuit applies a controller interface block including flip-flop so that balancing can start and end at any time when balancing is required. And finally the balancing circuit was designed to operate with a low power consumption of 180uW, minimizing the battery power consumption required to perform the balancing operation.**

Keywords— BMS, balancing, battery, voltage, isolation, low consumption, semiconductor

I. INTRODUCTION

Recent study reports that batteries used in electric vehicles are at risk of thermal runaway or fire [1]. This is due to the continuous charging and discharging of batteries, resulting in differences in voltage between cells or overcharging in any cell. A battery pack consists of a series or parallel combination of multiple cells. Battery pack should be charged and discharged at the same voltage for all battery cells under any serial/parallel conditions in the pack. If a voltage difference occurs in any battery cell, power loss may be concentrated due to the current difference. In particular, electric vehicle batteries consist of more than 7,000 battery cells, and even an inaccurate voltage state in just one battery cell during charging and discharging can accelerate cell degradation and failure [2], leading to battery explosion. In contrast, it is important that all battery cells are charged and discharged at the same voltage, which requires battery balancing operation. To solve this problem, papers related to balancing have been published. The paper [3] suggests a balancing operation circuit that controls the current through a function of selecting the resistance connection of the battery cell, but the method risks the high voltage of the battery breaking of a circuit operating at a low voltage of 5V or less. The other paper [4] implements self-balancing based on the resistance and current operation characteristics of the battery cell. Although it was confirmed that the amount of electricity moves from a high normal cell to a low deteriorated voltage when charging multiple battery cells so that current and voltage operate evenly, [4] assumes self-balancing of cells connected in parallel to overcharging. In particular, if [3-4] technology is a battery cell that does not follow the resistance and current operation characteristics of the cell among aged battery cells, there is a problem that incorrect self-balancing operation occurs. The paper [5] implemented an active balancing circuit using a photo detector. The active balancing operation structure using photo detector still has the problem of breaking BMS balancing circuit with up to thousands of volts available in electric vehicles because high voltage battery voltage is inputted to CMOS semiconductors. This work proposed that the balancing circuit is isolated from high-voltage batteries to ensure a stable balancing operation to overcome the heating problem caused by voltage difference in an environment that combines batteries operating up to thousands of volts in electric vehicles.

II. IOLATED BALANCING CIRCUIT

A circuit whose battery balancing operation is isolated from the high-voltage battery and operates safely was presented with a structure that separates the control part and the balancing part.

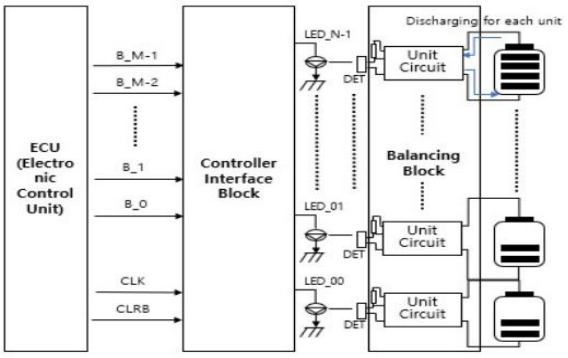


Figure 1. The proposed isolated balancing circuit

As shown in FIG. 1, the balancing circuit of the high voltage battery may be divided into a controller interface block and a balancing block to operate as an indirect connection, and thus may be independently controlled in a single cell voltage unit. In other words, the controller interface block inputted by synchronizing the CLK with the M-bit signal command to select the battery required for discharge from the electronic controller unit (ECU) transmits information indirectly to the balancing block circuit.

A. Controller Interface BLOCK

A controller interface block that supports a control operation was designed to receive a command of M bit from ECU to select a battery requiring discharging and perform a balancing operation. In Equation (1), N means the number of batteries, and M means the number of digital bits for selecting N batteries. The relationship between N and M can be expressed by equation (1).

$$N \leq 2^{M} \quad (1)$$

As illustrated in Fig. 2, the controller interface block has the controller interface root cell and the flip-flop corresponding to the number of battery cells. N battery cell values (C_0, C_1, and C_N-1) output from the controller interface root cell were input to the flip flop according to the command B_0, B_1,..., B_M-1 of M bits coming from the ECU.

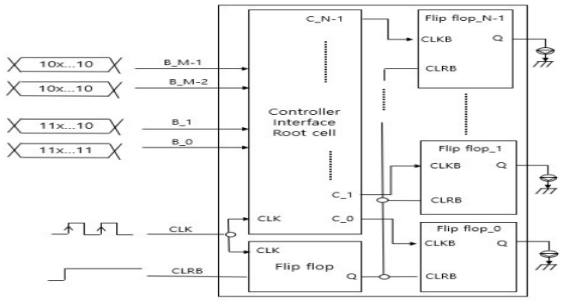


Figure 2. Controller Interface Block Circuit Brief Block Diagram

During the initial balancing operation, the CLRB signal helps to prevent discharging until the voltage of all battery cells is monitored and a problematic battery voltage is discharged. This was also operated as a flip-flop with the CLK signal so that there was no problem with the noise signal.

B. Balacing Block

The balancing circuit supports each cell to be charged and discharged at the same voltage in a battery pack including a combination of all of battery cells. Recently, electric vehicles have been able to obtain voltage up to thousands of volts by stacking them in series with battery cells.

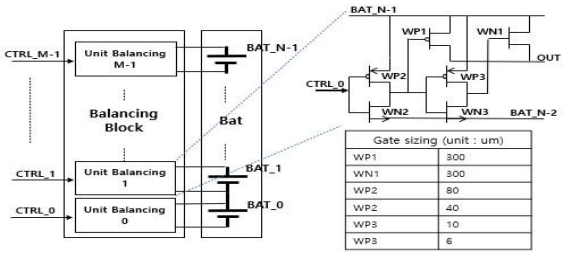


Figure 3. Balancing Block Structure

For battery cells whose voltages do not match in this environment, all cells must operate at the same voltage while performing the balancing operation, and the voltage of thousands of volts affects the balancing circuit [3-5]. To solve this problem, the controller interface block and the balancing block were presented to operate in isolation. Figure 3. is a structural diagram illustrating a balancing block.

to 100mA battery cell voltage. WP1 and WN1 are operated through WP2, WN2, WP3, and WN3 by control voltage (CTRL) so that balancing operation can be performed safely.

Where BAT block is to be configuring a battery group of N. Figure 3 shows the structure of the circuit of the unit balancing circuit_M-1 that makes up the balancing block. It is embodied so that the controller interface block operates to select one of the N battery cells for balancing by receiving an M-bit input from the external ECU. The balancing block consists of as many unit balancing circuits_M-1 as the number of battery cells. As shown in Figure 4, the unit balancing circuit_M-1 is directly coupled to a battery operating at up to several thousand volts, while the battery cells and the unit balancing circuit_M-1 are configured as a cell-by-cell circuit. The controller interface block also operates with the high-voltage balancing block, but can operate stably.WP1 and WN1 of the unit balancing circuit_M-1 are designed with a gate width of 300u (micro) to allow balancing operation up to 100mA battery cell voltage. WP1 and WN1 are operated through WP2, WN2, WP3, and WN3 by control voltage (CTRL) so that balancing operation can be performed safely.

III. THE RESULT

In this paper, the controller interface block and the high-voltage balancing block are operated in a separate manner to achieve stable cell-unit battery balancing. By designing a controller interface block including a flip-flop, the circuit was configured so that balancing starts and ends when balancing is required at any time. When balancing is required due to the voltage difference of the battery cells, it was designed to balance at 100 mA. This balancing operation is designed to operate with a small power consumption of 180 uW with a balancing block operation with power consumption of nearly 0 uW and a controller interface block operation with 180 uW consumption. Accordingly, it was designed so that the battery hardly consumes power to perform the balancing operation function. Figure 5. shows the layout drawing of the designed balancing circuit.

ACKNOWLEDGMENT

The chip fabrication was supported by the IC Design Education Center(IDEC), Korea.

REFERENCES

[1] Sun Peiyi Bisschop, Roeland, Niu Huichang, Huang Xinyan, “A Review of Battery Fires in Electric Vehicles.” Fire Technology 56, p5, 2020.

[2] Jin Uk Yeon, Ji Whan Noh, Innyeal Oh,“Temperature-Compensated Overcharge Protection Measurement Technology,” Advances in Science, Technology and Engineering Systems Journal, Vol. 8, No. 2, pp24-29.2023.

[3] Jeong Lee, Tae-Ho Eom, Jun-Mo Kim, Min-Ho Shin, Chung-Yuen Won, “Battery lifetime prediction using the cell-balancing method of BMS,” Proceedings of KIIEE Annual Conference, pp126-126, 2016.

[4] Min-Chul Jung, Ji-Myung Kim, Dong-Hyun Tae, Dae-Seok Rho, “Operation Algorithm of Battery Management System for Li-ion Battery Considering Self-energy Balancing Phenomenon,” The transactions of The Korean Institute of Electrical Engineers, Vo. 70, No. 10, pp1443-1451, 2021

[5] Younghee Kim, Hongzhou Jin, Yoongyu Ha, Panbong Ha, Juwon Baek, “Design of a gate driver driving active balancing circuit for BMSs.” Journal of Korea Institute of Information Electronics and Communication Technology, Vol. 11, No. 6, pp732-741, 2018

RS-DeepNet: A Machine Learning Aided RSSI Fingerprinting for Precise Indoor Localization

Fawad, Arif Ullah, Iftikhar Ahmad and Wooyeol Choi*
Department of Computer Engineering, Chosun University
Gwangju, Republic of Korea
{fawad, arifullah, iftikhar, wyc}@chosun.ac.kr

***Abstract*—Intelligent recommendation applications in smart cities require the precise location of the users. The traditional global positioning system (GPS) uses satellite signals for the precise positioning of the user but is vulnerable to signal blockage in the complex indoor environment. The unforeseeable propagation losses due to multipath effects and the permittivity and permeability difference of materials in the indoor environment result in non-linear attenuation in the electromagnetic (EM) beam generated by the beacon devices. Therefore, a robust indoor localization algorithm is required to precisely localize the users in the indoor environment with severe EM blockages. In this paper, we propose a novel hybrid RS-DeepNet framework that uses received signal strength (RSS) from WiFi devices for indoor localization of users. The proposed RS-DeepNet is a deep learning architecture that utilizes multiple gated recurrent layers (GRU) and a K-nearest neighbors (KNN) classifier to estimate the precise location of the user in the indoor setup. Simulation results show that the proposed RS-DeepNet outperforms the state-of-the-art approaches and efficiently localizes the users in two different indoor scenarios with the lowest mean absolute error of 4.81 and 1.68 meters, respectively.**

***Index Terms*—Indoor localization, machine learning, RSSI fingerprinting, feature extraction, classification**

I. Introduction

Most of the internet users in the network are located in the indoor environment,therefore, network densification is considered to provide service to these users by deploying user-centric small base station deployment [1], [2]. To enable different location-based services, acquiring the accurate position information of the users in the indoor environment is critical [3]. The global navigation satellite system (GNSS) is capable of localizing the user and can achieve a sub-meter localization accuracy in some scenarios. However, it does not perform well in the indoor setup [4]. The emerging Internet-of-Things (IoT) sensor technology and the advancement in artificial intelligence (AI) have recently captured the attention of researchers to develop an intelligent and efficient indoor localization framework for smart cities. Visual-based localization (VBL) and wireless sensor-based localization (WSBL) uses computer vision and wireless sensor network (WSN) for precise indoor localization [5]. However, visual biometric-based localization of individuals through indoor CCTV cameras suffers from various photometric variations such as intra-class occlusions, illumination differences, scale variations, and orientation changes [6]. The wireless sensor-based localization is broadly classified into time-based, angle-based, and receive-power-based techniques [7]. The time-based indoor localization relies on the time of arrival (ToA), time difference of arrival (TDoA), and round trip time (RTT), while the angle-based method relies on the radio signal's angle of arrival (AoA) at a specified grid of location coordinates and the receive-power-based methods depend on the received signal strength (RSS) fingerprint and the radio propagation model of the environments. The RSS fingerprint consists of the received signal power from the anchor nodes at uniformly distributed grid locations and carries useful location information. Fingerprint includes an RSS indicator (RSSI) and is directly accessed from the application layer without requiring extra hardware or software changes. However, a huge fingerprint data transmission is needed in fingerprint-based localization, which makes it inefficient in terms of power, memory, and computational resources, especially in complex environments [8].

Machine learning (ML) techniques are capable of tackling these challenges. The ML model either utilizes the raw RSSI data or extracts statistical features from the input data. The classification frameworks utilized for indoor localization so far include the multi-class support vector machine (SVM) [9], K-nearest neighbors (KNN) [10], probabilistic decision tree, and an ensemble model. However, these models result in lower precision and higher mean absolute error (MAE). The deep neural network (DNN) comprising long-short term memory (LSTM) [11], bi-directional LSTM (BiLSTM) [12], and one-dimensional convolutional neural network (1D-CNN) [13] are used to accurately determine the indoor object's location based on the RSS from beacon nodes. ML frameworks also utilize the AoA and TDoA data for low-cost indoor positioning. Instead of raw data by the statistical handcrafted methods, a DNN generates the distinctive feature values for the raw input data and results in improved localization performance. The authors in [13] transformed the RSSI fingerprint data into 2D images using Continuous Wavelet Transforms (CWT) and applied two-dimensional (2D) DNN models to estimate indoor object coordinates. Although the 2D DNN is robust against photometric noises, the limited number of beacons and the limited difference in RSSI signal strengths within

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) under a grant (No. NRF-2021R1I1A3050535).
*Corresponding author

the fingerprint result in lower precision compared to 1D CNN models.

Unlike isolated feature extraction and classification models, fused frameworks are utilized, including multiple ML algorithms to combine and localize indoor objects. The Unsupervised Fusion of Extended Candidate Location Set (UFL-ECLS) [14] is a fused model consisting of multiple trained classifiers. The UFL-ECLS performance is highly dependent on the individual classifier accuracy, and any single classifier in-accurate results could lead to degradation of the overall localization performance. To address this issue, the authors in [15] propose a SmartLoc framework that utilizes multiple ML algorithms along with alignment probabilistic procedures to enhance offline training processes. Although the SmartLoc framework offers greater precision by employing multiple ML frameworks and probabilistic alignment methods, it suffers from the increased computational cost issue, which makes it unsuitable for real-time applications.

Motivated by the aforementioned works, the goal of this paper is to propose an efficient ML framework for precise indoor localization which utilizes RSSI fingerprint data to estimate the location coordinates of the users in the complex indoor environment. The existing frameworks for indoor localization need more accuracy, which is attributed either to the vulnerability of the estimation model or the limitations of training samples for the model. To address this issue and achieve high accuracy while reducing computational complexity, this paper proposes RS-DeepNet, which leverages the advantages of both DNN features and clustering by first extracting features from the RSSI data and then using these features for clustering-based classification.

The main contributions of the paper are outlined as follows.

- We propose an RS-DeepNet which is a DNN architecture for indoor localization. The proposed framework is based on feature extraction from the input RSSI data followed by an isolated classification of location information from the extracted features. The feature extraction is performed using multiple cascaded GRU blocks, each comprising a GRU layer, normalization layer, and dropout layer. The isolated KNN classifier is used to classify deep features collected through the cascaded GRU-DNN to estimate the user's coordinates in order to achieve highly precise localization.
- Considering two different indoor scenarios, we demonstrate that the proposed RS-DeepNet precisely predicts the user location information from the RSSI fingerprint data. We further show that the proposed learning framework outperforms the other existing techniques by achieving the lowest mean absolute error (MAE).

The rest of the paper is organized as follows: Section II presents the system model, including the procedure for generating fingerprint data. Section III presents the methodology of the proposed feature extraction and categorization. Section IV provides the simulation results and comparison of the proposed framework with the benchmark. Finally, Section V summarize the key findings and concludes the paper.

II. System Model

The considered system model has K' number of WiFi access points (APs) with the known location distributed over a square area of dimension $S \times S$ and the location coordinate of the ith reference user point (UP) denoted by $\mathcal{U}_i = (x_i, y_i)\ \forall i = \{1, 2, ..N\}$, where N is the number of users. The RSSI-fingerprint-based indoor localization framework consists of an offline training stage and an online testing process. During the offline training stage, the RSSI fingerprint is created by dividing the serving area into a square grid of size $X \times Y$ with grid markers across the x and y-axis, respectively. In this study, we consider two indoor environment scenarios as shown in Fig. 1 (a) and (b), respectively. For scenario 1: $S = 30$m with $K' = 4$ distributed APs while for scenario 2: $S = 50$m with $K' = 5$ distributed APs. The UPs are distributed across the grid points in such a way that each grid position belongs to one UP.

The RSSI fingerprint is collected and stored in the database for the particular scenario. The RSSI fingerprint and the location information stored are given by

$$\begin{aligned}\Omega &= \begin{bmatrix}\mathcal{U} & \mathbf{\Psi}\end{bmatrix} \\ &= \begin{bmatrix}(x_i, y_i) & (\Psi_{i,1}, \Psi_{i,2}, ...\Psi_{i,K}) \\ : & : \\ (x_N, y_N) & (\Psi_{i,1}, \Psi_{i,2}, ...\Psi_{N,K})\end{bmatrix},\end{aligned} \tag{1}$$

where the $\Psi_{n,k}$ refers to RSSI at the nth location across the grid with signal strength for the kth AP.

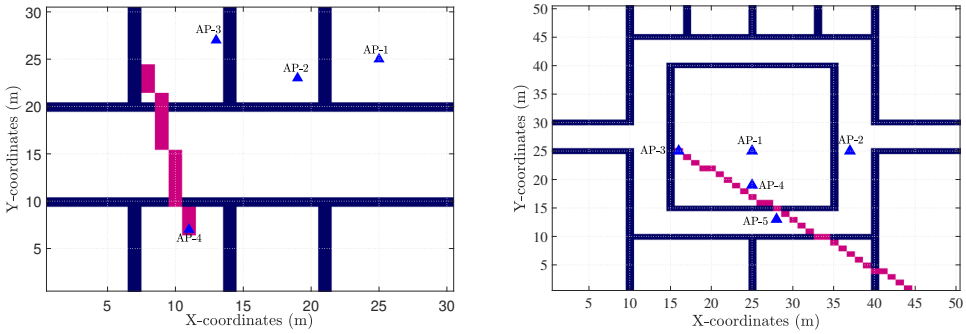
The signal propagating between the kth AP and the nth UP follows the Log-Normal path loss model and is highly dependent on obstacles within the line-of-sight (LOS) between APs and UPs. In the simulated environment, we consider the random shadowing effect and noise due to the presence of randomly moving objects within the environments [16]. As the RSS randomly fluctuates at each time instant, therefore, the average RSS received from the k_{th} AP at location (x, y) can be expressed as

$$\Psi_{n,k} = \frac{\sum_{t=1}^{\mathcal{T}} s_{n,k}(t)}{\mathcal{T}}, \tag{2}$$

where $s_{n,k}(t_j)$ denotes the RSSI received from the kth AP at the nth UP at time instant t and $\Psi_{n,k}^{t}$ denotes the RSSI values set for a total of $\mathcal{T}$ time samples.

III. Proposed DNN architecture for RS-DeepNet

The proposed RS-DeepNet model extracts the GRU DNN features from the RSSI fingerprint and classifies them into geographic location coordinates of the indoor site. The RSSI fingerprint data, along with the respective geographic location coordinates of the grid, is used as a training dataset for the KNN classifier. The KNN classifier estimates the nearest locations of the test RSSI set, and the mean location is considered as the position of the test reference UP. The



(a) (b)

Fig. 1: Floor map of RSSI-Fingerprint environment with dark blue represents wall structures while light blue triangle denotes the AP position, and the pink color denotes User trajectory (a) Scenario 1: square grid with $S = 30$m and $K' = 4$ (b) Scenario 2: square grid with $S = 50$m and $K' = 5$.

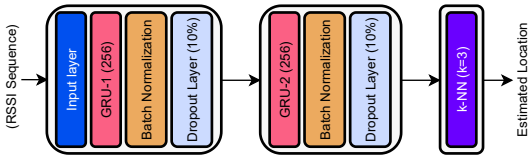


Fig. 2: Mean Square Error of the proposed RS-DeepNet in comparison with the other DNN models.

location coordinates of APs and the RSS fingerprint values of the kth AP at nth reference UPs are, respectively, given as

$$\mathcal{I} = \mathbf{\Psi} \begin{bmatrix} (\Psi_{1,1}, \Psi_{1,2}, .., \Psi_{1,K}) \\ (\Psi_{2,1}, \Psi_{2,2}, .., \Psi_{2,K}) \\ \vdots \\ (\Psi_{N,1}, \Psi_{N,2}, .., \Psi_{N,K}) \end{bmatrix}, \tag{3}$$

$$\mathcal{L} = \mathcal{U} = \begin{bmatrix} \mathcal{U}_1 = (x_1, y_1) \\ \vdots \\ \mathcal{U}_N = (x_N, y_N) \end{bmatrix}, \tag{4}$$

where $\mathcal{I}$ represents the Input data and $\mathcal{L}$ is the corresponding labels matrix used during training of the proposed RS-DeepNet architecture.

The proposed RS-DeepNet comprises two GRU blocks followed by an isolated KNN classifier, as shown in Fig. 2. Each GRU block consists of a GRU layer with 256 cell units, a normalization layer, and a dropout layer with a probability of 0.1. The model is trained with $[\mathcal{I}\ \mathcal{L}]$ where $\mathcal{I}$ is fed at the input of RS-DeepNet. The final output of the proposed RS-DeepNet can be mathematically expressed as

$$\mathcal{O} = \mathcal{C}_{knn}\left(\mathcal{G}_2\left(\mathcal{G}_1\left(\mathcal{I}\right)\right)\right), \tag{5}$$

where $\mathcal{G}_1(\cdot)$, $\mathcal{G}_2(\cdot)$, and $\mathcal{C}_{knn}(\cdot)$ denote the first GRU block, second GRU block, and KNN classifier, respectively. A fixed number of data samples, $\mathcal{T}$ are generated for each scenario and is set to 250. During training, 70% and 20% of the total data are used for training and validation, respectively, while for the KNN classifier, we set $K = 3$. The model is trained for 1000 epochs with a batch size of 64 and a learning rate of 0.005, and an ADAM optimizer is used. The performance of the proposed RS-DeepNet is evaluated in terms of mean absolute error (MAE). Let us denote the actual true locations and estimated locations of the UPs by $\mathcal{U}$ and $\tilde{\mathcal{U}}$, respectively. Then, the MAE can be written as

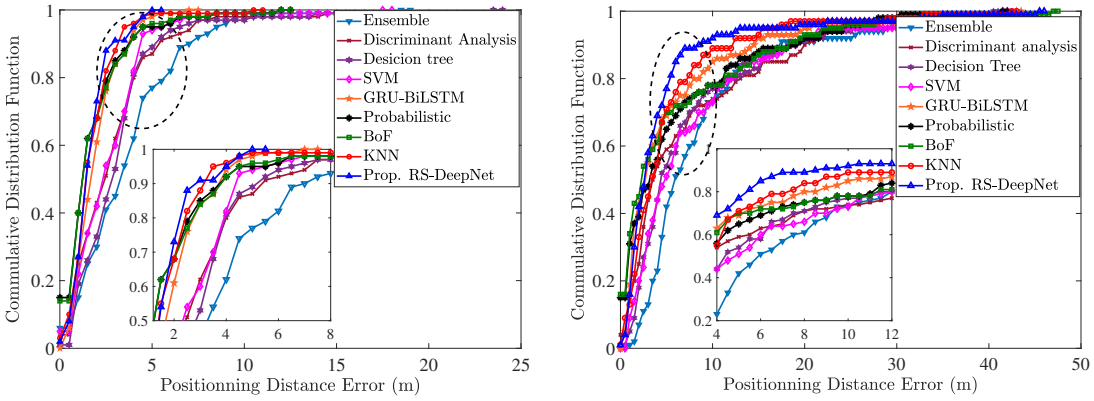
$$\text{MAE} = \frac{1}{\xi_{test}} \sum_{z=1}^{\xi_{test}} |\mathcal{U}_z - \tilde{\mathcal{U}}_z|, \tag{6}$$

where ξ denotes the total number of test data points considered for evaluation.

IV. Simulation Results

The performance of the proposed RS-DeepNet model is evaluated on two different simulated virtual environments, as shown in Fig. 1. Both environments have different indoor structures and dimensions. Random environmental noise, path loss, shadowing, and multi-path effects are considered in both environments. We evaluate the performance of the proposed RS-DeepNet in terms of the cumulative distribution function (CDF) of the positioning distance error and the MAE for the tested dataset.

Fig. 3 presents the CDF performance of the proposed RS-DeepNet in comparison with other learning techniques for two indoor scenarios. Fig. 3 (a) and (b) shows that compared to the traditional models, the proposed RS-DeepNet provide superior performance at lower positioning distance errors for indoor scenario 1 and 2, respectively. The comparison with the other classification models that utilize raw RSSI fingerprints, such as SVM, KNN, decision tree, discriminant analysis, ensemble, and probabilistic models, is also provided. Furthermore, the Bag-of-Features (BoF) method and the stacked GRU-BiLSTM model, which extracts robust features from the raw RSSI, are also considered for a fair comparison. Fig. 3 depicts that the proposed RS-DeepNet achieves superior performance compared to all benchmark methods. The second and third best performing indoor localization algorithms are KNN and GRU-BiLSTM, respectively.



(a) (b)

Fig. 3: Cumulative distribution function of the positioning distance error for (a) Scenario 1: 30m x 30m geographic grid area (b) Scenario 2: 50m x 50m geographic grid area.

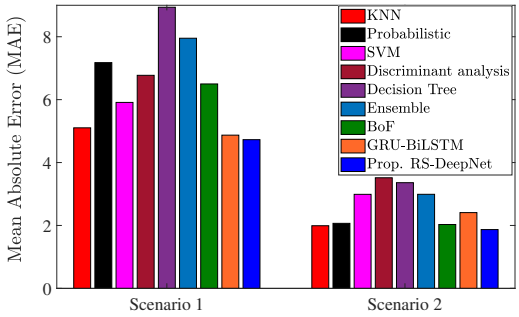


Fig. 4: Mean Square Error of the proposed RS-DeepNet in comparison with the other DNN models.

Fig. 4 presents the MAE of the true and the estimated location coordinates by the proposed RS-DeepNet. Fig. 4 shows that the proposed RS-DeepNet achieves the lowest MAE of 4.81 and 1.68 compared to the other reported benchmarks in the considered scenario 1 and scenario 2, respectively.

V. Conclusions

In this paper, we proposed a novel RS-DeepNet framework for indoor localization, which is a DNN architecture that utilizes two GRU blocks for features extraction and an isolated KNN for classification. The proposed RS-DeepNet is tested for two different indoor residential apartment scenarios. The performance of the proposed RS-DeepNet is compared with the other DNN models. It is observed from simulation results that for both indoor scenarios, the proposed model precisely estimates the location of UPs during the online testing and provides superior performance compared to the other DNN models with reduced MAE of 4.81 m and 1.68 m in scenario 1 and scenario 2, respectively.

References

[1] Z. H. Abbas, A. Ullah, G. Abbas, F. Muhammad, and F. Y. Li, "Outage probability analysis of user-centric SBS-based HCNets under hybrid Rician/Rayleigh fading," *IEEE Communications Letters*, vol. 24, no. 2, pp. 297–301, 2019.

[2] A. Ullah, Z. H. Abbas, G. Abbas, F. Muhammad, and J.-M. Kang, "Hybrid millimeter wave heterogeneous networks with spatially correlated user equipment," *Digital Communications and Networks*, 2022.

[3] J. Wang, X. Zhang, Q. Gao, H. Yue, and H. Wang, "Device-free wireless localization and activity recognition: A deep learning approach," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 7, pp. 6258–6267, 2016.

[4] X. Guo, L. Li, N. Ansari, and B. Liao, "Knowledge aided adaptive localization via global fusion profile," *IEEE Internet of Things Journal*, vol. 5, no. 2, pp. 1081–1089, 2017.

[5] S. B. Altaf Khattak, M. M. Nasralla, M. A. Esmail, H. Mostafa, and M. Jia, "WLAN RSS-based fingerprinting for indoor localization: A machine learning inspired bag-of-features approach," *Sensors*, vol. 22, no. 14, p. 5236, 2022.

[6] B. Yang, J. Li, and H. Zhang, "Resilient indoor localization system based on UWB and visual-inertial sensors for complex environments," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1–14, 2021.

[7] S. M. Asaad and H. S. Maghdid, "A comprehensive review of indoor/outdoor localization solutions in IoT era: Research challenges and future perspectives," *Computer Networks*, p. 109041, 2022.

[8] J. Yan, G. Qi, B. Kang, X. Wu, and H. Liu, "Extreme learning machine for accurate indoor localization using RSSI fingerprints in multifloor environments," *IEEE Internet of Things Journal*, vol. 8, no. 19, pp. 14 623–14 637, 2021.

[9] A. Chriki, H. Touati, and H. Snoussi, "SVM-based indoor localization in wireless sensor networks," in *2017 13th international wireless communications and mobile computing conference (IWCMC)*. IEEE, 2017, pp. 1144–1149.

[10] I. T. Haque and C. Assi, "Profiling-based indoor localization schemes," *IEEE Systems Journal*, vol. 9, no. 1, pp. 76–85, 2013.

[11] Z. Chen, H. Zou, J. Yang, H. Jiang, and L. Xie, "WiFi fingerprinting indoor localization using local feature-based deep LSTM," *IEEE Systems Journal*, vol. 14, no. 2, pp. 3001–3010, 2019.

[12] N. Liu, T. He, S. He, and Q. Niu, "Indoor localization with adaptive signal sequence representations," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 11 678–11 694, 2021.

[13] P. Ssekidde, O. Steven Eyobu, D. S. Han, and T. J. Oyana, "Augmented CWT features for deep learning-based indoor localization using wifi rssi data," *Applied Sciences*, vol. 11, no. 4, p. 1806, 2021.

[14] X. Guo, S. Zhu, L. Li, F. Hu, and N. Ansari, "Accurate WiFi localization by unsupervised fusion of extended candidate location set," *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2476–2485, 2018.

[15] L. Li, X. Guo, and N. Ansari, "Smartloc: Smart wireless indoor localization empowered by machine learning," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6883–6893, 2019.

[16] M. Jia, S. B. A. Khattak, Q. Guo, X. Gu, and Y. Lin, "Access point optimization for reliable indoor localization systems," *IEEE Transactions on Reliability*, vol. 69, no. 4, pp. 1424–1436, 2019.

Implementation of a Household Energy Management System Based on Deep Learning

Jeongeun Park
School of Digital Media Engineering
Tongmyong University
Busan, Korea
pakrose0108@nate.com

Jeonghyun Shin
School of Digital Media Engineering
Tongmyong University
Busan, Korea
sjhs1203@gmail.com

Soyeon Park
School of Digital Media Engineering
Tongmyong University
Busan, Korea
wkdtks3@hanmail.net

Eunjeong Park
School of Digital Media Engineering
Tongmyong University
Busan, Korea
dmswjdwjddjs@naver.com

Hyuntae Cho
Dept. of Digital Content
Tongmyong University
Busan, Korea
marine@tu.ac.kr

***Abstract*—As interest in energy conservation has recently increased, efforts to reduce energy consumption in various environments are also increasing. In particular, households consume a huge amount of power due to various home appliances. Nevertheless, residents cannot confirm how much energy they are using. Therefore, this paper proposes a household energy management system using deep learning. The proposed system collects power consumption data of home appliances by using smart plugs. Then, the application monitors the power consumption of each home appliance, and predicts the power leakage and anomality of the appliance based on deep learning. The system provides meaningful information about electricity to customers or managers.**

Keywords—power management, energy management, smart home, deep learning, smart plug

I. INTRODUCTION

As the world develops economically, power consumption is also increasing, and in some countries, things such as shutdown are happening due to power shortages. In addition, excessive power use and power leakage in the home cause expensive management cost. According to the Enerdata [1] 2022 report, global power consumption reached about 25,000 TWh in 2021, of which 12,165 TWh was consumed in Asia. By country, China, the United States, India, Russia, and Japan used a lot of electricity. According to KOSIS [2] in 2022, about 47 TWh of electricity was used in Korea, of which 48% was used in the manufacturing industry, 13.7 TWh was used in the service industry, and 6.3 TWh was used for household use. Efficient use, management, and consumption of electricity is a very important issue for individuals, domestically, and nationally. This paper proposes a household energy management system (EMS) that can efficiently manage power and monitor usage, and detect abnormal situations in home appliances at an early stage by adopting deep learning technology.

This research was supported by the MISP (Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP (Institute for Information & communications Technology Promotion) (2018-0-018740301001).

II. REALTED WORK

Zhou [3] introduced a conceptual overview on the architecture and functional modules of smart household energy management system (HEMS), and reviewed the HEMS infrastructures and home appliances in smart houses. Then, they investigated various home appliance scheduling strategies to reduce the residential electricity costs and improve the energy efficiency of power generation utilities. Adika [4] presented an energy management system that autonomously execute all tasks of appliances without the prompting of the customers. In the paper, authors proposed a demand side energy management for a grid connected household using locally generated photovoltaic energy. To ensure efficient household energy management, smart scheduling of electrical appliances has also been presented.

III. HOUSEHOLD ENERGY MANAGEMENT SYSTEM USING DEEP LEARNING

Figure 1 illustrates a conceptual overview of a household energy management system based on deep learning.

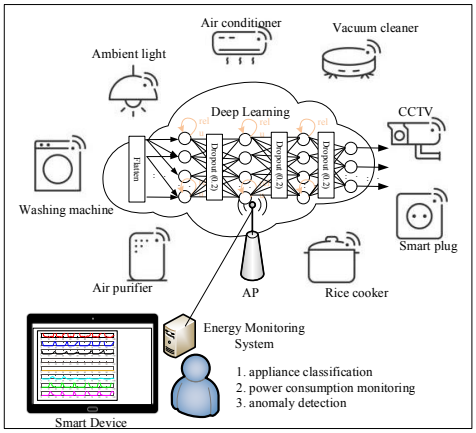


Fig. 1. Concpetual overview of the household energy management system.

In this paper, in order to measure the amount of power consumed at home, appliances receive the necessary power through smart plugs. Commercial off-the-shelf (COTS) smart plugs were used to power appliances and measure power consumption. [5]. The smart plug can continuously measure the power consumption of the home appliance inserted into the plug and transmit it wirelessly to a server or directly to a smart device using Bluetooth. The power consumption is continuously monitored and anomalies of each home appliance are detected by the application using on-device deep learning. In order to detect an abnormal state, it is also necessary to classify the type of home appliance using deep learning, and to monitor the abnormal state based on this.

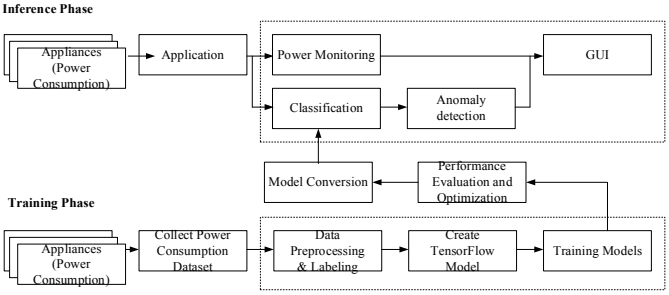


Fig. 2. Algorithm flow for Classification, Power monitoring, and anomalous detection.

Figure 2 shows the entire process of monitoring power usage and detecting anomalies. The application of a smart device periodically collects power consumption data from home appliances. And, it goes through a pre-processing process for the data and selectively displays hourly, daily, weekly, and monthly data on the GUI of the application. Then, deep learning technology is used to detect abnormal conditions of home appliances. First, the received power consumption data is preprocessed and normalized, and then input into the trained deep learning model. To diagnose the failure of each home appliance, the application first classifies the appliance based on the data appliances consumed. For home appliance classification, we use Kaggle's household appliances power consumption dataset [6] in which includes 3D printer, air conditioner, air purifier, boiler, coffee machine, computer, drier, fan, freezer, fridge, Internet router, laptop, etc.. We trained the dataset with deep learning model and evaluated its performance. The trained model is used for real-time inference of the power consumption data later. Accordingly, the type of home appliance is determined, and an abnormal state of each appliance is detected. To detect anomalily, only normal data of each home appliance is used and inference is made again through deep learning. In normal state, only power consumption results are output and in abnormal state, warnings are delivered along with past data history.

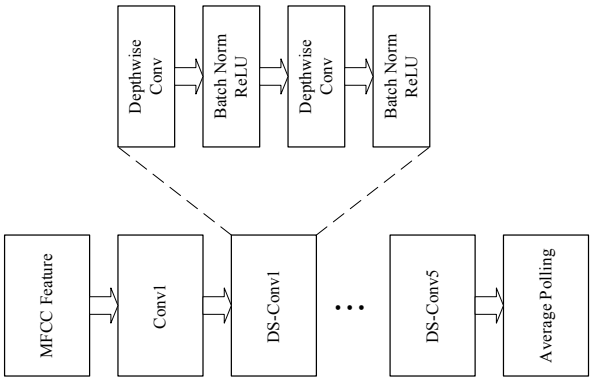


Fig. 3. An example of MFCC and depthwise separable CNN model to detect anomalous state.

In addition, in this work, in order to compare and analyze the performance of deep learning, various deep learning models can be designed and analyzed to find the optimal model. Models such as dense neural network (DNN), convolution neural network (CNN), and bidirectional long short-term memory (B-LSTM) can be used for classification, and autoencoder and depth-wise separable CNN (DS-CNN) can be for detecting anomalies. Figure 3 shows one example of many deep learning models. After performance evaluation, the optimal model will be converted to a tflite model via TensorFlow lite converter, which can then be loaded and used on Android or iOS.

IV. Conclusion

In this Work-In-Progress paper, we discussed with the implementation of an energy management system to save unnecessary and wasted power in the home and to diagnose home appliance failures at an early stage. An application running on a smart phone or tablet PC monitors the power consumption of home appliances and provides graphical information about the power consumption to the user. In addition, it uses two-step machine learning technology to detect abnormal conditions of each home appliance. First, the type of each home appliance can be recognized using deep learning technology for classification and secondary deep learning technology can be used to detect abnormal conditions. It is believed that such a home energy management system can solve problems such as fire and power leakage due to failure of home appliances as well as energy saving.

References

[1] Enerdata, “Report – 2022 Edtion: Annual benchmarks and Long-term impact,” https://yearbook.enerdata.net/total-energy/world-consumption-statistics.html, available on Sep. 01, 2022.

[2] Korean Statistical Information Service, “Result of Power Supply by year (2022), KOSIS, 2022.

[3] B. Zhou, W. Li, K. W. Chan, Y. Cao, Y. Kuang, X. Liu, X. Wang, “Smart home energy management system: concept, configurations, and scheduling strategies,”Renewable and Sustainable Energy Reviews, 61, 30-40, 2016.

[4] C. O. Adika, L. Whag, “Autonomous Appliance Scheduling for Household energy managemt,” IEEE Transactions on Smart Grid, Vol. 5, Iss. 2, pp. 673-382, Mar. 2014.

[5] https://www.tp-link.com/kr/home-networking/smart-plug/tapo-p100/, available on March 11 2023.

[6] https://www.kaggle.com/datasets/ecoco2/household-appliances-power-consumption, available on March 11 2023.

Improvements of Patient Waiting Time using BLE-based Real-time Location System in Emergency and Trauma Department

Mohd Shafarudin Osman
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
shafarudin.osman@gmail.com

Azizul Azizan
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
azizulazizan@utm.my

Siti Sophiayati Yuhaniz
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
sophia@utm.my

Suriani Mohd Sam
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
suriani.kl@utm.my

Noraimi Shafie
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
noraimi.kl@utm.my

Yusnaidi Md Yusof
Advanced Informatic Department, Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
yusnaidi.kl@utm.my

***Abstract*—The delay in public hospitals delivering services to patients waiting in Emergency and Trauma department (ETD) are severe in the Green Zone. Green Zone patients experience long waiting time, resulting in overcrowding and affecting the quality of services provided in the ETD. Hospitals have already adopted Hospital Information System (HIS) to improve clinical workflow however patient waiting time are still captured manually within the system, lacking real-time patient location monitoring and automated service tracking. This paper proposes a Real-time Location System (RTLS) adopting low-cost Bluetooth Low Energy (BLE) based nodes that track and automates patient tracking based on proximity detection. The results obtained from the semi-integrated HIS and RTLS systems shows better accuracy and improvements of both Arrival to Consult (ATC) waiting time and Length of Stay (LOS) compared to non-RTLS based tracking.**

Keywords— Real-time Location System, Patient Waiting Time, Bluetooth Low Energy (BLE)

I. INTRODUCATION

Emergency and Trauma Department (ETD) provides 24 hours treatment for various emergency cases and patients are typically classified into 3 zones depending on severity of the cases i.e. Red Zone - critical patients, Yellow Zone - semi-critical patients and Green Zone - non-critical patients. The large number of patients, results in patient congestion and overcrowding affecting the quality of services in ETD especially for non-critical patients.

To improve the overcrowding of patients and enhance the service, the Malaysian Ministry of Health (MOH) has implemented interventions such as LEAN healthcare that improves the process workflow and digital records of patients within the Hospital Information System (HIS) or Electronic Medical Records (EMRs) [1-2]. The HIS stores patient data such as treatments, clinical results, and billing, however there is no real-time tracking on service time implemented to improve the efficiency of the clinical services delivered to patients.

Although HIS can tackle long patient waiting time and improve healthcare service delivery, it is unable to automatically track patients in real-time, where users need to click multiple buttons in multiple pages to capture and notify patient arrival time at certain counters or rooms.

Such workflow causes inaccurate data capture, requiring certain amount of time to be completed. And can lead to data error such as acknowledging the wrong patient in the wrong place. By introducing real-time data capture capability in HIS, patients can be traced and notified in real time reducing waiting time [3].

This paper examines the impact of a low-cost Real-time Location System (RTLS) using proximity-based detection that eliminates manual patient tracking to improve patient waiting time in ETD Green Zone at public hospitals.

II. RTLS IN HEALTHCARE SERVICES

A. RTLS Implementations

RTLS is a technology where physical tags, sensors and digital middleware captures data, process it to visualize the location of non-stationary objects, people or animals in real-time [4]. It has been used in logistics and transportation, manufacturing, agriculture and healthcare sectors to enable traceability of humans and assets.

Although RTLS technology has been implemented in various sectors, for the healthcare sector there has been no real-time implantation of such system especially in the ETD Green Zone to detect and reduce patient waiting time. Several notable research implements RTLS to track medical devices [5,6], staff hand washing compliance and safety of dementia patients wandering around hospital compounds [4].

Studies have also been conducted to analyse the impact on patient waiting time using RTLS in non-emergency out-patient clinic by [7] found that average patient waiting time was reduced by 31 minutes when using RTLS. Meanwhile patients who were prioritized, saw an additional reduction of

their waiting time by an average of 9 minutes. [8] found that RTLS reduced patient location time by 73% in the physical room and reduced the number of user-system interactions (mouse clicks) by 75%.

B. *Wireless protocol for RTLS*

Table I lists the RTLS wireless protocols ranging from large scale outdoor based Global Positioning System (GPS), medium range Wireless Fidelity (Wi-Fi), to short proximity tracking protocols such as Radio Frequency Identification (RFID), Ultra-Wideband (UWB) and Bluetooth Low Energy (BLE).

TABLE I. IMPORTANT CHARACTERISTICS OF RTLS AND RELATED TECHNOLOGIES

Technology	Location Methods	Algorithm	Accuracy	Scalability| Complexity | Cost
GPS	Trilateration	TOA	6-10m	L|H|H[a]
UWB	Trilateration Angulation	RSS TOA TDOA AOA	1cm-1m	L|L|H
IR	Proximity Trilateration	Cell-ID TOA	1cm-5m	L|L|M[a]
Wi-Fi	Proximity Trilateration Angulation Scene Analysis	Cell-ID TOA TDOA AOA RSS	10-100m	H|L|L[a]
RFID	Proximity Trilateration Scene Analysis	Cell-ID RSSI	1-5m	M|L|L[a]
Bluetooth/ BLE	Proximity Trilateration Scene Analysis	Cell-ID RSS TOA	1-5m	H|L|L[a]

[a] Note: H=High, M=Medium, L=Low

From Table I each wireless protocol is examined in terms of scalability, complexity, and costs. Although GPS, UWB and IR technologies have been used in other environments such as construction, industrial and logistic, they are low in scalability and have a high cost when implemented in the ETD environment. As for the most common indoor location tracking protocol i.e., the Wi-Fi, it suffers from low accuracy for proximity tracking applications.

The RFID-based RTLS meanwhile suffers from a higher cost of implementation where active RFID is required to enable accurate proximity detection in comparison to BLE-based RFID. BLE based systems have the most cost-effective protocol to be used to track ETD Green Zone patients as it is highly scalable, low complexity and low-cost, where BLE is pervasive in small passive tags with low power operation.

III. BLE BASED RTLS IMPLEMENTATION

A. *Proximity and Received Signal Strength (RSSI) based BLE*

The BLE technology is an IEEE 802.15 wireless standard for wireless personal area network (WPAN), enabling a shorter range of detection compared to common Wi-Fi based wireless local area network (WLAN). The current BLE technology consumes lower power where it can operate for few months without battery charging [9].

The proximity-based detection method is implemented in this research as it is the most power efficient method to enable detection of patients in the crowded ETD spaces in the hospital. It is easier to be implemented compared to Wi-Fi based fingerprinting solutions and does not require advanced Time of Arrival (TOA), Time Difference of Arrival (TDOA), Angle of Arrival (AOA) calculations [10].

The setup for the BLE-based RTLS in the ETD is shown in Figure 1 below.

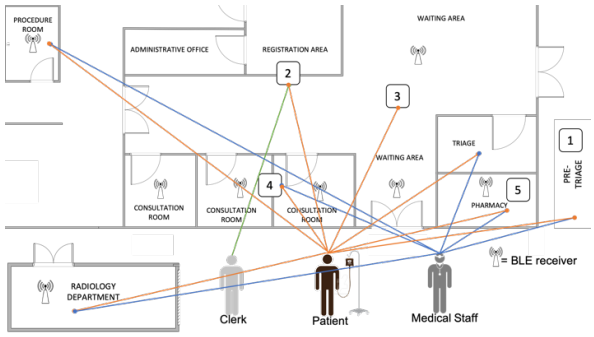


Fig. 1. BLE-based RTLS.

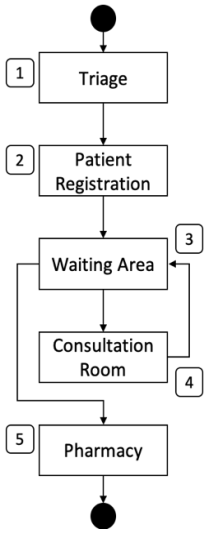


Fig. 2. ETD workflow

The RTLS infrastructure setup is based on ETD workflow consists of 5 different stages: 1) Triage, 2) Patient Registration Area, 3) Waiting Area, 4) Consultation Room, and 5) Pharmacy as shown in Figure 2.

B. *Patient waiting time performance measurements*

The arrival to consult (ATC) and length of stay (LOS) are two main parameters to measure the performance of the RTLS.

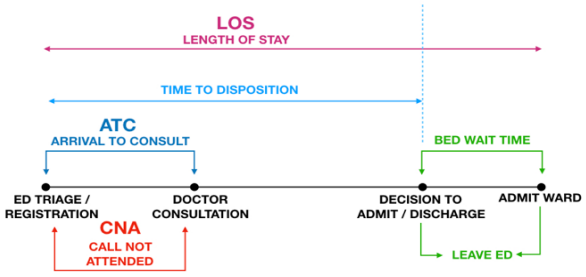


Fig. 3. Relationship between patient flow and waiting time in ETD

Figure 1 shows that as patients arrive at ETD, they are screened at the triage and later registered. They wait for doctor consultation; the time between this stage is called Arrival to Consult (ATC) waiting time [11]. The Length of Stay (LOS) meanwhile measure the total servicing time and a prolonged ATC contributes to overcrowding resulting in low patient satisfaction.

The data was collected for three months where 1174 patient were included in the study. There were 587 data for RTLS scenario, and 587 data for non-RTLS scenario.

IV. RESULTS AND ANALYSIS

A. *Findings*

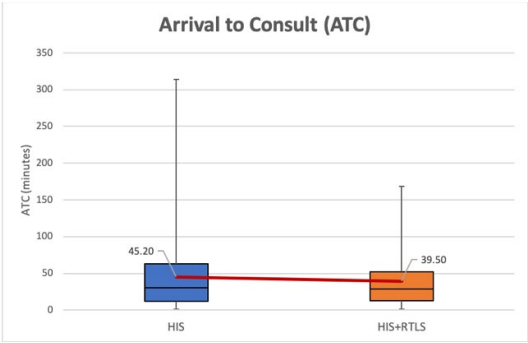


Fig. 4. ATC time Mean Comparison for non-RTLS (HIS) and RTLS-based system (HIS+RTLS)

The ATC time mean has dropped from the mean of *45*.20 minutes to 39.50 minutes whereas the standard deviation reduces from 45.17 minutes 35.77 minutes when HIS+RTLS is introduced in the ETD workflow. The data obtained in Figure 4 showed that the RTLS has reduced the measured ATC, and the data spread is more distributed within the mean compared to non-RTLS.

Fig. 5. LOS Mean Comparison for non-RTLS (HIS) and RTLS-based system (HIS+RTLS)

The LOS mean has dropped from 3067.59 minutes to 152 minutes with a reduced standard deviation of 110.74 minutes from 1127.48 minutes for HIS+RTLS based ETD workflow as illustrated in Figure 5. The reduction is significant due to the inaccuracy and time lapses of manual time updates in an HIS without RTLS within the current workflow.

Figure 6 and figure 7 shows the time distribution of the ATC waiting time. The waiting time has a as condensed distribution for HIS+RTLS as compared to the non-RTLS system. The histograms show a representation of the results in Figure 4 where the spread for the HIS only system is around 200 ms while for HIS+RTLS is around 150 ms.

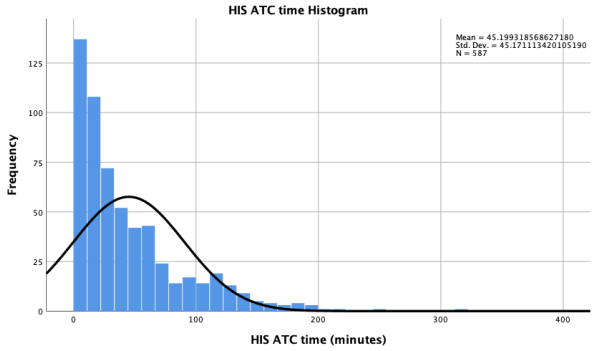


Fig. 6. ATC time Histogram for non-RTLS (HIS)

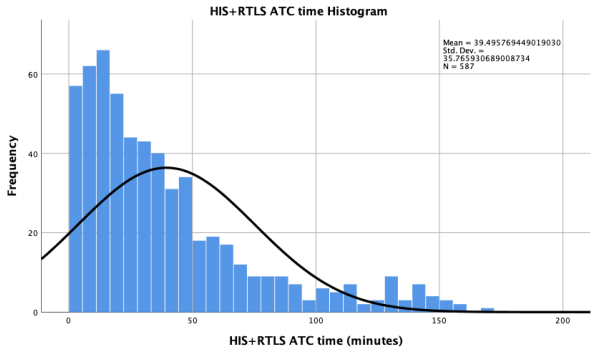


Fig. 7. ATC time Histogram for RTLS (HIS+RTLS)

V. CONCLUSION

The implementation of BLE-based RTLS patient tracking to monitor patient flow has shown substantial improvements in ETD operation. The system enables real-time patient ATC and LOS monitoring to improve overcrowding of the ETD.

ACKNOWLEDGMENT

We expressed our appreciation to the staff in Emergency and Trauma Department, Hospital Putrajaya, Malaysia for hosting the RTLS-HIS prototype that enables data collection for this research.

REFERENCES

[1] A. Ismail, A. T. Jamil, A. F. A Rahman, J. M. Abu Bakar, N. Mohd Saad, and H. Saadi, ―The implementation of Hospital Information System (HIS) in tertiary hospitals in Malaysia,‖ Malaysian J. Public Heal. Med. 2010, 10(2), 16–24, 2010.

[2] N. Azliza, N. Ariffin, A. M. Yunus, and Z. C. Embi, ―Improving Electronic Medical Records (EMRs) Practices through a Clinical Microsystem in the Malaysian Government Hospitals,‖ Commun. IBIMA, 5, 50–64, 2008

[3] Ewing, A., Rogus, J., Chintagunta, P., Kraus, L., Sabol, M., & Kang, H. (2017, April). A systems approach to improving patient flow at UVA Cancer Center using Real-Time Locating System. In 2017 Systems and Information Engineering Design Symposium (SIEDS) (pp. 259-264). IEEE.

[4] Oude Weernink, C. E., Felix, E., Verkuijlen, P. J. E. M., Dierick-van Daele, A. T. M., Kazak, J. K., & van Hoof, J. (2018). Real-time location systems in nursing homes: state of the art and future applications. Journal of Enabling Technologies, 12(2), 45-56.

[5] Ramirez, M. R. (2018). A Business Case for Using Real Time Locating Systems for Medical Equipment. *CMBES Proceedings*, *41*.

[6] Youn, J. H., Ali, H., Sharif, H., Deogun, J., Uher, J., & Hinrichs, S. H. (2007, October). WLAN-based real-time asset tracking system in healthcare environments. In *Third IEEE International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob 2007)* (pp. 71-71). IEEE.

[7] Stübig, T., Suero, E., Zeckey, C., Min, W., Janzen, L., Citak, M., ... & Gaulke, R. (2013). Improvement in the workflow efficiency of treating non-emergency outpatients by using a WLAN-based real-time location system in a level I trauma center. Journal of the American Medical Informatics Association, 20(6), 1132-1136.

[8] King, K., Quarles, J., Ravi, V., Chowdhury, T. I., Friday, D., Sisson, C., & Feng, Y. (2018). The Impact of a Location-Sensing Electronic Health Record on Clinician Efficiency and Accuracy: A Pilot Simulation Study. Applied clinical informatics, 9(04), 841-848.

[9] Omre, A. H., & Keeping, S. (2010). Bluetooth low energy: wireless connectivity for medical monitoring. *Journal of diabetes science and technology*, *4*(2), 457-463.

[10] Khudhair, A. A., Jabbar, S. Q., Sulttan, M. Q., & Wang, D. (2016). Wireless indoor localization systems and techniques: survey and comparative study. *Indonesian Journal of Electrical Engineering and Computer Science*, *3*(2), 392-409.

[11] Public Service Delivery Transformation (PSDT), Lean Agile for 16 hospitals across Malaysia (state and major hospitals) – Emergency Department and Medical Ward, Jan 2015-Sept 2015.

Estimation and Correction of Document-Table Skew by Deep-Learning-based Text Detection

1st Vo Sy Hung
Hanoi University of Science and Technology
Hanoi, Vietnam

2nd Nguyen Sy Duy
Hanoi University of Science and Technology
Hanoi, Vietnam

3rd Vo Le Cuong
Hanoi University of Science and Technology
Hanoi, Vietnam
cuong.vole@hust.edu.vn

4th Doan Tien Dat
Hanoi University of Science and Technology
Hanoi, Vietnam

5th Do Doan Khue
Hanoi University of Science and Technology
Hanoi, Vietnam

***Abstract*— Skew estimation and correction play a crucial role in detecting and recognizing text from scanned-document images. In many documents such as financial reports, social insurance documents, etc., tables occupy mainly its contents. In this paper, we propose a skew estimation and correction method for scanned tables. As the tables have many different structures with/without bounding lines, the method is based on information of detected bounding boxes of words inside each table by deep learning techniques. The approach is divided into 3 steps. Firstly, 4-sided polygons of words are located by using deep-learning-based text detection methods. Then, the ten longest lines of ten 4-sided polygons are used to create angles with the image's horizontal line. Noisy angles, which have variances higher than a threshold compared to average one, are removed. Then, a skew angle of each table is created by averaging the remaining ones which are not noise. To solve the challenge of up-side-down table after rotation, it is based on a certainty of text recognition using another PP-OCR model. Our self-developed method has demonstrated the Average Error Deviation of 0.21° in detecting the angle of rotation for skew table ranging from –175° to 180°.**

Keywords—correction, OCR, skew detection, table, text detection

I. INTRODUCTION

Latterly, document images have witnessed an exceptional expansion in terms of quantity. Consequently, Document Image Processing (DIP) system becomes progressively crucial [1]. One of the most common techniques for document digitalization is scanning [2]. The document scanning process produces skewed document images which will have negative effects on the performance of subsequent phases such as Optical Character Recognition (OCR) [3], Table Extraction (TE) [4]. For years, researchers have been conducting a variety of estimation-and-correction methods for document image skew [5] adopting different approaches namely Projection Profile (PP) analysis, Hough transform, Fourier transform, etc.

The PP, which is also called a histogram, is generated by projecting a binary image onto a one-spatial vector [5]. Preponderantly, a scanned document image achieves maximum peaks of the histogram if its content is not skewed. To detect the inclination, a document image has to be rotated and examined the peaks of its projection profile continuously. This procedure caused the fundamental problem of traditional PP-based approaches which is high computational cost [6]. To overcome the obstacle, various researchers computed the histogram based on only a portion of the entire documents [7]. Additionally, another drawback of the conventional PP-based approach is the instability of accuracy when encountering dense-content images. Recently, to resolve the mentioned issue, Bao et al [2] proposed piecewise projection profile method in 2022. The method introduces valley, which is the number of black pixels belonging to only a specific rectangle area of horizontal projection profile, to determine the skew angle. The rotation angle, which minimizes the valley value, is the estimated inclination angle of input document. The approach could decrease by approximately 80% of time complexity while still achieving similar accuracy as the traditional projection profile method. Furthermore, the effectiveness of PP-based approaches is also challenged by document images with graphics [6]. In 2020, to estimate tilt angle effectively on graphic document images, Khuman et al. [8] combines horizontal projection profile and entropy. The efficiency of the proposed algorithm was proved on well-known benchmarks such as ICDAR 2013 and Meitei/Meetei document images with achieved precision of 0.5. Finally, yet significantly, PP-based approaches for document image skew estimation and correction are also applicable for woven fabric images [9].

Hough transform, which is introduced by Paul Hough in 1962, is a competent method for line detection on two-level images [10]. The method employs a mathematical characteristic in which intersection points in Polar coordinates is corresponding to lines in Cartesian coordinates [6]. For years, the Hough transform has been effectively adopted for de-skewing document images suffering inclination angles [11]. In 2022, Bao et al. [2] utilized the Hough line detection method for estimating tilt angle of form images such as document images with bordered tables. To overcome noise in skew images, Probabilistic Hough Transform (PHT) and clustering are combined by Ahmad et al. [6] in order to estimate the inclination of document images. Firstly, PHT is utilized for locating lines. Subsequently, cluster with the maximum number of parallel lines is identified and adopted for determining the tilt angle. Furthermore, Morphological Skeleton and Progressive Probabilistic Hough Transform (PPHT) are employed to de-skew historical documents [5]. Firstly, the Morphological Skeleton is used for removing redundant pixels. Afterward, PPHT is utilized for locating image lines in order to estimate global inclination angle of the input document images.

In recent times, deep learning approaches are employed for estimate inclination angle of scanned images. In 2020, Akhter

et al. [12] proposed a ConvNet architecture with an output layer of 360 classes corresponding to slant angles between 0 and 359°. A limitation of this method is that it only results in discrete skew angles. On the other hand, the actual inclination angles can take any continuous value within the range of -180° to 180°.

In document image analysis, understanding table document images has attracted considerable attention over the recent decade [13]. However, to the best of our knowledge, there is no test set built specifically for the problem of table document image skew estimation and correction. Additionally, based on our survey, most of the document image skew estimation and correction approaches are only evaluated on or capable of tackling limited inclination angle range such as (- 90°; 90°) [14], (-45°; 45°) [2], (-44.9°; 44.9°) [15]. Furthermore, deep learning approaches have not received enough attention. In this research, we propose a novel approach to skew estimation and correction for table document images which is built on robust deep learning models for text detection. Our proposed method has the capability of handling any continuous inclination angle from -180° to 180°. Moreover, we also introduce a new test set (called the AICS Table Skew) specifically designed for evaluating methods of table document image skew estimation and correction. The results of our proposed approach can be the baseline for future studies on the AICS Table Skew test set.

Overall, the main contributions of our research are:

- A new test set named AICS Table Skew is created, which contains only scanned tables. The tables in the dataset have a high accuracy of rotated angles and a wide diversity of contents and characteristics.
- A novel method using information extracted by text detection methods is introduced to de-skew table document images.

II. Proposed Method

A. Method Overview

In this paper, a skew estimation and correction method is proposed for scanned tables, which have diverse contents and structures. For example, tables have variants of bordered or borderless. Fig.1 shows 3 main stages namely text detection, skew angle estimation, and rotation & post processing. At the beginning of the process, text detection is performed on scanned skew table, resulting in 4-sided polygons for text. In the skew angle estimation step, skew angle is detected by using the ten longest lines of ten 4-sided polygons. Then in the final step, document table images are rotated by two possible skew angles. Subsequently, PP-OCR model is used to detect texts of the two rotated images and the image with a higher OCR confidence score is selected. Finally, the three steps are applied to the selected image again to achieve higher accuracy.

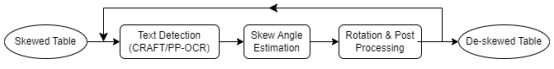


Fig. 1. Block diagram of proposed method

Our proposed approach for skew estimation and correction is illustrated in Algorithm 1 as follows.

Algorithm 1: Pseudo code for our proposed method

INPUT: Scanned document image with skewed table

TEXT DETECTION

Step 1:
CRAFT model, which is used for detecting text in the input image, returns 4_sided_text_polygons

SKEW ANGLE ESTIMATION

Step 2:
intermediate_lines = []
for each polygon in 4_sided_text_polygons **do**
polygon_longest_line = the longest line of the polygon
intermediate_lines.append(polygon_longest_line)
end for
selected_lines = list of the 10 longest lines of immediate_lines

Step 3:
selected_angles = [], angles of the longest lines selected
for each line in selected_lines **do**
θ = angle of line with the image's horizontal line
selected_angles.append(θ)
end for

Step 4:
Calculate average angle α and standard deviation σ of selected_angles

Step 5:
if σ > 0.5 **then**
Remove the line in selected_lines which has the maximum $|\theta - \alpha|$ and its θ in selected_angles then return to **Step 4**
end if

ROTATION & POST PROCESSING

Step 6:
if α < 0 **then**
α' = α + 180
else
α' = α - 180
end if
rotated_img_a = The input image is rotated by angle -α
rotated_img_a' = The input image is rotated by angle -α'

Step 7:
Determine the confidence score of each table using PP-OCR
x = confidence_score(rotated_img_a)
y = confidence_score(rotated_img_a')
if x > y **then**
output_img = rotated_img_a
else
output_img = rotated_img_a'
end if

repeat steps 1 – 7 once using PP-OCR in step 1

return output_img

B. Method Details

1) Text Detection

Step 1: In our method, the image goes through two separate text detection steps, with each process being performed once. In the first process, we utilized CRAFT (Character Region Awareness for Text Detection) [16], which provides 4-sided polygons for individual words and performs exceptionally well when dealing with significant skew. This allowed us to

obtain precise results in such cases. Then at this stage of the second process, we use PP-OCR [17], which provides 4-sided polygons for words that are close together. Thus, it performs effectively when dealing with minor skew. By adopting the two text detection methods, our system achieved promising accuracy across various skew conditions.

2) Skew Angle Estimation

Based on our observation, as the length of the longest line in 4-sided polygons increases, the accuracy of estimating skew angles improves. Therefore, text polygons which have the longest lines are employed to estimate skew angle of the table image. The specific details of this stage are described below.

Step 2: Determining the length of each side of each text polygon using (1):

$$line = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \quad (1)$$

where (x_1, y_1) and (x_2, y_2) represent the coordinates of the vertices of the text polygons determined before. For each polygon, find its longest line.

Step 3: Calculating the angles θ_i formed by the longest lines of each text polygon with the Ox axis, as shown in Figure 2:

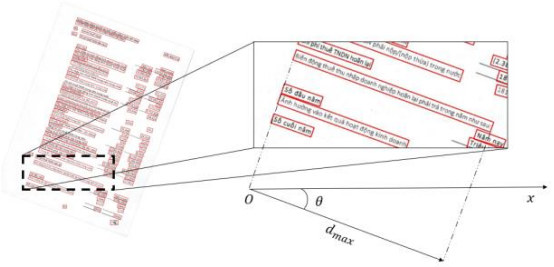


Fig. 2. Calculating the angle θ_i

Step 4: Calculating average angle α and standard deviation σ formed by the 10 angles of the 4-sided polygons with the longest lines using (2) and (3):

$$\alpha = \frac{\sum_{i=1}^{n} \theta_i}{n} \quad (2)$$

$$\sigma = \sqrt{\frac{\sum_{i=1}^{n} (\theta_i - \alpha)^2}{n - 1}} \quad (3)$$

where n represents the number of selected polygons.

Step 5: To achieve the highest accurate result, the permissible standard deviation is set as STD_THRESH = 0.5°. Therefore, if σ is greater than STD_THRESH, the text polygons and its angle, which have maximum $|\theta - \alpha|$, will be removed. The average angle α is updated again and the process is repeated until σ does not exceed STD_THRESH.

Finally, when the standard deviation reaches the permissible level, the last calculated angle α will be considered as the skew angle of the table.

3) Rotation & Post Processing

Step 6: At this stage, the document table image is rotated by two different angles, denoted as $-\alpha$ and $-\alpha'$, where angle α' is calculated using (4) because if the image is only rotated by $-\alpha$, there is a chance that the image will be upside-down.

$$\alpha' = \begin{cases} \alpha - 180^0, if\ \alpha \geq 0 \\ \alpha + 180^0, otherwise \end{cases} \quad (4)$$

Step 7: After the original table image is rotated by the two angles, two different tables are returned: one is correctly rotated, and the other is upside-down. Through the experimental process, the confidence score when applying PP-OCR on the correctly rotated tables is higher than the upside-down tables. Therefore, the image with a higher confidence score is selected to avoid upside-down situation.

Finally, the above process is repeated one time with the use of PP-OCR in step 1 to increase the accuracy.

III. EXPERIMENTS AND RESULTS

In this section, we will describe how the AICS Table Skew test set is collected and present the evaluation results of the proposed method on this test set. The method's performance was evaluated using standard evaluation metrics such as Average Error Deviation and Standard Deviation. Furthermore, an analysis was conducted to examine the distribution of images based on the skew angles after the correction.

A. Test set preparation

Our test set includes financial statements and social insurance table images collected from the official websites of banks, companies, and insurance organizations in Vietnam. To ensure that the tables are not skew, for each image, 6 points named A, B, C, D, E, and F along the bottom table line are created (or another horizontal line on the document if the table has no bottom line) to calculate three angles formed by the axis Ox and the sides AB, CD, and EF. Images with a mean value of the three angles less than 0.05 degrees are considered as the non-skewed images. Then, these non-skewed images are augmented, resulting in a total of 720 images with angles ranging from -175 to 180 degrees, with a 5-degree step.

B. Experimental Results

Following the test set's preparation, the proposed method was applied to each document image for evaluation purposes.

For every document table image j of our test set, the distance $E(j)$ between the ground-truth and the estimation of our method was calculated for evaluation. In order to measure the performance of our method, the following three criteria were used: the Average Error Deviation (AED), the Standard Deviation (STD), and the percentage of Correct Estimations (CE). The definition of the above criteria is given in the rest of this section.

The AED, STD and CE criterion are described by the following (5), (6) and (7) where n equals to 720 denotes the number of table images of our test set:

$$AED = \frac{\sum_{j=1}^{n} E(j)}{n} \quad (5)$$

$$STD = \sqrt{\frac{\sum_{j=1}^{n} \big(AED - E(j)\big)^2}{n - 1}} \quad (6)$$

$$CE = \frac{\sum_{j=1}^{n} K(j)}{n} \ where\ K(j) = \begin{cases} 1\ if\ E(j) \leq 1^\circ \\ 0\ otherwise \end{cases} \quad (7)$$

TABLE I. AED, STD, AND CE OF PROPOSED METHOD ON AICS TABLE SKEW

Metric	***AED*(°)**	***STD*(°)**	***CE*(%)**
Value	0.21	0.24	98.47

As shown in Table I, our algorithm demonstrates exceptional performance by achieving an AED of only 0.21 degrees. This result on the AICS Table Skew test set is equivalent to some other methods on the DISEC'13 [18] benchmark data set. Additionally, when applying a threshold of 1 degree, our method accomplished a Correct Estimations (CE) rate of 98.47%.

TABLE II. TABLE OF DISTRIBUTION OF ANGLE COUNTS BY VALUE RANGES

Angle ranges	**[0°;0.1°)**	**[0.1°;0.5°)**	**[0.5°;1°)**	**[1°;1.8°)**
Number of images	300	348	61	11

For deeper analysis, a statistical examination of the distribution of skew angles after the correction is conducted, as shown in TABLE II. Out of the total of 720 images, 709 images (98.47%) are less than 1 degree. This result demonstrates the success of the skew correction by the proposed method, leading to accurate results.

IV. CONCLUSION AND FUTURE WORK

In our study, an accurate method based on text detection to estimate and correct skew angles in table images, which has been evaluated on our AICS Table Skew test set is proposed. Furthermore, our approach is independent of table characteristics and relies exclusively on the textual content within the image. This feature enables our method to be applicable not only to table images but also to diverse types of text images. However, the existence of a few images (1.53%) with skew angles exceeding 1 degree indicates the limitation of our method. Future research is required to handle these cases as well as optimize the accuracy and speed of our method, which allows us to achieve better results and advancements in various applications.

REFERENCES

[1] Liu, L., Wang, Z., Qiu, T., Chen, Q., Lu, Y., & Suen, C. Y. (2021). Document image classification: Progress over two decades. *Neurocomputing*, *453*, 223-240.

[2] Bao, W., Yang, C., Wen, S., Zeng, M., Guo, J., Zhong, J., & Xu, X. (2022). A Novel Adaptive Deskewing Algorithm for Document Images. *Sensors*, *22*(20), 7944.

[3] Rice, S. V., Jenkins, F. R., & Nartker, T. A. (1995). *The fourth annual test of OCR accuracy* (Vol. 3). Technical Report 95.

[4] Yin, Z. C., Wang, M., Xu, G., Mo, C., Tang, W. H., & Zhang, L. L. (2020, November). Recognition and Extraction of Information from Image-based Tables for Electric Power System Operation and Maintenance. In *2020 International Conference on Smart Grids and Energy Systems (SGES)* (pp. 533-537). IEEE.

[5] Boudraa, O., Hidouci, W. K., & Michelucci, D. (2020). Using skeleton and Hough transform variant to correct skew in historical documents. *Mathematics and computers in simulation*, *167*, 389-403.

[6] Ahmad, R., Naz, S., & Razzak, I. (2021). Efficient skew detection and correction in scanned document images through clustering of probabilistic hough transforms. *Pattern Recognition Letters*, *152*, 93-99.

[7] Kavallieratou, E., Fakotakis, N., & Kokkinakis, G. (2002). Skew angle estimation for printed and handwritten documents using the Wigner–Ville distribution. *Image and Vision Computing*, *20*(11), 813-824.

[8] Khuman, Y. L. K., Devi, H. M., & Singh, N. A. (2021). Entropy-based skew detection and correction for printed meitei/meetei script ocr system. *Materials Today: Proceedings*, *37*, 2666-2669.

[9] Yildirim, B. (2014). Projection profile analysis for skew angle estimation of woven fabric images. *The Journal of The Textile Institute*, *105*(6), 654-660.

[10] Hassanein, A. S., Mohammad, S., Sameer, M., & Ragab, M. E. (2015). A survey on Hough transform, theory, techniques and applications. *arXiv preprint arXiv:1502.02160*.

[11] Srihari, S. N., & Govindaraju, V. (1989). Analysis of textual images using the Hough transform. *Machine vision and Applications*, *2*, 141-153.

[12] Akhter, S. S. M. N., & Rege, P. P. (2020, July). Improving Skew Detection and Correction in Different Document Images Using a Deep Learning Approach. In *2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT)* (pp. 1-6). IEEE.

[13] Hashmi, K. A., Liwicki, M., Stricker, D., Afzal, M. A., Afzal, M. A., & Afzal, M. Z. (2021). Current status and performance analysis of table recognition in document images with deep neural networks. *IEEE Access*, *9*, 87663-87685.

[14] Cai, C., Meng, H., & Qiao, R. (2021). Adaptive cropping and deskewing of scanned documents based on high accuracy estimation of skew angle and cropping value. *The Visual Computer*, *37*, 1917-1930.

[15] Pham, L., Hoang, P. H., Mai, X. T., & Tran, T. A. (2022, October). Adaptive Radial Projection on Fourier Magnitude Spectrum for Document Image Skew Estimation. In *2022 IEEE International Conference on Image Processing (ICIP)* (pp. 1061-1065). IEEE.

[16] Baek, Y., Lee, B., Han, D., Yun, S., & Lee, H. (2019). Character region awareness for text detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 9365-9374).

[17] Du, Y., Li, C., Guo, R., Yin, X., Liu, W., Zhou, J., ... & Wang, H. (2020). Pp-ocr: A practical ultra lightweight ocr system. *arXiv preprint arXiv:2009.09941*.

[18] Papandreou, A., Gatos, B., Louloudis, G., & Stamatopoulos, N. (2013, August). ICDAR 2013 document image skew estimation contest (DISEC 2013). In *2013 12th International Conference on Document Analysis and Recognition* (pp. 1444-1448). IEEE.

Automated Selection of Optimal Number of Topics in Latent Dirichlet Allocation using Jensen Shannon Divergence for Text Classification

Noor Ashikin Abdullah
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
noor.ashikin@graduate.utm.my

Siti Sophiayati Yuhaniz
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
sophia@utm.my

Syahid Anuar
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
syahid.anuar@utm.my

Azizul Azizan
Razak Faculty of Technology and Informatics
Universiti Teknologi Malaysia
Kuala Lumpur, Malaysia
azizulazizan@utm.my

***Abstract*— Latent Dirichlet Allocation (LDA) is a popular statistical model used for topic recognition by extracting keywords from unread documents. Researchers and practitioners face the challenge of selecting the optimal number of topics, denoted as 'k', as the ground truth is unknown and the chosen number of topics may exceed the optimal one. Perplexity and topic coherence are commonly used methods for selecting 'k', but they require manual steps. This research proposes an implementation using Jensen Shannon Divergence (JSD), a heuristic method that estimates the number of topics in a given query from retrieved documents. The goal of this study is to enhance the LDA model by automatically selecting the optimal number of topics based on JSD, reducing the computational burden while maintaining good performance. The proposed model, demonstrated in this research, utilizes maximum likelihood for fitting the data, with JSD providing precise measurement of the goodness-of-fit and efficient properties. The research consists of three phases: determining characteristics to improve LDA performance, developing and enhancing the LDA model for efficient topic selection using JSD, and validating the proposed model through classification evaluation and computational efficiency analysis. The datasets are generated from Reddit and Twitter posts on depression and non-depression. Bag of Words (BoW) vector, one of Vector Space Model (VSM) and optimal number of topics are used to generate the latent topic distribution with feature vectors that can be categorized using machine learning algorithms. LDA-JSD outperform other methods with ET classifier that are 99% accuracy and 0.995 F1-score. The result show that our proposed method may effectively improve performance accuracy. When compared to other existing methodologies, LDA-JSD execution is more efficient, and the resulting features can be applied to text classification.**

Keywords*— *Latent Dirichlet Allocation, Topic Detection, Natural Language Processing, Text Classification, Artificial Intelligence

I. Introduction

Latent Dirichlet Allocation (LDA) is a widely used statistical model for topic recognition and keyword extraction from unread documents. The selection of the optimal number of topics (k) poses a challenge as the ground truth is typically unknown, and choosing a higher number of topics than necessary can lead to suboptimal results. Existing methods for determining the optimal number of topics, such as perplexity and topic coherence, have limitations and require manual intervention. This research aims to address this issue by proposing an approach that utilizes Jensen Shannon Divergence (JSD) to automatically select the optimal number of topics in LDA. By estimating the number of latent concepts (topics) in a given query from retrieved documents, JSD offers the potential for accurate and efficient topic selection. Additionally, JSD has shown promise in performing big-data Genome-wide association studies (GWAs) for identifying complex diseases' genome-wide multi-locus interactions. The objective of this research is to enhance the LDA model by leveraging JSD for efficient and reliable topic selection, reducing computational burden while maintaining good performance.

II. Literature Review

The literature on topic modeling and selection methods for Latent Dirichlet Allocation (LDA) [1] reveals various approaches and challenges. Despite the availability of techniques such as perplexity and topic coherence for determining the optimal number of topics, it is noteworthy that many researchers still rely on manual approaches to select the appropriate value of k [2]. The manual selection process raises concerns regarding subjectivity and the need for more efficient and automated methods. Researchers have also investigated alternative measures, such as Jensen Shannon Divergence (JSD) [4], which accurately estimates the number of latent concepts (topics) based on retrieved documents. Moreover, JSD has shown promise in big-data Genome-wide association studies (GWAs) by detecting complex diseases' genome-wide multi-locus interactions. However, there is a need for an efficient and accurate method that combines the benefits of JSD with LDA to automatically select the optimal number of topics. This literature review highlights the gap in research and the necessity for an improved LDA model that leverages JSD for efficient and reliable topic selection, while

reducing the computational burden and maintaining good performance.

III. Methodology

In this research, a comprehensive methodology is applied to enhance the Latent Dirichlet Allocation (LDA) model for efficient and accurate selection of the optimal number of topics. The methodology consists of three main phases.

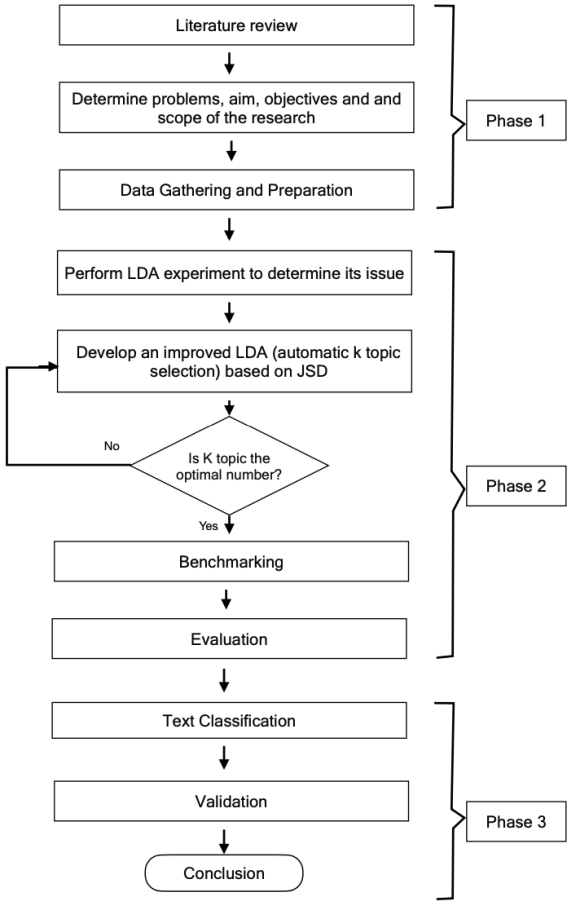


Fig. 1. The research methodology

In the first phase, we aim to determine the characteristics that can improve the performance of the LDA model. This involves an in-depth analysis of the LDA algorithm and its underlying assumptions, as well as exploring various factors that may impact the quality of topic modeling. By identifying these characteristics, we can better understand the strengths and limitations of the LDA model and lay the foundation for its improvement.

The second phase focuses on the development and enhancement of the LDA model to efficiently select the optimal number of topics using Jensen Shannon Divergence (JSD) as a key component. JSD, a heuristic method, has the potential to accurately estimate the number of latent concepts (topics) in a given query based on retrieved documents. We will integrate JSD into the LDA model, leveraging its ability to measure the goodness-of-fit precisely and efficiently for maximum likelihood estimation. By incorporating JSD, we aim to provide a more robust and automated approach for selecting the optimal number of topics in LDA.

In the third phase, we will validate the proposed model through classification evaluation and computation efficiency analysis. Classification evaluation will assess the performance of the enhanced LDA model in terms of topic quality, coherence, and interpretability. We will compare the results with baseline LDA models and other existing methods for topic selection. Additionally, we will conduct computation efficiency analysis to measure the reduction in computational burden achieved by the proposed model. Performance time analysis will be conducted to evaluate the efficiency of the model in handling large-scale datasets and complex scenarios.

Throughout the methodology, we utilize relevant datasets and benchmark tests to ensure the reliability and validity of the proposed enhancements. The results obtained from these experiments will provide insights into the effectiveness and practicality of the proposed methodology in improving the LDA model for topic selection.

IV. The Experimental Design

Our research includes a technical overview of methods used for identifying depression via NLP and text classification techniques. Data is cleaned and features are extracted using Bag of Words (BoW) vector, one of vector space models (VSM) [4]–[7] and number of topics, k to generate the latent topic distribution with feature vectors that can be categorized using machine learning classifiers, and finally experimental setup are shown in Fig. 2.

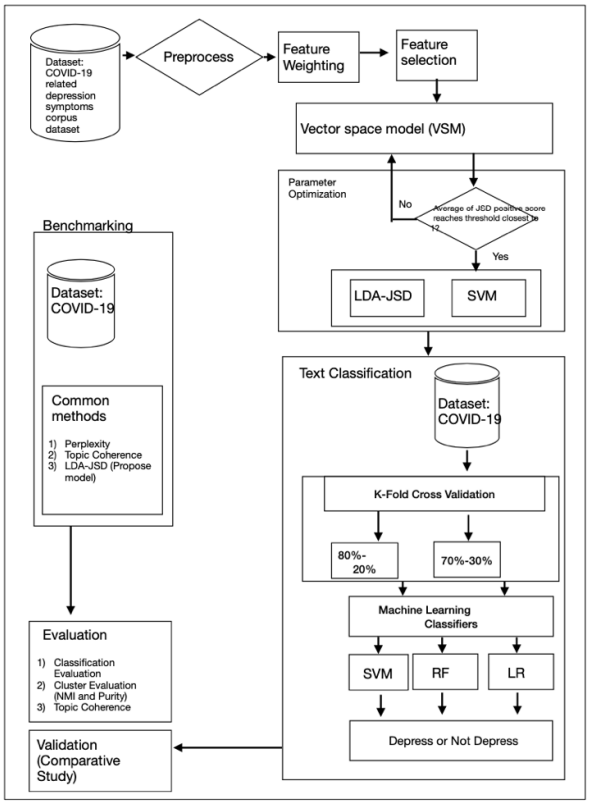


Fig. 2. The Experimental Setup

A. Benchmark Evaluation Datasets

For benchmarking, our suggested model is compared with baseline LDA following research by [8], we used D1 Reddit-based created by [9]. This dataset contains a list of depressed and non-depressed users from Reddit. The data corpus comprises 1293 depression posts and 548 non-depression posts. Posts submitted by non-depressed people are gathered from subreddits about education and personal finance.

B. Validation Datasets

For training, we used D2 Reddit-based dataset created by [10], which comprises of postings known as subreddits that were extracted with a focus on particular mental health communities of Reddit users during COVID-19, such as depression subreddit (r/depression) and non-mental health subreddits, such as education (r/teaching) and personal finance (r/personalfinance). We collected 3000 depression posts and 603 non-depression ones.

For testing, we used D3 Twitter-based dataset. The dataset, which was compiled by [11], [12], includes a collection of Twitter users' posts during COVID-19 . We sample 29752 negative sentiments that can be classified as depress posts and 246 positive sentiments as non-depress posts [13], [14].

V. Experiments and Analysis for Benchmarking

In this section, we conduct experiments to perform benchmarking with methods such as baseline LDA model and our proposed model, LDA-JSD using 10-fold cross validation to classify depression posts. We run each model on D1 dataset and report the accuracy, f1-score, precision and recall.

A. Classification Accuracy

Table I shows performance results of the classification models. We applied three machine learning classifiers: AdaBoost, Support Vector Machine (SVM) and Random Forest (RF) following [8]. Our proposed model (LDA-JSD) outperforms baseline LDA with Random Forest classifier to detect depression with 70% accuracy and F1-score 0.80 score as well as AdaBoost classifier with 71% and F1-score 0.81.

TABLE I. PERFORMANCE RESULTS OF THE CLASSIFICATION MODELS

Met.	*Number of Selected Features*	*Classifier*	*Acc.*	*F1-score*	*Precision*	*Recall*
LDA	70	AdaBoost	66%	0.74	0.61	0.95
		SVM	72%	0.80	0.75	0.88
		RF	68%	0.67	0.68	0.68
LDA-JSD	5	AdaBoost	71%	0.81	0.75	0.89
		SVM	72%	0.83	0.73	0.96
		RF	70%	0.80	0.76	0.84

VI. Experiments and Analysis for Comparative Study

In this section, we conduct experiments to compare our proposed method with other methods to select optimal number of topics. We validate each method on D3 dataset using three evaluation metrics that are classification accuracy, clustering evaluation and topic coherence. First metric is using classification accuracy used accuracy, f1-score, precision and recall. Second metric is Purity [15] and Normalized Mutual Information (NMI) [16] are used for clustering evaluation. Next is Normalised Pointwise Mutual Information (NPMI) [17] is used for evaluating topic coherence, an improved formulation replacing PMI with log conditional probability. After that, we compare the performance time of various topic selection methods.

A. Classification Accuracy

Table II depicts performance results of the classification models. We apply SVM and Extra Trees (ET) for this experiment following research by [10]. LDA-JSD outperform other methods with ET classifier that is 99% accuracy and 0.995 F1-score.

TABLE II. PERFORMANCE RESULTS OF THE CLASSIFICATION MODELS BASED ON SELECTED FEATURES

Met.	*Number of selected features*	*Classifier*	*Acc. (%)*	*F1-score*	*Precision*	*Recall*
Topic Coherence	7	SVM	100.0	0.996	0.993	1.00
		ET	98.9	0.994	0.994	0.99
Perplexity	2	SVM	100.0	0.996	0.993	1.00
		ET	98.6	0.992	0.994	0.99
LDA-JSD	4	SVM	100.0	0.996	0.993	1.00
		ET	99.0	0.995	0.993	1.00

B. Clustering Evaluation and Topic Coherence

Table III depicts the evaluation results of the clustering based on the different number of optimal k. We apply Purity and NMI to perform cluster evaluation and NPMI for topic coherence evaluation. Perplexity outperforms other methods with NMPI score 0.0044.

TABLE III. CLUSTERING EVALUATION AND TOPIC COHERENCE

Method	*Number of Selected Features/ Optimal 'k'*	*Purity*	*NMI*	*NPMI*
Topic Coherence	7	0.9917	1.0	-0.0392
Perplexity	2	0.9917	1.0	0.0044

LDA-JSD	4	0.9917	1.0	0.0023

C. Performance Time

Fig. 3 depicts performance time executed on training data between LDA-JSD, Perplexity and Topic Coherence. Based on the figure, LDA-JSD is more efficient than other methods.

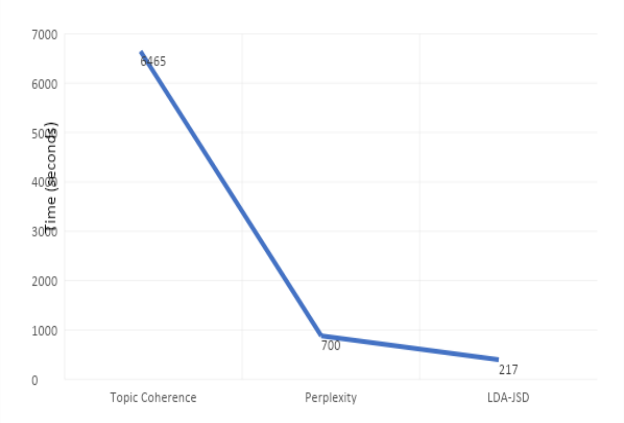


Fig. 3. The performance time

VII. DISCUSSION

For benchmarking section, LDA-JSD achieved higher accuracy compared to baseline LDA. LDA-JSD automatically select true optimal 'k' that provide accurate number of selected features to classify the depression posts.

Validation results show that LDA-JSD excels in overall evaluation metrics, classification accuracy and clustering evaluation. However, for topic coherence (NMPI) evaluation, perplexity excels. We compare the efficiency of various methods using performance time to select optimal 'k'. Among all methods evaluated, LDA-JSD is the most efficient as expected. Even though, our proposed method falls behind perplexity in topic coherence evaluation but it still indicates as effective because number of topics when reach optimal threshold value 1, JSD able to detect full differences between topics and not overlap in terms of their probability distributions. Even though Topic coherence with ET classifier that is 98.9% accuracy that selects optimal number of topics or achieve number of selected features, 7 achieves negative score, -0.0392 in NPMI compared to other methods. Moreover, topic coherence also spends much more time compared to the other two methods.

VIII. CONCLUSIONS

This paper presents a new method for selecting a number of LDA topics for text classification based on JSD. It proposes a new method for automatically detecting optimal number of topics from social media in order to classify depression-COVID-19 related posts. Our research indicates that selecting the optimal number of topics can lead to appropriate feature extraction and increase the accuracy of text classification results. For future work, we would like to increase and test the sample size of datasets to take advantage of distributed or parallel environment that can be used in handling large-scale datasets.

ACKNOWLEDGMENT

This work was financially supported by the Ministry of Higher Education Malaysia under Fundamental Research Grant Scheme FRGS/1/2021/ICT06/UTM/02/6).

REFERENCES

[1] D. M. Blei, A. Y. Ng, and J. B. Edu, "Latent Dirichlet Allocation Michael I. Jordan," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, 2003.

[2] S. Sbalchiero and M. Eder, "Topic modeling, long texts and the best number of topics. Some Problems and solutions," *Qual. Quant.*, vol. 54, no. 4, pp. 1095–1108, Aug. 2020, doi: 10.1007/s11135-020-00976-w.

[3] F. Nielsen, "On the Jensen–Shannon Symmetrization of Distances Relying on Abstract Means," Entropy, vol. 21, no. 5, May 2019, doi: 10.3390/E21050485.

[4] C. C. Aggarwal and C. X. Zhai, "Mining text data," *Mining Text Data*, vol. 9781461432234, pp. 1–522, Aug. 2013, doi: 10.1007/978-1-4614-3223-4/COVER.

[5] [2] F. Wang and K. Orton, "Community verification with topic modeling," Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), vol. 10251 LNCS, pp. 278–288, 2017, doi: 10.1007/978-3-319-60033-8_25/COVER.

[6] [3] Y. Fang, Y. Guo, C. Huang, and L. Liu, "Analyzing and Identifying Data Breaches in Underground Forums," *IEEE Access*, vol. 7, pp. 48770–48777, 2019, doi: 10.1109/ACCESS.2019.2910229.

[7] [4] C. Badenes-Olmedo, "Semantically-enabled Browsing of Large Multilingual Document Collections," 2021.

[8] [5] M. M. Tadesse, H. Lin, B. Xu, and L. Yang, "Detection of depression-related posts in reddit social media forum," *IEEE Access*, vol. 7, pp. 44883–44893, 2019, doi: 10.1109/ACCESS.2019.2909180.

[9] [6] I. Pirina and Ç. Çöltekin, "Identifying Depression on Reddit: The Effect of Training Data," pp. 9–12, Jun. 2018, doi: 10.18653/V1/W18-5903.

[10] [7] D. M. Low, L. Rumker, T. Talkar, J. Torous, G. Cecchi, and S. S. Ghosh, "Natural language processing reveals vulnerable mental health support groups and heightened health anxiety on reddit during COVID-19: Observational study," *J Med Internet Res*, vol. 22, no. 10, Oct. 2020, doi: 10.2196/22635.

[11] [8] R. Lamsal, "Design and analysis of a large-scale COVID-19 tweets dataset," *Applied Intelligence*, vol. 51, no. 5, pp. 2790–2804, May 2021, doi: 10.1007/S10489-020-02029-Z/FIGURES/9.

[12] [9] R. Lamsal, M. R. Read, and S. Karunasekera, "CORONAVIRUS GEO-TAGGED TWEETS DATASET," 2023, Accessed: Jul. 17, 2023. [Online]. Available: https://ieee-dataport.org/open-access/coronavirus-covid-19-geo-tagged-tweets-dataset

[13] [10] S. Ghosh and T. Anwar, "Depression Intensity Estimation via Social Media: A Deep Learning Approach," *IEEE Trans Comput Soc Syst*, vol. 8, no. 6, pp. 1465–1474, Dec. 2021, doi: 10.1109/TCSS.2021.3084154.

[14] [11] X. Tao, R. Dharmalingam, J. Zhang, X. Zhou, L. Li, and R. Gururajan, "Twitter Analysis for Depression on Social Networks based on Sentiment and Stress," *BESC 2019 - 6th International Conference on Behavioral, Economic and Socio-Cultural Computing, Proceedings*, Oct. 2019, doi: 10.1109/BESC48373.2019.8963550.

[15] [12] Y. Zhao and G. Karypis, "Criterion Functions for Document Clustering * Experiments and Analysis", Accessed: Apr. 09, 2023. [Online]. Available: http://www.cs.umn.edu/

[16] [13] C. Xue, J. Liu, X. Li, and Q. Dong, "Normalized-Mutual-Information-Based Mining Method for Cascading Patterns," *ISPRS International Journal of Geo-Information 2016, Vol. 5, Page 174*, vol. 5, no. 10, p. 174, Sep. 2016, doi: 10.3390/IJGI5100174.

[14] J. H. Lau, D. Newman, and T. Baldwin, "Machine Reading Tea Leaves: Automatically Evaluating Topic Coherence and Topic Model Quality," *14th Conference of the European Chapter of the Association for Computational Linguistics 2014, EACL 2014*, pp. 530–539, 2014, doi: 10.3115/V1/E14-1056.

Modelling and Optimization of a Couple of Block-shaped Metasurface for Q-factor Enhancement.

1st D, Mahmudin
Research Center for Telecommunication
National Research and Innovation Agency
Bandung, Indonesia
Faculty of Electrical Engineering
Universiti Teknologi Malaysia
Johor Bahru, Malaysia
dadin@graduate,utm.my,
dadi005@brin.co.id

2nd S.M. Idrus
Faculty of Electrical Engineering
Universiti Teknologi Malaysia, Johor Bahru, Malaysia
sevia@utm.my

3rd M.Y.M. Noor
Faculty of Electrical Engineering
Universiti Teknologi Malaysia, Johor Bahru, Malaysia
yusofnor@utm.my

4th Y. N. Wijayanto
Research Center for Electronics
National Research and Innovation Agency
Bandung, Indonesia
yusu008@brin.co.id

***Abstract*— Research on metasurface-based terahertz(THz) technology is in great demand because of its advantages and uniqueness. One of the advantages of all-dielectric metasurfaces over metallic metasurfaces is their lower ohmic loss. The Q-factor is often a consideration in designing metasurface devices. Based on these descriptions, an All-dielectric metasurface structure of a block-shaped silicon-based unit cell is proposed. In this paper, the variables of the unit cell structure are optimized using electromagnetic analysis software. The trend of changing the geometric size of the unit cell to the Q-factor is analyzed. In the results of the simulation, a curve is obtained that describes the trend of changes in the Q factor when the unit cell variable size is changed. These results are expected to guide the design of metasurfaces in the THz range with the structures and materials discussed in this paper.**

Keywords—all-dielectric metasurface, terahertz, Q-factor, electromagnetic analysis software.

I. INTRODUCTION

In the last decade, electromagnetic waves in the range of 0.1 to 100 THz, known as the Terahertz spectrum, have been investigated in various applications [1]. Similarly, to electromagnetic wave radiation in general, THz has the same properties such as reflection, refraction, interference, and diffraction. In addition to having its uniqueness, such as being non-pengion, it can identify many chemical and biological molecules through its characteristic resonance peaks and its ability to detect hidden objects [2]. In the field of telecommunications technology, the THz frequency range can meet the needs of high-speed wireless communications with wide bandwidth and high data throughput [3]. Although it has great potential in a wide variety of application fields, the terahertz spectrum has some drawbacks. In some cases, the utilization of the THz spectrum requires high costs, large dimensions, and difficult fabrication processes [4].

On the other hand, research in the field of metasurfaces has been carried out over all frequency ranges in many applications. [5] The properties of electromagnetic waves can be arbitrarily engineered as desired by adjusting the dimensions, shape, and material of the metasurface structure, thus allowing for the expansion of the field of application that can be explored. The high-quality factor (Q factor) as a consequence of sharp resonance is often one of the essential goals of metasurface design for various applications such as narrow band filters, ultra-sensitive sensors, and slow light devices. [6-8] Due to the high Q factor of the resonator, the device produces strong light-matter interactions so that an efficient platform in various application fields can be provided [8].

In the THz frequency range, the exploration of artificial "metasurface" materials to increase the Q-factor resonance is widely carried out. Initially, the material used to construct the metasurface structure was metal [9–10]. However, because the metallic metasurface causes ohmic loss, it is difficult to achieve high Q-factor resonance levels. This causes limited performance of metal-based metasurfaces in terms of sensitivity, operating speed, and transmission efficiency [11].

The high-index dielectric metasurface is the solution to metallic metasurface problems. In addition to being free from ohmic loss, low thermal conductivity, and efficient electromagnetic control [8, 11, 12]. Dielectric metasurface, known as All-Dielectric Metasurface (ADM), can generate mie resonance [12, 14]. Various advantages of mie resonance, such as effective control of the phase, amplitude, and polarization of the incident wave, resulting in perfect transmission, a perfect reflector, perfect damping, and nonlinear optics with strong electromagnetic energy localization, can be provided [13–15].

Based on the previous description, in this study, a metasurface structure with a unit cell of a couple of blocks shaped and attached obliquely with an angular orientation to a central point is proposed. The proposed metasurface structure can be seen in Fig. 1. The material used to build this metasurface is all dielectric. The basic material and unit cell are constructed by silicon dioxide (SiO2) and silicon, respectively. The more the variety of variables in the structure metasurface that can be changed, the more variations can be adjusted to find the optimum structure. Optimization of the proposed structure is employed by simulation using Electromagnetic Analysis Software (CST). The frequency response obtained from the simulation results will be calculated for the Q-factor value. From the simulation results obtained by varying the size of each variable in the unit cell, a graph is obtained that illustrates the tendency to change the Q factor when the size of the unit cell variable is changed. These results are expected to be a guide for designing metasurfaces in the THz range with the structures and materials discussed in this paper.

II. Proposed Design

A. Design

The metasurface design proposed in this study uses silicon dioxide as a substrate with a refractive index (n) of 1.94. A couple of block-shaped unit cells are placed on a substrate with an orientation angle of phi to the centre point. The unit cell material used is silicon (S_i) with a refractive index value (n) of 3.42. The schematic design of the proposed metasurface structure can be seen in Fig. 1. Referring to Zhong et al., the dimension variable size is P_1 = 176 mm, P_2 = 110 mm, h_1 = 40 mm, h_2 = 45 mm, L_y = 88 mm, and L_x = 35 mm [7].

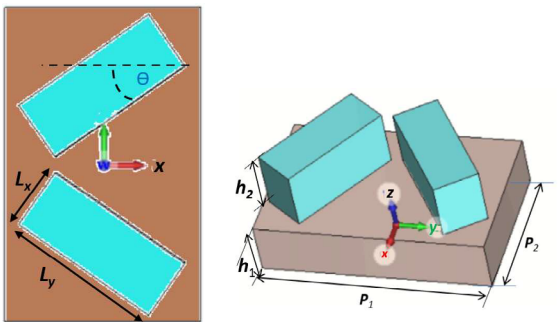


Fig. 1. Proposed Design

These variables are dimensional variables whose values will be varied to determine the effect of changes on the Q-factor and its working frequency. Based on these data, it will be possible to determine the value of the dimensional variable that produces the most optimal Q-factor and frequency shift. The unit cell dimensional variables can be seen in Table I and Table II.

TABLE I. List of Unit Sel Dimensional Variables

Variable	
h_1	Substrate thickness (µm)
h_2	Resonator thickness (µm)
P_1	Substrate length (µm)
P_2	Substrate Width (µm)
L_x	Resonator Width (µm)
L_y	Resonator Length (µm)
phi (θ)	The angle of the orientation of the resonator to the center of the resonator (^{0})
R	The distance between the two center points of the resonator (µm)

B. Simulation

In this study, the simulation was carried out using the Electromagnetic Analysis (CST) software. If the input beam has impinged on the metasurface structure, then the reflected output beam will be investigated. Furthermore, all variables affecting the size of the metasurface structure will be varied. The effect of changes in metasurface structural variables on the output beam frequency response will be investigated.

From the simulation results, the frequency response will be obtained in the form of the S - parameter, and then the Q-factor value can be obtained. The Q-factor value can be obtained by the formula,

$$Q_{Factor} = \frac{\lambda}{FWHM} \quad (1)$$

where FWHM means full width at half the maximum of the peak value of the graph [14], meanwhile λ is the position of the peak of Resonance. After that, the tendency of changes in the size of each variable to the Q factor will be analyzed.

TABLE II. Variation Range of Simulation Dimensional Variable

Variable		*Width step*
θ	0-25^0	-
h_2	45-80 µm	5 µm
L_x	35-65 µm	5 µm
L_y	72-96µm	5 µm
R	0-15 µm	5 µm
h_1	30-60 µm	5 µm
P_1	168-188 µm	4 µm
P_2	102-126 µm	4 µm

III. Results and Discussion

In this section, we will discuss the simulation results of phenomena that occur in the output beam due to changes in the environment of the metasurface structure. Fig. 2 shows the different frequency responses due to changes in the size of the geometry of the metasurface structure. It is seen that this change affects the Q-factor, resonant frequency, and reflectance. The spectral colours indicate the frequency response of different unit cell sizes. This means that the output beam properties can be manipulated by changing the environment of the metasurface structure.

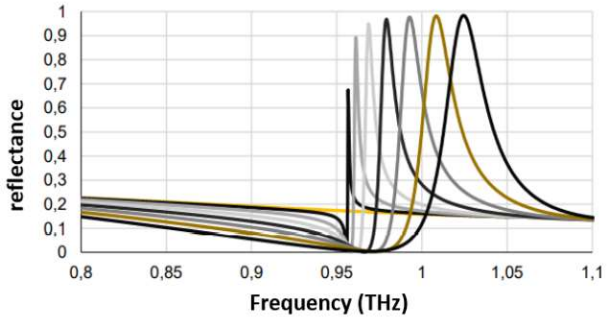


Fig. 2. Frequency response of the proposed Metasurface.

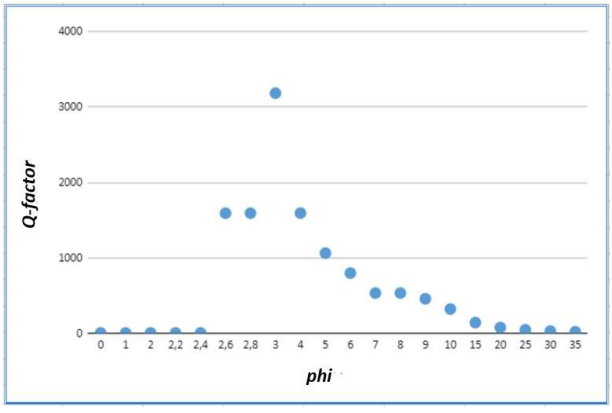


Fig. 3. Q-factor curve to changes in the value of *phi.*

After that, we discuss changes in Q-factors due to environmental changes in the metasurface

structure. Environmental changes occur due to geometric changes in the metasurface structure. The Q-factor is obtained from the spectrum of the output beam using formula 1. Fig. 3-10 is a graph that explains the tendency of the Q-factor value due to changes in the geometry of each parameter.

Fig. 3 shows the curve of the Q-factor change when the angular θ position of the unit cell is varied. The unit cell angle position between 2.5^0 and 5^0 produces a good Q-factor, with a Q-factor above 1000. The best Q-factor is in the 3^0 position, which is 3183. After 50, the Q-factor continues to decrease; even after a position angle of 15^0, the Q-factor approaches 0. Fig. 4 shows the curve for the trend of the Q factor value due to changes in the thickness of the unit cell h_2. The value of the Q factor decreases with increasing thickness. The Q-factor also decreases with increasing length and width of the unit cell, although there is a slight increase in the Q-factor with increasing L_x width between 35 µm and 45 µm. These observations can be seen in Fig. 5 and 6.

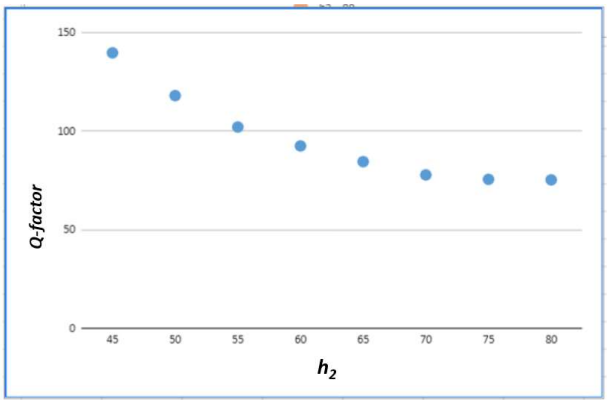


Fig. 4. Q-factor curve to changes in the value of h_2.

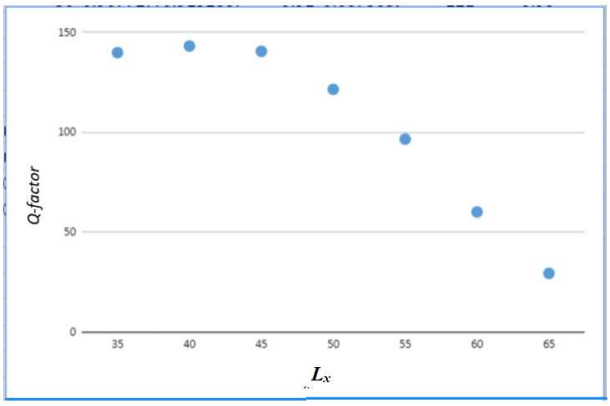


Fig. 5. Q-factor curve to changes in the value of L_x.

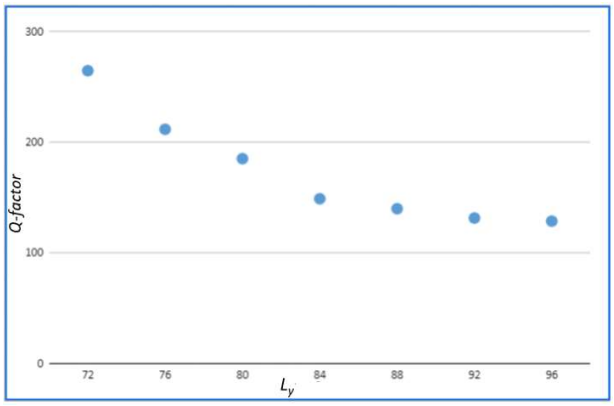


Fig. 6. Q-factor curve to changes in the value of L_y.

Fig, 7,8 and 9, are curves of the results of changes in the substrate. Varied parameters were height h1, length P1, and width P2. The change in h1 produces a curve that shows a value that decreases as the height of the substrate increases. While the increase in the value of the Q-factor occurs when the length of the substrate is increased. Changes also occur when the width of the P2 substrate is varied. From Fig. 9, it can be seen that changes in substrate width have a slight effect on the Q-factor.

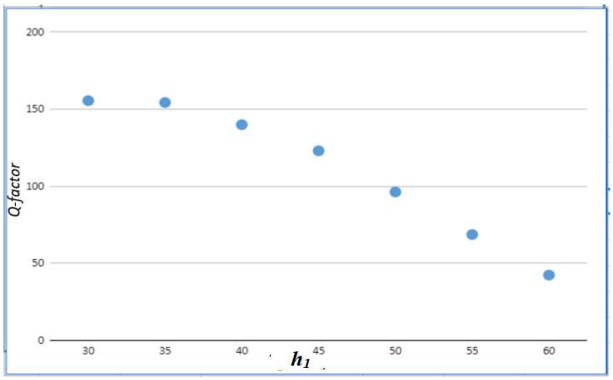


Fig. 7. Q-factor curve to changes in the value of h_1

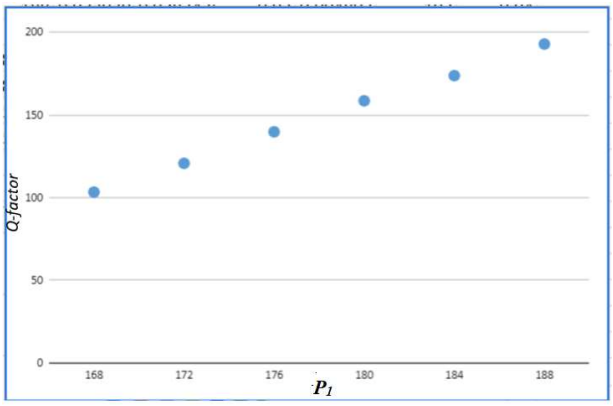


Fig. 8. Q-factor curve to changes in the value of P_1.

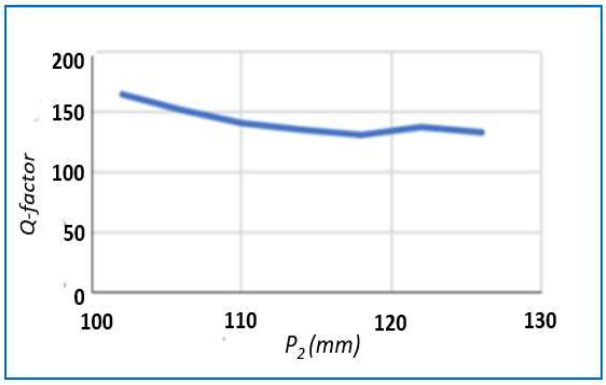


Fig. 9. Q-factor curve to changes in the value of P_2.

A conclusion about the relationship between changes in the value of the dimensional variable and changes in the values of the Q-factor, resonant frequency, and peak reflectance can be drawn from a series of tests of variations in the value of the metasurface unit cell dimensional variable carried out in the previous experiment. The average change in the variable's value produces the same effect, but the changes in phi and P2 get slightly different results. The two variables have the potential to be further varied in different combinations of variables. The relationship between the Q-factor parameter and the resonant frequency with an increase in the value of the dimensional variable is summarized in the following Table III to facilitate reading.

TABLE III. SUMMARY

Dimensional variables	Q-factor	Resonant frequency
φ	Decrease when φ = 3	Shift to the right
h_2	Decrease	Shift to the left
L_x	Decrease	Shift to the left
L_y	Decrease	Shift to the left
R	Increase when close to R = 0	Shift to the left when close to R = 0
h_1	Decrease	Shift to the left
P_1	Increase	Shift to the left
P_2	Decrease	Shift to the right

IV. CONCLUSION

An all-dielectric metasurface with a couple of block-shaped silicon-based unit cell structures is proposed. The optimization of the proposed structure is simulated using Electromagnetic Analysis Software (CST) by optimizing the variables of the unit cell structure, resulting in varying frequency responses. The trend curve for the change in the Q-factor when the unit cell variable size is changed has been obtained. These results can be used as a guide in designing metasurfaces in the THz range with the structures and materials discussed in this paper.

ACKNOWLEDGEMENT

This work was conducted as a part of Universiti Teknologi Malaysia (UTM) and Badan Riset Inovasi Nasional, Indonesia (BRIN) collaborative research grant vot R.J130000.7351.4B734. This work also partly supported by Organisasi Riset Elektronika dan Informatika, Badan Riset Inovasi Nasional (BRIN), in Rumah Program Sistem Kendaraan Listrik Otonom, 2023, NOMOR 2/III.6/HK/2023.

REFERENCES

[1] H.-J. Song and N. Lee, "Terahertz Communications: Challenges in the Next Decade," *IEEE Transactions on Terahertz Science and Technology*, vol. 12, no. 2, pp. 105–117, Mar. 2022, doi: 10.1109/tthz.2021.3128677.

[2] [1]S. I. Gusev, P. S. Demchenko, E. A. Litvinov, O. P. Cherkasova, I. V. Meglinski, and M. K. Khodzitsky, "Study of glucose concentration influence on blood optical properties in THz frequency range," *Nanosystems: Physics, Chemistry, Mathematics*, pp. 389–400, Jun. 2018, doi: 10.17586/2220-8054-2018-9-3-389-400.

[3] Y. Huang, Y. Shen, and J. Wang, "From Terahertz Imaging to Terahertz Wireless Communications," *Engineering*, vol. 22, pp. 106–124, Mar. 2023, doi: 10.1016/j.eng.2022.06.023.

[4] [1]H. W. Tian, H. Y. Shen, X. G. Zhang, X. Li, W. X. Jiang, and T. J. Cui, "Terahertz Metasurfaces: Toward Multifunctional and Programmable Wave Manipulation," *Frontiers in Physics*, vol. 8, Oct. 2020, doi: 10.3389/fphy.2020.584077.

[5] M. Chen, M. Kim, A. M. H. Wong, and G. V. Eleftheriades, "Huygens' metasurfaces from microwaves to optics: a review," *Nanophotonics*, vol. 7, no. 6, pp. 1207–1231, Jun. 2018, doi: 10.1515/nanoph-2017-0117.

[6] Z. Zheng *et al.*, "Planar narrow bandpass filter based on Si resonant metasurface," *Journal of Applied Physics*, vol. 130, no. 5, p. 053105, Aug. 2021, doi: 10.1063/5.0058768.

[7] Y. Zhong *et al.*, "Ultrasensitive specific sensor based on all-dielectric metasurfaces in the terahertz range," *RSC Advances*, vol. 10, no. 55, pp. 33018–33025, 2020, doi: 10.1039/d0ra06463g.

[8] F. Yan, Q. Li, Z. Wang, H. Tian, and L. Li, "Extremely high Q-factor terahertz metasurface using reconstructive coherent mode resonance," *Optics Express*, vol. 29, no. 5, p. 7015, Feb. 2021, doi: 10.1364/oe.417367.

[9] S. I. Maslovski, "Subwavelength imaging with arrays of plasmonic scatterers," *Optics Communications*, vol. 285, no. 16, pp. 3363–3367, Jul. 2012, doi: 10.1016/j.optcom.2011.12.079.

[10] H. Tao, W. J. Padilla, X. Zhang, and R. D. Averitt, "Recent Progress in Electromagnetic Metamaterial Devices for Terahertz Applications," *IEEE Journal of Selected Topics in Quantum Electronics*, vol. 17, no. 1, pp. 92–101, Jan. 2011, doi: 10.1109/jstqe.2010.2047847.

[11] X. Liu, K. Fan, I. V. Shadrivov, and W. J. Padilla, "Experimental realization of a terahertz all-dielectric metasurface absorber," *Optics Express*, vol. 25, no. 1, p. 191, Jan. 2017, doi: 10.1364/oe.25.000191.

[12] Q. Zhao, J. Zhou, F. Zhang, and D. Lippens, "Mie resonance-based dielectric metamaterials," *Materials Today*, vol. 12, no. 12, pp. 60–69, Dec. 2009, doi: 10.1016/s1369-7021(09)70318-9.

[13] H. Chen, C. Zhou, Y. Li, S. Liang, and J. Li, "Dielectric tetrahedrons as terahertz resonators switched from perfect absorber to reflector," *Scientific Reports*, vol. 10, no. 1, Oct. 2020, doi: 10.1038/s41598-020-74252-0.

[14] Y. Kivshar, "All-dielectric meta-optics and non-linear nanophotonics," *National Science Review*, vol. 5, no. 2, pp. 144–158, Jan. 2018, doi: 10.1093/nsr/nwy017.

[15] L. Yang *et al.*, "High-efficiency all-dielectric transmission metasurface for linearly polarized light in the visible region," *Photonics Research*, vol. 6, no. 6, p. 517, Apr. 2018, doi: 10.1364/prj.6.000517.

Design of a Narrow Channel Navigation Assistance System with Real-Time Path Deviation Monitor(PDM) and Obstacle Avoidance Function

Dayoung Kang
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
myjoo119@naver.com

Junbum Park
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
jb03park@navy.ac.kr

Wansik Choi
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
akaman0222@navy.ac.kr

Minseok Han
dept. of electronics and control engineering
Republic of Korea Naval Academy
Changwon, South Korea
mshan1024@navy.ac.kr

***Abstract*— This paper proposes a system that utilizes radar to search for stationary objects (such as lighthouses, fixed structures, etc.) in narrow channels. It performs real-time identification of the distance and angle to the detected objects, enabling determination of the optimal point for maintaining or altering a specific course. Furthermore, in the presence of moving ships and obstacles, the system assesses the speed vector of the obstacle to evaluate the risk of collision. It actively avoids obstacles classified as dangerous and restores the vessel to its original course.**

Keywords—narrow channel, navigation assistance system, Path Deviation Monitor(PDM), obstacle avoidance

I. INTRODUCTION

Narrow channels are defined as waterways with limited width, often serving as the primary access points to naval bases. Navigating through these channels requires heightened caution due to the confined space and high concentration of ships and structures. Collisions involving ships or obstacles frequently occur in such areas. To mitigate these accidents, the navy has implemented various measures, including increased personnel deployment and extended duty hours. This paper aims to contribute to accident prevention in narrow channels by proposing a Narrow Channel Navigation Assistance System.

Current autonomous vessel technology, classified under the International Maritime Organization's (IMO) marine autonomy levels, has reached the second stage out of four, which involves the vessel's ability to autonomously determine the optimal path [1]. However, due to the unique security and operational mission requirements of warships, directly adopting civilian ship technologies that rely on vision sensors for obstacle detection is unsuitable. Moreover, considering the navy's need to rapidly assess and respond to the dynamic maritime conditions within narrow channels, relying solely on path optimization technology presents challenges in assisting warship navigation.

Traditionally, when deviating from the designated optimal course, experienced officers rely on subjective judgment to determine the deviation point and avoid moving obstacles. While this subjective judgment may not burden seasoned officers, it can pose challenges for junior officers who are still developing their navigation skills. Therefore, the need for an assistance system that alleviates the burden on junior officers while they refine their navigation techniques becomes imperative.

This study proposes a radar-based system for navigating fixed markers such as lighthouses and buoys within narrow channels. It calculates the real-time distance and angle to these markers to guide the deviation point or maintain the course, with the system automatically implementing the required adjustments. Additionally, when encountering moving ships or obstacles, the system assesses collision risks based on the obstacle's velocity vector. It actively avoids obstacles categorized as dangerous and subsequently returns to the original course. This paper suggests the implementation of an automatic control algorithm to assist warship operations in narrow channels, aiming to enhance operational stability and ensure the safety of the crew during navigation. Furthermore, it will reduce the burden on junior officers who face the complexities of navigating through narrow channels.

II. THEORETICAL BACKGROUND

A. The Concept of CPA, DCPA, TCPA

Closest Point of Approach (CPA) refers to the minimum distance between the two vessels when the own and target ships maintain the current course and speed, and gradually approach and sail. Further, Time to Closest Point of Approach (TCPA) refers to the time required for two vessels to reach the CPA and Distance at Closest Point of Approach (DCPA) related on the minimum distance between them. VTSO utilizes CPA and TCPA as major indicators to select priorities among vessel relationship pairs in the process of vessel traffic management. CPA and TCPA can be obtained using the principle shown in Fig. 1.

If the coordinates, course, and speed of one's own ship are (x_0, y_0), φ_o, and V_o respectively, and those of the other ship are

(x_t, y_t), φ_t, and V_t respectively, then the time it takes to reach the CPA and the distance between one's own ship and the other ship at the CPA are calculated as follows.

$$D_r = \sqrt{(x_t - x_0)^2 - (y_t - y_0)^2} \quad (1)$$

$$V_r = V_o \times \sqrt{1 + \left(\frac{V_t}{V_o}\right)^2 - 2 \times \left(\frac{V_t}{V_o}\right) \times \cos(\varphi_O - \varphi_t)} \quad (2)$$

$$\varphi_r = \cos^{-1} \times \left(\frac{V_o - \cos(\varphi_O - \varphi_t)}{V_r}\right) \quad (3)$$

$$T_{CPA} = D_r \times \cos(\varphi_r - \alpha_t - \pi)/V_r \quad (4)$$

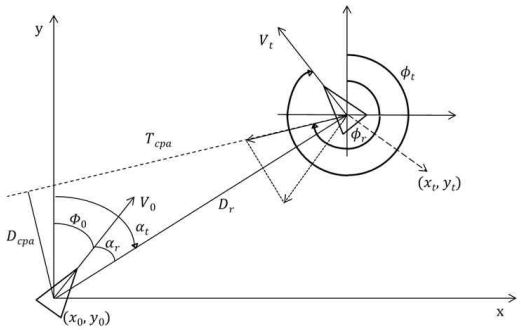


Fig. 1. Concept of CPA, DCPA, and TCPA

In this context, D_r is the relative distance between one's own ship and the other ship, V_r is the relative speed, φ_r is the relative course, α_t is the bearing of the other ship, and α_r is the relative bearing. The closer both TCPA and DCPA are to 0, the higher the collision risk [2],[3].

B. Determing the Safety Distance

There have been numerous studies aimed at defining a safety zone and reducing the risk of collision between vessels. According to Shaobo, over 90% of respondent sailors did not take any evasive action if the DCPA was more than 1.5NM [4]. Lyu and Sawada used a zone with a radius of 0.3NM as a risk area in their avoidance experiments, while Sawada [5] set it at 0.5NM.

According to Article 22 of the International Regulations for Preventing Collisions at Sea (COLREGS), the intensity of the mast lights of vessels over 50m long must be set to be visible from a minimum distance of 6NM. Therefore, the range of the safety zone used in this paper will be set as a circle with a radius of 6NM. Zaccone defined collision zones, COLREG application zones, and free navigation zones as concentric circles, distinguishing areas of different risk levels [6]. The detection range in this paper is also composed of two circular zones as shown in Fig. 1. The smaller zone around the target ship is set as a safety zone, with the assumption that ships will collide within a 0.5NM range. The larger zone around one's own ship is the detection circle, where measures need to be taken to prevent collisions if a risk is present within a 6NM radius.

III. System Architecture and Methodology

A. Proposal for the Narrow Channel Autonomous Navigation Assistance System Algorithm

The narrow channel for navigation is defined, and the pre-defined optimal course is loaded. The navigation assistance can be divided into two situations: automatically determining the course alteration point and, when encountering an approaching vessel, avoiding it and returning to the original course.

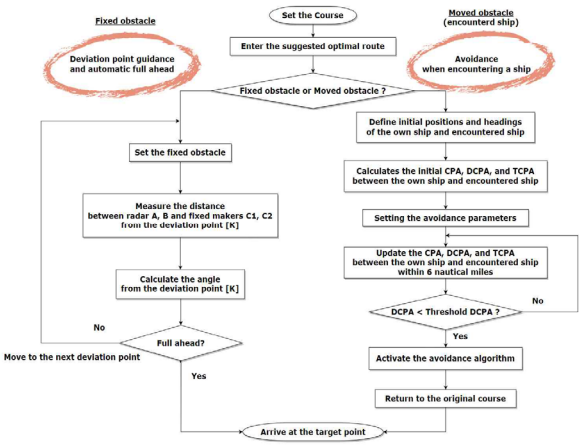


Fig. 2. Flowchart of proposed narrow channel navigation assistance system with real-time path deviation monitor(PDM) and obstacle avoidance function

B. Guiding Course Alteration Points

As shown in Fig. 3, Radar A and B are installed at the bow and stern of the ship. Fig. 3 represents the estimated position when the actual position is measured at 0800. The actual course of the ship changes every moment due to external forces, and the estimated course alteration point also changes as each actual position is updated. Let's say the angles formed by Radar A, B and fixed marks C1, C2 are $\angle AC_1B = \theta_1$, $\angle AC_2B = \theta_2$, respectively. At the course alteration point K, measure the distance between Radar A, B and fixed marks C1, C2 and calculate the angle $\angle AC_1B = \theta_{k1}$ and $\angle AC_2B = \theta_{k2}$. As the ship navigates along the given course, changes continuously, and the course alteration can be started when $\theta_1 = \theta_{k1}$ *and* $\theta_2 = \theta_{k2}$ are simultaneously satisfied.

By inputting the distance from Radar A, B to marks C1, C2, the angle is automatically generated and set as a threshold angle. The system that automatically changes the coordinates and makes the turn when reaching the threshold angle was implemented through MATLAB. Once the turn is completed,

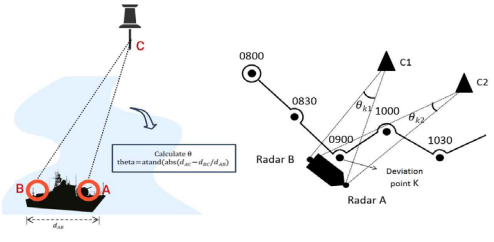


Fig. 3. Modeling of calculate θ and guiding course alteration points

the mark is changed and the angle at the next course alteration point is set as the threshold angle in the same way as before to determine the moment of the turn.

C. Evasion when encountering an approaching vessel

When encountering an approaching ship, if we know the velocity vector of our own ship and the other ship, and the distance between the ships, we can calculate CPA, DCPA, and TCPA as shown in Equation <1>. We set a high collision risk DCPA as the threshold DCPA, and when DCPA < threshold DCPA, we activate the avoidance algorithm, complete the avoidance action, and designed an algorithm that deactivates the avoidance action and returns to the original course when DCPA > threshold DCPA.

IV. RESUTLS AND DISCUSSIONS

A. Scenario Design and Algorithm Application Narrow Channel and Buoy Settings

Fig. 4 is based on the narrow channel of Jinhae Port provided by the National Oceanographic Research Institute, on which we have arbitrarily set an optimal route. We used the GPS Coordinates website to set the position of the buoys and our own ship and measure the distances. Positions are noted as (latitude, longitude). Let the coordinates of Buoy C1 be (35.0632,128.7777), Buoy C2 be (35.0565,128.7886), Radar A be (35.05338, 128.7236), and Radar B be (35.0540,128.7228). The distance between the two radars is said to be 100m. The approximate distance from the ship's radar to Buoys C1 and C2 is $AC_1 = 2.72NM$, $BC_1 = 2.75NM, AC_2 = 3.2NM,$ *and* $BC_2 = 3.24NM.$

Let's also identify the position and distance at the estimated point of course alteration. If we designate the positions of Radars A and B at course alteration point K as $A_k(35.0459,128.733)$, $B_k(35.04528,127.734)$, respectively, then $A_kC_1 = 2.40NM$, $B_kC_1 = 2.42NM, A_kC_2 = 0.87NM,$ *and* $B_kC_2 = 0.82NM.$. If we input this into the system we designed with MATLAB, the results will be $\angle A_kC_1B_k = 1.1924^o$, $\angle A_kC_2B_k = 1.3828^o$. We can proceed with course alteration by hitting the target when the values of $\angle AC_1B$, $\angle AC_2B$ at real-time changing positions become the same as the angle values at the course alteration point.

In actual ship operations, it's difficult to accurately measure the coordinates of the exact location, so it suffices to input only the distance value directly measured from the buoy by the radar. At this time, the distance values to be considered are the length of the ship (the distance between Radar A and B) and the distance to the buoy. After the course alteration is completed, we capture Buoys C3 and C4 again and repeat the same process.

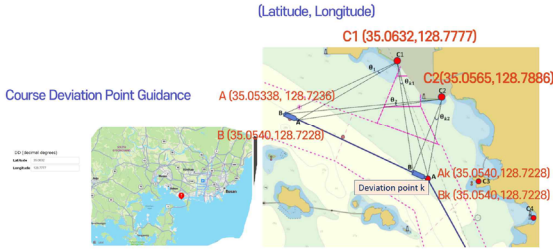


Fig. 4. Modeling of scenario design and algorithm application narrow channel and buoy settings

B. Automatic Avoidance When Encounting Approaching Vessels and Obstacles in a Narrow Channel

In this part, Matlab based code simulates a ship encounter scenario and calculates the CPA, DCPA, and TCPA between two ships. Let's go through the code and explain each part:

The initial positions and headings of the own ship and encountered ship are defined. The own ship's initial position is [0, 0], and its heading is 0 radians (facing right). The encountered ship's initial position is [50, 0], and its heading is π radians (180 degrees) (facing left). The speeds of the own ship and encountered ship are defined. The own ship's speed is 15 knots, and the encountered ship's speed is 10 knots. The threshold CPA value is set to 25 nautical miles. If the calculated CPA is less than or equal to this threshold, an avoidance maneuver will be activated. Simulation parameters are defined, including the time step and duration of the simulation. Time and positions arrays are initialized to store the positions of the own ship and encountered ship at each time step. The code calculates the initial CPA, DCPA, and TCPA between the own ship and encountered ship based on their initial positions, headings, and speeds. The avoidance parameters are set, including the duration of the avoidance maneuver (avoidance Time) and the offset in the y-axis during the avoidance maneuver (avoidance Offset).

Step 1: Initial encounter. The positions of the own ship and encountered ship at the first time step are set to their initial positions.

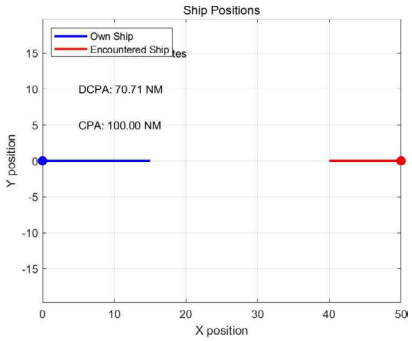
Step 2: CPA < threshold CPA, activate avoidance maneuver. The code iterates through each time step and calculates the relative position and speed between the own ship and encountered ship.

If the avoidance Start Time is 0 (avoidance maneuver not activated) and the calculated CPA is less than or equal to the threshold CPA, the avoidance Start Time is set to the current time step. If avoidance Start Time is greater than 0, it means the avoidance maneuver is activated. The avoidance Active flag is set to true, and the avoidance End Time is calculated by adding the avoidance Time to the avoidance Start Time. During the avoidance maneuver, the new positions of the own ship and encountered ship are calculated based on their speeds, headings, and the avoidance Offset in the y-axis.

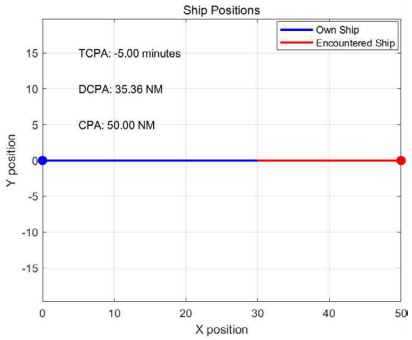
Step 3: Threshold CPA reached, deactivate avoidance maneuver. If the avoidance Start Time is greater than 0 and the current time step is greater than or equal to the avoidance End Time, the avoidance Active flag is set to false.

Step 4: Resume original course. If the avoidance maneuver is not active (avoidance Active is false), the own ship and encountered ship continue on their original course by updating their positions based on their speeds and headings.

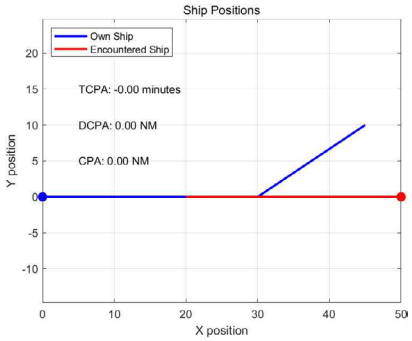
The simulation loop continues until all time steps are processed. After the simulation, the code plots the positions of the own ship and encountered ship throughout the simulation.



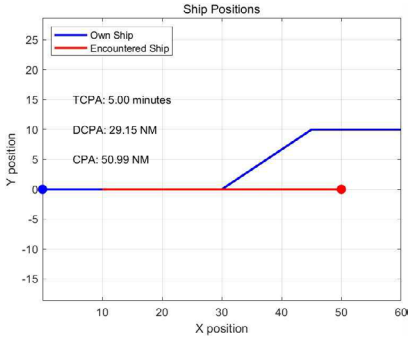
(a) Initial encounter



(b) CPA > threshold CPA



(c) Threshold CPA reached (activate avoidance maneuver)



(d) deactivate avoidance maneuver and original course by updating their positions based on their speeds and headings

Fig. 5. Result of automatic avoidance when encounting approaching vessels and obstacles in a narrow channel in five situations

The CPA, DCPA, and TCPA information is displayed on the plot as show in Fig. 5. It simulates a ship encounter scenario and demonstrates the activation of an avoidance maneuver when the CPA is below a specified threshold. The positions of the ships are plotted, and the CPA, DCPA, and TCPA values are displayed in five situations.

V. Conclusion

The algorithm proposed in this study has successfully implemented a system that facilitates course adjustments based on the angle formed by the radar and buoys when a ship follows a recommended course. Additionally, an auxiliary system has been developed to automatically avoid incoming ships by considering the closest point of approach (CPA) when encountering them. The application of this technology in actual vessel operations is expected to alleviate the burden on officers responsible for steering the ship during navigation. Moreover, it will enhance ship safety by minimizing collision risks with other vessels and obstacles during entry and exit, even in hazardous situations where actual collisions are likely to occur.

Furthermore, in alignment with Navy Vision 2045, the automation of ships is anticipated to address the issue of decreasing naval personnel. Currently, ensuring safety in narrow channel navigation relies on deploying separate personnel for ship position calculation and assigning watchmen. However, meeting these requirements with the existing naval personnel is challenging. By implementing the system proposed in this paper, the flexibility to reallocate personnel within the evolving navy structure is expected. Nonetheless, direct application of the aforementioned technology to the navy still presents challenges, considering radar accuracy, the unique mission of each ship, and maritime conditions.

Therefore, further research is necessary to comprehensively consider various technical aspects. By appropriately adapting radar usage according to the situation, defining conditions that suit the ship's characteristics, and accounting for external forces in the actual maritime environment through simulation, the applicability of the technology can be further enhanced..

References

[1] M. Y. Arafat, M. M. Alam, and S. Moh, “Vision-Based Navigation Techniques for Unmanned Aerial Vehicles: Review and Challenges”, Journal of Drones, Vol. 7, No. 89, pp.1-41, 2023.

[2] H. Namgung, “Local Route Planning for Collision Avoidance of Maritime Autonomous Surface Ships in Compliance with COLREGs Rules”, Journal of sustainability, Vol. 14, No. 198, pp.1-41, 2022.

[3] H. Lyu and Y. Yin, “Fast Path Planning for Autonomous Ships in Restricted Waters”, Journal of applied sciences, Vol. 8, No. 2592, pp.1-41, 2018.

[4] W. Shaobo, Z. Yingjun, and L. Lianbo, “A Collision Avoidance Decision-Making System for Autonomous Ship Based on Modified Velocity Obstacle Method”, Journal of Ocean Engineering, Vol. 215, No. 107910, pp.1-21, 2020.

[5] R. Sawada, K. Sato, and T. Majima, “Automatic Ship Collision Avoidance Using Deep Reinforcement Learning with LSTM in Continuous Action Spaces, Journal of Marine Science and Technology, Vol. 26, pp.509–524, 2021.

[6] R. Zaccone, “COLREG-Compliant Optimal Path Planning for Real-Time Guidance and Control of Autonomous Ships”, Journal of Marine Science and Technology, Vol. 9, No. 405, pp.1-22, 2021.

Downlink Transmission Using Homomorphic Encryption for Secure Satellite Communication in Multiple Ground Station Environments

1st Jaehyoung Park
Mission Operation and Infrastructure Service Division
Korea Aerospace Research Institute
Daejeon, Republic of Korea
jhpark90@kari.re.kr

2nd Gyeoul Lee
Mission Operation and Infrastructure Service Division
Korea Aerospace Research Institute
Daejeon, Republic of Korea
winter@kari.re.kr

3rd Yonggang Kim
Division of Computer Science and Engineering
Kongju National University
Cheonan, Republic of Korea
ygkim@kongju.ac.kr

4th Myeoungshin Lee
Mission Operation and Infrastructure Service Division
Korea Aerospace Research Institute
Daejeon, Republic of Korea
mslee@kari.re.kr

***Abstract*—Advances in satellite development and launch technology have reduced costs, which has led to an increase in the number of satellites present in Earth's orbit. When the launched satellite safely reaches Earth's orbit, it performs missions such as observing Earth, relaying data, and observing climate. Secure satellite communication must be ensured for reliably performing the missions. Owing to the characteristics of satellite communication with a long physical distance, it is vulnerable to radio signal eavesdropping. In this paper, we introduce a technique that can securely perform downlink transmission in an environment where multiple ground stations exist in the downlink range.**

***Index Terms*—Secure Satellite Communication, Multi Ground Station, Secure Multi-Party Computation, Homomorphic Encryption.**

I. Introduction

Space-related technologies have developed remarkably in recent decades, and several institutions and companies have recently launched a large number of satellites and successfully settled them into Earth's orbit. These institutions and companies are continuing to enter the space industry and produce tangible results, and these efforts have reduced the cost of developing and launching satellites [1]. Therefore, more satellites are expected to operate in Earth's orbit soon [2]. Satellites perform important tasks such as earth observation, remote observation for scientific missions, climate observation, and data relay. Because these missions contain important data, secure satellite communication for information delivery is a very important factor. However, satellite communications have a wide downlink range because it transmits wireless signals from hundreds to tens of thousands of kilometers away. Due to these characteristics, it has the disadvantage of being vulnerable to wireless signal eavesdropping [3]. To overcome the security vulnerability of satellite communication, several studies have been conducted, including anti-disruption schemes, anti-spoofing strategies, etc [4].

As the number of satellites operating in Earth's orbit increases rapidly, the role of ground stations that control satellites is growing. Ground stations are responsible for controlling satellites reliably and processing and distributing data transmitted from satellites [5]. The advancement of ground station operation technology and the development of new ground stations are in progress to reliably control and utilize many satellites. Therefore, the probability that multiple ground stations exist within the satellite's downlink transmission range also increases. This paper applies a simple secure multi-party computation technique using learning with error (LWE) based cryptosystem in multi-ground station environments to increase the security level of satellite communication.

II. Preliminary

A. Learning with error (LWE) based cryptosystem

The LWE-based cryptosystem is based on the difficulty of calculating on a lattice and is a technique of hiding plaintext by inserting random noise during encryption. The encryption function is represented as follows:

$$\mathrm{Enc}(m) = (-A \cdot sk + L \cdot m + e) \bmod q, \tag{1}$$

where m is a plaintext, A is a random matrix, sk is a secret key in the form of an integer vector, L is a power of 10, q is a large integer number, and e is a random noise [6]. The decryption function is represented as follows:

$$\mathrm{Dec}(c) = \left\| \frac{(c \cdot sk \bmod q)}{L} \right\| \tag{2}$$

where c is a ciphertext.

B. Secure multi-party computation

Secure multiparty computation is an algorithm that can achieve desired results while preserving data privacy in a distributed system environment [7]. For easier understanding, we can take the problem of calculating the average salary as an example. Considering the three participants, each individual arbitrarily divides the salary into three values. Distribute one of the three divided values to the other two. The shared split salaries are incomplete information. Each individual adds three shared values and sends them to the rest. Finally, the average salary can be calculated without leaking sensitive personal information. This secure multiparty computation can be utilized in various fields such as blockchain, genetic testing, and auction.

III. Simple secure multi-party computation

This paper proposes a secure downlink transmission technique by applying the LWE-based cryptosystem. If only one ground station is located within the satellite's downlink range, the data is protected by applying the cryptosystem. If two or more ground stations are located within the downlink range, a proposed noise cancellation method using homomorphic arithmetic operations based on the LWE-based cryptosystem is applied. Since the location of the ground stations is fixed, the number of ground stations existing within the downlink range can be predicted in advance. Therefore, downlink transmission data may be prepared according to the number of ground stations.

The proposed noise cancellation method uses an LWE-based homomorphic encryption technique. Using the LWE-based homomorphic encryption technique proposed in [8], an addition operation between ciphertexts is possible. The explanation of the assumptions considered in this paper and the procedure for the proposed scheme are as follows:

- The proposed technique considers an environment consisting of one satellite, one primary ground station, and one auxiliary ground station.
- Only the satellite and the primary ground station own the secret key. The auxiliary ground station receives encrypted random noise periodically the satellite through a secure channel. The primary ground station does not know the noise.
- The satellite adds a noise value to the downlink data and encrypts it to transmit it. The auxiliary ground station removes the encrypted noise value from the received data by homomorphic arithmetic operations. The primary ground station acquires data by decrypting the corresponding noise-removed ciphertext.
- If attackers eavesdrop on the satellite's downlink signal, they cannot access the data because it is encrypted. Even if an attacker steals the secret key, data is not accessible because of the noise value. Furthermore, even the primary ground station cannot access the data because the random noise cannot be removed without the auxiliary ground station.
- Therefore, the proposed scheme can increase the security level of downlink transmission in multiple ground stations environments.

When there are two auxiliary ground stations in the downlink range of the satellite, each ground station receives the encrypted random noise from the satellite. The satellite adds two noise values to the downlink data and then performs downlink transmissions. Each auxiliary ground station removes its own noise from the received downlink data by homomorphic operation and then delivers the data to the primary ground station. When d is downlink data, r_1 and r_2 are random noise, the primary ground station can obtain encrypted negative downlink data R by calculating the following equation:

$$R = D_s - D_{a_1} - D_{a_2} = Enc(-d), \tag{3}$$

where D_s is $\mathrm{Enc}(d+r_1+r_2)$, D_{a_1} is $\mathrm{Enc}(d+r_2)$, and D_{a_2} is $\mathrm{Enc}(d+r_1)$. If the auxiliary ground station is N, the proposed noise cancellation scheme can be generalized by referring to the algorithm proposed in [9].

IV. Conclusion

We propose a homomorphic encryption-based noise cancellation scheme that can securely deliver data in an environment where multiple ground stations exist within the downlink range of satellites. The security level of satellite communication has been increased by protecting data even when the secret key is stolen or when the ground station is compromised. In the future, we will try to find ways to further improve performance by comparing the security level and computational complexity with existing security techniques.

References

[1] G. Denis, D. Alary, X. Pasco, N. Pisot, D. Texier, and S. Toulza, "From new space to big space: How commercial space dream is becoming a reality," *Acta Astronautica*, vol. 166, pp. 431–443, 2020.

[2] I. Del Portillo, B. G. Cameron, and E. F. Crawley, "A technical comparison of three low earth orbit satellite constellation systems to provide global broadband," *Acta astronautica*, vol. 159, pp. 123–135, 2019.

[3] A. Kalantari, G. Zheng, Z. Gao, Z. Han, and B. Ottersten, "Secrecy analysis on network coding in bidirectional multibeam satellite communications," *IEEE Transactions on Information Forensics and Security*, vol. 10, no. 9, pp. 1862–1874, 2015.

[4] O. Kodheli, E. Lagunas, N. Maturo, S. K. Sharma, B. Shankar, J. F. M. Montoya, J. C. M. Duncan, D. Spano, S. Chatzinotas, S. Kisseleff, *et al.*, "Satellite communications in the new space era: A survey and future challenges," *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70–109, 2020.

[5] M. Manulis, C. P. Bridges, R. Harrison, V. Sekar, and A. Davis, "Cyber security in new space: analysis of threats, key enabling technologies and challenges," *International Journal of Information Security*, vol. 20, pp. 287–311, 2021.

[6] J. Kim, H. Shim, and K. Han, "Comprehensive introduction to fully homomorphic encryption for dynamic feedback controller via lwe-based cryptosystem," *Privacy in Dynamical Systems*, pp. 209–230, 2020.

[7] Y. Lindell, "Secure multiparty computation," *Communications of the ACM*, vol. 64, no. 1, pp. 86–96, 2020.

[8] C. Gentry, "Fully homomorphic encryption using ideal lattices," in *Proceedings of the forty-first annual ACM symposium on Theory of computing*, pp. 169–178, 2009.

[9] J. Park, N. Y. Yu, and H. Lim, "Privacy-preserving federated learning using homomorphic encryption with different encryption keys," in *2022 13th International Conference on Information and Communication Technology Convergence (ICTC)*, pp. 1869–1871, IEEE, 2022.

A Smart Parking Space Detection System for University Campuses

Indranath Chatterjee*[1]
Dept. of Computer Engineering, Tongmyong University
Busan, South Korea
indranath@tu.ac.kr

Jun Hyeon An
Dept. of Computer Engineering, Tongmyong University
Busan, South Korea
junhyun4554@gmail.com

Min Woo Kim
Dept. of Computer Engineering, Tongmyong University
Busan, South Korea
kmw365849@gmail.com

Gyo Sung Bae
Dept. of Computer Engineering, Tongmyong University
Busan, South Korea
bae749121@gmail.com

Min June Kim
SW Industry-Academy Convergence Center, Tongmyong University
Busan, South Korea
narsiaceo@gmail.com

***Abstract*—To reduce traffic and enhance the parking experience for students, employees, and visitors, it is essential to allocate parking spaces effectively on university campuses. It is a unique deep learning (DL) and Internet of Things (IoT)-based system for real-time parking spot distribution on university campuses, is presented in this study. The system captures real-time photographs of parking spots using IoT devices equipped with Raspberry Pi and web cameras. These images are sent to a cloud server for storage and processing. Using video footage from the parking lot, a sophisticated DL model is built, allowing for reliable detection and prediction of parking spot availability. A smartphone application is being developed to give users real-time information on parking spot availability around campus. Extensive testing is performed on the system to demonstrate its usefulness in accurately identifying and forecasting parking spot availability. The suggested system provides several advantages, such as real-time parking information, improved parking management, and increased customer satisfaction. Future work might concentrate on tackling issues like changing lighting conditions and occlusion of parking spots, as well as expanding the system to include additional features such as parking reservations and payment integration. Overall, this study is a promising solution for intelligent parking space distribution on university campuses, demonstrating the capabilities of DL and IoT technologies in tackling parking management difficulties.**

Keywords— parking space, deep learning, transfer learning, IoT, computer vision, YOLO v5, mobile app development

I. INTRODUCTION

Parking management is an important part of university campuses since it ensures effective use of parking spots while also delivering a seamless experience for students, professors, and visitors. However, a lack of real-time parking space information and improper parking spot distribution can cause congestion and irritation [1]. Emerging technologies such as the Internet of Things (IoT) and Deep Learning (DL) provide promising answers to these difficulties [2].

Recent research has demonstrated the usefulness of deep learning algorithms in a variety of computer vision applications, including object identification and recognition [3]. Furthermore, IoT devices have been widely deployed in smart systems for data gathering and transmission [4]. The integration of DL and IoT technologies in parking systems, particularly in the setting of university campuses, hasn't been fully researched yet [5, 6].

In this study, we present a parking spot allocation system for university campuses based on DL and IoT. Using camera footage, the system uses DL algorithms to discover and find unoccupied parking spaces in real time. In addition, IoT devices such as Raspberry Pi and webcams are used for data collecting and transmission. The suggested method intends to improve parking space usage and user experience on university campuses.

II. RELATED WORK

Efficient parking space distribution systems have received a lot of attention in the literature. Rule-based algorithms or optimization models are frequently used in traditional techniques [7]. These techniques, however, may not take into account real-time information and changing parking demands.

Recent advances in machine learning and deep learning have resulted in the creation of data-driven parking spot allocation systems. These systems improve parking spot allocation by utilizing historical parking data, sensor networks, and real-time occupancy information [8-10]. However, there has been little study into merging DL algorithms with IoT devices for real-time parking spot distribution on university campuses.

Due to its capacity to deliver real-time parking spot information, IoT-based parking systems have attracted a lot of attention. These systems gather and distribute parking occupancy data using sensor networks and communication technologies [11, 12]. These systems provide increased monitoring capabilities and promote optimal parking space distribution by connecting IoT devices with parking infrastructure.

In object detection tasks, deep learning methods, notably convolutional neural networks (CNNs), have exhibited outstanding performance. Several DL-based parking systems

*[1] Corresponding author: Dr. Indranath Chatterjee
(Email: indranath.cs.du@gmail.com)

for reliably detecting and recognizing parking spots have been proposed [13, 14]). YOLO (You Only Look Once) is a popular DL model for object identification that accomplishes real-time detection by splitting the input picture into a grid and predicting bounding boxes and class probabilities [15].

III. Methodology

We used a methodical strategy to construct our DL and IoT-based parking spot allocation system, which included data gathering, model training, IoT framework setup, and integration of DL models with real-time data. Figure 1 shows the schematic diagram for the entire proposed approach.

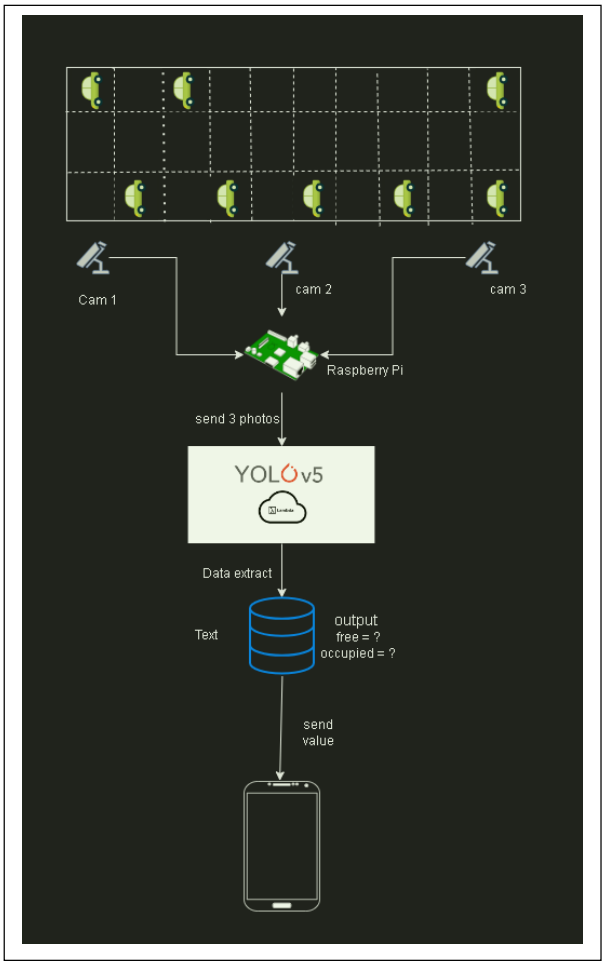


Figure 1: A schematic diagram of the proposed framework for parking spot recognition using DL and IoT

A. Data Collection

Using high-resolution cameras, we captured video data of the parking space at our university campus parking zone. The camera data recorded a variety of parking circumstances, including varying lighting conditions and vehicle kinds. This dataset will be used to train our deep learning model.

B. DL Model Training and Hyperparameter Tuning

As our DL framework for parking spot identification, we choose the YOLOv5 model. To serve as ground truth, the collected video footage was annotated with bounding boxes around parking spots. We next used the annotated dataset to train the YOLOv5 model, optimizing for accuracy and generalization. To improve the model's performance, hyperparameter tweaking approaches such as learning rate scheduling and weight decay were used.

C. IoT Framework Setup

We built an IoT framework utilizing Raspberry Pi and webcams to enable real-time data gathering and transmission. We used a technique called Motion Eye [16] to show the video with the Raspberry Pi's IP address, capture it, and store it directly to the Raspberry Pi. This eliminated the necessity for an intermediary cloud server. The Raspberry Pi operated as the central controller, gathering photos of the parking lot from the attached web cameras regularly, generally every 10 seconds. These collected photos were then easily accessible for additional processing and integration with the YOLO model.

D. Integration of DL Model for Parking Space Detection

The trained YOLOv5 model has been deployed on the cloud server to find parking spaces in real-time using the collected pictures. The algorithm effectively finds empty parking spots and their particular positions inside the parking area by using its capacity to recognize and locate things.

IV. Experimental Results

We conducted thorough experiments and performance assessments to assess the efficacy of our DL and IoT-based parking spot allocation system. We used precision, recall, and F1-score measures to assess the trained YOLOv5 model's performance in spotting unoccupied parking spots. The

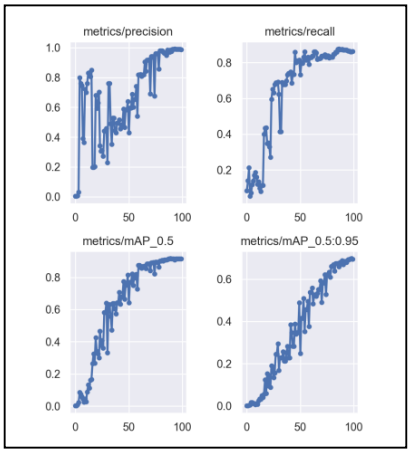


Figure 2: Metrics showing loss, precision, and recall values during the traning process

findings from the study proved the practicality of our approach, with the YOLOv5 model finding unoccupied parking spaces with a training accuracy of roughly 92%. Figure 2 shows the curves for all the metrics being calculated during the training process. Here we have considered the training precision, recall and mean average precision (mAP). Similarly, Figure 3 demonstrate the curves for the metrics being calculated during the testing process. The figures show the F1-Confidence curve, Precision-Confidence curve, Precision-Recall curve, and Recall-Confidence curve. This high level of accuracy assures precise and real-time information on parking spot availability. Figure 4 displays

V. Discussions

Several major advantages and consequences have been revealed by the proposed DL and IoT-based parking spot allocation system for university campuses. In this part, we explain our study's findings and contributions, as well as

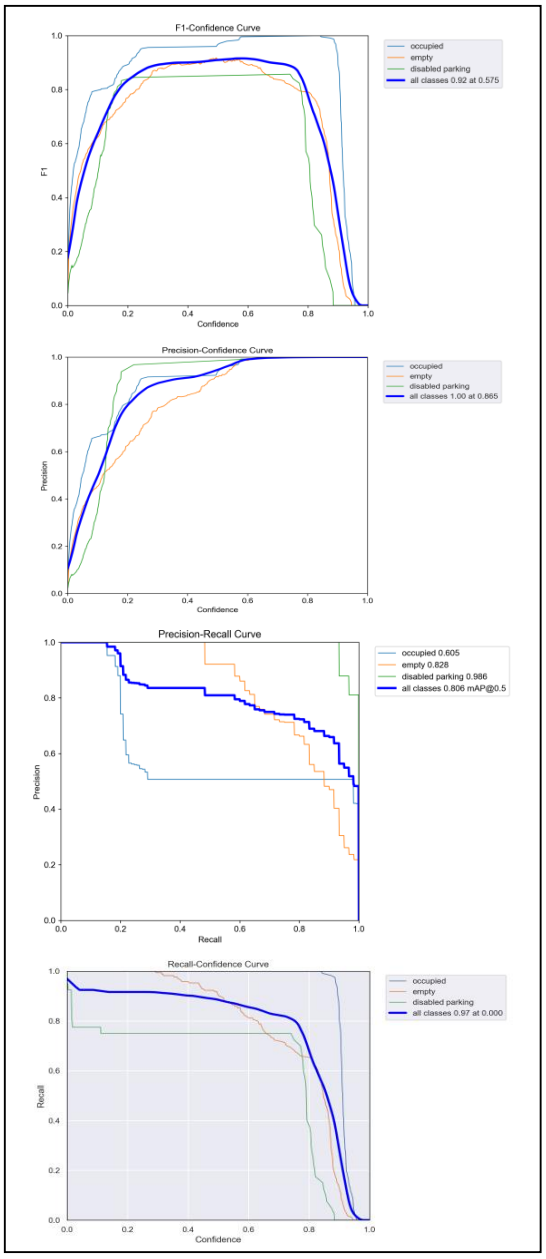


Figure 3: Curves showing the metrics during the testing process. The figures show the F1-Confidence curve, Precision-Confidence curve, Precision-Recall curve, and Recall-Confidence curve.

potential limits and topics for further research and development. The use of DL algorithms and IoT devices in parking management has important implications for improving efficiency and user experience on university campuses. We accomplished precise and real-time identification of unoccupied parking spots by employing DL methods, especially the YOLOv5 model. The trial findings

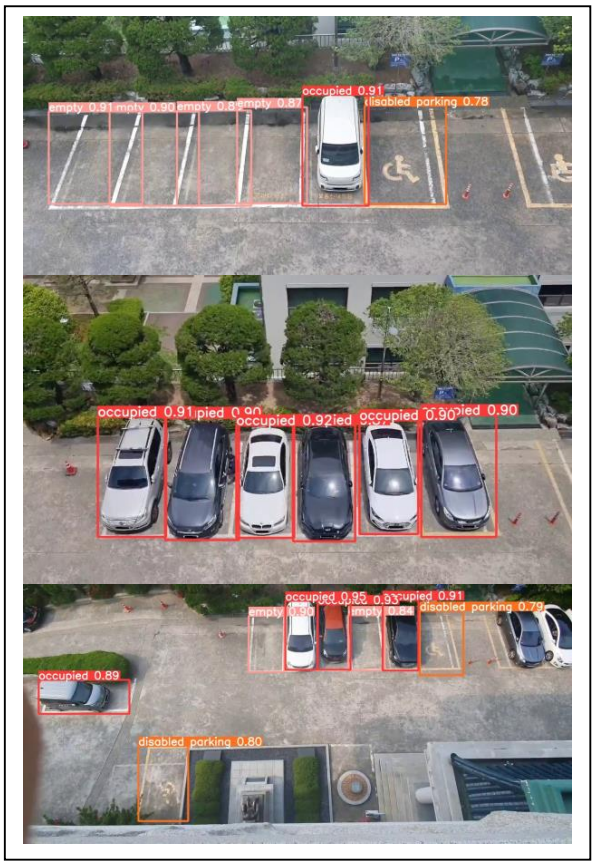


Figure 4: The outcome of our proposed approach showing its efficacy by identifying the empty, occupied and disabled parking spot with high precision.

demonstrated a high accuracy rate in recognizing unoccupied parking spaces, confirming the system's dependability and efficacy.

The use of IoT devices for data collecting and communication, such as Raspberry Pi and webcams, proved to be an efficient and cost-effective option. The IoT framework's implementation enables real-time data gathering and transfer, removing the need for extra cloud servers. We were able to collect and save parking area photos straight to the Raspberry Pi by exploiting Motion Eye technology [1], simplifying the system architecture and lowering processing latency. This simplified technique helps to provide a more seamless and responsive parking spot allocation system.

While our DL and IoT-based system demonstrated promising results, several limitations should be noted. To begin, the DL model's accuracy is strongly dependent on the quality and variety of the training dataset. In this investigation, we used video footage from our university campus parking zone to train the YOLOv5 model. Variations in illumination, weather, and vehicle kinds, on the other hand, may alter the model's performance in different settings. As a result, it is critical to update and increase the training dataset regularly to enhance the model's resilience and generalization.

Furthermore, the suggested system's scalability and deployment issues should be investigated. Our research concentrated on a university campus location, which normally has a reasonable quantity of parking places. In larger-scale

contexts, such as city-wide parking systems, the efficacy, and efficiency of the system must be assessed. Additional problems, like network congestion and computational resource constraints, may occur, necessitating adaptive data collecting and model deployment solutions. The mobile app can provide real-time parking space information to users, allowing them to easily navigate and find available parking spots within the university campus. Additionally, the integration of advanced AI techniques, such as reinforcement learning, can further optimize parking space allocation by considering user preferences and dynamic parking demands.

Finally, by presenting a DL and IoT-based parking space distribution system for university campuses, our research adds to the field of parking management. The technology detects and locates unoccupied parking spaces in real-time, increasing parking space usage and improving overall user experience. While the study emphasizes the potential and advantages of combining DL and IoT technologies, more research and development are required to solve limits, scale the system, and investigate new optimization methodologies.

VI. Conclusions

We introduced a DL and IoT-based parking spot allocation system for university campuses in this study. Our technology detects and locates unoccupied parking spaces in real-time by combining DL algorithms with IoT sensors. The suggested system provides a comprehensive solution for improving parking management, optimizing parking spot distribution, and overall user experience on university campuses. Future work will concentrate on the creation of a user-friendly mobile app that will allow users to get real-time parking spot information. Scalability and deployment issues will also be addressed to assure the system's efficacy in large-scale university campus situations.

Acknowledgment

All the authors are thankful to Tongmyong University, Busan, South Korea for providing the platform for conducting this research work. This research was supported by the MISP (Ministry of Science, ICT, and Future Planning), Republic of Korea, under the National Program for Excellence in SW, supervised by the Institute of Information & Communications Technology Promotion (IITP) with project-sanctioned number 2018-0-018740301001.

References

1. Ali, G., Ali, T., Irfan, M., Draz, U., Sohail, M., Glowacz, A., Sulowicz, M., Mielnik, R., Faheem, Z.B. and Martis, C., 2020. IoT-based smart parking system using deep long short memory network. Electronics, 9(10), p.1696.
2. Tekler, Z.D. and Chong, A., 2022. Occupancy prediction using deep learning approaches across multiple space types: A minimum sensing strategy. Building and Environment, 226, p.109689.
3. Srinivas, J., Das, A.K., Wazid, M. and Vasilakos, A.V., 2020. Designing secure user authentication protocol for big data collection in IoT-based intelligent transportation system. IEEE Internet of Things Journal, 8(9), pp.7727-7744.
4. Ren, S., He, K., Girshick, R. and Sun, J., 2015. Faster r-cnn: Towards real-time object detection with region proposal networks. Advances in neural information processing systems, 28.
5. Adamu, A.A., Wang, D., Salau, A.O. and Ajayi, O., 2020. An integrated IoT system pathway for smart cities. International Journal on Emerging Technologies, 11(1), pp.1-9.
6. Nagowah, S.D., Ben Sta, H. and Gobin-Rahimbux, B., 2021. A systematic literature review on semantic models for IoT-enabled smart campus. Applied Ontology, 16(1), pp.27-53.
7. Oh, S.K., Pedrycz, W. and Park, B.J., 2002. Hybrid identification of fuzzy rule‐based models. International journal of intelligent systems, 17(1), pp.77-103.
8. Vlahogianni, E.I., Kepaptsoglou, K., Tsetsos, V. and Karlaftis, M.G., 2016. A real-time parking prediction system for smart cities. Journal of Intelligent Transportation Systems, 20(2), pp.192-204.
9. Tamrazian, A., Qian, Z. and Rajagopal, R., 2015. Where is my parking spot? Online and offline prediction of time-varying parking occupancy. Transportation Research Record, 2489(1), pp.77-85.
10. Mohapatra, H. and Rath, A.K., 2021. An IoT based efficient multi-objective real-time smart parking system. International journal of sensor networks, 37(4), pp.219-232.
11. Baroffio, L., Bondi, L., Cesana, M., Redondi, A.E. and Tagliasacchi, M., 2015, December. A visual sensor network for parking lot occupancy detection in smart cities. In 2015 IEEE 2nd World Forum on Internet of Things (WF-IoT) (pp. 745-750). IEEE.
12. Atif, Y., Ding, J. and Jeusfeld, M.A., 2016. Internet of things approach to cloud-based smart car parking. Procedia Computer Science, 98, pp.193-198.
13. Rasheed, F., Saleem, Y., Yau, K.L.A., Chong, Y.W. and Keoh, S.L., 2023. The Role of Deep Learning in Parking Space Identification and Prediction Systems. Computers, Materials and Continua, 75(1), pp.761-784.
14. Neupane, D., Bhattarai, A., Aryal, S., Bouadjenek, M.R., Seok, U.M. and Seok, J., 2023. SHINE: Deep Learning-Based Accessible Parking Management System. arXiv preprint arXiv:2302.00837.
15. Redmon, J., Divvala, S., Girshick, R. and Farhadi, A., 2016. You only look once: Unified, real-time object detection. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 779-788).
16. Motion Project Wiki. (n.d.). Motion Eye. Retrieved June 10, 2023, from https://github.com/ccrisan/motioneye/wiki

Robust Long-Term Prediction of Vehicle Trajectory with Adaptive Link Projection

Minsung Kim and Taewoon Kim*
Pusan National University
Busan, South Korea
msungkim,taewoon@pusan.ac.kr

***Abstract*—This paper proposes a robust long-term prediction of vehicle trajectory, which is important in intelligent transportation system. Compared to the short-term prediction, forecasting a long-term sequence is especially challenging since the error accumulates as the sequence becomes longer. In this paper, we propose to utilize the actual road information to minimize the prediction errors. The proposed method projects the predicted trajectory onto the nearest link geometry so that the adjusted prediction is placed on the road. The proposed method has been validated by using the actual vehicle trajectory, and has shown the reduction of root mean squared error by 53.4 percent.**

***Index Terms*—Vehicle Trajectory Prediction, Long-Term Prediction, Error Minimization, LSTM, Intelligent Transport System**

I. Introduction

In the intelligent transportation system of modern society, due to the development of communication technology, a large amount of vehicle trajectory data can be collected [1]. However, predicting the long-term vehicle trajectory by leveraging the collected data requires further research in order to avoid the accumulation of errors [2]. Barrios et al. [3] has proposed a Kalman-Filter-based trajectory prediction model by using the Geographic Information System and Global Positioning System data. Alzyout et al. has proposed an ARIMA-based trajectory prediction method, which is using the location, speed and bearing information along with an enhanced model selection approach [4]. A. Ip et al. has proposed a LSTM-based trajectory prediction approach and then, proposed to divide a map into a 2D grid [5]. The aforementioned studies achieved a high accuracy on predicting short-term trajectories, but may not be applied to forecasting long-term sequences as they are since the prediction error accumulates as the sequence becomes longer.

In this paper, we propose a low-complex and robust trajectory prediction method. We propose to use a light-weight LSTM model to predict the future trajectory of a vehicle. Due to the simplicity of the proposed model, it can predict the future sequence in real-time. However, it may yield and accumulate prediction errors as the prediction sequence becomes longer. To iteratively minimize the prediction error, we propose to project the predicted trajectory onto the link geometry in an adaptive manner, with which the estimated vehicle location is placed on the real road which is important for some applications such as autonomous driving. The rest of the paper is organized as follows. Section 2 describes the proposed method in detail. Section 3 evaluates and validates the model proposed in the study. Finally, in Section 4, we discuss the conclusions and future research directions.

* Corresponding author

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. 2021R1F1A1059109).

Algorithm 1: Link Projection-Based Trajectory Prediction

input : $T := Trajectory, |T| = N$
input : $M :=$ number of samples to predict
input : $L :=$ Geometry (link) set
output : $P := Predicted\ Trajectory$

1 $P \leftarrow Queue(T)$;
2 **for** t *in* M **do**
3 $\quad p \leftarrow \texttt{LSTM}(P[t:t+N])$
4 $\quad l \leftarrow \texttt{getMinDistIndex}(L, p)$
5 $\quad P[t+1] \leftarrow \texttt{linkProjection}(l,p)$
6 **end**

II. Proposed Algorithm

The goal of the proposed algorithm is to minimize the prediction errors when forecasting a long sequence of vehicle trajectory. To achieve the goal, we propose to adaptively project the predicted sequence onto the link geometry. Also, we propose to use a light-weight LSTM model to reduce the inference time, making it applicable to real-time services. Due to the simplicity of the model, it may yield large errors, but the proposed projection method can effectively reduce such undesired deviations iteratively. The proposed algorithm is described in Algo. 1.

The proposed approach pre-processes the acquired data (i.e., GNSS trajectory of a vehicle), including the adjustment of the coordinate system. Then, it uses the trained LSTM model to predict the following future coordinate which can be deviated from the correct location (Line:3 in Algo. 1). The proposed adaptive adjustment algorithm (Line:4-5 in Algo. 1) searches for the closest link geometry onto which the LSTM prediction is to be projected. It can effectively reduce the prediction errors in the long run. The above process will be repeated

Parameter	Value
Learning Rate	0.0001
Input Size	2
Hidden Layer Size	256
Number of Hidden Layers	1
Epoch	2000
Optimizer	Adam
Window Size (N)	8

TABLE I
LSTM MODEL PARAMETERS

M times to predict M coordinates (Line:2 in Algo. 1). The proposed adjustment approach is called Bounding Box Link Projection. This is because when it searches for the closest link geometry it gradually increases the area of a virtual box centered around the LSTM prediction. Once the box touches a link, then the predicted coordinate will be projected on the link. This approach can prevent the accumulation of errors (i.e., deviations from the road) in long-term projections. In this study, we use a light-weight, low-complex LSTM model to achieve a real-time estimation of the future trajectory. The proposed method can be easily applied to other deep learning models for trajectory prediction.

III. Evaluation

A. LSTM Structure

In this study, we use the LSTM model, which is known to be effective for sequence prediction. The values of the hyper-parameters and its structure are shown in Table 1.

B. Dataset

For the dataset, we used the data acquired from Seoul Transportation Big Data [6]. In particular, only the location information (coordinate data) of vehicles was used. The collected data consists of coordinate data of vehicles collected during a certain day.

C. Result and Comparison

Both Fig. 1 and Fig. 2 show the comparison among the actual trajectory (black) and the predicted trajectories of the proposed model (green) and the vanilla LSTM (red). As it can be seen from Fig. 1 both the proposed and vanilla LSTM can effectively predict the trajectory. However, the proposed approach outperforms the vanilla LSTM to a large extent especially when the road has a complex structures such as intersections (see Fig. 2).

For the comparison between the original and predicted trajectory, the RMSE (Root-Mean-Squared-Error) method was used to measure the similarity between the two. The RMSE of proposed model is 17.1301, and that of the vanilla LSTM is 36.7747. As a result, the proposed model has enhanced the prediction accuracy by 53.4% compared to the vanilla LSTM. The reason for the performance enhancement was due to the adaptive link projection. It can iteratively places the predicted coordinate onto the actual road, and thus even when the sequence of prediction becomes longer, the proposed method can effectively minimize the prediction errors.

Fig. 1. Green dots are the predicted vehicle trajectory of the proposed model, red dots are predicted trajectory using the vanilla LSTM, and black dots are actual vehicle traveled trajectory.

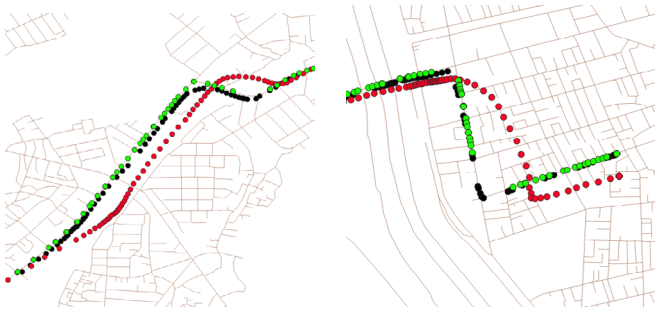


Fig. 2. Prediction results at intersections, where green, red and black dots are the proposed, vanilla LSTM and the actual trajectory result, respectively.

IV. Conclusion

In this paper, we have proposed a robust, light-weight vehicle trajectory prediction method. By projecting the estimated vehicle location onto the link geometry, the errors in the predicted long-term sequence can be effectively reduced even with a light-weight LSTM model. As a result, the predicted long-term trajectory can be much closer to the original dataset compared to the vanilla LSTM model.

References

[1] W. Wang et al. "Trust-enhanced collaborative filtering for personalized point of interests recommendation," in *IEEE Transactions on Industrial Informatics*, vol. 16, no. 9, pp. 6124-6132, Sept., 2019.

[2] F. Altché and A. de La Fortelle, "An LSTM network for highway trajectory prediction," in 2017 IEEE 20th International Conference on Intelligent Transportation Systems (ITSC), Yokohama, Japan, 2017, pp. 353-359.

[3] C. Barrios and Y. Motai, "Improving Estimation of Vehicle's Trajectory Using the Latest Global Positioning System With Kalman Filtering," in IEEE Transactions on Instrumentation and Measurement, vol. 60, no. 12, pp. 3747-3755.

[4] M. S. Alzyout and M. A. Alsmirat, "Performance of design options of automated ARIMA model construction for dynamic vehicle GPS location prediction." in Simulation Modelling Practice and Theory 104, vol. 104, pp. 102148. 2020.

[5] A. Ip et al. "Vehicle trajectory prediction based on LSTM recurrent neural networks," in *2021 IEEE 93rd Vehicular Technology Conference (VTC2021-Spring)*, pp. 1-5, April., 2021.

[6] "Vehicle Velocity and Location Inforamation Service, Seoul Transportation Big Data," https://t-data.seoul.go.kr/ (accessed Dec. 15, 2022)

MBTI-type Travel Plan Based on Emotional Personality

Hyunwoo Seo and Amsuk Oh
Department of Digital Media Engineering
University of Tongmyong
Busan 48520, Korea
shw1801@naver.com, asoh@tu.ac.kr

***Abstract*— Research aims to develop an emotional travel planning app using Android Studio. The app focuses on providing personalized travel plans by considering various factors such as user preferences, interests, emotions, time constraints, and budget limitations. Users can set their individual preferences and travel goals through the app, which then offers optimized travel routes and activity suggestions accordingly.**

Keywords—, Android Studio, Emotional travel planning app, Personalized travel planning

I. INTRODUCTION

Traveling is an important activity that provides new experiences and relaxation to many people. However, many travelers often face difficulties in planning their trips due to complex itineraries and numerous choices. To overcome these challenges and provide personalized travel plans considering user preferences and emotions, this research aims to develop an emotional travel planning app using Android Studio [1].

The objective of this research is to develop an emotional travel planning app using Android Studio that offers customized travel plans by considering user preferences, interests, emotions, time constraints, and budget limitations. This will enhance users' travel experiences and increase their satisfaction [2].

Existing travel planning apps mainly focus on scheduling and place recommendations, often neglecting the consideration of users' personal preferences and emotions. However, incorporating emotional aspects into travel planning can provide more personalized experiences and higher satisfaction. Therefore, users can enjoy more meaningful trips through customized travel plans that consider emotional elements [3-4].

II. EMOTIONAL TRAVEL PLANNING

A. Emotional Travel Planning App Design

In this research, an emotional travel planning app will be implemented using Android Studio. Android Studio is an integrated development environment (IDE) for developing Android applications, providing efficient tools for both app development and design. The app design process follows the following steps:

a. Requirements Analysis: User surveys and literature reviews are conducted to understand the desired features and requirements of travelers. Based on this, the key functionalities and their priorities for the app are determined.

b. UI/UX Design: Intuitive and user-friendly interfaces (UI) are designed to enhance the user experience (UX). The UI is designed to provide easy access to features such as travel itinerary creation, preference settings, and emotion analysis.

c. Database Design: A database is designed to store and manage information on travel destinations, activities, budgets, etc. The database structure is designed to efficiently store and retrieve user inputs and analyzed data.

d. App Development: The app development is carried out using Android Studio. The key functionalities of the app are implemented, including the algorithm for generating customized travel plans based on user inputs and data analysis. Integration with the database enables real-time retrieval of relevant information.

By designing the app using Android Studio, the research aims to optimize the core features of the emotional travel planning app and enhance the user experience by considering user preferences and emotions in generating personalized travel plans. Fig. 1 shows the design of emotional travel planning.

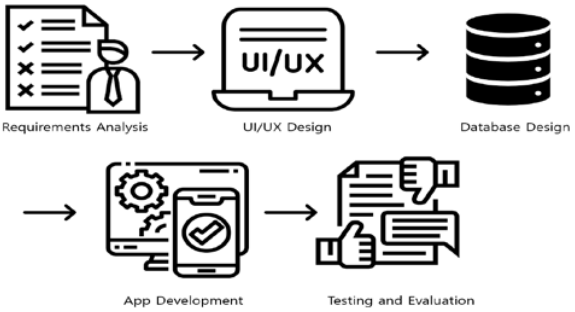


Fig. 1. Design of Emotional Travel Planning

B. Key Features of the Emotional Travel Planning App Using Android Studio

MBTI-based Travel Recommendations

One of the key features of the emotional travel planning app is the MBTI (Myers-Briggs Type Indicator)-based travel recommendations. MBTI is an indicator of the user's personality traits, and each MBTI type may have different travel preferences and interests. Therefore, the app collects the user's MBTI and provides personalized travel recommendations based on it. This allows users to select the most suitable travel destinations, activities, and itineraries according to their individual preferences.

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP(Institute for Information & communications Technology Promotion)(2018-0-018740301001).

TABLE I.

FEATURES OF MBTI-BASED TRAVEL RECOMMENDATIONS

Feature	Description
Personalized Recommendations	Provides personalized travel recommendations based on the user's MBTI. Takes into account the user's personality traits and preferences to recommend the optimal travel destinations, activities, and itinerary.
Diverse Choices	Offers a variety of travel destinations and activities, providing users with diverse options. Recognizes that travel preferences and interests may vary based on MBTI types, allowing users to make choices that best suit their preferences.
Individualized Experience	Tailors travel recommendations to each user's personality and preferences, offering an individualized travel experience. Enables users to enjoy trips that align with their own tastes and preferences.

UI Design Utilizing Data Visualization Techniques for UX Enhancement

Enhancing user experience (UX) is a crucial aspect of the emotional travel planning app. To improve UX, the app employs data visualization techniques to design user-friendly interfaces (UI). For instance, visual representation of travel destinations through images, reviews, ratings, etc., makes it easier for users to make choices. Additionally, presenting travel itineraries through graphs or charts helps users visually grasp information such as budget and time constraints. Through these efforts, users can easily adjust and review their travel plans using an intuitive and convenient UI.

TABLE II.

UX IMPROVEMENT USING DATA VISUALIZATION TECHNIQUES FOR UI DESIGN

UX Improvement	Data Visualization Techniques
Visualizing Destination Information	Displaying destination images, ratings, and reviews to provide a visual representation of the travel options.
Interactive Maps	Incorporating interactive maps with markers to showcase various travel destinations, allowing users to explore and select their preferred locations.
Graphs and Charts	Utilizing graphs and charts to represent travel itineraries, including budget breakdowns, time allocations, and activity schedules. This helps users easily understand and modify their travel plans.

The emotional travel planning app developed using Android Studio incorporates the core feature of MBTI-based travel recommendations and utilizes data visualization techniques to enhance user experience. By providing personalized travel recommendations based on MBTI and offering a user-friendly UI, the app enables users to enjoy more enriching travel experiences and efficiently plan their trips. Fig. 2 shows the key technologies for emotional travel planning.

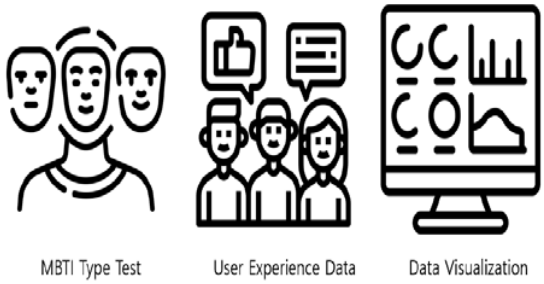


Fig. 2. Key Technologies for Emotional Travel Planning Apps

III. CONCLUSION

The objective of this study was to develop an emotional travel planning app using Android Studio. This app provides personalized travel recommendations based on the user's MBTI (Myers-Briggs Type Indicator) and enhances the user experience through the utilization of data visualization techniques. The core feature of the app is MBTI-based travel recommendations. By collecting the user's MBTI, the app suggests personalized travel destinations, activities, and itineraries to help users plan their trips according to their individual preferences. Additionally, the UI design incorporates data visualization techniques to improve user experience, allowing users to easily adjust and review their travel plans in an intuitive and convenient environment. The developed emotional travel planning app is expected to enhance user satisfaction and engagement by offering personalized travel experiences and improved travel planning. Furthermore, it provides practical insights into the utilization

ACKNOWLEDGMENT

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP(Institute for Information & communications Technology Promotion)(2018-0-018740301001).

REFERENCES

[1] Se, H., & Moon, J. (2020). Structural Relationships Among Service Characteristics, User Resistance, Attitude, and Satisfaction in Mobile Travel App: Focusing on Perceived Risk and Consumer Knowledge through Multi-Group Analysis. Tourism Management Studies, 96, 653-676.

[2] Kim Young-sook. "a study of leisure life according to MBTI personality type." studies of consumption culture 7.2 (2004): 83-106.

[3] Moon, Y., Kim, G. S., Jung, M. Y., & Kim, K. S. (2012). Consumer Emotional Responses Formed by Travel Agency Experiences. Journal of Venture and Entrepreneurship, 7(4), 33-40

[4] Chung Moon-young, Kim Gye-seok, Jung Moon Young, and Kim Gye Seok. " the emotional response of consumers formed by travel agency experience." venture start-up study 7.4 (2012): 33-40.

Coordinate Personal Schedules Using Group Calendars and Support Group Collaboration

Hanwon Lee and Amsuk Oh
Department of Computer Media Engineering
University of Tongmyong
Busan 48520, Korea
l8409@naver.com, asoh@tu.ac.kr

***Abstract*— Collaboration and work sharing among members are essential for smooth work progress, and in some cases, data is exchanged and shared through e-mail and messenger programs. However, alternatives to solving this problem are often found because data transmission methods, inquiries are inconvenient, and work history management is not well managed. Therefore, in this paper, we would like to study the schedule management application for group-type mutual cooperation and information sharing with these functions, and we would like to create a calendar, add events, and design and build an event list. In addition, a Firebase-based database is also intended to be used to build and store user information and group schedule contents.**

Keywords—, Group Calendar, Group Collaboration, Personal Schedule

I. INTRODUCTION

In modern society, collaboration with various groups is often required, but whenever a new schedule is planned in groups, it is difficult to coordinate each other's time and schedule [1]. In this situation, the functions of the group calendar are urgently needed to share the time to cooperate with each other so that individual schedules do not interfere with new plans[2]. Therefore, in this paper, we would like to study a group calendar application with these functions and produce it using Google Calendar API based on Android Studio. Detailed features use Firebase to create calendars, add events, design and build event lists, and store databases such as user information and group schedules [3].

II. PROBLEMS AND REQUIREMENTS

In collaboration between groups, data transmission and inquiry are inconvenient and task history management is difficult, so this paper proposes a schedule management plan for group-type mutual cooperation and information sharing to solve this problem. First, in order to activate the function of the group calendar more easily and conveniently, the characteristics of problems and requirements that may arise in a group-based schedule collaboration environment are important, and Table 1 shows these contents.

TABLE I.

PROBLEMS AND REQUIREMENTS OF GROUP-BASED SCHEDULE COORDINATION ENVIRONMENT

Field	Problem	Solution
User interface	Difficulty in coordinating complex schedules	Provide an intuitive and user-friendly interface
Schedule Conflict Check	Possibility of schedule conflicts within a group	Implement automated conflict detection and warning system
Team member communication	Lack of real-time communication between group members	Provides real-time chat or message functionality
Manage Priorities	Critical schedules may be ignored	Provides scheduling priorities and notifications
Time Zone Management	Difficulty in coordinating group members in different time zones	Include time zone conversion and schedule display features
Security and Privacy	Risk of exposing schedules and personal information	Enhanced security and privacy with user authentication with Firebase Authentication

III. COORDINATE PERSONAL SCHEDULES

The functions that enable collaboration of the group calendar are as follows.

① The ability to see group members' schedules at a glance by each group and user: The account name or ID is briefly displayed on the schedule registered by the user to share in the group calendar, providing convenience to the user by allowing a quick view of who is participating in the schedule.

② Ability to add and modify schedules: Members of a group can express their intention to participate or comment on a registered schedule, giving the user who registered the schedule the right to modify the details of the schedule in real time.

③ Notification of new or changed schedules to group members: When a group member registers a new schedule or changes the schedule, a push notification is sent to members who belong to the group or who decide to participate in the schedule, allowing faster and more accurate collaboration.

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP(Institute for Information & communications Technology Promotion)(2018-0-018740301001).

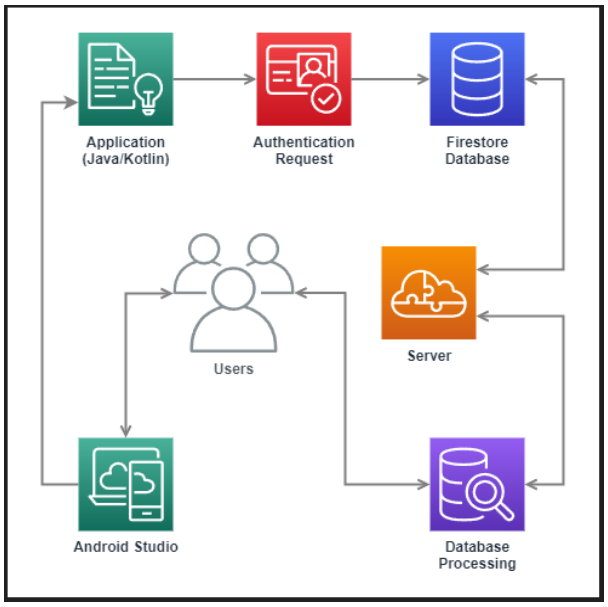


Fig. 1. Execution Process of Coordinate Personal Schedules

The Android application, developed using Java or Kotlin, allows users to interact with the platform. It incorporates authentication mechanisms to ensure secure access and user verification. Authentication is facilitated through communication with Firestore, a cloud-based NoSQL database provided by Firebase. Firestore acts as a server, handling data synchronization and communication with the backend. The server component, supported by appropriate backend technologies, processes requests from Firestore, managing tasks such as schedule integration, conflict resolution, and data storage. The database, under server management, stores and organizes user data, ensuring consistency and availability. Fig. 1 shows the execution process of coordinate personal schedules.

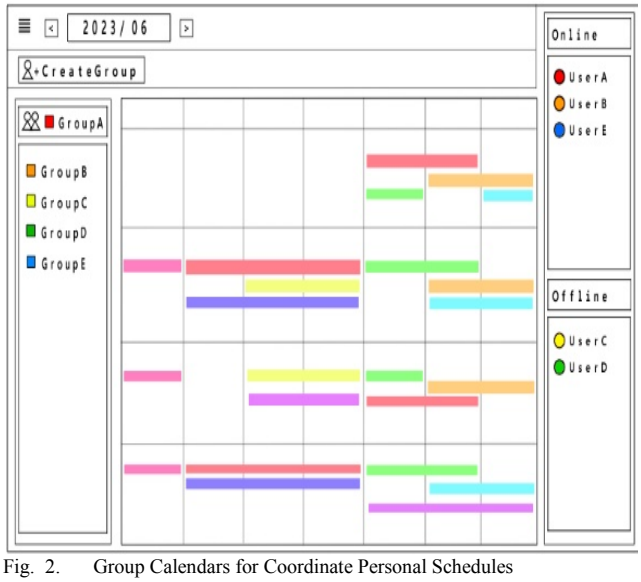


Fig. 2. Group Calendars for Coordinate Personal Schedules

For each group, you can set whether to publish a schedule and participate in events in multiple groups at the same time. This feature enhances communication between group members by checking the online and offline status of group calendar users in real time. A schedule displayed in a different color in the Calendar section indicates the schedule that group members are participating in. It provides scheduling priorities and notifications because many schedules can cause critical schedules to be ignored. Fig. 2 shows the group calendars for these coordinate personal schedules.

IV. CONCLUSION

This paper describes a schedule management application for group-type mutual cooperation and information sharing. When using the group creation and participation function of the group calendar, the group creation function and participation function of the group calendar register their work schedule in the calendar, and the individual's work schedule can be systematically managed and individual's work management ability is naturally improved. Since the schedule of members in the group can be identified at a glance in the group calendar, it saves time to check the schedule individually for each member, and checks the shared schedule in advance to prevent errors such as duplicate reservations. Since the group calendar can be updated in real time while sharing the schedule, it is possible to adjust the schedule by reflecting the update immediately whenever each schedule changes.

ACKNOWLEDGMENT

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW) supervised by the IITP(Institute for Information & communications Technology Promotion)(2018-0-018740301001).

REFERENCES

[1] Jae-Hwan Jin, Jung-Jeong Jo, Myung-Joon Lee, “G Calendar : Group Calendar Application Supporting Elgg Open Source Social Network Service”, Korean Institute of Information Scientists and Engineers a collection of academic presentations, pp. 1100-1102. June 2014.

[2] Sung-Ho Yu, "Study on the usability of UX designs in calendar apps - With a focus on the Google calendar app and Naver calendar app", Communication Design Study, vol 53, pp. 72-82, 2015.

[3] Jung-Won Yoon, "A Study on the Using Characteristics of Calendar : Mainly focused on scheduling strategy and sociality on time management", PROCEEDINGS OF HCI KOREA, pp.227-232, January 2016.

Poster Session

Data encryption survey on AI+X in public area

Jiae Lee
Urban Strategy Research Division
Seoul Institute of Technology
Seoul, Republic of Korea
jiaelee@sit.re.kr

Yonggang Kim*
Division of Computer Science and Engineering
Kongju National University
Cheonan, Republic of Korea
ygkim@kongju.ac.kr

***Abstract*—AI+X is a compound word of AI and convergence, and means a system that converges various industries based on AI. It enables the use of data collected by governments, local governments, and companies in various domains such as medicine, politics, and transportation. However, since this information includes personal information, confidentiality, integrity, and availability of data must all be guaranteed. Data encryption, a corporation for this purpose, is a method of converting plain text through cryptography for data confidentiality. AI+X seems to require data security so that it can be used more actively in various domains, and in this paper, domains and encryption technologies for this are investigated.**

***Index Terms*—Artificial Intelligence, Convergence, Security, Data, Encryption**

I. Introduction

Representative technologies to solve the security issue of shared data include homomorphic encryption, differential privacy, federated learning, and K-anonymity. Homomorphic encryption is when the result of analyzing the ciphertext is the same as that of the plaintext. Differential privacy is a technique that prevents inference of the original text through the query result statement. Federated learning is a method of learning in a distributed state of personal information. Each technology has its pros and cons. AI+X requires accurate and efficient calculations because domain services must be performed smoothly through a system that utilizes artificial intelligence.

II. AI+X and Data

Artificial intelligence imitates human intelligence and is developing machine learning, deep learning, and generative artificial intelligence as sub-concepts. Along with these technological developments, artificial intelligence is being applied in various industries and public fields such as medical care, transportation, welfare, and administration. Here, AI+X means convergence service using AI and data of each province. Data is actively used in the public and industrial areas. In the public sector, infrastructure and personal information are dealt with mainly in administration, government, medical care, welfare, safety, city, and transportation. Information such as personal information, physical, mental, social, property, communication, location, etc. belongs to the type of personal information and must be calculated without exposure. Such information can become an issue throughout society if confidential information is exposed or shared. Considering these points, data technology that can be flexibly analyzed using AI is needed. However, open data is limited in scope, and disclosure and sharing of actual data is limited. As a solution to this, various technologies are being used.

III. Data Security on Public Area

Homomorphic encryption is a method that can be analyzed without exposing it to the outside, and has the advantage that the result of calculating the encrypted data is the same as the result of analyzing the plain text. $(m_1, m_2 \in \mathcal{M} \rightleftarrows E(m_1), E(m_2) \in \mathcal{C})$

$$m_1 + m_2 \in \mathcal{M} \rightleftarrows E(m_1) + E(m_2) \in \mathcal{C} \tag{1}$$

Starting with cryptography, it was developed into fully homomorphic encryption(FHE), semi-homomorphic encryption(SHE), and partial homomorphic encryption(PHE). Among them, fully homomorphic encryption has no restrictions on the type and number of operations, and bootstrapping, squashing, approximate value calculation and search, and statistical processing are possible. Recently, if applied to big data in the public sector, machine learning for analysis, and information search, it can be efficiently used for public service development. It seems that it can be freely divided even in a cloud-based environment where complex calculations such as machine learning are supported. In particular, since public data has limitations in that it is difficult to utilize due to the openness of personal information and data and the low recognition rate in the security area, it is expected that the effectiveness will increase with gradual application.

Acknowledgment

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2022-00166739).

References

[1] Fontaine, Caroline, and Fabien Galand. "A survey of homomorphic encryption for nonspecialists." EURASIP Journal on Information Security 2007 (2007): 1-10.
[2] Cheon, Jung Hee, et al. "Homomorphic encryption for arithmetic of approximate numbers." Advances in Cryptology–ASIACRYPT 2017: 23rd International Conference on the Theory and Applications of Cryptology and Information Security, Hong Kong, China, December 3-7, 2017, Proceedings, Part I 23. Springer International Publishing, 2017.
[3] CRYPTOLAB, https://www.cryptolab.co.kr/
[4] Ministry of Government Legislation, https://www.laws.or.kr/

A Machine Learning Model Based on Complex Biological Signals for Fatigue Detection of Carrying Workers

Jeong-Su Kim
Chief Executive
EasyGeo Co.
Busan, Republic of Korea
js.kim@ezgeo.com

Dong Myung Lee
Department of Computer Engineering
Tongmyong University
Busan, Republic of Korea
dmlee@tu.ac.kr

Abstract—In this paper, a machine learning model that can determine the mental and physical fatigues based on electroencephalogram (EEG) and electromyogram (EMG) biological signals is proposed in order to reduce safety accidents caused by physical and psychological fatigue of carrying workers at industrial sites. The proposed model was confirmed to have high precision and reproducibility as the accuracy improved by 0.98 compared to other models and machine learning models through related studies.

Keywords-Complex Biological Signals, Fatigue Detection, Machine Learning, Carrying Worker, Industrial Safety Accident

I. INTRODUCTION

If the psychological fatigue of workers is high, the possibility of being insensitive to safety increases, and if the level of physical fatigue is high, the coping ability decreases due to the decline of physical and cognitive functions. In particular, there are many problems in determining fatigue using the existing driver fatigue evaluation model because the signal values of the brain wave and the muscle strength used in many cases appear differently depending on the work of the worker. Therefore, there is a need for an improved fatigue evaluation model that can increase the preventive effect of industrial accidents and safety accidents.

II. DESIGN OF MACHINE LEARNING MODEL BASD ON COMPLEX BIOLOGICAL SIGNALS

The overall composition and procedure of the machine learning model based on complex biological signals for fatigue detection of carrying workers is shown in Fig. 1. A single data set is constructed by preprocessing the electroencephalogram (EEG) and electromyogram (EMG) biological signals to match the characteristics of each signal.

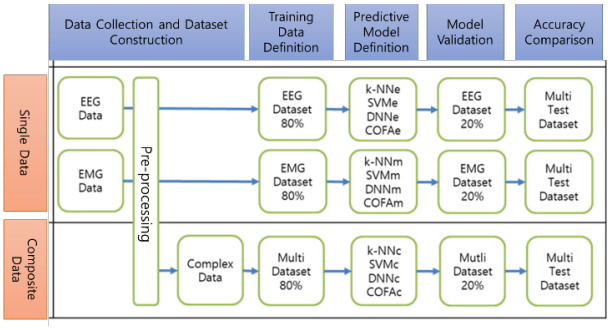


Fig. 1. A design of machine learning model based on complex biological signals.

Then, the pre-processed EEG and EMG single datasets are mixed and a composite dataset is created after the preprocessing process. The single dataset and composite dataset generated in this way are classified into a training dataset and a validation dataset at a ratio of 80:20, respectively [1-2].

These datasets are applied to K-nearest neighbor (k-NN), support vector machine (SVM), deep neural network (DNN), convolutional neural network (CNN), long short-term memory (LSTM) and proposed COFA model, respectively, and compare the discrimination accuracy of the fatigue prediction model [3]. Finally, the prediction model with the best performance is selected and applied to the practical fatigue detection system.

III. EXPERIMENTS AND RESULTS

The performance of the proposed model was compared to that of machine learning model such as k-NN, SVM, DNN, CNN and LSTM. The proposed model was confirmed to have high precision and reproducibility as the accuracy improved by 0.98 compared to other models and machine learning models through related studies. Therefore, the proposed complex dataset configuration and model can be used for highly accurate fatigue determination considering physical and mental fatigue, and this model will be expected to be useful in reducing industrial accidents and safety accidents.

IV. CONCLUSION

The proposed complex dataset configuration and proposed model can be used for highly accurate fatigue determination considering physical and mental fatigue, and this model will be expected to be useful in reducing industrial accidents and safety accidents. Although the proposed model shows relatively good performance for fatigue discrimination, it is necessary to apply a method to prevent overfitting to improve generalization performance.

ACKNOWLEDGMENT

This work was supported by a grant from Tongmyong University Innovated University Research Park (I-URP) funded by Busan Metropolitan City, Republic of Korea. (IURP2301)

REFERENCES

[1] Gao, Z., Wang, X., Yang, Y., Mu, C., Cai, Q., Dang, W., & Zuo, S., "EEG-Based Spatio-Temporal Convolutional Neural Network for Driver Fatigue Evaluation", IEEE Transactions on Neural Networks and Learning Systems, pp. 1–9, 2019.

[2] Wang, J., Sun, Y., & Sun, S., "Recognition of Muscle Fatigue Status Based on Improved Wavelet Threshold and CNN-SVM", IEEE Access, 8, pp. 207914–207922, 2020.

[3] Jeong-Su Kim and Dong Myung Lee, "A Study on the Fatigue Evaluation Platform using Bio-signal based on Artificial Neural Network," Proceedings of the 5th International Conference on Artificial Intelligence in Information and Communication (ICAIIC 2023), IEEE, KICS, pp. 030-032, 20-23, Feb. 2023.

A Combinational Approach to Robust Robot Localization in Dynamic Environments

Suhyeon Kang, Doyeon Kim, Heoncheol Lee*

Abstract— **This paper addresses the localization problem in dynamic environments. While the localization can be accurately conducted in static environments, it is difficult to acquire accurate localization results in dynamic environments because dynamic objects may cause. To resolve the problem, this paper proposes a combinational approach which adaptively combines the results of Monte Carlo localization, LiDAR (light detection and ranging) odometry, and visual odometry according to the number of dynamic objects. The proposed approach is implemented based on ROS (robot operating system) and tested in dynamic environments. Experimental results showed that the proposed approach is more accurate than other methods.**

I. INTRODUCTION

Robot localization predicts the robot's current position and is vital in areas like autonomous and mobile robots. Robots use sensors like camera [1], LiDAR (light detection and ranging), GPS (global positioning system), and IMU (inertial measurement unit) for localization in static environments. In addition, methods combining sensors [2], such as cameras and LiDAR, have been developed to enhance localization accuracy. However, these methods overlook dynamic objects, leading to subpar performance.

Therefore, we propose a new approach that uses a combination of localization results from three algorithms (Monte Carlo localization, LiDAR odometry, visual odometry) using camera and LiDAR for robust localization in dynamic environments. The three types of localization results are used in a contextualized combination depending on the number of dynamic objects recognized by yolov5. An average consensus filter is used for the combination of location information. In summary, this paper proposes an adaptive consensus filter (ACF) with the number of people recognized by yolov5 as a parameter for robust robot localization.

The contributions of this paper are as follows:

- An combinational approach considering dynamic number of objects has been proposed for accurate localization in dynamic environments.
- The proposed approach has been quantitatively compared with other methods and showed more accurate results than the others.

The remainder of this paper is organized as follows. In Section II, problems such as robot, sensor and system configuration are explained. Section III describes the proposed method. Section IV presents the evaluation results of the proposed method. Finally, Section V presents the conclusion.

II. PROBLEM DESCRIPTION

A. System configuration and limitations

In this paper, we address the problem of improving the localization accuracy in indoor dynamic environment. The overall system is implemented using ROS. Algorithms used for position estimation are Monte Carlo localization (MCL), LiDAR odometry, and visual odometry. Localization results from each algorithm is averaged by a consensus filter (CF) according to the number of people. The implementation environment consists of an indoor dynamic environment. Dynamic objects recognized by yolov5 are limited to people.

B. Robot and sensor models

In this paper, we experimented using a robot with six differential wheels. The mini-pc used to implement all the algorithms is inside the robot. The robot's position is represented by x, y coordinates and an angle θ in the global reference frame. Depth camera is used to perform visual odometry and yolov5. 2D LiDAR is used for the cartographer [3] and EMCL2 [4] algorithms.

III. PROPOSED METHOD

A. Overview of the proposed method

After each EMCL2, cartographer, and ORB SLAM2 [5] algorithm is executed, ACF node is executed. Object detection is then performed when yolov5 is executed. A consensus filter is adaptively applied based on the number of people recognized in yolov5.

B. Combinational approach

Combinatorial localization approach [6] that combines the localization results from each algorithm have been studied previously. In this paper, we go further than the existing methods and use an ACF approach for the localization results. There are currently many different ways to apply a consensus filter. The method used in this paper is the average consensus algorithm [7]. The average consensus algorithm is a popular distributed algorithm for computing the average of some values.

In this paper, we add an adaptive approach to the average consensus algorithm to improve the accuracy of localization. In the case of no or one person, we apply the consensus filter to the results of MCL and cartographer using LiDAR, which show relatively more accurate results than visual odometry. When there are 5 or more people, we only use the results of visual odometry, which is more accurate than LiDAR in dynamic situations. Finally, when there are at least 2 and no

S. Kang and H. Lee are with the Department of IT Convergence Engineering, Kumoh National Institute of Technology, Korea. (e-mail: {20236101, hclee}@kumoh.ac.kr)

D. Kim is with the School of Electronic Engineering, Kumoh National Institute of Technology, Korea. (e-mail: dodea@kumoh.ac.kr)

more than 4 people, the CF is applied to all algorithm results (MCL, LiDAR odometry, and visual odometry).

- MCL

MCL algorithm is one of the most widely used methods of robot position estimation using particle filter. Particle filters use particles that contain probabilities to indicate where the robot is. Among the various MCL algorithms, EMCL2 (MCL with expansion resetting) is used. EMCL2 uses pre-drawn 2D map to perform localization.

- LiDAR odometry

Cartographer is a 2D LiDAR odometry method developed by Google. Cartographer consists of two subsystems: local SLAM and global SLAM. The robot pre-draws a 2D map using the cartographer. These drawn 2D map is then applied to simulation tool and used for localization.

- Visual odometry

ORB SLAM2 is one of the representative visual odometry methods using a camera. The feature point is extracted using the ORB (oriented fast and rotated brief) extraction method. There are several build versions of the ORB SLAM2, build version used is RGB-D in this paper. To obtain more accurate robot position information, we adjusted the camera parameters and scale factor.

IV. Experimental Results

A. Environmental setups

All algorithms are implemented in ROS. An intel NUC 11 mini-pc is used to implement the algorithms. We obtained a dataset containing LiDAR and camera data to validate the performance. Error rate (ER, %) is expressed as the ratio of the actual coordinates minus the estimated coordinates divided by the actual coordinates.

B. Results of localization in dynamic environments

The dataset consists of a dynamic environment with randomly moving people, gathered personally. ACF dynamically applies the three algorithms as a consensus filter and calculates the average value. Table I presents quantitative results when yolov5 detects more than 5 people, with bold indicating the best result and underlined indicating the second best result. In crowded dynamic environments with over 5 people, only visual odometry results are utilized. ACF demonstrates the lowest error rate, matching visual odometry. Figure 1 in rviz (ROS visualization tool) shows the localization results: yellow (ACF, proposed.), red (EMCL2), blue (cartographer), and green (ORB SLAM2). Overall, ACF exhibits the closest coordinates to the ground truth compared to CF, EMCL2, cartographer, and ORB SLAM2.

TABLE I. Comparison of Errors with More than Five People

-	x	y	θ
ACF (Proposed.)	**0.00303**	**0.00105**	**0.0052**
CF	0.012	0.00431	0.0122
EMCL2	0.015	0.019	0.0233
Cartographer	0.024	0.0327	0.00809
ORB SLAM2	**0.00303**	**0.00105**	**0.0052**

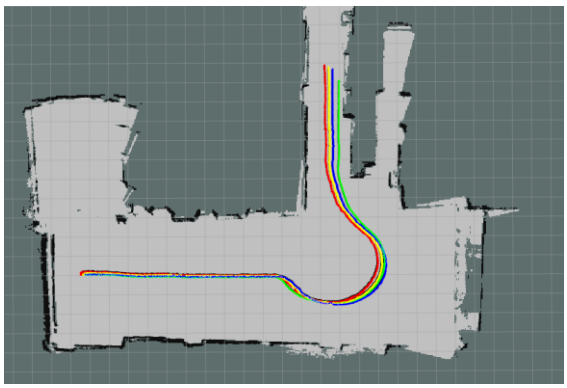


Figure 1. Results of each algorithm's localization in rviz

V. Conclusion

In this paper, we proposed a combinational approach for robust robot localization in dynamic indoor environments. Experimental results in dynamic environments show that the ACF algorithm proposed in this paper has the lowest error. In future research, we plan to refine the proposed algorithms by removing dynamic objects detected by LiDAR or excluding localization results that significantly deviate from neighboring nodes.

Acknowledgment

This work was supported in part by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2021R1F1A1064358), and in part by the Government-wide R&D Fund for Infections Disease Research (GFID), funded by the Ministry of the Interior and Safety, Republic of Korea (grant number : 20014854).

References

[1] C. Jennings, D. Murray and J. J. Little, "Cooperative robot localization with vision-based mapping," *Proceedings 1999 IEEE International Conference on Robotics and Automation (Cat. No.99CH36288C)*, Detroit, MI, USA, 1999, pp. 2659-2665 vol.4.

[2] P. Zingaretti and E. Frontoni, "Vision and sonar sensor fusion for mobile robot localization in aliased environments," *2006 2nd IEEE/ASME International Conference on Mechatronics and Embedded Systems and Applications*, Beijing, China, 2006, pp. 1-6.

[3] Y. Huang, G. Wu and Y. Zuo, "Research on slam improvement method based on cartographer," *9th International Symposium on Test Automation & Instrumentation (ISTAI 2022)*, Online Conference, Beijing, China, 2022, pp. 349-354.

[4] R. Ueda, T. Arai, K. Sakamoto, T. Kikuchi and S. Kamiya, "Expansion resetting for recovery from fatal error in Monte Carlo localization - comparison with sensor resetting methods," *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, Sendai, Japan, 2004, pp. 2481-2486 vol.3.

[5] Raúl Mur-Artal and Juan D. Tardós. ORB-SLAM2: "an Open-Source SLAM System for Monocular, Stereo and RGB-D Cameras." *IEEE Transactions on Robotics,* Vol. 33, No. 5, 2017, pp. 1255~1262.

[6] D. kim, H. Lee. "Combinatorial Position Estimation Techniques for Robotic Autonomy in Dynamic Indoor Environments" *Artificial Intelligence Conference,* 2022.

[7] R. Olfati-Saber, J. A. Fax and R. M. Murray, "Consensus and cooperation in networked multi-agent systems", *Proc. IEEE*, vol. 95, no. 1, pp. 215-233, Jan. 2007.

High-Quality Foley Sound Synthesis using Monte Carlo Dropout

Chae-Woon Bang[1], Nam Kyun Kim[2], and Chanjun Chun[1, *]
[1]*Dept. of Computer Engineering, Chosun University*
Pilmun-daero 309, Dong-gu Gwangju 61452, South Korea,
[2]*Dept. of Automotive Electronics R&D center, Korea Automotive Technology Institute*
Jingoksandanjungang-ro 55, Gwangsan-gu, Gwangju 62465, South Korea,

bcw4045@chosun.ac.kr, kimnk@katech.re.kr, cjchun@chosun.ac.kr

***Abstract*—This paper describes the foley sound synthesis system. Thus, it aims to creates foley sound, which is widely utilized as various sound effects in multimedia contents. To accomplish this, it uses sound synthesis technique, generating a 4-second audio clip of one of seven classes. Specifically, we fine-tuned the baseline model provided by DCASE 2023 such that improves the performance. After that, we ensemble the models using Monte Carlo Dropout. The performance of the proposed system was compared with the baseline using Frechet Audio Distance(FAD), which is referred as an audio evaluation metric. As a result, it was confirmed that both the single model and the ensemble model outperform the existing baseline system.**

Keywords— Foley sound synthesis, Monte Carlo Dropout, model ensemble

I. INTRODUCTION

Foley sound refers to sound effects generated by events occurring in radio or movies. This foley sound is employed to add various sound effects in multimedia contents. The conventional foley sound synthesis was manually recorded and mixed by foley artists. However, recently, with advancements in generative models, research is being conducted to utilize sound synthesis techniques to generate foley sounds [1].

DCASE opens various challenges based on audio data. Among them, Task7 is to utilize sound synthesis technology to generate foley sounds. It consists of seven classes and generates sounds of 4 seconds in length.

In this paper, we propose suitable hyperparameters based on the baseline provided by DCASE to perform high-quality sound generation. Furthermore, we suggest an ensemble system using Monte Carlo Dropout [2]. In other words, we fine-tuned the base-line model very sensitively to improve model behavior. Moreover, we utilized the Monte Carlo Dropout technique to facilitate ensemble training for models with long training times.

This paper is structured as follows. Section 2 describes the model structure and training method of the proposed system. Section 3 describes the performance comparison of the baseline provided by DCASE and our proposed model. Finally, Section 4 describes the conclusion of this technical report.

II. PROPOSED METHOD

A. Dataset

We used the dataset provided by DCASE. The data provided by DCASE consists of a total of 4,850 sounds divided into seven classes (*DogBark, Footstep, Gunshot, Keyboard, MovingMotorVehicle, Rain, Sneeze/Cough*). The dataset was collected from UrbanSound8K, FSD50K, and BBC Sound Effects1, and seven classes were selected considering urban sound taxonomy [3]. All audio files have been converted to mono 16-bit format and have a sampling rate of 22,050Hz. In addition, the length of each sound is four seconds, and the number of samples in each class is different. In this study, a window size of 1024 and a hop length of 256 were set to extract 80-dimensional mel-spectrograms.

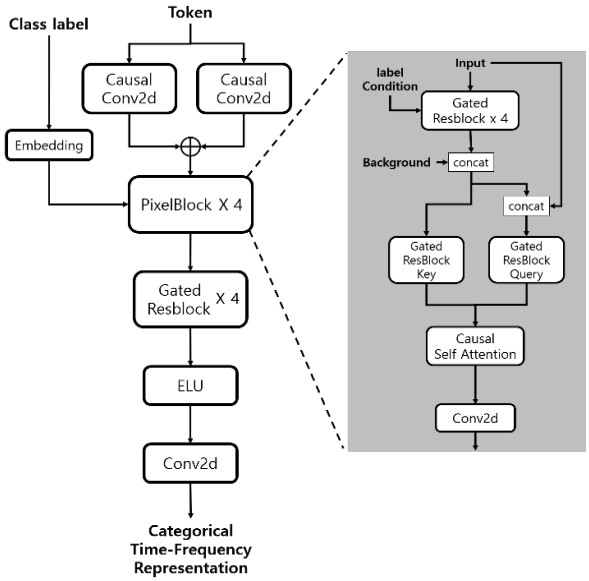


Fig. 1. Overall architecture of PixelSNAIL

B. Model Architecture

The baseline system consists of a total of 3 modules: One of them includes PixelSNAIL, which is known as a generative model that combines causal convolution and self-attention mechanism to generate high-quality distributions [4]-[6]. PixelSNAIL takes the class label of the sound to be generated as input and generated a discrete time-frequency

representation(DTFR). The following module is a VQ-VAE model that performs effective representation learning through vector quantization [7]. We use a trained VQ-VAE to acquire the DTFR generated from a mel-spectrogram [6]. In order to reconstruct a time-domain audio signal from the log mel-spectrogram, The HiFi-GAN, which is widely known as a high-performance neural Vocoder, was utilized [8].

Figure 1 shows the model structure of PixelSNAIL [6]. Here, The token is a vector filled with all zeros in the same shape as the DTFR in VQ-VAE. The PixelBlock consists of a combination of gated residual block and causal self-attention mechanism. The gated residual block regulates the information flow between layers using a gated activation unit and residual block [6]. It allows for controlled communication of information between different layers. The causal self-attention is utilized to extract crucial information by capturing the relationships and dependencies among the elements of the data. In other words, it helps capture important dependencies and patterns in the data by considering the causal relationships within the sequence [9].

The convolutional encoder used in VQ-VAE captures context-related acoustic information at various scales through a multi-scale convolution layer and then converts a discrete T-F representation(DTFR). The convolutional decoder is responsible for reconstructing the extracted time-frequency representation into a mel-spectrogram. The structure of the decoder uses a structure similar to that of the encoder, but the only difference is that it does not use a multi-scale convolution layer [7].

C. Training Method

Adam was used as an optimizer for model training. By setting the Cycle Scheduler, the learning rate is increased up to 3e-4, and the learning rate is adjusted periodically through the cosine period function [10].

Moreover, PixelSNAIL used 4 PixelBlocks, and Dropout was set to 0.1. In order to an appropriate model for PixelSNAIL, we performed several training with different channel information.

TABLE I. TEST RESULTS FOR EACH MODEL

Model	Frechet Audio Distance
Baseline Model	10.665
Fine-Tuned Model	9.439
Ensemble Model (N=2)	9.213
Ensemble Model (N=5)	9.160

III. EXPERIMENTS

In order to demonstrate the effectiveness of our foley sound system, the objective evaluation was conducted using Frechet Audio Distance(FAD), which is widely employed as an audio evaluation metric [11]. Note that there was no overlap between the training and evaluation data. For each class, we utilized approximately 50 samples for evaluation.

Table 1 shows the FAD result for each model. The baseline model in Table 1 indicates 256-channel PixelBlock, while the fine-tuned model implies 512-channel PixelBlock. To ensemble several models, we utilized the Monte Carlo Dropout, where N represents the number of inferencing iterations. The experimental result indicates that our fine-tuned model and ensemble model outperform the baseline model. Among them, the ensemble model with 5 inferencing iterations showed the highest performance.

IV. CONCLUSION

In this paper, we attempted to fine-tune the baseline model for high quality sound synthesis, and ensembled the systems via Monte Carlo Dropout. The results of the evaluation through the validation dataset show that the single model and the ensemble model have high performance compared to the baseline model.

ACKNOWLEDGMENT

This work was supported by the "Science and Technology Project Opens the Future of the Region" program of Innopolis Founda-tion funded by Ministry of Science and ICT(2022-DD-UP-0312).

REFERENCES

[1] K. Choi, J. Im, L. Heller, B. McFee, K. Imoto, Y. Okamoto, M. Lagrange, and S. Takamichi, "Foley sound synthesis at the DCASE 2023 challenge," *arXiv: 2304.12521*, Apr. 2023.

[2] Y. Gal and Z. Ghahramani, "Dropout as a bayesian approximation: Representing model uncertainty in deep learning," in *Proc. International Conference on Machine Learning*(*ICML*), pp. 1050–1059, June 2016.

[3] J. Salamon, C. Jacoby, and J. P. Bello, "A dataset and taxonomy for urban sound research," in *Proc. the 22nd ACM International Conference on Multimedia,* pp. 1041–1044, Nov. 2014.

[4] A. V. D. Oord, N. Kalchbrenner, L. Espeholt, O. Vinyals, A. Graves, and K. Kavukcuoglu, "Conditional image generation with PixelCNN decoders," in *Proc. Advances in Neural Information Processing Systems*(*NeurIPS*), pp. 4790–4798, June 2016.

[5] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, and A. N. Gomez, "Attention is all you need," in *Proc. Advances in Neural Information Processing Systems(NeurIPS)*, vol. 30, pp. 6000-6010, Dec. 2017.

[6] X. Chen, N. Mishra, M. Rohaninejad, and P. Abbeel, "PixelSNAIL: An improved autoregressive generative model," in *Proc. International Conference on Machine Learning (ICML)*, pp. 864–872, June 2018.

[7] A. V. D. Oord, O. Vinyals, and K. Kavukcuoglu, "Neural discrete representation learning," in *Proc. Advances in Neural Information Processing Systems(NeurIPS)*, vol. 30, pp. 6306–6315, Dec. 2017.

[8] J. Kong, J. Kim, and J. Bae, "HiFi-GAN: Generative adversarial networks for efficient and high fidelity speech synthesis," in *Proc. Advances in Neural Information Processing Systems(NeurIPS)*, vol. 33, pp. 17022–17033, Oct. 2020.

[9] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conference on Computer Vision and Pattern Recognition(CVPR),* pp. 770-778, June 2016.

[10] L. N. Smith, "Cyclical learning rates for training neural networks," in Proc. IEEE Winter Conference on Applications *of Computer Vision(WACV)*, pp. 464–472, Apr. 2017.

[11] K. Kilgour, M. Zuluaga, D. Roblek, and M. Sharifi, "Fréchet audio distance: A metric for evaluating music enhancement algorithms," *arXiv: 1812.08466*, Dec. 2018.

Predicting Power Consumption in an Industrial Complex in Gumi using XGBoost

Eun Jeong Son
ICT Convergence Research Center
Kumoh national Institute of Technology
Gumi, Republic of Korea
nadojo1227@kumoh.ac.kr

Jae Gwang Ahn
ICT Convergence Research Center
Kumoh national Institute of Technology
Gumi, Republic of Korea
jkan17@kumoh.ac.kr

Jae-Min Lee
Dept. of IT Convergence Engineering
Kumoh national Institute of Technology
Gumi, Republic of Korea
ljmpaul@kumoh.ac.kr

Dong-Seong Kim*
Dept. of IT Convergence Engineering
Kumoh national Institute of Technology
Gumi, Republic of Korea
dskim@kumoh.ac.kr

***Abstract*—Recently, eco-friendly technologies have been emerging to address global warming and pollution. As a result, smart energy platforms (SEPs) have been implemented in industrial complexes in Korea. This study focuses on predicting power consumption in an industrial complex located in Gumi, South Korea, using the XGBoost machine learning algorithm. Accurate forecasting of power consumption is crucial for efficient energy management and resource allocation in industrial settings. The proposed approach utilizes historical power consumption data collected by the SEP to develop a predictive model. XGBoost, a powerful gradient-boosting algorithm, is employed to capture complex relationships and patterns within the data. The trained model demonstrates the ability to predict the trend of power consumption, enabling proactive decision-making for energy optimization and cost reduction. The findings of this study contribute to enhancing energy efficiency and sustainability in industrial complexes, providing valuable insights for energy managers and policymakers.**

***Index Terms*—XGBoost, Prediction, Regression, Energy Consumption, Industrial Complex, Smart Energy Platform**

I. Introduction

Recently, eco-friendly technologies have been emerging to address global warming and pollution. As a result, smart energy platforms (SEPs) have been implemented in industrial complexes in Korea [1], [2]. The industrial sector is a significant consumer of energy, and accurate prediction of power consumption plays a crucial role in effective energy management and resource allocation. In recent years, advancements in machine learning techniques have provided new opportunities for developing accurate predictive models in various domains. In this study, we focus on predicting power consumption in an industrial complex located in Gumi, South Korea, leveraging the eXTreme Gradient Boosting (XGBoost) machine learning algorithm [3].

Efficient energy management in industrial complexes requires proactive decision-making based on accurate power consumption forecasts [4]–[6]. By accurately predicting power consumption, energy managers can optimize energy usage, improve cost-effectiveness, and reduce environmental impact. Traditional forecasting methods often rely on simple statistical techniques that may not capture the complex relationships and patterns present in the data. In contrast, machine learning algorithms offer the potential to capture nonlinear relationships and make more accurate predictions.

To address this challenge, we propose a predictive modeling approach that utilizes historical power consumption data, which is collected by the SEP in Gumi, to train an XGBoost model. XGBoost, a powerful gradient boosting algorithm, has shown excellent performance in various prediction tasks. It can effectively capture complex interactions between input variables and produce accurate predictions. By leveraging historical power consumption data and the capabilities of XGBoost, we aim to develop a predictive model that accurately forecasts power consumption trends in the industrial complex.

The remainder of this paper is organized as follows. Section 2 presents the methodology, including data collection, preprocessing, and the XGBoost modeling approach. Section 3 discusses the experimental setup and presents the results of the predictive model. Finally, Section 4 concludes the paper with a summary of the findings and potential avenues for future research.

II. Proposed Method

To predict the energy consumption of the industrial complex in Gumi, historical data from three randomly chosen factories, collected by the Smart Energy Platform (SEP) as shown in Figure 1, is used as the dataset. However, the available historical data is limited to the year 2020. Therefore, we restricted the training data to the period from January to November and used the data from December for testing. The data collected by SEP is provided in hourly intervals and measured in kilowatt-hours (kWh). Hence, we performed preprocessing on the data to prepare it for training the XGBoost model. The input features include hours, day of the week, quarter, month, day of the year, day of the month, and week of the year. After preprocessing the data, separate datasets were created for training and testing.

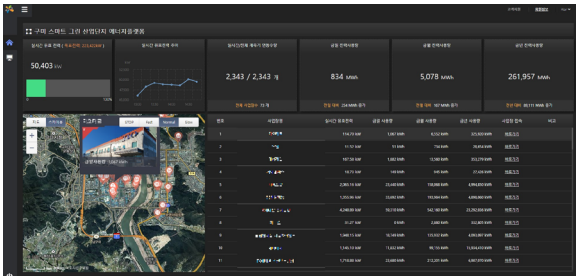


Fig. 1. Smart Energy Platform monitoring display

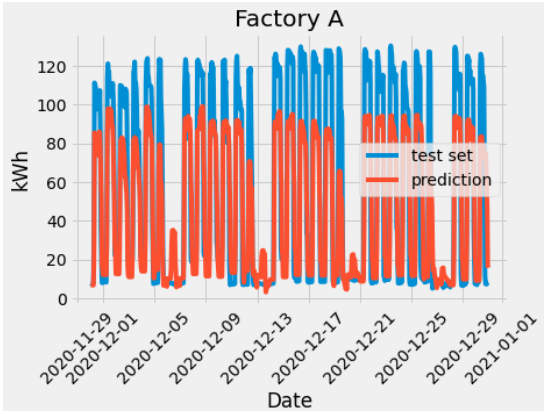


Fig. 2. Prediction result1

Using these datasets, the XGBoost model was trained and employed for predicting energy consumption.

III. Evaluation Results

This section presents the prediction results obtained from the trained XGBoost model. As described in Section II, the training dataset covers the period from January to November 2020, while the test dataset covers the period of December 2020. Figures 2, 3, and 4 depict the comparisons between the test dataset and the prediction results generated by the XGBoost model. As demonstrated in Figures 2, 3, and 4, the predictions effectively capture the trend of future energy consumption.

IV. Conclusion

In this study, we focused on predicting power consumption in an industrial complex located in Gumi, South Korea, using the XGBoost machine learning algorithm. Accurate forecasting of power consumption is crucial for efficient energy management and resource allocation in industrial settings. By leveraging historical power consumption data collected by the Smart Energy Platform (SEP) and employing the powerful XGBoost algorithm, we developed a predictive model that effectively captured the complex relationships and patterns within the data.

Acknowledgment

This work was supported in part by the Energy Self-Sustaining Infrastructure Development Project, funded by Korea Industrial Complex Corporation (KICOX), Korea, in part by the Priority Research Centers Program under Grant 2018R1A6A1A03024003, and in part by the National Research Foundation of Korea (NRF) and the Center for Women in Science, Engineering and Technology (WISET) Grant WISET-2023-040 funded by the Ministry of Science and ICT (MSIT) under the Program for Returners into R&D.

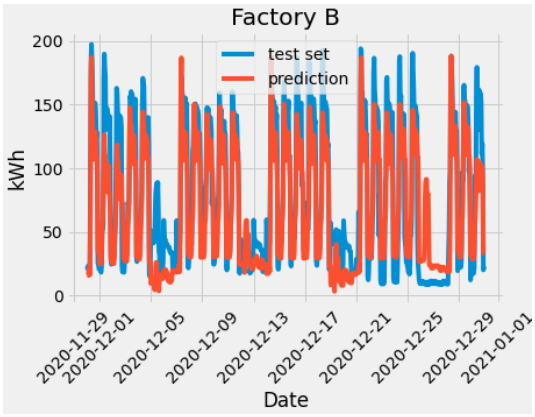


Fig. 3. Prediction result2

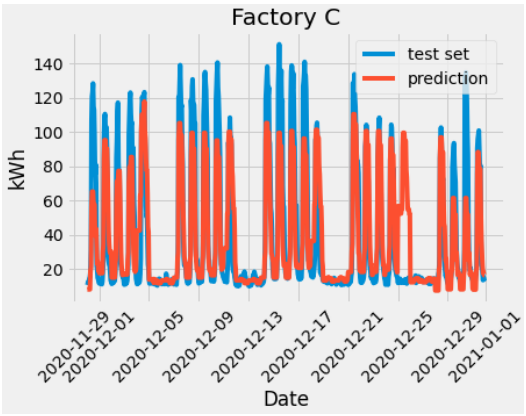


Fig. 4. Prediction result3

References

[1] Kerdsuwan, Somrat, Krongkaew Laohalidanond, and Woranuch Jangsawang. "Sustainable development and eco-friendly waste disposal technology for the local community." Energy Procedia 79 (2015): 119-124.

[2] J. W. Kim, J. M. Lee, D. -S. Kim, W. J. Ryu, G. Yang and Y. Kim, "Design and Implementation of Smart Energy Platform for Industrial Complex," 2021 International Conference on Information and Communication Technology Convergence (ICTC), Jeju Island, Korea, Republic of, 2021, pp. 1418-1420, doi: 10.1109/ICTC52510.2021.9621178.

[3] Chen, Tianqi, et al. "Xgboost: extreme gradient boosting." R package version 0.4-2 1.4 (2015): 1-4.

[4] Abdukhakimov, Asatilla, et al. "Reliability analysis in smart grid networks considering distributed energy resources and storage devices." International Journal of Electrical and Electronic Engineering & Telecommunications 8.5 (2019): 233-237.

[5] Landsiedel, Olaf, Klaus Wehrle, and Stefan Gotz. "Accurate prediction of power consumption in sensor networks." The Second IEEE Workshop on Embedded Networked Sensors, 2005. EmNetS-II.. IEEE, 2005.

[6] Abdukhakimov, Asatilla, Kumar Nalinaksh, and Dong-Seong Kim. "A Reliability Perspective of Distribution Systems in Smart Grid Communication Networks." J. Commun. 14.10 (2019): 926-935.

IEEE 802.11be Wi-Fi 7: A Study on Multi-Access Point (AP) Coordination Schemes

Dabin Kim
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
k_siwoo@yonsei.ac.kr

Wonsik Yang
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
yangws95@yonsei.ac.kr

Dongjun Jung
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
dongjunjun@yonsei.ac.kr

Minsoo Joo
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jooms91@yonsei.ac.kr

Jong-Moon Chung
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jmc@yonsei.ac.kr

***Abstract*— The demand for high throughput, low-latency, and fast response times from applications such as augmented reality (AR), cloud gaming, virtual reality (VR), 4K/8K videos, extended reality (XR), and others is rapidly increasing. These requirements exceed the performance bounds of the current standards, IEEE 802.11ax, and necessitate a high performing wireless local area network (WLAN) standard. To meet this demand, task group BE (TGbe) established a new standard known as Extremely High Throughput (EHT), also referred to as IEEE 802.11be (Wi-Fi 7), which can provide a throughput of over 30 Gbps. EHT has achieved significant performance improvements in wireless networks through new technologies at the lower layers, and these improvements are expected to make a significant contribution that can improve the performance of aerial and maritime wireless communications. In maritime environments, there are various access points (APs) such as ships, docks, and ports, where it is essential to minimize the interference between these APs to enhance the communication efficiency. Accordingly, this paper provides a comprehensive overview of the multiple access point (multi-AP) transmission technology among various technologies supported by Wi-Fi 7.**

Keywords— EHT, Wi-Fi 7, 802.11be, multi-AP coordination, transmission, WLAN.

I. INTRODUCTION

IEEE 802.11 (Wi-Fi) is a standard specification for wireless local area networks (WLAN) that are utilized by over 18 billion Wi-Fi devices worldwide [1]. It is widely adopted in various fields, including personal devices and industrial IoT sensors. The advancement of WLAN technology has led to the exploration of IEEE 802.11be (EHT), which focuses on enhancing medium access control (MAC) layers and the physical (PHY). This aims to improve Quality of Service (QoS) by increasing data transfer rates and decreasing network latency. Furthermore, the next-generation IEEE 802.11 standard will be able to meet the communication needs between various ships, ports, and control facilities through high throughput, low latency, and minimal interference in maritime wireless communication environments. EHT provides WLAN throughput five times larger than Wi-Fi 6 using various technologies, such as, up to 320 MHz bandwidth, increasing modulation speed through 4096-QAM, and increasing the number of spatial streams through MU-MIMO. Especially, EHT supports multi-AP coordination, which optimizes channel selection and improves the WLAN performance through various AP solutions. In [2], the transmission techniques of multi-AP coordination are summarized as joint transmission (JXT) and coordinated beamforming (CBF) at the PHY layer and coordinated orthogonal frequency-division multiple access (Co-OFDMA) and coordinated spatial reuse (CSR) at the MAC layer. In addition, various methods have been proposed to implement efficient multi-AP coordination to mitigate interference and improve system performance, including methods that combine techniques such as frequency/time allocation and spatial reuse, while others seek to address complexity through machine learning.

II. MULTI-AP COORDINATION

In this section, an overview of the schemes related to multi-AP coordination that is supported by 802.11be. When demand increases in a close wireless network environment, channel interference inevitably occurs, resulting in poor resource utilization. To overcome this, TGbe proposes to introduce multi-AP coordination that can optimize channel selection and perform efficient communication by adjusting load through close cooperation between nearby APs. First, in order to coordinate between independent APs, channel state information (CSI) is first provided from the STA to AP through multi-AP channel sounding, and after performing this procedure, multi-AP transmission can start. The multiple-AP systems considered by TGbe are coordinated and jointed in two ways [3], and four schemes are introduced in the subsection below.

A. Coordinated spatial reuse (CSR)

The previous standard, 802.11ax, does not perform multiple coordination between APs, so some APs that coordinate their transmit power will have a lower SINR. Therefore, EHT proposes a coordination scheme that minimizes interference through CSR, thus improving resource utilization and saving power. To improve this uncoordinated approach, EHT minimizes interference by coordinating the transmit power between APs for proper SINR of all stations (STAs) in a multi-AP Wi-Fi network as shown in Fig. 1 (a). The CSR scheme can be adjusted with less feedback between APs and is more resistant to interference than SR in 802.11ax.

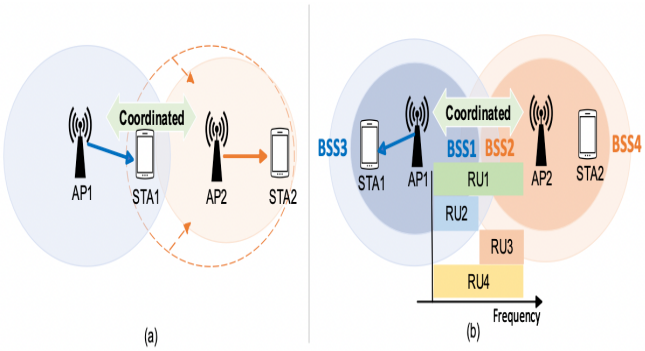


Fig. 1. Wi-Fi network example of (a) CSR and (b) C-OFDMA.

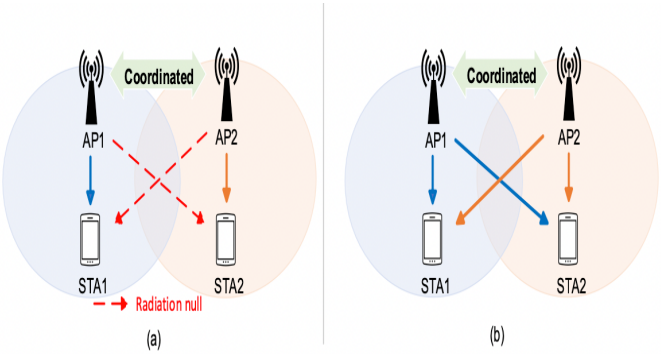


Fig. 2. Wi-Fi network example of (a) CBF and (b) JTX (D-MIMO).

B. Coordinated-OFDMA (C-OFDMA)

OFDMA is a technology that allows multiple access by dividing the available bandwidth into resource unit (RU) and allocating the available RUs to each STA. EHT extends this capability to multi-AP, allowing APs to cooperatively coordinate their schedules in the domains of time and frequency. In C-OFDMA, APs can allocate the same or different RU to share all OFDMA resources and coordinate with each other to avoid interference, as shown in Fig. 1 (b). This allows APs to jointly allocate resources to transmit data simultaneously, minimizing collisions between basic service sets (BSSs) and enabling more efficient use of spectrum.

Due to their simplicity and flexibility, CSR and C-OFDMA schemes are considered to be the most likely schemes to be supported in the new revision [4].

C. Coordinated-Beamforming (CBF)

CBF (or Null steering) is a scheme that suppresses interference by using spatial radiation nulls toward STAs that are not served by the AP, as shown in Fig. 2 (a). In current WLANs, inter-AP interference cannot be controlled because no coordination is performed. To implement this, the CBF checks the CSI of neighboring STAs in order to ignore signals for STAs that are not in service. Therefore, CBF does not perform coordination through data exchange between APs, and the scheme demonstrates a superior concurrency performance compared to single AP and CSR [5].

D. Joint Transmission (JTX)

JTX (or D-MIMO) is a scheme that supports multiple APs to transmit data simultaneously, reducing data transmission collisions and improving the reliability and bandwidth. Although JTX shows high levels of transmit and receive gains compared to other coordination schemes, JTX has complexity issues, such as, accurate time/phase synchronization and high-speed backhaul, to transmit data from multiple APs simultaneously. However, with proper backhauling and synchronization, this is a scheme that can increase the frame rate through spatial diversity [6].

III. Coordination Open Issue

There are various open issues for coordination between APs in real Wi-Fi networks. First, coordination in multi-AP networks requires significant signal overhead and processing complexity in the process of communicating with neighboring APs and acquiring CSI in the process of channel sounding. There is also the question of whether coordination between different networks will be possible given the heterogeneous environment of APs produced by different vendors in real-world environments. Next, coordination methods like JXT are complex because strict synchronization between APs can be affected by a variety of factors in the real world. Finally, the need for advanced scheduling techniques also exists because of the multiple information exchanges that occur to make this coordination happen [4].

IV. Conclusion

This paper summarizes several technologies for multi-AP coordination, which is one of the key features of next-generation Wi-Fi 7. Based on the introduced schemes, many studies are being conducted to optimize and utilize their performance in various ways. In future works, there is still research that needs to be done on various optimization schemes and new implementations to overcome various open issues that need to be addressed, which will contribute to an improved performance in Wi-Fi networks.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2021-0-00040, Development of intelligent stealth technology for information and communication resources for public affairs and missions).

References

[1] (2022). *Wi-Fi Alliance Wi-Fi mometum in 2022. [Online]*. Available : https://www.wi-fi.org/beacon/the-beacon/wi-fi-momentum-in-2022

[2] C. Deng, X.Fang, X.Han, X.Wang, L.Yan, R.He, Y.Long, and Y.Guo , "IEEE 802.11be Wi-Fi 7: New challenges and opportuniities ," *IEEE Commun. Surveys Tuts*., vol. 22, no. 4, pp. 2136–2166, 4th. Quart., 2020.

[3] S. Vermani, B. Tian, and J. Liu, *Terminology for AP Coordination, document IEEE 802.11-18/1926r2*, Nov. 2018.

[4] E. Khorov, I. Levitsky, and I. F. Akyildiz, "Current status and directions of IEEE 802.11be, the future Wi-Fi 7," *IEEE Access*, vol. 8, pp. 88664–88688, May. 2020.

[5] R. Doostnejad, *Multi-AP Collaborative BF in IEEE 802.11, document IEEE 802.11*, IEEE, Piscataway, NJ, USA, May 2019. [Online].Available: https://mentor.ieee.org/802.11/documents?is_dcn= 0772&is_group=0eht

[6] T. Adame, M. Carrascosa, and B. Bellalta, "Time-sensitive networking in IEEE 802.11 be: On the way to low-latency WiFi 7," 2019. *arXiv:1912.06086*. [Online]. Available: http://arxiv.org/abs/1912.06086

Overview of Multi-Link Operation in IEEE 802.11be Wi-Fi Networks

Kwang Min Jeong
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jkm1123@yonsei.ac.kr

Minsu Choi
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
slau1019@yonsei.ac.kr

Byounghoon Son
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
sonhun99@yonsei.ac.kr

Jeongyoon Shin
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
sjy970528@yonsei.ac.kr

Jong-Moon Chung
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jmc@yonsei.ac.kr

***Abstract*— The demand for faster and more reliable wireless communication is increasing with the growing number of connected devices and data-intensive applications. Wi-Fi 7, the latest standard of the IEEE 802.11 family, aims to address these challenges by offering unprecedented speed, capacity, and efficiency improvements. Wi-Fi 7 can be used in the maritime area to provide high-speed wireless connectivity for various applications, such as navigation, communication, entertainment, and security. Wi-Fi 7 can also support multi-user and multi-device scenarios, such as multiple ships or drones communicating with each other or with a base station. This survey paper provides an overview of the key features and element technologies of Wi-Fi 7, specifically Multi-link Operation (MLO).**

Keywords*— *IEEE 802.11be, Wi-Fi 7, Multi-link operation, Alternating multi-link, High band simultaneous multi-link, Preamble puncturing

I. INTRODUCTION

In recent years, the maritime industry has seen a growing demand for high-speed internet connectivity for both commercial and personal use. Wi-Fi is one of the most popular wireless technologies used in maritime environments, providing connectivity to devices on ships, offshore platforms, and other maritime facilities. With the introduction of Wi-Fi 7, the maritime industry can potentially benefit from improved data rates, increased capacity, and reduced latency. And, various research efforts are underway in companies and research labs, and papers (including white papers from IT companies [1],[2]) for implementing IEEE 802.11be for this technology have been provided, with some products expected to be released soon, even in 2023. This paper provides an overview of the technology elements of multi-link operation (MLO) used in Wi-Fi 7. The paper concludes with directions for future research on Wi-Fi 7 in maritime environments. Section II describes MLO and discusses the concepts of its elementary technologies: alternating multi-link (AML), high-band simultaneous multi-link (HBSML), and preamble puncturing (PP). Finally, Section III discusses recent research trends.

II. MULTI-LINK OPERATION

A. Introduction

Wi-Fi technology's wider channels offer high throughput and low latency, but those come with some inefficiencies. To avoid interference, transmissions need to be fully synchronized when modulating to wider Orthogonal Frequency Division Multiple Access (OFDMA). If the underlying 20 MHz channel is busy, it can block the entire wide channel, leaving sub-channels idle. Wi-Fi 7 overcomes this inefficiency by introducing an advanced multi-link operation.

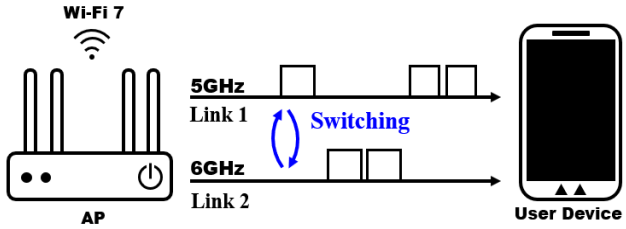


Fig. 1. Alternating multi-link operation example

MLO is a technique that allows multiple independent links to operate simultaneously between two or more devices. With MLO, users can expect high speeds and low latency as multilink devices (MLDs) can use the 2.4GHz, 5GHz, and 6GHz bands all at the same time in a variety of situations, such as load balancing to address traffic demand spikes or aggregating data across multiple bands. In cluster ship operations that use IoT device control and anything else that is latency sensitive, this technology can serve as a big advantage. Through bandwidth aggregation, Wi-Fi 7 MLO stations (STAs) can increase the throughput and have more opportunities to transfer data, which reduces latency. Those are also able to coordinate between frequencies so that those don't interfere with each other, resulting in a stable connection.

B. Alternating Multi-Link(AML)

Access Points (APs) support a low-band channel of 2.4 GHz and two high-band channels of 5 and 6 GHz. Wi-Fi 7 can support different ways for user devices to utilize these channels. Figure 1 shows a device using MLO to freely switch between each frequency band. One is to allow devices to automatically select a less congested frequency band and transmit data over that frequency channel. This prevents congestion on the link, reducing latency. Devices using older Wi-Fi technology can connect to all three bands, but those can only send data over one band at a time, and the frequency hopping switching overhead

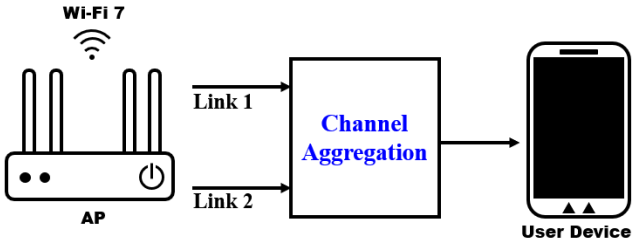


Fig. 2. High band simultaneous Multi-Link

can result in delays of up to 100 ms. However, with MLO technology, the delay is expected to be reduced to 1 ms using a unified connection

C. *High band simultaneous Multi-Link (HBSML)*

One of the advancements in Wi-Fi 7 is HBSML. As shown in Figure 2, HBSML allows devices to transmit data on all available frequency bands simultaneously, maximizing throughput and reducing latency. By operating on two bands simultaneously, devices can combine available frequency bands and utilize the throughput of each. This reduces congestion on each band, resulting in faster data transfer and a better performance. Using multiple bands simultaneously also provides redundancy and ensures a more reliable connection. In addition to utilizing multiple frequency bands, it is possible to use the 2.4 GHz frequency band more efficiently. IoT devices that require lower transmission rates can be assigned to this band, freeing up more bandwidth in the 5 GHz and 6 GHz bands for other user devices. This frequency band allocation helps reduce interference and improve overall network performance. Wi-Fi 7 can also combine four streams through HBSML to create a single, efficient 320 MHz band, which can deliver higher data rates and better performance, especially in high network demand environments.

D. *Preamble Puncturing (PP)*

In previous Wi-Fi systems, if an AP occupies a free contiguous channel in the spectrum, it is not allowed to use that frequency. PP is a smart interference avoidance technology to solve this problem. Wi-Fi 7 uses its smart solution to prevent interference in the spectrum, as shown in Figure 3. By avoiding parts of the spectrum and allocating contiguous channels within that spectrum, the overall bandwidth is lowered by the punctured amount, but in the end, a wider channel is available than if PP had not been used. In addition, when multiple bands are operating simultaneously, such as with HBSML as described above, devices accessing an interfering channel can be directed to an alternate channel.

III. Open Issues and Current Direction

Despite the promise of MLO, there are several open issues, including Non-simultaneous transmit and receive (NSTR) and Legacy Blindness, Spectrum Inefficiency, Channel Access Fairness, and Load Balancing, which are pointed out in [3] and direction is given on how to address them. In addition, several conference papers and journal articles have proposed measures to further improve MLO. In [4], a new analytical model for NSTR and Legacy Blindness issues is presented, showing a scenario where NSTR MLD does not result in fairness problems

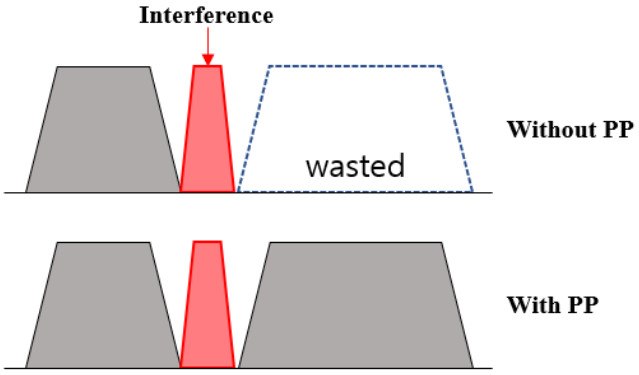


Fig. 3. Preamble Puncturing

in heterogeneous Wi-Fi networks. Among them, implementing a traffic manager to perform traffic allocation as a method for load balancing issues is proposed in [5]. The authors of [6] analyzed scenarios to achieve higher performance when building a dense network using MLO and showed that it is important to choose the right perspective according to the network or device type when determining channel access in dense deployments. It is proposed that artificial intelligence techniques such as reinforcement learning could be the solution.

IV. Conclusion

In this paper, a description of MLO and its component technologies are presented. It is shown that some published papers have addressed this technology focusing on open technical issues and corresponding methods that the authors have tried to overcome and improve the performance.

Acknowledgment

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2021-0-00040, Development of intelligent stealth technology for information and communication resources for public affairs and missions)

References

[1] D. Tseng, “Key Advantages of Wi-Fi 7: Performance, MRU & MLO” Mediatek, White paper, February 2022. [Online]. Available https://www.mediatek.com/blog/mediatek-wifi-7-whitepaper-detailing-key-advantages-performance-mru-and-mlo

[2] A. Hsu, “Wi-Fi 7 Multi-Link Operation (MLO)”, Mediatek, White paper, January 2022. [Online]. Available, https://www.mediatek.com/blog/wifi7-mlo-white-paper

[3] Á. López-Raventós and B. Bellalta, "Multi-Link Operation in IEEE 802.11be WLANs," *IEEE Wireless Commun*, vol. 29, no. 4, pp. 94-100, Aug. 2022

[4] N. Korolev, I. Levitsky and E. Khorov, "Analytical Model of Multi-Link Operation in Saturated Heterogeneous Wi-Fi 7 Networks," *IEEE Wireless Commun. Lett.*, vol. 11, no. 12, pp. 2546-2549, Dec. 2022

[5] Á. López-Raventós and B. Bellalta, "Dynamic Traffic Allocation in IEEE 802.11be Multi-Link WLANs," in *IEEE Wireless Commun. Lett.*, vol. 11, no. 7, pp. 1404-1408, July 2022

[6] D. Medda, A. Iossifides, P. Chatzimisios, F. José Velez and J. -F. Wagen, "Investigating Inclusiveness and Backward Compatibility of IEEE 802.11be Multi-link Operation," *2022 IEEE Conference on Standards for Communications and Networking (CSCN)*, Thessaloniki, Greece, 2022, pp. 20-24

Analysis of Energy Efficient Operation using Target Wake Time (TWT) in IEEE 802.11be

Youngwook Kim
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
kyw8738@yonsei.ac.kr

Yunyeoung Goh
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
rhdbsdud@yonsei.ac.kr

Jungmin Seo
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
sjm1335@yonsei.ac.kr

Minseung Park
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
alex35@yonsei.ac.kr

Jong-Moon Chung
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
jmc@yonsei.ac.kr

***Abstract*— IEEE 802.11be has been proposed to support extremely high-throughput Wi-Fi services, which include real-time applications and high-resolution video streaming. In maritime, especially, many crew members can connect to access points (APs) on a ship, enabling them to use the high-throughput and low-latency services for video monitoring, remote control, and entertainment purposes. However, the energy efficiency in battery-powered mobile devices is also an important consideration. This article introduces IEEE 802.11be and related features. Furthermore, this article provides an overview of Target Wake Time (TWT) for power-saving. Finally, we analyze and summarize state-of-the-art research related to power-saving methods in IEEE 802.11be.**

Keywords—IEEE 802.11be, Power-saving, Target Wake Time

I. INTRODUCTION

Innovations in wireless communication technology have created a tremendous demand for high-throughput, reliability and low-latency in the network. The growth of network service, such as, extended reality (XR) and 4K/8K ultra-high-resolution video streaming, is further increasing this demand. Especially in maritime applications, there are many use cases that require an enhanced network. For example, high-resolution CCTV is being streamed in real-time as a way to check the external and internal conditions of a ship, and many containers are being remotely controlled from a ship. Additionally, many crew members use high-throughput network for entertainment purposes. To address these needs, Task Group BE (TGbe) of the IEEE 802.11 working group has defined a standard, named IEEE 802.11be Extremely High Throughput (EHT), to support high-speed wireless networks. IEEE 802.11be EHT aims to provide superior performance such as high-throughput [1]. The standard also includes Target Wake Time (TWT) in order to enhance battery efficiency on high-throughput networks. Energy efficiency is important in high-throughput networks, as most network devices (e.g., smartphone and IoT devices) are powered by batteries and have a limited lifetime. For this reason, TGbe has included TWT in the IEEE 802.11be specifications to increase energy efficiency. In this article, we introduce IEEE 802.11be and TWT from a high-throughput network and energy efficiency perspective. In the final section, we will investigate how to use TWT with IEEE 802.11be.

II. IEEE 802.11BE EHT AND ENERGY EFFICIENCY

To support extremely high throughput above 30 Gbps, enhanced PHY and MAC protocols are proposed in 802.11be EHT. In the IEEE 802.11be EHT specification [2], IEEE 802.11be features mainly the following characteristics.

- Multi-RU
- Multi-link operation (MLO)
- Channel Bandwidth up to 320 MHz
- Enhanced Modulation scheme 4096-QAM

IEEE 802.11be features support enhanced network performance using enhanced resource allocation, increased network capacity, and multi-link operation. However, since most of the stations (STAs) are battery-powered devices, the impact of improved throughput on battery lifetime is the key factor to be considered. In [3], research has shown that, due to the characteristics of MLO and multi-AP collaboration, EHT-devices can drain batteries while in listening mode for a long period of time, which is needed to handle multi-link connections. Moreover, the authors suggested that battery management in 802.11be can be one of the research objectives in the 802.11be study.

III. TARGET WAKE TIME (TWT)

TWT is a feature of IEEE 802.11 for energy efficiency. Since the introduction of TWT in IEEE 802.11ah, TWT has improved power-saving performance. Eventually, TGbe included further improved TWT, Restricted TWT, in IEEE 802.11be. TWT allows networked devices to enter doze mode to save power consumption while ensuring that they wake up in time to receive or send data packets. IEEE 802.11ax supports several types of TWT mechanism as listed below.

- Broadcast TWT
- Individual TWT

For individual TWT operations, each STA needs to establish individual TWT agreements with the AP. In contrast, Broadcast TWT doesn't require an agreement of TWT between the scheduled STAs and the scheduling STA [4]. Fig. 1 shows an overview of the broadcast TWT mechanism. During the Target Beacon Transmission Time (TBTT) negotiation period, STAs agree on the target beacon to be received. As a result, STAs can

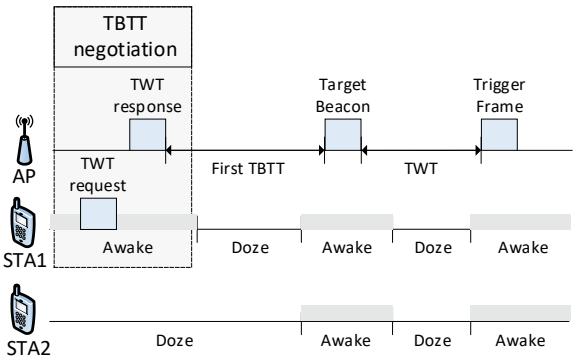


Fig. 1. Broadcast Target Wake Time mechanism

save the battery by entering doze mode until the negotiated time to listen for a target beacon. When STAs listen a target beacon containing information about the time to wake up, STAs can enter doze mode again until the time to listen for a trigger frame. In doze mode, STAs can save much battery energy [5]. For this reason, E. Stepanova, *et al*. proposed collaborated ways to use TWT and Wake Up Radio (WUR) based on IEEE 802.11ba for improved power-saving of IoT devices [6]. TWT is the most effective power-saving methods in IEEE 802.11, but if TWT does not work properly, it can cause performance degradation and battery drain due to transmission collisions. For this reason, a lot of studies to optimize TWT has been conducted based on IEEE 802.11ax. For example, Chen, Q., *el at*. proposed a TWT scheduling strategy in IEEE 802.11ax features, such as, the number of active UEs or resource units (RUs), which have a great impact on throughput and energy efficiency [7]-[9], and the author of [10] proposed a MAC protocol based on grouping using TWT scheduling for improvement of energy efficiency.

IV. IEEE 802.11BE WITH TARGET WAKE TIME

TGbe considers TWT one of the key features for energy efficiency in IEEE 802.11be. In [11], the authors mentioned that TGbe suggested TWT with Traffic Indication Map (TIM) as a power-saving method for MLO features in 802.11be. In addition, 802.11be has extended the functionality of TWT under the name of Restricted TWT (rTWT). TWT is not only a power-saving method by minimizing medium contention, but also a deterministic low latency way by explicitly identifying the time for STA to wake up. rTWT is a IEEE 802.11be feature that APs can reduce jitter and volatility of latency using enhanced medium access protection and resource reservation operations. Consequently, rTWT increases reliability of latency sensitive services. In [12]-[14], the authors featured rTWT as a key feature that reduces latency volatility in IEEE 802.11be. Table I summarizes the TWT proposals based on target standards and related features.

V. CONCLUSIONS

This article studied the characteristics of IEEE 802.11be and battery drain issues caused by IEEE 802.11be features. In addition, this article presented TWT mechanism and TWT related studies including rTWT. In summary, TWT in IEEE 802.11be will become the most popular topic in IEEE 802.11 for energy efficiency and reliable networks.

TABLE I. TWT PROPOSALS DRIVEN BY FEATURES

Paper	Target Standard	Problems or challenges	Approach
[4]	801.11ax	Battery efficiency in 802.11ax	TWT
[6]	802.11ba	Energy efficient scheme for IoT devices	TWT + WUR
[7]-[9]	802.11ax	Improper scheduling of TWT and Energy efficient scheduling	Feature based TWT scheduling
[10]	802.11ax	Optimization of the MAC layer with TWT	MAC protocol with TWT
[11]	802.11be	Power-saving scheme for Multi-Link Operation in 802.11be	TWT + TIM
[12]-[14]	802.11be	Deterministic worst latency and low jitter in 802.11be or future	Restricted TWT

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government (MSIT) (2021-0-00040, Development of intelligent stealth technology for information and communication resources for public affairs and missions).

REFERENCES

[1] D. Lopez-Perez, A. Garcia-Rodriguez, L. Galati-Giordano, M. Kasslin and K. Doppler, "IEEE 802.11be Extremely High Throughput: The Next Generation of Wi-Fi Technology Beyond 802.11ax," *IEEE Commun. Mag.*, vol. 57, no. 9, pp. 113-119, Sep. 2019

[2] IEEE 802.11 WG, "Ammendment 8: Enhancements for Extremely High Throughput (EHT)," IEEE P802.11be D1.0, May 2021..

[3] Deng, C., Fang, X., Han, X., Wang, X., Yan, L., He, R., Long, Y., and Guo, Y., "IEEE 802.11 be Wi-Fi 7: New challenges and opportunities," *IEEE Commun. Surveys Tuts*., vol. 22, no. 4, pp. 2136-2166, 4th Quart., 2022.

[4] Deng, D. J., Lin, Y. P., Yang, X., Zhu, J., Li, Y. B., Luo, J., and Chen, K. C. "IEEE 802.11 ax: highly efficient WLANs for intelligent information infrastructure," *IEEE Commun. Mag.*, vol. 55, no. 12, pp. 52-59, Dec. 2017.

[5] M. Nurchis and B. Bellalta, "Target Wake Time: Scheduled Access in IEEE 802.11ax WLANs," *IEEE Wirel. Commun.*, vol. 26, no. 2, pp. 142-150, Apr. 2019

[6] E. Stepanova, D. Bankov, E. Khorov and A. Lyakhov, "On the Joint Usage of Target Wake Time and 802.11ba Wake-Up Radio," *IEEE Access*, vol. 8, pp. 221061-221076, 2020

[7] Chen, Q., Liang, G., and Weng, Z., "A target wake time based power conservation scheme for maximizing throughput in IEEE 802.11 ax WLANs," in *Proc. IEEE 25th Int. Conf. Parallel Distrib. Syst. (ICPADS)*, Tianjin, China, Dec. 2019, pp. 217-224

[8] Chen, Q., and Zhu, Y. H., "Scheduling channel access based on target wake time mechanism in 802.11 ax WLANs," *IEEE Trans. Wireless Commun.*, vol. 20, no. 3, pp. 1529-1543, Mar. 2020.

[9] Chen, Q., Weng, Z., and Chen, G., "A target wake time scheduling scheme for uplink multiuser transmission in IEEE 802.11 ax-based next generation WLANs." *IEEE Access*, vol. 7, pp. 158207-158222, 2019.

[10] Chen, Q., "An Energy-Efficient Channel Access With Target Wake Time Scheduling for Overlapping 802.11 ax Basic Service Sets." *IEEE Internet of Things J.*, vol. 9, no. 19, pp. 18973-18986, Oct. 2022.

[11] Á. López-Raventós and B. Bellalta. "Multi-link operation in IEEE 802.11 be WLANs." *IEEE Wirel. Commun.*, vol. 29, no. 4, pp. 94-100, Aug. 2022.

[12] E. Reshef and C. Cordeiro. "Future Directions for Wi-Fi 8 and Beyond" *IEEE Commun. Mag.*, vol. 60, no. 10, pp. 50-55, Oct. 2022.

[13] C. Chen, X. Chen, D. Das, D. Akhmetov and C. Cordeiro, "Overview and performance evaluation of Wi-Fi 7." *IEEE Commun. Stand. Mag.*, vol. 6, no. 2, pp. 12-18, Jun. 2022

[14] D. Cavalcanti, C. Cordeiro, M. Smith and A. Regev, "WiFi TSN: Enabling Deterministic Wireless Connectivity over 802.11." *IEEE Commun. Stand. Mag*., vol. 6, no. 4, pp. 22-29, Dec. 2022

Implementation of Monitoring System for Mission Critical Network Verification

Jae-Woo Kim
ICT Convergence Research Center
Kumoh national institute of Technology
Gumi, Republic of Korea
jaewookim@kumoh.ac.kr

Gi-Hyeob Kwon
ICT Convergence Research Center
Kumoh national institute of Technology
Gumi, Republic of Korea
navkwon@kumoh.ac.kr

Dong-Seong Kim
Dept. of IT Convergence Engineering
Kumoh national institute of Technology
Gumi, Republic of Korea
dskim@kumoh.ac.kr

***Abstract*—This paper proposes a method for designing and implementing a network testing tool that can efficiently measure network performance. The proposed method defines the necessary functionalities for testing network performance and designs a tool for network testing that meets these requirements. The proposed testing tool utilizes the open-source network traffic testing tool, iPerf, to measure network throughput. The proposed tool incorporates the necessary functionalities for user convenience and network testing, as based on iperf. This paper outlines the design process, implementation method, and results of a network performance testing tool that utilizes iperf.**

***Index Terms*—Traffic Test, Network Test, iperf, implementation**

I. Introduction

Network testing tools are widely used not only in research but also in the IT industry. These procedures are essential for identifying various problems that may occur in the network environment before designing and configuring a network. It also helps to verify the network's maximum performance. Network testing tools can be used to identify and address issues such as insufficient bandwidth, packet loss, and latency in the network environment. In particular, for networks that prioritize stability, it is essential to proactively identify and resolve any potential failures. There are several commercial network testing tools available, including Wireshark, which offers packet capture and analysis capabilities. Nmap is another tool that measures security-related network performance. Additionally, there are various tools developed to measure specific performance parameters such as network bandwidth, utilization, and packet loss. [1] [2].

On the other hand, evolving network systems such as 5G and 6G require pre-testing of networks due to their new structures and demands. For example, during a field trial using the 28GHz band in a 5G network, the iperf network testing tool was used for network performance verification [3]. Networks have transitioned from a structure in which each node used only one band to a structure in which each node can simultaneously use multiple bands. To accurately test the performance of a network, it is essential to generate multiple types of traffic simultaneously. This ensures that the network can handle various types of data and traffic loads, providing a more comprehensive evaluation of its capabilities. Additionally, a feature is needed to remotely test traffic instead of directly managing intricate network nodes. Moreover, given that network nodes utilize a mix of Windows, Linux, and macOS operating systems, the network testing tool must be compatible with all three operating systems.

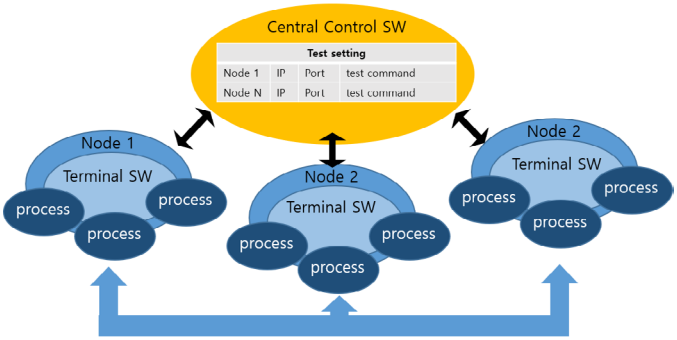


Fig. 1. Network Test tool architecture

In this paper, we present a GUI-based network traffic testing tool designed to meet the aforementioned requirements. The tool is based on iperf, a command-line-based network traffic testing tool, but with added user convenience and functionality. We have enhanced the network performance testing software to enable remote simultaneous testing on multiple nodes, which was previously limited to testing on a single node. We utilized a cross-compile framework as a development tool to run desktop applications on any operating system [4]. Figure 1 depicts the structure of the network performance testing tool proposed in this paper. The network performance testing tool implemented in this paper is divided into terminal software and central control software. The terminal software and central control software communicate with each other by exchanging control data and test result data.

II. System Design

A. Terminal Software

The terminal software is installed on each node of the network, and it executes the iperf process to measure network traffic. The terminal software is capable of measuring the maximum available bandwidth between network nodes by creating multiple processes independently. It can also test

specific bandwidths based on the settings. It can measure multiple bands simultaneously by performing multiple processes. Figure 2 shows the operational flow of the terminal software. Moreover, the terminal software is designed to function in network mode, enabling it to be controlled remotely by the central control software. In network mode, this system has the ability to configure connections with the central control software, differentiate control data received from the central control software, and execute and oversee the network traffic testing tool processes.

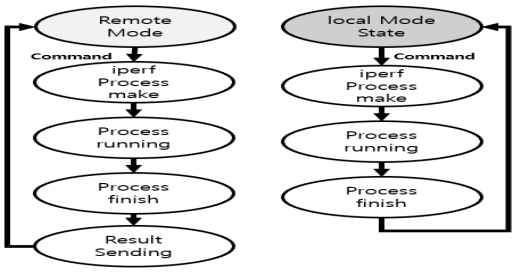


Fig. 2. The Terminal SW Flowchart

B. Central Control Software

The central control software enables efficient network testing by registering multiple tests simultaneously and allowing remote testing of each node. Figure 3 displays the operational flow of the central control software. It transitions from the "Wait" state to perform tasks such as configuring traffic testing scenarios, conducting traffic tests, and analyzing test results as per the user's request. The software also manages relevant lists based on connection requests from the terminal software (clients). When a test is conducted, the list of test commands, results of each test, and settings are stored in a database format.

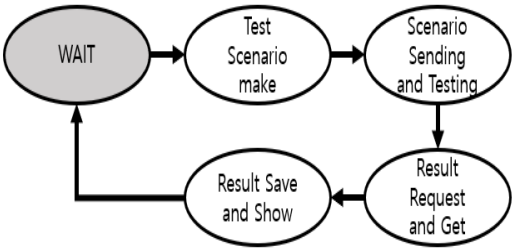


Fig. 3. Central Control SW Flowchart

III. Implementation Results

The proposed network performance testing tool was implemented using the cross-compile framework Qt6 C++. It was ported to both Windows and Linux platforms, and the implementation results were verified. Terminal software was loaded onto separate nodes running Windows and Linux (Ubuntu) respectively, and the central control software conducted the tests from the Windows node. In the traffic test, a certain amount of bandwidth was tested between the two nodes where the terminal software was loaded. Figure 4 shows the results of the test performed in standalone mode by the terminal software. A bandwidth of approximately 500 Mbps was tested using a loopback address, and the test results can be observed through graphs and terminal logs. Figure 5 represents the implementation results of the central control software. It was confirmed that the proposed design method functions correctly, including remote node access, test command configuration, execution, result collection, and verification. The result verification displays graphs and summarizes key test results similar to the terminal software.

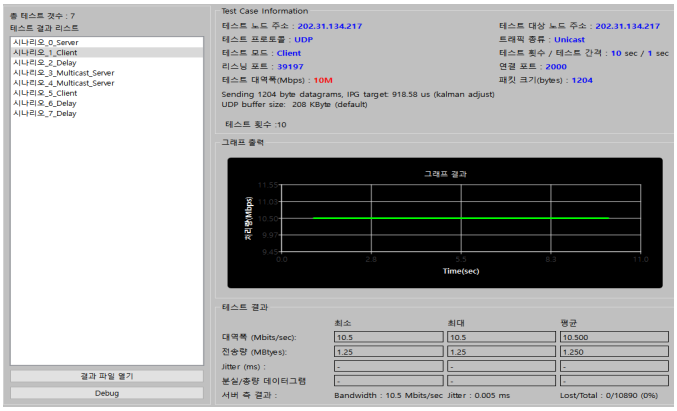


Fig. 4. Central Control SW Flowchart

IV. Conclusions

This paper describes the method and implementation results of developing a testing tool to measure network performance in the next-generation network architecture. The proposed network performance testing tool is capable of simultaneously performing multiple tests required for network testing, providing operating system independence, remote control capability, and user convenience. The implementation results demonstrate the ability to measure the available bandwidth between end-to-end nodes in the network and satisfy the requirements of network testing. Future research plans include expanding the functionality of the proposed tool and incorporating additional features required in various scenarios.

Acknowledgment

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2022R1I1A1A01071058,2018R1A6A1A03024003).

References

[1] Lamping, Ulf, and Ed Warnicke "Wireshark user's guide," Interface 4.6 2024.

[2] Orebaugh, Angela, and Becky Pinkard. Nmap in the enterprise: your guide to network scanning. Elsevier, 2011.

[3] Tirumala, A. (1999). Iperf: The TCP/UDP bandwidth measurement tool. http://dast. nlanr. net/Projects/Iperf/.

[4] The Qt Framework, 2023, [online] https://www.qt.io/ product/framework

A Study on Applying DL to MAS-based Multiple UAVs for Disaster Management

Sung Hyun Kim
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
leothecat@yonsei.ac.kr

Juyeong Hwang
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
wndud85v@yonsei.ac.kr

Gangwoo Lee
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
gang5541@yonsei.ac.kr

Jong-Moon Chung
School of Electrical and Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jmc@yonsei.ac.kr

***Abstract*—When facing disastrous situations, it is important to provide emergency supplies in an efficient way to the area of disaster occurrence. Additionally, preventing the high risk of casualties and victims is another reason why using multiple Unmanned Aerial Vehicle (UAV) could play an essential role in this field. This paper review focuses on how multi-agent system (MAS) based multiple UAVs could play an essential role in disaster management.**

Keywords— UAV, multi-agent system, artificial intelligence, machine learning, prediction

I. Introduction

Recently, Unmanned Aerial Vehicles (UAV) are given many essential roles in military, security, IoT, agriculture, and network operations. UAVs are controlled remotely, capturing images along the way when piloted [1]. Drone image processing is used in many industrial fields, including oil, gas, agriculture, coastline mapping, and urban planning. According to the United States Coast Guard's Boating Safety Partners, in 2021, there were 4439 accidents, 2641 injuries, and 658 deaths in total. Wind, water conditions, water temperature, and visibility is an important factor for boating. Due to these boating conditions, using drones for coastline mapping is vital due to the rise of sea level, natural disasters, and climate change [2]. Also, there are still countires, such as Maldives and Indonesia, that face challenges regards to localized level Early Warning Systems (EWS). Thus, UAVs could be used in the coastline environment for preventing unprecedented accidents, such as casualties and boating instruments damages. It can be used to provide images and live videos from the area of operations [3]. Using a machine learning-based maritime accident prevention prediction system, the availability to interpret the reason of accident becomes possible.

According to [4], risk factors are categorized into four groups based on the common characteristics of boating accidents: environment, human, vessel, and management [4]. With the unique characteristics explained above, this could correlate to analyzing accident data to predict the high risk of casualties and victim occurrence. In order to cover disaster relief and reduce the loss risk, the use of multi-agent cooperative systems consisting of multi-UAVs will provide early warnings for disaster management (DM) [5].

The focus of this paper is to analyze and overview the following questions:

1) What is the role of a multi-agent systems in DM?

2) What are the impact and benefits of Multi-agent systems (MAS) applying multiple UAVs for predicting a disastrous situation?

3) How can multiple UAVs become essential when equipped with Deep Learning (DL)?

II. Background and Related Works

According to [4], MAS is defined as a numerous autonomous agents that can accomplish complex tasks and interact information from the surroundings. There are various types of MAS, such as multiple UAVs, multiple Autonomous Underwater Vehicles (AUV), and multiple Unmanned Ground Vehicles (UGV), depending on the system purpose and the type of task.

When it comes to large area applications, such as disaster monitoring and management, it is a better option to use multiple UAVs rather than a single UAV. It is known that multiple UAVs have the advantage of scalability as well. Additionally, when it comes to the broadness of network and communication range, a MAS will enable multiple agents to cooperate as a team to work in a broader era. When it comes to the purpose of achieving DM, using multiple agents to cooperate and exchange information with neighboring agents will reach the purpose to manage and prevent disastrous events.

Reliability is one of the advantage that MAS-based multiple UAV operations can provide [5]. When there is an agent who failed to accomplish a task in a MAS for whatever reason, the task will not fail due to the backup of the other agents by reconfiguring it. Since each of the agent works autonomously in a multi-agent structure [4], when considering team behavior and communication, a persistent connectivity between the agents is essential.

Additionally, a MAS has various characteristics that supports the idea of disaster management in a multiple UAV environment [5]. The behavior-based strategy, also known as a structured network, is when the behavior coordinator takes the decision making. In this manner, each behavior gets inputs from the agent's sensor and provide outputs to the agent's actuators. This would give advantages during disaster management by keeping formation, collision avoidance, and goal-seeking. Fig. 1

depicts the concept of the MAS-based multiple UAV environment for an emergency communication system.

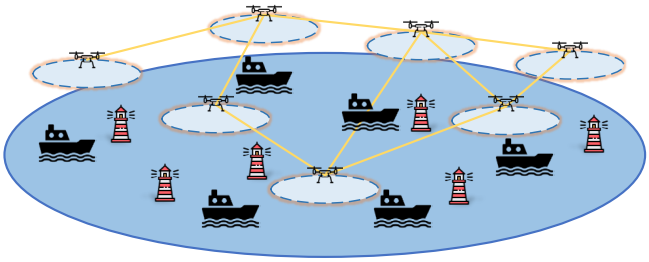


Fig. 1. MAS-based multiple UAV environment for an emergency communication system

When multiple UAVs are applied in a MAS-based environment, there are various applications following with limitations as well. If the MAS-based multiple UAVs were to be applied to the prediction and warning system, then it will gather information of the health monitoring of the environment, and present an autonomous early warning system in order to fly regularly along the established routes [6]. For example, when it comes to a wildfire disaster, by exchanging data with the local administration of forests, then data can be accumulated, thus by using DL and data analysis, the prediction patterns will be more accurate.

Regardless of the disastrous situation, the effectiveness and convenience of using MAS-based multiple UAVs is essential when transporting emergency supplies [7]. When it comes to solving disastrous situations, it is important to deliver emergency supplies to shelters to meet the requirements of DM. Reinforcement learning (RL) is a method of learning by selecting the action that maximizes the reward. In this manner, it is important to consider the battery consumption of multiple UAVs.

III. MULTIPLE UAV TRANSPORTATION COST FUNCTION

Each UAV consumes its battery at a different rate depending on its traveling distance and payload. Referring to the study of [8], the energy consumption function (Equation (1)) for the amount of energy consumed by the UAV to travel from shelter a to b with payload v is described as follows.

$$R(v) = \delta_0 + \delta v + h_{ab}(\rho_0 + \rho v) \quad (1)$$

When it comes to using the transportation cost function in this environment, Q-learning (QL) can be applied, having the agent to be the UAV. The agent chooses an action A_t on the state S_t at a certain time t [7]. In QL, the agent has an action-value matrix, also known as the Q matrix (Equation (2)). This matrix represents the value of being in a specific state S_t when choosing an action A_t at an instant t. Based on this, the agent, being the UAV, will update a value of the Q matrix, represented as $Q(state, action)$, for each action processed. This is defined as follows:

$$Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha[r_{t+1} + \gamma \max_a Q(s_{t+1}, a) - Q(s_t, a_t)] \quad (2)$$

where $Q(s_t, a_t)$ is the current Q matrix and α is the learning rate with a range of $(0 \leq \alpha \leq 1)$. Parameter γ is the discount factor $(0 \leq \gamma \leq 1)$, r_{t+1} is the expected reward at instant t, and $\max_a Q(s_{t+1}, a)$ is the maximum action value among every possible actions in state s_{t+1}.

According to [7], a Q matrix which is shared among the agents is updated consecutively after each transportation. The state of the environment in the Q matrix and action of the agent based on a UAV were defined as follows:

- States: $S_t = Y_t = \{D_{1t}, D_{2t}, \dots, D_{st}\}$, where D_{at} is the set of the demands of shelter j at an instant t.
- Actions: An action at an instant t, A_t, is denoted as $A_t = \{L_m, w_{iablmn}\}$.

IV. CONCLUSION

Due to the high risk of casualties and management of disastrous events, it is highly important to have a prevention and backup plan for these type of situations. In this paper, the concept of the MAS-based multiple UAV concept was introduced and a few ideas were proposed to help manage and prevent the disastrous events by using DL technologies. In addition, future studies need to be focused on how the MAS-based multiple UAV concept could be applied to satisfy real-time emergency cases more reliably.

ACKNOWLEDGMENT

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143782) of the Republic of Korea.

REFERENCES

[1] N. Vanitha, G. Padmavathi, V. Indu priya, and S. Lavanya, "An Investigation of Unmanned Aerial Vehicle Surveliance Data Processing with Big Data Analysis," *Springer*, vol. 111, pp. 219–230, Sep. 2022.

[2] E. Djunarsjah, A. B. Harto, M. M. Julian, F. S. A. Sinurat, and N. S. Lubis, "Optimization of Drone for Coastline Mapping," in *2021 IEEE International Conference on Aerospace Electronics and Remote Sensing Technology (ICARES)*, Bali, Indonesia, Nov. 2021, pp. 1-7.

[3] E. Vasilopoulos, G. Vosinakis, M. Krommyda, L. Karagiannidis, E. Ouzounoglou, and A. Amditis, "A Comparative Study of Autonomous Object Detection Algorithms in the Maritime Environment Using a UAV Platform," *Computation*, vol. 10, no. 3, pp. 42-61, Mar. 2022.

[4] G. Kim and S. Lim, "Development of an Interpretable Maritime Accident Prediction System Using Machine Learning Techniques," *IEEE Access*, vol. 10, pp. 41313-41329, Apr. 2022.

[5] A. Khan, S. Gupta, and S. K. Gupta, "Cooperative Control between Multi-UAVs for Maximum Coverage in Disaster Management: Review and Proposed Model," in *2022 2nd International Conference on Computing and Information Technology (ICCIT)*, Tabuk, Saudi Arabia, Feb. 2022, pp. 271-277.

[6] S. Li, "Wildfire early warning system based on wireless sensors and unmanned aerial vehicle," *Journal of Unmanned Vehicle Systems*, vol. 7, no. 1, pp.76-91, Nov. 2019.

[7] D. Hachiya, E. Mas, and S. Koshimura, "A Reinforcement Learning Model of Multiple UAVs for Transporting Emergency Relief Supplies," *Applied Sciences*, vol. 12, no. 20, pp. 12-34, Oct. 2022.

[8] B. Rabta, C. Wankmüller, and G. Reiner, "A drone fleet model for last-mile distribution in disaster relief operations," *International Journal of Disaster Risk Reduction*, vol. 28, pp. 107-112, Jun. 2018.

A Survey of Conditional Handover in Non-Terrestrial Networks

Jehun Lee
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
loiloi123@yonsei.ac.kr

Jaewook Jung
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
qazaq9669@yonsei.ac.kr

Jeongeon Park
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
pje1996@yonsei.ac.kr

Shinhyeok Oh
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
ohsh99@yonsei.ac.kr

Jong-Moon Chung
School of Electrical & Electronic Engineering
Yonsei University
Seoul, Republic of Korea
jmc@yonsei.ac.kr

***Abstract*— This research paper discusses the challenges of handover (HO) in non-terrestrial networks (NTN) and proposes conditional HO (CHO) as a potential solution. NTN, including satellites and unmanned aerial vehicles, have advantages over terrestrial networks in mobility and coverage. However, due to the fast movement of satellites and the large distance between user equipment (UE) and satellites, measurement-based HO faces issues such as frequent HO and signaling overhead. This paper informs CHO as a mobility enhancement, which enables the UE to evaluate the HO execution conditions and initiate HO faster than conventional methods. Fast CHO (FCHO) is also discussed as a potential solution.**

Keywords—Non-terrestrial networks, Conditional handover, Fast conditional handover, LEO satellite, UAV communications.

I. Introduction

Due to the advancements in communication and production technology of satellites and UAVs, research on non-terrestrial networks (NTN) has been actively conducted to address the limitations of existing cellular networks. NTN can be defined as 'Networks, or segments of networks, using an airborne or space-borne vehicle to embark a transmission equipment relay node or base station.' NTN consists of satellites, high altitude platform systems (HAPs), and unmanned aerial vehicles (UAVs) [1], [2]. 3GPP is continuing to standardize on the use of NTNs in 5G environments starting with Release 15. One of the many benefits of NTN is the ability to communicate from anywhere on the planet, and in particular, low-earth orbit (LEO) satellites which can provide low-latency, high throughput communications anywhere in the world. Several satellite communication companies are competitively orbiting LEO satellites, and SpaceX, which is leading the industry in related technologies such as rocket relaunch, has a goal of launching about 42,000 LEO satellites [3], with about 3,500 satellites in use as of 2023. Compared to satellites, UAVs have the advantage of being very fast and flexible to deploy and operate, and can support wireless broadcasting and high rate transmission through strong LOS connectivity links [4], [5], [6]. Based on the advantages mentioned above, NTN can be a good solution to address various challenges in the maritime environment. Because the use of terrestrial networks (TN) in this environment is limited due to the requirement for terrestrial infrastructure to enable mobile communication. This research paper aims to describe the mobility challenges and enhancement of NTNs and provide a brief overview of HO strategies.

II. Maritime Use cases of NTN

There are various use cases for NTN in maritime environments. One such use case is satellite-based Automatic Identification System (AIS) [7], which can overcome the terrestrial VHF limitations of the AIS by utilizing satellites to share ship identification and movement-related information in any maritime location to prevent accidents. Another use case is Automatic Dependent Surveillance-Broadcast (ADS-B), which automatically collects and wirelessly broadcasts aircraft location and status information to related organizations, and utilizes LEO constellations to overcome the physical limitations of ground stations [8]. Furthermore, NTN can be utilized for military applications, such as providing line-of-sight (LOS) using UAVs, improving the accuracy of long-range missiles using satellite communication, continuously checking sea conditions using maritime sensors, and tracking underwater acoustic signals.

III. Challenges in Mobility for HO in NTN

In wireless communication, HO refers to the change of a serving cell to another cell while the UE is connected to the network [9]. In NTN, successful HO is essential to ensure the mobility of UEs and provide continuous and reliable services. HO in NR, as specified in 3GPP Release 15, is measurement-based, comparing measurements taken by the UE and moving it from the source cell to the target cell if certain conditions are met. However, in NTN, especially in LEO satellites, the distance between the satellite and the UE is very large and the LEO satellites move much faster compared to the UEs on the ground, so this HO method has the following challenges [2], [10].

A. Measurement validity

Geostationary orbit (GEO) satellites can use Release-15 HO mechanisms due to their high altitude, large cell size, and small signal variation. LEO satellites have smaller propagation delays than GEO satellites, but their rapid movement can lead to measurement validity issues, causing additional latency and early or late HO.

B. Cell overlap and reduced signal strength variation

TN has a large difference in Reference signal received power (RSRP) between the cell center and edge, but this is not the case for NTN especially for LEO satellites. Because the cell size and UE-satellite distance are very large compared to TN [2], [10]. Therefore, during measurement-based HO (e.g., A3 event), it may be difficult to find a cell that satisfies the HO condition, which may lead to HO problems such as UE ping-pong effects.

C. Frequent and unavoidable HO

In the LEO scenario, the UEs can suffer frequently due to unavoidable HOs because of satellite's high speed movement. Assuming that the LEO satellite has a velocity of 7.56 km/s, the time to HO can be calculated to be as low as 6.61 seconds and as high as 132.28 seconds in the LEO scenario, even when the UE's velocity is neglected as it is typically much slower than that of the satellite [2]. These frequent HOs cause signaling overhead and additional energy consumption of the UE, which can lead to long service interruptions and a poor user experience [10].

D. HO for a large number of UEs

The aforementioned characteristics of NTN indicate that it has the potential to result in a significantly higher number of UE HOs compared to TN cases. Assuming a uniform distribution of UEs, where the number of UEs leaving and entering a cell is equal, and if the number of UEs in a cell is 65,519, it can be calculated that at least 990 to 19,824 HOs occur per second, depending on the cell diameter [2]. Such a large number of HOs can cause signaling overhead and service continuity.

IV. Conditional HO as a Mobility Enhancement

3GPP introduced conditional HO (CHO) as a mobility enhancement in Rel-16. CHO is defined as a HO executed by the UE when one or more HO execution conditions are met. The UE starts evaluating the execution conditions as soon as it receives the CHO configuration and stops evaluating the execution conditions once the HO is executed [11]. Therefore, CHO initiates HO faster than traditional methods, but may introduce a delay until the HO candidate cell satisfies the conditions. Various HO triggering conditions have been proposed for CHO: Measurement-based triggered HO considers the NTN environment to determine the HO trigger threshold and performs HO based on it. Location-triggered HO is an HO method that utilizes the location information of the UE and satellite based on the deterministic position change of the satellite, either alone or in combination with other conditions. In addition, time (r)-based triggered HO that utilizes time information, HO that utilizes the time advance value (TAV) of the target cell and UE, and HO that conditions the elevation angle of the satellite have been proposed in 3GPP's standards documents [2], [10].

Recently, fast CHO (FCHO) has also been studied, which can compensate for the shortcomings of CHO [14]. While CHO has the advantage of overcoming mobility issues and reducing HO failures, it has the disadvantage of discarding HO preparations when it ends, so if HO fails, the HO process must be restarted from the beginning. This causes overhead and long latency. To overcome these problems, FCHO retains HO preparations in the HO preparation phase for some time after HO to reuse it when needed [12], [13], [14], [15]. Compared to CHO, FCHO has the advantage of reducing signaling overhead and lowering the HO failure rate [14], [15].

V. Conclusion

The mobility challenges and improvements in NTN have been summarized in this paper. NTN will become an essential service in next-generation communication systems due to its wide service coverage and flexibility, and seamless HO is very important when considering service continuity. However, the conventional HO method has problems due to mobility, and to solve this problem, 3GPP standardized CHO. In addition recently, Fast CHO has also been studied to compensate for the shortcomings of CHO, which can be effective for NTNs.

Acknowledgment

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143782) of the Republic of Korea.

References

[1] Study on New Radio (NR) to Support Non-Terrestrial Networks (Release 15), Standard 3GPP, TR 38.811 V15.4.0, 2020.

[2] Solutions for NR to Support Non-Terrestrial Networks (NTN) (Release 16), Standard 3GPP, TR 38.821 V16.1.0, 2021

[3] Di Vruno, Federico, and Vincenza Tornatore. "Large Satellite Constellations and Their Potential Impact on VGOS Operations," *International VLBI Service for Geodesy and Astrometry 2022 General Meeting Proceedings*. 2023.

[4] Oltjon Kodheli, et al, "Satellite communications in the New space era : a survey and future challenges," in *IEEE Communications Surveys & Tutorials*, vol. 23, no. 1, pp. 70-109, Firstquarter 2021.

[5] B. Li, Z. Fei and Y. Zhang, "UAV Communications for 5G and Beyond: Recent Advances and Future Trends," in *IEEE Internet of Things Journal*, vol. 6, no. 2, pp. 2241-2263, April 2019.

[6] M. Marchese, A. Moheddine, and F. Patrone, "IoT and UAV integration in 5G hybrid terrestrial-satellite networks," *Sensors*, vol. 19, no. 17, p. 3704, 2019.

[7] G. Sahay, P. Meghana, V. V. Sravani, T. P. Venkatesh and V. Karna, "SDR based single channel S-AIS receiver for satellites using system generator," *2016 IEEE International Conference on ANTS*, India, 2016.

[8] M. Strohmeier, M. Schäfer, V. Lenders and I. Martinovic, "Realities and challenges of nextgen air traffic management: the case of ADS-B," in *IEEE Communications Magazine*, vol. 52, no. 5, pp. 111-118, 2014.

[9] Evolved Universal Terrestrial Radio Access (E-UTRA) and Evolved Universal Terrestrial Radio Access Network (E-UTRAN) (Release 17), Standard 3GPP, TS 36.300 V17.4.0, 2023.

[10] E. Juan, M. Lauridsen, J. Wigard and P. Mogensen, "Handover Solutions for 5G Low-Earth Orbit Satellite Networks," in *IEEE Access*, vol. 10, pp. 93309-93325, 2022.

[11] E. Juan, M. Lauridsen, J. Wigard and P. Mogensen, "Performance Evaluation of the 5G NR Conditional Handover in LEO-based Non-Terrestrial Networks," *2022 WCNC*, USA, 2022, pp. 2488-2493.

[12] 3GPP, *Enhancements for Conditional Handover*, document R2-1907216, 3GPP TSG RAN WG2 Meeting #106, Apple, Reno, NV, May, 2019.

[13] 3GPP, *Consecutive Conditional Handover*, document R2-1909862, 3GPP TSG RAN WG2 Meeting #107, Apple, Czech Republic, Aug. 2019.

[14] 3GPP, *New WID on Further NR Mobility Enhancements*, document RP-213565, 3GPP TSG RAN Meeting #94e, E-Meeting, Dec. 2021.

[15] S. Bin Iqbal, A. Awada, U. Karabulut, I. Viering, P. Schulz and G. P. Fettweis, "On the Modeling and Analysis of Fast Conditional Handover for 5G-Advanced," *2022 IEEE 33rd Annual International Symposium on PIMRC*, Japan, 2022, pp. 595-601.

Design of Autonomous Unmanned Vehicle for Rapid Deployment

Seungjae Baek
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
baeksj@kiost.ac.kr

SangKi Jeong
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
jeongsk313@kiost.ac.kr

SukHyung Lee
Future Strategy Department
BMtech System Co., LTD
Seoul, Korea
gototop1@bmtsys.com

Sungmin Koo
Maritime ICT·Mobility Research Department
KIOST
Busan, Korea
smkoo@kiost.ac.kr

Bongchool Jeong
R&D Center
DAEYANG Electric. Co., LTD
Busan, Korea
bcjeong@daeyang.co.kr

Kyuchul Cho
R&D Center
Hangook Prop Inc.
Jinju, Korea
kccho@hangookprop.com

***Abstract*— In this paper, we introduce the concept of flying autonomous unmanned vehicles, which combines the functionalities of conventional autonomous unmanned vehicles and unmanned aerial vehicles into a single platform for rapid deployment. The primary focus of this paper is to conduct a study on the conceptual design of such a device and outline the essential components required for a unified unmanned underwater/aerial vehicle that can operate seamlessly in both air and water.**

Keywords—Autonomous Underwater Vehicle, Unmanned Aerial Vehicle, Rapid Deployment, Maritime Accident

I. Introduction

Traditional maritime unmanned vehicles such as AUVs(Autonomous Unmanned Vehicle) or underwater gliders rely on vessel operation, rendering them impractical during severe weather conditions when vessel operation is not feasible. Moreover, their slow movement speed prevents swift deployment to target areas, resulting in exorbitant operational costs [1, 2, 3]. This underscores the urgent need to develop a high-speed, constantly deployable unmanned marine vehicle for prompt responding to maritime accidents. In this paper, we introduce the concept of a rapid deployable AUV, called FlyingAUV, an autonomous unmanned vehicle capable of immediate deployment at the scene of a maritime accident. FlyingAUV's primary objective is to minimize human casualties by swiftly tracking and providing initial response to maritime and underwater targets across the entire Korean waters. Specifically, within one hour of dispatch, it can conduct a thorough search for maritime targets within a 100 km radius, transmit optical information, and in the event of a target sinking, maneuver at depths of up to 100 meters with a maximum speed of 3 knots for up to one hour. It will then transmit the sinking coordinates and optical/sonar information acquired underwater before returning to base.

II. Related Works

A maritime unmanned vehicle capable of both underwater and aerial mobility is a nascent technology that has garnered significant attention and research primarily from unmanned vehicle advanced nations due to its potential and profound impact. The U.S. Naval Research Laboratory has developed a maritime unmanned vehicle called Flimmer, which enables 'flight, surface landing, and underwater mobility' to perform

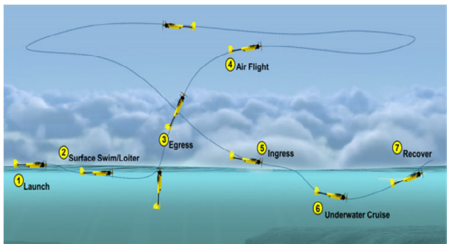


Fig. 1. EagleRay operation concept diagram.

missions such as submarine tracking by transitioning from long-range aerial flight to underwater submergence [2,5]. It has been confirmed to have a maximum takeoff weight of 9.9 Kg, a length of 1.06 m, and a wingspan of 1.82 m. The subsequent model, Flying Wanda Flimmer, is reported to have a maximum airspeed of 111 Km/h and a maximum underwater speed of 6 Knots [4].

North Carolina State University and Teledyne Scientific have developed a maritime unmanned vehicle called EagleRay, designed for high-speed aerial movement to enable underwater exploration at multiple locations with seamless aerial-to-underwater transitions [6]. It has a maximum takeoff weight of 5.7 Kg, a length of 1.4 m, a wingspan of 1.5 m, an airspeed of 19.6 Km/h, and an underwater speed of 1.32 Knots.

Imperial College in the United Kingdom has developed an underwater-to-aerial capable unmanned vehicle named AquaMAV. It utilizes a carbon dioxide-based water ejection system to escape from underwater and utilizes fixed-wing during aerial movement while folding its wings for underwater maneuverability [7]. It can fly at a speed of 10 m/s for up to 14 minutes in the air and is constructed using carbon fiber and kevlar to withstand the impact during transition.

III. Mission Requirement

Through an analysis of the actual requirements of end-users, we have developed scenarios for responding to maritime accidents and meticulously reviewed them to create a mission requirements document. The FlyingAUV is designed to vertical takeoff and landing (VTOL), allowing it to ascend vertically and reach the target location at a minimum speed of 100 Km/h in the event of a maritime accident. The departure point can be either on land or on a vessel, and it

Fig. 2. AquaMAV's underwater-to-air maneuvers.

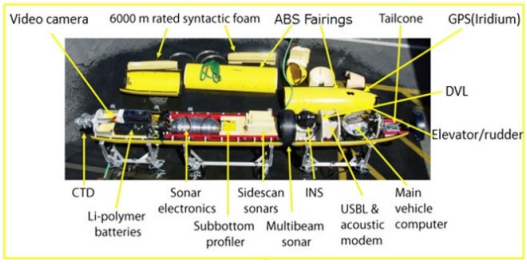


Fig. 3.AUV main components.

swiftly moves to the accident scene before the arrival of rescue teams or ships, transmitting optical images and location information. When the target object sinks, the FlyingAUV employs VTOL capabilities to descend from 5 meters in the air to a depth of 1 meter underwater within 30 seconds to begin underwater submergence. Underwater, it can submerge for up to 1 hour at a maximum depth of 100 meters and a speed of up to 3 Knots, capturing the final sinking location and acquiring sonar images of the sinking shape. If it is not possible to determine the location of the target object from the surface, the FlyingAUV can conduct a 30-minute search within a radius of 25 Km or a 60-minute search within a radius of 1.5 Km underwater. After securing the final sinking location and shape, it moves to the surface to transmit the acquired information. Within 30 seconds, it ascends to a height of 5 meters in the air and returns to the departure location or a designated location, concluding the mission.

IV. Design of FlyingAUV

FlyingAUV provides the both functionalities of conventional AUVs and UAVs in a single platform. Therefore, we conducted a weight analysis by examining the key components of conventional AUVs and UAVs. The key components of AUVs are presented in Fig. 3, while the key components of conventional UAVs are shown in Fig. 4.

After eliminating redundant and unnecessary elements from the components of AUVs and UAVs, we selected the key components of FlyingAUV based on the current mission requirements. The selected key components are illustrated in Fig. 5. They consist of a multi-beam imaging sonar for acquiring underwater sonar images, an IMU (Inertial Measurement Unit)/DVL (Doppler velocity log) for underwater positioning, and an EO (Electric Optica)/IR (Electric Infrared) system for capturing optical images in the air. Additionally, the mission system includes a sensor/main computer, batteries, and a propulsion system.

To optimize the placement of the EO/IR system, we designed FlyingAUV in a pusher-type configuration, which is commonly used in various unmanned aircraft. Considering the underwater-to-air environment where large propellers cannot be utilized, we opted for a ducted propeller to ensure sufficient thrust.

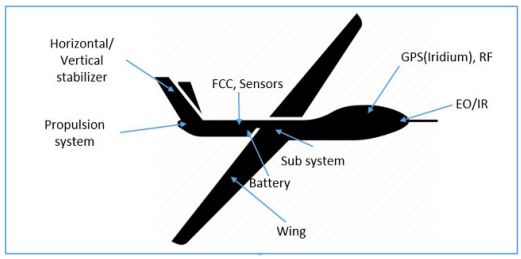


Fig. 4.UAV main components.

Carbon-fiber composite body	EO/IR
Multibeam image sonar	FCC, sensors, RF
DVL	Sub system
Main vehicle computer	Motor
Battery	Aileron
GPS(Iridium)	Duct & propeller

Fig. 5. FlyingAUV main components.

V. Conclusion

In this paper, we conducted a conceptual design of an immediate-response airborne and underwater capable autonomous unmanned vehicle called FlyingAUV. Through the development of a rapid-deployment airborne and underwater mobile marine robot, we aim to minimize human casualties by reducing rescue time and promptly reaching the target area in the event of a maritime accident. In the future, to realize this concept, we plan to pursue in-depth research on lightweight compact/pressure-resistant design technologies and a unified underwater/airborne propulsion system to facilitate the successful development of FlyingAUV.

Acknowledgment

This research was supported in part of the project titled `Development and Demonstration of Seawater Secondary Battery Module over 1kWh for Marine Equipment', funded by Ministry of Trade, Industry and Energy

References

[1] KOREA COAST GUARD, "2020 Maritime Distress Accident Statistical Yearbook, 2020.

[2] J. S. Kang, "US and European drone industry policy and regulatory policy walking different paths," The Optical Journal, Serial no.158, pp.61-64, Sep. 2015.

[3] Jinhong Noh, Kyon-Mo Yang, Min-Ro Park, Ji-won Lee, Min-Gyu Kim, and Kap-Ho Seo, "LiDAR Point Cloud Augmentation for Mobile Robot Safe Navigation in Indoor Environment," Journal of Institute of Control, Robotics and Systems, vol, 28, no, 1, pp. 52-58, Jan. 2021.

[4] Dasol Lee, "A Real-time Algorithm of Multiple Task Planning and Trajectory Planning for Drone Swarms," Journal of Institute of Control, Robotics and Systems, vol, 28, no, 4, pp. 371 - 376, Apr. 2021.

[5] Edwards Dan, "Flimmer: A flying submarine," Updates on NRL's Autonomy Research, Spectra: the Magazine of The Navy's Corporate Laboratory, no.4, pp.6-9, Aug. 2014.

[6] Warren Weisler, William Stewart, Mark B. Anderson, Kara J. Peters, Ashok Gopalarathnam, Matthew Bryant, "Testing and characterization of a fixed wing cross-domain unmanned vehicle operating in aerial and underwater environments," IEEE Journal of Oceanic Engineering, vol.43, no.4, pp.969-982, Aug. 2018.

[7] R. Siddal and M. Kovac, "Launching the AquaMAV: Bioinspired Design for Aerial-Aquatic Robotic Platforms," Bioinspiration and Biomimetics, vol.9, no.3, pp.1-15, Aug.SS

A Survey on Task Offloading System Modeling for Multi-Access Edge Computing

Myoungbo Kim
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
kmb0168@yonsei.ac.kr

Sangdo Kim
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
sangdokim96@yonsei.ac.kr

Wonsuk Yoo
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
wsbm34@yonsei.ac.kr

Jong-Moon Chung
School of Electrical & Electronic Engineering
Yonsei University
Seoul, South Korea
jmc@yonsei.ac.kr

***Abstract*— As technology advances, new applications such as extended reality (XR), augmented reality (AR), virtual reality (VR), autonomous driving (AD), industrial internet of things (IoT) has emerged. As the demand for data throughput increases, it is necessary to maximize the computational power and minimize latency. In addition, as technology gradually began to be used in various fields, data processed on the ground is extended to various ranges, such as, unmanned aerial vehicles (UAV), satellite, and maritime. As the distance between the Edge device and the server extends, the propagation delay naturally increases. As more User Equipment (UE) access services, more data needs to be processed by the server. To solve these computational power and delay problems, a new network technology called Multi-Access Edge Computing (MEC) has emerged, which can reduce the propagation delay and the workload of the server by processing the task near the UE side, at a location close to the network edge. The most important issue that needs to be addressed in MEC technology is Task offloading. The task offloading process is focused on determining how much of the task the UE device will offload to the MEC server. In this paper, we will introduce the methodologies to solve task offloading issues in many recent studies, classifying them according to the algorithm paradigm.**

Keywords—Multi-Access Edge Computing (MEC), Task Offloading

I. INTRODUCTION

With the rapid development of the mobile communications in recent years, people are no longer limited to places, but can use the service. Many mobile users use data not only on the ground, but also on aerial and maritime domains. It is hoped to be used huge computational power such as extended reality (XR), augmented reality (AR), virtual reality (VR), and streaming service. To satisfy this, MEC technology will be a good solution. MEC is a recently introduced technique that simply brings the task of the server to the Edge network closest to the UE. MEC has many advantages: it can reduce network congestion and increase utility by distributing computational power and reducing data communication. For Mobile Devices, which has limited computing and storage resources, the delay decreases and service quality increases. In particular, the task-offloading problem can benefit from appropriately distributing the task to an important issue within the MEC [1]. Task offloading is a problem that determines how much the mobile device will load its task to the MEC server, which benefits from various aspects such as minimizing latency, maximizing utility, bandwidth, and saving resources and energy [2]. Afterwards, the recent techniques will be introduced according to the algorithm paradigm of Task offloading.

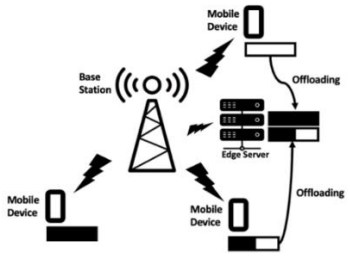


Fig. 1. Task offloading system model of MEC.

II. TASK OFFLOADING

If you define the concept of a task to analyze the task offloading algorithm, a task is the minimum unit of work to provide a service. In other words, the service must properly distribute the task to the server to satisfy Quality of Service (QoS) through a smooth service with a series of tasks. However, the task offloading problem is a representative optimization problem and at the same time a NP-hard problem [3]. To solve this problem, since the task offloading algorithm that has been recently being studied is an optimization problem, many studies have applied Greedy Heuristics Algorithm, Machine Learning, and the like to optimize the task offloading problem.

TABLE I. TASK OFFLOADING ALGORITHM PARADIGM

Algorithm	Proposed approach	Reference
Greedy heuristics	Game theoretical task offloading	[4]
	Iterative heuristic MEC resource allocation	[5]
	A greedy Select Maximum saved Energy First algorithm	[8]
Machine learning	Deep Reinforcement Learning	[6]
	Federated Deep Reinforcement Learning resource allocation	[7]
	DRL based task offloading algorithm to minimize latency	[9]

A. Greedy heuristics

Task offloading is a representative optimization problem and is also a NP-hard problem. Therefore, a greedy heuristic approach can be used to solve this problem. This method can find near optimal solutions without requiring special algorithms or more computations.

In [8], A greedy Select Maximum Saved Energy First (SMSEF) algorithm was proposed to minimize the energy consumption of mobile devices. This algorithm proceeds in a way that adopts the most energy-saving task at each iteration. In [4], The authors proposed Game-based Distributed Task Allocation and Scheduling (GDTAS) algorithm, which first convert the multi-server multi-task allocation problem to a game model and then prove the existence of Nash equilibrium of this game. Then, they approached the solution by proposing the Greedy Heuristic Better Response Update (GH-BRU) algorithm. In [5], also an iterative heuristic MEC resource allocation (IHRA) algorithm is presented to solve the NP-hard problem, and based on this, the greedy heuristic method is still being studied.

B. Machine Learning

With the development of machine learning, algorithms using machine learning and deep learning are being applied in various fields, and task offloading problems can be solved by using these algorithms. The authors of [6] proposed a learning method for selecting whether to load through deep reinforcement learning, and [7] proposed a federated Deep Reinforcement Learning (DRL) resource allocation algorithm that utilizes servers located additionally close to the edge in DRL.

In [9], the authors proposed the DRL model with three layers: application, MEC, and cloud. They trained the model based on the Directed Acyclic Graph (DAG) collected from the cloud layer. Then, based on the new S2S (Sequence-to-Sequence) neural network and Markov Decision Process (MDP), a task offloading problem is set up and the problem is solved. There are many ways to apply reinforcement learning that continues to learn and change due to the characteristics of networks that are difficult to change and predict.

III. Conclusion

With the development of XR, AR, VR, and autonomous driving, mobile devices require a considerable amount of computation. In addition, delays are also becoming an important issue as they work far away from land servers, such as Aerial and Maritime. Therefore, to meet this, MEC, a new network method, is drawing attention. Among MECs, task offloading problems are emerging in tasks where delay-sensitive or computational power is important, and to solve this problem, several papers are introduced according to the task-offloading algorithm. Although the field of MEC research is still in its infancy, more papers will be published in the future, and we hope this survey will help with future task offloading research.

Acknowledgment

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143782) of the Republic of Korea.

References

[1] H. Lin, *et al*., "A survey on computation offloading modeling for edge computing," *J. Netw. Comput. Appl.*, vol.169, no.1, Nov.2020, doi: 10.1016/j.jnca.2020.102781

[2] Ahmed. E., Ahmed. A., Yaqoob. I., Shuja. J., Gani. A., Imran. M., Shoaib. M., "Bringing computation closer toward the user network: Is edge computing the solution?," *IEEE Commun. Mag.* 55(11), pp. 138-144, Nov. 2017, doi: 10.1109/MCOM.2017.1700120

[3] Q.-V. Pham, L.B. Le, S.-H. Chung, W.-J. Hwang, "Mobile Edge Computing With Wireless Backhaul: Joint Task Offloading and Resource Allocation," *IEEE Access*., vol.7, pp. 16444-16459, Jan. 2019, doi: 10.1109/ACCESS.2018.2883692

[4] H. Teng, Z. LI, K. Cao, S. Long, S. Guo and A. Liu, "Game Theoretical Task Offloading for Profit Maximization in Mobile Edge Computing," *IEEE Trans. Mobile Comput*., early access, May. 2022, doi: 10.1109/TMC.2022.3175218

[5] Z. Ning, P. Dong, X. Kong, and F. Xia, "A cooperative partial computation offloading scheme for mobile edge computing enabled Internet of Things," *IEEE Internet Things J*., vol.6, no. 3, pp. 4804–4814, Jun. 2019, doi: 10.1109/JIOT.2018.2868616

[6] M. Elbayoumi, W. Hamouda, A. Youssef, "Edge Computing and Multiple-Association in Ultra-Dense Networks: Performance Analysis", *IEEE Trans. Commun.*, vol.70, no.8, pp. 5098-5112, Aug. 2022, doi: 10.1109/TCOMM.2022.3186989

[7] L. Zang, X. Zhang, and B. Guo, "Federated Deep Reinforcement Learning for Online Task Offloading and Resource Allocation in WPC-MEC Networks," *IEEE Access*, vol.10, pp. 9856–9867, Jan. 2022, doi: 10.1109/ACCESS.2022.3144415

[8] F. Wei, S. Chen, W. Zou, "A greedy algorithm for task offloading in mobile edge computing system," *China Commun.* vol.15, no.11, pp. 149-157, Nov. 2018, doi: 10.1109/CC.2018.8543056

[9] L. Huang, X. Feng, C. Zhang, L. Qian, Y. Wu, "Deep reinforcement learning based joint task offloading and bandwidth allocation for multi-user mobile edge computing," *Digit. Commun. Netw.* vol.5, no.1, pp. 10-17, Feb.2019, doi: 10.1016/j.dcan.2018.10.003

ICMIC 2023 The 2nd International Conference on Maritime IT Convergence

Proceedings of International Conference on Maritime IT Convergence (eISSN: 2983-3132)

http://icmic-conf.org/