ISSN 2093-0542






KOREAN SOCIETY FOR INTERNET INFORMATION

The 18th Asia Pacific International Conference on Information Science and Technology

(APIC-IST 2023)

June 25-28, 2023, Acros Fukuoka 7F, Fukuoka, Japan

http://www.apicist.org

Proceedings of APIC-IST 2023

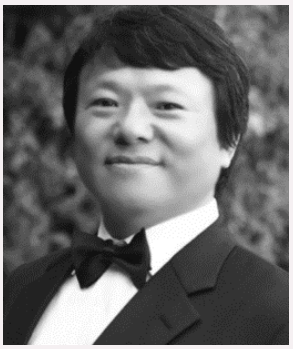
| Organized by |

Korean Society for Internet Information (KSII)

| Sponsored by |

Fukuoka Convention & Visitors Bureau

A Message from Honorary Chair



I express my warm heartfelt welcome to all of attendants of the 18th Asia Pacific International Conference on Information Science and Technology (APIC-IST 2023) that will be held on June 25~28, 2023 in Acros Fukuoka, Fukuoka city, Japan

I deeply appreciate all contributors for submitting papers, chairing sessions, and reviewing papers. And I must thank for the support we received from many companies.

Over the past 23 years, KSII (Korea Society for Internet Information) has increased its academic impact and influence in the area of Information Science, Computer Engineering and Internet Technology both worldwide and nationwide through this international conference and its SCIE-indexed journal, the *KSII Transactions on Internet and Information Systems* (TIIS).

As many of you have witnessed, APIC-IST is one of important International Conferences which KSII have provide annually. The amount of interest towards the internet is never decreasing among both the international academic and practitioners' communities. Topics covered in APIC 2023 truly represent the comprehensiveness of this conference. We are sure that APIC-IST 2023 will provide you with a wealth of information and many opportunities for future discussion.

Many people have worked hard to make APIC 2023 successful. A variety of conference arrangements and special events as well as well organized program have been prepared by the Conference Chair Prof. Jang-Hyun Kim (Sungkyunkwan Univ., ROK), Program Chairs Prof. Ji-Hwan Kim (Sogang Univ., ROK), Prof. Imran Ghani (Virginia Military Institute, USA). I also express my special thanks to the Keynote Speaker Dr. Jim Jansen (PhD) who is Qatar Computing Research Institute (QCRI), Hamad Bin Khalifa University, Doha, Qatar and Editor-in-chief, Information Processing and Management. In addition, my special thanks go to Advisory Board members, Program Committee members, Steering Committee members.

Last but not least, I would like to thank all authors and speakers again for their valuable research papers and efforts to make the 18th APIC-IST 2023 a success!
I wish a great success in their future endeavor and look forward to meeting you all again at ICONI 2023 that will be held in December.

Dr. Kwanghoon Pio Kim

Honorary Chair of APIC-IST 2023

A Message from Conference Chair



Dr. Jang-Hyun Kim
Sungkyunkwan Univ., ROK

Welcome to the 18th Asia Pacific International Conference on Information Science and Technology (APIC-IST 2023). which will be held on June 25-28, 2023, at Acros Fukuoka, Fukuoka city, Japan. APIC-IST is an annual conference organized by KSII providing an open forum for researchers, engineers, policy makers, and service providers in the Information Science and Technology.

In this year, we are glad to invite the prospective authors to submit oral & poster papers in the following areas:
Network/Internet of Things/Mobile Computing, Intelligent ICT Convergence/Smart Building, Big data platform /Analysis / Application,
Cloud Computing/Digital Twin/Smart Factory, Security System/ Information Security, Artificial Intelligence/Time Series Machine Learning,
Defense Information Technology/High Performance Computing, Cyber
Education/Social Media/Metaverse/UX/SNS, Scientific Public Security /
Crime Prevention/Intelligent CCTV Application, Technology Management
/e-commerce/Information System/ERP/BPM/DX, Graphics/Image Processing/Video Deep Learning/Video Control, Software High Growth Club, KETI Workshop and KDSA- Special Session.

As time goes on, annual APIC-IST conference becomes more qualified conference in which many honorable researchers participate to share their valuable research results. APIC-IST 2023 offer 1 invited talk, 19 sessions and a special workshop that present the latest results in various research fields of ICT (Information Communication Technology).

I must thank to all the authors and speakers for the research papers on their precious achievements. Fukuoka city is one of most popular tourist destinations in Japan. It is known as a beautiful and historical place in the world. I hope all participants of APIC-IST take this opportunity to explore and experience encounter beautiful nature simultaneously while staying in Fukuoka, Japan.

Dr. Jang-Hyun Kim

Conference Chair of APIC-IST 2023

A Message from Program Chairs

It is a pleasure to welcome you to the 18th Asia Pacific International Conference on Information Science and Technology (APIC-IST 2023). The APIC-IST 2023 will be held in Fukuoka, Japan on June 25-28. APIC-IST is one of the premier conferences in the Asia Pacific region. The hosting organization (KSII) is one of the largest ICT related academic societies in Korea.

Over the past 18 years, APIC-IST has grown to be one of the major international conferences in Information Science and Technology area.

APIC-IST aims to provide a high-quality forum for researchers and practitioners to share research findings, practices and ideas on key issues in ICT related technologies, services, and administrations. The APIC-IST 2023 Program Committee organizes 15 tracks and presentations of completed research papers as well as research-in-progress papers, panel discussion, and posters.

We very appreciate the talents, dedication and time of many reviewers from USA, Singapore, Canada, Vietnam, Pakistan, China and Republic of Korea. We would like to express special gratitude and appreciation to Dr. Kwanghoon Pio Kim, Dr. Jang-Hyun Kim and session chairs.

All submitted papers have gone through a fair paper review process and excellent 101 papers will be presented at APIC-IST 2023. We would like to thank to all the authors and speakers for submitting and presenting their valuable research works at APIC-IST 2023.



Dr. Ji-Hwan Kim
Sogang Univ.,ROK



Dr. Imran Ghani
Virginia Military Institute, USA

Dr. Ji-Hwan Kim & Dr. Imran Ghani
Program Chairs of APIC-IST 2023

Keynote Speaker



Keynote Speaker Dr. Jim Jansen (PhD)

Principal Scientist

Qatar Computing Research Institute (QCRI), Hamad Bin Khalifa Univer
Doha, Qatar

Editor-in-chief
Information Processing and Management

Title: Understanding Audiences, Customers, and Users via Analytics

Abstract of the talk:

Using digital analytics from the laboratory, social media platforms, and the Web is an industry-standard approach for understanding audiences, customers, and users. The people that assume these roles of audiences, customers, and users often have complex behaviors and motivations that are challenging to understand. The use of analytics can aid in this understanding. After conceptualizing the analytics process, this presentation will skim analytics foundations, including strategies, data-gathering approaches, and validation methods. We will also discuss specific methods, including Eye Tracking, Mouse Tracking, Social Media APIs, Search Logs, and Analytics Triangulation both in and outside the lab. We will also discuss the validity of online analytics. The presentation aims to offer a primer on contemporary lab, Web, and social media analytics topics.

Short Bio :

Dr. Jim Jansen is a Principal Scientist in the Social Computing Group of the Qatar Computing Research Institute leading a research team working on the automatic generation of personas (see https://persona.qcri.org/). He is a graduate of West Point and has a Ph.D. in computer science from Texas A&M University. Professor Jansen is editor-in-chief of the journal, Information Processing & Management (Elsevier), a member of the editorial boards of seven international journals, and former editor-in-chief of the journal, Internet Research (Emerald). He has received several awards and honors, including an ACM Research Award, six application development awards, and a university-level teaching award, along with other writing, publishing, research, teaching, and leadership honors. Dr. Jansen has authored or co-authored 350 or so research publications, with articles appearing in a multi-disciplinary range of journals and conferences. He is the author of the book, Understanding Sponsored Search: A Coverage of the Core Elements of Keyword Advertising (Cambridge University Press).

Organizing Committee

| Honorary Chair |

Kwanghoon Pio Kim, Kyonggi University, ROK

| Conference Chair |

Jang-Hyun Kim, Sungkyunkwan University, ROK

| Program Chairs |

Ji-Hwan Kim, Sogang University, ROK

Imran Ghani, Virginia Military Institute, USA

| Publication Chairs |

Jong-Moon Chung, Yonsei University, ROK

Minho Jo, Korea University, ROK

Moonseong Kim, Seoul Theological University, ROK

| Financial Chair |

Junho Ahn, Kyonggi University, ROK

| Workshop Chair |

Kiwon Kwon, KETI, ROK

| Steering Chairs |

Namgi Kim, Kyonggi University, ROK

Jongho Paik, Seoul Women's University**,** ROK

Organizing Committee

| Program Committee |

Jin Xiaolong, Chinese Academy of Sciences, China
Jiao Hao, Beijing Normal University, China
Jian Mou, Pusan National University, ROK
Junchul Chun, Kyonggi University, ROK
Seung Ryul Jeong, Kookmin University, ROK
Keon Chul Park, Advanced Institute of Convergence Technology, ROK
Woochan Lee, Incheon National University, ROK
JeongHyeon Chang, Kyonggi University, ROK
Jeongwook Seo, Hanshin University, ROK
Dohoon Kim, Kyonggi University, ROK
Kyungyong Chung, Kyonggi University, ROK
DongKyoo Shin, Sejong Univ., ROK
Hangbae Chang, Chung-Ang University, ROK
Hoill Jung, Daelim College, ROK
Jieun Han, Hanyang University, ROK
Kyung Han You, Jeonbuk National University, ROK
Sangmin Lee, CHA University, ROK
Woochun Jun, Seoul National University of Education, ROK
Yong-Suk Park, Korea Electronics Technology Institute, ROK
Yongwon Kim, Korea Expressway Corporation Research Institute, ROK
Younkyu Lee, Hongik University, ROK
Yoosi Kim, Airdeep Ltd., ROK
Jisu Kim, Singapore Institute of Technology, Singapore
Ogan Gurel, L'Institut des Sciences et Arts, Canada

Conference Program

Sun / Jun. 25, 2023	Room 1
17:00-19:00	VIP Reception & Greetings

Mon / Jun. 26, 2023	Room 1	Room 2	Room 3	Room 4
08:30-	Registration			
09:00-10:30	Session 1	Session 2	Session 3	Session 4
10:40-11:10	Opening Ceremony & Keynote Speech			
11:20-12:20		Session 5	Session 6	Session 7
12:20-14:00	Lunch (12:20-14:00)			
14:10-15:40	Session 8	Session 9	Session 10	Session 11
18:00-19:30	Award Ceremony & Banquet (The Luigans Spa and Resort)			

Tue / Jun. 27, 2023	Room 1	Room 2	Room 3	Room 4
09:00-10:30	Session 12	Session 13	Session 14	Session 15
10:40-12:10	Session 16	Session 17	Session 18	Session 19
12:10-14:00	Lunch (12:10-14:00)			
14:10-15:00	Panel Discussion : Collaborative Research			

Wen / Jun. 28, 2023	
10:00-13:00	Conference Special Program “teamLap Forest Fukuoka“ / paypaydom

Contents

1-1	Proposal of a Hardware-in-the-Loop Simulation System for Automatic Charging Systems for Battery Electric Buses **Koog-Hwan Oh, Hyun-Chang Cho (Korea Electronics Technology Institute, ROK)**	1-3
1-2	Antibiotic Risk Assessment System in Korean Agricultural Environment based on Monte Carlo Simulation **Han Jong Gil, Il Ryong Kweon, Min Ho Son (PODO Inc, ROK)**	4-6
1-3	A Haptic Data Glove for VR Interactions **Sang-Woo Seo(Electronics and Telecommunications Research Institute, ROK), Yejin Kim (Hongik Univ., ROK)**	7-8
1-4	An Off-Chain Methodological Approach for Ensuring Authenticity of Non-Fungible Tokens **Heesang Kim, Jaewook Koo, Dohoon Kim (KyonggiUniv., ROK)**	9-11
1-5	Automated Dysarthria Severity Classification Using Diadochokinetic test and Speech Intelligibility Based on LightGBM **MinSeo Kim, Junseok Oh, Ji-Hwan Kim (SogangUniv., ROK)**	12-13
1-6	UdiOlle: An OlleTrail Course Recommendation Mobile Application Reflecting User Preferences **Nayoung Kim, Dongho Yang (Jeju National Univ., ROK)**	14-16
1-7	An Algorithm based on Infant Abnormal Posture Detection through AI Pose Estimation **Jeonghyo Kim, Seungha Noh, Sunwoo Pi, Wonbeen Im, Chanui Hong, Sangjin Kim (KyonggiUniv., ROK)**	17-18
1-8	Privacy-Preserving DTxwith Differential Privacy **Geehee Yun, Kyoungjin Kim (SungshinWomen's Univ., ROK)**	19-20
1-9	A Research on the Highway New Business Models Corresponding to Changes in Future Transportation Environment **Yong-won Kim, Chan-bok Lee (Korea Expressway Corporation Research Institute, ROK)**	21-22

Contents

2-1	Shipboard Indoor Positioning: A Controllable Feature Selection for IQ Fingerprint Map of Bluetooth 5.1 **Qianfeng Lin, Jooyoung Son (Korea Maritime and Ocean Univ., ROK)**	23-25
2-2	A Visualization Method of Abnormal Sound Using Acoustic Measurement System Based on Socket Network Communication for Web Client Users **Jiin Shin, Suyoung Kim, Jaehyun Lee, Jongho Paik (Seoul Women's Univ., ROK)**	26-29
2-3	A Study on Vulnerability Analysis and Memory Forensics of ESP32 **Jiyeon Baek, Jiwon Jang, Seongmin Kim (SungshinWomen' s Univ., ROK)**	30-32
2-4	Generating multicast routing trees method in SDN using Pearl **Heechan Kim, Janggun Jeon, Yejin Lee, Namgi Kim (Kyonggi Univ., ROK)**	33-35
2-5	Implementation of Graph Neural Network for Link Prediction in Software-Defined Networks **Syed M. A. H. Bukhari, AfaqMuhammad, Wang-Cheol Song (JejuNational Univ., ROK)**	36-38
3-1	Road Hazardous Object Recognition System for the Visually Impaired **Sung-Hyun Kang, Yong-Jin Joo, Ji-Min Son, Yoon-Jung Yu, Sang-Jin Lee, Jun-ho Ahn (Kyonggi Univ., ROK)**	39-41
3-2	Enhancing Additive Manufacturing Processes with Machine Learning Algorithms: Challenges and Opportunities **Oakyoung Han (Sungkyunkwan Univ., ROK)**	42-45
3-3	Mutually Exclusive Decision-Making Learning of AI based on Contexts **Hyun Jung Lee (Goyang Research Institute, ROK)**	46-49

Contents

3-4	Retrieval-based Question-Answer Generation **Seonmin Koo, Jungwoo Lim, Suhyune Son, Yuna Hur, Heuiseok Lim (Korea Univ., ROK)**	50-52
3-5	Performance Analysis of Mixed Vibration Using RNN Models **Ju-Hyung Lee, Jun-Ki Hong (Kongju National Univ., ROK)**	53-55
4-1	Convergence of Blockchain and Artificial Intelligence: New Opportunities for Secure and Reliable Systems **Nakhoon Choi, Heeyoul Kim (Kyonggi Univ., ROK)**	56-58
4-2	Design of Insider Technology Leakage Sign Management System using Blockchain **Giwan Hong, Yuna Han, Woojoo Kwon, Hangbae Chang (Chung-Ang Univ., ROK)**	59-62
4-3	An Approach to Detect SMS Spam Using ChatGPT **Sooyon Seo (Sungkyunkwan Univ., ROK), Jaeho Hwang (Yonsei Univ., ROK), Moohong Min (Sungkyunkwan Univ., ROK)**	63-65
4-4	Nurturing Personal Information Protection Specialists: An Integrated Approach to Curriculum Development **Yeji Choi, Juno Lee, Hangbae Chang (Chung-Ang Univ., ROK)**	66-68
4-5	Cybersecurity Challenges in Blockchain-based Social Media Networks: A Review **Muhammad Hasnain (Lahore Leads Univ., Pakistan), Imran Ghani, David Smith (Virginia Military Institute, USA), Akbar Hussain (Lahore Leads Univ., Pakistan), Ali Daud (Abu Dhabi School of Management, United Arab Emirates), Sardar Usman (Grand Asian Univ., Sialkot, Pakistan), Seung Ryul Jeong (Kookmin Univ., ROK)**	69-71

Contents

4-6	Proposing Controls for a Secure M&A of National Core Technology with Smart Contract **Yurim Choi, Hangbae Chang (Chung-Ang Univ., ROK)**	72-74
4-7	A Study on Factors Derivation of Influencing Security R&D Programs for Economic Research Performance **Jawon Kim, Hangbae Chang (Chung-Ang Univ., ROK)**	75-77
5-1	Development and Application of Computer Science Textbooks: Focusing on Digital Therapeutics **Eunsun Choi, Namje Park (Jeju National Univ., ROK)**	78-81
5-2	Next Activity Prediction and Analysis Based Process Mining **Jin-Hyeong Lee, Eunbi Cho, Tsagaantsooj Batzaya, So-hyang Bak, Thanh-hai Nguyen, Kwanghoon Pio Kim (Kyonggi Univ., ROK)**	82-84
5-3	Similarity estimation between distributed inference environments using vision transformer embedding vectors **Jongbin Park (KETI, ROK)**	85-87
5-4	A framework to Support the Development of Safety Guidelines in the Manufacturing fields using Ontolog **Taehoon Kim, Jinmyeong Lee, Jungmin Yun, Jungnyun Lee (Korea Institute of Industrial Technology, ROK)**	88-90
5-5	Implementation of a Mixed Reality Service with Real-time Interaction of Virtual and Real Worlds **Jong-gak Choe, Yongjin Kwon (Korea Aerospace Univ., ROK)**	91-94
6-1	An Analysis on the Impact of Consumer Needs on Consumer Behavior: A Case Study based on Design of a Smart Wearable Product under E-Commerce **Jialei Ye, Xiaoyou He (Nanjing Univ. of the Arts, China), Ziyang Liu (Kyonggi Univ., ROK)**	95-97

Contents

6-2	Theorizing User' s Attitudes Towards AI-Based Systems by Using Philosophical Reasonings **Jongyoon Won, KC Lee (Sungkyunkwan Univ., ROK)**	98-101
6-3	A Web-based Business Process Discovery System **Thanh-Hai Nguyen, Sohyang Bak, Kyoung-Sook Kim (Kyonggi Univ., ROK), Dinh-Lam Pham (Vietnam National Univ., Vietnam), Kwanghoon Pio Kim (Kyonggi Univ., ROK)**	102-104
7-1	Pine Wilt Disease Detection using Synthetic Data and Object Detection Techniques **Yonghoon Jung, Wonseop Shin, Sanghyun Seo (Chung-Ang Univ., ROK)**	105-107
7-2	Automatic 3d Mesh Deformation Using Anime Character Image Information **Sanghyun Byun, Bumsoo Kim, Sanghyun Seo (Chung-Ang Univ., ROK)**	108-111
7-3	An Interdisciplinary Study on the Acoustic Perception Characteristics of TV Speakers: OLED Flat Panel Speaker and Dynamic Speaker Acoustic Focus **Kundong Choi, Sangwoo Hahm (Semyung Univ., ROK), Hyungwoo Park (Dong-Seoul Univ., ROK)**	112-114
7-4	Anomaly Detection from X-ray Images based on Region-Specific and Global PatchCore **Hyunbin Kim, Junchul Chun (Kyonggi Univ., ROK)**	115-117
8-1	Comprehensive Statistical Data Analysis for Production Decision in Aquaculture **Juhyoung Sung, Kyungwon Park, Byoungchul Song, Kiwon Kwon, Sungyoon Cho (Korea Electronics Technology Institute, ROK)**	118-120
8-2	Data communication dualization system between drone and GCS using RF and LTE **Wonseok Jung, Jong-Hong Park, Il-Yeop Ahn (Korea Electronics Technology Institute, ROK)**	121-122

Contents

8-3	Integrated Data Control Platform for Building a National Park Drone Sky Road **Jong-Hong Park, Wonseok Jung, Il-Yeop Ahn (Korea Electronics Technology Institute, ROK)**	123-124
8-4	Implementation of Centimeter-level Maritime Precise PNT Service Platform **Junsu Kim, Juhyoung Sung, Wongi Jeon, Kyungwon Park (Korea Electronics Technology Institute, ROK)**	125-126
8-5	Improvement of Ion Mobility Spectrometry System Performance through Structural Design Change of Drift Tube **Gi-Seop Im, Won-Young Lee, Sun-Nam Hwang (NewOne S&T Ltd, ROK)**	127-129
8-6	AI-enabled AGV service management with AIoT platform **SeungMyeong Jeong, Dae-Geun Heo, ChungJae Choe (Korea Technology Institute, ROK)**	130-131
8-7	Development of a drone-based low-light environment real-time crack detectio system **Sung-Chan Choi, Jong-Hong Park and Sung-Wook Jung, Il-Yeop Ahn (Korea Electronics Technology Institute, ROK)**	132-133
9-1	Structural relationships of various factors on the intention to use a mobile anger management training application: focusing on the mediating effect of anger management utility **Kyunghan You, Minhee Lee (Jeonbuk National Univ., ROK), Jaehee Cho (Sogang Univ., ROK), Sehwan Kim (Dankuk Univ., ROK)**	134-136
9-2	Investigation of Potential Relationships between Actual Behaviors of Using Smartphone Applications and Trait Anger **Jaehee Cho (Sogang Univ., ROK), Kyung Han You (Jeonbuk National Univ., ROK), Sewhwan Kim (Dankook Univ., ROK), Jin Xianmei (Sogang Univ., ROK), Dongbin Kim (Dankook Univ., ROK)**	137-139

Contents

9-3	Comparative Study on the Advertising Effectiveness of Real and Virtual Influencers on Healthcare Products : An Eye-Tracking Approach **Mengqi Du, Kyunghan Ryu (Jeonbuk National Univ., ROK)**	140-142
9-4	Design of Location-based Traveler-tailored Crime Information Fusion Application **Youjin Seo, Junho Ahn (Kyonggi Univ., ROK)**	143-145
9-5	Exploring the Relationship between Metaverse Usage Patterns and Cultural Tastes: A Study of Avatar Formation among Generation Z **Hyun-Ah Park, Kyung Han You (Jeonbuk National Univ., ROK)**	146-148
10-1	Multi Object Action Classification using Vision Transformer-based YOLO **Seo-El Lee, So-Hyun Park, Joo-Chang Kim, Sangeun Ahn, Kyungyong Chung, Hyun Yoo (Kyonggi Univ., ROK)**	149-151
10-2	An Architecture for Content Search in Video Retrieval System **Dinh-Lam Pham, Sang-eun Ahn, Kyunghee Sun, Kyong-Sook Kim, Kwanghoon Pio Kim (Kyonggi Univ., ROK)**	152-154
10-3	The Mediating Effect of Fear of Crime on Korean Women's Defensive Behaviors: The Necessity of CCTV **Sooyun Lee, Hyewon Jeon, Song Kang, Euigab Hwang (Kyonggi Univ., ROK)**	155-157
10-4	Measurement of Social Disorder based on Distance Between Objects in CCTV Video **Eunbi Cho, Thanh-Hai Nguyen, Jeong-Hyeon Chang, Tsagaantsooj Batzaya, Kwanghoon Pio Kim (Kyonggi Univ., ROK)**	158-160
10-5	Continual Learning Algorithms for Enhanced CCTV Effectiveness **YaeEun Lee, Hyunki Lim, JeongHyeon Chang (Kyonggi Univ., ROK)**	161-163
10-6	Action Classification using SlowFast and Automated Machine Learning **Gyu-Il Kim, Seong-Un Cho, Jeong-Hyeon Chang, Kyungyong Chung, Hyun Yoo (Kyonggi Univ., ROK)**	164-166

Contents

11-1	Machine Learning vs. Deep Learning on a Tabular Traffic Dataset in Big Data **Dalyapraz Manatova (Indiana Univ., USA), Jongwook Woo (California State Univ., Los Angeles, USA)**	167-169
11-2	Occlusion Object Model using PVT based Mask-RCNN **Ji-Won Baek, Ye Yeon Kang, Geon Park, Kyungyong Chung (Kyonggi Univ., ROK)**	170-172
11-3	Design and implementation of user participated SNS through voting function **JiHoon Lee, MoonSung Kim, Jinju Jeong, HeeYeon Cho, Sieun Park, Sangin Kim (Kyonggi Univ., ROK)**	173-175
11-4	Customer phase identification based on sequential permutation search with voltage time series similarity **Ji-Hyun Choi (Hanshin Univ., ROK), Haesung Lee, Han-Min Lim, Byungsung Lee (Korea Electric Power Research Institute, ROK), Hyun Ahn (Hanshin Univ., ROK)**	176-178
11-5	Blockchain-Based Platform for Secure and Patient-Centric Personal Health Record Management **Sangmin Lee (CHA Univ., ROK)**	179-181
11-6	Intermedia Network Agenda Setting of Environmental News in Korea and China: Focusing on The Relationship Between Media and Public Agenda **Yifan Wang, Kyung Han You (Jeonbuk National Univ., ROK)**	182-184
12-1	Development of comprehensive information service on plant viruses in domestic and international areas **Minho Son, Ilryong Kweon (PODO inc. ROK)**	185-187
12-2	Improving Blockchain Security with T-OTP-Based Node Access Control **Sungbeen Kim, Inwuk Yeo, Dohoon Kim (Kyonggi Univ., ROK)**	188-190
12-3	Web-based 3D Genome Visualization System for Whole Genome Alignments **Min Jae Kang, Hwan Gue Cho (Pusan National Univ., ROK)**	191-192

Contents

12-4	The Use of Grammar Transducer in Korean End-to-end Automatic Speech Recognition Models **Eunsoo Cho, Hosung Park, Hayoung Kim, Ji-Hwan Kim (Sogang Univ., ROK)**	193-195
12-5	Improving Conformer-CTC based Korean ASR Performance with Data Augmentation and Language Model **Yoonseo Chung, Hosung Park, Eunsoo Cho, Ji-Hwan Kim (Sogang Univ., ROK)**	196-197
12-6	Predicting Winding Machine Quality based on Deep Time-Series Data Analysis **Seung-Joon Lee (Korea Institute of Industrial Technology, ROK)**	198-200
12-7	A Study on Playbook Automatic Generation Method for Threat Analysis Automation **Seul-Ki Choi, Taeeun Kim, Kijong Son, Saewoom Lee, Taehyeon Kim (Korea Internet & Security Agency, ROK)**	201-202
12-8	Wearables for Physical Activity Verification **Ye Won Park, Yubin Yun, Hyoseok Yoon (Hanshin Univ., ROK)**	203-204
13-1	Beamforming with Generalized Power Iteration to Maximize Sum Rate in Maritime Relay Communications System **Tae-Woo Kim, Kyeongjea Lee(Yonsei Univ., ROK), Sung-yoon Cho, Ki-won Kwon(Korea Electronics Technology Institute, ROK), Dong Ku Kim (Yonsei Univ., ROK)**	205-207
13-2	Design and implementation of predictive system with AI/ML dataset management supported IoT platform **Dae-Geun Heo, Seung-Myeong Jeong (Korea Electronics Technology Institute, ROK)**	208-210
13-3	A study on the risk management plan of domestic coastal fishing vessels using big data **Hyunjeong Lim (Korea Maritime Transportation Safety Authority, ROK), Sooyeon Kwon (Inha Univ., ROK)**	211-213

Contents

13-4	Data system for Fishing Vessel Design **Dongheum Ryu, Yongbin Lee, Changhyun Park (PIDOTECH Inc., ROK), Sooyeon Kwon (Korea Maritime Transportation Safety Authority, ROK)**	214-215
13-5	Aids to Navigation (AtoN) Trend Analysis using LDA Topic Modeling **Narae Yu, Jooyoung Lee (Korea Science Policy Platform, ROK)**	216-218
13-6	Expanding Communication Coverage in Maritime Communications: Outage Capacity Analysis of two-Hop Relay and Beamforming **Kyeongjea Lee, Tae-Woo Kim (Yonsei Univ., ROK), Sung-yoon Cho, Ki-won Kwon (Korea Electronics Technology Institute, ROK), Dong Ku Kim (Yonsei Univ., ROK)**	219-221
14-1	JenTrip: A Travel Platform for Solo Travelers **Honglae Kim, Seyoung Kim (HUPLE Co., Ltd., ROK)**	222-224
14-2	Developing Travel Information Systems Using Accessability Index and VR **Eugene Choi (Kangnam Univ., ROK)**	225-226
15-1	Rethinking the Locus and Focus of Citizen Participation: Role of UX and UI for Urban Planning Application **Woomin Nam, DaeHwan Kim, Keon Chul Park (Advanced Institute of Convergence Technology, ROK)**	227-229
15-2	Where and Why? A Novel Approach for Prioritizing Implementation Points for Public CCTVs using Urban Big Data **Daehwan Kim, Keon Chul Park (Advanced Institute of Convergence Technology, ROK)**	230-232
16-1	Leverage time-based data to improve recommendation system accuracy **Sangmin Lee, Jeon Janggun, Lee Yejin, Namgi Kim (Kyonggi Univ., ROK)**	233-235
16-2	Speech Intelligibility Prediction of Dysarthria Using Deep Convolutional Networks **Junseok Oh, Hosung Park, Ji-Hwan Kim (Sogang Univ., ROK)**	236-237

Contents

16-3	Analysis of H.266 Versatile Video Coding Streaming and Decoding on Smartphones **Youngwook Kim, Jong-Moon Chung (Yonsei Univ., ROK)**	238-239
16-4	Electoencephalography-based Fatigue Assessment for Internet of Medical Things System **Jung-Hyok Kwon, Sol-Bee Lee, Eui-Jik Kim (Hallym Univ., ROK)**	240-241
16-5	A Study on Approach for Automating Field Identification in Heterogeneous Security Events Analysis **Taehyeon Kim, Kijong Son, Saewoom Lee, Seul-Ki Choi, Taeeun Kim (Korea Internet & Security Agency, ROK)**	242-243
16-6	Cross-Device Extended Reality Experiences **Sihyeon Paik, Jiseong Park, Seoi Na, Hyoseok Yoon (Hanshin Univ., ROK)**	244-245
16-7	Air Care IoT Platform for Air Quality and Energy Saving **Yoosin Kim (AirDeep Inc. ROK)**	246-247
16-8	Factorized 3D CNNs based Human Action Recognition **Seho Park, Taehyeon Kim (Korea Electronics Technology Institute, ROK)**	248-249
17-1	The Design of IT Project Community Service Platform for Exploring Human Resources Using Recommendation Algorithm **Yoonjung Jang, Yumin Jo, Hyebin Lee, Jiyeon Jeong, Jongho Paik, Jun Hwang (Seoul Women's Univ., ROK)**	250-253
17-2	Avatar-Chatbot Making for High School-College Major Experience Program **Tai Hoon KIM, In Seop NA (Chonnam National Univ., ROK)**	254-256
17-3	A Deep Learning Method for IQ Fingerprint Map Based Indoor Positioning using Bluetooth 5.1 Technology in Ship Environments **Yootae Shin, Jooyoung Son, Qianfeng Lin (Korea Maritime and Ocean Univ., ROK)**	257-259

Contents

17-4	An Empirical Study on the Length Granularity in Long-Form Question Answering **Hyeonseok Moon, Jinsung Kim, Kinam Park, Jeongbae Park, Heuiseok Lim (Korea Univ., ROK)**	260-263
17-5	A Vision-based Unmanned Store System Fusing User Behavior Patterns **Yeongjun Choi, Harin Jo, Jisoo Cheon, Yunkyung Hwang, Junho Ahn (Kyonggi Univ., ROK)**	264-266
17-6	Solving University Course Scheduling Problems Using Adaptive Parallel Tabu Search and Simulated Annealing **Xiaorui Shao, Chang Soo Kim (Pukyong National Univ., ROK)**	267-270
18-1	Workshop for Efficient NLP Analysis Using TEXT for Korean, English, & Chinese **Taehwan Rhee, Chaenam Jun (The IMC)**	-
19-1	Revolutionizing ESG Evaluation: A Study on the Potential of Earnings Call Transcript Data **Haein Lee, Hae Sun Jung, Heungju Park, Jan Hyun Kim (Sungkyunkwan Univ., ROK)**	271-274
19-2	LSTM-based Prediction of Bitcoin Prices Using Technical Indicators, Sentiment Analysis, and Google Trends **Hae Sun Jung, Haein Lee, Jang Hyun Kim (Sungkyunkwan Univ., ROK)**	275-278
19-3	Public Opinion Analysis of E-sports in China Based on LDA Model: A Case Study of EDG Winning the Championship **ShaoPeng Che, Kai Kuang (Tsinghua Univ., China)**	279-282
19-4	Ontological metaphor in human-computer interaction: a study of the user's conceptual perception on machines **Youngsang Kim (Sungkyunkwan Univ., ROK)**	283-286
19-5	The application of affordance theory in understanding social media users' perception of algorithms **Jisu Kim (Singapore Institute of Technology, Singapore)**	287-289

Copyright ⓒ 2023 KSII

Proposal of a Hardware-in-the-Loop Simulation System for Automatic Charging Systems for Battery Electric Buses

Koog-Hwan Oh and Hyun-Chang Cho*
Smart Electrics Research Center, Korea Electronics Technology Institute
Gwangju, South Korea
[e-mail: ohkhwan@keti.re.kr, hccho@keti.re.kr]
*Corresponding author: Koog-Hwan Oh

Abstract

This paper proposes a hardware-in-the-loop (HIL) simulation system for automatic charging systems for battery electric buses (BEBs). In this paper, several types of automatic charging system for the BEBs are briefly introduced and a simulating system applicable to them, which estimates the state information of them in accordance with the operation scenarios inputted by users such as the power consumption, the electric fee, the charging time, etc., is proposed based on the HIL simulation.

Keywords: HIL Simulation, Automatic Charging System, Battery Electric Bus

1. Introduction

Adopting the Paris Agreement in December 2015, the ecofriendly means of transportation has received considerable interest. This trend makes not only private cars but also public buses rapidly electrified. Several automatic charging systems for battery electric buses (BEBs) have actively been researched and developed for efficient operation of a BEB fleet. However, the initial cost to construct them is still high and it is not easy for users to choose a suitable type of automatic charging system according to routes or operating hours of their own BEB fleets. If the users can simulate the operation of the automatic charging systems, they can be easier to choose a type of the charging system than without the simulation. It can also offer the development convenience for developers of the charging systems. So, this paper proposes a hardware-in-the-loop (HIL) simulation system for the automatic charging systems for BEBs.

2. Proposed system

2.1 Automatic Charging Systems for Electric buses

This paper briefly introduces several types of automatic charging system for the BEBs, prior to propose the HIL simulation system. Most BEBs have been used a plug-in charging type for depot charging in common with other electric vehicles. The plug-in charging method has a disadvantage that a person is required to connect and disconnect the charger to the BEB when the charge begins and ends. So, several automatic charging systems for BEBs have been implemented such as pantograph charging system (PCS), battery swapping system (BSS), and wireless charging system (WCS) [1]. The PCS quickly charges a BEB through the physical connection between the charger and the roof of the BEB using a conductive element at stop.

This work was supported by the Korea Institute of Energy Technology Evaluation and Planning(KETEP) grant funded by the Korea government(MOTIE) (20212020800110, Development and demonstration of a transfer type automatic charging system for electric bus connection with roof connection method)

There are several configurations for this system and the most common are the pantograph-up and the pantograph-down according to the location of the pantograph [2]. The BSS charges a BEB which has the swappable battery system through automatically replaces a discharged battery with a fully charged one [3]. The WCS transfers electric power to a BEB through the wireless power transfer system installed under the road surface [4].

2.2 Proposed HIL Simulation System

The automatic charging systems mentioned above should include additional modules for automatically charging the BEBs without a person who connects and disconnects the charger to the BEBs. For instance, the modules which automatically transfer the pantographs and the swappable batteries are required to charge the BEBs in the case of PCS and BSS respectively. All automatic charging systems mentioned above might also involve the modules to distinguish whether the BEB is placed in proper position for automatically charging. These additional modules for automatically charging should be controlled by one or more control modules. And there should be a main control system that controls and manages all additional modules and their control modules. So, the main control system should be considered to simulate automatic charging systems in their development and real-time testing. This paper proposes a HIL simulation system including the main control system and **Fig. 1** shows its block diagram.

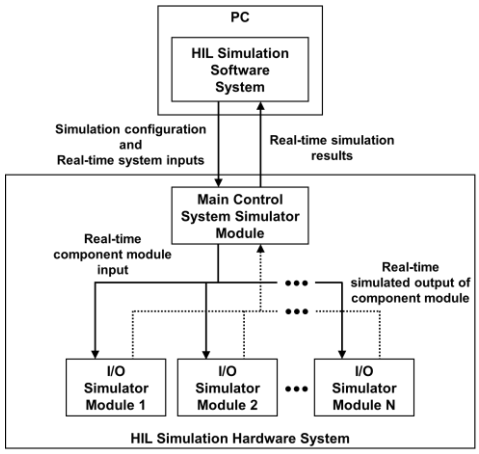


Fig. 1. Block diagram of the proposed HIL simulation system

The proposed HIL simulation system can be divided into two main systems: HIL simulation software system (HILSSS) and HIL simulation hardware system (HILSHS). The HILSSS is an application software for PC and the HILSHS consists of a main control system simulator module (MCSSM), which is identical to the main control system of an automatic charging system, and one or more I/O simulator modules (IOSMs). The number of IOSMs is determined by the number of modules that are controlled by the main control system in the automatic charging system. Each IOSM has a microcontroller unit and simulates input and output signals of different module controlled by the main control system. It also has an identical I/O interface between the main control system and the module which it simulates. The proposed system performs the following steps: (1) the user inputs an operation scenario of the charging system into the HILSSS, then the HILSSS generates a system input sequence for the MCSSM in accordance with the operation scenario and transmits each system input in real-time to the MCSSM at the appropriate time; (2) the MCSSM identically operates as the main control system of charging system in accordance with the system inputs from HILSS and transmits input signals, which are equal to input signals of the component modules in the charging system, to IOSMs in real-time; (3) each IOSM simulates output signal of the component modules and transmits the simulated output to the MCSSM; (4) the MCSSM operates in accordance with the simulated outputs; (5) the HILSSS finally monitors and estimates the state information of the charging system such as the SOC of the BEBs, the power consumption, the electric fee, the charging time.

3. Conclusions

In this paper, several types of automatic charging system for the BEBs are briefly introduced. We propose a HIL simulation system including simulator modules that simulates main control system and other component modules in the automatic charging systems. Nowadays we have been developing a HIL simulation system for the electric bus pantograph charging system. Our future works include expansion of our system into other automatic charging systems.

References

[1] Clairand, J. M., Guerra-Terán, P., Serrano-Guerrero, et al., "Electric Vehicles for Public Transportation in Power Systems: A Review of Methodologies," *Energies*, Vol 12, No. 16, p. 3114, 2019.

[2] López, I., Calvo, P. L., Fernández-Sánchez, G., et al., "Different Approaches for a Goal: The Electrical Bus-EMT Madrid as a Successful Case Study," *Energies*, Vol 15, No. 17, p. 6107, 2022.

[3] Ahmad, F., Saad Alam, M., Saad Alsaidan, I., et al., "Battery swapping station for electric vehicles: opportunities and challenges," *IET Smart Grid*, Vol 3, No. 3, pp. 280-286, 2020.

[4] Luo, X., and Fan, W., "Joint design of electric bus transit service and wireless charging facilities," *Transportation Research Part E: Logistics and Transportation Review*, Vol 174, p. 103114, 2023.

Copyright ⓒ 2023 KSII

Antibiotic Risk Assessment System in Korean Agricultural Environment based on Monte Carlo Simulation

Han Jong Gil[1*], Il Ryong Kweon[1] and Min Ho Son[1]
Digital Platform HQ, PODO Inc.[1]
Seongnam-si, Gyeonggi-do, Republic of Korea
[e-mail: hjgil@ipodo.co.kr, irk@ipodo.co.kr, mhson@ipodo.co.kr]
*Corresponding author: Han Jong Gil

Abstract

Overseas, antibiotics are recognized as a major threat to human health. Identifying potential sources and transmission routes in the agricultural environment, and conducting research on the impact assessment on the human body and the transfer process. In Korea, the demand for systematic management of antibiotic information is increasing. For effective management and communication of antibiotics that can be exposed through various channels, a system that delivers and manages risk-based information is essential. Through this study, an antibiotic risk assessment system is established and a web-based app is created that reflects result of this study. The app manages various data related to antibiotics and performs risk assessment based on them. In addition, GIS is used to visualize the results on the map so that they can be easily recognized visually. We implement a chi-square test to determine the suitability of sample data and a Monte Carlo simulation function to predict the risk distribution based on it.

Keywords: Web-based, Antibiotics, Statistics, Chi Square, Monte Carlo, Geoserver, GIS, Java, Spring

1. Introduction

As it becomes a major threat to human health, researches on antibiotics are increasing. Such studies are focusing on potential sources of antibiotics and transmission routes in the environment, evaluating the effects of antibiotics on the human body, and conducting research on migration pathways and transfer processes.

Likewise, research on antibiotics is increasing in Korea, and accordingly, the need for systematic management of antibiotic information is emerging. Statistics on the production and sales of domestic agricultural antibiotics are reported by the Korea Crop Protection Association annually, but they are classified only for limited purposes. There are no statistics for usage on each crop or risk based assessment.

It is planning to study on the use of Antibiotics for agriculture, establish of residual DB information system, establish of a risk assessment system for each exposure route and substance of antibiotics and develop integrated evaluation techniques based on antibiotic resistance index and risk index through this research project.

2. Objective

This research project establishes a risk assessment system through three years of research and develops the system as a result.

This study was carried out with the support of the Rural Development Administration's Agricultural Environment Antibiotics Risk Assessment System Establishment Research Project (PJ015958).

Year 1 (2021)	• Set of soil ecosystem risk indicators(for micro organisms and resistant genes) for organic matter and agricultural antibiotics • Calculation of the Predicted Exposure Concentration (PEC) of antibiotics in agricultural soil, crops, and river water • Securing land and aquatic toxicity data for antibiotics and evaluating the risk effects on land and living organisms • Establishment of a database on the status of agricultural antibiotics use and residual information
Year 2 (2022)	• Verifying of soil ecosystem risk indicators (for micro organisms and resistant genes) for organic matter and agricultural antibiotics • Investigation of the risk effects of land and aquatic creatures on antibiotics and the study of changes of microbial bacteria in the biological intestinal tract • Development of risk assessment formulas for antibiotics by exposure route and substance • Design integrated antibiotics risk assessment system and development prototype of the system
Year 3 (2023)	• Selection of final soil ecosystem risk indicators for antibiotics and establishment of soil ecosystem risk diagnosis techniques • Establishment techniques for assessing risk effects of terrestrial and aquatic organisms on antibiotics • Integration with agricultural hazard assessment system • Completion of the development of an integrated antibiotic risk assessment system and transfer to research institutes

Table 1. Annual Objective

3. Risks Assessment

The main pollutant medium is largely divided into three types: agricultural products, soil, and agricultural water. Detailed exposure pathways are divided into ingestion/inhalation/contact according to direct/indirect exposure. In this system, a total of 12 exposure routes are set and risk assessment is performed.

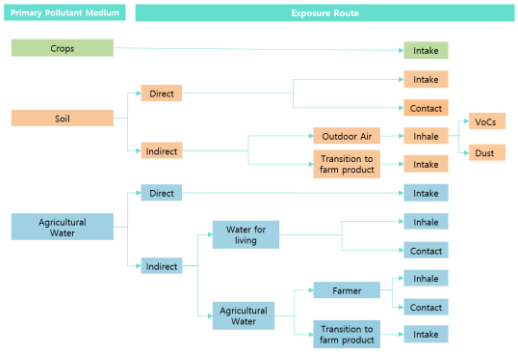


Fig. 1. Exposure Routes in Evaluation

Category		Assessment Formulas
Crops	Intake	$ADD = \frac{(CR_p \times Cs + CF_1 \times CR_L) \times ABS_{GI} \times EF \times ED}{BW \times AT}$
Soil	Intake	$ADD = \frac{CR_s \times CF_1 \times ABS_{GI} \times C_s \times EF \times ED}{BW \times AT}$
	Contact	$ADD = \frac{C_s \times (AF \times ABS_D \times CF_1) \times SA_s \times EF \times ED}{BW \times AT}$
	Volatile inhalation in outdoor air	$ADE(mg/m^3) = \frac{C_s \times EF \times ED \times 1/VF}{AT}$
	Inhalation of scattering dust in outdoor air	$ADE(mg/m^3) = \frac{C_s \times ABS_{INH} \times (TSP \times frs \times Fr) \times EF \times ED \times CF_2}{AT}$
	Intake of soil-derived crops	$ADD = \frac{(BCF_{s-plant} \times CR_p + CR_s \times CF_1) \times ABS_{GI} \times C_s \times EF \times ED}{BW \times AT}$
Agricultural Water	Intake	$ADD = \frac{C_W \times CR_W \times EF \times ED}{BW \times AT}$
	Water for living – Inhale	$LADD_{Inhal} = \frac{BR \times ET \times EF \times ED}{BW \times AT} \times VF \times C_{GW}$
	Water for living – Contact	$LADD_{Derm} = \frac{SA \times SP \times ET \times EF \times ED}{BW \times AT} \times C_{GW}$
	Water for agriculture – Inhale	$LADD_{Inhal} = \frac{BR \times ET \times EF \times ED}{BW \times AT} \times VF \times C_{GW}$
	Water for agriculture – Intake	$LADD_{Derm} = \frac{SA \times SP \times ET \times EF \times ED}{BW \times AT} \times C_{GW}$
	Intake of water-derived crops	$ADD = \frac{(BCF_{w-plant} \times CR_p + CR_w \times CF_1) \times ABS_{GI} \times C_w \times EF \times ED}{BW \times AT}$

Fig. 2. Risks Assessment

An average value of data is used when evaluating point evaluation. When performing distribution prediction, this system uses distribution of data (norm, lognorm, etc.) and simulates 10K times using Monte-Carlo method. In case of farm product, exposure prediction of distribution is determined by weight of consumer(BW), amount of intake(CR_P), residue level of antibiotics(HR). User can select appropriate data distribution for the prediction.

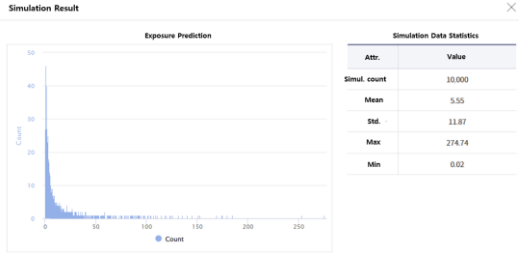


Fig. 3. Exposure Simulation

4. System Configuration

The system operates on a web-based basis and is configured as a server/client side. On the server side, the main app is implemented using a Java-based Spring Boot framework. It interfaces with DBMS using JPA. The DBMS uses Postgres and GIS Server uses Geoserver to express Geospatial data. The user may check the risk information through a web page using OpenLayers on the client side.

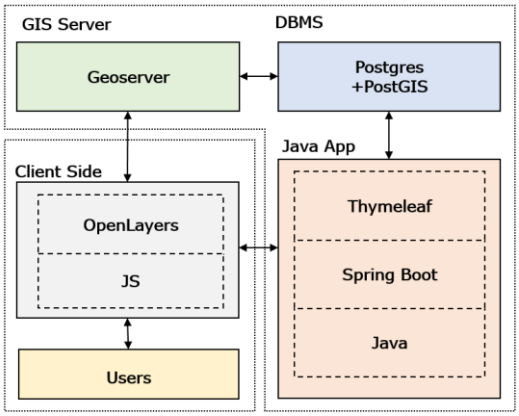


Fig. 4. System Architecture

Each component is distributed as a docker container to operate individually. It is configured as a docker compose so that it can be used as a single package in the same environment.

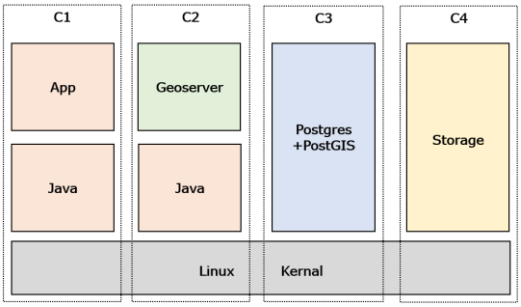


Fig. 5. Docker Container Configuration

5. Conclusions

Through this research project, the foundation for the Korean antibiotic risk assessment system was established, and it was reflected in the assessment web system so that a meaningful evaluation could be made. By developing Korean risks assessment techniques and indexes, the establishment of a pesticide safety management system is promoted. It is expected that the developed risk management system will contribute to minimizing the cost of safety management in agricultural environment by expanding the use of related ministries.

The goal of future research is to improve the process of risks assessment that selecting the distribution of data directly by users. An algorithm that automatically selects an appropriate distribution through MSE value comparison will be developed and applied.



Fig. 6. Main Web Page



Fig. 7. Risks Assessment Section

References

[1] Young Han Kim, “Java ROM Standard JPA Programing,” *aCORN Publishing Co.,* 2015.

[2] Asa Shiho, “Programmer no tameno Docker Kyoukasho dai2han,” *SHOEISHA Co., Ltd .,* 2018.

[3] The Ministry of Information and Communication, “One Health Antibiotic Resistance Bacteria Multi-Ministry Joint Response Project Report,” 2018.

[4] Thomas M. Carsey and Jeffrey J. Harden, “Monte Carlo Simulation and Resampling Methods for Social Science,” *SAGE Publications Inc.,* 2013.

Copyright ⓒ 2023 KSII

A Haptic Data Glove for VR Interactions

Sang-Woo Seo[1] and Yejin Kim[2*]
[1] Creative Content Research Division, Electronics and Telecommunications Research Institute
Daejeon, 34129 - Republic of Korea
[e-mail: swseo@etri.re.kr]
[2] School of Games, Hongik University
Sejong, 30015 - Republic of Korea
[e-mail: yejkim@hongik.ac.kr]
*Corresponding author: Yejin Kim

Abstract

This paper introduces a novel device that can track hand motion and provide haptic force feedback for VR interactions in real time. The proposed device estimates the 3D positions of hands by using a combination of the position estimation of ultrasonic sensors and the 3D acceleration and pose estimation of inertial sensors. The multichannel ultrasonic sensors in the 2D plane are used to localize the 3D positions of hands at a relatively low speed while a set of inertial sensors on hands are used to accurately estimate the hand positions at a relatively high speed. The proposed device provides haptic force feedback using a set of actuators and overcomes an overheat problem of the actuators by adopting an efficient power control. We present a capability of the proposed device in the VR training application such that a user can practice handling a real fire extinguisher in an immersive environment.

Keywords: Data glove, hand motion, haptic force feedback, inertial sensor, ultrasonic snesor, source localization

1. Introduction

From design to deployment, VR training applications are developed from the specific learning purposes. For this reason, previous systems [1] modified hand-held VR controllers or built special tools for user interactions during training. On the other hand, the data gloves are used to track the hand motion and provide user interactions with virtual objects in a precious way [2,3]. However, these commercial devices are lack of capturing physical phenomena such as the feeling of gripping a real object through force feedback. Moreover, they require an external tracker with base stations to recognize hand motion in a 3D space. In contrast to the development of tracking methods with visual feedback, studies on haptic force feedback for VR interactions are scarce.

In this paper, we propose a data glove that can track hand motion and provide haptic force feedback for VR interactions in real time. The proposed device estimates the 3D positions of hands by using a combination of the position estimation of ultrasonic sensors and the 3D acceleration and pose estimation of inertial sensors. The multichannel ultrasonic sensors in the 2D plane are used to localize the 3D positions of hands at a relatively low speed while a set of inertial sensors on hands are used to accurately estimate the hand positions at a relatively high speed. The proposed device provides haptic force feedback using a set of actuators and overcomes an overheat problem of the actuators by adopting an efficient power control. We present a capability of the proposed device in the VR training application such that a user can practice handling a real fire extinguisher in an immersive environment.

This work was supported by a grant (2019-MOIS34-001) from Preventive Safety Service Technology Development Program funded by Korean Ministry of Interior and Safety (MOIS) and a grant from the National Research Foundation of Korea (NRF) funded by the Korea government (MSIT) (No. 2021R1F1A1046513).

2. Design and Implementation

2.1 Hand position estimation

To measure the 3D positions of an object, we applied the ultrasound-based position estimation and inertial sensor-based 3D acceleration and pose estimation. However, there are well-known limitations on these methods such that the ultrasound-based device suffers from the high computational cost while the drifting problem makes the inertial measurement unit (IMU) sensors difficult to estimate the 3D positions continuously. For these reasons, we combined the two methods into one system that integrates the data of acceleration and rotation estimated by the inertial sensor and the data of position and velocity estimated by the ultrasonic sensor into the output data using the Kalman filter.

2.2 Haptic force feedback

In the proposed device, we introduce haptic force feedback using shape-memory alloy (SMA) springs that are lightweight to wear and compatible with the hand motion estimation. Unlike the motor-based kinesthetic feedback, a major challenge in the SMA-based device is the low cooling rate of the heated springs when kinesthetic senses are transmitted to users. Generally, seconds to several tens of seconds are required to cool down the springs depending on the activation and surrounding conditions. We adopted a flexible thermoelectric device [4] and thermal grease with a high thermal conductivity to resolve the cooling problem. Furthermore, the proposed device applies the SMA-based actuators to the hand dorsal and palm to provide bidirectional feedback.

3. Experimental Results

Figure 1 shows the prototype of the data glove implemented and compares the performance of tracking hand positions between the proposed device and the VIVE tracker [5]. The total distance offset of the proposed device compared to the tracker was measured less than 5 cm without a drift problem. **Fig. 2** shows the application of the proposed device to a virtual training application that exercises the use of a fire extinguisher. Here, various kinesthetic feedbacks are provided such as squeezing and releasing the handle of the fire extinguisher.

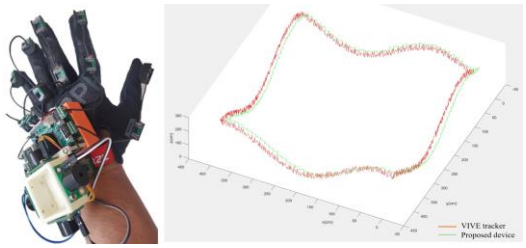


Fig. 1. Prototype of the data glove and its tracking accuracy



Fig. 2. VR training application

4. Conclusions

In this paper, we introduced a novel device that can track 3D hand motion and provide haptic force feedback for VR interactions. Unlike previous systems, the proposed device does not rely on an external tracker to recognize various hand poses. Currently, we are conducting a user test to evaluate the effectiveness of the proposed device in the virtual training programs.

References

[1] B. Xie, H. Liu, R. Alghofaili, Y. Zhang, Y. Jiang, F.D. Lobo, C. Li, W. Li, H. Huang, M. Akdere, C. Mousas, and L.-F. Yu, "A Review on Virtual Reality Skill Training Applications," *Frontiers in Virtual Reality*, Vol 2, 645153, 2021.

[2] VRgluv. Available online: https://kickstarter.com (accessed on 1 March 2023).

[3] Teslasuit Glove. Available online: https://teslasuit.io (accessed on 1 March 2023).

[4] TEGWAY Available online: http://tegway.co/tegway/ (accessed on 1 March 2023).

[5] VIVE Tracker. Available online: https://www.vive.com/kr/accessory/tracker3/ (accessed on 1 March 2023).

Copyright ⓒ 2023 KSII

An Off-Chain Methodological Approach for Ensuring Authenticity of Non-Fungible Tokens

Heesang Kim, Jaewook Koo, and Dohoon Kim*
Department of Computer Science, Kyonggi University
Gyeonggi, South Korea
[e-mail: skewed@kyonggi.ac.kr, paeter3@naver.com, karmy01@kyonggi.ac.kr]
*Corresponding author: Dohoon Kim

Abstract

In recent years, the use of Non-Fungible Tokens (NFTs) has gained popularity in various industries, such as art, music, and sports. However, there have been concerns about the authenticity and ownership of NFTs, which are traded on the blockchain network. This research paper proposes an off-chain methodological approach to ensure the authenticity of NFTs. The proposed approach utilizes off-chain storage to store additional information about NFTs that cannot be stored on the blockchain. This approach will provide an additional layer of security and protection against fraud and counterfeiting. The proposed methodology has the potential to significantly improve the credibility and trust of NFT transactions, thereby increasing the adoption of NFTs in various industries.

Keywords: Blockchain, Non-Fungible Tokens (NFTs), Authenticity, Off-Chain, Methodological Approach, Fraud Prevention

1. Introduction

The use of Non-Fungible Tokens (NFTs) [1-3] has significantly increased in recent years, providing a unique digital ownership experience in various industries such as art [1,3,4], music [5,6], and sports [3,7,8]. However, the trustworthiness and authenticity of NFTs remain a major concern [9-14], as they rely on the blockchain network, which is vulnerable to various cyber-attacks and security breaches. To address this concern, the proposed research paper aims to explore the utilization of an off-chain database approach to ensure the exact NFTs recorded on the blockchain. The proposed approach entails storing additional information about NFTs in a separate off-chain database, which is then referenced by the blockchain. This approach ensures that any tampering with the NFTs' records is easily detectable and can be prevented, ensuring the authenticity and legitimacy of the NFTs' ownership and transactions.

The proposed approach provides a viable solution to the inherent limitations of the blockchain in recording vast amounts of information and reduces the need for a high degree of complexity in smart contract programming. By incorporating an off-chain database, the proposed approach offers an additional layer of security and flexibility, as the database can be designed to allow for easy retrieval and management of NFT information. Ultimately, this research will contribute to the improvement of the credibility and trustworthiness of NFT transactions, which will enhance the adoption of NFTs in various industries and revolutionize the digital

"This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)"(2021-0-01393)

ownership experience.

2. Methodological Approaches

2.1 NFT Warranty Modelling

To link on-chain NFTs and off-chain users to an oracle for recording purposes, we propose the steps below:

First, we create an off-chain user account. We need to create an off-chain user account for each user who wants to link their NFT to the oracle. This account will be used to store the user's data and any relevant information about the NFT.
Then we generate a unique identifier. To link the user and their NFT to the oracle, we need to generate a unique identifier. This identifier can be a hash of the user's account and the NFT's metadata. We can use a hashing algorithm like SHA-256 to generate this identifier.
We create an Oracle Contract. An oracle contract on the blockchain will be responsible for receiving and processing data from off-chain sources, such as the user accounts and NFT metadata.
Thirdly, we link NFT to Oracle Contract. We can link the NFT to the oracle contract by calling a function in the contract that takes the unique identifier generated in step 2 as input. This function will store the NFT's metadata and link it to the unique identifier.
We also link user to Oracle Contract. Similarly, we can link the user to the oracle contract by calling a function in the contract that takes the user's account and the unique identifier as input. This function will store the user's data and link it to the unique identifier.

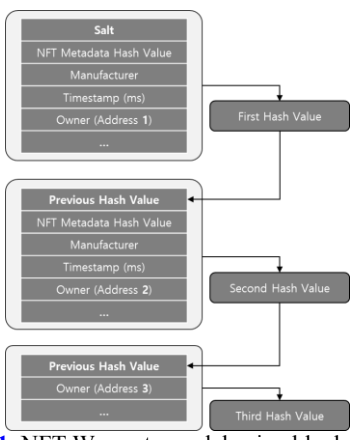


Fig. 1. NFT Warranty model using blockchain.

To ensure the integrity of the data, we can add a verification step where the oracle contract checks the authenticity of the data before storing it on the blockchain. For example, we can use a digital signature to verify that the data is coming from the correct source.
Finally, once the data has been verified, the oracle contract can record it on the blockchain (**Fig. 1**). This will ensure that the data is immutable and tamper-proof.

By following the above steps, we can link on-chain NFTs and off-chain users to an oracle for recording purposes. This will provide a guarantee of the authenticity and integrity of the data, which can be useful for various applications such as digital asset management, supply chain management, and identity verification.

2.2 Using an Off-chain Method

The main idea of linking a database system (off-chain) with a ledger (on-chain) to certify the authenticity of NFTs is to create a system that ensures that the NFTs recorded on the ledger are accurate and can be traced back to their original owners. This can be achieved through the use of a unique identifier that is assigned to each NFT, which can be recorded on both the ledger and the database system.
The database system can be used to store additional information about the NFTs, such as their provenance, ownership history, and any other relevant details. This information can be linked to the unique identifier recorded on the ledger, creating a secure and transparent system that can be used to verify the authenticity of the NFTs.
To ensure the accuracy and security of the system, it is important to implement appropriate security measures, such as encryption and access controls, to protect the data stored in the database system. Additionally, regular audits and checks should be conducted to ensure that the data recorded on the ledger and the database system match and that there are no discrepancies.
By linking the database system and the ledger, it is possible to create a system that provides a secure and transparent way to verify the authenticity of NFTs, making them more valuable and trustworthy in various industries such as art, music, and sports.

3. Conclusion

In conclusion, the use of Off-Chain databases in conjunction with On-Chain NFTs offers a viable solution to the authenticity and trustworthiness challenges posed by blockchain networks. The Off-Chain databases are used to store data that is not suitable for blockchain storage, and users can verify the authenticity of the digital assets, thereby preventing duplication or counterfeit. By providing a secure storage solution, the Off-Chain databases prevent unauthorized access and tampering.

Despite the potential benefits, the use of Off-Chain databases presents several challenges, such as security vulnerabilities and scalability issues. To mitigate these challenges, researchers have proposed the use of decentralized and distributed Off-Chain databases that leverage blockchain technology to ensure a secure and tamper-proof storage solution.

The continuous development of blockchain technology is expected to provide more advanced and secure methods for NFTs in various industries. Further research is necessary to explore the use of Off-Chain databases and their potential to solve digital ownership challenges.

References

[1] Kugler, Logan. "Non-fungible tokens and the future of art." Communications of the ACM 64.9, 19-20, 2021.

[2] Chohan, Usman W. "Non-fungible tokens: Blockchains, scarcity, and value." Critical Blockchain Research Initiative (CBRI) Working Papers, 2021.

[3] Wang, Qin, et al. "Non-fungible token (NFT): Overview, evaluation, opportunities and challenges," arXiv preprint arXiv:2105.07447, 2021.

[4] Samudra, Aparna. "Non-Fungible Tokens (NFTs): New Renaissance for the Artists." Available at SSRN 4112932, 2022.

[5] Folgieri, Raffaella, Paul Arnold, and Alessandro Giuseppe Buda. "NFTs In Music Industry: Potentiality and Challenge." Proceedings of EVA London 2022, 63-64, 2022.

[6] Behal, Pranav. "Listen-To-Earn: How Web3 Can Change the Music Industry." Available at SSRN 4150998, 2022.

[7] Kiong, Liew Voon. DeFi, NFT and GameFi Made Easy: A Beginner's Guide to Understanding and Investing in DeFi, NFT and GameFi Projects. Liew Voon Kiong, 2021.

[8] Rehman, Wajiha, et al. "NFTs: Applications and challenges," 2021 22nd International Arab Conference on Information Technology (ACIT). IEEE, 2021.

[9] Das, Dipanjan, et al. "Understanding security issues in the NFT ecosystem," Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security, 2022.

[10] Rehman, Wajiha, et al. "NFTs: Applications and challenges," 2021 22nd International Arab Conference on Information Technology (ACIT), IEEE, 2021.

[11] Martinod, Nicolas, et al. "Towards a secure and trustworthy imaging with non-fungible tokens." Applications of Digital Image Processing XLIV. Vol. 11842. SPIE, 2021.

[12] Musamih, Ahmad, et al. "NFTs in Smart Cities: Vision, Applications, and Challenges," IEEE Consumer Electronics Magazine, 2022.

[13] Adler, Amy, "Artificial Authenticity: Art, NFTs, and the Death of Copyright." NFTs, and the Death of Copyright (September 18, 2022) 98, 2022.

[14] Rafli, Dw Putu Alit Denbagus, "NFT Become a Copyright Solution." Journal of Digital Law and Policy 1.2 (2022): 87-96, 2022.

Copyright ⓒ 2023 KSII

Automated Dysarthria Severity Classification Using Diadochokinetic test and Speech Intelligibility Based on LightGBM

MinSeo Kim[1], Junseok Oh[2] and Ji-Hwan Kim[2*]
[1] Department of Artificial Intelligence, Sogang University Seoul, South Korea
[2] Department of Computer Science and Engineering, Sogang University Seoul, South Korea
[e-mail: minseokim@sogang.ac.kr, ohjs@sogang.ac.kr, kimjihwan@sogang.ac.kr]
*Corresponding author: Ji-Hwan Kim

Abstract

There has been a growing interest in developing digital treatments for dysarthria recently. Assessing the severity of dysarthria and identifying its cause are crucial for personalizing patient treatment. In this paper, we use Neurospeech source code to extract Diadochokinetic (DDK) features. We propose a severity classification model based on LightGBM, using the extracted DDK features, gender, and intelligibility. DDK regularity, duration, energy, and 8 other features were extracted from the diadochokinetic test, where /pa/, /ta/, and /ka/ were repeatedly pronounced. We trained a model to classify severity level with the gender, intelligibility and DDK features extracted from the speech of 320 patients, and it showed an accuracy of 79.17%. Intelligibility, number of pauses per second and average of the syllables duration have the high importance on the severity level.

Keywords: Diadochokinetic Test, Dysarthria, LightGBM, NeuroSpeech

1. Introduction

Dysarthria refers to a group of speech impairments caused by damage to the peripheral nervous system or central nervous system [1]. Some people who have had a stroke may suffer from dysarthria. These symptoms can often be alleviated through ongoing speech rehabilitation therapy. Research has been underway to develop digital treatments that can be customized to address individualized need or challenges. In order to provide personalized treatment, it is necessary to assess the severity of dysarthria and determine which assessment factors are problematic. A diadochokinetic(DDK) test can provide information about the condition of patients. We can use the features extracted by DDK test to assess severity and provide personalized care.
Speech intelligibility refers to the extent to which listeners understand what the speaker is trying to communicate. Intelligibility offers information about the overall severity of dysarthria. We utilized intelligibility as a classification feature for assessing severity with DDK features.
In this paper, we extract severity diagnostic features from the diadochokinetic speech of stroke patients with dysarthria and propose a model that can classify severity level based on lightGBM.

2. Related Works

2.1 Diadochokinetic

A diadochokinetic test requires an individual to repeatedly and quickly pronounce /pa/, /ta/, and /ka/ within a specified timeframe. The diadochokinetic test is a fundamental component of motor speech assessment, providing quantitative measurements of the speech articulators' motor abilities [1].

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00621,Development of artificial intelli- gence technology that provides dialog-based multi-modal explainability)

2.2 NeuroSpeech

NeuroSpeech is an open source software that analyzes speech signals from dysarthria patients with Parkinson's disease [2]. The software can analyze speech signals in five dimensions : phonation, articulation, prosody, DDK, and intelligibility [2]. The features extracted from the DDK analysis are related with DDK regularity, duration, energy, pause rate, and 8 other features.

2.3 LightGBM

LightGBM is a powerful machine learning algorithm that can be utilized for various tasks, including feature selection and regression analysis [3]. It utilizes a unique combination of gradient-based one-side sampling (GOSS) and histogram-based gradient boosting techniques, allowing for faster training and prediction times while effectively handling high-dimensional features.

3. Experiments

The dataset[1] used in this study consists of 957 diadochokinetic speech recordings from 320 stroke patients with dysarthria. Severity level was measured on a scale of mild, moderate, and severe, and intelligibility was measured on a scale of 1-5.

We extracted DDK features using the NeuroSpeech source code and trained the lightGBM model using these features, along with gender, age, and intelligibility. We used 80% of the dataset as train data, 10% as validation data, and 10% as test data. Since the data is unbalanced (794 mild, 109 moderate, and 54 severe), we weighted each class ratio when training the LightGBM model.

We measured the model accuracy with test data and it shows 79.17% accuracy. We used SHAP value as the visualization method. **Fig. 1** shows which features are important for classifying severity level. Intelligibility, number of pauses per second (Pause_rate), and average of the syllables duration (avg_dur_ddk) have high importance. Intelligibility, not DDK features, has the highest importance. When evaluating the test data with the model trained using DDK features except intelligibility, it showed 61.46% accuracy. This shows that intelligibility has the highest importance, but DDK features are also important for classifying severity.

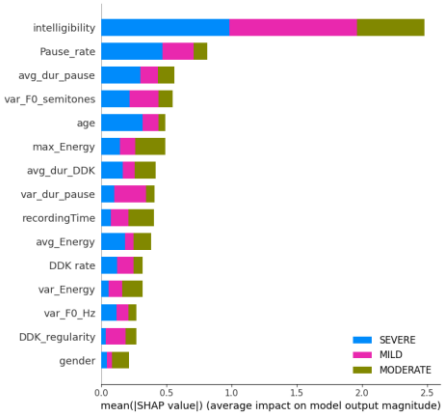


Fig. 1. Feature importance for severity classification

4. Conclusions

In this paper, we extracted features from diadochokinetic speech of stroke patients with dysarthria and use them to classify severity level with lightGBM model. We also investigated which features have high importance in classifying severity. The model demonstrated 75% accuracy, with intelligibility being the most important factor. However, the model performance improved when DDK features were included. Future work will focus on studying the automatic assessment of intelligibility using DDK speech, with the aim of classifying severity using only speech data.

References

[1] F. L. Darley, A. E. Aronson, and J. R. Brown, "Differential diagnostic patterns of dysarthria," *Journal of Speech and Hearing Research*, vol. 12, no. 2, pp. 246-269, 1969.

[2] J. R. Orozco-Arroyave et al., "NeuroSpeech: An open-source software for Parkinson's speech analysis," *Digital Signal Processing*, vol. 77, pp. 207-221, 2018.

[3] G. Ke et al., "LightGBM: A highly efficient gradient boosting decision tree," *Advances in Neural Information Processing Systems*, vol. 30, 2017.

[1] This paper used datasets from 'The Open AI Dataset Project (AI-Hub, S. Korea)'. All data information can be accessed through 'AI-Hub (www.aihub.or.kr)'.

Copyright ⓒ 2023 KSII

UdiOlle: An Olle Trail Course Recommendation Mobile Application Reflecting User Preferences

Nayoung Kim[1] and Dongho Yang[2*]
[1]Department of Computer Education, Jeju National University
[2]Software Convergence Education Institute, Jeju National University
Jeju, South Korea
[e-mail: [1]nayy6974@gmail.com, [2]paris23@jejunu.ac.kr]
*Corresponding author: Donghoo Yang

Abstract

In Korea, the travel style of the general public has recently focused on special events and emotional meanings that suit individual tastes and preferences, with walking tours the most representative reflection of these tourism trends. Opened by the Jeju Olle Corporation, Olle Trail, where one can experience the beautiful natural scenery and characteristics of Jeju, is the best walking tour in Korea. Twenty-three courses are on Jeju Island, including the courses on the surrounding Udo Island, Gapado Island, and Chujado Island, making up a total of 27 courses. Among the various Olle Trail courses, travelers must consider several factors for selecting a course, ranging from weather to time and the purpose of the visit. However, existing mobile applications released by Ollegil do not provide services that recommend courses to reflect these trends. Accordingly, a study was conducted to develop an Olle Trail course recommendation mobile application that reflects user travel tastes and the characteristics of the Jeju Olle Trail courses.

Keywords: Olle Trail, Travel app, Customized tour

1. Introduction

Travel preferences differ widely among individuals as personal tastes in travel have become individualized and diversified. The definition of travel has expanded from merely moving to a different place to experiencing personal tastes and preferences.[1] Walking tours are the most representative reflection of these tourism trends. Novel walking tour programs have been implemented throughout South Korea and have significantly influenced related industries and economies of neighboring areas.

The Olle Trail consists of courses on Udo Island, Gapado Island, and Chujado Island, in addition to 23 other courses on Jeju Island. As a result, tourists must consider various factors, such as the weather, time, and the purpose of the visit, when selecting their preferred Olle Trail course.

Jeju's attractions have the advantage of creating the value of Jeju's image in increasing originality and favorability. The place where the brand is embodied provides a symbolic means of lifestyle and self-expression.[2] Each course in the Olle Trail has distinct characteristics, and their respective subareas have several unique locations that can assist in image branding.

Jeju expanded the scope of the use of big data beyond the level of understanding the tourism phenomenon to the establishment of a real-time customized service platform.[3] Among the big data provided by Jeju as public data, various data are provided to help reflect user preferences, including location information related to Olle Trail, wheelchair sections, and real-time weather information.

Therefore, in this study, a mobile application was developed for recommending Olle Trail courses that reflect visitor preferences in the post-COVID era.

2. Method

Fig. 1 shows an overall schematic diagram of the system.

The database contains two groups of JavaScript Object Notation (JSON) data. The information on each Olle Trail course includes the name, latitudes and longitudes of the starting and end points, landscape, total distance (km), reviews, and ratings.

The location data of the Olle Trail courses provided as public data are in a CSV file that contains basic information, such as the course number, locations of the starting and end points, and average walking time for each course. Based on this information, the data were divided into segments, with data collected through Internet surfing added and processed into JSON files.

The course recommendation algorithm based on user preference matches the information data on the travel preferences of the user already stored in the database with the Olle Trail course data, while the course recommendation algorithm based on alignment and filtering matches the travel preference information input by the user in real-time with the Olle Trail course data and data that can be provided in real-time, such as detailed weather information, and recommends courses that suit user conditions.

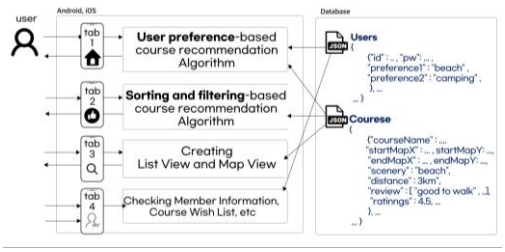


Fig. 1. Schematic diagram

3. Results

The UdiOlle application developed in this study consists of three tab screens—main screen, course recommendation screen, and course search screen.

The main screen provides course recommendation services by theme and atmosphere, as shown in **Fig. 2**. The information on user travel preferences stored in the database is accessed to recommend various subareas of the Olle trail by theme and atmosphere.



Fig. 2. Main Screen

The course recommendation screen recommends the most suitable course based on user preferences by subdividing the factors considered while selecting the walking tour route or the search conditions, as shown in **Fig. 3**.

If the user searches for the name of an area in the search window, as shown in **Fig. 4**, the course search screen lists all the courses related to that area.

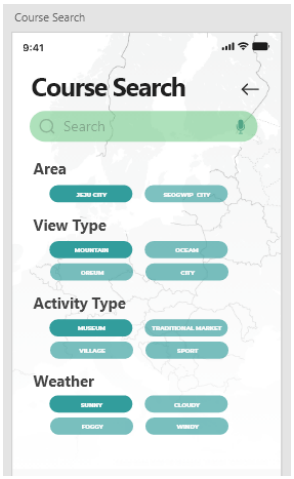


Fig. 3. Course search window



Fig. 4. Recommendation window

4. Discussion

The analysis of the behaviors of Jeju Olle Trail users for each course through text and location information showed that courses that allowed tourists to indulge in experiential tourism, such as the Jeju Olle Market, and visits to nearby tourist sites were popular.[4]

Based on these results, the following alternative solutions were found to be necessary: First, in addition to the walking tour preferences of the user, information on safety and accessibility should be provided to users for selection using a search condition. Second, when providing information on each walking route, it is necessary to provide such information as the location of trash cans on the Olle Trail and precautions for the walking tour in connection with the Jeju Environmental Protection Campaign. Third, the community function should be implemented to enable users to access ratings or reviews given by other users; in addition, companion matching services should be considered to enable users to find walking companions that meet their given conditions.

5. Conclusion

A mobile application for recommending Olle Trail courses was developed, improving upon the limitations of existing Olle Trail applications. This application can recommend walking tour routes depending on user preferences and provides a potential solution to revitalize Jeju tourism centered around the Olle Trail course.

To expand the scope of recommendation, the official information on the 23 Olle Trail courses provided as public data was subdivided into 207 courses and constructed as new JSON data, which are expected to be helpful in future research.

Acknowledgement

"This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW)(2018-0-01863) supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) in 2023"

References

[1] Korea Tourism Organization Press Release, "2022 Tourism Trend Is My Special Moment", Korea Tourism Consulting Team, 1-2, 2021.12.14.

[2] Seong, Yong Seok, & Lim, Kyung Ho, "A Study on the Value Creation of Olle Trail for the Place Branding in Jeju," *A Journal of Brand Design Association of Korea*, 18(1), 115–130, 2020. https://doi.org/10.18852/BDAK.2020.18.1.115

[3] Koh, Sun-Young, & JEONG, GEUNOH, "Policy Achievements and Tasks for Using Big-Data in Regional Tourism -The Case of Jeju Special Self-Governing Province," *Journal of the Korea Academia-Industrial cooperation Society*, 22(3), 579-586, 2021. https://doi.org/10.5762/KAIS.2021.22.3.579

[4] Jeong, S.-H., Yang, J.-D., Kim, M., & Kim, G.-H., "A Study on the Behavior of Jeju Olle Trail Users Using Python and QGIS : Focusing on pre-and post-COVID-19 pandemic. Journal of Tourism Enhancement," *The Korea Tourism Enhancement Society*, 10(3), 277-292, August 31 2022. https://doi.org/10.35498/kotes.2022.10.3.277

Copyright ⓒ 2023 KSII

An Algorithm based on Infant Abnormal Posture Detection through AI Pose Estimation

Jeonghyo Kim*, Seungha Noh, Sunwoo Pi, Wonbeen Im, Chanui Hong and Sangjin Kim
Department of Computer Engineering, Kyonggi University
Seoul, South Korea
[e-mail: longkjh@kyonggi.ac.kr, seung38231@kyonggi.ac.kr, ppssww1202@naver.com, hcu55@naver.com, deemo5551@kyonggi.ac.kr, drsjkim@kgu.ac.kr]
*Corresponding author: Jeonghyo Kim

Abstract

Sudden Infant Death Syndrome (SIDS) accounts for a high proportion of the causes of infant death in Korea. In order to solve this problem, it is important to correct the sleeping posture of infants. However, there is a limit for parents to keep an eye on their child's sleeping posture. Therefore, the present study proposes an infant pose estimation algorithm for correcting children's sleeping posture.

Keywords: pose estimation, computer vision, SIDS

1. Introduction

Sudden infant death syndrome (SIDS) refers to sudden death due to unknown causes in infants under 12 months old.[1] SIDS accounts for a high percentage of infant mortality in Korea. As a result of conducting a study on the child's sleep posture in Korea to solve this syndrome, it is recommended that the angst posture is better than lying the baby on its side or sleeping in an abdominal posture.[2] Based on these studies, this paper proposes an algorithm to detect abnormal sleep postures by measuring the posture of a baby. The structure of the proposed system is as follows. Seventeen body key points coordinate information is obtained by Pose Estimation using the YOLOv8m pose model. Through the obtained key points coordinates, an algorithm using the distance and correlation between coordinates is executed to detect whether it falls into the abnormal posture category set in this paper. It is hoped that our system can prevent SIDS.

2. Pose Estimation

To detect abnormalities, it is necessary to collect information on the body to be used in the algorithm before the algorithm. We obtain 17 key points coordinates from the body of the detection target using the pre-trained YOLOv8m Pose model to obtain body information.[4] Since the model was not pre-trained for children, tests using child data confirmed that it was well detected. It also confirmed that the model works well in the image at an average speed of 50 frames.

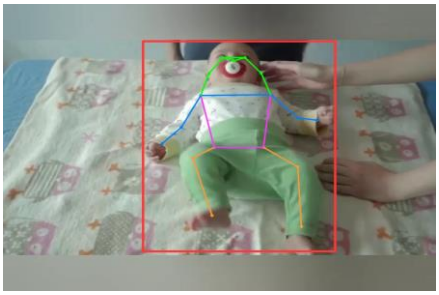


Fig. 1. Skeleton results of baby pose detection using SyRIP Dataset

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW(2021-0-01393) supervised by the IITP(Institute of Information & communications Technology Planing & Evaluation)(2021-0-01393)

3. Infant Posture Detection Algorithm

We developed an algorithm using 17 key points coordinates of the body obtained through Pose Estimation using the YOLOv8m model. Our study judged that the category of abnormal posture was larger than that of normal posture, and based on this judgment, an algorithm was developed to detect normal posture rather than abnormal posture and judge other postures as abnormal posture. There are a total of three posture classes judged by the algorithm. Normal posture, prone posture, abnormal posture. The reason for separately distinguishing the sleeping posture on the back is that the risk of suffocation is higher than that of other abnormal postures, so the risk was determined considerably, and the class was set separately.

3.1 Normal Posture

Among the 17 body coordinates, the coordinates to be used are coordinates 5 to 10, which are the coordinates of both arms and shoulders. If No. 8 (left elbow) and No. 10 (left hand) are more left than No. 6 coordinate (left shoulder), and No. 7 (right elbow) and No. 9 (right hand) are more right than No. 5 coordinate (right shoulder), it is detected in a normal posture.

3.2 Prone Posture

The YOLOv8m is a 2D Pose Estimation model. In 2D Pose Estimation, we cannot distinguish between the front and back of the body, so we use a method of judging the process of reversing the body. If the correlations of coordinates 5 to 10 are opposite to each other, 1 is accumulated in the bad stack per frame, and if the bad stack reaches 100, it is judged as Prone Posture

3.3 Abnormal Posture

If case is non-2.1 and 2.2, 1 is accumulated in the danger stack per frame, and if the danger stack reaches 100, it is judged as Abnormal Posture

4. Conclusions

To solve SIDS, which accounts for a high proportion of infant mortality, this paper proposes an infant abnormal sleep posture detection algorithm using key points coordinates extracted with Pose Estimation. The posture class is divided into three categories to distinguish the degree of posture risk so that risk detection can be performed more carefully. We expect that using our algorithm will reduce the number of SIDS deaths.

References

[1] Kim CH. "Sudden Infant Death Syndrome," J Korean Med Assoc, 44(9), 976-982, Sep. 2001. https://doi.org/10.5124/jkma.2001.44.9.976

[2] Kinney, H. C., & Thach, B. T., "The sudden infant death syndrome," The New England journal of medicine, 361(8), 795–805, 2009. https://doi.org/10.1056/NEJMra0803836

[3] Jocher, G., Chourasia, A., & Qiu, J., "YOLO by Ultralytics (Version 8.0.0)," Computer software, 2023. https://github.com/ultralytics/ultralytics

[4] Huang, Xiaofei and Fu, Nihang and Liu, Shuangjun and Ostadabbas, Sarah, "Invariant Representation Learning for Infant Pose Estimation with Small Data," IEEE International Conference on Automatic Face and Gesture Recognition (FG), 2021.

Copyright ⓒ 2023 KSII

Privacy-Preserving DTx with Differential Privacy

Geehee Yun and Kyoungjin Kim*
Department of Convergence Technology Engineering, Sungshin Women's University
Seoul, South Korea
[e-mail: geehee_yun@naver.com, kyongjin@sungshin.ac.kr]
*Corresponding author: kyoungjin Kim

Abstract

With the increasing adoption of Digital Therapeutics (DTx), concerns arise about the involvement of non-medical entities in processing personal health information. Personal health information encompasses patient prescriptions, identity details, lifestyle logs, and more. As the DTx industry expands, research on privacy protection becomes imperative. Therefore, this study explores the application of the well-established privacy model, Differential Privacy(DP), to adjust the ε value and examine the distribution divergence from the original data. By ensuring a certain level of preservation in patterns shared with the original data through the adjusted ε value, this research highlights the need for further investigation on the trade-off between privacy preservation and data utility.

Keywords: Differential Privacy, Digital Therapeutics, Privacy, Data Utility

1. Introduction

The COVID-19 pandemic has emphasized the significance of health management, leading to a rise in the demand for diverse health management solutions. Digital Therapeutics (DTx) offers a convenient and cost-effective alternative to traditional healthcare practices. However, ensuring privacy protection for patient prescription records and real-time lifelogs from wearable devices is crucial.

2. Background

2.1 Related Work

[1] investigates efficient determination of ε in ε-DP using medical big data, while [2] proposes a method to safeguard wearable device-generated lifelog data through DP. The current study focuses on applying DP to secure patient data management within the DTx platform, estimating an appropriate ε value for privacy during data storage. The current study focuses on applying DP to secure patient data management within the DTx platform, estimating an appropriate ε value for privacy during data storage.

2.2 Gaussian Differential Privacy

DP, initially proposed by D. Work in 2006, is

This work is partly supported by the National Research Foundation of Korea (NRF) grant funded by the Ministry of Science and ICT (MSIT) (No. 2022R1F1A1074038), the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korean Government (MOTIE) (P0008703, The Competency Development Program for Industry Specialist), and the MSIT under the ICAN (ICT Challenge and Advanced Network of HRD) program (No. IITP-2022-RS-2022-00156310) supervised by the Institute of Information & Communication Technology Planning & Evaluation (IITP) and Korea Foundation for Women In Science, Engineering and Technology (WISET) grant funded by the Ministry of Science and ICT(MSIT) under the team research program for female engineering students (WISET-2023-140)..

a privacy mechanism that quantitatively models and measures privacy protection while preserving data utility and considering data distribution characteristics. The basic idea is to add noise using the Laplace distribution. Our study focuses on applying Gaussian-DP to enhance data utility by minimizing noise variance. We investigate the appropriate values of ε by analyzing the distribution difference between the original and perturbed data, emphasizing the need for future research.

3. Research Methodology and Result

Test dataset [3] is a file containing mental health information for tens of thousands of individuals, with 15 columns of data. It includes quasi-identifiers such as residence, age under the 'subgroup' column, as well as numerical information with different distributions, namely the 'Time Period Label' and 'value' columns. The structure of the sample data is presented in **Table 1**, and the distributions are depicted in **Fig. 1**.

Table 1. Test data

Type	Colum name	Datatype
Quasi-Identifiers	Subgroup	object
Sensitive	Time Period Label	float64
Sensitive	Value	float64

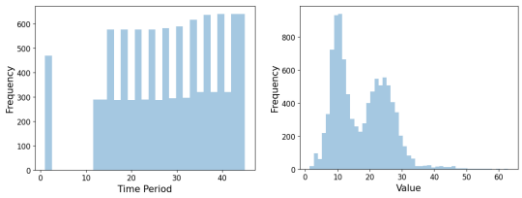
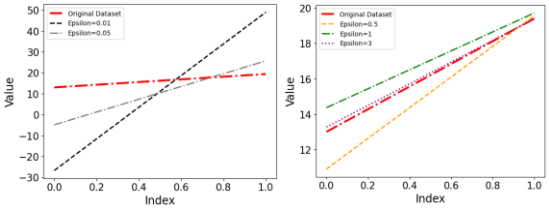


Fig. 1. Data Distribution

The mechanism of adding noise using Gaussian distribution has been implemented in Python. Five values of ε ranging from 0.01 to 3 were selected, and the distribution changes between the original and the data with added noise were represented in **Fig. 2**. The x-axis represents the index of the values, while the y-axis represents the values with DP applied and the values from the original data. According to the experimental results, **Fig. 2(A)** shows that more noise is added depending on the ε values (0.01, 0.05), resulting in significant distribution distortion. On the other hand, **Fig. 2(B)** demonstrates that the values exhibit a certain level of adherence to the original pattern while introducing differences based on the ε values (0.5, 1, 3). Therefore, it is necessary to calibrate the ε values belonging to (B) in order to improve data utility while maintaining security, indicating the need for further research in this area.



(A) (B)

Fig. 2. Original and Noisy Datasets with Different Epsilon Values

4. Conclusions

The paper examines the distribution changes between original data and data with noise by specifying ε values ranging from 0.01 to 3 for DP. Two values with different distributions and a quasi-identifier were selected for the experiments using lifelog data during a patient's DTx usage period. Results show that adding certain levels of noise ($\varepsilon \geq 0.5$) maintains the pattern of the original data to some extent while generating different data. Future research aims to enhance both the privacy protection and utility of real-time patient information by exploring approaches such as ε, δ-DP with statistical calibration of noise based on ε values. This approach can contribute to increasing the usability of the data.

References

[1] Jeongmuk Lim, Taesung Kim, *Proceedings of the Autumn Conference of the Korean Society of Management Information Systems*, pp. 12-15, 2020.

[2] Jonghyun Lim and Jongwook Kim, "IoT Data Collection Method in Privacy-Preserving Environment through Distributed Differential Privacy," *Journal of the Korean Society of Computer Information*, vol. 25, no. 6, pp. 65-72, 2020.

[3] data.gov, "Mental Health Care in the Last 4 Weeks." Available: https://data.gov/.

Copyright ⓒ 2023 KSII

A Research on the Highway New Business Models Corresponding to Changes in Future Transportation Environment

Yong-won Kim
Korea Expressway Corporation Research Institutes
Gyeonggi-do, South Korea
[e-mail: ywkim@ex.co.kr]
*Corresponding author: Yong-won Kim

Abstract

The scope of this research analyzes trends in major changes such as 4th Industrial Revolution and Mobility 4.0, market demand survey, response level diagnosis and comparison analysis of major users due to diversification of future transportation, analysis of systems related to the introduction of eco-friendly vehicles and self-driving vehicles, and analysis of business readiness, promotion direction, and foremost, to discover 2030 mid-to long-term business items (autonomous driving service, etc) in response to changes in the future automobile industry and to derive specific implementation plans.

In order to analyze the future transportation environment, the analysis was conducted with the PEST(Political, Economic, Social and Technological) analysis. For example, to analyze the policy environment, the 「Government 120 National Task」 and other involving policies such as 「Mobility Innovation Roadmap」 by Ministry of Land, Infrastructure and Transport were examined as well as overseas future transportation policy trends. Keywords for industry trend analysis were defined as 2CASE, meaning Connection and Convergence, Automatic and Application, Sharing and Safety, Electrification and Energy Harvesting.

We analyzed key variable factor for Korea Expressway Corporation through surveys such as market size, growth prospects, and participation status of major domestic and foreign companies for each industry trend keyword. These were summarized as keywords for 41 future projects, and 4 core projects (‘EX-EMS: Highway Smart Charging’, ‘EX-Sharing: Highway Shared Mobility Hub’, ‘EX-Mode: Highway Integrated Mobility System’, and ‘Ex-Safety: Highway Future Traffic Safety System’) and a total of 17 new business pools were derived through PEST perspective analysis In addition, 9 core projects that should be promoted primarily by the Korea Expressway Corporation were derived through the AHP analysis, focusing on the importance of each project in perspective of value and growth contribution.

The above research results are expected to have a great effect in deriving practical results in the future business promotion process through the understanding and consensus formation of executives and employees on the new business model unique to Korea Expressway Corporation and the provision of specific roadmaps and resource input plans corresponding to the business promotion stage.

Keywords: Autonomous Vehicle, Electric Vehicle, Business Model

1. Introduction

The Korea Expressway Corporation, as a public institution and highway manager, is being required to discover new businesses that preactively respond to transformation of future transportation environment including electric vehicle distribution, self-driving car commercialization, collaborative economy, Industry and Mobility 4.0

Therefore, the scope of this research analyzes trends in major changes such as 4th Industrial

This research was supported by a research grant from Korea Expressway Corporation Research Institutes.

Revolution and Mobility 4.0, market demand survey, response level diagnosis and comparison analysis of major users due to diversification of future transportation, analysis of systems related to the introduction of eco-friendly vehicles and self-driving vehicles, and analysis of business readiness, promotion direction, and foremost, to discover 2030 mid-to long-term business items (autonomous driving service, etc) in response to changes in the future automobile industry and to derive specific implementation plans.

2. Methodology

In this research, necessary data for the research were collected by applying techniques such as literature review, benchmarking similar institutions, surveys, and interviews. Based on these data, a problem-solving approach that suggests the best way for Korea Expressway Corporation was applied through research and analysis of various methods.

In order to analyze the future transportation environment, the analysis was conducted with the PEST(Political, Economic, Social and Technological) analysis. For example, to analyze the policy environment, the 「Government 120 National Task」and other involving policies such as「Mobility Innovation Roadmap」by Ministry of Land, Infrastructure and Transport were examined as well as overseas future transportation policy trends.

Keywords for industry trend analysis were defined as 2CASE, meaning Connection and Convergence, Automatic and Application, Sharing and Safety, Electrification and Energy Harvesting. We analyzed key variable factor for Korea Expressway Corporation through surveys such as market size, growth prospects, and participation status of major domestic and foreign companies for each industry trend keyword. These were summarized as keywords for 41 future projects, and 4 core projects (‘EX-EMS: Highway Smart Charging’, ‘EX-Sharing: Highway Shared Mobility Hub’, ‘EX-Mode: Highway Integrated Mobility System’, and ‘Ex-Safety: Highway Future Traffic Safety System’) and a total of 17 new business pools were derived through PEST perspective analysis In addition, 9 core projects that should be promoted primarily by the Korea Expressway Corporation were derived through the AHP analysis, focusing on the importance of each project in perspective of value and growth contribution.

3. Research Results

The 9 core projects of Korea Expressway Corporation are categorized and specified as projects according to project goals and expected effects. ‘EX-EMS: Development of Central Control System’, to reach ‘EX-EMS: Highway Smart Charging’, ‘EX-EMS Demand-Customized Charging Infrastructure Expansion’, ‘EX-EMS: Diversifying Energy Harvesting System’, ‘Highway Mobility Hub’ to reach ‘EX-Sharing: Highway shared Mobility Hub’, ‘Shared Parking Platform Development’, ‘EX-Mode Solution Development’ to reach ‘Ex-Mode: Highway Integrated Mobility System’, ‘EX-Blockchain Digital Asset Development’, ‘EX-ARS Advancement’ and ‘Establishment of future traffic safety comprehensive system’ to reach ‘EX-Safety: Highway future traffic safety system.’ [1].

4. Conclusion

This research attempted to do more specific business modeling by considering the 8 elements including: business definition, business goals, business necessity, domestic and foreign cases, business contents and process, role of consultative group, manpower and budget for the core businesses. In the process, factors such as initiative, sustainability, profitability, and publicity were considered important. Especially, to raise awareness of the understanding of the project and the legitimacy of the project, a poster-type image showing the future of the highway to which the new project was applied was produced, focusing on the benefits from the user's point of view.

REFERENCES

[1] Y. Kim et al., “A study on the proposal of a new 2030 road business model according to the change of future automobile”, Korea Expressway Corporation, 2022.

Copyright ⓒ 2023 KSII

Shipboard Indoor Positioning: A Controllable Feature Selection for IQ Fingerprint Map of Bluetooth 5.1

Qianfeng Lin[1], and Jooyoung Son[2]*
[1]Department of Computer Engineering, Korea Maritime and Ocean University
Busan, Republic of Korea
[e-mail: linqianfeng@g.kmou.ac.kr]
[2]Division of Marine IT Engineering, Korea Maritime and Ocean University
Busan, Republic of Korea
[e-mail: mmlab@kmou.ac.kr]
*Corresponding author: Jooyoung Son

Abstract

This paper introduces an indoor positioning scheme using IQ fingerprint maps based on Bluetooth 5.1 and a Random Forest classifier specifically designed for ship environments. The proposed scheme achieves centimeter-level accuracy in ship environments, where Bluetooth 5.1 Angle of Arrival technology faces performance degradation due to complex metal structures and dynamic hull deformation caused by engines, waves, and payload pressure. The scheme makes three contributions: proposing a novel IQ fingerprint map that better characterizes ship environments, developing an indoor positioning scheme using Random Forest classifier, and driving a controllable feature selection to remove the impact of noise. The experimental results indicate that when the number of features directly using IQ samples reaches 35, the test accuracy of the Random Forest classifier achieves 96.95%.

Keywords: IQ fingerprint map, Bluetooth 5.1, random forest classifier, ship environments

1. Introduction

Over 300 travelers and crew members have reportedly experienced symptoms of vomiting and diarrhea during the ongoing journey of the Ruby Princess ship, operated by Princess Cruises. A spokesperson from Princess tells USA TODAY through an email that the mild illnesses are suspected to be caused by norovirus [1]. Implementing close contact tracing is an effective way to mitigate the risk of viral transmission. Indoor Positioning Systems (IPS) greatly facilitate this process. These systems primarily utilize navigation sensors, including accelerometers, gyroscopes, and compasses, as well as RFID, Wi-Fi, and Bluetooth technologies [2], to produce positioning signals. These signals are then integrated with positioning algorithms for accurate position tracking. The energy-efficient and cost-effective Bluetooth Low Energy (BLE) tag is employed for indoor positioning within buildings [3]. By attaching the BLE tag to a smartphone and assigning it a unique ID, close contact tracing can be effectively accomplished for individual identification. A Constant Tone Extension (CTE) is a feature in wireless communication systems, particularly in Bluetooth 5.1 technology, which is used to enhance the performance of Angle of Arrival (AoA) and Direction Finding (DF) techniques. CTE contains a series of in-phase (I) and quadrature-phase (Q) samples that together represent the complex IQ data essential for

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2021R1I1A3056125).

accurate signal processing and analysis. The metallic surroundings inside the vessel affect signal reflection paths, causing signal distortion with noise and subsequently altering the IQ samples [4]. Furthermore, during navigation, the ship's walls are in motion, which exacerbates the influence of moving walls on signal reflection paths and results in even greater IQ sample distortion.

2. Related Works

Bluetooth 5.1 direction-finding is based on the concept of AoA, which is used to determine the direction of a device such as a BLE tag. The BLE tag transmits a special direction-finding signal with a single antenna, and the receiving device (e.g., an AoA Locator) has an array of multiple antennas. As the signal passes through the array, the receiving device detects a phase difference due to the varying distances between the transmitting antenna and each received antenna in the array. By taking IQ samples of the signal and switching between received antennas in the array, the relative direction of the signal can be calculated [5]. Our proposed approach is to directly input the signal samples as features using a Controllable Feature Selection (CFS) into the classification algorithm, reducing the time consumption of parameter estimation and improving the accuracy of estimating position.

3. Shipboard Indoor Positioning

The fingerprint map is a commonly used technique for indoor positioning that relies on measuring the received signal strength from beacons or sensors to determine the user's position. This approach involves generating a venue signature by recording signal data at each position while walking. By mapping the venue to a grid, a database of signal data can be created and linked to specific positions within the building, enabling the position to be encoded as a label.

We propose that the IQ fingerprint map can be utilized as training data for developing a classifier using supervised learning methods, such as the Random Forest (RF) classifier. Fig. 1 provides a technical overview of the ship indoor positioning scheme. The IQ fingerprint map is fed into CFS, which employs LightGBM [6] to calculate the feature importance of IQ samples.

Enhancing a RF classifier involves sorting features by importance, in descending order. This method, known as feature ranking, helps identify crucial variables that significantly contribute to prediction accuracy. Prioritizing these features, the classifier optimizes performance, while this ranking method fosters a more transparent feature selection process.

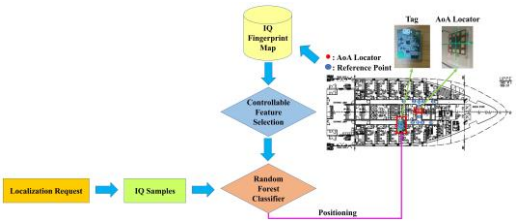


Fig. 1. Workflow of the indoor positioning scheme for ships

4. Results and Discussion

We conduct an indoor positioning experiment, collecting IQ samples from a test area located on the upper deck of the HANNARA ship, which is used by students. The test area measures 7 meters by 7.14 meters. Within this area, we designate 12 positions as reference points, as illustrated in Fig. 2. A total of 21859 CTE are collected from the 12 reference points. Of these, 15301 CTE are used to create the IQ fingerprint map, which serves as the train set. The remaining 6558 CTE are used as the test set.

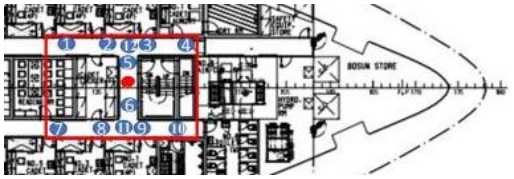


Fig. 2. Deployment of reference points on the upper deck of HANNARA ship

To compare the results of CFS, this paper employs feature mining on IQ samples using tsfel packet [7]. Feature importance is calculated using information gain calculated by LightGBM. Fig. 3 shows the train and test accuracy of the RF classifier on the 44 features selected by CFS. These 44 features are the original IQ samples. As shown in Fig. 3, the train accuracy changes significantly when the number of features

increases from 1 to 3, indicating that using the original IQ samples as features is helpful for the classification accuracy of RF classifier. When the number of features reaches 35 (the number of optimal feature set), the test accuracy of the RF classifier is 96.95%.

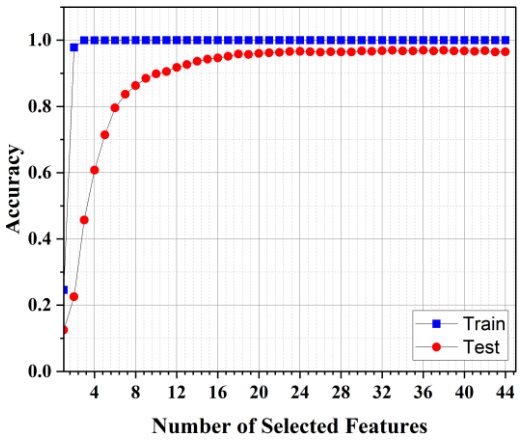


Fig. 3. Accuracy of train and test for RF Classifier using 44 features obtained by original IQ samples

5. Conclusions

This paper evaluates Bluetooth 5.1 AoA indoor positioning in ship environments and find the result of estimating unsatisfactory due to high RMSE and inaccurate positioning because of distorted IQ samples. Therefore, we proposes a Controllable Feature Selection (CFS) for IQ fingerprint map of Bluetooth 5.1 that overcomes the challenges of distorted IQ samples. We test CFS in a real ship environment. The experimental result show that when the number of features using directly IQ samples reaches 35 (the number of optimal feature set), the test accuracy of the RF classifier is 96.95%.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2021R1I1A3056125).

References

[1] USA Today. More than 300 passengers, crew fall ill on Ruby Princess cruise ship. https://www.usatoday.com/story/travel/cruises/2023/03/09/ruby-princess-cruise-ship-300-passengers-sick/11434650002/, 2023.

[2] Liu H, Darabi H, Banerjee P, et al. Survey of wireless indoor positioning techniques and systems[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2007, 37(6): 1067-1080.

[3] Memon S, Memon M M, Shaikh F K, et al. Smart indoor positioning using BLE technology[C]//2017 4th IEEE International Conference on Engineering Technologies and Applied Sciences (ICETAS). IEEE, 2017: 1-5.

[4] Cominelli M, Patras P, Gringoli F. Dead on arrival: An empirical study of the Bluetooth 5.1 positioning system[C]//Proceedings of the 13th international workshop on wireless network testbeds, experimental evaluation & characterization. 2019: 13-20.

[5] Nordic Semiconductor, nWP036 – Direction Finding, https://infocenter.nordicsemi.com/index.jsp?topic=%2Fnwp_036%2FWP%2Fnwp_036%2Fintro.htm, 2022.

[6] Ke G, Meng Q, Finley T, et al. Lightgbm: A highly efficient gradient boosting decision tree[J]. Advances in neural information processing systems, 2017, 30.

[7] Barandas M, Folgado D, Fernandes L, et al. TSFEL: Time series feature extraction library[J]. SoftwareX, 2020, 11: 100456.

Copyright ⓒ 2023 KSII

A Visualization Method of Abnormal Sound Using Acoustic Measurement System Based on Socket Network Communication for Web Client Users

Jiin Shin[1], Suyoung Kim[1], Jaehyun Lee[3], Jongho Paik[2]*
[1]Major in Software Fusion*, [2]Department of Software Convergence,
[3]Seoul Women's University and SoundCam Korea
Seoul, South Korea
[e-mail: {jiin99, rlatndud0515, paikjh}@swu.ac.kr, christian.lee@soundcam-kr.com]
*Corresponding author: Jongho Paik

Abstract

In this paper, we propose a visualization method of abnormal sound using acoustic measurement system based on socket network communication for web client users. The proposed method visualizes the data collected from device based on socket communication and areas where abnormal sound occurs and makes it possible to check the data visualization in a web browser. Furthermore the proposed method allows users to conveniently access the web from anywhere in an environment connected to the Internet and intuitively detect the source location of abnormal sound in a specific frequency range.

Keywords: Socket Communication, Computer Vision, Server, Flask

1. Introduction

Cars and home appliances that generate noise and vibration on their own are subjected to quality inspections at the end of the production process. Quality inspections are typically conducted by an inspector or quality control specialist who listens to the noise or uses a machine to measure it after operating the product, in order to determine whether it meets the required performance standards[1]. However, the method of determining how a person perceives the noise inevitably depends on the experience of the measurer. Therefore, the judgment may vary depending on factors such as the emotional state of the measurer and the noise at the site. Instrumental determination methods can also make noise measurement difficult due to background noise. In this case, a sensor is attached to measure the vibration and indirectly measure the noise[2]. However, this method requires attaching sensors directly to the device one by one. In addition, it can be applied only to small home appliances, and even if a large object such as a pipe in a factory fails, it is difficult to determine where the failure occurs.

To solve this problem, it is necessary to collect acoustic data and visualize the point where a specific abnormal sound frequency occurs. However, since the existing abnormal sound measurement system is based on socket communication, a program for each OS suitable for the user's device is required to access it. Also, monitoring is not possible on a tablet.

Therefore, in this study, we propose a visualization method of abnormal sound using acoustic measurement system based on socket network communication for web client users. The proposed method visualizes the data

collected from a device with one camera and multiple microphones so that it can be checked on the web.

2. Acoustic Measurement System

The abnormal sound measurement system comprises a device equipped with one camera and multiple microphones, firmware installed on the device, and a host program. The PC and device with the host program installed communicate with each other through Ethernet using socket communication. The host program receives data from the device through socket communication, and performs preprocessing on the collected data. The processed data is then used to generate an image and a spectrogram, which helps to locate the point of sound occurrence and the corresponding sound frequency. Preprocessing is performed by classifying the data according to the type.

The incoming binary data is converted into unsigned positive integer values and processed into 240 x 320 black and white images. Collected sound data is converted into a two-dimensional array, with values above a certain decibel threshold identified as anomalies. To convert the sound image into a format that can be displayed along with the video image, all values in the 2D array except for those between the maximum value and the value four steps lower than the maximum are set to 0 and converted to the same size as the video image. Each value of the video image array and the acoustic image array is multiplied by α and 1-α, respectively, and the two arrays are added to synthesize the two images. Spectrum data converts incoming binary data into 1024 real numbers. The converted data is inserted into the first row of the two-dimensional array containing the spectrum data, and the last row is removed before saving. The preprocessed data is output through the GUI.Through this process, not only can the data collection area be adjusted to a specific frequency range, but also the point where a specific sound or abnormal sound is generated is visualized. **Fig. 1** shows a visualizaion example of abnormal sound using acoustic measurement system[3].



Fig. 1. A visualization example of Abnormal sound [4]

3. Visualization Method for Web Client Users

In the visualization method proposed in this paper, the client first accesses the web and requests a UDP connection with the device from the host program. The host program that received the request establishes a UDP and TCP connection and transmits a connection completion response code with the device to the web. After confirming that the host program and the device are connected, the client requests the host program to transmit device setting data. The client requests the host program to collect data. The host program receiving the request from the client requests the device to start collecting and starts receiving data. The received data is transmitted to the web after parsing and preprocessing. The sequence diagram for the entire process is shown in **Fig. 2**.

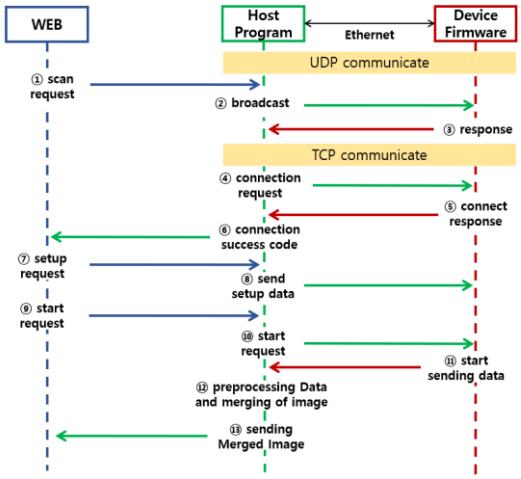


Fig. 2. Sequence diagram of visualization method of abnormal sound for web client users

3.1 Socket programming for sending and receiving data.

The device firmware quickly sends various types of data to the host program. In this system, among various data, acoustic data, video data, and spectrum data are used. And each data comes in binary form at 100, 50, and 25 per second. Therefore, it is necessary to convert the data into a form in the host program. However, while converting the data, it is suggested to transfer the data from the device. To solve this problem, it was implemented by dividing the thread that receives data and the thread that converts the received data. The thread receiving the data classifies and stores the received data according to the object ID. **Fig. 3** shows the structure of the data.

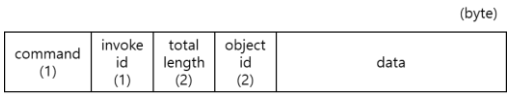


Fig. 3. Structure of socket data

The device transmits a lot of data in addition to the data required by the host program. The data required by the host program are video data, acoustic video data, and spectrum data in data messages. Therefore, only those whose command is “data message” and whose data object id is “video data”, “acoustic video data”, and “spectrum data” are classified and saved.

Among grammars for classification according to various conditions, the performance of match-case and if-else, which play the role of switch-case in Python, were compared. As a result, branches with if-else received more data, so we used them. However, as shown in Table 1, there is a problem that the number of received data varies depending on the order of conditions in the if statement. This is because if the most incoming acoustic video data is given as the third condition, it is saved after going through two if statements. Therefore, the acoustic video data with the most incoming data was entered as the condition of the if statement in order. **Table 1** shows the number of data received during 1 second according to conditions.

Table 1. Number of data received by condition

conditional statement	Order of conditions	video	acoustic	spectrum
match-case	video, acoustic, spectrum	30	88	22
if-else	video, acoustic, spectrum	50	97	24
if-else	acoustic, video, spectrum	50	100	25

3.2 Data parsing and preprocessing

The host program visualizes and displays video data, acoustic video data, and spectrum data to the user. This data is received in binary form from the device and stored in a list based on its type. To ensure proper visualization, the received data needs to be preprocessed and parsed sequentially. To prevent delays, separate threads were implemented for data parsing and conversion. **Fig. 4** illustrates the process of receiving, parsing, and preprocessing data from the device.

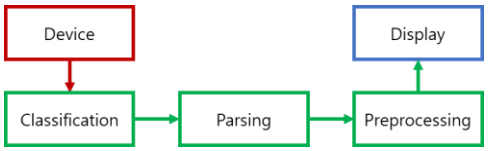


Fig. 4. Data processing process block diagram

First, we compared the performance of methods that can be used to retrieve data stored in a list. The test was conducted by creating an array filled with 100,000 zeros and measuring the time required to retrieve all elements. The most representative method used to retrieve data stored in a list one by one is pop(). There are two ways to output data through pop(): a method of passing the index value of the data to be retrieved as a parameter, and a method of not inserting an index. As a result of testing the two methods, better performance was shown when indexes were not inserted. Next, we tested the next() function of the iterator object in the list data type of python. As a result, it took less time than when I tested pop(). Therefore, when fetching data stored in a list, next() of the iterator that takes the least amount of time is used. However, unlike pop(), next() of iterator does not delete the

retrieved data. Therefore, after retrieving all the data stored in the array, initialize the array to an empty array. **Table 2** shows the time required to search all data for each method.

Table 2. Data retrieval time by method

method	time(sec)
pop(0)	1.3443870544433594
pop()	0.013965845108032227
__iter__().__next__()	0.006989717483520508

The retrieved data is parsed and stored in an array. The stored data is preprocessed in a new thread before being sent to the web server. The preprocessing process is the same as the existing abnormal sound measurement system.

3.3 Web browser implementation

Flask is a micro-framework written in Python, and is suitable in that it has high interoperability with Python. **Fig.s 5** shows the screen output through cv2 of python and the web browser implemented through Flask, respectively.

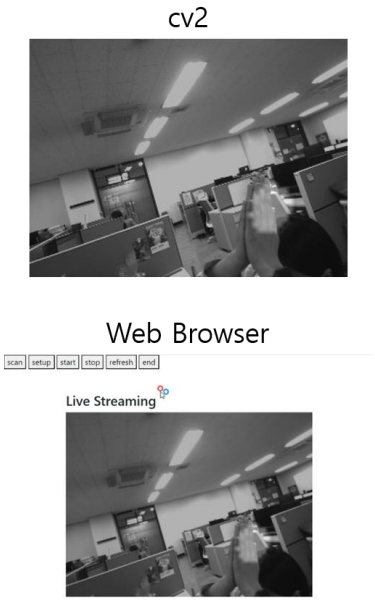


Fig. 5. Screen of cv2 and web browser

The web's role in this study is to display a series of images on the screen, resembling a video. The server configures the response header and Content-Type as JPEG to enable proper display. The images are transmitted as encoded bytes and displayed in the order they are received. This allows users to access and view the visualized data on internet-connected devices, even if they are not directly connected to the device.

4. Conclusions

In the previous abnormal sound measurement system, users accessed the device through a host program to collect and visualize data. However, it had a limitation where the measurement system could only be accessed and monitored on a PC directly connected to the device firmware during the socket communication process.

To address this issue, this paper proposed a visualization method of abnormal sound using acoustic measurement system based on socket network communication for web client users. The Flask framework was used for this purpose, and the image frames collected by the system were periodically transmitted as response data and displayed on the web screen. Through the proposed method, users could access the web from anywhere in an Internet-connected environment and identify the points where sound was generated within a specific frequency range. Furthermore, it is expected that in the future, the system can automatically recognize the corresponding points without the need for manual modification of the frequency domain by the user when a specific sound is generated.

References

[1] Youngki Kim, Sunghyo Park, Jaesun Lee, “Deep Learning for Noise Vibration Quality Inspection”, Journal of KSNVE, Vol. 30, No. 1, pp. 26-31, 2020

[2] C. K. Chun, S. J. Park, G. S. Yi, Y. S. Ma, “A study on Vibration Monitoring for Inferior Window Regulator Selection”, Journal of Korea Academia-Industrial cooperation Society, Vol. 8, No. 1, pp. 18-24, 2007

[3] https://www.soundcam-kr.com/

[4] Jiin Shin, Suyoung Kim et al., “Design of Acoustic Image Analysis System based on Socket Communication”, Korean Society for Internet Information(KSII), Vol24. No.1, pp. 9-10, 2023

Copyright ⓒ 2023 KSII

A Study on Vulnerability Analysis and Memory Forensics of ESP32

Jiyeon Baek[1], Jiwon Jang[1], and Seongmin Kim[2*]
[1]Department of Future Convergence Technology Engineering, Sungshin Women's University
[2]Department of Convergence Security Engineering, Sungshin Women's University
Seoul, South Korea
[e-mail: hesedbaek98@gmail.com, nebulaboratory@naver.com, sm.kim@sungshin.ac.kr]
*Corresponding author: Seongmin Kim

Abstract

As the Internet of Things (IoT) has gained significant prominence in our daily lives, most IoT devices rely on over-the-air(OTA) technology to automatically update firmware or software remotely via the network connection to relieve the burden of manual updates by users. However, preserving security guarantees for OTA interface legs behind the requirement to defend against potential threats. This paper presents a simulation of an attack scenario on the commoditized System-on-a-chip, ESP32 chip, utilized for drones during their OTA update process. We demonstrate three types of attacks, WiFi cracking, ARP spoofing, and TCP SYN flooding techniques to postpone the automatic OTA update procedure. Finally, we review the chip to obtain traces of attacks from a forensics perspective and acquire memory forensic artifacts to indicate the SYN flooding attack.

Keywords: Memory forensics, Over-The-Air (OTA), ESP32, Attack scenario

1. Introduction

The Internet of Things(IoT) has become increasingly ubiquitous across various domains. However, security for IoT devices has been less of a priority than functionality and performance due to its limited resources. This creates challenges in implementing robust security measures, resulting in various security incidents.

In this paper, we study IoT security by simulating an attack scenario on an over-the-air (OTA) update with a ESP32 chip. The OTA is a technology that updates firmware or software remotely via the network, and the ESP32, a low-cost System on Chip (SoC), offers rich integrations with the OTA feature. In particular, we target the IoT drone based on the ESP32 DOIT DEVKIT board (hereinafter ESP32 drone). We attack the OTA update of the ESP32 drone and perform a memory forensics on the chip to obtain traces of the attack.

2. Simulation of OTA attack scenario

This section demonstrates a simulation of the OTA attack scenario as Fig. 1. We assume that a victim is updating the ESP32 drone firmware via OTA, and an attacker attempts network attacks during the update. First, we crack the victim's WiFi network to identify the IP address of the ESP32 drone. Two attacks are then carried out during the update process: ARP spoofing to sniff the OTA packet data and TCP SYN flooding to disrupt the update. To implement the scenario, Kali Linux is used as the attacker's PC, and IPTIME N604R plus router is used as the victim's access point to which the ESP32 drone

This work is partly supported by the Korea Institute for Advancement of Technology (KIAT) grant funded by the Korean Government (MOTIE) (P0008703, The Competency Development Program for Industry Specialist), and the MSIT under the ICAN (ICT Challenge and Advanced Network of HRD) program (No. IITP-2022-RS-2022-00156310) supervised by the Institute of Information & Communication Technology Planning & Evaluation (IITP).

is connected.

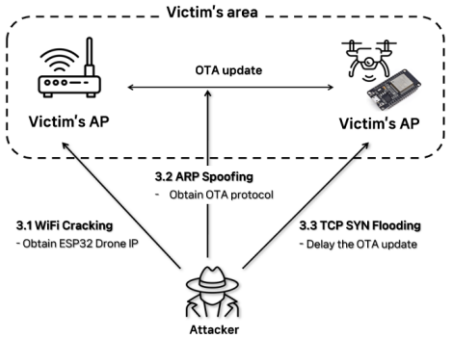


Fig. 1. The OTA attack scenario on the ESP32 drone.

2.1 Cracking Wireless Access Point

In the initial stage of the attack scenario, we employed various network hacking tools in the aircrack-ng package, which crack WEP/WPA networks, to infiltrate the victim's router [1]. First, we set the wireless interface as monitor mode using airmon-ng. Next, we scanned the victim's network with airodump-ng and utilized aireplay-ng to capture the WPA handshake of victim's network. Subsequently, we employed aircrack-ng with the password dictionary rockyou.txt [2] to crack the WiFi network. After accessing the victim's network, we identified the IP address of the ESP32 drone, which was utilized for ARP spoofing attack.

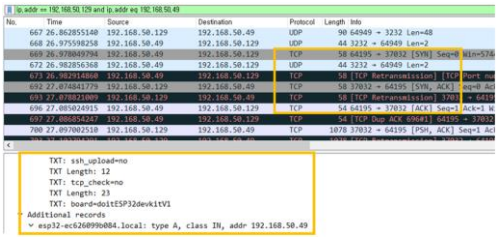


Fig. 2. Discovered TCP 3-way handshake connection and Drone information during OTA update.

2.2 OTA update packet sniffing

To implement OTA packet sniffing, we performed ARP poisoning supported by Ettercap [3], and concurrently captured the OTA packet using Wireshark [4] while the OTA update was underway on the ESP32. Through this, we uncovered that the TCP protocol is used during the OTA update. Furthermore, we observed device informations such as board type and user-designated strings, as depicted in **Fig. 2**.

2.3 TCP SYN Flooding attack during OTA update

After identifying the OTA protocol, we conducted the TCP SYN flooding attack with the hping3 tool [5] to disrupt the OTA update. We observed that the update server ceased to establish a connection as shown in **Fig. 3**.

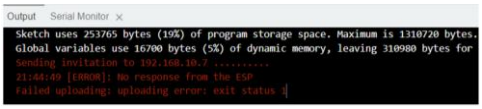


Fig. 3. The OTA update process fails when a TCP SYN packet is sent to the ESP32 drone.

In summary, our study presented the attack scenario including WiFi cracking, ARP spoofing and TCP SYN flooding attack on the OTA update process of the ESP32 drone, and successfully disrupted the update. This scenario highlights that if an attacker continues to hinder OTA updates with malicious intent, the vulnerable software version may persist, leaving the possibility of future hacking incidents.

3. ESP32 Memory Analysis

To excavate the forensic evidence of TCP SYN flooding attack, we perform a memory forensics on the flash memory in the ESP32, using the UART (Universal Asynchronous Receiver /Transmitter) interface and ESP-IDF (Espressif IoT Development Framework).

3.1 ESP32 Memory Structure

Our target ESP32 was set up with one of the Espressif's pre-defined partition tables which contains OTA partitions. Once an OTA update is performed, the updated data is written in the *app* partitions, and the partition ID is stored in the *otadata* area. After the update, the ESP32 is automatically rebooted, and the bootloader refers to the ID stored in the *otadata* to execute the corresponding partition [6].

3.2 Memory Dump Analysis

In order to dump the flash memory, we established a connection between the chip and a PC via UART communication and extracted the flash memory dump using esptool.py, an

open-source tool supplied by ESP-IDF. Next, we performed a total of three memory dumps: two for the normal update attempts, each of which was given the name *esp32_dump.bin* and *esp32_dump2.bin*, and one for the memory dump taken after the SYN flooding attack, which we named *synflooding.bin*.

1) Analyzing *esp32_dump.bin*

In order to identify the memory related to the OTA update, we first analyzed the *otadata* partition in the *esp32_dump.bin*. We figured out that this area contained a historical record of WiFi connections, including a set of SSID and passwords, as depicted in **Fig. 4**. Furthermore, we compared the updated firmware in the chip with the sniffed OTA packets, and discovered that the same values with the packet data were stored in the *app* partition area.

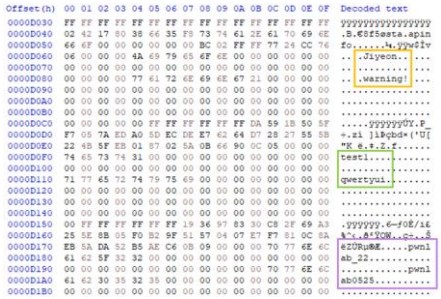


Fig. 4. WiFi data is stored in *otadata* partition.

This analysis implies that memory forensics enables to obtain of the complete history of WiFi connections of the ESP32 drone. Not only does this apply to the attack case presented in this paper, but we believe this information can also serve as a trace of an attacker in various attack cases and provide valuable clues for identifying potential artifacts related to the transmission of fake firmware.

2) Detecting evidence of TCP SYN Flooding

To identify the memory artifacts of the SYN flooding attack, we compared the three dump files, as illustrated in Fig. 5. The analysis showed that the *otadata* partition recorded data block size of 0x40 during a regular OTA update. However, when the SYN flooding attack was executed on the device, the memory registered five times the amount of data compared to the normal OTA update. The result states that an abnormally high number of the data blocks written in this partition can be considered an artifact of an SYN flooding attack. Moreover, it is possible to consider this a potential sign of other types of DoS attacks.



Fig. 5. Comparison between *esp32_dump.bin*, *esp32_dump1.bin*, and *synflooding.bin* (list order).

4. Conclusion

In this study, we selected the ESP32 drone and simulated the OTA attack scenario during an OTA update. We discovered that TCP SYN flooding could interrupt OTA updates, which leads to the ESP32 being vulnerable as the firmware remains unpatched. Also we performed a memory analysis from a forensic perspective and identify the artifacts of the attack stored in the non-volatile memory partition.

References

[1] Aircrack-Ng. [Online]. Available: https://www.kali.org/tools/aircrack-ng/, accessed May 8, 2023.

[2] Wordlists. [Online]. Available: https://www.kali.org/tools/wordlists/, accessed May 8, 2023.

[3] Ettercap Project. [Online]. Available: https://www.ettercap-project.org/, accessed May 8, 2023.

[4] Wireshark. [Online]. Available: https://www.wireshark.org, accessed May 8, 2023.

[5] Hping3. [Online]. Available: https://www.kali.org/tools/hping3/, accessed May 8, 2023.

[6] ESP32 Tutorials. [Online]. Available: http://www.lucadentella.it/en/2016/12/22/esp32-4-flash-bootloader-e-freertos/, accessed May 8, 2023.

Copyright ⓒ 2023 KSII

Generating multicast routing trees method in SDN using Pearl

Heechan Kim, Janggun Jeon, Yejin Lee and Namgi Kim
Department of Computer Science, Kyonggi University
Suwon, South Korea
[e-mail: train1312@naver.com, 201812285@kyonggi.ac.kr, sukiouob@kyonggi.ac.kr, ngkim@kyonggi.ac.kr]
*Corresponding author: Namgi Kim

Abstract

With the recent proliferation of various IT devices and the internet, network traffic is steadily increasing at an exponential rate. In particular, Multimedia services occupy a significant portion of overall network traffic due to their high bandwidth demands. To efficiently maintain and manage networks, SDN(Software Define Network) is being extensively researched. For efficient processing of multimedia traffic, multicast routing is often used. In this paper, we propose a method for generating multicast routing trees using meta-reinforcement learning in SDN environment.

Keywords: Meta Reinforcement Learning, Deep Learning, Software Define Network, Network

1. Introduction

With the prolifetation of various digital devices and the internet, network traffic is steadily increasing. In particular, Multimedia services occupy a significant portion of overall network traffic due to ther high bandwidth demands. Recently, with the spread of high-quality video streaming and virtual reality services, the rate of increase in network traffic has been even steeper. In this paper, we propose a method for generating multicast routung trees in SDN environment that can effectively handle multimedia traffic.

2. Related Work

2.1 Software Define Network(SDN)

The conventional network architecture has both the data plane and control plane in a single network device, which makes it difficult to manage the entire network. To address this limitation, Software-Define Network(SDN) separates the control plane from the data plane, with the SDN controller responsible for the control plane functionality and the data plane functionality executed as in the conventional networks. As a result, the entire network can be managed in a software-based manner, making network management more efficient and flexible compared to the conventional network architecture.

2.2 Multicast Routing

There are two types of network traffic processing methods: unicast routing and multicast routing. [1] Unicast routing is a method of transmitting packets from a single source node to a single destination node.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2020R1A6A1A03040583), This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research], “This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)”(2021-0-01393)

Multicast routing is a method of transmitting the same packet from a single source node to multiple destination nodes. Therefore, multicast routing can efficiently utilize network bandwidth since a single packet can be transmitted to multiple destination nodes.

2.3 Probabilistic Embeddings for Actor critic RL(PEARL)

On-policy meta reinforcement learning algorithms have very low sample efficiency based on on-polict methods. [2] Therefore, in this paper, we used an off-policy meta reinforcement learning algorithm that disentangles task inference and control. [3] This method demenstartes 20-100 times better sample efficiency and 50-100% improved performance compared to on-policy meta reinforcement learning algorithms.

3. Method

3.1 Markov Decision Process(MDP)

The Problem of generating multicast routing trees can be defined as a sequential decision-making problem where the links and nodes from the source to the destination must be determined in a certain order. This problem can be modeled as a Markov Decision Process(MDP) since it involves sequential decision-making. Therefore, in this paper, the problem situation was modeled as shown in **Table 1** using an MDP.

Table 1. Markov Decision Process Modeling

Component	Description
State	Network topology information, Source/destination information, Generated multicast routing tree connection information
Action	The set of nodes that are connectable
Reward	If the destination node is reached, a reward is provided
Discount factor	The relationship between immediate reward and future reward

3.2 Proposal method structure and learning process

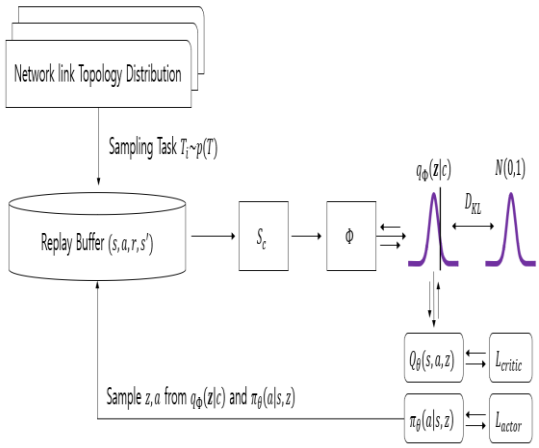


Fig. 1. Proposal method structure

The following data is stored in the replay buffer. First, a task is sampled from the network link topology distribution. Then, z is sampled from q, and an action is extracted from π using z. The extracted action is used to store (s,a,r,s') data in the replay buffer.

The training process involoves using the set data for training. To update the encoder with S, c is sampled, and to update the SAC agent with B, b is sampled. Next, the encoder q samples the context variable b and uses z and b to calculate the actor and critic network loss functions. Then, the regularzizing term D of the overall object function obtained from the variational inference is calculated to output the loss function. The obtained loss function value is used to update the encoder parameters, actor parameters, and critic parameters.

4. Conclusion

In this paper, we propose a method to generate multicast routing trees in SDN using Pearl. We will conduct experiments using the proposed method to demonstrate its efficiency compared to existing routing tree creation methods. After that, we plan to add Graph Neural Network into this structure using the recently emerging Graph Neural Network method. [4]

References

[1] Steve Deering, Deborah Estrin, "Multicast Routing in Datagram Internetworks and Extended LANs," *ACM Transactions on Computer Systems*, Vol. 8 No. 2, 1990.

[2] Yan Duan, John Schulman, Xi Chen, Peter L. Bartlett, llya Sutskever, Pieter Abbeel, "RL²: Fast Reinforcement Learning via Slow Reinforcement Learning," *Proceedings of the 34th International Conference on Machine Learning(ICML)*, 2017.

[3] Kate Rakelly, Aurick Zhou, Deirdre Quillen, Chelsea Finn, Sergey Levine, "Efficient Off-Policy Meta-Reinforcement Learning via Probabilistic Context Variables," *In Advances in Neural Information Processing Systems*, Vol 33, pp. 10131-10142, 2020.

[4] Thomas N. Kipf, Max Welling, "Semi-Supervised Classification with Graph Convolution Network," *International Conference on Learning Representations(ICLR)*, 2017.

Copyright ⓒ 2023 KSII

Implementation of Graph Neural Network for Link Prediction in Software-Defined Networks

Syed M. A. H. Bukhari[1], Afaq Muhammad[1], and Wang-Cheol Song[1*]
[1] Department of Computer Engineering, Jeju National University
Jeju, South Korea
[E-mail: ammar@office.jejunu.ac.kr, afaq24@gmail.com, philo@jejunu.ac.kr]
[*]Corresponding author: Wang-Cheol Song

Abstract

Software Defined Networking (SDN) provides great flexibility and scalability in managing and controlling the network. Identifying and managing the resources in an SDN is a challenging task. This paper presents the idea of leveraging the Graph Neural Network (GNN) to predict links between the resources inside the network. The proposed framework can be utilized in many ways, such as load balancing, network management, anomaly detection, topology prediction, and network evolution analysis. This paper discusses the techniques which can be utilized to solve the link prediction problem using the GNN. The GNN can take input in the form of node features, adjacency matrix, and edge features. More precisely, this paper emphasizes the implementation details and advantages of utilizing GNN for link prediction in SDNs. Moreover, we also provide the potential benefits of utilizing the proposed approach. The GNN takes advantage of graph structure and node features to effectively capture complex relational patterns, outperforming traditional methods in link prediction tasks.

Keywords: Software-Defined Network, Graph Neural Network, link prediction, network management

1. Introduction

With the ever-increasing internet and the advancements in network devices, it has become difficult to configure each network device individually. The network administrator has to visit each site to deploy and configure each device, which is time-consuming and cost-inefficient. However, with the introduction of Software-Defined Networking (SDN) and virtualization of network resources, it has become easier to deploy the configurations remotely [1]. The SDN is a modern approach that separates the control plane from the data plane, enabling the efficient management and configuration of network resources. Separating the control and data plane enables the flexibility and scalability of centralized management through network programmability.

An SDN typically consists of physical or virtual network devices, such as switches, routers, and bridges, connected to an SDN controller that manages these devices. The SDN controller manages these resources through an OpenFlow protocol standardized by Open Networking Foundation (ONF). The SDN controller uses the OpenFlow protocol to configure the network devices and implement the network policies [2]. The network devices are connected through communication links, either physical or virtual, to transmit and receive data between the network devices.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2022R1I1A1A01064556). This research was supported by Brain Pool program funded by the Ministry of Science and ICT through the National Research Foundation of Korea (2019H1D3A1A01102980).

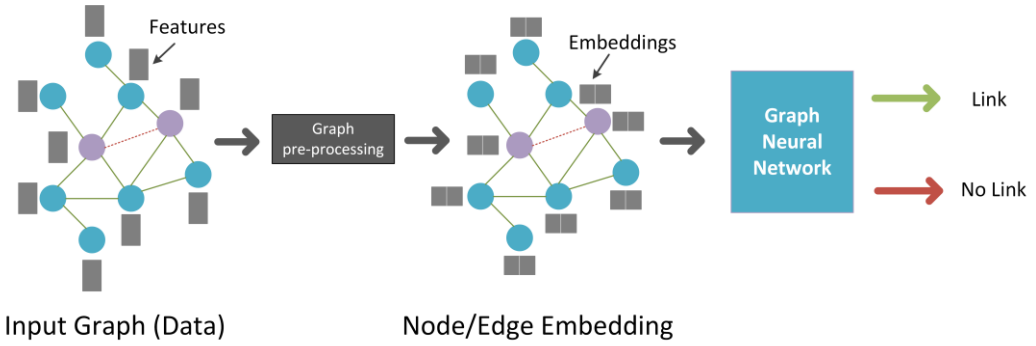


Fig. 1. Link Prediction using Graph Neural Network

In an SDN, the network topology changes dynamically due to overutilization of resources, network device failure, and increased traffic over the network device. To address the aforementioned challenges, the SDN controller should determine about re-routing the traffic and allocate resources by applying network policies [3]. For this purpose, the SDN controller should know about the existence of links between different resources of the network. The link prediction between different network resources is essential in optimizing performance and increasing network reliability [4]. However, the link prediction between different network resources is challenging due to the dynamic nature of the network. For this purpose, many techniques have been proposed for predicting links, which can be utilized to address the aforementioned challenge [5].

The Graph Neural Network (GNN) has shown potential in the area where data is in the form of graphs, such as network topologies [6]. For example, GNN can be used for link prediction tasks by taking advantage of complex network data, as shown in **Fig. 1**. In SDNs, the GNN can capture complex network relations and predict the missing links between network resources. The GNNs have better performance in predicting the link than the traditional machine learning algorithms [7]. This paper presents the implementation details of how GNN can be utilized in link prediction tasks to manage an SDN network efficiently. The aim of utilizing GNN is to achieve better resource utilization, improved network management, and increased accuracy.

2. Implementation details

This section discusses the implementation details of various techniques which can be utilized for link prediction in an SDN.

2.1 Node Embedding-based methods

In node embedding-based methods, features of nodes (switches and routers) are used as embeddings or feature vectors to capture the relations between nodes and predict the link between nodes. The feature vectors are input to GNN, and GNN finds patterns between different nodes inside the network. In the case of node embedding-based methods, network device properties, network topology, network traffic metrics, and performance metrics can be used as embedding features. The GNN generates a likelihood score, which indicates the link between nodes in the network.

2.2 Edge embedding-based methods

The edge embedding-based methods are similar to node embedding-based methods. The difference between these two methods is that the edge embedding methods consider the edge features instead of the node features to predict the links between nodes. In the context of edge embedding methods, the metrics such as bandwidth, delay, jitter, and packet loss can be used as feature vectors for link prediction. Although in the link prediction task, using the edge features to find the link between two network nodes is unnecessary. However, the edge embedding methods are useful in finding the possibility that the link between two nodes can change depending on the network condition.

2.3 Hybrid embedding-based methods

The hybrid embedding-based methods utilize node and edge embeddings as feature vectors. The main purpose of hybrid embedding is to consider local and global features for the link prediction task. Hybrid embedding has improved performance by considering both node and edge embedding. During the initial step of finding the link between two nodes, the node embedding of the two nodes is calculated. After that, the edge embedding of the corresponding edge of the two nodes is estimated. Both embeddings are concatenated and passed to neural network layers to calculate the likelihood of a link between two nodes.

2.4 Attention-based methods

Attention-based methods are another technique to estimate the existence of a link between two nodes of the network. Attention-based methods work based on the importance of the node and edge inside the network. These methods allow the GNN to assign weights to each link and node based on their importance. Those nodes and edges are then used to predict the link between any two nodes in the network.

2.5 Spatial-temporal methods

Spatial-temporal methods rely on spatial and temporal information to predict links between two nodes. Spatial information refers to the static network topology, whereas temporal information means the information of the network when the network changes dynamically. These methods are used to predict the network link where the network information changes over time. These methods are useful in scenarios where the nodes are mobilized, as in mobile ad-hoc networks or vehicular networks.

3. Conclusions

This paper discusses the methods for utilizing the GNN in the problem area of link prediction in complex networks. For this purpose, this paper presented five methods that can be used for the link prediction task by utilizing the potential of GNN. Each method is discussed with its significance and the data type it takes as input. The discussed methods can be utilized in an SDN network for link prediction between two nodes. Moreover, these methods can also be utilized for topology prediction, node classification, and graph classification.

References

[1] W. Xia, Y. Wen, C. H. Foh, D. Niyato, and H. Xie, "A survey on software-defined networking," *IEEE Communications Surveys & Tutorials,* vol. 17, no. 1, pp. 27-51, 2014.

[2] A. Lara, A. Kolasani, and B. Ramamurthy, "Network innovation using openflow: A survey," *IEEE communications surveys & tutorials,* vol. 16, no. 1, pp. 493-512, 2013.

[3] L. Zhu, M. M. Karim, K. Sharif, C. Xu, F. Li, X. Du, and M. Guizani, "SDN controllers: A comprehensive analysis and performance evaluation study," *ACM Computing Surveys (CSUR),* vol. 53, no. 6, pp. 1-40, 2020.

[4] N. N. Daud, S. H. Ab Hamid, M. Saadoon, F. Sahran, and N. B. Anuar, "Applications of link prediction in social networks: A review," *Journal of Network and Computer Applications,* vol. 166, pp. 102716, 2020.

[5] A. Kumar, S. S. Singh, K. Singh, and B. Biswas, "Link prediction techniques, applications, and performance: A survey," *Physica A: Statistical Mechanics and its Applications,* vol. 553, pp. 124289, 2020.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE transactions on neural networks and learning systems,* vol. 32, no. 1, pp. 4-24, 2020.

[7] M. Zhang, and Y. Chen, "Link prediction based on graph neural networks," *Advances in neural information processing systems,* vol. 31, 2018.

Copyright ⓒ 2023 KSII

Road Hazardous Object Recognition System for the Visually Impaired

Sung-Hyun Kang, Yong-Jin Joo, Ji-Min Son, Yoon-Jung Yu, Sang-Jin Lee, Jun-ho Ahn*
Department of AI Computer Science and Engineering, Kyonggi University
Suwon, South Korea
[e-mail: tsugu00@kyonggi.ac.kr, jyj8943@kyonggi.ac.kr, annie1591@kyonggi.ac.kr, opi2982@gmail.com, vj852741@kyonggi.ac.kr, jha@kyonggi.ac.kr]
*Corresponding author: Junho Ahn

Abstract

The visually impaired, particularly those with severe vision impairment, faces difficulty when going outside due to limited vision. For overcoming such difficulty, this study designed and implemented a system that detects obstacles using YOLO and notifies users of the presence of obstacles through audible signals such as TTS to enable users to preemptively avoid obstacles when transporting. An obstacle detection program and pedestrian walkway guidance API are integrated to provide the following services for securing the mobility rights of the visually impaired: walkway guidance for avoiding high-slope ways and stairs on which the persons with limited vision cannot easily walk, obstacle avoidance, reconfiguring the guidance for detour routes if users cannot pass through the originally given path, and providing information of public transportation and crosswalks through beacons.

Keywords: Visually impaired, Object recognition, Artificial Intelligence camera, Real-time detection

1. Introduction

The advancement of technology in general has enabled miniaturization of chipsets and sensors, performance improvement, and cost reduction. Technological development has allowed advanced services to be readily used by the public on smartphone and wearable devices. In particular, research and development efforts are consistently put forth toward services that provide assistance to the elderly or the disabled who are mobility disadvantaged.

Certain products have been developed where information is identified through a braille display since the visually impaired can read through braille, but images or graphs cannot be easily deciphered without visual information. Furthermore, eye glasses products that output auditory information into visual information using AI speech-to-text technology have been developed for the hearing impaired persons who have difficulty in recognizing auditory information in conversations.

This study thus aims to design a system that detects obstacles and provides the optimal guidance using the detected data for the visually impaired persons who have limited vision.

2. Reconstruction of Volumetric Models

2.1 Related Work/Information

Insufficient management of braille blocks has a negative impact on the visually impaired. Accordingly, research is currently underway to find solutions using the YOLO systems [1][2].

This study proposes an algorithm based on YOLO for a system that delivers the information of braille blocks to the visually impaired in a smartphone environment. Focusing on the Go block and Stop block which directly affect walking, the system is designed to capture braille blocks using a smartphone camera and send them to a PC in real time, and recognize the images using a braille block recognition device including GPU to convert into voice data to be transmitted to a smartphone. In addition, the

algorithm is designed to be adjustable according to the angle because images need to be captured by the camera at the height similar to a user's height. Virtual braille block data were used to train the YOLOv3 based algorithm according to angle, and research is currently being conducted on an algorithm with high image recognition accuracy and performance through a numerical analysis on loss, mAP, precision, recall, and IoU based on YOLO.

Insufficient management of braille blocks has a negative impact on the visually impaired. Accordingly, research is currently underway to find solutions using the YOLO systems [1].

This study proposes an algorithm based on YOLO for a system that delivers the information of braille blocks to the visually impaired in a smartphone environment. Focusing on the Go block and Stop block which directly affect walking, the system is designed to capture braille blocks using a smartphone camera and send them to a PC in real time, and recognize the images using a braille block recognition device including GPU to convert into voice data to be transmitted to a smartphone. In addition, the algorithm is designed to be adjustable according to the angle because images need to be captured by the camera at the height similar to a user's height. Virtual braille block data were used to train the YOLOv3 based algorithm according to angle, and research is currently being conducted on an algorithm with high image recognition accuracy and performance through a numerical analysis on loss, mAP, precision, recall, and IoU based on YOLO.

2.2 Previous studies (Previous data)

There is a need for a function that recognizes and notifies users of any obstacles that may hinder walking. For such function, this study uses a YOLOv5 model having the highest computation capacity among various deep-learning based image recognition models for obstacle object detection. The YOLO model is trained by having obstacles easily found on streets such as power poles, trees, bollards, and stairs as a dataset.



Fig. 1. Vision-based stair detections

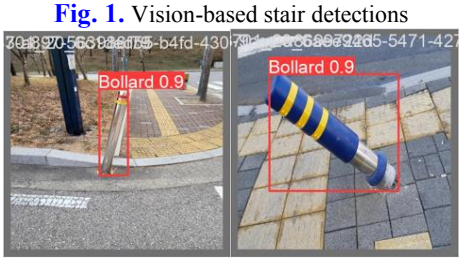


Fig. 2. Vision-based bollard detections

Executing YOLO using the model trained with the dataset produces the results shown in **Fig.s 1** and **2**. As shown in the figures, the objects included in the dataset were recognized by the YOLO model through rectangular boxes and labels.

2.3 The proposed architecture

The visually impaired frequently experiences restrictions when walking or transporting due to limited vision. For example, the obstacles impeding walking include stairs, curbs, streetlights, trees, and crosswalks; the visually impaired also have difficulty in finding bus numbers, subway lines, and terminal stops or stations when using public transportation. Stairs and roadside curbs can cause even the persons with disabilities to trip over easily, resulting in accidents and injuries. Such obstacles may result in even more severe accidents or injuries for the visually paired with limited vision; streetlights and trees can cause serious injuries when hit head-on. The visually impaired also have to use public transportation in some cases; however, since they cannot visually identify bus numbers, they inevitably have to ask the drivers for bus numbers one by one.

Thus, this study aims to apply YOLO and beacon to improve the difficulties faced by the visually impaired persons when

transporting.

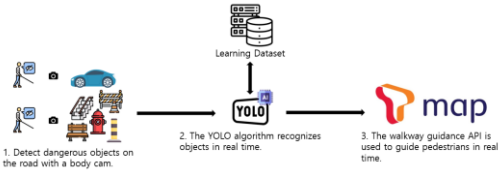


Fig. 3. Architecture for the road hazardous object recognition system

As shown in **Fig.s 1** and **2**, the YOLO model that has learned obstacle objects can be used to detect obstacles through real-time image recognition. The distance between an obstacle and a user is calculated through a real-time obstacle detection system that has been constructed by installing a body camera that replaces a user's eyes, and then if the detected obstacle is considered a risk factor for walking, it is notified through text-to-speech (TTS) technology. The object detection rate may decrease if images captured at night are not clear. Infrared sensors of a camera can be used to calculate the distance from an obstacle for enhanced reliability of nighttime obstacle detection. Accidents occurring during walking can be reduced by notifying users of the presence of obstacles through voice.

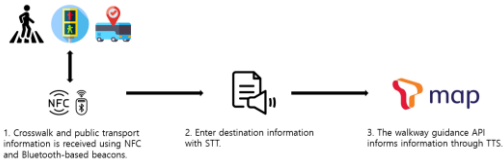


Fig. 4. Beacon-based traffic infrastructure detections

When a beacon is installed in public transportation means or near crosswalks, information related to public transportation or crosswalks such as bus numbers or traffic signals can be transmitted through the beacon and notified to users through TTS. The tagless public transportation system using NFC and beacon has recently been commercialized in Korea [3]. Therefore, the proposed service can be implemented without additionally installing beacons because the tagless public transportation system has already been established.

The walkway guidance service provides a guidance with the least amount of slope ways and alleys in which the visually impaired persons typically have difficulty in walking. The Google Map [4] result show the slope of the path, while the Naver Map [5] result shows the path mostly incorporating main roads. The optimal path avoiding high-slope way and alleyways can be derived using the information.

4. Conclusions

This study designed a walkway guidance system using YOLOv5 and beacon for guaranteeing mobility rights of the visually impaired. The designed system provides a guidance through TTS for paths mostly involving flatland and main roads on which the visually impaired person can walk relatively more easily, notifies users when obstacles that hinder walking are detected, and guides users through a detour route if they cannot pass through. Beacons are utilized to inform the visually impaired of the public transportation information through TTS. Resultantly, the mobility rights of the visually impaired are expected to be guaranteed by lowering the accident rates.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2021-0-01393)

References

[1] Oh, S. R., Bae, Y. C., "Braille Block Recognition Algorithm for the Visually Impaired Based on YOLO V3“, Korean Institute of Intelligent Systems, Journal of Korean Institute of Intelligent Systems Volume 31 Issue 1, pp.60-67, 2021.

[2] Cho, S, h., Kim, H. J., Park, S. s., Choi, Y. j, Lee, S. w. “A Mobile-based Walking Danger Notification System for Visually Impaired People”, Proceedings of the 2021 Korean Computer Information Society Conference, Vol. 29, No. 2, pp.25-28, 2021

[3] Kim, S. J., Kim J. S., “Tagless Bus Fare Payment System: Evaluation of Pilot Project and Policy Implication” Gyeonggi Research Institute.

[4] Google, “Google map”, https://www.google.com/maps

[5] Naver “Naver map” https://map.naver.com/

Copyright ⓒ 2023 KSII

Enhancing Additive Manufacturing Processes with Machine Learning Algorithms: Challenges and Opportunities

Oakyoung Han*
University College, Sungkyunkwan Univ*ersity
Seoul, South Korea
[e-mail: oakyoung@skku.edu]
*Corresponding author: Oakyoung Han

Abstract

Additive manufacturing, commonly known as 3D printing, is a fast-developing technology that has diverse applications across various industries. The integration of machine learning algorithms has proved to be a powerful approach for enhancing the accuracy and efficiency of additive manufacturing processes. This paper provides a comprehensive overview of the different types of machine learning algorithms used in additive manufacturing, including supervised, semi-supervised, unsupervised, and reinforcement learning. Moreover, this study highlights the key challenges faced during the implementation of machine learning algorithms in additive manufacturing, including the need for high-quality data, complex manufacturing processes, and error detection mechanisms. Finally, the research concludes with a discussion of the current state of the art and future research directions in the field of machine learning for additive manufacturing.

Keywords: Additive manufacturing, 3D printing, machine learning algorithms, manufacturing processes, error detection

1. Introduction

Additive manufacturing has witnessed remarkable growth in recent years, revolutionizing the manufacturing industry by enabling the production of complex geometries with high precision and customization. However, to fully unlock the potential of additive manufacturing, there is a need to optimize the manufacturing processes and ensure the production of high-quality and reliable products[1]. One promising approach to address these challenges is the integration of machine learning algorithms in additive manufacturing[2].

In the context of additive manufacturing, machine learning techniques offer opportunities to enhance the manufacturing processes by leveraging the power of data-driven decision-making and predictive modeling. By analyzing large volumes of data generated during additive manufacturing processes, machine learning algorithms can extract patterns, identify correlations, and make predictions that can lead to process optimization and improved product quality[3].

The objectives of the paper is to explore the challenges and opportunities of enhancing additive manufacturing processes through the application of machine learning algorithms. I will review the existing literature, discuss the current state of research, and identify potential areas for leveraging machine learning techniques[4]. Additionally, I will address the challenges and limitations associated with machine learning in additive manufacturing.

2. Challenges in Additive Manufacturing Processes

Additive manufacturing processes, while offering numerous advantages, also present

This research was supported by Seoul Campus Town Technology R&D Project (Real-time Detection System for Product Defects During the 3D Printing Process)

several challenges that need to be addressed for optimal performance and reliability. Understanding and mitigating these challenges is essential to fully unlock the potential of additive manufacturing[5].

2.1 Geometrical Complexity

Additive manufacturing enables the production of highly complex geometries with intricate designs that may be difficult or impossible to manufacture using traditional methods[6]. However, the inherent complexity of these designs poses challenges in terms of process planning, support structures, and post-processing. Ensuring the accurate reproduction of complex geometries and maintaining dimensional accuracy throughout the additive manufacturing process is an ongoing challenge.

2.2 Characterization

The selection of suitable materials for additive manufacturing is critical to achieve desired mechanical properties, functionality, and performance of the final product. However, identifying and characterizing materials suitable for specific additive manufacturing processes can be challenging[7]. The wide range of available materials, variations in material behavior during processing, and limited data on material properties for additive manufacturing are obstacles that must be overcome to optimize material selection and ensure consistent quality.

2.3 Process Parameters Optimization

Additive manufacturing processes involve a multitude of process parameters, such as laser power, scanning speed, layer thickness, and build orientation, which significantly affect the final part quality. Determining the optimal combination of these parameters is a complex task that requires iterative experimentation and process optimization. Efficiently identifying the optimal process parameters to achieve desired mechanical properties, surface finish, and build time remains a significant challenge[8].

2.4 Defect Detection

Additive manufacturing processes are susceptible to various defects, including surface roughness, porosity, layer misalignment, and thermal distortions. Detecting and mitigating these defects in real-time during the build process is crucial for ensuring the production of high-quality parts[9]. However, the dynamic nature of the additive manufacturing process, limited visibility, and complex defect detection algorithms pose challenges in implementing effective in-process quality control strategies.

2.5 Failure Prediction

Build failures can lead to wasted time, materials, and resources. Predicting and preventing build failures is crucial to enhance the reliability and efficiency of additive manufacturing processes[10]. However, accurately predicting failure modes and understanding the complex interactions between process parameters, material behavior, and part geometry remains a challenge, requiring advanced modeling and simulation techniques.

2.6 Post Processing

While additive manufacturing offers the advantage of near-net-shape manufacturing, post-processing steps are often necessary to achieve the desired surface finish, dimensional accuracy, and functional properties. Post-processing challenges include the development of efficient support removal methods, surface treatment techniques, and integration with traditional machining or assembly processes. Automation and optimization of post-processing steps are crucial for streamlining the overall manufacturing workflow[11].

Addressing these challenges requires interdisciplinary research efforts, including advancements in materials science, process optimization techniques, sensing and monitoring technologies, and data-driven decision-making. Overcoming these challenges will enhance the reliability, scalability, and adoption of additive manufacturing processes across various industries and applications[12].

3. Machine Learning Algorithm

3.1 Supervised Learning Algorithms

Supervised learning algorithms are trained on labeled data, where the input features and corresponding output labels are known. In additive manufacturing, these algorithms can be used for tasks such as defect detection, quality control, material behavior prediction, and part classification. Examples of supervised learning algorithms include support vector machines, decision trees, random forests, and gradient boosting.

3.2 Unsupervised Learning Algorithms

Unsupervised learning algorithms are applied to unlabeled data, where the goal is to discover patterns, clusters, or relationships within the data. These algorithms can be used in additive manufacturing for tasks such as anomaly detection, clustering similar parts, or identifying hidden structures in the data. Examples of unsupervised learning algorithms include k-means clustering, hierarchical clustering, and self-organizing maps.

3.3 Reinforcement Learning Algorithms

Reinforcement learning algorithms involve an agent interacting with an environment to learn optimal actions based on rewards and punishments. In additive manufacturing, reinforcement learning can be employed for process optimization, adaptive control, and autonomous decision-making. reinforcement learning algorithms, such as Q-learning or deep Q-networks (DQN), can learn optimal policies by exploring and exploiting the additive manufacturing process space.

3.4 Deep Learning and Neural Networks

Deep learning techniques, specifically neural networks, have gained significant attention in recent years due to their ability to handle complex patterns and large amounts of data. In additive manufacturing, neural networks can be used for tasks such as defect detection, material behavior modeling, generative design, and predictive maintenance. Convolutional neural networks (CNNs) are commonly employed for image-based tasks, while recurrent neural networks (RNNs) are suitable for sequential data analysis.

By leveraging these machine learning algorithms, additive manufacturing can benefit from enhanced process optimization, defect detection, material behavior prediction, generative design, and predictive maintenance[13]. The selection of appropriate machine learning techniques depends on the specific task, available data, and desired outcomes in additive manufacturing. Continued research and development in machine learning for additive manufacturing will further advance the field and unlock new possibilities.

4. Enhancing Process

Generative Adversarial Networks have emerged as a powerful tool for enhancing additive manufacturing processes. The following is a system developed by applying GAN. Fig. 1 represents the layered images and 3D image of the additive manufacturing.

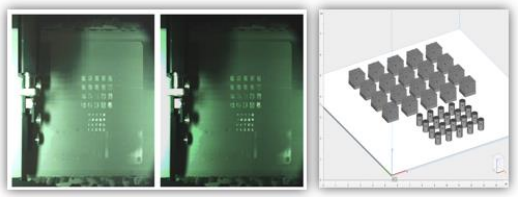


Fig. 1 .Layered images and 3D image of the product

After generating a predictive model by applying optical flow, error detection is improved through application of GAN. It is determined through GAN that defect detection occurs, and in this case, re-printing is executed after solving the problem for a defect-free product like in Fig. 2.

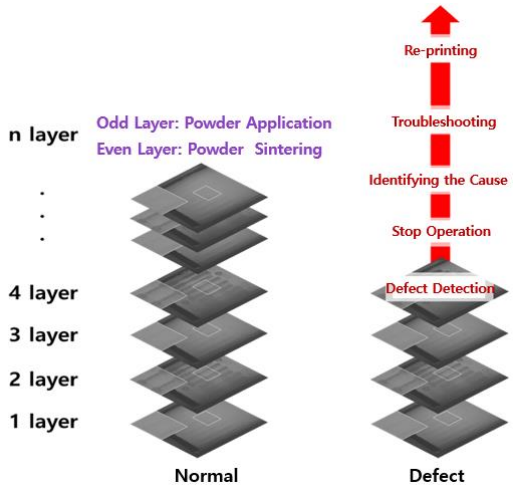


Fig. 2 .Normal and defect case fo AM

5. Opportunities

One significant opportunity for the application of machine learning in additive manufacturing is the integration of real-time sensor data. By incorporating sensor data into machine learning models, it becomes possible to monitor and control the additive manufacturing process in real-time.

Understanding how machine learning models make decisions or predictions is crucial for gaining trust and acceptance in the manufacturing industry. Developing techniques to explain the reasoning behind machine learning predictions and providing interpretability will facilitate better adoption of machine learning in additive manufacturing and enable experts to validate and refine the models. In conclusion, the opportunities for machine learning in additive manufacturing are vast and evolving. The integration of real-time sensor data, development of explainable and interpretable machine learning models, data sharing and collaboration, ensuring cybersecurity and data privacy, and machine learning are key areas that hold promise for the future of additive manufacturing. The system demonstrated the potential of GANs in enhancing the AM process by improving defect detection and enabling real-time corrective actions. The combination of optical flow techniques and GANs provides a comprehensive approach to optimize the AM workflow and ensure the production of high-quality, defect-free products.

References

[1] Z. Liu, D. Zhao, P. Wang et al., "Additive manufacturing of metals: Microstructure evolution and multistage control," *Journal of Materials Science & Technology*, Vol. 100, pp. 224-236, 2022.

[2] J. Qin, F. Hu, Y. Liu et al., "Research and application of machine learning for additive manufacturing," *Additive Manufacturing*, Vol. 52, pp. 1-25, 2022.

[3] J. Jiang, Y. Xiong, Z. Zhang, & D. W. Rosen, "Machine learning integrated design for additive manufacturing," *Journal of Intelligent Manufacturing*, Vol. 33, No. 4, pp. 1073-1086, 2022.

[4] R. Raju, N. Manikandan, D. Palanisamy et al., "A review of challenges and opportunities in additive manufacturing," *Recent Advances in Materials and Modern Manufacturing: Select Proceedings of ICAMMM 2021*, pp. 23-29, 2022.

[5] B. A. Praveena, N. Lokesh, A. Buradi et al., "A comprehensive review of emerging additive manufacturing (3D printing technology): Methods, materials, applications, challenges, trends and future potential," *Materials Today: Proceedings*, Vol. 52, pp. 1309-1313, 2022.

[6] R. Scimone, T. Taormina, B. M. Colosimo et al., "Statistical modeling and monitoring of geometrical deviations in complex shapes with application to Additive Manufacturing," *Technometrics*, Vol. 64, No. 4, pp. 437-456, 2022.

[7] J. Jiang, & Y. Ma, "Path planning strategies to optimize accuracy, quality, build time and material use in additive manufacturing: a review," *Micromachines*, Vol. 11, No. 7, pp. 1-18, 2020.

[8] J. Plocher, & A. Panesar, "Review on design and structural optimisation in additive manufacturing: Towards next-generation lightweight structures," *Materials & Design*, Vol. 183, pp. 1-20, 2019.

[9] Y. Chen, X. Peng, L. Kong et al., "Defect inspection technologies for additive manufacturing," *International Journal of Extreme Manufacturing*, Vol. 3, No. 2, pp. 1-21, 2021.

[10] S. Afazov, A. Serjouei, G. J. Hickman et al., "Defect-based fatigue model for additive manufacturing," *Progress in Additive Manufacturing*, pp. 1-8, 2022.

[11] X. Peng, L. Kong, J. Y. H. Fuh, & H. Wang, "A review of post-processing technologies in additive manufacturing," *Journal of Manufacturing and Materials Processing*, Vol. 5, No. 2, pp. 1-23, 2021.

[12] O. Abdulhameed, A. Al-Ahmari, W. Ameen, & S. H. Mian, "Additive manufacturing: Challenges, trends, and applications," *Advances in Mechanical Engineering*, Vol. 11, No. 2, pp. 1-27, 2019.

[13] J. Jiang, "A survey of machine learning in additive manufacturing technologies," *International Journal of Computer Integrated Manufacturing*, pp. 1-23, 2023.

Copyright ⓒ 2023 KSII

Mutually Exclusive Decision-Making Learning of AI based on Contexts

Hyun Jung Lee*
Department of Urban Policy Research, Goyang Research Institute
Goyang, South Korea
[e-mail: hjlee5249@gmail.com]
*Corresponding author: Hyun Jung Lee

Abstract

This paper presents Mutually Exclusive Decision-Making Learning (MEDML) which is based on contextual conditions including different temporal perspectives based on the specific timeframe, considerable level of process and decision-making criteria. ChatGPT using LLM as generative AI is trained by Reinforcement Learning with Human Feedback (RLHF) based on probabilistic data. However, probabilistic data don't guarantee that it will always lead to the accurate results of decision makings. Especially, in mutually exclusive decision process, it is necessary to consider contextual variables which can be dynamically changed by considerable conditions like the extent or considerable process level, specific timeframe to be considered, decision-making criteria, etc. In mutually exclusive decision making process, it is necessary to determine the perspectives based on the considerable contexts by decision-maker's specific criteria. In addition, it can provide explanation of the derived results in decision making process by AI. In this paper, the proposed MEDML is to support the critical decision making depending on criteria of AI and human interaction. It is possible to support decision-making in mutually exclusive decision processes through the human involvement by consideration of contextual variables and provide the explanation of the derived results in decision-making process.

Keywords: Mutually Exclusive Decision-Making Learning, Contexts, Explainable Artificial Intelligence

1. Introduction

The chatbot ChatGPT, developed by OpenAI, demonstrated the potential for generalization to solve problems using Artificial Intelligence (AI) based on Reinforcement Learning with Human Feedback (RLHF) using Large Language Model (LLM). Generative AI, as opposed to simple computation or factual presentation, is utilized for dynamic and complex problem-solving tasks such as generating results based on making context-based decisions.

In ChatGPT, reward function using reinforcement learning is activated by previously trained-agent by probabilistic data and human feedback as in “AlphaGo”. However, in mutually exclusive decision learning, reward function can be activated by contextual variables such as the extent or level of considerable process, specific timeframe to be considered, decision-making criteria, etc. When mutually exclusive decision making is needed like resource allocation, competitive situaitons, priority determination, etc., it is not easy to derive decision-making by the trained-agent which is learned by probabilistic data. The mutually exclusive choices may dynamically vary depenging on the cotextual variables and values based on the specific context and circumstances. In addition, there are other criteria that can be considered such as user’s intent, ethiccal considerations, risk-reward tradeoff, etc. Therefore, it does not guarantee that the drived decision-mking based on probabilities is always right. This paper focuses on Mutually Exclusive Decision-Making Learning (MEDML) of Artifical Intelligence(AI) under consideration of contexts and circumstances.

2. Literature Review

Generative AI aims to generate approriate resposponses depending on context according to user's intent. Therefore, ChatGPT[5] as generative AI aims to generate coherent and contextually appropriate responses. To do this, ChatGPT is based on reinforcement learning with human feedback (RLHF) in Large Language Model (LLM). RLHF is adopted to reduce the incorrectness and inaccuracy of the information provided by ChatGPT through the verification and cross-reference with human. Even if, OpenAI adopted RLHF to filter inappropriate, offensive, or harmful content, ChatGPT is still generating inaccurate or irrelevant responses using plausible but incorrect information such as hallucination[5] and does not provide any explanation as black-box[1]. To overcome these kind of problems, researches focus on Explainable Artificial Intelligence (XAI)[1,6] to provide discussions, evidence and resources for decison-making[2,3,4,7,8].

In ChatGPT as a kind of LLM, RLHF is provided to enhance the learning process of an RL agent [5]. In RLHF, human feedback is used to guide the agent's decision-making and improve its performance[5]. However, all of decision-making can not make a same decision.

In this research, the proposd MEDML is for the mutually exclusive learning of AI agent for deicision-making under consideration of contextual variables and values. When the AI derive different decisions acoording to the considerable contexts, the conditions and variables as contexts given in the decision-making process of AI can be explanations or justifications for the derived results[1,4,6].

3. Framework of MEDML

3.1 MEDML Problem-Solving Process

When mutually exclusive decisions are required, it is not easy to derive making-decision to use RLHF based on probabilistic learning in problem solving process. To make mutually exclusive decisions, it is necessary to learn the mutually exclusive decision with Human feedback depending on contextual variables and values in the decision process.

The proposed Mutually Exclusive Decision-Making Learning (MEDML) is based on contextual conditions like the specific timeframe, considerable depth of process and decision-making criteria.

In mutually exclusive decision-making process depending on the contexts, AI-agent needs to be trained by human decision feedback when it needs to make a mutually exclusive decision. Furtherly, the proposed mutually exclusive learning with human decision feedback can provide explanations of the derived results by AI.

Fig. 1 shows the problem-solving process of the MEDML that is activated by mutually exclusive decision-making process which needs pruning branches by human depending on contexts like tmeframe, criteria, the considerable depth of process.

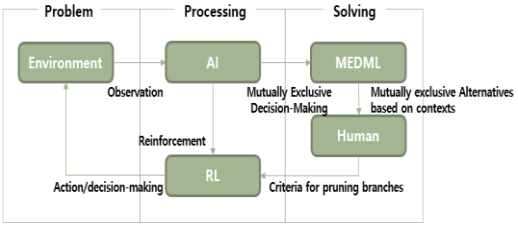


Fig. 1. MEDML Problem-Solving Process

3.2 The features of MEDML

It is possible to consider two types of learning in AI. One is reinforcement learning, which is based on probabilistic data and is suitable for goal-seeking in graph structures. So, it can achieve the desired goal regardless of which processes are selected in the reasoning process. The other involves pruning branches in learning for mutually exclusive decision-making. If the predicted reward is mutually exclusive, it is necessary to prune the branches in the decision tree based on contexts, as shown in **Fig. 2**.

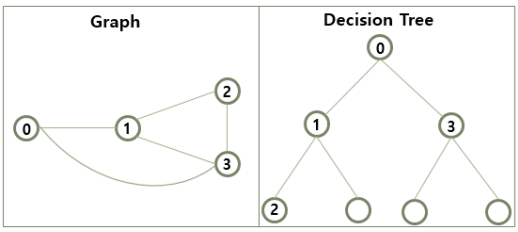


Fig. 2. Graph and Decision Tree for AI Learning

Table 1 presents the comparison of the features in reinforcement and mutually exclusive learning.

Table 1. The Comparison Mutually Exclusive and Reinforcement Learning

Features	Reinforcement Decision-Making	Mutually Exclusive Decision-Making
Conceptual Structure	Graph	Decision Tree
Process	Fine-tuning	Pruing-branches
Reward function	Statically predicted based on probability	Dynamically changes based on contexts
Problem-Solving	Predicted like "AlphaGo"	Unpredicted be changed
Interaction	Data	Human
AI Model	Blak-box	White-box (XAI)

In reinforcement learning, a graph structure is suitable for solving decision-making problems where the agent learns through trial and error, adjusting its actions based on observed rewards to improve decision-making abilities over time. This process involves fine-tuning, static prediction of reward functions, and interaction with data. On the other hand, in mutually exclusive decision-making processes, a decision tree is commonly used to construct the data or processing structure. The agent learns to prune branches, interacts with humans, and receives feedback as decision-making criteria based on contextual variables and values. This process includes pruning branches and dynamically predicting reward functions. The agent interacts with humans to achieve the goal of decision-making. When AI requires mutually exclusive decision-making, it is possible to provide explanations about the decision-making process at each point of interaction between the agent and humans.

3.3 Contexts for MEDML

For mutually exclusive decision-making, the agent needs to interact with humans to determine the decision criteria based on contextual variables and values. The contextual variables encompass specific timeframes, a substantial depth of the decision process, and the criteria for decision-making.

Fig. 3 shows an example of timeframe as a contextual variable for MEDML. Decision tree has two-branch such as *A* and *B* which are mutually exclusive. As inclusive variables, *A* an d *B* are comprised of variables to maximize the cumulative reward such as $A=\{a_i^+ \mid i = 1, \ldots, p\}$ and $B=\{b_j^- \mid j = 1, \ldots, q\}$. However, *A* and *B* are completely mutually exclusive relationship like $A \cap B = \emptyset$ which is comprised of exclusive variables. *T* is timeframe which is comprised of *t, t-1, t+1*. *t* is present which is related current moment. *t-1* is past which is referreing to events, experiences, or phenomena that have already occurred. *t+1* is future which is pertaining to events, possibilities, or conditions that are yet to happen. According to the specific timeframe being considered, it lead to different decision-making like *A* or *B*. *A* is the decision-making at time *t-1* and *B* is at time *t+1*.

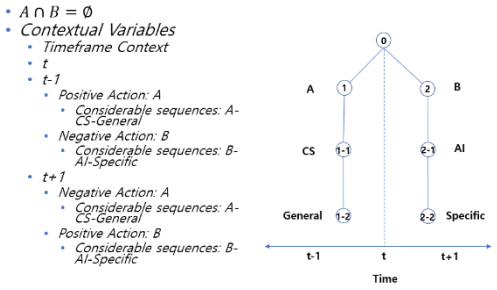


Fig. 3. An Example of Timeframe as a Contextual Variable in MEDML

Timeframe $= \{t \pm i \mid i \geqq 0\}$, *where t is time condition, i is a considerable level of time status to construct timeframe.*

Process $= \{p \pm j \mid j \geqq 0\}$, *where p is process or event in decsion-making process, j is a considerable depth of process.*

Criteria $= \{f_k \mid k \geqq 1\}$, *where* f_k *is* k^{th} *feature which affect decsion-making, k is a number of feature.*

Timeframe is comprised of past, present and future. Process is the level of decision-making progress. Criteria is applied to prune the mutually exclusive decision tree branches which are comprised of specific timeframe and depth of decision process. As the considerable criteria, there are demographic data, family relationships, etc. Each feature is weighted for the pruning process and tiemframe, the feature and weight are defined by human feedback.

To support the proposed mutually exclusive decision-making process, utility function is adopted for the specific timframes and decision-process levels associated with different possible decisions or choices. As in **Fig. 3**, it can be presented by $U = wA + (1\text{-}w)B$. Utility function is constructed by selection of the decision-process depending on contextual criteria and the weighted criteria for maximizing their expected utility.

$$U = \sum_i w_i e_i, where \sum_i w_i = 1$$

w_i is a weight of i^{th} event as e_i .

$$U = \{e_i | \, e_i \; is \; i^{th} \; decision \; event\}$$

U is comprised of e_i as i^{th} decision event which is determined by interaction with human. When mutually exclusive decision-making is needed, agent is affected by interaction with contextual criteria based on human feedback. The event is recorded like e_i. U is a set of e_i and can give the explanations of the decision-making consequences. In the process of selection of e_i, yield variables (y_j) are generated in mutually exclusive decision-making. If the criteria is changed by contextual variables and values, then new utility functions as alternatives will be generated including the y_j.

4. Conclusions and Further Study

In this paper, MEDML (Mutually Exclusive Decision-Making Learning) is proposed for decision-making processes that require mutually exclusive choices based on contextual variables and values. The contextual variables considered in this study include the specific timeframe, a significant level of decision process, and decision-making criteria.

For mutually exclusive decision-making, it is crucial to determine the temporal perspective using a time frame and the level of decision process, which are based on contextual criteria influenced by human interaction. A utility function is adopted to guide different possible decisions or choices based on contextual variables. The utility function represents a set of events that provide explanations for AI's decision-making. If the decision-making criteria with human interaction change, the utility function may be constructed differently by incorporating events from previous decision-making processes.

The proposed MEDML focuses on training for mutually exclusive decision-making. It differs from reinforcement learning in terms of data conceptual structure, training process, reward values, and functions, among other aspects. MEDML is expected to have applications in various areas of everyday life where mutually exclusive decisions are required.

In further studies, it is necessary to implement mutually exclusive decision-making in AI agents. Considering the contextual level in the decision-making process, it is not effective to consider all alternatives due to the enormous search space. Therefore, it is important to determine the appropriate search level for finding feasible or optimal solutions effectively. Additionally, it is crucial to define whether an agent should engage in interaction with humans.

References

[1] V. Belle, and I. Papantonis, "Principles and Practice of Explainable Machine Learning," *Frontiers in Big Data*, Vol 4, 2021.

[2] A. Bertrand, R. Belloum, J. R. Eagan, and W. Maxwell, "How Cognitive Biases Affect XAI-assisted Decision-making: A Systematic Review," *in Proc. of the 2022 AAAI/ACM Conference on AI, Ethics, and Society (AIES)*, 2022.

[3] G. Campitelli, and F. Gobet, "Herbert Simon's Decision-Making Approach: Investigation of Cognitive Processes in Experts," *Brunel University*, 2010.

[4] V. Harish, F. Morgado, A. D. Stern, and S. Das, "Artificial Intelligence and Clinical Decision Making: The New Nature of Medical Uncertainty," *Academic Medicine*, Vol 96, No. 1, pp. 31-36, 2021.

[5] OpenAI, "GPT-4 Technical Report," 2023.

[6] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the Game of Go with Deep Neural Networks and Tree Search," *Google DeepMind*, 2016.

[7] H. A. Simon, "Administrative Behavior: A Study of Decision-making Processes in Administrative Organizations," *The Tree Press*, 1997.

[8] J. Simonsen, "Herbert A. Simon: Administrative Behavior How organizations can be understood in terms of decision processes," *Roskilde University*, 1994.

Copyright ⓒ 2023 KSII

Retrieval-based Question-Answer Generation

Seonmin Koo[1], Jungwoo Lim[1], Suhyune Son[1], Yuna Hur[2], and Heuiseok Lim[1*]
Department of Computer Science and Engineering, Korea University[1], Human-Inspired AI Research[2]
Seoul, South Korea
[e-mail: fhdahd@korea.ac.kr, wjddn803@korea.ac.kr, ssh5131@korea.ac.kr, yj72722@korea.ac.kr, limhseok@korea.ac.kr]
*Corresponding author: Heuiseok Lim

Abstract

Research on question-answer generation (QAG) is widely conducted since it derives an effective tool to help educators. Previous QAG studies utilize neural network model to improve the question generation quality. In this paper, we introduce QAG models utilizing retrieved information. We also analyze the results of improving question quality by retrieving more relevant information in answer generation. In addition, we conduct experiments on different units of information in the retrieving process. From the results, it is shown that different level of information from the retriever affects the question quality in end-to-end architecture. Moreover, we find that emphasizing the information at the passage level may enhance the question generation quality.

Keywords: Question-answer generation, Relevant information, Retrieval unit, Context-aware

1. Introduction

Question-answer generation (QAG) is a task that generates both question and answer based on input text. QAG contributes to important real-world applications, such as assisting educators by suggesting question-answer (QA) automatically for educators and their students [1,2]. However, obtaining effective QA that can measure how well one understands the given content is challenging since it requires an enhanced level of understanding capability of the human.

Despite the previous studies on neural QAG, there exists a significant challenge in obtaining high-quality QA pairs from the unstructured text at a large scale. One solves the task by providing relevant information when gaining the question. The study formulates the task as a sequence-to-sequence problem while designing various encoders and decoders to improve the quality of the generated questions [3,4]. On the other hand, [5] provide the question according to the type of the questions; implicit and explicit. However, distinguishing explicit and implicit questions lets the model confuse generating the questions due to the ambiguity between explicit and implicit questions. From the previous research, it is implied that capturing the relevant information within the passage is crucial when obtaining the question.

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01405) supervised by the IITP(Institute for Information & Communications Technology Planning & Evaluation). This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques). This research was supported by the Core Research Institute Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2021R1A6A1A03045425). This work was supported by Institute for Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2022-0-00369, (Part 4) Development of AI Technology to support Expert Decision-making that can Explain the Reasons/Grounds for Judgment Results based on Expert Knowledge).

Table. 1. Performance of the QAG system according to the number of beams in the Retrieval

Model	Beam K	Phrase		Sentence		Passage	
		BLEU	ROUGE	BLEU	ROUGE	BLEU	ROUGE
BART	1	0.184	0.517	0.113	0.415	0.177	0.527
	5	0.180	0.515	0.117	0.418	0.177	0.527
	10	0.179	0.516	0.112	0.412	0.174	0.525
T5	1	0.064	0.241	0.198	0.064	0.056	0.237
	5	0.070	0.252	0.217	0.070	0.068	0.262
	10	0.072	0.26	0.212	0.072	0.069	0.267

Since we assume extracting significant information from the given context produces effective QA pairs, we concentrate on capturing important chunks within the text. To this end, we utilize the dense retrieval component to find the key information from the given context. In order to produce a high-quality answer, supplementary information is inserted within the context. Generating questions based on the answers can lead to maintaining the quality of QA pairs. Also, we analyze the results to investigate the proper amount of information that is effective in generating the QA pairs according to retrieval units.

3. Proposed Method

We propose QAG system using retrieval results for crucial information to obtain both answers and questions. Our QAG system contains three components: retrieval model, question-based answer generation (AG) model, and answer-based question generation (QG) model. We use dense phrase retrieval (DPHR) to extract key information from the context. 3.1 Retrieval Model

3.1 Retrieval Model

We utilize DPHR [10] to retrieve important positions based on questions in a given passage. In the inference phase, phrase encoder and question encoder trained on Wikipedia were used to retrieve phrases by using the inner product as a similarity function between a phrase and a question. Also, we replace retrieval pool to FairyTaleQA dataset [12]. DPHR retrieve a slightly larger number of phrases, compute the score for each passage, and return top-k various unique units. We select units of representation from the phrases and sentences. The retrieval result is passed to the subsequent process using top-1.

3.2 Answer Generation (AG) Model

The extracted key information is used to obtain the answer. We use fine-tuned BART and T5 [6,11] to obtain the answer by inserting the retrieval output. The input format of the AG model is the concatenation of context, retrieved key information, and questions. We additionally utilize the retrieved output in the form of a phrase or sentence.

3.3 Question Generation (QG) Model

After obtaining the context-aware answers, we use the top-1 answer in the QG model. In other words, we generate a question with the context and the obtained answer. BART and T5 models are fine-tuned and used to gain questions. Therefore, we train the AG model and the QG model respectively, and finally provide the QA pairs.

4. Experiments

We compare the ability to generate QA pairs of the QAG framework according to the critical information units. The proposed system consists of three components: retrieval model, AG model, and QG model. we train BART and T5 [6,7] models to generate question-based AG and answer-based QG. During inference, we set the beam size to 1, 5, and 10. We evaluate the generated QA pairs on BLEU and ROUGE [8,9] scores by varying the number of beams to 1, 5, and 10. The results of automatic evaluation on generated QA pairs are shown in Table 1. In this experiment, the retrieved phrase is the top-1 phrase similar to the question in the given passage, and the sentence is the entire sentence containing a phrase as input.

For the BART model, the retrieval units of passages and sentences achieve the highest score when the number of beams is set to 5, while the

retrieval unit of phrases achieves the highest score when the beam size is set to 1. In the case of the T5 model, there is a tendency to obtain generally high scores when the beam size is set to 10, regardless of the range of the retrieval unit.

5. Conclusions

In this paper, we introduced QAG models utilizing retrieved information. Our method contains three components: retrieval, question-based AG, and answer-based QG. We evaluated our system using the BART and T5 models to examine their effectiveness. We found out that emphasizing the information at the passage level may enhance the question generation quality.

References

[1] Snyder, Lynn, Donna Caccamise, and Barbara Wise, "The assessment of reading comprehension: Considerations and cautions," *Topics in Language Disorders* 25.1, pp. 33-50, 2005.

[2] Xu, Ying, et al., "Same benefits, different communication patterns: Comparing Children's reading with a conversational agent vs. a human partner," *Computers & Education 161*, 104059, 2021.

[3] Xinya Du, Junru Shao, and Claire Cardie, "Learning to Ask: Neural Question Generation for Reading Comprehension," *in Proc. of the 55th Annual Meeting of the Association for Computational Linguistics,* vol 1, pp 1342–1352, 2017.

[4] Bang Liu, Mingjun Zhao, Di Niu, Kunfeng Lai, Yancheng He, Haojie Wei, and Yu Xu, "Learning to generate questions by learningwhat not to generate," *In The World Wide Web Conference,* pp 1106–1118, 2019.

[5] Zhenjie Zhao, Yufang Hou, Dakuo Wang, Mo Yu, Chengzhong Liu, and Xiaojuan Ma, "Educational question generation of children storybooks via question type distribution learning and event-centric summarization. In Proc. of the 60th Annual Meeting of the Association for Computational Linguistics, vol 1, pp 5073–5085, 2022.

[6] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer, "BART: Denoising sequence-to-sequence pre-training for natural language generation, translation, and com- prehension," *in Proc. of the 58th Annual Meet- ing of the Association for Computational Linguistics*, pp 7871–7880, 2020.

[7] Danqi Chen, Adam Fisch, Jason Weston, and Antoine Bordes, "Reading Wikipedia to Answer Open-Domain Questions," *in Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics,* vol 1, pp 1870–1879, Vancouver, 2017.

[8] Wenhan Xiong, Mo Yu, Shiyu Chang, Xiaoxiao Guo, and William Yang Wang, "Improving Question Answering over Incomplete KBs with Knowledge-Aware Reader," *in Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, pp 4258–4264, 2019.

[9] Akari Asai, Kazuma Hashimoto, Hannaneh Hajishirzi, Richard Socher, and Caiming Xiong, "Learning to retrieve reasoning paths over wikipedia graph for question answering," *arXiv preprint arXiv,* 1911.10470, 2019.

[10] Jinhyuk Lee, Alexander Wettig, and Danqi Chen, "Phrase Retrieval Learns Passage Retrieval, Too," *in Proc. of the 2021 Conference on Empirical Methods in Natural Language Processing*, pp 3661–3672, 2021.

[11] Colin Raffel, Noam Shazeer, Adam Roberts, Katherine Lee, Sharan Narang, Michael Matena, Yanqi Zhou, Wei Li, and Peter J Liu, "Exploring the limits of transfer learning with a unified text-to-text trans-former," *The Journal of Machine Learning Research*, 21(1):5485–5551, 2020.

[12] Ying Xu, Dakuo Wang, Mo Yu, Daniel Ritchie, Bing- sheng Yao, Tongshuang Wu, Zheng Zhang, Toby JiaJun Li, Nora Bradford, Branda Sun, et al., "Fantastic questions and where to find them: Fairytaleqa– an authentic dataset for narrative comprehension," *in Proc. of Annual Meeting of the Association for Computational Linguistics,* vol 1, pp 447 – 460

Copyright ⓒ 2023 KSII

Performance Analysis of Mixed Vibration Using RNN Models

Ju-Hyung Lee[1] **and Jun-Ki Hong**[2*]
[1]Department of Computer Engineering, Pai Chai University
Daejeon, South Korea
[2]Department of Smart Information Technology Engineering, Kongju University
Cheonan, South Korea
[e-mail: 2237012@pcu.ac.kr, jkhong@kongju.ac.kr]
*Corresponding author: jkhong@kongju.ac.kr

Abstract

This study investigates the performance of six recurrent neural network (RNN) models in predicting mixed time-series vibration data that includes both stationary and unsteady vibrations. To analyze the prediction accuracies of the six RNN models, the RNN models are trained on varying learning intervals of the mixed vibration. The simulation results show that increasing the learning interval improves the prediction accuracy for all six RNN models. Notably, the long short-term memory (LSTM) model exhibits the highest rate of accuracy improvement as the training interval increases. Furthermore, attention LSTM (Attn.-LSTM) and attention gated recurrent unit (Attn.-GRU) models achieve impressive accuracy in vibration prediction by effectively pre-calculating the similarity between input and predicted values.

Keywords: Deep learning, recurrent neural network (RNN), time-series mixed vibration prediction, long short-term memory (LSTM), gated recurrent unit (GRU), attention mechanism, bidirectional RNN

1. Introduction

Drones are a type of unmanned aerial vehicle that can fly without a pilot, and it is one of the core machineries that are in the spotlight in the era of the 4th industrial revolution. According to Global Drone Market Report by Drone Industry Insights, the drone market is valued at $55.8 billion in 2026, with the commercial marketing growing at 8.3% [1]. As such, research is being conducted in many fields related to drones, and drones are used in various industries such as agriculture [2], transportation [3], and lifesaving [4]. However, since drones are flying objects, human casualties may occur if the drone crashes. For example, if a drone crashes in an area with a large floating population, enormous human casualties may occur. In addition, there is a risk of secondary accidents such as traffic accidents caused by fragments of a crashed drone, and when a drone crashes into a mountain, there is a risk of a large-scale fire due to battery explosion. In fact, according to the results of a drone safety survey announced by the Korea Consumer Agency in 2017, human casualties caused by drones are increasing every year, and drone crashes accounted for about 20% of the causes of human casualties caused by drones [5]. Therefore, it is very important to prevent drone crashes by predicting future vibrations of the mixed vibrations from normal and abnormal vibrations generated by drone motors.

2. Methodology

2.1 Comparative Analysis Framework

This subsection presents the analysis framework used to compare the prediction performance of the recurrent neural network (RNN) models for time-series mixed vibration data, including long

short-term memory (LSTM) [6], attention-LSTM (Attn.-LSTM) [7], bidirectional-LSTM (Bi-LSTM) [8], gated recurrent unit (GRU) [9], attention-GRU (Attn.-GRU) [7], and bidirectional GRU (Bi-GRU) [8]. Figure 1 illustrates the proposed analysis framework, which consists of collecting normal and abnormal motor vibration data from drones, building training datasets using 40% and 60% of the collected data, and predicting vibrations in the remaining 60% and 40% segments using the six RNN models.

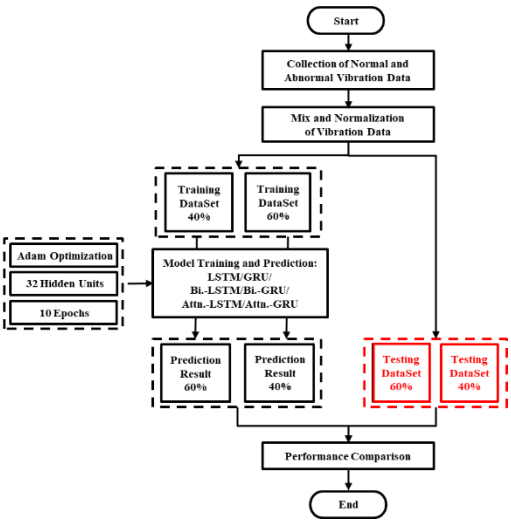


Fig. 1. Flowchart of comparative analysis framework

According to **Fig. 1**, steady and unsteady vibration data are collected from the drone's motor. Then, mixed vibration is generated from the steady and unsteady vibration data with the ratio of 5:5. This mixed vibration data is used to train and test six different RNN models to compare the prediction accuracies. To compare the prediction accuracy of the six RNN models for future mixed vibrations, the mixed vibration data is first divided into training datasets of 40% and 60% segments. Next, the six RNN models are trained on the training datasets and used to predict the remaining 60% and 40% segments. Finally, the prediction accuracies of the six RNN models are compared and analyzed to determine the most suitable model for predicting the mixed vibrations of drones.

2.2 Collection of Vibration Data

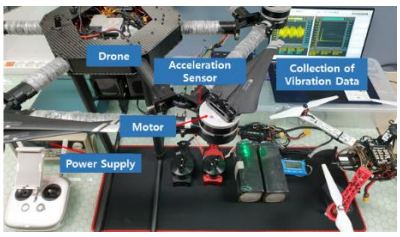


Fig. 2. Configuration of collecting time series vibration data from motor

Fig. 2 shows the environment in which vibration data is collected from motor of the drone. The mixed vibration was generated by combining the collected steady and unsteady vibration data. Further, **Fig. 3** illustrates the waveform of the mixed vibration obtained by combining the steady and unsteady vibrations.

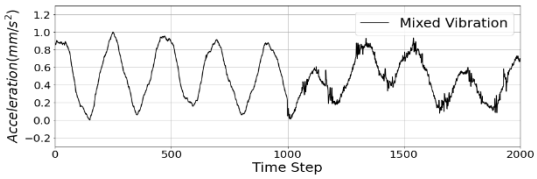


Fig. 3. Mixed vibration from steady and unsteady vibrations

As shown in **Fig. 3**, the ratio of normal vibration and abnormal vibration is 5:5 for mixed vibration. The prediction accuracy of mixed vibration with changes in training intervals is analyzed by LSTM, Attn.-LSTM, Bi-LSTM, GRU, Attn.-GRU, and Bi-GRU models.

3. Simulation Results

3.1. Simulation Parameters

This subsection describes the simulation parameters employed for predicting mixed vibrations using the six RNN models. **Table 1** describes the simulation parameters.

Table 1. Simulation parameters

Parameter	Value
No. of hidden units	32
Initial learning rate	0.0002
No. of epochs	10
Minimum batch size	32

Optimizer	Adam [10]
Iterations	20

3.2. Waveform of the Predicted Vibrations

This subsection aims to compare and analyze the waveforms of the vibrations predicted by the six RNN models from the mixed vibration data. The waveforms of the predicted vibrations from the mixed vibration data using the six RNN models are presented in **Fig. 4**.

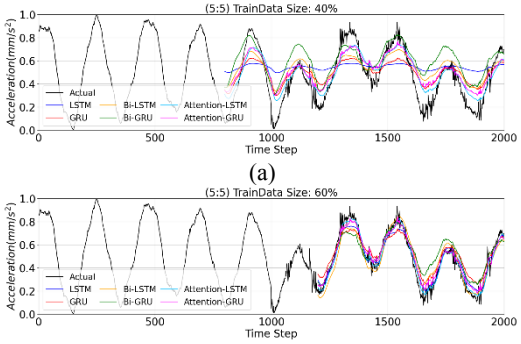


Fig. 4. Waveforms of the vibrations predicted from the mixed vibrations: (a) 40%, (b) 60% segments of training data

The results in **Fig. 4** demonstrate that the vibration prediction accuracy of all six RNN models increased as the size of the training set increased from 40% to 60%. Moreover, the accuracy of all RNN models improved with larger training segments. Specifically, when the training data size was 40%, the vibration prediction accuracy using LSTM was notably lower than the other five RNN models. However, LSTM exhibited the most significant improvement in vibration prediction accuracy when the training data size was increased to 60%.

4. Conclusions

This paper presents a comparative analysis of the prediction accuracy of six RNN models using mixed vibration data. The simulation results reveal that the prediction accuracy of mixed vibration increases with an increase in the size of the training data for all six RNN models. Moreover, it has been confirmed that the Attn.-LSTM and Attn.-GRU models, which calculate the similarity between the input value and the value to be predicted in advance and reflect it in the prediction, demonstrate the highest accuracy. These findings suggest that the attention mechanisms with the RNN models can effectively improve the prediction accuracy of mixed vibrations, and highlight the importance of selecting appropriate RNN models for accurate predictions.

References

[1] L. Schroth, "Global Drone Market Report 2022-2030," *Drone Industry Insights*, 2020.

[2] V. Mazzia, L. Comba, A. Khaliq, M. Chiaberge, P. Gay, "UAV and Machine Learning Based Refinement of a Satellite-Driven Vegetation Index for Precision Agriculture." *Sensors*, Vol. 20, No.9, p. 2530, 2020.

[3] L. Gupta, R. Jain, G. Vaszkun, "Survey of Important Issues in UAV Communication Networks," *IEEE Communications Surveys and Tutorials*, Vol. 18, No. 2, pp. 1123-1152, 2016.

[4] P. Yao, Z. Xie and P. Ren, "Optimal UAV Route Planning for Coverage Search of Stationary Target in River," *IEEE Transactions on Control Systems Technology*, Vol. 27, No. 2, pp. 822-829, 2019.

[5] Product Safety Team, "Drone Safety Status Survey Results," Korea Consumer Agency, 2017.

[6] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput*., Vol. 9, No. 8, pp. 1735-1780, 1997

[7] D. Bahdanau, K. Cho, Y. Bengio, "Neural Machine Translation by Jointly Learning to Align and Translate," " in *Proc*. of *the 3rd International Conference on Learning Representations*, 2015.

[8] M. Schuster and K. Paliwal, "Bidirectional Recurrent Neural Networks," *IEEE Trans. Signal Processing*, Vol. 45, No. 11, pp. 2673-2681. 1997.

[9] J. Chung, C. Gulcehre, K. Cho and Y. Bengio, "Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling," in *Proc*. of *Neural Information Processing Systems Workshop on Deep Learning*, 2014.

[10] D.P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," In *Proc*. of *the 3rd International Conference for Learning Representations*, 2015.

Copyright ⓒ 2023 KSII

Convergence of Blockchain and Artificial Intelligence: New Opportunities for Secure and Reliable Systems

Nakhoon Choi and Heeyoul Kim*
Department of Computer Science, Kyonggi University,
Gyeonggi-do, South Korea
[e-mail: skg4463@gmail.com, heeyoul.kim@kyonggi.ac.kr]
*Corresponding author: Heeyoul Kim

Abstract

Blockchain technology has recently gained attention for its ability to provide a secure and decentralized system. Blockchains provide consensus between untrusted parties, while enabling the integrity and persistent storage of all data. At the same time, artificial intelligence (AI) has emerged as a powerful tool for analyzing and predicting large amounts of data. It is said that combining these two technologies can build a more efficient and reliable system, and there is great potential for improvement in various aspects such as data privacy, security, and accuracy. This paper explores the possibility of fusing blockchain and AI technologies to achieve better technological advancements. It first introduces the basic concepts of blockchain and AI and discusses how they can be combined. It then explores several cases where blockchain and AI technologies can be applied and identifies some of the challenges that may arise from their convergence, such as scalability and interoperability. In conclusion, the paper suggests that the convergence of blockchain and AI can revolutionize the way we interact with data and build more secure and reliable systems. The paper also suggests the development of new business models and services that the convergence of blockchain and AI can provide. This technological convergence is expected to provide new opportunities for businesses, governments, and individuals alike to improve processes, increase efficiency, reduce costs, and enhance data management and reliability.

Keywords: Blockchain, Artificial Intelligence

1. Introduction

In the modern world, blockchain and artificial intelligence technologies are becoming innovative technologies. Artificial intelligence is used in various fields based on its prediction and data analysis capabilities, while blockchain technology is used to implement a secure and reliable system based on a distributed system. Blockchain and AI are expected to provide new opportunities through technological convergence. Large-scale data processing on various platforms is increasingly occurring, and reliability and personal information security issues are arising as a result. The technical convergence of blockchain and artificial intelligence can provide effective solutions in systems where security and reliability are important.

This paper aims to explore the technical convergence of blockchain and AI and its potential applications. After providing a technical overview of blockchain and AI, we explore how each technology can be converged to provide new opportunities. To this end, we categorize recent research into two groups according to the nature of the convergence and identify research trends. Finally, we discuss the new opportunities and advantages of the convergence of blockchain and AI, as well as future research directions. In doing so, we identify the potential for developing secure and reliable systems.

2. Overview of Blockchain and AI Technology

2.1 Blockchain

Blockchain is part of a family of distributed ledger technologies, starting with Bitcoin, which was unveiled by Satoshi Nakamoto in 2009.

This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research]

Blockchain is a decentralized network protocol that elects and records a common overall opinion through consensus between mutually untrusted participants in a distributed network context. This process is defined by a consensus algorithm, which is categorized as Proof of Work (PoW), Proof of Stake (PoS), etc. depending on the nature of the consensus. A blockchain's distributed record repository, the Ledger, is created by a consensus-elected block producer and linked to the previous block, and is owned by all participants in duplicate. Through the process of record comparison and verification between participants, the integrity of all information registered on the blockchain can be verified. In addition, the connectivity between blocks and the irreversibility of the hash algorithm ensures the irreversibility of previous blocks, ensuring permanence and transparency. All participants in the blockchain can perform privacy protection by disclosing only the blockchain account created based on public key cryptography.

2.2 Artificial Intelligence

Artificial intelligence is a technology designed to mimic the learning and decision-making abilities of humans through computer programs. Machine learning, a classification of AI, aims to find specific patterns in data to create predictive models. Furthermore, deep learning uses artificial neural networks to extract features from large amounts of data and create graphs of patterns that cannot be represented by ordinary mathematical expressions. In doing so, AI mimics human decision-making abilities to perform quickly and at scale. These deep learning techniques are used in a variety of fields, including image processing, natural language processing, speech recognition, and finance. While AI is capable of processing data from large systems, it has raised controversy over security and privacy, transparency, and fairness. The data of the learning process, as well as the integrity of the process and the transparency of the prediction entity, are being reconsidered.

3. Cases and Applications of Convergence of Blockchain and AI Technology

We present recent research that explores new value and opportunities through the technological convergence of blockchain and artificial intelligence. In doing so, we identify complementary convergence technologies and potential applications that blockchain and AI can create.

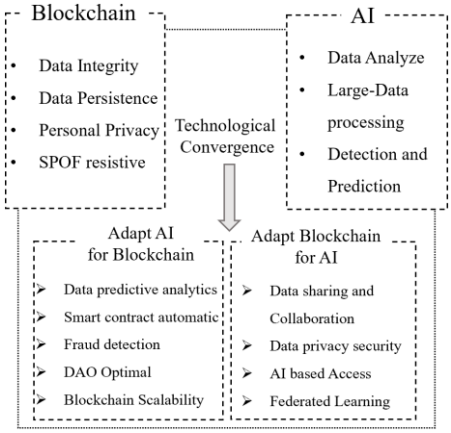


Fig. 1. Technological Convergence Specification

3.1 Adapt AI for Blockchain

This section identifies the role and potential of artificial intelligence as a technical complement to blockchain. A blockchain is a large-scale data platform that records all transactions and liquid data. As with any system, large-scale transactions on a blockchain can involve fraudulent activity by malicious users, and Zibin Zheng [1] studied anomaly detection using deep learning, data mining, and data visualization to identify transaction risks and identify vulnerabilities in smart contracts. M Deepa [2] also proposes a solution to prevent fake transaction detection through anomaly detection. BD Deebak [3] proposed an analysis method based on XGBoost to optimize network load by analyzing transactions and shares occurring on the blockchain. Among the studies using artificial intelligence in terms of blockchain protocols, T Wang [4] proposes a strategy to derive an optimization method for block mining through consensus in blockchain through reinforcement learning. Attempts to use AI in the contract and application layers have also been proposed. W Xiao [5] proposes a platform that supports the establishment of transaction contracts through a model that predicts the likelihood of resource delivery in IoT-based blockchain systems.

As a technical complement to blockchain, AI is being used for smart contract automation, transaction anomaly detection, protocol predictive analysis, DAO optimization, and protocol scalability. Highlighting the

characteristics of a large-scale data platform and managing the data produced, AI is highly suitable for blockchain expansion and new opportunities.

3.2 Adapt Blockchain for AI

This section describes examples of how blockchain has been used to complement AI technologies and how it can be applied. In AI, it is common to train a specific model from large-scale data, and this process requires complementing data and training integrity and model transparency. V Mugunthan [6] proposes an incentive system to support distributed learning without centralized datasets and auditors in the learning process. Furthermore, K Toyoda [7] and Dun Li [8] propose a platform to support federated learning via blockchain. Wei Liang [9] presents a detection model for real-time attack detection based on collaborative clustering in large-scale data systems such as blockchain.

AI is gradually developing models with high accuracy. However, there are problems with the integrity and reliability of training data and the learning process, and AI models provided as a service are not guaranteed to be reliable and transparent. In addition, there are problems with learning data, such as data containing personal information that is difficult to share. Blockchain can work as a suitable solution to the above problems and can be used to complement the reliability and integrity of AI. As a platform for learning sensitive information, blockchain can provide a learning process that ensures the integrity and transparency of learning information without data sharing, and can be used as a transparent incentive system for members. As a technical complement to AI, blockchain can be used for data sharing and collaboration, data privacy and security, AI-based identity assurance, distributed learning, etc. It will make a great contribution to accelerating the market development and adoption of AI technology.

4. Conclusions

This paper explored the technological convergence, applications, and new opportunities of blockchain and artificial intelligence. With the increasingly large-scale data on the platform, reliability and privacy are becoming increasingly important. The convergence of blockchain and AI will provide important opportunities to ensure the security and reliability of systems that deal with large amounts of data. Each technology has the ability to grow as a complementary technology, and we believe that through technological convergence, they will become the foundation of mega-scale systems.

To extend the literature on the technical convergence of blockchain and AI, we would like to add a fine-grained categorization and precise comparison of convergence application research cases. Through this, we hope to propose and explore the future direction of blockchain and AI research.

References

[1] Zheng, Zibin, Hong-Ning Dai, and Jiajing Wu. "Blockchain intelligence: When blockchain meets artificial intelligence," *arXiv preprint arXiv:1912.06485*, 2019).

[2] Deepa, M., and D. Akila. "Cost-effective anomaly detection for blockchain transactions using unsupervised learning," in *Proc. of Intelligent Computing and Innovation on Data Science ICTIDS*, 2021.

[3] Deebak, Bakkiam David, and AL-Turjman Fadi. "Privacy-preserving in smart contracts using blockchain and artificial intelligence for cyber risk measurements," *Journal of Information Security and Applications*, 58, 102749, 2021.

[4] Wang, Taotao, Soung Chang Liew, and Shengli Zhang. "When blockchain meets AI: Optimal mining strategy achieved by machine learning," *International Journal of Intelligent Systems*, 36.5, 2183-2207, 2021.

[5] Xiao, Wenjing, et al. "Blockchain for secure-GaS: Blockchain-powered secure natural gas IoT system with AI-enabled gas prediction and transaction in smart city," *IEEE Internet of Things Journal*, 8.8, 6305-6312, 2020.

[6] V. Mugunthan, R. Rahman, and L. Kagal. "Blockflow: An accountable and privacy -preserving solution for federated learning," *arXiv preprint arXiv:2007.03856*, 2020.

[7] Toyoda, Kentaroh, et al. "Blockchain-enabled federated learning with mechanism design," *Ieee Access*, 8, 219744-219756, 2020.

[8] Li, Dun, et al. "Blockchain for federated learning toward secure distributed machine learning systems: a systemic survey," *Soft Computing* 26.9, 4423-4440, 2022.

[9] Liang, Wei, et al. "Data fusion approach for collaborative anomaly intrusion detection in blockchain-based systems," *IEEE Internet of Things Journal*, 9.16, 14741-14751. 2021.

Copyright ⓒ 2023 KSII

Design of Insider Technology Leakage Sign Management System using Blockchain

Giwan Hong[1], Yuna Han[1], Woojoo Kwon[1] and Hangbae Chang[2,*]
[1]Department of Security Convergence, Graduate School, Chung-Ang University
Seoul, South Korea
[e-mail: ghd9201@cau.ac.kr, juliahan094@cau.ac.kr, tnswo2425@cau.ac.kr]
[2]Department of Industrial Security, Chung-Ang University
Seoul, South Korea
[e-mail: hbchang@cau.ac.kr]
*Corresponding author: Hangbae Chang

Abstract

The incidents of technology leaks instigated by insiders, including both present and past employees as well as collaborating companies, is increasing. These leaks are becoming more sophisticated and difficult to prevent with traditional methods. Security solutions for leak detection, largely based on abnormal behavior-based detection, are insufficient due to the high incidence of false positives and negatives Current research in the field of insider technology leak detection frequently resorts to synthetic datasets which are largely confined to specific actions or scenarios. In this study, we propose a leakage sign management system focused on leakage signs derived from actual technology leak activities. The proposed system aims to identify and detect leakage signs that can be collected from information systems and is based on opensource security software and blockchain. We expect that the dataset generated from the system is expected to serve as basic data for research of technology leak detection. The proposed system is also anticipated to provide a foundation for establishing reliable industrial technology protection environments in various organizations and corporations.

Keywords: Technology Leakage, Insider Threat, Open Source Security Software, Blockchain

1. Introduction

The number of technology leakage incidents caused by authorized insiders, such as current and former employees or partnering companies, is on the rise. These incidents are becoming more sophisticated and intelligent, making it difficult to handle with traditional method. Security solutions for preventing technology leaks by insiders largely hinge on post-response methods or anomaly detection methods based on abnormal behavior. Yet, these solutions are known to generate high rates of false positives and negatives. Notably, a majority of studies for detecting insider technology leak activities have largely relied on leak threat datasets. These datasets, however, are largely confined to specific actions or scenarios and pose a significant challenge when applied to real-world work environments.

In this paper, we present a leakage sign management system focused on leakage signs derived from actual technology leak activities. we design the leakage sign management system based on open-source security software and blockchain to identify and detect leakage signs that can be collected from information systems. We expect that the data generated from our proposed system could lay a solid groundwork for further research into the detection of technology leaks. Also, The proposed system based on open-source security software and

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01799) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation) and also supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0008703, The Competency Development Program for Industry Specialist)

blockchain is anticipated to serve as basic data for constructing industrial technology protection environments with reliability in various companies and organizations.

2. Related Works

2.1 Research and security solution trends for insider technology leakage

Internal leakage has increasingly been regarded as a vital factor in determining an organization's competitive edge and survival, thus gaining prominence as a crucial issue within the security field [1, 2]. However, existing enterprise security solutions have shown limitations in addressing this issue, particularly with high false positive and negative rates that stem from policy-based approaches that focus on espionage identification through the timing of leaks and anomaly detection that fails to account for specific industry business features [2].

In order to solve these problems, various studies have been conducted to investigate insider leakage behavior [1, 2]. For instance, foundational research using focus group interviews with security experts has been carried out to understand leakage behavior [2]. In addition, the risk associated with each behavior has been examined through n-gram scenarios and risk degree clustering [1, 2].

2.2 Insider threat datasets

Several studies on insider threat behavior have utilized synthetic data generated by a virtual office environment due to the lack of industrial datasets [3]. For example, the CERT r5.2 dataset has been adopted in many research based on insider leakage scenarios [2]. In addition, APEX, a synthetic log-based dataset containing 5 out of 13 malicious users, and RUU, a dataset with 14 out of 34 masqueraders, have been used for insider leakage analysis and detection [4]. However, synthetic datasets may be challenging to apply across industries with diverse variables in real-world scenarios [3]. Consequently, the significance of using actual datasets in research has grown. In recent years, the Enron dataset has been utilized for traitor-based analysis and classification [5]. Therefore, in this study, we focus on constructing a dataset and designing its system by referencing the actual insider leakage symptoms from previous studies.

2.3 Open Source Security Software

In the current era of advancing computerization and informatization, most organizational tasks are carried out through information systems, which typically operate in the form of personal computers (PCs). In particular, Windows systems, which are predominantly used by individuals and organizations, generate a vast array of log data, including but not limited to, system logs, application logs, network logs, and device logs [6, 7]. Open Source Security Software such as Wazuh, osquery, and sysmon have the capability to amass a wide range of data related to systems, event logs, processes, users, and file systems [8, 9, 10].

Open-source software, recognized for their openness and adaptability, offers a safety net consistently verified by a global network of developers and security experts who routinely improve the software to meet emerging security threats [11]. By utilizing open-source software, organizations can mitigate challenges such as high costs and limited expertise encountered when implementing comprehensive security systems, thereby reducing security costs and benefiting from a solution that is continually improved and verified [12]. Open Source Security Software such as Wazuh, osquery, and sysmon are well-maintained and updated continually by active communities for new functionalities, bug fixes, and adjustments to improve performance or reliability as an open source projects [8, 9, 10]. Accordingly, we design a leakage sign management system using Wazuh and other open source security software which are largely used for security monitoring system.

2.4 Trends of Blockchain in Security Sector

Security solutions, including Data Loss Prevention (DLP), Digital Rights Management (DRM), and Security Information and Event Management (SIEM), are commonly employed across various enterprises and organizations. However, these centralized security systems have certain shortcomings such as vulnerability to a single point of failure and susceptibility to sophisticated insider threats that can manipulate log data [13, 14].

Leveraging distributed ledger technology, blockchain offers advantages like data integrity, transparency, and scalability, which are superior to those of existing centralized security systems. Studies on blockchain-based log and data management have been conducted to leverage its advantages, thereby enhancing security control and accountability within organizations [15, 16, 17]. In this paper, we propose a leakage sign management system based on blockchain technology for insider threat accountability and data integrity and transparency, thereby overcoming the limitations of traditional centralized security data mechanisms.

3. Design of Insider Technology Leakage Sign Management System using Blockchain

3.1 Insider Technology Leakage Sign

Electronic leakage signs derived from the in-depth interview in the technical leakage scenario collected through the Focus Group Interview for security experts who have experienced technical leakage incidents are shown in **Table 1** [2]. Each sign of leakage can be detected through a variety of data, including web history, email, jumplists, prefatches, and registry generated from the information system.

Table 1. Detection Methods of Leakage Signs [2]

Technology Leakage Signs	Data for Detection
Use of unauthorized external (commercial) e-mail	Web History
Send e-mails from other (internal) employee computers	Web History, Email
Copy key document using USB memory	Jumplist
Excessive reading of key documents	Prefetch, Registry, Link File
Installation of SW (Key Logger) that records computer usage logs	Registry
Attempting to access key documents with multiple accounts (IDs) on a specific computer (few)	Application and System Log
Attempt to access key document outside of business hours (including holidays)	Prefetch, Registry, Link File
Leak key document through access to unauthorized sites	Web History

Thus, we design the technical leakage sign management system for the electronic detection of leakage signs, as outlined in **Table 1**.

3.2 Design of Insider Technology Leakage Sign Management System

We identify leakage sign data that can be collected on information systems based on actual technology leakage activities, and propose a leakage sign management system using blockchain, as shown in **Fig. 1**.

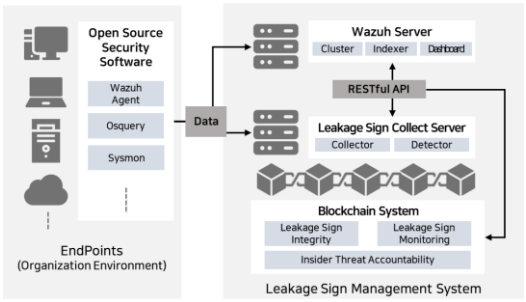


Fig. 1. The Overview of Insider Technology Leakage Sign Management

The proposed system can be composed of an information system including PCs, laptops and Servers and Cloud Systems in an organizational environment, and a leakage sign management system including Wazuh server, leakage sign collection server, and blockchain system. The endpoint information system uses the agent of open source security software to collect data from the information system, and most of it is sent to the Wazuh server, and some of it can be sent to the leakage sign collection server. The leakage sign collection server collects electronic data through the RESTful API provided by the Wazuh server, analyzes the electronic data, and detects technology leakage signs. The detected technology leakage signs are sent to the blockchain system, and the blockchain system supports data integrity management, leakage signs monitoring, and insider threat accountability.

4. Conclusions

Internal technology leakage incidents pose significant economic threats to businesses and organizations, undermine their industrial competitiveness, and can even jeopardize the economic security of a country. In this paper, we

aimed to identify leakage signs based on actual technology leakage activities and design the system to detect real-world signs of technology leakage that may occur within companies and organizations.

We anticipate that the proposed system which collects electronic information from information systems and is based on real-world technology leakage incidents can be used as a basic data for data collection in technology leakage research. Furthermore, we envisage the proposed system based on open source security software to serve as a foundational resource for establishing industrial technology protection environments in various companies and organizations. As part of our future research, we aim to make a dataset, derived from our proposed system, for technology leakage detection research, and verify feasibility and applicability of the dataset.

References

[1] J. Kim, J. Kim, and H. Chang, "Research on behavior-based data leakage incidents for the sustainable growth of an organization," *Sustainability*, Vol. 12, No. 15, pp. 6217, 2020.

[2] J. Kim and H. Chang, "A Development of Technology Leakage Sign Evaluation System by Insider Threats," *Doctoral Dissertation, Chung-Ang University Graduate School*, Seoul, Korea, 2023.

[3] W. Yoon and H. Chang, "Design and Implementation of Blockchain System for Sharing Insider Threat of University Labs," *Master's Thesis, Chung-Ang University Graduate School*, Seoul, Korea, 2023.

[4] M. N. Al-Mhiqani et al., "A Review of Insider Threat Detection: Classification, Machine Learning Techniques, Datasets, Open Challenges, and Recommendations," *Applied Sciences*, vol. 10, no. 15. MDPI AG, p. 5208, 28-Jul-2020.

[5] F. Janjua, A. Masood, H. Abbas, I. Rashid, and M. M. Z. M. Khan, "Textual analysis of traitor-based dataset through semi supervised machine learning," *Future Generation Computer Systems*, Vol. 125, pp. 652-660, 2021.

[6] GlobalStat statcounter, "Desktop Operating System Market Share Worldwide" https://gs.statcounter.com/os-market-share/desktop

[7] Song Joong-ho, Kim Hak-min, and Yunjin. "a window - based PC log collection system using open source." practical paper on information science society computing 22.7 (2016): 332-337.

[8] Wazuh, https://wazuh.com/

[9] Osquery, https://www.osquery.io/schema/5.8.2/

[10] Sysmon, https://learn.microsoft.com/ko-kr/sysinternals/downloads/sysmon

[11] R. Kitchin, "The Data Revolution: Big Data, Open Data, Data Infrastructures & Their Consequences." *SAGE Publications Ltd*, 2014.

[12] S. Chauhan, M. Jaiswal, S. Rai, L. Motiwalla, and L. Pipino, "Determinants of adoption for open-source office applications: A plural investigation," *Information Systems Management*, vol. 35, no. 2. Informa UK Limited, pp. 80–97, Mar-2018.

[13] F. Zafar et al., "Trustworthy data: A survey, taxonomy and future trends of secure provenance schemes," *Journal of Network and Computer Applications*, vol. 94. Elsevier BV, pp. 50–68, Sep-2017.

[14] X. Li, P. Jiang, T. Chen, X. Luo, and Q. Wen, "A survey on the security of blockchain systems," *Future Generation Computer Systems*, vol. 107. Elsevier BV, pp. 841–853, Jun-2020.

[15] A. Ali, A. Khan, M. Ahmed, and G. Jeon, "BCALS: Blockchain- based secure log management system for cloud computing," *Transactions on Emerging Telecommunications Technologies*, vol. 33, no. 4. Wiley, 17-Apr-2021.

[16] M. S. Ozdayi, M. Kantarcioglu, and B. Malin, "Leveraging blockchain for immutable logging and querying across multiple sites," *BMC Medical Genomics*, vol. 13, no. S7. Springer Science and Business Media LLC, Jul-2020.

[17] T. Guimarães, R. Duarte, B. Pinheiro, D. Faria, P. Gomes, and M. F. Santos, "Blockchain Analytics - Real-time Log Management in Healthcare," *Procedia Computer Science*, vol. 201. Elsevier BV, pp. 702–707, 2022.

Copyright ⓒ 2023 KSII

An Approach to Detect SMS Spam Using ChatGPT

Sooyon Seo, Jaeho Hwang and Moohong Min*
Department of Global Business Administration, Sungkyunkwan University, South Korea
Department of Computer Science and Industrial System Engineering, Yonsei University, South Korea
School of Convergence/Convergence Program for Social Innovation, Sungkyunkwan University, South Korea
[e-mail: sooyon1119@g.skku.edu, ashiereki@yonsei.ac.kr, iceo@skku.edu]
*Corresponding author: Moohong Min

Abstract

The abundance of spam text messages received through SMS poses a persistent challenge in our daily lives. To combat this issue, researchers have explored various machine learning techniques for SMS spam filtering. This paper proposes an approach that leverages ChatGPT, an AI chatbot based on the GPT architecture developed by OpenAI. ChatGPT is known for its language generation capabilities and has gained popularity for its accurate responses across diverse domains. The research team aims to assess ChatGPT's ability to differentiate between spam and legitimate SMS messages using both Korean and English spam datasets. This study contributes to the ongoing efforts in improving spam detection and enhancing the effectiveness of SMS spam filtering mechanisms.

Keywords: Recommender Systems, SMS Spam, ChatGPT

1. Introduction

Spam messages pose risks and persist across communication channels, including SMS, despite regulations. ChatGPT, an advanced AI language model, generates human-like responses and has been widely used in diverse domains. This paper presents a novel SMS spam detection system using ChatGPT, evaluated on a large Korean and English message dataset [1]. Results show the system accurately identifies SMS spam, addressing the challenge and highlighting ChatGPT's efficacy in this domain.

2. Related Works

Researchers have been making efforts to classify SMS spam messages using machine learning techniques [2-4]. Roy et al. achieved high accuracy in English SMS spam filtering but focused only on English messages [2]. Liu et al. proposed a spam Transformer model and emphasized the importance of testing larger datasets and advanced models [3]. Ghourabi et al. introduced a hybrid CNN-LSTM model for Arabic and English SMS spam classification, outperforming other techniques [4].

This paper utilizes ChatGPT to develop a robust SMS spam detection system. It evaluates the system on a dataset of recent Korean and English SMS spam messages, advancing SMS spam detection techniques.

3. Methodology and Experiments

In this research, Korean and English SMS messages were tested using data provided by the Korea Internet & Security Agency (KISA) and the UCI Machine Learning Depository [5].

The experiments in this study were conducted using the OpenAI platform's API. The response parameters were configured as follows:

- model: “gpt-3.5-turbo”
- tempreture: “0.9”
- max_tokens: “2048”
- top_p: “1”
- frequency_penalty: “0”
- presence_penalty: “0.6”

This work was supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. RS-2022-00166729).

The 'model' parameter selects the underlying model for generating responses, while 'temperature' controls the output's randomness. 'max_tokens' limits the response length, and 'top_p' (previously 'nucleus') sampling focuses on the most probable tokens. Lastly, 'frequency_penalty' is to discourage repetitive phrases and 'presence_penalty' to encourage topic adherence [6].

The research team asked various forms of questions to determine if a message was spam or ham, including distinguishing scamming spam from regular advertising spam. The team also examined ChatGPT's consistency in generating the same decision for the same SMS spam message by asking repeated questions in new chats to avoid contextual errors.

In order to minimize errors caused by the context, both input Korean inquiries and English inquiries had the same meaning. To be precise, the team used ChatGPT for the translation of those questions. In addition, it was tested whether ChatGPT can discriminate spam messages that are trying to scam from those that are not. Questions written on the **Table 1** are the examples.

Korean	English
이것은 스팸문자일까?	*Is this a spam message?*
이것은 스팸일까 햄일까?	*Is this a spam or ham?*
이것은 사기 스팸 문자일까? 단순 광고성 문자일까?	*Is this a scamming spam message or just a regular advertisement message?*

Table 1. Inquiry Examples

Additionally, in order to examine whether ChatGPT generates the same decision for the same SMS spam message, the research team has inquired the same question many times and observed the resemblances and differences. Since ChatGPT remembers what the user said earlier in the conversation, each question was asked in a new chat.

4. Result

It was found that ChatGPT is capable of effectively distinguishing SMS spam messages. The research team tested 50 randomly selected Korean and English SMS messages and observed that ChatGPT could not only classify them accurately but also provide reasoning for its decisions. Results are as below.

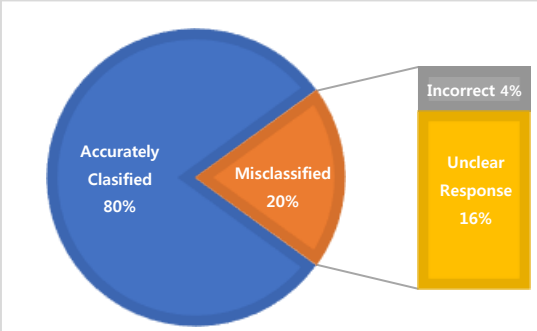


Fig. 1. The Result of Korean Spam Message Classification

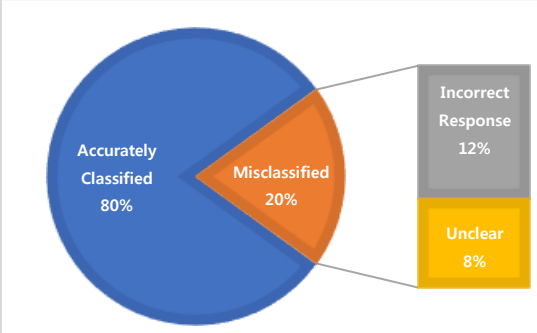


Fig. 2. The Result of English Spam Message Classification

However, despite its generally high accuracy, there were instances where ChatGPT failed to recognize or misclassified SMS spam messages. Nevertheless, it still acknowledged the possibility of the message being spam or inappropriate and provided explanations for its reasoning. Moreover, it was discovered that ChatGPT occasionally generated different judgments for the same SMS, particularly for those initially misclassified. When the same SMS was inquired again, ChatGPT produced varying decisions over time.

5. Discussion

5.1 Implications

The research team discovered that ChatGPT demonstrates a high level of accuracy in recognizing SMS spam messages, making it a potential tool for SMS spam detection. It

provides logical explanations for its decisions.

However, there were cases where ChatGPT misclassified or couldn't classify SMS messages. Despite this, it still advised users to be cautious and consider their next actions. The depth of the explanation varied based on the input messages. Some messages received detailed consideration, with ChatGPT summarizing the text and highlighting suspicious parts. In contrast, for other messages, ChatGPT compared textual features with typical characteristics of SMS spam messages. This feature can be utilized in constructing a framework for SMS spam detection, grouping received spam messages based on the specific words generated by ChatGPT and evaluating the reliability of its replies to provide accurate information about the SMS message.

5.2 Limitations and Potential for Further Research

First, this paper specifically tested Korean and English SMS spam data and could not assess the general applicability to other languages due to limited data. However, ChatGPT's accessibility with multiple languages suggests the potential for SMS spam detection in other languages, although accuracy may vary.

Second, SMS spam messages evolve over time, particularly those with scamming purposes. The research team observed longer and more detailed spam messages in the recent data used. It raises questions about ChatGPT's ability to identify new forms of SMS spam in the future, considering the ever-changing trends. ChatGPT's knowledge is also limited to events and information after 2021.

Lastly, the research focused solely on detecting SMS spam messages and did not test other types such as email spam or social networking service (SNS) spam.

6. Conclusions

This paper focuses on utilizing ChatGPT to filter SMS spam messages, specifically in Korean and English. The research team tested ChatGPT's ability to distinguish between spam and ham by posing various questions. The findings reveal that ChatGPT can identify SMS spam messages, although not always with complete accuracy. It provides valid reasons to support its decisions, but the depth of explanation varies depending on the input message. By categorizing SMS messages based on specific phrases generated by ChatGPT, a framework for assessing message reliability can be developed. However, there are limitations to consider. The research only tested Korean and English SMS spam data, raising uncertainty about its applicability to other languages. ChatGPT may struggle to identify new forms of SMS spam as they evolve over time. Additionally, the research solely focuses on SMS spam and does not cover other types of spam like email or SNS spam.

This research highlights the potential of ChatGPT in detecting SMS spam messages and lays the groundwork for further exploration in the field of spam detection using natural language processing techniques. It could inspire the development of automated systems for filtering unwanted SMS messages and encourage investigations into identifying and filtering other types of spam messages. Ultimately, this research contributes to advancing more efficient and robust spam detection systems.

References

[1] A. Hughes, “ChatGPT: Everything you need to know about OpenAI's GPT-4 tool,” BBC Science Focus Magazine, https://www.sciencefocus.com/future-technology/gpt-3/.

[2] P. K. Roy, J. P. Singh, and S. Banerjee, “Deep learning to filter SMS SPAM,” Future Generation Computer Systems, Vol 102, pp. 524–533, 2019.

[3] X. Liu, H. Lu, and A. Nayak, “A spam transformer model for SMS spam detection,” IEEE Access, Vol. 9, pp. 80253–80263, 2021.

[4] A. Ghourabi, M. A. Mahmood, Q. M. Alzubi, “A hybrid CNN-LSTM model for SMS spam detection in Arabic and English messages,” Future Internet, Vol. 12, No. 9, pp. 156, 2020.

[5] Center for Machine Learning and Intelligent Systems, “SMS Spam Collection Data Set - UCI ML Repository,” https://archive.ics.uci.edu/ml/datasets/SMS+Spam+Collection, 2012.

[6] OpenAI, “API Reference – Completions – Create completion,” https://platform.openai.com/docs/api-reference/completions/create.

Copyright ⓒ 2023 KSII

Nurturing Personal Information Protection Specialists: An Integrated Approach to Curriculum Development

Yeji Choi, Juno Lee, and Hangbae Chang
Department of Security Convergence, Chung-Ang University
Seoul, South Korea
[e-mail: yjji9783@cau.ac.kr, iamjuno95@cau.ac.kr]
Department of Industrial Security, Chung-Ang University
Seoul, South Korea
e-mail: hbchang@cau.ac.kr
*Corresponding author: Hangbae Chang

Abstract

As the Fourth Industrial Revolution unfolds, the transition into a data-centric economy underlines the increasing importance of data in modern society. A significant portion of the data utilized today is derived from personal information. Particularly emphasized is the secure utilization of sensitive information, which holds potential for innovative value. Consequently, demand grows for balanced personal information protection specialists capable of managing both breach responses and safe data usage. However, there is a global scarcity of departments or curricula focusing exclusively on 'personal information protection' to address these issues. Therefore, this study examines the trends in personal information protection-related courses in the curricula of universities in the United States, Japan, and Korea, and identifies essential courses commonly offered in personal information protection. Based on an analysis of industry demand, this study identified four practical competencies required. Ultimately, a curriculum for nurturing personal information protection specialists is developed by integrating these essential courses and practical competencies. This research is significant as it lays the foundation for an academic field aimed at nurturing personal information protection specialists and provides direction for a curriculum aimed at cultivating personal information protection capabilities.

Keywords: Personal Information Protection Specialists, curriculum development, practical competency

1. Introduction

The Fourth Industrial Revolution has ushered in a data-centric era where core technologies like AI and blockchain rely heavily on data. A significant proportion of this data comprises personal information, particularly sensitive information such as medical records. Given the potential value and risks associated with this information, the concept of personal information protection has broadened beyond preventing breaches to also include safe usage. As a result, there's an increasing demand for specialists in personal information protection. However, a workforce mismatch persists in this field due to a significant supply-demand gap.

Currently, there are no universities offering a department solely dedicated to personal information protection. While some departments include related courses, they typically combine elements of information security and personal information protection. Though related, these disciplines have different focuses[1]; information security aims to defend against system, network, and data threats, focusing on technical areas such as encryption and network

This paper was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE) (P0008703, The Competency Development Program for Industry Specialist)

security[2]. In contrast, personal information protection is grounded on laws and regulations, aiming to safeguard individual-related information and covering areas like personal information protection laws, technologies, and management systems[3]. Given these differences, focusing only on information security measures is insufficient for personal information protection's objectives. A dedicated system for nurturing personnel in personal information protection is necessary to lay the groundwork for a data-centric society.

In this study, a tailored curriculum for nurturing specialists in personal information protection is aimed to be established. Trends in related courses from universities in the United States, Japan, and Korea are analyzed, and the core courses most frequently offered are identified. Alongside an analysis of industry demand, four key practical competencies are determined. These findings are then integrated to develop a specialized curriculum.

2. Terminology

2.1. Personal Information

"Personal Data"[4] pertains to any information associated with an identified or identifiable living individual. This term is commonly used in the EU under the GDPR. In contrast, 'Personal Information,' a term originating from OECD privacy principles, is predominantly used in countries like South Korea, Canada, and Australia. Both terms encompass a broad range of information related to individuals, with this study primarily using “Personal Information.”

2.2. Personal Information Protection(PIP)

Various terms are used to denote activities associated with the protection of personal information, such as information privacy, data privacy, data protection, and personal information protection. These terms are used interchangeably depending on the user, context, and conceptual differences, but they generally include elements such as guaranteeing individual rights, implementing safety measures, and considering legal aspects. South Korean government agencies primarily use “personal information protection (or PIP)”, a term this study adopts."

3. Analysis of Global Trends

3.1 United States

Carnegie Mellon University and the University of Texas both offer tailored master's programs focusing on the integration of information security and personal information protection. Carnegie Mellon's program focuses on technology-oriented subjects, while the University of Texas includes a specialized segment on 'Information Security and Privacy Courses' within their graduate program.

3.2 Japan

Both the University of Tokyo's Department of Information Science and Waseda University's Department of Communications and Computer Engineering offer curricula emphasizing technical protection courses. They commonly cover topics such as computer network security, cryptography, and personal information-related subjects, thus highlighting their focus on safeguarding data and personal information.

3.3 South Korea

Among 11 Korean universities, Konyang University's Department of Cybersecurity offers the most comprehensive personal information protection courses. This curriculum includes personal information encryption technology, safety measures, impact assessment, next-generation protection technology, and management systems for information protection. Korea University, Busan University of Foreign Studies, and Hanseo University each provide a lecture on personal information protection.

4. Practical Competencies

This study derived practical competencies in personal information protection from the analysis of various educational programs and job postings within South Korea. These competencies fall into four main areas: Rights Protection Technology, Safe Utilization Technology, Exposure Minimization Technology, and Legal-Regulatory Competency. Rights Protection Technology and Safe Utilization Technology focus on securing personal information rights and using personal

information safely, respectively, through technological means. Exposure Minimization Technology concentrates on reducing unnecessary exposure of personal information and overall information security. Lastly, Legal-Regulatory Competency encompasses the comprehension and adherence to the legal and regulatory aspects of personal information protection.

5. Development of a Specialized Curriculum for PIP

From the analysis of personal information-related curricula at universities worldwide, essential courses for nurturing specialists in this field were identified. Along with four practical competencies in technology, management, and law/regulations that were identified, a specialized curriculum for personal information protection was developed. This curriculum includes 14 distinct subjects, as shown in **Fig. 1**.

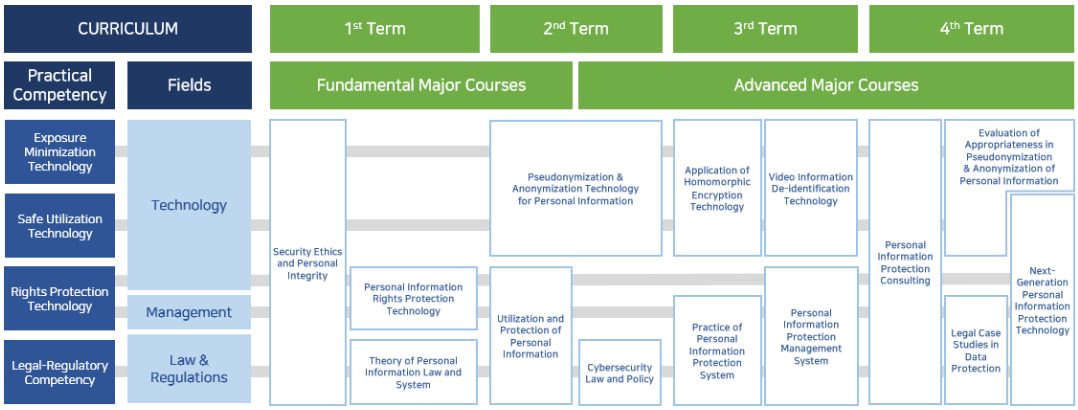


Fig. 1. A Specialized Curriculum for Personal Information Protection(PIP)

6. Conclusions

In this study, a curriculum has been developed to nurture specialists in the field of personal information protection. This process involved a comprehensive analysis of not only related curricula from the United States, Japan, and South Korea, but also various educational programs and job postings within South Korea. This analysis led to the identification of common courses and practical competencies that align with industry demands. The findings of this study highlight the importance of personal information protection and pave the way for future curriculum development, thereby expected to enhance the credibility of technologies that use personal information. This curriculum balances theoretical knowledge and practical application, and its effectiveness is set to be evaluated in future research post-implementation.

References

[1] P. Jain, M. Gyanchandani, and N. Khare, "Big data privacy: a technological perspective and review," *Journal of Big Data*, Vol 3, pp. 1-25, 2016.

[2] W. Lim and S. Ahn, "A Study on Improvements of the Information Security epartment via the Curriculum Analysis," *The Journal of Korean Association of Computer Education*, Vol 17, No. 6, pp. 71-80, 2014.

[3] E. Bertino, "Data Security and Privacy in the IoT," *EDBT*, Vol 2016, pp. 1-3, 2016.

[4] EU Commission, "What is personal data?," *Reform of EU data protection rules*. Retrieved May 13, 2023, https://commission.europa.eu/law/law-topic/data-protection/reform/what-personal-data_en.

Copyright ⓒ 2023 KSII

Cybersecurity Challenges in Blockchain-based Social Media Networks: A Review

Muhammad Hasnain[1], Imran Ghani [2], David Smith [3], Akbar Hussain [4], Ali Daud [5], Sardar Usman [6], and Seung Ryul Jeong [7,*]

[1] Department of Computer Science, Lahore Leads University, Lahore, 42000, Pakistan
[2, 3] Computer and Information Sciences, Virginia Military Institute, Lexington, VA 24450, USA
[4] Department of Information Technology, Lahore Leads University, Lahore, 42000, Pakistan
[5] Abu Dhabi School of Management, Abu Dhabi, United Arab Emirates
[6] Department of Computer Science and Software Engineering, Grand Asian University Sialkot, 51040, Pakistan
[7] Graduate School of Business IT, Kookmin University, Seoul 136, Korea
[e-mail: drhasnain.it@leads.edu.pk, ghanii@vmi.edu, smithdtr@vmi.edu, drakbar.it@leads.edu.pk, alimsdb@gmail.com, Sardar.usman@gaus.edu.pk, srjeong@kookmin.ac.kr]
*Corresponding author: Seung Ryul Jeong

Abstract

Blockchain as a disruptive technology has gained great deal of attention from scholars. Blockchain technology is now widely used in several fields, such as social media networks, financial transactions, healthcare etc. However, users of social media networks are still concerned about the information security and privacy regarding blockchain technology. Literature on social media networks and blockchain technology is rapidly increasing these days. The published literature is focused on users' trust in social media networks and role of blockchain technology in enhancing security and privacy aspects of users' information. This review aims to present an overview of the existing studies on the role of blockchain technology in social media networks. The popular digital libraries such as IEEE Xplore, ScienceDirect, Springer, and Google Scholars were used to search for recent literature. This review's findings indicate that several attempts are made by researchers to secure users' information in social media networks with the support of blockchain technology. This study also identifies a few challenges to blockchain technology and suggests innovative future research directions to eliminate the users' concerns by overcoming reliability, privacy, security, scalability and trust concerns about social media networks.

Keywords: decentralized technology, privacy and security challenges, vulnerability detection.

1. Introduction

Online social Networks (OSNs) are popular among users due to their widespread use for information sharing and other daily activities. More than 4 billion users use intenet across the globe, and 3 billions out of them use social media networks [1]. Currently, all social media networks are centralized that means they use the central servers to keep the information of users. In the meantime users believe that OSNs protect their shared information. The centralized structure of these networks causes several drawbacks as data can be stolen, or tampered without the control of oweners of data [2].

Blockchain as a disruptive technology provides protection against malicious activities on OSNs. Decentralized social medial plateforms have been propoposed to overcome the privacy issue and store users' information. In this regard, Blockchain-based Online Social Networks (BOSNs) address the privacy and fake news issues by using blockchian technology [3].

Recent literature is focused on securing OSNs and information exchanged over these networks. Review articles presented the state of art on cybersecurity [4] and blockchain deployment in Internt of Things (IoT) and healthcare systems [5]. However, exisitng reviews show limitations in considering cyber security challenges faced by researchers in blockchain based social media networks. This review presents an overview of the exisitng literature and contributes to the literature by identifying several challenges of social media networks, which were not addressed by the traditional socila media (TSM) techniques.

2. Blockchain Technology Based Approaches

Several approaches have been propsoed to secure the users' confidential information. When users' information is collected and analysed the data leakage may occur and users' privacy is compropmised. To overcome this issue, researchers used differential keys to exhange data in blockchains [6-7]. Ushare approach uses permissionless and permissioned methods to validate the users' participation by a third party [8]. However, this technqique is less effective when blockchain network itself is enabled to validate the participating users.

Social media networks have widespread postings of fabricated screenshots and fake news. To overcome these issues, blockchain technology can notarize the information in a tamper-proof way [9]. However, challenges of verifying the information before notarizing persists and need to be addressed in future works.
In order to secure users' information over social media networks, many studies highlighted the key chraracteristics such as reliability, trust, which are widely examined in the literature [10-11]. On the other hand, privacy and security characteristics are discussed only in study [12].

3. Open Issues and Future Research Directions

This review identified the areas, which have not received sufficient investigation in the context of recent proposed approaches. The following section provides a brief description of these areas.

3.1 Private smart contracts

Smart contract is the primary element of a blockchain that runs networks smoothly. One of the challenges is to deploy the private smart contract that requires enough knowledge to write a program. Furthermore, once a smart contract is deployed, it cannot be modified and updated, which means any flaws in it exists forever. Vulnerability is another challenge to researchers for its detection in blockchain [6]. Bothe syntactic and sematic errors may cause vulnerabilities in smart contracts [12]. Therefore, vulnerability detection and prevention approaches can be proposed to overcome the issues in smart contracts.

3.2 Users' reaction to blockchain technology

Most users of blockchain technology are concerned on its technological aspects. People's reaction towards the usage of blockchain technology is not well studied and remains an open challenge for researchers. Unwanted consequences may arise from the privacy concerns of users [13]. Therefore, the organization using blockchain technology can be targeted and its reputation can be negatively affected.

3.3 Validation of security systems

This review article identifies the issue of validating the security systems. A system can suffer failure and may be exposed to cyber-attacks [14]. Therefore, a secure system is required to combat the cyber-attacks by adopting countermeasures. Sometimes complex scenarios require tight security for systems [15], and therefore, future researches endeavors can be undertaken to address the robustness and security issues in a comprehensive manner.

4. Conclusions

In this paper, we present an overview of the proposed approaches using blockchain technology in social media networks. Differential privacy based approaches have been widely examined in the literature. The proposed approaches extensively examined privacy and security concerns of users' information on social media networks. Moreover, this review

identified several open challenges to researchers regarding the users' information privacy over social media networks and blockchain technology. Scalability and storage issues of blockchain technology can be addressed in future works. In addition, researchers may propose methods to address the identified open issues in this study. In future works, researchers can examine the challenges of security in BCT identified in this study. Hacking issue can be tackled to demonstrate the blockchain technology users' trustworthiness.

References

[1] B. Guidi, "When blockchain meets online social networks," *Pervasive and Mobile Computing,* vol. 62, p. 101131, 2020.

[2] L. Jiang and X. Zhang, "BCOSN: A blockchain-based decentralized online social network," *IEEE Transactions on Computational Social Systems,* vol. 6, no. 6, pp. 1454-1466, 2019.

[3] M. U. Rahman, B. Guidi, and F. Baiardi, "Blockchain-based access control management for decentralized online social networks," *Journal of Parallel and Distributed Computing,* vol. 144, pp. 41-54, 2020.

[4] M. Gimenez-Aguilar, J. M. De Fuentes, L. Gonzalez-Manzano, and D. Arroyo, "Achieving cybersecurity in blockchain-based systems: A survey," *Future Generation Computer Systems,* vol. 124, pp. 91-118, 2021.

[5] E. M. Adere, "Blockchain in healthcare and IoT: A systematic literature review," *Array,* p. 100139, 2022.

[6] M. U. Hassan, M. H. Rehmani, and J. Chen, "Differential privacy in blockchain technology: A futuristic approach," *Journal of Parallel and Distributed Computing,* vol. 145, pp. 50-74, 2020.

[7] X. Dong, B. Guo, Y. Shen, X. Duan, Y. Shen, and H. Zhang, "A self-controllable and balanced data sharing model," *IEEE Access,* vol. 7, pp. 103275-103290, 2019.

[8] A. Chakravorty and C. Rong, "Ushare: user controlled social media based on blockchain," in *Proceedings of the 11th international conference on ubiquitous information management and communication*, pp. 1-6, 2017.

[9] G. Li, L. You, G. Hu, and L. Hu, "Recoverable Private Key Scheme for Consortium Blockchain Based on Verifiable Secret Sharing," *KSII Transactions on Internet and Information Systems (TIIS),* vol. 15, no. 8, pp. 2865-2878, 2021.

[10] H.-W. Kim and Y.-S. Jeong, "Secure authentication-management human-centric scheme for trusting personal resource information on mobile cloud computing with blockchain," *Human-centric Computing and Information Sciences,* vol. 8, no. 1, pp. 1-13, 2018.

[11] Y. Qi, M. S. Hossain, J. Nie, and X. Li, "Privacy-preserving blockchain-based federated learning for traffic flow prediction," *Future Generation Computer Systems,* vol. 117, pp. 328-337, 2021.

[12] A. Singh, R. M. Parizi, Q. Zhang, K.-K. R. Choo, and A. Dehghantanha, "Blockchain smart contracts formalization: Approaches and challenges to address vulnerabilities," *Computers & Security,* vol. 88, p. 101654, 2020.

[13] P. V. Ingold and M. Langer, "Resume= Resume? The effects of blockchain, social media, and classical resumes on resume fraud and applicant reactions to resumes," *Computers in Human Behavior,* vol. 114, p. 106573, 2021.

[14] N. Elisa, L. Yang, F. Chao, and Y. Cao, "A framework of blockchain-based secure and privacy-preserving E-government system," *Wireless networks,* pp. 1-11, 2018.

[15] P. Sun, "Security and privacy protection in cloud computing: Discussions and challenges," *Journal of Network and Computer Applications,* vol. 160, p. 102642, 2020.

Copyright ⓒ 2023 KSII

Proposing Controls for a Secure M&A of National Core Technology with Smart Contract

Yurim Choi and Hangbae Chang*
Department of Convergence Security, Chung-Ang University
Department of Industrial Security, College of Business and Economics, Chung-Ang University*
Seoul, South Korea
[e-mail: julie330@cau.ac.kr, hbchang@cau.ac.kr*]
*Corresponding author: Hangbae Chang

Abstract

As the importance of economic security increases, the protection of national core technologies is being paid attention such as strengthening related laws. Nevertheless, national core technology leakages continue to occur, especially in the process of mergers and acquisitions(M&A). According to this problem, we analyze security risks of technology leakage and propose needs of prerequisite to evaluate whether the smart contract is secure enough to accomplish smart contract securely as controls for successful M&A through literature reviews. Based on this, in future research, we will develop the study about detail controls and organizing checklists as prerequisite for accomplishing secure M&A with smart contract based on reliability to prevent the leakage of national core technologies.

Keywords: Mergers and Acquisitions (M&A), national core technology, technology leakage, smart contract

1. Introduction

As the competition among nations in which the economy, technology and security are coordinated deepens, the need to prevent the outflow of technology from companies to others is increasing. More attention should be paid to the protection of national core technology, which could lead to national losses if leaked. However, a total of 83 cases were exposed from 2018 to July 2022, of which about 40% involved leaks of national core technologies [1]. The risk of national core technology leakage is now changing from simple technology theft to joint ventures and M&A that directly have corporate management rights and acquire technologies at once [2]. In addition, the government announced that the national core technology protection period will be abolished in April 2023, which means the maintenance or cancellation of national core technology will be determined through annual technology re-evaluation to flexibly reflect global technological development trends [3]. However, the industry is concerned if technology is leaked due to attempts by foreign companies' M&A with the excluded technology which had been cared by national core technology. To prevent this, we analyze the types of security risks in the process of M&A of national core technology and what are essential controls for secure M&A.

2. Literature Reviews

2.1 National Core Technology Outflow

National core technologies have high technological and economic value in domestic and foreign markets or have high growth

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01799) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)

potential in related industries, and if leaked abroad, they will have a serious negative impact on national security and the development of the national economy. Currently, a total of 75 technologies in 13 technical fields, including semiconductors, displays, and robots have been designated as national core technologies. In 2019, the legal regulation for exporting national core technology was strengthened, and the system was mainly modified from reporting to being approved when conducting overseas M&A with national core technologies to block attempts to take over the technology. In addition, by expanding the scope of M&A for foreigners, the government strengthened controls to prevent M&A for the purpose of technology theft in advance.

NIS analyzed ways of extorting national core technology are broadly divided into six categories: abusing M&A, acquisition of core personnel, abusing cooperating companies, the outflow of pretending technology from joint research, etc. [1]. It could derive a negative effect such as reducing morale even though M&A is being actively pursued for the benefit of companies such as technological development and entering the global market.

2.2 Security Risk of the M&A Process

M&A, which aims to steal technology, acquires and merges not only companies but also already developed technologies to mitigate potential risks with developing technology. According to Forescout's research about the role of cybersecurity in M&A diligence and analyze about 65% of respondents answered they regretted making an M&A deal due to cybersecurity concerns [4].

UpGuard, a cybersecurity corporate, analyzes the 6 cyber risks in M&A process: Technology disruption, Dormant threats, Information technology resiliency, Data security, Lack of information, Organizational disruption [5].

Erwän Keräudy, the CEO of CybelAngel, claims that integrating different IT systems without well-defined policies and strategies can increase network security management challenges. In addition, there is a significant risk of inheriting data breach during the M&A process. Lastly, even with the best security policies and tools in place, it's likely that organizations might already have weak spots in their attack surface, potentially resulting in stolen credentials, or leaked [6]. Thus, corporate must assess a company's level of risk exposure across the external attack surface by infiltrating through vulnerable points before M&A.

2.3 Methods for Preventing Industrial Technology Leakage in M&A

According to An's study, they derive risk factors and causes of data breaches for each M&A stage and design management plan for risk factors by M&A [7]. They consider all the important data of corporates including national core technology but don't for unique characteristics of national core technology such as pre-declaration and acceptance before M&A.

In Kim's study, he discusses the advantages of utilizing smart contract in M&A, such as transparency, reliability, cost-time-efficiency, and automatizing Earn-out by adding contract codes. However, to activate transactions of smart contract, the legal system related to smart contract must be supplemented and it is necessary that prior consultation between companies that want to merge. Also, smart contract platforms should have the security function as a kill-button, which suspends all transactions in a problem [8]. For national core technologies that proceed with mergers between countries, safe and appropriate countermeasures should be prepared to deal with complex interests arising from differences in legal regulations between countries to which companies want to M&A belong.

To prevent technology leakage during the M&A process, research on methods using blockchain and smart contract is underway overseas either. Clairfield International, a multinational investment bank based in Geneva, Switzerland, posted "The potential of blockchain and smart contracts in M&A." in 2019 [9]. It presents how smart contract can be used effectively at each stage of M&A. It says smart contract can enhance process quality, reduce costs, efforts and improve deal certainty until the signing including post-transaction activities.

Most of the related literature analyzes risks in the overall process of M&A, not each step. In addition, there is little literature that deals with national core technologies. In this study, thus, we propose smart contract and the standards of evaluating it as requirements for secure M&A

with national core technology.

3. Controls for Secure M&A Using Smart Contract

We analyzed the risks of national core technology leakage in M&A. To lessen those risks and secure M&A, following factors are essentially assured: secure transaction, accountability, mutual consent etc. Smart contract can automate overall process of M&A not only if certain obligations and conditions are met but also Earn-out in post transaction which can reduce risks and serve higher convenience. Thus, we propose controls to lessen national core technology leakage risks as a condition to accomplish smart contract for secure M&A with national core technology. These following controls are to check if there is mutual countries' legislation in contract, if they declared and were accepted the technology previously and if there is function as the Kill-button, which can stop the transaction immediately when there is any problem in M&A. These controls are needed as standards to accomplish smart contract securely in M&A.

4. Conclusions and Future Research

In this study, we analyzed the leakage risks of national core technologies that may occur in M&A. These risks can be decreased with smart contract that assures integrity and confidentiality. It is expected to help companies export technology and national economic security securely by trustworthy M&A such as accountability using smart contract used to sign it transparently and fairly. However, we conducted only conceptual study this time. In future research, thus, we would like to develop the study detail controls for checklist as a prerequisite to accomplish smart contract securely in detail with national core technologies M&A in statistical way.

References

[1] Maeil news, "Over the past five years, 83 technology leaks, 40% of which are national core technologies," (Aug. 2022), https://www.m-i.kr/news/articleView.html?idxno=946248(last visited May, 15, 2023)

[2] Business newdaily, "Display Association Launches 3rd Industrial Technology Security Council... "Protecting the nation's core technology," (Apr. 2023), https://biz.newdaily.co.kr/site/data/html/2023/04/25/2023042500091.html(last visited May, 15, 2023)

[3] ChosunBiz, "The government eliminated the national core technology protection period and re-evaluated every year...Industry, Concerns about the Second Highness Crisis" (Feb. 2023), https://biz.chosun.com/policy/policy_sub/2023/02/07/YCSBZL5F5JB53CNJRQFIODDQPU/(last visited May, 15, 2023)

[4] Forescout, "Cybersecurity in merger and acquisition report," (2019) https://www.forescout.com/resources/cybersecurity-in-merger-and-acquisition-report (last visited May, 15, 2023)

[5] UpGuard, "The Role of Cybersecurity in Mergers and Acquisitions (M&A)," (May. 2023),https://www.upguard.com/blog/the-role-of-cybersecurity-in-mergers-and-acquisitions#toc-1(last visited May, 15, 2023)

[6] Global banking finance review, "The security risks of M&A : What are the biggest pitfalls," https://www.globalbankingandfinance.com/the-security-risks-of-ma-what-are-the-biggest-pitfalls/(last visited May, 15, 2023)

[7] YB. An, and HB. Chang, "A study on minimization of leakage of important information in M&A," *Journal of Society for e-Business Studies*, Vol 25, No. 1, pp. 215-228, 2020.

[8] BJ. Kim, and CY. Lee, "The Availability of Smart Contracts in M&A Transactions," *Dong-A Law Review*, Vol 84, pp. 297-324, 2019.

[9] Clairfield International, "The potential of blockchain and smart contract in M&A", (Mar. 2019), https://www.clairfield.com/en/ (last visited May, 15, 2023)

Copyright ⓒ 2023 KSII

A Study on Factors Derivation of Influencing Security R&D Programs for Economic Research Performance

Jawon Kim and Hangbae Chang
BlockChain Business Service Research Center, Chung-Ang University, Seoul, South Korea
Department of Industrial Security, Chung-Ang University, Seoul, South Korea
[e-mail: jjawon@cau.ac.kr, hbchang@cau.ac.kr]
*Corresponding author: Hangbae Chang

Abstract

South Korea's national research and development project investment scale is about 102 trillion won, 4.9% of GDP, the highest level among major OECD countries. In order to effectively manage and operate national research and development projects funded by the national treasury, national research and development projects are classified and managed according to various classification standards. However, many projects generate only fragmentary results, making it difficult to develop into a virtuous cycle structure that leads from scientific performance to economic performance to social performance. Therefore, this paper aims to analyze the influencing factors for performance creation leading to economic performance based on the technological information (research and development level, research and development character, technology life cycle) of the national research and development project. As a research method, this research is conducted based on the research results disclosed in the National Science & Technology Information Service and analyzed according to the classification of technology information of the research and development project. The results of this study can be used as basic data for future research and development project performance management plans through blockchain technology.

Keywords: Security R&D Programs, Research result, Advanced research result

1. Introduction

It is difficult to carry out efficient national research and development projects due to the sluggish commercialization rate of research and development project performance and research and development results in Korea. In order to successfully carry out research and development projects, it is necessary to secure the feasibility of devlopment technologies first. And then, it will led to the scientific results such as research paper, patent, software etc.. These days the scientific research results are fragmently occured. In order to more efficiently investment for national research and development project, they must need opertion method for can be promoted to economic results. Therefore, this study aims to investigate the characteristics of national research and development projects that have produced economic outcomes.

2. The classification method of security R&D programs for economic research outcomes

Research and Development is defined as a systematic and creative activity that acquires new knowledge about objects, phenomena, and functions, or applies already acquired knowledge [1]. Among various projects that conduct

"This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the ITRC(Information Technology Research Center) support program(IITP-2023-2018-0-01799) supervised by the IITP(Institute for Information & communications Technology Planning & Evaluation)“

research and development, “Projects supported by budget or funds by central administrative agencies for research and development projects in the field of science and technology [2]” In the case of the “Act on Performance Evaluation and Management of National research and development projects, etc.”, It is limited to “National research and development project” and is managed and operated. In order to effectively manage and operate these national research and development projects, national research and development projects are classified and managed according to various classification standards.

National research and development projects can be classified according to basic project information, technology field information that classifies project types, and technology level information. Thinking about these research and development stages at outcomes side, the main research results of the basic research stage are spread through thesis writing. And technological outcomes, such as research paper, are converted into intellectual property rights, such as software registration, patent, etc. and become the basement of technology transfer. The applied research stage or development research stage can create new results that can contribute to related industries, including the results of papers and patents in the basic research stage. The technology life cycle consists of five categories, first is introductory stage, second is growth stage, third is maturity stage, fourth is decline stage, and last one is others.

3. Analysis the Character of Research Project with Economic Outcomes

In this study, for research and analysis, “technology transfer” and “commercialization” that can be inquired and aggregated in National Science & Technology Information Service was set a s the confirmation index of economic research outcomes. Research and analysis were conducted for copyright protection and use activation technology development projects. we intend to conduct research and analysis on the projects of the last three years (2019-2021). The number of tasks performed in the last 3 years is confirmed as 52 projects, and each project consists of an average of 496 million won. In addition, the information protection core source technology development project is a core source technology development project in the information protection field to create a safe national cyber environment and respond to emerging security threats such as ransomware. The number of tasks performed over the past three years was confirmed as 126 tasks, and each task consisted of an average of KRW 715 million. Descriptive statistical analysis is performed based on the technology level information (research and development level, research and development characterization, technology life cycle) of the investigated research project.

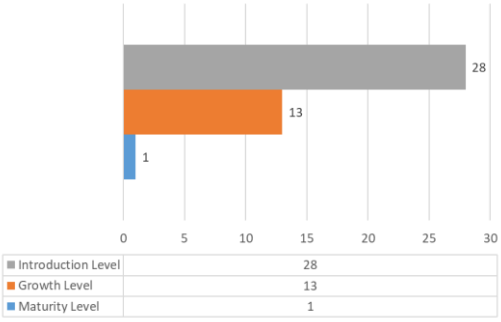


Fig. 1. Technology Life Cycle (Copyright)

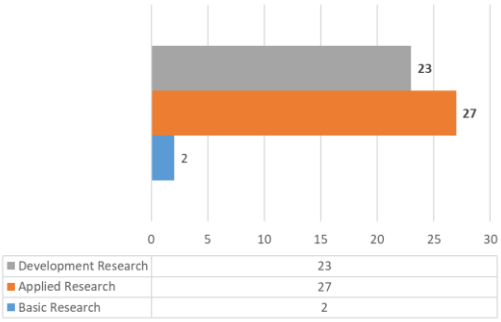


Fig. 2. Technology Development Step (Copyright)

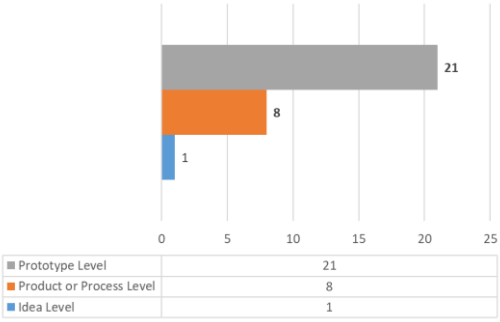


Fig. 3. Feature of Technology (Copyright)

This study analyzes technology information of research and development projects. **Fig. 1** is analysis by technology life cycle, **Fig. 2** is analysis by technology development step, **Fig. 3** is analysis by feature technology. 34 projects have economic research results (technology transfer, industrialization). Result of analysis by technology life cycle is 28 projects in Introduction Level. Result of analysis by technology development step is 23 projects in Development Research. Result of analysis by feature of technology is 21 projects in Prototype Level.

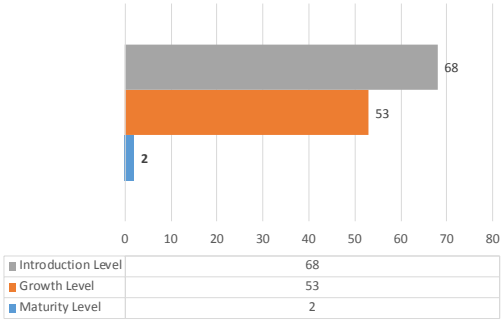


Fig. 4. Technology Life Cycle (Information Security)

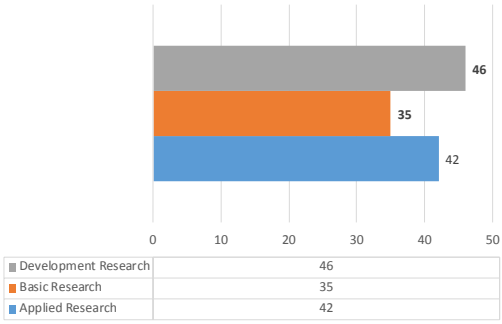


Fig. 5. Technology Development Step (Information Security)

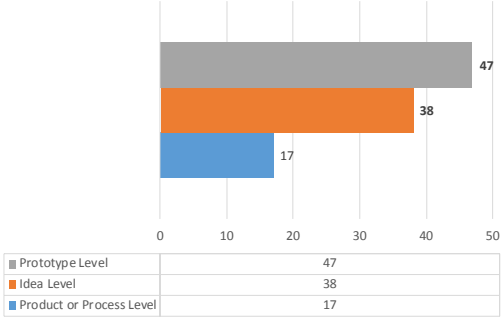


Fig. 6. Feature of Technology (Information Security)

From **Fig. 4** to **Fig. 6** is analysis result in information security research project. Result of analysis by technology life cycle is 68 projects in Introduction Level. Result of analysis by technology development step is 46 projects in Development Research. Result of analysis by feature of technology is 47 projects in Prototype Level.

4. Conclusions

In this paper, the technical information of the research project was analyzed in order to investigate the measures for creating economic performance of the developed technology. In general, it is predicted that it will be possible to create economic results in applied research rather than basic research, growth/maturity research rather than introductory research, and prototype level rather than idea level. However, judging from the results of recent research projects, more economic research results have been created in basic research. As a result of the analysis, 17 of the 46 development studies in the information security technology development project produced economic results, but 20 of the 35 basic research projects produced economic results. This means that the technical information classification of a research project cannot represent the performance characteristics of a research project and cannot serve as a criterion for managing R&D projects to create economic results. As future research, we will analyze the characteristics of tasks that have generated economic results. Based on the research results of this paper, it is expected that it will be possible to design a blockchain service that can manage research projects from the viewpoint of deriving economic results.

References

[1] OECD, “Frascati Manual 2015: Guidelines for collecting and reporting data on research and experimental development,” *The Measurement of Scientific, Technological and Innovation Activities, OECD Publishing*, 2015.

[2] Act On The Performance Evaluation And Management Of National Research And Development Projects, 2021.

Copyright ⓒ 2023 KSII

Development and Application of Computer Science Textbooks: Focusing on Digital Therapeutics

Eunsun Choi and Namje Park*
Department of Computer Education, Teachers College, Jeju National University
Jeju, South Korea
[e-mail: choi910624@jejunu.ac.kr, namjepark@jejunu.ac.kr]
*Corresponding author: Namje Park

Abstract

Computer science is being conducted as primary literacy education entering the intelligent information society. In this paper, we developed textbooks that teach elementary and secondary school teachers the principle of digital therapeutics (DTx). Textbooks for elementary school teachers were designed with the principle of DTx for Attention-Deficit Hyperactivity Disorder (ADHD), and textbooks for middle school teachers were developed with DTx corresponding to digital dramas as the central theme. The textbooks were developed based on the ASSURE model and included teaching and learning plans, worksheets, and reading materials for immediate use in the school field. In addition, textbooks can be used in non-face-to-face classes by using customized online teaching tools. In order to confirm the effectiveness of the developed textbooks, we applied the textbooks to 33 Korean teachers. As a result, teachers' information teaching efficacy using textbooks was statistically significantly improved.

Keywords: Textbook development, computer science education, digital therapeutics, information teaching efficacy

1. Introduction

Intelligence information technology, combining artificial intelligence with the Internet of Things, big data, cloud, and mobile-based information, has emerged as a critical keyword of the fourth industrial revolution. Therefore, it is becoming essential to cultivate talents who understand it. Digital therapeutics (DTx) was selected as a promising future technology to pay attention to by world-renowned research institutes such as the World Economic Forum [1] and McKinsey & Company [2]. Also, in South Korea, the Korea Health Industry Promotion Agency focused on DTx as one of the top 10 technology trends 2021 [3]. This paper shows the results of developing textbooks that can teach students the principles of DTx for elementary and secondary school teachers to foster future talents with an understanding of information technology. The textbooks were designed with the theme of DTx for Attention Deficit/Hyperactivity Disorder for elementary school teachers and DTx for digital drama, a type of cyber violence that occurs among teenagers, for middle school teachers.

2. Materials and Methods

The textbooks proposed in this paper were based on the ASSURE model that is classically used in developing computer science textbooks. This model suggests designing education across six stages: Analyze Learners, State Objects, Select

This work was supported by the Ministry of Education of the Republic of Korea and the National Research Foundation of Korea (NRF-2022S1A5C2A04092269). And, this work was supported by Electronics and Telecommunications Research Institute(ETRI) grant funded by the Korean government. [23ZD1160, Regional Industry ICT Convergence Technology Advancement and Support Project in Daegu-GyeongBuk(Mobility)]

Methods, media and materials, Utilize media and materials, Require Learner Partnership, Evaluate and Revise [4]. We used this model to develop textbooks for elementary and middle school teachers. The textbooks were provided to 18 elementary and middle school teachers in Korea to analyze their effectiveness. In addition, we conducted an information teaching efficacy test on elementary and secondary teachers before and after using textbooks. To this end, the test paper presented by Park Heejung et al. (2021) was used, and it consisted of class values, teaching strategies, and infrastructure utilization factors [5]. In this test tool, Cronbach's alpha coefficient was 0.867, and the tool's validity was verified. Furthermore, as an analysis method, the Shapiro -Wilk test was performed for the normality test, and technical statistics and the Wilcoxon code ranking test, a non-parametric test, were performed.

3. Results

3.1 Development of Computer Science Textbooks focusing on DTx

The textbooks proposed in this paper were developed according to the ASSURE model. We proposed the composition, content, and characteristics according to the subject of the textbooks. This model considers learning media and teaching materials as the main items that determine the quality of learning content. Accordingly, the textbooks also propose media and tools tailored to the changing classroom environment, and various types of class materials, such as unplugged and online worksheets.

To analyze the learner stage, the first procedure of the ASSURE model, we used data from the 2020 Elementary and Secondary Career Education Survey by the Ministry of Education (2021) [6]. Among the top 10 jobs desired by elementary school students across the country, there were no jobs directly connected to computer science, and middle school students were hoping for the bottom. According to these results, it was recognized that education and textbooks for computer science were urgently needed. The second procedure is state objects. The goal of the textbooks was primarily to improve teachers' future information class teaching capabilities, and the ultimate goal was to contribute to the growth of students into talents required by the future society. The third procedure is the selection of teaching methods, media, and materials. In the textbooks, we selected EdApp, an online customized learning tool, and adopted gamification that can increase interest or unplugged education techniques for various educational environments. The fourth procedure is utilizing media and materials. Fig. 1 shows part of the first screen of the Attention-Deficit Hyperactivity Disorder (ADHD) DTx class using EdApp, one of the media selected in the third step. The fifth step is the step of inducing learner participation. The textbooks encouraged learners to participate in classes by increasing learners' interest in DTx, a relatively unfamiliar concept. The final step is evaluation and revision. Evaluation tools used when teachers conduct classes with the textbooks were developed as creativity evaluation, self evaluation, peer evaluation, and reflection log.



Fig. 1. Online learning using EdApp

The ADHD DTx principle textbook for elementary school teachers provided a curriculum proposal and materials that can be taught in 4 classes. The first session focuses on understanding the basic principles of ADHD DTx and conducts a debate on the pros and cons of ADHD treatment through games. The second session is about the application of ADHD DTx, and students experience attention improving games. The third session examines the application case of ADHD DTx and performs an unplugged learning game that can understand the Brain Computer Interface. The fourth session is based on the future of ADHD DTx, and students draw their future ADHD treatments using artificial intelligence tools.

The DTx principle textbook corresponding to digital dramas for middle school teachers presented three classes. The first session is designed to understand the principles of

cyberbullying and DTx, with the theme of understanding DTx in response to digital dramas. In the second session, students experience creating cyber violence prevention videos to create DTx corresponding to digital dramas indirectly. In the third session, students share opinions on how to respond to digital dramas effectively. **Table 1** shows the main contents of each session proposed in the textbook.

Table. 1. Curriculum for Understanding DTx Principles

Level	Topic	Class	Contents
Elementary School	Understanding ADHD DTx Basic Principles	1/4	• ADHD and DTx concepts • How to treat neuro developmental disorders through digital technology and gaming with Selective Stimulus Management Engine
	Application of ADHD DTx	2/4	• ADHD DTx cases • Effective communication skills using sandwich methods
		3/4	• Three dimensional multi object tracking principles • BCI as a game to explore cognitive improvements
	Future of ADHD DTx	4/4	• Prospect of development of ADHD DTx • Effective ways to treat ADHD
Secondary School	Understanding DTx Against Digital Drama	1/3	• Differences between digital drama and cyberbullying • DTx concepts
	Development of DTx Against Digital Drama	2/3	• DTx cases to prevent cyberbullying • Examples of DTx that treat victims of digital drama or prevent digital drama
	Proposal for Digital Drama Response Plan	3/3	• Production of digital drama prevention education video using artificial intelligence platform • Discuss ways to effectively respond to and prevent digital dramas

3.2 Results Analysis of the Information Teaching Efficacy Test

We first performed normality verification to analyze the results of the information teaching efficacy test of elementary school teachers. Shapiro-Wilk test was performed on the pre-test results, and the normality of all factors was not satisfied (p<.05). Thus, the results were analyzed using the Wilcoxon code ranking test, a non-parametric statistical method. As a result of the verification, among the three factors, the factor with the highest prior-average was the factor of infrastructure utilization (mean = 3.33, standard deviation = 0.686). Similarly, the factor with the highest post-average was the factor of infrastructure utilization (mean = 4.00, standard deviation 0.767). In all factors of class value, teaching strategies, and infrastructure utilization, the average score increased afterward, and only the infrastructure utilization factor showed a significant change in statistics (p<.05).

In order to analyze information teaching efficacy in secondary school teachers, normality verification was conducted in the same way, and the Shapiro-Wilk test showed that normality was not satisfied (p<.05). Therefore, the Wilcoxon code ranking test was performed. As a result of

the test, the factor with the highest prior average was infrastructure utilization (mean = 3.53, standard deviation = 0.516), and the factor with the highest post-average was infrastructure utilization (mean = 4.53, standard deviation = 0.516). In contrast, the factor with the lowest prior-average was the teaching strategies (mean = 2.60, standard deviation = 0.507), and the factor with the lowest post-average was the class value factor (mean = 3.67, standard deviation 0.617). For secondary teachers, statistically significant improvement occurred afterward rather than before in all factors. **Table 2** shows the pre-post average change in information teaching efficacy of elementary and secondary teachers.

Table. 2. Information Teaching Efficacy of Pre-Post in Elementary and Secondary teachers

Level	No.	Factor		M	SD	z	p
Elementary School	18	C.V.	Pre	3.06	0.76	-1.43	0.15
			Post	3.39	0.61		
		T.S.	Pre	2.83	0.86	-1.89	0.06
			Post	3.33	0.77		
		I.U.	Pre	3.33	0.69	-2.17*	0.03
			Post	4.00	0.77		
Secondary School	15	C.V.	Pre	3.13	0.64	-2.28*	0.02
			Post	3.67	0.62		
		T.S.	Pre	2.60	0.51	-3.27**	0.00
			Post	3.93	0.60		
		I.U.	Pre	3.53	0.52	-3.63***	0.00
			Post	4.53	0.52		

C.V. = Class Value, T. S. = Teaching Strategies, I. U. = Infrastructure Utilization, $^{*}p < .05$, $^{**}p < .01$, $^{***}p < .001$

4. Conclusions

DTxs are relevant to information technology and the future medical field. As information and communication technology develops in the future, the DTx market is expected to grow further. For students to grow into talents who can lead the future industry, they must develop their ability to think creatively and solve problems with an understanding of promising technologies in the future [7]. For this, teacher guidance is paramount. It is because an education system's quality cannot exceed its teachers' quality [8]. The actual change in education in the field depends on the teacher.

This paper proposes computer science textbooks that can increase the understanding of the principles of DTx for elementary and secondary teachers and show the results of analyzing the educational effectiveness thereof. It was analyzed that textbooks developed by applying the ASSURE model enhance teachers' information teaching efficacy. However, since the number of study subjects was small and analyzed using non-parametric statistical techniques, it is required to increase the number of study subjects and to conduct detailed studies through demographic classification in future studies.

References

[1] World Economic Forum, “Top 10 Emerging Technologies of 2020,” pp. 1-26, 2020.

[2] McKinsey Global Institute, “Ten innovations that can improve global health,” 2020. https://www.mckinsey.com/industries/healthcare/our-insights/ten-innovations-that-can-improve-global-health

[3] Korea Health Industry Development Institute, “An In-depth Survey of Global Health Industry Trends in the First Half of 2021,”, pp. 1-39, 2021.

[4] R. Heinich, M. Molenda, J. Russell, et al., “Instructional Media and Technologies for Learning,” 7th edn. Prentice Hall, Englewood Cliffs, 2001.

[5] H. Park, H. Kim, J. Choi, et al., “Development of teaching efficacy instrument in Informatics(Software and AI) subject,” *The Journal of Korean Association of Computer Education*, Vol. 24, No. 4, pp. 39-52, 2021.

[6] Ministry of Education (South Korea), “2020 Announcement of the Results of the Survey on the Status of Career Education in Elementary and Secondary Schools,” pp. 1-31, 2021.

[7] E. Choi, N. Park, “Digital therapeutics from data analytics and information security perspectives,” *Mathematical Statistician and Engineering Applications*, Vol. 71, No. 3s, pp. 521-526, 2022.

[8] M. Barber, M. Mourshed, “How the world’s best-performing school systems come out on top,” 2007. https://www.mckinsey.com/industries/education/our-insights/how-the-worlds-best-performing-school-systems-come-out-on-top

Copyright ⓒ 2023 KSII

Next Activity Prediction and Analysis Based Process Mining

Jin-Hyeong Lee[1], Eunbi Cho[2], Tsagaantsooj Batzaya[2], So-hyang Bak[3], Thanh-hai Nguyen[3], and Kwanghoon Pio Kim[4]
[1]Department of Computer Engineering, Kyonggi University
[2]Department of Public Safety Bigdata, Kyonggi University
[3]Department of Computer Science, Kyonggi University
[4]Division of AI Computer Engineering, Kyonggi University
Seoul, South Korea
[e-mail: {jaims2289, eunbee0508, tsbaachka95, sa7291, nguyenthanhhai, kwang}@kyonggi.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Abstract

Accurately predicting the next activity is an important problem in predictive business process research. Recently, research on predicting the next activity through predictive business processes using deep learning techniques has been actively conducted. In this paper, each feature of the process is found through process mining from the process execution history data, The next activity prediction model for predictive business process monitoring is experimentally verified based on LSTM and transformer models specialized for time series characteristics. To analyze the reasons for the different performances of the two models in the next activity prediction experiment, we find patterns in the process data based on process mining.

Keywords: next activity, Predictive monitoring, LSTM, Transformer, deep learning

1. Introduction

Due to recent rapid changes in the global economy, companies are establishing and utilizing more efficient business processes to reduce costs, secure competitiveness, and improve productivity. A business process is the procedure of performing business in a company, and there are various steps. Therefore, To build an efficient business process, companies record these business processes in the form of event logs in the system. Since the process should be delayed or stopped because of various factors inside and outside the company, the company can quickly identify and respond to problems by continuously monitoring them through process execution history data analysis, which records the overall process in the form of a log. Existing process monitoring analyzes previously recorded data to detect abnormal processes in the past or monitor current processes, whereas Predictive Business Process Monitoring predicts the future based on past process data. Especially in predictive business process monitoring, next activity prediction can provide a path to the next step of an instance based on historical data, which can increase work efficiency. Since process execution history data has a sequential nature and deep learning models such as LSTM and transformer models are currently being actively used for the next activity prediction, we experimentally verify the next activity prediction based on LSTM and transformer models specialized for time series characteristics. In this paper, we confirm that LSTM and transformer have similar prediction performance for a specific dataset through the next activity prediction experiment. Therefore, we would like to analyze the specific dataset through process mining.

2. Related Work

2.1 LSTM

LSTM(Long Short-Term Memory)[1] is a type of RNN, a time series deep learning model that processes inputs and outputs sequentially. LSTM is developed to alleviate the problem of the RNN

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT), Republic of Korea (Grant No. NRF-2022R1A2C2093002).

model, and the problem of RNN which has difficulty in handling long-term dependence. Long-term dependence is a phenomenon in which the past information in the hidden layer is not transferred to the last layer, which is related to the vanishing gradient. The vanishing gradient is the error through backpropagation, where the gradient value becomes very small or large during the backpropagation. LSTM is composed of three gates that allow them to remember old information. These gates only remember important information and pass it on to the next layer. Business process data is characterized by sequential data, and LSTM can learn correlations between events that are far apart in time.

2.2 Transformer

Transformer[3] consists of self-attention and multi-head attention to extract relationships between input sequences. Since the transformer considers information from more distant time periods and it doesn't has a sequential model structure, Transformer can process information in parallel for faster and more accurate predictions. Furthermore, Transformer can effectively analyze long sequence data compared to RNN-based models, which degrade when the input sequence is long. Therefore, this paper uses the Process Transformer[4] model, which is specialized for process data.

3. Next activity prediction

3.1 Data Structure Analysis

The dataset used in this study is BPIC 2012, Request for payment, review_example_large, and helpdesk, which are frequently used in the field of predictive monitoring of business processes. Data preprocessing uses process mining from process execution history data to discover the characteristics of each process. The preprocessing results are shown in **Table 1**.

Table 1. Preprocessing data

Dataset / Attribute	BPIC2012	Request for payment
Min activity length	3	1
Max activity length	175	20
The number of activities	24	19
The number of traces	13087	6886

Dataset / Attribute	Review_example_large	helpdesk
Min activity length	11	2
Max activity length	86	15
The number of activities	14	14
The number of traces	10000	4580

In the case of business processes, patterns in the form of Sequential, OR, AND, and LOOP occur. The four patterns that occur in business process data are shown in **Fig. 1**. The result of business process pattern mining on the actual Review_example_large data is shown in **Fig. 2**. These patterns affect the prediction of the next activity. Therefore, our team analyzes three of these four data sets.

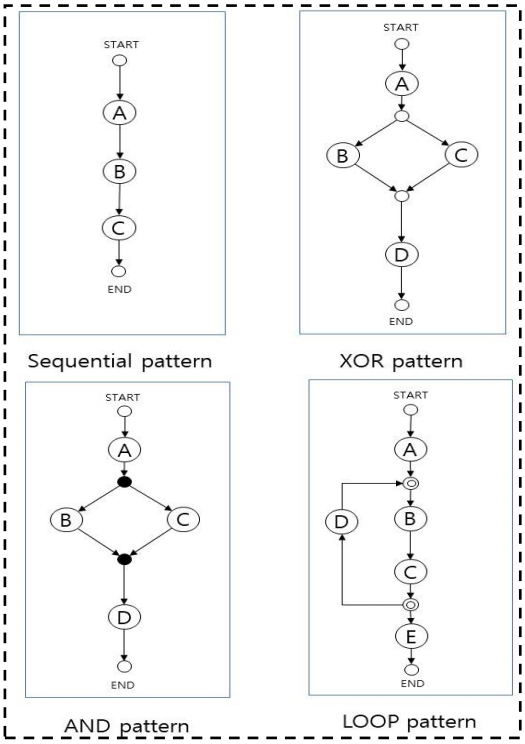


Fig. 1. Business process Patterns

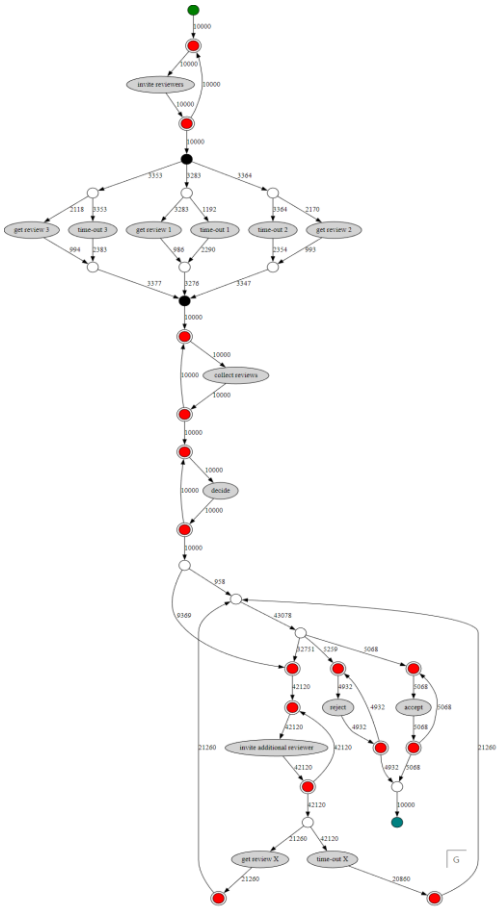


Fig2. Review_example_large patterns mining

3.2 Next activity prediction

The model hyper-parameter values used to predict the following activities are shown in **Table 2**. The results of predicting the next activity based on LSTM and Transformer for the four datasets are shown in **Table 3**.

Table 2. Model Hyperparameters

	LSTM	Transformer
Epoch	100	10
Batch_size	15	12
Learing_rate	0.002	0.001
Loss_function	Cross entropy	Cross entropy
optimizer	Adam	Adam

Table 3. Next Activity Predicted Performance

Model / Data	LSTM	Transformer
BPIC 2012	64.51	82.96
Request for payment	88.52	88.22
Review_example_large	78.27	77.34
helpdesk	70.39	84.56

The accuracy analysis shows that BPIC 2012 and helpdesk have higher accuracy on the Transformer model, But Request for payment and review_example_large have slightly higher accuracy on the LSTM model. To see what factors in the data affected the accuracy, we used the BPIC 2012 dataset, which has higher accuracy on Transformer, and the review_example_large dataset with higher accuracy on LSTM using a process mining tool to identify the features of the datasets through process mining. As a result of process mining, there are 15 LOOP patterns in BPIC2012 dataset and 22 LOOP patterns in the review_example_large dataset.

4. Conclusions

In this paper, we experimentally validate the use of LSTM and transformer models to predict the next activity for predictive business process monitoring. In previous studies, the transformer model often performs better than the LSTM model. However, this paper shows that both models have similar next activity prediction performances on the Request for payment and Review_example_large datasets. For the analysis of this performance difference, we analyzed the structural patterns of the datasets through process mining of the datasets. In future work, we would like to further improve the performance of the Transformer model by changing the values of the model's hyperparameters such as loss function, optimizer, and layers to better fit the process data.

References

[1] Sepp Hochreiter, Jurgen Schmidhuber, "Long short-term memory", Neural Computation, Vol.9, pp.1735-1780, 1997.

[2] Niek Tax, Ilya Verenich, Marcello La Rosa, Marlon Dumas, "Predictive Business Process Monitoring with LSTM Neural Networks", Springer, CAiSE, pp.477-492, 2017.

[3] Vaswani, Shazeer et al., "Attention is all you need", NIPS, pp.5998-6008,2017.

[4] Zaharah A. Bukhsh, Aaqib Saeed, Remco M. Dijkman, "ProcessTransformer: Predictive Business Process Monitoring with Transformer Network", arXiv, 2021.

Copyright ⓒ 2023 KSII

Similarity estimation between distributed inference environments using vision transformer embedding vectors

Jongbin Park*
Information Media Research Center
Korea Electronics Technology Institute
Seoul, South Korea
jpark@keti.re.kr
*Corresponding author: Jongbin Park

Abstract

Edge computing is a distributed computing approach that utilizes computing resources in a decentralized manner to process and store data, providing faster response times and enabling large-scale data processing from IoT devices and mobile devices. Edge computing is widely applied in various fields such as IoT, smart cities, autonomous vehicles, medical devices, and manufacturing. However, the diverse input data from multiple devices can lead to uncertainty in model performance. To address this issue, a proposed solution involves measuring the similarity between heterogeneous edge inference environments using a Vision Transformer to extract embedding vectors from images. This approach clusters information to select suitable models for edge inference environments and can be used as prior information for federated learning.

Keywords: Edge Computing, Edge Node Clustering, Vector Space Embedding, Vision Transformer

1. Introduction

Edge computing is a decentralized computing architecture where data processing and storage are performed closer to the source of the data, rather than relying on centralized cloud computing infrastructure [1,2]. This computing approach can reduce data interactions because it processes and stores data locally on devices that are close to the source of data generation, rather than transmitting the data to a centralized cloud server for processing. By processing and analyzing data locally, edge computing reduces the amount of data that needs to be transmitted back and forth between the device and the cloud, which can improve overall system performance and reduce network congestion. This is especially important for applications that require real-time processing and low latency, such as autonomous vehicles, industrial control systems, and healthcare devices [1,2].

Even though edge devices use the same inference model, the prediction performance varies with the characteristics of the data. This causes uncertainty in model performance, and even models that perform well in the learning environment can suffer performance degradation in the inference environment [1,2].

A common approach to solving this problem is usually to focus on training models using rich datasets, or designing models that can handle different data distributions. We are also making efforts to optimize the size and structure of the model for the inference environment, such as using knowledge distillation. Nevertheless, these methods can also cause performance degradation if data characteristics change over time or data bias occurs [1,2].

This paper proposes a technique for measuring the similarity between heterogeneous edge

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No. 2021-0-00907, Development of Adaptive and Lightweight Edge-Collaborative Analysis Technology for Enabling Proactively Immediate Response and Rapid Learning).

inference environments as a solution to these problems. Similarity analysis among multiple edges can be utilized to select models that are well-suited for edge inference environments or to leverage them as prior information for federated learning [3-6].

In particular, this paper utilizes Vision Transformer techniques [7,8] to extract embedding vectors from inference images. These vectors are then transmitted to a cloud server called Edge Vision Cluster (EVC) server, which measures the similarity among multiple edge computing environments using these vectors.

2. Embedding Vector Extraction at the Inference Edge Nodes

2.1 Vision Transformer

Vision Transformer (ViT) is a type of neural network architecture that applies the Transformer model to image recognition tasks [7]. The Transformer model was proposed for natural language processing [8], and it achieved state-of-the-art performance in many natural language-related tasks [9, 10].

ViT extends the Transformer model to process image data by splitting an input image into patches and treating each patch as a sequence of tokens. The model then processes these patches using self-attention mechanisms to learn spatial relationships between different patches [7,11].

The key advantage of ViT over traditional convolutional neural networks (CNNs) is its ability to learn long-range dependencies between pixels in an image without the use of explicit convolutional operations [7,11].

2.2 Embedding Vector Extraction

From a vector embedding perspective, the ViT model can be thought of as learning a mapping between image patches and high-dimensional embedding vectors. The embedding vectors are learned in such a way that they capture important visual features of the image patches, which can then be used to recognize objects and scenes in the image [7,11].

Fig. 1 illustrates the summarized process of performing ViT on the input image and extracting the embedding vector. **Fig. 2** depicts the process of transmitting the extracted embedding vectors from multiple edge nodes to the cluster for analysis.

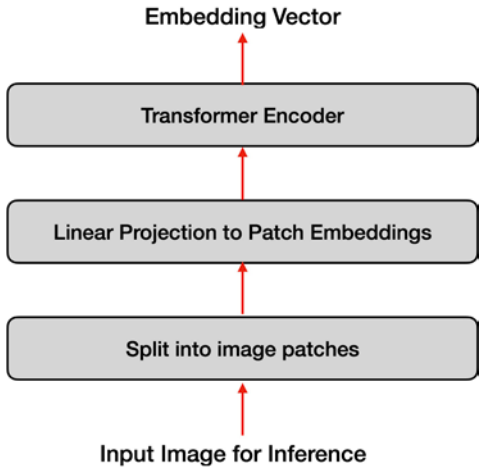


Fig. 1. Embedding vector extraction process for single input image

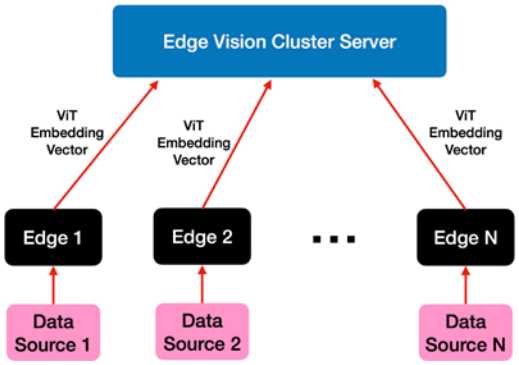


Fig. 2. Transfer the extracted embedding vector from edge nodes to the cloud server

3. Similarity Analysis between Edge Nodes in the Cloud Server

The transferred embedding vectors are used to calculate similarities between nodes in the Edge Vision Cluster Server in **Fig. 2**. In this paper, we use Canonical Correlation Analysis (CCA) metric [12] to measure the similarity among multiple edge domains. **Fig. 1** shows the similarity comparison results based on CCA for five edge nodes. In addition, with the received embedding vector, EVC server can calculate the mean of cosine similarity among domains, and

can also apply principal component analysis (PCA) to perform inter-domain clustering in low-dimensional space.

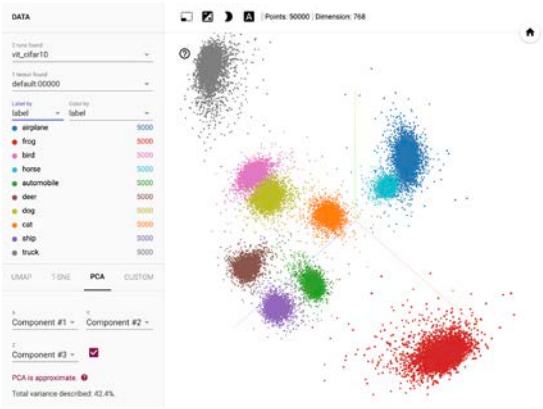


Fig. 3. Example of 3D space projection of ViT embedding vectors for 50,000 CIFAR-10 images

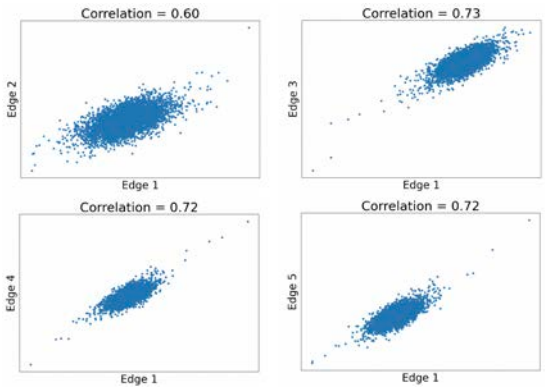


Fig. 4. Example of CCA matric based similarity comparison for 5 edge environments

4. Conclusions

In this paper, we applied a method to extract embedding vectors from inference images to measure the similarity of inference environments among multiple edge nodes. These embedding vectors were obtained using Vision Transformer. The similarity information among edge nodes can be utilized in various ways in the future. It can be used to cluster edge nodes, select models suitable for edge inference environments, or serve as prior information for federated learning.

References

[1] S. Weisong, et al. “Edge computing: Vision and challenges,” *IEEE internet of things journal*, vol. 3, no. 5, pp. 637-646, 2016.

[2] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. y Arcas, “Communication-efficient learning of deep networks from decentralized data,” in *Artificial intelligence and statistics. PML*R, pp. 1273–1282, 2017.

[3] M. Aledhari, R. Razzak, R. M. Parizi, and F. Saeed, “Federated learning: A survey on enabling technologies, protocols, and applications,” *IEEE Access*, vol. 8, pp. 140 699–140 725, 2020.

[4] K. Bonawitz, H. Eichner, W. Grieskamp, D. Huba, A. Ingerman, V. Ivanov, C. Kiddon, J. Kone˘cn`y, et al., “Towards federated learning at scale: System design,” *Proceedings of Machine Learning and Systems*, vol. 1, pp. 374–388, 2019.

[5] T. Li, A. K. Sahu, A. Talwalkar, and V. Smith, “Federated learning: Challenges, methods, and future directions,” *IEEE Signal Processing Magazine*, vol. 37, no. 3, pp. 50–60, 2020.

[6]L. U. Khan, S. R. Pandey, N. H. Tran, W. Saad, Z. Han, M. N. Nguyen, and C. S. Hong, “Federated learning for edge networks: Resource optimization and incentive mechanism,” *IEEE Communications Magazine*, vol. 58, no. 10, pp. 88–93, 2020.

[7] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, N. Houlsby, “An image is worth 16x16 words: Transformers for image recognition at scale,” In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 3156-3164, 2020.

[8] A. Vaswani, et al. “Attention is all you need,” *Advances in neural information processing systems*, vol.30, pp. 1-15, 2017.

[9] URL, https://paperswithcode.com/

[10] URL, https://www.kaggle.com/

[11] S. Khan, et-al, “Transformers in vision: A survey”, *ACM computing surveys (CSUR)*, no. 10, vol. 54, pp. 1–41, 2022.

[12] A. Beyer, G. Kauermann, H. Schütze, “Embedding Space Correlation as a Measure of Domain Similarity,” *Proceedings of the 12th Language Resources and Evaluation Conference*, pp. 2431-2439, 2020.

Copyright ⓒ 2023 KSII

A framework to Support the Development of Safety Guidelines in the Manufacturing fields using Ontology

Taehoon Kim, Jinmyeong Lee, Jungmin Yun and Jungnyun Lee
Research Institute of Industrial Technology Convergence, Korea Institute of Industrial Technology
Ansan 15588, South Korea
[e-mail: thkim216@kitech.re.kr, ljm2910@kitech.re.kr, jmyun@kitech.re.kr, blueleen2@kitech.re.kr]
*Corresponding author: Taehoon Kim

Abstract

This paper proposes an ontology-based framework to support the development of safety guidelines in manufacturing. Safety guidelines are crucial in managing and executing manufacturing fields, but developing them can be time-consuming and costly, especially for Small and Medium-sized Enterprises (SMEs). The proposed framework comprises two modules: a hazard identification module and a safety guideline suggestion module. The hazard identification module infers risk factors and situations by using relationships defined in the manufacturing ontology and work environment information. The safety guideline suggestion module identifies the machine and related accidents based on inferred risk factors and suggests suitable safety guidelines by calculating safety devices and minimum distances. The proposed framework aims to prevent accidents and save time and costs by considering safety standards, toxic materials, and operation methods. The framework is expected to assist safety managers in making safety guidelines with economic cost and time, and to suggest safety management standards suitable for each work environment. The study contributes to accident prevention and to the development of work safety guidelines for SMEs.

Keywords: Ontology, Safety guideline, Decision support framework, Manufacturing fields

1. Introduction

Governments of many countries are making great efforts to prevent safety accidents. However, in the case of small and medium-sized enterprises that lack the capacity to establish and operate safety management systems, it is difficult to establish and operate safety management guidelines suitable for each company with limited time and expenses. Consequently, SMEs pose a higher risk of accidents due to low education level, lack of safety awareness among workers, and poor implementation of informal safety procedures and fundus management [1].

There have been other previous studies to reduce such safety accidents. For example, studies on analyzing the relationship between safety accidents and human conditions and experiences were conducted[3, 4], and also studies on safety operation for collaborative robots, which have shown high occurrence of field accidents, are recently carried out[5,6].

However, it is difficult to create and operate safety guidelines by quantifying human behavior patterns by reflecting these studies. In addition, analysis methods for the safe operation of collaborative robots are difficult to apply commonly in practice due to the characteristics of SMEs with various sizes and configurations of actual manufacturing sites.

This paper was supported by Korea Institute for Advancement of Technology (KIAT) grant funded by the Korea Government (MOTIE)(P0014286, Industry Innovation Foundation Building Program)

Therefore in this study, we propose a framework that can support the development of safety guidelines suitable for each company's situation through the information of each small and medium-sized company's workplace by using ontology. At this time, the safety guidelines include preliminary safety precautions necessary for the manufacturing field, possible accidents, and guides for quantitative values to prevent them.

The proposed framework consists of two modules. The purpose of the first module is to conceptualize field information and risk information based on ontology. That of the second module is to suggest guidelines to reduce risks of similar accidents by managing accident cases as instances. In the first module, possible hazards are identified based on the manufacturing field information. The second module considers each risk factors and suggest safety guidelines by listing the equipment that need to be prepared before manufacturing operations, and by identifying the potential field areas vulnerable to accidents.

By acknowledging each various manufacturing field, the modules will help develop guidelines that can prevent safety accidents for small and medium-sized enterprises that cannot spend a lot of time and money in constructing appropriate guidelines.

2. Framework to Support the Development of Safety Guidelines

The system proposed in this study consists of two modules and a defined manufacturing safety ontology including workplace information, risk factors, and accident information. The first module is the Risk Factor Identification Module based on Work Environment Information, and the extracted risk factors are sent to the second module, which is the Safety Guideline Suggestion Module based on Risk Factor. This system flow is shown in **Fig. 1**.

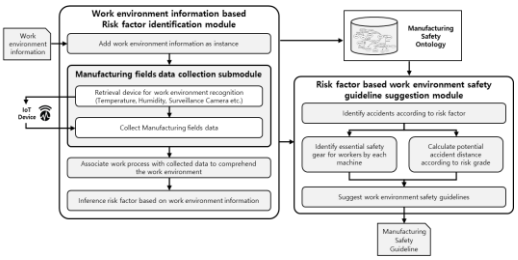


Fig. 1. Overall framework to support the development of safety guidelines

2.1 Manufacturing Safety Ontology

Ontology utilizes Protégé, an open source editor and reasoner, and the ontology model is constructed by using the following steps.
Conceptualize the "Material" and "Worker" required for the workplace based on the "Resource" class.
Conceptualize the "Machine", "Devices", and "Gear" installed in the workplace based on the "Equipment" Class.
Conceptualize the "Process" of mechanical equipment and related accidents based on the "Machine" class.
Conceptualize "Risk factor" and "Accident" based on "Process" class and "Accident" class.
Define the relationship attribute according to the relationship of each class and add the instance based on the manufacturing accident case.
If the proposed ontology is used, risk factor inference and concept expansion for manufacturing safety are possible with low cost and effort.



Fig. 2. Example of manufacturing safety ontology

2.2 Risk Factor Identification Module based on Work environment Information

Each enterprise has a different size and composition of the manufacturing field. In order to reflect the characteristics of the field, the work environment information (machine, device, process, worker, materials, etc.) is instantiated. At this time, the equipment must be configured

as individual instances so that the data generated by the machine can be collected and utilized.
Next, in order to collect data generated at the manufacturing field, the system searches for the devices that can collect the environment data, and the data are collected by using them. This sub-module is used only when data collection is possible at the field.
Collectable data is divided into two types: (1)essential data, which is the data generated by equipment, and (2)optional data, which is manufacturing environment data. By linking the data to the process information, understanding of field condition can be enhanced.

2.3 Safety Guideline Suggestion Module based on Risk Factor

This module suggests safety guidelines based on work environment information. For this purpose, possible accidents are identified based on the risk factors inferred from the previous module. Each accident is used to calculate the possible radial distance from the machine and to identify essential safety gear that can reduce the severity of the accident. As a result, it is possible to calculate the hazardous area for each work area and proposes a work environment guideline that can provide information on the safety gear that must be worn in each hazardous area.
When these safety guidelines are used together with spatial information, it will help identify the work area and identify the information of objects to be detected in each area.

4. Conclusions and Future Work

In this paper, we propose an architecture to support the development of safety guidelines for small and medium-sized enterprises using a manufacturing safety ontology in which the relationship between workplace characteristics and risk factors is defined.
To serve the purpose, we developed a manufacturing safety ontology model using five steps, using the knowledge based on manufacturing workplace characteristics and manufacturing accident cases. Based on this, we propose optimized safety guidelines for the workplace through identification of possible risks based on manufacturing field, safety gear according to possible risks, and calculation of hazardous areas.
In the future empirical studies including services such as detection of dangerous objects and notifications according to intrusion of dangerous areas will be conducted by using a deep learning-based analysis model in the proposed architecture.

References

[1] Masi, D., & Cagno, E. "Barriers to OHS interventions in small and medium-sized enterprises." *Safety science*, 71, 226-241, 2015.
[2] Ansori, N., & Widyanti, A. "The role of safety silence motives to safety communication and safety participation in different sectors of small and medium enterprises: investigation results on two kinds of industries in Indonesia." *Safety and health at work*, 12.2, 192-200, 2021.
[3] Perales Gomez, A. L., Fernández Maimó, L., Huertas Celdran, A., Garcia Clemente, F. J., Gil Pérez, M., & Martínez Pérez, G. "SafeMan: A unified framework to manage cybersecurity and safety in manufacturing industry." *Software: Practice and Experience*, 51.3, 607-627, 2021.
[4] Oah, S., Na, R., & Moon, K. "The influence of safety climate, safety leadership, workload, and accident experiences on risk perception: A study of Korean manufacturing workers." *Safety and health at work*, 9.4, 427-433, 2018.
[5] Bi, Z. M., Luo, C., Miao, Z., Zhang, B., Zhang, W. J., & Wang, L. “Safety assurance mechanisms of collaborative robotic systems in manufacturing.” *Robotics and Computer-Integrated Manufacturing*, 67, 102022, 2021.
[6] Wang, P., Liu, H., Wang, L., & Gao, R. X. “Deep learning-based human motion recognition for predictive context-aware human-robot collaboration.” *CIRP annals*, 67(1), 17-20, 2018.

Copyright ⓒ 2023 KSII

Implementation of a Mixed Reality Service with Real-time Interaction of Virtual and Real Worlds

Jong-gak Choe[1] and Yongjin Kwon[2*]
[1] GRRC, Korea Aerospace University
Goyang-si, Gyeonggi-do 10540 - Korea
[e-mail: zonx@tikwon.hangkong.ac.kr]
[2] School of Electronics and Information Engineering, Korea Aerospace University
Goyang-si, Gyeonggi-do 10540 - Korea
[e-mail: yjkwon@kau.ac.kr]
*Corresponding author: Yongjin Kwon

Abstract

Recently, the advancement of virtual reality and augmented reality technologies has led to attention to the metaverse - an extended virtual world that expands reality. However, the metaverse has not been widely utilized in our daily lives because of the difficulty to find connections between the virtual and real worlds. This makes users face challenge of incorporating it into everyday life. One model has been proposed [1], which aims to develop services that generate and utilize virtual information in the real world. In this paper, we propose three key elements required to implement Mixed Reality (MR) with a connection between the virtual and real worlds: location, physics engine, and real-time interaction. By developing a service based on the elements and the model, we show that users can naturally generate and utilize virtual information in the real world. In this study, we demonstrate an example of an MR service implemented using Unreal Engine and Cesium, based on the model and the proposed elements. This service allows users to store information in virtual spaces in real-time, such as lifelogs, where it was generated in the real world. It also allows the users to retrieve that information in virtual spaces based on the real-world location.

Keywords: Mixed Reality, Augmented Reality, Location Based Service, Mirror World, Metaverse

1. Introduction

As computing technology advances, 3D virtual reality-based services such as games have become popular. With the development of mobile communication and positioning technologies, location-based services such as google maps have also become widespread, and Mixed Reality (MR) services are now being attempted. Based on these environments, the metaverse is emerging.

The metaverse is generally defined as “the notional environment in which users of networked computers interact [2]”. Current metaverse-based services emphasize social activities among users and highlight the ability to do things that are impossible in the real world. However, because existing services have difficulty finding connections between the virtual and real worlds, the metaverse is not well-utilized in our daily lives.

In this paper, we propose a metaverse-based service that can be utilized in the real world: a Mixed Reality Service with Real-time Interaction of Virtual and Real Worlds. We also propose three key elements required to implement the service: location, physics engine, and real-time interaction. Based on an MR

This work was supported by the GRRC program of Gyeonggi province. [GRRC2017-B05, The development of intelligent service platform converged with cm-level location information]

service model [1] and the elements, we show an example of the implemented MR service using Unreal Engine and Cesium. This service allows users to store information such as lifelogs created in the real world in 3D virtual spaces in real-time, and search for the information in 3D virtual spaces based on the location of the real world.

2. Related Research

2.1 Second life

Second Life is an early metaverse-based service. Users can create and place virtual objects with special features in dedicated spaces called "sims". Users can customize their avatars to express individuality. Also, they can participate in economic activities using a virtual currency called Linden Dollar. Although interactions between users and spatial expansions are flexible, the service requires significant computer hardware and network resources.



Fig. 1. Second Life

2.2 Roblox

Roblox is a popular metaverse-based service with over 50 million daily active users in 2022 [3]. For example, Roblox users can engage in activities that are difficult to achieve in the real world. For example, a famous American rapper, Lil Nas X, held a concert in Roblox, attracting an audience of 33 million [4]. The users can create their games with their own rules using "Roblox Studio" and play them with other users. They can also create virtual items and trade them using a virtual currency called Robux.



Fig. 2. Roblox

3. Proposal of a Mixed Reality Service with Real-time Interaction between Virtual and Real Worlds

Although the metaverse is gaining attention, it is not widely used in our daily lives. For example, in Roblox, users can trade items to earn Robux and exchange them for cash, but it is difficult to find other applications in daily life. In other words, metaverse-based services are mainly perceived as games where users can spend or earn money. This is because it is difficult to find a connection between the metaverse and reality.

On the other hand, one existing study for creating and utilizing virtual information in the real world is the "Multi-layer Mixed Reality Service Model based on Real Space Location [1]". This study defines Mixed Reality as a space where virtual information interacts with real-world objects. The "space" here refers to both real and virtual spaces.

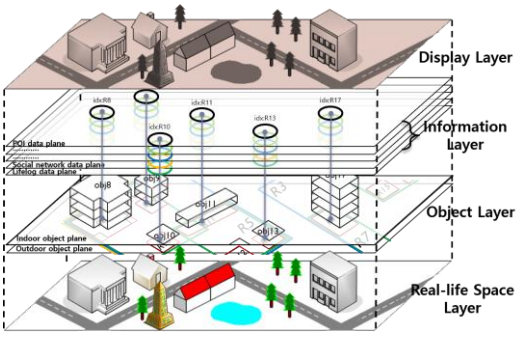


Fig. 3. A Multi-layer Mixed Reality Service Model Based on Real Space Location

In this study, the real world's "location" is used to organize real-world objects and related virtual information. It also searches for virtual information based on real-world objects. In other words, to implement Mixed Reality, it is necessary to recognize real-world objects. To achieve this, the study proposes using "location" and a digital map. This study can be seen as a way to create a connection between the metaverse and reality. There is a service based on this study: "Ambient Browser for Goyang City [5]".



Fig. 4. Ambient Browser for Goyang City

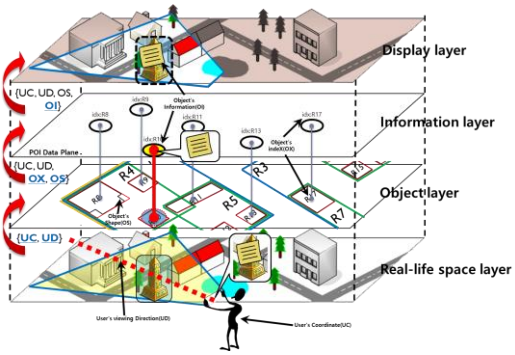


Fig. 5. Layers of Ambient Browser based on the MR Service Model

In this paper, we propose a "Mixed Reality Service with Real-time Interaction between Virtual and Real Worlds" based on the service model. We also propose three elements necessary for implementing the proposed service: location, physics engine, and real-time interaction. The "location" refers to the "real-world location." The "physics engine" means "implementing the physical constraints of the real world in the virtual space," for example, recognizing objects that are inaccessible in the user's real-world location as also inaccessible in the virtual space, thereby filtering out unnecessary information for the user. The "Real-time interaction" refers to "the real world and virtual space being interconnected in real-time." Based on these three elements and the service model, the proposed service creates a connection between the virtual and real worlds, allowing users to naturally generate and utilize virtual information in the real world.

4. Implementation of the Proposed Service Using Unreal Engine and Cesium

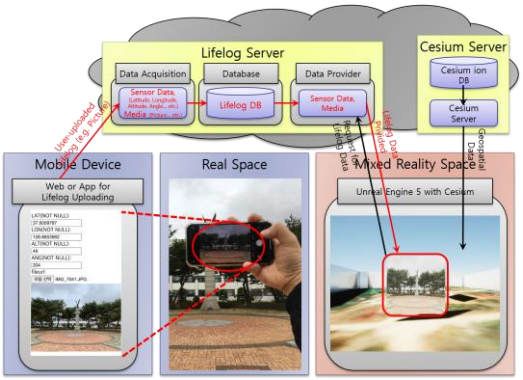


Fig. 6. Overview of Implemented Service

The main features of the service example presented in this paper are as follows: Users can upload photos taken in the real world to a 3D virtual space, and the photos will appear in the same location in the virtual space. Users can use an avatar to navigate a 3D virtual space that mimics the real world, and access these photos based on their location. The proposed service is implemented with the following server-client structure:

4.1 Lifelog Server

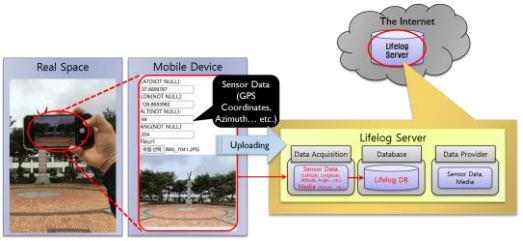


Fig. 7. Process of Uploading a Picture to Lifelog Server in Real-time

The lifelog server corresponds to the "Information Layer" in the service model. The server is implemented using web-based Node.js, and users can upload photos taken with their smartphones in real-time.

4.2 Cesium Server

The Cesium Server corresponds to the "Object Layer" in the service model. Cesium is a platform that provides global geospatial data, offering 2D and 3D maps and a plug-in compatible with Unreal Engine.

4.3 Client - Unreal Engine with Cesium and Custom Module

The client corresponds to the "Information Layer" in the service model. To implement the proposed service, the location relationship between the real world and the 3D virtual space must be consistent. To achieve this, we use Unreal Engine 5 to create a 3D virtual space that mimics the real world, utilizing the Cesium plug-in. This space was then used as the client.

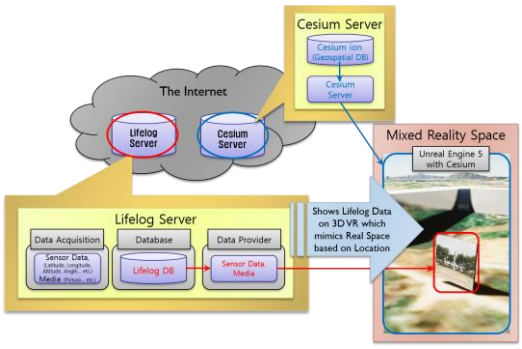


Fig. 8. Process of Displaying the Picture in the Service

Unreal Engine allows developers to create and implement desired features using C++. Based on this, we have developed a custom module that checks the lifelog server every 10 seconds for new photo uploads, and when new photos are uploaded, it retrieves them and places them in the corresponding location within the client, the 3D virtual space.

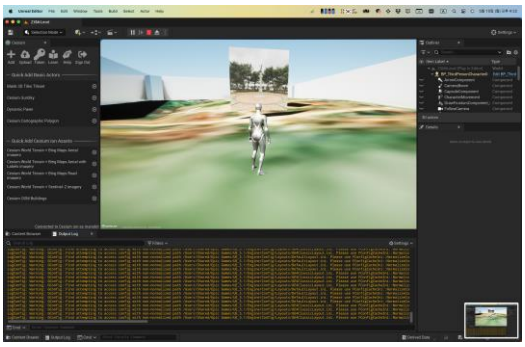


Fig. 9. Avatar in Implemented 3D Virtual Space

5. Conclusions

In this paper, we proposed a Mixed Reality service that can be utilized in the real world. We also proposed three key elements required for implementing the proposed service. We also demonstrated an implementation of the proposed service. Future research will enhance the implementation by adding lifelog upload features, such as GPS logs.

References

[1] Jong-gak Choe, & Yongjin Kwon, "Multi-layer Mixed Reality Service Model based on Real Space Location", *Graduate School of Korea Aerospace University*, Ph.D. Thesis, 2021.

[2] Oxford English Dictionary, https://www.oed.com/view/Entry/271922 (Accessed on May, 2023)

[3] Roblox, https://blog.roblox.com/2022/01/year-roblox-2021-data/ (Accessed on May, 2023)

[4] Wacom Blog of Korea, https://wacomkoreablog.com/1490 (Accessed on May, 2023)

[5] Jong-gak Choe, & Yongjin Kwon, "Outdoor Mixed Reality Model based on Location and Its Application", *KSII The 14th Asia Pacific International Conference on Information Science and Technology 2019*, pp. 107-109, 2019.

KSII The 18th Asia Pacific International Conference on Information Science and Technology(APIC-IST) 2023.
Copyright ⓒ 2023 KSII

An Analysis on the Impact of Consumer Needs on Consumer Behavior: A Case Study based on Design of a Smart Wearable Product under E-Commerce

Jialei Ye[1], Xiaoyou He[2] and Ziyang Liu[3*]
1 Design Postdoctoral Research Station, Nanjing University of the Arts, 210013, Nanjing, China
[e-mail: b314666@nua.edu.cn]
2 School of Industrial Design, Nanjing University of the Arts, Nanjing 210013, China
[e-mail: 532102483@qq.com]
3 Kyonggi University, Suwon-si, Gyeonggi-do, 16227, South Korea
[e-mail: morninglzy@hotmail.com]
*Corresponding author: Ziyang Liu

Abstract

This paper focuses on the correlation between consumer demand-oriented product design and consumer behavior under E-commerce. Since its emergence, the consumption of smart wearable products has been growing rapidly, so the representative smart bracelet is set as the example in this study. By classifying and organizing consumer needs from the perspective of product design, this paper expects to prove that consumers' consumption behavior can be promoted through design of products on the basis of consumers' needs in E-commerce. This study uses three methods: big data semantic analysis, KANO model analysis, and satisfaction analysis. The research materials were collected through big data search and questionnaire. The final results show that: 1. There is a correlation between consumer demand and consumer behavior. 2. There are categorial differences in consumers' needs. 3. To improve consumer satisfaction has positive influence on consumer behavior. This paper also proposes a method of interpreting and designing with consumer needs as the core and summarizes a design strategy that exploits consumer needs to guide consumer behavior in E-commerce. In future research, we can try to use different data analysis methods to compare and study the transformation of consumer demand and the influencing factors therein, so as to improve the research on the influencing factors between consumer demand and consumer behavior under E-commerce.

Keywords: Product design in E-commerce, consumer demand, consumption behavior, smart wearable products

1. Introduction

This study takes smart wearable products as a case to study on consumer demand in E-commerce and their intention and attitudes. Product design is the directed study of industrial consumer products by art and design professionals, who seek to increase the added value of products through product appearance, service systems and processes, and interaction methods, etc. This study analyzes consumer demand from the perspective of product design and explores the relationship between it and consumers' attitudes toward consumption. Sustained consumption behavior in E-commerce is a definition of consumption behavior from the perspective of development. Consumer behavior is a combination of consumers' mental activities such as the psychology of needs, purchase motivation, and willingness to consume, and real behavior.

2. Research Background and Methodology

Design is an effective tool to improve competitiveness in business competition. There is a close connection between product design and consumer behavior, with the help of product design consumers gradually identify products, evaluate products, and consume products. Continuous attention to consumer needs from the designer's perspective is a reliable way to promote sustained consumption behavior in E-commerce.

TF-IDF is a common weighting technique used in information retrieval and data mining during semantic analysis of big data, and the resulting values can be used to rate the importance of words. This study used the big data mining platform TEXTOM for semantic analysis to obtain the TF-IDF core vocabulary. And the information needed for the study was obtained by sorting and summarizing the vocabulary.

KANO model analysis is a structured questionnaire and analysis method developed by Noriaki Kano based on the principle of segmentation of customer needs. KANO model is a typical qualitative analysis model, and the analysis method is mainly used to identify user acceptance of fuctions of products, to help companies understand user needs at different levels, to identify the contact points between customers and companies, and to identify the most important factors that can satisfy customers. The features of the products were classified into five categories in the analysis: Must-be Attribute, One-dimensional Attribute, Attractive Attribute, Indifferent Attribute, and Reverse Attribute. Based on the nature of five attributes, the product attributes are linked to their respective groups, as shown in **Fig. 1**.

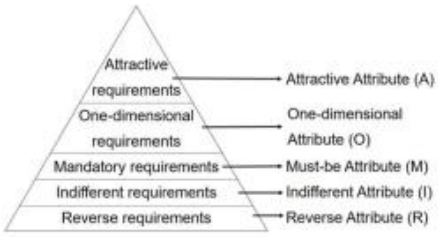


Fig. 1. Classification of Attributes

3. Materials and Methods

3.1 Semantic Analysis

We collect the TF-IDF core vocabulary centered on "smart wearable products in E-commerce" through text mining, and then summarize and interpret them. The data was collected for the period from October 1, 2020 to October 1, 2022 using the keyword "smart wearable products in E-commerce". The vocabulary sorted out from the perspective of consumer demand and user demand. With the guidance of design theories, this paper classified the vocabulary as shown in **Table 1**.

Table 1. Classification of Demands and Definitions

Variable	Definition	Vocabulary	Bibliography
Sensory Requirements	To categorize the ways in which product information is communicated. Mainly: visual, auditory and tactile.	Image, reminders, experience, sound	Ou, J., Tan, Y, Fang, J.Y., Kang, Q.C., Xu, Y. (2022).
Appearance Identifiability Requirements	To evaluate the product from two perspectives: identifiability of the product's appearance and features and users' trust to use.	Exercise, health, quality, style, medical, portability, life	Li, A. (2007).
Functional Service Requirements	The functional services of the product are evaluated from four perspectives: ease of decision making, ease of access, ease of benefit, and ease of maintenance of the product.	Heart rate, data, monitoring, analysis, operation, price, service, function, privacy	Berry, L.L., Seiders, K., & Grewal, D. (2002).

3.2 KANO Model Analysis

In this section, it is hoped that the KANO model will guide the typology of consumer needs in E-commerce and guide the attention and input to consumer needs when designing according the attribute classification of different consumer needs. The questionnaire method was used in this study. The questionnaire consists of 36 questions in total, and is composed of 2 parts: the first part is the basic demographic data of the respondents; the second part investigates the consumers' requirement，as shown in **Table 2**.

Table 2. Evaluation Objectives

Requirement	Objectives	Content
Sensory Requirements	Vision	D1 Look for Information
	Hearing	D2 Listen for Message
	Touch	D3 Vibrating Alert
Appearance Identifiability Requirements	Convenience	D4 Easy to Wear, D5 Compact Size
	Product CMF	D6 Fashionable Color, D7 Lightweight Material, D8 Sturdy Structure
	Product Shape	D9 Simple Shape
Functional Service Requirements	Function	D10 Real-Time Monitoring, D11 Information Prompts, D12 Ease to Recharge
	Interactive Mode	D13 Operating Mode, D14 Voice Control, D15 Personalized Settings
	Privacy Protection	D16 Data Security Management

We calculate the demand index coefficients, the results are shown in **Table 3**. To draw a scatter plot can present the indices in Table 8 more visually as shown in **Fig. 2**.

Table 3. Analysis of Consumer Satisfaction with Smart Bracelets in E-commerce

Categories	No.	SI	DSI	Optimized Classification
Sensory Requirements	D1	0.147	0.480	M
	D2	0.235	0.559	M
	D3	0.265	0.529	M
Appearance Identifiability Requirements	D4	0.235	0.529	M
	D5	0.578	0.608	O
	D6	0.208	0.188	I
	D7	0.854	0.537	O
	D8	0.196	0.516	M
	D9	0.126	0.116	I
Functional Service Requirements	D10	0.168	0.442	M
	D11	0.167	0.521	M
	D12	0.113	0.083	I
	D13	0.405	0.810	O
	D14	0.529	0.088	A
	D15	0.612	0.102	A
	D16	0.511	0.170	A

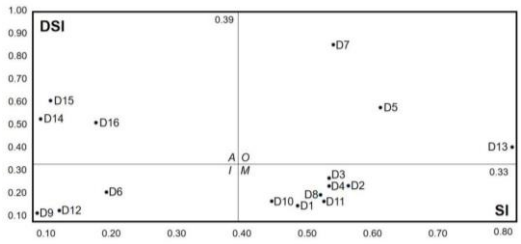


Fig. 2. SI-DSI Requirements Satisfaction Scatter Plot

4. Discussion

4.1 Summarizing the Results

This study has scientifically classified the consumer requirements in E-commerce of smart bracelets, as shown in **Table 4**. A comprehensive analysis of them shows that there is a categorial relation among consumers' requirements in E-commerce for smart bracelets, and different classification indicators will have different effects on consumers' consumption attitudes.

Table 4. Requirements Classification Table

Requirements Classification	Attributes of the Requirements	Items
Attractive Requirements	A	Voice control, personalized settings, data security control
One-dimensional Requirements	O	Small size, light material, operating and control mode
Mandatory Requirements	M	Information observation, listening to information, vibration alert, easy to wear, sturdy structure, real-time monitoring, message alert
Indifferent requirements	I	Fashionable color, simple shape, convenient to recharge

4.2 Consumer Requirements

Consumer requirements studied in this research are divided into four main groups, and different methods of design are used for different types of requirements to promote consumers make continuous consumption behavior under E-commerce. Attractive requirements guide product design to bring differences and attract the attention of new users. One-dimensional requirements-oriented products design can maintain user loyalty. To optimize the product based on mandatory requirements, so that the product conforms to the trend of mainstream. Indifferent requirements can help products become more recognizable.

4.3 Design Strategy

The strategy is to focus on consumer needs, and scientifically analyze data to guide product design in E-commerce, so to achieve sustainable consumption of products. Consumer needs have a direct impact on consumer behavior, so the criteria for evaluating the excellence of products in E-commerce should be based on consumer needs.

The paper confirms that consumer needs in E-commerce have an important influence on consumption behavior. To guide the direction of the development of consumer requirements in E-commerce by means of design can facilitate the formation of good and sustainable consumer behavior. Designing around consumer needs leads consumers to create new needs and develop expectations for new products [70].

5. Conclusions

This study is a case study of smart wearable products in E-commerce, which addresses the relationship between consumer demand and consumer behavior from the perspective of product design. The study developed design methods for smart wear products in respond to the five classifications of consumer requirements, giving full play to the value of the analysis results. Finally, a product design strategy under E-commerce centered on consumer needs is summarized and summarized to improve consumers' recognition of the product through design and promote sustained consumption behavior.A variety of data analysis techniques were used in this study to analyze consumer demand in E-commerce and to confirm the link between it and product design and consumer behavior.

Copyright ⓒ 2023 KSII

Theorizing User's Attitudes Towards AI-Based Systems by Using Philosophical Reasonings

Jong Yoon Won and Kun Chang Lee*
SKK Business School, Sungkyunkwan University
Seoul, South Korea
[e-mail: yoonjbest1@gmail.com, kunchanglee@gamil.com]
*Corresponding author: Kun Chang Lee

Abstract

In this study, we utilize philosophical justification to imagine how people feel about AIBS (AI-Based Systems). It takes a philosophical look at how AI technologies have changed our lives and how Heidegger's ideas have helped to shed light on previously obscured areas. The issue at hand is tackled by investigating the impact of AI on user choice. Humanity can be preserved and advanced through human-AI contact, which is why the proposed framework uses the 4O classification (Overabundance, Opportunity, Obstruction, Objection) to look for philosophical alternatives. Theorizing is the process by which researchers examine and analyze the social implications of new AI technologies. This work seeks to provide a comprehensive knowledge of the topic by combining philosophical and AI-related discourse.

Keywords: IS Philosophy, Theorizing, AI-Based Systems (AIBS), AI Technology

1. Introduction

Heidegger's view of technology is that it is always connected to human existence because it has a big effect on our lives and is part of who we are [1]. Technology is a bridge between humans and the rest of the world. It affects how we act, think, and see the world, which in turn affects our existence. But he warned that the progress of technology does not mean that people will be free. Instead, he said that technology could make us forget what makes us human, limit our freedom, and push us toward a reckless pursuit of efficiency and calculation [3]. So, in some interpretations, Heidegger is sometimes seen as someone who is against technology. But, despite what people think, he did not completely turn down technology [4]. Instead, he wanted to show us how we should think about and use technology. He suggested ways to respect human nature and values while avoiding over-reliance on technology and mindless pursuit of it. Heidegger's main goal was to figure out what technology is all about in a way that "does not force us to blindly move forward with technology or, what amounts to the same thing, to rebel against it" [5].

As artificial intelligence (AI) technology has gotten better, services and smart devices with AI-based systems or AIBS have become more common. AIBS-powered apps like Google Maps, self-driving cars, automated marketing, e-commerce recommendation systems, and anti-fraud detection, for example, have made our lives a lot easier. But AI, which mimics human intelligence by using human responses, choices, and preferences, can appeal to and change the inner parts of people, such as their thoughts and feelings. From Heidegger's point of view, it is important to think about how much AI technology is a part of our daily lives and how it affects how we act and how we relate to each

This paper was supported by SKKU Excellence in Research Award Research Fund, Sungkyunkwan University, 2022.

other. Even though Heidegger's idea of "parting" was first mostly associated with the industrial realm, it is important to realize that the current form of "parting," where AI technology reigns, has control over humans not only in the industrial realm but also in their private lives and leisure activities. For example, when we use the internet, do we really think of ourselves as active users, or do we think the internet is watching and influencing us? So, we have to ask philosophical questions about whether we are being taken over by Ge-stell and absorbed into the framework of this technology. In our research, we want to find out how people feel about AIBS by using philosophical reasoning and trying to find answers to these basic questions.

2. Theorizing User's Attitudes Towards AIBS

Even though there is a lot of literature on theory and how theories are made in the human sciences, there are still different views on theory in IS-related fields [7]. In order to deal with these problems, we want to look at the conversation about social issues caused by the rise of AI technology from the point of view of the IS field.

2.1 Foundational practices

Heidegger's idea of "revealing" helps us learn more about how AI technology is hidden or hard to see. It also helps us figure out what we should pay attention to. As AI technology gets better, we often do not notice that it is making our lives easier and more efficient to the point where we forget that we are letting ourselves be controlled and turned into mere consumers. So, we need to ask, "What are the mechanisms behind the ease of use that AI technology provides? What kinds of materials make up these machines, and how much do they affect our lives?" You should also think about these things. By learning about these hidden facts, we can bring to light the hidden problems and make our relationship with technology and the rest of the world more clear. We are aware of the problem that AIBS is constantly watching us through smart devices and services, and our goal is to bring this issue to light in a clear way. We suggest the 4O framework, which is made up of Overabundance, Opportunity, Obstruction, and Objection, as a way to reach this goal.

Aristotle's concept of Phronesis

Phronesis is closely connected to human behavior and moral judgment. It is the ability to judge and act in a morally good way in the context of human life [8]. Phronesis refers to a person's ability or wisdom to help them make ethical decisions and act on them, based on the practical wisdom they have gained from experience and real-life situations. In our framework, we use Phronesis to figure out what it means to have too much. This means we have to be wise enough to see how AIBS is everywhere in our daily lives and figure out what it means for us as a whole.

Heidegger's concept of Challenging Forth

The term "Challenging Forth" is used to describe the difficulties that arise whenever humans interact with technology [4]. Heidegger argues that these difficulties threaten our very existence because of the weight they carry. Heidegger did not see technology in a wholly negative light [9], despite the fact that it often dominates our daily lives and obscures our values and purpose. He thought that the technological obstacles we face are actually opportunities for expansion, creativity, and the maturation of our species. In light of this, we build our methodology around the idea of Opportunity. A user's perspective on AIBS is encapsulated in the term "opportunity," which includes the possibility of exciting new adventures, amusement, novelty, and benefits.

Heidegger's concept of Gestell

Gestell describes how technology imposes a particular framework or structure on our lives and the world [10]. Heidegger holds a critical view of Gestell, emphasizing how it can lead to our over-reliance on technology, causing us to forget our own essence and surrender our freedom and creativity [4]. Heidegger contends that it is essential for humans to preserve their humanity and overcome their reliance on technology by critically recognizing the impact of Gestell. In our framework, we include the concept of Objection, which represents a form of resistance to excessive AI technology and can be viewed as a component of this resistance.

Heidegger's concept of Geschick

According to Heidegger, the term "Geschick" describes the predetermined course that humanity will take in the future. To him, technology governs and controls human existence, so it is crucial that we realize and develop our own destinies [11]. From this vantage point, we conceive of Obstruction, which investigates the challenges posed by AIBS when it comes to the possibility of technological advancements dictating human destinies.

2.2 Generative practices

The role of generative practices in encouraging and advancing discourse is crucial. To be more precise, we investigate the social framework, tensions, and inconsistencies surrounding the problem of "AIBS constantly following us." It is common knowledge that major tech firms have grown rapidly thanks to AI and provided us with a wealth of free services. To that end, we plan to investigate whether or not users and these businesses engage in fair transactional dynamics. In particular, we lay out a structure for the 4O idea by conducting empirical research into the impact that unveiling AI's underpinning practices has on users' ability to make informed decisions.

3. Conclusions

This paper explores how philosophical justification can be applied to understanding user's atttidues towards AIBS. By including philosophical considerations, we hope to increase user's insight into and appreciation for technology's complex role in human life. With this method, we can examine the profound effects of AI technology on our daily lives and the underlying problems that arise from it from a philosophical perspective.

To this end, we develop a conceptualization of users' attitudes by drawing on Heidegger's concept of "revealing" to expose previously hidden facets of AI technology. This stress on serious thought and heightened awareness of technology's implications is crucial. We also consider the potential restrictions on individuality and originality that technology may impose, drawing on Heidegger's concept of "Gestell."

As an added bonus, this paper delves into the conceptualization of the "problems that AI continues to haunt us," which includes social structures, conflicts, and contradictory aspects. We conduct an empirical investigation into how users' choices are affected by analyzing the underlying practices that AI technology reveals. The proposed framework, built on the 4O idea, helps us understand how people make choices when confronted with the revealed facets of AI technology.

By incorporating these philosophical perspectives, we gain a more thorough comprehension of the problems caused by our interactions with AI technology and can more effectively pursue philosophical alternatives and potential solutions. This study aims to provide a road map for resolving issues related to AI technology through a scientific investigation of the philosophical factors influencing users' decisions and attitudes. In the end, it provokes philosophical consideration of the principles that ought to underpin our use of AI technologies, with an eye toward retaining and enhancing our humanity.

References

[1] D. Ihde, "Has the philosophy of technology arrived? A state-of-the-art review." *Philosophy of Science*, Vol 71, No. 1, pp 117-131, 2004.

[2] G. B. Smith, "Heidegger, technology and postmodernity." *The Social Science Journal*, Vol 28, No. 3, pp. 369-389, 1991.

[3] C. Mitcham, "What is the Philosophy of Technology?" *International Philosophical Quarterly*, Vol 25, No. 1, pp. 73-88, 1985.

[4] D. I. Waddington, "A field guide to Heidegger: Understanding 'the question concerning technology'." *Educational Philosophy and Theory*, Vol 37, No. 4, pp. 567-583, 2005.

[5] H. L. Dreyfus & C. Spinosa, "Further reflections on Heidegger, technology, and the everyday." *Bulletin of Science, Technology & Society*, Vol 23, No. 5, pp. 339-349, 2003.

[6] Y. C. Yang, H. Ying, Y. Jin, H. K. Cheng, & T. P. Liang, "Special issue editorial: Information systems research in the age of

smart services." *Journal of the Association for Information Systems*, Vol 22, Vo. 3, 2021.

[7] N. R. Hassan, L. Mathiassen, & P. B. Lowry, "The process of information systems theorizing as a discursive practice." *Journal of Information Technology*, Vol 34, No. 3, pp. 198-220, 2019.

[8] O. Eikeland, "Phronesis, Aristotle, and action research." *International Journal of Action Research*, Vol 2, No.1, pp. 5-53, 2006.

[9] J. Zwier, & V. Blok, "Saving Earth: Encountering Heidegger's Philosophy of Technology in the Anthropocene." *Techne: Research in Philosophy & Technology*, Vol 21, 2017.

[10] M. Blitz, "Understanding Heidegger on technology." *The New Atlantis*, pp. 63-80, 2014.

[11] H. D. Bahr, "Das Wesen der Technik und das, andere Geschick ': Ge-stell und Gegnet im Denken Martin Heideggers." *Heidegger Studies*, Vol 29, pp. 89-120, 2013.

Copyright ⓒ 2023 KSII

A Web-based Business Process Discovery System

Thanh-Hai Nguyen[1], Sohyang Bak[1], Kyoung-Sook Kim[2], Dinh-Lam Pham[3], and Kwanghoon Pio Kim[1*]

[1]Division of AI Computer Science and Engineering, Kyonggi University, Republic of Korea
[e-mail: {nguyenthanhhai, sa7291, kwang}@kgu.ac.kr]
[2]Contents Convergence Software Research Institute, Kyonggi University, Republic of Korea
[e-mail: khmjmc@kgu.ac.kr]
[3]Information Technology Institute, Vietnam National University, 144 Xuan Thuy Street,Cau Giay, Hanoi, Vietnam
[e-mail: phamdinhlam @vnu.edu.vn]
*Corresponding Author: Kwanghoon Pio Kim

Abstract

The Structure Information Control Net (SICN) mining framework is powerful for representing business process models in workflow mining or rediscovery. This paper presents a web-based system using the SICN-oriented process mining framework along with open-source platforms such as Django, Graphviz, and Cytoscape to rediscover and visualize business process models. The discovered SICN-oriented process models from the process instances are extracted from the IEEE XES-formatted process enactment event logs dataset by ρ(rho)-algorithm and visualized by transfer data function from this model to a JSON or DOT formatted file that is used respectively by Cytoscape or Graphviz. The business process discovery system leverages the powerful Django framework to build a friendly web interface providing a clear, concise, modern, and interactive view of the business processes rediscovered. The development of the system can help readers, students, and practitioners interested in process mining access and use a completely free process miner to learn how to rediscover and visualize a process model from event logs.

Keywords: Business process discovery, Web-based application, Process mining

1. Introduction

In today's Industry 4.0, Organizations must continuously innovate and optimize their business processes to stay competitive in the rapidly evolving business environment. Business process discovery and mining tools have emerged as essential systems support for organizations to achieve these goals [6]. Business process mining miners are software systems that leverage process mining techniques to extract data from event logs generated from process-aware information systems [8]. Process discovery tools allow discovery of hidden patterns and complicated relationships in the workflow log to apprehend more information about what is happening in the organization [7], lead to increase transparency, reduce costs, and improve organizational performance.

In the business process discovery, more and more tools support different aspects of business operations, such as ProM, BPMN.io, Celonis, Disco, etc. Business process discovery engines use techniques and algorithms for process mining, such as Alpha Miner, Inductive Miner, and Heuristic Miner. These tools use modeling languages to represent process models, among which Petri net, BPMN, or Process Tree, or DFG [5], is one of the most well-known modeling languages used in modeling business processes. Besides Petri net, the Information Control Net (ICN) model [1] , along with the ρ-algorithm, is a popular approach to discovering and rediscovering business process patterns that enable the representation and analysis effectively of simple to complex workflows. This paper proposes a web-based business process discovery system that supports the mining process model

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT, Ministry of Science and ICT), Republic of Korea (Grant No. NRF-2022R1A2C2093002).

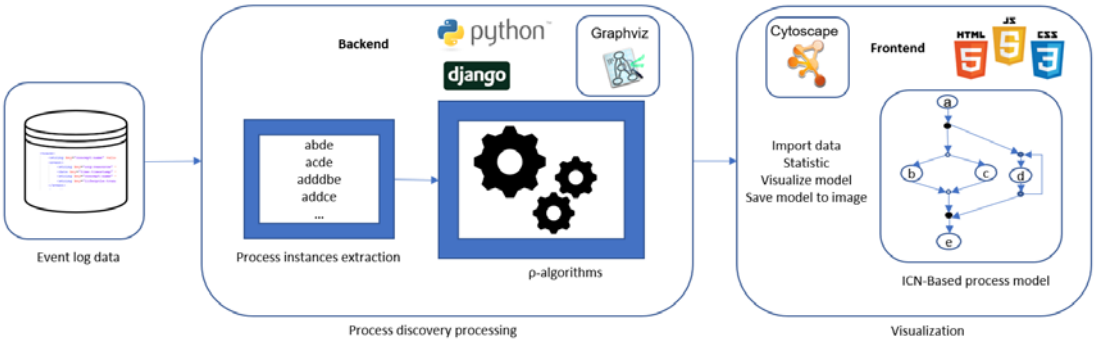


Fig. 1. Architecture system

based on SICN-oriented process mining framework [3]. We deployed this system using Django Python web framework [10], Graphviz [11], and Cytoscape [12] to visualize the SICN-oriented process models.

2. System architecture

Our system, illustrated in **Fig. 1**, consists of two main components. The first component is the backend which extracts information from the event log, then implements the ρ-algorithm to build a SICN-oriented process model. The second component is the frontend performs modeling of process models and statistical results, and data analysis.

In the first component, the application backend is built using Django, a powerful web framework that handles requests, database operations, and server-side logic. To allow users to upload XES-IEEE Formatted file event log data, we have integrated the Django web framework with a file upload library, providing safe and efficient file handling. Python programming language is used to programmatically build a SICN-oriented process model using the ρ-algorithm [3] and then convert the data to a JSON and DOT data file for Cystocape and Graphviz to generate the respective model as a graph.

The second component, the application's user interface, is built using HTML, CSS, and JavaScript, providing a visualization and analysis view of the discovered process model, allowing users to explore the SICN-oriented in various ways, such as drag and drop nodes and highlighting specific nodes or edges. The graph is displayed in the HTML canvas element, allowing for smooth and responsive interaction. The process model is also visualization as a static image that can be downloaded in SVG format or as a single DOT language file.

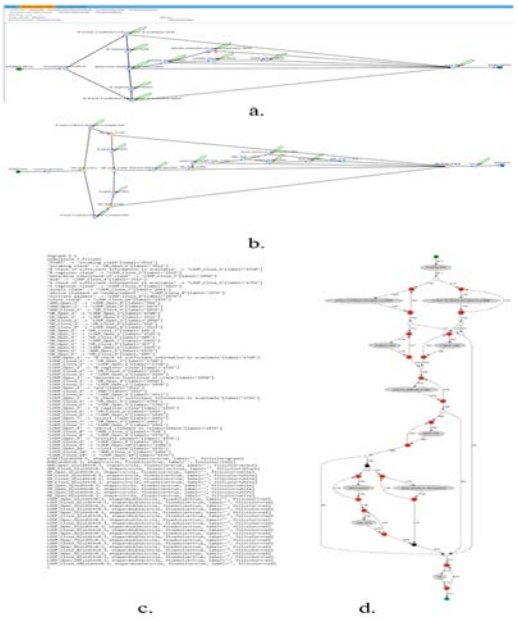


Fig. 2. The discovered SICN-oriented process model

3. Results and visualization

To describe the results and evaluate the system's effectiveness, we used the publicly available Teleclaims dataset [9]. This dataset contains information about an insurance company's claims handling process and is commonly used in process mining research. The

dataset contains more than 46138 events from over 3512 instances, each containing attributes such as case ID, activity name, timestamp, and resource ID.

3.1 Discover SICN-oriented Process model

The SICN-oriented process model is the primary output of our system and represents the flow of information in a business process. **Fig. 2** shows the SICN-oriented process model generated by our system. The obtained SICN-oriented process model provided a detailed view of the relationship between the activities regarding the process helping reveal some key insights into the TELECLAIMS process. We can easily observe the flow of the process and identify several activities that are closely related and seem to be very important to the process, such as these activities "check if sufficient information is available" at two facilities B and S. Or, by observe AND-gate Transition in the generated graph, we can easily realize the activities "initiate payment activities"; "close claim"; "advise claim" on the process be made simultaneously through in the process.

3.2 Discover Affiliate Network Model

The system also generates the Affiliate network [2], providing detailed information on the relationships between the performers and the activities, roles or groups, and organizations in which they participate. This network model help identify potential areas for cooperation or improvement of business workflow. **Fig. 3** shows the results of the Perfomers-Activities Affiliate Network discovered by our system from the Teleclaims process event log dataset.

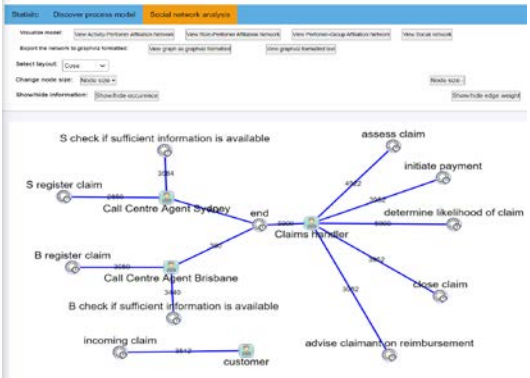


Fig. 3. The discovered Performer-Activity Affiliate Network model

4. Conclusions

In this research, we presented a web application that leverages the SICN-oriented business processes framework with the support of Django, Cystoscape, and Graphviz platforms to discover business processes model from event logs. Our testing on data sets shows that the model works well, achieving accuracy, relevance, and simplicity. We plan to expand the system's functionality to incorporate deep learning models to perform workload predictions in resource allocation and predict business operating costs.

References

[1] Kim, K. and Ellis, C. A., ICN-based Workflow Model and its Advance. In Handbook of Research on Business Process Modeling, pages 142–171. IGI Global, 2009.

[2] Kim, K. P., Discovering activity-performer affiliation knowledge on icn-based workflow models. J. Inf. Sci. Eng., 29:79–97, 2013.

[3] Kim, K.-S., et al., ρ-Algorithm: A SICN-Oriented process mining framework. IEEE Access, 9:139852–139875, 2021.

[4] Kim, K.-S., et al, Experimental verificatio.n and validation of the SICNoriented process mining algorithm and system. Journal of King Saud University - Computer and Information Sciences, 34(10, Part B):9793–9813, 2022.

[5] Panagiotis, D. and Apostolou, D., On the selection of process mining tools. Electronics, 10:451, 2021.

[6] van der Aalst et al., Workflow mining: discovering process models from event logs. IEEE Transactions on Knowledge and Data Engineering, 16(9):1128–1142, 2004

[7] van der Aalst, W. M. P. Process mining: A 360 degree overview. In Process Mining Handbook, pages 3–34. Springer International Publishing, Cham, 2022.

[8] van Dongen, et al. The ProM framework: A new era in process mining tool support. In Ciardo, G. and Darondeau, P., editors, Applications and Theory of Petri Nets 2005, pages 444–454.

[9] Teleclaim dataset, can be found in chapter 8.zip file download form website locate at: https://processmining.org/oldversion/files/chapter 8.zip

[10] https://www.djangoproject.com/

[11] https://graphviz.org/

[12] https://js.cytoscape.org/

Copyright ⓒ 2023 KSII

Pine Wilt Disease Detection using Synthetic Data and Object Detection Techniques

Yonghoon Jung, Wonseop Shin and Sanghyun Seo*
Dept. of Imaging, The graduate school of Advanced Imaging Science, Chung-Ang University
Seoul, South Korea
School of Art and Technology, Chung-Ang University
Anseong, South Korea
[e-mail: dydgns2017@cau.ac.kr, wonseop218@gmail.com, sanghyun@cau.ac.kr]
*Corresponding author: Sanghyun Seo

Abstract

Pine Wilt Disease (PWD) is a significant concern in forest management due to the lack of current treatment options. Although recent research has aimed to detect PWD using a deep learning-based detection model, the acquisition of large-scale, high-quality datasets for deep learning training remains a challenging issue. In this paper, we propose a PWD detection pipeline that utilizes a method for generating synthetic data based on 3D rendering and deep learning-based object detection techniques. We used Unreal Engine 5 to create forest scenes containing suspect, infected, and dead trees due to PWD, and extracted images from these scenes. The extracted images were segmented into the size of the deep learning input and used to train the deep learning model. Experimental results showed that the model pre-trained on synthetic data and then fine-tuned on real data demonstrated improved performance compared to the model trained solely on real data.

Keywords: Pine Wilt Disease, Deep Learning, Object Detection, Synthetic Data

1. Introduction

Pine Wilt Disease (PWD) causes damage to forest ecosystems worldwide, especially in East Asia's pine forests, such as those in Korea, Japan, and China. It is a fatal disease that inflicts severe damage[1]. To address this, recent advances in deep learning technologies have led to the development of deep learning-based PWD detection studies considering the external characteristics of PWD. However, for object detection and image classification required for learning, large-scale training datasets are needed. While many public datasets exist, due to the domain's specificity, there are instances where it is necessary to collect new datasets or improve existing ones.

In this study, we aimed to detect PWD using YOLOv7[2], a state-of-the-art real-time object detection model, and generated synthetic data based on 3D rendering techniques to compensate for the lack of real data. In addition, we conducted learning strategy experiments for performance improvement using the collected real and synthetic data.

2. Methodology

As shown in Fig 1, we aimed to detect PWD using a deep learning model by collecting real and synthetic data. This section describes the methods and results of collecting real and synthetic data, as well as the training strategies used with the collected data.

2.1 PWD Data Collection

This study was carried out with the support of R&D Program for Forest Science Technology (Project No.2021338C10-2123-CD02) provided by Korea Forest Service (Korea Forestry Promotion Institute).

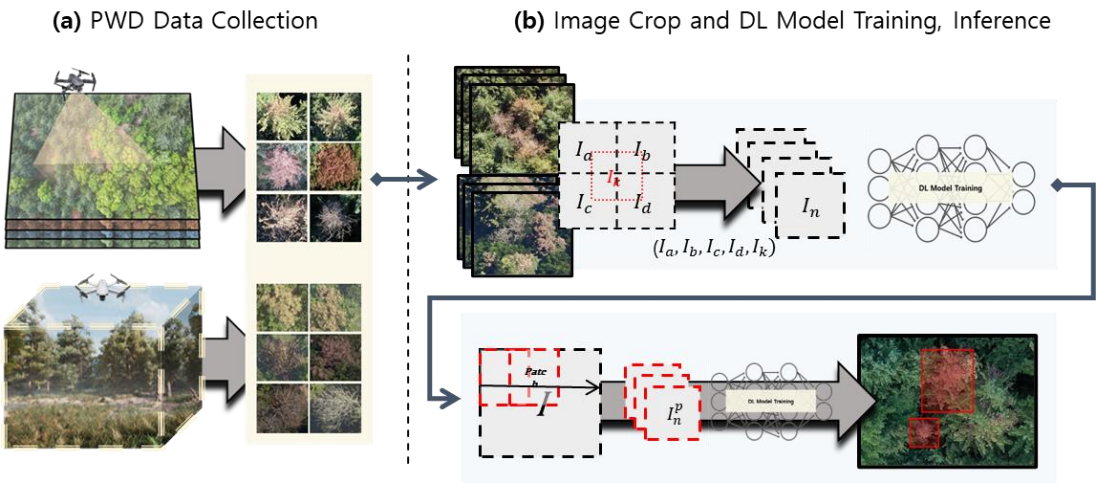


Fig. 2. Our PWD data collection and model training pipeline are as follows: (a) Process of collecting real and synthetic data using UAV (Unmanned Aerial Vehicles) and Unreal Engine 5. (b) After collecting the data, the images are sliced to create training data, and then the deep learning model is trained. In addition, during inference, a single source image is sliced, the results of each image are calculated, and then merged into a single image.



Fig. 2. shows the inferred results of the YOLOv7 model. The green line represents the ground truth, and the red line represents the prediction of the trained YOLOv7 model.

To collect real data, we utilized Unmanned Aerial Vehicles (UAVs) and extracted still images from video footage recorded over the period of August to October in Sangsa-ri, Hasa-ri, Gaegok-ri (Gapyeong County, Gyeonggi Province, South Korea), Jicheon-myeon (Chilgok County, Gyeongsangbuk-do, South Korea), Hadong County and Geochang County (Gyeongsangnam-do, South Korea). The size of the footage is 3840x2160, and the distance between the UAV and the pine tree object ranges from a minimum of 30m to a maximum of 50m.

For synthetic data, we created a pine tree model using Blender, a commercial 3D software. The created 3D model was imported into Unreal Engine 5 and positioned within a virtual space. Next, we performed material work on the tree objects in the virtual space, classifying the trees and creating materials according to the progression of PWD infection. Afterward, we used procedural foliage techniques in Unreal Engine to randomly distribute the trees. We then extracted synthetic data at a size of 3840x2160.

Approximately 600 images of real data and around 900 images of synthetic data were collected. The real data was manually labeled as suspected, infected, and dead trees according to the progression of PWD infection. For the synthetic data, automatic labeling was performed using the location information in Unreal Engine 5.

2.2 Training Strategy

Typically, strategies to improve performance involve either mixing synthetic data with real data or using synthetic data for pre-training followed by fine-tuning with real data [3, 4]. Based on this, we created three datasets: real, synthetic, and mixed. To verify the performance of each dataset, we used YOLOv7, one of the real-time object detection models. Moreover, we proceeded with the process of pre-training with synthetic data and fine-tuning with real data.

3. Experiment Results

For using the YOLOv7 model, we utilized an already implemented open-source library. The

model training was performed with a batch size of 32, a learning rate of 0.001, and 200 epochs, with other hyperparameters set the same as the original YOLOv7 network. The hardware used for training included 48 Intel Xeon (Cascadelake) 2.4Ghz processors and 8 GeForce RTX 2080 Ti GPUs.

For testing, we randomly split the real data into training and test sets at an 8:2 ratio. To verify the utility of synthetic data, all tests were conducted solely with the real data test set. Also, for the original image to be inputted into the deep learning model, both training and testing proceeded with images sliced into 640x640 sizes. During inference, a process of merging the images back together was conducted.

The evaluation metrics used were Precision, Recall, and mAP50. The learning results of each dataset are as shown in **Table 1**. The performance was relatively low when only synthetic data was used. However, mixed training showed improved performance compared to using only real data. Moreover, the best performance was shown in the case of *P*→*F*, where synthetic data was used for pre-training, followed by fine-tuning with real data. **Fig. 2** shows the result when inference was performed using the *P*→*F* model.

Table 1. Training results for each dataset, where *R* represents real data, *S* represents synthetic data, *R*+*S* represents a combination of real and synthetic data, and *P*→*F* represents training with synthetic data pretrained and fine-tuned with real data.

Method	***P**recision*	***R**ecall*	*mAP*^{50}
R	71.2%	70.4%	66.0%
S	66.3%	62.0%	60.4%
R+*S*	74.2%	70.5%	70.2%
P*→*F	**75.5%**	**74.2%**	**71.5%**

4. Conclusions

We conducted a process to detect PWD using real and synthetic data and aimed to improve model performance through our training strategy. Among these, the use of *P*→*F* showed improved performance compared to using only real data. This suggests the utility of our synthetic data creation technique for PWD. In future research, it may be necessary to consider rich annotation information and vegetation index when generating synthetic data. This is expected to show improved performance in detecting PWD. Also, it is essential to examine the statistical distribution of real and synthetic data and minimize the differences. To this end, we plan to study and apply domain adaptation techniques. We anticipate our approach will assist in constructing large-scale data necessary for deep learning training in the field of forestry.

References

[1] Ichihara, Y., Fukuda, K., and Suzuki, K. "Early symptom development and histological change associated with migration of Bursaphelenchus xylophilus in seeding tissues of Pinus thunbergii, ", *Plant Disease*, Vol 84, No. 6, pp. 675-680, 2000.

[2] Wang, C., Bochkovskiy, A., & Liao, H.M. "YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors," *ArXiv*, abs/2207.02696, 2022.

[3] Nowruzi, F.E., Kapoor, P., Kolhatkar, D., Hassanat, F.A., Laganière, R., & Rebut, J. "How much real data do we actually need: Analyzing object detection performance using synthetic and real data," *ArXiv*, abs/1907.07061, 2019.

[4] Erfanian Ebadi, S., Dhakad, S., Vishwakarma, S., Wang, C., Jhang, Y.C., Chociej, M., Crespi, A., Thaman, A., & Ganguly, S. "PSP-HDRI+: A Synthetic Dataset Generator for Pre-Training of Human-Centric Computer Vision Models, " ArXiv, abs/2207.05025, 2022.

Copyright ⓒ 2023 KSII

Automatic 3d Mesh Deformation Using Anime Character Image Information

Sanghyun Byun, Bumsoo Kim, and Sanghyun Seo*
Graduate School of Advanced Imaging Science, Multimedia & Film, Chung-Ang University
VIVE STUDIOS
College of Art and Technology, Chung-Ang University
[e-mail: egoist12276@cau.ac.kr, bskim@vivestudios.com, sanghyun@cau.ac.kr]
*Corresponding author: Sanghyun Seo

Abstract

In this paper, we propose a new framework for automatic transformation and creation of non-realistic 3D avatars in response to the growing demand for avatars with diverse body types in the metaverse. Previous studies have mainly focused on creating realistic human avatars, but this approach has limitations in creating avatars with unbalanced or unrealistic shapes. To address this issue, we introduce a new shape representation parameter and use a deep learning-based method to extract character body information and perform automatic template mesh transformation. The resulting outputs has been visually presented and evaluated by users to demonstrate the effectiveness of the proposed method. This approach provides an automatic mesh transformation method tailored to the non-photorealistic domain and extends the existing method to the creation of avatars with various body types. Overall, our work contributes to the advancement of avatar creation in the metaverse by enabling creation of non-realistic and diverse avatars in a more efficient and cost-effective manner.

Keywords: Digital Human Modeling, 3D Cartoon Stylized Avatar, Automatic Method, Metaverse application

1. Introduction

The increasing demand for diverse and distinctive avatars in multimedia, virtual reality, gaming, and the metaverse has highlighted the significance of digital human modelling technology. Despite the recent advancements in 3D human modeling techniques, the traditional methods still suffer from high initial investment costs, time-consuming manual work, and limited scalability. Moreover, most of the existing research has focused on creating realistic human-shaped avatars, leaving a gap in the development of non-realistic avatars with various shapes that are not necessarily human. Our work aims to improve these limitations by proposing an innovative framework to automatic transformation template meshes using a deep learning-based partial transformation technique. Research in this area primarily focuses on tracking and registration quality as well as specific visualizations in support of medical procedure.

Existing methods for 3D human modeling, such as full-body scanning [1-4] and deep learning-based technology [5-7], have made significant contributions to the field but have still exhibited several shortcomings. They mostly focus on reducing reconstruction errors and increase realism, leading to limited scalability when applied to non-realistic avatars with different shapes. Additionally, these techniques often perform batch transformation processes based on parameters, which restrict their ability to express detailed body proportions. Our research addresses these limitations by developing a general methodology for obtaining non-realistic human meshes and extending the approach to represent characters with imbalanced or non-human body shapes. In this paper, we propose a novel framework that extracts shape and proportion information of a character from a single image using deep learning-based methods and applies it to perform automatic transformation of human meshes.

Our approach targets non-realistic illustrated or game characters and performs non-rigid 3D transformations on an explicit template mesh input

This research was supported by Culture Technology R&D Program through the Korea Creative Content Agency grant funded by Ministry of Culture, Sports and Tourism in 2021 (A Specialist Training of Content Research and Development Based on Virtual Production) under Grant R2021040044, Contribution Rate: 100%)

to match the reference image of character's proportions and shape. We also demonstrate how our methodology can be easily applied to animation, opening new avenues for further research in creating avatars with diverse shapes and appearances, as well as developing more advanced animation and control methods. The key contributions of our study include the development of an extended approach for representation of characters with non-realistic or imbalanced body shapes, a framework capable of obtaining various 3D human meshes using an automated method requiring only a single image as input and pavement of the way for future research in creating diverse avatars and advancing animation and control techniques. Our proposed methodology addresses the limitations of traditional 3D human modeling processes, providing a practical and effective framework for reconstructing avatars with various shapes for a wide range of applications.

2. Method

We aim to obtain a 3D mesh that can represent characters with non-realistic shapes and proportions based on the input image and template mesh. In our methodology, we identify a series of steps including the process of defining parameters for extracting information from images, transfer learning and adjustment of the keypoint detector, and an automatic mesh processing workflow.

2.1 Define Character Parameter

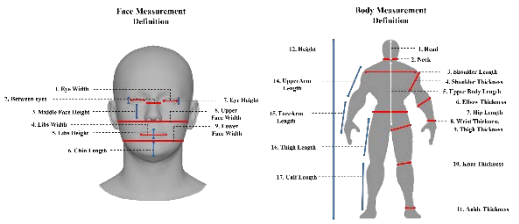


Fig. 1. Full body parameter definition

In this work, we aim to accurately define each part of a character based on body and facial keypoint formats. To achieve this, we use the MPII-TRB [8] keypoint format for defining body keypoints and select character samples from diverse media types, ensuring a variety of body types and proportions. High-quality images of each character are obtained, and specific landmarks are marked according to the MPII-TRB standard. For facial keypoints, we preprocess the dataset by cropping and resizing facial images, defining each part through 28 landmarks. This methodology enables consistent and accurate results in extracting information about human anatomy, providing valuable insights into body expression across various character types. **Fig. 1** depicts the defined parameters.

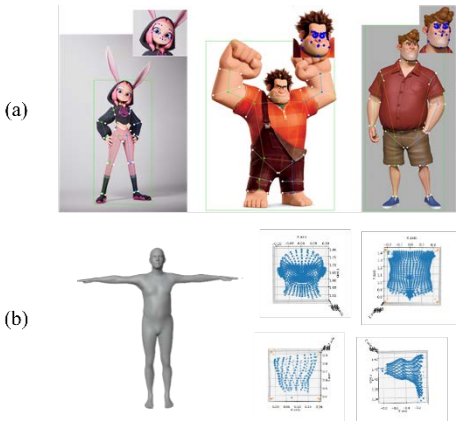


Fig. 2. Examples of keypoint detection and mesh segmentation

2.2 Transfer Learning and Fine Tuning

To obtain information from images using keypoint detector, transfer learning and fine-tuning of pre-trained detectors are required. We gather high-quality images of diverse 2D and 3D characters in various styles, poses, angles, and perspectives, along with annotation files containing keypoint coordinates. The pre-trained model for human pose estimation is fine-tuned on the character dataset, allowing it to learn the specific features of character body structures. This approach expands skeleton and contour keypoint estimation over existing methods, resulting in a powerful and highly scalable model for character pose and contour estimation, with practical applications in animation, gaming, and virtual reality industries. **Fig. 2 (a)** shows an example of the result of character body keypoint detection.

2.3 Automatic Mesh Processing

We automate mesh deformation by calculating Euclidean distances between each keypoint segmenting the mesh, and determining deformation points and weights. We calculate relative body proportions based on head length and apply this

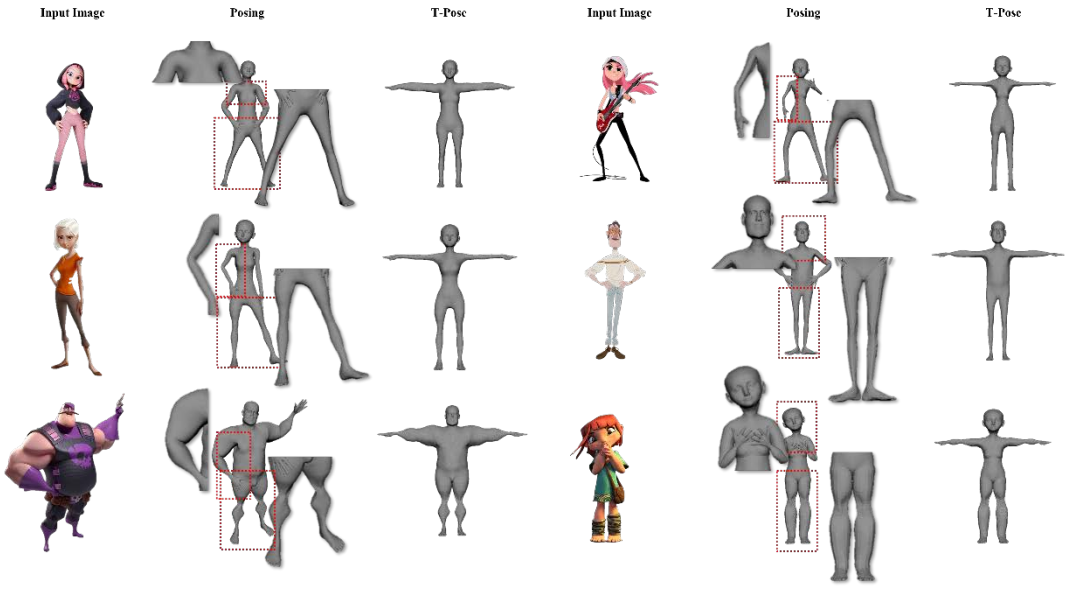


Fig. 3. The resulting 3D mesh obtained through our methodology.

approach to facial proportions. To partially deform the input 3D
template mesh, we use graph convolutional networks for automatic segmentation, and apply Free Form Deformation (FFD) [9] for mesh deformation. We calculate deformation weights and volumes for both body and face, enabling flexible and efficient control of the deformation process. Finally, we adopt FLAME (Faces Learned with an Articulated Model and Expressions) [10] head models and SMPL (Skinned Multi-Person Linear) [11] body models, which are mainly used in human modeling studies, as template meshes for deformation. We obtain an integrated human 3D model by aligning and transforming those two models. Fig 2. (b) shows SMPL template mesh and mesh segmentation examples.

3. Results

We have implemented the proposed methodology in a series of step-by-step processes, including character image body dataset collection and keypoint detection as well as human mesh partial deformation. Through this, we extracted keypoints from various input images of anime, game, and illustration characters, and calculated proportions and shape parameters. Subsequently, we were able to generate 3D meshes. Table 1 shows the accuracy results of keypoint detectors (face and body) where transfer learning was conducted. Also, **Fig. 3** shows the results derived from our methodology. As can be seen from the table, the keypoint detector has the accuracy to obtain information on the shape and ratio of characters in the image. In addition, the derived full-body 3D mesh can express the shape of the target character of the image well.

Table 1. Simple Results of Keypoint Detection

Body Kepoint Detector			Face Keypoint Detector
Arch	PCKh@0.1	PCKh@0.5	NME
ResNet-50	0.3468	0.8815	0.054
ResNet-152	0.3482	0.8856	0.052
SCNet-50	0.3353	0.8794	0.056
SCNet-101	0.3471	0.8832	0.053
HRNet-w32	0.3557	0.8885	0.054
HRNet-w48	0.3665	0.8902	0.049

4. Conclusions

In this paper, we presented a novel framework for automatically deforming 3D human avatars with varying shapes and proportions using deep learning-based keypoint estimation and 3D freeform deformation. This approach enables the creation of unique avatars, moving beyond traditional methods that focus on realistic human shapes. The reconstructed mesh simplifies the 3D character modeling process and enhances the

diversity of avatars, contributing to a more inclusive virtual environment. However, limitations include dependency on the template mesh and potential quality degradation due to keypoint extraction failures. Future work will focus on improving keypoint estimation accuracy, refining 3D deformation techniques, and developing methods to generate meshes from scratch using image-based information extraction.

References

[1] Tneb, Rainer, et al. "3D body scanning-systems, methods and applications for automatic interpretation of 3D surface anthropometrical data." *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*. Vol. 44. No. 38. Sage CA: Los Angeles, CA: Sage Publications, 2000.

[2] Ma, Yong-You, Hui Zhang, and Shou-Wei Jiang. "Realistic modeling and animation of human body based on scanned data." *Journal of Computer Science and Technology* 19.4: 529. 2004.

[3] Xiao, Yijun, and J. Siebert. "Building superquadric men from 3-D whole-body scan data." *4th IEEE Chapter Conference on Applied Cybernetics*. 2005.

[4] Taylor Andrew, Ertu. Unver. An experimental study to test a 3D laser scanner for body measurement and 3D virtual garment design in fashion education. *In Wearable futures—hybrid culture in the design and development of soft technology*, pp. 1–14. 2005.

[5] Saito, Shunsuke, et al. "PIFUHD: Multi-level pixel-aligned implicit function for high-resolution 3D human digitization." *in Proc. of* the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2020.

[6] Xiu, Yuliang, et al. "ICON: Implicit clothed humans obtained from normals." 2022 *in Proc. of* IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE. 2022.

[7] Zheng, Zerong, et al. "PAMIR: Parametric model-conditioned implicit representation for image-based human reconstruction." *in Proc. of* IEEE Transactions On Pattern Analysis and Machine Intelligence 44.6 : 3170–3184. 2021.

[8] Duan, Haodong, et al. "TRB: A novel triplet representation for understanding 2 D human body." *in Proc of.* the IEEE/CVF International Conference on Computer Vision. 2019.

[9] Sederberg, Thomas W., and Scott R. Parry. "Free -form deformation of solid geometric models." *in Proc. of* the 13th Annual Conference on Computer Graphics and Interactive Techniques. 1986.

[10] Li, Tianye, et al. "Learning a model of facial shape and expression from 4D scans." ACM Trans. Graph. 36.6 : 194:1 -194:17 . 2017

[11] Loper, Matthew, et al. "SMPL: A skinned multi -person linear model." ACM Transactions on Graphics (TOG) 34.6 : 1 - 16. 2015.

[12] Baek, Seung-Yeob, and Kunwoo Lee. "Parametric human body shape modeling framework for human-centered product design." Computer-Aided Design 44.1 : 56-67. 2012.

[13] Lee, WonSook, Jin Gu, and Nadia Magnenat- Thalmann. "Generating animatable 3D virtual humans from photographs." Computer Graphics Forum. Vol. 19. No. 3. Oxford, UK and Boston, USA: Blackwell Publishers Ltd. 2000.

[14] Wang, Charlie C.-L., et al. "Virtual human modeling from photographs for garment industry." Computer-Aided Design 35.6 : 577-589. 2003.

[15] Loper, Matthew, et al. "SMPL: A skinned multi-person linear model." ACM Transactions on Graphics (TOG) 34.6 : 1-16. 2015.

[16] Kanazawa, Angjoo, et al. "End-to-end recovery of human shape and pose." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2018.

[17] Kocabas, Muhammed, et al. "PARE: Part attention regressor for 3D human body estimation." Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021.

Copyright ⓒ 2023 KSII

An Interdisciplinary Study on the Acoustic Perception Characteristics of TV Speakers: OLED Flat Panel Speaker and Dynamic Speaker Acoustic Focus

Kundong Choi[1], Sangwoo Hahm[1] and Hyungwoo Park[2]*
[1] Department of Business Administration, Semyung University, Jecheon, Republic of Korea(27136)
[2] Department of Information and Communication, Dong-Seoul University, Republic of Korea(13117)
[e-mail: Michael.park@du.ac.kr]
Corresponding author: Hyungwoo Park

Abstract

Previous studies have focused on the importance of audiovisual elements in information transmission, particularly in technologies like TV. In previous studies have explored the use of exciter speakers in OLED panels to improve the synchronization of video and sound and enhance realism. To investigate the difference in sound quality between down-firing and front-firing TVs, an experiment was conducted with 164 participants. After watching TV, participants completed a questionnaire assessing their perception. The Likert-5 scale was used to measure participants' cognitive responses, and statistical analysis was performed based on the survey results. Participants generally preferred the front-firing speaker for various reasons.

Keywords: Direct Drive Speaker, Exciter Speaker of OLED TV, Sound Improvement, front-down fire seoun compare

1. Introduction

People get a lot of information around them through visual, auditory, and various senses. In particular, audiovisual data is a very important information transfer factor and has developed various technologies and methods for a long time [1]. And with the development of these technologies, audiovisual information such as TV can be transmitted far away. In the meantime, the technological development of the display device has been focused on improving the image quality and improving the design of the device [2]. In previous studies, we have found that a better effect is achieved when the focus of the screen matches the focus of the sound [3]. To that end, it is now possible to create a sound by vibrating a modern display panel, such as an OLED (Organic Light-Emitting Diode) [4]. In our previous research, we have studied the research of creating a stereo sound by directly vibrating a single plane, the reason why the sound of this method is better than the existing sound, the position of the exciter optimized for sound quality, the shape of the exciter and the enclosure [5]. This study aimed to compare the cognitive aspects of existing down-firing sound TVs with a proposed sound TV system. The proposed system utilized a sound radiation system that addressed the drawback of reflected sound associated with down-firing speakers, resulting in improved clarity of the voice for viewers. By considering previous research findings, this study confirmed the advantages of

Following are results of a study on the "Leaders in INdustry-university Cooperation 3.0" Project, supported by the Ministry of Education and National Research Foundation of Korea

the proposed system from the viewers' cognitive perspective.

2. Previous Research and principal Background

After a speaker was developed as a transducer that converts an object into an acoustic signal and converts an electrical signal into an acoustic signal, magnetic materials such as permanent magnets have been developed after the improvement of high-temperature adhesives and related materials. Each was developed in the early 1930s [6]. OLED is a single layer with excellent vibration transmission, so the sound from the exciter is played along with the sound along the glass. The structure in **Fig. 1** can be modeled as a one-degree-of-freedom oscillation system in which an exciter, back cover, and plate are harmoniously connected [7,8].

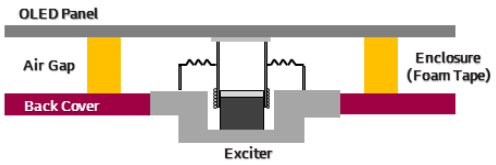


Fig. 1. Proposed model of Flat panel speaker

3. Experiment and Conclusions

A 10W direct drive exciter speaker mounted on an amplifier capable of channel-independent control to realize improved sound, that is, a TV of the same size with a single-core stereo exciter speaker TV and a conventional down-firing dynamic speaker. has been installed. Viewers for evaluation repeatedly compare and evaluate videos of various genres in terms of sound quality and direction of sound between 15 and 30 seconds. The evaluators of 164 participants compared the two types of TVs by completing a questionnaire on five items. The experiment was conducted in a group of about 10 people, and they were allowed to watch two types of TV without prior information. Questionnaires about participants' cognition were measured using the Likert-5 scale. Information on the survey questions is as follows.

Purchase intention
Wang, C. L., Li, D., Barnes, B. R., & Ahn, J. (2012). "Country image, product image and consumer purchase intention: Evidence from an emerging economy." International Business Review, 21(6), 1041-1051.
Choi K. M., & Doo S. J. (2007). "A Multivariate Analysis of Variance Applied to the Subjective Test of the Sound Quality of the Car Audio System." The Korean Journal of Applied Statistics, 20(3), 475-485.
Kim S. Y., Oh J. S. & Kim K. S. (2019), "Sensory Evaluation of Soun Quality for Vehicle Power Seat Slide Adjuster." Trans. Korean Soc. Noise Vib. Eng., 29(3), 318-326.
Sang-Woo Hahm*, Hyungwoo Park** "An Interdisciplinary Study about Voice Change of the Presidential Candidate and Cognition Change of the Voters" The Journal of The Institute of Internet, Broadcasting and Communication (IIBC) Vol. 18, no. 3, pp. 193-200, Jun. 30, 2018.
And 6 variables related to sound quality

And statistical analysis was conducted based on the survey results. The order of statistical analysis is as follows.
1. Factor analysis and reliability – Validation of survey validity and reliability
2. t-test – analysis of differences between two groups
3. Anova - analysis of differences over three groups
4. Corre – Correlation of variables
5. Regre – causal relationship of variables

According to the statistical analysis results, 1. People perceived the sound quality of the front firing TV set to be better than that of the down firing TV set. 2. There was no difference in the sound quality of the front firing TV set according to gender or sitting position. 3. There was a difference between Koreans and Chinese in terms of the sound quality of the front firing TV set. The Chinese gave higher scores than the Koreans. (Especially, comfort and body) 4. In the case of feeling good sound quality for the front firing, the intention to purchase was also high. 5. In particular, People who preferred front firing TV set in the order of dynamic, clearness, and body were more likely to purchase this product.

References

[1] Jun Tai Kim, Jung Ho Kim, Jin Oh Kim, Jin Ki Min. "Acoustic Characteristics of a

Loudspeaker Obtained by Vibroacoustic Analysis." Trans. Korean Soc. Mech. Eng. A vol. 21, No. 10, pp. 1742-1756, 1997.

[2] H. W. Choi, Y. J. Kim, Y. W. Park. "Acoustic analysis and vibration modeling for design of flat vibration speaker." Proceedings of Korean Society for Precision Engineering, 2008 Fall Conference (2008.11): pp. 545-546.

[3] D.J. Choi, Y. W. Park, H. J. Park, “Research of Relation Between Sound Pressure and Magnetostrictive Speaker For 2ch Flat Display,” Proceedings of Korean Society for Precision Engineering, 2010 Fall Conference (2010.11): pp. 603-604.

[4] Hong jung Beon, “Trends in speaker manufacturing technology at domestic and abroad,” The Magazine of the IEEE Korea, Vol. 13, No. 6, (1986.12): pp.513-520.

[5] SungTae Lee, Hyungwoo Park, Myungjin Bae and Kwanho Park, "Sound Quality Improvement of a Flat-Panel Display Front Speaker: Emotional and EEG Test of Exciter Speaker." SID Symposium Digest of Technical Papers. Vol. 48. No. 1. 2017.

[6] S.H. Nam, D.J. Choi, J.S. Chae, “Objective Audio Quality Assessment," proceedings of EEE Korea, (1995.1): 108-111.

[7] J.H. Kim, J.W. Kim, "Thin speakers and transparent speakers." CERAMIST, 17.2 (2014.6): 60-66.

[8] S.J. Oh, Theory and design of loudspeaker, SuckHakDnag press, 2011.

[9] S. Lee, Basic Properties of Sound and Application, Chung-Moon-Gak Publisher, 2004.

[10] U. Park, S. B. Lee, and S. H. Lee, General theory of sound technology, Cha Song Publisher, 2009.

Copyright ⓒ 2023 KSII

Anomaly Detection from X-ray Images based on Region-Specific and Global PatchCore

Hyunbin Kim and Junchul Chun*
Department of Computer Science, Kyonggi University
Suwon, South Korea
[e-mail: nu_start@naver.com, jcchun@kgu.ac.kr]
*Corresponding author: Junchul Chun

Abstract

This paper presents a method for diagnosing either presence or absence of lesions by detecting anomalies in chest X-Ray images. The method proposed in this paper is based on the PatchCore anomaly detection method, which extracts a feature vector containing the location information of an image patch from normal image data and calculates the distance from the normal vector as an anomaly. Since PatchCore defines normal features range from entire frame area, it is difficult to apply it simply to medical image processing, which may include diseases that occur only in specific organs and may include image noise unrelated to lesions. In this paper, we present an image alignment method through affine transformation parameter prediction that standardizes already captured X-ray images into a specific composition. We also present a region-specific abnormality detection method that requires affine transformed chest X-ray images. In addition, We present method for improving application efficiency and performance through feature map hard masking. From the experiments, we can show that the proposed approach achieved maximum AUROC of 0.774. Compared to the previous study conducted on the same dataset, the method achieved 6.9% higher performance and improved accuracy.

Keywords: Anomaly detection, Image alignment, PatchCore, Semi-Supervised Learning

1. Introduction

Vision based image analysis techniques such as classification, localization, object detection and semantic segmentation have been used for X-ray images. PA(Posterior-Anterior) X-ray images have been widely used for diaognising lesion. In previous work [1], we proposed an anomaly detection method using extracted patch features from pre-trained model. This method was designed to detect anomalies from entire area. However, any specific locations where the anomaly can exist should be considered to improve the accuracy of the anomaly detection method. Therefore, in this paper we propose a new approach. to detect anomalies from local and global area of the X-ray image.

2. The Proposed Method

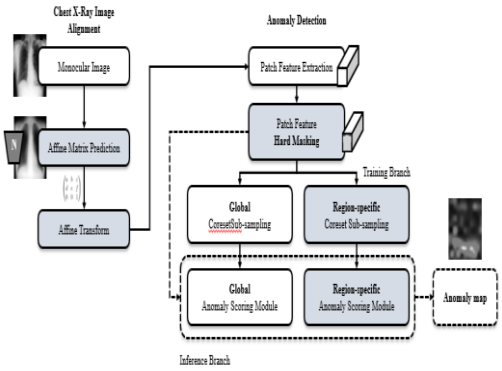


Fig. 1. Schematic diagram for anomaly detection from X-ray images

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2020R1A6A1A03040583)

The overall procedure for detecting anomaly from X-ray images consist of two major phases as illustrated in **Fig. 1**. The first is the step for alignmenting chest X-ray images and the second is the step to detecting anmalies based on global coreset sub-sampling and region-specfic coreset subsampling. In anomaly detecting, the proposed method detects both region-specific lesion from the specific area of the chest image and global lesion which are commonly occurred in various area based on patch feature hard masking.

2.1 X-ray Image Alignment

In general, PA(Postero-Anterior) approach to aquire the chest x-ray images uses same obtaining process however, the x-ray images show some different structual alignment. As shown in **Fig. 2**, the sample images are different form the canonnical image, it is necessary to adjust the sample images to an aligned sample.

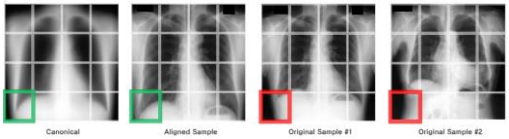


Fig. 2. Comparion of x-ray images: canonical, aligned, original sample x-ray images (from left to right)

The region-specific PatchCore is constructed by extracting features of the local areas from sub-coreset groups, which are properly aligned each other. Image alignment method basically adopts the method by Liu J. et al [2]. The canonnical image is directly made by 500 x-ray images using rotataion, perspection, translataion, scaling, shearing and bicubic interpolation. Consequently, the parameters of an affine transform matrix are evaluated for aligned images using the canonical and sample images. The similarity measure between two images is the joint loss of consistency, Reversed SSIM and VGG perecptual loss.

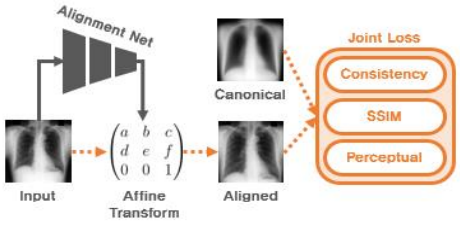


Fig. 3. X-ray image alignment using alignment network

2.2 Region-Specific Anomaly Detection

The currently available global PatchCore identifies the core vector based on the distance between vectors and does not consider local information in the training phase. The proposed Region specific PatchCore identify the nomality based on the local simlarity in the training phase like [3]. To obtain the region-specific anomaly detection, the input images are divided into n x n feature maps and the coreset subsampling is performed for each sub image. In this work, we consider n=4.

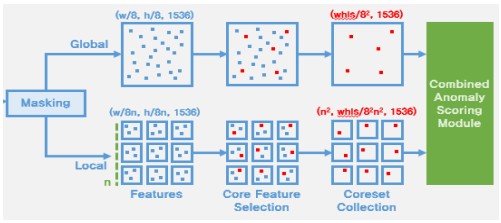


Fig. 4. The process of global and region-specific anomlay detection

2.3 Patch Feature Hard Masking

The parallel processing of the global and region-specfic anomaly detection needs huge memory and computational burden compared to global only PatchCore method. In this work, in order to improve the memory efficiency in training and computational efficiency in reasoning phase, we utilize Hard Masking by applying the anomaly detected area to extracted feature as shown in **Fig. 5**.

In this work, in order to only utilize the necessary chest area from the x-ray image while detecting anomaly, the binnary mask is used to remove the areas under a certain intensity level, 10 in our case, based on the average value of 500 x-ray images. The hard making can improve the nearest neighbor searching speed in evaluating a large numbers of feature vectors. In our study, we can reduce about 21% of feature vectors which are unnecessary for the process.

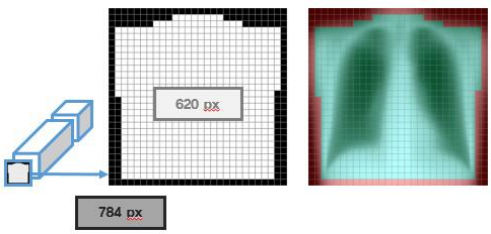


Fig. 5. An example of feature map hard masking: binary mask(left) and featuremap mak(right)

3. Experimental Results

In the experiment, 1,777 normal training data and 2364 testing x-ray images(half normal and half abnormal) are used from PadChest data set [4]. The evaluation is performed by using AUROC (Area Under Curve-Receiver Operating Characteristic). The developing environments are CPU (Intel i9-10980XE 18c36t, GPU(Nvidia Geforce RTX 3090), RAM (128GB), OS (Ubuntu 20.04), Language (Python 3.8), DL Framework (pytorch 1.8.1), CUDA(CUDA 11.1,cudnn8.0) and faiss-gpu.

Table 1-2 show the comparative AUROC results when the anomaly detection methods under different conditions such as utilization of image alignments, hard masking, combination of region and global PatchCore.

Table 1. Comparative results of anomaly detection according to either image aligned or not
(1. Not/ PatchCore, 2. Aligned/ PatchCore, 3. Not /Region-specific PatchCore, 4. Aligned/ Region-specific PatchCore, 5. Not/ Combined, 6. Aligned/Combined)

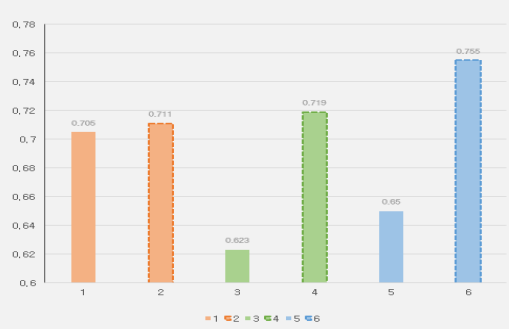


Table 2. Comparative result of annomaly detection according to either patch feature hard masking or not
(1. Not/ PatchCore, 2. Masking/ PatchCore, 3. Not /Region-specific PatchCore, 4. Masking/ Region-specific PatchCore, 5.Not/ Combined, 6. Masking/Combined)

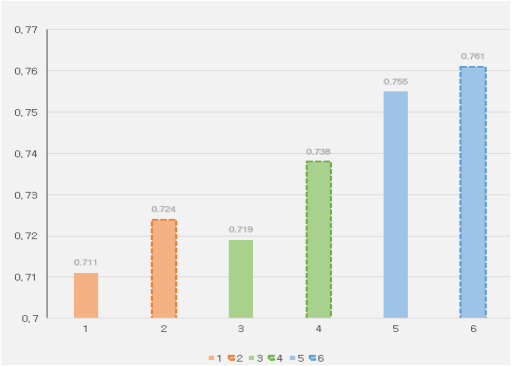
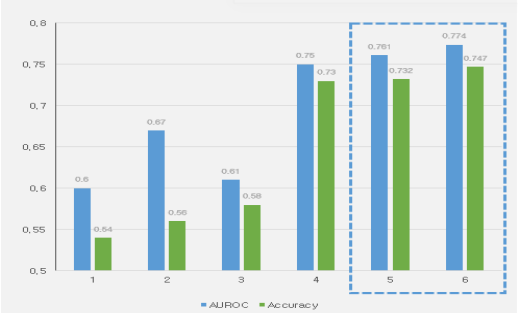


Table 3. Comparison of annomaly detection between the existing methods and the proposed method
(1. F-AnoGAN 2. Patch-SVDD, 3. Pix2pixHD, 4. Patch-SVDD MR Ensambled, 5.Proposed Combined(45GB), 6. Proposed Combined (128GB))



In **Table 3**, we can verify that the effiency and accuracy of the proposed method is better than other methods.

4. Concluding Remarks

In this paper, we present an image alignment method through affine transformation parameter prediction that standardizes already captured x-ray images into a specific composition. We also introduce a region-specific abnormality detection that requires affine transformed chest x-ray images. From the experiments, the method improves application efficiency and performance through feature map hard masking. The proposed approach shows maximum AUROC of 0.774 which is 6.9% higher performance of currently available methods using same data set.

References

[1] Hyunbin Kim and Chun, "Lesion detection in chest X-ray images based on Coreset of patch feature," *Journal of Internet Computing and Services*, Vol. 23, No. 3, pp 35-45, 2022.

[2] Liu, J. et al. "Align, Attend and Locate: Chest X-ray Diagnosis via Contrast Induced Attention Network with Limited Supervision," *Proceedings of the IEEE/CVF International Conference on Computer Vision*, pp.10632-10641, 2019.

[3] Yi, J. et al. "Patch SVDD: Patch-level SVDD for Anomaly Detection and Segmentation," *Proceedings of the Asian Conference on Computer Vision,* 2020.

[4] Bustos, A. et al. "Padchest: A Large Chest X-ray Image Dataset with Multi-label Annotated Reports," *Medical Image Analysis*, Vol. 66, pp.101797, 2020.

Copyright ⓒ 2023 KSII

Comprehensive Statistical Data Analysis for Production Decision in Aquaculture

Juhyoung Sung, Kyungwon Park, Byoungchul Song, Kiwon Kwon and Sungyoon Cho*
Korea Electronics Technology Institute (KETI)
Seongnam-si, South Korea
[e-mail: {jh.sung, kwpark, songbc, kwonkw, sycho}@keti.re.kr]
*Corresponding author: Sungyoon Cho

Abstract

In the aquaculture industry, as in any other business, a variety of factors have determined the price. In order to maximize production proficiency, fish farm managers try to predict fish prices. However, the prices are highly affected by a lot of unpredictable internal and external variables such as biological environments and market demand which makes decision making so difficult for the managers. In this paper, we carry out the analysis focusing on the halibut prices in the aquaculture. We collect various time-series statistical data related to the halibut production like local market price, the amount of shipment, and operation costs for farm maintenance from different data sources. Moreover, we integrate the statistical data through the data refinement. After the data pre-processing, we employ statistical methods including correlation analysis to explore the price decision factor. The approach using real-world statistical data related to the halibut production identifies relationships among different variables affecting the halibut price.

Keywords: correlation, decision making, profit enhancement, statistical data analysis, time-series

1. Introduction

As with various types of products, a large number of variables combine to influence fish prices. Fish farm mangers want to forecast the fish price trends because they are directly related to profits. Especially in aquaculture industry, highly fluent biological factors affecting fish growth also should be considered to determine the price. Therefore, the production decision making which includes from fish breeding to sales is significant in aquaculture operation. There have been several attempts to predict and analyze the fish price dynamics [1, 2]. However, most of typical studies have focused on the price itself raher than considering the relationship with external factors. Since the price is formed through interaction with the external environments, supplementary research is needed. In this paper, we try to analyze comprehensive variables which affect the halibut prices and identify the relationship among the different factors applying a statistical approach.

The remainder of this paper is organized as follows. In section 2, we describe the sturcture of collected real-world statistical data. The statistical analysis is presented in section 3. We conclude the paper in section 4.

2. Statistical data structure

To analyze the halibut price decision maker, we collect various statistical factors including local market prices and aquaculture maintenance costs from 2008 to 2022. Moreover, we divided the factors into 2 groups, production & sales (*Group I*) and maintenance cost (*Group II*) according to the data property. We summarize the detailed collected statistical data information in **Table 1**.

This research was supported by the project No. 20220579 “Big data-based aquaculture productivity improvement technology”, funded by the Ministriy of Oceans and Fisheries, Republic of Korea.

Table 1. Statistical data information

Data name		Collection Period
Production & sales (*Group I*)	Amount of shipments	Monthly
	Amount of exports	
	500g price	
	700g price	
	1000g price	
	1100g price	
	2000g price	
Maintenance cost (*Group II*)	Seed	Yearly
	Facility	
	Person	
	Operation	
	Feed	
	Etc	
	Total cost	

3. Result analysis

Since the data collection period is different as described in **Table 1**, we process the data collected on a monthly basis to a yearly basis to synchronize the data cycle. During the data refinement, we take the average of the values only within 1st quartile to 3rd quartile to mitigate the effect caused by outliers. We apply self-correlation within a group and cross-correlation between different groups. The correlation coefficient has a value between -1 and 1. As absolute value of the coefficient approaches 1, there is a strong correlation among variables.

Fig. 1 shows the correlation result among factors within *Group I*. As depicted in **Fig. 1**, there is a positive correlation between the amount of shipments and that of exports. In contrast, as the weight of halibut increases, the price generally tends to decrease when the amount of exports increases. Moreover, the halibut price has a similar trend regardless of the weight.

Fig. 2 illustrates the correlation result among factors within *Group II*. We can check the trends of the cost spent on feed and the total maintenance cost are very similar from the result of **Fig. 2**. In contrast, labor and operating costs show the opposite treand to the total maintenance cost.

Fig. 3 describes the correlation among variables in *Group I* and *Group II*. The halibut prices are

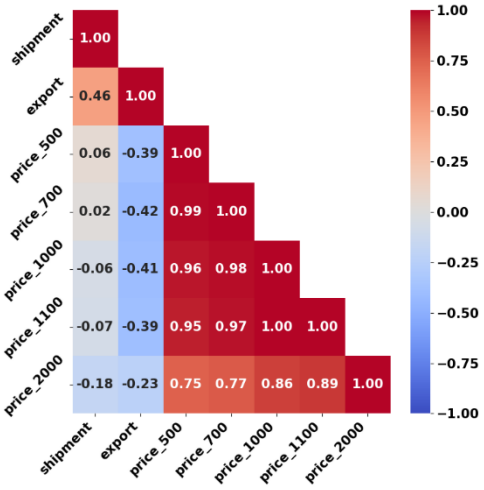


Fig. 1. Self-correlation result within a *Group I*

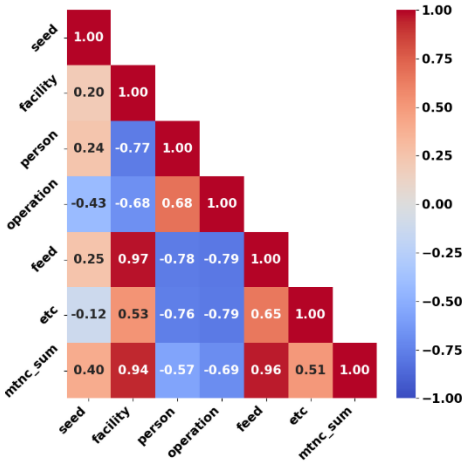


Fig. 2. Self-correlation result within a *Group II*

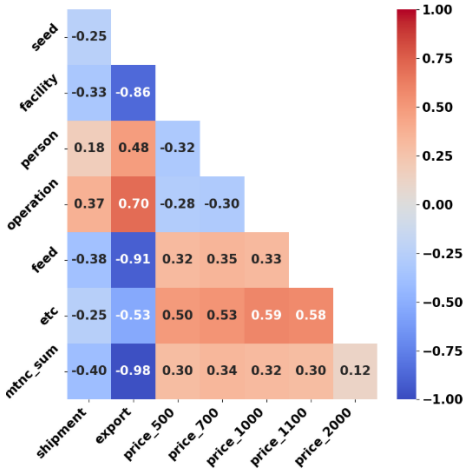


Fig. 3. Cross-correlation result between *Group I* and *Group II*

formed similarly to the total maintenance cost change. Interestingly, we can also find that the various factors consisting of the maintenance cost have different effects on the prices.

4. Conclusion

In this paper, we presented the relationship among variables affecting the halibut price through statistical approach. The results provided production decision makers with a quantitative understanding of price variation, enabling fish farm managers to optimize production planning.

References

[1] R. Xiang, C. Jones, R. Mamon, and M. Chavez, "Modelling exchange-driven fish price dynamics," *J. Model. Manag.*, vol. 16, no. 4, pp. 1054-1069, Aug. 2021.

[2] R. Dahl, and M. Yahya, "Price volatility dynamics in aquaculture fish markets," *Aquac. Econ. Manag.*, vol. 23, no. 3, pp. 321-340, Jul. 2019.

Copyright ⓒ 2023 KSII

Data communication dualization system between drone and GCS using RF and LTE

Wonseok Jung*, Jong-Hong Park and Il-Yeop Ahn
Autonomous IoT Research Center, Korea Electronics Technology Institute
Seongnam, South Korea
[e-mail: wonseok@keti.re.kr, jonghong@keti.re.kr, iyahn@keti.re.kr]
*Corresponding author: Wonseok Jung

Abstract

In this paper, a Ground Control System(GCS) using both Radio Frequency(RF) communication and Long-Term Evolution(LTE) was proposed to solve the problem of communication loss due to communication distance and obstacles, which are problems in the existing RF communication-based drone and GCS control system. The proposed system prevents data redundancy by preferentially allowing data transmitted through RF communication through a data communication redundancy method that utilizes RF communication and LTE communication simultaneously, and immediately utilizing data transmitted through LTE communication when RF communication becomes impossible. Through this, it solves the problems of existing RF communication and increases the stability of drone operation. The proposed system was verified by actual application and flight.

Keywords: Data communication dualization, RF, LTE, Drone, GCS, stable drone operation, oneDrone-S

1. Introduction

Drones and ground control systems (GCS) were first started for military purposes, but with recent advances in communication technology, they are rapidly developing into public service fields such as infrastructure monitoring, disaster response, and search and rescue. Drones used in various fields are being used for one service by collaborating with many drones, and the operating range is expanding, and non-visual operation is the main focus [1,2].

In order to safely operate multiple drones used in public services, research on means to monitor and control drones on the ground is being actively conducted at home and abroad. In the case of RF communication, the operating radius is limited, it is difficult to avoid interference in the same band, and if there is an obstacle between the drone and the ground, frequent communication disruptions are encountered. In the case of cellular communication, it is possible to utilize the existing commercial network, but due to the construction of a network that does not consider drones in the sky, it is not optimized for application to drones. In addition, actual drone operators still prefer drone control through existing RF communication, making it difficult to replace drone-ground communication with other communication technologies.

Accordingly, this paper proposes a data communication redundancy system, uses RF communication and LTE communication at the same time to secure stability in data communication between the drone and the ground, and tries to prevent data duplication by prioritizing RF communication. Section 2 of this paper introduces the data communication redundancy system and explains data redundancy processing. A conclusion is drawn in Chapter 3.

This work was supported by the R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT under Grant NRF-2022M3E8A2089225.

2. Data communication dualization system

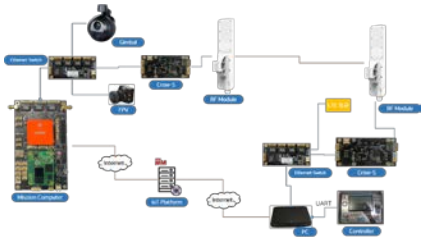


Fig. 1. Data communication dualization system structure diagram using RF and LTE

Fig. 1. shows the system architecture of the oneDrone-S platform, which supports data duplication, prevents data duplication, and enables RF communication encryption by attaching and detaching the encryption/decryption module. The hardware of the oneDrone-S platform is largely divided into the drone part, server, and GCS part. In the drone part, there is a Raspberry Pi Compute Module 4-based mission computer (MC) with a built-in LTE module called Crow and a flight controller (FC) called Pixhawk Cube Orange, and the MC and FC are connected through UART and SBUS. The MC is connected to the RF module via Ethernet with the Ethernet switch in between. At the center is an Ubuntu-based server with Mobius, an IoT server platform. Lastly, the GCS part includes a controller capable of generating 32-channel SBUS control signals and a mini PC connected by UART, and the mini PC is connected to the RF module through Ethernet with an Ethernet switch in between. If encryption is required, Crow-S, an encryption module, is installed between the Ethernet switch and the RF module of the drone part and the GCS part, respectively. The software includes oneDrone-TELE and oneDrone-RC on the MC board of the drone part. oneDrone-TELE receives drone data and transmits this data to Mobius on the server via the LTE module and to Mobius on GCS via the RF module. In addition, it receives drone control commands sent from GCS through the LTE module and RF module. In this process, data redundancy treatment methods are applied. oneDrone-RC receives RC data from the Mobius of the server through the LTE module and Mobius of the GCS through the RF module and transfers it to the FC after processing data redundancy. Mobius installed on the server is developed with Node.js based on the oneM2M standard and collects data using MySQL Database. Mini PCs in the GCS part include oneDrone-RC-HUB and oneDrone-TELE-HUB. oneDrone-RC-HUB receives RC data from the controller connected to Mini PC through UART. RC data is transmitted to the Mobius of the server through the LTE module, and transmitted to the Mobius of GCS through the RF module. Also, oneDrone-TELE-HUB receives drone data through the LTE module and RF module. The received drone data is transferred to GCS S/W after data redundancy processing. The data redundancy processing method declares a data object for each drone data, drone control command, and RC data. and after checking the sequence included in the data, the data received through RF communication is transmitted after assigning the sequence as a key value to the object, and the data received through LTE communication is deleted if there is a sequence in the object and is not forwarded to the next destination. But if there is no sequence, it passes to the next destination. Finally, Crow-S operates the Korea Cryptographic Module Validation Program(KCMVP) and encrypts and decrypts data sent and received through RF communication.

3. Conclusions

In this paper, the proposed RF communication and LTE communication-based data communication redundancy system was explained, and the data redundancy processing method was explained to prevent data redundancy. The proposed system was verified by actual application and flight at the Hwaseong drone flight test site.

References

[1] M. Alwateer, S. W. Loke and N. Fernando, "Enabling Drone Services: Drone Crowdsourcing and Drone Scripting," in *IEEE Access*, vol. 7, pp. 110035-110049, 2019.

[2] P. Hell, M. Mezei and P. J. Varga, "Drone communications analysis," *Proc. of 2017 IEEE 15th International Symposium on Applied Machine Intelligence and Informatics (SAMI)*, Herl'any, Slovakia, pp. 000213-000216, 2017.

Copyright ⓒ 2023 KSII

Integrated Data Control Platform for Building a National Park Drone Sky Road

Jong-Hong Park*, Wonseok Jung and Il-Yeop Ahn
Autonomous IoT Research Center, Korea Electronics and Technology Institute
Gyeonggi-do, South Korea
[e-mail: jonghong@keti.re.kr, dnjstjr93@keti.re.kr, iyahn@keti.re.kr]
*Corresponding author: Jong-Hong Park

Abstract

This paper relates to a drone-based system in national parks for efficient regular patrol and accident response in case of emergency. In particular, for the efficient use of the drone system in national parks, a platform for integrated control in the central control system and the construction of a drone sky road is proposed. The drone sky road construction includes the creation of an optimal route through the development of a three-dimensional communication quality coverage map for a mobile communication network including LTE or 5G and setting of a geo-fence for the communication shadow area. This paper aims to develop technology and system for resolving communication shadow areas in case the flight of the established geo-fence is unavoidable. In addition, it aims to develop a platform-based central integrated control system for the control of all drones owned by national parks across the country.

Keywords: drone sky road, mobile communication network, ground control system, unmanned aerial vehicle, data platform

1. Introduction

South Korea has 22 national parks, covering approximately 6% of the country's total land area. These parks are managed by the Korea National Park Service, which was established to protect and preserve the natural beauty and cultural heritage of these areas. In 2019, the Korea National Park Service conducted research to establish a mid- to long-term plan for national park spatial information management, which led to the discovery of customized drones for park management. Since then, the agency has introduced and operated 64 drones for various purposes in national parks.[1] The use of drones has allowed for more efficient and effective park management, including monitoring and mapping of park areas, wildlife surveys, and search and rescue operations. The Korea National Park Service introduced a multipurpose drone system, and after reviewing its operation status, several challenges were identified. One of the challenges is the difficulty in designating a flight path due to the absence of topographic and communication information. Additionally, there is a communication blind spot between the ground control center and drones, which can lead to unconditional mission suspension and return when flying in this area. Another challenge is the lack of a drone video control and monitoring system, which makes it difficult to accurately grasp the on-site situation and control the drones in real-time. These challenges need to be addressed to ensure the safe and effective use of the multipurpose drone system in national parks. The Korea National Park Service is working to develop solutions and improve the system accordingly.

This work was supported by the R&D Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science and ICT under Grant NRF-2022M3E8A2089225.

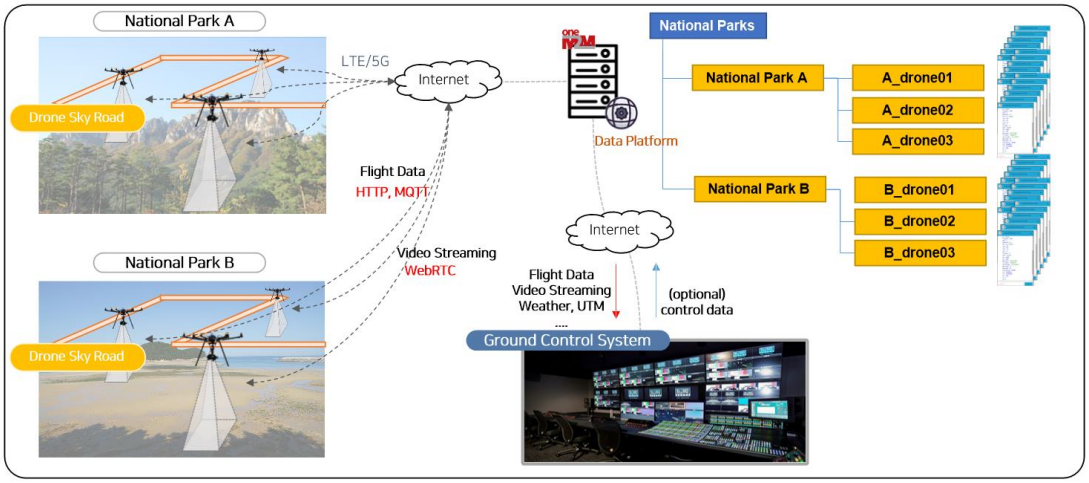


Fig. 1. National park drone sky road and integrated control platform structure diagram

2. Necessity to Promote Development

The Korea National Park Service recognizes the need for a national park drone sky road and an integrated control platform. Accurate information on the communication status of the commercial network within the national park is crucial[2]. Drones are mainly used to fly in rugged mountainous terrain and invisible areas, and a prompt response and regular patrol are essential tasks in case of a safety accident. Therefore, when using drones, there is a high dependence on communication status. In addition, it is necessary to find a solution when communication with the ground control system is restricted (GCS). The possibility of falling due to uncontrollability when communication is restricted is a concern. Currently, only the return function (Return to Launch, RTL) is available when communication is restricted, and utilization in case of an emergency, such as delivery of emergency goods, is low due to the impossibility of continuous mission performance. To ensure stability in the operation of the demand side, it is necessary to develop an integrated control system. There is a need to develop an integrated control platform for managing not only drone flight data but also various data, considering the passive operation due to communication status problems and surrounding environmental problems in operating expensive drones.

3. Conclusions

In this paper, a detailed analysis of the current state of national parks in Korea, with a focus on the use of drones for park management has presented. Several challenges and limitations in the current system have identified and have proposed potential solutions for future improvements. In particular, we have highlighted the necessity of building a national park drone sky road and developing an integrated control platform to ensure safe and effective drone operations within the park. Additionally, we have emphasized the importance of accurate communication information and the need for solutions to address restricted communication with the ground control system. Overall, our dissertation represents a significant contribution to the field of national park management, and we hope that it will serve as a valuable resource for future research and policy development.

References

[1] Y.-G Han, Y. Cho, and O. Kwon, "The use of conservation drones in ecology and wildlife research," *Journal of Ecology and Environment,* 38.1, 113-118, 2015.

[2] J.-H. Park and K.-H. Won, "Implementation of UAS Traffic Management System in LTE Environments," *in proc. of 7th Asian-Australian Rotorcraft Forum*, 2018.

Copyright ⓒ 2023 KSII

Implementation of Centimeter-level Maritime Precise PNT Service Platform

Junsu Kim, Juhyoung Sung, Wongi Jeon, and Kyungwon Park*
Smart Network Research Center, Korea Electronics Technology Institute (KETI)
Seoul, South Korea
[e-mail: {jesus0084, jh.sung, jeonwg, kwpark}@keti.re.kr]
*Corresponding author: Kyungwon Park

Abstract

There are two methods of generating correction information for centimeter-level GPS services. One is the observation space representation (OSR) in which the reference station generates the correction information based on the current user position. The other is the state space representation (SSR) which generates the correction information for each error factor in the entire service region in advance and broadcasts it by using ground stations or satellites. Among them, it is effective to use the SSR method to obtain the correction information for centimeter-level maritime precise position, navigation, and time (PNT) services in maritime environment since the SSR is relatively free from the use of the reference station. In this paper, we implement a centimeter-level maritime precise PNT service platform based on the SSR. The proposed platform consists of a GPS module, a wireless communication module that receives the SSR-based correction information, and a CPU module that combines location data and correction information to calculate precise location information and provides the results to users.

Keywords: GNSS, maritime, PNT, positioning, SSR

1. Introduction

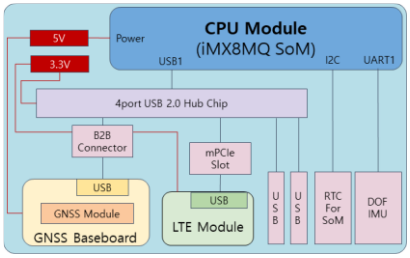
The correction information for centimeter-level GPS services which is generated by the SSR method in advance for each error factor [1] is broadcast through ground stations or satellites. The SSR method is free from the use of the reference station compared to the OSR method in which the reference station receives the current location of the receiver and generates correction information based on the receiver's position [2]. In this paper, we design a PNT service platform providing centimeter-level GPS services at maritime using the correction information generated by the SSR method.

2. Hardware Design

The PNT service platform consists of a GNSS module, a wireless communication module, and a CPU module. The GNSS module receives GPS signal from satellites to obtain an approximate current position. The wireless communication module receives correction information from the ground station through the network transport of RTCM via internet protocol (NTRIP). The CPU module parses the correction information generated by the SSR method from the correction information received through the network module and combines it with the GPS location information received by the GNSS module to calculate centimeter-level location information.

Fig. 1 shows a block diagram and its hardware implementation for the PNT service platform.

This research was supported by a grant from National R&D Project “Development of ground-based centimeter-level maritime precise PNT technologies” funded by the Ministry of Oceans and Fisheries(1525013759)



(a) A hardware block diagram



(b) Hardware implementation

Fig. 1. A block diagram and hardware implementation for the PNT service platform

3. Software Design

Ths software of the PNT service platform is divided into two parts. One receives radio technical commission for maritime services (RTCM) data through NTRIP in a wireless communication module, parses it, and extracts correction information in SSR format.

The other obtains centimeter-level location by combining the SSR correction information, PPP-RTK data that combines RTK (Real-Time Kinematic) and PPP (Precise Point Positioning) data, and the GPS location obtained through the GNSS module. The block diagram of the PNT service platform software is shown in **Fig. 2**.

Additionally, the PNT service platform embeds a web-based monitoring application that displays the current location and PNT service platform status to the user. The web-based monitoring application is programmed in Flutter web and displays the current location and status information of the PNT service platform to the user in a dashboard view. The GUI of the

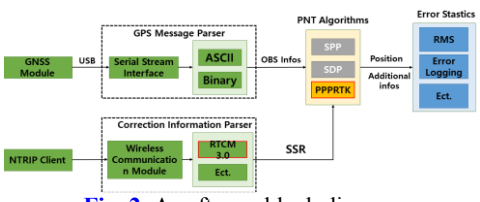
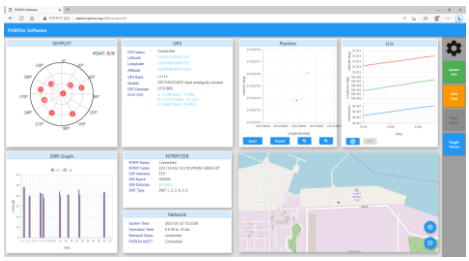


Fig. 2. A software block diagram



(a) The dashboard GUI of the monitoring application



(b) The scatter plot and line graph for a receiver

Fig. 3. The GUI of the web-based monitoring application

web-based monitoring application is shown in **Fig. 3.**

4. Conclusions

In this paper, we designed a PNT service platform providing centimeter-level GPS services at maritime using the correction information generated by the SSR method. We expect that the designed platform can provide the precise location information in a marine environment.

References

[1] M. Kim and K. Park, “Development and positioning accuracy assessment of single-frequency precise point positioning algorithms by combining GPS code-pseudorange measurements with real-time SSR corrections,” *Sensors*, vol. 17, no. 6, Jun. 2017.

[2] J. Park, “Development of gridded-VRS using real-time SSR messages provided by Korean National Geographic Information Institute,” Inha University, 2023.

Copyright ⓒ 2023 KSII

Improvement of Ion Mobility Spectrometry System Performance through Structural Design Change of Drift Tube

Gi-Seop Im*, Won-Young Lee, and Sun-Nam Hwang
NewOne S&T Ltd
Seoul, South Korea
[e-mail: semizeus@hanmail.net, 1_lee_3@naver.com]
*Corresponding author: Gi-Seop Im

Abstract

This paper presents a method to improve the performance of Ion Mobility Spectrometry technology by changing the drift tube structure. IMS is a valuable technique applied to the trace detection of explosives and narcotics through the characterization of the ion mobility in a drift tube formed electric field. The Drift tube consists of an ionization module for target material, a drift region for ions, and collector module that transmits electrical signals via ions and is designed as an integral structure. The performance of the IMS system can be enhanced by changing the design factor of the drift tube. This study evaluates the performance improvement of the IMS system by measuring narcotic samples. The result is expected to enhance the measurement reliability and reproducibility of the trace detection device.

Keywords: Ion Mobility Spectrometry, Drift tube, Ionization, Collector,

1. Introduction

The Ion Mobility Spectrometry (IMS) system de tects substances by measuring the difference in mobility of ions generated by inhaling a target s ubstance under atmospheric pressure. In most cases, narcotics to be tested generate positive ions in the ionization process, and the production process is shown in the equation.

$M + H^+(H_2O)_n \rightarrow MH^+(H_2O)n \rightarrow MH^+(H_2O)_{n-x} + xH_2O$
Sample + Reaction ion→ cluster ion→ Product ion + water

Ion formation in IMS has been accomplished through chemical reactions between molecules of the sample and a reservoir of ions known as reactant ions. The Product ions generated in this process move to the collector with the inherent mobility in the electric field of the drift tube and generate an electric signal. In addition, The product ions collide with the collector, consequently generating an electric signal of about 10 ~ 100pico ampere. This electric signal is converted into a voltage signal of 0 to 10 volts through an amplifier [1].

In this paper, the improvements are described through structural changes in the ionization module and collector module.

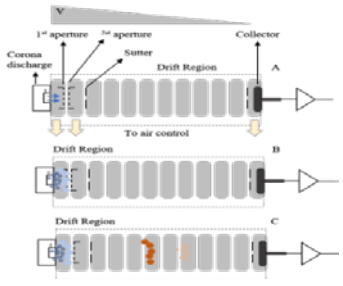


Fig. 1. Schematic representation of drift tube for IMS

This research was supported by a research grant from KIMST(Korea Institute of Marine & Technology Peomotion)

2. Reconstruction of Drift Tube structure

2.1 Drug samples for evaluation

Three kinds of narcotic samples: Fentanyl, Ketamine, and MDMA were used to evaluate the effect of the structural changes in the drift tube. These substances have analgesic and hallucinogenic effects and are rapidly spreading recently.

2.2 Structural change to the ionization

The ionization module is located at the front end of the integral drift tube, and structural changes were made to the ionization module.
The reactivity of the system was tested using sample materials to gauge the design change. The drift tube used in the experiment is shown below



Fig. 2. Drift tube structure in IMS

In order to generate more ions, a structure was designed to secure a space to collect gas in the ionization module, and the reactivity was confirmed using a drug sample with the new structure. The same volume of samples is used for the test. it was confirmed that the reaction intensity is increased in the design-changed structure [2,3].

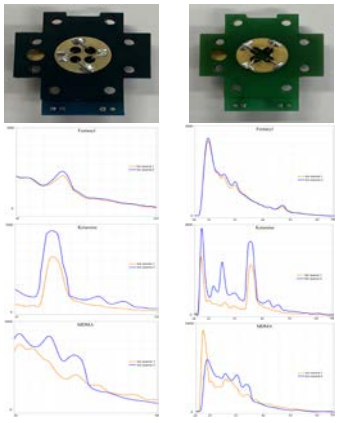


Fig. 3. Reaction intensity depending on the ionization

2.3 Structural change to the collector

The collector module is located at the end of the drift tube, and the effect of the structural change of the collector was tested through the change in reactivity intensity using a narcotic samples. The collector module was designed to collect signals by directly connecting it to the base ring of the drift tube, and by applying this structure, the change in reactivity of the narcotic samples converted into an electrical signal was confirmed [2,3].

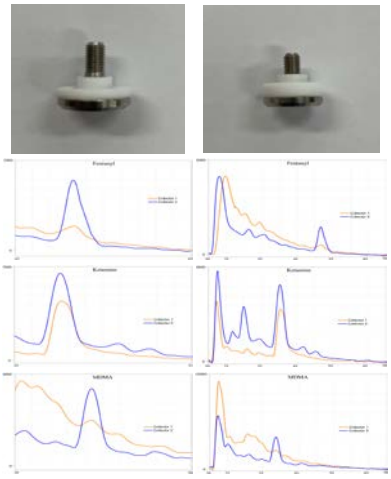


Fig. 4. Reaction intensity depending on the collector

3. Conclusions

Security device IMS technology should be able to perform more accurate detection by increasing the reactivity to extremely small amounts of hazardous substances. It was confirmed that the reactivity to the drug samples was improved by securing space in the ionization module. This has the effect of increasing the pulse width of the collected signal, which may cause overlap with interfering substances. However, since the range is not large, it is not expected to be a major problem for hazardous substance identification.
In order to confirm the effect of the structural change of the collector module, there was no change in the structure of the grid mounted on the front end of the base ring, and the sample material of the same volume was applied to test.
The results showed that the structural position of the collector module improved the response to detected substances.

References

[1] S. A. Eiceman, Z. Karpas, “Ion Mobility Spectormetry”, Second Edition, pp. 3-9, 2005.

[2] Sung-seen Choi, Chae-eun Son, Myung-won Shin, Gyu-seop Choi “Influence of Smear Matrix Type on Detection Efficencies of Explosives in Corona Discharge-Ion Mobility Spectrometer” *Bulletin of the korean Chemical Society*, Vol 37, No.4, 2016.

[3] Sung-seen Choi, Yun-ki Kim, Ok-bae Kim, Seung-geon An, Myung-won Shin, Seug-jin Maeng, Gyu-seop Choi “Comparison of Cocaine Detections in Corona Discharge Ionization-Ion Mobility Spectrometry and in Atmospheric Pressure Chemical Ionization-Mass Spectrometry” *Bulletin of the korean Chemical Society*, Vol 31, No.8, 2010.

Copyright ⓒ 2023 KSII

AI-enabled AGV service management with AIoT platform

SeungMyeong Jeong*, Dae-Geun Heo and ChungJae Choe
Autonomous IoT Research Center, Korea Electronics Technology Institute, Seongnam, South Korea
[e-mail: sm.jeong@keti.re.kr, hdkyon96@keti.re.kr, cjchoe1@keti.re.kr]
*Corresponding author: SeungMyeong Jeong

Abstract

As one of the AIoT (AI + IoT) use cases, an AI-enabled AGV (Automated Guided Vehicle) in a warehouse can detect materials and count objects with radar sensors and AI models on it. In such systems, AI model deployment is necessary since models are running on AGVs while they are created by AI modelers and service admins manages the system remotely. This paper proposes the AIoT platform with oneM2M standard extension APIs, so AI model management solution can be interoperable with AI devices and service applications.

Keywords: AGV, AI, AIoT, IoT, oneM2M

1. Introduction

IoT has been evolved as AIoT with AI capabilities on devices and platforms. For AIoT service development, the platform supports AI features such as AI training and inference dataset management instead of AI modelers [1].

AI/ML is being adopted to radar sensors on AGV so they can move autonomously [2]. Lately, the radar sensor on also can be used for material detection and counting which cannot be done by camera visions [3].

In this paper, we present an AIoT platform, which is the AI/ML model management extended capability on oneM2M standard compatible IoT platform. With the proposal we show the new AIoT capabilty can be beneficial for the AI-enabled AGV services in terms of remote AI model management in the field.

2. AI/ML model management

2.1 system design

The main idea of this paper is to provide the AI/ML model management capability on the AIoT platform so AI-enabled AGVs and the management portal for service users can focus more on service implementations.

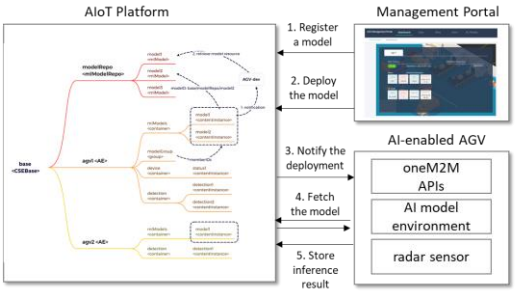


Fig. 1. System configuration with AIoT platform

Fig. 1 depicts the system configuration and the overall call flows. When a user uploads a new model on the portal, the portal registers the model as a oneM2M <*mlModel*> resource, as defined in the Table 1.

By the service admin on the user portal, a model is chosen and deployed to an AGV. The deployment command includes the resource ID of the registered model and it is notified to the AGV. In this scenario, each AGV is subscribing

This work was supported by the Technology Innovation Program (20016243) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea)

an model deployment comand resource (i.e. *mlModels*) so it can get the push notification and fetch the model resource from the platform.

Table. 1. Attributes of an <*mlModel*> resource

attribute	description	sample data
name	model name	"material detection"
version	model version	2
platform	ML platform	"tensorFlow"
description	model description	"counting model for radar sensor"
inputSample	inferent input data signature sample	"{\"keys\": [[11.0], [2.0]], \"features\": [[1, 1, 1, 1, 1, 1, 1, 1, 1], [1, 1, 1, 1, 1, 1, 1, 1, 1]]}"
model	model file	(base64 encoded binary file)
modelUrl	URL of a remote	http://mydomain/models/model1

Once the model is retrieved and served on the AGV, radar sensing data is provided into the model and the inference result (e.g. "glass" detected) is stored on the AIoT platform.

Fig. 2 represents the proposed oneM2M resource tree design in the AIoT platform with the new resource types (i.e. *mlModelRepo* and *mlModel*) yet in oneM2M.

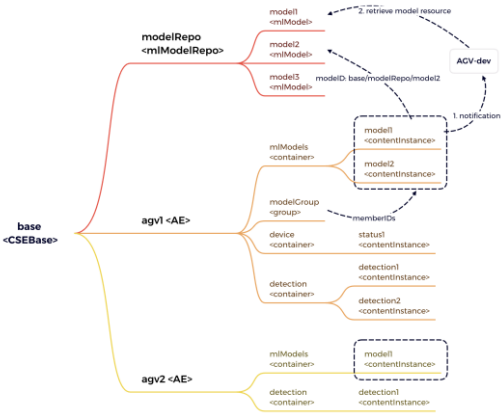


Fig. 2. oneM2M resource tree design

New ML model registration by the portal creates new <*mlModel*> resources as the children of <*mlModelRepo*> resource. When a registered model is deployed to an AGV, the portal creates a model information under *mlModels* resource under an AGV resource (e.g. "*agv1*").

2.2 Implementation with Portal

In this system, there are two type of users that use the portal, which communicates with the AIoT platform with the proposed interfaces. One is the AI modelers who registers the model to the AIoT platform via the portal. The other user, the service admin, deploys AI models to the field machines as well as manages sensing and detection data on the portal.

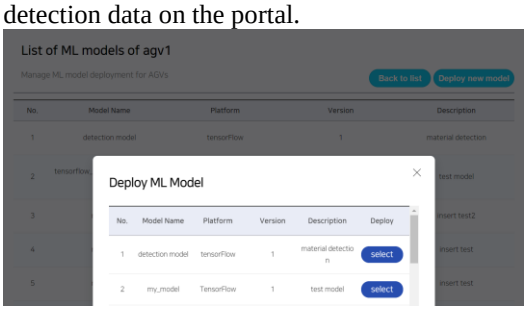


Fig. 3. AI/ML model managent portal

4. Conclusions

In this paper, we present an AI/ML model management for systems with AI-enabled devices. The AIoT platform serves extended API on top of oneM2M standard platform, so it could be interoparable solutions for AI models when it gets adopted in the specifications.

References

[1] S. Jeong, Y. Lee, D. H. Park, "Machine Learning dataset creation common service for real-time data platform," *International conference on Internet 2021*.

[2] C. Rafal, J. C. Lin, and J. H. Syu, "Automated Guided Vehicles Challenges for Artificial Intelligence," *2022 IEEE International Conference on Big Data*, Osaka, Japan, pp. 6281-6289, 2022.

[3] T.S. Gunasekaran, R, Hajika, Y.S. Pai, E. Hayashi, "RaITIn: Radar-Based Identification for Tangible Interactions," *CHI Conference on Human Factors in Computing Systems Extended Abstracts*. 2022.

[4] A. Paleyes, R.G. Urma, N.D. Lawrence, "Challenges in deploying machine learning: a survey of case studies," *ACM Computing Surveys*. 55.6, 1-29, 2022.

Copyright ⓒ 2023 KSII

Development of a drone-based low-light environment real-time crack detection system

Sung-Chan Choi*, Jong-Hong Park, Sung-Wook Jung and Il-Yeop Ahn
Autonomous IoT Research Center, Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: csc@keti.re.kr]
*Corresponding author: Sung-Chan Choi

Abstract

In this study, we developed an advanced solution for automatically segmenting cracks in tunnels. Our approach integrates image processing technology with deep learning-based artificial intelligence, considering challenging variables and environmental conditions in tunnels. To address low illumination, we employed sophisticated techniques including SID, Zero-DCE, and EnlightenGAN. These methods enhance image quality by adjusting gamma and introducing Gaussian noise based on artificially generated low-light images. For crack segmentation, we implemented a state-of-the-art CNN-based U-Net model. To objectively evaluate our solution's performance, we used widely used metrics such as Dice score, IoU, and F1 Score. These metrics quantitatively measure the agreement between the predicted crack image and the ground truth label image. Through rigorous evaluation, we validated the effectiveness of our proposed approach.

Keywords: Deep Learning, Object Detection, Unmanned Vehicle Communication Networks

1. Introduction

Cracks in concrete structures, roads, and tunnels raise concerns and cause anxiety. Current methods for inspecting concrete cracks in tunnels rely heavily on human inspection, which has drawbacks such as difficulties in accurate quantification and identification due to low illumination and limited accessibility. Hence, a solution is needed to accurately segment crack areas in low-light environments. [1] This study addresses these issues by using a U-Net model based on CNN to generate AI-based predictive images of cracks under different illumination conditions. The developed deep learning model was deployed on Jetson embedded hardware devices, and a socket communication software was implemented for transmitting detection results wirelessly. The remaining sections of this paper include the system design in Section II, experimental results in Section III, and conclusions in Section IV.

2. System Design

The system includes a drone that operates in tunnels, a low-light camera mounted on the drone, and a deep learning crack detection algorithm. It detects cracks in low-light tunnel environments using drones.

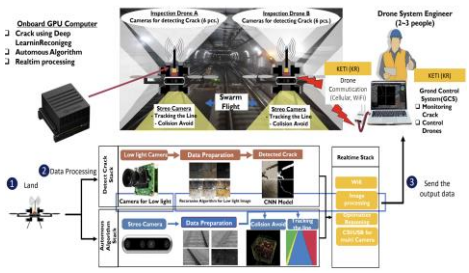


Fig. 1. Drone-based Crack Detection System

(Support note) This research was supported by Unmanned Vehicles Core Technology Research and Development Program through the National Research Foundation of Korea(NRF) and Unmanned Vehicle Advanced Research Center(UVARC) funded by the Ministry of Science and ICT, the Republic of Korea (NRF-2020M3C1C1A02086427).

2.1 U-Net Deep Learning Model

The U-Net model, introduced by Olaf et al. in 2015, is widely used for biomedical image segmentation. [2] It employs convolution filters, pooling layers, up-convolution layers, and skip connections. **Fig. 2** shows the model's structure, which can handle two or three dimensional image data. It consists of an encoder, decoder, and skip connections. The encoder extracts features and reduces the image size, while the decoder reconstructs the image. Skip connections enable direct information transfer, improving segmentation accuracy.

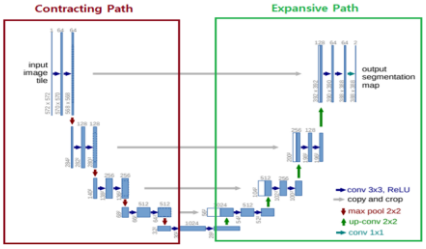


Fig. 2. U-Net Architecture

2.2 Models for Improving Low-Light Environments

Zero-DCE is a deep learning-based method that enhances image brightness by estimating the light curve and adjusting the dynamic range. EnlightenGAN enhances illumination and color using an illumination map and attention map, employing an encoder-decoder structure similar to U-Net. SID is an FCN model that processes low-light images in a single step, combining multiple processing steps and enhancing image quality.

3. Experiment Results

Table 1 compares performance metrics of different models using 50 test videos. Preprocessed test images were evaluated using Precision, Recall, F1, and Dice scores commonly used for segmentation models. Results, summarized in **Table 1** and shown in figures, align with the model's training performance. Low-light enhanced images perform similarly or worse than low-light trained images. Fig. 3 shows automatically segmented predicted images from each model, except for SID, which closely resembles the correct videos.

Table 1. Performance evaluation metrics for each model using test footage

	Precision	Recall	F1 Score	Dice Score
Low-light model	0.8945	0.7921	0.8231	0.7256
Zero-DCE	0.8867	0.8267	0.8501	0.6962
SID	0.4895	0.4558	0.4526	0.3892
EnlightenGAN	0.8937	0.8296	0.8571	0.7823

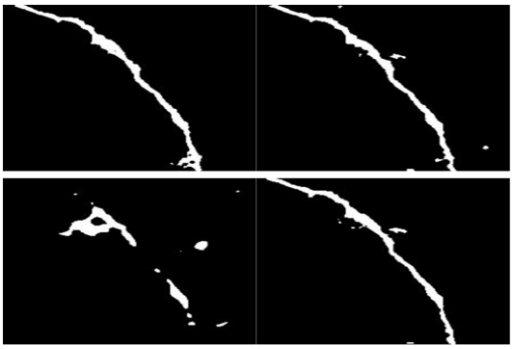


Fig. 3. Predictive video using test footage

4. Conclusions

In this paper, we developed an automatic tunnel crack segmentation solution using image processing and deep learning AI. We applied SID, Zero-DCE, and EnlightenGAN to enhance image illumination. The CNN-based U-Net model was used for crack segmentation. Evaluation metrics such as Dice score, IoU, and F1 Score were used to assess model performance. In the future, we aim to improve the model by acquiring more training data and exploring weight transfer learning for compensating data limitations.

References

[1] Liang, Jianbing, Suxia Wang, Shuang Zhao, and Shuhui Chen. "FECC: DNS Tunnel Detection model based on CNN and Clustering," Computers & Security 128, 103132, 2023.

[2] Li, Xiaomeng, Hao Chen, Xiaojuan Qi, Qi Dou, Chi-Wing Fu, and Pheng-Ann Heng. "H-DenseUNet: hybrid densely connected UNet for liver and tumor segmentation from CT volumes," IEEE transactions on medical imaging 37, no. 12, 2663-2674, 2018.

Copyright ⓒ 2023 KSII

Structural relationships of various factors on the intention to use a mobile anger management training application : focusing on the mediating effect of anger management utility

Kyunghan You and Minhee Lee
Department of Media Communications Jeonbuk National University Jeonju, Korea
Jaehee Cho*
School of Media, Arts, & Science , Sogang University Seoul, South Korea
Sehwan Kim
School of Medical Science, Dankuk University Cheon-An , Korea
[e-mail : kuy114@jbnu.ac.kr,minhee@jbnu.ac.kr,paul.kim@dankook.ac.kr,jcho76@sogang.ac.kr]
*Corresponding author : Jaehee Cho

Abstract

This study developed a mobile app for anger management training and recruited 330 participants to use it for three weeks. Pre- and post-surveys were carried out to assess three gratifications obtained (GO), subjective anger reflection, objective state anger, and intent to repeat the program. The major findings reveal that among the three gratifications, knowledge seeking had a considerable impact on rage levels. Post-state rage was found to be inversely associated to information seeking. Subjective anger reflection was favourably influenced by all three forms of gratifications, with knowledge seeking having a rather high effect. All three types of pleasure influenced users' inclination to reuse the application. The mediating impact of objective state anger on reuse intention was identified solely in the connection between information seeking and reuse intention. Subjective anger reflection was found to have a mediation impact in the link between all independent components and reuse intention. The obtained gratifications had significant effects on the intention to reuse, but the mediational effects of users' anger management utility were varied.

Keywords: Anger Management, entertainmnet, information seeking, social utility, state anger, uses and gratifications

1. Introduction

The growth of mobile applications has sparked interest in understanding user behavior and effects. This study focuses on anger management training apps and identifies factors influencing their usage intention. The use and gratification theory is used to explain user motivations and satisfaction derived from apps. Three common motivations explored are information seeking, entertainment, and social utility. The study aims to investigate the effects of gained satisfaction (GO) and state anger levels on intention to reuse the app. A mediation model is proposed, decomposing anger management utility into subjective anger reflection and objective state anger.

2. Literature review

2.1 Theory of use and gratification

The use and gratification theory suggests that

individuals have specific motivations for using media, including information seeking, entertainment, and social utility. Information seeking is a prominent motivation for internet usage, while entertainment and social utility play significant roles in new media services.

2.2 State-anger and Subjective anger reflection

State anger and subjective anger reflection are important aspects of anger management. Various studies have explored anger management programs, including mobile apps. Subjective anger reflection refers to users' reflection on their anger after app usage, which can help regulate negative emotions.

3. Research Question

The study develops research questions to evaluate anger management training apps.
RQ1 examines if gained satisfaction reduces state anger.
RQ2 investigates if gained satisfaction influences subjective anger reflection.
RQ3 explores if gained satisfaction increases intention to reuse the app.
RQ4 analyzes if changes in anger states impact app usage intention.

4. Research methods

The study recruited 330 users to utilize an anger management app for three weeks. Pre- and post-surveys were conducted to measure gained satisfaction, subjective anger reflection, objective state anger, and intention to reuse the app. The State-Trait Anger Expression Inventory-Korean Version and a self-report questionnaire were used.



Fig. 1. Anger management application using humanities classic content.

5. Result

Confirmatory factor analysis confirmed a good model fit. Regarding RQ1, information-seeking satisfaction significantly reduced anger levels, while entertainment and social utility had no significant impact. RQ2 findings revealed that all three gratifications increased subjective anger reflection. RQ3 results showed that all three gratifications positively influenced intention to reuse the app. RQ4 found that both anger indicators mediated the association between gratifications and reuse intention.

6. Conclusions

This study developed an anger management app and investigated motivations and anger-related metrics. Meeting information-seeking satisfaction reduced state anger but increased subjective anger reflection. Objective and subjective anger levels showed contrasting effects on app reuse. To enhance app reuse, managing state anger while promoting reflection is crucial. Developing gamification strategies to minimize frustration and increase willingness to use the app is recommended for companies.

References

[1] Chan, K., & Fang, W., “Use of the internet and traditional media among young people,” Young Consumers, 8(4), 244–256, 2007.
[2] D. W. Hahn, K. K. K. Chon, C. D. Spielberger, “Korean Adaptation of the State-Trait Anger Expression Inventory : Anger and Blood Pressure,” Korean Journal of Health Psychology, Vol. 2, No. 1, pp. 60-78, 1997.
[3] Fatima Alsalem, “Why Do They Post? Motivations and Uses of Snapchat, Instagram and Twitter among Kuwait College Student,” Media Watch, 10 (3), 550-567, 2019.
[4] G. Falgoust, E. Winterlind, P. Moon et al., “Applying the uses and gratifications theory to identify motivational factors behind young adult’s participation in viral social media challenges on TikTok,” Human Factors in Healthcare(2), 2022.
[5] Jamal J. Al-Menayes, “Motivations for Using Social Media: An Exploratory Factor

Analysis," International Journal of Psychological Studies, Vol. 7, No. 1, 43-50, 2015.

[6] J. Y. Cha, "The Effects of Parent-child Relationship Program through Enhancing Anger Management Abilities of Mothers with Preschoolers: An Integrative Approach of CPRT and AL," Journal of Future Early Childhood Education, Vol. 28, No. 2, 79-114, 2021.

[7] J. h. Cho, K. H. Yoo, Y. R. Jang, S.H Kim, "Development of Anger Management Training App Based on Classical Texts: Focusing on State Anger and Depression," Journal of Digital Contents Society Vol. 23, No. 9, 1713-1721, 2022.

Copyright ⓒ 2023 KSII

Investigation of Potential Relationships between Actual Behaviors of Using Smartphone Applications and Trait Anger

Jaehee Cho*
School of Media, Arts, & Science, Sogang University
Seoul, South Korea
[e-mail: jcho76@sogang.ac.kr]
Kyung Han You
Department of Media and Communication Studies, Jeonbuk National University
Jeonju, South Korea
Sewhwan Kim
Dankook Medical School, Dankook University
Cheonan, South Korea
Jin Xianmei
Department of Mass Communications, Sogang University
Seoul, South Korea
Dongbin Kim
Dankook Medical School, Dankook University
Cheonan, South Korea
*Corresponding author: Jaehee Cho

Abstract

This research aimed at exploring potential relationships between individuals' 'actual behaviors' of using smartphone applications and their trait anger. This research obtained smartphone application usage data collected by a research company through a panel for one year in 2021 and measured research participants' trait anger by using 10 items. Data from a total of 527 panels were collected, and smartphone usage data was collected at the interval of '1 week'. 24 major applications were classified based on application genres provided by the App Store and Google Play, and game applications were coded by dividing them into 8 detailed genres. The GLM analyses were conducted to analyze the behavioral and survey data, and for this analyses research participants were categorized into either low or high group of trait anger. The results from GLM provided the following major research findings. There were significant interaction effects of 1) trait anger and age on using applications of 'board game', 'navigation', 'entertainment', 'music', and 'magazine & newspaper', 2) trait anger and gender on uses of 'adventure game', 'navigation', 'photo & video', 'music', and 'magazine & newspaper' applications, 3) trait anger, age, and gender on using 'sports game', 'puzzle game', 'education', and 'travel' applications. Among those results, it was considerable that people's uses of all genres of game applications were higher among people with relatively lower level of trait anger. It was also interesting that while 'book' and 'magazine & newspaper' applications use was more active in the higher trait anger group.

Keywords: Digital phenotype, smartphone application uses, general linear modeling, trait anger

1. Introduction

In Korea, notable increases of society-wide competitions as well as various conflicts has push people to be exposed to a lot of stressors. A considerable point in regards to such increased stressors must be their negative effects on people's psychological conditions such as depression, anger, and so on[1-3]. In the area of media studies, researchers have dig into the close relationships between media uses and negative

psychological conditions[4,5].

Although previous studies have provided scholars invaluable opportunities to thoroughly understand the roles of media uses on individuals' psychological conditions, the main limitation of those studies should be that they have rarely measured people's actual uses of media. In most cases, previous research asked research participants to self-report their perceptions of their own behaviors of using specific types of media [6]. Therefore, this research paid the major attention to understanding how smartphone applications uses would be related to people's trait anger.

2. Digital Phenotype, Smartphone Uses, & Trait Anger

Paying attention to close connections between digital media uses and people's mental conditions, scholars began to argue that thorough monitoring of people's behaviors of using digital media is helpful for predicting their mental conditions [7]. This has been known as the concept of 'digital phenotype'. This present research was also based on the main argumentation of digital phenotype by measuring people's actual behaviors of using smartphone applications as well as trait anger.

As described above, media uses including SNS uses can be negatively associated with individuals' psychological conditions [8]. Anger is also often closely related to people's use of diverse media [9-10]. It should be also considered that people use media in order to overcome their bad feelings [11]. When people feel depression, they tend to listen to peaceful or hardcore musics, play games, watch funny movies, or use social media [12]. This means that media use can be efficient solutions of overcoming bad psychological and mental conditions. Considering these arguments, this present research pursued to explore potential associational relationships between people's behaviors of using smartphone applications and their trait anger.

According to Spielberger and colleagues [13], anger can be categorized into state anger and trait anger. While the former means anger that people are currently feeling under a certain context, the latter is related to people's personality of explaining the extent to which they are likely to feel easily anger in their everyday lives. Therefore, this research explored how people's smartphone uses for their everyday lives would be associated with their trait anger as a specific type of personality.

4. Methods

In order to explore the research question, this research analyzed smartphone application usage data collected by a research company through a panel for one year in 2021. Data from a total of 527 panels were collected, and smartphone usage data was collected at the interval of '1 week'. In addition to such behavioral data 527 research participants, this research measured their trait anger by using 10 items originally proposed by Spielberger and colleagues [13]. An average score for those 10 items was used for the further analyses.

In terms of smartphone applications, 24 major applications were classified based on application genres provided by the App Store and Google Play. In addition, game applications were coded by dividing them into 8 detailed genres. Application uses were coded into 31 genres in total.

5. Results

The GLM(general linear modeling) analyses were conducted to analyze the behavioral and survey data in order to explore the research question. The results from GLM(general linear modeling) provided the following major research findings.

First, it could be found that there were significant differences in the use of applications of 'book'($F=1166.4$, $p=0.019$). The use of those applications was higher among people with higher trait anger.

Second, there were significant interaction effects of trait anger and age on using applications of 'board game'($F=3437.2$, $p=0.011$), 'navigation' ($F=1152.9$, $p=0.019$), 'entertainment'($F=1.81$, $p=0.076$), 'music'($F=321.7$, $p=0.035$), and 'magazine & newspaper'($F=427.4$, $p=0.031$).

Next, there were also significant interaction effects of trait anger and gender on uses of 'adventure game'(F=145.0, p=0.053), 'navigation'(F=91.3, p=0.066), 'photo & video'(F=51.2, p=0.088), 'music'(F=3336.1, p=0.011), and 'magazine & newspaper' applications(F=480.2, p=0.029).

Last, this research could find the interaction effects of trait anger, age, and gender on using 'sports game'(F=2.99, p=0.087), 'puzzle game'(F=5.62, p=0.019), 'education'(F=3.32, p=0.069), and 'travel'(F=11.94, p=0.001) applications.

4. Discussion

This research aimed at exploring how people's use of smartphone applications would be associated with their trait anger. In total, 31 genres of applications were included into the GLM analyses. Among those genres of smartphone applications, ten applications were significantly affected by anger, age, and/or gender. Although this research's main findings cannot explain the causal relationships between smartphone uses and trait anger, they provide researchers with insights for establishing more evidence-driven hypotheses of exploiting smartphone use behaviors' effects on psychological conditions.

References

[1] W. G. Moons, N. I. Eisenberger, and S. E. Taylor, "Anger and fear responses to stress have different biological profiles," *Brain, Behavior, and Immunity*, Vol 23, No. 2, pp. 215-219, 2010.

[2] C. J. Holahan, and R. H. Moos, "Life stressors, personal and social resources, and depression: A 4-year structural model," *Journal of Abnormal Psychology*, Vol 100, No. 1, pp. 31-38, 1999.

[3] C Fiorilli, T. G. Gapitello, D. Barni, I. Buonomo, and S. Gentile, "Predicting adolescent depression: The interrelated roles of self-esteem and interpersonal stressors," *Frontiers in Psychology*, Vol 10, pp. 1-10, 2019.

[4] D. Liu, R. F. Baumeister, C. Yang, and B. Hu, "Retracted: Digital communication media use and psychological well-being: a Meta-analysis," *Journal of Computer-Mediated Communication*, Vol 24, No. 5, pp. 259-273, 2019.

[5] J. M. Twenge, and W. K. Campbell, "Media use is linked to lower psychological well-being: Evidence from three datasets," *Psychiatric Quarterly*, Vol 90, pp. 311-331, 2019.

[6] H. Lee, H. E. Lee, J. Choi, J. H. Kim, and H. L. Han, "Social media use, body image, and psychological well-being; a cross-cultural comparison of Korean and the United States," *Journal of Health Communication*, Vol 19, No. 12, pp. 1343-1358, 2014.

[7] J. Onnela, and S. L. Rauch, "Harnessing smartphone-based digital phenotyping to enhance behavioral and mental health," *Neuropsychopharmacology*, Vol 41, pp. 1691-1696, 2016.

[8] E. A. Vogel, J. P. Rose, L. R. Roberts, and K. Eckles, "Social comparison, social media, and self-esteem," *Psychology of Popular Media Culture*, Vol 3, No. 4, pp. 206–222, 2014.

[9] G. J. Devilly, R. P. O'Donohue, and K. Brown, "Personality and frustration predict aggression and anger following violent media," *Psychology, Crime & Law*, Vol 29, No. 1, pp. 83-119, 2023.

[10] C. Fitzpatrick, M. Binet, E. Harvey, R. Barr, M. Couture, and G. Garon-Carrier, "Preschooler screen time and temperamental anger/frustration during the COVID-19 pandemic," *Pediatric Research*, 2023.

[11] M. Bae, "Coping strategies initiated by COVID-19-related stress, individuals' motives for social media use, and perceived stress reduction," *Internet Research*, Vol 33, No. 1, pp. 124-151, 2023.

[12] L. N. Wolfers, and S. Utz, "Social media use, stress, and coping," *Current Opinion in Psychology*, Vol 45, 2022.

[13] C. D. Spielberger, "*The State-Trait Anger Expression Inventory-2 (STAXI-2): Professional manual*," Odessa, FL: Psychological Assessment Resources, 1999.

Copyright ⓒ 2023 KSII

Comparative Study on the Advertising Effectiveness of Real and Virtual Influencers on Healthcare Products : An Eye-Tracking Approach

Mengqi Du and Kyunghan Ryu*
Department of Media Communication, Jeonbuk National University
Jeonju, South Korea
[e-mail: dumengqi08@gmail.com, kuy114@jbnu.ac.kr]
*Corresponding author: Kyunghan Ryu

Abstract

As millennials are among the main consumers of vitamins, advertisers try to promote their products through influencers who are favoured by young people. In this paper, we explore the relationship between the type of influencer, visual attention, perceived characteristics of the influencer and the advertising effectiveness of healthcare products. To determine this, we conducted an experiment with 120 participants with an average age of 24.7 years. Data on their visual attention when viewing influencer advertisements was collected and analysed using eye-tracking technology. The results confirm that the perceived characteristics of influencers and visual attention have an impact on advertising effectiveness; there are interactive effects of influencer type and visual attention on perceived characteristics and advertising effectiveness. Finally, we discuss the implementation of the proposed influencer marketing for health products and describe its implications and future considerations.

Keywords: Virtual influencer, Healthcare product, Perceived characteristics, Visual attention, Eye-tracking, Advertising effectiveness

1. Introduction

In the wake of global economic growth and demographic change, a new market for consumer-oriented products with significant health benefits is emerging, the so-called "consumer healthcare products" market. Healthcare products include a wide range of products, the most common of which are vitamins, minerals, etc. According to a new study by Mintel, of the 78% of Americans who consume vitamins, one-third (34%) have increased their consumption since the start of the pandemic COVID -19, including nearly half (47%) of Millennials [1]. As young people are among the main consumers of vitamins, marketers are taking a more strategic approach to promoting their brands, especially with the help of influencers, who are favoured by Millennials. In recent years, virtual influencers created using CGI technology, whose appearance and actions are almost indistinguishable from those of real people, have proven to be effective endorsers in marketing [2].

In the traditional advertising study, the source of information conveying a message was considered a key factor in persuasion and it was emphasized that the characteristics of the source influence the effectiveness of the advertisement [3]. In the context of the Internet, where a large number of online advertising stimuli compete for consumers' visual attention, it is worthwhile to investigate how to attract consumers' attention in a limited period of time and what factors in advertising images may have an impact on consumers' attitudes and behavior. Existing research on advertising effectiveness has shown that visual attention is an important factor in the degree of information processing, and offers a new approach to accurately measure advertising effectiveness by detecting vision and perception [4]. Since virtual CGI influencers are very similar to real people, they are more difficult to

identify than real influencers, which complicates the process of information intake. Therefore, the user's visual attention when viewing an advertisement may also vary depending on the type of influencer, which may affect the evaluation of the influencer's characteristics and advertising effectiveness.

2. Research Questions and Methods

2.1 Research Questions

[RQ1] How do user demographics (gender, income), influencer factors (type, gender), perceived influencer characteristics (attractiveness, authenticity, social presence), and visual attention influence advertising effectiveness (advertising attitude, brand attitude, purchase intention)?
[RQ2] Is there an interactive effect of influencer type (real vs. virtual) and visual attention on perceived influencer's characteristics?
[RQ3] Is there an interactive effect of influencer type and visual attention on advertising effectiveness?

2.2 Experimental Design and analysis methods

We conducted the experiment with 120 undergraduate students living in Korea. Based on manipulation checks, 2 real influencers and 2 dummy influencers were selected with similar likability scores. The experimental setups of the four selected species are shown in Figure 1. In the experiment, 30 participants were randomly assigned to the four types of advertising presentation (A, B, C, and D). For the product, vitamin C was chosen as the experimental stimulus and "UISICARI" was used as the fictitious brand to exclude the influence of previous experience. In this study, the total time of fixation duration for each area of interest (AOI) was chosen as a measure of visual attention. The AOIs were set to the influencer's face, product, and ad copy. As part of an experiment to obtain data, Tobii Pro Lab was used to track participants' visual attention while viewing influencer advertisements. After the experiment, demographic information, Instagram usage characteristics, perceived influencer characteristics, and ad effectiveness were confirmed through a survey.



Fig. 1. Selected virtual and real influencer models

3. Results

3.1. The Perceived Characteristics of Influencer and The Impact of Visual Attention on Advertising Effectiveness

Perceived authenticity (B=.356, SE =.111, p <.01), social presence (B=.455, SE =.075, p <.001), and visual attention to the product (B=.000, SE =.000, p <.05) were found to have a positive effect on ad attitude. Second, significant results were found for attractiveness (B=.215, SE =.086, p <.05) and authenticity (B=.283, SE =.094, p <.01) in terms of their effects on brand attitude. Third, perceived authenticity of the influencer (B=.212, SE =.083, p <.05) and advertising attitude (B=.235, SE =.091, p <.05) were found to have a significant positive effect on purchase intentions, while influencer type (B=-.31, SE =.129, p <.05) had a negative effect.

3.2. The Interaction Effect of Influencer Type and Visual attention on Perceived Characteristics

Influencer type* fixation time (B=-.001, p <.001) had a significant effect on perceived attractiveness. As time increased, the rating of the influencer's attractiveness also increased. The interaction term between influencer type and fixation time had a significant effect on authenticity (B=-.001, p <.001) and social presence (B=-.002, p <.001). As visual attention increased, evaluations of authenticity and social presence of the real influencer increased sharply, whereas the opposite was true for virtual influencers. However, we did not find an interaction effect between the type of influencer

and the fixation time of the product on characteristics.

3.3. The Interaction Effect of Influencer Type and Visual Attention on Advertising Effectiveness

It was found that there was a significant interaction effect between fixation time on the face of the influencer and influencer type in terms of ad attitude (B=-.001, p <.001), brand attitude (B=-.001, p <.001), and purchase intention (B=-.001, p <.001). With increasing fixation time on the face of a real influencer, the evaluation of advertising effectiveness increased sharply, while for a virtual influencer, a gradual decrease was observed with increasing fixation time.

4. Conclusions

In this study, we identify the factors that influence the effectiveness of advertising from the perspective of visual attention, which can be considered a new finding in the study of the advertising effectiveness of virtual influencers.

First, we confirm that authenticity has a significant positive impact on ad attitude, brand attitude, and purchase intention when advertising healthcare products. We also found that the more virtual the influencer, the more negative the impact on purchase intention. Among the factors of visual attention, more than the face of the influencer, the ad copy, attention to the product are the key to forming a good ad attitude. Therefore, advertisers should pay more attention to the visual representation of the products in the advertising images, and attracting consumers' visual attention to the products is important for forming a good advertising attitude. Besides, real influencers are more persuasive than virtual influencers as endorsers for healthcare products.

Second, at different levels of visual attention, depending on the type of influencer, there were differences in participants' perceived characteristics of the influencer and evaluations of the ad effectiveness. Since more information about the stimulus can be obtained when attention is focused on the stimulus for a longer period of time, this helps participants to make an objective and comprehensive evaluation. at higher levels of information processing, the identity of virtual influencers is recognised. If participants perceive that a virtual influencer, rather than a human being, is promoting a dietary supplement, they may negatively evaluate the reliability and authenticity of the advertisement. Therefore, to improve the evaluation of influencer characteristics and advertising effectiveness, advertisers must strategically direct consumers' visual attention to the face of the influencer, depending on the type of influencer. These findings provide efficient marketing strategies and basic frameworks for businesses in today's visualisation-based influencer marketing environment.

References

[1] Mintel. "Vitamin boost: Half of Millennials are using vitamins and supplements more often than they did before the pandemic" Retrieved from: https://www.prnewswire.com/news-releases/

[2] de Brito Silva M J, de Oliveira Ramos Delfino L, Alves Cerqueira K, et al. "Avatar marketing: A study on the engagement and authenticity of virtual influencers on Instagram"[J]. *Social Network Analysis and Mining*, 2022, 12(1): 130.

[3] Yilmaz, Cengiz, et al. "Source characteristics and advertising effectiveness: The roles of message processing motivation and product category knowledge," *International Journal of Advertising* 30.5 : 889-914, 2011.

[4] S J. Hwang & G. Kim, "Analyzing consumer response to fashion brand advertisements using eye trackers," *Brand Design Research*, 11(4), 203-220, 2013.

Copyright ⓒ 2023 KSII

Design of Location-based Traveler-tailored Crime Information Fusion Application

Youjin Seo and Junho Ahn
Department of AI Computer Science and Engineering, Kyonggi University
Suwon, South Korea
[e-mail: youjinseo1@gmail.com, jha@kyonggi.ac.kr]
*Corresponding author: Junho Ahn

Abstract

The proposed study is an application UI/UX study to increase user safety and efficiency when traveling. When traveling, both domestically and internationally, users actively use applications to get information about travel destinations, airline tickets, accommodation, and restaurants, but the most important safety information, such as crime information, is missing. This paper aims to enable users to travel safely by automatically providing local crime information based on the user's location when the user searches for the necessary information in the application.

Keywords: Crime, Location-based services, Intelligent system, Traveler, UX, Application Design

1. Introduction

With the recent lifting of the COVID-19 emergency state, physical movement has become more frequent, not only within the country but also internationally. When visiting or traveling to new places, mobile users seek information through real-time search. They obtain the information they need by searching keywords on websites or applications for tourist attractions, restaurants, and transportation facilities they want to visit. However, when visiting a new place, the keywords they search for are limited because they have limited prior information about the place and may be missing information that the user has not thought of. Information related to safety is missing in the information users usually search for, even though it is the most important information. In a situation where the state of emergency has been lifted and people are actively moving between countries, safety information is crucial because it is related to life and death. Travel information services that provide information based on user searches or existing user information are limited because they focus on providing users with the information they are interested in. With the application proposed in this study, an intelligent service is proposed to provide tailored safety information when danger is detected by automatically collecting the user's information based on GPS and collecting crime information in the area through the police agency's cloud server when the user uses the application, even if the user does not search for the information. This study aims to design an application that provides smartphone app-based customized services for travelers by fusing sensors such as activity sensors, audio sensors, GPS information, smartphone touch sensors, and Wi-Fi and LTE/5G embedded in a smartphone, and integrating crime-related information collected from the police agency's cloud server and a cloud server that manages internal personal information.

2. Research and Investigation

When traveling, travelers usually search for information such as accommodations, plane tickets, and directions to their destinations. We will take a look at widely used services among applications for travelers.

2.1 Case studies of applications that travelers mainly use

When traveling to a new place, users of conventional websites and smartphones use websites and applications to find the information they need. According to the ranking of travel-related applications in the App Store [1], the most downloaded apps are for booking accommodations, flights, and vehicles. Some paid services provide more detailed information, such as vegan restaurant search results, camping spot search results, and translators. We will analyze three of the most popular applications used by travelers. [2]

Google Maps is a popular application used by travelers. The typical reason why users use Google Maps is its reliability.[3] When the user selects an area, it provides information about it. If the user selects Wikipedia, descriptions such as history and photographs are shown. If the user selects a starting point and a destination when looking for directions while traveling, he/she can easily find the directions through arrows. In addition to descriptions of the area and information about its attractions, it also provides information about accommodations and even offers a booking service. Google Maps provides indispensable services as an application that gives trust to its users with accurate maps.

Agoda [4] is a hotel reservation application. It provides information on various types of accommodation, and it is convenient to compare prices. It is easy to find deals or inexpensive accommodations, and it is trustable since it has a lot of information and reviews. However, it does not provide as much comprehensive information as Google Maps or other travel applications.

The Dining code [5] uses big data technology to collect restaurant information from websites such as Naver blogs, Daum, and T-story and create a list of highly rated restaurants. It provides personalized restaurant recommendations to users based on their location and preferences. Furthermore, it improves reliability by reflecting reviews written by restaurant visitors.

3. Research Description

3.1 Proposed information service ideas that reflect user experience

Security and safety are the most important issues when travelers travel. Traditional applications have focused on providing information for directions, accommodations, flight tickets, restaurants, and tourist attractions. Users need to download and use a specific application to search for local crime information.[6] However, they do not search for crime information when they search for destination information or looking for highly rated restaurants. [7]

My proposed application is designed to collect GPS information from the smartphone while the user is searching for information of interest, and simultaneously collect information from the local police website to provide the necessary information in real-time. It aims to provide tailored information related to the safety in the area the user is visiting based on the user's age or categorized group using geolocation sensors. [8]



Fig. 1. Architecture design

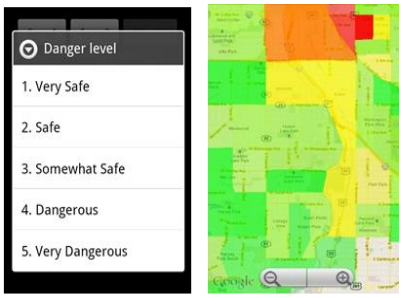


Fig. 2. UI design for crime rates on map based on mobile app

The image below shows a map of Boulder, Colorado, USA, divided into five risk levels based on the crime rate information from the Colorado Police Department cloud. The safest areas are visualized in green and the most dangerous in red, immediately allowing users to see quickly which areas are dangerous and which are safe. By fusing crime information received through the police cloud server and sensors such as smartphone built-in activity sensors, audio sensors, GPS information, smartphone touch sensors, Wi-Fi, and LTE/5G, intelligence-based user-tailored safety information will be provided. The service that provides real-time location-based safety information when users search for travel information protects travelers, and it is efficient because it automatically alerts users without having to search separately for safe travel.

4. Conclusion

The combination of user information, location information, big data, and artificial intelligence (AI) has become technologically stable, and users want to receive more accurate and customized services for different user situations and propensities. In this demand, an application service that can provide the user's location information, real-time data of the located area, and crime information in the area in a customized and efficient manner is proposed. When moving or traveling to a new place, this application automatically provides not only useful information such as directions, search research results for tourist attractions, and restaurant/accommodation/transportation selection, but also essential but easy-to-miss information related to life and death such as safety-related crime information and incidents and accidents occurring in real-time. Going beyond applications that provide information that users search for or are curious about, it can become a security measure to not only provide the information they need in a customized manner, but also to supplement the information they have not searched for. The ever-evolving AI technology can rapidly accelerate the research, development, and commercialization of new and useful applications. Location-based customized safety information services using AI can be used not only for travelers but also for people who need help in real life, such as children and patients when they go out.

Acknowledgement

This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research]

References

[1] App Store, https://apps.apple.com/us/charts/iphone/travel-apps/
[2] Yunsik Kim, et al. Safety report app installation and safety report activation plan research, TRKO201900000547, Ministry of Public Safety and Security, 2016-12
[3] Travel Tips, Why Traverllers Should Use Google Maps, Travel and Destinations, https://www.travelanddestinations.com/why-travellers-should-use-google-maps/
[4] Agoda, https://www. agoda.com
[5] Dining code, https://www.diningcode.com/
[6] Competitive Matrix and Research, Crime Finder Application, https://nicholash.com/crimes.html
[7] Kosis , https://kosis.kr/statHtml/statHtml.do?orgId=132&tblId=DT_13204_3101
[8] Andy Yuan Xue; Rui Zhang; Yu Zheng; Xing Xie; Jin Huang; Zhenghua Xu, Destination prediction by sub-trajectory synthesis and privacy protection against such prediction, 2013 IEEE 29th International Conference on Data Engineering (ICDE), 2013

KSII The 18th Asia Pacific International Conference on Information Science and Technology(APIC-IST) 2023.
Copyright ⓒ 2023 KSII

Exploring the Relationship between Metaverse Usage Patterns and Cultural Tastes : A Study of Avatar Formation among Generation Z

Hyun Ah Park and Kyung Han You*
Department of Media Communication, Jeonbuk National University
Jeonju, South Korea
[e-mail: nunbyung@naver.com, kuy114@jbnu.ac.kr]
*Corresponding author: Kyung Han You

Abstract

The metaverse is a new interactive space that allows users to engage in a wide range of activities using advanced technologies. It is seen as an extension of reality, particularly by Generation Z, who perceive no significant difference between their virtual and physical selves. This study explores the metaverse from a cultural sociology perspective, analyzing how users' cultural tastes influence their avatars and experiences within the metaverse. The research conducted focus group interviews and individual in-depth interviews with Generation Z users.
The study found that while Generation Z users operate freely in the metaverse, their behavior is influenced by their socio-economic status. Cultural tastes, shaped by various experiences with their parents, play a significant role in their engagement with cultural activities in the metaverse. The recognition of avatars was categorized into Desire satisfaction type, Image exaggeration type, and expanded self type, with higher cultural capital associated with perceiving avatars as an extension of the self. Participants' perceptions of the metaverse varied, with some viewing it as a fantasy world or a space to do anything.

Keywords: Metaverse, Cultural Taste, Avatar, Z generation

1. Introduction

Can the metaverse be considered an extension of reality? The metaverse is gaining increasing attention as a novel interactive space, and it is anticipated that it will enable users to experience a wide range of activities, including economic transactions, using enhanced 3D graphics, augmented reality, and cryptocurrency technologies. Notably, Generation Z, who have grown up in the era of ubiquitous technology, perceive the metaverse as their playground and attribute no significant difference between virtual and physical selves [1].

Hence, this study acknowledges the significance of the metaverse environment as a media environment shift and aims to analyze the metaverse through the lens of cultural sociology, building upon Bourdieu's discourse on cultural tastes. If the metaverse can be regarded as an extension of reality, the cultural preferences of reality would impact the avatar, which represents oneself in the metaverse.

2. The avatar and cultural taste of Generation Z in the metaverse

Over the past half-century, research on cultural taste has been divided into two major streams. The first hypothesis is Pierre Bourdieu's theory of cultural taste, which argues that an individual's tastes are not innate but are shaped by their social and economic position. According to this theory, individuals in higher social classes consume highbrow culture, while those in lower classes

consume lowbrow culture. The second hypothesis is the omnivore-univore hypothesis proposed by Peterson [2], which suggests that individuals with higher social class not only enjoy high culture but also have a broad and diverse taste for culture, resulting in a flattening of the taste hierarchy.

Taste is formed, developed, and refined through socialization in childhood, interaction with cultural and educational institutions, and education, which is the most important cultural transmission format. The lifestyle of modern individuals is closely related to their social background, educational capital, and cultural tastes [3]. In a democratic society where visible social distinctions are abolished, cultural taste plays an important role as a mechanism for class formation [4].

Taste has the power to create an unconscious unity of the class, reinforcing social divisions and functioning as an exclusive resource based on one's social position. Those with higher cultural tastes tend to prefer socializing with those who share their taste, while those with lower cultural tastes may not recognize the social reproductive significance of their preferences. It is important to note that current power relations operate through unconscious mechanisms that are not readily apparent in cultural and symbolic dimensions [5].

In the metaverse, an individual's cultural taste is implemented through avatars, which are closely related to the real-world self in the technologically advanced era. Avatars are an extension of the self that the user experiences in virtual space [6]. The real-world self and the avatar are interdependent, symbiotic relationships intertwined with shared experiences that influence the properties and experiences of the avatar.

Generation Z is a major user of the metaverse, actively engaging in online activities. I

They were born and raised with current technology, and technology is their identity [7]. To them, virtual reality (VR) is not just science fiction, but rather a part of their daily lives. Their "digital fluency" is a characteristic that has grown with the advancement of technology, often connecting to virtual environments, and technology plays a significant role in shaping their identity [7].

3. Research methodology

To achieve the research objectives, this study conducted Focus Group Interviews (FGIs) and individual in-depth interviews with Generation Z, who are the primary users of the Metaverse. The interviews aimed to investigate how Generation Z's avatar formation and usage behavior in the Metaverse is influenced by their socio-economic status and cultural tastes. Furthermore, the study explored how their cultural tastes, shaped by diverse cultural experiences since childhood, manifest in their engagement with the Metaverse.

4. Cultural taste and avatar types

First, it was found that cultural tastes influence the social positioning of the Generation Z users who use the metaverse. Users with higher levels of cultural tastes demonstrated a clear distinction in terms of their manners, creative use of the metaverse, and adherence to cultural tastes, habits, and attitudes. Users with higher cultural tastes tend to exhibit better manners, more creative use of the metaverse, and a broader range of cultural tastes and experiences. However, access to cultural experiences and opportunities still tends to be skewed toward a minority of people. The study also found that the cultural tastes of Generation Z users are influenced by both their own cultural tastes and that of their parents and that their avatars tend to reflect a fluid combination of the two. In conclusion, the factors that determine cultural tastes are complex and multidimensional, involving both Bourdieu's theory of cultural tastes and Peterson's Omnivorousness hypothesis.

Secondly, individual differences were found in avatar recognition in the metaverse according to personal cultural tastes. Avatar recognition was classified into three types: "Desire-satisfaction type" in which individuals projected unfulfilled desires onto their avatars, "Image exaggeration

type" in which individuals experimented with new identities unrelated to their real-life selves, and "Extended self type" in which individuals recognized their avatars as an extension of themselves in reality. Participants with higher levels of cultural tastes tended to recognize their avatars as an "expanded self." This group maintained a respectful attitude even in the virtual world where their faces were not exposed and utilized the metaverse creatively, not just for gaming purposes.

Avatars of the "Expanded self-type" exhibit highbrow cultural tastes and share the characteristic of constructing and maintaining their identity similarly in both reality and the metaverse. Users of the "Desire-satisfaction type" express the desires they wanted to achieve in reality through avatars in the metaverse, with intermediate cultural taste levels. On the other hand, "Image exaggeration type" users tend to present avatars that are unrelated to reality, appear as exaggerated images, and have cultural tastes that are mostly mainstream. In other words, it was discovered that when the cultural taste is mainstream, users tend to use avatars in a manner that is too estranged from reality, while users with highbrow cultural tastes tend to utilize avatars in a way that connects both.

5. Conclusions

This study challenges the notion that cultural and sociological research is no longer relevant in a postmodern era where taste hierarchies have supposedly disappeared. Our findings demonstrate the relevance of cultural sociology in understanding the metaverse and its users.

Generation Z users of the metaverse exhibit varying types of avatar and metaverse perceptions according to their cultural tastes, despite the metaverse's appearance of relative autonomy.

Users with higher cultural tastes were found to utilize various technologies available in the metaverse in a more diverse and innovative manner beyond gaming. It was also revealed that cultural biases inherited unconsciously from parents during upbringing persisted in their metaverse activities. The study participants with greater cultural tastes had access to more cultural and educational experiences during their formative years, and these cultural factors were observed to be reproduced at home. It has been found that cultural experiences and perspectives still vary depending on the environment in which people grow up, which is largely determined by their parents. This suggests that unconsciously acquired artistic experiences play a role in reproducing the social class structure.

This study is significant in that it takes a user-centered approach to the study of the metaverse, which has mainly been studied from a technological and industrial perspective, and expands the scope of the real world into the virtual space of the metaverse from a socio-cultural perspective.

References

[1] Filho, E. J. M. A., Gammarano, I. D. J. L. P., Barreto, I. A., “Technology-driven consumption: digital natives and immigrants in the context of multifunctional convergence,” *Journal of Strategic Marketing*, Vol 29, No. 3, pp. 181-205, 2021.

[2] Peterson, R. A., “Understanding audience segmentation: From elite and mass to omnivore and univore.” *Poetics, 21*(4), 243-258, 1992.

[3] Hyun Tak Soo, Munhwa wa kwonruk, Nanam, 1998.

[4] Lee, H.Y., Seo, “Rethinking Cultural Capital and Inequality in the Digital Age,” *The Journal of Cultural Policy* 23. pp. 69-95, U.S. 2010.

[5] Bourdieu, P., Distinction: A social critique of the judgement of taste. Harvard university press, 1987.

[6] Behm-Morawitz, E., “Self-presence and the effects of the avatar on health and appearance,” *Games for Health: Research, Development, and Clinical Applications, 2*(1), 50-52, 2013.

[7] Gaidhani, S., Arora, L., & Sharma, B. K., “Understanding the attitude of generation Z towards workplace.” *International Journal of Management, Technology and Engineering, 9*(1), 2804-2812, 2019.

Copyright ⓒ 2023 KSII

Multi Object Action Classification using Vision Transformer-based YOLO

Seo-El Lee[1], So-Hyun Park[2], Joo-Chang Kim[3], Sangeun Ahn[4], Kyungyong Chung[5], Hyun Yoo[6*]
[1]Department of Pulbic Safety Bigdata, Kyonggi University, Suwon, South Korea
[2,5]Division of AI Computer Science and Engineering, Kyonggi University, Suwon, South Korea
[3,4,6*]Contents Convergence Software Research Institute, Kyonggi University, Suwon, South Korea
[e-mail: [1]lse_1031@kyonggi.ac.kr, [2]pshsh0516@naver.com, [3]kjc2232@naver.com,
[4]sangeun1804@kyonggi.ac.kr, [5]dragonhci@gamil.com, [6]rhpa0916@gmail.com]
*Corresponding author: Hyun Yoo

Abstract

Due to the widespread deployment of CCTV cameras, it is impractical for humans to monitor and manage all CCTV footage, and capturing all events in real-time is also challenging. Therefore, we propose a method for detecting and classifying the behaviors of multiple objects appearing in video using a ViT-based object detection model. To classify behaviors as a set of poses, we train a ViT-based object detection model on pose image datasets. Then, we use an object-tracking model to generate pose vectors for each object for tracking purposes. We train a pattern classification model on the generated pose vectors for each object in each frame of the video to classify the behaviors of each object. This allows for real-time video analysis in areas such as transportation, social safety, and crime and enables rapid response to emergencies that may arise when two or more objects appear in a single frame.

Keywords: Human Action Classification, Pattern Analysis, Vision Transformer, Video Analysis

1. Introduction

Crime and security issues are among the most important in modern society. In particular, as urbanization progresses, the importance of CCTV monitoring systems increases as the number of crimes and accidents occurring in cities increases. However, the widespread use of CCTV cameras incurs high economic costs to manage them all directly. There are limitations to detecting and dealing with incidents occurring in multiple locations simultaneously. Therefore, research has emerged combining CCTV monitoring and object detection and tracking technologies using artificial intelligence in crime and security [1]. Various points must be considered to detect criminal and accident behaviors of objects in videos. First, crimes and accidents between two or more people occur more frequently than those caused by individuals. Therefore, to detect behaviors, it is necessary to understand the interrelationships between various objects appearing in the video and between objects and their surroundings and objects. In addition, videos have a temporal context as well as spatial information.

K. Muhammad et al. classify object behavior in videos using attention mechanism-based bi-directional long short-term memory (BiLSTM) and a dilated convolutional neural network (DCNN) [2]. Feature extraction is performed through DCNN, which is used as input to BiLSTM to ultimately derive the behavior of objects in the video, considering the temporal context. Although this method shows high accuracy, it has limitations in that it exhibits low accuracy for similar context actions such as diving and jumping.

This study proposes Multi-Object Action Classification using Vision Transformer-based YOLO to address this issue. Each object's pose pattern vector is extracted using a Vision Transformer-based YOLO and object tracking model DeepSORT in the video. This is then used

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

as input to a classification model of the LSTM series to classify the behavior of objects appearing in the video.

This paper is organized as follows. Section 2 describes the trend of behavior classification models, Section 3 describes the proposed method, and Section 4 concludes the paper.

2. Related Work

2.1 ViT-YOLO

The original YOLO (You Only Look Once) has limitations in detecting objects in images with diverse scales and complex backgrounds. Therefore, ViT-YOLO [3] emerges as a combination of the Vision Transformer, which allows for information preservation through dependencies between image feature patches, and YOLO for object detection. This model integrates multi-head self-attention into the CSP-DarkNet backbone of YOLO. This enables the model to learn various feature scales, utilizing not only local information but also global contextual information, resulting in the accurate detection of multiple objects in the image.

2.2 DeepSORT

DeepSORT (Deep Simple Online and Real-Time Tracking) [4] is a deep learning-based object-tracking algorithm that works with object detection models. Objects detected in the first frame are identified using an object detection model. The Kalman filter is then used to predict the positions of objects in the next frame. A deep learning-based association metric is employed to determine the consistency of objects between frames and decide whether to track them. DeepSORT performs real-time object tracking and is robust to occlusions and noise.

2.3 BiLSTM

BiLSTM [5] is a model that can process time-series data bi-directionally. It uses two independent LSTMs to capture long-range dependencies in sequences in both forward and backward directions. The outputs from the two LSTMs are concatenated to generate the final output. This enables the model to utilize information from both the past and future to predict and classify the current time step.

3. Multi Object Action Classification using Vision Transformer-based YOLO

This paper proposes a model that uses a Vision Transformer-based YOLO to classify the actions of multiple objects. "Pose" refers to the stationary pose of an object, and "action" represents the collection of such poses that occur across various frames. Therefore, an image dataset of these poses is trained on ViT-YOLO. ViT-YOLO extracts features for each pose and learns them. Afterward, DeepSORT tracks multiple objects in the video to extract action pattern vectors for each object. The action patterns consist of class numbers representing the pose of each object in each frame. The video is time-ordered sequential data. Hence, the action pattern data is input to BiLSTM to classify what actions the evolving set of poses represents for each frame. **Fig. 1** shows the process of the proposed model.

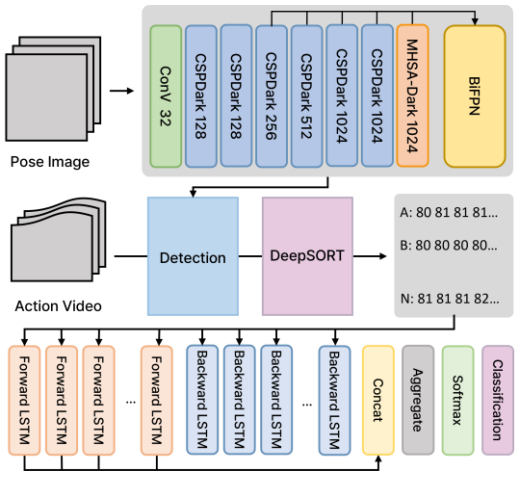


Fig. 1. The process of the multi object action classification using vision transformer-based YOLO

The dataset used for training the ViT-YOLO model on pose recognition is derived from the COCO dataset [6], explicitly focusing on the 'person' class. This dataset contains various images where human poses can be clearly distinguished. To prepare the dataset, the objects labeled as 'person' in the images are assigned

new class numbers corresponding to different poses. This labeling enables the model to learn and recognize different pose variations during training. In addition to the pose-labeled image dataset, a person action video dataset provided by AI Hub is utilized to further enhance the model's understanding of pose patterns. This video dataset contains sequences of frames capturing various human actions. These videos serve as a valuable resource for generating pose pattern vectors. The integration of ViT-YOLO with DeepSORT allows for robust object tracking in videos while considering the temporal consistency of object positions. Despite changes in frames and object positions, the combined algorithm effectively identifies and tracks connected objects throughout the video. This approach enables the capture of the association between objects appearing on a frame-by-frame basis.

To represent the detected poses, the class numbers outputted by the ViT-YOLO model are used to form pose pattern vectors for each tracked object. These vectors encapsulate the sequential information of pose variations exhibited by the objects across frames. Furthermore, action label values are assigned to the constructed pose pattern vectors. This step associates each pose pattern vector with a specific action performed by the object in the video. The assigned action labels provide the ground truth for training a pattern recognition model, specifically the Bidirectional Long Short-Term Memory. The BiLSTM model is trained using the pose pattern vectors to classify the actions performed by each object in the video. By considering the sequential nature of the pose patterns, the BiLSTM model learns to capture the temporal dependencies and variations in the objects' actions, leading to accurate action classification.

4. Conclusions

In this paper, we proposed multi object action classification using vision transformer-based YOLO. We trained ViT-YOLO on a dataset consisting of images from the 'person' class in the COCO dataset, focusing on distinguishing different poses. By assigning new class numbers to labeled 'person' objects, ViT YOLO was enabled to learn pose-specific features. To capture temporal dependencies and track objects across frames, DeepSORT was integrated with ViT-YOLO. This allowed us to associate objects and extract pose pattern vectors based on their trajectories. These vectors, representing the temporal evolution of poses, were then fed into a BiLSTM for action classification. This study contributes to the field of computer vision by addressing the challenging task of recognizing the actions of multiple objects in video sequences. Particularly, it has the potential to make significant contributions to CCTV surveillance systems by enabling real-time detection of abnormal and risky behaviors, allowing for prompt intervention and prevention of serious damages.

References

[1] H. Yan, B. Hu, G. Chen and E. Zhengyuan, "Real-time continuous human rehabilitation action recognition using OpenPose and FCN," in *AEMCSE*, Shenzhen, China, pp. 239-242, 2020.

[2] K. Muhammad, Mustaqeem, A. Ullah, A. S. Imran, M. Sajjad, M. S. Kiran et al., "Human action recognition using attention based LSTM network with dilated CNN features," in *Future Generation Computer Systems*, vol. 125, pp. 820-830, 2021.

[3] Z. Zhang, X. Lu, G. Cao, Y. Yang, L. Jiao et al., "ViT-YOLO: Transformer-Based YOLO for Object Detection," in *Proc. of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 2799-2808, 2021.

[4] S. Kapania, D. Saini, S. Goyal, N. Thakur, R. Jain et al., "Multi Object Tracking with UAVs using DeepSORT and YOLOv3 RetinaNet Detection Framework, in *Proc. of the 1st ACM Workshop on Autonomous and Intelligent Mobile System*, no. 1, pp. 1-6, 2020.

[5] Z. Hameed and B. Garcia-Zapirain, "Sentiment Classification Using a Single-Layered BiLSTM Model," in *IEEE Access*, vol. 8, pp. 73992-74001, 2020.

[6] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona et al., "Microsoft COCO: Common objects in context", in *ECCV*, Zurich, Switzerland, pp. 740-755, 2014.

Copyright ⓒ 2023 KSII

An Architecture for Content Search in Video Retrieval System

Dinh-Lam Pham[1,2], Sang-eun Ahn[2], Kyunghee Sun[2], Kyong-Sook Kim[2], Kwanghoon Pio Kim[2,*]
[1]Information Technology Institute, Vietnam National University, Hanoi
144 Xuan Thuy street, Cau Giay district, Hanoi, Vietnam.
[2]Contents Convergence Software Research Institute, Kyonggi University
154-42 Gwangkyosan-ro, Youngtong-gu, Gyeonggi-do, 16227, Republic of Korea
[e-mail: phamdinhlam@vnu.edu.vn, sangeun1804@kgu.ac.kr, sunkh0507@gmail.com, khmjmc@kgu.ac.kr, kwang@kgu.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Abstract

As we know, searching and retrieving content in video data generated from CCTV are consuming tasks and require the concentration of the officer in charge of managing and controlling these systems. With the rapidly developed machine learning techniques, several strategies were introduced to support these kinds of activities, and hence improving the accuracy of search results and getting faster response time. This paper presents an approach and architecture for searching content in CCTV videos using human natural language through a repository containing extracted features in each frame of the video. More precisely, we propose a method using deep learning and computer vision algorithms to detect objects and their features from each frame in the video generated by the CCTV cameras. Subsequently, the entire data is stored in a JSON file format following a specific structure to organize information on each frame in the video. Later, in the searching procedure, for retrieving the content in the video with human natural language, we use some natural language processing (NLP) techniques to analyze the input text (in English, string format) and then shift the data to JSON file to retrieve a list of relevant frames in the video.

Keywords: Video retrieval system, Video frame features extraction, CCTV camera, Deep learning-based video retrieval system.

1. Introduction

In our model life nowadays, security cameras are ubiquitous and widely used in several areas for keeping people out of potential risks, monitoring traffic, support in emergency response with the aim of improving public security. Accompanied by the popularity of CCTV cameras, there is a growing demand for managing and controlling these systems effectively as such activities require time consuming through a huge amount of data generated every hour [1]. An effective solution to address this matter contributes to improving public security management as providing high accuracy of return information retrieval results meanwhile reduces time consuming [2]. Numerous approaches were proposed for the purpose of developing functions that support retrieval content in CCTV video, especially the approaches using deep learning techniques [3] [4][5].

This paper proposes an architecture for searching content in CCTV video using natural human language through a set of data files containing features of frames in the videos that are generated using deep learning and computer vision techniques. The system promises to be an effective solution that contributes to improving public security management.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number 2020R1A6A1A03040583)

We organize the remaining of the paper as follows. The next section discusses some research studies related to the topic of content searching and video content retrieval systems. In section 3, we describe the proposed architecture and functions in detail. Finally, section 4 gives some conclusions of the paper and future research direction.

2. Related Works

In [6], Muhling et al. presented a method using deep learning techniques to support professional media production. They introduced some algorithms in combination with similarity search, face detection, face recognition, and face clustering for retrieving content in videos. Furthermore, in the study, they described a novel multi-task learning CNN for concurrent concept detection and similarity search. Results of the study demonstrated an excellent quality of retrieving content with mAP 90% of detecting 58 concepts in the videos, and the similarity search results achieved 93.5% mAP.

In [7], Rosetto et al. conducted a survey on interactive video retrieval methods presented in the Video browser showdown (VBS) 2019. Six different methods presented in the VBS 2019 were analyzed to determine the main strategies, popular features, and problems retrieving content from video files. The analysis indicated that textual search and content browsing are crucial user experience interactions. A shift towards deep learning-driven features for detecting and classifying content from frames in the video, especially using neural networks for creating labels of objects in the videos.

In [8], Lin et al. introduced a study using cloud-based system for retrieving frames containing identified face in videos. They trained three different models (VGGFace, ArcFace, and FaceNet) and choose the most appropriate model (FaceNet with accuracy of 99.96% and 4.6 millisecond per image) to implement the cloud-based face video retrieval system.

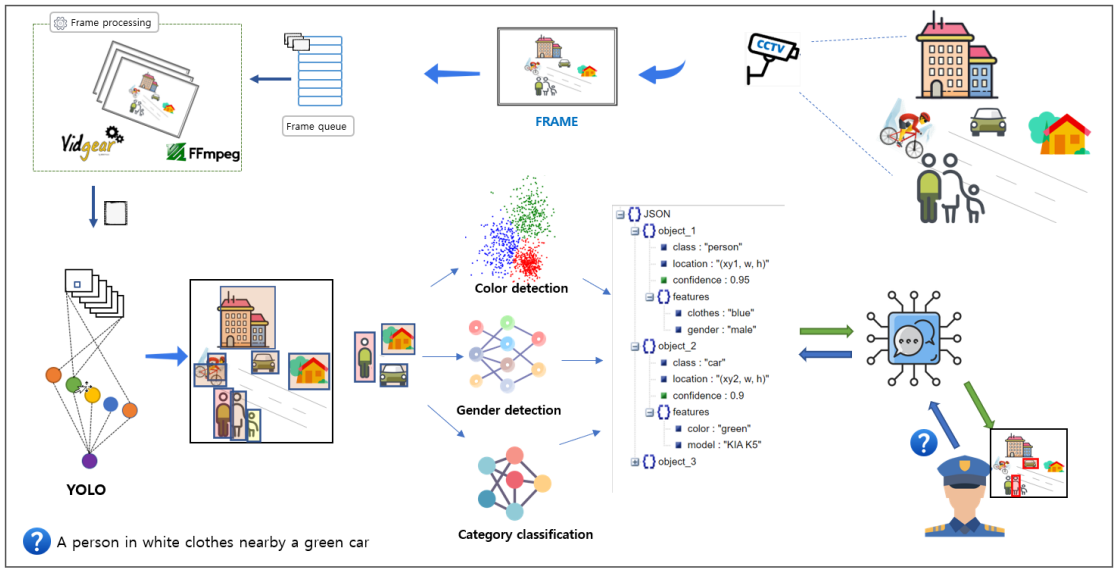


Fig. 1. Proposed architecture for content search in video retrieval system

3. Proposed Architecture

Figure 1 describes the proposed architecture for searching video content with detected object features in detail. As we can see in the architecture, to process searching content with natural human language, the officer put the searching sentence to the natural language processing module to classify nouns and adjectives in the sentence to get the name of objects and their features for retrieving in JSON data. The JSON data file is created following the JSV data format described in [9] to store all the detected objects and features. For creating the JSON file containing object's location, we use the YOLO algorithm to detect objects in frames

after getting them from the Frame processing module. After that, these objects will be put into different modules to detect features of color, gender, and category. To classify the object's primary color, we convert the object's bounding box to HSV color space to separate color information from brightness and intensity and then apply the K-Means clustering algorithm to classify main colors of the flattened pixel values of the bounding box. We use the VGG16 neural network for creating models to classify the gender of persons and category models of vehicles.

5. Conclusions

This paper proposed a methodology combination of some different techniques for searching content in a video retrieval system. The proposed architecture consists of YOLO neural network for detecting objects and their location (bounding boxes) on the frame. After that, different methods were applied to detect features of different objects depending on their classes. Consequently, detailed information about each object is stored in a JSON file following a specific structure. Later, a natural language processing module is used to classify input text from the officer managing the CCTV system in the searching function. Notably, the natural language processing module analyzes the nouns and adjectives from the input text and then passes this query to the JSON file to locate the video's frame position.

In future studies, we will implement a system based on this architecture to bring the proposal into practice and evaluate its effectiveness. Using this architecture, we aim to create a comprehensive system for helping officers to manage CCTV camera systems and improve public security management.

References

[1] N. Spolaôr, H. D. Lee, W. S. R. Takaki, L. A. Ensina, C. S. R. Coy, and F. C. Wu, "A systematic review on content-based video retrieval," *Eng. Appl. Artif. Intell.*, vol. 90, p. 103557, Apr. 2020, doi: 10.1016/J.ENGAPPAI.2020.103557.

[2] S. R. Dubey, "A Decade Survey of Content Based Image Retrieval Using Deep Learning," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 32, no. 5, pp. 2687–2704, May 2022, doi: 10.1109/TCSVT.2021.3080920.

[3] Y. Liu, S. Albanie, A. Nagrani, and A. Zisserman, "Use What You Have: Video Retrieval Using Representations From Collaborative Experts," *30th Br. Mach. Vis. Conf. 2019, BMVC 2019*, Jul. 2019, Accessed: May 10, 2023. [Online]. Available: https://arxiv.org/abs/1907.13487v2

[4] G. Kordopatis-Zilos, C. Tzelepis, S. Papadopoulos, I. Kompatsiaris, and I. Patras, "DnS: Distill-and-Select for Efficient and Accurate Video Indexing and Retrieval," *Int. J. Comput. Vis.*, vol. 130, no. 10, pp. 2385–2407, Oct. 2022, doi: 10.1007/S11263-022-01651-3/TABLES/10.

[5] B. S. Kumar and K. Seetharaman, "Content based video retrieval using deep learning feature extraction by modified VGG_16," *J. Ambient Intell. Humaniz. Comput.*, vol. 13, no. 9, pp. 4235–4247, Sep. 2022, doi: 10.1007/S12652-022-03869-Y/METRICS.

[6] M. Mühling *et al.*, "Deep learning for content-based video retrieval in film and television production," *Multimed. Tools Appl.*, vol. 76, no. 21, pp. 22169–22194, Nov. 2017, doi: 10.1007/S11042-017-4962-9/METRICS.

[7] L. Rossetto *et al.*, "Interactive Video Retrieval in the Age of Deep Learning - Detailed Evaluation of VBS 2019," *IEEE Trans. Multimed.*, vol. 23, pp. 243–256, 2021, doi: 10.1109/TMM.2020.2980944.

[8] F. C. Lin, H. H. Ngo, and C. R. Dow, "A cloud-based face video retrieval system with deep learning," *J. Supercomput.*, vol. 76, no. 11, pp. 8473–8493, Nov. 2020, doi: 10.1007/S11227-019-03123-X/METRICS.

[9] D. L. Pham *et al.*, "A Deep Learning-Based Real-Time Video Object Contextualizing and Archiving System," *Int. Conf. Adv. Commun. Technol. ICACT*, vol. 2023-February, pp. 137–144, 2023, doi: 10.23919/ICACT56868.2023.10079454.

Copyright ⓒ 2023 KSII

The Mediating Effect of Fear of Crime on Korean Women's Defensive Behaviors
: The Necessity of CCTV

Sooyun Lee, Hyewon Jeon, Song Kang and Euigab Hwang*
Department of Criminology, Kyonggi University
Seoul, South Korea
[e-mail: lee_soosoosoo@kyonggi.ac.kr, jlana1022@gmail.com, js8998s@naver.com, flyhwang@gmail.com]
*Corresponding author: Euigab Hwang

Abstract

Fear of crime is an important issue for the safety of women. This research attempts to investigate the factors of defensive behaviors among women in South Korea, focusing on the mediating role of fear of crime. This approach is based on vulnerability, disorder, community control, and victimization model. We conducted an AMOS by using mediating variables and the Korean Crime Victimization Survey of 2020. It also includes policy recommendations describing a crime prevention system through CCTV. The findings provided helpful implications to understand defensive behaviors of women in Korea. Implications for future research are discussed based on the findings.

Keywords: Vulnerability, Fear of Crime, Defensive Behavior, Disorder, Women's Safety

1. Introduction

Fear of crime is considered a significant social issue because it not only affects individuals' physical and mental health but also the sense of community among neighbors. Gender inequality, social disparities, and crime rates in South Korea may contribute to the women's perceptions towards personal safety. According to the Korean Statistical Office in 2010, it has been confirmed that the fear of crime in South Korea is significantly higher compared to the other countries and this is likely due to the physical differences. There have been various policy measures to reduce women's fear of crime and specifically, the necessity of CCTV. However, there has been a lack of active research on this topic. Therefore, this research aims to investigate the strategies employed to safeguard personal safety as well as the mediating effect of fear of crime on Korean women's defensive behaviors.

2. Theoretical Background and Review of Research

2.1 Vulnerability & Disorder

This study focuses on the physical vulnerability of "women," and previous international studies related to vulnerability suggest that unlike men, "age" does not have a significant impact on fear of crime[1]. In contrast, previous domestic studies argue that women, specifically older women, have higher levels of fear of crime compared to younger women, indicating a difference in the factors influencing fear of crime between genders and age groups in international and domestic research. The physical devastation and behavioral disorder of the community affect individual experience of crime damage, fear of crime, and crime avoidance and defense behavior.

"This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

2.2 Social control & Confidence in Police

In areas where community disintegration occurs, informal social controls among residents weaken, leading to a higher perception of subjective social and physical disorder[2]. Moreover, a high level of perceived disorder is associated with increased fear of crime[3]. In studies on public trust in the police, it has been found that citizens tend to attribute responsibility for high crime rates, caused by social disorder or weak community bonds, to the police[4]. While it is commonly assumed that fear of crime leads to low trust in the police, the reverse causal relationship is less commonly examined[5].

2.3 Fear of Crime

Individuals who are afraid of crime will avoid activities at specific times and spaces that are perceived to be dangerous, take evasive actions accordingly, and prepare for defensive actions and measures[6]. Research findings indicate that women have higher levels of fear of crime compared to men, despite lower actual victimization rates[7]. This difference is attributed to variations in self-protective abilities based on gender, which can also be interpreted as a result of physical vulnerability, supporting the contradictory statistical results between men and women.

2.4 Defensive Behaviors

Lack of protective capabilities leads to defensive behaviors. Women and younger age groups engage more in defensive activities. Higher education, income, and neighborhood decay relate to increased defensive behaviors. Those with more victimization experience, higher perceived risks and fears of crime tend to engage in more defensive actions[8]. When women perceive their residential areas to be characterized by rampant disorder and an increasing trend in crime, they are inclined to take more defensive actions.

3. Method

3.1 Research Methods

The data used in this study was obtained from the "National Crime Victim Survey 2020" conducted by the Korean Institute of Criminology. The survey consisted of a basic questionnaire completed by all respondents aged 14 and older, and an incident questionnaire completed by respondents who experienced crime victimization. The data analyzed in this study only included responses from 7,218 women, excluding men and duplicate responses.

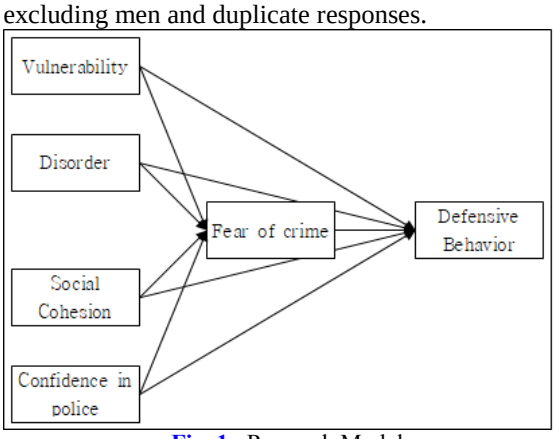


Fig. 1. Research Model

3.2 Hypothesis

Table 1. Hypothesis

Variable			Fear of Crime	Defensive behavior
Independent variable	Vulnerability	Age	+	-
		Education level	+	+
		Social economic Status	+	+
		Victimization	+	+
	Disorder		+	+
	Social Cohesion	Formal control	-	+
		Informal control	-	+
	Confidence in Police		-	+
Mediating variable	Fear of crime		*	+

4. Result

4.1 Analysis of the Modified Structural Model

Table 3. Revised Structural Model's goodness of fit

	Chi -square	RMS EA	CFI	TLI	NFI
Revised model	1933. 885***	0.052	0.983	0.945	0.983

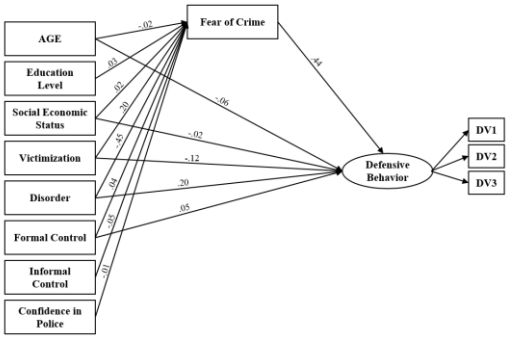


Fig. 3. Revised structural model

After connecting the error terms using Modification Indices (M.I.), the fit indices of the modified structural model are as follows, as shown in the **Table 3**. The RMSEA value has decreased and falls between 0.05 and 0.08, and the CFI, TLI, and NFI values have improved and are above 0.9.

5. Discussion & Suggestion of the application of CCTV

Recent incidents of violent crimes targeting women continue to occur, causing not only concern among women but also nationwide anxiety. Gender-based violence against women is recognized as a serious social problem in our society. Given that crimes specifically targeting women often highlight their vulnerability, there is a growing need for analysis of defensive behavior that arises from the fear of crime experienced by women and their situations that lead to fear.

In South Korea, due to the abundant installation of CCTV cameras on the streets by public institutions, policy recommendations focusing on utilizing CCTV to reduce fear of crime are the most practical. Originally, the primary purpose of CCTV was to assist in post-incident measures rather than proactive crime prevention. Therefore, some studies aim to propose environmental methods for maximizing the effectiveness of CCTV in the Korean context.

This study is significant in examining the mediating effect of fear of crime on defensive behavior and exploring the impact of defensive behavior on women's daily life changes. Analyzing the factors influencing fear of crime in the unique context of Korea, different from previous overseas studies, adds to its value. However, there were no specific survey questionnaires regarding CCTV which it suggests to include those questions in future studies in order to make it possible to develop policy implications and sophisticated analysis.

References

[1] SuHong Min, *"Fear of Crime in Korea: Analysis of the Social Survey data."* The Institute of Police Science, Vol. 6, No.1, pp. 155-176, 2011.

[2] Bursik, R. J. & Grasmick, H. G. *"Neighborhoods and crime: the deminsions of effective community control."* Macmillan, New York, 1993.

[3] Taylor, R. B. & Hale, M. *"Testing alternative models of fear of crime,"* Journal of Criminal Law and Criminology, 77, pp. 151-189, 1986

[4] Jackson, J. & Bradford, B. *"Crime, Policing and Social Order: On the Expressive Nature of Public Confidence in Policing."* British Journal of Sociology, 60: pp. 493-521, 2009.

[5] YoonHwan Park, HyunSeok Jang, *"A study on the determinants of regional-level fear of crime victimization: Focusing on the official crime rate, disorder, collective efficacy, and confidence in the police."* Korean Association Of Police Science, Vol. 15, No.6, pp. 59-88, 2013.

[6] SunGho Noh & SoRa Shin, *"A Study focus on Examination of gender differences in fear of crime."* Korean Association Of Police Science, Vol.23, No.6, 2021.

[7] M Warr & M Stafford, *"Fear of Victimization: A look at the proximate causes."* Social Forces, Vol 61, No.4, 1983

[8] KwangDo Ki, *"The Empirical Study on Avoidance and Defensive Behaviors against Crime Victimization."* Korean Association Of Victimology, Vol.25, No.2, pp.61-87, 2017

Copyright ⓒ 2023 KSII

Measurement of Social Disorder based on Distance Between Objects in CCTV Video

Eunbi Cho[1], Thanh-Hai Nguyen[2], Jeong-Hyeon Chang[3], Tsagaantsooj Batzaya[1], and Kwanghoon Pio Kim[3*]

[1]Department of Public Safety Big-data Analytics, Kyonggi University
[2]Department of Computer Science, Kyonggi University
[3]Contents Convergence Software Research Institute, Kyonggi University
[e-mail: eunbee0508@gmail.com, nguyenthanhhai@kgu.ac.kr, tiger564@nate.com, tsbaachka95@kyonggi.ac.kr, kwang@kgu.ac.kr]
*Corresponding author: Kwanghoon Pio Kim

Abstract

CCTV videos have a lot of information about risks, so they are efficiently used to predict and prevent crimes and accidents. The goal of this paper is to measure and visualize the degree of social disorder. Specifically, deep learning algorithms are used to detect objects in video files captured by CCTV, and object information detected in each video frame is contextualized and stored. The distance between objects is estimated using the stored object information. By normalizing the distance range between objects, normalized data values are divided by grade and social disorder by grade is identified and visualized in real time.

1. Introduction

The need for safety and security in our society is increasing due to various issues. These social problems are occurring in various fields such as crime and disasters, and we need ways to prevent them. In recent years, new technologies and innovative ideas to respond to these problems are constantly emerging and are successfully applied to various fields. Among them, CCTV technology is recognized as an effective tool for preventing crimes, disasters, and accidents. However, while CCTV is useful for identifying crimes and accidents that have already occurred, it has limitations when it comes to preventing them. In particular, in traffic, investigation, and disasters, video information must be processed in real time and preventive measures must be taken based on it. Therefore, this study aims to build a system that measures the degree of disorder using the distance between objects and visualizes it by dividing and grading the level of disorder. This system can process CCTV video information in real time and provide information to prevent and deal with dangerous situations in advance, which helps preventive use of CCTV technology, which is important in modern society where the demand for safety and security is increasing.

2. Related work

2.1 Disorder

Disorder is associated with community disintegration[1], the fear of crime[2], and the occurrence of crime[3] itself. The concept of disorder can be broken down into two aspects: physical disorder and social disorder. Physical disorder relates to the physical environment of a community and is manifested by signs such as graffiti on walls, piles of unattended trash, broken streetlights, broken windows, unmaintained buildings, and deteriorating infrastructure. Social disorder is related to human behavior, such as street hoodlums, people roaming the streets, prostitution stings, drunks in public, drug addicts, homeless people, and others who are not criminals but who, if left unchecked, can lead to crime.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

2.2 Systematic Social Observation (SSO)

The measurement of disorder has been conducted by human observers through "Systematic Social Observation" (SSO). SSO[4][5] is a method in which an observer follows a certain route and observes the community with their own eyes in a certain place, such as a street or alleyway, and counts the signs of disorder. However, this method is subject to the subjective judgment of the observer, and the observation results may vary depending on the observer. This is a time-consuming and labor-intensive method that requires expert participation and training, as the collection and analysis of data requires specialized knowledge. A major disadvantage of this method that it is a snapshot in time. It is difficult to consider the time of day, day of the week, and seasonal factors that change over time. Especially when it comes to crime and accident prevention, real-time is important. The downside of these methods is that they cannot detect time-varying disorder.

3. Method

The system works in five steps. First, it detects objects from CCTV images using deep learning algorithms and saves them as JSON files containing frame ID, number of objects, object type, object location, etc. through contextualization. Second, to calculate the distance between objects, the distance from the center to the edge of the frame is calculated first, and the real-time distance is calculated through the distance between the x, y coordinates of the object closest to the center. Third, to investigate the influence of objects on the occurrence of disorder, we consider that objects move closer to each other through each frame, so we base it on the change in the standard deviation of object distances. We also consider the mean absolute deviation, which is a more direct measure of the mean. Fourth, we level the degree of social disorder based on deviation. Due to the nature of movement, the distances between objects captured in a frame are highly variable, so the data has a wide distribution range, which we normalize using a Min-Max scaling method. The distance deviation (SD and MD) data is normalized to a range from 0 to 1, and the values of this normalized data are divided into 7 levels. These 7 levels represent the degree of social disorder from low to high, with level 1 representing a high risk of occurrence and level 7 representing a safe state with no risk at all. Fifth, visualize the leveling of disorder in CCTV footage based on the previous steps.

$$sd = \sqrt{\frac{\sum_{i=1}^{n}(t_i - t)^2}{n-1}} \quad (1)$$

$$md = \frac{\sum_{i=1}^{n}|t_i - \bar{t}|}{n} \quad (2)$$

Table 1. Seven levels of social disorder

Level	Values of deviation	Color
1	(0, 1/7)	Red
2	(1/7, 2/7)	Orange
3	(2/7, 3/7)	Yellow
4	(3/7, 4/7)	Light Green
5	(4/7, 5/7)	Green
6	(5/7, 6/7)	Light Blue
7	(6/7, 1)	Blue

4. Experiment and Results

To test the proposed system, we used an kidnap video from the anomalous behavior CCTV video provided by AI HUB. The video depicts a girl being kidnapped by two criminals. The left side of each figure is the captured CCTV image, the bottom right is the shortest distances objects graph (SDOG), and the top right is the top three deviation graphs. In **Fig. 1**, the color of the bars representing the deviation of the three graphs is mainly green or blue, indicating that a social disorder situation is not occurring. **Fig. 2** is a video of the moment when two criminals forcibly drag a girl into a car, and the bars are colored red, indicating that a social disorder situation may occur. **Fig. 3** shows the results calculated as mean absolute deviation (MD), which is similar to the results calculated as SD.

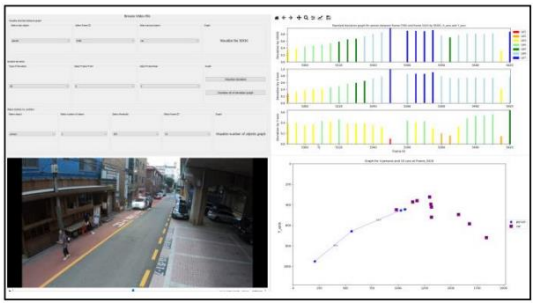


Fig. 1. Safe case calculated by SD



Fig. 2. Unsafe situation calculated by SD

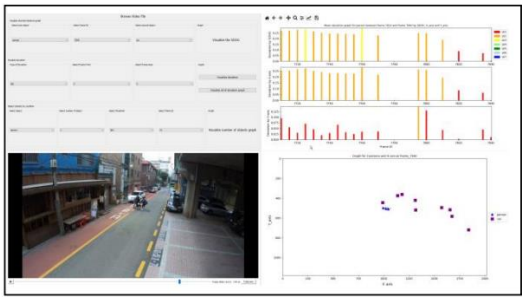


Fig. 3. Unsafe situation calculated by MD

5. Conclusions

In this study, we proposed a system to identify social disorder by calculating the distance between detected objects, especially human objects, in each frame of CCTV images in real time. It compensates for the shortcomings of CCTV, which has limitations in terms of proactive prevention, by processing video information in real time to provide information to prevent and deal with dangerous situations in advance. By replacing the existing method of measuring social disorder, which must be measured and analyzed by humans, with deep learning technology, it is expected to provide a basis for disorder leveling based on real-time crime and safety accident prediction and implement preemptive control. In future research, we will consider de-identification technology or a fully contextualized system to prevent privacy violations, and focus on recognizing and analyzing behaviors and environmental factors captured in CCTV images rather than face recognition technology that the general public is reluctant to use. We will also add other elements of social disorder to detect and analyze, and improve the system to measure physical disorder as well.

References

[1] W. Steenbeek, & D. Weisburd, "Where the action is in crime? An examination of variability of crime across different spatial units in The Hague, 2001–2009," *Journal of quantitative criminology*, Vol 32, pp. 449-469, 2016.

[2] G. L. Kelling, C. M. Coles, Fixing broken windows: Restoring order and reducing crime in our communities, Simon and Schuster, 1997.

[3] W. G. Skogan, Disorder and decline: Crime and the spiral of decay in American neighborhoods. Univ of California Press, 1992

[4] Taylor, R. B., Gottfredson, S. D., & Brower, S., "Block crime and fear: Defensible space, local social ties, and territorial functioning," *Journal of Research in crime and delinquency*, Vol 21. No.4, pp. 303-331, 1984.

[5] D. D. Perkins, J. W. Meeks, J, R. B. Taylor, "The physical environment of street blocks and resident perceptions of crime and disorder: Implications for theory and measurement," *Journal of environmental psychology*, Vol 21. No.1, pp. 21-34, 1992.

Copyright ⓒ 2023 KSII

Continual Learning Algorithms for Enhanced CCTV Effectiveness

YaeEun Lee, HyunKi Lim, and JeongHyeon Chang*
Contents Convergence Software Research Institute, Kyonggi University
Suwon, South Korea
[e-mail: yaeeunlee@kyonggi.ac.kr, hlim20@ kyonggi.ac.kr, crime_tiger564@kyonggi.ac.kr]
*Corresponding author: JeongHyeon Chang

Abstract

CCTV installation is increasing for public safety, and crime prevention through systematic monitoring is now being aided by recent developments in AI technology. However, the current neural network-based machine learning approach for automated monitoring is not well-suited for learning the diverse environments captured by CCTV and the infinite behavioral patterns of criminals. Retraining the model requires significant computing resources, making it challenging to adapt to updates in a flexible manner. To address this, we propose the use of Continual Learning Algorithms for Abnormal Detection and aims to compare the performance of the model using abnormal behavior data suitable for situations in Korea. The research results demonstrate the potential of using Continual Learning in different CCTV environments.

Keywords: CCTV, Continual Learning, Crime Prevention, Fear of Crime, Monitoring

1. Introduction

CCTV-VIDEO is purpose at crime prevention and criminal investigation evidence collection. Therefore, It can reduce victim and fear of crime. Crime prevention and reduction of crime damage through the use of CCTV depend on active monitoring by the managing institution. [1,2]. Computer vision is improve and developed by various algorithms such as object recognition, division, and tracking through automated processing of visual information of images. [3].

However, detecting abnormal behavior in CCTV images is a challenging task due to the unpredictable diversity of behavior patterns. In particular, as the capacity of data increases, physical devices and time for complex calculations are required. General neural network-based machine learning is not suitable for a CCTV environment where continuous updates are required to detect various types of anomalies [4]. In this study, we propose detecting anomalies in CCTV video through continual learning algorithms.

2. Continual Learning

In neural network learning, when learning by adding new data after completing the learning, Catastrophic Forgetting occurs. Unlike the human learning process, which does not forget the previous content even when new knowledge is acquired, neural networks perform poorly on previous tasks because parameters tailored to the previous learning are newly updated according to the added data [5,6].

However, models are required to learn from a stream of data over a long period of time. Therefore, various continual learning algorithms have been developed to alleviate this catastrophic forgetting problem and to allow the model to maintain performance and proceed with learning.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (No. 2020R1A6A1A03040583)

3. Application of Continual Learning Algorithms in CCTV environment

3.1 Continual Learning Algorithms

Five Continual Learning Algorithms were used to demonstrate their feasibility in CCTV environments. Each algorithm uses a unique approach to prevent catastrophic forgetting.
Firstly, iCaRL (Incremental Classifier and Representation Learning), selectively saves previous data in exemplars to retain knowledge from previous tasks [7]. Secondly, DGR (Deep Generative Replay) replays previous tasks like the hippocampus of human brain does [8]. Thirdly, DER (Dynamic Expandable Repression) expands neural network by adding nodes as the new data comes in [9]. Fourthly, LWF (Learning Without Forgetting) limits the weight updates to keep performance on previous task and enable the training of new ones [10]. Finally, LSIL (Large Scale Incremental Learning) uses both methods of saving previous data and regulating updates, making it suitable for dealing with massive data without the dimensional curse [11].

3.2 Data

Abnormal Behavior CCTV Video Data from AI Hub was used. There are 10 classes for train and test set including abnormal behavior of falling, kicking, pulling, punching, pushing, threatening, climbing walls, stealing, wandering, and tottering. Each set of train data includes 500 images, and each set of test data includes 100 images.

3.3 Experimental Settings and Results

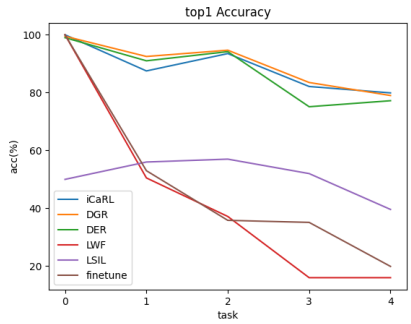


Fig. 1. Performance Evaluation of Continual Learning Algorithms

A total of 10 classes for 5 tasks were trained, with 2 classes for the first task and an additional 2 classes for each subsequent task. The baseline model used for training and classification was Resnet18, with an epoch set to 50, a learning rate of 0.001 and memory capacity 1000.
Fig. 1 shows the results of the performance evaluation of Continual Learning Algorithms. Overall, compared to the fine-tuning method, which does not include any method for preventing catastrophic forgetting, Continual Learning Algorithms maintained their accuracy over the tasks.
Notably, the memory-based Continual Learning Algorithms such as iCaRl, DER, and DGR demonstrated high performance throughout the continual learning process. It seems that sample data, which is selectively saved and applied to the next tasks, contains the robust features, and effective enough in maintainig performance over time. On the other hand, Continual Learning Algorithms using regulization method such as LWF, and LSIL showed lower performance. It appears that regularization method was not adequate for learning and adapting representations that are effective for all tasks.

4. Conclusions

Continual Learning is an efficient machine learning method that can help flexible adaptation and coping in the CCTV environment where the amount of data is overwhelming and diverse. In this paper, high performance of Continual Learning Algorithms was confirmed, and the feasibility of using memory-based continual learning algorithms was demonstrated.

References

[1] Welsh, Brandon C., David P. Farrington, and Sema A. Taheri. "Effectiveness and social costs of public area surveillance for crime prevention." *Annual Review of Law and Social Science,* Vol 11, pp. 111-130, 2015.
[2] Yavuz, Nilay, and Eric W. Welch. "Addressing fear of crime in public space: Gender differences in reaction to safety

measures in train transit." *Urban studies,* Vol 47.12, pp. 491-2515. 2010.

[3] Feng, Xin, et al. "Computer vision algorithms and hardware implementations: A survey." *Integration*, Vol. 69, pp.309-320, 2019.

[4] Doshi, Keval, and Yasin Yilmaz. "Rethinking video anomaly detection-a continual learning approach." in *Proc. of the IEEE/CVF winter conference on applications of computer vision*. 2022.

[5] Doshi, Keval, and Yasin Yilmaz. "Any-shot sequential anomaly detection in surveillance videos." in *Proc. of* the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020.

[6] Qu, Haoxuan, et al. "Recent advances of continual learning in computer vision: An overview." *arXiv preprint arXiv*:2109.11369, 2021.

[7] Rebuffi, Sylvestre-Alvise, et al. "icarl: Incremental classifier and representation learning." in *Proc. of* the IEEE conference on Computer Vision and Pattern Recognition. 2017.

[8] Shin, Hanul, et al. "Continual learning with deep generative replay." *Advances in neural information processing systems* 30. 2017.

[9] Yan, Shipeng, Jiangwei Xie, and Xuming He. "Der: Dynamically expandable representation for class incremental learning." in *Proc. of* the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2021.

[10] Li, Zhizhong, and Derek Hoiem. "Learning without forgetting." *IEEE transactions on pattern analysis and machine intelligence* Vol. 40.12. pp. 2935-2947. 2017.

[11] Wu, Yue, et al. "Large scale incremental learning." in *Proc. of* the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2019.

Copyright ⓒ 2023 KSII

Action Classification using SlowFast and Automated Machine Learning

Gyu-Il Kim[1], Seong-Un Cho[2], Jeong-Hyeon Chang[3], Kyungyong Chung[2] and Hyun Yoo[3*]
[1]Department of Computer Science, Kyonggi University, Suwon-si, South Korea
[2]Division of AI Computer Science and Engineering, Kyonggi University, Suwon-si, South Korea
[3]Contents Convergence Software Research Institute, Kyonggi University, Suwon-si, South Korea
[e-mail: ran6491@kyonggi.ac.kr, jodory9524@gmail.com, crime_tiger564@kyonggi.ac.kr, dragonhci@daum.net, rhpa0916@gmail.com]
*Corresponding author: Hyun Yoo

Abstract

Video analysis is a useful technology for crime prevention and accident detection, but faces challenges such as data size and slow processing. The combination of AI and video recognition, including deep learning and the SlowFast model, has shown high accuracy. However, manual labor and cost remain issues. To address these, recent developments combine automated technologies like AutoML. This paper proposes a methodology to enhance the SlowFast model's accuracy through AutoML and transform its multi-class classification into a multi-label problem.

Keywords: Artificial intelligence, Convolutional neural network, Video analysis, Action Classification

1. Introduction

Video analysis is a technology that extracts and analyzes various types of information from videos, including text, images, audio, and behavioral data. In particular, in video surveillance systems, it is essential to detect and analyze human actions or object movements to proactively prevent or respond to risks such as crimes and accidents. However, video analysis currently faces significant challenges due to limitations in data size and slow processing speeds. As a result, the application of artificial intelligence technology is increasingly emphasized in the field of video analysis. AI technology proves highly beneficial for processing and analyzing massive amounts of video data, and various techniques, including deep learning, have been applied, resulting in high classification accuracy. Therefore, the combination of video analysis and deep learning is crucial [1].

The SlowFast model has shown excellent performance in the field of video recognition recently and has demonstrated outstanding results in video classification and recognition tasks. However, training such models requires a significant amount of manual labor and cost. To address these issues, automated technologies such as AutoML have gained attention, and the need for automated model training techniques in the field of video recognition is also emerging [2].

Therefore, in this paper, we propose the construction of an AutoML system utilizing the action classification probabilities of the SlowFast model. We demonstrate the potential of automated model training in the field of video recognition. Through this approach, we present the possibilities of automated technologies replacing human labor in video recognition and describe techniques for improving accuracy through AutoML. We provide a methodology for achieving these goals.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(No. 2020R1A6A1A03040583)

2. Related Works

2.1 SlowFast and AutoML

The SlowFast model is a popular 3D convolutional neural network model in the field of video recognition [3]. It operates by processing slow frames and fast frames in separate neural networks and combining the results. This allows the SlowFast model to consider both slow and fast actions, enabling it to adapt to rapid changes in video data. The internal structure of SlowFast consists of the Slow Pathway and the Fast Pathway. These pathways process slow frames and fast frames, respectively. Training is conducted by processing different frames in different neural networks and combining the results. The slow frames primarily handle slowly changing actions, while the fast frames handle rapidly changing actions. By combining these two pathways, the SlowFast model can accommodate various actions, resulting in high classification accuracy. However, it has the drawbacks of slow processing speed, relatively large memory usage, and the requirement for a large amount of data for effective training.

AutoML is a technology that automates the selection of machine learning model architecture and hyperparameters. Traditional approaches for model configuration and hyperparameter selection often rely on the expertise and knowledge of domain experts. However, AutoML reduces the reliance on human experts and enables more efficient model configurations in machine learning. AutoML models automatically search for the optimal model configuration by trying out various combinations of hyperparameters. The selected models can be further enhanced through ensemble techniques, resulting in improved performance. This allows for more accurate action classification and more efficient results compared to manual methods. AutoML is particularly useful in handling large-scale datasets, enabling the identification of optimal machine learning model configurations based on the characteristics of the data. By utilizing the action classification probabilities obtained from the SlowFast model, training an AutoML model can be implemented to achieve high-performance video analysis technology.

3. Action Classification using SlowFast and Automated Machine Learning

This study proposes Action Classification using SlowFast and Automated Machine Learning. It is a method that utilizes AutoML techniques to enhance the accuracy of the existing SlowFast model. **Fig. 1** illustrates the Action Classification process of the proposed model in this paper.

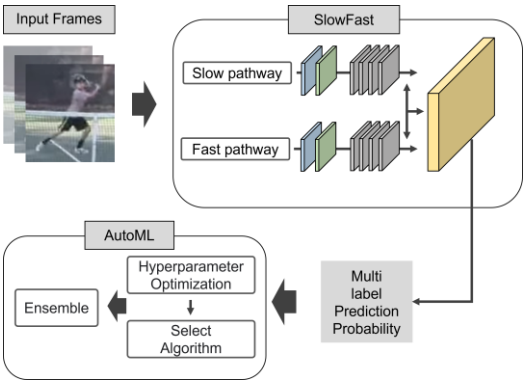


Fig. 1. Process of Action Classification using SlowFast and Automated Machine Learning

3.1 Data preprocessing

In this study, we trained the SlowFast model using the Kinetics-400 dataset. The Kinetics-400 dataset is a large-scale video dataset consisting of 240,000 videos with 400 action classification labels. Training the model with the entire Kinetics-400 dataset would be time-consuming, so we randomly sampled 8 frames from the original videos. Each frame was resized to 256x256 pixels and then normalized to a range of 0-1. During this process, since the number of video samples corresponding to each action classification label in the dataset varies, we need to perform uniform data sampling for each class. In the first step of video processing, unnecessary frames were discarded, and only about 64 frames were used from each video. This allows for video processing within the limited GPU memory. When feeding the preprocessed video data into the model, it needs to be structured as mini-batches. Each mini-batch consists of a 4D tensor with the size of TxCxHxW, where T is the number of frames, C is the number of channels, H is the image height, and W is the image width. The SlowFast model is trained using the Kinetics-400 dataset that has undergone these

data preprocessing steps.

3.2 Multi-label Classification

The existing SlowFast model solves the Action Classification problem by predicting one action for each frame, treating it as a multi-class classification problem. To improve the accuracy of the original SlowFast model by incorporating AutoML techniques, we transform the multi-class classification problem into a multi-label classification problem. In a multi-label classification problem, instead of selecting one class among multiple classes, we predict the probability values for each class. Through this approach, we can calculate the probability values for each class and apply a threshold to make classification decisions. This allows us to assign one or more labels to each action. Since human actions are not mutually exclusive, addressing the multi-label classification problem can improve the accuracy of SlowFast by considering multiple labels for each action.

3.3 AutoML for Action Classification

To transform SlowFast into a multi-label classification problem and leverage AutoML, several steps are required. First, the loss function and activation function need to be modified. The loss function should be adapted to the multi-label setting, typically using Binary Cross-Entropy Loss. Additionally, the activation function commonly used in multi-label classification tasks, such as the sigmoid function, should be employed. Lastly, the output layer needs to be adjusted. The original output layer, tailored for multi-class classification problems, should be modified to suit the multi-label classification task. To explore and apply suitable multi-label classification algorithms, AutoML is used. AutoML techniques automatically search and apply the optimal model architecture and hyperparameters. The multi-label classification problem of the SlowFast model provides the probability values for each label, which are then used as inputs for AutoML. Through the AutoML process, the model is trained and evaluated using the provided data. Hyperparameter adjustments, sample loss weight adjustments, and other techniques are applied to improve the model's performance. Finally, the top three performing multi-label classification algorithms are combined into an ensemble model. Through this process, SlowFast can be transformed into a multi-label classification problem and integrated with AutoML, ultimately enhancing the accuracy of the original SlowFast model.

4. Conclusions

This paper proposes the method of Action Classification using SlowFast and Automated Machine Learning. The proposed approach transforms the original SlowFast model, which predicts actions through a multi-class classification problem, into a multi-label classification problem. The probabilities for each label obtained through multi-label classification are then utilized as input data for AutoML. AutoML explores the optimal model and hyperparameters based on the provided data and constructs a new model by ensembling the top-performing three models. Through this process, the accuracy of the original SlowFast model can be enhanced.

This paper utilizes machine learning techniques in combination with existing deep learning models. However, one limitation is that it does not incorporate the latest trend of transformer-based action prediction models. In future research, we plan to improve the accuracy of the existing action prediction network by ensembling the networks of transformer-based action prediction models.

References

[1] V. Sharma, M. Gupta, A. Kumar, and D. Mishra, "Video processing using deep learning techniques: A systematic literature review," *IEEE Access*, Vol. 9, pp. 139489-139507, 2021.

[2] A. Galanopoulos, J. A. Ayala-Romero, D. J. Leith, and G. Iosifidis, "AutoML for video analytics with edge computing," in *Proc. of INFOCOM 2021 - IEEE Conference on Computer Communications (INFOCOM)*, pp. 1-10, 2021.

[3] C. Feichtenhofer, H. Fan, J. Malik, and K. He, "Slowfast networks for video recognition," in *Proc. of IEEE/CVF international conference on computer vision (ICCV),* pp. 6202-6211, 2019.

Copyright ⓒ 2023 KSII

Machine Learning vs. Deep Learning on a Tabular Traffic Dataset in Big Data

Dalyapraz Manatova, Jongwook Woo
Luddy School of Informatics, Indiana University, USA
Department of Information Systems, California State University Los Angeles, USA
[e-mail: dmanato@iu.edu, jwoo5@calstatela.edu]
*Corresponding author: Jongwook Woo

Abstract

This paper compares the performance of traffic prediction models in machine learning and deep learning capable of classifying a multiclass traffic dataset. The dataset comprises five different traffic jam levels of tabular format. Its data size exceeds 1.8 GB, making storing and processing using conventional single-node machines impractical. To address this issue, we developed classification models on a Big Data platform using machine learning algorithms, specifically Random Forest and XGBoost. Additionally, we implemented a deep learning model on the platform using feed-forward multilayer artificial neural network. Then, we evaluated the performance of the models in terms of accuracy and computing time. The data is in a tabular format, and the deep learning model did not perform better in the predictive analysis than the machine learning models. The Random Forest model presented the highest accuracy and efficiency in computing time. Our experimental results indicate that the deep neural network model was less effective than tree-based machine learning models in predicting tabular format multiclass traffic data.

Keywords: Spark, Big Data, GPU, Rapids, Deep Learning, Tabular Data, Scalable Computing

1. Introduction

Deep Learning algorithms are known as superior to Machine Learning algorithms, especially with large-scale datasets. However, some show it is not always true [2, 3].

Spark platforms is linearly scalable as distributed parallel computing systems that store and process large-scale datasets. It also supports multiclass classification algorithms: logistic regression, decision trees, random forests, and naive Bayes. In this paper, we compare the models to predict traffic jams using a traffic dataset collected from smartphone device apps. The dataset is over 1.8 GB in a tabular format. The dataset has five-level traffic jams, which requires a multiclass classification model. We also study if the Deep Learning (DL) model could perform better than the Machine Learning (ML) model with tabular datasets [2, 3].

2. Related Work

Dalya et al. improved the performance of machine learning in distributed parallel computing systems with GPU to predict binary-class traffic jam [1]. In this paper, we develop multiclass traffic jam prediction models in Big Data systems, including a DL model.

Ravid et al. compared the accuracy of an XGBoost and various Deep Learning models with 11 tabular data sets. It showed that XGBoost model performs better than deep learning models [2]. Léo et al. investigated the performance of Deep Learning and Tree-based models with 24 tabular datasets. They found that tree ensemble models are superior to deep learning. They stated that it is because the tabular dataset has irregular patterns in the target function, uninformative features, and non-rotationally-invariant data [3]. Our paper's models are multiclass classification in the big

This research was supported by a research credit from Amazon AWS and Oracle Cloud

data platform with over 1.8 GB size, while they compared binary classification and regression models with small data sets.

3. Big Data Machine Learning and Deep Learning with GPU

Apache Spark has been a popular Big Data solution. GPU chip uses multi-cores for parallel computing to achieve high performance and accelerates the development of the deep learning applications. The Rapids suite is an open-source software libraries to utilize GPU [5]. We can leverage Spark Big Data cluster with Rapids in multiple nodes that have both CPUs and GPU.

4. Dataset and Specifications

We collaborated with the Information Technology Agency of the Los Angeles City Department with the traffic dataset. The dataset comprises JSON files covering information reported by app users and information captured from users' smartphones. The dataset is of size 1.8 GB in Dec 31, 2017 – Jan 8, 2018. The data has two significant files: alerts (information reported by users) and jams (information captured by the user's device). We store the dataset in the Amazon cluster, EMR 6.8, which is consisted of 3 nodes: one *m3.xlarge* and two *g4dn.xlarge*. Each node has 4 cores and 16 GB memory. It supports Spark and Hadoop File Systems. The dataset comprises columns from the device and the app users: *[location_x, location_y, address, road_type, type, pub_date, date_pst, month, day, hour, min, sec, weekday, speed, length, delay, report_description, level]*. The column, *level*, is the multiclass label showing traffic jam level. It ranged from 1 to 5, where 1 is "almost no jam" and 5 is "standstill jam". We build machine learning and deep learning models to predict the *five-level* (*five-classes*) traffic jam.

5. Multiclass Classification Models

We develop multiclass classification models to classify five-level traffic jams: RF (Random Forest), XGBoost, and FF (Feed-Forward Multilayer Neural Network). FF contains many hidden layers with neurons and stochastic gradient descent using back-propagation.

5.1 Data Engineering and Modeling in Machine Learning

We partitioned our dataset into 80% training and 20% testing sets. Besides, we put the hyper-parameters of *RF* model as *numTrees*: [10, 20], *minInfoGain*: [0.0, 0.01], and *maxDepth*: [5, 10]. For *XGBoost*, we set the hyper-parameters as *ntrees*: 50, and *max_depth*: 5. In FF, we put *hidden layer*: (200, 200), and *epochs*:100. We evaluate the performance of the models by employing the metrics, *Precision*. It denotes the ratio of correctly identified positive records out of the total records identified as positive. Another metric is *Recall*, which represents the ratio of correctly identified positive records out of the total actual positives. We also measure Macro- and Micro-Average Precision & Recall, as the label has five classes/levels. We focus on Macro-Average more because it is affected by the majority class.

5.2 Experimental Result

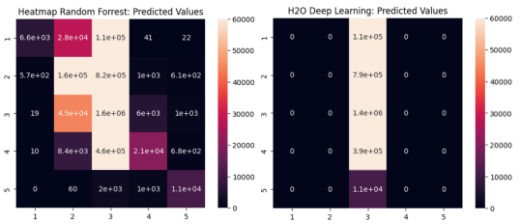


Fig. 1. Confusion Matrix Heat Maps of Random Forest and Feed-Forward Deep Learning

We calculate the metrics of multiclass classification models with a *Confusion Matrix* composed of Actual and Predicted values of the five levels. Its rows are Actual Values, and the columns are Predicted.

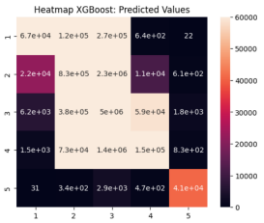


Fig. 2. Confusion Matrix Heat Map of XGBoost

RF of **Fig. 1** shows that classes 2 and 3 have higher TP (True Positive) values than other

classes. It also shows that class 3 of FF has the only highest TP while other predicted valuse are 0s. Figure 2 illustrates that XGBoost has more stable and high TP values than other models. TP is displayed in the diagonal of the matrix. Table 1 shows macro-average Precision & Recall of the models and the computation time to build a model. RF has the shortest computing time. The computation time using GPU in XGBoost is 9 % faster than CPU.

Table 1. Performance Comparison of the Models

Macro -Avg	RF	XGBoost		FF (Epochs)	
		CPU	GPU	50	100
Prec.	0.733	0.690		0	
Recall	0.400	0.468		0.2	
Comp. (mins)	38	66	60	59	120

Table 2. Accuracy of Random Forest Model

	Precision	Recall
Class 1	0.91697	0.04601
Class 2	0.66767	0.16480
Class 3	0.54003	0.96947
Class 4	0.71825	0.04177
Class 5	0.82419	0.77846

Tables 2 and **3** list each Precision & Recall of the five classes in RF and XGBoost models. We focus more on the high Precision in the low classes. Class 1 with the highest Precision is helpful for the drivers, where the higher the Precision of Class 1, the smaller the FP (False Positive) of Class 1 is. We can translate the FP as follows: a model predicts that a place is Class 1 ("almost no jam"), but it is false. So, it has a standstill jam, which will annoy the drivers who follow the prediction. The RF model has higher *Precision* of 91.7 % and 66.8 % in Classes 1 and 2 than the XGBoost model with 69.3 % and 59.3 %. It is challenging to use the FF model that has only a precision of 50.9 % in Class 1, while other classes have Precisions of 0s.

Table 3. Accuracy of XGBoost Model

	Precision	Recall
Class 1	0.69281	0.14494
Class 2	0.59282	0.26392
Class 3	0.55945	0.91817
Class 4	0.67765	0.09465
Class 5	0.92680	0.91651

6. Conclusions

This paper compares the performance of traffic prediction models capable of classifying a multiclass traffic dataset with five traffic jam levels exceeding 1.8 GB in size. We evaluated the performance of the multiclass classification models in terms of accuracy and computing time using ML and DL algorithms on a Big Data Spark platform: RF, XGBoost, and FF. The FF model did not exhibit higher accuracy than the ML models, as the data was in a tabular format. The RF model demonstrated the highest *Precision* (91.7 %) with the most efficient computing time (38 mins), especially in predicting low traffic jam levels 1 and 2. The XGBoost model showed generally higher *Precision* in all traffic jam levels with a 9% faster computation time with GPU acceleration. In contrast, the DL model efficiently predicted only traffic jam level 3. Our experimental results show that tree-ensemble ML models effectively predict multiclass traffic jam levels in tabular data format than a DL FF model.

References

[1] Dalyapraz Dauletbak, Junghoon Heo, Sooyoung Kim, Yeon Pyo Kim and, Jongwook Woo, "Scalable Traffic Predictive Analysis for Smart City using GPU in Big Data," KSII The 16th Asia Pacific International Conference on Information Science and Technology (APIC-IST), pp144-148, ISSN 2093-0542 June 20-22 2021.

[2] Ravid Shwartz-Ziv, and Amitai Armon, "Tabular data: Deep learning is not all you need," Journal of Information Fusion, vol 81, pp. 84-90, 2022, Elsevier

[3] Léo Grinsztajn, Edouard Oyallon, and Gaël Varoquaux, "Why do tree-based models still outperform deep learning on typical tabular data?," Thirty-sixth Conference on Neural Information Processing Systems (neurIPS) Dec 2-4 2022.

[4] M. Schnuerle, "Louisville and Waze: Applying Mobility Data in Cities," Harvard Civic Analytics Network Summit on Data-Smart Government, 2017.

Copyright ⓒ 2023 KSII

Occlusion Object Model using PVT based Mask-RCNN

Ji-Won Baek[1], Ye Yeon Kang[2], Geon Park[3], Kyungyong Chung[4,*]
[1]Department of Computer Science, Kyonggi University, su-won, South Korea
[2,3,4]Division of AI Computer Science and Engineering, Kyonggi University, su-won, South Korea
[e-mail: jwbaek@kyonggi.ac.kr[1], bboya0517@kyonggi.ac.kr[2], qkrrjs0408@naver.com[3], dragonhci@gmail.com[4]]
*Corresponding author: Kyungyong Chung

Abstract

Occlusion is a phenomenon in which some areas of an object are covered by other objects or lights and becomes invisible. This affects the accuracy and detection rate of object detection. To address this, we improve the performance against occlusion detection by using a pyramid vision transformer network as a backbone to enable superior detection of objects of various sizes with Mask RCNN models capable of occlusion detection.

Keywords: Occlusion Object, Pyramid Vision Transformer, YOLO, Deep Learning

1. Introduction

Artificial intelligence for image data analysis is actively being researched, and the Transformer model, originally used in natural language processing, has shown promise in the field of image analysis due to its high computational power and accuracy [1]. Since Transformers use spatial information in a limited way, they are useful in detecting and analyzing the relationships and interactions between objects in images without placing excessive emphasis on location information, making them versatile in object detection tasks where position information may not be critical. Additionally, Transformer models produce a fixed-size output matrix, independent of the input sequence length, which facilitates effective learning regardless of input image size. However, the resource demands of these models are high, requiring significant training time and computing and memory resources for processing large datasets.

Object detection is a technique for identifying and localizing objects in images or videos, performing well when objects are fully visible in the image [2,3]. However, in real-world scenarios, objects of interest are often partially occluded or hidden behind other objects, making detection difficult. Research is underway in the field of occlusion detection, which has potential applications in autonomous driving, healthcare, security systems, and more. Mask R-CNN is a popular method used for occlusion detection, performing object detection and segmentation simultaneously, allowing for accurate detection and classification of partially occluded objects [4]. However, the model architecture is complex and computationally demanding, with reduced accuracy when the object density is high.

2. Related Work

2.1 Occlusion Object using Mask RCNN

Object detection detects objects of interest in image data such as photos, videos, and CCTV [5,6]. This localizes objects in image data using bounding boxes and performs classification based on the identified regions. Object detection is classified by classification and regional proposal. Classification is an algorithm that distinguishes classes for a specific object, while regional proposal is an algorithm that quickly finds areas where objects are likely to be present. One-stage detectors perform classification and regional proposal simultaneously after inputting

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(2020R1A6A1A03040583)

image data into the model and extracting features through convolutional layers. In contrast, two-stage detectors perform classification and regional proposal sequentially, which can lead to slower processing times. Object detection can be utilized in various fields such as autonomous driving, healthcare, and facial recognition. However, the accuracy of object detection is low when the object is partially occluded [7]. Therefore, occlusion detection is necessary. Occlusion detection enables object detection even when objects are partially occluded and is divided into three stages. The first stage is segmentation, which divides the image into multiple parts and predicts objects in each region. The second stage is occlusion detection, which identifies occluded areas based on the results extracted in the segmentation stage and identifies the object in the occluded region. The last stage is detection result correction, which corrects the detection results by identifying occluded objects and adding their areas to the detection results. Therefore, occlusion detection can accurately detect occluded objects and improve the reliability of object detection. Mask R-CNN is a representative method used for occlusion detection [8]. Mask R-CNN performs object detection and segmentation simultaneously, and segments occluded areas of objects to improve detection rates.

A. Mahmoud et al. [9] proposed an object detection model using adaptive Mask RCNN for optical remote sensing images. Their method efficiently reduces the overlap of detection boxes and allows for multi-scale targets in complex background images by using adaptive Mask RCNN to detect objects in multi-scales in optical remote sensing images. P. Ammirato et al. [10] proposed a Mask RCNN for probabilistic object detection. Their method models uncertainty at object locations by treating detection scores as actual probabilities based on probability-based detection quality (PDQ) for object detection. This provides a strong baseline for probabilistic object detection using Mask RCNN object detectors.

3. Occlusion Object Model using PVT based Mask-RCNN

3.1 Data Collection and Pre-processing

The data used in this paper is the NIH Chest X-ray Dataset [11], consisting of 112,120 images of 14 classes. The 14 class labels are Atelectasis, Cardiomegaly, Effusion, Infiltration, Mass, Nodule, Pneumonia, Pneumothorax, Fibrosis, Consolidation, Edema, Emphysema, Pleural Thickening, and Hernia. The size of the images is 1024 x 1024. However, using images of this size requires a lot of time and computational resources for training the model. Therefore, the size of the images is resized to 224 x 224 to reduce the training time and computational cost.

3.2 PVT-based Mask RCNN for Occlusion Detection

When occlusion occurs, object recognition or tracking becomes difficult. For example, in the healthcare field, when a part of the lung is obscured by another organ in a CT or X-ray image used to diagnose lung diseases, it is difficult to see. Occlusion detection can improve the accuracy of diagnosis, so Mask RCNN is used for occlusion detection [12,13]. Mask RCNN performs object detection and classification tasks using a region proposal network (RPN) and ROI (Region of Interest) pooling, and additionally performs the task of generating a segmentation mask. Therefore, it accurately identifies the object's area, which results in high accuracy for occlusion detection.

The size of objects in an image varies. The pyramid network is a form of stacking feature maps with different resolutions generated by a convolutional network. This allows for better feature extraction at various scales, including capturing small objects or detailed features. Therefore, to detect various sizes, a pyramid network is used, and the PVT (Pyramid Vision Transformer), which has superior performance, is used as the backbone network of Mask RCNN. **Fig. 1** shows the structure of the PVT-based Mask RCNN.

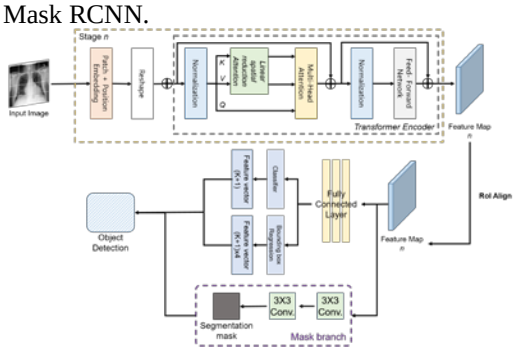


Fig. 1. The structure of the PVT-based Mask RCNN

In the structure of **Fig. 1**, the input image is divided into patches, and patch embedding and position embedding are performed. After reshaping, the feature map is obtained from the Transformer encoder. In order to reduce the computational cost of the attention operation in the Transformer encoder, a Linear Spatial Reduction Attention layer is added. As a result, both computational and memory costs can have linear complexity. The feature map generated in the previous stage is used as input to generate the next stage's feature map, which is then used for ROI Align to extract features and perform classification and bounding box regression. In addition, a mask is output for the ROI, allowing for more precise object detection.

4. Conclusions

In this paper, we propose the Occlusion Object Model using PVT based Mask-RCNN. The proposed method is to improve Mask RCNN for high-performance occlusion detection. Occlusion detection is important for improving the accuracy of object recognition or tracking, such as in healthcare imaging. Mask RCNN is used for occlusion detection, performing object detection and classification tasks using RPN and ROI pooling, and generating a segmentation mask for accurate object area identification. To detect various object sizes, a pyramid network is used with PVT as the backbone network for superior performance in feature extraction at various scales.

References

[1] K. Han, Y. Wang, H. Chen, X. Chen, J. Guo, Z. Liu, Y. Tang and A. Xiao, "A Survey on Vision Transformer," *TPAMI*, Vol 45, No. 1, pp. 87-110, 2022.

[2] N. Zeng, P. Wu, Z. Wang, H. Li, W. Liu, and X. Liu, "A small-sized object detection oriented multi-scale feature fusion approach with application to defect detection," *IEEE Trans Instrum Meas IEEE T INSTRUM MEAS*, Vol 71, pp. 1-14, 2022.

[3] J. Kang, S. Tariq, H. Oh, and S. S. Woo, "A survey of deep learning-based object detection methods and datasets for overhead imagery," *IEEE Access*, Vol 10, pp. 20118-20134, 2022.

[4] B. Ahmed, T. A. Gulliver, and S. alZahir, "Image splicing detection using mask-RCNN," *Signal, image and video processing*, Vol 14, pp. 1035-1042, 2020.

[5] R. Usamentiaga, D. G. Lema, O. D. Pedrayes, and D. F. Garcia, "Automated surface defect detection in metals: A comparative review of object detection and semantic segmentation using deep learning," *IEEE Trans. Ind. Appl. IEEE T IND APPL*, Vol 58, No. 3, pp. 4203-4213, 2022.

[6] M. Afif, R. Ayachi, Y. Said, E. Pissaloux, and M. Atri, "An efficient object detection system for indoor assistance navigation using deep learning techniques," *Multimed. Tools. Appl.*, Vol 81, No. 12, pp. 16601-16618, 2022.

[7] Q. Li, Y. Su, Y. Gao, F. Xie, and J. Li, "OAF-Net: An Occlusion-Aware Anchor-Free Network for Pedestrian Detection in a Crowd," *IEEE trans Intell Transp Syst.*, Vol 23, No. 11, pp. 21291-21300, 2022.

[8] L. Zuo, P. He, C. Zhang, and Z. Zhang, "A robust approach to reading recognition of pointer meters based on improved mask-RCNN", *Neurocomputing*, Vol 388, pp. 90-101, 2020.

[9] A. Mahmoud, S. Mohamed, R. El-Khoribi, and H. AbdelSalam, "Object detection using adaptive mask RCNN in optical remote sensing images," *Int. J. Intell. Eng. Syst*, Vol 13, No. 1, pp. 65-76, 2020.

[10] P. Ammirato, and A. C. Berg, "A mask-rcnn baseline for probabilistic object detection," arXiv preprint arXiv:1908.03621.

[11] NIH Chest x-ray, "https://nihcc.app.box.com/v/ChestXray-NIHCC"

[12] Y. Yuan, J. Chu, L. Leng, J. Miao, and B. G. Kim, "A scale-adaptive object-tracking algorithm with occlusion detection," *EURASIP Journal on Image and Video Processing*, pp. 1-15, 2020.

[13] D. Zeng, R. Veldhuis, and L. Spreeuwers, "A survey of face recognition techniques under occlusion," *IET biometrics*, Vol 10, No. 6, pp. 581-606, 2021.

[14] A. Wang, Y. Sun, A. Kortylewski, and A. L. Yuille, "Robust object detection under occlusion with context-aware compositionalnets," *In Proc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 12645-12654, 2020.

Copyright ⓒ 2023 KSII

Design and implementation of user participated SNS through voting function

JiHoon Lee, MoonSung Kim, Jinju Jeong, HeeYeon Cho, Sieun Park and Sangin Kim*
Department of Computer Engineering, Kyonggi University
Suwon, South Korea
[e-mail: wlgnsdl1016@kyonggi.ac.kr, moonsung0331@kyonggi.ac.kr, wjdwlswn848@kyonggi.ac.kr, 202015006@kyonggi.ac.kr, teetoos@kyonggi.ac.kr]
*Corresponding author: Sangjin Kim
[e-mail: drsjkim@kgu.ac.kr]

Abstract

Among Internet services, the most developed and popular system is SNS(Social Network Service). Recently, many experts in various field are also using SNS as a means of promoting themselves, and many companies are using SNS to communicate with consumers and promote their products. Even individuals are using SNS services for personal branding by communicating with acquaintances and sharing their lives and hobbies. This paper proposes a new SNS service design and implementation method centered on voting. The service provides analytical information on voting participation and results using big data, while having the function of an existing SNS that shares interests and forms relationships with each other. Through this, a more active user participation effect is expected than that of the existing SNS services. The service is implemented using the Spring framework, and the development environment is constructed using Docker containers to optimize the application for the cloud and maintain the same security as individual systems. The build and deployment environment is built using GitHub Actions, Docker Compose, and AWS cloud services, enabling automatic CI/CD.

Keywords: SNS, Personal Branding, Vote, Analysis of results, Marketing

1. Introduction.

With the development of the Internet, social network services (SNS) have become an essential means of communication in our daily lives[1]. Beyond communication between individuals, SNS is used by experts and companies in various fields for their promotion and marketing. In addition, individuals share their lives and hobbies and use SNS as a space for personal branding [2].

This paper proposes a new SNS service. The proposed service has the functions of existing SNS, adds new functions centered on voting, and uses big data analysis technology to provide participants with active participation and analysis. Through this, we expect more active communication and active user participation than existing SNS services and aim to provide analysis data for various votes [3].

2. Main subject

1. Voting-based SNS development

1-1. Account verification

The service uses the login API provided by Kakao, Naver, and Google as a login method to provide user convenience. A simple login process is required to use the voting service. During the initial log in process, additional information is collected to provide the user with various categories of analysis data. Additional information collection collects information from the user through a screen for entering additional

This research was supported by the MISP(Ministry of Science, ICT & Future Planning), Korea, under the National Program for Excellence in SW(2021-0-01393) supervised by the IITP(Institute of Information & communications Technology Planing & Evaluation)(2021-0-01393)

information such as birthday, gender, occupation, and address.



Fig. 1. Screen for account verification

1-2. Timeline

After account authentication, users can view the timeline. Timeline provides existing SNS functions such as sorting, commenting, searching, and bookmarking, so users can search, filter, and save voting posts on topics of interest. Through this, new ideas can be derived in the process of sharing opinions on various topics.

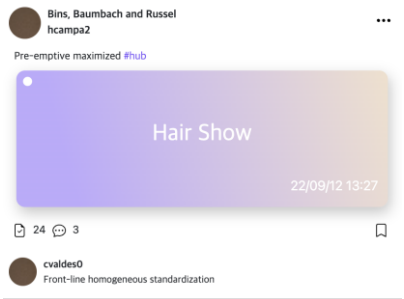


Fig. 2. Screen for timeline

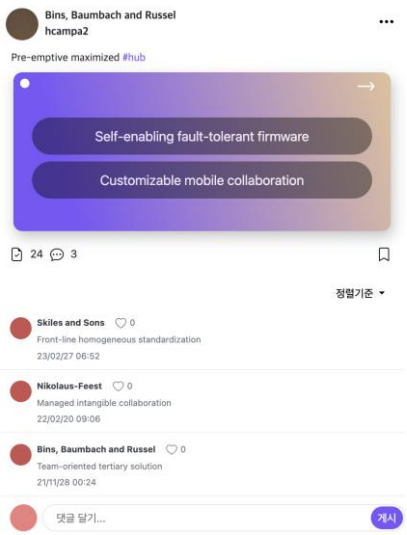


Fig. 3. Screen for detail page

1-3. Voting

When a user selects a post with a topic of interest among various posts existing on the timeline, the user can check the candidate list entered by the author. Voting participants can check the percentage of votes for each candidate when they vote for a candidate that matches their opinion.

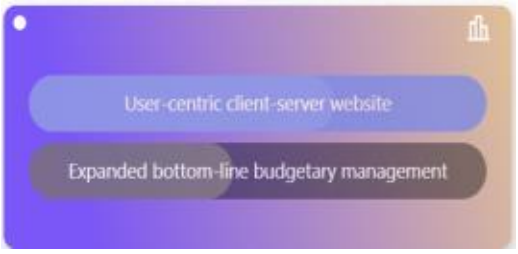


Fig. 4. Screen for voting participation

1-4. Voting analysis

Users who participated in voting can view analysis data on voting results. Click the icon in the upper right corner of Figure 4 to view the results. The voting result analysis function was implemented through the Hadoop Framework, and the results of the users who participated in the vote are aggregated through slicing, shuffle, and reducing, then placed in a model and transmitted to the user screen. To show this visually to the user, the d3.js library is provided in a graph format, and analysis data is downloaded for secondary use. The purpose of the voting analysis statistics provided to users is to be used as information for establishing corporate marketing strategy plans and government policies.

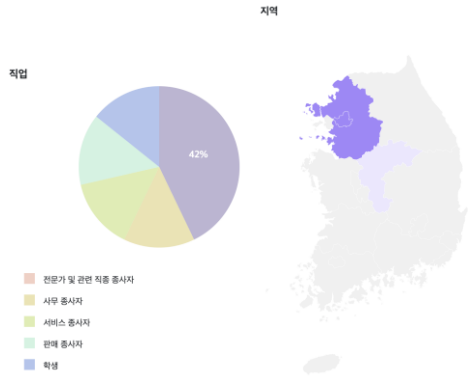
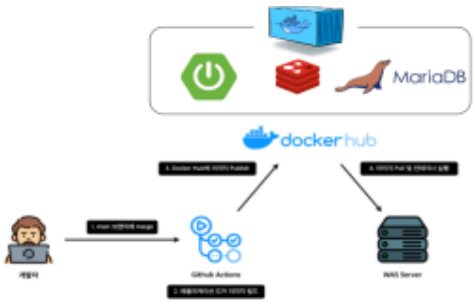


Fig. 5. Screen for Result analysis

2. Voting system configuration diagram

The configuration diagram of the developed system is as follows.



2-1. Client

Web applications were used as clients for users. UI design was carried out using Figma, and the MVC architecture was adopted to effectively develop business processing logic (Model) and interface elements (Views) as internal architectures. You can communicate with the Amazon server through a web browser to get the information you want from the client.

2-2. Docker

It uses a docker that combines the spring framework, Redis, and Maria DB into one container. The project is executed by imaging the project through GitHub Actions, uploading it to Docker Hub, and receiving it from the Amazon server and executing it.

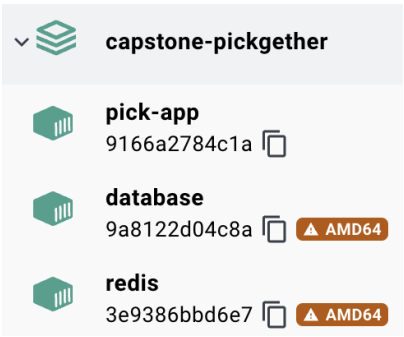


Fig. 6. Screen for docker container configuration

3. Conclusion

In this paper, a new type of SNS service centered on voting was proposed and implemented to create a space for publicity, marketing, and personal branding. This is characterized by providing analysis information on voting participation and results by applying big data analysis while maintaining the function of existing SNS services. Through this, you can expect the following. First, easy personal branding. It supports easier personal branding through participatory content called voting. Second, support for single-person households. For the increasing number of single-person households, new SNS services can be used as an effective means to form social networks and reduce negative experiences [4]. Third, enabling data utilization. It analyzes, visualizes, and provides voting results in various categories to support secondary use. Afterwards, various voting formats are applied and big data is processed. As an additional research task, improving the performance of Hadoop do.

References

[1] Cole, David A et al. “Online Social Support for Young People: Does It Recapitulate In-person Social Support; Can It Help?,” *Computers in human behavior*, Vol 68, 456-464, 2017.

[2] Byung Uk Bae,“In the era of the 4th Industrial Revolution, create me with personal branding,” *Korean Association of Converging Business Review*, Vol 9, 49-21, 2020.

[3] Mira Mayrhofer, Jörg Matthes, Sabine Einwiller & Brigitte Naderer, “User generated content presenting brands on social media increases young adults’ purchase intention,” *International Journal of Advertising*, Vol 39, No. 1, 166-186, 2020.

[4] Sangmin Kim, Hyowon Lee, “Establishing Digital Trust: The Analysis of SNS and Citizen Participation,” *Information Society & Media*, Vol 23, No. 2, 185-219, 2022.

Copyright ⓒ 2023 KSII

Customer phase identification based on sequential permutation search with voltage time series similarity

Ji-Hyun Choi[1], Haesung Lee[2], Han-Min Lim[2], Byungsung Lee[2] and Hyun Ahn[1*]
[1]Department of Information and Communication, Hanshin University, Gyeonggi, South Korea
[2]Smart Power Distribution Lab, Korea Electric Power Research Institute, Daejeon, South Korea
[e-mail: {wlgus2391, hyunahn}@hs.ac.kr, {seastar, hanmin, bysung}@kepco.co.kr]
*Corresponding author: Hyun Ahn

Abstract

With the emergence of concerns over the depletion of fossil fuels, there has been an increasing interest in the aspect of energy efficiency. As a result, smart grids, which are the next-generation power grids incorporating ICT and big data technologies, are receiving significant attention. Smart grids optimize power efficiency by enabling bidirectional communication between power providers and consumers, allowing for the monitoring of power usage. However, currently, there are difficulties in accurately assessing power consumption due to various factors. Therefore, in this research, we propose a sequential sequence search based on time series similarity as a means to identify patterns that form the foundation of power consumption measurement.

Keywords: Smart grid, phase identification, sequential permutation search

1. Introduction

In the existing power systems, energy production and supply were based on predicted values without real-time consideration of power consumption, as the system operated on a unidirectional information delivery system. As a result, there was a problem of energy waste when the actual power consumption of customers was lower than the anticipated production by power suppliers. Smart grids, as the next-generation power grids, address this issue by enabling real-time information exchange between power suppliers and consumers [1]. The Advanced Metering Infrastructure (AMI), a key technology of smart grids, automatically collects and transmits consumption data in both directions. Through this information provision, customers can monitor their power usage in real time and voluntarily conserve energy. Power suppliers, based on customer data, can prevent energy waste and optimize energy efficiency by building power generation facilities based on demand.

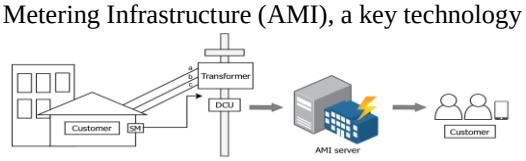


Figure 1. AMI Information communication process in smart grid

While the smart grid system offers various advantages, such as the integration of renewable energy and real-time communication, it faces difficulties in reliable power supply due to power losses caused by factors such as communication errors and the integration of renewable energy in the distribution process. To address these issues, maintaining phase balancing in the power distribution circuit is crucial, and this requires proper phase connections [2]. However, currently, errors frequently occur during the manual input of phase connection information by

This work was supported by the Korea Electric Power corporation(KEPCO) (No. R22XO05-08). This work was also supported by the National Research Foundation of Korea(NRF) grant funded by the Korea government(MSIT) (No. NRF-2021R1F1A1063160).

workers, leading to incorrect connections [3]. Therefore, the aim of this research is to identify the correct phase connection information by sequentially searching permutations based on the time series similarity of customer power data.

2. Measuring Voltage Time Series Similarities

2.1 Data configuration and background

This research utilizes TDU load values and AMI power consumption data provided by the Korea Electric Power corporation (KEPCO) Institute. The dataset was collected from a transformer in the SG (Smart Grid) pilot district in Jeju Island and includes data from 10 households with a three-phase four-wire system. Both datasets consist of time series data measured at 15-minute intervals and include variables for each of the three-phase.

The three-phase four-wire system refers to a power distribution system that utilizes three-phase (A, B, C) along with a neutral wire (N). In this system, a total of four power conductors are used. The three-phase are supplied with alternating current (AC) power, with a phase difference of 120 degrees and a time delay between them [4]. The neutral wire provides a return path for the current and helps balance the system. This configuration allows for efficient and balanced distribution of AC power in three-phase systems

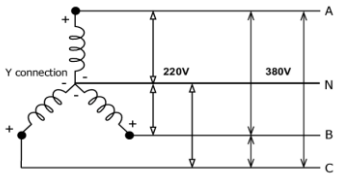


Fig. 2. Three-phase four-wire wiring method (Y connection)

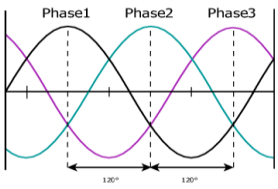


Fig. 3. Phase difference of three-phase

To normalize the TDU and AMI values, a conversion formula is applied. In the three-phase four-wire system, each phase has a voltage of 380V. For single-phase customers, the voltage is 220V, while for three-phase customers, both single-phase power at 220V and three-phase power at 380V can be used depending on the neutral wire connection method. When measuring power consumption, the AMI, which corresponds to the customer's power usage, is measured based on the line voltage (380V), which is the maximum available power. On the other hand, the TDU, which monitors the power supply and load, is measured based on the phase voltage (220V). Therefore, the following formula is applied to convert the phase voltage to the line voltage.

$$V_l = \sqrt{3}\, V_p \tag{1}$$

2.2 Sequential Permutation Search

We explore the interconnection information based on customer data. Each customer's AMI data $X^{(m)}$, is a time-series data measured over time step T, representing three-phase voltages as variables. The phases, represented by each time series x_i, correspond to A, B, and C with i values of 1, 2, and 3, respectively. In this research, considering the possibility of phase connection inconsistencies, all possible permutations of the phases for the AMI are taken into account. Therefore, a single AMI can have a total of 2^3 permutations for phase connections, and we sequentially access the AMI data of 10 customers to explore the interconnection information.

$$X^{(m)} = \begin{cases} x_1^{(m)}, x_2^{(m)}, x_3^{(m)} & 0 \\ x_1^{(m)}, x_3^{(m)}, x_2^{(m)} & 1 \\ x_2^{(m)}, x_1^{(m)}, x_3^{(m)} & 2 \\ x_2^{(m)}, x_3^{(m)}, x_1^{(m)} & 3 \\ x_3^{(m)}, x_1^{(m)}, x_2^{(m)} & 4 \\ x_3^{(m)}, x_2^{(m)}, x_1^{(m)} & 5 \end{cases} \tag{2}$$

For each permutation, calculate the time series similarity with TDU and determine the permutation with the highest correlation as the confirmed connectivity information for that AMI. The AMI with confirmed connectivity is excluded from the search target, and the above method is applied repeatedly to the remaining AMIs.

2.3 Measuring time series similarity

To measure the similarity between TDU and AMI, the Pearson correlation coefficient is used. As electricity is transmitted through wires for power supply, power loss due to resistance inevitably occurs. Therefore, when measuring similarity based on absolute values, even if the actual trends are the same, the similarity can be measured as low. On the other hand, the Pearson correlation coefficient effectively captures correlation based on trends, even in cases where there are differences in values or scale. Thus, in this research, the Pearson correlation coefficient is applied as the similarity measurement method.

$$r(X_i, Y_i) = \frac{\sum_t^T (X_{t,i} - \overline{X_i})(Y_{t,i} - \overline{Y_i})}{\sqrt{\sum_t^T (X_{t,i} - \overline{X_i})^2} \sqrt{\sum_t^T (Y_{t,i} - \overline{Y_i})^2}} \quad (3)$$

3. Experiments

To perform a performance comparison, three experiments were conducted, including the proposed content of this paper. The details of the experiments are as follows:

· Case 1: Original data
· Case 2: Phase B fixed (A-C phase variable exchange)
· Case 3: Sequential Permutation Search

Table 1. Experimental setup

Attribute	Description
GPU	NVIDIA GeForce RTX 3080 Laptop GPU 8GB
CPU	AMD Ryzen 9 6900HX
RAM	16GB
OS	Windows 11
Python libraries	pandas, numpy, scikit-learn, Matplotlib, plotly etc.

In Case 2, considering all three-phase, the total number of possibilities is 6^{10}(= 60,466,176). To reduce the search space, we took into account the expert's suggestion that Phase B shows a relatively consistent pattern compared to Phases A and C. Therefore, we conducted the experiment with a total of 1,024 search spaces.

The average of the total sum of the time series for each connected state in the finalized AMI, and the measured time series correlation coefficient of the TDU, showed a similarity of 0.958237 in the case of sequential permutation search. Additionally, in the original data (Case1), a similarity of 0.958237 was observed, while fixing state B (Case2) resulted in a similarity of 0.958596, indicating the highest similarity in the results of sequential permutation search.

Table 2. Experimental results

Case	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9	X_{10}	Corr.
1	0	0	0	0	0	0	0	0	0	0	0.9582
2	5	0	0	5	5	5	0	5	5	0	0.9585
3	2	2	5	2	0	2	0	1	5	0	0.9589

4. Conclusions

In this paper, sequential access to the connectivity permutations was performed to identify the AMI's connectivity information. Although the method proposed in this research showed the most dominant similarity value in terms of similarity, it is difficult to fully rely on it as there is no secured label and it is not a result of a comprehensive survey, thus global optimum is not guaranteed. However, despite these limitations, the proposed method in this research has successfully reduced the search space from 60,466,176 to 60, which suggests a feasible approach even with an increase in the number of meters. It is expected that more reliable results will be obtained by applying further improvements such as incorporating connected state and time lag information, as well as reducing the interval between readings in the future.

References

[1] Fang, Xi, et al., "Smart grid—The new and improved power grid: A survey.", *IEEE communications surveys & tutorials*, 14.4, pp. 944-980, 2011.
[2] Arya, Vijay, et al., "Phase identification in smart grids.", *IEEE International Conference on Smart Grid Communications (SmartGridComm)*, pp. 25-30, 2011.
[3] Blakely, Logan, Matthew J. Reno, and Wu-chi Feng. "Spectral clustering for customer phase identification using AMI voltage timeseries.", *IEEE Power and Energy Conference at Illinois (PECI)*, pp. 1-7, 2019.
[4] Gruzs, Thomas M. "A survey of neutral currents in three-phase computer power systems.", *IEEE Transactions on Industry Applications,* 26.4, pp. 719-725, 1990.

Copyright ⓒ 2023 KSII

Blockchain-Based Platform for Secure and Patient-Centric Personal Health Record Management

Sangmin Lee*
Department of Business Administration and Data Science, CHA University
Pocheon, South Korea
[e-mail: slee@cha.ac.kr]
*Corresponding author: Sangmin Lee

Abstract

To provide optimized healthcare services, an open ecosystem that allows the free exchange and utilization of personal health records (PHRs) is necessary. However, due to the data management systems being centered on medical institutions, the sharing of PHRs remains a challenge. To address this, our study designed and implemented a patient-centered PHR platform using blockchain technology. This platform safeguards the transparency and stability of information by mitigating the risk of personal information leakage and falsification that could occur during the exchange and utilization of existing PHRs. It also incorporates a consent system for PHRs primarily used by individual medical institutions, thus empowering patients with sovereignty over their records. Through this blockchain-based PHR platform, we anticipate a transformation in the medical data management system where disparate medical information from various hospitals can be consolidated and stored in a personal health record platform, and personal health records can be freely utilized in everyday life and emergency situations.

Keywords: Personal Health Record, Patient Consent, Dynamic Consent, Patent Authentication

1. Introduction

Currently, most personal health record platforms face limitations in exchanging and utilizing personal health records due to the centralized data management systems of medical institutions [1]. Specifically, personal health records that incorporate blockchain technology encounter legal constraints regarding alteration, deletion, and destruction [2].

Typically, patients' personal health records are divided into on-chain data based on automated medical information and off-chain data based on hospital medical records [3]. The Personal Information Protection Act requires 'freely given consent' for the utilization of personal information contained in personal health records. This necessitates a dynamic consent system that provides real choice and control to the information subject and allows changes to the consent choice in real time [4].

To address this issue, this study propose a personal health record (PHR) platform based on a patient-centric medical data management system, incorporating a personal consent mechanism using blockchain technology. By leveraging blockchain, safer management and utilization of sensitive personal health records requiring a high level of reliability and security can be enables [5].

2. Related Works

2.1 Blockchain Technology

Blockchain is a decentralized data storage technology that transparently records transaction history in a ledger that anyone can view, and

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00621,Development of artificial intelligence technology that provides dialog-based multi-modal explainability)

replicates and stores it on multiple computers, creating blocks with data at regular intervals and connecting them to previous blocks like a chain [6]. Due to this characteristic, arbitrary modification of data is not possible, and data forgery and tampering are also impossible based on distributed computing, where anyone can view and confirm the results of changes [7]. To understand blockchain, we can compare it to a database, a traditional way of managing data. **Fig. 1** below compares the characteristics of blockchain and database [8].

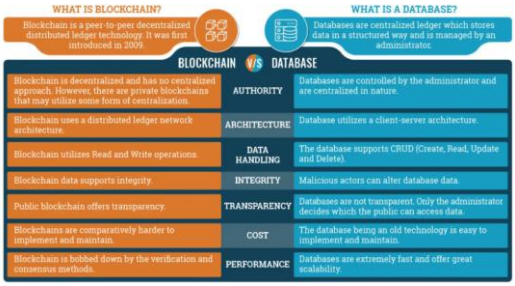


Fig. 1. Difference between blockchain and database

2.2. Dynamic Consent

Consent methods can be divided into opt-in and opt-out depending on when consent is given by the individual. The opt-in method involves obtaining consent before the personal information handler collects and uses the individual's information. Conversely, the opt-out method allows the personal information handler to stop using the individual's information if they express a desire to refuse after their information has been collected and used [9].

Currently, the consent system in the domestic healthcare field is document-based, making it inconvenient and virtually impossible for the individual to withdraw consent after it has been obtained via written agreement, verbal explanation of the consent content, and handwritten signature. Therefore, it's necessary to develop a platform centered on the individual – the patient – using information technology and UI/UX to realize dynamic consent.

3. Research Model

3.1 Design Goal

In this paper, we designed and developed a patient-centered personal health record platform to realize the following major issues. First, improving the ease of information sharing and utilization through the integration of personal health records. Second, promoting data security and transparency by utilizing blockchain. Third, utilization of patient-centered personal health records through consent system. The personal health record platform designed and developed in this paper aims to achieve the goals of security, secure autonomy, and data integration.

3.2 Sharing Personal Health Record Procedure

The personal health record sharing data structure consists of a unique value sequence of data whose numerical value is incremented by 1 when newly created from a unique value called MEDI1, a user sequence, which is a unique value sequence of identity authentication data that can identify which user is requesting to share, an agree sequence, which is a unique value sequence of consent data that can identify which consent is requested to share personal health records, and a status, which stores the request status value.

In addition, there is detail data that can store the hash value of personal health record data, which is a unique value sequence of data that increases the numerical value by 1 when newly created from the unique value called MEDIDETAIL1, and MEDI Sequence, a unique value sequence of personal health record sharing data that can identify which personal health record sharing data is, User Address, which is a user's address value to identify which user is requesting to share, Agree Sequence, which is a unique value sequence of consent data that can be checked based on which consent is requested to share personal health records, Medi Info, which hashes and stores personal health information data, and finally Status, which records the transmission status value.

4. Implementation

In this study, the configuration of three virtual servers of blockchain is based on Ubuntu. They are composed of two orders and two organizations, and each organization comprises two peer nodes. The one who issues the order uses validated transactions to ensure the secure

delivery of transactions to peer nodes. The peer node is the most basic node in the Hyperledger Fabric with chain (ledger); the chaincode (smart contract) developed in this study is also included in the peer node. The peer validates the requested transaction and shares it with other peers by updating the chain (ledger). Among the orders, Raft is a distributed consensus algorithm of the Hyperledger Fabric, and it is used to provide an ordering service to the blockchain framework, similar to the configuration of a cluster, where multiple computers are connected and operate as one system [11].

5. Conclusion

In this paper, a blockchain-based personal health record (PHR) platform for patient-centered PHR utilization was designed and implemented using Hyperledger Fabric. We designed and implemented a patient-centered data utilization platform based on a consent system for the utilization of personal health records. We also conducted a performance analysis of blockchain response time and block size in the PHR platform to show the feasibility of the PHR platform. The significance of this platform is that it integrates personal health records that were dispersed by each hospital into one, and after receiving personal health record consent in advance, it is designed and implemented to be utilized safely and conveniently when necessary.

References

[1] Hyeob Kim, “Blockchain, Protection and Application of Medical Information”, Communications of the Korean Institute of Information Scientists and Engineers 9

[2] Sung won Oh, Su Min Park, Seng Phil Hong, “A Case Study for the Safe Implementation of Blockchain: Focused on Hospital Information System,” in Proc. of Symposium of the Korean Institute of communications and Information Sciences , pp. 131-132, June, 2017.

[3] Jacob, E. and S. Tai, On or Off the Blockchain? Insights on Off-Chaining Computation and Data, TU Berlin, Germany, p.13, 2017

[4] Jane Kaye, Edgar A Whitley, David Lund, Michael Morrison, Harriet Teare & Karen Melham, “Dynamic consent: a patient interface for twenty-first century research networks”, European Journal of Human Genetics (2015) 23, pp. 141–146, 2015

[5] Ryeong Eun Choi, See Yeon Kim, Seo Jeong Jang, Joo Hwan Jeong, Kyung Jae Won, Ah Ra Lee, Il Kon Kim, “Patient-centered consent system for using healthcare data based on blockchain technology”, in Proc. of the Korea Software Congress 2020, pp. 1309-1311, July, 2020.

[6] Kyung Ho Lee, “A Study on the type of BlockChain (Ethereum, Hyperledger Fabric)”, in Proc. of Symposium of the Korean Institute of communications and Information Sciences, pp. 442-443, Nov, 2018.

[7] Zaya Sukhbat, Jaewon Choi, “Blockchain technology usage on health equity: Is Blockchain technology a panacea for global health equity issues?”, Knowledge Management Research v.19, no.4, pp. 187-201, 2018.

[8] Ho Hyeorn Han, “Authentication for Blockchain Service”, Communications of the Korean Institute of Information Scientists and Engineers 38(7), pp.19-24, July, 2020.

[9] Taejin Son, “Under the Revised Personal Information Protection Act, whether blockchain-related information are personal information,” Communications of the Korean Institute of Information Scientists and Engineers 38(7), pp. 40-43, July, 2020.

[10] Ho Hyeorn Han, “Authentication for Blockchain Service”, Communications of the Korean Institute of Information Scientists and Engineers 38(7), pp.19-24, July, 2020.

[11] B. Min, “An Improvement of Hyperledger Fabric Raft Algorithm toward Enhancing Availability,” The Journal of KING Computing, vol. 17, no. 2, pp. 80–91, April, 2021

Copyright ⓒ 2023 KSII

Intermedia Network Agenda Setting of Environmental News in Korea and China: Focusing on The Relationship Between Media and Public Agenda

Yifan Wang and Kyung Han You*
Department of Media Communication, Jeonbuk National University
Jeonju, South Korea
[e-mail: yf970616@gmail.com, kuy114@jbnu.ac.kr]
*Corresponding author: Kyung Han You

Abstract

The purpose of this study is to find out what are the main issues that have been formed in Korean and Chinese societies on the topic of carbon neutrality, and whether the effect of issue setting among the media and public opinion is visible. News (49.227 articles) and commentaries (338.981 articles) in Korea and news (105.680 articles) and Weibo blog posts (136,422 articles) in China were collected from the time when carbon neutrality was announced in both countries in 2020, up to May 9, 2022, and subjected to topic modeling analysis and concor analysis with these texts. This was followed by further QAP analysis to explore Network agenda-setting effects of media and public opinion in both countries. The results found that Korean media preferred to report on carbon neutrality from the citizens perspective, while Chinese media preferred to report on carbon neutrality from the national perspective. Korean netizens often express their dissatisfaction with carbon neutrality through news comments, while Chinese netizens prefer to make statements about financial investments related to carbon neutrality. Finally, it was confirmed that the media in both countries have a weak agenda-setting effect on public opinion.

Keywords: carbon neutrality, network agenda-setting, topic modeling, concor analysis, QAP analysis

1. Introduction

With the advent of social media in the WEB2.0 age, the public now has access to a multitude of news gathering channels. In this environment, the effectiveness of agenda-setting was called into doubt. At the same time, the emergence of network agenda-setting theory, which focuses on the interaction between agenda attributes, has opened up a new avenue of research for agenda-setting. According to the network agenda-setting theory, connections between media agenda features can be transferred to public agendas. Therefore, this study analyzes the relationship between media issues and public issues in carbon-neutral Korea and China, based on the theory of Network Agenda Setting and focusing on Korea and China, which have different political systems.

2. Literature Review

2.1 Carbon neutrality

Public perceptions and attitudes toward environmental issues such as carbon neutrality are shaped by a variety of media [1]. Among these, the public receives most of its information about environmental issues from the news [2]. Faced with the same issue of carbon neutrality, media coverage in Korea and China exhibits different characteristics, and the

differences in media coverage between the two countries are closely related to differences in agenda, perspective, and actual national community interests [3].

2.2 Network Agenda Setting

Network Agenda Setting is the view that news media can indeed bundle various objects and attributes, and at the same time make these bundles of elements salient in the public mind. The NAS model suggests that the more likely a news outlet is to mention two elements at the same time, the more likely the audience is to think that they are connected [4]. Wang examined there was no network agenda-setting effect between Chinese traditional media and Weibo opinion leaders, and they were not even related [5]. Cha examined confirmed the network agenda-setting effects of media coverage on comments [6].

3. Methods

In terms of methodology, this study gathered carbon neutrality reports published in Korean and Chinese media, as well as user comments on them. A total of 49,227 Korean news, 338,981 Korean news comments, 105,680 Chinese news, and 136,422 Chinese Weibo articles were collected. The collected data was divided into three time periods depending on when the issue of carbon neutrality was first addressed and when the policy was formally stated, and media reports and user comments were examined using topic modeling approaches. Then, for each period, we extracted the co-occurrence frequency of the 50 words with the highest co-occurrence frequency to perform concor analysis, and finally, we selected the same words from the press and public agenda to construct metrics for QAP and MR-QAP analysis.

4. Results

4.1 Analysis results of Topics modeling

We employed topic modeling to investigate Korean and Chinese media reports and discovered that community development and presidential election were related topics in the Korean media, while Chinese dream, building an ecological civilization, and Winter Olympics were related topics in the Chinese media, reflecting contextual differences between the two countries. Policy, meetings, financial markets, and corporate strategy were highlighted as major issues in both Korean and Chinese media. Lastly, practical activities and market economy were among the top topics in the Korean media, while industrial development and international society were prominent in the Chinese media. Topic modeling of news comments in South Korea and Weibo in China revealed that discussions about carbon neutrality were frequently tied to people's interests. Dissatisfaction with the government, President Moon Jae-in, tax hikes, policies such as nuclear phase-out and electricity privatization, and presidential election-related topics were prevalent in South Korea, while Weibo topics in China included the stock market, personal value realization, and corporate and energy industry development.

4.2 Results of Discourses analysis

An analysis of Korean and Chinese media reveals distinct discourses focuses during different periods. In the period 1, Korean media emphasized environmental discourse, government-led industrial development, hydrogen, international cooperation, and various energy strategies. In contrast, Chinese media concentrated on prospects of carbon neutral development, innovation, and government-led energy industry advancement. During the period 2, South Korean media witnessed a decline in international cooperation rhetoric, while emphasizing corporate strategies in the carbon neutral field. Chinese media, on the other hand, focused on green development, companies' creation of carbon neutral strategies, and international cooperation. In the period 3, the discourse observed in the second period continued, with Chinese media featuring increased coverage of financial market investments and discourse related to carbon-neutral power generation.

Reviewing Korean news comments, the period 1 focused on public opinions, addressing energy growth, corruption concerns, and apprehensions about corruption in carbon neutral measures. The period 2 focused on the green energy development strategy of the Moon Jae-in

administration, the execution of nuclear power generation policies, concerns regarding negative consequences of energy challenges, and discourse about presidential contenders. In the period 3, candidates accepted carbon neutrality as a policy direction and addressed nuclear power generation strategies. On Weibo in China, the period 1 highlighted hot news, industrial development, development directions, and financial investments associated with carbon neutrality. The period 2 explored carbon neutrality in the realms of green energy, power generation, and its impact on corporate growth and financial market investments. Green energy continued to be discussed in the period 3, with the public increasingly interested in engaging in carbon neutrality, and corporations launching substantial environmental protection products.

4.3 Results of QAP analysis

The QAP correlation analysis of Korean media and comments found no correlation between coverage in periods 1 and 2 and comments in periods 2 and 3, but coverage in period 1 influenced coverage in period 2, and comments in period 1 influenced both coverage and comments in period 2. Furthermore, coverage in Period 2 only had an impact on coverage in Period 3, whereas comments in Period 2 had an impact on both coverage and comments in Period 3. The QAP analysis in China confirms the positive relationship between media coverage and Weibo. In the MR-QAP regression, however, media coverage in period 1 influenced just media coverage in period 2, whereas Weibo in period 1 influenced both media coverage and Weibo in period 2. The second period's media coverage influenced both media coverage and Weibo in the third period, whereas the second period's Weibo only influenced Weibo in the third period, not media coverage in the second period.

5. Discussion and conclusion

South Korean media comments mainly express dissatisfaction with the problems caused by carbon neutrality. China's Weibo is mostly composed of topics such as stock market trends and personal value realization. This demonstrates that the carbon neutrality draws attention from policy, market, and personal interests, and represents the fact that the public agenda is oriented on personal interests.

There is no mutual agenda-setting effect between Korean carbon neutrality media coverage and news user comments. Traditional media in China have little influence on Weibo's public agenda. Arguably, for environmental policy issues like carbon neutrality, traditional media have a weak agenda-setting effect on public opinion.

References

[1] R. Cox, Environmental communication and the public sphere: Sage, 2013.

[2] T. Wagner, “Reframing ecotage as ecoterrorism: News and the discourse of fear,” Environmental Communication, vol. 2, no. 1, pp. 25-39, 2008.

[3] S. H. Dickson, “Press and US policy toward Nicaragua, 1983–1987: A study of the New York Times and Washington Post,” Journalism Quarterly, vol. 69, no. 3, pp. 562-571, 1992.

[4] L. Guo, H. T. Vu, and M. McCombs, “An Expanded Perspective on Agenda-Setting Effects. Exploring the third level of agenda setting,” Revista de comunicación, vol. 11, no. 1, pp. 51-68, 2012.

[5] Q. Wang, “A comparative case study: Network agenda setting in crisis and non-crisis news,” Global media and China, vol. 1, no. 3, pp. 208-233, 2016.

[6] Cha, Hyeon-Ju, Tian, Zai Min, Kweon, Sang Hee, " Intermedia Network Agenda Setting Analysis on Fine Dust Issue: Focused on News Report and Naver Blog " Journal of Cybercommunication Academic Society, vol. 37, no. 4, pp. 5-52, 2020.

Copyright ⓒ 2023 KSII

Development of comprehensive information service on plant viruses in domestic and international areas

Minho Son[1]*, and Ilryong Kweon[1]
[1] DigitalPlatform HQ PODO inc.
Seoul, South Korea
[e-mail: mhson@ipodo.co.kr, irk@ipodo.co.kr]
*Corresponding author: Minho Son

Abstract

Due to various climate changes and increasing international trade in agricultural products, high-risk problems are occurring in domestic crops caused by various pests and diseases that are increasing day by day. However, it is difficult to quickly respond to control measures such as pest management without accurate diagnosis by experts. This study surveyed information on domestic and international plant viruses and compiled them into a database. Through this, we aim to provide a comprehensive information system for monitoring high-risk virus outbreaks in domestic crops by establishing a DB of monitoring information and management manuals for high-risk virus outbreaks that are suspected to be introduced from overseas.

Keywords: comprehensive management information, crop virus, GIS, monitoring

1. Introduction

If high-risk pest problems that occur in crops are not accurately diagnosed, it is difficult to systematically reduce crop damage and mitigate significant losses such as safe agricultural production. Farmers need to develop technologies that can diagnose and address high-risk pest problems directly and accurately on-site, without relying on professional agencies, to minimize the damage caused by these pests. Plant viruses are increasingly prevalent in domestic crops due to various climate changes and increased international agricultural trade, but without a specialist, it is difficult to make a rapid and appropriate response to accurately diagnose them. Unlike other pathogens, plant viruses do not have pesticides, so if the infecting strain is not removed early based on accurate diagnosis, they can quickly spread and cause significant damage to agricultural production. Since 2010, the National Crop Pest Management System (NCPMS) has been DBifying the information on major pests and providing it as a national service, but the diagnostic images such as old photos are insufficient. Some typical symptoms of plant virus diseases, such as yellowing and withering, are difficult to distinguish from physiological disorders caused by cultivation conditions and weather changes. Therefore, it is necessary to develop technologies that allow farmers to diagnose and respond directly and accurately on-site without going through professional agencies to minimize the damage caused by virus diseases. An AI-based agricultural R&D integrated platform is needed to utilize the experience, technology, and data accumulated in the agricultural field for new business innovations. Due to climate change, changes in the agricultural ecosystem, and increased international trade, new crop virus diseases continue to emerge, requiring active monitoring activities in the agricultural field to address the risk of high-risk viruses entering the country.

This study was carried out with the support of the Virus monitoring concerns over domestic inflow and development of a comprehensive information system for viruses managed by the state (PJ016389).

There is increasing demand from farmers for nationwide surveillance of domestically occurring nationally managed viruses and comprehensive management information web services to prevent crop damage, requiring comprehensive information services.

2. Objective

This study aims to develop technologies for reducing damages caused by current pest problems, such as fruit tree viruses, by developing a national virus surveillance and management system through virus disease surveys. To achieve this, first, we developed a DB map for identifying high-risk viruses and their diagnostic methods, symptoms, range, occurrence environment conditions, and management techniques, by developing a DB map for monitoring information on high-risk virus occurrence caused by foreign introductions, as well as national virus and non-quarantine virus management DB maps by country. Second, we aim to develop a comprehensive virus management information web service system for national management-level viruses in domestic farming areas. We will develop distribution maps of occurrences by region, virus lists by crop, and standard diagnostic manual DB maps, and develop comprehensive management manual DB maps for resistance varieties, prevention of occurrences, and minimization of damage, with the ultimate goal of developing a web service system. [**Fig. 1**]

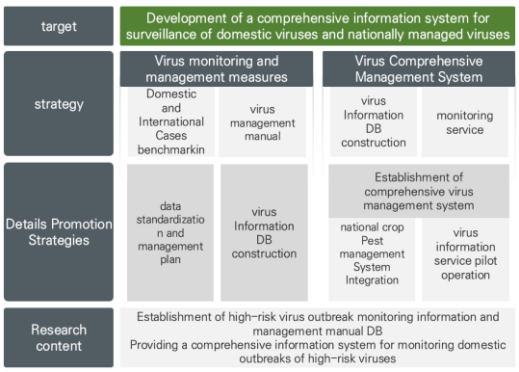


Fig. 1. Establishment Process and Objective

3. System Configuration

The system consists of App, DBMS, and GIS servers. The App is developed using a Java-based Spring Framework. The GIS server uses Geoserver, an open source, and the DBMS uses Postgresql. Space data is stored in the PostGIS in the form of a shape. Spatial representation of virus data is implemented through OpenLayers. The dataset utilizes Natural Earth's National Boundary and Land Geographic Information Service's municipal Boundary Dataset and refers to the virus survey and statistical data of the Rural Development Administration "Development of National Management Virus Distribution Survey and Management Manual" report.[1,2,3] [**Fig. 2**]

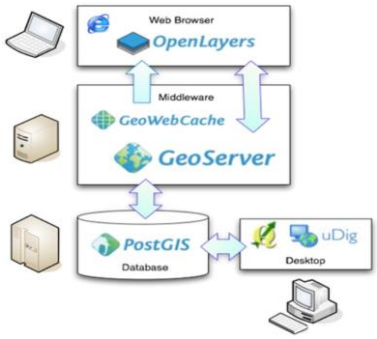


Fig. 2. System Configuration

4. Case study

Through this study, we have developed a comprehensive information system for nationally managed plant viruses and plant viruses of concern for domestic introduction. [**Fig. 3**]

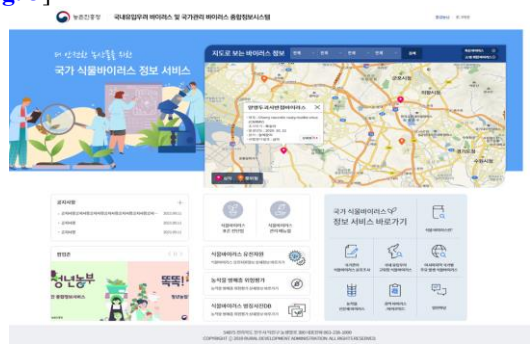


Fig. 3. Domestic Invasive Plant Viruses and National Management Plant Virus Integrated Information System

The development of a comprehensive management information web service system for plant viruses at the national management level in areas of domestic farmland involved the development of distribution maps of virus occurrence by region, lists of viruses by crop, and a database map of standard diagnostic manuals. The system also includes the development of a management manual database map for resistance varieties by crop, prevention and minimization of damage caused by virus outbreaks, and the construction of a web service system. [Fig. 4]

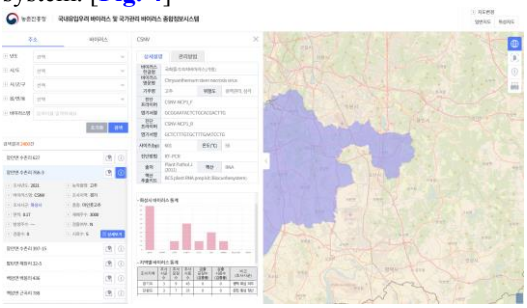


Fig. 4. Comprehensive information on plant viruses for national management

We have developed a database map for monitoring information and management manuals related to the occurrence of high-risk plant viruses that are of concern for introduction into the country. Through this effort, we have developed a database map of identification methods, rapid diagnostic tests, symptoms of damage by strain, range of infection by strain, environmental conditions for occurrence, and management measures for high-risk plant viruses. In addition, we have developed a database map of the current situation of high-risk plant virus occurrence by crop for each country, as well as management viruses for quarantine and non-quarantine viruses. [Fig. 5]

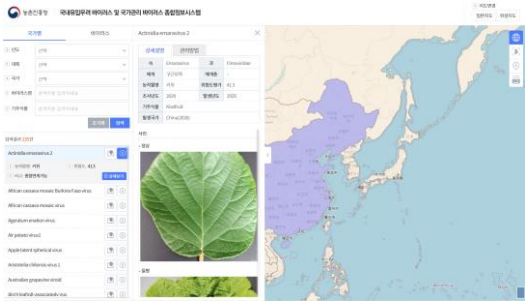


Fig. 5. Comprehensive information on plant viruses with concerns about domestic influx

In addition, we have shared various cases (18~21) of handling plant virus complaints that occur domestically, and we have maximized the utilization of plant virus information by constructing diverse plant virus information. [Fig. 6]



Fig. 6. Comprehensive information on cases of plant virus complaints

5. Conclusions

We provide comprehensive information on plant viruses with high risk of introduction into Korea through this study, and offer a service to support systematic decision-making based on such information. Additionally, by providing comprehensive information on nationally managed plant viruses, we can prevent the spread of plant viruses through faster pest control and response in agricultural fields.

References

[1] Rural Development Administration, "Development of National Management Virus Distribution Survey and Management Manual" report, 2020.

[2] Rural Development Administration, "Development of National Management Virus Distribution Survey and Management Manual" report, 2021.

[3] Rural Development Administration, "Development of National Management Virus Distribution Survey and Management Manual" report, 2022.

Copyright ⓒ 2023 KSII

Improving Blockchain Security with T-OTP-Based Node Access Control

Sungbeen Kim, Inwuk Yeo and Dohoon Kim*
Department of Computer Science, Kyonggi University
Suwon, South Korea
[e-mail: beensk@kyonggi.ac.kr, inwu03@kyonggi.ac.kr, karmy01@kyonggi.ac.kr]
*Corresponding author: Dohoon Kim

Abstract

Blockchain technology has gained tremendous popularity in various industries in recent years due to its unique properties of transparency, immutability, and decentralization. However, security remains a significant challenge in the widespread adoption of this technology. One of the key security issues is the lack of access control for data recorded on permissioned blockchain ledgers used in various industries, leading to reliability and security concerns. In this paper, we propose using Time-based One-Time Password(T-OTP) to perform access control for transaction data access by nodes participating in a blockchain network. Our research proposes a structure that authenticates nodes that can access based on the security level of the data itself using T-OTP and allows access accordingly. As a result, this enhances the security of confidential data recording in industries and improves the confidentiality of personal information managed by industries.

Keywords: Blockchain, access control, permissioned, T-OTP, security, confidentiality, information

1. Introduction

Blockchain[1] technology has revolutionized the way data is stored and managed in various industries. Blockchain is a distributed network system in which different nodes share the same ledger, ensuring integrity and reliability of stored data. However, security is one of the major challenges to the widespread industry adoption of this technology.[2] One of the major security issues with blockchain technology is the lack of access control to data recorded on permissioned blockchain ledgers. In this paper, we propose to strengthen the confidentiality of confidential data records in the industry and improve the confidentiality of personal information managed by the industry by performing access control on transaction data access of nodes participating in the blockchain network using T-OTP.[3]

Section 2 of this paper introduces the node participation method of the existing blockchain and the concept of T-OTP. Section 3 describes the access control structure of nodes participating in the T-OTP-based blockchain proposed in this paper, and finally, Section 4 discusses the contribution of this study and future research.

2. Backgrounds

2.1 Method of Node Participation in Blockchain

Node participation in a permissioned blockchain[4] refers to the process by which a node or participant is granted access to a private blockchain network. Unlike public blockchains like Bitcoin[5] or Ethereum[6], which are open to anyone, permissioned blockchains are designed for a specific group of participants, such as corporations or government agencies, that require a higher level of control and privacy. To participate in a permissioned blockchain network, a node must first obtain permission

"This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)"(2021-0-01393)

from the network administrator. This usually includes a screening process to ensure that each node meets certain criteria, such as regulatory compliance or technical requirements. When a node is approved, it is given a unique identity on the network, such as a public key used to sign and verify transactions. In a permissioned blockchain, access to the ledger is restricted to approved nodes only, and each node normally validates transactions and maintains a copy of the ledger. This allows greater control of the network and makes it easier for participants to comply with regulatory requirements and other legal obligations. However, since access control levels are not assigned to authorized nodes, all participating nodes can access all data in the ledger. In this study, to solve this problem, the concept of T-OTP is introduced to control access to data.

2.2 Access Control using T-OTP

Time-based One-Time Password (T-OTP)[3] is an authentication method that uses cryptographic algorithms to generate one-time passwords that are valid for short periods of time, typically 30 seconds. In the past, access control was performed using the concept of T-OTP for user authentication, but there was no process for controlling access to nodes participating in the blockchain network. However, T-OTP can be used to enforce node access control over data access in blockchain networks. In the context of data access control in permissioned blockchains, T-OTP can be used to authenticate which nodes have access to transactional data based on the security level of the data itself. In this paper, this enhances the overall security of data stored on a blockchain network by ensuring that only authorized nodes with appropriate security permissions can access sensitive transactional data.

3. Blockchain Node Access Control

3.1 Structure of Node Access Control in Permissioned Blockchain

The structure of the T-OTP-based blockchain node access control proposed in this study is shown in **Fig. 1**. Blockchain CA gives permission for the network through PKI[7] to nodes participating in the blockchain network, and distributes seeds according to the level of the node. The Blockchain CA creates a list of nodes that have distributed seeds in the form of a seed table and records it in the blockchain network. Only nodes whose access control has been completed based on the recorded table can perform T-OTP authentication based on the seed and view transactions in the blockchain. Below is an explanation of the role of each component that composes the structure.

Blockchain CA: Grants permission by using PKI to nodes that want to participate in the blockchain network, creates seeds according to the node level, and controls access to blockchain transactions through the structure of T-OTP.

Node: Nodes participating in the blockchain network have a level according to each position and security level, and the authority to access transactions and data recorded in the blockchain network varies depending on the level.

Blockchain Network: It refers to a network in which data of actual institutions and companies are recorded. Blocks in which transactions are recorded exist in the blockchain network, and when Blockchain CA provides seeds to each

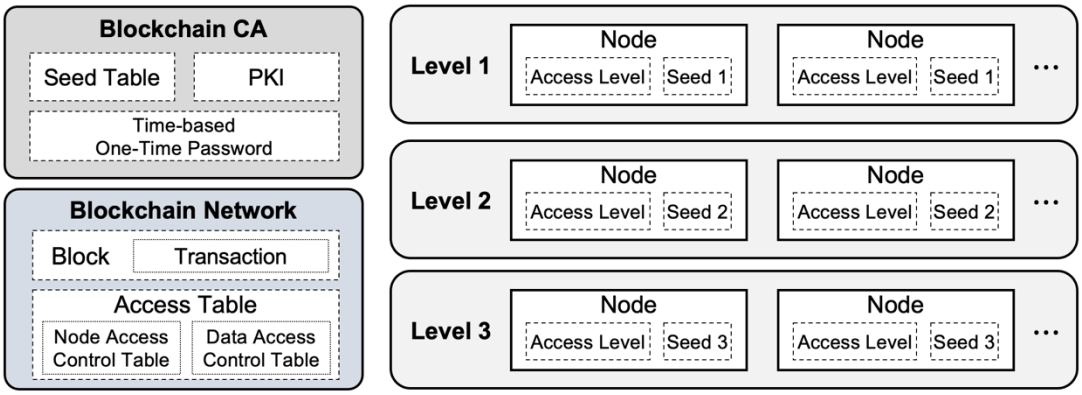


Fig. 1. Structure of Node Access Control in Blockchain Node Participation

node, it is recorded in the Access Table to ensure integrity and reliability of access control according to the level of the node.

3.2 Access Control Process using T-OTP

Only nodes authorized through the Blockchain CA participate in the blockchain network. In order to access transactions in the blockchain network, the T-OTP authentication procedure through Seed must be performed on the Blockchain CA. **Fig. 2** shows the process of controlling access to transactions recorded in the blockchain network using T-OTP. Node performs T-OTP authentication through Blockchain CA and initially issued Seed to access transaction data. Nodes that have been authenticated in the process determine the transactions that can be accessed through the Access Table recorded in the Blockchain. After that, the Blockchain CA records the access control execution result in the Blockchain Network. When authentication is complete, access is granted to the node. Nodes can access the blockchain network and view or record data if access is permitted.

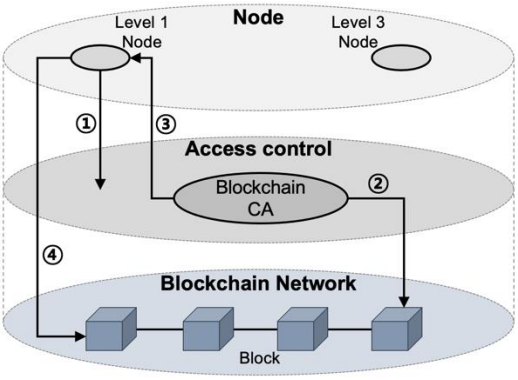


Fig. 2. Process of Node Access Control in Blockchain Node Participation

4. Conclusions

In this paper, we proposed an access control structure according to the security level for nodes participating in the blockchain network using T-OTP. The structure improves security by controlling access to confidential data stored by the industry. In conclusion, through this study, we perform transaction data access control for non-existent permissioned blockchain nodes. Since various data are shared in each industry, the company's confidentiality can be guaranteed by controlling access to these data, and security of data is secured by assigning security levels to each department and employee. Future research will evaluate the performance of the blockchain network with the corresponding structure to evaluate the usability and usability of whether it can be used in the actual industry. In addition, it is expected that the confidentiality of data shared in the industry can be further improved by applying the confidential data tracking process using hash chain.

References

[1] Nofer, Michael, et al. "Blockchain." *Business & Information Systems Engineering* 59 (2017): 183-187.

[2] Zhang, Rui, Rui Xue, and Ling Liu. "Security and privacy on blockchain." in *Proc. of ACM Computing Surveys (CSUR)* 52.3 (2019): 1-34.

[3] Sudar, Chandramohan, S. K. Arjun, and L. R. Deepthi. "Time-based one-time password for Wi-Fi authentication and security." in *Proc. of 2017 International Conference on Advances in Computing, Communications and Informatics (ICACCI)*. IEEE, 2017.

[4] Helliar, Christine V., et al. "Permissionless and permissioned blockchain diffusion." *International Journal of Information Management* 54 (2020): 102136.

[5] Nakamoto, Satoshi, and A. Bitcoin. "A peer-to-peer electronic cash system." *Bitcoin.–URL: https://bitcoin. org/bitcoin. pdf* 4.2 (2008). (Accessed on 5 May)

[6] Maurer, Ueli. "Modelling a public-key infrastructure." *Computer Security—ESORICS 96: 4th European Symposium on Research in Computer Security Rome, Italy, September 25–27, 1996 Proceedings 4*. Springer Berlin Heidelberg, 1996.

[7] Antonopoulos, Andreas M., and Gavin Wood. *Mastering ethereum: building smart contracts and dapps*. O'reilly Media, 2018.

Copyright ⓒ 2023 KSII

Web-based 3D Genome Visualization System for Whole Genome Alignments

Min Jae Kang and Hwan Gue Cho*
Department of Computer Science, Pusan National University
Pusan, South Korea
[e-mail: minjae_kang@pusan.ac.kr, hgcho@ pusan.ac.kr]
*Corresponding author: Hwan Gue Cho

Abstract

Research has been continuously conducted on a tool or web browser that visualizes a synteny block representing a homologous region between two genomes. Previous studies have used methods such as displaying only parts or simplifying the entire flow to represent a large synteny block, but it is difficult to express the relationship between chromosomes due to the limited space in two dimensions, especially when comparing multiple sequences. Therefore, this paper presents a 3D visualization approach to overcome the aforementioned problems. After processing the synteny block file generated through an external tool, a chromosome object is created in response to user manipulation, and a 3D scene is configured. Finally, the spatial advantage of the 3D scene allows users to quickly grasp the relationship between the homologous regions from various angles due to free-viewpoint changes.

Keywords: 3D visualization, Comparative genomics, Synteny block

1. Introduction

Synteny blocks play an important role in representing similar genomic regions in the chromosomes of different ancestral species. However, due to the nature of the synteny block, it is difficult to check the state as it is if there are very large or multiple regions between the sequences data. Accordingly, various techniques have been developed to explore and analyze synteny blocks such as web applications[1] and program packages[2] and interface tools[3] to confirm the similarity between genome regions at a glance through visual representations.

However, since the vast sequence information of the synteny block cannot be fully represented in 2D space, the ability to display only a part of the sequence or by simplifying the entire area has been developed. but There is a loss of detail due to simplification and a spatial shortage that still exists. Therefore, in this paper, we attempt to a web-based visualization system that deviates from the constraints of 2D space and does not require biological prior knowledge by expressing synteny blocks in 3D canvas.

2. System Overview

Sequence alignment data between 15 marine organisms was generated in the form of Pairwise mApping Format (PAF) through minimap2[4] and used as an input file. For web-based systems, we used JavaScript to implement a GUI interface for file I/O and rendering option throttling, and 3D canvas was created by implementing data parsed into 3D objects using Three.js, a 3D Open Library based on WebGL.

The system sequentially performs the processes listed below to perform a 3D visualization. First, database is created by reading the user-uploaded synteny block file. Then, user selects the desired function and its corresponding target from a list of chromosomes found in the synteny block data. Based on the selected function and target, the similar region data bet is parsed. The parsed results are used to create and sort objects and canvas, or to provide relationship information for

This research was supported by Korea Institute of Marine Science & Technology Promotion(KIMST) funded by the Ministry of Oceans and Fisheries, Korea(20220517).

selected objects. Finally, the 3D object's position and state are changed in real-time on the screen according to the user's manipulation.

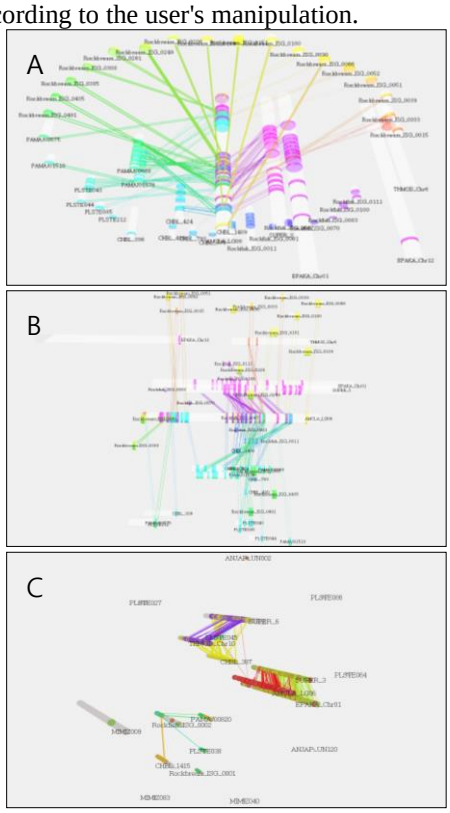


Fig. 1. 3D Visualization images
(a) Main features about AMCLA_LG06 related number of chromosomes are 42.
(b) position arrangement with (a)
(c) multi-comparison view.

Table. 1. Species abbreviations Table in **Fig. 1**.

AMCLA	Amphiprion_clarkii	PLSTE	Platichthys_stellatus
ANJAP	Anguilla_japonica	Rockbream	Oplegnathus_fasciatus
CHBL	Planiliza_haematoucheilus	Rockfish	Sebastes_schlegelii
EPAKA	Epinephelus_akaara	SUPER	Scomber_japonicus
MIMIZ	Misgurnus_mizolepis	THMOD	Thamnaconus_modestus
PAMAJ	Pagrus_major		

3. 3D Visualization Demo

The function of expressing uploaded data in 3D can be freely selected according to the object that the user wants to check among the two below, the canvas configuration is shown in the example in **Fig. 1**, and the abbreviation of marine species in the example is shown in **Table 1**.

3.1 Main Features

After the user selects the reference chromosome, all the data on the similar regions existing on other chromosomes associated with it are represented on the 3D canvas, as shown in **Fig. 1**. (a). The user can adjust the cut-off length and rendering options of the displayed regions to change the screen configuration. As can be seen from **Fig. 1** (b), when placing chromosomes with regions corresponding to the reference chromosome, they are placed according to the reference chromosome region, so the overall distribution of matching chromosomes can be confirmed.

3.2 multi-comparison in 3D

Users select several chromosomes to know how closely different chromosomes are related. Computes the proximity matrix by performing hierarchical clustering based on the size of the interrelated regions. In **Fig. 1** (c), related chromosomes are grouped together based on how similar they are, and the overall similarity pattern can be confirmed through the distance between the elements in each cluster.

4. Conclusions

This study attempted to express the relationship between synteny blocks between different genomes on a 3D canvas using WebGL, and similar regions of chromosomes were expressed according to the user's choice to easily confirm what the data means. In future research, external functions such as user-friendly interfaces and detailed information display functions, and internal functions such as database and 3D object management optimization will be developed.

References

[1] M. Johnson, et, “Ncbi blast: a better web interface,” *Nucleic acids research*, vol. 36, pp. W5–W9, 2008.
[2] Y. Wang et al “Mcscanx: a toolkit for detection and evolutionary analysis of gene synteny and collinearity,” *Nucleic acids research*, vol. 40, no. 7, p. e49, 2012.
[3] C. Chen et al “Tbtools: an integrative toolkit developed for interactive analyses of big biological data,” *Molecular plant*, vol. 13, no. 8, pp. 1194–1202, 2020.
[4] LI, Heng. “Minimap2: pairwise alignment for nucleotide sequences,” *Bioinformatics*, vol. 34, no. 18, pp. 3094-3100, 2018.

Copyright ⓒ 2023 KSII

The Use of Grammar Transducer in Korean End-to-end Automatic Speech Recognition Models

Eunsoo Cho, Hosung Park, Hayoung Kim, and Ji-Hwan Kim*
Department of Computer Science, Sogang University
Seoul, South Korea
[e-mail: choeunsoo@sogang.ac.kr, hosungpark@sogang.ac.kr, hayoung@sogang.ac.kr, kimjihwan@sogang.ac.kr]
*Corresponding author: Ji-Hwan Kim

Abstract

Fundamental object of Automatic Speech Recognition (ASR) is to recognize most likely word segments for given audio inputs. Basic approach to decoding most probable output to input features requires lexicon dictionary and language model applied to acoustic model. They are constructed in forms of weighted finite state transducer (WFST), which can incorporate given information efficiently through composition and optimization operations. Recent development of end-to-end model and its application to ASR made it possible to train ASR model with only transcription to audio files. However, End-to-end ASR models based on CTC loss function have an inherent issue, which does not allow the model to train language information provided in lexicon and grammar. In order to overcome lack of language information, we propose use of Grammar Transducer, langauge model in WFST to further improve ASR performance. In testset of Zeroth Korean news data corpus, character error rate (CER) for NVIDIA Korean Citrinet and Conformer were 4.95 % and 5.39 %. Lexicon Transducer improved End-to-end ASR performance to 2.10 % in Citrinet and 2.73 % in Conformer. Integration of grammar information improved problems with out of context vocabularies or inaccurate grammar in end-to-end ASR models hypothesis. It resulted in CER of 0.66 % and 0.78 % in NVIDIA Korean Citrinet and Conformer.

Keywords: automatic speech recognition, end-to-end model, weighted finite-state transducer

1. Introduction

The objective of Automatic Speech Recognition (ASR) is to derive the most probable sequence of word segments(W) or other model output units for given input acoustic features(O). End-to-end ASR model maps acoustic feature sequences directly to output layer label sequences. This is made possible by Connectionist Temporal Classification (CTC) [1] loss function.

While end-to-and ASR model greatly simplifies ASR framework, and is able to train context information from input feature sequence, it cannot train context information, or grammar and lexicon information from transcriptions. Due to this structural limitation of CTC, end-to-end model may output grammatically incorrect inference results or vocabularies with spelling errors.

In this paper, we suggest composition with Grammar Transducer with end-to-end ASR model to integrate external grammar information in order to resolve these inference errors.

2. Related Works

2.1 End-to-end Model

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00621,Development of artificial intelligence technology that provides dialog-based multi-modal explainability)

Transformers [2] introduced self-attention mechanism which made it possible to learn connections between all input elements regardless of distance or time discrepancy between them. Variations of Transformers for sequence based tasks were suggested, including Conformer[3]. Conformer uses the same architecture of decoder as Transformer, but with an encoder which combines CNN layers with self-attention mechanism to capture both local and global features of input audio features.

Citrinet [4] is a ResNet-based ASR model which showed close performance to autoregressive ASR models. The performance of the two models on the Librispeech test-clean shows a word error rate (WER) of 3.6%.

2.2. Connectionist Temporal Classification

Initially, ASR required segmentation process, which labelled each feature to corresponding phoneme. Development of CTC made ASR model without segmentation possible. The process of decoding the most probable output to given acoustic feature includes Path Probability Calculation, the process in which model calculates the generation probability of each path. A path (π) is a sequence of probable CTC output labels, which includes blank labels and repeated labels. In this stage, CTC assumes conditional independence for the output probability of the label in each frame. This assumption is what makes learning grammar in CTC based end-to-end models impossible.

$$p(\pi|x) = \prod_{t=1}^{T} y_{\pi_t}^{t} \quad (3)$$

x	input sequence of acoustic feature
t	possible label in path for time step t
$y_{\pi_t}^{t}$	probability of output label t in time step t

2.3. Weighted Finite-State Transducer

Weighted Finite-State Transducer is a finite automata in which each state transition is associated with an input/output label pair and a weight [5]. It is a finite state automaton which accepts an input sequence and transforms it into an output sequence by a structure which consists of states and transitions. Each transition has an input label, output label, and weight which denotes probability of transition to the connected state. This allows separate training of necessary components and flexible integration through WFST operations.

3. Model Architecture

3.1. Acoustic Model

Acoustic model's role is to accept acoustic features as input and convert them into sequence of CTC output labels. In this paper, we used CTC based end-to-end models such as Citrinet and Conformer, currently well-known models that were widely used for their accuracy.

3.2. Utterance Transducer

Utterance Transducer models output probability of each CTC output label in every time step. Every time step, there are edges which represent all U labels of CTC output layer tokens. The weight of each edge is the probability of each token for every time step. Therefore, Utterance Transducer can represent every possible path for given acoustic features extracted from given input.

3.3. Lexicon Transducer

Lexicon Transducer is a lexicon in form of WFST. Its input sequence is sequence of CTC output layer tokens and the output is vocabulary or word segment which consists of given tokens as well as weight, which is output probability of the calculated output sequence.

3.4. Grammar Transducer

The final component of the architecture is Grammar Transducer. Its purpose is to convert the sequence of output vocabularies from lexicon transducer to grammatically correct sentences. Grammar Transducer is a language model trained on text corpus in given language mapped into form of WFST.

3.5. Search Graph

All components are combined into a single search graph to calculate the most probable word segments sequence for given input. Individually trained components are then combined into a search graph through WFST operations

including composition, determinization, and minimization.

4. Experiment

4.1. Models

In this paper, we experimented using pretrained automatic speech recognition models. To maintain objectivity, we use pretrained models from NVIDIA model catalogue, Conformer CTC Large[1] and Citrinet 1024[2].

4.2. Datasets

The test dataset to measure performance or accuracy of ASR model is Zeroth Korean [3] dataset. Zeroth Korean dataset includes utterances of 105 different people of 3,000 sentences derived from scripts of news broadcast. There are total of 22,263 utterances.

4.3. Results

Among Zeroth Korean testset, we selected the 100 sample sentences to analyze the category of errors in inference of news broadcast domain. It revealed that most common error is space error. Spelling error, meaning error, and grammar error follow.

Table 1. Categories of Error in Korean End-to-end ASR Models

Model	Space (%)	Spelling (%)	Meaning (%)	Grammar (%)
Conformer	41.5	29	22	13
Citrinet	29	25.5	18	12.5

Grammar Transducer was created from utterances of training dataset transcriptions trained as tri-gram models, and then converted into form of WFST graph. The inference results are shown in **Table 2**. While Lexicon Transducer greatly reduced spelling, space, and grammar errors, context and grammar information and raised performance in news domain from CER 5.39 % to 0.78 %, furthur information provided by grammar transducer, grammar and context information also contributed to slight increase in performace, from CER 0.78 % to 0.66 %.

Table. 2 Performance of NVIDIA pretrained Korean ASR models (L and G each denotes Lexicon Transducer and Grammar Transducer)

Model	Character Error Rate
Conformer Large	5.39 %
Citrinet 1024	4.95 %
Conformer + L + G	0.78 %
Citrinet + L + G	0.66 %

5. Conclusions

In this paper, we improved inference errors in Korean Automatic Speech Recognition models due to shortcomings caused by inherent structure of Connectionist Temporal Classification. By using grammar and lexicon transducers, which provided these information, it showed significant increase in accuracy, as shown in **Table 2**.

References

[1] A. Graves, *et al.*, “Connectionist temporal classification. Supervised sequence labelling with recurrent neural networks”, in *Proc. of International Conference on Machine Learning (ICML)*, pp. 369-376, 2006.

[2] A. Vaswani, *et al.*, “Attention is All You Need”, *Advances in neural information processing systems*, Vol. 30.

[3] A. Gulati, *et al.*, “Conformer: Convolution Augmented Transformer for Speech Recognition”, in *Proc. Interspeech*, pp. 5036-5040, 2020.

[4] S. Majumdar, *et al.*, "Citrinet: Closing the gap between non-autoregressive and autoregressive end-to-end models for automatic speech recognition," *arXiv preprint arXiv:2104.01721*, 2021.

[5] M. Mohri, *et al.*, “Weighted finite-state transducers in speech recognition,” *Computer, Speech & Language*, Vol. 16, No. 1, pp. 69-88, 2002.

[1] catalog.ngc.nvidia.com/orgs/nvidia/teams/riva/models/speechtotext_ko_kr_Conformer

[2] catalog.ngc.nvidia.com/orgs/nvidia/teams/riva/models/speechto text_ko_kr_Citrinet

[3] http://www.openslr.org/40/

Copyright ⓒ 2023 KSII

Improving Conformer-CTC based Korean ASR Performance with Data Augmentation and Language Model

Yoonseo Chung, Hosung Park, Eunsoo Cho and Ji-Hwan Kim
Department of Computer Science and Engineering, Sogang University
Seoul, South Korea
[e-mail: dbstj1231@sogang.ac.kr, hosungpark@sogang.ac.kr, choeunsoo@sogang.ac.kr, kimjihwan@sogang.ac.kr]
*Corresponding author: Ji-Hwan Kim

Abstract

In this paper, we present a Korean ASR that improves performance by applying data augmentation and language models. ASR model is trained on AIHub's in-vehicle conversation and command dataset. We extracted log-Mel spectrograms from audio signals and applied data augmentation techniques including speed perturbation, adding white noise, pitch shift, and specaugment. We used a tokenizer using sentencepiece to create subword segments of the transcription for training. We built a Conformer-CTC based ASR model using Nemo toolkit and trained it with the augmented data. In the decoding process, a language model was applied. This method showed improved speech recognition performance and finally achieved a CER of 0.15%.

Keywords: Automatic Speech Recognition, Conformer-CTC, Data Augmentation, Language Model

1. Introduction

Automatic Speech Recognition (ASR) is the process of converting speech signals into word sequence. Recently, the performance of the Korean speech recognizer has improved as various Korean datasets have become accessible from AIHub e.g. In-Vehicle Conversation and Command dataset [1]. However, it is difficult to improve the performance of Korean ASR by simply training on dataset alone because even a large dataset might not cover all the possible variation in Korean speech. Therefore, this paper applies data augmentation and external language models to improve performance of Korean ASR for in-vehicles conversation and command.

The model training process is as follows: First, log-Mel spectrograms were extracted from audio signals, and data augmentation techniques including speed perturbation, white noise addition, pitch shift, and specaugment [2] were applied to generalize the dataset. We also used sentencepiece as tokenizer to create segments of the transcription for training. We built a Conformer-CTC based model [3,4] using the Nemo toolkit and the model is trained with AIHub's in-vehicle conversation and command dataset. For the decoding process, we applied a language model through beam search [5] decoding using shallow fusion [6]. As a result, this study showed improved speech recognition performance and finally achieved a Character Error Rate (CER) of 0.15%.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00377, Development of Intelligent Analysis and Classification Based Contents Class Categorization Technique to Prevent Imprudent Harmful Media Distribution)

2. ASR Architecture

2.1 Dataset

We used 284 hours of in-vehicle conversation and command from AIHub [1] as train and test dataset. 89.39% of the audio data is between 3 and 5 seconds long. Out of 284 hours of data, 270 hours were used as training data and 14 hours as test data. And we split the training set to training and validation set in a 95:5 ratio. Both training and test datasets are provided as audio files in PCM format, with a sampling rate of 16k and mono channel.

2.2 Data Augmentation

We applied 4 data augmentation methods. First, we used speed perturbation. This method generates new data by slightly adjusting the speed of the audio signal. Second, we add white noise. Third, we shift the pitch of audio data, which augments the data by changing the pitch of the voice. Last, we applied specaugment [2]. This method augments data by applying random distortions to the spectrogram. These techniques allow the model to respond to different audio signal variations.

2.3 Conformer-CTC

Conformer [3] is a model that combines the advantages of Transformer and Convolutional Neural Networks (CNN). By combining the self-attention mechanism of Transformer with the locality-sensitive nature of CNNs, Conformer can learn both long-range and short-range dependencies of audio sequences. Connectionist Temporal Classification (CTC) [4] is used because CTC loss allows the model to learn without the exact alignment of the output sequence. In this paper, we utilized the Nemo toolkit to build a Conformer-CTC based speech recognition model. And we build tokenizer using sentencepiece to split the transcription into 4000 subword segments.

2.4 Language model

Decoding was performed by combining the trained model with an external language model. We used the training data to generate a 6-gram language model with Witten-Bell Smoothing. The generated language model was applied using beam search [5] decoding with shallow fusion [6].

3. Results

We used in-vehicle conversation and commands dataset from AIHub to train and test the model. We used the CER metric as the evaluation metric. In our experiments, before applying data augmentation and language models, the ASR model had a CER of 0.18%, and after applying them, it showed improved result as 0.15%.

4. Conclusions

In this paper, we present Conformer-CTC based Korean ASR for in-vehicle conversation and command. We improve performance of model by applying data augmentation and language models. The performance evaluation results show a CER of 0.15%, which proves the effectiveness of data augmentation and language model.

References

[1] AIHub Homepage, Available online: https://aihub.or.kr/aihubdata/data/view.do?currMenu=115&topMenu=100&aihubDataSe=realm&dataSetSn=112 (accessed on 15 May 2023)

[2] D.S. PARK, et al., “SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition,” arXiv preprint arXiv:1904.08779, 2019.

[3] A. Gulati, et al., "Conformer: Convolution-augmented Transformer for Speech Recognition," in Proc. of the Interspeech, pp. 25-29, 2020.

[4] A. Graves, et al., "Connectionist Temporal Classification: Labelling Unsegmented Sequence Data with Recurrent Neural Networks," in Proc. of the 23rd Int. Conf. on Machine Learning, pp. 369-376, 2006.

[5] S. Wiseman, et al., "Sequence-to-Sequence Learning as Beam-Search Optimization," in Proc. of the EMNLP, pp. 1296-1306, 2016.

[6] J. Chorowski, et al., "Towards Better Decoding and Language Model Integration in Sequence-to-Sequence Models," in Proc. of the Interspeech, pp. 523-527, 2017.

Copyright ⓒ 2023 KSII

Predicting Winding Machine Quality based on Deep Time-Series Data Analysis

Seung-Joon Lee*
Korea Institute of Industrial Technology
Ansan, South Korea
[e-mail: sjlee10@kitech.re.kr]
*Corresponding author : Seung-Joon Lee

Abstract

This paper is about improving the quality of the production equipment system.The equipment used in this paper is a winding machine equipment that produces coil-wound slots, which are parts for hybrid vehicle motors, it deals with time series data analysis technology that can obtain real-time speed, acceleration, and position information from a number of motors that make up the winding machine and predict output quality by obtaining real-time tension data from the winding machine control system. The technology proposed in this paper is not a conventional method of obtaining images of products produced through industrial cameras and evaluating the quality of products , it is about artificial intelligence technology that can evaluate the quality of production products through changes in tension data of windings and real-time driving data of motors that are driven when parts are produced, and it is a study using LSTM algorithm. It also requires a very low level of complexity compared to traditional image-based quality evaluation techniques in terms of computational and memory complexity, ensuring high efficiency.

Keywords: Deep Time-Series Data Analysis, Machine Learning , Production and manufacturing system, Bidirectional LSTM

1. Introduction

As the manufacturing industry becomes more competitive, companies are looking for new ways to optimize production processes and reduce costs. One of the promising technologies that has attracted attention in recent years is artificial intelligence (AI). Using AI algorithms, manufacturers can automate repetitive tasks, improve quality control, and implement predictive maintenance. However, adopting artificial intelligence in manufacturing equipment can be a complex and difficult process[1] and requires careful planning and execution. This paper identifies the production processor of the production product to introduce AI technology from the actual operating equipment, explains the data acquisition method, and through the development of AI models, we will discuss ways to improve the quality of production products and predictions. Through this paper, we provide a comprehensive guide for companies that want to introduce AI into manufacturing equipment based on real-world cases. By following these steps, companies expect to improve their manufacturing processes, increase efficiency, and remain competitive in increasingly challenging industries.

2. Acquisition of data from a winding equipment controller

2.1 Motor parts production equipment for hybrid vehicles

In order to acquire 10 motor data (speed, torque, position) and tension information of coil windings in the winding coil production equipment, data was acquired and analyzed in

This research was supported by a research grant from the Korea Evaluation Institute Of Industrial Technology (Grant No: 20015802).

real time from the control system of the existing equipment.

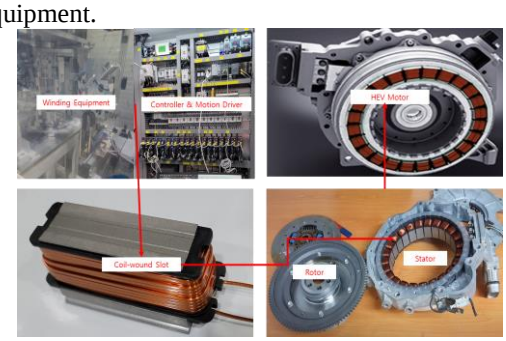
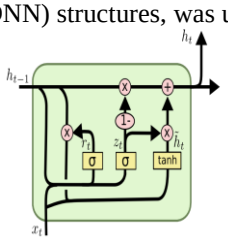


Fig. 1. Winding Equipment for HEV Motor

3. Time Series Data Analysis

3.1 Winding machine equipment data analysis and visualization

Time-series data analysis refers to extracting meaningful features from data accumulated over a specific time set by the user and analyzing the results through inference, Long Short Term Memory (LSTM), a type of Recurrent Neural Network (RNN) among Deep Neural Network (DNN) structures, was utilized. [2]



$$z_t = \sigma(W_z \cdot [h_{t-1}, x_t])$$
$$r_t = \sigma(W_r \cdot [h_{t-1}, x_t])$$
$$\tilde{h}_t = \tanh(W \cdot [r_t * h_{t-1}, x_t])$$
$$h_t = (1 - z_t) * h_{t-1} + z_t * \tilde{h}_t$$

Fig. 2. The overall structure and reasoning process of the LSTM

While analyzing the data of the winding machine, data unrelated to the actual quality was removed. In addition, in order to utilize the Dynamic Time Warping (DTW) algorithm, data normalization of each measurement was performed to solve the problem of being excessively dependent on a specific measurement. As shown in the table below, the measurements that are important to quality were deduced.

Table. 1. Euclid distance using DTW from time series data showing normal quality of winding machine output

Data ID	OK #1 VS OK #2	OK #1 VS FAIL #1	OK #1 VS FAIL #2
1	67.7485	145.5859	143.2867
10	11.3112	8.9636	22.7429
24	0.2266	0.2279	0.8856

Data ID No. 1 means "Torque", No. 10 means "AXIS 3 Actual Velocity", No. 24 means "AXIS 10 Actual Torque".

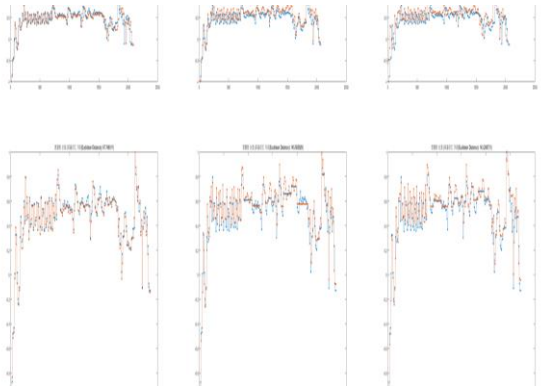


Fig. 3. Comparison of DWT results for "Torque" measurements

This is an image visualizing the DWT comparison result for "Torque" among 31 types of measurements. Unlike the existing simple distance comparison method, the DWT method can be applied regardless of the length of time series data, so even if each data has a different time series length, quantitative time series data comparison can be performed. The three columns in **Fig. 3** compare normal and normal, normal and bad 1, normal and bad 2, and in each detailed image, the first row expresses a pattern on time series data before applying the alignment algorithm, and the second row shows the result of comparing the two signals after applying the alignment algorithm. Therefore, the smaller the difference between the two signals in the second row, the lower the Euclidean distance value applied with DWT, which can be expressed as similar between the two signals. As can be seen in **Fig. 4**, different normal winding machine time series data show similar signal types. However, the trend of normal and bad data seems similar even in the comparison of aligned signals, but differences in the timing or size of the inflection point can be seen. In conclusion, it can be seen that these "Torque" measurements are significant measurements in the quality classification of winding output.

4. Winding machine output quality prediction

4.1 Prediction of output quality of winding machine based on deep neural network

In this study, it is a sequence classification technology that classifies output quality from a given overall trend of time series data, and data analysis in only directions that match the passage of time can cause unnecessary bias in the inference results of neural networks. Therefore, the algorithm used in this paper used Bidirectional LSTM to perform a comprehensive data analysis of a given winding time series data [3], considering the forward and reverse directions of time-steps.

4.2 Bidirectional LSTM -based winding machine output quality classification learning results

Experimental results after application to winding output quality classification and prediction technology using deep neural networks confirmed that the introduction technology using both L2 normalization and Bidirectional LSTM achieved 91.38% classification accuracy.

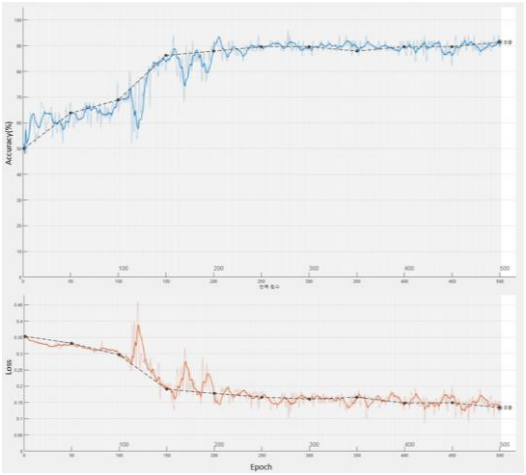


Fig. 4. Learning and Verification Graphs for Bidirectional LSTM-Based Models

As shown in **Fig. 4**, by using normalization, we can see that the accuracy (black dotted line) graph of the represented verification dataset is placed in an advantage over the learning graph. This confirms that no overfitting problem occurred during the entire learning epoch, and in this study, we can demonstrate that the selection of appropriate learning configuration parameters according to the amount of data set achieves high-quality classification performance by utilizing Bidirectional LSTM based on actual acquired data from the winding machine.

4. Conclusions

In this paper, quality analysis and prediction were possible using two-way LSTM algorithms, and the direction of how artificial intelligence technology can be applied and better products can be produced through existing real-world equipment was presented. In the future, I intend to increase the performance of quality analysis and prediction by securing more data by applying it to actual production sites.

References

[1] Dr. P. Karuppusamy, "Machine Learning Approach to Predictive Maintenance in Manufacturing Industry," *Journal of Soft Computing Paradigm (JSCP)*, Vol. 2, No. 4, pp. 246-255, 2020.

[2] A. Essien and C.Giannetti, "A Deep Learning Model for Smart Manufacturing Using Convolutional LSTM Neural Network Autoencoders," *IEEE Transactions on Industrial Informatics*, Vol. 16, Issue 9, pp. 6069-6078, 2020.

[3] Junchuan Shi, Dikang Peng, Zhongxiao Peng, Ziyang Zhang, Kai Goebel and Dazhong Wu, "Planetary gearbox fault diagnosis using bidirectional-convolutional LSTM networks," Mechanical Systems and Signal Processing, Vol 162, 2022.

Copyright ⓒ 2023 KSII

A Study on Playbook Automatic Generation Method for Threat Analysis Automation

Seul-Ki Choi*, Taeeun Kim, Kijong Son, Saewoom Lee and Taehyeon Kim
Korea Internet & Security Agency
Republic of Korea
[e-mail: skchoi@kisa.or.kr, tekim31@kisa.or.kr, skj@kisa.or.kr, sprout2@kisa.or.kr, thkim@kisa.or.kr]
*Corresponding author: Seul-Ki Choi

Abstract

The manpower required to analyze and respond to rapidly increasing security events is currently at its limit. Recently, security operation centers are introducing SOAR technology that automates analysis and response tasks with playbooks. Although the SOAR solution has reduced the manpower required for analysis and response, manpower is being wasted in generating and managing playbooks. Therefore, in this paper, we propose a method that automatically generate playbooks using artificial intelligence technology.

Keywords: Security Monitoring, Security Orchestration Automation Response, Threat Analysis, Automation

1. Introduction

Recently, Security Operation Centers (SOCs) are facing the threat of cyber attacks that are evolving into very complex and deadly forms. In addition, threats to SOC operations are rapidly increasing as artificial intelligence technology is used to create new types of security threats and attempt to automate attacks[1]. On the other hand, there were limitations in responding to cyber threats due to the lack of manpower and passive response. Therefore, recently in order to respond to these threats a security control technology called SOAR (Security Orchestration Automation and Response) [2] which can automate the passive response part as much as possible, has recently been in the spotlight.

Through the introduction of SOAR, resource consumption for responding to security events could be reduced, but as the existing personnel were assigned to create the playbook [3], which is the core of SOAR operation, the workload of SOC operation personnel did not change significantly.

Therefore, in this paper, we propose a method that can automatically generate playbooks save a lot of time and human resource consumption for playbook creation when introducing and operating SOAR and increase the efficiency of SOC operation. In addition, by generating the automatic generated result as an independent file(JSON), the proposed method is intended to be utilized in various SOAR solutions.

2. Playbook Automatic Generation Method

2.1 Playbook Template and Playbook

- Playbook Template

It shows an approximate procedure for analyzing and responding to detection events, and there is no definition for component information in the form of SW and API that can actually be executed. The more detailed the expression of the action defined in the playbook template, the more detailed the playbook template can be expressed.

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grand funded by the Korea government(MSIT)(No.2021-0-00358, AI·Big data based Cyber Security Orchestration and Automated Response Technology Development).

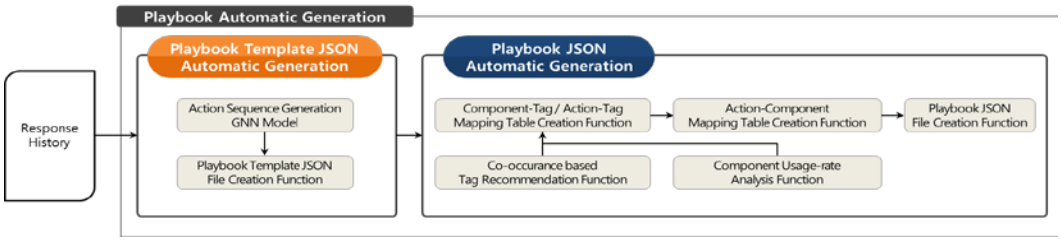


Fig. 1. Response history-based playbook automatic generation process

▪ Playbook

As a procedure in which executable components (SW or API) are mapped to analyze and respond to detection events, it is possible to perform analysis and response tasks only if they are matched with detection events.

2.2 Playbook Generation Process

Fig. 1 shows the process of automatically generating playbooks based on response history.

▪ Response history preprocessing

This is the process of analyzing and extracting the action of the playbook template and the sequence information of actions by analyzing the syntax of the response history (including ticket sending data) managed by the SOC.

▪ Generate a playbook template

This is the step of using the action and action sequence information extracted from the response history data as node and edge of the graph, respectively. For each detection name or detection event, the graph type with the strongest connection between Action and Action sequence information is recommended. Finally, the recommended result is used as a playbook template.

▪ Generate a playbook

This is the process of creating a playbook that can actually be executed in the SOAR system by using playbook template information. This is the step of mapping components connected to action of playbook template and tag information that represents characteristics of each action and component is used for the information to connect the action and component. The result of playbook creation also creates a JSON file containing playbook information as a result file so that it can be used in various SOAR systems.

2.3 Tag Information Management

▪ Co-occurrence-based tag recommendation function

When the SOC personnel inputs tag information representing the characteristics of action and component, the tag information entered by the SOC personnel and frequently appearing tag information are recommended so that the tag information for action and component with similar characteristics can be similarly managed.

▪ Component utilization analysis function

When creating a playbook through playbook template information, you can use this function when you want to create a playbook centered on components that are frequently used in the SOC. Although there may be multiple components that perform the same role, playbooks that perform the same role can be prevented from being created by creating playbooks based on components mainly used in the SOC.

3. Conclusions

This paper proposed an automatic playbook generation method to solve one of the limitations of SOAR technology, which is that a lot of resources are consumed in playbook creation. The proposed method can reduce the time-consuming resource consumption that occurs in the playbook creation process when SOAR is introduced by quickly generating an initial playbook using the data possessed by the existing SOC.

References

[1] J.K.Hong, B.Y.Lee, “Artificial Intelligence-based Security Control Construction and Countermeasures,” *Journal of Korea Contents Association*, vol.21, no.1, pp.531-540, Jan. 2021.
[2] Channi Islam, Muhammad Ali Babar and Surya Nepal, “A Multi-Vocal Review of Security Orchestration,” ACM Computing Surveys, Vol.52, No.2, Article 37, Apr. 2019.
[3] Yan Zhao, Yuanbo Guo, “Playbook-Centric Scalable SOAR System Architecture,” 7th International Conference on Eletronic Technology and Information Science, Jan. 2022.

Copyright ⓒ 2023 KSII

Wearables for Physical Activity Verification

Ye Won Park[1], Yubin Yun[1] and Hyoseok Yoon[2*]
[1]Division of Computer Engineering, Hanshin University
[2]School of Computing and Artificial Intelligence, Hanshin University
Osan, Gyeonggi-do, South Korea
[e-mail: ywpark0124@hs.ac.kr, yyb09@hs.ac.kr, hyoon@hs.ac.kr]
*Corresponding author: Hyoseok Yoon

Abstract

We present the design and implementation of a wearable physical activity verification method using smart glasses and smartwatches. Our aim is to provide an in-depth analysis of the benefits and challenges associated with this approach. The benefits highlighted include improved accuracy and reliability in tracking physical activity, a better understanding of user behavior, enhanced user experience, and the versatility of the devices. However, the potential problems discussed include the cost of these devices, privacy concerns related to data collection, and potential discomfort for users during extended wear. To address these challenges, we emphasize the importance of considering additional verification methods, ensuring data security measures, and monitoring user feedback to refine the design. Ultimately, our work contributes to the field of HCI by showcasing the potential of wearable devices in tracking physical activity, and by highlighting the factors that should be considered to optimize user experience and data collection.

Keywords: Physical activity, smart glasses, smartwatch, wearable device

1. Introduction

Human-Computer Interaction (HCI) researchers have shown a growing interest in wearable technologies that help individuals monitor and track their physical activities. A form of self-tracking known as quantified self using wearable devices such as smartwatches and fitness trackers. These devices have shown to be effective in tracking various physical activities, such as eating, walking, running, cycling, and swimming. However, there are still some limitations to these devices, particularly in their accuracy and reliability of verifying more complex and procedural routine executions. Therefore, there is a need for a more practical, reliable, and accurate method to track physical activity. One potential solution is to use an intimate connection between smart glasses [1] and smartwatches to verify, monitor, and record physical activity of the wearer through a coordinated computation.

2. Wearable Activity Verification

We propose a wearable physical activity verification (WPAV) method to help improve the accuracy, reliability, and practicality of physical activity tracking. A conceptual design of WPAV method is depicted in **Fig. 1**.



Fig. 1. A conceptual design of WPAV method

For example, a user may practice basic dance steps with visualized instructions on their smart glasses with timely feedback as shown in **Fig. 2**.

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1F1A1063340).



Fig. 2. Dance basic steps examples excerpt from bounce (https://youtu.be/6INnoBQsXrM) and isolation (https://youtu.be/OXujvIc9_mg)

There are several benefits to designing and implementing a wearable exercise verification method using smart glasses and smartwatches in HCI studies.

Improved accuracy and reliability: Smart glasses and smartwatches can measure a wider range of physical activities, such as weightlifting, yoga [2], Pilates, dance, and exergame [3] which are not easily tracked by traditional fitness trackers.

Better understanding of user behavior: By tracking physical activity in real-time, researchers can observe how individuals interact with their environment and adjust their activity level.

Enhanced user experience (UX): Smart glasses and smartwatches can provide a more seamless UX. With these devices, users can receive real-time feedback on their physical activity, which can help motivate them to stay active.

Fig. 3 shows another example of visualized instructions on the wearer's smart glasses so that the user can better know the appropriate placement of her feet on the equipment.



Fig. 3. Pilates examples excerpt from leg pull, visualized instruction on the user's smart glasses

With the proposed WPAV, there are also some technical challenges to overcome.

Cost: Smart glasses and smartwatches can be expensive, and not everyone may be able to afford these devices. This could limit the participation of certain groups of people in HCI studies that require the use of smart glasses and smartwatches.

Privacy concerns: Smart glasses and smartwatches can collect a significant amount of personal data, such as location, heart rate, and activity level. Researchers must ensure that the data collected is secure and that participants' privacy is protected.

Comfort: Wearable devices can be uncomfortable to wear for extended periods, particularly during physical activity [4]. This discomfort could lead to a decrease in participation and data accuracy.

3. Conclusions

The use of WPAV methods using smart glasses and smartwatches offers numerous benefits in HCI research. It provides researchers with accurate and objective data on user behavior during exercise, allows for more efficient and effective data collection, and offers potential for new insights into user behavior and preferences. Yet, to mitigate various technical concerns, additional verification methods, data security measures, and UX monitoring are required.

References

[1] O. Amft, F. Wahl, S. Ishimaru et al., "Making regular eyeglasses smart," *IEEE Pervasive Computing*, Vol 14, No. 3, pp. 32-43, 2015.

[2] H.-Y. Jo, L. Seidel, M. Pahud et al., "FlowAR: How different augmented reality visualizations of online fitness videos support flow for at-home yoga exercises," *in Proc. of CHI 2023*, 2023.

[3] S. Cmentowski, S. Karaosmanoglu, L. E. Nacke et al., "Never skip leg day again: Training the lower body with vertical jumps in a virtual reality exergame," *in Proc. Of CHI 2023*, 2023.

[4] Y. Lee, S. K. Kim, H. Yoon et al., "Smart glasses and telehealth services by professionals in isolated areas in Korea: Acceptability and concerns, *Technology and Health Care*," Vol 31, No. 3, pp. 855-865, 2023.

Copyright ⓒ 2023 KSII

Beamforming with Generalized Power Iteration to Maximize Sum Rate in Maritime Relay Communications System

Tae-Woo Kim[1], Kyeongjea Lee[1], Sung-yoon Cho[2], Ki-won Kwon[2] and Dong Ku Kim[1]*
[1] Department of Electrical engineering, Yonsei University, Seoul, South Korea
[2] Smart Network Research Center, Korea Electronics Technology Institute Seoul, South Korea
[e-mail: tw.kim@yonsei.ac.kr, kjlee92@yonsei.ac.kr, sycho@keti.re.kr, kwonkw@keti.re.kr, dkkim@yonsei.ac.kr]
*Corresponding author: Dong Ku Kim

Abstract

This paper shows the maximized sum spectral efficiency with the generalized power iteration method. We consider a system that communicates from the ground base station on the coast to vessels floating in the sea far from the coast through the several relay buoys in the near-coast sea. We find the beamforming vector of the ground base station for maximizing the sum spectral efficiency of vessels by applying the generalized power iteration method. We show the formulation, derive the optimization problem, and show the performance with simulation results.

Keywords: Maritime relay communication, Ocean channel model, Beamforming, Generalized power iteration

1. Introduction

As far as we know, the following papers studied the maritime relay communications system, increasing coverage based on a multi-hop repeater with a maximum ratio combining technique [1], increasing the sum rate and the energy efficiency by jointly calculating the beamforming vector of a ground base station and several relays with several vessels [2].

Sum rate optimization is a non-convex problem. The generalized power iteration [3] method offers practicality, lower complexity, and guaranteed performance compared to conventional methods for optimization problems. This paper applies the generalized power iteration method to the maritime relay communications system to determine the optimal relay beamforming vector.

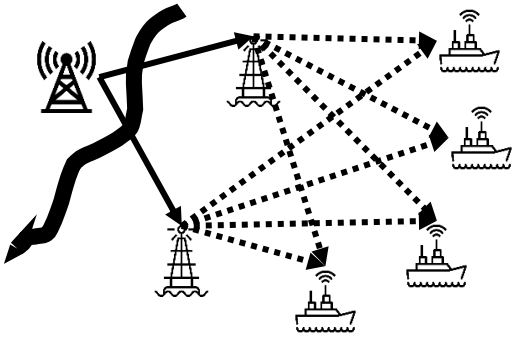


Fig. 1. Maritime multicast scenario with one ground station, multiple relay buoys, and vessel users.

2. System Model

In this paper, As shown in **Fig. 1**, we consider a system model consisting of a single ground base station (BS) with N antennas on the coast, R Amplify-and-Forward (AF) relay buoys (RN), each with M antennas on the sea, and K vessel users (UE) with a single antenna on the sea far from the coast. We assume the BS perfect channel state information (CSI) for simplicity.

This research was a part of the project titled 'The advancement of smart aids to navigation facilities(20210636)', funded by the Ministry of Oceans and Fisheries, Korea.

2.1 Maritime Channel Model

In our maritime relay model, since the near-coast communication channel is dominated by the Line-of-Sight (LoS) path from the BS to the RN with low sea-level reflection, we adopt the empirical loss Rician fading model [2,4] to account for the first hop channel. The first hop channel from the BS to the r-th RN is defined as

$$\left[\mathbf{H}_r\right]_{m,n} = h_{r,m,n}, \quad (1)$$

where $\mathbf{H}_r \in \mathbb{C}^{M\times N}$, r is the RN index, n is the antenna index at the BS, m is the antenna index at the r-th RN, and $h_{r,m,n}$ is the Rician distributed. The second hop, far from the coast, uses a modified two-ray reflection model subject to Rayleigh fading [2,5]. Then, the channel from the r-th RN to the k-th UE is defined as

$$\mathbf{g}_{k,r} \in \mathbb{C}^{M\times 1} = \left[g_{1,k,r},\ldots,g_{m,k,r},\ldots,g_{M,k,r}\right]^T, \quad (2)$$

where k is the UE index, and $g_{m,k,r}$ is the coefficient between the m-th antenna at the r-th RN and the k-th UE, which is the Rayleigh distributed. The equivalent channel from the BS to the k-th UE via the r-th RN is

$$\tilde{\mathbf{h}}_{k,r} \in \mathbb{C}^{N\times 1} = \left[\mathbf{g}_{k,r}^H \mathbf{H}_r\right]^T. \quad (3)$$

2.2 The signal Model

The received signal at the k-th UE is

$$y_k = \sum_{r=1}^{R}\left(\mathbf{g}_{k,r}^H \mathbf{H}_r\left(\sum_{k=1}^{K}\mathbf{f}_k s_k\right)+\mathbf{g}_{k,r}^H \mathbf{n}_r\right)+n_k, \quad (4)$$

where s_k is the message for user k, $\mathbf{f}_k \in \mathbb{C}^{N\times 1}$ is the precoding vectors for maximization sum rate at the BS, $\mathbf{n}_r \in \mathbb{C}^{M\times 1} \sim CN\left(\mathbf{0},\sigma_r^2\mathbf{1}\right)$ is additive white Gaussian noise (AWGN) at each RN, and $n_k \sim CN\left(0,\sigma_k^2\right)$ is AWGN at each UE.

Hence, the received SINR of the k-th UE is

$$\mathrm{SINR}_k = \frac{\sum_{r=1}^{R}\left|\tilde{\mathbf{h}}_{k,r}^H \mathbf{f}_k\right|^2}{\sum_{j\neq k}^{K}\sum_{r=1}^{R}\left|\tilde{\mathbf{h}}_{k,r}^H \mathbf{f}_j\right|^2 + \sum_{r=1}^{R}\sigma_r^2\left\|\mathbf{g}_{k,r}^H\right\|^2 + \sigma_k^2}. \quad (5)$$

The information theoretical achievable spectral efficiency for UE k is $R_k = \log_2\left(1+\mathrm{SINR}_k\right)$.

We assume $\sigma^2 = \sigma_r^2 = \sigma_k^2$ for simplicity.

3. Problem Formulation

Maximizing the sum spectral efficiency that the optimization problem is formulated as

$$\max_{\mathbf{f}_1,\ldots,\mathbf{f}_K}\sum_{k=1}^{K}R_k \quad (6)$$

$$\text{subject to } \sum_{k=1}^{K}\left\|\mathbf{f}_k\right\|^2 \le 1. \quad (7)$$

We rewrite the precoding vectors in a higher dimensional vector by stacking each vector as $\bar{\mathbf{f}} \in \mathbb{C}^{NK\times 1} = \left[\mathbf{f}_1^T,\ldots,\mathbf{f}_K^T\right]^T$. With this, we rewrite (6) by Rayleigh quotients representation as [5]

$$\max_{\bar{\mathbf{f}}}\sum_{k=1}^{K}\log_2\left(\frac{\sum_{r=1}^{R}\bar{\mathbf{f}}^H\mathbf{A}_{k,r}\bar{\mathbf{f}}}{\sum_{r=1}^{R}\bar{\mathbf{f}}^H\mathbf{B}_{k,r}\bar{\mathbf{f}}}\right), \quad (8)$$

where

$$\mathbf{A}_{k,r} = \mathrm{blkdiag}\left(\tilde{\mathbf{h}}_{1,r}\tilde{\mathbf{h}}_{1,r}^H,\ldots,\tilde{\mathbf{h}}_{k,r}\tilde{\mathbf{h}}_{k,r}^H,\ldots,\tilde{\mathbf{h}}_{K,r}\tilde{\mathbf{h}}_{K,r}^H\right) + \frac{\sigma^2\left(1+\left\|\mathbf{g}_{k,r}^H\right\|^2\right)}{P}\mathbf{I} \in \mathbb{C}^{NK\times NK} \quad (9)$$

and

$$\mathbf{B}_{k,r} = \mathbf{A}_{k,r} - \mathrm{blkdiag}\left(\mathbf{0},\ldots,\underbrace{\tilde{\mathbf{h}}_{k,r}\tilde{\mathbf{h}}_{k,r}^H}_{\text{the } k\text{ th block}},\ldots\mathbf{0}\right). \quad (10)$$

By Rayleigh quotients property, we rewrite (8) as

$$\max_{\bar{\mathbf{f}}}\sum_{k=1}^{K}\log_2\left(\frac{\bar{\mathbf{f}}^H\mathbf{A}_k\bar{\mathbf{f}}}{\bar{\mathbf{f}}^H\mathbf{B}_k\bar{\mathbf{f}}}\right), \quad (11)$$

where $\mathbf{A}_k = \sum_{r=1}^{R}\mathbf{A}_{k,r}$ and $\mathbf{B}_k = \sum_{r=1}^{R}\mathbf{B}_{k,r}$.

4. Generalized Power Iteration

For the solution of the transformed problem (11), we derive a first-order optimality condition of (11) that is satisfied if the following holds [5]:

$$\mathbf{B}_{\mathrm{KKT}}^{-1}\left(\bar{\mathbf{f}}\right)\mathbf{A}_{\mathrm{KKT}}\left(\bar{\mathbf{f}}\right)\bar{\mathbf{f}} = \lambda\left(\bar{\mathbf{f}}\right)\bar{\mathbf{f}}, \quad (12)$$

where

$$\mathbf{A}_{\mathrm{KKT}}\left(\bar{\mathbf{f}}_{(t)}\right) = \lambda_{num}\left(\bar{\mathbf{f}}\right)\times\sum_{k=1}^{K}\frac{\mathbf{A}_k}{\bar{\mathbf{f}}^H\mathbf{A}_k\bar{\mathbf{f}}}, \quad (13)$$

$$\mathbf{B}_{\mathrm{KKT}}\left(\bar{\mathbf{f}}_{(t)}\right) = \lambda_{denum}\left(\bar{\mathbf{f}}\right)\times\sum_{k=1}^{K}\frac{\mathbf{B}_k}{\bar{\mathbf{f}}^H\mathbf{B}_k\bar{\mathbf{f}}}, \quad (14)$$

and

$$\lambda\left(\bar{\mathbf{f}}\right) = \prod_{k=1}^{K}\left(\frac{\bar{\mathbf{f}}^H\mathbf{A}_k\bar{\mathbf{f}}}{\bar{\mathbf{f}}^H\mathbf{B}_k\bar{\mathbf{f}}}\right)\frac{\lambda_{num}\left(\bar{\mathbf{f}}\right)}{\lambda_{denum}\left(\bar{\mathbf{f}}\right)}. \quad (15)$$

We summarize this process in **Table 1**.

Table 1. Generalized Power Iteration Algorithm

Algorithm 1 Generalized Power Iteration

Initialize: $\bar{\mathbf{f}}_{(0)} = \text{MRT}$

- Set the iteration count $t = 1$.

While $\left\|\bar{\mathbf{f}}_{(t)} - \bar{\mathbf{f}}_{(t-1)}\right\| > \varepsilon$ **do**

- Construct the matrices $\mathbf{A}_{\text{KKT}}\left(\bar{\mathbf{f}}_{(t-1)}\right)$ and $\mathbf{B}_{\text{KKT}}\left(\bar{\mathbf{f}}_{(t-1)}\right)$ by using (12) and (13).
- Update $\bar{\mathbf{f}}_{(t)} = \dfrac{\mathbf{B}_{\text{KKT}}^{-1}\left(\bar{\mathbf{f}}_{(t-1)}\right)\mathbf{A}_{\text{KKT}}\left(\bar{\mathbf{f}}_{(t-1)}\right)\bar{\mathbf{f}}_{(t-1)}}{\left\|\mathbf{B}_{\text{KKT}}^{-1}\left(\bar{\mathbf{f}}_{(t-1)}\right)\mathbf{A}_{\text{KKT}}\left(\bar{\mathbf{f}}_{(t-1)}\right)\bar{\mathbf{f}}_{(t-1)}\right\|}$ (16)
- $t \leftarrow t+1$

End while

5. Simulation Results

We compare the results with the following methods below:

- GPI: algorithm by table 1.
- RZF: regularizing zero-forcing (ZF)

$$\mathbf{f}_k = \left(\sum_{r=1}^{R}\left(\tilde{\mathbf{h}}_{k,r}\tilde{\mathbf{h}}_{k,r}^H + \beta\mathbf{I}_N\right)\right)^{-1}\sum_{r=1}^{R}\tilde{\mathbf{h}}_{k,r}, \quad (17)$$

where $\beta = \dfrac{\sigma^2\left(1+\left\|\mathbf{g}_{k,r}^H\right\|^2\right)}{P}$.

- MRT: maximum ratio transmission (MRT),

$$\mathbf{f}_k = \frac{\sum_{r=1}^{R}\tilde{\mathbf{h}}_{k,r}}{\left\|\sum_{r=1}^{R}\tilde{\mathbf{h}}_{k,r}\right\|}. \quad (18)$$

Fig. 2 shows that the MRT scheme has higher sum spectral efficiency than the RZF in a low SNR region, and the RZF scheme has higher sum spectral efficiency than the MRT scheme in a high SNR region.

Nevertheless, the performance of GPI is the highest in all SNR ranges. This is because it is the result of finding and optimizing the precoding vector that maximizes the sum spectral efficiency, so the optimal precoding solution was found by considering the noise signals coming into the relay buoys and the AWGN noise of each user. We can see from the simulation result that it has a higher sum spectral efficiency in the entire SNR range than any other method.

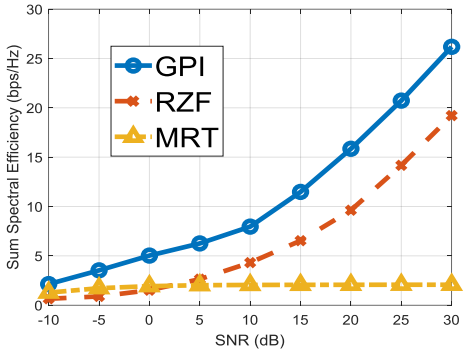


Fig. 2. The sum spectral efficiency comparison per SNR. When $N = 32$, $R = 2$, $M = 4$, and $K = 4$.

6. Conclusion

In this paper, it is shown that the sum spectral efficiency increases with the generalized power iteration method. The optimization for energy efficiency with the generalized power iteration method would also be promising future work.

References

[1] H.-J. Kim, S. V. Tiwari, and Y.-H. Chung, "Multi-hop relay-based maritime visible light communication," *Chinese Optics Letters*, Vol. 14, No. 5, pp. 050 607(1)–050 607(5), 2016.

[2] R. Duan, J. Wang, H. Zhang, Y. Ren, and L. Hanzo, "Joint multicast beamforming and relay design for maritime communication systems," *IEEE Transactions on Green Communications and Networking*, Vol. 4, No. 1, pp. 139–151, 2020.

[3] J. Park, J. Choi, N. Lee, W. Shin, and H. V. Poor, "Rate-splitting multiple access for downlink MIMO: A generalized power iteration approach," *IEEE Transactions on Wireless Communications*, Vol. 22, No. 3, pp. 1588–1603, 2023.

[4] J. Wang, H. Zhou, Y. Li, Q. Sun, Y. Wu, S. Jin, T. Q. S. Quek, and C. Xu, "Wireless channel models for maritime communications," *IEEE Access*, vol. 6, pp. 68 070–68 088, 2018.

[5] P. S. Balkees, K. Sasidhar, and S. Rao, "A survey based analysis of propagation models over the sea," *in Proc. of IEEE Int. Conf. Adv. Comput. Commun. Informat. (ICACCI)*, 2015.

Copyright ⓒ 2023 KSII

Design and implementation of predictive system with AI/ML dataset management supported IoT platform

Dae-Geun Heo and Seung-Myeong Jeong*
Autonomous IoT Research Center, Korea Electronics Technology Institute, Seongnam, South Korea
[e-mail: hdkyon96@keti.re.kr, sm.jeong@keti.re.kr]
*Corresponding author: SeungMyeong Jeong

Abstract

When considering AI supported parking services to alleviate the parking space shortage problem, it requires a dataset created by merging collected data for each parking lot to utilize AI/ML models. However, creating a process to merge data requires additional time for developers. In order to mitigate this shortcoming, this paper proposes a parking prediction system applying the oneM2M standard IoT platform with a dataset management interface. The proposed system uses the interface extending the oneM2M IoT platform to create a dataset from several types of data collected to predict parking congestion level, and provides datasets for AI/ML training and inference. As the result, for the proposed AI-based predictive system, the dataset preparation by application developers can be easily done by the oneM2M platform with the extended dataset management capability.

Keywords: Internet of Things, Artificial Intelligence, oneM2M

1. Introduction

To alleviate the parking space shortage problem, various measures such as the shared parking control system based on the IoT platform and the parking occupancy prediction service are being proposed [1][2]. When building an IoT platform-based parking prediction service, creating a dataset with data of each parking lot and using an AI/ML model based on it can improve the prediction accuracy [3]. However, for AI/ML model developers, it takes a lot of time to create datasets for the model and create dataset input and output scenarios.

To easily create datasets per parking lot to build AI/ML models by the deloper, this paper proposes and describes a parking congestion prediction system based on the oneM2M standard IoT platform. The platform extended an interface to integrate data within the oneM2M platform using the <datasetPolicy> resource according to the specified dataset generation [4]. Through this, it aims to minimize the time to prepare datasets when building a parking service using AI/ML models.

2. Proposed system

2.1 Parking Congestion Prediction Model

The parking congestion prediction model to be used in the proposed system is based on the Long Short Term Memory (LSTM) model to derive predictions by inputting data with time series characteristics [5].

Table 1 lists the data that be used to predict parking congestion. Temperature, precipitation, fine dust data from open APIs and parking congestion level are stored as individual oneM2M resources on the oneM2M platform. The congestion level is processed with parking spot status that are provided from image recognition(i.e. YoLo5)-based virtual parking

This work was supported by the Technology Innovation Program (20016243) funded By the Ministry of Trade, Industry & Energy (MOTIE, Korea)

sensors and radar sensors [6].

Table. 1. data features for prediction model

Features	Description	Periodicity	oneM2M Resource
congestion	Actual parking congestion level (occupied spot / total parking spot)	every hour	"congestion"
temp	Temperature	every hour	"weather"
precipitation	Precipitation status ("0"=NoRain; "1"=Rain; "2"=Snow)		
finedust	Level of Air Quality ("0"=Good; "1"=Bad; "2"=Very Bad)		
temp forecast	Temperature forecast	every dat at 0, 3, 6, 9, 12, 15, 18, 21 o'clock	"weather forecast"
precipitation forecast	Precipitation status forecast ("0"=NoRain; "1"=Rain; "2"=Snow)		
finedust forecast	Level of Air Quality forecast ("0"=good; "1"=bad; "2"=very bad)	every day at 6, 18 o'clock	

Each oneM2M resource also contains "time_stamp" property so all features from different resources can be merged with the same timestamp. **Table 2** shows an example of the created dataset which is used to train an LSTM model.

Table. 2. Example of the training dataset

time_stamp	congestion	precipitation	temp	finedust
"2023-05-05 09:00:00"	0.473	"1"	14	"0"
"2023-05-05 10:00:00"	0.759	"0"	16	"0"
"2023-05-05 11:00:00"	0.981	"0"	19	"0"

2.2 Predictive System without Dataset Management Interface

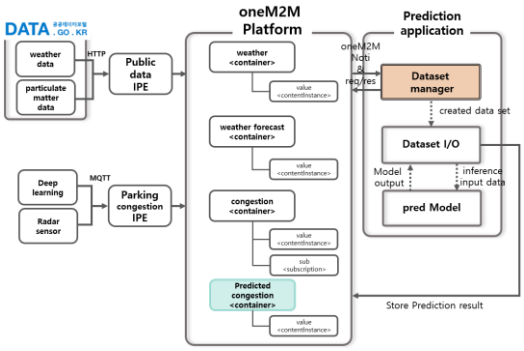


Fig. 1. System architecture of the predictive system without a dataset management interface

The components of the prediction system in **Fig. 1** are as follows.

- Interworking Proxy Entity(IPE) that performs the conversion of non oneM2M resource data to ingest as oneM2M resource
- oneM2M platform
- prediction application that aggregates data from the platform into datasets like shown in **Table 2** and operate prediction model based on them.

The prediction application subscribes to the child resource creation event of the "congestion" resource. When the application receives a notification message from the platform upon "congestion" resource creation [7], It requests the temperature, precipitation, fine dust data loaded on the platform and combines the data received in responses into a dataset. After this process, the predictive congestion level is generated from the model inference and stored on the platform by the prediction application.

2.3 Predictive System with Dataset Management Interface

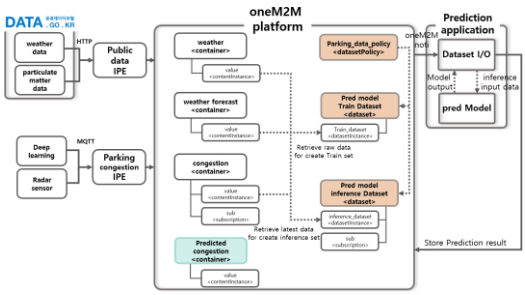


Fig. 2. System architecture of the predictive system with a dataset management interface

Unlike the system in **Fig. 1**, the proposed platform in **Fig. 2** has an extended dataset management interface on the oneM2M IoT platform. Predictive application development gets simpler because the application does not implement a dataset management functionality, but leveraging the dataset inteface provided by the platform.

The system follows the scenario shown in **Fig. 3**. The prediction application subscribes to the child resource creation event of the "inference_dataset" resource. When there is a new resource creation among original resources (e.g. "congestion") of the dataset, the platform create a new inference data resource and it gets sent as a notification to the application. The new inference input data is served to the model and the parking congestion level is stored on the oneM2M platform.

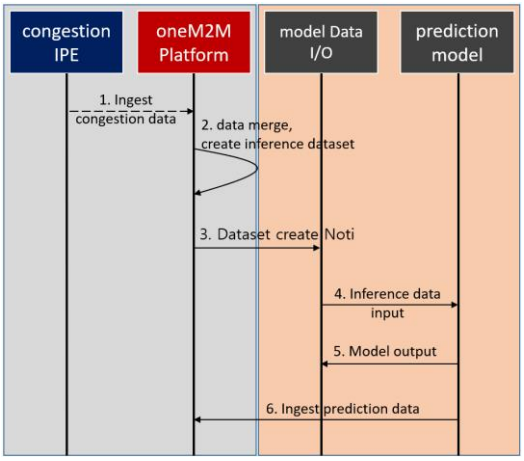


Fig. 3. Scenario with the dataset management interface

As a result of making dataset management a common feature of IoT platforms, The application can eliminate the need for applications to request data from the platform and assemble the received data into datasets. This reduces the application development process.

3. Conclusions

This paper proposes the oneM2M IoT platform-based parking congestion prediction system that has been extended through the training and inference dataset management interface of AI/ML models. Since the proposed system enables automation of training and inference dataset generation of AI/ML models, it can be easier to develop AI/ML applications when building an IoT platform-based predictive services using individual parking congestion prediction models.

For the future work, we would like to enhance the interface to support complex data having objects and arrays, while current work support simple data types. Then it can be used in more general AI/ML services which are still oneM2M compatible.

References

[1] M. Nam and H. Lee, "Design of an Open Platform-based Smart Parking System," *Journal of the Korean Institute of Electrical Engineers*, Vol 56, No. 11, pp. 1105-1110, 2019.

[2] K. Kang, Y. Koh, D. Kim and H. Lee, "A Study on the Utilization of Parking Sharing Control System for the Railway Service," *The Korean Society for Railway*, pp. 89-101, 2019.

[3] S. Inam, A. Mahmood, S. Khatoon, M. Alshamari and N. Nawaz, "Multisource Data Integration and Comparative Analysis of Machine Learning Models for On-Street Parking Prediction," *Sustainability*, Vol 14, No. 12, pp. 272-287, 2022.

[4] S. Jeong and C. Choe, "Machine Learning Dataset Management Interface for AIoT Platforms Supporting AI/ML Services," *Journal of the Korean Institute of Communications and Information Sciences*, pp. 421-422, 2022.

[5] W. Shao, Y. Zhang, B. Guo et al., "Parking Availability Prediction with Long Short Term Memory Model" *Green, Pervasive, and Cloud Computing: 13th International Conference*, Vol 13, pp. 124-137, 2019.

[6] H. Son, J. Hwang, I. Lee and J. Song, "A Study on the Development of Integrated Control System and the Problem of Parking Space Using Deep Learning in the Internet of Things Environment," *Journal of the Korean Institute of Communications and Information Sciences*, pp. 939-940, 2020.

[7] oneM2M, "TS-0001, Functional Architecture TS v4.0.0".

Copyright ⓒ 2023 KSII

A study on the risk management plan of domestic coastal fishing vessels using big data

Hyunjeong Lim[1] and Sooyeon Kwon[2*]
[1]Department of Fishing Vessel Safety Research, Korea Maritime Transportation Safety Authority(KOMSA),
[2]Department of Naval Architecture & Ocean Engineering, Inha University
[e-mail: hjlim02@komsa.or.kr, kissiny@komsa.or.kr]
* Corresponding author: Sooyeon Kwon

Abstract

Fishing vessels account for more than 88% of domestic vessels, and more than 90% of them are domestic coastal fishing vessels. Domestic coastal fishing vessels are small fishing vessels of less than 10 tons, and safety is very important because they are used by fishermen engaged in livelihood. However, about 2,000 marine accidents occur every year. Therefore, in order to prepare measures to prevent accidents of domestic coastal fishing vessels, it is necessary to analyze big data on the current status of accidents by sea area, type of fishery, and ton class in Korea and derive the cause of the accident. In addition, by analyzing the safety management policy and accident status of Norway, an advanced European maritime country, we intend to prepare applicable plans for reducing accidents of domestic coastal fishing vessels.

Keywords: Big data, Safety management, Coastal fishing vessels, Marine accidents

1. Introduction

A coastal fishing vessel is a fishing vessel with a gross tonnage of less than 10 tons according to Article 40 (Permitted Fishery) Paragraph 2 of the Fisheries Industry Act. According to the fishing vessel data registered with the Ministry of Oceans and Fisheries it was investigated that it accounted for about 94% of the total number of fishing vessels. In addition, looking at the status of marine accidents by type of fishery by the Korea Maritime Safety Tribunal, the accident rate of fishing vessels compared to the total number of accidents is about 65% , and among them, the total number of accidents for fishing vessels under 10 tons It was surveyed at about 73.5% of the comparison. The types of marine accidents were classified into 16 categories: collision, grounding, sinking, fire, explosion, casualty accident, contact, capsize, and others. According to the current status of marine accidents by vessel use and accident type (2018-2022) provided by the Korea Maritime Safety Tribunal, the highest types of marine accidents for fishing vessels were engine damage, collision, flotation, and safety accidents, in order. And since about 2,000 accidents occur every year, it is necessary to analyze accidents and search for ways to reduce accidents.

Therefore, in this study, a data form containing information on the accident fishing vessel was created through the ruling of the Korea Maritime Safety Tribunal for the last five years (2018-2022). However, it is not possible to generalize all accidents because it only deals with written decision, not all accident data. In addition, by utilizing the safety management policy and big data of Norway, an advanced maritime country in Europe, we plan to prepare domestically applicable measures for accident reduction.

2. Domestic coastal fishing vessel data research

2.1 subject of investigation

This research was supported by the ‘Development and demonstration of data platform for AI-based safe fishing vessel design (20220210)’ of the Ministry of Oceans and Fisheries, Republic of Korea.

The data of coastal fishing vessels registered for ship inspection by the Korea Maritime Transportation Safety Authority is the subject of investigation.

2.2 analysis method

For analysis by type of fishery, sea area, vessel material, and tonnage class, a total of 3 regions were clustered and then classifying by sea area. In addition, a specific sea area (southwest coast) where fishery management vessels are distributed was also analyzed.

2.3 analysis result

The current status by sea area, ship age, and ton class is shown in **Fig. 1**.

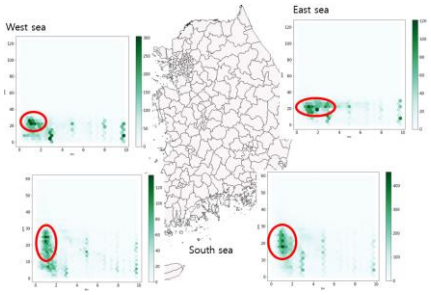


Fig. 1. Distribution of fishing vessels by sea area, age, and ton class

Through **Fig. 1**, by sea area, fishing vessels older than 20 years were the most, and the ton class was less than 2 tons, and more than 2 tons and less than 5 tons. It is shown in **Table 1** in more detail, including ship material and type of fishery

Table 1. Distribution of major type of fishery by sea region

sea area	Ton class	age of the ship	type of ship	ship material
West sea	less than 2 tons	more than 20 years	coastal gillnet, coastal complex	FRP
East sea	More than 2 tons and less than 5 tons	more than 20 years	coastal gillnet	FRP
South sea	less than 2 tons	more than 20 years	coastal complex	FRP
South sea (fishing ground management vessel)	less than 2 tons	more than 20 years	fishing ground management vessel	FRP

3. Domestic coastal fishing vessel accident status data investigation

3.1 subject of investigation

Maritime accidents were analyzed for the major type of fishery of each of the three sea areas obtained from the previous analysis, and the ruling of the Korea Maritime Safety Tribunal was used as data.

3.2. analysis method

The causes by type of accidents of major coastal fishing vessels by sea area were classified into 4M (Man, Machine, Media, Management) and analyzed. Based on the analysis results, a plan to reduce marine accidents was derived.

3.3 analysis result

For coastal fishing vessel accidents, the distribution of fishing vessel accidents in the order of collision (about 64%), casualty accident (about 14.8%), and stranding (about 8.8%) compared to the total number of vessels (453ships). Previously, the accident types for the analysis results of 2.3 can be shown in Table 2.

Table 2. Types of accidents by major type of fishery by sea area

Sea area	type of fishery	Types of Marine Accidents
West sea	Coastal Complex, Coastal Gillnet	Collision, casualty accident, stranding
East sea	Coastal Gillnet	Collision, casualty accident, capsize
South sea	Coastal Complex	Collision, stranding, casualty accident
	fishery management vessel	Collision, fire, casualty accident

Table 3 shows the common types of accidents by sea area, such as collisions and casualty accident. The causes of marine accidents were divided into man, machine, media, management, and 4M [1], and detailed causes were also described.

Table 3. major accident causes

accident type	Main cause	detailed cause
collision	Guard negligence	navigation by sight(C)
		Poor performance of the radar(A)
		Neglect of vigilance during fising(C)

	Insufficient avoidance cooperation	Lack of warning signs(C)
	Improper navigation	Safety distance not maintained(C)
		No voyage plan established(D)
casualty accident	Lack of captain's safety management	Inappropriate work order(D)
		simple caution(D)
	Sailor's carelessness	dangerous work position(B)
		Inappropriate use of fishing equipment(C)

4M : A(Machine), B(Media), C(Man), D(Management))

Since most of them are caused by human factors, it is necessary to prepare measures to prevent marine accidents by reducing the risk factors of the above factors.

4. Norway Fishing Vessels Risk Management Policy and Accident Investigation

4.1 accident status

Looking at the accident data of Norwegian fishing boats over the past 10 years, the most frequent marine accidents occurred in the order of Occupational accident accident, groundings, fires/explosions, and collisions. 3.3 Analysis results (Korea) and Norway's major marine accident types are similar. Therefore, it is necessary to find a way to reduce Korea's fishing vessel maritime accidents by utilizing Norway's fishing vessel risk management policy.

4.2 Policy

Currently, Norway provides a risk assessment checklist template for each fishery so that fishermen can independently evaluate the risk level and establish an action plan for safe navigation in a systematic way

4.3 Derivation of applicable domestic policies

In Korea, there is no system to check the self-management of fishermen to reduce accidents, and safety education and publicity for fishermen are being conducted, but they are still at an insufficient level. Therefore, Korea needs to establish a system that can prevent accidents through self-diagnosis of fishermen upon departure and arrival through a service similar to Norway. Therefore, Korea needs to establish a system that can prevent accidents through self-diagnosis of fishermen upon departure and arrival through a service similar to Norway. Therefore, based on the analysis of accident data, if the system with the process shown in Fig.4 is provided so that fishermen can easily use it, and the data is analyzed and used for standard revision, it is thought that it can contribute to reducing fishing vessle accidents.

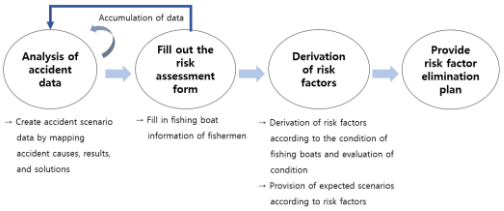


Fig. 4. Risk assessment process based on accident data analysis

5. Conclusions

This study examined the current status of domestic coastal fishing vessels, collected accident-related data based on the ruling of the Korea Maritime Safety Tribunal, and analyzed the types of accidents and causes of accidents by 4M method. As a result, it was confirmed that most of the accidents were caused by humans . Therefore, it is necessary to provide templates suitable for the characteristics of each type of fishery to check the status of their ships so that fishermen can manage risks on their own so that they can improve their safety awareness. In addition, it is thought that domestic coastal fishing vessel accidents can be reduced if the prepared template contents are dataized and analyzed to provide policies for accident prevention.

References

[1] W.-S. KIM, Y.-K. HYUN, and Y.-W. LEE, "*Risk factors of fisher on stow net fishing vessel using analysis of adjudication*," Journal of the Korean Society of Fisheries Technology, vol. 56, no. 2. The Korean Society of Fisheries and Ocean Technology, pp. 155–162, 31-May-2020.

[2] Y.-W. LEE, S.-H. KIM, and K.-J. RYU, "*Analysis of occupational accidents for fisher's on gillnet fishing vessel using the written verdict*," Journal of the Korean Society of Fisheries Technology, vol. 58, no. 4. The Korean Society of Fisheries and Ocean Technology, pp. 367–373, 30-Nov-2022.

Copyright ⓒ 2023 KSII

Data system for Fishing Vessel Design

Dongheum Ryu[1], Yongbin Lee[1], Changhyun Park[1,*] and Sooyeon Kwon[2,3]
[1] PIDOTECH Inc., Seoul, South Korea
[2] Korea Maritime Transportation Safety Authority, Sejong, South Korea
[3] Inha University, Incheon, South Korea
[e-mail: lyusky@pidotech.com, eyongbin@pidotech.com, chpark@pidotech.com, kissiny@komsa.or.kr]
*Corresponding author: Changhyun Park

Abstract

Data system capable of efficiently designing fishing vessel has been developed. A large amount of engineering data was generated for AI Model learning, and a dedicated preprocessor for extracting features was developed. After obtaining the performance data of the fishing vessel according to the module combination of fishing vessel, AI algorithms autonomously generate metamodels of fishing vessel performance. The generated metamodels are used to predict the performance of the fishing vessel and to determine the optimal module combination. The design system for fishing vessel design will be mounted on the web platform and provided to various users.

Keywords: AI, data system, fishing vessel, metamodel, optimization

1. Introduction

Fishermen need a safe fishing vessel that is fast and can load many fishing equipment. In recent years, fishermen have tended to increase the length of fishing vessels to increase the speed of fishing vessels and lower the depth of fishing vessels to maintain the licensed tonnage of fishing vessels. In the end, the freeboard is lowered, which inevitably increases the risk of accidents, and there is a problem that illegal additives that have not been considered for stability are attached.

To design a fishing vessel with enhanced performance and safety, various considerations for hull, superstructure, engine, propulsor, equipment, etc. and the requirements of fishermen should be reflected. In addition, we have developed an AI algorithm that can autonomously generate metamodels of fishing vessel performance so that ship designers and shipbuilders can design fishing vessel easily and quickly. These AI algorithms are mounted on data systems along with optimization algorithms, which are used to quickly predict the performance of the combined modules and determine the optimal module combination that can satisfy all design requirements.

2. Standardized Modules and Sampling Techniques

2.1 Standardized modules

Various requirements for fishing vessels were collected through a living lab conducted for Fishermen, ship designers, and shipbuilders, and the scope of fishing vessel design was determined by analyzing KOMSA's fishing vessel statistics data. Modules were optimized to reflect the design requirements by using design optimization. In addition, the self-propulsion, stability, structural, and safety performance of the fishing vessel were evaluated by simulation [1], and the simulation performance was verified through towing tank experiments[2]. Through this process, standardized modules for hull, superstructure, engine, propulsor, and equipment that can guarantee performance and safety were developed.

This research was supported by the 'Development and demonstration of data platform for AI-based safe fishing vessel design (20220210)' of the Ministry of Oceans and Fisheries, Republic of Korea

2.2 Sampling methodology

The level value of each standardized module is expressed in the form of a categorical rather than a numerical value. Accordingly, a new sampling technique was developed to combine categorical variables. For efficient design, standardized modules were combined for each sampling point, and the performance of fishing vessels combined with standardized modules was evaluated through simulation and experiments.

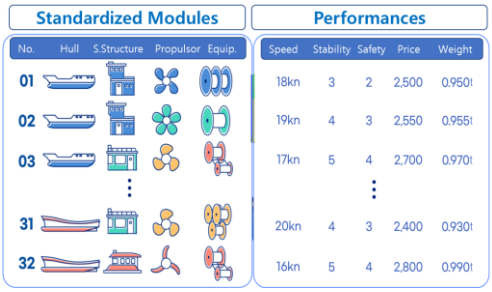


Fig. 1. Standardized modules and performances of fishing vessel

3. AI Algorithm for Autonomous Generation of Metamodel

Simulations of the same size as fishing vessels are inevitably time-consuming. On the other hand, the metamodel can predict the performance of fishing vessels in real time and is very useful for design optimization that requires repetitive calculations[3]. However, determining the metamodel type and hyperparameters that can accurately predict fishing vessel performance values requires expert-level technical knowledge.

In this study, we selected more than 200 engineering research benchmark problems to generate 300,000 engineering data, and developed a preprocessor for automatically extracting characteristic values from the engineering data. The data represented by characteristic values are used for AI model learning that can automatically determine metamodel types. After learning, the AI model is mounted on the AI algorithm and used to autonomously generate metamodels for each fishing vessel performance, and the standardized module combination and performance values for each combination are used as input data.

4. Data System for Optimal Fishing Vessel Design

Metamodels are used to predict the performance of a fishing vessel combined with modules. In addition, AI algorithm is used in conjunction with optimization algorithm to determine the optimal module combination that can satisfy design requirements. The optimization algorithm is a global optimization algorithm that can handle categorical and discrete design variables.

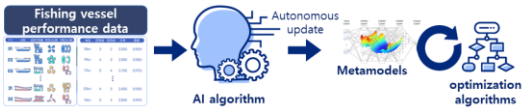


Fig. 2. Data system for fishing vessel design

5. Conclusions

In this study, a data system for fishing vessel design was developed. A sampling technique capable of efficiently combining input variables expressed in a categorical form was developed. Engineering data were created for AI model learning, a dedicated preprocessor was developed for characteristic extraction, and the learned AI model was mounted on the AI algorithm so that the optimal metamodel could be autonomously generated.

The data system will be mounted on a web-based platform after upgrading. This is expected to provide effective services for fishing vessel design and safety to fishermen, ship designers, shipbuilders, and system managers.

References

[1] J. Yu, M. Lee, Y. Kim, S. Suh, and I. Lee, "An optimization study on the hull form and stern appendage for improving resistance performance of a coastal fishing vessel," *Appl. Sci*, 2021.

[2] M. Lee, S. Kim, J. Yu, and Y. Lee, "An experimental study on the optimization of stern appendix for new generation korean fishing vessels," *Journal of the Society of Naval Architects of Korea*, Vol. 58, No. 1, pp. 32-39, 2021.

[3] PIDOTECH Inc., *PIAnO Signature User's Manuals and Tutorials*; PIDOTECH Inc.: Seoul, Republic of Korea, 2023.

Copyright ⓒ 2023 KSII

Aids to Navigation (AtoN) Trend Analysis using LDA Topic Modeling

Narae Yu and Jooyoung Lee*
Korea Science Policy Platform co-op
Seoul, South Korea
[e-mail: nryou@kspp.re.kr, jylee@kspp.re.kr]
*Corresponding author: Jooyoung Lee

Abstract

As autonomous vessels and smart ports in the field of maritime transport rapidly develop with advances in ICT technology, there is also a need to promote the digital transformation of "Aids to Navigation (AtoN)". Previous research has been focused on technology transition towards the future, but there is lack of research on AtoN trends. This study aims to identify trends, key technologies, and future directions in response to the current trend of technology transition. To achieve this, we conducted a keywords and LDA topic analysis based on 14,004 data collected from BigKinds news database, between January 1, 1990 and March 31, 2023, to summarize the trends of changes in AtoN and present future prospects, along with keywords.

Keywords: Aids to Navigation, Lighthouse, LDA, Trend Analysis, Big Data

1. Introduction

AtoN refers to facilities that use light, color, sound, radio, waves, and others to inform vessels of their location, direction, and the position of obstacles to assist in safe navigation [1]. AtoN needs to respond to the changing future environment of maritime transportation, including the development of autonomous vessel and smart port, in line with the era of ICT technology development and transformation. However, there was a lack of trend research on major issues, core technologies, and policies according to the trend of the times compared to the long history of AtoN. In this study, news data on AtoN from January 1, 1990, and March 31, 2023, were collected to extract keywords and select main topics through LDA analysis. LDA analysis is commonly used in research to extract new information from unstructured text documents and is proposed as an effective means of understanding topics and trends in specific fields.

2. LDA Topic Modeling

Topic modeling is an indispensable algorithm utilized in text mining technique which for finding a topic by discovering a certain pattern in a large number of unstructured documents [2]. The topic is inferred by clustering words with similar meanings using context-related clues [3]. One of the topic modeling, LDA is a topic modeling algorithm that estimates two probability distributions, assuming that documents consist of the probability distributions of topics and the probability distributions of words within the topic. LDA is a popular text mining technique that identifies topics in a document collection by estimating a Dirichlet distribution in large amounts of document data. Due to its effectiveness, LDA has been applied in various fields. In this study, we would like to extract topics through news articles on AtoN using LDA topic modeling and present how these topics change according to trends and times.

This research was supported by the project No. 20210636 “The advancement of smart aids to navigation facilities,” funded by the Ministry of Oceans and Fisheries, Republic of Korea.

3. AtoN Trend Analysis using LDA

3.1 Data and Methods

This study was a total of 14,004 AtoN-related articles published in Korean online news articles from January 1, 1990, to March 31, 2023 based on 14,004 data collected from BigKinds news database. The criteria for selecting AtoN-related articles included words of maritime or ocean but excluded duplication and photo/video articles. In addition, this study excluded sports, cooking, fishing, and publishing news unrelated to the study.

The collected data were analyzed in the order of (1) keyword extraction and preprocessing process, (2) keyword analysis, and (3) topic modeling.

In this study, keywords in the form of nouns were extracted to extract morphemes, the minimum semantic unit, from news articles, and synonyms and exclusions were selected based on the extracted keywords and went through a keyword preprocessing process. For example, spacing words, region name, institutional name, etc. were named as one word to refine the unification of words, and words that did not have much meaning, such as 'maximum', 'degree', 'as much', 'while', and 'today', were excluded as stopword.

In order to understand the characteristics of the extracted keywords, TF(term frequency), which is the frequency of keywords appearing throughout the news article, was obtained.

However, excluding keywords with very high TF values, such as search terms, TF-IDF (term frequency-inverse document frequency) was analyzed, and in this study was excluded the highest words based on TF-IDF 0.02 which are 'lighthouse' and 'sea'.

TF-IDF is a method of identifying keywords with important meanings within a document by multiplying the frequency of a specific word appearing in the entire document by the reciprocal of the value indicating how many documents a specific word appears in [4].

In addition, the top 50 keywords that appeared in news articles were identified based on DF (document frequency) in which keywords appeared. Since the TF-IDF weight is the value of words by document, the sum of the TF-IDF weights was used in this study to evaluate the importance of keywords in all news articles.

3.2 Results

Following this procedure, **Table 1** shows the results of TF, TF-IDF, and DF analysis by collecting news data. Among the keywords extracted from news articles throughout the entire period, the top keywords based on TF-IDF were derived in the order Marine, Installation, Vessel, Tokdo Island, Aids to Navigation, Busan Safety, Construction, and Facility.

Table 1. TF-IDF, DF, TF results

Rank	Keywords	TF-IDF	DF	TF	Rank	Keywords	TF-IDF	DF	TF
1	해양	0.0213	5408	13665	26	문화	0.0088	2476	5057
2	설치	0.0205	4941	11146	27	마을	0.0083	1930	6215
3	선박	0.0169	3725	8213	28	위치	0.0082	2779	4541
4	독도	0.0160	970	7525	29	경북 포항	0.0081	1221	3216
5	항로표지	0.0159	2081	5056	30	관리	0.0081	2094	4053
6	부산	0.0145	2556	8696	31	이용	0.0080	2913	4645
7	사업	0.0141	2930	8735	32	선정	0.0077	1774	3524
8	안전	0.0135	3162	6270	33	점검	0.0076	952	2247
9	조성	0.0121	3033	7139	34	항로	0.0074	1327	2430
10	시설	0.0119	3293	6101	35	수산	0.0074	1863	3135
11	계획	0.0116	4091	6945	36	해역	0.0073	1465	2710
12	제주도	0.0115	1617	6836	37	관계자	0.0072	2716	3197
13	체험	0.0114	2037	5106	38	해안	0.0072	2224	4843
14	지역	0.0114	4271	8557	39	공간	0.0072	1938	3643
15	해양수산부	0.0108	1703	5118	40	파도	0.0072	1891	4040
16	울산	0.0097	1493	4574	41	일본	0.0072	1472	4626
17	운영	0.0096	3014	5245	42	정보	0.0071	1365	2987
18	해상	0.0095	2412	4159	43	사고	0.0070	1360	2801
19	방파제	0.0094	1736	4036	44	주민	0.0070	1966	4468
20	추진	0.0093	2592	5124	45	등표	0.0069	551	1363
21	인천	0.0093	1517	3875	46	세계	0.0068	2221	4528
22	행사	0.0092	1715	3772	47	항만	0.0068	1232	2635
23	관광	0.0091	2286	5134	48	해수욕장	0.0067	1217	3627
24	관광객	0.0089	2408	4725	49	항해	0.0066	1658	2474
25	전남 여수	0.0088	1003	3788	50	사진	0.0066	2119	3676

When LDA topic modeling, consistency scores are derived for each topic number using the npmi method as shown in **Fig 1.** and are optimally determined when the score is the highest. Furthermore, the nine topics were the most consistent.

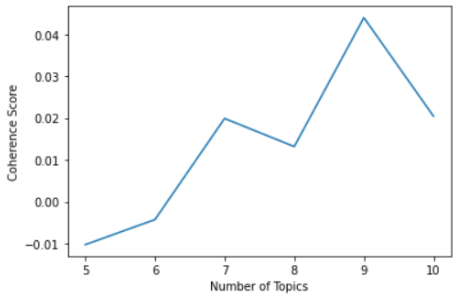


Fig 1. Coherence score by number of topics

Accordingly, LDA was conducted through nine topics in each 10-year cycle from 1990 after extracting the main keywords. As a result, from 1990 to 1999 topics named "Lighthouse and sea route safety," "Innovation and technological development in the marine industry," "Improvement of marine safety operation and information-oriented system," etc. which main

topics are related to AtoN technology and system. Next, from 2000 to 2009, the main topics suggested pollution of the marine environment, positioning technology and utilization, lighthouse cultural heritage, for example, "measures for pollution in the marine environment," "build a marine safety facility," "positioning technology and utilization," "Lighthouse tourism activation relation to coastal scenery and cultural heritage," etc. From 2010 to 2019, topics are presented region-related, such as Busan, Ulsan, and coastal region tourist attractions. Especially, "Busan marine cultural festival," Plans to revitalize marine tourism in Ulsan,", "Sewol ferry disaster and accidents at sea," etc. Lastly, after 2020, the topics are emphasized about government and lighthouse, for example, "Utilizing the lighthouse heritage," "government's role in the promotion of marine policy," "Urban environment improvement project using lighthouse," etc.

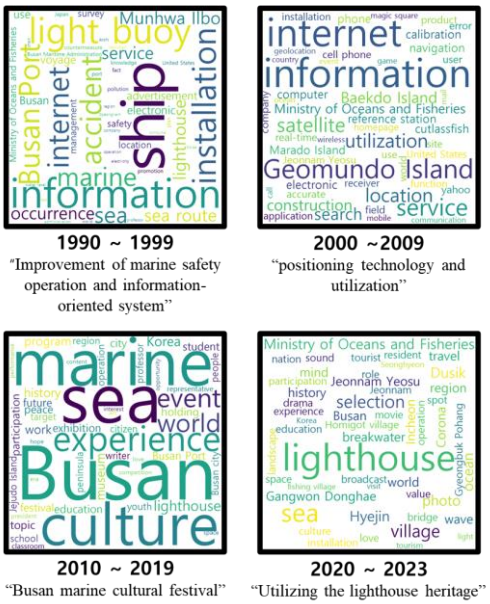


Fig 2. News article keyword Word Cloud by main topics each period

4. Conclusions

This study collected news databases using AtoN as keywords and quantified them using LDA analysis, a text mining technique. As a result of the analysis, installation, vessel, projects, safety, and build made up the main keywords, and marine safety, promotion of projects, and cultural events were the main topics.

In particular, as a result of the rapid between the four generations increases in keywords related to marine safety, the creation of a lighthouse, and the promotion of marine tourism, the need to drive related policies and R&D, such as streng thening safety and creating cultural infrastructure, was confirmed from a long-term perspective.

Based on this study, it is expected that it will be possible to analyze changes in annual issues of AtoN in more detail through AtoN's history and derive key technology keywords.

References

[1] Ministry of Oceans and Fisheries, Aids to Navigation Act, Chapter I General Provisions, Article 2 Definitions, 2022.

[2] D. Blei, "Probabilistic topic models," *Communications of the ACM*, Vol 55, No 4, pp. 77-84, 2012.

[3] C. Calabrese, J. Ding, B. Millam and G. Barnett, "The Uproar Over Gene-Edited Babies: A Semantic Network Analysis of CRISPR on Twitter," *Environmental Communication*. Vol 14, No 7, pp. 954-970, 2020.

[4] SJ. Lee, HJ. Kim, "Keyword extraction from news corpus using modified TF-IDF," *The Journal of Society for e-Business Studies*. Vol 14, No. 4, pp. 59-73.

Copyright ⓒ 2023 KSII

Expanding Communication Coverage in Maritime Communications: Outage Capacity Analysis of two-Hop Relay and Beamforming

Kyeongjea Lee, Tae-Woo Kim, Sung-yoon Cho, Ki-won Kwon and Dong Ku Kim
Department of Electrical engineering, Yonsei University Seoul, South Korea
Smart Network Research Center, Korea Electronics Technology Institute Seoul, South Korea
[e-mail: kjlee92@yonsei.ac.kr, tw.kim@yonsei.ac.kr, sycho@keti.re.kr, kwonkw@keti.re.kr, dkkim@yonsei.ac.kr]
*Corresponding author: Dong Ku Kim

Abstract

Maritime communication networks face numerous challenges, such as long distances, high mobility, and harsh environmental conditions. These factors necessitate the development of efficient techniques to enhance coverage, reliability, and throughput. In this paper, system model is a two-hop downlink scenario involving base station, relay, and destination. We first analyze ocean channel condition, including path loss, ocean wave and buoy movement. Then, we evaluate relay's receive beamforming under various channel conditions in terms of the outage probability and capacity. Via simulations, we confirmed that a wide beamwidth provides better performance in rough ocean conditions, whereas a narrow beamwidth outperforms the wide beam in calm ocean conditions. It is worth noting that appropriate selection of beamwidth in maritime communication systems plays a crucial role in minimizing the outage probability while simultaneously enhancing the capacity for communication.

Keywords: Maritime communication, Relay, Beam forming, outage capacity

1. Introduction

Maritime communication systems play a crucial role in ensuring the safety, security, and efficiency of maritime operations. However, maritime communication networks face unique challenges due to the harsh environmental conditions, long distances, and high mobility of ships [1]. These factors lead to the degradation of communication links and increased outage probability. To address these challenges, various techniques have been proposed in the literature, including relay-assisted communication and adaptive beamforming [2]. In this paper, we investigate the application of relay selection and beamforming techniques to enhance the outage capacity of maritime communication networks in two-hop downlink scenarios. Building upon the recent advancements in relay-assisted communication [3] and adaptive beamforming [4], we propose an integrated system model that combines the benefits of these techniques.

2. Channel model

According to [5,6], the communication channel between a BS and relay is considered line-of-sight (LoS) path. On the other hand, the second hop in the offshore communication channel is considered 2-ray path channel.

This research was a part of the project titled 'The advancement of smart aids to navigation facilities(20210636)', funded by the Ministry of Oceans and Fisheries, Korea, and was supported by the National Research Foundation of Korea (NRF) Grant through the Ministry of Science and ICT (MSIT), Korea Government, under Grant 2022R1A5A1027646

2.1 Direct path channel model

We assume that the base station to the first relay is approximated to the LoS characteristics [5]. The path loss of LoS single path can be represented as follows:

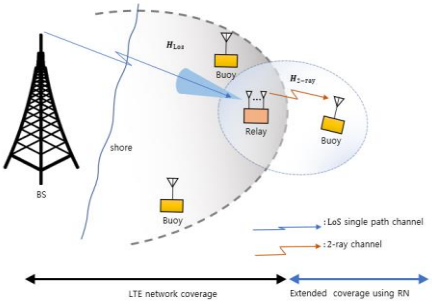


Fig. 1. System architecture in maritime communication

$$PL_L = -27.56 + 20\,log_{10}(d) + log_{10}(f)$$

where d is communication distance and f is frequency.

2.2 two-ray path channel model

In this paper, for simplicity, the roughness of the sea level was regarded as 0, and the reflection coefficient is to be -1[6].The path loss of two-ray path can be represented as follows:

$$PL_T = -20\,log_{10}\left\{\left(\frac{\lambda}{4\,\pi d}\right)\left|1 + R\,\exp\left(\frac{j\alpha\ 2\,\pi\Delta d}{\lambda}\right)\right|\right\}$$

where represents the wavelength, R is the reflection coefficient.

2.3 Buoy angle

The angle θ between the antenna and the sea level is an important parameter. Equation equation is used to approximate the resulting angle θ.

$$\theta = \tan^{-1}\left(\frac{\gamma(x+1) - \gamma(x-1)}{2}\right)$$

where γ represents the surface elevation [7].

3. Beamforming and outage capacity

In this paper, there is no direct (source destination) link, The signal transmitted by the relay in the AF mode is $x_i = \sqrt{K}y_i$, where K is power gain. Thus, the input-output relationship of the whole 2-hop AF relay channel is [8]

$$\boldsymbol{Y} = \sqrt{K}\boldsymbol{H_T F H_L^{\dagger} x} + \sqrt{K}\boldsymbol{H_T \xi_1} + \xi_2$$

where x and Y are the transmitted (BS) and received (Buoy) signal vectors, $\boldsymbol{H}$ is the hop channel, $\boldsymbol{F}$ is the by i-th relay's receive beamforming matrix. To find the SNR at the destination can be expressed as:

$$\gamma = \frac{|\boldsymbol{H_T}|^2 \boldsymbol{F H_L^{\dagger} R_x H_L}}{\boldsymbol{\sigma_N^2}}$$

where $\boldsymbol{R_x} = E\{xx^{+}\}$is the covariance matrix of the source transmitted signal and σ_N^2 is the rely noise and UE noise part. The AF relay capacity error-free data transmission in the AF protocol, can now be succinctly expressed as

$$C = log(1+\gamma)$$

Then, to find the outage capacity, it need to find the largest R for which P_{out} does not exceed a given threshold (e.g. 0.1, which means an outage probability of 10%). This can be using the cumulative distribution function (CDF) of the channel capacity C. CDF of the channel capacity C be $F_C(c)$, the outage probability can be expressed as $P_{out} = F_c(R)$. Thus, for a given outage probability value, we can find the outage capacity by solving the equation:

$$R_{out} = (1 - P_{out})B\log_2(1+\gamma_{min})$$

In the above equation, R_{out} represents outage capacity and B is bandwidth.

4. Simulation

In this section, the beam width performance of the relay is numerically evaluated in the situation shown in the Fig.1. System parameters are summarized in Table 1. Fig. 2 depicts the outage probability in 2-hop AF relay systems, specifically related to the threshold data rate criteria that is proposed in the Smart Navigational Aid project [9]. The simulation sh ow that the narrow beamwidth (7°) is more advantageous with low variance in the buoy's angle of movement, whereas The wide beamwidth (32°) depicts that the best performance is observed when the angle variance exceeds 7 degrees.

Parameters	Values
Center Frequency	2 GHz
Bandwidth	10 MHz
BS antenna Height	30 m
Relay antenna height	10 m
Destination antenna height	5 m
Beamwidth	7°, 15°, 32°
BS power	46 dBm
Relay power	23 dBm
Threshold data rate	3kbps
Ocean wave period	10 sec
Distance BS – Relay	10 km
Distance Relay– destination	5 km

Table. 1. Simulation parameter

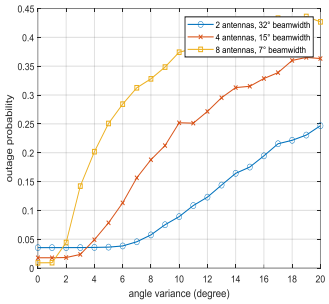


Fig. 2. Outage probability according to buoy angle

4. Conclusions

In this paper, we gain valuable insights into the behavior of wireless signals in marine settings, which can greatly assist in the design of effective communication systems for such envi-ronments. Our research demonstrates that adaptive beamforming techniques can achieve higher capacity and energy efficiency in near-shore communication systems by adjusting the beamwidth based on the prevailing sea state conditions.

References

[1] Zhou, Z., Cui, JH., Zhou, S. (2007). Localization for Large-Scale Underwater Sensor Networks. In: Akyildiz, I.F., Sivakumar, R., Ekici, E., Oliveira, J.C.d., McNair, J. (eds) NETWORKING 2007. Ad Hoc and Sensor Networks, Wireless Networks, Next Generation Internet. NETWORKING 2007. Lecture Notes in Computer Science, vol 4479. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-72606-7_10

[2] J. N. Laneman, D. N. C. Tse and G. W. Wornell, "Cooperative diversity in wireless networks: Efficient protocols and outage behavior," in IEEE Transactions on Information Theory, vol. 50, no. 12, pp. 3062-3080, Dec. 2004. https://doi.org/10.1109/TIT.2004.838089.

[3] R. Mudumbai, D. R. Brown Iii, U. Madhow and H. V. Poor, "Distributed transmit beamforming: challenges and recent progress," in IEEE Communications Magazine, vol. 47, no. 2, pp. 102-110, February 2009. https://doi.org/10.1109/MCOM.2009.4785387.

[4] Lee, Yee Hui et al. "Near Sea-Surface Mobile Radiowave Propagation at 5 GHz: Measurements and Modeling." Radioengineering 23 (2014): 824-830.

[5] N. Mehrnia and M. K. Ozdemir, "Novel maritime channel models for millimeter radiowaves," 2016 24th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), 2016, pp. 1-6, https://doi.org/10.1109/SOFTCOM.2016.7772162.

[6] M. S. Choi, B.-m. Kim, K. H. Kim, M. A. Jeong, and S. R. Lee, "Ship to ship maritime communication for e-navigation using ieee 802.16 j," Proc. IEEE Int'l Advanced Science and Technology Letters (ASTL), pp. 110–115, 2013.

[7] Y. Huo, X. Dong and S. Beatty, "Cellular Communications in Ocean Waves for Maritime Internet of Things," in IEEE Internet of Things Journal, vol. 7, no. 10, pp. 9965-9979, Oct. 2020. https://doi.org/10.1109/JIOT.2020.2988634.

[8] S. Borade, L. Zheng, R. Gallager, Amplify-and-Forward inWireless Relay Networks: Rate, Diversity, and Network Size, IEEE Trans. Info. Theory, v. 53, N. 10, pp. 3302-3318, Oct. 2007.

[9] K.-J. Lee, T.-W. Kim, Sungyoon Cho, Kiwon Kwon and D.-K Kim, "Extension of communication coverage through amplify-and-forward relay for Maritime Communication Systems," in KSII The 14th International Conference on Internet (ICONI) 2022.

Copyright ⓒ 2023 KSII

JenTrip: A Travel Platform for Solo Travelers

Honglae Kim, and Seyoung Kim*
HUPLE Co., Ltd.
Seoul, South Korea
[e-mail: honglae@huple.kr, lacrosse28@huple.kr]
*Corresponding Author: Seyoung Kim

Abstract

This article describes the development and servitization of the data system for 'JenTrip (mobile application), a local travel platform created by HUPLE Co., Ltd. JenTrip is equipped with travel information and features tailored to solo travelers. In particular, in data-centric development, it introduces the construction of data for under-tourable travel destinations and content production, and the safety map structure based on 21 types of local and region-specific data, such as safety, facilities, weather, and transportation, for real understanding of the travel areas. It also presents the planner function that links nearby local stores based on the route of the traveler. Moreover, to facilitate travel planning and ensure easy usability for subscribers, content development for providing a travel route integrated with recent generative AI has been introduced. HUPLE Co., Ltd., as the operator of JenTrip, strives to solve social issues of local extinction through regional dispersion and balanced travel routes, while observing the active movements of travelers according to the emotional excellence of the travel destinations.

Keywords: Solo Traveler, Under-tourism, Travel Planner, Local Travel Content Development, Travel Platform

1. Introduction

Surveys show that 25% of potential global trips are self-planned and taken solo. Of these, 80% are led by female populations. The biggest concern for them is undoubtedly safety. Nevertheless, the rate of solo travels is increasing every year, and the current travel market in the era of hyper-personalization solidly distinguishes between services for family and group travelers and those for individual and small group travelers. Our smart tourism business 'JenTrip' has been developed in reflection of these trends, and to provide the necessary information and features for solo travelers via mobile application.

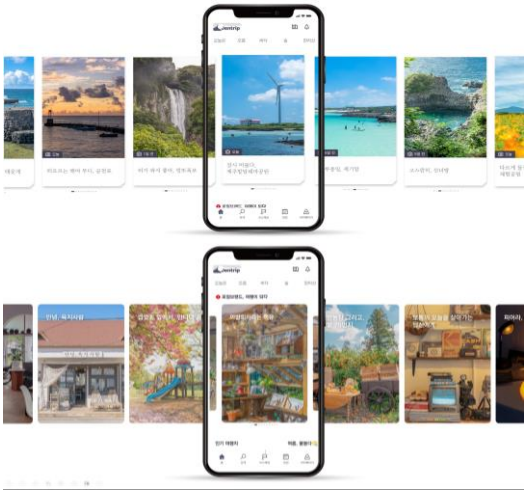


Fig. 1. Information screen displaying JenTrip's local, and local brand travel

2. Development of Tourism Data System

2.1 Informationalization for Under-Tourism

The concept of 'undertour,' contrasted with 'overtour,' is a type of travel favored by modern people who are wearied from urban competition and seek personal healing.

Responding to such demand, JenTrip has informationalized and built Points of Interest (POIs) of Jejudo, Korea's largest tourist destination, making it easier for them to search a variety of attractions in their individual travel itineraries. The constructed information focuses on elements of nature's sentimental landscapes,

themes like flowers, animals, sunsets and sunrises; activity options like hiking, trekking, and surfing; as well as spatial aspects like galleries, museums, and villages that encapsulate local art and culture, aiming to enhance the satisfaction of healing.

The informatization aims to extend from travel destinations to nearby local shops, cafés, restaurants, etc. Because these stores have distinctive features in architecture, interior design, and unique items for sale, JenTrip defines them as an extension of the trip and refers to them as 'local brand travel.'

2.2 Safety Map Construction via Region-Specific Data

For female solo travelers and tourists with special needs, JenTrip has constructed a hub of safety information in the form of an electronic map. This allows them to find public restrooms, convenience stores, public transportation, as well as to quickly respond to emergencies by directly connecting to law enforcement agencies and indicating locations of nearby hospitals and pharmacies.

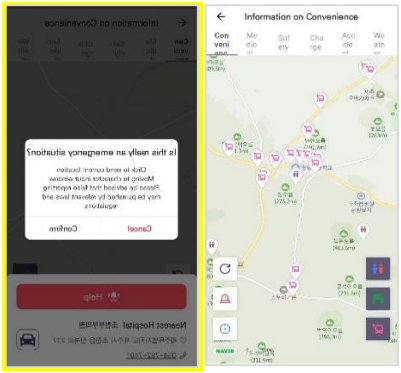


Fig. 2. Information screen displaying JenTrip's Safety Information System

We have built 21 types of local-intensive data, providing information about charging facilities for electric wheelchair users, frequent accident spots, and the expected changes in real-time weather.

2.3 Tourism Data System Development That Allows Radius Unit Queries Based on Geospatial Information

The realization of such information services involved the construction of spatial information based on geographic location in the database design, and the layer configuration of local-intensive data, publicly provided tourism data, as well as directly constructed tourist data on top.

This has been developed to provide information in radius units from a specific tourist spot and has usability that can be displayed on the interface.

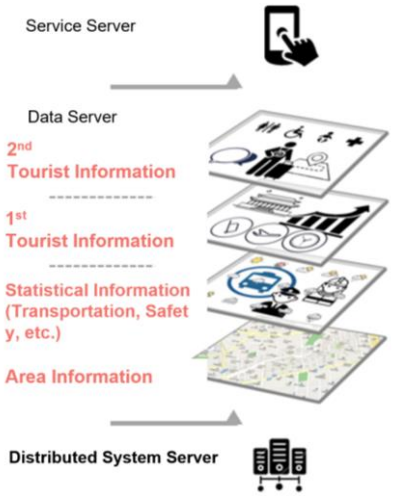


Fig. 3. JenTrip's Database Configuration System

3. Development of the Travel Planner Feature

2.1 Connecting API Based on Travel Route

The planner feature of JenTrip was developed not only to visualize the route, but also to show the current weather at each travel destination and the distance between them, and to be able to sync the itinerary with a navigation application.
Most importantly, considering that many travelers venture to places they have never been to, we provided information that allows users to determine the feasibility of daily plans by utilizing visitor entry data for each destination and calculate the average time spent.

The routes configured by users provide a search function that facilitates discovering local shops, cafés, restaurants, etc. in the area, enabling an emotionally friendly and deeply enriched travel experience even in unfamiliar areas.

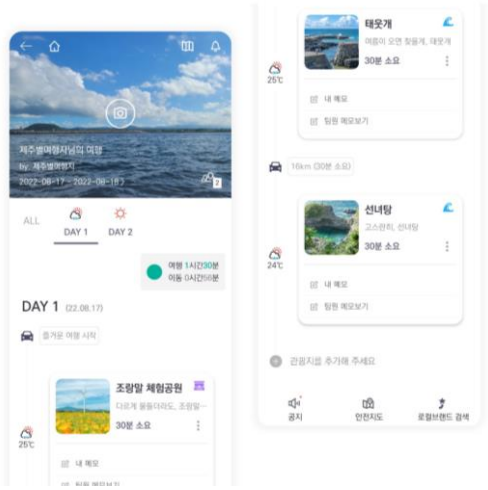


Fig. 4. JenTrip's Travel Planner Screen

4. JenTrip's Social Mission

Through the operation of JenTrip for the past two years, HUPLE CO., Ltd. has provided information on more than 1,400 travel destinations in Jeju, helping users prepare and execute their own travel itineraries. Analyzing the logs, we were able to confirm the following:

① Users under the age of 40 tend to visit travel destinations with high emotional value regardless of their popularity or awareness.

② There was clear preference for under-tours in ecological regions like Jeju.

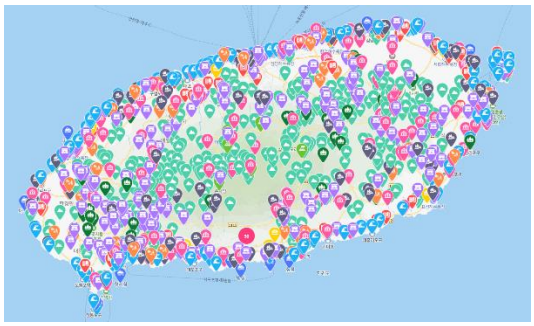


Fig. 5. Over 1,400 Jeju Travel POIs provided by JenTrip

According to this analysis, we expect to be able to escape the phenomenon of over-tourism, where there is concentration on specific tourist destinations, and achieve balanced distribution of tourism across diverse regions.

Some rural regions, despite their well-preserved ecosystems, face extinction, thus in response, we aim to effectively foster the value of under-tourism in these areas so as to spark interest and more frequent visits by travelers, providing equal opportunities for regional revitalization. Through the operation of JenTrip, we will continue to make our best efforts to achieve this goal.

HUPLE
People-Centered Data,
becoming a Valuable Platform for Travel.

Jentrip Jentrip

Fig. 5. <Above> HUPLE CO., Ltd. CI, and Slogan <Below> JenTrip BI

Copyright ⓒ 2023 KSII

Developing Travel Information Systems using Accessibility Index and VR

Eugene Choi
School of Social Science, Kangnam University
Seoul, South Korea
[e-mail: echoi0816@icloud.com]

Abstract

Leisure activities such as tourism are very important factors that determine the quality of life, and it is necessary to identify the barriers to tourism activities to secure equal leisure rights for everyone and improve the quality of life of the socially weak. The United Nations World Tourism Organization (UNWTO) declared 'accessible tourism for all' in 1991, recommending that tourism rights be guaranteed for the socially weak. The purpose of this study is to develop travel information systems using accessibility index for the barrier-free tourism for the tourism weak. In this study, the accessibility index is divided into three sub-indexes: congestion index, amenity index and physical accessibility index. This study utilizes Delphi surveys and AHP (Analytic Hierarchy Process) which are widely used in the field of social science to develop various indexes.

Keywords: Travel Information System, Accessibility Index, Barrier-free Tourism, AHP, Delphi Survey

1. Introduction

Leisure activities such as tourism are very important factors that determine the quality of life, and it is necessary to identify the barriers to tourism activities to secure equal leisure rights for everyone and improve the quality of life of the socially weak. The United Nations World Tourism Organization (UNWTO) declared 'accessible tourism for all' in 1991, recommending that tourism rights be guaranteed for the socially weak. According to the Ministry of Health and Welfare in Korea (2019), registered disabled people account for 5.1% of the total population, and the elderly population aged 65 or older is expected to increase to 20.3% by 2025. As can be seen from these projections, the number of people vulnerable to tourism in Korea on the rise.

The purpose of this study is to develop travel information systems using accessibility index for the barrier-free tourism for the tourism weak. In this study, the accessibility index is divided into three sub-indexes: congestion index, amenity index and physical accessibility index. This study utilizes Delphi surveys and AHP (Analytic Hierarchy Process) which are widely used in the field of social science to develop various indexes.

2. Methodology

2.1 Participants in the Surveys

A Total of 20 expert panels participated in the Delphi surveys including eight male panels and twelve female panels. The panels were all employees of institutions serving the socially weak or experience researchers conducting related projects.

2.2 Significance of Delphi Surveys

The Delphi survey methodology or Delphi Technique was developed in 1948 by researchers at the Rand Institute in the United States and has been widely used in forecasting by the public and private sectors. This technique is an intuitive

This research is part of the Korea Ministry of Culture, Sports and Tourism (Operating Institution: Korea Creative Content Agency) designated research project for tourism service innovation growth R&D project (Research No.: R2022020049)

prediction method that repeatedly acquires, exchanges, and develops expert opinions on future events. In this study, the purpose of running Delphi survey is to measure absolute significances of proxies of each three sub-index.

2.3 Utilization of AHP

AHP developed by Saaty is a method for determining a more valuable alternative among given alternatives. Through pairwise comparison, the decision maker's value system can be easily extracted, and the importance or weight can be calculated by objectively and consistently judging the value of a given alternative, so it is often applied to the development of indicators in the social sciences.

3. AHP Modeling

Top Indicator	Sub Indicator
congestion indicator	average number of visitors
	waiting time
	alternative destination
	real-time number of visitors
amenity indicator	equipment list
	disabled toilet
	communication equipment and personnel
	support facilities in tourist destination
physical accessibility	empty parking space
	wheelchair path
	public parking lot
	tourist attraction entrance
	recommended travel route

4. Conclusions

The conclusions obtained after conducting two Delphi surveys and AHP analysis are as follows: among given information, information on ‘support facilities for the socially weak in tourist destinations’ was revealed as the most valuable information followed by ‘communication equipment and personnel’ and ‘disabled toilet’; on the other hand, information on ‘public parking lots’ and ‘entrances to tourist destinations’ was found to have relatively low informational value for the barrier-free tourism for the socially weak.

References

[1] Ko, M. Yang, S., “Content Analysis of Welfare Tourism Research Trend in KCI Journals,” *Northeast Asia Tourism Research*, Vol 12. No. 4, pp. 191-201. 2016.

[2] Lee, J. Kim, B., “An Evidence-Based Assessment of Policy Decision Tools: A close look at AHP, Delphi, and CBA,” *Korean Society and Public Administration*, Vol 34. No.1, pp. 29-62, 2023.

Copyright ⓒ 2023 KSII

Rethinking the Locus and Focus of Citizen Participation: Role of UX and UI for Urban Planning Application

Woomin Nam, DaeHwan Kim and Keon Chul Park*
Urban Computing & Innovation Lab, Advanced Institute of Convergence Technology
[e-mail: applewoods@snu.ac.kr, dhbit@snu.ac.kr, parkkc07@snu.ac.kr]
*Corresponding author: Keon Chul Park

Abstract

This paper aims to analyze the UX/UI of a map-based citizen participation platform in smart city and find the features of citizen-friendly UX/UI. With the global urbanization rate and urban population continue to rise, smart city has received attention as one of the solutions to address various urban problem. Smart city has transitioned from a technology-driven model, city-led model to currently adopting a citizen co-created model. Additionally, efforts are being made to activate citizen participation by utilizing ICT technologies. However, examining the web traffic of citizen participation platforms reveals an average bounce rate of 69.23%, indicating that citizen participation is not actively taking place on these platforms. This study analyzes the UX/UI of citizen participation platforms using heuristic evaluation and honeycomb model, aiming to explore the features of citizen-friendly UX/UI. By analyzing citizen-friendly UX/UI, this paper provides evidence that it can increase citizen participation in online platform and contribute as basic data for the development of citizen participation platform in the urban planning.

Keywords: Citizen Participation Platform, UX/UI, Heuristic Evaluation, Honeycomb Model

1. Introduction

Smart city is receiving attention as one of the ways to solve urban problems that caused by accelerating urbanization and the growth of urban populations around the world. Smart city is undergoing a transformation in the order of technology-driven, city-led, and citizen co-creation[1]. Currently, smart city adopting a citizen co-creation[2], and efforts are being made to activate citizen participation through the utilization of ICT technologies to online platform[3]. However, according to Lee[4], the web traffic of the urban planning selected as an operation case shows a bounce rate of an average of 69.23%, as show in **Table 1**[5]. This means that approximately 7 out of 10 people leave the platform without navigating to other pages within the platform due to issues with user experience(UX)/user interface(UI) and other related problems. Therefore, in order to activate citizen participation in the online space through citizen participation online platforms, it is urgent to establish user-friendly UX/UI improvement measures that can lead citizen participation.

Table 1. Urban Planning Platform's Web Traffic

Platform	Total Visits	Bounce Rate
Seoul Urban Planning Platform	58.1K	50.57%
Decidim Barcelona	54.9K	52.78%
Decide Madrid	18.1K	78.38%
Democratic Seoul	27.4K	95.20%

Therefore, this paper aims to analyze the UX/UI of geo-based citizen participation platform being used in advanced cities that are creating successful examples of smart cities. For example, S-Map, LA GeoHub, etc. Through this

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143404)

analysis, we intened to investigate the characteristics of citize-friendly UX/UI and provide directions for promoting citizen participation in the development process of citizen participation platforms in urban planning.

2. Heuristic Evaluation and Honeycomb Model

Huristic evaluation was first proposed by Jacob Nielsen, and it is an evaluation method where domain professional find and solve problems based on ten heuristic guidelines[6]. By using this method and conducting evaluations with a minimum of five professionals, it is possible to find 85% of the issues related to UX/UI[7].

The Honeycomb model is a tool for measuring the overall UX based on seven UX factors: useful, desirable, findable, credible, valuable, accessible, usable[8]. Furthermore, the Honeycomb diagram helps define the priorities of each factor[9].

3. UI/UX Analysis

In this paper, we aim to analyze the UX/UI of geo-based citizen participation platforms, namely S-Map, LA Geohub, and Spatial Big Data Analysis Platform, which are being utilized in advanced cities that are creating successful cases in the field of smart cities. However, due to the limited precedent research on UX/UI in the field of citizen participation platforms, it is challenging to derive academic consensus on the objective validity of survey questions. Therefore, we have chosen the heuristic evaluation, which is professional-based exploratory assessment approach[10]. Furthermore, we have selected the honeycomb model to analyze the overall user experience of citizen participation platforms. In Appendix 1, we have constructed a questionnaire for this purpose.

4. Conclusions

This paper aims to analyze the UX/UI of a geo-based citizen participation platform by using heuristic evaluation and honeycomb model. Although this research provides guidance for citizen participation in user-friendly online interfaces, ultimately it is expected to contribute to the establishment of citizen demand-based city administration by promoting citizen participation in the urban planning process.

References

[1] B. Cohen., “The 3 Generations of Smart Cities”, Inside the Development of the Technology Driven City, No.2, 2015.

[2] W. M. Nam, and K. C. Park. “Civic Participation in Smart City: A Role and Direction”. Journal of Internet Computing and Service, Vol 23. No. 6, pp. 76-86, 2022.

[3] Y. J. Ahn., et al, “Establishment of Smart City Model Based on Citizen Participation”, Daejeon Sejong Research Institute, 2019.

[4] D. Y. Lee., “Case Studies and Implications for Operation of Online Participation Platform for Urban Planning: Focus on Information Provision and Policy Proposal Online Platform”, Korea Research Institute for Human Settlements Working Paper, No. 23, 2021.

[5] Similarweb, “Analyzed websit or app”, https://www.similarweb.com/ (Searched at April 23, 2023.)

[6] J. Nielsen & T. K. Landauer., “A mathematical model of the finding of usability problems”, In Proceedings of the INTERACT ’93 and CHI ’93 conference on Human factors in computing systems, pp. 206-213, 1993.

[7] J. Nielsen, “Why You Only Need to Test with 5 Users”, Nielsen Norman Group(Online), 2000.

[8] P. Morville, “User Experience Design”, Semantic Studios 21, 2004.

[9] Quiñones. D., and Rusu. C., “Applying a methodology to develop user eXperience heuristics”, Computer Standards & Interfaces, No. 66, 2019.

[10] H. J. An., S. H. Lee., “UX Analysis of Advertising Platform Mobile Applications for Culture and Arts Contents: Critical case study based on the UX Honeycomb model”, The Korea Contents Association, Vol 22. No. 9, pp. 1-18, 2022.

[11] G. Y. Rhree, S. C. Park, S. Y. Ryoo, “Performance Measurement Model for Open Big Data Platform”, Knowledge Management Research, vol 21. No. 4, pp. 243-263, 2020.

Appendix 1. Question by Evaluation Element [10,11]

Useful	Is the map-based visualization function of (platform A) user friendly?
	Does (Platform A) have well-organized by item?
	Does (Platform A) assist users in easily finding the information they want?
	Does the information provided by (Platform A) prove to be useful to user
	Does (Platform A) present data analysis results in a way that user can easily understand?
Usable	Does (Platform A) have a reasonable loading speed when using features such as map-based visualization and data analysis tools?
	Is it easy to navigate from the main menu to specific menu such as map-based visualization and data analysis tools?
	Does (Platform A) provide functions that are easy to understand and clear?
	Is the usage of functions, such as map-based visualization, data analysis tools, in (Platform A) convenient?
Desirable	Does using (Platform A) evoke a sense of visual appeal?
	Is the usage of text size and color harmonious in (Platform A)?
	Does (Platform A) spark user's curiosity during the citizen participation process
	Is there consistency in the application of style, layout, and functionality throughout the entire pages of (Platform A)
Findable	Is it easy to find the menu in (Platform A)?
	Is (Platform A) designed with menu layout that is easy for user to understand?
	Does (Platform A) indicate search result when conducting keyword searches?
	Is it possible(convenient) to restrict categories for search on (Platform A)
Accessible	Does (Platform A) provide information that can be easily and quickly understand?
	Is (Platform A) using a simple word?
	Can user easily find the menu they want in (Platform A)
Credible	Does (Platform A) provide functions that can be trusted? (In terms of errors or issues during usage)
	Does (Platform A) offer clear explanations for errors that occur within the platform?
	Is the information provided by (Platform A) reliable enough?
Valuable	Does the information provided by (Platform A) feel like it is delivering the necessary information to the user?
	Does (Platform A) give a citizen participate experience?

Copyright ⓒ 2023 KSII

Where and Why? A Novel Approach for Prioritizing Implementation Points for Public CCTVs using Urban Big Data

Daehwan Kim and Keon Chul Park*
Advanced Institute of Convergence Technology, Urban Computing and Innovation Lab
[e-mail: dhbit@snu.ac.kr, parkkc07@snu.ac.kr]
*Corresponding author: Keon Chul Park

Abstract

Citizens' demand for public CCTVs continues to rise, along with an increase in various crimes and social problems in cities. In line with the needs of citizens, the Seoul Metropolitan Government began installing CCTV cameras in 2010, and the number of new installations has increased by over 10% each year. As the large surveillance system represents a substantial budget item for the city, decision-making on location selection should be guided by reasonable standards. The purpose of this study is to improve the existing related models to establish a methodology to analyze priority regions for Seoul-type public CCTVs and propose new mid- to long-term installation goals. Additionally, using the improved methodology, we determine the CCTV priority status of 25 autonomous districts across Seoul and calculate the goals. Through its results, this study suggests improvements to existing models by addressing their limitations, such as the sustainability of input data, the conversion of existing general-purpose models to urban models, and the expansion of basic local government-level models to metropolitan government levels. The results can also be applied to other metropolitan areas and are used by the Seoul Metropolitan Government in its CCTV operation policy.

Keywords: CCTV, Priority Regions, Public Data, Spatial Analysis, Seoul Metropolitan City

1. Introduction

As social interest in safety increases, citizens' demand for the expansion of the provision of public CCTVs for crime prevention is also increasing. In the past, when selecting new installation sites for public CCTVs, there were no systematic standards or manuals. However, from an objective point of view, the fact that there have been many requests (or complaints) does not necessarily mean that CCTVs are urgently needed in that area. Additionally, a single location includes at least three to four cameras, and considering such things as electrical wiring, additional construction, and subsequent operations and maintenance, each installation location costs about 20–30 million won.

Considering the impact on the budget, the reality is that public CCTVs cannot be installed at all demand points. Therefore, an objective and reasonable selection process for new public CCTV installation regions has to be applied, rather than selecting locations through an unstructured and subjective method.

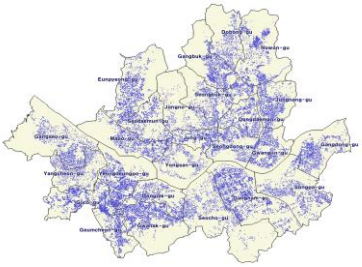


Fig. 1. Public CCTV Installation Distribution in Seoul

This work is supported by the Korea Agency for Infrastructure Technology Advancement(KAIA) grant funded by the Ministry of Land, Infrastructure and Transport (Grant RS-2022-00143404)

In this context, this study aims to review existing relevant studies and establish a methodology for analyzing Seoul-type public CCTV priority regions to overcome the limitations of existing models. The results derived by applying real data throughout Seoul are empirically shown, while new mid-to-long-term installation goals of Seoul are presented.

2. Literature Review

Harris et al. [1] studied the cost-effectiveness of CCTVs and the direction of the introduction of systems, including personal privacy issues, amid the growing use of CCTVs in the 1990s for crime prevention in the United States and Britain. Welsh and Farrington [2] found that CCTV-equipped areas saw crime reduction of more than 20% compared with control groups. It was most effective in reducing crime in parking lots, especially when coupled with improved street lighting and when focused on vehicle-related crime. On the other hand, a study by Gil et al. [3] showed that unlike the public's positive perception of CCTV installation and operation in the public domain, the strength of positivity was relatively low in neighborhoods with housing, where the installation areas are private.

Furthermore, Kim [4] proposed a method for calculating the risk points for each building by utilizing the location and property of the road and the building's location and property to improve location analysis. Previously, crime-related information was provided only in administrative units; thus, the details of vulnerable areas could not be identified. In particular, Jang et al. [5] turned away from CCTV location analysis based on crime occurrence data by administrative dong (district) unit; they proposed a method to improve CCTV location analysis by combining the installation locations for crime prevention with address information of crime occurrence points and basic spatial data of case areas.

Lee and Kim [6] proposed the selection of locations for Ansan City, one of the local governmental areas, by applying floating population data. Park and Kim [7] demonstrated how to analyze CCTV locations using crime statistics with no address details. Identifying vulnerable areas using dangerous road extraction, building use, building type, income distribution, population distribution, and existing CCTV concentration areas was highly accurate compared with the crime-ridden areas presented in the national safety map. Meanwhile, Park and Lee [8] proposed an optimal location selection method for CCTVs, considering factors for three-dimensional filming, including the size and focal distance of the "Charge Coupled Device" that affects the angle of view of cameras, the height of installation related to the filming range, and the angle.

In terms of location analysis conducted at the government level, the Ministry of Public Administration and Security, South Korea, developed in 2016 and distributed in 2017 standardized public CCTV priority location analysis manuals to prevent overlapping and inefficiency of similar projects among agencies. This model supports existing decision-making systems that relied on civil complaints or personnel experience and intuition to be reasonable. The primary data used in the model are CCTV installation information, crime occurrence address, store business data, floating population data, resident population data, housing type data, and registered foreign address. Based on the above data, two sub-indexes and three weights are calculated, and the final priority index [PI] by grid cell is derived [9]. Since its development and distribution, this model has been frequently used by local governments.

Prior research was limited in that it did not cover the characteristics of the Seoul Metropolitan Governmental region, big data of urban sustainability, and policy usability. Therefore, this study proposes a methodology to improve these three aspects and proposes a new number of mid- and long-term CCTV installation targets achieved early by the Seoul Metropolitan Government through the developed methodology. We also show the results of the application throughout Seoul as an example.

3. Methodology

The analysis was carried out in the following six steps. First, the original data were secured and refined for analysis. The grid cells, which were not subject to analysis, were excluded at this stage.

Table 1. List of Databanks

Collected Data	Level
Seoul Metropolitan Area Grid (map.ngii.go.kr, 2020. 4.)	Grid Cell (100m*100m)
Existing CCTV installation locations (Each district of Seoul, 2020. 4.)	Latitude, Longitude
Status of the five major crimes (data.seoul.go.kr, 2019. 10.)	Districts (Gu)
Commercial store data (data.go.kr, 2019. 12.)	Address
Location of financial establishment (Crawling, 2020. 5.)	Latitude, Longitude
Living population data* *Population existing in a specific area at a unit time (data.seoul.go.kr, 2020. 1.)	Districts (Dong)
Resident registration population (2020. 4.)	Grid Cell (100m*100m)
Building register headings (cloud.eais.go.kr, 2020. 3.)	Address
Number of registered foreigners (data.seoul.go.kr, 2020. 1st quarter.)	Districts (Dong)
Number of productive populations (map.ngii.go.kr, 2020. 4.)	Grid Cell (100m*100m)
National Building Integration Information (Ministry of Land, Infrastructure, and Transport, South Korea, 2020. 9.)	Building

Next, to calculate the CCTV Monitoring Vulnerability Index (MVI), monitoring coverage was set based on the location of installed CCTV cameras; the coverage area of the device was removed from the corresponding grid cells. Three factors(Occurrence Factor, Environmental Factor and Floating Population Factor) were calculated and added together for the Crime Vulnerability Index (CVI) in step 3. In step 4, the three weights were calculated to reflect the other characteristics of each grid cell. The final PI for each grid cell was calculated by adding the CCTV Monitoring Vulnerability Index and the CVI (calculated in steps 2 and 3) and multiplying the weights. Finally, the PI was graded into five levels to derive the status of Seoul Metropolitan Government districts and utilized to set standards for mid-to-long-term CCTV installation goals

4. Conclusions

This study produced a proposed systematic and objective method for analyzing CCTV priority regions and a method for calculating the number of mid- to long-term installation goals from the perspective of the Seoul Metropolitan Government. his study overcomes the limitations of existing models by taking into account the view of the Seoul Metropolitan Government, the sustainability of urban big data, and policy-utilization aspects. It also provides as a basis for the Seoul Metropolitan Government's mid- and long-term goals for installing CCTVs and the differential allocation of subsidies to autonomous districts.

References

[1] Harris, C., Jones, P., Hillier, D., & Turner, D., "CCTV surveillance systems in town and city centre management," *Property Management*, 16(3), 160–165, 1998.

[2] Welsh, B., & Farrington, D., "Evidence-based crime prevention: The effectiveness of CCTV," *Crime Prevention and Community Safety*, 6(2), 21–33, 2004. https://doi.org/10.1057/palgrave.cpcs.8140184

[3] Gill, M., Bryan, J., & Allen, J., "Public perceptions of CCTV in residential areas: "It Is Not as Good as We Thought It Would Be," International Criminal Justice Review, 17(4), 304–324, 2007. https://doi.org/10.1177/10575677073

[4] Kim E., "Methodology of Identifying Crime Vulnerable Road and Intersection Using Digital Map Version 2.0," *Journal of Korean Society for Geospatial Information Science*, 22(4), 135-142, 2014.

[5] Jang H., Kim K., & Lee J., "A study on the improvement of CCTV location for crime prevention by citizens' daily activity pattern," *The Korean Urban Geographical Society*, 17(1), 101–112, 2014.
UCI: G704-001360.2014.17.1.011

[6] Lee M., & Kim Y., "Location modeling of CCTV for crime prevention in Ansan using floating population and population density," *The Geographical Journal of Korea*, *48*(4), 533–546, 2014.
UCI : G704-001284.2014.48.4.011

[7] Park J., & Kim, E., "Methodology of extraction of crime vulnerable areas through grid-based analysis," *Journal of the Korean Society of Surveying, Geodesy, Photogrammetry and Cartography*, *33*(4), 221–229, 2015.

[8] Park J., Lee S., "Optimal location allocation of CCTV Using 3D Simulation," *Journal of the Korean Association of Geographic Information Studies*, 19(4), 92–105, 2016. https://doi.org/10.11108/kagis.2016.19.4.092

[9] Ministry of the Interior and Safety, South Korea, *Public big data standard analysis model manual: CCTV*, National Information Society Agency, 2017.

Copyright ⓒ 2023 KSII

Leverage time-based data to improve recommendation system accuracy

Sangmin Lee[1], Jeon Janggun[2], Lee Yejin[3] and Namgi Kim[3]*
[1]Department of Computer Science, Kyonggi University
[2]Department of Electric Engineering, Kyonggi University
[3]Department of AI Computer Engineering, College of Software Management, Kyunggi University
Suwon, South Korea
[e-mail: d9249@kyonggi.ac.kr, 201812285@kyonggi.ac.kr, sukiouob@kyonggi.ac.kr, ngkim@kyonggi.ac.kr]
*Corresponding author: Namgi Kim

Abstract

This study shows that the industry is developing rapidly with the development of mobile devices, and most users can easily receive various recommendations through mobile devices. In order to maintain the quality of user experience, predicting the recommendations that current users want is the main takeaway from this study. Through our experiments, we found that making recommendations based on recent user data rather than the amount of data can improve the overall recommendation accuracy. However, we also found that the amount of data is important in terms of recall score. Further research is needed to improve both accuracy and recall scores.

Keywords: Graph Neural Network, Recommendation System, Link prediction.

1. Introduction

In order to improve the quality of these services, companies have provided services to recommend items for users to buy or watch by referring to their purchase history and viewing history, and these recommendation systems are being researched in various fields.

Representative success stories include Amazon, Netflix, Facebook, Baidu, and Alibaba, which have adopted personalized recommendation systems to increase sales and brand loyalty, differentiate themselves from competitors by meeting individual customer preferences, and apply recommendations to relevant customer experiences, new users, or new catalog items in real time, with 35% of their revenue coming from recommendations, and Netflix, which is motivated by the fact that "Netflix members lose interest after 60 to 90 seconds if they cannot decide on a movie after reviewing 10 to 20 titles". In order to improve the accuracy and personalization of these services, Netflix held a technology competition called "Netflix Prize" and offered $1 million to anyone who could improve the quality of Netflix's recommendation system "CineMatch" by 10% [1].

These recommendation systems are a major source of revenue for many industries, and the development of these systems requires efficient ways to use large amounts of data.

2. Recent trends in recommendation systems

The existing recommendation system has been developed based on technologies such as Collaborative Filtering, Content-based Filtering,

This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2020R1A6A1A03040583) and This work was supported by the GRRC program of Gyeonggi province. [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research] and This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)(2021-0-01393)

and Hybrid Recommendation, but these methods have the disadvantage of not showing high accuracy because they do not properly capture the interaction between data users and items, and as a way to capture such interactions better than the existing ones, research on graph-based recommendation methods that express the relationship between users and items in the form of a graph is actively being conducted, and by applying a type of graph neural network (GNN), it is possible to express the interaction between users and items in the form of a graph, and based on this, personalized recommendations can be made more accurately [2, 3, 4, 5, 6, 7].

In this study, GNN, which can capture the relationship between users and items better than existing methods, was used in the recommendation system, and experiments were conducted on how to use the data to predict the items that users want at the current time.

3. Experiments

The goal of this experiment is to predict "recommendation of items that the current user wants", and since the experiment requires predicting the current time, the main data attribute is time, and "MovieLens 100k", which contains temporal information about user behavior, was used for the experiment. The timestamp attribute refers to the time when the user who consumed the item left the item rating, and we sorted the data in chronological order based on this attribute and divided the data to conduct the experiment. We conducted the experiment based on the idea that recommending items consumed by recent users would be more effective in performance in order to accurately predict "recommend items that the current user wants." The test data for evaluating the model's performance was assumed to be the latest data based on the timestamp, and the same data was used for all experiments. The process of the data divided by period and the experimental results by data are visually represented as A, B, C, D, and R in **Fig.s 1, 2,** and **3**, where R is the data for conducting experiments using the same number of data as A, which separates only the most recent data, and is randomly extracted from the entire training data and used as training data. The performance results of the overall experimental results show that the experimental results of A using only the most recent data with the smallest number of data have higher F1, NDCG@20, MRR@20, and Precision@20 scores than learning more data.

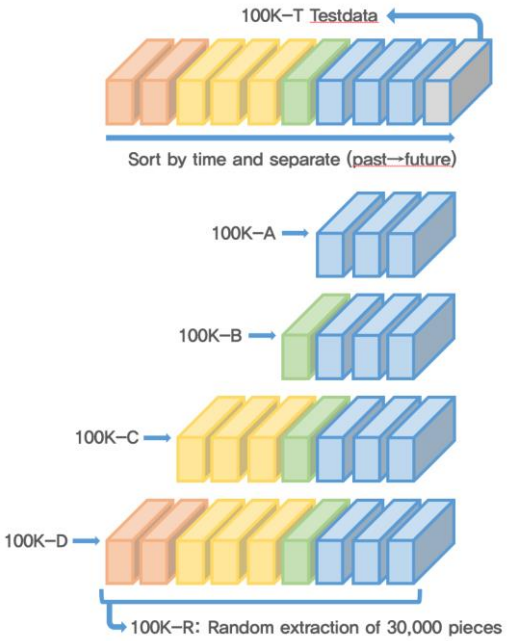


Fig. 1. MovieLens 100k Data separation schematic



Fig. 2. Graph comparing NDCG@20 and MRR@20 performance by data

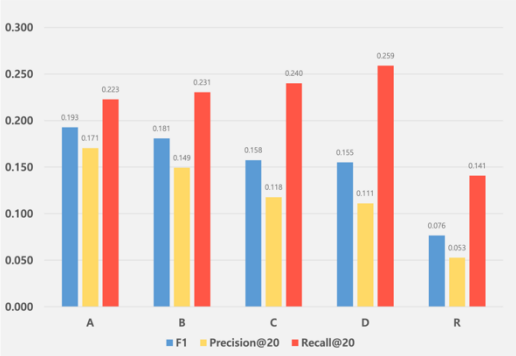


Fig. 3. Graph comparing Precision@20 and Recall@20 performance by data

4. Conclusion and future research

Most of the existing research on Collaborative Filtering, Content-based Filtering, Hybrid Recommendation, Graph-based Recommendation, etc. is based on capturing meaningful user-item relationships by applying the entire data and new methodologies. However, in this study, we studied how to utilize data to more accurately predict the "current user's desired item recommendation" and proved through experiments that using only recent user records is more accurate. However, by learning only recent data, we found that the recall score was less accurate than learning using all data.

We believe that it is necessary to apply a methodology to improve this, and for this purpose, instead of not learning the user's past information at all, we believe that applying a method to predict the user's interests by gradually changing the user's interests based on past information is a way to improve both precision and recall scores, and we plan to conduct research for this.

Acknowledgement

"This research was supported by Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education (No. 2020R1A6A1A03040583)" and "This work was supported by the GRRC program of Gyeonggi province." [GRRC KGU 2020-B04, Image/Network-based Intellectual Information Manufacturing Service Research] and "This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation)"(2021-0-01393)

References

[1] F. Ricci, L. Rokach, and B. Shapira, "Introduction to recommender systems handbook." Springer, Boston, MA, pp. 1-35, 2011. DOI: 10.1007/978-0-387-85820-3_1.

[2] X. Su, and T. M.K. hoshgoftaar, "A survey of collaborative filtering techniques." Advances in artificial intelligence, 2009. DOI: 10.1155/2009/421425.

[3] R. Burke, "Hybrid recommender systems: Survey and experiments." User model and user-adapted inter, 12(4), 331-370, 2002. DOI: 10.1023/A:1021240730564.

[4] P. Resnick and H. R. Varian, "Recommender systems." Communications of the ACM, vol. 40 no. 3, pp. 56-58. 1997.

[5] R. Ying, R. He, K. Chen, P. Eksombatchai, W. L. Hamilton, and J. Leskovec, "Graph Convolutional Neural Networks for Web-Scale Recommender Systems." In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '18). Association for Computing Machinery, New York, NY, USA, 974–983, 2018. DOI: 10.1145/3219819.3219890.

[6] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and P. S. Yu, "A comprehensive survey on graph neural networks." IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 1, pp. 4-24, 2020. DOI: 10.1109/TNNLS.2020.2978386.

[7] W. L. Hamilton, R. Ying and J. Leskovec, "Inductive Representation Learning on Large Graphs." Advances in neural information processing systems, 30, 2017. DOI: 10.48550/arXiv.1706.02216.

Copyright ⓒ 2023 KSII

Speech Intelligibility Prediction of Dysarthria Using Deep Convolutional Networks

Junseok Oh, Hosung Park and Ji-Hwan Kim*
Department of Computer Science and Engineering, Sogang University
Seoul, South Korea
[e-mail: ohjs@sogang.ac.kr, hosungpark@sogang.ac.kr, kimjihwan@sogang.ac.kr]
*Corresponding author: Ji-Hwan Kim

Abstract

Speech intelligibility, the degree to which a listener understands speech, is a crucial factor in diagnosing and managing speech impairments. This study proposes an objective and efficient method for its prediction using a Convolutional Neural Network, specifically the ResNext101 model. The method relies on a dataset of Diadochokinetic (DDK) task voices from 319 patients, previously assessed by experts. The model takes the Mel-spectrogram of DDK task speech as input and classifies it into five levels of speech impairment. A balanced categorical cross-entropy was used to mitigate data imbalance during model training. The model achieved a macro-average accuracy of 59.6% and a micro-average accuracy of 72.1%, demonstrating its effectiveness in predicting speech intelligibility. This work paves the way for future research aimed at improving accuracy and generalizability across a broader range of speech disorders.

Keywords: speech intelligibility, Convolutional Neural Network, dysarthria

1. Introduction

Speech intelligibility is vital for diagnosing and managing speech impairments [1]. It is affected by factors such as articulation, speech rate, rhythm, and voice quality. Dysarthria, a motor speech disorder, presents significant challenges in speech intelligibility [2]. Expert evaluations, though essential, can be subjective, time-consuming, and costly. Therefore, an automatic prediction method using a Convolutional Neural Network (CNN) model is proposed to classify speech impairment levels [3].

2. Related Works

CNNs have been widely applied in speech and audio processing due to their ability to learn robust features directly from raw data [4]. CNNs have been used in hybrid NN-HMM models for speech recognition, demonstrating their efficiency in learning hierarchical features.

3. Dataset

The dataset [1] consists of recordings from 319 patients with language disorders performing a Diadochokinetic (DDK) task [5]. A DDK task is a standard speech pathology assessment

[1] This research (paper) used datasets from 'The Open AI Dataset Project (AI-Hub, S. Korea)'. All data information can be accessed through 'AI-Hub (www.aihub.or.kr)'.

This work was supported by Institute of Information & communications Technology Planning & Evaluation (IITP) grant funded by the Korea government(MSIT) (No.2022-0-00621,Development of artificial intelligence technology that provides dialog-based multi-modal explainability)

involving the rapid, clear repetition of a series of simple syllables, typically /pa/, /ta/, and /ka/. These tasks are used to evaluate the motor abilities of the speech articulators and can reveal potential movement limitations. The patients' performances were previously assessed by experts on a 1-5 speech impairment scale. In total, 957 files were compiled and divided into training, validation, and testing sets at an 8:1:1 ratio.

4. Intelligibility Prediction Model

The ResNext101 [6] architecture with Mel-spectrogram input was employed to predict speech intelligibility automatically. This model incorporates a 2D convolution in the time-frequency domain, enabling it to capture the intricate patterns present in the Mel-spectrogram data efficiently. The model takes a 224-dimensional Mel-spectrogram of DDK task speech as input and classifies it into one of five speech intelligibility levels [7].

To mitigate data imbalance, weighted categorical cross-entropy was used as the loss function. Weighted cross-entropy is a variant of the cross-entropy loss function, which assigns different weights to different classes. This is particularly useful when dealing with imbalanced datasets.

Additionally, data augmentation techniques were utilized to enhance the robustness of the model further. Techniques such as SpecAugment [8] and speed perturbation were employed to create variations in the dataset, thus aiding in better generalization of the model by exposing it to more diverse scenarios within the data.

This combined approach of using a weighted loss function and data augmentation helps effectively deal with data imbalances and enhance the model's performance.

5. Experiment

The model was trained and evaluated using the collected dataset, achieving a macro-average accuracy of 59.6% and a micro-average accuracy of 72.1%. These results suggest the model's effectiveness at predicting speech intelligibility based on DDK task voices.

5. Conclusion

In this paper, we have proposed a CNN-based model for the prediction of speech intelligibility. The proposed model utilizes the ResNext101 architecture, proving itself an efficient, objective tool for automatically predicting speech intelligibility in patients with language disorders. In future work, we will focus on enhancing the accuracy of this model and broadening its generalizability.

References

[1] R. D. Kent *et al.*, "Acoustic studies of dysarthric speech: methods, progress, and potential," *Journal of Communication Disorders*, vol. 32, no. 3, pp. 141-186, 1999.

[2] J. R. Duffy, "Motor Speech Disorders: Substrates, Differential Diagnosis, and Management," Elsevier Mosby, 2005.

[3] S. G. Fletcher, "Time-by-count measurement of diadochokinetic syllable rate," *Journal of Speech and Hearing Research*, vol. 15, no. 4, pp. 763-770, 1972.

[4] O. Abdel-Hamid *et al.*, "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition," *in Proc. of 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 4277-4280, 2012.

[5] S. G. Fletcher, "Time-by-count measurement of diadochokinetic syllable rate," *Journal of Speech and Hearing Research,* vol. 15, no. 4, pp. 763-770, 1972.

[6] S. Xie *et al.*, "Aggregated Residual Transformations for Deep Neural Networks," *in Proc. of the IEEE Conference on computer vision and pattern recognition (CVPR)*, pp. 1492-1500, 2017.

[7] S. B. Davis and P. Mermelstein, "Comparison of parametric representations for monosyllabic word recognition in continuously spoken sentences," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 28, no. 4, pp. 357-366, 1980.

[8] D. S. Park *et al.*, "SpecAugment: A Simple Data Augmentation Method for Automatic Speech Recognition," *in Proc. of the Interspeech 2019*, pp. 2613-2617, 0Graz, Austria, Sep. 2019.

Copyright ⓒ 2023 KSII

Analysis of H.266 Versatile Video Coding Streaming and Decoding on Smartphones

Youngwook Kim[1,2] and Jong-Moon Chung[1*]
[1]School of Electrical & Electronics Engineering, Yonsei University
Seoul, South Korea
[e-mail: kyw8738@yonsei.ac.kr, jmc@yonsei.ac.kr]
[2]MX Division, Samsung Electronics Co., Ltd.
Suwon, South Korea
*Corresponding author: Jong-Moon Chung

Abstract

The demand for ultra-high-resolution and 360° video streaming is increasing. As a result, a highly efficient video compression standard to support these services is required, in which the H.266 Versatile Video Coding Streaming (VVC) standards were created. The performance and energy consumption of VVC encoding and decoding on smartphones are critical because much of the video demand comes from mobile devices. In addition, due to the limited battery capacity and processor performance of mobile devices, the decoding performance of VVC on smartphones is a key indicator. This paper presents a summary of the VVC and streaming protocols. In addition, the major factors for measuring the VVC decoding performance on smartphones are introduced.

Keywords: Versatile Video Coding, VVC, H.266, Smartphone

1. Introduction

The Joint Video Experts Team (JVET) group recently announced the H.266 Versatile Video Coding (VVC) standards that can handle enhanced high-resolution and 360° video. The decoding performance of VVC on smartphones is very important because many video streaming services are consumed on smartphones. In this article, the smartphone elements that can be used in analyzing and decoding the performance of VVC are presented.

2. Versatile Video Coding (VVC) and Video Streaming

VVC supports versatility with a higher level of compression than HEVC. In particular, VVC is designed to support high-resolution video (8K+) and immersive video such as 360° video. The main features of the VVC coder include a wide range of block partitioning from 128×128 to 4×4 CU sizes, subpicture-level processing, and support for combined inter/intra prediction (CIIP), as compared to HEVC. These features provide a higher level of compression and versatility, but also much higher complexity than HEVC [1]. Therefore, it inevitably requires more energy consumption and processing time for decoding. Furthermore, video streaming specs are used in analyzing a smartphone's decoding performance. Because streaming video over the 5G network is the most common video service. The most popular protocols for video streaming are the Real-Time Streaming Protocol (RTSP) and Dynamic Adaptive Streaming over HTTP (DASH). Fig. 1 shows an overview of DASH. DASH is a method in which the content provider stores the video in small chunks at different rates in advance, and then delivers the chunks at the most suited rate to the client, depending on the connected network status.

3. Energy Modeling for Video Streaming and Test Environment

There are several studies on energy modeling for video streaming. In [2], energy modeling was conducted according to CPU frequency, communication latency, throughput, and video resolution. In [3], energy modeling of decoding is based on the energy factor of the decoder operating features (e.g., Islice, PBslice) of HEVC and the number of execution times. In [4], details of the HEVC decoder are provided, however, since Android does not support VVC, this study used the same decoder by building it for smartphones. The tests were conducted on a Galaxy S20 smartphone with a heterogeneous core AP, Exynos990 (4×2.0 GHz, 2×2.5 GHz, and 2×2.7 GHz). Prior to testing, the CPU affinity and CPU clock were fixed for testing under control conditions. In this paper, the decoding energy model measures the battery consumption while monitoring the CPU clock and thermal status. This is because modern smartphones can utilize features such as DVFS to forcibly control the CPU and GPU performance which influences the heat and battery [5]. In the test, 6 HEVC videos were decoded for more than an hour, and an accurate energy model was obtained through a regression model as shown in **Fig. 2**. Feature-based modeling of VVC is possible in the same way [6]. For a more practical energy model, the energy model can add networking, HDMI video/speaker, and audio processing energy to the decoder energy [7]. To set up a streaming server network, the RTSP or DASH protocols are considered, in which the server/client VVC model includes FFmpeg, FFserver, Apache Web Server, and VLP Player.

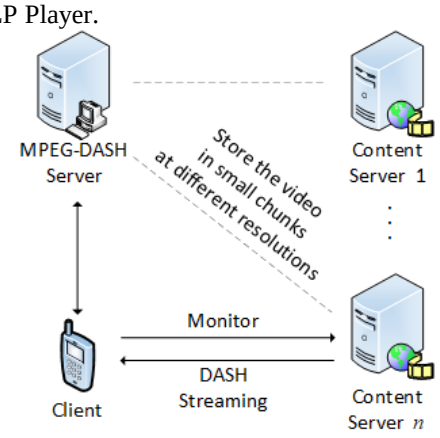


Fig. 1. Smartphone DASH measurement model.

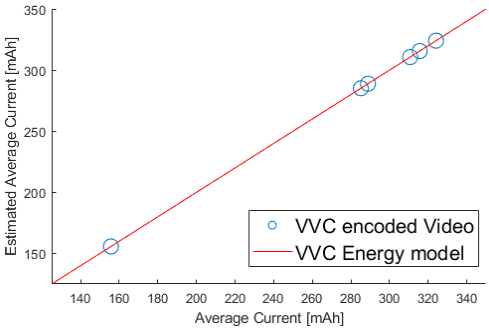


Fig. 2. Energy modeling based on decoder features.

4. Conclusions

The study proposes a new method including multi-factors to accurately measure the H.266 VVC streaming performance on smartphones.

References

[1] A. Mercat *et al.*, "Comparative Rate-Distortion-Complexity Analysis of VVC and HEVC Video Codecs," *IEEE Access*, vol. 9, pp. 67813-67828, 2021.

[2] H. Wang, B. Kim, J. L. Xie and Z. Han, "LEAF + AIO: Edge-Assisted Energy-Aware Object Detection for Mobile Augmented Reality," *IEEE Trans. Mobile Comput.*, early access, Jun. 2022.

[3] C. Herglotz *et al.*, "Modeling the energy consumption of the HEVC decoding process," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 28, no. 1, pp. 217–229, Jan. 2018.

[4] "HEVC analyzer repository" Accessed: Apr. 23, 2023. [Online]. Available: https://gitlab.lms.tf.fau.de/LMS/HM-XX_analyzer

[5] S. Kim *et al.*, "zTT: Learning-based DVFs with zero thermal throttling for mobile devices," in *Proc. ACM 19th Annu. Int. Conf. Mobile Syst. (MobiSys'21)*, New York, USA, Jun. 2021, pp. 41–53.

[6] M. Kranzler, C. Herglotz, and A. Kaup, "Decoding energy modeling for versatile video coding," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Abu Dhabi, United Arab Emirates, Oct. 2020, pp. 3144–3148.

[7] C. Herglotz *et al.*, "Power modeling for video streaming applications on mobile devices," *IEEE Access*, vol. 8, pp. 70234–70244, 2020.

Copyright ⓒ 2023 KSII

Electoencephalography-based Fatigue Assessment for Internet of Medical Things System

Jung-Hyok Kwon[1] , Sol-Bee Lee[2], and Eui-Jik Kim[2*]
[1] Smart Computing Laboratory, Hallym University
Chuncheon, South Korea
[2] Division of Software, Hallym University
Chuncheon, South Korea
[e-mail: {jhkwon, thfqla3535, ejkim32}@hallym.ac.kr]
*Corresponding author: Eui-Jik Kim

Abstract

This paper presents an electroencephalography (EEG)-based fatigue assessment, abbreviated EFA, for the Internet of Medical Things (IoMT) System. The EFA is designed to classify personal fatigue into four levels by analyzing the EEG data collected from a wearable device equipped with EEG sensors. To this end, the IoMT server analyzes the variation patterns of the EEG data periodically received from the wearable device. Specifically, it estimates the number of eye blinks per minute by counting the number of EEG data whose value is greater than a specific threshold for a certain duration. Then, the IoMT server determines the fatigue levels according to the estimated number of eye blinks. To verify the EFA feasibility, we conducted an experimental implementation using Revolution Board Kit BT Model and Raspberry Pi 4 Model B. The results showed that EFA is one of the practical IoMT applications to assess personal fatigue using EEG with a simple operational procedure.

Keywords: Data analysis, electroencephalography, fatigue assessment, IoMT, wearable device

1. Introduction

Recently, a variety of industries have been interested in the Internet of Medical Things (IoMT) system to efficiently reduce industrial accidents caused by poor physical and mental conditions of workers [1]. In particular, fatigue assessment using electroencephalogram (EEG) has been considered one of the promising IoMT applications. This is because the IoMT system can address the drawback of conventional EEG-based fatigue assessments such as complex operational procedures and expensive machinery. Specifically, the IoMT system can support a simple operating procedure for EEG-based fatigue assessment using wearable devices equipped with biomedical sensors and analysis servers.

In this paper, we focus on the development of an EEG-based fatigue assessment (EFA) for the IoMT system. In order to assess fatigue, EFA analyzes EEG data received from the wearable device to estimate the number of eye blinks per unit of time. To this end, it counts the number of EEG data samples whose value is greater than the threshold for a certain duration. Then, the EFA classifies fatigue into four levels considering the estimated number of eye blinks. To evaluate the EFA feasibility, we conducted implementation using open-source hardware Revolution Board Kit BT Model and Raspberry Pi 4 Model B. The results showed that EFA is a

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2020R1I1A3052733). This work was also supported by the NRF grant funded by the Korea government (MSIT) (NRF-2021R1C1C2095696).

practical IoMT application supporting a simple operation.

2. EEG-based Fatigue Assessment

The wearable device periodically generates and transmits the EEG data to the analysis server. The analysis server creates EEG dataset ($\mathbf{X}$) expressed by Eq. (1) every minute.

$$\mathbf{X} = \{x_1, x_2, x_3, .., x_n\} \quad (1)$$

where x_n is the n-th EEG data and the unit of each EEG data is expressed in voltage. The number of elements in $\mathbf{X}$ (i.e., n) is equal to the number of transmissions per minute.

After creating an EEG dataset, the analysis server conducts EFA consisting of the eye blink estimation and the fatigue level classification. Algorithm 1 shows the procedure of the eye blink estimation in which the number of eye blinks per minute is counted. In the algorithm, T is the predefined threshold, k is the size of the sliding window, and cnt is the estimated number of eye blinks per minute. As shown in Algorithm 1, the estimated number of eye blinks increases by one when the mean of k elements is greater than the threshold.

Algorithm 1. Eye blink estimation

```
1:  INITIALIZE i to 1, cnt to 0
2:  WHILE k(i+1) < n
3:    IF avg(X[ik : k(i+1)−1]) > T
4:      cnt ← cnt ++
5:    ELSE
6:      cnt ← cnt
7:    ENDIF
8:      i ← i ++
9:  ENDWHILE
10: RETURN cnt
```

where $avg(\cdot)$ is the average function. k is determined by $\lceil n/120 \rceil$.

Table 1. Fatigue level calssification rule

Fatigue level	Classification rule
No fatigue [1]	cnt < 20
Low fatigue [2]	20 ≤ cnt < 40
Moderate fatigue [3]	40 ≤ cnt < 60
High fatigue [4]	60 ≤ cnt < 80

Then, the analysis server conducts the fatigue level classification. Specifically, to classify the fatigue level according to the estimated number of eye blinks, it refers to the fatigue level classification rule in **Table 1**.

3. Implementation Results

The wearable device and data receiver were implemented using the Revolution Board Kit BT Model and Raspberry Pi 4 Model B, respectively. In addition, the EFA was implemented using Java and it is executed on the analysis server. The threshold was set to 700 mV, and the EEG data generation period was set to 100 ms. **Fig. 1** shows changes in EEG data for each fatigue level. As shown in the figure, the average value of the EEG data increases as the fatigue level increases.

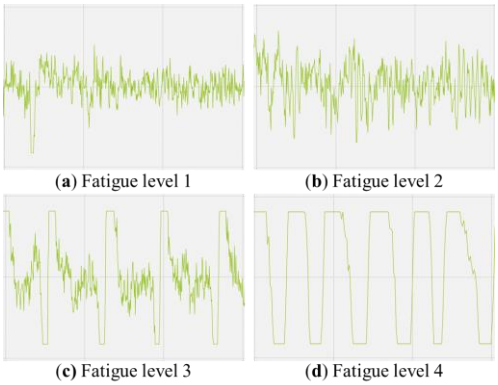


Fig. 1. EEG data for each fatigue level

4. Conclusions

In this paper, we propose an EFA for the IoMT system that consists of two steps: 1) eye blink estimation and 2) fatigue level classification. To verify the feasibility of EFA, the experimental implementation was conducted using open-source hardware. The results showed that EFA has a simple operational procedure and is feasible for a real environment.

References

[1] A. Ghubaish, T. Salman, M. Zolanvari, D. Unal, A. Al-Ali and R. Jain, "Recent Advances in the Internet-of-Medical-Things (IoMT) Systems Security," *IEEE Internet of Things Journal*, Vol. 8, No. 11, pp. 8707–8718, 2021.

KSII The 18th Asia Pacific International Conference on Information Science and Technology(APIC-IST) 2023.
Copyright ⓒ 2023 KSII

A Study on Approach for Automating Field Identification in Heterogeneous Security Events Analysis

Taehyeon Kim, Kijong Son, Saewoom Lee, Seul-Ki Choi and Taeeun Kim*
Infrastructure Security Technology Team, Korea Internet & Security Agency,
South Korea
[e-mail: thkim@kisa.or.kr, skj@ kisa.or.kr, sprout2@ kisa.or.kr, skchoi@ kisa.or.kr, tekim31@kisa.or.kr]
*Corresponding author: Taeeun Kim

Abstract

With the development of various detection technologies, the amount of security logs has increased rapidly. As a result, devices for analyzing security data have also emerged. However, these security devices face difficulties in analyzing the association of different types of devices and limitations in responding to new threats. To overcome these limitations, this paper proposes an automated field identification technique using a combination of regular expression-based essential field identification techniques and NER techniques. The proposed system is expected to increase the accuracy and reliability of identifying and extracting security events. Additionally, it is expected to increase the efficiency of maintenance and response.

Keywords: Security event, Named Entity Recognition, Identification

1. Introduction

With the recent development of various ICT technologies and communication networks, the amount of security logs in companies and organizations is rapidly increasing. As high-quality data becomes increasingly important in various fields, the need to thoroughly record and analyze security event data is increasing. Therefore, research and development is actively underway to improve devices that can store, analyze, and utilize security logs such as ESM, SIEM, and SOAR. However, the use of security devices from various companies creates new problems. It is difficult to integrate and analyze heterogeneous security event analysis device by reflecting the unique format and characteristics of each security device. In addition, logs are often integrated using defined patterns or rules, which have limitations for new security threats and disadvantages that are difficult to maintain.

Therefore, to overcome these limitations, this paper proposes an automated field identification technique that combines required field identification techniques based on regular expressions and named entity recognition (NER). This system stores the structure of identified security events in a separate file, allowing for the extraction of only desired information by comparing the stored file with security events. Therefore, it has the advantage of not having to repeat the identification process using regular expression and NER techniques. Additionally, by facilitating analysis tasks such as converting to a unified format, this system can help prevent potential security threats.

2. Technology for automated security event field identification

2.1 Data and Processing Models

The data studied in this paper has the specificity of the security event log. Considering the difficulty and specificity of collecting security event logs, suitable data used the BERT-NER model that went through fine turning method[1]. To collect learning data, a corpus suitable for the purpose was constructed

using three data argumentation techniques: SR (Synonymous Replacement), RI (Random Insertion), and RS (Random Swap)[2].

Table 1. Example of data recognition

INPUT DATA	OUTPUT DATA
<11>1 2022-05-01T00:28:07.820+09:00 Aj@u_S@#_A RT_F%W - RT_FLOW_SESSION_CLOSE [test@2345.67.890.123.4.111 reason=\"TCP SERVER RST\" source-address= \"111.1.234.198\" source-port=\"1313\" destination-address= \"123.456.123.51\" destination-port=\"443\" c@@nection-tag=\"0\" ~~~	111.1.234.198 -> IP 1313 -> PORT 123.456.123.51 -> IP 443 - PORT

Table 1 shows the trained model as an example of data recognition, and then created and entered security device data with input data. The input data is taken sequentially, converted to the input format of the BERT Tokenzer, and the IP, port, are extracted. Finally, the corresponding values and order are stored and outputted.

2.2 Configuring Automated Field Identification Techniques

Since a technique for identifying various field names and field values of heterogeneous security events was needed, field identification using regular expressions and field identification methods of the learned BERT-NER model were introduced. For efficient use such as data processing resources, it was configured as follows.

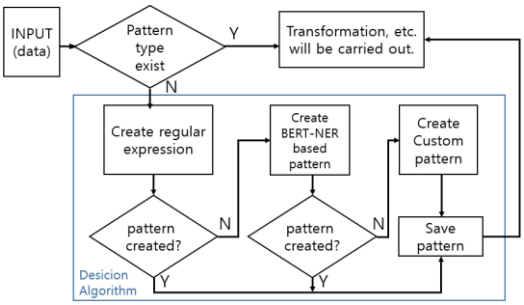


Fig. 1. Flowchart of data field identification

In this paper, the format in which the field order and field format of each security event are stored is expressed as a pattern, and when data comes in, a separate file (txt, py, etc.) is referred to check the existence of the pattern. If the pattern does not exist, the desired field is extracted through a regular expression, and if it is impossible to extract by a regular expression, the desired field is extracted using a finely adjusted BERT-NER. Finally, a user-defined method was established to reflect deficiencies If the same structure as in **Fig. 1** is used, there is an advantage that additional resources such as data identification through the use of AI models are not consumed because event identification is carried out based on the stored pattern files.

3. Conclusions

In order to solve the difficulty of analysis and utilization in heterogeneous events, we paper proposed an event field identification method that uses a combination of regular expression-based field identification technology and BERT-NER-based field identification technology. If the desired essential fields can be extracted from the event, it is expected that the response to security incidents will also be faster, such as collecting IPs of specific malicious users from various devices and tracking the malicious behavior of those users.

In the future, in addition to basic information such as IP, we will expand the scope of information essential for security threats, such as information on malicious behavior, to form a corpus and study an extraction model optimized for security threat extraction.

ACKNOWLEDGMENT

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grand funded by the Korea government(MSIT)(No.2021-0-00358, AI·Big data based Cyber Security Orchestration and Automated Response Technology Development).

References

[1] Devlin, Jacob, et al.,“Bert: Pre-training of deep bidirectional transformers for language understanding,” arXiv:1810.04805, 2018.

[2] Jason Wei and Kai Zou, “EDA: Easy Data Augmentation Techniques for Boosting Performance on Text Classification Tasks,” EMNLP-IJCNLP 2019 short paper, arXiv:1901.11196, 2019.

Copyright ⓒ 2023 KSII

Cross-Device Extended Reality Experiences

Sihyeon Paik[1], Jiseong Park[1], Seoi Na[1] and Hyoseok Yoon[2*]
[1]Division of Computer Engineering, Hanshin University
[2]School of Computing and Artificial Intelligence, Hanshin University
Osan, Gyeonggi-do, South Korea
[e-mail: tlgus000912@hs.ac.kr, pjssoj07193@hs.ac.kr, nsi1022@hs.ac.kr, hyoon@hs.ac.kr]
*Corresponding author: Hyoseok Yoon

Abstract

We propose a concept of wearable cross-device interaction for extended reality contents in human-computer interaction studies. The proposed concept involves the use of a combination of smart glasses, smartwatch, and other wearable devices to enable cross-device interaction and control of extended reality contents. We explore the benefits and challenges of using this concept in HCI research. The benefits include increased user engagement and immersion, improved interaction techniques, and enhanced accessibility, while foreseeable challenges include increased complexity and potential user discomfort. To overcome these difficulties, we provide actionable design guidelines such as the need for user-centered design and testing, seamless integration of devices and software, and consideration of user comfort and safety. Moreover, we discuss potential future directions for wearable cross-device interaction for extended reality contents, such as the use of artificial intelligence and multi-modality to enhance cross-device interaction and the exploration of novel interaction techniques.

Keywords: Cross-device interaction, wearable device, extended reality, human-computer interaction

1. Introduction

Extended reality (XR) technologies, such as virtual and augmented reality, have seen significant growth in recent years. However, current XR systems are often limited using handheld controllers, which can hinder the immersive experience and disrupt user engagement. A potential solution is to employ wearable cross-device interaction for XR contents, which enables users to interact with XR contents seamlessly across multiple devices [1]. In this paper, we explore the benefits as well as technical challenges of this concept in HCI research and provide actionable design guidelines and takeaways for implementation.

2. Wearable Cross-Device Interaction

The use of wearable cross-device interaction for XR contents offers numerous benefits for HCI research.

Enhanced user experience (UX): Wearable cross-device interaction enables users to interact seamlessly across multiple devices, providing a more immersive and engaging experience. Users can interact with XR contents using natural hand and body gestures, enabling a more intuitive and fluid interaction [2].

Increased flexibility: Wearable cross-device interaction allows users to interact with XR contents using various devices, including smartwatches, smart glasses [2], and smartphones, providing greater flexibility and adaptability.

Improved data collection: use of wearable cross-device interaction for XR contents enables researchers to collect more accurate and objective data on user behavior and preferences during XR interactions [3].

This work was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2021R1F1A1063340).

While wearable cross-device interaction for XR contents offers numerous benefits, it also has some potential disadvantages. Designing and implementing wearable cross-device interaction systems can be challenging, requiring careful consideration of device compatibility, network connectivity, and synchronization.

2.3 Privacy and Security Concerns

Wearable cross-device interaction systems require access to personal data, including user location and biometric data, raising concerns around privacy and data security.

3. Actionable Design Guidelines

To mitigate the potential detriments of wearable cross-device interaction for XR contents in HCI research, researchers should consider the following design guidelines and takeaways.

Compatibility: Design wearable cross-device interaction systems that are compatible with a wide range of devices, including smartwatches, smart glasses, and smartphones, to provide users with greater flexibility and adaptability.

Synchronization: Ensure synchronization of XR contents across multiple devices is seamless, providing users with a more immersive and engaging experience.

Privacy and security: Implement appropriate data security measures to protect user privacy, including user consent and data encryption.

User feedback: Collect and analyze user feedback on wearable cross-device interaction systems to identify and address potential issues and areas for improvement.

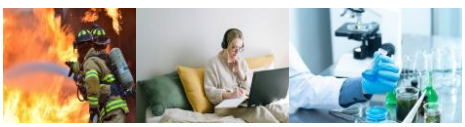


Fig. 1. Scenarios for wearable cross-device XR

Table 1 and **Fig. 1** represent how the wearable cross-device interaction can be employed in 6 XR-based contents. An eyewear device such as smart glasses or HMD is used with smartwatch and hearables to fully immerse users to the presented scenarios.

Table 1. Wearable cross-device interaction examples

Contents	Eyewear (E), Smartwatch (S), Hearables (H)
Fire evacuation	(E) Display an escape path, (S) Vibration as paint feedback, (H) Sound effect for alarm, fire, screams
Remote Lecture	(E) Eye gaze to assess the user's attention, (S) Hand-writing gesture, user's physiological sensing, (H) Noise-cancelling, wake-up sound for the sleeping students
Chemistry Experiment	(E) Chemical identification, (S) Vibration when an object is dropped, (H) Sound effect for fire and alarming

4. Conclusions

The use of wearable cross-device interaction for XR contents offers numerous benefits in HCI research, including enhanced user experience, increased flexibility, and improved data collection. However, it also has some potential limitations, including technical challenges and privacy and security concerns. To maximize the benefits and mitigate the potential detriments of this concept, researchers should consider the proposed design guidelines when implementing wearable cross-device interaction systems.

References

[1] F.-C. Tseng, Z.-Y. Chiou, H.-H. Chuang et al., "Multiple device users' actual and ideal cross-device usage for multi-stage notification-interactions: An ESM study addressing the usage gap and impacts of device context," *in Proc. of CHI 2023*, 2023.

[2] L. -H. Lee and P. Hui, "Interaction methods for smart glasses: A survey," *IEEE Access*, Vol 6, pp. 28712-28732, 2018.

[3] H. Yoon, "Multi-Sensor, third-party API, and cross-device integrated toolkit for enriching participatory research," *Journal of Multimedia Information System*, Vol 9, No. 4, pp. 287-292, 2022.

Copyright ⓒ 2023 KSII

Air Care IoT Platform for Air Quality and Energy Saving

Yoosin Kim*
* AirDeep Inc. Jeju, Republic of Korea
[e-mail: yoosin@airdeep.co.kr]
*Corresponding author: Yoosin Kim

Abstract

Indoor Air pollution including fine-dust and micro-dust are increasing day by day and pose a major threat to our society and human health. Therefore, the systems and algorithms are required to monitor and control indoor air quality (IAQ) to assure the people about the environmental conditions they are living and working on. In this paper, we conducted indoor air quality monitoring and controlling system using IoT sensor data, which collects sensor data PM10, PM2.5, CO2, VOCs, Temperature, and Humidity which has direct or indirect impact to Indoor Air Quality. The platform is composed of three parts: 1) IAQ Sensor, 2) Big Data System on Cloud, and 3) Data analysis and monitoring application. It can collect, process, analyze, visualize, and control the indoor air quality information in real time. Additionally, it can detect the fraud air pollution such as smoking cigarette in a car and indoor space.

Keywords: Air Quality, IoT, Cloud, Big Data

1. Introduction

As the air pollution with fine-dust & micro-dust has become a serious social issue, out-door activities are too much limited and uncomfortable because of the health problem and disease. And also, it causes much concern and fear about in-door life in the office, home, class, car, etc. To detect and reduce in-door air pollution, many ideas, and goods such as ventilation systems are invented and adopted into in-door facilities. [1]. To monitor and reduce the impact to the human health of fine-dust and micro-fine-dust s to human health, and improve and create fresh indoor environment, require the real time integrated intelligent system to collect, measure and improve the indoor air quality, and control the IoT sensor such as air filter or ventilator remotely as well as automatically [2]. In this paper we present an Indoor Air Quality(IAQ) monitoring platform based on IoT and Cloud computing, to control, monitor and improve indoor as well as outdoor temperature, humidity and amibient particulate matter, that are harmful to human health. This real time air quality information can be used to improve indoor air quality as well as save the energy such as electricity usage at house, resort, hotel, etc.

2. Indoor Air Quality Monitoring

2.1 Design Air-Care service Platform

The proposed air-care service platform (ACP) is designed for the IoT based indoor air quality monitoring service, as shown in **Fig. 1**, which enables to extend the IoT network and different legacy system from different location. Previous research only focuses on specific places to measure and visualize the air quality information, but ACP can collect process and analyze data from different IoT network installed in different places.

This research is supported by Korea Institute of Energy Technology Evaluation and Planning (KETEP) grant funded by the Korean government (MOTIE) (Project No. 20192020101170)

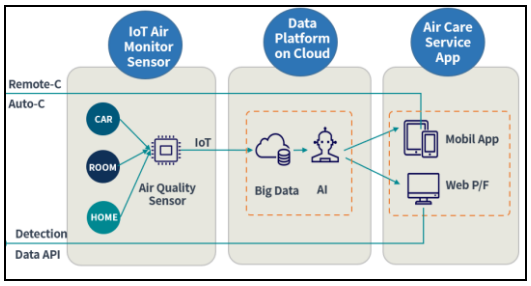


Fig. 1. Air-Care Platform architecture

Therefore, ACP can monitor air quality devices such as the air conditioner, the ventilation, the humidity remover, etc., and operate them remotely either manually or automatically with data analysis result. The air device control functions can be optimized artificial intelligent to improve indoor air quality and save the energy cost like electricity usage.

2.2 Data Processing and Monitoring

We are collecting data from 4 different sensor packs. Sensor pack 1 measures indoor air quality before filtering air pollution. Sensor pack 2 measures indoor air quality after purifying air using purifier. Sensor pack 3 measures outside air quality information and sensor pack 4 also measures indoor air quality. Senor packs measure temperature, humidity, fine dust, carbon dioxide, voc and sends to our proposed system. Also, air filter device sends its running status, air volume, on/off status, error, and filter replacement information. Including real time data processing, data analysis, and visualization services of ACP cloud platform is designed as **Fig. 2**.

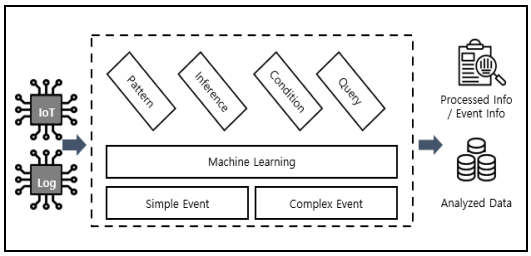


Fig. 2. Complexed Air Data Processing

The real time data processing part of proposed system is implemented on AWS cloud platform for experimental analysis and the benefits such as expandability, productivity, reliability, feasibility is gained using Cloud services. As the outcomes, ACP provides the dash-board service both with web and mobile applications (**Fig. 3**).

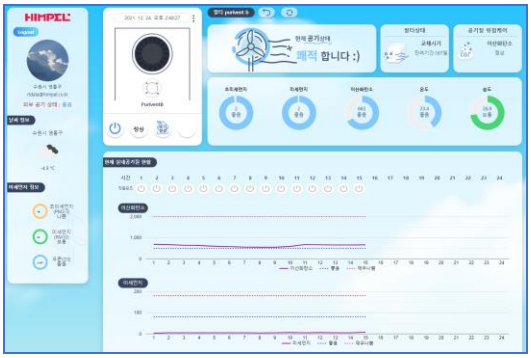


Fig. 3. Dash-board of ACP

Various modules are composed in the dash-board to show the current air-quality conditions, the trend of air-quality, the devices running conditions, and operating conditions. Additionally, the users can turn it on/off on online applications.

3. Conclusions

In this paper, we proposed the ACP to monitor and control the indoor air quality. The main purpose of this platform is to provide the integrated system for collecting IoT Sensor data, processing and predicting the air quality on cloud platform and visualizing the data analysis results via web and mobile application.

References

[1] Yang, Chao-Tung, et al. "Implementation of an intelligent indoor environmental monitoring and management system in cloud." *Future Generation Computer Systems*, Vol. 96, pp. 731-749, 2019.

[2] Luo, X. J., Oyedele, L. O., Ajayi, A. O., Monyei, C. G., Akinade, O. O., & Akanbi, L. A. "Development of an IoT-based big data platform for day-ahead prediction of building heating and cooling demands". *Advanced Engineering Informatics*, 41, 100926, 2019.

Copyright ⓒ 2023 KSII

Factorized 3D CNNs based Human Action Recognition

Seho Park and Taehyeon Kim*
Contents Convergence Research Center
Korea Electronics Technology Institute
Seoul, South Korea
[e-mail: sehopark@keti.re.kr, taehyeon.kim@keti.re.kr]
*Corresponding author: Taehyeon Kim

Abstract

In this paper, we introduce a novel approach to enhance the efficiency of action recognition by utilizing a factorized 3D convolutional neural network. Our proposed method aims to decrease the computational complexity and memory requirements of 3D convolutional neural networks while ensuring negligible performance degradation in recognizing human actions within video data. The proposed technique involves representing the 3D convolutional layer, which is designed to extract and represent spatiotemporal features from input videos, as a series of interconnected sub-convolutional layers using Tucker decomposition. Through experiments using a human action recognition benchmark, it was confirmed that the proposed method achieves a high level of computational and memory efficiency while maintaining a similar level of performance compared to existing approaches.

Keywords: 3D Convolutional Neural Networks, Human Action Recognition, Tensor Decomposition

1. Introduction

Human action recognition refers to the problem of classifying actions of individuals or multiple people depicted in video, based on user-defined class established beforehand. Neural network-based human action recognition research and development are primarily based on three methodologies. The first approach utilizes a 3D convolutional layer-based convolutional neural network to capture spatiotemporal features from videos [1]. The second approach employs pose estimation techniques to extract joints and key points of human objects in the video, and then uses neural networks capable of time-series analysis, such as recurrent neural networks [2]. The third approach assumes the pose estimation results as a graph and performs action recognition using graph neural networks [3]. Among these approaches, this paper focuses on the first method, which utilizes a 3D convolutional neural network (3D-CNN). This is because, compared to other methodologies, it does not require additional computer vision processes like pose estimation and allows for a single neural network to extract and process spatiotemporal features from videos. A video can be considered as a collection of image frames sequentially connected over time. In general, for videos without scene changes due to editing, there is a statistical dependency between adjacent frames when viewed with respect to a specific frame.

Convmolution operations in convolutional layers are defined as operations that re-define the relationship between adjacent data points in a specific Euclidean data space by utilizing a weighted sum and non-linear operations, where locality connectivity exists. In the case of 2D convolutional layers, this operation primarily refers to a technique for extracting features concerning the relationship between a specific pixel in a given image and its neighboring pixels.

This work was supported by Korea Evaluation Institute of Industrial Technology(KEIT) grant funded by the Korea government(MOTIE). (No 20009760. Development of human friendly multipurpose service robot and new market creation by applying human robot interaction design)

Therefore, a 3D convolutional layer is defined as a convolution operation that simultaneously considers not only the relationships between pixels within each frame but also the relationships between different frames. This makes it a suitable neural network structure for video analysis. However, the 3D convolutional layer is not a particularly suitable convolution operation for devices with limited computational power, such as edge devices. To address this issue, this paper proposes a technique that efficiently transforms 3D convolutional layers by utilizing Tucker decomposition, which is one of the high-dimensional principal component analysis methods.

2. Proposed Algorithm

The weight tensor of a 3D convolutional layer is composed of a 4-dimensional tensor $\mathcal{W} \in \mathbb{R}^{t \times d \times d \times c \times s}$, where t represents the number of video frames processed in the 3D convolutional layer, d denotes the spatial filter size, c corresponds to the number of channels in each frame, and s signifies the total number of 3D convolutional filters composing the layer. Therefore, the operation involves converting a video of n frames with a total of c channels into a collection of feature tensors, each consisting of n frames with s channels.

The proposed method aims to reduce the computational complexity that arises during the conversion from c channels to s channels by utilizing Tucker decomposition to decompose the weight tensor and transform it into two sub 3D convolutional layers. The first 3D convolutional layer is composed of a spatiotemporal size of $t \times d \times d$ and has a rank size based on the Tucker decomposition. The weight tensor for this layer is as $\mathcal{W}_1 \in \mathbb{R}^{t \times d \times d \times c \times r}$. Subsequently, a 3D feature expansion layer is utilized to restore the original channel values. At this point, the weight tensor of the convolutional layer has a size of $1 \times 1 \times 1 \times r \times s$, where r represents the rank of the Tucker decomposition. Therefore, the memory efficiency of the proposed method in this paper can be defined as $\frac{tddcr+rs}{tddcs}$, where the denominator representing the memory consumption of the original 3D convolutional layer and the numerator corresponding to the memory consumption of the factorized layers..

3. Experimental Result

We evaluated human action recognition performance using the UCF-101 dataset. The UCF-101 dataset is a widely used benchmark for human action recognition in videos. It consists of 13,320 video clips that span 101 action categories, such as walking, running, jumping, swimming, and various sports. These clips are collected from YouTube and have a diverse range of variations in terms of camera motion, pose, and illumination. The dataset is designed to evaluate the performance of algorithms for human action recognition and provides a challenging testbed for the development of new methods in this field. The experimental results show that the proposed method achieves approximately 30% higher memory efficiency compared to the original 3D CNNs, while attaining an accuracy of 87.5%. This corresponds to a negligible accuracy decrease of only 0.3% compared to the pre-compression performance.

4. Conclusion

This paper presents a 3D convolutional layer compression technique using tensor decomposition for human action recognition. The proposed method compresses filter dimensions rather than factorizing spatial and temporal information, enabling higher efficiency when combined with existing methods.

References

[1] Tran, Du, et al. “A closer look at spatiotemporal convolutions for action recognition,” *Proceedings of the IEEE conference on Computer Vision and Pattern Recognition*, 1-10, 2018

[2] Liu, Mengyuan, et al. “Enhanced skeleton visualization for view invariant human action recognition,” *Pattern Recognition*, 68, 346-362, 2017.

[3] Shi, Lei, et al. “Skeleton-based action recognition with directed graph neural networks,” *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 7912-7921, 2019.

Copyright ⓒ 2023 KSII

The Design of IT Project Community Service Platform for Exploring Human Resources Using Recommendation Algorithm

Yoonjung Jang[1], Yumin Jo[1], Hyebin Lee[2], Jiyeon Jeong[2], Jongho Paik[2]* and Jun Hwang[2]
[1] Department of Software Fusion, [2] Department of Software Convergence
Seoul Women's University
Seoul, South Korea
[e-mail:{dbs920313, rosemin97, lhb8106, lilyandcat, paikjh, hjin}@swu.ac.kr]
*Corresponding author: Jongho Paik

Abstract

Demand for jobs for developers in the IT industry is increasing, but the shortage of developer manpower is not being resolved. One of the main causes of the shortage of developer manpower in IT industry is the difficulty in hiring developers who fit the ability required by companies. In addition, most of the current employment platforms are operating in a way that companies post announcements and job seekers apply. This has the problem that the range of choices in hiring the capable appliance needed by the company is limited. Therefore, in this paper, we propose an IT project community mediation service that can be used by companies to search for human resources using a recommendation algorithm. The proposed service not only announces a project meeting, but also supports various search methods such as keyword search and technology stack search so that companies can hire the talent, they want based on the project information the user has participated in. In addition, it supports the function of identifying the talent who suitable for project and recommending it to project creators. Through this, we expect individuals will be able to manage portfolios through project experiences of interest, and companies will be able to recruit reliable team members suitable for the company.

Keywords: human resources, recommendation algorithm, service platform

1. Introduction

Despite the fact that IT industry has emerged as a field that attracts attention, the shortage of developer manpower that companies appeal to is serious. The Ministry of SMEs and Startups predicted that the number of developers lacking from 2021 to 2025 would be at least 30,000. They point out the cause of this shortage of developer manpower is mismatching of a job connection between companies and job seeker[1].

To solve this problem, more and more companies are choosing to outsource work by dividing it into tasks or hiring freelancers through online platforms such as Wishket[2].

However, the current freelance recruitment method, in which freelancers apply after seeing the company's announcement, has difficulty of solving the mismatching problem completely. First, if there are no applicants for the announcement, companies still have difficulties in recruiting. Second, among the applicants, there may be no applicants with the specifications desired by the company.

To solve these problems, in this paper, we

This research was supported by the Office of Software-Centered University, Seoul Women's University (2023)

propose a service platform that enables two-way recommendation for individual users and enterprise users. The proposed service aims to alleviate the mismatching problem by adding a contact function to enable active recruitment of companies.

2. Implementation of Service Platform

2.1 Environment of Service Platform Implementation

The proposed service platform is designed based on the MVC (Model-View-Controller) pattern. As DBMS (Database Management System) corresponding to the model, MySQL is used to support Full-Text Search (FTS). The server, corresponding to the Controller, we use Node.js. It is suitable for processing text/image data and operating database using JSON objects. Also, to perform convenient MySQL CRUD (Create, Read, Update, Delete) operations, the Sequelize library is used [3].

Since this service platform focuses on connectivity between users, Android was selected as the View. It uses the MVVM (Model-View-ViewModel) pattern that separates the model and the view through the ViewModel. This separates the application's business logic and presentation logic from the user interface (UI). Using the MVVM design pattern makes it easier to test, to maintain, and to evolve applications and enables code reuse[4].

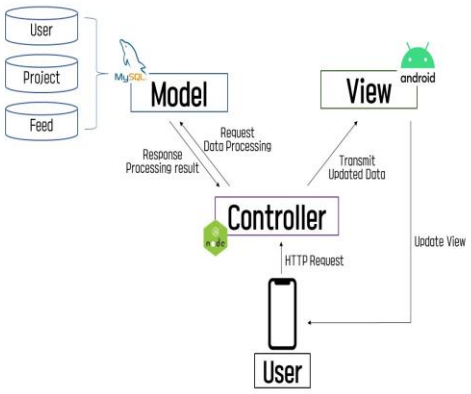


Fig. 1. MVC pattern of the proposed service platform

2.2 Main Features of the Service Platform

The proposed service platform supports not only the function of simply announcing a project for recruiting team members, but also a portfolio function through a feed service. The feed allows individual users to upload the projects they have participated in by adding descriptions of the technology stack they have used. Searching the feeds also makes it easy for enterprise users to find developers they want to hire. Service components and each related functions are shown in **Table 1**.

Table 1. Components of the service platform and related functions

Service Components	Related Functions	
Users	Individual	• Collect information such as available technologies and preferred project types. • Information is used for project recommendation algorithms.
	Enterprise	Recommend individual users based on the individual users' information.
Project	Information on the project, such as the project field and scale, setting conditions for applicants to be recruited, and managing applicants.	
Feed	Management of information about completed projects, such as technology used and participants.	
	Individual	use as a portfolio.
	Enterprise	use as a user navigation through searching.

2.2.1 Project Announcement and Recruitment of Participants

In this service platform, subject creating the project can be an enterprise user for hiring freelancers and individual user, so it is marked as the project creator. They create project including the information about the project and upload it to the service. When an individual user announces

for the purpose of strengthening individual capabilities, the project type is set as personal portfolio. When an enterprise user announces a project for the purpose of outsourcing or hiring freelances, the project type is set as corporate or assignment.

Additionally, as project information, project creators enter the project field and duration, project size, number of recruits, and detailed conditions for applicants. After checking project information, applicants apply to the project they want to support. The project creator can check the applicants' information and decide whether to approve or to reject the applicants as a team member. Fig. 2 shows this process.

If fewer people applied than it required for the project or there are no applicants satisfying the creator, the project creator can suggest to apply for the project to the desired user through a feed search. Details of offers are described in Section 2.2.3.

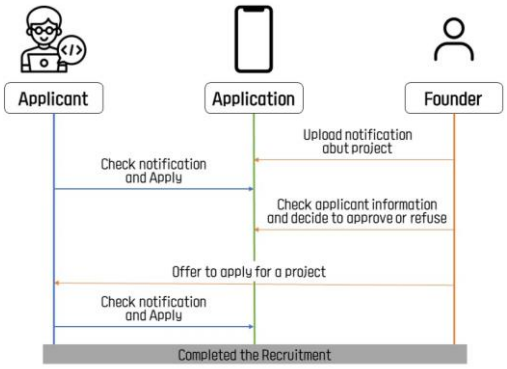


Fig. 2. The sequence diagram of announcement the project meeting and recruiting participant

2.2.2 Creating and Managing Feeds

Users can upload feeds as a result of projects they have participated in. The feed includes project images, service descriptions, participant information, and technologies used. In the detailed view of the feed, the screen is configured to easily check the participant information of the project. Project participants tagged in the feed can check the feed on the user's My Page. Therefore, the project creator can easily check the information about the project that applicant participated in at the one's My Page. The corresponding screen is shown in **Fig. 3**. This makes project creators easy to check if the applicant whom matches to the project member.



Fig. 3. The screen of checking projects ones has participated in on the user's My Page

2.2.3 Exploring Human Resources

The proposed service platform provides function that searching project, feed, and user. User searching is organized based on the user's position and available technology stack. For example, if someone want to find an Android developer who knows how to handle the JAVA language, researcher can select a position as developer in the search filter and enter "Java" or "JAVA" at the search box. Feed searching and project searching can be searched by selecting project type keyword as well as technology stack keywords. For example, if someone wants to find users who have worked on web or application projects, researchers can select "Web" and "Application" from the project type. Due to this function, enterprise users can easily find individual users with the capabilities and experiences they need. The corresponding screen is shown in Fig 4. If researcher touch the participant's profile, page move to the user's My Page, and enterprise users can offer to apply the project or recruitment to the user through the email address and chat room address listed on the My Page. Through this process, enterprise users can actively recruit who they want.



Fig. 4. The screen of participants on the feed detail view screen

3. Design of The Recommendation function of Projects and Applicants

The recommendation function provided by the proposed service is divided into two types according to the subject. First, project creators are recommended users for each project and position through this service. Parameters used for recommendation include applicants whose progress times do not overlap as much as possible, and who have previously participated in project similar to announced projects and have had good ratings and who have participated so far among projects corresponding to the field of interest set by the user. Project creators receive an automatically sorted list of applicants through this service. However, in the case of a cold start, applicants for this project are recommended among users who have received high ratings in other similar project announcements. In other words, when recommending applicants to the project founder, hybrid recommendation filtering, which combines user-based recommendation filtering and item-based recommendation, is used.

In addition, among the announced projects, the user can select the field of interest, position, and progress period. Projects are recommended considering average working hours per day and project size. Since each user prioritizes different items among the items used for recommendation, this method uses a method of sorting the values selected by the user directly from the top bar rather than automatic sorting. That is, when recommending a project to a user, item-based recommendation filtering is used.

4. Conclusions

In this paper, we proposed an IT project community service platform that can be used by companies to search for talent by combining the recruitment service, which was previously operated as a separate service, and the project meeting announcing service. The proposed service is structured to easily find users with similar experiences through project and feed search. This provides a forum for companies or individuals to directly gather project applicants, and at the same time, allows them to search for the desired talent. In addition to direct search, users are recommended through the recommendation function within the service that is suggested according to the subject. The project creator receives a more detailed recommendation by applying various parameters among the applicable applicants for each project and position. The user also directly selects a project that meets the conditions desired by the user within the parameters provided by this service from among the projects and receives a recommendation. Through this, it is expected that individuals will be able to experience the desired project and manage their portfolio through participation in various projects, and companies will be able to search for and hire talent suitable for the project they want to proceed.

References

[1] Ministry of SMEs and Startups, *"220223 Venture Startup Academy Support Project"*, 2022.
[2] Wishket, *"How are major IT service companies in Korea dealing with the lack of developer recruitment?"*, 2022.
[3] Hyunyoung Cho, *"Node.js Textbook"*, Vol 2, Gilbut, 2020.
[4] Microsoft official document, *"MVVM"*

Copyright ⓒ 2023 KSII

Avatar-Chatbot Making for High School-College Major Experience Program

Tai Hoon KIM and In Seop NA*
School of Elec. And Comp. Engineering, Division of Culture Contents, Chonnam National University
Yeosu, South Korea
[e-mail: taihoonn@jnu.ac.kr, ypencil@hanmail.net]
*Corresponding author: In Seop NA

Abstract

The High School-College Major Experience Program aims to increase high school students' understanding of university life by providing hands-on experience with college majors. In this paper, we developed an "Avatar Chatbot for Customer Support" and prepared a microphone, speaker, and camera. We obtained voice through the microphone, converted it into text using STT (Speech-to-Text) software, and processed it with NLP (Natural Language Processing) techniques. After NLP processing, we conducted voice recognition. Then, we played a pre-made avatar video that resembled me according to the relevant content among the customer support program. The video had generated lip movements based on the pre-designed scenarios, and the TTS (Text-to-Speech) was used to create voice, which was output to the screen and speaker for conversation. The avatar was created using a personal photo, and the video pronouncing according to the pre-made scenario was generated. We evaluated the participants' understanding of university majors through pre- and post-program surveys."

Keywords: Avatar, Chatbot, Customer Support, Natural Language Processing, TTS

1. Introduction

The High School-College Major Experience Program is a program designed to give high school students the opportunity to explore different college majors and gain a better understanding of what university life is like. Through this program, students are able to get hands-on experience with college majors by attending classes, participating in labs or workshops, and engaging with faculty and students in their chosen field of study. This allows them to gain a better understanding of the academic requirements, expectations, and career prospects associated with different majors. The program is aimed at increasing high school students' understanding of university life, as well as helping them make informed decisions about their academic and career paths. By providing students with the opportunity to explore different majors and experience college-level coursework, the program aims to prepare them for the challenges and opportunities of higher education.

In this paper, we developed an "Avatar Chatbot for Customer Support" and prepared a microphone, speaker, and camera. We obtained voice through the microphone, converted it into text using STT (Speech-to-Text) software, and processed it with NLP (Natural Language Processing) techniques. After NLP processing, we conducted voice recognition. Then, we played a pre-made avatar video that resembled me according to the relevant content among the customer support program. The video had generated lip movements based on the pre-designed scenarios, and the TTS (Text-to-Speech) was used to create voice, which was output to the screen and speaker for conversation. The avatar was created using a personal photo, and the video pronouncing according to the pre-made scenario was generated.

This study was financially supported by Chonnam National University(Grant number: 2023-0933)

2. Avatar Chatbot for Customer Support

The process involves developing an avatar chatbot that can interact with customers through a combination of text and audio. Voice input is captured and converted to text, which is then processed using NLP techniques. Voice recognition is then conducted to understand the customer's needs. As shown in **Fig. 1**, the avatar video is played with pre-made lip movements and voice generated using TTS, which is output to the screen and speaker for conversation. The avatar is created using a personal photo and is animated according to pre-made scenarios to provide customer support.

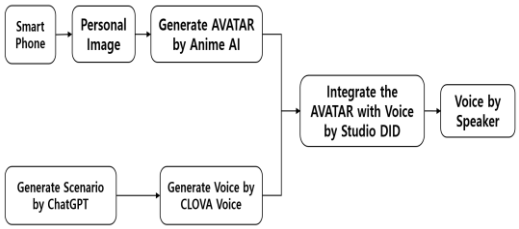


Fig. 1. The Flow-chart of Avatar-Chabot Generation

2.1 AVATAR Generation

The process of creating an avatar using Anime AI **[1]** involves the following steps:

1. *Downloading and installing Anime AI from the app store.*
2. *Once installed, you can use Anime AI to create an avatar by inputting a keyword or uploading a photo.*
3. *After inputting the keyword or uploading a photo, you can then select the desired art style from a range of available options.*
4. *Anime AI will then automatically transform the input image into an avatar based on the selected art style.*
5. *You can further customize the avatar by adjusting features such as the background, pose, and expression.*

The process of creating an avatar using Anime AI is relatively straightforward and involves using the app to transform a photo into an art-style avatar. The app uses advanced algorithms and machine learning to transform the image, providing a quick and easy way to create unique avatars with different styles. **Fig. 2** shows example of generated avatar by Anime AI.



Fig. 2. Example of Generated Avatar by Anime AI

2.2 Voice Generation and Lip-Sync with AVATAR

The scenario involves using ChatGPT **[2]** to generate responses, converting these responses into speech using CLOVA Voice **[3]**, and using Studio DID **[4]** to create avatars and lip-sync based on the speech.

1. *Generate appropriate responses to user inputs or requests using the ChatGPT model.*
2. *Use CLOVA Voice to convert the generated responses into speech.*
3. *Create avatars and lip-sync using Studio DID based on the speech, allowing the response to be presented on screen alongside the avatar.*
4. *The combination of the avatar and speech provides a natural conversational experience for the user.*

By using ChatGPT, CLOVA Voice, and Studio DID together, it is possible to create innovative and engaging interfaces that offer a more immersive experience for users. Overall, this process allows for the creation of avatars and lip-sync based on speech generated by ChatGPT, offering a more natural and immersive conversational experience for users.

3. Experimental Results

3.1 Model Design

On April 21, 2023, the High School-College Major Experience Program was conducted with 24 first-year students from Seoncheon Maesan High School who had previously learned the

Python language as part of their regular high school curriculum and had basic prior knowledge of Python programming. The program was designed to provide hands-on experience with college majors and was organized by the department. The program's effectiveness was evaluated through pre- and post-program surveys, which were used to collect and analyze information on student satisfaction, as well as suggestions for future program improvements and operational direction.

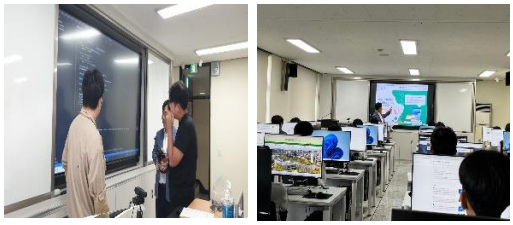


Fig. 3. Activity of High School-College Major Experience Program

3.2 Analysis of Effectiveness and Satisfaction

Regarding the satisfaction of participation (time, location, content), survey results showed that participants were highly satisfied (66.7%, 75%, 75%) and satisfied (33.3%, 25%, 25%) with the satisfaction level regarding the participation experience (time, location, content). Overall program management was rated as very satisfactory (70.8%) and satisfactory (29.2%) with all participating students being satisfied.

The number of students who hope to attend Chonnam National University increased after participating in the program. Specifically, the number of students who answered "very much" increased from 2 (8.3%) to 4 (16.7%), "yes" increased from 10 (41.7%) to 12 (50%), "average" remained the same with 5 students (20.8%), and "no" decreased from 7 (29.2%) to 3 (12.5%).

The program was well-received by the participants, with 41.7% answering "very much", 41.7% answering "yes", 12.5% answering "average", and 4.2% answering "no" when asked if they would recommend the program to others.

The main reasons for recommending the program were that the professor was kind and interesting in programming, that the program provided hands-on practice that made programming and AI more approachable, that the class was easy to understand and enjoyable even for those with no prior experience, that participants could create new videos using various AI applications related to their daily lives, and that the program was useful for the future.

4. Conclusions

Based on the analysis of the survey results, it can be concluded that the Avatar-Chatbot Making Program can be a successful high school-college major experience program. The program received positive feedback from the participating students, who found it interesting, informative, and engaging. The practical hands-on approach of the program and the exciting topic of AI and programming make it an ideal choice for high school students who are interested in pursuing a career in this field.

The program can also be recommended to high school students who may not have a strong background in programming or AI. The user-friendly and easy-to-understand course content and fun practical exercises make the program accessible to all students.

Overall, the Avatar-Chatbot Making Program can be a valuable addition to the high school-college major experience programs. It provides an opportunity for high school students to gain exposure to the field of AI and programming and develop an interest in pursuing it as a career path.

References

[1] Anime AI, https://animeai.app/
[2] ChatGPT, https://openai.com/blog/chatgpt
[3] CLOVA Voice, https://clova.ai/voice/
[4] Studio D-ID, https://studio.d-id.com/

Copyright ⓒ 2023 KSII

A Deep Learning Method for IQ Fingerprint Map Based Indoor Positioning using Bluetooth 5.1 Technology in Ship Environments

Yootae Shin[1], Jooyoung Son[2]*, and Qianfeng Lin[3]
[1]Devision of Marine Information Technology, Korea Maritime and Ocean University
Busan, Republic of Korea
[e-mail: sytvip1@gmail.com]
[2]Division of Marine IT Engineering, Korea Maritime and Ocean University
Busan, Republic of Korea
[e-mail: mmlab@kmou.ac.kr]
[3]Department of Computer Engineering, Korea Maritime and Ocean University
Busan, Republic of Korea
[e-mail: linqianfeng@g.kmou.ac.kr]
*Corresponding author: Jooyoung Son

Abstract

The importance of indoor positioning has grown in numerous application areas such as emergency response, logistics, and industrial automation. In ships, indoor positioning is also needed to provide services to passengers on board. Due to the complex structure and dynamic nature of ship environments, conventional positioning techniques have limitations in providing accurate position. Compared to other indoor positioning technologies, Bluetooth 5.1-based indoor positioning technology is highly suitable for ship environments. Bluetooth 5.1 attains centimeter-level positioning accuracy by gathering IQ samples from wireless signals. However, distorted IQ samples can lead to significant errors in the final estimated position using deep learning. Therefore, we propose a deep learning method based on IQ fingerprint map to overcome these challenges and achieve accurate positioning.

Keywords: Deep learning, IQ fingerprint map, Bluetooth 5.1, ship environments

1. Introduction

Indoor positioning has become increasingly important in various application domains, including emergency response, logistics, and industrial automation. In ship environments, Bluetooth offers one advantage over RFID, WIFI indoor positioning technologies. Bluetooth provides low energy consumption makes BLE (Bluetooth Low Energy) Tag suitable for long-term deployment in ship environments. BLE tags can be attached to user devices, such as smartphones, to provide accurate indoor positioning [1].

In wireless communication systems, the signal is often represented as a complex number in the form of in-phase (I) and quadrature (Q) components, collectively known as IQ samples. Bluetooth 5.1 achieves centimeter-level positioning by collecting IQ samples of wireless signals. In Bluetooth 5.1 technology, IQ samples play a crucial role in accurately estimating the Angle of Arrival (AoA) [2]. However, the complex structure and dynamic nature of ship environments pose significant challenges to existing indoor positioning techniques. Due to

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2021R1I1A3056125).

the metallic environment of ships, multipath effects may occur. Multipath effects refer to the phenomenon where wireless signals are transmitted outward and produce multiple copies upon encountering walls. Distorted IQ samples can lead to significant errors in the final estimated position. Specifically, during a ship's voyage. This paper aims to address these challenges by proposing a deep learning for IQ fingerprint map based on indoor positioning using Bluetooth 5.1 technology in ship environments.

2. Related Works

Many indoor positioning techniques have been proposed, including Wi-Fi fingerprinting, ultra-wideband (UWB), and radio-frequency identification (RFID). While these methods have shown success in some environments, their performances in ship environments are limited due to the complex structure and dynamic nature of these spaces. Bluetooth 5.1 technology has been proposed as a promising solution for ship environments due to its low energy consumption, high accuracy, and robustness against environmental changes. The impact of metallic environments on IQ samples has been a topic of significant interest in the field of wireless communications and indoor positioning. The presence of metallic structures in ship environments can result in various challenges, such as multipath propagation, which may lead to distortions in IQ samples [3].

3. Deep Learning for Bluetooth 5.1 Indoor Positioning with IQ Fingerprint Map

In deep learning, optimization algorithms play a crucial role in determining the performance of Deep Neural Network (DNN). Adam, RMSprop, SGD, Adagrad, and Adamax are some of the popular optimization algorithms used in training DNN [4]. Here, we discuss the impact of these algorithms on the performance of DNN and explain why Stochastic Gradient Descent (SGD) often yields the best results.

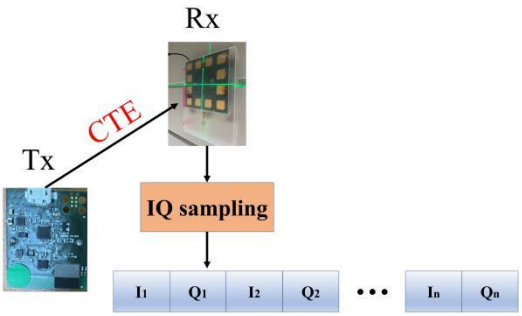


Fig. 1. IQ samples from a BLE tag (Tx) received by an AoA Locator (Rx)

In ship environments, dealing with distorted IQ samples can be challenging. In this paper, a well-trained DNN can be employed to classify the distorted IQ samples effectively. SGD can help DNNs achieve desirable performance. SGD is a simple yet efficient optimization algorithm that computes the gradient of the loss function with respect to each weight by using a small subset of the training data.

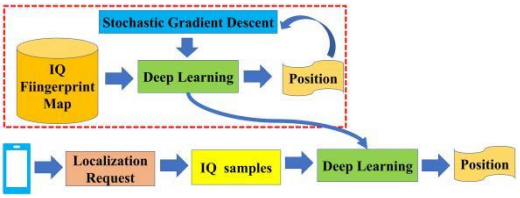


Fig. 2. Workflow of deep learning with IQ fingerprint map for estimating position

4. Performance Evaluation

We conduct an indoor positioning experiment to collect IQ samples from a test area situated on the upper deck of HANNARA ship used by students. The test area measures 7 meters by 7.14 meters. Within test area, we design 12 positions as reference points, as shown in Fig 3.

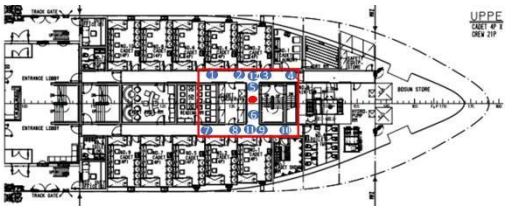


Fig. 3. Reference points deployment diagram in upper deck of HANNARA ship

Table 1. Best Cases of Optimizers

Optimizer	Learning rate	Batch Size	Neuron Size	Stop Epochs	Accuracy
SGD	0.1	32	[960, 480, 120]	100	0.977
Adagrad	0.1	64	[960, 480, 120]	89	0.975
RMSprop	0.01	128	[480, 240, 60]	86	0.966
Adam	0.01	128	[480, 240, 60]	75	0.944
Adamax	0.01	32	[480, 240, 60]	93	0.974

SGD showed the highest accuracy compared to other optimizers, reaching 0.977. Its simplicity also contributes to its stable performance. The best hyperparameters for SGD were observed at a high neuron size, a learning rate of around 0.1, and a low batch size. Due to the simplicity of SGD, a lower batch size allows for more detailed steps with each epoch, leading to increased accuracy. On the other hand, SGD cannot control its step size, and a low learning rate can result in large steps (causing the callback function to stop training due to high loss) and excessive resource consumption.

We collect IQ samples at 12 reference points and create an IQ fingerprint map. The IQ fingerprint map consists of IQ samples with location indices. We use the IQ fingerprint map as the training set, while the other IQ samples collected at reference points serve as the test set. We input the training set into the DNN and then evaluate the model's performance on the test set. By testing various batch sizes, learning rates, and layer sizes, we eventually obtain the best result. We find that when the batch size is 32, the learning rate is 0.1, and the layer sizes are [960, 480, 120], the test set's accuracy reaches its highest at 97.76%.

5. Conclusions

We evaluate the proposed method in a real ship environment. Firstly, we collect IQ samples at 12 reference points. Secondly, we use the collected IQ samples as the training set and the other collected IQ samples as the test set. Finally, we input the training set into the DNN and evaluate the training performance on the test set. We find that when the batch size is 32, the learning rate is 0.1, and the layer sizes are [960, 480, 120], the test set's accuracy reaches its highest at 97.76%.

Acknowledgement

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (grant number: 2021R1I1A3056125).

References

[1] Zafari F, Gkelias A, Leung K. A survey of indoor localization systems and technologies[J]. IEEE Communications Surveys & Tutorials, 2019, 21(3): 2568-2599.

[2] Lin Q, Son J, Shin H. A self-learning mean optimization filter to improve Bluetooth 5.1 AoA indoor positioning accuracy for ship environments[J]. Journal of King Saud University-Computer and Information Sciences, 2023, 35(3): 59-73.

[3] Sambu P, Won M. An experimental study on direction finding of Bluetooth 5.1: Indoor vs outdoor[C]//2022 IEEE Wireless Communications and Networking Conference (WCNC). IEEE, 2022: 1934-1939.

[4] Kingma D P, Ba J. Adam: A method for stochastic optimization[J]. arXiv preprint arXiv:1412.6980, 2014.

Copyright ⓒ 2023 KSII

An Empirical Study on the Length Granularity in Long-Form Question Answering

Hyeonseok Moon, Jinsung Kim, Kinam Park, Jeongbae Park, Heuiseok Lim*
Dept. of Computer Science and Engineering, Korea University
Seoul, South Korea
[e-mail: glee889@korea.ac.kr, jin62304@korea.ac.kr, superkn@naver.com, insmile@korea.ac.kr, limhseok@korea.ac.kr]
*Corresponding author: Heuiseok Lim

Abstract

Long-form question answering (LFQA) task aims to generate a descriptive answer for a given question based on a long-form source context such as Wikipedia. Existing studies generate appropriate answers based on retrieval of passages related to questions. Because it works in this way, processing long sequence inputs with passages and generating longer-length answers to make them easier to understand are two important factors in the LFQA task. However, recent work figured out that LFQA model does not adequately understand long input document. We interpret this to be due to information sparsity at the retrieval stage and at the input construction stage. To verify this, we introduce three retrieval methods with different granularity. This means the retrieval method of passage, sentence, and phrase units, and we verify the suitability and effectiveness of the LFQA task by verifying the performance of the LFQA model to which each methodology is applied. Eventually, we discuss whether the method of adjusting the intensity of information by adjusting the granularity differently is the appropriate method for LFQA.

Keywords: Artificial Intelligence, Information Retrieval, Long Form Question Answering, Natural Language Processing, Question Answering

1. Introduction

Long-form question answering (LFQA) task aims to generate a descriptive answer for a given question based on a long-form source context such as Wikipedia. The dataset used in this task is based on the long-form explanation information for Reddit platform. The purpose of the subreddit for this dataset is to explain like the questioner is five-years-old. As an open-domain question that requires a detailed explanation is given, LFQA task has two typical characteristics over the length and diversity; one is that it has answers that span multiple sentences and the other is that it has contents from knowledge sources required to answer the questions. Through achieving to handle this task effectively, it is expected to expand the capability to deal with long-form textual data.

However, recent works figured out the limitations of the LFQA task. Representatively, [1] proposed three hurdles that make challenging in proceeding LFQA task: answers are not grounded in retrieval, train/valid/test data overlapping, and naive performance of ROUGE evaluation metric in gauging the actual performance of the LFQA models.

We focus on the first challenge pointed out above. In particular, the reason that a model for LFQA task does not generate the answer grounded in the input document is derived from the excessively long length of the input sequence and output sequence. We attribute this

phenomenon to the study that routing transformer certainly understands long context but has limitations on understanding excessively long sequences. Moreover, we also hypothesize that a decoder does not generate long sequences by concentrating on significant information in the given context [2].

To alleviate this challenge regarding understanding the excessively long contexts and generating entirely explanatory answers, we prove the effectiveness of exploring in terms of the following three granularities, i.e., passage, sentence, and phrase level. Through these experiments, we demonstrate the effective document retrieval method in terms of retrieval unit granularity and the proper method for the answer generation in the retrieved documents.

2. Proposed Method

2.1 Dense Passage Representation

The DPR[3] is a paragraph-level information retrieval model which specializes in open-domain question answering. It uses the 2018 Wikipedia dump for the retrieval pool and split each document into a set of paragraphs. Each paragraph P consists of multiple text blocks of 100 words. We use the retriever pool provided in the KILT [4] and fed the question Q to the DPR to get the top-K relevant paragraphs P.

2.2 Dense Contextual Sentence Representation

The [5] is a sentence-level information retrieval model. DCSR uses the same encoder structure which was proposed in DPR but it retrieves sentences by adding <sent> token to the head and the tail for each sentence, then gets those special token's representation for the dot product between question representation. By fedding Q to the DCSR, we first get top-K relevant paragraphs P with respect to the maximum dot product score of the sentence within each passage. We also get the top 200 relevant sentences for each question to compare the effectiveness of the retrieval unit to paragraph-level and phrase-level in the LFQA task.

2.3 Dense Phrase Representation

For the phrase-level information retrieval model, we use DensePhrases [6]. DensePhrases train phrase encoder, which is different from the DPR and DCSR. The phrase is a contiguous text segment up to 20 words. To get the phrase representation, each phrase has <p_start> and <p_end> special token to get the phrase representation. In spite of the phrase retrieval unit, DensePhrase can retrieve top-K paragraphs using maximum phrase representation score within each paragraph. Following this setting, we retrieve the top-K paragraphs P

2.4 Long Form Answer Generation

In the answer generation phase, we followed the previous approaches, in particular, we fine-tuned an encoder-decoder architecture, i.e., BART model [7] by constructing the input sequences according to the granularity. In detail, the passage-level model used a set of retrieved passages, the sentence-level model used a set of retrieved sentences, and the phrase-level model used a set of retrieved phrases as the input sequence. The number of components that consists of each set is different, and they are limited by max sequence length. Therefore, it is expected that the generated answer will be different when each input sequence is fed in a different form in terms of granularity.

3. Experiment

3.1 Experimental Setting

Table 1. Dataset details

		Train	Valid	Test
# of Data		271,304	1,507	1,507
Query	**Avg** length	76.319	76.879	93.447
	Max length	299	283	292
	Min length	0	5	7
Answer	**Avg** length	1766.799	1660.636	6781.946
	Max length	219,504	48,213	84,557
	Min length	30	39	317

Table. 2. Performance comparison between different retrieval strategies. All the retrieved information are in the passage formats

	Max			Mean		
	ROUGE-1	**ROUGE-2**	**ROUGE-L**	**ROUGE-1**	**ROUGE-2**	**ROUGE-L**
DPR	0.3435	0.0884	0.1872	0.2737	0.0548	0.1401
DSCR	0.3388	0.0883	0.1879	0.2687	0.0536	0.1396
DensePhrase	0.3374	0.0867	0.1852	0.2671	0.0526	0.1382

Table. 3. Performance comparison between different retrieval units. Retrieval strategy is fixed to DensePhrase, and we differ only its granularity in generating input sequence

	Max			Mean		
	ROUGE-1	**ROUGE-2**	**ROUGE-L**	**ROUGE-1**	**ROUGE-2**	**ROUGE-L**
Passage	0.3374	0.0867	0.1852	0.2671	0.0526	0.1382
Sentence	0.3376	0.0856	0.1838	0.2796	0.0552	0.1434
Phrase	0.3374	0.0867	0.1852	0.2712	0.0538	0.1405

The corpus used in the experiment is Explain Like I'm Five (ELI5) dataset [8]. This dataset was constructed based on the Reddit ELI5 forum up to July 2018. The question-answer pairs were selected according to the number of votes, i.e., only the QA pairs with at least two more up-votes than down-votes, resulting in the 272K questions. The number of sentences per passage is 6.6, and that of words is 130 on average. The supporting documents are constructed through CommonCrawl dataset as this is because Wikidata alone was not enough to build data, by utilizing TF-IDF with query method crawling 100 most relevant web sources for each question excluding reddit.

As a baseline model, we adopt the vanilla BART-base model. The model is based on an encoder-decoder architecture.

The hyperparameters for retrieval are as follows: 256 for batch size, 256 for sequence length, 0.1 for dropout rate, and Adam optimizer is used, and for answer generation are 512 for sequence length.

3.2 Main Results

To compare the effectiveness of the retrieval unit in the LFQA task, we fix the retrieval output unit to paragraph-level but use different retrieval models. The result is shown in Table 2. We measure the mean and max ROUGE score of the predicted answer with reference answers among all retrieved cases for each question. The DPR showed the best performance in all cases. In case of the DSCR and DensePhrases, the difference between the two models was marginal, which showed poor performance compared to the DPR. To sum up, our assumption didn't fit when it comes to the retrieval unit granularity in the LFQA task.

We conduct an additional experiment on decoder input granularity. For this experiment, we only use the DensePhrases for the retrieval model and get 3 retrieval outputs with respect to different retrieval units. The result is shown in Table 3. In this experiment, all the performances of the retrieval unit for the different retrieval outputs showed similar results. We assume this result is related to the maximum input length of BART model [7], which limits the full capability of passage retrieval output.

4. Conclusions

Through this study, we verified the effects of information sparsity in the input construction stage and information sparsity in the retrieval stage in the LFQA task. However, contrary to our assumption, granularity did not have a significant effect. We interpret the reason for the slight difference in performance between the granularity methodologies as that the retrieval pool, Wikipedia, is not informative. This is based

on the assumption that Wikipedia itself, which contains descriptive information, will not provide enough information since the LFQA task focuses on generating easily interpretable answers. As a future study, we plan an LFQA study with different retrieval pools.

Acknowledgement

This work was supported by Institute of Information & communications Technology Planning & Evaluation(IITP) grant funded by the Korea government(MSIT) (No. 2020-0-00368, A Neural-Symbolic Model for Knowledge Acquisition and Inference Techniques) and was supported by Basic Science Research Program through the National Researc/ h Foundation of Korea (NRF) funded by the Ministry of Education (NRF-2022R1A2C1007616). This research was supported by the Core Research Institute Basic Science Research Program through the National Research Foundation of Korea(NRF) funded by the Ministry of Education(NRF-2021R1A6A1A03045425)

References

[1] Kalpesh Krishna, Aurko Roy, and Mohit Iyyer. Hurdles to progress in long-form question answering. arXiv preprint arXiv:2103.06332, 2021.

[2] Simeng Sun, Kalpesh Krishna, Andrew Mattarella-Micke, and Mohit Iyyer, Do long-range language models actually use long-range context? arXiv preprint arXiv:2109.09115, 2021.

[3] Vladimir Karpukhin, Barlas Oguz, Sewon Min, Patrick Lewis, Ledell Wu, Sergey Edunov, Danqi Chen, and Wen-tau Yih, “Dense passage retrieval for open-domain question answering,” In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pages 6769–6781, 2020.

[4] Fabio Petroni, Aleksandra Piktus, Angela Fan, Patrick Lewis, Majid Yazdani, Nicola De Cao, James Thorne, Yacine Jernite, Vladimir Karpukhin, Jean Maillard, et al., “Kilt: a benchmark for knowledge in-tensive language tasks,” In Proceedings of the 2021 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pages 2523–2544, 2021.

[5] Wu Hong, Zhuosheng Zhang, Jinyuan Wang, and Hai Zhao, “Sentence-aware contrastive learning for open-domain passage retrieval,” In Proceedings of the 60th Annual Meeting of the Association for Computational Linguistics, Volume 1: Long Papers, pages 1062–1074, 2022.

[6] Jinhyuk Lee, Mujeen Sung, Jaewoo Kang, and Danqi Chen, “Learning dense representations of phrases at scale,” In Proceedings of the 59th Annual Meeting of the Association for Computational Linguistics and the 11th International Joint Conference on Natural Language Processing, Volume 1: Long Papers, pages 6634–6647, 2021.

[7] Mike Lewis, Yinhan Liu, Naman Goyal, Marjan Ghazvininejad, Abdelrahman Mohamed, Omer Levy, Veselin Stoyanov, and Luke Zettlemoyer, “Bart: Denoising sequence-to-sequence pre-training for natural language generation, translation, and comprehension,” In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 7871–7880, 2020.

[8] Angela Fan, Yacine Jernite, Ethan Perez, David Grang-ier, Jason Weston, and Michael Auli, “Eli5: Long form question answering,” In Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, pages 3558–3567, 2019.

Copyright ⓒ 2023 KSII

A Vision-based Unmanned Store System Fusing User Behavior Patterns

Yeongjun Choi, Harin Jo, Jisoo Cheon, Yunkyung Hwang and Junho Ahn*

Department of AI Computer Science and Engineering, Kyonggi University
Suwon, South Korea
[e-mail: dudwnszero99@gmail.com, gkflsgg@kyonggi.ac.kr, noori4543@gmail.com, yunnn426@kyonggi.ac.kr, jha@kyonggi.ac.kr]
*Corresponding author: Junho Ahn

Abstract

Due to the outbreak of COVID-19 in 2019, the number of unmanned stores such as “Emart-24” and “Amazon Go” has been continuously increasing recently. These unmanned stores use various sensors for stable operation and detect the location and movement of users in the store without any special manager. However, the artificial intelligence (AI) cameras currently used in unmanned stores are limited in technology, thus requiring a large number of sensors and cameras, which increases the cost burden on store operations. To solve this problem, this paper proposes an AI system that uses AI cameras to detect movements by analyzing user behavior patterns in various ways, and also reduces the number of sensors that need to be used in the store, which helps to reduce the cost burden on operation.

Keywords: Unmanned store, Artificial intelligence, AI cameras, Behavior patterns, cost

1. Introduction

Recently, as the COVID-19 pandemic has spread around the world, Amazon has opened “Amazon Go”, an unmanned offline grocery store. By using various artificial intelligence (AI) technologies and sensors to understand the behavior of objects, Amazon Go helps consumers check out without human assistance and prevents theft.[1] In South Korea, “Emart-24” has been established as the South Korean version of Amazon Go. Emart-24 also uses various sensors such as LiDAR sensors and AI cameras to track customer behavior and weight changes of shelves.[2] However, despite these efforts, unmanned stores are still vulnerable to theft compared to manned stores.[3] Furthermore, there is a large burden in terms of cost to manage stores with numerous sensors and AI cameras, such as Amazon Go and Emart-24. This paper proposes a system that helps prevent incidents and operate unmanned stores at a low cost by applying AI technology to cameras instead of using many sensors to detect human behavior patterns.

2. Research on AI-based Unmanned Management System

2.1 Unmanned Stores of Amazon Go and Emart-24

Amazon Go is a chain of unmanned offline grocery stores operated by Amazon and provides consumers with a new shopping experience through its "Just Walk Out" technology.

When a customer enters the store with the recognition of a QR code registered on the Amazon app, multiple cameras on the ceiling recognize the object and track its behavioral information until the moment it leaves the store. When the customer picks up a product item, the sensor on the shelf detects the change in weight and places the item in a virtual cart. Afterward, when the customer leaves the store, the items in the virtual cart are paid for using the payment

information on file with Amazon, without any checkout process.[4]



Fig. 1. 'Just Walk Out' System

"Emart-24 Self Store" is a chain of unmanned convenience stores operated by adopting the technology of Amazon Go. Like Amazon Go, it uses various sensors, such as LiDAR (Light Detection and Ranging) sensors, and computer vision technology to perform AI-based video analysis.

To operate the "Just Walk Out" technology applied to these large supermarkets, it requires sensor and camera technology for not only simple CCTV video data but also consumer identification and behavioral analysis. This can be costly as it requires dozens of sensors to implement, making it difficult for smaller retailers to adopt as a drawback.

2.2 Design for User Behavior Pattern Detection

This paper proposes the following AI model to reduce the cost burden of unmanned store operations.

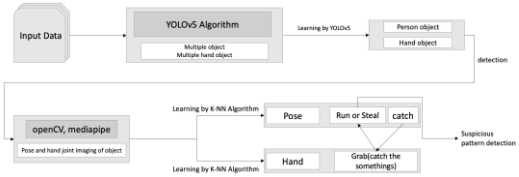


Fig. 2. Artificial intelligence recognition architecture for unmanned store systems

As shown in **Fig. 2**, when image data is input through CCTV, the YOLOv5 algorithm is used to recognize objects in the image and classify them into human objects and hand objects. For the classified objects, OpenCV and MediaPipe are used to predict the pose of each object and the joint coordinates of the hands in the image, based on which the movement of the object is predicted and then determined as "catch", "grab", "run", or "steal". Learning about the behavior patterns of objects is learned using the K-NN algorithm. If the movement is determined to be "run" or "steal" after "grab", the movement is determined to be a suspicious behavior and the scene is saved as a photo. If this AI model is used, the number of sensors required in the unmanned store will be greatly reduced, and AI cameras will take over those empty spaces, thus reducing the cost burden on store operation.

2.3 AI Detection Based on User Behavior Patterns

In this paper, a study is conducted to detect objects in the store and analyze user behavior patterns based on the aforementioned design.

First, the YOLOv5 algorithm was used to detect the user's object and the user's hand objects in the store. This algorithm tracks objects in real-time and preserves their uniqueness so that they can be easily distinguished, enabling the system to distinguish and detect objects even in spaces with a high density of users, such as unattended stores.[5]

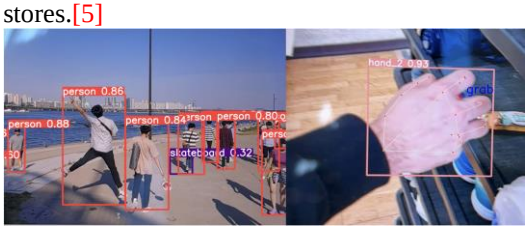


Fig. 3. Vision-based multi-object detections

After detecting the presence of each object, their behavior patterns are detected. The OpenCV standard library is used to facilitate the image processing function, and the MediaPipe' algorithm is used to learn specific behavior patterns and then place all joint coordinates of the user object and the joint coordinates of the hand objects on the image. In particular, the MediaPipe algorithm facilitates the calculation of the angle for each joint coordinate, thus providing more detailed detection, which is useful for distinguishing the specific behavior of the object.[6] The learned behavior patterns are divided into three types: catch, grab, and steal. When these patterns are detected, the "catch", "grab", and "steal" labels are assigned, respectively. If the first two behaviors are detected, the stealing behavior is considered last to finally determine the criminal behavior of the object.



Fig. 4. Catching behavior detection using Mediapipe



Fig. 5. Grabbing behavior detection using Mediapipe



Fig. 6. Stealing behavior detection using Mediapipe

3. Conclusions

Currently, limited AI technology is used in unmanned stores, because of which a large number of sensors are used. This can lead to a cost burden on store operations. As a remedy for this problem, this paper uses algorithms and sources such as YOLOv5 and MediaPipe to enable multi-object detection and movement detection using joint coordinates to analyze more diverse behavior patterns. This can reduce the use of high-cost sensors such as LiDAR sensors, which in turn can reduce the cost burden on store operations. Therefore, this paper will supplement the contents of this study by detecting more diverse behaviors that can appear in stores and studying image analysis technologies that can supplement them.

Acknowledgement

This research was supported by the MSIT(Ministry of Science and ICT), Korea, under the National Program for Excellence in SW supervised by the IITP(Institute of Information & communications Technology Planning & Evaluation) (2021-0-01393)

References

[1] The New York Times, 2018, https://www.nytimes.com/2018/01/21/technology/inside-amazon-go-a-store-of-the-future.html
[2] Business Post, 2022, https://www.businesspost.co.kr/BP?command=article_view&num=294354
[3] The JoongAng, 2018, https://www.joongang.co.kr/article/22442595#home
[4] Youtube, 2016, https://www.youtube.com/watch?v=NrmMk1Myrxc
[5] Taeju Park, Teukseop Song, “A Method of Companion Animal Breed Classification Using Deep Learning YOLOv5 Object Recognition”, Journal of the Korea Institute of Information and Communication Engineering, Vol 27, No. 4, pp. 565-571, 2023.
[6] Rakbin Song, Yuna Hong, Noyoon Kwak, “User Interface Using Hand Gesture Recognition Based on MediaPipe Hands Model”, Journal of Korea Multimedia Society, Vol 26, No. 2, pp. 103-115, 2023.

Copyright ⓒ 2023 KSII

Solving University Course Scheduling Problems Using Adaptive Parallel Tabu Search and Simulated Annealing

Xiaorui Shao[1] and Chang Soo Kim[2*]

[1] Industrial Science Technology Research Center, Pukyong National University, Busan 608737, Korea

[2] Information Systems, Pukyong National University,
Busan, 608737 South Korea
[e-mail: shaoxiaorui@pukyong.ac.kr, cskim@pknu.ac.kr]

*Corresponding author: Chang Soo Kim

Abstract

This paper presents a novel and effective approach based on the Tabu search and simulated annealing algorithm for solving the university course scheduling problem (USCP). Firstly, the initial solution of the UCSP instance is generated by one construction heuristic algorithm, the first fit algorithm. Secondly, we defined one union move selector to control the moves to provide diverse solutions, which are combined with two changing move selectors. Tabu search and simulated annealing (SA) are used to filter out unacceptable moves in a parallel mode. Besides, the acceptable moves are selected by one adaptive decision algorithm, which is used as the next step to construct the final solving path. Benefits from the good design of the union move selector, parallel tabu search and SA, and adaptive decision algorithm, the proposed method could effectively solve UCSP. We designed and tested our proposed algorithm in one real UCSP instance. The experimental results confirmed its effectiveness. Moreover, the in-depth analysis confirmed each component's effectiveness.

Keywords: Timetabling, scheduling, metaheuristic algorithm, university course scheduling problem.

1. Introduction

USCP aims to arrange teaching resources consist of professors, students, and rooms within a fixed timeslot to satisfy serval constraints [1], which has been proved to be an NP-complete problem [2]. These constraints could be divided into hard and soft constraints. The hard constraints cannot break out during solving and we should satisfy soft constraints at most to attach optimization.

The current methods for solving UCSP mainly contains two branches. One is the exact approach, which aims at finding the best solution through modeling one complex mathematical function. The most used exact approaches for UCSP is integer linear programming (ILP) [3]. Another one is metaheuristic approaches such as genetic algorithm (GA) [4], Tabu search [5], simulated annealing (SA) [6], etc. However, ILP is only suitable for easy UCSP instances and is very time-consumption, which cannot satisfy modern universities' needs.

Although current metaheuristic approaches obtained acceptable solutions, they mainly use one single algorithm to select the next step in the solution, which limits its performance. Moreover, they lack one comprehensive analysis to demonstrate their effectiveness. Therefore, this paper proposes a new approach that combined Tabu search and SA to select the most suitable moves for constructing a solution path. Besides, the in-depth analysis showed the proposed method could satisfy all hard constraints and most soft constraints.

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF), funded by the Ministry of Education(2021R1I1A3052605).

2. Proposed Method

2.1 UCSP

UCSP is to arrange teaching materials to a set of rooms and timeslots, which requires meeting all hard constraints and most soft constraints. We utilized a set of problem cases to describe our proposed method. Firstly, five hard constraints are given:

H1: One room can accommodate at most one course at the same time.

H2: The room should meet the course's functional requirements.

H3: The room size must be equal to or larger than the applied student's numbers.

H4: One professor can teach at most one course at the same time.

H5: One student only can attend one course at most at the same time.

Secondly, the five soft constraints are listed as follows:

S1: The professor prefers to teach in one fixed room.

S2: The professor has his preferable time duration.

S3: The professor prefers to teach in a sequential mode.

S4: The student dislikes studying in a sequence mode.

S5: The course and room should be consistent with its department.

Accordingly, we defined serval optimizing objects, as **Table 1** shown. We did not give all contents for each object because of page limitation. Notice that the type of Fact indicates that it cannot change during solving.

Table 1. Defined objects

Object	Type	Content
Room	Fact	$\{\text{id}, \text{name}, \text{room}_{\text{type}} \ldots \text{size}\}$
Timeslot	Fact	$\{\text{id}, \text{dayofweek} \ldots \text{end}_{\text{time}}\}$
Course	Entity	$\{\text{id}, \text{subject} \ldots \text{room}\}$
Timetable	Decision	$\{\text{room}_{\text{list}}, \text{course}_{\text{list}}, \text{score}\}$

2.2 The Proposed Method

The structure of the proposed method is given in **Fig. 1**. It consists of five steps: UCSP instance construction, giving initial solution, generating possible moves, Tabu+SA filters out unacceptable moves and selects the next move using an adaptive decision algorithm, respectively.

Step 1: Define the UCSP instance, which consists of $\text{Timesolt}_{\text{list}}$, $\text{Room}_{\text{list}}$, and $\text{Course}_{\text{list}}$.

Step 2: Generating the initial solution using the first fit algorithm, which can save a lot of time to find a better solution.

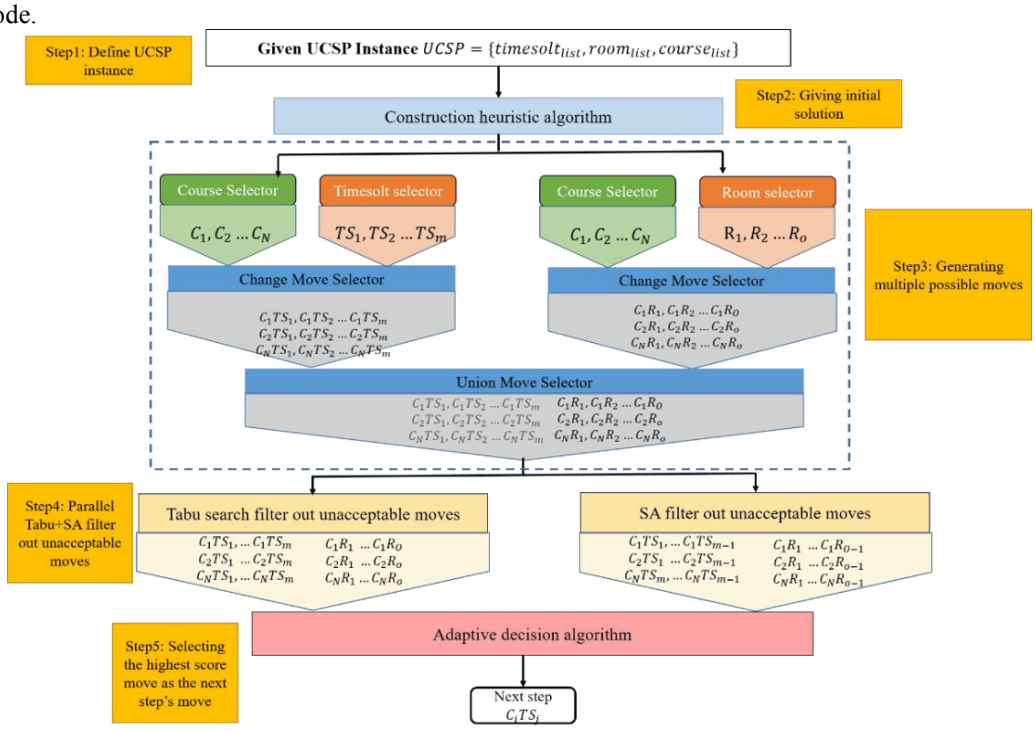


Fig. 1. The proposed parallel Tabu+SA for solving UCSP.

Step 3: We define a union of move selectors to generate the possible moves for the next step. Especially, one is for course object $C = \{C_1, C_2 \ldots C_N\}$ and value timeslot $TS = \{TS_1, TS_2 \ldots TS_m\}$, another one is for object course and value room $R = \{R_1, R_2 \ldots R_o\}$. The principle of change move is automatically changing the value into another one. For example, from C_1TS_1 to C_1TS_2. After two move selections, one union move selector is used to combine C_iTS_j and C_iR_k for the next step.

Step 4: Parallel Tabu+SA algorithm to filter out unacceptable moves generated by step 3. Tabu and SA work separately to receive the potential next step's solution. A detailed explanation of each algorithm is shown in algorithms 1 and 2.

Algorithm 1: Tabu search

Input: Possible next steps solution $PS = \{C_1TS_1, C_1TS_2 \ldots C_1TS_N, \ldots, C_1R_1 \ldots C_1R_o\}$
Output: Accpectable next step solution $AP \in PS$.
Define current solution i, best solution $s = i$, and tabu list H.

1. Randomly generate one solution $i \in PS$, and calculate its fitness score f(i).
2. While not stop:
3. $A = N(i, H)$ #generate the neighbors of solution I from H.
4. $i = SelectionBestSolution(A)$
5. Update the H.
6. if $f(i) < f(s)$:
7. $f(s) = f(i)$
8. End if
9. $k = k + 1$
10. End While
11. $AP = s$

Algorithm 2: SA algorithm

Input: Possible next steps solution $PS = \{C_1TS_1, C_1TS_2 \ldots C_1TS_N, \ldots, C_1R_1 \ldots C_1R_o\}$
Output: Accpectable next step solution $AP \in PS$.
Define starting temperature $T_s = 1000$, initial solution i, best solution $s = i$, and the cooling ratio is 0.99.

1. Randomly generate one solution $i \in PS$, and calculate its fitness score f(i).
2. While not stop
3. While not stop
4. Generate a random neighbor s'
5. $\Delta E = f(s) - f(s')$
6. If $\Delta E < 0$ then $s = s'$
7. Else accept s' with probability $e^{\frac{-\Delta E}{T_s}}$
8. Until condition
9. Update $T_s = G(T_s)$
10. Until T_s<1
11. $AP = s$

Step 5: Select the next move using one adaptive decision algorithm. The algorithm selects the move with the highest score. If both Tabu and SA have similar highest scores, will randomly select one to execute.

3. Experimental Verification

We implement the proposed approach using Python 3.6 to test its effectiveness on one real UCSP instance, the Department of Information Systems from Pukyong National University (PKNU).

3.1 Results Analysis

The instance consists of eight rooms $R = \{R_1, R_2, .., R_8\}$, where R_1 and R_4 are for practice rooms and others are teaching rooms. We only arrange courses from Monday to Friday, each day has three courses starting from 9:00 am and ending at 21:30, and each course takes three and half hours. Besides, one week has 17 courses $C = \{C_1, C_2, C_3, \ldots, C_{17}\}$ touched by six professors $Prs = \{Pr_1, Pr_2, \ldots, Pr_6\}$, where $C_7, C_8, C_9, C_{10}, C_{14}$ and C_{15} are practice courses and others are not. Notice that the professor's preferable time duration is not given as the page limitations, which will be explained in the extension version.

Table 2. Solving PKNU UCSP instance using the proposed method

Day	Timeslot		
	9:00-12:30	13:00-16:30	18:30-21:30
Monday	$C_4(Pr_4\|19th\|R_1)$ $C_{14}(Pr_6\|17th\|R_4)$	$C_8(Pr_4\|21st\|R_1)$ $C_2(Pr_2\|18th\|R_4)$	$C_{11}(Pr_2\|19th\|R_4)$
Tuesday	$C_{13}(Pr_6\|22nd\|R_4)$	$C_{10}(Pr_6\|18th\|R_4)$	
Wednesday			
Thursday	$C_{16}(Pr_2\|18th\|R_4)$	$C_3(Pr_3\|17th\|R_1)$ $C_7(Pr_2\|19th\|R_4)$ $C_5(Pr_3\|20th\|R_3)$	$C_{15}(Pr_5\|18th\|R_4)$
Friday	$C_{17}(Pr_1\|19th\|R_1)$	$C_1(Pr_1\|17th\|R_1)$ $C_{12}(Pr_2\|20th\|R_4)$	$C_6(Pr_1\|17th\|R_1)$ $C_9(Pr_2\|20th\|R_4)$

The timetable solved by the proposed approach is shown in **Table 2**. Where the results are shown with the format of the $Course(Professor|Level|Room)$. The results showed that the proposed method satisfied all hard constraints and most of soft constraints. It arranges all practice courses correctly and only uses three rooms, which reduces the management cost for cleaning the room, delivering keys, etc. In which, each professor token courses in a sequence mode while students taken courses are

not. Especially, professor Pr_1 likes teaching on Friday in the same room R_1 in a sequence mode, but the students are not.

3.2 Ablation and Comparative Study

We designed an ablation study to explore the inner effect of the proposed method. Especially, designed Tabu search and SA alone to explore their effectiveness; Designed the Tabu+SA without first fit construction heuristic algorithm ($\text{Proposed}_{\text{wo}}$) to validate its effectiveness. The comparative results as shown in **Fig. 2**.

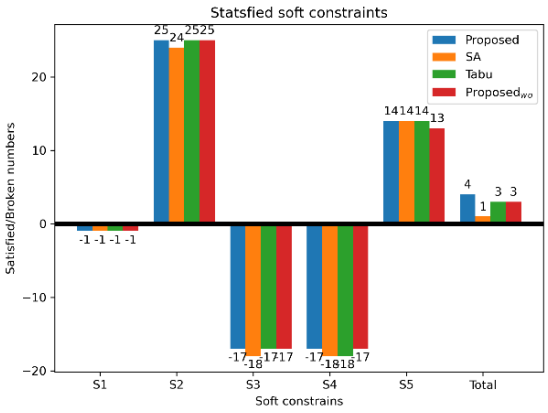


Fig. 2. The comparative results in terms of satisfied soft constraints

The results showed that the proposed parallel Tabu+SA performs the best, it obtains the highest soft constraints score (4) in total. All methods broke one time for S1 and received the same score of 25 for S2 except SA is 24. For S3 and S4, all methods broke 17 or 18 times, and the proposed method outperforms the others. We can conclude that the usage of Tabu has improved by 3 scores; and SA has improved by 1 score, which is conducted by comparing the proposed method with SA and Tabu alone accordingly. Besides, the first fit could give the initial solution to help the proposed method to meet S5.

4. Conclusions

This paper presents one novel and effective approach, parallel Tabu+SA, to solve UCSP. Firstly, it utilizes the first fit algorithm to give the initial solutions, and then one union change move selector is designed to increase the diversity of potential moves. The Tabu search and SA algorithms are designed in a parallel mode to select the next move with the highest score. This good design makes full use of Tabu search and SA to help the proposed method increase accuracy. Besides, an adaptive decision selection algorithm is designed to select the next move from Tabu and SA solutions automatically. The experimental verification based on the PKNU real case verified the proposed method's effectiveness. Moreover, one ablation study confirmed each component's effectiveness in the proposed method.

In summary, the proposed method could effectively solve USCP, it also could be easily transferred to other scheduling problems. Besides, we have detected the proposed method on other benchmarks (not shown because of the page limitations), and the results confirmed the proposed method's effectiveness.

References

[1] S. Imran Hossain, M. A. H. Akhand, M. I. R. Shuvo, N. Siddique, and H. Adeli, "Optimization of University Course Scheduling Problem using Particle Swarm Optimization with Selective Search," *Expert Syst. Appl.*, vol. 127, pp. 9–24, 2019,

[2] M. Azmi Al-Betar and A. Tajudin Khader, "A hybrid harmony search for university course timetabling," *MISTA 2009*, no. August, pp. 10–12, 2009.

[3] N. Boland, B. D. Hughes, L. T. G. Merlot, and P. J. Stuckey, "New integer linear programming approaches for course timetabling," *Comput. Oper. Res.*, vol. 35, no. 7, pp. 2209–2233, 2008.

[4] S. Abdullah and H. Turabieh, "Generating university course timetable using Genetic Algorithms and local search," Proc. - 3rd Int. Conf. Converg. *Hybrid Inf. Technol. ICCIT 2008*, vol. 1, no. November, pp. 254–260, 2008.

[5] F. H. Awad, A. Al-Kubaisi, and M. Mahmood, "Large-scale timetabling problems with adaptive tabu search," J. *Intell. Syst.*, vol. 31, no. 1, pp. 168–176.

[6] R. M. Chen and H. F. Shih, "Solving university course timetabling problems using constriction particle swarm optimization with local search," *Algorithms*, vol. 6, no. 2, pp. 227–244, 2013.

Copyright ⓒ 2023 KSII

Revolutionizing ESG Evaluation: A Study on the Potential of Earnings Call Transcript Data

Haein Lee[1], Hae Sun Jung[2], Heungju Park[4] and Jang Hyun Kim[3*]
[1]Department of Applied Artificial Intelligence/Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University, Seoul, 03063, Korea
[2]Department of Applied Artificial Intelligence, Sungkyunkwan University, Seoul, 03063, Korea
[3] Department of Interaction Science/Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University, Seoul, 03063, Korea
[4]SKK Business School, Sungkyunkwan University, Seoul, 03063, Korea
[e-mail: lhi00034@g.skku.edu, jestiriel@g.skku.edu, hj.park@skku.edu, alohakim@skku.edu]
*Corresponding author: Jang Hyun Kim

Abstract

While the inclusion of ESG criteria is recognized as essential for sustainability and increased firm value, inconsistent disclosure of ESG data and vague assessment standards have been key challenges. To address these issues, this study proposes a novel strategy using text-based automated techniques to rate ESG. Earnings call transcript data were classified as E, S, or G using the Refinitiv-Sustainable Leadership Monitor's 450+ metrics. The study employed advanced natural language processing techniques such as BERT, RoBERTa, and ALBERT models to accurately categorize preprocessed ESG documents, using a voting ensemble model. The results demonstrate the effectiveness of this approach in enhancing ESG assessment criteria established by different rating agencies.

Keywords: ESG, Earnings call transcript, Natural language processing (NLP), BERT

1. Introduction

The importance of environmental, social, and governance (ESG) has grown significantly over the past few years, with international cooperation aimed at reducing carbon emissions and combating climate change [1]. However, the impartiality of ESG ratings remains a debated issue. ESG ratings are highly inconsistent, with a lack of transparency regarding the evaluation criteria by ESG rating agencies. Furthermore, the absence of standards for reporting ESG information makes it considerably harder to ensure objectivity [2,3]. With the increasing volume of data produced in the ESG space, it is crucial to employ advanced natural language processing (NLP) techniques to effectively manage and analyze the unstructured textual data.

To address these challenges, this study proposed a strategy for objective ESG scoring of corporate open data using text-based automated techniques. Earnings call transcript data were collected from Thomson Reuters and labeled E, S, or G using ESG documents provided by the Refinitiv-Sustainable Leadership Monitor, which employs over 450 scales. The study employed advanced NLP techniques, including BERT, RoBERTa, and ALBERT models, to accurately classify preprocessed ESG documents using a voting ensemble model. This approach can help companies understand the factors behind their high ESG scores, providing insights for better decision-making.

This study was supported by a National Research Foundation of Korea (NRF) (http:// nrf.re.kr/eng/index) grant funded by the Korean government (RS-2023-00208278).

2. Related Works

2.1 ESG Value

As ESG continues to gain importance, related studies are emerging. According to [4], companies that neglect their social responsibilities or have ineffective governance face substantial "hidden" risks of economic losses or payment of costs due to environmental litigation. According to [5], ESG was found to be effective in reducing operational costs such as raw material or carbon costs, which could impact an enterprise's operating income by up to 60%. Additionally, research [6,7] has shown that strong ESG practices within a company can attract competent employees, enhance motivation, and ultimately increase overall productivity. Moreover, companies that prioritize social responsibility have shown a positive correlation with shareholder returns. Overall, research has demonstrated the potential financial benefits of incorporating ESG criteria into business operations. These benefits range from enhanced corporate performance indicators to cost savings and increased productivity. As a result, ESG is becoming increasingly important for companies as they strive for long-term sustainability and success.

2.2 ESG Rating

ESG indicators continue to be a topic of debate due to their perceived lack of impartiality. [2] notes that investors face significant challenges when incorporating ESG information into their investment strategies because there are no standardized guidelines for managing ESG information reporting. In addition, [3] compared various ESG ratings and information providers and highlighted the limitations of identifying objectivity or equity due to the lack of transparency regarding the evaluation criteria used by ESG rating agencies. According to [8], there is a relationship between a company's credit rating and its corporate social responsibility (CSR). However, different ESG rating agencies consider different individual components of CSR to be relevant, leading to different ESG ratings. To address these issues, there is a need for a standardized framework that can objectively measure ESG ratings regardless of the data used.

2.3 ESG and Natural Language Processing (NLP)

This section summarizes existing research that has utilized natural language processing (NLP) techniques to extract insights from ESG-related text data. Reference [9] employed NLP to convert unstructured text data extracted from social media into ESG scores and utilized BERT to identify ESG risks. Reference [10] analyzed big data from news articles and academic papers using BERTopic to discover major keywords and topics associated with ESG and proposed a strategic direction for successful ESG management. In Reference [11], past tendencies in ESG discourses were analyzed by examining the records of corporate earning calls, and the authors found that 15 percent of statements made during earning calls over the past five years were related to ESG, indicating the growing importance of ESG in business strategy.

3. Method

3.1 Data Collection and Preprocessing

The authors collected earnings call transcripts from Thomson Reuters between January 1, 2014, and December 31, 2018, resulting in a dataset of 3,574,923. They then used the ESG pillar scores from the Refinitiv-Sustainable Leadership Monitor [12,13] to label the data. This metric is designed to assess the relative environmental, social, and governance performance of companies within a sector, measuring ESG achievement and efficiency.

3.2 Model

The NLP approach suggested in the study uses a methodology based on transformer architecture language modeling, which is capable of recognizing linguistic features and learning general language characteristics [14]. This approach comprises two stages: pretraining and downstream, where a large corpus is preprocessed to capture the general attributes of the language. The authors evaluated different models, including BERT, RoBERTa, and ALBERT, to determine their suitability for the ESG domain.

4. Results

In this section, the authors compared the performance of single and ensemble models that use language models to classify ESG. They conducted experiments with different batch sizes and found that, in general, the ensemble models outperformed the individual classifiers. The results indicated that the ensemble model using BERT and RoBERTa had the highest accuracy of 96.84%, with a batch size of 32.

5. Conclusions

This article discusses the importance of ESG ratings in assessing a company's long-term investment potential. However, the lack of uniformity among companies and standards for handling ESG information reporting have been identified as primary obstacles in incorporating ESG ratings into investment procedures. To address these limitations, the authors collected earnings call transcript from Thomson Reuters and used advanced NLP techniques to extract pertinent information from extensive textual data to supplement ESG evaluation criteria. They used a transformer-based language model and ensemble models to classify ESG ratings, which proved to be effective. This article contributes significantly to the advancement of knowledge in the field of ESG assessment. It addresses the scientific gap by tackling the challenges associated with inconsistent ESG data disclosure and vague assessment standards. The novel strategy presented in this study not only provides a systematic framework for categorizing ESG-related documents but also offers a more objective and reliable approach to evaluating companies' ESG performance. By leveraging advanced NLP techniques, this research establishes a foundation for standardizing ESG assessment and enables comparisons across companies and industries. However, The model was validated using only earnings call transcript data, and future research is necessary to verify it using ESG data from various sources.

Acknowledgements

This study was supported by a National Research Foundation of Korea (NRF) (http://nrf.re.kr/eng/index) grant funded by the Korean government (RS-2023-00208278).

References

[1] A. G. Hoepner, A. A. Majoch and X. Y. Zhou, "Does an asset owner's institutional setting influence its decision to sign the principles for responsible investment?," *Journal of Business Ethics*, vol. 168, no. 2, pp. 389–414, 2021.

[2] A. Amel-Zadeh and G. Serafeim, "Why and how investors use ESG information: Evidence from a global survey," *Financial Analysts Journal*, vol. 74, no. 3, pp. 87-103, 2018.

[3] E. Escrig-Olmedo, M. Á. Fernández-Izquierdo, I. Ferrero-Ferrero, J. M. Rivera-Lirio and M. J. Muñoz-Torres, "Rating the raters: Evaluating how ESG rating agencies integrate sustainability principles," *Sustainability*, vol. 11, no. 3, pp. 915, 2019.

[4] V. Díaz, D. Ibrushi and J. Zhao, "Reconsidering systematic factors during the COVID-19 pandemic–The rising importance of ESG," *Finance Research Letters*, vol. 38, pp. 101870. 2021.

[5] W. Henisz, T. Koller and R. Nuttall, "Five ways that ESG creates value," *McKinsey*, 2019.

[6] A. Edmans, "Does the stock market fully value intangibles? Employee satisfaction and equity prices," *Journal of Financial economics*, vol. 101, no. 3, pp. 621-640, 2011.

[7] A. Edmans, "The link between job satisfaction and firm value, with implications for corporate social responsibility," *Academy of Management Perspectives*, vol. 26, no. 4, pp. 1-19, 2012.

[8] N. Attig, S. El Ghoul, O. Guedhami and J. Suh, "Corporate social responsibility and credit ratings," *Journal of business ethics*, vol. 117, pp. 679-694, 2013.

[9] A. Sokolov, J. Mostovoy, J. Ding and L. Seco, "Building machine learning systems for automated ESG scoring," *The Journal of Impact and ESG Investing*, vol. 1, no. 3, pp. 39-50, 2021.

[10] H. Lee, S. H. Lee, K. R. Lee and J. H. Kim, "Esg discourse analysis through bertopic: comparing news articles and academic papers," Computers, Materials & Continua,

vol. 75, no.3, pp. 6023-6037, 2023.
[11] N. Raman, G. Bang and A. Nourbakhsh, "Mapping ESG trends by distant supervision of neural language models," *Machine Learning and Knowledge Extraction*, vol. 2, no. 4, pp. 453-468, 2020.
[12] F. Berg, K. Fabisik and Z. Sautner, "Is history repeating itself? The (un) predictable past of ESG ratings," *The (Un) Predictable Past of ESG Ratings (August 24, 2021). European Corporate Governance Institute–Finance Working Paper*, vol. 708, 2020.
[13] Refinitiv, [Online]. Available: https://www.refinitiv.com/en/products/sustainability-reporting-on-leadership
[14] J. Howard and S. Ruder, "Universal language model fine-tuning for text classification," 2018. [Online]. Available: https://doi.org/10.48550/arXiv.1801.06146

Copyright ⓒ 2023 KSII

LSTM-based Prediction of Bitcoin Prices Using Technical Indicators, Sentiment Analysis, and Google Trends

Hae Sun Jung[1], Haein Lee[2], and Jang Hyun Kim[3*]
[1]Department of Applied Artificial Intelligence, Sungkyunkwan University, Seoul, 03063, Korea
[2]Department of Applied Artificial Intelligence/Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University, Seoul, 03063, Korea
[3] Department of Interaction Science/Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University, Seoul, 03063, Korea
[e-mail: jestiriel@g.skku.edu, lhi00034@g.skku.edu, alohakim@skku.edu]
*Corresponding author: Jang Hyun Kim

Abstract

Predicting Bitcoin price trends is crucial as they serve as indicators of the overall cryptocurrency market trend. Due to Bitcoin's short history and high volatility, numerous studies have investigated the factors that influence its price fluctuations. In this study, we gathered data from Reddit and LexisNexis spanning over six years. We employed Long Short-Term Memory (LSTM) with inputs that included technical and sentiment indicators, as well as Google trends data, to make predictions. This approach yielded an RMSE of 0.0188, MSE of 0.0004, and an R-squared score of 0.8641. As a result, it is shown that the use of sentiment index, technical indicators such as moving average, and relative intensity index (RSI) with Google trends data can achieve compliant results. Consequently, the input function employed in this study can be applied to forecast Bitcoin prices, assisting investors in making informed decisions regarding Bitcoin-related investments.

Keywords: Bitcoin, cryptocurrency, sentiment analysis, price prediction, LSTM

1. Introduction

Bitcoin is a digital currency that utilizes blockchain technology, allowing for fast payments and transfers between parties without the need for financial institutions [1]. Various cryptocurrencies have emerged since the late 2000s, and the market's interest in cryptocurrencies, particularly Bitcoin, increased significantly in late 2017 and early 2018, recording unprecedented returns [2,3]. Cryptocurrencies, including Bitcoin, are recognized as a type of asset, and regulations related to cryptocurrencies are gradually settling in. According to previous studies, cryptocurrencies can be utilized for portfolio diversification and hedging purposes [4,5]. Since its launch on January 3, 2009, Bitcoin has become a symbolic asset that leads the cryptocurrency market [6]. According to [7], other cryptocurrencies tend to move in the same direction as Bitcoin's price trend [8].

However, the cryptocurrency market has a short history and high volatility. Therefore, achieving good prediction performance based solely on past sequential price data is impossible. Unlike stock investors who can obtain various forms of information to estimate a company's value, cryptocurrency investors tend to rely on media information due to the lack of evidence to judge the value of cryptocurrencies. According to [9,10], the cryptocurrency market strongly reacts to news and social media.

This study presents a methodology that utilizes sentiment analysis, technical indicators and Google Trends data. Ultimately, Bitcoin price prediction considering investment sentiment and interset with technical analysis was performed using Long Short-Term Memory (LSTM).

2. Related Works

2.1 Factors Affecting Changes in the Price of Bitcoin

Due to its limited historical data and volatile nature, researchers have focused their attention on identifying the factors that impact the price of Bitcoin. According to [11], external factors such as fluctuations in interest rates and exchange rates can influence Bitcoin's price. Moreover, [12] determined that there is either no correlation or a negative correlation between Bitcoin's price and other financial assets like gold, the US dollar, and major stock market indices. Through the analysis of Google search term queries, [13] discovered that global interest in Bitcoin typically precedes price increases, while prices tend to decline when there is an increase in market collapse fears. Additionally, [14] established a strong correlation between public interest and Bitcoin prices. Lastly, [15] utilized Pearson's correlation coefficient to demonstrate that Twitter sentiment can be useful in predicting fluctuations in Bitcoin prices.

2.2 Study on Predicting Bitcoin Prices Using Sentiment Analysis

In recent years, there has been an increasing focus on utilizing social media platforms, online forums, and news posts to forecast Bitcoin's price by examining the sentiments expressed by investors and the public. According to [16], the sentiment score derived from the valence aware dictionary and sentiment reasoner (VADER) during the COVID-19 pandemic displayed a noteworthy correlation with short-term price trends of Bitcoin. Similarly, [17] employed support vector machines (SVM), regression models, and Twitter sentiments to predict fluctuations in Bitcoin's price. In [18], logistic regression was applied using tweets and news headlines to forecast price changes for Bitcoin and Ethereum. Previous research indicates that utilizing social media or news data can be an effective approach in assessing overall investor sentiment. Furthermore, the authors emphasized the potential for more accurate predictions by incorporating technical indicators commonly used in stock or cryptocurrency trading into the analysis.

3. Method

3.1 Data Collection

News, Google Trends, and social media data were collected between March 1, 2017, and March 1, 2023. A total of 17,230 news articles were obtained from LexisNexis by searching for terms such as "Bitcoin," and "BTC." On Reddit, submissions, and comments from the r/Bitcoin, r/BTC subreddits were collected using the Pushshift API, resulting in 10,699,720 data. Data on the price of Bitcoin in U.S. dollars, including closing prices, opening prices, high prices, low prices, and transaction volumes, were collected from Investing.com.

3.2 Technical Indicators

To identify appropriate technical indicators for predicting Bitcoin prices, the authors referred to previous research on stock price prediction [19,20]. The Python TA-lib library was used to calculate the technical indicators from the collected Bitcoin historical price data. The relative strength index (RSI), simple moving average (SMA), exponential moving average (EMA), moving average convergence and divergence (MACD), Stochastic RSI were employed.

3.3 Sentiment Index

To obtain a compound score, the lexicon-based unsupervised method called VADER was employed. A threshold of 0.1 was established for the compound score. Based on this threshold, each post was categorized as positive if the score was equal to or greater than the threshold, and negative if it fell below the threshold. Based on [21], the sentiment index was then calculated as the difference between the number of negative and positive posts on a given day.

$$Sentiment\ Index = \frac{M_{tpos} - M_{tneg}}{M_{tpos} + M_{tneg}} \quad (1)$$

Where M_{tpos} is the total amount of positive text, and M_{tneg} is the total number of negative text on the day t, respectively

3.4 Model

The regression results were derived through LSTM by adding concatenated data as input in the form of input window size of 20 days and output window size of 1 day. In conclusion, we conducted a price regression analysis on the

closed price of Bitcoin for 300 days before March 1, 2023.

4. Result

Since Bitcoin price prediction is a regression task, the prediction performance was evaluated through root mean squared error (RMSE), mean squared error (MSE) and R-squared scores. The regression analysis results are presented in **Fig. 1.** And **Table 1.**



Fig. 1. Bitcoin price prediction for last 10 month

Table 1. Regression results

RMSE	0.0188
MSE	0.0004
R-squared score	0.8641

5. Conclusion

In this research, data was gathered from Reddit, which serves as a representative social media platform [22]. To overcome the limitations of Twitter's data collection using the API, we obtained all submissions and comments from Reddit through the Pushshift API. Utilizing Reddit data offers advantages in terms of both volume and noise reduction compared to other platforms, as users interact within specific topic-oriented bulletin boards known as subreddits. The collected data spanned a period of four years, enabling comprehensive experiments to be conducted. Three indicators were used as inputs for the model: technical indicators derived from Bitcoin price data, a sentiment index calculated from Reddit and news text, and Google Trends data. By implementing LSTM, the model achieved an RMSE of 0.0188, MSE of 0.0004, and an R-squared score of 0.8641. This study demonstrated that combining technical indicators with a sentiment index and Google Trends data can enhance the accuracy of Bitcoin price predictions. These input features can be applied in future research on Bitcoin price forecasting. However, it is important to note that the current study has limitations that should be addressed in future work. Firstly, lexicon-based unsupervised learning tends to have lower accuracy compared to approaches using labeled data. Additionally, the performance of lexicon-based learning may not be optimal when applied to unstructured text data. Secondly, the prediction of trends was conducted on a daily basis rather than in real-time using inputs at an hourly or minute level.

Acknowledgements

This study was supported by a National Research Foundation of Korea (NRF) (http://nrf.re.kr/eng/index) grant funded by the Korean government (RS-2023-00208278).

References

[1] R. Böhme, N. Christin, B. Edelman and T. Moore, “Bitcoin: Economics, technology, and governance,” Journal of Economic Perspectives, vol. 29, no. 2, pp. 213-38, 2015.

[2] CoinDesk, 2017. [Online]. Available: https://www.coindesk.com/markets/2017/12/29/from-900-to-20000-bitcoins-historic-2017-price-run-revisited/

[3] Fortune, 2018. [Online]. Available: https://fortune.com/2018/01/03/ripple-price-crypto-currency/

[4] E. Bouri, P. Molnár, G. Azzi, D. Roubaud and L. I. Hagfors, “On the hedge and safe haven properties of Bitcoin: Is it really more than a diversifier?,” *Finance Research Letters*, vol. 20, pp. 192-198, 2017.

[5] K. Guesmi, S. Saadi, I. Abid and Z. Ftiti, “Portfolio diversification with virtual currency: Evidence from bitcoin,” *International Review of Financial Analysis*, vol. 63 pp. 431–437, 2019.

[6] S. Nakamoto, “Bitcoin: A peer-to-peer electronic cash system,” *Decentralized Business Review*, pp. 21260, 2008.

[7] MakeUseOf, 2022. [Online]. Available:

https://www.makeuseof.com/what-is-btc-dominance/

[8] P. Ciaian and M. Rajcaniova, "Virtual relationships: Short-and long-run evidence from BitCoin and altcoin markets," *Journal of International Financial Markets, Institutions and Money*, vol. 52, pp. 173-195, 2018.

[9] O. Kraaijeveld and J. De Smedt, "The predictive power of public Twitter sentiment for forecasting cryptocurrency prices," *Journal of International Financial Markets, Institutions and Money*, vol. 65, pp. 101188, 2020.

[10] M. Poongodi, T. N. Nguyen, M. Hamdi and K. Cengiz, "Global cryptocurrency trend prediction using social media," *Information Processing & Management*, vol. 58, no. 6, pp. 102708, 2021.

[11] T. Panagiotidis, T. Stengos and O. Vravosinos, "The effects of markets, uncertainty and search intensity on bitcoin returns," *International Review of Financial Analysis*, vol. 63, pp. 220-242, 2019.

[12] A. H. Dyhrberg, "Bitcoin, gold and the dollar–A GARCH volatility analysis," *Finance Research Letters*, vol. 16, pp. 85-92, 2016.

[13] R. Hakim das Neves, "Bitcoin pricing: Impact of attractiveness variables," *Financial Innovation*, vol. 6, no. 1, pp. 1-18, 2020.

[14] Y. Kaya, "Analysis of cryptocurrency market and drivers of the Bitcoin price: Understanding the price drivers of Bitcoin under speculative environment," M.S. thesis, Department of Industrial Engineering and Management, KTH Royal Institute of Technology, Stockholm, Sweden, 2018.

[15] T. R. Li, A. S. Chamrajnagar, X. R. Fong, N. R. Rizik and F. Fu, "Sentiment-based prediction of alternative cryptocurrency price fluctuations using gradient boosting tree model," *Frontiers in Physics*, vol. 7, no. 98, 2019.

[16] T. Pano and R. Kashef, "A complete VADER-based sentiment analysis of bitcoin (BTC) tweets during the era of COVID-19," *Big Data and Cognitive Computing*, vol. 4, no. 4, pp. 33, 2020.

[17] I. Georgoula, D. Pournarakis, C. Bilanakos, D. Sotiropoulos and G. M. Giaglis, "Using time-series and sentiment analysis to detect the determinants of bitcoin prices," 2015. [Online]. Available: SSRN-id2607167.

[18] C. Lamon, E. Nielsen and E. Redondo, "Cryptocurrency price prediction using news and social media sentiment," *SMU Data Sci. Rev*, vol. 1, no. 3, pp. 1-22, 2017.

[19] Y. Zhai, A. Hsu and S. K. Halgamuge, "Combining news and technical indicators in daily stock price trends prediction," in *Proc. ISNN*, Nanjing, China, pp. 1087-1096, 2007.

[20] R. Rosillo, D. De la Fuente and J. A. L. Brugos, "Technical analysis and the Spanish stock exchange: Testing the RSI, MACD, momentum and stochastic rules using Spanish market companies," *Applied Economics*, vol. 45, no. 12, pp. 1541-1550, 2013.

[21] S. Wu, Y. Liu, Z. Zou and T. H. Weng, "S_I_LSTM: Stock price prediction based on multiple data sources and sentiment analysis," *Connection Science*, vol. 34, no.1, pp. 44-62, 2022.

[22] S. Priya, R. Sequeira, J. Chandra and S. K. Dandapat, "Where should one get news updates: Twitter or Reddit," *Online Social Networks and Media*, vol. 9, pp.17-29, 2019.

Copyright ⓒ 2023 KSII

Public Opinion Analysis of E-sports in China Based on LDA Model: A Case Study of EDG Winning the Championship

ShaoPeng Che[1] and Kai Kuang[1]
[1]School of Journalism and Communication, Tsinghua University, Beijing, China
[e-mail: cheshao@tsinghua.edu.cn, kaikuang17@gmail.com]
*Corresponding author: Kai Kuang

Abstract

Public opinion plays a crucial role in developing and influencing the e-sports industry in China. Understanding and analyzing public opinion on e-sports in China can help us uncover the attitudes, viewpoints, and sentiments of the public toward this industry. However, there is still a research gap in the study of public opinion on e-sports in China. Therefore, this study took the case of EDward Gaming winning the 2021 League of Legends World Championship as the research subject, utilized Sina Weibo as the data source, and conducted a preliminary exploration of public opinion in the Chinese e-sports industry based on the Latent Dirichlet Allocation model. The results indicated that the optimal number of topics was eight, and the distribution exhibited diversity, reflecting that participants in topic discussions came from different domains.

Keywords: E-sports, public opinion, LDA, EDG, League of Legends, Weibo

1. Introduction

Being an emerging form of competition, e-sports is rapidly growing worldwide and has garnered tremendous attention and growth in China. As one of the largest e-sports markets globally, China's e-sports industry exhibits a solid upward trajectory regarding event scale, viewership, and market value [1]. In particular, the successful organization of e-sports events, such as the League of Legends World Championship, coupled with outstanding performances by players, has further propelled the development of e-sports in China [2]. However, despite the increasing public interest in e-sports, there remains a limited understanding of the public's perspectives, attitudes, and sentimental inclinations [3].

In the realm of Chinese e-sports, the significance of public opinion cannot be ignored [4]. Understanding public perceptions and attitudes towards e-sports, and exploring key topics in public opinion, is of great importance to policymakers and managers in the e-sports industry [5]. Therefore, a question worthy of investigation is how to employ data analysis to conduct an in-depth excavation of public opinion on e-sports in China and uncover the public's views, attitudes, and sentimental tendencies.

This study aims to conduct a preliminary analysis of public opinion on e-sports in China using the Latent Dirichlet Allocation (LDA) model. We selected EDward Gaming (EDG)' s victory in the 2021 League of Legends World Championship (LOLWC) as a case study, collected relevant social media text data, and applied the LDA model to identify the topics discussed by the public during this event. By exploring public opinion characteristics, we aim to discover significant topics in the Chinese e-sports landscape.

2. Related Works

2.1 E-sports Public Opinion Research

By utilizing advanced technologies such as big

This project is supported by the Computational Communication and Intelligent Media Lab Research Support Plan at the School of Journalism and Communication, Tsinghua University

data analysis, social network mining, and natural language processing, the research on e-sports public opinion aims to unveil the patterns of generation, dissemination, and influence of e-sports public opinion.

Through the analysis of indicators such as the sentiment tendency of public opinion, key issues, and user participation, researchers can identify the popularity, trends, and evolution of topics. This analysis allows for the construction of dynamic maps of public opinion and models of network influence, enabling the identification of opinion leaders and critical communication pathways. These research findings give practitioners and decision-makers valuable decision support and strategic guidance [6].

E-sports public opinion research also explores the interaction between public opinion and the e-sports ecosystem. It analyzes the influence of public opinion on audience behavior, player image formation, and brand promotion, which serves as an essential basis for developing the e-sports industry chain and decision-making by stakeholders [7].

In summary, the study of e-sports public opinion delves into the phenomenon of public opinion in the e-sports domain and its impact on industrial development through data analysis methods.

2.2 Application of LDA in e-sports public opinion

In recent years, the rapid growth of e-sports has led to an increased public awareness of this field. In this context, the LDA model has proven to be an effective tool for studying e-sports public opinion.

By analyzing extensive social media data and online discussions, the LDA model can unearth the public's focal points, popular discussion topics, and topic distribution [8]. Through the mining of sentimental tendencies, researchers can gain insights into the public's preferences and needs concerning e-sports. This information provides a scientific foundation for developing the e-sports industry and serves as a guide for shaping public opinion [9].

Furthermore, the LDA model can be combined with machine learning techniques to construct predictive models and recommendation systems, that support personalized recommendations and e-sports product marketing strategies [10].

In brief, the LDA model plays a critical role in studying e-sports public opinion, serving as a powerful tool and methodology for deepening our understanding of the public's perceptions and demands related to e-sports.

3. Method

3.1 Data Collection and Preprocessing

In this study, we utilized Python to crawl all the comments posted on the official account of the EDG team on the Chinese Sina Weibo platform, explicitly focusing on the comments made under the 2021 LOLWC post published on November 7, 2021. A total of 22,695 comment data were collected and used as research objects.

To ensure the quality and accuracy of the data, we implemented data-cleaning techniques. This involved removing noisy information, such as special characters, punctuation marks, and web links. Subsequently, we employed the Jieba word segmentation tool to segment the original text into a sequence of words, enabling further analysis [11].

A stopword list was constructed to mitigate the influence of irrelevant and common words on the analysis results. This list was applied to the results of the word segmentation process, filtering out the stopwords [12].

3.2 Feature Extraction and Weight Calculation

To capture and quantify the textual feature information, the TfidfVectorizer was employed as the feature extraction method [13]. This allowed us to convert the segmented text into a feature matrix. Additionally, several parameters were set, including the maximum number of features and the stopwords list, to enhance the informativeness and distinctiveness of the feature vector representation. By computing the weights of words using a combination of their frequency within the text and their inverse document frequency across the entire corpus, we derived the feature vector representation for each text [14].

3.3 Model Construction and Training

The LDA model was employed to train and model the feature matrix [15]. The parameters of the LDA model, such as the number of topics and iterations, were set to obtain the best-fitting results for the topic model. Through an iterative optimization algorithm, we inferred the distribution of topics for each text and the word distribution for each topic. This process ultimately led to the training of the LDA model [16].

3.4 Perplexity Calculation and Model Evaluation

Perplexity serves as a vital indicator to assess the effectiveness of the LDA model and the reliability of the interpreted results. It measures the model's ability to predict unseen data, with lower values indicating better performance on new data. In this study, we computed the perplexity of the trained LDA model using the reserved test data. Furthermore, we comprehensively compared and evaluated various models to identify the optimal one for our analysis [17].

4. Result

Fig. 1 illustrates the perplexity curve, while **Table 1** showcases the top five words associated with each of the eight topics.

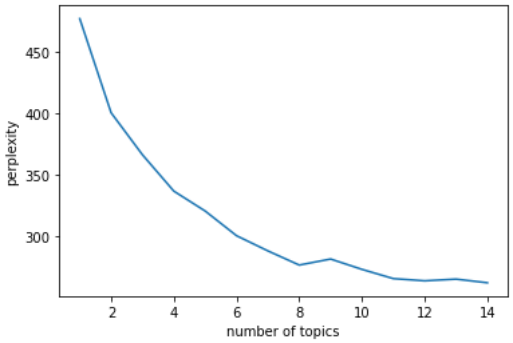


Fig. 1. Perplexity curve

Table 1. Top five words in each of the eight topics

Topic	Top five words
LPL is NO.1	Final, world, champion, defeat, NO.1
Replacing the national flag with the EDG team flag	Students, national flag, university, team flag, video
Boyfriend is excited	Live streaming, boyfriend, play games, mood, adore a star
Setting up flag	Flag, disgusting, win, speechless, quality
A historic moment	E-sports club, summit, moment, trophy, dream
Celebrations disturb public	Boy, disturbance, social issues, affect, rational
Missing Clearlove	Fans, Clearlove, stopped in the quarterfinals, gold content, spirit
This event is hot on WeChat Moments	WeChat Moments, flood the screen, heat, young people, influence

We found significant differences among the topics, and through text tracing, we discovered that this is because the fact that the public participating in topic discussions comes from various domains.

5. Conclusion

This paper presented a preliminary exploration of public opinion in the Chinese e-sports industry using the LDA model. We have identified several important topics associated with the event by investigating the public opinion event triggered by EDG's victory in the 2021 LOLWC on Weibo. The low correlation between these topics reflects the diverse civic identities within the e-sports community.

The findings of this study provided important insights and recommendations for policymakers, brand marketers, and academic researchers in the e-sports industry, facilitating the development and management of the e-sports field in China.

Future research can build upon this work by incorporating larger data samples and longer periods to explore public opinion evolution trends in Chinese e-sports.

References

[1] Qian, T. Y., Sonkeng, K., & Luo, L., Exploring the Dark Side of Esports Online Spectatorship: Passion as a Mediator and Collective Narcissism as a Moderator. Communication & Sport,

21674795231164305, 2023.

[2] Lord, J. J., How Digital Content Marketing Affects Viewership and Engagement of Multiplayer Online Battle Arena Esport Events: An Examination of League of Legends, 2023.

[3] Zhao, S., & Liu, Y., "From industry development to social influence: video games in Chinese newspaper coverage, 2010–2020," Chinese Journal of Communication, 16(1), 35-52, 2023.

[4] Wong, D., & Meng-Lewis, Y., "Esports diplomacy–China's soft power building in the digital era," Managing Sport and Leisure, 1-23, 2022.

[5] Shu, T., Wang, Z., Jia, H., Zhao, W., Zhou, J., & Peng, T., "Consumers' Opinions towards Public Health Effects of Online Games: An Empirical Study Based on Social Media Comments in China," International Journal of Environmental Research and Public Health, 19(19), 12793, 2022.

[6] Wang, Y., Zheng, X., & Fan, T., "Public Outlook on E-sports in China: A Content and Sentiment Analysis of Social Media," In 2022 IEEE 23rd International Conference on Information Reuse and Integration for Data Science (IRI) (pp. 270-277). IEEE. August 2022.

[7] Sabtan, B., Cao, S., & Paul, N., "Current practice and challenges in coaching Esports players: An interview study with league of legends professional team coaches," Entertainment Computing, 42, 100481, 2022.

[8] Wang, W., & Fan, J., "Topic mining of real-time discussions: what catches the attention of live-streaming esports viewers?," European Sport Management Quarterly, 1-22, 2022.

[9] Konstantinova, K., Bulygin, D., Okopny, P., & Musabirov, I., "Online Communication of eSports Viewers: Topic Modeling Approach," In Advances in Computer Entertainment Technology: 14th International Conference, ACE 2017, London, UK, December 14-16, 2017, Proceedings 14, pp. 608-613, Springer International Publishing, 2018.

[10] Yu, Y., Nguyen, B. H., Yu, F., & Huynh, V. N., "Esports Game Updates and Player Perception: Data Analysis of PUBG Steam Reviews," In 2021 13th International Conference on Knowledge and Systems Engineering (KSE), pp. 1-6. IEEE, November 2021.

[11] Li, X., Guo, X., Kim, S. K., & Lee, H., "Analysis of Social Media Utilization based on Big Data-Focusing on the Chinese Government Weibo," KSII Transactions on Internet & Information Systems, 16(8), 2022.

[12] Che, S., Kamphuis, P., Zhang, S., & Kim, J. H., "RECP from the Perspective of Chinese Mainlanders and Taiwanese Netizens: A Comparative Semantic Network Analysis," In 2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM), pp. 1-4, IEEE, January 2022.

[13] Che, S., Nan, D., Kamphuis, P., Jin, X., & Kim, J. H., "A cluster analysis of Lotte young plaza using semantic network analysis method," In 2021 15th International Conference on Ubiquitous Information Management and Communication (IMCOM), pp. 1-6, IEEE, January 2021.

[14] Che, S., Nan, D., Kamphuis, P., Zhang, S., & Kim, J. H., "Examining Crisis Communication Using Semantic Network and Sentiment Analysis: A Case Study on NetEase Games," Frontiers in Psychology, 13, 176, 2022.

[15] Che, S., Nan, D., Kamphuis, P., & Kim, J. H., "A comparative analysis of attention to facial recognition payment between China and South Korea: a news analysis using Latent Dirichlet allocation," In HCI International 2021-Posters: 23rd HCI International Conference, HCII 2021, Virtual Event, July 24–29, 2021, Proceedings, Part II 23, pp. 75-82, Springer International Publishing, 2021.

[16] Choi, J., Kim, J., & Lee, H., "Hybrid Fraud Detection Model: Detecting Fraudulent Information in the Healthcare Crowdfunding," KSII Transactions on Internet & Information Systems, 16(3), 2022.

[17] Nam, D., & Choi, G., "The Identification of Emerging Technologies of Automotive Semiconductor," KSII Transactions on Internet & Information Systems, 17(2), 2023.

KSII The 18th Asia Pacific International Conference on Information Science and Technology(APIC-IST) 2023.
Copyright ⓒ 2023 KSII

Ontological metaphor in human-computer interaction: a study of the user's conceptual perception on machines

Youngsang Kim
Department of Human-Artificial Intelligence Interaction, Sungkyunkwan University
Seoul, South Korea
[e-mail: mynameis03@skku.edu]

Abstract

This study aimed to analyze users' perceptions of personified chatbots using the conceptual metaphor theory. The conceptual metaphor theory provides a robust framework for understanding speakers' conceptual understanding. By examining users' metaphorical perceptions of chatbots, this research sought to uncover their genuine perceptions of personified chatbots. The study investigated users' cognitive perceptions of personified chatbots by analyzing previous research on personified chatbots and employing the conceptual metaphor theory.

Keywords: conceptual metaphor, chatbot, human machine interaction, personification, conversational interface

1. Introduction

The objective of this study is to analyze users' perceptions of the Chatbots they possess, specifically focusing on their personification, using the conceptual metaphor theory. While there have been numerous discussions on the personification of AI agents, the existence of the uncanny valley theory suggests that a human-like appearance of AI agents in graphical user interfaces may cause user aversion. However, in general, the personification of AI agents is considered to have a positive impact on user interface. Additionally, personified AI agents enhance social presence. According to the CASA(computers are social actors) theory, users tend to perceive virtual entities such as computers as members of society and apply social norms applicable to humans to them (Lee, Nass, 2010). While it is important to consider the culturally relative perspective regarding users' perception of AI agents as personified beings (Fraune et al., 2020), the personification of AI agents not only increases users' trust but also leads them to perceive AI agents as collaborative entities (Lee, Nass, 2010).

Previous studies have predominantly focused on enhancing the personification of AI agent in various areas of human-computer interface (HCI), including graphical user interface, voice user interface, and conversational user interface. These studies have also explored the impact of AI agent personification on usability. Additionally, some research has investigated the influence of users' cultural backgrounds on their perspectives of personification. However, limited attention has been given to the universal perception of personified chatbot.

To address this gap, the present study aims to analyze users' cognitive perspectives on personified Chatbots by employing the conceptual metaphor theory, with a specific focus on ontological metaphors. Conceptual metaphors used by individuals often reveal their cognitive thought processes. Consequently, analyzing the metaphorical perspectives that users hold about Chatbots appears appropriate for examining the universal perception of Chatbots among users.

2. Related work

2.1 Personification of AI agent

Individuals tend to perceive personified AI agents as social entities and express a desire for social interaction with them (Breazeal, 2003). However, AI agents inherently possess inherent uncertainty and ambiguity due to their virtual nature. When AI agents acquire personified characteristics, they are perceived as more predictable and controllable by users, resulting in increased user trust (Jung et al., 2022). Various factors contribute to enhancing the personification of AI agents, including behavioral aspects such as emotional responsiveness, interpersonal warmth, cognitive openness, individuality, depth, and visual appearance resembling humans (Kuzminykh et al., 2000; Brandt & Reyna, 2011).

However, people's perceptions of AI agent personification vary across cultural contexts, influenced by religious and economic factors. For instance, individuals from Western cultures with Christian backgrounds may express reservations about humanizing non-human entities due to the belief in the exclusive possession of souls by humans within Christian teachings. In contrast, East Asian cultures influenced by Buddhism exhibit a less distinct separation between humans and machines, resulting in a lower resistance to the humanization of machines (Samuel, 2019). Additionally, individuals from economically less developed countries often hold a positive attitude towards the advent of new technologies, including the personification of AI agents (Trovato et al., 2017).

Existing research findings suggest that users' perceptions of AI agent personification exhibit cultural relativism, highlighting the need for comprehensive studies on the universal perception of this aspect by users.

2.2 Conceptual metaphor

Conceptual metaphor refers to the cognitive process by which abstract domains such as emotions, thoughts, beliefs, and worldviews are understood through mapping them onto concepts from the material world and expressing them in language. It is widely recognized as possessing universal characteristics that reflect the cognitive system of humans (Lakoff & Johnson, 1980). Ontological metaphor, a specific type of conceptual metaphor, involves metaphorical representations that embody human experiences associated with physical objects or entities and articulate them through language (Lakoff & Johnson, 1987). For instance, one example of an ontological metaphor is "THE MIND IS A MACHINE." This conceptual metaphor represents the conceptualization of the abstract domain of the mind as a physical object, specifically a machine. It gives rise to metaphorical expressions such as "My mind isn't functioning properly today."

Moreover, Santibanez (1999) proposed a connection between ontological metaphor and the theory of the great chain of being, suggesting that they share a common conceptual foundation. The theory of the great chain of being, as formulated by Lovejoy (1963), is rooted in three fundamental principles concerning the nature of existence in the universe. First, the principle of plenitude asserts that various types of existence exhibit maximum diversity and fill every conceivable gap. Second, the principle of continuity posits that the universe consists of interconnected and infinite entities, with each entity sharing at least one attribute with its neighboring entities. Lastly, the principle of linear gradation states that interconnected entities exist in a hierarchical order, ranging from the lowest unit of existence to God as the absolute and complete being. According to this theory, all entities in the world form a hierarchical system, with each entity occupying a specific level in the chain as an essential and unchanging entity (Krikmann, 2007). Applying this theory to the classification of existing entities, God occupies the highest position, followed by angels, humans, animals, plants, and inanimate objects in ascending order (see Figure 1 below).

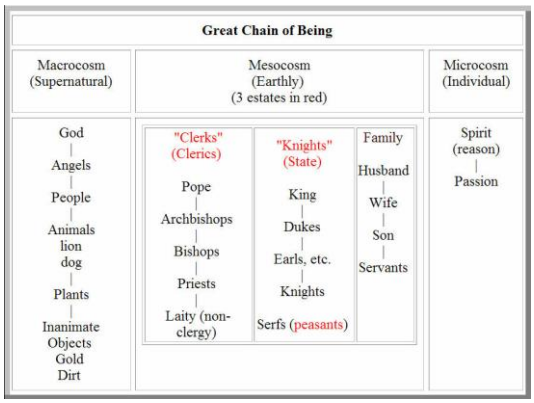


Fig. 1. The great chain of being

2.3 Conceptual metaphor in conversational interface

Conceptual metaphor plays a crucial role in facilitating a seamless interface between users and machines within the field of human-machine interaction. For instance, in the context of Chatbots, conceptual metaphor serves as a cognitive lens through which users can frame and comprehend the abstract domain of Chatbots. An illustrative example is the conceptual metaphor "COMPUTER IS A DESKTOP." Through this metaphor, the abstract concept of a computer is conceptualized and understood as a physical object, specifically a desktop. This enables users to easily grasp the functionality of computers and interact with them more effectively.

Simultaneously, metaphor serves as a powerful mechanism for shaping user expectations and mediating experiences with AI systems. Depending on the user's intention, a Chatbot can be metaphorically personified as a friend, a psychotherapist, a personal assistant, and so on. Metaphorical personification serves the purpose of providing users with a visual representation of the abstract entity of a Chatbot and can also be employed as a strategy to enhance usability (Khadpe et al., 2020). For instance, research has indicated that Chatbots with a female voice are more effective in delivering information (Nass & Lee, 2007), and introverted Chatbots are better suited for introverted users (Mou & Xu, 2017). Moreover, robots metaphorically personified as animals have demonstrated effectiveness in the education and emotional therapy of elderly individuals and young children (Lee & Cho, 2022; Jung et al., 2022). However, it remains unclear whether these effects are attributable to users perceiving Chatbots as personified entities, and there is a dearth of existing research on this subject matter.

3. Conclusions

This study examines users' perceptions of anthropomorphism in chatbots using the conceptual metaphor theory. The conceptual metaphor theory is a valuable tool for understanding how speakers conceptualize and perceive chatbots. To further advance research in this area, it is crucial to investigate users' cognitive perceptions of chatbots by analyzing real language data and conducting experimental studies.

References

[1] Ammari, T., Kaye, J., Tsai, J. Y., & Bentley, F., "Music, Search, and IoT: How People (Really) Use Voice Assistants," *ACM Transactions on Computer-Human Interaction*, 26(3), 17-1, 2019.

[2] Brandt, M. J., & Reyna, C., "The chain of being: A hierarchy of morality," *Perspectives on Psychological Science*, 6(5), 428-446, 2011.

[3] Breazeal, C., "Toward sociable robots," *Robotics and autonomous systems*, 42(3-4), 167-175, 2003.

[4] Fraune,M.R., Oisted, B. C., Sembrowski, C. E., Gates, K. A., Krupp, M. M., & Šabanović, S., "Effects of robot-human versus robot-robot behavior and entitativity on anthropomorphism and willingness to interact," *Computers in Human Behavior*, 105, Article 106220, 2020.

[5] Jung, J. Y., Qiu, S., Bozzon, A., & Gadiraju, U., "Great chain of agents: The role of metaphorical representation of agents in conversational crowdsourcing," *In Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems,* pp. 1-22, 2022.

[6] Kuzminykh, A., Sun, J., Govindaraju, N., Avery, J., & Lank, E., "Genie in the bottle: Anthropomorphized perceptions of conversational agents," *In Proceedings of the 2020 CHI conference on human factors*

in computing systems, pp. 1-13, 2020.

[7] Khadpe, P., Krishna, R., Fei-Fei, L., Hancock, J. T., & Bernstein, M. S., "Conceptual metaphors impact perceptions of human-AI collaboration," *Proceedings of the ACM on Human-Computer Interaction,4(CSCW2)*, 1-26, 2020.

[8] Krikmann, A., "the great chain of being as the background of personificatory and depersonificatory metaphors in proverbs and elsewhere", *In Kertas kerja dalam 1st Interdisciplinary colloquium on proverbs*, Tavira, Portugal. 2007.

[9] Lakoff, G., Johnson, M., "The metaphorical structure of the human conceptual system", *Cognitive science*, 4(2), pp. 195-208, 1980.

[10] Lakoff, G., Women, fire, and dangerous things: What categories reveal about the mind, University of Chicago press. Volume 8 Issue 6, Pages.911-919, 1987.

[11] Lee, J. E. R., & Nass, C. I., "Trust in computers:The computers-are-social-actors (CASA) paradigm and trustworthiness perception in human-computer communication,*" In Trust and technology in a ubiquitous modern environment: Theoretical and methodological perspectives*(pp. 1-15). IGI Global, 2010.

[12] Lovejoy, A. O.(1963), the great chain of being, Harvard University Press.

[13] Nass, and K. M. Lee, "Does computer-synthesized speech manifest personality? Experimental tests of recognition, similarity-attraction, and consistency-attraction," *Journal of experimental psychology: applied*, vol.7, no.3, pp.171-181, 2007.

[14] Pradhan, A., Findlater, L., & Lazar, A., " Phantom Friend" or" Just a Box with Information" Personification and Ontological Categorization of Smart Speaker-based Voice Assistants by Older Adults. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW), 1-21. 2019.

[15] Samuel, J. L., "Company from the uncanny valley: A psychological perspective on social robots, anthropomorphism and the introduction", *Ethics in Progress*, 10(2), 8-26, 2019.

[16] Santibáñez, F., "Constraints on metaphor: some notes on the role of the invariance principle in metaphoric mappings", *Revista española de lingüística aplicada*, 13, pp. 177-188, 1999.

[17] Trovato, G.,Ramos, J. G.,Azevedo, H.,Moroni, A.,Magossi, S.,Simmons, R.,…Takanishi, A., "A receptionist robot for Brazilian people: Study on interaction involving illiterates. Paladyn", *Journal of Behavioral Robotics*, 8(1), 1–17. 2017.

[18] Y. Mou, and K. Xu, "The media inequality: Comparing the initial human-human and human-AI social interactions," *Computers in Human Behavior*, vol.72, pp.432-440, 2017.

Copyright ⓒ 2023 KSII

The application of affordance theory in understanding social media users' perception of algorithms

Jisu Kim
Business, Communication and Design Cluster, Singapore Institute of Technology
Singapore
[e-mail: jisu.kim@singaporetech.edu.sg]
*Corresponding author: Jisu Kim

Abstract

This paper explains affordance theory and discusses how the theoretical framework can be used in understanding social media users' perception of algorithms. Due to the opacity of social media algorithms, users have a limited understanding, and thus have different perceptions and attitudes toward these algorithms. Affordance theory can be useful in explaining these differences, given that the theory assumes differences in every use of technology based on users' goals and their capabilities in using the technology.

Keywords: Algorithms, Affordance Theory, Platforms, Technology, Social Media, User Perception

1. Introduction

Social media algorithms have been used by different social media platforms to sort and filter content on their main feeds according to their predetermined rules. Most content that social media users see on platforms is chosen by their algorithms, and thus social media algorithms play a crucial role in structuring our online experience.[1] However, due to the opacity of social media algorithms, our understanding of platforms is limited, and users have different perceptions and attitudes toward algorithmic decision-making on social media.[2]

Although a few scholars have examined how social media users make sense of and engage with algorithms based on different concepts and theories,[2,3,4] we still do not have a clear concept or theoretical framework to understand how social media users have different perceptions of social media algorithms. Thus, this paper proposes how affordance theory can be applied in understanding user interaction with social media algorithms.

2. The Affordance Theory

2.1 The Affordance Theory Framework and Its Usefulness

The concept of affordance was first proposed by Gibson.[5] Originally, it refers to the action possibilities available in the environment regardless of users' intentions or perceptions of using the technology. Later, scholars differentiated perceived versus actual affordances and focused more on the relational aspect of affordance.[6] That is, the concept of affordance is referred to as possibilities for goal-oriented action afforded by the technology to a specific user.[6] Different from designs and mere characteristics of technology, the concept of affordance reflects users' intention to use the technology based on their goals and technological capabilities that make a particular action possible. Thus, even though social media users use the same platforms, social media affordance could be different, depending on their goals. For example, for the purpose of knowledge creation, Wagner et al.[7] identified

reviewability, experimentation, editability, and authoring as social media affordances, whereas Sæbø et al.[8] discovered adapting rules, making decisions collectively, and circulating information as affordances for collective action.

Affordance theory proposes a useful framework in understanding how and when users perceive and actualize the potential action that a technology affords them.[9] In its framework, there are four distinct phases for perceiving and actualizing affordances: affordance existence, affordance perception, affordance actualization, and effects. First, *affordance exists* based on the interaction between users and the technology. When users start to become aware of some of these affordances, we call this *affordance perception*. Users' perception of affordances would be affected by their goals, capabilities in using the technology, and information about the affordance existence. Users perceive affordances by learning from other users or external sources, or by interpreting cues from the technology that signals the existence of affordances.[10] Then, users engage in various processes to actualize these perceived affordances. *Affordance actualization* refers to the actions taken by users to employ perceived affordances through their use of the technology to achieve their goals.[9,10] After actualizing these affordances, there would be consequences attributed to the actualization of the affordances.[6,9]

This theoretical framework helps examine the relationships between users' perception of affordances and subsequent behaviors and (potential) consequences. The framework can also be employed to explain why there are differences in user perceptions and the use of technological affordances, including users' goals, capabilities, and so on.

2.2. Application of Affordance Theory to Users' Perception of Algorithms

As previously mentioned, general social media users have a limited understanding of social media algorithms due to their technical sophistication and non-transparency. For this reason, social media users have different perceptions and attitudes toward social media algorithms.

Since these differences are based on users' goals to use the platforms and their ability and capacities in obtaining information about social media algorithms, affordance theory can provide a useful framework to explain the dynamic nature of user perceptions and behaviors related to social media algorithms.

First, affordance theory can explain how users' goals to use a platform would affect their perception and actualization of the affordances. For example, social media users' actualization behaviors can differ when they use the platform for their own enjoyment or for monetary reasons. Second, affordance theory helps differentiate users' perception of affordances from actualization behaviors and examines the relationships between them. The differentiation between how social media users understand algorithms and their behaviors in taking advantage of algorithms for their goal achievement is crucial: the former would be an important factor for the latter. Without the theoretical framework, the two have not been clearly conceptualized or operationalized. Lastly, affordance theory helps us understand platform differences in users' perception of algorithms based on the relationships between users' goals to use certain platforms and the platform's particular technical features.

3. Conclusions

In this paper, the author proposes how affordance theory could be useful in understanding users' perception of social media algorithms. Thus far, most scholars have used qualitative approaches for affordance theory; however, there have been some efforts to use a mixed-method or quantitative approach to examine the theory. As new and emergent technology continues to appear in our lives (e.g., generative AI, VR/AR platforms, etc.), this theoretical framework can be more useful in understanding different perceptions and behaviors related to new technology.

References

[1] T. Gillespie, "The Relevance of Algorithms," *Media Technologies: Essays Communication, Materiality, and Society*,

pp. 167-194, 2014.
[2] J. Swart, "Experiencing algorithms: How Young People Understand, Feel About, and .Engage With Algorithmic News Selection on Social Media," *Social Media + Society*, Vol 7, No. 2, pp. 1-12, 2021.
[3] S. Min, "From algorithmic disengagement to algorithmic activism: Charting social media users' responses to news filtering algorithms," *Telematics and Informatics*, Vol 43, pp. 101-251, 2019.
[4] M. A. DeVito, "How transfeminine TikTok creators navigate the algorithmic trap of visibility via folk theorization," in *Proc. of the ACM on Human-Computer Interaction*, Vol 6, pp. 1-31, 2022
[5] J. J. Gibson, "The Ecological Approach to Visual Perception," Houghton Mifflin, 1979.
[6] M.L. Markus and M.S. Silver, "A Foundation for the Study of IT Effects: A New Look at DeSanctis and Poole's Concepts of Structural Features and Spirit," *Journal of the Association for Information Systems*, Vol 9, pp. 609-632, 2008.
[7] D. Wagner, G. Vollmar, and H.-T. Wagner, "The impact of information technology on knowledge creation: An affordance approach to social media," *Journal of Enterprise Information Management*, Vol 27, No. 1, pp. 31-44, 2014.
[8] Ø. Sæbø, T. Federici, and A. M. Braccini, "Combining social media affordances for organising collective action," *Information Systems Journal*, Vol 30, No. 4, pp. 699-732. 2020.
[9] E. Bernhard, J. Recker, and A. Burton-Jones, "Understanding the actualization of affordances: A study in the process modeling context," in *Proc. of the 34th International Conference on Information Systems (ICIS 2013)*, 2013.
[10] T. Lehrig, O. Krancher, and J. Dibbern, "How Users Perceive and Actualize Affordances: An Exploratory Case Study of Collaboration Platforms," in *Proc. of the 38th International Conference on Information Systems*, 2017.