Com SIS
Computer Science and Information Systems
Volume 21, Number 2, April 2024

Contents

ISSN: 2406-1018 (Online)

Computer Science and Information Systems

Vol 21, No 2, April 2024

Com SIS

Computer Science and Information Systems

Published by ComSIS Consortium

Special Issue on Deep Learning Techniques in Intelligent Internet of Things and 5G Communication Networks

Volume 21, Number 2
April 2024

ComSIS is an international journal published by the ComSIS Consortium

ComSIS Editorial Office:
University of Novi Sad, Faculty of Sciences,
Department of Mathematics and Informatics
Trg Dositeja Obradovića 4, 21000 Novi Sad, Serbia
Phone: +381 21 458 888; **Fax:** +381 21 6350 458
www.comsis.org; Email: comsis@uns.ac.rs

Volume 21, Number 2, 2024
Novi Sad

Computer Science and Information Systems

Special Issue: Deep Learning Techniques in Intelligent Internet of Things and 5G Communication Networks

ISSN: 2406-1018 (Online)

The ComSIS journal is sponsored by:

Ministry of Education, Science and Technological Development of the Republic of Serbia
http://www.mpn.gov.rs/

Computer Science and Information Systems

AIMS AND SCOPE

Computer Science and Information Systems (ComSIS) is an international refereed journal, published in Serbia. The objective of ComSIS is to communicate important research and development results in the areas of computer science, software engineering, and information systems.

We publish original papers of lasting value covering both theoretical foundations of computer science and commercial, industrial, or educational aspects that provide new insights into design and implementation of software and information systems. In addition to wide-scope regular issues, ComSIS also includes special issues covering specific topics in all areas of computer science and information systems.

ComSIS publishes invited and regular papers in English. Papers that pass a strict reviewing procedure are accepted for publishing. ComSIS is published semiannually.

Indexing Information

ComSIS is covered or selected for coverage in the following:

- Science Citation Index (also known as SciSearch) and Journal Citation Reports / Science Edition by Thomson Reuters, with 2022 two-year impact factor 1.4,
- Computer Science Bibliography, University of Trier (DBLP),
- EMBASE (Elsevier),
- Scopus (Elsevier),
- Summon (Serials Solutions),
- EBSCO bibliographic databases,
- IET bibliographic database Inspec,
- FIZ Karlsruhe bibliographic database io-port,
- Index of Information Systems Journals (Deakin University, Australia),
- Directory of Open Access Journals (DOAJ),
- Google Scholar,
- Journal Bibliometric Report of the Center for Evaluation in Education and Science (CEON/CEES) in cooperation with the National Library of Serbia, for the Serbian Ministry of Education and Science,
- Serbian Citation Index (SCIndeks),
- doiSerbia.

Information for Contributors

The Editors will be pleased to receive contributions from all parts of the world. An electronic version (LaTeX), or three hard-copies of the manuscript written in English, intended for publication and prepared as described in "Manuscript Requirements" (which may be downloaded from http://www.comsis.org), along with a cover letter containing the corresponding author's details should be sent to official journal e-mail.

Criteria for Acceptance

Criteria for acceptance will be appropriateness to the field of Journal, as described in the Aims and Scope, taking into account the merit of the content and presentation. The number of pages of submitted articles is limited to 20 (using the appropriate LaTeX template).

Manuscripts will be refereed in the manner customary with scientific journals before being accepted for publication.

Copyright and Use Agreement

Computer Science and Information Systems

Volume 21, Number 2, April 2024

CONTENTS

Computer Science and Information Systems 21(2):i–v https://doi.org/10.2298/CSIS240200iC

Guest Editorial: Deep Learning Techniques in Intelligent Internet of Things and 5G Communication Networks

Jia-Wei Chang[1], Nigel Lin[2], Qingguo Zhou[3], Yi-Zeng Hsieh[4], Mirjana Ivanovic[5]

[1] Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taichung City, Taiwan
jwchang@nutc.edu.tw

[2] Department of Computer Science and Engineering, Santa Clara University, Santa Clara, CA, USA
nigel@mail.topwise.com

[3] School of Information Science & Engineering, Lanzhou University, Gansu Province, China
zhouqg@lzu.edu.cn

[4] Department of Electrical Engineering, National Taiwan University of Science and Technology, Taipei City, Taiwan
yzhsieh@mail.ntust.edu.tw

[5]University of Novi Sad, Faculty of Sciences, Novi Sad, Serbia

mira@dmi.uns.ac.rs

In the rapidly evolving digital transformation landscape, the synergy between Deep Learning (DL), the Internet of Things (IoT), and 5G communication networks heralds a new era of technological innovation. This guest editorial delves into the pivotal role of DL in enhancing the capabilities of IoT ecosystems and the performance of 5G networks, thereby paving the way for a more intelligent, more connected world. The advent of IoT has brought about a paradigm shift in how devices communicate, collect, and process data. With billions of connected devices generating vast data, DL techniques are adept at handling and interpreting the complexity and volume of IoT data, enabling advanced analytics, decision-making, and automation. In the context of IoT, DL facilitates the realization of truly intelligent systems. The integration of IoT with 5G communication networks further amplifies these benefits. 5G, known for its high speed, low latency, and massive connectivity, is a perfect match for the IoT, providing the necessary infrastructure for seamless data transmission. DL algorithms enhance 5G network management by optimizing resource allocation, improving network security, and facilitating the efficient handling of the increased data traffic by IoT devices. In conclusion, DL, IoT, and 5G convergence hold tremendous potential for transforming various industries. As we stand on the brink of this technological revolution, it is imperative to navigate the associated challenges wisely, ensuring that the benefits of these advanced technologies are realized securely and efficiently.

This special issue received 49 submissions where the corresponding authors were majorly counted by the deadline for manuscript submission with an open call-for-paper. All these submissions are considered significant in the field, but however, only one-third of them passed the pre-screening by guest editors. The qualified papers then went through double-blinded peer review based on a strict and rigorous review policy. After a totally three-round review, 13 papers were accepted for publication. A quick overview to the papers in this issue can be revealed below, and we expect the content may draw attentions from public readers, and furthermore, prompt the society development.

The first paper, titled "Implementation of Multimedia Search & Management System Based on Remote Education," by Byeongtae Ahn, addresses the need for efficient management and retrieval of video information in remote education. Highlighting the critical role of real-time processing of compressed video data, it introduces a system leveraging MPEG-4, the leading video compression standard. The paper develops a management and search solution designed explicitly for multimedia in distance learning, emphasizing the importance of MPEG-4 compression for real-time video handling. This work contributes significantly to the field by enhancing the accessibility and effectiveness of video resources in educational environments.

The second paper, titled "Automatic Voltage Stabilization System for Substation using Deep Learning," by Jiyong Moon et al., introduces an innovative solution to automate voltage regulation, which is traditionally reliant on manual intervention and prone to inefficiencies. By employing a deep learning approach with a stacked LSTM model, the system predicts the necessary input capacity for stabilization, overcoming the uncertainties of human-based regulation and enhancing operational efficiency with economic considerations. It further optimizes regulation plans and incorporates a user interface for algorithm operation visualization and model prediction communication. Tested with real substation data, the findings reveal the system's capability to significantly improve the automation of the voltage regulation process, marking a notable advancement in power facility management.

The third paper, titled "The Effects of Process Innovation and Partnership in SCM: Focusing on the Mediating Roles," by Yoonkyo Cho et al., explores the influence of supply chain management (SCM) components on organizational performance, highlighting process innovation and partnerships as essential mediators. Analyzing responses from 193 workers in smartphone manufacturing, the study identifies the positive effects of information systems, top management support, and performance management on process innovation and the fostering of partnerships. These elements, in turn, significantly enhance both the financial and non-financial outcomes for firms. The findings suggest that bolstering process innovation and partnerships is crucial for advancing a firm's SCM efficiency, offering insights into leveraging these dynamics in the context of Industry 4.0's technological shifts.

The fourth paper, titled "Navigation Control of Autonomous Ackerman Robot Using a Lidar-sensing-based Fuzzy Controller in Unknown Environments," by Cheng-Jian Lin et al., introduces a novel lidar-sensing-based navigation control system for autonomous Ackerman robots operating in uncharted territories. Utilizing a behavioral controller, this system enables effective obstacle avoidance and goal-directed movement without reliance on global map data. A Wall-Following Fuzzy Controller's core mechanism processes lidar-derived distance measurements to adjust the robot's steering angle, ensuring safe passage through diverse settings without collisions. Additionally, a

specialized escape strategy has been incorporated to circumvent potential endless looping. Experimental validation in simulated and real-world scenarios confirms the system's proficiency in guiding Ackerman robots through unfamiliar environments, highlighting its practical utility and efficiency.

The fifth paper, titled “Application of Item Response Theory and the revised Girvan–Newman Clustering for Estimating Learning Ability in Cooperative Programming Learning,” by Wen-Chih Chang, explores an innovative approach to enhance programming education through cooperative learning. Recognizing the wide range of student abilities in comprehending complex programming concepts, this study introduces a novel grouping methodology that combines item response theory with social network analysis clustering. This method strategically groups students by their learning abilities, aiming to optimize educational outcomes. The effectiveness of this approach was empirically tested in a programming course for beginners, with results indicating significant improvements in learning achievements. This paper presents a promising direction for tailoring cooperative learning experiences better to meet the diverse needs of students in programming education.

The sixth paper titled “A Study of Identity Authentication Using Blockchain Technology in a 5G Multi-Type Network Environment,” by Jui-Hung Kao, examines the application of blockchain for identity authentication within the nuanced landscape of 5G networks. Highlighting 5G's potential for facilitating rapid digital transformation through its low latency and ability to support a multitude of connections, this study addresses the challenge of limited indoor penetration by integrating 5G with Wi-Fi 6 for enhanced mobile connectivity. The paper proposes an innovative authentication framework utilizing Mobile Edge Computing and blockchain to manage access in a 5G Local Breakout network, ensuring secure and efficient user authentication across 5G and Wi-Fi 6 networks. Real-world validation confirms the effectiveness of this approach in improving user access control and network service quality, promising advancements in mobile network security and user experience through edge computing and blockchain technologies.

The seventh paper titled “An Empirical Study on Success Factors of the Game Industry,” by Jun-Ho Lee et al., explores the dynamic growth of Korea's game industry, particularly its expansion in China and Southeast Asia. This research delves into how the interplay of management, technology, market, and industry characteristics influences the success of Korean game companies both domestically and internationally. Through empirical analysis, it identifies key factors such as cutting-edge technological advancements, managerial competencies, market trends, and industry insights as pivotal to achieving growth and global market penetration. Moreover, the study highlights the critical role of intellectual property rights in sustaining performance and facilitating market expansion. Distinguishing itself from prior work that focused mainly on the external impacts of games, this study offers a holistic view of the internal workings, market dynamics, and industry strategies, underscoring the multifaceted approach needed for game companies to thrive.

The eighth paper titled “Design of TAM-based Framework for Credibility and Trend Analysis in Sharing Economy: Behavioral Intention and User Experience on Airbnb as an Instance,” by Yenjou Wang et al., investigates the pivotal role of trust in the sharing

economy, using Airbnb as a case study. Addressing the inherent uncertainties of pre-purchase conditions in such a market, this research employs the Technology Acceptance Model to identify factors influencing consumer behavior and intentions. Through a comprehensive three-year survey and data collection from Airbnb users, the study applies Partial Least Squares-Structural Equation Modeling for hypothesis testing. It further explores the effects of user experience variations on trust and purchasing intentions via Multi-Group Analysis, revealing that Airbnb's ease of use significantly shapes consumer attitudes more than any specific platform information, thereby positively affecting overall behavioral intentions. This work underscores the importance of trust in the sharing economy and highlights the critical impact of user experience on consumer engagement and platform credibility.

The ninth paper titled "Robust Compensation with Adaptive Fuzzy Hermite Neural Networks in Synchronous Reluctance Motors," by Chao-Ting Chu et al., introduces an innovative robust compensation scheme for synchronous reluctance motors (SRMs) utilizing adaptive fuzzy Hermite neural networks (RCAFHNN). Addressing the challenges posed by parameter variations, external disturbances, and nonlinear dynamics inherent in SRMs, this study leverages the adaptive neural fuzzy interface system (ANFIS) framework to refine motor control. RCAFHNN distinguishes itself through three primary advancements: incorporation of fuzzy logic and neural network-based online estimation for dynamic adjustment, the adoption of Hermite polynomial functions to expedite membership function training, and the assurance of system convergence and robustness through Lyapunov stability analysis. Experimental comparisons between RCAFHNN and traditional ANFIS approaches demonstrate RCAFHNN's enhanced performance, marking a significant step forward in precise motor control technologies.

The tenth paper titled "Machine Learning Based Approach for Exploring Online Shopping Behavior and Preferences with Eye Tracking," by Zhenyao Liu et al., investigates the evolving landscape of consumer behavior in the digital age, particularly the shift towards online shopping accelerated by the COVID-19 pandemic. This research integrates eye-tracking technology to understand better how visual stimuli influence online shopping decisions. By analyzing the eye movements of 60 participants engaged in online shopping activities, the study leverages statistical and machine learning techniques to examine the impact of visual complexity and consumer preferences on purchasing behavior. The findings reveal that when analyzed with machine learning algorithms, eye-tracking data can effectively predict consumer choices and improve e-commerce recommendation systems. The research also differentiates between hedonic and utilitarian purchasing behaviors, noting distinct patterns in visual attention. This study provides valuable insights for enhancing e-commerce platforms and tailoring marketing strategies to meet consumer needs better.

The eleventh paper, titled "A Novel Multipath QUIC Protocol with Minimized Flow Complete Time for Internet Content Distribution," by Lin Hui, addresses the challenges of scaling Internet content distribution efficiently amidst surging data flows. It critically evaluates the Quick UDP Internet Connections (QUIC) protocol, renowned for enhancing media transfer through flow-controlled streams, reduced latency in connection setup, and flexible network path migration. Despite QUIC's advancements over TCP in connection and transmission efficiency, its performance is often bottlenecked by the bandwidth limitations and variability of single network paths. This study introduces an innovative multipath QUIC strategy designed to leverage multiple

network paths concurrently to optimize bandwidth usage and circumvent congestion. Unlike previous approaches that rely on simplistic round-robin or shortest-time-first data scheduling, this method applies a sophisticated algorithm considering path delay and packet loss rate, significantly improving flow completion times. The proposed scheme demonstrates marked superiority in experimental comparisons with conventional QUIC, Lowest-RTT-First (LRF) QUIC, and Pluginized QUIC (PQUIC), offering a promising avenue for enhancing the robustness and efficiency of internet content distribution networks.

The twelfth paper titled "A study on how to augment fire data from video/image using the Similar-label and F-guessed method," by Jong-Sik Kim et al., tackles the challenge of enhancing fire detection capabilities with limited datasets. In fire detection, where data scarcity often hampers the improvement of detection rates, the research delves into semi-supervised learning as a solution, acknowledging its effectiveness yet highlighting the pitfalls of pseudo-label methods that can introduce false labels and biases. To counteract these issues, the study introduces a novel approach that generates similar-labeled data during the initial learning phase using the F-guessed method combined with the Region of Interest (ROI) expression in videos. This technique aims to maintain accuracy in label distribution, preventing the introduction of bias early on. The methodology proved substantially effective, enlarging the dataset by approximately 6.5 times, from 5,565 to 41,712 entries, and significantly enhancing the mean Average Precision (mAP@0.5) by 26.1%, from 65.9% to 92.0%, while also improving the loss from 3.347 to 1.69. This innovative approach presents a significant advancement in the field of fire detection research, offering a scalable and more accurate method for data augmentation and model training.

The thirteenth paper titled "Multi-language IoT Information Security Standard Item Matching based on Deep Learning," by Yu-Chih Wei, addresses the complexity of navigating through various information security standards applicable to IoT and other domains, such as ISO/IEC 27001 and the IEC 62443 series. With the proliferation of standards, the task of identifying and matching specific controls relevant to particular scenarios has become increasingly challenging and labor-intensive. This paper introduces a novel approach that leverages text mining and deep learning techniques to analyze and match similar control items across different security standards, regardless of language barriers. By utilizing translations of domestic and international standards as a foundation, the study aims to streamline the process of finding correspondences between controls, facilitating a more efficient implementation and research of information security standards. This method promises to significantly reduce the effort required to compare and locate applicable controls, thus enhancing the security posture of businesses and organizations in the rapidly evolving digital landscape.

Acknowledgments. The guest editors are thankful to authors who submitted interesting and challenging papers, to reviewers for their effort in reviewing the manuscripts and inspiring authors to improve quality of their papers. We also thank the Editor-in-Chief, Prof. Mirjana Ivanovic, and editorial assistants for their supportive guidance and help during the entire process of preparation this special issue.

Computer Science and Information Systems 21(2): vii–vii https://doi.org/10.2298/CSIS240200viiE

CORRIGENDUM

Mirjana Ivanović

University of Novi Sad, Faculty of Sciences
Novi Sad, Serbia
mira@dmi.uns.ac.rs

The authors of the article: *Samarbakhsh, L., Tasić, B.: What makes a board director better connected? Evidence from graph theory. Computer Science and Information Systems, Vol. 17, No. 2, 357–377. (2020), https://doi.org/10.2298/CSIS190628045S* have informed the Editorial Office that they missed acknowledging two facts:

1. The authors Dr. Laleh Samarbakhsh and Dr. Boza Tasic would like to thank Ted Rogers School of Management for funding support as part of the TRSM Research Development Grant.
2. The authors would like to acknowledge the contribution of Dr Hamid Ebrahimi and ask that he be added as a third co-author to the paper. Their decision is based on a recent reflection on the unique circumstances regarding Dr Ebrahimi's involvement in the paper.

Therefore, the journal is publishing this Corrigendum. The authors of this article should be listed as follows: *Laleh Samarbakhsh, Bože Tasić, Hamid Ebrahimi.*

Computer Science and Information Systems 21(2):419–436 https://doi.org/10.2298/CSIS220509007A

Implementation of Multimedia Search & Management System Based on Remote Education*

Byeongtae Ahn

Liberal & Arts College, Anyang University, 22, 37-Beongil, Samdeok-Ro, Manan-Gu, Anyang 430-714, South Korea
ahnbt@ anyang.ac.kr

Abstract. In order for remote education using multimedia to be effective, an efficient management technique for video information needs to be developed. Therefore, for real-time processing of moving images, it is necessary to manage and search image data in a compressed state. MPEG-4 is the most widely used video compression technology. In order to process video in real time in distance education using multimedia, it is very important to develop a technique for managing and retrieving video information compressed with MPEG-4. Therefore, in this paper, a multimedia information management system and search technology were developed using MPEG-4 compression technology used for real-time distance education.

Keywords: Multimedia, MPEG, Remote Education, Video search, Compressed video.

1. Introduction

With the recent development of the Internet and the Web, the demand for multimedia, especially video information, is rapidly increasing. As object-oriented database-based multimedia database systems are being developed, they are in the stage of utilizing them in various multimedia authoring systems. Among them, many studies are being conducted on the storage and retrieval of multimedia information, especially video information [1].

However, the management of moving picture information using such a multimedia DBMS is a method of managing the search target bitmap or wave pattern in an uncompressed state. However, due to the nature of video, it is difficult to store, retrieve, or transmit uncompressed natural video as it is [2].

Therefore, in order to solve these problems and put it into practice in the video management system, a technology that compresses and stores video information, and searches and transmits the video information in real time in a compressed state, is required [3].

In this paper, we develop a compressed video management system that compresses video information with MPEG-4 technology, stores it in a database system, and searches

* The paper is an extended version of a conference paper(https://www.fronticomp.com/ic2022-metaverse).

using query words or representative images (key frames). In addition, annotation-based search and content-based search for video information search were used. Content-based search can automatically extract shape, texture, and movement from data features. Annotation-based search uses natural language processing to extract and provide semantic information of video data. This technique is modeled in various ways by easily extracting various features of moving picture data [4-6].

In this paper, we propose an Integrated Video Data Model (IVDM) that can use both annotation-based search and content-based search by analyzing general video data for real-time video search. This study made it possible to search various real-time images. And to improve the real-time image processing capability, MPEG-4 technology was applied to improve the compression capability.

This paper designed and implemented Compressed Video Information Management System (CVIMS) using Mpeg-4 compression technology. CVIMS consists of 3 layers. Layer 1 consists of user interface and layer 2 consists of video processing. Video processing is classified into movie display, caption/figure description editor, and query processor. Layer 3 consists of DBMS to store data.

In Section 2 of this paper, a related study is proposed, and in Section 3, an improved video data model is proposed. In Section 4, we designed a compressed image data management system, and in Section 5, we implemented the system. In Section 6, an integrated data model system was proposed, and in Section 7, a new image scheme was designed. Finally, in Section 8, conclusions and future tasks are presented.

2. Related Studies

Methods for storing and retrieving video data can be broadly divided into content-based search and annotation-based search.

Content-based search is a method of searching for the meaning of video data by extracting color, shape, and movement from each frame of the video, and searching based on this. Although this method shows good search results for a specific domain, it is difficult to extract the general meaning contained in the video data, and in the case of a compressed map image, it is inefficient in terms of performance because it must be decompressed and searched for image extraction [7-10].

The video data is largely composed of image, audio, and writing data. The image is what allows the listener to see the instructor's face. In this case, the visual effect can be increased. Audio and image data processing uses a compressed file after compression using a multimedia compression tool. For other writing data, when the lecturer draws a line or draws a rectangle on the screen, the actions are expressed by objectifying them. The act of changing the currently active page is also made possible to be expressed as a single object.

Annotation-based search is a method in which a person first grasps the meaning of video data, expresses it using natural language, and searches based on this. This method makes it possible to easily model various meanings of video material that are difficult to find with an automated method, and make them available for search. On the other hand, it is easy to lose the consistency of the video material because annotations can be given or interpreted differently depending on the user's point of view. In particular, it becomes

more and more difficult to maintain consistency when it is intended to give a very detailed comment rather than a comprehensive comment [11-14].

Therefore, it is necessary to find a way to integrate these two techniques. At this time, in order to support the two search methods in a form suitable for the user's needs, it is necessary to develop an integrated data model above all else.

Recently, a multi-layered video model (MLVD) has been proposed for a search that integrates these two techniques. The MLVM model maintains independence for each layer and implements a query processor that does not depend on a specific method of content-based or annotation-based search, and suggests a model related to video data search. However, this paper proposes an integrated model for video management, accepts a part of the MLVM model in the search, and presents a method to approach the user's needs even though it has a step-by-step dependency. In this paper, a general video data model is proposed for efficient management of video documents, and research has been conducted on the development of an MPEG-4 compressed video document management system that supports only annotation-based search based on this [15-18].

However, in this paper, an integrated video data model that supports and manages annotation-based search and content-based search at the same time is presented, and based on this model, a system for managing compressed video information using an object-relational database in a client/server environment is developed. do. In this case, a plug-in technique is also used for use on the web.

3. Extended Video Data Model

Movies are stored in a movie database as successive groups of frames called storage movie segments. Thus, a drawing image is represented by a video stream mapped into one or more stored video segments [19-22].

IVDM was created with the concept of structural components related to the semantic units of a moving picture document. The concept of structural component is subdivided into compound unit, sequence, scene and shot, and these subclasses are defined in a hierarchical relationship with each other. A shot consists of one or more consecutive frames, and appears as a temporal and spatial sequence of actions. A scene is made up of several sets, and a sequence is made up of these scenes. A collection of related sequences constitutes a compound unit again, and the compound unit can refer to itself at an arbitrary level. The video search structure is divided into two stages: a stage that supports annotation-based search and a stage that supports content-based search.

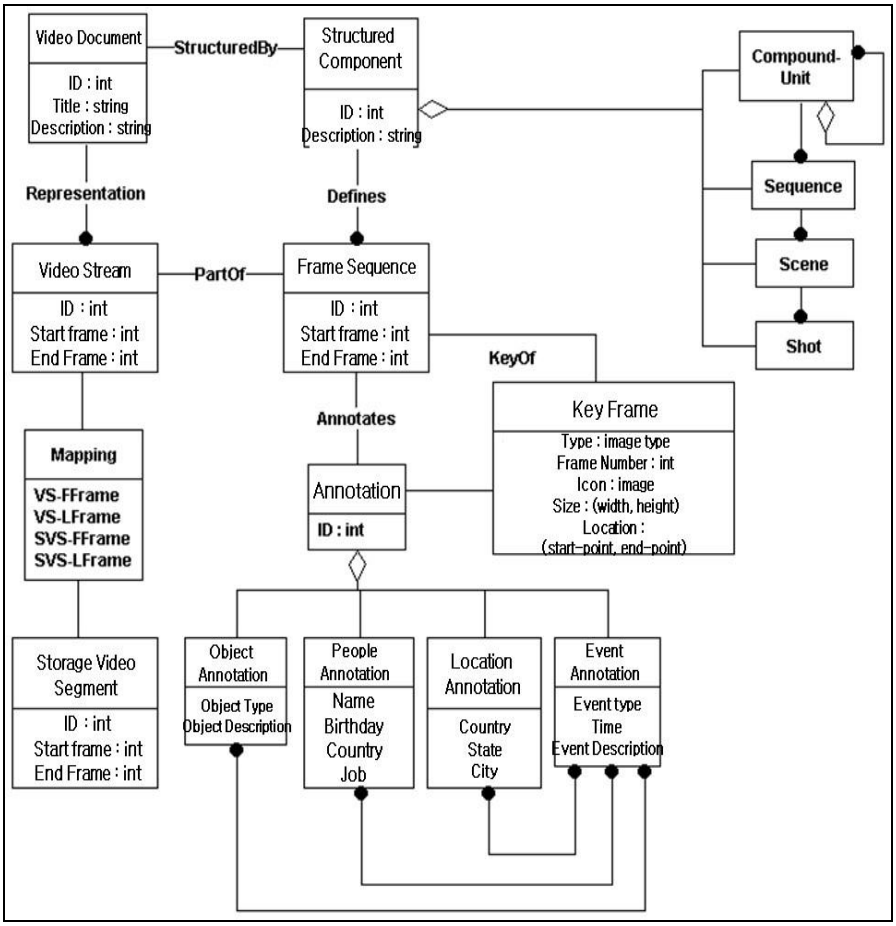


Fig. 1. Data Model of Video Information using Object Diagram

Fig. 1 shows a data model of moving picture information using an object diagram. The object diagram (OMT) shown in Fig. 1 was supplemented by adding a key frame management module and use of Dublin Core metadata [23].

(1) Key frame management module

(2) Utilize Dublin Core metadata for object annotation management

The Enhanced Generic Video Data Model (EGVDM) is a frame that provides functions for structuring video data, free annotation of video data, and sharing and reusing video data. it is work in EGVDM, moving picture data is a continuous group of frames called stored video segments.

The frame sequence in Fig. 1 is classified into an annotation object and a key frame object. A key frame object extracts a specific representative image from a frame sequence and consists of image, image type, frame number, size, and location information about it. Annotation objects include object annotation, person annotation, and location annotation. It consists of subclasses of location annotation and event

annotation. Object annotation consists of object type and object description. In this paper, object annotation is defined using Dublin Core-based metadata. That is, an object annotation is defined by a title, a subject, an identifier, a relation, a right, a language, a document format, and the like [24].

4. Design of Compressed Video Information Management System (CVIMS)

Compressed video information management that extracts key frames from MPEG-4 compressed video data based on the video data model presented above, adds captions and picture descriptions, and stores them in the database in text format for management The system (CVIMS) was designed [25].

4.1. Index Structure for MPEG-4

MPEG-4 compressed video files are mainly composed of three types: I-frames, P-frames, and B-frames [26]. A double I-frame is a frame compressed using only a spatial compression technique without using a temporal compression technique. Therefore, since the I-frame can be independently decoded and can be accessed randomly, it can be a reference frame. Therefore, CVIMS assumes that all I-frames in the MPEG-4 compressed video file can be key frame candidates. Based on this assumption, CVIMS provides a way for users to directly select key frames by extracting I-frames from MPEG-4 compressed video. In addition, the search for each video is not performed by actual frame, but caption information for each key frame is created, structured together with the key frame, stored in the database, and then search is performed using the caption information [27].

Fig. 2 shows the relationship between the index structure and caption for MPEG-4 moving pictures. Technically, after extracting a key frame from a video, it was searched by attaching a caption that processes the contents of the key frame through image recognition.

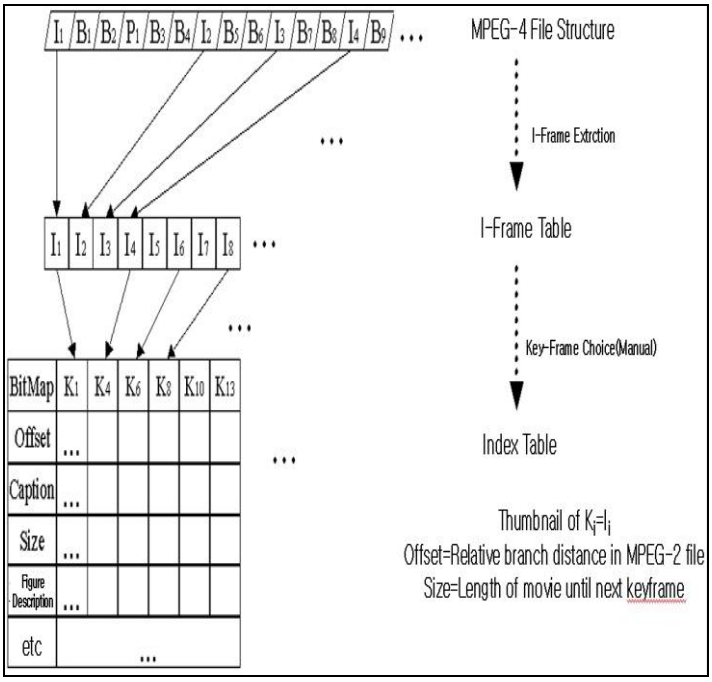


Fig. 2. Structure of MPEG-4 File and Index Information

4.2. Overall Structure of CVIMS

CVIMS includes a user interface and a caption and picture description editor that can index MPEG-4 video, a query processor that processes various user queries, a video display that displays query results, a database that manages index data and video data, and it is composed of a storage server that stores MPEG-4 video [28].

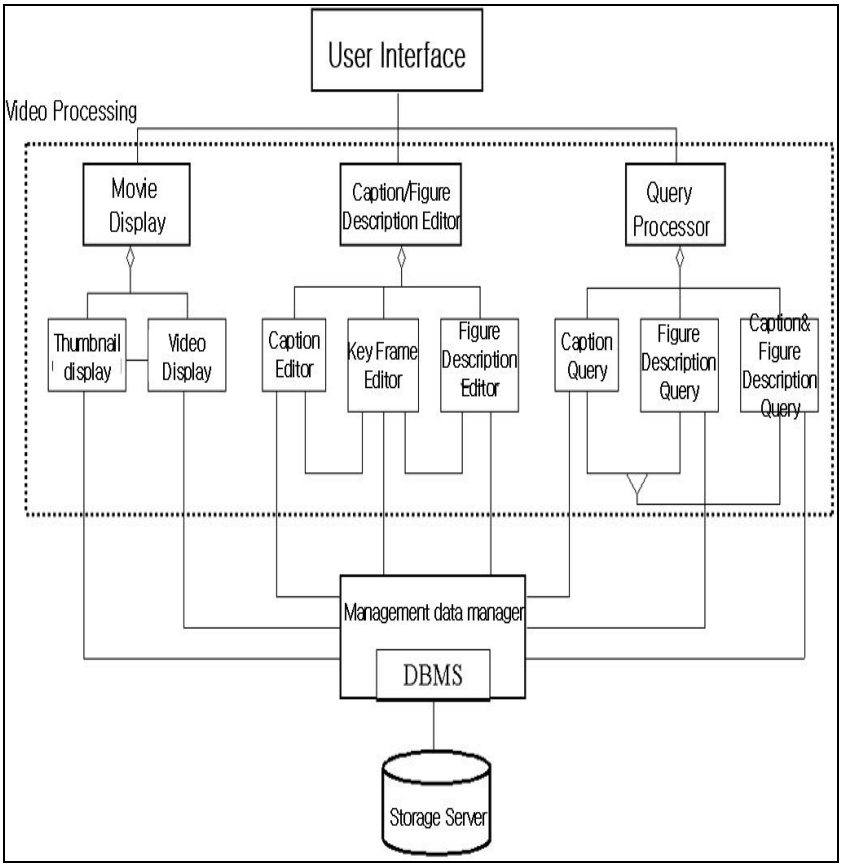


Fig. 3. CVIMS design diagram

Fig. 3 shows the components of CVIMS and subclasses and relationships of each component. VIMS is largely composed of user interface, video processor, and management data manager. The video processor again consists of a query processor, caption/picture description editor, and video display. According to the query type, the query processor consists of a caption query, a picture description query machine, and a query machine that combines captions and picture descriptions. It accepts the user's query, searches each object managed by the management data manager, and brings the desired result. The caption/picture description editor selects keyframes from the list of pre-decoded I-frames, writes caption information and picture description information for each keyframe, and stores them in the database. The video display is a part that displays the query result and is classified into a thumbnail display that outputs an icon in the form of a thumbnail picture and a video display that displays an actual video. Lastly, the management data manager manages the information stored in the database, and manages various index information and caption/picture description information created in the editor [29].

5. Implementation Result of CVIMS

Section 5 shows the implementation result of CVIMS designed in Section 4, focusing on the user interface screen.

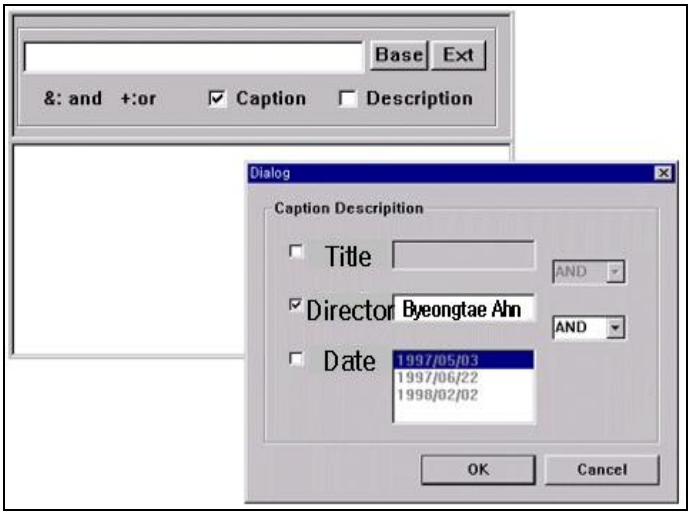


Fig. 4. Keyword Input Screen

Fig. 4 shows a screen in which search conditions and search keywords are entered after clicking basic search in the search window. In this window, set the items to be searched using check boxes and list boxes, and enter keywords for each selected item.

The user's query processing in the search window makes a query to the actual database through the following SQL statement [30].

▶ Simple query

```
  select * from caption_info
where title = search term [and(or) author = search term]
[and(or) madeday = search term]
    select * from picture_desc
where main1_content = search term [and(or)
main2_content = search term] [and(or) content = search word]
```

▶ Complex query

```
    Join Caption_info and Picture_desc tables
```

The results of query processing are shown in Fig. 5.

It appears in the form of a compressed picture as shown in Figure 5.

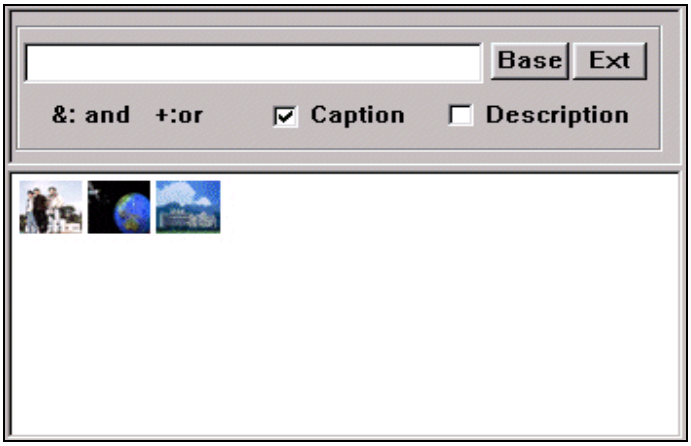


Fig. 5. Output Screen of Retrieval Result

Searching through the web also performs a search using caption information for a video, in the form of first selecting a desired item and then entering a search term for the selected item. At this time, the input search word is transmitted to the database in the form of a query word through the CGI program, and the search result is displayed on the web as an html document. If you click the compressed picture of the video you want here, the video with this thumbnail as a key frame is displayed, and the displayed video system is operated as a Netscape plug-in.

6. Unified Video Data Model (IVDM)

In order to search and manage videos efficiently, it is necessary to share the compressed video itself, related annotations, and image analysis results as an integrated database. To do this, it is necessary to create a general standard model and manage various and vast amounts of compressed video. Therefore, this section proposes an Integrated Video Data Model (IVDM) for video information management. By structuring video data, IVDM supports free annotation-based search for various video data at a high level and content-based search at a lower level.

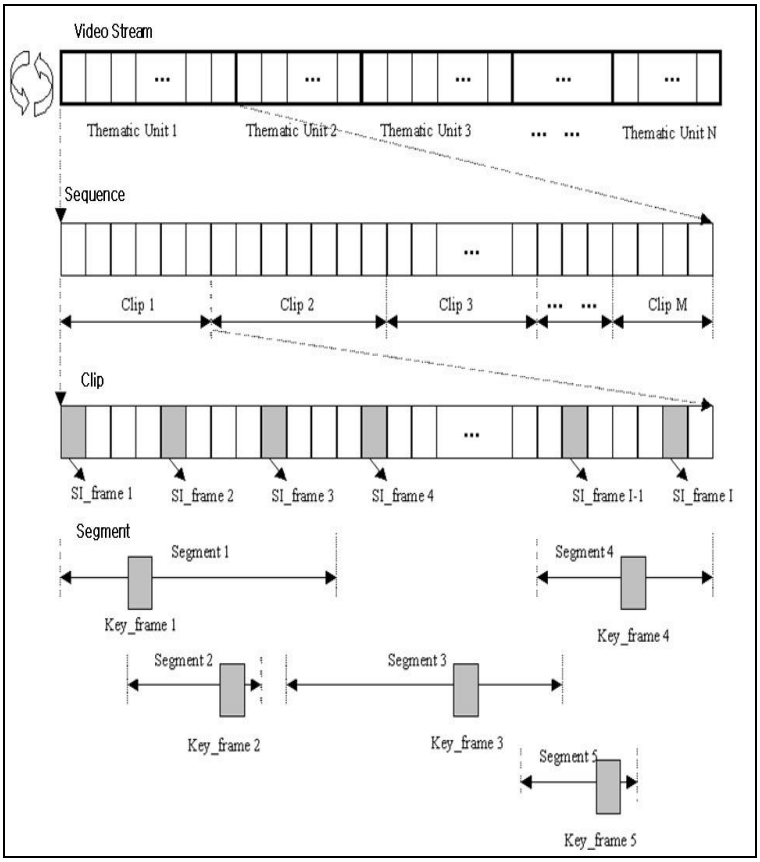


Fig. 6. Partition process for extraction characteristic of video

Fig. 6 shows the process of dividing a general video. The segmentation process is performed on the premise that the video stream belongs to one of these categories when moving pictures are classified into movies, news, dramas, video conferences, and the like. The whole news becomes a video stream, and sharing by topic, event, or reporter becomes a topic unit, and the circular arrow on the left can be repeatedly divided up to several levels. In this case, the number of repetitions may vary depending on the type of video, and the number or size of subject units may also vary depending on the type or subject of the video. In Figure 6, the smallest subject unit is expressed as a sequence. In connection with the previous example, the sequence becomes the content of a reporter's coverage. The sequence is again divided into scenes, where the scene corresponds to the part divided according to whether the reporter's coverage is a simple incident scene or an interview scene. Frames are extracted at regular intervals from this scene to search the flow order of moving pictures, that is, time dimension, and are called SI(Same Interval)_frames. For spatial-dimensional search, each scene is divided into segments where the target object exists, and the frame in which the target object appears most clearly is used as the key-frame. In the extracted SI_frame, the movement of the camera

or object is analyzed, and in the key-frame, color, shape, texture, etc. are analyzed and used for search.

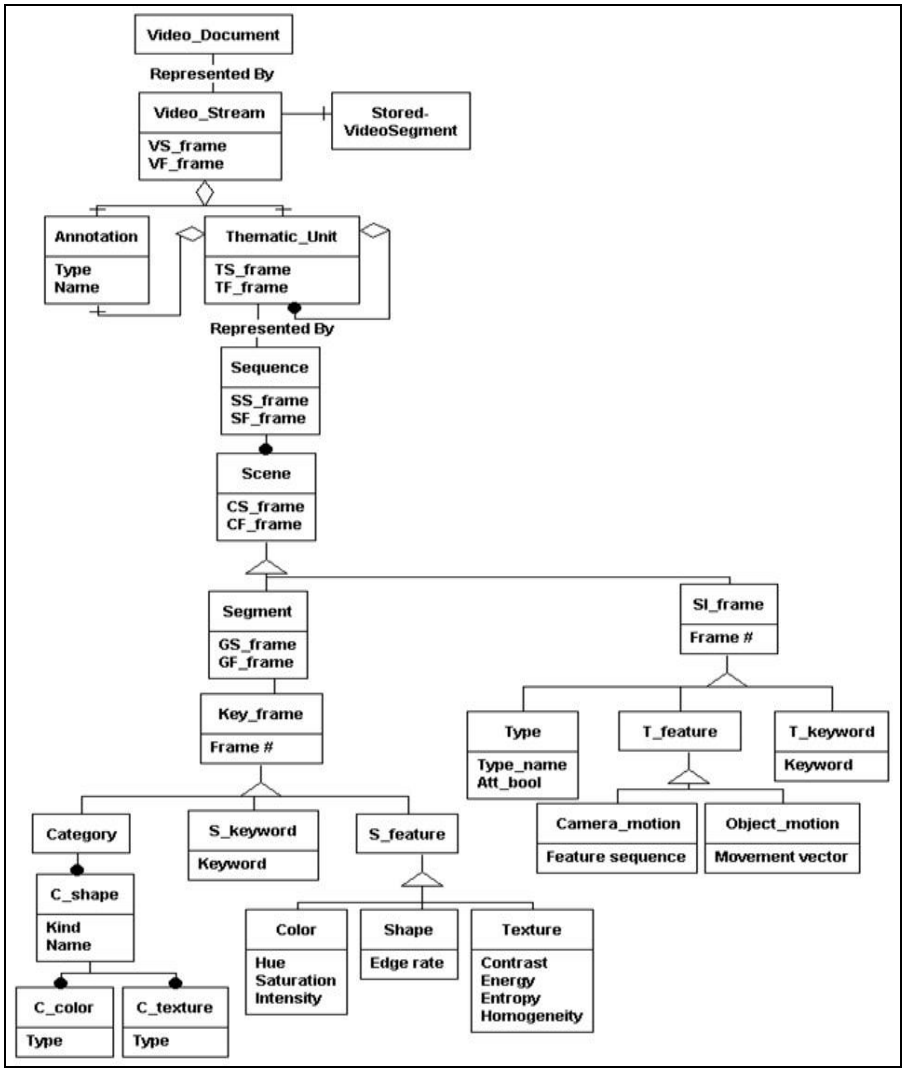


Fig. 7. OMT Object Diagram of IVDM

Through the process shown in Fig. 6, we propose an integrated video data model (IVDM) as shown in Fig. 7. Fig. 7 shows the OMT object diagram of the Integrated Video Data Model (IVDM). The OMT object diagram represents classes and their relationships, which is well suited to the design of databases.

Video_Document is expressed 1:1 again as Video_Stream, which is composed of one or more Stored_Video Streams and is stored in the database. In this case, Video_Stream has two attributes indicating the start frame and the end frame. As a part relationship

(part_of) between Video_Steam and Annotation and Thematic_Unit, the set of one or more Annotation and Thematic_Unit becomes Video_Stream. Thematic_Unit may or may not contain a smaller Thematic_Unit again, and like Video_Stream, it is composed of one or more Annotations. As a generalization relationship (is_a) between a Scene and a Segment and SI_frame, the Scene can be expressed again as a Segment or SI_frame. SI_frame can be expressed as Type, T_feature, and T_keyword, respectively, and T_feature is generalized to Camera_Motion and Object_Motion again. Segment is in reference relationship with Key_frame, and this Key_frame can be expressed again as Category, S_keyword, and S_feature.

7. Schema Design for News Videos

In Section 7, we design the schema structure and query type of the news video based on the IVDM model and examine the processing process. In actual implementation, Informix, an object-relational DBMS, was used to manage index information, and the user interface was implemented using Visual C++ [15].

7.1. Schema Structure of News Videos

In Section 1, based on the IVDM model, a news video that can be a representative example of a video was designed to be implemented in an object-oriented database.

Fig. 8 shows the structure of the news video schema. The upper part of each square box is the class name of the database, and the lower part is the properties of each class. In the news video schema, each subclass inherits the properties of the top video class, such as start frame, end frame, and oid of actual video data. And news, theme, event, reporter, and scene classes are connected by properties with oid in order as classes for annotation-based search. Classes below scene are for content-based search. Key_frame and lower are for spatial search, and SI_frame and lower are classes for time dimension. The class for using the automated method through the actual image analysis algorithm is the class below s_feature or t_feature.

7.2. The Process of Searching for News Videos

Fig. 9 shows the actual processing process of news video search. Fig. 9 is a case of news video compressed with MPEG-4, and it is largely composed of a user interface, video processing module, and data storage. The user interface is again divided into index editor, video searcher and video player. The video processing module is the process from the user interface to accessing the actual video data or related information in the DBMS, the data storage, in order to respond to the user's request. The role of the index editor is to annotate each topic for content-based video search in later comments and I-frames, and extract SI_frame and Key_frame.

In the video searcher, there is a difference in the search method depending on whether the data input for the search is in the form of text, an image, or a video. When the search

word is in the form of text, it is searched when the word exactly matches the data given in the form of a keyword, movement type, or category among annotation data assigned to each subject or content-based search. However, when image or video data is input as a query, color, shape, texture, and movement are analyzed in the query image as in the case of analyzing frames for content-based search and compared with the characteristic data stored in the database.

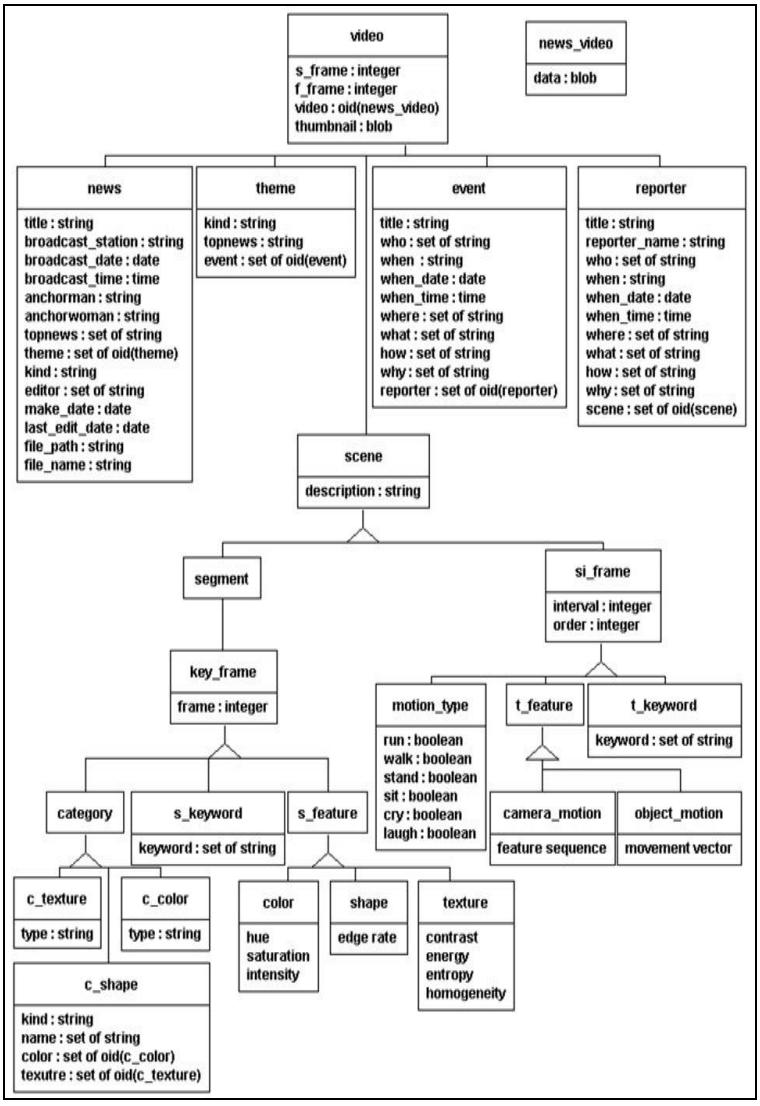


Fig. 8. Schema Structure of News Video

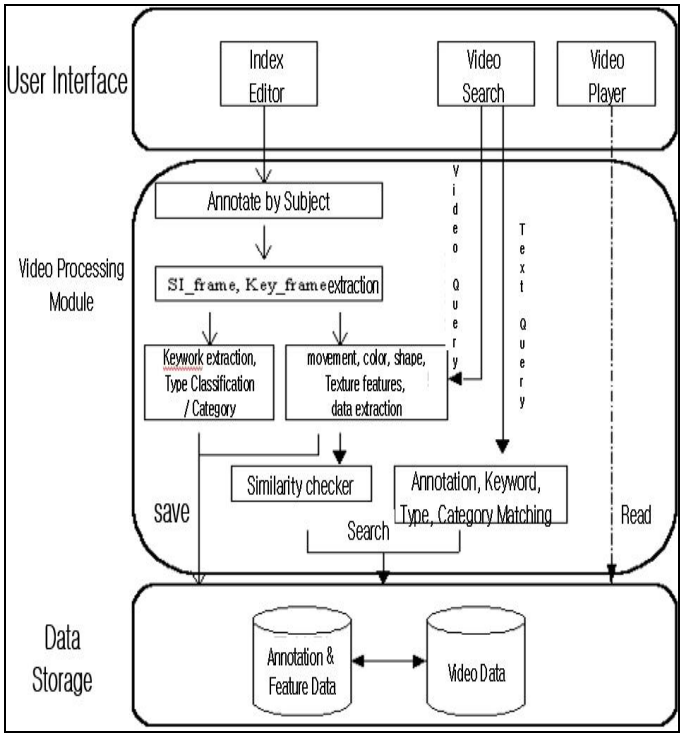


Fig. 9. News Video Search Process

The following shows the actual query, its processing process, and search results in order.

▶ Question: Among sports events that took place in December 1998, search for information about Se-ri Pak swinging and Chan-ho Park pitching.

▶ Query processing process:

```
 Ref1 := SELECT * FROM theme WHERE (when_date>='12/1/1998') AND
(when_date<='12/31/1998') AND (kind = 'sport');
 Ref2 := SELECT * FROM motion_type WHERE (swing = True) OR
(throw = True);
 Ref3 := SELECT * FROM c_shape WHERE (name IN 'Park Chan-ho') OR
(name IN 'Seri Pak');
Temp := Compare(Ref1, Ref2);
Result := Compare(Temp, Ref3);
```

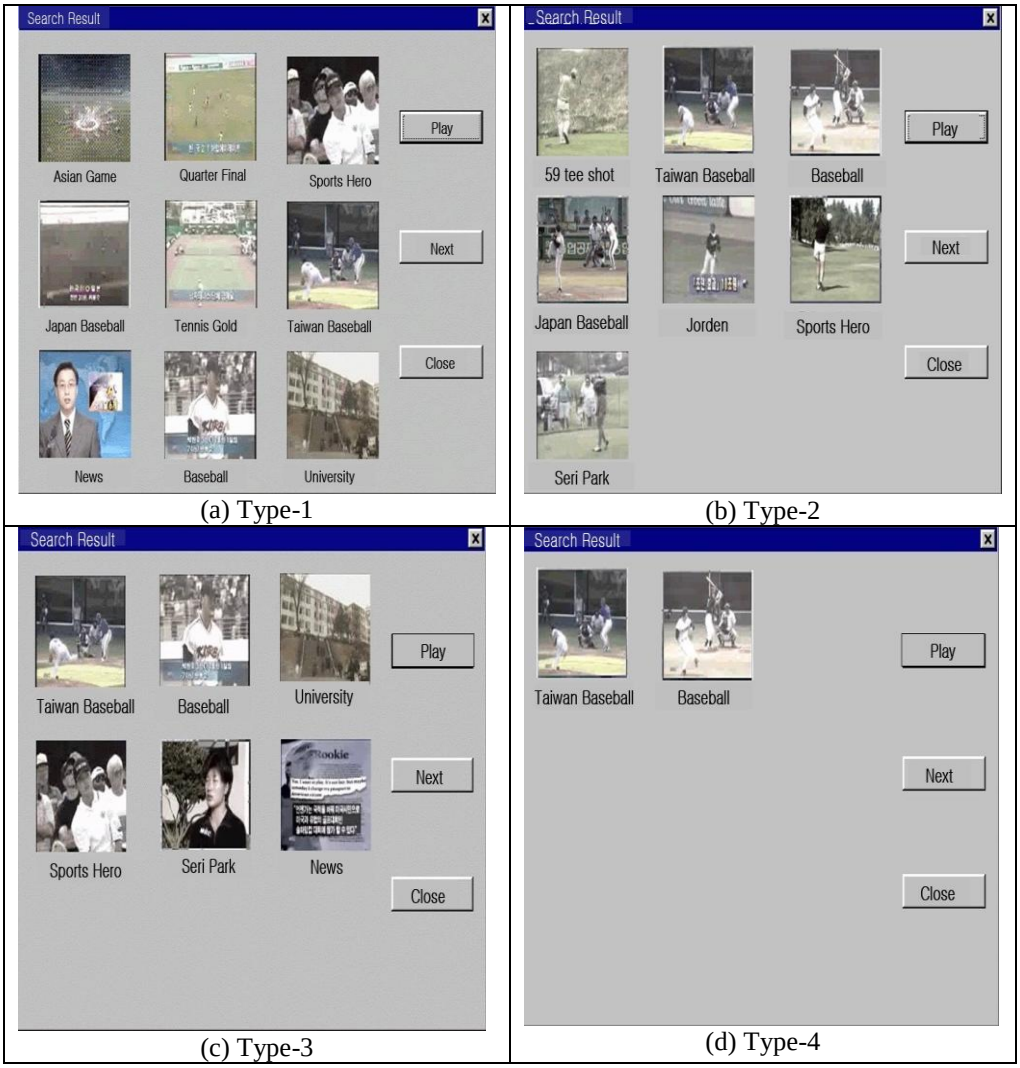


Fig. 10. Query Results for Videos

▶ Fig. 10(a) is the contents of Ref1, which is a Type-1(annotation-based) query result, and Fig. 10(b) is the contents of Ref2, which is a Type-2(SI_frame based) query result. Fig. 10(c) is the contents of Ref3 of Type-3(Key-frame based) query result. And Fig. 10(d) is the content of the final result of the Type-4 query that integrates the above three.

8. Conclusions and Future Challenges

In this paper, EGVDM (Enhanced Generic Video Data Model) is implemented by extending the video data model, and based on this, a prototype of compressed video information management system (CVIMS) that can manage MPEG-4 compressed video is designed and implemented. And based on this model, we designed an object-oriented database schema using news video as an example.

As a future task, not only index information but also image data itself should be structured and stored in the database. In order to structure the image data itself, the image data must be objectified, and an object-oriented database supporting objectification must be further developed. In addition, it is necessary to standardize the category of Key_frame and the type of SI_frame among the sub-scene structures for content-based search.

References

1. Mark Anthony Camilleri & Adriana Caterina Camilleri, The Acceptance of Learning Management Systems and Video Conferencing Technologies: Lessons Learned from COVID-19, Technology, Knowledge and learning 27, (2022) 1311-1333.
2. W. Zhu, et al., Evaluation of sino foreign cooperative education project using orthogonal sine cosine optimized kernel extreme learning machine, IEEE Access 8 (2020) 61107–61123, https://doi.org/10.1109/ACCESS.2020.2981968.
3. O Kuchai, K Skyba and A Demchenko, The importance of multimedia education in the information of society, IJCSNS VOL 22, No. 4, (2022) 797-803.
4. L.O. Seman, G. Gomes, R. Hausmann, CC-CCjs: a javascript web based application for education on basic converters, IEEE Latin Am. Trans. 13 (Aug. (8)) (2021) 2715–2722, https://doi.org/10.1109/TLA.2015.7332154.
5. Z. Gingl, G. Makan, J. Mellar, G. Vadai, R. Mingesz, Phonocardiography and photoplethysmography with simple arduino setups to support interdisciplinary STEM education, IEEE Access 7 (2019) 88970–88985, https://doi.org/10.1109/ACCESS.2019.2926519.
6. Y.A.M. Qasem, R. Abdullah, Y.Y. Jusoh, R. Atan, S. Asadi, Cloud computing adoption in higher education institutions: a systematic review, IEEE Access 7 (2019) 63722–63744, https://doi.org/10.1109/ACCESS.2019.2916234.
7. N.J. Martarelli, M.S. Nagano, Socioeconomic class of brazilian cities for health, education and employment & income IFDM: a clustering data analysis, IEEE Latin Am. Trans. 14 (3) (2016) 1513–1518, https://doi.org/10.1109/TLA.2016.7459643.
8. M.A. Cohen, G.O. Niemeyer, D.S. Callaway, Griddle: video gaming for power system education, IEEE Trans. Power Syst. 32 (July (4)) (2020) 3069–3077, https://doi.org/10.1109/TPWRS.2016.2618887.
9. J. Zheng, Q. Zhang, S. Xu, H. Peng, Q. Wu, Cognition-based context-aware cloud computing for intelligent robotic systems in mobile education, IEEE Access 6 (2018) 49103–49111, https://doi.org/10.1109/ACCESS.2018.2867880.
10. S.D. Assimonis, V. Fusco, RF energy harvesting with dense rectenna-arrays using electrically small rectennas suitable for IoT 5G embedded sensor nodes, in: 2018 IEEE MTT-S International Microwave Workshop Series on 5G Hardware and System Technologies (IMWS-5G), Dublin, 2018, pp. 1–3, https://doi.org/10.1109/IMWS-5G.2018.8484384.

11. Andreu Vaillo, Y., Murgui Perez, S., Martínez Lopez, P., Romero Retes, R., 2021. Mini-mental adjustment to cancer scale: construct validation in Spanish breast cancer patients. J. Psychosom. Res. 114, 38–44. https://doi.org/10.1016/j. jpsychores.2018.09.004.
12. Annunziata, M.A., Muzzatti, B., Bidoli, E., Flaiban, C., Bomben, F., Piccinin, M., Gipponi, K.M., Mariutti, G., Busato, S., Mella, S., 2022. Hospital Anxiety and Depression Scale (HADS) accuracy in cancer patients. Support. Care Canc 28, 3921–3926. https://doi.org/10.1007/s00520-019-05244-8.
13. K.T.D. Nguyen, C. Huang, An intelligent parallel algorithm for online virtual network embedding, in: 2019 International Conference on Computer, Information and Telecommunication Systems (CITS), Beijing, China, 2019, pp. 1–5, https://doi.org/10.1109/CITS.2019.8862072.
14. N. Ericsson, T. Lennvall, J. Åkerberg, M. Bj ¨orkman, A flexible communication stack design for time sensitive embedded systems, in: 2017 IEEE International Conference on Industrial Technology (ICIT), Toronto, ON, 2017, pp. 1112–1117, https://doi.org/10.1109/ICIT.2017.7915518.
15. D. Punia, B. Singh, Study of high-performance RFIC designs with efficient PA architectures for 5G networks, in: 2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT), Kanpur, India, 2022, pp. 1–5, https://doi.org/10.1109/ICCCNT45670.2019.8944582.
16. J. Cui, X. Zhang, H. Zhong, Z. Ying, L. Liu, RSMA: reputation system-based lightweight message authentication framework and protocol for 5g-enabled vehicular networks, IEEE Internet Things J. 6 (4) (2019) 6417–6428, https://doi.org/10.1109/JIOT.2019.2895136. Aug.
17. A. Fendt, C. Mannweiler, L.C. Schmelz, B. Bauer, An efficient model for mobile network slice embedding under resource uncertainty, in: 2019 16th International Symposium on Wireless Communication Systems (ISWCS), Oulu, Finland, 2019, pp. 602–606, https://doi.org/10.1109/ISWCS.2019.8877372.
18. M. Dawson, F.G. Martinez, P. Taveras, Framework for the development of virtual labs for industrial internet of things and hyperconnected systems, in: 2019 IEEE Learning With MOOCS (LWMOOCS), Milwaukee, WI, USA, 2019, pp. 196–198, https://doi.org/10.1109/LWMOOCS47620.2019.8939660.
19. G. Chen, J. Tang, J.P. Coon, Optimal routing for multihop social-based D2D communications in the internet of things, IEEE Internet Things J. 5 (3) (2018) 1880–1889, https://doi.org/10.1109/JIOT.2018.2817024. June.
20. Z. Han, A. Xu, Ecological evolution path of smart education platform based on deep learning and image detection, Microprocess. Microsyst. (2020), 103343, ISSN 0141-9331.
21. C. Wei, 5G-oriented IOT coverage enhancement and physical education resource management, Microprocess. Microsyst. (2020), 103366. ISSN 0141-9331.
22. T. Yamakawa, M. Hashiba, T. Koyama, K. Akazawa, Amethod to convert HDTV videos of broadcast satellite to RealSystem multimedia contents, J. Med. Syst. 26 (2002) 439—444.
23. Golaghaie, F., Esmaeili-Kalantari, S., Sarzaeem, M., Rafiei, F., 2021. Adherence to lifestyle changes after coronary artery bypass graft: outcome of preoperative peer education. Patient Educ. Counsel. 102, 2231–2237. https://doi.org/10.1016/j. pec.2019.07.019.
24. M.T. Chou, P. McGinnis, R. Tello, A web-based video tool for MR arthrography, Comput. Biol. Med. 33 (2003) 113—117.
25. R. Friedl, M.B. Preisack, W. Klas, T. Rose, S. Stracke, K.J. Quast, A. Hannekum, O. Godje, Virtual reality and 3D vi-sualizations in heart surgery education, Heart Surg. Forum 5 (2002) E17—E21.
26. T. Boudier, D.M. Shotton, Video on the Internet: an intro-duction to the digital encoding, compression, and trans-mission of moving image data, J. Struct. Biol. 125 (1999) 133—155.
27. M.J. Garcia, J.D. Thomas, N. Greenberg, J. Sandelski, C.Herrera, C. Mudd, J. Wicks, K. Spencer, A. Neumann, B.Sankpal, J. Soble, Comparison of MPEG-1digital videotape with

digitized sVHS videotape for quantitative echocardio-graphic measurements, J. Am. Soc. Echocardiogr. 14 (2001) 114—121.
28. K. Spencer, L. Solomon, V. Mor-Avi, K. Dean, L. Weinert, M.Gulati, A. Herle, A. Spiegel, B. Balasia, T. Pionke, L. Sieb,R.M. Lang, Effects of MPEG compression on the quality and diagnostic accuracy of digital echocardiography studies, J.Am. Soc. Echocardiogr. 13 (2000) 51—57.
29. K. Spencer, L. Weinert, V. Mor-Avi, K. Dean, B. Bala-sia, L. Solomon, T. Pionke, L. Sieb, R.M. Lang, Electronic transmission of digital echocardiographic studies: effects of MPEG compression, Int. J. Cardiol. 15/75 (2000) 141—145.
30. J.S. Soble, G. Yurow, R. Brar, T. Stamos, A. Neumann, M. Garcia, M.F. Stoddard, P.K. Cherian, B. Bhamb, J.D. Thomas, Comparison of MPEG digital video with super VHS tape for diagnostic echocardiographic readings, J. Am. Soc. Echocardiogr. 11 (1998) 819—825.

Byeongtae Ahn works at Faculty of Liberal Arts College at Anyang University, Korea. He was assistant professor, Dept of Computer Information of Catholic University in 2006~2012. His research interests include: Image Processing, Video Analysis, IoT, BlockChain, Multimedia Database and MPEG-7. His address is: 37-22, Samduck Minahn-gu Anyang-City Gyeonggi-do, 430-714 South Korea.

Received: May 09, 2022; Accepted: September 19, 2023.

Computer Science and Information Systems 21(2):437–452 https://doi.org/10.2298/CSIS220509050M

Automatic Voltage Stabilization System for Substation using Deep Learning

Jiyong Moon[1], Minyeong Son[2], Byeongchan Oh[3], Jeongpil Jin[4], and Younsoon Shin[5]

[1] Department of Business Administration, Dongguk University,
30, Pildong-ro 1-gil, Jung-gu, Seoul, Korea
asdwldyd@dongguk.edu
[2] Department of Medical Biotechnology, Dongguk University,
30, Pildong-ro 1-gil, Jung-gu, Seoul, Korea
smya0930@dongguk.edu
[3] Department of Statistics, Dongguk University,
30, Pildong-ro 1-gil, Jung-gu, Seoul, Korea
oxox97@dongguk.edu
[4] Department of Industrial System Engineering, Dongguk University,
30, Pildong-ro 1-gil, Jung-gu, Seoul, Korea
chin9510@dongguk.edu
[5] Department of Computer Science, Dongguk University,
30, Pildong-ro 1-gil, Jung-gu, Seoul, Korea
ysshin@dongguk.edu

Abstract. The operating voltage in the substation must be maintained at its rated voltage within the specified standard because a voltage outside the specified range may cause a malfunction of the power facility and interfere with the stable power supply. Therefore, the voltage regulation process to maintain the rated voltage of the substation is essential for the stability of the power system. However, the voltage regulation process is currently performed manually by resident staff. Voltage regulation based on human judgment increases the uncertainty of voltage stabilization and makes efficient operation in consideration of the economic feasibility of power facilities difficult. Therefore, this paper proposes an automatic voltage stabilization system that can automatically perform voltage regulation. Instead of predicting the electrical load or overvoltage conditions studied so far, we focus on more direct, scalable input capacity prediction for an automatic voltage stabilization system. First, the proposed system predicts the input capacity required for a given situation through a trained stacked LSTM model. Second, an optimal regulation plan is derived through an optimization process that considers the economic feasibility of power facility operation. Additionally, the development of the user interface makes it possible to visualize the operation of algorithms and effectively communicate the models' predictions to the user. Experimental results based on real substation data show that the proposed system can effectively automate the voltage regulation process.

Keywords: automatic voltage stabilization system, energy system, input capacity prediction, deep learning, optimal regulation plan

1. Introduction

The operating voltage in the substation must be maintained at its rated voltage within the specified standard for the stability of the power system. If the voltage exceeds (**overvoltage**) or falls below (**undervoltage**) the rated voltage range, it may cause a malfunction of the power facility and interfere with the stable power supply. Therefore, the voltage regulation process to maintain the rated voltage of the substation is essential. The voltage regulation process is done through a voltage stabilization system (VSS). The voltage stabilization system refers to a system that can sequentially control the operating conditions of reactors constituting a substation [16]. A reactor is an absorber of reactive power, therefore compensating for high voltage transmission [14]. When the reactor is operated, the voltage decreases due to the consumption of reactive power, and when the reactor is stopped, the voltage increases.

However, most of existing voltage stabilization systems are manually operated by resident staff. In other words, various decisions for voltage regulation, such as deciding whether to operate a reactor, are made solely according to the personal judgment of the resident staff. There are two main problems with the voltage regulation process performed by humans. First, continuous monitoring is difficult. In the case of manual work, real-time response may be difficult due to breaks or shift hours, and inconsistent response may occur because each employee has a different handling method [24]. Second, efficient operation considering economic feasibility is difficult. In general, the more a reactor is used, the more likely it is to fail. When the reactor is operated at a high frequency, very fast transient overvoltage (VFTO) occurs more frequently, and when the voltage exceeds the basic impulse insulation level (BIL), it leads to the failure of the reactor [16]. Therefore, when performing voltage regulation, it is necessary to distribute the frequency of use of each reactor. Still, if the voltage regulation process is done manually, it isn't easy to properly consider this by personal judgment.

In order to solve the above problems, automation of the voltage stabilization system is required. Therefore, in this paper, we propose a prediction-based automatic voltage stabilization system using a stacked long-short term memory (stacked LSTM) model. Beyond statistical or mathematical methods [26,6,2,27], many prediction-based methods have been proposed for the stability of the power system. Recently, machine learning or deep learning-based methods have been mainly proposed. The main object of prediction is to predict the overvoltage situation for voltage stabilization [5,4,11,37], electrical loads [36,19,13], and reactive power [14]. Overvoltage situations, electrical loads, and reactive power are all important for the stability of a power system. However, developing an automatic voltage stabilization system requires a different approach. Overvoltage situations, electrical load, and reactive power can be used as indicators of the stability of a power system, but in terms of automatic voltage regulation, their purpose is different. This is because even with predictions for overvoltage situations, electrical load, and reactive power, it is not known how to adjust the power facility for a given situation. In other words, to implement automatic voltage stabilization, an additional prediction process is inevitable. In order to automatically regulate the reactors to the situation through an automatic voltage stabilization system, it is necessary to predict a more direct value to regulate. Therefore, we developed a model to predict the input capacity required for a given situation. Input capacity means the maximum amount of reactive power that one reactor can consume, but it can be used as a standard to regulate the reactor. For example,

if the model predicts that an input capacity of $400\,\mathrm{Mvar}$ is needed in a given situation, it can respond by operating two shunt reactors (Sh.R) with an input capacity of $200\,\mathrm{Mvar}$. Predicting the input capacity indicates information about the level of danger expressed by overvoltage situations, electrical load, and reactive power. Predicting the input capacity also makes it easy to infer how to adjust the power facility in a given case. This is because input capacity is the most basic and direct basis for power facility operation. Therefore, the task of predicting input capacity is more suitable for implementing an automatic voltage stabilization system than simply predicting overvoltage, reactive power, and electrical load. In addition, this method does not require a prior definition of the applied system. This is because the required input capacity is fixed regardless of what kind of power facility the system consists of or the number of reactors constituting the system. This means that predicting input capacity is also beneficial for expansion and application.

In this study, we design an input capacity prediction model that is more suitable for automatic voltage stabilization systems and propose a solution that can be directly applied to the actual work site. The model was evaluated based on the data extracted from the actual substation to ensure reliability. We develop not only the algorithm but also the user interface and integrate them into one system so that it can be applied easily in the actual field.

2. Related Works

2.1. Prediction-based Methods for Voltage Stabilization

Prediction-based methods for voltage stabilization are mainly aimed at predicting overvoltage conditions, electrical loads, and reactive power. Various machine learning and deep learning algorithms were used for prediction.

Bulac et al. [4] proposed a method to perform real-time voltage stabilization monitoring using a multi-layer perceptron (MLP). The target class is divided into stable, unstable, and dangerous.The proposed MLP model predicts the risk level of overvoltage in a given situation by receiving voltage-related features as input.

Zhu et al. [37] proposed a method of identifying a class imbalance problem [30] in which a situation corresponding to 'unstable' in a voltage stabilization system is very rare when predicting an overvoltage situation and improving performance using an imbalance learning. The class imbalance problem was solved by amplifying the unstable situation class data through the synthetic minority oversampling technique (SMOTE) [7], and the weighted cost was set to make the model learn more focused on a small number of unstable classes. In addition, they tried to improve the model's generalization performance and increase its applicability by allowing the model to learn with new data through incremental learning continuously. Similarly, since deep learning-based methods cause a high dependence on data and annotations for high performance, Li et al. [21] proposed combining data augmentation methods to lower this dependence.

Gomez et al. [11] tried to predict the overvoltage condition early using one of the powerful classification models, the support vector machine (SVM) [25], based on the idea that it is important to quickly predict how much the voltage will be affected immediately after the situation causing the overvoltage. The significant errors that can cause overvoltage include features such as generator voltage, speed, or rotation angle, and these variables

are used as inputs for the proposed SVM model. Also, a support vector regressor (SVR), which applied SVM to a regression problem, was used to predict the electrical load, and a chaotic genetic algorithm (CGA) [34] was used in the hyperparameter determination process of SVR [13].

Cao et al. [5] proposed a method combining convolutional neural networks (CNN) [1] and deep reinforcement learning (DRL) [15] to predict overvoltage stability in the energy internet. The proposed method predicts overvoltage stability by performing a convolution operation on time-series information composed of a two-dimensional matrix and determines whether the voltage can be stabilized within a given time in the current state through DRL.

Jiapeng et al. [31] proposed a method for identifying overvoltage types of high-voltage electrical systems of multiple units based on lightweight ShuffleNet [35]. The six overvoltage types are mapped to grayscale images by the B2G algorithm, and ShuffleNet takes them as input and classifies the overvoltage types.

Ko et al. [19] proposed a hybrid model that combines a radial basis function neural network (RBFNN) [3] and a dual extended Kalman filter (DEKF) [7] with SVR for electrical load prediction. SVR and DEKF are used in the initial value setting and learning process of RBFNN, respectively.

Zheng et al. [36] used a time-series deep learning model, recurrent neural networks (RNN) [23], and an improved version, long-short term memory (LSTM) [12], for electrical load prediction. The proposed model proposes a model that predicts the electrical load of the next 12 steps with the electrical load data of the past 12 steps through the RNN architecture using the LSTM cell. The LSTM architecture was also used in the reactive power prediction study and showed better performance as the length of the input sequence length increased [14].

Like our objective, Yin et al. [32] proposed an automatic voltage stabilization method using an emotional deep neural network (EDNN) structure and an artificial emotional Q-learning algorithm. Jiajun et al. [9] proposed GridMind using deep reinforcement for autonomous voltage control in the power grid. Hanchen et al. [29] proposed the use of computationally efficient Batch Reinforcement Learning (BRL), along with a formulation strategy using the Markov Decision Process (MDP) for voltage regulation in power distribution systems.

Our study is similar to that of Yin et al. [32], Jiajun et al. [9] and Hanchen et al. [29] in that it considers automatic voltage stabilization. However, since these studies are mainly aimed at minimizing the voltage deviation across the system, they differ from ours, focusing on solving the overvoltage situation. We also paid attention to practical aspects, including the user interface. Additionally, our study is similar to that of Hossain et al. [14] and Zheng et al. [36] in that it uses RNN and LSTM architectures. However, there is a difference in that the prediction target of our proposed method is input capacity. We predict the input capacity using RNN and LSTM architectures, given that voltage and input capacity have time-series characteristics. The following subsection provides a brief introduction to RNN and LSTM.

2.2. Recurrent Neural Networks

LSTM has the architecture of a RNN. RNN is a deep learning architecture specialized for time series data processing [18,23,32]. The most straightforward architecture of RNN is

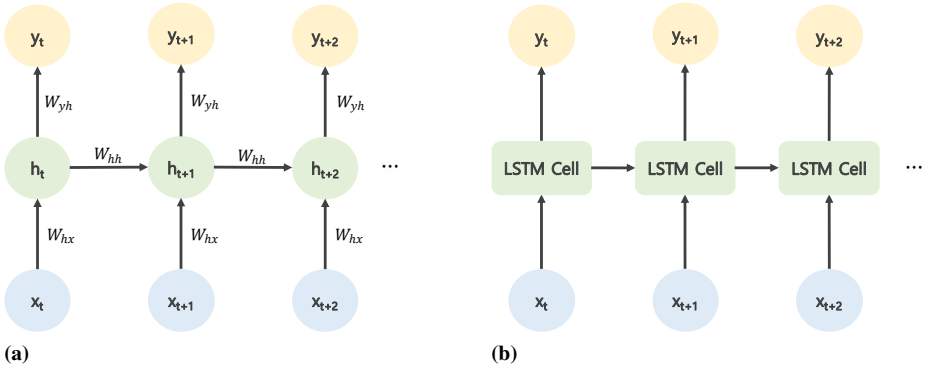


Fig. 1. A simple RNN and LSTM architecture. (a) RNN architecture. (b) LSTM architecture

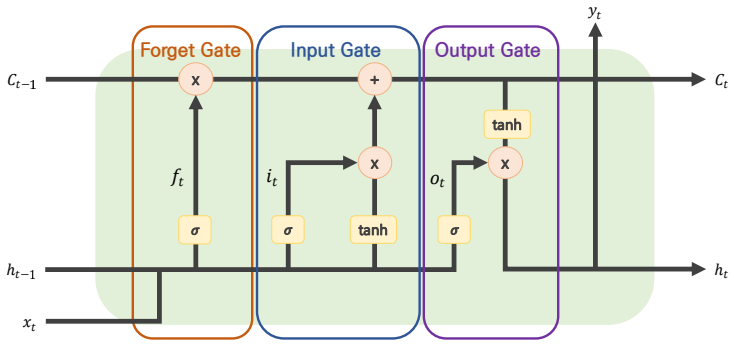


Fig. 2. The internal structure of the LSTM cell

shown in Fig. 1 (a). Like other deep learning models, the RNN goes through one or more hidden layers for a given input and returns the output. However, the unique feature of the RNN architecture is that the output of the hidden layer comes back into the input of the corresponding hidden layer. This structure considers the characteristic of sequence data that the data point of each time step is not independent of the data point of the previous time step. Information of each time step is accumulated, which is is reflected in the next time step processing to process sequence data.

The LSTM refers to an architecture in which the part corresponding to the hidden layer in the RNN is replaced with an LSTM cell [12]. A simple LSTM architecture is shown in Fig. 1 (b). Although the purpose of processing sequence data is the same, LSTM operates slightly differently from general RNN processing due to this structural change. The internal structure of the LSTM cell is shown in Fig. 2. Unlike the previous RNN, the LSTM has a cell state indicated by C_{t-1} and C_t. The cell state is the path of information passing through the entire time step. By not only using the hidden state for information accumulation and reflection but by defining a separate cell state to flow information that can be utilized in the entire time step, LSTM can process longer sequences than general RNN structures and has superior performance [20]. In LSTM, the flow of information through the cell state is controlled by three gates. Forget gate determines how much information in the cell state to forget. The input gate decides how much to reflect the current input and hidden state in the cell state. The output gate determines how much of the cell state to send as the current output and hidden state. We used this LSTM architecture for input capacity prediction.

3. Proposed Method

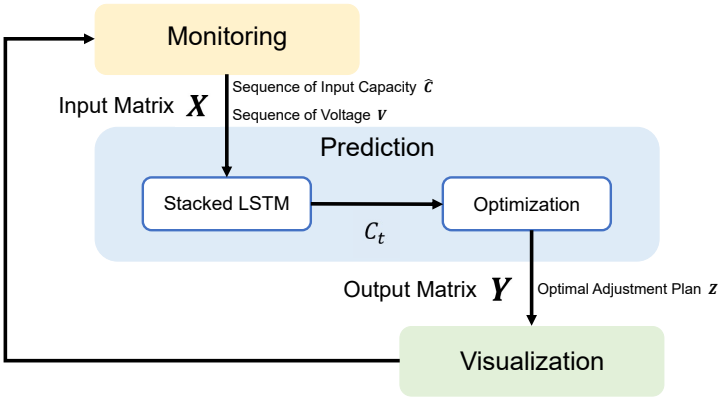


Fig. 3. The proposed automatic voltage stabilization system

In this paper, we implemented an automatic voltage stabilization system based on the input capacity prediction. The overall flow chart of the proposed system is presented in Fig. 3. First, the proposed system monitors the voltage of the applied substation. At the

same time, a time-series input matrix **X** consisting of the monitored voltage and past input capacity is extracted for input capacity prediction. The stacked LSTM model predicts C_t, the required input capacity at the current time t, through the **X**. Based on the predicted C_t, an optimal regulation plan for whether to operate each reactor is derived through the optimization process. This information constitutes the output matrix **Y**. In addition, the monitored voltage and optimal regulation plan are visualized through the designed user interface. This process is repeated at fixed time intervals. The voltage regulation process can be automated through the proposed system, so the problems of existing manual operation can be solved.

The proposed system is largely divided into two parts: optimal regulation plan prediction (Section 3.1) and visualization (Section 3.2). First, a trained stacked LSTM model predicts the required input capacity from a given input. Next, a final optimal regulation plan is derived through the optimization process. Finally, information such as the derived optimal adjustment plan and voltage is visualized through the user interface.

3.1. Deriving the Optimal Regulation Plan

Input Capacity Prediction

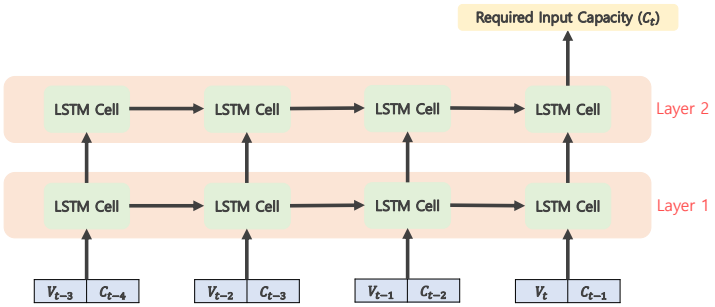


Fig. 4. Proposed input capacity prediction model architecture

The corresponding voltage and input capacity also have a time-series feature because electricity demand has a time-series characteristic. Therefore, a statistical time series model using time as a variable can be used to predict the input capacity [6]. However, given that electricity demand is a non-linear time series, a more robust prediction model than a statistical model is needed [19]. Additionally, it is also necessary to consider additional variables such as past voltages rather than using time as the only variable. Therefore, in this paper, we use the stacked LSTM, a deep learning model specialized in the sequence data processing. Through this, it is possible to consider the time series characteristics of input capacity, further improve performance by considering non-linearity, and consider additional variables other than time.

The proposed input capacity prediction model is shown in Fig. 4. The model has an LSTM architecture. In addition, by stacking two hidden layers composed of LSTM cells, more non-linearities can be considered. The stacked LSTM architecture has the advantage of learning various characteristics of time series data at each time step over the basic

LSTM architecture [33]. The input is composed of the past voltage and the input capacity along with the current voltage. In addition, the length of sequence data coming in as input is 4 (The details of the hyperparameter setting are described in Section 4.2). Therefore, the input matrix **X** described in Fig. 3 is composed as follows:

$$\mathbf{X} = \begin{bmatrix} V_t & C_{t-1} \\ V_{t-1} & C_{t-2} \\ V_{t-2} & C_{t-3} \\ V_{t-3} & C_{t-4} \end{bmatrix} \in \mathbb{R}^{4\times 2} \tag{1}$$

In (1), V_t means the voltage at each time point, and C_t means the input capacity at each time point. Since the purpose of prediction is C_t, which is the required input capacity at the current time t, note that C is composed of 4 starting at $t-1$ instead of at t like V. The model predicts the currently required input capacity C_t by sequentially processing the input matrix **X**.

Optimization

After predicting the required input capacity through the model, it is necessary to decide how to regulate the power facilities (i.e., reactors). In this paper, the optimal regulation plan is derived through the optimization formula. The optimization formula was designed considering economic feasibility and efficiency. As mentioned in Section 1, the probability of failure increases as the number of operations of the power facility increases [16]. Therefore, it is necessary to distribute the number of operations for each power facility, which can be a basis for deriving an optimal regulation plan.

The defined optimization formula is as follows:

$$\begin{aligned} &\text{minimize}_{z_1,\ldots,z_n} \quad \textstyle\sum_{i=1}^{n} \gamma_i z_i \\ &\text{subject to} \quad \textstyle\sum_{i=1}^{n} C_i z_i \geq C_t \\ &\text{subject to} \quad \textstyle\sum_{i=1}^{n} C_i z_i - C_t \geq C_{min} \end{aligned}$$

(3.1) is the objective function of the optimization formula. In (3.1), z_i means the operating state of each power facility of the applied system and has a value of 0 or 1. γ_i means the cumulative number of uses of the corresponding power facility. The optimization process treats the sum of the cumulative use times of each power facility as a cost, and aims to determine whether to operate each power facility in which the cost can be minimized. (3.1) is the first constraint. In (3.1), C_i means the input capacity of the corresponding power facility, and C_t means the predicted required input capacity. If there is no constraint, the optimization process will minimize the cost to zero by disabling all power facilities. Therefore, (3.1) solves this problem by forcing the optimization process to input the power equipment as much as the predicted required input capacity. (3.1) is the second constraint. In (3.1), C_{min} means the input capacity of the power facility with the smallest input capacity among all power facilities. If there is no constraint, the optimization process will try to keep the previous state when the previous input capacity is greater than the currently needed input capacity. Therefore, (3.1) solves this problem by forcing the optimization process to change the state within the expressible input capacity range. In summary, the optimization process means considering economic feasibility and

efficiency by lowering the power facility management cost and the chance of damage by forcing the power facility to operate first with the lowest cumulative use frequency.

Through the optimization, an optimal regulation plan is derived. The derived optimal regulation plan becomes the output matrix **Y** of Fig. 3, and its composition is as follows:

$$\mathbf{Y} = \begin{bmatrix} z_1 \\ \vdots \\ z_n \end{bmatrix} \in \mathbb{R}^n \quad (z_i \in \{0, 1\}) \tag{5}$$

In (5), **Y** means the optimal regulation plan and contains information on whether each optimized power facility operates.

3.2. Visualization

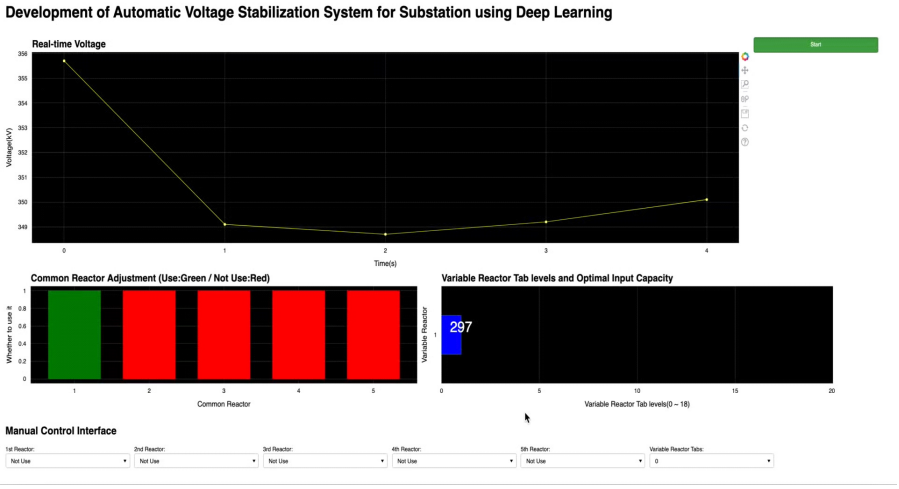


Fig. 5. Designed user interface

The derived optimal regulation plan is visualized through a designed user interface and information on the recorded voltage sequence. The user interface makes it easy to see how the system works and its results. The designed user interface is shown in Fig. 5.

In Fig. 5, when the start button at the top right is pressed, the user interface is operated. The user interface consists of three elements. First, the voltage graph appears at the top. The voltage at the latest 20-time points is expressed, and the overall flow of the voltage can be checked. Second, the optimal regulation plan is visualized in the center. Whether each of the derived power facilities operates and the predicted input capacity value are simultaneously expressed. The green bar means active, and the red bar means inactive. Finally, at the bottom is a manual operation button. In addition to the results automatically predicted by the system, it can be applied when a manual operation is required. As mentioned earlier, the user interface is updated according to a predefined time interval, and prediction and visualization are executed sequentially.

4. Experimental Results

In this section, the performance of the proposed system is evaluated. It is divided into the evaluating input capacity prediction model and the actual operation analysis.

4.1. Experimental Environment and Dataset

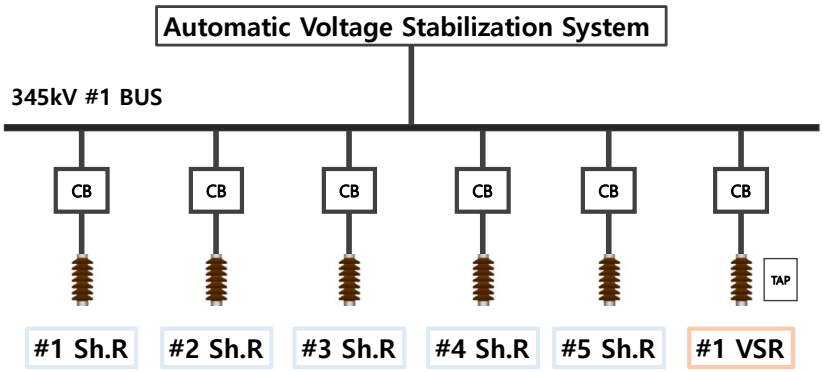


Fig. 6. Assumed substation environment

Before evaluation, it is necessary to assume the environment of the substation to which the system is applied. The considered substation environment is shown in Fig. 6. We assume that the applied substation consists of one $345\,\mathrm{kV}$ bus. Additionally, it consists of five Sh.R and one variable shunt reactor (VSR), each with an input capacity of $200\,\mathrm{Mvar}$. VSR is a reactor that can control power more delicately through a tap device. The tap of the VSR consists of a total of 18 stages [16]. Unlike the existing Sh.R, VSR operates on a tap basis, so the output matrix in (2) should be changed as follows:

$$\mathbf{Y} = \begin{bmatrix} z_1 \\ \vdots \\ z_5 \\ \hat{z}_1 \end{bmatrix} \in \mathbb{R}^6 \quad (z_i \in \{0,1\}, \quad \hat{z}_1 \in \{0,...,18\}) \tag{6}$$

In (6), $\hat{z}_1$ means the operating state of the VSR and has a value between 0 and 18.

The experimental data were collected in the real substation environment defined above. The substation automatically saves various information, including voltage, according to defined intervals (*i.e.,* one minute). The data contains operation information for each power facility constituting the substation system between 2019 and 2021. Data features include uptime, generation load, transmission load, input capacity, and ancillary information such as temperature, wind speed, and precipitation. We extracted only information about voltage and input capacity required for prediction. The total data size is about 450,000 data points; we used 25% as test data and the rest as training data.

4.2. Input Capacity Prediction Performance

Table 1. Input capacity prediction model performance (RMSE)

Model	Input Combination					
	$\mathbf{X}_1$	$\mathbf{X}_2$	$\mathbf{X}_3$	$\mathbf{X}_4$	$\mathbf{X}_5$	$\mathbf{X}_6$
XGBoost	183.13	79.07	182.22	20.57	20.41	14.32
LightGBM	183.12	124.93	179.04	21.13	19.59	14.01
RandomForest	183.13	78.55	182.15	20.60	20.00	14.14
GradientBoost	183.15	148.68	182.08	20.44	19.66	19.72
ElasticNet	184.12	180.35	184.10	67.93	32.63	32.54
DNN	184.90	138.97	183.09	20.53	19.54	13.97
LSTM	183.69	143.28	183.39	20.54	19.25	13.31
Stacked LSTM	183.69	185.30	182.85	20.56	19.30	**12.86**

First, we evaluated the performance of the input capacity prediction model. The purpose of the model is to predict the required input capacity given the appropriate inputs. Several models were trained and evaluated to find the optimal model and input combinations. Root mean squared error (RMSE) was used as the evaluation metric.

The overall result is shown in Table 1. A total of eight machine learning and deep learning models were trained and evaluated. XGBoost [8], LightGBM [17], and GradientBoost [22] are machine learning algorithms that show strong performance as tree boosting ensemble methods. RandomForest [10] is an ensemble model using the bagging method, and it is a model that reinforces the randomness of data and features. ElasticNet [38] is a regulated regression model that combines L1 and L2 regulation into linear regression. DNN is a structure in which several hidden layers are stacked in general artificial neural networks (ANN) [28], and we constructed a model with four hidden layers. As input combinations, six combinations were evaluated. $\mathbf{X}_1$ means only the current voltage at t is used as an input. $\mathbf{X}_2$ means using input time information such as a month, day, hour, etc., considering seasonal characteristics along with the current voltage at t. $\mathbf{X}_3$ means using the voltage sequence of the past time as an input together with the current voltage of time t. $\mathbf{X}_4$ means that only the input capacity at $t-1$ is used as input. $\mathbf{X}_5$ means that the past input capacity of the same length as $\mathbf{X}_3$ is used as input. $\mathbf{X}_6$ means to use a sequence composed of the input capacity of the past time as an input together with the voltage sequence of $\mathbf{X}_3$.

All models achieved the best performance when $\mathbf{X}_6$ was used as the input combination. As can be seen when $\mathbf{X}_4$ and $\mathbf{X}_5$ are used as inputs, the model's performance is significantly improved when it can explore the past input capacity or input capacity sequence rather than when voltage alone is used. However, the performance is further enhanced when the past voltage and the current voltage are used together with the past input capacity sequence ($\mathbf{X}_6$). Additionally, the model performance of the RNN architecture specialized for sequence data processing was the best among all models, and the performance of the stacked LSTM model was the best with RMSE 12.86. Therefore, stacked LSTM was selected as the final model, and it was decided to use current voltage, past voltage, and past input capacity together as the input combination.

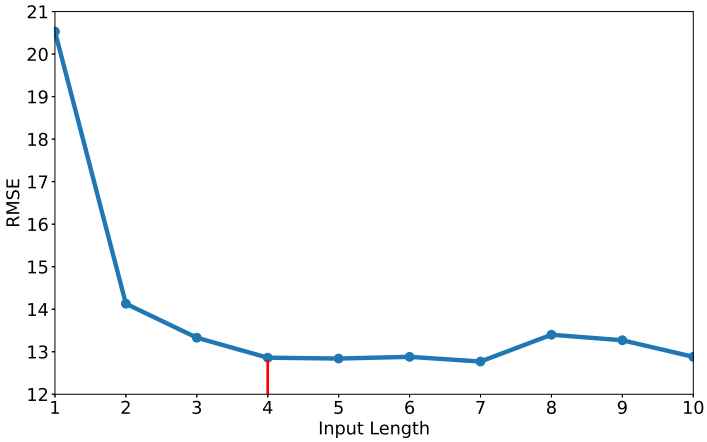


Fig. 7. Performance of the stacked LSTM model by the length of the input sequence

When using a sequence of voltage and input capacity as input, additional evaluation was performed to select the optimal length of the sequence, that is, to what point in the past voltage and input capacity information will be used. The results are shown in Fig. 7. In Fig. 7, the model showed significant performance improvement until the sequence length reached 4. After this, there was no significant performance improvement thereafter, so we set the optimal sequence length to 4.

4.3. Operation Analysis

Table 2. Operational Analysis Results

Item	Time Point				
	t	$t+1$	$t+2$	$t+3$	$t+4$
Voltage (kV)	353.8	349.8	354.5	349.7	345.5
Predicted Input Capacity (Mvar)	735	697	897	697	497
#1 Sh.R	**1**	**1**	**1**	**1**	**1**
#2 Sh.R	0	0	0	0	0
#3 Sh.R	0	0	**1**	**1**	**1**
#4 Sh.R	**1**	**1**	**1**	**1**	0
#5 Sh.R	**1**	**1**	**1**	0	0
#1 VSR (Tap Position)	**9**	**1**	**1**	**1**	**1**

Second, we conducted an operational analysis to see if the system actually works well. The results are shown in Table 2. In addition, Table 3 shows the assumed cumulative numbers of uses for each reactor in the optimization process.

Table 3. Assumed Cumulative Use Count

	#1 Sh.R	#2 Sh.R	#3 Sh.R	#4 Sh.R	#5 Sh.R	#1 VSR
Cum. Num. of Uses	200	150	100	80	50	20

In Table 2, system operation results for five consecutive time points from t to $t + 4$ are presented. Additionally, information on items such as voltage, input capacity, and reactor operation status at each time point is presented together. In more detail, at time t, the voltage was observed to be $353.8\,\mathrm{kV}$, and the model predicted that an input capacity of $735\,\mathrm{Mvar}$ was required. For the predicted input capacity, the operating state of each Sh.R and the tap position of the VSR were determined through an optimization process. After that, the observed voltage at time $t+1$ is $349.8\,\mathrm{kV}$, which is lower than before. This is because reactors consumes reactive power equal to previously input capacity to lower the voltage. Additionally, for the lower voltage, the model predicted that an input capacity of $697\,\mathrm{Mvar}$ lower than the time t was required. This shows that the input capacity prediction model predicts the appropriate input capacity considering the level of voltage. When checking the operation state of each reactor at time $t + 1$, it can be seen that the tap position of the VSR has changed from 9 to 1. This means that voltage adjustment was performed by changing the tap of the least frequent VSR through the optimization process in consideration of the number of uses for each reactor assumed in Table 3. In other words, it can be seen that the optimization process is properly distributing the operation for each reactor with the number of uses as a cost as intended. These results are also the same at all time points thereafter, including time $t + 2$. According to the experimental results, it can be seen that an automatic voltage regulation system can be effectively implemented through the designed system, and it can be confirmed that the goal of the study and the required performance requirements can be met.

5. Conclusion

This paper covered the development of an automatic voltage stabilization system for voltage regulation automation. First, a trained stacked LSTM model was designed to predict the input capacity required for a given situation using actual voltage and input capacity data. In addition, it was possible to derive the optimal regulation plan considering the economic feasibility of power facility operation by using the optimization method. Finally, the user interface shows how the model works as intended.

In this paper, only two variables of time-series voltage data and input capacity were used as inputs when training the model to predict the optimal input capacity. However, in addition to these two variables, there are other variables that could affect voltage changes, such as weather, season, temperature, and humidity. It is expected that future studies can use these variables to improve model performance considering complex voltage environments.

This automatic voltage stabilization system is more effective and economic than the conventional voltage control system. This not only enables a stable power supply but also increases the lifespan of power facilities and reduces the cost burden on the company for facility failure. Additionally, this paper can also contribute to the goals of informatization and securing big data in the substation field.

Acknowledgments. This research was supported by the MSIT (Ministry of Science, ICT), Korea, under the High-Potential Individuals Global Training Program) (2021-0-01549) supervised by the IITP (Institute for Information & Communications Technology Planning & Evaluation).

References

1. Albawi, S., Mohammed, T.A., Al-Zawi, S.: Understanding of a convolutional neural network. In: 2017 International Conference on Engineering and Technology (ICET). pp. 1–6 (2017)
2. Almeshaiei, E., Soltan, H.: A methodology for electric power load forecasting. Alexandria Engineering Journal 50(2), 137–144 (2011), https://www.sciencedirect.com/science/article/pii/S1110016811000330
3. Broomhead, D.S., Lowe, D.: Radial basis functions, multi-variable functional interpolation and adaptive networks. Tech. rep., Royal Signals and Radar Establishment Malvern (United Kingdom) (1988)
4. Bulac, C., Triştiu, I., Mandiş, A., Toma, L.: On-line power systems voltage stability monitoring using artificial neural networks. In: 2015 9th International Symposium on Advanced Topics in Electrical Engineering (ATEE). pp. 622–625 (2015)
5. Cao, J., Zhang, W., Xiao, Z., Hua, H.: Reactive power optimization for transient voltage stability in energy internet via deep reinforcement learning approach. Energies 12(8) (2019), https://www.mdpi.com/1996-1073/12/8/1556
6. Chakhchoukh, Y., Panciatici, P., Mili, L.: Electric load forecasting based on statistical robust methods. IEEE Transactions on Power Systems 26(3), 982–991 (2011)
7. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: Smote: synthetic minority oversampling technique. Journal of artificial intelligence research 16, 321–357 (2002)
8. Chen, T., Guestrin, C.: Xgboost: A scalable tree boosting system. In: Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining. pp. 785–794 (2016)
9. Duan, J., Shi, D., Diao, R., Li, H., Wang, Z., Zhang, B., Bian, D., Yi, Z.: Deep-reinforcement-learning-based autonomous voltage control for power grid operations. IEEE Transactions on Power Systems 35(1), 814–817 (2020)
10. Friedman, J.H.: Greedy function approximation: a gradient boosting machine. Annals of statistics pp. 1189–1232 (2001)
11. Gomez, F.R., Rajapakse, A.D., Annakkage, U.D., Fernando, I.T.: Support vector machine-based algorithm for post-fault transient stability status prediction using synchronized measurements. IEEE Transactions on Power Systems 26(3), 1474–1483 (2011)
12. Hochreiter, S., Schmidhuber, J.: Long short-term memory. Neural Computation 9(8), 1735–1780 (1997)
13. Hong, W.C.: Hybrid evolutionary algorithms in a svr-based electric load forecasting model. International Journal of Electrical Power & Energy Systems 31(7), 409–417 (2009), https://www.sciencedirect.com/science/article/pii/S0142061509000507
14. Hossain, N., Hossain, S.R., Azad, F.S.: Univariate time series prediction of reactive power using deep learning techniques. In: 2019 International Conference on Robotics,Electrical and Signal Processing Techniques (ICREST). pp. 186–191 (2019)
15. Hua, H., Qin, Y., Hao, C., Cao, J.: Optimal energy management strategies for energy internet via deep reinforcement learning approach. Applied Energy 239, 598–609 (2019), https://www.sciencedirect.com/science/article/pii/S0306261919301746
16. Kang, Y.W., Seo, C.S., Han, B.J., Jang, Y.H., Song, B.C., Kim, D.H.: The development of voltage stability system(vss) device for variable shunt reactor(vsr). Proceedings of the Korean Electrical Society Conference pp. 881–882 (2021)

17. Ke, G., Meng, Q., Finley, T., Wang, T., Chen, W., Ma, W., Ye, Q., Liu, T.Y.: Lightgbm: A highly efficient gradient boosting decision tree. Advances in neural information processing systems 30 (2017)
18. Kim, K., Lee, J.H., Lim, H.K., Oh, S.W., Han, Y.H.: Deep rnn-based network traffic classification scheme in edge computing system. Computer Science and Information Systems 19(1), 165–184 (2022)
19. Ko, C.N., Lee, C.M.: Short-term load forecasting using svr (support vector regression)-based radial basis function neural network with dual extended kalman filter. Energy 49, 413–422 (2013), https://www.sciencedirect.com/science/article/pii/S0360544212008766
20. Lee, M.C., Chang, J.W., Hung, J.C., Chen, B.L.: Exploring the effectiveness of deep neural networks with technical analysis applied to stock market prediction. Computer Science and Information Systems 18(2), 401–418 (2021)
21. Li, Y., Zhang, M., Chen, C.: A deep-learning intelligent system incorporating data augmentation for short-term voltage stability assessment of power systems. Applied Energy 308, 118347 (2022), https://www.sciencedirect.com/science/article/pii/S0306261921015944
22. Liaw, A., Wiener, M., et al.: Classification and regression by randomforest. R news 2(3), 18–22 (2002)
23. Medsker, L.R., Jain, L.: Recurrent neural networks. Design and Applications 5, 64–67 (2001)
24. Park, J.H.: Voltage regulating device (Sh. C) automatic operation system expansion and construction. Journal of Electrical World Monthly Magazine , 28–34 (2013), http://www.dbpia.co.kr/journal/articleDetail?nodeId=NODE02084990
25. Sain, S.R.: The nature of statistical learning theory (1996)
26. Viawan, F.A., Karlsson, D.: Combined local and remote voltage and reactive power control in the presence of induction machine distributed generation. IEEE Transactions on Power Systems 22(4), 2003–2012 (2007)
27. Viawan, F.A., Karlsson, D.: Voltage and reactive power control in systems with synchronous machine-based distributed generation. IEEE Transactions on Power Delivery 23(2), 1079–1087 (2008)
28. Wang, S.C.: Artificial Neural Network, pp. 81–100. Springer US, Boston, MA (2003), https://doi.org/10.1007/978-1-4615-0377-4_5
29. Xu, H., Dominguez-Garcia, A.D., Sauer, P.W.: Optimal tap setting of voltage regulation transformers using batch reinforcement learning. IEEE Transactions on Power Systems 35(3), 1990–2001 (2020)
30. Xu, Y., Dong, Z.Y., Zhao, J.H., Zhang, P., Wong, K.P.: A reliable intelligent system for real-time dynamic security assessment of power systems. IEEE Transactions on Power Systems 27(3), 1253–1263 (2012)
31. Yang, J., Yang, S., Song, K., Liu, Z.: Research on overvoltage identification method of emus high voltage electrical system based on deep learning. In: 2021 IEEE 4th Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC). vol. 4, pp. 1985–1990. IEEE (2021)
32. Yin, L., Zhang, C., Wang, Y., Gao, F., Yu, J., Cheng, L.: Emotional deep learning programming controller for automatic voltage control of power systems. IEEE Access 9, 31880–31891 (2021)
33. Yu, L., Qu, J., Gao, F., Tian, Y.: A novel hierarchical algorithm for bearing fault diagnosis based on stacked lstm. Shock and Vibration 2019 (2019)
34. Yuan, X., Yuan, Y., Zhang, Y.: A hybrid chaotic genetic algorithm for short-term hydro system scheduling. Mathematics and Computers in Simulation 59(4), 319–327 (2002), https://www.sciencedirect.com/science/article/pii/S0378475401003639
35. Zhang, X., Zhou, X., Lin, M., Sun, J.: Shufflenet: An extremely efficient convolutional neural network for mobile devices. In: Proceedings of the IEEE conference on computer vision and pattern recognition. pp. 6848–6856 (2018)

36. Zheng, J., Xu, C., Zhang, Z., Li, X.: Electric load forecasting in smart grids using long-short-term-memory based recurrent neural network. In: 2017 51st Annual Conference on Information Sciences and Systems (CISS). pp. 1–6 (2017)
37. Zhu, L., Lu, C., Dong, Z.Y., Hong, C.: Imbalance learning machine-based power system short-term voltage stability assessment. IEEE Transactions on Industrial Informatics 13(5), 2533–2543 (2017)
38. Zou, H., Hastie, T.: Regularization and variable selection via the elastic net. Journal of the royal statistical society: series B (statistical methodology) 67(2), 301–320 (2005)

Jiyong Moon, Department of Business Administration, Dongguk University, Seoul, Korea.

Minyeong Son, Department of Medical Biotechnology, Dongguk University, Seoul, Korea.

Byeongchan Oh,Department of Statistics, Dongguk University, Seoul, Korea.

Jeongpil Jin, Department of Industrial System Engineering, Dongguk University, Seoul, Korea.

Younsoon Shin, Department of Computer Science, Dongguk University, Seoul, Korea.

Received: May 09, 2022; Accepted: October 26, 2023.

Computer Science and Information Systems 21(2):453–472 https://doi.org/10.2298/CSIS220514051C

The Effects of Process Innovation and Partnership in SCM: Focusing on the Mediating Roles

Yoonkyo Cho[1] and Chunsu Lee[2,⋆]

[1] Dept. of K-Internet Business Management
Halla University, Wonju 26404, Korea
yoonkyo.cho@halla.ac.kr
[2] Dept. of International Trade,
Pukyong National University, Busan 48513, Korea
leecs@pknu.ac.kr

Abstract. In this study, we examined the impact of supply chain management factors on firm performance, and we focused on the mediating role of process innovation and partnerships. For the analysis, we surveyed 193 workers working in smartphone manufacturing companies. We found that information systems, support of top management, and performance management have positive impacts on a company's process innovation. The factors that affect partnership are the support of top management and performance management. Process innovation and partnership also positively affect a firm's financial and nonfinancial performance. Nonfinancial performance also shows effectiveness. Thus, to improve a firm's supply chain management (SCM) performance, companies should focus on enhancing process innovation and partnerships that positively affect firm performance. Furthermore, this research can serve as a stepping stone for the development of SCM in line with the technological innovation of Industry 4.0.

Keywords: process innovation, partnership, SCM factors, industry 4.0.

1. Introduction

The industrial environment is changing rapidly. In this environment, efficient supply chain management (SCM) is essential for companies to achieve high performance. Especially in the smartphone market, the life cycle of products—smartphones and their components—is shortening. Short life cycles increase the risk of product loss. This leads to intense global competition in the industry.

The smartphone manufacturing industry is a system of producing finished products in cooperation with each other, from raw material companies to parts manufacturers and finished goods-producing companies. This means that organic activities between companies on the supply chain (SC) line are critical to securing corporate competitiveness. Therefore, research on partnerships between companies is needed to ensure competitiveness in a complex business environment.

In addition, process innovation is perceived as an essential factor of the company's management strategy and performance. Process innovation is studied by many researchers to achieve and maintain an edge in competition over competitors [1,2,50]. Therefore, this

⋆ Corresponding author

study's results will demonstrate the relationships between SCM's key elements and firm performance.

To have a competitive advantage, firms need to solve the various difficulties in management. SCM performance is economically inefficient in the smartphone industry, as shown in Figure 1. Consumers' lack of awareness and understanding was the highest at 29%. Conditions that make it difficult to hire experts came in second with 18.1%. Other reasons include insufficient initial investment, lack of awareness by executives, and current systems' incompatibilities.

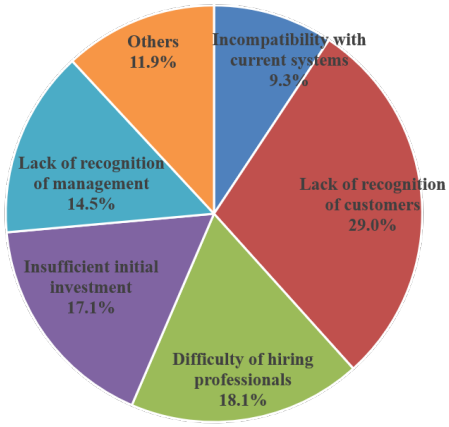


Fig. 1. The reasons for ineffective SCM performance

There are three contributions made in this study. First, we focus on intermediate companies (suppliers) in the smartphone industry. Prior research focused on companies dealing with complete products. However, it is essential for companies dealing with intermediate goods to link SCM with raw material companies located in the front of the SC and for SCM cooperation to work with final product companies in the rear. Thus, dealing with intermediate parts companies can demonstrate the importance of process innovation and intercompany partnerships to a firm's performance in its SCM operations.

Second, we suggest that both internal and external factors are important for a firm's performance. Because of the nature of smartphone parts companies with short product life cycles, it is necessary to reduce time and cost to survive and be competitive. Process innovation is what makes this possible. Thus, companies can improve their performance through process innovation internally. In contrast, collaboration between forward and backward companies is an essential factor because of the nature of intermediate parts companies. Therefore, improving and developing these matters can lead to high management performance. Therefore, to have superior performance regarding SCM, both process innovation (an internal factor) and partnership (an external factor) are important.

Third, we take a balanced approach to performance measurement. We examine the effects of SCM factors, process innovation, and partnership on both financial and non-financial performance. In most cases, management performance deals with nonmonetary performance or only monetary performance [8,31,44,45,65]. However, we have addressed both management performances and found that nonmonetary management performance positively affects monetary management performance. Therefore, both types of management performance can be crucial factors for the survival and growth of a company. This measurement can present a clear picture of organizational performance.

The order of the remainder of this paper is as follows. Section 2 introduces the theory and concept behind this research. Section 3 presents the data and methodology used. Section 4 presents the main results of the study. Section 5 concludes the study.

2. Theoretical Background and Hypotheses

2.1. Supply Chain Management

Ellram and Cooper [14] stated that reducing inventory investment, increasing customer service, and gaining a competitive advantage on the supply chain are the core of SCM. Lambert et al. [35] stated that SCM is a strategy that creates added value across the supply chain—such as products, services, and information—by integrating and operating processes from the initial supplier to the end user to the related businesses and customers.

Before the mid-1990s, when SCM was introduced in earnest, the concept of logistics was widely used [56]. This concept included the integration of other functions as part of an effort to achieve an entity's overall performance [46]. At this time, production-oriented planning and management, procurement of parts and raw materials, and sales and distribution processes were operated separately. Thus, manufacturers had to comply with delivery times on their own in the operation of manufacturing lines, increasing productivity and reducing inventory.

Entering the mid-1990s, SCM evolved into a concept that could create value through the coordination of functions outside the enterprise and various business functions within the organization. Currently, many companies have adopted SCMs to integrate logistics, information, and financial-related businesses and to build improved systems that have been limited within the firm-specific optimization. This has led to competition as a supply chain for several companies beyond a single enterprise. In addition, by improving the efficiency of business processes through information sharing between different businesses and organizations in the supply chain, inventory is reduced, and unnecessary logistics costs are minimized. Moreover, this increases customer satisfaction by improving management speed [25].

Most researchers' definition of SCM is subtly different, but in most studies, researchers define the core of SCM as a management technique that can increase customer satisfaction by connecting and managing all processes from the production stage of the product to the delivery to the consumer.

2.2. SCM and Industry 4.0

Industry 4.0 technology is developing rapidly around the world. The fourth industrial revolution predicts that artificial intelligence (AI) in the supply chain will gradually increase

the use of AI automation [41]. Because of the benefit of new technology, transportation and communication charges will be reduced, logistics and global supply chains will be operated more efficiently, and transaction costs will be reduced. All of this is expected to open new markets and trigger economic growth. This shows that the impact of the fourh industrial revolution will play a big role in supply chain management as well. The characteristics of the fourth industrial revolution affecting supply chain management are as follows.

First, robotics affects the supply chain process [13]. Many production processes already use pick-and-place robots that pick up objects and place them in designated locations. Daniela Rus, director of MIT's Computer Science–Artificial Intelligence Lab, predicts customized robots automating tasks in a wide range of areas. AI custom robots differ from conventional robots and reduce the time needed to equip automation in industries that rely on custom orders and short product life cycles. The robots know where to store data and how to assemble products, thereby increasing the efficiency of SCM.

The second is the use of big data. Big data refers to large-scale data with a shorter generation period and includes text and image data as well as numerical data. In the supply chain process, big data can be used to identify transportation information that identifies real-time transportation locations and problems based on past and present data. In addition, big data can predict traffic congestion or risk and identify expected arrival and delay times, weather events, and natural disasters. The use of such big data can greatly contribute to the efficiency of the supply chain by providing an optimal environment for logistics operations [61].

Third is the application of the Internet of Things (IoT). The IoT refers to intelligent technologies and services connecting all things based on the Internet to communicate information between people and things and between things and things. In other words, things establish a relationship with humans based on interconnected technology. The IoT is most widely used in remote monitoring technology. In the case of the transportation industry, companies can attach sensors to all boxes, trucks, and containers to obtain location information whenever they move. Consumers can also check when and where the goods they have purchased arrive in real time. With the development of the IoT, collecting various data generated in the logistics process is possible, and information that was difficult to grasp in the past supply chain management system can be grasped [32].

Fourth is the advent of unmanned transportation. Recently, drones have been in the spotlight as unmanned autonomous vehicles (UAVs), and more and more companies are using them. With the development of UAV technology, drones, boats, and aircraft have emerged as unmanned transportation means. UAVs in particular are developing quickly. UAVs will dramatically replace the role of existing transportation means. The use of suitable unmanned transportation means enables companies to increase supply chain management's performance (i.e., efficiency and effectiveness [48]).

2.3. SCM Factors, Process Innovation, and Partnership

Information System The introduction and utilization of an integrated information system for supply chain management not only increases quality, shortens delivery time, and reduces costs but also ultimately enhances the competitiveness of supply chain management for the continuous growth of firms [20,27,34]. Active use of information technology

and the standardization of products and data are required to increase the introduction effect of this system and enhance the competitiveness of firms. It is necessary to establish information systems such as point of sales, electronic data interchange, and electronic ordering systems for smooth information exchange between business organizations in the supply chain. The information system constructed in this way is premised on the accuracy of information sharing and information delivery between members and aims to standardize information systems and information linkage among organizational members. The maturity level of an organization's information system depends on how well it can be used for business applications or strategic purposes after the organization's information system is built [28]. Therefore, the higher the maturity of an information system, the easier it will be to use the system without difficulty, and the spread of this information system will have a greater impact on firm performance after SCM implementation. Companies with high information technology (IT) capabilities can be more active in information sharing between business processes. When business processes between companies are integrated along the value chain through information sharing, firm performance can be maximized. IT solutions are critical in realizing the abundant benefits of supply chain management implementation [39]. To exchange and share information flawlessly both inside and outside of the company, building a sound information system infrastructure and utilizing information technology are necessary. Therefore, the company's advanced information system will play a positive role in corporate performance by integrating internal and external SC processes of the company.

Hypothesis 1a *Information system has a positive effect on process innovation.*

Hypothesis 1b *Information system has a positive effect on partnership.*

Support from Top Management The will of the CEO plays a vital role in shaping the direction and values of the organization [33], is essential for cooperation between companies [47], and has a significant impact on the performance of the company [12]. The CEO's will, leadership, and commitment to change are major antecedents influencing successful SCM implementation [35]. For the same reason, the lack of the will of the CEO is a significant obstacle to the implementation of SCM [38]. The will of the CEO has a significant impact on the adoption and utilization of strategic systems such as interorganizational information systems and is also important for overcoming barriers and resistance to change and innovation [57].

As an innovation leader within the organization, the top management should properly recognize the characteristics and factors of SCM. If a new SCM is introduced in the existing organizational work process, it may face opposition from organizational members because it will bring about innovative changes. Because it is necessary to establish a new SCM through continuous support from the CEO, the CEO plays an essential role in the introduction and diffusion of information systems [9]. In particular, the introduction of the intercompany information system in SCM is a large-scale project that requires innovation of intercompany relationships and complex supply chains, so continuous investment is necessary for a certain period of time. In this process, the top management's support is most important to minimize the opposition of organizational members and to induce the participation of members in the innovation process. In addition, the CEO's support

is necessary to successfully establish a cooperative relationship that maintains a lasting relationship among several business partners outside the company.

Hypothesis 2a *Support from top management has a positive effect on process innovation.*

Hypothesis 2b *Support from top management has a positive effect on partnership.*

Planning For effective supply chain management, the accuracy and appropriateness of demand planning that leads the entire supply chain are essential [42,60]. Recent advances in IT are rapidly shortening the planning cycle for the supply chain. For example, the current trend is for SC plans to be implemented on a weekly, daily, and even shift basis. Rapid response to demand fluctuations through optimization can generate plans closer to market conditions by reflecting the constraints of the entire supply chain in real time. This plan is optimized to meet the supply chain demand, considering the limitations of equipment and materials for each base.

The results of a company's effective planning are no longer dependent on individual company profits or growth but rather on how well its members collaborate throughout the supply chain. Therefore, it is necessary to strengthen competitiveness based on collaboration among members of the supply chain [62]. As the need for such a collaboration to implement efficient planning systems increases, the supply chain has been developing by gradually expanding the exchange of information. Moreover, the development of IT and the emergence of e-business allow members to cooperate by forming a supply chain on the Web [36].

To establish a supply chain management system for a rapid market response, planning should play a role in improving the accuracy of demand planning and extending the range of collaboration, leading to a positive effect on the company's performance.

Hypothesis 3a *Planning has a positive effect on process innovation.*

Hypothesis 3b *Planning has a positive effect on partnership.*

Performance Management Even if a company uses an appropriate SCM, it cannot operate effectively if the performance achieved is not properly monitored and measured. An effective SCM performance measurement system improves the understanding of SCM, influences the behavior of organizational members, and provides information about the system's performance. Ultimately, measuring supply chain performance improves overall performance [52, 53]. Through the performance measurement system, it is important to set or improve the company's target by comparing it with the performance measurement of other companies in the same business category. And it is a necessary element for the growth of a company to establish an improvement direction and strategy to overcome the inferior performance on the part of the company through comparison with promising companies in the same industry.

Performance management can also be implemented through performance sharing between partners. Performance sharing reduces potential risks in the supply chain and increases profits through goal-oriented relationship building by forming common goals [49, 51]. If a shared performance goal is set and personnel inside or outside the company can perform joint production and research and development activities to achieve the goal, more open innovation can be achieved, positively affecting the company's performance.

Hypothesis 4a *Performance management has a positive effect on process innovation.*

Hypothesis 4b *Performance management has a positive effect on partnership.*

2.4. Process Innovation, Partnership, and Performance

Process Innovation Process innovation is a change to establish an efficient and effective organizational system, enabling a company to quickly respond to customer needs and flexibly respond to distribution channels and new environments. When the added value is generated transparently in the process from the purchase stage of a product to the final consumer, the efficiency of corporate management can be increased, and competitiveness can be achieved [59]. For this, innovation in the entire SC line should be organically developed and should prompt the processing of customer orders. For process innovation to be successful, it is necessary to coordinate and manage an efficient system through information sharing using information technology in its internal organization and the connected chain outside the company.

Today's process innovation reduces time and cost from the input of goods to the final output [66], achieves customer satisfaction by improving product quality, and eliminates various obstacles in inventory management through rapid transportation. Accuracy of demand forecasting due to process innovation can reduce delays by confirming actual sales based on manufacturers' point-of-sale information. As a result, reasonable inventory adequacy can be maintained. A company's efficient inventory management reduces excessive inventory levels by improving production technology [6]. In particular, prompt provision of sales information allows manufacturers to effectively maintain proper inventory and dramatically reduce the lead time required to produce items.

Rapid response due to SCM process innovation also affects efficiency increase. To respond quickly to customer orders, supply chain integration between the company's internal and external organizational structures is performed to increase customer satisfaction, thereby affecting the company's performance [18, 37]. Companies that have incorporated process innovation into their supply chain can secure a competitive advantage over their competitors and increase the efficiency of corporate management [29].

Process innovation enables changes in the entire process, from designing or introducing a new idea to practical use of skills and technologies by organization members. Through the propagation of new technology by the adoption of potential innovation, an organization advances over time. In addition, process simplification, standardization, and integration develop high-quality strategic systems and improve the quality of information support services for users, which will positively affect corporate performance.

Hypothesis 5a *Process innovation has a positive effect on a firm's financial performance.*

Hypothesis 5b *Process innovation has a positive effect on a firm's nonfinancial performance.*

Partnership One of the topics highlighted in recent supply chain management research is a collaboration among members of the supply chain [40]. This is because instead of maximizing its own profits, it is possible for a company to seek opportunities for greater business performance by forming cooperative relationships with partners.

Companies' efforts to form collaborative relationships and maintain close relationships with key partners can usually be discussed from three perspectives. The first is the point of view of transaction costs theory. The theory suggests that a company that increases investment in the specificity of transactions between companies can improve the performance of the supply chain by reducing the coordination cost and motivation cost associated with the transaction compared to a company that does not [64]. The second perspective is information processing theory. With this theory, a company seeks to overcome market uncertainty and improve corporate performance by sharing or integrating information or resources among companies. The third perspective is a sociopolitical approach. This is when a company seeks to increase the supply chain's performance by establishing intercompany relationships such as intercompany partnerships or strategic alliances [3, 30].

When the concept of collaboration is expressed from a sociopolitical approach, it is understood that collaboration is from the perspective of behaviors that appear according to the relationship between companies and the establishment of partnerships or strategic alliances between companies, joint coordination efforts on an equal footing, and flexibility according to the occurrence of situations. Collaborative behavior can be divided into two main aspects according to the integrated form of decision-making. The first is to jointly deal with problems that may arise in business-to-business transactions, such as joint task-solving actions. These actions appear throughout SCM, including production planning and operations, procurement, order processing, engineering design, and business integration. In particular, the more buyers and suppliers participate at the beginning of the planning process, the higher the opportunity to innovate a product or service.

The second aspect is the flexibility of collaboration between the parties to the transaction. In general, flexibility refers to the ability to cope with changes in an uncertain environment [24]. The uncertainty associated with the transaction between the buyer and the supplier is considerable. For example, unforeseen circumstances may result in a change in order quantity, additional costs, a request for service that exceeds the contractual terms, a request for replacement of new material, or delivery to a particular specification. In this case, if new contract terms can be concluded between companies or if the two companies can solve such problems jointly, the performance of the supply chain can be significantly improved compared to the case where it is not possible.

Partnerships between companies are important to design; they integrate the management of supply chain activities and enhance efficiency in supply chain operations. When partnerships between companies are formed, companies participating in the supply chain expand the scope of collaboration, such as information sharing, synchronization of plans, integration of business processes, and the creation of new business models, and increase the utilization of business processes between companies. Considering that this collaboration takes place within the supply chain, the performance of the supply chain will appear through the performance of the partnership.

Hypothesis 6a *Partnership has a positive effect on a firm's financial performance.*

Hypothesis 6b *Partnership has a positive effect on a firm's nonfinancial performance.*

2.5. Nonfinancial Performance and Financial Performance

Nonfinancial performance plays an important role in SCM as well as a company's financial performance. Reputation can be considered part of the nonfinancial performance of a company. It is the cumulative result of the perceived image of an organization's management-related attitudes and activities over a long period of time [15]. Reputation is a comprehensive evaluation according to the perception of stakeholders outside the organization. Also, as a thorough evaluation of the organization's consistent reliability and integrity, it is a term that includes expectations for future activities and evaluations of past activities of the organization.

When a company seeks to establish a business relationship with a new company, it pursues minimizing risks associated with the transaction. Adverse selection and moral hazard caused by a transaction have negative results for a company, so reputation can be used in a sufficient monitoring process before a transaction. For example, a company may refer to the evaluations or rumors circulated about the new counterparty by a third party who has done business with the company. Therefore, organizations strive to have a favorable reputation among their members. They also try to establish relationships and networks with companies with favorable reputations, exclude organizations with unfavorable reputations, or cut off business relationships [4]. A good reputation can be a valuable asset, and a considerable amount of time and money is invested in an individual company to have a favorable reputation. Reputation plays a role in limiting opportunistic behaviors in business relationships and affects trust. When the level of reputation of the trading company is good or excellent, the level of credit toward the trading company is also improved [19]. Transactions with partners that have good reputations lead to minimized transaction costs, including the financing cost. Therefore, good nonfinancial performance will positively affect a company's financial performance.

Hypothesis 7 *Nonfinancial performance has a positive effect on financial performance.*

3. Methodology

3.1. Research Model

In this study, we analyzed the effect of SCM factors on a company's performance, focusing on the mediating effect of process innovation and partnership. The research model of this study based on the hypotheses is shown in Figure 2.

3.2. Data

The subjects of this study were smartphone parts manufacturing companies operating SCM. We directly visited the companies located in Busan. We explained the purpose of the questionnaire to the other parts manufacturing companies through email and distributed 230 copies of the questionnaires. We collected a total of 206 questionnaires. Of these, we used 193 as the data for this study, excluding the questionnaires containing missing responses. We measured all study variables on a 5-point Likert scale. The characteristics of the 193 smartphone component manufacturers surveyed are shown in Figures 3–5.

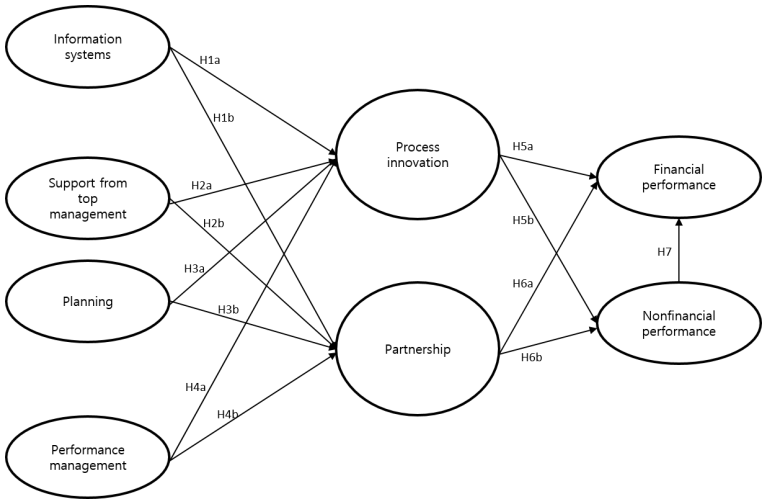


Fig. 2. Research model

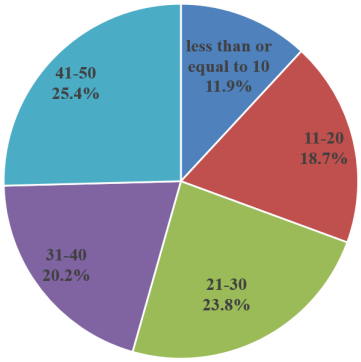


Fig. 3. Firm age (Year)

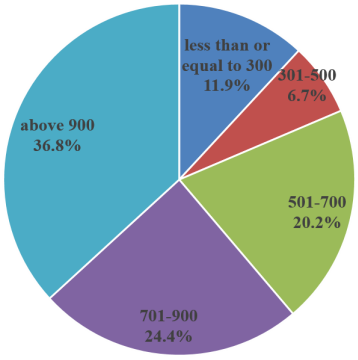


Fig. 4. Firm size (Number of employees)

First, 25.4% of companies were 41–50 years old, followed by 23.8% for 21–30 years, 20.2% for 31–40 years, and 18.7% for 11–20 years, and 11.9% for less than 10 years. The following companies were in the order of 11.9%. The number of employees with more than 900 employees was the highest at 36.8%, followed by 24.4% with between 700 and 900 employees, 20.2% with between 500 and 700 employees, 11.9% with fewer than 300 employees, and 6.7% with between 300 and 500 employees. In the case of sales in the previous year, results showed that sales amounted to 31.6% of the companies with more than 100 billion won, followed by 25.4% of companies with more than 70 billion won, 22.8% of companies with more than 50 billion won, 11.4% of companies with more than 10 billion won, and 8.8% of companies with less than 10 billion won.

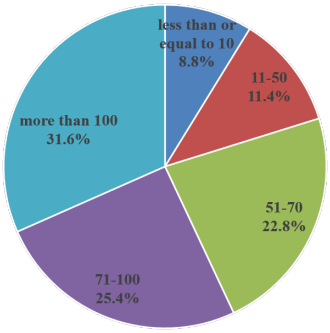


Fig. 5. Sales (Billion won)

3.3. PLS Structural Equation Research Model

In this study, we applied the partial least square structural equation model (PLS-SEM) to analyze the effect of SCM factors on company management performance. The structural equation model is a more powerful analytical method than traditional multivariate analysis. It can indirectly measure nonobservable potential variables through observable measurement variables and explain the measurement error of observed variables. It is widely used in the field of social science research. Because PLS-SEM estimates the path coefficients to maximize the explanatory power (R^2) by minimizing the error term of endogenous latent variables, it focuses on the explanation and prediction of intrinsic latent variables corresponding to dependent variables rather than the structural characteristics of the model. Therefore, it is more suitable for theory development and exploratory research. PLS-SEM shares all assumptions in multiple regression analysis and creates a predictive model when there are a large number of factors or very high multicollinearity. The PLS-SEM can be effectively applied to small sample sizes and complex models with virtually no assumptions regarding the distribution of the data to be analyzed and can easily include formative measurement models and reflective measurement models. Single-item potential variables can also be applied without model identification problems [23].

4. Results

4.1. Validity and Reliability

Validity refers to how accurately a measurement instrument measures the concept or property that it is trying to measure. The purpose of this study was to examine the validity of SCM factors of smartphone component makers as independent variables, process innovation and intercompany partnerships of smartphone component manufacturers as mediators, and nonmonetary and monetary management performance measures as dependent variables. To verify the validity, the research factors were composed of a measurement model and the confirmatory factor analysis of the research factors. Table 1 represents the measurement of each variable.

Table 1. Measurement items for study constructs

Constructs	Measurement	Literature
Information systems	IS01 IT is implemented in various services and functions. IS02 Information is shared across functions. IS03 Expense of operating IT technology is reasonable.	[39]
Support from top management	TS01 CEO highly pays attention to SCM initiatives. TS02 CEO actively invests in SCM adoption and utilization.	[26, 43]
Planning	PN01 Implementable plans are established for production/sales at the supply chain level. PN02 Plans for SCM are set periodically. PN03 Expectations for SCM are clearly stated, understood, and agreed to up front.	[16, 58]
Performance management	PM01 Measures are established systematically for performance. PM02 Activities of employees are reported for performance management. PM03 Roles and responsibilities and incentives are specified clearly.	[21, 52]
Process innovation	PI01 SCM improves and manages processes in an enterprise. PI02 Top management is actively involved in the exploration of challenges for process innovation. PI03 The company possesses a mechanism by which process innovation can be applied to other functions. PI04 The company possesses systems to maintain and manage changes in processes.	[5, 63]
Partnership	PS01 There is on-time delivery to partner firms. PS02 Our partner initiates contracts. PS03 We share information with partner firms in timely manner.	[7, 54]
Financial performance	FP01 Revenue is increased. FP02 Marginal profit is increased. FP03 Inventory costs is reduced.	[21]
Nonfinancial performance	NF01 Flexibility is improved in SCM. NF02 Rate of damage-free in the production is increased. NF03 Reputation is improved.	[22]

Through the confirmatory factor analysis, items that lowered the factor load or impaired the fit of the measurement model were removed, and the factors of SCM consisted of three items of information system, two items of support from top management, three items of planning, and three items of performance management. The final metrics consisted of four items for process innovation and three items for partnership among companies. In addition, three questions each consisted of nonmonetary and monetary outcomes as dependent variables. Because all the extracted values show more than 0.6, there seems to be no problem with the validity of the variables. Table 2 shows the results of the factor analysis conducted with the validation.

We also performed reliability verification. Table 3 shows the results of the reliability analysis. As a result of reviewing the reliability of the final metric, Cronbach's α coefficient was 0.637 for the information management factor, 0.771 for the activation support factor, 0.727 for the planning and collaboration factor, and 0.744 for the process innovation factor. The partnership factor between companies was 0.642, the nonmonetary performance was 0.664, and the monetary performance factor was 0.715. Every coefficient of Cronbach's α is above 0.6, and the constructive reliability is acceptable [11, 55].

Next, the concept reliability and average variance extraction (AVE) were reviewed to examine the concentration validity of latent factors.

Table 2. The value of cross-loading

		Info	Support	Plan	Perform	Innov	Partner	Fin	Nonfin
Info	IS01	**0.823**	0.273	0.326	0.431	0.352	0.297	0.338	0.196
	IS02	**0.777**	0.225	0.214	0.352	0.265	0.198	0.257	0.093
	IS03	**0.667**	0.204	0.313	0.350	0.200	0.227	0.213	0.137
Support	TS01	0.294	**0.920**	0.315	0.362	0.317	0.315	0.320	0.232
	TS02	0.268	**0.883**	0.271	0.292	0.237	0.289	0.318	0.180
Plan	PN01	0.362	0.274	**0.869**	0.451	0.272	0.323	0.355	0.140
	PN02	0.299	0.316	**0.839**	0.382	0.245	0.224	0.210	0.125
	PN03	0.226	0.189	**0.692**	0.287	0.198	0.182	0.244	0.122
Perform	PM01	0.428	0.333	0.388	**0.837**	0.332	0.376	0.422	0.208
	PM02	0.335	0.266	0.389	**0.792**	0.230	0.310	0.259	0.222
	PM03	0.451	0.286	0.381	**0.808**	0.346	0.281	0.331	0.240
Innov	PI01	0.284	0.200	0.156	0.267	**0.692**	0.211	0.209	0.189
	PI02	0.227	0.166	0.230	0.137	**0.747**	0.255	0.230	0.178
	PI03	0.283	0.201	0.270	0.305	**0.681**	0.298	0.227	0.129
	PI04	0.257	0.290	0.202	0.331	**0.725**	0.293	0.299	0.292
Partner	PS01	0.276	0.177	0.272	0.315	0.311	**0.756**	0.339	0.199
	PS02	0.259	0.353	0.229	0.341	0.238	**0.801**	0.384	0.136
	PS03	0.203	0.227	0.215	0.255	0.319	**0.731**	0.315	0.226
Financial	FP01	0.367	0.359	0.340	0.409	0.328	0.400	**0.861**	0.320
	FP02	0.311	0.308	0.276	0.349	0.286	0.397	**0.880**	0.327
	FP03	0.168	0.150	0.190	0.238	0.202	0.285	**0.642**	0.293
Nonfinancial	NF01	0.182	0.235	0.073	0.219	0.271	0.159	0.344	**0.779**
	NF02	0.082	0.092	0.112	0.198	0.218	0.207	0.238	**0.802**
	NF03	0.192	0.213	0.203	0.221	0.166	0.199	0.323	**0.735**

First, concentration validity represents the degree of correlation between two or more measurement items for a potential factor. If the concept reliability is 0.7 or more [10] and the AVE index is 0.5 or more, the concentration validity is acceptable. The concept reliability is more than 0.7 in all variables, and the AVE value is more than 0.5, which proves the validity of potential factors.

4.2. Validation of Research Hypotheses

Correlation Analysis The correlations among potential factors, such as SCM factors, process innovation, partnerships between companies, and management performance of smartphone parts manufacturing companies, are shown in Table 4. Numbers in bold type with diagonal lines represent the squared root of AVE. Because this number is larger than the other nondiagonal numbers, the component has a reasonable level of discriminant validity [17].

Table 3. Reliability

	Item	Mean	SD	Weight	Cronbach's α	Composite Reliability	AVE
InfoSys	IS01	3.83	0.93	0.823	0.637	0.802	0.576
	IS02	3.89	0.82	0.777			
	IS03	3.86	0.82	0.667			
TmtSupport	TS01	3.77	0.86	0.920	0.771	0.897	0.813
	TS02	3.63	0.89	0.883			
Plan	PN01	4.01	0.97	0.869	0.727	0.844	0.646
	PN02	3.89	0.83	0.839			
	PN03	3.84	0.89	0.692			
PerfMgt	PM01	4.03	0.88	0.837	0.744	0.853	0.660
	PM02	3.80	0.92	0.792			
	PM03	3.82	0.87	0.808			
ProcessInnov	PI01	3.67	0.88	0.692	0.679	0.804	0.507
	PI02	3.50	0.93	0.747			
	PI03	3.58	0.93	0.681			
	PI04	3.68	0.87	0.725			
Partner	PS01	3.85	0.92	0.756	0.642	0.807	0.583
	PS02	3.95	0.89	0.801			
	PS03	3.89	0.85	0.731			
Nonfinancial	NF01	3.77	1.01	0.779	0.664	0.816	0.596
	NF02	3.78	0.96	0.802			
	NF03	3.90	1.03	0.735			
Financial	FP01	3.62	0.88	0.861	0.715	0.842	0.643
	FP02	3.70	0.77	0.880			
	FP03	3.95	0.84	0.642			

Table 4. Correlations of constructs

	Info	Support	Plan	Perform	Inno	Partner	Nonfin	Fin
Info	**0.759**							
Support	0.312	**0.902**						
Plan	0.376	0.326	**0.804**					
Perform	0.502	0.366	0.474	**0.812**				
Inno	0.370	0.311	0.300	0.377	**0.712**			
Partner	0.323	0.336	0.312	0.400	0.375	**0.763**		
Nonfin	0.194	0.231	0.160	0.274	0.287	0.242	**0.773**	
Fin	0.364	0.353	0.343	0.423	0.345	0.455	0.389	**0.802**

Note: Bold numbers show square root of AVE

Empirical Analysis In this study, the SCM factors of smartphone parts manufacturers were designed as independent variables, and the dependencies were designed to verify the causality of the SCM factors and management factors. An SEM analysis was conducted to look at the causal relationship between SCM factors, process innovation, intercompany partnerships, and management performance factors. Figure 6 shows the results.

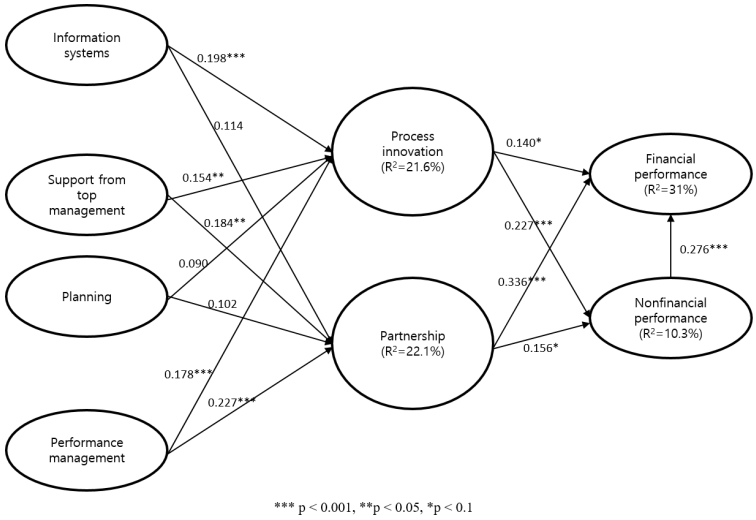


Fig. 6. Results

First, the information system had a significant effect on process innovation ($\beta = 0.198, p < 0.001$). Therefore, Hypothesis 1a was supported. The effect on partnership showed a positive signal but did not show a significant effect. Second, support from top management was found to have a significant effect on process innovation and partnership, respectively ($\beta = 0.154, p < 0.05; \beta = 0.184, p < 0.05$). Therefore, both Hypotheses 2a and 2b were supported. Here, we once again discover that the role of top management is important for improving SCM performance. Third, it was found that planning had no effect on process innovation and partnership. Fourth, performance management was found to have a significant effect on process innovation and partnership ($\beta = 0.178, p < 0.001; \beta = 0.227, p < 0.001$). Therefore, Hypotheses 4a and 4b were supported. Process innovation was found to have a positive effect on both the financial and nonfinancial performance of a company ($\beta = 0.140, p < 0.1; \beta = 0.227, p < 0.001$), and Hypotheses 5a and 5b were supported. Partnership was also found to have a positive effect on both financial and nonfinancial performance of a company ($\beta = 0.336, p < 0.001; \beta = 0.156, p < 0.1$), and Hypotheses 6a and 6b were supported. Finally, nonfinancial performance was found to have a positive effect on financial performance ($\beta = 0.276, p < 0.001$), and Hypothesis 7 was supported. The results for each hypothesis are summarized in Table 5.

Table 5. Summary of results

Hypothesis	Relationship	Beta	Std. Error	T Statistics	Support
1a	InfoSys → ProcessInnov	0.198	0.030	6.514	O
1b	InfoSys → Partner	0.114	0.037	3.034	x
2a	TmtSupport → ProcessInnov	0.154	0.033	4.699	O
2b	TmtSupport → Partner	0.184	0.031	5.983	O
3a	Plan → ProcessInnov	0.090	0.033	2.737	x
3b	Plan → Partner	0.102	0.036	2.862	x
4a	PerfMgt → ProcessInnov	0.178	0.029	6.110	O
4b	PerfMgt → Partner	0.227	0.036	6.243	O
5a	ProcessInnov → Financial	0.140	0.034	4.097	O
5b	ProcessInnov → Nonfinancial	0.227	0.032	7.205	O
6a	Partner → Financial	0.336	0.028	12.193	O
6b	Partner → Nonfinancial	0.156	0.036	4.357	O
7	Nonfinancial → Financial	0.276	0.028	9.777	O

5. Conclusions

To improve corporate performance, we examined how SCM factors affect corporate performance using two intermediates: process innovation and partnership. The results are as follows. First, top management support and performance management have positive significant effects on both process innovation and partnership. Second, an information system has a positive significant effect on process innovation. Third, both process innovation and partnership have a positively significant effect on financial and nonfinancial performance. Forth, nonfinancial performance has a positive effect on financial performance. Fifth, information systems have an insignificant effect on partnerships. Information sharing can have a positive effect on partnerships; however, if general staff answered the survey, it may be difficult to gain a detailed understanding of whether information sharing has a positive effect on the partnership. Lastly, planning has an insignificant effect on both process innovation and partnership. First, we conjecture that planning is related to maintenance and may not have much to do with process innovation and partnership. Second, if general staff answered the survey, the results would be insignificant because general staff members do not have much knowledge about the planning process.

Companies that produce fast-changing high-end products or components have different characteristics than those in other industries. In particular, high-tech goods companies change their cycles quickly because of the short life of the products they produce. As new technology development speeds up throughout the industry, these companies will likely survive if they can follow the faster cycle through internal process innovation. In addition, parts companies take raw materials, make intermediate parts, and deliver them to finished product companies. If there is a problem with the company supplying the raw material or if there is a problem with the company that produces the finished product, the company will interfere with the production schedule. Therefore, partnership with other companies is also crucial for companies producing intermediate goods.

From the result of this study, we provided important implications for managers. To have a good performance through SCM, companies need to focus more on the support of top management and performance management. Also, process innovation and partnership are critical factors that affect firms' performances. Although prior research does not equally weigh the importance of internal and external factors, there are of the same importance. Therefore, firms need to invest in process innovation and make appropriate relationships with their partners.

In process innovation, it is necessary to consider the following points. Depending on the degree of establishment and development of a company's production process, the extent to which process innovation is affected by SCM factors will be different. Also, different level of company's production process development may have different effects on the company's business performance. In this study, we verified the effect of process innovation on business performance, but we did not make a detailed classification of process innovation itself, which is a limitation. Therefore, future research is needed to systematically classify differences in the process establishment and development level of smartphone parts manufacturing companies and to investigate their performance.

Many fields of industry are facing changes due to the fourth industrial revolution—in particular, the advanced technologies of the 4th Industrial Revolution. Robotics, the IoT, big data, and unmanned transportation are expected to have a major impact on the overall SCM. For a company to achieve sustainable growth with a competitive advantage by utilizing this phenomenon, it is necessary to understand the existing SCM's characteristics and performance and to use that data to implement a new strategy.

In this study, we examined the performance of SCM for companies that currently produce high-tech products. Findings from this research can further serve as an important foundation for future research that measures the performance of other high-tech products or processes applied by Industry 4.0, such as artificial intelligence, the IoT, robotics, and big data within the SCM model.

References

1. Al-Sa'di, A.F., Abdallah, A.B., Dahiyat, S.E.: The mediating role of product and process innovations on the relationship between knowledge management and operational performance in manufacturing companies in jordan. Business Process Management Journal (2017)
2. Arshad Ali, A., Mahmood, A., Ikram, A., Ahmad, A.: Configuring the drivers and carriers of process innovation in manufacturing organizations. Journal of Open Innovation: Technology, Market, and Complexity 6(4), 154 (2020)
3. Bensaou, M.: Portfolios of buyer-supplier relationships. MIT Sloan Management Review 40(4), 35 (1999)
4. Berg, A., Gottschalg, O.F.: Understanding value generation in buyouts. Journal of Restructuring Finance 2(01), 9–37 (2005)
5. Blakeslee Jr, J.A.: Implementing the six sigma solution. Quality progress 32(7), 77 (1999)
6. Cachon, G.P., Fisher, M.: Supply chain inventory management and the value of shared information. Management science 46(8), 1032–1048 (2000)
7. Cadilhon, J.J., Fearne, A.P., Moustier, P., Poole, N.D.: Modelling vegetable marketing systems in south east asia: phenomenological insights from vietnam. Supply Chain Management: an international journal (2003)
8. Cao, M., Zhang, Q.: Supply chain collaboration: Impact on collaborative advantage and firm performance. Journal of operations management 29(3), 163–180 (2011)

9. Chandren, S., Qaderi, S.A., Ghaleb, B.A.A.: The influence of the chairman and ceo effectiveness on operating performance: Evidence from malaysia. Cogent Business & Management 8(1), 1935189 (2021)
10. Chin, W.W.: Commentary: Issues and opinion on structural equation modeling (1998)
11. Cossío-Silva, F.J., Revilla-Camacho, M.Á., Vega-Vázquez, M., Palacios-Florencio, B.: Value co-creation and customer loyalty. Journal of Business Research 69(5), 1621–1625 (2016)
12. Day, D.V., Lord, R.G.: Executive leadership and organizational performance: Suggestions for a new theory and methodology. Journal of management 14(3), 453–464 (1988)
13. Demir, S., Paksoy, T.: Ai, robotics and autonomous systems in scm. Logistics 4.0: Digital Transformation of Supply Chain Management p. 156 (2020)
14. Ellram, L.M., Cooper, M.C.: Supply chain management, partnership, and the shipper-third party relationship. The international journal of logistics management 1(2), 1–10 (1990)
15. Eltantawy, R.A., Fox, G.L., Giunipero, L.: Supply management ethical responsibility: reputation and performance impacts. Supply Chain Management: An International Journal (2009)
16. Fisher, M.L.: What is the right supply chain for your product? Harvard business review 75, 105–117 (1997)
17. Fornell, C., Larcker, D.F.: Evaluating structural equation models with unobservable variables and measurement error. Journal of marketing research 18(1), 39–50 (1981)
18. Frohlich, M.T., Westbrook, R.: Arcs of integration: an international study of supply chain strategies. Journal of operations management 19(2), 185–200 (2001)
19. Ganesan, S.: Determinants of long-term orientation in buyer-seller relationships. Journal of marketing 58(2), 1–19 (1994)
20. Goldhar, J.D., Lei, D.: The shape of twenty-first century global manufacturing. The Journal of Business Strategy 12(2), 37 (1991)
21. Gunasekaran, A., Kobu, B.: Performance measures and metrics in logistics and supply chain management: a review of recent literature (1995–2004) for research and applications. International journal of production research 45(12), 2819–2840 (2007)
22. Gunasekaran, A., Patel, C., Tirtiroglu, E.: Performance measures and metrics in a supply chain environment. International journal of operations & production Management (2001)
23. Hair Jr, J.F., Sarstedt, M., Hopkins, L., Kuppelwieser, V.G.: Partial least squares structural equation modeling (pls-sem): An emerging tool in business research. European business review (2014)
24. Heide, J.B., Miner, A.S.: The shadow of the future: Effects of anticipated interaction and frequency of contact on buyer-seller cooperation. Academy of management journal 35(2), 265–291 (1992)
25. Heikkilä, J.: From supply to demand chain management: efficiency and customer satisfaction. Journal of operations management 20(6), 747–767 (2002)
26. Higginson, J.K., Alam, A.: Supply chain management techniques in medium-to-small manufacturing firms. The International Journal of Logistics Management 8(2), 19–32 (1997)
27. Huggins, J.W., Schmitt, R.G.: Electronic data interchange as a cornerstone to supply chain management. In: Annual Conference Proceedings of the 1995 Council of Logistics Management (1995)
28. Imache, R., Izza, S., Ahmed-Nacer, M.: An enterprise information system agility assessment model. Computer science and information systems 9(1), 107–133 (2012)
29. Jeon, S.S., Lee, R.: Impact of scm system operation strategy on scm performance and mediating effect of process innovation. Journal of Theoretical and Applied Information Technology 100(5) (2022)
30. Ke, W., Wei, K.K.: Factors affecting trading partners' knowledge sharing: Using the lens of transaction cost economics and socio-political theories. Electronic Commerce Research and Applications 6(3), 297–308 (2007)

31. Kim, D., Cavusgil, S.T., Calantone, R.J.: Information system innovations and supply chain management: channel relationships and firm performance. Journal of the academy of marketing science 34(1), 40–54 (2006)
32. Kothari, S.S., Jain, S.V., Venkteshwar, A.: The impact of iot in supply chain management. International Research Journal of Engineering and Technology 5(8), 257–259 (2018)
33. Kotter, J.: A force for change: How management differs from leadership. New York: FreePress (1990)
34. LaLonde, B.J., Masters, J.M.: Logistics: perspectives for the 1990s. The International Journal of Logistics Management (1990)
35. Lambert, D.M., Cooper, M.C., Pagh, J.D.: Supply chain management: implementation issues and research opportunities. The international journal of logistics management 9(2), 1–20 (1998)
36. Le Tan, T., Thi Dai Trang, D.: Issues of implementing electronic supply chain management (e-scm) in enterprise. European Business & Management 3(5), 86–94 (2017)
37. Lee, S.M., Lee, D., Schniederjans, M.J.: Supply chain innovation and organizational performance in the healthcare industry. International Journal of Operations & Production Management (2011)
38. Loforte, A.J.: The implications of multicultural relationships in a transnational supply chain. In: National association of purchasing management annual conference proceedings. pp. 69–77 (1991)
39. Marien, E.J.: The four supply chain enablers. SUPPLY CHAIN MANAGEMENT REVIEW, V. 2, NO. 3 (FALL 1998), P. 60-68: ILL (2000)
40. Mentzer, J.T., Min, S., Zacharia, Z.G.: The nature of interfirm partnering in supply chain management. Journal of retailing 76(4), 549–568 (2000)
41. Min, H.: Artificial intelligence in supply chain management: theory and applications. International Journal of Logistics: Research and Applications 13(1), 13–39 (2010)
42. Min, H., Yu, W.B.: Collaborative planning, forecasting and replenishment: demand planning in supply chain management. International Journal of Information Technology and Management 7(1), 4–20 (2008)
43. Premkumar, G., Ramamurthy, K.: The role of interorganizational and organizational factors on the decision mode for adoption of interorganizational systems. Decision sciences 26(3), 303–336 (1995)
44. Qrunfleh, S., Tarafdar, M.: Supply chain information systems strategy: Impacts on supply chain performance and firm performance. International journal of production economics 147, 340–350 (2014)
45. Quang, H.T., Sampaio, P., Carvalho, M.S., Fernandes, A.C., An, D.T.B., Vilhenac, E.: An extensive structural model of supply chain quality management and firm performance. International Journal of Quality & Reliability Management (2016)
46. Quinn, F.J.: What's the buzz. Logistics Management 36(2), 43–47 (1997)
47. Rai, A., Borah, S., Ramaprasad, A.: Critical success factors for strategic alliances in the information technology industry: an empirical study. Decision Sciences 27(1), 141–155 (1996)
48. Rejeb, A., Rejeb, K., Simske, S.J., Treiblmaier, H.: Drones for supply chain management and logistics: a review and research agenda. International Journal of Logistics Research and Applications pp. 1–24 (2021)
49. Ryals, L.J., Humphries, A.S.: Managing key business-to-business relationships: what marketing can learn from supply chain management. Journal of Service research 9(4), 312–326 (2007)
50. Salvador, F., Villena, V.H.: Supplier integration and npd outcomes: Conditional moderation effects of modular design competence. Journal of Supply Chain Management 49(1), 87–113 (2013)
51. Sarmah, S.P., Acharya, D., Goyal, S.: Coordination and profit sharing between a manufacturer and a buyer with target profit under credit option. European Journal of Operational Research 182(3), 1469–1478 (2007)

52. Shepherd, C., Günter, H.: Measuring supply chain performance: current research and future directions. Behavioral operations in planning and scheduling pp. 105–121 (2010)
53. Stefanović, N., Stefanović, D.: Supply chain performance measurement system based on scorecards and web portals. Computer Science and Information Systems 8(1), 167–192 (2011)
54. Su, Q., Song, Y.t., Li, Z., Dang, J.x.: The impact of supply chain relationship quality on cooperative strategy. Journal of Purchasing and Supply Management 14(4), 263–272 (2008)
55. Taber, K.S.: The use of cronbach's alpha when developing and reporting research instruments in science education. Research in science education 48(6), 1273–1296 (2018)
56. Tan, K.C.: A framework of supply chain management literature. European Journal of Purchasing & Supply Management 7(1), 39–48 (2001)
57. Teo, T.S., Tan, M., Buk, W.K.: A contingency model of internet adoption in singapore. International Journal of electronic commerce 2(2), 95–118 (1997)
58. Tyndal, G., Gopal, C., Partsch, W., Kamauff, J.: Making it happen: the value producing supply chain. Ernst & Young, available at: www. ey. com/global/gcr. nsf/US/Supercharging_Supply_Chains_-_Think_Tank_-_ Ernst_% 26_Young _LLP (accessed 10 January 2001) (2000)
59. Un, C.A., Asakawa, K.: Types of r&d collaborations and process innovation: The benefit of collaborating upstream in the knowledge chain. Journal of Product Innovation Management 32(1), 138–153 (2015)
60. Uzsoy, R., Fowler, J.W., Mönch, L.: A survey of semiconductor supply chain models part ii: demand planning, inventory management, and capacity planning. International Journal of Production Research 56(13), 4546–4564 (2018)
61. Waller, M.A., Fawcett, S.E.: Data science, predictive analytics, and big data: a revolution that will transform supply chain design and management (2013)
62. Wankmüller, C., Reiner, G.: Coordination, cooperation and collaboration in relief supply chain management. Journal of Business Economics 90(2), 239–276 (2020)
63. White, R.E., Hamermesh, R.C.: Toward a model of business unit performance: An integrative approach. Academy of Management Review 6(2), 213–223 (1981)
64. Williamson, O.E.: Assessing contract. Journal of Law, Economics, & Organization 1(1), 177–208 (1985)
65. Wisner, J.D.: A structural equation model of supply chain management strategies and firm performance. Journal of Business logistics 24(1), 1–26 (2003)
66. Yip, G., McKern, B.: China's many types of innovation. Forbes, Sept 19, 2014 (2014)

Yoonkyo Cho received the Doctorate (Ph.D.) in Management from The State University of New York at Buffalo, USA in 2017. She is currently working as an assistant professor at Halla University in Wonju, Korea. Her research interests are in the fields of strategic management, international business, and entrepreneurship.

Chunsu Lee received the Doctorate (Ph.D.) in International Business Management from the Korea University at Seoul, Korea in 2006. He is currently working as a professor at Pukyong National Unversity in Pusan, Korea. His research interests are in the fields of international business, international marketing and strategic management.

Received: May 14, 2022; Accepted: December 26, 2022.

Computer Science and Information Systems 21(2):473–490 https://doi.org/10.2298/CSIS220826008L

Navigation Control of an Autonomous Ackerman Robot in Unknown Environments by Using a Lidar-Sensing-Based Fuzzy Controller

Cheng-Jian Lin[1], Jyun-Yu Jhang[2,⋆] and Chen-Chia Chuang[1]

[1] Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung 411, Taiwan
cjlin@ncut.edu.tw
3a617090@gmail.com
[2] Department of Computer Science and Information Engineering, National Taichung University of Science and Technology, Taichung 404, Taiwan
jyjhang@nutc.edu.tw

Abstract. In this paper, a real-time navigation control system based on lidar sensing is proposed for use in unknown environments. The proposed system comprises a behavioral controller for controlling an autonomous Ackerman robot for obstacle avoidance in the absence of global map information when moving toward a goal. The adopted obstacle avoidance method is selected by a wall-following fuzzy controller. The input parameter of this controller is the distance between the robot and the wall, which is determined by the lidar sensor, and the output parameter of the controller is the steering angle of the robot for it to reach the destination without collision. To prevent the robot from entering an endless loop, an endless loop escape mechanism is added to the proposed system. The simulation and experimental results of this study indicate that the proposed navigation control system can effectively assist an Ackerman robot to complete the navigation task successfully in unknown environments.

Keywords: Ackerman robot, fuzzy logic controller, lidar, navigation system, unknown environment.

1. Introduction

Autonomous mobile robots is key in the trend toward automation, due to labor shortages, in factories. However, autonomy is difficult to achieve in these robots because of unknown environments and uncertain dynamic obstacles[12], as evident in applications such as self-driving cars [25] and large object manipulation [28],[17]. The navigation control of autonomous mobile robots involves the two steps of goal finding and obstacle avoidance, where are performed using a robust controller [26]. For unknown environments, autonomous robots must perceive environmental information and control the angle and speed of robot movement to reach the destination and automatically avoid obstacles [2].

Many methods have been proposed to solve problems related to robot navigation control; these methods include artificial potential field [11], vector field histogram [5], behavior-based [21], and fuzzy logic [7] methods. Behavior-based methods are widely

⋆ Corresponding author

used in the navigation of autonomous mobile robots [10], and these methods can handle various situations without a global map. In behavior-based methods, autonomous mobile robots engage in wall-following behavior to explore an unknown environment. These robots can move by following the contour and distance information of an object to avoid obstacles and move toward the destination [24],[15]. To control a robot efficiently and stably, fuzzy logic control (FLC) has been incorporated into robot navigation controllers.

Fuzzy theory is used to express the knowledge and experience of experts in the form of language rules to construct a knowledge base and handle uncertain situations [1]. Fuzzy control systems have been used in many domains, including control engineering, signal processing, information processing, and machine intelligence technology [13],[14],[3], [22],[27]. In addition, FLC has proven to be a successful control method for many complex nonlinear systems and has replaced traditional control methods [9]. Mamdani and Assilian [18],[19] designed a fuzzy controller system for controlling a small steam engine. Their experimental results indicate that a fuzzy controller system can achieve better control performance than can a classical controller.

Autonomous mobile robots mostly rely on sensors to measure their relative distance from objects in the environment [6] for perceiving an unknown environment, analyzing and processing environmental information, and making relevant movement decisions. The sensors commonly used in autonomous mobile robots include infrared cameras, sonar, radar, and ultrasonic sensors. However, in a real environment, noise affects the signal captured by a sensor and might lead to wrong decisions. In contrast to the aforementioned sensors, lidar sensors can measure the distance between objects with high precision, identify the shapes of objects, and construct a three-dimensional geographic information model of the surrounding area without being affected by the weather. In the present study, a lidar sensor was adopted to obtain accurate environmental information.

The mobile chassis of autonomous mobile robots are mostly designed with a two-wheel differential structure or omnidirectional wheel structure. The radius and speed of a two-wheel differential structure during steering are determined by the speeds of the two wheels, which can enable circular objects, such as wheels, to be turned on the spot. This structure has relatively strong flexibility but low control precision. The omnidirectional wheel structure can realize omnidirectional walking without changing the body posture [20]. This structure results in very smooth movement but cannot be used in uneven environments. Compared with the aforementioned structures, the Ackerman chassis architecture has higher control precision and smoother movement. Moreover, this architecture allows the robot to move freely in different types of terrain. When the Ackerman architecture is turning, each wheel rotates around the same center; thus, this architecture is not prone to slippage and tire position misalignment [4].

In this paper, a navigation control method is proposed for an autonomous Ackerman robots in unknown environments. The proposed system comprises a behavior controller for controlling an Ackerman robot to achieve obstacle avoidance when heading toward the destination in the absence of global map information. To achieve obstacle avoidance, a wall-following fuzzy controller (WFFC) is used. Furthermore, an escape mechanism is used to prevent the robot from entering an endless loop. Experimental results indicate that the proposed navigation method can complete the navigation task in simulated and real environments. The remainder of this paper is structured as follows. Section 2 illustrates related work. Section 3 introduces the proposed navigation method. Section 4 presents the

experimental results obtained in a simulated environment and real environment. Finally, section 5 concludes this study.

2. Related Work

In recent years, the development of autonomous Ackermann robot controller has gained significant attention. This section provides an overview of the related work and advancements in this field.

• Lidar-based Perception and Mapping: Lidar sensors are widely used in autonomous robotics for environment perception and mapping. Researchers have explored the integration of Lidar sensors with Ackermann steering robots to enable accurate and real-time perception of the surroundings. Through Lidar data, the generated 3D environment map was then used for localization and obstacle detection, facilitating autonomous navigation of the Ackermann robot [23].

• Fuzzy Logic Control for Autonomous Navigation: Fuzzy logic controllers have been applied to achieve autonomous navigation in various robotic systems. When combined with Lidar sensing, these controllers can effectively handle uncertainties and variations in the environment. The controller utilized fuzzy rules to interpret Lidar data and generate steering and speed commands, enabling safe and efficient navigation in dynamic environments [16].

• Obstacle Avoidance and Collision Detection: Autonomous Ackermann robots require practical obstacle avoidance and collision detection capabilities to ensure safe navigation. Researchers have developed a fuzzy logic-based collision avoidance system for Ackermann steering robots. By analyzing Lidar data, their system made real-time decisions to avoid obstacles and maintain a safe distance during navigation [8].

In summary, the development of autonomous Ackermann robot controller has seen significant progress. Researchers have focused on perception, control, obstacle avoidance, path planning, and real-world applications. The integration of Lidar sensors with fuzzy control systems provides a powerful approach to achieving autonomous navigation in diverse environments.

3. Navigation Control System for the Autonomous Ackerman Robot

This section introduces the proposed navigation control system for an autonomous Ackerman robot. The proposed navigation system includes a behavior controller that makes an Ackerman robot move toward the destination while avoiding obstacles. The flowchart of this navigation system is shown in Fig. 1. When the behavior controller does not detect any obstacle, it instructs the robot to move toward the destination. By contrast, if this controller detects an obstacle, it switches to the wall-following mode for the robot to avoid the obstacle. However, in the wall-following mode, the robot might encounter an endless terrain loop, which makes the robot unable to successfully complete the navigation task. Therefore, an endless loop escape mechanism is designed to assist the robot to escape an endless loop terrain. Navigation control is completed when the autonomous Ackerman robot reaches its destination.

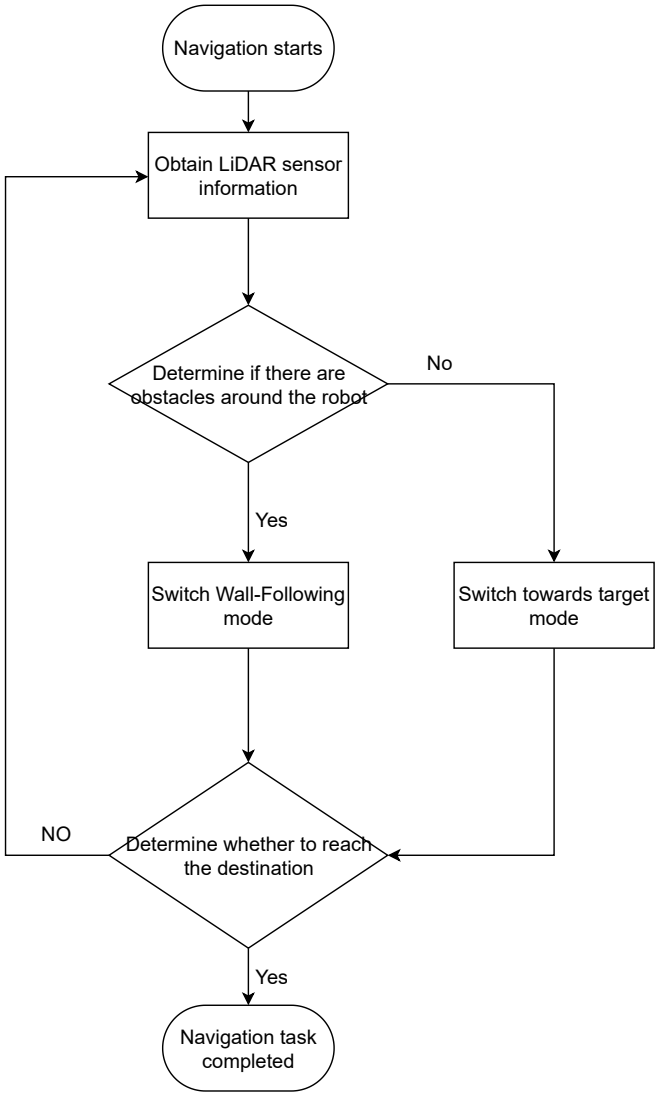


Fig. 1. Flowchart of the proposed navigation system

3.1. Autonomous Ackerman Robot

We independently developed the autonomous Ackerman robot used in this study. The adopted robot uses a Velodyne Puck (VLP-16) lidar sensor to scan for surrounding obstacles and an edge-embedded device [NVIDIA Jetson AGX Xavier (AGX)] to conduct real-time data processing. The sensing range of VLP-16 is 0.5 to 5 m, and its horizontal angular measurement range is 360°. AGX uses Ubuntu 16.04 and the Robot Operating System (ROS) to drive the robot's motor system. Through control commands, the movement speed and turning angle of the robot are controlled. In addition, the robot chassis has the Ackerman architecture for it to move smoothly when handling heavy objects and uneven terrain. The designed autonomous Ackerman robot is shown in Fig. 2.

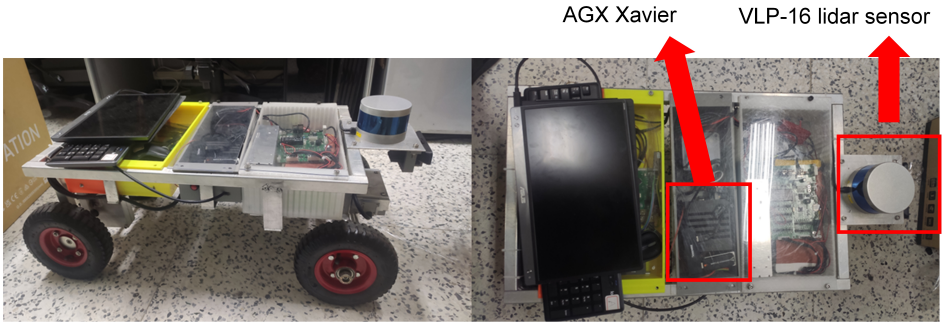


Fig. 2. Designed autonomous Ackerman robot

3.2. Behavior Controller

The behavioral controller plays a decision-making role in the proposed navigation system. This controller switches between the toward-goal mode and wall-following mode depending on the robot position and environment. To detect the position of an obstacle, the sensing area of the lidar sensor is divided into three areas, denoted A1, A2 and A3 (Fig. 3). A1 is the front-right area of the robot, A2 is the front-left area of the robot, and A3 is the rear of the robot. When an obstacle is detected in A1 or A2, the behavior controller switches from the toward-goal mode to the wall-following mode. The behavior controller remains in the toward-goal mode as long as the lidar sensor does not detect any obstacles in A1 and A2.

A. Toward-Goal Mode

When no obstacle is detected in front of the autonomous Ackerman robot, the robot moves toward the goal. As displayed in Fig. 4., the autonomous Ackerman robot calculates the steering angle according to its current position and the goal position, then turns toward the goal position, and then moves straight toward the goal. The designed Ackerman architecture has a turning angle between 45° and −45°; thus, the maximum angle of left and right turns is 45°.

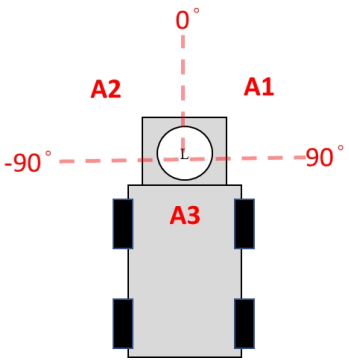


Fig. 3. Three obstacle detection areas for the autonomous Ackerman robot

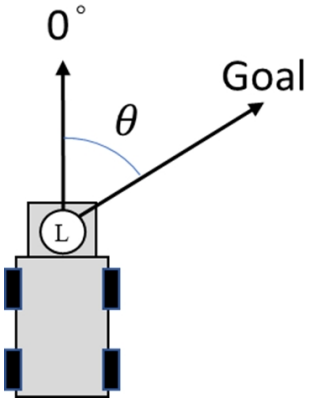


Fig. 4. Angle between the autonomous Ackerman robot and the goal

B. Wall-Following Mode

If an obstacle is detected in front of the autonomous Ackerman robot, the behavior controller switches to the wall-following mode to instruct the robot to move along the object until the object has been passed. To achieve this behavior, a fuzzy controller with a wall-following function, namely a WFFC, is designed. Fig. 5 displays the system flow of the wall-following mode. First, the lidar sensor detects the distance to obstacles around the robot. Subsequently, the distance information is used as the input of the controller to obtain the steering angle of the robot as the output. The proposed WFFC contains four parts: a fuzzifier, fuzzy rule base, fuzzy inference engine, and defuzzifer. The basic structure of the proposed WFFC is displayed in Fig. 6.

The various parts of the WFFC are detailed in the following text.

• Fuzzifier

A fuzzifier maps a crisp value to a fuzzy number (i.e., a real number between 0 and 1). This process is called fuzzification, and fuzzy logic better accords with human cognition

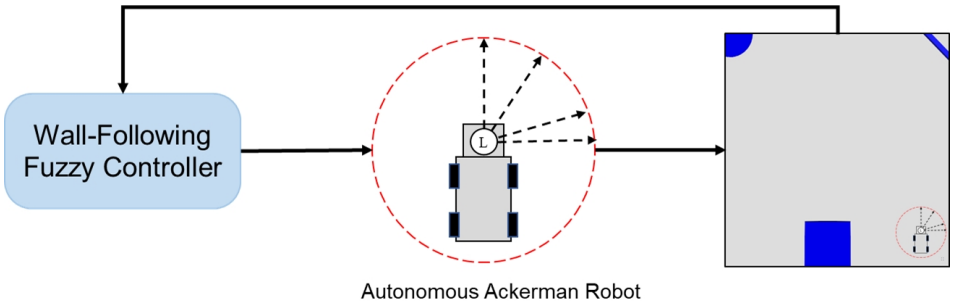


Fig. 5. System flow of the wall-following mode

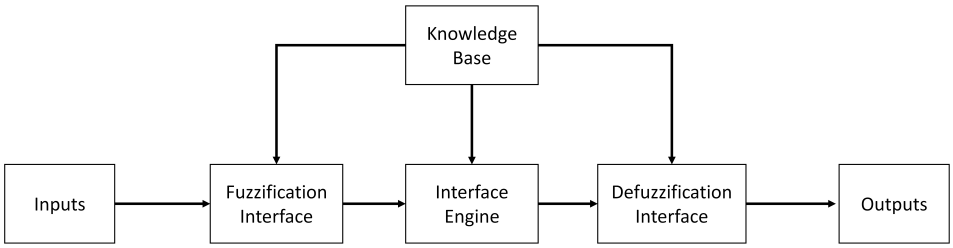


Fig. 6. Architecture of the proposed WFFC

relative to classical (bivalent) logic. Membership functions are used to evaluate the degree of each input of a system. Triangular or trapezoidal membership functions are most commonly used membership functions in fuzzy systems (Figs. 7 and 8). These membership functions are constructed using straight lines. Please review whether the edits convey your intended meaning accurately.Compared with Gaussian membership functions, linear membership functions are simpler and thus enable the design of simpler and more computationally lightweight robot controllers. Therefore, triangular and trapezoidal membership functions were used to design the robot controller in this study. The membership function of fuzzy set A can be defined as $\mu_A(x)$, where $\mu_A(x)$ denotes the degree of input x from fuzzy set A. A triangular membership function contains three parameters, namely a_1, a_2, and b_1, which denote the positions of the left boundary, Please specify which vertex of the triangle is being referred to here.triangle vertex, right boundary, respectively. The definition of a triangular membership function is provided in Equation 1. Trapezoidal membership functions contain four parameters, namely a_1, a_2, b_1, and b_2, which represent the positions of the left boundary, right boundary, left triangle vertex, and right triangle vertex, respectively. The definition of trapezoidal membership is provided in Equation 2.

$$\mu_A(x) = \begin{cases} 0 & x \leq a_1 \\ \frac{x-a_1}{a_2-a_1} & a_1 < x < a_2 \\ \frac{b_1-x}{b_1-a_2} & a_2 < x < b_1 \\ 0 & b_1 \leq x \end{cases} \tag{1}$$

$$\mu_A(x) = \begin{cases} 0 & x \le a_1 \\ \frac{x-a_1}{a_2-a_1} & a_1 < x < a_2 \\ 1 & a_2 < x < b_1 \\ \frac{b_1-x}{b_1-a_2} & b_1 < x < b_2 \\ 0 & b_2 \le x \end{cases} \tag{2}$$

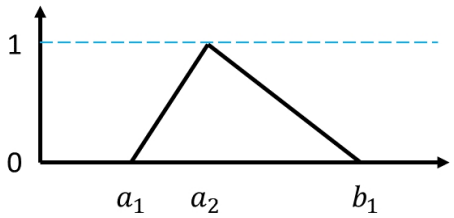


Fig. 7. Triangular membership function

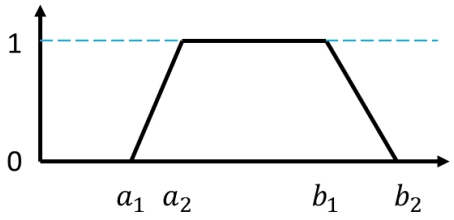


Fig. 8. Trapezoidal membership function

The scanning angle range of a lidar sensor is 360°, and A1 covers an angular range of 90°. Therefore, if lidar information is acquired for each degree in A1, 90 input data are obtained. To reduce the quantity of input data, only the lidar sensing data at 0°, 30°, 75°, and 90° are used as input (denoted as L1, L2, L3, and L4, respectively). Fig. 9 shows the four directions sensed by the lidar in A1.

Three membership functions can be used to define the distance of objects sensed by the lidar sensor as near, normal, or far. The detection angle of L1 is close to that of L2, and the detection angle of L3 is close to that of L4. Therefore, the same membership function is used for the forward (L1) and obliquely forward (L2) directions, and the same membership function is used for the forward-right (L3) and right (L4) directions. The membership functions of lidar for forward and rightward sensing are displayed in Figs. 10 and 11, respectively. The Ackerman architecture requires a large radius of gyration when turning. Therefore, in the membership function for forward sensing (Fig. 10), a sensing distance of greater than 4.25 m indicates that the obstacle is located far away from the

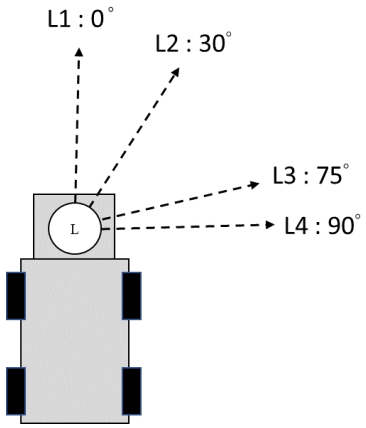


Fig. 9. Four directions sensed by the lidar sensor

robot. By contrast, a sensing distance of less than 3 m indicates that the object is close to the robot. In the membership function for rightward sensing (Fig. 11), a sensing distance of less than 2 m indicates that the robot is close to the obstacle. Moreover, a sensing distance of greater than 2 m indicates that the robot is located far from the obstacle.

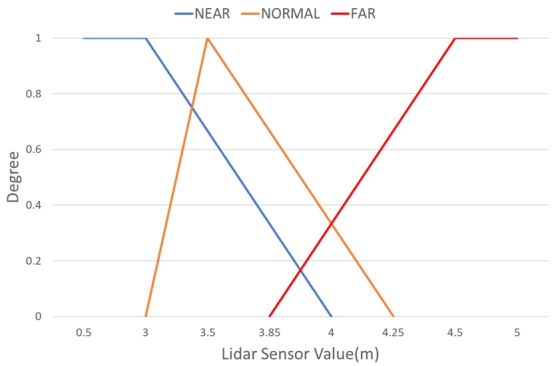


Fig. 10. Membership functions for forward sensing by the lidar sensor (L1 and L2)

• Fuzzy Rule Base and Fuzzy Inference Engine

If–then rule statements are adopted to construct a fuzzy rule base. A fuzzy rule can be defined as follows:

R_i: If x_1 is A_1 and x_2 is $A_2 \ldots$ and x_n is A_n, then y is b_i,

where x and y are linguistic variables. Because the robot might encounter different terrains during its movement, appropriate fuzzy rules must be designed for different conditions. Fig. 12 illustrates the obstacles located ahead of the robot at an acute angle, a right

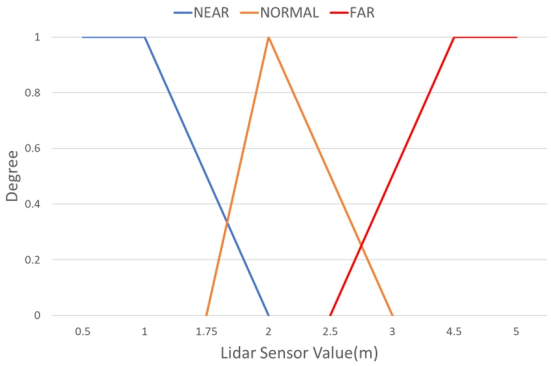


Fig. 11. Membership functions for rightward sensing by the lidar sensor (L3 and L4)

angle, and an obtuse angle. According to the distances sensed by L1 and L2, the state of the obstacle can be obtained to determine the turning range. Fig. 13 depicts the state of the robot and obstacles. In Fig. 13, three robot–obstacle conditions are observed: robot is parallel to the obstacle, robot is located close to the obstacle, and robot is located far away from the obstacle. On the basis of the values sensed by L3 and L4, the angle of the robot body can be adjusted.

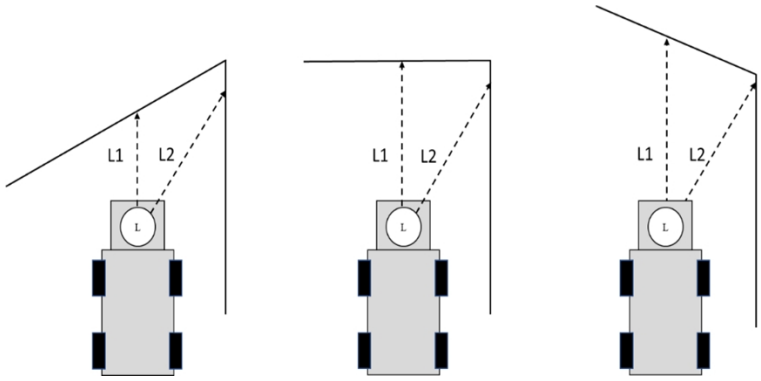


Fig. 12. Obstacles located ahead of the robot at three angles

On the basis of the aforementioned conditions, 21 fuzzy rules were designed (Table 1). The inputs of the proposed wall-following controller are four distance variables, namely those sensed by L1–L4, and the output is the Ackerman turning angle (θ), which is between 45° and −45°. This angle is mapped to a value between 1 and −1 (ω) by using Equation 3. In this study, the AND operation is used for fuzzy rule computation.

$$\omega = \frac{\theta}{45^\circ} \tag{3}$$

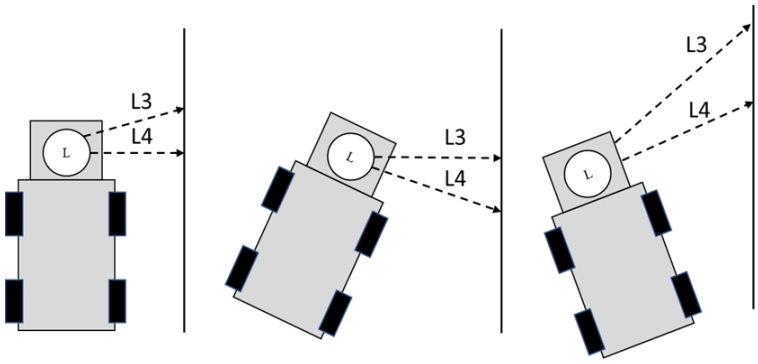


Fig. 13. State of the robot and obstacles

Table 1. Twenty-one fuzzy rules of the proposed wall-following controller

Intput				Output
L1	L2	L3	L4	ω
ANY	ANY	ANY	NEAR	0.7
ANY	ANY	ANY	NORMAL	0
ANY	ANY	ANY	FAR	-0.7
NEAR	NEAR	ANY	ANY	0.7
NEAR	NORMAL	ANY	ANY	0.7
NEAR	FAR	ANY	ANY	0.7
NORMAL	NEAR	ANY	ANY	0.7
NORMAL	NORMAL	ANY	ANY	0.6
NORMAL	FAR	ANY	ANY	0.6
FAR	NEAR	ANY	ANY	0.5
FAR	NEAR	ANY	ANY	0
FAR	FAR	ANY	ANY	-0.3
ANY	ANY	NEAR	NEAR	0.5
ANY	ANY	NEAR	NORMAL	0.3
ANY	ANY	NEAR	FAR	0.3
ANY	ANY	NORMAL	NEAR	0.1
ANY	ANY	NORMAL	NORMAL	0
ANY	ANY	NORMAL	FAR	-0.1
ANY	ANY	FAR	NEAR	0.1
ANY	ANY	FAR	NORMAL	-0.2
ANY	ANY	FAR	FAR	-0.6

• **Defuzzifier**

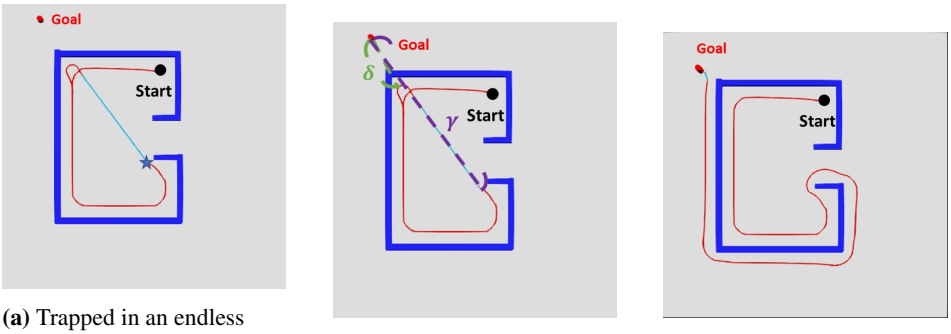
The output of a defuzzifier is a crisp value. The centroid defuzzification process can be expressed as follows:

$$y = \frac{\sum_{i=1}^{21} \mu_A(x_i)\omega_i}{\sum_{i=1}^{21} \mu_A(x_i)} \tag{4}$$

where y represents the output of the wall-following controller, $\mu_A(x_i)$ is the firing strength of the ith rule, and ω is the fuzzy value of the robot turning angle (between -1 and 1).

3.3. Endless Loop Escape Mechanism

To prevent the autonomous Ackerman robot from falling into an endless loop, an endless loop escape mechanism is proposed (Fig. 14). In Fig. 14(a), the robot faces no obstacle when it moves toward the star point along the wall. At this time, the behavior controller switches to the toward-goal mode and cause the robot to fall into an endless loop. In the proposed endless loop escape mechanism, the shortest distance (δ) between the robot and the goal is recorded to determine which behavioral mode should be employed. When the robot moves to the star point, the current distance from the robot to the goal (γ) is greater than δ, and the behavior controller executes the wall-following mode. Until δ is greater than γ, the behavior controller functions in the toward-goal mode to escape the endless loop, as displayed in Fig 14(c).



(a) Trapped in an endless loop

(b) Escape mechanism

(c) Escaping the endless loop

Fig. 14. Proposed endless loop escape mechanism

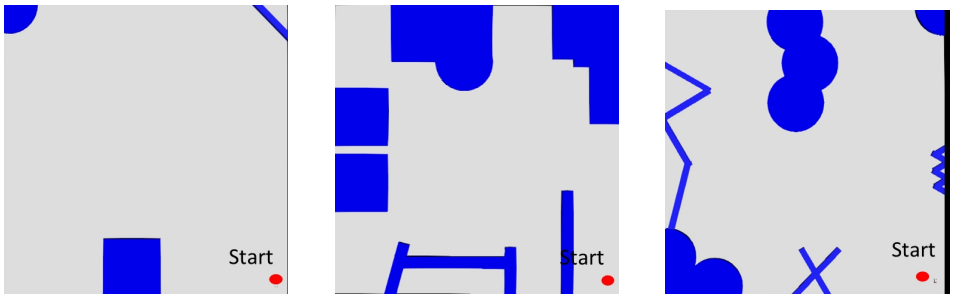
4. Simulation and Experimental Results

To verify the performance of the proposed navigation method, an open-source robot simulator (Webots) was used to construct a testing environment. Webots is a development

environment software for the modeling, programming, and simulation of mobile robots, and this software program can run on Linux, Windows, and macOS. The proposed robot controller can be programmed in C, C++, Python, Java, MATLAB, or the ROS by using a simple application programming interface that covers all basic robot control techniques. Simulation experiments were conducted in six environments to test the effectiveness of the proposed method, and the size of each simulated environment was 40 m 40 m. Finally, the proposed method was used to complete a navigation task with an autonomous Ackerman robot in a real environment.

4.1. Simulation Results Obtained for the WFFC

We designed three environments with small numbers of square and circular obstacles to test the performance of the proposed WFFC (Fig. 15). The first environment consisted of simple circular objects, hypotenuses, and square obstacles. The second environment was mainly composed of square obstacles to test the robot's obstacle avoidance performance at right angles. Finally, the third environment was composed of circular obstacles and special concave corners to test whether the robot could effectively avoid concave corners.



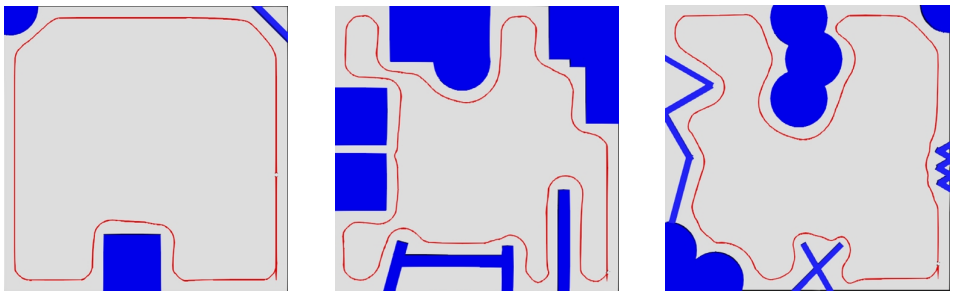
(a) Few obstacles **(b)** Square obstacles **(c)** Circular obstacles

Fig. 15. Three testing environments for the proposed WFFC

Fig. 16 illustrates the paths of the autonomous Ackerman robot in the aforementioned three testing environments when using the proposed WFFC. The robot successfully circumnavigated the three environments without collision.

4.2. Simulation Results for Navigation Control

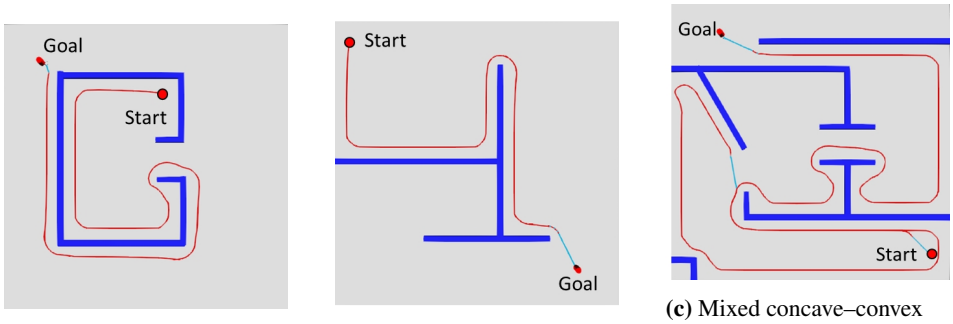
Three testing environments, namely a concave environment, convex environment, and mixed concave–convex environment, were designed to evaluate the navigation control performance achieved with the proposed method. These environments are displayed in Fig. 17. The red movement path in Fig. 17 represents the behavior controller executing the wall-following mode, whereas the blue movement path denotes this controller executing the toward-goal mode. As displayed in Fig. 17(b), when the navigation started, the behavior controller executed the toward-goal mode because no obstacles were detected



(a) Environment 1 **(b)** Environment 2 **(c)** Environment 3

Fig. 16. Movement paths of the autonomous Ackerman robot in the three testing environments designed for the proposed WFFC

ahead of the robot (blue path). When an obstacle was detected, the controller switched to the wall-following mode and moved the robot along the red path. Fig. 17(c) indicates that the proposed endless loop escape mechanism effectively assisted the robot to escape an endless loop terrain and complete the navigation task in an unknown environment.



(a) Concave environment **(b)** Convex environment **(c)** Mixed concave–convex environment

Fig. 17. Navigation paths of the autonomous Ackerman robot in three testing environments

4.3. Experimental Results in a Real Environment

The proposed navigation control method was implemented for an autonomous Ackerman robot in a real environment. Fig. 18 shows the floor plan of the real testing environment. Several square and circular obstacles were included in the real environment to verify the effectiveness of the proposed navigation control method. As displayed in Fig 19, the autonomous Ackerman robot sensed obstacles in detection area A1. Therefore, the behavior controller executed the wall-following mode. Fig. 20 displays the computation time for each time step. As displayed in Fig. 20, the computation time for each step was between

0.275 and 0.295 s. Thus, the proposed navigation control method can realize real-time computation. The results also indicate that the proposed method can be effectively applied in unknown environments without the need for complex global map construction and model training.

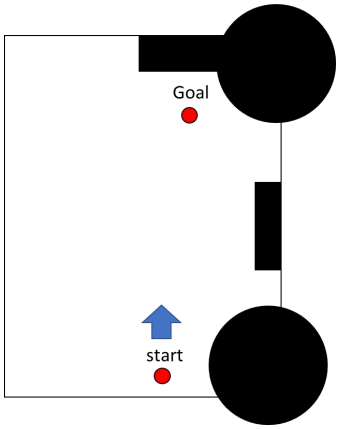


Fig. 18. Floor plan of the real testing environment

Fig. 19. Navigation control in the real environment

5. Conclusion

In this paper, an effective navigation control method is proposed for autonomous Ackerman robots moving in unknown environments. The proposed method can accomplish

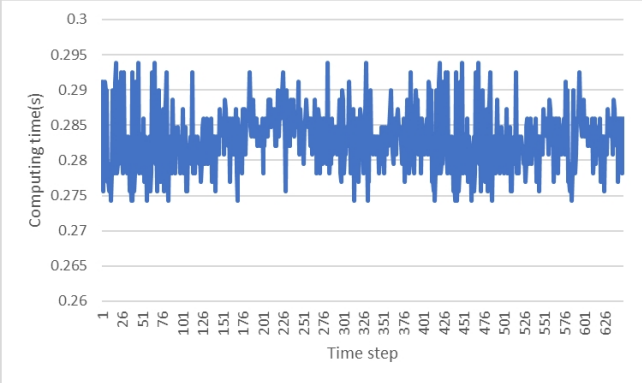


Fig. 20. Computation time for each time step

the navigation task without the construction of a global map or the training of a complex model. The designed behavior controller enables an autonomous Ackerman robot to undertake obstacle avoidance and complete the navigation task automatically according to the current environment state. Furthermore, the computation time per time step of the proposed method is less than 0.3 s, which indicates that the proposed method has real-time computation capability. Simulation and experimental results indicated that the proposed navigation control method can enable an autonomous Ackerman robot to complete the navigation task effectively without collision in an unknown environment. In a future study, we will consider applying the developed autonomous Ackerman robot to practical applications.

Acknowledgments. The authors would like to thank the National Science and Technology Council of the Republic of China, Taiwan for financially supporting this research under Contract No. NSTC 111-2222-E-025-001.

References

1. Abdelazim, T., Malik, O.: An adaptive power system stabilizer using on-line self-learning fuzzy systems. In: 2003 IEEE Power Engineering Society General Meeting (IEEE Cat. No.03CH37491). vol. 3, pp. 1715–1720 Vol. 3 (2003)
2. Bao, Q.y., Li, S.m., Shang, W.y., An, M.j.: A fuzzy behavior-based architecture for mobile robot navigation in unknown environments. In: 2009 International Conference on Artificial Intelligence and Computational Intelligence. vol. 2, pp. 257–261 (2009)
3. Batur, C., Kasparian, V.: Model based fuzzy control. Mathematical and Computer Modelling 15(12), 3–14 (1991)
4. Carpio, R.F., Potena, C., Maiolini, J., Ulivi, G., Rosselló, N.B., Garone, E., Gasparri, A.: A navigation architecture for ackermann vehicles in precision farming. IEEE Robotics and Automation Letters 5(2), 1103–1110 (2020)
5. Chen, W., Wang, N., Liu, X., Yang, C.: Vfh based local path planning for mobile robot. In: 2019 2nd China Symposium on Cognitive Computing and Hybrid Intelligence (CCHI). pp. 18–23 (2019)

6. Discant, A., Rogozan, A., Rusu, C., Bensrhair, A.: Sensors for obstacle detection - a survey. In: 2007 30th International Spring Seminar on Electronics Technology (ISSE). pp. 100–105 (2007)
7. Doitsidis, L., Valavanis, K., Tsourveloudis, N.: Fuzzy logic based autonomous skid steering vehicle navigation. In: Proceedings 2002 IEEE International Conference on Robotics and Automation (Cat. No.02CH37292). vol. 2, pp. 2171–2177 vol.2 (2002)
8. Elsayed, H., Abdullah, B.A., Aly, G.: Fuzzy logic based collision avoidance system for autonomous navigation vehicle pp. 469–474 (2018)
9. Feng, G.: A survey on analysis and design of model-based fuzzy control systems. IEEE Transactions on Fuzzy Systems 14(5), 676–697 (2006)
10. Ganapathy, V., Yun, S.C., Ng, J.: Fuzzy and neural controllers for acute obstacle avoidance in mobile robot navigation. In: 2009 IEEE/ASME International Conference on Advanced Intelligent Mechatronics. pp. 1236–1241 (2009)
11. Hwang, Y., Ahuja, N.: Path planning using a potential field representation. In: Proceedings. 1988 IEEE International Conference on Robotics and Automation. pp. 648–649 vol.1 (1988)
12. Kan, X., Teng, H., Karydis, K.: Online exploration and coverage planning in unknown obstacle-cluttered environments. IEEE Robotics and Automation Letters 5(4), 5969–5976 (2020)
13. Lee, C.: Fuzzy logic in control systems: fuzzy logic controller. i. IEEE Transactions on Systems, Man, and Cybernetics 20(2), 404–418 (1990)
14. Lee, C.: Fuzzy logic in control systems: fuzzy logic controller. ii. IEEE Transactions on Systems, Man, and Cybernetics 20(2), 419–435 (1990)
15. Li, Y.J., Chou, W.C., Chen, C.Y., Shih, B.Y., Chen, L.T., Chung, P.Y.: The development on obstacle avoidance design for a humanoid robot based on four ultrasonic sensors for the learning behavior and performance. In: 2010 IEEE International Conference on Industrial Engineering and Engineering Management. pp. 376–379 (2010)
16. Lin, C.J., Chang, M.Y., Tang, K.H., Huang, C.K.: Navigation control of ackermann steering robot using fuzzy logic controller. Sensors and Materials 35, 781 (2023)
17. Liu, O., Yuan, S., Li, Z.: A survey on sensor technologies for unmanned ground vehicles. In: 2020 3rd International Conference on Unmanned Systems (ICUS). pp. 638–645 (2020)
18. Mamdani, E.: Application of fuzzy algorithms for control of simple dynamic plant. Proceedings of the Institution of Electrical Engineers 121, 1585–1588(3) (December 1974)
19. Mamdani, E., Assilian, S.: An experiment in linguistic synthesis with a fuzzy logic controller. International Journal of Man-Machine Studies 7(1), 1–13 (1975)
20. Morales, S., Magallanes, J., Delgado, C., Canahuire, R.: Lqr trajectory tracking control of an omnidirectional wheeled mobile robot. In: 2018 IEEE 2nd Colombian Conference on Robotics and Automation (CCRA). pp. 1–5 (2018)
21. Motlagh, O.R.E., Hong, T.S., Ismail, N.: Development of a new minimum avoidance system for a behavior-based mobile robot. Fuzzy Sets and Systems 160(13), 1929–1946 (2009), theme: Information Processing and Applications
22. Pedrycz, W.: Fuzzy Control and Fuzzy Systems (2nd, Extended Ed.). Research Studies Press Ltd., GBR (1993)
23. Pendleton, S.D., Andersen, H., Du, X., Shen, X., Meghjani, M., Eng, Y.H., Rus, D., Ang, M.H.: Perception, planning, control, and coordination for autonomous vehicles. Machines 5(1) (2017)
24. Peng, J., Yumei, H.: Behavior-based avoiding barriers system of mobile robot. In: 2009 WRI World Congress on Software Engineering. vol. 3, pp. 106–112 (2009)
25. Said, H.B., Marie, R., Stéphant, J., Labbani-Igbida, O.: Skeleton-based visual servoing in unknown environments. IEEE/ASME Transactions on Mechatronics 23(6), 2750–2761 (2018)
26. Shayestegan, M., Din, S.: Fuzzy logic controller for robot navigation in an unknown environment. In: 2013 IEEE International Conference on Control System, Computing and Engineering. pp. 69–73 (2013)
27. Sugeno, M.: On stability of fuzzy systems expressed by fuzzy rules with singleton consequents. IEEE Transactions on Fuzzy Systems 7(2), 201–224 (1999)

28. Tsuru, M., Escande, A., Tanguy, A., Chappellet, K., Harad, K.: Online object searching by a humanoid robot in an unknown environment. IEEE Robotics and Automation Letters 6(2), 2862–2869 (2021)

Cheng-Jian Lin received the B.S. degree in electrical engineering from the Ta Tung Institute of Technology, Taipei, Taiwan, in 1986, and the M.S. and Ph.D. degrees in electrical and control engineering from National Chiao-Tung University, Taiwan, in 1991 and 1996, respectively. Currently, he is a Chair Professor of the Computer Science and Information Engineering Department, National Chin-Yi University of Technology, Taichung, Taiwan, and the Dean of the Intelligence College, National Taichung University of Science and Technology, Taichung. His current research interests include machine learning, pattern recognition, intelligent control, image processing, intelligent manufacturing, and evolutionary robots.

Jyun-Yu Jhang received the B.S. and M.S. degrees from the Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung, Taiwan, in 2015, and the Ph.D. degree in electrical and control engineering from National Chiao-Tung University, Taiwan, in 2021. He is currently an Assistant Professor with the Computer Science and Information Engineering Department, National Taichung University of Science and Technology, Taichung. His current research interests include fuzzy logic theory, type-2 neural fuzzy systems, evolutionary computation, machine learning, and computer vision and application.

Chen-Chia Chuang received the B.S. and M.S. degrees from the Department of Computer Science and Information Engineering, National Chin-Yi University of Technology, Taichung, Taiwan, in 2023. His current research interests are machine learning, pattern recognition, and image processing.

Received: August 26, 2023; Accepted: July 06, 2023.

Computer Science and Information Systems 21(2):491–505 https://doi.org/10.2298/CSIS220830069C

A revised Girvan–Newman Clustering Algorithm for Cooperative Groups Detection in Programming Learning

Wen-Chih Chang

International Master Program in Information Technology and Applications,
National Pingtung University, Pingtung City 900, Taiwan
yilan.earnest@mail.nptu.edu.tw

Abstract. Learning to program is a challenging task for novices. Students vary substantially in their ability to understand complex and abstract topics in computer programming logic, such as loop logic, function recursion, arrays, passing parameters, and program structure design. Cooperative learning is an effective method of learning and teaching programming. In traditional cooperative learning, students group themselves, or teachers group students intuitively. This paper proposes a clustering method based on item response theory (IRT) and the revised Girvan–Newman clustering for clustering students by learning ability. Item response theory calculated the learner's ability and interpersonal relationship questionnaire generated by the social network analysis. The proposed method was validated by conducting a quasi-experimental test in a freshmen programming course, and the method significantly improved learning outcomes in this course.

Keywords: Learner ability, Girvan–Newman clustering, Social Network Analysis, Programming.

1. Introduction

Cooperative learning is a form of learning in which students learn and work together to accomplish shared goals. It has been applied in numerous fields. In most cases, cooperative learning is performed in small groups. Students in these groups discuss topics; through these discussions, all students learn and achieve beneficial outcomes. Cooperative learning can also be competitive; for example, groups might compete to see which group can answer the most questions in a limited time. Competitive group goals require all group members to work together to improve their learning. If the conditions in which competitive and purely cooperative learning should be applied are determined, a cooperative learning course can be designed for any subject.

Cooperative base groups are long-term, heterogeneous cooperative learning groups with stable membership [1]. Heterogeneous cooperative learning groups include students with different learning abilities. The term "stable membership" indicates that group members can work together over a long time or have good relationships. However, selecting people with good relationships in a class is challenging.

Girvan and Newman (2002) [2] proposed the Girvan–Newman clustering method for investigating communities. The authors test the method with computer-generated

communities and real-world community structures. The result showed high sensitivity and reliability.

There are some studies, which applied AI and metaverse methods to support education. Omonayajo, Fadi, and Nadire (2022) [3] examined the smart technologies that have assisted smart education in achieving educational goals. These smart technologies enhanced the teaching and learning process in today's education. Yu and Lin (2022) [4] explained the data mining status and the college students' psychological health problems. This research used the decision tree to analyze the psychological health problem data.

Innovation thinking and computational thinking affect students' learning, which promotes students learning performance. Dagienė, Jevsikova, Stupurienė, and Juškevičienė (2022) [5] surveyed 52 countries with a qualitative study of 15 countries, which helped them to identify teachers' understanding level of computational thinking and its integration approach in the class activities. It is useful for e-learning systems and content developers to improve teachers' computational thinking. In the other research Zheng et al. (2022) [6] made a training system, that made the major in computer science students have better academic performance and significantly improved compared with the performance before the innovative thinking.

Dale's Cone of Learning [7] model states that activities in which students experience, discuss, do, and participate cause greater retention than simply reading, watching, or hearing. In cooperative learning, students must be active participants in discussions and must support their team members. Thus, cooperative learning activities improve student learning, understanding, and retention.

A teacher can flexibly modify their lecturing style or learning material to maximize teaching quality based on student feedback. However, teachers typically prepare their teaching materials before classes begin. Thus, predicting student learning ability is key for preparing appropriate class activities. However, measuring learner ability is challenging for teachers. Therefore, a method that can be used to estimate learner ability and cluster students appropriately to obtain learning groups comprising heterogeneous members would be of considerable benefit to teachers and student outcomes. Assessments are typically used to measure and analyze student performance and learning skills. These assessments also can be used as feedback for teachers and students, which is crucial in learning and development.

The remainder of this paper is organized as follows. Section 2 describes item response theory (IRT) and the adopted clustering method. Section 3 presents details regarding how IRT and clustering are used to estimate student learning ability and identify cluster learners. The experimental results are presented in section 4, and section 5 provides the conclusions of this study and suggestions for future research.

2. Related Studies

With the increased acceptance of e-learning, numerous researchers have proposed various student assessment methods. For example, the researchers [8] designed a teaching for students to assess the smartphone to study Geography. With simple test items, the proposed system provides individual learning profile and test analysis report

for each student. The result shows an interesting approach and reveals the learning profile and test analysis for students is a good reply and suggestion for students. Some teachers applied the social network analysis clustering method [9] for cooperative learning in programming courses at the university. The relationship among all the classes is considered the connection between students. It shows significant differences in students' performance and scores. Some teachers used the combining flipped learning and online formative assessment platforms to enhance students' learning performance [10]. Research has increasingly focused on assessments to assist learning and teaching. IRT is often used to estimate learner ability. In IRT, the probability that a student answers a particular question (item) correctly is expressed using a continuously increasing graph called the item characteristic curve. The item characteristic curve is defined in terms of one, two, or three of the following parameters: item discrimination, item difficulty, and student guessing. Item discrimination refers to the extent to which an item discriminates between high- and low-ability students. Item difficulty indicates whether an item is easy or difficult, and student guessing can be included as a corrective factor if students are likely to guess the correct answer. Figure 1 presents the three-item characteristic curves of three items with the same discrimination of 1 and distinct difficulties of 1, 3, and 5.

The characteristic curve of each item in IRT is a logistic function that is expressed as follows.

In this function, e is Euler's number, b is the difficulty parameter (typically $-3 \leq b \leq 3$), a is the discrimination parameter (typically $-2.8 \leq a \leq 2.8$), $L = a(\Theta - b)$ is the logistic deviate (logic), and Θ indicates student ability level. A one-parameter item characteristic curve presents only the difficulty of the problem; the discrimination and guessing are ignored (set to 1). A one-parameter model (Equation 1) is expressed as follows:

$$P(\theta) = \frac{1}{1+e^{-a(\theta-b)}} \text{.....Equation 1}$$

The two-parameter logistic model considers the discrimination and item difficulty (Equation 2), and the three-parameter logistic model (Equation 3) considers the discrimination, item difficulty, and the probability that a guess is correct c. A three-parameter model is expressed as follows:

$$P(\theta) = \frac{1}{1+e^{-a(\theta-b)}} = \frac{1}{1+e^{-1(\theta-b)}} \text{.....Equation 2}$$

$$P(\theta) = c + (1-c)\frac{1}{1+e^{-a(\theta-b)}} \text{.....Equation 3}$$

The parameter c can theoretically range between 0 and 1; in practice, values greater than 0.35 are rarely used.

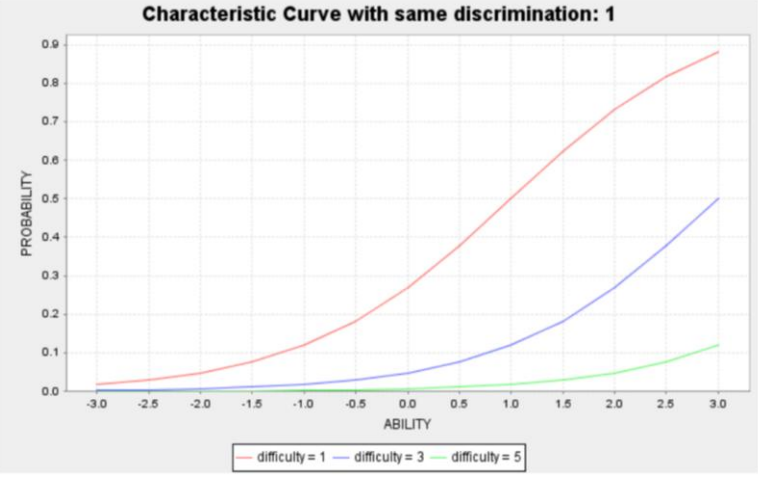


Fig. 1. Same discrimination of 1 with distinct difficulties in one-parameter models

Clustering is a method of organizing a collection of unlabeled data by grouping similar items. Clustering algorithms have been applied in biology, marketing, earthquake studies, and city planning. The K-means method is one of the most commonly used clustering techniques. This method is used to group a collection of data samples into *k* clusters based on a distance measurement. Distance is usually determined according to a data attribute, such as the price of a product, the score of a student, or the time and location of an earthquake. In this study, we clustered students in a learning community by using the Girvan–Newman (GN) community clustering algorithm. In social network analysis (SNA), social relationships between members of a social structure of any scale are considered to define nodes, ties, groups, and betweenness centrality [15]. In simple terms, SNA is a method of surveying all relationships among actors in a community [16]. Betweenness centrality indicates the extent to which a vertex or edge lies on a path between vertices. Nodes or vertices with high betweenness might have considerable influence on a network. Because of their presence on numerous paths, nodes or vertices can control considerable information flowing through a network.

3. Research Method

This paper proposes a methodology that combines K-means clustering with the GN community clustering algorithm, and the proposed methodology involves considering the distance (the betweenness value) between communities. Moreover, we propose a grouping algorithm combined with IRT for estimating learner ability to achieve heterogeneous groupings for cooperative learning.

3.1. Pretest

A learner's ability can be approximated by their test scores. However, the difficulty and discrimination of items differ; thus, students with the same score might still have different abilities. We applied the two-parameter logistic model based on IRT. As we mentioned in Section 2, the Two-parameter logistic model considers the discrimination and item difficulty (Equation 2). Using the discrimination and difficulty, we can get Θ which indicates the student's ability level.

We adopted Kelly's method to determine the item difficulty and discrimination indices. The best percentage for subsequent calculations was 27%, and acceptable percentages were 25%–33% [17]. We selected a percentage of 25% for these calculations. We then sorted students by their exam scores and defined the top and bottom 25% of students by test score. The total number of correct answers in the higher and lower groups for each question are denoted as PH and PL, respectively. The item difficulty index for each problem was calculated using the equation b = (PH + PL)/2, and the item discrimination index for each problem was calculated using the equation a = PH − PL. The default learner ability θ was set as 1. The parameters were input into the item characteristic equation to obtain P for item 1. For any student, P was calculated for the 20 items to calculate the student's learning ability.

3.2. Learner Clustering in Cooperative Programming Learning

Learner clustering is critical for cooperative programming learning. We revised a social network clustering method (GN iteration) (Figure 2), a heterogeneous function, and then used a grouping algorithm for clustering.

3.3. GN iteration [14]:

(1) Compute the betweenness of every edge in the graph. For node X, perform a breadth-first search to determine the number of shortest paths from node X to each node, and assign these numbers as scores to each node.
(2) Beginning at the leaf nodes, calculate the credit of an edge as [1 + (sum of the edge credits)] × (score of the destination node/score of the starting node).
(3) Compute the credits of all edges in the graph G and repeat from step 1 until all nodes have been selected.
(4) Sum all the credits computed in step 2 and divide by 2. The result is the betweenness of each edge.
(5) Remove the edges with the highest betweenness.
(6) Compute the modularity Q of the communities split.
(7) If Q > 0.3–0.7, repeat from step 1. (0.3-0.7 is the experimental result for better performance)

Heterogeneous function is used to make sure learner ability is distributed in different levels. We applied Equation 2, P(θ) is the learner ability. With the discrimination index

and difficulty index, the learner ability P(θ) can be calculated. Learner ability was classified as high, middle, and low. The most appropriate candidates were selected into teams according to the betweenness centrality and learner ability. Learner ability is calculated by item response theory.

```
IF(N >5)
        REPEAT
              FOR i=0 to n-1
                    LET B[i] BE betweenness centrality of edge i
                    IF B[i] > max_B
                    THEN max_B = B[i]
                              max_B_edge = i
                    ENDIF
              ENDFOR
               remove edge i from graph
         UNTIL number of edges in graph is 0
        //Divided into 2 groups
         Heterogeneous( );
ELSE IF ( 0< N && N<=5)
         Heterogeneous( );
```

*N is the number of nodes in the group graph, n is the number of edges in the group graph

Fig. 2. A revised GN algorithm

3.4. Quasi-Experimental Method and Posttest

This study referenced the research [9], which is designed based on a mixed approach. The difference part between the research [9] is the algorithm design and algorithm complexity comparison. This study also optimizes the Grouping algorithm. The study includes experimental and control groups. The experimental group has 34 male students and 10 female students. The control group has 38 male students and 6 female students. Two groups received the same teaching material and teaching progress in the semester. However, the clustering method in cooperative learning is different. The experimental group was clustered by social network analysis results, and the control group was clustered by the students they chose by themselves.

The experimental group of students was designed to answer two questions. The first question is "Who you will choose to be the team members?". The second question is "Who is the person you will ask or discuss when you encounter some problems in learning programming course?". Students can write 1~3 students' names. The study applied 1st question SNA clustering result and a little modified based on 2nd question answer to generate the cooperative learning team members.

The course taught variables, control commands, loop, pointer, array, function, recursion, and project. It took 18 weeks, including preparation, pretest (week 1~week 2), clustering of team members, posttest (week 18), answer questionnaire, and interview

procedures. The pretest is composed of five programming questions (such as int, double, calculate BMI, string decomposition, and if command operation).

T-test measures the difference between two means, which may or may not be related to each other. It also indicates the probability of the differences to have happened by chance. A T-test is usually a test for two experimental numbers, which has a difference between them. For example, the experimental result is better than the control result.

Paired Sample is the hypothesis testing conducted when two groups belong to the same group or population. In this experiment, P is a statistical measure that helps to determine whether the hypothesis is correct or not. Furthermore, it assists in demonstrating the significance of the results. In the experimental design, the null hypothesis is a default situation that which there is no relationship between two measured phenomena. H0 denotes the null hypothesis. The other hypothesis H1, is the researcher's belief that the null hypothesis is false. P-value is a number between 0 and 1. The significance level is a predefined threshold, which is set at 0.05 generally.

The assumption of statistics test is performed below:

Null Hypothesis: $\mu_d = 0$: There is no significance between our revised GN clustering algorithm and the students' willingness group.

Alternative Hypothesis: $H_1{:}\mu_d \neq 0$: There is significance between our revised GN clustering algorithm and the students' willingness group.

The pretest scores of the experimental and control groups were not significantly different (p = 0.804, Table 1). However, the posttest scores of the experimental group were significantly higher than their pretest scores (p = 0.0001, Table 2) and the posttest scores of the control group (p = 0.024, Table 3).

Table 1. Pretest scores [9]

Group	Average Score	Standard Deviation	t	p	Significance
Experimental Group	52.93	11.38	-0.248	0.804	No significance
Control Group	53.45	7.86			

Table 2. Pretest and post-test scores of the experimental and control groups [9]

Group	Test	Average Score	Standard Deviation	t	p	Significance
Experimental Group	Pretest	52.93	11.38	-3.796	0.0001***	No significance
	Posttest	63.72	16.94			
Control Group	Pretest	53.45	7.86	-0.737	0.465	
	Posttest	55.43	16.89			

*: p ≤ 0.05, **: p ≤ 0.01, ***: p ≤ 0.001

Table 3. Posttest scores of the experimental and control groups [9]

Group	Average Score	Standard Deviation	t	p	Significance
Experimental Group	63.72	16.94	2.298	0.024*	Significance
Control Group	53.43	16.89			

*: p ≤ 0.05, **: p ≤ 0.01, ***: p ≤ 0.001

The statistic test shows that there is significance. We reject the null hypothesis, or it means that the alternative hypothesis is accepted. The average mean between using our revised GN clustering algorithm and the students' willingness group is a significant difference of 0.52. Moreover, the standard deviation between the two groups is similar at 11.38 and 7.86. This implies that the learning performance in the pretest is quite the same, however, some students in our revised GN clustering algorithm can improve the average mean from 52.93 to 63.72. This concludes that using our revised GN clustering algorithm has an efficiency to apply in programming learning.

The final T-test interpretation could be obtained in either of the two ways:

A null hypothesis signifies that the difference between the means is zero and where both the means are shown as equal.

An alternate hypothesis implies the difference between the means is different from zero. This hypothesis rejects the null hypothesis, indicating that the data set is quite accurate and not by chance.

3.5. Comparison of Clustering Method

Our proposed revised GN clustering algorithm has better clustering result for teaching and learning, cost less time than K-means clustering, and is significant in the quasi-experimental method described in section 3.4. The following introduces the compared clustering results for teaching need, time complexity comparison, pretest, and posttest learning effectiveness comparison.

There are three clustering algorithms, such as k-means clustering, our revised GN clustering, and students' willingness clustering. Figure 3 shows the solid line in the k-means clustering algorithm, which shows the result in 6 groups. Each group with large, varied student number (1,3,2,3,15,20). K-means results are not appropriate for real classroom teaching. The second method is our revised GN clustering, shown in the dashed line in Figure 3, which generated 9 groups with close numbers (4,5,5,5,5,5,5,5,5). Our revised GN clustering result is the best case for cooperative learning. The third is traditional teaching, which is grouped by students' willingness to cooperate in learning with a dotted line in Figure 3 with large, varied numbers (3,4,5,5,6,10,11). Most teachers need to negotiate with students in groups again.

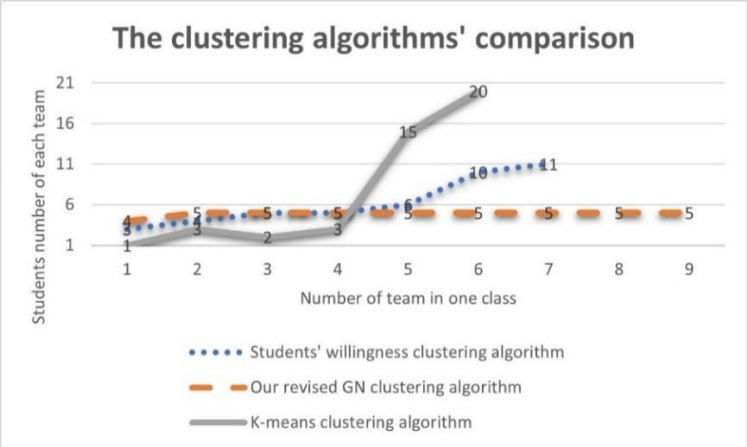


Fig. 3. The cluster difference among Weka K-means clustering [18], students' willingness clustering, and our revised GN clustering

Time complexity

The K-Means algorithm is a good example, which is one of the most widely used in literature. K-Means algorithm time complexity is O(N) [19]. The Girvan-Newman algorithm time complexity is $O(N^3)$ and $O(m^2n)$ [19], which we adapted in our research. In this experiment and most teaching experience, the number of the class will not be bigger than 100 students. Therefore, the cost time will not have a large influence.

Pretest and posttest learning effectiveness comparison

Figure 4 shows the mean score in our revised GN algorithm makes students' scores improve from pretest 53.45 to posttest 63.72 (Figure 4, dashed line). The student's willingness mean score improved from the pretest 52.93 to the posttest 55.43 (Figure 4, dotted line). Section 3.4 concludes that using our revised GN clustering algorithm is more efficient than the other method in programming learning.

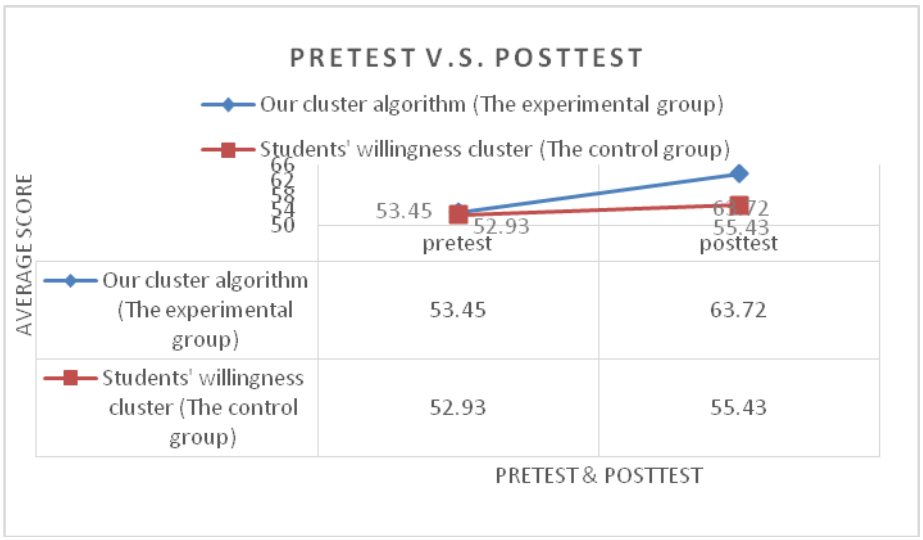


Fig. 4. The comparison between students' willingness clustering and our clustering

Social Network Analysis & Interpersonal Relationship

Social network analysis applied the student's interpersonal relationship questionnaire to generate the SNA (social network analysis) graph. The first step in our algorithm is using the student's interpersonal relationship to produce the SNA graph.

In Figure 5(a), the number of each group is too different. In a cooperative learning environment, it is not easy to arrange more than 5 students in a group. The more students in a group, the learning efficiency becomes lower. The cooperating learning suggested number is four to five. All the teams are arranged with high, middle, and low-score group students. Figure 5(b) shows the final clustering result.

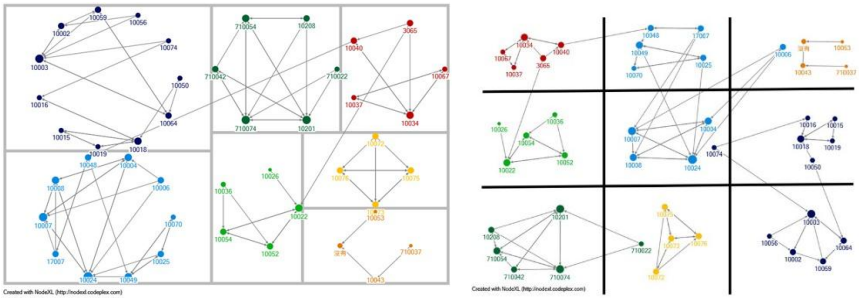


Fig. 5. (a) Original experimental group clustering graph (b) The experimental group clustering graph after our revised GN clustering

4. Lag Sequential Analysis of Programming Exam Videos

Content analysis involves the study of documents and communication methods, including text, images, audio, or video. Scientists have applied content analysis to investigate communication patterns in a replicable and systematic manner. The noninvasive nature of content analysis is a crucial advantage when using it for examining social phenomena. Researchers can simulate social situations, collect survey questionnaires, or record videos to reveal patterns. Computer-based content analysis methods are being increasingly used [20-23]. Video, answers to open-ended questions, newspaper articles, online discussions, medical records, or experimental observations can be systematically analyzed after conversion to a machine-readable format. The input is analyzed and coded into categories to reveal patterns. Some computer-assisted methods can reduce the time required to analyze large digital data sets. Certain studies have eliminated the need to establish intercoder reliability for multiple human coders. However, human coders are still critical in content analysis because they are superior to computers for recognizing nuance and latent meanings in text.

Table 4. Coding scheme

Code	Phase / Description
C1	Coding/Debug: The process of students writing programs or debugging, and it also includes debugging, copying and pasting code, and compilation and testing.
C2	Search for information: Search for information on the internet, watch programming videos, or read other programs. It involves Internet references, assignments previously uploaded to the platform, reference materials, files on the platform, or recorded teaching videos.
C3	Review questions/code/Debug information: Viewing or reading the exam questions, the student's program, the debug information, or program execution results.
C4	Thinking: Think about how to code or what to do next.
C5	Others: Other than the above four codes. For example, asking a teacher a question on the platform, opening a folder or file, saving a file, saving as a new file, switching windows quickly with no obvious action, and other miscellaneous actions not covered by the other four categories.

Our coding schema is introduced in Table 4. The problem-solving behaviors displayed by students in our recorded videos were analyzed and labeled using five codes. The recorded videos are recorded on students' computer screens, which is automatic recording. We can record students' movements when they are solving problems and writing programs.

Lag sequential analysis [24,25] has become an important tool for researchers of interpersonal interaction. This method [26] enables one to explore and summarize cross-dependencies occurring in complex interactive sequences of behavior.

Lag sequential analysis of individual interactions was explored as a tool to generate hypotheses regarding the social control of inappropriate classroom behavior of students with severe behavior disorders. Gunter et al. [27] proposed three coded events (student hand raise, teacher attention, and the "stop code") that were identified as highly related to the student's disruptive behavior. The results are discussed in terms of the usefulness

of the analysis procedures in contributing to the functional analysis of students' classroom behavior.

This study [28] then discusses the different learning behavior patterns based on the theoretical framework of Hofstede's National Cultural Dimensions (NCD). The obtained results highlighted that students from each culture behave differently due to several interconnecting factors, such as educational traditions.

This study [29] examines it in the context of 83 elementary schoolers' mobile serious game-playing behaviors. Lag-sequential analysis of the participants' observed behavioral patterns, and of differences in such patterns between two performance subgroups (i.e., students with high vs. low academic performance), yielded two main findings. First, all these young learners exhibited knowledge construction, and moved smoothly from lower to higher phases of it in the mobile environment; and second, the high-performing group attained a deeper level of knowledge construction through the negotiation of meaning than the low-performing group did. Some theoretical and practical implications of these results are also discussed.

This study applied lag sequential analysis to find out the obvious transition of the programming actions. The five codes are discussed and referenced [30] the problem-solving code. The coding scheme definition is listed in the following.

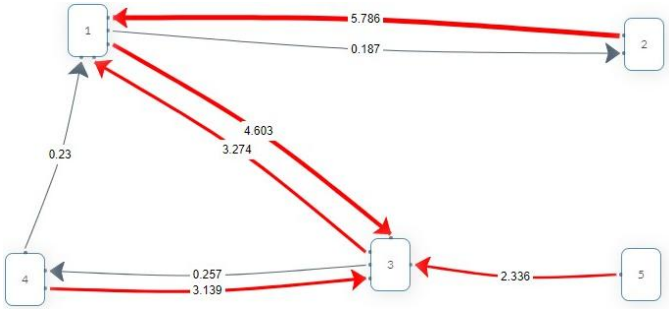


Fig. 6. The experimental group's sequential analysis

This section describes the sequential analysis of the experimental group and compares the group to check if there is obvious movement from one state to the other state. The distribution of the experimental group content analysis is as follows: C1 1223 times, 51.3%; C2 325 times, 13.6%; C3 862 times 36.2%; C4 525 times, 22%; C5 130 times, 5.5%. The distribution of the experimental group sequential analysis is as follows: C1 to C3 4603 times; C2 to C1 5786 times; C3 to C1 3274 times; C4 to C3 3139 times; C5 to C3 2336 times. According to the [31] Allison and Liker (1982) used the z score to calculate. We obtain the following obvious transition in Figure 6.

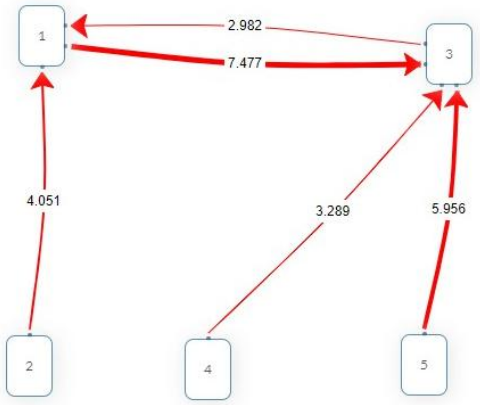


Fig. 7. The compared group's sequential analysis

The distribution of the experimental group content analysis is as follows: C1 968 times, 45.8%; C2 386 times, 18.3%; C3 996 times 47.1%; C4 315 times, 14.9%; C5 128 times, 6.1%. The distribution of the experimental group sequential analysis is as follows: C1 to C3 7477 times; C2 to C1 4051 times; C3 to C1 2982 times; C4 to C3 3289 times; C5 to C3 5956 times. According to [31] Allison and Liker (1982) used the z score to calculate. We obtain the following obvious transition in Figure 7.

The Experimental group has two obvious loops, C1C3, C3C4 and C1C2, even C1 to C2 and C3 toC4 are not so obvious. However, it shows the experimental group with our revised GN clustering makes more learning efficiency in programming.

5. Conclusion

In this study, GN clustering based on the betweenness value between students was combined with a grouping algorithm based on IRT to develop a combined methodology for estimating learner ability to achieve heterogeneous student grouping for cooperative learning. An experimental group of students clustered using our proposed SNA approach had significantly higher post-test scores than did a control group of students who grouped themselves.

Acknowledgment. I thank the Taiwanese Ministry of Education for financially supporting this study under the Teaching Practice Research Program. And thanks to Prof. Yang Hsin-Che for discussing this study idea, and Chang An-Ray for data analysis.

References

1. 1. Johnson, D. W., Johnson, R., Houbec, E.: Cooperation in the classroom (7th ed.). Edina, MN: Interaction Book Company. (2008)

2. Girvan M., Newman M. E. J.: Community structure in social and biological networks, in Proceedings of the National Academy of Sciences, 99 (12) pp.7821-7826. (2002)
3. Omonayajo, B., Al-Turjman, F., Cavus, N.: Interactive and Innovative Technologies for Smart Education. Computer Science and Information Systems, 19(3), pp.1549-1564. (2022)
4. Yu, J., Lin, J.: Data mining technology in the analysis of college students' psychological problems. Computer Science and Information Systems, 19(3), pp.1583-1596. (2022)
5. Dagienė, V., Jevsikova, T., Stupurienė, G., Juškevičienė, A.: Teaching Computational Thinking in Primary Schools: Worldwide Trends and Teachers' Attitudes. Computer Science and Information Systems, Vol. 19(1), pp.1-24. (2022)
6. Zhang G., Wang X., Zhao R.L., Wang C., Wang C.: Construction of Innovative Thinking Training System for Computer Majors under the Background of New Engineering Subject. Computer Science and Information Systems, 19(3), pp.1499-1516. (2022)
7. Dale, E.: Audio-Visual Methods in Teaching (3rd ed., p. 108). Holt, Rinehart & Winston, New York: Dryden Press. (1969)
8. Yang H.-C., Chang W.C.: Ubiquitous smartphone platform for K-7 students learning geography in Taiwan. Multimedia Tools Application 76, pp.11651–11668 (2017).
9. Chang W.-C.: Integrating Social Network Analysis, Content Analysis and Sequence Analysis with Cooperative Learning in the Programming Courses: A Case Study, International Journal of Engineering Education, Vol. 39 No. 2, pp.397-408. (2022)
10. Cheng S.-C., Cheng Y.-P., and Huang Y.-M., Enhancing Students' Learning Performance by Combining Flipped Learning and Online Formative Assessment Platform, International Journal of Engineering Education, Vol. 39 No. 2, pp.409–419. (2022)
11. Otte, E.; Rousseau, R.: Social network analysis: a powerful strategy, also for the information sciences". Journal of Information Science. 28(6), pp.441-453. (2002)
12. Milgram, S., The Small World Problem. Psychology Today, 2, pp.60-67. (1967)
13. Watts, D., Strogatz, S.: Collective dynamics of 'small-world' networks, Nature 393, pp.440-442. (1998)
14. Barnes, J.A.: Class and Committees in a Norwegian Island Parish. Human Relations, 7, pp.39-58. (1954)
15. Pattison P.: Algebraic Models for Social Network, Cambridge University Press, (1993)
16. Tseng, C. H.: The Type of Network Organization of MNC"s Subsidiary in Taiwan and the Management Mechanisms Adopted by MNC"s Headquarters, National Cheng-chi University, Unpublished doctoral thesis (1995)
17. Kelly T.L.: The selection of upper and lower groups for the validation of test items. Journal of Educational Psychology, 30, pp.17-24. (1939)
18. Weka official website at the University of Waikato in New Zealand, *https://www.cs.waikato.ac.nz/~ml/weka/*
19. Elbattah, M., Roushdy, M., Aref, M., Salem M.A.: Large-Scale Entity Clustering Based on Structural Similarity within Knowledge Graphs, Book chapter of Big Data Analytics: Tools, Technology for Effective Planning, pp.311-334. (2017)
20. Jeffery Chiang, *https://medium.com/analytics-vidhya/girvan-newman-the-clustering-technique-in-network-analysis-27fe6d665c92*, Last Accessed 02/07/2022.
21. Pfeiffer, S., Stefan F., Wolfgang E.: Automatic audio content analysis, Technical Reports 96 (1996).
22. Grimmer, J., Brandon M. S.: Text as data: The promise and pitfalls of automatic content analysis methods for political texts, Political Analysis 21(3), pp.267-297. (2013)
23. Nasukawa, Tetsuya, Jeonghee Yi.: Sentiment analysis: Capturing favorability using natural language processing, in Proceedings of the 2nd international conference on Knowledge capture. ACM (2003)

24. Sackett, G. P.: The lag sequential analysis of contingency and cyclicity in behavioral interaction research. In J. D. Osofsky (Ed.), Handbook of infant development (pp. 623-649). New York: Wiley. (1979)
25. Sackett, G. P.: Lag sequential analysis as a data reduction technique in social interaction research. In D. B. Sawin, R. C. Hawkins. (1980)
26. Faraone S.V., Dorfman D.D.: Lag sequential analysis: Robust Statistical Methods, Psychological Bulletin, 101(2), pp.312-323. (1987)
27. Gunter, P. L., Jack, S. L., Shores, R.E., Carrell, D.E., Flowers, J.A: Lag Sequential Analysis as a Tool for Functional Analysis of Student Disruptive Behavior in Classrooms, Journal of Emotional and Behavioral Disorders, SAGE Publications Inc, 1(3), pp.1063-4266. (1993)
28. Tlili A., Wang H.H., Gao B.J., Shi Y.H., Nian Z.Y., Looi C.-K., Huang R.H.: Impact of cultural diversity on students' learning behavioral patterns in open and online courses: a lag sequential analysis approach, Interactive Learning Environments, pp.1-20. (2021)
29. Sun, Z., Lin, CH., Lv, K., Song J.: Knowledge-construction behaviors in a mobile learning environment: a lag-sequential analysis of group differences. Education Tech Research Dev 69, pp.533–551 (2021).
30. Hou, H. T., Sung Y. T., & Chang, K. E.: Exploring the behavioral patterns of an online knowledge sharing discussion activity among teachers with problem-solving strategy. Teachers and Teachers Education, 25(1), pp.101-108. (2009)
31. Allison P. D., Liker, J.K.: Analyzing sequential categorical data on dyadic Interaction, Psychological Bulletin, 91, pp.393-403. (1982)

Wen-Chih Chang received B.S., M.S., and Ph.D. degrees in computer science and information engineering from Tamkang University, Taipei, Taiwan. He has been an Associate Professor in the International master program in Information Technology and Applications at the National Pingtung University, Taiwan. His current research interests include social network analysis and AI in e-learning.

Received: August 30, 2022; Accepted: September 19, 2023.

Computer Science and Information Systems 21(2):507–524 https://doi.org/10.2298/CSIS221115070K

A Study of Identity Authentication Using Blockchain Technology in a 5G Multi-Type Network Environment

Jui-Hung Kao[1,*], Yu-Yu Yen[2,3], Wei-Chen Wu[4], Horng-Twu Liaw[5], Shiou-Wei Fan [6], and Yi-Chen Kao[7]

[1] Department of Information Management, Shih Hsin University, Taipei, Taiwan
kjhtw@mail.shu.edu.tw
[2] Center of General Education, Shih Hsin University, Taipei, Taiwan
melyen@mail.shu.edu.tw
[3] Department of Biomedical Engineering, National Yang Ming Chiao Tung University, Taipei, Taiwan
sheepkelly19.be11@nycu.edu.tw
[4] Department and Graduate Institute of Finance, National Taipei University of Business, Taipei, Taiwan
weichen@ntub.edu.tw
[5] Department of Information Management, Shih Hsin University, Taipei, Taiwan
htliaw@mail.shu.edu.tw
[6] Department of Information Management, Shih Hsin University, Taipei, Taiwan
fan@mail.shu.edu.tw
[7] Department of Information Management, Shih Hsin University, Taipei, Taiwan
i110925102@mail.shu.edu.tw

Abstract. The 5G technology, known for its large bandwidth, high speed, low latency, and multi-connection capabilities, significantly accelerates digital transformation in enterprises, especially in addressing factory automation challenges. It facilitates efficient machine-to-machine (M2M) and device-to-device (D2D) connectivity, ensuring rapid data transfer and seamless process convergence under 5G standards. Although 5G offers substantial communication and low latency benefits, its limited indoor coverage requires the deployment of decentralized antennas or small base stations. In contrast, Wi-Fi 6 seamlessly complements 5G, providing superior indoor mobile connectivity. This integration is crucial for businesses looking to accelerate digital transformation. To optimize 5G, the deployment of devices such as bypass switches, SDN switches, and MEC in the 5G Local Breakout network enables user access control and fast authentication. Real-world validation confirms the effectiveness of these measures, which are expected to lead to the future of 5G mobile networks.

Keywords: Fifth-Generation Mobile Communication, Blockchain, Identity Authentication.

* Corresponding Author

1. Introduction

Literat In recent years, the era of fifth generation mobile networks (5G) has arrived, and countries around the world are competing to invest in 5G development resources. The International Telecommunication Union (ITU) has compiled trends released by a variety of organizations and proposed the 5G system specification (International Mobile Telecommunications 2020, IMT 2020), which emphasizes that future communications must meet eight indicators of technical requirements (8KPIs) and three application requirements. The overall system capacity of Extreme Mobile Broadband (eMBB), Massive Machine Type Communication (mMTC), and Ultra Reliable Low Latency Communication (uRLLC) is 1000 times that of 4G (4th Generation Mobile Networks) to meet the bandwidth requirements of 5G communication [1].

As a new technology, 5G has some limitations in terms of physical constraints. Wireless 5G signals will be transmitted at a significantly smaller distance than 4G; that is, to serve devices within the same range, 5G will require more base stations than 4G, which is undoubtedly a barrier to 5G adoption, whether due to the impact of deployment time or increased cost. To overcome these limitations, a possible solution is to use the free unlicensed spectrum available in Wi-Fi (Wireless Fidelity) technology. As such, a complementary solution is proposed to have 5G and Wi-Fi 6 coexist, so that the two technologies can complement each other to provide better service quality and higher speed, lower latency, and higher capacity for end users. However, in this multi-type 5G network environment, how to enable IoT devices to have a unique and identifiable identity, with undeniability and privacy, and the ability to authenticate each other and switch connections in different network environments without interruption has become an important issue [2, 3].

The core of the 5G system is secure identity management, where only users who have passed identification and authentication can access network services. 5G inherits the powerful cryptographic components (e.g., key generation functions and interdevice and internetwork authentication) and security features of the original 4G system. It should be mentioned that a new security function in the 5G system is the identity authentication framework, which provides mobile service operators with the flexibility to choose the identity authentication credentials, logo format and authentication method for users and IoT devices, unlike previous mobile networks that required a physical SIM (Subscriber Identity Module) card as the credential. The different authentication methods available are called the 5G Authentication and Key Agreement (5G-AKA) and the Extensible Authentication Protocol (EAP) [4, 5].

The purpose of this study is to investigate the interoperability between the fifth-generation mobile communication network and the new wireless LAN technology of Wi-Fi 6. Both Wi-Fi 6 and 5G have improved transmission efficiency, bandwidth, and quality, which is of great help for manufacturing automation, telemedicine, and other critical IoT devices in many industries. Regarding the issues of how to retain original characteristics and also care for information security in these fields, this study focuses on how to use blockchain technology to identify IoT devices when switching between Wi-Fi 6 and 5G signals for research and discussion.

2. Materials and Methods

2.1. Introduction to the 5G network environment

There are two environmental modes of 5G network architectures [6]: NSA and SA. The first is the NSA architecture, which is the 5G network formed by LTE (long-term evolution) 4G technology and the 5G radio access architecture; the second is the SA architecture. Before discussing the SA architecture, we should first introduce 5G NR. NR is the name of 5G New Radio, which is a global standard for 5G with OFDM (Orthogonal Frequency Division Multiplexing) and this standard was approved as a 5G connectivity standard by the international standards organization 3GPP (The Third Generation Partnership Project), which is composed of enterprises such as Huawei and Samsung, etc. Therefore, the 5G SA is composed of new radio access technology (RAT), which is different from the 5G process made up of the NSA.

The difference between 5G and 4G technologies (as shown in Fig. 1.) is that 5G NR uses a large number of Parallel Narrowband Subcarriers instead of Single Broadband Carriers to transmit data, so NR can cover low frequencies (450 MHz to 6000 MHz), lower frequencies than 6 GHz and higher frequencies (24250 MHz to 52600 MHz), higher frequencies than 24 GHz and millimeter wave range; that is, it can fully cover the spectrum from 6 GHz to 100 GHz in the millimeter wave (mmWave) band to meet the standard required for 5G. The emergence of NR technology is very helpful for the three main characteristics of 5G, eMBB, mMTC and URLLC, allowing for a new specification and standard for 5G [**7**].

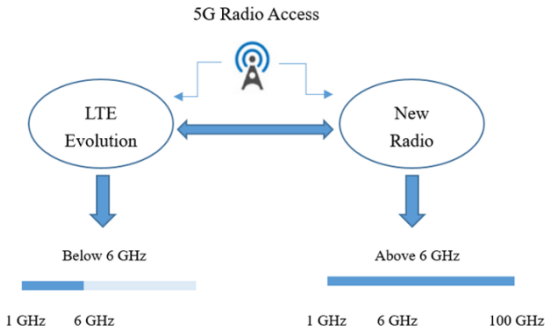


Fig. 1. NSA and SA Environments

2.2. Introduction to the 5G authentication mechanism

The 5G authentication mechanism is a continuation of the 4G authentication mechanism with improvements including the 3 identity authentication mechanisms of EAP-AKA

(Extensible Authentication Protocol-Authentication and Key Agreement), 5G-AKA and EAP-TLS (Extensible Authentication Protocol-Transport Layer Security), of which EAP-TLS [8] is defined in 5G for limited use conditions (e.g., IOT environments).

The Service Based Architecture (SBA) is proposed for 5G in the definition of the core network (as shown in Fig. 2.) [9]. The physical architecture and service request definition are very different from those of the 4G core network, so we will not discuss the 4G architecture here but focus on the 5G architecture, which has the following five major parts:

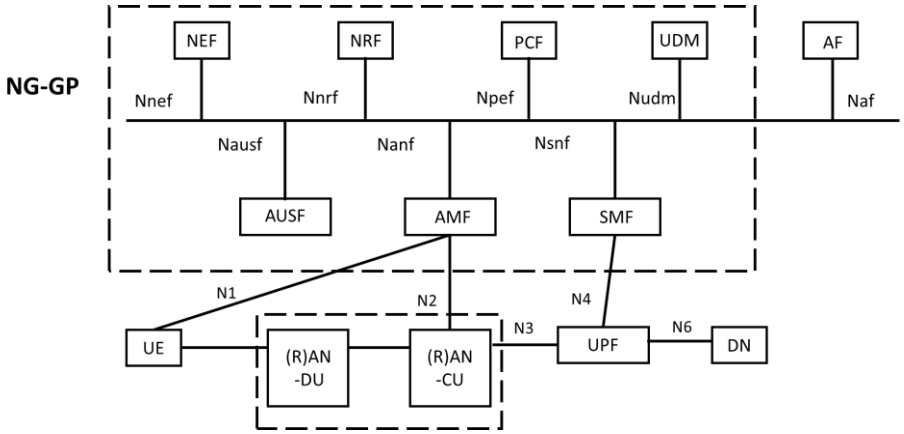


Fig. 2. SBA Services for the 5G Core Network Architecture

(1) Security Anchor Function (SEAF)

UE and its local network need to communicate through it during the identity authentication process. It can reject UE identity authentication, but relies on the local network of UE to accept identity authentication.

(2) Authentication server function (AUSF)

It has the function to decide whether or not to allow the connection for UE authentication, and it relies on the back-end service to calculate authentication data and keys if 5G-AKA or EAP-AKA is used.

(3) Unified Data Management (UDM)

It is an entity that carries functions related to data management, such as the Authentication Credential Repository and Processing Function (ARPF), which selects identity authentication methods based on subscriber identity and configured policies and calculates identity authentication data and keys as needed.

(4) Subscription Identifier Deconcealing Function (SIDF)

The subscription concealed identifier (SUCI) is decrypted to obtain its long-term identity; that is, the subscription permanent identifier (SUPI), such as the International Mobile Subscriber Identity (IMSI). In 5G, transmission is always done encrypted over the wireless port. More specifically, public-key-based encryption is used to protect the SUPI (Subscription Permanent Identifier). Therefore, only SIDF (Subscriber Identity Deconcealing Function, SIDF) has access to the private key associated with the public key assigned to UE (User Equipment) to encrypt its SUPI.

2.3. 5G-AKA

AUSF (Authentication Server Function) provides the identity authentication service via Nausf_UE authentication, while UDM provides the identity authentication service via Nudm_UE authentication [10]. A brief description of the 5G authentication procedure is shown in Fig. 3. [11]:

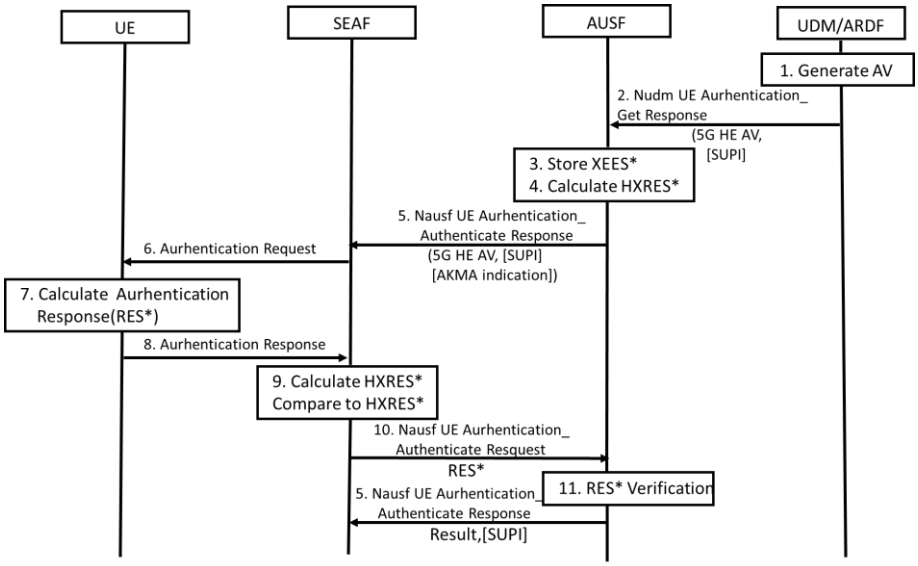


Fig. 3. 5G authentication procedure

2.4. Wi-Fi 6 identity authentication

Wireless security is an important issue for WLAN systems. Since wireless networks use the so-called open medium, which uses public electromagnetic waves as carriers for transmitting data signals, and there is no physical line connecting the two communicating parties, the risk of data theft is very high if proper encryption or other protection measures are not taken during the transmission process. Therefore, it is particularly important to ensure data security in the wireless network environment of WLAN [12, 13].

(1) Basic Concepts

802.11i is the latest wireless network security standard. IEEE has proposed additional amendments to compensate for the insecure encryption functions of 802.11 with the concept of RSN (Robust Security Network) added to the 802.11i standard to enhance the encryption and authentication functions of wireless network data transmission, and to address the shortcomings of WEP (Wired Equivalent Privacy) encryption mechanism with many corrections [14]. The proposed solution for identity authentication in the

802.11i standard is based primarily on the 802.1X framework and the extensible authentication protocol (EAP), while the encryption algorithm is based on the encryption algorithm of the Advanced Encryption Standard (AES) [15, 13].

(2) Introduction to the Link Authentication Method

The so-called link authentication refers to the 802.11 identity authentication, which is a low-level authentication method. It occurs when an STA is associated with an AP over 802.11, which precedes access authentication. Any STA must be authenticated using the 802.11 identity authentication method before trying to connect to the network, and 802.11 identity authentication can be thought of as the starting point of the handshake process when an STA (station) connects to the network, which is the first step in the so-called network connection process [16]. The IEEE 802.11 standard defines two types of link layer authentication: Open System Authentication and Shared Key Authentication, which are briefly described below:

1) Open System Authentication

This means that any user is allowed to access the wireless network in the sense that no data protection is actually provided, that is, no authentication. In other words, if the authentication type is set to open-system authentication, all STA requests for authentication will pass 802.11 authentication. The open-system authentication consists of two steps: The first step is to request authentication from the STA, and its data contain the STA's ID (Identity) (Media Access Control Address) after the STA sends the authentication request. The second step is for the AP (Access Point) to send back the authentication results, and the content of the authentication reply issued by the AP contains whether or not the authentication result is a success or failure. If the authentication result is "success", then STA and AP have passed the two-way authentication.

2) Shared Key Identity Authentication

The so-called shared key authentication refers to another authentication mechanism in addition to the open system authentication mentioned above. Shared-key authentication requires both STA and AP to be configured with the same key, and the authentication process is as follows. Step 1, STA first sends an authentication request to AP; Step 2, after receiving the authentication request, AP randomly generates a Challenge packet (i.e., a string) and then transmits the string to STA; Step 3, STA copies the string received from AP into a new message and then encrypts it with the key and sends it back to AP; Step 4, after receiving the message from STA, AP will decrypt the message with the key and then compare the decrypted string with the one given to STA at the beginning; if they are the same in accordance with the comparison, it means STA has the same shared key at the wireless device, that is, it has passed the shared key authentication requirement; otherwise, the shared key authentication result is "failure."

2.5. Integration of Two Access Technologies

IOTA (Integration of Two Access Technologies) was founded in 2015 by David Sønstebø and others with the goal of enabling the communication of various devices on the IoT, which is faster, used by more people, and can withstand a larger number of transactions than traditional blockchain, and is currently the most popular decentralized ledger technology in Europe. In recent years, more and more cities are moving smart and

providing many citizen-friendly smart city services, among which the number of services that users have to pay for is increasing. IOTA proposes a block-chain technology solution for Internet of Things (IoT) systems that aims to over-come the limitations or problems of existing IoT systems mentioned above. As de-scribed above, the rarely mentioned characteristics of blockchain technology, such as decentralization, invariance, availability, tracking and tracing, and integrity, smart contracts make it a disruptive technology for IoT applications [17, 18].

IOTA is a kind of revolutionary public distributed ledger of the new generation with a new invention called "Tangle" at its core. Tangle is a new data architecture based on the Directed Acyclic Graph (DAG) [19]. Therefore, it has no blocks, no chains, and no miners. Due to this radical new architecture, IOTA works completely differently from other blockchains [20].

The main difference worth mentioning (other than DAG versus blockchain) is how IOTA reaches consensus and how it conducts transactions. As mentioned earlier, there is no miner role exists. This means that every participant in the network who wants to conduct transactions must actively participate in the network consensus by approving 2 past transactions. This proof of the validity of two previous transactions ensures that the entire network reaches consensus on the current status of approved transactions and enables a variety of unique functions that can only be seen in IOTA [21].

IOTA is the missing puzzle for the machine economy to fully emerge and play out its intended potential. We envision IOTA as the public, permission-exempt backbone of IoT, enabling true interoperability between all devices.

Due to its architecture, IOTA has a unique series of functions [22]:

Scalability

IOTA can achieve high transaction throughput thanks to parallel validation of transactions, with no limit to the number of transactions that can be validated at a given interval.

No transaction fees

With the launch of smart contracts in November 2021, IOTA does not charge transaction fees, which is a great advantage and is a good choice for data transaction validation.

Decentralization

IOTA has no miner role, and every participant who performs transactions on the network is actively involved in the consensus. Therefore, IOTA is more decentralized than any blockchain.

Quantum Immunity

IOTA uses a new technique, called Curl's ternary hash function, which can resist quantum attacks and avoid brute-force cracking attacks.

3. Results

The number of 5G users is growing and the trend is to have multiple heterogeneous wireless network interfaces on mobile devices. Many smart mobile phones are al-ready equipped with wireless LAN interfaces. However, these mobile devices often lack an effective mobility management mechanism to take full advantage of these heterogeneous

network interfaces at the same time. To solve this problem, we use blockchain technology to design a set of intermediary mechanisms that can integrate and roam efficiently among heterogeneous networks, making the identity authentication of end devices more convenient and secure.

3.1. IOTA decentralized ledger technology

IOTA's underlying ledger architecture, Tangle, is not designed in terms of blocks and chains, but rather in terms of a decentralized architecture. When the data are placed on the IOTA's decentralized ledger [23], they are copied and distributed to numerous network nodes to achieve the characteristic that the data cannot be tampered with. In addition, Tangle does not have a mining mechanism [24], but rather validates transactions through IOTA users and, therefore, does not require transaction fees. The nature of Tangle architecture is that the larger the transaction size, the higher the availability, so it is more suitable for the quantitatively large IoT industry than traditional blockchain technology [25].

Tangle

Tangle, as mentioned by IOTA, has a data structure of a directed acyclic graph (DAG) where each message is attached to 2 to 8 previous messages, and anyone can attach messages at different locations in front of Tangle, and the protocol can process these different messages in parallel. There is no cost to send a message on Tangle, because the network has no miners or pledgers. In Tangle, PoW (Proof-of-Work) is not used to protect the network; instead, PoW is only used to block spam, and all IOTA nodes validate messages and use different functions to reach consensus when confirming messages [26].

Directed Acyclic Graph

In general, IOTA operates in such a way that there is no domain-wide blockchain, but a directed acyclic graph (DAG), which is the Tangle described in the previous section. All transactions issued through the nodes constitute the Tangle, the set of ledgers in which all transactions are stored. When a new transaction is created, it must validate two previously completed transactions, and these validation relationships are represented by the directed nodes. If there is no direct-connected directed node from transaction A to transaction H, but there is a directed node path of length at least greater than two, we say that transaction A indirectly validates transactions B and D. Furthermore, there is a Genesis transaction that is validated directly or indirectly by all transactions (as shown in Fig. 4.). Assuming that H is the Genesis Transaction, the following description is given in the IOTA technology: At the beginning there is an address that has all the tokens. Then, through the behavior, the Genesis Transaction will transfer the IOTA coins to other founder’s addresses, stating that all tokens are generated by the Genesis Transaction, which means that no new token will be generated, and this is also the reason why the DAG will not loop. In fact, simply put, it is the concept of receiving IOTA coins without the need for mining behavior [27].

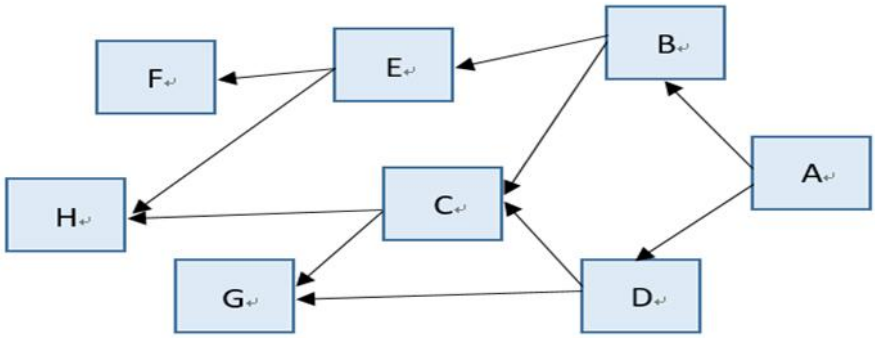


Fig. 4. 5G authentication procedure

3.2. The Cross-network Authentication Information Security Framework for 5G and Wi-Fi 6

With the proliferation of fifth-generation mobile communication technology (5G) and Wi-Fi 6, there has been an increasing demand for faster communication speeds, greater capacity, and enhanced data security. Against this backdrop, ensuring efficient and secure data exchange between these two major communication technologies has become a paramount concern. This article elucidates how IOTA's Keccak-384 and the sponge function facilitate secure data exchanges in cross-network authentication communication between 5G and Wi-Fi 6.

First, let us explore the background of IOTA and the rationale behind its adoption of Keccak-384 [28]. As mentioned previously, IOTA initially employed the SHA-3-384 algorithm, but due to potential security vulnerabilities, shifted its preference to Keccak-384[1]. The foundation of this algorithm is the sponge function. Thus, by using the Keccak-384 sponge function, IOTA ensures the security of its blockchain transactions, making it resistant to various collision attacks [29].

Introduced by Guido Bertoni's team in 2007, the sponge function is based on the properties of sponges, with the ability to absorb and squeeze out large amounts of data. This trait allows the sponge function to accept input of any length, undergo specific algorithmic processing, and produce output of the desired length [30].

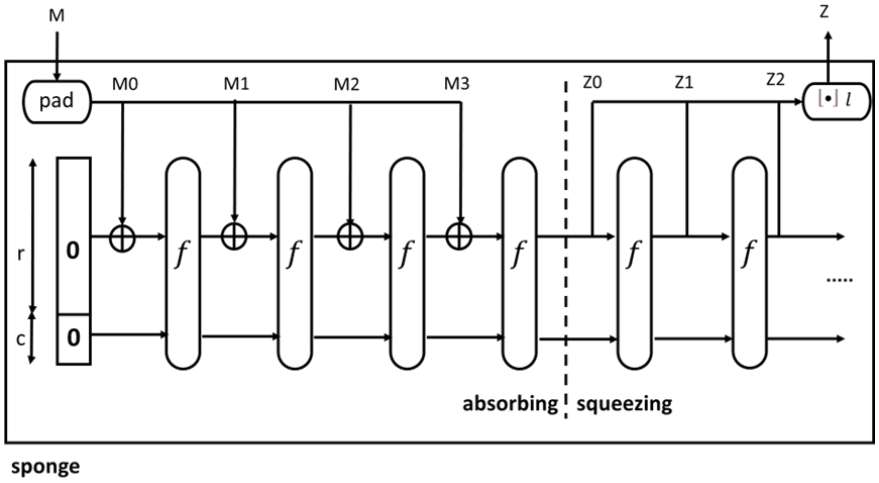


Fig. 5. Schematic diagram of the operation of the sponge function

As described by the IOTA Foundation, this function is made up of three main components:

Memory State (S): This particular state encompasses 'b' bits and is segmented into two sections: R (comprising 'r' bits) and C (amounting to b-r bits or 'c' bits). Here, 'r' denotes the Bitrate, while 'c' signifies the Capacity.

Transformation function (f): This function transforms S to a fixed size. IOTA has chosen Keccak-384 as its transformation function, which is a variant of the sponge function from the Keccak family.

Padding function (P): This function ensures that the input M is a multiple of the bitrate 'r'. It continually adds data until this condition is met, after which the padded data are segmented into multiple chunks of size 'r'.

In the high-speed communication environment of 5G and Wi-Fi 6, ensuring data integrity and security becomes paramount. Utilizing the sponge function allows efficient encryption of communication data. During the absorption phase, the original data are XORed with data that is a multiple of the bitrate 'r', and through the function 'f', are transformed multiple times, resulting in a complex data set. In the squeezing phase, the desired length data is extracted from the transformed data.

However, depending solely on the sponge function fails to meet all the security requirements of cross-network authentication communication between 5G and Wi-Fi 6. An effective mechanism to ensure the authenticity and integrity of communication between the two parties is essential. This is where IOTA's Keccak-384 comes into play. Through Keccak-384, IOTA can generate a unique address and signature, ensuring the authenticity and integrity of the data during transmission.

5G local breakout private network

Local breakout private networks focus on data offload processing and are typically deployed between the base station and the core network. Mobile communication operators can choose to deploy MECs at the appropriate location based on requirements such as business type, processing capacity, network planning, etc. to achieve transparent deployment of terminals and networks. According to ETSI White Paper No. 28 [10], if an MEC server is deployed in the core network, MEC can be integrated with S/P-GW, and when MEC is deployed near the RAN side of the wireless network, the MEC server can be a standalone network element, or MEC functions can be integrated into the hub node or eNodeB. If the MEC server is a standalone element, it can be a device of a different vendor from that of HubNode and eNodeB.

The MEC mobile edge computing network provides application developers and content service providers with cloud computing capabilities and the IT service environment for the mobile edge network to achieve ultralow latency, large bandwidth, and real-time access to network information with the following key technologies [11]:

(1) Temporary storage of content on the wireless side: the MEC server can obtain the hotspot content in the service, including video, pictures, documents, etc., through interconnection with the service system and carry out local temporary storage. During the service process, the MEC server performs real-time deep packet parsing of the data on the base station and can directly push the content in the temporary storage to the terminal if the service content applied by the terminal is already in the local temporary storage.
(2) Local diversion: users can access the local network directly through the MEC platform, and the local service data stream does not need to go through the core network, but is directly diverted to the local network by the MEC platform, which can reduce the return bandwidth attrition and service latency and improve the user service experience.
(3) Business optimization: through the MEC server near the wireless side, information from the wireless network can be collected and analyzed in real time, and the network conditions can be obtained to perform dynamic and quick optimization of services, select the appropriate service rate, content diversion mechanism, congestion control strategy, etc.
(4) Through the MEC platform, mobile networks can provide network resources and capabilities to third parties (MVNOs), open up capabilities such as network monitoring, network infrastructure services, QoS control, positioning, big data analysis, and others to the outside world, identify the development potential of network services, and achieve a win-win situation with partners.

4. Discussion

This study builds a 5G Local Breakout Private Network System environment and combines with Wi-Fi 6 to extract the MAC of the terminal carrier to build a security validation loop to validate that the IOTA transaction achieves a fast identity authentication mechanism to provide high-speed computing and reduce the transmission

latency, focusing on the local offload of information services to reach the near-side service access, and how to manage specific users accessing the field and provide a flexible and customized field management mechanism since the provision of application services in the field and the management of users accessing the field have high information security requirements.

4.1. 5G local breakout standalone private network

If an enterprise requires high network autonomy and privacy, it is provided with a standalone private network, from the base station, the MEC to the core network, with a complete set of mobile network deployment placed on the enterprise client side, and a standalone private network is set up to provide the dedicated base station for the enterprise, and the signal coverage is based on the area of the private network of the enterprise (as shown in Fig. 6.).

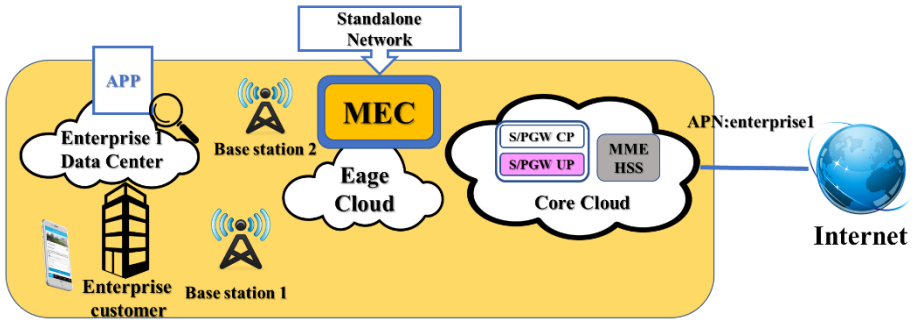


Fig. 6. MEC Standalone Private Network Architecture

In order to provide high-speed computing and reduce transmission latency, MEC mobile edge computing focuses on the local offload of information services to reach the near-side service access, and enterprises often deploy MEC directly in the fields of the enterprises. Since enterprise customers have high information security requirements for the provision of application services in the field and the management of users accessing the field, how to manage specific users accessing the field and provide a flexible and customized field management mechanism is the core of the problem to be discussed in this paper.

Through the implementation of MEC type private access, this architecture can access the signals and services flowing through the network by connecting in series base stations and core networks in series in a transparent and pass-through manner, and it can conduct in-depth tests of the MEC mobile edge computing network, develop dynamic service content, application host management, and content diversion mechanism, provide control of the users accessing the fields, and site equipment network management such as real-time alarm notification, remote monitoring, etc. It is the most flexible MEC mobile edge computing architecture.

4.2. Process Design for Identity Authentication of 5G and Wi-Fi Combined with IOTA

In this study, we simulate how to authenticate the terminal identity through the integration of two access technologies in a hybrid multi-type network environment of a standalone 5G Local Breakout private network combined with Wi-Fi 6. The architecture and flow of the system are shown in the system authentication flow chart. The main authentication mechanism is divided into three parts. The first part is the terminal device, the second part is the identity authentication web page, and the third part is the IOTA node (as shown in Fig. 7.).

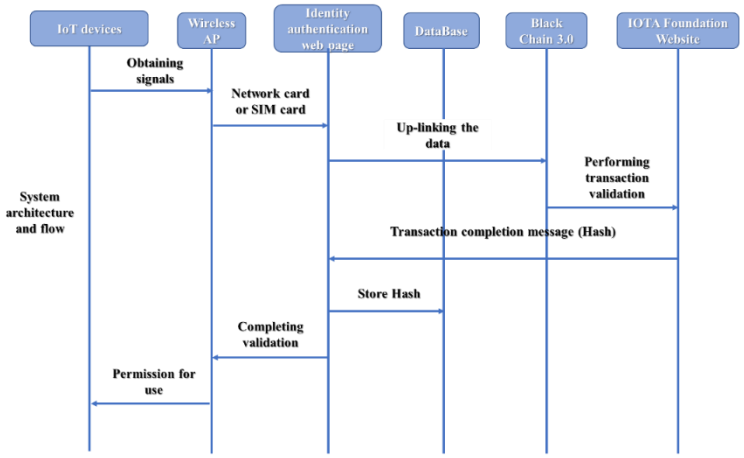


Fig. 7. System Authentication Flow Chart

4.3. IOTA node transaction validation

The 5G and Wi-Fi identity authentication process combined with IOTA includes a Hash that is obtained after the transaction is completed. To validate whether or not the data in this hash is the data of the original network card, this hash can be entered into the IOTA node website to query it (as shown in Fig. 8.).

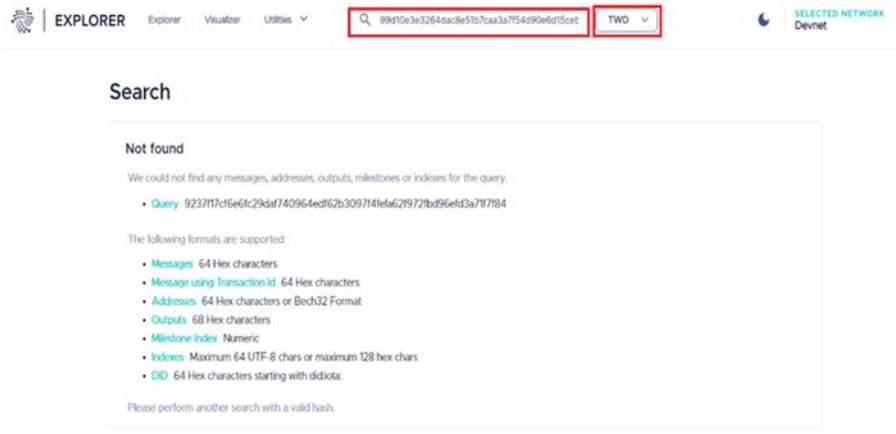


Fig. 8. IOTA Foundation Validation Screen

The data validation screen clearly shows that the Hash of this transaction can be used to find the data of the original transaction on the network card, which confirms that the result of this study is correct (as shown in Fig. 9.).

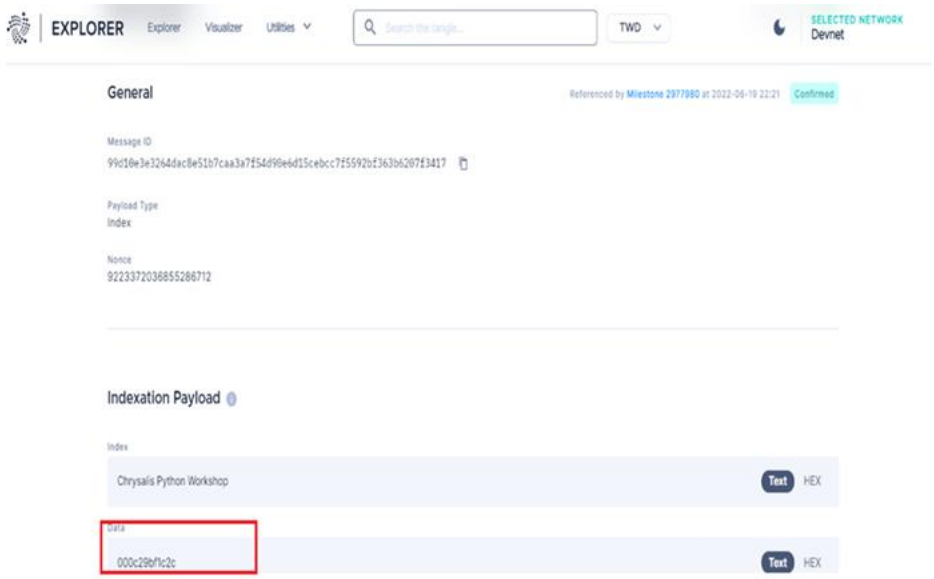


Fig. 9. Data Query Screen of the Node Transaction

From this validation website, we can see the transaction validation record screen (as shown in Fig. 10.). Before completing the validation of this transaction, we have validated the two-transaction data on the right side and know the transaction data we have completed and the transactions that are provided for validation.

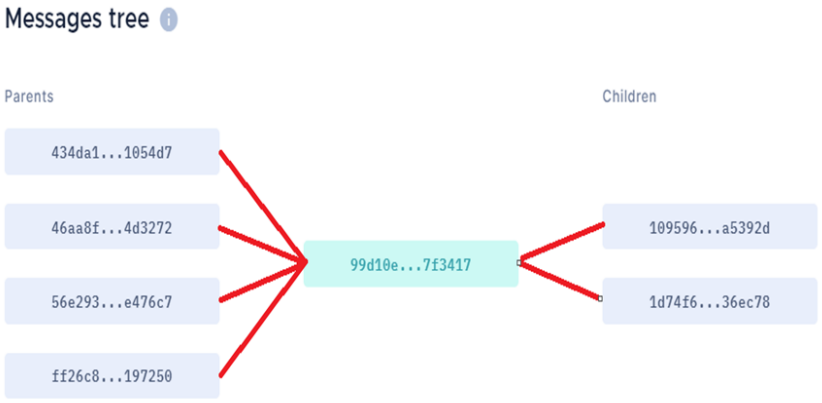


Fig. 10. Validation record screen of the node transaction

5. Conclusions

The lack of penetration of 5G signals requires the use of free unlicensed spectrum available in Wi-Fi technology to compensate for this problem. Therefore, the coexistence of 5G and Wi-Fi 6 in the future environment makes the two technologies complement each other, which has become the new trend in wireless communication in the future. The development of blockchain technology is no longer based primarily on mining; instead, it is replaced by the application of IoT and the validation through smart contracts. After blockchain 3.0 technology has overcome the problem that the more people use the blockchain, the slower it becomes and no longer has the role of miners. The IOTA Foundation created a new decentralized ledger technology called Tangle, which solves the current problem of blockchains 1.0 and 2.0 that the more people use it, the less efficient it is, and creates a new consensus method in a decentralized peer-to-peer solution. In other words, as long as two transactions are validated, it is no longer the mining ability that determines the trading partner.

The IOTA technology uses bundles to organize several transactions, including the output to the receiving address and the input from the sending address; in the IOTA technology transaction validation behavior, the transaction signature can be simply converted to the terminal MAC, so that the IOTA transaction mode is used for validation, and private nodes are set up in the 5G multi-type network. Then the relevant functions provided by IOTA are used to solve the identity authentication (network card) problem of IoT devices in WI-Fi -6 and 5G network environment through IO-TA technology.

Simply put, the characteristic that the IOTA signature can be converted to the terminal MAC is used to package the IOT network cards into a transaction using the Python function developed by IOTA, and the transaction is sent to the IOTA node for validation using the IOTA validation function. After validation, the HASH value of a transaction is sent back, completing a transaction. The next step is to prove the validity

of the obtained HASH value. In the IOTA node validation function, the above HASH value can be entered to decode the network data from the initial validation to achieve fast identity authentication.

This study enables the establishment of terminal devices that can allow WI-Fi-6 and 5G network environments to have unique and identifiable identities, with non-repudiation and privacy, and with the function of mutual authentication, for authentication in heterogeneous network environments. This study has already addressed this problem using IOTA technology, which can also be combined with blockchain technology in the new heterogeneous wireless network environment and is very convenient. Their applications can be very diversified, and we believe that more and more studies related to blockchain and 5G environment will be conducted to maximize the potentials of relevant technologies.

References

1. S. Henry, A. Alsohaily and E. S. Sousa.: 5g is real: Evaluating the compliance of the 3gpp 5g new radio system with the itu imt-2020 requirements, IEEE Access, 8, 42828-42840. (2020)
2. C.-C. LIU.: A lightweight security scheme with mutual authentication in mobile edge computing. Master's Thesis, Department of Information and Communication Engineering,Chaoyang University of Technology. (2020)
3. W. Serrano, The blockchain random neural network for cybersecure iot and 5g infrastructure in smart cities, Journal of Network and Computer Applications, 175, 102909. (2021)
4. K.-H. LIN.: An efficient group-based service authentication and session key negotiation scheme for mmtc devices in 5g. Master's Thesis, Computer Science and Information Engineering,National Chung Cheng University. (2019)
5. K. Yue, Y. Zhang, Y. Chen, Y. Li, L. Zhao, C. Rong and L. Chen.: A survey of decentralizing applications via blockchain: The 5g and beyond perspective, IEEE Communications Surveys & Tutorials , 23 , no. 4, 2191-2217. (2021)
6. M. Hirzallah, M. Krunz, B. Kecicioglu and B. Hamzeh.: 5g new radio unlicensed: Challenges and evaluation, IEEE Transactions on Cognitive Communications and Networking ,7, no. 3, 689-701. (2020)
7. E. Al Abbas, M. Ikram, A. T. Mobashsher and A. Abbosh.: Mimo antenna system for multi-band millimeter-wave 5g and wideband 4g mobile communications, IEEE Access ,7, 181916-181923. (2019)
8. Q. Hao, L. Sun, S. Guo, H. Liu, D. Qian and X. Zhu.: Improvement of eap-tls protocol based on pseudonym mechanism, 2021 International Conference on Wireless Communications and Smart Grid (ICWCSG), IEEE, 23-28. (2021)
9. Y. Siriwardhana, P. Porambage, M. Liyanage and M. Ylianttila.: A survey on mobile augmented reality with 5g mobile edge computing: Architectures, applications, and technical aspects, IEEE Communications Surveys & Tutorials ,23, no. 2, 1160-1192. (2021)
10. A. Ghosh, A. Maeder, M. Baker and D. Chandramouli.: 5g evolution: A view on 5g cellular technology beyond 3gpp release 15, IEEE access ,7, 127639-127651. (2019)
11. A. Yazdinejad, R. M. Parizi, A. Dehghantanha and K.-K. R. Choo.: Blockchain-enabled authentication handover with efficient privacy protection in sdn-based 5g networks, IEEE Transactions on Network Science and Engineering ,8 ,2019, no. 2, 1120-1132.
12. E. Mozaffariahrar, F. Theoleyre and M. Menth.: A survey of wi-fi 6: Technologies, advances, and challenges, Future Internet ,14, no. 10, 293. (2022)

13. K. Ramezanpour, J. Jagannath and A. Jagannath.: Security and privacy vulnerabilities of 5g/6g and wifi 6: Survey and research directions from a coexistence perspective, arXiv preprint arXiv:2206.14997. (2022)
14. J.-D. Li and C.-P. Fan.: Design and vlsi implementation of low latency ieee 802.11 i cryptography processing unit, Journal of Advances in Computer Networks ,8, no. 1. (2020)
15. B. Dey, S. Vishnu and O. S. Swarnkar.: An efficient dynamic key based eap authentication framework for future ieee 802.1 x wireless lans, Proceedings of the 2nd International Conference on Digital Signal Processing,125-131. (2018)
16. M. Rudenkova.: A methodology of modeling the ieee 802.11 wireless lan using ns-3, 2020 V International Conference on Information Technologies in Engineering Education (Inforino), IEEE, 1-4. (2020)
17. L. Tennant.: Improving the anonymity of the iota cryptocurrency, Univ. Cambridge, Cambridge, UK, Tech, 1-20. (2017)
18. B. Goswami and H. Choudhury.: A blockchain-based authentication scheme for 5g-enabled iot, Journal of Network and Systems Management ,30, no. 4, 1-33. (2022)
19. S. Popov.: The tangle, White paper ,1, no. 3. (2018)
20. M. Bhandary, M. Parmar and D. Ambawade.: A blockchain solution based on directed acyclic graph for iot data security using iota tangle, 2020 5th International Conference on Communication and Electronics Systems (ICCES), IEEE, 827-832. (2020)
21. S. Sicari, A. Rizzardi and A. Coen-Porisini.: 5g in the internet of things era: An overview on security and privacy challenges, Computer Networks,179, 107345. (2020)
22. I. Guidebook.: What is iota? (2022)
23. X. Xiao, F. Guo and A. Hecker.: A lightweight cross-domain proximity-based authentication method for iot based on iota, 2020 IEEE Globecom Workshops GC Wkshps, IEEE, 1-6. (2020)
24. C. Igiri, D. Bhargava, C. Udanor and A. Sowah.: Blockchain versus iota tangle for internet of things: The best architecture, Blockchain Technology, 259-278. ,(2022)
25. Y.-C. Wang.: Implementation of iot service management by applying blockchain and the study of peering management scheme in the chain. Master's Thesis, Department Communication Engineering,National Central University. (2019)
26. W. F. Silvano and R. Marcelino.: Iota tangle: A cryptocurrency to communicate internet-of-things data, Future Generation Computer Systems ,112, 307-319. (2020)
27. P. Gangwani, A. Perez-Pons, T. Bhardwaj, H. Upadhyay, S. Joshi and L. Lagos.: Securing environmental iot data using masked authentication messaging protocol in a dag-based blockchain: Iota tangle, Future Internet ,13, no. 12, 312. (2021)
28. I. Dinur, O. Dunkelman and A. Shamir.: Collision attacks on up to 5 rounds of sha-3 using generalized internal differentials, International Workshop on Fast Software Encryption, Springer, 219-240. (2013)
29. F. Liu, T. Isobe, W. Meier and Z. Yang.: Algebraic attacks on round-reduced keccak, Information Security and Privacy: 26th Australasian Conference, ACISP 2021, Virtual Event, December 1–3, 2021, Proceedings 26, Springer, 91-110. (2021)
30. M. Colavita and G. Tanzer.: A cryptanalysis of iota's curl hash function, White paper, 1-13. (2018)

Jui-Hung Kao is an assistant professor at Shih Hsin University since 2020. During his tenure as project manager at the Research Center for Humanities and Social Sciences in 2014, he was responsible for the administrative business of research and program execution, which combined statistical methods with spatial information visualization, and is good at writing programs and data analysis. The topics of empirical research focus on three parts: spatial data analysis, medical management research, and long-term medical policy.

Yu-Yu Yen has been working as adjunct assistant professor in the Center for General Education at Shih Hsin University since 2022, and has also been assisting in the 5G Education Network Industry-Academia Collaboration Project at the Center for Cloud & IOT research in the College of Management, Shih Hsin University. She also currently enrolled in a PhD program in the Department of Biomedical Engineering, National Yang Ming Chiao Tung University.

Wei-Chen Wu was born in Taipei, Taiwan R.O.C. He is Assistant Professor in the Department of Finance at the National Taipei University of Business. He received his Ph.D. degree in Information Management from National Central University in 2016. From 2020-2021, He was Assistant Professor in the Department of Finance at the Feng Chia University. From 2008-2016, he was also an Assistant Professor and Director of the Computer Center at the Hsin Sheng College of Medical Care and Management. His teaching interests lie in the area of programming languages, ranging from theory to design to implementation, and his current research interests include blockchain technology, fintech cybersecurity, network security, and deep learning. Wei-Chen Wu has collaborated actively with researchers in several other disciplines of computer science. He has served on many conference and workshop program committees and served as the workshop chair for Frontier Computing Conference (FC2017~FC2021) and Machine Learning on FinTech, Security and Privacy Conference (MLFSP2019~MLFSP2022).

Horng-Twu Liaw is a professor of Information Management at Shih Hsin University since 2004, and he is the Vice President of Shih Hsin University since 2018. He has studied e-Commerce and networking communities, service-oriented information technology and management, information system development and project management, network management and information security, and information security management. The focus of empirical research in recent years has focused on spatial data analytics, information security, big data analytics, and artificial intelligence.

Shiou-Wei Fan is a full-time lecturer at Shih Hsin University, has served as the Chief of network management division in the Office of Library and Information Services for 27 years. His main expertise is network management, network security, and cloud services.

Received: November 15, 2022; Accepted: September 19, 2023.

Computer Science and Information Systems 21(2):525–545 https://doi.org/10.2298/CSIS221117009L

An Empirical Study of Success Factors in Korea's Game Industry

Jun-Ho Lee[1], Jae-Kyu Lee[2], and Seung-Gyun Yoo[3,*]

[1] Division of Public Affairs and Police Administration, Dongguk University-WISE, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea
juno@dongguk.ac.kr
[2] Department of Information Management, Dongguk University-WISE, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea
duckjk89@dongguk.ac.kr
[3] Department of International Trade and Airline Service, Dongguk University-WISE, 123, Dongdae-ro, Gyeongju-si, Gyeongsangbuk-do, 38066, Republic of Korea
bluetrade@dongguk.ac.kr

Abstract. Korea's game industry is enjoying remarkable growth along with China and Southeast Asia. This study proposes and analyzes the relationships among characteristics of the basic environment, such as management, technology, marketing, and industry trends, among Korea's game companies. Through this analysis, game companies can attempt to achieve growth and expansion into global markets. From this study, these achievements can be made through leadership in technological development, by identifying competence in managers, and from awareness of the trends in markets and the game industry. Securing intellectual property rights to sustain performance and market expansion is one of the most important strategies in the game industry. In other words, the performance of a game company depends on the ability of managers to provide the newest story and user services, and to apply research and development in technology, marketing, and related industries. Because previous research has focused on the external aspects of games, including their effectiveness and impacts, this study differs in that it comprehensively considers internal aspects of the game company, the market, and the industry. This study explores the key success factors for improving corporate performance in Korea's game industry by setting up environmental, strategic, and performance models to investigate relevant factors. We also parameterize the market adaptation and R&D functions of companies. Through this research, we expect to support strategic decision-making in the game industry and contribute to enhancing the performance of game companies.

Keywords: Game Industry, R&D, Intellectual Property, Performance, Adaptation.

* Corresponding Author

1. Introduction

The game industry has expanded with the development of computers. Due to the influence of the Internet, which has increased since the 1990s, the game industry has moved away from time and space constraints and has become a global industry. Through its recent combination with virtual reality, the game industry is perceived as a business with a low risk of failure, unlike other cultural industries. In particular, it has become a culture beyond simple leisure activities through a combination of knowledge and technology.

Gaming is a comprehensive industry that encompasses various fields. This makes it a high value-added knowledge industry that has both cultural and industrial characteristics in movies, broadcasting, characters, and advertising. Just like other industries, it is important for the gaming industry to gain market leadership to improve performance. These initiatives stem from a variety of strategies that apply and extend the enterprise's internal and external capabilities.

The nature of knowledge-based industries recognizes the rights that can be obtained with new technologies and knowledge, because property rights are recognized as management resources of an entity through legal rights and protection schemes. In other words, companies are expanding their intellectual property rights and enhancing their competitiveness through research and development.

The fourth industrial revolution (Industry 4.0) strengthened the technical characteristics of the game industry through factors such as AI, deep learning, and big data. These environmental changes have led to the protection of intellectual property rights and patent rights. Industry 4.0 has also enhanced corporate performance by increasing profit through marketing activities.

A company's marketing serves as a driving force to improve performance and expand markets. Achieving standardization through continuous technology development can reduce costs and expand user services. Moreover, various activities enhancing management performance, such as a differentiation strategy, improvement of the distribution structure, and establishment of service centers, have made the game industry independent. Even though COVID-19 is causing a slowdown in the global economy, gaming remains a growing sector. The game industry is creating a virtual world that is interconnected through online access. Growing enterprises have significant success factors. By analyzing these success factors, we can identify factors that drive corporate growth. However, few studies have explored success factors in the game industry.

The purpose of this study is to explore the success factors for improving corporate management performance in Korea’s game industry. Through this, differentiated strategies and growth plans are identified that could help companies in the game industry looking for new growth engines. It is also meaningful from the fact that this research has been conducted in terms of management and strategy.

To this end, this study explores relevant factors by setting up environment, strategy, and performance models. The performance of enterprises was analyzed by setting management, technology, market, and industrial factors as environmental variables. A company's market adaptation and R&D capabilities have been parameterized.

In Chapter 1, the significance and expansion of the gaming industry are elucidated. Additionally, the criticality of innovations and market adaptability within the gaming industry is discussed, with the research objectives being delineated. Chapter 2 delves

into extant literature concerning the trends and determinants of success in the gaming industry, elucidating its linkage and differentiation from prior studies. Chapter 3 delineates the research objectives, posits hypotheses, and introduces the research framework. The methodology, encompassing data collection and analytical techniques congruent with the research aims, is also expounded. In Chapter 4, the amassed data is systematically analyzed, and the research outcomes are presented. Chapter 5 discerns the factors of success based on the analytical findings and proffers insights pertinent to innovation and market adaptability in the gaming sector. Lastly, Chapter 6 encapsulates the research's conclusions, highlighting its limitations and suggesting avenues for future inquiry.

2. Trends and Prior Studies in the Game Industry

2.1. Trends in Korea's Game Industry

Games create value in terms of economics, merchantability, and diligence through the combination of software and hardware. Gaming is a knowledge-based industry in which high value can be achieved even from a small amount of capital investment, and it is an industry that can create synergy by merging with other industries. In addition, gaming is largely divided into hardware-oriented markets (e.g., PC games, mobile games, console games, and arcade games), commodity production (e.g., software, characters, and peripherals), and service providers such as PC rooms and complex game venues.

Table 1. The Size and Prospects of Korea's Game Market (2017-2021)

(Units: 100 million won)

Division	2017		2018		2019 (E)		2020 (E)		2021 (E)	
	Sales	Growth (%)	Sales	Growth (%)	Sales	Growth (%)	Sales	Growth (%)	Sales	Growth (%)
PC Games	45,409	-2.9	50,236	10.6	51,929	3.4	53,210	2.5	52,399	-1.5
Mobile Games	62,102	43.4	66,588	7.2	70,824	6.4	72,579	2.5	76,757	5.8
Console Games	3,734	42.4	5,285	41.5	5,467	3.4	5,334	-2.4	7,042	32.0
Arcade Games	1,798	121.0	1,854	3.1	1,908	2.9	1,881	-1.4	1,992	5.9
PC Rooms	17,600	20.0	18,283	3.9	19,879	5.6	19,879	2.9	19,527	-1.8
Arcades	780	4.0	686	-12.0	691	6.5	691	-5.5	703	1.7
Total	131,423	20.6	142,902	8.7	153,575	5.1	153,575	2.3	158,421	3.2

Source: Korea Creative content Agency, WHITE PAPER ON KOREAN GAMES 2019, p.5.

Korea's game market continues to grow in console games and mobile games. In particular, most game segments showed marked growth in 2017. This indicates that the

combination of game-related hardware and content has resulted in completion of the growth engine (Table 1).

Considering the trend in international trade for Korea's game industry, the export sector continued to grow from US$2.6 billion to US$6.4 billion, but imports increased by between US$180 million and US$300 million.

China had the largest share of all trading partners for Bishop Games Studio, inc., followed by North America and Japan. What is noteworthy is that the proportion of imports and exports for the Chinese region and other regions are in contrast to each other, which is believed to be due to the different characteristics of game stories and the local game infrastructure. It also shows that game companies need to identify these market characteristics and local technical and potential needs (Table 2).

Korea's game industry has a global market share of about 30%. Among them, PC games and mobile games have a high proportion due to the development of domestic IT technology, consumer marketing, and gaming market managers' capabilities (Table 3).

Table 2. Import and Export Status of Korea's Game Industry

(Units: US$1,000)

Division		2012	2013	2014	2015	2016	2017	2018
Export	**Export Amount**	2,638,916	2,715,400	2,973,834	3,214,627	3,277,346	5,922,998	6,411,491
	Increase/ Decrease (%)	11.0	2.9	9.5	8.1	2.0	80.7	8.2
Import	**Import Amount**	179,135	172,229	165,558	177,492	147,362	262,911	305,781
	Increase/ Decrease (%)	-12.6	-3.9	-3.9	7.2	-17.0	78.4	16.3

Source: http://www.kocca.kr/cop/bbs/view/B0000146/1841389.do?menuNo=201826&delCode=0&pageIndex=1, Search May 27, 2020.

Table 3 Global Market Share for 2019

(Units: US$1 million)

Division	PC Game	Mobile Game	Console Game	Arcade Game	Total
World Game Market	32,807	63,884	48,968	32,709	178,368
Korea Game Market	4,566	6,049	480	231	11,326
Share (%)	13.9	9.5	1.0	0.7	6.3

Source: Korea Creative content Agency, WHITE PAPER ON KOREAN GAMES 2019, 2019, p.26.

2.2. Prior Studies

Game-related research has been conducted from technical aspects (e.g., IT and programs), but research into the game industry in terms of corporate performance is rare.

Park understood the game industry by linking it to content, arguing that to enhance corporate performance, it is necessary to converge family-oriented content strategically with content syndication, IT and story combination, and consumer-led content development. That study also stressed the need for market-adaptable corporate management to spread game platforms [55]. Kim predicted that the future of the game industry would emerge from genre specialization, technology monopolization, and the expansion of online games, noting that collaboration between companies, securing professional technical personnel, and marketing can determine performance in the game industry [28].

Choi et al. argued that the game industry should be fostered through value chain models. They stressed the importance of distribution through the global value chain, suggesting that growth of the game industry has increased significantly in the entertainment sector. They emphasized the need for development tailored to cultural background and market consumer characteristics in order for Korea's game companies to enter global markets. The marketing and management aspects of companies, such as Chinese consumers' preferences, distribution networks, market customs, and service management, are important if Korean companies try to enter the Chinese market [68].

Oh and Kim suggested measures to enhance corporate performance through environmental and industrial factors. They found that environmental factors such as consumer sentiment and related laws on games, market distribution structures, and industrial factors such as R&D, facilities, and marketing should be overcome. To this end, the Commission requested cooperation among businesses, the sharing of distribution networks, and government support to establish infrastructure for the game industry [54].

Factors such as corporate research and development investment and corporate performance have a causal relationship with patented technology[29, 38, 42, 56], and patent information well represents the technical ability to link corporate research and development investment, innovation activities, and corporate performance[16, 40, 49, 71]. Additionally, the realization of reality by computer technology has begun to provide a degree of reality to things like traditional card games, Go, and chess, and to activities like flying a fighter jet, firing missiles, and exploring space[22, 35, 60, 73].

Lee and Huh pointed to a need to foster the game industry through the introduction of industrial technology, and through management and administrative perspectives from an interdisciplinary point of view. They showed that various institutions and government support are needed for the development of certain industries [39]. Jung et al. argued that government policy support is important to small game companies in order to address their lack of technology. They proposed a government funding and technology evaluation system as an improvement plan, and demanded government support for the game industry, which requires continuous R&D [24].

The game industry requires not only continuous R&D investment but protection of intellectual property rights such as patents and copyright. In particular, Choi et al. showed that government support could raise R&D spending and patent registrations by firms [68].

Ayaz and Li argued that consumer preferences and user demand are indicative of R&D, and taking them into account can lead to an increase in corporate performance. This shows that R&D is a major factor in gaining a competitive advantage, helping companies grow and expand their market share [4].

Lee et al. looked at R&D activities based on the size of the enterprise. Their findings indicated that the larger the company and the higher the sales, the more likely they are to engage in R&D activities and secure property rights. This shows that expanding the size of game companies and/or collaboration among them is a way to secure competitiveness [40].

Koo stated that when firms are willing to spend on R&D and when internal capabilities are well-equipped, if technology procurement is internationalized, then corporate performance is positive. In addition, the characteristics of corporate managers and overseas market activities have a positive impact on R&D performance, and overseas collaboration and marketing have a positive impact on corporate performance [37].

Liu and Kwon explored the difference between the content business and the entertainment business in terms of corporate performance. Because the nature of knowledge is strong in the content business, the willingness and management strategies of corporate managers are important, and in the entertainment business, the improvement of R&D and market adaptation is more likely to enhance corporate performance [45]. This encourages relatively small businesses to expect aggregation through M&A for qualitative development. They also proposed multi-use management through the establishment of a consumer-oriented game network and distribution platform, rather than a supplier-oriented management method.

The investigation into the proportion of patent value to a country's total research and development investment has verified that factors such as corporate research and development investment and corporate performance have a causal relationship with patented technology. Technological innovation often utilizes patent data to measure the direction of spillover effects, and the spillover effects of technological innovation include the social benefits from ideas or information resulted from research and development investment and the non-competitive goods affecting other research[2, 8, 15, 29, 41, 45].

Choi et al. claimed that the establishment of a platform for item trading through an analysis of the game market affects the performance of game companies. They stressed the need to develop a transaction-based platform based on a Chinese market analysis, which should lead to market-oriented corporate management, including consumer-oriented marketing strategies and market distribution [68].

Goyal pointed out that the world's top companies have read the future of the game content industry and have invested in R&D. In addition, online payment can improve the game industry's performance, and online payment systems need to be overhauled through R&D [71]. Choi et al. called for technology development to improve performance, referring to managers' abilities to apply new technologies such as mobile payment platforms and to adapt to market trends in corporate competitiveness [68].

3. Hypotheses

3.1. Managers

The internal environment of a company is a controllable area. In particular, various studies have identified the ability to enhance corporate competitiveness through human resource management [4, 18, 24, 36, 44, 49, 50]. In addition, corporate managers' global interests and capabilities lead to government support, and affect adaptations to local markets and R&D [40, 64].

In terms of a strategy for enhancing a company's performance, manager-related characteristics have an impact [2, 6, 10, 11, 16], and empirical studies have shown that management's characteristics have a significant impact on innovation activities [32, 50, 75].

Therefore, the experience, attitudes, and know-how of game company managers influence R&D activities when adapting to local markets using marketing to strengthen market share.

H 1-1 Manager competence will have a positive effect on market adaptation.

H 1-2 Manager competence will have a positive effect on R&D.

3.2. Technology

Technology in the game industry is a very important means of enhancing competitiveness. A corporate entity may have technical capabilities by developing technology on its own or by purchasing it. In particular, companies with professional resources have a high R&D ratio, and property rights are actively protected [14, 20, 21, 42, 47, 67].

In particular, a company is able to drive changes in the market by advancing the industry with improved technology and by improving infrastructure, securing a competitive advantage through its value chain [37, 41, 70].

Therefore, the technology and the technical professionals of game companies can pursue market changes, strengthen market adaptation, and influence R&D activities to improve performance.

H 2-1 Technology factors will have a positive effect on market adaptation.

H 2-2 Technology factors will have a positive effect on R&D.

3.3. Markets

The game industry has different technology levels and growth rates depending on the size of the market. Domestic and overseas markets differ in size, sales, and consumer preferences [62, 74]. Markets with high-income consumers are well-equipped with laws and systems, and respond quickly to technical demand [3, 49, 60].

Additionally, consumer preferences increase the demand for items with a related technology [5, 11, 52, 72, 63].

Strategic choices and the necessary R&D activities will vary depending on market factors such as when products are released, product levels, and customer satisfaction. In other words, game companies should implement various forms of marketing according to consumer demand [15, 46, 51].

Therefore, companies adapt to the market according to positive market conditions, such as game recognition, institutional devices, and the level of market competition, strengthening the R&D capabilities needed.

H 3-1 Market factors will have a positive effect on market adaptation.

H 3-2 Market factors will have a positive effect on R&D.

3.4. The Industry

The game industry consists of small and medium-sized enterprises engaged in various activities such as planning, development, storytelling, and distribution. This shows that industry growth can bring about corporate growth. Recognition from the industry is particularly important in the early stages of products offering new technologies to meet consumer demand [27, 29, 30, 44]. However, due to consumer loyalty and market infrastructure in the growth phase of a product, there is a strong tendency to make conservative choices rather than novel ones [9, 15, 73].

Regulations and support for R&D and marketing activities vary depending on the industry [33, 36, 57]. In industries where management resources can easily be combined, the phenomenon of a shared economy through strategic networks and synergies through marketing and R&D activities can be expected [22, 48].

Increasing performance in the gaming industry requires a consumer technology and platform that integrates tightly with time of product release onto the market [12, 69].

Therefore, the industrial environment, such as distribution, government support, and market entry barriers, has a positive impact on a company's market adaptation and R&D.

H 4-1 Industrial factors will have a positive effect on market adaptation.

H 4-2 Industrial factors will have a positive effect on R&D.

3.5. Adaptation and R&D

The ability to execute marketing that is tailored to local consumers and intended to increase demand has a significant impact on corporate performance [41, 58, 56]. Performance improvement through government support and market systems [7, 8, 26, 34, 35, 44, 49, 53, 58, 60, 61] along with active funding and technology evaluation systems in the market enable performance improvement beyond the company's scale constraints [27, 29].

Game companies can seek continued market competition and market leadership by investing in R&D, which has a positive impact on corporate performance by securing intellectual property such as technological innovations and patent rights [25, 33, 36, 57, 59].

Therefore, market adaptation and R&D, such as enterprise marketing activities and consumer preferences, have a positive impact on management performance.

H 5-1 Market adaptation will have a positive effect on management performance.

H 5-2 R&D will have a positive effect on management performance.

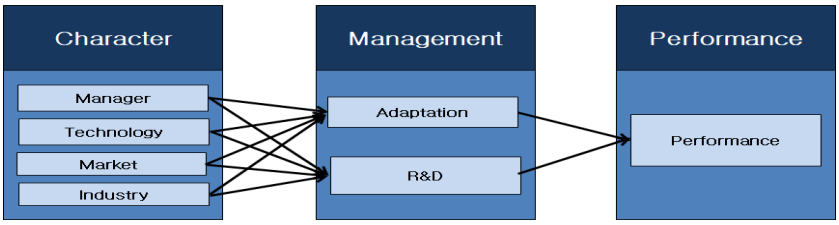


Fig. 1. The research model

3.6. Data Collection

This study surveyed companies in Korea's game industry. The method of selecting the companies to be surveyed utilized the list of companies registered in the Game Marketing Forum (an Internet gathering of the Korea Game Industry Promotion Agency, the Game Developers Council, game company marketing companies, and distributors).

The data included interviews with the person in charge, plus e-mailed and direct surveys of game companies that joined the game association.

From May 25 to August 25, 2019, 900 copies of a questionnaire were distributed via e-mail and given offline via interpersonal interviews and group interviews (Table 4). After 350 responses were collected (a response rate of 38.9%), 336 were used for the study, excluding 14 that were incomplete.

Table 4. Sample Aggregation of the Questionnaire

Division	Online		Offline		Total		
Companies	Targeted	Answered	Targeted	Answered	Targeted	Answered	Response Rate
	150	150	750	200	900	350	38.9%

4. Research Results

4.1. Characteristics of the Sample

1) Major activities of game companies

The scope of the development, distribution and service offerings, and planning activities cited by game companies were 37.8%, 25.9%, and 26.8%, respectively (Table 5). These percentages can be attributed to the fact that the growth cycle of Korea's game industry spans the product development period to the maturity period, with major tasks performed in each cycle.

Table 5. Major activities by game companies

Activity	Count	Percentage
Planning	90	26.8
Development	127	37.8
Distribution and Service Offerings	87	25.9
Marketing	23	6.8
Management	9	2.7
Total	336	100.0

2) Major platforms of game companies

The main platform for Korea's game companies is online, accounting for 128 out of 336 companies (38.1%), followed by PC games (80 of 336, or 23.7%) (Table 6). Other respondents did not have one clear platform, but are engaged in the game industry for the dispatch of human resources, as development agencies, and as distribution companies. The recent growth of online and mobile games has led to an increase in R&D for many game companies.

Table 6. Game Company Platforms

Platform	Count	Percentage
Arcades	24	7.2
PCs	80	23.7
Online	128	38.1
Video	7	2.1
Mobile Devices	66	19.6
Other	31	9.3
Total	336	100.0

3) Number of employees

Regarding the number of employees, 153 companies (45.6%) had between 11 and 50 employees, followed by 93 (27.8%) with fewer than 10 (Table 7). Most of the game companies operate as small and medium-sized companies and venture companies, resulting in a shortage of professionals in areas such as R&D and marketing. These results indicate the need for government policies and for fostering professional workers who can work in the gaming industry.

4) Import and export values over three years

Looking at the average annual import and export values over the previous three years, 234 companies (69.6%) earned under US$1 million, revealing the small scale of operations for many game companies (Table 8). However, in interviews with the people in charge, the reason given for trade volumes slowly increasing every year was that they are interested in overseas markets. Moreover, despite a lack of information on overseas

sites and poor marketing capabilities, competitiveness in IT-based technologies is potentially playing a role in enhancing the competitiveness of games by Korean companies in overseas markets.

Table 7. Company Employees

Range	Number of Employees	Percentage
Under 10	93	27.8
11 to 50	153	45.6
51 to 100	38	11.4
101 to 300	26	7.6
More than 300	26	7.6
Total	336	100.0

Table 8. Average Annual Import and Export Values over Three Years

Value (US$)	Number of Companies	Percentage
Under 500,000	119	35.4
510,000 to 1,010,000	115	34.2
1,010,000 to 2,500,000	64	19
2,510,000 to 5,000,000	8	2.5
5,010,000 to 10,000,000	17	5.1
More than 10,010,000	13	3.8
Total	336	100.0

5) Major export areas

China and Southeast Asia were the major export destinations for 80 companies (23.7%) and 58 companies (17.3%), respectively. Other export markets include South America and Central Asia (Table 9).

Table 9. Major Export Area

Export Destination	Count	Percentage
America	46	13.6
EU	43	12.7
Japan	49	14.5
China	80	23.7
Southeast Asia	58	17.3
Middle East, Africa	15	4.6
Other	46	13.6
Total	336	100.0

4.2. Validity and Reliability Analysis

In this study, the reliability of variables constituting each factor was tested using Cronbach's alpha, the most common method for reliability analysis. Analysis results

exceeded the threshold of 0.7 or higher (Table 10). In general, questionnaire analysis acknowledges that a confidence coefficient of 0.70 or higher is relatively high.

Table 10. Reliability Coefficients of Variables

Variable	Measurement	Start	Erase	Use	Cronbach's Alpha
Manager	Executive experience and capability	6	3	3	.773
	Management attitude				
Technology	Innovation in technology	6	3	3	.876
	Technical mimicry potential				
	R&D personnel				
Market	Degree of market competition	10	7	3	.708
	Institutional protection				
	Game recognition				
Industry	Economic level	10	7	3	.729
	GDP				
	Game Industry Growth				
	Product Life Cycle				
	Network				
Adaptation	Marketing	6	3	3	.928
	Service				
	Platform				
R&D	R&D	5	3	2	.829
	Human Resources				
Performance	Profit Amount	3	0	3	.856
	Sales Profit				
	Export Profit				

Table 11. Factor Analysis

Measured Items		Components			
		Manager	Techno-logy	Market	Industry
X1	Management's ability to develop products and services	.817			
X2	Cognition of products and services by managers	.814			
X3	Professional competence of a manager	.758			
X7	High cooperation with relevant departments		.785		
X8	Standardized products of technological superiority		.774		
X9	Main axis of products with high differentiation		.605		
X15	Overseas market larger than domestic market			.761	
X18	Help from government-related research institutes			.881	
X20	Timely product supply			.729	
X25	Growth of the game industry				.843
X26	The higher the GDP, the higher the adaptation				.799
X27	The higher the GDP, the higher the R&D				.706
Characteristic		22.644	5.287	5.865	9.444
Total sample dispersion ratio		9.921	7.104	7.609	9.618
Cronbach's alpha		.773	.876	.708	.729
KMO		.713			
Bartlett's test		Chi-Square=1246.946, df=496			
Significance probability		.000			

** Value of the variable with the largest amount of factor load, significance level =0.05

Table 12. Correlation by Factor

Classify	A	B	C	D	E	F	G
Manager (A)	1						
Technology (B)	.105(**)	1					
Market (C)	.392	.006(*)	1				
Industry (D)	.068(*)	.541	.462	1			
Adaptation (E)	.008(**)	.006(**)	.036(**)	.545	1		
R&D (F)	-0.732	.054(*)	.598	.002(**)	.635	1	
Performance (G)	.196	.050(*)	.004(**)	.694	.013(**)	.005(*)	1
**(Significance at the 0.01 level), *(Significance at the 0.05 level)							

Unnecessary factors were eliminated, and factors were extracted through factor analysis. Eigenvalues of 1.000 or less were excluded. In exploratory factor analysis, the principal component method was used, and factor rotation ensured interdependence between the factors using the varimax orthogonal rotation method. The factor analysis results showed that the Kaiser-Mayer-Olkin (KMO) measure of sample adequacy (MSA) was 0.713 > α=0.5; chi-square in Bartlett's test was 1246.946, and the significance probability was 0.000 < <=0.05. The cumulative distribution of the four factors accounted for 43.24% of the total data (Table 11).

Correlation analysis between variables provides an overview of the relationships between variables introduced in the study, and predicts the results from verification of an established hypothesis. Correlation values are used to interpret the analysis, and it is common to assume the following: 1.0 to 0.7 (very relevant), 0.7 to 0.4 (significant), 0.4 to 0.2 (slightly relevant), and 0.2 to 0.0 (irrelevant). Correlation analysis results are shown in Table 12.

Conformity assessment of the study model is a procedure to examine how well the covariance structural model fits the hypotheses in the study (Table 13).

Table 13. Conformity Assessment Index

Classification	Model Conformity Assessment Index		Result
Absolute Conformity Index	χ^2	Chi-square (degree of freedom)	44.462 (39df)
	P	Significance probability	*.101 ≥ .05*
	Q	Chi-squared/degree-of-freedom ratio ≤ 3	1.01 3
	GFI	Goodness of Fit Index ≥ 0.9	.942 ≥ .9
	AGFI	Adjusted GFI ≥ 0.9	.793 ≤ .9
	RMR	Root Mean Square Residual	.029 ≤ .05
	RMSEA	Root Mean Square Error of Approximation	.04 ≤ .05
Incremental Conformity Index	NFI	Normed Fit Index ≥ 0.9	.992 ≥ .9
	RFI	Relative Fit Index ≥ 0.9	.805 ≤ .9
	CFI	Comparative Fit Index ≥ 0.9	.952 ≥ .9
Simplicity Conformity Index	PNFI	Parsimonious Normed-of-Fit Index	.593

4.3. Route Analysis Results

In this study, the results of structural equation modeling used to test the hypotheses are shown in Table 14.

First, the hypothesis that manager factors have a positive effect on market adaptation was supported, but an effect on R&D was not. According to this study, manager confidence in the company is a significant factor in both information technology and relationships [1, 52, 53]. However, this study found that market adaptation linked to relationships was supported, but manager factors did not show any effect on R&D [63,72].

Second, the hypotheses that technology factors have a positive effect on market adaptation and R&D were supported. This is consistent with prior studies [21, 37, 41]. In other words, a well-equipped entity achieves effective performance, and enhances performance through market adaptation and R&D. Therefore, in the game industry, it is very important to enhance the technical competence of the enterprises.

Third, the hypothesis that market factors have a positive effect on market adaptation was supported. This is consistent with a study that showed changes in market demand require rapid responses [63, 12, 72]. However, the hypothesis that market factors have a positive effect on R&D was not supported. This hypothesis did not match prior studies, which is believed to be due to negative factors such as technology imitation in the game industry, or unauthorized use of patents [49, 60].

Fourth, the hypothesis that industrial factors have a positive effect on market adaptation was not supported. This hypothesis is not consistent with prior studies [9, 28, 73] and perhaps it is because it is difficult to drive the flow of the market for companies that have items that are pioneering new markets. However, the hypothesis about them having a positive effect on R&D was supported [12, 64, 69].

Fifth, the hypothesis that market adaptation factors have a positive effect on corporate performance was established as consistent with prior studies [5, 26, 27, 29]. Also, the hypothesis that R&D factors have a positive impact on corporate performance was supported [33, 36]. Therefore, to enhance corporate performance, it is necessary to continuously strengthen R&D and adapt to markets. The route analysis results are as shown in Table 14.

Table 14. Route Analysis Results

Hypothesis	Path	Path coefficient	Standard error	t	p	Result
H 1-1	Manager→Adaptation	.231	.238	1.942	.001**	Accepted
H 1-2	Manager→R&D	-.257	.241	-2.008	.118	Rejected
H 2-1	Technology→Adaptation	.289	.148	.902	.009*	Accepted
H 2-2	Technology→R&D	.651	.232	2.721	.007*	Accepted
H 3-1	Market→Adaptation	.245	.431	2.541	.002	Accepted
H 3-2	Market→R&D	-.191	.435	-.491	.515	Rejected
H 4-1	Industry→Adaptation	.101	.145	.254	.581	Rejected
H 4-2	Industry→R&D	.513	.269	3.375	.003*	Accepted
H 5-1	Adaptation→Performance	.393	.171	2.571	.001**	Accepted
H 5-2	R&D→Performance	.338	.145	2.581	.001**	Accepted

**Significance at the 0.01 level, *Significance at the 0.05 level

5. Implications

This study looked at Korea's game companies to determine factors that affect a firm's performance in the game industry. To that end, internal and external factors of the enterprises were identified, and empirical analysis was performed using market adaptation and R&D capabilities as parameters. The analysis results are as follows.

First, the experience of corporate managers, their management know-how, and attitudes toward the introduction of external technologies showed significant impacts on R&D. In addition, in the game industry, where creative perspectives and timing are important, the subjective will of managers is an obstacle to R&D and market adaptation.

Second, a company's discriminatory technology capabilities showed significant effects on sustained R&D and market adaptation. However, imitation by latecomers and the lack of corporate size and technical expertise were shown to be obstacles to R&D and market adaptation.

Third, product awareness and time of release onto the market have a significant impact on market adaptation, but were shown to be a barrier to R&D. This is because the game market attracts consumer choices through marketing, rather than technology and creative approaches.

Fourth, the nature of the game industry has a significant impact on R&D, but not on market adaptation. This means that technology changes are required to meet environmental characteristics such as consumer demand and game environment infrastructure, but such characteristics are somewhat too much to lead market changes.

Fifth, R&D and market adaptation by enterprises have a significant impact on performance. This shows that companies improve corporate profits and secure market stability by strengthening product competitiveness through R&D and from consumer marketing through market adaptation.

At the same time, a negative perception about copying technology has emerged in the game industry. To solve this problem, strengthening intellectual property rights to prevent the theft or copying of creative ideas, plus indirect support through intergovernmental negotiations, is required when exporting to underdeveloped countries.

R&D and property rights management vary depending on the size of the enterprise. Based on the results of this research, the following measures are proposed to maintain the competitiveness of game companies.

First, steady support for R&D is needed to enhance corporate performance. R&D should be handled as a corporate policy, not as changes in R&D budgets and support only at the discretion of managers. In other words, securing R&D competitiveness should be prioritized in budgeting and policy decisions.

Second, in order to maintain and develop technology, qualitative management through the recruitment of professionals and performance-linked incentives are required. If the company is large, it is necessary to set up and operate a dedicated department. However, if an entity is small, it is necessary to establish inter-enterprise cooperation or clusters.

Third, developmental imitation, not simple imitation, can reduce R&D costs. It is necessary to identify ongoing technology and market trends, and to strengthen mutual cooperation through cross-licensing if necessary.

Fourth, it is necessary to utilize R&D capabilities in companies as a key strategic objective. Marketing should be carried out in a technology-driven market, and policies

should be formulated to protect property rights in cooperation with government. If necessary, market dominance should be secured through M&A and clusters.

Finally, it is necessary to seek market access and expansion to meet the life cycle of the game industry. Each country has a different game environment and infrastructure, so there is no need to pursue fast R&D. Stable management performance and enhancement of enterprises can be secured in various markets.

With online growth and technological advances, the game industry is globalizing. Games are no longer a mere tool of amusement but a tool of learning. For development in the game industry, it is necessary to consider transformation of developer awareness, standardization of technology, and development of links with other industries. The game industry can be a new growth engine driving a country's economic growth.

6. Conclusion

The game industry is becoming a new growth engine in the Industry 4.0 paradigm. In other words, the game industry requires continuous management and investment, including identifying market trends, R&D, and monitoring of foreign technologies. This study explored success factors that can enhance market performance among Korea's game companies. The implications of the empirical analysis are as follows.

First, managers should strengthen their capabilities and pursue cooperation with other companies. The game industry requires collaboration to enhance performance in technology development, marketing, and services. Administrators need to invest more in ongoing collaborative networks to reflect the nature of the enterprise and achieve its goals.

Second, R&D sharing through clustering is required because it differentiates technology according to the size of the game company. It is necessary to build a cluster that can have a significant impact on market performance, such as the retention of professionals and capital liquidity.

Third, since marketing is deeply related to customer service, it needs to be sensitive to changes in the game market environment. In addition, adequate market adaptation is necessary for new game environments such as video, arcades, PCs, and mobile devices. Consequently, it can lead to the release and distribution of games and to the expansion of game-related items, thereby enhancing corporate performance.

Fourth, it is necessary to grow gaming into a strategic industry through government support and policy development. The game industry can be fostered through policies such as R&D support, funding for the distribution of games, and by protecting property and patent rights.

Consequently, the gaming industry is emerging as a central sector in the Industry 4.0 paradigm, necessitating sustained management, investment, and vigilance towards global technological trends. This research delineated pivotal success factors for enhancing market performance among Korean gaming corporations. Noteworthy findings advocate for managers to augment their competencies and seek collaboration with external entities, underscoring the significance of R&D sharing and clustering contingent upon the firm's size. Additionally, adaptability to market shifts across diverse gaming

platforms is imperative. Ultimately, governmental support and policy initiatives are crucial for the strategic advancement of the industry.

However, this study did not deal with administrative procedures such as obtaining intellectual property rights or protecting patent rights. Also, analysis of individual items in terms of the effects of R&D investment and consumer awareness was insufficient. Besides, the number of game companies in the sample is relatively small, making it less valid to generalize these research findings. Therefore, in follow-up studies, we want to supplement the humanities approach that companies and consumers create together, rather taking than a technical approach, and intend to increase the number of sample companies to conduct in-depth industry-specific research on game companies and games. In addition, a comparative study of R&D and intellectual property management strategies in the game industry is necessary.

References

1. Abdullah, M. F., Khan, N. R. M., & Ibrahim, M. A.: Exploring the Influence of SERVQUAL Dimensions of Reliability, Responsiveness and Assurance towards Consumers Loyalty: The Mediating Effect of Commitment-Trust Relationship Marketing Theory. International Journal of Academic Research in Business and Social Sciences, Vol. 12,No. 11, 1580 -1591. (2022)
2. Akcay, D. D. S.: Causality Relationship between Total R&D Investment and Economic Growth. The Journal of Faculty of Economics and Administrative Sciences, Vol. 16, No. 1, 79-92. (2011)
3. Anderson, C., Narus, A.: A model of distributor firm and manufacturer firm working partnerships. Journal of Marketing, Vol. 54, No. 1, 48-58. (1990)
4. Ayaz Ahmed, Jia Li: Effect of In-App Purchases on Consumer Purchasing Behavior in Mobile Games. Journal of Business Research, Vol. 117, No. 9, 222-235. (2020)
5. Baek Yeong-ki: Competiton, Collaboration and Innovation Networks in Regional Economic Development: the Case of Chonbuk. Journal of the Economic Geographical Society of Korea, Vol. 9, No. 3, 459-472. (2006)
6. Barney, J. B.: Firm Resources and Sustained Competitive Advantage. Journal of management, Vol. 17, No. 1, (1991)
7. Bravo-Ortega, Claudio and A.G. Marin: R&D and Productivity: A Two Way Avenue?. World Development, Vol. 39, No. 7, 1090-1107. (2011)
8. Caillaud, Bernard, Jullien, Bruno: Chicken and Egg Competition among intermediation service providers. Rand Journal of Economics. Vol. 34, No. 2, 309-328. (2003)
9. Camargo P., Piggin J., Mezzadri F.: The politics of sport funding in Brazil: a multiple streams analysis. International Journal of Sport Policy and Politics, Vol. 12, No.4, 599-615. (2020)
10. Chang, C. H., Lee, K. H., Noh, K. S.: A Study on Comparative Analysis for Competitiveness of Success Factors of the Platform Business. Journal of Digital Convergence, Vol. 14, No. 3, 243-250. (2016.)
11. Choi Jinah, Kim Dongweon: Global Marketing Strategies of Online Game Companies: The Case of Korean Companies' Localization in Japan. International Business Reviews, Vol. 15, No. 3, 175-200. (2011)
12. Choi Sung: A Study of IT competitiveness of SMEs by Cloud Services. Journal of Digital Convergence, Vol. 11, No. 3, 59-71. (2013)

13. Choi, Byong-Sam, Kim Joo Han: A Study of the Effect of the General Definition of Platforms on the Firm's Economic and Strategy Decision-Making. The Journal of Business Education, Vol. 25, No. 3, 157-176. (2011)
14. Cockburn, I., and Griliches, Z.: Industry Effects and Appropriability Measures in the Stock Market's Valuation of R&D and Patents. The American Economic Review, Vol. 78, No. 2, 419-423, (1988)
15. Covin, J. G., & Slevin, D. P.: A Conceptual Mode of entrepreneurship as Firm Behavior. Entrepreneurship theory and Practice, Vol. 16, No. 1, 7-26. (1991)
16. Daradkeh, M.: Exploring the Boundaries of Success: A Literature Review and Research Agenda on Resource, Complementary, and Ecological Boundaries in Digital Platform Business Model Innovation. Informatics, Vol. 10, No. 2, 1-30. (2023)
17. Freighthub, Available online: https://freighthub.com/en/ (accessed on 17 June 2019).
18. Freightos, Available online: https://www.freightos.com/ (accessed on 16 June 2019).
19. Griliches, Z. and F. Lichtenberg: Interindustry Technology Flows and Productivity Growth: A Re-examination. The Review of Economics and Statistics, Vol. 66, No. 2, 324-29. (1984)
20. Griliches, Z.: Productivity, R&D, and Basic Research at the Firm Level in the 1970s. American Economic Review, Vol. 76, No. 1, 141-154. (1986)
21. Griliches, Z.: The Discovery of the Residual: A Historical Note. Journal of the Economic Literature, Vol. 34, No. 3, 1324-1330. (1996)
22. Guichard, L., Stepanok, I.: International Trade, Intellectual Property Rights and the (Un)employment of Migrants. The World Economy, Vol. 46, No. 7, 1940-1966. (2023)
23. Hambrick, D. C. and Phylls , Manson: Upper Echelons: The Organization as a Reflection of Its Top Managers. Academy of Management Review, Vol. 9, No. 2, 193-206. (1984)
24. Hyunseung Jung, Kiyoon Kim, Daiwon Hyun: Analysis of Priorities of Policy Implementation Tasks for Revitalizing Virtual Reality(VR) and Augmented Reality(AR) Industries. JOURNAL OF THE KOREA CONTENTS ASSOCIATION, Vol. 21, No. 9, 12-23. (2021)
25. Jin, Dongsu: Exploratory Research on Success & Failure of Platform business. International Commerce and Information Review, Vol. 15, No. 2, 387-410. (2013)
26. Joo, Hyun Woo: A Study on the Strategic Development Model of the Logistics Platform. Master's Thesis, Chung-Ang University, Seoul, (2018)
27. Kim Changwook, Lee Sangkyu: Fragmented industrial structure and fragmented resistance in Korea's digital game industry. Television & New Media, Vol. 5, No. 4, 354-371. (2020)
28. Kim, J.: Meta-verse platforms and content business trends in domestic and international markets. KOREA COMMUNICATIONS AGENCY, Media Issue & Trend, Vol. 45, 32-42. (2021)
29. Kim Mi-na: A Study on the institutional delay and the path dependent change of the game industrial policy in korea. The Korea Association for Policy Studies, Vol. 12, No. 3, 143-170. (2003)
30. Kim Min Kyu: Reflection and Suggestions on the Effects of Game Culture Policy. Journal of Korea Game Society, Vol. 18, No. 6, 95-110. (2018)
31. Kim Sun-nam, Kang Kyong Sik: Empirical study on the acceptance intention of online service platform - Focused on international logistics. Journal of Korea Safety Management & Science, Vol. 18, No. 2, 101-107. (2016)
32. Kim Sun-Nam: An Empirical Study on the Important Factors Increasing the Acceptance Intention of Online International Logistics Platform. Doctoral Dissertation, Myongji University, Seoul, (2016)
33. Kim Yeon Jeong: The Inter Industrial Competency Analysis of Game Industry and Character Industry. Journal of Korea Technology Innovation Society, Vol. 16, No. 4, 1187-1204. (2013)

34. Kim Yunkyung: A Study on the Current Status and Forecast on Chinese Game in Relation to Game Platform-Focus on SNG. Journal of Korea Game Society, Vol. 10, No. 2, 81-88. (2010)
35. Kim, Mie-Jung and Thunt, Htut-Oo: An Analysis of Export Competitiveness in Myanmar: Measuring Revealed Comparative Advantage. Journal of International Trade & Commerce, Vol. 13, No. 2, 149-172. (2017)
36. Kim, Sung-Chul: Bilateral Trade Intensity Analysis and Implications on Korean Computer Industry in U.S. Market. Journal of Industrial Economics and Business, Vol. 28, No. 5, 2087-2104. (2015)
37. Koo Hoon Young: The Role of Cooperative R&D and Intangible Assets in Innovation and Corporate Performance of R&D Investment in Manufacturing Sectors. Journal of Society of Korea Industrial and Systems Engineering, Vol. 43, No. 1, 79-86. (2020)
38. Korea Creative content Agency, WHITE PAPER ON KOREAN GAMES 2019, (2019)
39. Kyoung-Mi Lee, Hoon Huh: An Empirical Study on the Entry of the Korean Game Industry into Emerging Markets and Strategy - Focusing on the Gravity Model –. Journal of the Korea Entertainment Industry Association, Vol. 16, No. 8, 165-175. (2022)
40. Lee Jongho, Kim Tae Hwan, Jung Woo-jin: R&D Investment Effect through Patent on IT firms using Panel Structural Equations. Knowledge Management Research, Vol. 21, No. 1, 137-150. (2020)
41. Lee Seung-Hee, "Nori" Culture in the Age of the Fourth Industrial Revolution ― Focusing on Korea-China Game Industry, The Association Of Chinese Language, Literature And Translation In Korea, Vol. 41, 217-239, (2017)
42. Legris, P., Ingham, J., Collerette, P.: Why do people use information technology? A critical review of the technology acceptance model. Information & Management, Vol. 40, No. 3, 191-204. (2003)
43. Lim Jun-Hyeong: Export Competitiveness among Korea, China and Japan in the Electronic Integrated Circuits Industry under the HS. Journal of Korea Research Society for Customs, Vol. 12, No. 3, 133-149. (2011)
44. Lim Jun-Hyung: Competitiveness Comparison between Korea and China in the Household Electronic Appliances Industry. Journal of Industrial Economics and Business, Vol. 22, No. 2, 905-918. (2009)
45. Liu Yu, Kwon Sang Jib: Entertainment Contents Corporation Tencent's Growth Strategy : Focusing on Imitative Innovation and M&A. Journal of Korea Entertainment Industry Association, Vol. 14, No. 3, 1-13. (2020)
46. Mamuneas, T. P.: Spillovers from Publicly Financed R&D Capital in High-Tech Industries. International Journal of Industrial Organization, Vol. 17, No. 2, 215-239. (1999)
47. Mansfield, Edwin: Basic Research and Productivity Increase in Manufacturing. American Economic review, Vol. 70, No. 5, 863-873. (1980)
48. Maskus, K. E., & M. Penubarti, "How Trade -related Are Intellectual Property Rights?", Journal of International Economics, (1995)
49. Maskus, Keith E. & Guifang Yang: Intellectual Property Rights, Foreign Direct Investment, and Competition Issues in Developing Countries. International Journal of Technology Management, Vol. 19, 1-18. (1999)
50. Mayer, Roger C., Davis, James H., Schoorman, F. David.: An Integrative Model of Organizational Trust. The Academy of Management Review, Vol. 20, No. 3, 709-734, (1995)
51. Mitja Ruzzier, E. Douglas, Maja Konečnik Ruzzier, Jana Hojnik: International Entrepreneurial Orientation and the Intention to Internationalize. Sustainability, Vol. 12, No. 14, 1-20. (2020)
52. Morgan, R.M. and Hunt, S.D.: The commitment-trust theory of relationship marketing. Journal of Marketing, Vol. 58, No. 3, 20–38. (1994)

53. O'Connor, P.: User-generated content and travel: A case study on tripadvisor.com. Inf. Commun. Technol. Tour., 47-58. (2008)
54. Oh, Sun Jung, Kim, Taeyoung: Analyzing the employment impact of the Korean Wave support policies using the Contingent Valuation Method. Journal of Korea Culture Industry, Vol. 21, No. 1, 51-59. (2021)
55. Park Jong-Sam: A Study on the K Contents Industry.Journal of Korea Culture Industry. Vol. 21, No. 3, 193-200. (2021)
56. Pankaj M. Madhani: Deploying a "Good Jobs" Strategy in Service Sectors for Enhancing Competitive Advantage. International Journal of Business Strategy and Automation, Vol. 2, No. 1, 29-53. (2021)
57. Petri Kettunen, Janne Järvinen, Tommi Mikkonen, Tomi Männistö: Energizing collaborative industry-academia learning: a present case and future visions. European Journal of Futures Research, Vol. 10, No. 8, 1-16. (2022)
58. Pfeffer, J.: Competite Advantage through People, Boston. Harvard Business School Press, (1994)
59. Pontiggia A., Virili, F.: Network effects in technology acceptance: laboratory experimental evidence. International Journal of Information Management, Vol. 30, No. 1, 68-77. (2010)
60. Quan Dong, Juan Carlos Bárcena-Ruiz, María Begoña Garzón: Intellectual property rights and North-South trade: Exports vs. foreign direct investment. Estudios de Economía, Vol. 49, No. 2, 145-160. (2022)
61. Rapp, Richard T., Richard P. Rozek: Benefits and Costs of Intellectual Property Protection in Developing Countries. Journal of World Trade Vol. 24, No. 5, 75-102. (1990)
62. Rosenberg, Larry J., Czepiel, John A.: A marketing approach for customer retention. Journal of Consumer Marketing, Vol. 1, No. 2, 45-51, (1984)
63. Shin Kwangyong, Zhongtian Shen, Haekun Shin, Shiyao Zhang, Ke Chen, & Lijie Li: The Mechanism of How Integrated Marketing Communications Influence on the Chinese Online Customer's Repurchase Intention. Science Journal of Business and Management, Vol. 9, No. 1, 26-38. (2021)
64. Shim Sang-min, Content Industry - New Trends in Content Business and Response Strategies. Digital Contents, Vol. 113, 114-121. (2002)
65. Smith, B., Barclay, W.: The effects of organizational differences and trust on the effectiveness of selling partner relationships. The Journal of Marketing, Vol. 61, No. 1, 3-21. (1997)
66. Snow, C. C., L. G. Hrebiniak: Measuring Oraganizational Strategies, & Organizational Performance. Administrative Science Quarterly Vol. 25, No. 2, (1980)
67. Sorce J, Issa, R.: Extended Technology Acceptance Model (TAM) for adoption of Information and Communications Technology (ICT) in the US Construction Industry. ITcon, Vol. 26, 227-248. (2021)
68. Sung-Wone Choi, Sung-Mok Lee, Joong-Eon Koh, Hyun-Ji Kim, Jeong-Soo Kim,: A Study on the elements of business model innovation of non-fungible token blockchain game : based on 'PlayDapp' case, an in-game digital asset distribution platform. Journal of Korea Game Society, Vol. 21, No. 2, 123-137. (2021)
69. Tamzil, F., Anwar, N., & Hadi, M. A.: Security Utilization of Cloud Computing in The World of Business For Small Medium Enterprises (SMEs). International Journal of Science, Technology & Management, Vol. 3, No. 1, 41-49. (2022)
70. Venkatesh, Viswanath, Morris, Michael G., Davis, Gordon B., Davis, Fred D.: User Acceptance of Information Technology: Toward a Unified View. MIS Quarterly, Vol. 27, No. 3, 425-478, (2003)
71. Vishal Goyal: The COVID-19 Pandemic and Digital Gaming: A Boon or a Bane?. International Journal of Advanced Research in Computer Science and Software Engineering, Vol. 10, No. 10, 286-290. (2020)

72. Wen, C., Prybutok, V.R. and Xu, C.: An integrated model for customer online repurchase intention. J. Comput. Inf. Syst. Vol. 52, No.1, 14–23. (2011)
73. Yoo Sang-Keon, Kim Yong-Eun, Won-Jae Seo: Sports Celebrities as a Determinant of Sport Media Distribution Contents: Focusing on Tacit Premise of Agenda Setting Theory. Journal of Distribution Science, Vol. 17, No. 10, 83-91. (2019)
74. Yang Q, Hayat N, Al Mamun A, Makhbul ZKM, Zainol NR: Sustainable customer retention through social media marketing activities using hybrid SEM-neural network approach. PLoS ONE, Vol. 17, No. 3, 1-23. (2022)
75. Zheng Qiumei, Li Chenglong, Bai Shizhen: Evaluating the couriers' experiences of logistics platform: The extension of expectation confirmation model and technology acceptance model. Frontiers in Psychology, Vol. 13, 1-20. (2022)

Jun-Ho Lee works at a Faculty of Division of Public Affairs and Police Administration at Dongguk University- WISE. He received his Ph. D. in Public Management from Renmin University of China. His current research interests include: Industrial Clusters, Strategic Alliances between Companies.

Jae-Kyu Lee is a Researcher of the Smart Contents Institute at Dongguk University-WISE. He received his Ph. D. in Engineering, specializing in Information Technology from Dongguk University-WISE in 2024. His research interests include: Virtual Reality, Artificial Intelligence, Educational Technology, Realistic Media, and IT Convergence Software.

Seung-Gyun Yoo (Correspondence*) works at a Faculty of Department of International Trade and Airline Service at Dongguk University-WISE. He has worked on the editorial board of the journal of the Korea Society for Customs from 2017 and the e-Business Studies from 2019. His research interests include: Global Marketing, IoT, FDI, Cross-cultural Management.

Received: November 17, 2022; Accepted: October 05, 2023.

Computer Science and Information Systems 21(2):547–568 https://doi.org/10.2298/CSIS230323010W

Design of TAM-based Framework for Credibility and Trend Analysis in Sharing Economy: Behavioral Intention and User Experience on Airbnb as an Instance

Yenjou Wang[1], Jason C. Hung[2], Chun-Hong Huang[3], Sadiq Hussain[4], Neil Y. Yen[5], and Qun Jin[6]

[1] Waseda University, 2-579-15 Mikajima, Tokorozawa, 359-1192 Saitama, Japan
yjwjennifer2021@ruri.waseda.jp
[2] National Taichung University of Science and Technology, No. 129, Section 3, Sanmin Rd, North District, 404Taichung, Taiwan
jhung@gm.nutc.edu.tw
[3] Lunghwa University of Science and Technology, No. 300, Section 1, Wanshou Rd, Guishan District, 333 Taoyuan, Taiwan
ch.huang@mail.lhu.edu.tw
[4] Dibrugarh University, Dibrugarh, 786004 Dibrugarh Assam, India
sadiq@dibru.ac.in
[5] Aizu University, Aizuwakamatsu City, 965-8580 Fukushima Prefecture, Japan
neil219@gmail.com
[6] Waseda University, 2-579-15 Mikajima, Tokorozawa, 359-1192 Saitama, Japan
jin@waseda.jp

Abstract. Sharing economy redefines the meaning of share. Thanks to it, products provided by suppliers may have rather different standards due to their subjective consciousness. This situation brings high pre-purchase uncertainties to consumers, therefore, trust between suppliers and consumers then becomes a key to succeed in the era of sharing economy. Airbnb, one of the platforms that best describes the concept of sharing economy, is taken as an example in this study. Our team designs a series of scenarios and assumptions that follow the criteria of the Technology Acceptance Model (TAM) to find out various factors that affect customer behavioral intentions and prove that trust is the most important factor in the Sharing economy. Both parties, including host and user on the platform, are considered as subjects, and a three-year-long questionnaire test is implemented to collect data from end-users in order to reach an objective conclusion. Partial Least Squares-Structural Equation Modeling is then applied to verify the hypothesis. In addition, consumption is a continuous action, personal experience may also affect trust in the Airbnb and even consumption propensity. Therefore, Multi-Group Analysis (MGA) is used to explore the impact of consumer experience differences on trust and purchase intention. Finally, the results show that the ease of use of the Airbnb Platform has a greater impact on consumer attitude than all of the information on Airbnb, and then have a positive impact on overall behavioral intentions.

Keywords: sharing economy, behavior and trend analysis, TAM model, confirmatory factor analysis, multi-group analysis.

1. Introduction

Rapid development of sharing economy model prompts the redefinition of ownership. This model relates to exploiting profits using the Internet to link between the idling resources and demands of individuals [1]. The ongoing development of web-based information technology has boosted the accessibility of individuals reaching out those who are seeking goods and services [2]. And such phenomenon not only prompts economy development but also changes consumption patterns.

The sharing economy is an economic model in which individuals are able to borrow or rent assets owned by someone else. In other words, the sharing economy has redistributed the resources and promotes its re-sharing and re-use to create new value. However, it is not a new idea, but the reassembly of existing concepts widely applied in many fields. Some of the sharing organizations appeared early before Capitalism, such as charities and religious organizations or in the form of flea markets, swap meets, and second-hand shops...etc. This kind of transaction method can promote the exchange of goods or services between people. However, hinder we didn't have good communication channels and technology in the past, people just focus on face-to-face consumption and didn't pay attention to it. But it has regained a new impetus through information technology now, especially Web 2.0, mobile technology, and social media now. In 2000, Online peer-to-peer (P2P) marketplaces are growing at a rapid rate, people try to use the Internet to achieve the best use of things. These marketplaces comprise individuals (consumers) who transact directly with other individuals (sellers), while the marketplace platform itself is maintained by a third party [3]. In 2011, the TIME nominated what is now commonly understood as the Sharing Economy as one of "10 ideas that will change the world". Sharing economy marketplaces have flourished because of network communication technology. And in Europe and the US, the "Sharing Economy", the new concept of network service technology innovation model, is already flourishing spread globally. For example, Just Park shared the parking space in England, Ola rent transportation in India, Time Republic exchange extra time of labor, and Chegg not only rent textbooks but also help students to find tutors. And in recent years, the sharing economy concept also spread into Asia. Because of the high population and high consumption ability, the Asia market has become an important global economic leader. People also start to have many remaining resources. If these resources can be used well, they must bring Potential risks also are found due to the rapid growth of the sharing economy. Relative to the traditional business model, the warmth of people-to-people conversations cannot be felt. The more important is that the personal subjective cognition or information asymmetry between the seller and the buyer cannot be efficiently solved by actual touch with the product in sharing economy. Under these factors, it becomes more difficult to convince the buyer to trust that this is a good product and to purchase it. So, building trust a strategically important issue at the beginning of the B2C relationship [4]. Especially in the sharing economy, most suppliers mostly operated by individuals, they do not have a strong brand to support their reputation. Therefore, trust is particularly important in the sharing economy. In the era of sharing economy, 'trust' has become a kind of quasi-money. It would be difficult in sharing the economy without trust. When strangers shared with each other, greater information transparency let trust stronger. Therefore, the success of the sharing economy depends on establishing mutual trust.

Unlike the vendor-client relationship in the traditional business model, Information Technology (IT) becomes the intermediate, and the only one, that connects buyers and sellers. This intermediate is supposed to prompt interactions between buyers and sellers like a bridge. Therefore, whether this platform is accepted by customers becomes an important key to the success of a transaction. The Technology Acceptance Model (TAM) is a model of user acceptance of information systems technology based on the theory of reasoned action. The first school of thought considers a Web site to be information technology, and as such argues that the same use-antecedents that apply across IT, namely Perceived Usefulness and Perceived Ease of Use as identified by TAM [5-6]. TAM has been used in a variety of studies to explore the factors affecting an individual's use of new technology. Casalo' et al. (2010) also pointed out that consumer participation in online travel communities is affected by Perceived Ease of Use and Perceived Usefulness. Although there are many studies on Airbnb, an explicit and comprehensive understanding of the sharing economy, some literature focus on reputational feedback mechanism or topics related to the nature of peer-to-peer markets. [7-8]. Others focus on the exploratory study in sharing economy or the topic of legal [9-10], but the role of effect on trust to user's Behavioral Intention, is limited [11-12]. C Consequently, the TAM is used to explain whether a trust has an impact on the consumption intentions of Airbnb users in this study. In addition, we believe that consumption is a repetitive behavior, the consumer's attitude and consumption behavior are also indirectly affected by personal experience. E.g., if a customer is cheated in the previous transaction, it may cause customers to avoid using the same platform to consume. Therefore, personal experience will be additionally added to the TAM model to explore whether the difference in experience indirectly affects trust and consumption intentions. Three purposes are pursued as follows.

- To design a theoretical model that explores the effect of the perceived belief for antecedents (i.e., Perceived Usefulness and Perceived Ease of Use) and consequences (i.e., trust and Behavioral Intention)
- To identify the correlations among contexts, provided by/via sharing economy platform (i.e., Airbnb), trustworthiness, and users' experience, and their implicit impacts on user behavior and future purchase intention

The rest of this article is arranged as follows. An overview of the related work is described in Section II. Section III details the proposed model of the research. From the beginning, The TAM model was established according to the research objectives, and hypotheses were established based on this model. Partial Least Squares-Structural Equation Modeling (PLS-SEM) is used to analyze the relationship between hypotheses. Finally, Multi-Group Analysis (MGA) is used to analyze the trust degree and difference according to different user experiences. Section IV goes ahead to discuss the results, and Section V then concentrates on the findings from hypotheses and experiment results. Finally, Section VI then concludes the work and points out potential directions.

2. Related Work

2.1. The importance of trust in Sharing Economy

Sharing economy has become an emerging platform and its growth in various sectors especially in the tourism sector is phenomenal [13-15]. As people 's attention rises, Various surveys and models had been deployed in this area [13-27]. M. Abdar et al. proposed a universal user model to reflect differentially of internal (gender, age, nationality etc.) and external (social media, time, device etc.) factors on crowd's behavior and preference [14]. The statistical and machine learning approach divulged that the users' internal and external factors shared similarity with their behavior pattern. They found that Airbnb users are interested in interactions with host, local culture and unique accommodations of atmosphere and interiors. These three aspects have significance impact on the Airbnb users. Wu et al. [16] explored the purchases made on one of the top short-term rental sites in China called Xiaozhu.com to find the effects of host attributes on such purchases. The data was collected from 935 hosts from Beijing during the period 18th November 2015 to 14th February 2016. The host attributes and their rental characteristics were collected through python powered crawler program. The effects of the attributes were estimated by using Poisson regression model. They found that the key host attributes were gender of the host, personal profile, the number of owned listings, time of reservation confirmation and the acceptance rate. From the sheer volume of reviews about a product on the web, it is difficult to find the true quality of it [17].

Through the above research, Host has a certain influence rate on customer behavior in sharing economy, but as a consumer, it is not easy to perceive real evaluation. Under highly uncertain factors, trust plays a crucial role in developing relationships with customers on this platform [18]. The study by [18] suggested that experience in using the web and a higher degree of trust in e-commerce were the influencing factors of customer's trust. The key factors in this area are user's web experience, technical trustworthiness, site quality and perceived market orientation. Higher level trust in e-commerce makes the people participate in e-commerce. According to their study, the top three risk reduction strategies were partnerships with well-known business partners, money back warranty and positive 'word of mouth'.

The authors in [19] integrated the economic and sociological theories about institution-based trust to recommend that three IT-enabled institutional tactics - credit card guarantees, third-party escrow services and feedback mechanisms – created buyer trust in the group of online auction sellers. Their structural model was supported by the data collected from Amazon's online auction marketplace comprising of 274 buyers. Their study showed that self-reported and actual buyer behaviors were correlated with transaction intentions. Their findings also encompassed that both "strong" (legally binding) and "weak" (market-driven) mechanisms derived from perceived effectiveness of institutional mechanisms. Yang et al. [20] devised a research model to understand the continuance use intention in trust in sharing economy. They integrated Trust Building Model (TBM) with attachment theory and identified trust initiators- affect and cognitive

based trust. Their work demonstrated the mediating role of attachment in the relationship between behavioral outcome and trust.

The researchers in [21] revealed that review scores were impossible to differentiate in Airbnb as all hosts obtained maximum values. They investigated the Airbnb databases and found that the guests relied on host's photo as communicating trustworthiness. The hosts who had personal photos were perceived as more trustworthy and had more likely to be booked. In sharing economy transactions, members of both sides must trust one another to perform in good faith. Cheng et al. [22] empirically explored potential guests` trust perceptions in Airbnb via online review contents. They discovered six thematic characteristics of accommodation experiences from the review contents. They found that prominent cognitive themes were repurchase intention, location, host attributes, room description, overall evaluation and room aesthetics. They predicted the trust perceptions by utilizing Convolutional neural network. Zloteanu et al. [23] engendered an artificial sharing economy accommodation platform to study how reputation information and community-generated trust impacted user judgment. They varied the elements concerned to hosts' digital identity, exploiting users' decisions to interact and their perceptions. They came to a conclusion that reputation and trust not only enhanced users` credibility, perceived trustworthiness of hosts but also proclivity to rent a room in their home. Complete profiles or profiles with user selected information had done that effect.

The authors in [24] investigated the trust concept and its temporal C2C relationships with users of Airbnb from the viewpoint of an accommodation provider. They exploited the formation of trust by integrating two antecedents- 'Familiarity with Airbnb.com' and 'Disposition to trust'. Further, they discriminated between 'Trust in renters' and 'Trust in Airbnb.com' and scrutinized their inference on two provider intentions. Their results exhibited that both trust constructs were critical to instigate a sharing deal successfully between two parties. Tussyadiah et al. [25] conducted a multi-stage study to examine how Airbnb hosts eloquent themselves online and how consumer retort to varied host self-presentation blueprint. They found that hosts in Airbnb presented themselves as (1) an individual of a certain profession or (2) a well-traveled individual, enthusiastic to meet new guest. They utilized text mining methods comprising of Airbnb hosts' descriptions from 14 major cities. Consumers responded to the two host self-presentation techniques in a different way and well-traveled hosts demonstrated elevated levels of perceived trustworthiness. The study in [26] investigated sources of distrust in the context of Airbnb. They reviewed the negative comments posted by Airbnb customers on Trustpilot's website. They searched for the keyword 'trust' to mine the negative impact of trust with Airbnb. They extracted 216 negative reviews from the 2733 online reviews. They employed the grounded theory approach which derived two themes that presented the source of distrust: the hosts' unpleasant behavior and Airbnb's poor customer service. The managerial implications were that the customers' concerns should be addressed with positive actions, with prompt apologies and to compensate these customers to negate their distrust. Penz et al. [27] recognized vital aspects of the sharing economy to illustrate its potential in fostering sustainability. It was disparity to applications and definitions of sharing economy models which did not focus on sustainability. Their qualitative and quantitative research examined edifice of communities on consumer side as well as accomplishment of regulations and trust-building in the interaction between consumers and providers in Europe and Asia.

2.2. Trust Analysis Model based on TAM for Sharing Economy

TAM is one of the most commonly-applied theories in the field of information system (IS)/information technology (IT) to examine issues related to usability [28]. Major concepts include: 1) Perceived Ease of Use (PEOU) that presents the extents of user's believe in a system free of effort to use; 2) Perceived Usefulness (PU) that presents the extents of user's believe in a particular system that improves the performance at job; and 3) dependent variable behavioral intention (BI) that presents the extents to which one has devised conscious plans to execute or not in some future behavior.

TAM can be served as a starting point for scrutinizing the effect of external variables that can demonstrate on behavioral intentions [29]. TAM has progressed because of its flexibility via a meticulous development process. The simplicity and the understandability have made TAM one of the extensively used models in the IT research. It can be used to explore user requirement and key features vital for e-services because of its adaptability. Bielefeldt et al. [30] investigated the barriers to participation in the sharing economy. They accomplished a survey in Germany on car sharing. They found that society, personality and firm-related barriers had noteworthy effects on behavioral intention and Attitude that determined participation by employing PLS with structural equation modeling. The authors in [31] devised an empirical analysis model by taking into account the features of sharing economy services. They extended TAM by incorporating perceived enjoyment, reliability and price sensitivity to TAM to derive the key factors that had an effect on the use intention and distinctiveness of services on sharing economy. Their results asserted that use intention, perceived enjoyment, Perceived Ease of Use, Perceived Usefulness, reliability, technology innovation, self-efficacy and price sensitivity exhibited and affected in different ways. The researchers in [32] interviewed 50 drivers who provided service and cars in a digital car-sharing platform. They integrated TAM and Social Exchange Theory (SET) to examine salient motivators in this regard. They presented a motivation model of users' sharing opinion based on Self-determination Theory (SDT) in digital platform besides it. Sun et al. [33] examined the critical factors for lack of adoption in peer-to-peer indirect exchange services. They investigated the usage and attitudes towards peer-to-peer resource sharing sites among 37 New York City residents. Furthermore, they conducted a survey consisting of 195 respondents to determine the function of trust on willingness to lend. They also discussed the non-monetary and monetary structure issues related to adoption. They employed prior research on peer economies and critical mass theory to devise a TAM for indirect exchange systems that incorporated ease of coordination and generalized trust.

Two theoretical models [34] were employed TAM and Diffusion of Innovation Theory to examine consumer adoption of the Uber mobile application. Their results illustrated that social influence, observability, complexity, compatibility and relative advantage had crucial influence on both Perceived Ease of Use and Perceived Usefulness that led to consumer adoption intentions and Attitudes. They combined the two ad-hoc adoption theories. Wang et al. [35] investigated the key factors of the consumers' intention to use ride-sharing services and to promote such services. They extended TAM by utilizing three novel constructs: perceived risk, environmental awareness and personal innovativeness. They surveyed 426 participants with questionnaire and their model based on it was empirically tested. The experimental

results showed that Perceived Usefulness, environmental awareness and personal innovativeness had positive association with consumers' intention to hire ride-sharing services while there was negative association between perceived risk and Perceived Usefulness and the intention. Furthermore, personal innovativeness is negatively related to perceived risk, but positively related to Perceived Usefulness.

The current research, including it, has been introduced in the introduction. In addition to the basic discussion of ease of use and usefulness in the application of TAM in Sharing Economics, most of the research focuses on discoursing the impact of the social environment on consumer behavior and recognizing the changes in the overall environment on consumer behavior Most of them show a positive correlation. However, the very important "trust" in e-commerce is rarely discussed. Therefore, this study focuses on whether consumer behavior is affected by trust.

3. Related Method

This section discusses the method applied to examine the effect of the antecedents of the TAM. In addition, discussions on the research model and hypotheses development, data collection, sampling, and questionnaire design, and analytical methods are presented.

3.1. Research Model and Hypotheses

This study proposes a research model and hypotheses development to verify the importance of trust in consumers. To make the research results close to actual consumer behavior, personal experiences are also considered an important factor while discussing. E.g., the satisfaction of previous use, whether the previous transaction encountered a situation, etc. To verify these hypotheses, some constructs are proposed to establish TAM. Fig. 1 shows the research framework based on TAM. Under each of these constructs, there are several indicators with similar properties, which are used to analyze the values of the construct. Table 1 explains the term definition for the pre-defined constructs. Since the purpose of this study is to explore the impact of trust on consumer behavior. This study mainly discusses trust-related issues and understands their relationship with other corresponding constructs. We assume that the results can verify that trust is one of the most important factors affecting consumer behavior in sharing economy.

H1: Airbnb context has a significant positive effect on Perceived Usefulness.
H2: Airbnb context has a significant positive effect on Perceived Ease of Use.
H3: Personal Experience has a significant positive effect on Perceived Usefulness.
H4: Personal Experience has a significant positive effect on Perceived Ease of Use.
H5: Perceived Ease of Use has a significant positive effect on Perceived Usefulness.
H6: Perceived Ease of Use has a significant positive effect on Attitude.
H7: Perceived Usefulness has a significant positive effect on Attitude.
H8: Perceived Usefulness has a significant positive effect on Behavioral Intention.
H9: Attitude has a significant positive effect on Trust on Airbnb and Trust on Host.
H10: Perceived Usefulness has a significant positive effect on Behavioral Intention.

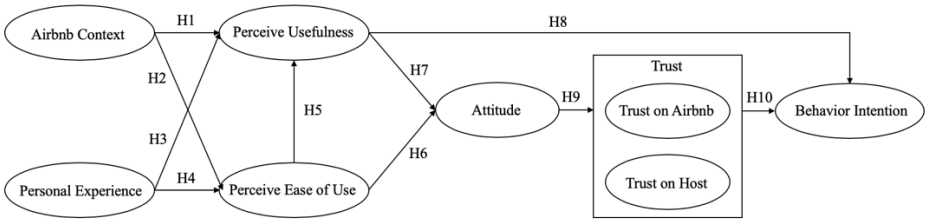


Fig. 1. Research Framework

Table 1. Term Definition of Constructs

Construct	***Definition***
Airbnb Context	All of the hostel Information on the Airbnb platform, including room description, pictures and etc.
Personal Experience	The experience of previous use and motivation of use.
Perceived Ease of Use	Convenience and operation feelings of using the Airbnb interface
Perceived Usefulness	Recognition of using Airbnb to book a room
Attitude	Satisfaction with the Airbnb
Trust	Trust in Airbnb platform and Airbnb Host
Behavioral Intention	Willingness level to use Airbnb again in the future

3.2. Data Collection and Questionnaire Design

Data Collection. In this study, users who had experience in using Airbnb are selected as subjects to conduct the questionnaires. Airbnb is a well-known platform, however, it has been used by fewer people than we expected, so finding people who have used it and are willing to do a questionnaire is also less than we expected. A survey period, in general, is set to three months, but however, the survey period of this study extends to a total period of 21 months (July 2018 to March 2020) by reviewing the collected data every three months, to obtain complete and objective data. Since such data is strongly required because of frequent releases and updates on user interface and service provisions by Airbnb during the above period. The updates may cause especially changes in users' experience and their subjective thoughts on the platform, and wills to continuously stay with the same platform or move to other platforms. The situation in that Airbnb does the updates based on its users' feedback is also taken into consideration. This means that requirements and user experience for the platform have constantly been changing by users. With data collection after a longer period and doing trend analysis, analysis results are more discriminative than the one with data collected within a short period. The questionnaires were mainly distributed via online survey services (i.e.,

SurveyMonkey, Google Form), but considering recent changes in human behavior that is daily time spent on SNS (Social Network Service) has a sharp growth, reaching more than 2 hours/day [36]. SNS, like Line, Facebook, and WeChat, was also applied to reach as many as possible potential subjects. The questionnaire was distributed to 16 SNS groups. There are about 50-100 members in a group, therefore a total of about 1,200 questionnaires are sent. Among them, about one-fourth of the questionnaires in each group will be completed. Fortunately, a total number of 268 questionnaires were collected and about half of them were confirmed valid to conduct further analysis after excluding those samples with extreme statistical significance.

Questionnaire and Constructs. The research model is based on the extended version of Davis' TAM and is developed to derive the Exogenous variables that affect user Behavioral intention. The TAM model will be used to explain how external variables affect the user acceptance process. In addition, path analysis is applied to explore the empirical strength of the relationship in the proposed model.

Questionnaire Design. Based on the hypothesized model developed, and a detailed review of the related literature on user acceptance of technology, information content, a 37-item questionnaire was devised as a measurement scale for the research. This study uses the Likert seven-point scale, with one - seven points where lower point stands for negative feedback and higher point stands for the positive ones showing as "Strongly Disagree", "Disagree", "Somewhat Disagree", "Neutral", "Somewhat Agree", "Agree", "Strongly Agree".

3.3. Analytical Methods

To assess the overall model of the study, Hair et al. (2017) [37] stages in structural equation modeling (SEM), were adapted. From the result of that literature review, the study incorporated those stages, and the following steps were adopted and implemented in this study. Statistical analysis for the study included descriptive statistics, Confirmatory Factor Analysis (CFA), SEM, and Multi-Group Analysis (MGA). Detailed information for each analysis method is as follows.

Descriptive Statistics. First, this study starts with descriptive statistical analysis that includes gender, age, education, occupation, and annual income. Besides, the study focuses on Asians, therefore questionnaire respondents need to respond to their National. Descriptive analyses are used to determine items of measurement. The mean and standard deviation of variables are used to identify measurement items that are tested on the survey questionnaire in the next stage for overall model testing.

Confirmatory Factor Analysis. CFA is one of the most applied methods which implement by a process to identify the consistency, and relationship as well, between scientific hypotheses and obtained results through the research. CFA is usually implemented by several sequential stages. Different discriminant indicators are usually adopted due to different research purposes and statistical software. This method is

applied to models that already have preliminary settings to confirm the fitness between the hypothetical model and the data [38]. Factor loading, convergent validity, and discriminant validity are used to gradually analyze and study the model. In addition, some models, such as Path Analysis/SEM, PLS-SEM etc., are often paired to conduct in the analysis. With the main targets on measurement and structural data model, PLS-SEM is then adopted to conduct the analysis in this research by statistical software, SmartPLS [39].

As above, evaluating the hypothetical model usually begins with factor loading which is the process for observation of correlation(s) between constructs and indicators [40]. Factors that are less relevant to this study have been eliminated. Secondly, the average variance extracted (AVE) is used to identify the convergent validity of the model [41], which is checking the attribute of indicators in each construct is consistent or not. According to the definition, the value of factor loading shall be higher than 0.60, and the value of AVE shall be 0.50 or higher to reach a valid analysis. To show how much variation per node, the square of the indicator's outer loadings which can also show the reliability of indicator is calculated. For exploratory research, we expect the value should close to 0.70, and the higher the better [42-43]. The final step of the CFA process is discriminant validity. The AVE value is checked again. All outer loading must be higher than cross-loadings in models with discriminant validity. This implies that the direct correlation between constructs must be higher than the indirect correlation.

Partial Leased Squares-Structural Equation Modeling. PLS-SEM is a widely applied multivariate analysis method to estimate variance-based structural equation models and become a popular data analysis technique in success factor studies, especially in the application of information system is the most widely used [44]. It also has been used in areas of marketing, enterprise resource planning systems, and knowledge management systems.

PLS-SEM fits especially to those cases with small size of samples, and it meets the requirement of reflective and formative models that contain multiple or single item construct indicators. This method is allowed to model complex relationships among multiple variables. Researchers often use this approach to identify relationships among variables. In short, PLS-SEM is a variance-based method that estimates composites representing latent variables in path models. Based on the information provided in the literature and the intent of the research study, PLS-SEM was used to analyze the data. The significance of the path coefficients was determined by comparing these to the critical t values for significance levels of 0.05 and 0.10. And then the assessment of the structural model, started from obtaining the coefficient of determination (R^2) achieved in the relationship between the independent variables and the dependent variable ranges from 0 to 1 and the closer to 1, the greater the proportion explained. Before testing the model, the data was checked for common method bias. Then, measurement model was examined, followed by structural model.

Multi-Group Analysis. MGA is used to determine whether there are obvious differences in different parameters in the data set (e.g., outer weights, outer loadings, and path coefficients). SmartPLS used in this research provides many MGA methods, e.g., Confidence Intervals (Bias Corrected), Partial Least Squares Multi-Group Analysis (PLS-MGA), etc. Among them, PLS-MGA is often used to determine the difference in

path coefficients between different data groups [45]. Therefore, in this study, the PLS-MGA is used to divide the data into two groups (accidents / no accidents when using Airbnb) and investigate whether trust and attitude are significantly affected by different personal experiences.

4. Research Result

4.1. Confirmatory Factor Analysis Results

CFA, as discussed earlier, is to verify the consistency of the hypothetical model and the experimental results. Therefore, we must confirm that each indicator and construct meet the validity standard before verification. The first stage is factor loading that must be measured and used to delete the indicator associated with the lower relationship in construct. Every indicator is analyzed by CFA and must meet the preferred threshold at 0.60. We observed that all indicators reached the boundary threshold except two indicators, PER4 and PER5 with obtained scores at 0.440 and -0.275 respectively. An outer loading relevance test is conducted to determine whether the indicator should be excluded by evaluating each indicator's contribution to the effectiveness of the content [46]. Table 2 presents the results after factor loadings.

Testing internal consistency reliability is the next step. The double verification method [47] is applied to ensure consistency reliability through the values of Cronbach 's Alpha and AVE. Cronbach's Alpha has a required threshold value of 0.70 and higher to show reliability, while the threshold value of AVE should be above 0.50. In terms of consistency reliability, the composite reliability (CR) threshold value, say 0.70 or higher, is used for discrimination.

To ensure the convergent validity is one of the bases of the evaluation model, therefore it should take place in the beginning. Table 3 presents the results of each construct at the convergent validity evaluation. The results indicate that all the constructs fulfill the minimum requirement. The value of Cronbach's Alpha of all constructs are greater than the basic value of 0.70, while the value of AVE reaches 0.80 in average. Although AVE value for construct the personal experience touches 0.582, which is considered lower than others, its value still passes the standard value at 0.50. In addition, it is found that all values reach 0.90, the baseline for CR, and all of them are higher than corresponding Cronbach's Alpha value. This proofs that the model has internal consistency reliability, indicators' properties for all constructs have no direct conflicts in between and demonstrate that our model has discriminant validity.

4.2. Path Analysis

Examining the proposed hypotheses is then conducted after the results of CFA were obtained. Before examining the proposed hypotheses, all constructs that accurately interpret given indicators must be ensured to confirm the predictive capability of our

model. Therefore, the value of R2 is used in this step to check the interpretation capability of each construct of our model. As shown in Fig. 2, the R2 value of all structures reaches the given threshold at 0.26 [48]. Next, the PLS-SEM was used to do Path Analysis. To ensure the accuracy of the results, subsamples are used to estimate the PLS path model.

Table 2. Factor Loading for Model

Construct	***Indicators***	***Factor Loading***
Airbnb Context	AC1	0.843
	AC2	0.820
	AC3	0.908
	AC4	0.882
	AC5	0.804
Personal Experience	PER1	0.847
	PER2	0.889
	PER3	0.795
Perceived Ease of Use	EOU1	0.941
	EOU2	0.946
	EOU3	0.918
Perceived Usefulness	PU1	0.957
	PU2	0.962
Attitude	ATT1	0.949
	ATT2	0.933
Trust	TA1	0.891
	TA2	0.892
	TA3	0.909
	TA4	0.826
	TA5	0.862
	TH1	0.890
	TH2	0.912
	TH3	0.827
Behavior Intention	BI1	0.942
	BI2	0.953
	BI3	0.900

Table 3. Measure that Discriminant Validity for Model

Construct	*Cronbach's Alpha*	*AVE*	*CR*
Behavioral Intention	0.924	0.802	0.952
Perceived Ease of Use	0.928	0.811	0.954
Perceived Usefulness	0.914	0.843	0.959
Personal Experience	0.801	0.582	0.882
Airbnb Context	0.906	0.651	0.930
Attitude	0.871	0.778	0.939
Trust	0.962	0.759	0.968

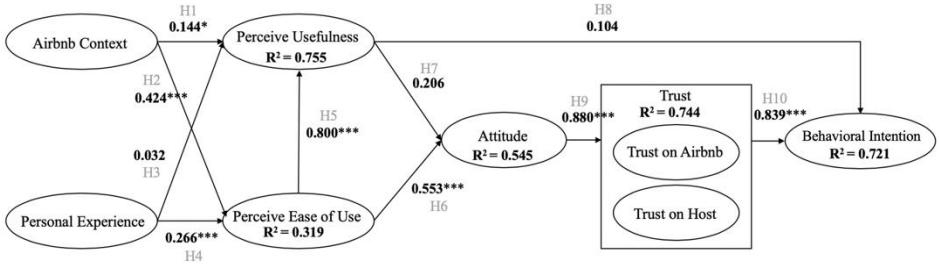


Fig. 2. Result of Path Analysis (*p < 0.05, **p < 0.01, ***p < 0.001)

Table 4. The Result of Path Analysis

		Path Coefficient	t-value	p-value	Hypothesis Testing Result
H1	AC → PU	0.144	1.995	0.046	Accept
H2	AC → EOU	0.424	4.979	0.000	Accept
H3	PER → PU	0.032	0.605	0.545	Reject
H4	PER → EOU	0.266	3.796	0.000	Accept
H5	EOU → PU	0.800	13.313	0.000	Accept
H6	EOU → ATT	0.553	3.350	0.001	Accept
H7	PU → ATT	0.206	1.252	0.211	Reject
H8	PU → BI	0.014	0.134	0.893	Reject
H9	ATT → Trust	0.880	30.889	0.000	Accept
H10	Trust → BI	0.839	9.343	0.000	Accept

Table 5. Constructs and Measurement Items

Indicators		***Measure***
Airbnb Context	AC1	Host provides the number of room's photos and the resolution of those photos which are important information.
	AC2	The brief overviews of the room are important information such as the type of rooms, available number of people, the number of bathrooms/bedrooms, time of check in/out.
	AC3	The amenities of the room that host will provide or not, ex: WIFI, toiletry, breakfast are also an appropriate information.
	AC4	Host set the pricing of room including discounts, extra people, cleaning fee, cancellations fee are important information.
	AC5	The rules of house are important reference.
Personal Experience	PER1	Interface of Airbnb is similar to the website I used before.
	PER2	I use it before and I am satisfied .
	PER3	I use Airbnb because my friend are also using it.
	PER4	Have you ever met the situation below? Advertisement does not match corresponding product
	PER5	Have you ever met any accident during your stays?
Perceived Ease of Use	EOU1	Airbnb is easy to use even for the first time.
	EOU2	Booking rooms on Airbnb is easy.
	EOU3	Information provided by Airbnb makes booking rooms easier.
Perceived Usefulness	PU1	Information provided by Airbnb is useful for users to search and book rooms.
	PU2	Information provided by Airbnb allows me to know that how to search and book rooms more efficiently.
Attitude	ATT1	I think Airbnb is worthy to use for booking rooms.
	ATT2	Using Airbnb for booking hotel is a good idea.
Trust on Airbnb	TA1	Booking on Airbnb is reliable.
	TA2	Accommodation options of Airbnb is trustworthy.
	TA3	Room information is consistent with the facts which is provided by Airbnb.
	TA4	If I required help, Airbnb would do its best to help me.
	TA5	I believe Airbnb would do its best to support me Immediately.
Trust on Host	TH1	The room information is trustworthy which provided by host in Airbnb.
	TH2	The room information with the facts provided by host in Airbnb is consistent.
	TH3	I believe that host in Airbnb can keep its promises and commitments.
Behavior Intention	BI1	I would like to choose Airbnb to collect information when I want to search rooms or make a reservation.
	BI2	I will still choose Airbnb for booking rooms in the future.
	BI3	In the future, I will intend to increase the use of sharing economy platforms.

Table 6. Comparison of MGA Result

	Path Coefficient in no Accident	*Path Coefficient in had Accident*	*Impact percentage*
AC → PU	0.184	0.104	-0.04%
AC → EOU	0.435	0.469	+1.07%
PER → PU	0.049	0.014	-0.72%
PER → EOU	0.292	0.123	-0.42%
EOU → PU	0.791	0.812	+1.02%
EOU → ATT	0.483	0.562	+1.16%
PU → ATT	0.314	0.190	-0.41%
PU → BI	0.111	0.048	-0.57%
ATT → Turst	0.888	0.904	+1.02%
Trust → BI	0.805	0.980	+1.22%

In general, there are 5,000 subsamples randomly generated at this stage. According to the result of Path Analysis, the closer the obtained path coefficient score is to 1, the stronger the relationship will be. Based on the definition, we may find that the weakest relationship is H8 (Perceived Usefulness → Behavioral Intention) with a path coefficient of 0.014 and the strongest relationship is H9 (Attitude → Trust) with a path coefficient of 0.880. This means that in our model, Perceived Usefulness has the least influence on Behavioral Intention, while Attitude has the most influence on Trust.

This concludes that the analysis of the structural model and the hypothesis findings is discussed. According to the definition of SEM analysis, the t-value is required to be greater than the significance level of 1.96 and the p-value shall be less than 0.05 if a hypothesis can be considered valid. This study has a total number of 10 hypotheses. All hypotheses are accepted except H3 ($t = 0.605$, $p = 0.545$), H7 ($t = 1.252$, $p = 0.211$), and H8 ($t = 0.134$, $p = 0.893$). Table 4 provided the results of path analysis and the proposed hypotheses. All results were produced based on bootstrapping with 5000 subsamples. The full names of the abbreviations in the table are as follows: AC= Airbnb Context; PER= Personal Experience; PU= Perceived Usefulness; EOU = Perceived Ease of Use; ATT= Attitude; Trust= Trust on Airbnb and Trust on Host; BI= Behavior Intention. The result show that the most significant hypothesis is H9 ($t = 30.889$, $p = 0.000$), followed by H10 ($t = 9.343$, $p = 0.000$), there is a joint relationship between the two hypotheses, and both are main objective of this study. The detailed questionnaire content is presented on Table 5.

4.3. Multi-Group Analysis

In all consumption behaviors, the user's Attitude will be affected by Personal Experience, especially in the sharing economy that emphasizes trust. Any accident may affect the Trust and consumption tendency. Therefore, this study divides Airbnb consumer data into two categories: (1) Unexpected accidents have been encountered in the use of Airbnb, and (2) No accidents have occurred in the use of Airbnb. And the use

of MGA analysis to explore that the Path Coefficient will be affected whether under different personal. In the result of MGA as Table 6, except for the received usefulness in the aforementioned research results, the impact on the model is relatively lower. After consumers encounter unexpected situations in using Airbnb, the ratio of Personal Experience affecting various constructs is low, but the impact of Airbnb content and Ease of Use on construct has improved overall. And Trust has the most impact on Behavioral Intention, which has increased by 1.22%. After encountering an accident, users will rely more on the content and convenience of Airbnb to affect their perception of the Airbnb platform. Behavioral Intention will be affected more by Trust than before.

5. Discussion

The research model for this study was based on the TAM. Airbnb context, Perceived beliefs, Trust are the independent variables, and Behavioral Intention is an outcome variable. For all constructs, there was a combined total of 28 indicators that were analyzed through CFA and PLS-SEM with SmartPLS. Although some revealed issues with factor loading, after amended all the indicators that factor loading, composite reliability, convergent validity, and discriminant validity were in line with the minimum threshold requirements. The findings showed that the model's predictive accuracy and overall significance.

5.1. Research Results and Hypothesis Discussion

As a result, H1, H2, and H4 were accepted, but hypothesis H3 was rejected. The mining of result is that the context of the Airbnb context had a positive and direct effect on Perceived Usefulness and Perceived Ease of Use. Personal Experience has no positive effect on received usefulness. This means that whether before or after use, current users pay more attention to the convenience and ease of use of the platform. The easy-to-use platform makes it easy for users to produce satisfactory Attitude, and then promote the next consumption behavior. These results support previous studies that perceived beliefs are affected by external variables.

Among H5 - H8, only H5 and H6 were accepted, while H7 and H8 were rejected. This shows that received ease of use and attitude and Perceived Usefulness a positive and direct effect. This result means that although received ease of use is helpful to improve the received usefulness, as mentioned earlier, the convenience is paid more attention by consumer now. We believe that such a result is produced because, in this generation of Sharing Economic, the usefulness of the platform has become a basic condition. To win in this fierce competition, the fluency of the platform must be strengthened and improved. Ease of use. E.g., it allows consumers to easily search for the target product during use, and easily go to the checkout page.

H9 and H10 is the focus of this study, and it is also the two most significant assumptions. This shows that the user's Attitude plays a huge role in consumption. It will affect the Trust of the landlord and the platform, which in turn Behavioral Intention will be affected. This result also shows the importance of the platform. The mechanism

instant feedback and evaluation are provided in most of the sharing economy platforms. This is to enable the platform to improve according to feedback and reduce information asymmetry with consumers. In turn, Trust is increased. When the platform is Trusted by more and more consumers, the evaluation will be relatively improved, and the willingness to consume will also increase. This result is also consistent with the aforementioned theory, Trust is an important key to affect the sharing economy.

5.2. Result Change Discussion

This study has been updated from the beginning of 2018 to the present. Although the construct of the model has been revised in the course of two years, the basic construct remains unchanged. Therefore, trend analysis was used to analyze the research results of these years to understand the change in consumer behavior. As shown in Fig. 3 below, the main goal in this research, "Trust", has always played an important influencing factor in consumer behavior. And, as the platform grows. As mentioned in related work, a large number of evaluations make it more difficult for consumers to judge the true evaluations, and Trust is increasingly valued in the sharing economy.

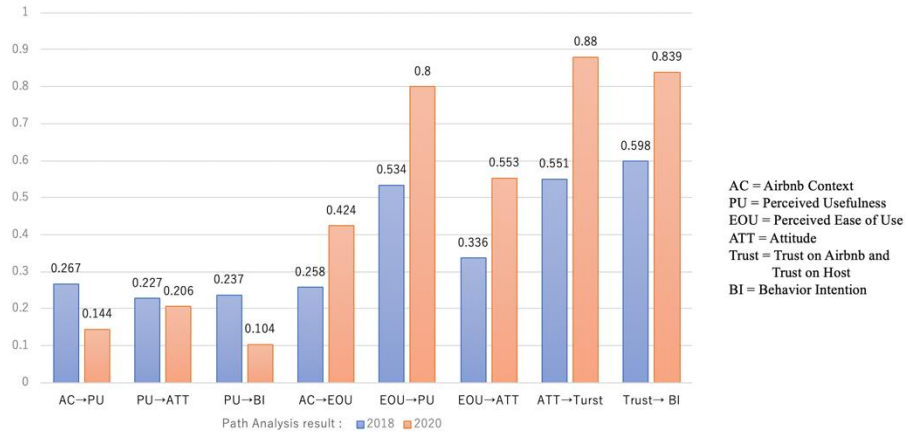


Fig. 3. Compare with Research Result of 2018 and 2020

Interestingly, on the platform side, the impact of Perceived Usefulness in the overall model has declined, and even Airbnb Context has not to effect on the improvement of Perceived Usefulness. However, the impact of the Perceived Ease of Use on model has increased significantly. Except as mentioned earlier, the integrity of all platforms has been improving, Perceived Usefulness has become basic. It also shows changes in consumer behaviors. Accurate and convenient, the impact on consumer behavior is gradually increasing.

This change is also consistent with Airbnb's platform changes in recent years. Airbnb has simplified search content in recent years, with a more intuitive user interface (UI) representation. After searching for listings, in addition to the display of basic listings, the evaluation of the landlord is also replaced with a star rating. If you want to view a

more recent review, you need to click on it again. This is to make the booking process faster, the quality of the listing can be understood by the user in a short time, reduce the consumer's consideration time, avoid being reviewed, and increase the chance of the booking being booked.

6. Conclusion

This research builds a theoretical model based on the TAM model. CFA and PLS-SEM statistical methods were used to explore whether several factors such as Trust affects consumer behavior in Airbnb. Factors such as contexts, user experience, perceived beliefs, attitude, trust, and behavioral intention that may cause the changes in usability were especially concentrated. Especially 'Trust' is considered the key that decides whether users accept to use sharing economy platforms according to past studies.

To estimate how trust influences Airbnb users, a hybrid TAM model with personal experience as one of the external factors is applied in this study. What can be known is that we usually conducted statistical analysis for a short period of time to conclude their assumptions in the past. However, user behavior should be treated as continuously changing trends from a statistical point of view. The results will be limited if the short period of data collecting. To reach a more objective result close to the situation, this study conducted the trend analysis for a period of approximately 2 years from Summer 2018 to Summer 2020. The issue of trust, according to the obtained results, is still the key factor that affects consumer behavior during the whole period. In addition, the impact of Perceived Ease of Use on consumer intention has significantly grown. While Perceived Usefulness is least impact of consumer intention.

Although Airbnb cannot stand for all the platforms of sharing economy, it indeed shows that it can be one of the most significant platforms in the field. It is foreseen that more and more similar platforms will be developed to meet the various needs of users. Through the results of this study, we are firm that preferences of consumers continue to change, so every platform need to be constantly changed to increase consumer preferences. However, increasing consumer trust level is the best way to increase consumer loyalty to the platform in sharing economy.

References

1. Rinne, A.: Circular Economy Innovation & New Business Models Initiative. Young Global Leaders Sharing Economy Working Group. (2013) [online]. Available: https://www3.weforum.org/docs/WEF_YGL_CircularEconomyInnovation_PositionPaper_2013.pdf (current February 2023)
2. Felsom. M., Spaeth, J. L.: Community Structure and Collaborative Consumption: A routine activity approach. American Behavioral Scientist, Vol. 21, No. 4. 614-624. (1978)
3. Botsman, R., Rogers, R.: What's mine is yours: how collaborative consumption is changing the way we live. Collins. (2011)
4. McKnight, D. H., Cummings, L. L., Chervany, N. L.: Initial trust formation in new organizational relationships. In: Academy of Management Review, Academy of Management, 473-490. (1988)

5. Davis, F. D., Bagozzi, R. P., Warshaw, P. R.: User Acceptance of Computer Technology: A Comparison of Two Theoretical Models. Management Science, Vol. 35, No.8, 982-1003. (1989)
6. Davis, F. D.: Perceived Usefulness, Perceived Ease of Use and User Acceptance of Information Technology. MIS Quarterly, Vol. 13, No. 3, 319-340. (1989)
7. Kohda, Y., Masuda, K.: How do Sharing Service Providers Create Value?. In: Semantic Scholar, [online]. Available: https://www.semanticscholar.org/paper/How-do-Sharing-Service-Providers-Create-Value-Kohda-Masuda/144bd1060ac4159e2afe08af1d3458d95a559554. (current February 2023)
8. Slee, T.: Some Obvious Things About Internet Reputation Systems. [online]. Available: https://www.semanticscholar.org/paper/Some-Obvious-Things-About-Internet-Reputation-Slee/410091be5c14ba4b17dcbd74f7e594f0c22782b5. (current February 2023)
9. Kassan, J., Orsi. J.: The Legal Landscape of the Sharing Economy. Journal of Environmental Law & Litigation, Vol. 27, No.1. (2012)
10. Guttentag, D.: Airbnb: disruptive innovation and the rise of an informal tourism accommodation sector. Current Issues in Tourism, Vol 18, 1192-1217. (2013)
11. Casalo, L.V., Flavi, C., Guinaliu, M.: Determinants of the intention to participate in firm-hosted online travel communities and effects on consumer behavioral intentions. Tourism Management, Vol.31, No. 6, 898-911. (2010)
12. Tsai, H. T., Pai, P.; Why do newcomers participate in virtual communities? An integration of self-determination and relationship management theories. Decision Support Systems, Vol. 57, 178-187. (2014)
13. Abdar, M., Lai, K. H., Yen, N. Y.: Crowd Preference Mining and Analysis Based on Regional Characteristics on Airbnb. In: Proceedings of the 3rd IEEE International Conference on Cybernetics (CYBCONF), Exeter, UK, 21-27. (2017)
14. Abdar, M., Yen. N.Y.: Design of A Universal User Model for Dynamic Crowd Preference Sensing and Decision-Making Behavior Analysis. EEE Access, Vol. 5, 24842-24852. (2017)
15. Tussyadiah, I. P., Pesonen, J.: Impacts of Peer-to-Peer Accommodation Use on Travel Patterns. Journal of Travel Research, Vol.55, 1022-1040. (2016)
16. Wu, J., Ma, P., Xie, K. L.: In sharing economy we trust: the effects of host attributes on short-term rental purchases. International Journal of Contemporary Hospitality Management, Vol. 29, No.11, 2962-2976. (2017)
17. Lee, S., Choeh, J.Y.: Predicting the helpfulness of online reviews using multilayer perceptron neural networks. Expert Systems with Applications, Vol. 41, No. 6, 3041-3046. (2014)
18. Corbitt, B. J., Thanasankit, T., Yi, H.: Trust and e-commerce: a study of consumer perceptions. Electronic commerce research and applications, Vol. 2, No. 3, 203-215. (2003)
19. Pavlou, P. A., Gefen, D.: Building effective online marketplaces with institution-based trust. Information systems research, Vol. 15, No. 1, 37-59. (2004)
20. Yang, S. B., Lee, K., Lee, H., Chung, N., Koo, C.: Trust Breakthrough in the Sharing Economy: an Empirical Study of Airbnb 1. In: Proceedings of the Pacific Asia Conference on Information Systems , Chiayi, Taiwan, Vol.27, 131. (2016)
21. Ert, E., Fleischer, A., Magen, N.: Trust and reputation in the sharing economy: The role of personal photos in Airbnb. Tourism Management, Vol. 55, 62-73. (2016)
22. Cheng, X., Fu, S., Sun, J., Bilgihan, A., Okumus, F.: An investigation on online reviews in sharing economy driven hospitality platforms: A viewpoint of trust. Tourism Management, Vol.71. 366-377. (2019)
23. Zloteanu, M., Harvey, N., Tuckett, D., Livan, G.: Digital Identity: The effect of trust and reputation information on user judgement in the Sharing Economy. PloS one, Vol. 13, No. 12, (2018).
24. Mittendorf, C.: What Trust means in the Sharing Economy: A provider perspective on Airbnb. Com. In: Semantic Scholar, Available via Semantic Scholar. (2016) [Online].

Available: https://www.semanticscholar.org/paper/Some-Obvious-Things-About-Internet-Reputation-Slee/410091be5c14ba4b17dcbd74f7e594f0c22782b5. (current February 2023).
25. Tussyadiah, I. P., Park, S.: When guests trust hosts for their words: Host description and trust in sharing economy. Tourism Management, Vol. 67, 261-272. (2018)
26. Sthapit, E., Björk, P.: Sources of distrust: Airbnb guests' perspectives. Tourism Management Perspectives, Vol.31, 245-253. (2019)
27. Davis, F. D.: Perceived usefulness, perceived ease of use, and user acceptance of information technology. MIS quarterly, 319-340. (1989)
28. Davis, F. D., Bagozzi, R. P., Warshaw, P. R.: User acceptance of computer technology: a comparison of two theoretical models. Management science, Vol. 35, No. 8, 982-1003. (1989)
29. Bielefeldt, J., Poelzl, J., Herbst, U.: What's mine isn't yours–barriers to participation in the sharing economy. Die Unternehmung, Vol.70, No.1, 4-2. (2016)
30. Lee, J. S., Jeon, H. S., Jeong, M. S.: An Empirical Study on the Use Intention to Sharing Economy Services: Focusing on Price Sensitivity Reliability and Technology Acceptance Model. Journal of Digital Convergence, Vol. 14, No. 7, 57-72. (2016)
31. Cheng, X., Zhu, R., Fu, S.: Modeling the Motivation of Users' Sharing Option: A Case Study Based on A Car-Sharing Digital Platform. In Proceedings of WHICEB, Wuhen, China, 64. (2016)
32. Sun, E., McLachlan, R., Naaman, M.: TAMIES: a study and model of adoption in P2P resource sharing and indirect exchange systems. In Proceedings of the 2017 ACM conference on Computer supported cooperative work and Social Computing, Portland Oregon, USA, 2385-2396. (2017)
33. Min, S., So, K. K. F., Jeong, M.: Consumer adoption of the Uber mobile application: Insights from diffusion of innovation theory and technology acceptance model. Journal of Travel and Tourism Marketing, Vol. 36, No. 7, 770-783. (2019)
34. Wang, Y., Wang, S., Wang, J. Wei, J., Wang, C.: An empirical study of consumers' intention to use ride-sharing services: using an extended technology acceptance model. Transportation, Vol. 47, 397- 415. (2020)
35. Hair, J. F., Hult, G. T.M., Ringle, C.M., Sarstedt, M. A primer on partial least squares structural equation modeling (PLS-SEM). SAGE, CA, USA. (2017)
36. J. Clement: Daily social media usage worldwide 2012-2022. In: Statista via Statista website. (2022) [online]. Available: https://www.statista.com/statistics/433871/daily-social-media-usage-worldwide/. (current February 2023).
37. Williams, B., Onsman, A., Brown, T.: Exploratory factor analysis: A five-step guide for novices. Australasian Journal of Paramedicine, Vol. 8, No. 3. (2010)
38. Vinzi, V. E., Chin, W. W., Henseler, J., Wang, H. (Eds.): Handbook of partial least squares: Concepts, methods and applications. Springer Science & Business Media, Germany, Berlin. (2010)
39. Ringle, C. M., Wende, S., Becker, J.-M.: SmartPLS 3, SmartPLS GmbH: Boenningstedt. (2015)
40. Hair, J. F., Ringle, C. M., Sarstedt, M.: PLS-sem: Indeed a silver bullet. The Journal of Marketing Theory and Practice, Vol.19, 139-151. (2011).
41. Hulland, J.: Use of partial least squares (PLS) in strategic management research: a review of four recent studies. Strategic Management Journal, Vol. 20, No. 2, 195-204. (1999)
42. Hair, J. F., Hult, G. T. M., Ringle, C. M., Sarstedt, M.: A primer on Partial Least Squares Structural Equation Modeling (PLS-SEM). SAGE, USA, Thousand Oaks. (2014)
43. Hair, J. F., Sarstedt, M., Ringle, C. M., Mena, J. A.: An assessment of the use of partial least squares structural equation modeling in marketing research. Journal of the Academy of Marketing Science, Vol. 40, No.3, 414-433. (2012)

44. Henseler, J., Ringle, C.M., Sinkovics, R. R.: The Use of Partial Least Squares Path Modeling in International Marketing. Advances in International Marketing, Emerald, Bingley, 277-320. (2009).
45. Ringle, C. M., Sarstedt, M., Straub, D. W.: A critical look at the use of PLS-SEM. MIS Quarterly, Vol.36, No. 1. 3-39. (2012)
46. Hair, J. F., Ringle, M., Sarstedt, M.: PLS-sem: Indeed a silver bullet. In: The Journal of Marketing Theory and Practice, Vol. 19, 139-151. (2011)
47. Peterson, R. A., Kim, Y.: On the Relationship between Coefficient Alpha and Composite Reliability. Journal of Applied Psychology, Vol. 98, No. 1, 194-8. (2013)
48. Cohen, J. (2nd ed.): Statistical Power Analysis for the Behavioral Sciences. Hillsdale, Lawrence Erlbaum Associates, USA. (1988)

Yenjou Wang (Corresponding author) is a doctoral student specializing in Human Sciences at Waseda University, having completed her Master's in Computer Science at the University of Aizu in 2021. Her interdisciplinary research spans computer science, engineering, and human informatics. Her areas of interest include big data analysis, optimization of machine learning models, social network analysis, and blockchain application in health analytics. She actively contributes to the academic community, publishing papers at various international conferences organized by IEEE. She serves as a chair at international conferences sponsored by IET. Additionally, she reviews for various academic journals.

Jason C. Hung is a Professor of Department of Computer Science and Information Engineering at National Taichung University of Science and Technology, Taiwan, ROC. His research interests include e-Learning, Intelligent System, Social Computing, Affective Computing, Multimedia System, Artificial Intelligence. Dr. Hung received his BS and MS degrees in Computer Science and Information Engineering from Tamkang University, in 1996 and 1998, respectively. He also received his Ph.D. in Computer Science and Information Engineering from Tamkang University in 2001. He is the founder of International Conference on Frontier Computing- Theory, Technologies and Applications, In April of 2014, he was elected as Fellow of the Institution of Engineering and Technology (FIET).

Chun-Hong Huang is an Assistant Professor of the Department of Computer Information and Network Engineering at Lunghwa University of Science and Technology. His research interests encompass the Information Analysis and Applications of Multimedia, as well as Human-Computer Interaction and Virtual/Argument Reality. Currently, his research is directed towards on the fields of Data Science and Acritical Intelligence.

Sadiq Hussain is System Administrator at Dibrugarh University, Assam, India. He received his PhD degree from Dibrugarh University, India. His research interest includes data mining, machine learning, medical analytics and deep learning. He is associated with Computerization Examination System and Management Information System of Dibrugarh University. He published various research and conference papers of international repute.

Neil Y. Yen is an Associate Professor at the University of Aizu, specializing in interdisciplinary research in computer science, information management, and human informatics. He earned his doctorate in Human Sciences from Waseda University in Japan and in Engineering from Tamkang University in Taiwan. He has been involved extensively in an inter-disciplinary field of research, where the themes are big data science, computational intelligence, and human-centered computing. He has been actively involved in the research community by serving as a Guest Editor, an Associate Editor, and a Reviewer for international referred journals and as the Organizer/Chair of the ACM/IEEE-sponsored conferences, workshops, and special sessions. He is now a member of IEEE Computer Society, IEEE System, Man, and Cybernetics Society, and technical committee of awareness computing (IEEE SMC).

Qun Jin is a professor in the Department of Human Informatics and Cognitive Sciences, Faculty of Human Sciences, Waseda University, Japan. He has been extensively engaged in research works in the fields of computer science, information systems, and human informatics, with a focus on understanding and supporting humans through convergent research. His recent research interests cover intelligent and comprehensive data analytics, personal analytics and individual modeling, trustworthy platforms for data federation, sharing, and utilization, cyber-physical-social systems, and applications in healthcare and learning support and for the realization of a carbon-neutral society. He is a foreign fellow of the Engineering Academy of Japan (EAJ).

Received: March 23, 2023; Accepted: October 23, 2023.

Computer Science and Information Systems 21(2):569–592 https://doi.org/10.2298/CSIS230803076C

Robust Compensation with Adaptive Fuzzy Hermite Neural Networks in Synchronous Reluctance Motors

Chao-Ting Chu[1] and Hao-Shang Ma[2,⋆]

[1] Chunghwa Telecom Laboratories, Internet of Things Laboratory,
No.99, Dianyan Rd., Yangmei District, Taoyuan City 32661, Taiwan, R. O. C.
chaot@cht.com.tw
[2] Department of Computer Science and Information Engineering,
National Taichung University of Science and Technology,
No. 129, Section 3, Sanmin Road, North District, Taichung City 404336, Taiwan, R. O. C.
hsma@nutc.edu.tw

Abstract. In this paper, a robust compensation scheme using adaptive fuzzy Hermite neural networks (RCAFHNN), for use in synchronous reluctance motors (SRMs), is proposed. SRMs have a simple underlying mathematical model and mechanical structure, but are affected by problems related to parameter variations, external interference, and nonlinear dynamics. In many fields, precise control of motors is required. Although the use of neural network and fuzzy are widespread, such controllers are affected by unbound nonlinear system model. In this study, RCAFHNN, based on an adaptive neural fuzzy interface system (ANFIS), was used to bound motor system model controller algorithm. RCAFHNN can be characterized in three parts. First, RCAFHNN offers fuzzy expert knowledge, a neural network for online estimation, and recursive weight estimation. Second, the replacement of the Gaussian function by the Hermite polynomial in RCAFHNN enables reduced membership function training times. Third, the system convergence and robustness compensation of RCAFHNN were confirmed using Lyapunov stability. RCAFHNN ameliorates the problems of external load and system lump uncertainty. The experimental results, in which the output responses of RCAFHNN and ANFIS (adaptive neural fuzzy interface systems) were compared, demonstrated that RCAFHNN exhibited superior performance.

Keywords: Synchronous reluctance motors, Lyapunov stability, Robust, Adaptive control, Neural network estimator, Adaptive laws.

1. Introduction

In recent years, motor control has gained significant popularity [6, 16, 19, 23]. A three-phase motor is typically supplied by a three-phase AC power source. This means there are three cables providing power, and each cable's voltage has a phase difference of 120 degrees. These three phases are referred to as Phase A, Phase B, and Phase C. However, calculating the three-phase system involves complex mathematical equations and issues related to mutual inductance coupling. Traditionally, we can employ coordinate transformation to convert the system from three-phase to two-phase, simplifying the calculations and also addressing the mutual inductance coupling issues associated with the

⋆ Corresponding author

motor. It need to address various uncertainties generated during the actual operation of the motor. Therefore, controlling alternating current in synchronous reluctance motors (SRMs) [1, 2, 8] has become a central concern. SRMs have a simple underlying mathematical model and mechanical structure but are affected by nonlinear problems such as parameter variations, external loads, and nonlinear friction. Numerous studies have explored using the controller to mitigate these problems such as robust control [27]. Moreover, a instead controller is used for controlling a SRMs using Hermite neural networks. Hermite polynomials replace traditional Gaussian functions, eliminating the need to select the vertices and widths of Gaussian functions, thereby simplifying the computational complexity. Additionally, recursive weights are employed to increase the parameters of the neural network. The Lyapunov method is used to prove that the system overcomes the cumulative uncertainties, ensuring the stability of the motor system control.

Inspiring the success of deep learning on many fields, various neural network structures have been proposed [5, 7, 10, 13, 15, 18, 28]. These research utilize the non-linear capabilities of neural network to learn and adapt to auto control. For example, an adaptive NN dynamic surface controller design for nonlinear pure-feedback switched systems with time-delays and quantized input, showing that the system's output response had satisfactory performance. A wavelet neural network sliding-mode controller [7] was used in a permanent magnet synchronous motor, where the width of the wavelet function improved neural network function. In addition, fuzzy controllers and neural networks each have distinctive advantages. Some studies have combined these two controllers to create adaptive neural fuzzy interface systems (ANFISs) [14,25,26]. ANFIS combines fuzzy expert knowledge with online neural network learning, resulting in better performance than using a simple fuzzy controller or neural network controller. In neural networks-based control systems, Gaussian functions are commonly employed. However, Gaussian functions have a limitation in that they rely on peak and width parameters, which necessitates more intricate calculations to ascertain the most suitable values for these parameters.

In this work, an RCAFHNN is proposed for use in SRMs, exhibiting satisfactory output responses in experimental results that include Laypunov functions to ensure system stability. Control inputs do not require nonlinear system parameters, and Hermite polynomials replace traditional Gaussian functions, eliminating the need to choose optimal vertices and widths. From the experimental results, we can observe that RCAFHNN offered satisfactory performance in handling lumped uncertainty and nonlinear dynamics.

The main contributions of this work are as follows.

- We propose a controller which utilizes Hermite neural networks to control synchronous reluctance motors. Hermite polynomials replace traditional Gaussian functions to simplify the computational complexity.
- A recursive weighting is used to increase neural network parameters in AFHNN, and a Lyapunov-based approach is employed to demonstrate the system's ability to overcome total uncertainty, ensuring stable control of the motor system.
- Experimental tests are conducted under various challenging conditions, including unloaded, loaded, and rotational wave commands, to evaluate the performance of the proposed controller.

The remainder of this paper is organized as follows. The mathematical model of the SRM system is presented in Section 3. The RCAFHNN is described in Section 4. The experimental results are shown in Section 5, and they demonstrate that the proposed

RCAFHNN offers fast performance and satisfactory dynamic responses when handling parameter variations and external loads. Finally, the conclusion is presented in Section 6.

2. Related Works

Studies have sought to improve the stability of nonlinear systems in robust control with neural networks [9–12]. Hsiao et al. [10] employed a neural-network-based approach with delay-dependent robust stability criteria, and they analyzed dithered chaotic systems with multiple time-delays. Huang et al. [11] presented an evolutionary radial basis function neural network combined with robust genetic-based immune computing, achieving precise command tracking in autonomous robots. In the field of motors, precise position control of sensorless PMSM [12] servo drives is required. Adaptive robust speed control with a recurrent Elman neural network can offer more precise control of a system and decrease system position errors. Gong et al. [9] also proposed robust state estimation for delayed complex-valued neural networks to consider available output measurements containing nonlinear Lipschitz-like terms.

Work environments demand precise control of drilling machines [4,29]. Self-optimizing algorithms [4] and switched-control algorithms [29] have been employed in pressure drilling and have demonstrated satisfactory performance results. Viola et al. [22] also propose a parallel enabled and stability-aware self optimizing control for using numerical twin instances during the most computationally intensive steps. Several studies have investigated fuzzy neural network sliding-mode controllers [3, 10, 30]. Fuzzy neural networks can reduce the system chattering phenomenon and can train parameters online to increase the precision of the system. Castaneda et al. [3] and Song et al. [21] used neural sliding-mode controllers in motors, and online neural network training enabled the system to overcome lumped uncertainties.

Various neural network structures have been proposed [5, 7, 10, 13, 15, 18, 28]. Hsiao et al. [10] proposed a neural-network-based approach for delay-dependent robust stability criteria for dithered chaotic systems with multiple time-delays. Niu et al. [18] proposed an adaptive NN dynamic surface controller design for nonlinear pure-feedback switched systems with time-delays and quantized input, showing that the system's output response had satisfactory performance. Additionally, Chen et al. [5] researched a rotor fault diagnosis system based on sGA-based individual neural networks, utilizing GA algorithms to search for optimal parameters to address nonlinear system issues.

A wavelet neural network sliding-mode controller [7] was used in a permanent magnet synchronous motor, where the width of the wavelet function improved neural network function. Yin et al. [24] used a Hermite neural network as an activation function. Similarly to the wavelet function, the width of the Hermite function enabled satisfactory system performance. Studies have also utilized diagonal neural networks with second-order learning algorithms [13] in system identification [20, 28] due to the faster convergence speed of second-order algorithms compared to that of first-order algorithms.

Fuzzy controllers and neural networks each have distinctive advantages. Some studies have combined these two controllers to create adaptive neural fuzzy interface systems (ANFISs) [14, 25, 26]. Yun et al. used RBFNN and ANFIS to predict the market price of electricity [25] and demonstrated that ANFIS offered accurate predictions. The power amplifier modeling conducted in [26] incorporated ANFIS to identify various effects and

different rules. Liu et al. proposed a new ANFIS structure [14] using numerical analysis and classification.

3. SynRM mathematical model

The voltage equations of the $d-q$ axis equivalent architecture in a SynRM are expressed as

$$V_{ds} = R_s i_{ds} - \omega_R L_{qs} i_{qs} + L_{ds} \frac{di_{ds}}{dt} \tag{1}$$

$$V_{qs} = R_s i_{qs} - \omega_R L_{ds} i_{ds} + L_{qs} \frac{di_{qs}}{dt} \tag{2}$$

where V_{ds} and V_{qs} are the direct and quadrature axis voltages, respectively. i_{ds} and i_{qs} are the direct and quadrature axis currents, respectively. L_{ds} and L_{qs} are the direct and quadrature inductances, respectively. R_s is the copper loss resistor. ω_R is the rotor velocity in SynRM.

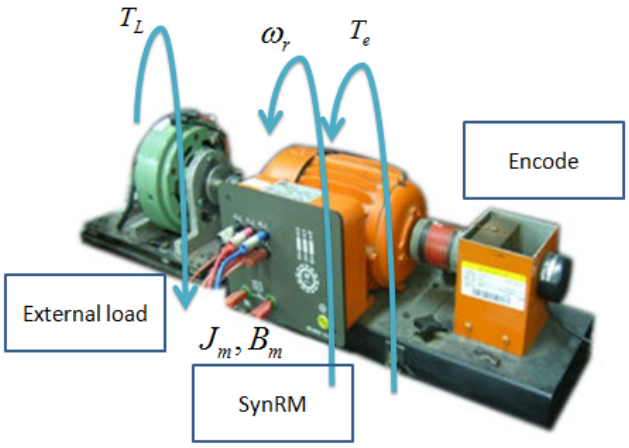


Fig. 1. Torque architecture of SRM

The torque architecture of SynRM in the mechanical equation that shows in Figure 1, and the equation is expressed as

$$T_e = J_m \frac{d\omega_R}{dt} + B_m \omega_R + T_L \tag{3}$$

where T_e is the torque of SynRM, T_L is the external load of torque, J_m is the moment of inertia, B_m is the coefficient of friction. We can rewrite the dynamic equation 3 as

$$\dot{\omega}_r = -\frac{B_m}{J_m} \omega_R + \frac{1}{J_m}(T_e - T_L) \tag{4}$$

We have electromagnetic torque equation in the rotating $d-q$ reference axis as

$$T_e = \frac{3}{4}P(L_{ds} - L_{qs})i_{ds}i_{qs} \tag{5}$$

where P is the poles in the SynRM. Therefore, the system model of SynRM is showed in Figure 2.

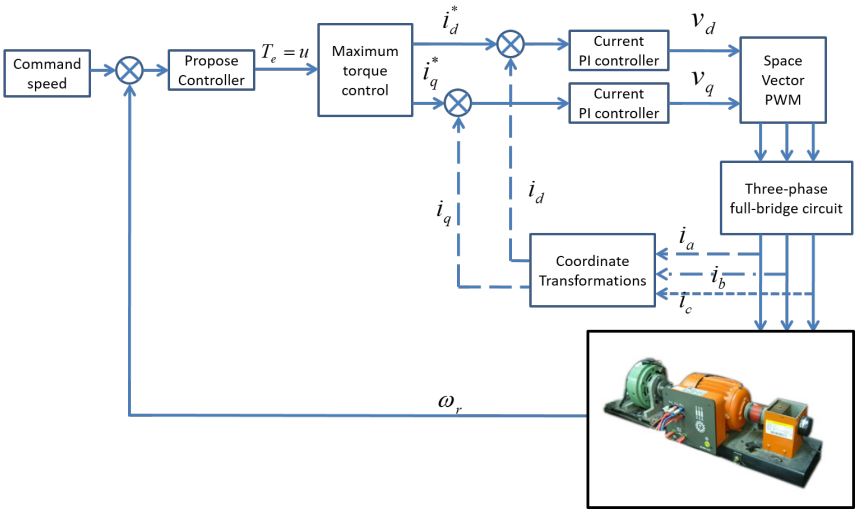


Fig. 2. Gaussian basis function neural network

4. Design of robust compensation with adaptive fuzzy Hermite neural network

4.1. SynRM nonlinear system equation

Consider a nonlinear system equation as

$$\dot{x} = f(x) + b(x)u, \tag{6}$$

$$y = x, \tag{7}$$

where $f(x)$ and $b(x)$ are unknown real continuous nonlinear functions, $u \in R$ is the control input, $y \in R$ is the system output, and $x = \omega_d \in R$ is the state vector of the system, which we assume to be available for measurement. In order to be controllable for the dynamic system, function $b(x)$ must be nonzero for vector x in certain controllability regions. Without loss of generality, we assume that $0 < b(x) < \infty$. We can rewrite the dynamic equation 6 as

$$\dot{x} = f_1(x) + b_1(x)u + E(x) \tag{8}$$

where $f(x) = f_1(x) + f_2(x)$, $b(x) = b_1(x) + b_2(x)$, $E(x) = f_2(x) + f_2(x)$, $f_1(x)$ and $b_1(x)$ are the known real continuous parameters. $f_2(x)$ and $b_2(x)$ are the unknown real continuous parameters.

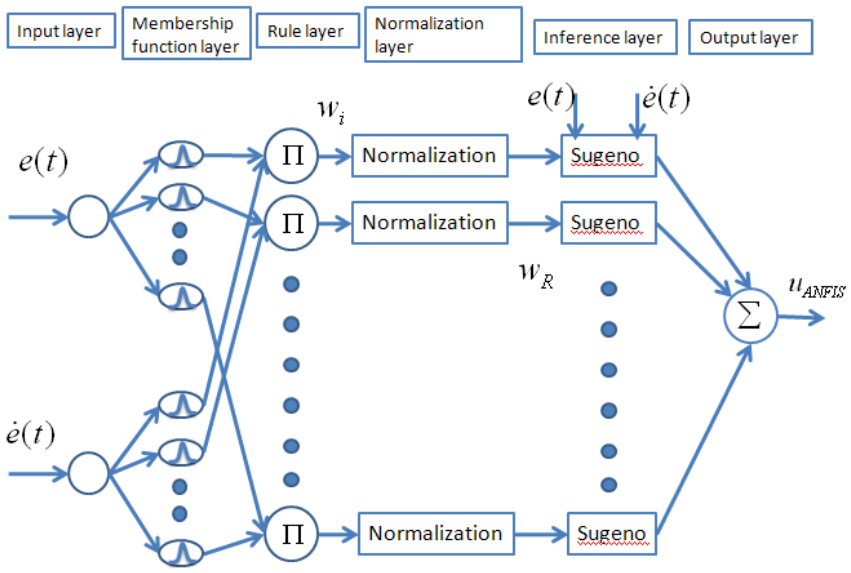


Fig. 3. The structure of ANFIS

4.2. Adaptive neural fuzzy inference system

ANFIS (Adaptive Neuro-Fuzzy Inference System) combines fuzzy expert knowledge with online neural network learning, resulting in better performance than using a simple fuzzy controller or neural network controller. The ANFIS structure is depicted in Figure 3, consisting of six layers. The first layer serves as the input layer, receiving the error signal into the network. This can be expressed by the equation:

$$y_1^1 = e(t), \quad y_2^1 = \dot{e}(t) \tag{9}$$

where $e(t) = x_d - x$, x_d is the command speed. Superscript is the n-th network, and subscript is n-th input.

The second layer is membership function layer, which is used fuzzificationcan to first layer, and the equation can be expressed as

$$y_j^2 = \exp\left[\frac{-(e(t) - v_j)^2}{2d_j^2}\right] \tag{10}$$

$$y_{j+\max j}^2 = \exp\left[\frac{-(\dot{e}(t) - v_{j+\max j})^2}{2d_{j+\max j}^2}\right] \tag{11}$$

where $\exp$ is the function of exponent, $\max j$ is the maxima of j, v_j is the Gauss function vertex, d_j is the Gauss function width, j is the j-th node.

The third layer is the rule layer, which is used logical product operator to second layer, so the output can be expressed as

$$y_i^3 = w_i = y_j^2 y_{j+\max j}^2, i = 1, 2, ..., Q \tag{12}$$

where Q is the rule number.

The fourth layer is normalization layer which is normalize to weight, and we can be expressed as

$$y_{R}^{4} = w_{R} = \frac{w_i}{\sum_{i=1}^{Q} w_i}, R = 1, 2, ..., Q \tag{13}$$

The fifth layer is the inference system, which is used Sugeno and average weighting method to defuzzification. The output can be expressed as

$$y_{R}^{5} = w_{R} f_{ANFIS}(e(t), \dot{e}(t)) = w_{R}(a_{R} e(t) + b_{R} \dot{e}(t) + c_{R}) \tag{14}$$

where $a_{R}, b_{R}, c_{R} > 0, R = 1, 2, ..., Q$ is the inference function.

The sixth layer is the output layer, which is used the linear combination of fifth layer, and the output can be expressed as

$$u_{ANFIS} = y^{6} = \sum_{R=1}^{Q} y_{R}^{5} \tag{15}$$

This paper is used the Lyapunov stability and steepest gradient method to convergence the network in ANFIS, in which we search optimal value of a_{R}, b_{R}, c_{R}. First define the Lyapunov function as

$$V_1 = \frac{1}{2} S^2 \tag{16}$$

where $S = h_1 \dot{e} + e, h_1 > 0$.

Stability criteria by the Lyapunov function, we must be $V < 0$, so that we has update equation of weight as follows

$$\Delta a_{R} = -\eta_{11} \frac{\partial \dot{V}_1}{\partial a_{R}} = -\eta_{11} \frac{\partial S \dot{S}}{\partial a_{R}} = -\eta_{11} \frac{\partial \dot{S}}{\partial a_{R}} \tag{17}$$

where η_{11} is the learning rate, $\eta_{11} > 0$,and we can rewrite equation 17 by calculus chain law as

$$\frac{\partial \dot{S}}{\partial a_{R}} = \frac{\partial \dot{S}}{\partial u_{ANFIS}} \frac{\partial u_{ANFIS}}{\partial a_{R}} \tag{18}$$

And equation 8 into equation 18, we obtain

$$\frac{\partial \dot{S}}{\partial u_{ANFIS}} = \frac{\partial h_1 \ddot{e} + \dot{e}}{\partial u_{ANFIS}} = \frac{\partial(\ddot{\omega}_d - f_1(x) - b_1(x) u_{ANFIS} - E(x) + h_0 \ddot{e})}{\partial u_{ANFIS}} = -b_1, \tag{19}$$

where $-b_{R} u_{ANFIS} = -b_{R} u$, and $\frac{\partial u_{ANFIS}}{\partial a_{R}} = \frac{w_i e(t)}{\sum_{i=1}^{R} w_i}$. Hence,

$$a_{R}(t+1) = a_{R}(t) + \Delta a_{R}(t) = a_{R}(t) + \eta_{11} S b_1 \frac{w_i e(t)}{\sum_{i=1}^{R} w_i}, \tag{20}$$

Therefore, we has update equation by b_{R} and c_{R} as

$$\Delta b_{R} = -\eta_{12} \frac{\partial \dot{V}}{\partial b_{R}} = -\eta_{12} \frac{\partial S \dot{S}}{\partial b_{R}} = -\eta_{12} S \frac{\partial \dot{S}}{\partial b_{R}}, \tag{21}$$

$$b_R(t+1) = b_R(t) + \Delta b_R(t) = b_R(t) + \eta_{12} S b_1 \frac{w_i \dot{e}(t)}{\sum_{i=1}^{R} w_i}, \tag{22}$$

where η_{12} is the learning rate, $\eta_{12} > 0$.

$$\Delta c_R = -\eta_{13} \frac{\partial \dot{V}}{\partial c_R} = -\eta_{13} \frac{\partial S\dot{S}}{\partial c_R} = -\eta_{13} S \frac{\partial \dot{S}}{\partial c_R}, \tag{23}$$

$$c_R(t+1) = c_R(t) + \Delta c_R(t) = c_R(t) + \eta_{13} S b_1 \frac{w_i}{\sum_{i=1}^{R} w_i}, \tag{24}$$

where η_{13} is the learning rate, $\eta_{13} > 0$.

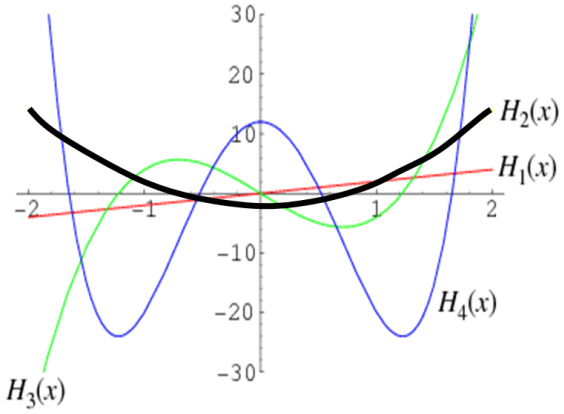


Fig. 4. Orthogonal Hermite polynomials

4.3. Robust Compensation with Adaptive Fuzzy Hermite Neural Networks

In neural networks applied to control systems, Gaussian functions are commonly employed. However, Gaussian functions have a drawback as they require parameters for their peak and width, necessitating more complex calculations to determine the optimal values for these parameters. In contrast, Hermite Polynomials have the advantage of expanding their input range with increasing order, eliminating the need for complex calculations to determine the optimal width. This not only simplifies the computational burden during system implementation but also reduces overall computational complexity.

Figure 4 displays the Orthogonal Hermite polynomials, with H_1 through H_4 representing polynomials of first to fourth order. Orthogonal Hermite polynomials exhibit a broader range compared to Gaussian functions. The paper proposes the Adaptive Fuzzy Hermite Neural Network (AFHNN), which incorporates Orthogonal Hermite polynomials, dynamic weight feedback, and robustness compensation. Finally, we employ Lyapunov stability to demonstrate system convergence. The AFHNN structure, depicted in

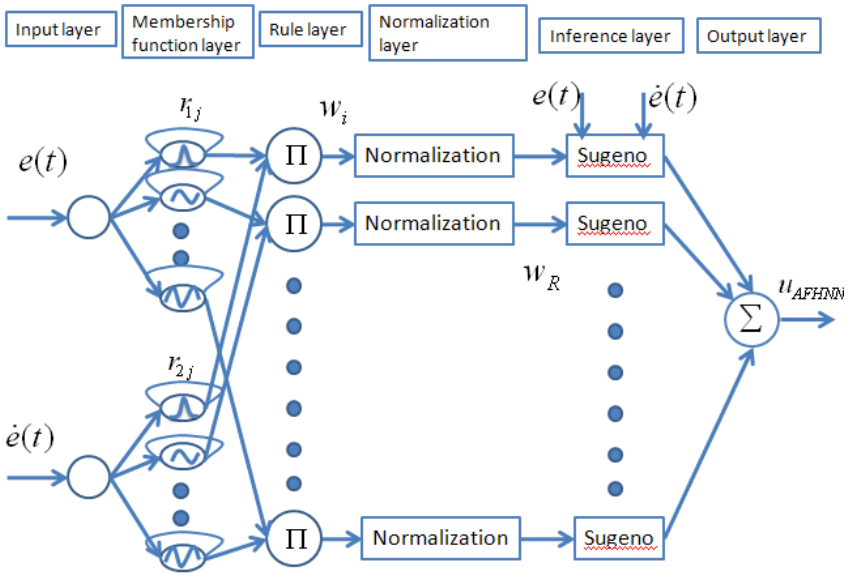


Fig. 5. The structure of AFHNN

Figure 5, consists of six layers. The first layer serves as the input layer, receiving external signals into the network. This can be expressed by the equation:

$$y_1^1 = e(t), \quad y_2^1 = \dot{e}(t), \tag{25}$$

where Superscript is the n-th network, and subscript is n-th input.

The second layer is membership function layer, in which is used fuzzification from first layer. The equation we can be expressed as

$$y_j^2 = \sigma_{j,k}(x) \tag{26}$$

$$y_{j+\max j}^2 = \sigma_{j+\max j,k}(x) \tag{27}$$

$$\sigma_{j,k}(x) = \frac{1}{\sqrt{2^j j! \sqrt{\pi}}} \exp^{-\vartheta_{j,k}^2/2} H_j(\vartheta_{j,k}) \tag{28}$$

$$\sigma_{j+\max j,k}(x) = \frac{1}{\sqrt{2^j j! \sqrt{\pi}}} \cdot \exp^{-\vartheta_{j+\max j,k}^2/2} \cdot H_{j+\max j}(\vartheta_{j+\max j,k}) \tag{29}$$

where j is the note. k is the simple time. $\vartheta_{j,k} = e(t) + r_{1j}\sigma_{j,k-1}$. exp is the exponential function. $\vartheta_{j+\max j,k} = \dot{e}(t) + r_{2j}\sigma_{j,k-1}$. $H_j(\vartheta_{j,k})$ is the Orthogonal Hermite polynomials. r_{1j} and r_{2j} is the recursive weight. $H_1(\vartheta_{j,k}) = 1, H_2(\vartheta_{j,k}) = 2\vartheta_{j,k}$, $H_n(\vartheta_{j,k}) = 2\vartheta_{j,k}H_{n-1}(\vartheta_{j,k}) - 2(n-1)H_{n-1}(\vartheta_{j,k})$ when $n \geq 3$.

The third layer is the rule layer, which is used logical product operator to second layer, so the output can be expressed as

$$y_i^3 = w_i = y_j^2 y_{j+\max j}^2, i = 1, 2, ..., Q \tag{30}$$

where Q is the rule number. The fourth layer is regularization layer, which is regulated to weight, and we can be expressed as

$$y_R^4 = w_R = \zeta_{R,k} = \frac{w_i}{\sum_{i=1}^{Q} w_i}, R = 1, 2, ..., Q \tag{31}$$

where k is the simple time of k-th. Fifth layer is the inference system, which is used Sugeno and average weighting method to defuzzification. The output can be expressed as

$$y_R^5 = w_R f_{AFHNN}(e(t), \dot{e}(t)) = \zeta_{R,k}(a_R e(t) + b_R \dot{e}(t) + c_R) = \zeta_{R,k} \varpi_R \tag{32}$$

where $a_R, b_R, c_R > 0, R = 1, 2, ..., Q$ is the inference function. The sixth layer is the output layer, which is used the linear combination of fifth layer, and the output can be expressed as

$$u_{AFHNN} = \sum_{R=1}^{n} y_R^5 = \mathbf{W^T}(\mathbf{A}, \mathbf{B}, \mathbf{C}) \cdot \boldsymbol{\varphi}(\mathbf{R_1}, \mathbf{R_2}) \tag{33}$$

where $\mathbf{W^T} = \left[\varpi_1, \ldots, \varpi_Q\right]_{1\times Q}$, $\boldsymbol{\varphi}^{\mathbf{T}} = \left[\zeta_1, \ldots, \zeta_Q\right]_{1\times Q}$, $\mathbf{A^T} = \left[a_1, \ldots, a_Q\right]_{1\times Q}$, $\mathbf{B^T} = \left[b_1, \ldots, b_Q\right]_{1\times Q}$, $\mathbf{C^T} = \left[c_1, \ldots, c_Q\right]_{1\times Q}$. $\mathbf{R_1^T} = \left[r_{11}, \ldots, r_{1j}\right]_{1\times j}$, $\mathbf{R_2^T} = \left[r_{21}, \ldots, r_{2j}\right]_{1\times j}$, $a_Q, b_Q, c_Q > 0$.

RCAFHNN used the Lyapunov function and feedback learning algorithms [24] to compensation output distribution. The control input define as

$$u = \frac{-1}{b_1(x)}\Big(-\dot{x}_1 + f_1(x) + E(x) - h_1\ddot{e}(t) + \dot{e}(t) + h_2 e(t) + h_3 \int_0^t e(t)dt\Big) = \hat{u} + \varepsilon_1 \tag{34}$$

where $\hat{u}$ is the output of RCAFHNN, ε_1 is the error between u and $\hat{u}$. In the formula of equation 34, the SRM parameters and lumped uncertainty are unknown. Therefore, we use AFHNN to track u. Substituting equation 34 to equation 8 can be obtained

$$\dot{e}(t) = -h_1\ddot{e}(t) + h_2 e(t) + h_3 \int_0^t e(t)dt + (u - \hat{u} - \varepsilon_1) \tag{35}$$

where $u - \hat{u} - \varepsilon_1 = 0$

Define the estimate error of AFHNN as

$$\begin{aligned} \widetilde{u} &= u - \hat{u} = \mathbf{W}^{*\mathbf{T}}(\mathbf{A}^*, \mathbf{B}^*, \mathbf{C}^*)\,\phi^*(\mathbf{R}_1^*, \mathbf{R}_2^*) - \hat{\mathbf{W}}^{\mathrm{T}}(\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}})\hat{\boldsymbol{\varphi}}\left(\hat{\mathbf{R}}_1, \hat{\mathbf{R}}_2\right) - u_{ss} \\ &= \mathbf{W}^{*\,\mathrm{T}}\widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{W}}^{\mathrm{T}}\hat{\boldsymbol{\varphi}} - u_{ss} \end{aligned} \tag{36}$$

where $\hat{u} = u_{AFHNN} + u_{ss}, u_{ss}$ is the control output of robustness compensation. $\widetilde{\mathbf{W}} = \mathbf{W}^* - \hat{\mathbf{W}}, \widetilde{\boldsymbol{\varphi}} = \boldsymbol{\varphi}^* - \hat{\boldsymbol{\varphi}}, \mathbf{A}^*, \mathbf{B}^*, \mathbf{C}^*$ are the approximation weight of default control input. $\mathbf{R}_1^*, \mathbf{R}_2^*$ are the approximation recursive weight of default control input. $\hat{\mathbf{A}}, \hat{\mathbf{B}}, \hat{\mathbf{C}}$ are the weight of AFHNN. $\hat{\mathbf{R}}_1, \hat{\mathbf{R}}_2$ are the recursive weight of AFHNN.

Define as

$$\begin{aligned}\tilde{\mathbf{W}} = \begin{bmatrix}\widetilde{\varpi}_1 \\ \vdots \\ \widetilde{\varpi}_Q\end{bmatrix} = \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial \mathbf{A}^{\mathbf{T}}} \\ \vdots \\ \frac{\partial \bar{\varpi}_Q}{\partial \mathbf{A}^{\mathbf{T}}}\end{bmatrix}\right|_{\mathbf{A}=\hat{\mathbf{A}}} \left(\mathbf{A}^* - \hat{\mathbf{A}}\right) + \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial \mathbf{B}^{\mathbf{T}}} \\ \vdots \\ \frac{\partial \bar{\varpi}_Q}{\partial \mathbf{B}^{\mathbf{T}}}\end{bmatrix}\right|_{\mathbf{B}=\hat{\mathbf{B}}} \left(\mathbf{B}^* - \hat{\mathbf{B}}\right) \\ + \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial \mathbf{C}^{\mathbf{T}}} \\ \vdots \\ \frac{\partial \bar{\varpi}_Q}{\partial \mathbf{C}^{\mathbf{T}}}\end{bmatrix}\right|_{\mathbf{C}=\hat{\mathbf{C}}} \left(\mathbf{C}^* - \hat{\mathbf{C}}\right) + \boldsymbol{\varphi}_{\mathbf{H2}} = \mathbf{W}_{\mathrm{A}}^{\mathrm{T}}\widetilde{\mathbf{A}} + \mathbf{W}_{\mathrm{B}}^{\mathrm{T}}\widetilde{\mathbf{B}} + \mathbf{W}_{\mathrm{C}}^{\mathrm{T}}\widetilde{\mathbf{C}} + \varphi_{\mathrm{H2}}\end{aligned} \tag{37}$$

$$\begin{aligned}\widetilde{\boldsymbol{\varphi}} = \begin{bmatrix}\widetilde{\varsigma}_1 \\ \vdots \\ \widetilde{\varsigma}_Q\end{bmatrix} = \left.\begin{bmatrix}\frac{1}{2}\left(\frac{\partial \zeta_1}{\partial \mathbf{R}_1^{\mathrm{T}}} + \frac{\partial \zeta_1}{\partial \zeta_{j,k-1}}\frac{\partial \zeta_{j,k-1}}{\partial \mathbf{R}_1^{\mathrm{T}}}\right) \\ \vdots \\ \frac{1}{2}\left(\frac{\partial \zeta_Q}{\partial \mathbf{R}_1^{\mathrm{T}}} + \frac{\partial \zeta_Q}{\partial \zeta_{Q,k-1}}\frac{\partial \zeta_{Q,k-1}}{\partial \mathbf{R}_1^{\mathrm{T}}}\right)\end{bmatrix}\right|_{\mathbf{R}_1-\hat{\mathbf{R}}_1} \left(\mathbf{R}_1^* - \hat{\mathbf{R}}_1\right) \\ + \left.\begin{bmatrix}\frac{1}{2}\left(\frac{\partial \zeta_1}{\partial \mathbf{R}_2^{\mathrm{T}}} + \frac{\partial \zeta_1}{\partial \zeta_{\mathrm{j},k-1}}\frac{\partial \zeta_{j,k-1}}{\partial \mathbf{R}_2^{\mathrm{T}}}\right) \\ \vdots \\ \frac{1}{2}\left(\frac{\partial \zeta_Q}{\partial \mathbf{R}_2^{\mathrm{T}}} + \frac{\partial \zeta_Q}{\partial \zeta_{\mathrm{Q},k-1}}\frac{\partial \zeta_{Q,k-1}}{\partial \mathbf{R}_2^{\mathrm{T}}}\right)\end{bmatrix}\right|_{\mathbf{R}_2-\hat{\mathbf{R}}_2} \left(\mathbf{R}_2^* - \hat{\mathbf{R}}_2\right) + \boldsymbol{\varphi}_{\mathrm{H1}} \\ = \varphi_{\mathrm{R}_1}^{\mathrm{T}}\widetilde{\mathbf{R}}_1 + \varphi_{\mathrm{R}_2}^{\mathrm{T}}\widetilde{\mathbf{R}}_2 + \varphi_{\mathrm{Hl}}\end{aligned} \tag{38}$$

where

$$\mathbf{W}_{\mathbf{A}} = \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial a_1} & \frac{\partial \varpi_2}{\partial a_1} & \cdots & \frac{\partial \varpi_Q}{\partial a} \\ \frac{\partial \varpi_1}{\partial a_2} & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \frac{\partial \varpi_1}{\partial a_Q} & \frac{\partial \varpi_2}{\partial a_Q} & \cdots & \frac{\partial \varpi_Q}{\partial a_Q}\end{bmatrix}_{Q\times Q}\right|_{A=\hat{A}};$$

$$\mathbf{W}_{\mathbf{B}} = \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial \mathbf{B}} & \frac{\partial \varpi_2}{\partial \mathbf{B}} & \cdots & \frac{\partial \varpi_Q}{\partial \mathbf{B}}\end{bmatrix}_{Q\times Q}\right|_{\mathbf{B}=\hat{\mathbf{B}}};$$

$$\mathbf{W}_{\mathbf{C}} = \left.\begin{bmatrix}\frac{\partial \varpi_1}{\partial \mathbf{C}} & \frac{\partial \varpi_2}{\partial \mathbf{C}} & \cdots & \frac{\partial \varpi_Q}{\partial \mathbf{C}}\end{bmatrix}_{Q\times Q}\right|_{\mathbf{C}=\hat{\mathbf{C}}};$$

$$\varphi_{\mathbf{R_1^T}} = \left[\begin{array}{c} \frac{1}{2}\left(\frac{\partial \zeta_1}{\partial \mathbf{R_1}} + \frac{\partial \zeta_1}{\partial \zeta_{1,k-1}} \frac{\partial \zeta_{1,k-1}}{\partial \mathbf{R_1}}\right) \\ \frac{1}{2}\left(\frac{\partial \zeta_2}{\partial \mathbf{R_1}} + \frac{\partial \zeta_2}{\partial \zeta_{2,k-1}} \frac{\partial \zeta_{2,k-1}}{\partial \mathbf{R_1}}\right) \\ \vdots \\ \frac{1}{2}\left(\frac{\partial \zeta_Q}{\partial \mathbf{R_1}} + \frac{\partial \zeta_Q}{\partial \zeta_{Q,k-1}} \frac{\partial \zeta_{Q,k-1}}{\partial \mathbf{R_1}}\right) \end{array} \right]_{Q \times j} \Bigg|_{R_1 = \hat{R}_1} ;$$

$$\varphi_{\mathbf{R_2^T}} = \left[\begin{array}{c} \frac{1}{2}\left(\frac{\partial \zeta_1}{\partial \mathbf{R_2}} + \frac{\partial \zeta_1}{\partial \zeta_{1,k-1}} \frac{\partial \zeta_{1,k-1}}{\partial \mathbf{R_1}}\right) \\ \frac{1}{2}\left(\frac{\partial \zeta_2}{\partial \mathbf{R_2}} + \frac{\partial \zeta_2}{\partial \zeta_{2,k-1}} \frac{\partial \zeta_{2,k-1}}{\partial \mathbf{R_1}}\right) \\ \vdots \\ \frac{1}{2}\left(\frac{\partial \zeta_Q}{\partial \mathbf{R_2}} + \frac{\partial \zeta_Q}{\partial \zeta_{Q,k-1}} \frac{\partial \zeta_{Q,k-1}}{\partial \mathbf{R_1}}\right) \end{array} \right]_{Q \times j} \Bigg|_{R_2 = \hat{R}_2} ;$$

$$\widetilde{\mathbf{A}} = \mathbf{A}^* - \hat{\mathbf{A}}; \widetilde{\mathbf{B}} = \mathbf{B}^* - \hat{\mathbf{B}}; \widetilde{\mathbf{C}} = \mathbf{C}^* - \hat{\mathbf{C}}; \widetilde{\mathbf{R}}_1 = \mathbf{R}_1^* - \hat{\mathbf{R}}_1; \widetilde{\mathbf{R}}_2 = \mathbf{R}_2^* - \hat{\mathbf{R}}_2$$

$\varphi_{\mathrm{H1}}, \varphi_{\mathrm{H2}}$ are the higher-order error in Taylor expansion.

Substituting equation 37-38 to equation 36 can be obtained

$$\begin{aligned} \widetilde{u} &= \mathbf{W}^{\star \mathrm{T}} \widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{W}}^{\mathrm{T}} \hat{\boldsymbol{\varphi}} - u_{ss} = \hat{\mathbf{W}}^{\mathrm{T}} \widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{W}}^{\mathrm{T}} \widetilde{\boldsymbol{\varphi}} + \widetilde{\mathbf{W}}^{\mathrm{T}} \hat{\boldsymbol{\varphi}} - u_{ss} \\ &= \hat{\mathbf{W}}^{\mathrm{T}} \left(\boldsymbol{\varphi}_{\mathrm{R}_1}^{\mathrm{T}} \widetilde{\mathbf{R}}_1 + \boldsymbol{\varphi}_{\mathrm{R}_2}^{\mathrm{T}} \widetilde{\mathbf{R}}_2 \right) + \left(\mathbf{W}_{\mathrm{A}}^{\mathrm{T}} \widetilde{\mathbf{A}} + \mathbf{W}_{\mathbf{B}}^{\mathrm{T}} \widetilde{\mathbf{B}} + \mathbf{W}_{\mathrm{C}}^{\mathrm{T}} \widetilde{\mathbf{C}} \right)^{\mathrm{T}} \hat{\boldsymbol{\varphi}} - u_{ss} + L_1 \end{aligned} \tag{39}$$

where $L_1 = \widetilde{\mathbf{W}}^{\mathrm{T}} \widetilde{\boldsymbol{\varphi}} + \boldsymbol{\varphi}_{\mathrm{H2}}^{\mathrm{T}} \hat{\boldsymbol{\varphi}} + \hat{\mathbf{W}}^{\mathrm{T}} \varphi_{\mathrm{Hl}}$ is the total estimation error in AFHNN. Define Lyapunov function as

$$\begin{aligned} V_2 = \frac{1}{2} S^2 &+ \frac{1}{2n_1} \widetilde{\mathbf{A}}^{\mathrm{T}} \widetilde{\mathbf{A}} + \frac{1}{2n_2} \widetilde{\mathbf{B}}^{\mathrm{T}} \widetilde{\mathbf{B}} + \frac{1}{2n_3} \widetilde{\mathbf{C}}^{\mathrm{T}} \widetilde{\mathbf{C}} \\ &+ \frac{1}{2n_4} \widetilde{\mathbf{R}}_1^{\mathrm{T}} \widetilde{\mathbf{R}}_1 + \frac{1}{2n_5} \widetilde{\mathbf{R}}_2^{\mathrm{T}} \widetilde{\mathbf{R}}_2 + \frac{1}{2n_6} \widetilde{L}^2 \end{aligned} \tag{40}$$

where $\widetilde{L} = L - \hat{L}$, L is the lump uncertainty of RCAFHNN and system, $S(t) = h_1 \dot{e}(t) + e(t), \quad n_1, n_2, n_3, n_4, n_5, n_6 > 0$.

Differential the equation (38), and subsisting equation (37), we get

$$\begin{aligned}
\dot{V}_2 &= S\dot{S} - \frac{1}{n_1}\widetilde{\mathbf{A}}^{\mathrm{T}}\dot{\hat{\mathbf{A}}} - \frac{1}{n_2}\widetilde{\mathbf{B}}^{\mathrm{T}}\dot{\hat{\mathbf{B}}} - \frac{1}{n_3}\widetilde{\mathbf{C}}^{\mathrm{T}}\dot{\hat{\mathbf{C}}} - \frac{1}{n_4}\widetilde{\mathbf{R}}_1^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_1 - \frac{1}{n_5}\widetilde{\mathbf{R}}_2^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_2 - \frac{1}{n_6}\widetilde{\hat{L}}\dot{\hat{L}} \\
&= S\left(-h_1\ddot{e}(t) + h_2 e(t) + h_3\int_0^t e(t)dt + (u - \hat{u} - \varepsilon_1) + h_1\ddot{e}\right) - \frac{1}{n_1}\widetilde{\mathbf{A}}^{\mathrm{T}}\dot{\hat{\mathbf{A}}} - \frac{1}{n_2}\widetilde{\mathbf{B}}^{\mathrm{T}}\dot{\hat{\mathbf{B}}} \\
&- \frac{1}{n_3}\widetilde{\mathbf{C}}^{\mathrm{T}}\dot{\hat{\mathbf{C}}} - \frac{1}{n_4}\widetilde{\mathbf{R}}_1^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_{\mathbf{1}} - \frac{1}{n_5}\widetilde{\mathbf{R}}_2^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_2 - \frac{1}{n_6}\tilde{\hat{L}}\dot{\hat{L}} \\
&= S\left(h_2 e(t) + h_3\int_0^t e(t)dt + \widetilde{u} - \varepsilon_1\right) - \frac{1}{n_1}\widetilde{\mathbf{A}}^{\mathrm{T}}\dot{\hat{\mathbf{A}}} - \frac{1}{n_2}\widetilde{\mathbf{B}}^{\mathrm{T}}\dot{\hat{\mathbf{B}}} - \frac{1}{n_3}\widetilde{\mathbf{C}}^{\mathrm{T}}\dot{\hat{\mathbf{C}}} - \frac{1}{n_4}\widetilde{\mathbf{R}}_1^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_1 \\
&- \frac{1}{n_5}\widetilde{\mathbf{R}}_2^{\mathrm{T}}\dot{\hat{\mathbf{R}}}2 - \frac{1}{n_6}\widetilde{\hat{L}}\dot{\hat{L}} \\
&= S\Big(\varepsilon_1 + \hat{\mathbf{W}}^{\mathrm{T}}\left(\boldsymbol{\varphi}_{\mathbf{R}_1}^{\mathrm{T}}\widetilde{\mathbf{R}}_1 + \boldsymbol{\varphi}_{\mathbf{R}_2}^{\mathrm{T}}\widetilde{\mathbf{R}}_2\right) + \left(\mathbf{W}_{\mathbf{A}}^{\mathrm{T}}\widetilde{\mathbf{A}} + \mathbf{W}_{\mathbf{B}}^{\mathrm{T}}\widetilde{\mathbf{B}} + \mathbf{W}_{\mathbf{C}}^{\mathrm{T}}\widetilde{\mathbf{C}}\right)\hat{\boldsymbol{\varphi}} - u_{ss} + h_2 e(t) \\
&+ h_3\int_0^t e(t)dt + L_1 - \varepsilon_1\Big) - \frac{1}{n_1}\widetilde{\mathbf{A}}^{\mathrm{T}}\dot{\hat{\mathbf{A}}} - \frac{1}{n_2}\widetilde{\mathbf{B}}^{\mathrm{T}}\dot{\hat{\mathbf{B}}} - \frac{1}{n_3}\widetilde{\mathbf{C}}^{\mathrm{T}}\dot{\hat{\mathbf{C}}} - \frac{1}{n_4}\widetilde{\mathbf{R}}_1^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_1 - \frac{1}{n_5}\widetilde{\mathbf{R}}_2^{\mathrm{T}}\dot{\hat{\mathbf{R}}}_2 \\
&- \frac{1}{n_6}\tilde{\hat{L}}\dot{\hat{L}}
\end{aligned} \tag{41}$$

Define as

$$L = (-\varepsilon_1 + L_1) \tag{42}$$

Therefore, we can get the adaptive law and robust compensation as

$$u_{ss} = h_2 e + h_3\int_0^t e(t)dt + k_v S + \hat{L} \tag{43}$$

$$\dot{\hat{\mathbf{A}}} = n_1 S\mathbf{W}_{\mathbf{A}}\hat{\boldsymbol{\varphi}} \tag{44}$$

$$\dot{\hat{\mathbf{B}}} = n_2 S\mathbf{W}_{\mathbf{B}}\hat{\boldsymbol{\varphi}} \tag{45}$$

$$\dot{\hat{\mathbf{C}}} = n_3 S\mathbf{W}_{\mathbf{C}}\hat{\boldsymbol{\varphi}} \tag{46}$$

$$\dot{\hat{\mathbf{R}}}_1 = n_4 S\boldsymbol{\varphi}_{R_1}\hat{\mathbf{W}} \tag{47}$$

$$\dot{\hat{\mathbf{R}}}_2 = n_5 S\boldsymbol{\varphi}_{R_2}\hat{\mathbf{W}} \tag{48}$$

$$\dot{\hat{L}} = n_6 S \tag{49}$$

As we can observe in equations (43) to (49), the input control variables used do not depend on system parameters. In other words, the proposed controller in this paper can be applied to parameterless systems as well as nonlinear systems. The use of Lyapunov convergence criteria ensures the updating of neural network parameters, overcoming uncertainties during the operation of the motor system. Replacing traditional Gaussian functions with Hermite Polynomials eliminates the need to calculate optimal peak and width parameters. Substituting equation 43-49 to 41, we have

$$\dot{V} = -k_v S^2 \leq 0 \tag{50}$$

We can know the SRM is convergence of Lyapunov function by 50. Then define as

$$\xi(t) = k_v S^2 \tag{51}$$

Integrating equation 51, we have

$$\int_0^t \xi(\tau)d\tau = V\Big(S(0)\Big) - V\Big(S(t)\Big) \tag{52}$$

Because $V\Big(S(0)\Big)$ and $V\Big(S(t)\Big)$ are bounded, hence

$$\lim_{t\to\infty} \int_0^t \xi(\tau)d\tau < \infty \tag{53}$$

According Barbalat lemma [17], we have

$$\lim_{t\to\infty} \xi(\tau) = 0 \tag{54}$$

When $t \to \infty$, then $S \to 0$ and height error $e(t) \to 0$.

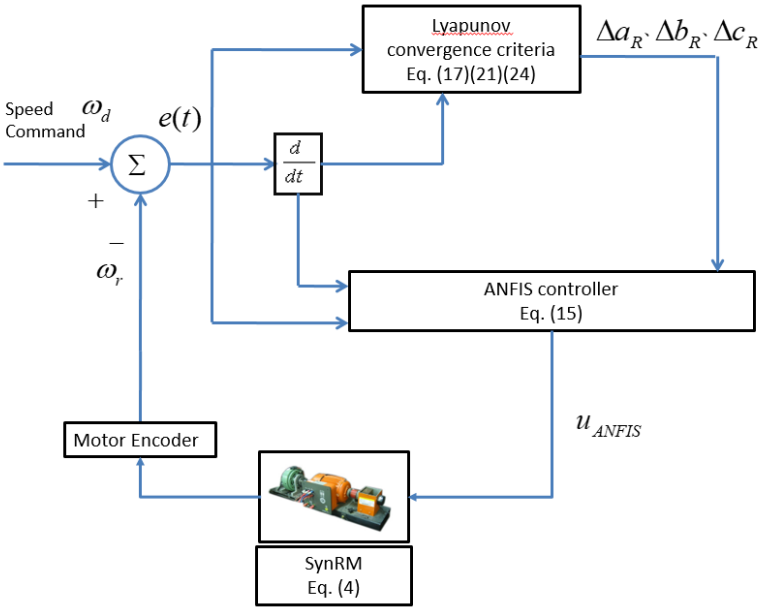


Fig. 6. Block of ANFIS

5. Experimental results

In the experiments, we aim to compare the differences between using ANFIS and the proposed neural control method in SRMs (Synchronous Reluctance Motors).

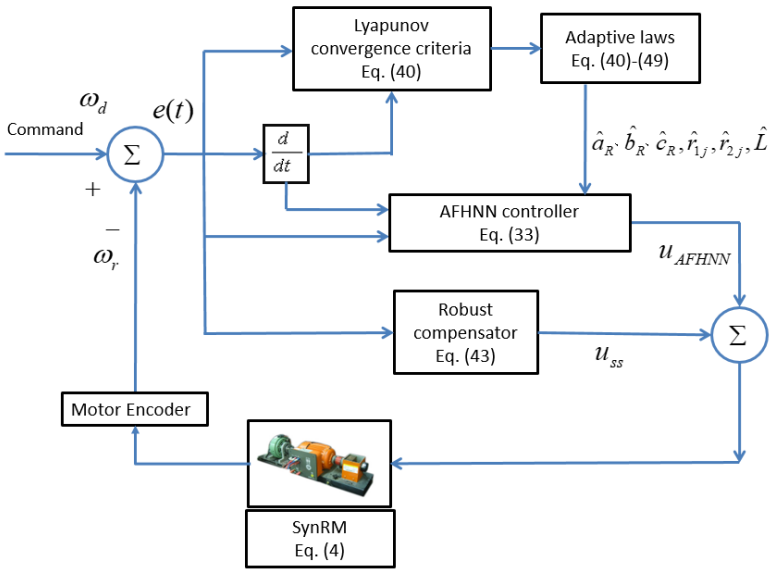


Fig. 7. Block of RCAFHNN

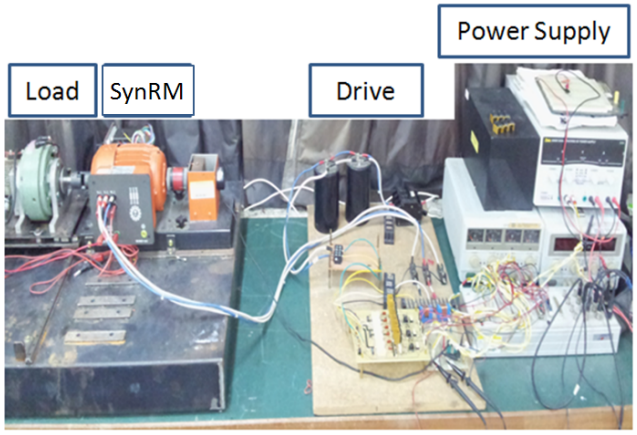


Fig. 8. Synchronous reluctance motors equipment

We have designed experiments to track motor speed errors in various demanding control scenarios during experimental testing. These scenarios include motor operation under no-load conditions, loaded conditions, and with different speed commands. We will assess the performance of the velocity controller in response to these scenarios.

The ANFIS work environment is illustrated in Figure 6. Initially, the command speed is set using a computer, and the system calculates the error between the command speed and the system output. The error signal is then fed into ANFIS, and the control input is calculated. Finally, the Lyapunov function is utilized to adjust the ANFIS weight values until the error approaches zero.

The RCAFHNN work environment is depicted in Figure 7. Similarly, the command speed is set using a computer, and the system calculates the error between the command speed and the system output. This process yields both the error and differential error signals. These signals are then input into AFHNN, and the control output is calculated to yield u_{AFHNN} and u_{ss}. Finally, the Lyapunov function is employed to adjust the AFHNN weight values until the error approaches zero, and the robust composition controller compensates for the lump uncertainty of SRM.

The RCAFHNN demonstrates an improvement in handling lump uncertainty, parameter variations, and external load in SRMs. Figure 8 illustrates the experimental SRM equipment. The controller was implemented using the ds1104 Card from dSPACE Company. The parameters utilized in this study are presented in Table 1.

Table 1. controller and SynRM parameters

Control Methods	Motor Parameters	Controller Parameters	Public Parameters
ANFIS	$J_m = 0.00076$	$\eta_1 = 0.01, \quad \eta_2 = 0.01$	$a_{1\sim9} = 2, \quad b_{1\sim9} = 50$
	$B_m = 0.00012$	$\eta_3 = 0.01$	$c_1 = -0.1, \quad c_2 = -0.1$
RCAFHNN		$h_2 = 60, \quad h_3 = 1$	$c_3 = 0, \quad c_4 = -0.1$
		$k_v = 100 \quad n_1 = 0.01$	$c_5 = 0, \quad c_6 = 0.1$
		$n_2 = 0.01 \quad n_3 = 0.001$	$c_7 = 0, \quad c_8 = 0.1$
		$n_4 = 0.001 \quad n_5 = 0.001$	$c_9 = 0.1, \quad j = 3$
		$n_6 = 0.001 \quad n_7 = 0.001$	$h_1 = 90, \quad R = 9$
		$n_8 = 20$	

Figure 9 shows the simulation output responses, error responses, A phase current comparison for initial command speed 800rpm at $0 \leq t < 5$ sec, and the changed command speed 1200rpm at $t \geq 5$ sec of ANFIS and RCAFHNN. In Figure 9, RCAFHNN can track command speed faster than ANFIS at transient response, and accurate steady-state tracking speed when the command speed is changed.

Figure 10 are the simulation output responses, error responses, A phase current comparison for time varying command speed $800+100\sin(2\pi t)$ rpm of ANFIS and RCAFHNN. In Figure 10, RCAFHNN has better tracking ability and error faster convergent.

Figure 11 is shown command speed 600rpm and initial external load is added 0.35NT-m, then we change external load is added 0.9NT-m at $t \geq 10$ of ANFIS and RCAFHNN. In figure 11, we show output response, output amplifier response, error response, A phase current comparison, neural network output and Phase plane for the error and differen-

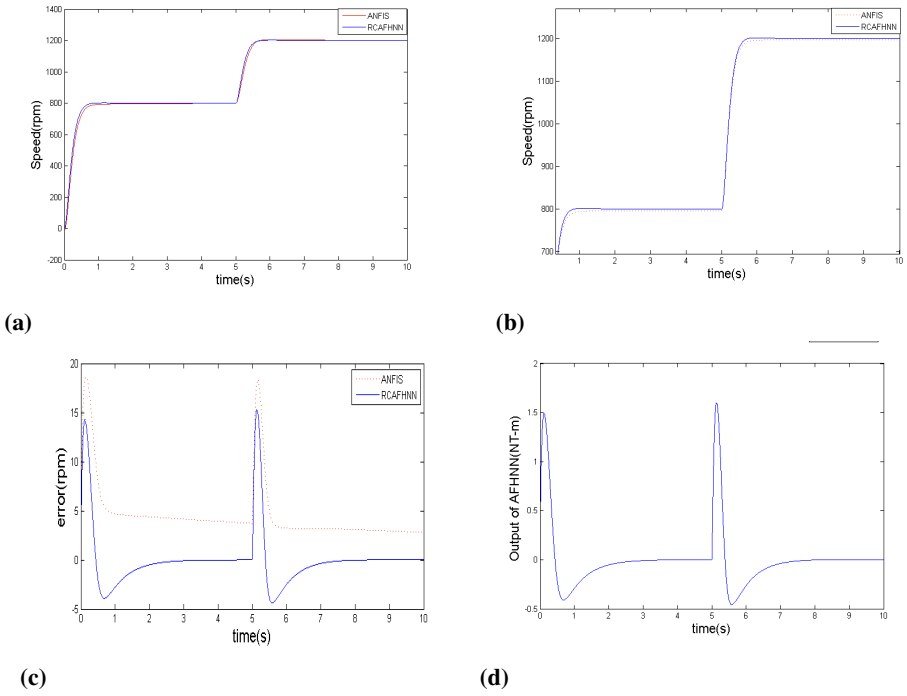


Fig. 9. Simulation responses of RCAFHNN and ANFIS for command speed 800rpm at $0 \leq t < 5$ and 1200rpm speed command at $t \geq 5$ sec (a) comparison of output responses, (b) zoomed-in comparison of output responses, (c) comparison of error responses, (d) output of AFHNN, (e) robust compensation

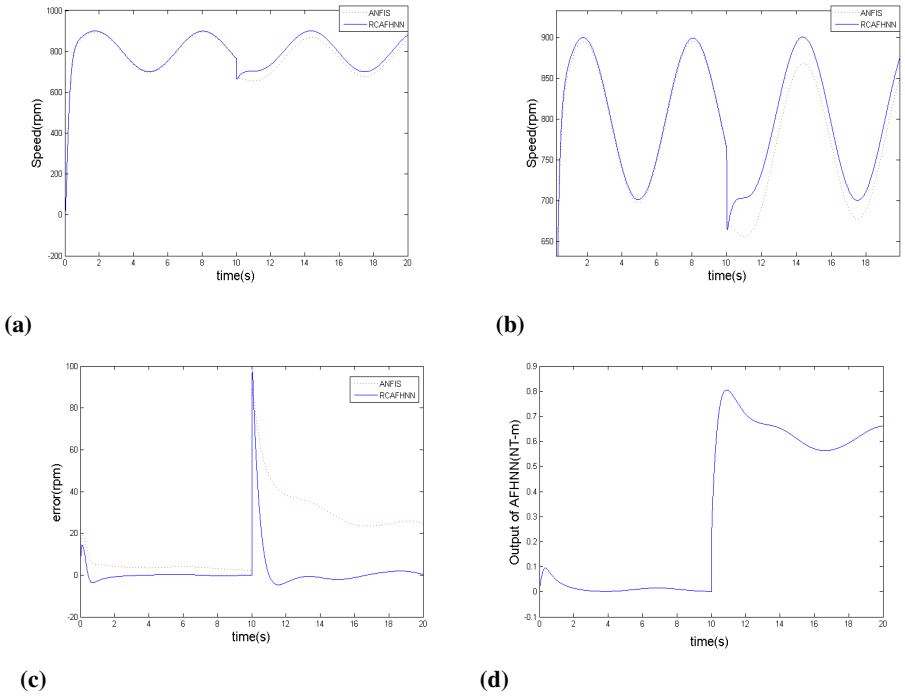


Fig. 10. Simulation responses of RCAFHNN and ANFIS at command speed $800 + 100\sin(2\pi t)$ and added external load 0.8 NT-m at $t \geq 10$ seconds (a) comparison of output responses, (b) zoomed-in comparison of output responses, (c) comparison of error responses, (d) output of AFHNN

tial error. Figure 11 (a)-(d), ANFIS tracking slowly of command speed at transient state. RCAFHNN has the faster tracking error and stability control output.

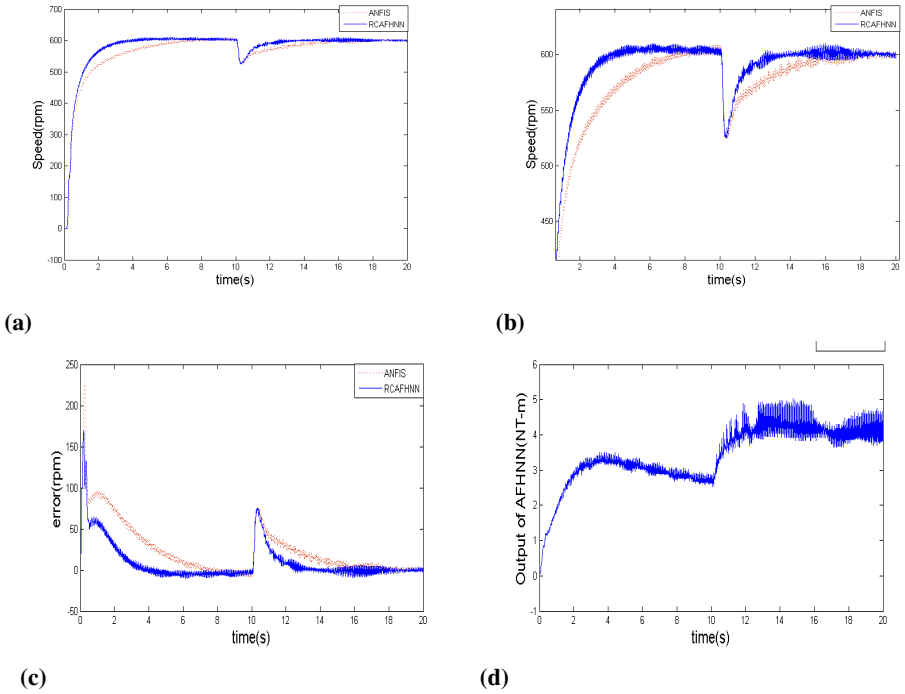


Fig. 11. Experimental responses of RCAFHNN and ANFIS at command speed 600rpm an 0.35NT-m external load is added at initial. At $t \geq 10$ seconds, an 0.9NT-m external load is added. (a) comparison of output responses, (b) zoomed-in comparison of output responses, (c) comparison of error responses, (d) Output of AFHNN

Figure 12 is shown command speed 600rpm and initial external load is added 0.35NT-m, then we change command speed 800rpm at $t \geq 5$ and external load is added 0.9NT-m at $t \geq 10$ of ANFIS and RCAFHNN. In figure 12, we show output response, output amplifier response and error response, A phase current comparison and neural network output. In figure 12, we can know that RCAFHNN has better tracking error when change the command speed and external load.

Figure 13 is shown time varies command speed $700+100\sin(2\pi t)$ rpm and initial external load is added 0.35NT-m, then we change external load is added 0.9NT-m at $t \geq 10$ of ANFIS and RCAFHNN. In figure 13, we show output response, output amplifier response, error response, A phase current comparison, and neural network output. In figure 13, RCAFHNN track the sine wave has the better ability than ANFIS, and RCAFHNN has faster track error when change external load.

Table 2 and Table 3 compares the experimental RMSEs. The performance index, RMSE, is defined as follows:

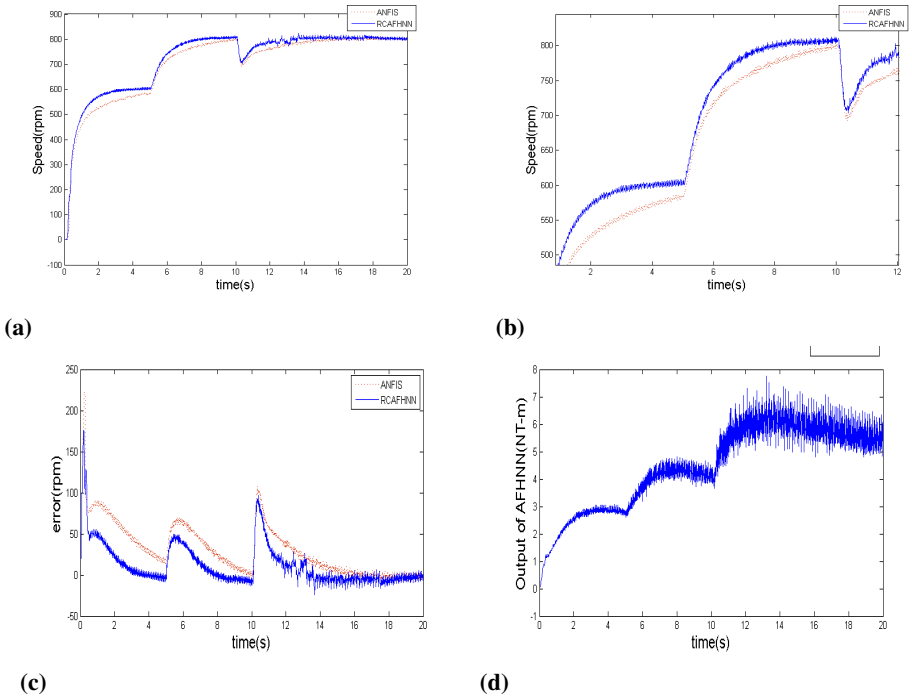


Fig. 12. Experimental responses of RCAFHNN and ANFIS at command speed 600rpm is $0 \leq t < 5$ seconds and 800rpm speed command in with an 0.35NT-m external load is added at initial. At $t \geq 10$ seconds, an 0.9NT-m external load is added. (a) comparison of output responses, (b) zoomed-in comparison of output responses, (c) comparison of error responses, (d) Output of AFHNN

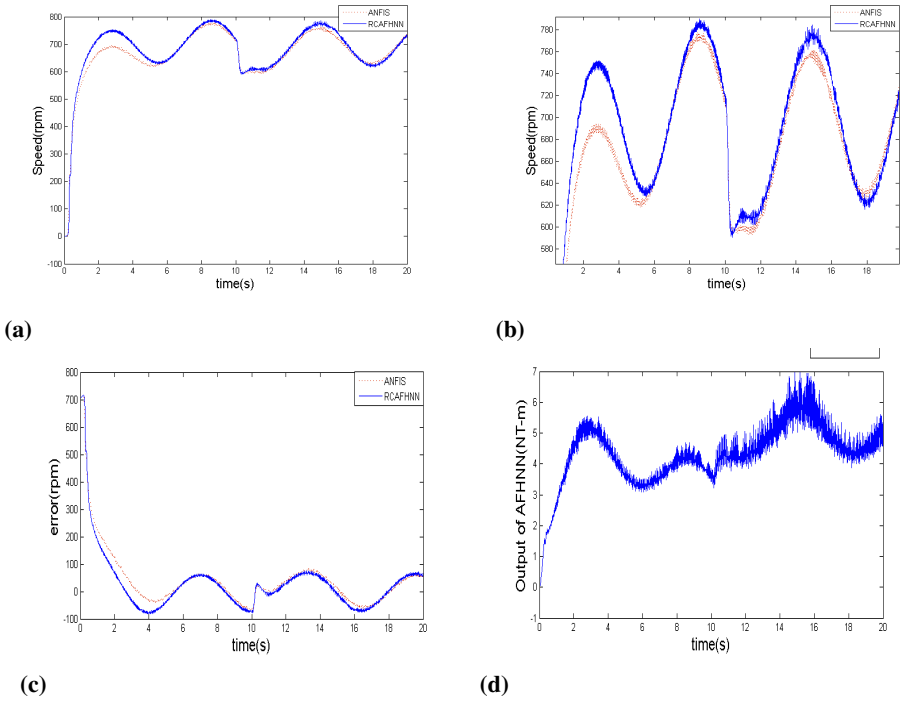


Fig. 13. Experimental responses of RCAFHNN and ANFIS at command speed $700 + 100\sin(2\pi t)$ rpm and an 0.35NT-m external load is added at initial. At $t \geq 10$ seconds, an 0.9NT-m external load is added. (a) comparison of output responses, (b) zoomed-in comparison of output responses, (c) comparison of error responses, (d) Output of AFHNN

$$RMSE = \sqrt{\frac{\sum_{i=1}^{\alpha} e^2\,[i]}{\alpha}} \tag{55}$$

where α is the number of the sampled points. Table 2 and Table 3 clearly demonstrates that RCAFHNN outperforms the ANFIS schemes under all operational conditions because of its energy control input is consider in controller. The experimental results conclusively establish the regulation ability of the proposed RCAFHNN over a wide range of speeds, its dynamic tracking capability, and its robustness.

Table 2. Simulation Comparsion of RMSE

Control Methods	800rpm to 1200rpm	800+100sin$(2\pi t)$ rpm with load
ANFIS	0.3268	1.5906
RCAFHNN	0.1831	1.2920

Table 3. Implement Comparsion of RMSE with load

Control Methods	600rpm	600rpm to 800 rpm	700+100 sin$(2\pi t)$rpm
ANFIS	40.0064	44.9906	120.2712
RCAFHNN	25.4595	28.2062	117.5688

6. Conclusion

This study successfully implemented the RCAFHNN (robust compensation scheme using adaptive fuzzy Hermite neural networks) in an SRM (synchronous reluctance motor). The RCAFHNN used adaptive laws to train weights online. Lyapunov stability was used to confirm the stability of the SRM. Moreover, the RCAFHNN offered satisfactory performance in handling lumped uncertainty and nonlinear dynamics. Finally, it can adapt to and track changes in speed and external load at transient and steady states, in spite of sine waves. Simulation and experimental results demonstrated the advantages of the proposed method.

References

1. Abootorabi Zarchi, H., Soltani, J., Arab Markadeh, G.: Adaptive input–output feedback-linearization-based torque control of synchronous reluctance motor without mechanical sensor. IEEE Transactions on Industrial Electronics 57(1), 375–384 (2010)
2. Barcaro, M., Bianchi, N., Magnussen, F.: Permanent-magnet optimization in permanent-magnet-assisted synchronous reluctance motor for a wide constant-power speed range. IEEE Transactions on Industrial Electronics 59(6), 2495–2502 (2012)

3. Castaneda, C.E., Loukianov, A.G., Sanchez, E.N., Castillo-Toledo, B.: Discrete-time neural sliding-mode block control for a dc motor with controlled flux. IEEE Transactions on Industrial Electronics 59(2), 1194–1207 (2012)
4. Cavanough, G.L., Kochanek, M., Cunningham, J.B., Gipps, I.D.: A self-optimizing control system for hard rock percussive drilling. IEEE/ASME Transactions on Mechatronics 13(2), 153–157 (2008)
5. Chen, C.S., Chen, J.S.: Rotor fault diagnosis system based on sga-based individual neural networks. Expert Systems with Applications 38(9), 10822–10830 (2011)
6. Choi, H.H., Jung, J.W., Kim, R.Y.: Fuzzy adaptive speed control of a permanent magnet synchronous motor. International Journal of Electronics 35(6) (2012)
7. El-Sousy, F.F.M.: Robust wavelet-neural-network sliding-mode control system for permanent magnet synchronous motor drive. IET Electric Power Applications 5(1), 113–132 (2011)
8. Ghaderi, A., Hanamoto, T.: Wide-speed-range sensorless vector control of synchronous reluctance motors based on extended programmable cascaded low-pass filters. IEEE Transactions on Industrial Electronics 58(6), 2322–2333 (2011)
9. Gong, W., Liang, J., Kan, X., Nie, X.: Robust state estimation for delayed complex-valued neural networks. Neural Processing Letters 46, 1009–1029 (2017)
10. Hsiao, F.H.: Neural-network based approach on delay-dependent robust stability criteria for dithered chaotic systems with multiple time-delay. Neurocomputing 191, 161–174 (2016)
11. Huang, H.C., Chiang, C.H.: An evolutionary radial basis function neural network with robust genetic-based immunecomputing for online tracking control of autonomous robots. Neural Process. Lett. 44(1), 19–35 (aug 2016)
12. Jon, R., Wang, Z., Luo, C., Jong, M.: Adaptive robust speed control based on recurrent elman neural network for sensorless pmsm servo drives. Neurocomputing 227, 131–141 (2017)
13. Kazemy, A., Hosseini, S.A., Farrokhi, M.: Second order diagonal recurrent neural network. In: 2007 IEEE International Symposium on Industrial Electronics. pp. 251–256 (2007)
14. Liu, M., Dong, M., Wu, C.: A new anfis for parameter prediction with numeric and categorical inputs. IEEE Transactions on Automation Science and Engineering 7(3), 645–653 (2010)
15. Ma, L., Khorasani, K.: Constructive feedforward neural networks using hermite polynomial activation functions. IEEE Transactions on Neural Networks 16(4), 821–833 (2005)
16. Nam, K.T., Kim, H., Lee, S.J., Kuc, T.Y.: Observer-based rejection of cogging torque disturbance for permanent magnet motors. Applied Sciences 7(9) (2017)
17. Narendra, K.S., Annaswamy, A.M.: Stable adaptive systems. Prentice-Hall (1989)
18. Niu, B., Li, H., Qin, T., Karimi, H.R.: Adaptive nn dynamic surface controller design for nonlinear pure-feedback switched systems with time-delays and quantized input. IEEE Transactions on Systems, Man, and Cybernetics: Systems 48(10), 1676–1688 (2018)
19. Rafaq, M.S., Lee, H., Park, Y., Lee, S.B., Fernandez, D., Diaz-Reigosa, D., Briz, F.: A simple method for identifying mass unbalance using vibration measurement in permanent magnet synchronous motors. IEEE Transactions on Industrial Electronics 69(6), 6441–6444 (2022)
20. Sharma, P., Ajjarapu, V., Vaidya, U.: Data-driven identification of nonlinear power system dynamics using output-only measurements. IEEE Transactions on Power Systems 37(5), 3458–3468 (2022)
21. Song, J., Wang, Y.K., Zheng, W.X., Niu, Y.: Adaptive terminal sliding mode speed regulation for pmsm under neural-network-based disturbance estimation: A dynamic-event-triggered approach. IEEE Transactions on Industrial Electronics 70(8), 8446–8456 (2023)
22. Viola, J., Chen, Y.: Parallel enabled and stability-aware self optimizing control with globalized constrained nelder-mead optimization algorithm. IEEE Journal of Radio Frequency Identification 7, 178–181 (2023)
23. Wang, H., Wang, J., Wang, X., Lu, S., Hu, C., Cao, W.: Detection and evaluation of the interturn short circuit fault in a bldc-based hub motor. IEEE Transactions on Industrial Electronics 70(3), 3055–3068 (2023)

24. Yin, K.L., Pu, Y.F., Lu, L.: Hermite functional link artificial-neural-network-assisted adaptive algorithms for iov nonlinear active noise control. IEEE Internet of Things Journal 7(9), 8372–8383 (2020)
25. Yun, Z., Quan, Z., Caixin, S., Shaolan, L., Yuming, L., Yang, S.: Rbf neural network and anfis-based short-term load forecasting approach in real-time price environment. IEEE Transactions on Power Systems 23(3), 853–858 (2008)
26. Zhai, J., Zhou, J., Zhang, L., Zhao, J., Hong, W.: Dynamic behavioral modeling of power amplifiers using anfis-based hammerstein. IEEE Microwave and Wireless Components Letters 18(10), 704–706 (2008)
27. Zhang, G., Cai, Y., Zhang, W.: Robust neural control for dynamic positioning ships with the optimum-seeking guidance. IEEE Transactions on Systems, Man, and Cybernetics: Systems 47(7), 1500–1509 (2017)
28. Zhang, J., Hou, G.: Diagonal recurrent neural networks with application to multivariable temperature control. In: 2006 1ST IEEE Conference on Industrial Electronics and Applications. pp. 1–4 (2006)
29. Zhou, J., Stamnes, O.N., Aamo, O.M., Kaasa, G.O.: Switched control for pressure regulation and kick attenuation in a managed pressure drilling system. IEEE Transactions on Control Systems Technology 19(2), 337–350 (2011)
30. Zilong, L., Guozhong, L., Jie, L.: Neural adaptive sliding mode speed tracking control of a dc motor. Journal of Systems Engineering and Electronics 15(3), 304–308 (2004)

Chao-Ting Chu graduated from the Ph.D. program in the Graduate Institute of Engineering Science and Technology at National Yunlin University of Science and Technology in 2015. Since 2016, he has been an integral part of Chunghwa Telecom Co., Ltd., specializing in IoT product service development, firmware integration, and cloud system research. Dr. Chu has been involved in a diverse range of projects, including the development of smart home systems (SmartLife), home appliances, cross-border IoT platforms, and connectivity management platforms. His contributions have significantly advanced the landscape of interconnected technologies.

Hao-Shang Ma received the B.S. and M.S. degree in Computer Science and Engineering from Yuan Ze University at 2010 and 2013 respectively. He studied in the institute of Computer and Communication Engineering for PhD in National Cheng Kung Unversity and received the PhD degree in July 2022. Currently, he is an assistant professor in Department of Computer Science and Information Engineering at National Taichung University of Science and Technology. Since January 2021, he is the Young Professionals Secretary of the Institution of Engineering and Technology (IET) - Taipei Network. His research interests include Artificial Intelligence, Data Mining, Social Network Analysis, Recommender Systems, and Nature Language Processing.

Received: August 03, 2023; Accepted: September 19, 2023.

Computer Science and Information Systems 21(2):593–623 https://doi.org/10.2298/CSIS230807077L

Machine Learning Based Approach for Exploring Online Shopping Behavior and Preferences with Eye Tracking

Zhenyao Liu[1,⋆], Wei-Chang Yeh[1*], Ke-Yun Lin[1], Hota Chia-Sheng Lin[2] and Chuan-Yu Chang[3]

[1] Integration & Collaboration Laboratory
Department of Industrial Engineering and Management Engineering
National Tsing Hua University, Hsinchu, Taiwan
liuzhenyao49@gmail.com
yeh@ieee.org
keyun924@gmail.com
[2] Department of Department of Leisure and Recreation Administration
Ming Chuan University, Taoyuan, Taiwan
hota.c.s.lin@gmail.com
[3] Medical Image Processing Laboratory
Department of Computer Science and Information Engineering
National Yunlin University of Science and Technology, Yunlin, Taiwan
chuanyu@yuntech.edu.tw

Abstract. In light of advancements in information technology and the widespread impact of the COVID-19 pandemic, consumer behavior has undergone a significant transformation, shifting from traditional in-store shopping to the realm of online retailing. This shift has notably accelerated the growth of the online retail sector. An essential advantage offered by e-commerce lies in its ability to accumulate and analyze user data, encompassing browsing and purchase histories, through its recommendation systems. Nevertheless, prevailing methodologies predominantly rely on historical user data, which often lack the dynamism required to comprehend immediate user responses and emotional states during online interactions. Recognizing the substantial influence of visual stimuli on human perception, this study leverages eye-tracking technology to investigate online consumer behavior. The research captures the visual engagement of 60 healthy participants while they engage in online shopping, while also taking note of their preferred items for purchase. Subsequently, we apply statistical analysis and machine learning models to unravel the impact of visual complexity, consumer considerations, and preferred items, thereby providing valuable insights for the design of e-commerce platforms. Our findings indicate that the integration of eye-tracking data into e-commerce recommendation systems is conducive to enhancing their performance. Furthermore, machine learning algorithms exhibited remarkable classification capabilities when combined with eye-tracking data. Notably, during the purchase of hedonic products, participants primarily fixated on product images, whereas for utilitarian products, equal attention was dedicated to images, prices, reviews, and sales volume. These insights hold significant potential to augment the effectiveness of e-commerce marketing endeavors.

Keywords: recommender systems, eye tracking, shopping preferences, machine learning, consideration factors.

⋆ Corresponding authors

1. Introduction

In recent years, the COVID-19 pandemic and the widespread adoption of computer equipment and the internet have led to a significant shift in consumer behavior, with a growing preference for e-commerce over physical retail shopping. The Ministry of Economy of Taiwan reports a steady annual increase in online sales, reaching NT$430.3 billion in 2021, a 24.5% year-on-year growth that constituting 10.8% of the total retail industry, a record high. The e-commerce sector shows continuous growth potential. Understanding consumers is critical for the success of e-commerce, which relies on three key elements: quality products, well-designed websites, and effective marketing. Successful platforms like Amazon and Netflix owe part of their triumph to their recommendation systems, which employ vast amounts of data (e.g., product data, user interactions, behavior, and personal information) and robust algorithms to predict products of interest to customers. Personalized recommendations contribute to increased sales, user satisfaction, and platform traffic, as evidenced by approximately 35% of Amazon purchases and 75% of Netflix content views originating from personalized recommendations [1]. Visual stimuli significantly impact consumer purchase intentions, accounting for 87% of sensory information received by humans [2–4]. Eye-tracking technology, utilizing advanced sensors and instruments, enables the detection of human visual activity, providing insights into consumer interests. Most e-commerce platforms rely on historical shopping and browsing data to create recommendation systems [5]. However, for new platforms or customers without such data, the absence of effective recommendations remains a challenge. Eye-tracking addresses this limitation by analyzing real-time consumer visual activity, offering precise insights into consumer psychology and behavior, thus enhancing recommendations for new customers and platforms. Recent developments in eye-tracking systems using webcams have reduced costs, making eye-tracking more prevalent [6, 7]. However, the vast amount of consumer data collected by e-commerce platforms burdens the system, prompting a shift towards machine learning and deep learning methods for more efficient data processing and analysis. This study aims to employ statistical analysis and machine learning with eye-tracking data to analyze consumers' shopping preferences and factors influencing their behavior, providing valuable insights for e-commerce platform development [8–16]. The study will collect visual activity data during online shopping using eye-tracking technology, aiming to establish a model for analyzing consumer shopping interests and validate conclusions from the literature review. Participants will wear eye-tracking devices while browsing shopping websites, and their desired purchase items will be documented. The recorded eye movement indicators and purchase choices will help achieve the study's objectives. The purpose of this study is as follows:

1. Utilize eye-tracking data combined with personal input information from participants to employ machine learning techniques in predicting participants' desired products. This would provide a reference for integrating eye-tracking data into future recommendation systems.
2. Investigate whether the complexity of product images affects eye movement indicators when participants view products. It is hypothesized that when participants view products with higher image complexity, their fixation count, fixation duration, visit duration, and visit frequency will be higher compared to products with lower image complexity.

3. Use eye movement indicators to explore participants' attention allocation to different product information during online shopping. Generally, attention level is positively correlated with fixation duration and fixation count. Therefore, this study anticipates analyzing participants' level of interest in various product information based on fixation duration and fixation count.

This research utilizes eye-tracking technology to investigate consumers' online shopping behavior and preferences, aiming to provide insights and recommendations for e-commerce platforms.
Participants in this study will wear eye-tracking devices to record eye movement data during the shopping process. After product selection, they will complete a survey to indicate their intended purchases. The data analysis section will involve examining and discussing the collected eye movement data.The research comprises six chapters. Chapter One serves as an introduction, providing background information and motivation for using eye-tracking analysis in online shopping and outlining the research objectives. Chapter Two presents a literature review, discussing past relevant studies, including eye-tracking technology and its commercial applications, machine learning classification algorithms, related eye movement classification research, and effectiveness, as well as basic recommendation system algorithms and eye-tracking applications. Chapter Three outlines the research methodology, detailing the participants, equipment, experimental procedures, data analysis, and the analysis model framework. Chapter Four showcases the experimental results, presenting the predictive effectiveness of eye movement data in determining shopping preferences, analyzing the impact of product image complexity on eye movement indicators, and exploring consumers' attention allocation during online shopping. Chapter Five discusses the results from Chapter Four, speculating on potential reasons for findings and addressing study limitations. Finally, Chapter Six presents the conclusion, summarizing the experimental findings and suggesting future research directions.

2. Related Work

2.1. Eye-Tracking Technology and Relevant Research in Business Behavior

Eye-Tracking Technology and Indicators Eye Tracker is a device that utilizes high-resolution cameras to capture human eye images at different intervals. Computer analysis software processes the eye data, allowing researchers to record human visual activity. Eye-tracking enables the observation of eye fixations, saccades (rapid eye movements between fixations), and changes in pupil size, among other information. Its applications are widespread, being used in neuroscience , human factors engineering , sports science , user experience research , and other fields to conduct further studies and investigations.This section introduces the important indicators of eye-tracking [17–23], eye-tracking technology has already been applied in a lot of different fields, Stember et al. found that eye tracking technology can generate segmentation masks for deep learning semantic segmentation in healthcare, achieving similar results to manually annotated masks, with the potential to enhance efficiency in radiology clinical workflow [24]. Nugrahaningsih et al. explored the use of gaze data to distinguish between Visual and Verbal learning styles, demonstrating a significant correlation when presenting information graphically and in

text, offering valuable insights into the application of eye tracking technology in learning styles research [25]. Eye tracking, integrated into specialized eye-tracking devices and incorporated into PC/Pad, AR/VR/XR, automobiles, and other specific equipment, has found extensive applications in fields such as scientific research, healthcare, gaming, market research, education and training, design, and manufacturing.
Area of Interest (AOI) refers to the region of interest where researchers intend to observe participants' visual movements. Saccades are the rapid movements of both eyes between fixations, while fixations involve focusing on a specific location for a certain period. Fixations are vital indicators in eye-tracking research and are closely related to attention.
Eyes possess powerful communicative abilities, and eye contact and gaze direction are central to human communication. In various fields, the above-mentioned eye-tracking indicators can be used to study and explore human behavior. Recent years have seen extensive use of eye-tracking in the field of Human-Computer Interaction (HCI) and it holds significant development potential [26]. Therefore, this research aims to utilize eye-tracking technology to investigate consumers' online shopping behavior and gain insights into human psychology through visual communication.

Eye-Tracking and Consumer Behavior Eye movement indicators, documenting consumers' visual engagement during shopping, can reveal valuable insights into their purchasing decisions. Past research has highlighted a direct correlation between high eye movement metrics (like Number of Fixations, Total Fixation Duration, Total Visit Duration) and consumer engagement, especially with particular products [27]. Furthermore, studies using these metrics have successfully predicted product attractiveness and potential purchases [28, 29].
It's noteworthy, however, that the utilization of predictive models with these metrics remains under-explored. Likewise, studies have identified gender-based differences in consumers' attention to product information and their opinion through eye movement indicators [30]. Consequently, this study aims to leverage eye movement data like fixation count and duration, and visit duration to predict consumer product interest, providing businesses with critical insights for strategic development.

Image Complexity and Eye Movement Data The eyes, acting as information conduits to the brain, are influenced by visual stimuli, affecting interpretation time and eye movement data. Visual stimuli intensity, related to stimulus complexity, can be divided into feature complexity (e.g., color, brightness), element complexity (diversity of elements, irregularity), and arrangement complexity (irregular or asymmetric arrangement). Studies show that on e-commerce platforms, product image background complexity impacts consumer attention; products with high complexity garner higher attention, while medium complexity enhances purchase intent [31]. Likewise, images with more elements increase fixation count and visit duration due to their information-rich complexity [32].
Therefore, this study investigates whether image complexity affects eye movement data, validating prior research consistency. The results will help determine image complexity as a potential factor when integrating eye movement data into recommendation systems.

2.2. Machine Learning

Supervised Learning Supervised Learning, a key machine learning branch known for its accuracy, utilizes training and test datasets [33–35]. The training dataset, comprising features and corresponding labels, aids in developing a model that can map these inputs to outputs and predict new data. This iterative learning model constantly adjusts its structure for enhanced performance. The test dataset measures the model's proficiency in predicting unknown data and checks for overfitting. Supervised learning includes regression and classification models, with the former predicting numerical values and the latter categorizing data. Given this research aims to classify consumer-interest products, a classification model is employed. Subsequent sections will explore machine learning classification models, including Decision Trees, Support Vector Machine (SVM), Random Forest, and Gradient Boosting Trees.

Decision Tree, DT The structure of a decision tree resembles an upside-down tree, composed of nodes and branches. Starting from the root node, which represents the entire sample set, each internal node represents a rule. Based on the rule's conditions, the data is branched out, and decisions are made. This process is repeated until all data is classified, and the nodes with completed branches become the leaf nodes [36]. For classification problems, decision trees often use metrics such as Information Gain, Gain Ratio, and Gini Index to evaluate the quality of branches. These metrics are explained as follows:

1. Information Gain
   First, we need to define the measure of uncertainty for a random variable, which is called entropy. Let's assume a dataset D, and the entropy of D is given by Equation 1:
   $$Entropy(D) = -\sum_{k=1}^{K} p_k \log_2 p_k \tag{1}$$
   Here, p_k represents the proportion of class k in the dataset D, and log_2 is the logarithm with base 2, which ensures that the entropy falls within the range of 0 to 1. Information Gain represents the change in entropy before and after a split. It is calculated based on a rule A that partitions the sample data D into j nodes. The number of samples in the $i-th$ node is denoted by number of D_i. The formula for Information Gain, as given by Equation 2, is used to measure the effectiveness of rule A in partitioning the samples:
   $$Gain(D, A) = Entropy(D) - \sum_{i=1}^{j} \frac{numberofD_i}{numberofD} Entropy(D) \tag{2}$$
   A larger Information Gain indicates that the rule A results in greater purity of sample partitioning. Consequently, the rule with the highest Information Gain is selected to perform the split in the decision tree.
2. Information Gain Ratio
   Information Gain prefers choosing rules that can branch into more subsets of data to maximize data purity. However, using Information Gain as an evaluation criterion for branching can lead to decision trees with reduced generalization ability, resulting in

adverse effects on classification problems. Therefore, the Information Gain Ratio is introduced as an improvement to address this issue, showing a preference for rules that branch into fewer subsets of data, as shown in Equation 3.

$$GainRatio(D, A) = \frac{Gain(D, A)}{-\sum_{i=1}^{j} \frac{numberofD_i}{numberofD} \log_2 \frac{numberofD_i}{numberofD}} \tag{3}$$

3. Gini Coefficient
   The Gini coefficient is another method for calculating impurity.

$$Gini(\mathrm{D}) = 1 - \sum_{k=1}^{K} p_k^2 \tag{4}$$

Support Vector Machine SVM's key principle involves using kernel functions to project low-dimensional inseparable data into high-dimensional space, where it locates an optimal hyperplane that efficiently distinguishes different classes of data [37, 38]. Additionally, SVM strives to optimize the margin of separation, ensuring the largest possible boundary region. Its mathematical solution is as follows:

$$\max_{\mathrm{w}} \left\{ \frac{2}{\|w\|} \right\} subjecttoy_i \left(w^T x_i + \gamma I \right) \geq I, \forall i = 1, \ldots, n \tag{5}$$

The support vector machine (SVM) model can be viewed as an optimization problem, where the equation $w^T x_i + \gamma I$ represents the separating hyperplane. The objective is to maximize the margin of separation while ensuring the ability to classify different types of data, as shown in Equation 5.

Random Forest Random Forest's classification result of each tree is resolved via majority voting, determining the final outcome [39]. As part of the bagging algorithm [40], Random Forest applies the law of large numbers and random ensembles, significantly mitigating the risk of decision tree overfitting.

Extreme Gradient Boosting XGBoost generates trees in a sequential manner. The decision trees generated later are focused on reinforcing the learning and correcting errors from the previous trees, creating interdependence among the trees. Additionally, XGBoost incorporates regularization terms $L1/L2Regularization$ into its objective function to control the model's complexity and reduce the risk of overfitting [41]. Below is a brief explanation of the objective function used in XGBoost:

$$Obj(t) = \sum_{i=1}^{n} l \left(y_i, \widehat{y}_i^{t-1} + f_t \left(x_i \right) \right) + \Omega \left(f_t \right) + constant \tag{6}$$

The objective function of Extreme Gradient Boosting (XGBoost) model consists of two components, namely the loss function l and the regularization function Ω. The loss function is used to measure the error between actual values and predicted values, while the regularization function serves as a penalty term to control the model's complexity and prevent overfitting.

Eye Tracking Data and Machine Learning Eye tracking data analysis has increasingly incorporated machine learning algorithms in recent years. Schweikert et al. employed AdaBoost, Mixed Group Ranks (MGR), RF, and Multi-layer Combinatorial Fusion (MCF) to predict image attractiveness using visual data such as the final 200 milliseconds of fixation time, total visit duration, and movement count between facial features. The precision of AdaBoost and RF was 0.938 and 0.949, respectively, signifying both ensemble algorithms' accuracy in predicting such data. The MCF algorithm also outperformed MGR, indicating its potential for further refinement [42]. Additionally, machine learning has been used with eye tracking data in business, with Pfeiffer et al. utilizing algorithms like LR, RF, and SVM to differentiate between goal-directed and exploratory search behaviors in physical and VR shopping scenarios. Notably, SVM excelled in classification accuracy, with all three algorithms achieving over 70% accuracy and demonstrating efficacy in small sample sizes [16]. The studies underscore machine learning's competence in classifying eye tracking data and its enhanced interpretability relative to deep learning. These algorithms not only rank indicator importance, aiding in identifying critical predictive factors, but also offer profound managerial insights. Hence, this study seeks to use machine learning to classify eye tracking data in consumer research.

2.3. Recommendation System

Collaborative Filtering Recommendation Collaborative Filtering (CF) is a recommendation method based on a user's past purchase behavior and ratings given to products within the system, as well as the collaborative behavior and ratings of other users. Through algorithms, it calculates the similarity of users' preferences and provides product recommendations accordingly [43–45]. As shown in Figure 1, if User A has purchased and rated products 1, 2, and 4 positively, and User C has purchased products 1 and 4 and given similar ratings as User A, then collaborative filtering algorithms identify the similarity in preferences between User A and User C. As a result, the system automatically recommends product 2 to User C. Collaborative filtering recommendation has the advantage of being able to recommend suitable products based on consumers' preferences. However, it also faces two major problems: firstly, if the majority of users have not rated the products, there will be a lack of essential recommendation basis, resulting in the sparsity problem; secondly, when new users enter the system without past purchase and rating history, or when new products have no ratings, collaborative filtering recommendation lacks historical information and becomes ineffective, which is known as the cold start problem [5].

Content-Based Filtering Recommendation Content-Based Filtering (CBF) is a recommendation method that relies on the features of products themselves, the products that users search for, the features of previously purchased products, and the information provided by users when they join the platform. Through algorithms, it calculates the preferred product features of users and generates recommendations for products that users might be interested in, as depicted in Figure 2. Content-Based Filtering does not require the use of other users' data and solely relies on the comparison and recommendation of individual users' preferences and product features. Therefore, during the early stages of platform construction with limited user data and product ratings, Content-Based Filtering can effectively address the sparsity problem and cold start problem encountered in collaborative

filtering recommendations. However, Content-Based Filtering has two main drawbacks: Firstly, since it calculates recommendations based on product features, it necessitates the appropriate and comprehensive definition of features for each product [46]. Secondly, Content-Based Filtering's primary limitation lies in using consumer preferences for product features as the basis for recommendations, which tends to recommend products of the same type. Consequently, new products with unique features might not be effectively recommended, leading to a lack of exposure to diverse products, known as the Over Specialization Problem [47].

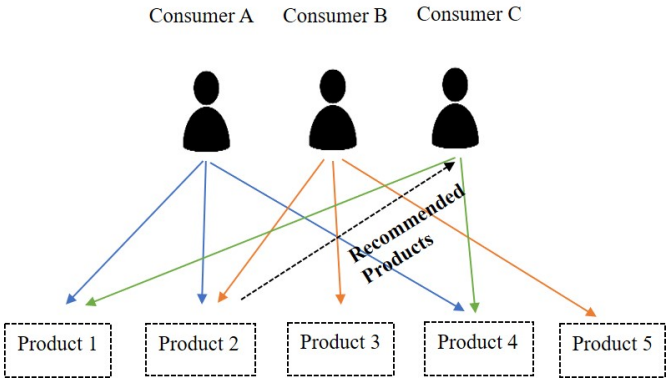


Fig. 1. Collaborative filtering recommendation

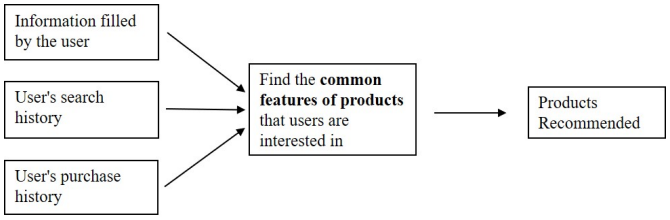


Fig. 2. Content-based filtering recommendation

Hybrid Recommendation and Eye Tracking Applications In response to the inherent constraints of single recommendation methods, research has focused on Hybrid Recommendation [48], combining different algorithms to improve basic systems. For instance, Basiri et al. utilized the Ordered Weighted Averaging (OWA) algorithm [49] to calculate weights for five classifiers, effectively addressing the cold start problem for new users or products [50]. Walek and Fojtik, in 2020, introduced a hybrid method incorporating an expert system for final ranking, which outperformed traditional methods in movie recommendations [51]. While existing website-based recommendation systems lack dynamic

channels for capturing real-time user experiences, the maturation of eye-tracking technology has offered deeper insights into user behavior. Hence, recent studies have begun integrating eye-tracking indicators into systems for more precise recommendations. For example, Song and Moon combined gaze indicators and social behavior data into their recommendation model [52], and other researchers have used webcams to record users' eye movements and facial expressions while viewing products to offer tailored recommendations [53].

These studies highlight the evolution of recommendation systems, incorporating multiple methods, including eye-tracking, to improve accuracy and user satisfaction. This integration offers a unique approach to predicting consumer interests, ensuring a more personalized user experience.

3. Research Method

3.1. Research Subjects

60 participants, devoid of eye disease history and color blindness, aged 18-35 with a minimum corrected visual acuity of 0.8, were recruited for this study, regardless of gender. Participants were sourced via social media networks. Prospective participants filled out an online form detailing the experiment's location, content, procedures, and potential risks. This ensured participant understanding prior to commitment to participation. Additionally, the form surveyed participant's eye health, contact information, and experiment scheduling availability. Suitable participants were chosen based on the form responses, and subsequently contacted for further arrangements. The study was ethically approved by the Research Ethics Committee of National Tsing Hua University.

3.2. Experimental Equipment

The Ergoneers Dikablis Glasses 3 eye-tracking system, depicted in Figure 3, was employed to monitor eye positions and document eye movements in this study. The eye-tracking system comprises a front camera (field/scene camera) capturing the environment and dual eye cameras recording binocular movements. The front camera, operating at 30 fps, records the participant's field of view with a resolution of $1920*1080$ pixels. The eye cameras, functioning at $60\,\mathrm{Hz}$ with a $648*488$ pixel resolution, permit exact participant eye movement tracking. The eye-tracking system is connected to the computer, and the information recorded by the front and eye cameras is transmitted to the computer. The system utilizes two-dimensional barcode (Marker) technology as the calibration reference for Areas of Interest (AOI). In this study, we aim to observe participants' visual activities during online shopping, with the focus on their gaze within the computer screen. Therefore, AOIs will be set on the information displayed on the computer screen.

3.3. Experimental Procedure

The experimental setup, conducted in an indoor laboratory, is depicted in Figure 4. The primary experimenter readies the experimental environment before participant involvement. Participants are subsequently familiarized with the eye-tracking device and calibrated to ensure precise eye movement capture. Upon verification of successful visual

Fig. 3. Ergoneers Dikablis Glasses 3 Eye Tracker (Source: Ergoneers)

activity capture, the team delineates the experimental purpose, procedure, potential risks, benefits, and data management to participants. Participants are requested to sign a consent form following explanation, preceding the actual experiment.
The experiment primarily aims to amass eye-tracking data and evaluate participant product choices during a shopping task. Three product categories, shoes, clothes, and earphones, are utilized to gauge the performance of machine learning models across diverse product categories. Test groups for shoes and earphones are segmented into low and high image complexity subgroups, to further explore image complexity impact. Participants don the eye-tracking device during the experiment, recording their eye movements while making product selections, before proceeding to subsequent product tests. Upon completion of all product category experiments, the research team facilitates eye-tracking device removal, signifying the conclusion of the experiment.

3.4. Experimental Design

Experimental Material Selection Three daily-use products, shoes, clothes, and earphones, were selected for this study, categorized based on their type. Dhar and Wertenbroch's research shows that consumer buying decisions are influenced by hedonic and utilitarian consideration [54], thus allowing for a classification into hedonic and utilitarian goods. Utilitarian goods, including items like earphones, are characterized by functional utility, with consumers prioritizing aspects such as functionality, quality, and price. In contrast, hedonic goods, such as clothes, offer experiential consumption, providing pleasure and enjoyment. The experimental products, shoes and clothes (hedonic goods) and earphones (utilitarian goods), were classified to investigate the variation in consumers' attention to product information due to differing product attributes. To account for previous research showing the influence of image complexity on eye-tracking data and to ensure the accuracy of machine learning eye-tracking models, three product images of similar complexity were selected for each product category. The study aims to assess the impact of image complexity on eye-tracking metrics. The test groups for shoes and earphones

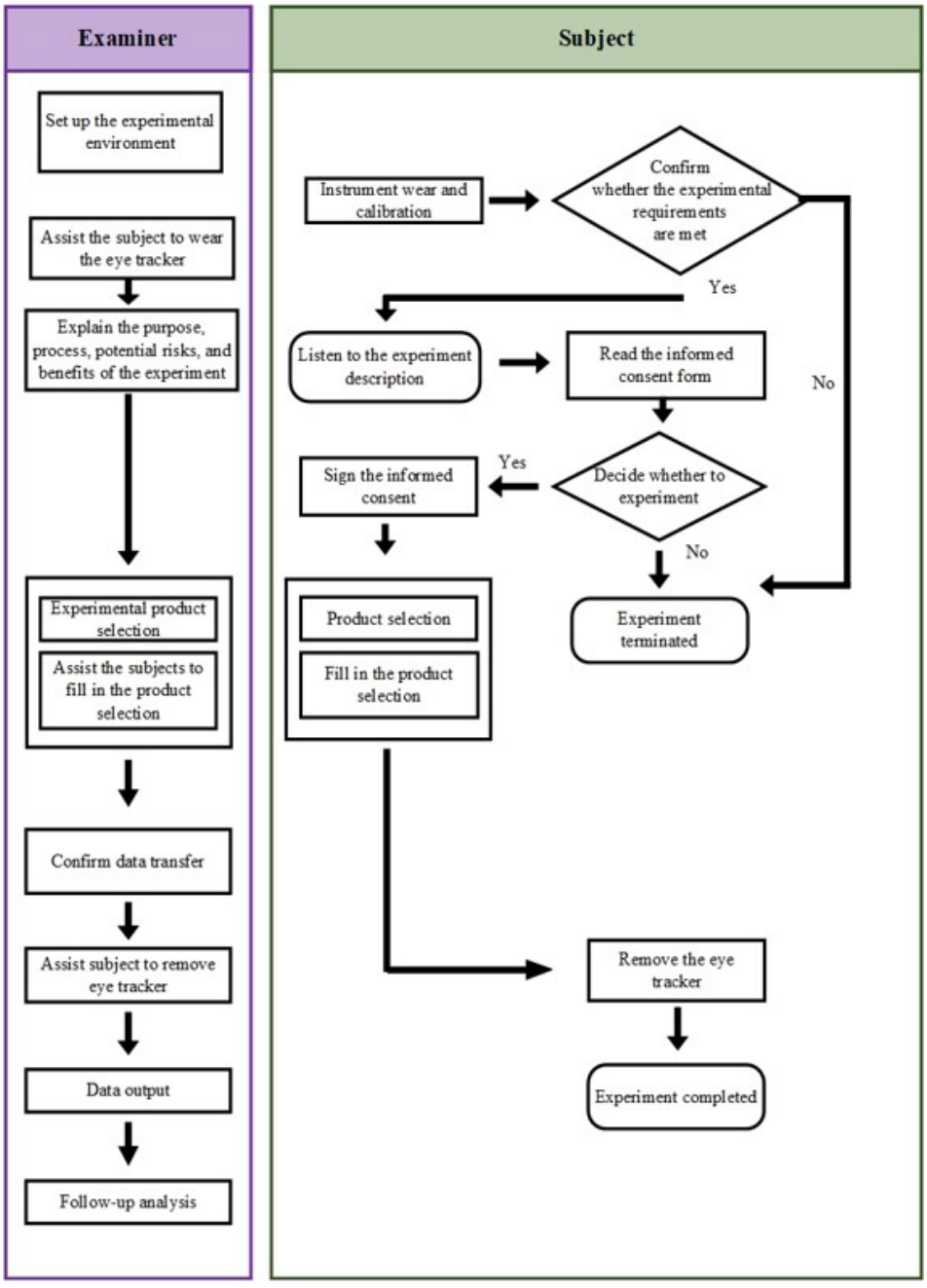


Fig. 4. Experiment Flowchart

were subdivided into low and high image complexity subgroups for comparative analysis of eye-tracking data. Following Qiuzhen et al.'s research, this study defines image complexity through feature complexity, element complexity, and arrangement complexity. Images with low complexity feature only the product in Figure 5, while those with high complexity contain more than four elements and colors, arranged irregularly and diversely in Figure 6.

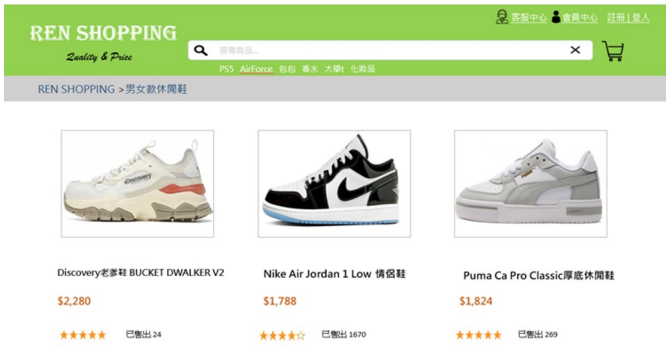


Fig. 5. Products with low image complexity

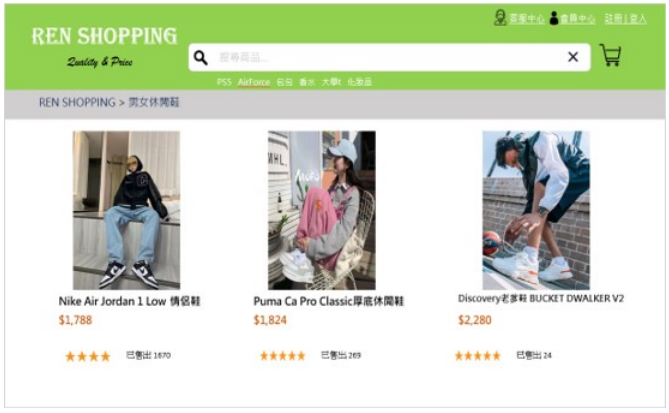


Fig. 6. Products with high image complexity

Eye-tracking Data The main Areas of Interest (AOIs) in this experiment will be set to the product information displayed on the screen, which can be divided into four major areas: all product information, product images, product prices, and product ratings and sales volume, as shown in Figure 7. The shaded regions represent the AOI areas. Subsequent

analysis will utilize participants' gaze and visit data within these AOIs to observe their visual activities and attention allocation during the shopping task. Specifically, for the eye-tracking recommendation data, the large AOI covering all product information (gray region in Figure 7) will be selected as the basis for analysis. For the experiment on image complexity and eye-tracking data, the data within the AOI of product images (blue region in Figure 7) will be used for analysis. For the experiment analyzing attention allocation with eye-tracking data, data from three AOIs will be used: product images (blue region in Figure 7), product prices (orange region in Figure 7), and product ratings and sales volume (green region in Figure 7). The data used for analysis in this experiment were

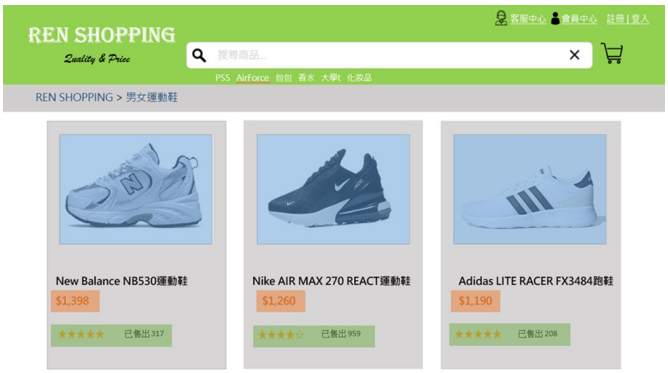


Fig. 7. Product information AOI

obtained from the D-LAB analysis software. The data description is as follows:

1. Session Duration: The time taken to complete a task, which in this study can be considered as the time taken for product selection.
2. Number of Glances: The frequency of visits to the Areas of Interest (AOIs).
3. Total Glance Time: The overall time spent visiting the AOIs.
4. Glance Location Probability: This metric compares the attention distribution among different AOIs as the formula 7 shows:

$$GlanceLocationProbability = \frac{NumberofGlancestoanAOI}{\sum NumberofGlancestoAOI1, AOI2} \tag{7}$$

5. Number of Fixations: The frequency of fixations or instances where the gaze is fixated on a particular point.
6. Total Fixation Time: The cumulative duration of all fixations, representing the total time spent with gaze fixed on various points of interest.

Experimental Environment Design The present experiment simulates the environment of consumers shopping online and to ensure that the eye tracker can accurately capture the entire website interface, participants' eye distance from the screen is controlled to be approximately 50 centimeters. Additionally, the height of the chair will be adjusted according to the participants' different heights, as shown in Figure 8.

Fig. 8. Experimental Environment

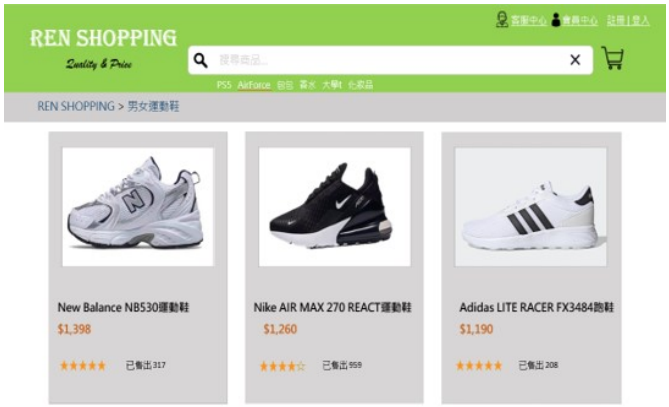


Fig. 9. AOI of the forecasting model

3.5. Machine Learning Eye Movement Prediction Model

This study aims to predict participant product choices using eye movement data within various product Areas of Interest (AOIs). Three products, each associated with its AOI, are selected for the test group. Eye movement data within these AOIs will inform the machine learning model, facilitating the development of a recommendation model grounded in eye movement behavior in Figure 9.

Given the diverse input data and categorical product choice data, a machine learning model is employed for classification prediction. A supervised learning approach, utilizing a Multiclass Classification model, is employed to predict product interest. Feature data comprises preprocessed eye movement data, including total glance time, total glance frequency, AOI attention ratio, and other relevant metrics aligned with product eye movement behavior. Participant purchase choices serve as model labels during analysis, resulting in preprocessed eye movement data as input and predicted product interest as output in Figure 10. The study classifies products of interest into three categories, influenced by participant preferences, potentially leading to imbalanced data and lower prediction performance. To address potential imbalance, the Synthetic Minority Oversampling Technique (SMOTE) will be used to augment minority class data before applying machine learning classification models. Previous literature reveals promising results in classifying eye-tracking data using Decision Trees (DT), Support Vector Machines (SVM), and Random Forests (RF) [16, 42, 55, 56]. In addition, the XGBoost (XGB) classification model is commonly used in recent machine learning competitions. This study will apply and compare four different classifiers - DT, SVM, RF, and XGB. Model validation will be executed using a Confusion Matrix to assess performance. In the Confusion Matrix, True Positive (TP) and True Negative (TN) represent correctly predicted positive and negative instances, respectively, while False Positive (FP) and False Negative (FN) indicate incorrect positive and negative predictions. The study uses Accuracy, Precision, Recall Rate, and F1-Score as performance metrics. Accuracy represents the ratio of correctly classified instances, Precision indicates the ratio of correct positive predictions, Recall Rate defines the proportion of correct positive classifications among actual positives, and F1-Score is the harmonic mean of Precision and Recall Rate.

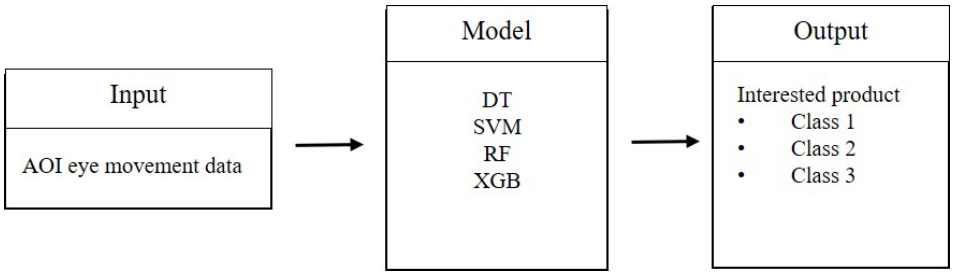


Fig. 10. Diagram of machine learning data

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \tag{8}$$

$$Precision = \frac{TP}{TP + FP} \tag{9}$$

$$Recall = \frac{TP}{TP + FN} \tag{10}$$

$$F1 - Score = \frac{2}{\frac{1}{Precision} + \frac{1}{Recall}} \tag{11}$$

3.6. Statistical Analysis

Eye-tracking data and Image Complexity Wang's web design study suggests product images with greater background complexity draw more consumer attention, due to the multitude of features influencing consumer cognitive processing and fluency, resulting in extended time spent understanding the product. Thus, products with higher background complexity yield greater fixation duration and frequency than those with less complexity [31]. Vu et al. observed a significant increase in both fixation frequency and visit duration as the number of image elements increased, as larger and more complex information requires increased processing time [32]. Building upon these findings, this experiment seeks to explore the impact of image complexity on eye-tracking data within e-commerce platforms. Image complexity is thus categorized into low and high groups, with experiments performed using images from each complexity level. The study analyzes eye-tracking indicators including fixation duration, fixation frequency, visit duration, and visit frequency. Eye-tracking data from the two complexity groups are compared to discern differences in eye movement patterns. For the eye-tracking data and shoe image complexity, the following hypothesis H1 is proposed: Participants will focus more attention on shoe images with higher background complexity. Subsequently, the following individual hypotheses (H1a, H1b, H1c, H1d) are proposed for the shoe group eye-tracking data:

1. *H1a: As the complexity of shoe images increases, consumers' visit duration also increases.*
2. *H1b: As the complexity of shoe images increases, consumers' fixation duration also increases.*
3. *H1c: As the complexity of shoe images increases, consumers' visit frequency also increases.*
4. *H1d: As the complexity of shoe images increases, consumers' fixation frequency also increases.*

Likewise, for the eye-tracking data and earphone image complexity, the hypothesis H2 is proposed: Participants will focus more attention on earphone images with higher background complexity. Subsequently, the following individual hypotheses (H2a, H2b, H2c, H2d) are proposed for the earphone group eye-tracking data:

1. *H2a: As the complexity of earphone images increases, consumers' visit duration also increases.*
2. *H2b: As the complexity of earphone images increases, consumers' fixation duration also increases.*
3. *H2c: As the complexity of earphone images increases, consumers' visit frequency also increases.*

4. *H2d: As the complexity of earphone images increases, consumers' fixation frequency also increases.*

For the observation of image complexity and eye-tracking data, this study employs paired-samples t-tests. The eye-tracking data for each group are obtained by summing the visit duration, fixation duration, visit frequency, and fixation frequency of the three product images in each group. A significance level of 0.05 is used for the comparison of eye-tracking data between the groups, as shown in Figure 11.

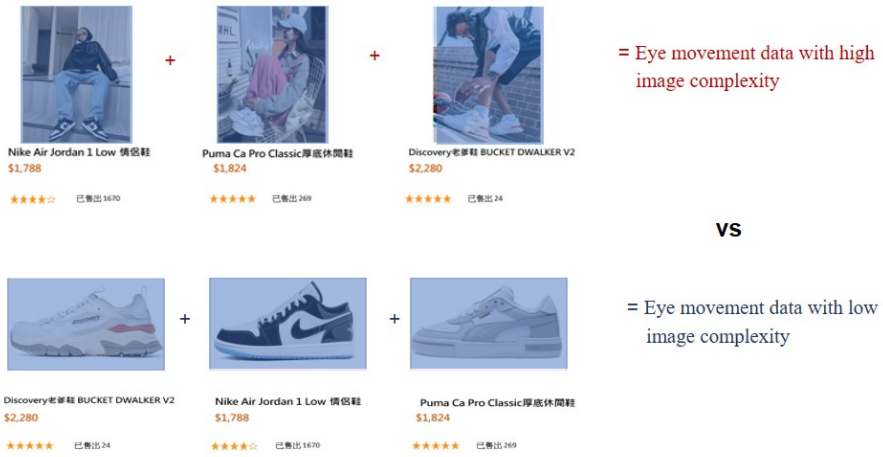


Fig. 11. Diagram of image complexity calculation

Eye Movement Data and Purchase Consideration Factors Hwang and Lee conducted eye-tracking research to investigate consumer attention allocation during online shopping. The results showed that consumers' highest attention was on product information, including product images, product prices, and product descriptions. The next highest attention was on consumer opinions [28], but there was no further exploration of individual product information such as product images and prices. Therefore, this experiment aims to use eye-tracking to further study consumer attention allocation to individual product information when shopping online. Individual product information includes product images, product prices, product ratings, and sales volume, these three major aspects. In this study, we defined separate Areas of Interest (AOIs) for these three pieces of information, as shown in Figure 12. We intend to use Total Fixation Duration (TFD) and Number of Fixations (NF) within these three AOIs as indicators of participant attention to observe their attention allocation during shopping.

Since this study categorizes shoes and clothing as hedonic products, it is hypothesized that when consumers shop for these two categories, they primarily consider the appearance of the product, followed by factors such as price and ratings. Hypotheses H3 and H4 are proposed: For shoes and clothing, participants' attention to product images will be

greater than their attention to product prices, ratings, and sales volume. Attention to product prices and ratings, as well as sales volume, will be equal. Furthermore, since attention is composed of both the time spent viewing and the number of times viewed, hypotheses are proposed for Total Fixation Duration (TFD) and Number of Fixations (NF): For shoes (H3a, H3b) and clothing (H4a, H4b):

1. *H3a: TFD for images > TFD for prices = TFD for ratings and sales volume*
2. *H3b: Number of fixations (NF) for images > NF for prices = NF for ratings and sales volume*
3. *H4a: TFD for images > TFD for prices = TFD for ratings and sales volume*
4. *H4b: Number of fixations (NF) for images > NF for prices = NF for ratings and sales volume*

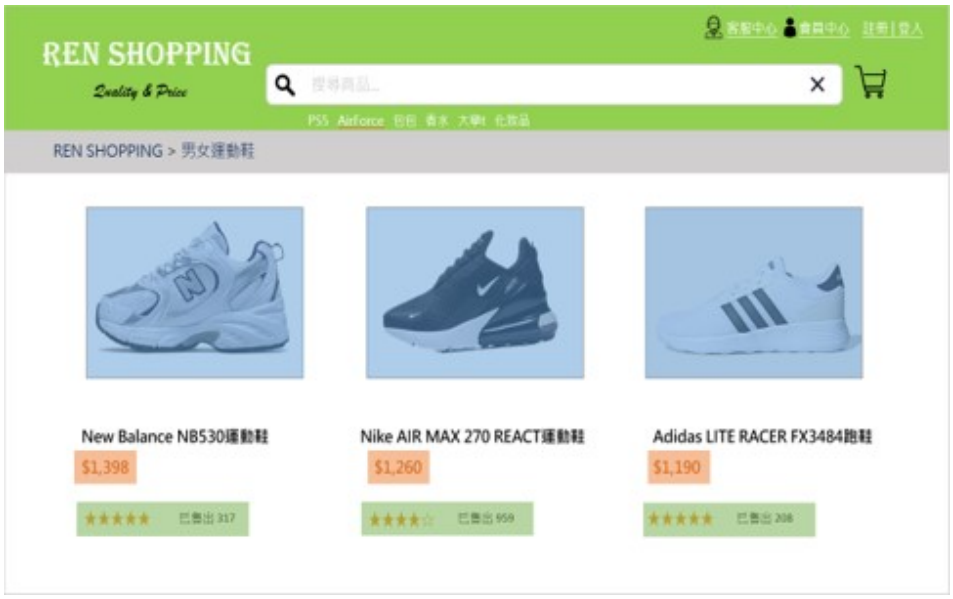


Fig. 12. Product Information AOI

In this study, earphones are categorized as utilitarian products, where consumers prioritize product quality and functionality when purchasing earphones. Factors that reflect product quality during online shopping include product price and product ratings and sales volume. Therefore, it is hypothesized that when consumers shop for earphones, they will primarily consider product price and product ratings and sales volume, with product appearance being of secondary importance. Consequently, the following attention allocation hypotheses for earphones are proposed:

1. *H5a: TFD for prices = TFD for ratings and sales volume > TFD for images*
2. *H5b: NF for prices = NF for ratings and sales volume > NF for images*

This study employs a one-way analysis of variance (One-Way ANOVA) with different product information AOIs as groups, as depicted in Figure 13. It aims to compare fixation duration and fixation count separately, with a significance level set at 0.05. If there are significant differences in attention allocation among the three product information categories, post-hoc comparisons will be conducted to analyze the hierarchy of attention allocation among them.

Fig. 13. Diagram of eye-tracking data calculation for commodity information

4. Experimental Results

The experiment recruited a total of 60 participants, comprising 30 males and 30 females, with an average age of 22.7 ± 2.68 years. All participants were college students without any eye-related disorders.

4.1. Eye-tracking Machine Learning Predictive Model Performance

This study utilizes eye-tracking data as input for a machine learning predictive system to forecast participants' purchase intentions. The experiment focuses on predicting purchases within three categories: shoes, clothing, and earphones. The models used encompass Decision Trees (DT), Support Vector Machines (SVM), Random Forest (RF), Extreme Gradient Boosting (XGB), and statistical-based models. Eye-tracking data, comprising total visit time, visit frequency, visit ratio, total fixation time, and fixation count, were employed to segregate products into three categories - highest, intermediate, and lowest eye-tracking data. These categories were utilized as machine learning features, culminating in a total of 15 features. Each product test group contained 60 data samples, split into training and testing sets at an $8:2$ ratio. Imbalanced minority class data were counterbalanced using the SMOTE technique during training. Table 1 illustrates the performance of the models within the shoe test set. The SVM model displayed superior performance with an accuracy of 0.80256, trailed by the RF model with an accuracy

of 0.78974. The statistical-based voting model demonstrated the lowest accuracy, at just 0.64358.

Table 1. Prediction performance table for the test set of shoes

Dataset	Method	Accuracy	Precision	Recall	F1-Score
Shoes	DT	0.69487	0.73333	0.69333	0.68667
	SVM	0.80256	0.82333	0.80333	0.79667
	RF	0.78974	0.81333	0.80667	0.78667
	XGB	0.77692	0.81	0.77667	0.77333
	Statistical	0.64358	0.76333	0.65667	0.63667

In the clothing test set, the performance of the models was not as prominent as in the shoe test set, as shown in Table 2. The RF model achieved the highest predictive performance with an accuracy of 0.71538, followed by the XGB model with an accuracy of 0.70513. The DT model showed the lowest accuracy, with only 0.63333.

Table 2. Prediction performance table for the test set of clothes

Dataset	Method	Accuracy	Precision	Recall	F1-Score
Clothes	DT	0.63333	0.62667	0.63	0.60333
	SVM	0.66410	0.68667	0.67	0.64667
	RF	0.71538	0.72667	0.70333	0.69
	XGB	0.70513	0.72333	0.69667	0.68
	Statistical	0.68205	0.67333	0.68333	0.64333

In the earphone test set, the performance of the models falls between that of the shoe test set and the clothing test set, as shown in Table 3. Among the models, the SVM model achieved the highest predictive performance with an accuracy of 0.74359, followed by the RF model with an accuracy of 0.73333. The Statistical model showed the lowest accuracy, with only 0.61538.

4.2. Impact of Image Complexity on Eye-tracking Data

This section examines the influence of images on attention through eye-tracking data. It compares groups with high-complexity images and groups with low-complexity images in terms of eye-tracking data, including glance time, fixation time, glance numbers, and fixation numbers, to determine whether significant differences exist. The experiment included two types of products, shoes and earphones, and presented the comparative results of image complexity between the shoe group and the earphone group. The experimental results are presented in the table 4-13 below:

Table 3. Prediction performance table for the test set of earphones

Dataset	Method	Accuracy	Precision	Recall	F1-Score
Earphones	DT	0.65128	0.6	0.57	0.56
	SVM	0.74359	0.79333	0.72667	0.72667
	RF	0.73333	0.73667	0.72333	0.72
	XGB	0.70256	0.73333	0.69333	0.69333
	Statistical	0.61538	0.61333	0.56	0.53667

Table 4. t-test of shoe selection time for different complexity groups

	Group	N	Mean	Std.	T-Value	P-Value
Product selection time	High complexity	60	19.60	9.85	2.13	0.037*
	Low complexity	60	16.82	9.45		

Table 5. t-test for total glance time between high-complexity shoe group and low-complexity shoe group

	Group	N	Mean	Std.	T-Value	P-Value
Total Glance Time	High complexity	60	7.449	6.145	0.52	0.605
	Low complexity	60	7.026	5.769		

Table 6. t-test of total fixation time between high-complexity shoe group and low-complexity shoe group

	Group	N	Mean	Std.	T-Value	P-Value
Total Fixation Time	High complexity	60	7.026	5.894	-0.85	0.400
	Low complexity	60	7.719	7.070		

Table 7. t-test for the number of glances between the high-complexity shoe group and the low-complexity shoe group

	Group	N	Mean	Std.	T-Value	P-Value
Number of Glances	High complexity	60	10.467	4.928	0.28	0.780
	Low complexity	60	10.233	6.596		

Table 8. t-test for the number of fixations between the high-complexity shoe group and the low-complexity shoe group

	Group	N	Mean	Std.	T-Value	P-Value
Number of Fixations	High complexity	60	14.07	7.97	-0.63	0.530
	Low complexity	60	14.92	10.38		

Table 9. t-test of the product selection time earphones between different complexity groups

	Group	N	Mean	Std.	T-Value	P-Value
Product Selection Time	High complexity	60	19.74	9.80	1.81	0.075
	Low complexity	60	17.93	8.44		

Table 10. t-test for total glance time between high-complexity earphones group and low-complexity earphones group

	Group	N	Mean	Std.	T-Value	P-Value
Total Glance Time	High complexity	60	4.994	3.413	-0.13	0.895
	Low complexity	60	5.047	3.561		

Table 11. t-test for total fixation time of high-complexity earphones group and low-complexity earphones group

	Group	N	Mean	Std.	T-Value	P-Value
Total Fixation Time	High complexity	60	4.579	3.190	-1.46	0.150
	Low complexity	60	5.170	4.383		

Table 12. t-test for the number of glances between the high-complexity earphones group and the low-complexity earphones group

	Group	N	Mean	Std.	T-Value	P-Value
Number of Glances	High complexity	60	7.867	4.102	-1.61	0.112
	Low complexity	60	8.800	4.977		

Table 13. t-test for the number of fixations between the high-complexity earphones group and the low-complexity earphones group

	Group	N	Mean	Std.	T-Value	P-Value
Number of Fixations	High complexity	60	11.40	6.43	-1.27	0.208
	Low complexity	60	12.70	8.67		

4.3. Product Information Attention Allocation

This study examines consumers' attention allocation during the process of purchasing products using eye-tracking data. Specifically, we compare the eye movement data related to three types of product information: product images, product prices, and product ratings and sales volume. The product categories include hedonic products such as shoes and clothing, as well as utilitarian products like earphones. The following experiment will present the attention allocation results for shoes, clothing, and earphones.The experimental results are presented in the table 14-24 below:

Table 14. ANOVA table of fixation time for three product information in the shoes test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Total Fixation Time	2	1013	506.58		
Error	177	2538	14.59	34.73	0.000*
Total	179	3551			

Table 15. Post-hoc comparative analysis of total fixation time of the shoe test group

info	N	Mean	Std.
Image	60	7.295	6.021
Price	60	2.312	1.783
Reviews and sales	60	2.132	2.080

Table 16. ANOVA table of the number of fixations of the three product information in the shoe test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Number of Fixations	2	1840	919.77		
Error	177	6050	34.77	26.45	0.000*
Total	179	7890			

4.4. Hypothesis Consolidation Table

The hypotheses and results of this study according to the experiment results are presented in Table 25. Subsequently, in Chapter 5, a further discussion and explanation will be provided regarding the experimental outcomes for each research item.

Table 17. Post-hoc comparative analysis of number of fixations of the shoe test group

info	N	Mean	Std.
Image	60	14.720	6.509
Price	60	8.924	5.242
Reviews and sales	60	7.178	5.871

Table 18. ANOVA table of total fixation time for three product information in the clothes test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Total Fixation Time	2	774.4	387.179		
Error	177	1691.5	9.557	40.51	0.000*
Total	179	2465.9			

Table 19. Post-hoc comparative analysis of total fixation time of the clothes test group

info	N	Mean	Std.
Image	60	6.404	4.791
Price	60	2.195	1.688
Reviews and sales	60	1.835	1.692

Table 20. ANOVA table of the number of fixations of the three product information in the clothes test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Number of Fixations	2	2634	1317.04		
Error	177	8674	49.01	26.87	0.000*
Total	179	11308			

Table 21. Post-hoc comparative analysis of the number of fixations in the clothes test group

info	N	Mean	Std.
Image	60	15.08	9.47
Price	60	8.367	5.810
Reviews and sales	60	6.067	4.857

Table 22. ANOVA table of total fixation time for three product information in the earphone test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Total Fixation Time	2	106.6	53.311		
Error	177	1370.7	7.744	6.88	0.001*
Total	179	1477.3			

Table 23. Post-hoc comparative analysis of total fixation time in the earphone test group

info	N	Mean	Std.
Image	60	4.506	3.263
Price	60	3.067	2.258
Reviews and sales	60	2.732	2.735

Table 24. ANOVA table of the number of fixations on the three product information in the earphones test group

	DF	Adj SS	Adj MS	F-Value	P-Value
Number of Fixations	2	237.9	118.95		
Error	177	9687.1	54.73	2.17	0.117
Total	179	9925.0			

Table 25. Hypothesis Consolidation Table

Hypothesis	Valid
H1a: The higher the complexity of the shoes image, the longer the consumer's glance time	No
H1b: The higher the complexity of the shoes image, the longer the consumer's fixation time	No
H1c: The higher the complexity of the shoes image, the higher the number of glances by consumers	No
H1d: The higher the complexity of the shoes image, the higher the number of fixations by consumers	No
H2a: The higher the complexity of the earphones image, the longer the consumer's glance time	No
H2b: The higher the complexity of the earphones image, the longer the consumer's fixation time	No
H2c: The higher the complexity of the earphones image, the higher the number of glances by consumers	No
H2d: The higher the complexity of the earphones image, the higher the number of fixations by consumers	No
H3a: TFD of images TFD of prices = TFD of reviews and sales	Yes
H3b: NF of images NF of prices = NF of reviews and sales	No
H4a: TFD of images TFD of prices = TFD of reviews and sales	Yes
H4b: NF of images NF of prices = NF of reviews and sales	Yes
H5a: TFD of prices = TFD of reviews and sales TFD of images	No
H5b: NF of prices = NF of reviews and sales NF of images	No

5. Discussion

This study utilizes eye tracking metrics, such as visit and gaze duration and frequency, to enhance understanding of consumer attention in e-commerce engagements. It explores the use of machine learning techniques to predict purchasing decisions based on categorized participant eye tracking data across three product categories - shoes, clothing, and earphones. Findings suggest a promising 70% prediction accuracy, demonstrating the potential of eye tracking data in estimating consumer interest. The Random Forest (RF) and Extreme Gradient Boosting (XGB) models have been particularly successful, outperforming traditional statistical models in terms of majority voting. This indicates the benefits of these models for predicting consumer preferences using eye tracking data, especially under limited training data conditions [57]. Among them, RF shows superior performance, making it an ideal model for eye tracking recommendation systems. The experiment results suggests that eye tracking data can effectively predict consumer interests, providing a valuable tool for e-commerce platforms. The RF model, capable of integrating various features for prediction, could be combined with additional data types, such as demographics or purchase history, to enhance personalization of product recommendations. Contrary to prior literature [31], we found no significant variance in eye tracking data for different product images, irrespective of their complexity. These results could be attributed to the experimental stimuli, as high complexity images were employed. However, these findings underline the need for further research into the role of image complexity in consumer gaze behavior [58–60, 31, 32]. This research categorizes products as hedonic and utilitarian and assesses differences in consumer focus across product types. It found that product images tend to command greater attention than other elements, such as price or rating, across both product types. This emphasis on images underscores their importance in e-commerce platforms and suggests that improvements in image quality could enhance consumer engagement [61, 62]. The gaze frequency data indicates variations in consumer focus depending on the product type. For instance, consumers prioritized product appearance for shoes and clothing, while price, ratings, and sales volume were equally important for high-priced products like earphones. These findings suggest tailored promotional activities could enhance consumer engagement with different product types. Despite the insights provided, this study acknowledges certain limitations, particularly the lack of diversity among the participant pool and the experimental setting, which excluded valuable contextual information, such as browsing history. Additionally, the impact of individual differences in decision-making styles on the effectiveness of eye tracking data requires further exploration.

In conclusion, this study highlights the potential of eye tracking data in e-commerce recommendation systems. However, further research is required to overcome the existing limitations and optimize the integration of eye tracking data with other forms of data for more precise and practical recommendations.

6. Conclusions

Our proposed approach integrates eye-tracking data and machine learning algorithms to predict consumer purchasing behavior on e-commerce platforms. Notably, the Random Forest (RF) model demonstrated exceptional performance, achieving a precision rate exceeding 70%, thereby outperforming other methods when utilizing eye-tracking metrics

for forecasting. Additionally, this study unveils distinct consumer preferences for hedonic and utilitarian products, providing valuable insights to guide differentiated marketing strategies aimed at enhancing consumer engagement. Product images emerge as pivotal in shaping consumer understanding, underscoring the critical role of effective design on e-commerce platforms. The integration of eye-tracking data for predicting individual product preferences holds the potential to significantly enhance e-commerce personalization, albeit necessitating adaptability due to varying levels of product page complexity. Moreover, the observed variability in browsing patterns and decision-making times across different personality traits suggests the prospect of refining predictive models through the inclusion of personality traits as predictive factors. While it is acknowledged that current webcam-based eye tracking systems have certain limitations, ongoing advancements in technology are anticipated to enhance precision, thereby making their widespread adoption increasingly feasible. The judicious utilization of eye-tracking data empowers e-commerce platforms with profound customer insights, ultimately leading to heightened customer satisfaction and increased sales by enabling more accurate tailoring of the shopping experience.

References

1. I. MacKenzie, C. Meyer, and S. Noble, "How retailers can keep up with consumers," *McKinsey & Company*, vol. 18, no. 1, 2013.
2. R. Rathee and P. Rajain, "Sensory marketing-investigating the use of five senses," *International Journal of Research in Finance and Marketing*, vol. 7, no. 5, pp. 124–133, 2017.
3. L. N. van der Laan, I. T. Hooge, D. T. De Ridder, M. A. Viergever, and P. A. Smeets, "Do you like what you see? the role of first fixation and total fixation duration in consumer choice," *Food Quality and Preference*, vol. 39, pp. 46–55, 2015.
4. S. Jantathai, L. Danner, M. Joechl, and K. Dürrschmid, "Gazing behavior, choice and color of food: Does gazing behavior predict choice?," *Food Research International*, vol. 54, no. 2, pp. 1621–1626, 2013.
5. L. Sharma and A. Gera, "A survey of recommendation system: Research challenges," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 5, pp. 1989–1992, 2013.
6. A. L. Montgomery, S. Li, K. Srinivasan, and J. C. Liechty, "Modeling online browsing and path analysis using clickstream data," *Marketing science*, vol. 23, no. 4, pp. 579–595, 2004.
7. A. Papoutsaki, "Scalable webcam eye tracking by learning from user interactions," in *Proceedings of the 33rd Annual ACM Conference Extended Abstracts on Human Factors in Computing Systems*, pp. 219–222, 2015.
8. I. Portugal, P. Alencar, and D. Cowan, "The use of machine learning algorithms in recommender systems: A systematic review," *Expert Systems with Applications*, vol. 97, pp. 205–227, 2018.
9. K. Tsuji, F. Yoshikane, S. Sato, and H. Itsumura, "Book recommendation using machine learning methods based on library loan records and bibliographic information," in *2014 IIAI 3rd International Conference on Advanced Applied Informatics*, pp. 76–79, IEEE, 2014.
10. S. Zahra, M. A. Ghazanfar, A. Khalid, M. A. Azam, U. Naeem, and A. Prugel-Bennett, "Novel centroid selection approaches for kmeans-clustering based recommender systems," *Information sciences*, vol. 320, pp. 156–189, 2015.
11. M. Nilashi, K. Bagherifard, M. Rahmani, and V. Rafe, "A recommender system for tourism industry using cluster ensemble and prediction machine learning techniques," *Computers & industrial engineering*, vol. 109, pp. 357–368, 2017.

12. M. Krol and M. Krol, "A novel approach to studying strategic decisions with eye-tracking and machine learning," *Judgment and Decision Making*, vol. 12, no. 6, pp. 596–609, 2017.
13. Y. Lou, Y. Liu, J. K. Kaakinen, and X. Li, "Using support vector machines to identify literacy skills: Evidence from eye movements," *Behavior research methods*, vol. 49, pp. 887–895, 2017.
14. S. Hoppe, T. Loetscher, S. A. Morey, and A. Bulling, "Eye movements during everyday behavior predict personality traits," *Frontiers in human neuroscience*, p. 105, 2018.
15. Z. Zhao, H. Tang, X. Zhang, X. Qu, X. Hu, and J. Lu, "Classification of children with autism and typical development using eye-tracking data from face-to-face conversations: Machine learning model development and performance evaluation," *Journal of Medical Internet Research*, vol. 23, no. 8, p. e29328, 2021.
16. J. Pfeiffer, T. Pfeiffer, M. Meißner, and E. Weiß, "Eye-tracking-based classification of information search behavior using machine learning: evidence from experiments in physical shops and virtual reality shopping environments," *Information Systems Research*, vol. 31, no. 3, pp. 675–691, 2020.
17. M. Shepherd, J. M. Findlay, and R. J. Hockey, "The relationship between eye movements and spatial attention," *The Quarterly Journal of Experimental Psychology Section A*, vol. 38, no. 3, pp. 475–491, 1986.
18. H. Deubel and W. X. Schneider, "Saccade target selection and object recognition: Evidence for a common attentional mechanism," *Vision research*, vol. 36, no. 12, pp. 1827–1837, 1996.
19. J. L. Orquin and S. M. Loose, "Attention and choice: A review on eye movements in decision making," *Acta psychologica*, vol. 144, no. 1, pp. 190–206, 2013.
20. L. Katus, N. J. Hayes, S. McCann, L. Mason, A. Blasi, M. K. Darboe, M. de Haan, S. E. Moore, S. Lloyd-Fox, and C. E. Elwell, "Implementing neuroimaging and eye tracking methods to assess neurocognitive development of young infants in low-and middle-income countries," *Gates Open Research*, vol. 3, 2019.
21. S. P. Devlin, N. L. Brown, S. Drollinger, C. Sibley, J. Alami, and S. L. Riggs, "Scan-based eye tracking measures are predictive of workload transition performance," *Applied ergonomics*, vol. 105, p. 103829, 2022.
22. A. Moran, M. Campbell, and D. Ranieri, "Implications of eye tracking technology for applied sport psychology," *Journal of Sport Psychology in Action*, vol. 9, no. 4, pp. 249–259, 2018.
23. M. Kuhar and T. Merčun, "Exploring user experience in digital libraries through questionnaire and eye-tracking data," *Library & Information Science Research*, vol. 44, no. 3, p. 101175, 2022.
24. J. N. Stember, H. Celik, E. Krupinski, P. D. Chang, S. Mutasa, B. J. Wood, A. Lignelli, G. Moonis, L. Schwartz, S. Jambawalikar, *et al.*, "Eye tracking for deep learning segmentation using convolutional neural networks," *Journal of digital imaging*, vol. 32, pp. 597–604, 2019.
25. N. Nugrahaningsih, M. Porta, and A. Klašnja-Milićević, "Assessing learning styles through eye tracking for e-learning applications," *Computer Science and Information Systems*, vol. 18, no. 4, pp. 1287–1309, 2021.
26. P. Majaranta and A. Bulling, "Eye tracking and eye-based human–computer interaction," in *Advances in physiological computing*, pp. 39–65, Springer, 2014.
27. B. K. Behe, M. Bae, P. T. Huddleston, and L. Sage, "The effect of involvement on visual attention and product choice," *Journal of Retailing and Consumer Services*, vol. 24, pp. 10–21, 2015.
28. P. Chandon, J. Hutchinson, E. Bradlow, and S. H. Young, "Measuring the value of point-of-purchase marketing with commercial eye-tracking data," *INSEAD Business School Research Paper*, no. 2007/22, 2006.
29. J. N. Sari, L. E. Nugroho, P. I. Santosa, and R. Ferdiana, "The measurement of consumer interest and prediction of product selection in e-commerce using eye tracking method," *Int. J. Intell. Eng. Syst*, vol. 11, no. 1, 2018.

30. Y. M. Hwang and K. C. Lee, "Using an eye-tracking approach to explore gender differences in visual attention and shopping attitudes in an online shopping environment," *International Journal of Human–Computer Interaction*, vol. 34, no. 1, pp. 15–24, 2018.
31. Q. Wang, D. Ma, H. Chen, X. Ye, and Q. Xu, "Effects of background complexity on consumer visual processing: An eye-tracking study," *Journal of Business Research*, vol. 111, pp. 270–280, 2020.
32. T. M. H. Vu, V. P. Tu, and K. Duerrschmid, "Design factors influence consumers' gazing behaviour and decision time in an eye-tracking test: A study on food images," *Food Quality and Preference*, vol. 47, pp. 130–138, 2016.
33. M. I. Jordan and T. M. Mitchell, "Machine learning: Trends, perspectives, and prospects," *Science*, vol. 349, no. 6245, pp. 255–260, 2015.
34. E. G. Learned-Miller, "Introduction to supervised learning," *I: Department of Computer Science, University of Massachusetts*, vol. 3, 2014.
35. Z. Liu, L.-M. Hu, and W.-C. Yeh, "Risk-averse two-stage stochastic programming-based closed-loop supply chain network design under uncertain demand," *Applied Soft Computing*, p. 110743, 2023.
36. Y.-Y. Song and L. Ying, "Decision tree methods: applications for classification and prediction," *Shanghai archives of psychiatry*, vol. 27, no. 2, p. 130, 2015.
37. S. Suthaharan and S. Suthaharan, "Support vector machine," *Machine learning models and algorithms for big data classification: thinking with examples for effective learning*, pp. 207–235, 2016.
38. W.-C. Yeh, "A two-stage discrete particle swarm optimization for the problem of multiple multi-level redundancy allocation in series systems," *Expert Systems with Applications*, vol. 36, no. 5, pp. 9192–9200, 2009.
39. L. Breiman, "Random forests," *Machine learning*, vol. 45, pp. 5–32, 2001.
40. L. Breiman, "Bagging predictors," *Machine learning*, vol. 24, pp. 123–140, 1996.
41. T. Chen and C. Guestrin, "Xgboost: A scalable tree boosting system," in *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794, 2016.
42. C. Schweikert, L. Gobin, S. Xie, S. Shimojo, and D. Frank Hsu, "Preference prediction based on eye movement using multi-layer combinatorial fusion," in *Brain Informatics: International Conference, BI 2018, Arlington, TX, USA, December 7–9, 2018, Proceedings 11*, pp. 282–293, Springer, 2018.
43. D. Das, L. Sahoo, and S. Datta, "A survey on recommendation system," *International Journal of Computer Applications*, vol. 160, no. 7, 2017.
44. P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, and J. Riedl, "Grouplens: An open architecture for collaborative filtering of netnews," in *Proceedings of the 1994 ACM conference on Computer supported cooperative work*, pp. 175–186, 1994.
45. J. B. Schafer, D. Frankowski, J. Herlocker, and S. Sen, "Collaborative filtering recommender systems," in *The adaptive web: methods and strategies of web personalization*, pp. 291–324, Springer, 2007.
46. P. B. Thorat, R. M. Goudar, and S. Barve, "Survey on collaborative filtering, content-based filtering and hybrid recommendation system," *International Journal of Computer Applications*, vol. 110, no. 4, pp. 31–36, 2015.
47. A. B. Barragáns-Martínez, E. Costa-Montenegro, J. C. Burguillo, M. Rey-López, F. A. Mikic-Fonte, and A. Peleteiro, "A hybrid content-based and item-based collaborative filtering approach to recommend tv programs enhanced with singular value decomposition," *Information Sciences*, vol. 180, no. 22, pp. 4290–4311, 2010.
48. R. Burke, "Hybrid recommender systems: Survey and experiments," *User modeling and user-adapted interaction*, vol. 12, pp. 331–370, 2002.
49. D. Filev and R. R. Yager, "On the issue of obtaining owa operator weights," *Fuzzy sets and systems*, vol. 94, no. 2, pp. 157–169, 1998.

50. J. Basiri, A. Shakery, B. Moshiri, and M. Z. Hayat, "Alleviating the cold-start problem of recommender systems using a new hybrid approach," in *2010 5th International Symposium on Telecommunications*, pp. 962–967, IEEE, 2010.
51. B. Walek and V. Fojtik, "A hybrid recommender system for recommending relevant movies using an expert system," *Expert Systems with Applications*, vol. 158, p. 113452, 2020.
52. H. Song and N. Moon, "Eye-tracking and social behavior preference-based recommendation system," *The Journal of Supercomputing*, vol. 75, pp. 1990–2006, 2019.
53. S. Jaiswal, S. Virmani, V. Sethi, K. De, and P. P. Roy, "An intelligent recommendation system using gaze and emotion detection," *Multimedia Tools and Applications*, vol. 78, pp. 14231–14250, 2019.
54. R. Dhar and K. Wertenbroch, "Consumer choice between hedonic and utilitarian goods," *Journal of marketing research*, vol. 37, no. 1, pp. 60–71, 2000.
55. A. Gere, L. Danner, N. de Antoni, S. Kovács, K. Dürrschmid, and L. Sipos, "Visual attention accompanying food decision process: An alternative approach to choose the best models," *Food Quality and Preference*, vol. 51, pp. 1–7, 2016.
56. T.-J. Hsieh, H.-F. Hsiao, and W.-C. Yeh, "Mining financial distress trend data using penalty guided support vector machines based on hybrid of particle swarm optimization and artificial bee colony algorithm," *Neurocomputing*, vol. 82, pp. 196–206, 2012.
57. O. Sagi and L. Rokach, "Ensemble learning: A survey," *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, vol. 8, no. 4, p. e1249, 2018.
58. K. Humphrey and G. Underwood, "The potency of people in pictures: Evidence from sequences of eye fixations," *Journal of Vision*, vol. 10, no. 10, pp. 19–19, 2010.
59. Q. Wang, Y. Yang, Q. Wang, and Q. Ma, "The effect of human image in b2c website design: an eye-tracking study," *Enterprise Information Systems*, vol. 8, no. 5, pp. 582–605, 2014.
60. A. Furnham and H. C. Boo, "A literature review of the anchoring effect," *The journal of socio-economics*, vol. 40, no. 1, pp. 35–42, 2011.
61. P. W. Miniard, S. Bhatla, K. R. Lord, P. R. Dickson, and H. R. Unnava, "Picture-based persuasion processes and the moderating role of involvement," *Journal of consumer research*, vol. 18, no. 1, pp. 92–107, 1991.
62. Y. Li and Y. Xie, "Is a picture worth a thousand words? an empirical study of image content and social media engagement," *Journal of Marketing Research*, vol. 57, no. 1, pp. 1–19, 2020.

Zhenyao Liu is currently a Ph.D. candidate in the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan. His research areas are soft computing and machine learning.

Wei-Chang Yeh received the M.S. and Ph.D. degrees from the Department of Industrial Engineering, University of Texas at Arlington. He is currently a Chair Professor of the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan. Most of his research is focused around algorithms, including exact solution methods and soft computing. He has published more than 250 research articles in highly ranked journals and conference papers.

Ke-Yun Lin received the M.S. degree from the Department of Industrial Engineering and Engineering Management, National Tsing Hua University, Taiwan.

Hota Chia-Sheng Lin is currently an assistant professor of the Department of Leisure and Recreation Management, Ming Chuan University, Taiwan.

Chuan-Yu Chang received the Ph.D. degree in electrical engineering from the National Cheng Kung University, Taiwan, in 2000. He is currently the Deputy General Director of the Service Systems Technology Center, Industrial Technology Research Institute, Taiwan. He is a Distinguished Professor with the Department of Computer Science and Information Engineering, National Yunlin University of Science and Technology, Taiwan. His current research interests include computational intelligence and its applications to medical image processing, automated optical inspection, emotion recognition, and pattern recognition.

Received: August 07, 2023; Accepted: October 06, 2023.

Computer Science and Information Systems 21(2):625–643 https://doi.org/10.2298/CSIS230818078L

A Novel Multipath QUIC Protocol with Minimized Flow Complete Time for Internet Content Distribution ⋆

Fang-Yi Lin[1], Wu-Min Sung[1], Lin Hui[2,⋆⋆], Chih-Lin Hu[1], Nien-Tzu Hsieh[1], and Yung-Hui Chen[3]

[1] Department of Communication Engineering, National Central University
Taoyuan City 320317, Taiwan
{fangyi.lin, wumin.sung, neintzu.hsieh}@g.ncu.edu.tw
clhu@ce.ncu.edu.tw
[2] Department of Computer Science and Information Engineering, Tamkang University
New Taipei City 25137, Taiwan
amar0627@gms.tku.edu.tw
[3] Department of Computer Information and Network Engineering, Lunghwa University of Science and Technology, Taoyuan City 333326, Taiwan
cyh@mail.lhu.edu.tw

Abstract. The rapid growth of network services and applications has led to an exponential increase in data flows on the internet. Given the dynamic nature of data traffic in the realm of internet content distribution, traditional TCP/IP network systems often struggle to guarantee reliable network resource utilization and management. The recent advancement of the Quick UDP Internet Connect (QUIC) protocol equips media transfer applications with essential features, including structured flow-controlled streams, quick connection establishment, and seamless network path migration. These features are vital for ensuring the efficiency and reliability of network performance and resource utilization, especially when network hosts transmit data flows over end-to-end paths between two endpoints. QUIC greatly improves media transfer performance by reducing both connection setup time and transmission latency. However, it is still constrained by the limitations of single-path bandwidth capacity and its variability. To address this inherent limitation, recent research has delved into the concept of multipath QUIC, which utilizes multiple network paths to transmit data flows concurrently. The benefits of multipath QUIC are twofold: it boosts the overall bandwidth capacity and mitigates flow congestion issues that might plague individual paths. However, many previous studies have depended on basic scheduling policies, like round-robin or shortest-time-first, to distribute data transmission across multiple paths. These policies often overlook the subtle characteristics of network paths, leading to increased link congestion and transmission costs. In this paper, we introduce a novel multipath QUIC strategy aimed at minimizing flow completion time while taking into account both path delay and packet loss rate. Experimental results demonstrate the superiority of our proposed method compared to standard QUIC, Lowest-RTT-First (LRF) QUIC, and Pluginized QUIC schemes. The relative performance underscores the efficacy of our design in achieving efficient and reliable data transfer in real-world scenarios using the Mininet simulator.

⋆ This paper is an extended version of a conference paper, which was published in The 13th International Conference on Frontier Computing (FC 2023), July 10 -14, 2023, after a thorough enhancement.

⋆⋆ Corresponding author

Keywords: Quick UDP internet connect (QUIC), multipath transport, HTTP, content distribution, internet protocol, internet services.

1. Introduction

The HTTP protocol family [1] is the basis for global internet data communications, enabling the rapid development of Web browsers and internet services. HTTP/1.1 and HTTP/2 are two major web protocols. With the proliferation of user demands and mobile services, particularly mobile media streaming and AR/VR flows to an increasing user population, the functions provided by HTTP/1.1 and 2 are no longer sufficient. In 2013, the IETF organization proposed the RFC 9000 [2], i.e., Quick UDP Internet Connect (QUIC) – a UDP-based multiplexed and secure transport protocol. QUIC is often known as the transport layer for HTTP/3. It is recommended to develop HTTP/3 with QUIC and UDP in place of conventional HTTP/1.1 and 2 with TCP or UDP for internet services and applications in wireless and mobile environments.

QUIC provides internet applications with flow-controlled streams for encrypted, multiplexed and reliable communication, low-latency connection establishment, and network path migration. It can sustain high dynamics of traffic loading and resource provision on network hosts, rather than HTTP/2 based on TCP, TLS 1.2, and other HTTP derivatives. Compared with TCP, QUIC need not the 3-way handshake mechanism, so it can greatly reduce the time of network connection establishment and transmission latency. With multiplexing and path migration, it can strengthen the control of congested networks, making it more suitable for emerging mobile services in Wi-Fi and 4G/5G environments.

Prior studies argued that the performance of QUIC can be affected in the case of delivering large-size data between two endpoints [3]. This is because the packet pacing policy is basically used to vary the transmission speed of each stream when numerous packets enter into that stream. The overall completion time of a data flow in a stream, so-called *flow complete time* briefly, will vary as well. Thus, data throughput of each flow going on a link may not reach to the full bandwidth capacity. Moreover, internet operators may operate any self-protection controls by limiting the transmission rate of UDP flows. Regarding the safety of a network system, a self-imposed constraint can be understood to defend against unpredictable threats to the system, although the bandwidth resource along with those links between two endpoints is not used fully.

As the literature review will be mentioned in Section 2, recent studies used Multipath QUIC (MPQUIC) to deal with the above concerns arising from the restrictions of a single path. Similar to Multipath TCP (MPTCP) [4], MPQUIC sends data through different paths and uses the aggregate bandwidth of different paths. It also likely modifies the *path scheduler* policy for increasing the transmission speed and thence decreasing the path delay that definitely corresponds to the end-to-end transmission delay of a QUIC stream between two endpoints in a network. In light of the aforementioned concept of MPQUIC, our study in this paper leverages the functionality of MPQUIC to devise a novel MPQUIC-based path selection strategy for internet content delivery. The contributions of our study are outlined as follows:

- We introduce a novel Multipath QUIC scheme. Its functionality is distinguished by considering both path latency and packet loss rate to identify the most efficient paths.

As a result, network transmission performance is enhanced by minimizing the flow completion time.
- We present the mathematical formulation of the proposed MPQUIC-based path selection strategy and detail the algorithmic procedures. Subsequently, we create a proof-of-concept implementation on the Mininet simulation platform.
- We evaluate the relative performance of our proposed strategy against several standard schemes, including standard QUIC, Lowest-RTT-First (LRF) QUIC [5], and Pluginized QUIC (PQUIC) [6] scheduling strategies, using the Abilene topology on the Mininet emulator. Performance results underscore the superiority of our strategy. Notably, our scheme consistently achieves stable and efficient outcomes in terms of the cumulative distribution function (CDF) of flow completion time. Furthermore, our strategy results in lower path delay and packet loss rates compared to other schemes.
- Performance results demonstrate the superiority of our scheme because this scheme can achieve stable and efficient effects in measure of the cumulative distribution function (CDF) of flow completion time. In addition, this scheme obtains lower path delay and packet loss rate than the other schemes.

The rest of this paper is organized as follows. Section 2 describes background knowledge ad related work. Section 3 details the problem formulation and the path selection algorithm. Section 4 describes the relative performance. Finally, the conclusion is given in Section 5.

2. Background Knowledge & Related Work

Section 2.1 briefly introduces the QUIC protocol to ease the understanding of special functional extensions by contrast to the conventions of TCP and HTTP protocols. Section 2.2 mentions the recent studies on the QUIC-based media transfer techniques in the literature.

2.1. Preliminary Knowledge

With the increasing demand for real-time applications, HTTP/2 shortcomings came to the fore; HTTP/3 aims to provide fast, reliable, and secure web connections. Figure 1 illustrates the architecture of the HTTP/3 protocol. HTTP/3 uses a new transport layer network protocol called QUIC, which runs over the UDP internet protocol instead of the ordered message exchange by TCP. The goal of HTTP/3 is to improve the overall web experience, suitable for IoT, real-time applications, and micro services. In addition, UDP provides greater flexibility, so that it can enables QUIC to exist completely in the user space. When QUIC can be independent of the operating system on the host, users only need to update a Web browser version with QUIC supported to experience the improved network performance by HTTP/3 [7]. Explicitly, QUIC serves as a new message transport layer, featuring Zero Round-Trip Time (0-RTT), flow control, congestion control, multiplexing, built-in security through TLS 1.3, and multiple paths. To aid comprehension, the following describes some of the essential features of QUIC:

Firstly, QUIC's primary attribute is the reduction in connection establishment latency. Unlike traditional TCP connections, QUIC eliminates the need for a three-way handshake, allowing for swift connection establishment. Thus, QUIC can lower initial latency

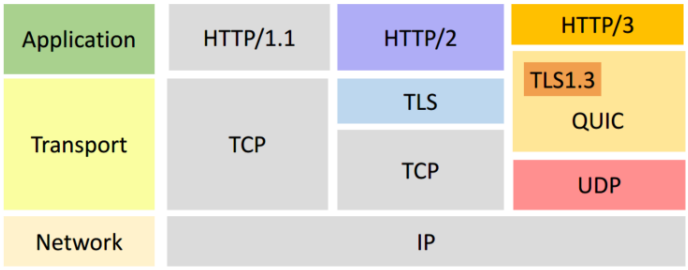


Fig. 1. HTTP architecture

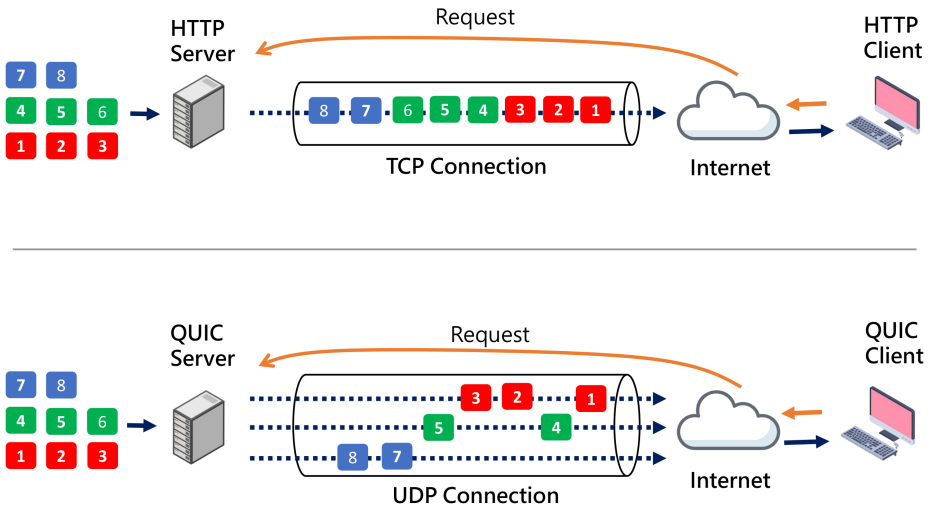


Fig. 2. Difference of the multiplexing features between TCP and QUIC

and quickly respond to end users through 0-RTT by sending data in the very first packet of a connection. Secondly, multiplexing allows for the concurrent transmission of multiple data streams over a single connection, as shown in Figure 2. This improves the efficiency of data transfer and overall performance while addressing the head-of-line blocking problem commonly encountered in HTTP/1.1. Additionally, QUIC possesses built-in error correction mechanisms that swiftly handle corrupted or lost data packets, enhancing the reliability of data transfer in the network. Thirdly, while congestion control in TCP commonly uses the CUBIC algorithm [8], it is not the most optimal for transmitting latency-sensitive network traffic. QUIC offers both the CUBIC and the Bottleneck Bandwidth and Round-trip propagation time (BBR) [9] schemes to address congestion-related issues. BBR actively probes and groups recently sent data, establishing a network model based on the current maximum bandwidth and round-trip time, allowing for the adjustment of transmission rates based on dynamic network conditions, effectively preventing flow con-

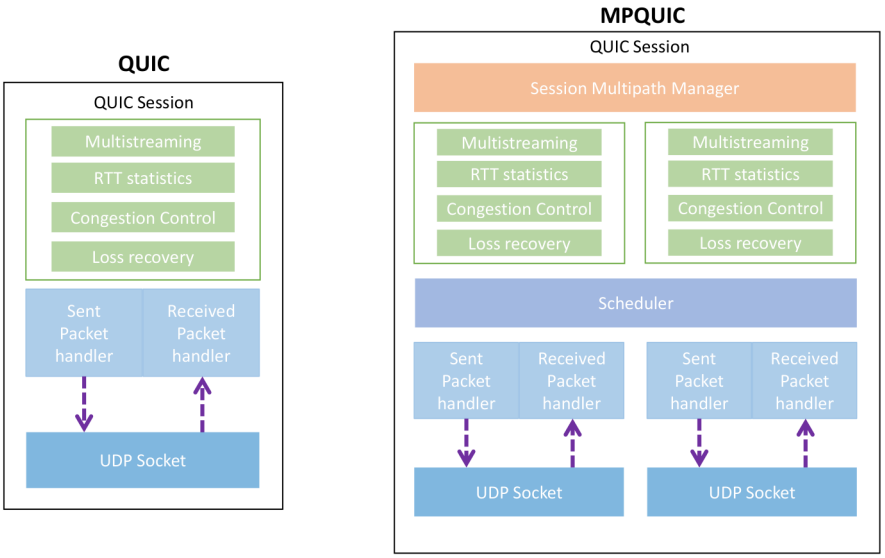


Fig. 3. Comparison between QUIC and MPQUIC network protocols

gestion and optimizing network performance. Fourthly, QUIC integrates Transport Layer Security (TLS) version 1.3 by default, ensuring that data is encrypted and secure during transmission. QUIC's adaptability is notable, allowing for dynamic path and protocol version selection in response to real-time changes in network conditions. Finally, QUIC enables the simultaneous utilization of multiple paths in a network, bolstering network robustness and performance by sending data through various routes, reducing latency, and maximizing bandwidth utilization, as illustrated in the comparison between QUIC and MPQUIC network architectures in Figure 3.

In summary, QUIC offers a comprehensive suite of features that collectively improve internet communication by enhancing speed, reliability, and security, making it well-suited for a wide range of network applications and effectively addressing the demands of modern internet usage, including real-time communication, mobile networks, and high-performance scenarios.

2.2. Related Work

Lots of studies on MPQUIC was inspired by MPTCP. As addressed in [10], the concept of MPQUIC can arrange QUIC connections to go on different paths according to network characteristics. There are two main reasons for the use of the multi-path function. The first is to collect the network resources of different paths to transmit data. Automatically selecting the best path becomes an interesting idea. The second is to maintain user experience against network failures. Given a device with multiple ports, if one of the network interfaces/ports/paths run to failed, the way of immediately switching to another one will not affect the user experience. Using multi-path designs can thus ensure the reliability

and stability of network transport services because such designs can distribute and schedule streams to reduce the overall completion time with respect to media transfer in the Internet.

Our literature review summarizes recent studies that proposed various MPQUIC scheduling methods based on a variety of design aspects, such as transmission completion time, path characteristics, data priority, congestion control, and machine learning to enhance the performance of multipath transmission. In course of MPQUIC scheduling, the path selection is crucial for determining the network throughput, reliability, and load balancing with respect to different service requirements. In what follows, we classify prior studies into five categories corresponding to different design aspects.

1. Transmission completion time: [11] investigated MPQUIC's performance on different paths using the proposed Estimated Transfer Completion time (ETC) scheduling method. It considers transmission time and path congestion, reducing the overall file transmission time. However, as a large scheduling unit is used, this method has a drawback of being inefficient for transferring short files. [12] proposed the Delay-based In-Order Decode (DIOD) method which combines Forward Error Correction (FEC) and MPQUIC for reliable and delay-sensitive applications. While it reduces the influence of packet loss, it does not guarantee deadlines and will necessitate a precise loss estimation method for scheduling flexibility.
2. Path characteristics: [13] showed a PStream scheduling method that efficiently matches stream-path characteristics and avoids stream competition for the fastest path. [14] proposed a Nine Tails scheduler that can selectively use redundant paths to reduce latency as sending data in the tail part. By switching between redundant and non-redundant scheduling policies, it can have higher overall throughput and loss recovery. [15] designed an optimal bandwidth allocation strategy which can prioritize streams with a combination of priority and size. However, it underutilizes multi-path aggregation, resulting in suboptimal bandwidth allocation for time-critical stream. When network topology and bandwidth changes, this strategy may fall in performance degradation and flow congestion.
3. Data priority: [16] emphasized prioritized fair bandwidth allocation based on stream priority to prevent delays of individual streams due to varying network paths. [17] developed a Priority Bucket method that categorizes streams into priority-based buckets. Streams with the same priority in a bucket are served in a first-come-first-served order using HTTP/2 expressions. [18] designed Xlink, which is a user-perceived video Quality of Experience (QoE) control mechanism for MPQUIC scheduling. It showed the feasibility and deployability of MPQUIC in the Taobao platform, while substantial bandwidth is potentially required. [19] assumed the server-side has prior knowledge of the web page's dependency tree. It used a flow-aware downlink packet scheduling with stream priorities to optimize the transmission order of streams. This way can reduce flow completion time, page loading time, and expedite loss recovery, but may have efficiency implications for low-priority flows.
4. Congestion control: [20] developed MM-QUIC within the ITSN architecture, utilizing regular satellite orbits for rapid transmission. It also incorporated a basic multi-path model for congestion management but noted potential packet loss during handovers in weak signal areas. [21] extended MPQUIC to SR-MPQUIC for 5G networks, improving latency and reliability for prioritizing traffic with redundant path

replication. With the primary focus on delay-sensitive traffic, it may slightly increase bandwidth usage and latency. [22] focused on congestion control and packet scheduling in multipath scenarios. It proposed a Delay-BBR algorithm that complements rate control to reduce packet loss and transmission delay for real-time video.

5. Machine learning: In [23], a reinforcement learning-based scheduling method, Peekaboo, was proposed. It considered temporal certainty and randomness in path characteristics for decision-making. [24] proposed MPQUIC schedulers using the deep reinforcement learning, this design which emphasized fairness to concurrent TCP flows in multipath protocols. [25] introduced a reinforcement learning-based MPQUIC scheduler using Deep Q-Network (DQN) to improve multimedia streaming performance and reduce video download time.

Our study considers the flow completion time in related with two network-oriented factors, i.e., delay and packet loss rate of a path. Accordingly, we formulate a weighting normalization method to calculate the weights of paths, which can be used to facilitate path selection and thus minimize the flow completion time over MPQUIC streams.

3. Design of Path Selection Scheme

This section first describes the problem formulation and then specifies a novel MPQUIC-based path selection scheme for efficient content delivery in the internet.

3.1. Problem Formulation

Give a network topology $G(V, L)$. For every link $l_{i,j} \in L$ from v_i to v_j, the available bandwidth, the delay of the link, and the packet loss rate w.r.t $l_{i,j}$ are denoted as $b_{i,j}$, $t_{i,j}$ and $o_{i,j}$, respectively. Then, $b_{i,j}^{max}$ denotes the maximum amount of bandwidth that $l_{i,j}$ can use.

Let F contain a set of all streams in $G(V, L)$, $\mathcal{P}_f^*$ represent a multipath set in use for a stream $f \in F$, $P_f^*[m]$ be the set of links in the m^{th} path, and likewise $\mathcal{P}_f^*[m][n]$ be the n^{th} link of the m^{th} path. Thus, for the stream and path selection, we take $x_{l_{i,j}}^f$ to be a binary indicator, defined as follows:

$$x_{i,j}^f = \begin{cases} 1, & \text{if a stream } f \text{ passes through a link } l_{i,j}\ , \\ 0, & \text{other conditions.} \end{cases} \tag{1}$$

We further define several expressions regarding the relationship between links and paths, as follows:

$$b_f^P = min\Big(b_{i,j} \times x_{i,j}^f\Big), \quad \forall l \in l_{i,j},\ x_{i,j}^f \neq 0,\ f \in F \tag{2}$$

$$b_{i,j}^{max} \geq \sum_{f \in F} b_{i,j} \times x_{i,j}^f, \quad \forall l \in l_{i,j} \tag{3}$$

$$t_f^P = \sum_{l_{i,j} \in L} t_{i,j} \times x_{i,j}^f, \quad \forall\ f \in F,\ l \in l_{i,j} \tag{4}$$

Table 1. Notations used in MPQUIC domain

Symbol	Description
$G(V, L)$	a graphic representation of a MPQUIC system
V	a set of all nodes in $G(V, L)$
L	a set of all links between two adjacent nodes in $G(V, L)$
$b_{i,j}$	available bandwidth of a link $l_{i,j}$
$b_{i,j}^{max}$	the maximum bandwidth of a link $l_{i,j}$
$t_{i,j}$	transmission delay of a link $l_{i,j}$
$o_{i,j}$	packet loss rate of a link $l_{i,j}$
F	a set of all data flows in the network
b_f	amount of bandwidth required for data stream f
t_f	transmission delay tolerance of a data stream f
o_f	packet loss rate tolerance for a stream f
$\mathcal{P}$	a set of all routing paths between any two nodes in $G(V, L)$
$\mathcal{P}_f$	a set of available paths for a data stream f
$\mathcal{P}_f[m]$	a set of links for the m^{th} path available to the data stream f
$\mathcal{P}_f[m][n]$	n^{th} link of the m^{th} path available to the data stream f
$P[n]$	n^{th} link of path P
$\mathcal{P}_f^*$	a set of multipath that the system ultimately uses for the data stream f
$\mathcal{P}_f^*[m]$	a link set of the m^{th} path in a set of multipath used by the system for stream f
$\mathcal{P}_f^*[m][n]$	the n^{th} link in the m^{th} path in the set of multipath used by the system for the data stream f

$$o_f^P = 1 - \prod_{l_{i,j} \in L} \left(1 - o_{i,j} \times x_{i,j}^f\right), \quad \forall l \in l_{i,j},\ x_{i,j}^f \neq 0,\ f \in F \tag{5}$$

$$y(\mathcal{P}_f^*) = \begin{cases} 1, & \bigcup P_f^* \neq \emptyset, \\ 0, & \bigcup P_f^* = \emptyset. \end{cases} \tag{6}$$

Formula (2) indicates the available bandwidth of a stream f in the set of paths P, and then takes the minimum value. (3) indicates that the bandwidth passed by a link cannot be greater than the maximum bandwidth available of the link. (4) means the sum of transmission delays on a link w.r.t. a stream f. (5) is to multiply the successful rate of each link to get the overall successful rate on a path, so as to obtain the packet loss rate of this path.

To transform a single-path stream into a multipath stream by (6), $y(\mathcal{P}_f^*)$ indicates whether any link and path in the set of paths $\mathcal{P}_f^*$ can be reused or not. Here, we further discuss two cases, as follows.

Case 1 When the links and paths in $\mathcal{P}_f^*$ are not reused.

Since links are not reused, the sum of the available bandwidth of each path can be calculated by (7). Then, for $y(\mathcal{P}_f^*) = 0$ and $\forall v_j \in V$, we can formulate (8) to check the link condition of v_i and v_j: (i) the total number of positive multipaths, (ii) the total number of negative multipaths, and (iii) a balanced state if both v_i and v_j are intermediate relays.

$$b_f^* = \sum_{P \in \mathcal{P}_f^*} b_f^P, \quad \forall f \in F,\ y(\mathcal{P}_f^*) = 0 \tag{7}$$

$$\sum_{l_{i,j} \in L} x_{i,j}^f - \sum_{l_{j,i} \in L} x_{j,i}^f = \begin{cases} |\mathcal{P}_f^*|, & \text{if } v_i \text{ is a start point of } f\ , \\ -|\mathcal{P}_f^*|, & \text{if } v_i \text{ is a target point of } f, \\ 0, & \text{if } v_i \text{ is a relay point of } f. \end{cases} \tag{8}$$

□

Case 2 When the links and paths in $\mathcal{P}_f^*$ can be reused

Let $z_{l_{i,j}}^{\mathcal{P}_f^*}$ indicate whether $l_{i,j}$ is reused in $\mathcal{P}_f^*$:

$$z_{l_{i,j}}^{P_f^*} = \begin{cases} 1, & l_{i,j} \subseteq \bigcup \mathcal{P}_f^*, \\ 0, & l_{i,j} \nsubseteq \bigcup \mathcal{P}_f^*. \end{cases} \tag{9}$$

$n(l_{i,j}, P_f^*)$ indicates the number of times that $l_{i,j}$ is reused by some paths in $\mathcal{P}_f^*$:

$$n(l_{i,j}, P_f^*) = \begin{cases} \sum\limits_{m \in |\mathcal{P}_f^*|} \sum\limits_{n \in |\mathcal{P}_f^*[m]|} l_{i,j} \wedge P_f^*[m][n] - 1, & \forall f \in F,\ l_{i,j} \in L,\ z_{l_{i,j}}^{P_i^*} = 1, \\ 0, & \forall f \in F,\ l_{i,j} \in L,\ z_{l_{i,j}}^{P_i^*} = 0. \end{cases} \tag{10}$$

Then, the bandwidth of a link is divided into two parts: the link bandwidth that has been reused $\bar{b_f^*}$, and the link that has not been reused $\hat{b_f^*}$, as follows.

$$b_f^* = \bar{b_f^*} + \hat{b_f^*}, \quad \forall f \in F, \quad \text{subject to} \tag{11a}$$

$$\bar{b_f^*} = min(b_f^P), \quad \forall f \in F,\ P \in \mathcal{P}_f^*,\ y(\mathcal{P}_f^*) = 1,\ z_{l_{i,j}}^{P_i^*} = 1. \tag{11b}$$

$$\hat{b_f^*} = \sum_{P \in \mathcal{P}_f^*} b_f^P, \quad \forall f \in F,\ y(\mathcal{P}_f^*) = 1,\ z_{l_{i,j}}^{P_i^*} = 0, \tag{11c}$$

Formula (11a) adds the two parts together, which yields the total amount of bandwidth that a path set can provide.

Formula (12) clarifies the link relation in three conditions. (i) If v_i is a start point of a stream f, the total of paths that a steam can still use is given by $|\mathcal{P}_f^*|$ minus the number of times $l_{i,j}$ that is currently used by some paths in $\mathcal{P}_f^*$, i.e., $n(l_{i,j}, P_f^*)$. (ii) If v_i is a target point, the calculation is in opposition to (i). (iii) Finally, if v_i is a relay w.r.t. $\forall\ y(\mathcal{P}_f^*) = 1$ and $v_j \in V$, there are three sub-cases (a)(b)(c). Explicitly, (a) multiple paths converge at this relay point, then $n(l_{j,i}, P_f^*) - n(l_{i,j}, P_f^*)$ is negative. (b) multiple paths to divert from this point, this outcome is positive. (c) in a balanced state, the outcome equals to 0.

$$\sum_{l_{i,j} \in L} x_{l_{i,j}}^f - \sum_{l_{j,i} \in L} x_{l_{j,i}}^f = \begin{cases} |\mathcal{P}_f^*| - n(l_{i,j}, P_f^*), & \text{if } v_i \text{ is a start point of } f\ , \\ -|\mathcal{P}_f^*| + n(l_{j,i}, P_f^*), & \text{if } v_i \text{ is a target point of } f, \\ n(l_{j,i}, P_f^*) - n(l_{i,j}, P_f^*), & \text{if } v_i \text{ is a relay point of } f. \end{cases} \tag{12}$$

□

Note that under the multipath scenario, the delay time and packet loss rate of a path are not affected by whether a path is reused subject to (2). Regardless of the value of (6),

the delay time and packet loss rate w.r.t. any $P \in \mathcal{P}_f^*$, denoted as t_f^* and o_f^*, can be given below.

$$t_f^* = \max(t_f^P)\,, \quad \forall f \in F,\ P \in \mathcal{P}_f^*,\ y(\mathcal{P}_f^*) = 0 \tag{13}$$

$$o_f^* = \sum_{P \in \mathcal{P}_f^n} \frac{o_f^P}{|\mathcal{P}_f^*|}, \quad \forall\ f \in F, y(\mathcal{P}_f^*) = 0 \tag{14}$$

According to (13), given a set of final selected multipaths, the delay time is represented by the maximum delay time on the path for $\forall P \in \mathcal{P}_f^*$. The outcome of (14) indicates the average of packet loss rate for those selected paths in $\mathcal{P}_f^*$. After calculating the available bandwidth, delay time, and packet loss rate, now, it is able to figure out the comparison between user requirements and actually available provision, as explained below.:

$$b_f \leq b_f^*, \quad \forall f \in F \tag{15}$$

$$t_f \geq t_f^*, \quad \forall f \in F \tag{16}$$

$$o_f \geq o_f^*, \quad \forall f \in F \tag{17}$$

Particularly, (15) ensures that the multipath bandwidth is available for streaming f, while (16) and (17) enforce that both transmission delay and packet loss rate in the selected path need to be smaller than the tolerable bounds as requested by f.

Hence, in accordance with the above formulae and constraints of the multipath provision, we develop an optimal multipath selection problem of minimizing the flow completion time subject to user requirements, as expressed below:

$$\begin{aligned}
&\arg\ \min \sum_{f \in F} t_f^*, \\
&\text{s.t.} \\
&\quad x_{i,j}^f = 1, && \forall l_{i,j} \in L, \\
&\quad z_{l_{i,j}}^{\mathcal{P}_f^*} \in (0,1), && \forall \mathcal{P}_f^*,\ l_{i,j} \in L, \\
&\quad y(\mathcal{P}_f^*) \in (0,1), && \forall \mathcal{P}_f^* \in \mathcal{P}, \\
&\quad \text{Eqs. } (15), (16), (17).
\end{aligned} \tag{18}$$

3.2. MPQUIC-Based Path Selection and Algorithmic Procedures

Our study refers to the research efforts in [26][27], and learns that such a multipath selection problem for QoS-based data streaming is known as NP-Complete [28]. Instead of finding a static optimization in theory, our study in this paper attempts to develop an optimal-approximate solution to figure out a set of appropriate multipaths using heuristic strategies with two design factors, i.e., path delay and packet loss rates. Particularly, we describe a weighting normalization method in 19 with two tuning parameters α and β to change the relative influence of path delay and packet loss rate over MPQUIC streams.

$$p_w = \alpha \times \frac{t_f}{t_f^*} + \beta \times \frac{o_f}{o_f^*}. \tag{19}$$

In what follows, we specify the algorithmic procedures for finding the paths for MPQUIC streams.

Algorithm 1 Path Set Selection with Joint Path Delay and Packet Loss Rate

input : $G(V, L)$: network topology,
k: the number of paths in the multipath,
α: a coefficient of path delay,
β: a coefficicient of packet loss.
output: $\mathcal{P}_f^*$: the set of multipath.

while *Flow f comes into the system* **do**
  $\mathcal{P}_f = \{\emptyset\}$;
  $A[\,][\,] =$ null;
  while $\mathcal{P}_f = \{\emptyset\}$ **do**
    $\mathcal{P}_f \leftarrow getDefaultPathSet(\mathcal{P}, f)$;
    foreach $p \in \mathcal{P}_f$ **do**
      $A[p][0] \leftarrow getPathBW(\mathcal{P}[p])$; ▷ (2)
      $A[p][1] \leftarrow getPathDelay(\mathcal{P}[p])$; ▷ (4)
      $A[p][2] \leftarrow getPathPL(\mathcal{P}[p])$; ▷ (5)
    end foreach
  end while
  if *($P_f = \{\emptyset\}$ or $|P_f| < k$)* **then**
    $Reject\ f$;
  else
    $\mathcal{P}_f^* \leftarrow getkPath(\mathcal{P}_f, \alpha, \beta, f, k, A)$; ▷ Go to Alg. 2
    if $\mathcal{P}_f^* = \emptyset$ **then**
      $\mathcal{P}_f^* \leftarrow getShorestkPath \in \mathcal{P}_f$;
    end if
  end if
end while

Algorithm 1 Path Set Formation with Joint Path Delay and Packet Loss Rate

When the stream enters the MPQUIC, the system initializes the set of available paths $\mathcal{P}_f$ for a data stream f, as well as prepares an empty two-dimensional matrix $A[\,][\,]$. At first, when $\mathcal{P}_f$ is empty, the system refers to (2), (4) and (5) to determine the values of data stream bandwidth, delay, and packet loss rate, which are stored in $A[\,][\,]$. Then, the system checks a condition of whether the set of available paths for f contains equal to or more than k paths. As this condition is valid, the system proceeds to Algorithm 2 with a set of candidate paths for f. Later soon, Algorithm 2 will figure out k shortest paths to form a set of $\mathcal{P}_f^*$.

Algorithm 2 Finding *k* Shortest Paths over MPQUIC Streams

Algorithm 2 is the path selection procedure for finding the k-shortest paths based on QoS requirements. This procedure refers to Yen's k-shortest path algorithm [29] with QoS-specific conditions. To find the k-shortest paths, the procedure runs several routes sequentially: (a) define variables p_w, b_f^* and $\mathcal{P}_f^*[\,][\,]$, (b) calculate the weight value p_w of a stream by (19), (c) sort the weights of streams in descending order, and (d) update the available bandwidth of each link according to (7) and (11a). Then, the procedural

Algorithm 2 Finding k Shortest Paths over MPQUIC Streams

Function $getkPath(\mathcal{P}_f, \alpha, \beta, f, k, A)$ **is**
  $p_w[\,] = \text{null};$
  $b_f^* = 0;$
  $\mathcal{P}_f^*[\,][\,] = \text{null};$
  foreach $p \in \mathcal{P}_f$ **do**
    $p_w[p] \leftarrow getPathWeight(\mathcal{P}[p], \alpha, \beta, A)$ ▷ (19)
  end foreach
  $p_w \leftarrow sortByDescendingOrder(p_w)$;
  $\mathcal{P}_f^* \leftarrow selectPathTopk(p_w, k)$;
  $b_f^* \leftarrow getMultiPathBW(\mathcal{P}_f^*)$ ▷ (7) and (11a)
  while $b_f^* \leq b_f$ **do**
    if $minBWPath(\mathcal{P}_f^*) \geq maxBWPath(\mathcal{P}_f - \mathcal{P}_f^*)$ **then**
      $\mathcal{P}_f^* = \emptyset$
      break;
    end if
    $\mathcal{P}_f^* \leftarrow \mathcal{P}_f^* - minBWPath(\mathcal{P}_f^*)$
    $\mathcal{P}_f^* \leftarrow \mathcal{P}_f^* + maxBWPath(\mathcal{P}_f - \mathcal{P}_f^*)$
    $b_f^* \leftarrow getMultiPathBW(\mathcal{P}_f^*)$ ▷ (7) and (11a)
  end while
  return $\mathcal{P}_f^*$
end

routine goes into a while-loop with a condition as b_f^* is smaller than the bandwidth b_f asked by a stream f. If the minimum bandwidth of $\mathcal{P}_f^*$ exceeds the currently available path $\mathcal{P}_f$, $\mathcal{P}_f^*$ is still to be null. Then, the routine updates the set of available paths $\mathcal{P}_f^*$ and the bandwidth b_f^*, remove the path of the smaller bandwidth from $\mathcal{P}_f^*$, add a path with the larger bandwidth, update b_f^*, and then push the value of $\mathcal{P}_f^*$ back to Algorithm 1 to allocate available paths. Eventually, the data flow is passed through those suitable and multiple paths in the current network. To better explore the effects of Algorithms 1 and 2, we will present experiments and performance results in Section 4.

4. Performance Results

This section shows the performance of our proposed method in comparison with QUIC, multipath QUIC LRF [5] and the PQUIC schemes [6].

4.1. Experimental Setting

We conducted experiments on the Mininet simulation platform that runs on a computer equipped with an Intel Core i7 processor, 16GB memory, and Ubuntu 18.04.6 LTS. All the algorithmic programs are coded in C language. We used the Wireshark packet analyzer to trace the data flows during the experiments. The following table 2 defines the simulation parameters used in this paper.

Experiments were divided into three sorts with different sizes per data flow: 100, 200, and 400 MB, and produced three measure results of the overall flow completion time, path

Table 2. Simulation Parameter

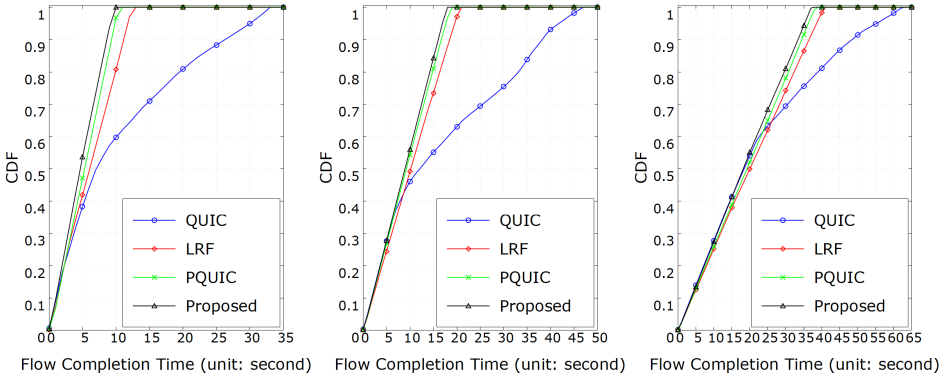
Parameter	Value
Number of nodes in Abilene	11
Number of links in Abilene	14
Data packet size	960-1200 bytes
Transmission bandwidth capacity of a link	100 Mbps
Transmission delay of a link	0-100 ms
Packet loss rate of a link	0.001%
α, coefficient of the measure in 19	0.5
β, coefficient of the measure in 19	0.5
Transmission data size	100MB, 200MB, 400MB

delay, and packet loss rate. We employed the Mininet to adjust simulation parameters. Explicitly, we set $k = 3$, delay coefficient $\alpha = 0.5$ and packet loss coefficient $\beta = 0.5$ as calculating the weighted value p_w. We adopted the Abilene topology [30]: there are 11 nodes and 14 links, the size of each packet is between 960 and 1200 bytes, the path bandwidth is set to 100 Mbps, the delay is from 0 to 100 ms by the binomial distribution, and packet loss rate is set to 0.001%. All experimental cases were run in 20 times to have the results on average.

4.2. Flow Completion Time

Figures 4a, 4b and 4c exhibit the flow completion time in terms of the cumulative distribution function (CDF). As observed, the performance by naive QUIC is the worst, because QUIC only transmits data through a single path, as compared with the other schemes that take multiple paths. Hence, distributing data across multiple paths can obtain better network performance, redundancy, and fault tolerance. It is visible that our scheme outperforms LRF and PQUIC. Explicitly, LRF is based on finding the path with the minimum RTT for transmitting the top-priority data first. Thus, LRF behaves like a greedy way and only focuses on the RTT condition without referring to other network characteristics. The above observations indicate the importance of taking a more comprehensive method for improving network performance.

PQUIC switches between multipaths to ensure that data packets are sent to the receiver fairly. Although PQUIC likely increases the complexity of managing multipath transmissions in dynamic networks, it still suffers from minor performance degradation as path characteristics often change, and as the data size becomes larger. Relatively, our proposed scheme considers both path delay and packet loss rate of path candidates. Such a sophisticated path selection method can lead to better performance to network applications that concern the packet loss. By using a weighting normalization method, it is able to calculate P_w. The higher P_w, the higher priority the data needs to be scheduled for transmission first. This method can dynamically adjust the priorities of data transmission according to network conditions. Our proposed scheme with weighting effects can minimize the flow completion time, resulting in a remarkable comparison with LRF and PQUIC. Thus,



(a) Data size 100 MB **(b)** Data size 200 MB **(c)** Data size 400 MB

Fig. 4. Flow Completion Time

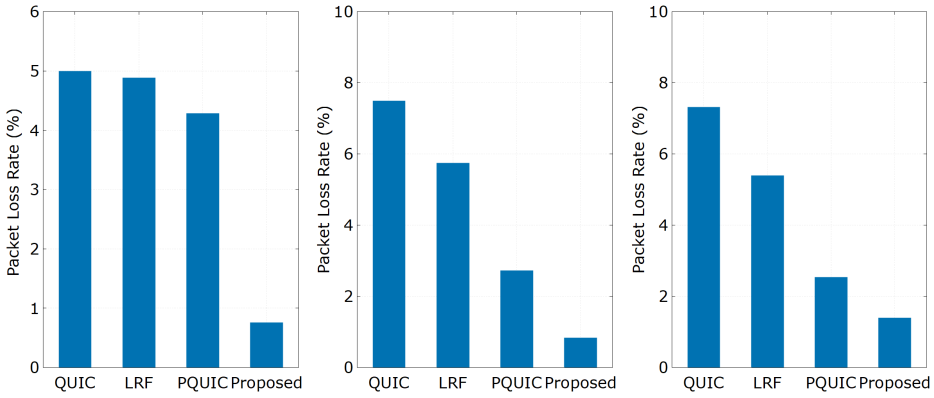
this remarkable result can highlight the importance of intelligent path selection and data prioritization in efficient data transmission and better user experience.

4.3. Packet Loss Rate

Figures 5a, 5b and 5c present the packet loss rate of the overall system performance. As observed, the packet loss rate of QUIC is higher than the other multipath schemes. This is because only the resource allocation of a single path is used, so that the packet loss is apparently affected to a sensitive extent. In the case of data size 100 MB per stream, the packet loss rates of QUIC, LRF, and PQUIC are similar, but become different when the data size per stream increases to 200 MB and even 400 MB. LRF searches for the path of the minimum RTT, which may cause the problem of packet loss in the rear tail of data stream. As examined, this problem is not easy to be resolved when RTT is solely concerned in path selection. PQUIC is fairer as allocating multiple paths to a data stream. Its packet loss rate is lower than the LRF's result. By contrast, our scheme can distribute the data to multiple paths efficiently, thereby being less susceptible to the increase of data size per stream. As seen, our scheme is able to cope with the packet loss rate to be lower than 1 % regardless the increasing data size from 100 MB to 400 MB. Therefore, experimental results provide insights into the relationship between different packet loss rates, data sizes, and transmission schemes. The above findings help in understanding the variance of network performance during the multipath data transmission.

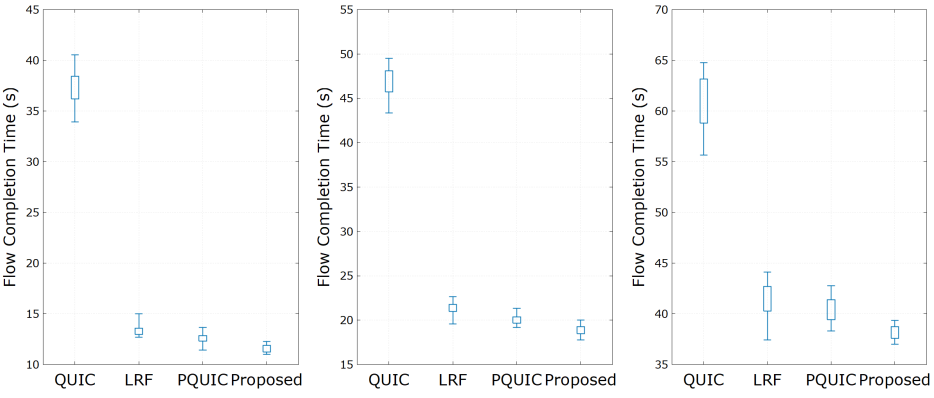
4.4. Overall system stability

Figures 6a, 6b and 6c depict the quartile distribution of flow completion time when the experiment launched 20 data flows one by one repeatedly. Obviously, QUIC needs to take much more time to accomplish the transmission of per data flow. The time gap between QUIC and three multipath QUIC scheme is apparent. This phenomenon shows that employing multiple paths schemes can bring a positive influence on reducing the flow



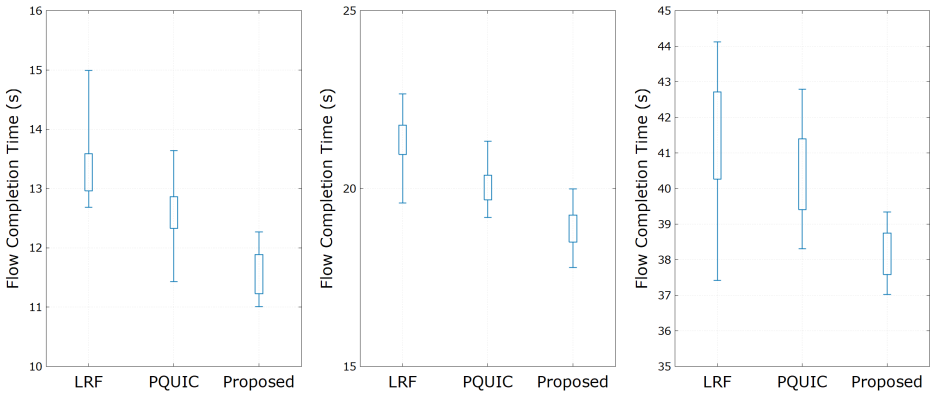
(a) Data size 100 MB **(b)** Data size 200 MB **(c)** Data size 400 MB

Fig. 5. Packet Loss Rate



(a) Data size 100 MB **(b)** Data size 200 MB **(c)** Data size 400 MB

Fig. 6. Overall system stability with QUIC



(a) Data size 100 MB **(b)** Data size 200 MB **(c)** Data size 400 MB

Fig. 7. Overall system stability without QUIC

completion time. Instead, Figures 7a, 7b and 7c exhibit a clear view on the time gap of three multipath QUIC schemes. LRF has not only a larger completion time but also a wider quartile distribution than PQUIC and our scheme. That is, LRF's flow complete time is inconsistent with high variance. We examined that as compared with our scheme, PQUIC cannot perfectly allocate data packets to paths. As the amount of data packets increases rapidly, the probability of head-of-line blocking will increase and then affect the data throughput. Therefore, the results by our scheme are obvious with a minor quartile distribution and the lowest flow completion time. In other words, our scheme can offer stable transport performance since data flows are completed efficiently and with relatively low variability.

5. Conclusion

This paper designs a novel data transport scheme based on MPQUIC. Compared with the traditional network protocol TCP, MPQUIC is based on UDP and keeps the advantages of QUIC from a single-path to multi-path data transport. Our proposed MPQUIC scheme is able to joint sustain transmission delay and packet loss rate with respect to data flows. Performance study is conducted by comparing the proposed scheme with three prior schemes, i.e., QUIC, LRF, and PQUIC. It is remarkable that our proposed scheme performs efficiently and stably in terms of the flow completion time in the system. When the flow completion time is reduced significantly, this scheme also exhibits the effectiveness of reducing path delay and lower packet loss rate under comparative cases with different sizes of data flows.

Our future research will continue to implement MPQUIC and measure the network transport performance in more complicated network scenarios with emerging AR/VR applications, particularly in mobile environments. On the other hand, we notice the adoption of machine learning techniques in internet traffic engineering and management. Our study will further incorporate edge intelligence to network hosts for pro-actively allocat-

ing network resources to data flows and streams. Potential effects on network throughput, security, load balancing and user experiences will be investigated.

Acknowledgments. This work was supported in part by the National Science and Technology Council, Taiwan (R.O.C.), under Contracts MOST-109-2221-E-008-051, NSTC-111-2221-E-008-064 and NSTC-111-2410-H-262-001.

References

1. R. Fielding, J. Gettys, J. Mogul, H. Frystyk, L. Masinter, P. Leach, and T. Berners-Lee, "Hypertext transfer protocol–http/1.1," Tech. Rep., 1999.
2. J. Iyengar and M. Thomson, "QUIC: A udp-based multiplexed and secure transport," https://datatracker.ietf.org/doc/html/rfc9000, May 2021, accessed: 2023-7-30.
3. P. Megyesi, Z. Krämer, and S. Molnár, "How quick is quic?" in *2016 IEEE International Conference on Communications (ICC)*. IEEE, 2016, pp. 1–6.
4. C. Raiciu, C. Paasch, S. Barre, A. Ford, M. Honda, F. Duchene, O. Bonaventure, and M. Handley, "How hard can it be? designing and implementing a deployable multipath {TCP}," in *9th USENIX symposium on networked systems design and implementation (NSDI 12)*, 2012, pp. 399–412.
5. T. Viernickel, A. Froemmgen, A. Rizk, B. Koldehofe, and R. Steinmetz, "Multipath quic: A deployable multipath transport protocol," in *2018 IEEE International Conference on Communications (ICC)*. IEEE, 2018, pp. 1–7.
6. Q. De Coninck, F. Michel, M. Piraux, F. Rochet, T. Given-Wilson, A. Legay, O. Pereira, and O. Bonaventure, "Pluginizing quic," in *Proceedings of the ACM Special Interest Group on Data Communication*, 2019, pp. 59–74.
7. R. Marx, J. Herbots, W. Lamotte, and P. Quax, "Same standards, different decisions: A study of quic and http/3 implementation diversity," in *Proceedings of the Workshop on the Evolution, Performance, and Interoperability of QUIC*, 2020, pp. 14–20.
8. S. Ha, I. Rhee, and L. Xu, "Cubic: a new tcp-friendly high-speed tcp variant," *ACM SIGOPS operating systems review*, vol. 42, no. 5, pp. 64–74, 2008.
9. N. Cardwell, Y. Cheng, C. S. Gunn, S. H. Yeganeh, and V. Jacobson, "Bbr: Congestion-based congestion control," *Communications of the ACM*, vol. 60, no. 2, pp. 58–66, 2017.
10. Q. De Coninck and O. Bonaventure, "Multipath quic: Design and evaluation," in *Proceedings of the 13th international conference on emerging networking experiments and technologies*, 2017, pp. 160–166.
11. H. Zeng, L. Cui, F. P. Tso, and Z. Zhang, "Optimizing multipath quic transmission over heterogeneous paths," *Computer Networks*, vol. 215, p. 109198, 2022.
12. V. A. Vu and J. Wolff, "Supporting delay-sensitive applications with multipath quic and forward erasure correction," in *Proceedings of the 17th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, 2021, pp. 95–103.
13. X. Shi, L. Wang, F. Zhang, B. Zhou, and Z. Liu, "Pstream: Priority-based stream scheduling for heterogeneous paths in multipath-quic," in *2020 29th International Conference on Computer Communications and Networks (ICCCN)*. IEEE, 2020, pp. 1–8.
14. V. A. Vu and B. Walker, "On the latency of multipath-quic in real-time applications," in *2020 16th International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob)*. IEEE, 2020, pp. 1–7.
15. X. Shi, L. Wang, F. Zhang, and Z. Liu, "Fstream: Flexible stream scheduling and prioritizing in multipath-quic," in *2019 IEEE 25th International Conference on Parallel and Distributed Systems (ICPADS)*. IEEE, 2019, pp. 921–924.

16. A. Rabitsch, P. Hurtig, and A. Brunstrom, "A stream-aware multipath quic scheduler for heterogeneous paths," in *Proceedings of the Workshop on the Evolution, Performance, and Interoperability of QUIC*, 2018, pp. 29–35.
17. X. Shi, F. Zhang, and Z. Liu, "Prioritybucket: a multipath-quic scheduler on accelerating first rendering time in page loading," in *Proceedings of the Eleventh ACM International Conference on Future Energy Systems*, 2020, pp. 572–577.
18. Z. Zheng, Y. Ma, Y. Liu, F. Yang, Z. Li, Y. Zhang, J. Zhang, W. Shi, W. Chen, D. Li *et al.*, "Xlink: Qoe-driven multi-path quic transport in large-scale video services," in *Proceedings of the 2021 ACM SIGCOMM 2021 Conference*, 2021, pp. 418–432.
19. J. Wang, Y. Gao, and C. Xu, "A multipath quic scheduler for mobile http/2," in *Proceedings of the 3rd Asia-Pacific Workshop on Networking 2019*, 2019, pp. 43–49.
20. W. Yang, S. Shu, L. Cai, and J. Pan, "Mm-quic: Mobility-aware multipath quic for satellite networks," in *2021 17th International Conference on Mobility, Sensing and Networking (MSN)*. IEEE, 2021, pp. 608–615.
21. R. S. Mogensen, C. Markmoller, T. K. Madsen, T. Kolding, G. Pocovi, and M. Lauridsen, "Selective redundant mp-quic for 5g mission critical wireless applications," in *2019 IEEE 89th Vehicular Technology Conference (VTC2019-Spring)*. IEEE, 2019, pp. 1–5.
22. S. Zhang, W. Lei, W. Zhang, Y. Guan, and H. Li, "Congestion control and packet scheduling for multipath real time video streaming," *IEEE Access*, vol. 7, pp. 59 758–59 770, 2019.
23. H. Wu, Ö. Alay, A. Brunstrom, S. Ferlin, and G. Caso, "Peekaboo: Learning-based multipath scheduling for dynamic heterogeneous environments," *IEEE Journal on Selected Areas in Communications*, vol. 38, no. 10, pp. 2295–2310, 2020.
24. E. Quevedo Caballero, M. Donahoo, and T. Cerny, "Fairness analysis of deep reinforcement learning based multi-path quic scheduling," in *Proceedings of the 38th ACM/SIGAPP Symposium on Applied Computing*, 2023, pp. 1772–1781.
25. S. Lee and J. Yoo, "Reinforcement learning based multipath quic scheduler for multimedia streaming," *Sensors*, vol. 22, no. 17, p. 6333, 2022.
26. Z. Wang and J. Crowcroft, "Quality-of-service routing for supporting multimedia applications," *IEEE Journal on selected areas in communications*, vol. 14, no. 7, pp. 1228–1234, 1996.
27. C.-L. Hu, C.-Y. Hsu, and W.-M. Sung, "Fitpath: Qos-based path selection with fittingness measure in integrated edge computing and software-defined networks," *IEEE Access*, vol. 10, pp. 45 576–45 593, 2022.
28. R. M. Karp, *Reducibility among combinatorial problems*. Springer, 2010.
29. J. Y. Yen, "Finding the k shortest loopless paths in a network," *management Science*, vol. 17, no. 11, pp. 712–716, 1971.
30. S. Knight, H. X. Nguyen, N. Falkner, R. Bowden, and M. Roughan, "The internet topology zoo," *IEEE Journal on Selected Areas in Communications*, vol. 29, no. 9, pp. 1765–1775, 2011.

Fang-Yi Lin received the M.S. degree from the Department of Communication Engineering, National Central University, Taiwan.

Wu-Min Sung is currently a Ph.D. student in the Department of Communication Engineering, National Central University, Taiwan. Her research areas are mobile computing and distributed networks.

Lin Hui is currently an associate professor with the department of computer science and information engineering, Tamkang University, Taiwan. Her research interests include machine learning, multimedia applications, and mobile information systems. She has published some journal articles, book chapters, and conference papers related to these re-

search fields. She had served as journal guest editor/reviewer, and program co-chair/chair for many international conferences and workshops.

Chih-Lin Hu received the PhD degree in electrical engineering from the National Taiwan University, in 2003. He was a researcher with BenQ Advanced Technology Center, Taiwan in 2003–2007. In 2008, he joined with the National Central University, Taiwan, and has been a full professor since August 2022. His research interests include mobile and pervasive computing, distributed networks, and Internet of Things.

Nien-Tzu Hsieh is currently a Ph.D. student in the Department of Communication Engineering, National Central University, Taiwan. Her research areas include mobile computing and Internet of Things.

Yung-Hui Chen is currently a full professor and vice-chairman of the Department of Computer Information and Network Engineering, LungHwa University of Science and Technology, Taiwan.

Received: August 18, 2023; Accepted: October 06, 2023.

Computer Science and Information Systems 21(2):645–661 https://doi.org/10.2298/CSIS230820011K

A study on fire data augmentation from video/image using the Similar-label and F-guessed method

Jong-Sik Kim and Dae-Seong Kang*

Dept. of Electronics Engineering, Dong-A University,37 Nakdong-daero 550 beon-gil Saha-gu, Busan, Korea
{kjsluke, dskang}@dau.ac.kr

Abstract. When data collection is limited, such as in the case of fire detection, improving the detection rate with only number of small labeled data is difficult. Therefore, researchers have conducted many related studies, among which semi-supervised learning methods have achieved good results in improving detection rates. Most recent semi-supervised learning models use the pseudo-label method. But there is a problem, which is that it is difficult to label accurately in samples that deviate from the true label distribution due to false labels. In other words, due to the pseudo-label used for data augmentation, erroneous biases can be accumulated and adversely affect the final weights. To improve this, we proposed a method of generating Similar-labeled data (prediction result labeling value and correct answer value are similar), which was used through the F-guessed method and the Region of Interest (ROI) expression method in the video during initial learning. This has the effect of preventing the bias from being distorted in the initial stages. As a result, data generation increased by about 6.5 times, from 5,565 to 41,712, mAP@0.5 increased by about 26.1%, from 65.9% to 92.0%, and loss improved from 3.347 to 1.69, compared to the initial labeled data.

Keywords: semi-supervised learning, deep learning, pseudo-labeling, fine-tuning, Similar-label, F-guessed.

1. Introduction

The semi-supervised learning method has developed increasingly in computer vision over the past few years. Currently, the most advanced methods introduce hybrid methods by simplifying previous work or combining them with other formulas in the aspect of architectures and loss functions [1]. However, supervised learning is the most used method in the field of deep learning. Supervised learning is a learning method for memorizing learning patterns. It is not easy to identify data that has never been learned before. A lot of labeled data must be required for better generalization [2]. In addition, obtaining large numbers of labeled data in areas where labeling requires expertise or the labeling process takes a long time may be difficult. To improve this problem, Dong- Hyun Lee proposed a pseudo labeling method [3]. The pseudo labeling method is a simple method that can be used for both classification and regression. But there is a limit to improving performance and challenging to match the correct label if a sample is out of the distribution of the labeled answer [4].

* Corresponding authors

However, numerous Semi-Supervised Learning (SSL) papers inspired by the pseudo-labeling method have been published [5, 6, 7, 8, 9]. Among them, MixMatch [10], ReMix-Match [11] and FixMatch [12] announced by Google tried various methods to supplement the problems of pseudo label. MixMatch is training by applying entropy minimization to labeled and unlabeled data. Unlabeled data is labeled using the pseudo labeling method. Pseudo-labeling is sensitive to parameter tuning as it is a method of combination of various mechanisms. Therefore, it requires careful parameter tuning. Nowadays, semi-supervised learning models are mainly using the pseudo- labeling method. When pseudo labels are used, incorrect bias will be stacked due to the pseudo labels. If not solving the data bias, it will learn a biased decision boundary of a specific data sample unlikely the actual labeled data. It can be complicated to use current methods when there are constraints on labeled data, such as in the case of a fire event. Sometimes, there may be errors in recognizing data if it was not included in the learning data. This means that the collected answer label data distribution may not be able to cover all the data.

In this paper, we suggested the following ways to minimize data bias when collecting the data. Instead of the pseudo-labeling method, apply the Similar-labeling method, which uses Region of Interest (ROI) on a video to get labels which are close to the answer. To classify no correct answer label data more precisely, using guessed label after fine-tuning the existing method. Instead of learning all the data at once, extracting guessed labels from half quantities (2,187 pcs) of the initial data (5,565 pcs) and using the extracted data for the next step learning model. To improve the fire recognition rate and significantly reduce the time required for human labeling by minimizing the training bias in several steps. Fig. 1 is a diagram of fire data creation that extracts Similar-label by setting the ROI of suggested algorithms and using Intersection Over Union (IOU) comparison.

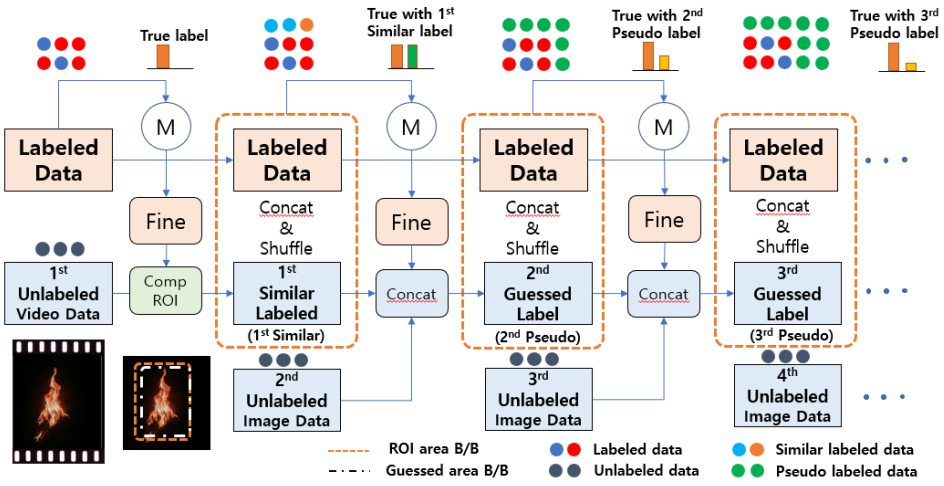


Fig. 1. Conceptual diagram of fire data augmentation using Similar-label and F-guessed comparison method

2. Related work

Semi-supervised learning can be considered if there are few correct answer-labeled data and many labeled data without correct answers. Semi-supervised learning aims to improve performance by applying supervised learning for a few correct answer labels and applying Unsupervised learning for many labeled data without correct answers. Various semi-supervised learning methods have appeared from the perspective of using labeled data without answers for learning. Semi-supervised learning has emerged to collect correct answer data and reduce the resources and costs for labeling work. Objective Function of semi-supervised learning can be expressed as minimizing the sum of supervised learning loss Ls and unsupervised learning loss Lu as in equation (1).

$$Loss = Ls + Lu \tag{1}$$

Semi-supervised learning can be seen as modeling the essential characteristics of the data itself, moving away from the model of the correct answer of the label. It means that the generalization performance can be improved with a small number of learning through a small number of true-label data. Studies similar to the currently proposed technology include pseudo-labeling, MixMatch and FixMatch.

2.1. Pseudo-labeling

Pseudo-label is a popular method because it is very simple. Based on the predicted values of the models sufficiently learned by supervised learning, we attach pseudo-label to the unlabeled data with simple rules such as threshold. The model is then re-learned by combining labeled data and pseudo-labeled data [5]. Fig. 2 shows the basic concept of the pseudo-label method very well.

2.2. MixMatch

Recently, semi-supervised learning algorithms get supervised loss for labeled data and unsupervised loss for unlabeled data. A method of learning a model using these two losses is widely used. Entropy minimization, Consistency loss and MixUp methods were suggested for Unsupervised loss. MixMatch is a supervised learning algorithm that encompasses the three methods. In Fig. 3 shows the MixMatch operation.

- Entropy minimization: The classifier minimizes the predictive entropy of labeled data without an answer, and one of the methods of entropy minimization is pseudo-labeling.

- Mixup: Mixup is a method that mixes augmented answer labels and without answer labels and overlaps the answer and without answer labeled data images for the data.

- Consistency regularization: Using answer labels and without answer labels for learning the data. When similar or modified data are offered to learn, the result has to present similar results.

The algorithm performed better than existing semi-supervised learning algorithms even when using only a small number of labeled data. When correct answer labeled data (X) and labeled data without answers (u) provide for the MixMatch algorithm, it will

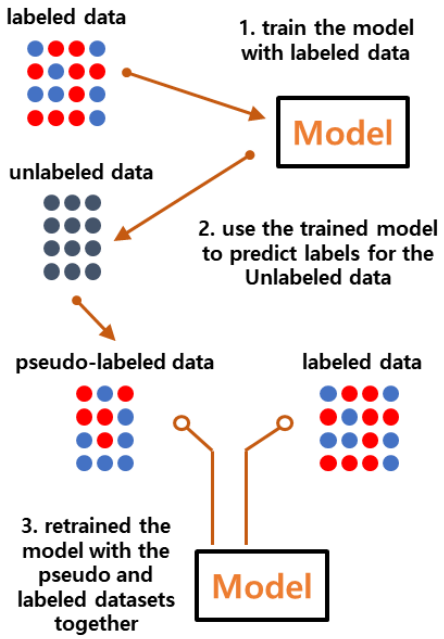


Fig. 2. Pseudo-labeling operation

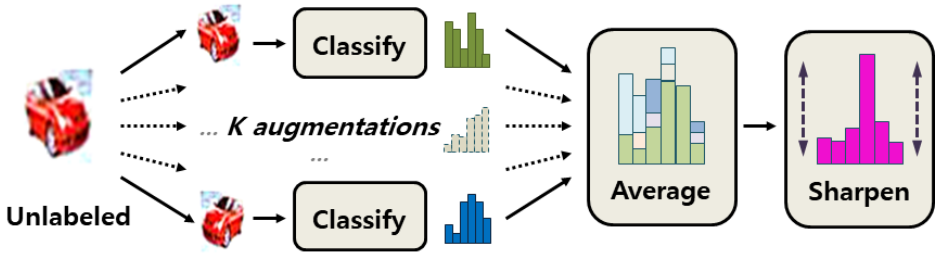


Fig. 3. MixMatch operation

generate processed answer labeled samples ($X^{'}$) and predicted guessed labeled ($u^{'}$). Officially, coupling loss L for semi-supervised learning is defined as equation (2) [10][11].

$$\begin{aligned} X', u' &= MixMatch\left(X, u, T, K, a\right) \\ Lx &= \frac{1}{|X'|} \sum_{xpeX'} H\left(P, Pmodel\left(y|x;\theta\right)\right) \\ Lu &= \frac{1}{L|u'|} \sum uQeu' ||q - Pmodel\left(y|u;\theta\right)||_2^2 \\ L &= Lx + \lambda uLu \end{aligned} \tag{2}$$

H(p, q) is the cross entropy between distributions p and q, and T, K, α, λu are hyperparameters.
T : sharpening temperature.
K : number of unlabeled augmentations.
α : Beta distribution for MixUp.
λu: unsupervised loss of weight.

2.3. FixMatch

FixMatch is a method of training a supervised learning model from correct answer-label images using cross-entropy loss. To get two images by applying weak and strong augment methods for each image of labels without a correct answer. Weakly augmented images are passed on to the model, prediction for the class is obtained, and the probability of the most confident class is compared to a threshold. Use the class as the basic label (pseudo-label) if it is higher than the threshold. After that, strongly augmented images are passed on to the model and proceed with predictions for the class. The predictions can be used as cross-entropy loss to compare with the answer pseudo-label. At this point, combining two losses and optimizing the model. In Fig. 4, the FixMatch Realization method is schematized [12].

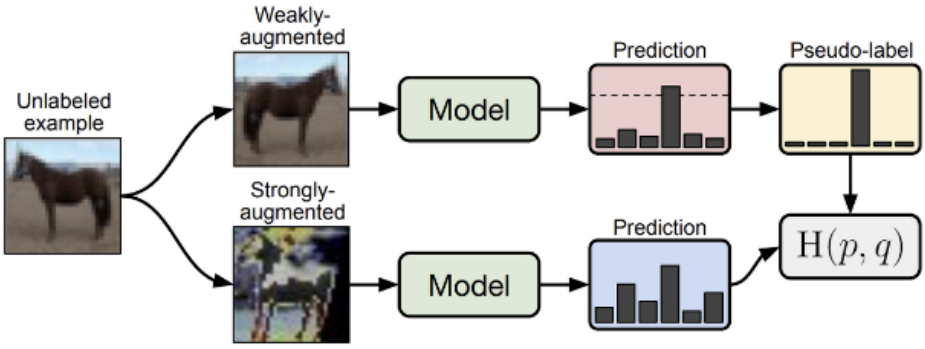


Fig. 4. FixMatch operation

2.4. Fine-tuning

Fine-tuning transforms an architecture to fit image data for new purposes based on previously learned models and updating learning from already learned model weights. In deep learning, fine-tuning means injecting additional data into the existing model to update parameters. For more detail, fine-tuning can be considered as precise parameter tuning. To finish the Fine-tuning, the existing learned layer data must be additionally trained to update the parameters. If it uses completely random initial parameters or a less abstracted layer that learns general features, this will collapse the entire parameters because of overfitting. To change the purpose of the pre trained model for needs, fine-tuning is required with one strategy from four strategies in Fig. 5 [13].

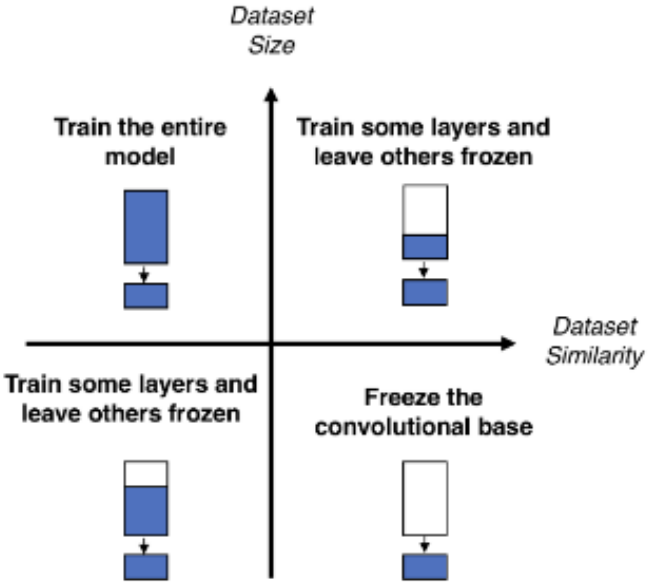


Fig. 5. Types of fine-tuning

The first quadrant is a big dataset but differs from the pre-trained model dataset. Because the dataset is big, the dataset can train a model from the beginning and proceed with all works. The second quadrant uses a big dataset similar to the dataset of the pre-trained model. Since the dataset is large, overfitting will not be an issue and can be learned effectively. The third quadrant uses a small dataset which is opposed to the dataset of the pre-trained model. It is hard to find a balance between the quantity of trainable layer and the same amount of layer, and it could be overfitting. The fourth quadrant is the small dataset but uses all the pre-trained models' datasets. This method changes only the last Fully Connected (FC) and trains a new classifier [14].

3. Proposal method

As mentioned in the introduction, the weakness of the pseudo-label is when the learning model is overfitted to one side and has a bias, and the bias is also applied when generating the pseudo-label. In other words, since the weights are shared, learning through potentially false pseudo-labels is risky. In case of limited data collection, such as fire, it is inevitable to have more distorting bias. In addition, " A Study on Fire Data Generation and Recognition Rate Improvement using F-guessed and Semi-supervised Learning " previously studied by the author [15] is also a model trained by the pseudo-labeling method. Which extracts images per frame from fire videos and uses fire pseudo-labeled, so overfitting to one side, we had no choice but to have the bias accumulated.

3.1. Similar-labeled data using ROI

In this study, the Region of Interest (ROI) was set in the Fire image to prevent false biases from being included in the weights during initial learning. When generating F-guessed, the decision boundary detected within this ROI area obtains the pseudo-labeled data most similar to the labeled data (correct answer or true label). In other words, since the existing pseudo-label data utilizes an unlabeled dataset, it is impossible to know how much wrong bias it has for which class because there is no label information [12]. However, Similar-labeled data has the most similar class and decision boundary to the labeled data. Fig. 6 shows a process of setting an ROI using unlabeled video and extracting Similar-labeled data. For more details, set the ROI for the fire part in the video images and calculate the IOU of the decision boundary (B_d) and boundary of ROI (B_{ri}), occurring near the ROI. If the difference is less than 50%, use for Similar-labeled data. Equation (3) shows the calculation method.

$$B_d \, and \, B_{ri} \, of \, IOU = \frac{\mathrm{B}_d \cap B_{ri}}{\mathrm{B}_d \cup B_{ri}} \tag{3}$$

And as shown in Fig. 7, a fire gradually increases over a certain period when it is ignited. This means that the shape of the fire will vary as long as the camera is not moving, but the size of the fire will remain similar to its size until the fire expands. Based on this, when extracting a decision boundary from a fire video, set ROI on the video of the fire point. Until the fire expands significantly, the shape and form of the fire mostly change within these ROIs. As a result, gathering a considerable amount of fire data similar to labeled data without the need for separate labeling tasks each time is possible.

The disadvantage of this study is that the Region of Interest (ROI) must be drawn once on the fire image. However, the initial ROI display has more advantages than disadvantages in improving overfitting due to incorrect fire labeling in a state with little fire-labeled data at the beginning of learning. In this study, relabeling was performed closer to labeled data to minimize mislabeling that may occur when the number of true labeling data is small. As a result, a similar labeling technique improved the recognition rate to minimize misrecognition when predicting fire image data.

3.2. Fine-tuning

The reason for applying fine-tuning is to transform the architecture to suit the image data for a new purpose based on the previously learned model and to update the learning from

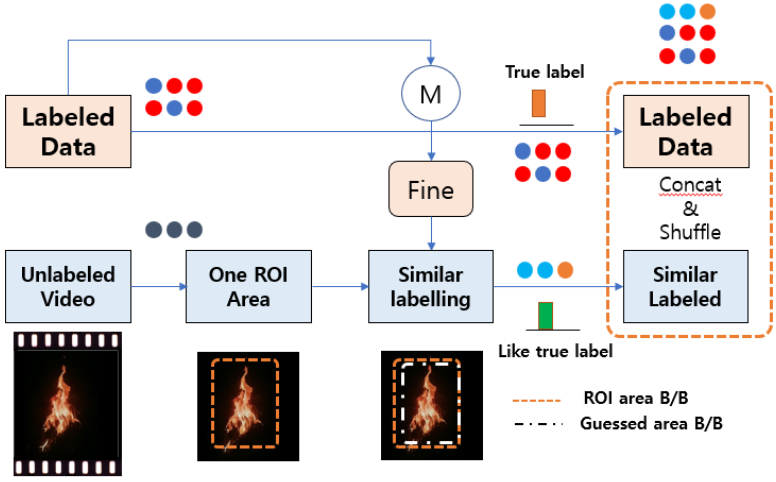


Fig. 6. Conceptual diagram of initial fire data generation using Region of Interest (ROI) comparison method

Fig. 7. The shape and size of fire in the ROI(Region of Interest) in the video

the already learned model weights. The parameters of the less abstract layer that learned the general features were added to prevent overfitting. An optimization process is added by learning a previously learned layer and updating parameters. Fine-tuning means re-learning and optimizing processes using existing neural networks. This is because labels that are more similar to the true labels can be predicted if label data without correct answers is predicted(guessed) after precise parameter tuning of the existing learning model [13].

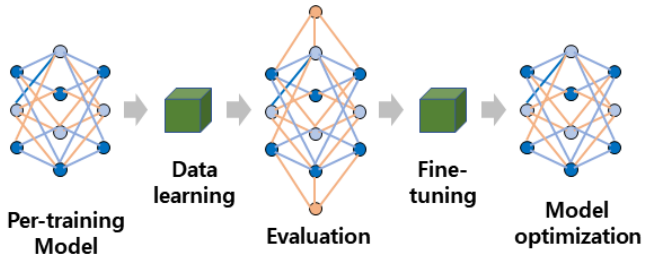


Fig. 8. Fine-tuning optimization method

Fig. 8 shows an optimization method through fine-tuning. It is designed to perform additional fine-tuning each time new data is added, and a new prediction model is created by re-mixing the existing labeled dataset, similarly labeled dataset, and guessed labeled dataset using the additional fine-tuning learning result.

3.3. Step-by-step data growth and redundant labeling

Instead of learning all data at once, it is a method of extracting a guessed label with about half the quantity (2187 pcs) of the initial labeled data quantity (5,565 pcs) and using it as the next step of the learning model. The label was continuously increased by about half its initial quantity. This is because the training process is divided into stages to minimize initial overfitting [7]. Also, the initial labeled data (true label) was used only for learning purposes and was not used as F-guessed data. In other words, for semi-supervised learning, labeled data is always used only as learning data (labeled data) regardless of the learning order, and no transformation is made by labeling. Unlabeled data and Similar-labeled data are designed so that final prediction labeling is always applied according to the learning order for semi-supervised learning. It means the true label + prediction (guessed + similarity) data combines and mixes the true label and the correct answer prediction label to create a new step model for semi-supervised learning. A new fine-tuning is performed using this learning weight value [8]. Fig. 9. is a conceptual diagram for redundant labeling.

4. Experimental Results

A research experiment on how to generate fire data from a video using the F-guessed method was conducted in a computer environment with CPU: AMD Ryzen 7 3700X 8-

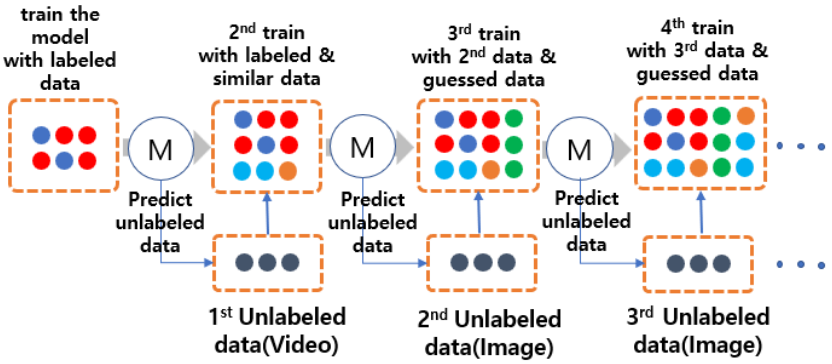


Fig. 9. Conceptual diagram of step-by-step data increase and redundant labeling

Core Processor 3.6 GHz, GPU: NVIDIA GeForce RTX 8000, and 32GB of RAM. Moreover, CNN used Darknet 53, and an object detector has experimented with yolov4 [16]. Table 1. shows the initial labeled dataset information. The numbers in this table mean the number of images, and even in the actual fire image, Person, Smoke, and Spark also include a considerable number of overlapping labels depending on the image. In addition, these images secured data using the Internet [17], fire department site photos, and self-data augmentation methods (Using its own DA-FSL augmentation method [18]).

Table 1. Basic labeled data set information.

Data	Fire	Person	Smoke	Spark	Total
Q'ty	2585	1500	634	846	5565

As shown in Table 1, the experiment was conducted to determine the impact of false bias on pseudo-label during learning when there is not enough initial data. Fig 10 shows false labeling image results from an experiment using unlabeled video data.

To prevent false bias from being included in the weight when initial learning data is insufficient, a region of interest (ROI) was marked on the fire video to obtain pseudo-labeled data most similar to the labeled data when generating pseudo labels. Then, the decision boundary detected within the ROI area was checked to exclude incorrect Labeled data or change Labeling to secure Similar labeled data that was most similar to the correct answer. Since the existing pseudo-label data uses an unlabeled dataset and does not have labeled data information, it was hard to know how much incorrect bias it had for which class. However, Similar labeled data has the most similar class and decision boundary to labeled data using the ROI method.

Table 2 shows the quantity of data augmentation and total image quantity at each stage of fire data generation using the Similar-label and F-guess method. 5,565 pcs correct answer labels used in the initial learning are labeled by humans (labeled data). Similar labeled data close to the correct answer labels were generated using the ROI in the video.

Fig. 10. Red color B/B indicates ROI, Top Left (TL) image is incorrectly recognized as a spark, Top Right (TR) image is incorrectly recognized as Fire and Person, and Bottom Left (BL) image is Fire and Person. In the case of Bottom Right (BR) images, it is mistakenly recognized as smoke

By using the unlabeled data images, table 2 shows the F-guessed quantities guessed by the labeled data. F-guessed quantity increases as it repeats its steps with the final weight values obtained from F-guessing, learning, and labeling on video/image. Except for existing labeled data, added Unlabeled data will repeat learning and labeling in every step. Minus numbers in F-guessed columns are numbers of deleted images with no label in labeling steps.

Table 2. F-guessed labeled data set augmentation information

Data	Labeled Q'ty	Unlabeled Q'ty	F-guessed Q'ty	Division
Basic labeled data	5,565	0	0	image
1st augmentation	5,565	2,783	2,783	Similar label(video)
2nd augmentation	5,565	4,175	6,956	image
3rd augmentation	5,565	6,261	12,976 (-242)	image
4th augmentation	5,565	9,391	22,609 (-416)	image
5th augmentation	5,565	14,087	36,696 (-548)	image

In Table 3, the results of the change in fire recognition rate over five times by applying the Similar-label and F-guess method based on the learning model of the initial answer labeled data are displayed in the order of Loss, mIOU, and mAP. Compared to the ini-

tial correct label data, Loss decreased by up to 1.66%, mIOU increased by 26.6% and mAP@0.5 improved by 27.1% as a result of the test. Additional learning was not conducted after the fifth round because the standard for finishing the program was set based on a small change in loss. It was judged that the low loss meant that the consistency of the labeling data was secured.

Table 3. Object precision rate test results based on max batch = 8,000.

Mode		Loss(%)	mIOU(%)	mAP(%)
Basic labeled data	Train	3.347	52.23	65.93
(True labeled)	Fine-tuning	3.060	56.12	70.67
1st augmentation	Train	2.783	56.35	67.48
(Similar labeled)	Fine-tuning	2.63	59.64	75.42
2nd augmentation	Train	2.70	65.88	75.03
(F-Guess labeled)	Fine-tuning	2.413	65.53	77.22
3rd augmentation	Train	1.958	69.33	78.7
(F-Guess labeled)	Fine-tuning	1.828	70.09	79.30
4th augmentation	Train	1.66	73.44	87.00
(F-Guess labeled)	Fine-tuning	1.516	76.16	87.45
5th augmentation	Train	1.815	76.57	90.67
(F-Guess labeled)	Fine-tuning	1.69	78.84	92.0

Fig. 10 shows the effect of the wrong bias on pseudo labels during learning with a lack of primary learning data. And Fig. 11 compares and displays the results of the labeling image that has changed since applying F-guessed with Similar-labeled data. In more detail, the initial learning model learned with early primary labeled data inevitably results in mislabeling, which in turn causes misrecognition. Therefore, to minimize erroneous labeling at the beginning of learning, the program was modified to exclude images for erroneous labeling within the Region of Interest (ROI) or automatically change them to fire classification labels. This proposed method is named similar labeling because it re-labels similar to the correct answer. As a result, the mislabeling that occurs in Basic labeled data is significantly improved after using Similar-labeled data, as shown in Fig. 11.

In Fig. 12, each stage's change in fire image recognition rate is displayed from 1st to fifth. The image data used for each order results from testing by randomly selecting general images not used for learning from the Internet. The result shows many things that could be improved when initially proceeding with a small number of labeled data. However, it shows stable results as the additional labels continue to increase. Then, only the images showing the greatest difference among several images were selected.

Image No.1 identified fire correctly but kept changing the smoke direction during the learning processes. Image No. 2 correctly identified fire but struggled with recognizing smoke at first. However, through the learning process, it improved recognition precisely over time by smoke and clouds. Image No.3 also recognized fire correctly and smoke kept changing through the learning process. Initially, fire recognition was accurate even with a small amount of data. However, due to limited data, both misrecognition and unrecog-

Fig. 11. Comparison of labeling image results changed after applying Similar-labeled data and F-guessed

nition occurred. However, increasing the data using the F-guessed method resolved these issues.

Table 4 presents the experimental results for ”F-guessed” and ”Similar-label and F-guessed”. The results are based on 36,749 manually labeled labels by humans and 5,565 initial answer labels. Comparing manual labeling with Similar-label, the result improved Loss by 0.69, mIOU by 9.42% and mAP by 13.66% as a result. Also, compared with the existing F-guessed method, Similar-label improved performance considerably.

Table 4. Manual labeled, F-guessed and Similar-labeled data comparison experiment tables

Data	Q'ty	Loss(%)	mIOU(%)	mAP(%)
Basic labeled data	5,565	3.347	52.23	65.93
Manual labeled	36,749	2.38	69.42	78.34
F-guessed labeled	35,633	1.41	78.22	82.49
F-guessed + Similar-labeled	41,712	1.69	78.84	92.0

In comparison to the previously studied F-guessed labeled method, incorrect bias significantly affects the recognition rate improvement in the initial stages. However, the Similar-labeled method enhanced recognition rate accuracy by approximately 10% compared to the present method.

5. Conclusions

In this paper, if data collection is limited, such as in a fire or disaster, the paper proposes a Similar labelling method to improve recognition rates when only a small amount of labeled data is available. The current pseudo-labeling method has limitations in improving performance because it is difficult to accurately label samples that are out of the distribution of correct labels. Therefore, a method of marking a Region of Interest (ROI) in a fire video was used to prevent false biases from being included in the weights during initial learning. This is method automatically changes to a fire class label when the decision boundary detected within the ROI area is recognized as an incorrect class label when the initial pseudo label is created. In this way, Similar-labeled data most similar to the true labeled data can be obtained. As a result, loss decreased by up to 1.66% compared to the initial basic label data, mIOU increased by 26.6%, and mAP@0.5 improved by 26.1%. Also, the number of secured data was 41,712 F-guessed data, which increased by 6.5 times based on the initial true label data of 5,565. And, through additional research in the future, we plan to further study the false recognition rate of fire through uncertainty distribution by using the Bayesian Neural Network to improve false recognition of fire.

Acknowledgments. This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government(MSIT)(No.RS-2023-00247045).

Fig. 12. Comparison of labeling image results changed after applying Similar-labeled data and F-guessed

References

1. Amit Chaudhary.: Semi-Supervised Learning in Computer Vision", https://amitness.com /2020/07/ semi-supervised-learning [accessed: Sep. 10, 2022]
2. Yassine Ouali, Céline Hudelot, and Myriam Tami.: An Overview of Deep Semi-Supervised Learning, Machine Learning (cs.LG), arXiv:2006.05278, Jul. 2020.
3. Dong-Hyun Lee.: Pseudo-label: The simple and efficient semi-supervised learning method for deep neural networks", In ICMLW, 2013.
4. Vinko Kodžoman: Pseudo-labeling a simple semi-supervised learning method, https://dataw hatnow.com/pseudo-labeling-semi-supervised-learning [accessed: Apr. 10, 2023]
5. Hieu Pham, Zihang Dai, Qizhe Xie and Quoc V. Le.: Meta Pseudo Labels, Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11557-11568, 2021.
6. Baixu Chen, Junguang Jiang, Ximei Wang, Pengfei Wan, Jianmin Wang, and Mingsheng Long: Debiased Self-Training for Semi-Supervised Learning, Advances in Neural Information Processing Systems 35 (NeurIPS 2022), arXiv:2202.07136v5, Nov 2022.
7. Barret Zoph, Golnaz Ghiasi, Tsung-Yi Lin, Yin Cui, Hanxiao Liu, Ekin Dogus Cubuk, and Quoc Le.: Rethinking Pre-training and Self-training, Neural Information Processing Systems 33, 2020.
8. Mengde Xu, Zheng Zhang, Han Hu, Jianfeng Wang, Lijuan Wang, Fangyun Wei, Xiang Bai, and Zicheng Liu.: End-to-End Semi-Supervised Object Detection with Soft Teache, IEEE/CVF International Conference on Computer Vision (ICCV), pp. 3060-3069, 2021.
9. Xiaokang Chen, Yuhui Yuan, Gang Zeng, and Jingdong Wang.: Semi-Supervised Semantic Segmentation with Cross Pseudo Supervision, Computer Vision and Pattern Recognition (CVPR), pp. 2613-2622, Jun. 2021.
10. David Berthelot, Nicholas Carlini, Ian Goodfellow, Nicolas Papernot, Avital Oliver, and Colin A Raffel.: MixMatch: A Holistic Approach to Semi-Supervised Learning, Neural Information Processing Systems 32, 2019.
11. David Berthelot, Nicholas Carlini, Ekin D. Cubuk, Alex Kurakin, Kihyuk Sohn, Han Zhang and Colin Raffel.: ReMixMatch: Semi-Supervised Learning with Distribution Alignment and Augmentation Anchoring, Machine Learning (stat.ML), arXiv:1911.09785, Feb. 2020.
12. Kihyuk Sohn, David Berthelot, Nicholas Carlini, Zizhao Zhang, Han Zhang, Colin A. Raffel, Ekin Dogus Cubuk, Alexey Kurakin, and Chun-Liang Li.: FixMatch: Simplifying Semi-Supervised Learning with Consistency and Confidence, Neural Information Processing Systems 33, 2020.
13. Ananya Kumar, Aditi Raghunathan, Robbie Jones, Tengyu Ma, and Percy Liang.: Fine-Tuning can Distort Pretrained Features and Underperform Out-of-Distribution, Computer Vision and Pattern Recognition(cs.CV), arXiv:2202.10054, Feb. 2022.
14. Marius Mosbach, Maksym Andriushchenko, and Dietrich Klakow.: On the Stability of Fine-tuning BERT: Misconceptions, Explanations, and Strong Baselines, Machine Learning (stat.ML), arXiv:2006. 04884, Mar. 2021.
15. Jong-Sik Kim and Dae-Seong Kang.: A Study on Fire Data Generation and Recognition Rate Improvement using F-guessed and Semi-supervised Learning, The Journal of Korean Institute of Information Technology, Vol. 20, pp. 123-134, Dec 2022.
16. Alexey Bochkovskiy, Chien-Yao Wang, Hong- Yuan, and Mark Liao.: YOLOv4: Optimal Speed and Accuracy of Object Detection, Computer Vision and Pattern Recognition (cs.CV), arXiv:2004. 10934, Apr. 2020.
17. AI Hub data, https://aihub.or.kr [accessed: Apr. 10, 2023]
18. Hye-Youn Lim, Jun-Mock Lee and Dae-Seong Kang.: A Method for Improving Learning Convergence Curve and Learning Time of DA-FSL Model using Knowledge Distillation, The Journal of Korean Institute of Information Technology, Vol. 18, pp. 25-32, Oct. 2020.

Jong-SiK KIM received the B.S. in Electronic Engineering from Pukyong National University, South Korea, in 1991; received his master's degree in electronics engineering at Dong-A University, South Korea, in 2020; currently as a doctoral student in Department of Electronic Engineering at Dong-A University; His current research interests Image processing and AI.

Dae-Seong Kang (Corresponding author: dskang@dau.ac.kr) received his Ph.D. in electrical engineering at Texas T&M University, USA, in 1994; He is a professor at electronics engineering in Dong-A University; Main research directions: Image processing, pattern recognition and machine learning.

Received: August 20, 2023; Accepted: November 21, 2023.

Computer Science and Information Systems 21(2):663–683 https://doi.org/10.2298/CSIS230822012W

Multi-language IoT Information Security Standard Item Matching based on Deep Learning ⋆

Yu-Chi Wei[1], Yu-Chun Chang[2], and Wei-Chen Wu[3]

[1] National Taipei University of Technology
Taipei, Taiwan
vickrey@mail.ntut.edt.tw
[2] National Taipei University of Technology
Taipei, Taiwan
t109ab8013@ntut.org.tw
[3] Department of Finance, National Taipei University of Business
Taipei, Taiwan
weichen@ntub.edu.tw

Abstract. In the realm of IoT information security and other domains, various information security standards exist, such as the IEC 62443 series standards published by the International Electrotechnical Commission and ISO/IEC 27001 by the International Organization for Standardization. Business organizations are striving to improve and protect their operations through the implementation and study of these information security standards. However, comparing or pinpointing applicable control measures is becoming increasingly labor-intensive and prone to errors or deviations, especially given the plethora of information standards available. Identifying specific control measures scattered across different information security standards is gradually becoming an important issue. In this research, we utilise a range of domestic and international information security standards as the foundation, employing text mining and deep learning methods to map the similar parts of control measures between standards, thereby enhancing the efficiency of comparison tasks and allowing human resources to be allocated to more pertinent issues.

Keywords: Information Security, Information Security Standards, IoT Security, Text mining, Deep Learning.

1. Introduction

With the proliferation of Internet of Things (IoTs) technologies, everyday life has become increasingly digitized. IoT devices have a wide range of practical applications, whether in office environments, transportation, financial transactions, healthcare, or even in standard household smart appliances [22]. Broadly speaking, any device that can connect to the internet falls into this category, from those with basic network functionality to those combining various sensor devices, specialized software, or even capable of receiving and transmitting data from other complex IoT devices. The advent of IoT and the digital economy is a double-edged sword, on one hand making our lives more convenient to some extent, but on the other hand, escalating the information security threats associated with IoT devices and applications.

⋆ An extended version of The 12th Frontier Computing Conference/FC2022 paper

In the current era where the Internet of Things (IoT) is burgeoning and the interconnection of all things is becoming a trend, the potential risks behind its applications warrant our deep reflection and assessment. Revising new standards is a time-consuming and labour-intensive project, requiring information security professionals to reference, organise, and summarise the contents of various different standards. In this research, based on textual data exploration, existing international IoT standards are automatically preprocessed into numerous features, and then trained using deep learning models. This enables the automatic analysis of existing standards' information security requirements and their alignment with those of other international IoT information security standards. Finally, members of the standards drafting unit can directly refer to and assess whether the automatically generated corresponding results are suitable for use, thus saving a substantial amount of labour and time costs.

This research aims to utilise text mining to automatically translate and reference various existing international IoT standards. After textual preprocessing, these standards are trained using machine learning and deep learning models. The objective is to segment and automatically analyse the information security requirements of the existing standards, matching them with the requirements listed in other international IoT security standards. This not only assists the Mobile Application Security Alliance in continually updating IoT security verification standards, but also allows for the practical examination of whether domestic information security standard-setting processes comply with international IoT security standards. This study uses both domestic and international information security standard content as its dataset, with the capability to swiftly identify similar content. Furthermore, the content is not limited to being in the same language, and the overall output process can be finely tuned based on the input dataset to achieve the best matching results. This application is not limited to comparing and analysing the content of information security standards alone. It can also be based on other existing data and literature to explore and analyse their similarities, providing a reference for researchers looking to implement text processing, text analysis, machine learning, deep learning, and information security standards in their workflow.

2. Related Works

2.1. IEC 62443 Standards

IEC 62443 is a series of international standards for Industrial communication networks - IT security for networks and systems, which contains a series of technical procedures for the security of control systems, and the standard classified the user roles into operator, integrator and manufacturer, designs risks and potential problems for each role to help users of the standard to design and evaluate their own industrial automation systems and improve network security. The IEC 62443 series of standards is divided into four parts. The first part includes terminology and explanations of concepts related to automated industrial control systems, as well as examples of their use; the second part describes the security planning, operation, and management of the structure of industrial automation and control systems; the third part details technologies related to information security, information security risk assessments, and other definitions concerning information security; the fourth part focuses on the description of various security requirements, including

the product security development lifecycle, components, and technologies. In the process of developing the Mobile Application Security Alliance IoT security certification series, IEC 62443 Part 4-2 [8] (IEC 62443-4-2) is also one of the key reference standard and will be introduced separately in subsequent sections. The table below, Table 1, shows the structure and orientation of each content of the ”IEC 62443” series of standards.

Table 1. The list of IEC 62443 series of standards

Standard structure	Parts	Standard content
General: Defines the standard concepts, models, terminology interpretation and examples, etc.	1-1 TS	Concepts and models
	1-2 TR	Master glossary of terms and abbreviations
	1-3	System security compliance metrics
	1-4	IACS security life cycle and use-cases
Policies and Procedures: Provide defined system management requirements for IACS asset owners and services.	2-1 TS	Secure program requirements for IACS asset owners
	2-2	Security Protection Rating
	2-3 TR	Patch management in the IACS environment
	2-4 IS	Requirements for IACS service providers
	2-5 TR	Implementation guidance for IACS asset owners
System: Security risk assessment and security requirements defined for industrial control systems.	3-1	Security technologies for IACS
	3-2	Security risk assessment and system design
	3-3	System security requirements and security levels
Component: The safety product development process and component safety requirements as defined by the product supplier.	4-1 IS	Secure product development lifecycle requirements
	4-2	Technical security requirements for IACS components

2.2. OWASP Top 10

OWASP, known as the Open Web Application Security Project, is an open, non-profit organization dedicated to helping governments and businesses improve web software security, tools, and technical documentation, as well as gain practical insight into the vulnerabilities and security of the information assets they use. Every few years, OWASP produces a list of the top 10 web application security vulnerabilities and provides some easy ways and directions to educate users on how to avoid these vulnerabilities. Table 2. below shows the ten web application security vulnerabilities pro-posed in ”OWASP Top 10:2021 [18]”.

Despite all the vulnerabilities presented in the OWASP Top 10 are carefully organized and filtered to the top ten most common web application security vulnerabilities of our time, there is still a ranking hierarchy among the vulnerabilities, and the higher the ranking, the more important the web application security vulnerability is in the current information environment.

Among the existing web application security vulnerabilities, there are several items that have appeared in the previous version of the ”OWASP Top 10”, but their ranking has been changed in response to the changing times and environment. For example, A01:

Table 2. The list of OWASP Top 10:2021

Vulnerabilities ID	Vulnerabilities Top 10 of application security ver.2021
A01:2021	Broken Access Control
A02:2021	Cryptographic Failures
A03:2021	Injection
A04:2021	Insecure Design
A05:2021	Security Misconfiguration
A06:2021	Vulnerable and Outdated Components
A07:2021	Identification and Authentication Failures
A08:2021	Software and Data Integrity Failures
A09:2021	Security Logging and Monitoring Failures
A10:2021	Server-Side Request Forgery

Access Control Failure in "OWASP Top 10:2021", which was ranked fifth in the previous version of OWASP Top 10:2017, was moved from fifth to first in the latest version. According to the officials, more than 90% of the applications they tested had a category access failure problem, and the number of occurrences was much higher than other vulnerability categories.

In addition to the ten most common security weaknesses of web applications, OWASP also has responded to the increasing use of APIs and Internet of Things devices in the industry, they presented the "OWASP API Security Top 10" and "OWASP IoT Top 10", which includes ten most common security vulnerabilities of network applications. Despite there is a newer version of "OWASP IoT Top 10", which is the version 2018, but overall and detailed information of "OWASP IoT Top 10:2014 [16]" is relatively more abundant than the 2018 version on the official OWASP website, more information is definitely more helpful for deep learning model to classify information security controls into similar categories, that was the main reason we chose to use "OWASP IoT Top 10:2014 [16]" instead of "OWASP IoT Top 10:2018 [17]". Table 3 below shows the list of top 10 security vulnerabilities of "OWASP IoT Top 10:2014".

Table 3. Introduction to the OWASP IoT Top 10:2014

Vulnerabilities ID	Vulnerabilities Top 10 of Internet of Things ver.2014
I01:2014	Insecure Web Interface
I02:2014	Insufficient Authentication/Authorization
I03:2014	Insecure Network Services
I04:2014	Lack of Transport Encryption
I05:2014	Privacy Concerns
I06:2014	Insecure Cloud Interface
I07:2014	Insecure Mobile Interface
I08:2014	Insufficient Security Configurability
I09:2014	Insecure Software/Firmware
I10:2014	Poor Physical Security

For the showcase of this study, we attempted to make IEC 62443-4-2 controls automatically classified into the closest of the ten specified "OWASP IoT Top 10" categories through text mining and deep learning methods, thus saving the time and cost required for manual comparison of information security standards.

2.3. Text Similarity Matching

Text similarity matching methods are becoming increasingly important in many applications. Existing methods often compute similarity based on shallow syntax or POS tagging or by comparing basic syntax similarity, generating vectors, and then inferring similarity from this set of vectors. However, due to the variability in natural language expression, these methods often struggle to predict actual semantic content and implications. To address these issues, researchers have attempted various approaches. At the beginning, using a lexicon to note the positions of words within sentences, forming a one-hot encoding vector representation. However, this method couldn't link related words. Mikolov et al. citeMikolov2013 tried combining neural networks in their research. Turian et al. [23] tried using pre-trained word representations in conjunction with supervised learning methods as extra features, which showed significant improvements over traditional word embedding methods. These methods evolved into larger frameworks, like sentence embedding [10] or paragraph embedding [11]. Matthew et al. [19] extracted context-sensitive features from language models, integrated these features into training for specific tasks, and gradually began to understand the variability and actual semantic content in language expressions.

Researchers also considered the context and situation of sentences. The Skip-gram model [5] is a renowned method that trains and identifies using the context of target words. WordNet [4] is a network primarily focused on the "word-semantics" in English, storing the structures and potential relationships between words, quantifying the semantic relationship between two different words. ConceptNet [12] uses a dictionary-based embedding model, aligning with the hierarchical structure of predefined words in WordNet, defining various relationships between words. Emrah [7] proposed a method focused on calculating sentence similarity without using machine learning, relying on dependency parsers and lexical embedding models, achieving results better than most traditional methods.

In research on machine learning for text similarity analysis and comparison, Ji and Eisenstein [9] introduced a supervised machine learning method that measures semantic similarity between sentences using a discriminative term or proper noun, in conjunction with a set weighting index, giving higher importance to certain features, then computing sentence similarity. The authors claimed their new method outperforms the widely-used TF-IDF weighting method. Mohamed and Oussalah [15] proposed a similarity calculation method that uses WordNet to obtain dependency relationships for words which based on instances extracted from Wikipedia and normalized Google distance. The normalized Google distance calculates the hit count returned for a set of keywords using the Google search engine. Hassan [6] proposed a method based on Wikipedia's content for context determination, called Salient Semantic Analysis (SSA). Mihalcea et al. [13] combined corpus-based semantic similarity with knowledge-based semantic similarity, using data from WordNet and the British National Corpus, which reduced the error rate compared to traditional methods. Wang et al. [24] focused on the similar and dissimilar parts of

sentences. They constructed a similarity matrix and corresponding vectors for each word meaning, decomposing the resulting matching vectors to identify similar and dissimilar parts, eventually using matrix decomposition to extract sentence vectors to compute sentence similarity.

BERT [2], introduced by Google's AI team in 2018, used BooksCorpus and over 800 million entries and data from Wikipedia for pre-training. The operation is divided into two stages: pre-training and fine-tuning. In the pre-training phase, there are two training methods: Masked LM and Next Sentence Prediction. Then in the fine-tuning phase, the model is adjusted based on specific tasks. BERT performs well in sentence classification, tagging, and text classification. However, Reimers and Gurevych [20] found that while BERT and RoBERTa achieve effects in many sentence regression tasks, such as text semantic similarity, they need to input two sentences to be compared into the model repeatedly until the closest two sentences are found. The excessive computational cost makes BERT unsuitable for semantic similarity searches. Hence, they introduced SBERT (Sentence-BERT). SBERT, unlike BERT, which repeatedly attempts to combine two sentences, calculates the similarity distance between two sentences directly by matching their word embedding representations, significantly reducing computation. This model also achieves good results in some STS and transfer learning tasks.

3. Research Methodology

3.1. Data Pre-processing

In this study, we use python and jupyter notebook as the test environment. In the data pre-processing progress, we first need to retrieve the contents of the information security standard, and split the contents into each column, including its control number, control name and control description as a spreadsheet. After this, the contents of the information security standard form are stored in memory and ready to go.

These manually retrieved control contents in the spreadsheet does not require data pre-processing, they can simply import into the deep learning model in their original format for training. The pre-trained models provided by SBERT [21] are already trained from various types of datasets, familiar with the original word patterns, so there is no need to perform steps such as words and sentences segmentation, word lemmatization, stemming or other data pre-process methods you can find in other NLTK tasks to filter the features.

As shown in the figure above, the following figure is a screenshot of the jupyter notebook after importing and reading the information security standard content into the spreadsheet. This experiment uses IEC 62443-4-2 [8] content as the training set, and tries to classify the content of the controls in each of the ten categories of "OWASP IoT Top 10:2014 [16]" as the test set.

It is also possible to match similar contents between different language information security standards. In the data pre-processing state, while keeping the unique control id number field legible, translation modules can be used to translate control descriptions into the specified language and then perform a similarity comparison exercise with other information security standards. Usually, it is better to translate other languages into English and perform similarity matching between the two standards using English as the common language, because most of the pre-training data for deep learning models are trained from English data as shown in the Fig. 2 below.

```
In [17]: train_blist[0]

Out[17]: 'Components shall provide the capability to identify and authenticate all human
         users according to IEC 62443-3-3 SR 1.1 on all interfaces capable of human user
         access. This capability shall enforce such identification and authentication on
         all interfaces that provide human user access to the component to support segre
         gation of duties and least privilege in accordance with applicable security pol
         icies and procedures. This capability may be provided locally by the component
         or by integration into a system level identification and authentication syste
         m.'

In [18]: test_blist[0]

Out[18]: 'Default passwords and ideally default usernames to be changed during initial s
         etup'
```

Fig. 1. A schematic diagram of train/test data content

	影像監控系統標準_zh	影像監控系統標準_translated
0	產品預設不應透過實體介面存取產品作業系統之除錯模式。 若需經實體介面存取，則應通過身分鑑別作...	Product presets should not be removed by the p...
1	產品應具有實體埠插拔操作記錄功能。	The product should have the function of the ph...
2	產品應具備相關警示功能於實體操作發生斷訊時。	The product should have the relevant warning f...
3	產品外部不應有徒手即可還原預設通行碼的功能。	The outside of the product should not have the...
4	產品應支援安全啟動(Secure Boot)功能，不應以未經授權的韌體、驅動程式及作業系統執...	The product should support the Security Boot f...
5	產品之作業系統與網路服務，不應存在美國國家弱點資料庫所公開的常見弱點與漏洞資料，且漏洞評鑑系...	The operating system and network services of t...
6	產品之作業系統與網路服務，不應存在美國國家弱點資料庫所公開的常見弱點與漏洞資料，且漏洞評鑑系...	The operating system and network services of t...
7	產品開啟之網路服務應為廠商提供必要服務之所需，防止產品因啟用網路介面而被侵入的可能性，且廠商...	The online service of the product opening shou...
8	產品所收集之遙測資料應告知使用者，且未告知之遙測資料不應被收集。	The remote measurement data collected by the p...
9	韌體應具備更新機制。	The ligament should have a update mechanism.
10	產品若支援離線手動更新，則更新檔案應加密保護以確保機密性，且應採用NIST SP 800-1...	If the product supports offline manual update,...
11	產品若支援線上更新，其更新路徑應通過安全通道，且安全通道版本應符合「附錄 A」的要求，同時金...	If the product supports online update, its upd...
12	產品應具備驗證韌體之完整性及真確性的功能。	The product should have the function of verify...
13	產品應具備備援更新功能，即發生更新失敗時，系統能回復至更新前之狀態。	The product should have the renewal function, ...
14	產品所儲存的敏感性資料，應僅由獲授權個體存取。	The sensitive data stored in the product shoul...
15	產品所儲存之身分鑑別因子、加解密用之金鑰(不含非對稱加密用之公鑰)不應明文儲存，而保護資料的...	The identity of the product's identity identit...
16	產品應提出金鑰管理程序，以確保金鑰管理的品質。	The product should propose the golden key mana...

Fig. 2. A schematic diagram illustrating the successful prediction of control items between two different standards

In the Fig. 2 above, we used a local IoT security standard for this showcase, which is "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]" from Mobile Application Security Alliance, which is an IoT product certification alliance dedicated to the promotion of domestic IoT information security in Taiwan. According to the figure, any standards in different languages can be translated into English by the translation module and then start the comparison process of information security standards directly, this allows the process to be able to compile information security standards in different languages without any limitation due to language.

3.2. Model Training

BERT [1], the abbreviation of Pre-training of Deep Bidirectional Transformers for Language Understanding, which is already highly characterized by the endless training data based on the Google search engine, so that BERT only needs to specify the form of its output data, and then fine-tune it according to the task, finally, it can be used for various common natural language processing tasks.

But Reimers and Gurevych found that although both BERT and RoBERTa achieve some good results in many sentences regression tasks, such as textual semantic similarity, they both need to pass both sentences to be compared into the model and repeat this process until the two most similar sentences are found. This is a very costly process, especially when the data is large. According to this, they considered BERT is not suitable for the task of semantic similarity search because of the limitation of the algorithm, so they proposed SBERT [21] (Sentence-BERT), which does not need to try to combine two sentences repeatedly like BERT, but by directly matching and calculating the words similarity distance of two sentences using word's embed-ding representations, which greatly reduces the computational effort and achieves very good results in some STS and migration learning tasks.

The deep learning method SBERT provides a number of pre-training models, which allow users to train their own research data directly to make further predictions. In the official guidance document of SBERT, 13 pre-training models are provided. The 13 pre-training models are listed with their performance of sentence embeddings, performance of semantic search, average overall performance, running speed and model size, so that users can select them according to their task requirements. The five models with the best performance based on the above five indicators were shown as Table 4.

In this study, the best average overall performance one: *all-mpnet-base-v2*, were selected, which was an all-round model tuned for many use-cases, trained on a large and diverse dataset of over 1 billion training pairs.

As shown in the figure above, IEC 62443-4-2 controls were successfully classified by deep learning models into the ten categories of OWASP IoT Top 10:2014. Matching similar contents including controls or descriptions between several information security standards, which often requires a lot of labor and time, but this study showed that it is totally possible to quickly generate similarity comparison results between certain information security standards by using text mining and deep learning methods. It can also be said that this study, corresponding contents between information security standards and standards is also one of the typical NLP tasks, i.e., the application of semantic textual similarity tasks.

Table 4. Comparison of SBERT best performance pre-trained models

Model name	Performance of sentence embeddings	Performance of semantic search	Average overall performance	Encoding speed	Model size
all-mpnet-base-v2	69.57	57.02	63.30	2800	420 MB
multi-qa-mpnet-base-dot-v1	66.76	57.60	62.18	2800	420 MB
distiluse-base-multilingual -cased-v2	60.18	27.35	43.77	4000	480 MB
paraphrase-MiniLM-L3-v2	62.29	39.19	50.74	19000	61 MB
paraphrase-multilingual-mpnet-base-v2	65.83	41.68	53.75	2500	970 MB

	IEC 62443 Controls	OWASP Top 10:2014 Category	Distance
0	5.3 CR 1.1	I5	0.515
1	5.4 CR 1.2	I5	0.519
2	5.5 CR 1.3	I5	0.385
3	5.6 CR 1.4	I5	0.416
4	5.7 CR 1.5	I2	0.469
...	...	...	...
83	15.9 NDR 3.12	I5	0.575
84	15.10 NDR 3.13	I5	0.580
85	15.11 NDR 3.14	I10	0.528
86	15.12 NDR 5.2	I3	0.455
87	15.13 NDR 5.3	I3	0.530

Fig. 3. A comparative schematic illustrating the distance between control measures across different standards

3.3. Evaluation Methodology

In section 3.2, we have demonstrated that it is possible to perform similarity comparisons between information security standards using deep learning methods. But, how was the predictive accuracy? To find out the predictive accuracy of the model, first, a reference answer that cross-validates the model prediction results is necessary. For example, a table which providing an official mapping of the controls of a standard itself to the controls of another standard, such as a table which maps IEC 62443-4-2 [8] controls to EN 303-645 [3] controls. But unfortunately, no such mapping table is provided in the official documents of these two parties.

For this reason, we use the official mapping of Appendix D of IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14], which is a Taiwanese IoT information security standard focused on image monitoring systems, includes a mapping table to the OWASP IoT Top 10:2014 [16], these two information security standard have built a explicit relations between their controls, which allows this study to use the information in this table as a reference for the accuracy of automated comparisons with deep learning models. Fig. 4 below shows a screenshot of the controls in Appendix D of the standard "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements" against each standard specification.

附錄 D

(參考)

技術要求事項與各標準規範對照表

Control's serial number of the standard

表 D.1 技術要求事項與各標準規範對照表

技術要求	OWASP 對應項目(4)	ANSI/UL 2900-1 對應項目	ONVIF 對應項目(19-20)
5.1.1.1	I10： Poor Physical Security Ensuring only required external ports such as USB are required for the product to function. Ensuring the product has the ability to limit administrative capabilities.		-

Fig. 4. A screenshot of the standard " IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements, Appendix D " against OWASP IoT Top 10:2014

The reference answer mapping of the standard comparison is based on the ”IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements” standard, and the official mapping table of the standard to OWASP IoT Top 10:2014 in Appendix D of the standard as the reference answer. In other words, a total of 38 screened security items in the Image Monitor System Information Security Standard will actually be classified into the ten corresponding categories of ”OWASP IoT Top 10:2014”. Although each category of ”OWASP IoT Top 10:2014” has from 4 to 14 information security controls, it was found that it is difficult to match the information security control specified in the reference answer for information security standards from different sources. In addition to the difference in terminology between different standards, it is assumed that the accuracy of the wording of the original Chinese standard will be affected after translation. Therefore, in this section, we choose to convert the accuracy of the base standard information security control into reference values by whether they are correctly classified or not. Figure 5 below shows the schematic diagram of the two experimental approaches.

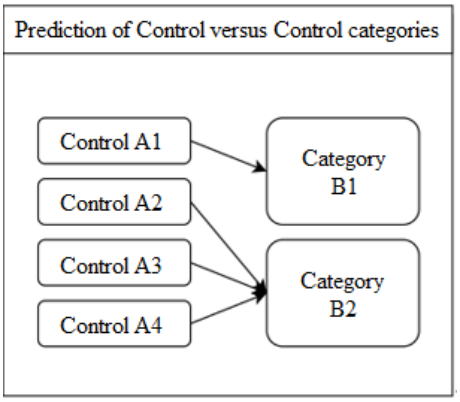


Fig. 5. A schematic diagram of controls versus control categories comparison

As shown in the Fig. 4, which shows that the control numbered 5.1.1.1 of ”IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]” can actually corresponded to category I10 of ”OWASP IoT top 10:2014 [16]”.

However, since the standard itself is written in Chinese, it needs to be translated into English and then fed into a deep learning model for comparison, so we have used the translation module mentioned in section 3.1 to automatically complete this task for us. After checking the table, a total of 38 filtered security controls in the ”IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements” will actually be classified into the ten corresponding categories of ”OWASP IoT Top 10:2014”.

4. Evaluation Results

4.1. Initial Evaluation Results

We used the five models with the best performance in Table 4. and *distiluse-base-multilingual-cased-v2*, which is a multilingual model that supports more than 50 different languages, and more balanced in the scores of the indicators, were selected and compared with the "OWASP IoT Top 10: 2014", and the following Table 5 shows the experimental results.

Table 5. Comparison of experimental results with different SBERT models

Exp. No.	Model	k=1	k=2
S1	all-mpnet-base-v2	61 %	68 %
S2	multi-qa-mpnet-base-dot-v1	50 %	66 %
S3	paraphrase-MiniLM-L3-v2	39 %	50 %
S4	distiluse-base-multilingual-cased-v2	68 %	68 %
S5	paraphrase-multilingual-mpnet-base-v2	50 %	74 %

In the above table, the k represents the prediction of the k most similar outcomes at the end of each prediction. In other words, when the model can output the least number of predictions, the more accurate it can hit the same category of predictions, which means that the model has a better performance on the task of matching information security standards. The number of successful hits is one of the important indicators of the effectiveness of the reference model for this task.

Under this condition, experiment number S1 and S4 have the best performance, which are *all-mpnet-base-v2* and multilingual model *distiluse-base-multilingual-cased-v2*, achieving 61% and 68% hit rate respectively under the restriction of k=1, and 68%, 74% hit rate respectively under the restriction of k=2, which means at least three quarters of controls in the standard were successfully predicted to the correct categories by the deep learning models.

In addition to the difference in terminology between different standards, the accuracy of the wording of the original standard will also be affected if it is translated, not to mention the fact that there are also controls or requirements that meet several OWASP IoT Top 10 categories after review and analysis, but the reference answer only has a given category and thus cannot be included. However, when it comes to the actual use for the information standards, even though they are for the same domain-oriented information security standards, there are some parts that are not similar. In practice, when an information security consultant is looking for controls or requirements that are suitable for a particular case, the items that are suitable for the case may be scattered in different information security standards, or different categories inside the same standard. Among those that are not successful, there must be some items that are not in the same category but have similar practical applications and application methods.

4.2. Discussion of Evaluation Results

In Section 3.2 of this paper, the five SBERT models that performed better on average were compared with the results of the comparison experiments between their scores provided in

the official guidance documents and the English standards. This means that nearly a quarter of the information security items are difficult to classify correctly by the model. The actual list of information security item numbers that were not predicted by each model shows that these unpredictable information security item numbers are specific numbers, as shown in Fig. 6 below shows the prediction status of each model for the specified corresponding security item at k=3. The dark squares indicate that the number was successfully predicted by the specified model, while the light squares indicate that the number was not successfully predicted by the specified model.

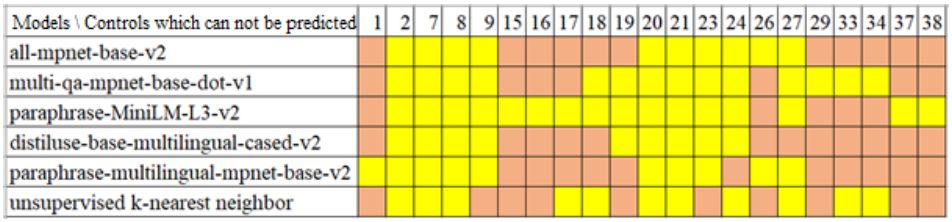


Fig. 6. Standard controls for which none of the plural models can be predicted

The distribution of the light-colored squares in the above figure shows that the information security items that cannot be successfully predicted by the specified models are very similar for the above five deep learning models, especially for Experiment Numbers. 2, 7, 8, 9, 20, 21, and 23. Experiment number 2, 7, 8, 9, 20, 21, 23, these experiment numbers correspond to the following items in the original standard: "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]".

Based on the above table, it can be inferred that it is more difficult for the deep learning models to classify the contents of information security controls into two categories, category 2 and category 8. When encountering the above information security controls in practice, both category 2 and category 8 will not be the first choice of the deep learning models, but other categories. In" OWASP IoT Top 10:2014" [16], category 2 is Insufficient Authentication/Authorization, which translates to unreliable authentication mechanism, and category 8 is Insufficient Security Configurability, which translates to unreliable security configuration.

From the evaluation results, the two deep learning models with the best prediction results, *all-mpnet-base-v2* and *distiluse-base-multilingual-cased-v2*, are the best predicted models for the translated "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]", which corresponds to the prediction of "OWASP IoT Top 10:2014", achieved 61% and 68% hit rate at k=1 respectively. The prediction results are shown in Table 7 below.

According to the above table, the prediction results can be easily classified into two categories, one is the category where Model 1, denoted as $M1$, and Model 2, denoted as $M2$, have the same prediction results for the information security sub-category, but both predict failure. Table 8 explores the potential causes of prediction errors in the classification results. In Experiment Number 2, the correct category should in Category 2, the

Table 6. List of standard controls that cannot be predicted by the majority of models.

Exp. No.	Control Number	Correct Category	Control Description
2	5.1.3.1	2	The product should not have the ability to restore the default pass code with your bare hands.
7	5.2.4.1	8	Sensitive data stored in the product shall be accessible only by authorized individuals.
8	5.2.4.2	8	The identity authentication factor and key for encryption and decryption (excluding the public key for asymmetric encryption) stored in the product should not be stored in clear text, and the data should be protected by the security functions approved by NIST SP 800-140C, CMVP Approved Security Functions.
9	5.2.4.4	8	Sensitive data should be stored in the security domain of the product, isolated from the normal operating environment.
20	5.3.3.1	8	The product should provide the user to turn on/off the WPS PIN function of ”Wi-Fi Protected Setup (WPS)” and its default value should be off.
21	5.3.3.2	8	By default, the Wi-Fi security mechanism should be ”Wi-Fi Protected Access (WPA)” and the version of Wi-Fi Protected Access should meet the requirements of Appendix C.
23	5.4.1.1	2	Before accessing the product resources, the identity identification mechanism with protection against retransmission attacks should be adopted.

Table 7. Prediction results of the two models for the specified standard controls

Exp.No.	Control	$M1$ prediction result	$M2$ prediction result
2	5.1.3.1	10	10
7	5.2.4.1	5	5
8	5.2.4.2	4	4
9	5.2.4.4	5	10
20	5.3.3.1	7	5
21	5.3.3.2	7	7
23	5.4.1.1	10	7

corresponding information security subdivision is that the product should not have the ability to restore the default passcode externally with bare hands. It should be the word ”external” that causes the deep learning model to predict this information security item as Category 10: Poor physical security. In Experiment Number 7, the corresponding information security breakdown is that sensitive information stored in the product should only be accessed by authorized individuals. In terms of this information security control, it is reasonable to predict to Category 5: privacy concerns because it also describes user privacy. In Experiment Number 8, The corresponding information security itemized content is: the identity authentication factor and key for encryption and decryption (excluding the public key for asymmetric encryption) stored in the product should not be stored in clear text, and the data protection method should be used with the security functions approved by NIST SP 800-140C, CMVP Approved Security Functions. In terms of this information security category, the prediction to Classification 4: Lack of Transport Encryption is reasonable because it contains key words in the field of encryption such as encryption and decryption, key and plaintext. In Experiment Number 21, for the information security control, the default security mechanism for Wi-Fi is ”Wi-Fi Protected Access (WPA)” and the version of Wi-Fi Protected Access should meet the requirements of Appendix C. In terms of this information security control, the predicted classification is Category 7: Insecure Mobile Interface, which is not accurate. It is guessed that in the pre-training data of the two models, Wi-Fi usually appears together with key words such as cell phone and mobile, so the models classified it as Category 7.

Table 8. The two models jointly classify the error causes of the false security item

No.	Control	Result ($M1\&M2$)	Correct category
2	5.1.3.1	10	2
7	5.2.4.1	5	8
8	5.2.4.2	4	8
21	5.3.3.2	7	8

The reason behind the inaccurate predicts are that different deep learning models have different prediction judgments for the same information security item, but the classification is basically similar to the former one: it is influenced by specific wording, or the information security item may apply to both plural ”OWASP IoT Top 10:2014” classification, resulting in its misclassification. The actual results of the respective predictions are listed for analysis, and the reasons for the wrong classification results are speculated in Table 9. In Experiment Number 9, the corresponding information security sub-section is: sensitive data should be stored in the security domain of the product, isolated from the normal operating environment. Model 1 predicts that Category 5: Privacy Concerns are reasonable, and sensitive data are indeed related to user privacy; Model 2 predicts that Category 10: Poor Physical Security is not reasonable, and the model presumes that the information security item is not related to physical security because of the terms ”operating environment”, ”isolation”, and ”security domain”. The model predicts that the information security item is related to the description of physical security because of the terms ”operating environment,” ”isolation,” and ”secure area. In Experiment Number 20,

the information security control is: the product should provide users to turn on/off the WPS PIN function of ”Wi-Fi Protected Setup (WPS)”, and the default value should be off. Model 1 predicts Category 7: Insecure Mobile Interface, which is a relatively inaccurate classification. It is guessed that in the pre-training data of both models, Wi-Fi usually appears together with key words such as cell phone and mobile, so the model classifies it as Category 7. After all, if Wi-Fi is automatically connected to public networks, it may cause user privacy leakage, which is a user privacy concern.

In Experiment Number 23, the corresponding information security control is: Before accessing product resources, identity authentication mechanism with protection against retransmission attacks should be used. Model 1 predicts a classification of 10: Poor Physical Security, which is inaccurate. Model 2 predicts a classification of 7: Insecure mobile interface, which is more reasonable than the prediction of Model 1, but not correct. In Experiment Number 21, for the information security control, the default security mechanism for Wi-Fi is ”Wi-Fi Protected Access (WPA)” and the version of Wi-Fi Protected Access should meet the requirements of Appendix C. In terms of this information security control, the predicted classification is Category 7: Insecure Mobile Interface, which is not accurate. It is guessed that in the pre-training data of the two models, Wi-Fi usually appears together with key words such as cell phone and mobile, so the models classified it as Category 7.

Table 9. The two models each classify the wrong security category of the error cause speculation

No	Control	Result($M1$)	Result($M2$)	Correct category
9	5.2.4.4	5	10	8
20	5.3.3.1	7	5	8
23	5.4.1.1	10	7	2
21	5.3.3.2	7	7	8

From Table 9 and the speculation on the failure of the prediction of the security category for which none of the plural models could be predicted, it is clear that at least half of the security categories that failed to be predicted may also apply to the plural "OWASP IoT Top 10:2014[16]" classification, plus the fact that in the standard "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]", the corresponding OWASP In Appendix D of the original Top 10:2014 mapping table, the security subcategory does not specify a mapping to another subcategory, even though the subcategory is similar for that security subcategory, resulting in model prediction failure. By actually viewing the table and the information security controls that failed for the seven security controls that could not be predicted by the plural model, if the predictions that were judged to be reasonable were categorized as correct predictions, with model 1 representing *all-mpnet-base-v2* and model 2 representing *distiluse-base-multilingual-cased-v2*, the two models The final revised prediction results for these seven information security controls are shown in Table 10 below.

Table 10. Results of the error analysis of the two models for the unpredictable security controls breakdown

Exp. No. / Ctrl. No.	2 / 5.1.3.1	7 / 5.2.4.1	8 / 5.2.4.2	9 / 5.2.4.4	20 / 5.3.3.1	21 / 5.3.3.2	23 / 5.4.1.1
Model 1	X	V	V	V	X	X	X
Model 2	X	V	V	X	V	X	X

According to the above table, model 1: *all-mpnet-base-v2* and model 2: *distiluse-base-multilingual-cased-v2* achieve 61% and 68% hit rate respectively for k=1. If the predictions with reasonable classification are classified as correct and recalculated, the hit rate will increase to 69% and 76%.

Finally, the experimental results proved that the use of deep learning models for fast and automated comparison of information security standard content has good accuracy and retains considerable room for improvement.

4.3. Final Evaluation Results

SBERT [21], as an enhanced version of BERT [1] for text similarity search task, provides a pre-training model with higher accuracy than the native pre-training model provided by BERT. Table below shows the best two models in SBERT, *all-mpnet-base-v2*

and *distiluse-base-multilingual-cased-v2*, with the same k=1, i.e., each information security sub-prediction only outputs one closest information security sub-prediction, and this output value is the only consideration for accuracy. Under the condition that the translated "IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements [14]" corresponds to the prediction of "OWASP IoT Top 10:2014"[16].

Table 11. Deep Learning Approach to Information Security Standard Prediction Implementation Results

Exp. No.	Model name	Predict accuracy	Predict / All
SS1	all-mpnet-base-v2	69%	26 / 38
SS2	distiluse-base-multilingual-cased-v2	76%	29 / 38

According to the above table, the better model, *distiluse-base-multilingual-cased-v2*, successfully predicted 26 of the 38 information security items with k=1, while the remaining items failed to be predicted by the plural model in sub section 4.2 of this study. In this study, we examined the seven information security items for which both models failed to predict, and confirmed that three of the items were also applicable to the plural "OWASP IoT Top 10:2014" classification, although they did not match the answers.

In spite of the information security standards targeting the same aspect, there will still be parts where they differ significantly from each other. In practice, when information security consultants are looking for suitable sub-items or control measures for a specific case, the relevant items may be spread across different categories. Among the items that don't align perfectly, there are bound to be some that, while not in the same category, are very similar in practical application and usage. This suggests that, in practical terms, using deep learning models for comparing information security standards has shown, from experimental results, to be not only faster but also fairly accurate. Its performance surpasses the implementation using traditional machine learning. In the future, this research will experiment with generative AI, attempting to produce more general terms related to the control items of different standard, and then apply the SBERT method for further experimentation to enhance the readiness of successful classification.

5. Conclusion

This study utilises the contents of multiple international information security standards and translated domestic standards as its dataset, possessing the ability to rapidly identify similar control items. The content is not restricted to a single language and demonstrates good predictive accuracy. The study also proposes an automated process, streamlining a workflow that would otherwise require significant labour to review and compare. Ultimately, this can serve as a reference for scholars wishing to conduct future research in text processing, text mining, deep learning, and information security standards.

Although this research has achieved commendable results in comparing similarities among different information security standards, there are still many areas that warrant

further exploration in the future. For instance, automated data processing procedures or the application of machine learning methods such as Few-Shot Learning for data with lower volume, greater diversity, and insufficient annotations. Additionally, the use of generative AI represents another avenue to explore. Some standards may feature different customary terminologies across various standards organisations or publishers. Generating more general terms related to control and then utilising the SBERT method for further experiments might enhance the accuracy of successful classifications.

Acknowledgments. This research was partially funded by National Science and Technology Council (NSTC 112-2221-E-027-067-).

References

1. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.: Bert: Pre-training of deep bidirectional transformers for language understanding. arXiv preprint arXiv:1810.04805 (2018)
2. Devlin, J., Chang, M.W., Lee, K., Toutanova, K.N.: Bert: Pre-training of deep bidirectional transformers for language understanding (2018), https://arxiv.org/abs/1810.04805
3. European Telecommunications Standards Institute: EN 303 645:CYBER; Cyber Security for Consumer Internet of Things: Baseline Requirements, v2.1.1 edn. (2020)
4. Fellbaum, C.: WordNet, pp. 231–243. Springer Netherlands, Dordrecht (2010), https://doi.org/10.1007/978-90-481-8847-5_10
5. Guthrie, D., Allison, B., Liu, W., Guthrie, L., Wilks, Y.: A closer look at skip-gram modelling. In: Proceedings of the Fifth International Conference on Language Resources and Evaluation (LREC'06). European Language Resources Association (ELRA), Genoa, Italy (May 2006), http://www.lrec-conf.org/proceedings/lrec2006/pdf/357_pdf.pdf
6. Hassan, S.: Measuring semantic relatedness using salient encyclopedic concepts. Ph.D. thesis (2011), https://www.proquest.com/dissertations-theses/measuring-semantic-relatedness-using-salient/docview/1011651248/se-2
7. Inan, E.: Simit: a text similarity method using lexicon and dependency representations. New Generation Computing 38(3), 509–530 (2020)
8. International Electrotechnical Commission: Security for industrial automation and control systems - Part 4-2: Technical security requirements for IACS components. (2019)
9. Ji, Y., Eisenstein, J.: Discriminative improvements to distributional sentence similarity. In: Proceedings of the 2013 conference on empirical methods in natural language processing. pp. 891–896 (2013)
10. Kiros, R., Zhu, Y., Salakhutdinov, R.R., Zemel, R., Urtasun, R., Torralba, A., Fidler, S.: Skip-thought vectors. In: Cortes, C., Lawrence, N., Lee, D., Sugiyama, M., Garnett, R. (eds.) Advances in Neural Information Processing Systems. vol. 28. Curran Associates, Inc. (2015), https://proceedings.neurips.cc/paper_files/paper/2015/file/f442d33fa06832082290ad8544a8da27-Paper.pdf
11. Le, Q., Mikolov, T.: Distributed representations of sentences and documents. In: Xing, E.P., Jebara, T. (eds.) Proceedings of the 31st International Conference on Machine Learning. Proceedings of Machine Learning Research, vol. 32, pp. 1188–1196. PMLR, Bejing, China (22–24 Jun 2014), https://proceedings.mlr.press/v32/le14.html
12. Liu, H., Singh, P.: Conceptnet—a practical commonsense reasoning tool-kit. BT technology journal 22(4), 211–226 (2004)

13. Mihalcea, R., Corley, C., Strapparava, C.: Corpus-based and knowledge-based measures of text semantic similarity. In: Proceedings of the 21st National Conference on Artificial Intelligence - Volume 1. p. 775–780. AAAI'06, AAAI Press (2006)
14. Mobile Application Security Alliance: IoT-1001-1 v2.0 Image Monitor System Information Security Standard - Part 1: Information Security Requirements (2021)
15. Mohamed, M., Oussalah, M.: A hybrid approach for paraphrase identification based on knowledge-enriched semantic heuristics. Language Resources and Evaluation 54, 457–485 (2020)
16. OWASP IoT Security Team: OWASP Internet of Things (IoT) Top 10 2014. (2014)
17. OWASP IoT Security Team: OWASP Internet of Things (IoT) Top 10 2018. (2018)
18. OWASP IoT Security Team: OWASP Top 10 vulnerability 2021. (2021)
19. Peters, M.E., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L.: Deep contextualized word representations. CoRR abs/1802.05365 (2018), http://arxiv.org/abs/1802.05365
20. Reimers, N., Gurevych, I.: Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In: Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP). pp. 3982–3992. Association for Computational Linguistics, Hong Kong, China (Nov 2019), https://aclanthology.org/D19-1410
21. Reimers, N., Gurevych, I.: Sentence-bert: Sentence embeddings using siamese bert-networks. arXiv preprint arXiv:1908.10084 (2019)
22. Swamy, S.N., Kota, S.R.: An empirical study on system level aspects of internet of things (iot). IEEE Access 8, 188082–188134 (2020)
23. Turian, J., Ratinov, L., Bengio, Y.: Word representations: a simple and general method for semi-supervised learning. In: Proceedings of the 48th annual meeting of the association for computational linguistics. pp. 384–394 (2010)
24. Wang, Z., Mi, H., Ittycheriah, A.: Sentence similarity learning by lexical decomposition and composition. In: Proceedings of COLING 2016, the 26th International Conference on Computational Linguistics: Technical Papers. pp. 1340–1349. The COLING 2016 Organizing Committee, Osaka, Japan (Dec 2016), https://aclanthology.org/C16-1127

Yu-Chih Wei is an Associate Professor in the Department of Information and Finance Management at the National Taipei University of Technology. He holds a Ph.D. in Information Management from National Central University, and a B.S. and a M.S. in Information Management from YuanZe University. His research interests include FinTech security, health informatics security, ISRA, SupTech, VANET security, information security management, and business continuity management. Before pursuing an academic career, Dr. Wei was a researcher at the Information & Communication Security Laboratory of Chunghwa Telecom Co., Ltd.

Yu-Chun Chang received his M.S. degree in Department of Information and Finance Management, National Taipei University of Technology in 2023. His research interests include information security and text mining.

Wei-Chen Wu is Assistant Professor in the Department of Finance at the National Taipei University of Business. He received his Ph.D. degree in Information Management from National Central University in 2016. From 2020-2021, He was Assistant Professor in the Department of Finance at the Feng Chia University. From 2008-2016, he was also

Assistant Professor and Director of the Computer Center at Hsin Sheng College of Medical Care and Management. His teaching interests lie in the area of programming languages, ranging from theory to design to implementation and his current research interests include blockchain technology, fintech cybersecurity, network security, and deep learning. Wei-Chen Wu has collaborated actively with researchers in several other disciplines of computer science. He has served on many conference and workshop program committees and served as the workshop chair for Frontier Computing Conference (FC2017 FC2021) and Machine Learning on FinTech, Security and Privacy Conference (MLFSP2019 MLFSP2023).

Received: August 22, 2023; Accepted: November 21, 2023.