Complexity

# Theory, Applications, and Challenges of Cyber-Physical Systems 2021

Lead Guest Editor: Hang Su
Guest Editors: Bo Xiao, Mingchuan Zhou, Wen Qi, Juan Sandoval, and Seong Tae Kim

# Theory, Applications, and Challenges of Cyber-Physical Systems 2021

Complexity

# Theory, Applications, and Challenges of Cyber-Physical Systems 2021

Lead Guest Editor: Hang Su
Guest Editors: Bo Xiao, Mingchuan Zhou, Wen Qi, Juan Sandoval, and Seong Tae Kim

# Contents

**Theory, Applications, and Challenges of Cyber-Physical Systems 2021**
Hang Su, Bo Xiao, Mingchuan Zhou, Wen Qi, Juan Sandoval, and Seong Tae Kim
Editorial (3 pages), Article ID 9861298, Volume 2022 (2022)

**The Influence of Figures in Warning Signs at the Manual Toll Station on the Lane Change Timing of Drivers in the Context of Virtual Reality of High-Proportion ETC Vehicles**
Su Zhi, Peng Qunjie, Wang Chaolun, Xiang Wang, and Ling Chen
Research Article (8 pages), Article ID 7121952, Volume 2022 (2022)

**Planning the Emergency Collision Avoidance Strategy Based on Personal Zones for Safe Human-Machine Interaction in Smart Cyber-Physical System**
Thanh Phuong Nguyen, Hung Nguyen, and Ha Quang Thinh Ngo
Research Article (21 pages), Article ID 2992379, Volume 2022 (2022)

**A Robot Human-Like Learning Framework Applied to Unknown Environment Interaction**
Xianfa Xue, Lei Zuo, and Ning Wang
Research Article (10 pages), Article ID 5648826, Volume 2022 (2022)

**Simulation Testing of Maritime Cyber-Physical Systems: Application of Model-View-ViewModel**
Dong-Chul Lee, Kyung-Min Seo, Hee-Mun Park, and Byeong Soo Kim
Research Article (14 pages), Article ID 1742772, Volume 2022 (2022)

**Optimization of Noise Transfer Path Based on the Composite Panel Acoustic and Modal Contribution Analysis**
Qiang Liu, Xuanyi Zhou, Jianxin Zhu, and Xiaoping Gong
Research Article (15 pages), Article ID 3059865, Volume 2021 (2021)

**An Improved Framework for Content- and Link-Based Web-Spam Detection: A Combined Approach**
Asim Shahzad, Nazri Mohd Nawi, Muhammad Zubair Rehman, and Abdullah Khan
Review Article (18 pages), Article ID 6625739, Volume 2021 (2021)

**Approximate Inertial Manifold-Based Model Reduction and Vibration Suppression for Rigid-Flexible Mechanical Arms**
Lisha Xu, Hua Deng, Chong Lin, and Yi Zhang
Research Article (17 pages), Article ID 8290978, Volume 2021 (2021)

**LSKE: Lightweight Secure Key Exchange Scheme in Fog Federation**
Yashar Salami and Vahid Khajehvand
Research Article (9 pages), Article ID 4667586, Volume 2021 (2021)

**Complex Electrical Stimulation Systems in Motor Function Rehabilitation after Spinal Cord Injury**
Jiange Kou, Maolin Cai, Fei Xie, Yixuan Wang, Na Wang, and Meng Xu
Review Article (16 pages), Article ID 2214762, Volume 2021 (2021)

**Pursuer Navigation Based on Proportional Navigation and Optimal Information Fusion**
Shulin Feng, Zhanxin Li, Li Liu, Hongyong Yang, Yuanhua Yang, Gaohuan Lv, and Liang Yang
Research Article (11 pages), Article ID 6245168, Volume 2021 (2021)

**Flow Field Analysis of Adult High-Flow Nasal Cannula Oxygen Therapy**
Jingen Xia, Jiaqi Chang, Jixiang Liang, Yixuan Wang, and Na Wang
Research Article (11 pages), Article ID 4981691, Volume 2021 (2021)

Hindawi
Complexity
Volume 2022, Article ID 9861298, 3 pages
https://doi.org/10.1155/2022/9861298

*Editorial*

# Theory, Applications, and Challenges of Cyber-Physical Systems 2021

**Hang Su,[1] Bo Xiao,[2] Mingchuan Zhou,[3] Wen Qi,[1] Juan Sandoval,[4] and Seong Tae Kim[5]**

[1]*Department of Electronics, Information and Bioengineering, Politecnico di Milano, 20133 Milan, Italy*
[2]*Hamlyn Centre for Robotic Surgery and the Department of Computing, Imperial College London, London, UK*
[3]*College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou, China*
[4]*Department of GMSC, Pprime Institute, CNRS-ENSMA-University of Poitiers, UPR 3346, Poitiers, France*
[5]*Department of Computer Science and Engineering, Kyung Hee University, Yongin, Republic of Korea*

Correspondence should be addressed to Wen Qi; wen.qi@ieee.org

Received 20 December 2021; Accepted 20 December 2021; Published 23 June 2022



Cyber-physical systems (CPS) is a field that encompasses various computational and engineering disciplines, such as complex networks and communications, control, information fusion, signal processing, optimization, and computer science, as well as their applications [1–3]. The increasing number of complex system applications, together with more requirements and demands for system intelligence, flexibility, adaptability, robustness, and resilience, is revealing new theoretical and technological challenges for advanced cyber-physical systems [4–6].

A major challenge is the need for innovative frameworks and algorithms to deal with complex cyber-physical systems [7, 8]. Increasing complexities can be due to complex components, for example, some applications of advanced robot manipulators, sensor networks, multiagent systems, human-in-the-loop systems, and also complex environments. In particular, as unstructured environments in manufacturing with multi-spatial and temporal scales, uncertain human-robot interaction in the complex healthcare system, or intelligent control and multimodal information fusion in large-scale complex distributed systems, is the urgent problem to be solved [9–11]. These theoretical, methodological, or practical focus complexities make high-precision and reliable real-time performance difficult to achieve and therefore require the design and development of novel methods for cyber-physical systems [12]. Advancements in cyber-physical systems will provide engineering systems with new capabilities that far exceed today's levels of intelligence, autonomy, functionality, adaptability, robustness, and cybersecurity.

This Special Issue aims to provide a platform for researchers and engineers to report their recent results, exchange research ideas, and overlook emerging research and application directions in "Theory, Applications, and Challenges of Cyber-Physical Systems." After a strict peer-review process, seven papers were selected out of the submissions. The selected papers can be categorized into the following: adaptive and robust control for complex cyber-physical systems, computational methods for internet-of-things-based cyber-physical systems, agent-based modeling and activity recognition in complex human-robot interaction, analytical methods of design, modeling and evaluation of complex cyber-physical systems, artificial intelligence-based analysis, modeling, optimization, and evaluation for complex cyber-physical systems, smart interactions in complex cyber-physical systems, applications of data science for complex cyber-physical systems, intelligent hardware devices and software algorithms in complex cyber-physical systems, multimodal and multichannel signal processing, and information fusion in complex cyber-physical systems.

The paper by Qiang Liu et al. entitled "Optimization of Noise Transfer Path Based on the Composite Panel Acoustic and Modal Contribution Analysis" [13] aims to propose a novel method based on composite panel acoustic and modal contribution analysis and noise transfer path optimization in a vibroacoustic model. Indeed, most studies only consider the panel acoustic contribution of a single frequency, without considering the contribution of major frequencies synthesis to confirm the optimized panels. The effectiveness

of the proposed method is applied and verified in an excavator cab. The sound pressure level (SPL) of the driver's right ear (DRE) decreased obviously. The acoustic analysis of the composite panel acoustic contribution and modal acoustic contribution can more accurately recognize an optimized area than the traditional PACA. This method can be applied in the optimization of the structure-borne transmission path for construction machinery cab and vehicle body.

The paper by Asim Shahzad et al. entitled "An Improved Framework for Content- and Link-Based Web-Spam Detection: A Combined Approach" [14] aims to propose an improved framework for content- and link-based web-spam identification. The framework uses stopwords, keywords' frequency, part of speech (POS) ratio, spam keywords database, and copied-content algorithms for content-based web-spam detection. This paper focused on the issue of web spamming, which is one of the most significant issues encountered by SEs because it dramatically affects the quality of SE results. To address these challenges, they initially exposed the relationship network behind the link-based web spamming and then used the paid-link database, neighbour pages, spam signals, and link-farm algorithms. Finally, they combined all the content- and link-based spam identification algorithms to identify both types of spam for link-based web-spam detection. WEBSPAM-UK2006 and WEBSPAM-UK2007 datasets were used to conduct experiments and to obtain threshold values. A promising F-measure of 79.6% with 81.2% precision shows the applicability and effectiveness of the proposed approach.

The paper by Lisha Xu et al. entitled "Approximate Inertial Manifold-Based Model Reduction and Vibration Suppression for Rigid-Flexible Mechanical Arms" [15] aims to accurately position and control the rigid-flexible mechanical arms with a complex rigid-flexible structure. It proposed a model reduction method of rigid-flexible mechanical arms based on the approximate inertial manifold. To repress the residual vibration of the end of the mechanical arm, a feedforward control strategy is designed. The simulation results depict the superiority of the proposed method, which greatly suppresses the end residual vibration of the mechanical arm and realizes the accurate positioning of the end of the mechanical arm. In addition, the hardware experimental device of the rigid-flexible mechanical arms is constructed, and the experimental verification of the above method is put into effect. The simulation results of angular displacement and end vibration of the reduced model are accordant which is shown by the experimental results of the hardware platform.

The paper by Jiange Kou et al. entitled "Complex Electrical Stimulation Systems in Motor Function Rehabilitation after Spinal Cord Injury" [16] aims to using spinal cord electrical stimulation to alleviate spinal cord injury in an effective way. Hence it summarizes several different spinal cord electrical stimulation methods, analyzes the stimulation effect, and briefly describes the current understanding of their origin and mechanism of action. Finally, the possible development direction and corresponding challenges of spinal cord electrical stimulation in the future are proposed.

The paper by Yashar Salami and Vahid Khajehvand entitled "LSKE: Lightweight Secure Key Exchange Scheme in Fog Federation" [17] focuses on the fog computing architecture which allows data exchange with the vehicle network, sensor networks, etc. They have proposed a lightweight, secure key exchange scheme for the fog federation to reduce computational overhead instead of with a high computational overhead. To prove the lightweight, they have compared the proposed scheme with the Yashar design in terms of computing, and communication cost AVISPA Tool was used for the formal analysis of the proposed scheme. Then, it simulated the proposed scheme with the NS3 tool and compared it with throughput, packet loss, packet delivery, and end-to-end delay with Yashar et al.'s scheme. The results show that the proposed design reduced 3.2457 ms of computational overhead and 1,024 transmitted data bits.

The paper by Shulin Feng et al. entitled "Pursuer Navigation Based on Proportional Navigation and Optimal Information Fusion" [18] focuses on the pursuer navigation based on the three-dimensional proportional navigation law. This method presents a family of navigation laws resulting in a rich behavior for different parameters. Based on point-to-point navigation, obstacle avoidance is implemented by adjusting the control parameters, and the combination can enrich the application range of obstacle avoidance and guidance laws. Finally, simulation results verify the availability of the proposed navigation law.

The paper by Jingen Xia et al. entitled "Flow Field Analysis of Adult High-Flow Nasal Cannula Oxygen Therapy" [19] focuses on the mechanical ventilation of human body, which is treated as a complex human-computer interaction process. The purpose of this paper is to analyze the pressure, flow, and strain rate of the upper respiratory tract with different flow and oxygen concentrations by using finite element simulation, to guide professionals to adjust the appropriate flow and oxygen concentration parameters of the HFNC machine. This paper studies the complex human-computer interaction environment of the human respiratory tract and ventilation airflow. The 3D model of the respiratory tract established by the conversion of image scanning data was taken as the research object. The flow state of the gases in the respiratory tract was judged by Reynolds equation. After that, RNG K-$\varepsilon$ model was applied to the research object, and the simulation diagram of airway pressure, flow rate, and strain rate and trace diagram of flowing particles were obtained under the finite element method. The results explain some clinical phenomena in HFNC and guide people to make better use of mathematical tools to study the human-computer complex environment.

## Conflicts of Interest

The Guest Editors declare that they have no conflicts of interest regarding the publication of this Special Issue.

*Hang Su*
*Bo Xiao*
*Mingchuan Zhou*
*Wen Qi*
*Juan Sandoval*
*Seong Tae Kim*

## Acknowledgments

The Guest Editors would like to express great appreciation to all authors who supported this Special Issue by contributing papers. The Guest Editor team is deeply indebted to all reviewers for their insightful and constructive comments that helped shape this Special Issue.

## References

[1] H. Su, Y. Hu, H. R. Karimi, A. Knoll, G. Ferrigno, and E. De Momi, "Improved Recurrent Neural Network-Based Manipulator Control with Remote Center of Motion Constraints: Experimental Results," *Neural Networks*, vol. 131, pp. 291–299, 2020.

[2] W. Qi, H. Su, and A. Aliverti, "A Smartphone-Based Adaptive Recognition and Real-Time Monitoring System for Human Activities," *IEEE Transactions on Human-Machine Systems*, vol. 50, no. 5, pp. 414–423, 2020.

[3] B. Xiao, H. K. Lam, H. Zhou, and J. Gao, "Analysis and Design of Interval Type-2 Polynomial-Fuzzy-Model-Based Networked Tracking Control Systems," *IEEE Transactions on Fuzzy Systems*, vol. 29, pp. 2750–2759, 2020.

[4] W. Qi, S. E. Ovur, Z. Li, A. Marzullo, and R. Song, "Multi-Sensor Guided Hand Destures Recognition for Teleoperated Robot Using Recurrent Neural Network," *IEEE Robotics and Automation Letters*, vol. 2021, Article ID 3089999, 2021.

[5] B. Xiao, H. K. Lam, Z. Zhong, and S. Wen, "Membership-Function-Dependent Stabilization of Event-Triggered Interval Type-2 Polynomial Fuzzy-Model-Based Networked Control Systems," *IEEE Transactions on Fuzzy Systems*, vol. 28, no. 12, pp. 3171–3180, 2019.

[6] H. Su, A. Mariani, S. E. Ovur, A. Menciassi, G. Ferrigno, and E. De Momi, "Toward Teaching by Demonstration for Robot-Assisted Minimally Invasive Surgery," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 484–494, 2021.

[7] B. Xiao, H. K. Lam, G. Song, and H. Li, "Output-Feedback Tracking Control for Interval Type-2 Polynomial Fuzzy-Model-based Control Systems," *Neurocomputing*, vol. 242, pp. 83–95, 2017.

[8] L. Zhang, W. Qi, Y. Hu, and Y. Chen, "Disturbance-Observer-based Fuzzy Control for a Robot Manipulator using an EMG-Driven Neuromusculoskeletal Model," *Complexity*, vol. 2020, Article ID 8814460, 2020.

[9] X. Zhang, J. Li, W. Qi et al., "A Novel Human-like Control Framework for Mobile Medical Service Robot," *Complexity*, vol. 2020, Article ID 2905841, 2020.

[10] H. Su, W. Qi, Y. Hu, H. R. Karimi, G. Ferrigno, and E. De Momi, "An Incremental Learning Framework for Human-Like Redundancy Optimization of Anthropomorphic Manipulators," *IEEE Transactions on Industrial Informatics*, vol. 18, Article ID 3036693, 2020.

[11] M. Zhou, Q. Yu, K. Huang et al., "Towards Robotic-Assisted Subretinal Injection: A Hybrid Parallel–Serial Robot System Design and Preliminary Evaluation," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 8, pp. 6617–6628, 2019.

[12] W. Qi and A. Aliverti, "A Multimodal Wearable System for Continuous and Real-time Breathing Pattern Monitoring during Daily Activity," *IEEE journal of biomedical and health informatics*, vol. 24, no. 8, pp. 2199–2207, 2019.

[13] Q. Liu, X. Zhou, J. Zhu, and X. Gong, "Optimization of Noise Transfer Path based on the Composite Panel Acoustic and Modal Contribution Analysis," *Complexity*, vol. 2021, Article ID 3059865, 2021.

[14] A. Shahzad, N. M. Nawi, M. Z. Rehman, and A. Khan, "An Improved Framework for Content-and Link-Based Web-Spam Detection: A Combined Approach," *Complexity*, vol. 2021, Article ID 6625739, 2021.

[15] L. Xu, H. Deng, C. Lin, and Y. Zhang, "Approximate Inertial Manifold-Based Model Reduction and Vibration Suppression for Rigid-Flexible Mmechanical Arms," *Complexity*, vol. 2021, Article ID 8290978, 2021.

[16] J. Kou, M. Cai, F. Xie, Y. Wang, N. Wang, and M. Xu, "Complex Electrical Stimulation Systems in Motor Function Rehabilitation after Spinal Cord Injury," *Complexity*, vol. 2021, Article ID 2214762, 2021.

[17] Y. Salami and V. Khajehvand, "LSKE: Lightweight Secure Key Exchange Scheme in Fog Federation," *Complexity*, vol. 2021, Article ID 4667586, 2021.

[18] S. Feng, Z. Li, L. Liu et al., "Pursuer Navigation based on Proportional Navigation and Optimal Information Fusion," *Complexity*, vol. 2021, Article ID 6245168, 2021.

[19] J. Xia, J. Chang, J. Liang, Y. Wang, and N. Wang, "Flow Field Analysis of Adult High-flow Nasal Cannula Oxygen Therapy," *Complexity*, vol. 2021, Article ID 4981691, 2021.

Hindawi
Complexity
Volume 2022, Article ID 7121952, 8 pages
https://doi.org/10.1155/2022/7121952

*Research Article*

# The Influence of Figures in Warning Signs at the Manual Toll Station on the Lane Change Timing of Drivers in the Context of Virtual Reality of High-Proportion ETC Vehicles

**Su Zhi**,[1] **Peng Qunjie**,[2] **Wang Chaolun**,[1] **Xiang Wang**,[1] **and Ling Chen**[3]

[1]*Changsha University of Science and Technology, Changsha 410076, China*
[2]*Shenzhen Transportation Design & Research Institute CO., LTD, Shenzhen 518003, China*
[3]*College of Engineering and Design, Hunan Normal University, Changsha 410081, China*

Correspondence should be addressed to Ling Chen; chenling81@hunnu.edu.cn

Received 6 August 2021; Accepted 10 March 2022; Published 5 May 2022

Academic Editor: Wen Qi



The increase of ETC (electronic toll collection system) vehicles on expressways has changed the proportion of ETC/MTC (manual toll collection system) lanes at toll stations. Based on a driving simulator, three toll gate lane warning sign schemes (scheme for present situation, MTC guidance scheme, and arrow + MTC guidance scheme) were proposed in this study. Driving simulation experiments were conducted to study the influence of figures in warning signs at the manual toll station on the lane change timing of drivers. It was found that the addition of arrows to the warning signs can significantly shorten the response time and guide the driver to make lane change decisions earlier to reduce the congestion between MTC vehicles and the mainline ETC vehicles at the toll plaza, thereby improving the traffic capacity and safety.

## 1. Introduction

In recent years, under the promotion of the Ministry of Transport, China, the proportion of ETC (electronic toll collection system) vehicles on expressways has gradually increased, thus adjusting the proportion of ETC/MTC (manual toll collection system) lanes at toll stations, which has correspondingly changed the behaviors and preferences of drivers. Therefore, with the rapid increase in the proportion of ETC vehicles, the targeted design of warning signs is particularly important.

The design of toll station signs involves many disciplines, among which human factors engineering (HFE) takes an essential part that investigates the optimization of man-machine-environment systems from psychology and physiology to improve system efficiency and ensure human safety, health, and comfort [1]. When designing warning signs, information-rich ones are more in line with the ideas of cognitive ergonomics [2, 3]. Studies have optimized the design or evaluation of warning signs from the perspective of drivers according to the HFE, such as age, emotion, and preference, so as to reduce the traffic accidents caused by the misinterpretation to warning signs [4, 5] (He W [6–8]). In addition, studies have also found that gender, education level, monthly income, and nationality of drivers are related to the understanding of road signs, and male drivers have a better understanding than female ones (Hashim et al., 2002 [9, 10]). Some studies have analyzed the design of warning signs from the aspect of driving state, such as driver distraction (David, 2015 [8]). Furthermore, studies have explored the recognition, understanding, and response of drivers to warning signs in different contexts [11, 12] and investigated among dyslexia, understanding of sign, and situational awareness [10, 13].

Research on the design of warnings signs at toll stations has been helping the drivers to find the correct toll lane at the toll station by optimizing figures and text in the signs. By summarizing the results of previous field and laboratory experiments, Ullman et al. [14] found that graphic signage is easier to interpret for drivers who have difficulty

Table 1: Other information of the participants.

| Information | Category | N | Proportion (%) |
|---|---|---|---|
| *Average number of toll booths passing per year* | 0 | 10 | 23.26 |
| | 1~5 | 11 | 25.58 |
| | 6~12 | 8 | 18.6 |
| | ≥13 | 14 | 32.56 |
| *Average annual driving mileage (k*km)* | <5 | 9 | 20.93 |
| | 5~10 | 9 | 20.93 |
| | 11~30 | 15 | 34.88 |
| | ≥31 | 10 | 23.26 |

understanding words. Zwahlen et al. (2000) analyzed the eye movement data and lane change behavior data and found that the graphic signs on the highway ground can guide the driver to change lanes earlier. Skowronek [15] improved the graphic signs on the Houston expressway overpass and found that graphic signs can improve drivers' understanding of signs and reduce the cost of sign production. Wang et al. [16] assessed the effects of adding graphics to changeable message signs and found most people responded to graphics significantly faster than text messages, particularly for elderly drivers. Huang and Bai [17] found the addition of graphic-assisted portable changeable message signs can reduce the speed of vehicles in front of the work zone through driving simulator experiments. The purpose of the traditional warning signs is to remind ETC vehicles to turn left in advance to enter the ETC lane on the left side of the toll station. However, under the background of a high proportion of ETC vehicles, ETC vehicles can enter ETC lanes without changing lanes. Therefore, in order to avoid congestion and collisions in the toll gate plaza, the signs to ETC vehicles should be replaced with those to MTC vehicles that should turn right and change lanes in advance. In order to simplify the sign content and enhance the guidance, the semiotic theory was introduced in this study [18, 19]. The arrow is the most directional graphic symbol in the traffic guidance design [20, 21]. Meng [22] pointed out that arrows have high functionality and good visual recognition capabilities, which makes people receive clear traffic guidance information at a certain distance in a short period of time. Zhang et al. [23] compared the fixation count (FC) and found that the triangular design of the arrow was the most attractive figure for the driver. The larger the FC in the observation area, the more important and prominent the area for the observer. Due to access issues and security concerns with field testing, virtual reality (VR) has become an important means of testing the efficacy of warning signs, which allows for effective verification of sign content and location settings. Previous studies have shown that virtual reality (VR) simulations can be used as an effective tool for studying driving behavior, such as driver visual demand [24], testing proposed positioning of road signs, and testing traffic-control devices [25, 26]. And, in a study, sponsored by the Federal Highway Administration, the American Association of State Highway and Transport Officials, and the National Cooperative Highway Research Program [27], the usefulness of virtual reality (VR) simulation in the road design process was also verified and the use of driving simulators was recommended to be promoted in the road design community.

## 2. Experimental Design

*2.1. Experimental Participants.* A total of 45 participants were recruited for this experiment, and finally, 43 participants completed the experiment for all scenarios, including 33 males and 10 females, aged 18–53 years (AV = 33.7, SD = 8.7), with 1–17 years of driving experience (AV = 4.3, SD = 3.5). Other information of the participants is shown in Table 1 below.

*2.2. Experimental Equipment.* The driving simulators from Changsha University of Science and Technology was adopted to experimentally study the effect of combined warning signs on driving behavior at mainline toll stations. The CSUST simulator is a high-performance, high-fidelity driving simulator with a linear motion base capable of operating with 3 degrees of freedom. It is composed of a full-size vehicle cabin (Ford Focus), environmental noise and shaking simulation system, digital video replay system, and vehicle dynamic simulation system. as shown in Figure 1.

*2.3. Experimental Scenario Design.* The toll station facilities and the exit warning signs were redesigned in this study. Nine ETC toll lanes and four MTC toll lanes were set at the toll station. Luo [28] points out that when the ETC utilization rate is greater than 90%, the capacity of the mixed toll station will increase with the number of ETC lanes and then smoothly, while when the ETC utilization rate is less than 90%, the capacity of the mixed toll station will decrease with the increase of the number of ETC lanes. Therefore, to make the capacity of the mixed toll station optimal, this paper chooses the background in which the proportion of ETC vehicles is above 90% for the study. According to the national standard (GB5768.2–2009), 2 km, 1 km, and 500m exit warning signs and exit warning (action point) signs should be set considering the actual design of the Yuehlin Expressway. Huang et al. [29] pointed out that, within 500m before the exit was an important location for vehicle lane change behavior, so the design of the exit warning signs mainly focused on the content of 500 m and the 1 km, 300 m, and 0 m signs. The content design of the signs should follow the principles of simplicity, guidance, and advance. In this test, a total of three combative mainline toll station warning

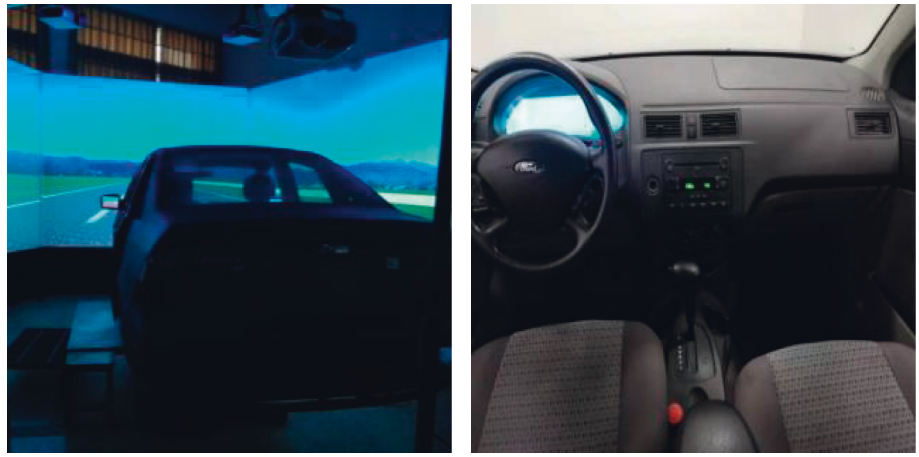


Figure 1: Driving simulator.

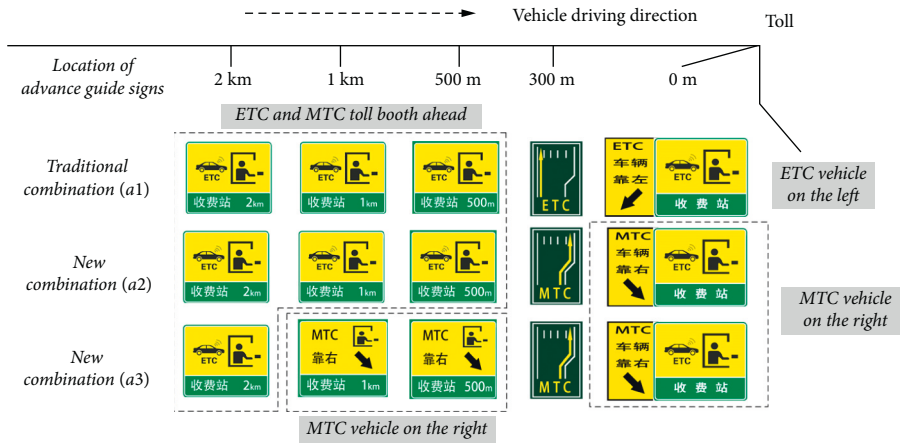



Figure 2: The three sign design schemes.

signs schemes were designed with the control variate method.

A status quo group (a1) was set up to study whether the existing scheme can still meet the traffic demand of a high proportion of ETC vehicles, which is the combination of the traditional signs. Experimental group 2 (a2), based on the status group, replaced the 300 m and 0 m signage ETC guidelines with MTC guidelines and replaced the arrow symbols inside the signage in order to have a better visual recognition function based on semiotic theory. Experimental group 3 (a3), based on experimental group 2 (a2), replaced the ETC guidance in the advance signage at 1 km and 500 m with MTC to the right and added corresponding text guidance, while guiding the text content of MTC to the right through the arrow to the right to enhance the readability of the message in accordance with the principle that arrows are the strongest image symbols in traffic guidance design in semiotic theory. The three sign design schemes are shown in Figure 2.

In this study, the experiment simulated a highway scenario in daytime with sufficient light, and to reduce the influence of extraneous environmental factors on driving behavior, the vehicle was an enclosed space and the interior temperature was 19~24°C, which is the appropriate temperature for humans. Given that there are two types of vehicle users in reality, each driver should complete the experiment as an ETC user and MTC user, respectively. Three experimental scenarios (a1–a3) and six experimental numbers (s1–s6) were set up in this experiment, as shown in Table 2.

### 2.4. Experimental Procedure

(1) Pre-experimental phase: participants adjusted their own seats to a suitable position and familiarized themselves with the simulated vehicle's throttle, brake, and steering, drove the simulated scenario, and checked whether they had any physical discomfort.

(2) Experiment preparation stage: the staff explained the precautions of the experiment for the participants and calibrated the driving simulators

(3) Formal experiment phase: one participant operated the driving simulation software. The whole experiment has three scenarios, the driver took part in the experiment with a different designated identity (ETC

Table 2: Experimental design grouping.

| Experiment no. | Experimental identity | Experimental scheme | Warning sign combination |
|---|---|---|---|
| s5 | ETC | a1 | 2 km, 1 km, 500 m, 300 m (ETC), 0 m |
| s1 | MTC | a1 | 2 km, 1 km, 500 m, 300 m (ETC), 0 m |
| s6 | ETC | a2 | 2 km, 1 km, 500 m, 300 m (MTC), 0 m (new) |
| s2 | MTC | a2 | 2 km, 1 km, 500 m, 300 m (MTC), 0 m (new) |
| s4 | ETC | a3 | 2 km, 1 km (new), 500 m (new), 300m (MTC), 0m (new) |
| s3 | MTC | a3 | 2 km, 1 km (new), 500 m (new), 300 m (MTC), 0 m (new) |

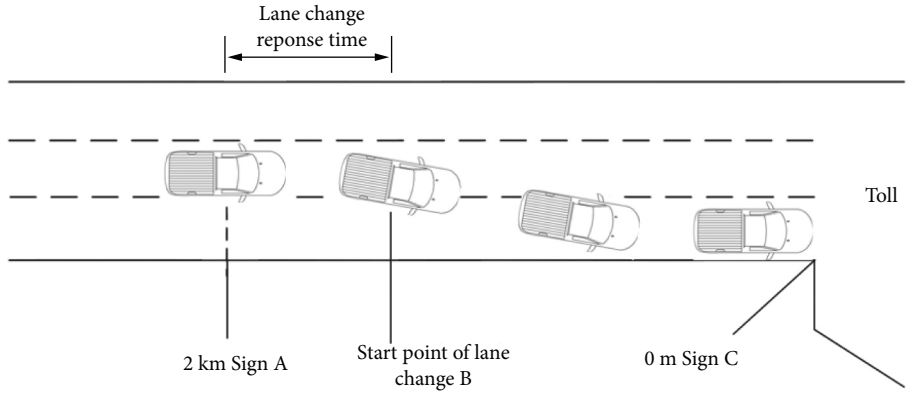



Figure 3: Lane change diagram.

Table 3: Results of descriptive statistical analysis of response time and analysis of variance test.

| Driver | Category | $N$ | Response time | |
|---|---|---|---|---|
| | | | Mean (s) | F value ($P$ value) |
| *Age* | 18~25 | 26 | 45.51 | 0.545 (0.653) |
| | 26~35 | 35 | 49.31 | |
| | 36~50 | 25 | 40.37 | |
| | ≥51 | 10 | 46.15 | |
| *Gender* | Male | 72 | 44.32 | 0.698 (0.406) |
| | Female | 24 | 49.55 | |
| *Driving age* | 1~3 | 18 | 43.56 | 0.066 (0.936) |
| | 4~10 | 58 | 46.13 | |
| | ≥11 | 20 | 46.02 | |
| *Average number of toll booths passing per year* | 0 | 19 | 47.32 | 0.440 (0.725) |
| | 1~5 | 25 | 49.84 | |
| | 6~12 | 21 | 41.43 | |
| | ≥13 | 31 | 44.03 | |
| *Average annual driving mileage (k*km)* | <5 | 17 | 51.33 | 0.897 (0.446) |
| | 5~10 | 19 | 51.41 | |
| | 11~30 | 35 | 42.28 | |
| | ≥31 | 25 | 42.04 | |
| *Combination of warning signs* | 1 | 19 | 43.16 | 5.620 (0.005) |
| | 2 | 38 | 55.95 | |
| | 3 | 39 | 36.77 | |

user or MTC user), with a 20-minute break after each experiment.

(4) At the end of the experiment, the participants filled in the evaluation questionnaire of the effectiveness of the combination of warning signs and the subjective evaluation form of the realism of the driving simulator according to their subjective feelings.

*2.5. Data Collation.* The amount of raw data generated by the driving simulator was very large and contained a large number of nonclosely related variables, so it cannot directly reflect the experimental results, and key variables were extracted from the original simulator data for the analyses. During the driving process from 2500m to 0m from the toll booth, the driver would notice five warning signs, which would lead to different driving behaviors due to decision-

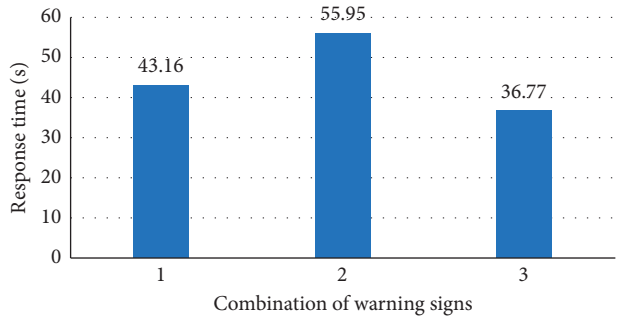



Figure 4: Mean value chart of time-based indicators.

Table 4: Results of descriptive statistical analysis of start position of lane change and analysis of variance test.

| Driver | Category | $N$ | Start position of lane change | |
|---|---|---|---|---|
| | | | Mean (m) | $F$ value ($P$ value |
| *Age* | 18~25 | 26 | 818.71 | 1.452 (0.233) |
| | 26~35 | 35 | 675.18 | |
| | 36~50 | 25 | 708.93 | |
| | ≥51 | 10 | 442.87 | |
| *Gender* | Male | 72 | 712.19 | 0.214 (0.644) |
| | Female | 24 | 658.00 | |
| *Driving age* | 1~3 | 18 | 895.19 | 1.811 (0.169) |
| | 4~10 | 58 | 661.60 | |
| | ≥11 | 20 | 629.17 | |
| *Average number of toll booths passing per year* | 0 | 19 | 780.11 | 1.540 (0.210) |
| | 1~5 | 25 | 549.18 | |
| | 6~12 | 21 | 620.42 | |
| | ≥13 | 31 | 794.21 | |
| *Average annual driving mileage (k*km)* | <5 | 17 | 749.81 | 1.104 (0.352) |
| | 5~10 | 19 | 521.28 | |
| | 11~30 | 35 | 713.26 | |
| | ≥31 | 25 | 778.20 | |
| *Combination of warning signs* | 1 | 19 | 691.33 | 13.142 (0.000) |
| | 2 | 38 | 439.19 | |
| | 3 | 39 | 955.01 | |

making. Humans do not respond to external stimuli directly, with delay and uncertainty. Therefore, participants' motivation to change lanes is a cumulative stimulus response to the warning signs. The differences in driver motivation to change lanes are related to individual differences and scenario factors.

The driver's motivational behavior for lane change was analyzed in the simulation experiments, so the parameters measured included response time, lane change start position, and speed.

The response time refers to the time from the 2 km warning sign to the start point of the lane change, denoted as AB, as shown in Figure 3. The start position of the lane change is the distance from the 0 m sign at the moment the driver starts the lane change, and the speed at the start of the lane change is the instantaneous speed at the start of the lane change.

To analyze whether different levels of a single independent variable (combination of warning signs and driver characteristics) have an effect on the dependent variable (motivation to change lane), the dependent variables were analyzed using one-way ANOVA. The hypothesis testing in the following analyses was based on a significance level of 0.05.

## 3. Result

The experiment recorded 129 complete participants of 43 participants experiencing three combination of warning signs. The vast majority of ETC users would choose to maintain their lanes because there was no lane change motivation, so this study only investigated the driving behavior of lane change participants. For the MTC participants, 96 times of advanced lane change were observed; therefore, the total sample (N) is 96.

*3.1. Response Time.* As shown in Table 3, the response time of lane change was only significantly related to the different combinations of signs ($F$ (2.93) = 5.620, $P$ = 0.005). After LSD postinspection, there was a significant difference between scheme 3 and scheme 1 ($P$ = 0.001), as shown in

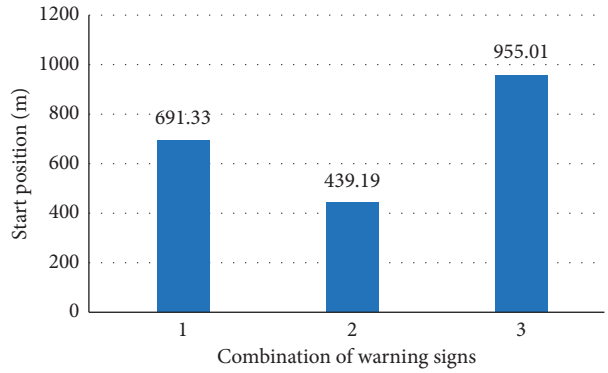



Figure 5: Mean value of start position under different combinations of sign schemes.

Table 5: Results of descriptive statistical analysis of speed at the start of a lane change and analysis of variance test.

| Driver | Category | $N$ | Speed at the start of a lane change | |
|---|---|---|---|---|
| | | | Mean (km/h) | $F$ value ($P$ value) |
| *Age* | 18~25 | 26 | 92.68 | 1.361 (0.260) |
| | 26~35 | 35 | 78.84 | |
| | 36~50 | 25 | 102.68 | |
| | ≥51 | 10 | 85.99 | |
| *Gender* | Male | 72 | 89.18 | 0.017 (0.896) |
| | Female | 24 | 90.62 | |
| *Driving age* | 1~3 | 18 | 101.07 | 0.317 (0.813) |
| | 4~10 | 58 | 82.81 | |
| | ≥11 | 20 | 98.68 | |
| *Average number of toll booths passing per year* | 0 | 19 | 91.03 | 1.540 (0.210) |
| | 1~5 | 25 | 84.77 | |
| | 6~12 | 21 | 85.16 | |
| | ≥13 | 31 | 95.45 | |
| *Average annual driving mileage (k*km)* | <5 | 17 | 96.31 | 0.518 (0.671) |
| | 5~10 | 19 | 88.62 | |
| | 11~30 | 35 | 93.32 | |
| | ≥31 | 25 | 80.36 | |
| *Combination of warning signs* | 1 | 19 | 81.52 | 0.677 (0.511) |
| | 2 | 38 | 87.18 | |
| | 3 | 39 | 95.76 | |

Figure 4, and the response time of lane change in scheme 3 (36.77s) was significantly shorter than that of scheme 1 (43.16 s). This indicated that guiding the MTC vehicles can shorten the response time of MTC vehicles.

*3.2. Start Position of Lane Change.* As shown in Table 4, the initial position of the lane change was only significantly related to the different combination of sign schemes ($F(2.93) = 13.142$, $P < 0.001$), and after LSD postinspection, there was a significant difference between scheme 3 and scheme 1 ($P = 0.035$) and scheme 2 ($P < 0.001$), as shown in Figure 5. The initial lane change in scheme 3 (995.01 m) position was significantly longer than the initial position in scheme 1 (691.33 m) and scheme 2 (439.19 m), indicating that the guidance to MTC vehicles could make vehicles make lane changes earlier.

*3.3. Speed at the Start of a Lane Change.* As shown in Table 5, there was no significant relationship between the speed at the start of the lane change and either driver characteristics or different combination of sign schemes.

## 4. Discussion

Under the background that the proportion of ETC vehicles exceeds 90%, the existing notice signs in front of toll stations reduce the operational efficiency and safety of the toll lanes, cause traffic congestion, and pose a safety hazard.

In this paper, by replacing the warning sign with those for MTC vehicles to change lanes in advance and introducing semiotics theory, the content of the warning sign combination in front of the expressway toll station was redesigned. It can be seen from the experimental results that the combination of different warning signs has different effects on the response time and the starting position of the lane change.

The combination of different warning signs has different effects on the response time and the starting position of the lane change. Among the three schemes, the driver's response

time under scheme 2 is the longest, 29.6% longer than the original scheme. This indicated that scheme 2 has a negative impact on the driver's content design. Comparing schemes 1 and 2, the relevance of signs within 1 km of scheme 2 is poor. There are not only MTC vehicle guidance but also ETC vehicle guidance. According to GB5768.2–2009, expressway road signs should follow the association principle. Because the poor content relevance between the various signs makes it hard for the driver to understand the content of the signs during driving, thereby increasing the response time. However, scheme 3, which satisfies the correlation and order between signs and adds arrow guidance, has the best driver response, 14.8% lower than the original scheme. This demonstrated that setting up reasonable and scientific warning signs can satisfy drivers' visual recognition needs for traffic information on expressways [30].

The combination of different warning signs has different effects on the response time and the starting position of the lane change. The lane change position of scheme 3 is the farthest from the toll plaza, followed by scheme 1. Scheme 2 is the worst. However, the response time of scheme 2 is prolonged, which leads to the driver's lane change starting position being closer to the toll station. When the lane change position is closer to the end of the lane, the driver may adopt more aggressive and dangerous forced lane changes [31, 32]. Compared with schemes 1 and 2, scheme 3 with arrow guidance allows the driver to start the lane change earlier due to the reduction in response time, thus making the initial position of the lane change earlier. This shows that the graphic sign with the arrow guidance can guide drivers to make lane-changing decisions earlier and reduce traffic accidents [14].

Regarding the speed at the beginning of the lane change, different schemes have no significant effect on this. The reason may be that the traffic flow in the experiment was free traffic, so the driver's perception of speed was not obvious. Therefore, there was no significant difference between the speed at the beginning of the lane change under different schemes.

Limitations of the study should be pointed out. Since this experiment was actually carried out under free traffic flow, the driver's perception of speed was not strong. It is recommended to further study the driver's motivation for lane changing under different traffic flows and vehicle types in future experiments, for example, under high, medium, volume traffic, to verify how the toll warning graphic affects the timing of driver lane changes.

## 5. Conclusion

The influence of figures in warning signs at the manual toll station on the lane change timing of drivers in the context of virtual reality of high-proportion ETC vehicles was investigated based on driving simulator. It was found that the addition of arrows to the warning signs can significantly shorten the response time and guide the driver to make lane change decisions earlier to reduce the congestion between MTC vehicles and the mainline ETC vehicles at the toll plaza, thereby improving the traffic capacity and safety.

## Data Availability

All data are received based on driving simulation experiment and did not relate to publicly archived datasets.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References

[1] Y. Feng, "Ergonomics Keeping Abreast of the Times," *Journal of Nanjing University of Technology (social science edition*, vol. 2004, no. 4, pp. 71–75+81, 2004.

[2] L. Vilchez Jose, "Mental representation of traffic signs and their classification: informative signs," *Theoretical Issues in Ergonomics Science*, vol. 22, no. 4, pp. 441–456, 2021.

[3] Z. Li, B. Huang, A. A. Jourdain, C. Yang, C. Y. Su, and A. Bicchi, "Asymmetric bimanual control of dual-arm exoskeletons for human-cooperative manipulations," *IEEE Transactions on Robotics*, vol. 34, no. 1, 2017.

[4] M. Zahabi, P. Machado, C. Pankok et al., "The role of driver age in performance and attention allocation effects of roadway sign count, format and familiarity," *Applied Ergonomics*, vol. 63, pp. 17–30, 2017.

[5] D.-W. Koh, J.-K. Kwon, and S.-G. Lee, "Traffic sign recognition evaluation for senior adults using EEG signals," *Sensors*, vol. 21, p. 4607, 2021.

[6] J. M. Purduski and M. J. Rys, "Evaluations of a new Advance Flagger traffic sign," *International Journal of Industrial Ergonomics*, vol. 24, no. 1, pp. 107–114, 1999.

[7] Z. Li, C. Deng, and K. Zhao, "Human cooperative control of a wearable walking exoskeleton for enhancing climbing stair activities," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 4, 1 page, 2019.

[8] H. Su, W. Qi, C. Yang, A. Aliverti, and E. D. Momi, "Deep Neural Network Approach in Human-like Redundancy Optimization for Anthropomorphic Manipulators," *IEEE Access*, vol. 7, no. 99, p. 1, 2019.

[9] H. Al-Madani and A.-R. Al-Janahi, "Assessment of drivers' comprehension of traffic signs based on their traffic, personal and social characteristics," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 5, no. 1, pp. 63–76, 2002.

[10] Z. Li, B. Huang, Z. Ye, M. Deng, and C. Yang, "Physical human-robot interaction of a robotic exoskeleton by admittance control," *IEEE Transactions on Industrial Electronics*, vol. 65, p. 1, 2018.

[11] H. Su, S. Li, J. Manivannan, L. Bascetta, G. Ferrigno, and E. D. Momi, "Manipulability Optimization Control of a Serial Redundant Robot for Robot-Assisted Minimally Invasive Surgery," in *Proceedings of the 2019 International Conference on Robotics and Automation (ICRA)*, IEEE, Montreal, QC, Canada, May2019.

[12] S. G. Charlton, "Conspicuity, memorability, comprehension, and priming in road hazard warning signs," *Accident Analysis & Prevention*, vol. 38, no. 3, pp. 496–506, 2006.

[13] B. Taylor, E. Chekaluk, and J. Irwin, "Reading the situation: the relationship between dyslexia and situational awareness for road sign information," *Transportation Research Part F: Traffic Psychology and Behaviour*, vol. 36, no. JAN, pp. 6–13, 2016.

[14] B. R. Ullman, N. D. Trout, and C. L. Dudek, *Use of Symbols and Graphics on Dynamic Message Signs. Report No. FHWA/*

*TX-08/0-5256-1*, Texas Transportation Institute, College Station, Texas, 2009.

[15] D. A. Skowronek, *An Investigation of Potential Urban Freeway Guide Sign Problem Locations in Houston, Texas*, Master of Science Thesis, Texas A&M University, College Station, Bryan, Texas, 1990.

[16] J.-H. Wang, S. Hesar, and C. Collyer, "Adding Graphics to Dynamic Message Sign Messages," *Transportation Research Record*, vol. 2018, pp. 63–71, 2007.

[17] Y. Huang and Y. Bai, "Effectiveness of graphic-aided portable changeable message signs in reducing vehicle speeds in highway work zones," *Transportation Research Part C: Emerging Technologies*, vol. 48, pp. 311–321, 2014.

[18] Y. Bai, *The Research of Pictogram Design in the Sign System (Doctoral Dissertation*, Wuhan University of Technology, Hubei, China, 2006.

[19] X. Zhou, W. Qi, S. E. Ovur, L. Zhang, and E. D. Momi, "A novel muscle-computer interface for hand gesture recognition using depth vision," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 1, 2020.

[20] C. Calori and D. Vanden-Eynden, *Signage and Wayfinding Design (A Complete Guide to Creating Environmental Graphic Design Systems*, John Wiley & Sons, Hoboken, New Jersey, U.S, 2015.

[21] H. Su, S. E. Ovur, Z. Xuanyi, W. qi, G. Ferrigno, and E. De Momi, "Depth vision guided hand gesture recognition using electromyographic signals," *Advanced Robotics*, vol. 34, pp. 1–13, 2020.

[22] J. Meng, *A Comparative Study on Signs Design of Public Urban Road Traffic between Chinia and Germany*, Doctoral dissertation, Hunan Normal University, Hunan–Changsha, 2010.

[23] K. Zhang, C. Cui, G. Niu, and G. Jing, "Research on the visual attention and identification feature of the shapes and colors of the safety signs," *Journal of Safety and Environment*, vol. 14, no. 6, pp. 18–22, 2014.

[24] S. Easa and C. Ganguly, "Modeling driver visual demand on complex horizontal alignments," *Journal of Transportation Engineering-asce J TRANSP ENG-ASCE*, p. 131, 2005.

[25] D. McAvoy, K. Schattler, P. Assistant, and T. Datta, "Driving Simulator Validation for Nighttime Construction Work Zone Devices," *Transportation Research Record*, vol. 2015, no. 1, 2007.

[26] D. Noyce and C. Smith, "Driving simulators for evaluation of novel traffic-control devices: protected-permissive left-turn signal display analysis," *Transportation Research Record*, vol. 1844, pp. 25–34, 2003.

[27] K. Keith, M. Trentacoste, L. Depue et al., "Roadway human factors and behavioral safety in Europe," *Report No.: FHWA-PL-05-005*, 2005.

[28] Z. Luo, *Study on the Capacity and ETC Lane Configuration of Expressway Toll Station. Master Dissertation*, East China Jiaotong University, Nanchang, China, 2019.

[29] L. Huang, X. Zhao, Y. Li, J. Ma, and Y. Wang, "Optimal design alternatives of advance guide signs of closely spaced exit ramps on urban expressways," *Accident Analysis & Prevention*, vol. 138, Article ID 105465, 2020.

[30] L. Dai, *Research on the Visual Design of the Freeway Guidance System (Master Dissertation*, Guangxi Norma University, Guilin , Guangxi , China, 2016.

[31] P. G. Gipps, "A model for the structure of lane-changing decision," *Transportation Research Part B: Methodological*, vol. 20, no. 5, pp. 403–414, 1986.

[32] T. Toledo, H. N. Koutsopoulos, and M. E. Ben-Akiva, "Modeling integrated lane-changing behavior," *Transportation Research Record Journal of the Transportation Research Board*, no. 1, p. 1857, 2003.

[33] W. He, *Research on the Influence of Guidance Signs Situation Awareness of Drivers in Highway Interchange*, Master Dissertation, Fuzhou University, Fuzhou, China, 2018.

[34] H. Al-Madani and A. R. Al-Janahi, "Role of drivers' personal characteristics in understanding traffic sign symbols," *Accident Analysis & Prevention*, vol. 34, no. 2, pp. 185–196, 2002a.

[35] Z. Li, C. Xu, W. Qiang, S. Chao, and C. Y. Su, "Human-inspired control of dual-arm exoskeleton robots with force and impedance adaptation," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 50, no. 99, pp. 1–10, 2018b.

[36] D. Kaber, C. Pankok, B. Corbett, W. Ma, J. Hummer, and W. Rasdorf, "Driver behavior in use of guide and logo signs under distraction and complex roadway conditions," *Applied Ergonomics*, vol. 47, pp. 99–106, 2015.

[37] H. Su, W. Qi, Y. Hu et al., "Towards model-free tool dynamic identification and calibration using multi-layer neural network," *Sensors*, vol. 19, no. 17, 2019a.

[38] H. Zwahlen, A. Russ, J. Roth, and T. Schnell, "Viewing ground-mounted diagrammatic guide signs before entrance ramps at night: driver eye scanning behavior," *Transportation Research Record Journal of the Transportation Research Board*, vol. 1843, pp. 61–69, 2003a.

[39] H. Zwahlen, A. Russ, J. Roth, and T. Schnell, "Effectiveness of ground-mounted diagrammatic advance guide signs for freeway entrance ramps," *Transportation Research Record Journal of the Transportation Research Board*, vol. 1843, pp. 70–80, 2003b.

[40] Thirteen General Administration of quality supervision, *Inspection and Quarantine of the People's Republic of China. (2009). National Standards of the People's Republic of China: Road Traffic Signs and Markings. Part 3, Road Traffic Markings*, China Standards Press, Beijing, China, 2009.

Hindawi
Complexity
Volume 2022, Article ID 2992379, 21 pages
https://doi.org/10.1155/2022/2992379

*Research Article*

# Planning the Emergency Collision Avoidance Strategy Based on Personal Zones for Safe Human-Machine Interaction in Smart Cyber-Physical System

**Thanh Phuong Nguyen,[1] Hung Nguyen,[1] and Ha Quang Thinh Ngo [2,3]**

[1]*HUTECH Institute of Engineering, HUTECH University, Ho Chi Minh, Vietnam*
[2]*Department of Mechatronics, Faculty of Mechanical Engineering, Ho Chi Minh City University of Technology (HCMUT), 268 Ly Thuong Kiet Street, District 10, HCMC, 700000, Vietnam*
[3]*Vietnam National University Ho Chi Minh City, Linh Trung Ward, Thu Duc District, Ho Chi Minh, 700000, Vietnam*

Correspondence should be addressed to Ha Quang Thinh Ngo; nhqthinh@hcmut.edu.vn

Received 12 May 2021; Revised 16 February 2022; Accepted 24 February 2022; Published 30 March 2022

Academic Editor: Hang Su



Human contact is a key issue in social interactions for autonomous systems since robots are increasingly appearing everywhere, which has led to a higher risk of conflict. Particularly in the real world, collisions between humans and machines may result in catastrophic accidents or damaged goods. In this paper, a novel stop strategy related to autonomous systems is proposed. This control method can eliminate the vibrations produced by a system's movement by analysing the poles and zeros in the model of autonomous vehicles and goods. Using the pole placement technique, the motion of a system is guaranteed to be more stable, more flexible and smoother. Moreover, several control profiles are employed in the switching mechanism to choose the proper vibration-free effect. The main contributions of this paper are (i) the recommendation of an active stopping planner using different smooth generators from a modelling study, (ii) the validation of their physical characteristics and (iii) the launching of a switching algorithm based on the socially aware navigation framework of a robot. This theoretical work is based on the virtual environment of MATLAB, and the experiment is implemented in the practical platform of an automated guided vehicle. From these results, it can be seen that the proposed approach is robust, effective and feasible for applications in storehouse management, public transportation or factory manufacturing.

## 1. Introduction

Together with the growth of science and technology, autonomous systems have increased considerably in the context of the industrial revolution 4.0. As a result, machines have appeared in the surrounding space of humans. In the manufacturing environment, shared space is necessary for cooperation, as it enhances industrial productivity. Typically, an autonomous robot executes its tasks in the real world. Even in the absence of men, robots should attempt to avoid conflict with other objects. As a matter of fact, the autonomous machine is able to behave more intelligently in real life.

Since social consciousness has become a popular topic, human-oriented studies have attracted not only researchers but also practitioners. All kinds of transportation models are required to complete the steering function among humans. Not all of them know how to respond in the event of a crash. One of the means of public transportations in our society is the bus, which serves many passengers. An extended optimal velocity traffic flow model on two lanes that includes a bus stop and a bus deceleration area was suggested in [1]. Two new traffic states were found such that lane-changing occurred frequently, whether ahead of or behind the bus, and the stop-and-go wave would occur on both lanes instantaneously. With the growing density of vehicles, the lane-changing region around a bus varies. Considering impact due to abnormal deceleration of the current vehicle, the traditional criticality of the working conditions is hardly able

to ensure safety [2]. A new calculation method was established in the event of a sudden reaction to maintain the criticality safety distance. It matches the demands for no injuries on an expressway as well as prevents rear-end collisions. Regarding personal characteristics, the health of a driver has an effect on driving protection. The cognitive dysfunction of a patient [3] on public roads is a subject to study. They are occasionally unable to drive owing to symptoms of executive dysfunction. Each tester was equipped with wearable sensors to analyse his/her driving behaviour. It was found that there exists a difference in deceleration posture between patients and healthy adults at an intersection. This indication can guide medical treatment for individual healthcare. Normally, most transportation on a road is carried out by private cars. As the density of traffic has increased, the risk of harm to people has also increased. To indicate regions of deceleration, the individual road markers notifying drivers of the need for deceleration have been input into a PC-based simulator [4]. The interface between the computer and participants generates a three-dimensional scene presenting visual, auditory and tactile effects. A car user's psychological feelings and visual illusions are fed back to confirm the authenticity of the situation. Regarding a person's hostility, an automatic braking system has been derived with consideration of a pleasant car owner [5]. From graphical inputs, a driver's perceived risk of collision was defined to formulate the braking support system. The findings show the need to output smooth deceleration profiles uniformly with very simple calculations. The action of braking initializes at the time of judgment to reduce speed. More generally, a driver risk evaluation method leans on the analysis of recorded driving data [6]. An estimation covering the acceleration, deceleration and steering actions is observed and scored by a risk consultant. Other types of carriers have also drawn attention from investigators. The common aero vehicle slows down its speed by performing a coning motion without a velocity profile [7]. A one-order approximate model of the velocity and long range is extracted; then, the features of the coning motion target velocity and position precision are obtained by orthogonally allocating the trajectory acceleration and velocity over the load. On the prototype of a train, the construction of an air brake and its nonlinear performance are examined [8]. The deceleration control of the pneumatic brake system is compared with different PID schemes to display its feasibility and accuracy.

## 2. Background Works

Safety is a constant topic of interest in numerous fields. The problem of collision avoidance mainly involves two modes of transportation: two objects moving in either opposite directions or in the same direction. Currently, the methods of ensuring safety could be categorized into passive plans and active plans. In the passive schedule, one object selects a stopping stratagem with respect to the behaviours of another. This rule offers the advantages of a dynamic braking motion, conserves kinetic energy and can afford movement in a short time. However, when faced with several objects, a time delay in communication or braking could result in serious harm. The time restriction becomes stricter because the response may not occur in time. Particularly, when an object is moving in the direction opposite to that of the driver, the human body may be wounded. Active deceleration results in a higher level of protection since one decides to halt independently. If an objector person is shown to prevail in the moving space, the driving speed of the will vehicle robustly slow down. The core of the strategy is how to break the whole system to maintain the current status.

There have been several studies concentrated on the investigation and implementation of human-machine interaction safety system which are summarized in Table 1. Some reports researched the vision-based approach that issued from tracking human's direction and computing the time or distance between operator and machine [9, 30]. The sharing fenceless space where there is no physical barrier is monitored by optical sensors. The major point is that the mentioned cyber-physical system architecture could assess the collision risk on the scale of millisecond. Based on real-time evaluation of safety distance, this leads to fast response time, trigger the safety policies and ensure the coexistence of human-machine. However, the loading status must be known and the relation between response time and image processing should be tested. In some technical investigations, the actively invisible distance from operator to robot is measured by vision-based method [31], laser scanner [10], or both [32]. Most of these studies fuse data from different sources to define the relative position, configuration of upper body or pose. Similar to the distance-based or time-based method, the space separation approach relies on the classification of working space. In the manufacturing environment, the surrounding zones of any object are categorized as three levels, such green zone, yellow zone and red zone [9]. The farthest one, named as green, is quite safe and there is no need of actions. Closing to the centre, it must deliver the alerting message or warning signal to human and robot controller must reduce speed in yellow zone. In emergency area or red zone, the notifications to inform the potential collision and, stop command or moving away command would force robot in order to prevent the collision. In the same idea but in different work, authors [33] presented a real-time safety system by allowing safe human-robot interaction at very low distances. It required to leverage known robot joint angle values and precise estimation of human's position in the working area. There is not much modification in robot platform, for instance, computer vision tool is additionally attached in body of robot and implementation in robot software consents to manipulate.

Following the concept of robot's awareness, developers in [11] represented a collision-free interaction model which is not only control the robot in collision-avoidance path, but also perceive information from human's activities. To support the mentioned requirements, a context-aware collision avoidance interaction system design that consists of sensing module, path planning module and context-aware human pose recognition module was mainly described. In the similar manner, some researchers [12] suggested the

Table 1: The summary of the state-of-the-art research topics.

| Classifications | Author(s) | Methodology | Advantage(s) | Limitation(s) |
|---|---|---|---|---|
| Collision-free approach based on human-in-loop in cyber-physical system (CPS) | Nikolaos N. et al. [9] | The proposed CPS architecture enables collision risk assessment on the scale of millisecond regarding the triggering of safety policies | (i) Low cost and efficient time performance<br>(ii) Method of human detection and enabling reaction signal | (i) Loading status may impact on robot's action<br>(ii) Relationship between response time and image processing is unknown |
| | Mohammad S. et al. [10] | The proposed framework from data fusion between laser scanner and IMU to calculate the human-robot minimum distance on-the-fly is investigated | (i) Accuracy and reliability are manageable for collision avoidance success | (i) A limited number of industrial tasks was carried out and residual vibration from robot's motion which acted on results, was not evaluated |
| Collision-free approach based on environmental context-awareness in CPS | Hongyi L. et al. [11] | The context-aware collision-free system with sensing module for path planning based on human pose recognition is introduced | (i) Detecting potential collision and planning a path to reach target simultaneously<br>(ii) Be able to distinguish human pose and assembly context | (i) Burden computation, an exact human pose recognition algorithm and various assembly sequence are challenges |
| | Yingzhong T. et al. [12] | A universal control system which allows an operator to communicate his motions to the robot manipulator, is proposed on the self-adaption workspace mapping method | (i) No need to find parameters by trial-and-error<br>(ii) It adapts to different operators and robots in workspace | (i) The recognition and classification of human movements were not focused<br>(ii) The effective distance between operator and system was not discussed. |
| | Zong C. et al. [13] | A study for the adaptability of the tracked robot in complex working environment is explored. It includes the mechanical structure, static stability in three terrains for human-robot interaction | (i) Appropriate dimension, good mass distribution and limited velocity is advantageous to maintain the system stability on terrains | (i) Load capacity and financial effectiveness are crucial to investigate |
| | Wang H. et al. [14] | The investigation on improving the performance of RFID robot system by anticollision scheme is exemplified. The tag collisions in the current slot are detected by proposed method, then further resolve each small tag collision to enhance system | (i) This approach is beneficial to boost the moving speed and identification reliability of the RFID robots in complex environments | (i) The solution did not fully consider in the context of collaboration, stability or dynamical control<br>(ii) The phenomenon of weak signal, tags on multi-RFID robots and identification efficiency might occur in practice |
| | Zong C. et al. [15] | The dynamic process of climbing stairs for the tracked mobile robot is analysed by reason of on the novelty of mechanical structure and working principle | (i) The feasibility of design and flexible motion could be gained in unknown environment | (i) The advanced algorithms should be considered to deal with different sizes of stairs |

Table 1: Continued.

| Classifications | Author(s) | Methodology | Advantage(s) | Limitation(s) |
|---|---|---|---|---|
| Collision-free approach based on mathematical computing control in CPS | Andrea M. Z. et al. [16] | An avoidance strategy based on depth camera that suggests the robot alternative paths to be traversed, is achieved both collision free and minimum traversing time | (i) This scheme could predict unintended contacts with operator and maintain the speed of robot as much as possible close to its max value | (i) The traffic conditions should be studied more if ethernet connection is utilized |
| | Mingcong C. et al. [17] | An artificial potential field-based novel model predictive control path replanner is implemented to provide ample space and sufficient time for proper steering/accelerate/brake in such hazardous scenarios | (i) The proposed method is still deemed capable of tracking aggressive collision-free trajectories while maintaining vehicle stability | (i) A model of four-wheel-independently-actuated is not practical |
| | Xiangkun H. et al. [18] | A novel emergency steering control strategy consisting of decision-making layer and motion control layer performs a collision avoidance maneuver | (i) The physical limits of driving actuators and nonlinear factors are included | (i) Human's behaviour was not concerned |
| | Joseph F. et al [17] | A new control structure using model predictive and feedback controller with tire nonlinearities provides firstly collision avoidance ability, then temporarily violate stabilization criteria | (i) All of vehicle's performance capability are utilized to avoid an accident<br>(ii) Prioritizing targets are defined | (i) Be inappropriate with cargo-transportation vehicle |
| Collision-free approach based on human emotion and activity recognition in CPS | Huang X. et al. [19] | Due to the fundamental emotion modulation theory and the neural mechanisms of generating complex motor patterns, a model of emotion generation and modulation to train a recurrent neural network for robot control to perform goal-directed tasks | (i) The emotion-modulating method is able to control manipulator with higher accuracy and faster learning rate | (i) Various faces with many emotions involve the excellent training process |
| | Guo S. et al. [20] | A novel idea for nonlinear multiview Laplacian least squares (MvLL) which construct a global Laplacian weighted graph in order to introduce category discriminant information as well as protect the local neighbourhood information, is proposed. | (i) The effectiveness and robustness of MvLL approach is approved although the tasks of multipose and multifeature are taken | (i) The largest disadvantage of this method is the high time complexity |
| | Zheng E. et al. [21] | The new musculoskeletal-based-method manipulated by the electrical Impedance Tomography (EIT) signals for continuously estimating wrist flexion/extension angles is mentioned in the field of wearable robot | (i) The model-based method performed better with small training data sizes | (i) It is limited in extracting motion information on large data and multi-degree-of-freedom |

Table 1: Continued.

| Classifications | Author(s) | Methodology | Advantage(s) | Limitation(s) |
|---|---|---|---|---|
| Collision-free approach based on skill-learning/skill-transfer method in CPS | Su H. et al. [22] | A novel methodology by integrating the cognitive learning techniques and the developed control techniques, which allow the robot to be highly intelligent to learn senior surgeons' skills and to perform the learned surgical operations in semiautonomous surgery, was discussed | The control algorithm does not only provide the learning ability of the human operation skill from multiple demonstrations on a specific task, but also can transfer the learned motion from open surgery by guaranteeing the robot constraints | More complicated motions entail advanced sensors and burden computation |
| | Su H. et al. [23] | An improved human-robot collaborative control scheme was proposed in teleoperated minimally Invasive surgery scenario | The contributions are not only the improvement of surgical task accuracy and robot constraint, but also the computation efficiency without trajectory planning | Adding one more force sensor for external forces sensing should be conceived and performed with the physical interaction between the surgical tip and the organs |
| | Su H. et al. [24] | The swivel motion reconstruction approach was applied to imitate human-like behaviour using the kinematic mapping in robot redundancy in this research | The algorithm gains the highly online regression prediction for accuracy enhancement and fast computation | It is important to integrate both human upper limb kinematic models and joint limits |
| Collision-free approach based on data fusion for navigation method in CPS | Yang Q. et al. [25] | A novel approach of outdoor localization with 3D-laser scanner is proposed to solve the problem of poor localization accuracy in GPS-denied environments. A path planning strategy based on geometric feature analysis and priority evaluation algorithm is also adopted to ensure the safety and reliability of mobile robot's autonomous navigation and control | The robot has high accuracy of localization without GPS so that the mean error of position is 0.1 m and the mean error of path angle is 6° in the experiment. The function of obstacle avoidance includes static obstacle avoidance, dynamic vehicle avoidance and going through narrow regions | The mechanical structure of mobile platform does not guarantee to overcome any outdoor environment |
| | Sanders D. et al. [26] | The purpose of this investigation is to study the effect on time to complete a task depending on how a human operator interacts with a mobile robot. This kind of interaction utilizes two teleoperated mobile-robot systems, three different ways of interacting with robots and several different environments | It may perform better without a sensor system to assist in more complicated environments. Sometimes, it performs better with a camera mounted on the robot compared with premounted cameras observing the environment | It relies heavily on visual feedback and experienced operators |
| | Chen D. et al. [27] | A control framework was used and consisted of two levels: one was a decision level to give the possible positions of all nodes in sensor networks, while the other was a control level to move along the edge of obstacles such that the problem of obstacle avoidance can be transformed into a coordination problem of multiple robots | The proposed control approach can guide the mobile robot to avoid obstacles and deal with the corresponding dynamical events so as to locate all sensor nodes for an unknown wireless network | If the number of obstacles increases, the average localization errors between the actual locations and the estimated locations significantly increase |

Table 1: Continued.

| Classifications | Author(s) | Methodology | Advantage(s) | Limitation(s) |
|---|---|---|---|---|
| Collision-free approach based on trajectory online adaptation in tele-robot | Luo J. et al. [28] | An innovative method of human motion prediction according to an autoregressive model for teleoperation was motivated. The robot's motion trajectory could be updated in real time through updating the parameters of the proposed model | This approach combines both the prediction of human movement as a feedforward component and a virtual force as feedback | The actively interactive performance of robot as well as the existing trouble in time delay should be considered |
| | Jing L. et al. [29] | A hybrid shared control approach based on EMG signals sensor and artificial potential field is invented to avoid obstacles owing to the repulsive force and attractive force from human perception | It provides an alternative force feedback solution along with muscle activation and human's control of intention and prediction | The robustness and long distance control of the proposed work could be noticed |

universal control system that naturally releases an object manipulation mission from combining the human's gesture to robot manipulator. The advanced functions, for example recognition and classification of shoulder, elbow, wrist or tip of fingers are utilized to control the robot end point. By providing the learning abilities for robot after some iteration processes, a cooperative robot for pick-and-place tasks becomes more intelligent in its workplace. Toward an optimal solution, the research on collaborative robot [16] investigated an avoidance strategy that generated the alternative trajectory to possibly traversed. The superior outcomes are both to prevent collision and to minimize travelling time. Exploiting information receiving from these sensors properly, the robot controller lonely decides motion planning which is the best option in the presence of operator.

Another research topic of safety human-machine interaction is the analysis of dynamic characteristics so as to give a final decision by itself in an emergency case. For these circumstances, a hierarchical model predictive control method was depicted in [17] for four-wheel independent-driving vehicle. An integrated framework that comprises the artificial potential field module for path replanner and feedback compensation control module for path follower, is still capable to track the collision-free trajectories in the ill-conditioned states and supervise the other driving behaviours on the neighbour lane. To deal with uncertain factors in driving actuators and unexpected problems in emergency steering situation, a decision-making layer is implemented into the hierarchical control architecture [18]. The risk indexes associated with collision and destabilisation, are continuously estimated in threat assessment model. An alternative approach has explicitly incorporated vehicle stabilization into path tracking and collision avoidance framework [17]. Sometimes, it permits that the stable motion of vehicle could be violated temporarily if needed to keep away from potential collisions. As a result, this method differs from any others since prioritizing collision avoidance is higher than.

Though, most of studies does not focus on interaction between autonomous grounded vehicle and human while there is a considerable increasing number of automated solutions in e-commerce, logistics and supply chain. In general, when mobile robot or autonomous vehicle leaves research laboratory and join in manufacturing environment, it must be able to deal with emergency circumstances which occur suddenly. Some of them might necessitate manoeuvrings, i.e., emergency collision-free that requires short time response and maintain vehicle stability. The techniques of emergency collision avoidance for autonomous vehicle are demonstrated in [34].

The situation discussed in this paper is that in which a human interacts with a machine, where the robot is equipped with intelligent awareness to evade concurrences. The engine must determine which stop policies are proper while still preserving the balance and keeping the freight undamaged. By investigating the theory of a mass and spring system, the constraints needed to stabilize the whole structure are achieved robustly. Using the actual context of a distribution centre, where workers are obliged to share their workspace, the automated guided vehicle is employed as the platform to prove the proposed approach. Some smooth motion profiles are used in this hardware test to verify the effectiveness and success of our policy.

## 3. Problem Statement

The target hardware in our research [35] is based on the automated vehicle shown in Figure 1, which could track a reference trajectory or move freely. Its mission is to lift up the shelf, bring cargo to its destination and set the shelf down. Depending on the applications, navigation tools (radio frequency identification, laser sensor or magnetic guide sensor) are additionally attached to the platform. It can work autonomously in a warehouse while the operator stays along distance away. Nevertheless, in developing countries or in some scenarios where a man needs, without a stop policy, the availability of a machine presents nearby, which may cause an unexpected accident as in Figure 2.

We assume that $m_1$ and $m_2$ are the mass of the machine hardware and freight, respectively. Once the robot elevates

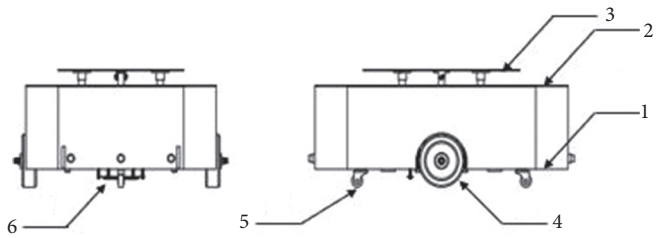



Figure 1: The illustration of our platform, which comprises (1) base, (2) upper cover, (3) lifting part, (4) actively driving wheel, (5) caster wheel, and (6) sensors (line tracking, magnetic tracking and so on).

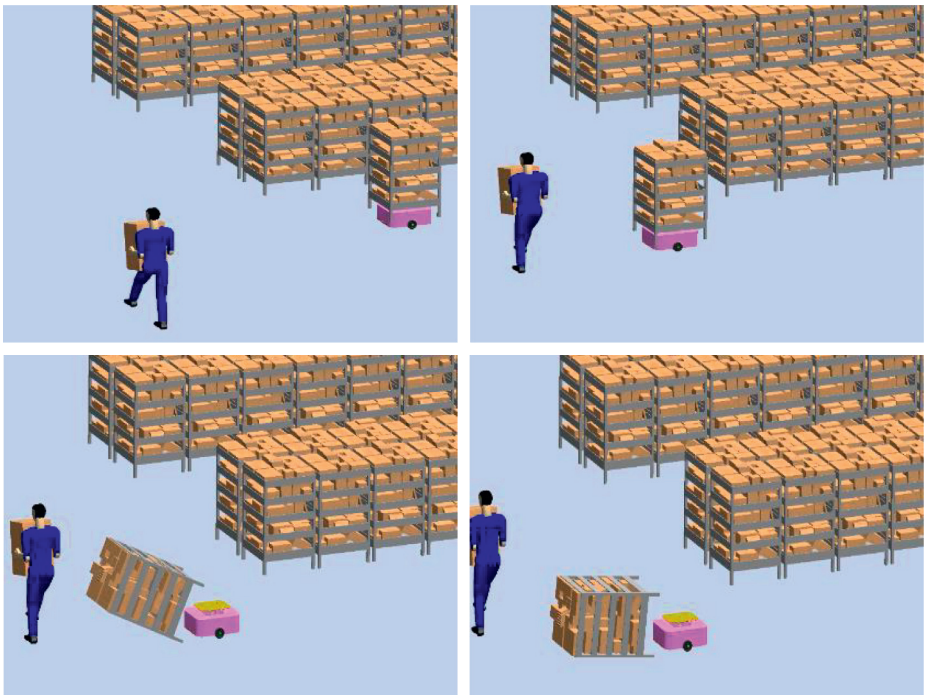


Figure 2: Simulation of physical collision between human and robot sharing the same workspace (see in http://tiny.cc/pawrgz).

the shelf, the two become a rigid body. Without loss of generality, the freight $m_2$ is connected to the frame via two springs with stiffnesses of $k_1$ and $k_2$. The whole modelling procedure is described in Figure 3.

tTo govern the dynamics characteristic, the interacting forces on the system are examined in Figure 4. If the other forces, such as friction forces, exist, they can be ignored. $\vec{\mathbf{v}}$ is the velocity of the vehicle for the same direction of movement. $\vec{\mathbf{F}}_{\mathbf{qt}}$ is the inertial force from the vehicle to the cargo. The state variable $\mathbf{x}$ of the system denotes displacements of each element in vector form: $\mathbf{x} = [\, x_1 \;\; x_2 \,]^T$. Applying Newton's law, the relationship among them can be realized as.

$$m_2 \vec{\mathbf{a}}_2 = k_1 \vec{\mathbf{x}}_2 + k_2 \vec{\mathbf{x}}_2 + \vec{\mathbf{F}}_{\mathbf{qt}}. \tag{1}$$

(1) is rewritten in the positive direction,

$$m_2 a_2 = -k_1 x_2 - k_2 x_2 + m_1 a_1, \tag{2}$$

$$\Leftrightarrow m_2 \ddot{x}_2 = -k_1 x_2 - k_2 x_2 + m_1 \ddot{x}_1. \tag{3}$$

The above equations are related to the time domain. To convert them into the frequency domain, by taking the Laplace transformation, we obtain the following

$$m_2 s^2 X_2 = -k_1 X_2 - k_2 X_2 + m_1 s^2 X_1, \tag{4}$$

where $X_1, X_2$ mean the Laplace transformations of. $x_1, x_2$

By using mathematical manipulation, the related constraints in the proposed model are expressed as

$$X_2 = \frac{m_1 s^2 X_1}{k_1 + k_2 + m_2 s^2}, \tag{5}$$

or

$$\frac{X_2}{X_1} = \frac{s^2}{s^2 + \omega_n^2}, \tag{6}$$

$\omega_n$ is the natural frequency of the damped system.

If the base $X_1$ is utilized precisely, the performance of mass $X_2$ can be measured via (4). In the initial issue, the

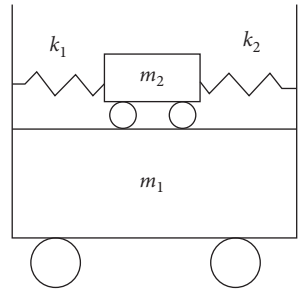



Figure 3: Model of the vehicle and cargo that is assumed in standby status.

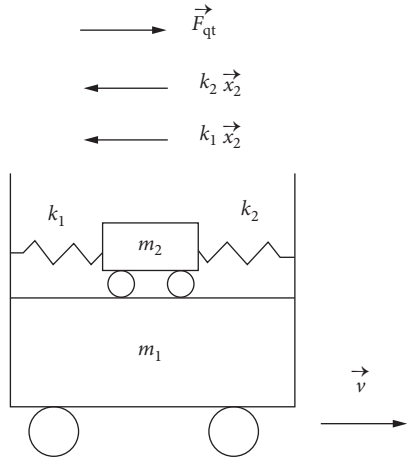



Figure 4: The forces acting on the model are estimated simply in deceleration status.

velocity of $X_1$ is governed by the trapezoidal scheme shown in Figure 5. Three phases occur for this shape: acceleration, constant speed and deceleration. We are focused on the relationship between the deceleration period and stop status.

$v_{\max}, A_{\max}$: maximum value of the velocity $v$ and acceleration $At_a, t_c, t_d$: time period of acceleration, constant speed and deceleration, respectively

$t_0, t_1, t_2, t_3$: time slice at the beginning of acceleration, constant speed, deceleration and end of motion, respectively.

Analysing the above equations, the natural characteristic of the overall system is revealed through poles and zeros which are depicted in the real-imagine coordinate as Figure 6. Since these points locate at the vertical axis, they are usually symmetrical about the horizontal axis.

Poles: $k = 1, 2, 3, \ldots$

$$s_{P_{1,2}} = \pm j\sqrt{\frac{k_1 + k_2}{m_2}} = \pm j\omega_n = \pm j2\pi f_n. \tag{7}$$

Zeros:

$$s_{Z_{1,2}} = \pm j\frac{2\pi k}{t_3 - t_2}. \tag{8}$$

Theoretically speaking, the autonomous system tends to oscillate if the system states are at the poles. In detail, when the conditions in (5) are satisfied, the unexpected vibration might occur and ascend inside this system. As a result, the incident event could cause harm to workers quickly.

## 4. Proposed Stop Strategy

To overcome these challenges, a pole placement for cancelling the unstable points is introduced. The advantages of this approach are that it suppresses the residual vibration, carries out the tuning rule regarding physical factors and the actuator specification, and does not require the exact information of the model. Therefore, free-vibration motion can be achieved in a robust manner.

*4.1. Robust Pole Placement Technique.* Our motivation is to present the novel approach for stable control according to the analysis of system state. Deriving from the idea of pole placement technique, the pole-zero cancelation method is applicable if the zero points are duplications of the locations of the poles. In Figure 7, the outcome of this work is to identify the condition for relocating zero points to positions of poles. By substituting into (5)-(6), the vibration-less conditions are as follows.

$$\sqrt{\frac{k_1 + k_2}{m_2}} = \frac{2\pi k}{t_3 - t_2}, \tag{9}$$

$$\Leftrightarrow T_3 = \frac{2\pi k}{\omega_n}, \tag{10}$$

$$\Leftrightarrow f_n T_3 = k. \tag{11}$$

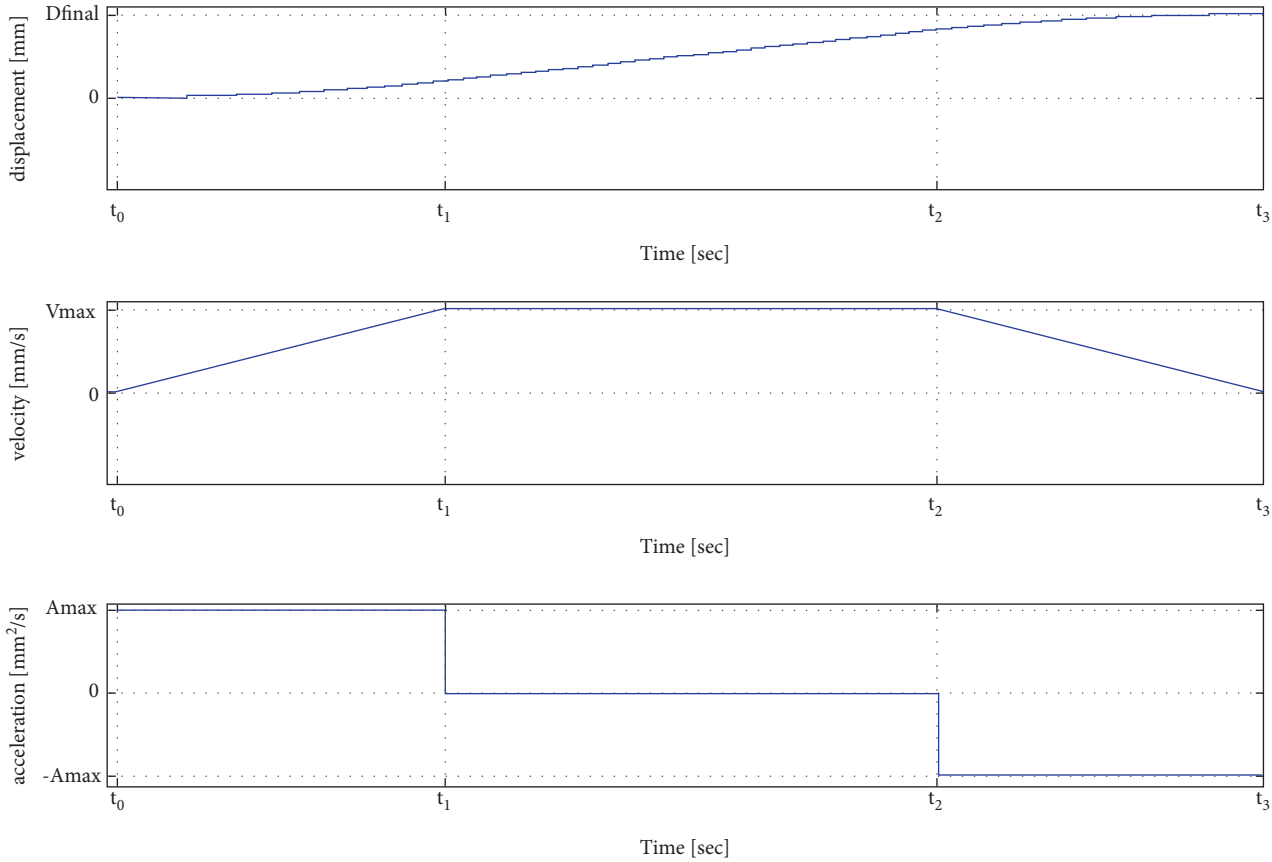



Figure 5: Motion profile of the velocity and acceleration in the trapezoidal strategy.

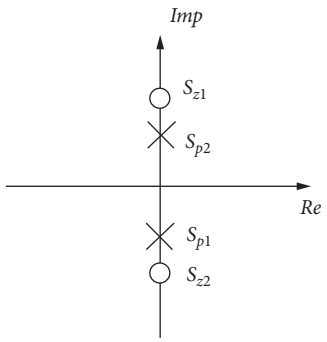



Figure 6: The poles and zeros are placed in the imaginary plane; *o*: zero, *x*: pole.

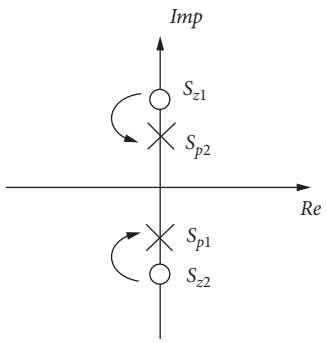



Figure 7: The free-vibration motion by the pole placement technique; *o*: zero, *x*: pole.

$$V(t) = \ddot{x}(t) = \begin{cases} \frac{1}{2}J_{\max}t^2 + V_o, & 0 \le t < t_1, \\ A_{\max}(t - t_1) + V_1, & t_1 \le t < t_2, \\ A_{\max}(t - t_2) - \frac{1}{2}J_{\max}(t - t_2)^2 + V_2, & t_2 \le t < t_3, \\ V_3, & t_3 \le t < t_4, \\ -\frac{1}{2}\frac{J_{\max}}{\gamma^2}(t - t_4)^2 + V_4, & t_4 \le t < t_5, \\ -\frac{A_{\max}}{\gamma}(t - t_5) + V_5, & t_5 \le t < t_6, \\ -\frac{A_{\max}}{\gamma}(t - t_6) + \frac{1}{2}\frac{J_{\max}}{\gamma^2}(t - t_6)^2 + V_6, & t_6 \le t < t_7. \end{cases} \tag{12}$$

In the physical view, when duration $T_3$, from $t_2$ to $t_3$, is the integer multiples of the period of oscillation, the vibration suppression is entirely effective after the vehicle completes its motion. This key constraint guarantees that there is no sudden vibration during the deceleration stage and that the vehicle actively breaks at the specified position. From this idea, various stop profiles are researched to determine which is proper for different cases.

*4.2. Stop Profiles.* The criteria to rate the proper profile are able to suppress the vibration of both the vehicle and cargo; the policy of the generating mechanism depends on the system constraints and ensures the real-time performance. First, the trapezoidal profile is chosen due to its simplicity but effectiveness, and it does not cost much in terms of the timing generation. Later, without loss of generality, the strategy is exerted on numerous profiles for comparison. If the initial velocity $V_0$ is given, then the expression of the asymmetric S-curve speed profile [36] can be manifested as follows:

With

$$V_1 = \frac{1}{2}J_{\max}t_1^2 + V_0, \tag{13}$$

$$V_2 = A_{\max}(t_2 - t_1) + V_1, \tag{14}$$

$$V_3 = A_{\max}(t_3 - t_2) - \frac{1}{2}J_{\max}(t_3 - t_2)^2 + V_2, \tag{15}$$

$$V_4 = V_3, \tag{16}$$

$$V_5 = -\frac{1}{2}\frac{J_{\max}}{\gamma^2}(t_5 - t_4)^2 + V_4, \tag{17}$$

$$V_6 = -\frac{A_{\max}}{\gamma^2}(t_6 - t_5) + V_5, \tag{18}$$

$$V_7 = -\frac{A_{\max}}{\gamma}(t - t_6) + \frac{1}{2}\frac{J_{\max}}{\gamma^2}(t - t_6)^2 + V_6, \tag{19}$$

$$\tau_m = \frac{V_{\max}}{A_{\max}}, \tag{20}$$

$$t_j = \beta\tau_m, \tag{21}$$

$$t_a = (1 - \beta)\tau_m, \tag{22}$$

$$J_{\max} = \frac{A_{\max}}{t_j}. \tag{23}$$

In Figure 8, the asymmetric S-curve motion profile of the mobile base is written with the unit step function $u(t)$. Totally, there are seven segments with $[t_i, t_{i+1}], i \in \{0, 1, 2, 3, 4, 5, 6, 7\}$. $V_{\max}, A_{\max}, J_{\max}$ are maximum values of velocity, acceleration and jerk correspondingly. $V_i$ is the final value of velocity in the time slice $[t_{i-1}, t_i]$. $\tau_m$ is ratio between velocity and acceleration while $t_j$ and $t_a$ are the acceleration/deceleration time and constant acceleration respectively. Thus, equation (12) depicts the time-varying velocity of system. From equations (13) to (19), the components of velocities in each period are identified. The relationship between maximum acceleration and maximum velocity as well as tuning parameters is illustrated from equations (20) to (23).

$$\begin{aligned} s^2X_1 = &-\frac{J_{\max}}{\gamma^2}(t - t_4)[u(t - t_4) - u(t - t_5)] \\ &-\frac{A_{\max}}{\gamma^2}[u(t - t_5) - u(t - t_6)] \\ &+\left(\frac{J_{\max}}{\gamma^2}(t - t_2) - \frac{A_{\max}}{\gamma}\right)[u(t - t_6) - u(t - t_7)]. \end{aligned} \tag{24}$$

Substituting the above into (4), the information of mass $m_2$ can be measured as

$$s^2 X_2 = -\frac{m_1 J_{\max}\left(e^{-t_4 s} - e^{-t_5 s} - e^{-t_6 s} + e^{-t_7 s}\right)}{m_2 \gamma^2 \left(s^2 + \omega_n^2\right)}. \tag{25}$$

In the time domain, (10) is expressed in time-varying form.

$$\begin{aligned}\ddot{x}_2(t) &= -\frac{m_1 J_{\max}}{m_2 \gamma^2 \omega_n} \times \left[\sin\left(\omega_n\left(t - t_4\right)\right) - \sin\left(\omega_n\left(t - t_5\right)\right)\right. \\ &\quad \left. -\sin\left(\omega_n\left(t - t_6\right)\right) + \sin\left(\omega_n\left(t - t_7\right)\right)\right].\end{aligned} \tag{26}$$

It has been indicated that the performance of a vehicle in lifting cargo $m_2$ is discoursed via four sine functions. Their mutual characteristics equalize the frequency but shift the phase. To acquire the desired output, these terms of (11) are forced to dismantle each other. In the case where the control algorithm is valuable, the oscillating performance of $m_2$ approaches zero. The concern in this paper is to develop a policy for stabilizing the movement of a system with different deceleration phases. Again, using the asymmetric S-curve profile, the general conditions needed to eliminate the vibration of cargo are demonstrated.

$$\frac{2\pi k}{\beta \gamma \tau_m} = 2\pi f_n, \tag{27}$$

$$\Leftrightarrow \beta \gamma \tau_m = k T_n, \tag{28}$$

$$\frac{2\pi m}{\gamma \tau_m} = 2\pi f_n, \tag{29}$$

$$\Leftrightarrow \gamma \tau_m = m T_n. \tag{30}$$

From the above results, it is clear that the fluctuation can be destroyed when the motion planning period equals an integer multiplied by the period of natural oscillation. It is desired that at the end of mobile platform movement, the vibration of $m_2$ is completely cancelled as soon as possible. Some simulations with a trapezoid, an S-curve, and an asymmetric S-curve or without a strategy are shown in Figures 9–12, respectively.

$$\delta_{\mathrm{SPB}} = \frac{V_{\max}^2}{2A_{\max}}(1 + \beta), \quad \forall \beta. \tag{31}$$

Generally, it is critical to find the minimum distance for vehicle to generate the smooth stop profile. If $\Delta P_i$ symbolizes the moving distance during a period, later the number of displacements in deceleration duration are expressed as follows:

$$\Delta P_5 = J_{\max} \gamma \left(\beta \tau_m\right)^2 \left(\frac{5}{6}\beta \tau_m + t_a\right), \tag{32}$$

$$\Delta P_6 = \frac{1}{2} J_{\max} \gamma \beta \tau_m t_a \left(\beta \tau_m + t_a\right), \tag{33}$$

$$\Delta P_7 = \frac{1}{6} J_{\max} \gamma \left(\beta \tau_m\right)^3. \tag{34}$$

**MDB** (minimum distance barrier): the cyber limitation ($\delta_{\min}$) in awareness of autonomous system to decelerate safely and smoothly minimizes to be required.

$$\delta_{\mathrm{MDB}} = \frac{A_{\max} \gamma \tau_m^2}{2}(1 + \beta). \tag{35}$$

In dissimilar profiles, the restrictions of minimum distances are different. Each type exists a predetermined range from vehicle to human which avoid to interrupt.

**TPB** (trapezoidal psychological barrier): $\beta = 0$

$$\delta_{\mathrm{TPB}} = \frac{V_{\max}^2 \gamma}{2A_{\max}}, \quad \forall \gamma. \tag{36}$$

**SPB** (S-curve psychological barrier): $\gamma = 1$

**APB** (AS-curve psychological barrier):

$$\delta_{\mathrm{APB}} = \frac{V_{\max}^2 \gamma}{2A_{\max}}(1 + \beta), \quad \forall \beta, \gamma. \tag{37}$$

*4.3. Active Policies with Respect to Safety.* In the paper [37], a unified framework that enables a robot to safely and socially reach both a dynamic human and a human group was developed. Basically, the real-world environment around a human is personal space, where people interact only with close objects such as relatives, and a machine is obligated to inviolate this barrier. In further space, named social space, humans will communicate with others, for example, shaking hands and engaging in oral discussion. A machine should not stay in the social space so that humans feel comfortable and safe. Representing the furthest distance, public space is a cyber interval in which humans come across each other. This is a commonly acceptable gap to preserve a complete stop strategy.

To adapt to the virtual spaces around humans, it is indispensable to launch a switching mechanism among the stop strategies. Given a target distance $\delta$, the mechanism initially ranges from autonomous machines to men.

Here, $H = \{h_1, h_2, \ldots, h_n\}$, where $h_i$ is the ith person. The human states of person $h_i$ are regulated as $h_i = (x_i^h, y_i^h, \theta_i^h, v_i^h)$, where $(x_i^h, y_i^h)$ is the position, $\theta_i^h$ is the orientation with $-180^\circ \le \theta_i^h \le 180^\circ$, and $v_i^h$ is the linear velocity. For the vehicle, its states are denoted as $\vartheta_i = (x_i^\vartheta, y_i^\vartheta, \theta_i^\vartheta, v_i^\vartheta)$, where its position is $(x_i^\vartheta, y_i^\vartheta)$ and its orientation is $\theta_i^\vartheta$, with $-180^o \le \theta_i^\vartheta \le 180^o$.

In Figure 13, Algorithm 1 is utilized whenever an autonomous vehicle carries out a decision before one is made by a human. In this study, the direction a person faces and the location of the right hand or left hand are not our concern. It is noted that the estimation of the relative distance is not the key discussion. We use Kinect digital camera to gain the depth map from the environment, with far distances, and a skeleton option in tracking mode to calculate the current travelling distance. We input the parameters; the values of $\delta_{\mathrm{APB}}, \delta_{\mathrm{SPB}}, \delta_{\mathrm{TPB}}$ are computed in advance. Because of these data, it can be known which stop policy is best to activate.

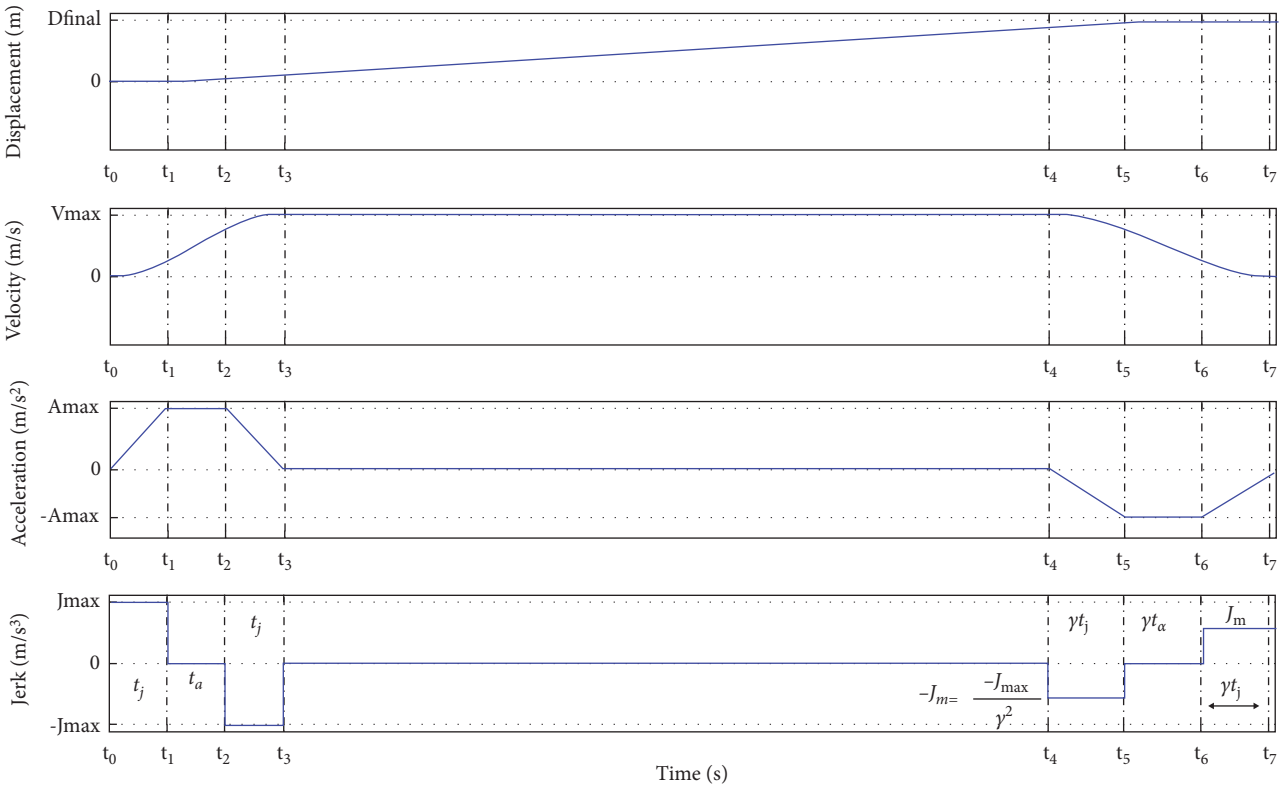



Figure 8: Theoretical motion profile of the asymmetric S-curve.

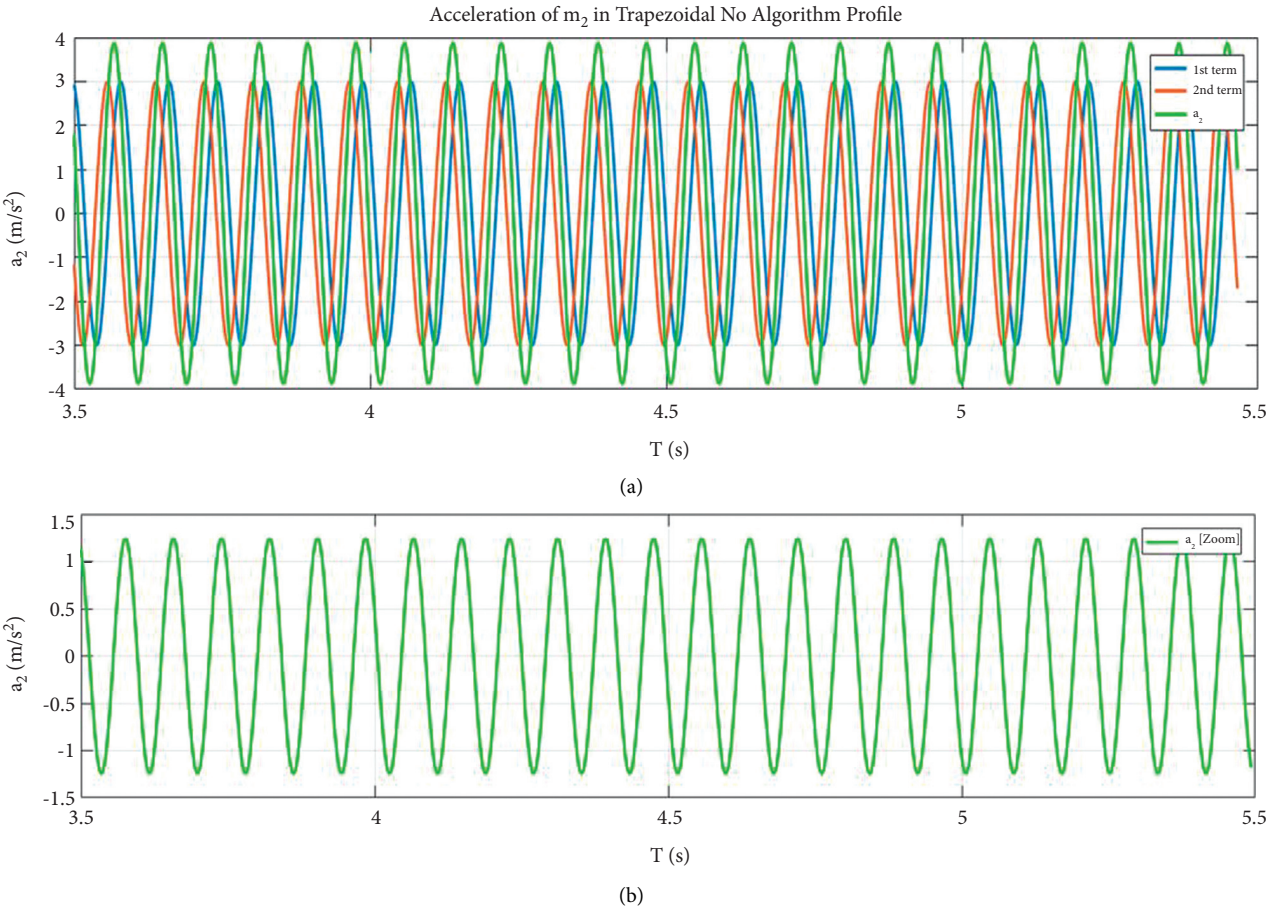



Figure 9: Simulation of the vibration suppression of $m_2$ (a) and its enlarged plot (b) without a stable strategy. Blue: 1st term, orange: 2nd term, green: $m_2$ acceleration.

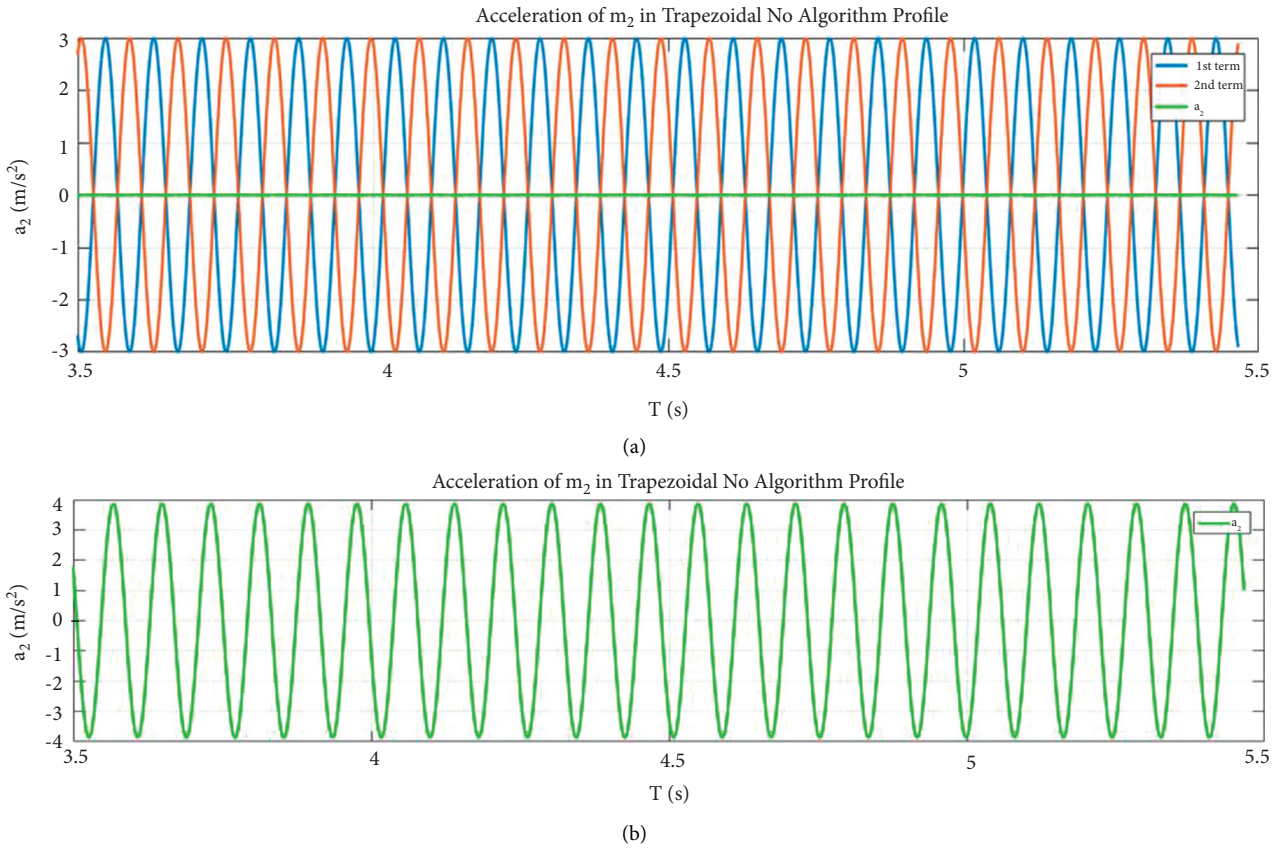



Figure 10: Simulation of the vibration suppression of $m_2$ (a) and its enlarged plot (b) with the trapezoidal strategy. Blue: 1st term, orange: 2nd term, green: $m_2$ acceleration.

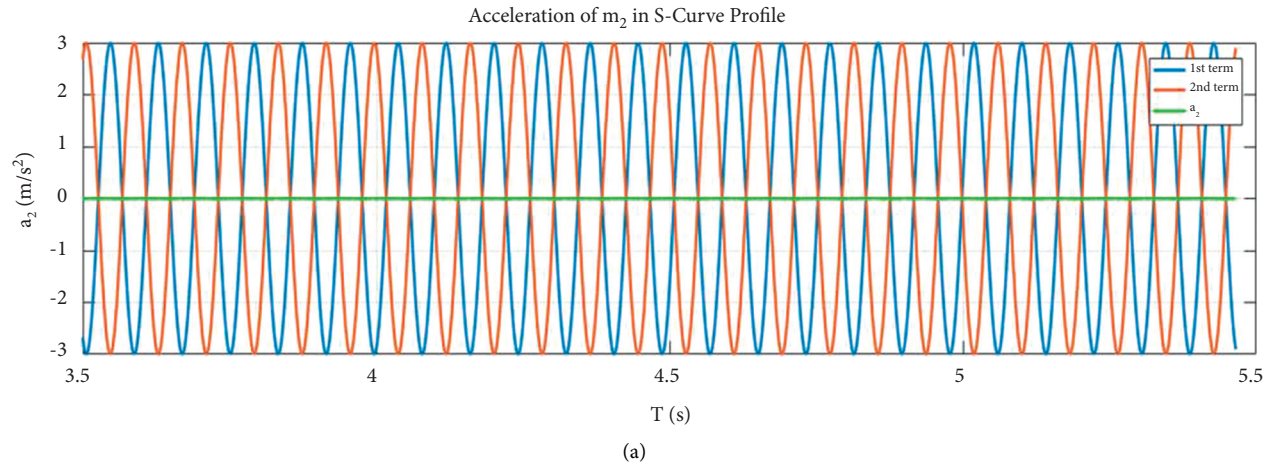
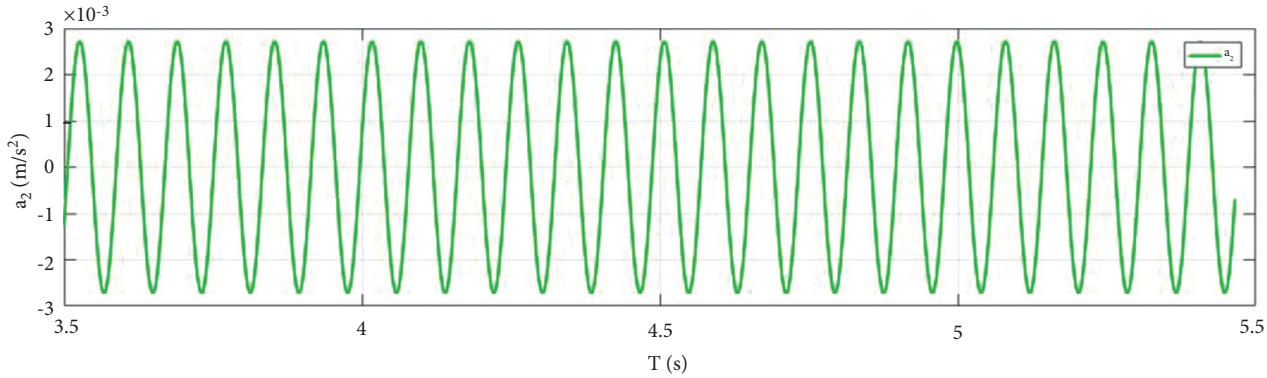
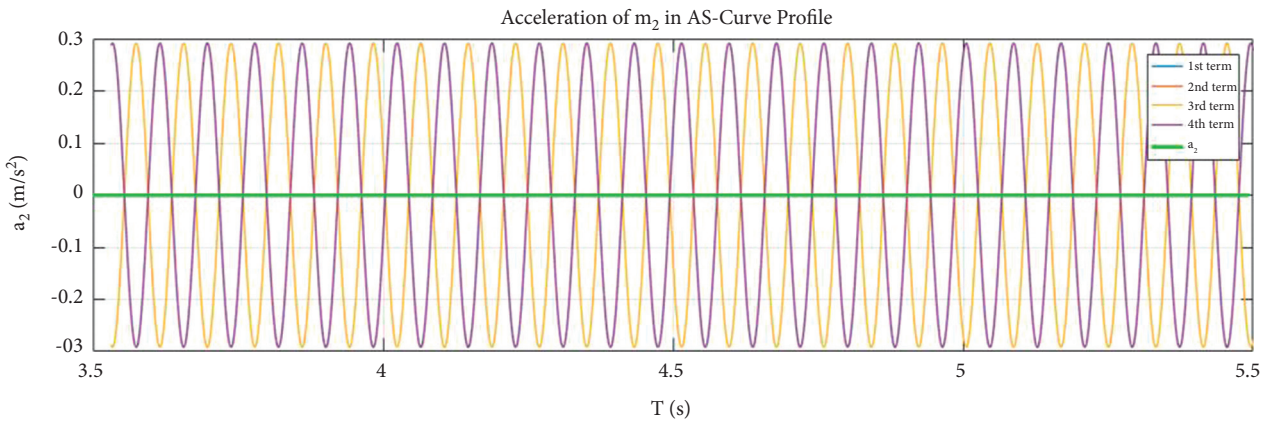



Figure 11: Continued.

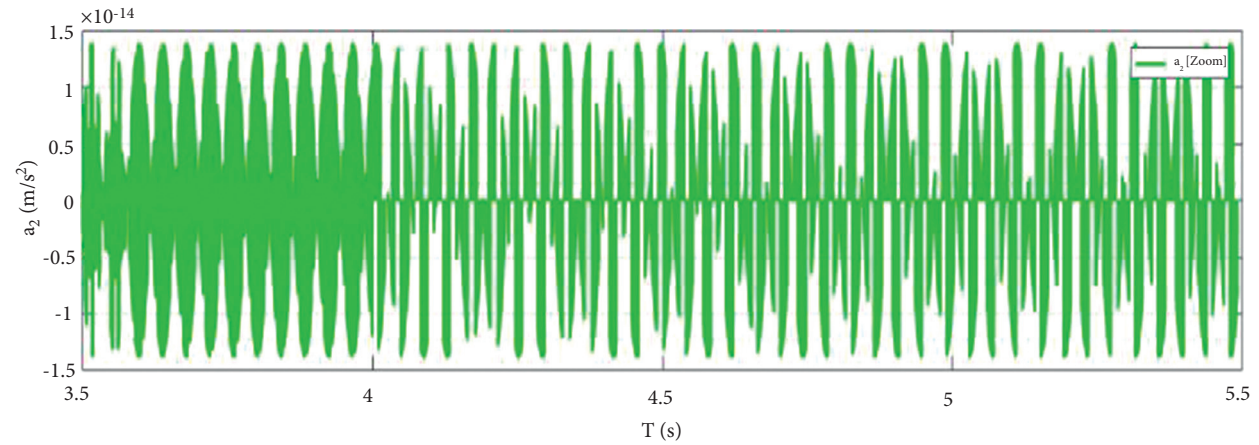



(b)

FIGURE 11: Simulation of vibration suppression of $m_2$ (a) and its zoom plot (b) with S-curve strategy. Blue: 1st term, orange: 2nd term, green: $m_2$ acceleration.




(a)




(b)

FIGURE 12: Simulation of the vibration suppression of $m_2$ (a) and its enlarged plot (b) with the AS-curve strategy. Blue: 1st term, orange: 2nd term, green: $m_2$ acceleration.

The context of human-machine interaction is illustrated as Figure 14. In this case, a scenario of the possibility humans reaching machines in global coordinate $xy$ consists of $v_i^p$ and $v_r$ denoting velocity vectors of $i^{\text{th}}$ person and machine, locations $(x_i^p, y_i^p, \theta_i^p)$ for $i^{\text{th}}$ person and $(x_r, y_r, \theta_r)$ for machine respectively. Besides, the spaces around human are also described to clarify the approaching zone.

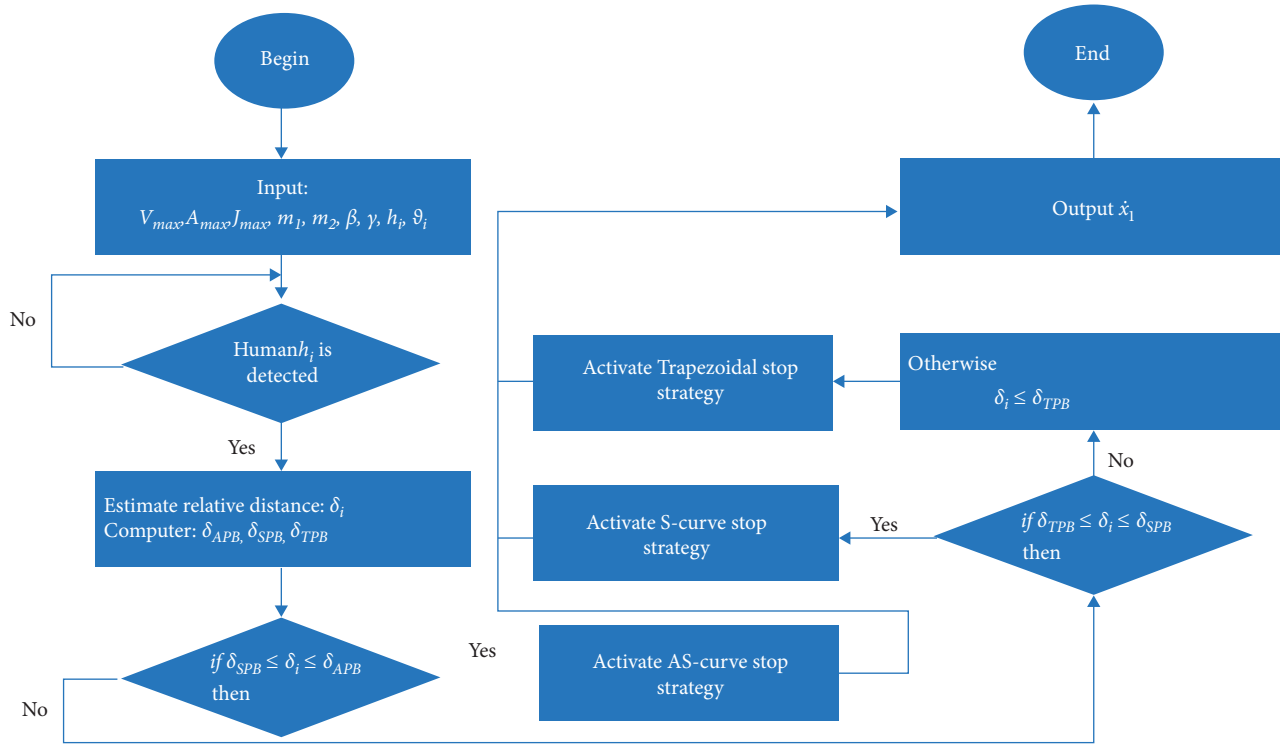



Figure 13: Description of the proposed switching mechanism for stop profile generation.

(i) **Input:** $V_{\max}, A_{\max}, J_{\max}, m_1, m_2, \beta, \gamma, h_i = (x_i^h, y_i^h), \vartheta_i = (x_i^\vartheta, y_i^\vartheta)$
(ii) **Output:** $\dot{x}_1$
(iii) **while** *Human* $h_i$ *is detected* **do**
(1) **Estimate relative distance** $\delta_i$
(2) **Compute** $\delta_{APB}, \delta_{SPB}, \delta_{TPB}$
(3) **if** $\delta_{SPB} \le \delta_i \le \delta_{APB}$
(4) **Activate AS-curve top strategy;**
(5) **else if** $\delta_{TPB} \le \delta_i \le \delta_{SPB}$
(6) **Activate S-curve top strategy;**
(7) **else if** $\delta_i \le \delta_{TPB}$
(8) **Activate Trapezoidal top strategy;**

Algorithm 1: Switching mechanism of stop profile generation.

## 5. Experimental Validation

In this paper, we suggest two classifications of the experiments. In the first laboratory test, the performance of the separated deceleration scheme needs to be measured. The machine is installed with a consciousness regarding the stop behaviour. The vibration results of different strategies are compared to discover the proper profiles for each situation. The second one tests whether to switch among the strategies in real life.

An experimental platform was setup as in Figure 15 to validate our approach. We developed one practical automated guided vehicle to serve in manufacturing factories or distribution centres. To imitate the reaction of cargo, a mass $m_2$ connected to a vehicle by means of a linear aluminium ruler is replaced above the lifting component of the vehicle. A sensing board boostXL Texas Instrument, is placed together with the mass to weigh some physical parameters

along the $x, y, z$ axes. To process data in real-time mode, the development kit LaunchPad MSP432P401 R is utilized to record vibrating outputs. This device supports a powerful 32 bit ARM Cortex-M4F microcontroller up to a 48 MHz system clock, 256 KB of flash memory, and 64 KB of SRAM. The data are then transmitted to the embedded computer to acknowledge the system status. To control two servo motors, the TM4C evaluation board and DCS3T-27 driver are used to close the feedback loop. Tiva C series TM4C123 G LaunchPad including the TM4C123GH6PMI microcontroller with two motion control PWM modules capable of generating 16 PWM outputs, two quadrature encoder interfaces, and several driving platforms.

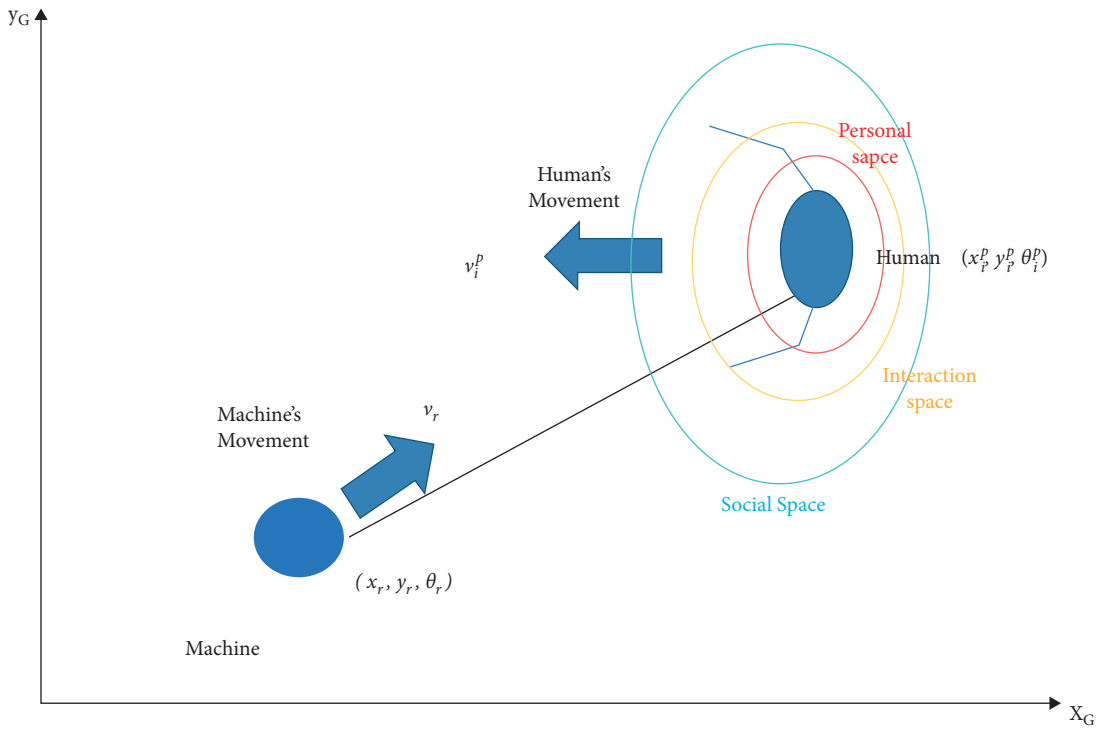



Figure 14: Relative constraints between a human and machine.

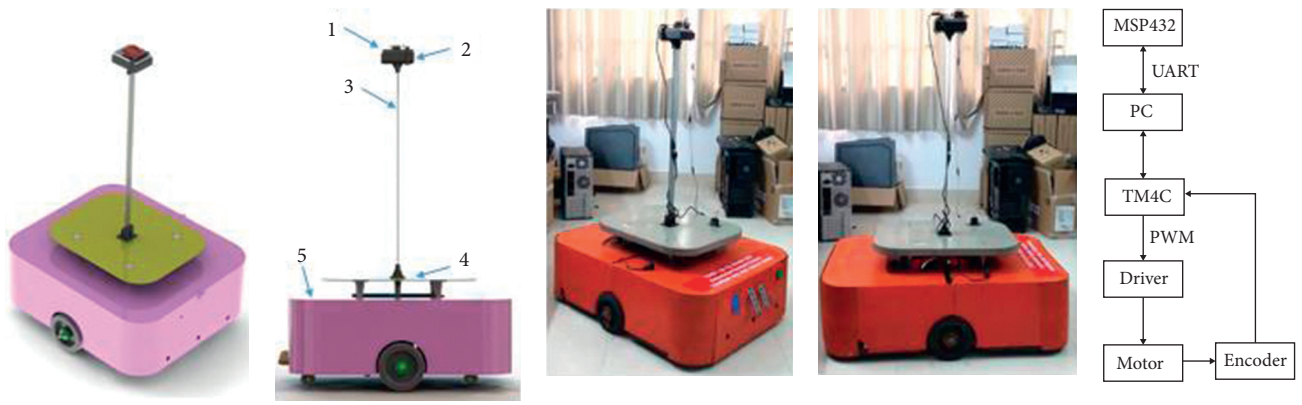
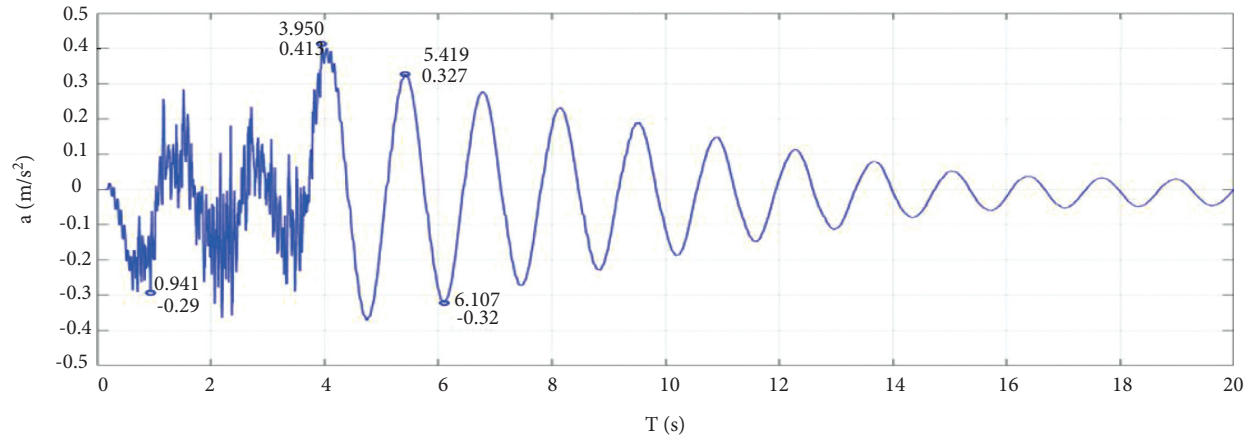
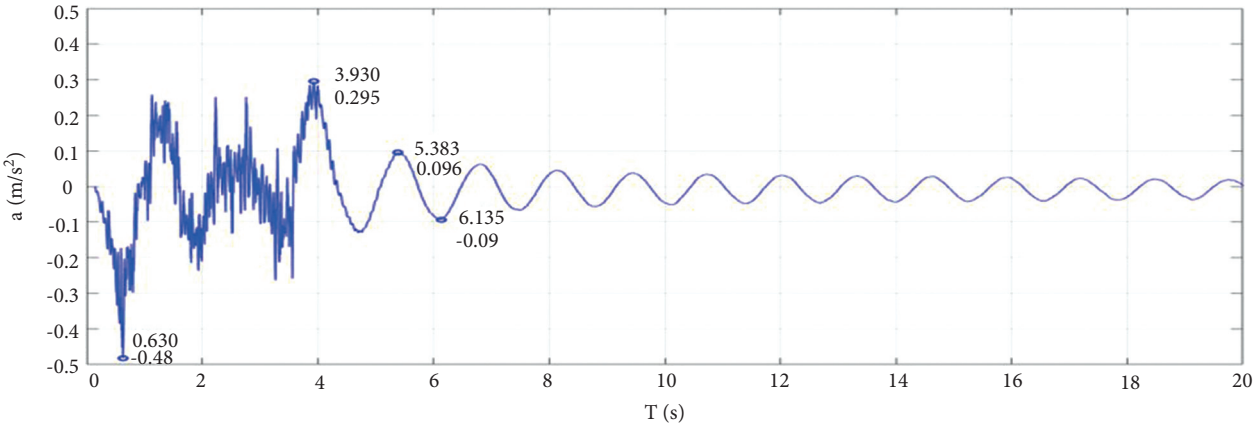



Figure 15: Experimental setup test: (a) virtual 3D view, (b) virtual side view: board sensor[1], $m_2$ = mass[2], linear ruler[3], lifting component of the vehicle[4], and practical automated guided vehicle[5], (c) real 3D view, (d) real side view and (e) flowchart of the data transmission (see in http://tiny.cc/eewrgz).

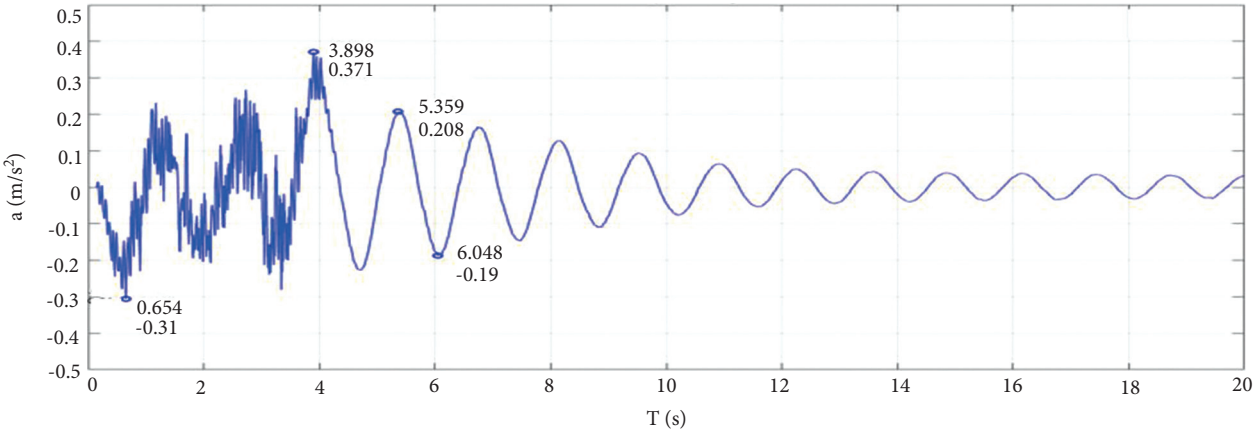
In the first validation test, our target is to categorize the performance of these three policies and that without a policy. The separated results in each case should be shown in Figure 16. In the initial stage, because of a sudden break while travelling, the largest swinging wave appears indicating unbalanced movement. Without any policy, the amplitude of the vibration reaches its maximum height after a moment. It takes more time to decline the output variation. Applying the TPB scheme, the result becomes better in terms of a small swing and fast response to suppress vibration. For the effect of the SPB algorithm, the achieved performance is of smaller magnitude for oscillation and suppresses it more quickly. Lastly, the most superior performance of the smallest level of vacillation and the fastest response are achieved by APB scheme while the powerful computation of micro-processor is required. To visualize the differences of their impact, a combined result representing a comparison of the proposed strategies is displayed as in Figure 17. The deep impression on the diminishing vibration of the APB method has been observed. However, its constraints require a machine far enough away to decelerate. On the other hand, the schemes of SPB and TPB satisfy the conditions of vibration suppression while not need more spaces to execute.




(a)




(b)




(c)

Figure 16: Continued.

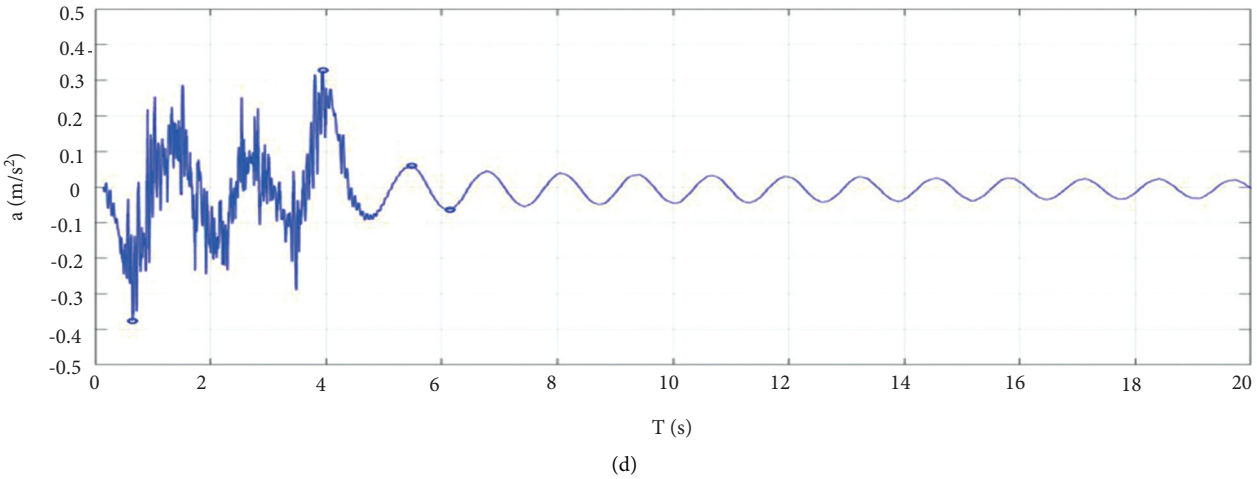



Figure 16: Experimental results of different test scenarios: (a) without a strategy, (b) with SPB, (c) with TPB, and (d) with APB.

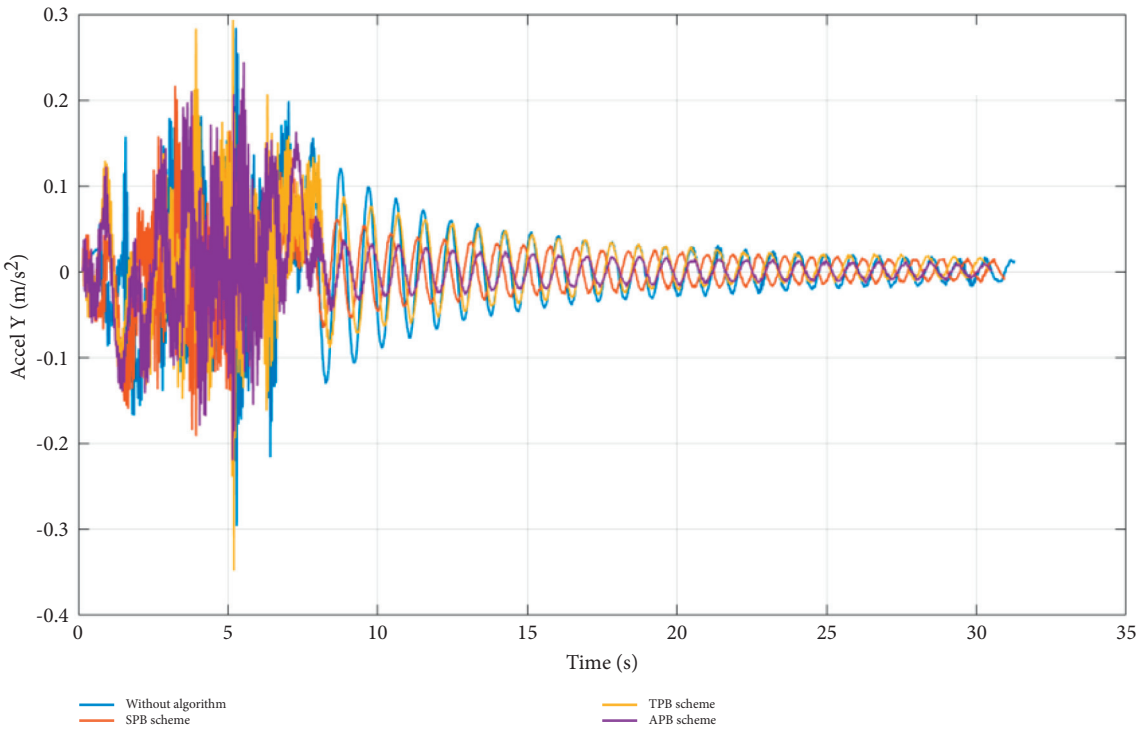



Figure 17: Comparison of test cases with various proposed strategies.

Table 2 reveals the comparative performance for all laboratories. This research regards how fast the oscillation is terminated and how many times the largeness of the oscillation decreases. When the proposed control techniques are not considered, the acceleration of mass $m_2$ tends to decrease the vibrating magnitude by 1.26 times after 1.47 s. In the primary scheme, TPB helps to lessen the number of times to 1.78 after 1.46 s from the moment of the unexpected stop. In other tests, under the power of the SPB and APB schemes, the vibration declines by 3.07 times and 5.56 times after 1.41 s and 1.27 s, correspondingly. It is concluded that the APB technique can offer the best performance in terms of antioscillation even though it is expected that more space will be required to complete the stop strategy. The others show that a smaller oscillating range and faster response of the system can be attained, and the travelling distance desired by a mobile vehicle can be increased.

Equations (28) and (30) are named E28 and E30, respectively. An autonomous vehicle can be driven with a trapezoidal, an S-curve or an AS-curve speed profile. Depending on the type, the constraints required to address the vibration-free case are dissimilar. For the TPB case, only E30 is able to suppress the oscillation, although it does not violate E28. The APB and SPB schemes are both able to

Table 2: Comparative results of experimental laboratories.

| Acceleration of second mass | Min | Max |
|---|---|---|
| Without proposed algorithm | −0.323 | −0.327 |
| TPB | −0.187 | 0.208 |
| SPB | −0.09 | 0.096 |
| APB | −0.064 | 0.059 |

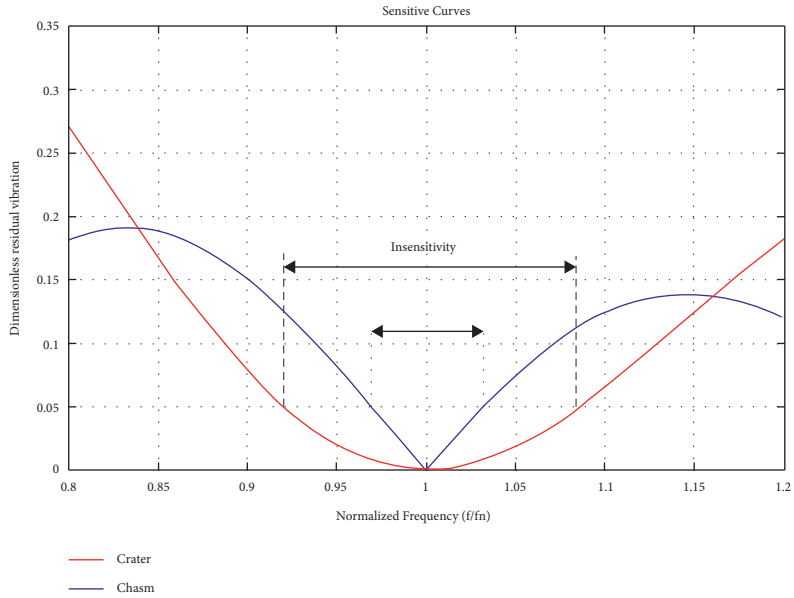



Figure 18: Sensitivity curves of variation in the vibration magnitude against the natural frequency error with $f_n\tau_m = 1.2229$.

operate under the E28 and E30 conditions. After several experiments, it is known that the tolerance of the vibrating frequency in TPB is smaller than that in APB or SPB. The sensitive curve in Figure 18 denotes the variation in amplitude when the estimated natural frequency differs from the practical natural frequency. In this study, it is assumed that the natural frequency error is ±20% of the real value. The motion that satisfies both E28 and E30 is defined as a crater. If it satisfies only one of the above vibration-less conditions, the movement is categorized as a chasm. Notably, the impacts needed to reduce the residual vibration of a crater are larger than those of a chasm.

In general, if the tuning parameters satisfy both E28 and E30, then the residual vibration amplitude will decline in comparison with only E28 or E30 whenever the error modelling is small enough.

## 6. Conclusions

An active stop strategy with vibration cancelation control and collision avoidance between humans and machines (i.e., material transportation) was proposed in this paper. The antivibration control constraint is analytically well-defined based on a dynamic modelling of a vehicle and cargo. Various profiles for activating the stop policy were inspected, and the specific characteristics of each profile were analysed. In particular, the physical characteristics of TPB, SPB and APB were analytically compared to the case without an active stop planner. From the points of view of smoothness and robustness, both investigated strategies drive the whole platform to halt under flexible motion while the system is constantly stabilized. Owing to the socially aware navigation framework of a mobile platform in dynamic surroundings, we developed a switching mechanism to select the proper profile for different social situations. We certified the effectiveness of the proposed approach through both simulation work and a real-world laboratory test. Clearly, the newly investigated stop method could be incorporated into any path planning technique to enhance the performance of human-awareness of an autonomous system. It is clear that our approach has undergone successful integration such that the robot behaviour is improved to a socially acceptable response, achieving safe restrictions while conserving system parameter constraints.

Compared to the works invented by authors [23], although the interaction force is not mentioned and the natural frequency seems to be difficult to estimate, our study focusses on stable control for the autonomous vehicle. Compared to our previous topic [35], not only the potential human-machine interaction is examined, but also the internal system state is mainly concerned. Future work is required. We will investigate the proposed approach in different scenarios of social interaction (i.e., in front of group discussion) and dynamic environments. Additionally, various types of intelligent control should be considered to enhance the decision making for autonomous system.

Furthermore, it is raised a promising research issue that the constrained parameter set could be optimized by using machine learning techniques.

Note to practitioners. In general, the robot trajectory is scheduled and planned in host computer and human could wear the tracking device to avoid collision. However, in some emergent cases, robot must be clearly aware the human appearance and actively prevent collision. Thus, the proposed method achieves the purpose of collision avoidance and socially interactive action. Additionally, the switching mechanism to decide proper strategy based on the distancing communication between human and robot is suggested in this work. It is potentially utilized in logistic warehouse, service restaurant, and so on.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors acknowledge the support of time and facilities from Ho Chi Minh City University of Technology (HCMUT), VNU-HCM, for this study.

## References

[1] W. Shi, H. Kuang, L.-Y. Dong, and S.-Q. Dai, "Characteristics of lane changing induced by bus stop and deceleration area," in *Proceedings of the Fourth International Joint Conference on Computational Sciences and Optimization*, pp. 1064–1068, Yunnan, China, April 2011.

[2] Y. Wu, J. Xie, L. Duand, and Z. Hou, "Analysis on traffic safety distance of considering the deceleration of the current vehicle," in *Proceedings of the Second International Conference on Intelligent Computation Technology and Automation*, pp. 491–494, Hunan, China, June 2009.

[3] T. Yamamoto, M. Nakamura, and A. U. T. Toriyama, "A study of deceleration behavior for cognitive dysfunction drivers on public road," in *Proceedings of the International Conference on Intelligent Informatics and Biomedical Sciences*, pp. 376-377, Okinawa, Japan, 2015.

[4] Y. Ying and Z. Yuhui, "Virtual simulation experiment analysis of Chevron deceleration marking based on driving simulator," in *Proceedings of the Fourth International Conference on Intelligent Computation Technology and Automation*, pp. 991–994, Guangdong, China, 2011.

[5] T. Wada, S. Doi, N. Tsuru, K. Isaji, and H. Kaneko, "Formulation of deceleration behavior of an expert driver for automatic braking system," in *Proceedings of the International Conference on Control, Automation and Systems*, pp. 2908–2912, Seoul, Republic of Korea, 2008.

[6] C. Miyajima, H. Ukai, A. Naito, H. Amata, N. Kitaoka, and K. Takeda, "Driver risk evaluation based on acceleration, deceleration, and steering behavior," in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 1829–1832, Prague, Czech Republic, 2011.

[7] D. Ji, B. Wu, C.-W. Min, Z. Guo, Z. Xiao, and F.-H. He, "Deceleration control by the coning motion for CAV terminal guidance without velocity profile," in *Proceedings of the 35th Chinese Control Conference*, pp. 5747–5751, Chengdu, China, 2016.

[8] Z. Jianyong and L. Yuechao, "Simulation on pneumatic brake control of train based on deceleration feedback," in *Proceedings of the Fourth International Conference on Intelligent Computation Technology and Automation*, pp. 928–931, Shenzhen, China, March 2011.

[9] N. Nikolakis, V. Maratos, and S. Makris, "A cyber physical system (CPS) approach for safe human-robot collaboration in a shared workplace," *Robotics and Computer-Integrated Manufacturing*, vol. 56, pp. 233–243, 2019.

[10] M. Safeea and P. Neto, "Minimum distance calculation using laser scanner and IMUs for safe human-robot interaction," *Robotics and Computer-Integrated Manufacturing*, vol. 58, pp. 33–42, 2019.

[11] H. Liu and L. Wang, "Collision-free human-robot collaboration based on context awareness," *Robotics and Computer-Integrated Manufacturing*, vol. 67, p. 101997.

[12] Y. Tian, G. Wang, L. Li, T. Jin, F. Xi, and G. Yuan, "A universal self-adaption workspace mapping method for human-robot interaction using Kinect sensor data," *IEEE Sensors Journal*, vol. 20, 2020.

[13] C. Zong, Z. Ji, J. Yu, and H. Yu, "An angle-changeable tracked robot with human-robot interaction in unstructured environments," *Assembly Automation*, vol. 40, no. 4, pp. 565–575, 2020.

[14] H. Wang, S. Wang, J. Yao et al., "Effective anti-collision algorithms for RFID robots system," *Assembly Automation*, vol. 40, 2019.

[15] C. Zong, Z. Ji, and H. Yu, "Dynamic stability analysis of a tracked mobile robot based on human–robot interaction," *Assembly Automation*, vol. 40, 2019.

[16] A. M. Zanchettin, P. Rocco, S. Chiappa, and R. Rossi, "Towards an optimal avoidance strategy for collaborative robots," *Robotics and Computer-Integrated Manufacturing*, vol. 59, pp. 47–55, 2019.

[17] J. Funke, M. Brown, S. M. Erlien, and J. C. Gerdes, "Collision avoidance and stabilization for autonomous vehicles in emergency scenarios," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 4, pp. 1204–1216, 2016.

[18] X. He, Y. Liu, C. Lv, X. Ji, and Y. Liu, "Emergency steering control of autonomous vehicle for collision avoidance and stabilisation," *Vehicle System Dynamics*, vol. 57, no. 8, pp. 1163–1187, 2019.

[19] X. Huang, W. Wu, H. Qiao, and Y. Ji, "Brain-inspired motion learning in recurrent neural network with emotion modulation," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 4, pp. 1153–1164, 2018.

[20] S. Guo, L. Feng, Z.-B. Feng et al., "Multi-view laplacian least squares for human emotion recognition," *Neurocomputing*, vol. 370, pp. 78–87, 2019.

[21] E. Zheng, J. Wan, L. Yang, Q. Wang, and H. Qiao, "Wrist angle estimation with a musculoskeletal model driven by electrical impedance tomography signals," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 2186–2193, 2021.

[22] H. Su, A. Mariani, S. E. Ovur, A. Menciassi, G. Ferrigno, and E. De Momi, "Toward teaching by demonstration for robot-assisted minimally invasive surgery," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 484–494, 2021.

[23] H. Su, C. Yang, G. Ferrigno, and E. De Momi, "Improved human-robot collaborative control of redundant robot for t minimally invasive surgery," *IEEE Robotics and Automation Letters*, vol. 4, no. 2, pp. 1447–1453, 2019.

[24] H. Su, W. Qi, Y. Hu, H. R. Karimi, G. Ferrigno, and E. De Momi, "An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 3, pp. 1864–1872, 2020.

[25] Q. Yang, D. Qu, F. Xu, F. Zou, G. He, and M. Sun, "Mobile robot motion control and autonomous navigation in GPS-denied outdoor environments using 3D laser scanning," *Assembly Automation*, vol. 39, 2019.

[26] D. Sanders, "Comparing speed to complete progressively more difficult mobile robot paths between human tele-operators and humans with sensor-systems to assist," *Assembly Automation*, vol. 29, 2009.

[27] D. Chen, Q. Lu, D. Peng, K. Yin, C. Zhong, and T. Shi, "Receding horizon control of mobile robots for locating unknown wireless sensor networks," *Assembly Automation*, vol. 39, 2018.

[28] J. Luo, D. Huang, Y. Li, and C. Yang, "Trajectory online adaption based on human motion prediction for teleoperation," *IEEE Transactions on Automation Science and Engineering*, 2021.

[29] J. Luo, Z. Lin, Y. Li, and C. Yang, "A teleoperation framework for mobile robots based on shared control," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 377–384, 2019.

[30] C. Morato, K. N. Kaipa, B. Zhao, and S. K. Gupta, "Toward safe human robot collaboration by using multiple kinects based real-time human tracking," *Journal of Computing and Information Science in Engineering*, vol. 14, no. 1, 2014.

[31] A. Mohammed, B. Schmidt, and L. Wang, "Active collision avoidance for human-robot collaboration driven by vision sensors," *International Journal of Computer Integrated Manufacturing*, vol. 30, no. 9, pp. 970–980, 2017.

[32] S. Kumar, D. Gupta, and S. Yadav, "Sensor fusion of laser & stereo vision camera for depth estimation and obstacle avoidance," *International Journal of Computer Application*, vol. 1, no. 26, pp. 22–27, 2010.

[33] P. A. Lasota, G. F. Rossano, and J. A. Shah, "Toward Safe Close-Proximity Human-Robot Interaction with Standard Industrial Robots," in *Proceedings of the IEEE International Conference on Automation Science and Engineering (CASE)*, pp. 339–344, Taipei, Taiwan, August 2014.

[34] M. Galvani, F. Biral, B. M. Nguyen, and H. Fujimoto, "Four wheel optimal autonomous steering for improving safety in emergency collision avoidance manoeuvres," in *Proceedings of the IEEE 13th International Workshop on Advanced Motion Control (AMC)*, pp. 362–367, IEEE, Yokohama, Japan, March 2014.

[35] H. Q. T. Ngo, V. N. Le, V. D. N. Thien, T. P. Nguyen, and H. Nguyen, "Develop the socially human-aware navigation system using dynamic window approach and optimize cost function for autonomous medical robot," *Advances in Mechanical Engineering*, vol. 12, no. 12, 2020.

[36] K. H. Rew and K. S. Kim, "A closed-form solution to asymmetric motion profile allowing acceleration manipulation," *IEEE Transactions on Industrial Electronics*, vol. 57, no. 7, pp. 2499–2506, 2010.

[37] X.-T. Truong and T.-D. Ngo, ""To approach humans?": a unified framework for approaching pose prediction and socially aware robot navigation," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 10, no. 3, pp. 557–572, 2018.

Hindawi
Complexity
Volume 2022, Article ID 5648826, 10 pages
https://doi.org/10.1155/2022/5648826

*Research Article*

# A Robot Human-Like Learning Framework Applied to Unknown Environment Interaction

**Xianfa Xue,[1] Lei Zuo ,[2] and Ning Wang [3]**

[1]*Key Laboratory of Autonomous Systems and Networked Control, School of Automation Science and Engineering, South China University of Technology, Guangzhou 510641, China*
[2]*School of Electronic and Control Engineering, Chang'an University, Xi'an 710064, China*
[3]*Bristol Robotics Laboratory, University of the West of England, Frenchay, Coldharbour Ln, Bristol BS34 8QZ, UK*

Correspondence should be addressed to Ning Wang; nwang0902@gmail.com

Received 13 July 2021; Accepted 4 December 2021; Published 1 March 2022

Academic Editor: Atila Bueno



Learning from demonstration (LfD) is one of the promising approaches for fast robot programming. Most learning systems learn both movements and stiffness profiles from human demonstrations. However, they rarely consider the unknown environment interaction. In this paper, a robot human-like learning framework is proposed, where it can learn human skills through demonstration and complete the interaction task with an unknown environment. Firstly, the desired trajectory was generated by dynamic movement primitive (DMP) based on human demonstration. Then, an adaptive optimal admittance control scheme was employed to interact with environments with the reference adaptation method. Finally, the experimental study was conducted, and the effectiveness of the framework proposed in this paper was verified via a group of curved surface wiping experiments on a balloon with unknown model parameters.

## 1. Introduction

Robot learning from demonstration (LfD) has recently drawn much attention because of its high efficiency in robot programming [1]. Thus, robots can quickly program the robots to perform operating variable skills and replace human tutors from such tasks in a complex industrial environment [2]. Compared to conventional programming methods using a teaching pendant, LfD is an easier and more intuitive way for people who are unfamiliar with programming. Besides, human characteristics involved in the demonstrations are available for robots to further improve the flexibility and compliance of motions [3, 4].

After the demonstration, how to use the information of the human tutor is very important. Dynamic movement primitive (DMP) is a kind of common method in human-robot skill transfer tasks [5]. DMP has many advantages; for example, the DMP model is so simple that we only need to adjust a few parameters to achieve trajectory modeling. Besides, we can use regression algorithm to quickly learn model parameters in the online trajectory planning of robots [6]. In addition, the DMP model is also easy to generalize; we can quickly generalize a trajectory with the same style as the original trajectory by simply adjusting the starting and ending coordinates of the trajectory [7, 8]. Because of the above advantages, the DMP has been widely used in human-robot skill transfer tasks [9].

Appropriate control strategies help robots reproduce human skills more accurately and stably. In some specific tasks such as surface cleaning, cargo handling, and environment identification, robots are required to track a task trajectory and achieve compliance in the interaction with environments [10]. In the previous literature on interaction control, two main methods have been studied widely: impedance control [11] and hybrid position/force control [12]. The admittance control, which is regarded as the position-based impedance control, can achieve good interaction performance by trajectory adaptation [13–15]. According to admittance models, the external forces received by the robot will be transformed to the position of

the end-effector, and then, the desired interaction performance can be ensured by trajectory adaptation and tracking [16]. The control strategies mainly include proportion integration differentiation (PID) control, adaptive control, adaptive control using neural networks, and fuzzy control [17–20]. When robots perform different tasks in an unknown, complex, and dynamic environment, it is usually difficult to obtain accurate task models and environmental information, and various errors may have a serious impact on the final control results [21]. In recent years, control methods based on neural network learning have shown better adaptability to the system and environmental uncertainty, but this method requires a large amount of system data samples, and it is difficult to integrate various constraints in unknown environments in real-time [22, 23].

In this paper, firstly, the desired trajectory is generated by human demonstration, and then, an adaptive admittance control scheme is applied to interact with environments using the reference adaptation method. The contributions can be summarized as follows:

(1) An adaptive optimal admittance controller is developed to take into account the unknown interaction environment dynamics. Combining with the generalization ability of the DMP model and the compliance control ability of the adaptive optimal admittance model, the interaction performance between the robot and the unknown environment is improved.

(2) A complete human-like learning framework is developed. In the beginning, the desired trajectory can be obtained quickly and accurately by human teaching and generalization. And then, the online adaptive controller recalculates and updates the originally desired trajectory to obtain a new reference trajectory. The framework can update the new reference trajectory combined with different interaction environments, which greatly enhances the interaction accuracy

The rest of the article is organized as follows: In Section 2, the methods of desired trajectory generation and adaptive optimal admittance controller used in this paper are introduced. In Section 3, the experimental study is presented and then the effectiveness of the framework proposed in this paper is verified via balloon surface wiping experiments. Finally, Section 4 summarizes the whole paper.

## 2. Preliminaries and Methods

*2.1. Overview of the Framework.* The scheme of the proposed framework is shown in Figure 1. In the proposed learning framework, the human tutor presents a demonstration at first. The trajectory learned from the DMP model is regarded as the desired trajectory. Then, the desired trajectory and the interaction force measured by the force sensor are input into an adaptive admittance controller to obtain the modified reference trajectory. Here, $x$ and $\dot{x}$ represent the current position and velocity, respectively. $x_d$ and $x_d$ represent the desired position and velocity, respectively. $x_r$ represents the reference position. $q$, $\dot{q}$, and $\tau$ represent the current angle, angular velocity, and torque, respectively. $q_r$ and $\tau_r$ represent the reference angle and reference torque, respectively. $f_{\text{int}}$ represents the interaction force. Finally, new manipulation motions are implemented by the robot joint controller, and the collected new data are taken as a new demonstration for the repetitive training.

*2.2. Dynamic Movement Primitives (DMPs).* In this paper, motion DMP can be obtained by using the DMP model to fit the motion trajectory. The principles of motion DMP used in this paper are stated as follows [24, 25].

The essence of DMP is a second-order nonlinear dynamical system including spring and damper. A single-degree-of-freedom motion can be expressed by the following equations:

$$\tau\dot{\beta}_2 = a(g - \beta_1) - b\beta_2 + f(s;\omega), \tag{1}$$

$$\tau\dot{\beta}_1 = \beta_2, \tag{2}$$

$$\tau\dot{s} = -k_1 s, \tag{3}$$

where we ignore the time variable for the sake of simplicity. For example, $\beta_1(t)$ is represented by $\beta_1$; $a$ and $b$ represent the damping coefficient and spring constant of the system, respectively. And $a$ is usually set as $a = b^2/4$; $g$ is the target value of the motion trajectory; and $\tau$ represents the time scaling constant. $\beta_1$ and $\beta_2$ represent the position and velocity of motion trajectories, respectively. And the relationship between these two variables is shown in equation (2). $\omega$ means the weight of the Gaussian model. $s$ is the phase variable of the system, which is calculated by the regular system of equation (3). And $k_1$ is a positive constant. The nonlinear function $f(s;\omega)$ is defined as

$$f(s;\omega) = \frac{\sum_{i=1}^{N}\varphi_i\omega_i}{\sum_{i=1}^{N}\varphi_i}(g - \beta_0)s,$$
$$\varphi_i = \exp\left(-d_i(s - c_i)^2\right), \tag{4}$$

where $c_i$, di, and $\omega_i$ are the centre, width, and weight of the $i$-th kernel function, respectively. $\beta_0$ is the initial value of the motion trajectory. $N$ is the total amount of Gaussian models.

In general, the initial value of $s$ is set to 1, which gradually decays to zero. Since the value of $s$ tends to zero, the nonlinear function $f(s;\omega)$ is bounded, and the model becomes a stable second-order spring-damped system.

In general, supervised learning algorithms such as the local weighted regression (LWR) algorithm are used to determine model parameters $\omega$ [26]. Given the teaching trajectory $\beta(t)$, where $t = [1, 2, ..., T]$, and $g = \beta(T)$, the force function can be determined according to the following equation

$$f_{\text{target}} = \tau\dot{\beta}_2 - K(g - \beta_1) - D\beta_2, \tag{5}$$

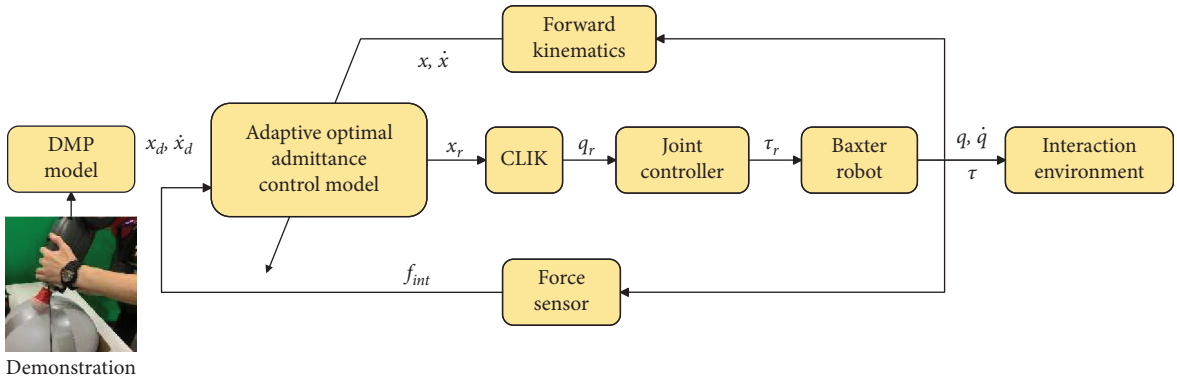



Figure 1: Overview of the proposed framework.

where $K$ and $D$ represent the stiffness and damping of the system, respectively. $\omega$ can be determined by the following equation:

$$\text{argmin}_{\omega} = \sum\left(f_{\text{target}} - f(s;\omega)\right)^2. \tag{6}$$

*2.3. Adaptive Optimal Admittance Control.* In this section, an adaptive task-specific admittance controller is developed. This adapts the parameters of the prescribed robot admittance model so that the robot system assists the human to achieve task-specific objectives. The task information is modeled by DMP so that the controller can adapt to the human tutor characteristics. The designed adaptive admittance controller will be used in the reproduction phase.

As shown in Figure 1, the process of adaptive admittance control in this article is as follows: the robot obtains the desired trajectory $x_d$, $\dot{x}_d$, $\ddot{x}_d$ through LfD and DMP generalization; then, the force sensor collects the interaction force between the robot end-effector and the environment in real-time. They are used as the input of the adaptive admittance model, and the expected trajectory $x_d$ is improved according to the admittance model. At this time, a new reference trajectory $x_r$ is obtained, and $x_r$ is transmitted to the controller as the control signal to ensure the fast and accurate tracking of the actual trajectory to the reference trajectory. Among them, the core of adaptive admittance control is that the model parameters are not fixed but can be optimized online with the help of an adaptive algorithm according to the real-time position and interaction force information, in order to minimize the quadratic cost function [13].

The prescribed admittance model is defined as follows:

$$M_E\left(\ddot{x} - \ddot{x}_d\right) + C_E\left(\dot{x} - \dot{x}_d\right) + K_E x = -f_{\text{int}}, \tag{7}$$

where $x$, $\dot{x}$, and $\ddot{x}$ represent the current position, velocity, and acceleration, respectively. $\dot{x}_d$ and $\ddot{x}_d$ represent the desired velocity and acceleration, respectively. $M_E$, $C_E$, and $K_E$ represent the unknown mass, damping, and stiffness matrices in the model, respectively. However, the mass matrix $M_E$ is usually high nonlinear. In this study, the mass-damping-stiffness model is simplified as the damping-stiffness model, which is used to interact with a balloon as a kind of flexible object. The simplified model is as follows:

$$C_E\left(\dot{x} - \dot{x}_d\right) + K_E x = -f_{\text{int}}. \tag{8}$$

Suppose the following continuous-time linear system:

$$\dot{\xi} = A\xi + Bu(t), \tag{9}$$

where $\xi = [x^T, x_d{}^T]^T$, $A = \text{diag}\{-C_E{}^{-1}K_E, I_n\}$, $B = [-C_E{}^{-1}, 0]^T$, and $u(t) = f_{int}$ is system input variable, which is related to the dynamic model of the interactive environment (8).

The optimal control input of the system is designed as $u = -K\xi$, and the control objective is to minimize the cost function by designing the control system. The cost function is defined as follows:

$$\begin{aligned} V(\text{t}) &= \int_0^{\infty}\left(\left(x - x_d\right)^T Q\left(x - x_d\right) + f_{\text{int}}^T R f_{\text{int}}\right)\text{dt}, \\ &= \int_0^{\infty}\left(\xi^T Q'\xi + u^T R u\right)\text{dt}, \end{aligned} \tag{10}$$

where $Q$ is a constant matrix, $Q' = [Q, -Q]^T$ [1, −1], which represents the weight matrix of tracking error. $R$ represents the weight matrix of the external force. In this paper, the design of the cost function takes into account the both robot system state and the external environment to evaluate the interactive control effect.

In the case that $A$ and $B$ are unknown constant matrices, an algorithm to obtain the optimal control signal by online learning is proposed. First, some variables are defined as follows:

$$\begin{aligned} \overline{\xi} &= \left[\xi_1^2, \xi_1\xi_2, \ldots, \xi_1\xi_m, \xi_2^2, \xi_2\xi_3, \ldots, \xi_m^2\right]^T, \\ \Delta_{\xi\xi} &= \left[\overline{\xi}\left(t_1\right) - \overline{\xi}\left(t_0\right), \overline{\xi}\left(t_2\right) - \overline{\xi}\left(t_1\right), \ldots \overline{\xi}\left(t_h\right) - \overline{\xi}\left(t_{h-1}\right)\right]^T, \\ I_{\xi\xi} &= \left[\int_{t_0}^{t_1}\xi\otimes\xi, \int_{t_1}^{t_2}\xi\otimes\xi, \ldots, \int_{t_{h-1}}^{t_h}\xi\otimes\xi\right]^T, \\ I_{\xi u} &= \left[\int_{t_0}^{t_1}\xi\otimes u, \int_{t_1}^{t_2}\xi\otimes u, \ldots, \int_{t_{h-1}}^{t_h}\xi\otimes u\right]^T, \end{aligned} \tag{11}$$

where $\xi$, $\Delta_{\xi\xi}$, $I_{\xi\xi}$, and $I_{\xi u}$ are the intermediate variable matrices used to calculate the state feedback gain. $h$ is the

degree of integration and ⊗ represents the Kronecker product. We define that $P_K$ which transforms matrix into vector form, and $P_K$ is a symmetric matrix:

$$P_K = \left[P_{11}, 2P_{12} \ldots 2P_{1m} \ldots P_{mm}\right]^T. \tag{12}$$

The principle of the adaptive optimal admittance scheme is summarized in Algorithm 1 [27, 28].

$$\begin{aligned} \Theta_k &= \left[\delta_{\xi\xi}, -2I_{\xi\xi}\left(I_m \otimes K_k^T R\right) - 2I_{\xi u}\left(I_m \otimes R\right)\right], \\ \Xi_k &= -I_{\xi\xi}\mathrm{vec}\left(Q_k\right), \\ Q_k &= Q + K_k^T R K_k. \end{aligned} \tag{13}$$

where $I_m$ is the $m$-dimension identity matrix and $vec(\star)$ is the function that transforms the matrix into a vector. Through equation (17), we can obtain the optimal feedback control gain $K_K*$. Substituting $K_K*$ into $u = -K\ \xi$, the optimal feedback control signal $u$ can be obtained.

*2.4. Inverse Kinematics Using CLIK.* The closed-loop inverse kinematics (CLIK) algorithm is employed to resolve the Cartesian reference trajectory $x_r$ into $q_r$ in joint space [29–31]. The solution error is $e = k(q_r) - x_r$, where $k(\star)$ denotes the forward kinematics and $e$ is given by

$$\dot{e} = -K_e e. \tag{14}$$

where $K_e$ is a positive user-defined matrix that decides the convergent rate of $e$. Expanding the above equations and combining with $x = J_i\ q$ and $J_i = \partial k(q)/q$, Ji is the Jacobian matrix of the robot. The following equation is obtained:

$$\dot{q}_r = J_i^+\left(\dot{x}_r - K_e\left(k\left(q_r\right) - x_r\right)\right). \tag{15}$$

Furthermore, we obtain the CLIK method:

$$q_r = \int_0^t \left(J_i^+ \dot{x}_r - J_i^+ K_e\left(k\left(q_r\right) - x_r\right)\right)\mathrm{dt}, \tag{16}$$

where $J_i^+ = J_i^T\left(J_i J_i^T + \delta I_n\right)^{-1}$ .

## 3. Experiment and Analysis

In this section, the performance of the proposed learning framework was validated by conducting experiments on a 7-DOF Baxter robot, as shown in Figure 2. The manipulator was equipped with the ATI Mini45 Force/Torque force sensor. The end effector was wrapped in a towel used to wipe the drawn curve on the surface of the balloon. The force sensor and the system controller are communicated by the UDP protocol whose sampling rate and control rate are set as 100 and 50 Hz, respectively. To prevent the displacement of the balloon from affecting the experimental results, the balloon to be wiped was fixed in the fixing box. The box is a paper carton with a size just big enough to hold the balloon. The paper carton is of size 43 cm × 32 cm × 18 cm and is fixed on the test bench with adhesive tape.

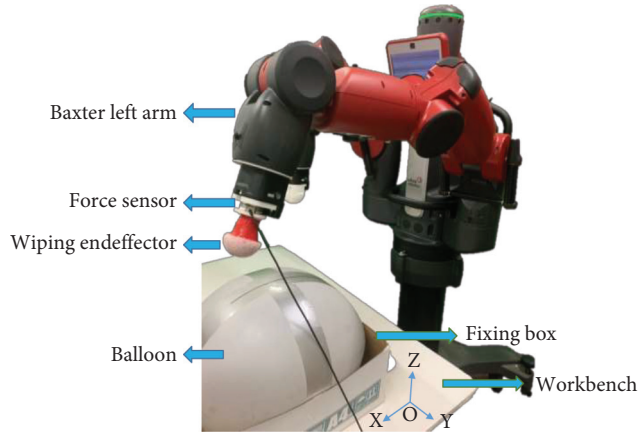



Figure 2: Experimental platform based on Baxter robot.

*3.1. Demonstration Stage.* In the teaching stage, first, a curve was drawn on the surface of the balloon with a whiteboard pen, and then, the human tutor dragged the left arm of the Baxter robot to complete the teaching task—wiping. In the meantime, the teaching trajectory information was recorded and input into the DMP model through the program. Then, the system learnt and generalized to obtain the desired trajectory $x_d$. At the same time, the force sensor recorded the interactive forces in the $X$, Y, and $Z$ directions for subsequent analysis.

*3.2. Reproduction of the Wiping Task.* In the beginning, a new curve was drawn on the surface of the balloon. The robot end-effector was controlled to move to the starting point of the desired trajectory at [0.992, 0.280, 0.227] $m$. At this time, the end of the robot arm had interacted with the environment and changed from free space motion to constrained space motion. Since balloon was used as an interactive environment, its parameters are unknown. So an adaptive optimal admittance control was proposed to solve this problem. According to the set cost function, online adaptive learning of the interactive environment model parameters could help to achieve our desired control effects and complete the wiping task of the new curve.

The first wiping experiment process was that the trajectory obtained by a demonstration was directly used as the reference trajectory, and the admittance model parameters were specified as $C_E = [-0.5,\ 0.01,\ -0.8]$ and $K_E = [7,\ 2,\ 10]$. In the second experiment, the trajectory obtained by teaching was input into the DMP model for learning and generalization. The generated trajectory was used as the reference trajectory and then applied to the admittance model of the first experiment. In the third experiment, the expected trajectory obtained by DMP learning generalization of the teaching trajectory was used as the input of the adaptive optimal admittance controller, and finally, a new reference trajectory was obtained and then input to the Baxter joint controller. The initial value of state feedback gain in the $X$, $Y$, and $Z$ directions was set to [−10, 1], and the weight matrix of cost function was $Q =$ [200, −200], $R = 5$. Finally, the optimal state feedback gains

**Input:** the set of initial feedback gain $K_0$ and state variable $\xi$;
**Output:** optimal feedback gain $K_K^*$;
phase 1: set $f_{int0} = K_0\xi$ as the initial input while the manipulator is contacting with the environment;
**Repeat:** compute $\Delta_{\xi\xi}$, $I_{\xi\xi}$, and $I_{\xi u}$;
**Until:** rank$[I_{\xi\xi}, I_{\xi u}] = m(m+1)/2 + mr$;
Repeat:
phase 2: solve $P_K$ and $K_{K+1}$ according to:

$$\begin{bmatrix} \widehat{P}_K \\ \mathrm{vec}(K_{K+1}) \end{bmatrix} = (\Theta_k^T\Theta_k)^{-1}\Theta_k^T\Xi_k.$$

**Until:** $\|P_K\text{-}P_{K\text{-}1}\| < \varepsilon$;
**Return:** $K_K^*$;

Algorithm 1: Framework of the adaptive optimal admittance scheme.

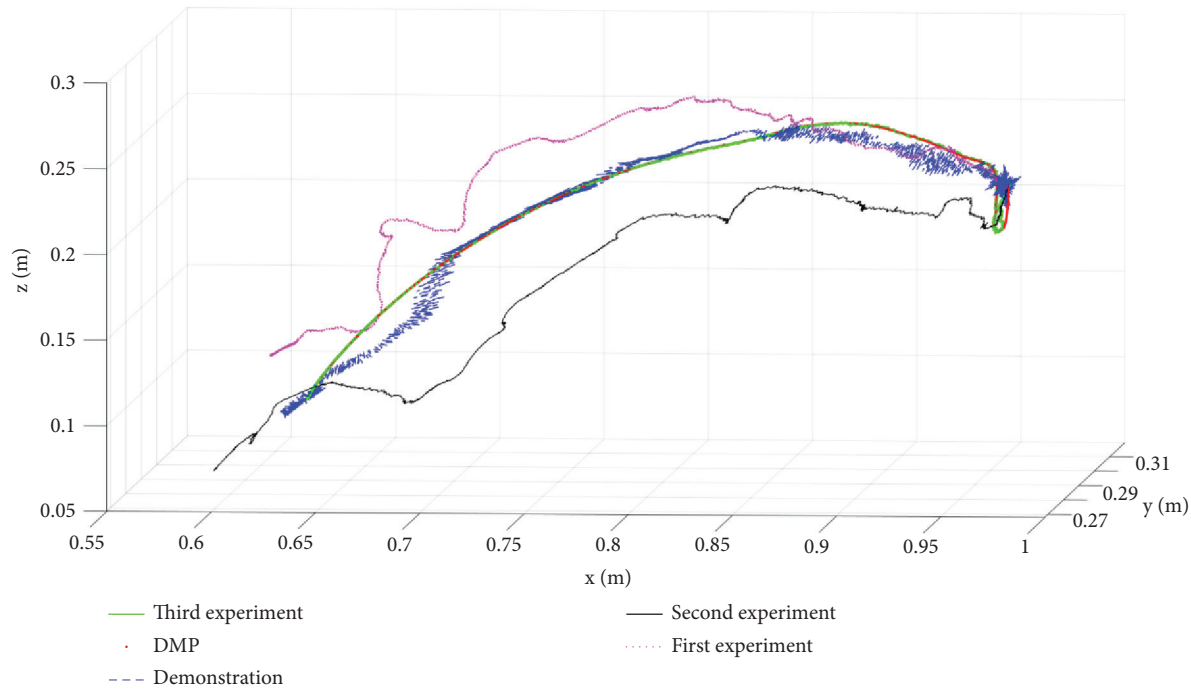



Figure 3: Three-dimensional curves of teaching trajectory, DMP generalization trajectory, and three times experimental trajectory.

in the $X$, $Y$, and $Z$ directions were $K^*_{Kx} = [-16.1083, 10.7866]$, $K^*_{Ky} = [2.3554, 9.6812]$, and $K^*_{Kz} = [10.7259, 80.5781]$, respectively. Next, to verify the effectiveness of the proposed framework, the effects of the above three experiments were compared and the trajectory tracking error and interaction force changes were analyzed.

*3.3. Experimental Results and Discussion.* First of all, we give the three-dimensional curves of teaching trajectory, DMP generalization trajectory, and three experimental trajectories in the same space rectangular coordinate system, as shown in Figure 3. As can be seen from the figure, the unprocessed teaching trajectory has serious jitter, and the expected trajectory after DMP generalization becomes smoother.

In the first experiment, the teaching trajectory is used as the reference input of the joint controller of the Baxter robot. The wiping effect is shown in Figure 4(b). It can be seen that the wiping task is not successfully completed under these experimental conditions. The time-varying curve of the interaction force during this process is shown in Figure 5. It shows that the robot performed the wiping task between 6 s and 18 s. However, its interaction force is not large enough and later becomes smaller and smaller close to 0. So, it fails to wipe the handwriting clean.

In the second experiment, although the curve is able to be erased from the surface of the balloon, it can be seen from the force trajectory that the interaction force in this case is very large. As shown in Figure 6, the maximum force in the $Z$ direction is up to 25 N. The actual task figure (Figure 4(c)) also shows that a very serious inward depression occurs on the balloon at this time. Assuming that the interactive environment is not a flexible object as a balloon and the rigidity is very large, it may cause some damage to the robot arm or the interactive object. Therefore, it is obvious that this experiment fails to complete the wiping task well.

(a) (b) (c) (d)

Figure 4: The actual effect of the wiping experiment. (a) Object to be wiped. (b) The wiping effect of the first experiment. (c) The wiping effect of the second experiment. (d) The wiping effect of the third experiment with the proposed framework.

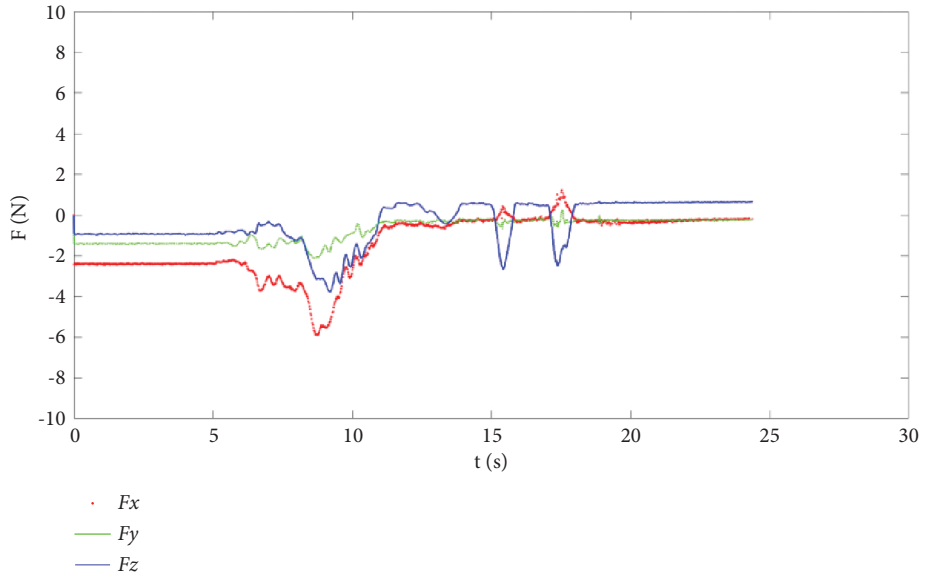



Figure 5: Change curve of interaction force in the first experiment. *Fx*, *Fy*, and *Fz* represent interaction forces in the *X*, *Y*, and *Z* directions.

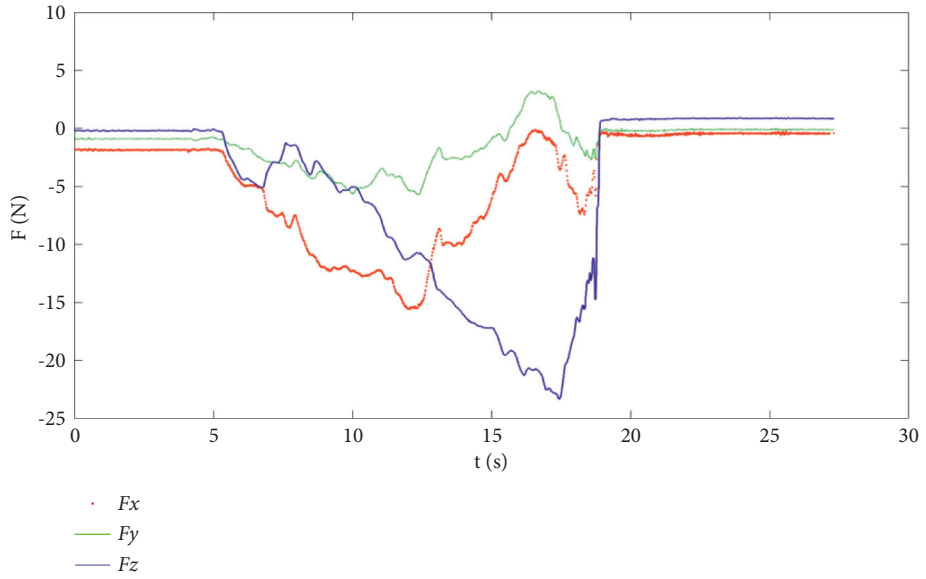



Figure 6: Change curve of interaction force in the second experiment. *Fx*, *Fy*, and *Fz* represent interaction forces in the *X*, *Y*, and *Z* directions.

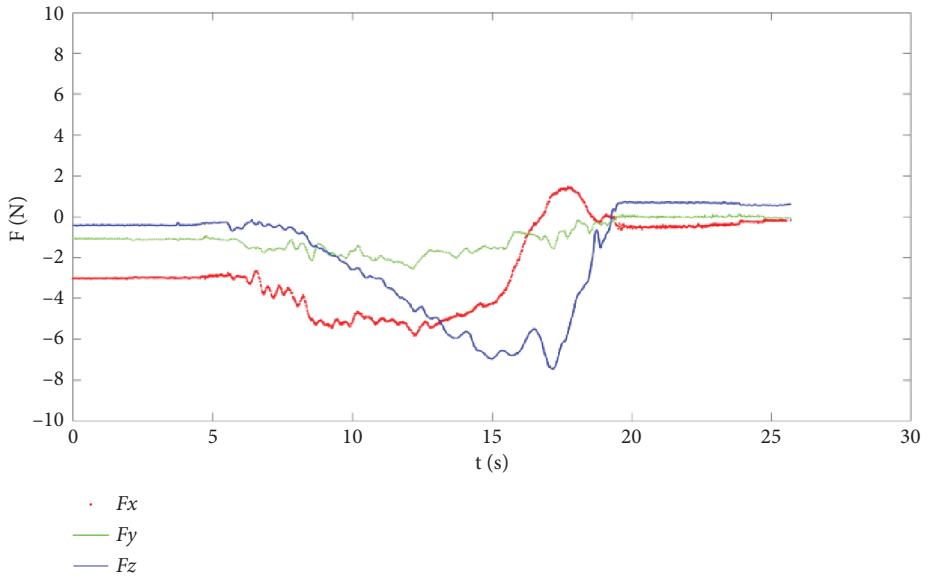



Figure 7: Change curve of interaction force in the third experiment. *Fx*, *Fy*, and *Fz* represent interaction forces in the *X*, *Y*, and *Z* directions.

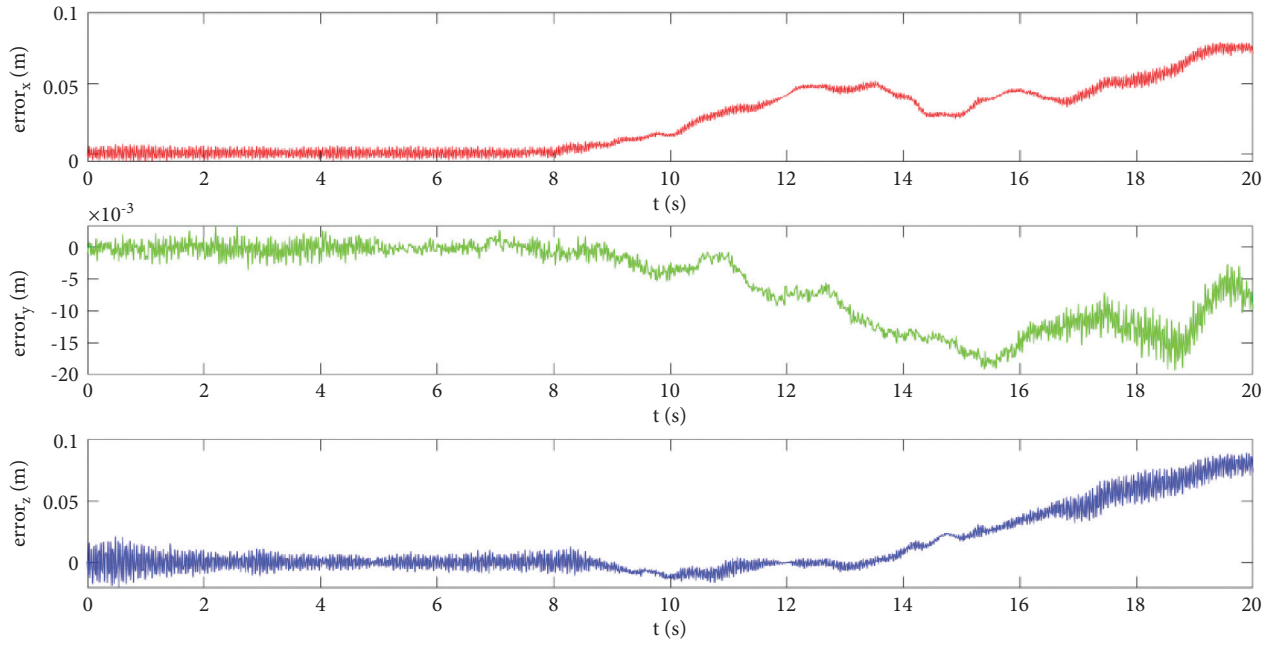



Figure 8: Position tracking error curve of the first experiment. The $error_x$, $error_y$, and $error_z$ represent the position errors in the *X*, *Y*, and *Z* directions, respectively.

The third experiment is based on the learning framework proposed in this paper. It can be seen from Figure 4(d) that the wiping effect is greatly improved compared with the previous two, and the handwriting curve can be basically wiped clean. The interactive force graph (Figure 7) shows that the robot interacted with the balloon in about 6–20 s, during which the force in the *Z* direction, as the dominant force of the wiping task, changes smoothly between 0 N and 8 N. Compared with the previous two experiments, it is obvious that the interactive force is relatively optimal.

From the three-dimensional trajectory in Figure 3, it can be seen intuitively that there are large errors between the first experiment and the third experiment in terms of the expected trajectory. The interaction between the manipulator end effector and the balloon is basically completed within 20 seconds. The following is to analyze and discuss the change of trajectory tracking error in this stage. As shown in Figure 8, in the first experiment, the tracking error reaches the maximum in the later stage of the wiping task, and the maximum value in the *Z* direction has exceeded 0.1 m. Likewise, the trajectory tracking error of the second experiment (Figure 9) is also relatively large. Only the third experiment can track the expected trajectory well. The tracking error curves in the *X*, *Y*, and *Z* directions

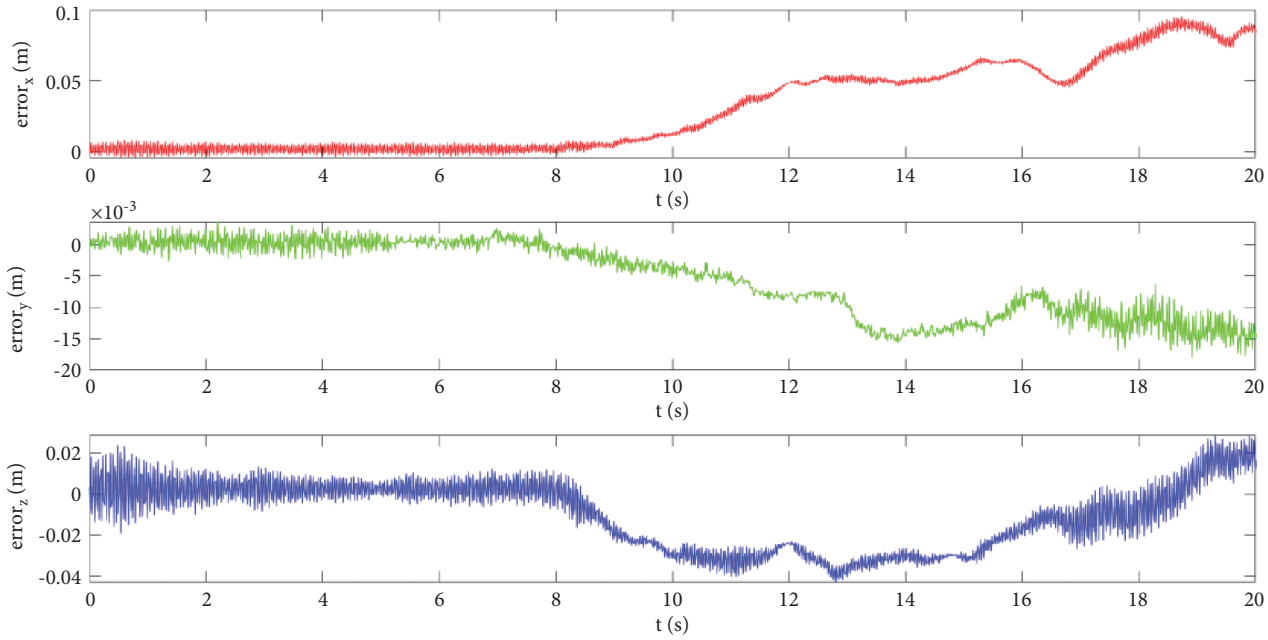



Figure 9: Position tracking error curve of the second experiment. The $error_x$, $error_y$, and $error_z$ represent the position errors in the $X$, $Y$, and $Z$ directions respectively.

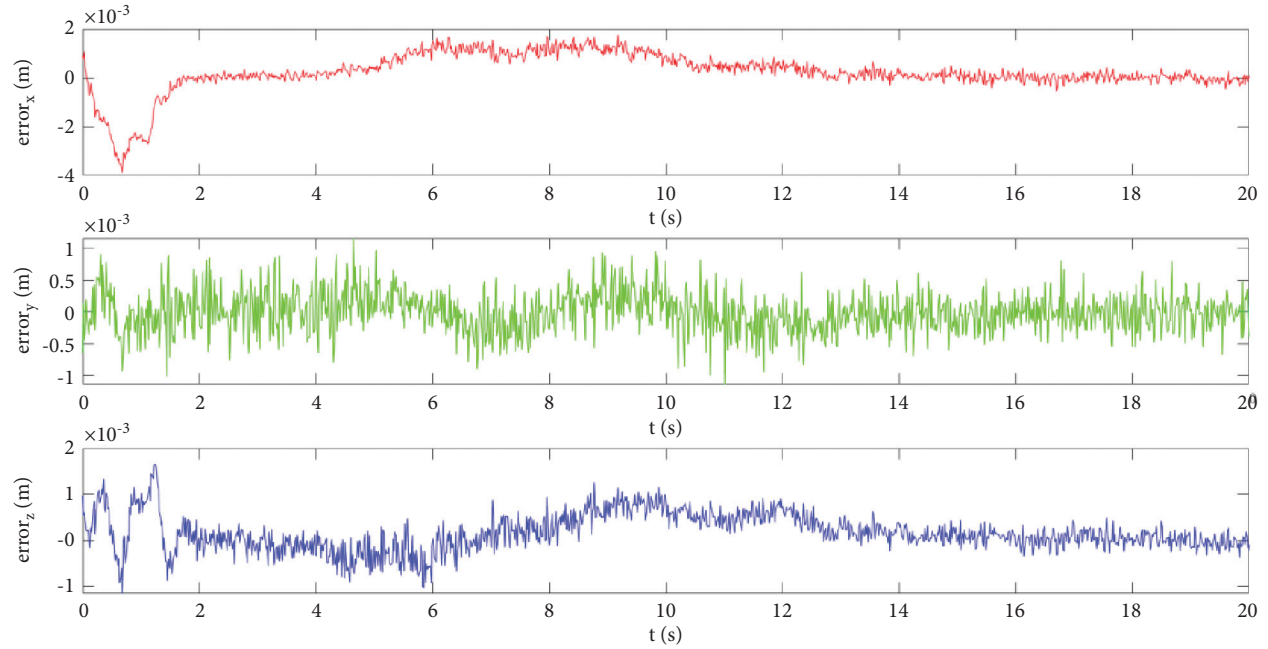



Figure 10: Position tracking error curve of the third experiment. The $error_x$, $error_y$, and $error_z$ represent the position errors in the $X$, $Y$, and $Z$ directions, respectively.

in the third experiment are shown in Figure 10. The error values all lie within the range of ±0.04 m and, in the process of wiping, the error is basically stable around 0. It also proves that the learning framework proposed in this paper is able to well track the reference trajectory.

## 4. Conclusions

In this paper, a robot human-like learning framework based on robot and unknown environment interaction was proposed. The LfD approach can make the robot obtain the

reference input more quickly and accurately. At the same time, combined with the generalization ability of the DMP model and the compliance control ability of the adaptive optimal admittance model, the interaction performance between the robot and the unknown environment was enhanced. At last, the effectiveness of the proposed framework was verified by the wiping experiment of the balloon surface. Our future work will apply the proposed framework to different complex tasks and environments such as writing on an unknown curved surface, and the learning and generalization of interaction force will also be considered in force control.

## Data Availability

The detailed parameters of used model and controller are given in the article. The results are computed on the Matlab 2018a software, while the relevant results are also given in the manuscript.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was partially supported by the Industrial Key Technologies R&D Program of Foshan (2020001006308) and the National Natural Science Foundation of China under Grant no. 61803039.

## References

[1] J. Aleotti and S. Caselli, "Robust trajectory learning and approximation for robot programming by demonstration," *Robotics and Autonomous Systems*, vol. 54, no. 5, pp. 409–413, 2006.

[2] C. Yang, C. Chen, N. Wang, Z. Ju, J. Fu, and M. Wang, "Biologically inspired motion modeling and neural control for robot learning from demonstrations," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, no. 2, pp. 281–291, 2018.

[3] R. Dillmann, T. Asfour, M. Do et al., "Advances in robot programming by demonstration," *KI - Künstliche Intelligenz*, vol. 24, no. 4, pp. 295–303, 2010.

[4] M. G. Han and J. C. Koo, "DMP accuracy improvement to facilitate learning from demonstration for industrial cooperative robots," *Journal of Institute of Control*, vol. 26, no. 12, pp. 1062–1066, 2010.

[5] A. J. Ijspeert, J. Nakanishi, H. Hoffmann, P. Pastor, and S. Schaal, "Dynamical movement primitives: learning attractor models for motor behaviors," *Neural Computation*, vol. 25, no. 2, pp. 328–373, 2013.

[6] Y. Cohen, O. Bar-Shira, and S. Berman, "Motion adaptation based on learning the manifold of task and dynamic movement primitive parameters," *Robotica*, vol. 39, no. 7, pp. 1299–1315, 2020.

[7] M. Karlsson, A. Robertsson, and R. Johansson, "Autonomous interpretation of demonstrations for modification of dynamical movement primitives," in *Proceedings of the 2017 IEEE International Conference on Robotics and Automation (ICRA)*, Singapore, May 2017.

[8] A. B. Pehlivan and E. Oztop, "Dynamic movement primitives for human movement recognition," in *Proceedings of the IECON 2015 - 41st Annual Conference of the IEEE Industrial Electronics Society IEEE*, Yokohama, Japan, November 2015.

[9] Z. Chao, C. Yang, Z. Chen, and S. L. Dai, "Robot learning human stiffness regulation for hybrid manufacture," *Assembly Automation*, vol. 38, no. 5, pp. 539–547, 2018.

[10] Z.-Y. Liu and H. Qiao, "Multiple ellipses detection in noisy environments: a hierarchical approach," *Pattern Recognition*, vol. 42, no. 11, pp. 2421–2433, 2009.

[11] N. Hogan, "Impedance control: an approach to manipulation: Part I-theory," *Journal of Dynamic Systems, Measurement, and Control*, vol. 107, no. 1, pp. 1–7, 1985.

[12] M. H. Raibert and J. J. Craig, "Hybrid position/force control of manipulators," *Journal of Dynamic Systems, Measurement, and Control*, vol. 103, no. 2, pp. 126–133, 1981.

[13] K. P. Tee, R. Yan, and H. Li, "Adaptive Admittance Control of a Robot Manipulator under Task Space Constraint," in *Proceedings of the IEEE International Conference on Robotics & Automation*, Anchorage, AK, USA, May 2010.

[14] H. Seraji, "Adaptive Admittance Control: An Approach to Explicit Force Control in Compliant Motion," in *Proceedings of the IEEE International Conference on Robotics & Automation*, Washington, DC, USA, May 2002.

[15] K. Yamane, "Admittance control with unknown location of interaction," *IEEE Robotics and Automation Letters*, vol. 6, no. 2, pp. 4079–4086, 2021.

[16] I. Ranatunga, F. L. Lewis, D. O. Popa, and S. M. Tousif, "Adaptive admittance control for human–robot interaction using model reference design and adaptive inverse filtering," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 1, pp. 278–285, 2016.

[17] G. Lai, Z. Liu, Y. Zhang, and C. L. Philip Chen, "Adaptive fuzzy tracking control of nonlinear systems with asymmetric actuator backlash Based on a new smooth inverse," *IEEE Transactions on Cybernetics*, vol. 46, no. 6, pp. 1250–1262, 2017.

[18] H. Huang, T. Zhang, C. Yang, and C. L. P. Chen, "Motor learning and generalization using broad learning adaptive neural control," *IEEE Transactions on Industrial Electronics*, vol. 67, no. 10, pp. 8608–8617, 2020.

[19] C. Yang, X. Wang, L. Cheng, and H. Ma, "Neural-learning-based telerobot control with guaranteed performance," *IEEE Transactions on Cybernetics*, vol. 47, no. 10, pp. 3418–3159, 2017.

[20] H. Huang, C. Yang, and C. L. P. Chen, "Optimal robot-environment interaction under broad fuzzy neural adaptive control," *IEEE Transactions on Cybernetics*, vol. 51, no. 7, pp. 3824–3835, 2021.

[21] H. Zhan, D. Huang, Z. Chen, M. Wang, and C. Yang, "Adaptive dynamic programming-based controller with admittance adaptation for robot-environment interaction," *International Journal of Advanced Robotic Systems*, vol. 17, no. 3, 11 pages, 2020.

[22] C. Mei, J. Yuan, and R. Guan, "Adaptive Unified Impedance and Admittance Control Using Online Environment Estimation," in *Proceedings of the IEEE International Conference on Robotics and Biomimetics (ROBIO)*, Kuala Lumpur, Malaysia, December 2018.

[23] L. Roveda, N. Iannacci, F. Vicentini, N. Pedrocchi, F. Braghin, and L. M. Tosatti, "Optimal impedance force-tracking control design with impact formulation for interaction tasks," *IEEE Robotics and Automation Letters*, vol. 1, no. 1, pp. 130–136, 2016.

[24] C. Yang, C. Chen, W. He, R. Cui, and Z. Li, "Robot learning system based on adaptive neural control and dynamic

movement primitives," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 3, pp. 777–787, 2019.

[25] A. Ude, B. Nemec, T. Petrić, and J. Morimoto, "Orientation in Cartesian Space Dynamic Movement Primitives," in *Proceedings of the 2014 IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong, China, May 2014.

[26] C. G. Atkeson, "Using locally weighted regression for robot learning," in *Proceedings of the - IEEE International Conference on Robotics and Automation*, vol. 2, pp. 958–963, Sacramento, CA, USA, April 1991.

[27] C. Yang, G. Peng, Y. Li, R. Cui, L. Cheng, and Z. Li, "Neural networks enhanced adaptive admittance control of optimized robot-environment interaction," *IEEE Transactions on Cybernetics*, vol. 49, no. 7, pp. 2568–2579, 2018.

[28] Y. Jiang and Z.-P. Jiang, "Computational adaptive optimal control for continuous-time linear systems with completely unknown dynamics," *Automatica*, vol. 48, no. 10, pp. 2699–2704, 2012.

[29] J. Marti-Saumell, A. Santamaria-Navarro, C. Ocampo-Martinez, and J. Andrade-Cetto, "Multi-task closed-loop inverse kinematics stability through semidefinite programming," in *Proceedings of the 2020 IEEE International Conference on Robotics and Automation (ICRA) IEEE*, Paris, May 2020.

[30] D. Huang, H. Zhan, and C. Yang, "Impedance model-based optimal regulation on force and position of bimanual robots to hold an object," *Complexity*, vol. 2020, no. 3, pp. 1–13, 2020.

[31] D. Siciliano, "A closed-loop inverse kinematic scheme for online joint-based robot control," *Robotica*, vol. 8, no. 3, pp. 231–243, 1990.

Hindawi
Complexity
Volume 2022, Article ID 1742772, 14 pages
https://doi.org/10.1155/2022/1742772

WILEY | Hindawi

*Research Article*

# Simulation Testing of Maritime Cyber-Physical Systems : Application of Model-View-ViewModel

**Dong-Chul Lee,[1] Kyung-Min Seo,[2] Hee-Mun Park,[2] and Byeong Soo Kim[3]**

[1]*Daewoo Shipbuilding & Marine Engineering (DSME), Siheung 15011, Republic of Korea*
[2]*Korea University of Technology and Education (KOREATECH), Cheonan 31253, Republic of Korea*
[3]*Myongji University, Seoul 03674, Republic of Korea*

Correspondence should be addressed to Byeong Soo Kim; kimbs@mju.ac.kr

Received 5 August 2021; Revised 12 October 2021; Accepted 29 October 2021; Published 21 January 2022

Academic Editor: Bo Xiao



From the perspective of the system of systems development, system-level functional testing is required for designing subsystems. This study utilizes modeling and simulation techniques to analyze the operational behaviors of the subsystems and confirm data communication between them. The targeted system in the study is a naval combat system (NCS), which is a typical type of defense cyber-physical system (CPS). Three types of models were designed for the simulation testing of the NCS: a combat-management model for simulating the overall computational activities, physical models to confirm the intrasubsystem behaviors, and data integration models to test the intersubsystem communications. These models are realized with the Model-View-ViewModel design pattern, which strongly facilitates graphical user interfaces being decoupled from model logic and data. We consider underwater combat scenarios as an application. Six significant physical subsystems within the NCS are simulated and tested: a ship-steering system, an inertial navigation system, a global navigation satellite system, a periscope, sonar systems, and a plotting board. We expect that the proposed work will play a principal role when analyzing the behaviors and communications of defense CPSs and providing an environment for functional testing as a digital twin.

## 1. Introduction

A military combat system is a type of maritime cyber-physical system (CPS) [1–3] that collects physical data through sensor subsystems [4]. In turn, it makes real-time decisions by feeding the data into computing subsystems [5]. As a naval combat system (NCS) for a surface ship or underwater vessel is highly complex, its functionalities regarding sensors, computers, and communications should be tested individually in system of systems engineering [6, 7].

For the functional testing of an NCS, the following real-world requirements at the subsystem level must be considered. Firstly, most physical subsystems of an NCS are highly distributed [8]. As the subsystems operate separately, they have multiple console panels to identify and control their operations. Next, the subsystems usually have different data formats and communication protocols [9]. To communicate between the subsystems normally, NCS requires data integration systems (DISs) to convert their input/output (I/O) data into interpretable forms. Thus, DIS provides the loose connections of subsystems, which guarantees flexible system adjustment, e.g., by allowing the addition, removal, and change of the subsystems.

Because of massive costs in resources and time, it is difficult to test an entire NCS with actual subsystems. For such cases, modeling and simulation (M&S) techniques have become good alternatives. Although CPSs containing an NCS are intrinsically complicated, well-categorized models for them are understandable, analyzable, and certifiable [10–12]. Therefore, they give helpful simulation results and insights to develop CPSs successfully [13–15]. Based on the requirements of an NCS, the functional scope of the simulation testing conducted for this study is two-fold: (1) the subsystems' operational behaviors and (2) data conversion between them.

Various M&S methods for defense CPSs have been used over the last decade. Some researchers have proposed classification methods for CPS models [16–19], and others have developed simulation systems with diverse modeling languages [20–22]. The collective contribution of these studies has been to suggest meaningful modeling methods and simulation software for defense CPSs. Despite this contribution, the methods still warrant some improvements. For example, some CPS models lack a realistic configuration of their target systems. They focus on evaluating a CPS's integrated performance as a whole rather than analyzing the operational behaviors of subsystems. In addition, some researchers developed simulation systems with a centralized approach so that the communications between the subsystems could not be simulated or expanded. Thus, the previous techniques were not suitable for subsystem-level simulation testing.

This study presents a practical simulation approach to testing defense CPSs, specifically an NCS, at the subsystem level. We functionally classify CPS models that closely resemble the components of the real NCS. Then, we develop a distributed simulation system to help with analyzing intrasubsystem behaviors and confirm the conversion of intersubsystem data.

The designed models were implemented using the Model-View-ViewModel (MVVM) design pattern [23] in Windows Presentation Foundation (WPF). The main advantage of the MVVM pattern is to decouple the graphical user interface (GUI) significantly from the model logic and data [24]. We created GUIs for the subsystems' console panels with extensible application markup language [25]. The logical parts of the models—e.g., the decisions, dynamics, and data conversions—are implemented in the models of the MVVM pattern. This separation between data and logic increases the transparency, modularity, and flexibility of the overall modeling, i.e., the external graphical interfaces and the internal model logic. For this reason, the MVVM design pattern has been used widely to develop software for human-computer interaction (HCI) in many industrial fields [26, 27]. To the best of our knowledge, this study is the first attempt to develop a practical implementation of the MVVM design pattern for simulating a naval CPS.

As a practical case study, we simulated underwater combat operations. The physical subsystems to be modeled in this simulation included a ship-steering system, an inertial navigation system (INS), a global positioning system (GPS), a periscope, sonar systems, and a plotting board. A DIS model was developed to convert simulated data from the INS model into real-system data comprising hexadecimal values with a predefined header. The empirical results show how the physical models behaved during the simulation and how the DIS model transformed the collected data into compatible information for system scalability. With several discussions, we expect that the developed simulation system can be utilized to test an NCS at the subsystem level successfully.

The study is organized as follows: Section 2 describes the NCS briefly to provide the background, and Section 3 analyzes previous works. Section 4 proposes the overall simulation architecture and the realization of the three model types based on the MVVM pattern. Section 5 explains and discusses an application for naval combat simulation. Finally, Section 6 presents our conclusions.

## 2. Background

An NCS, as a defense CPS, evaluates real-time threat situations and assigns appropriate weapons based on complex computations. Figure 1 shows a simplified configuration of an NCS. It comprises the multiple families of physical subsystems that perform similar functions (the red part in Figure 1). For example, the navigation sensor systems containing a GPS and an INS calculate the position, orientation, and velocity of their vessel. Underwater sensor systems, as typified by sonar systems, mainly detect objects in the water, and surface sensor systems, such as radar or periscope systems, perceive objects on the water's surface [28]. The underwater and surface sensor systems must be integrated with the navigation sensors because the accuracy of target detection requires precise self-localization. Weapon systems, e.g., missiles or torpedoes, contain independent objects launched from the vessel against offensive and defensive tactics.

As a computational part of the CPS, a combat management system (CMS) utilizes all information gathered from sensor and weapon systems [29]. It continuously performs the information-processing stages in the computational domain. For example, the CMS perceives the surroundings of its vessel and identifies approaching threats. After tracking them, it launches appropriate weapons and controls them, if necessary [30]. The CMS also contains several subsystems. As shown in the green part of Figure 1, it has multifunction consoles for HCI. The multifunction consoles display the overall tactical situation upon awareness by data integration from signal-processing units. They assist vessel operators with understanding the situation and commanding their decision-making.

An NCS configures a combat system data bus (CSDB) and a central bus network to connect all the physical and computational subsystems. Note that the physical subsystems do not plug into the CSDB directly. The subsystems within the same group link to a data-integration system (DIS), which is interconnected across the CSDB sequentially [31]. In this case, the DIS (the blue part in Figure 1) performs essential roles as a gateway [32]. As each subsystem has a different data format and communication type, the DIS converts its I/O data to compatible forms with the CSDB. It also allows new subsystems to be added to the same system group without directly affecting the CSDB. Therefore, the physical subsystems are connected through interoperability arrangements, which guarantees that they do not require strong coupling or tight integrations [33].

To summarize, the NCS is a defense CPS that has physical and computational subsystems connected via a centralized network [34]. For data conversion and system scalability, NCS also contains several DISs according to the subsystem types. In this regard, we developed three models for simulation testing: combat management (CM), physical, and DIS models.

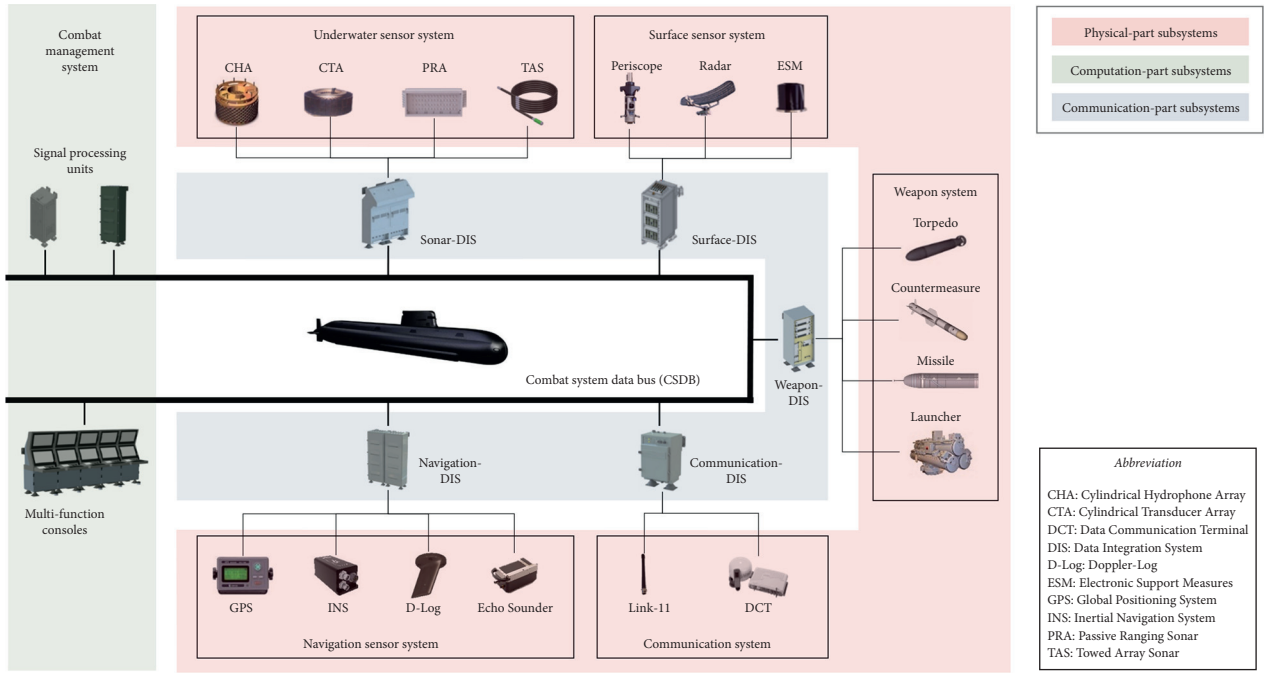



Figure 1: Simplified configuration of the naval combat system classified into three components: physical, computation, and communication subsystems.

## 3. Literature Review

Simulation testing utilizes the computerized models of a real system to evaluate the corresponding system's functionality under a given set of conditions [35]. Over the last decade, various studies have been conducted to provide indications for the simulation testing of defense CPSs. In this section, we categorize them by aspects of their modeling and simulation, as summarized in Table 1.

From a modeling perspective, several classification methods have been proposed for defense CPS models. For example, Sung and Kim [16] developed a collaborative modeling methodology for a domain-specific discrete event-simulation system, which categorizes models as discrete event-level models or behavioral-level models. Li et al. [17] similarly explained two-level modeling (physical and decision modeling) for defense CPSs. They focused on decision modeling rather than physical modeling because combat effectiveness is closely related to the tactics of the defense CPS under certain conditions. Zhu et al. [18] introduced domain-specific metamodeling with two orthogonal dimensions. Ontological modeling locates a model element from a domain-definition perspective, and linguistic modeling is concerned with modeling language definition. Park et al. [19] split a defense system into two parts for CPS modeling: core parts representing the system's physical and performance characteristics and shell parts for its operational attributes.

Although the above modeling methods provide transparent classification systems, they focused more on analyzing the integrated functionalities of the target systems. They tend to explain a holistic view of their modeling rather than the modeling itself. Thus, they have a weakness when evaluating the functionalities of subsystems, such as navigation, radar, and sonar systems. In addition, as the methods were developed to enhance general modeling indices, e.g., model composability and reusability, the authors were not interested in basing the model components on a realistic system configuration.

Simulation studies have been conducted to represent defense CPS models graphically and analyze them. Squarcia et al. [20] developed a numerical simulation tool named Integrated Combat Systems Simulation (ICS-Sim), which allows the integration of simulation models to analyze and evaluate systems' performance for studying innovative algorithms. As these studies were fundamentally developed with a centralized simulation environment, it is difficult to apply them to geographically dispersed simulation systems with a common simulation platform.

Similar to our simulation approach, some researchers focused on distributed simulation with realistic components of defense CPSs. For example, Xu et al. [21] federated one computational simulation system at the engagement level and three physical simulation systems at the engineering level. Cheng et al. [22] also developed a distributed simulation system containing a general simulation platform, several observation systems, and a navigation system. They focused on how to construct interoperable simulations with legacy systems and achieve data flows according to real-time constraints. Despite their contributions, however, these works do not provide behaviors or communication data at the subsystem level.

In recent years, our research team has carried out various M&S studies for defense CPSs. For successful model-based systems engineering (MBSE), we have studied M&S activities across all phases of MBSE. For example, we have

Table 1: Related works for modeling and simulation studies of defense cyber-physical systems.

| Previous work | Motivation | Method |
|---|---|---|
| Sung and Kim [14–16] | A collaborative modeling process was proposed. It formally defines the roles and responsibilities of domain engineers, M&S engineers, and platform engineers. | The models can be switched at the run-time or the load-time of a simulation run by the dynamically linked library technology. |
| Li et al. [15–17] | Domain-specific modeling containing physical and intelligent aspects was developed. | Code generation mechanisms transform domain-specific decision models to python code. |
| Zhu et al. [16–18] | The proposed two-level metamodeling was based on petri nets to support continuous state transitions and event triggering. | MetaDepth was used for supporting the two-level textual modeling and implementing the deep semantics. |
| Park et al. [17–19] | A defense CPS model is classified into core and shell parts. The two parts are specified with the DEVS (discrete event systems specification) formalism. | The Delta3D engine was used to represent combat entities graphically within a centralized simulation environment. |
| Squarcia et al. [18–20] | A multiscale method was applied for modeling and simulation of systems/subsystems with a different degree of detail and accuracy. | An integrated simulation environment was developed for the performance evaluation of entire naval CPSs. |
| Xu et al. [19–21] | The proposed warship model contains one computational model for overall tactical behaviors and three engineering models regarding weapons. | A specific software-hardware environment was required to interconnect legacy systems. |
| Cheng et al. [20–22] | The proposed model for a CPS contains navigation, radar, and measurement submodels. | Three key technologies are resolved in their simulation system: general simulation platform, distributed simulation, and data exchange. |

conducted the requirement analyses of next-generation systems [36] and created several modeling methods for CPSs, such as conceptual modeling [37], discrete event modeling [38], multifidelity modeling [39], and distributed simulation for heterogeneous systems [40]. Based on this methodological background, this study develops a practical simulation system with CPS models that are more realistic. The critical points for this study are (1) how to evaluate behaviors of subsystem models graphically for performance analysis and (2) how to convert their communication data for system scalability.

## 4. Proposed Work

This section, firstly, proposes the overall simulation architecture for testing the naval CPSs. Then, we explain the three main components: the CM, physical, and DIS simulators.

*4.1. Overall System Architecture.* Simulation testing is the process of designing and creating a computerized model of an engineered system for conducting tests to evaluate the behavior of the corresponding real system under a given set of conditions [35]. The proposed CPS simulation confirms intrasystem behavior and intersystem communication, which is required for simulation testing. To this end, the simulation architecture in Figure 2 utilizes three types of simulators: a CM simulator for overall simulation activities and physical and DIS simulators to achieve the above-stated requirements.

The CM simulator supports the continuous execution of the information-processing stages in the tactical domain and builds a situational picture of the ship's surroundings [29]. To simulate the overall actions of all the combat entities, it contains simulation engines and model constructions, which will be explained in the following subsection. After simulating the main activities, the simulator sends the simulated data to the relevant physical simulators.

The physical simulators realize their dynamics and behaviors by imitating the console panels of the corresponding system. As the real NCS has various types of physical subsystems, we selected several subsystems developed in this study, as shown in Table 2. The vessels usually use multiple navigation systems, which provides greater accuracy than possible when using any single system. For example, physical simulators control the propulsion, check their own behaviors with navigation simulators, detect threats with sonar and periscope simulators, and check the surroundings with a plotting board simulator. In this study, a periscope simulator was developed for virtual-constructive simulation, which will be explained in the discussion section.

Finally, the DIS simulator performs a gateway role. It links the existing physical simulators to external CPSs or their simulators by generating compatible I/O data on both sides. In other words, it stores received simulation data, converts them to real-world data, and sends the converted data externally.

*4.2. Combat Management Simulator Development.* Figure 3 shows the configuration of the CM simulator. It consists of simulation engines and a simulation model, which means that models can be developed independently of a simulation engine and interfaced with the simulator using an API.

The simulator layer contains platform-level models. The submarine model has various submodels, as well as submarine and target models. The combat-management simulator enables the users to interact with and control the physical systems for the large-scale systems [42]. In the model layer, the space module simulates the physical signals to be transferred within the environment. The acoustic

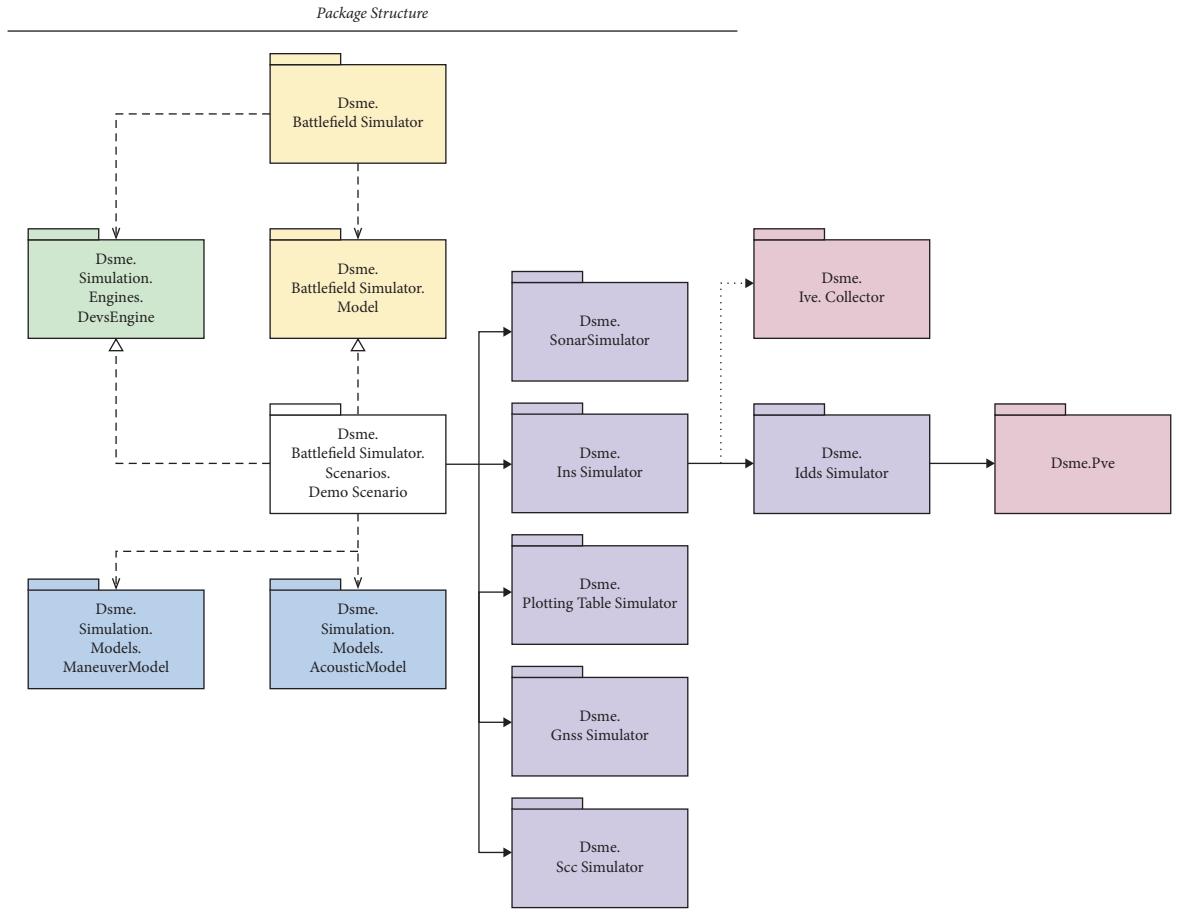



Figure 2: An overall simulation architecture for testing cyber-physical systems.

Table 2: Physical subsystems targeted in this study.

| Physical subsystems | System group | Subsystem role |
|---|---|---|
| Ship steering system | Propulsion system | It controls positional and postural information. |
| GPS | Navigation sensor system | It detects and calculates positional and postural data of own ship on the water. |
| INS | Navigation sensor system | It detects and calculates positional and postural data of own ship underwater. |
| Sonar system | Underwater sensor system | It detects and identifies underwater and surface ship threats. |
| Plotting board | Surveillance system | It checks overall situations with battlefield condition. |
| Periscope | Surface sensor system | It checks sea-level situations. |

module simulates the acoustic signals generated from the engines and propellers of the combat entities.

It is a closed-loop system that repeatedly performs object detection and data processing, makes command decisions, and determines how to respond to actions by appropriately using composite equipment over the course of a war. Each submodel is designed by the system taxonomy. In this study, we used a discrete event system and a discrete time system. For example, the data from the sensor and weapon models are fed into a control system simulation that models the sequence of events, from detection of the threat by various sensors to its engagement by different weapons.

Figure 4 shows the class diagram for the combat-management simulator to implement the concept of Figure 3. To resolve this issue from the software side, we use the MVVM design pattern in the WPF framework. We divide the class diagram into View, ViewModel, and Model horizontally according to the MVVM pattern.

In Model, the combat-management simulator conducts a simulation algorithm and model logic. In View, the user can configure the scenario. The battlefield simulator helps the user to configure new scenarios. In the scenario, the user sets the model structure and values of the model's attributes, observes the simulation's trajectories, and monitors the specified attributes. The user handles the simulation object's spatial information, such as its position and speed. Mutual influences among the subsystems and the nonindependent behavior of the subsystems are evaluated from the battlefield simulator. The combat-management simulator has references for specific scenarios. It only uses interface classes. Thus, it is not affected by changes to the scenario. The modeler specifies which model attributes can be changed in

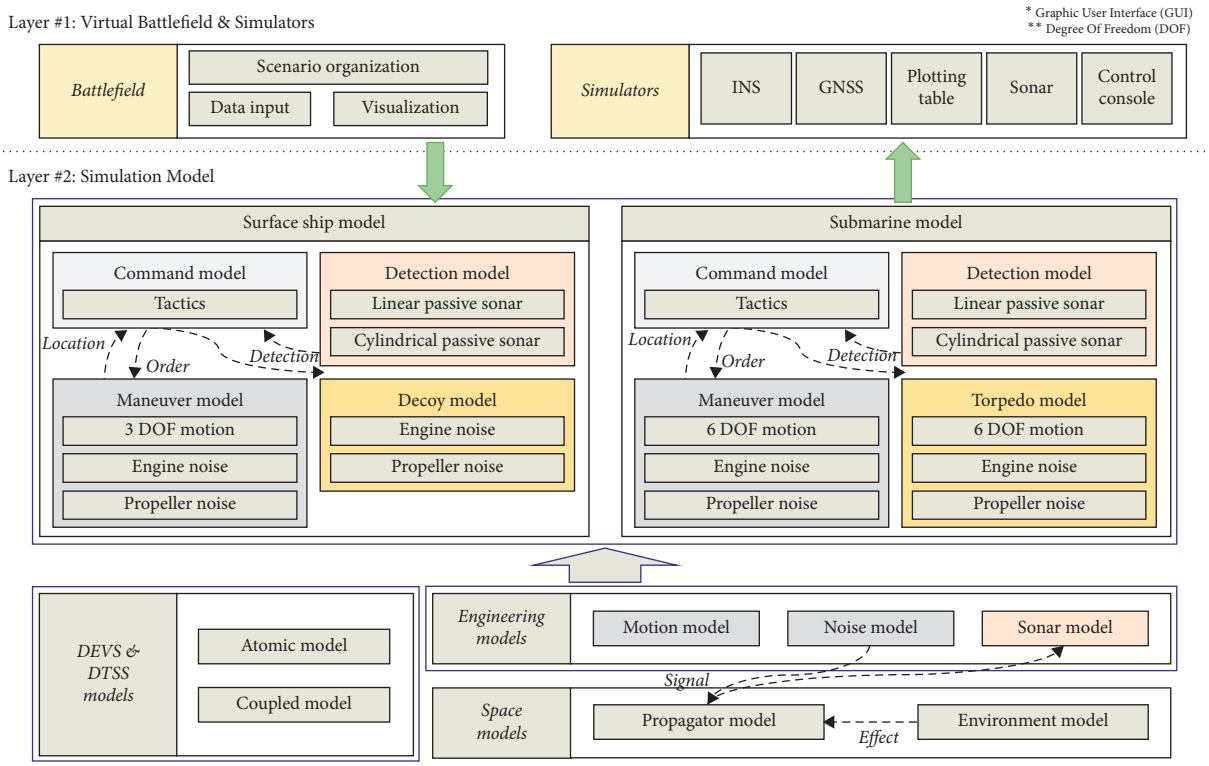



Figure 3: Simulation architecture of combat management simulator [41].

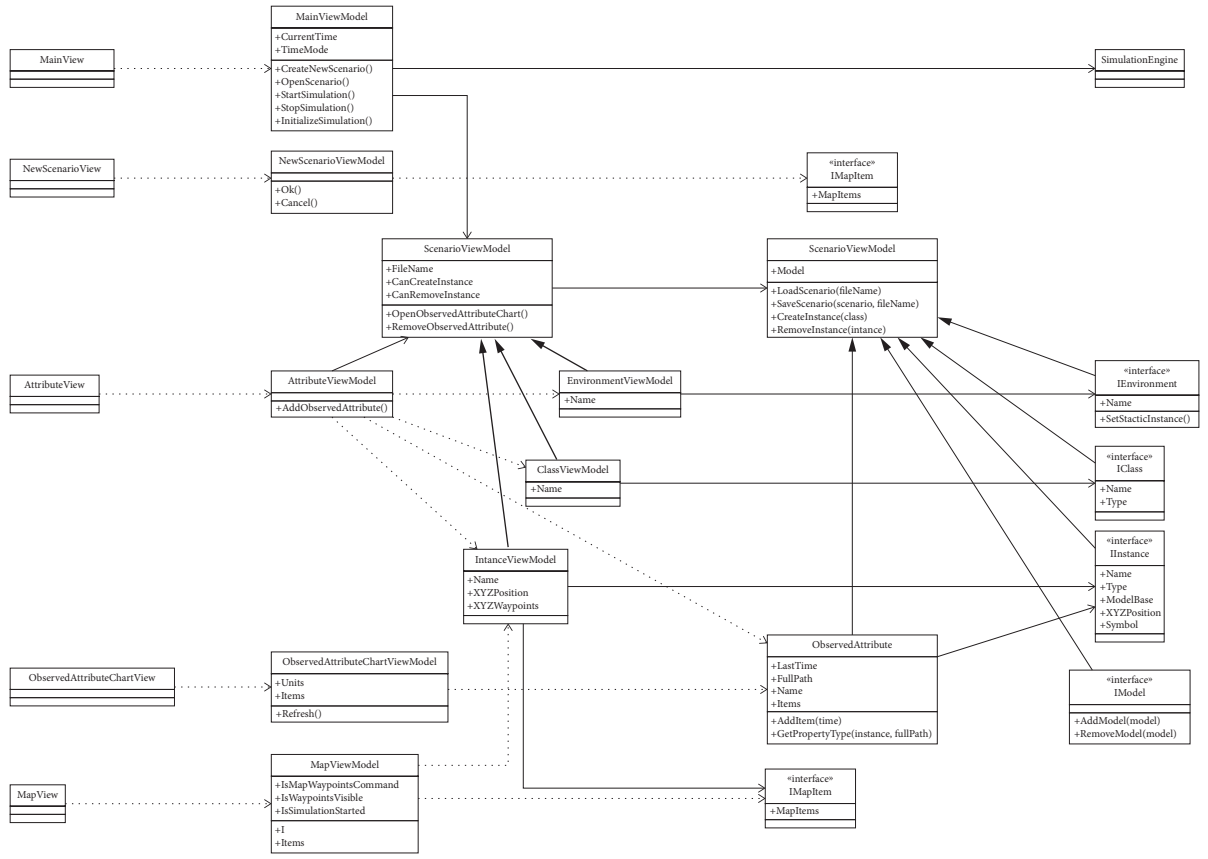



Figure 4: Simplified class diagram of the combat management simulator.

the view and specifies values for those attributes. It is desirable to allow a modeler to change the model's structure in view. The software architecture for the battlefield simulator encapsulates (hides from the outside) their implementations (i.e., their internal behaviors) while providing similar interfaces (i.e., access methods).

*4.3. Physical Simulators Development.* The physical simulators provide physical subsystem aids and data and meteorological information for physical stabilization. Figure 5 shows the underwater sensor simulator in Table 2. Figure 5(a) shows that, firstly, the targets generate noise, and the noise radiates in all directions. The radiated noise propagates far away, undergoing the loss of acoustic energy and addition to background noise. Once the radiated noise arrives at sonars, the sonars analyze the radiated noise. Finally, sonar operators determine whether the noise represents targets (Figure 5(b)).

The sonar equation systematically estimates the expected signal-to-noise ratios for sonar systems. The signal-to-noise ratio determines whether a sonar will detect a signal in the presence of background noise in the ocean. It considers the source level, sound spreading, sound absorption, reflection losses, ambient noise, and receiver characteristics. Such signal-to-noise can be expressed as a passive sonar equation.

$$\mathrm{SE} = \mathrm{SL} - \mathrm{PL} - \mathrm{NL} - \mathrm{BW} + \mathrm{AG} - \mathrm{DT}. \quad (1)$$

The first term, source level (SL), is the magnitude of the radiated noise, whose major sources are engines and propellers. The characteristics of the radiated noise vary, depending on the type and operational conditions of the naval ship, which becomes the clue for the classification of targets. Propagation loss (PL) refers to the loss of acoustic energy that the radiated noise undergoes when propagating. Noise level (NL) is the magnitude of the background noise. Bandwidth (BW) and array gain (AG) are related to signal processing. Sonars analyze the radiated noise by signal processing, and the results of the signal processing are the magnitudes of the radiated noise according to bearings and frequencies. Finally, detection threshold (DT) is the standard for determining whether there are targets.

*4.4. DIS Simulators Development.* The combat system links together various network topologies, treated as subnetworks, in a larger topology. In this case, the DIS is designed to act as a proxy for combat systems residing on the central data bus, which enhances interface compatibility. The DIS simulator verifies that the realized simulated model accurately reflects the authentic behavior of the real system to be tested. The primary role of the DIS simulator is to convert simulation data into a middleware data format.

## 5. Application: Naval Combat Simulation

This section introduces a practical application to maritime warfare. The developed simulation system was used to test the behaviors of the NCS's subsystems and verify data conversion in a specific subsystem.

*5.1. Simulation Scenario.* Modern warfare has two conflicting goals, depending on the side: (1) how the attacker can hit targets accurately, and (2) how the defenders can incapacitate the attackers effectively [36]. For example, during a maritime engagement between a submarine and a surface ship, the submarine fires a torpedo, which sends out sounds and seeks the surface target with echoes. In contrast, the surface ship employs countermeasures to deceive the threat, i.e., the torpedo, and implements evasive maneuvering [36].

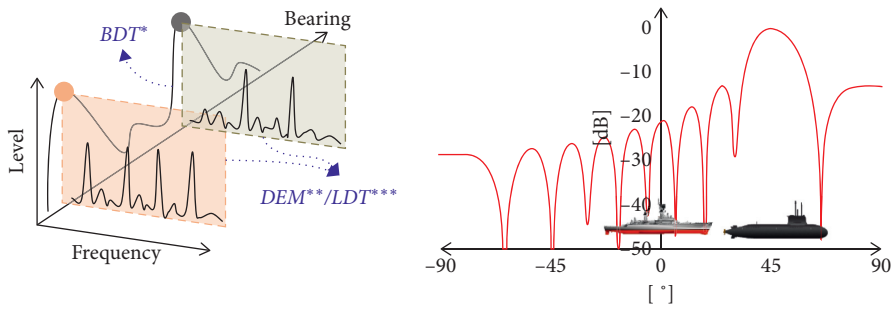
We assumed underwater warfare in this application: a friendly submarine attacks a hostile surface ship with a torpedo. Simulation models for the submarine were built based on the proposed CPS modeling. The CPS models specify when the submarine would launch a torpedo and how it would control the torpedo to hit the target. The simulation models for the surface ship also contained when and how to launch various types of countermeasures according to the defensive tactics used. The simulation would terminate when the surface ship has gone down or when the torpedo is discharged.

*5.2. Simulation Environment.* We developed three types of CPS simulators for the submarine. Five physical simulators, as shown in Figure 6, were distributed across four desktops. The CM and DIS simulators were set up on the same desktop as that of the plotting board simulator. They interacted with the developed physical simulators. For example, the CM simulator computed the overall actions of all of the combat objects. The submarine's actions were decided based on the sonar, GPS, INS, ship steering, and plotting board simulators. The DIS simulator transformed simulation data from the INS simulator into real-world data.

The simulators were developed in .NET framework 4.7 using the language C#. We utilized WCF to realize service-oriented communications between them. As a real INS has a serial interface to connect to external systems, the INS and DIS simulators were connected based on RS-232 serial communications. To retain the consistency of the GUI with the MVVM design pattern, we used DevExpress, which provides best-in-class user-interface controls for WinForms, ASP.NET, and WPF [43].

*5.3. Simulation Execution.* Figures 7–10 show the GUI results of the simulators developed in this application. Figure 7(a) shows a screenshot of the CM simulator, which comprises of five submenus. The main menu (Part I in Figure 7(a)) helps the user to control the overall simulation settings. A user can manage the simulation scenarios, such as by creating, saving, and opening them. When a simulation scenario is chosen, the user constructs the simulation models and configures their variables and parameters via Parts II and III. After completing the configuration, the user can start the simulation and pause, resume, and stop it, if necessary.

Part IV is the main screen to visualize the current simulation situation. It provides all of the trajectories of the combat objects, which are represented with North Atlantic Treaty Organization joint military symbols. In this




(a) (b)

Figure 5: Sonar simulator descriptions: (a) analysis of the radiated noise; (b) sonar operation based on radiated noise.

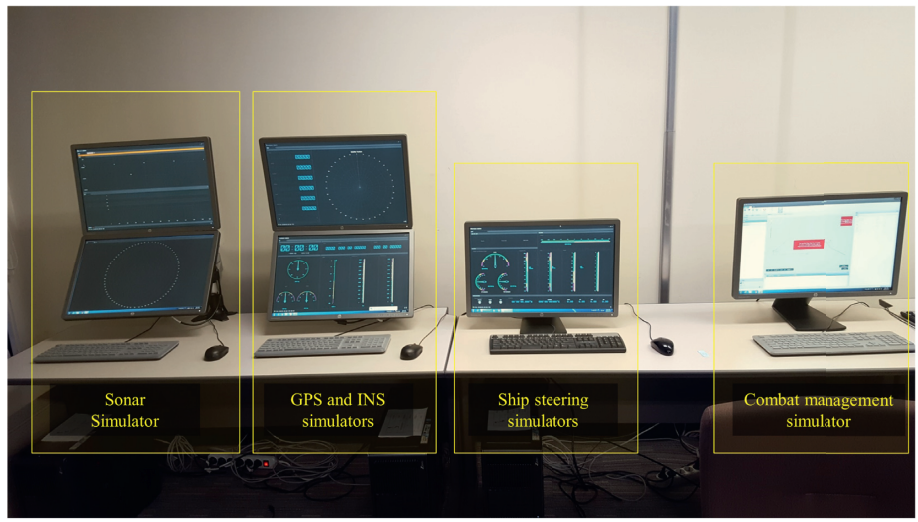
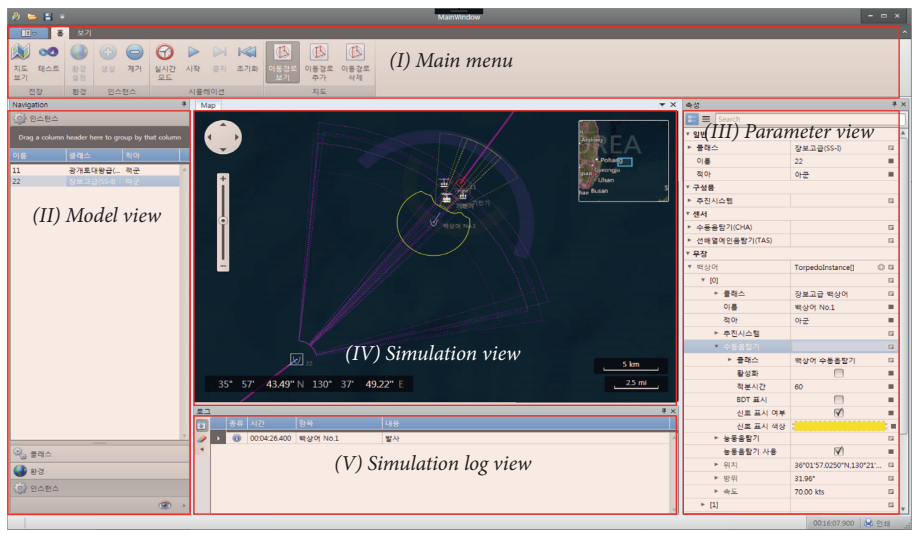
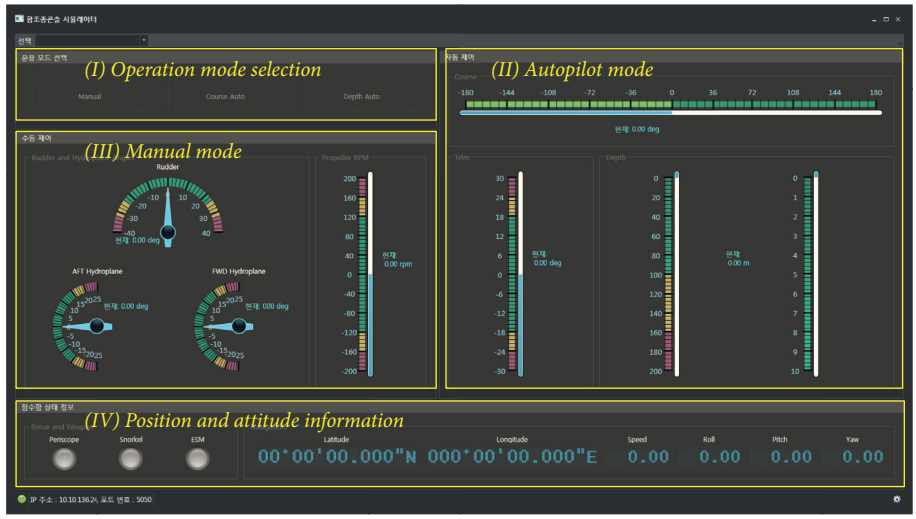



Figure 6: Simulation architecture of combat management simulator.



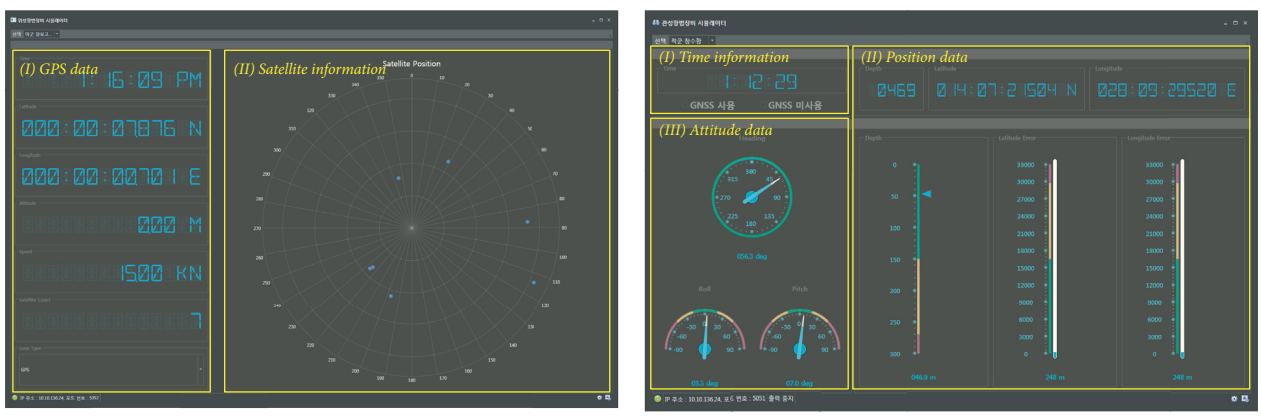

(a)

Figure 7: Continued.



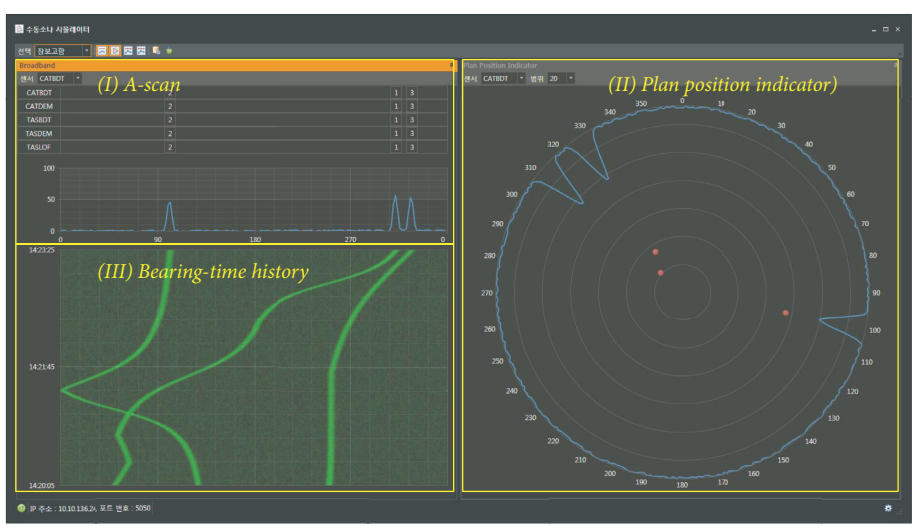

(b)

Figure 7: Graphical user interfaces (GUIs) of the developed simulators: (a) CM simulator; (b) ship steering simulator.



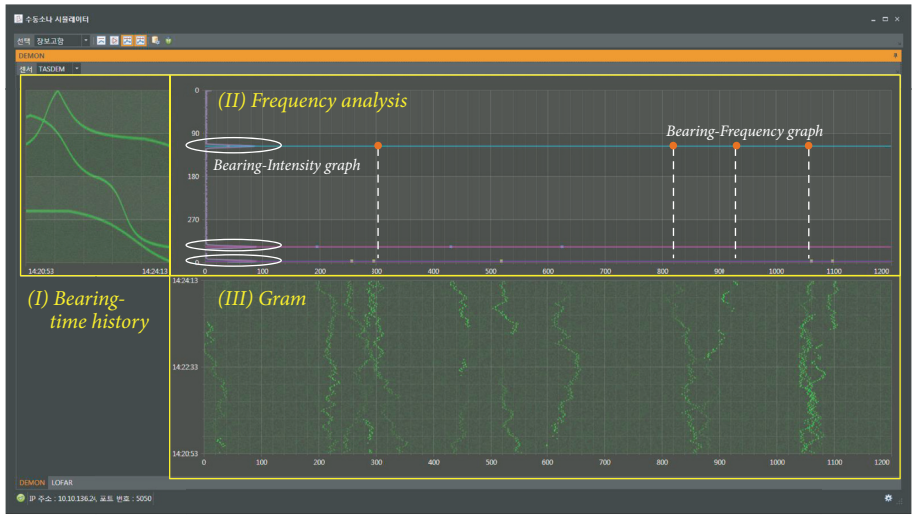

(a) (b)

Figure 8: GUIs of the navigation simulators: (a) global positioning system (GPS) simulator; (b) inertial navigation system (INS) simulator.




(a)

Figure 9: Continued.




(b)

Figure 9: GUIs of the sonar simulator with two views: (a) broadband view; (b) narrowband view.

screenshot, a friendly submarine fires a torpedo at multiple threats. The sensing ranges of the submarine and the torpedo are displayed as purple and yellow fan-shaped sectors, respectively. Part V shows specific events during the simulation, e.g., threat detection, target evaluation, weapon assignment, and weapon launch. As the combat-management simulator provides an integrated situation, the individual behaviors are verified by the physical simulators.

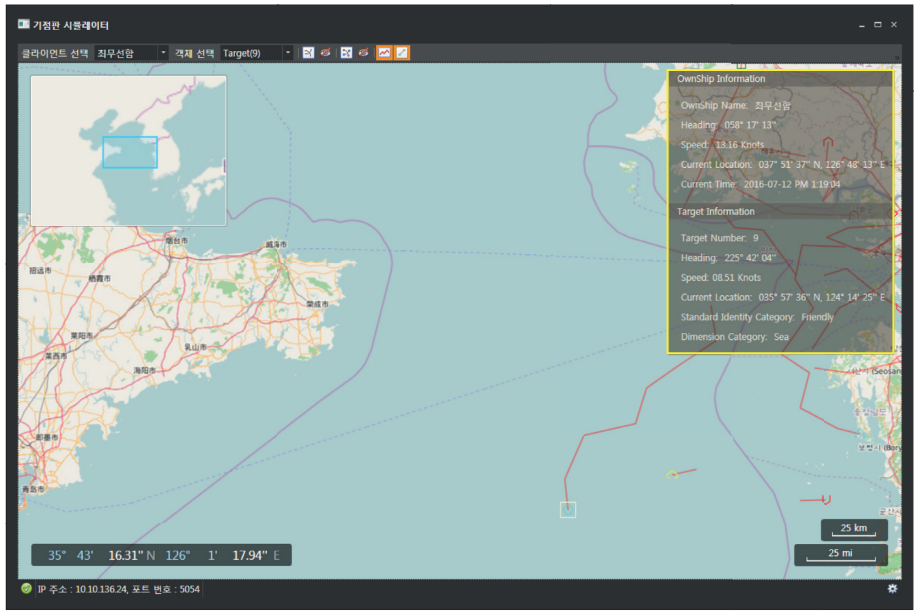
Figure 7(b) is a screenshot of the ship-steering simulator, which shows the six degrees of freedom for the vessel, i.e., latitude, longitude, depth, heading, pitch, and roll. The simulator was developed with autopilot and manual operation modes. The user can select the mode from part I, and the selected mode is displayed and managed via parts II and III. For example, in the manual mode, the user can control a pair of hydroplanes at the forward and afterward sides to change the vessel's horizontal direction and two rudders mounted in the vertical plane to change the lateral movement. The current positional and attitude information is displayed in part IV. All of the views in this simulator are displayed as bar charts.

Figure 8 shows the results from the two navigation sensor simulators. The GPS simulator in Figure 8(a) provides positional data based on satellite information. Part I displays the current latitude, longitude, altitude, and speed with text, and part II illustrates the number of satellites needed to calculate the position. In this application, the navigation data are processed with the global marine navigation protocol, which is a combined electrical and data specification for communication between marine navigation equipment [44].

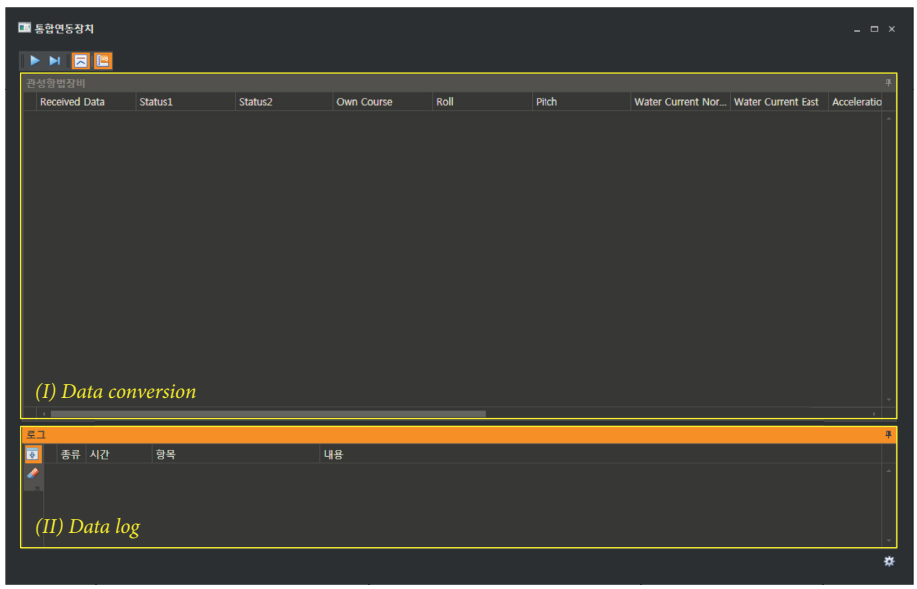
When the vessel is operated underwater, it cannot receive a GPS signal; therefore, the INS mainly is activated, instead of the GPS simulator. In contrast to the GPS simulator, the INS simulator in Figure 8(b) provides textual and graphical information for the vessel's position and altitude (Parts II and III). The INS simulator uses the GPS signal to calibrate navigation inaccuracies when the GPS signal is available. Thus, through Part I, the user can choose whether to use the GPS signal while checking the current simulation time.

Figure 9 provides the screenshots of the sonar simulator, which displays all of the undersea-detection information as broadband and narrowband views. The broadband view in Figure 9(a) contains thee subviews: the bearing diagram, plan-position indicator, and bearing time history. Firstly, the bearing diagram displays the amplitude envelope curve of the signal according to the directed sensors. If a target is detected in a specific sensor, the target is displayed as a raised section and assigned a specific number. Next, the plan-position indicator view (part II) is a type of radar display using a polar coordinate system. In this view, our vessel is in the center, and the detected targets are marked as red points. The view displays the targets' sound levels with a crushed blue circle. In parts I and II of Figure 9(a), three targets—numbered 1, 2, and 3—were detected at about 100°, 310°, and 330° angles, respectively, with the 310° target having the highest sound level. Finally, the bearing time history (part III), referred to as a waterfall diagram, displays the output of the beamforming processor by time. The newest information is at the top of part III, which is matched with the bearing diagram.

Although the broadband view finds the targets' bearings and sound levels, it does not indicate which types of targets have been detected. The target type is specified by the sounds emitted and can be identified from the narrowband view. Figure 9(b) shows two narrowband views for the frequency analysis (II) and gram (III). The frequency analysis, which contains bearing-intensity and bearing-frequency graphs, provides all of the targets' detected sound frequencies. The frequency signals represent vital information about the targets' outward shapes and operations. Similar to the bearing time history in the broadband view, the gram (part III) in the narrowband view displays several sound beams simultaneously for time-series analysis. As shown in




(a)




(b)

Figure 10: GUIs of the developed simulators: (a) plotting board simulator; (b) DIS simulator.

Figure 9(b), twelve beam arrays for three targets were detected and temporarily faded away.

In Figure 10(a), the plotting-board simulator displays various types of information from the platforms currently in operation. The main view shows the overall trajectories of the submarine, surface ship, and torpedoes with different-colored lines. The right part of the main view offers two types of information: (1) numerical values such as latitude, longitude, heading, and speed and (2) status information containing the targets' categories and identification. Finally, Figure 10(b) is a screenshot of the DIS simulator, which provides the real-world data to be converted from simulation data. Thus, it provides increased transparency for the data set, to both the existing and connected systems.

*5.4. Discussion.* The simulation experiments show functional testing results at the subsystem level. The CM simulator simulates real-world combat scenarios, computes overall activity, and sends simulated data to the relevant physical simulators. The five physical simulators display specific behaviors via the user interface, and the DIS simulator shows the data communication between the simulators.

Here, we discuss two further applications for the developed simulators. First, the simulators were utilized to analyze effectiveness and performance. The ultimate goal of NCS is to maximize fighting and defensive strengths based on optimal control of sensor and weapon systems. The NCS should be optimized, and performance should be evaluated

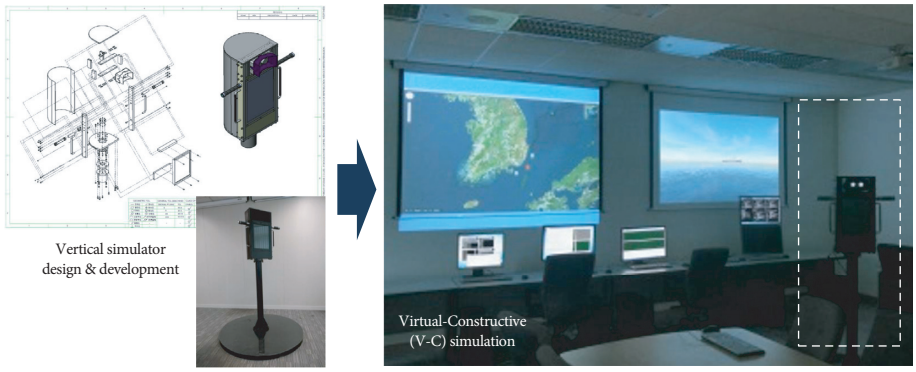



Figure 11: Application of virtual-constructive simulation with periscope simulator.

before a real war is conducted, to measure the system's reliability [42]. Such early feedback during the system-design phase improves to the product's quality by allowing a new understanding of emergent behavior. The proposed simulation can reduce the time it takes to introduce new products, as has been very effectively demonstrated in the automotive industry [41].

Next, the present study supports a testing environment for other studies. It is common practice in military simulations to federate systems at different levels so that they can interoperate to provide their functionalities in a new context. MBSE, a paradigm in which the model becomes the actual software, provides not only a safe and efficient testbed but also provides an immersive environment in which CPSs can learn to adapt and optimize their behavior [45–47]. A build-and-test approach is insufficient because the complexity, cost, and design times become more challenging [48].

Figure 11 shows virtual and constructive simulations, a further application for the developed system [49, 50]. We developed a periscope simulator with HCI interaction. The operator, in turn, interacts with the system or performs these activities, linked with the help of specific HCI panels. The DIS simulator in this advanced simulation produces new real-system data from received simulation data, which extends to live-virtual-constructive simulation.

## 6. Conclusion

In this study, we propose a layered simulation architecture for functional testing that differentiates intrasubsystem behaviors and intersubsystem communications. We designed and modularized three types of simulators. The CM simulator simulates real-world combat scenarios and sends simulated data to the relevant physical simulators. With the simulated data, each physical simulator realizes its dynamics and behaviors by imitating the console panels of its corresponding system. Finally, the DIS simulator performs a gateway role. It links the existing physical models to external CPSs or their models by generating compatible I/O data on both sides.

The proposed *s* increases the automation's transparency, both to the designer and to the user, by separating the domain processing needed to build the system view from the processing. The developed software provides parallel management of the external interfaces to the combat system and internal interfaces between subsystems within the combat system for simulation testing.

Specifically, we discussed several findings, including the balancing of physical and computational abilities as well as the importance of information technology and the statistical trends between the models' inputs and outputs. Models enable simulation and analysis, which can result in earlier identification of design defects than prototyping can. MBSE is now a critical enabler for managing the complexity when developing complex modern systems like CPSs. We believe that the proposed systems also will allow developers and the research community, in general, to better understand open problems and their impact on the broad applicability of model-based design technologies.

## Abbreviations

AG: Array gain
BW: Bandwidth
CM: Combat management
CMS: Combat management system
CPS: Cyber-physical system
CSDB: Combat system data bus
DIS: Data-integration system
DT: Detection threshold
GPS: Global positioning system
GUI: Graphical user interface
HCI: Human-computer interaction
NL: Noise level
I/O: Input/output
ICS-Sim: Integrated combat systems simulation
INS: Inertial navigation system
M&S: Modeling and simulation
MBSE: Model-based systems engineering
MVVM: Model-View-ViewModel
NCS: Naval combat system
PL: Propagation loss
SE: Signal excess
SL: Source level
WPF: Windows Presentation Foundation.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by 2020 Research Fund of Myongji University.

## References

[1] Z. Liu, D.-s. Yang, D. Wen, W.-M. Zhang, and W. Mao, "Cyber-physical-social systems for command and control," *IEEE Intelligent Systems*, vol. 26, no. 4, pp. 92–96, 2011.

[2] V. Gunes, S. Peter, T. Givargis, and F. Vahid, "A survey on concepts, applications, and challenges in cyber-physical systems," *KSII Transactions on Internet & Information Systems*, vol. 8, no. 12, 2014.

[3] Y. Liu, Y. Peng, B. Wang, S. Yao, and Z. Liu, "Review on cyber-physical systems," *IEEE/CAA Journal of Automatica Sinica*, vol. 4, no. 1, pp. 27–40, 2017.

[4] P. Helle, I. Giblett, and P. Levier, "An integrated executable architecture framework for System of Systems development," *The Journal of Defense Modeling and Simulation: Applications, Methodology, Technology*, vol. 10, no. 4, pp. 435–448, 2013.

[5] B. Ge, K. W. Hipel, K. Yang, and Y. Chen, "A novel executable modeling approach for system-of-systems architecture," *IEEE Systems Journal*, vol. 8, no. 1, pp. 4–13, 2014.

[6] N. Guertin, B. Womble, P. Bruhns, D. Schmidt, A. Porter, and B. Antypas, "Management strategies for software infrastructure in large-scale cyber-physical systems for the US Navy," *Cutter IT Journal*, vol. 28, no. 5, 2015.

[7] W. Kang, K. Kapitanova, and S. H. Son, "RDDS: a real-time data distribution service for cyber-physical systems," *IEEE Transactions on Industrial Informatics*, vol. 8, no. 2, pp. 393–405, 2012.

[8] S.-C. Shin, J.-G. Shin, and D.-K. Oh, "Development of data analysis tool for combat system integration," *International Journal of Naval Architecture and Ocean Engineering*, vol. 5, no. 1, pp. 147–160, 2013.

[9] K.-M. Seo and K.-P. Park, "Interface Data Modeling to Detect and Diagnose Intersystem Faults for Designing and Integrating System of Systems," *Complexity*, vol. 2018, Article ID 7081501, 7081501.21 pages, 2018.

[10] I. Akkaya, P. Derler, S. Emoto, and E. A. Lee, "Systems engineering for industrial cyber-physical systems using aspects," *Proceedings of the IEEE*, vol. 104, no. 5, pp. 997–1012, 2016.

[11] B. Xiong, B. Li, R. Fan, Q. Zhou, and W. Li, "Modeling and Simulation for Effectiveness Evalution of Dynamic Discrete Military Supply Chain Networks," *Complexity*, vol. 2017, Article ID 6052037, 6052037.9 pages, 2017.

[12] P. J. Nicholas and D. L. Alderson, "Fast design of wireless mesh networks to defend against worst-case jamming," *Military Operations Research*, vol. 23, no. 3, pp. 5–20, 2018.

[13] P. Derler, E. A. Lee, and A. S. Vincentelli, "Modeling cyber-physical systems," *Proceedings of the IEEE*, vol. 100, no. 1, pp. 13–28, 2012.

[14] Y. Gao, J. Cheng, J. Huang, and L. Quan, "Simulation analysis and experiment of variable-displacement asymmetric axial piston pump," *Applied Sciences*, vol. 7, no. 4, p. 328, 2017.

[15] B. G. Kang, K.-M. Seo, and T. G. Kim, "Model-based design of defense cyber-physical systems to analyze mission effectiveness and network performance," *IEEE Access*, vol. 7, Article ID 42063, 2019.

[16] C. Sung and T. G. Kim, "Collaborative modeling process for development of domain-specific discrete event simulation systems," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 42, no. 4, pp. 532–546, 2012.

[17] X. Li, Y. Lei, H. Vangheluwe, W. Wang, and Q. Li, "Domain-specific decision modelling and statistical analysis for combat system effectiveness simulation," *Journal of Statistical Computation and Simulation*, vol. 84, no. 6, pp. 1261–1279, 2014.

[18] Z. Zhu, Y. Lei, A. Alshareef, H. Sarjoughian, and Y. Zhu, "Domain specific MetaModeling for deep semantic composability," *IEEE Access*, vol. 6, Article ID 18276, 2018.

[19] S. C. Park, Y. Kwon, K. Seong, and J. Pyun, "Simulation framework for small scale engagement," *Computers & Industrial Engineering*, vol. 59, no. 3, pp. 463–472, 2010.

[20] E. Squarcia, V. Gaeta, and F. Ciambra, "4.2. 1 Multi-scale performance evaluation in Naval Combat Systems domain," in *Proceedings of the INCOSE International Symposium*, Wiley Online Library, Chicago, Illinois, USA, July 2010.

[21] X. Xu, G. Li, M. Yang, R. Ju, and K. Huang, "Towards the integration of engagement and engineering-level simulation over real-time and heterogeneous systems," *Concurrency and Computation: Practice and Experience*, vol. 28, no. 12, pp. 3390–3408, 2016.

[22] H.-H. Cheng, Q. Liu, J. Liu, D. Wang, J. Zhou, and J. Li, "Design of key technologies for simulation training system of the multi-element observation system," in *Proceedings of the International Conference on Computer, Electronics and Communication Engineering*, Sanya, China, June 2017.

[23] X. Li, D. Chang, H. Pen, X. Zhang, Y. Liu, and Y. Yao, "Application of MVVM design pattern in MES," in *Proceedings of the Cyber Technology in Automation, Control, and Intelligent Systems (CYBER), 2015 IEEE International Conference on*, June 2015.

[24] A. Weil, *Learn WPF MVVM-XAML, C# and the MVVM Pattern*, Lulu. com, North Carolina, NC, USA, 2016.

[25] F. Jarnjak, "Flexible GUI in robotics applications using windows presentation foundation framework and model view ViewModel pattern," in *Proceedings of the 4th International Conference on New Trends in Information Science and Service Science*, Gyeongju, Korea (South), May 2010.

[26] D. Shan, Dong, Jun et al., "Application OF contemporary programming techniques IN health monitoring software for suspension bridges," *NED University Journal of Research*, vol. 12, no. 2, 2015.

[27] E. Salvador, S. Tidon, and B. E. V. Comendador, "A configurable real-time auto-locked software for electronics company (RT-ASEC)," *Journal of Software*, vol. 10, no. 10, pp. 1170–1175, 2015.

[28] A. K. K, J. George, J. George, and V. P. Jagathyraj, "R equirements analysis of an integrated sonar suite for surface ships: systems engineering perspective," *Defence Science Journal*, vol. 64, no. 4, pp. 366–370, 2014.

[29] T. Kim, D. Lee, and S. Lee, "Hierarchical multilevel on/off source traffic modeling for a warship combat system," *IEEE Journal of Oceanic Engineering*, vol. 39, no. 2, pp. 226–242, 2014.

[30] H. F. R. Arciszewski, T. E. De Greef, and J. H. Van Delft, "Adaptive automation in a naval combat management system," *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans*, vol. 39, no. 6, pp. 1188–1199, 2009.

[31] K.-M. Seo, K.-P. Park, and B.-J. Lee, "Achieving data interoperability of communication interfaces for combat system engineering," *IEEE Access*, vol. 5, Article ID 17938, 2017.

[32] M. R. Khaefi, J.-Y. Im, and D.-S. Kim, "An efficient DDS node discovery scheme for naval combat system," in *Proceedings of the Emerging Technologies & Factory Automation (ETFA), 2015 IEEE 20th Conference on*, September 2015.

[33] M. L. Butterfield, J. S. Pearlman, and S. C. Vickroy, "A system-of-systems engineering GEOSS: architectural approach," *IEEE Systems Journal*, vol. 2, no. 3, pp. 321–332, 2008.

[34] I. Stojmenovic, "Machine-to-machine communications with in-network data aggregation, processing, and actuation for large-scale cyber-physical systems," *IEEE Internet of Things Journal*, vol. 1, no. 2, pp. 122–128, 2014.

[35] A. Engel, *Verification, validation and testing of engineered systems*, John Wiley & Sons, Hoboken, NJ, USA, 2010.

[36] J. H. Kim, K. M. Seo, T. E. Lee, and B. W. Choi, "Achieving new insights into combat engagement analysis via simulation-based sequential experimentation," *Military Operations Research*, vol. 23, no. 4, pp. 51–80, 2018.

[37] K.-M. Seo, W. Hong, and T. G. Kim, "Enhancing model composability and reusability for entity-level combat simulation: a conceptual modeling approach," *SIMULATION*, vol. 93, no. 10, pp. 825–840, 2017.

[38] K.-M. Seo, C. Choi, T. G. Kim, and J. H. Kim, "DEVS-based combat modeling for engagement-level simulation," *SIMULATION*, vol. 90, no. 7, pp. 759–781, 2014.

[39] S. H. Choi, K.-M. Seo, and T. G. Kim, "Accelerated simulation of discrete event dynamic systems via a multi-fidelity modeling framework," *Applied Sciences*, vol. 7, no. 10, 1056.1 pages, 2017.

[40] B. G. Kang, K.-M. Seo, and T. G. Kim, "Communication analysis of network-centric warfare via transformation of system of systems model into integrated system model using neural network," *Complexity*, vol. 2018, Article ID 6201356, 6201356.16 pages, 2018.

[41] D.-H. Jeong, M.-I. Roh, S.-H. Ham, and C.-Y. Lee, "Performance analyses of naval ships based on engineering level of simulation at the initial design stage," *International Journal of Naval Architecture and Ocean Engineering*, vol. 9, no. 4, pp. 446–459, 2017.

[42] T. J. Paciencia, D. J. Richmond, J. J. Schumacher, and W. L. Troy, "A framework and system for theater air attrition modeling," *Military Operations Research*, vol. 23, no. 2, pp. 41–60, 2018.

[43] P. T. Kimmel, *Professional DevExpress ASP. NET Controls*, John Wiley & Sons, Hoboken, NJ, USA, 2010.

[44] B. Park, J. Lee, Y. Kim, H. Yun, and C. Kee, "DGPS enhancement to GPS NMEA output data: DGPS by correction projection to position-domain," *Journal of Navigation*, vol. 66, no. 2, pp. 249–264, 2013.

[45] B. Zeigler, S. Mittal, and M. Traore, "MBSE with/out simulation: state of the art and way forward," *Systems*, vol. 6, no. 4, p. 40, 2018.

[46] S. Mittal and J. L. R. Martín, *Netcentric System of Systems Engineering with DEVS Unified Process*, CRC Press, Florida,FD,USA, 2016.

[47] H. Lu and H. You, "Roadmap modeling and assessment approach for defense technology system of systems," *Applied Sciences*, vol. 8, no. 6, p. 908, 2018.

[48] T. A. Bapty, J. Scott, S. Neema, and R. Owens, "Integrated modeling and simulation for cyberphysical systems extending multi-domain M&S to the design community," in *Proceedings of the Symposium on Model-Driven Approaches for Simulation Engineering*, Virginia, VB, USA, April 2017.

[49] W. I. Jung, L. Lowe, L. Rabelo, G. Lee, and O. Kwon, "A methodology on guiding effectiveness-focused training of the weapon operator using big data and VC simulations," *SAE International Journal of Aerospace*, vol. 10, no. 2, pp. 57–67, 2017.

[50] Q. Qi and F. Tao, "Digital twin and big data towards smart manufacturing and industry 4.0: 360 degree comparison," *IEEE Access*, vol. 6, pp. 3585–3593, 2018.

Hindawi
Complexity
Volume 2021, Article ID 3059865, 15 pages
https://doi.org/10.1155/2021/3059865

*Research Article*

# Optimization of Noise Transfer Path Based on the Composite Panel Acoustic and Modal Contribution Analysis

**Qiang Liu**,[1] **Xuanyi Zhou**,[2] **Jianxin Zhu**,[1,3] **and Xiaoping Gong**[1]

[1]*State Key Laboratory for High Performance Complex Manufacturing, Central South University, Changsha 410100, China*
[2]*Key Laboratory of Special Purpose Equipment and Advanced Processing Technology, Zhejiang University of Technology, Hangzhou, China*
[3]*National Enterprise R&D Center, Sunward Intelligent Equipment Co Ltd, Changsha 410100, China*

Correspondence should be addressed to Xuanyi Zhou; zhouxuanyi@zjut.edu.cn

Received 5 August 2021; Revised 29 September 2021; Accepted 9 October 2021; Published 20 November 2021

Academic Editor: Hang Su



The noise of a cab directly affects the comfort and labor efficiency of the operators. The optimization of the structure-borne transmission path can obviously reduce the cab noise. The method of panel acoustic contribution analysis (PACA) is used to reduce structure noise. However, most studies only consider the panel acoustic contribution of a single frequency, without considering the contribution of major frequencies synthesis to confirm the optimized panels. In this paper, a novel method is proposed based on composite panel acoustic and modal contribution analysis and noise transfer path optimization in a vibro-acoustic model. First, the finite element model (FEM) and the acoustic model are established. Based on the acoustic transfer vector (ATV) method, a composite panel acoustic contribution analysis method is proposed to identify the panels affecting the noise of the field point. Combined with the modal acoustic contribution of the modal acoustic transfer vector (MATV) method, the noise field point is confirmed in the area which has the most significant influence. Second, the optimization algorithm NLOPT which is a nonlinear optimization is applied to design the areas. The noise transfer path optimization with vibroacoustic coupling response can quickly determine the optimal thickness of the panels and reduce low-frequency noise. The effectiveness of the proposed method is applied and verified in an excavator cab. The sound pressure level (SPL) the driver's right ear (DRE) decreased obviously. The acoustic analysis of the composite panel acoustic contribution and modal acoustic contribution can more accurately recognize an optimized area than the traditional PACA. This method can be applied in the optimization of the structure-borne transmission path for construction machinery cab and vehicle body.

## 1. Introduction

Noise, vibration, and harshness (NVH) is usually used to evaluate the drivers' comfort of a vehicle [1, 2]. With the increasing academic attention to the environment and the technical improvement of vibration and noise standards of construction machinery, the NVH performance of construction machinery has been adopted by major manufacturers in recent years. Drawing on the achievement of automotive NVH, analysis methods and the design of construction machinery NVH have been gradually formed with the finite element and boundary element as the core [3, 4].

For analysis and control of vibration and noise in a cab, continuous efforts have been made. Sheng and Tijs used the wavenumber finite and boundary element methods to obtain a numerical solution [5, 6]. Li used the experimental method to obtain a structural modal of the body-in-white of a car, which provides an excellent way to solve similar problems [7, 8]. Mohanty identified that the frequency of structural noise was less than 1 kHz, which belongs to the low and middle frequency range, and can be controlled by suppressing structural vibration [9]. Cheung performed finite element modeling of the vehicle cab after measuring the random excitation. The low-frequency structural noise of the cab was predicted when the contribution of the body panels was analyzed [10]. Hyunwoo and Siano calculated the acoustic response and predicted the noise by inputting excitation [11–13]. Guo and Cheng reduced the structure-

borne noise since the automotive body panel vibration was controlled [14, 15]. Bao analyzed and predicted the sound field inside the cab of the shovel loader using acoustic vibration coupling, and the design was optimized by confirming the panel with the largest acoustic contribution [16]. Long and Wolff analyzed acoustic response, panel contribution volume, and modal participation factor by applying ATV and MATV methods for the cab of a truck, and the panels were confirmed to be improved [17, 18]. Wu established HELS-based near-field acoustic holography (NAH) formulations that can analyze and rank the panel acoustic contributions toward the acoustic field and reconstructed the acoustic field inside the vehicle passenger compartment [19]. At present, the traditional method for PACA aims at the maximum sound pressure level peak at a field point. Then the panels which contribute the most to the maximum sound pressure level are recognized. Finally, the panels are ameliorated to reduce noise. However, in practice, there is often more than one sound pressure peak with small differences in each other. The contributions of a panel are also different at main frequencies. While the maximum peak of the sound pressure is reduced at a frequency, the other one may be increased at another frequency. Therefore, the effectiveness of noise reduction is unsatisfactory. In this study, a method of composite panel acoustic contribution analysis is proposed, which takes the multiple frequencies into consideration.

After recognition of the regions that contribute most to the interior noise, it is necessary to optimize the panels. Han compared three different conditions to affirm the optimum proposal [20]. Liang used structural topology optimization to modify the panels [21]. Most researchers used experiments or topology optimization methods to optimize panels to reduce the interior noise of a cab but rarely used optimization algorithms in construction machinery. Optimization algorithms have been studied widely to solve multiobjective optimization problems, such as the global optimization algorithm [22, 23], the local optimization algorithm without gradient [24–26], the local optimization algorithm based on gradient [27–29], and so on. The optimization of reducing noise in a cab is a constrained, multiobjective, and multivariable problem, which is suitable for the optimization algorithm without gradient. Chan and Kumar studied algorithms in the NLOPT library and selected appropriate algorithm to solve the multiobjective problem [30, 31]. The optimization algorithm in the NLOPT library can meet the requirements.

In summary, the current research mostly focuses on the single-frequency PACA while rarely considers the composite contribution of the multiple frequencies when recognizing the optimized panels. Furthermore, most of the transfer path optimization methods are topology optimization or multischeme experimental comparison, while the ways of combining the optimization algorithm with simulation software are rarely used in a cab. In this paper, the theory and method of composite panel acoustic and modal contribution analysis are constructed. The composite panel acoustic and modal contribution are integrated to determine the improved panels and areas. Then, the transfer path optimization coupling acoustic software, Python language, and algorithm are used to optimize the panels and areas. Finally, the effectiveness of the proposed method is verified in an excavator cab.

## 2. Basis Analysis Theory

*2.1. Composite Panel Acoustic Contribution Analysis Method.* Among the energy transferring from the vibration source to the cab panel, the bending wave is the main fluctuation form. The periodic average power transmitted by a simple harmonic bending wave is expressed as

$$P = V_{\max,n}^2 K_s \omega k^3, \tag{1}$$

where $V_{\max,n}$ is the maximum normal velocity of a thin panel structure, $K_s$ is the bending stiffness, $\omega$ is the angular frequency, and $k$ is the wave number. The vibration energy per unit wavelength of a simple harmonic bending wave can be expressed as

$$E = \frac{1}{2} V_{\max,n}^2 K_s k^4. \tag{2}$$

The transmission of vibration energy within a thin panel can be determined by the maximum normal velocity at different locations.

The wave equation governing sound field is as follows:

$$\nabla^2 p = \frac{1}{v_0^2} \frac{\partial^2 p}{\partial t^2}, \tag{3}$$

where $p$ denotes the sound pressure, $v_0$ denotes the speed of sound, $t$ denotes the time, and $\nabla^2$ denotes the Laplace operator. Refer to the panel theory, the vibration differential equation of the panel structure in the cab is as follows:

$$K_s \nabla^4 u + \rho_s h_s \frac{\partial^2 u}{\partial t^2} = f_s - p, \tag{4}$$

where $\rho_s$ denotes the density of the thin panel, $h_s$ denotes the thickness of the thin panel, and $f_s$ denotes the normal load acting on the surface of the structure per unit area.

The finite element representation is established by applying the structural dynamics equation of the panel with the wave equation and using the Galerkin method.

$$\begin{bmatrix} M_s & 0 \\ \rho_f A^T & M_f \end{bmatrix} \begin{Bmatrix} \ddot{u}_s \\ \ddot{p}_f \end{Bmatrix} + \begin{bmatrix} C_s & 0 \\ 0 & C_f \end{bmatrix} \begin{Bmatrix} \dot{u}_s \\ \dot{p}_f \end{Bmatrix} + \begin{bmatrix} K_s & -A \\ 0 & K_f \end{bmatrix} \begin{Bmatrix} u_s \\ p_f \end{Bmatrix} = \begin{Bmatrix} F_s \\ F_f \end{Bmatrix}, \tag{5}$$

where $u_s$ denotes the displacement column vector of the node; $p_f$ denotes the sound pressure column vector of the node; $M_s$, $C_s$, and $K_s$ are the mass matrix, damping matrix, and stiffness matrix, respectively; $F_s$ denotes the column vector of external load acting on the structure or the sound cavity; the subscripts $s$ and $f$ denote the structure body and the sound cavity, respectively; and $A$ denotes the coupling matrix of sound pressure and displacement. Neglecting the sound absorption materials and damping materials, the acoustic cavity damping matrix and structural damping matrix are zero. Applying Fourier transform, the above equation is transformed as follows:

$$\left(-\omega^2 M_f + i\omega C_f + K_f\right)P_f = -\omega^2 \rho_f A^T U_s + Q_f, \tag{6}$$

where $i = \sqrt{-1}, U_s$, $P_f$, and $Q_f$ are frequency-domain parameters, while $u_s, p_f$, and $F_f$ are the time-domain parameters. Since the acoustic cavity has no load other than the wall panel vibration, the external load $Q_f = 0$.

Let

$$Z = \frac{1}{-\omega^2 M_f + i\omega C_f + K_f}. \tag{7}$$

Then equation (6) is transformed to

$$P_f = i\omega \rho_f A^T Z \dot{U}_s. \tag{8}$$

It can be inferred that controlling the vibration speed of the thin cab panel can affect the vibration energy and acoustic radiation.

ATV illustrates the relation between sound pressure response at a field point and surface vibration velocity.

$$P_{\text{site}} = A_{\text{ATV}}^T(\omega) V_n(\omega), \tag{9}$$

where $P_{\text{site}}$ denotes the sound pressure at a field point, $A_{\text{ATV}}^T(\omega)$ denotes the acoustic transfer vector, and $V_n(\omega)$ denotes the normal vibration velocity of a panel. The acoustic contribution is the algebraic sum of the acoustic contribution of every element in the panel to the field points. The mathematical formula of the element acoustic contribution is as follows:

$$P_i = \sum_{j=1}^{n} P_{i,j} = \sum_{j=1}^{n} \text{ATV}_j \cdot v_n, \tag{10}$$

where $P_{i,j}$ denotes the acoustic contribution of the $i$ th element in the $j$ th panel to the field point.

Based on the structure-acoustic coupling equation, the synthetic sound pressure at the field point within the acoustic cavity is expressed as follows:

$$P_N = \sum_{j=1}^{n} P_{N,j}^{\text{panel}}, \tag{11}$$

where $P_{N,j}^{\text{panel}}$ is the sound pressure generated by panel $j$ at site $N$ and $n$ denotes the number of all relevant panels. To quantify the contribution of the panel vibration to the sound pressure level at the field point, the sound pressure component caused by any panel $j$ at site $N$ is projected in the direction of the complex vector of the combined sound pressure to obtain the conventional contribution of the panel to the synthetic sound pressure.

$$(P_c)_{N,j} = \frac{P_{N,j}^{\text{panel}} \cdot P_N}{|P_N|}. \tag{12}$$

The composite panel acoustic field contribution analysis takes all sound pressure peaks into consideration, i.e., the same importance is given to all sound pressure peaks. According to the contributions to the overall sound pressure level of the field point, it is necessary to give different weighting to the sound pressure peaks, and the correction coefficient $\beta$ is introduced and defined as follows:

$$\beta = \frac{P_i}{P_{\text{site}}^{\text{RMS}}}, \tag{13}$$

where $P_i$ is the peak sound pressure of interest frequency and $P_{\text{site}}^{\text{RMS}}$ is the root mean square (RMS) value, and the composite panel acoustic contribution is obtained as follows:

$$(P_c)_{N,j}^{\text{sum}} = \sum_{i=1}^{l} \beta \cdot (P_c)_{N,j}^{\text{peak}}. \tag{14}$$

*2.2. Modal Contribution Analysis Method.* The modal acoustic transfer vector illustrates the relationship between the sound pressure at the field point and the modal participation factor, which takes into account the structural modal on the basis of ATV. In accordance with the vibration displacement vector of the surface, the normal vibration velocity vector is projected to the normal direction of the surface.

$$V_n = j\omega \cdot \Phi_n \cdot \text{MRSP}(\omega), \tag{15}$$

where $\Phi_n$ denotes the matrix of structural modal and $\text{MRSP}(\omega)$ denotes the matrix of modal participation factors. The sound pressure at the field point can be calculated as

$$\begin{aligned} p &= \{\text{ATV}(\omega)\}^T \times j\omega \times \Phi_n \text{MRSP}(\omega) \\ &= \{\text{MATV}(\omega)\}^T \times \text{MRSP}(\omega), \end{aligned} \tag{16}$$

where $\text{MATV}(\omega)$ denotes the modal acoustic transfer vector and can be calculated as

$$\{\text{MATV}(\omega)\}^T = j\omega \cdot \{\text{ATV}(\omega)\}^T \cdot \Phi_n. \tag{17}$$

The modal contribution of each frequency is calculated by extracting the modal matrix to recognize the optimized area.

## 3. Composite Panel Acoustic and Modal Contribution Analysis Method

This paper uses a composite panel acoustic and modal contribution analysis method to recognize the optimized area and then uses a noise transfer path optimization with vibroacoustic coupling response to determine the optimal area to reduce low-frequency noise.

Figure 1 shows the process of composite panel acoustic and modal contribution analysis. First, the FEM and the acoustic model of a cab are established and coupled. The structure model and the acoustic cavity model are calculated for the modal participation factor. Based on the vibroacoustic coupling model and external excitation, the SPL at the field point is calculated. The correctness of the vibroacoustic model is verified by comparison with experimental and simulation results. Second, based on the ATV method, the composite panel acoustic contribution analysis method is proposed to identify the main panels affecting the field point. Composite panel acoustic contribution analysis gives different weighting to the sound pressure peaks. Combined with the modal acoustic contribution of the MATV method, the areas which have the largest influence noise at the field point are confirmed in panels. Finally, the noise transfer path optimization with vibroacoustic coupling response analysis can determine the panel thickness and reduce low-frequency noise. The optimization algorithm NLOPT is applied to design the areas.

## 4. Application and Verification

This paper takes an excavator cab as the application target of the new method. Excavators are the most widely applied equipment in construction. For the purpose of reducing the low-frequency noise caused by panel vibration in an excavator cab, the composite panel acoustic and modal contribution analysis and noise transfer path optimization are applied. The optimization results are verified by an experiment.

### 4.1. Numerical Model

*4.1.1. The Cab Structural Model.* The excavator cab is mainly made of the welded panel and shell structure. Some structural parts such as small holes, chamfered corners, and tabs that have less impact on the cab model are applied to simplify the process. The material of the body in white is Q235 with 210 GPa elastic modulus, 7850 kg/m$^3$ density, and 0.3 Poisson's ratio. The glass is added around the FEM of the body in white to constitute an enclosed cab. The elastic modulus of the glass is 72 GPa, density is 2500 kg/m$^3$, and the Poisson's ratio is 0.22. The cab is meshed with an average mesh cell size of 10 mm. Figure 2(a) shows the FEM of the cab with a cell number of 324072 and a node number of 319817. According to the procedure of solving the structural model of the cab, the first four structural modes of the enclosed cab are obtained, as shown in Figure 3. Most large displacements in modal shapes are concentrated in the glasses, which demonstrates that the glasses are prone to be excited vibration.

*4.1.2. The Cab Acoustic Model.* The acoustic model is established on the basis of the structural model. The material of the acoustic cavity is air with 340 m/s velocity and 1.22 kg/m$^3$ density. The frequency range caused by the panel vibration is mainly 20–200 Hz. To improve simulation accuracy, the length of the acoustic grid should meet the requirements $L \le c_0/6f_{\max}$, where $L$ denotes the length of the divided acoustic grid cell, $c_0$ denotes the sound velocity, and $f_{\max}$ denotes the maximum frequency. The maximum length of the acoustic grid is 283.33 mm. The length of the acoustic grid is set as 40 mm. Moreover, the acoustic model is obtained with 9453 elements, as shown in Figure 2(b). Figure 4 shows the first four acoustic modes of the excavator cab. The structural modal and acoustic modal can be imported to the vibroacoustic model for acoustic response analysis.

*4.1.3. The Cab Vibroacoustic Model.* The structural model of the enclosed cab is first inputted into the simulation software. Then the acoustic cavity model also is imported. The field points are set up in the acoustic cavity elements. The excitation points are created under the cab. Finally, the structure model and acoustic cavity model have achieved mutual mapping by way of establishing the coupling face. The vibroacoustic coupling model of the excavator cab is established, as shown in Figure 5.

### 4.2. Numerical Analysis

*4.2.1. Baseline Measurements.* Figure 6 shows the process of data acquisition in the excavator cab. The vibration acceleration in the $Z$ direction of the cab suspension and the corresponding one-third octave right ear sound pressure of a medium excavator at idle speed (1050 rpm) are measured as input to the panel acoustic contribution analysis, as shown in Figures 7 and 8, respectively. A low-frequency band, mainly 20–200 Hz, possesses the main energy of sound pressure. The frequencies corresponding to the sound pressure peaks are 50 Hz, 100 Hz, and so on.

*4.2.2. Acoustic Response Analysis.* By acoustic response analysis, the sound pressure at the field point connects with the elements in FEM. The acceleration excitation of the suspension is imported to the vibroacoustic coupling model. The frequency range of the simulation is set as 20–200 Hz, and the step size is 1 Hz. After the simulation calculation, Figure 9 shows the comparison simulation results with experimental results about SPL of DRE.

In the frequency range of 20–200 Hz, the tendency of the simulation curve and the experiment curve is roughly accordant and the frequency corresponding to the peak also coincides. It shows that the simulation results can be reflected the acoustic response and verified the accuracy of the coupling model.

*4.3. Composite Panel Acoustic Contribution Analysis.* The excavator cab is divided into 12 panels in the acoustic model, and the panel numbers and the corresponding panel names are shown in Table 1.

As an example, Figures 10 and 11 show the panel acoustic contribution at 50 Hz and 100 Hz, respectively. The panels have a great difference in contribution at diverse

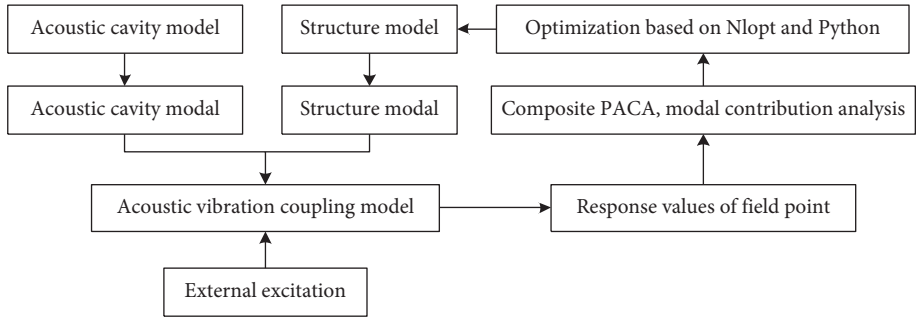
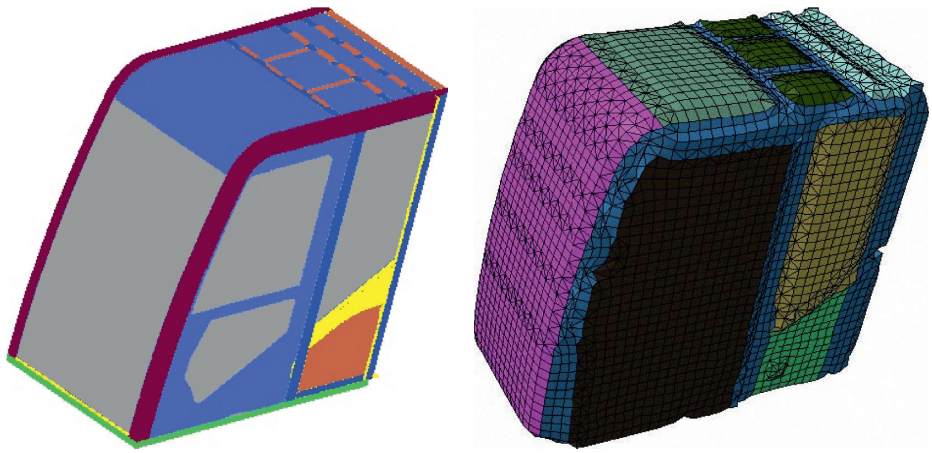



Figure 1: The process of composite panel acoustic and modal contribution analysis.



(a) (b)

Figure 2: The structural and acoustic model of an enclosed cab. (a) The structural model. (b) The acoustic model.

frequencies, or even opposite contributions. Hence, it is important to consider the composite contribution of the main frequencies. The composite panel acoustic contribution analysis for the partitioned panels is calculated in accordance with equations (13) and (14). The histogram is shown in Figure 12. From the figure, it can be seen that panels 1, 3, 4, 5, 6, 7, and 12 have a positive contribution to the overall SPL at the field point and the rest of the panels have a negative contribution. Among the panels contributing positively, panels with a contribution degree greater than 0.05 are the right glass, door glass, and bottom panel, which need to be optimized.

*4.4. Modal Contribution Analysis.* The enclosed structure modal obtained from the MSC Nastran solution is imported into the acoustic software. Moreover, the calculation results of ATV are also imported to map the structure modal to the acoustic mesh, and the modal participation factor is calculated. Then, the MATV matrix is calculated using ATV as the modal transfer matrix and combining the modal participation factors. Afterward, the MATV matrix is extracted by combining the frequency corresponding to sound pressure peaks in the DRE sound pressure response curve. According to Figure 8, the frequency of the higher peak value of sound pressure is 50 Hz and 100 Hz. Figures 13 and 14 show the value of the modal contribution of the cab modal at 50 Hz and 100 Hz, respectively.

The top five-mode orders in modal contributions at 50 Hz are $3^{rd}$, $4^{th}$, $6^{th}$, $36^{th}$, and $40^{th}$ orders, while the top five-mode orders in modal contributions at 100 Hz are the $3^{rd}$, $4^{th}$, $6^{th}$, $23^{rd}$, and $25^{th}$ orders. The related mode vibration patterns are shown in Table 2, and the representative mode vibration patterns are shown in Figures 15, 16, and 17. From the representative modal vibration patterns, it can be concluded that the regions with greater influence on the peak sound pressure corresponding to 50 Hz and 100 Hz are mainly the center of the door glass, the center of the right glass, and the center of the bottom panel.

### *4.5. Structure Optimization and Verification*

*4.5.1. Structure Optimization.* The method of composite panel acoustic and modal contribution analysis identified the key locations for improvement i.e. the door glass, the right side glass, and the bottom panel. The optimization method of the noise transfer path is adopted to reduce cab noise. Since the upper part of the bottom panel is the seat and the bottom is the hydraulic pipeline, it is not easy to change the structure. Thus, the thickness optimization of the center of the bottom panel is not considered. The structure of the excavator cab is optimized

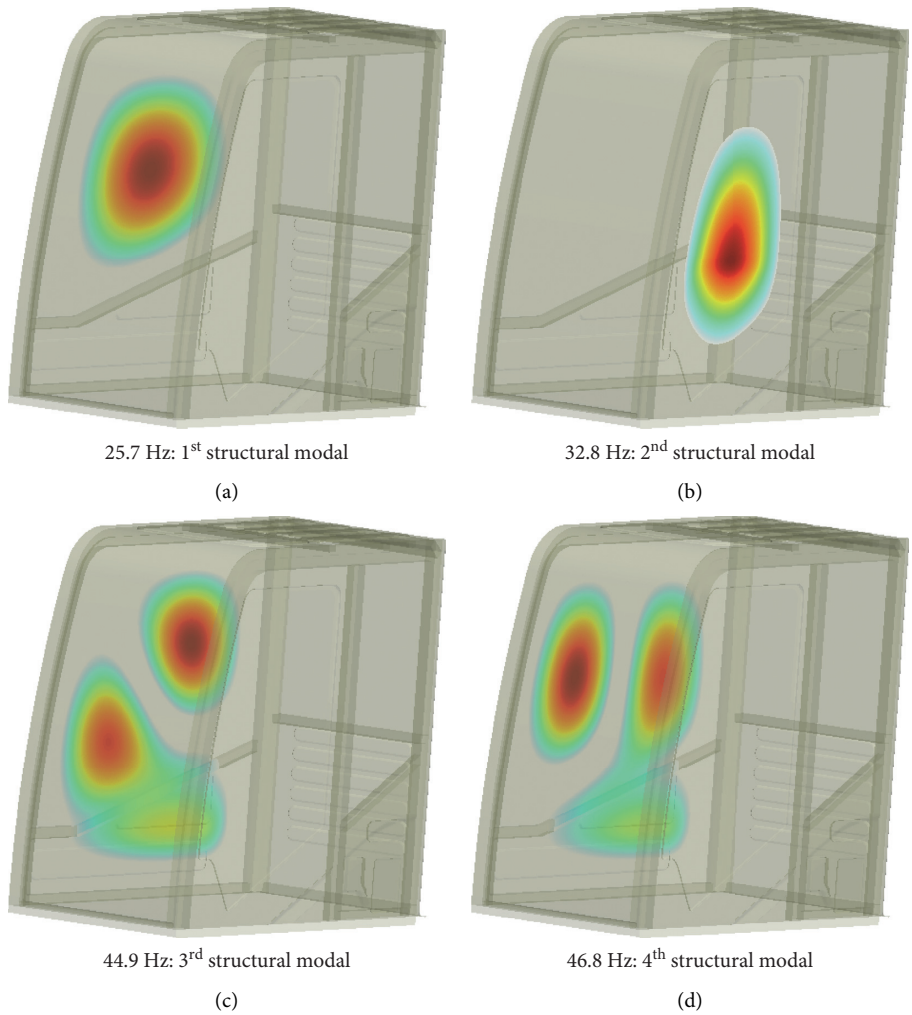



Figure 3: The structural modal of an enclosed cab. (a) 25.7 Hz: 1st structural modal. (b) 32.8 Hz: 2nd structural modal. (c) 44.9 Hz: 3rd structural modal. (d) 46.8 Hz: 4th structural modal.

using the NLOPT optimization program combining with the acoustic software. The NLOPT script is shown in Figure 18. The left side door glass center and the right glass center can be optimized. NLOPT involves various kinds of optimization algorithms with a uniform interface; the derivative-free optimization with LN_COBYLA function is suitable for the accuracy and speed of the optimization.

Applying the same boundary conditions and excitation conditions to the optimized model, the optimization program is terminated after 20 iterations. Moreover, an optimal design is found that the total weight of the whole board is less than 35 kg, and the sound pressure level is low. After the optimization, the thickness of the glass on the right is 0.00566 m, and the thickness of the left side door glass is 0.0062 m. The optimized sound response in the cab is analyzed. Figure 19 shows the optimized DRE sound pressure curve and the original DRE sound pressure curve. It can be seen that the SPL at the frequency of 100 Hz is reduced from 53.06 dB to 51.36 dB. Although the SPL in the frequency range from 120 Hz to 200 Hz increased, the total sound pressure has been reduced by 2.3 dB, which indicates the direction of noise reduction for the cab.

*4.5.2. Experimental Verification.* According to the results of the structure optimization, it can be concluded that the driver's right ear noise can be reduced by increasing the thickness of the right glass and left side door glass.

The manual scheme is experimented by adding a mass block on the right glass and left side door glass, as shown in Figures 20 and 21.

When the right ear noise is collected under the 1050 rpm working condition, compared with baseline measurement results, as shown in Figure 22, the sound pressure level is reduced by 3.18 dB in the frequency band of 20 Hz–200 Hz, and the overall sound pressure level is reduced by 2.25 dB. The gap is because the sound pressure level of 315 Hz and 400 Hz and other medium and high-frequency noise increases, which leads to the noise level reduction of the whole frequency band. The experiment proves that the low-frequency noise can be reduced by thickening the panels.

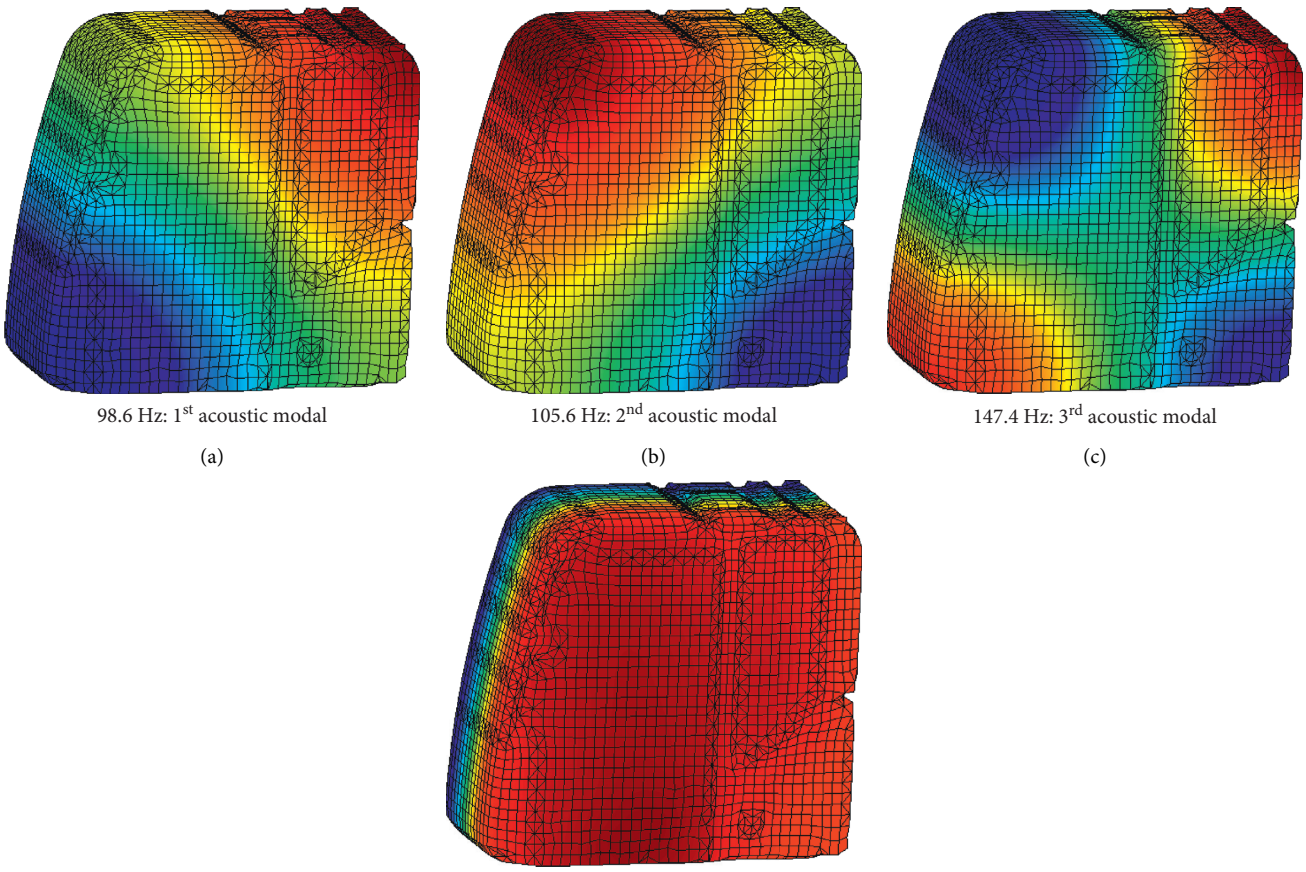



Figure 4: The acoustic modal of the cab. (a) 98.6 Hz: 1st acoustic modal. (b) 105.6 Hz: 2nd acoustic modal. (c) 147.4 Hz: 3rd acoustic modal. (d) 184.7 Hz: 4th acoustic modal.

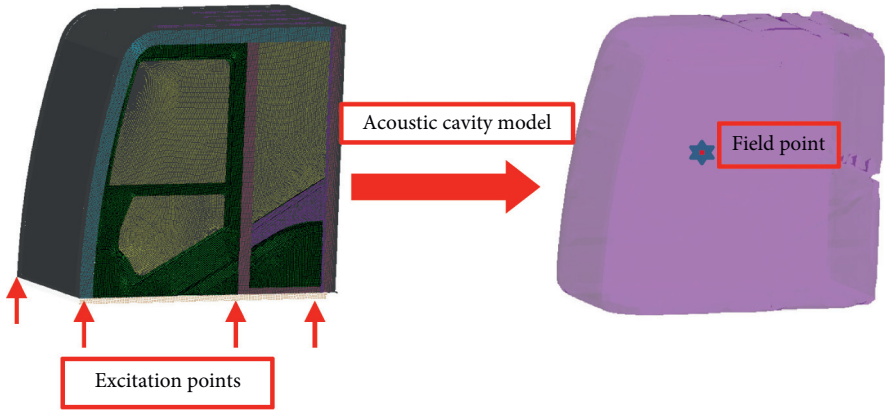



Figure 5: Cab vibroacoustic coupling model.

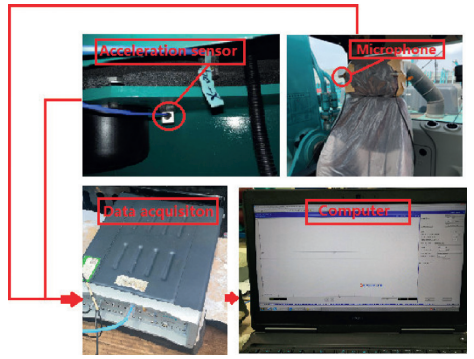



Figure 6: Baseline measurements for vibration and noise.

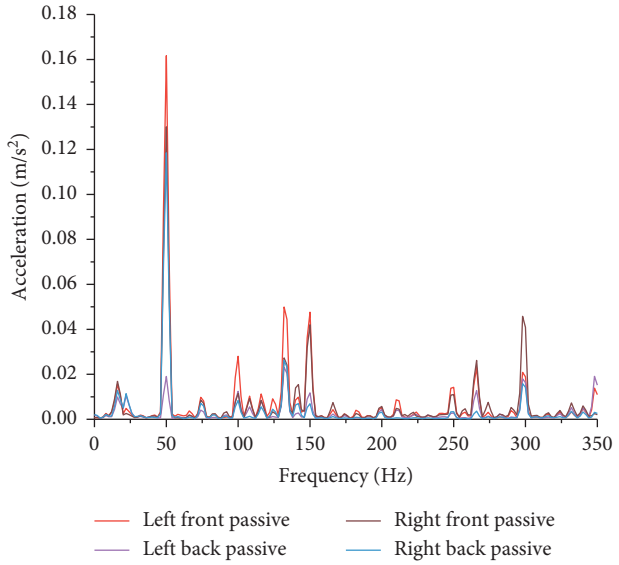



Figure 7: Acceleration excitation at 1050 rpm.

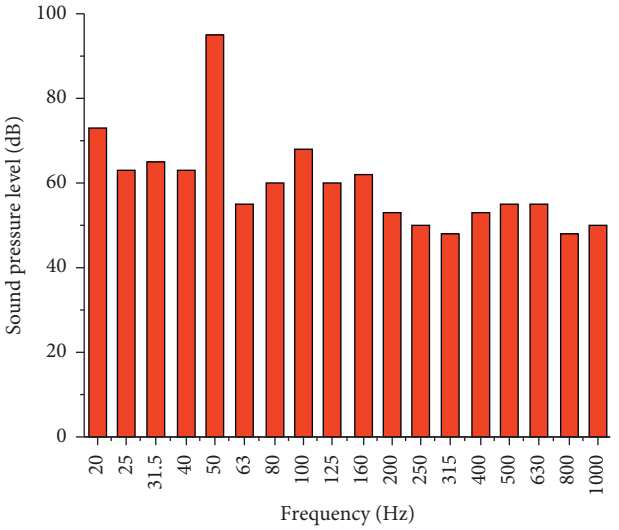



Figure 8: 1/3 octave band at the right ear of the cab.

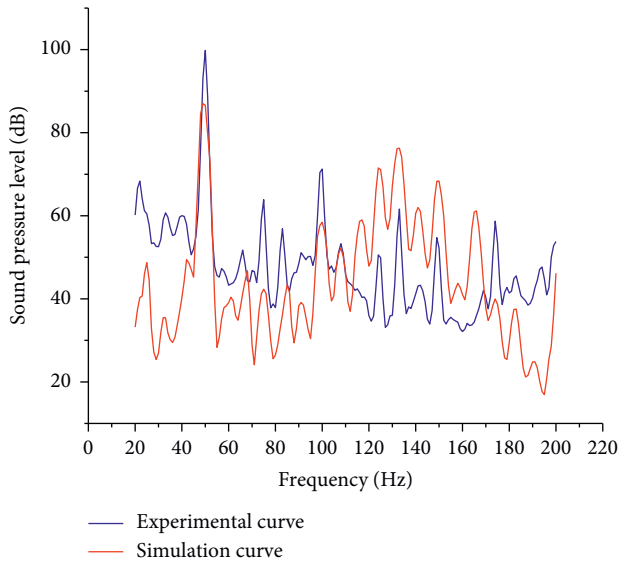



Figure 9: Simulation and experimental SPL curve.

Table 1: Cab panel division.

| Panel number | Panel name | Panel number | Panel name |
|---|---|---|---|
| 1 | Right side glass | 7 | Left side door glass |
| 2 | Right fence panel | 8 | Left rear top glass |
| 3 | Front glass | 9 | Left rear bottom glass |
| 4 | Roof front glass | 10 | Rear side glass |
| 5 | Central roof enclosure | 11 | Rear side panel |
| 6 | Roof rear enclosure | 12 | Base panel |

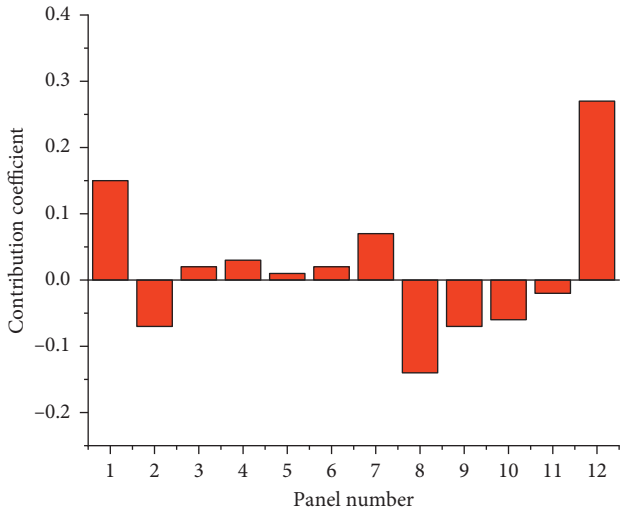



Figure 10: The panel acoustic contribution at 50 Hz.

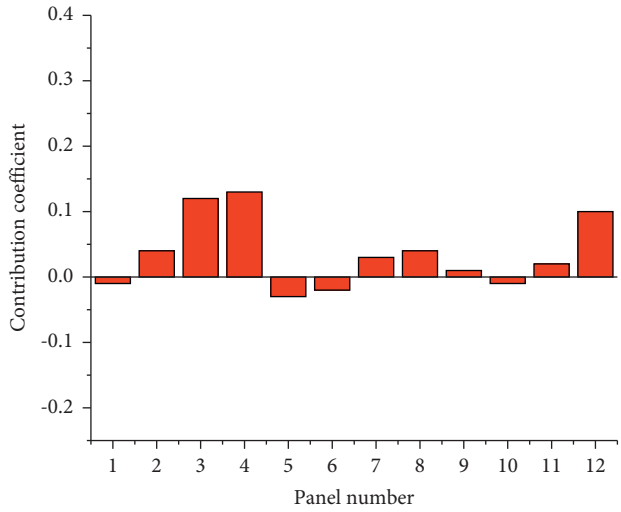



Figure 11: The panel acoustic contribution at 100 Hz.

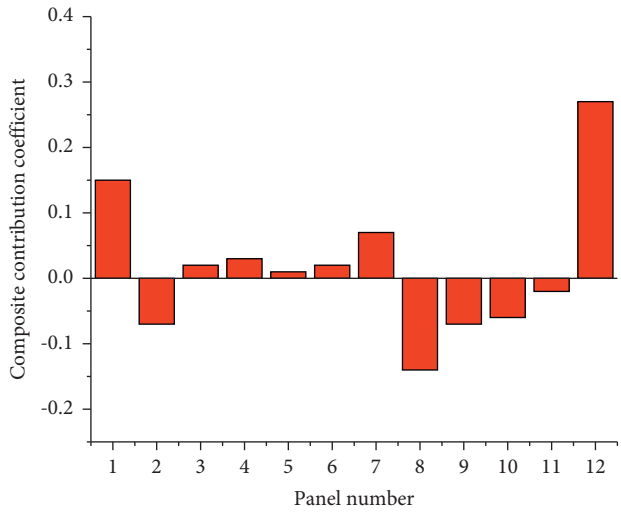



Figure 12: The composite panel acoustic contribution.

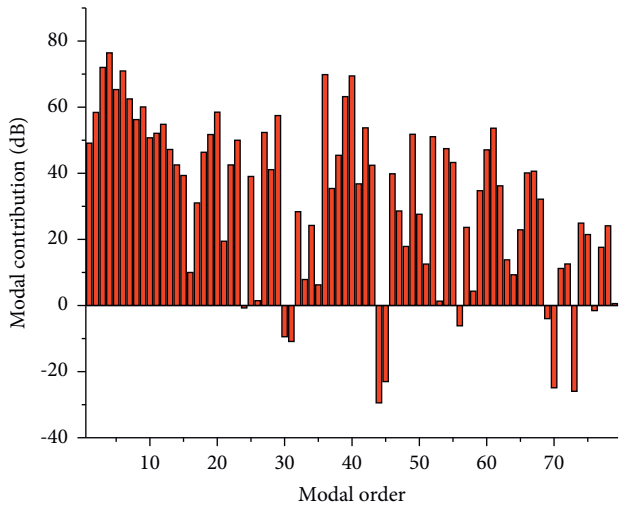



Figure 13: Modal contribution of the cab at 50 Hz.

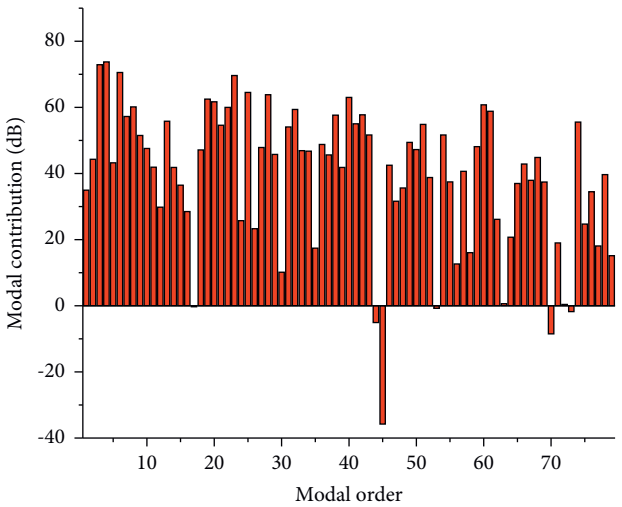



Figure 14: Modal contribution of the cab at 100 Hz.

Table 2: Manifestation of the relevant mode vibration pattern.

| Modal order | Modal vibration expression |
|---|---|
| 3 | Base panel |
| 4 | Right side glass |
| 6 | Door panel glass |
| 23 | Front glass |
| 25 | Door panel glass |
| 36 | Front glass, roof glass |
| 40 | Roof glass |

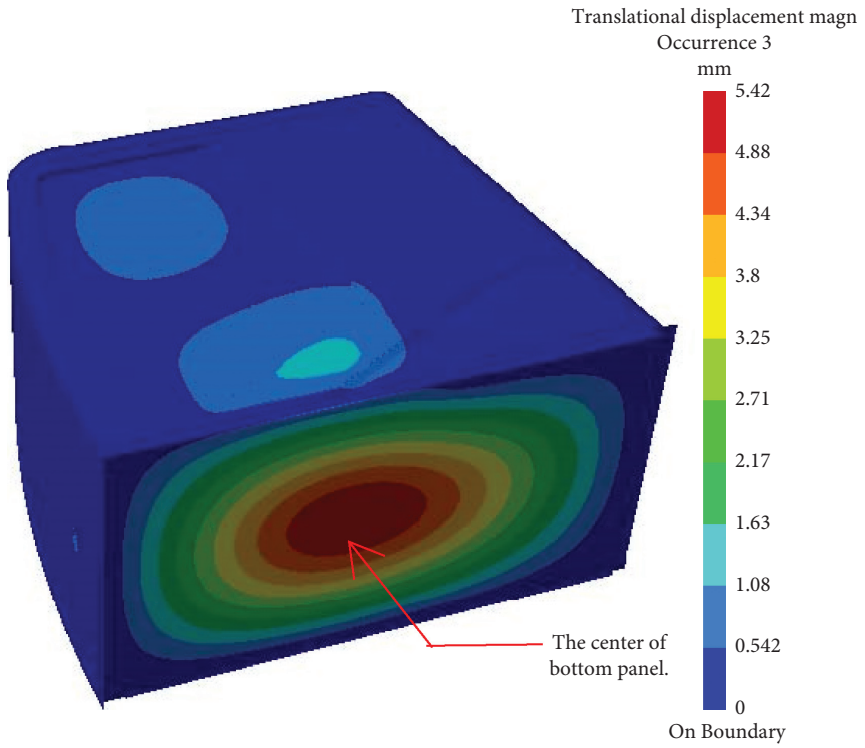



Figure 15: Third-order vibroacoustic coupling modal.

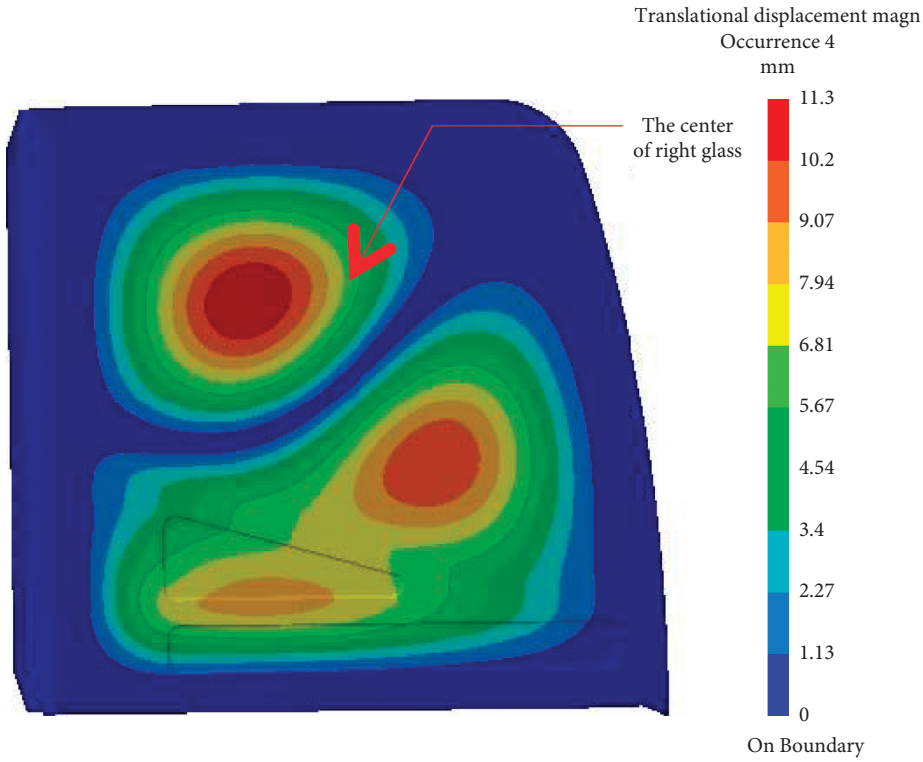



Figure 16: Fourth-order vibroacoustic coupling modal.

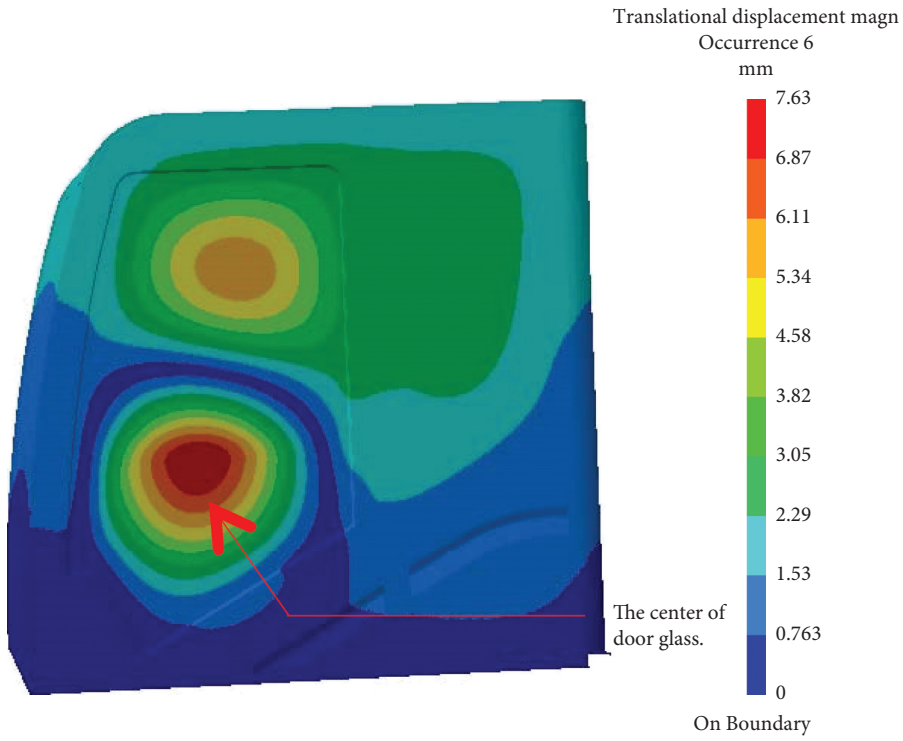



Figure 17: Sixth-order vibroacoustic coupling modal.

```
#USER INPUT: Actran main input file
actran_MFR_file='coupled_MFR.edat'
structure_modal_file='structure_mode.edat'

#USER INPUT: number of iteration
n_iteration = 20

#USER_INPUT: the initial value
initial_thickness =[5,5]

#USER_INPUT: lower and upper band for the
variables
lower_band=[3 , 3]
upper_band=[7 , 7]

#USER_INPUT: mass limit
mass_limit = 35
```

Figure 18: Part of the optimized program script.

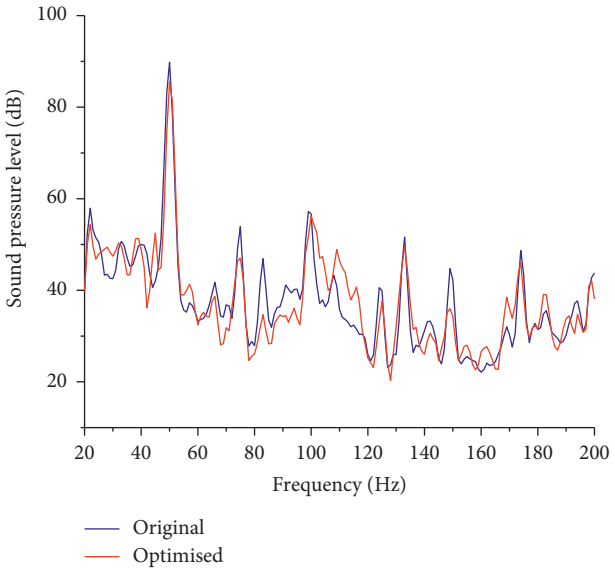



Figure 19: Original and optimized sound response curve.

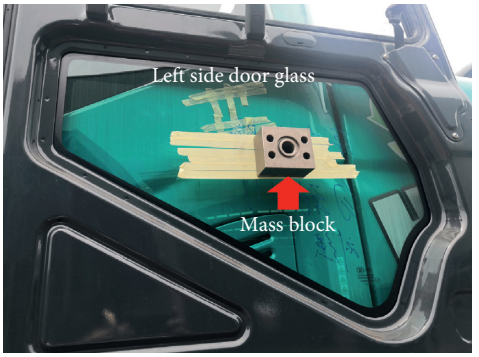



Figure 20: Manual verification on the left side door glass.


Figure 21: Manual verification on the right glass.

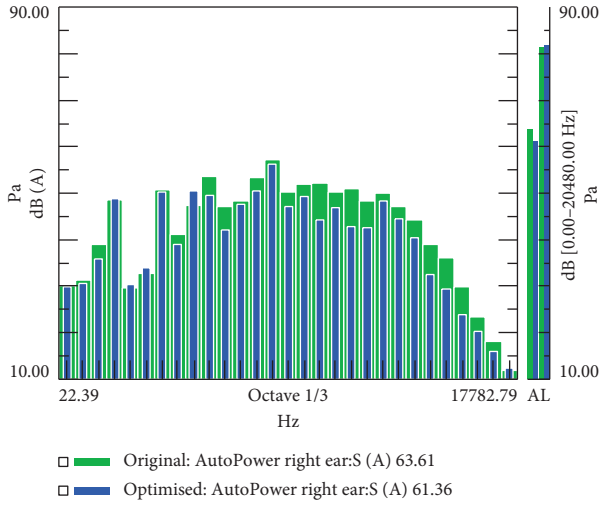



Figure 22: Comparison of right-ear noise experiment values of the cab before and after optimization.

## 5. Conclusion

Reducing the vibration noise usually achieved by improving the structure, but how to quickly and accurately identify and optimize the structure with a large contribution, is a challenge. In this paper, a method of noise transfer path optimization based on the composite panel acoustic and modal contribution analysis is proposed. The acoustic transfer vector method is used to analyze the acoustic characteristics in the vibroacoustic model. To avoid the lack of considering a single frequency contribution in the traditional panel acoustic contribution analysis, the composite panel acoustic contribution analysis that takes the multiple frequencies into consideration is proposed. Synthesizing panel contribution and modal contribution results, the panels and areas that need to be optimized identified. The parameter optimization method coupling acoustic software, Python language, and algorithm is adopted for the target panel and areas. The method is verification in an excavator cab. The acoustic response of the excavator cab is consistent with the experimental sound pressure curve at the field point. By optimizing the thickness of the right glass and left side door glass, the overall sound pressure level at DRE is reduced by 2.25 dB. The experimental result verifies the feasibility of analysis and optimization of the noise based on the composite panel contribution and modal acoustic contribution methods, which provide a reference for passive noise control in a cab.

## Data Availability

The data used to support the findings of this study have not been made available because the data are one of the research results of Sunward Intelligent Equipment Co., Ltd.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This study was supported by Top Ten Technical Research Projects in Hunan Province (Research on digital prototype and digital twin technology of the rotary drilling), the 111 Project (No. D16004), Hunan Provincial Natural Science Foundation (2020JJ5173), and the National Natural Science Foundation of China (51809091).

## References

[1] B. Challen, "Automotive NVH career needs," *Sound and Vibration*, vol. 47, pp. 5-6, 2013.

[2] J. Gomez and J. Williams, "Analysis of acoustic characteristics of a car cabin using computer-aided engineering," *Applied Mechanics and Materials*, vol. 268-270, p. 883, 2013.

[3] Y. Z. Mi, Y. Q. Zhou, and L. Wang, "Vibration and noise damping treatment for an excavator cab," *Advanced Materials Research*, vol. 889-890, pp. 455–458, 2014.

[4] M. Hosseini Fouladi, M. J. M. Nor, and A. K. Ariffin, "Spectral analysis methods for vehicle interior vibro-acoustics identification," *Mechanical Systems and Signal Processing*, vol. 23, no. 2, pp. 489–500, 2009.

[5] X. Sheng, C. J. C. Jones, and D. J. Thompson, "Prediction of ground vibration from trains using the wavenumber finite and boundary element methods," *Journal of Sound and Vibration*, vol. 293, no. 3–5, pp. 575–586, 2006.

[6] E. Tijs, J. Wind, and D. Fernández Comesaña, "High resolution panel noise contribution method," *SAE Technical Paper*, vol. 01, pp. 94–104, 2011.

[7] H. Li and F. Shibo, "Experimental modal analysis of car's body-in-white," *Sustainable Construction Materials and Computer Engineering*, vol. 346, pp. 627–633, 2012.

[8] S. Marburg, H.-J. Beer, J. Gier, H.-J. Hardtke, R. Rennert, and F. Perret, "Experimental verification of structural-acoustic modelling and design optimization," *Journal of Sound and Vibration*, vol. 252, no. 4, pp. 591–615, 2002.

[9] A. R. Mohanty, B. D. St Pierre, and P. Suruli-Narayanasami, "Structure-borne noise reduction in a truck cab interior using numerical techniques," *Applied Acoustics*, vol. 59, no. 1, pp. 1–17, 2000.

[10] K. Cheung, R. G. Southwick, and J. P. Campbell, "Low frequency noise prediction and control," in *Proceedings of the 2012 IEEE 11th International Conference on Solid-State and Integrated Circuit Technology*, pp. 531–534, Xi'an, October 2012.

[11] Y. Hyunwoo and S. Kwon Lee, "Analysis of influence of oil pan structure modification on radiation noise of an engine using fem," *Transaction of the Korean Society of Automotive Engineers*, vol. 25, pp. 573–580, 2017.

[12] D. Siano and S. Giacobbe, "Radiated engine noise prediction using multi-body simulation," in *Proceedings of the Eurodyn-International Conference on Structural Dynamics*, pp. 3110–3117, Leuven, Belgium, 2011.

[13] S. Chen, D. Wang, J. Song, and G. Tan, "Interior noise prediction and analysis of heavy commercial vehicle," *SAE Technical Papers*, vol. 13, pp. 263–272, 2011.

[14] R. Guo, L. J. Zhang, J. Zhao, and H. Zhou, "Interior structure-borne noise reduction by controlling the automotive body panel vibration," *Journal of Automobile Engineering*, vol. 226, no. 7, pp. 943–956, 2012.

[15] G. Cheng, D. W. Herrin, J. Liu, and J. Stencel, "Determination of acoustic emissions using panel contribution analysis and scale modeling," *Applied Acoustics*, vol. 155, pp. 63–74, 2019.

[16] G. Bao and G. Shao, "Identification and contribution analysis of vehicle interior noise based on acoustic array technology," *Advances in Mechanical Engineering*, vol. 9, pp. 1–12, 2017.

[17] M. Long and W. Xia, "Analysis of vehicle interior low-frequency noise based on ATV," *Applied Mechanics and Materials*, vol. 494-495, pp. 78–81, 2014.

[18] O. Wolff and R. Sottek, "Panel contribution analysis-an alternative window method," *SAE Paper*, vol. 22, pp. 74–84, 2005.

[19] S. F. Wu, M. Moondra, and R. Beniwal, "Analyzing panel acoustic contributions toward the sound field inside the passenger compartment of a full-size automobile," *Journal of the Acoustical Society of America*, vol. 137, no. 4, pp. 2101–2112, 2015.

[20] X. Han, Y.-J. Guo, H.-D. Yu, and P. Zhu, "Interior sound field refinement of a passenger car using modified panel acoustic contribution analysis," *International Journal of Automotive Technology*, vol. 10, no. 1, pp. 79–85, 2009.

[21] X. H. Liang, P. Zhu, and Z. Q. Lin, "Acoustic analysis and topology optimization of auto-body using CAE Method," *Machine Design & Research*, vol. 6, pp. 64–66, 2006.

[22] H. Su, C. Yang, H. Mdeihly, A. Rizzo, G. Ferrigno, and E. D. Momi, "Neural network enhanced robot tool identification and calibration for bilateral teleoperation," *IEEE Access*, vol. 9, pp. 70–84, 2019.

[23] Y. Hu, H. Su, L. Zhang, S. Miao, G. Chen, and A. Knoll, "Nonlinear model predictive control for mobile robot using varying-parameter convergent differential neural network," *Robotics*, vol. 8, no. 3, pp. 64–85, 2019.

[24] Z. Li, B. Huang, Z. Ye, M. Deng, and C. Yang, "Physical human-robot interaction of a robotic exoskeleton by admittance control," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, pp. 1–11, 2018.

[25] X. Zhou, W. Qi, S. E. Ovur et al., "A novel muscle-computer interface for hand gesture recognition using depth vision," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, pp. 1–12, 2020.

[26] H. Su, W. Qi, Y. Hu et al., "Towards model-free tool dynamic identification and calibration using multi-layer neural network," *Sensors*, vol. 19, pp. 36–53, 2019.

[27] Z. Li, B. Huang, A. Ajoudani, C. Yang, C. Yi-Su, and A. Bicchi, "Asymmetric bimanual control of dual-arm exoskeletons for human-cooperative manipulations," *IEEE Transactions on Robotics*, vol. 7, pp. 264–271, 2017.

[28] X. Zhou, J. He, Q. He, C. Ren, Bhushan, and M. He, "Motion kinematics analysis of a horse inspired terrain-adaptive unmanned vehicle with four hydraulic swing arms," *IEEE Access*, vol. 8, pp. 194351–194362, 2020.

[29] Z. Li, Y. Yuan, L. Luo et al., "Hybrid brain/muscle signals powered wearable walking exoskeleton enhancing motor ability in climbing stairs activity," *IEEE Transactions on Medical Robotics and Bionics*, vol. 1, no. 4, pp. 218–227, 2019.

[30] U. Kumar, S. Soman, and J. Deva, "Benchmarking NLopt and state-of-the-art algorithms for continuous global optimization via [formula omitted]," *Swarm & Evolutionary Computation*, vol. 27, pp. 116–131, 2015.

[31] B. P. Chan, "YAM2: yet another library for the M2 variables using sequential quadratic programming," *Computer Physics Communications*, vol. 264, pp. 107967–107979, 2021.

Hindawi
Complexity
Volume 2021, Article ID 6625739, 18 pages
https://doi.org/10.1155/2021/6625739

*Review Article*

# An Improved Framework for Content- and Link-Based Web-Spam Detection: A Combined Approach

**Asim Shahzad**,[1] **Nazri Mohd Nawi**,[2] **Muhammad Zubair Rehman**,[3] **and Abdullah Khan**[4]

[1]*Faculty of Computer Science, Abbottabad University of Science and Technology, KPK, Abbottabad, Pakistan*
[2]*Soft Computing & Data Mining Centre (SMC), Faculty of Computer Science and Information Technology, Universiti Tun Hussein Onn Malaysia, Parit Raja, Johor 86400, Malaysia*
[3]*Faculty of Computing and IT, Sohar University, Sohar 311, Oman*
[4]*Institute of Computer Sciences and Information Technology, Faculty of Management and Computer Sciences, University of Agriculture, Peshawar, Pakistan*

Correspondence should be addressed to Abdullah Khan; abdullah_khan@aup.edu.pk

Received 29 July 2021; Revised 25 September 2021; Accepted 15 October 2021; Published 15 November 2021

Academic Editor: Bo Xiao



In this modern era, people utilise the web to share information and to deliver services and products. The information seekers use different search engines (SEs) such as Google, Bing, and Yahoo as tools to search for products, services, and information. However, web spamming is one of the most significant issues encountered by SEs because it dramatically affects the quality of SE results. Web spamming's economic impact is enormous because web spammers index massive free advertising data on SEs to increase the volume of web traffic on a targeted website. Spammers trick an SE into ranking irrelevant web pages higher than relevant web pages in the search engine results pages (SERPs) using different web-spamming techniques. Consequently, these high-ranked unrelated web pages contain insufficient or inappropriate information for the user. To detect the spam web pages, several researchers from industry and academia are working. No efficient technique that is capable of catching all spam web pages on the World Wide Web (WWW) has been presented yet. This research is an attempt to propose an improved framework for content- and link-based web-spam identification. The framework uses stopwords, keywords' frequency, part of speech (POS) ratio, spam keywords database, and copied-content algorithms for content-based web-spam detection. For link-based web-spam detection, we initially exposed the relationship network behind the link-based web spamming and then used the paid-link database, neighbour pages, spam signals, and link-farm algorithms. Finally, we combined all the content- and link-based spam identification algorithms to identify both types of spam. To conduct experiments and to obtain threshold values, WEBSPAM-UK2006 and WEBSPAM-UK2007 datasets were used. A promising F-measure of 79.6% with 81.2% precision shows the applicability and effectiveness of the proposed approach.

## 1. Introduction

Spamdexing or web spamming is described as "an intentional act intended to trigger illegally favourable importance or relevance for a page, considering the web page's true significance" [1]. Several researchers studied spamdexing issues and found that twenty per cent of web hosts are spam [2]. Web spamming is a well-known challenge for search engines [3]. It massively reduces the SE's results quality. Several users get frustrated while searching for important information when they end up with inappropriate information due to web spamming. As the WWW is growing at an unprecedented rate, the size of available textual data has become huge for any end-user. Worldwide web size's recent survey shows that the web consists of 5.41 billion web pages. Thousands of web pages are being added every day to the web corpus, and many of them are either spam or duplicated [3]. Web spammers use several creative spamming methods for dragging the Internet users to their websites for taking different benefits from them. The goal behind creating the spam pages is to cheat the SE so that it presents spam pages that are entirely nonbeneficial and irrelevant to the web users. The web spammer's ultimate

target is to improve the spam page's rank on a search engine's results page. Besides, there is a substantial economic impact of web spamming; a website with a higher page rank can qualify for a large volume of web traffic and free advertisement. During the past couple of decades, researchers from industry and academia were working hard to propose some advanced techniques for web-spam detection, but web-spamming methods are also evolving and spammers are introducing new spamming methods every day [4]. The research on web-spam detection has become an arms race to challenge an opponent who is consistently introducing new and advanced methods [4]. If one can detect and remove all spam web pages, building a robust and efficient Information Retrieval System (IRS) will be possible. More efficient and advanced SEs, providing results consistent with the user's search query, are currently required. The next significant task would be to order the retrieved web pages by their content or semantic similarity between the search query entered by the user and retrieved pages. The last step would be to present the arranged web pages to the user. Web spamming has several adverse effects on both search engines and end-users [5]. Spamdexing is not only wasting storage space and processing resources but also wasting time. As a search engine needs to index and store many pages, extra storage space is required. Moreover, based on the user's search query, SE needs to search the vast corpus of web pages, so more time is required for searching and retrieving the relevant data. This weakens the search engine's effectiveness and reduces the end-user's confidence in the search engine [6]. Enhancement in antispamdexing techniques is required to overcome web-spam attacks. Any method that is used to obtain an undeservedly high rank for a web page is web spamming. Generally, there are three types of spamdexing techniques: cloaking, link-based spamdexing, and content-based spamdexing. In cloaking, spammers offer content to end-users that is entirely different from the content submitted to the search engine spiders [7]. However, most commonly used web-spamming techniques are content- and link-based web spamming which are investigated in this research work. All techniques used by web spammers to change the logical view that a search engine has over the page content are known as content-based web spamming [1]. It is a widespread type of web spamming [8]. Most of the search engines use information retrieval (IR) models, such as BM25 [9], probabilistic models, statistical-language models [10], and vector-space models [11]. These models are applied to a web page's content for ranking the web page, so content-based web spamming is very popular among web spammers. To manipulate the spam web pages' content, web spammers utilise vulnerabilities of these models [12]. For example, they might use famous keywords many times on a spam web page to increase the keywords' frequencies, copy legitimate website's content, produce the content for spam web pages using machine-generated techniques, and add dictionary's words on a spam web page, giving those words the colour of the background so that these words cannot be seen on the spam web page by the user and are only visible to search engine spiders. These are some of the content-based web-spamming techniques used by web spammers to obtain a higher page rank on search engine's results. Based on the structure of a web page, content-based web spam can be divided into five subcategories: title spam, anchor-text spam, meta-tag spam, body spam, and URL spam. Several other spamming techniques which can target different algorithms used by search engines exist [13]. Another widespread type of spamdexing is link-based spamdexing. Davison [14] described the link-based spamming as the links among several pages that are present for a reason other than merit. To get recognition from link-based algorithms, web spammers build link structures using link-based spamdexing techniques. For example, the PageRank (PR) algorithm will allot a higher page rank to a web page if several other top-ranked websites point to this page with backlinks.

This research work focuses on the detection of both the link- and content-based spamdexing techniques. An improved framework for content- and link-based web-spam detection is proposed. In the proposed approach, stopwords, keywords frequency, part of speech, spam keywords database, and copied-content algorithm are used to detect the content-based web spamming. For the detection of link-based spamdexing, we initially exposed the relationship network behind the link spamming and then used the neighbour pages and spam signals for link-based spam identification. For this purpose, WEBSPAM- UK2006 and WEBSPAM-UK2007 datasets are used. The results with a promising F-measure and better precision show the proposed framework's effectiveness and applicability.

## 2. Literature Review

In response to the web-spamming challenge [12], different researchers proposed various methods for identifying content-based spamdexing. A group of researchers [3] suggested a few critical content features. They introduced HTML-based characteristics and text compressibility to recognise the content spam. Piskorski et al. [15] explored a considerable number of linguistic features. Latent Dirichlet Allocation (LDA) [16] is widely used for text classification tasks. Biro et al. enhanced the LDA and introduced modified versions, linked LDA [17] and multicorpus LDA [18] models, to detect spamdexing more efficiently. To detect content-based spam web pages, Ntoulas et al. [3] applied the decision-tree classifier. They proposed several features, such as the number of words, the average length of words, the visible portion of the content, and the anchor-text amount within a web page. Another group of researchers introduced the semisupervised method and combinatorial feature fusion [17]. They used semisupervised learning to exploit unlabelled samples and used a combinatorial feature fusion technique to create new features and reduce the term frequency-inverse document frequency (TF-IDF) of content-based features. The results determined the effectiveness of their strategy. Ntoulas et al. [3] did the most fundamental work on the detection of content-based spamdexing algorithms [18–20]. In their research article [20], they suggested the use of statistical analysis. Spam web pages can easily be classified using statistical analysis because web spammers usually create spam pages automatically using the phrase stitching and weaving techniques [1]. These spam web pages

are specifically designed for search engine spiders to obtain a higher rank in search engine results. These web pages are not intended for real humans, so one can observe the abnormal behaviour of these web pages. The researchers also reported that there are a massive number of dots and dashes in the URL of spam web pages and its length is exceptional. During their analysis, they observed that out of hundreds of longest hostnames, eleven were pointing to financial credit-related web pages and eighty were referring to adult content. They also found that these web pages themselves have a duplicating nature; spam web pages hosted by the same host almost contain the same content with minimal variance in the word count. Spammers change the content on these web pages very quickly. They gravely observed the content changing feature for a specific host every week and they tracked the changes on these spam web pages. Finally, they concluded that, based only on content changing features, 97.2 per cent of the most active spam hosts can be detected. Their research identified several other features, which can be seen in [20]. In another research study, these researchers worked on content duplication and concluded that large clusters with similar content are spam [18, 19]. To identify such clusters and duplicate content, they used the shingling technique [21] which is based on Rabin fingerprints [22,23]. Another group of researchers [24] worked on machine-learning models and various other features. They explained how machine-learning models and different features could help in web-spam detection. They achieved the best classification results by using easy to calculate content features, LogitBoost, RandomForest, and several learning models. During this work, they identified computationally demanding and global features; for instance, PageRank (PR) could help very little in quality enhancement. Based on their findings, they claimed that the selection of a proper machine-learning model is critical. Urvoy et al. [25] proposed some features to identify the script- or machine-generated spam web pages. The features are based on the structure of the HTML page. For data preprocessing, these researchers used a unique and nontraditional technique. They removed all the content from the web page and only retained its layout. Instead of examining the content of a web page, they just examined the layout to identify the web page duplication. They used fingerprinting technique [22, 23] followed by clustering and identified groups of spam pages which are structurally near-duplicate. By matching the language models [26] for web pages, a group of researchers [27] proposed a technique for spamdexing detection in blogs.

Web spammers use link-spamming techniques for various reasons ranging from malware propagation to monetary activities. Due to the fast growth and volatile nature of the Internet, spammers are introducing more sophisticated techniques to generate more revenue [28–30]. Unfortunately, these practices have several negative impacts on both search engines and user's experience. For instance, users get annoyed when they cannot find what they are looking for and their systems face possibly high-security threats due to malicious content on these spam web pages. Similarly, spam web pages cause issues for search engines by wasting useful resources of SEs. In the past few decades, researchers developed several new antispamdexing techniques to overcome this issue [31–33]. However, web spammers keep a close eye on antispamdexing techniques and continuously improve their spamming techniques to avoid detection. Usually, search engines use the number of inbound links and their ranking to determine a web page's popularity and reputation. If a large number of popular web pages will link to a web page $p$, it will get a higher rank in search engine results pages [34]. Though it is an excellent technique to define the page ranking of $p$, spammers exploit this technique to boost their page rank by increasing the inbound links to it. Web spammers often use the so-called Facebook search engine optimisation (SEO) groups, SEO forums, paid-link services, subreddits, and SEO service providers for this reason even though every search engine provides guidelines to website owners for achieving a good rank [35]. Several researchers from academia and industry also worked on different techniques used for link-based spamdexing detection. Based on these techniques' working mechanisms, all web-spam detection algorithms can be subdivided into five categories. The first category deals with recognising suspicious links, nodes, and their subsequent downweighting [36]. The working mechanism of algorithms in this category identifies weak and suspicious links and then penalises them. Several issues in the HITS algorithm were identified by Bharat et al. [37], for instance, the neighbour-graph topic drift and the dominance of mutually reinforcing relationship. To solve these issues, they proposed a method in which they augmented content analysis with link analysis. Another group of researchers [38] studied the same problem and suggested a technique known as the projection-based technique, and they used this technique for calculating authority scores. They changed the eigenvector in the HITS algorithm to get better results. In the second category, spamdexing detection algorithms deal with the topological relationship between a set of web pages with unknown and known labels by applying various propagation rules for computing the labels of other nodes [36]. One of the earliest algorithms from this group is TrustRank (TR). TR uses personalised PR to propagate the trust from a small seed set of excellent web pages [39]. The third category represents the graph regularisation techniques of link-based spamdexing detection algorithms [40]. The algorithms utilise web graph in this category for smoothing the predicted labels. Some results proved the effectiveness of these algorithms. The fourth group of spamdexing detection algorithms is based on the label refinement on a web graph topology [41]. Usually, the researchers use machine-learning methods for label refinement to classify general problems [42], but several researchers used this method for web-spam detection. Finally, the last category of spamdexing detection algorithms is extracting the link-based features for each node and applying various machine-learning methods for spamdexing detection [43]. As they are supported by most of the link-based techniques, HITS and PR are the most critical ones among all the above algorithms.

Abernethy et al. [40] proposed a combined approach of content- and link-based features for spamdexing detection. Their research offered a WITCH algorithm to detect

spamdexing, and they used a graph regularisation classifier (GRC) with a support-vector machine (SVM). To differentiate legitimate web pages from spam web pages, Egele et al. [44] introduced a new technique and used a J48 decision-tree classifier in their approach. They could detect one spam page out of five spam web pages by decreasing the false positive to zero. Prieto et al. [42] proposed a SAAD (spam analyser and detector) spamdexing detection technique based on a heuristic set. For testing and comparing the SAAD with other existing benchmark techniques, they used two public datasets, Yahoo! and e-mail spam (web-spam corpus). Finally, after getting the results, they announced that their proposed approach could generate a secure client environment and protect the users from attacks. For detecting spam web pages using weight properties, Goh et al. [45] proposed a link-based approach. They defined weight properties as the influence of one web node on the other. They used the WEBSPAM-UK2007 dataset for their experiments, and their results outperformed the benchmark algorithms by up to 6.11 per cent improvement at the page level and 30.5 per cent improvement at the host level. Roul et al. [13] proposed a combined approach of content- and link-based techniques for spamdexing detection. They used part of speech (POS) ratio and term density to detect content-based spam and explored the personalised PR to classify web pages as nonspam or spam. For conducting their experiments, they used the WEBSPAM-UK2006 dataset. Finally, they compared their results with the existing techniques. An excellent F-measure of 75.2 per cent shows the effectiveness of their approach. The search engines' performance for spamdexing detection can be improved by combining content- and link-based spamdexing identification techniques. However, few researchers have worked on it, and very little work has been done using an integrated approach. In our framework, for detecting the spam web pages, we combined both content-based and link-based features. Our results demonstrate the significance of the combined approach and provide better results by obtaining excellent F-measure as compared to existing standard techniques.

## 3. Content-Based Spamdexing Detection

Content-based web-spamming techniques are easy to use and are favoured among web spammers. To detect content-based web spamming, we proposed an improved framework [46]. To obtain the most appropriate threshold values for conducting experiments, we first preprocessed the data and then proposed and improved several different content-based web-spam-detection algorithms. Each proposed or improved algorithm can detect different types of content-based web-spamming techniques, such as the proposed stopword density technique and an improved POS technique that can detect the machine-generated content. The keyword density technique can identify the keyword stuffing on any web page. The keyword density technique identifies spam web pages based on the number of spam keywords on the page, and the copied-content technique can identify the spam web pages created by copying the content from other useful websites. After proposing several different content-based techniques, we combined all techniques to propose a content-based web- spam-detection framework. The details regarding all of the improved and proposed techniques and the proposed framework can be seen in [46].

## 4. Revealing the Hidden Relationship behind Link-Spam Network

Detecting newly evolved link-based spamdexing is a continuous process for SEs which is a hidden process from end-users. Usually, in SEO communities, link spammers exchange the links of their web pages to create global link farms. The standard web-spam-detection datasets (WEBSPAM- UK2007 and WEBSPAM-UK2006) are quite old, and to propose an improved link-based spamdexing detection framework, it is essential to identify and understand the currently practised spamdexing techniques. To understand new spamdexing techniques and to reveal the secret relationship behind the link-spam network, we performed some experiments to collect the data through data-collection architecture. Several channels are used by individuals to communicate with each other, and the most common of them are Facebook SEO groups, SEO forums, and subreddits. Individuals also contact SEO service providers to improve their visibility in SERPs. Therefore, using traditional antispamdexing methods, it is impossible to detect these link farms. We studied the spam web pages and the behaviour of web spammers who are using Facebook SEO groups, SEO forums, subreddit, and paid SEO services during our experiments for data collection. We also maintained the database of current spam web pages and revealed the secret link-based spamming networks. For revealing the secret network, we proposed an architecture for data collection that allows us to identify the web pages participating in link farms and to determine the web pages using Facebook SEO groups, SEO forums, and subreddits to improve their rank.

Moreover, we were able to reveal the secret relationship between these web pages. We proposed this architecture to study link spam and kept in mind that we will use this data and knowledge to suggest and improve the web-spam detection framework for link-based web spamming. The basic concept behind our system is that link spammers mutually exchange reciprocal links. More specifically, when a web spammer X wants to increase the inbound links to X's website, X will contact other web spammers, and all of them will mutually exchange backlinks. Usually, web spammers use Facebook SEO groups, SEO forums, SEO subreddits, and some other platforms to communicate with each other for link exchange and other services. They exchange the links in three different ways:

(i) When a web page x links to another web page y and asks for payment or any other benefits instead of the backlink in return, the link exchange is said to be one-way. Usually, paid-link services sell the backlinks using this technique.

(ii) When a web page x links to a web page y and in return y links back to x, the link exchange is said to be two-way. Two-way link exchange is popular among web spammers.

(iii) In the three-way link exchange technique, web page x does not link directly to web page y. Instead of a direct link, a third web page z is used to create a loop. For example, the web page x links to the web page z, and the web page z links to x. Detecting this type of link exchange is more problematic as compared to the first two. By having this information in advance, we can precisely identify all web pages trying to improve their visibility in SERPs using fraudulent techniques. Therefore, we have designed crawlers for link exchange identification, honeypot accounts for obtaining information and data from web spammers in private messages, and databases for storing URLs of all spam web pages. We analysed ten leading Facebook SEO groups, ten SEO forums, five subreddits, and two SEO service providers. After one year, we identified 437,811 web pages involved in link spamming. We discovered several categories of web spammers during our studies and determined that web spammers are using almost all available platforms for increasing their web pages' ranks. They offer backlinks services and provide Facebook likes, Twitter followers, Instagram followers, YouTube subscribers, post boosting, and video views for different social media accounts connected to a specific website for which they want to increase the page rank. Several web spammers offered us their services for approving Google AdSense for spam web pages during our experiments. You can get the cracked version of all SEO software on these platforms and authorise Google AdSense to earn income from Google Ads. After analysing all the spamming categories, we observed two significant types of spammers with unique features. Both the varieties behave differently, so other techniques are required to identify their spam pages. Besides, a more in-depth analysis of the data exposed evident variations in the type of link exchange and helped us understand the secret-relationship network. The secret-relationship network between URLs displayed in public posts and threads is different from the URLs sent through personal messages. In short, our main contributions are as follows:

(1) We gathered the corpus of web-spam links from Facebook SEO groups, SEO forums, subreddits, purchased links, and SEO service providers. We also used the honeypot accounts for harvesting the web-spam links from the web spammer's messages.

(2) We analysed all the links in our corpus. Instead of entirely depending on the data gathered from different platforms, we crawled respective pages to confirm the link exchange.

(3) We revealed several strategies and techniques used by inexperienced and advanced spammers.

(4) We exposed different types of link farms, wheels, and pyramids currently used by web spammers.

(5) We performed an in-depth analysis of spam web pages and identified spam signals in spam pages.

(6) We used these spam signals in our framework for the identification of link-based web spamming.

*4.1. Data-Collection Architecture.* Figure 1 represents the general structure of our technique. The main components of our architecture are (i) honeypot accounts; (ii) SEO groups, SEO forums, and SEO services; (iii) web page crawlers; and (iv) spam-links and paid-links databases.

Initially, we planned to execute SEO crawlers for harvesting URLs from Facebook groups, SEO forums, and other platforms used in our experiments. However, we manually collected the links from all platforms due to (i) some policy restrictions of platforms, (ii) diverse technologies used by different platforms for implementing their systems, and (iii) difficulty faced by SEO crawlers in differentiating between legitimate and spam URLs in different discussion groups. These were the main reasons for collecting the spam URLs manually and carefully instead of executing SEO crawlers. The spam URLs are manually collected from public posts, author's signatures, and private messages.

*4.1.1. Honeypot Accounts.* Honeypot accounts are used to attract web spammers and getting secret information from them which is not available publicly. Hence, technically, honeypot accounts are fake accounts created by us for collecting unrevealed data from web spammers. Web spammers only present this kind of information in private messages, so honeypot accounts are advantageous in exposing web spammers.

*4.1.2. Spam- and Paid-Links Databases.* We created two types of databases: paid-links and spam-links databases. We observed that some spammers are not interested in backlinks and only offer one-way link exchange for money or other benefits. Therefore, we added all such URLs to the paid-links database. We also used several search engines and different keywords to identify the services and URLs that are selling the backlinks. The spam URLs, collected from public threads and private messages, interested in two- and three-way link exchange, were added to the spam-links database.

*4.1.3. Web Page Crawler.* We used the web page crawler to map the link structure of the secret networks behind the collected URLs. After defining a specific depth for crawling, our web page crawler followed all outgoing links. For more in-depth analysis, it stored the crawled URLs in our spam-links database. We also defined the mechanism to check if a URL is already crawled. During our experiment, we observed that most of the spam web pages link other pages

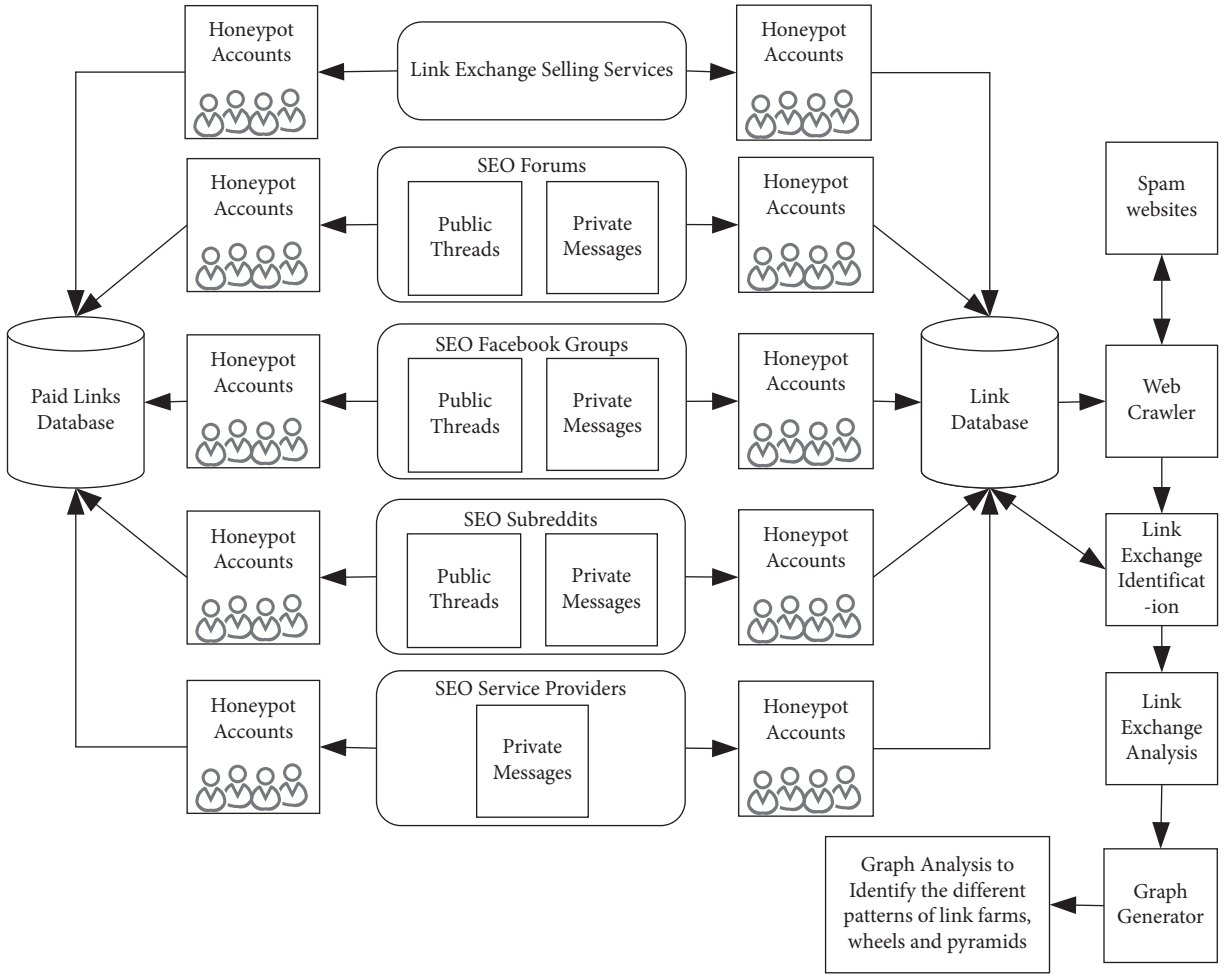



Figure 1: Data-collection architecture for the spam web page.

from their main page. In short, we used a web crawler to validate the link exchange.

*4.1.4. Moz Pro.* To obtain more accurate results, we used the Moz Pro tool for identifying incoming links of initially collected URLs from public threads, private messages, and SEO services. Moz Pro identified the inbound links, and all the links are stored in spam-links databases. Most web-spam-detection techniques detect spam based on the rule that spam pages usually point to other spam pages. Identifying inbound links will help us in generating a more accurate spam-network graph.

*4.1.5. Link Exchange Analysis.* After gathering and storing all the necessary information in the databases, we performed the link exchange analysis and correlated the relationships among crawled web pages. More specifically, we identified the relationship between various entities and observed how web spammers interact in link farms. Our approach can recognise all three types of link exchange techniques. We identified different types of link farms, link circles, and link pyramids during our analysis, although three-way link exchange was more difficult to locate.

*4.1.6. Web Graph.* Representing spam web pages in graphical form can provide a clear view. Therefore, we generated a web graph to visualise the structure of different link farms, link circles, and link pyramids. We recognised significant players on the link exchange and how they create combined link farms, link circles, and link pyramids. All this information helped us in designing the improved framework for the detection of link-based spamdexing.

*4.2. Analysis of Facebook Groups, SEO Forums, Subreddits, and SEO Service Providers.* For our experiment, we selected the best ten Facebook SEO groups, ten SEO forums, five subreddits, and two SEO service providers and collected the data for one year. We have selected the most active groups, forums, subreddits, and two SEO service providers using black-hat techniques. SEO service providers were carefully chosen after analysing several services and applying the selection criteria. The users use all these places for discussing the SEO boosting methods, so we only target those discussion threads where users were talking about the link exchange and SEO techniques. In total, we analysed 32,674 threads and extracted 137,842 unique URLs posted by 16,827 users and extracted 5,000 unique URLs from the documents

provided by SEO service providers. They provided us with these documents containing backlinks to our website as proof of their work. Our results show that the ratio of replies to each thread is 7.2, which means that for every web page trying to increase its page rank, seven other web pages are available for participation to achieve this target. During this study, we identified two different types of link spammers (beginners and experts). Beginners freely exchange their websites in publicly available posts, while expert spammers only exchange their websites in private messages. Our experiments proved that spammers who exchange their URLs through private messages are more suspicious and are part of bigger link farms and pyramids. Another exciting aspect of understanding the link-based spam network is calculating the page authority, domain authority, and page rank of the website requesting the link exchange. We used SEO PageRank and Moz Pro for calculating the page authority, domain authority, and page rank. After calculating these values, we noticed that most spam web pages have on average the domain authority of seven, page authority of fifteen, and page rank of two, while more than 92% of spam web pages have on average the domain authority of twelve, page authority of seventeen, and page rank of four. This means that the owners of low-ranked web pages behave differently from the owners of high-ranked web pages. Usually, high-ranked web page owners do not exchange links with low-ranked web pages; in some cases, we noticed they are selling the backlinks for money. Every website has a theme of content and the subject of most of the spam pages is earning quick money, finance, travel, adult content, and health. More in-depth analysis revealed that web spammers exchange links with high-ranked or identical-ranked web pages. In public threads, the spammers prefer a two-way exchange while only a small percentage of spammers speak about the one-way or three-way exchanges. On the other hand, in the case of honeypot accounts, most of the spammers are interested in one-way exchange, three-way exchange, and link farms. Initially, through a manual approach, we collected 142,842 unique URLs from honeypot accounts and SEO service providers. For our web crawler, we used these URLs as seeds. In one year, we crawled more than twenty-five million pages, and we were able to identify 437,811 web pages that were engaged in all three types of link exchange. We, by analysing the data, found that the percentage of two-way link exchange is 90.86%, the three-way link exchange percentage is 7.43%, and the one-way link exchange (paid-links) percentage is 1.71%.

## 5. Data Preprocessing and Experiments for Calculating the Threshold Values for Link-Based Spamdexing

To perform our experiments, we used well-known datasets, namely, WEBSPAM-UK2006 and WEBSPAM-UK2007, and the data was collected using data-collection architecture. To get more accurate results, we manually selected web pages labelled as nonspam/spam by humans. We extracted some link-based features for link-based spamdexing identification. Finally, we obtained the most appropriate threshold values that provided the fewest false positive ratios and high F-measure through different experiments.

*5.1. Paid-Links Database.* We identified and collected websites that were selling paid links during experiments conducted through data-collection architecture. We also crawled the Internet to search for services that are selling paid links that might increase the ranking and visibility of undeserving websites in search engine result pages. We stored all such services and sites in the paid-links database. We will use this database in our framework for link-based spamdexing detection. Besides, we also stored all spam-web-page links identified through data-collection architecture in another spam-page database. Whenever our framework marks a web page as spam, it will automatically store the spam page's link in the database for reuse in the future. This will help us identify spam web pages and improve the accuracy of our framework.

*5.2. Spam-Signal Identification.* We manually selected 2,000 spam/nonspam web pages identified during the data-collection experiment and 3,000 web pages labelled as spam/nonspam by humans from the datasets (WEBSPAM-UK2006 and WEBSPAM-UK2007) to analyse web pages for spam signals. We carefully examined all the web pages one by one and identified the following potential spam signals:

(1) Single page website: many spam websites consist of a single web page or very few pages. Though this does not prove that every website with a single or few pages is spam, several spam websites have one or a few pages; hence, there is a correlation.

(2) Thin content: web pages having no or very little content are known as thin content web pages. These pages have no or little value to the user. Usually, these pages consist of copied, machine-generated, or low-quality affiliate content.

(3) No contact information: almost all the spam websites do not have any contact information on their pages. These websites rarely have real phone numbers, e-mail addresses, or any other contact information on their pages.

(4) Presence of spammy keywords: spam sites regularly use particular words directly linked to spam topics like earning money from home, getting free cash, gaming, pharmaceuticals, adult content, and others.

(5) No SSL certificates: most of the spam web pages did not invest in SSL certificates, though HTTPS is an excellent trust signal.

(6) No links to social media accounts: usually, spammers do not associate social media pages with their spam web pages. During our analysis, we found that nearly all spam sites had not associated LinkedIn pages, and Facebook tracking pixel was rarely present on these spam pages.

(7) External outgoing link: the outgoing link points to an external or targeted domain and is different from the link present on the source domain. Spam sites host a massive number of nonrelevant external outgoing links.

(8) Content to links ratio: on legitimate sites, the ratio of the content and external links is balanced, while spam sites have abnormal ratios of content to links. We found that some spam pages consist of external links only and there was no content at all.

(9) The ratio of incoming links: a widespread perception is that only legitimate sites backlink other legitimate sites. In spam sites, we noticed that the ratio of spam incoming links to legitimate incoming links is higher, which is a significant spam signal.

(10) External links in navigation: spam sites host and, in most cases, try to hide many external links in toolbars, footers, and sidebars.

(11) A few internal links: during our page analysis, we found that spam sites have very few internal links. Usually, legitimate sites heavily link to themselves internally to correlate the content. The absence of internal links and navigation is one of the spam signals.

(12) URL length: spam pages character count (URL length) is abnormally long and higher than average size. For example, https://getcheapmedicines.freeshipping.cheappharmacy.com shows the keyword stuffing.

(13) Numerals in domain name: most spam websites contain numbers in their domain names. There is a strong possibility that domain names with numbers in them are generated automatically; therefore, it is a spam signal.

(14) Top-level domains (TLD): some of the top-level domains are correlated with spam domains, for instance, .cc, .pw, .pl, etc. Several spam sites use these TLDs.

(15) Huge proportion of anchor text: in spam sites, the content text proportion is minimal compared to anchor text. A large amount of anchor text in a website is a spam signal.

(16) Site markup proportion: in spam sites, the markup proportion is abnormally small compared to JavaScript and HTML where the proportion of visible text is higher. Typically, legitimate sites invest in a rich user experience with extensive markup CSS and JavaScript.

(17) Broken links: typically, spammers do not maintain and update their websites. We found a lot of broken links in spam pages, which is a spam signal.

(18) Favicon: in today's world, the favicon is considered an essential part of any website, and legitimate websites do use the favicon. Most spam websites do not use the favicon because they change their websites very quickly after getting benefits from it. Missing favicon is a spam signal.

(19) Page not found 404 error: during our analysis, we frequently saw the 404 error on spam sites and the reason behind this is that spammers change their websites very often and they add and remove pages daily, but they do not maintain the site properly. Usually, they do not design a customised 404 error page. The occurrence of too many 404 errors on a website is a sign of a spam site.

(20) Meta description length: standard meta description length is 44–164 characters, while spam sites contain a too short or very long meta description, which is a potential spam signal.

(21) Length of the title: the recommended length of a title is 65–70 characters and spaces are included in it, while spam sites use too short or very long titles, which is a spam signal.

After identifying the spam signals above, we performed a more in-depth analysis to determine spam signals' effects on any website. We randomly selected 5,000 nonspam/spam pages from the datasets WEBSPAM-UK2006 and WEBSPAM-UK2007, and the data was collected using data-collection architecture. To get more accurate results, we manually selected web pages labelled as nonspam/spam by humans. Then, we crawled the URLs and removed all the URLs that did not return a 200 OK. Each spam signal is a potential warning which indicates that a website might be spammy. If a website contains more spam signals, this increases the probability that a website is a spam site, so the total number of spam signals on a site is a powerful predictor of spam. Therefore, we used spam signals to predict a page because several legitimate sites might have one or a few spam signals. We analysed the pages selected for this experiment to show the relationship between the number of spam signals and spam sites. Table 1 shows all the spam signals with the odds ratio for each spam signal. For every spam signal, we can calculate two different percentages: first, the percentage of sites that contain the spam signal and are marked as spam; second, the percentage of sites that contain spam signals and are not marked as spam. The odds ratio increases the possibility that if a site has a spam signal, it is spam. For instance, the first row of Table 1 says that the site with a single page spam signal is 5.2 times more likely to be spam than the site without this spam signal.

The overall spam score is an aggregate of twenty-one different spam signals. Table 2 shows the relationship between the number of spam signals and the percentage of sites containing those spam signals that humans labelled as spam. Danger levels are divided into four different categories: low (<10%), moderate (11–50%), borderline (51–89%), and spam (≥90%).

## 6. Link-Based Spamdexing Detection

In this section, we present an improved framework for link-based spamdexing detection. Figure 2 shows the flowchart of the proposed framework. This framework uses three different methods, and every technique uses a unique feature to identify link-based spamdexing. The training process is not required as no machine-learning technique is used in the proposed framework. The methods are as follows:

Table 1: Percentage of different spam signals.

| Spam signals | Spam vs. nonspam ratio (odds ratio) | Percentage of sites with spam signals and marked as spam (%) | Percentage of sites with spam signals and marked as nonspam (%) |
|---|---|---|---|
| Single page website | 5.2 | 18.87 | 5.39 |
| Thin content | 4.08 | 26.67 | 5.09 |
| No contact information | 6.78 | 19.28 | 2.07 |
| Presence of spammy keywords | 11.32 | 23.59 | 2.32 |
| No SSL certificates | 13.01 | 27.30 | 3.83 |
| No links to social media accounts | 9.43 | 36.18 | 10.98 |
| External outgoing links | 4.60 | 21.47 | 6.75 |
| Content to links ratio | 3.78 | 36.48 | 5.89 |
| The ratio of incoming links | 9.17 | 20.06 | 1.96 |
| External links in navigation | 12.52 | 16.08 | 5.89 |
| A few internal links | 3.06 | 32.76 | 6.60 |
| URL length | 6.90 | 21.69 | 3.63 |
| Numerals in domain name | 7.36 | 35.44 | 4.69 |
| Top-level domains | 11.17 | 21.29 | 11.18 |
| Huge proportion of anchor text | 2.15 | 26.78 | 7.66 |
| Site markup proportion | 4.41 | 10.36 | 8.87 |
| Broken links | 1.19 | 17.50 | 7.96 |
| Favicon | 2.45 | 11.37 | 5.09 |
| Page not found 404 error | 2.39 | 15.27 | 9.47 |
| Meta description length | 1.72 | 13.78 | 5.59 |
| Length of the title | 2.70 | 10.70 | 3.88 |

Table 2: Spam score based on the number of spam signals in a website.

| Number of spam signals | Probability of spam (%) |
|---|---|
| 0 | 0.68 |
| 1 | 1.10 |
| 2 | 2.50 |
| 3 | 6.98 |
| 4 | 8.10 |
| 5 | 13.04 |
| 6 | 19.09 |
| 7 | 26.89 |
| 8 | 33.79 |
| 9 | 34.68 |
| 10 | 54.05 |
| 11 | 63.20 |
| 12 | 74.77 |
| 13 | 91.75.00 |
| 14 | 100.00 |
| 15 | 100.00 |
| 16 | 100.00 |
| 17 | 100.00 |
| 18 | 100.00 |
| 19 | 100.00 |
| 20 | 100.00 |

(1) link-based spamdexing detection using paid-links database

(2) Link-based spamdexing detection using spam signals

(3) Link-farm detection using backlinks

All three techniques are discussed in detail in the following sections.

*6.1. Link-Based Spamdexing Detection Using Paid-Links Database.* Backlinks play an essential role in ranking a website high in search engine result pages, and they also have a high impact on search engine ranking algorithms. That is why backlinks to a website are essential. Web spammers use different techniques to build these links and pay for the backlinks to third-party domains. Paying for the backlinks is strictly banned by search engines. This kind of web spamming can be detected by using the paid-links database. A web crawler is used to identify all backlinks of a web page W*pi*. After identifying all the backlinks, it will compare each backlink of W*pi* with links present in the paid-links database. As paid links are strictly banned and considered spam, a web page with backlinks found to match any link in the paid-links database will be marked as paid-link spam. Otherwise, the page W*pi* will be tested by the spam signals module. Figure 3 shows the algorithm for paid-link spamdexing detection using the paid-link database.

*6.2. Link-Based Spamdexing Detection Using Spam Signals.* Every single spam signal is a potential warning which might indicate that a website is spammy. Therefore, if the number of spam signals is high on a web page, there are higher

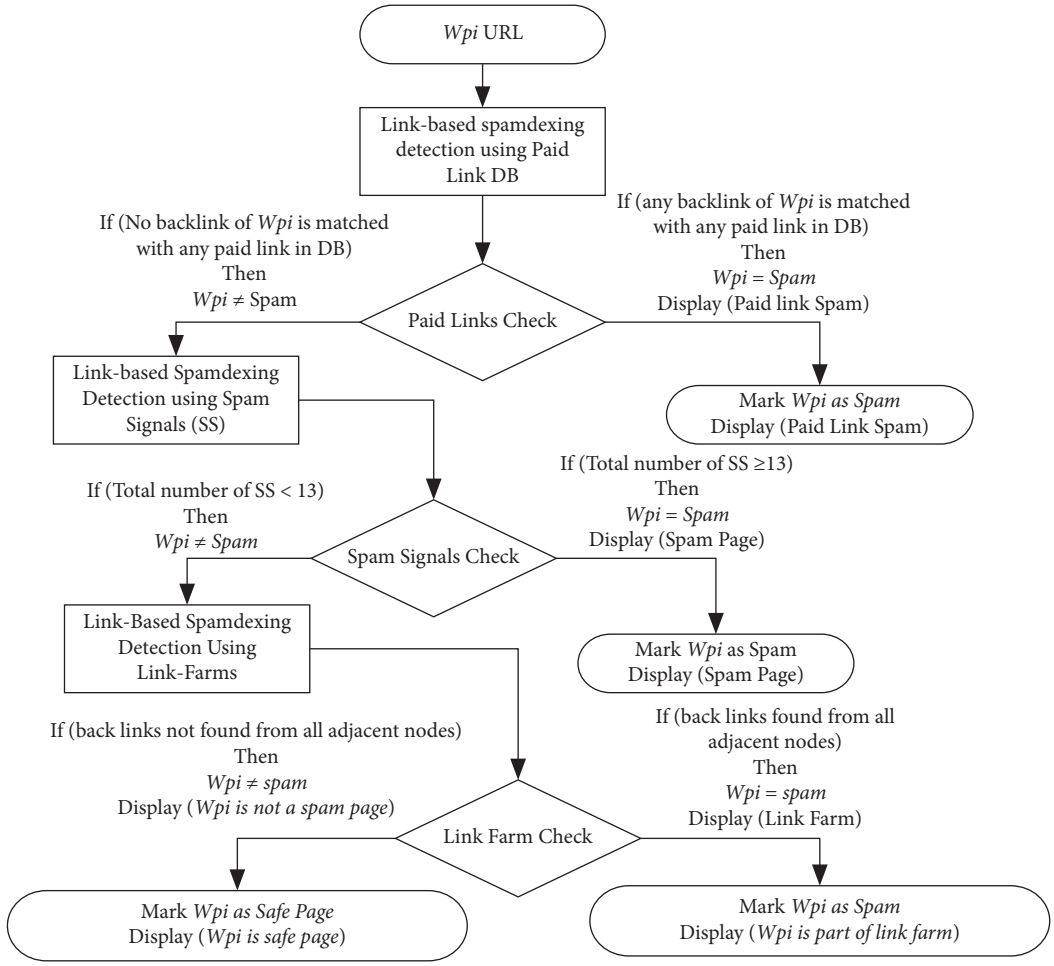



Figure 2: Improved framework for link-based spamdexing detection.

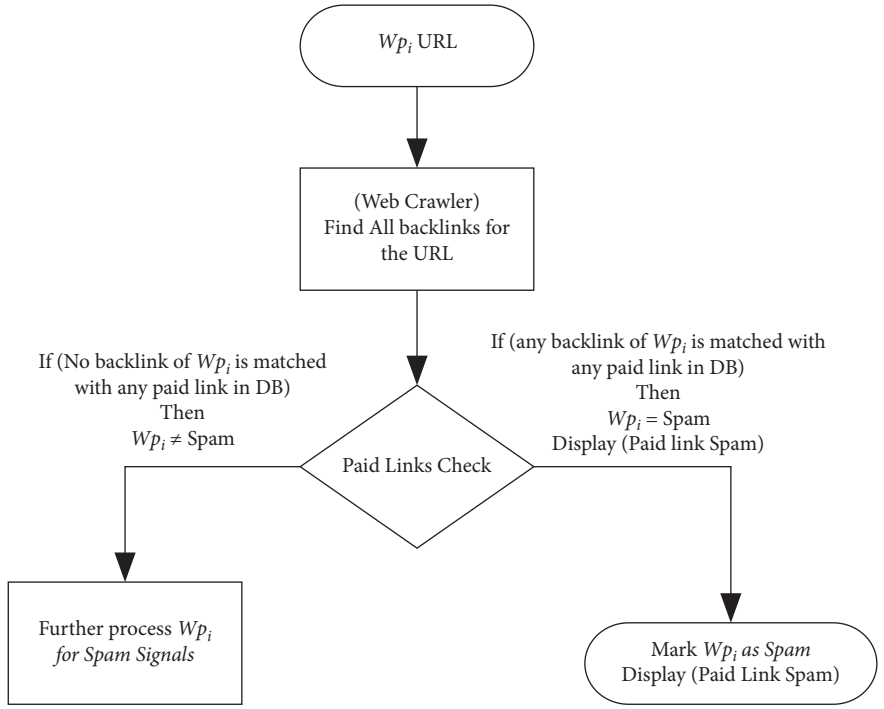



Figure 3: Algorithm for link-based spamdexing detection using the paid-link database.

chances that the page is a spam page. During our analysis and experiments described in Section 5, we identified twenty-one different spam signals and also found the effects of each spam signal on a web page. Our results in Section 5 showed that if the number of spam signals on a web page W*pi* is equal to or greater than thirteen, the web page is spam. Moreover, if the number of spam signals is less than thirteen, there are chances that the page might not be spam, so we need to run further link-based spam tests on all such pages. Since outliers always exist in our experiments, we observed some web pages with very few spam signals but nonetheless marked as spam. There were some web pages with a high number of spam signals on them, but they were marked as nonspam. We have designed a custom web crawler that can identify spam signals on any web page W*pi*. The working mechanism of the spam-signal-detection module for spam-page identification is described as follows:

(1) The spam-signal module will obtain URL of the page W*pi* from the previous module and then submit it to the web crawler.

(2) The web crawler will analyse the page W*pi* and will identify the spam signals.

(3) After the spam-signal identification, it will count spam signals on the web page W*pi*.

(4) Finally, the spam-signal module will check if the total number of spam signals on the page W*pi* is equal to or greater than thirteen. If the number of spam signals is equal to or greater than thirteen, it will mark the page as spam; otherwise, it will forward the URL of the page W*pi* to the next module for further link-based spam-detection test. Figure 4 shows the algorithm for spam signals detection using a web crawler.

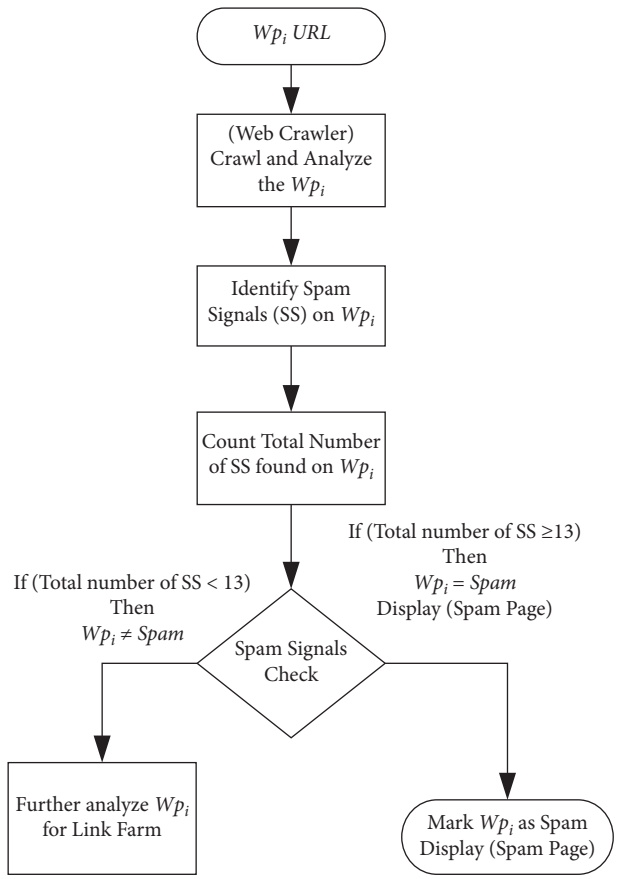



Figure 4: Algorithm for link-based spamdexing detection using spam signals.

*6.3. Link-Based Spamdexing Detection Using Link Farm.* Any group of websites on the Internet that hyperlink to every other website in the group is known as a link farm. A link farm is a clique in graph-theoretic terms. Although some web spammers create link farms manually, many use automated tools and services to create link farms. Web spammers use private-blog networks and link farms to boost ranks of their websites.

The working mechanism of the link-farm-identification module is as follows:

(1) The link-farm module will obtain the URL of the web page W*pi* from the previous module and submit it to the web crawler

(2) The web crawler will crawl all the adjacent nodes (outgoing links) of the web page W*pi*

(3) The web crawler will crawl adjacent nodes one by one to the defined depth

As W*pi* and all its adjacent nodes are crawled, a web graph is obtained and this web graph can be represented with mathematical notations as shown below. Figure 5 shows the algorithm for link-based web-spam detection using link farm.

Let us say $G = (V, E)$ is a directed graph, $V$ represents the nodes or vertices, and E represents the directed edges. G can be described as a web graph in which $V$ denotes the set of web pages and $E$ represents a set of links between pages. Moreover, it is worth mentioning that if there are several incoming or outgoing links between two pages, our algorithm will consider only one link as proof of a connection between two pages. We used the following notations and definitions in this research.

The out-degree *(od)* of a web page W*pi* is the total number of outgoing links:

$$od(Wpi) = \sum_{Wpi} E_{ij}. \quad (1)$$

The in-degree (id) of a web page W*pi* is the total number of inbound links:

$$id(Wpi) = \sum_{Wpi} E_{ji}. \quad (2)$$

The formula in (1) and (2) can be used to create an adjacency matrix. An element Aij = 1 if web page W*pi* has a link to web page W*pj*; otherwise, Aij = 0. We can call it the link/connectivity matrix.

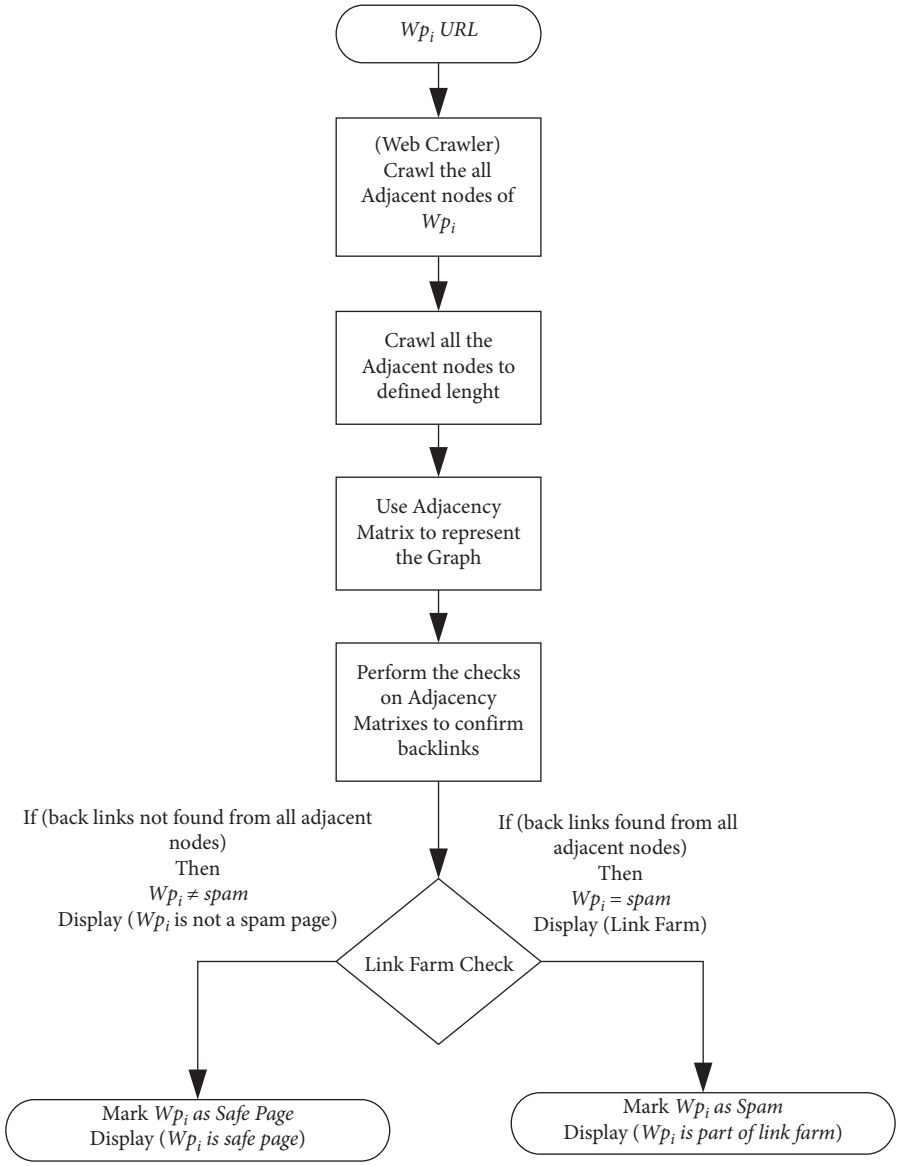



Figure 5: Algorithm for link-based spamdexing detection using link farm.

$$A_{ij} = \begin{cases} 1, & \text{if}\left(Wp_i, Wp_j\right) \in E, \\ 0, & \text{Otherwise.} \end{cases} \tag{3}$$

For a directed graph, the generalised $n$ ч $n$ adjacency matrix $A$ is shown as follows:

$$A = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{n,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{n,2} \\ \vdots & \vdots & \ddots & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{n,n} \end{pmatrix}. \tag{4}$$

For checking reciprocal links between two web pages, the algorithm will follow the following procedures. The procedures for link-farm identification are explained with the help of the diagram in Figure 6. Consider that our web crawler identified all the adjacent nodes and found all the reciprocal links between the web page W$pi$ and all its neighbouring nodes so that a web graph, shown in Figure 6, is obtained.

Initially, it will create the adjacency matrix. If adjacent nodes also have adjacent nodes other than the nodes adjacent to W$pi$, they will be ignored and will not be added to the adjacency matrix. For instance, in the web graph in Figure 6, W$pi$ has *A*, *B*, *C*, *and D* adjacent nodes, so only these nodes will be considered for the adjacency matrix; the nodes adjacent to *A*, namely, *H* and *G*, will be ignored. Similarly, if a web page links back to itself, it will also be ignored and will be represented with *0* in the adjacency matrix. For instance, in the example above, node *B* is linking back to itself and is, therefore, ignored. Accordingly, the following will be the adjacency matrix for the web graph in Figure 6.

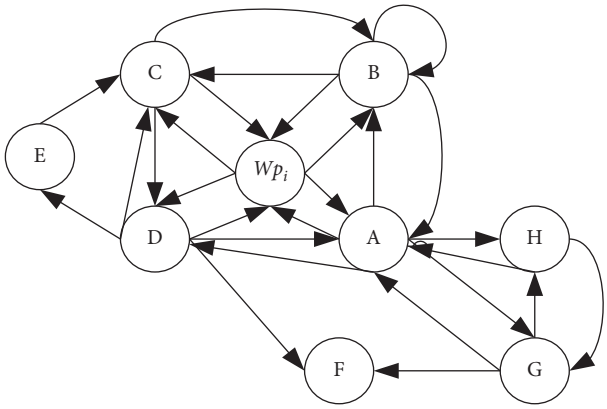



Figure 6: Web graph representing the websites and links.

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 0 \end{pmatrix}. \quad (5)$$

If $a_{1,2} = a_{2,1} = 1$, this means there exists a reciprocal link between W*p1* and W*p2*; 1 in the adjacency matrix represents the backlink and zeros represent no link. To confirm that every web page in the group links to all other pages of the group, the program will replace all 0s with 1s and all 1s with 0s in the adjacency matrix. The following will be the new adjacency matrix.

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}. \quad (6)$$

Finally, the resultant adjacency matrix will be checked. If the resulting matrix is an identity matrix, each web page in the group links to every other page in the group, and W*pi* is a member of a link farm and will be marked as spam. If the resultant matrix is not an identity matrix, the link-farm module will mark every adjacent node of W*pi* as W*p2* W*p3*,..., W*pn* and will apply the same link-farm detection technique on every W*pn*. If any adjacent node W*pn* is a part of a link farm, W*pi* is not directly connected to any link farm, but W*pi* is a part of the link pyramid. Wpi will be marked as spam.

## 7. Combined Approach for Content- and Link-Based Spamdexing Detection

In this section, we discuss a combined approach used for the identification of spam web pages. We combined the content-based framework with a link-based framework so that the final output of the content-based approach will be the input of the link-based approach. During our experiments and analysis, we observed that web spammers are practising different spamdexing strategies. Some spammers only focused on content-based spamdexing, some targeted the link-based spamdexing techniques, and several were involved in both types of spamdexing at the same time. Therefore, the combined approach for spamdexing detection is better in that it can check any web page for both techniques of spamdexing. It is possible that a page is involved in content-based spamdexing but does not involve link-based spamdexing techniques and vice versa. For instance, if a technique is designed for content-based spamdexing detection, the technique can only detect the pages involving content-based spamdexing. It will mark the web page as clean, which involves link-based spamdexing. Unfortunately, several researchers in the field focused on a single detection technique only, and a few worked on a combined approach. The working mechanism of the combined approach is described as follows:

(1) The combined approach will accept a web page as input and check the web page for content-based spamdexing using the five different methods discussed in Section 6. If the web page involves content-based spamdexing, it will mark the page as spam and the process will stop there. However, if the web page passed all the content-based spamdexing identification checks, the URL of the web page will be forwarded to the link-based spamdexing detection section for further analysis.

(2) There are possibilities that a web page does not involve content-based spamdexing but involves link-based spamdexing only. Now, whether the web page is engaged in link-based spamdexing will be checked by scanning it using link-based spamdexing methods. If the page is involved in link spamming, then it will be marked as spam; otherwise, it will be marked as nonspam. Figure 7 shows the complete framework of the proposed combined approach.

## 8. Results

To conduct experiments, our verification set consists of randomly chosen web pages labelled as spam and nonspam. These web pages are selected from the dataset obtained through experiments in Section 4, using WEBSPAM-UK2006 and WEBSPAM-UK2007 datasets. These datasets are well known and are most suitable for web-spam detection due to the following properties:

(1) The datasets are a mixture of spam and nonspam web pages practising several different web-spamming techniques.

(2) All the researchers in the field can freely access the datasets from the official website and use them as a benchmark measure to detect spam web pages.

(3) The sample web pages in the datasets are random and uniform.

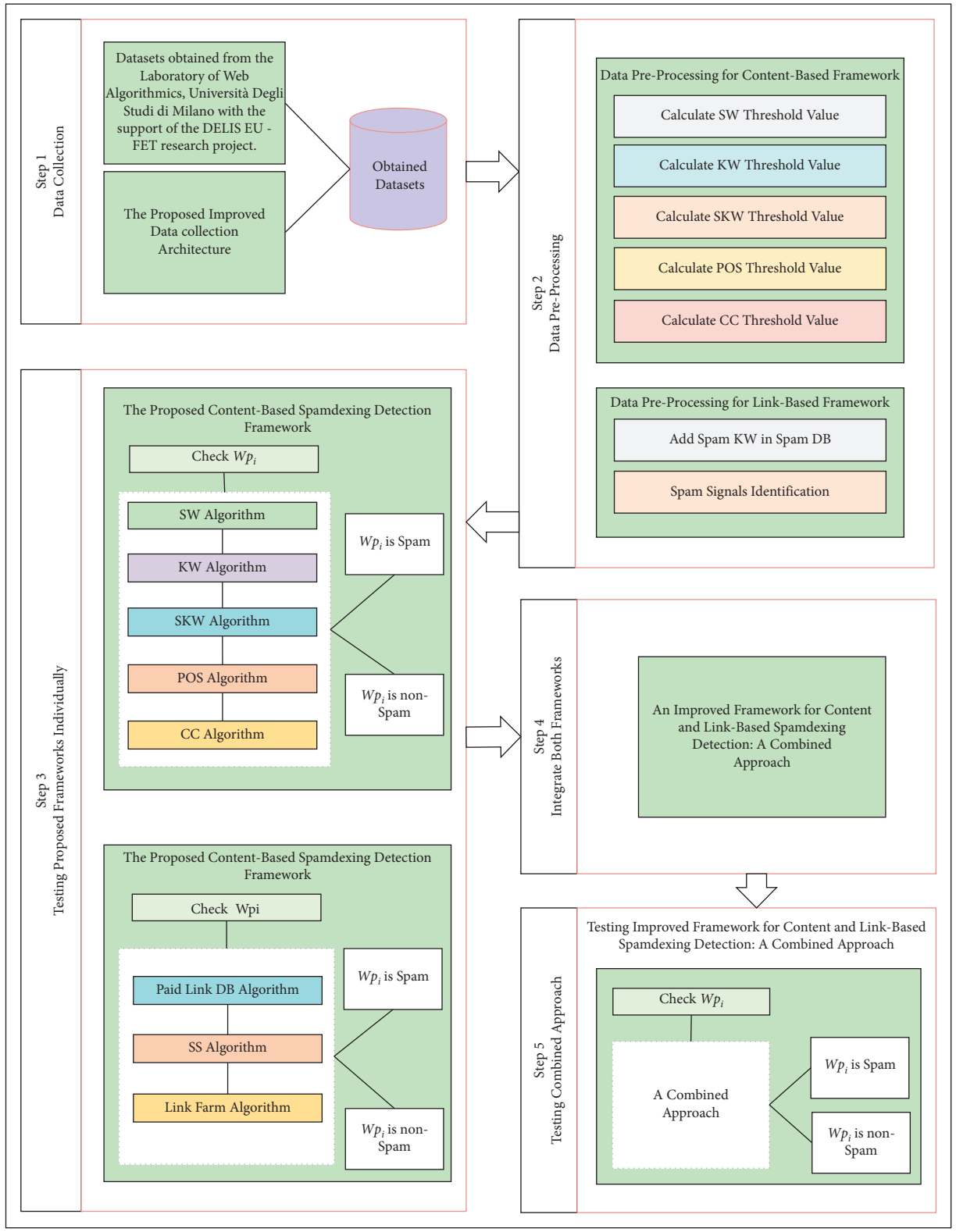



Figure 7: The proposed framework for the combined approach.

(4) The datasets consist of several different kinds of spam pages created using many types of web-spamming techniques.

(5) The web pages are split into training and testing sets with both spam and nonspam labels. Therefore, these datasets can be used to perform experiments and to

detect both content- and link-based web-spamming techniques.

(6) We used these datasets to obtain the optimised threshold values for our proposed combined approach.

There are 11,402 hosts in WEBSPAM-UK2006, but only 7,473 are labelled, while WEBSPAM-UK2007 contains 114,529 hosts in total, and 6,479 are labelled. The following preprocessing steps are performed to obtain a dataset of five thousand web pages:

(1) The web pages labelled as spam or nonspam by real humans are considered

(2) We further filtered out the human label pages and only considered the web pages that are currently working/existing links

(3) Among the currently existing web pages, we only considered web pages having at least one KB of content

(4) Furthermore, the content of these web pages is extracted and stored in a text file format

To implement the proposed combined approach, Python was used, and a machine with 128 GB DDR3, 2x Intel Xeon E5-2670 V2 2.5 GHz 10 Core, and operating system Ubuntu 14.04 was used for the execution of algorithms. As F-measure is a standard approach for combining precision and recall, we used the F-measure to compare our work with other similar related works and to evaluate the proposed framework. Our proposed combined approach for spamdexing detection achieved the results shown in Table 3.

Table 3: Performance evaluation of improved framework.

| Technique | Precision (%) | Recall (%) | F-measure (%) |
|---|---|---|---|
| Content-based | 78.3 | 75.6 | 77 |
| Link-based | 73.5 | 69.7 | 71.5 |
| Combined | 81.2 | 78 | 79.6 |

Table 4: Comparison of the proposed combined approach with other standard techniques.

| Combined techniques | Precision | Recall | F-measure |
|---|---|---|---|
| The proposed framework | 81.2 | 78 | 79.6 |
| Roul et al. | 72.9 | 77.6 | 75.2 |
| Dia et al. | 65 | 44.3 | 52.7 |
| Benczúr et al. | 67.1 | 76.7 | 71.6 |
| Egele et al. | 51.2 | 35.6 | 41.9 |
| Becchetti et al. | 68.8 | 76.2 | 72.3 |

## 9. Comparison with Existing Approaches

We compared our results with the following existing techniques. The comparison results in Table 4 clearly show that the proposed combined approach surpasses other spam-detection techniques.

*9.1. The Proposed Framework vs. Roul et al. [13].* We compared the results of our framework with the work done by Roul et al. [13]. For the identification of spam web pages, they proposed a combined approach for content- and link-based spamdexing. To identify the content-based spamdexing, they used part of speech ratio test and term density, while for link-based spamdexing detection, they used the personalised page ranking to categorise the websites as spam and nonspam. For their experiments, they used the WEBSPAM-UK2006 dataset. Finally, they combined both of their spamdexing detection techniques to achieve 72.9% precision and 75.2% F-measure, listed in [47], which are significantly less than those found in our results.

*9.2. The Proposed Framework vs. Dia et al. [47].* Next, we compared our empirical results with Dia et al. [47]. For spam identification, they considered the historical web page information in their work. For improvement in spam classification, they used content features from the old version of pages. They combined the classifiers based on the temporal characteristics and the current page content by applying supervised learning techniques. With their method, they extracted several temporal features from archival copies of the web presented by Internet Archive's Wayback Machine. For their experiments, they used the WEBSPAM-UK2007 dataset. Dai et al. achieved an F-measure of about 52.7 and a precision of approximately 65, which are less than what we achieved [47].

*9.3. The Proposed Framework vs. Benczúr et al. [48].* Further, we compared the results obtained from our framework with the work of Benczúr et al. [48]. They introduced several features for web-spam filtering based on the appearance of keywords. The keywords they used were either highly advertised or frequently used in web spamming. Their web-spam features used for spamdexing detection include "the distribution of Google AdSense ads over pages of a site," "Google AdWords advertisement keywords suggestions," "the Yahoo Mindset classification of web pages," and "online commercial intention." To perform their experiments, they used the WEBSPAM-UK2006 dataset. They achieved an F-measure of 71.6% with a precision of 67.1, which are less than the results we obtained using our framework.

*9.4. The Proposed Framework vs. Egele et al. [44].* Moreover, we compared our work with Egele et al. They introduced a technique to recognise the spam web pages on the search engine result pages. In the first step, the importance of different page features is determined by them to rank higher in the search engine result pages. Then, they developed a classification technique based on the page features to identify the spam web pages. For their experimental work, they used the J48 classifier. Finally, they listed the results of their experiments in Table 3 of Egele et al. [44]. They have achieved 51.2% precision and 41.9% F-measure, which are less than the results we obtained through our experiments.

*9.5. The Proposed Framework vs. Becchetti et al. [49].* Finally, we compared our work with the results of Becchetti et al. [49]. Their technique used different content-based features, for instance, redirection to other pages, presence of unrelated keywords in URL, hidden text proportion on web pages, and duplicate content. Similarly, they plotted a web graph for link-based features and identified several page ranks like "trust rank", "degree-based measure," and "truncated page rank" of pages to identify the spam. As a base classifier, they have used the C4.5 decision tree. To perform their experiments, they used WEBSPAM-UK2002 and WEBSPAM-UK2006 datasets [49]. By combining both content- and link-based techniques, they achieved 66.8% precision and 72.3% F-measure, significantly less than those found in our results.

The comparisons above clearly show that our proposed framework for content- and link-based spamdexing detection is better than the above-mentioned five techniques.

## 10. Conclusion

Web spamming is a huge issue for people searching for information on the Internet using search engines. It also causes significant financial losses. Researchers have proposed many web-spam-detection methods to overcome this issue, but until now there is no single effective detection method that can detect all types of web spam available on the World Wide Web with high accuracy. In this research paper, we presented an improved combined approach for content- and link-based web-spam detection. We explored five different techniques for a content-based framework: stopword density, keyword density, spam keyword density, part of speech ratio, and copied-content test. Similarly, our link-based spamdexing detection framework used a paid-link database and spam-signal and link-farm identification technique with collaborative detection to detect a nonspam or spam page. We have used two datasets, WEBSPAM-UK2006 and WEBSPAM-UK2007, and the dataset obtained through our experimental work. An excellent and very promising F-measure of 79.6% compared to other existing approaches shows our framework's robustness. We will extend this research work by adding more content- and link-based spamdexing detection features to this framework. We believe that we can enhance our framework's power to identify a wide range of spam web pages by combining techniques and using more features.

## Data Availability

WEBSPAM-UK2006 and WEBSPAM-UK2007 benchmark datasets are used in this paper.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

The authors would like to thank the Office of Research, Innovation, Commercialization and Consultancy Management (ORICC), Universiti Tun Hussein Onn Malaysia (UTHM), and Ministry of Higher Education (MOHE), Malaysia, for financially supporting this Research under Tier-1 Research Grant vote no. H938.

## References

[1] Z. Gyongyi and H. Garcia-Molina, "Web spam taxonomy," in *Proceedings of the First International Workshop on Adversarial Information Retrieval on the Web (AIRWeb 2005)*, Chiba, Japan, 2005.

[2] M. R. Henzinger, R. Motwani, and C. Silverstein, "Challenges in web search engines," in *Proceedings of the 18th International Joint Conference on Artificial Intelligence*, Acapulco, Mexico, 2003.

[3] A. Ntoulas, M. Najork, M. Manasse, and D. Fetterly, "Detecting spam web pages through content analysis," in *Proceedings of the 15th International Conference on World Wide Web -WWW '06*, p. 83, Scotland, UK, 2006.

[4] N. Z. J. MCA and P. Prakash, "Document content based web spam detection using cosine similarity," *International Journal of Intelligence Research (IJOIR)*, vol. 7, 2016.

[5] A. Shahzad, N. Mohd Nawi, E. Sutoyo et al., "Search engine optimization techniques for malaysian university websites: a comparative analysis on google and bing search engine," *International Journal on Advanced Science, Engineering and Information Technology*, vol. 8, no. 4, pp. 1262–1269, 2018.

[6] Y. Li, X. Nie, and R. Huang, "Web spam classification method based on deep belief networks," *Expert Systems with Applications*, vol. 96, pp. 261–270, 2018.

[7] Z. Guo and Y. Guan, "Active probing-based schemes and data analytics for investigating malicious fast-flux web-cloaking based domains," in *Proceedings of the 2018 27th International Conference on Computer Communication and Networks (ICCCN)*, pp. 1–9, Hangzhou, China, August 2018.

[8] N. Spirin and J. Han, "Survey on web spam detection," *ACM SIGKDD Explorations Newsletter*, vol. 13, no. 2, pp. 50–64, 2012.

[9] S. Robertson, H. Zaragoza, and M. Taylor, "Simple BM25 extension to multiple weighted fields," in *Proceedings of the Thirteenth ACM International Conference on Information and Knowledge Management*, pp. 42–49, Washington, DC, USA, November 2004.

[10] C. Zhai, "Statistical language models for information retrieval," *Synthesis Lectures on Human Language Technologies*, vol. 1, no. 1, pp. 1–141, 2008.

[11] G. Salton, A. Wong, and C. S. Yang, "A vector space model for automatic indexing," *Communications of the ACM*, vol. 18, no. 11, pp. 613–620, 1975.

[12] N. El-Mawass and S. Alaboodi, "Data quality challenges in social spam research," *Journal of Data and Information Quality*, vol. 9, no. 1, pp. 1–4, 2017.

[13] R. K. Roul, S. R. Asthana, M. Shah, and D. Parikh, "Detecting spam web pages using content and link-based techniques," *Sadhana*, vol. 41, no. 2, pp. 193–202, 2016.

[14] B. Davison, "Recognising nepotistic links on the web," *Artificial Intelligence Web Search*, pp. 23–28, 2000.

[15] J. Piskorski, M. Sydow, and D. Weiss, "Exploring linguistic features for web spam detection: a preliminary study," in *Proceedings of the 4th International Workshop on Adversarial Information Retrieval on the Web*, pp. 25–28, Beijing, China, April 2008.

[16] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent dirichlet allocation," *Journal of Machine Learning Research*, vol. 3, no. Jan, pp. 993–1022, 2003.

[17] Y. Tian, G. M. Weiss, and Q. Ma, "A semi-supervised approach for web spam detection using combinatorial feature-fusion," in *Proceedings of the Graph Labelling Workshop and Web Spam Challenge at the European Conference on Machine Learning and Principles and Practice of Knowledge Discovery*, pp. 16–23, New York, NY, USA, September 2007.

[18] D. Fetterly, M. Manasse, and M. Najork, "Detecting phrase-level duplication on the world wide web," in *Proceedings of the 28th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 170–177, Salvador, Brazil, August 2005.

[19] D. Fetterly, M. Manasse, and M. Najork, "On the evolution of clusters of near-duplicate web pages," in *Proceedings of the IEEE/LEOS 3rd International Conference on Numerical Simulation of Semiconductor Optoelectronic Devices (IEEE Cat. No. 03EX726)*, pp. 37–45, IEEE. Tokyo, Japan, 2003.

[20] D. Fetterly, M. Manasse, and M. Najork, "Spam, damn spam and statistics: using statistical analysis to locate spam web pages," in *Proceedings of the 7th International Workshop on the Web and Databases: colocated with ACM SIGMOD/PODS 2004*, pp. 1–6, Paris, France, June 2004.

[21] A. Z. Broder, S. C. Glassman, M. S. Manasse, and G. Zweig, "Syntactic clustering of the web," *Computer Networks ISDN System*, vol. 29, no. 8, pp. 1157–1166, 1997.

[22] A. Z. Broder, "Some Applications of Rabin's Fingerprinting Method," in *Sequences II*, pp. 143–152, Springer, New York, NY, USA, 1993.

[23] M. O. Rabin, "Fingerprinting by random polynomials," *Center for Research in Computing Technology*, Report TR-15-81, Harvard University, Cambridge, MA, USA, 1981.

[24] M. Erdélyi, A. Garzó, and A. A. Benczúr, "Web spam classification: a few features worth more," in *Proceedings of the 2011 Joint WICOW/AIRWeb Workshop on Web Quality*, pp. 27–34, Hyderabad, India, 2011.

[25] T. Urvoy, T. Lavergne, and P. Filoche, "Tracking web spam with hidden style similarity," in *Proceedings of the Second International Workshop on Adversarial Information Retrieval on the Web*, pp. 25–31, Seattle, WA, USA, August 2006.

[26] D. Hiemstra, *Language Models BT—Encyclopedia of Database Systems*, L. LIU and M. T. ÖZSU, Eds., Springer US, Boston, MA, USA, 2009.

[27] G. Mishne, D. Carmel, and R. Lempel, "Blocking blog spam with language model disagreement," *AIRWeb*, vol. 5, pp. 1–6, 2005.

[28] Z. Li, K. Zhang, Y. Xie, F. Yu, and X. Wang, "Knowing your enemy: understanding and detecting malicious web advertising," in *Proceedings of the 2012 ACM Conference on Computer and Communications Security*, pp. 674–686, Raleigh, NC, USA, October 2012.

[29] B. Stone-Gross, R. Stevens, A. Zarras, R. Kemmerer, C. Kruegel, and G. Vigna, "Understanding fraudulent activities in online ad exchanges," in *Proceedings of the 2011 ACM SIGCOMM Conference on Internet Measurement Conference*, pp. 279–294, Berlin, Germany, November 2011.

[30] A. Zarras, A. Kapravelos, G. Stringhini, T. Holz, C. Kruegel, and G. Vigna, "The dark alleys of madison avenue: understanding malicious advertisements," in *Proceedings of the 2014 Conference on Internet Measurement Conference*, pp. 373–380, Vancouver, Canada, November 2014.

[31] S. Antonatos, I. Polakis, T. Petsas, and E. P. Markatos, "A systematic characterization of IM threats using honeypots," in *Proceedings of the ISOC Network and Distributed System Security Symposium (NDSS)*, San Diego, CA, USA, February 2010.

[32] J. L. Ledford, *Search Engine Optimisation Bible*, John Wiley & Sons, Hoboken, NY, USA, 2015.

[33] P. T. Metaxas and J. DeStefano, "Web spam, propaganda and trust," in *Proceedings of the First International Workshop on Adversarial Information Retrieval on the Web*, Chiba, Japan, May 2005.

[34] S. Sethi and A. Dixit, "A novel page ranking mechanism based on user browsing patterns," *Software Engineering*, Springer, Berlin, Germany, pp. 37–49, 2019.

[35] D. Sharma, R. Shukla, A. K. Giri, and S. Kumar, "A brief review on search engine optimisation," in *Proceedings of the 2019 9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, pp. 687–692, Noida, India, 2019.

[36] M. D. Oskuie and S. N. Razavi, "A survey of web spam detection techniques," *International Journal of Computer Applications Technology and Research*, vol. 3, no. 3, pp. 180–185, 2014.

[37] K. Bharat and M. R. Henzinger, "Improved algorithms for topic distillation in a hyperlinked environment," *ACM SIGIR Forum*, vol. 51, no. 2, pp. 194–201, 2017.

[38] S. Nomura, S. Oyama, T. Hayamizu, and T. Ishida, "Analysis and improvement of hits algorithm for detecting web communities," *Systems and Computers in Japan*, vol. 35, no. 13, pp. 32–42, 2004.

[39] J. J. Whang, Y. S. Jeong, I. S. Dhillon, S. Kang, and J. Lee, "Fast asynchronous anti-trust rank for web spam detection," in *Proceedings of the WSDM Worskshop MIS2*, Marina Del Rey, CA, USA, 2018.

[40] J. Abernethy, O. Chapelle, and C. Castillo, "Graph regularisation methods for web spam detection," *Machine Learning*, vol. 81, no. 2, pp. 207–225, 2010.

[41] Q. Gan and T. Suel, "Improving web spam classifiers using link structure," in *Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web*, pp. 17–20, Banff, Canada, May 2007.

[42] V. M. Prieto, M. Álvarez, and F. Cacheda, "SAAD, a content based web spam analyser and detector," *Journal of Systems and Software*, vol. 86, no. 11, pp. 2906–2918, 2013.

[43] E. Amitay, D. Carmel, A. Darlow, R. Lempel, and A. Soffer, "The connectivity sonar: detecting site functionality by structural patterns," in *Proceedings of the Fourteenth ACM Conference on Hypertext and Hypermedia*, pp. 38–47, Nottingham, UK, August 2003.

[44] M. Egele, C. Kolbitsch, and C. Platzer, "Removing web spam links from search engine results," *Journal in Computer Virology*, vol. 7, no. 1, pp. 51–62, 2011.

[45] K. L. Goh, R. K. Patchmuthu, and A. K. Singh, "Link-based web spam detection using weight properties," *Journal of Intelligent Information Systems*, vol. 43, no. 1, pp. 129–145, 2014.

[46] A. Shahzad, H. Mahdin, and N. M. Nawi, "An improved framework for content-based spamdexing detection," *International Journal of Advanced Computer Science and Applications(IJACSA)*, vol. 11, no. 1, 2020.

[47] X. Dai, B. D. Davison, and X. Qi, "Looking into the past to better classify web spam," in *Proceedings of the 5th International Workshop on Adversarial Information Retrieval on the Web*, pp. 1–8, Madrid, Spain, April 2009.

[48] A. Benczúr, I. Bíró, K. Csalogány, and T. Sarlós, "Web spam detection via commercial intent analysis," in *Proceedings of the 3rd International Workshop on Adversarial Information Retrieval on the Web*, pp. 89–92, Banff, Canada, May 2007.

[49] L. Becchetti, C. Castillo, D. Donato, S. Leonardi, and R. Baeza-Yates, "Web spam detection: link-based and content-based techniques," *European Integral Project Dynamic Evolution Large Scale Information System DELIS Proceeding Final Work.*vol. 222, pp. 99–113, 2008, https://www.chato.cl/papers/becchetti_2008_link_spam_techniques.pdf.

Hindawi
Complexity
Volume 2021, Article ID 8290978, 17 pages
https://doi.org/10.1155/2021/8290978

*Research Article*

# Approximate Inertial Manifold-Based Model Reduction and Vibration Suppression for Rigid-Flexible Mechanical Arms

**Lisha Xu,[1,2] Hua Deng,[1,2] Chong Lin,[1,2] and Yi Zhang[1,2]**

[1]*School of Mechanical and Electrical Engineering, Central South University, Changsha 410083, China*
[2]*The State Key Laboratory of High Performance and Complex Manufacturing, Changsha 410083, China*

Correspondence should be addressed to Yi Zhang; zhangyicsu@csu.edu.cn

Received 26 July 2021; Revised 25 August 2021; Accepted 12 September 2021; Published 31 October 2021

Academic Editor: Hang Su



The dynamic characteristics of the mechanical arm with a rigid-flexible structure are very complex. The reason is that it is a complex DPS (distributed parameter system) with infinite dimension and nonlinearity in essence due to the rigid-flexible coupling. So, accurately positioning and controlling the rigid-flexible mechanical arms could be difficult. Therefore, a model reduction method of rigid-flexible mechanical arms based on the approximate inertial manifold is put forward. To repress the residual vibration of the end of the mechanical arm, a feedforward control strategy is designed. The high-dimensional solution of the vibration equation of the rigid-flexible mechanical arms is projected into the complete space composed of orthogonal decomposition modes. By using Galerkin's method, the system is simplified and the approximate solution is obtained through the interaction between high-order and low-order modes. The truncated finite mode is also used to construct a lowest-order dynamic model on the basis of approximate inertia manifold. Given the reduced-order rigid-flexible mechanical arms dynamic model, dynamic response analysis is conducted to optimize the target position error and end residual vibration. A limited number of sinusoidal signals approximately combine the input signal, by using the particle swarm optimization algorithm to optimize the input signal, and the amplitude of the sinusoidal signal is corrected. The simulation results depict the superiority of the proposed method, which greatly suppresses the end residual vibration of the mechanical arm and realizes the accurate positioning of the end of the mechanical arm. In addition, the hardware experimental device of the rigid-flexible mechanical arms is constructed, and the experimental verification of the above method is put into effect. The simulation results of angular displacement and end vibration of the reduced model are accordant which is shown by the experimental results of the hardware platform.

## 1. Introduction

A flexible mechanical arm has the advantages of a high weight ratio, low energy consumption, high efficiency, less inertia, and safe operation. At present, most researchers are focusing on flexible mechanical arms. However, less research involves rigid-flexible mechanical arms [1, 2]. Rigid-flexible mechanical arms have stronger operability and versatility, higher payload, and more low total cost than traditional flexible mechanical arms [3, 4]. They have been widely used in scientific fields, such as medicine and space [5–8]. Just like the cyber-physical systems, they also have more nonparametric uncertainties [9] and complex factors, such as dynamic/static friction, joint flexibility, structural resonance modes, and drive dynamics, than the traditional flexible mechanical arms. Given the complex coupling relationship between the state variables of the rigid-flexible mechanical arms [10], the difficulty of modeling and vibration suppression is greatly increased.

So far, many research studies on the modeling and deformation theory have been done by many researchers and achieved a series of results [11–13], for example, using finite element method, virtual modal method, and other methods to model flexible mechanical arm [14–17]. Various dimensionality reduction methods are used in this field. Researchers transform nonlinear partial differential equations from infinite dimension to finite dimension through models. Some of the commonly used methods include the

traditional Galerkin's method (TGM) [18], inertial manifold method [19], intrinsic orthogonality decomposition method [20], and central manifold method [21]. When these methods are applied to nonlinear dynamic analysis, the question of how many modes should be included in the solution will arise. When insufficient modes are considered, the simplified system may not be topologically equivalent to the original system. By contrast, excessive considerations will bring difficulties to system analysis. In this way, the sensitivity of the solution to disturbance is the key factor affecting the accurate modeling of the nondynamic model.

This disturbance may come from the numerical error and mode truncation of the above method, and this slight disturbance may affect the system performance [22, 23]. The approximate inertial manifold (AIM) has been widely used in recent years. The asymptotic properties of some infinite-dimensional dissipative dynamic systems will evolve. They turn into a finite-dimensional compact set containing nonlinear phenomena called global attractors [24, 25]. To the best of the researchers' knowledge, no research has introduced the AIM into flexible mechanical arms or rigid-flexible coupling mechanical arms for dimensionality reduction.

Given that the flexible mechanical arm has underactuated variables, the controller for the underactuated variables of the system is difficult to design directly to converge in the study of position control. These difficulties lead to the residual vibration of the end of the mechanical arm in the movement process, which cuts down the accurate positioning of the end of the mechanical arm. Therefore, the residual vibration generated during the movement of the flexible mechanical arm needs to be suppressed, which is the key to achieving high-precision position control. At present, the main vibration suppression methods are active control and passive control [26, 27]. Passive control is mainly achieved by adding a damping structure to the flexible structure. As early as the 1980s, the authors in [28] discovered that damping materials could be used to speed up the attenuation of residual vibration. They then proposed to increase the damping of the robot to achieve structural vibration suppression. However, this will change the structure of the mechanical arms. Active control includes feedforward control and closed-loop control [29–31]. Closed-loop control is the measurement of vibration signals by sensors and feedback to the control system. The control system outputs instructions to control the end vibration, including PID control, sliding mode control, robust control, state feedback control, fuzzy control, adaptive control, and neural network control [32–43]. The closed-loop control method mentioned above has a certain effect on the vibration restrain of the end of the flexible mechanical arm, but it needs to increase the sensor to measure the end deformation.

Moreover, constructing a complex feedback loop is necessary for the control, which makes its implementation in practical applications quite difficult. Feedforward control does not need to add additional sensors, and the control structure is very simple. Among them, input shaping is the most widely studied [44, 45]. Yet, it needs to design the input shaping filter under some constraints. Many parameters are also needed to be considered. For example, one should consider the period, the number of pulses, and the amplitude of the pulse signal. To address the above problem, the authors in [46] proposed a feedforward control strategy that does not consider the input shaping filter to restrain the remains vibration of the long and flexible mechanical arm. They expressed the starting and braking part of the input signal as a Fourier series and optimized the coefficients of the Fourier series with the goal of zero remains vibration. Nevertheless, their methods only optimize the starting and braking parts of the input signal. The coefficients of the Fourier series were also modified through trial and error. Moreover, the selection of the higher mechanical harmonics is undetermined, and whether precise positioning can be achieved is not considered.

On the basis of the above problems, the infinite-dimensional solution of the vibration equation of the rigid-flexible mechanical arms is projected into the complete space formed by the POD (proper orthogonal decomposition) mode by using the AIM [47–49] in this study. According to AIM, the vibration equation of an infinite-dimensional rigid-flexible mechanical arm can be changed into a finite-dimensional equation to extract nonlinear phenomena. Then, the TGM is used to solve the simplified system to obtain the interaction between the high-order mode and the low-order mode. The proposed method reduces the free degree of the system and improves the calculation efficiency without significantly losing the accuracy of the solution comparing with the TGM. Given the reduced-order rigid-flexible mechanical arms dynamic model, dynamic response analysis is conducted to optimize the target position error and end residual vibration. The input signal is approximately combined by a limited number of sinusoidal signals. The NF (natural frequency) of the system determines the highest frequency of the sinusoidal signal, to optimize the input signal, and the amplitude of the sinusoidal signal is corrected using the particle swarm optimization (PSO) algorithm. The simulation results show that this method cannot only achieve precise positioning but also effectively suppress the residual vibration generated during the movement of the rigid-flexible mechanical arms.

The main dedications of this study are as follows:

(1) Based on the previous detailed description of the approximate inertial manifold theory, it is applied to the nonlinear model of rigid-flexible mechanical arms, and the model is simplified to obtain its lowest dimensional approximate model. This dimension reduction method not only retains the main characteristics of the original system but also reduces the degree of freedom of the system without significantly reducing the accuracy of the solution. Therefore, it brings great convenience for system analysis and subsequent controller design.

(2) To restrain the residual vibration of the rigid-flexible mechanical arms, a feedforward control strategy based upon PSO is proposed. The input shaping filter does not need to be considered.

(3) The input signal is approximated by a combination of a finite number of sinusoidal signals, and the highest frequency of the sinusoidal signal is determined using the NF (natural frequency) of the system. The natural frequency of the system is obtained by physical analysis and simulation.

(4) According to the dynamic response of the rigid-flexible mechanical arms, the PSO algorithm is also used to adjust the amplitude of the sinusoidal signal to achieve the purpose of optimizing the input signal. While realizing the precise positioning of the rigid-flexible mechanical arms, it effectively suppresses the end residual vibration of the mechanical arm.

(5) For the lowest dimensional model of the rigid-flexible mechanical arms, the feedforward control method is used, and the input signal based on particle swarm optimization is used as the input of the model controller. The results show that this method improves the calculation efficiency of the model, realizes the accurate positioning of the end of the flexible arm, and suppresses the residual vibration.

The other parts of this paper are arranged as follows. Part II builds the dynamics high-order model of rigid and flexible mechanical arm. Part III reduces the model based on approximate inertial manifold to obtain the lowest-order output model. Part IV conducts dynamic simulation analysis. An improved feedforward strategy about vibration suppression based upon PSO to optimize input signals is put forward in Part V. Part VI summarizes the achievements.

*1.1. Dynamic Modeling of the Rigid-Flexible Mechanical Arms.* The structure of the research object is shown in Figure 1. The rigid mechanical arm is connected to the fixed support base, and the flexible mechanical arm is connected to the end of the rigid mechanical arm by clicking the drive shaft. The following assumptions are made in modeling: the flexible mechanical arm can bend freely, after elastic deformation, the intersecting surface is vertical to the deformation axis, and the longitudinal deformation and gravitational potential energy are ignored in Figure 1, and the cyber-physical system composition diagram is depicted in Figure 2.

As depicted in Figure 1, the parameters are as follows:

$l_1$ is the length of the rigid mechanical arm

$l_2$ is the length of the flexible mechanical arm;

$OXY$ is the inertial rectangular coordinate system with the base as the origin

$o_1x_1y_1$ is the local coordinate system of the rigid mechanical arm

$o_2x_2y_2$ is the local coordinate system of the flexible mechanical arm, and it turns with the turn of the mechanical arm

$\theta_1$ is the rotary angle of the coordinate system $o_1x_1y_1$relative to the coordinate system $OXY$

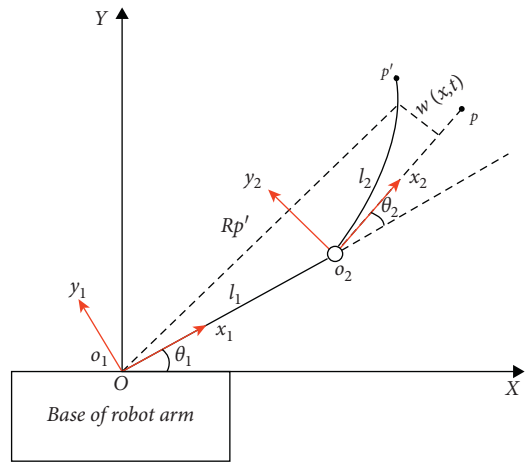



Figure 1: Structure diagram.

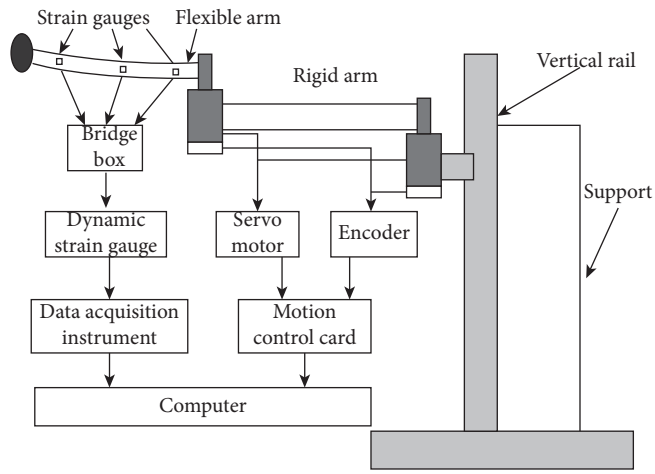



Figure 2: System composition diagram.

$\theta_2$ is the rotary angle of the coordinate system $o_2x_2y_2$ relative to the coordinate system $o_1x_1y_1$

According to the theory of Euler–Bernoulli beam and the results of reference [50], the vibration equation of the rigid-flexible mechanical arm can be obtained as follows:

$$\ell_2\ddot{w}(x,t) + 2\delta EI_2\dot{w}^{(4)}(x,t) + \ell_2 EI_2 w^{(4)}(x,t) = F(x,t), \tag{1}$$

where $w(x, t)$ represents shift of the flexible mechanical arm at the deformed position.

The dynamic boundary constraints of the flexible mechanical arm are as follows:

$$\begin{aligned} F(x,t) = {} & \rho_2\left[l_1 \sin\theta_2\dot{\theta}_1^2 + l_1\cos\theta_2\ddot{\theta}_1^2 - w(x,t)\left(\dot{\theta}_1+\dot{\theta}_2\right)^2\right] \\ & - M_p\left[l_1\cos\theta_2 w(l_2,t)\ddot{\theta}_1 - w(l_2,t)\left(\dot{\theta}_1+\dot{\theta}_2\right)^2\right. \\ & \left. + l_2\left(\dot{\theta}_1+\dot{\theta}_2\right) + l_1\dot{\theta}_1^2\sin\theta_2\right], \end{aligned} \tag{2}$$

where $F(x,t) = \ell_2[l_1 sin\theta_2\dot{\theta}_1^2 + l_1 \cos\theta_2\ddot{\theta}_1^2 - w(x,t)(\dot{\theta}_1+\dot{\theta}_2)^2] - M_p[l_1 \cos\theta_2 w(l_2,t)\ddot{\theta}_1 - w(l_2,t)(\dot{\theta}_1+\dot{\theta}_2)^2 + l_2(\dot{\theta}_1+\dot{\theta}_2) + l_1\dot{\theta}_1^2 sin\theta_2]$.

The influence of joint friction factors should be considered to describe the dynamic model accurately. The four common static friction models are as follows: the Coulomb viscous friction model, Coulomb friction model, static friction Coulomb viscous friction model, and Stribeck friction model. The Coulomb viscous friction model adopts a continuous function to approximate the Stribeck friction model, which is conducive to the realization of the control of each joint motor. Therefore, the Coulomb viscous friction model is used to obtain the friction torque of each joint of the n-degree-of-freedom mechanical arm as follows [51]:

$$\nabla^4 = \frac{\partial^4}{\partial x^4}, \tag{3}$$

where $i = 1,2$, $f_{ci}$ represents the Coulomb friction torque coefficient, and $f_{vi}$ represents the viscous friction torque coefficient. In this way, the modified dynamic model of rigid-flexible mechanical arms can be obtained:

$$\nabla^4 = \frac{\partial^4}{\partial x^4}. \tag{4}$$

The coefficients of each matrix are given in Appendix A.

*1.2. Model Reduction Based on Approximate Inertial Manifold.* The approximate inertial manifold is essentially a Lipshitz manifold. When the time is large enough, the solutions of all the original equations will be attracted into a neighborhood of the manifold because its existence does not need the spectral interval condition. Using the approximate inertial manifold, it is possible for us to approximate and study the long-time dynamic behavior of the original equation with a finite-dimensional ordinary differential system; this ordinary differential equation is the approximate inertial form to be studied in this paper.

Equation (4) is a system of partial differential integral equations with nonlinear and strong coupling characteristics, and the exact analytical solution cannot be obtained directly. Therefore, dimensionality reduction is needed for the infinite-dimensional nonlinear distributed parameter system.

The vibration equation (1) of the flexible mechanical arm can be rewritten as follows:

$$\frac{\partial^2 w(x,t)}{\partial t^2} = 2\delta\nabla^4\dot{w}(x,t) + \ell_2\nabla^4 w(x,t) + f(x,t), \tag{5}$$

where $\nabla^4 = \partial^4/\partial x^4$ and $f(x,t) = F(x,t)/\ell_2$ is equivalent to the external force.

The linear operator of Equation (5) can be defined as $Aw = \partial^4 w/\partial x^4$. Following the TGM, in Equation (5), $w(x,t)$ is the displacement at the deformation point $x$ of the flexible mechanical arm, namely:

$$w(x,t) = \sum_{i=1}^{\infty}\Gamma_i(x)q_i(t), \tag{6}$$

where $\Gamma_i(x)$ is a characteristic function and $q_i = q_i(t)$ is an undetermined variable. Substituting (6) into (5) and considering the boundary conditions in (2), we have [50]

$$\Gamma_i(x) = k_i(\sin\lambda_i x - \sinh\lambda_i x) - \cos\lambda_i x + \cosh\lambda_i x, \tag{7}$$

where $\lambda_i$ is the eigenvalue, is the modal order, and

$$k_i = \frac{\cosh\lambda_i l_2 + \cos\lambda_i l_2}{sinh\lambda_i l_2 + \sin\lambda_i l_2}. \tag{8}$$

The solution of equation (5) is expanded according to the orthogonal basis and using program of Galerkin, and the approximate solution of equation (5) can be obtained.

$$\begin{aligned} z = &\sum_{i=1}^{\infty} q_i(t)\cdot\Gamma_i(x) + EI_2\sum_{i=1}^{\infty} q_i(t)\frac{d^4\Gamma_i(x)}{dx^4} \\ &+ \frac{2\delta EI_2}{\ell_2}\sum_{i=1}^{\infty}\dot{q}_i(t)\frac{d^4\Gamma_i(x)}{dx^4} - f(x,t). \end{aligned} \tag{9}$$

Let $U = [q_1, q_2, \ldots, q_n]^T$ represent the orthogonal projection in the Hilbert space (H) onto the space strided over the first $n$ eigenfunctions of $A$ and $V = [q_{n+1}, q_{n+2}, \ldots, q_\infty]^T$ represent the rest, where $U$ is the low module subspace and $V$ is the high module subspace with infinite dimensions. It must be balance truncated, and the first **N** modes in all modes are used for the approximate solution of the model [22], which can be expressed by the inner product in **H**, and then the following projections are obtained:

$$(z,\Gamma_k) = 0, \quad k = 1,2,\ldots\ldots,n, \tag{10}$$

$$(z,\Gamma_k) = 0, \quad k = n+1, n+2, \cdots N. \tag{11}$$

According to the above projections, we have

$$\ddot{U} + A_1(\dot{U},\dot{V},U,V) = 0, \tag{12}$$

$$V + A_2(\dot{U},\dot{V},U,V) = 0, \tag{13}$$

where $A_1(\dot{U},\dot{V},U,V)$ and $A_2(\dot{U},\dot{V},U,V)$ are nonlinear functions of $\theta, q, \dot{\theta}, \dot{q}$.

By the spectral method combined with the Galerkin truncation criterion, the infinite-dimensional rigid-flexible mechanical arms system can be approximated to a third-dimensional model (i.e., $N = 3$) after truncation [50]. Then, equation (6) can be written as follows:

$$w(x,t) = \sum_{i=1}^{3} q_i(t)\Gamma_i(x). \tag{14}$$

Combining (4) and (12)–(14), a five-dimensional model of the rigid-flexible mechanical arms can be obtained as follows:

$$M\begin{bmatrix}\ddot{\theta}\\ \ddot{q}\end{bmatrix} + \begin{bmatrix}F_1(\theta,q,\dot{\theta},\dot{q})\\ F_2(\theta,q,\dot{\theta},\dot{q})\end{bmatrix} + \begin{bmatrix}E_1\dot{\theta}\\ E_2\dot{q} + Kq\end{bmatrix} = \begin{bmatrix}u\\ 0\end{bmatrix}, \tag{15}$$

where $M = M^T$ is the generalized symmetric inertia matrix, $\theta = (\theta_1, \theta_2)^T$ is the rotation angle vector of the mechanical

arm joint, $q = (q_1, q_2, q_3)^T$ represents the modal value of flexible mechanical arm deformation, $u$ is the input resultant torque to the joint, $\left[ F_1(\theta, q, \dot{\theta}, \dot{q}) \;\; F_2(\theta, q, \dot{\theta}, \dot{q}) \right]^T$ is a vector of nonlinear terms, $E_1 = \mathrm{diag}(\mu_1, \mu_2)$ is the positive and definite damping matrix, and $E_2 = \mathrm{diag}(\delta_1, \delta_2, \delta_3)$ is the structural damping matrix [52]. Equation (14) represents a five-dimensional strongly coupled nonlinear system based on generalized coordinates. The values of elements in the equation are shown in Appendix B.

Given that the nonlinear system (15) is a five-dimensional system with ten-order, it is a very high-order nonlinear system and difficult for system simulations, analyses, and controller design. Therefore, further truncation should be applied to (14). In equations (11) and (12), $U$ and $V$ contain low and high mode subspaces, and the modes in lower and higher subspaces can be thought of as "active" and "inactive," respectively [22]. Thus, a low-dimensional model may be obtained through the traditional Galerkin method by setting the first derivative of V and V in equation (12) to 0 then only the following equation can be obtained without considering equation (13)

$$\ddot{U} + A_1(\dot{U}, U) = 0. \tag{16}$$

Equation (15) means that the interaction between low modes and high modes is ignored when following the traditional Galerkin method. Nevertheless, on the basis of AIM theory, the higher-order modes of the flexible deformation displacement will soon decayed relatively to the lower-order modes under the action of viscous damping [22]. Thus, the derivatives of the higher-order mode V can be set to be zero, that is, $\ddot{V} \approx 0$ and $\dot{V} \approx 0$. From (13), a form of approximate inertial manifold can be built as

$$A_2(\dot{U}, U, V) = 0. \tag{17}$$

Substituting the solution of (17), i.e., $V \approx \Psi(\dot{U}, tU)$, into (16), (17) becomes

$$A_2(\dot{U}, U, \Psi(\dot{U}, U)) = 0. \tag{18}$$

Equation (18) means that the low modes are used to capture the behavior of the high modes in equation (13), and the interaction between low modes and high modes is maintained. Substituting $V \approx \Psi(\dot{U}, U)$ and $\dot{V} \approx 0$ into (12), we have

$$\ddot{U} + A_1(\dot{U}, U, \Psi(\dot{U}, U)) = 0. \tag{19}$$

Letting $n = 1$ in equation (12) and substituting (18) and (19) into (13), a model with the lowest dimension (i.e., three-dimension) can be obtained as follows:

$$M\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 00 \\ \ddot{q}_1 \end{bmatrix} + \begin{bmatrix} J_1(\theta, q, \dot{\theta}, \dot{q}) \\ J_2(\theta, q, \dot{\theta}, \dot{q}) \\ J_3(\theta, q, \dot{\theta}, \dot{q}) \\ J_4(\theta, q, \dot{\theta}, \dot{q}) \\ J_5(\theta, q, \dot{\theta}, \dot{q}) \end{bmatrix} + \begin{bmatrix} \mu_1 & 0 & 0 & 0 & 0 \\ 0 & \mu_2 & 0 & 0 & 0 \\ 0 & 0 & \delta_1 & 0 & 0 \\ 0 & 0 & 0 & \delta_2 & 0 \\ 0 & 0 & 0 & 0 & \delta_3 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{q}_1 \\ 0 \\ 0 \end{bmatrix} + K \begin{bmatrix} 0 \\ 0 \\ q_1 \\ q_2 \\ q_3 \end{bmatrix} = \begin{bmatrix} \tau_1 - \tau_{f1} \\ \tau_1 - \tau_{f2} \\ 0 \\ 0 \\ 0 \end{bmatrix}, \tag{20}$$

where

$$J_1(\theta, q, \dot{\theta}, \dot{q}) = -\left[\frac{1}{2}\ell_2 l_1 l_2^2 \sin\theta_2 + \ell_2 l_1 \cos\theta_2 \sum_{j=1}^{3} h_j(1) q_j(t)\right]\dot{\theta}_2^2 - \left[\ell_2 l_1 l_2^2 \sin\theta_2 + \ell_2 l_1 \cos\theta_2 \sum_{j=1}^{3} h_j(1) q_j(t)\right]\dot{\theta}_1\dot{\theta}_2$$

$$- M_p\left(l_1 \cos\theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) dx + 2 l_1 l_2 \sin\theta_2\right)$$

$$\dot{\theta}_2^2 - \left[2\ell_2 L_1 \sin\theta_2 \sum_{j=1}^{3} h_j(1)\dot{q}_j(t) - 2\ell_2 L_2 \sum_{j=1}^{3} q_j(t)\dot{q}_j(t)\right]\left(\dot{\theta}_1 + \dot{\theta}_2\right)$$

$$+ M_p\left(2 l_1 \cos\theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) dx + 2 l_1 l_2 \sin\theta_2\right)\dot{\theta}_1\dot{\theta}_2,$$

$$J_2(\theta, q, \dot{\theta}, \dot{q}) = \left[\frac{1}{2}\ell_2 l_1 l_2^2 \sin\theta_2 + \ell_2 l_1 \cos\theta_2 \sum_{j=1}^{3} h_j(1) q_j(t)\right]\dot{\theta}_1^2 + 2\ell_2 l_2\left(\dot{\theta}_1 + \dot{\theta}_2\right)$$

$$\left(\sum_{j=1}^{3} q_j(t)\dot{q}_j(t) + M_p\left(l_2^2 + \sum_{j=1}^{3} h_j(1)q_j(t)\cdot\sum_{j=1}^{3} q_j(t)\dot{q}_j(t)\right)\right)$$
$$+ M_p\left(l_1\cos\theta_2\sum_{j=1}^{3} h_j(1)q_j(t) + l_1 l_2\sin\theta_2\right)\dot{\theta}_1^2,$$
$$J_{i+2}(\theta, q, \dot{\theta}, \dot{q}) = -\ell_2\left(\dot{\theta}_1+\dot{\theta}_2\right)^2 q_i + \ell_2 l_2\dot{\theta}_1^2\sin\theta_2\cdot h_i(1), \quad i = 1, 2, 3. \tag{21}$$

And solution (18) gives the following:

$$q_2 = \tau_2 m_{33} - \left\{\left[2m_{33}\ell_2 l_2\left(\dot{\theta}_1+\dot{\theta}_2\right)q_1 - k_3 m_{23}\right]\dot{q}_1 + \ell_2 l_1\cos\theta_2\cdot h_1\dot{\theta}_1^2 m_{33} - \left(\delta_3 - \ell_2\left(\dot{\theta}_1+\dot{\theta}_2\right)^2 m_{23}\right)q_1\right\}$$
$$+ \ell_2 l_1\cos\theta_2\cdot h_3\dot{\theta}_1^2 m_{33}\cdot\frac{\ell_2 l_2\dot{\theta}_1^2\sin\theta_2\cdot h_3}{\ell_2\left(\dot{\theta}_1+\dot{\theta}_2\right)^2 - \delta_5} \tag{22}$$
$$+\left[k_2\dot{\theta}_2 + \delta_2\theta_2 + \frac{1}{2}\ell_2 l_1 l_2^2\sin\theta_2\dot{\theta}_1^2\right]m_{33} - \frac{\ell_2 l_2\dot{\theta}_1^2\sin\theta_2\cdot h_1 m_{23}}{\ell_2 l_1\cos\theta_2\cdot h_2\dot{\theta}_1^2 m_{33}},$$

$$q_3 = \tau_2 m_{33} - \left\{\left[2m_{33}\ell_2 l_2\left(\dot{\theta}_1+\dot{\theta}_2\right)q_1 - k_3 m_{23}\right]\dot{q}_1 + \ell_2 l_1\cos\theta_2\cdot h_1\dot{\theta}_1^2 m_{33} - \left(\delta_3 - \ell_2\left(\dot{\theta}_1+\dot{\theta}_2\right)^2 m_{23}\right)q_1\right\}$$
$$+ \ell_2 l_1\cos\theta_2\cdot h_2\dot{\theta}_1^2 m_{33}\cdot\frac{\ell_2 l_2 h_2\dot{\theta}_1^2\sin\theta_2}{\ell_2\left(\dot{\theta}_1+\dot{\theta}_2\right)^2 - \delta_4} \tag{23}$$
$$+\left[k_2\dot{\theta}_2 + \delta_2\theta_2 + \frac{1}{2}\rho_2 l_1 l_2^2\dot{\theta}_1^2\sin\theta_2\right]m_{33} - \frac{\rho_2 l_2\dot{\theta}_1^2\sin\theta_2\cdot h_1 m_{23}}{\rho_2 l_1\cos\theta_2\cdot h_3\dot{\theta}_1^2 m_{33}},$$

where $m_{23} = \ell_2\int_0^{l_2} x\cdot\Gamma_1(x)\mathrm{d}x + M_p l_2\int_0^{l_2}\Gamma_3(x)\mathrm{d}x$ and $m_{33} = \ell_2\int_0^{l_2}\Gamma_1(x)\mathrm{d}x$.

Finally, substitute $w(x,t) = \Phi_1(x)q_1 + \Phi_2(x)q_2 + \Phi_3(x)q_3$ and $w(x,t) = \Phi_1(x)q_1 + \Phi_2(x)q_2 + \Phi_3(x)q_3$ into equation (14) to get the low-dimensional output model of rigid-flexible mechanical arms as follows:

$$w(x,t) = \Gamma_1(x)q_1 + \Gamma_2(x)q_2 + \Gamma_3(x)q_3. \tag{24}$$

## 2. Dynamic Simulations

*2.1. Model Simulation and Analysis.* The initial state of the rigid and flexible mechanical arms is assumed in a horizontal position; the basic initial parameters are as follows: $\theta_1 = \theta_2 = q_1 = q_2 = q_3 = 0$. The joint input torque is $\tau_1 = \tau_2 = 0.412$ N/m. The length, width, and height of the flexible mechanical arm are 300 mm, 20 mm, and 15 mm, respectively. Other parameters are shown in Table 1. The relationship between modal and energy is obtained by simulation as depicted in Figure 3. The first-order mode accounts for 94.12% of the total energy, and the first three-order modes account for 99.78% of the total energy. This finding shows that the accuracy of the first three models is very high, but only taking the first mode is not enough to ensure the accuracy of the model.

*2.2. Free-Response Dynamic Simulations.* When the input torque is 0 and only the initial state is given, the response results of the approximate model obtained by the spectral method and the exact solution are compared. The initial time is $t_0 = 0$, and the end time is $t_f = 10$ s. The initial state of the system is $\theta_1 = 0$, $\theta_2 = \pi/4$, $\dot{\theta}_1 = \dot{\theta}_2 = 0$, $q_i = 0$, and $i = 1, 2, 3$. The initial value of the first derivative of all other generalized variables is 0. The fourth-order Runge–Kutta formula is used to solve the AIM-based low-order model and the TGM-based truncated third-order model. The simulation results are depicted in Figures 4–7.

In the free-response without input torque, $\theta_1(t)$ and $\theta_2(t)$ are finally stabilized, and the mutual influence between the rigid-flexible coupling can be obtained. The end deformation of the lowest-order model based on AIM is almost similar to that of the truncated flexible model based on TGM, as shown in Figures 4 to 6. They are all stabilized at zero in the end because of the features of structural damper of the mechanical arm itself. The simulation results of the first-order system and the third-order system after using the AIM approximation are the same, as shown in Figure 7. This conclusion can further prove that the AIM method is effective for model dimensionality reduction and can retain the main characteristics of the original system dynamics. Compared with the TGM-based truncation method, the lowest-order system based on AIM has a

Table 1: Other parameter values of the rigid-flexible mechanical arms.

| Parameter name | Rigid mechanical arm | Flexible mechanical arm |
|---|---|---|
| Length (m) | $L_1 = 0.33$ | $L_2 = 0.30$ |
| Moment of inertia (kg·m$^2$) | $J_1 = 0.0812$ | $J_2 = 0.138$ |
| Line density (kg·m$^3$) | | $\ell_2 = 0.4865$ |
| Elastic modulus (N·m$^2$) | | $EI_2 = 26.055$ |
| End mass (kg) | $M_t = 0.721$ | $M_p = 0.5$ |

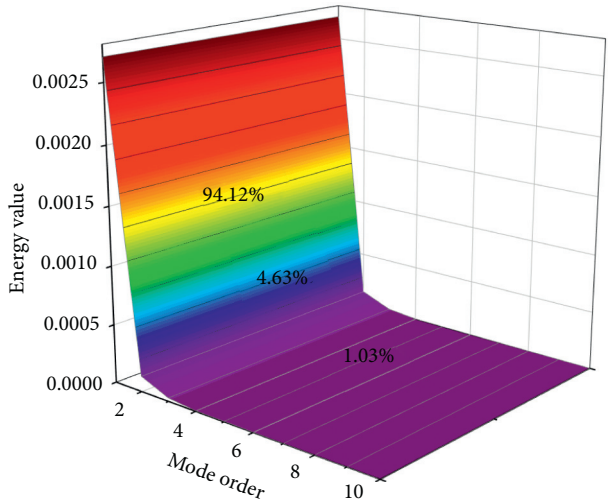



Figure 3: Relationship of modes and energy.

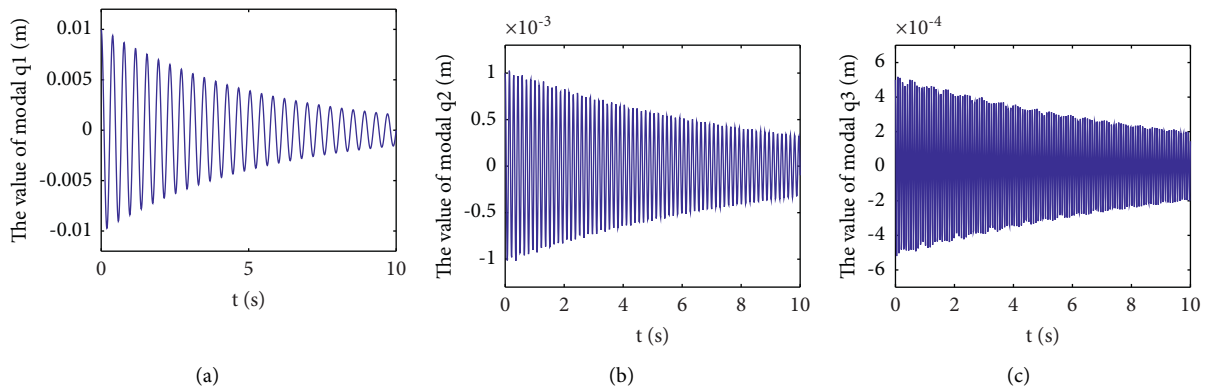



Figure 4: Generalized coordinate curve of flexible deformation: (a) $q1$; (b) $q2$; (c) $q3$.

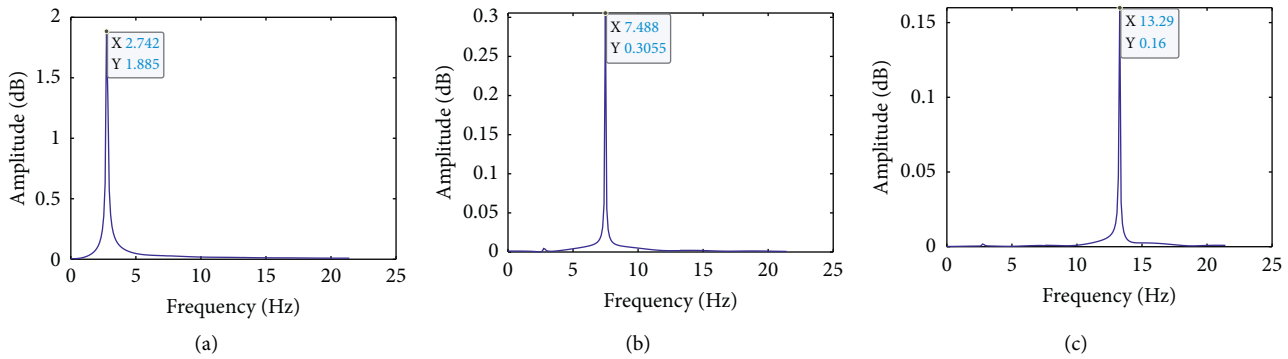



Figure 5: Spectrum of flexible generalized coordinates: (a) spectrum of $q_1$; (b) spectrum of $q_2$; (c) spectrum of $q_3$.

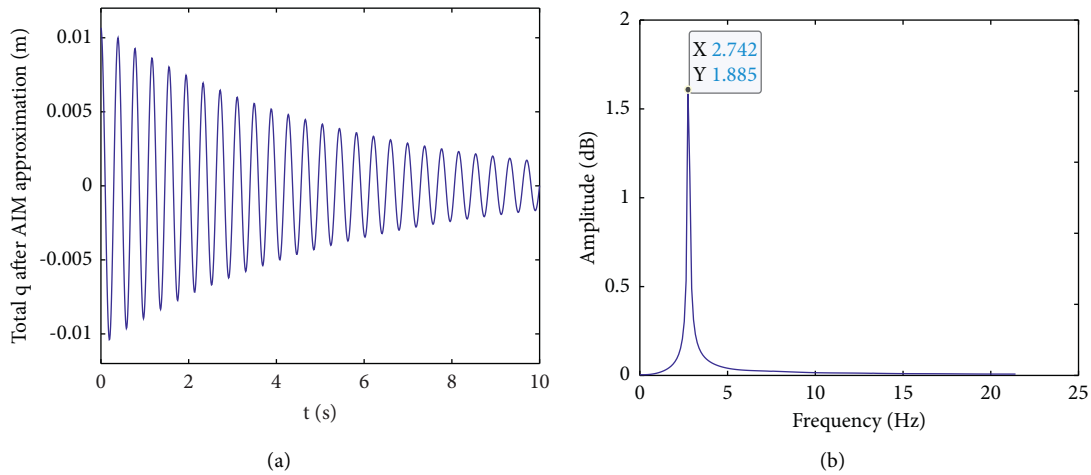



Figure 6: (a) Curve (q) of AIM; (b) spectrum (q) of AIM

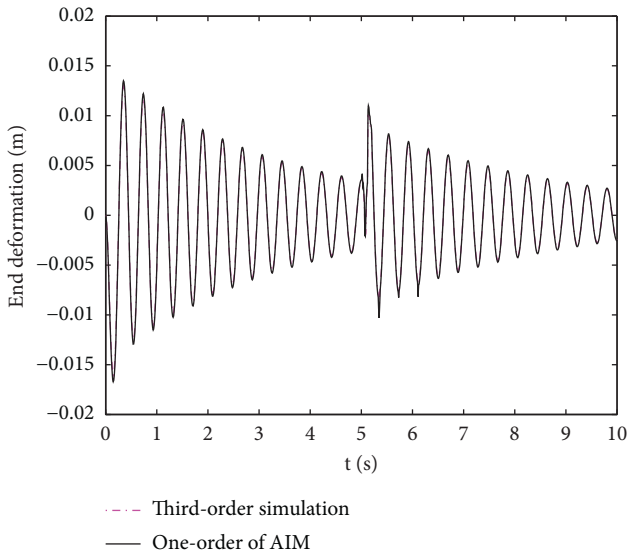



Figure 7: End vibrations with the first-order model and the third-order model.

smaller amount of calculation; according to this method, it is very convenient to analyze the system and design the controller.

## 3. Vibration Suppression Based on PSO

The AIM approximation model is used for numerical calculation in this study, and the function coefficients of the approximate input signal are optimized to suppress the end vibration. Firstly, a series of continuous sinusoidal functions approximate the input signal. The selection of the numbers of the approximation function is determined by the natural frequency, which is applied to the flexible mechanical arm to minimize the target position error and the displacement of end vibration. Then, using the PSO algorithm, the coefficients of each input function are optimized without considering the input shaping filter [44, 45]. In this way, the calculation efficiency of complex dynamic response is significantly improved, and the residual vibration is minimized when the target position is reached. The control structure diagram is depicted in Figure 8.

*3.1. Optimization-Based Vibration Suppression and Precise Positioning.* In Figure 1, the input signals acting on Joint 1 and Joint 2 are represented by trapezoidal signals, as shown in Figure 9. The expression is shown as follows.

$$u(t) = \begin{cases} \dfrac{At}{\varpi}, & 0 \le t < \varpi, \\ A, & \varpi \le t < (T - \varpi), \\ \dfrac{A(T-t)}{\varpi}, & (T - \varpi) \le t < T. \end{cases} \tag{25}$$

It is assumed that the operating cycle of rigid-flexible mechanical arms is $T$. Therefore, the signal $u(t)$ can be approximately expressed by a linear combination of infinitely many sinusoidal functions of different frequencies on $[0, T]$ and as depicted by the following formula:

$$u(t) = \frac{k_0}{2} + \sum_{n=1}^{\infty} k_n \sin w_n t, \tag{26}$$

where $k_n$ is the coefficient to be optimized and $w_n = 2\pi n/T$ is the angular frequency. Theoretically, the greater the value of $n$, the higher the fitting accuracy of the function. The time acting on the signals of Joint 1 and Joint 2 is taken as 4 s, and $d$ is taken as 0.1 s. The fundamental frequency is $2\pi/T$. When $n$ is taken as 11, $w_n$ is close to the natural frequency of the mechanical arms (as shown in Figure 6(b)). In this case, the end vibration will not be suppressed but increased. Thus, $n$ is selected as 9. The parameters of the input signal of Joint 1 to be optimized are $k_{10}$, $k_{11}$, $k_{12}$, ……, $k_{19}$. The parameters of the input signal of Joint 2 to be optimized are $k_{20}$, $k_{21}$, $k_{22}$, ……, $k_{29}$. These parameters are combined as follows:

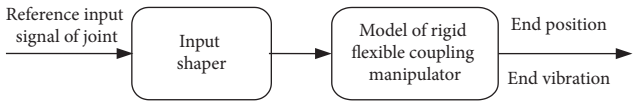



Figure 8: The control structure diagram of feedforward.

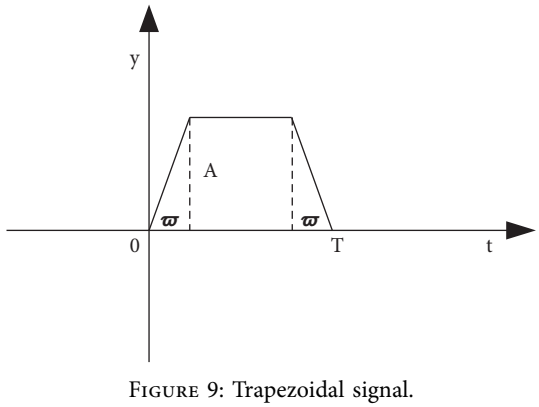



Figure 9: Trapezoidal signal.

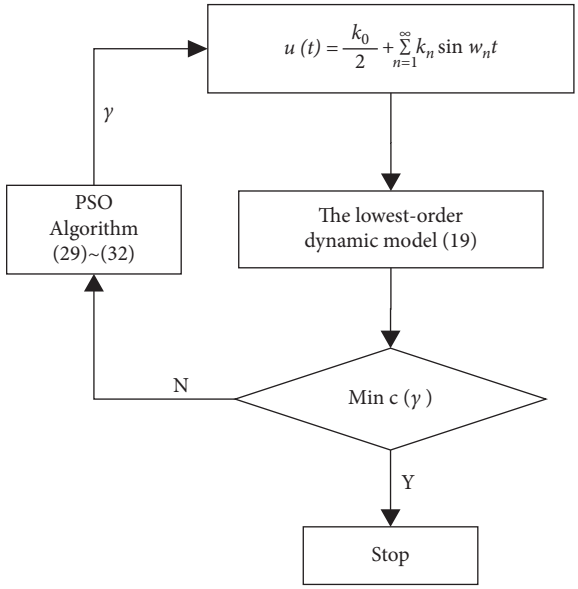



Figure 10: Optimization process diagram.

$$\gamma = [k_{10}, k_{11}, \ldots\ldots, k_{19}, k_{20}, k_{21}, \ldots\ldots, k_{29}] \in R^{20}. \quad (27)$$

When $\gamma$ is determined, the input signals applied to Joint 1 and Joint 2 can also be determined. Using equations (21)–(24), the relationship between the end vibration of the rigid-flexible mechanical arms and the applied torque of each joint is established. Therefore, the displacement of the end vibration of the flexible mechanical arm can be shown by the undetermined parameter $\gamma$ as follows:

$$w(t, \gamma) = \sum_{i=1}^{N} \Gamma_i(x) q_i(t, \gamma). \quad (28)$$

Moreover, the final target position of the mechanical arms after applying the input signal (25) is supposed to be $(\sigma_{tf}, \delta_{tf})$, and the end position error function of the flexible mechanical arm in the movement is built as follows:

$$e_p(t, \gamma) = \sqrt{\left[\sigma_{tf} - \sigma(t, \gamma)\right]^2 + \left[\delta_{tf} - \delta(t, \gamma)\right]^2}. \quad (29)$$

The objective of this study is to minimize the error of the mechanical arms reaching the target position and the residual vibration displacement at the end of the mechanical arms. The objective optimization function is established as follows:

$$c(\gamma) = \xi_1 |w(t, \gamma)| + \xi_2 e_p(t, \gamma), \quad (30)$$

where $\xi_1$ and $\xi_2$ are two weighted coefficients, and $\xi_1 + \xi_2 = 1$. The objective of the optimization is to minimize the objective function (30) by adjusting these parameters to be optimized. In this way, the precise positioning of the end position of the mechanical arms and the suppression of residual vibration can be transformed into solving the optimal value of each parameter in equation (27).

The PSO algorithm [46] is a study based on predation behavior of birds proposed by Kennedy et al. [46], because the algorithm is simple in structure and easy to implement. It is extensively applied since it can remember the current individual optimal and global optimal and requires only a few adjustment parameters [53]. It is a global optimization algorithm of probability type. The advantage of the nondeterministic algorithm is that the algorithm has more chances to solve the global optimal solution. It does not depend on the strict mathematical properties of the optimization problem itself. The PSO algorithm is applied to optimize the parameters in equation (27). In a searching space with $N$-dimension, let the quantity of particles be $m$, where the position of the $i$th particle is $x_i \in R^n$, and its velocity is $v_i \in R^n$. $p\text{best}_i \in R^n$ represents the best position of the $i$th particle. $g\text{best}_i \in R^n$ represents the best position of all particles. The velocity and position of the particles are updated as follows [46]:

$$\begin{cases} v_i^{n+1} = v_i^n + c_1 r_1 (p\text{best}_i^n - x_i^n)) + c_2 r_2 (g\text{best}_i^n - x_i^n), \\ x_i^{n+1} = x_i^n + v_i^{n+1}, \quad (n = 1, 2, \ldots, N, i = 1, 2, \ldots, m), \end{cases} \quad (31)$$

where $c_1$ and $c_2$ are the contraction factors and $r_1$ and $r_2$ are the arbitrary values between [−1, 1]. The detailed steps of parameter optimization based on particle swarm optimization for the rigid-flexible mechanical arms are as follows.

*Step 1.* In these solutions space, 20-dimensional particles are initialized as $\{x_1^l, x_2^l, \ldots, x_N^l\}$, and their corresponding flight velocities are recorded as $\{v_1^l, v_2^l, \ldots, v_N^l\}$.

*Step 2.* Each particle is substituted into equation (26), and the input signals of Joint 1 and Joint 2 are calculated. Then, the modal coordinates of the flexible mechanical arm are calculated according to equation (6). Subsequently, the displacement of end vibration is calculated according to equations (21)–(23) and (28). Finally, the fitness value of

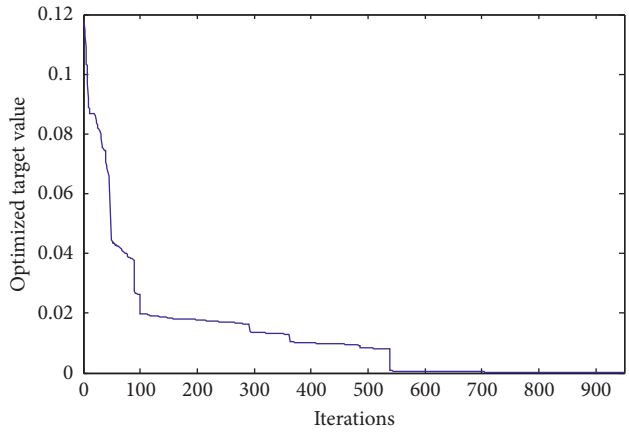



Figure 11: Optimization curve of the optimal objective function.

each particle is calculated according to the following equation:

$$f_i(l) = c\left(x_i^l\right), \quad i = 1, 2, ..., N. \tag{32}$$

*Step 3.* The local optimal values and the globally optimal values are initialized, and the local optimal and the global optimal positions are determined at the same time. If $l = 1$, then

$$\begin{cases} f_{p\text{–best}} = f_{g\text{–best}} \\ P_p = P_g \end{cases} \tag{33}$$

*Step 4.* The local optimal and the global optimal positions are recorded. The position and flight speed of particles are updated in accordance with the following formula:

$$\begin{cases} x_{iN}^{l+1} = x_{iN}^l + v_{iN}^{l+1}, \\ v_{iN}^{l+1} = v_{iN}^l + c_1 r_1\left(p_{pN}^l - x_{iN}^l\right) + c_2 r_2\left(p_{gN}^l - x_{iN}^l\right). \end{cases} \tag{34}$$

*Step 5.* Continue to execute the program until the minimum value set by the optimization goal or the maximum number of iterations is met.

*3.2. Parameter Optimization Based on the PSO Algorithm.* Table 1 depicts the parameters of the rigid-flexible mechanical arms. The time of the input signal acting on Joint 1 and Joint 2 is 4 s. The parameter setting of the PSO algorithm is as follows: $c_1$ and $c_2$ are constant values, which are set by 2.25; $r_1$ and $r_2 \in [-1,1]$, where $r_1$ is −0.3 and $r_2$ is 0.5. The particle quantity is 60, the maximum iterations are 950, and the minimum value of the optimization function is set by $10^{-5}$. The weight of the optimization function is set by $\xi_1 = 0.43$ and $\xi_2 = 0.57$ to ensure that the mechanical arm can reach the target position.

The curve of the optimization process based on the above algorithm is shown in Figure 11. It decreases monotonically until meeting the conditions when the algorithm stops. In this study, the objective function value is already less than the minimum setting value when the number of iterations reaches 540 generations. The final parameters optimized by the PSO algorithm are depicted in Table 2. The complete input signal is depicted in Figures 12(a) and 12(b). The optimized input signal is similar to the original input signal, but the transition of the optimized input signal is smoother.

The original input signal and the optimized input signal are applied to the rigid-flexible mechanical arms to compare the end vibration of the flexible mechanical arm. The movement of the end of the mechanical arms is depicted in Figures 13(a) and 13(b). After applying the original input signal for 4 seconds, the maximum amplitude of the residual vibration is 16.12 mm. After applying the optimized input signal for 4 seconds, the end position error is 0.56 mm, the maximum amplitude of the terminal residual vibration is 4.76 mm, and the end residual vibration decreased by 70.22%. This process not only realizes to suppress the residual vibration of the end of the mechanical arms obviously but also realizes to control the end position of the mechanical arms accurately and the improvement of the positioning accuracy of the mechanical arms.

## 4. Experimental Verification

A hardware experimental device of rigid-flexible mechanical arms is developed so as to verify the effectiveness of the method of dimension reduction. The sensors consist of three resistance strain gauges, which can measure the deformation of different parts accurately. For the accurate measurement of the strain force, the connection mode of the half bridge power supply circuit of the mechanical arms is selected. The strain force amplifier is used to accurately measure the voltage signal and amplify the measured voltage signal. Then, the fast data acquisition instrument is used for data acquisition and A/D conversion for the collected data. Finally, the terminal deformation of the mechanical arms is calculated by computer, as shown in Figure 14.

According to the software of the control card, a pulse signal with a width of 5 seconds is given to the servo motor. According to the motor encoder, the angular velocity of the motor is obtained, and observing the voltage value of the resistance strain gauge, then can obtain the deformation of the end of the flexible arm through transformation. The simulation results of the reduced dimension model are compared with the experimental results, as depicted in Figures 15 and 16.

(1) As depicted in Figure 15, the simulation conclusions of the angular velocity of the flexible mechanical arm are mainly accordant with the experimental conclusions, which demonstrate that the selection of the spectral method and the dimension reduction method of Galerkin truncation theory can accurately reflect the characteristics of the dynamic model of the flexible mechanical arm.

(2) As depicted in Figure 16, in the first 5 seconds, the simulation conclusions of the first-order model by AIM are basically consistent with those of the

Table 2: Coefficients optimized by the PSO algorithm.

| Joint 1 signal | | Joint 2 signal | |
|---|---|---|---|
| $k_{10} = 0.4135$ | $k_{15} = 0.0540$ | $K_{20} = 0.4152$ | $k_{25} = 0.0613$ |
| $k_{11} = 0.2684$ | $k_{16} = 0.0006$ | $k_{21} = 0.2757$ | $k_{26} = 0.0034$ |
| $k_{12} = 0.0003$ | $k_{17} = 0.0375$ | $k_{22} = 0.0006$ | $k_{27} = 0.0406$ |
| $k_{13} = 0.0886$ | $k_{18} = 0.0019$ | $k_{23} = 0.0950$ | $k_{28} = 0.0002$ |
| $k_{14} = 0.0004$ | $k_{19} = 0.0301$ | $k_{24} = 0.0004$ | $k_{29} = 0.0308$ |

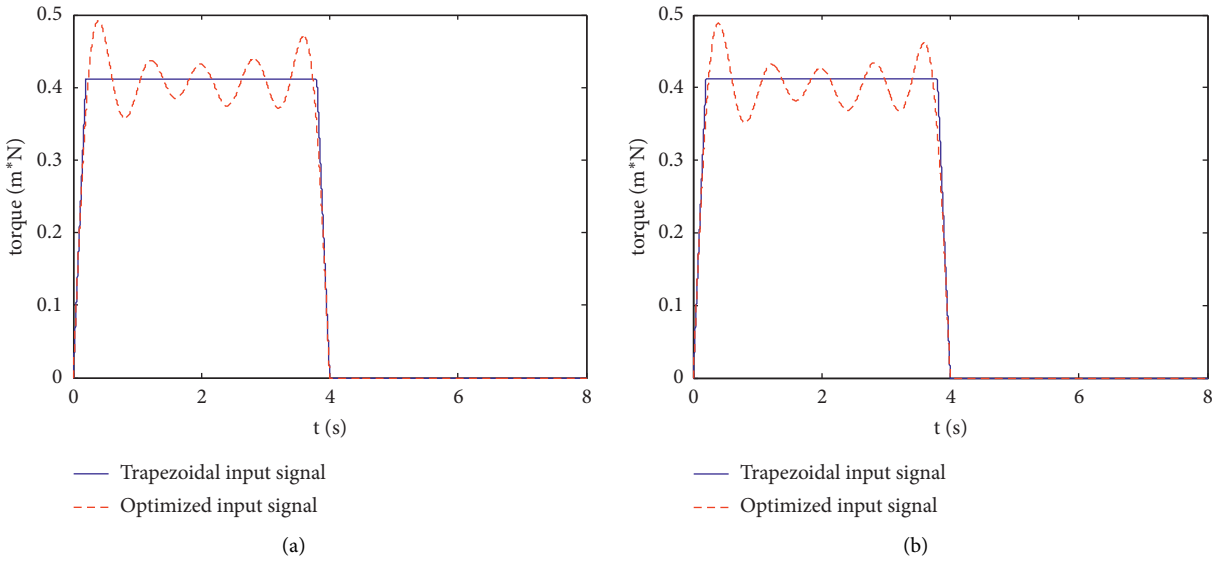



Figure 12: Comparison of the input signal: (a) Joint 1 input signal; (b) Joint 2 input signal.

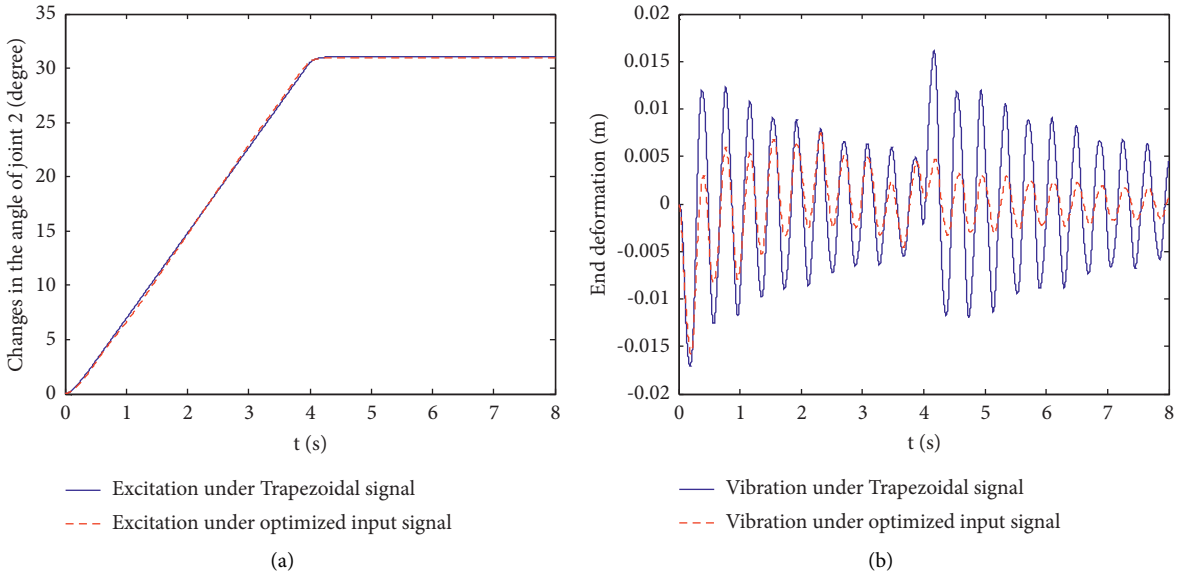



Figure 13: Comparison of end angle and end vibration: (a) end angle; (b) end vibration.

directly truncated third-order modal model. When the voltage signal is removed after the fifth second, the vibration mutation in the terminal deformation is removed slowly under the effect of structural damping, which confirms that the overall goal can further simplify the third-order system and lays a foundation for the next research on the control of rigid-flexible mechanical arms.

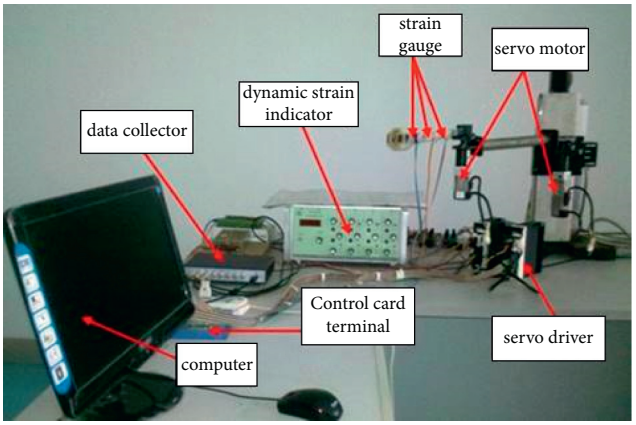



Figure 14: Experiment platform.

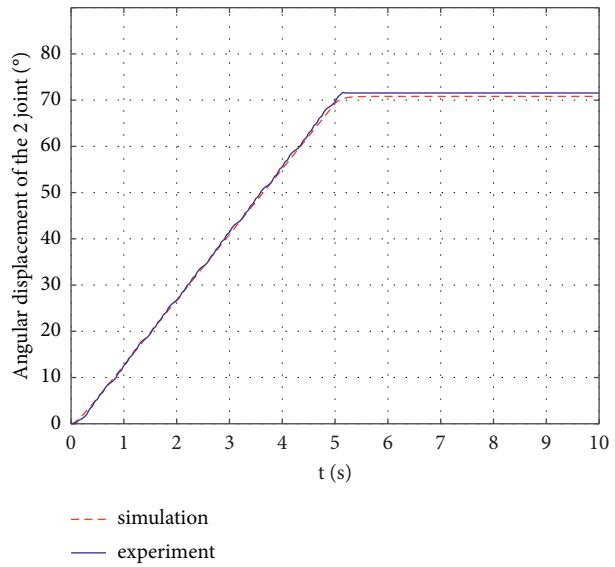



Figure 15: Comparison of flexible joint angular displacement.

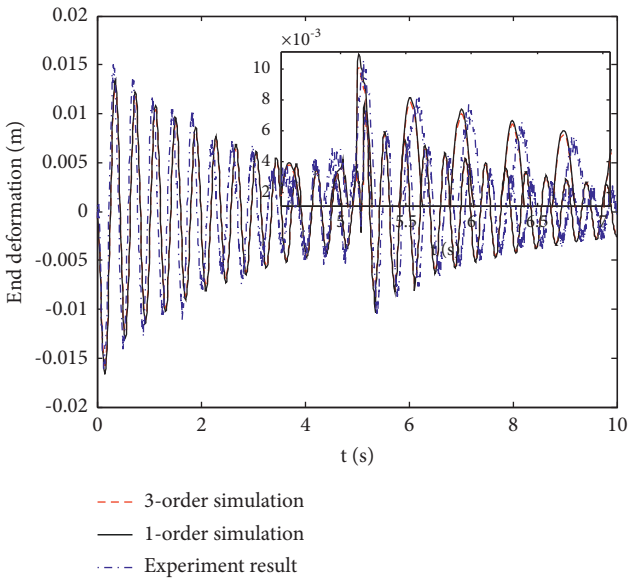



Figure 16: Comparison of deformation of flexible joint end.

(3) The experimental verification of vibration suppression is underway.

## 5. Conclusion

A dynamic model reduction method of rigid-flexible mechanical arms based on the approximate inertial manifold and a new method of terminal residual vibration suppression based on particle swarm optimization feedforward control are established. Firstly, the approximate inertial manifold theory is used to approximate the model of the rigid-flexible mechanical arms whose high-order mode of TGM is truncated and acts on the low-order mode, and its low-order system dynamics model is obtained. Then, the feedforward control method based on particle swarm optimization is applied to the dynamic model of the lowest-order rigid-flexible mechanical arm. The particle swarm optimization algorithm is used to modify the amplitude and frequency of the input signal composed of a limited number of sinusoidal signals with uncertain parameters so as to reduce the target position error and end residual vibration. The simulation results show that compared with the original input signal before optimization, the maximum amplitude of end residual vibration of rigid-flexible mechanical arms is reduced by 70%. Under the action of optimizing the input signal, the target position error is relatively small in the specified operation cycle. Simulation results show that this method can achieve accurate positioning and improve the suppression effect of residual vibration. The simulation results of angular displacement and end vibration of the simplified model are basically consistent with the experimental results of the hardware platform. The future research work is to verify the feedforward control method based on particle swarm optimization on the experimental platform and compare it with the simulation results to verify the effectiveness of the method [54].

## Appendix

## A

$$A_{11} = J_1 + J_t + M_t l_1^2 + \ell_2 l_1^2 l_2 + \frac{1}{3}\ell_2 l_2^3 + M_p\left(l_1^2 + l_2^2 + 2l_1 l_2 cos\theta_2 + w^2 - 2l_1 w \sin\theta_2\right) + \ell_2 l_1^2 l_2 \cos\theta_2 + \ell_2 \int_0^{l_2} w^2 \mathrm{d}x - 2\ell_2 l_1 \sin\theta_2 \int_0^{l_2} w \mathrm{d}x,$$

$$A_{12} = A_{21} = J_t + \frac{1}{3}\ell_2 l_2^3 + M_p\left(l_2^2 + l_1 l_2 cos\theta_2 + w^2 - l_1 w \sin\theta_2\right) + \frac{1}{2}\ell_2 l_1 l_1^2 \cos\theta_2 + \ell_2 \int_0^{l_2} w^2 \mathrm{d}x - \ell_2 l_1 \sin\theta_2 \int_0^{l_2} w \mathrm{d}x,$$

$$A_{22} = J_t + \frac{1}{3}\ell_2 l_2^3 + M_p\left(l_2^2 + w^2\right) + \ell_2 \int_0^{l_2} w^2 \mathrm{d}x,$$

$$G_1\left(\dot{\theta}_1, \dot{\theta}_2, \theta_2\right) = -\left(\frac{1}{2}\ell_2 l_1 l_2^2 \sin\theta_2 + M_p\left(l_1 l_2 \sin\theta_2 + l_1 w \cos\theta_2\right) + \ell_2 l_1 \cos\theta_2 \int_0^{l_2} w \mathrm{d}x\right)\dot{\theta}_2^2 + \left(-\left[\ell_2 l_1 l_2^2 \sin\theta_2 + 2\ell_2 l_1 \cos\theta_2 + M_p\left(2l_1 l_2 \sin\theta_2 + 2l_1 w \cos\theta_2\right) + \ell_2 l_1 \cos\theta_2 \int_0^{l_2} w \mathrm{d}x\right]\right)\dot{\theta}_1\dot{\theta}_2 + M_p\left(l_2^2 + 2w\cdot\dot{w} - 2l_1\dot{w}\sin\theta_2\right) + 2\ell_2 \int_0^{l_2} w\cdot\dot{w}dx - 2\ell_2 l_1 \sin\theta_2 \int_0^{l_2} \dot{w}dx\left(\dot{\theta}_1 + \dot{\theta}_2\right) + M_p w\cdot\left(l_1\cos\theta_2 + l_2\right) + \ell_2 l_1 \cos\theta_2 \int_0^{l_2} w \mathrm{d}x + \ell_2 \int_0^{l_2} x\cdot w \mathrm{d}x,$$

$$G_2\left(\dot{\theta}_1, \dot{\theta}_2, \theta_2\right) = -\left(\frac{1}{2}\ell_2 l_1 l_2^2 \sin\ \theta_2 + M_p\left(l_1 l_2 \sin\ \theta_2 + l_1 w\ \cos\ \theta_2\right) + \ell_2 l_1 \cos\ \theta_2 \int_0^{l_2} w \mathrm{d}x\right)\dot{\theta}_2^2$$

$$+ M_p\left(l_2^2 + w \cdot \dot{w}\right) + 2\ell_2 \int_0^{l_2} w \cdot \dot{w} \mathrm{d}x\left(\dot{\theta}_1 + \dot{\theta}_2\right) + M_p w \cdot \left(l_1 \cos\ \theta_2 + l_2\right), \tag{A.1}$$

where $w$ replaces $w\ (x,\ t)$.

## B

$$M = \begin{matrix} m_{11} & m_{12} & m_{13} & m_{14} & m_{15} \\ m_{12} & m_{22} & m_{23} & m_{24} & m_{25} \\ m_{13} & m_{23} & m_{33} & 0 & 0 \\ m_{14} & m_{24} & 0 & m_{44} & 0 \\ m_{15} & m_{25} & 0 & 0 & m_{55} \end{matrix},$$

$$m_{11} = J_1 + J_t + M_t l_1^2 + \ell_2 l_1^2 l_2 + \frac{1}{3}\ell_2 l_2^3 + \ell_2 l_1^2 l_2 \cos\ \theta_2 + \ell_2 l_2 \sum_{i=1}^{3} q_i^2(t) - 2\ell_2 l_1 \sin\ \theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x$$

$$+ M_p l_1^2 + l_2^2 + 2l_1 l_2 \cos\ \theta_2 - 2\ell_2 l_1 \sin\ \theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x + \left(\sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x^2\right)$$

$$m_{12} = J_t + \frac{1}{3}\ell_2 l_2^3 + \frac{1}{2}\ell_2 l_1 l_2^2 \cos\ \theta_2 + \ell_2 \sum_{i=1}^{3} q_i^2(t) - \ell_2 l_1 \sin\ \theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x + M_p l_2^2$$

$$+ \left(l_1 l_2 \cos\ \theta_2 + \left(\sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x\right)^2 - l_1 \sin\ \theta_2 \sum_{i=1}^{3} \int_0^{l_2} \Gamma_i(x) q_i(t) \mathrm{d}x\right),$$

$$m_{22} = J_t + \frac{1}{3}\rho_2 l_2^3 + \rho_2 l_2 \sum_{i=1}^{3} q_i^2(t) + M_p\left(l_2^2 + \left(\sum_{i=1}^{3} \int_0^{l_2} \Phi_i(x) q_i^2(t) \mathrm{d}x\right)^2\right) m_{1(i+2)}$$

$$= \rho_2 \int_0^{l_2} x \cdot \Phi_i(x) \mathrm{d}x + \rho_2 l_1 \cos\ \theta_2 \int_0^{l_2} \Phi_i(x) \mathrm{d}x + M_p l_1 \cos\ \theta_2 \int_0^{l_2} \Phi_i(x) \mathrm{d}x, \quad i = 1, 2, 3,$$

$$m_{2(i+2)} = \rho_2 \int_0^{l_2} x \cdot \Phi_i(x) \mathrm{d}x + M_p l_2 \int_0^{l_2} \Phi_i(x) \mathrm{d}x, \quad i = 1, 2, 3,$$

$$m_{(i+2)(i+2)} = \rho_2 \int_0^{l_2} \Phi_i(x) \mathrm{d}x, \quad i = 1, 2, 3,$$

$$
\begin{aligned}
K &= \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & k_3 & 0 & 0 \\ 0 & 0 & 0 & k_4 & 0 \\ 0 & 0 & 0 & 0 & k_5 \end{bmatrix}, \\
k_{i+2} &= EI_2\left(H_i^0\right)^4, \quad i = 1, 2, 3 \\
h_i(1) &= \int_0^{l_2} \Gamma_i(x)\mathrm{d}x, \\
h_i(x) &= \int_0^{l_2} x \cdot \Gamma_i(x)\mathrm{d}x, \quad i = 1, 2, 3, \\
F &= \begin{bmatrix} f_1 & f_2 & f_3 & f_4 & f_5 \end{bmatrix}^T \\
F_1(\theta, q, \dot{\theta}, \dot{q}) &= \begin{bmatrix} f_1 & f_2 \end{bmatrix}^T, \\
F_2(\theta, q, \dot{\theta}, \dot{q}) &= \begin{bmatrix} f_3 & f_4 & f_5 \end{bmatrix}^T.
\end{aligned} \tag{B.1}
$$

## Data Availability

The simulation and experimental data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

## Acknowledgments

This work was supported in part by the National Key Research and Development Project under Grant 2018YFB1307203 and the Project of State Key Laboratory of High Performance Complex Manufacturing, Central South University, under Grant ZZYJKT2021-17.

## References

[1] Z. Zhao, X. He, and C. K. Ahn, "Boundary disturbance observer-based control of a vibrating single-link flexible mechanical arm," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 49, no. 5, pp. 1–9, 2019.

[2] H. Karagülle, L. Malgaca, M. Dirilmiş, M. Akdağ, and Ş. Yavuz, "Vibration control of a two-link flexible manipulator," *Journal of Vibration and Control*, vol. 23, no. 12, pp. 2023–2034, 2015.

[3] D. Liang, Y. Song, T. Sun, and X. Jin, "Rigid-flexible coupling dynamic modeling and investigation of a redundantly actuated parallel manipulator with multiple actuation modes," *Journal of Sound and Vibration*, vol. 403, no. 9, pp. 129–151, 2017.

[4] W. Congqing, W. Pengfei, and Z. Xin, "Composite sliding mode control for a free-floating space rigid-flexible coupling mechanical arm system," *International Journal of Advanced Robotic Systems*, vol. 124, no. 10, pp. 1–10, 2013.

[5] S. Gharooni, B. Heller, and M. O. Tokhi, "A new hybrid spring brake orthosis for controlling hip and knee flexion in the swing phase," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 9, no. 1, pp. 106-107, 2001.

[6] R. Kumar, P. Berkelman, and P. Gupta, "Preliminary experiments in cooperative human/robot force control for robot assisted microsurgical manipulation," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 610–617, Montreal, Canada, January 2000.

[7] Z. Ni, J. Liu, Z. Wu, and X. Shen, "Identification of the state-space model and payload mass parameter of a flexible space manipulator using a recursive subspace tracking method," *Chinese Journal of Aeronautics*, vol. 32, no. 2, pp. 513–530, 2019.

[8] D. Meng, X. Wang, W. Xu, and B. Liang, "Space robots with flexible appendages: dynamic modeling, coupling measurement, and vibration suppression," *Journal of Sound and Vibration*, vol. 396, no. 5, pp. 30–50, 2017.

[9] H. N. Rahimi and M. Nazemizadeh, "Dynamic analysis and intelligent control techniques for flexible manipulators: a review," *Advanced Robotics*, vol. 28, no. 2, pp. 63–76, 2014.

[10] X. Zhou, H. Wang, Y. Tian, and G. Zheng, "Disturbance observer-based adaptive boundary iterative learning control for a rigid-flexible manipulator with input backlash and endpoint constraint," *International Journal of Adaptive Control and Signal Processing*, vol. 34, no. 9, pp. 1220–1241, 2020.

[11] Y. Zhao, C. M. Ruan, and S. .W. Wang, "Accurate modeling and experimental research on the second-order theory of rigid-flexible flexible mechanical arms," *Chinese Journal of Mechanical Engineering*, vol. 29, no. 2, pp. 205–210, 2018.

[12] D. Liang, Y. Song, T. Sun, and X. Jin, "Dynamic modeling and hierarchical compound control of a novel 2-DOF flexible parallel manipulator with multiple actuation modes," *Mechanical Systems and Signal Processing*, vol. 103, no. 3, pp. 413–439, 2018.

[13] F. Cao and J. Liu, "Three-dimensional modeling and input saturation control for a two-link flexible manipulator based on infinite dimensional model," *Journal of the Franklin Institute*, vol. 357, no. 2, pp. 1026–1042, 2020.

[14] G. Piras, W. L. Cleghorn, and J. K. Mills, "Dynamic finite-element analysis of a planar high-speed, high-precision

parallel manipulator with flexible links," *Mechanism and Machine Theory*, vol. 40, no. 7, pp. 849–862, 2005.

[15] C. Sun, W. He, and J. Hong, "Neural network control of a flexible robotic mechanical arm using the lumped spring-mass model," *IEEE Transactions on Systems, Man, and Cybernetics*, vol. 47, no. 8, pp. 1863–1874, 2016.

[16] M. Vakil, R. Fotouhi, and P. N. Nikiforuk, "Trajectory tracking for the end-effector of a class of flexible link manipulators," *Journal of Vibration and Control*, vol. 17, no. 1, pp. 55–68, 2011.

[17] X. Xiaoping Zhang, W. Wenwei Xu, and V. Chellaboina, "PDE modeling and control of a flexible two-link manipulator," *IEEE Transactions on Control Systems Technology*, vol. 13, no. 2, pp. 301–312, 2005.

[18] G. Natasha, S. Suharjito, and V. Noviantri, "Saint-venant model analysis of trapezoidal open channel water flow using finite difference method," *Procedia Computer Science*, vol. 157, no. 11, pp. 6–15, 2019.

[19] M. Jiang and H. Deng, "Optimal combination of spatial basis functions for the model reduction of nonlinear distributed parameter systems," *Communications in Nonlinear Science and Numerical Simulation*, vol. 17, no. 12, pp. 5240–5248, 2012.

[20] L. Chen and H. Deng, "Model reduction of rigid-flexible manipulators with experimental validation," *Advanced Materials Research*, vol. 655-657, no. 1, pp. 1101–1107, 2013.

[21] Y. C. Zhou, Z. Lu, H. Hu, and Y. S. Hu, "Surrogate modeling of high-dimensional problems via data-driven polynomial chaos expansions and sparse partial least square," *Computer Methods in Applied Mechanics and Engineering*, vol. 364, no. 1, pp. 1–24, 2020.

[22] J.-Z. Zhang, Y. Liu, and P.-H. Feng, "Approximate inertial manifolds of Burgers equation approached by nonlinear Galerkin's procedure and its application," *Communications in Nonlinear Science and Numerical Simulation*, vol. 16, no. 12, pp. 4666–4670, 2011.

[23] K. Ito and K. Kunisch, "Reduced-order optimal control based on approximate inertial manifolds for nonlinear dynamical systems," *SIAM Journal on Numerical Analysis*, vol. 46, no. 6, pp. 2867–2891, 2008.

[24] S. Hidetsugu, "Chaotic dynamics of an unstable Burgers equation," *Physica D*, vol. 129, no. 2, pp. 57–67, 1999.

[25] B. Mario, S. Viriato, and C. Francisco, "Dynamics and synchronization of numerical solutions of the Burgers equation," *Computational and Applied Mathematics*, vol. 231, no. 2, pp. 793–806, 2009.

[26] A. Abe, "Trajectory planning for residual vibration suppression of a two-link rigid-flexible manipulator considering large deformation," *Mechanism and Machine Theory*, vol. 44, no. 9, pp. 1627–1639, 2009.

[27] C. T. Kiang, A. Spowage, and C. K. Yoong, "Review of control and sensor system of flexible manipulator," *Journal of Intelligent and Robotic Systems*, vol. 77, no. 1, pp. 187–213, 2015.

[28] J. S. Lane and S. L. Dickerson, "Contribution of passive damping to the control of flexible mechanical arm," in *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pp. 175–180, Montreal, Canada, January 1984.

[29] Z.-c. Qiu, C. Li, and X.-m. Zhang, "Experimental study on active vibration control for a kind of two-link flexible manipulator," *Mechanical Systems and Signal Processing*, vol. 118, no. 7, pp. 623–644, 2019.

[30] E. Lu, W. Li, X. Yang, Y. Wang, and Y. Liu, "Optimal placement and active vibration control for piezoelectric smart flexible manipulators using modal H2 norm," *Journal of Intelligent Material Systems and Structures*, vol. 29, no. 11, pp. 2333–2343, 2018.

[31] S. Zhang, Y. Zhang, and X. Zhang, "Fuzzy PID control of a two-link flexible mechanical arm," *J. Vibroeng.*vol. 18, no. 1, pp. 250–256, 2016.

[32] H.-J. Yang and M. Tan, "Sliding mode control for flexible-link manipulators based on adaptive neural networks," *International Journal of Automation and Computing*, vol. 15, no. 2, pp. 239–248, 2018.

[33] A. Izadbakhsh, "Robust control design for rigid-link flexible-joint electrically driven robot subjected to constraint: theory and experimental verification," *Nonlinear Dynamics*, vol. 85, no. 2, pp. 751–765, 2016.

[34] J. O. Pedro and T. Tshabalala, "Hybrid NNMPC/PID control of a two-link flexible mechanical arm with actuator dynamics," in *Proceedings of the Asian Control Conference (ASCC)*, Jeju Island, Korea, May 2015.

[35] X. Xing and J. Liu, "PDE model-based state-feedback control of constrained moving vehicle-mounted flexible manipulator with prescribed performance," *Journal of Sound and Vibration*, vol. 441, no. 2, pp. 126–151, 2019.

[36] Z. Li, B. Huang, Z. Ye, M. Deng, and C. Yang, "Physical human-robot interaction of a robotic exoskeleton by admittance control," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 12, pp. 9614–9624, 2018.

[37] Z. Li, J. Li, S. Zhao, Y. Yuan, Y. Kang, and C. L. P. Chen, "Adaptive neural control of a kinematically redundant exoskeleton robot using brain-machine interfaces," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 12, pp. 3558–3571, 2019.

[38] Y. L. Liu, W. Su, Z. Li et al., "Motor-imagery-based teleoperation of a dual-arm robot performing manipulation tasks," *IEEE Transactions on Cognitive and Developmental Systems*, vol. 11, no. 3, pp. 414–424, 2019.

[39] H. Su, Y. Hu, H. R. Karimi, A. Knoll, G. Ferrigno, and E. De Momi, "Improved recurrent neural network-based manipulator control with remote center of motion constraints: experimental results," *Neural Networks*, vol. 131, pp. 291–299, 2020.

[40] H. Su, W. Qi, Y. Hu, H. R. Karimi, G. Ferrigno, and E. De Momi, "An incremental learning framework for human-like redundancy optimization of anthropomorphic manipulators," *IEEE Transactions on Industrial Informatics*, vol. 9, p. 1, 2020.

[41] H. Su, W. Qi, C. Yang, J. Sandoval, G. Ferrigno, and E. D. Momi, "Deep neural network approach in robot tool dynamics identification for bilateral teleoperation," *IEEE Robotics and Automation Letters*, vol. 5, no. 2, pp. 2943–2949, 2020.

[42] X. Zhou, W. Qi, S. E. Ovur et al., "A novel muscle-computer interface for hand gesture recognition using depth vision," *Journal of Ambient Intelligence and Humanized Computing*, vol. 11, no. 11, pp. 5569–5580, 2020.

[43] X. Y. Zhou, J. He, Q. He, C. Ren, and M. He, "Motion kinematics analysis of a horse inspired terrain-adaptive unmanned vehicle with four hydraulic swing mechanical arms," *IEEE Access*, vol. 8, no. 2, pp. 194350–194362, 2020.

[44] H. Yavuz, S. Mıstıkoğlu, and S. Kapucu, "Hybrid input shaping to suppress residual vibration of flexible systems," *Journal of Vibration and Control*, vol. 18, no. 1, pp. 132–140, 2012.

[45] K. Shin and M. J. Brennan, "Two simple methods to suppress the residual vibrations of a translating or rotating flexible

cantilever beam," *Journal of Sound and Vibration*, vol. 312, no. 1, pp. 140–150, 2007.

[46] Y.-x. Zheng and Y. Liao, "Parameter identification of nonlinear dynamic systems using an improved particle swarm optimization," *Optik*, vol. 127, no. 19, pp. 7865–7874, 2016.

[47] R. Mandel, A. Shooshtari, and M. Ohadi, "A "2.5-D" modeling approach for single-phase flow and heat transfer in manifold microchannels," *International Journal of Heat and Mass Transfer*, vol. 126, no. 1, pp. 317–330, 2018.

[48] M. Jolly, I. Kevrekidis, and E. Titi, "Approximate inertial manifolds for the Kuramoto-Sivashinsky equation: analysis and computations," *Physica D: Nonlinear Phenomena*, vol. 44, no. 2, pp. 38–60, 1990.

[49] X. Li and C. Sun, "Inertial manifolds for the 3D modified-Leray-$\alpha$ model," *Journal of Differential Equations*, vol. 268, no. 4, pp. 1532–1569, 2020.

[50] Y. Pan and H. Deng, "Model reduction of a two-link rigid-flexible manipulators based on spectral appximation method." in *Proceedings of the International Conference on Electrical Engineering, Computing Science and Automatic Control (ICEEAC)*, pp. 26–28, Mexico City, Mexico, November 2010.

[51] S. Baglioni, F. Cianetti, C. Braccesi, and D. M. De Micheli, "Multibody modelling of N DOF robot arm assigned to milling manufacturing. Dynamic analysis and position errors evaluation," *Journal of Mechanical Science and Technology*, vol. 30, no. 1, pp. 405–420, 2016.

[52] X. chen and Z. C. Fang, "Feedback control and numerical simulation of flexible mechanical arm vibration," *Journal of Vibration and Shock*, vol. 25, no. 1, pp. 1–4, 2006.

[53] D. H. Kim, "A swarm system design based on a modified particle swarm algorithm for a self-organizing scheme," *Advanced Robotics*, vol. 20, no. 8, pp. 913–932, 2006.

Hindawi
Complexity
Volume 2021, Article ID 4667586, 9 pages
https://doi.org/10.1155/2021/4667586

*Research Article*

# LSKE: Lightweight Secure Key Exchange Scheme in Fog Federation

**Yashar Salami and Vahid Khajehvand**

*Faculty of Computer and Information Technology Engineering, Qazvin Branch, Islamic Azad University, Qazvin, Iran*

Correspondence should be addressed to Vahid Khajehvand; vahidkhajehvand@gmail.com

Received 16 July 2021; Revised 13 August 2021; Accepted 29 September 2021; Published 25 October 2021

Academic Editor: Bo Xiao



The fog computing architecture allows data exchange with the vehicle network, sensor networks, etc. However, before exchanging data, the nodes need to know each other and key exchange. Yashar et al. recently proposed a secure key exchange scheme for the fog federation. However, their proposed scheme has a high computational overhead and is not suitable for fog federation. Therefore, we have proposed a lightweight, secure key exchange scheme for the fog federation to reduce computational overhead. To prove the lightweight, we have compared the proposed scheme with the Yashar design in terms of computing, and communication cost AVISPA Tool was used for the formal analysis of the proposed scheme. Then, we simulated the proposed scheme with the NS3 tool and compared it with Throughput, packet loss, Packet Delivery, and end-to-end delay with Yashar et al. scheme. The results show that the proposed design reduced 3.2457 ms of computational overhead and 1,024 transmitted data bits.

## 1. Introduction

The spread of distributed systems such as the cloud [1] has made it possible for users to access their data from anywhere and share or process their data. Furthermore, with the expansion of various branches of computer science and the relationship between these sciences, the development of distribution systems has accelerated. Today, the Internet of Things is connected to the fog layer and can generate thousands of data at any time that are sent to the fog layer for processing [2].

However, the fog layer needs to provide the necessary security for data processing between its nodes before processing the data. One of the most important challenges in maintaining security is how to exchange the key from the other side so that it is resistant to known attacks in the fog layer.

Novel remote user authentication and key agreement scheme for mobile client-server environment scheme in 2013 were proposed by Sun et al. [3]. This scheme was not secure and could not support the fog federation. In 2015, Li et al. [4] proposed smart card-based mutual authentication schemes in cloud computing. This scheme was not secure and could not support the fog federation. Security and privacy preservation scheme of face identification and resolution framework using fog computing in the Internet of things was presented by Hu et al. [5] in 2017; this scheme did not support fog federation and key exchange. The scheme proposed by Jia et al. [6], Wazid et al. [7], Chen et al. [8], Zheng and Chang [9], and Chen et al. [10] in 2019, 2020, and 2021 were all safe and supportive of mutual authentication and key exchange. However, they are not suitable for fog federation environments. Yashar et al. [11] proposed a secure key exchange scheme in the fog federation in 2021. This scheme supported mutual authentication and key exchange; however, this scheme is not lightweight. Table 1 shows a comparison of related work. Providing a secure and lightweight key exchange scheme in a fog federation environment is a challenge in this area.

### 1.1. Paper Contribution

(i) In this paper, we propose a secure lightweight key exchange scheme for the fog federation

(ii) For formal security analysis, the proposed scheme uses the AVISPA tool

(iii) The proposed scheme is compared with Yashar et al. regarding computing cost, communication cost, and security requirement

Table 1: Comparison of related work.

| Related work | Fog federation | Secure | Mutual authentication | Key exchange | Lightweight |
|---|---|---|---|---|---|
| Sun et al. [3] | x | x | ✓ | ✓ | x |
| Li et al. [4] | x | x | ✓ | ✓ | x |
| Hu et al. [5] | x | ✓ | ✓ | x | x |
| Jia et al. [6] | x | ✓ | ✓ | ✓ | x |
| Wazid et al. [7] | X | ✓ | ✓ | ✓ | x |
| Chen et al. [8] | x | ✓ | ✓ | ✓ | x |
| Zheng et al. [9] | x | ✓ | ✓ | ✓ | x |
| Chen et al. [10] | x | ✓ | ✓ | ✓ | x |
| Yashar et al. [11] | ✓ | ✓ | ✓ | ✓ | x |
| Proposed scheme | ✓ | ✓ | ✓ | ✓ | ✓ |

✓, the scheme is supported; X, the scheme is not supported.

(iv) The proposed scheme and Yashar et al. scheme are simulated with the NS3 tool and examined in terms of throughput, packet loss, packet delivery, and end-to-end delay criteria

*1.2. Paper Organization.* The rest of the paper is organized as follows. Section 2 reviews the Yashar et al. and network model. The proposed scheme has been presented in Section 3. Section 4 provides a security analysis of the proposed scheme with the AVISPA tool. Section 5 presents the performance analysis and security requirements. Section 6 compares the proposed scheme's simulation results with Yashar et al. Finally, conclusions have been presented in Section 7.

## 2. The Background

This section provides the ECC and network model and problem statement and scheme of Yashar et al.

*2.1. Review of ECC.* The elliptic curve cryptography (ECC) is a public key encryption method, which has been designed based on an algebraic structure of elliptic curves on the finite fields. The curves of the elliptic equations are in the form of $y + axy + by = x^3 + cx^2 + dx + e$. In this equation, $\mathbb{R} = \{a, b, c, d, e\}$. These are real numbers that must satisfy simple conditions. In these curves, a point is zero or a point in infinity. For more information, you can refer to [12].

*2.2. Network Model and Problem Statement.* The network model presented in Figure 1 shows that cloud servers are at the top tier and can communicate with each other. In the network model, there is a middle layer of fog nodes. In this layer, there is a central fog whose main task is to manage other fog nodes. The middle layer can be connected to the top layer and the low layer. The purpose of developing the haze layer was to reduce latency for bottom layer processing. At the low layer are IOV, IOS, IOE, and M2M devices. If these devices require high processing, they send their data to the fog layer for processing. In the fog layer, the central node needs to be aware of the identity of the nodes so that they can exchange data with each other. Furthermore, because the central node is being processed and managed, a secure, lightweight key exchange scheme is needed that can withstand known attacks.

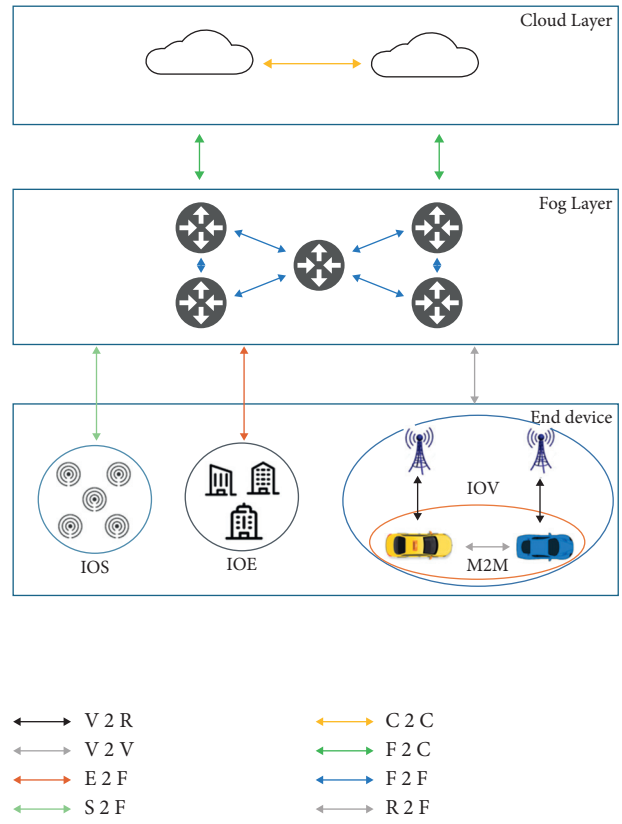



Figure 1: Network model [11].

*2.3. Notations.* The list of notations used in this paper is shown in Table 2.

*2.4. Review of Yashar Et Al.* The key exchange request steps are as follows.

Step 1: Bob generates an RB message to request the key exchange and transmits it to Alice. RB is calculated as follows:

Table 2: Notations used for the proposed work.

| No. | Notations | Description |
|---|---|---|
| 1 | IDA | Identity of Alice (fog center) |
| 2 | IDB | Identity of Bob (fog node) |
| 3 | NCa | Nonce of Alice (fog center) |
| 4 | NCB | Nonce of Bob (fog node) |
| 5 | Cha | Challenging of Alice (fog center) |
| 6 | CHb | Challenging of Bob (fog node) |
| 7 | Kas | The public key of Alice (fog center) |
| 8 | Kbs | The public key of Bob (fog node) |
| 9 | TA | Timestamp of Alice (fog center) |
| 10 | TB | Timestamp of Bob (fog node) |
| 11 | $\Delta T$ | Expiration time |
| 12 | $h$ | Hash |
| 13 | $\|\|$ | Concatenation |

Bi = IDB, NCb, NHa, Kbs, TB,
HBi = $h$(Bi),
RB = HBi‖Bi.

Step 2: Alice separates the file contents upon receiving RB and then hashes Bi and compares the Bi` hash with HBi. Then, if the hash of Bi and HBi are the same, she checks the packet timestamp with the predetermined ΔT. If the timestamp of the received packet is smaller than ΔT, the packet is valid. Next, Alice generates RBi, and Ai sends it to Bob. RB and Ai are calculated as follows:

Bi‘ = $h$(Bi),
Check if Bi‘ = HBi,
Check if TB ≤ $\Delta T$,
HRBi = $h$(IDB, NCb, CHb, Kbs, TA, ),
RBi = (IDB, NCb, CHb, Kbs, TA, HRBi)Kbs.

Key generation by Alice is as follows:

(i) To generate a Galois field, Alice selects a large prime number and calls it $p$. The field Zp might have $p-1$ generators.
(ii) Alice selects one of the generators of Zp and calls it $G$.
(iii) Alice selects an arbitrary number and calls it a, and keeps it secrete. Then, the selected numbers are substituted in equation (1) to generate A:

$$A = G^a \bmod P,$$
$$\text{HAi} = h\,(\text{IDA, NCa, CHa, Kas, } (P, G, A), \text{ TA}),$$
$$\text{Ai} = (\text{IDA, NCa, CHa, Kas, } (P, G, A), \text{TA, HAi})\text{Kbs}. \tag{1}$$

Step 3: Upon receiving Ai and RBi, Bob separates the contents of RBi with his public key and hashes the contents except for HRBi and compares RBi` with HRBi. Then, continue the calculation as follows:

RBi decrypt by key Kbs,
RBi‘ = $h$(IDB, NCb, CHb, Kbs, TA),
Check if RBi‘ = HRBi,
Check if TA ≤ $\Delta T$.

Key generation by Bob are as follows:

Bob uses equation (2) to generate B. Next, to obtain the shared key with Alice empowers A by $b$ in the modulus of P, according to equation (3), the result would be the shared key agreed upon by Alice. In the next step, it calculate RAi from the following relation and sends it to Alice:

$$B = G^b \bmod P, \tag{2}$$

$$\begin{aligned} K &= A^b \bmod P = (G^a)^b \bmod P \\ &= G^{a.b} \bmod p, \text{ Ai decrypt by key Kbs, Ai‘} \\ &= h\left(\text{IDA, NCa, CHa, Kas, } G^{\text{Na}}, \text{ TA}\right), \text{Check if Ai‘} \\ &= \text{HAi, Check if TA} \le \Delta T, \text{HAi} \\ &= h\,(\text{IDA, NCa, CHa, Kas, } B, \text{ TB}), \text{RAi} \\ &= (\text{IDA, NCa, CHa, Kas, } B, \text{HAi, TB})\text{Kas}. \end{aligned} \tag{3}$$

Step 4: Alice opens RAi with her public key, hashes the packet contents except for HAi, and compares RAi' with HAi. Then, Alice empowers B by $a$ in the modulus of P according to equation (4) to obtain the shared key. Figure 2 shows the key exchange scheme of Yashar et al. The shared key calculation is as follows:

$$K = B^a \bmod P = \left(G^b\right)^a \bmod P = G^{a.b} \bmod p$$
$$\text{RAi decrypt by key Kas}$$
$$\text{RAi‘} = h\,(\text{IDA, NCa, CHa, Kas, } B, \text{ TB})$$
$$\text{Check if } RAi\text{‘} = RAi$$
$$\text{Check if TB} \le \Delta T. \tag{4}$$

## 3. Proposed Scheme

This section presents the proposed scheme. The key exchange request steps are as follows:

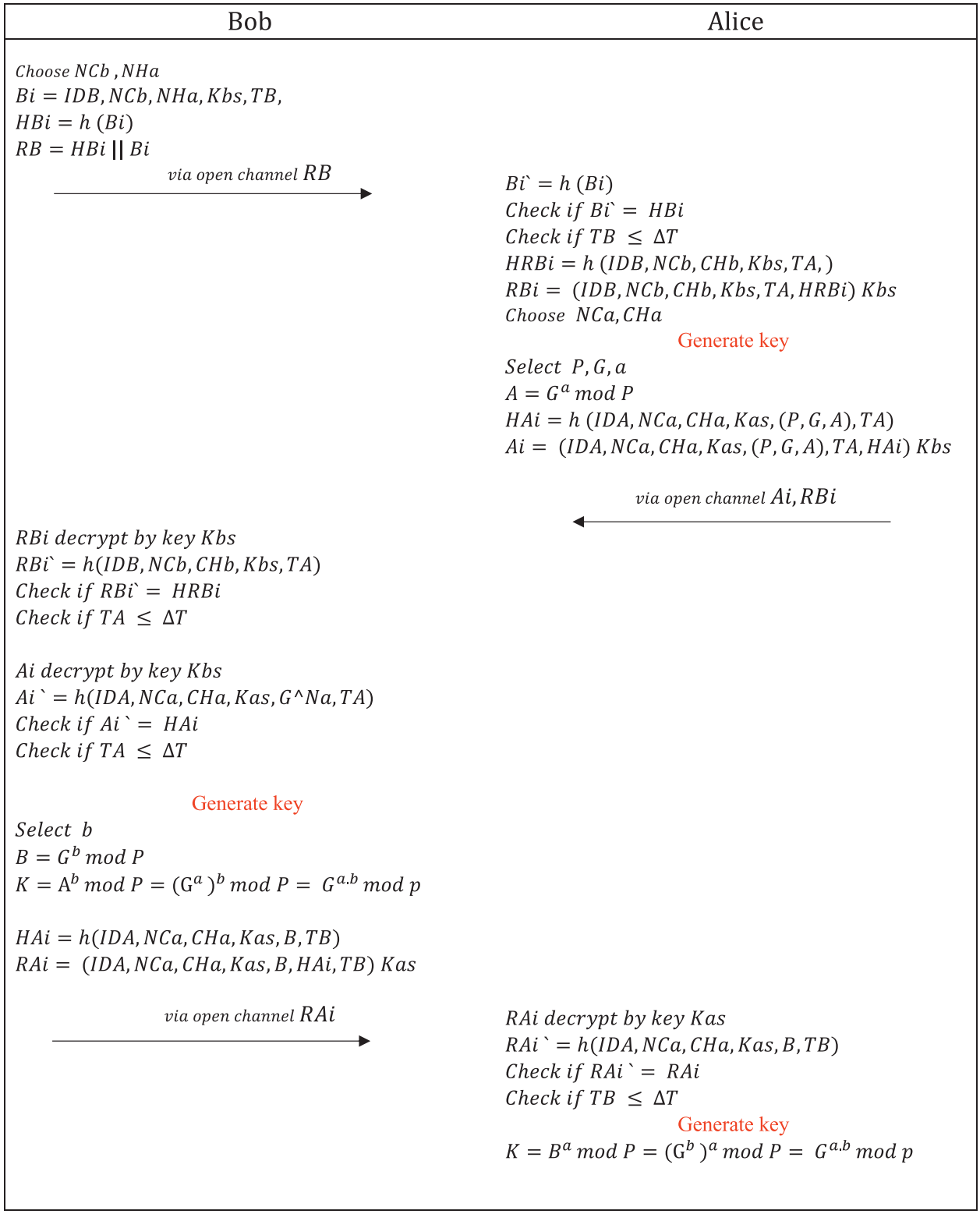



Figure 2: The key exchange scheme of Yashar et al.

Step 1: calculates the fog node of the equation $A1 = (\mathrm{IDA}, \mathrm{IDB}, \mathrm{TA}, \mathrm{Kas})$ and send it to the fog center.

Step 2: the Fog center first checks the time stamp with the expiration time; if the timestamp is shorter than the expiration time, it stores the Kas key.

In the next step, he chooses the numbers $a, b, p, R1, R2$, and NB and calculates them through equations $H1, B1, H2, B2, H3$, and PB, through equation (5). Finally, it sends B3 to the fog node:

$$\begin{aligned} H1 &= h(\mathrm{IDA}, \mathrm{IDB}, \mathrm{Kbs}),\\ B1 &= (\mathrm{IDA}, \mathrm{IDB}, \mathrm{Kbs}, H1),\\ H2 &= h(a, b, p, R1, R2),\\ B2 &= (a, b, p, R1, R2, H2),\\ \mathrm{PB} &= \mathrm{NB} * G(R1, R2),\\ H3 &= h(\mathrm{PB}),\\ B3 &= (\mathrm{PB}, \mathrm{TB}, H3, B1, B2)\mathrm{Kas}. \end{aligned} \tag{5}$$

Step 3: the Fog node first checks the time stamp with the expiration time; if the timestamp is shorter than the expiration time, it stores the Kas key. It then hashes B1, B2, and B3 and compares it to H1, H2, and H3 to ensure that the message is not tampered with. It then saves a, b, p, R1, R2, and PB. The fog node selects a random number in the next step, places it in equation (6), and obtains PA.

$$\mathrm{PA} = \mathrm{NA} * G(R1, R2). \tag{6}$$

After the following calculations, it sends A2 and H4 to the fog center:

$$\begin{aligned} H4 &= h(\mathrm{PA}),\\ A2 &= (\mathrm{PA}, \mathrm{TA}, H4)\mathrm{Kbs}. \end{aligned}$$

Fog node through equation (7) calculates the common key:

$$K = \mathrm{NA} * \mathrm{PB}. \tag{7}$$

Step 4: the Fog center first checks the time stamp with the expiration time; if the timestamp is shorter than the expiration time, it first hashes A2 and compares it to H4. Then, it checks the time stamp with the expiration time; if the timestamp is shorter than the expiration time, the fog center calculates the common key through equation (8). Figure 3 shows the proposed scheme.

$$K = \mathrm{NB} * \mathrm{PA}. \tag{8}$$

## 4. Security Analysis

This section presents the simulation results of the proposed scheme with the AVISPA tool.

The AVISPA tool is a formal simulation to assess whether a secure or insecure protocol [13]. AVISPA uses an HLPSL language to describe and display the security specifications of protocols. HLPSL is a role-oriented language in which each entity plays an independent role during the protocol implementation [14]. In HLPSL, a legal role is conceived for the attacker, modeled by Dolew-yao [15]. AVISPA has four built-in tools OFMC (On-the-Fly Model-Checker) [16], CL-AtSe (Constraint Logic-based Attack Searcher) [17], SATMC (SAT-based Model-Checker) [18], and TA4SP (Tree Automata based on Automatic Approximations for the Analysis of Security Protocols) [19] that are used for security analysis. After parsing, the output results indicate whether the protocol is secure or insecure.

*4.1. Analysis of Simulation Results.* Figures 4 and 5 show the simulation results showing the proposed design with security tools OFMC and CL-AtSe. The simulation results in the OFMC show that the total number of nodes visited for the proposed scheme was 17 and with a depth of 4 in 0.14 seconds. The simulation results in the CL-AtSe show that the total number of analyzed and reachable for the proposed scheme was four states in translation time was 0.04 seconds. Furthermore, the security analysis results with tools OFMC and CL-AtSe show that the proposed scheme is secure.

## 5. Performance Analysis

In this section, the performance analysis of the proposed scheme and security requirements are compared with Yashar et al. The following symbols are defined to evaluate the computing cost of the proposed scheme. Th is the execution number of a hash operation. Pm is the execution number of Point Multiplication. Pe is the execution number of public key encryption. Pd is the execution number of public key decryption. Se is the execution number of symmetric key encryption. Sd is the execution number of symmetric key decryption. The execution time to perform the computation is as follows. Th ≈ **0.0023**, Pm ≈ **2.226**, Pe ≈ **3.8500**, Pd ≈ **3.8500**, Se ≈ **0.0046**, and Sd ≈ **0.0046.** The proposed scheme uses 1024 bit RSA.

*5.1. Computation Cost.* Table 3 shows a comparison of the computing cost of the proposed scheme and the Yashar et al. scheme. Our observations show that the Yashar scheme consists of 7 Th, 3Pe, and 3Pd; the total cost is 23.1161 ms. On the contrary, our proposed scheme consists of 8Th, 2Pm, 2Pe, and 2Pd; the total cost is 19.8704 seconds. Thus, compared to Yashar et al., the proposed scheme reduced the calculation by 3.2457.

*5.2. Communication Cost.* Table 4 shows a comparison of the communication cost of the proposed scheme, and the Yashar et al. scheme has a communication cost of 3, and the total number of bits used is 3072. In our proposal, the communication cost is three, and the total number of bits used is 2048. Thus, we reduced 1,024 bits sent over the scheme of Yashar et al.

*5.3. Security Requirements' Comparison.* Our observations show that the proposed scheme is resistant to defined attacks. However, our proposal also cannot support device

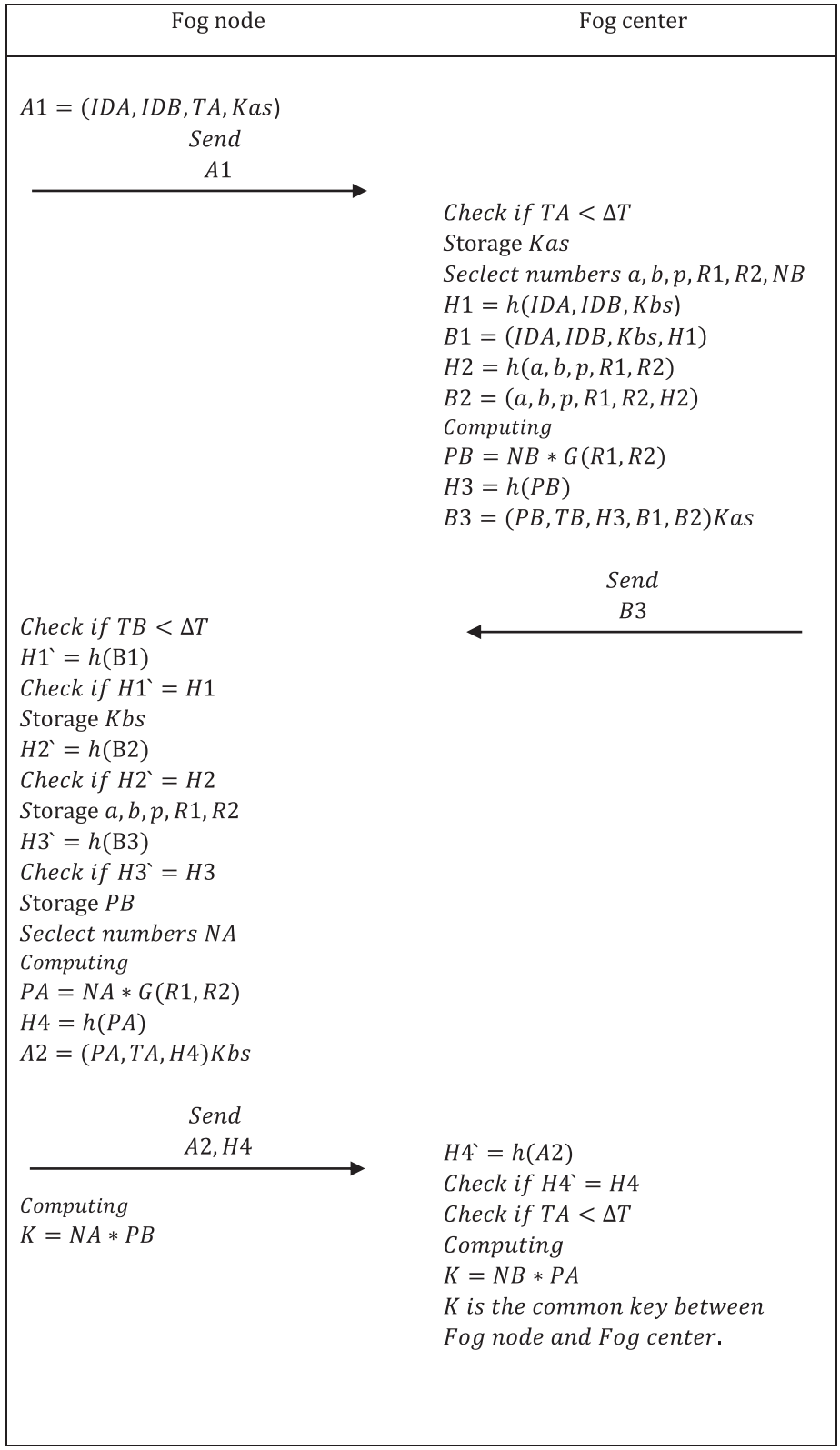



Figure 3: The proposed scheme.

```
% OFMC
% Version of 2006/02/13
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
PROTOCOL
 /home/span/span/testsuite/results/hlpslGenFile.if
GOAL
  as_specified
BACKEND
  OFMC
COMMENTS
STATISTICS
  Parse Time: 0.00s
  Search Time: 0.14s
  Visited Nodes: 17 nodes
  depth: 4 plies
```

Figure 4: Simulation results of the proposed scheme under OFMC.

```
SUMMARY
  SAFE
DETAILS
  BOUNDED_NUMBER_OF_SESSIONS
  TYPED_MODEL
PROTOCOL
  home/span/span/testsuite/results/hlpslGenFile.if
GOAL
  As Specified
BACKEND
  CL-AtSe
STATISTICS
  Analysed : 4 states
  Reachable : 4 states
  Translation: 0.04 seconds
  Computation: 0.00 seconds
```

Figure 5: Simulation results of the proposed scheme under CL-ATSe.

Table 3: Comparison of computation cost.

| No. | Schemes | Hash function | Point multiplication (Pm) | Public key encryption | Public key decryption | Symmetric key encryption (Se) | Symmetric key decryption | Total cost | TC (ms) |
|---|---|---|---|---|---|---|---|---|---|
| 1 | Yashar et al. | 7Th | 0 | 3Pe | 3Pd | 0 | 0Sd | 7Th + 3Pe + 3Pd | 23.1161 |
| 2 | Proposed | 8Th | 2 | 2Pe | 2Pd | 0 | 0Sd | 8Th + 4Pm + 2Pe + 2Pd | 19.8704 |

anonymity and session key agreement. Table 5 shows the security requirements' comparison of the proposed scheme with the Yashar et al. scheme. Note: AF1: replay attack; AF2: man-in-the-middle attack; AF3: insider attack; AF4: impersonation attack; AF5: brute force attack; AF6: offline password guessing attack; AF7: device anonymity; AF8: mutual authentication; AF9: session key agreement; AF10: key exchange, AF11: fog federation; AF12: OFMC; AF13: CL-ATSE; ✓: the scheme is supported; X: the scheme is not supported.

## 6. Simulation and Result

In this section, a simulation of the proposed design with the Yashar design is provided. In addition, simulation by network simulation tool (NS 3 2.29 simulator) on the Ubuntu-20.04.1 platform is provided. The hardware environment for carrying out NS3 simulation [20] was on Dell Inspiron 5110 machine with Intel Core i5 2410 M/ 2.30 GHz processor having 4 GB RAM and 1 TB HDD (Hard Disk Drive).

Table 4: Comparison of communication cost and the number of bits.

| No. | Schemes | No. of messages | Total cost (in bits) |
|---|---|---|---|
| 1 | Yashar et al. | 3 | 3072 |
| 2 | Proposed | 3 | 2048 |

Table 5: Comparison of security requirements.

| Security requirements | Scheme | |
|---|---|---|
| | Yashar et al. | Proposed |
| AF1 | ✓ | ✓ |
| AF2 | ✓ | ✓ |
| AF3 | ✓ | ✓ |
| AF4 | ✓ | ✓ |
| AF5 | ✓ | ✓ |
| AF6 | ✓ | ✓ |
| AF7 | X | x |
| AF8 | ✓ | ✓ |
| AF9 | X | x |
| AF10 | ✓ | ✓ |
| AF11 | ✓ | ✓ |
| AF12 | ✓ | ✓ |
| AF13 | ✓ | ✓ |

Table 6: Simulation parameters.

| Parameters | Description |
|---|---|
| Platform | Ubuntu-20.04.1 |
| Hardware platform | Dell 5110, Intel Core i5, 4 GB RAM, 1 TB HDD |
| Tool used | NS 3 2.29 |
| Number of fog node | 20 |
| Number of fog center | 10 |
| Mobility of fog node | 0 |
| Mobility of fog center | 0 |
| Simulation environment area | 300 ∗ 1500 M |
| Loss model | Friis loss |
| Transmit power | 7.5 dB |
| Routing protocol | OLSR |
| Medium access control type | IEEE 802.11 |
| Wireless protocol | 802.11 p |
| Communication range of fog node to fog center | 100 M |
| Simulation time | 1800 seconds |

*6.1. Simulation Environment and Settings.* The various parameters used in the NS3 simulations are provided in Table 6. The simulation time of the proposed scheme was 1800 seconds. The number of fog centers is ten, and the fog node is 20. Other parameters are as follows: the mobility of the fog centers and fog node is 0 m/s, loss model is Friis loss, transmit power is 7.5-dBm, medium access control type IEEE 802.11, wireless protocol 802.11 *p*, routing protocol: OLSR, and Simulation Environment Area is 300 ∗ 1500M.

*6.2. Simulation Results.* The simulation results show that the proposed scheme performs better in terms of throughput than the Yashar et al. scheme. Figure 6 shows a comparison

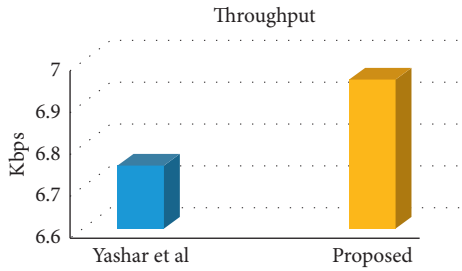



Figure 6: Comparison of throughput.

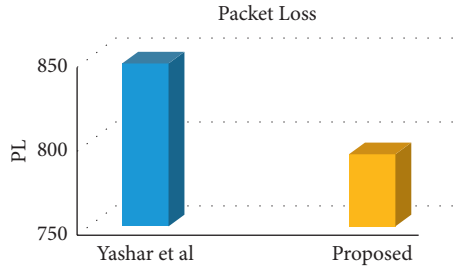



Figure 7: Comparison of packet loss.

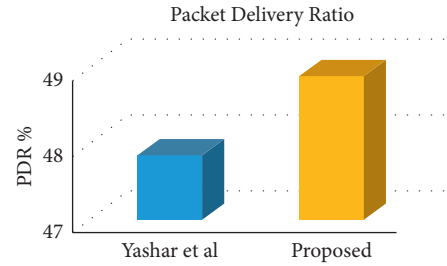



Figure 8: Comparison of packet delivery.

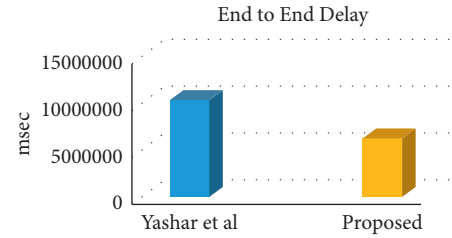



Figure 9: Comparison of end-to-end delay.

of the proposed scheme and Yashar et al. scheme in terms of throughput. The proposed scheme has a much better performance in packet loss than Yashar et al. Figure 7 compares the proposed scheme and Yashar et al. scheme in packet loss. In terms of packet delivery rate, the proposed scheme has shown better performance than Yashar et al. Figure 8 compares the proposed scheme and Yashar et al. scheme in terms of the packet delivery rate. Finally, in terms of end-to-end delay, the performance of the proposed design is better than that of Yashar et al. Figure 9 shows a comparison of the proposed scheme and the Yashar et al. scheme in terms of end-to-end delay.

## 7. Conclusion

The secure key exchange in fog federation environments is a major challenge. This paper presents a lightweight, secure key exchange scheme based on ECC for fog federation environments. The results of the AVISPA tool show that the proposed scheme is safe, and the proposed scheme is compared with Yashar et al. Comparison results show that the proposed scheme has a lower computational and byte cost. The proposed scheme is then simulated with the NS3 tool. The simulation results show that the proposed scheme performs better in terms of throughput, packet loss, packet delivery, and end-to-end delay than Yashar et al. In future work, our goal is to provide a three-way key exchange scheme in fog federation.

## Data Availability

Data used to support this novel scheme are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## References


[1] K. Zhang, Y. Li, and L. Lu, "Privacy-preserving attribute-based keyword search with traceability and revocation for cloud-assisted IoT," *Security and Communication Networks*, vol. 2021, Article ID 9929663, 13 pages, 2021.

[2] B. Zheng, Z. Mei, L. Hou, and S. Qiu, "Application of internet of things and edge computing technology in sports tourism services," *Security and Communication Networks*, vol. 2021, Article ID 9980375, 10 pages, 2021.

[3] H. Sun, Q. Wen, H. Zhang, and Z. Jin, "A novel remote user authentication and key agreement scheme for mobile client-server environment," *Applied Mathematics & Information Sciences*, vol. 7, no. 4, pp. 1365–1374, 2013.

[4] H. Li, F. Li, C. Song, and Y. Yan, "Towards smart card based mutual authentication schemes in cloud computing," *KSII Transactions on Internet and Information Systems*, vol. 9, no. 7, pp. 2719–2735, 2015.

[5] P. Hu, H. Ning, T. Qiu, H. Song, Y. Wang, and X. Yao, "Security and privacy preservation scheme of face identification and resolution framework using fog computing in internet of things," *IEEE Internet of Things Journal*, vol. 4, no. 5, pp. 1143–1155, 2017.

[6] X. Jia, D. He, N. Kumar, and K.-K. R. Choo, "Authenticated key agreement scheme for fog-driven IoT healthcare system," *Wireless Networks*, vol. 25, no. 8, pp. 4737–4750, 2019.

[7] M. Wazid, A. K. Das, N. Kumar, and A. V. Vasilakos, "Design of secure key management and user authentication scheme for fog computing services," *Future Generation Computer Systems*, vol. 91, pp. 475–492, 2019.

[8] C.-M. Chen, Y. Huang, K.-H. Wang, S. Kumari, and M.-E. Wu, "A secure authenticated and key exchange scheme for fog computing," *Enterprise Information Systems*, vol. 75, pp. 1–16, 2020.

[9] Y. Zheng and C.-H. Chang, "Secure mutual authentication and key-exchange protocol between PUF-embedded IoT endpoints," in *Proceedings of the 2021 IEEE International Symposium On Circuits And Systems (ISCAS)*, pp. 1–5, Daegu, South Korea, May 2021.

[10] Y. Chen, J. Yuan, and Y. Zhang, "An improved password-authenticated key exchange protocol for VANET," *Vehicular Communications*, vol. 27, p. 100286, 2021.

[11] Y. Salami, Y. Ebazadeh, and V. Khajehvand, "CE-SKE: cost-effective secure key exchange scheme in Fog Federation," *Iran Journal of Computer Science*, vol. 4, no. 3, pp. 1–13, 2021.

[12] S. Kalra and S. K. Sood, "Secure authentication scheme for IoT and cloud servers," *Pervasive and Mobile Computing*, vol. 24, pp. 210–223, 2015.

[13] L. Viganò, "Automated security protocol analysis with the AVISPA tool," *Electronic Notes in Theoretical Computer Science*, vol. 155, no. 1, pp. 61–86, 2006.

[14] D. Von Oheimb, "The high-level protocol specification language HLPSL developed in the EU project AVISPA," in *Proceedings of the APPSEM 2005 Workshop*, pp. 1–17, Frauenchiemsee, Germany, September 2005.

[15] D. Dolev and A. Yao, "On the security of public key protocols," *IEEE Transactions on Information Theory*, vol. 29, no. 2, pp. 198–208, 1983.

[16] D. Basin, S. Mödersheim, and L. Viganò, "An on-the-fly model-checker for security protocol analysis," *Computer Security-ESORICS 2003*, vol. 2808, no. 3, pp. 253–270, 2003.

[17] M. Turuani, "The CL-Atse protocol analyser," in *Term Rewriting and Applications*, F. Pfenning, Ed., pp. 277–286, Springer, Berlin Germany, 2006.

[18] A. Armando and L. Compagna, "SATMC: a SAT-based model checker for security protocols," in *Logics in Artificial Intelligence*, J. J. Alferes and J. Leite, Eds., Springer, Berlin Germany, 2004.

[19] Y. Boichut, N. Kosmatov, and L. Vigneron, "Validation of Prouvé protocols using the automatic tool TA4SP," *TFIT*, vol. 6, pp. 467–480, 2006.

[20] "NS3." https://www.nsnam.org.

Hindawi
Complexity
Volume 2021, Article ID 2214762, 16 pages
https://doi.org/10.1155/2021/2214762

*Review Article*

# Complex Electrical Stimulation Systems in Motor Function Rehabilitation after Spinal Cord Injury

**Jiange Kou,[1] Maolin Cai,[1] Fei Xie,[2] Yixuan Wang,[1] Na Wang,[3] and Meng Xu[4]**

[1]*School of Automation Science and Electrical Engineering, Beihang University, Beijing 100191, China*
[2]*Department of Pulmonary and Critical Care Medicine, Chinese PLA General Hospital, Beijing 100039, China*
[3]*Engineering Training Center, Beihang University, Beijing 100191, China*
[4]*Department of Orthopedics, Chinese PLA General Hospital, Beijing 100039, China*

Correspondence should be addressed to Fei Xie; xiefei0522@163.com

Received 24 May 2021; Revised 16 July 2021; Accepted 30 September 2021; Published 25 October 2021

Academic Editor: Wen Qi



Spinal cord injury (SCI) is an existing incurable disease that brings great pain and life obstacles to patients. Spinal cord electrical stimulation is an effective means to alleviate spinal cord injury. However, its complicated mechanism of action is still unclear. This article aims to summarize several different spinal cord electrical stimulation methods, analyze the stimulation effect, and briefly describe the current understanding of its origin and mechanism of action. In recent years, several application cases of the electrical stimulation system of stimulation methods have confirmed its positive effects in spinal cord injury diseases and provided new perspectives for the improvement of spinal cord injury. Finally, the possible development direction and corresponding challenges of spinal cord electrical stimulation in the future are proposed.

## 1. Introduction

*1.1. Spinal Cord Injury.* Spinal cord injury (SCI) is a transverse injury of the spinal cord caused by various pathogenic factors (trauma, tumor, inflammation, etc.). The spinal cord is part of the central nervous system and consists of a large bundle of nerves that allow the brain to communicate with the rest of the body (through the peripheral nerves). Once SCI occurs, it will lead to spinal cord nerve dysfunction below the injury level, such as sensory and autonomic dysfunction, impaired motor function, and sphincter and autonomic nerve dysfunction [1]. This brings great life pressure and psychological burden to patients and their families. Traumatic SCI has the characteristics of acute onset, serious condition, high disability-fatality rate, and so on. The treatment of traumatic SCI is also a worldwide problem [2–4], and the incidence of SCI is increasing year by year [5]. The high incidence of SCI is mostly caused by traffic accidents, falls, violence, and other factors [6]. After SCI, how to restore nerve function to enable paraplegic patients to obtain motor ability is still a difficult problem in life science.

There are surgical treatment [7], drug therapy [8], physical rehabilitation [9], and other treatment methods for patients with SCI. Surgical treatment is generally concentrated in the acute stage of spinal cord injury, but the influence of the timing of surgical decompression is still discussed, and there are a large number of changes in clinical practice [10]. Drug therapy generally plays a role in reducing inflammation but does not play a significant role in the improvement of motor function [8]. As a common treatment, physical rehabilitation has been used in patients with spinal cord injury and achieved certain results. However, the individual specificity of physical rehabilitation treatment is strong, and the universality needs to be improved [9]. In addition, there are new treatments such as biomaterials or stem cell transplantation [11, 12]. But the pathophysiological mechanism of spinal cord injury is complex, changeable, and multi-inducing, and the mechanisms are still unclear [13], which brings great obstacles to the treatment of spinal cord injury.

So far, great efforts have been made to solve this problem in various fields around the world. SCI is no longer a single

medical problem, but a life science difficulty to be overcome in biomedical, mechanical and electronic, computer, information, and other fields. Engineering methods that combine multiple fields are used to treat spinal cord injuries, such as electrical stimulation techniques.

*1.2. Electrical Stimulation Technique.* Electric current can cause responses in excitable cells, including neurons, and can be used to produce sensory or motor feedback. Through this principle, the researchers invented the electrical stimulation technology and carried out the initial clinical application. The spinal cord has become an attractive stimulation target, because it retains motor and sensory pathways below the injury level [14, 15]. The existing methods of spinal cord electrical stimulation include epidural electrical stimulation, transcutaneous stimulation, and intraspinal electrical stimulation. Its effectiveness can be verified by several examples, including hand grip [16], walking [17, 18], protecting spinal motor neurons [19], and bladder management [20].

Most of the early electrical stimulation techniques set the stimulation parameters through the stimulation system and did not form a complete closed-loop control. However, with the development of science and technology, such as brain-computer interface technology, electrical stimulation technology is gradually improved. As shown in Figure 1, Depending on the recording method, neuroscientists distinguish between electroencephalogram (EEG), electrocorticogram (ECoG), and local field potential (LFP). High-frequency collection of single nerve cell discharges peak potential spike [21]. Researchers can collect EEG as control signals to extract the characteristic information contained in the signals. By means of machine learning, the feature information is classified, and the mapping relationship between the feature information and the action is established and transformed into stimulus signals. After that, when this type of signal is collected again, the receptor can be stimulated to produce the corresponding action, and finally the motor function can be restored. This technique has also been applied to solve the problem of SCI, and a breakthrough has been achieved [22]. The purpose of this paper is to summarize the different methods of spinal cord electrical stimulation used to restore motor function, analyze their advantages and disadvantages, and briefly introduce the potential future of this technique.

## 2. Key Components of Electrical Stimulation System

*2.1. Electrical Stimulator.* The electrical stimulator is used to generate different stimulating currents, which are connected to the electrode to achieve the purpose of transmitting the current. In general, electrical stimulators can connect one or more channels, each consisting of a pair of electrodes (anode and cathode), and each channel is independent of each other. The electrical stimulator can set different pulse parameters (amplitude, pulse width) and the stimulation sequence of each channel, resulting in different stimulation strategies.

*2.2. Electrodes.* The electrode is used to make direct contact with the stimulated target to transmit electric current. According to its placement position, it can be divided into the implantable electrode and nonimplantable electrode. Implantable electrodes are generally flaky (mostly epidural electrical stimulation) or needle (mostly intraspinal electrical stimulation), which need to be surgically implanted in the body. It can be closer to the stimulus position but also requires a smaller stimulus current, which is generally less than 25 mA. Nonimplantable electrodes are generally attached to the skin above the spinal cord, do not require surgical participation, are very suitable for temporary use, and can adjust the stimulation position to produce the desired effect (movement). Because the electrode and the stimulation position are separated from the skin, it generally requires a large current and will be controlled within the 120 mA, as shown in Table 1.

In addition to the electrode position and current, the key factors affecting the effect of electrical stimulation are frequency, pulse width, and so on. The stimulation frequency is the rate of pulse transmission, which affects the speed of muscle contraction. Exceeding 40 Hz will cause tetanic contraction of muscles, and spinal cord electrical stimulation is generally used at 20–25 Hz. Current amplitude and pulse width will affect the stimulation effect at the same time, and there is an inverse correlation between them, and they work together to produce enough energy to produce a response. In order to achieve the comfort of stimulating effect, the current amplitude is generally controlled within 100 mA [23].

*2.3. Sensors.* There may be sensors as auxiliary devices to obtain better stimulation effects for different stimulation strategies. For example, stimulus strategies based on inertial sensors are more smoothly controlled [24], trigger algorithm based on inertial sensor to assist paraplegic patients to swim [25], using EMG signal electrical stimulation feedback control [26], and so on. The prediction of motor behavior based on sensor information [27, 28] and physical rehabilitation training can help patients return to normal life to the greatest extent.

*2.4. Control Systems.* It is difficult to control the pulse frequency, amplitude, pulse width, and other stimulation parameters for the electrical stimulation system. It is also the most important part of the stimulation system. While producing a stimulating effect, ensure the safety, comfort, and individual differences of the system. These parameters will affect the stimulation effect and ultimately affect the patient's rehabilitation progress. Appropriate stimulation parameters will improve the efficiency of stimulation, and unreasonable parameters are likely to hinder the patient's recovery.

The research on electrical stimulation control system focuses on (1) open-loop and closed-loop control, (2) control algorithm technology research, or (3) feedback information to the stimulator control unit through biological signals or sensors [29, 30]. The control objects are mainly

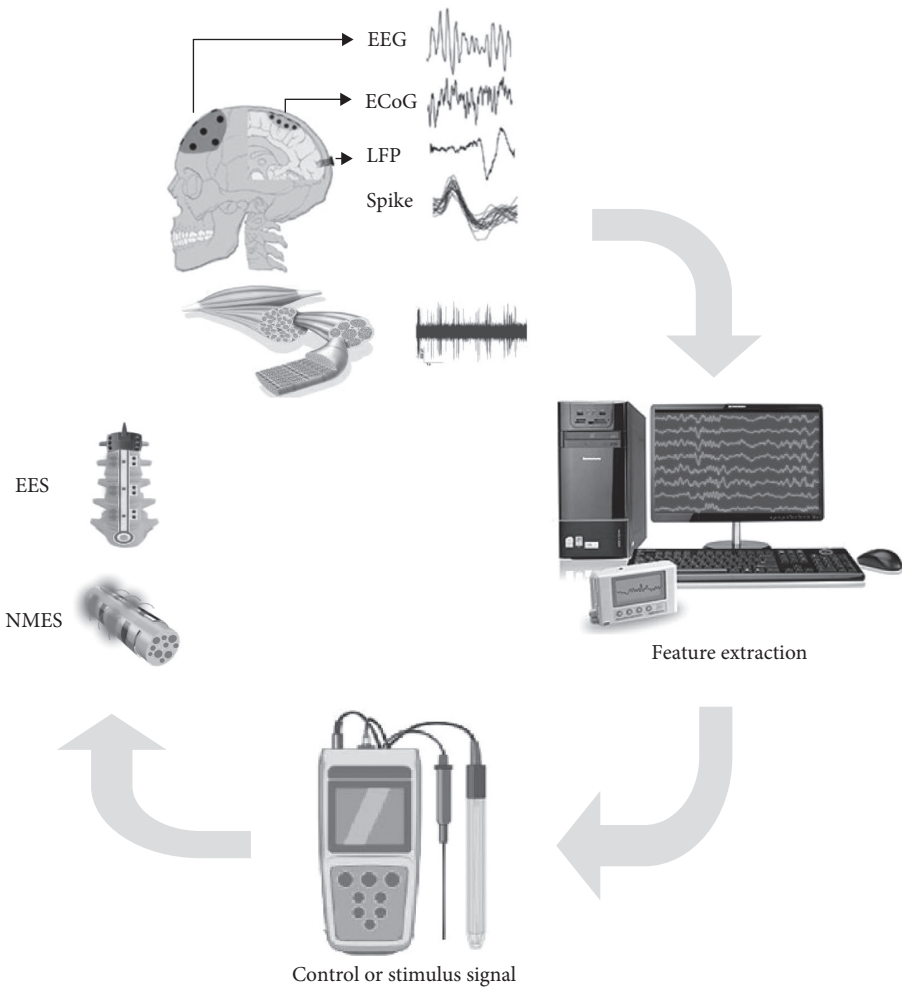



Figure 1: Electrical stimulation technology based on brain-computer interface.

Table 1: Different stimulation electrodes are compared.

| Electrode | Current (mA) | Advantage | Disadvantage |
|---|---|---|---|
| Implantable electrode | <25 | High specificity, can be used for a long time | Need surgery, may have inflammation |
| Nonimplantable electrodes | <120 | No surgery required | Not portable, prone to high current discomfort |

focused on stimulation frequency, stimulation amplitude, etc. [15]. Researchers have verified through experiments that closed-loop control may be more practical [31–33], as shown in Figure 2.

## 3. Types of Spinal Stimulation

Spinal cord electrical stimulation technology first appeared in the treatment of clinical cases in the 1970s. Initially, this technology was used to delay muscle atrophy and anti-inflammatory symptoms caused by muscle waste in patients with hemiplegia. It can also promote the speed of physical recovery [34–36]. Spinal cord electrical stimulation is a relatively new application for the recovery of motor function. Researchers use this technique to treat paraplegia, paralysis, or functional disorders caused by SCI or stroke [37–39].

Researchers have identified several effective forms of electrical stimulation, which can be divided into (1) epidural electrical stimulation, which places electrodes on the dural surface of the spinal cord, (2) transcutaneous electrical stimulation, where the electrode was placed on the skin above the vertebra, and (3) intraspinal electric stimulation, where the electrode was inserted into the spinal cord, as shown in Figure 3. Different stimulation methods require different stimulation parameters, and their mechanism of action will be different [40]. The current view is that epidural and transcutaneous electrical stimulation mainly stimulate the baseline excitability of the spinal cord and reactivate the intact circuit retained after SCI [41]. By activating the motor cistern and the neural network in the spinal cord, the electrical stimulation in the spinal cord can obtain the coordinated movement of the whole limb [42].

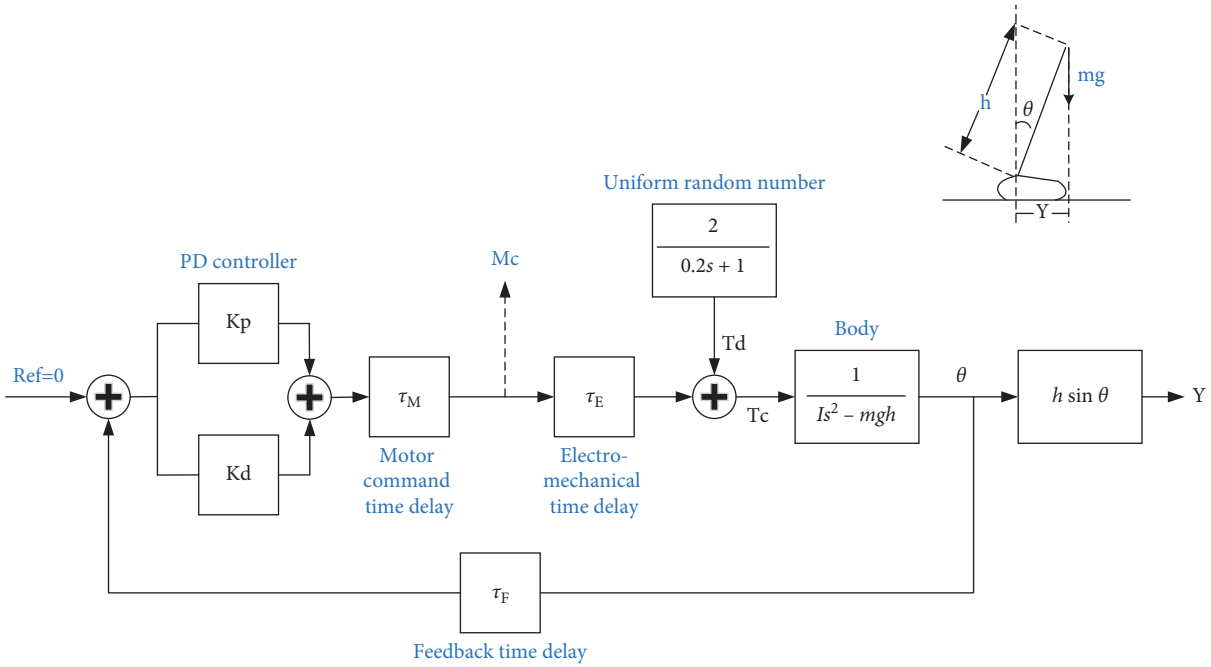



Figure 2: Inverted pendulum model and static attitude closed-loop control scheme. $Y$ is the horizontal center of gravity (COM) position, $h$ is the distance from COM to the ankle joint, $\theta$ is the body swing angle, $g$ is the acceleration caused by gravity, $Mc$ is the nerve movement command of the plantar flexor, and $Tc$ is the total torque of the ankle joint [31]. In order to maintain a safe standing, an effective and safe control stimulation method is necessary.

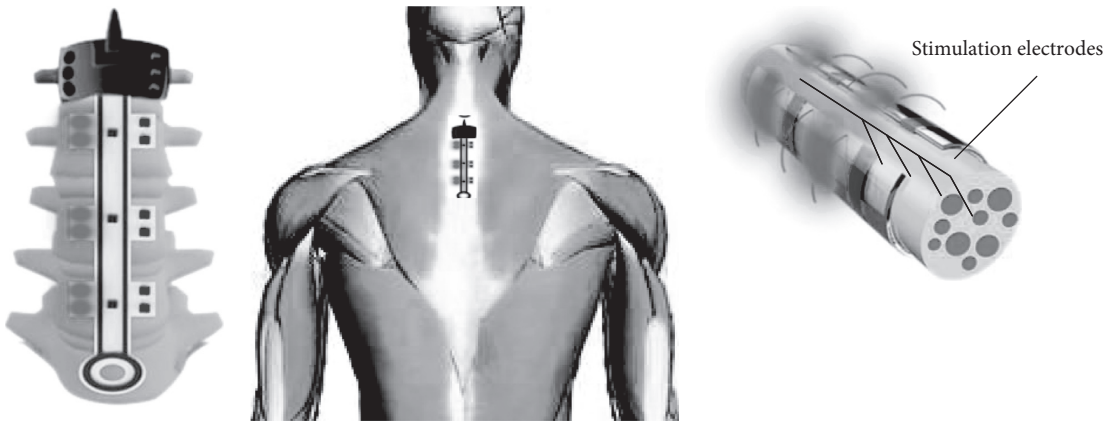



Figure 3: Types of spinal stimulation. (a) Epidural electrical stimulation; (b) transcutaneous electrical stimulation; (c) intraspinal electric stimulation.

*3.1. Epidural Electrical Stimulation.* The electrode is placed in the spinal cord epidural, through different current parameters to stimulate the electrode contact position, so as to achieve the effect of treatment or adjuvant therapy, and this technique is called epidural electrical stimulation. In the study of spinal cord function and its ecological environment, it is found that the functional modules of spinal cord can produce complex motor behavior without brain input [43, 44]. In fact, most of our daily actions escape our conscious attention. It also shows that it is possible to directly stimulate the downstream of the SCI plane beyond the SCI plane to cause the target movement of the moving organs.

The researchers applied epidural electrical stimulation to several animal models of SCI to verify the effectiveness of this technique [45–48]. It is undoubtedly a good choice to use small rodent models to verify this point of view. Gad et al. pointed out that, after complete resection of the middle thoracic spinal cord, adult rats can use spinal cord epidural stimulation to step on an electrically driven treadmill. At the same time, the effect of the combination of subthreshold stimulation intensity and spontaneous load proprioception to promote the walking and standing of hind limbs in paralyzed rats was studied [49]. Alam et al. used epidural electrical stimulation with different stimulation parameters (including different stimulation frequencies and single and

double stimulation electrodes) to train the forelimb grasping ability of rats with SCI. They found that cervical epidural stimulation increased the grip success rate compared with the nonstimulation condition, and bipolar stimulation (C6– C8+ and C6+ C8−) produced a larger spinal motor evoked potential and led to a higher success rate of arrival and mastery [50]. Studies in rats have found that epidural electrical stimulation of the lumbar spinal cord can reproduce the natural activation of synergistic muscle groups that produce motor function. Creating a suitable calculation model may be able to break through the complex mechanism of electrical stimulation. Capogrosso et al. designed a computational model and conducted in vivo experiments to study the types of fibers, neurons, and circuits that respond to EES recruitment. As shown in Figure 4, the model can predict the motor response induced by EES of various intensities and positions and provides a mechanical framework for the design of spinal nerve prosthesis system to improve standing and walking after nervous system diseases [51].

There is a certain gap between the anatomical structure of rodents and human beings. For rodents, spinal epidural electrical stimulation has achieved a good stimulation effect. However, in human testing, sometimes it does not work as well as in rodents. Studies have shown that EES cannot block proprioceptive input in rats but can block a large number of proprioceptive inputs in human body, which may be caused by different lengths of neural circuits [52], as shown in Figure 5. The results of Friedli et al. reveal that the fundamental differences in the essential anatomical and functional characteristics of the motor system between primates and rodents may lead to significant interspecific differences in the nature and degree of spinal cord repair mechanisms. It is suggested that although rodents are still very important in promoting regenerative therapy, the unique response of the corticospinal tract of primates after injury once again emphasizes the importance of primate models [53].

Some research teams set out to improve the electrical stimulation of primate models. Capogrosso et al. implanted a wirelessly controlled spinal cord stimulation system in rhesus monkeys with unilateral corticospinal tract lesions, which linked the online neural decoding of leg stretching and flexion movements to stimulation protocols that promote these movements. Without prior training for monkeys, the brain-spine interface restores the load-bearing movement of paralyzed legs on the treadmill and on the ground [54]. They made an in-depth study of the stimulation mechanism of the cervical spinal cord, showing that the lateral electrodes produce segmented recruitment of motoneurons in the upper limb and regulate the muscle response during exercise [55]. In order to explore the location of stimulation, Mesbah et al. carried out the research by means of statistics, discussed the mechanism of motor function recovery, and finally determined the exact position of the electrode [56]. Barra et al. recorded the muscle activity of rhesus monkeys during the hand-holding task and found that the task involved a stereotyped spatiotemporal map of motor neuron activation and characterized the specificity of EES delivery to the cervical vertebra segment. Finally, the team designed a stimulation scheme based on these results to reproduce the natural activation of motor neurons, thereby promoting the movement of the upper limb after injury [57]. This result also confirms the importance of spatiotemporal stimulation strategies for epidural electrical stimulation.

After exploration, epidural electrical stimulation can be systematically applied to human beings with SCI. As shown in Figure 6, patients with SCI are programmed into regular stimulation codes and sent to implantable pulse generators by collecting brain motor intention signals. The pulse generator connects a lamellar stimulation electrode attached to the dura matter of the L1~S2 segment of the spinal cord. By sending regular stimulation signals to the distributed contacts of the electrode, the lower limbs can be stimulated to make regular swinging movements.

In addition, there have been a number of cases to prove that spinal epidural stimulation has the effect of motor function recovery [59–63]. In 2011, Harkema et al. confirmed that a patient who has a complete motor injury but felt incomplete injury regained motor control after 2 years of complete paralysis by lumbosacral spinal cord epidural stimulation [64]. The David study showed that, 5 and 10 years after SCI, two adult women (48 years old and 52 years old, respectively) resumed a certain degree of voluntary exercise immediately after the implantation of the epidural spinal cord stimulation system, and there was no prescribed or obvious preconditioning training after the implantation of the stimulation system [65]. Possover stimulated the lumbosacral nerve in patients with chronic spinal cord injury for a long time. In the experiment, all patients with incomplete spinal cord injury regained voluntary control of previously paralyzed muscles [66]. The selection of appropriate stimulation strategy is very important to promote the standing of whole body weight-bearing stand with independent knee extension [67].

The Wagner team used an implantable pulse generator with real-time triggering to provide a series of spatially selective stimuli to the spinal cord in a certain time sequence. Within a week, this stimulation reestablished adaptive control of paralyzed muscles during ground walking. The exercise ability was improved during the rehabilitation period. A few months later, the participants regained voluntary control of previously paralyzed muscles without stimulation [58]. It is also shown that this kind of stimulation can produce long-term therapeutic effect.

Different stimulation parameters can produce different stimulation effects, and the influencing factors are complex and changeable, which brings difficulty to the control. Wenger et al. [46] developed a mechanical framework to optimize neural regulation in real time to achieve leg height simulation control during electrical stimulation. As shown in Figure 7, a parallel loop with a combination of robust feedback correction (PI controller) and predictive feedforward model is integrated by establishing a technical platform of embedded control strategy. The step height ($h_i$) is collected as the input, and the linear relationship between the EES frequency and the step height is refined and iterated. The controller combines the error ($e_i$) and the predicted value based on the current reference ($r_i$) to adjust the current

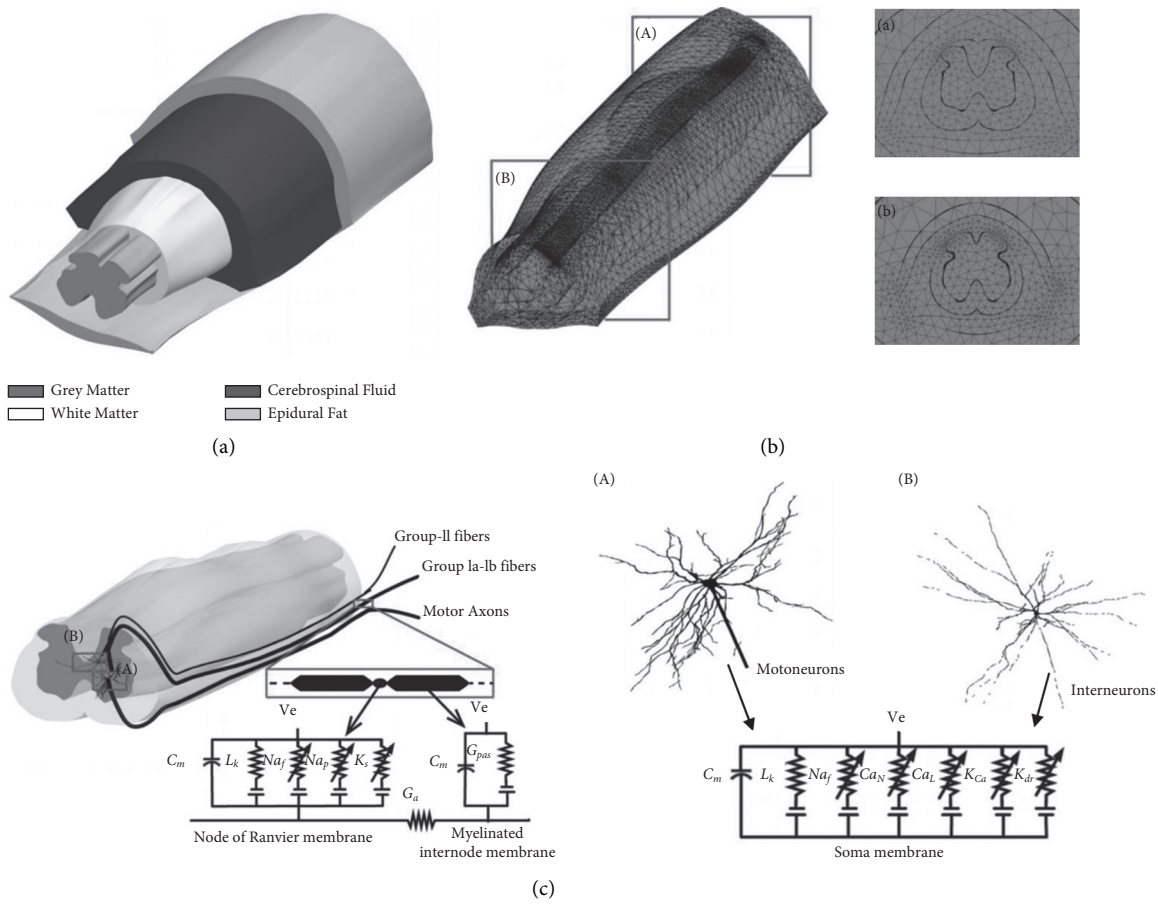



Figure 4: Characteristics of the computational model [51]. (a) Anatomically real spinal cord model, with different colors for gray matter and white matter, cerebrospinal fluid, and epidural fat. (b) The finite element structure of the spinal cord and the meshing of tetrahedral elements, as well as the cross-sectional schematic diagrams of the spine segments L2 (A) and S1 (B). (c) Modeling of afferent and efferent nerves, and establishing real geometric models of afferent and efferent fibers, $\alpha$ motor neurons and interneurons. The membrane potential was calculated by the Hodgkin-Huxley equation. The afferent fiber enters the spinal cord below the spine segment S1 and travels longitudinally before bending in the gray matter of its target segment. Interneurons are located in layers I to III and VII, and the efferent axons expand to the dorsal side or pass through the midline of the spinal cord, respectively.

frequency ($f_i$) to the stimulus frequency ($f_{r+1}$) in real time, so as to keep the step height within the set range of parameters.

*3.2. Transcutaneous Stimulation.* Transcutaneous stimulation refers to an electrode placed on the skin above the spinal cord, which penetrates the skin to the spinal cord to produce movement or grip [68]. To some extent, spinal transcutaneous stimulation increased the recovery of motor function in patients with SCI similar to epidural stimulation [69, 70].

Studies have shown that surface spinal stimulation, similar to epidural stimulation, can be used to enhance muscle recruitment without adversely affecting residual motion control in incomplete SCI [71]. The researchers performed percutaneous electrical stimulation of the spine in 15 SCI patients of different severity to restore their independent posture and found high levels of leg muscle activity during the standing weight shift achieved by spinal stimulation, depending on the muscle load. This suggests that transcutaneous stimulation can be used to regulate the lumbar spinal cord neural network to promote self-standing after chronic movement and complete paralysis [72]. Spinal cord stimulation and drugs can enhance the motor state of the subjects when they step into the exoskeleton. In addition, stimulation improved the coordination pattern of lower limb muscles, resulting in more continuous, smooth stepping movement of the exoskeleton [73]. This method can rejoin and train the spinal cord motor network of individuals with long-term complete motor paralysis. In some special cases, command selective activation of gastrocnemius and soleus was also observed [74].

The latest study has found that the use of painless transcutaneous electrical stimulation strategy to regulate the physiological state of the spinal cord can promote stampede in noninjured objects whose legs are in a gravity-neutral position, and multipoint stimulation can more effectively induce stepping movement [31]. Compared with the stampede driven by the robot alone, 30 Hz percutaneous spinal cord stimulation can increase the number of rhythmic

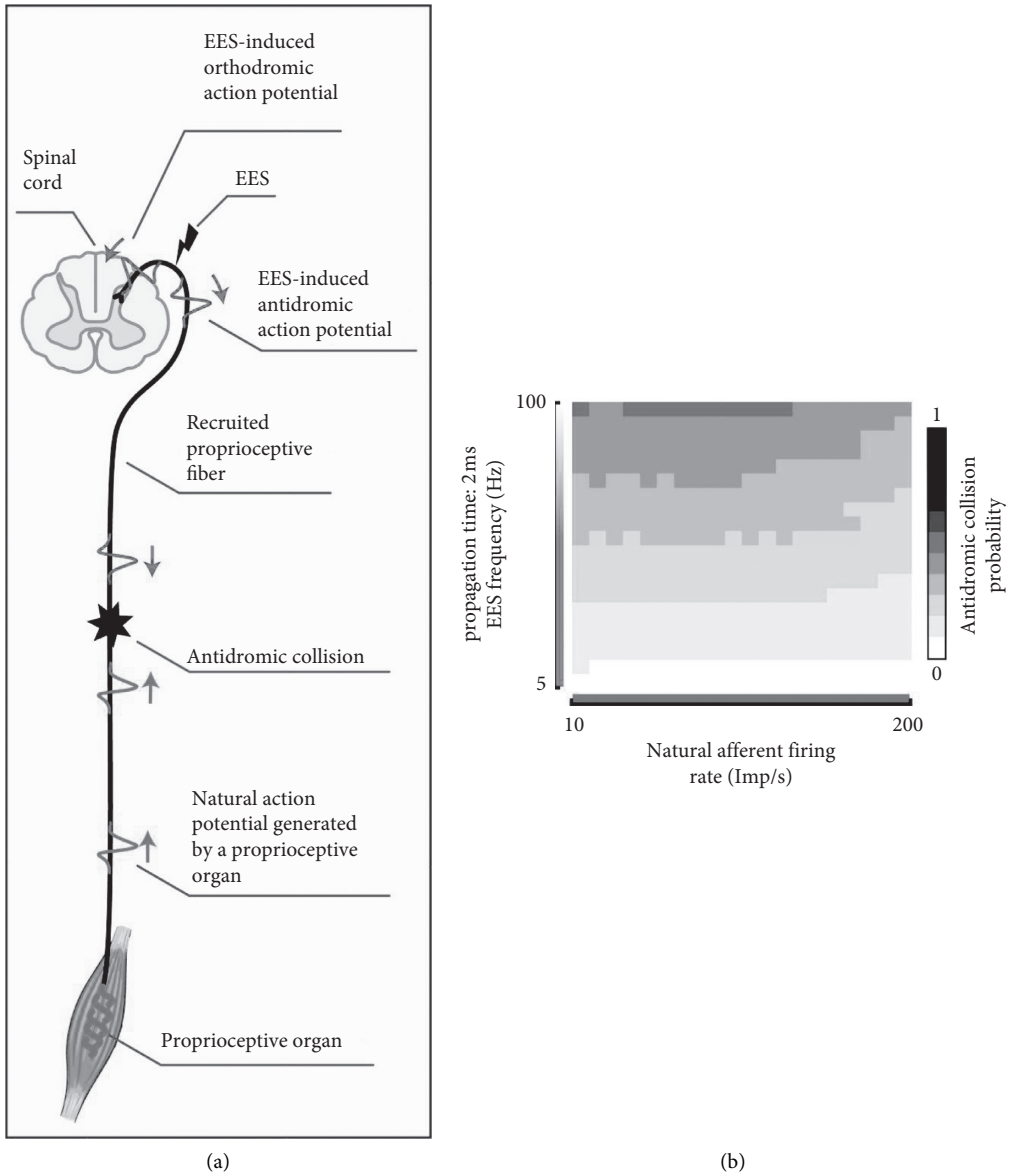
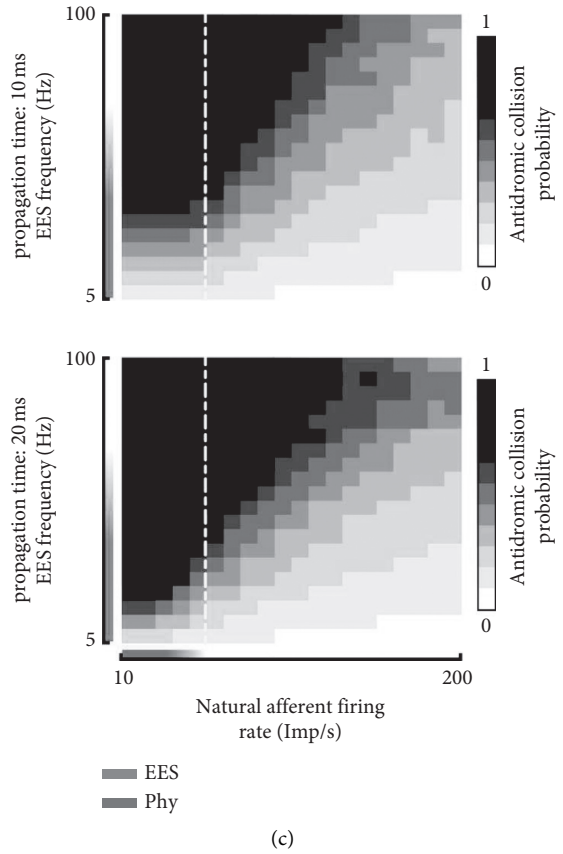



Figure 5: Continued.




(c)

Figure 5: Probability of antidromic collisions during EES in rats and humans [52]. (a) Schematic diagram of the conflict between EES-induced action potentials and naturally occurring action potential; (b) the probability of action potential conflict in mice (the transmission time of action potential in mice is about 2 ms); (c) the probability of action potential conflict in human body (the propagation time of action potential is about 10 ms in the proximal muscle circuit and about 20 ms in the distal muscle circuit).

muscles, enhance the activity of thigh muscles, and inhibit cloning. This shows that the rhythmic movement pattern produced by forced stampede is the response of spinal reflex circuit to cyclic sensory feedback. Providing additional stimulation and participation in spinal rhythm generation networks, combined with treadmill training, may increase rehabilitation outcome after severe SCI [75]. Similar to epidural electrical stimulation, transcutaneous stimulation also requires intensive treatment and physical exercise in the process of spinal cord stimulation to maximize motor function recovery [73]. After a long period of training, the noninvasive percutaneous electrical stimulation of the spinal network restores the movement and function of the hands and arms. This method of stimulation may be more suitable for patients with complete paralysis and long-term spinal cord injury. In addition, muscle spasm was reduced compared with invasive electrical stimulation, and the therapeutic effects of all six participants demonstrated long-term neuroplastic-mediated functional recovery [76].

Transcutaneous stimulation is a noninvasive stimulation that does not require surgery, so it is of sufficient interest to many people. However, percutaneous muscle stimulation has several practical limitations. Specifically, the skin has higher resistance than nerve tissue [77]. Therefore, most of them need higher current stimulation (>30 mA) to have an effect [78]. Sometimes, there is pain because of the location and individual differences of the electrodes [79]. Electrode placement and current intensity seem to be the key parameters to achieve selective response [80].

*3.3. Intraspinal Electric Stimulation.* The electrical stimulation in the spinal cord is different from the epidural electrical stimulation in that the electrode is implanted into the spinal cord, and the nerve region in the spinal cord is directly stimulated by the electrode power supply, because direct stimulation of the spinal cord requires a lower threshold current to induce movement than epidural stimulation, as shown in Table 2 [81].

So far, there are few studies on spinal cord electrical stimulation in humans, but the spinal cord stimulation experiments in animals provide a good research basis. Holinski used fine penetrating electrodes to stimulate the enlarged ventral horn of the waist of 5 adult anesthetized cats. The stimulation current of $<100\,\mu$A activates the inherent motor network in the spinal cord, which can produce functional ground walking with bilateral coordination. This suggests that ISES may be an effective intervention for functional walking after SCI [82], as shown in Figure 8.

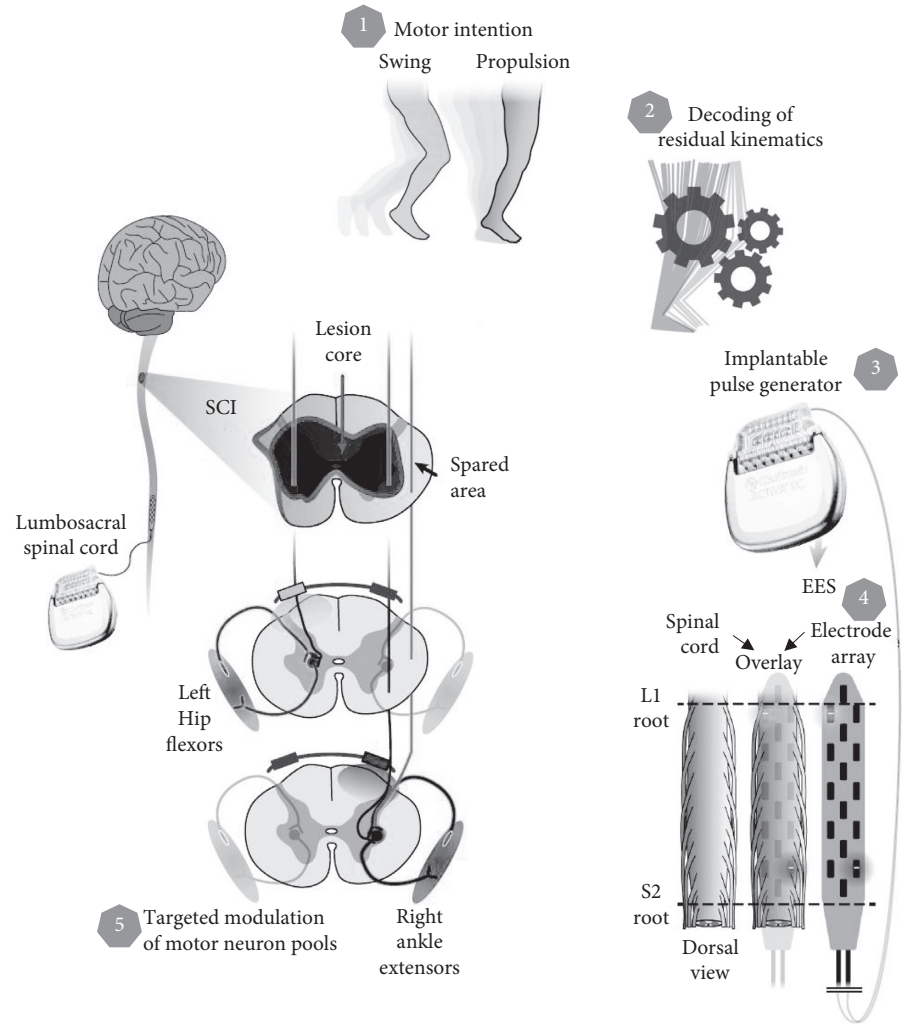



Figure 6: Schematic diagram of the structure of epidural electrical stimulation [58].

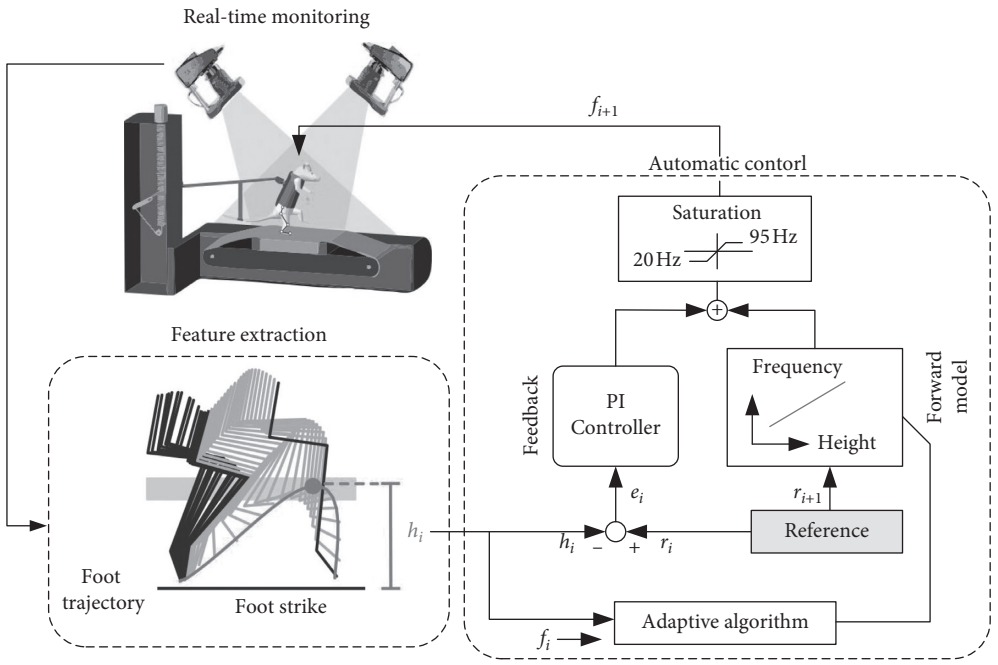



Figure 7: Rat real-time feedback control platform [46].

Table 2: Comparison of stimulation current threshold of different stimulation modes of spinal cord [81].

| Motion type | Threshold current ($\mu$A) | | Spine segment |
|---|---|---|---|
| | Epidural electrical stimulation (EES) | Intraspinal electric stimulation (ISES) | |
| Hip flexion | 250~320 | 88~110 | T13 |
| Hip extension | 180~282 | 65~90 | L2 |
| Hip adduction | 210~260 | 82~103 | L1 |
| Hip abduction | 230~278 | 80~90 | T13<br>L1 |

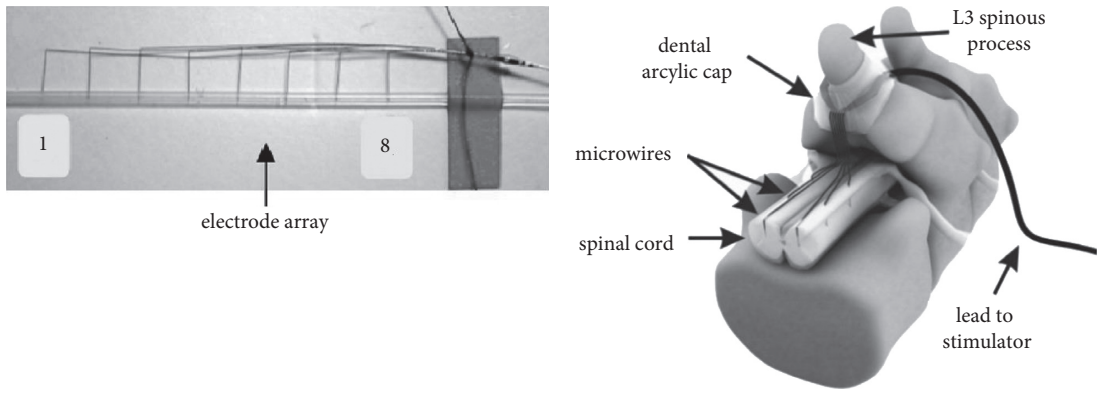



Figure 8: Electrode and implant location [82].

Dalrymple used electrical stimulation in the spinal cord and developed control strategies to achieve a cat model of hemisection SCI for weight-bearing pedaling [83]. In addition, in a number of rodent experimental studies, intraspinal cord stimulation is used after specific exercise [84, 85]. There will be a lasting improvement in forelimb movement a few weeks after the stimulation stops [85, 86]. These results suggest the potential long-term effect of intraspinal stimulation. Pairing intraspinal cord stimulation with rehabilitation physical training may have additional benefits, and a specific study of the joint approach will be useful.

Different spinal cord stimulation methods may have different mechanisms; the current view in this field is that intraspinal cord stimulation may activate the motor cistern and the inherent neural network in the spinal cord to complete coordinated limb movement [87]. The epidural electrical stimulation may increase the baseline excitability of the spinal cord, so as to keep the nerve triggering movement intact after SCI [41, 88].

## 4. Challenges and Prospects

Spinal cord electrical stimulation brings new treatment and hopes to patients with SCI, which is very commendable. With the improvement of technology, new stimulators, electrodes, and other devices will be created, and new stimulation strategies will also be proposed [55]. This will further promote the development of electrical stimulation technology and bring good news to patients. At present, although the results of spinal electrical stimulation are expected to play an important role in enhancing motor recovery, the potential for long-term and continuous improvement caused by this technique is likely to require further improvement, such as the biocompatibility of implants [89], noninvasive effectiveness [78, 79]. In addition, a combination of treatments, such as exoskeletons [90, 91], may be required, as well as medication [92], motor training [58, 93].

*4.1. Challenge.* The current spinal cord electrical stimulation technology needs to be further explored: first of all, the motor function repair of patients has not formed a complete circuit. The patient passively receives the stimulation signal or instruction to make the corresponding action, and it is difficult to receive the feedback signal of the stimulated muscle. Although Urbin et al. [94] try to collect the EMG of the target muscle and the foot ground-reaction forces (GRF) signals implanted in the body to build the somatosensory afferent pathway, there is still a long way to go before the recovery of sensory function. In addition, most of the existing motor function reconstruction systems are carried out in the laboratory environment, although some environments approximately simulate the reality. However, experiments conducted in a laboratory environment do not take into account all the availability factors that weaken the functional reconstruction system, including artificial factors, nonbrain effects, and the mental state of the user. Therefore, it is meaningful to add mental and psychological assessment options in future research plans, and a professional training team for the use of the functional reconstruction system should be prepared to make patients quickly and skillfully

operate the system and put it into practice as soon as possible. Finally, for the realization of motor function, all kinds of rehabilitation methods need an appropriate control strategy [83, 95–97]. At present, human action recognition is not accurate in dynamic situations. Using different algorithm recognition signals to recognize human activities and motion parameters accurately to guide or remotely guide electrical stimulation is a challenge. In the later development process, breakthroughs in signal processing will be an important development direction for electrical stimulation. It will be a long-standing challenge to be able to process stimulus signals efficiently and accurately, and to develop stable and efficient stimulus algorithms. In-depth discussion of the contraction mechanism and internal relationship of each muscle in the muscle group after electrical stimulation is indispensable for patients to achieve motor function. The related research on the mechanism of electrical stimulation should be more detailed and rich, which will be conducive to the development of motor function rehabilitation technology of electrical stimulation.

### 4.2. Future Direction

*4.2.1. Electrical Stimulation Technology Based on BCI.* Spinal cord electrical stimulation technology based on brain-computer interface will be an important development direction in the future, and the collected EEG signals can be used as control signals to guide stimulation strategies. A research team has applied the collected EEG signals to spinal cord electrical stimulation and made a breakthrough in the treatment of paralyzed patients. As an example, researcher Millan and others collected EEG signals from patients with upper limb paralysis caused by stroke and decoded motor consciousness signals. The research team combined motor consciousness signals based on BCI with spinal cord electrical stimulation to induce significant recovery of upper limb motor function, and the effect was still significant in 6 to 12 months after treatment [98].

This shows that this treatment is effective and persistent, but it is worth noting that the design of brain signal detection and control signal is difficult. This also opens up new doors and challenges for the research direction of spinal cord electrical stimulation technology, such as the detection of EEG signals [99], new control method [97], and inflammatory reaction of intrusive electrode [100, 101]. In addition, spinal cord electrical stimulation based on BCI requires patients to have good enough motor imagination to generate motor electrical signals; otherwise, it is a difficult problem to control spinal cord electrical stimulation [102–104]. The motor function recovery of FES technology based on BCI has achieved some results, but the curative effect and repair mechanism are not completely clear [98, 105, 106]. This requires further research by researchers, and it may be a good method to establish a good computational model to explore the signal loop [51, 107].

*4.2.2. Electrical Stimulation and Drug Therapy.* In some cases, these pharmacological treatments can enhance the effects of exercise training and epidural stimulation [108]. This combination intervention may target the internal environment of the injured spinal cord to further increase its excitability and enhance the effectiveness of therapeutic electrical stimulation. It was found that chondroitinase gene therapy could significantly enhance the upper limb function, improve the forelimb performance and grip, increase the spinal information transmission, and reduce the pathology of the injured spine after spinal cord contusion in rats [109]. Chondroitinase transmitted by canine mucosal olfactory ensheathing cells combined with rehabilitation can enhance the recovery after SCI [110]. In addition, stem cell transplantation has always been a hot topic in the treatment of SCI and has an early positive effect [111]. Electrical stimulation has a positive effect on stem cell transplantation, and stem cell and neural progenitor cell therapy can also improve hand stretching performance and hand function [112]. Based on these findings, it may be possible to use electrical stimulation to direct stem cells to damaged parts of the body, becoming an important treatment.

*4.2.3. Development of Neuroprosthesis.* Lagging science and technology and production capacity hinder the development of electrical stimulation system nerve prosthesis and cannot achieve the desired goal, but with the development of technology, there may be a big breakthrough. The composition or auxiliary devices of some electrical stimulation systems will be improved, such as electrodes for transmitting stimulus signals [113], stimulators for generating stimulus pulses [114], sensors for transmitting control source signals [25], and exoskeletons for sports training [115].

In addition, feedback control based on neural prosthesis may promote the progress of electrical stimulation technology. Many researchers use the collected needle electrode EMG for the detection and diagnosis of SCI related diseases, including muscle activation evaluation [116], improved EMG control [117, 118], evaluation of motor neuron loss [119], muscle fatigue detection [120], prosthetic EMG control [121], and other fields. The acquisition of surface EMG is closely related to the placement of electrodes, electrode materials, and the number of channels [122–125]. It can be used as a reference for the control of electrical stimulation signal and the evaluation of rehabilitation effect in exercise rehabilitation.

## 5. Summary

In summary, spinal cord injury brings tremendous physical and psychological pressure to patients, causing inconvenience and even loss of life. Researchers have made great efforts to improve the lives of patients with spinal cord injury in their respective fields. This article first introduces the different causes of spinal cord injury and various existing treatment methods and briefly introduces the current understanding of its origin and mechanism of action. Most importantly, this article provides a detailed review of the treatment methods of spinal cord electrical stimulation. Several different spinal cord electrical stimulation methods and corresponding electrical stimulation systems were

discussed, respectively, and the stimulation effects were analyzed. Finally, this article analyzes the challenges and future directions of the spinal cord electrical stimulation system and gives possible potential research programs. Whether it is advanced stimulation equipment or sophisticated stimulation strategy, spinal cord electrical stimulation technology provides practical help for people with mobility difficulties. This technology helps patients achieve things that were impossible before and even participate in daily life, greatly reducing the burden of life and improving self-confidence. Looking to the future, in order to solve the multifaceted effects of SCI and restore the complete function, a combination of multiple treatments and long-term treatment and training may be needed.

The latest treatment method combining cell transplantation and electrical stimulation has brought new treatment improvements and has made progress in animal models [126]. The pathophysiological mechanism of spinal cord injury is complicated. So far, the mechanism and process are not fully understood [127]. It is worth noting that comprehensive treatments such as surgical biomaterials, stem cell transplantation, drug therapy, and rehabilitation exercises have been proven effective. And the functional remodeling of neural circuits relies heavily on rehabilitation exercises, such as electrical stimulation, physical training, and brain-computer interfaces. In future research, we should pay attention to the individual differences of patients. For patients with incomplete injuries or no ruptures, more attention may be paid to physical rehabilitation training, and electrical stimulation systems are used for auxiliary treatment. For patients with large-scale fracture injuries, surgery or bioprosthetic transplantation may be required to help establish new neural circuits, maximize the effect of rehabilitation, and ultimately promote the research of spinal cord injury to have greater breakthroughs in treatment and rehabilitation and more fields.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was funded by a grant (No. 2019M660392) from China Postdoctoral Science Foundation and the Open Research Project of the State Key Laboratory of Media Convergence and Communication, Communication University of China (No. SKLMCC2020KF002).

## References

[1] F. S. Chang, Q. Zhang, M. Sun et al., "Epidemiological study of spinal cord injury individuals from halfway houses in Shanghai, China," *Journal of Spinal Cord Medicine*, vol. 41, no. 4, pp. 450–458, 2017.

[2] H. Chikuda, J. Ohya, H. Horiguchi et al., "Ischemic stroke after cervical spine injury: analysis of 11,005 patients using the Japanese diagnosis procedure combination database," *The Spine Journal*, vol. 14, no. 10, pp. 2275–2280, 2014.

[3] A. E. Smith, I. R. Molton, and M. P. Jensen, "Self-reported incidence and age of onset of chronic comorbid medical conditions in adults aging with long-term physical disability," *Disability and Health Journal*, vol. 9, no. 3, pp. 533–538, 2016.

[4] J. T. Street, V. K. Noonan, A. Cheung, C. G. Fisher, and M. F. Dvorak, "Incidence of acute care adverse events and long-term health-related quality of life in patients with TSCI," *The Spine Journal*, vol. 15, no. 5, pp. 923–932, 2015.

[5] H. Wang, X. Liu, Y. Zhao et al., "Incidence and pattern of traumatic spinal fractures and associated spinal cord injury resulting from motor vehicle collisions in China over 11 years: an observational study," *Medicine*, vol. 95, no. 43, pp. 1–5, 2016.

[6] National Spinal Cord Injury Statistical Center, "Spinal cord injury (SCI) facts and figures at a glance," *The Journal of the American Paraplegia Society*, vol. 39, no. 2, pp. 243-244, 2016.

[7] J. R. Wilson, C. D. Witiw, J. Badhiwala, B. K. Kwon, M. G. Fehlings, and J. S. Harrop, "Early surgery for traumatic spinal cord injury: where are we now?" *Global Spine Journal*, vol. 10, no. 1, pp. 84S–91S, 2020.

[8] E. Hayta and H. Elden, "Acute spinal cord injury: a review of pathophysiology and potential of non-steroidal anti-inflammatory drugs for pharmacological intervention," *Journal of Chemical Neuroanatomy*, vol. 87, pp. 25–31, 2018.

[9] M. Franz, L. Richner, M. Wirz et al., "Physical therapy is targeted and adjusted over time for the rehabilitation of locomotor function in acute spinal cord injury interventions in physical and sports therapy," *Spinal Cord*, vol. 56, no. 2, pp. 158–167, 2018.

[10] J. H. Badhiwala, J. R. Wilson, C. D. Witiw et al., "The influence of timing of surgical decompression for acute spinal cord injury: a pooled analysis of individual patient data," *The Lancet Neurology*, vol. 20, no. 2, pp. 117–126, 2021.

[11] S. Silvestro, P. Bramanti, O. Trubiani, and E. Mazzon, "Stem cells therapy for spinal cord injury: an overview of clinical trials," *International Journal of Molecular Sciences*, vol. 21, no. 2, 2020.

[12] B. Yang, F. Zhang, F. Cheng et al., "Strategies and prospects of effective neural circuits reconstruction after spinal cord injury," *Cell Death & Disease*, vol. 11, no. 6, p. 439, 2020.

[13] G. David, S. Mohammadi, A. R. Martin et al., "Traumatic and nontraumatic spinal cord injury: pathological insights from neuroimaging," *Nature Reviews Neurology*, vol. 15, no. 12, pp. 718–731, 2019.

[14] N. D. James, S. B. Mcmahon, E. C. Field-Fote, and E. J. Bradbury, "Neuromodulation in the restoration of function after spinal cord injury," *The Lancet Neurology*, vol. 17, no. 10, pp. 905–917, 2018.

[15] C. H. Ho, R. J. Triolo, A. L. Elias et al., "Functional electrical stimulation and spinal cord injury," *Physical Medicine and Rehabilitation Clinics of North America*, vol. 25, no. 3, pp. 631–654, 2014.

[16] E. F. Hodkin, Y. Lei, J. Humby et al., "Automated FES for upper limb rehabilitation following stroke and spinal cord injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 5, pp. 1067–1074, 2018.

[17] N. S. Makowski, L. M. Lombardo, K. M. Foglyano et al., "Walking after incomplete spinal cord injury with an implanted neuromuscular electrical stimulation system and a hinged knee replacement: a single-subject study," *Spinal Cord Series and Cases*, vol. 6, no. 1, p. 86, 2020.

[18] C. A. Angeli, M. Boakye, R. A. Morton et al., "Recovery of over-ground walking after chronic motor complete spinal

cord injury," *New England Journal of Medicine*, vol. 379, no. 13, pp. 1244–1250, 2018.

[19] Y.-T. Zhang, H. Jin, J.-H. Wang et al., "Tail nerve electrical stimulation and electro-acupuncture can protect spinal motor neurons and alleviate muscle atrophy after spinal cord transection in rats," *Neural Plasticity*, vol. 2017, Article ID 7351238, 11 pages, 2017.

[20] C. Daia, A. M. Bumbea, C. D. Badiu, C. Ciobotaru, and G. Onose, "Interferential electrical stimulation for improved bladder management following spinal cord injury," *Biomedical Reports*, vol. 11, no. 3, pp. 115–122, 2019.

[21] G. Buzsáki, C. A. Anastassiou, and C. Koch, "The origin of extracellular fields and currents-EEG, ECoG, LFP and spikes," *Nature Reviews Neuroscience*, vol. 13, no. 6, pp. 407–420, 2012.

[22] C. Moritz, "A giant step for spinal cord injury research," *Nature Neuroscience*, vol. 21, no. 12, pp. 1645-1646, 2018.

[23] M. R. Popovic, T. Keller, I. P. I. Papas, V. Dietz, and M. Morari, "Surface-stimulation technology for grasping and walking neuroprostheses," *IEEE Engineering in Medicine and Biology Magazine*, vol. 20, no. 1, pp. 82–93, 2001.

[24] C. Wiesener and T. Schauer, "The cybathlon RehaBike: inertial-sensor-driven functional electrical stimulation cycling by team hasomed," *IEEE Robotics and Automation Magazine*, vol. 24, no. 4, pp. 49–57, 2017.

[25] C. Wiesener, T. Seel, J. Axelgaard, R. Horton, A. Niedeggen, and T. Schauer, "An inertial sensor-based trigger algorithm for functional electrical stimulation-assisted swimming in paraplegics," *IFAC-PapersOnLine*, vol. 51, no. 34, pp. 278–283, 2019.

[26] T. Schauer, "Sensing motion and muscle activity for feedback control of functional electrical stimulation: ten years of experience in Berlin," *Annual Reviews in Control*, vol. 44, pp. 355–374, 2017.

[27] L. Bi, A. G. Feleke, and C. Guan, "A review on EMG-based motor intention prediction of continuous human upper limb motion for human-robot collaboration," *Biomedical Signal Processing and Control*, vol. 51, pp. 113–127, 2019.

[28] R. E. Singh, K. Iqbal, G. White, and T. E. Hutchinson, "A systematic review on muscle synergies: from building blocks of motor behavior to a neurorehabilitation tool," *Applied Bionics and Biomechanics*, vol. 2018, Article ID 3615368, 15 pages, 2018.

[29] G. P. Braz, M. Russold, and G. M. Davis, "Functional electrical stimulation control of standing and stepping after spinal cord injury: a review of technical characteristics," *Neuromodulation: Technology at the Neural Interface*, vol. 12, no. 3, pp. 180–190, 2009.

[30] A. R. Kralj and T. Bajd, *Functional Electrical Stimulation: Standing and Walking after Spinal Cord Injury*, CRC Press, Boca Raton, FL, USA, 1989.

[31] A. H. Vette, K. Masani, J.-Y. Kim, and M. R. Popovic, "Closed-loop control of functional electrical stimulation-assisted arm-free standing in individuals with spinal cord injury: a feasibility study," *Neuromodulation: Technology at the Neural Interface*, vol. 12, no. 1, pp. 22–32, 2009.

[32] C. L. Lynch, M. R. Popovic, and R. Engineering, "A comparison of closed-loop control algorithms for regulating electrically stimulated knee movements in individuals with spinal cord injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 20, no. 4, pp. 539–548, 2012.

[33] Y. Hu, H. Su, J. Fu, H. R. Karimi, G. Ferrigno, and E. D. Momi, "Nonlinear model predictive control for mobile medical robot using neural optimization," *IEEE Transactions on Industrial Electronics*, vol. 68, no. 12, pp. 12636–12645, 2020.

[34] O. Devinsky, "Electrical and magnetic stimulation of the central nervous system. Historical overview," *Advances in Neurology*, vol. 63, pp. 1–16, 1993.

[35] C. N. Shealy, J. T. Mortimer, and N. R. Hagfors, "Dorsal column electroanalgesia," *Journal of Neurosurgery*, vol. 32, no. 5, pp. 560–564, 1970.

[36] C. N. Shealy, J. T. Mortimer, and J. B. Reswick, "Electrical inhibition of pain by stimulation of the dorsal columns," *Anesthesia & Analgesia*, vol. 46, no. 4, pp. 489–491, 1967.

[37] N. Donaldson, T. Perkins, R. Fitzwater, D. E. Wood, and F. Middleton, "FES cycling may promote recovery of leg function after incomplete spinal cord injury," *Spinal Cord*, vol. 38, no. 11, pp. 680–682, 2000.

[38] A. S. Gorgey, D. R. Dolbow, J. D. Dolbow, R. K. Khalil, and D. R. Gater, "The effects of electrical stimulation on body composition and metabolic profile after spinal cord injury-Part II," *The Journal of Spinal Cord Medicine*, vol. 38, no. 1, pp. 23–37, 2015.

[39] A. Sivaramakrishnan, J. M. Solomon, and N. Manikandan, "Comparison of transcutaneous electrical nerve stimulation (TENS) and functional electrical stimulation (FES) for spasticity in spinal cord injury-a pilot randomized cross-over trial," *The Journal of Spinal Cord Medicine*, vol. 41, no. 4, pp. 397–406, 2018.

[40] A. N. Sharpe and A. Jackson, "Upper-limb muscle responses to epidural, subdural and intraspinal stimulation of the cervical spinal cord," *Journal of Neural Engineering*, vol. 11, no. 1, Article ID 016005, 2014.

[41] V. R. Edgerton, G. Courtine, Y. P. Gerasimenko et al., "Training locomotor networks," *Brain Research Reviews*, vol. 57, no. 1, pp. 241–254, 2008.

[42] C. H. Tator, K. Minassian, and V. K. Mushahwar, "Chapter 18-spinal cord stimulation: therapeutic benefits and movement generation after spinal cord injury," in *Handbook of Clinical Neurology*, J. Verhaagen and J. W. Mcdonald, Eds., Elsevier, Amsterdam, Netherlands, pp. 283–296, 2012.

[43] S. Grillner, "Biological pattern generation: the cellular and computational logic of networks in motion," *Neuron*, vol. 52, no. 5, pp. 751–766, 2006.

[44] L. R. Hochberg, D. Bacher, B. Jarosiewicz et al., "Reach and grasp by people with tetraplegia using a neurally controlled robotic arm," *Nature*, vol. 485, no. 7398, pp. 372–375, 2012.

[45] N. Wenger, E. M. Moraud, J. Gandar et al., "Spatiotemporal neuromodulation therapies engaging muscle synergies improve motor control after spinal cord injury," *Nature Medicine*, vol. 22, no. 2, pp. 138–145, 2016.

[46] N. Wenger, E. M. Moraud, S. Raspopovic et al., "Closed-loop neuromodulation of spinal sensorimotor circuits controls refined locomotion after complete spinal cord injury," *Science Translational Medicine*, vol. 6, no. 255, 2014.

[47] E. M. Moraud, M. Capogrosso, E. Formento et al., "Mechanisms underlying the neuromodulation of spinal circuits for correcting gait and balance deficits after spinal cord injury," *Neuron*, vol. 89, no. 4, pp. 814–828, 2016.

[48] P. Gad, J. Choe, P. Shah et al., "Sub-threshold spinal cord stimulation facilitates spontaneous motor activity in spinal rats," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, 2013.

[49] P. Gad, J. Choe, P. Shah et al., "Sub-threshold spinal cord stimulation facilitates spontaneous motor activity in spinal rats," *Journal of NeuroEngineering and Rehabilitation*, vol. 10, 2013.

[50] M. Alam, G. Garcia-Alias, B. Jin et al., "Electrical neuromodulation of the cervical spinal cord facilitates forelimb skilled function recovery in spinal cord injured rats," *Experimental Neurology*, vol. 291, pp. 141–150, 2017.

[51] M. Capogrosso, N. Wenger, S. Raspopovic et al., "A computational model for epidural electrical stimulation of spinal sensorimotor circuits," *Journal of Neuroscience*, vol. 33, no. 49, pp. 19326–19340, 2013.

[52] E. Formento, K. Minassian, F. Wagner et al., "Electrical spinal cord stimulation must preserve proprioception to enable locomotion in humans with spinal cord injury," *Nature Neuroscience*, vol. 21, no. 12, pp. 1728–1741, 2018.

[53] L. Friedli, E. S. Rosenzweig, and Q. Barraud, "Pronounced species divergence in corticospinal tract reorganization and functional recovery after lateralized spinal cord injury favors primates," *Science Translational Medicine*, vol. 7, no. 302, 2015.

[54] M. Capogrosso, T. Milekovic, D. Borton et al., "A brain-spine interface alleviating gait deficits after spinal cord injury in primates," *Nature*, vol. 539, no. 7628, pp. 284–288, 2016.

[55] N. Greiner, B. Barra, G. Schiavone et al., "Recruitment of upper-limb motoneurons with epidural electrical stimulation of the cervical spinal cord," *Nature Communications*, vol. 12, no. 1, p. 435, 2021.

[56] S. Mesbah, T. Ball, C. Angeli et al., "Predictors of volitional motor recovery with epidural stimulation in individuals with chronic spinal cord injury," *Brain*, vol. 144, no. 2, pp. 420–433, 2021.

[57] B. Barra, C. Roux, M. Kaeser et al., "Selective recruitment of arm motoneurons in nonhuman primates using epidural electrical stimulation of the cervical spinal cord," in *Proceedings of the Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 1424–1427, Honolulu, HI, USA, 2018.

[58] F. B. Wagner, J.-B. Mignardot, C. G. Le Goff-Mignardot et al., "Targeted neurotechnology restores walking in humans with spinal cord injury," *Nature*, vol. 563, no. 7729, pp. 65–71, 2018.

[59] G. Courtine and M. V. Sofroniew, "Spinal cord repair: advances in biology and technology," *Nature Medicine*, vol. 25, no. 6, pp. 898–908, 2019.

[60] J. D. Turtle, M. K. Henwood, M. M. Strain, Y.-J. Huang, R. C. Miranda, and J. W. Grau, "Engaging pain fibers after a spinal cord injury fosters hemorrhage and expands the area of secondary injury," *Experimental Neurology*, vol. 311, pp. 115–124, 2019.

[61] A. N. Herrity, C. S. Williams, C. A. Angeli, and S. J. Harkemaet, "Lumbosacral spinal cord epidural stimulation improves voiding function after human spinal cord injury," *Scientific Reports*, vol. 8, 2018.

[62] M. Rath, A. H. Vette, S. Ramasubramaniam et al., "Trunk stability enabled by noninvasive spinal electrical stimulation after spinal cord injury," *Journal of Neurotrauma*, vol. 35, no. 21, pp. 2540–2553, 2018.

[63] G. Taccola, D. Sayenko, P. Gad, Y. Gerasimenko, and V. R. Edgerton, "And yet it moves: recovery of volitional control after spinal cord injury," *Progress in Neurobiology*, vol. 160, pp. 64–81, 2018.

[64] S. Harkema, Y. Gerasimenko, J. Hodes et al., "Effect of epidural stimulation of the lumbosacral spinal cord on voluntary movement, standing, and assisted stepping after motor complete paraplegia: a case study," *The Lancet*, vol. 377, no. 9781, pp. 1938–1947, 2011.

[65] D. Darrow, D. Balser, T. I. Netoff et al., "Epidural spinal cord stimulation facilitates immediate restoration of dormant motor and autonomic supraspinal pathways after chronic neurologically complete spinal cord injury," *Journal of Neurotrauma*, vol. 36, no. 15, pp. 2325–2336, 2019.

[66] M. Possover, "Ten-year experience with continuous low-frequency pelvic somatic nerves stimulation for recovery of voluntary walking in people with chronic spinal cord injury: a prospective case series of 29 consecutive patients," *Archives of Physical Medicine and Rehabilitation*, vol. 102, no. 1, pp. 50–57, 2021.

[67] E. Rejc, C. A. Angeli, N. Bryant, and S. J. Harkema, "Effects of stand and step training with epidural stimulation on motor function for standing in chronic complete paraplegics," *Journal of Neurotrauma*, vol. 34, no. 9, pp. 1787–1802, 2017.

[68] S. Mangold, T. Keller, A. Curt, and V. Dietz, "Transcutaneous functional electrical stimulation for grasping in subjects with cervical spinal cord injury," *Spinal Cord*, vol. 43, no. 1, pp. 1–13, 2005.

[69] F. Inanici, S. Samejima, P. Gad, V. R. Edgerton, C. P. Hofstetter, and C. T. Moritz, "Transcutaneous electrical spinal stimulation promotes long-term recovery of upper extremity function in chronic tetraplegia," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 26, no. 6, pp. 1272–1278, 2018.

[70] K. Minassian, I. Persy, F. Rattay, M. R. Dimitrijevic, C. Hofer, and H. Kern, "Posterior root-muscle reflexes elicited by transcutaneous stimulation of the human lumbosacral cord," *Muscle & Nerve*, vol. 35, no. 3, pp. 327–336, 2007.

[71] P. K. Bedi and N. Arumugam, "Tapping the neural circuitry: surface spinal stimulation in spinal cord injury: a case report," *Journal of Exercise Science & Physiotherapy*, vol. 12, no. 1, pp. 69–75, 2016.

[72] D. G. Sayenko, M. Rath, A. R. Ferguson et al., "Self-Assisted standing enabled by non-invasive spinal stimulation after spinal cord injury," *Journal of Neurotrauma*, vol. 36, no. 9, pp. 1435–1450, 2018.

[73] P. Gad, Y. Gerasimenko, S. Zdunowski et al., "Weight bearing over-ground stepping in an exoskeleton with non-invasive spinal cord neuromodulation after motor complete paraplegia," *Frontiers in Neuroscience*, vol. 11, no. 333, p. 333, 2017.

[74] Y. P. Gerasimenko, D. C. Lu, M. Modaber et al., "Noninvasive reactivation of motor descending control after paralysis," *Journal of Neurotrauma*, vol. 32, no. 24, pp. 1968–1980, 2015.

[75] K. Minassian, U. S. Hofstoetter, S. M. Danner et al., "Spinal rhythm generation by step-induced feedback and transcutaneous posterior root stimulation in complete spinal cord-injured individuals," *Neurorehabilitation and Neural Repair*, vol. 30, no. 3, pp. 233–243, 2015.

[76] F. Inanici, L. N. Brighton, S. Samejima, C. P. Hofstetter, and C. T. Moritz, "Transcutaneous spinal cord stimulation restores hand and arm function after spinal cord injury," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 29, pp. 310–319, 2021.

[77] S. I. Bîrlea, P. P. Breen, G. J. Corley, N. M. Bîrlea, F. Quondamatteo, and G. ÓLaighin, "Changes in the electrical properties of the electrode-skin-underlying tissue composite during a week-long programme of neuromuscular electrical stimulation," *Physiological Measurement*, vol. 35, no. 2, pp. 231–252, 2014.

[78] J. L. Lujan and P. E. Crago, "Automated optimal coordination of multiple-DOF neuromuscular actions in

feedforward neuroprostheses," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 1, pp. 179–187, 2009.

[79] D. M. Niddam, T. Graven-Nielsen, L. Arendt-Nielsen, and A. C. N. Chen, "Non-painful and painful surface and intramuscular electrical stimulation at the thenar and hypothenar sites: differential cerebral dynamics of early to late latency SEPs," *Brain Topography*, vol. 13, no. 4, pp. 283–292, 2001.

[80] A. Megía García, D. Serrano-Muñoz, J. Taylor, J. Avendaño-Coy, and J. Gómez-Soriano, "Transcutaneous spinal cord stimulation and motor rehabilitation in spinal cord injury: a systematic review," *Neurorehabilitation and Neural Repair*, vol. 34, no. 1, pp. 3–12, 2019.

[81] C. Tao, X. Shen, L. Ma et al., "Comparative study of intraspinal microstimulation and epidural spinal cord stimulation," in *Proceedings of the 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pp. 3795–3798, Berlin, Germany, 2019.

[82] B. J. Holinski, K. A. Mazurek, D. G. Everaert et al., "Intraspinal microstimulation produces over-ground walking in anesthetized cats," *Journal of Neural Engineering*, vol. 13, no. 5, Article ID 056016, 2016.

[83] A. N. Dalrymple, D. G. Everaert, D. S. Hu, and V. K. Mushahwar, "A speed-adaptive intraspinal microstimulation controller to restore weight-bearing stepping in a spinal cord hemisection model," *Journal of Neural Engineering*, vol. 15, no. 5, p. 20, 2018.

[84] M. R. Kasten, M. D. Sunshine, E. S. Secrist, P. J. Horner, and C. T. Moritz, "Therapeutic intraspinal microstimulation improves forelimb function after cervical contusion injury," *Journal of Neural Engineering*, vol. 10, no. 4, 2013.

[85] J. G. Mcpherson, R. R. Miller, and S. I. Perlmutter, "Targeted, activity-dependent spinal stimulation produces long-lasting motor recovery in chronic cervical spinal cord injury," *Proceedings of the National Academy of Sciences*, vol. 112, no. 39, pp. 12193–12198, 2015.

[86] S. E. Mondello, M. R. Kasten, P. J. Horner, and C. T. Moritz, "Therapeutic intraspinal stimulation to generate activity and promote long-term recovery," *Frontiers in Neuroscience*, vol. 8, 2014.

[87] C. H. Tator, K. Minassian, and V. K. Mushahwar, "Spinal cord stimulation," *Handbook of Clinical Neurology*, vol. 109, pp. 283–296, 2012.

[88] S. J. Harkema, "Plasticity of interneuronal networks of the functionally isolated human spinal cord," *Brain Research Reviews*, vol. 57, no. 1, pp. 255–264, 2008.

[89] D. N. Heo, H.-J. Kim, Y. J. Lee et al., "Flexible and highly biocompatible nanofiber-based electrodes for neural surface interfacing," *ACS Nano*, vol. 11, no. 3, pp. 2961–2971, 2017.

[90] E. Y. Shapkova, E. V. Pismennaya, D. V. Emelyannikov, and Y. Ivanenko, "Exoskeleton walk training in paralyzed individuals benefits from transcutaneous lumbar cord tonic," *Electrical Stimulation*, vol. 14, no. 416, 2020.

[91] W. Guan, L. Zhou, and Y. Cao, "Joint motion control for lower limb rehabilitation based on iterative learning control (ILC) algorithm," *Complexity*, vol. 2021, Article ID 6651495, 9 pages, 2021.

[92] R. Van Den Brand, J. Heutschi, Q. Barraud et al., "Restoring voluntary control of locomotion after paralyzing spinal cord injury," *Science*, vol. 336, no. 6085, pp. 1182–1185, 2012.

[93] W. Cao, H. Yu, X. Wu, S. Li, Q. Meng, and C. Chen, "Development and evaluation of a rehabilitation wheelchair with multiposture transformation and smart control," *Complexity*, vol. 2021, Article ID 6628802, 14 pages, 2021.

[94] M. A. Urbin, M. Liu, E. C. Bottorff, R. A. Gaunt, L. E. Fisher, and D. J. Weber, "Hindlimb motor responses evoked by microstimulation of the lumbar dorsal root ganglia during quiet standing," *Journal of Neural Engineering*, vol. 17, no. 1, Article ID 016019, 2019.

[95] S. Shahdoost, S. B. Frost, D. J. Guggenmos et al., "A brain-spinal interface (BSI) system-on-chip (SoC) for closed-loop cortically-controlled intraspinal microstimulation," *Analog Integrated Circuits and Signal Processing*, vol. 95, no. 1, pp. 1–16, 2018.

[96] X.-y. Shen, W. Du, W. Huang, and Y. Chen, "Rebuilding motor function of the spinal cord based on functional electrical stimulation," *Neural Regeneration Research*, vol. 11, no. 8, pp. 1327–1332, 2016.

[97] C. Vidaurre, C. Klauer, T. Schauer, A. Ramos-Murguialday, and K.-R. Müller, "EEG-based BCI for the linear control of an upper-limb neuroprosthesis," *Medical Engineering & Physics*, vol. 38, no. 11, pp. 1195–1204, 2016.

[98] A. Biasiucci, R. Leeb, I. Iturrate et al., "Brain-actuated functional electrical stimulation elicits lasting arm motor recovery after stroke," *Nature Communications*, vol. 9, no. 1, p. 2421, 2018.

[99] M. Uma and S. Prabhu, "A feasibility study of BCI based FES model for differently abled people," *IOP Conference Series: Materials Science and Engineering*, vol. 402, Article ID 012009, 2018.

[100] A. Prasad, Q.-S. Xue, V. Sankar et al., "Comprehensive characterization and failure modes of tungsten microwire arrays in chronic neural implants," *Journal of Neural Engineering*, vol. 9, no. 5, Article ID 056015, 2012.

[101] C. Bennett, M. Samikkannu, F. Mohammed, W. D. Dietrich, S. M. Rajguru, and A. Prasad, "Blood brain barrier (BBB)-disruption in intracortical silicon microelectrode implants," *Biomaterials*, vol. 164, pp. 1–10, 2018.

[102] A. Thomschewski, A. Strohlein, and P. B. Langthaler, "Imagine there is no plegia. Mental motor imagery difficulties in patients with traumatic spinal cord injury," *Frontiers in Neuroscience*, vol. 11, p. 9, 2017.

[103] C. E. Bouton, A. Shaikhouni, N. V. Annetta et al., "Restoring cortical control of functional movement in a human with quadriplegia," *Nature*, vol. 533, no. 7602, pp. 247–250, 2016.

[104] C. Reynolds, B. A. Osuagwu, and A. Vuckovic, "Influence of motor imagination on cortical activation during functional electrical stimulation," *Clinical Neurophysiology*, vol. 126, no. 7, pp. 1360–1369, 2015.

[105] U. Chaudhary, N. Birbaumer, and A. Ramos-Murguialday, "Brain-computer interfaces for communication and rehabilitation," *Nature Reviews Neurology*, vol. 12, no. 9, pp. 513–525, 2016.

[106] S. Luo, H. Xu, Y. Zuo, X. Liu, and A. H. All, "A review of functional electrical stimulation treatment in spinal cord injury," *NeuroMolecular Medicine*, vol. 22, no. 4, pp. 447–463, 2020.

[107] M. Qian, G. Zhong, X. Yan, H. Wang, and Y. Cui, "A closed-loop brain stimulation control system design based on brain-machine interface for epilepsy," *Complexity*, vol. 2020, Article ID 3136715, 15 pages, 2020.

[108] G. Taccola, B. H. Salazar, R. Apicella, M. K. Hogan, P. J. Horner, and D. Sayenko, "Selective antagonism of A1 adenosinergic receptors strengthens the neuromodulation of the sensorimotor network during epidural spinal stimulation," *Frontiers in Systems Neuroscience*, vol. 14, p. 44, 2020.

[109] N. D. James, J. Shea, E. M. Muir, J. Verhaagen, B. L. Schneider, and E. J. Bradbury, "Chondroitinase gene

therapy improves upper limb function following cervical contusion injury," *Experimental Neurology*, vol. 271, pp. 131–135, 2015.

[110] J. Prager, D. Ito, D. R. Carwardine et al., "Delivery of chondroitinase by canine mucosal olfactory ensheathing cells alongside rehabilitation enhances recovery after spinal cord injury," *Experimental Neurology*, vol. 340, Article ID 113660, 2021.

[111] R. M. Grumbles, Y. Liu, C. M. Thomas, P. M. Wood, and C. K. Thomas, "Acute stimulation of transplanted neurons improves motoneuron survival, axon growth, and muscle reinnervation," *Journal of Neurotrauma*, vol. 30, no. 12, pp. 1062–1069, 2013.

[112] S. M. Schaal, B. M. Kitay, K. S. Cho et al., "Schwann cell transplantation improves reticulospinal axon growth and forelimb strength after severe cervical spinal cord contusion," *Cell Transplantation*, vol. 16, no. 3, pp. 207–228, 2007.

[113] A. D. Koutsou, J. C. Moreno, A. J. Del Ama, E. Rocon, and J. L. Pons, "Advances in selective activation of muscles for non-invasive motor neuroprostheses," *Journal of NeuroEngineering and Rehabilitation*, vol. 13, no. 1, p. 56, 2016.

[114] M. Alam, S. Li, R. U. Ahmed et al., "Development of a battery-free ultrasonically powered functional electrical stimulator for movement restoration after paralyzing spinal cord injury," *Journal of NeuroEngineering and Rehabilitation*, vol. 16, no. 1, p. 36, 2019.

[115] S. Mazzoleni, E. Battini, A. Rustici, and G. Stampacchia, "An integrated gait rehabilitation training based on functional electrical stimulation cycling and overground robotic exoskeleton in complete spinal cord injury patients: preliminary results," in *Proceedings of the International Conference on Rehabilitation Robotics (ICORR)*, pp. 289–293, London, UK, 2017.

[116] W. Kohei and A. Hiroshi, "Validity of surface electromyography for vastus intermedius muscle assessed by needle electromyography," *Journal of Neuroscience Methods*, vol. 198, no. 2, pp. 332–335, 2011.

[117] E. N. Kamavuako, E. J. Scheme, and K. B. Englehart, "Combined surface and intramuscular EMG for improved real-time myoelectric control performance," *Biomedical Signal Processing and Control*, vol. 10, pp. 102–107, 2014.

[118] L. Zhang, Z. Li, Y. Hu, C. Smith, E. M. G. Farewik, and R. Wang, "Ankle joint torque estimation using an EMG-driven neuromusculoskeletal model and an artificial neural network model," *IEEE Transactions on Automation Science and Engineering*, vol. 18, no. 2, pp. 564–573, 2021.

[119] M. Gawel, E. Zalewska, E. Szmidt-Salkowska et al., "Motor Unit Number Index (MUNIX) as a biomarker of motor unit loss in post-polio syndrome versus needle EMG," *Journal of Electromyography and Kinesiology*, vol. 46, pp. 35–40, 2019.

[120] G. Marco, B. Alberto, and V. Taian, "Surface EMG and muscle fatigue: multi-channel approaches to the study of myoelectric manifestations of muscle fatigue," *Physiological Measurement*, vol. 38, no. 5, pp. R27–R60, 2017.

[121] O. W. Samuel, H. Zhou, X. Li et al., "Pattern recognition of electromyography signals based on novel time domain features for amputees' limb motion classification," *Computers & Electrical Engineering*, vol. 67, pp. 646–655, 2018.

[122] M. Bhadane, J. Liu, W. Z. Rymer, P. Zhou, and S. Li, "Re-evaluation of EMG-torque relation in chronic stroke using linear electrode array EMG recordings," *Scientific Reports*, vol. 6, no. 1, Article ID 28957, 2016.

[123] Y. Liu, L. Tian, M. S. Raj et al., "Intraoperative monitoring of neuromuscular function with soft, skin-mounted wireless devices," *Npj Digital Medicine*, vol. 1, no. 1, p. 19, 2018.

[124] N. A. Chaya, B. R. Bhavana, S. B. Anoogna, M. Hiranmai, and N. Krupa B, "Real-time replication of arm movements using surface EMG signals," *Procedia Computer Science*, vol. 154, pp. 186–193, 2019.

[125] A. Botter, M. Beltrandi, G. L. Cerone, M. Gazzoni, and T. M. M. Vieira, "Development and testing of acoustically-matched hydrogel-based electrodes for simultaneous EMG-ultrasound detection," *Medical Engineering & Physics*, vol. 64, pp. 74–79, 2019.

[126] F. O. Fadeev, F. V. Bashirov, V. A. Markosyan et al., "Combination of epidural electrical stimulation with ex vivo triple gene therapy for spinal cord injury: a proof of principle study," *Neural regeneration research*, vol. 16, no. 3, pp. 550–560, 2021.

[127] B. Yang, F. Zhang, F. Cheng et al., "Strategies and prospects of effective neural circuits reconstruction after spinal cord injury," *Cell Death & Disease*, vol. 11, no. 6, pp. 439–514, 2020.

Hindawi
Complexity
Volume 2021, Article ID 6245168, 11 pages
https://doi.org/10.1155/2021/6245168

*Research Article*

# Pursuer Navigation Based on Proportional Navigation and Optimal Information Fusion

**Shulin Feng**,[1] **Zhanxin Li**,[2] **Li Liu**,[1] **Hongyong Yang**,[1] **Yuanhua Yang**,[3] **Gaohuan Lv**,[4] **and Liang Yang**[1]

[1]*School of Information and Electrical Engineering, Ludong University, Yantai 264025, China*
[2]*Tianjin Key Laboratory of Microgravity and Hypogravity Environment Simulation Technology, Tianjin 300000, China*
[3]*School of Mathematics and Statistics Science, Ludong University, Yantai 264025, China*
[4]*Ulsan Ship and Ocean College, Ludong University, Yantai 264025, China*

Correspondence should be addressed to Shulin Feng; fengshulin@ldu.edu.cn

Received 19 May 2021; Accepted 13 September 2021; Published 21 October 2021

Academic Editor: Wen Qi



Pursuer navigation is proposed based on the three-dimensional proportional navigation law, and this method presents a family of navigation laws resulting in a rich behavior for different parameters. Firstly, the kinematics model for the pursuer and the target is established. Secondly, the proportional navigation law is deduced through the kinematics model. Based on point-to-point navigation, obstacle avoidance is implemented by adjusting the control parameters, and the combination can enrich the application range of obstacle avoidance and guidance laws. Thirdly, information fusion weighted by diagonal matrices is used for decreasing the tracking precision. Finally, simulations are conducted in the MATLAB environment. Simulation results verify the availability of the proposed navigation law.

## 1. Introduction

tVarious navigation and obstacle avoidance methods are the important issues. The proportional navigation is a method well known and widely applied in the aerospace community. In [1], the augmented IPN is deduced for interception. In [2], the authors present a new homing guidance law using well-known BPN to perform an impact time constraint and impact angle constraint. Over and above the case of infinite maneuverability of the missile, the full condition that captures a nonmaneuvering target is deduced in [3]. Real-time navigation is given in [4] by integrating the backstepping method and neurodynamics model. In [5], the proportional navigation applied to missile guidance problems is tailored. In [6], the authors propose collision avoidance strategy for multiagent. A receding horizon control method for convergent navigation of the robot is given in [7], and this method includes a scientific procedure for the generation of potential control sequences. In [8], a modified cooperative proportional navigation is presented to avoid singularity, and the time-to-go control efficiency under the small leading angle is improved in this paper. The capturability of 3D PPN against the lower speed freely-maneuvering target for the homing phase is restudied in [9], extending the NOR method of the 2D PPN to 3D space. In the study of [10], pure proportional navigation (PPN) and a look angle-constrained guidance law consisting of PPN and look angle control are designed. In [11], a novel augmented proportional navigation (APN) is proposed for midrange autonomous rendezvous, and the midrange autonomous rendezvous can be absolutely implemented. The application of proportional navigation to the pursuer requires improvements. The presentation of this paper is different from the classical presentation. The proportional navigation is proposed by using the flight path and heading angles of the pursuer. This presentation is more proper for the pursuer than the classical presentation, where proportional navigation is proposed in accordance with the lateral and vertical acceleration. Then, the presentation of proportional navigation can be easily adapted to the collision avoidance mode since the

proportional navigation is written as a function of the flight path and heading angles of the pursuer. It allows a rapid change in the path of the pursuer under proportional navigation.

In the study of [12], obstacle avoidance and navigation are addressed by using the model-based control method. This method can be used for both online and offline. In [12], the proportional navigation is written based on the robotic steering angle. Moreover, the collision avoidance mode is implemented by using proportional navigation. Notwithstanding the method of [12] seems to be quite competent, and it sustains the following problems.

(1) The control law for the orientation angle can continue to be expanded

(2) The proportional navigation was proposed by assuming no sensor noise

(3) Information fusion is not combined with the proportional navigation to improve the tracking process

In this paper, the work is mainly motivated by the study in [12]. The aim of this paper is to consider the solution of the pursuer tracking toward a target in the 3D space. Based on the geometric relationship of pursuer-target, this paper presents the polar kinematics models. In this paper, the proportional navigation is given in terms of the flight path and heading angles for the pursuer. Moreover, the proportional navigation can also be adapted to the collision avoidance mode. The method can be used for indoor and outdoor navigation as well, especially to reach goals that are at a long distance from the pursuer, and as a result, they are out of the range of view of the sensors (such as the camera), but their position is known to the pursuer.

The proportional navigation of Belkhouche and Belkhouche [12] was proposed by assuming no sensor noise. For sensor noise, the filter method can be used to improve the tracking process. As studied in [13], a data-driven method combining the EKF and RBF neural network is given to estimate the internal temperature for the lithium-ion battery. In [14], the particle filter is applied to predict the aging trajectory of the lithium-ion battery. Even though the algorithm of Belkhouche and Belkhouche [12] seems to be quite efficient, it suffers from that Kalman filter techniques are not used for dynamic state estimation. Various multisensor fusion methods have been studied to solve this problem. Under the optimal fusion criterion of Sun and Deng [15], the multisensor fusion decentralized Kalman filter is obtained. In [16], the authors propose the two-sensor information fusion steady-state Kalman filter. In [17, 18], distributed optimal information fusion filter theory is presented under the classical Kalman filter. The device are argued in [19] in accordance with sensor data fusion methods, sensor design, and prototype setup. Based on multisensor fusion, a hybrid indoor localization system is given in [20]. In [21], the authors present the information fusion Kalman filter weighted by scalars. In [22], the functional equivalence of two optimal measurement fusion methods is proved under the steady-state Kalman filter. Information fusion weighted by diagonal matrices is proposed in [16–18]. As studied in [12], this method is not considered the negative influence of sensor noise. Based on the above control theory of information fusion, the control strategy of [12] can be further improved. In this paper, the proportional navigation combined with information fusion weighted by diagonal matrices are used to implement more reasonable tracking performance.

Control objective of this paper is to implement pursuer navigation and obstacle avoidance using a easy and valid model-based control law. It can be applied to both online and offline navigation and obstacle avoidance. This method consists of a family of methods for pursuer navigation under proportional navigation, where this paper applies the pursuer kinematics equations combined with the geometric rule. The challenge of this paper is how to design the control law for proportional navigation and implement obstacle avoidance. To deal with the challenge, this paper presents the polar kinematics models of pursuer-target. The control law of proportional navigation is given in terms of the flight path and heading angles for the pursuer. Under sensor noise, two-sensor information fusion is applied to improve the control law. Moreover, the proportional navigation can implement collision avoidance by using point-to-point navigation.

The contribution of this paper is mainly to present three-dimensional proportional navigation to implement tracking the target, outperforming the pure proportional navigation (PPN) in terms of interception time. Under sensor noise, two-sensor information fusion together with proportional navigation can enhance the tracking precision. Moreover, obstacle avoidance is implemented by using point-to-point navigation combined with proportional navigation.

The remainder of this paper is organized as follows. The dynamic model of pursuer-target is derived. Then, the proportional navigation law is discussed. This paper designs the control method of obstacle avoidance. Two-sensor information fusion is used in this paper. Simulation results are given to formulate the availability of the obtained results, and then, some conclusions are drawn.

## 2. Dynamic Model

The geometry of the navigation is illustrated in Figure 1. In the 3D coordinate system, the linear velocity of the pursuer is $v_P$. The flight path and heading angles are $\theta_P$ and $\phi_P$, respectively. LOS for target-pursuer is TP. $\sigma_{\mathrm{TP}}$ is the pitch angle of TP, and $\gamma_{\mathrm{TP}}$ is the yaw angle of TP. $r_{\mathrm{TP}}$ is the relative distance pursuer-target.

Based on [23, 24], the differential equations for $r_{\mathrm{TP}}$, $\sigma_{\mathrm{TP}}$, and $\gamma_{\mathrm{TP}}$ are

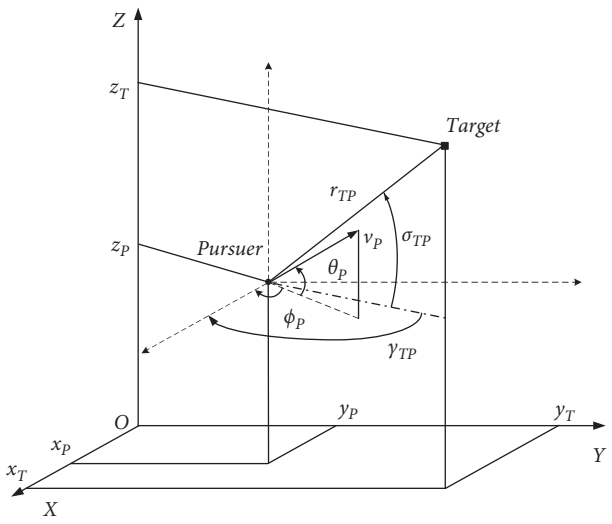



Figure 1: Geometric relationship of pursuer-target.

$$\begin{cases} \dot{r}_{\mathrm{TP}} = -v_P \cos(\sigma_{\mathrm{TP}} - \theta_P)\cos(\gamma_{\mathrm{TP}} - \phi_P) + v_T \cos(\sigma_{\mathrm{TP}} - \theta_T)\cos(\gamma_{\mathrm{TP}} - \phi_T), \\ r_{\mathrm{TP}}\dot{\sigma}_{\mathrm{TP}} = v_P \sin(\sigma_{\mathrm{TP}} - \theta_P)\cos(\gamma_{\mathrm{TP}} - \phi_P) - v_T \sin(\sigma_{\mathrm{TP}} - \theta_T)\cos(\gamma_{\mathrm{TP}} - \phi_T), \\ r_{\mathrm{TP}} \cos\ \sigma_{\mathrm{TP}}\dot{\gamma}_{\mathrm{TP}} = v_P \cos\ \theta_P \sin(\gamma_{\mathrm{TP}} - \phi_P) - v_T \cos\ \theta_T \sin(\gamma_{\mathrm{TP}} - \phi_T). \end{cases} \tag{1}$$

Since the target is motionless, one can have $v_T = 0$, and thus,

$$\begin{cases} \dot{r}_{\mathrm{TP}} = -v_P \cos(\sigma_{\mathrm{TP}} - \theta_P)\cos(\gamma_{\mathrm{TP}} - \phi_P), \\ r_{\mathrm{TP}}\dot{\sigma}_{\mathrm{TP}} = v_P \sin(\sigma_{\mathrm{TP}} - \theta_P)\cos(\gamma_{\mathrm{TP}} - \phi_P), \\ r_{\mathrm{TP}} \cos\ \sigma_{\mathrm{TP}}\dot{\gamma}_{\mathrm{TP}} = v_P \cos\ \theta_P \sin(\gamma_{\mathrm{TP}} - \phi_P). \end{cases} \tag{2}$$

The robustness of the method is a critical issue. It is should be noted that this method belongs to a family of methods in terms of the kinematics equation and geometric rule. These methods are famous for the robustness.

## 3. Three-Dimensional Proportional Navigation Law

This paper designs the proportional navigation law in conformity to the pursuer flight path and heading angles as follows:

$$\begin{cases} \theta_P(t) = G\sigma_{\mathrm{TP}}(t) + \delta, \\ \phi_P(t) = E\gamma_{\mathrm{TP}}(t) + \mu, \end{cases} \tag{3}$$

where $G$ and $E$ are the navigation constant with $(G \geq 1;\ E \geq 1)$ and $\delta$ and $\mu$ are the deviation angles.

Combining equation (2) with equation (3), the differential equations for $r_{\mathrm{TP}}$, $\sigma_{\mathrm{TP}}$, and $\gamma_{\mathrm{TP}}$ are

$$\begin{cases} \dot{r}_{\mathrm{TP}} = -v_P \cos[(G-1)\sigma_{\mathrm{TP}} + \delta]\cos((E-1)\gamma_{\mathrm{TP}} + \mu), \\ r_{\mathrm{TP}}\dot{\sigma}_{\mathrm{TP}} = -v_P \sin((G-1)\sigma_{\mathrm{TP}} + \delta)\cos[(E-1)\gamma_{\mathrm{TP}} + \mu], \\ r_{\mathrm{TP}} \cos\ \sigma_{\mathrm{TP}}\dot{\gamma}_{\mathrm{TP}} = -v_P \cos\ \theta_P \sin[(E-1)\gamma_{\mathrm{TP}} + \mu]. \end{cases} \tag{4}$$

Results in relation to the pursuer that tracks an immovable point are given as follows.

*Theory 1.* By using pure pursuit with $(G = E = 1;\ \delta = \mu = 0)$, the pursuer can reach the target from any original condition.

*Proof.* Combining $(G = E = 1,\ \delta = \mu = 0)$ with equation (4), it can be written as

$$\dot{r}_{\mathrm{TP}} = -v_P. \tag{5}$$

Since $\dot{r}_{\mathrm{TP}} < 0$, $r_{\mathrm{TP}}$ is decreasing and the pursuer can reach the target, with the final flight path angle $\theta_P(t_f) = \sigma_{\mathrm{TP}}(t_0)$ and heading angle $\phi_P(t_f) = \gamma_{\mathrm{TP}}(t_0)$.

This completes the proof. □

*Theory 2.* By using deviated pursuit with $(G = E = 1, \delta \neq 0, \text{and } \mu \neq 0)$, the pursuer can reach the target when

$$\delta, \mu \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \text{ or } \delta, \mu \in \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]. \tag{6}$$

*Proof.* On the basis of the first equation in the relative kinematics model, one can obtain

$$\dot{r}_{\mathrm{TP}} = -v_P \cos\,\delta\,\cos\,\mu. \tag{7}$$

$\dot{r}_{\mathrm{TP}}$ is a decreasing function when $\delta$ and $\mu \in [-\pi/2, \pi/2]$ or $\delta$ and $\mu \in [\pi/2, 3\pi/2]$.

This completes the proof. □

*Theory 3.* For $G > 1$ and $E > 1$, the pursuer navigating under equation (3) reaches the target for nearly all original states.

*Proof.* Combining $(G > 1; E > 1)$ with equation (4), it obtains

$$\begin{cases} \dot{\sigma}_{\mathrm{TP}} = -\dfrac{v_P}{r_{\mathrm{TP}}}\sin\left[(G-1)\sigma_{\mathrm{TP}} + \delta\right]\cos\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right] = f_{\mathrm{TP}}\left(\sigma_{\mathrm{TP}}, \gamma_{\mathrm{TP}}\right), \\ \dot{\gamma}_{\mathrm{TP}} = -\dfrac{v_P}{r_{\mathrm{TP}}}\dfrac{\cos\left(G\sigma_{\mathrm{TP}} + \delta\right)}{\cos\,\sigma_{\mathrm{TP}}}\sin\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right] = g_{\mathrm{TP}}\left(\sigma_{\mathrm{TP}}, \gamma_{\mathrm{TP}}\right). \end{cases} \tag{8}$$

This system has four equilibrium solutions, namely, $(\sigma^*_{\mathrm{TP}_1} = 2n\pi - \delta/G - 1, \gamma^*_{\mathrm{TP}_1} = 2n\pi - \mu/E - 1)$, $(\sigma^*_{\mathrm{TP}_2} = 2n\pi - \delta/G - 1, \gamma^*_{\mathrm{TP}_2} = 2n\pi + \pi - \mu/E - 1)$, $(\sigma^*_{\mathrm{TP}_3} = 2n\pi + \pi - \delta/G - 1, \gamma^*_{\mathrm{TP}_3} = 2n\pi - \mu/E - 1)$, and $(\sigma^*_{\mathrm{TP}_4} = 2n\pi + \pi - \delta/G - 1, \gamma^*_{\mathrm{TP}_4} = 2n\pi + \pi - \mu/E - 1)$.

After partial deviation, it can be obtained that

$$\begin{aligned} \frac{\partial f_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}} &= -\frac{v_P}{r_{\mathrm{TP}}}(G-1)\cos\left[(G-1)\sigma_{\mathrm{TP}} + \delta\right]\cos\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right], \\ \frac{\partial f_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}} &= \frac{v_P}{r_{\mathrm{TP}}}(E-1)\sin\left[(G-1)\sigma_{\mathrm{TP}} + \delta\right]\sin\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right], \\ \frac{\partial g_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}} &= \frac{v_P}{r_{\mathrm{TP}}}\sin\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right]\frac{(G-1)\sin\left(G\sigma_{\mathrm{TP}} + \delta\right)\cos\,\sigma_{\mathrm{TP}} + \sin\left[(G-1)\sigma_{\mathrm{TP}} + \delta\right]}{\cos^2\sigma_{\mathrm{TP}}}, \\ \frac{\partial g_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}} &= -\frac{v_P}{r_{\mathrm{TP}}}(E-1)\cos\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right]\frac{\cos\left(G\sigma_{\mathrm{TP}} + \delta\right)}{\cos\,\sigma_{\mathrm{TP}}} \\ &= -\frac{(E-1)v_P}{r_{\mathrm{TP}}}\cos\left[(E-1)\gamma_{\mathrm{TP}} + \mu\right]\frac{\cos\left[(G+1)\sigma_{\mathrm{TP}} + \delta\right] + \cos\left[(G-1)\sigma_{\mathrm{TP}} + \delta\right]}{2\,\cos^2\sigma_{\mathrm{TP}}}. \end{aligned} \tag{9}$$

By linearizing near each equilibrium solution, one has

$$
\begin{aligned}
T_1 &= \begin{bmatrix} \dfrac{\partial f_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_1},\gamma^*_{\mathrm{TP}_1}\right)} & \dfrac{\partial f_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_1},\gamma^*_{\mathrm{TP}_1}\right)} \\ \dfrac{\partial g_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_1},\gamma^*_{\mathrm{TP}_1}\right)} & \dfrac{\partial g_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_1},\gamma^*_{\mathrm{TP}_1}\right)} \end{bmatrix} \\
&= \begin{bmatrix} -\dfrac{(G-1)v_P}{r_{\mathrm{TP}}} & 0 \\ 0 & -\dfrac{(E-1)v_P}{2r_{\mathrm{TP}}}\dfrac{\cos\left[(G+1)\sigma_{\mathrm{TP}}+\delta\right]+1}{\cos^2\sigma_{\mathrm{TP}}} \end{bmatrix} = \begin{pmatrix} \lambda_{11} & 0 \\ 0 & \lambda_{12} \end{pmatrix}, \\
T_2 &= \begin{bmatrix} \dfrac{\partial f_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_2},\gamma^*_{\mathrm{TP}_2}\right)} & \dfrac{\partial f_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_2},\gamma^*_{\mathrm{TP}_2}\right)} \\ \dfrac{\partial g_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_2},\gamma^*_{\mathrm{TP}_2}\right)} & \dfrac{\partial g_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_2},\gamma^*_{\mathrm{TP}_2}\right)} \end{bmatrix} \\
&= \begin{pmatrix} \dfrac{(G-1)v_P}{r_{\mathrm{TP}}} & 0 \\ 0 & -\dfrac{(E-1)v_P}{2r_{\mathrm{TP}}}\dfrac{\cos\left[(G+1)\sigma_{\mathrm{TP}}+\delta\right]-1}{\cos^2\sigma_{\mathrm{TP}}} \end{pmatrix} = \begin{pmatrix} \lambda_{21} & 0 \\ 0 & \lambda_{22} \end{pmatrix}, \\
T_3 &= \begin{bmatrix} \dfrac{\partial f_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_3},\gamma^*_{\mathrm{TP}_3}\right)} & \dfrac{\partial f_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_3},\gamma^*_{\mathrm{TP}_3}\right)} \\ \dfrac{\partial g_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_3},\gamma^*_{\mathrm{TP}_3}\right)} & \dfrac{\partial g_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_3},\gamma^*_{\mathrm{TP}_3}\right)} \end{bmatrix} \\
&= \begin{pmatrix} \dfrac{(G-1)v_P}{r_{\mathrm{TP}}} & 0 \\ 0 & a-\dfrac{(E-1)v_P}{2r_{\mathrm{TP}}}\dfrac{\cos\left[(G+1)\sigma_{\mathrm{TP}}+\delta\right]+1}{\cos^2\sigma_{\mathrm{TP}}} \end{pmatrix} = \begin{pmatrix} \lambda_{31} & 0 \\ 0 & \lambda_{32} \end{pmatrix}, \\
T_4 &= \begin{bmatrix} \dfrac{\partial f_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_4},\gamma^*_{\mathrm{TP}_4}\right)} & \dfrac{\partial f_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_4},\gamma^*_{\mathrm{TP}_4}\right)} \\ \dfrac{\partial g_{\mathrm{TP}}}{\partial \sigma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_4},\gamma^*_{\mathrm{TP}_4}\right)} & \dfrac{\partial g_{\mathrm{TP}}}{\partial \gamma_{\mathrm{TP}}}\Big|_{\left(\sigma^*_{\mathrm{TP}_4},\gamma^*_{\mathrm{TP}_4}\right)} \end{bmatrix} \\
&= \begin{pmatrix} -\dfrac{(G-1)v_P}{r_{\mathrm{TP}}} & 0 \\ 0 & -\dfrac{(E-1)v_P}{2r_{\mathrm{TP}}}\dfrac{\cos\left[(G+1)\sigma_{\mathrm{TP}}+\delta\right]-1}{\cos^2\sigma_{\mathrm{TP}}} \end{pmatrix} = \begin{pmatrix} \lambda_{41} & 0 \\ 0 & \lambda_{42} \end{pmatrix}.
\end{aligned} \tag{10}
$$

The characteristic roots $\lambda_{11}$ and $\lambda_{12}$ of $T_1$ are with ($\lambda_{11}<0$ and $\lambda_{12}<0$), the characteristic roots $\lambda_{21}$ and $\lambda_{22}$ of $T_2$ are with ($\lambda_{21}>0$ and $\lambda_{22}>0$), the characteristic roots $\lambda_{31}$ and $\lambda_{32}$ of $T_3$ are with ($\lambda_{31}>0$ and $\lambda_{32}<0$), the characteristic roots $\lambda_{41}$ and $\lambda_{42}$ of $T_4$ are with ($\lambda_{41}<0$ and $\lambda_{42}>0$). According to Hartman and Grobman theorem of [25, 26], only $(\sigma^*_{\mathrm{TP}_1},\gamma^*_{\mathrm{TP}_1})$ is asymptotically stable.

Since the solutions for $\sigma_{\mathrm{TP}}$ and $\gamma_{\mathrm{TP}}$ approach their asymptotically stable equilibrium positions, one can get $\sigma_{\mathrm{TP}} \longrightarrow \sigma^*_{\mathrm{TP}_1} = 2n\pi - \delta/G - 1$ and $\gamma_{\mathrm{TP}} \longrightarrow \gamma^*_{\mathrm{TP}_1} = 2n\pi - \mu/E - 1$ with time. Since $\sigma_{\mathrm{TP}} \longrightarrow 2n\pi - \delta/G - 1$ and $\gamma_{\mathrm{TP}} \longrightarrow 2n\pi - \mu/E - 1$ with time, $\cos[(G-1)\sigma_{\mathrm{TP}} + \delta]\cos[(E-1)\gamma_{\mathrm{TP}} + \mu]$ is positive in $[t_1, t_f]$, $t_1 \geq t_0$; then, one can get $\dot{r}_{\mathrm{TP}} < 0$ after $t_1$.

This completes the proof.

At $t_0$, the proportional navigation law is

$$\left\{ \theta_P(t_0) = G\sigma_{\mathrm{TP}}(t_0) + \delta\phi_P(t_0) = E\gamma_{\mathrm{TP}}(t_0) + \mu. \right. \tag{11}$$

Then, two cases are

(i) Select $(G, E)$ and $(\delta, \mu)$ according to equation (11) on ($G$ and $E \geq 1$).

(ii) Put into use heading regulation which deduces $\theta_P$ and $\phi_P$ from their original values to the values which satisfy equation (11) for $(G, E)$ and $(\delta, \mu)$. This method which gives more adaptability for the selection of $(G, E)$ and $(\delta, \mu)$. □

## 4. Obstacle Avoidance

For simplicity and without loss of generality, obstacles are denoted by spheres $S_j$. Spheres $S_j$ has $d$ as a radius.

Points $E_1$ and $E_2$ are shown in Figure 2. $r_{j_1}$ and $r_{j_2}$ are the distances from the pursuer to $E_1$ and $E_2$, respectively. The differential equations for $r_{jk}$, $\sigma_{jk}$, and $\gamma_{jk}$ between the pursuer and the center of obstacle $S_j$ are

$$\begin{cases} \dot{r}_{jk} = -v_P \cos\left(G\sigma_{\mathrm{TP}} + \delta - \sigma_{jk}\right)\cos\left(E\gamma_{\mathrm{TP}} + \mu - \gamma_{jk}\right), \\ r_{jk}\dot{\sigma}_{jk} = -v_P \sin\left(G\sigma_{\mathrm{TP}} + \delta - \sigma_{jk}\right)\cos\left(E\gamma_{\mathrm{TP}} + \mu - \gamma_{jk}\right), \\ r_{jk}\cos\,\sigma_{jk}\dot{\gamma}_{jk} = -v_P \cos\left(G\sigma_{\mathrm{TP}} + \delta\right)\sin\left(E\gamma_{\mathrm{TP}} + \mu - \gamma_{jk}\right). \end{cases} \tag{12}$$

With $k = 1$ and 2, one can identify whether the pursuer is oncoming or deviating from the obstacle based on equation (12). The pursuer is in a collision when

$$\left\{ \theta_P(t) \in \left[\sigma_{j_1}, \sigma_{j_2}\right]\phi_P(t) \in \left[\gamma_{j_1}, \gamma_{j_2}\right], \right. \tag{13}$$

where $\sigma_{j_1} < \sigma_{j_2}$ and $\gamma_{j_1} < \gamma_{j_2}$. The avoidance course of the proportional navigation is

$$\left\{ G\sigma_{\mathrm{TP}}(t) + \delta \notin \left[\sigma_{j_1}, \sigma_{j_2}\right] E\gamma_{\mathrm{TP}}(t) + \mu \notin \left[\gamma_{j_1}, \gamma_{j_2}\right], \right. \tag{14}$$

when the pursuer is within a definite distance $d_0$ from the obstacle.

This section can provide free or obstacle directions, which is designed in consideration of the obstacles as follows:

$$q_j = \begin{cases} 1, & \theta_P(t) \in \left[\sigma_{j_1}, \sigma_{j_2}\right] \phi_P(t) \in \left[\gamma_{j_1}, \gamma_{j_2}\right], \\ 0, & \text{otherwise.} \end{cases} \tag{15}$$

With $Q = \cup_{j=1}^{K} q_j$, where $K$ is the total amount of obstacles.

Point $T_1$ corresponds to a free direction. $T_0$ is the point where the pursuer starts deviating from a possible obstacle. When $(r_i - d) < d_0$, the pursuer is driven to an intermediary target that occurs in a free direction.

$\sigma^0_{\mathrm{TP}_0}$ and $\gamma^0_{\mathrm{TP}_0}$ denote the pitch angle and yaw angle of pursuer-target measured at point $T_0$ at time $t_1^0$, and $\sigma^1_{\mathrm{TP}_0}$ and $\gamma^1_{\mathrm{TP}_0}$ the pitch angle and yaw angle pursuer-point $T_1$ measured at point $T_0$ at the same time. $(\sigma^0_{\mathrm{TP}_0}, \gamma^0_{\mathrm{TP}_0})$ and $(\sigma^1_{\mathrm{TP}_0}, \gamma^1_{\mathrm{TP}_0})$ are

$$\begin{cases} \tan\,\sigma^0_{\mathrm{TP}_0} = \dfrac{z_T - z_P\left(t_1^0\right)}{\sqrt{\left(x_T - x_P\left(t_1^0\right)\right)^2 + \left(y_T - y_P\left(t_1^0\right)\right)^2}} \tan\,\gamma^0_{\mathrm{TP}_0} = \dfrac{y_T - y_P\left(t_1^0\right)}{z_T - z_P\left(t_1^0\right)}, \\ \tan\,\sigma^1_{\mathrm{TP}_0} = \dfrac{z_{T_1} - z_P\left(t_1^0\right)}{\sqrt{\left(x_{T_1} - x_P\left(t_1^0\right)\right)^2 + \left(y_{T_1} - y_P\left(t_1^0\right)\right)^2}} \tan\,\gamma^1_{\mathrm{TP}_0} = \dfrac{y_{T_1} - y_P\left(t_1^0\right)}{z_{T_1} - z_P\left(t_1^0\right)}, \end{cases} \tag{16}$$

where $(x_{T_1}, y_{T_1}, z_{T_1})$ are the coordinates of point $T_1$. For the flatness of the path, then

$$\begin{cases} G_0\sigma^0_{\mathrm{TP0}} + \delta_0 = G_1\sigma^1_{\mathrm{TP0}} + \delta_1 \\ E_0\gamma^0_{\mathrm{TP0}} + \mu_0 = E_1\gamma^1_{\mathrm{TP0}} + \mu_1. \end{cases} \tag{17}$$

One can determine values of $(G_1, \delta_1)$ and $(E_1, \mu_1)$ that fulfill equation (17) and move the pursuer to point $T_1$. If the pursuer suffers other obstacles, this strategy is repeated.

## 5. Two-Sensor Information Fusion

Two-sensor discrete-time system is

$$\left\{ x(\tau + 1) = Bx(\tau) + w(\tau) y_i(\tau) = \Upsilon_i x(\tau) + v_i(\tau), \quad i = 1 \text{ and } 2, \right. \tag{18}$$

where

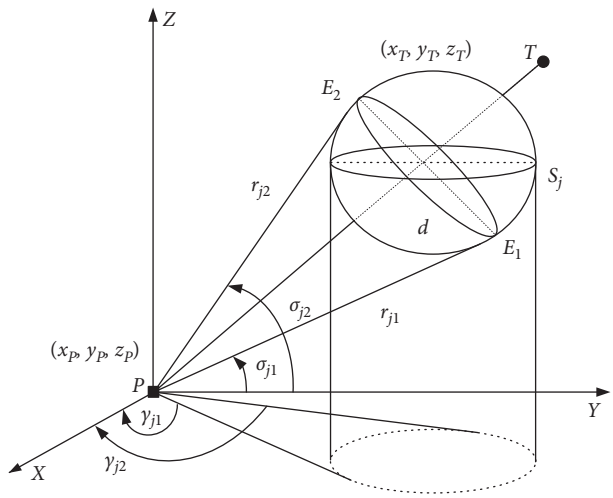



Figure 2: Representation of obstacle avoidance.

$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \quad \Upsilon_i = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{19}$$

Based on [27], the local optimal Kalman filter is

$$\begin{aligned} \widehat{x}_i(\tau|\tau) &= \left[I_n - K_i(\tau)\Upsilon_i\right]B\widehat{x}_i(\tau-1|\tau-1) + K_i(\tau)y_i(\tau), \\ K_i(\tau) &= P_i(\tau|\tau-1)\Upsilon_i^T\left[\Upsilon_i P_i(\tau|\tau-1)\Upsilon_i^T + R_i\right]^{-1}, \\ P_i(\tau|\tau-1) &= BP_i(\tau-1|\tau-1)B^T + \mathrm{GQG}^T, \\ P_i(\tau|\tau) &= \left[I_n - K_i(\tau)\Upsilon_i\right]P_i(\tau|\tau-1). \end{aligned} \tag{20}$$

Thus, the optimal Kalman filter is

$$\widehat{x}_0(\tau|\tau) = \zeta_1(\tau)\widehat{x}_1(\tau|\tau) + \zeta_2(\tau)\widehat{x}_2(\tau|\tau). \tag{21}$$

The optimal matrix of weight coefficients designed in [17, 18] can be calculated as

$$\zeta_i = \mathrm{diag}\left(\alpha_{i_1}, \alpha_{i_2}, \ldots, \alpha_{\mathrm{in}}\right), \quad i = 1 \text{ and } 2, \tag{22}$$

where the optimal weight coefficients are

$$\alpha_{1i} = \frac{P_2^{ii} - P_{12}^{ii}}{P_1^{ii} + P_2^{ii} - 2P_{12}^{ii}}, \quad \alpha_{2i} = \frac{P_1^{ii} - P_{12}^{ii}}{P_1^{ii} + P_2^{ii} - 2P_{12}^{ii}}, \tag{23}$$

where $P_i^{ii}$ and $P_{12}^{ii}$ are the diagonal element of $P_i$ and $P_{12}$. The error covariance matrix is

$$P_{0i} = \frac{P_1^{ii}P_2^{ii} - \left(P_{12}^{ii}\right)^2}{P_1^{ii} + P_2^{ii} - 2P_{12}^{ii}}, i = 1, 2, \ldots, n. \tag{24}$$

The trace of the error covariance matrix for information fusion is

$$\mathrm{tr}P_0(\tau|\tau) = P_{01} + P_{02} + \cdots + P_{0n}, \tag{25}$$

where $\mathrm{tr}P_0 \le \mathrm{tr}P_i$, $i = 1$ and 2.

By using the Kalman estimator, fusion position of the target $(x_{T_{\mathrm{Fu}}}, y_{T_{\mathrm{Fu}}}, z_{T_{\mathrm{Fu}}})$ and the pursuer $(x_{P_{\mathrm{Fu}}}, y_{P_{\mathrm{Fu}}}, z_{P_{\mathrm{Fu}}})$ can be obtained. Thus, $\sigma_{\mathrm{Fu}} = \arctan z_{T_{\mathrm{Fu}}} - z_{P_{\mathrm{Fu}}} / \sqrt{\left(x_{T_{\mathrm{Fu}}} - x_{P_{\mathrm{Fu}}}\right)^2 + \left(y_{T_{\mathrm{Fu}}} - y_{P_{\mathrm{Fu}}}\right)^2}$ and $\gamma_{\mathrm{Fu}} = \arctan y_{T_{\mathrm{Fu}}} - y_{P_{\mathrm{Fu}}} / x_{T_{\mathrm{Fu}}} - x_{P_{\mathrm{Fu}}}$. Under two-sensor information fusion, equation (3) is written as

$$\begin{cases} \theta_{P_{\mathrm{Fu}}}(t) = G\sigma_{\mathrm{Fu}}(t) + \delta, \\ \phi_{P_{\mathrm{Fu}}}(t) = E\gamma_{\mathrm{Fu}}(t) + \mu. \end{cases} \tag{26}$$

## 6. Simulation Results

This section proposes several simulations, where tracking can be implemented under proportional navigation. In this section, distances velocities and time have been with units to achieve realistic results.

*Example 1.* The original position of air vehicle is (12 m, 12 m, 12 m), with $\theta_P(t_0) = \pi$ and $\phi_P(t_0) = \pi/2$. The aerial target is situated at (110 m, 110 m, 150.6 m); thus, $\sigma_{\mathrm{TP}}(t_0) = \pi/4$ and $\gamma_{\mathrm{TP}}(t_0) = \pi/4$. The solutions for the two approaches discussed above are as follows.

(1) Since $(\theta_P(t_0) = \pi, \phi_P(t_0) = \pi/2)$ and $(\sigma_{\mathrm{TP}}(t_0) = \pi/4, \gamma_{\mathrm{TP}}(t_0) = \pi/4)$, $(G, \delta)$ and $(E, \mu)$ are calculated such that

$$\begin{cases} \pi = G\dfrac{\pi}{4} + \delta, \\ \dfrac{\pi}{2} = E\dfrac{\pi}{4} + \mu. \end{cases} \tag{27}$$

There exists an infinite number of methods for $(G, \delta)$ and $(E, \mu)$ which satisfy equation (27), and hence, there exists an infinite number of possible trajectories for the air vehicle. One can take [$(G = 3, \delta =$

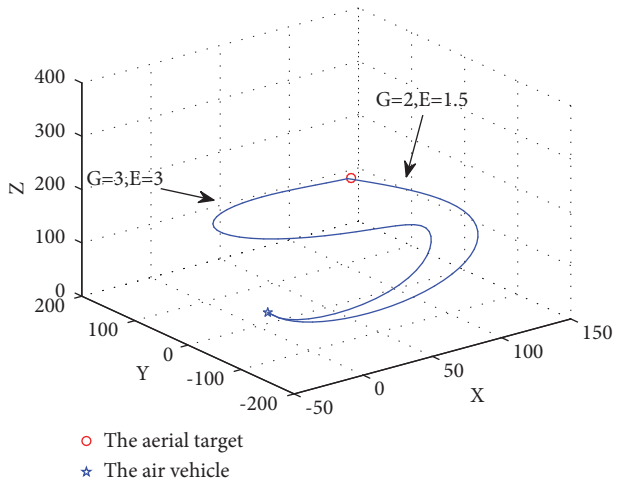



Figure 3: Configuration satisfied by selecting control parameters.

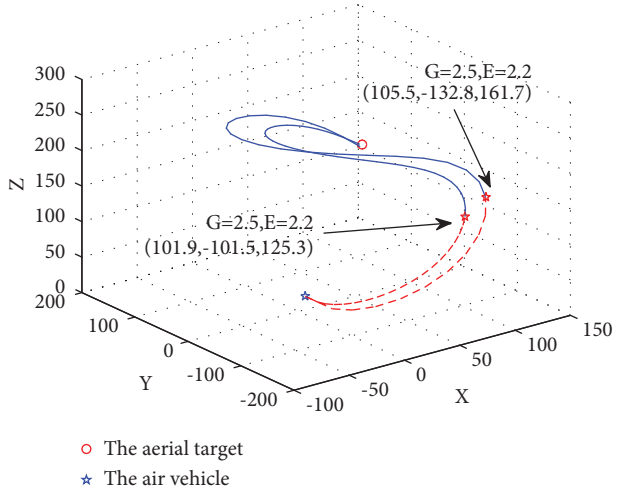



Figure 4: Illustration of the heading regulation.

$\pi/4)$, $(E = 3, \mu = -\pi/4)]$ or $[(G = 2, \delta = \pi/2), (E = 1.5, \mu = \pi/8)]$. Thus, the trajectory of the air vehicle is predefined by the original conditions and the values of $(G, \delta)$ and $(E, \mu)$. As illustrated in Figure 3, the air vehicle can reach the aerial target using this method.

(2) $(G, \delta)$ and $(E, \mu)$ are predetermined, and the heading regulation phase is necessary. Heading regulation is to take the air vehicle $(\theta_P^i, \phi_P^i)$ from the initial values to the intermediary values $(\theta_{P_0}^i, \phi_{P_0}^i)$ that satisfy equation (11). One can take $(G = 2.5, \delta = 3\pi/8)$ and $(E = 2.2, \mu = -\pi/20)$.

There exist various potentialities for the selection of $(x_P^i, y_P^i, z_P^i)$. One takes $(x_P^i, y_P^i, z_P^i) = (101.9\,\text{m}, -101.5\,\text{m}, 125.3\,\text{m})$; thus, $\theta_{P_0}^i = 1.61$ rad and $\phi_{P_0}^i = 3.2$ rad. One takes $(x_P^i, y_P^i, z_P^i) = (105.5\,\text{m}, -132.8\,\text{m}, 161.7\,\text{m})$; thus, $\theta_{P_0}^i = 1.33$ rad and $\phi_{P_0}^i = 3.26$ rad. Some methods from control theory are applied to the intention of heading regulation. The air vehicle navigation applying this method is shown in Figure 4. The dashed lines illustrate the path of the air vehicle under the heading regulation phase. This method can give more adaptability for the selection of $(G, \delta)$ and $(E, \mu)$.

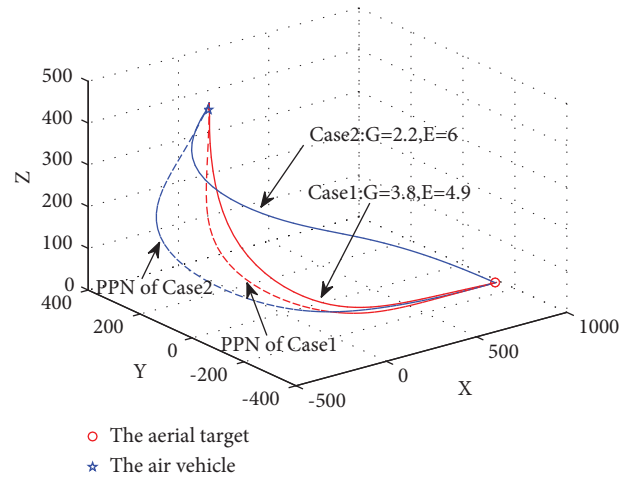



Figure 5: Comparison with proportional navigation and PPN.

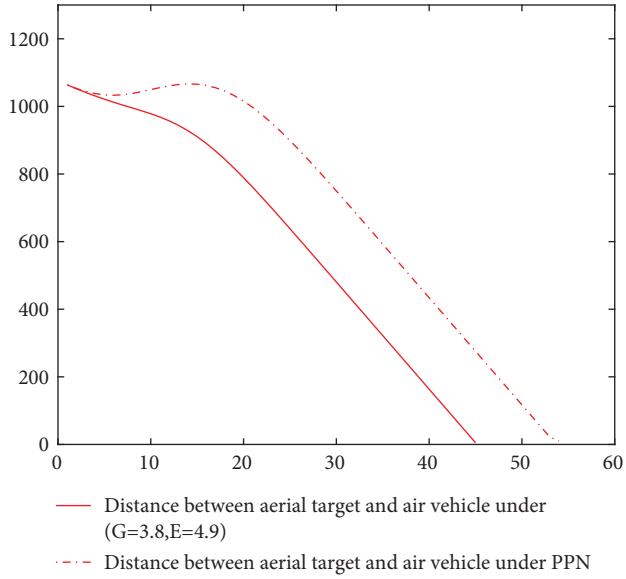



Figure 6: Distance between the aerial target and air vehicle under case 1.

*Example 2.* Comparison with the PPN, and this example will be considered. The original position of the air vehicle is (22 m, 300 m, 400 m), and the aerial target is situated at (880 m, −208 m, 30 m). In the first case, one can take $(G = 3.8, E = 4.9)$. From Figures 5 and 6, it is observed that the air vehicle navigating under proportional navigation reaches the aerial target before the PPN. The interception times are 45 s and 54 s for proportional navigation and PPN, respectively. In the second case, one can take $(G = 2.2, E = 6)$. From Figures 5 and 7, the interception times are 50 s and 63 s for proportional navigation and PPN, respectively. This example indicates that proportional

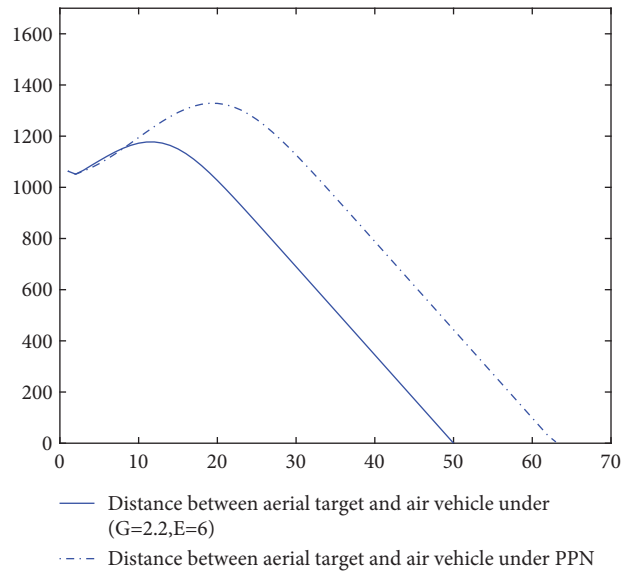



Figure 7: Distance between aerial target and air vehicle under case 2.

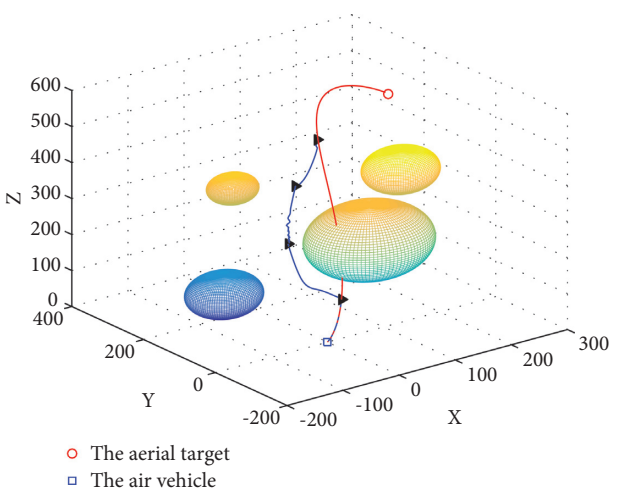



Figure 8: Online deviation at the appearance of spherical obstacles.

navigation can implement tracking the target, outperforming PPN in the field of interception time.

*Example 3.* At the appearance of spherical obstacles in a complicated environment, the point-to-point navigation method is applied to reach the aerial target and avoid the obstacles. The pursuer initiates from the initial position (0 m, 0 m, 0 m) and aims to reach the aerial target situated at (300 m, 300 m, 424.2 m). As illustrated in Figure 8, online deviation towards intermediary aerial targets $T_1$, $T_2$, $T_3$, and $T_4$ is applied with the following different control parameters. Phase $PT_1$: $(G = 2, \delta = \pi/4)$ and $(E = 1.2, \mu = 3\pi/40)$. Phase $T_1T_2$: $(G = 2.1, \delta = -10\pi/34)$ and $(E = 6.8, \mu = 10\pi/21)$. Phase $T_2T_3$: $(G = 2.1, \delta = -\pi/4)$ and $(E = 4, \mu = \pi/3)$. Phase $T_3T_4$: $(G = 1.3, \delta = -\pi/4)$ and $(E = 2, \mu = \pi/2)$. Phase $T_4T$: $(G = 2, \delta = \pi/4)$ and $(E = 1.2, \mu = 3\pi/40)$. The path of the air vehicle is $P \longrightarrow T_1 \longrightarrow T_2 \longrightarrow T_3 \longrightarrow T_4 \longrightarrow T$. These

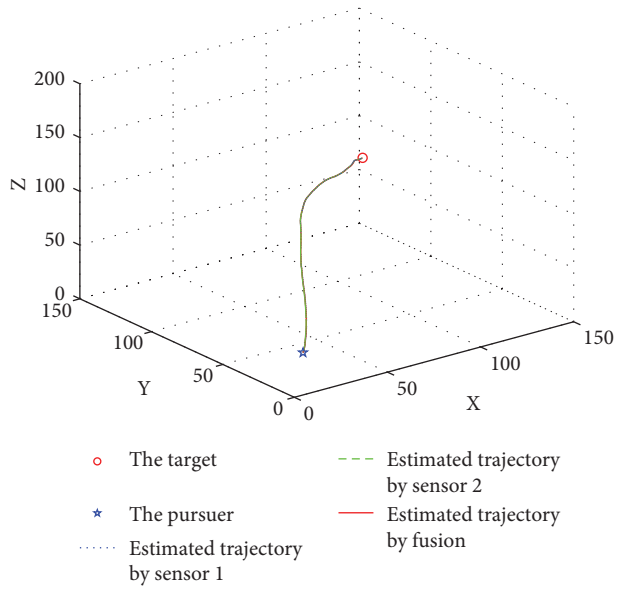



Figure 9: Filtered trajectory under proportional navigation.

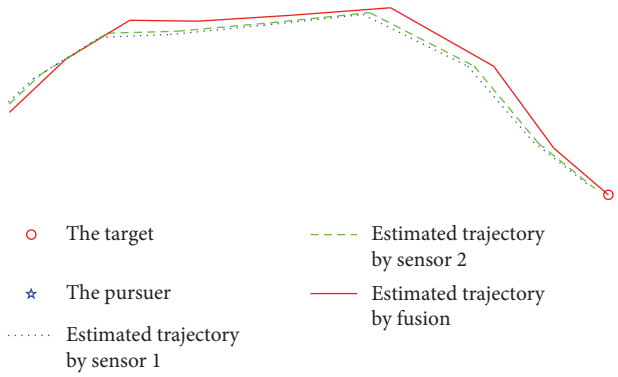



Figure 10: Enlargement in tracking result of Figure 9.

points can be selected so that the differences from the titular trajectory are small, which keeps smoothness of the trajectory. The air vehicle applies the point-to-point method to navigate towards the aerial target and avoid the obstacles.

*Example 4.* Under sensor noise, the initial position of the purser is (20 m, 20 m, 20 m) and the target is situated at (90 m, 90 m, 118.98 m). Thus, one can take $(G = 2, \delta = \pi/8)$ and $(E = 2, \mu = -\pi/8)$. In this case, two-sensor information fusion weighted by diagonal matrices together with proportional navigation is given to enhance the tracking precision. Figure 9 shows the filtered trajectory under proportional navigation. Enlargement in tracking result of Figure 9 is shown in Figure 10. The analysis of tracking performance indicates that the more higher the precision, the less the trace of the error covariance matrix. From Figure 11, one obtains $\mathrm{tr}P_i(k|k) > \mathrm{tr}P_0(k|k)$. As a result, the trace of error the covariance matrix under information fusion is lower than the value of the single sensor. Then, two-sensor information fusion provides proportional navigation with more accurate target estimates. This example shows

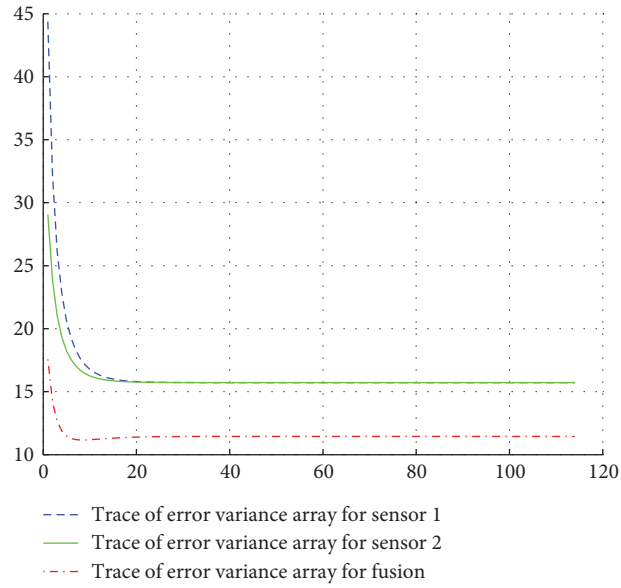



Figure 11: Trace of the error covariance matrix under two-sensor information fusion.

that the fusion result is better than that of the single sensor and the fusion method is effective.

## 7. Conclusion

This paper proposes a method for pursuer navigation under proportional navigation. The control strategy is primitive and depends on only the position of the target. For obstacle avoidance, this paper can avoid it by adjusting the control parameters. In the presence of sensor noise, the proportional navigation combined with information fusion weighted by diagonal matrices can achieve more reasonable interception performance. The method opens new directions for research, such as navigation using the proportional navigation under kinematics of pursuer and dynamics constraints and the influence of the control parameters, especially $G$ and $E$.

## Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

## Conflicts of Interest

The authors declare that there are no conflicts of interest.

## Acknowledgments

This work was supported by Key Laboratory of Cyber-Physical System and Intelligent Control in Universities of Shandong, Ph.D. Programs Foundation of Ludong University, National Natural Science Foundation of China (no. 61903172), and Open Project Program of Tianjin Key Laboratory of Microgravity and Hypogravity Environment Simulation Technology (no. TJWDZL2019KT003).

## References

[1] N. Cho and Y. Kim, "Optimality of augmented ideal proportional navigation for maneuvering target interception," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 52, no. 2, pp. 948–954, 2016.

[2] I. S. Jeon, J. I. Lee, and M. J. Tahk, "Impact-time-control guidance with generalized proportional navigation based on nonlinear formulation," *Journal of Guidance, Control, and Dynamics*, vol. 39, no. 8, pp. 1887–1892, 2016.

[3] F. Tyan, "Analysis of general ideal proportional navigation guidance laws," *Asian Journal of Control*, vol. 18, no. 3, pp. 899–919, 2016.

[4] S. X. Yang, A. Anmin Zhu, G. F. Guangfeng Yuan, and M. Q. Meng, "A bioinspired neurodynamics-based approach to tracking control of mobile robots," *IEEE Transactions on Industrial Electronics*, vol. 59, no. 8, pp. 3211–3220, 2012.

[5] S.-C. Han, H. Bang, and C.-S. Yoo, "Proportional navigation-based collision avoidance for UAVs," *International Journal of Control, Automation and Systems*, vol. 7, no. 4, pp. 553–565, 2009.

[6] P. Long, W. Liu, and J. Pan, "Deep-learned collision avoidance policy for distributed multiagent navigation," *IEEE Robotics and Automation Letters*, vol. 2, no. 2, pp. 656–663, 2017.

[7] M. Seder, M. Baotic, and I. Petrovic, "Receding horizon control for convergent navigation of a differential drive mobile robot," *IEEE Transactions on Control Systems Technology*, vol. 25, no. 2, pp. 653–660, 2017.

[8] Y. Chen, J. Wang, C. Wang, J. Shan, and M. Xin, "A modified cooperative proportional navigation guidance law," *Journal of the Franklin Institute*, vol. 356, no. 11, pp. 5692–5705, 2019.

[9] K. B. Li, H. S. Shin, A. Tsourdos, and M. Tahk, "Capturability of 3D PPN against lower-speed maneuvering target for homing phase," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 56, no. 1, pp. 711–722, 2019.

[10] S. Lee, S. Ann, N. Cho, and Y. Kim, "Capturability of guidance laws for interception of nonmaneuvering target with field-of-view limit," *Journal of Guidance, Control, and Dynamics*, vol. 42, no. 4, pp. 869–884, 2019.

[11] Y. Liu, K. Li, L. Chen, and Y. Liang, "Novel augmented proportional navigation guidance law for mid-range autonomous rendezvous," *Acta Astronautica*, vol. 162, no. 9, pp. 526–535, 2019.

[12] F. Belkhouche and B. Belkhouche, "Wheeled mobile robot navigation using proportional navigation," *Advanced Robotics*, vol. 21, no. 3-4, pp. 395–420, 2007.

[13] K. Liu, K. Li, Q. Peng, Y. Guo, and L. Zhang, "Data-driven hybrid internal temperature estimation approach for battery thermal management," *Complexity*, vol. 2018, Article ID 9642892, 15 pages, 2018.

[14] X. Tang, K. Liu, X. Wang, B. Liu, F. Gao, and W. D. Widanage, "Real-time aging trajectory prediction using a base model-oriented gradient-correction particle filter for lithium-ion batteries," *Journal of Power Sources*, vol. 440, Article ID 227118, 2019.

[15] S.-L. Sun and Z.-L. Deng, "Multi-sensor optimal information fusion kalman filter," *Automatica*, vol. 40, no. 6, pp. 1017–1023, 2004.

[16] Y. Gao, L. Mao, Z. J. Liang, and Z. L. Deng, "Two-sensor information fusion steady-state kalman filter weighted by diagonal matrices," *Science Technology and Engineering*, vol. 21, no. 2, pp. 52–54, 2004.

[17] Z. L. Deng, *Optimal Estimation Theory with Applications*, Harbin Institute of Technology Press, Harbin, China, 2005.

[18] Z. L. Deng, *Information Fusion Filtering Theory with Applications*, Harbin Institute of Technology Press, Harbin, China, 2007.
[19] J. J. Gu, M. Meng, A. Cook, and G. Faulkner, "Simulation study of artificial ocular movement with intelligent control," *Control Engineering Practice*, vol. 13, no. 4, pp. 509–518, 2005.
[20] Y. L. Shi, W. M. Zhang, Z. Yao et al., "Design of a hybrid indoor location system based on multi-sensor fusion for robot navigation," *Sensors*, vol. 18, no. 10, pp. 1–18, 2018.
[21] S. L. Sun and P. Y. Cui, "Two sensors optimal information fusion kalman filter and its application in tracking systems," *Journal of Astronautics*, vol. 24, no. 2, pp. 206–209, 2003.
[22] Z. L. Deng, C. X. Cui, and J. G. Bai, "On functional equivalence of two measurement fusion methods based on steady-state kalman filtering," *Science Technology and Engineering*, vol. 4, no. 11, pp. 897–902, 2004.
[23] S. A. Elgamel and A. H. Makaryous, "Performance of modern guided systems in presence of jamming signal," in *Proceedings of the International Conference on Aerospace Science and Aviation Technology*, Cairo, Egypt, 2001.
[24] S. A. Elgamel and J. J. Soraghan, "Target tracking enhancement using a kalman filter in the presence of interference," in *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, pp. 681–684, Cape Town, South Africa, July 2009.
[25] J. K. Hale and H. Kocak, *Dynamics and Bifurcation*, Springer-Verlag, Berlin, Germany, 1991.
[26] S. Wiggins, *Introduction to Applied Nonlinear Dynamical Systems and Chaos*, Springer-Verlag, Berlin, Germany, 1990.
[27] Y. Bar-Shalom, X. R. Li, and T. Kirubarajan, *Estimation with Application to Tracking and Navigation*, John Wiley, Hoboken, NJ, USA, 2001.

Hindawi
Complexity
Volume 2021, Article ID 4981691, 11 pages
https://doi.org/10.1155/2021/4981691

*Research Article*

# Flow Field Analysis of Adult High-Flow Nasal Cannula Oxygen Therapy

**Jingen Xia**,[1,2,3,4,5] **Jiaqi Chang**,[6] **Jixiang Liang**,[7] **Yixuan Wang**,[6] **and Na Wang**[8]

[1]*School of Biological Science and Medical Engineering, Beihang University, Beijing, China*
[2]*Department of Pulmonary and Critical Care Medicine, Center of Respiratory Medicine, China-Japan Friendship Hospital, Beijing, China*
[3]*National Center for Respiratory Medicine, Beijing, China*
[4]*Institute of Respiratory Medicine, Chinese Academy of Medical Sciences, Beijing, China*
[5]*National Clinical Research Center for Respiratory Diseases, Beijing, China*
[6]*School of Automation Science and Electrical Engineering, Beihang University, Beijing, China*
[7]*The State Key Laboratory for Manufacturing Systems Engineering, Xi'an Jiaotong University, Xi'an, Shaanxi, China*
[8]*Engineering Training Cent, Beihang University, Beijing, China*

Correspondence should be addressed to Jiaqi Chang; chang_jiaqi@buaa.edu.cn

Received 23 April 2021; Revised 22 June 2021; Accepted 3 July 2021; Published 22 July 2021

Academic Editor: Bo Xiao



The mechanical ventilation of human body is a complex human-computer interaction process. High flow nasal cannula oxygen therapy (HFNC) is a new type of ventilation, which is often measured by lung pressure, respiratory work, and other parameters. The purpose of this paper is to analyse the pressure, flow, and strain rate of upper respiratory tract with different flow and oxygen concentration by using finite element simulation, to guide professionals to adjust the appropriate flow and oxygen concentration parameters of HFNC machine. This paper studies the complex human-computer interaction environment of human respiratory tract and ventilation airflow. The 3D model of respiratory tract established by the conversion of image scanning data was taken as the research object. The flow state of the gases in the respiratory tract was judged by Reynolds equation. After that, RNG K-$\varepsilon$ model was applied to the research object, and the simulation diagram of airway pressure, flow rate, strain rate, and trace diagram of flowing particles were obtained under the finite element method. The results explain some clinical phenomena in HFNC and guide people to make better use of mathematical tools to study human-computer complex environment.

## 1. Introduction

The main clinical manifestations of patients with respiratory diseases are dyspnea, cough, expectoration, and other symptoms, which have serious adverse effects on the physical and mental health and quality of life of patients. At present, many measures and methods for the treatment of respiratory diseases are put forward. Oxygen therapy is the most commonly used treatment of respiratory diseases. Traditional oxygen therapy uses a nasal catheter or mask to inhale oxygen. It can help improve the respiratory function of patients with respiratory diseases, promote the functional indicators of patients to return to normal quickly, and delay and block the development trend of the disease, to improve the prognosis of patients. However, traditional oxygen therapy is difficult to achieve enough humidification degree and temperature, and the oxygen flow rate is limited. High flow nasal cannula oxygen therapy (HFNC) is the latest respiratory support technology, which can meet the needs of airway humidification, proper temperature, and oxygen flow rate at the same time. These advantages make it more comfortable for patients and are rapidly promoted in clinical practice [1]. The oxygen from the electromagnetic flow proportional valve is fully mixed with the air and heated and humidified by flowing through hot water and finally enters the human body with a high flow rate and appropriate temperature and humidity.

HFNC has the following advantages [2]: (1) it can provide more stable and higher inhalation oxygen concentration than the ordinary therapy, which can meet the needs of patients' spontaneous breathing [3]; (2) high flow airflow can reach or exceed the maximum inspiratory flow of patients' active inspiration, which can reduce inspiratory resistance and respiratory work [4]; (3) it can warm and humidify the gas to body temperature, which can reduce the heat and water consumption of patients with respiratory distress, and maintain the airway mucociliary function in the best state; (4) high flow airflow washes the dead space of upper airway, reducing anatomical dead space and improving patient ventilation efficiency [5, 6]; (5) high flow airflow provides a certain level of end-expiration positive airway pressure, which can open alveoli and increase lung volume; (6) HFNC does not need a completely closed path and is convenient to eat and communicate [7], while it can increase the comfort of patients [8]. These advantages were also evaluated in vitro under different flow rate, intubation size, and air leakage [9]. HFNC has also been shown to protect against respiratory failure and reintubation after extubation [10].

As a new way of ventilation, it has shown great practicality. Miguel-Montanes pointed out that HFNC significantly improved preoxygenation and reduced the incidence of severe hypoxemia [11]. Messika demonstrated that HFNC is effective for oxygenation in a small number of acute respiratory failure (ARF) through single centre study [12]. Nishimura cites others studies to conclude that HFNC is an effective early treatment for adult respiratory failure [13]. Spoletini suggested that HFNC is not only more comfortable than NIV but also relieves more pain of patients with dyspnea [14]. Manizheh compared the efficacy and safety of HFNC with nasal CPAP for respiratory support in premature infants with respiratory distress syndrome (RDS) [15]. Schlapbach used it in hospital transfer of children with critical disease and found that the treatment rate was significantly higher than that of noninvasive ventilation [16].

The flow field analysis involved in this study is the analysis of nasopharynx and trachea, including nasal cavity, nasopharynx, oropharynx, pharynx, and trachea. These parts are not only the channels of air, but also have the functions of defense, regulating air temperature, and humidity. Due to the complexity of the structure, the simulation results of computational fluid dynamics (CFD) technology have become an important reference for clinical research, which can be used to design and improve the high-flow nasal cannula oxygen therapy. Therefore, ensuring the accuracy of CFD simulation results and making a reasonable analysis of the results became the key to the study.

In this study, the flow field is analysed by FLUENT software, which replaces the physical model to grid model and calculates approximate solution of pressure or velocity in each grid so that the complex problem of the whole flow field state in respiratory tract is turned to a simple grid problem. It is not only highly accurate, but also suitable for the complex shape of respiratory tract. CFD to respiratory tract can not only effectively simulate the air distribution during oxygen therapy but also obtain the velocity and pressure of each point in the trachea. It is of great significance for diagnosing and treating respiratory diseases, adjusting HFNC equipment to the appropriate state accepted by patients, and reducing ventilator injury of patients.

This method is helpful for professionals to understand the airway condition of patients with different flow and oxygen concentration and can provide doctors with certain reference when adjusting the flow and oxygen concentration of HFNC. It can provide sufficient inspiratory flow and oxygen supply level, thus reducing inspiratory resistance and respiratory work, increasing ventilation efficiency, and avoiding airway damage to patients.

## 2. Methods

The method of flow field analysis is divided into the following steps: establishing geometric model, analysing flow form and flow model, setting material and boundary conditions, setting analysis steps, completing analysis, and postprocessing.

*2.1. Geometric Model of Upper Respiratory Tract.* In geometric model, two-dimensional (2D) model has simpler modelling, but for complex respiratory system, three-dimensional (3D) model is almost used to describe the structural characteristics of respiratory system in recent years. At present, the following methods are mainly used in the establishment of three-dimensional model: (1) the classic anatomical model is Weibel A model, which is an ideal geometric symmetry model. After that, some scholars put forward the lung model of asymmetric bifurcation region [17] and the asymmetric bifurcation bronchial tree model [18]. (2) The 3D model is established based on the transformation of image scanning data. 3D reconstruction software for computed tomography (CT) or magnetic resonance imaging (MRI) greatly facilitates the generation of 3D models, such as Mimics software. Obviously, the latter modelling method is closer to the real human respiratory tract and has more geometric details. However, due to the influence of noise and image resolution, the important parameters of smaller airway still need to be obtained by anatomical model [19].

The model studied is a 32-year-old female's upper respiratory tract and trachea. The surface model of respiratory tract wall was established by Mimics software (as shown in Figure 1), and then, SIMENS NX software was used to transform the surface model into a solid model. Then, the solid model of cavity was extracted (as shown in Figure 2) and finally meshed the model in FLUENT software (as shown in Figure 3). In the model of endonasal cavity, the diameter of nasal tube is 5 mm, the diameter of nostril is 10 mm, the axial length of nasal cavity is 73.29 mm, the transverse length of nasal cavity is 30.28 mm, the axial length of throat is 86.14 mm, the axial length of trachea is 109.38 mm, and the diameter of trachea end is 19.5 mm.

During oxygen therapy, the mixed gas flows from the nostril to the lung. The blue section is the air inlet, and the red section is the air outlet. The model has 653723 grids and 194538 nodes.

*2.2. Flow Model of Airflow during Ventilation.* To obtain the flow form, it is necessary to set the ventilation flow rate higher than the maximum flow rate of nature breathing,

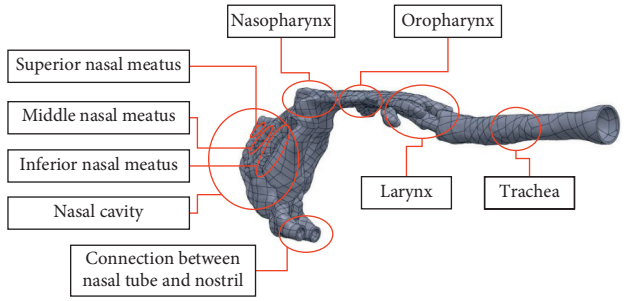



Figure 1: The structure model of airway wall based on CT scan.

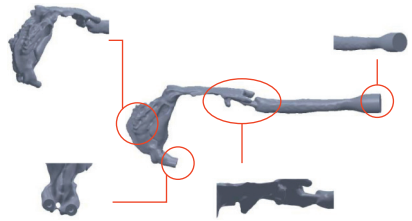


Figure 2: Solid details of airway filled based on the structure model of airway wall.

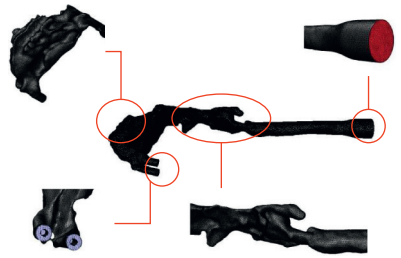


Figure 3: Mesh details of airway filled based on the structure model of airway wall.

which is among 40–60 L/min commonly. In the model, the diameter of nasal tube is assumed to be 5 mm, so the ventilation flow rate is selected as 32 m/*s*. Simultaneously, there are three kinds of operation conditions: 50 L/min and 60 L/min to describe the conditions of medium flow rate and high flow rate. Therefore, the gas flow rates under the three conditions are 32 m/*s*, 40 m/*s* and 48 m/*s* respectively. The calculation of Reynolds number is as follows:

$$\mathrm{Re} = \frac{v\rho d}{\mu}, \tag{1}$$

where $v$ is the velocity, $\rho$ is the density of air oxygen mixture, $d$ is the diameter of gas vent, and $\mu$ is the dynamic viscosity of air oxygen mixture.

The supply gas of the ventilator is a mixture of air and oxygen. In the finite element simulation, the density and dynamic viscosity of the gas mixture should be determined. These two variables are affected by the nature and composition of each molecule in the gas. The proportional valve of the oxygen therapy machine controls the amount of oxygen, and the fan controls the amount of air, which reflects the mixing degree of the gas mixture. The density and dynamic viscosity of the gas mixture can be calculated by the linear relationship of the mixing ratio of air and oxygen. When the mixing ratio of oxygen and air is $\alpha$ ($0\% < \alpha < 100\%$), the density and dynamic viscosity of mixed gas meet the following requirements:

$$\rho = 1.293\,(1-\alpha) + 1.429\alpha, \tag{2}$$

$$\mu = 18.448\,(1-\alpha) + 20.55\alpha, \tag{3}$$

$$\alpha = \frac{q_{v.o_2}}{q_{v.\mathrm{air}}}. \tag{4}$$

According to Reynolds number, laminar and turbulent models can be judged. The general fluid motion equation is Navier–Stokes (N–S) equation:

$$\rho\left(\frac{\partial V}{\partial t} + (V\cdot\nabla)V\right) = f - \nabla P + \mu\nabla^2 V. \tag{5}$$

It can be written in rectangular coordinates:

$$\begin{aligned}
\rho\left(\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x}\right) &= f_x - \frac{\partial P}{\partial x} + \mu\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right),\\
\rho\left(\frac{\partial v}{\partial t} + v\frac{\partial v}{\partial y}\right) &= f_y - \frac{\partial P}{\partial y} + \mu\left(\frac{\partial^2 v}{\partial x^2} + \frac{\partial^2 v}{\partial y^2} + \frac{\partial^2 v}{\partial z^2}\right),\\
\rho\left(\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial z}\right) &= f_z - \frac{\partial P}{\partial z} + \mu\left(\frac{\partial^2 w}{\partial x^2} + \frac{\partial^2 w}{\partial y^2} + \frac{\partial^2 w}{\partial z^2}\right),
\end{aligned} \tag{6}$$

where $\rho$ is the density of the fluid, $P$ is the pressure, $V$ is the velocity vector, $u$, $v$, and $w$ are the velocity components of the fluid at point (x, $y$, and z) at time $t$, $f$ is the external force on the fluid per unit volume, and $\mu$ is the dynamic viscosity.

According to the calculation, the Reynolds numbers of 40 L/min, 50 L/min, and 60 L/min at 70% mixing ratio of oxygen and air at the narrowest part in the airway are 11155, 13943.75, and 16732.5 respectively, indicating that the three cases are turbulent flow, and from the physiological point of view, the existence of turbulence can enhance the contact between the airflow and the mucosal boundary layer, and improve the heating and humidifying effects of the air passage on the air.

The solution of N–S equation is very dependent on initial and boundary conditions. For turbulence, the initial condition is to add random turbulence disturbance, which causes the uncertainty and randomness of the obtained turbulent solutions. Therefore, in addition to the general fluid formula, it is necessary to introduce specialized turbulence equations.

In the selection of turbulence model equations, some studies [20] have shown that: in the case of higher gas flow rate, several commonly used turbulence models (LES model, $K$-$\varepsilon$ model, standard $k$-$\omega$ model, SST $K$-$\omega$ model) are compared. Among them, the standard $k$-$\omega$ model has the least difference between the simulation and experimental results. In addition, compared with RNG $K$-$\varepsilon$ model [21], the traditional $K$-$\varepsilon$ model is a high Reynolds number model, but

it takes the whole upper respiratory tract as turbulence, which will lead to poor performance. RNG K-$\varepsilon$ model [22] is a low Reynolds number turbulence model, which considers the laminar flow and turbulent flow in the simulation process. The RNG model adds a condition to the $\varepsilon$ equation, which not only effectively improves the accuracy, but also takes the turbulent vortex into account, which has a high credibility in the numerical simulation of airflow in the respiratory tract. Jeong et al. showed that the RNG *K-εe* model was more consistent with the experimental values than the traditional *K-ε* model when simulating the upper respiratory tract of a patient. Therefore, RNG *K-εe* model equations are used in this study [23].

The RNG *K-εe* model is shown as follows. *K* equation is equation of turbulent kinetic energy and $\varepsilon$ equation is equation of turbulent dissipation rate:

$$\frac{\partial}{\partial t}(\rho k) + \frac{\partial}{\partial t}(\rho k u_i) = \frac{\partial}{\partial x_j}\left(\alpha_k \mu_{eff} \frac{\partial k}{\partial x_j}\right) + G_k + G_b - \rho\varepsilon - Y_M + S_k, \tag{7}$$

$$\frac{\partial}{\partial t}(\rho \varepsilon) + \frac{\partial}{\partial t}(\rho \varepsilon u_i) = \frac{\partial}{\partial x_j}\left(\alpha_\varepsilon \mu_{eff} \frac{\partial \varepsilon}{\partial x_j}\right) + C_{1\varepsilon}\frac{\varepsilon}{k}(G_k + G_{3\varepsilon}G_b) - C_{2\varepsilon}\rho\frac{\varepsilon^2}{k} - R_\varepsilon + S_\varepsilon. \tag{8}$$

Among them, $C_{1\varepsilon} = 1.42$, $C_{2\varepsilon} = 1.68$, $G_k$ is the generation of turbulent kinetic energy caused by mean velocity gradient, $G_b$ is the turbulent kinetic energy generated by buoyancy, $Y_M$ is the contribution of wave expansion to the total dissipation rate in compressible turbulence, $\alpha_k$ and $\alpha_\varepsilon$ are the reciprocal of the effective Prandtl numbers of K and $\varepsilon$, and $S_k$ and $S_\varepsilon$ are the source term. In this study, buoyancy and compressibility of flow are not considered.

The differential equation of turbulent viscosity is as follows:

$$d\left(\frac{\rho^2 k}{\sqrt{\varepsilon\mu}}\right) = 1.72\frac{\dot{\nu}}{\sqrt{\dot{\nu}^3 - 1 + C_\nu}}d\dot{\nu}. \tag{9}$$

In (9), the calculation of some parameters is shown as follows:

$$\dot{\nu} = \frac{\mu_{eff}}{\mu}, \tag{10}$$

$$C_\nu \approx 100. \tag{11}$$

Expression of turbulent viscosity is as follows:

$$\mu = \rho C_\mu \frac{k^2}{\varepsilon}, \tag{12}$$

where $C_\mu = 0.0845$.

The statistical method of renormalization group theory in RNG theory provides turbulent Prandtl numbers $\alpha_k$ and $\alpha_\varepsilon$ for formulas 7 and (8), which are set as fixed values in standard *K-ε* method. Reciprocals of effective Prandtl numbers $\alpha_k$ and $\alpha_\varepsilon$ are derived from RNG theory:

$$\left|\frac{\alpha - 1.3929}{\alpha_0 - 1.3929}\right|^{0.6321}\left|\frac{\alpha + 2.3929}{\alpha_0 + 2.3929}\right|^{0.3679} = \frac{\mu_{\text{mol}}}{\mu_{\text{eff}}}, \tag{13}$$

where $\alpha_0 = 1.0$. In this paper, the Prandtl numbers have little influence on the flow field, but they increase the calculation work, so $\alpha_k = \alpha_\varepsilon = 1.393$.

$R_\varepsilon$ is an additional term in the $\varepsilon$ equation, which is expressed as follows:

$$R_\varepsilon = \frac{C_\mu \rho \eta^3 (1 - \eta/\eta_0)}{1 + \beta\eta^3}\frac{\varepsilon^2}{k}, \tag{14}$$

where $\eta_0 = 4.38$ and $\beta = 0.012$.

## 3. Materials and Boundary Conditions

The material studied is a mixture of air and oxygen. The mixture density is calculated according to the formula with the oxygen and air mixing ratio mentioned previously. In this study, we analysed the oxygen concentration varying from 21% to 100%, and the oxygen and air mixing ratio is 0% to 100%. The density and dynamic viscosity of the gas mixture can be calculated by these data. In this study, the reduction of total density of mixed gas caused by the introduction of water vapor is not considered because of the very little pressure of saturated vapor pressure of water in the gas from HFNC equipment.

There are four flow surfaces in the nasal cavity and a flow surface at the end of the trachea. These five surfaces are the boundary interfaces. In most cases of inhalation, the nasal cavity does not get additional air from the outside but relies on HFNC ventilation. The mechanical ventilation surfaces are also the flow velocity input surface, and the surfaces from nostril and trachea are set as the pressure output surface and the pressure value is atmospheric pressure. On expiration, the space where the nostrils are not intubated is set as the pressure output surface, and the pressure value is atmospheric pressure. The surfaces of the HFNC ventilation surface and trachea are the flow rate input surface.

*3.1. Analysis Steps Setting.* The transient solver based on pressure is used in the finite element calculation. The initialization and calculation are started from the inlet. The SIMPLE method is used to solve the coupling of pressure and velocity. The time step is set as 0.1 s, while the number of time steps is 20 and the maximum number of iterations is 30. After calculation, the fluid movement time is 2 s, which is enough for single exhalation or single inhalation.

In the follow-up analysis, simulations are first taken out to compare the different effects of inhalation and exhalation

on the airway under the same conditions, and then, the pressure, velocity and path lines are compared to analyse the difference of airway flushing effect under different flow rates and the influence of different oxygen concentration on airway pressure and flow rate.

## 4. Results

*4.1. Comparison of Flow Field in Different Breathing States.* Through the comparison of inspiratory and expiratory parts, it can be seen that the parameters of flow field are quite different on inhaling and exhaling. Figure 4 shows the pressure contrast isoline of the airway at 70% oxygen ratio and 40 L/min flow rate. It can be seen that the pressure of 1.5 cm $H_2O$ can be reached on inhaling, and the minimum pressure is −1 cm $H_2O$. The highest pressure appears in the nasal cavity, especially at the top of the nasal cavity and the part near nasal concha, while the lowest pressure appears in the trachea. Generally speaking, the pressure from the nasal cavity to the trachea gradually decreases.

During exhalation, the highest pressure (about 2.7 cm $H_2O$) is produced at the end of the trachea, while the lowest pressure (about 1.7 cm $H_2O$) is produced at the place, where the nostrils connect with the outside world. However, the maximum pressure of expiration is still higher than that of inspiration. Simultaneously, the obvious irregular places, oropharynx and pharynx, have no pressure mutation. From the perspective of respiratory work, the pressure difference between the nasal cavity and trachea as well as the unchanged volume of the airway during inhalation provides part of the breathing work, reducing the extra work of the respiratory muscles of the lung and increasing patients' comfort.

Compared with the pressure, the flow rate shows a significant difference (shown as Figure 5). The maximum velocity (about 10 m/s) appears at the place near nostril. Except the velocity near the gas inlet and junction of larynx and trachea, velocity of all parts presents a low velocity state, most of which are between 0 and 1.5 m/s. During exhalation, the maximum flow rate also appears at gas vent (over 10 m/s). There is also a concentration of the flow velocity at the junction of the larynx and the trachea, and there is a larger velocity area in the trachea. That is because the gas coming out of the lung meets the narrow larynx when exhaling, which causes the velocity increase of this part. The excessive air oxygen mixture gas flows to the atmosphere through the space where the nostrils are not intubated, resulting in the high velocity of the place.

The strain rate can be multiplied by the viscosity to obtain the shear stress, which represents the force on airway wall. The strain rate is larger in the front of nasal cavity, oropharynx, and upper part of trachea, and the maximum value (about 13800) occurs at the junction of larynx and trachea, which is also prone to shear deformation. The maximum strain rate at the nasal cavity is 15000/*s*, which is generated in the front of the nasal cavity. The maximum stress at the larynx is 13500/*s* in the transition narrow area between the nasopharynx and oropharynx. The maximum strain rate of the trachea is 13000/*s* at the upper end of the trachea. In general, both inspiratory and expiratory strains in the front of the nasal cavity and the larynx should be paid more attention. Figure 6 also shows the strain rate of normal breathing on airway. It is found that the strain rate of HFNC at 40 L/min is increased compared with that of normal breathing. Although the maximum strain rate increase is not obvious, it is obvious that the range of larger strain region widens.

As shown in Figure 7, the left figure shows the path lines measured by particle ID during inhalation. In the process of inhalation, high flow gas is injected from nasal tube, and there are obvious gyrations in the vent and inferior nasal meatus, which is more due to the complexity of the air gap in the nasal cavity. Apart from the obvious turbulence at the junction of trachea and larynx, laminar flow features are more obvious in the trachea. The right figure shows the path lines during exhalation. In the process of exhalation, the gas from the trachea affects the airflow state of HFNC, making oxygen and air mixture unable to reach the nasopharynx. In addition, there is an obvious swirling turbulence in the nasal cavity, which finally flows out through the nostril. The gas also presents a similar laminar flow state in the trachea and an obvious turbulence in the nasal cavity. On exhaling, the oxygen and air mixture from the nasal tube forms turbulence in the nasal cavity, which then flows to the atmosphere from the nostril gap. Therefore, a certain amount of oxygen is maintained in the turbulent part, which dilutes carbon dioxide from the lungs and reduces the amount of carbon dioxide brought into the original nasal cavity during the next inhalation to achieve the purpose of flushing the dead zone of nasal cavity.

Figure 8 shows the time map of particles in the airway. When exhaling, the gas from the trachea conflicts with that from HFNC equipment, leading to a large convection in the front of the nasal cavity.

*4.2. Comparison of Flow Rate and Oxygen Concentration.* Figure 9 shows the maximum pressure comparison in different ventilation flow rates. Under the flow rate of 40 L/min, the pressure of a large area of upper respiratory tract is between 0.5 and 0.8 cm $H_2O$, and the smaller part reaches a pressure of 1.7 cm $H_2O$. The pressure of trachea (about −0.5 to 0.2 cm $H_2O$) is lower than that of upper respiratory tract. At 50 L/min and 60 L/min, the pressure of upper respiratory tract is also higher than that of the lower respiratory tract, whose maximum pressures are 2.0 cm $H_2O$ and 2.3 cm $H_2O$ and minimum pressures are −0.4 cm $H_2O$ and −0.1 cm $H_2O$, respectively. In addition, with the increase of ventilation flow, the maximum pressure and minimum pressure increased in varying degrees. This airway pressure difference increases with the increase of flow rate, which indicates that high flow rate gas can reduce the respiratory work required by respiratory muscles and increase the comfort of patients.

By comparing the airway velocity under different flow rates, the maximum velocity is generated at the contact part between the nasal cavity and the nasal tube, and the velocity is far lower than the ventilation velocity at the position with larger cross-sectional area. At same positions, the change

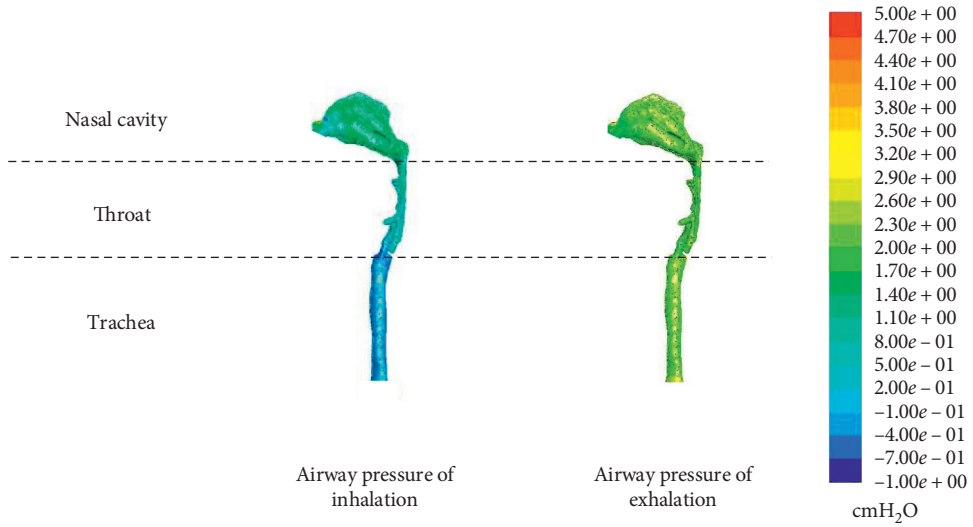



Figure 4: Comparison of pressure in different breathing states (cycle).

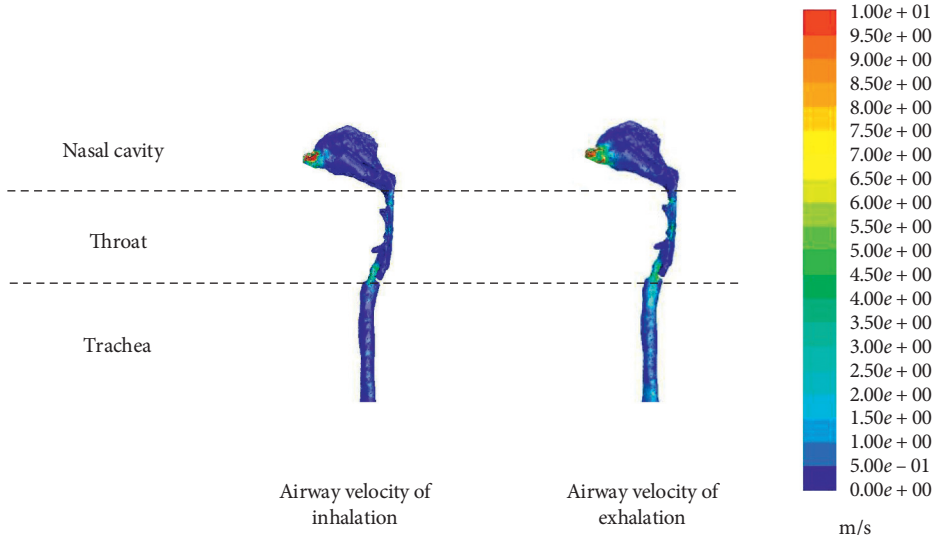



Figure 5: Comparison of velocity in different breathing states (cycle).

trend of the flow rate is the same. In addition, the second largest velocity is generated at the junction of the larynx and trachea, and the minimum is in the middle and rear part of nasal cavity. Figure 10 shows comparison of maximum velocity at different ventilation flow rates.

According to the pressure comparison of respiratory airway under different oxygen concentrations, it is found that the pressure of nasal cavity is greater than that of throat and trachea. Simultaneously, with the increase of oxygen concentration, the maximum pressure of each part decreases to varying degrees, which can be seen in Figure 11.

On choosing the median plane as the analysis plane to compare the oxygen concentration (as shown in Figure 12), it can be seen that the maximum velocity decreases with the increase of oxygen concentration and the area of high velocity area is also reduced, while the minimum flow rate is basically unchanged. This is because the viscosity of the gas mixture increases with the increase of oxygen concentration, and the friction force of intermolecular motion is large. When the initial velocity is the same, the instantaneous velocity at the same position will decrease with the increase of viscosity. The maximum velocity is produced at the junction of larynx and trachea, and there is no obvious change in other areas, indicating that the difference would be more obvious when the flow diameter is reduced.

## 5. Discussion

Based on the finite element analysis method, the cavity in the three-dimensional model of respiratory tract established by image scanning data conversion is taken as the research object. It is found that when inhaling, the greatest pressure is concentrated in the nasal cavity, and the highest velocity is concentrated in the vent and junction of the larynx and

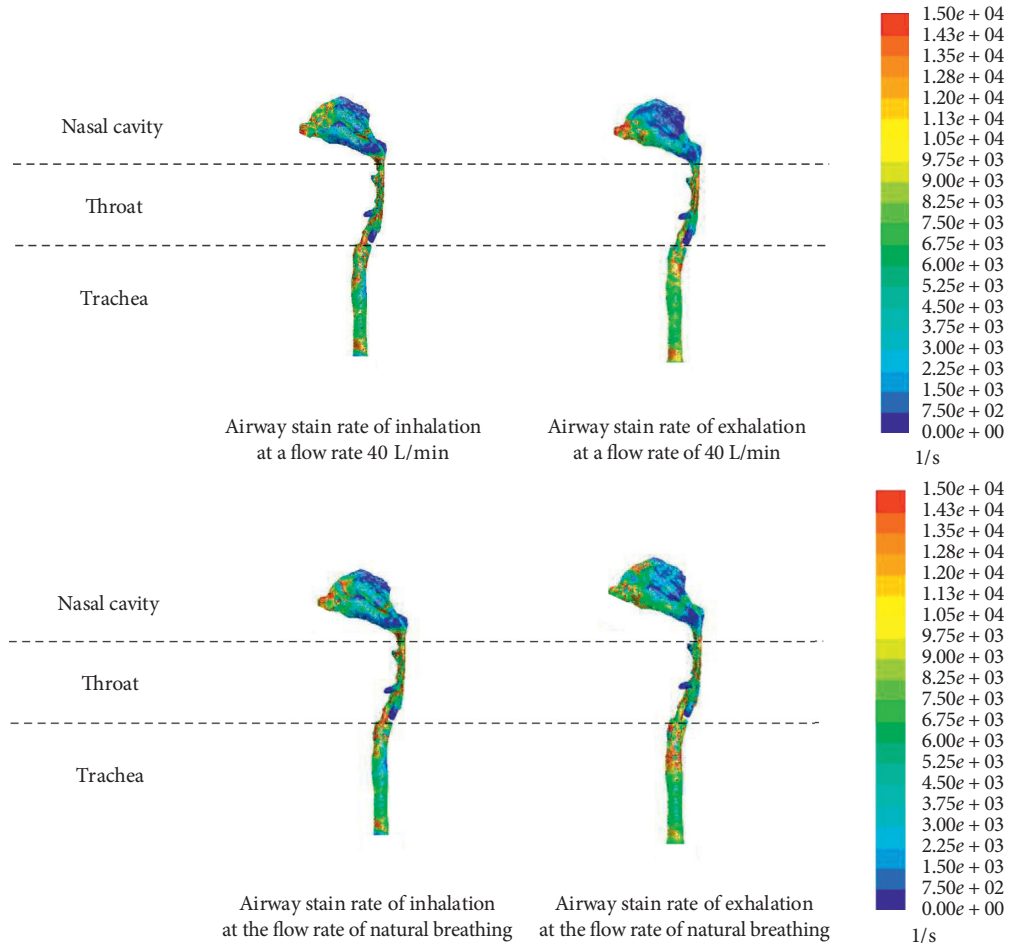



Figure 6: Comparison of stain rate in different breathing states (cycle).

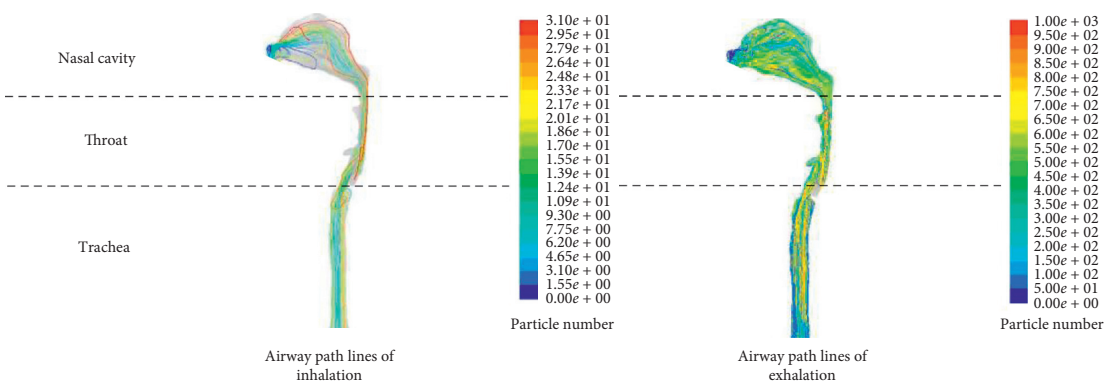



Figure 7: Comparison of path lines (measured by particle ID) in different breathing states.

trachea. The most easily damaged parts by shear stress are the front of the nasal cavity, near the vent, and junction of larynx and trachea. When exhaling, the greatest pressure is concentrated in the trachea, and the highest velocity is concentrated in the nasal vent. The most easily damaged parts by shear stress are the front of the trachea and the upper end of the trachea. In addition, the particle trace states and characteristics of different inlets are obtained by comparison. Then, different flow rates are compared, and it is concluded that even if the change of flow rate is not obvious, the increase of flow rate also leads to the increase of pressure. On comparing different oxygen concentrations, the flow rate decreases with the increase of the oxygen concentration.

In this study there are some assumptions. First, the flow of gas meets the simplified setting. The simplified design is that the fluid is in one-dimensional isentropic flow, and the physical properties are uniform in the inlet and outlet flow

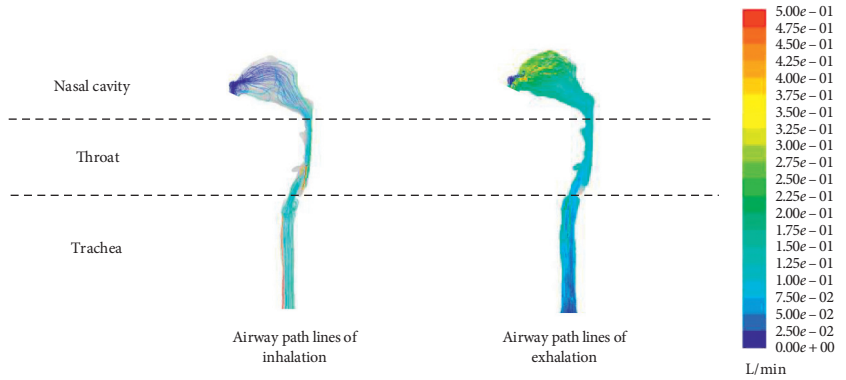



Figure 8: Comparison of path lines (measured by time) in different breathing states.

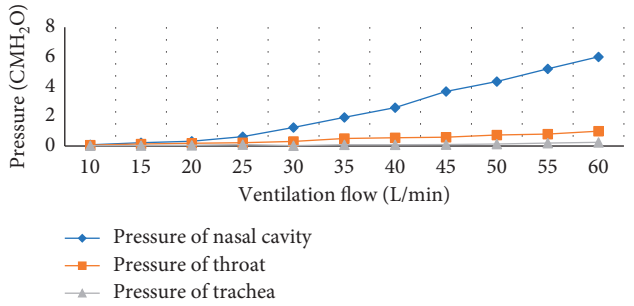



Figure 9: Comparison of maximum pressure in different ventilation flow rates.

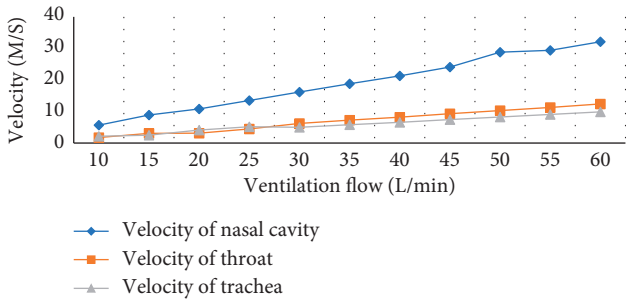



Figure 10: Comparison of maximum velocity in different ventilation flow rates.

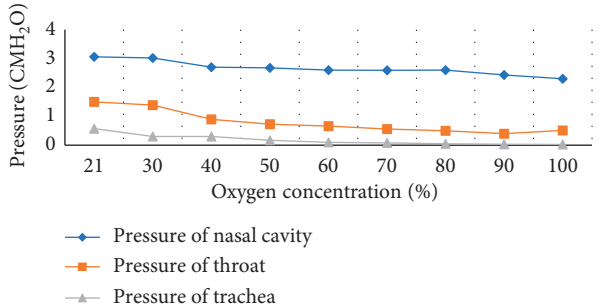



Figure 11: Comparison of maximum pressure comparison in different oxygen concentrations.

cross-section, the flow does not shrink at the nozzle, the gravity is ignored, and the specific heat of the fluid is constant. Second, the fluid-structure interaction between the gas and airway wall is not considered. Without considering the fluid-structure interaction, the elastic deformation caused by gas filling would not be considered, and the

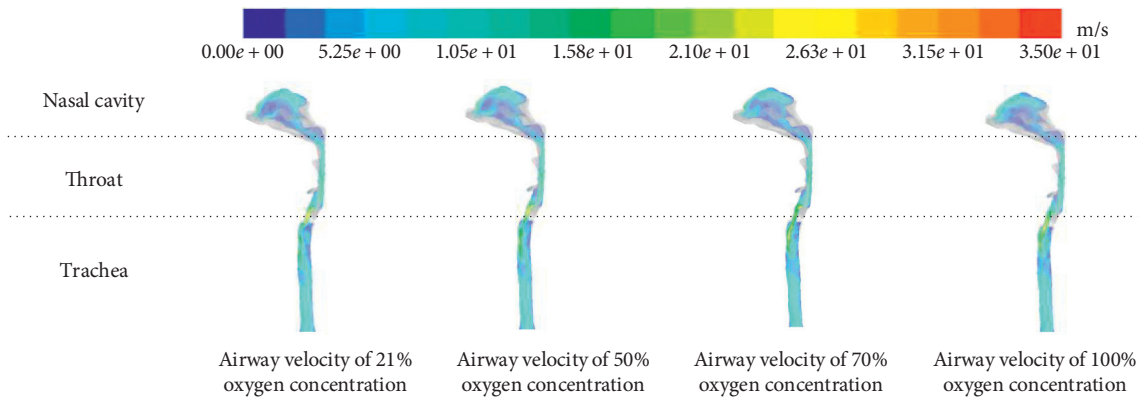



Figure 12: Comparison of velocity in different oxygen concentrations.

influence of actual elastic deformation on the airway would not be considered. The nasal cavity of human body is composed of bone and cartilage and the mucosa and skin covered on its surface. Some scholars regard the soft tissue around the upper respiratory tract as a linear elastic structure and simulate the airflow [24], which is widely used as a preliminary study on the interaction calculation of air flow and soft tissue [25]. However, this study focuses on the flow field state of the inner cavity, fully considering the internal velocity, pressure, strain rate, and path line. This paper also focuses on the possible pressure concentration, strain concentration position, internal pressure range, velocity range, and so on. Finally, the patient will not breathe through the mouth during ventilation.

In the acquisition of research objects, this study used Mimics software to convert the CT scanning data of airway into three-dimensional model, so that the position of ventilation port, nasal cavity, larynx and trachea can be clearly seen in the model. However, the obtained model is STL file, which is suitable for 3D printing directly but cannot be edited, so that it is impossible to extract the lumen directly. Therefore, it is necessary to convert it into a 3D model to the surface solid. SolidWorks software is used and the ideal effect is obtained. After that, the editable solid is obtained in the SIMENS NX software using the commands of combined surface, suture surface, and building solid. Finally, in the geometry module of ANSYS Workbench, the surfaces of inlet and outlet are generated, and the filling command is executed to generate the fluid part, which is the research object.

In terms of conclusion analysis, this study not only uses the overall contrast figure but also uses the longitudinal sectional contrast figure. However, the transverse sectional comparison is not selected because of its less distinctiveness. For example, when studying the velocity comparison of different oxygen concentrations, the maximum velocity at 0.01 m, 0.03 m, and 0.06 m away from the front end of nasal cavity changes slightly, so the middle longitudinal section is selected as the contrast section, which not only shows the whole character of the respiratory tract studied but also reflects the position of the maximum pressure.

The shortcomings of this study are as follows.

(1) The fluid-structure interaction of the oxygen air mixture flow on the respiratory tract is not considered, so the research can only focus on the flow field analysis but does not consider the shear damage of the gas flow on the trachea wall.

(2) This study is based on the finite element simulation. The mesh fineness can affect the simulation effect. It would be more visual and accurate if more details in the trachea could be analysed using a more refined mesh in the future.

(3) The RNG K-$\varepsilon$ model used in this paper has some advanced significance, but the complete flow field analysis needs to be verified by experiments. The follow-up study will be based on solving these shortcomings. It is hoped to obtain more accurate and realistic respiratory airway flow field under HFNC.

## 6. Conclusion

Based on the finite element analysis method, the three-dimensional model of trachea established by the conversion of image scanning data is taken as the research object, and the conclusion that the flow state presents turbulent state under different breathing states, flow rates, and oxygen concentrations is obtained. After that, RNG K-$\varepsilon$ model is used to solve the turbulence state of each position. The pressure, flow velocity, and strain rate simulation diagram of the airway under the finite element method are obtained.

(1) During inhalation, the greater pressure is concentrated in the nasal cavity, and the high flow velocity is concentrated at the junction of larynx and trachea. The parts easily damaged by shear stress are the front of nasal cavity, near the ventilation port, and junction of larynx and trachea. During exhalation, the greater pressure is concentrated in the trachea, the high flow rate is concentrated in the nasal vent and the front end of the trachea, and the severe shear stress is in the upper part of the trachea.

(2) The maximum movement time of particles in inhalation is less than that in exhalation. Oxygenated gases can reduce the dead space by increasing the oxygen content in nasal cavity.

(3) Flow rate and oxygen concentration affect the respiratory-related parameters. A large flow of gas can produce a large pressure difference, which can reduce inspiratory resistance and respiratory work. With the increase of oxygen concentration, the maximum flow rate decreases.

In future research, the shear damage of the air flow in the trachea wall will be considered, and a more precise grid will be used to analyse more details in the trachea intuitively and accurately.

## Data Availability

The data used to support the findings of this study are included within the article.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This research was funded by China Postdoctoral Science Foundation and the Open Research Project of the State Key Laboratory of Media Convergence and Communication (no. 2019M660392) and Communication University of China (no. SKLMCC2020KF002).

## References

[1] G. Y. Chang, C. A. Cox, and T. H. Shaffer, "Nasal cannula, CPAP, and high-flow nasal cannula: effect of flow on temperature, humidity, pressure, and resistance," *Biomedical Instrumentation & Technology*, vol. 45, no. 1, pp. 69–74, 2011.

[2] J. Q. Xu, L. X. Su, P. Yan et al., "Expert consensus on clinical standardized application of high-flow nasal cannula oxygen therapy in adults," *Chinese Medical Journal*, vol. 133, 2020.

[3] H. W. Lee, S. M. Choi, J. Lee et al., "Reduction of PaCO2 by high-flow nasal cannula in acute hypercapnic respiratory failure patients receiving conventional oxygen therapy," *Acute and Critical Care*, vol. 34, no. 3, pp. 202–211, 2019.

[4] T. Mauri, A. Galazzi, F. Binda et al., "Impact of flow and temperature on patient comfort during respiratory support by high-flow nasal cannula," *Critical Care*, vol. 22, no. 1, p. 120, 2018.

[5] O. Yu, R. Akimoto, H. Suzuki, N. Masaki, and K. Kawamae, "A high-flow nasal cannula system set at relatively low flow effectively washes out CO2 from the anatomical dead space of a respiratory-system model," *Korean Journal of Anesthesiology*, vol. 70, no. 1, 2017.

[6] O. Yu, R. Akimoto, H. Suzuki, M. Okada, M. Nakane, and K. Kawamae, "A high-flow nasal cannula system with relatively low flow effectively washes out CO2 from the anatomical dead space in a sophisticated respiratory model made by a 3D printer," *Intensive Care Medicine Experimental*, vol. 6, 2018.

[7] S. Nathalie, N. Charlene, V. Laurence et al., "Effects of nasal continuous positive airway pressure and high-flow nasal cannula on sucking, swallowing, and breathing during bottle-feeding in lambs," *Frontiers in Pediatrics*, vol. 5, p. 296, 2017.

[8] A. Motoyasu, K. Moriyama, H. Okano et al., "High-flow nasal cannula therapy reduced the respiratory rate and respiratory distress in a standard model simulator and in patients with hypoxemic respiratory failure," *Chronic Respiratory Disease*, vol. 16, 2019.

[9] A. L. Lampland, B. Plumm, P. A. Meyers, C. T. Worwa, and M. C. Mammel, "Observational study of humidified high-flow nasal cannula compared with nasal continuous positive airway pressure," *The Journal of Pediatrics*, vol. 154, no. 2, pp. 177–82, 2009.

[10] M. Lee, H. K. Ji, I. B. Jeong et al., "Protecting postextubation respiratory failure and reintubation by high-flow nasal cannula compared to low-flow oxygen system: single center retrospective study and literature review," *Acute and Critical Care*, vol. 34, no. 1, pp. 60–70, 2019.

[11] R. Miguel-Montanes, D. Hajage, J. Messika et al., "Use of high-flow nasal cannula oxygen therapy to prevent desaturation during tracheal intubation of intensive care patients with mild-to-moderate hypoxemia *," *Critical Care Medicine*, vol. 43, no. 3, pp. 574–583, 2015.

[12] J. Messika, K. Ben Ahmed, S. Gaudry et al., "Use of high-flow nasal cannula oxygen therapy in subjects with ards: a 1-year observational study," *Respiratory Care*, vol. 60, no. 2, pp. 162–169, 2015.

[13] M. Nishimura, "High-flow nasal cannula oxygen therapy in adults: physiological benefits, indication, clinical benefits, and adverse effects," *Respiratory Care*, vol. 61, no. 4, pp. 529–541, 2016.

[14] G. Spoletini, C. Mega, A. Khoja et al., "Better comfort and dyspnea scores with high-flow nasal cannula (HFNC) vs standard oxygen (SO) during breaks off noninvasive ventilation (NIV)," *European Respiratory Journal*, vol. 46, 2015.

[15] M. Mostafa-Gharehbaghi and H. Mojabi, "Comparing the effectiveness of nasal continuous positive airway pressure (NCPAP) and high flow nasal cannula (HFNC) in prevention of post extubation assisted ventilation," *Journal of Tropical Pediatrics*, vol. 64, 2015.

[16] L. J. Schlapbach, J. Schaefer, A.-M. Brady, S. Mayfield, and A. Schibler, "High-flow nasal cannula (HFNC) support in interhospital transport of critically ill children," *Intensive Care Medicine*, vol. 40, no. 4, pp. 592–599, 2014.

[17] K. Horsfield, G. Dart, D. E. Olson, G. F. Filley, and G. Cumming, "Models of the human bronchial tree," *Journal of Applied Physiology*, vol. 31, no. 2, pp. 207–217, 1971.

[18] H. Yeh and G. Schum, "Models of human lung airways and their application to inhaled particle deposition," *Bulletin of Mathematical Biology*, vol. 42, no. 3, pp. 461–480, 1980.

[19] C. Kleinstreuer and Z. Zhang, "Airflow and particle transport in the human respiratory system," *Annual Review of Fluid Mechanics*, vol. 42, no. 42, pp. 301–334, 2009.

[20] G. Mylavarapu, S. Murugappan, M. Mihaescu, M. Kalra, S. Khosla, and E. Gutmark, "Validation of computational fluid dynamics methodology used for human upper airway flow simulations," *Journal of Biomechanics*, vol. 42, no. 10, pp. 1553–1559, 2009.

[21] B. E. Launder and D. B. Spalding, *Lectures in MathemArticle Models of Turbulence*, pp. 1–169, London Academic Press, London, UK, 1972.

[22] V. Yakhot and S. A. Orszag, "Renormalization-group Analysis of turbulence," *Physical Review Letters*, vol. 57, no. 14, pp. 1722–1724, 1986.

[23] S.-J. Jeong, W.-S. Kim, and S.-J. Sung, "Numerical investigation on the flow characteristics and aerodynamic force of the upper airway of patient with obstructive sleep apnea using computational fluid dynamics," *Medical Engineering & Physics*, vol. 29, no. 6, pp. 637–651, 2007.

[24] F. Chouly, A. Van Hirtum, P.-Y. Lagrée, X. Pelorson, and Y. Payan, "Numerical and experimental study of expiratory flow in the case of major upper airway obstructions with fluid-structure interaction," *Journal of Fluids and Structures*, vol. 24, no. 2, pp. 250–269, 2008.

[25] C. Xu, M. J. Brennick, L. Dougherty, and D. M. Wootton, "Modeling upper airway collapse by a finite element model with regional tissue properties," *Medical Engineering & Physics*, vol. 31, no. 10, pp. 1343–1348, 2009.