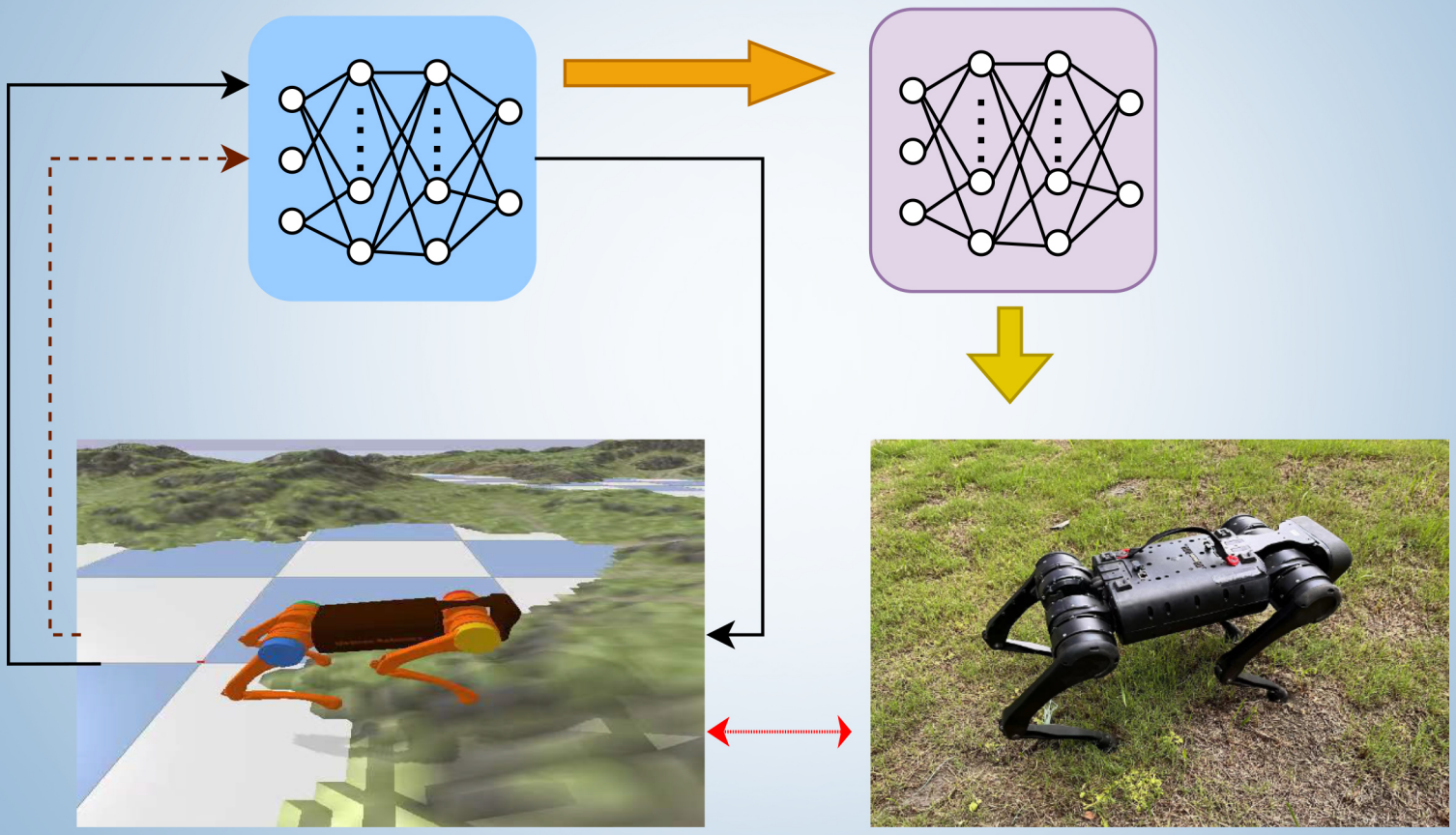
Volume 2, Issue 3, September 2022

INTELLIGENCE & ROBOTICS



Deep reinforcement learning for real-world quadrupedal locomotion: a comprehensive review

Hongyin Zhang, Li He, Donglin Wan

Open Access
ISSN 2770-3541 (Online)
www.intellrobot.com

Editor-in-Chief

Simon X. Yang

Prof. Simon X. Yang is currently the Head of the Advanced Robotics and Intelligent Systems Laboratory at the University of Guelph. His research interests include artificial intelligent, robotics, sensors and multi-sensor fusion, wireless sensor networks, control systems, bio-inspired intelligence, machine learning, neural networks, fuzzy systems, and computational neuroscience.

Our Features

(1) Gold Open Access
(2) Strong Editorial Board
(3) Rigorous Peer-review
(4) Free English Ianguage Editing Service
(5) Online First Once Accepted
(6) Free Publication Before 31 Dec 2024
(7) Wide Promotion (Twitter\LinkedIn\WeChat\Facebook)

Editorial Board

- 1 Editor-in-Chief
- 2 Executive Editor
- 5 Advisory Editorial Members
- 38 Associate Editors
- 21 Youth Editorial Board Members

Scope

Top-quality unpublished original technical and non-technical application-focused articles are welcome from intelligence and robotics, particularly on the inter-disciplinary areas of intelligence and robotics, including but not limited to the following areas:

- biological, bio-inspired, and artificial intelligence;
- neural networks, fuzzy systems, and evolutionary algorithms;
- sensing, multi-sensor fusion, localization, data analysis, modeling, planning, and control for various mobile, aerial, and underwater robotic systems;
- robot cooperation, teleoperation, and human-machine interactions.
- development and maintenance of real-world intelligent and robotic systems by multidisciplinary teams of scientists and engineers.

Journal Home
https://intellrobot.com/

Submission Link
https://oaemesas.com/login?JournalId=ir

EDITORIAL BOARD

Editor-in-Chief

Simon X. Yang
University of Guelph, Canada

Executive Editor

Lei Lei
University of New Brunswick, Canada

Hao Zhang
Tongji University, China

Advisory Board Members

Tianyou Chai
Northeastern University, China

Clarence W. De Silva
University of British Columbia, Canada

Toshio Fukuda
Nagoya University, Japan

Aike Guo
University of Chinese Academy of Sciences, China

Deyi Li
Chinese Academy of Engineering, China

Associate Editors

Zhuming Bi
Purdue University Fort Wayne, USA

Mohammad Biglarbegian
University of Guelph, Canada

Hicham Chaoui
Carleton University, Canada

Chao-Yang Chen
Hunan University of Science and Technology, Xiangtan, China

Guang Chen
Tongji University, China

Hongtian Chen
University of Alberta, Canada

Chao Cheng
Changchun University of Technology, China

Abdelghani Chibani
University of Paris-Est Creteil (UPEC), France

Carlos Renato Lisboa Francês
Federal University of Para, Brazil

Paulo Gonçalves
Polytechnic Institute of Castelo Branco, Portugal

Nallappan Gunasekaran
Toyoto Technological Institute, Japan

Zeng-Guang Hou
Institute of Automation, Chinese Academy of Sciences, China

Shaidah Jusoh
Princess Sumaya University for Technology, Jordan

Fakhri Karray
University of Waterloo, Canada

Zuojin Li
Chongqing University of Science and Technology,China

Ming Liu
The Hong Kong University of Science and Technology, China

Qiang Liu
Northeastern University, China

Chaomin Luo
Mississippi State University, USA

Jianjun Ni
Hohai University, China

Chen Peng
Shanghai University, China

Hong Qu
University of Electronic Science and Technology of China, China

Tao Ren
Chengdu University of Technology, China

Gerasimos Rigatos
Industrial Systems Institute, Greece

Ricardo Sanz
Universidad Politécnica de Madrid, Spain

Jinhua She
Tokyo University of Technology, Japan

Farhad Soleimanian Gharehchopogh
Islamic Azad University, Iran

Mariacarla Staffa
University of Naples Parthenope, Italy

Jindong Tan
University of Tennessee, USA

Mien Van
Queen's University Belfast, UK

Ying Wang
Kennesaw State University, USA

Wai Lok Woo
Newcastle University, Newcastle, UK

Gui-Song Xia
Wuhan University, China

Xin Xu
National University of Defense Technology, China

Huaicheng Yan
East China University of Science and Technology, China

Wen Yu
National Polytechnic Institute, Mexico

Anmin Zhu
Shenzhen University, China

Daqi Zhu
Shanghai Maritime University, China

Youth Editorial Board Members

Laith Abualigah
Amman Arab University, Jordan

Sawal Hamid Md Ali
Universiti Kebangsaan Malaysia, Malaysia

Yiyang Chen
Soochow University, China

Changxin Gao
Huazhong University of Science and Technology, China

Jianye Hao
Tianjin University, China

Manju Khari
Jawaharlal Nehru University, India

Dong Liu
Dalian University of Technology, Dalian, Liaoning, China

Anh-Tu Nguyen
Université Polytechnique Hauts-de-France, France

EDITORIAL BOARD

Farhad Pourpanah
Shenzhen University, China

Sangram Redkar
Arizona State University, USA

Bing Sun
Shanghai Maritime University, China

Xiaoqiang Sun
Jiangsu University, China

Yuxiang Sun
The Hong Kong Polytechnic University, China

Di Wang
Chongqing Jiaotong University, China

Donglin Wang
Westlake University, China

Zhongkui Wang
Ritsumeikan University, Japan

Guanglei Wu
Dalian University of Technology, China

Shuiqing Xu
Hefei University of Technology, China

Yu Xue
Nanjing University of Information Science and Technology, China

Peng Yao
Ocean University of China, China

Guoxian Yu
Shandong University, China

Zhiwei Yu
Nanjing University of Aeronautics and Astronautics, China

Zhiyao Zhao
Beijing Technology and Business University, China

GENERAL INFORMATION

About the Journal

Intelligence & Robotics (*IR*), ISSN 2770-3541 (Online), publishes top-quality unpublished original technical and non-technical application-focused articles on intelligence and robotics, particularly on the interdisciplinary areas of intelligence and robotics. The Journal seeks to publish articles that deal with the theory, design, and applications of intelligence and robotics, ranging from software to hardware. The scope of the Journal includes, but is not limited to, biological, bio-inspired, and artificial intelligence; neural networks, fuzzy systems, and evolutionary algorithms; sensing, multi-sensor fusion, localization, data analysis, modeling, planning, and control for various mobile, aerial, and underwater robotic systems; and robot cooperation, teleoperation and human-machine interactions. The Journal would be interested in distributing development and maintenance of real-world intelligent and robotic systems by multidisciplinary teams of scientists and engineers.

Information for Authors

Manuscripts should be prepared in accordance with Author Instructions.
Please check https://intellrobot.com/pages/view/author_instructions for details.
All manuscripts should be submitted online at https://oaemesas.com/login?JournalId=ir.

Copyright

Articles in *IR* are published under a Creative Commons Attribution 4.0 International (CC BY 4.0). The CC BY 4.0 allows for maximum dissemination and re-use of open access materials and is preferred by many research funding bodies. Under this license users are free to share (copy, distribute and transmit) and remix (adapt) the contribution for any purposes, even commercially, provided that the users appropriately acknowledge the original authors and the source.
Copyright is reserved by © The Author(s) 2022.

Permissions

For information on how to request permissions to reproduce articles/information from this journal, please visit www.intellrobot.com.

Disclaimer

The information and opinions presented in the journal reflect the views of the authors and not of the journal or its Editorial Board or the Publisher. Publication does not constitute endorsement by the journal. Neither the *IR* nor its publishers nor anyone else involved in creating, producing or delivering the *IR* or the materials contained therein, assumes any liability or responsibility for the accuracy, completeness, or usefulness of any information provided in the *IR*, nor shall they be liable for any direct, indirect, incidental, special, consequential or punitive damages arising out of the use of the *IR*. *IR*, nor its publishers, nor any other party involved in the preparation of material contained in the *IR* represents or warrants that the information contained herein is in every respect accurate or complete, and they are not responsible for any errors or omissions or for the results obtained from the use of such material. Readers are encouraged to confirm the information contained herein with other sources.

Published by

OAE Publishing Inc.
245 E Main Street Ste 107, Alhambra CA 91801, USA
Website: www.oaepublish.com

Contacts

E-mail: editorial@intellrobot.com
Website: www.intellrobot.com

Intelligence & Robotics

CONTENTS

Volume 2, Issue 3

Zhu *et al*. *Intell Robot* 2022;2(3):200-222
DOI: 10.20517/ir.2022.13

Intelligence & Robotics

Review

Open Access

Check for updates

Motion planning and tracking control of unmanned underwater vehicles: technologies, challenges and prospects

Danjie Zhu, Tao Yan, Simon X. Yang

School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada.

Correspondence to: Prof. Simon X. Yang, Advanced Robotics & Intelligent Systems (ARIS) Laboratory, School of Engineering, University of Guelph, Guelph, ON N1G 2W1, Canada. E-mail: syang@uoguelph.ca

How to cite this article: Zhu D, Yan T, Yang SX. Motion planning and tracking control of unmanned underwater vehicles: technologies, challenges and prospects. *Intell Robot* 2022;2(2):xx. http://dx.doi.org/10.20517/ir.2022.13

Received: 7 May 2022 **First Decision:** 27 May 2022 **Revised:** 15 Jun 2022 **Accepted:** 30 Jun 2022 **Published:** 8 Jul 2022

Academic Editor: Jianjun Ni, Mien Van **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

The motion planning and tracking control techniques of unmanned underwater vehicles (UUV) are fundamentally significant for efficient and robust UUV navigation, which is crucial for underwater rescue, facility maintenance, marine resource exploration, aquatic recreation, etc. Studies on UUV motion planning and tracking control have been growing rapidly worldwide, which are usually sorted into the following topics: task assignment of the multi-UUV system, UUV path planning, and UUV trajectory tracking. This paper provides a comprehensive review of conventional and intelligent technologies for motion planning and tracking control of UUVs. Analysis of the benefits and drawbacks of these various methodologies in literature is presented. In addition, the challenges and prospects of UUV motion planning and tracking control are provided as possible developments for future research.

Keywords: Unmanned underwater vehicles, motion planning, path planning, task assignment, tracking control

1. INTRODUCTION

More attention has been concentrated on underwater navigation this century due to the abundant resources buried in the deep-sea area, such as biological, mineral, and space resources[1]. Therefore, underwater vehicles (UV) have been applied due to their adaptiveness and safety when exploring undersea environments. The vehicle can tackle the problems of hardly -predictable obstacles, current flow, and hydraulic pressure as well as provide longer operating time and more functions than human divers[2,3].

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

Meanwhile, unmanned underwater vehicles (UUVs) are being developed, which can be divided into remote-operated vehicles (ROVs) and autonomous underwater vehicles (AUVs). The motion planning and tracking control of UUVs are assumed to be significant technologies for accomplishing efficient underwater navigation with a guaranteed response time and without direct contact with marine dangers. Research on the underwater motion planning and tracking control of UUVs originated decades ago and has been under the spotlight in recent years. The history of UUVs can be traced back to the mid-20th century when an unmanned vehicle was invented by the US Navy to recover a hydrogen bomb lost off the coast of Spain[4]. In 2009, the success of finding cracked pieces of Air France 447 realized by a UUV verified the vehicle's promising application in underwater navigation. In 2014, the search for flight MH370 also brought out the growing attention on demanding underwater navigation, which highly depends on UUV motion planning and tracking control technologies[5]. Nowadays, detecting deep into the ocean area for digging more available resources such as undersea oil development also requires the continuous progress of the UUV motion planning and tracking control[6].

According to the statistics collected by Web of Science, the number of organizations that devote efforts to underwater vehicle research has impressively increased in the past years. This trend corresponds to the growing demand for underwater navigation worldwide, where studies of underwater motion planning and tracking control are regarded as hot topics among UUV research projects. Underwater motion planning and tracking control form the most crucial part of underwater navigation. UUV motion planning is established on conventional or intelligent technologies of vehicle posture planning, task assignment, and path planning, while UUV tracking control is mainly about vehicle trajectory tracking. The paper focuses on UUV task assignment, path planning, and trajectory tracking controls. Task assignment is designed for the multi-UUV system, where multiple vehicles are arranged simultaneously to achieve the most efficient collective navigating plan without mutual interference; the path planning of UUVs aims at giving the optimal instruction to the vehicle for arriving at the target, which can largely save the time and reduce the energy consumption; and trajectory tracking study of UUVs guarantees the robustness and manner of the vehicle operation in practical cases. However, studies related to underwater motion planning and tracking control have not been thoroughly investigated due to the complexity of the ocean environment and the vehicle system[7,8]. In addition, to overcome the difficulty of accomplishing complex underwater operations, the multi-UUV system, which refers to the system of multiple UUVs and multiple targets, has pulled great attention owing to its high parallelism, robustness, and efficiency[9–11].

Motivated by the goal of realizing efficient and robust UUV navigation in ocean environments, studies related to the UUV motion planning and tracking controls should be systematically surveyed and discussed to address their potential in applications. Meanwhile, progress in this field can be promoted by analyzing the deficiency and possible developments of the relevant technologies. Therefore, the contribution of the paper is to collect and analyze technologies that have been and can be applied to the motion planning and tracking control of UUVs. The benefits and drawbacks of these technologies are discussed, and challenges and prospects are derived based on the gap in the literature. These analyses and conclusions are supposed to provide a brief overview of studies that can be developed on certain issues for researchers at the entry level in the field of UUV motion planning and tracking control.

In this paper, a brief review of the technologies regarding the UUV motion planning and tracking control is proposed. The review investigates the current development of the motion planning and tracking control achieved by UUVs and then derives the challenges as well as possible prospects of the study. The introduction is given in Section 1. In Sections 2 and 3, the current research status is described. Methodologies of UUV motion planning are organized into divisions of task assignment and path planning. In the tracking control section, the trajectory tracking methods of the UUV are surveyed. In Section 4, challenges and possible prospects are concluded and discussed.

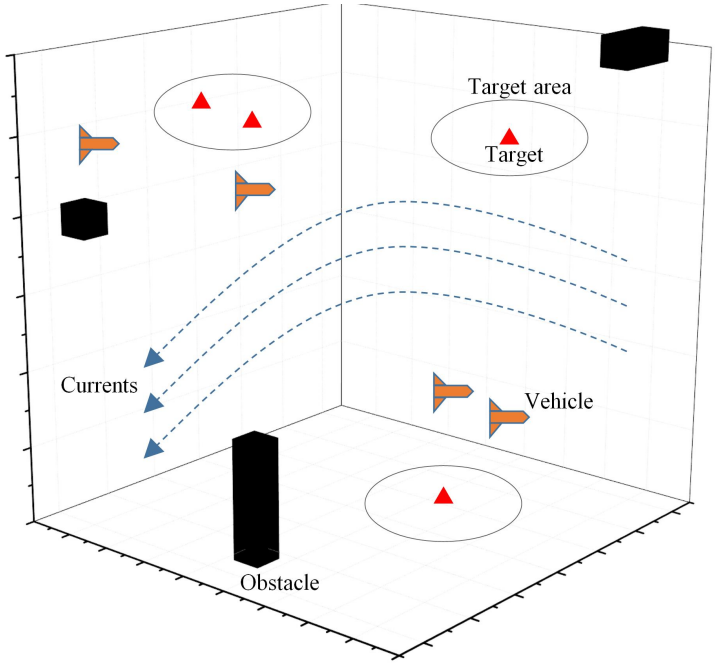


Figure 1. The underwater motion planning scenario of the UUV.

2. TECHNOLOGIES OF UUV MOTION PLANNING

In this section, technologies for motion planning of UUVs are presented. Motion planning of UUVs can be mainly categorized into steps of task assignment and path planning, where the path planning is usually split into point-to-point path planning and full-coverage path planning.

Underwater motion planning is the crucial part that decides the efficiency of a UUV navigation. The optimal vehicle motion has to be addressed in the requirement of the shortest total distance and time to arrive at the target. As shown by the underwater motion planning scenario in Figure 1, under the effect of ocean currents and obstacles, for the multi-UUV system, the optimal task assignment between multiple vehicles (in orange) upon multiple targets (in the red triangle) is considered as the preparation for assessing satisfactory planned paths. For the UUV path planning, the point-to-point path planning decides the initial navigation path from the vehicle to the target, while the full-coverage path planning instructs the vehicle's traversing operation after arriving at the target area (area within the black circle).

2.1. Task assignment of Multi-UUV system

Originated from the last century, strategies applied to the task assignment of the multi-UUV system are mostly realized by directly imitating animal behaviors. These assignments are designed through sensor-collected information, and the vehicle tasks are arranged referring to actual creature grouping behaviors[12,13]. Mataric *et al.* proposed a task assignment algorithm that imitates the animal grouping behaviors such as swarming and distributing[14]. Parkers established a distributed system that divides the assignment into smaller computing sections based on vehicle behaviors[12]. Miyata developed a behavior-based algorithm that independently assigned the task for vehicles based on the time priority[15]. These studies verify the directness, simple operating procedure, and no delays of behavior-based algorithms. However, they stay at the low administrative levels of imitation, which are short of self-regulation/optimization, and the unsatisfactory collaboration leads to the inefficiency of the algorithm and the requirement of intelligent task assignment methods.

Agent-based algorithms have been commonly applied to the task assignment of the multi-vehicle system[16–18].

In the agent-based task assignment algorithm, the whole system is assumed to be an economy entity, while each vehicle works as an agent. The agent-based algorithms are regarded as decentralized approaches, as each vehicle agent is supposed to know its requirement and limitation, and the final solution is deduced based on the balance between them. The task assigned to each agent is balanced after the repeated computation and comparison of the cost to their targets; therefore, the minor consumption and the largest profit for the whole entity can be obtained at the end[19]. The agent-based algorithms such as the auction algorithm resolve the task assignment problem of known targets efficiently; however, they do not work well in the vehicle assignment problem of unknown targets[20,21]. Yao applied the biased min-consensus (BMC) method, which introduces the edge weight into the standard min-consensus protocol. Yao achieved the path planning of simultaneous arrival for all UUVs through this agent-based task assignment algorithm, yet the situation of unknown targets is still not developed[22,23].

Intelligent methods such as swarm intelligence, genetic algorithm (GA), and neural network (NN) have been tested in solving the problem of the multi-vehicle task assignment[24–26]. These intelligent methods find the best task assignment solution by the objective function established on the total searching length and the heuristic cost through iteration algorithms. In recent years, the self-organizing map (SOM), an NN-based algorithm, was applied to the task assignment problem of a multi-vehicle system due to the competitiveness and self-improving features of the neural network[27]. The SOM-based task assignment algorithm guarantees that each vehicle in the multi-vehicle system can navigate along the shortest path to their target while maintaining the shortest total navigation cost for the whole system, whose structure is shown in Figure 2. The target locations serve as the inputs while the vehicle positions and paths are the outputs of the network, and the network is updated by the weights between layers that are deduced based on distances between targets and vehicles[28]. The turning direction angle and turning radius of the vehicle are then involved on the basis of the SOM method due to the vehicle's practical requirement of reducing the energy cost by reaching the target in a smooth curve in the task assignment problem[29,30].

However, the task assignment algorithms considering the underwater environment are still not thoroughly investigated due to the complex environmental factors and the nonlinear UUV system. Considering the complexity of the underwater environment such as the currents effect, Chow proposed an improved K-means algorithm to simultaneously resolve the task assignment and path planning problems for the multi-UUV system under the static ocean currents effect, where the vehicle successfully reached the target along smooth curves on the basis of optimal task assignment[31]. Nevertheless, the method does not work well for moving targets, and it lacks the discussion of applications under the 3D static ocean currents effect as well as the dynamic currents condition. Zhu *et al.* introduced SOM into the multi-UUV system and combined SOM with a velocity synthesis algorithm; hence, the task assignment and path planning problem for the multi-UUV system under time-varying ocean currents when chasing both static or dynamic targets could be addressed, which resolved the issues that existed in Chow's study. However, neither SOM-based methods could realize satisfactory collision avoidance[32].

Methods that have been applied to the task assignment of the multi-UUV system are listed in Table 1. Details of various intelligent methods for task assignment of UUVs can be found in Section 2.1. Gaps are still left for relative studies, which can be mainly concluded into two problems. The first problem is the difference among heterogeneous UUVs. They have different model parameters, navigating velocities or safe distances such that the assignment of parameters for every single UUV is not consistent in the practical application. The other problem is the effect of the underwater environmental factors such as the obstacles and the fluid effect, which may produce inevitable deviations or too many dynamic requirements for vehicles in the task assignment.

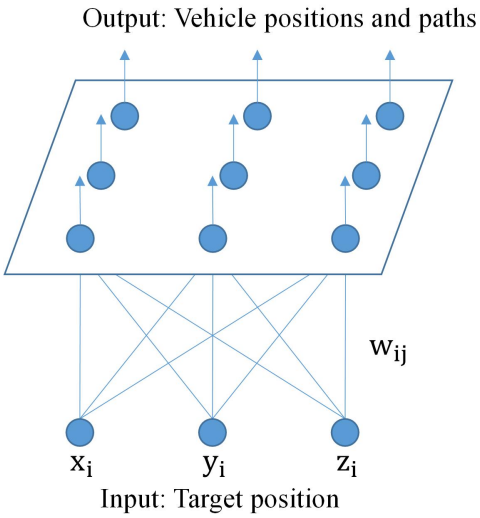


Figure 2. Structure of the SOM algorithm.

Table 1 Algorithms for task assignment of multi-vehicle system

Algorithms	Logic	Benefits	Drawbacks
Behavior imitation algorithms [12-15]	Simple imitation of the animal (including human) grouping behaviors such as swarming and distributing behaviors	(1) Easy to implement (2) React without lags	(1) Low efficiency (2) Cannot regulate themselves (3) Difficult to optimize
Agent-based algorithms [16-23]	(1) Assume the whole system as an economy entity while each vehicle works as an agent (2) Assign the task to each agent in the goal of gaining lowest cost for the whole entity	(1) Easy to implement (2) Satisfactory efficiency when resolving problems of known targets	Do not work well in the task assignment of unknown targets
Intelligent algorithms [24-30] (GA- or NN-based)	(1) Regard the task assignment as a search optimization problem (2) Take the searching distance as the objective function (3) Optimize through iterations	Outstanding adaptiveness due to consideration of the UUV system or environmental factors in the objective function	(1) Unsatisfactory real-time reaction owing to the computation complexity (2) Local minimum

2.2. Path planning of UUV

In this section, current methodologies developed for the path planning of the UUV system under different application cases are presented and concluded, divided into subsections on point-to-point path planning and full-coverage path planning.

2.2.1. Point-to-Point path planning

After completion of the task assignment, the UUV is required to navigate to the supposed destination position from its current position with: (1) an optimized path of shortest distance; and (2) avoidance of obstacles, which is described as the point-to-point path planning problem. Conventional map building methods such as grid-based modeling and topological approaches are used in the point-to-point path planning. Nowadays, typical methods that are applied in the UUV point-to-point path planning also include artificial potential field methods and a wide range of intelligent path planning algorithms.

Map building Method Map building methods plan the path by mapping the vehicle's surrounding area and then deriving the optimal solution accordingly. Based on the area information collected by the vehicle sensors such as the obstacle occupied status, different methodologies of mapping these areas can be addressed and deduce an efficient path solution accordingly. The fundamental part of map building methods such as mapping the vehicle searching area usually serves as the basis of most path planning algorithms, such as intelligent

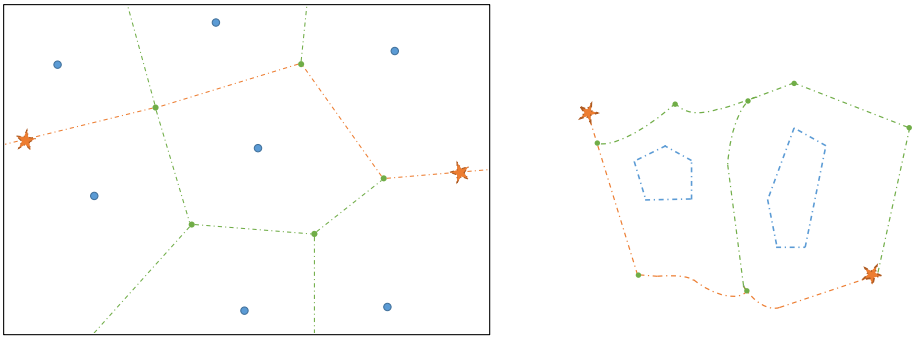


Figure 3. Typical Voronoi diagrams with the indicated graph branch nodes and optimal point-connected path: (A) obstacles in points; and (B) obstacles in convex polygons. Orange stars: Starts and ends of the path.

algorithms. However, in this section, the map building methods are limited to those that directly deduce on the map form without the combination of complex strategies such as self-regulation or self-evolution.

One typical map building method is called the visibility graph approach, where the graph is established on the connection of the vehicle, polygonal obstacle vertex, and the destination without crossing the obstacles[33]. The optimal path is determined by finding the route between the origin point and the destination point that has the shortest distance. The visibility graph approach derives the shortest path, yet it consumes long searching time and lacks flexibility, as the graph has to be reconstructed once the environmental information changes, such as the destination position or the obstacle shapes. Moreover, the visibility graph approach does not work for circular obstacles. The tangent graph method gives a more efficient path planning solution of shorter distance by controlling the vehicle to navigate along the tangent lines of obstacles[34]. However, the vehicle needs to approach the obstacle as close as possible when navigating along the tangent lines such that collisions might be produced in practical cases. The Voronoi diagram method resolves the collision problem through the combination of lines and parabolas, as shown in Figure 3, where the line is defined by the vertex of obstacles, while the parabola is defined by a vertex and a sideline of obstacle[35,36].

Grid-based path planning methods are also the widely used type of map building method. They decompose the surrounding area into nonoverlapping but connected cells, and then the optimal path is addressed between the origin and the destination cells without collisions. Dijkstra algorithm is one of the earliest grid-based path planning methods where a global search on all possible path solutions is required such that large computation is inevitable[37]. Therefore, A* algorithm is proposed with the advancement of adding the heuristic cost to reduce the searching space[38]. However, typical underwater disturbances such as the effect of currents might bring inevitable influence on UUV path planning; hence, the traditional grid-based path planning methods that need map of high accuracy and consistency are not appropriate to the UUV system[39,40].

Artificial Potential Field Method The artificial potential field (APF) method is established on a virtual artificial potential predefined field. The proposed destination is determined as the object that has the attraction to the vehicle, while the obstacles are regarded as the objects that generate repulsive force to the vehicle[41]. All the attractive and repulsive forces are quantified and presented in the form of gravity, where the positive gravity is correlated with the distance between the vehicle position and the destination, and negative gravity is performed within the domain of the obstacles. As the vehicle is closer to the destination, the gravity decreases until it reaches the destination. Deduced by the negative gradient of respective fields, the attractive force $\mathbf{F}_a$ and the repulsive force $\mathbf{F}_r$ are given by

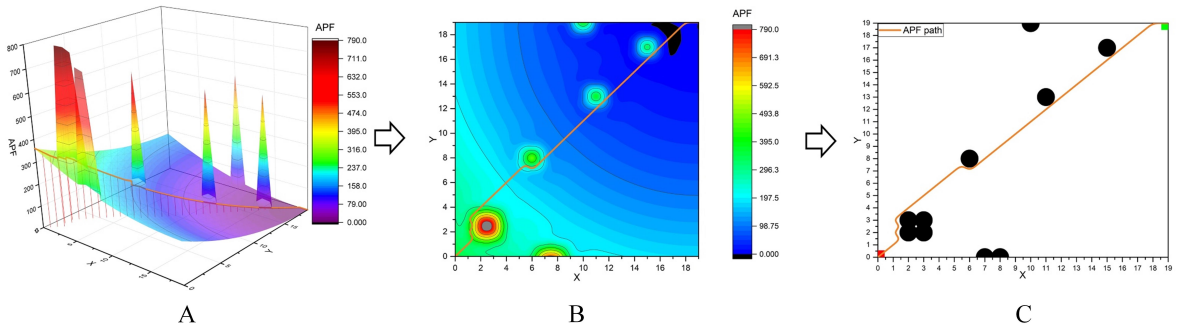


Figure 4. Path derived by the APF method on the 2D modeling map: (A) path planning based on the distribution of the APF on the map; (B) APF path on the contour map; and (C) final APF path presented on the 2D modeling map [42].

$$\mathbf{F}_a(x) = -\nabla \mathbf{U}_a(x) = k_a \rho(x, x_d) \tag{1}$$

$$\mathbf{F}_r(x) = -\nabla \mathbf{U}_r(x) = \begin{cases} k_r(\frac{1}{\rho(x,x_o)} - \frac{1}{\rho_o})\frac{1}{\rho^2(x,x_o)}\nabla\rho(x,x_o), \rho(x,x_d) < \rho_o \\ 0, \rho(x,x_d) \geq \rho_o \end{cases} \tag{2}$$

where $-\nabla \mathbf{U}_a$ represents the negative gradient of the attractive field; $-\nabla \mathbf{U}_r$ represents the negative gradient of the repulsive field; k_a is the coefficient for attraction; $\rho(x, x_d)$ represents the distance between the current position x and the destination position x_d; k_r is the repulsion coefficient; and $\rho(x, x_o)$ represents the distance between the current position to the obstacle position x_o and ρ_o is the radius of the obstacle.

Therefore, the destination has the lowest gravity field but the highest gravity force for attraction, while the gravity field for the obstacles performs higher such that the vehicle can flow along the gravity field descending route to complete the optimal path planning, as the path deduced from the point in Figure 4A to the one in Figure 4C.

The APF reduces the calculation complexity as well as performs outstanding real-time reactions, which is widely applied in the area of vehicle path planning. The virtual gravitational potential field realizes a fast calculation of the most optimal path to the target without collisions for the vehicle, by following the guidance of resultant forces given by the pre-designed attraction and repulsion [43]. Zhou *et al.* improved the APF method with a particle swarm algorithm to increase the pathfinding efficiency for tangent navigating robots [44]. Lin *et al.* designed a subgoal algorithm for the APF such that the path planning of the unmanned vehicle can overcome the local minimum and track the most optimal path [45]. The decision tree was added to the APF to form the efficient path planning algorithm without local minimum and collisions for vehicles [46]. Regarding the environmental factors, the effect of ocean currents was then involved in the path planning of the UUV while using the APF method [42].

However, most of the APF research do not involve environmental disturbance in the design, thus affecting the practical application of the APF. Moreover, the APF method for vehicle path planning often deduces the problem of local minimum, where the vehicle might stick at halfway instead of reaching the target position due to the larger resultant effect produced by the local minimum point [47]. The large computation complexity caused by the increasing obstacle numbers also affects the planning efficiency of the APF method.

Intelligent Path Planning Method More and more artificial intelligence methods have been applied in the studies of UUV path planning in recent years, covering the genetic algorithm, swarm intelligence, fuzzy logic, and

neural network algorithm.

The genetic algorithm (GA) and ant colony algorithm (ACO) were widely used in the early times for underwater path planning. The GA method imitates the natural selection and evolution procedure to provide the optimal solution through iterations, which has been involved in the path planning and obstacle avoidance under the underwater environment of dynamic currents effect[48,49]. The ACO method belongs to the swarm intelligence algorithm, where it is designed based on the swarm behavior of ant groups while chasing food, and the ant behavior-based intelligent method has been proved to work well in the UUV global path planning[50,51]. The swarm intelligence methods have been broadly applied in UUV path planning in recent years due to their simple implementation, fast convergence speed, and satisfactory robustness when modeling based on different swarming animal groups[52]. The swarm intelligence algorithms provide outstanding performance in the path planning of UUVs, yet the local minimum problem can be produced by this intelligent method, which finally leads to premature execution before reaching the destination.

Fuzzy logic performs well in the UUV path planning and obstacle avoidance owing to its expertise in processing the information uncertainty as the underwater environment is of high uncertainty and incompleteness[53,54]. Kim *et al.* used the fuzzy logic-based algorithm to deduce the turning direction and angle of the UUV to avoid collisions and complete the path planning[55]. Ali developed a fuzzy ontology modeling method to realize the UUV path planning[56]. The fuzzy logic-based intelligent algorithm does not need to establish accurate mathematical models as it is derived from the human cognitive experience. Thus, fuzzy logic can retune itself during the navigation and overcome the local minimum problem. However, the fuzzy logic rule relies heavily on experts' experience and approximations, and unverified errors cannot be thoroughly avoided. The complexity of the dynamic environmental factors also challenges the adaptiveness of the fuzzy logic design[57].

The application of neural networks in vehicle path planning has obtained wide attention in recent decades[58]. Ghatee applied the Hopfield neural network in the optimization of path planning distances[59]. Li *et al.* proposed a bio-inspired neural network for vehicle path planning, where both optimal planning paths and collision avoidance are realized with high efficiency[60]. The bio-inspired neural network helps to derive the optimal path that is composed of the continuous coordinates of the vehicle movement, based on a grid-based map and its corresponding neural network model, where each grid represents a single neuron, as shown in Figure 5. The bio-inspired neural network algorithm continuously updates the state of neurons by transmitting the information through the network to give an instant reaction and reduces the complexity by limiting the searching area to a certain range. Therefore, the bio-inspired neural network path planning utilizes the preserved information in the neurons to update its planning design while adjusting the network on time such that it is well suited to the dynamic underwater environment, providing an efficient and high adaptive approach for the UUV path planning[61].

In recent years, the application of reinforcement learning (RL) in UUV path planning has grown quickly. The RL method updates the vehicle's states and converges to the optimal path planning solution by making actions according to rewards set based on the environment. RL-based path planning combined with APF for intervention AUVs has been developed to remove sea urchins at an affordable cost[62]. AUV path planning in a complex and changeable environment is achieved through the combination of RL and deep learning[63]. Wang *et al.* proposed a multi-behavior critic RL algorithm for AUV path planning to overcome problems associated with oscillating amplitudes and low learning efficiency in the early stages of training, and they reduced the time consumed by the RL algorithm convergence for avoiding obstacles[64]. However, the slow convergence issue of RL-based path planning methods still needs further investigation.

The methods that are commonly used in the point-to-point path planning of the UUV are summarized in Table 2, where their implement theory, benefits, and drawbacks are described. Details of various intelligent

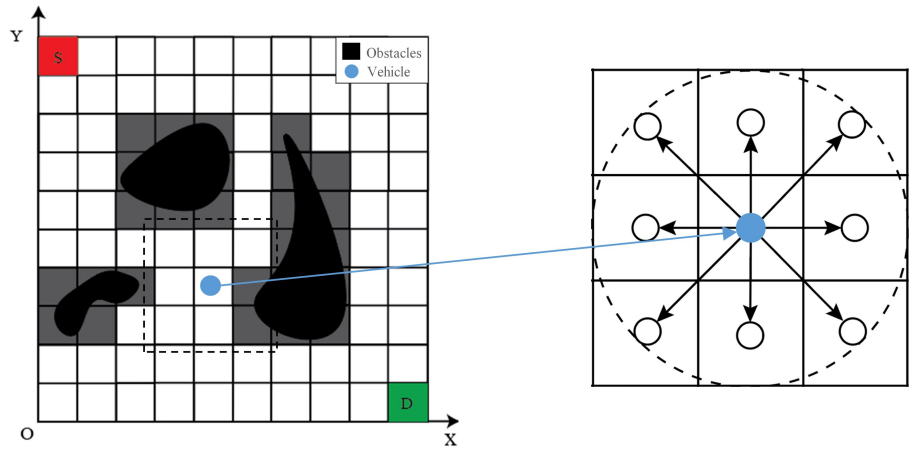


Figure 5. The 2D model of the bio-inspired neural network-based path planning algorithm. S, start; D, destination.

Table 2 Algorithms for UUV point-to-point path planning

Algorithms	Logic	Benefits	Drawbacks
Map Building Method [33-40]	Visibility graph-based: (1) Establish as graph on the connection of the vehicle, polygonal obstacle vertex, and the destination without crossing the obstacles (2) Find the optimal path between the origin point and the destination point that has the shortest distance Grid-based: (1) Decompose the surrounding area into nonoverlapping but connected cells (2) Address the optimal path between the origin and the destination cells without collisions	(1) Easy to implement (2) Direct because of visible mapping	Visibility graph-based: (1) Long time consumption when establishing the graph (2) Lack of flexibility (3) Do not work for circular obstacles Grid-based: (1) Large computation (2) Lack of consideration of environmental disturbance
Artificial Potential Field [41-47]	(1) Predefine a virtual artificial potential field (2) Assume the destination provides the attractive force while obstacles generate repulsive force to the vehicle (3) Address the optimal path for the vehicle through the field descending route	(1) Simple mechanism (2) High efficiency and realtime reaction	(1) Local minimum (2) Sometimes induce large computation
Intelligent Path Planning Algorithms [48-52,55-64] (GA, ACO, Fuzzy logic, NN, and RL)	(1) Regard the path planning as a search optimization problem (2) Take the searching cost as the objective function (3) Optimization through iterations	(1) Easy to implement (2) Adaptiveness.	(1) Unsatisfactory real-time reaction owing to the computation complexity (2) Local minimum

methods applied to the point-to-point path planning of a UUV can be found in the fourth part of Section 2.2.1.

2.2.2. Full-coverage path planning

The full-coverage path planning has to be considered when the vehicle reaches the designated search area, where the global area of the searching map shall be covered. The goal of the full-coverage path planning for the UUV is to simultaneously realize the high coverage rate, the low repetition route, and the short navigating distance.

The random coverage strategy was proposed at early times to complete the full-coverage path planning. Maxim proposed a full-coverage path planning algorithm for multi-robots in the unknown environment, which does not need to obtain the global map information in advance, and the vehicles would not produce collisions with each other [65]. However, the random coverage strategy is used in this algorithm to traverse the operating area

for each vehicle; hence, problems of the path clutter, the high repetition rate, and a not complete full-coverage path planning are induced.

The map building method based on sensor information is combined to achieve a complete full-coverage path planning for the vehicle. Parlaktuna developed a full-coverage path planning method based on the sensor system for multiple vehicles, where the generalized Voronoi diagram was applied for modeling and initializing a full-coverage path, and the path section is divided by the capacitated arc routing algorithm[66]. The full-coverage path planning is realized by the division of the vehicle navigation area; however, it only suits maps consisting of narrow paths such that the vehicle can cover the whole area through one-direction navigation and does not work well in a large space.

Based on the building map, many researchers have refined the full-coverage path planning method in the multi-vehicle system while resolving the collaboration problem, which is denoted as collaborating full-coverage problem. Janchiv applied cell decomposition to separate the operating area into several subareas and determine the suitable path planning result for each subarea. The vehicles can consume the least turning times and maintain a high efficiency to complete the full-coverage path navigation[67]. However, Janchiv's method did not consider the collaboration among the vehicle groups, and the method lacks proof of robustness. Rekletis introduced the boustrophedon cellular decomposition algorithm into the collaborating full-coverage path planning problem for multiple vehicles, where the domain decomposition method breaks the area and a greedy auction algorithm resolves the task assignment, as well as the collaboration of the vehicles[68]. This path planning method achieves the full coverage of the whole area, yet a large repetition of the navigated paths still cannot be avoided. Hazon proposed the multi-robots spanning-tree coverage algorithm (MSTC) that largely increased the robustness for the multiple vehicles to traverse the whole area, while it cannot guarantee the optimal coverage time[69]. Therefore, Zheng developed a multi-robots forest coverage (MFC) algorithm that realized the optimal coverage time[70].

With the advancement of intelligent algorithms, the full-coverage path planning method that can retune or optimize itself has been developed. For instance, Kapanoglu combined the genetic algorithm (GA) and template match approach into the collaborating full-coverage path planning problem, where GA is used to address the best match template for each single vehicle path planning such that both the fewest traversing paths and optimal coverage time can be promised, but the method lacks the adaption for dynamically changing environment, which is commonly seen for the underwater area[71]. The advantage of a bio-inspired neural network is to resolve the collaborating full-coverage path planning problem of ground cleaning robots, where each vehicle regards the others as obstacles such that the full-coverage with collision-free collaboration is realized. However, the large complexity of the neural network is still a big concern[72,73].

Moreover, to increase the efficiency of the full-coverage traversing algorithm, studies related to target search algorithms based on probabilistic priority map have been proposed. For example, Cai developed a full-coverage path planning algorithm depending on the bio-innovation such as animal behaviors, but considering the probabilistic priority, where the efficiency is increased, the method yet is not highly adaptive to the changing environment[74]. Yao proposed full-coverage path planning methods depending on the probability map of targets, where intelligent methods such as biased min-consensus (BN-BMC) algorithm, Gaussian-based analysis, or SOM are combined[75–77].

Generally, most full-coverage path planning methods are applied to land or aerial vehicles rather than UUVs. The problems of not completing full coverage and high repetition routes usually occur during the navigation process. The studies on full-coverage path planning for the UUV in underwater environments are summarized in Table 3, which are still at the very early stage and attention has to be paid to the concerns of enhancing the efficiency of full coverage and decreasing the repetition rate.

Table 3 Algorithms for UUV full-coverage path planning

Algorithms	Logic	Benefits	Drawbacks
Random Coverage Strategy [65]	Traverse the operating area with multiple vehicles following the random coverage strategy	(1) No need of initial environmental information (2) Collision avoidance	(1) Not complete full-coverage (2) High repetition
Sensor-based map building method [66-69]	(1) Build the map based on sensor information (2) Apply the diagram algorithms for modeling and initialize a full-coverage path by dividing the path into sections accordingly	(1) Complete full-coverage (2) Consider multi-vehicle collaboration	(1) Only work for narrow paths (2) Complete full-coverage cannot be realized in conditions of broad area (3) Lack of robustness (4) Lack of optimal multi-vehicle task assignment (5) High repetition
Intelligent method-based full coverage path planning [71-73]	Apply intelligent methods such as GA or NN for each single vehicle path planning	(1) Complete full-coverage; (2) Collision avoidance due to self-regulation; (3) High efficiency of shortest covering time and lowest energy cost	(1) Low adaptiveness to the dynamic environment (GA) (2) Large computation (NN)
Probabilistic priority-based full coverage path planning [74-77]	Plan the path due to the predefined probabilistic priority	(1) Easy to implement (2) Complete full coverage (3) Increasing efficiency	Not adaptive to dynamic environment

3. TECHNOLOGIES OF UUV TRACKING CONTROL

Due to the complex environmental factors of the deep-water space, such as the high pressure, invisibility, or unpredictable obstacles, UUVs are applied in most cases when operating underwater to guarantee the safety and efficiency [2,3,78]. Therefore, achieving the robustness and accuracy of controlling the UUV to track the desired trajectory is dramatically important for completing the real-time underwater navigation [79,80]. As mentioned in the Introduction, UUVs are mainly divided into ROV and AUV. ROV can be directly controlled through a control model for propagation, Robot operating system (ROS) modules, a visual processing pipeline, and a dashboard interface for the end-user, where the user gives commands remotely step by step [81]. This is known as remote control, and the ROV is controlled manually in this case, which is not the critical point of the section as the manual control strategy is direct and simple. For AUVs, the control is realized in an autonomous way, meaning the AUV has to recognize the surrounding areas and make the decision itself. Moreover, some ROVs also support the autonomous mode as a AUV, e.g., the "Falcon" ROV. Hence, in this review, the tracking control technologies emphasize the autonomy of UUVs, and applications on ROVs can also serve as examples of autonomous trajectory tracking control.

To realize the satisfactory trajectory tracking of the UUV, the vehicle must follow the desired path following the corresponding time period. In other words, the errors between the desired and actual trajectories have to be minimized at the different degrees of freedom [82]. However, different from common unmanned vehicles such as the land vehicle or the unmanned surface vehicle (USV), the UUV system contains more states, whose degrees of freedom (DOF) can be extended to six.

For the kinematic equation of the UUV, the velocity vector $\mathbf{v}$ can be transformed into the time derivative of position vector $\mathbf{p}$ by a transformation matrix $\mathbf{J}$ as

$$\dot{\mathbf{p}} = \mathbf{J}(\mathbf{p})\mathbf{v}. \tag{3}$$

where the velocity vector $\mathbf{v}$ is $[u\ v\ w\ r\ p\ q]^T$, as the velocity variable shown at each DOF in Figure 6.

The transformation matrix $\mathbf{J}(\mathbf{p})$ is

$$\mathbf{J}(\mathbf{p}) = \begin{bmatrix} \mathbf{J}_1 & \mathbf{O}_{3\times3} \\ \mathbf{O}_{3\times3} & \mathbf{J}_2 \end{bmatrix}, \tag{4}$$

where $\mathbf{J}_1$ and $\mathbf{J}_2$ are

$$\mathbf{J}_1 = \begin{bmatrix} \cos\psi\cos\theta & \cos\psi\sin\theta\sin\varphi - \sin\psi\cos\varphi & \cos\psi\sin\theta\cos\varphi + \sin\psi\sin\varphi \\ \sin\psi\sin\theta & \sin\psi\sin\theta\sin\varphi + \cos\psi\cos\varphi & \sin\psi\sin\theta\cos\varphi + \cos\psi\sin\varphi \\ -\sin\theta & \cos\theta\sin\varphi & \cos\theta\cos\varphi \end{bmatrix}, \tag{5}$$

$$\mathbf{J}_2 = \begin{bmatrix} 1 & \tan\theta\sin\varphi & \cos\varphi\tan\theta \\ 0 & \cos\varphi & -\sin\varphi \\ 0 & \sin\varphi/\cos\theta & \cos\varphi/\cos\theta \end{bmatrix}. \tag{6}$$

Among the six DOFs of the underwater vehicle, surge, sway, heave, roll, pitch, and yaw, roll and pitch can be neglected since these two DOFs barely have an influence on the underwater vehicle during practical navigation. Therefore, when establishing the trajectory tracking model to keep a controllable operation of the UUV, usually only four DOFs, namely surge, sway, heave, and yaw, are involved (see the DOFs shown in Figure 6). Hence, for the kinematic equation, the position vector can be simplified as

$$\mathbf{p} = \begin{bmatrix} \dot{x} \\ \dot{y} \\ \dot{z} \\ \dot{\psi} \end{bmatrix} = \mathbf{J}(\mathbf{p})\mathbf{v} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \mathbf{v} = \begin{bmatrix} \cos\psi & -\sin\psi & 0 & 0 \\ \sin\psi & \cos\psi & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u \\ v \\ w \\ r \end{bmatrix}, \tag{7}$$

where $\mathbf{J}$ is a transformation matrix derived from the physical structure of the UUV body, while $[u\ v\ w\ r]^T$ represents the velocities at the chosen four axes of a UUV, as presented in Figure 6.

For UUVs, a generally accepted dynamic model has be defined as

$$\mathbf{M}\mathbf{v} + \mathbf{C}(\mathbf{v})\mathbf{v} + \mathbf{D}(\mathbf{v})\mathbf{v} + \mathbf{g}(\mathbf{p}) = \boldsymbol{\tau}, \tag{8}$$

where M is the inertia matrix of the summation of rigid body and added mass; $\mathbf{C}(\mathbf{v})$ is the Coriolis and centripetal matrix of the summation of rigid body and added mass; $\mathbf{D}(\mathbf{v})$ is the quadratic and linear drag matrix; $\mathbf{g}(\mathbf{p})$ is the matrix of gravity and buoyancy; and $\boldsymbol{\tau}$ is the torque vector of the thruster inputs.

The torque vector of the thruster input is represented by

$$\boldsymbol{\tau} = \begin{bmatrix} \tau_x & \tau_y & \tau_z & \tau_k & \tau_m & \tau_n \end{bmatrix}^T, \tag{9}$$

where τ_x, τ_y, τ_z, τ_k, τ_m, and τ_n represent torques of the UUV in the surge, sway, heave, pitch, roll, and yaw directions, as shown in Figure 6. In addition, as mentioned in the previous section, in some practical cases, the torques in pitch and roll directions can be neglected.

Due to the nonlinearity of the UUV system, the typical linearization method, i.e., proportion–integration–differentiation (PID) control, does not work very well and is less studied for UUV trajectory tracking[83]. Hence, in this section, the major methods that are used for UUV trajectory tracking are discussed and categorized into conventional control (consisting of backstepping control, sliding mode control, and model predictive control), intelligent control, and fault-tolerant control.

3.1. Conventional control

In this section, some conventional control methods such as backstepping control, sliding mode control, and model predictive control are described. Studies regarding their applications in the trajectory tracking control of the UUV are stated. The summary of the features for these conventional controls is given in Table 4.

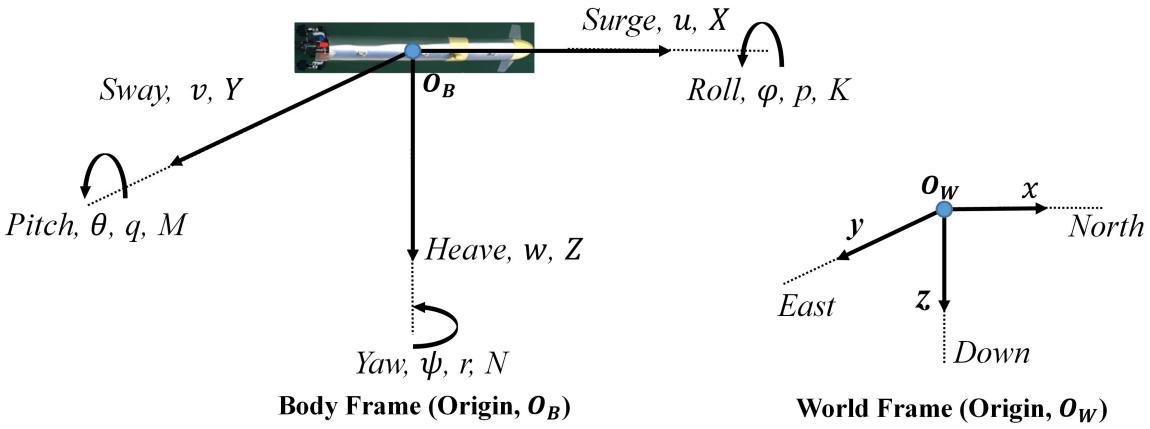


Figure 6. Degrees of freedom and corresponding axes for a UUV.

3.1.1. Backstepping control

In the backstepping method, control functions for each subsystem are designed based on the Lyapunov techniques and generated to form the complete control law[84]. However, the actuator saturation is induced by the speed-jump problem, which usually occurs in the backstepping control methods for trajectory tracking[85]. The excessive speed references affect the robustness of the UUV trajectory tracking by introducing excessive fluctuations of velocities at initial states or other large error states during the kinematic controlling procedure. Therefore, a sharp speed change is derived from the large errors accumulated from the generation of the subsystems, where speed-jump issues are induced when the deviation occurs. As the UUV cannot provide infinite driving inputs such as torques/forces due to its underwater workspace and limited electric power, actuator saturation, has to be considered during the trajectory tracking process of the vehicle, with the torques/forces constraints applied[86–88].

3.1.2. Sliding mode control

As one of the most basic robust controlling strategies, sliding mode control (SMC) is widely used due to its simple and robust mechanism; hence, SMC is often chosen to construct the trajectory tracking controller of the vehicle[89,90]. In SMC, a sliding surface mode is supposed to follow the desired tracking and keep the controlled outputs remaining on the surface. Once the controlled trajectory is out of the perfect sliding surface mode, SMC will push the trajectory slide back to the surface with addition or subtraction on the original controlling equation[91,92]. Therefore, SMC restricts the fluctuation of controlled outputs in an acceptable range through a simple operation, which is highly applicable in trajectory tracking problems[93].

However, SMC suffers chattering issues, although it is robust to variable changes, which is a critical factor that needs to be considered when designing the control strategy[94]. Xu refined SMC with a bio-inspired neural network algorithm such that the chattering problem can be alleviated, but it is limited to the application of land vehicles where fewer degrees of freedom are involved[95].

3.1.3. Model predictive control

Model predictive control (MPC) is appropriate for the UUV system that navigates in the mode of slow velocity. MPC is not demanding on the model accuracy and provides in-time feedback, and constraints can be added to the control strategy to alleviate the jumps of the speed. Therefore, motivated by the requirements of in-time reaction and restriction of velocities within physical constraints throughout the whole tracking process, MPC control stands out to be one of the most feasible solutions for constructing the tracking control for the UUV[96].

The MPC resolves the online optimization problem at each timeslot and derives in-time predictions with min-

imum errors[97]. The optimization process performs a receding horizon in MPC. When deducing the solution of the next timeslot, the optimization algorithm embedded in the control system first gives an optimized sequence within a pre-defined timeslot. The first result of the sequence is adopted as the solution and works as the basis for the next optimization loop while time is receding[98]. At the same time, constraints are added in the optimization to set the limitation to the optimized results as well as the variation of the sequence results[97,99]. By this receding optimization algorithm and the set constraints, online control can be realized and excessive velocity results are avoided. Sun *et al.* applied MPC as the vehicle trajectory tracking control, achieving satisfactory tracking results with fewer and gentler fluctuations, which demonstrates the effectiveness of MPC[85].

3.2. Intelligent control

Intelligent controls refer to the control strategies that can realize desired control goals without manual interventions, which are often used under situations of large uncertainties.

The fuzzy logic system is used as a component of the intelligent control, which addresses the uncertainties and gives a more flexible criterion for obtaining the optimized predictions within its conceptual framework[100,101]. It can also limit the output data and smoothen the kinematic error curves derived from the conventional backstepping method through its decision function. Compared to MPC, the fuzzy logic controller constructs a model that imitates human decision-making with inputs of continuous values between 0 and 1, which largely simplifies the computing process[53,54]. Some researchers have achieved successful tracking based on the fuzzy logic-refined backstepping method, yet their application is based on the underactuated surface vehicle (USV), with fewer states involved compared to the UUV[102]. Some researchers have applied synergetic learning in their controllers designed for vehicles and better performance is obtained, but they do not consider the practical constraints of the vehicle[103]. Li developed the fuzzy logic-based controller that provides satisfactory tracking results even with time-varying delays or input saturation, but the effectiveness of the algorithm on specific models such as the UUV has not been discussed[104]. Wang *et al.* developed a fuzzy logic-based backstepping method, yet it has not been experimented under specific application scenarios, with dynamic constraints applied[105].

As a typical intelligent method, the neural network-based models have been applied to the tracking control of the UUV for many years[2]. Due to the complex underwater work environment and limited electric power of UUVs, the excessive speed references as well as the actuator saturation problems have to be considered. The bio-inspired backstepping controller was introduced in the control design to give the resolution[87]. Based on the characteristics of the shunting model, the outputs of the control are bounded in a limited range with a smooth variation[106]. The bio-inspired backstepping controller has been applied to different UUVs under various conditions by combining with a sliding mode control that controls the dynamic component of the vehicle. An adaptive term is used in the sliding mode control to estimate the nonlinear uncertainties part and the disturbance of the underwater vehicle dynamics[107]. For example, the actuator saturation problem of a 7000 m manned submarine was resolved through this bio-inspired backstepping with sliding mode cascade control[108]. The control contains a kinematic controller that uses a bio-inspired backstepping control to eliminate the excessive speed references when the tracking error occurs at the initial state. Then, a sliding mode dynamic controller was proposed to reduce the lumped uncertainty in the dynamics of the UUV, thus realizing the adaptive trajectory tracking control without excessive speeds for the vehicle, as shown by the satisfactory curve and helix tracking results in Figure 7. Jiang accomplished the trajectory tracking of the autonomous vehicle in marine environments with a similar bio-inspired backstepping controller and adaptive integral sliding mode controller[109]. In the sliding mode controller, the chattering problem was alleviated, which increased the practical feasibility of the vehicle. However, more studies are needed to prove the effectiveness of the proposed control strategy, such as the tracking control based on the filtered backstepping method.

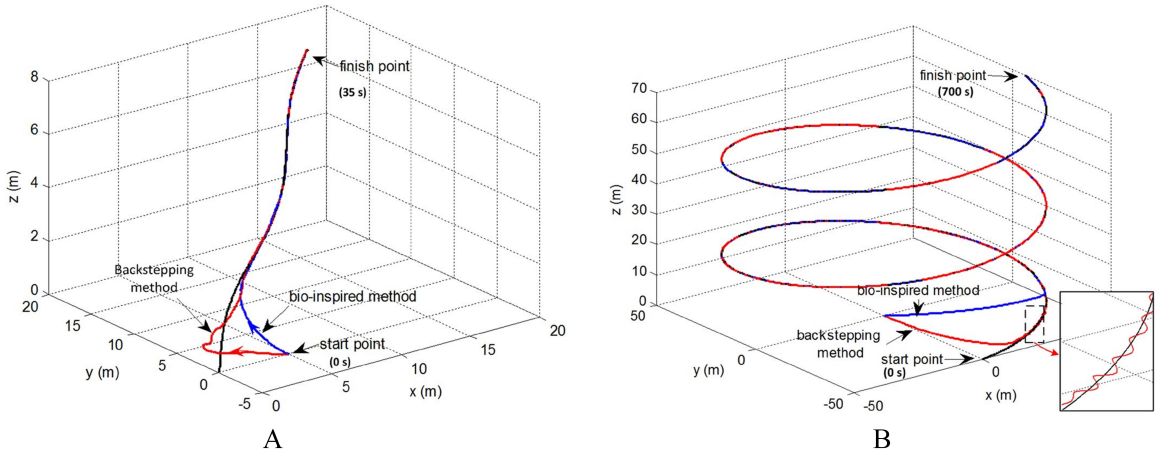


Figure 7. Tracking trajectory comparison of the bio-inspired model-based control and conventional backstepping control for the underwater robots: (A) curve tracking; and (B) helix tracking [108].

GA methods are also applied in the intelligent control of UUVs. They are usually applied based on the aim of addressing the most optimal solution during the control process owing to their feature of self-evolution. However, the computation cost of the GA methods always adds a burden to the tracking control algorithms such that they are usually combined with other intelligent algorithms to reach a more efficient control strategy. Tavanaei–Sereshki applied the quantum genetic algorithm (QGA), an optimization algorithm based on the probability that combines the idea of quantum computing and traditional genetic algorithm to realize the UUV's tracking along desired paths [110]. Zhang described a route planner that enables an AUV to selectively complete part of the predetermined tasks in the operating ocean area when the local path cost is stochastic through a greedy strategy-based GA (GGA), which includes a novel rebirth operator that maps infeasible individuals into the feasible solution space during evolution to improve the efficiency of the optimization and uses a differential evolution planner for providing the deterministic local path cost [111].

A brief summary of intelligent controls on UUVs can be found in Table 4. The details of various intelligent methods for tracking control of UUV are described in Section 3.2.

3.3. Fault-tolerant control

Regarding the unpredictability of the underwater environment, it is of high possibility for the UUV to meet unexpected accidents that affect the preset model of the vehicle. For example, in some cases, one or more of the UUV's thrusters are out of order, and the model needs to be modified to continue the desired trajectory tracking designed as before. Fault-tolerant control (FTC) is usually applied to alleviate abrupt errors and provides the most feasible solution when inevitable damages happen to the equipment in different fields [115]. However, the FTC of underwater vehicles has not been thoroughly investigated due to the complexity brought by the underwater environment and the UUV system [116–118].

Several techniques on the FTC have been developed in the 21st century [119–122]. Based on these studies, the design of the excessive number of thrusters compared to the number of degrees of freedom (DOF) is raised and accepted as a resolution to the UUV FTC problem, which is called the thruster control matrix reconfiguration [123,124]. For example, as shown in Figure 8, the Falcon and URIS UUVs have five thrusters while only four DOFs are considered such that the reconfiguration method can be applied. For example, when an unexpected fault of the vehicle's thrusters occurs, the thrusters installed on the vehicle that exceed the number of DOFs (six, i.e., surge, sway, heave, row, pitch, and yaw) have enough flexible space to be retuned to provide the required propulsion in the corresponding DOFs. To implement the thruster control matrix configuration theory

Table 4 Algorithms for UUV trajectory tracking

Algorithms	Logic	Benefits	Drawbacks
Backstepping control [84-88]	Design control function by generating the sub-system established based on the Lyapunouv theorem	(1) Easy to implement for nonlinear system (2) Real-time reaction	(1) Unsatisfactory adaptiveness (2) Produce excessive speed references and actuator saturation problems
Sliding mode control [89-95]	(1) Suppose a sliding surface mode to follow the desired tracking (2) Keep the controlled outputs on the surface.	(1) Simple algorithm; (2) Robust.	Chattering issue
Model predictive control [96-99,112]	(1) Resolve the online optimization problem in each timeslot and derive in-time predictions with minimum errors (2) Optimization algorithm embedded in the control system gives an optimized sequence within the pre-defined timeslot (3) The first result of the sequence is adopted as the solution and worked as the basis for the next optimization loop while time is receding	(1) High system accuracy not needed (2) In-time reaction (3) Adaptive	Long time consumption due to the recursive computation
Intelligent control [100-109] (Fuzzy logic, NN, and GA)	(1) Embed intelligent algorithms as a search optimization for the desired tracking result; (2) Take searching cost as the objective function; (3) Optimization through iterations.	(1) Easy to implement; (2) Adaptive.	Large computation.

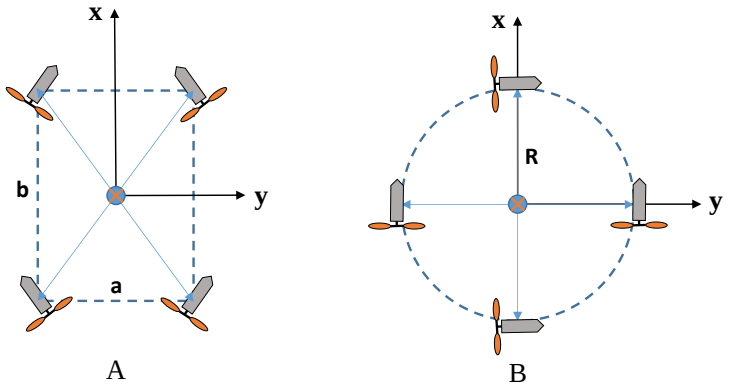


Figure 8. Two typical thruster configurations for the UUV: (A) Falcon UUV; and (B) URIS UUV.

in practical cases, the weighted pseudo-inverse matrix method was proposed, where the fault cases are quantified as degrees of damage and serve as the inputs to form the thruster control matrix configuration model [125]. By this method, the process of the FTC is largely simplified, as the required thruster propulsion can be deduced directly through a weighted pseudo-inverse matrix model. Nevertheless, physical constraints of the thruster outputs are rarely considered, thus inducing the over-actuated vehicle issue [126,127]. Additionally, among these studies, most of them work on eliminating the static errors induced by the fault cases. However, in UUV practical application, the realization of dynamic control on the vehicle's outputs in a real-time manner, which commonly refers to the trajectory tracking control for underwater vehicles, is of crucial importance [112,128].

4. CHALLENGES AND PROSPECTS OF UUV MOTION PLANNING AND TRACKING CONTROL

The motion planning and tracking control of the UUV has a promising future in the maritime projects of underwater rescue, detection, investigation, tube pavement, creature study, and military strategy. Hence, there is still a large requirement for thorough and systematic research in this field. In this section, the challenges and prospects of the UUV motion planning and tracking control for underwater navigation are listed and analyzed.

4.1. Multi-UUV collaboration

In this section, possible future studies on the multi-UUV collaboration are given, mainly divided into the multi-UUV collaboration considering the environmental effect and heterogeneous vehicles and the hunting of dynamic targets.

4.1.1. Environmental effect and heterogeneous vehicles

Most researchers discuss the UUV assignment in an ideal underwater environment and regard the vehicle as a pure particle, which lacks the consideration of the practical condition of the UUV operation. Therefore, the complex environmental factors such as ocean currents effect, unpredictable seamounts, and various moving obstacles should be involved in further studies. Moreover, the UUV system contains uncertainty that cannot be addressed initially, such that the heterogeneous vehicles of different model parameters, navigating velocities, or safe distance are required to be studied. In addition, the topics of the formation control of heterogeneous vehicles between UUVs and USVs (or even unmanned aerial vehicles) have become more attractive, as the USV can help to instruct the UUV in real-time positioning through its less-affected GPS system and efficient communication above the water surface[129,130].

4.1.2. Dynamic targets

Most current UUV motion planning and tracking control studies concentrate on tackling static targets, such as the underwater search of crashed vehicles, yet the following or hunting of dynamic targets for UUVs is a crucial issue in the military defense for the marine system[131]. The following or hunting of dynamic targets covers the topics of dynamic task assignment, intelligence of the moving target, path planning while chasing the target, and containment of the target completed by the UUVs, which are still waiting to be further investigated.

4.2. Efficient underwater positioning and path planning

The developed underwater positioning methods of real-time efficiency are required, as the conventional positioning and navigating systems such as GPS are not valid due to the poor communication affected by the underwater environment. In addition, path planning in the underwater condition is the most vital part of UUV motion planning, where some innovative intelligent planning methods can be applied in this field to improve efficiency, such as reinforcement learning.

4.3. Robust underwater trajectory tracking

Owing to the complexity and uncertainty of the underwater environment and the nonlinear UUV system, the robustness of UUV trajectory tracking has to be advanced to guarantee the UUV navigates as desired. Therefore, the problems that affect the robustness of UUV trajectory tracking such as excessive speed references, actuator saturation, and thruster damages are worth investigating[132].

4.4. Real-time underwater recognition

Current approaches that can attain the identity information of the underwater targets are limited due to the low invisibility and unpredictable obstacles of the surrounding environment; the inefficient communication through the fluids in the deep-sea area also restricts the real-time recognition of the underwater target, thus affecting the real-time UUV navigation to unknown targets. Hence, the advanced underwater target detection techniques such as the multi-sensor information infusion approach and deep learning-based image recognition can be applied in this field to process and achieve the results in time[133–136].

5. CONCLUSION

The paper discusses the methodologies that can be applied to perform satisfactory UUV motion planning and tracking control, as well as the authors' thoughts on the benefits or drawbacks of these methods. In general, motion planning and tracking control for UUVs require the vehicle to realize an efficient and robust underwater operation of addressing and approaching the targets, with the optimal planned paths, task assignment among multiple vehicles, and robust trajectory tracking procedure. The framework and current investigations of UUV motion planning and tracking control are given in Sections 2 and 3. Moreover, although researchers have developed some effective methodologies on these topics, challenges remain to be resolved, which are listed in Section 4, together with the possible developments of UUV motion planning and tracking control technologies.

DECLARATIONS

Authors' contributions

Made substantial contributions to the research and investigation process, reviewed and summarized the literature, wrote and edited the original draft: Zhu D, Yan T

Performed oversight and leadership responsibility for the research activity planning and execution as well as developed ideas and provided critical review, commentary and revision: Yang SX

Availability of data and materials

Not applicable.

Financial support and sponsorship

This work was supported by the Natural Sciences and Engineering Research Council (NSERC) of Canada.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Gafurov SA, Klochkov EV. Autonomous unmanned underwater vehicles development tendencies. *Procedia Eng* 2015;106:141-8. DOI
2. Li D, Wang P, Du L. Path planning technologies for autonomous underwater vehicles-a review. *IEEE Access* 2019;7:9745-68. DOI
3. Burdinsky IN. Guidance algorithm for an autonomous unmanned underwater vehicle to a given target. *Optoelectron Instrument Proc* 2012;48:69-74. DOI
4. Craven PJ, Sutton R, Burns RS. Control strategies for unmanned underwater vehicles. *J Navigation* 1998;51:79–105. DOI
5. Skaddan R, Alhashemi N, Zaini M, Khuraishi M. Design of an improved decision search system for missing aircrafts: MH370 case study the deconstruction of houdini's greatest act. In: *2017 Systems and Information Engineering Design Symposium (SIEDS)*. Piscataway, NJ, USA; 2017. pp. 73-8. DOI
6. Roberts GN, Sutton R. Advances in unmanned marine vehicles. Advances in unmanned marine vehicles. Institution of Engineering and Technology; 2006.
7. Sun P, Boukerche A. Modeling and analysis of coverage degree and target detection for autonomous underwater vehicle-based system. *IEEE Trans Veh Technol* 2018;67:9959-71. DOI

8. Huang H, Tang Q, Li J, et al. A review on underwater autonomous environmental perception and target grasp, the challenge of robotic organism capture. *Ocean Eng* 2020;195. DOI
9. Petillo S, Schmidt H. Exploiting adaptive and collaborative auv autonomy for detection and characterization of internal waves. *IEEE J Ocean Eng* 2014;39:150-64. DOI
10. Panda M, Das B, Subudhi B, Pati BB. A comprehensive review of path planning algorithms for autonomous underwater vehicles. *Int J Autom Comput* 2020;17:321-52. DOI
11. Hadi B, Khosravi A, Sarhadi P. A review of the path planning and formation control for multiple autonomous underwater vehicles. *J Intell Robot Syst* 2021;101. DOI
12. Parker L. Heterogeneous multi-robot cooperation [Ph.D. Thesis]. Massachusetts Institute of Technology; 1994.
13. Kulkarni IS, Pompili D. Task allocation for networked autonomous underwater vehicles in critical missions. *IEEE J Sel Areas Commun* 2010;28:716-27. DOI
14. Mataric MJ. Minimizing complexity in controlling a mobile robot population. In: *Proceedings. 1992 IEEE International Conference on Robotics And Automation (Cat. No.92CH3140-1)*. Los Alamitos, CA, USA; 1992. pp. 830-5. DOI
15. Miyata N, Ota J, Arai T, Asama H. Cooperative transport by multiple mobile robots in unknown static environments associated with real-time task assignment. *IEEE Trans Robot Autom* 2002;18:769-80. DOI
16. Turner RM. Context-mediated behavior for intelligent agents. *International Journal of Human Computer Studies* 1998;48:307-30. DOI
17. Dia H. An agent-based approach to modelling driver route choice behaviour under the influence of real-time information. *Transportation Research Part C: Emerging Technologies* 2002;10:331-49. DOI
18. Ahmed A, Patel A, Brown T, et al. Task assignment for a physical agent team via a dynamic forward/reverse auction mechanism. In: *International Conference on Integration of Knowledge Intensive Multi-Agent Systems (IEEE Cat. No.05EX1033)*. Piscataway, NJ, USA; 2005. pp. 311-7. DOI
19. Akkiraju R, Keskinocak P, Murathy S, Wu F. An agent-based approach for scheduling multiple machines. *Appl Intell, Int J Artif Intell Neural Netw Complex Probl-Solving Technol* 200;14:135-44. DOI
20. Atkinson ML. Results analysis of using free market auctions to distribute control of UAVs. In: *Collection of Technical Papers - AIAA 3rd "Unmanned-Unlimited" Technical Conference, Workshop, and Exhibit*. vol. 2. Chicago, IL, United states; 2004. pp. 803-11. DOI
21. Wahl T, Howell KC. Autonomous guidance algorithm for multiple spacecraft and formation reconfiguration maneuvers. In: *Advances in the Astronautical Sciences*. vol. 158. Napa, CA, United states; 2016. pp. 1939-56.
22. Yao P, Qi S. Obstacle-avoiding path planning for multiple autonomous underwater vehicles with simultaneous arrival. *Sci China Technol Sci* 2019;62:121 – 132. DOI
23. Yao P, Zhao Z, Zhu Q. Path planning for autonomous underwater vehicles with simultaneous arrival in ocean environment. *IEEE Systems Journal* 2020 Sep;14:3185-93. DOI
24. Tolmidis AT, Petrou L. Multi-objective optimization for dynamic task allocation in a multi-robot system. *Engineering Applications of Artificial Intelligence* 2013;26:1458-68. DOI
25. Boveiri HR. An incremental ant colony optimization based approach to task assignment to processors for multiprocessor scheduling. *Front Inform Technol Electron Eng* 2017;18:498-510. DOI
26. Liu C, Kroll A. Memetic algorithms for optimal task allocation in multi-robot systems for inspection problems with cooperative tasks. *Soft Comput* 2015;19:567-84. DOI
27. Kohonen T. Analysis of a simple self-organizing process. *Biol Cybern* 1982;44:135-40. DOI
28. Zhu A, Yang SX. A neural network approach to dynamic task assignment of multirobots. *IEEE Trans Neural Netw* 2006;17:1278-87. DOI
29. Zhu A, Yang SX. An improved SOM-based approach to dynamic task assignment of multi-robots. In: *Proceedings of the World Congress on Intelligent Control and Automation (WCICA)*; 2010. pp. 2168-73. DOI
30. Huang H, Zhu D, Ding F. Dynamic task assignment and path planning for multi-AUV system in variable ocean current environment. *J Intell Robot Syst* 2014;74:999-1012. DOI
31. Chow B. Assigning closely spaced targets to multiple autonomous underwater vehicles [Ph.D. Thesis]. University of Waterloo; 2009.
32. Zhu D, Huang H, Yang SX. Dynamic task assignment and path planning of multi-AUV system based on an improved self-organizing map and velocity synthesis method in three-dimensional underwater workspace. *IEEE Trans Cybern* 2013;43:504-14. DOI
33. D'Amato E, Nardi VA, Notaro I, Scordamaglia V. A Visibility Graph approach for path planning and real-time collision avoidance on maritime unmanned systems. In: *2021 IEEE International Workshop on Metrology for the Sea: Learning to Measure Sea Health Parameters, MetroSea 2021 - Proceedings.* Virtual, Online, Italy; 2021. pp. 400-5. DOI
34. Lam SK, Sridharan K, Srikanthan T. VLSI-efficient schemes for high-speed construction of tangent graph. *Robot Auton Syst* 2005;51:248-60. DOI
35. Magid E, Lavrenov R, Svinin M, Khasianov A. Combining voronoi graph and spline-based approaches for a mobile robot path planning. In: *Informatics in Control, Automation and Robotics. 14th International Conference, ICINCO 2017. Revised Selected Papers: Lecture Notes in Electrical Engineering (LNEE 495)*. Cham, Switzerland; 2020. pp. 475-96.
36. Wang J, Meng MQH. Optimal path planning using generalized voronoi graph and multiple potential functions. *IEEE Trans Ind Electron* 2020;67:10621-30. DOI
37. Dijkstra E. Communication with an automatic computer [Ph.D. Thesis]. University of Amsterdam, Netherlands; 1959.
38. Peter EH, Nils JN, Bertram R. A formal basis for the heuristic determination of minimum cost paths. *IEEE Transactions on Systems Science and Cybernetics* 1968;SSC-4:100-7. DOI
39. Wu Y, Low KH, Lv C. Cooperative path planning for heterogeneous unmanned vehicles in a search-and-track mission aiming at an

underwater target. *IEEE Trans Veh Technol* 2020;69:6782-87. DOI
40. Singh Y, Sharma S, Sutton R, Hatton D, Khan A. A constrained A* approach towards optimal path planning for an unmanned surface vehicle in a maritime environment containing dynamic obstacles and ocean currents. *Ocean Engineering* 2018;169:187-201. DOI
41. Khatib O. Real-time obstacle avoidance for manipulators and mobile robots. In: *Proceedings. 1985 IEEE International Conference on Robotics and Automation*. vol. 2; 1985. pp. 500-5. DOI
42. Zhu D, Yang SX. Path planning method for unmanned underwater vehicles eliminating effect of currents based on artificial potential field. *J Navig* 2021;74:955-67. DOI
43. Ralli E, Hirzinger G. Fast path planning for robot manipulators using numerical potential fields in the configuration space. In: *IROS '94. Proceedings of the IEEE/RSJ/GI International Conference on Intelligent Robots and Systems. Advanced Robotic Systems and the Real World (Cat. No.94CH3447-0)*. vol. vol.3. New York, NY, USA; 1994. pp. 1922-9. DOI
44. Zhou Z, Wang J, Zhu Z, Yang D, Wu J. Tangent navigated robot path planning strategy using particle swarm optimized artificial potential field. *Optik* 2018;158:639-51. DOI
45. Lin Z, Yue M, Wu X, Tian H. An improved artificial potential field method for path planning of mobile robot with subgoal adaptive selection. In: *Intelligent Robotics and Applications. 12th International Conference, ICIRA 2019. Proceedings: Lecture Notes in Artificial Intelligence (LNAI 11740). vol. pt.I*. Cham, Switzerland; 2019. pp. 211-20. DOI
46. Xin L, Zhan-Qing W, Xu-Yang C. Path planning with improved artificial potential field method based on decision tree. In: *2020 27th Saint Petersburg International Conference on Integrated Navigation Systems (ICINS)*. Piscataway, NJ, USA; 2020. p. 5 pp. DOI
47. Abdur Rahman M, Abul Kalam Azad M. To escape local minimum problem for multi-agent path planning using improved artificial potential field-based regression search method. In: *ACM International Conference Proceeding Series*. Singapore, Singapore; 2017. pp. 371-6. DOI
48. Alvarez A, Caiti A, Onken R. Evolutionary path planning for autonomous underwater vehicles in a variable ocean. *IEEE J Oceanic Eng* 2004;29:418–29. DOI
49. Cheng CT, Fallahi K, Leung H, Tse CK. A genetic algorithm-inspired UUV path planner based on dynamic programming. *IEEE Trans Syst Man Cybern, C, Appl Rev* 2012;42:1128-34. DOI
50. Ma YN, Gong YJ, Xiao CF, Gao Y, Zhang J. Path planning for autonomous underwater vehicles: an ant colony algorithm incorporating alarm pheromone. *IEEE Trans Veh Technol* 2019;68:141–54. DOI
51. Han G, Zhou Z, Zhang T, et al. Ant-colony-based complete-coverage path-planning algorithm for underwater gliders in ocean areas with thermoclines. *IEEE Trans Veh Technol* 2020;69:8959-71. DOI
52. Mo H, Xu L. Research of biogeography particle swarm optimization for robot path planning. *Neurocomputing* 2015;148:91-9. DOI
53. Lee CC. Fuzzy logic in control systems: fuzzy logic controller. I. *IEEE Trans Syst Man Cybern* 1990;20:404-18. DOI
54. Lee CC. Fuzzy logic in control systems: fuzzy logic controller. II. *IEEE Trans Syst Man Cybern* 1990;20:419-35. DOI
55. Kim YG, Bui LD. An obstacle-avoidance technique for autonomous underwater vehicles based on BK-products of fuzzy relation. *Fuzzy Sets Syst* 2006;157:560-77. DOI
56. Ali F, Kim EK, Kim YG. Type-2 fuzzy ontology-based semantic knowledge for collision avoidance of autonomous underwater vehicles. *Inf Sci* 2015;295:441-64. DOI
57. LeBlanc K, Saffiotti A. Multirobot object localization: a fuzzy fusion approach. *IEEE Trans Syst Man Cybern B, Cybern* 2009;39:1259-76. DOI
58. Ling S. A real-time collision-free path planning of a rust removal robot using an improved neural network. *J Shanghai Jiaotong Univ, Sci* 2017;22:633-40. DOI
59. Ghatee M, Mohades A. Motion planning in order to optimize the length and clearance applying a Hopfield neural network. *Expert Syst Appl* 2009;36:4688-95. DOI
60. Li H, Yang SX, Biletskiy Y. Neural network based path planning for a multi-robot system with moving obstacles. In: *2008 IEEE International Conference on Automation Science and Engineering (CASE 2008)*. Piscataway, NJ, USA; 2008. pp. 163-8. DOI
61. Zhu D, Yang SX. Bio-inspired neural network-based optimal path planning for UUVs under the effect of ocean currents. *IEEE Trans Veh Technol* 2021. DOI
62. Noguchi Y, Maki T. Path planning method based on artificial potential field and reinforcement learning for intervention AUVs. In: *2019 IEEE Underwater Technology (UT)*. Piscataway, NJ, USA; 2019. pp. 1-6. DOI
63. Li Z, Luo X. Autonomous underwater vehicles (AUVs) path planning based on Deep Reinforcement Learning. In: *2021 11th International Conference on Intelligent Control and Information Processing (ICICIP)*. Piscataway, NJ, USA; 2021. pp. 125-9. DOI
64. Wang Z, Zhang S, Feng X, Sui Y. Autonomous underwater vehicle path planning based on actor-multi-critic reinforcement learning. *Proc Inst Mech Eng, I, J Syst Control Eng* 2021;235:1787-96. DOI
65. Batalin MA, Sukhatme GS. Spreading out: a local approach to multi-robot coverage. In: *Distributed Autonomous Robotic Systems 5*. Tokyo; 2002. pp. 373-82.
66. Parlaktuna O, Sipahioglu A, Kirlik G, Yazici A. Multi-robot sensor-based coverage path planning using capacitated arc routing approach. In: *2009 IEEE International Conference on Control Applications (CCA)*. Piscataway, NJ, USA; 2009. pp. 1146-51. DOI
67. Janchiv A, Batsaikhan D, Kim Gh, Lee SG. Complete coverage path planning for multi-robots based on. In: *2011 11th International Conference on Control, Automation and Systems*; 2011. pp. 824–27.
68. Rekleitis I, New A, Rankin E, Choset H. Efficient boustrophedon multi-robot coverage: an algorithmic approach. *Ann Math Artif Intell* 2008;52:109-42. DOI
69. Hazon N, Kaminka GA. On redundancy, efficiency, and robustness in coverage for multiple robots. *Robotics and Autonomous Systems*

2008;56:1102-14. DOI
70. Zheng X, Koenig S. Robot coverage of terrain with non-uniform traversability. In: *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems. Piscataway, NJ, USA*; 2007. pp. 3757-64. DOI
71. Kapanoglu M, Alikalfa M, Ozkan M, Yazici A, Parlaktuna O. A pattern-based genetic algorithm for multi-robot coverage path planning minimizing completion time. *J Intell Manuf* 2012;23:1035-45. DOI
72. Yang SX, Luo C. A neural network approach to complete coverage path planning. *IEEE Trans Syst Man Cybern B Cybern* 2004;34:718–24. DOI
73. Yao P, Zhao Z. Improved Glasius bio-inspired neural network for target search by multi-agents. *Information Sci* 2021;568:40-53. DOI
74. Cai W, Zhang M, Zheng YR. Task assignment and path planning for multiple autonomous underwater vehicles using 3D dubins curves. *Sensors* 2017;17:1607-26. DOI
75. Yao P, Qiu L, Qi J, Yang R. AUV path planning for coverage search of static target in ocean environment. *Ocean Eng* 2021;241. DOI
76. Song D, Yao P. Search for static target in nonwide area by AUV: a prior data-driven strategy. *IEEE Syst J* 2021;15:3185-8. DOI
77. Yao P, Zhu Q, Zhao R. Gaussian mixture model and self-organizing map neural-network-based coverage for target search in curve-shape area. *IEEE Trans Cybern* 2022;52:3971–83. DOI
78. Sun P, Boukerche A. Modeling and analysis of coverage degree and target detection for autonomous underwater vehicle-based system. *IEEE Trans Veh Technol* 2018;67:9959-71. DOI
79. Bacha S, Saadi R, Ayad MY, Aboubou A, Bahri M. A review on vehicle modeling and control technics used for autonomous vehicle path following. In: *2017 International Conference on Green Energy Conversion Systems (GECS)*. Piscataway, NJ, USA; 2017. pp. 1-6. DOI
80. Liu X, Zhang M, Rogers E. Trajectory tracking control for autonomous underwater vehicles based on fuzzy re-planning of a local desired trajectory. *IEEE Trans Veh Technol* 2019;68:11657-67. DOI
81. Ray S, Bhowal R, Patel P, Panaiyappan AK. An overview of the design and development of a 6 dof remotely operated vehicle for underwater structural inspection. In: *2021 International Conference on Communication, Control and Information Sciences (ICCISC)*. Piscataway, NJ, USA; 2021. pp. 1-6. DOI
82. Shen C, Shi Y, Buckham B. Trajectory tracking control of an autonomous underwater vehicle using lyapunov-based model predictive control. *IEEE Trans Ind Electron* 2018;65:5796-805. DOI
83. Li J, Xu Z, Zhu D, et al. Bio-inspired intelligence with applications to robotics: a survey. *Intell Robot* 2022;1:58–83. DOI
84. Zhu D, Sun B. The bio-inspired model based hybrid sliding-mode tracking control for unmanned underwater vehicles. *Eng Appl Artif Intell* 2013;26:2260-9. DOI
85. Sun B, Zhang W, Song A, Zhu X, Zhu D. Trajectory tracking and obstacle avoidance control of unmanned underwater vehicles based on MPC. In: *IEEE 8th International Conference on Underwater System Technology: Theory and Applications (USYS)*. Piscataway, NJ, USA; 2018. pp. 1–6. DOI
86. Wan L, Sun N, Liao YL. Backstepping control method for the trajectory tracking for the underactuated autonomous underwater vehicle. *AMR*2013;798-799:484-8. DOI
87. Karkoub M, Wu HM, Hwang CL. Nonlinear trajectory-tracking control of an autonomous underwater vehicle. *Ocean Eng* 2017;145:188-98. DOI
88. Yang SX, Meng MQH. Real-time collision-free motion planning of a mobile robot using a Neural Dynamics-based approach. *IEEE Trans Neural Netw* 2003;14:1541-52. DOI
89. Li T, Zhao R, Chen CLP, Fang L, Liu C. Finite-time formation control of under-actuated ships using nonlinear sliding mode control. *IEEE Trans Cybern* 2018;48:3243-53. DOI
90. Qin J, Zhang G, Zheng WX, Kang Y. Adaptive sliding mode consensus tracking for second-order nonlinear multiagent systems with actuator faults. *IEEE Trans Cybern* 2019;49:1605-15. DOI
91. Zaihidee FM, Mekhilef S, Mubin M. Robust speed control of PMSM using sliding mode control (SMC)-a review. *Energies* 2019;12:1669-96. DOI
92. Dhanasekar R, Ganesh Kumar S, Rivera M. Sliding mode control of electric drives/review. In: *2016 IEEE International Conference on Automatica (ICA-ACCA). Piscataway, NJ, USA*; 2016. pp. 1–7. DOI
93. Liu H, Zhang T. Fuzzy sliding mode control of robotic manipulators with kinematic and dynamic uncertainties. *J DYN SYST-T ASME* 2012;134. DOI
94. Slotine JJE, Coetsee JA. Adaptive sliding controller synthesis for non-linear systems. *Int J Control* 1986;43:1631-51. DOI
95. Xu Z, X Yang S, Gadsden SA, Li J, Zhu D. Backstepping and sliding mode control for AUVs aided with bioinspired neurodynamics. In: *2021 IEEE International Conference on Robotics and Automation (ICRA)*. Xi'an, China; 2021. pp. 2113-9. DOI
96. Bai G, Meng Y, Liu L, Luo W, Gu Q. Review and comparison of path tracking based on model predictive control. *Electronics* 2019;8:1077 (32 pp.) . DOI
97. Dong L, Yan J, Yuan X, He H, Sun C. Functional nonlinear model predictive control based on adaptive dynamic programming. *IEEE Trans Cybern* 2019;49:4206-18. DOI
98. Liu L, He Y, Han C. Review of model predictive control methods for time-delay systems. In: *Proceedings of 2020 Chinese Intelligent Systems Conference. Lecture Notes in Electrical Engineering (LNEE 705)*. vol. 1. Singapore; 2021. pp. 624–33. DOI
99. Gutierrez B, Kwak SS. Modular multilevel converters (MMCs) controlled by model predictive control with reduced calculation burden. *IEEE Trans Power Electron* 2018;33:9176-87. DOI
100. Na J, Huang Y, Wu X, Su S, Li G. Adaptive finite-time fuzzy control of nonlinear active suspension systems with input delay. *IEEE Trans Cybern* 2020;50:2639-50. DOI

101. Wang F, Chen B, Sun Y, Gao Y, Lin C. Finite-time fuzzy control of stochastic nonlinear systems. *IEEE Trans Cybern* 2020;50:2617-26. DOI
102. Wang N, Karimi HR. Successive waypoints tracking of an underactuated surface vehicle. *IEEE Trans Ind Inf* 2020;16:898-908. DOI
103. Hayashibe M, Shimoda S. Synergetic learning control paradigm for redundant robot to enhance error-energy index. *IEEE Trans Cogn Dev Syst* 2018;10:573-84. DOI
104. Li Y, Tong S, Li T. Composite adaptive fuzzy output feedback control design for uncertain nonlinear strict-feedback systems with input saturation. *IEEE Trans Cybern* 2015;45:2299-308. DOI
105. Wang H, Liu PX, Niu B. Robust fuzzy adaptive tracking control for nonaffine stochastic nonlinear switching systems. *IEEE Trans Cybern* 2018;48:2462-71. DOI
106. Zhu D, Hua X, Sun B. A neurodynamics control strategy for real-time tracking control of autonomous underwater vehicle. *J Navig* 2013;67:113–27. DOI
107. Sun B, Zhu D, Ding F, Yang SX. A novel tracking control approach for unmanned underwater vehicles based on bio-inspired neurodynamics. *IJ Mar Sci Tech-japan* 2012;18:63–74. DOI
108. Sun B, Zhu D, Yang SX. A bioinspired filtered backstepping tracking control of 7000-m manned submarine vehicle. *IEEE Trans Ind Electron* 2014;61:3682–93. DOI
109. Jiang Y, Guo C, Yu H. Robust trajectory tracking control for an underactuated autonomous underwater vehicle based on bioinspired neurodynamics. *Int J Adv Robot Syst* 2018;15:172988141880674. DOI
110. Tavanaei-Sereshki Z, Ramezani-al MR. Quantum genetic sliding mode controller design for depth control of an underwater vehicle. *Meas* 2018;51:336-48. DOI
111. Zhang J, Liu M, Zhang S, Zheng R. Robust global route planning for an autonomous underwater vehicle in a stochastic environment. *Front Inf Technol Electron Eng* 2022. DOI
112. Shen C, Shi Y, Buckham B. Trajectory tracking control of an autonomous underwater vehicle using lyapunov-based model predictive control. *IEEE Trans Ind Electron* 2018;65:5796-5805. DOI
113. Li D, Wang P, Du L. Path planning technologies for autonomous underwater vehicles-a review. *IEEE Access* 2019;7:9745-768. DOI
114. Karkoub M, Wu HM, Hwang CL. Nonlinear trajectory-tracking control of an autonomous underwater vehicle. *Ocean Eng* 2017;145:188-98. DOI
115. Cao X, Tian Y, Ji X, Qiu B. Fault-tolerant controller design for path following of the autonomous vehicle under the faults in braking actuators. *IEEE Trans Transp Electrification* 2021;7:2530-40. DOI
116. Seto M, Svendsen K. Advanced AUV fault management. In: *Autonomous Underwater Vehicles: Design and practice. The Institution of Engineering and Technology*; 2020. pp. 419-45.
117. Kadiyam J, Parashar A, Mohan S, Deshmukh D. Actuator fault-tolerant control study of an underwater robot with four rotatable thrusters. *Ocean Eng* 2020;197. DOI
118. Zhu G, Ma Y, Li Z, Malekian R, Sotelo M. Event-triggered adaptive neural fault-tolerant control of underactuated MSVs with input saturation. *IEEE Trans Intell Transp Syst* 2021. DOI
119. Qi X, Qi J, Theilliol D, et al. A review on fault diagnosis and fault tolerant control methods for single-rotor aerial vehicles. *J Intell Robot Syst* 2014;73:535-55. DOI
120. Li T, Li G, Zhao Q. Adaptive fault-tolerant stochastic shape control with application to particle distribution control. *IEEE Trans Syst Man Cybern, Syst* 2015;45:1592-1604. DOI
121. Lu K, Xia Y, Yu C, Liu H. Finite-time tracking control of rigid spacecraft under actuator saturations and faults. *IEEE Trans Autom Sci Eng* 2016;13:368-81. DOI
122. Meyer RT, Johnson SC, DeCarlo RA, Pekarek S, Sudhoff SD. Hybrid electric vehicle fault tolerant control. *J Dyn Syst-T ASME* 2018;140. DOI
123. Martynova LA, Rozengauz MB. Approach to reconfiguration of a motion control system for an autonomous underwater vehicle. *Gyroscopy Navig* 2020;11:244-53. DOI
124. Liu X, Zhang M, Yao F. Adaptive fault tolerant control and thruster fault reconstruction for autonomous underwater vehicle. *Ocean Eng* 2018;155:10-23. DOI
125. Omerdic E, Roberts G. Thruster fault diagnosis and accommodation for open-frame underwater vehicles. *Control Eng Pract* 2004;12:1575-98. DOI
126. Podder TK, Antonelli G, Sarkar N. Fault tolerant control of an autonomous underwater vehicle under thruster redundancy: simulations and experiments. In: *Proc IEEE International Conference on Robotics and Automation (ICRA)* 2000;2:1251-6. DOI
127. Zhang S, Wu Y, He X, Liu Z. Cooperative fault-tolerant control for a mobile dual flexible manipulator with output constraints. *IEEE Trans Autom Sci Eng* 2021. DOI
128. Wen Q, Kumar R, Huang J. Framework for optimal fault-tolerant control synthesis: maximize prefault while minimize post-fault behaviors. *IEEE Trans Syst Man Cybern, Syst* 2014;44:1056-66. DOI
129. Stewart W, Weisler W, MacLeod M, et al. Design and demonstration of a seabird-inspired fixed-wing hybrid UAV-UUV system. *Bioinsp Biomim* 2018;13:056013-28. DOI
130. Yang Q, Parasuraman R. Needs-driven heterogeneous multi-robot cooperation in rescue missions. In: *IEEE International Symposium on Safety, Security, and Rescue Robotics (SSRR)*. USA; 2020. pp. 252-59. DOI
131. Liang H, Fu Y, Kang F, Gao J, Qiang N. A behavior-driven coordination control framework for target hunting by UUV intelligent swarm. *IEEE Access* 2020;8:4838-59. DOI

132. Shen C, Buckham B, Shi Y. Modified C/GMRES algorithm for fast nonlinear model predictive tracking control of AUVs. *IEEE Trans Control Syst Technol* 2017;25:1896-1904. DOI
133. Reed S, Ruiz IT, Capus C, Petillot Y. The fusion of large scale classified side-scan sonar image mosaics. *IEEE Trans Image Process* 2006;15:2049–60. DOI
134. Coiras E, Mignotte PY, Petillot Y, Bell J, Lebart K. Supervised target detection and classification by training on augmented reality data. *IET Radar Sonar Navig* 2007;1:83–90. DOI
135. Williams DP. Fast target detection in synthetic aperture sonar imagery: a new algorithm and large-scale performance analysis. *IEEE J Ocean Eng* 2015;40:71–92. DOI
136. Hou GJ, Luan X, Song DL, Ma XY. Underwater man-made object recognition on the basis of color and shape features. *J Coast Res* 2016;32:1135-1141. DOI

Yang *et al*. *Intell Robot* 2022;2(3):223-43
DOI: 10.20517/ir.2022.19

Intelligence & Robotics

Research Article

Open Access

$\mathcal{H}_\infty$ leader-following consensus of multi-agent systems with channel fading under switching topologies: a semi-Markov kernel approach

Haoyue Yang[1], Hao Zhang[1], Zhuping Wang[1], Xuemei Zhou[2]

[1]College of Electronic and Information Engineering, Tongji University, Shanghai 201804, China.
[2]College of Transportation Engineering, Tongji University, Shanghai 201804, China.

Correspondence to: Pro. Hao Zhao, College of Electronic and Information Engineering, Tongji University, 4800, Cao'an Road, Shanghai 201804, China. E-mail: zhang_hao@tongji.edu.cn

How to cite this article: Yang H, Zhang H, Wang Z, Zhou X. $\mathcal{H}_\infty$ leader-following consensus of multi-agent systems with channel fading under switching topologies: a semi-Markov kernel approach. *Intell Robot* 2022;2(3):xx. http://dx.doi.org/10.20517/ir.2022.19

Received: 6 Jun 2022 **First Decision:** 19 Jul 2022 **Revised:** 31 Jul 2022 **Accepted:** 10 Aug 2022 **Published:** 20 Aug 2022

Academic Editor: Nallappan Gunasekaran **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

This paper focuses on the leader-following consensus problem of discrete-time multi-agent systems subject to channel fading under switching topologies. First, a topology switching-based channel fading model is established to describe the information fading of the communication channel among agents, which also considers the channel fading from leader to follower and from follower to follower. It is more general than models in the existing literature that only consider follower-to-follower fading. For discrete multi-agent systems, the existing literature usually adopts time series or Markov process to characterize topology switching while ignoring the more general semi-Markov process. Based on the advantages and properties of semi-Markov processes, discrete semi-Markov jump processes are adopted to model network topology switching. Then, the semi-Markov kernel approach for handling discrete semi-Markov jumping systems is exploited and some novel sufficient conditions to ensure the leader-following mean square consensus of closed-loop systems are derived. Furthermore, the distributed consensus protocol is proposed by means of the stochastic Lyapunov stability theory so that the underlying systems can achieve $\mathcal{H}_\infty$ consensus performance index. In addition, the proposed method is extended to the scenario where the semi-Markov kernel of semi-Markov switching topologies is not completely accessible. Finally, a simulation example is given to verify the results proposed in this paper. Compared with the existing literature, the method in this paper is more effective and general.

Keywords: $\mathcal{H}_\infty$ leader-following consensus, multi-agent systems, channel fading, semi-Markov switching topologies, semi-Markov kernel.

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

1. INTRODUCTION

With the rapid development of computer technology and networks, distributed cooperative control has drawn increasing attention due to its application in various fields, especially computer science and automation control. A fundamental and important research topic of distributed system control is the consensus problem of multi-agent systems. It has attracted considerable interest among many researchers in different fields in the past few decades, due to its significant applications in civilians and militaries, such as unmanned air vehicles[1–3], autonomous underwater vehicles[4], multiple surface vessels[5], robot formation[6,7]. The consensus problem essentially refers to a team of agents reaching the same state by designing proper and available distributed control algorithms that only utilize local information exchange with neighbors. Over the past decade, there have been a wealth of interesting and instructive achievements focusing on consensus problem of multi-agent systems, including leaderless consensus[8–13] and leader-following consensus[14–21]. The leaderless output consensus problem of multi-agent systems composed of agents with different orders was studied by transforming the original system through feedback linearization. Static feedback and dynamic feedback controllers are designed to solve the consensus problem and sinusoidal synchronization problem under uniformly jointly strongly connected topologies[8]. Under cyber-attacks, literature[9] proposed a fully distributed adaptive control protocol to solve the leaderless consensus problem of uncertain high-order nonlinear systems. The work[11] discussed the event-triggered coordination problem for general linear multi-agent systems based on a Lyapunov equation method. Leader-following consensus means that the states of all follower agents are expected to approach the state of the leader agent. In many practical situations, leader-following consensus can accomplish more complex tasks by enhancing inter-agent communication. Compared with leaderless consensus, leader-following consensus can be beneficial in reducing control costs and save energy. The key to the leader-follower consensus problem is how to design a distributed control protocol to synchronize the states of all follower agents and the leader agent. The work[14] proposed a novel distributed observer-type consensus controller for high-order stochastic strict feedback multi-agent systems based only on relative output measurements of neighbors. The 1-moment exponential leader-following consensus of the underlying system is ensured by adopting appropriate state transformation. In[15], the sampled-data leader-following consensus problem for a family of general linear multi-agent systems was addressed, and the distributed asynchronous sampled-data state feedback control law was designed. The event-based secure leader-following consensus control problem of multi-agent systems with multiple cyber attacks, which contain reply attacks and DoS attacks simultaneously, was studied in[16]. The fixed-time leader-following group consensus of multi-agent systems composed of first-order integrators was realized under a directed graph[17]. By designing a nonlinear distributed controller, the follower agents of every group can reach an agreement with its corresponding leader within a specified convergence time. In[19], the author considered the problem of resilient practical cooperative output regulation of heterogeneous linear multi-agent systems, in which the dynamics of exosystem are unknown and switched under DOS attacks. A new cooperative output regulation scheme consisting of distributed controller, distributed resilient observer, auxiliary observer and data-driven learning algorithm was proposed to ensure the global uniform boundedness of the regulated output. More results can be seen in[18,20,21] and references cited therein.

It is well known that there is a large amount of data transmission in the control process of multi-agents. Data packets or signals between agents are usually transmitted through wireless communication networks. However, some special physical phenomena (such as reflection, refraction, diffraction) may occur during the transmission of a signal or data packet through a communication link or channel, which will result in the loss of signal energy and lead to the signal distorted. This type of phenomenon is often referred to the channel fading. Typically, fading effects are closely related to multipath propagation and shadows from obstacles. In practical applications, the factors that cause channel fading mainly include time, geographic location, and radar frequency. As a result, the phenomenon of channel fading may result in the degradation of signal quality due to the inability to receive accurate transmission information, thereby deteriorating the desired performance of the system. This also shows that it is meaningful to consider channel fading effects for the distributed control of multi-agent systems. In view of this, some results on channel fading have been published, such as chan-

nel fading of single systems[22–26], channel fading for multi-agent systems[27–29]. The reference[22] designed a nonparallel distribution compensation interval type-2 fuzzy controller to address dynamic event-triggered control problems for interval type-2 fuzzy systems subject to fading channel, where the fading phenomenon is characterized by a time-varying random process. The literature[23] focused on the finite-horizon H_∞ state estimation problem of periodic neural networks subject to multi-fading channels. By employing the stochastic analysis method and introducing a set of correlated random variables, sufficient criteria to ensure the stochastic stability of the estimation error system with correlated fading channels were obtained and the desired H_∞ performance was achieved. In[25], the event-triggered asynchronous guaranteed that cost control problem for Markov jump neural networks subject to fading channels could be addressed, where a novel rice fading model was established to consider the effects of signal reflections and shadows in wireless networks. The consensus tracking problem of second-order multi-agent systems with channel fading was investigated using the sliding mode control method, and the feasible distributed sliding mode controller was designed by introducing the statistical information of channel fading to the measure functions of the consensus errors[27]. It should be pointed out that most of the literatures mentioned above on channel fading in multi-agent systems only consider the fading effect among the follower agents and ignore the fading effect between the leader agent and the follower agents. As stated earlier, the leader plays a crucial role in the leader-following consensus problem. To improve the applicability of the controller and the ability to deal with the problem, it is reasonable and necessary to consider both the fading effect of leader-to-follower and follower-to-follower agents at the same time in the channel fading problem of multi-agent systems. This is one of the motivations of this paper.

On the other hand, the communication topology of multi-agent systems may change in practice due to various factors, such as sudden changes in the environment, communication range limitations, link failures, packet loss, malicious cyber attacks, etc. Given this, many researchers assume that the topology among agents is time-varying or Markov switching. Some good consensus results for multi-agent systems under time-varying topology and Markov switching topology have been reported in the past decade[30–34]. For example, the work[33] investigated the coupled group consensus problem for general linear time-invariant multi-agent systems under continuous-time homogeneous Markov switching topology. The designed linear consensus protocol can achieve coupled group consensus of the considered system under some algebraic and topological conditions. It is worth noting that since the transition probability in Markov jump process is constant and there is no memory characteristic, there are still some limitations in using Markov jump process to model topology switching among agents. Recently, a class of more general semi-Markov jump processes with a non-exponential distribution of sojourn-time (the time interval between two consecutive jumps) and time-varying transition probabilities has attracted interest of many scholars and has been used to characterize the topological switching among agents[35–38]. For example, the leader-following consensus of a multi-agent system under a sampled-data-based event-triggered transmission scheme was realized[35], where a semi-Markov jump process was employed to model the switching of the network topologies. The containment control problem concerning semi-Markov jump multi-agent systems with semi-Markov switching topologies was studied by designing static and dynamic containment controllers[37]. Under a semi-Markov switching topology with partially unknown transition rates, the H_∞ leader-follower consensus control of a class of nonlinear multi-agent systems with external perturbations was achieved, and sufficient conditions for ensuring system consistency and H_∞ performance were derived based on the linear matrix inequality form[38]. However, most of the above literature on semi-Markov switching topology is considered for the continuous system case. As a matter of fact, in the discrete-time case, the semi-Markov jump process can exert a stronger modeling ability and have a larger application range. The reason is that the probability density function of sojourn-time in the discrete semi-Markov jump process can be of different types in different modes, or of the same type but with different parameters. In order to make the modeling of switching topology more realistic, it is very necessary and valuable to employ discrete-time semi-Markov switching topologies. Naturally, how to deal with the consensus problem of multi-agent systems in this situation is a key point. Recently, the discrete semi-Markov jump process was adopted to model general linear systems and a semi-Markov kernel method was proposed to address

its stability and stabilization problems[39,40]. This also provides an idea for solving the consensus problem of multi-agent systems under discrete semi-Markov switching topology. To the best of our knowledge, to date, there have been few results on the leader-following consensus problem for discrete-time multi-agent systems subject to channel fading under semi-Markov switching topologies. Therefore, how to design a suitable distributed control protocol and how to establish leader-following consensus criteria for multi-agent systems with channel fading under discrete semi-Markov switching topologies are the key issues. This inspires us to carry out this work.

Motivated by the above discussion, in this paper, the $\mathcal{H}_\infty$ leader-following consensus problem of discrete multi-agent systems with channel fading is investigated under the premise of semi-Markov switching topology. The major contributions of this paper can be highlighted as follows: (I) Compared with the literature[27,41], a more general channel fading model based on discrete semi-Markov switching topologies is established to characterize the possible effects of inter-agent signal transmission. The influences of channel fading between leader and follower, follower and follower agents are simultaneously considered to explore the influence of channel fading on system consensus, rather than considering only channel fading among follower agents in literature[27,41]. (II) As mentioned in the previous paragraph, discrete semi-Markov processes are more powerful in modeling ability and application range than Markov and continuous-time semi-Markov processes. For this reason, different from the Markov switching and continuous-time semi-Markov switching topologies adopted in[32–38], this paper employs a discrete semi-Markov process to describe the network topology switching among agents and switching for channel fading. A set of novel sufficient conditions to guarantee that leader-following mean square consensus of multi-agent systems under semi-Markov switching topologies is derived via a semi-Markov kernel approach. (III) The distributed consensus controller design scheme based on fading relative states is proposed to solve the $\mathcal{H}_\infty$ leader-following consensus control problem when the semi-Markov kernel of switching topologies is fully accessible and incompletely accessible, respectively. The rest of this paper is organized as follows. The preliminaries and problem formulation are given in section II. Section III presents the main results. Then, simulation examples are provided in Section IV. Finally, the conclusion and future work are introduced in Section V.

Notations: Denote that $\mathbb{R}^n$ and $\mathbb{R}^{n\times n}$ be the sets of n-dimensional vectors and $n\times n$ real matrices. $\mathbb{N}$ represents the sets of nonnegative integers. The sets of positive integers is denoted by $\mathbb{N}^+$. $\mathbb{N}_{\geq b_1}$ and $\mathbb{N}_{[b_1,b_2]}$ stand for the sets $\{c\in\mathbb{N}|b\geq b_1\}$ and $\{b\in\mathbb{N}|b_1\leq b\leq b_2\}$, respectively. A matrix $Q>0(<0)$ indicates it is positive (negative) definite. A $N\times N$ identity matrix is defined as I_N. Denote a symmetric term in a matrix by $*$. $\otimes$ refers to the Kronecker product. Moreover, $\mathbb{E}\{\cdot\}$ and $\|x\|$ represent the mathematical expectation operator and the Euclidean norm of the vectors. If not specifically stated, matrices and vectors have appropriate dimensions.

2. PRELIMINARIES AND PROBLEM FORMULATION

2.1. Graph theory

In this paper, we employed an undirected graph $\mathcal{G}=\{\mathcal{V},\mathcal{E},\mathcal{A}\}$ to depict the information interaction topology among N agents. $\mathcal{V}=\{v_1,v_2,...,v_N\}$ stands for the node sets, in which v_i is the ith agent. $\mathcal{E}\subset\mathcal{V}\times\mathcal{V}$ represents a set of edges. The adjacency matrix associated with graph $\mathcal{G}$ is denoted by $\mathcal{A}=[a_{ij}]\in\mathbb{R}^{N\times N}$. If node v_i can receive information from node v_j, there is an edge (v_j,v_i) between node v_i and node v_j. The elements a_{ij} of matrix $\mathcal{A}$ is weighted coefficient of edge (v_j,v_i), and $a_{ij}>0$, if $(v_j,v_i)\in\mathcal{E}$, otherwise, $a_{ij}=0$. Self-loop is not considered. The set of neighbors of node v_i can be represented as $\mathcal{N}_i=\{v_j\in\mathcal{V}|(v_j,v_i)\in\mathcal{E}\}$. The degree matrix of graph $\mathcal{G}$ is denoted as $\mathcal{D}=diag\{d_i\}\in\mathbb{R}^{N\times N}$, where $d_i=\sum_{j\in\mathcal{N}_i}a_{ij}$. Then, one can obtain that the Laplacian matrix is $\mathcal{L}=\mathcal{D}-\mathcal{A}$. Denote matrix $\mathcal{M}=diag\{m_i\}\in\mathbb{R}^{N\times N}$, where m_i stands for the information exchange of node v_i and leader node. If node v_i can access the information of the leader, $m_i=1$, otherwise, $m_i=0$.

2.2. Problem formulation

Consider a linear discrete-time multi-agent system consist of one leader and N follower agents:

$$\begin{aligned} x_i(k+1) &= Ax_i(k) + Bu_i(k), i = 1, 2, ..., N \\ x_0(k+1) &= Ax_0(k) \end{aligned} \tag{1}$$

where $x_i(k) \in \mathbb{R}^{n_x}$, $u_i(k) \in \mathbb{R}^m$, $x_0(k) \in \mathbb{R}^{n_x}$ are the state, input of the ith follower, the state of leader, respectively. Matrices $A \in \mathbb{R}^{n_x \times n_x}$ and $B \in \mathbb{R}^{n_x \times m}$ represent known constant system matrices.

In the real world, the communication network topology among agents is more likely to be time-varying. In this paper, a switching signal $\gamma(k)$ is used to characterize the topology switching among agents. $\{\gamma(k),\ k \in \mathbb{N}^+\}$ represents a discrete-time semi-Markov chain with values in a finite set $\mathcal{O} = \{1, 2, ..., O\}$.

To describe semi-Markov chain more formally, the following concepts are introduced. (I) The stochastic process $\{U_n, n \in \mathbb{N}^+\} \in \mathbb{N}^+$ is denoted as the mode index of the nth jump, in which taking values in $\mathcal{O}$. (II) The stochastic process $\{k_n, n \in \mathbb{N}^+\} \in \mathbb{N}^+$ represents the time instant of at the nth jump. (III) The stochastic process $\{S_n, n \in \mathbb{N}^+\} \in \mathbb{N}^+$ stands for the sojourn-time of mode U_{n-1} between the $(n-1)$th jump and nth jump, where $S_n = k_n - k_{n-1}$.

Definition 1[39] The stochastic process $\{(U_n, k_n), n \in \mathbb{N}^+\}$ is said to be a discrete-time homogeneous Markov renewal chain (MRC), if the following conditions holds for all $p,\ q \in \mathcal{O}, \tau \in \mathbb{N}^+,\ n \in \mathbb{N}^+$:

$$\begin{aligned} Pr\{U_{n+1} = q, S_{n+1} = \tau | U_0, U_1, ...U_n = p; k_0, ...k_n\} &= Pr\{U_{n+1} = q, S_{n+1} = \tau | U_n = p\} \\ &= Pr\{U_1 = q, S_1 | U_0 = p\}, \end{aligned}$$

where $\{U_n, n \in \mathbb{N}^+\}$ is named as the embedded Markov chain (EMC) of MRC.

Denote the matrix $\Pi(\tau) = [\pi_{pq}(\tau)] \in \mathbb{R}^{O \times O}$ as the discrete-time semi-Markov kernel with

$$\begin{aligned} \pi_{pq}(\tau) &= Pr\{U_{n+1} = q, S_{n+1} = \tau | U_n = p\} \\ &= \frac{Pr\{U_{n+1} = q, U_n = p\}}{Pr\{U_n = p\}} \frac{Pr\{U_{n+1} = q, S_{n+1} = \tau, U_n = p\}}{U_{n+1} = q, U_n = p} \\ &= \theta_{pq}\omega_{pq}(\tau) \end{aligned} \tag{2}$$

where $\sum_{\tau=0}^{\infty} \sum_{q \in O} \pi_{pq}(\tau) = 1$ and $0 < \pi_{pq}(\tau) < 1$ with $\pi_{pq}(0) = 0$. The transition probability of EMC is defined by $\theta_{pq} = Pr\{U_{n+1} = q | U_n = p\}, \forall p, q \in \mathcal{O}$ with $\theta_{pp} = 0$, and the probability density function of sojourn-time is provided by $\omega_{pq}(\tau) = Pr\{S_{n+1} | U_{n+1} = q, U_n = p\}, \omega_{pq}(0) = 0$.

Remark 1 References[35,37,38] studied the leader-following consensus and containment control problems for multi-agent systems with semi-Markov switching topologies, respectively. A continuous-time semi-Markov jump process is employed to describe the switching of the topology. Accordingly, the probability density function of the sojourn-time can only be of a fixed probability distribution type for the different modes. This limits its practical application. In this paper, a discrete semi-Markov chain is introduced to characterize topology switching among agents. The introduced probability density function of sojourn time depends on both the current mode and the next mode, so that different parameters of the same distribution or different types of probability distributions can coexist. Hence, the probability density function introduced in this paper is more applicable than that in the literature[35,37,38].

Definition 2[39] The stochastic process $\{\gamma(k), k \in \mathbb{N}^+\}$ is said to be an semi-Markov chain associated with MRC $\{(U_n, k_n), n \in \mathbb{N}^+\}$, if $\gamma(k) = U_{\mathbb{N}(k)}, \forall k \in \mathbb{N}^+, \mathbb{N}(k) = \max\{n \in \mathbb{N}^+ | k_n \leq k\}$.

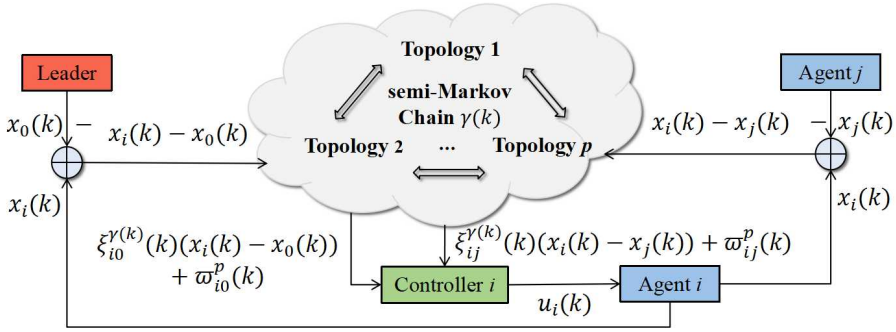


Figure 1. Multi-agent systems with channel fading under semi-Markov switching topologies

Based on the above introduction to the semi-Markov chain, the switching topology among agents in this paper can be denoted as $\mathcal{G}_{\gamma(k)}$. For convenience, let $\mathcal{G}_{\gamma(k)} = \mathcal{G}_p,\ p \in O$.

Every topological graph $\mathcal{G}_p$ is an undirected graph. Then, the Laplacian matrix of the graph $\mathcal{G}_p$, the adjacency matrix of the leader and the follower are denoted by $\mathcal{L}_p \in \mathbb{R}^{N\times N}$, $\mathcal{A}_p \in \mathbb{R}^{N\times N}$, and $\mathcal{M}_p \in \mathbb{R}^{N\times N}$.

In practice, the communication process between an agent and its neighbors is often affected by channel noise and fading. Motivated by the channel fading model in [28,29], in this paper we assume that each agent obtains relative state information from its neighbors through fading channels. Correspondingly, the channel fading model among agents can be expressed as

$$x_{ij}^{rel}(k) = \xi_{ij}(k)(x_i(k) - x_j(k)) + \varpi_{ij}(k),\ \ x_{i0}^{rel}(k) = \xi_{i0}(k)(x_i(k) - x_0(k)) + \varpi_{i0}(k) \tag{3}$$

where $\xi_{ij}(k)$, $\xi_{i0}(k)$ represent the channel fading of the follower-to-follower and the leader-to-follower, respectively. $\varpi_{ij}(k)$ and $\varpi_{i0}(k)$ are disturbances in the channel. Based on the above channel fading model, we design a distributed consensus controller as follows:

$$u_i(k) = K_p\{\sum_{j=1}^{N} a_{ij}^{p}(k)[\xi_{ij}^{p}(k)(x_i(k) - x_j(k)) + \varpi_{ij}^{p}(k)] + m_i^{p}(k)[\xi_{i0}^{p}(k)(x_i(k) - x_0(k)) + \varpi_{i0}^{p}(k)]\} \tag{4}$$

where $K_p \in \mathbb{R}^{m\times n_x}$ is the controller gain to be determined, $p \in O$.

Figure 1 shows the frame diagram of the system considered in this paper. Under the semi-Markov switching communication network topology, the channel fading between agents varies randomly with the switching of the topology. Each agent generates control inputs based on the relative information obtained from neighbors through fading and interference, thereby further controlling the entire system to achieve consensus.

Remark 2 Compared with the channel fading model established in the literature [27–29], the channel fading model introduced in this paper not only considers the fading influence from leader to follower and follower to follower, but also introduces the effect of channel interference on signal transmission and the influence of topology switching on the statistical characteristics of fading coefficients. This means that the models presented in this paper are more general than those in the previous literature. When $\varpi_{ij}(k) = \varpi_{i0}(k) = 0$, $\xi_{i0}(k) \equiv 1$ in equation (3), the model degenerates to the case in [27]. If the values of $\xi_{ij}(k)$ and $\xi_{i0}(k)$ are 0 or 1 and $\varpi_{ij}(k) = \varpi_{i0}(k) = 0$, the channel fading model (3) is reduced to a packet loss model. This also indicates that the channel fading model proposed in this paper is more general.

Define the consensus error as $\delta_i(k) = x_i(k) - x_0(k)$, $i = 1, 2, ..., N$. Combining system (1) and consensus

protocol (4), the dynamics of $\delta_i(k)$ can be given by

$$\begin{aligned}\delta_i(k+1) = A\delta_i(k) + BK_p\{\sum_{j=1}^{N} a_{ij}^p(k)[\xi_{ij}^p(k)(\delta_i(k)-\delta_j(k)) + \varpi_{ij}^p(k)] \\ + m_i^p(k)[\xi_{i0}^p(k)\delta_i(k) + \varpi_{i0}^p(k)]\}\end{aligned} \tag{5}$$

with $\delta_i(k) = [\delta_1^T(k), \delta_2^T(k), ..., \delta_N^T(k)]^T$. Assume that all channel fading are identical, ie., $\xi_{ij}(k) = \xi(k)$, $\xi_{i0}(k) = \xi_0(k)$, $\varpi_{ij}(k) = \varpi(k)$, and $\varpi_{i0}(k) = \varpi_0(k)$ for $k \geq 0, i, j = 1, 2, ..., N$. Then, the above equation can be rewritten in compact form as follows:

$$\delta(k+1) = [I_N \otimes A + (\xi_p(k)\mathcal{L}_p + \xi_{0p}(k)\mathcal{M}_p) \otimes BK_p]\delta(k) + [I_N \otimes (BK_p)]\omega(k) \tag{6}$$

where $\omega(k) = [\omega_1^T(k), \omega_2^T(k), ..., \omega_N^T(k)]^T$, $\omega_i(k) = \sum_{j=1}^N a_{ij}^p(k)\varpi_p(k) + m_i^p(k)\varpi_{0p}(k)$, $i = 1, 2, ..., N$.

Before the subsequent analysis, some definition and assumptions are introduced.

Definition 3 The systems (6) is said to achieve leader-following consensus in mean square sense, if the system (6) is σ-error mean square stable.

Definition 4[39] The dynamic system (6) is said to be σ-error mean square stable, if the following conditions hold

$$\lim_{k\to\infty} \mathbb{E}\left[\|\delta_i(k)\|^2\right]_{\delta_i(0),\gamma(0),S_{n+1}\leq T_{max}^p|U_n=p} = 0 \tag{7}$$

for given the upper bound of sojourn-time $T_{max}^p \in \mathbb{N}^+$ and any initial conditions $\delta_i(0)$, $\gamma(0) \in \mathcal{O}$, $i \in \{1, 2, ..., N\}$, $\omega(k) = 0$.

Assumption 1 Every possible undirected graph $\mathcal{G}_{\gamma(k)}$, $\gamma(k) = p \in \mathcal{O}$ is connected.

Assumption 2 The mean and variance of stochastic variables $\{\xi_p(k)\}$ and $\{\xi_{0p}(k)\}$ are $\mathbb{E}\{\xi_p(k)\} = \mu_p$, $\mathbb{E}\{\xi_{0p}(k)\} = \mu_{0p}$, $\mathbb{E}\{(\xi_p(k)-\mu_p)^2\} = \sigma_p^2$, and $\mathbb{E}\{(\xi_{0p}(k)-\mu_{0p})^2\} = \sigma_{0p}^2$.

According to the above analysis and discussion, it can be found that the leader-following consensus of system (1) under the semi-Markov switching topology $\mathcal{G}_p$ is equivalent to the mean square stability of system (6). Since $\varpi_p(k)$ and $\varpi_{0p}(k)$ are interference in the channel, we can treat the last term in equation (6) as a disturbance. To tackle with the disturbance, the following control output are given

$$z(k) = (I_N \otimes E)\delta(k) \tag{8}$$

with $z_i(k) = E(x_i(k) - x_0(k))$ and $z(k) = [z_1^T(k), z_2^T(k), ..., z_N^T(k)]^T$. $E \in \mathbb{R}^{n_x \times n_x}$ is a known constant matrix. Then, the consensus problem of system (1) is transformed into a $\mathcal{H}_\infty$ control problem of the system (6).

Consequently, the objective of this paper is to design the distributed consensus controller such that the following two conditions are satisfied:
(I) when $\omega(k) = 0$, the condition (7) holds;
(II) the inequality $\mathbb{E}\left\{\sum_{n=0}^{\infty}\sum_{k=k_n}^{k_{n+1}-1}[\|z(k)\|^2 - \hat{\gamma}^2\|\omega(k)\|^2]\right\} < 0$ holds for zero-initial condition and any nonzero $\omega(k) \in L_2(0, +\infty)$.

3. MAIN RESULTS

3.1. Consensus and $\mathcal{H}_\infty$ performance analysis

In the subsection, the leader-following consensus and $\mathcal{H}_\infty$ performance analysis of systems (6) and (8) are given. Sufficient conditions for mean square consensus are derived via stochastic Lyapunov function.

Theorem 1 Given scalar $T_{max}^p \in \mathbb{N}_{\geq 1}$, if there exist a scalar $\hat{\gamma} > 0$ and a set of symmetric matrices $P_p^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $P_q^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $\nu \in \mathbb{N}_{[0,T_{max}^p]}$ such that the following inequalities

$$\Psi = \begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * & * & * \\ 0 & -\hat{\gamma}^2 I_{nN} & * & * \\ \Xi & \Psi_{23} & \mathcal{P}_p^{(\nu)} & * \\ (I_N \otimes E) & 0 & 0 & I_{n_x N} \end{bmatrix} < 0 \tag{9}$$

$$\Phi = \begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \Xi & \tilde{\mathcal{P}}_q^{(0)} \end{bmatrix} < 0 \tag{10}$$

hold for any $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[1,T_{max}^p]}$, then the system (6) is leader-following consensus in mean square sense and possesses a $\mathcal{H}_\infty$ performance index $\hat{\gamma}$, where
$\Xi = [\bar{A}_p^T \ \sigma_p \bar{\mathcal{L}}_p^T \ \sigma_{0p} \bar{\mathcal{M}}_p^T]^T$, $\Psi_{23} = [(I_N \otimes BK_p)^T \ 0 \ 0]^T$, $\bar{A}_p = [I_N \otimes A + (\mu_p \mathcal{L}_p + \mu_{0p} M_p) \otimes BK_p]$,
$\bar{\mathcal{M}}_p = \mathcal{M}_p \otimes BK_p$, $\bar{\mathcal{L}}_p = \mathcal{L}_p \otimes BK_p$, $\mathcal{P}_p^{(\nu)} = diag\{-(I_N \otimes P_p^{-(\nu)}), -(I_N \otimes P_p^{-(\nu)}), -(I_N \otimes P_p^{-(\nu)})\}$,
$\tilde{\mathcal{P}}_q^{(0)} = diag\{-\tilde{\mathcal{P}}, -\tilde{\mathcal{P}}, -\tilde{\mathcal{P}}\}$, $\tilde{\mathcal{P}} = \sum_{\nu=1}^{T_{max}^p} \sum_{q \in \mathcal{O}} \frac{\pi_{pq}(\nu)}{\Omega_p} (I_N \otimes P_q^{-(0)})$, $\Omega_p = \sum_{\nu=1}^{T_{max}^p} \sum_{q \in \mathcal{O}} \pi_{pq}(\nu)$, $I_{n_x N} = I_{n_x} \otimes I_N$.

Proof Construct a stochastic Lyapunov function candidate as $V(\delta(k), \gamma(k), \nu(k)) = \delta^T(k)(I_N \otimes P_{\gamma(k)}^{(\nu(k))})\delta(k)$, where $\gamma(k) = p \in \mathcal{O}$, $k \in [k_n, k_{n+1})$. $\nu(k) = k - k_n$ represents the running time of the current mode for topology $\mathcal{G}_p$. Assume that $\gamma(k_n) = p$, $\gamma(k_{n+1}) = q$, $\omega(k) \equiv 0$. Then, one can obtain along the solution of systems (6) and (8)

$$\begin{aligned} &\mathbb{E}\Big\{V(\delta(k+1), \gamma(k+1), \nu(k+1))\Big\}|_{\delta(k),\gamma(k)=p} - V(\delta(k), \gamma(k), \nu(k)) \\ =&\mathbb{E}\Big\{\delta^T(k+1)(I_N \otimes P_p^{(\nu(k+1))})\delta(k+1)\Big\}|_{\delta(k),\gamma(k)=p} - \delta^T(k)(I_N \otimes P_p^{(\nu(k))})\delta(k) \\ =&\mathbb{E}\Big\{\delta^T(k)\tilde{\mathcal{A}}_p^T(I_N \otimes P_p^{(\nu(k+1))})\tilde{\mathcal{A}}_p\delta(k)\Big\} - \delta^T(k)(I_N \otimes P_p^{(\nu(k))})\delta(k) \\ =&\delta^T(k)\Big[\bar{A}_p^T(I_N \otimes P_p^{(\nu(k+1))})\bar{A}_p + \sigma_p^2\bar{\mathcal{L}}_p^T(I_N \otimes P_p^{(\nu(k+1))})\bar{\mathcal{L}}_p + \sigma_{0p}^2\bar{\mathcal{M}}_p^T(I_N \otimes P_p^{(\nu(k+1))}) \\ &\times \bar{\mathcal{M}}_p - (I_N \otimes P_p^{(\nu(k))})\Big]\delta(k) \end{aligned} \tag{11}$$

where $\tilde{\mathcal{A}}_p = I_N \otimes A + (\xi_p(k)\mathcal{L}_p + \xi_{0p}(k)\mathcal{M}_p) \otimes BK_p$, $\tilde{\mathcal{A}}_p \in \mathbb{R}^{n_x N \times n_x N}$. It can be found from the above equation that $\Delta V(\delta(k), \gamma(k), \nu(k)) < 0$ if and only if $\Sigma \in \mathbb{R}^{n_x N \times n_x N} < 0$, $\Sigma = \bar{A}_p^T(I_N \otimes P_p^{(\nu(k+1))})\bar{A}_p + \sigma_p^2\bar{\mathcal{L}}_p^T(I_N \otimes P_p^{(\nu(k+1))})\bar{\mathcal{L}}_p + \sigma_{0p}^2\bar{\mathcal{M}}_p^T(I_N \otimes P_p^{(\nu(k+1))})\bar{\mathcal{M}}_p - (I_N \otimes P_p^{(\nu(k))})$. Furthermore, it can be proved that the following inequality holds:

$$V(\delta(k+1), \gamma(k+1), \nu(k+1)) - V(\delta(k), \gamma(k), \nu(k)) < 0 \tag{12}$$

On the other hand, the stability of the system during mode switching needs to be considered. It can be derived along the trajectory of systems (6) and (8) that

$$\begin{aligned}
&\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))\Big\}-V(\delta(k_{n+1}-1),\gamma(k_{n+1}-1),\nu(k_{n+1}-1))\\
=&\mathbb{E}\Big\{\delta^T(k_{n+1}-1)\tilde{\mathcal{A}}_p^T(I_N\otimes P_q^{(\nu(k_{n+1}))})\tilde{\mathcal{A}}_p\delta(k_{n+1}-1)\Big\}-\delta^T(k_{n+1}-1)(I_N\otimes P_p^{(\nu(k_{n+1}-1))})\delta(k_{n+1}-1)\\
=&\mathbb{E}\Big\{\delta^T(k_{n+1}-1)\Sigma_{S_{n+1}=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}}\frac{\pi_{pq}(S_{n+1})}{\Omega_p}\tilde{\mathcal{A}}_p^T(I_N\otimes P_q^{(\nu(k_{n+1}))})\tilde{\mathcal{A}}_p\delta(k_{n+1}-1)\Big\}\\
&-\delta^T(k_{n+1}-1)(I_N\otimes P_p^{(\nu(k_{n+1}-1))})\delta(k_{n+1}-1)\\
=&\delta^T(k_{n+1}-1)\Big[\Sigma_{S_{n+1}=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}}\frac{\pi_{pq}(S_{n+1})}{\Omega_p}\big(\bar{A}_p^T\mathcal{P}\bar{A}_p+\sigma_p^2\bar{\mathcal{L}}_p^T\mathcal{P}\bar{\mathcal{L}}_p+\sigma_{0p}^2\bar{\mathcal{M}}_p^T\mathcal{P}\bar{\mathcal{M}}_p\big)\\
&-(I_N\otimes P_p^{(\nu(k_{n+1}-1))})\Big]\delta(k_{n+1}-1)
\end{aligned}\tag{13}$$

with $\mathcal{P}=(I_N\otimes P_q^{(\nu(k_{n+1}))})$. By means of the condition (10) and Schur complement lemma, one can get that $\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))\Big\}-V(\delta(k_{n+1}-1),\gamma(k_{n+1}-1),\nu(k_{n+1}-1))<0$.
Together with the condition (12), there exists a positive constant β such that gives

$$\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))-V(\delta(k_n),\gamma(k_n),\nu(k_n))\Big\}<-\beta\delta^T(k_n)\delta(k_n)\tag{14}$$

Then, it follows that

$$\mathbb{E}\big\{\delta^T(k_n)\delta(k_n)\big\}<-\frac{1}{\beta}\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))-V(\delta(k_n),\gamma(k_n),\nu(k_n))\Big\}$$

Summing both sides of the above equation from 0 to l yields

$$\mathbb{E}\big\{\Sigma_{n=0}^l\delta^T(k_n)\delta(k_n)\big\}<\frac{1}{\beta}\Big[V(\delta(k_0),\gamma(k_0),\nu(k_0))-\mathbb{E}\big\{V(\delta(k_{l+1}),\gamma(k_{l+1}),\nu(k_{l+1}))\big\}\Big]\tag{15}$$

Let $l\to\infty$, then $\lim\limits_{l\to\infty}\mathbb{E}\big\{\Sigma_{n=0}^l\delta^T(k_n)\delta(k_n)\big\}<\frac{1}{\beta}V(\delta(k_0),\gamma(k_0),\nu(k_0))$. Then, we have $\lim\limits_{n\to\infty}\mathbb{E}\{\delta^T(k_n)\delta(k_n)\}=\lim\limits_{k\to\infty}\mathbb{E}\{\delta^T)(k\delta(k)\}=0$. Furthermore, it can be shown that $\lim\limits_{k\to\infty}\mathbb{E}\big\{\|\delta_i(k)\|^2\big\}=0$. By Definition 3 and Definition 4, systems (6) and (8) with $\omega(k)\equiv 0$ is leader-following mean square consensus.

Next, the $\mathcal{H}_\infty$ consensus performance of the underlying system is discussed. Then, for $\omega(k)\neq 0$, we have

$$\begin{aligned}
&\mathbb{E}\Big\{V(\delta(k+1),\gamma(k+1),\nu(k+1))\Big\}|_{\delta(k),\gamma(k)=p}-V(\delta(k),\gamma(k),\nu(k))+z^T(k)z(k)-\hat{\gamma}^2\omega^T(k)\omega(k)\\
=&\mathbb{E}\Big\{\big[\delta^T(k)\tilde{\mathcal{A}}_p^T+\omega^T(k)[I_N\otimes(BK_p)]^T\big](I_N\otimes P_p^{(\nu(k+1))})\big[\tilde{\mathcal{A}}_p\delta(k)+[I_N\otimes(BK_p)]\omega(k)\big]\Big\}|_{\delta(k),\gamma(k)=p}\\
&-\delta^T(k)(I_N\otimes P_p^{(\nu(k))})\delta(k)+\delta^T(k)(I_N\otimes E)^T(I_N\otimes E)\delta(k)-\hat{\gamma}^2\omega^T(k)\omega(k)
\end{aligned}\tag{16}$$

Denote $\eta(k)=[\delta^T(k)\ \ \omega^T(k)]^T$, then the equation (16) can be transformed into

$$\begin{aligned}
&\mathbb{E}\{\Delta V(\delta(k),\gamma(k),\nu(k))\}+z^T(k)z(k)-\hat{\gamma}^2\omega^T(k)\omega(k)\\
=&\mathbb{E}\Big\{V(\delta(k+1),\gamma(k+1),\nu(k+1))\Big\}|_{\delta(k),\gamma(k)=p}-V(\delta(k),\gamma(k),\nu(k))+z^T(k)z(k)-\hat{\gamma}^2\omega^T(k)\omega(k)\\
=&\eta^T(k)\Theta\eta(k)
\end{aligned}\tag{17}$$

where $\Theta = \begin{bmatrix} \Sigma + (I_N \otimes E)^T (I_N \otimes E) & * \\ (I_N \otimes (BK_p))^T (I_N \otimes P_p^{(\nu(k+1))}) \bar{A}_p & [I_N \otimes (BK_p)]^T (I_N \otimes P_{(\nu(k))}) [I_N \otimes (BK_p)] - \hat{\gamma}^2 I_{nN} \end{bmatrix}$

According to Schur complement lemma and the condition (9), it gives that $\Theta \in \mathbb{R}^{2n_x N \times 2n_x N} < 0$. This implies that

$$\mathbb{E}\{\Delta V(\delta(k), \gamma(k), \nu(k))\} + z^T(k) z(k) - \hat{\gamma}^2 \omega^T(k) \omega(k) < 0$$

Further, we can obtain

$$\mathbb{E}\left\{\Sigma_{n=0}^{\infty} \Sigma_{k=k_n}^{k_{n+1}-1} [\|z(k)\|^2 - \hat{\gamma}^2 \|\omega(k)\|^2]\right\} < -(\mathbb{E}\left\{\lim_{k \to \infty} V(\delta(k), \gamma(k), \nu(k))\right\} - V(\delta(0), \gamma(0), \nu(0))) < 0$$

under the zero initial conditions. Thus, the $\mathcal{H}_\infty$ performance condition (II) holds. This proof is completed.

In Theorem 1, the leader-following consensus and $\mathcal{H}_\infty$ performance of the system (6) with channel fading is analyzed, in which data transmission between agents takes into account not only channel fading but also channel interference. Assuming that the channel interference $\varpi_{ij}^p(k) = 0$, $\varpi_{i0}^p(k) = 0$, that is $\omega(k) = 0$. Then, the system (6) can be reduced to

$$\delta(k+1) = [I_N \otimes A + (\xi_p(k) \mathcal{L}_p + \xi_{0p}(k) \mathcal{M}_p) \otimes BK_p] \delta(k) \tag{18}$$

In this case, the following corollary gives the leader-following mean square consensus analysis of the system (18) under no-channel interference fading model.

Corollary 1 Given scalar $T_{max}^p \in \mathbb{N}_{\geq 1}$, if there exist a sets of symmetric matrices $P_p^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $P_q^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $\nu \in \mathbb{N}_{[0, T_{max}^p]}$ such that the following inequalities

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \Xi & \mathcal{P}_p^{(\nu)} \end{bmatrix} < 0$$

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \Xi & \tilde{\mathcal{P}}_q^{(0)} \end{bmatrix} < 0$$

hold for $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[1, T_{max}^p]}$, $\tau \in \mathbb{N}_{[1, T_{max}^p]}$, then the system (18) is leader-following mean square consensus, where Ξ, $\mathcal{P}_p^{(\nu)}$, and $\tilde{\mathcal{P}}_q^{(0)}$ have been defined in Theorem 1.

The proof of Corollary 1 is similar to that of Theorem 1, so it is omitted.

3.2. Consensus controller gain design

Although Theorem 1 addresses a family of leader-following consensus conditions, these conditions cannot be directly utilized to solve controller gains K_p. Aiming at solving the leader-following consensus control problem of systems (6) and (8) under switching topologies, sufficient conditions on the existence of the desired controller gains are presented in the following theorem.

Theorem 2 Given a scalar $T_{max}^p \in \mathbb{N}_{\geq 1}$, if there exist a scalar $\hat{\gamma} > 0$ and sets of symmetric matrices $P_p^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $P_q^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, and matrices $Z_p \in \mathbb{R}^{n_x \times n_x}$, $Y_p \in \mathbb{R}^{n_x \times n_x}$, $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[0, T_{max}^p]}$ such that the following inequalities

$$\bar{\Psi} = \begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * & * & * \\ 0 & -\hat{\gamma}^2 I_{n_x N} & * & * \\ \bar{\Xi} & \bar{\Psi}_{23} & -\mathcal{P}_p^{-(\nu)} - \mathcal{Z}_p - \mathcal{Z}_p^T & * \\ (I_N \otimes E) & 0 & 0 & -I_{n_x N} \end{bmatrix} < 0 \tag{19}$$

$$\bar{\Phi} = \begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \bar{\Xi} & \Upsilon_q^{(0)} \end{bmatrix} \prec 0 \tag{20}$$

hold for any $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[1,T_{max}^p]}$, where

$\bar{\Xi} = [\tilde{A}_p^T \ \ \sigma_p \tilde{\mathcal{L}}_p^T \ \ \sigma_{0p} \tilde{\mathcal{M}}_p^T]^T$, $\bar{\Psi}_{23} = [(I_N \otimes Y_P) \ \ 0 \ \ 0]^T$, $\tilde{A}_p = I_N \otimes A + (\mu_p \mathcal{L}_p + \mu_{0p} \mathcal{M}_p) \otimes Y_p$,
$\tilde{\mathcal{L}}_p = \mathcal{L}_p \otimes Y_p$, $\tilde{\mathcal{M}}_p = \mathcal{M}_p \otimes Y_p$, $\mathcal{Z}_p = diag\{I_N \otimes Z_p, I_N \otimes Z_p, I_N \otimes Z_p\} \in \mathbb{R}^{3n_x N \times 3n_x N}$, $\Upsilon_q^{(0)} = diag\{\Upsilon, \Upsilon, \Upsilon\}$,
$\Upsilon = \Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}} \frac{\pi_{pq}(\nu)}{\Omega_p} (I_N \otimes P_q^{(0)} - I_N \otimes Z_p - I_N \otimes Z_p^T)$,

then the system (6) is leader-following consensus in mean square sense and possess a $\mathcal{H}_\infty$ performance index $\hat{\gamma}$. Moreover, the controller gains are given by $K_p = (B^T B)^{-1} B^T (Z_p^T)^{-1} Y_p$.

Proof Performing congruence transformations $diag\{I_{n_x N}, I_{n_x N}, \mathcal{Z}_p^T, I_{n_x N}\}$ to (9), $diag\{I_{nN}, \mathcal{Z}_p^T\}$ to (10), one can obtain

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * & * & * \\ 0 & -\hat{\gamma}^2 I_{n_x N} & * & * \\ \mathcal{Z}_p^T \Xi & \mathcal{Z}_p^T \Psi_{23} & \mathcal{Z}_p^T \mathcal{P}_p^{(\nu)} \mathcal{Z}_p & * \\ I_N \otimes E & 0 & 0 & -I_{n_x N} \end{bmatrix} \prec 0 \tag{21}$$

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \mathcal{Z}_p \Xi & \mathcal{Z}_p^T \tilde{\mathcal{P}}_q^{(0)} \mathcal{Z}_p \end{bmatrix} \prec 0 \tag{22}$$

According to the literature [40], we can obtain that for the positive definite matrix $U \in \mathbb{R}^{n \times n}$ and the real matrix $X \in \mathbb{R}^{n \times n}$, there must be $(U-X)^T U^{-1} (U-X) \geq 0$. Performing cholesky decomposition on the matrix U, there must be a lower triangular matrix $L \in \mathbb{R}^{n \times n}$ such that $U = LL^T$. Further, we can get $(U-X)^T U^{-1} (U-X) = Q^T Q \geq 0$, where $Q = L^T - L^{-1} X$. When the matrix Q is full rank, $(U-X)^T U^{-1} (U-X) \succ 0$ holds. Based on this, the inequalities

$$\begin{aligned} {}^T(I_N \otimes P_p^{-(\nu)})[(I_N \otimes P_p^{(\nu)}) - (I_N \otimes Z_p)] \geq 0 \\ [(I_N \otimes P_q^{(0)}) - (I_N \otimes Z_p)]^T (I_N \otimes P_q^{-(0)})[(I_N \otimes P_q^{(0)}) - (I_N \otimes Z_p)] \geq 0 \end{aligned}$$

can ensure

$$\begin{aligned} -(I_N \otimes Z_p^T P_p^{-(\nu)} Z_p) \preceq (I_N \otimes P_p^{(\nu)}) - (I_N \otimes Z_p) - (I_N \otimes Z_p^T), \\ -(I_N \otimes Z_p^T P_q^{-(0)} Z_p) \preceq (I_N \otimes P_q^{(0)}) - (I_N \otimes Z_p) - (I_N \otimes Z_p^T) \end{aligned} \tag{23}$$

for any $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[1,T_{max}^p]}$. Then, it can be shown that the inequalities (21) and (22) can ensure that the conditions (19) and (20) hold. Furthermore, it implies that the inequalities (9) and (10) are true. By Theorem 1, we can get that systems (6) and (8) is leader-following consensus in mean square sense with a $\mathcal{H}_\infty$ performance index $\hat{\gamma}$ under the controller (4) and the controller gains $K_p = (B^T B)^{-1} B^T (Z_p^T)^{-1} Y_p$. This proof is completed.

Supposing that the channel inferences $\varpi_{ij}^p(k)$ and $\varpi_{0i}^p(k)$ in (4) are ignored, that is, $\varpi_{ij}^p(k) = 0$, $\varpi_{i0}^p(k) = 0$. Then, we have $\omega(k) = 0$ in (6). Consequently, the leader-following consensus control of system (18) under semi-Markov switching topology is realized in the following corollary.

Corollary 2 Given a scalar $T_{max}^p \in \mathbb{N}_{\geq 1}$, if there exist a sets of symmetric matrices $P_p^{(\nu)} \in \mathbb{R}^{n_x \times n_x} \succ 0$, $P_q^{(\nu)} \in \mathbb{R}^{n_x \times n_x} \succ 0$, and matrices $Z_p \in \mathbb{R}^{n_x \times n_x}$, $Y_p \in \mathbb{R}^{n_x \times n_x}$, $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[0,T_{max}^p]}$ such that the inequalities hold:

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \bar{\Xi} & -\mathcal{P}_p^{-(\nu)} - \mathcal{Z}_p - \mathcal{Z}_p^T \end{bmatrix} \prec 0$$

$$\begin{bmatrix} -(I_N \otimes P_p^{(\nu-1)}) & * \\ \bar{\Xi} & \Upsilon_q^{(0)} \end{bmatrix} \prec 0$$

for any $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[1,T^p_{max}]}$, then the system (18) is leader-following consensus in mean square sense under the controller gains $K_p = (B^TB)^{-1}B^T(Z_p^T)^{-1}Y_p$, where $\mathcal{P}_p^{(\nu)}$, $\bar{\Xi}$, $\mathcal{Z}_p$, and $\Upsilon_q^{(0)}$ are defined in Theorem 1 and 2. By employing the same approach as Theorem 2, the Corollary 2 can be proved directly, omitting the proof process here.

Remark 3 Theorem 2 realizes the consistent controller design of systems (6) and (8) under the channel fading model (3). A set of sufficient conditions to ensure the consensus of systems and the existence of controller gains is established based on the linear matrix inequality form. To solve the controller gain matrix K_{pm}, some unknown variables are introduced into the inequality conditions of Theorem 2. The computational complexity of solving the inequality conditions in Theorem 2 can be analyzed according to the total number of unknown variables. It can be obtained by calculation that the total number of unknown variables in Theorem 2 is $var_t = \Sigma_{p=1}^{O}[(T^p_{max}+1)n_x^2]+\Sigma_{p=1}^{O}[Nn_x^2+2(Nn_x)^2]$. It can be found that with the increase of the number O of topological modes, the upper bound T^p_{max} of the sojourn time of mode p, and the number of agents N, the computational complexity of solving Theorem 2 also increases accordingly. Assuming that the upper bound $T^p_{max}, p \in \mathcal{O}$ of the sojourn time is the same in each topological mode, the total number of unknown variables is $var_t = O(T^p_{max}+1)n_x^2+O[Nn_x^2+2(Nn_x)^2]$. Similarly, the computational complexity of solving the inequality conditions in Theorem 3 can also be analyzed.

3.3. Extension results

In this subsection, it is assumed that the information of the semi-Markov kernel $\Pi(\tau)$ is not completely accessible. By using the similar method in [40], the index set $\mathcal{O}$ of the semi-Markov chain $\gamma(k)$ can be partitioned into the following form:

$\mathcal{O}_{ap} = \{q \in \mathcal{O} | \text{if } \pi_{pq}(\tau) \text{ is accessible }\}$,
$\mathcal{O}_{bp} = \{q \in \mathcal{O} | \text{if } \pi_{pq}(\tau) \text{ is inaccessible }\}$,
$\mathcal{O}_{cp} = \{q \in \mathcal{O} | \text{if } \omega_{pq}(\tau) \text{ is accessible, } \theta_{pq} \text{ is inaccessible }\}$,
$\mathcal{O}_{dp} = \{q \in \mathcal{O} | \text{if } \omega_{pq}(\tau) \text{ is inaccessible, } \theta_{pq} \text{ is accessible }\}$,
$\mathcal{O}_{ep} = \{q \in \mathcal{O} | \text{if } \omega_{pq}(\tau) \text{ is inaccessible, } \theta_{pq} \text{ is inaccessible }\}$,
where $\mathcal{O} = \mathcal{O}_{ap} \cup \mathcal{O}_{bp}$, $\mathcal{O}_{bp} = \mathcal{O}_{cp} \cup \mathcal{O}_{dp} \cup \mathcal{O}_{ep}$, $\mathcal{O}_{ap} \cap \mathcal{O}_{bp} = \emptyset$, $\mathcal{O}_{cp} \cap \mathcal{O}_{dp} = \emptyset$.

In this paper, only the case of $\mathcal{O} = \mathcal{O}_{ap} \cup \mathcal{O}_{ep}$ for incompletely accessible semi-Markov kernel is considered. In other words, the transition probability θ_{pq} of EMC and the probability density function $\omega_{pq}(\tau)$ of sojourn-time are partially accessible. Before presenting the results of this subsection, we make the following assumptions, which are crucial for subsequent derivations.

Assumption 3 Given a positive scalar ρ, the selection of the upper bounds T^p_{max} for sojourn time can be guaranteed by the following prerequisite:

$$\Sigma_{\tau=1}^{T^p_{max}}\omega_{pq}(\tau) \geq \rho, 0 < \rho < 1, \forall p \in \mathcal{O}, \forall q \in \mathcal{O}_{dp} \cup \mathcal{O}_{ep} \tag{24}$$

Then, the following Theorem proposes the leader-following mean-square consensus conditions for systems (6) and (8) under incompletely accessible semi-Markov kernel of switching topologies.

Theorem 3 Given a scalar $T^p_{max} \in \mathbb{N}_{\geq 1}$, if there exist a scalar $\hat{\gamma} > 0$ and sets of symmetric matrices $\tilde{P}_p^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, $\tilde{P}_q^{(\nu)} \in \mathbb{R}^{n_x \times n_x} > 0$, and matrices $\tilde{Z}_p \in \mathbb{R}^{n_x \times n_x}$, $\tilde{Y}_p \in \mathbb{R}^{n_x \times n_x}$, $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[0,T^p_{max}]}$ such that the following inequalities

$$\tilde{\Psi} = \begin{bmatrix} -(I_N \otimes \tilde{P}_p^{(\nu-1)}) & * & * & * \\ 0 & -\hat{\gamma}^2 I_{n_x N} & * & * \\ \check{\Xi} & \tilde{\Psi}_{23} & -\mathcal{P}_p^{-(\nu)} - \tilde{\mathcal{Z}}_p - \tilde{\mathcal{Z}}_p^T & * \\ (I_N \otimes E) & 0 & 0 & -I_{n_x N} \end{bmatrix} < 0 \tag{25}$$

$$\tilde{\Phi} = \begin{bmatrix} -(I_N \otimes \tilde{P}_p^{(\nu-1)}) & * \\ \tilde{\Xi} & \tilde{\Upsilon}_q^{(0)} \end{bmatrix} \prec 0 \tag{26}$$

hold for any $p \in \mathcal{O}_{ap}$, $q \in \mathcal{O}_{ep}$, $\nu \in \mathbb{N}_{[1,T_{max}^p]}$, where
$\check{\Xi} = [\check{A}_p^T \ \sigma_p \check{\mathcal{L}}_p^T \ \sigma_{0p} \check{\mathcal{M}}_p^T]^T$, $\tilde{\Psi}_{23} = [(I_N \otimes \tilde{Y}_P) \ 0 \ 0]^T$, $\check{A}_p = I_N \otimes A + (\mu_p \mathcal{L}_p + \mu_{0p} \mathcal{M}_p) \otimes \tilde{Y}_p$,
$\check{\mathcal{L}}_p = \mathcal{L}_p \otimes \tilde{Y}_p$, $\check{\mathcal{M}}_p = \mathcal{M}_p \otimes \tilde{Y}_p$, $\tilde{\mathcal{Z}}_p = diag\{I_N \otimes \tilde{Z}_p, I_N \otimes \tilde{Z}_p, I_N \otimes \tilde{Z}_p\} \in \mathbb{R}^{3n_x N \times 3n_x N}$,
$\tilde{\Xi} = [\check{A}_p^T \ \sigma_p \check{\mathcal{L}}_p^T \ \sigma_{0p} \check{\mathcal{M}}_p^T \ \check{A}_p^T \ \sigma_p \check{\mathcal{L}}_p^T \ \sigma_{0p} \check{\mathcal{M}}_p^T]^T$, $\tilde{\Upsilon}_q^{(0)} = diag\{\tilde{\Upsilon}_1, \tilde{\Upsilon}_1, \tilde{\Upsilon}_1, \tilde{\Upsilon}_2, \tilde{\Upsilon}_2, \tilde{\Upsilon}_2\}$,
$\tilde{\Upsilon}_1 = \Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ap}} \frac{\pi_{pq}(\nu)}{\bar{\omega}_p} (I_N \otimes \tilde{P}_q^{(0)} - I_N \otimes \tilde{Z}_p - I_N \otimes \tilde{Z}_p^T)$, $\tilde{\Upsilon}_2 = (1 - \bar{\eta}_p)(I_N \otimes P_q^{(0)} - I_N \otimes \tilde{Z}_p - I_N \otimes \tilde{Z}_p^T)$,
$\bar{\eta}_p = \Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ap}} \pi_{pq}(\nu)$, $\bar{\omega}_p = \bar{\eta}_p + (1 - \bar{\theta}_p)\rho$, $\bar{\theta}_p = \Sigma_{p \in \mathcal{O}_{ap}} \theta_{pq}$,

then systems (6) and (8) is leader-following mean square consensus under incompletely accessible semi-Markov kernel of switching topologies and possess a $\mathcal{H}_\infty$ performance index $\hat{\gamma}$. Moreover, the controller gains are given by $K_p = (B^T B)^{-1} B^T (\tilde{Z}_p^T)^{-1} \tilde{Y}_p$.

Proof Since the stability proof of the system at non-switching time of topologies is independent of the semi-Markov kernel, the corresponding proof is easy to obtain by Theorem 1. For incompletely accessible semi-Markov kernel, only the stability of the system at the switching time is given in this theorem.

Similar to inequality (13), we have

$$\begin{aligned}
&\mathbb{E}\Big\{V(\delta(k_{n+1}), \gamma(k_{n+1}), \nu(k_{n+1}))\Big\} - V(\delta(k_{n+1}-1), \gamma(k_{n+1}-1), \nu(k_{n+1}-1)) \\
=&\mathbb{E}\Big\{\delta^T(k_{n+1}-1)\tilde{\mathcal{A}}_p^T (I_N \otimes P_q^{(\nu(k_{n+1}))})\tilde{\mathcal{A}}_p \delta(k_{n+1}-1)\Big\} - \delta^T(k_{n+1}-1)(I_N \otimes P_p^{(\nu(k_{n+1}-1))})\delta(k_{n+1}-1) \\
=&\mathbb{E}\Big\{\delta^T(k_{n+1}-1)\Sigma_{S_{n+1}=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}} \frac{\pi_{pq}(S_{n+1})}{\Omega_p} \tilde{\mathcal{A}}_p^T (I_N \otimes P_q^{(\nu(k_{n+1}))})\tilde{\mathcal{A}}_p \delta(k_{n+1}-1)\Big\} \\
&- \delta^T(k_{n+1}-1)(I_N \otimes P_p^{(\nu(k_{n+1}-1))})\delta(k_{n+1}-1) \\
=&\mathbb{E}\Big\{\delta^T(k_{n+1}-1)\Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}} \frac{\pi_{pq}(\nu)}{\Omega_p} \tilde{\mathcal{A}}_p^T (I_N \otimes P_q^{(0)})\tilde{\mathcal{A}}_p \delta(k_{n+1}-1)\Big\} \\
&- \delta^T(k_{n+1}-1)(I_N \otimes P_p^{(\nu-1)})\delta(k_{n+1}-1)
\end{aligned} \tag{27}$$

for $\gamma(k_{n+1}) = q$, $\nu(k_{n+1}) = 0$, $\gamma(k_{n+1}-1) = p$, $\nu(k_{n+1}-1) = \nu - 1 = S_{n+1} - 1$. For $\Pi(\nu)$, $p \in \mathcal{O}_{ap}$, $q \in \mathcal{O}_{ep}$, the equation (27) can be written as

$$\begin{aligned}
&\mathbb{E}\Big\{V(\delta(k_{n+1}), \gamma(k_{n+1}), \nu(k_{n+1}))\Big\} - V(\delta(k_{n+1}-1), \gamma(k_{n+1}-1), \nu(k_{n+1}-1)) \\
=&\delta^T(k_{n+1}-1)\Big[\Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ap}} \frac{\pi_{pq}(\nu)}{\Omega_p} \big(\bar{A}_p^T \bar{\mathcal{P}} \bar{A}_p + \sigma_p^2 \bar{\mathcal{L}}_p^T \bar{\mathcal{P}} \bar{\mathcal{L}}_p + \sigma_{0p}^2 \bar{\mathcal{M}}_p^T \bar{\mathcal{P}} \bar{\mathcal{M}}_p\big) \\
&+ \Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ep}} \frac{\pi_{pq}(\nu)}{\Omega_p} \big(\bar{A}_p^T \bar{\mathcal{P}} \bar{A}_p + \sigma_p^2 \bar{\mathcal{L}}_p^T \bar{\mathcal{P}} \bar{\mathcal{L}}_p + \sigma_{0p}^2 \bar{\mathcal{M}}_p^T \bar{\mathcal{P}} \bar{\mathcal{M}}_p\big)\Big]\delta(k_{n+1}-1) \\
&- \delta^T(k_{n+1}-1)(I_N \otimes P_p^{(\nu-1)})\delta(k_{n+1}-1) \\
=&\delta^T(k_{n+1}-1)\Big[\Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ap}} \frac{\pi_{pq}(\nu)}{\Omega_p} \big(\bar{A}_p^T \bar{\mathcal{P}} \bar{A}_p + \sigma_p^2 \bar{\mathcal{L}}_p^T \bar{\mathcal{P}} \bar{\mathcal{L}}_p + \sigma_{0p}^2 \bar{\mathcal{M}}_p^T \bar{\mathcal{P}} \bar{\mathcal{M}}_p\big) \\
&+ \Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ep}} \frac{\pi_{pq}(\nu)}{\Omega_p - \bar{\eta}_p} \times \frac{\Omega_p - \bar{\eta}_p}{\Omega_p} \big(\bar{A}_p^T \bar{\mathcal{P}} \bar{A}_p + \sigma_p^2 \bar{\mathcal{L}}_p^T \bar{\mathcal{P}} \bar{\mathcal{L}}_p + \sigma_{0p}^2 \bar{\mathcal{M}}_p^T \bar{\mathcal{P}} \bar{\mathcal{M}}_p\big)\Big]\delta(k_{n+1}-1) \\
&- \delta^T(k_{n+1}-1)(I_N \otimes P_p^{(\nu-1)})\delta(k_{n+1}-1)
\end{aligned} \tag{28}$$

with $\bar{\mathcal{P}} = I_N \otimes P_q^{(0)} \in \mathbb{R}^{n_x N \times n_x N}$. Noting the fact $\Sigma_{\nu=1}^{T_{max}^p} \Sigma_{q \in \mathcal{O}_{ep}} \pi_{pq}(\nu) = \Omega_p - \bar{\eta}_p$, $p \in \mathcal{O}$, it can be found that

$\Sigma_{\nu=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}_{ep}}\frac{\pi_{pq}(\nu)}{\Omega_p-\bar{\eta}_p}=1, 0\le\frac{\pi_{pq}(\nu)}{\Omega_p-\bar{\eta}_p}\le 1, 0\le\bar{\omega}_p<\Omega_p\le 1$. Further, it can be derived that

$$\begin{aligned}
&\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))\Big\}-V(\delta(k_{n+1}-1),\gamma(k_{n+1}-1),\nu(k_{n+1}-1))\\
\le&\delta^T(k_{n+1}-1)\Sigma_{\nu=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}_{ep}}\frac{\pi_{pq}(\nu)}{\Omega_p-\bar{\eta}_p}\Bigg[\Sigma_{\nu=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}_{ap}}\frac{\pi_{pq}(\nu)}{\bar{\omega}_p}\big(\bar{A}_p^T\bar{\mathcal{P}}\bar{A}_p+\sigma_p^2\bar{\mathcal{L}}_p^T\bar{\mathcal{P}}\bar{\mathcal{L}}_p+\sigma_{0p}^2\bar{\mathcal{M}}_p^T\bar{\mathcal{P}}\bar{\mathcal{M}}_p\big)\\
&+(1-\bar{\eta}_p)\big(\bar{A}_p^T\bar{\mathcal{P}}\bar{A}_p+\sigma_p^2\bar{\mathcal{L}}_p^T\bar{\mathcal{P}}\bar{\mathcal{L}}_p+\sigma_{0p}^2\bar{\mathcal{M}}_p^T\bar{\mathcal{P}}\bar{\mathcal{M}}_p\big)-(I_N\otimes P_p^{(\nu-1)})\Bigg]\delta(k_{n+1}-1)
\end{aligned}\tag{29}$$

Then, $\mathbb{E}\Big\{V(\delta(k_{n+1}),\gamma(k_{n+1}),\nu(k_{n+1}))\Big\}-V(\delta(k_{n+1}-1),\gamma(k_{n+1}-1),\nu(k_{n+1}-1))<0$, if the following inequality holds:

$$\begin{aligned}
&\Sigma_{\nu=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}_{ap}}\frac{\pi_{pq}(\nu)}{\bar{\omega}_p}\big(\bar{A}_p^T\bar{\mathcal{P}}\bar{A}_p+\sigma_p^2\bar{\mathcal{L}}_p^T\bar{\mathcal{P}}\bar{\mathcal{L}}_p+\sigma_{0p}^2\bar{\mathcal{M}}_p^T\bar{\mathcal{P}}\bar{\mathcal{M}}_p\big)\\
&+(1-\bar{\eta}_p)\big(\bar{A}_p^T\bar{\mathcal{P}}\bar{A}_p+\sigma_p^2\bar{\mathcal{L}}_p^T\bar{\mathcal{P}}\bar{\mathcal{L}}_p+\sigma_{0p}^2\bar{\mathcal{M}}_p^T\bar{\mathcal{P}}\bar{\mathcal{M}}_p\big)-(I_N\otimes P_p^{(\nu-1)})<0.
\end{aligned}$$

By Schur complement lemma, the above inequality can be further transformed into

$$\tilde{\Phi}=\begin{bmatrix}-(I_N\otimes\tilde{P}_p^{(\nu-1)}) & *\\ \tilde{\Xi} & \hat{\mathcal{P}}_q^{(0)}\end{bmatrix}<0\tag{30}$$

where $\tilde{\Xi}=[\bar{A}_p^T\ \sigma_p\bar{\mathcal{L}}_p^T\ \sigma_{0p}\bar{\mathcal{M}}_p^T\ \bar{A}_p^T\ \sigma_p\bar{\mathcal{L}}_p^T\ \sigma_{0p}\bar{\mathcal{M}}_p^T]^T$, $\hat{\mathcal{P}}_q^{(0)}=diag\{-\hat{\mathcal{P}}_1,-\hat{\mathcal{P}}_1,-\hat{\mathcal{P}}_1,-\hat{\mathcal{P}}_2,-\hat{\mathcal{P}}_2,-\hat{\mathcal{P}}_2\}$, $\hat{\mathcal{P}}_1=\Sigma_{\nu=1}^{T_{max}^p}\Sigma_{q\in\mathcal{O}_{ap}}\frac{\pi_{pq}(\nu)}{\bar{\omega}_p}(I_N\otimes P_q^{-(0)})$, $\hat{\mathcal{P}}_2=(1-\bar{\eta}_p)(I_N\otimes P_q^{-(0)})$.
Applying congruence transformation $diag\{I_{nN},\hat{\mathcal{Z}}_p^T\}$ to (30), we can get

$$\begin{bmatrix}-(I_N\otimes\tilde{P}_p^{(\nu-1)}) & *\\ \tilde{\mathcal{Z}}_p\tilde{\Xi} & \hat{\mathcal{Z}}_p^T\hat{\mathcal{P}}_q^{(0)}\hat{\mathcal{Z}}_p\end{bmatrix}<0\tag{31}$$

with $\hat{\mathcal{Z}}_p=diag\{\tilde{\mathcal{Z}}_p,\tilde{\mathcal{Z}}_p\}$. According to condition (23), one can proof that inequality (31) can guarantee that condition (26) can hold. The rest of the proof can be directly derived in a similar way to Theorem 1 and Theorem 2. This proof is completed.

Remark 4 The controller design and consensus conditions proposed in Theorem 2 and Theorem 3 are based on the same channel fading. However, in practice, the fading variables and interference of communication channels between agents are more likely to be different, due to different complex external environments or different geographic locations of the agents. This restricts the issues considered in this paper to a certain extent. It is worth noting that although the problem of non-identical channel fading has been studied in [28,29], the above literature only considers leaderless multi-agent systems. They ignore the fading effects from leader to follower agents, and the edge Laplacian method introduced cannot be used to tackle the models considered in this paper. Therefore, it is interesting and meaningful to investigate the non-identical channel fading problem within the framework of the fading model proposed in this paper. No better method has been proposed to solve the problem of non-identical channel fading under model (3). This also encourages us to continue to study this issue in future work.

Remark 5 In Theorems 2 and 3, the fully known and incompletely available cases of the semi-Markov kernel for switching topologies are handled respectively, and the corresponding consensus conditions are also derived. Given parameters T_{max}^p and ρ, the minimum $\mathcal{H}_\infty$ performance index $\hat{\gamma}$ of the system can be calculated according to the solution of the following optimization problems:

$$\min\ \hat{\gamma}^2\quad\text{subject to (19) and (20)},\ p,\ q\in\mathcal{O},\ \nu\in\mathbb{N}_{[1,T_{max}^p]}$$

and

$$\min\ \hat{\gamma}^2 \quad \text{subject to (25) and (26)},\ p \in \mathcal{O}_{ap},\ p \in \mathcal{O}_{ep},\ \nu \in \mathbb{N}_{[1,T^p_{max}]}$$

for fully known semi-Markov kernel and incompletely available semi-Markov kernel.

Similar to Corollary 2, the following corollary gives the mean square consensus controller design for system (18) with incompletely accessible semi-Markov kernel.

Corollary 3 Given a scalar $T^p_{max} \in \mathbb{N}_{\geq 1}$, if there exist a sets of symmetric matrices $\tilde{P}^{(\nu)}_p \in \mathbb{R}^{n_x \times n_x} > 0$, $\tilde{P}^{(\nu)}_q \in \mathbb{R}^{n_x \times n_x} > 0$, and matrices $\tilde{Z}_p \in \mathbb{R}^{n_x \times n_x}$, $\tilde{Y}_p \in \mathbb{R}^{n_x \times n_x}$, $p, q \in \mathcal{O}$, $\nu \in \mathbb{N}_{[0,T^p_{max}]}$ such that the inequalities

$$\begin{bmatrix} -(I_N \otimes \tilde{P}^{(\nu-1)}_p) & * \\ \breve{\Xi} & -\mathcal{P}^{-(\nu)}_p - \tilde{\mathcal{Z}}_p - \tilde{\mathcal{Z}}^T_p \end{bmatrix} < 0$$

$$\begin{bmatrix} -(I_N \otimes \tilde{P}^{(\nu-1)}_p) & * \\ \tilde{\Xi} & \tilde{\Upsilon}^{(0)}_q \end{bmatrix} < 0$$

for any $p \in \mathcal{O}_{ap}$, $q \in \mathcal{O}_{ep}$, $\nu \in \mathbb{N}_{[1,T^p_{max}]}$ holds, then the system (18) with incompletely accessible semi-Markov kernel is leader-following mean square consensus under the controller gains $K_p = (B^T B)^{-1} B^T (\tilde{Z}^T_p)^{-1} \tilde{Y}_p$, where $\mathcal{P}^{(\nu)}_p$, $\breve{\Xi}$, $\tilde{\mathcal{Z}}_p$, $\tilde{\Xi}$, and $\tilde{\Upsilon}^{(0)}_q$ are defined in Theorem 1 and 3.

The proof of Corollary 3 can be obtained in a similar way to Theorem 3, which is omitted here.

Remark 6 Theorem 2 and Theorem 3 respectively realize the distributed consensus control of multi-agent systems under the condition that the semi-Markov kernel of switched topology is fully available and incompletely unavailable. In fact, event-triggered control and sampled data control are also excellent methods for dealing with problems related to multi-agent systems [6,11,16]. The advantages and disadvantages of these methods cannot be directly compared. Similarly, event-triggered control and sampled-data control methods can also be applied to the problems considered in this paper. Naturally, event-triggered control and sampled-data control can also be studied in a distributed framework.

4.SIMULATION RESULTS

In this section, a numerical example is provided to demonstrate the validity of the proposed results. Consider a multi-agent system consisting of four followers and one leader with the following parameter matrices

$$A = \begin{bmatrix} 1.0607 & 0.1881 & -0.4654 \\ -0.7425 & 1.1053 & 0.4284 \\ 0.1857 & 0.6231 & -1.0774 \end{bmatrix},\ B = \begin{bmatrix} -0.15 & 0.1 \\ 0.61 & -1.5 \\ -0.11 & -0.2 \end{bmatrix},\ E = \begin{bmatrix} 0.2 & 0 & 0 \\ 0 & 0.1 & 0.2 \\ 0.1 & 0 & 0.1 \end{bmatrix}.$$

In this paper, the information exchange between agents is represented by an undirected switching topology network. The topology graphs are shown in Figure 2. Correspondingly, the Laplacian matrices and the leader's adjacency matrices of each topology graph are given as

$$\mathcal{L}_1 = \begin{bmatrix} 2 & -1 & 0 & -1 \\ -1 & 2 & 0 & -1 \\ 0 & 0 & 1 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix},\ \mathcal{M}_1 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},\ \mathcal{L}_2 = \begin{bmatrix} 1 & 0 & -1 & 0 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix},\ \mathcal{M}_2 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\mathcal{L}_3 = \begin{bmatrix} 2 & 0 & -1 & -1 \\ 0 & 1 & 0 & -1 \\ -1 & 0 & 2 & -1 \\ -1 & -1 & -1 & 3 \end{bmatrix},\ \mathcal{M}_3 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

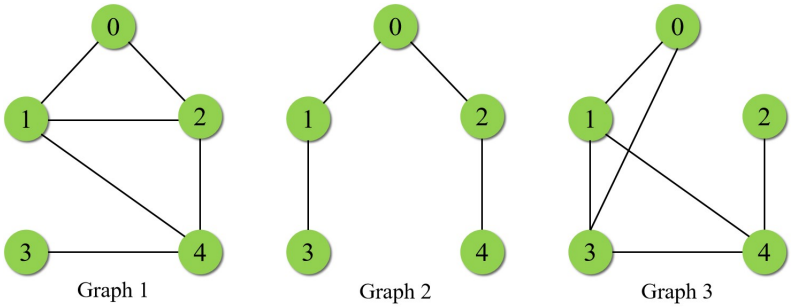


Figure 2. The communication topology graph.

A discrete semi-Markov jump process with semi-Markov kernel is employed to describe the switching of topologies. The transition probability matrix of EMC and the probability density function of sojourn time are provided by

$$[\theta_{pq}] = \begin{bmatrix} 0 & 0.8 & 0.2 \\ 0.26 & 0 & 0.74 \\ 0.5 & 0.5 & 0 \end{bmatrix}, \ [\omega_{pq}(\tau)] = \begin{bmatrix} 0 & \frac{0.6^{\tau}\cdot 0.4^{(10-\tau)}\cdot 10!}{(10-\tau)!\tau!} & \frac{0.7^{\tau}\cdot 0.3^{(10-\tau)}\cdot 10!}{(10-\tau)!\tau!} \\ 0.4^{(\tau-1)^{1.8}} - 0.4^{\tau^{1.8}} & 0 & \frac{0.5^{10}10!}{(10-\tau)!\tau!} \\ 0.5^{(\tau-1)^{1.3}} - 0.5^{\tau^{1.3}} & 0.4^{(\tau-1)^{1.5}} - 0.4^{\tau^{1.5}} & 0 \end{bmatrix}.$$

Let the upper bound of the sojourn-time for each topology mode be $T^1_{max} = T^2_{max} = T^3_{max} = 5$. The statistical characteristic parameters of channel fading are selected as $\mu_1 = 0.8, \mu_2 = 0.7, \mu_3 = 0.75, \mu_{01} = 0.75, \mu_{02} = 0.85, \mu_{03} = 0.6, \sigma_1 = 0.05, \sigma_2 = 0.15, \sigma_3 = 0.2, \sigma_{01} = 0.1, \sigma_{02} = 0.25, \sigma_{03} = 0.15$.

First, we assume that there is no channel fading phenomenon between the leader and the follower, that is, $\xi_{op}(k) = 1$ and $\varpi_{0p}(k) = 0$ in system (6). In this case, mean square consensus conditions in Theorem 2 will be somewhat simplified. At this time, the channel interference between the follower and the follower is selected as $\varpi_p(k) = [0.02\sin(k)e^{(-0.1k)}\ 0.01\cos(k)e^{(-0.1k)}\ -0.01\sin(k)e^{(-0.1k)}]^T$. The initial states of all agents and the initial mode of the communication topology are chosen as $x_0(0) = [0.2\ -0.1\ 0.2]^T, x_1(0) = [0.1\ -0.3\ 0.4]^T, x_2(0) = [-0.4\ 0.8\ -0.2]^T, x_3(0) = [0.5\ -0.2\ 0.4]^T, x_4(0) = [0.3\ -0.6\ -0.1]^T$, and $\gamma(k) = 1$.

By solving simplified consensus conditions and simulating, we can obtain the state response curves of the system consensus error in this case as the solid line in Figure 3 shown. It shows that the controller designed in this paper is still effective when there is no fading phenomenon between the leader and follower agents. Then, the channel fading between the leader and the follower is added to the system under the condition that the controller remains unchanged, and the simulation experiment is performed again. The state trajectory of the consensus error is shown as the dotted line in Figure 3. The simulation results indicate that the fading phenomenon between the leader and the follower will affect the consensus and performance of the system. This further shows that it is necessary and meaningful to study the coexistence of the channel fading phenomenon between the leader and the follower, and the follower and the follower agent.

Next, consider the simultaneous existence of channel fading between leader and follower and between follower and follower. Choosing $\varpi_{0p}(k) = [-0.01\sin(k)e^{(-0.1k)}\ 0.01\cos(k)e^{(-0.1k)}\ 0.02\sin(k)e^{(-0.1k)}]^T$. The rest of the parameters are the same as stated above. By solving the linear matrix inequality condition in Theorem 2, the controller gains K_p can be calculated as

$$K_1 = \begin{bmatrix} 2.3702 & 1.1042 & -2.6922 \\ 0.8871 & 0.8304 & -0.9780 \end{bmatrix}, \ K_2 = \begin{bmatrix} 3.7920 & 1.9250 & -4.7775 \\ 1.4275 & 1.3860 & -1.7717 \end{bmatrix}, \ K_3 = \begin{bmatrix} 2.2009 & 1.2900 & -3.1491 \\ 0.8326 & 0.9225 & -1.1972 \end{bmatrix}$$

In addition, the $\mathcal{H}_\infty$ performance index can be obtained as $\hat{\gamma} = 2.3705$. According to Theorem 2, the state trajectories of all agents in the system (6) are shown in Figure 4. Figure 5 shows the state-response curves of

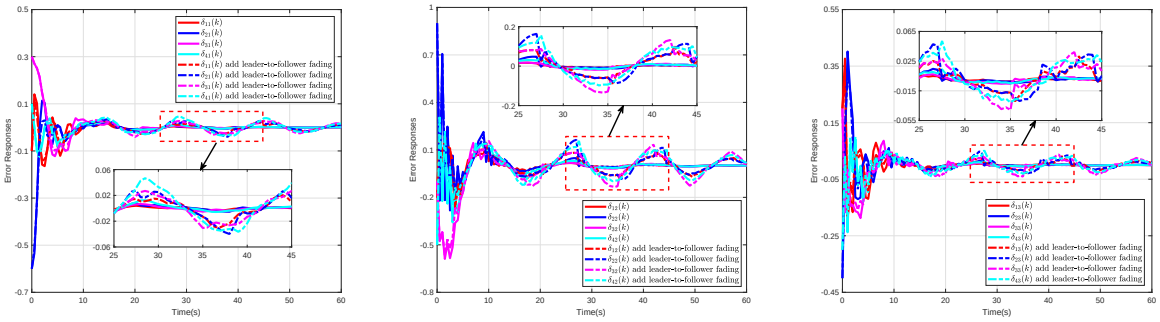


Figure 3. The consensus error responses under reduction conditions of Theorem 2 and additional leader-to-follower fading.

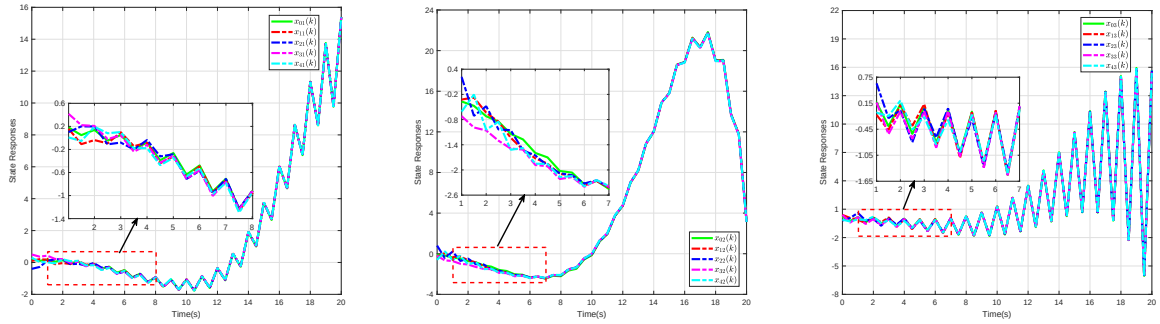


Figure 4. The state responses of $x_0(k)$ and $x_i(k)$ under switching topologies $\mathcal{G}_p$.

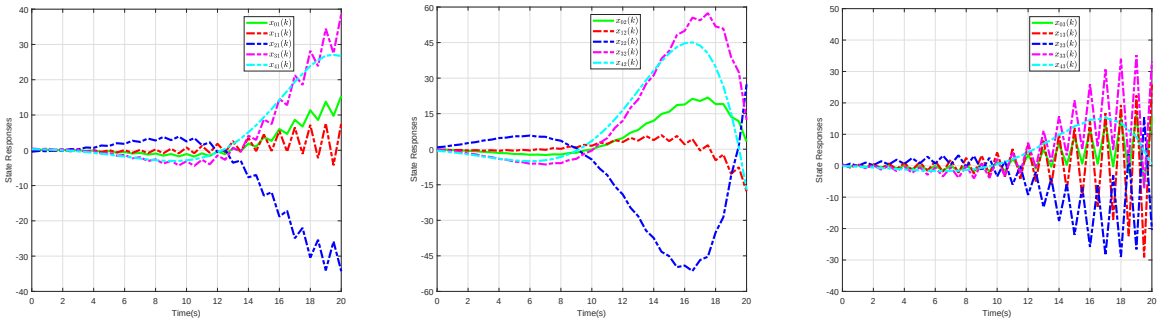


Figure 5. The state responses of $x_0(k)$ and $x_i(k)$ for the open-loop system.

each agent in the open-loop system. It can be seen that the designed control protocol (4) enables the system (6) to achieve leader-following mean square consensus under the premise of the simultaneous existence of channel fading and semi-Markov switching topologies.

In Figure 6, the norm-squared response curves of consensus error in two different cases are given. One is the consensus error of the system in Theorem 2, and the other is the influence of channel fading on system performance when leader-to-follower fading is not considered. From the observation of Figure 6, we can find that the fluctuation of the norm-squared response curves for the consensus error in Theorem 2 represented by the solid line is weaker than that represented by the dotted line. The latter refers to the error norm squared response of the system suffering from leader-to-follower fading in Theorem 2 without considering leader-to-follower fading. This indicates that the consensus error of the former must stabilize faster than the latter. This further means that the control protocol proposed in this paper is more efficient and general. Figure 7 shows a

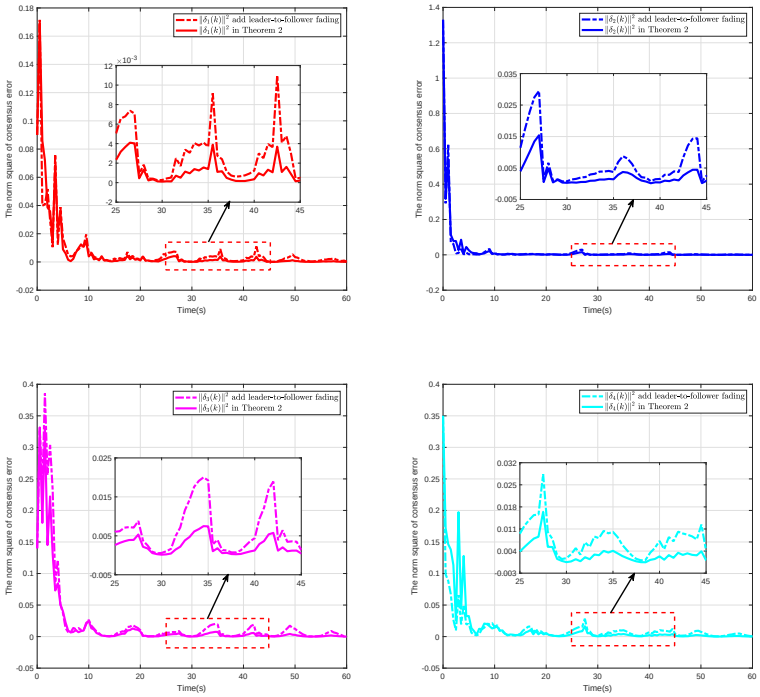


Figure 6. The norm-squared responses of $\delta_i(k)$ under two different scenarios.

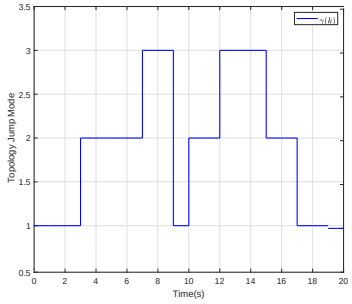


Figure 7. A possible mode jumping diagram of communication topologies.

possible topology jumping rule.

Then, the relationship between the statistical properties of channel fading and the $\mathcal{H}_\infty$ performance of the system (6) is discussed. We consider three different values for the variance of the channel fading coefficient variables $\xi_p(k)$ and $\xi_{0p}(k)$. By solving the conditions in Theorem 2, the value of the minimum $\hat{\gamma}$ can be calculated and shown in Table 1. Table 1 lists three groups of simulation experiments. Simultaneously change the variance of one or two or three groups of fading coefficient variables and calculate the minimum performance index $\hat{\gamma}$ of the system. It can be found from Table 1 that the greater the variance of the channel fading coefficient variables, the greater the $\mathcal{H}_\infty$ performance index $\hat{\gamma}$ of the system. At the same time, comparing the second experiment in the first group with the first experiment in the third group, as the number of variances increases, the $\mathcal{H}_\infty$ performance index $\hat{\gamma}$ of the system also increases. This means that the $\mathcal{H}_\infty$ performance of the system will deteriorate as the effect of the fading channel on the transmitted signal or information becomes greater.

Table 1. $\mathcal{H}_\infty$ **performance with different values of variance for** $\xi_p(k)$ **and** $\xi_{0p}(k)$

Number of variances σ_p and σ_{0p} that vary simultaneously	Values of the variance	Minimum $\hat{\gamma}$
1	$\sigma_1 = 0.05, \sigma_{01} = 0.1$	2.3705
	$\sigma_1 = 0.15, \sigma_{01} = 0.2$	2.4327
	$\sigma_1 = 0.25, \sigma_{02} = 0.3$	2.7232
2	$\sigma_2 = 0.15, \sigma_{02} = 0.25, \sigma_3 = 0.2, \sigma_{03} = 0.15$	2.3705
	$\sigma_2 = 0.1, \sigma_{02} = 0.2, \sigma_3 = 0.15, \sigma_{03} = 0.1$	2.2259
	$\sigma_2 = 0.05, \sigma_{02} = 0.1, \sigma_3 = 0.1, \sigma_{03} = 0.05$	2.1583
3	$\sigma_1 = 0.15, \sigma_2 = 0.25, \sigma_3 = 0.3, \sigma_1 = 0.2, \sigma_2 = 0.35, \sigma_3 = 0.25$	2.7265
	$\sigma_1 = 0.2, \sigma_2 = 0.3, \sigma_3 = 0.35, \sigma_1 = 0.25, \sigma_2 = 0.4, \sigma_3 = 0.3$	3.1614
	$\sigma_1 = 0.25, \sigma_2 = 0.35, \sigma_3 = 0.4, \sigma_1 = 0.3, \sigma_2 = 0.45, \sigma_3 = 0.35$	3.9035

Table 2. Comparative simulations with literature [27,41] for consensus performance

Method	Consensus performance under follower-to-follower fading	Consensus performance under leader-to-follower and follower-to-follower fading
Theorem 3.2 in [27]	Consensus	Inconsistent
Theorem 3.2 in [41]	Consensus	Inconsistent
Simplified Theorem 2 without leader-to-follower fading	Consensus	Consensus performance deterioration
Theorem 2	Consensus	Consensus

According to Table 2, it can be concluded that the consensus performance of the system will deteriorate or be destroyed, when Theorem 3.2 of[27], Theorem 3.2 of[41], and Simplified Theorem 2 in the absence of leader-to-follower fading additionally consider the channel fading of leader-to-follower. The control method proposed in Theorem 2 in this paper can still ensure the consensus performance of the system considering the channel fading between the leader-to-follower and follower-to-follower agents at the same time. This further proves that the model proposed in this paper is more general and the results are more effective than existing ones.

5. CONCLUSION AND FUTURE WORK

In this paper, the $\mathcal{H}_\infty$ leader-following consensus problem of discrete multi-agent systems subject to channel fading is solved under switching topologies with semi-Markov kernel. First, a fading model that takes into account all inter-agent channels (including leader-to-follower channels) is established based on a discrete semi-Markov switching topology. Then, new sufficient criteria have been developed to ensure the mean-square stability and $\mathcal{H}_\infty$ performance of the consensus error system (6) by means of stochastic analysis method and Lyapunov stability theory. Further, for the case where the semi-Markov kernel of switching topologies is not completely accessible, distributed consensus control protocols with fading states have been designed and the desired controller gains have been calculated based on linear matrix inequalities. Finally, a simulation example is presented to verify the effectiveness of the proposed approach. In future work, the problem of non-identical channel fading, adaptive fault-tolerant consensus and game optimization problems for heterogeneous or higher-order nonlinear multi-agent systems are interesting topics. In addition, how to reduce the number of decision variables in matrix inequality conditions and reduce the computational burden is also a problem worth studying.

DECLARATIONS

Authors' contributions

Made substantial contributions to the research, idea generation, algorithm design, simulation, wrote and edited the original draft: Yang H

Performed critical review, commentary and revision, as well as provided administrative, technical, and material support: Zhang H, Wang Z, Zhou X

Availability of data and materials

Not applicable.

Financial support and sponsorship

This work is supported by National Natural Science Foundation of China(61922063), Shanghai Hong Kong Macao Taiwan Science and Technology Cooperation Project (21550760900), Shanghai Shuguang Project (18SG18), Shanghai Natural Science Foundation(19ZR1461400), Shanghai Municipal Science and Technology Major Project (2021SHZDZX0100) and Fundamental Research Funds for the Central Universities. (Corresponding author: Zhang H).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Dong X, Yu B, Shi Z, Zhong Y. Time-varying formation control for unmanned aerial vehicles: theories and applications. *IEEE Trans Control Syst Technol* 2015;23:340-8. DOI
2. Dong X, Zhou Y, Ren Z, and Zhong Y. Timevarying formation tracking for second-order multi-agent systems subjected to switching topologies with application to quadrotor formation flying. *IEEE Trans Ind Electron* 2016;64:5014-24. DOI
3. Kladis GP, Menon PP, Edwards C. Fuzzy distributed cooperative tracking for a swarm of unmanned aerial vehicles with heterogeneous goals. *Int J Syst Sci* 2016;47:3803-11. DOI
4. El-Ferik S, Elkhider SM, Ghommam J. Adaptive containment control of multi-leader fleet of underwater vehicle-manipulator autonomous systems carrying a load. *Int J Syst Sci* 2019;50:1501-16. DOI
5. Do KD. Synchronization motion tracking control of multiple underactuated ships with collision avoidance. *IEEE Trans Ind Electron* 2016;63:2976989. DOI
6. Li X, Dong X, Li Q, Ren Z. Event-triggered time-varying formation control for general linear multi-agent systems. *J Franklin Inst* 2019;356:10179-95. DOI
7. Su H, Zhang J, Chen X. A stochastic sampling mechanism for time-varying formation of multiagent systems with multiple leaders and communication delays. *IEEE Trans Neural Netw Learn Syst* 2019;30:3699-3707. DOI
8. Ding C, Dong X, Shi C, Chen Y, Liu Z. Leaderless output consensus of multi-agent systems with distinct relative degrees under switching directed topologies. *IET Control Theory A* 2018;13:313-20. DOI
9. Gao R, Huang J, Wang L. Leaderless consensus control of uncertain multi-agents systems with sensor and actuator attacks. *Inf Sci* 2019;505:144-56. DOI
10. Cui Q, Huang J, Gao T. Adaptive leaderless consensus control of uncertain multiagent systems with unknown control directions. *Int J Rob Nonl Control* 2020;30:6229-40. DOI
11. Chen C, Lewis F L, Li X. Event-triggered coordination of multi-agent systems via a Lyapunov-based approach for leaderless consensus. *Automatica* 2022;136:109936. DOI
12. He W, Lv S, Wang X, Qian F. Leaderless consensus of multi-agent systems via an event-triggered strategy under stochastic sampling. *J*

Franklin Inst 2019;356:6502-24. DOI
13. He X, Wang Q. Distributed finite-time leaderless consensus control for double-integrator multi-agent systems with external disturbances. *Appl Math Comput* 2017;295:65-76. DOI
14. You X, Hua C, Yu H, Guan X. Leader-following consensus for high-order stochastic multi-agent systems via dynamic output feedback control. *Automatica* 2019;107:418-24. DOI
15. Liu W, Huang J. Leader-following consensus for linear multiagent systems via asynchronous sampled-data control. *IEEE Trans Automat Contr* 2019;65:3215-22. DOI
16. Liu J, Yin T, Yue D, Karimi H R, Cao J. Event-based secure leader-following consensus control for multiagent systems with multiple cyber attacks. *IEEE Trans Cybern* 2020;51:162-73. DOI
17. Shang Y, Ye Y. Leader-follower fixed-time group consensus control of multiagent systems under directed topology. *Complexity* 2017;2017:1-9. DOI
18. Zhang D, Deng C, Feng G. Resilient cooperative output regulation for nonlinear multi-agent systems under DoS sttacks. *IEEE Trans Automat Contr* 2022;Early Access. DOI
19. Zhang D, Deng C, Feng G. Resilient practical cooperative output regulation for MASs with unknown switching exosystem dynamics under DoS attacks. *Automatica* 2022;139:110172. DOI
20. Zhang D, Ye Z, Dong X. Co-design of fault detection and consensus control protocol for multi-agent systems under hidden DoS attack. *IEEE Trans Circuits Syst* 2021;68:2158-70. DOI
21. Wang J, Wang Y, Yan H, Cao J, Shen H. Hybrid event-based leader-following consensus of nonlinear multiagent systems with semi-Markov jump parameters. *IEEE Syst J* 2020;16:397-408. DOI
22. Zhang Z, Su S, Niu Y. Dynamic event-triggered control for interval type-2 fuzzy systems under fading channel. *IEEE Trans Cybern* 2020;51:5342-51. DOI
23. Li X, Zhang B, Li P, Zhou Q, Lu R. Finite-horizon H_∞ state estimation for periodic neural networks over fading channels. *IEEE Trans Neural Netw Learn Syst* 2019;31:1450-1460. DOI
24. Elia N. Remote stabilization over fading channels. *Systems and Control Letters* 2005;54:237-49. DOI
25. Yan H, Zhang H, Yang F, Zhan X, Peng C. Event-triggered asynchronous guaranteed cost control for Markov jump discrete-time neural networks with distributed delay and channel fading. *IEEE Trans Neural Netw and Learn Syst* 2018;29:3588-98. DOI
26. Xiao N, Xie L, Qiu L. Feedback stabilization of discrete-time networked systems over fading channels. *IEEE Trans Automat Contr* 2012;57:2176-89. DOI
27. Gu X, Jia T, Niu Y. Consensus tracking for multi-agent systems subject to channel fading: A sliding mode control method. *Int J Syst Sci* 2020;51:2703-11. DOI
28. Xu L, Xiao N, Xie L. Consensusability of discrete-time linear multi-agent systems over analog fading networks. *Automatica* 2016;71:292-9. DOI
29. Xu L, Zheng J, Xiao N, Xie L. Mean square consensus of multi-agent systems over fading networks with directed graphs. *Automatica* 2018;95:503-10. DOI
30. Liu Z-W, Wen G, Yu X, Guan Z-H, Huang T. Delayed impulsive control for consensus of multiagent systems with switching communication graphs. *IEEE Trans Cybern* 2020;50:3045-55. DOI
31. Jiang J, Jiang Y. Leader-following consensus of linear time-varying multi-agent systems under fixed and switching topologies. *Automatica* 2020;113:108804. DOI
32. Li B, Wen G, Peng Z, Huang T, Rahmani A. Fully distributed consensus tracking of stochastic nonlinear multiagent systems with Markovian switching topologies via intermittent control. *IEEE Trans Syst, Man, Cybern: Syst* 2021;52:3200-9. DOI
33. Shang Y. Couple-group consensus of continuous-time multi-agent systems under Markovian switching topologies. *J Franklin Inst* 2015;352:4826-44. DOI
34. Ge X, Han Q. Consensus of multiagent systems subject to partially accessible and overlapping Markovian network topologies. *IEEE Trans Cybernet* 2017;47:1807-19. DOI
35. Dai J, Guo G. Event-triggered leader-following consensus for multi-agent systems with semi-Markov switching topologies. *Inf Sci* 2018;459:290-301. DOI
36. Li K, Mu X. Containment control of stochastic multiagent systems with semi-Markovian switching topologies. *Int J Rob Nonlin Contr* 2019;29:4943-55. DOI
37. Liang H, Zhang L, Sun Y, Huang T. Containment control of semi-Markovian multiagent systems with switching topologies. *IEEE Trans Syst, Man, Cybernet: Syst* 2019;51: 3889-99. DOI
38. He M, Mu J, Mu X. H_∞ leader-following consensus of nonlinear multi-agent systems under semi-Markovian switching topologies with partially unknown transition rates. *Infor Sci* 2020;513:168-79. DOI
39. Zhang L, Leng Y, Colaneri P. Stability and stabilization of discrete-time semi-Markov jump linear systems via semi-Markov kernel approach. *IEEE Trans Automat Contr* 2016;61:503-8. DOI
40. Ning Z, Zhang L, Colaneri P. Semi-Markov jump linear systems with incomplete sojourn and transition information: Analysis and synthesis. *IEEE Trans Automat Contr* 2020;65:159-74. DOI
41. Li W, Niu Y, Cao Z. Event-triggered sliding mode control for multiagent systems subject to channel fading. *Int J Syst Sci* 2021;53:1233-44. DOI

Su *et al*. *Intell Robot* 2022;2(3):244-74
DOI: 10.20517/ir.2022.17

Intelligence & Robotics

Review

Open Access

A review of causality-based fairness machine learning

Cong Su[1], Guoxian Yu[1,2], Jun Wang[2], Zhongmin Yan[1], Lizhen Cui[1,2]

[1]School of Software, Shandong University, Jinan 250101, Shandong, China.
[2]SDU-NTU Joint Centre for Artificial Intelligence Research, Shandong University, Jinan 250101, Shandong, China.

Correspondence to: Prof. Guoxian Yu, School of Software, Shandong University, Jinan 250101, Shandong, China. E-mail: gxyu@sdu.edu.cn

How to cite this article: Su C, Yu G, Wang J, Yan Z, Cui L. A review of causality-based fairness machine learning. *Intell Robot* 2022;2(3):244-74. http://dx.doi.org/10.20517/ir.2022.17

Received: 8 Jun 2022 **First Decision:** 6 Jul 2022 **Revised:** 19 Jul 2022 **Accepted:** 25 Jul 2022 **Published:** 21 Aug 2022

Academic Editor: Simon X. Yang **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

With the wide application of machine learning driven automated decisions (e.g., education, loan approval, and hiring) in daily life, it is critical to address the problem of discriminatory behavior toward certain individuals or groups. Early studies focused on defining the correlation/association-based notions, such as statistical parity, equalized odds, etc. However, recent studies reflect that it is necessary to use causality to address the problem of fairness. This review provides an exhaustive overview of notions and methods for detecting and eliminating algorithmic discrimination from a causality perspective. The review begins by introducing the common causality-based definitions and measures for fairness. We then review causality-based fairness-enhancing methods from the perspective of pre-processing, in-processing and post-processing mechanisms, and conduct a comprehensive analysis of the advantages, disadvantages, and applicability of these mechanisms. In addition, this review also examines other domains where researchers have observed unfair outcomes and the ways they have tried to address them. There are still many challenges that hinder the practical application of causality-based fairness notions, specifically the difficulty of acquiring causal graphs and identifiability of causal effects. One of the main purposes of this review is to spark more researchers to tackle these challenges in the near future.

Keywords: Fairness, causality, fairness-enhancing mechanisms, machine learning, fairness notions

1. INTRODUCTION

Artificial intelligence (AI) techniques are widely applied in various fields to assist people in decision-making, such as hiring[1,2], loans[3,4], education[5], criminal risk assessment[6], etc. The motivation for using machine

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

learning models is that they can mine hidden laws and useful information from data with huge volumes and various structures more quickly and effectively than human beings. Most importantly, people often mix personal emotions when making decisions, making their decisions unfavorable to certain groups. It is canonically believed that the decisions made by automatic decision-making systems are more objective, and, thus, there will be no discrimination against specific groups or individuals. However, this assumption cannot always be met. Due to the biased training data and inherent bias of adopted models, machine learning models are not always as neutral as people expect.

Since many automated systems driven by AI techniques can significantly impact people's lives, it is important to eliminate discrimination embedded in the AI models so that fair decisions are made with their assistance. Indeed, in recent years, fairness issues of AI models have been receiving wide attention. For example, automated resume screening systems often give biased evaluations based on traits beyond the control of candidates (e.g., gender and race), which may not only discriminate against job applicants with certain characteristics but also cost employers by missing out on good employees. Early research on achieving fairness of algorithms focused on statistical correlation and developed many correlation-based fairness notions (e.g., predictive parity[7], statistical parity[8], and equalized odds[9]), which primarily focus on discovering the discrepancy of statistical metrics between individuals or sub-populations. However, correlation-based fairness notions fail to detect discrimination in algorithms in some cases and cannot explain the causes of discrimination, since they do not take into account the mechanism by which the data are generated. A classic example is Simpson's paradox[10], where the statistical conclusions are drawn from the sub-populations and the whole population can be different. On the other hand, discrimination claims usually require demonstrating causal relationships between sensitive attributes and questionable decisions, instead of the association or correlation between the sensitive attributes and decisions.

Consider the example of the graduate admissions at University of California, Berkeley in 1973, which confirms the importance of developing causal perspective admission to detect and eliminate discrimination. From the statistical results of historical data, roughly 44% of all men who applied were admitted, compared to 35% of women who applied. Then, a flawed conclusion may be drawn with the support of the difference in admission rates between males and females. That is, there exists discrimination towards women for their graduate admission. After an in-depth examination of this case, there is no wrongdoing by the educational institution, but a larger proportion of women applied to the most competitive departments, causing a lower admission rate than men. However, the question of discrimination is far from resolved, e.g., there is no way of knowing why women tended to apply to more competitive departments from the available data alone. Therefore, it is helpful to detect discrimination and interpret the sources of discrimination by understanding the data generating mechanism, namely the causality behind the problem of discrimination. In addition, causal models can be regarded as a mechanism to integrate scientific knowledge and exchange credible assumptions to draw credible conclusions. For this admission case, it seems that, due to women's socialization and education, they tend to toward fields of studies that are generally crowded. Therefore, it is necessary to explore the causal structure of the problem. Fortunately, more and more researchers have paid attention to detecting and eliminating discrimination from the perspective of causality, and various fairness concepts and fairness-enhancing methods based on causality have been proposed.

Compared with the fairness notions based on correlation, causality-based fairness notions and methods take additional consideration of the knowledge that reflects the causal structure of the problem. This knowledge reveals the mechanism of data generation and is helpful for us to comprehend how the influence of sensitive attributes change spreads in the system, which is conducive to improving the interpretability of model decisions[11–14]. Therefore, causality-based fairness machine learning algorithms help to enhance fairness. However, causality-based fairness approaches still face many challenges, one of which is unidentifiable situations of causal effects[15]. In other words, the causal effect between two variables cannot be uniquely computed from

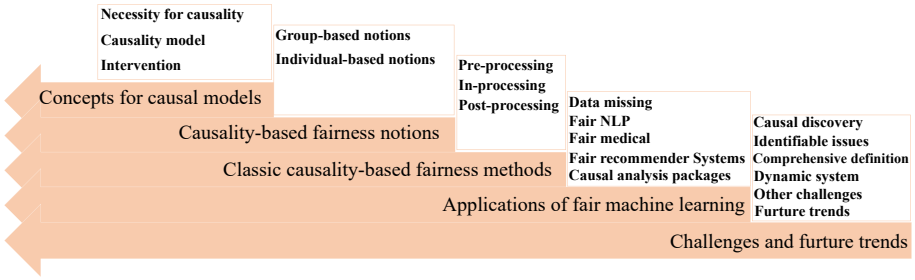


Figure 1. Organizational structure of this paper.

the observational data without extra assumptions.

The previous survey articles offer a high-level summary of technology to eliminate algorithm discrimination, and there is no detailed discussion on specific subareas[16–19]. Wan *et al.*[20] provided an exhaustive review of the methods of using the in-process mechanism to solve the fairness issues. Makhlouf *et al.*[21] summarized the advantages and disadvantages of each fairness notion and their suitability, aiming to help us select the fairness notions that are most suitable for a particular scenario. Makhlouf *et al.*[22] only focused on reviewing the concept of fairness based on causality. Instead of eliminating discrimination from a technical perspective, Wu *et al.*[23] summarized human-oriented fairness-enhancing methods as a way to explore the role of human beings in addressing algorithmic fairness.

To complement the previous survey papers on fairness-enhancing machine learning, this survey thoroughly reviews the concept of fairness based on causality and summarizes the core ideology behind the causality-based fairness-enhancing approaches. This survey aims to stimulate future exploration of causality-based fairness technology, because of the importance of causal modeling in improving algorithmic fairness. In this survey, the review of the concept and technology of fairness based on causality is conducted in several phases. The survey first reviews causality-based definitions and measures of fairness and summarizes the suitability of these causality-based fairness notions. Next, it provides a comprehensive overview of state-of-the-art methods to achieve fairness based on these causality-based fairness notions. The survey also discusses the practical applications, beyond classification, that the causality-based fairness methods are expected to benefit greatly. Finally, this survey discusses the challenges of eliminating discrimination from a causal perspective, including the acquisition of causal graphs and identifiable issues. It also reviews the efforts for addressing these challenges and summarizes the remaining issues, which provides some assistance to solve these problems for future research. Figure 1 shows the organizational structure of this survey.

The rest of this survey is structured as follows. Section 2 introduces an example to interpret the importance of causal modeling for addressing fairness issues. Section 3 presents the background of the causal model. Section 4 introduces definitions and measures of causality-based fairness notions and discusses the suitability or applicability of them. Section 5 discusses fairness mechanisms and causality-based fairness approaches, and compares these mechanisms. Section 6 introduces several typical applications of causality-based fairness methods. Section 7 analyzes the challenges and the research trends for applying causality-based fairness-enhancing methods in practical scenarios.

2. THE IMPORTANCE FOR CAUSALITY TO DETECT DISCRIMINATION: AN EXAMPLE

The importance of applying causal analysis to discrimination discovery is explored in this section. Consider a simple example that is inspired by a legal dispute about religious discrimination in recruitment[24]. To keep the situation simple, assuming that a company takes the religious belief S ($S = 1$ if an applicant has a religious

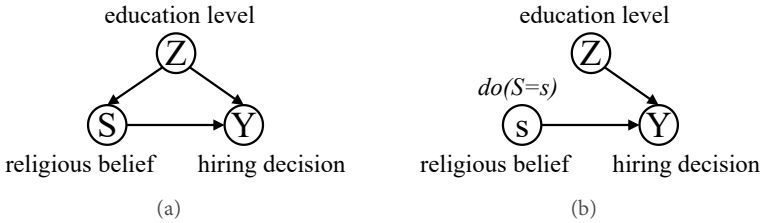


Figure 2. Two causal graphs for the hiring decision system. (a) A causal graph of the hiring decision system, where S (i.e., individual's religious belief) is the sensitive attribute and Y (i.e., hiring decision) is the decision. (b) A causal graph of this system after intervention on S.

belief, otherwise $S = 0$) and the education level Z ($Z = 1$ denotes a candidate has received higher education, while $Z = 0$ denotes the education level of am applicant is low) into account for deciding whether to hire applicant Y ($Y = 1$ represents the company decides to hire the applicant, otherwise $Y = 0$). The religious belief of an individual can be influenced by several variables, but, in this example, assume that S is only influenced by the individual's education level. For concreteness, we assume that, based on statistics on the collected data, we have the following conditional probabilities:

$$\begin{aligned} &P(S=0|Z=0)=0.2 \qquad && P(S=0|Z=1)=0.8 \\ &P(S=1|Z=0)=0.8 && P(S=1|Z=1)=0.2 \\ &P(Y=1|S=0,Z=0)=0.03 && P(Y=1|S=0,Z=1)=0.24 \\ &P(Y=1|S=1,Z=0)=0.02 && P(Y=1|S=1,Z=1)=0.25 \end{aligned} \tag{1}$$

Assume also that the proportions of applicants with faith are equal to those without religious beliefs, and the proportions of applicants with high education are the same as the one with low education, which means that $P(S = 1) = P(S = 0) = 0.5$ and $P(Z = 1) = P(Z = 0) = 0.5$. Then, the hiring decision made by the company is suspected of prejudice against applicants with religious beliefs because there are statistical differences in religious composition among employees (corresponding to statistical parity[8]):

$$\begin{aligned} P(Y=1|S=1) &= \sum_{z\in\{0,1\}} P(Y=1|S=1,Z=z)\cdot P(S=1|Z=z) \\ &= 0.02\times 0.8 + 0.25\times 0.2 = 0.066 \\ P(Y=1|S=0) &= \sum_{z\in\{0,1\}} P(Y=1|S=0,Z=z)\cdot P(S=0|Z=z) \\ &= 0.03\times 0.2 + 0.24\times 0.8 = 0.198 \end{aligned}$$

From such probabilities mentioned above, the company is more likely to hire the applicants without religious beliefs than ones with faith, which indicates that its hiring decisions are unfair. However, this conclusion is wrong, since it only considers the association between religious beliefs and hiring decisions, instead of the causal relationship. In fact, understanding causal mechanisms behind hiring decisions can avoid drawing such wrong conclusions, since it exposes the mechanism of data generation. Through the causal analysis for three variables in the above example, the education level of the individuals is an observable confounder, that is, an individual's educational background influences both his (or her) religious beliefs and hiring decisions made by the company. The higher the education level of individuals, the less willing they are to participate in religious activities. The causal relationships between these variables are shown in Figure 2(a). Based on the causal graph, further causal analysis of hiring decisions of this company is conducted to explore whether it is really discriminatory. In other words, intervention on S is performed to block the influence of Z on S to

evaluate the causal effect of S on Y (more details of intervention can be seen in Section 3). Figure 2(b) shows the causal structure of such an example after intervening on S. The hiring decisions made by the company are fair, if the hiring proportions when all applicants in the population have religious beliefs are the same as the hiring proportions when all applicants in the population have no religious beliefs, i.e., $P(Y = 1|do(S = 1)) = P(Y = 1|do(S = 0))$. Formally, these probabilities in this example are obtained as below:

$$\begin{aligned} P(Y = 1|do(S = 1)) &= \sum_{z\in\{0,1\}} P(Y = 1|S = 1, Z = z) \cdot P(Z = z) \\ &= 0.02 \times 0.5 + 0.25 \times 0.5 = 0.135 \\ P(Y = 1|do(S = 0)) &= \sum_{z\in\{0,1\}} P(Y = 1|S = 0, Z = z) \cdot P(Z = z) \\ &= 0.03 \times 0.5 + 0.24 \times 0.5 = 0.135 \end{aligned}$$

These values confirm that the hiring decisions made by the company do not discriminate against applicants with religious beliefs. Therefore, it is critical to conduct a causal analysis of the problem, since understanding the causal mechanisms behind the problem can not only help to detect discrimination but also help to interpret the sources of discrimination.

3. PRELIMINARIES AND NOTATION

In this review, an attribute is denoted by an uppercase letter, e.g., X; a subset of attributes is denoted by a bold uppercase letter, e.g., $\mathbf{X}$; a domain value of attribute X is denoted by a lowercase letter, e.g., x; and the value assignment of subset attributes $\mathbf{X}$ is denoted by a bold lowercase letter, e.g., $\mathbf{x}$. In particular, S represents the sensitive attribute (e.g., race) and Y represents the predicted result of the AI model system (e.g., loans).

One of the most popular causal model frameworks is Pearl's Structural Causal Model (SCM)[10]. A structural causal model $\mathcal{M}$ is represented by a quadruple $\langle \mathbf{U}, P(\mathbf{U}), \mathbf{V}, \mathbf{F} \rangle$:

1. $\mathbf{U}$ denotes exogenous variables that cannot be observed but constitute the background knowledge behind the model.
2. $P(\mathbf{U})$ represents the joint probability distribution of $\mathbf{U}$.
3. $\mathbf{V}$ denotes endogenous variables that can be observed.
4. $\mathbf{F}$ denotes a set of functions mapping from $\mathbf{U} \cup \mathbf{V}$ to $\mathbf{V}$, which reflects the causal relationship between variables. For each $X \in \mathbf{V}$, there is a mapping function $f_X \in \mathbf{F}$ from $\mathbf{U} \cup (\mathbf{V} \setminus X)$ to X, i.e., $X = f_X(Pa(X), U_X)$, where parent variables $Pa(X) \subset \mathbf{V} \setminus X$ are the endogenous variables that directly control the value of X, and U_X is a set of exogenous variables that directly determine X.

A causal model $\mathcal{M}$ is associated with a causal graph $\mathcal{G} = \langle \mathcal{V}, \mathcal{E} \rangle$, a directed acyclic graph, where $\mathcal{V}$ is a set of nodes, each of which represents an endogenous variable of $\mathbf{V}$, and each element in $\mathcal{E}$ indicates a directed edge $\rightarrow$, pointing from a node $X \in \mathbf{U} \cup \mathbf{V}$ to another node $Y \in \mathbf{V}$ if f_Y uses values of X as input, which represents a causal relationship between the corresponding variables. The exogenous variables $\mathbf{U}$ can be either independent or dependent. If the exogenous variables $\mathbf{U}$ are mutually independent, the causal model is called Markovian model. In this case, exogenous variables typically are not represented in the causal diagram. In the case, the exogenous variables are mutually dependent (hidden confounders), the causal model is called the semi-Markovian model. In the semi-Markovian model, dashed bi-directed edges are used to represent the hidden confounders between two variables. Figure 7 shows examples of causal graphs of Markovian model [Figure 3(a)] and semi-Markovian model [Figure 3(b)].

An intervention simulates the physical interventions that force some variable X to take certain values x regardless of the corresponding function f_x, denoted by $do(x)$. In the causal graph, it is shown as discarding all edges

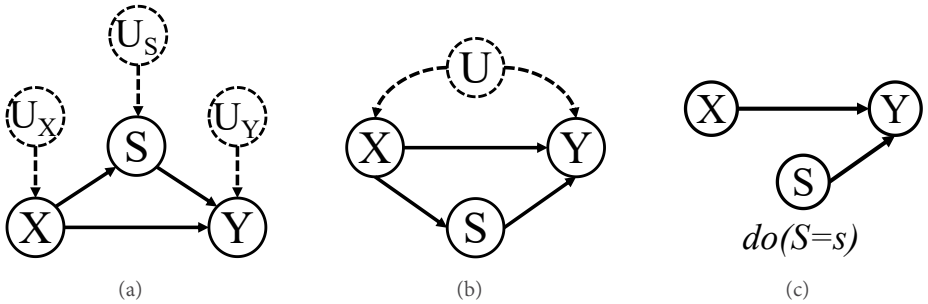


Figure 3. (a) An example causal graph based on Markovian assumption; (b) a causal graph based on semi-Markovian assumption; and (c) a causal graph after performing an intervention on S.

pointing to variable X. Figure 3(c) shows the causal diagram after the intervention $do(x)$. The mathematical meaning of $do(x)$ is defined as the substitution of equation $X = f_X(Pa(X), U_X)$ with $X = x$. For another endogenous variable Y which is affected by the intervention, its post-intervention distribution under $do(x)$ is denoted by $P(Y|do(x))$ or $P(y_x)$ for short. Intuitively, $P(Y|X = x)$ represents the population distribution of Y condition on observing attribute X value of individuals is x, while $P(Y|do(X = x))$ (i.e., $P(Y|do(x))$) represents the population distribution of Y if everyone in the population had their X value fixed at x. This post-intervention distribution $P(Y|do(x))$ is considered a counterfactual distribution since the intervention $do(x)$ forces X to take a certain value different from the one it would take in the actual world. For example, if S represents sex (s^+, male; s^-, female) and Y represents the hiring decision (y^+, hired; y^-, not hired), $P(y_{s^-}|Y = y^+, S = s^+)$ involves two worlds: a real world that a male applicant has been hired and a counterfactual world where the same applicant is female. Such expression means that, when a job-hunter whose gender is male has been observed to be hired, what is the probability that the same job-hunter would still be hired if this job-hunter were female.

Causality-based fairness notions aim to tell whether the outcome of a decision made by the AI decision model is discriminative, which are expressed by interventional and counterfactual probability distributions. The application of the causality-based fairness notions not only requires a dataset $\mathcal{D}$ as input but also relies on a causal graph $\mathcal{G}$. Causal approaches aim to limit the causal effects of sensitive attributes on decisions, which are computed by interventional and counterfactual probabilities. However, since these probabilities cannot be observed, they fail to be uniquely assessed from $\mathcal{D}$ and $\mathcal{G}$ in some cases, which is called the unidentifiable issue. In other words, if two variables have different causal effect measures resulting from different causal models that all agree with the observational distribution, the causal effects are unidentifiable.

4. CAUSALITY-BASED FAIRNESS NOTIONS

Pearl defined the causality as three rungs: correlation, intervention, and counterfactual[25]. The first rung (correlation) reflects the ability of observation, which aims to discover the patterns in the environment. The second rung (intervention) reflects the ability of action, which refers to the prediction of the results of deliberate changes to the environment. The third rung (counterfactual) refers to the ability to imagine the counterfactual world and speculate on the causes of the observed phenomena. The second and third rungs aim to expose the root causes of the patterns that we observe. Thus far, many fairness metrics have been proposed and all of them can be placed in such causal rungs. Figure 4 presents categorization of fairness notions. Early definitions of fairness are based on statistical correlations, all of which can be found at the first rung. All causality-based fairness notions can be found at the second and third rungs, each of which considers the mechanism of data generation. Thus, causality-based fairness notions can better reveal the causes of discrimination than statistics-based ones and have been attracting more and more attention.

Table 1. Typical causality-based fairness notions

Type	Fairness notion	Formulation	Description
Group	Total Effect [10]	$\|P(y\|S=s^+)-P(y\|S=s^-)\| \leq \tau$	The causal effects of the value change of the sensitive attribute S from s^+ to s^- on decision $Y=y$, where the intervention is transmitted along all causal paths and is within the fair threshold τ
	Effect of treatment on the treated [10]	$\|P(y_{s^+}\|S=s^-)-P(y_{s^-}\|S=s^-)\| \leq \tau$	The difference between the distribution of $Y=y$ had S been s^+ and that of $Y=y$ had S been s^-, given that S had been observed to be s^-
	Path-specific fairness [26,27]	$\|P(y\|do(s^+\|\pi, s^-\|\bar{\pi}))-P(y\|do(s^-))\| \leq \tau$	The causal effects of the value change of the sensitive attribute S from s^+ to s^- on decision $Y=y$ along specific causal paths, is within the fair threshold τ
	No unresolved discrimination [28]	-	It is satisfied when there is no directed path from sensitive attribute S to outcome Y allowed, except through a resolving variable
	No proxy discrimination [28]	$P(Y\|do(R=r_0))=P(Y\|do(R=r_1))$ $\forall r_0, r_1 \in dom(R)$	If it is satisfied, there is no path from the sensitive attribute S to the outcome Y blocked by a proxy variable
Individual	Counterfactual fairness [11]	$\|P(y_{s^+}\|\mathbf{O}=\mathbf{o}, S=s^-)-$ $P(y_{s^-}\|\mathbf{O}=\mathbf{o}, S=s^-)\| \leq \tau$	An outcome Y achieves counterfactual fairness towards an individual i (i.e., $\mathbf{O}=\mathbf{o}$) if the probability of $Y=y$ for such individual i is the same as the probability of $Y=y$ for the same individual, who belongs to a different sensitive group
	Individual direct discrimination [29]	$d(i,i')=\sum_{k=1}^{\|\mathbf{X}\|}\|CE(x_k,x_k')\cdot VD(x_k,x_k')\|$	It is based on situation testing where the causal reasoning is used to define the distance function $d(i,i')$
	Equality of effort [30]	$\Psi_{G^+}(\gamma)=\Psi_{G^-}(\gamma)$ where $\Psi_{G^+}(\gamma)=argmin_{t\in T}\mathbb{E}[Y_{G^+}^t]\geqslant\gamma$	It detects discrimination by comparing the effort required to reach the same level of outcome of individuals from advantaged and disadvantaged groups who are similar to the target individual
Hybrid	PC-fairness [31]	$\|P(\hat{y}_{s^+\|\pi,s^-\|\bar{\pi}}\|\mathbf{o})-P(\hat{y}_{s^-}\|\mathbf{o})\| \leq \tau$	It is a general fairness formalization for representing various causality-based fairness notions, which is achieved by differently tuning its parameters

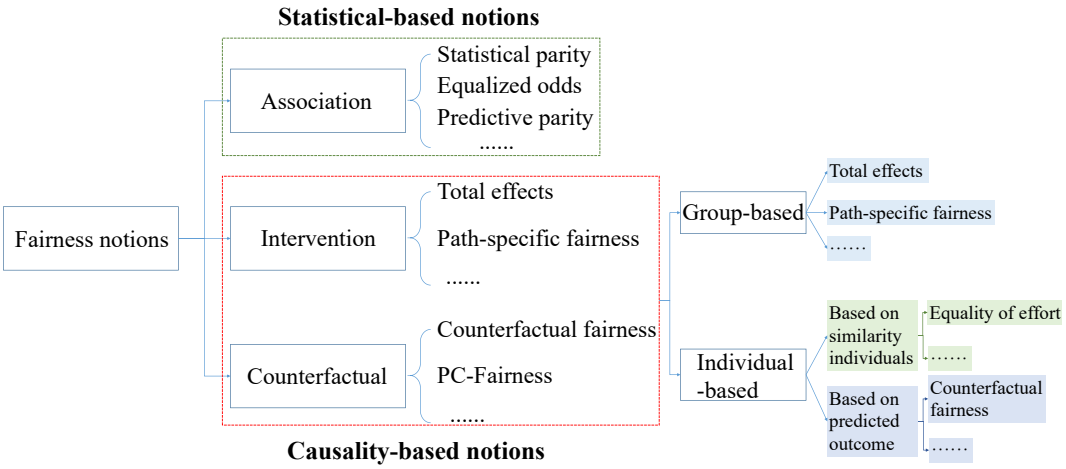


Figure 4. The categorization of fairness notions.

In the real world, the focus of different machine learning tasks is different, and thus, various causality-based fairness notions are proposed to detect discrimination in different scenarios. This section introduces several representative causality-based fairness measurements that quantify fairness from the perspective of groups or individuals, respectively. Without loss of generality, assume that the sensitive attribute S and the outcome of the automated decision making Y are binary variables where $S = s^+$ denotes the advantaged group (e.g.,

white men) and $S = s^-$ denotes the disadvantaged one (e.g., non-white men). Table 1 summarizes various causality-based fairness notions falling under different types.

4.1. Group causality-based fairness notions

Group fairness notions aim to discover the difference in outcomes of AI decision models across different groups. The value of an individual's sensitive attribute reflects the group he (or she) belongs to. Considered an example of salary prediction where s^+ and s^- represent male and female groups, respectively. Some representative group causality-based fairness notions are introduced as follows.

4.1.1. Total effect

Before defining total effect (TE)[10], statistical parity (SP) is first introduced, since it is similar to *TE* but is fundamentally different from *TE*. SP is a common statistics-based fairness notion, which denotes similar individuals treated similarly regardless of their sensitive attributes. Statistical parity is satisfied if

$$|SP(y)| = |P(y|S = s^+) - P(y|S = s^-)| \leq \tau \tag{2}$$

Intuitively, $SP(y)$ measures the conditional distributions of Y change of one's sensitive attribute S from s^+ to s^-, and it is considered to be fair if the difference between the conditional distributions is within the fair threshold τ. The main limitation of $SP(y)$ is that $SP(y)$ is unable to reflect the causal relationship between S and Y. Total effect is the causal version of statistical parity, which additionally considered the generation mechanism of the data. Formally, total effect can be computed as follows:

$$TE(y) = P(y|do(S = s^+)) - P(y|do(S = s^-)) \tag{3}$$

TE measures the difference between total causal effect of sensitive attribute S changing from s^+ to s^- on decision $Y = y$. Intuitively, statistical parity represents the difference in probabilities of $Y = y$ in the sampling population, while total effect represents the difference in probabilities of $Y = y$ in the entire population.

A more complex total effect considers the effect of changes in the sensitive attribute value on the outcome of automated decision making when we already observed the outcome for that individual, which is known as the effect of treatment on the treated (ETT)[10]. This typically involves a counterfactual situation which requires changing the sensitive attribute value of that individual at that time to examine whether the outcome changes or not. ETT can be mathematically formalized using counterfactual quantities as follows:

$$ETT(y) = P(y_{s^+}|s^-) - P(y_{s^-}|s^-) \tag{4}$$

where $P(y_{s^+}|s^-)$ represents the probability of $Y = y$ had S been s^+, given S had been observed to be s^-. $P(y_{s^-}|s^-) = P(y|s^-)$ represents the conditional distributions of $Y = y$ when we observe $S = s^-$. Such probability involves two worlds: one is an actual world where $S = s^-$ and the other is a counterfactual world where for the same individual $S = s^+$. Notice that $P(y_{s^-}|s^-) = P(y|s^-)$ for consistency.

Other fairness notions similar to *TE* are also proposed. For example, FACT (fair on average causal effect)[32] was proposed to detect discrimination of automated decision making, which is based on potential outcome framework[33,34]. It considers an outcome Y is fair, if the average causal effect over all individuals in the population of the value changes of S from s^+ to s^- on Y is zero, i.e., $\mathbb{E}[Y_i^{(s^+)} - Y_i^{(s^-)}] = 0$, where $Y_i^{(s^+)}$ denotes the potential outcome of an individual i had S been s^+.

TE and *ETT* both aim to eliminate the decision bias on all causal paths from S to Y. However, they cannot distinguish between direct discrimination, indirect discrimination, and explainable bias.

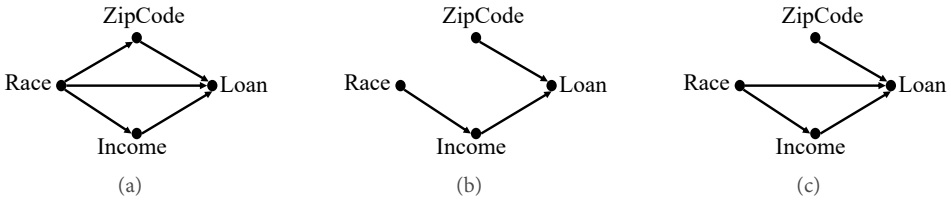


Figure 5. Two alternative graphs for the loan application system. (a) A causal graph of the loan application system, where *Race* is the sensitive attribute and *Loan* is the decision. (b) A causal graph of the system after removing unresolved discrimination. (c) A causal graph of the system that is free of proxy discrimination.

4.1.2. Path-specific fairness

The causal effect of sensitive attribute on the outcome can be divided into direct effect and indirect effect, and it can be deemed fair or discriminatory by an expert. Direct discrimination can be captured by the causal effects of S on Y transmitted along the direct path from S to Y, while indirect discrimination is measured using the causal effect of S on Y along causal paths from S to Y that pass through redlining/proxy attributes.

Figure 5(a) represents the causal graph of a simple example of a toy model of loan decision AI model, where *Race* is treated as the sensitive attribute and *Loan* is treated as the decision. Since *ZipCode* can reflect the information of *Race*, *ZipCode* is a proxy for the sensitive attribute, that is to say, *ZipCode* is a redline attribute. Thus, the causal effects spreading along the path *Race* $\rightarrow$ *Loan* are then considered to be direct discrimination, and the causal effects spreading along the path *Race* $\rightarrow$ *ZipCode* $\rightarrow$ *Loan* are considered to be indirect discrimination. Note that the causal effects spreading along the path *Race* $\rightarrow$ *Income* $\rightarrow$ *Loan* are explainable bias since it is reasonable to deny a loan to an applicant if he (or she) has a low income. That is to say, the partial difference in loan issuance across different race groups can be explained by the fact that some racial groups in the collected data tend to be underpaid.

Path-specific effect[10] is a fine-grained assessment of causal effects, that is, it can evaluate the causal effect transmitted along certain paths. Thus, it is used to distinguish among direct discrimination, indirect discrimination, and explainable bias. For any set of paths π, the π-specific effect can be computed as below:

$$PSF_{\pi}(y) = P(y|do(s^{+}|\pi, s^{-}|\bar{\pi})) - P(y|do(s^{-})) \tag{5}$$

where $P(y_{s^{+}|\pi,s^{-}|\bar{\pi}})$ denotes the distribution of $Y = y$ where the intervention $do(s^{+})$ (i.e., force S had s^{+}) is only transmitted along path π while the intervention $do(s^{-})$ (i.e., actual world $S = s^{-}$) is transferred along the other paths (denoted by $\bar{\pi}$). If π contains all direct edge from S to Y, $PSF_{\pi}(y)$ measures the direct discrimination. If π contains all indirect paths from S to Y that pass through redlining/proxy attributes, $PSF_{\pi}(y)$ evaluates the indirect discrimination. If π contains all indirect paths from S to Y that pass through explaining attributes, $PSF_{\pi}(y)$ assesses the explainable bias.

4.1.3. No unresolved/proxy discrimination

No unresolved discrimination[28] is a fairness notion which is based on Pearl's structural causal model framework and aims to detect indirect discrimination. This criterion is satisfied if there is no directed path from the sensitive attribute S to the outcome Y which is not blocked by the resolving variables. A resolving variable is any variable in a causal graph that is influenced by the sensitive attribute to a certain degree but accepted by practitioners as nondiscriminatory, which is very similar to the use of explanatory attributes in the statistics-based fairness notion. For example, Figure 5 shows three causal graphs of a simple loan example. There exists such discrimination in the causal graph shown in Figure 5(a) since the effects of *Race* on *Loan* can be transmitted along the causal paths *Race* $\rightarrow$ *Loan* and *Race* $\rightarrow$ *ZipCode* $\rightarrow$ *Loan*, while there is no unresolved discrimination, since the effects of *Race* on *Loan* can only be transmitted through resolved attribute *Income* along *Race*

$\rightarrow$ *Income* $\rightarrow$ *Loan*.

Similar to no unresolved discrimination, no proxy discrimination [28] also focuses on indirect discrimination. Given a causal graph, if this criterion is satisfied, the effects of the sensitive attribute S on the output Y cannot be transmitted through any proxy variable R (which is also denoted as redlining variable). A proxy variable is a descendant of sensitive attribute S and the ancestor of decision attribute Y. It is labeled as a proxy because it is exploited to capture the information of S. The outcome of an automated decision making Y exhibits no proxy discrimination if the equality of the following equation is valid for all potential proxies R:

$$P(Y|do(R = r_0)) = P(Y|do(R = r_1)) \quad \forall r_0, r_1 \in dom(R) \tag{6}$$

In other words, this notion implies that changing the value of R should not have any impact on the prediction. A simple example is shown in Figure 5. *ZipCode* is a redlining variable due to it reflects the information of the sensitive attribute *Race*. There is no proxy discrimination in causal graph shown in Figure 5(c), since the causal path *Race* $\rightarrow$ *ZipCode* $\rightarrow$ *Loan* has been blocked by intervening *ZipCode*.

No unresolved discrimination is a flawed definition of fairness. Specifically, no unresolved discrimination criterion is unable to identify some counterfactual unfair scenarios where some attributes are deemed as the resolved attributes. On the other hand, policy makers and domain professionals should carefully examine the relevance between sensitive variables and other endogenous variables so as to discover all resolving attributes and potential proxies that may lead to discrimination spread.

4.2. Individual causality-based fairness notions

Different from group fairness notions that measure the differences in the outcome of decision models between advantaged groups and disadvantaged ones, individual fairness notions aim to examine whether the outcome of decision models is fair to each individual in the population. Some representative group causality-based fairness notions are discussed here.

4.2.1. Counterfactual fairness

An outcome Y achieves counterfactual fairness towards an individual i (i.e., $\mathbf{O} = \mathbf{o}$) if the probability of the outcome $Y = y$ for such individual i is the same as the probability of $Y = y$ for the same individual whose value of sensitive attribute changing to another one. Formally, counterfactual fairness can be expressed as follows for any $\mathbf{O} = \mathbf{o}$:

$$|P(y_{s^+}|\mathbf{O} = \mathbf{o}, S = s^-) - P(y_{s^-}|\mathbf{O} = \mathbf{o}, S = s^-)| \leq \tau \tag{7}$$

where $\mathbf{O} \subseteq \mathbf{V} \setminus \{S, Y\}$ is the subset of endogenous variables except sensitive variables and decision variables. Any context $\mathbf{O} = \mathbf{o}$ represents a certain sub-group of the population, specifically, when $\mathbf{O} = \mathbf{V} \setminus \{S, Y\}$, it represents a specific individual. According to Equation (7), the decision model achieves counterfactual fairness if, for every possible individual ($\mathbf{O} = \mathbf{o}$, $S = s^-$) of the entire population, the probability distribution of the outcome Y is the same in both the actual ($S = s^-$) and counterfactual ($S = s^+$) worlds.

Counterfactual fairness was proposed by Kusner *et al.* [11]. They empirically tested whether the automated decision making systems are counterfactual fairness by generating the samples given the observed sensitive attribute value and their counterfactual sensitive value; then, they fitted decision models to both the original and counterfactual sampled data and examined the differences in the prediction distribution of predictor between the original and the counterfactual data. If an outcome Y is fair, the predictor is expected that the predicted results of actual and counterfactual distributions lie exactly on top of each other.

4.2.2. Individual direct discrimination

Individual direct discrimination[29] is a situation testing-based technique[35] guided by the structural causal model for analyzing the discrimination at the individual level. Situation testing is a legally grounded technique to detect the discrimination against a target individual by comparing the outcome of the individuals similar to the target one from both the advantaged group and the disadvantaged one in the same decision process. In other words, for a target individual i, select top-K individuals most similar to i from the group $S = s^+$ (denoted as G^+) and top-K individuals most similar to i from the group $S = s^-$ (denoted as G^-), and then perform one-to-one pairing according to the similarity ranking. The target individual is considered as discriminated if a significant difference is observed between the rate of positive decisions for all pairs from G^+ and G^- (typically, higher than fair threshold τ). The key issue for implementing situation testing is how to define a distance function $d(i, i')$ to measure similarity between individuals.

For the individual direct discrimination criterion, the distance function $d(i, i')$ is defined not only by adopting normalized Manhattan distance and overlap measurement but also by incorporating causal inference. Specifically, given a causal graph, only the variables that are direct parent nodes of the decision variable are considered to compute the similarity between individuals, which are denoted as $\mathbf{X} = Pa(Y) \setminus \{S\}$. The formal definition of $d(i, i')$ is as follows:

$$d(i, i') = \sum_{k=1}^{|\mathbf{X}|} |CE(x_k, x'_k) \cdot VD(x_k, x'_k)| \tag{8}$$

where $CE(x_k, x'_k)$ represents the causal effect of each of the selected variables ($X_k \in \mathbf{X}$) on the outcome and $VD(x_k, x'_i)$ is a distance function proposed by Luong *et al.*[36]. Specifically, the normalized Manhattan distance is employed for ordinal/interval variables (i.e., $VD(x_k, x'_k) = \frac{|x_k - x'_k|}{range}$, where $range$ denotes the difference between the maximum and minimum of the variable X_k) and the overlap measurement is employed for categorical variables (i.e., $VD(x_k, x'_k) = 0$ if $x_k = x'_k$, and $VD(x_k, x'_k) = 1$ otherwise).

For each selected variable X_k, the definition of $CE(x_k, x'_k)$ is as follows:

$$CE(y) = P(y|do(\mathbf{x})) - P(y|do(x'_k, \mathbf{x} \setminus x_k)) \tag{9}$$

where $P(y|do(\mathbf{x}))$ is the effect of the intervention that forces $\mathbf{X}$ to take the set of values $\mathbf{x}$, and $P(y|do(x'_k, \mathbf{x} \setminus x_k))$ is the effect of the intervention that forces X_k to take value x'_k and other variables in $\mathbf{X}$ to take the same values as $\mathbf{x}$.

4.2.3. Equality of effort

Equality of effort fairness notion[30] detects bias by comparing the effort required to reach the same level of outcome of individuals from advantaged and disadvantaged groups who are similar to the target individual. That is to say, given a treatment variable T, it quantifies how much this treatment variable T should change to make the individual achieve a certain outcome level in order to address the concerns of whether the efforts required to achieve the same level of outcome for individuals from the advantaged and disadvantaged groups are different. Following Rubin's potential outcome model framework[37], let Y_i^t be the potential outcome for individual i had T been t and $\mathbb{E}[Y_i^t]$ be the expectation of outcome for individual i. Then, for individual i, the needed minimal value of treatment variable T to achieve γ-level of outcome is defined as follows:

$$\Psi_i(\gamma) = argmin_{t \in T} \mathbb{E}[Y_i^t] \geqslant \gamma \tag{10}$$

Unfortunately, Y_i^t is not observable, which results in $\Psi_i(\gamma)$ being uncomputable. Situation testing is then used to estimate it, where the distance function $d(i, i')$ of equality of effort is the same as individual direct discrimination mentioned in Section 4.2.2. Let G^+ and G^- be the two sets of individuals with $S = s^+$ and $S = s^-$ that are similar to the target individual i, respectively, and $\mathbb{E}[Y_{G^+}^t]$ be the expected outcome under

treatment t for the subgroup G^+. The minimal effort required to achieve γ-level of outcome variable within the subgroup G^+ is computed as follows:

$$\Psi_{G^+}(\gamma) = argmin_{t \in T} \mathbb{E}[Y^t_{G^+}] \geqslant \gamma \tag{11}$$

Then, for a certain outcome level γ, individual γ-equal effort is satisfied for individual i if:

$$\Psi_{G^+}(\gamma) = \Psi_{G^-}(\gamma) \tag{12}$$

Equality of effort can also be extended to identify discrimination at any sub-population level or system level, when G^+ is extended to the entire group with $S = s^+$ and G^- denotes the entire group with $S = s^-$. To distinguish individual γ-equal effort, D^+ is used to denote the first set, while D^- denoted the second one. The γ-equal effort is satisfied for a sub-population if:

$$\Psi_{D^+}(\gamma) = \Psi_{D^-}(\gamma) \tag{13}$$

4.2.4. PC-Fairness

Path-specific Counterfactual Fairness (PC-fairness)[31] is used to denote a general fairness formalization for representing various causality-based fairness notions. Given a factual condition $\mathbf{O} = \mathbf{o}$ where $\mathbf{O} \in \mathbf{V}$ and a causal path set π, a predictor $\hat{Y}$ achieves the PC-fairness if it satisfies the following expression:

$$PC_\pi(\hat{y}_{s^- \to s^+} | \mathbf{o}) \leqslant \tau \tag{14}$$

where $PC_\pi(\hat{y}_{s^- \to s^+} | \mathbf{o}) = P(\hat{y}_{s^+|\pi, s^-|\bar{\pi}} | \mathbf{o}) - P(\hat{y}_{s^-} | \mathbf{o})$ and τ is a predefined fairness threshold (typically, 0.05). Intuitively, $PC_\pi(\hat{y}_{s^- \to s^+} | \mathbf{o})$ denotes when the value of the sensitive attribute S changes from s^+ to s^-, the causal effect of S on $\hat{Y}$ through the causal path set π and given the factual observation $\mathbf{o}$.

PC-fairness matches different causality-based fairness notions by tuning its parameters. For example, if the path set π contains all causal paths and $\mathbf{O} = \phi$, PC-fairness corresponds to the total effects in Equation (3). Apart from that, it also includes new types of fairness that have not been studied yet in the past. For example, PC-fairness can detect individual indirect discrimination by letting $\mathbf{O} = \mathbf{V} \setminus \{Y\}$ and the path set π containing all causal paths that pass through any redlining variables.

5. CAUSALITY-BASED FAIRNESS-ENHANCING METHODS

The need for causal models for detecting and eliminating discrimination is based on the intuition that the same individuals experience different outcomes due to innate or acquired characteristics outside of their control (e.g., gender). Therefore, causal models are useful for investigating which characteristics cannot be controlled by individuals and using the resulted understandings to identify and deal with discrimination. In other words, understanding the structure of root causes of the problem can assist in identifying unfairness and causes. Thus, there is a causal structure that must be considered rather than just the correlation between the sensitive attribute and outcome. Because of these advantages, many recent studies introduce fairness-enhancing approaches from the perspective of causality. According to the stages of training the machine learning algorithms, pre-processing, in-processing, and post-processing mechanisms can be used to intervene in the algorithm to achieve fairness. Therefore, causal-based methods can be divided into the above three categories. Figure 6 shows the general flow of different categorical causality-based approaches. This section provides an overview of studies for these categories, and then the advantages and disadvantages of these three types of mechanisms are summarized.

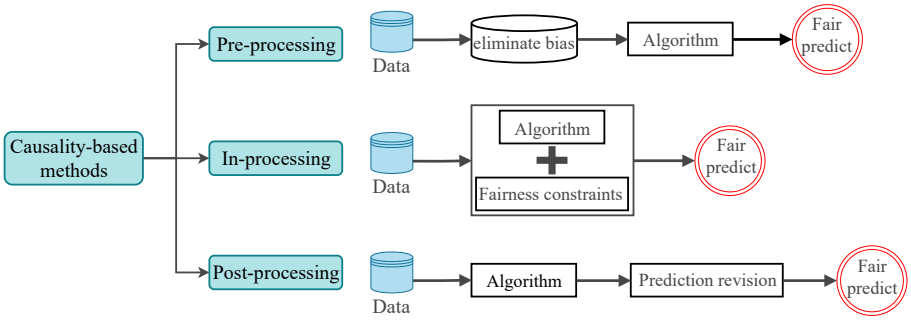


Figure 6. The categorization of causality-based fairness-enhancing approaches.

5.1. Pre-processing Causality-based methods

Pre-processing methods update the training data before feeding them into a machine learning algorithm. Specifically, one idea is to change the labels of some instances or reweigh them before training to limit the causal effects of the sensitive attributes on the decision. As a result, the classifier can make a fairer prediction[36]. On the other hand, some studies propose to reconstruct the feature representations of the data to eliminate discrimination embedded in the data[38,39].

For example, Zhang *et al.*[27] formalized the presence of discrimination as the presence of a certain path-specific effect, and then framed the problem as one maximizing the likelihood subject to constraints that restrict the magnitude of the PSE. To deal with unidentifiable cases, they mathematically bound the PSE. CFGAN[40], which is based on Causal GAN[41] to learn the causal relationship between the attributes, adopts two generators to separately simulate the underlying causal model that generates the real data and the causal model after the intervention and two discriminators to produce a close to real distribution and to achieve total effect fairness, counterfactual fairness, and path-specific fairness. Salimi *et al.*[42,43] leveraged the dependencies between sensitive and other attributes, which is provided by the causal knowledge, to add or remove samples from the collected datasets in order to eliminate the discrimination. Nabi *et al.*[44] only considered mapping generative models for $P(Y, \mathbf{V} \setminus \mathbf{W} | \mathbf{W})$ consisting of some attributes $\mathbf{W}$ to "'fair" versions of this distribution $P*(Y, \mathbf{V} \setminus \mathbf{W} | \mathbf{W})$ and ensured that $P*(\mathbf{W}) = P(\mathbf{W})$. PSCF-VAE[45] achieves path-specific counterfactual fairness by modifying the observed values of the descendant attribute of the sensitive attribute on the unfair causal path during testing, leaving the underlying data-generation mechanism unaltered during training.

5.2. In-Processing Causality-based methods

In-processing methods eliminate discrimination by adding constraints or regularization terms to machine learning models[46–50]. If it is allowed to change the learning procedure for a machine learning model, then in-processing can be used during the training of a model either by incorporating changes into the objective function or imposing a constraint.

For example, multi-world fairness algorithms[12] add constraints to the classification model that require satisfying the counterfactual fairness. To address the unidentifiable situation and alleviate the difficulty of determining causal models, it combines multiple possible causal models to make approximately fair predictions. A tuning parameter α is used to modulate the trade-off between fairness and accuracy. Hu *et al.*[51] proposed to learn multiple fair classifiers simultaneously from a static training dataset. Each classifier is considered to perform soft interventions on the decision, whose influence is inferred as the post-intervention distributions to formulate loss functions and fairness constraints. Garg *et al.*[52] proposed to penalize the discrepancy between real samples and their counterfactual samples by adding counterfactual logit pairing (CLP) to the loss function of the algorithm. Similarly, the authors of[53] proposed to add constraints terms during the training to eliminate the difference in outcome between two identical individuals, one from the real world and one

from the counterfactual world that belongs to another sensitive group. Kim *et al.*[54] addressed the limitation that some causality-based methods cannot distinguish between information caused by the intervention (i.e., sensitive variables) and information related to the intervention by decomposing external uncertainty into intervention-independent variables and intervention-related ones. They proposed a method called DCEVAE, which can estimate the total effect and counterfactual effects in the absence of full causal maps.

5.3. Post-processing Causality-based methods

Post-processing methods modify the outcome of the decision model to make fairer decisions[55,56]. For example, Wu *et al.*[57] adopted the c-component factorization to decompose the counterfactual quantity, identified the sources unidentifiable terms, and developed the lower and upper bounds of counterfactual fairness in unidentifiable situations. In the post-processing stage, they reconstructed the trained decision model so as to achieve counterfactual fairness. The counterfactual privilege algorithm[58] maximizes the overall benefit while preventing an individual from obtaining beneficial effects exceeding the threshold due to the sensitive attributes, so as to make the classifier achieve counterfactual fairness. Mishler *et al.*[59] suggested using doubly robust estimators to post-process a trained binary predictor in order to achieve approximate counterfactual equalized odds.

5.4. Which mechanism to use

We discuss the various mechanisms for enhancing fairness above. Here, we further compare these mechanisms and discuss the advantages and disadvantages of them, respectively. This section provides insights into how to select suitable mechanisms for use in different scenarios based on the characteristics of these mechanisms. Every type of mechanism has its advantages and disadvantages.

The pre-processing mechanism can be flexibly adapted to the downstream tasks since it can be used with any classification algorithm. However, since the pre-processing mechanism is a general mechanism where the extracted features can be widely applicable for various algorithms, there is high indeterminacy regarding the accuracy of the trained decision models.

Similar to the pre-processing mechanism, the post-processing mechanism also can be flexibly used in any decision model. Post-processing mechanisms are easier to fully eliminate discrimination for the decision models, but the accuracy of the decision models depends on the performance they obtained in the training stage[60]. Furthermore, post-processing mechanisms require access to all information of individuals during testing, which may be unavailable because of reasons of privacy protection.

The in-processing mechanism is beneficial to enable a balance between accuracy and fairness of the decision model, which is achieved by explicitly modulating the trade-off parameter in the objective function. However, such mechanisms are tightly coupled with the machine learning algorithm itself and are difficult to optimize in the application.

Based on the above discussion and the studies that attempt to comprehend which mechanism is best to use in certain situations[61,62], we can say that there is no single mechanism that outperforms the others in all cases, and the choice of suitable mechanisms depends on the availability of sensitive variables during testing, the characteristics of the dataset, and the desired fairness measure in the application. For example, when there exists evident selection bias in a dataset, it is better to select the pre-process mechanism for use than the in-process one. Therefore, more research is needed to develop robust fairness mechanisms or to design suitable mechanisms for practical scenarios.

6. APPLICATIONS OF FAIR MACHINE LEARNING

This section enumerates different domains of machine learning and the work that has been produced by each domain to combat discrimination in their methods.

6.1. Data missing

One major challenge for fairness-enhancing algorithms is to deal with the biases inherent in the dataset that is caused by missing data. Selection biases are due to the distribution of collected data, which cannot reflect the real characteristics of disadvantaged groups. Martínez-Plumed *et al.*[63] learned that selection bias is mainly caused by individuals in disadvantaged groups being reluctant to disclose information, e.g., people with high incomes are more willing to share their earnings than people with low incomes, which results in bias inference that training in the training institution helps to raise earnings. To address this problem, Bareinboim *et al.*[64] and Spirtes *et al.*[65] studied how to deal with missing data and repair datasets that contain selection biases by causal reasoning, in order to improve fairness.

On the other hand, the collected data represent only one side of the reality, that is, these data do not contain any information about the population who were not selected. Biases may arise that decide which data are contained or not contained in the datasets. For example, there is a dataset that records the information of individuals whose loans were approved and the information about their ability to repay their loans. Although the automatic decision system that satisfies certain fairness requirements is constrained based on this dataset to predict whether they repay their loan on time, such a predictor may be discriminatory when it is used to assess the credit score of further applicants, since populations not approved for loans are not sufficiently representative in the training data. Goel *et al.*[66] used the causal graph-based framework to model the causal process of possible missing data for different settings by which different types of decisions are made in the past, and proved some data distributions can be inferred from incomplete available data based on the causal graph. Although the practical scenarios they discussed are not exhaustive, their work shows that the causal structure can be used for determining the recoverability of quantities of interest in any new scenario.

A promising solution for dealing with missing data can be found in causality-based methods. We see that causality can provide tools to improve fairness when the dataset suffers from discrimination caused by missing data.

6.2. Fair recommender Systems

Recommenders are recognized as the most effective way to alleviate information overloading. Nowadays, recommender systems have been widely used in variable applications, such as ecommerce platforms, advertisements, news articles, jobs , etc. They are not only used to analyze user behavior to infer users' preferences so as to provide them with personalized recommendations, but they also benefit content providers with more potential of making profits. Unfortunately, there exist fairness issues in recommender systems[67], which are challenging to handle and may deteriorate the effectiveness of the recommendation. The discrimination embedded in the recommender systems is mainly caused by the following aspects. User behaviors in terms of the exposed items make the observational data confounded by the exposure mechanism of recommenders and the preference of the users. Another major cause of discrimination in recommender systems is that disadvantage items reflected in the observational data are not representative. That is to say, some items may be more popular than others and thus receive more user behavior. As a result, recommender systems tend to expose users to these popular items, which results in discrimination towards unpopular items and leads to the systems not providing sufficient opportunities for minority items. Finally, one characteristic of recommender systems is the feedback loop. That is, the systems exposes to the user for determining the user behavior, which is circled back as the training data for the recommender systems. Such a feedback loop not only creates biases but also intensifies biases over time, resulting in "the rich get richer" Matthew effect.

Due to the usefulness of causal modeling[10], removing discrimination for recommender systems from a causal perspective has attracted increasing attention, where the cause graph is used for exposing potentially causal relationships from data. On the one hand, most discrimination can be understood with additional confounding factors in the causal graph and the effect of discrimination can also be inferred through the causal graph. On the other hand, recommendation can be considered as an intervention, which is similar to treating a patient with a specific drug, requiring counterfactual reasoning. What happens when recommending certain items to the users? The causal model has the potential to answer this question. For example, Wu *et al.*[68] focused on fairness-aware ranking and proposed to use path-specific effects to detect and remove the direct and indirect rank discrimination. Zhao *et al.*[69] and Zheng *et al.*[70] considered the effect of item popularity on user behavior and intervened in the item popularity to make fair recommendations. Zhang *et al.*[71] attributed popularity bias in the recommender systems to the undesirable causal effect of item popularity on items exposure and suggested intervening in the distribution of the exposed items to eliminate this causal effect. Wang *et al.*[72] leveraged counterfactual reasoning to eliminate the causal effect of exposure features on the prediction. Li *et al.*[73] proposed generating embedding vectors independent of sensitive attributes by adversarial learning to achieve counterfactual fairness. Huang *et al.*[74] regarded causal inference as bandits and performed *do*-operator to simulate the arm selection strategy to achieve fairness towards individuals.

Nowadays, the explainability of recommender systems is increasingly important, which improves the persuasiveness and trustworthiness of recommendations. When addressing the fairness issues of recommender systems from the causal perspective, the explanation of recommendations can also be provided from the effects transmitted along the causal paths. Thus, we are confident that causal modeling will bring the recommendation research into a new frontier.

6.3. Fair natural language processing

Natural language processing (NLP) is an important technology for machines to understand and interpret human natural language text and realize human–computer interaction. With the development and evolution of human natural language, the natural language is characterized by a certain degree of gender, ethnicity, region, and culture. These characteristics are sensitive in certain situations, and inappropriate use can lead to prejudice and discrimination. For example, Zhao *et al.*[75] found that the datasets associated with multi-label object classification and visual semantic role labeling exhibit discrimination towards gender attribute, and, unfortunately, the model trained with these data would further amplify the disparity. Stanovsky *et al.*[76] provided multilingual quantitative evidence of gender bias in large-scale translation. They found that, among the eight target languages, all four business systems and two academic translation systems tend to translate according to stereotype rather than context. Huang *et al.*[77] used counterfactual evaluations to investigate whether and how language models are affected by sensitive attributes (e.g., country, occupation, and gender) to generate sentiment bias. Specifically, they used individual fairness metrics and group fairness metrics to measure counterfactual sentiment bias, conducted model training on news articles and Wikipedia corpus, and showcased the existence of sentiment bias.

Fair NLP is a kind of NLP without bias or discrimination with sensitive attributes. Shin *et al.*[78] proposed a counterfactual reasoning method for eliminating the gender bias of word embedding, which aims to disentangle a latent space of a given word embedding into two disjoint encoded latent spaces, namely the gender latent space and the semantic latent space, to achieve disentanglement of semantic and gender implicit descriptions. To this end, they used a gradient reversal layer to prohibit the inference about the gender latent information from semantic information. Then, they generated a counterfactual word embedding by converting the encoded gender into the opposite gender and used it to produce a gender-neutralized word embedding after geometric alignment regularization. As such, the word embedding generated by this method can strike a balance between gender debiasing and semantic information preserving. Yang and Feng[79] presented a causality-based post-processing approach for eliminating the gender bias in word embeddings. Specifically, their method was

based on statistical correlation and half-sibling regression, which leverages the statistical dependency between gender-biased word vectors and gender-definition word vectors to learn the counterfactual gender information of an individual through causal inference. The learned spurious gender information is then subtracted from the gender-biased word vectors to remove the gender bias. Lu *et al.* [80] proposed a method called CDA to eliminate gender bias through counterfactual data augmentation. The main idea of CDA is to augment the corpus by exchanging gender word pairs in the corpus and constructing matching gender word pairs with causal interventions. As such, CDA breaks associations between gendered and gender-neutral words and alleviates the problem that gender bias increases as loss decreases when training with gradient descent.

There exists a certain degree of bias and fairness issues in word embedding, machine translation, sentiment analysis, language models, and dialog generation in NLP. At present, most studies only focus on a single bias (such as gender bias), and there is a lack of research results on other biases or eliminating multiple biases at the same time. Therefore, how should we analyze and evaluate the mechanism and impact of multi-bias in word embedding and machine learning algorithms? Establishing effective techniques for eliminating various biases in word embedding and machine learning algorithms requires further research which needs to be carried out for fair NLP.

6.4. Fair medical

Electronic health records (EHRe) contain large amounts of clinical information about heterogeneous patients and their responses to treatments. It is possible for machine learning techniques to efficiently leverage the full extent of EHRs to help physicians make predictions for patients, thus greatly improving the quality of care and reducing costs. However, because of discrimination implicitly embedded in EHRs, the automated systems may introduce or even aggravate the nursing gap between underrepresented groups and disadvantaged ones. The prior works on eliminating discrimination for clinical predictive models mostly focus on statistics-based fairness-enhancing approaches [81,82]. In addition, they do not provide an effective evaluation of fairness to individuals, and the fairness metrics they used are difficult to verify. Some recent studies focus on assessing fairness for clinical predictive models from a causal perspective [83,84]. For example, Pfohl *et al.* [83] proposed a counterfactual fairness notion to extend fairness to the individual level and leveraged variational autoencoder technology to eliminate discrimination against certain patients.

6.5. Causal analysis packages

This section introduces some representative packages or software for causal analysis, which are helpful for us to develop causality-based fairness-enhancing approaches. These packages can be roughly divided into two categories: one for discovering potential causal structure in data and the other for making causal inferences. Table 2 summarizes typical packages or software for causal analysis.

TETRAD [85] is a full-featured software for causal analysis after considerable development where it can be used to discover the causal structure behind the dataset, estimate the causal effects, simulate the causal models, etc. TETRAD can accept different types of data as input, e.g., discrete data, continuous data, time series data, etc. The users can choose the appropriate well-tested causal discovery algorithms it integrates to search causal structure, as well as input prior causal knowledge to limit the search. In addition, TETRAD can parameterize the causal model and simulate the data according to the existing causal diagram. Causal-learn package [86] is the Python version of TETRAD. It provides the implementation of the latest causal discovery methods ranging from constraint-based methods, score-based methods, and constrained functional causal models-based methods to permutation-based methods. In addition, there are many packages for causal discovery [87–89]. Tigramite [88] focuses on searching causal structure from observational time series data. In addition to providing classic causal discovery algorithms, gCastle [89] provides many gradient-based causal discovery approaches.

CausalML [90] is a Python package which encapsulates many causal learning and causal inference approaches.

Table 2. Typical packages or software for causal analysis

Type	Package name	Program language	Description
Causal discovery	TETRAD [85]	Java	TETRAD is a full-featured software for causal analysis; after considerable development, it can be used to discover the causal structure behind the dataset, estimate the causal effects, simulate the causal models, etc
	Py-causal [87]	Python	Py-causal is a Python encapsulation of TETRAD, which can call the algorithms and related functions in TETRAD
	Causal-learn [86]	Python	Causal-learn is the Python version of TETRAD. It provides the implementation of the latest causal discovery methods ranging from constraint-based, score-based, and constrained functional causal models-based to permutation-based methods
	Tigramite [88]	Python	Tigramite focuses on searching causal structure from observational time series data.
	gCastle [89]	Python	gCastle provides many gradient-based causal discovery approaches, as well as classic causal discovery algorithms
Causal effect and Inference	CausalML [90]	Python	CausalML encapsulates many causal learning and inference approaches. One highlight of this software package is uplift modeling, which is used to evaluate the conditional average treatment effect (CATE)
	Causaleffect [92]	R	Causaleffect is the implementation of ID algorithm
	DoWhy [93]	Python	DoWhy takes causal graphs as prior knowledge and uses Pearl's do-calculus method to assess causal effects
	Mediation [91]	R	Mediation provides model-based method and design-based method to evaluate the potential causal mechanisms. It also provides approaches to deal with common problems in practice and random trials, that is, to handle multiple mediators and evaluate causal mechanisms in case of intervention non-compliance

One highlight of this package is uplift modeling, which is used to evaluate the conditional average treatment effect (CATE), that is, to estimate the impact of a treatment on a specific individual's behavior.

Mediation [91] is an R package which is used in causal mediation analysis. In other words, it provides model-based methods and design-based methods to evaluate the potential causal mechanisms. It also provides approaches to deal with common problems in practice and random trials, that is, to handle multiple mediators and evaluate causal mechanisms in case of intervention non-compliance.

Causaleffect [92] is an R package which is the implementation of ID algorithm. ID algorithm is a complete identification of causal effects algorithm, which outputs the expression of causal effect when the causal effect is identifiable or fails to run when the causal effect is unidentifiable. DoWhy [93], a Python package, also focuses on causal inference, that is, it takes causal graphs as prior knowledge and uses Pearl's *do*-calculus method to assess causal effects.

These packages used for causal analysis assist in developing causality-based fairness-enhancing methods, which are mainly reflected in exposing the causal relationship between variables and evaluating the causal effects of sensitive attributes on decision-making. However, they cannot be used directly to detect or eliminate discrimination. Although there are many software packages for detecting and eliminating discrimination, e.g., AI Fairness 360 Open Source Toolkit [94], Microsoft Research Fairlearn [95], etc, we are still lacking a package that integrates causality-based approaches.

7. CHALLENGES

Decision based on machine learning has gradually penetrated into all aspects of human society, and the fairness of its decision-making directly affects the daily life of individuals or groups, as well as users' trust and acceptance of machine learning application deployment. Recently, fair machine learning has received extensive attention, and researchers are gradually aware of the fact that relying solely on the observable data, with no additional causal information, is limited in removing discrimination, since the dataset only represents the selected population, without any information on the groups who were not selected, while such information

can be achieved using knowledge of a causal graph or by a controlled experiment making use of interventions. As such, causality-based fairness machine learning algorithms have attracted more and more attention and several causality-based fairness approaches have been proposed. Although causal fairness models can indeed help us overcome many of the challenges encountered with respect to fair prediction tasks, they still face many challenges, which are discussed in the following subsections.

7.1. Causal discovery

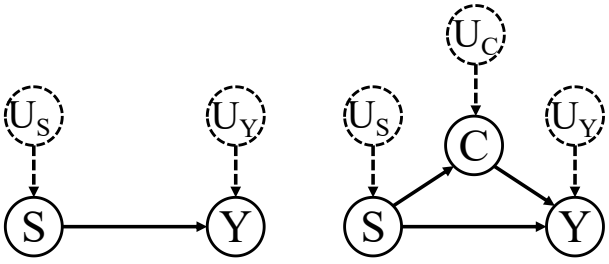
Causality-based fairness approaches require a causal graph as additional prior knowledge of input, where the causal graph describes the mechanism by which the data are generated, that is, it reveals the causal relationship between variables. However, in practice, it is difficult for us to obtain the correct causal graph. A basic way to discover the causal relationship between variables is to conduct randomized controlled trials. Randomized controlled trials consist of randomly assigning subjects (e.g. individuals) to treatments (e.g. gender), and then comparing the outcome of all treatment groups. Unfortunately, in many cases, it may not be possible to undertake such experiments due to prohibitive costs, ethical concerns, or they are physically impossible to carry out. For example, to understand the impact of smoking, it would be necessary to force different individuals to smoke or not smoke. As another example, to understand whether hiring decision models are gender-biased, it would be necessary to change the gender of a job applicant, which is an impracticality. Researchers are therefore often left with non-experimental, observational data, and they have developed numerous methods for uncovering causal relations, i.e., causal discovery. Causal discovery algorithms can be roughly classified into the following three categories: constraint-based, score-based, and those exploiting structural asymmetries.

Constraint-based approaches conduct numerous conditional independence tests to learn about the structure of the underlying causal graph that reflects these conditional independence. Constraint-based approaches have the advantage that they are generally applicable, but the disadvantages are that faithfulness is a strong assumption and that it may require very large sample sizes to get good conditional independence tests. Furthermore, the solution of this approach to causal discovery is usually not unique, and, in particular, it does not help determine the causal direction in the two-variable case, where no conditional independence relationship is available.

Score-based algorithms use the fact that each directed acyclic graph (DAG) can be scored in relation to the data, typically using a penalized likelihood score function. The algorithms then search for the DAG that yields the optimal score. Typical scoring functions include the Bayesian information criterion [96], Bayesian–Gaussian equivalent score [96], and minimum description length (as an approximation of Kolmogorov complexity) [97,98].

Structural asymmetry-based algorithms take into account the setting that it is impossible to infer the causal direction from observations alone when the data distribution admits structural causal models indicating either of the structural directions $X_i \to X_j$ or $X_i \leftarrow X_j$. To address this problem, structural asymmetry-based algorithms make some additional assumptions about the function of the underlying true data-generating structure, so that they can exploit asymmetries to identify the direction of a structural relationship. These asymmetries manifest in various ways, such as non-independent errors, measures of complexity, etc. Existing methods that exploit such asymmetries are typically local solutions, as they are only able to test one edge at a time (pairwise/bivariate causal directionality), or a triple (with the third variable being an unobserved confounder) [99].

In the absence of intervention and manipulation, observational data leave researchers facing a number of challenges: First, there may exist hidden confounders, which are sometimes termed the third variable problem. Second, observational data may exhibit selection bias. For example, younger patients may generally prefer surgery, while older patients may prefer medication. Third, most causal discovery algorithms are based on strong but often untestable assumptions, and applying these strong assumptions to structural or graphical models incites some harsh criticisms.



(a) A Simple causal graph with no confounder between sensitive attribute and outcome.

(b) A Simple causal graph with an observational confounder between sensitive attribute and outcome.

Figure 7. Simple causal graphs under Markovian assumption.

7.2. Identifiable issue

The identifiable issue is another main obstacle to the application of causal models in fair machine learning. The identifiability of causal effects, including total effect, path-specific effect, and counterfactual effect, has been extensively studied[10,100–103]. Many causality-based fairness methods have been proposed to solve this problem so as to achieve fairness more effectively. This section summarizes the main identifiability conditions, the causality-based fairness methods that try to overcome the unidentifiable situations, and their limitations.

7.2.1. Identifiability

This review starts with the simplest identifiability condition of total effects, where the causal effect of a sensitive attribute S on decision Y is computed by $P(Y|do(S = s))$, which denotes the distribution of decision Y after the intervention $S = s$. Similar to causal discovery, randomized controlled trials can also be used to assess the total effect by performing an intervention on the sensitive attributes to completely avoid unidentifiable situations of the total effect. A randomized controlled trial may allow us not only to identify causal relationships but also to estimate the magnitude of these relationships. However, as mentioned in Section 7.1, it is impossible for us to undertake such experiments in many cases. The causal effect of the sensitive attribute S on decision Y may not be uniquely assessed from observational data and causal graph alone, i.e., the unidentifiable situation. In Markovian models, the total effect is always identifiable, since $P(Y|do(S = s))$ is always identifiable[10]. The simplest case is when there is no confounder between S and Y (see Figure 7(a)). In this case, the causal effect $P(Y|do(S = s))$ is consistent with the conditional probability $P(y|S = s)$. As such, the total effect can be computed as follows:

$$TE(y) = P(y|do(S = s^+)) - P(y|do(S = s^-)) = P(y|s^+) - P(y|s^-) \tag{15}$$

If there exists an observational confounder between S and Y (see Figure 7(b)), the total effect is also identifiable by summing the probabilities of all values c in the domain of the observable confounder C:

$$TE(y) = P(y|do(S = s^+)) - P(y|do(S = s^-)) = \sum_C P(y|s^+, c)P(c) - \sum_C P(y|s^-, c)P(c) \tag{16}$$

where the formula $\sum_C P(y|s, c)P(c)$ in Equation (16) is also called the back-door formula. If there are no hidden confounders within the observable attributes, then the causal effect $P(y|do(S = s))$ can also be computed by the following formula:

$$P(y|do(s)) = \sum_{\mathbf{V}\setminus\{S,Y\}, Y=y} \prod_{V\in\mathbf{V}\setminus\{S\}} P(v|Pa(V)) \tag{17}$$

where $Pa(V)$ represents the parent variables of V. It is easy to see that the total effects computed by Equations (16) and (17) can produce the same result.

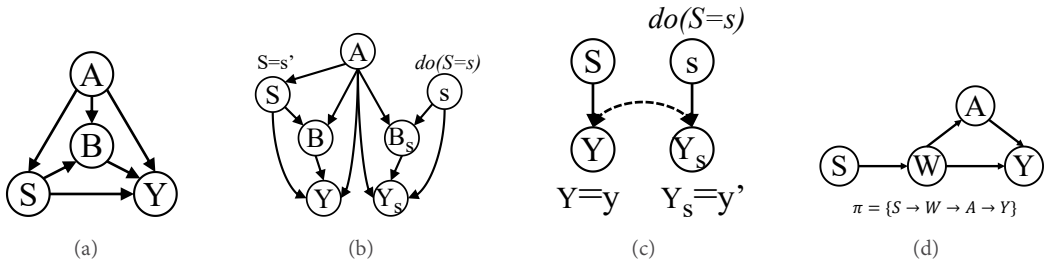


Figure 8. (a) The tony causal graph; (b) the counterfactual graph of (a); (c) the W-graph; and (d) the "kite" graph.

For semi-Markovian models, the causal effect $P(y|do(s))$ is not always identifiable, hence the total effect is not always identifiable. The causal effect $P(y|do(s))$ is identifiable if and only if it can be reduced to a do-free expression (i.e., turning the intervention operator $do(s)$ to observational probabilities) by do-calculus[10]. do-calculus is composed of three inference rules: (i) insertion/deletion of observations, i.e., $P(y_s|\mathbf{z}, w) = P(y_s|\mathbf{z})$ provided that Y and W are dependence-separated at fixed S and $\mathbf{Z}$ after all arrows leading to S have been deleted in causal graph; (ii) action/observation exchange, i.e., $P(y|do(s), \mathbf{z}) = P(y|s, \mathbf{z})$ if Y and S are probabilistically conditionally independent at fixed $\mathbf{Z}$ after deleting all arrows starting from $\mathbf{Z}$ in causal graph; and (iii) deletion of actions, i.e., $P(y|do(s)) = P(y)$ if there are no causal paths between S and Y.

do-calculus has been proven to be complete, that is, it is sufficient to derive all identifiable causal effects by do-calculus[100]. However, it is difficult to determine the correct order of application of these rules, and the wrong order may misjudge the identifiability of causal effects or produce a very complex expression. To address this issue, several studies attempt to give the explicit graphical criteria and map them to simple and concise do-free expressions[101,104]. A simple case of the identifiability of the causal effect $P(y|do(s))$ is when the sensitive attribute S is not influenced by any confounder[105]. In other words, the causal effect of S is identifiable, if all parents of S are observable. Graphically, there is no bi-directional edge connected to S. Formally, the causal effect $P(y|do(s))$ can be computed as follows:

$$P(y|do(s)) = \sum_{\mathbf{pa}(S)} P(y|s, pa(S))P(\mathbf{pa}(S)) \tag{18}$$

where $pa(S)$ represents the values of parent variables of S.

A complex case where the causal effect of S on $\mathbf{V}' = \mathbf{V} \setminus \{S\}$ is identifiable is that there may exist a bi-directed edge connected to the sensitive attribute S, but there are no hidden confounders connected to any direct child of S[105]. Graphically, there is no bi-directional edge connected to any child of S. If such criterion is satisfied, the causal effect $P(\mathbf{v}'|do(s))$ is identifiable and is given by:

$$P(\mathbf{v}'|do(s)) = \Big(\prod_{v_i \in Ch(s)} P(v_i|\mathbf{pa}(v_i))\Big) \sum_{s} \frac{P(\mathbf{v}')}{\prod_{v_i \in Ch(s)} P(v_i|\mathbf{pa}(v_i))} \tag{19}$$

where $Ch(s)$ denotes the set of S's children and $\mathbf{pa}(v_i)$ denotes the set of values of V_i's parents. Equation (19) can be easily adapted to assess the effect of the sensitive attribute S on outcome Y.

Tian *et al.*[105] also found that, although $P(\mathbf{v}'|do(s))$ is not identifiable, $P(\mathbf{w}|do(s))$ is still identifiable for some subsets $\mathbf{W}$ of $\mathbf{V}$. Specifically, causal effect $P(\mathbf{w}|do(s))$ is identifiable if there is no bi-directed path connecting S to any of its direct children in $\mathcal{G}_{An}(\mathbf{W})$, where $An(\mathbf{W})$ denotes the union of a set $\mathbf{W}$ and the set of ancestors of the variables in $\mathbf{W}$. $\mathcal{G}_{An}(\mathbf{W})$ denotes the sub-graph of $\mathcal{G}$ composed only of variables in $An(\mathbf{W})$.

Moreover, the causal effect of sensitive attribute S on outcome Y $P(y|do(s))$ can also be computed by using the front-door criterion, if $P(y|do(s))$ is identifiable. Before introducing the front-door criterion, the concepts of

the back-door path and front-door path are first introduced, which are helpful for the understanding of the front-door criterion. A causal path connecting S and Y which begins with an arrow leading to S is called a *back-door path* (e.g., $S \leftarrow ...Y$), while a path starting with an arrow pointing away from S is called a *front-door path* (e.g., $S \rightarrow ...Y$). If the front-door criterion is satisfied, there exits a mediator variable Z such that: (i) there are no backdoor paths from S to Z; and (ii) all backdoor paths from S to Z are blocked by S. Formally, $P(y|do(s))$ can be computed as follows:

$$P(y|do(s)) = \sum_{Z} P(y|s,z)P(z|s)P(s) \tag{20}$$

These criteria can be generalized to the case where there is no bi-directed edges connection between the sensitive attribute and its direct children. Given the fact that the observational distribution $P(\mathbf{v})$ can be decomposed to a product of several factors, c-component factorization was proposed to decompose the identification problem into smaller problems (i.e., c-components, each of which is a set of observational variables that are connected by a common confounder in the causal graph) to evaluate the causal effect[15,105].

Shpitser *et al.*[15] designed a sound and complete algorithm called *ID* to identify all identifiable causal effects where *ID* outputs the expression of the causal effect for the identifiable cases. In addition, they proved that all cases of unidentifiable causal effects $P(y|do(s))$ can be completely attributed to a graphical structure called *hedge*.

As to the identifiability of counterfactual effect, if complete knowledge of the causal model is known (including structural functions, $P(\mathbf{u})$, etc.), any counterfactual quantity can be exactly performed using three steps: (i) abduction, update $P(\mathbf{u})$ by observation $\mathbf{O} = \mathbf{o}$ to compute $P(\mathbf{u}|\mathbf{o})$; (ii) action, modify causal model $\mathcal{M}$ by intervention $do(s)$ to obtain the post-intervention model $\mathcal{M}_s$; and (iii) prediction, use post-intervention model $\mathcal{M}_s$ and $P(\mathbf{u}|\mathbf{o})$ to compute the counterfactual effect $P(y_s|s')$. However, the above method is usually infeasible in practice due to the lack of complete knowledge of the causal model. In most cases, we only have the causal graph and observational data, which makes the counterfactual effect not always identifiable. The simplest unidentifiable case of counterfactual effect is due to the unidentifiability of $P(Y = y, Y_s = y')$. Graphically, there exists "w-graph" in the causal graph (see Figure 8(c)).

The analysis of the identifiability of counterfactual effect $P(y_s|s', \mathbf{o})$ concerns the connection between two causal models, $\mathcal{M}$ and $\mathcal{M}_s$; thus, Shpitser *et al.*[15] proposed a *make-cg* algorithm to construct a counterfactual graph $\mathcal{G}'$ which depicts the independence relationship among all variables in $\mathcal{M}$ and $\mathcal{M}_s$. Specifically, *make-cg* first combines original causal graph and post-interventional causal graph by removing the same exogenous variables, and the duplicated endogenous variables that are not influenced by $do(s)$. The resultant graph is the so-called counterfactual graph, which can be considered as a typical causal graph for a larger causal model. For example, Figure 8(a) shows an example causal graph, while its counterfactual graph of the counterfactual effect $P(y_s|s')$ is shown in Figure 8(b). All the graphical criteria mentioned above for the identifiability of causal effects are applicable to the counterfactual graph and the c-component factorization of the counterfactual graph for performing the counterfactual inference[106]. Shpitser *et al.*[15] further developed *ID** and *IDC** algorithms to distinguish the identifiability of counterfactual effects and compute the counterfactual quantity. In addition, Pearl[10] further proved the results about the identifiability of counterfactual effects: if all the structural functions $\mathbf{F}$ are linear, any counterfactual quantity is identifiable whenever we have full knowledge about the causal model. Unfortunately, there is no single necessary and sufficient criterion for the counterfactual effects' identifiable issues in semi-Markovian models, even the linear causal model[107].

For the identifiability of path-specific effect $PSE_\pi(y)$, it depends on whether $P(y_{s^+|\pi,s^-|\bar{\pi}})$ is identifiable or not. Unfortunately, $P(y_{s^+|\pi,s^-|\bar{\pi}})$ is not always identifiable, even in Markovian models. Avin *et al.*[107] gave the

necessary and sufficient criterion for the identifiability of $P(y_{s^+|\pi,s^-|\bar{\pi}})$ in Markovian models called recanting witness criterion. The recanting witness criterion is satisfied when there is a variable W along the causal path π connected to Y through another causal path not in π. Consider an example causal model whose causal graph is shown in Figure 8(d), when the causal path that we follow with interest $\pi = \{S \rightarrow W \rightarrow A \rightarrow Y\}$ with W as witness, then the recanting witness criterion is satisfied. The corresponding graph structure is called "kite" graph. When this criterion is satisfied, $P(y_{s^+|\pi,s^-|\bar{\pi}})$ is not identifiable and $PSE_{\pi}(y)$ is not identifiable also. Shpitser *et al.*[102] generalized this criterion to semi-Markovian models known as recanting district criterion. Specifically, if there exists district D that represents the set of variables not belonging to the set of sensitive attributes S, but ancestral of decision Y via a directed path which does not intersect S, and nodes z_i, $z_j \in D$ (possibly $z_i = z_j$), such that there is a causal path $S \rightarrow Z_i \rightarrow \ldots \rightarrow Y$ in π and a causal path $S \rightarrow Z_j \rightarrow \ldots \rightarrow Y$ not in π, then the path-specific effect of S on Y is not identifiable.

7.2.2. Efforts for dealing with identifiable issues

In Section 7.2.1, we show that the causal effects are not always identifiable only from observational data and causal graphs. Several causality-based fairness methods have been proposed from different perspectives to deal with identifiable issues.

Most previous approaches tend to make simplified or even unrealistic assumptions to avoid unidentifiable situations. For example, to avoid the unidentifiable issue of the counterfactual effect, Kusner *et al.*[11] adopted three different assumptions: (i) only using non-descendants of the sensitive attributes to build the classifier; (ii) postulating and inferring the non-deterministic sub-situations of the hidden variables based on domain knowledge; and (iii) postulating the complete causal model, treating it as the additive noise model, and then estimating the errors. Zhang *et al.*[26] evaded the unidentifiable issue of path-specific effect caused by satisfying recanting witness criterion via changing the causal model, i.e., cutting off all causal paths from sensitive variables to the decision that pass through the redline variables. However, such simplified assumptions modify the causal model equivalent to "redefining success". Although these methods made simplified assumptions to avoid identifiable issues, such assumptions may severely damage the performance of the decision model and impose uncertainty on these methods. Besides, such simplified assumptions may modify the causal model equivalent to "redefining success", while any kind of repair is not expected within a modified model, which results in fair inferences in the real world.

Recently, some workarounds for dealing with unidentifiable situations aim to stay within the true causal model, but they obtain the true unidentifiable causal effects by developing the upper and lower bounds of the causal effects. For example, Wu *et al.*[57] mathematically developed the upper and lower bounds of counterfactual fairness in unidentifiable situations and used a post-processing method for reconstructing trained classifiers to make counterfactual fairness. Zhang *et al.*[27] mathematically bound indirect discrimination as the path-specific effect in unidentifiable cases and proposed a pre-processing method for reconstructing the observational data to remove the discrimination from the original dataset. Hu *et al.*[108] adopted implicit generative models and adversarial learning to estimate the upper and lower bound of average causal effect under unidentifiable cases.

One of the major reasons causal effects are not identifiable is the presence of hidden confounding. Most previous works[12,27,44,57] adopt the no hidden confounders assumption (i.e., Markovian model) to facilitate the assessment of the causal effects. However, in practical scenarios, the existence of hidden confounders is an inescapable fact, since measuring all possible confounders is impossible. For example, in many cases, we cannot measure variables such as personal preferences, most genetic factors, and environmental factors. In these cases, to deal with hidden confounders and identifiable issues, many studies adopt the potential outcome framework[33,34] and are devoted to finding so-called "proxy variables" that reflect the information of hidden confounders. For example, we cannot directly measure the socioeconomic status of patients, but patients' de-

mographic attributes, such as their zip code, consumption ability, or employment status, can be the proxies for socioeconomic status. Variational autoencoder has been widely used to learn causal models with hidden confounders, especially for approximately inferring the complex relation between the observational variables and hidden confounders[109]. It is a computationally efficient algorithm for learning the joint distribution of the hidden confounders and the observed ones from observational data. An alternative way to eliminate the confounding bias in causal inference is to utilize the underlying network information that is attached to observational data (e.g., social networks) to infer the hidden confounders. For example, Guo *et al.*[110] proposed the *network deconfounder* to infer the influence of hidden confounders by mapping the features of observational data and auxiliary network information into the hidden space. Guo *et al.*[111] leveraged the network information to recognize the representation of hidden confounders. Veitch *et al.*[112] remarked that merely partial information that hidden confounders contain affects both the treatment and the outcome. That is to say, only a portion of confounders is actually used by the estimator to estimate the causal effects. Therefore, if a good predictive model for the treatment can be built, then one may only need to plug the outputs into a causal effect estimate directly, without any need to learn all the true confounders. Since experimental data do not suffer from hidden confounders, another method is to combine experimental and observational data together. For example, Kallus *et al.*[113] used limited experimental data to correct the hidden confounders in causal effect models trained on larger observational data, even if the observational data do not fully overlap with the experimental ones, which makes strictly weaker assumptions than existing approaches.

Overall, these potential outcome framework-based methods mostly rely on proxy variables. Before selecting proxy variables for hidden confounders, we need a thorough understanding of what a hidden confounder is supposed to represent, and whether there is any proxy variable that actually represents it. However, a sufficiently clear understanding may be impossible to attain in some cases.

7.3. Comprehensive definition of fairness

The sources of unfairness in machine learning algorithms are diverse and complex, and different biases have different degrees of impact on unfairness. Since most fairness notions, including causality-based fairness notions, quantify fairness in a single dimension, when comparing the capabilities of different fairness machine learning algorithms, using different fairness measures will often lead to different results. This means that, whether the algorithm is fair or not is relative, which depends not only on the model and data but also on the task requirements. There is a lack of complete and multi-dimensional causality-based fairness definition and evaluation system for fairness, and it is not possible to effectively quantify the fairness risk faced by machine learning algorithms. Therefore, we need to further explore comprehensive causal-based fairness notions and establish a comprehensive multi-dimensional evaluation system for the fairness of machine learning algorithms. In addition, the definition of fairness needs to be combined with the laws and the concept of social fairness of various countries to avoid narrow technical solutions. The proposition of PC-fairness and causality-based fairness notion defined from both macro-level and individual-level[114] are useful explorations to solve this problem.

7.4. Achieving fairness in a dynamic environment

The existing works mainly focus on studying the fairness in machine learning in static, no feedback, short-term impact scenarios, without examining how these decisions affect fairness in future applications over time and failing to effectively adapt to evolutionary cycles. At present, the research on the fairness of machine learning shows a trend of dynamic evolution, which requires the definition of fairness and algorithms to consider the dynamic, feedback, and long-term consequences of decision-making systems. This is particularly evident in recommendation systems, loans, hiring, etc. Fortunately, some researchers are modeling the long-term dynamics of fairness in these areas[115–119]. D'Amour *et al.*[120] regarded dynamic long-term fair learning as a Markov decision process (MDP) and proposed simulation studies to model fairness-enhancing learning in a dynamic environment. They emphasized the importance of interaction between the decision system and the environ-

ment. A complementary work[121] shows the importance of causal modeling in dynamic systems. However, due to the complexity of the real-world environment, it is impossible to model the real environment at a high level. Besides, current studies are carried out on low-dimensional data. Therefore, how to highlight important dynamics in simulations and effectively use collected data to ensure an appropriate balance between results and real-world applicability and how to adapt to high-dimensional data are current challenges. In addition, future causality-based fairness-enhancing studies can be combined with dynamic game theory for improving fairness in the confrontation environment and research the detection mechanism of dynamic fairness.

7.5. Other challenges

AI has become more and more mature after rapid development. Although most of the research on AI thus far has focused on weak AI, the design of strong AI (or human-level AI) will be increasingly vital and receive more and more attention in the near future. Weak AI only focuses on solving the given tasks input into the program, while strong AI or human-level AI (HLAI) means that its ability of thinking and action is comparable to that of a human. Therefore, developing HLAI will face more challenges. Saghiri *et al.*[122] comprehensively summarized the challenges of designing HLAI. As they said, unfairness issues are closely related to other challenges in AI. There is still a gap between solving the unfairness problem in AI alone and building a trustworthy AI. Next, this review discusses the relationship between fairness and the other challenges in AI.

Fairness and robustness. The robustness of the AI model is manifested in its outer generalization ability, that is, when the input data change abnormally, the performance of the AI model remains stable. AI models with poor robustness are prone to crash and, thus, fail to achieve fairness. In addition, the attacker may obtain private information about the training data and even the training data themselves, although the attacker has no illegal access to the data. However, the research on robustness is still in its infancy, and the theory and notions of robustness are still lacking currently.

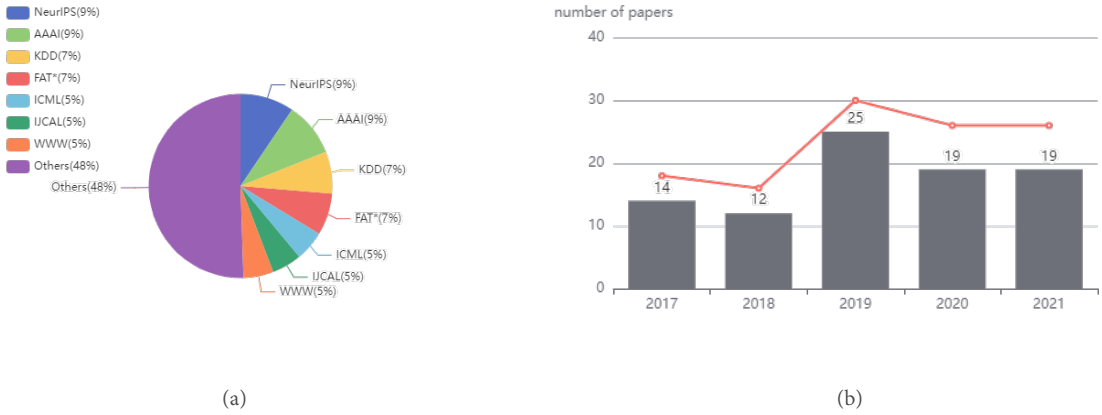
Fairness and interpretability. The explainability of discrimination is very important to improving users' understanding and trust in AI models, which is even required by law in many fields. On the other hand, interpretability can explain and judge whether the fairness of AI models is satisfied or not, which assists in improving the fairness of AI models. In some important areas, e.g., healthcare, this challenge becomes more serious because it requires that any type of decision-making must be fair and interpretable.

Causality-based methods are promising solutions to these challenges, as they can not only reveal the mechanisms by which data are generated but also enable a better understanding of the causes of discrimination. Of course, there are far more challenges faced by HLAI than those above, and more about the challenges of designing HLAI can be found in Saghiri *et al.*'s work[122].

7.6. Future trends

More realistic application scenarios. Most of the early studies are carried out under some strong assumptions, e.g., the assumption that there are no hidden confounders between observational variables. However, these assumptions are difficult to satisfy in practical applications, which leads to erroneous evaluation. Therefore, the trained model cannot guarantee that it satisfies the fairness requirement. The current studies tend to relax these assumptions and address the unfairness issue of algorithms in more general scenarios.

Privacy protection. Due to legal requirements, sensitive attributes are often inaccessible in real applications. Fairness constraints require predictors to be in some way independent of the attributes of group members. Privacy preservation raises the same question: Is it possible to guarantee that even the strongest adversary cannot steal an individual's private information through inference attacks? Causal modeling of the problem not only is helpful to solve the fairness issue but also enables stronger privacy-preserving than statistics-based methods[123]. Combining the existing fairness mechanism with differential privacy is a promising research



(a) (b)

Figure 9. Statistical charts of references in this survey: (a) a pie chart that shows the proportion of publications in journals or conferences; and (b) a bar chart about the number of publications per year (from 2017 to 2021).

direction in the near future.

Build a more complete ecosystem. There is an interaction of fairness between applications in the real world. For example, in bank loans, there exists discrimination in loan quotas for groups of different genders, and this unfairness may be caused by the salary level of groups of different genders in the workplace. Therefore, we need to further explore achieving cross-domain, cross-institution collaborative fairness algorithms.

8. CONCLUSION

This review presents the relevant background, typical causality-based fairness notions, an inclusive overview of causality-based fairness methods, and their applications. The challenges of applying causality-based fairness notions in practical scenarios and future research trends on solving fairness problems in algorithms are also discussed.

Papers related to the topic of addressing fairness issues based on causality are mostly reviewed in this survey. The statistics and analysis of these papers are also carried out in this survey and the results are presented in Figure 9. Figure 9(a) reports the proportion of papers published in reputable journals or conferences, while Figure 9(b) shows the number of publications from 2017 to 2021. The research community has only recently focused on defining, measuring, and mitigating discrimination in algorithms from a causal perspective, gradually realizing the importance of causal modeling of the problems to address fairness issues.

Therefore, we provide a relatively complete review of causality-based fairness-enhancing techniques to help researchers gain a deep understanding of this field, and we hope that more researchers will engage in this young but important field. On the one hand, discrimination detection and elimination from the causal perspective rather than statistics-based methods is more welcomed and trusted by the users of automated decision making systems, since causality-based fairness-enhancing methods consider how the data are generated and thus deeply understand the sources of discrimination. On the other hand, because of the completeness of the causal theory, it provides mathematical tools to discover discrimination when the dataset includes bias due to missing data. In addition, the main objective of this survey is to bridge the gap between the practical scenarios of discrimination elimination from the causal perspective and the ongoing theory problem. This is achieved by summing up causality-based fairness notions, approaches, and their limitations. Although the causal graph cannot be constructed without some untestable assumptions, it can still be used productively as well as serve

as an auxiliary tool to incorporate scientific domain knowledge. In addition, causal graphs can exchange the causal statements that are under plausible assumptions but lack grounding in established scientific knowledge for inferring plausible conclusions. To conclude, causality-based fairness-enhancing approaches are promising solutions to reduce discrimination despite having challenges to overcome.

DECLARATIONS

Authors' contributions

Project administration: Yu G, Yan Z
Writing-original draft: Su C, Yu G, Wang J
Writing-review and editing: Yu G, Yan Z, Cui L

Availability of data and materials

Not applicable.

Financial support and sponsorship

None.

Conflicts of interest

All authors declared that they have no conflicts of interest to this work.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Cohen L, Lipton ZC, Mansour Y. Efficient candidate screening under multiple tests and implications for fairness. arXiv preprint arXiv:190511361 2019. DOI
2. Schumann C, Foster J, Mattei N, Dickerson J. We need fairness and explainability in algorithmic hiring. In: International Conference on Autonomous Agents and Multi-Agent Systems; 2020. pp. 1716–20. DOI
3. Mukerjee A, Biswas R, Deb K, Mathur AP. Multi–objective evolutionary algorithms for the risk–return trade–off in bank loan management. *Int Trans Operational Res* 2002;9:583–97. DOI
4. Lee MSA, Floridi L. Algorithmic fairness in mortgage lending: from absolute conditions to relational trade-offs. *Minds and Machines* 2021;31:165–91. DOI
5. Baker RS, Hawn A. Algorithmic bias in education. *Int J Artif Intell Educ* 2021:1–41. DOI
6. Berk R, Heidari H, Jabbari S, Kearns M, Roth A. Fairness in criminal justice risk assessments: The state of the art. *Sociological Methods & Research* 2021;50:3–44. DOI
7. Chouldechova A. Fair prediction with disparate impact: A study of bias in recidivism prediction instruments. *Big Data* 2017;5:153–63. 10.1089/big.2016.0047] DOI
8. Dwork C, Hardt M, Pitassi T, Reingold O, Zemel R. Fairness through awareness. In: Proceedings of the Innovations in Theoretical Computer Science Conference; 2012. pp. 214–26. DOI
9. Hardt M, Price E, Srebro N. Equality of opportunity in supervised learning. In: Advances in Neural Information Processing Systems; 2016. pp. 3315–23. DOI
10. Pearl J. Causality: models, reasoning and inference. New York, NY, USA: Cambridge University Press; 2009.
11. Kusner MJ, Loftus J, Russell C, Silva R. Counterfactual fairness. In: Advances in Neural Information Processing Systems; 2017. pp. 4069–79. DOI
12. Russell C, Kusner MJ, Loftus J, Silva R. When worlds collide: integrating different counterfactual assumptions in fairness. In: Advances in Neural Information Processing Systems; 2017. pp. 6414–23. DOI

13. Pan W, Cui S, Bian J, Zhang C, Wang F. Explaining algorithmic fairness through fairness-aware causal path decomposition. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2021. pp. 1287–97. DOI
14. Grabowicz PA, Perello N, Mishra A. Marrying fairness and explainability in supervised learning. In: ACM Conference on Fairness, Accountability, and Transparency; 2022. pp. 1905–16. DOI
15. Shpitser I, Pearl J. Complete identification methods for the causal hierarchy. *J Mach Learn Res* 2008;9:1941–79. DOI
16. Caton S, Haas C. Fairness in machine learning: A survey. arXiv preprint arXiv:201004053 2020. Available from: https://arxiv.org/abs/2010.04053.
17. Du M, Yang F, Zou N, Hu X. Fairness in deep learning: a computational perspective. IEEE Intell Syst 2020;36:25–34. DOI
18. Mehrabi N, Morstatter F, Saxena N, Lerman K, Galstyan A. A survey on bias and fairness in machine learning. *ACM Comput Surv* 2021;54:1–35. DOI
19. Pessach D, Shmueli E. A Review on Fairness in Machine Learning. *ACM Comput Surv* 2022;55:1–44. DOI
20. Wan M, Zha D, Liu N, Zou N. Modeling techniques for machine learning fairness: a survey. arXiv preprint arXiv:211103015 2021. Available from: https://arxiv.org/abs/2111.03015.
21. Makhlouf K, Zhioua S, Palamidessi C. On the applicability of machine learning fairness notions. *ACM SIGKDD Explorations Newsletter* 2021;23:14–23. DOI
22. Makhlouf K, Zhioua S, Palamidessi C. Survey on causal-based machine learning fairness notions. arXiv preprint arXiv:201009553 2020. Available from: https://arxiv.org/abs/2010.09553.
23. Wu D, Liu J. Involve Humans in Algorithmic Fairness Issue: A Systematic Review. In: International Conference on Information; 2022. pp. 161–76. DOI
24. Zhang J, Bareinboim E. Fairness in decision-making—the causal explanation formula. In: AAAI Conference on Artificial Intelligence. vol. 32; 2018. pp. 2037–45. DOI
25. Pearl J, Mackenzie D. The book of why: the new science of cause and effect. Basic books; 2018. DOI
26. Zhang L, Wu Y, Wu X. A causal framework for discovering and removing direct and indirect discrimination. In: International Joint Conference on Artificial Intelligence; 2017. pp. 3929–35. DOI
27. Zhang L, Wu Y, Wu X. Causal modeling-based discrimination discovery and removal: Criteria, bounds, and algorithms. *IEEE Trans Knowl Data Eng* 2018;31:2035–50. DOI
28. Kilbertus N, Rojas-Carulla M, Parascandolo G, et al. Avoiding discrimination through causal reasoning. In: Advances in Neural Information Processing Systems; 2017. pp. 656–66. DOI
29. Zhang L, Wu Y, Wu X. Situation testing-based discrimination discovery: a causal inference approach. In: International Joint Conference on Artificial Intelligence; 2016. pp. 2718–24. DOI
30. Huan W, Wu Y, Zhang L, Wu X. Fairness through equality of effort. In: The Web Conference; 2020. pp. 743–51. DOI
31. Wu Y, Zhang L, Wu X, Tong H. Pc-fairness: a unified framework for measuring causality-based fairness. Advances in Neural Information Processing Systems 2019.
32. Khademi A, Lee S, Foley D, Honavar V. Fairness in algorithmic decision making: an excursion through the lens of causality. In: The Web Conference; 2019. pp. 2907–14. DOI
33. Rubin DB. Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology* 1974;66:688. DOI
34. Splawa-Neyman J, Dabrowska DM, Speed T. On the application of probability theory to agricultural experiments. Essay on principles. Section 9. *Statist Sci* 1990:465–72. DOI
35. Bendick M. Situation testing for employment discrimination in the United States of America. *Horizons stratégiques* 2007;5:17–39. DOI
36. Luong BT, Ruggieri S, Turini F. k-NN as an implementation of situation testing for discrimination discovery and prevention. In: ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2011. pp. 502–10. DOI
37. Imbens GW, Rubin DB. Causal inference in statistics, social, and biomedical sciences. Cambridge University Press; 2015.
38. Zemel R, Wu Y, Swersky K, Pitassi T, Dwork C. Learning fair representations. In: International Conference on Machine Learning; 2013. pp. 325–33. DOI
39. Feldman M, Friedler SA, Moeller J, Scheidegger C, Venkatasubramanian S. Certifying and removing disparate impact. In: proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2015. pp. 259–68. DOI
40. Xu D, Wu Y, Yuan S, Zhang L, Wu X. Achieving causal fairness through generative adversarial networks. In: International Joint Conference on Artificial Intelligence; 2019. pp. 1452–58. DOI
41. Kocaoglu M, Snyder C, Dimakis AG, Vishwanath S. CausalGAN: Learning causal implicit generative models with adversarial training. In: International Conference on Learning Representations; 2018. DOI
42. Salimi B, Howe B, Suciu D. Data management for causal algorithmic fairness. *IEEE Data Eng Bull* 2019:24-35. Available from: http://sites.computer.org/debull/A19sept/p24.pdf.
43. Salimi B, Rodriguez L, Howe B, Suciu D. Interventional fairness: Causal database repair for algorithmic fairness. In: International Conference on Management of Data; 2019. pp. 793–810. DOI
44. Nabi R, Shpitser I. Fair inference on outcomes. In: AAAI Conference on Artificial Intelligence; 2018. DOI
45. Chiappa S. Path-specific counterfactual fairness. In: AAAI Conference on Artificial Intelligence; 2019. pp. 7801–8. DOI
46. Agarwal A, Beygelzimer A, Dudík M, Langford J, Wallach H. A reductions approach to fair classification. In: International Conference on Machine Learning; 2018. pp. 60–69. [DOI: http://proceedings.mlr.press/v80/agarwal18a.html]
47. Bechavod Y, Ligett K. Learning fair classifiers: a regularization-inspired approach. arXiv preprint arXiv:170700044 2017. Available

from: http://arxiv.org/abs/1707.00044.
48. Kamishima T, Akaho S, Asoh H, Sakuma J. Fairness-aware classifier with prejudice remover regularizer. In: Joint European Conference on Machine Learning and Knowledge Discovery in Databases; 2012. pp. 35–50. DOI
49. Zafar MB, Valera I, Rogriguez MG, Gummadi KP. Fairness constraints: mechanisms for fair classification. In: Artificial Intelligence and Statistics; 2017. pp. 962–70. Available from: http://proceedings.mlr.press/v54/zafar17a.html.
50. Zafar MB, Valera I, Gomez Rodriguez M, Gummadi KP. Fairness beyond disparate treatment and disparate impact: learning classification without disparate mistreatment. In: The Web Conference; 2017. pp. 1171–80. DOI
51. Hu Y, Wu Y, Zhang L, Wu X. Fair multiple decision making through soft interventions. *Adv Neu Inf Pro Syst* 2020;33:17965–75. DOI
52. Garg S, Perot V, Limtiaco N, et al. Counterfactual fairness in text classification through robustness. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society; 2019. pp. 219–26. DOI
53. Di Stefano PG, Hickey JM, Vasileiou V. Counterfactual fairness: removing direct effects through regularization. arXiv preprint arXiv:200210774 2020. Available from: https://arxiv.org/abs/2002.10774.
54. Kim H, Shin S, Jang J, et al. Counterfactual fairness with disentangled causal effect variational autoencoder. In: AAAI Conference on Artificial Intelligence; 2021. pp. 8128–36. Available from: https://ojs.aaai.org/index.php/AAAI/article/view/16990.
55. Corbett-Davies S, Pierson E, Feller A, Goel S, Huq A. Algorithmic decision making and the cost of fairness. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2017. pp. 797–806. DOI
56. Dwork C, Immorlica N, Kalai AT, Leiserson M. Decoupled classifiers for group-fair and efficient machine learning. In: International Conference on Fairness, Accountability and Transparency; 2018. pp. 119–33. http://proceedings.mlr.press/v81/dwork18a.html
57. Wu Y, Zhang L, Wu X. Counterfactual fairness: unidentification, bound and algorithm. In: International Joint Conference on Artificial Intelligence; 2019. pp. 1438–44. DOI
58. Kusner M, Russell C, Loftus J, Silva R. Making decisions that reduce discriminatory impacts. In: International Conference on Machine Learning; 2019. pp. 3591–600. Available from: http://proceedings.mlr.press/v97/kusner19a/kusner19a.pdf.
59. Mishler A, Kennedy EH, Chouldechova A. Fairness in risk assessment instruments: post-processing to achieve counterfactual equalized odds. In: ACM Conference on Fairness, Accountability, and Transparency; 2021. pp. 386–400. DOI
60. Woodworth B, Gunasekar S, Ohannessian MI, Srebro N. Learning non-discriminatory predictors. In: Conference on Learning Theory; 2017. pp. 1920–53. Available from: http://proceedings.mlr.press/v65/woodworth17a.html.
61. Calders T, Verwer S. Three naive Bayes approaches for discrimination-free classification. Data Min Knowl Discov 2010;21:277–92. DOI
62. Friedler SA, Scheidegger C, Venkatasubramanian S, et al. A comparative study of fairness-enhancing interventions in machine learning. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019. pp. 329–38. DOI
63. Martínez-Plumed F, Ferri C, Nieves D, Hernández-Orallo J. Fairness and missing values. arXiv preprint arXiv:190512728 2019. Available from: http://arxiv.org/abs/1905.12728.
64. Bareinboim E, Pearl J. Causal inference and the data-fusion problem. *Proc Natl Acad Sci* 2016;113:7345–52. DOI
65. Spirtes P, Meek C, Richardson T. Causal inference in the presence of latent variables and selection bias. In: Conference on Uncertainty in Artificial Intelligence; 1995. pp. 499–506. DOI
66. Goel N, Amayuelas A, Deshpande A, Sharma A. The importance of modeling data missingness in algorithmic fairness: a causal perspective. In: AAAI Conference on Artificial Intelligence. vol. 35; 2021. pp. 7564–73. Available from: https://ojs.aaai.org/index.php/AAAI/article/view/16926.
67. Burke R. Multisided fairness for recommendation. arXiv preprint arXiv:170700093 2017. Available from: http://arxiv.org/abs/1707.00093.
68. Wu Y, Zhang L, Wu X. On discrimination discovery and removal in ranked data using causal graph. In: Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 2018. pp. 2536–44. DOI
69. Zhao Z, Chen J, Zhou S, et al. Popularity Bias Is Not Always Evil: Disentangling Benign and Harmful Bias for Recommendation. arXiv preprint arXiv:210907946 2021. Available from: https://arxiv.org/abs/2109.07946.
70. Zheng Y, Gao C, Li X, et al. Disentangling user interest and conformity for recommendation with causal embedding. In: The Web Conference; 2021. pp. 2980–91. DOI
71. Zhang Y, Feng F, He X, et al. Causal intervention for leveraging popularity bias in recommendation. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval; 2021. pp. 11–20. DOI
72. Wang W, Feng F, He X, Zhang H, Chua TS. Clicks can be cheating: counterfactual recommendation for mitigating clickbait issue. In: Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval; 2021. pp. 1288–97. DOI
73. Li Y, Chen H, Xu S, Ge Y, Zhang Y. Towards personalized fairness based on causal notion. In: International ACM SIGIR Conference on Research and Development in Information Retrieval; 2021. pp. 1054–63. DOI
74. Huang W, Zhang L, Wu X. Achieving counterfactual fairness for causal bandit. In: AAAI Conference on Artificial Intelligence; 2022. pp. 6952–59. DOI
75. Zhao J, Wang T, Yatskar M, Ordonez V, Chang KW. Men also like shopping: reducing gender bias amplification using corpus-level constraints. In: Conference on Empirical Methods in Natural Language Processing; 2017. pp. 2979–89. DOI
76. Stanovsky G, Smith NA, Zettlemoyer L. Evaluating gender bias in machine translation. In: Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics; 2019. pp. 1679–84. DOI
77. Huang PS, Zhang H, Jiang R,et al. Reducing sentiment bias in language models via counterfactual evaluation. arXiv preprint

arXiv:191103064 2019. DOI
78. Shin S, Song K, Jang J, et al. Neutralizing gender bias in word embeddings with latent disentanglement and counterfactual generation. In: Empirical Methods in Natural Language Processing Conference; 2020. pp. 3126–40. DOI
79. Yang Z, Feng J. A causal inference method for reducing gender bias in word embedding relations. In: AAAI Conference on Artificial Intelligence; 2020. pp. 9434–41. DOI
80. Lu K, Mardziel P, Wu F, Amancharla P, Datta A. Gender bias in neural natural language processing. In: Logic, Language, and Security; 2020. pp. 189–202. DOI
81. Chen IY, Szolovits P, Ghassemi M. Can AI help reduce disparities in general medical and mental health care? *AMA J Ethics* 2019;21:167–79. DOI
82. Zink A, Rose S. Fair regression for health care spending. *Biometrics* 2020;76:973–82. DOI
83. Pfohl SR, Duan T, Ding DY, Shah NH. Counterfactual reasoning for fair clinical risk prediction. In: Machine Learning for Healthcare Conference; 2019. pp. 325–58. Available from: http://proceedings.mlr.press/v106/pfohl19a.html.
84. Pfohl SR, Foryciarz A, Shah NH. An empirical characterization of fair machine learning for clinical risk prediction. *J Biomed Inform* 2021;113:103621. DOI
85. Ramsey JD, Zhang K, Glymour M, et al. TETRAD—A toolbox for causal discovery. In: International Workshop on Climate Informatics; 2018. Available from: http://www.phil.cmu.edu/tetrad/.
86. Zhang K, Ramsey J, Gong M, et al. Causal-learn: Causal discovery for Python; 2022. https://github.com/cmu-phil/causal-learn.
87. Wongchokprasitti CK, Hochheiser H, Espino J,et al.. bd2kccd/py-causal v1.2.1; 2019. https://doi.org/10.5281/zenodo.3592985.
88. Runge J, Nowack P, Kretschmer M, Flaxman S, Sejdinovic D. Detecting and quantifying causal associations in large nonlinear time series datasets. *Sci Adv* 2019;5:eaau4996. DOI
89. Zhang K, Zhu S, Kalander M, et al. gCastle: a python toolbox for causal discovery. arXiv preprint arXiv:211115155 2021. Available from: https://arxiv.org/abs/2111.15155.
90. Chen H, Harinen T, Lee JY, Yung M, Zhao Z. Causalml: Python package for causal machine learning. arXiv preprint arXiv:200211631 2020. Available from: https://arxiv.org/abs/2002.11631.
91. Tingley D, Yamamoto T, Hirose K, Keele L, Imai K. mediation: R package for causal mediation analysis. *J Stat Softw* 2014;59:1–38. DOI
92. Tikka S, Karvanen J. Identifying causal effects with the R Package causaleffect. *J Stat Softw* 2017;76:1–30. DOI
93. Sharma A, Kiciman E. DoWhy: an end-to-end library for causal inference. arXiv preprint arXiv:201104216 2020. Available from: https://arxiv.org/abs/2011.04216.
94. Bellamy RK, Dey K, Hind M, et al. AI Fairness 360: An extensible toolkit for detecting and mitigating algorithmic bias. *IBM J Res Dev* 2019;63:1–15. DOI
95. Bird S, Dudík M, Edgar R, et al. Fairlearn: A toolkit for assessing and improving fairness in AI. Technical Report MSR-TR-2020-32, Microsoft 2020.
96. Geiger D, Heckerman D. Learning gaussian networks. In: Conference on Uncertainty in Artificial Intelligence; 1994. pp. 235–43. DOI
97. Janzing D, Schölkopf B. Causal inference using the algorithmic Markov condition. *IEEE Trans Inf Theory* 2010;56:5168–94. DOI
98. Kalainathan D, Goudet O, Guyon I, Lopez-Paz D, Sebag M. Structural agnostic modeling: adversarial learning of causal graphs. arXiv preprint arXiv:180304929 2018. Available from: https://doi.org/10.48550/arXiv.1803.04929.
99. Hoyer PO, Shimizu S, Kerminen AJ, Palviainen M. Estimation of causal effects using linear non-Gaussian causal models with hidden variables. *Int J Approx Reason* 2008;49:362–78. DOI
100. Huang Y, Valtorta M. Identifiability in causal Bayesian networks: a sound and complete algorithm. In: National Conference on Artificial Intelligence; 2006. pp. 1149–54. DOI
101. Tian J. Identifying linear causal effects. In: AAAI Conference on Artificial Intelligence; 2004. pp. 104–11. DOI
102. Shpitser I. Counterfactual graphical models for longitudinal mediation analysis with unobserved confounding. *Cogn Sci* 2013;37:1011–35. DOI
103. Malinsky D, Shpitser I, Richardson T. A potential outcomes calculus for identifying conditional path-specific effects. In: International Conference on Artificial Intelligence and Statistics; 2019. pp. 3080–88. Available from: http://proceedings.mlr.press/v89/malinsky19b.html. [PMID: 31886462]
104. Shpitser I, Pearl J. Identification of conditional interventional distributions. In: Conference on Uncertainty in Artificial Intelligence; 2006. pp. 437–44. DOI
105. Tian J, Pearl J. A general identification condition for causal effects. eScholarship, University of California; 2002.
106. Shpitser I, Pearl J. What counterfactuals can be tested. In: Conference on Uncertainty in Artificial Intelligence; 2007. pp. 352–59. DOI
107. Avin C, Shpitser I, Pearl J. Identifiability of path-specific effects. In: International Joint Conference on Artificial Intelligence; 2005. pp. 357–63. DOI
108. Hu Y, Wu Y, Zhang L, Wu X. A generative adversarial framework for bounding confounded causal effects. In: AAAI Conference on Artificial Intelligence; 2021. p. 12104–12112. Available from: https://ojs.aaai.org/index.php/AAAI/article/view/17437.
109. Louizos C, Shalit U, Mooij JM, et al. Causal effect inference with deep latent-variable models; 2017. pp. 6446–56. DOI
110. Guo R, Li J, Liu H. Learning individual causal effects from networked observational data. In: International Conference on Web Search and Data Mining; 2020. pp. 232–40. DOI
111. Guo R, Li J, Liu H. Counterfactual evaluation of treatment assignment functions with networked observational data. In: Proceedings of the SIAM International Conference on Data Mining; 2020. pp. 271–79. DOI

112. Veitch V, Wang Y, Blei D. Using embeddings to correct for unobserved confounding in networks; 2019. pp. 13769–79. DOI
113. Kallus N, Puli AM, Shalit U. Removing hidden confounding by experimental grounding; 2018. pp. 10888–97. DOI
114. Mhasawade V, Chunara R. Causal multi-level fairness. In: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society; 2021. pp. 784–94. DOI
115. Lum K, Isaac W. To predict and serve? *Significance* 2016;13:14–19. DOI
116. Hu L, Chen Y. A short-term intervention for long-term fairness in the labor market. In: The Web Conference; 2018. pp. 1389–98. DOI
117. Mouzannar H, Ohannessian MI, Srebro N. From fair decision making to social equality. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019. pp. 359–68. DOI
118. Bountouridis D, Harambam J, Makhortykh M, et al. Siren: A simulation framework for understanding the effects of recommender systems in online news environments. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019. pp. 150–59. DOI
119. Kannan S, Roth A, Ziani J. Downstream effects of affirmative action. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2019. pp. 240–48. DOI
120. D'Amour A, Srinivasan H, Atwood J, et al. Fairness is not static: deeper understanding of long term fairness via simulation studies. In: Proceedings of the Conference on Fairness, Accountability, and Transparency; 2020. pp. 525–34. DOI
121. Creager E, Madras D, Pitassi T, Zemel R. Causal modeling for fairness in dynamical systems. In: International Conference on Machine Learning; 2020. pp. 2185–95. DOI
122. Saghiri AM, Vahidipour SM, Jabbarpour MR, Sookhak M, Forestiero A. A Survey of Artificial Intelligence Challenges: Analyzing the Definitions, Relationships, and Evolutions. *App Sci* 2022;12:4054. DOI
123. Tople S, Sharma A, Nori A. Alleviating privacy attacks via causal learning. In: International Conference on Machine Learning; 2020. pp. 9537–47. DOI

Zhang *et al*. *Intell Robot* 2022;2(3):275-97
DOI: 10.20517/ir.2022.20

Intelligence & Robotics

Review

Open Access

Check for updates

Deep reinforcement learning for real-world quadrupedal locomotion: a comprehensive review

Hongyin Zhang, Li He, Donglin Wang

School of Engineering, Westlake University, Hangzhou 310000, Zhejiang, China.

Correspondence to: Dr. Donglin Wang, School of Engineering, Westlake University, Dunyu Road No.600, Xihu District, Hangzhou 310000, Zhejiang, China. E-mail: wangdonglin@westlake.edu.cn; ORCID: 0000-0002-8188-3735

How to cite this article: Zhang H, He L, Wang D. Deep reinforcement learning for real-world quadrupedal locomotion: a comprehensive review. *Intell Robot* 2022;2(3):275-97. http://dx.doi.org/10.20517/ir.2022.20

Received: 30 Jun 2022 **First Decision:** 25 Jul 2022 **Revised:** 27 Jul 2022 **Accepted:** 22 Aug 2022 **Published:** 1 Sep 2022

Academic Editor: Simon X. Yang **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

Building controllers for legged robots with agility and intelligence has been one of the typical challenges in the pursuit of artificial intelligence (AI). As an important part of the AI field, deep reinforcement learning (DRL) can realize sequential decision making without physical modeling through end-to-end learning and has achieved a series of major breakthroughs in quadrupedal locomotion research. In this review article, we systematically organize and summarize relevant important literature, covering DRL algorithms from problem setting to advanced learning methods. These algorithms alleviate the specific problems encountered in the practical application of robots to a certain extent. We first elaborate on the general development trend in this field from several aspects, such as the DRL algorithms, simulation environments, and hardware platforms. Moreover, core components in the algorithm design, such as state and action spaces, reward functions, and solutions to reality gap problems, are highlighted and summarized. We further discuss open problems and propose promising future research directions to discover new areas of research.

Keywords: Deep reinforcement learning, quadrupedal locomotion, reality gap

1. INTRODUCTION

Wheeled and tracked robots are still unable to navigate the most challenging terrain in the natural environment, and their stability may be severely compromised. Quadrupedal locomotion, on the other hand, can greatly expand the agility of robot behavior, as legged robots can choose safe and stable footholds within their kinematic

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

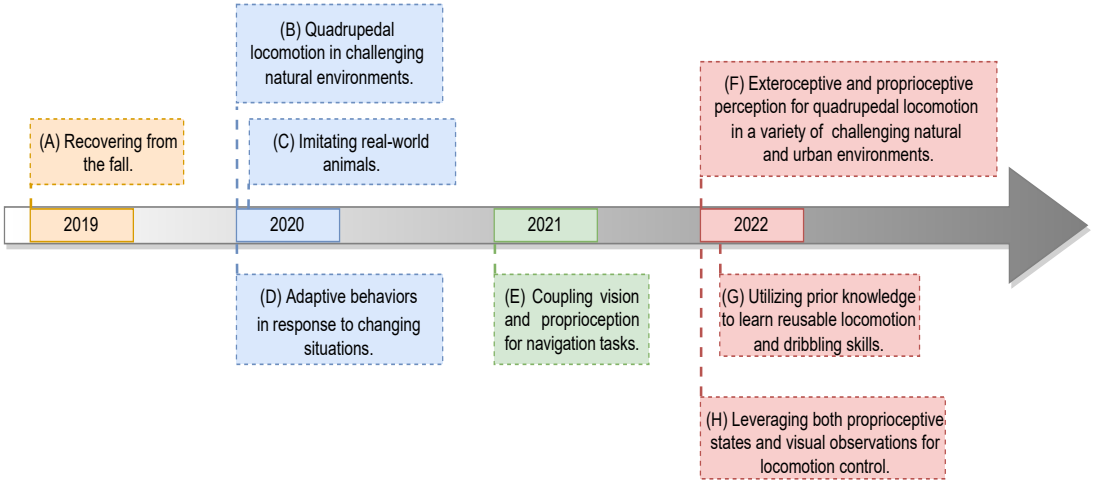


Figure 1. Several typical quadrupedal locomotion studies based on DRL algorithm: (A) recovering from a fall[6]; (B) a radically robust controller for quadrupedal locomotion in challenging natural environments[7]; (C) learning agile locomotion skills by imitating real-world animals[10]; (D) producing adaptive behaviors in response to changing situations[9]; (E) coupling vision and proprioception for navigation tasks[11]; (F) integrating exteroceptive and proprioceptive perception for quadrupedal locomotion in a variety of challenging natural and urban environments over multiple seasons[8]; (G) utilizing prior knowledge of human and animal movement to learn reusable locomotion and dribbling skills[12]; and (H) leveraging both proprioceptive states and visual observations for locomotion control[13].

reach and rapidly change the kinematic state according to the environment. To further study quadrupedal locomotion on uneven terrain, the complexity of traditional control methods is gradually increased as more scenarios are considered[1–4]. As a result, the associated development and maintenance becomes rather time-consuming and labor-intensive, and it remains vulnerable to extreme situations.

With the rapid development of the artificial intelligence field, deep reinforcement learning (DRL) has recently emerged as an alternative method for developing legged motor skills. The core idea of DRL is that the control policy learns to make decisions to obtain the maximum benefit based on the reward received from the environment[5]. DRL has been used to simplify the design of locomotion controllers, automate parts of the design process, and learn behaviors that previous control methods could not achieve[6–9]. Research on DRL algorithms for legged robots has gained wide attention in recent years. Meanwhile, several well-known research institutions and companies have publicly revealed their implementations of DRL-based legged robots, as shown in Figure 1.

Currently, there are several reviews on applying DRL algorithms to robots. Some works summarize the types of DRL algorithms and deployment on several robots such as robotic arms, bipeds, and quadrupeds[14]. They discuss in detail the theoretical background and advanced learning algorithms of DRL, as well as present key current challenges in this field and ideas for future research directions to stimulate new research interests. There is also a work summarizing some case studies involving robotic DRL and some open problems[15]. Based on these case studies, they discuss common challenges in DRL and how the work addresses them. They also provide an overview of other prominent challenges, many of which are unique to real-world robotics settings. Furthermore, a common paradigm for DRL algorithms applied to robotics is to train policies in simulations and then deploy them on real machines. This can lead to the reality gap[16] (also known as sim-to-real gap) problem, which is summarized for the robotic arm in[17]. These reviews introduce the basic background behind sim-to-real transfer in DRL and outline the main methods currently used: domain randomization, domain adaptation, imitation learning, meta-learning, and knowledge distillation. They categorize some of the most relevant recent works and outline the main application scenarios while also discussing the main opportunities and challenges of different approaches and pointing out the most promising directions. The closest work to our review simply surveys current research on motor skills learning via DRL algorithms[18],

without systematically combing through the relevant literature and without an in-depth analysis of the existing open problems and future research directions.

In this survey, we focus on quadrupedal locomotion research from the perspective of algorithm design, key challenges, and future research directions. The remainder of this review is organized as follows. Section 2 formulates the basic settings in DRL and lists several important issues that should be alleviated. The classification and core components of the current algorithm design (e.g., the DRL algorithm, simulation environment, hardware platform, observation, action, and reward function) are introduced in Section 3. Finally, we summarize and offer perspectives on potential future research directions in this field.

2. BASIC SETTINGS AND LEARNING PARADIGM

In this section, we first formulate the basic settings of standard reinforcement learning problems and then introduce the common learning paradigm.

Quadrupedal locomotion is commonly formulated as a reinforcement learning (RL) problem, which in the framework of Markov decision processes (MDPs) is specified by the tuple $M := (\mathcal{S}, \mathcal{A}, R, P, \rho_0, \gamma)$, where $\mathcal{S}$ and $\mathcal{A}$ denote the state and action spaces, respectively; $R : \mathcal{S} \times \mathcal{A} \to \mathbb{R}$ is the reward function; $P(\mathbf{s}'|\mathbf{s}, \mathbf{a})$ is the stochastic transition dynamics; $\rho_0(\mathbf{s})$ is the initial state distribution; and $\gamma \in [0, 1]$ is the discount factor. The objective is to learn a control policy π that enables a legged robot to maximize its expected return for a given task[19]. A state s_t is observed by the robot from the environment at each time step t, and an action $\mathbf{a}_t \sim \pi(\mathbf{a}_t \mid \mathbf{s}_t)$ is derived from robot's policy π. The robot next applies this action, which results in a novel state s_{t+1} and a scalar reward $r_t = R(\mathbf{s}_t, \mathbf{a}_t)$. As a result, a trajectory $\tau := (\mathbf{s}_0, \mathbf{a}_0, r_0, \mathbf{s}_1, \mathbf{a}_1, r_1, ...)$ is obtained by repeating applications of this interaction process. Formally, the RL problem requires the robot to learn a decision making policy $\pi(\mathbf{a}|\mathbf{s})$ that maximizes the expected discounted return:

$$\mathcal{J}(\pi) := \mathbb{E}_{\tau \sim p^{\pi}(\tau)} \left[\sum_{t=0}^{H-1} \gamma^t r_t \right], \tag{1}$$

where H denotes the time horizon of each episode and $p^{\pi}(\tau) = p(\mathbf{s}_0) \prod_{t=0}^{H-1} p(\mathbf{s}_{t+1} \mid \mathbf{s}_t, \mathbf{a}_t) \pi(\mathbf{a}_t \mid \mathbf{s}_t)$ represents the likelihood of a trajectory τ under a given policy π, with $p(\mathbf{s}_0)$ being the initial state distribution.

For quadrupedal locomotion tasks, most of the current research is based on a similar learning paradigm, as shown in Figure 2. First, we need to build a simulation environment (e.g., ground, steps, and stairs), and then design the state and action space, reward function and other essential elements. DRL-based algorithms are further designed and used to train policies in the simulation. The trained policy is finally deployed on the real robot to complete the assigned task.

3. DRL-BASED CONTROL POLICY DESIGN FOR QUADRUPEDAL LOCOMOTION

In this section, we detail the key components of a DRL-based controller. The classification results are presented in Tables 1 and 2 in the Appendix. After the most relevant publications in this field are summarized, their key parts are further condensed. As shown in Figure 3, we firstly review and analyze the general state and development trend of current research (e.g., DRL algorithms, simulators, and hardware platforms). Then, important components of DRL algorithm (state and action design, reward function design, solution to reality gap, *etc.*) are presented, as shown in Figure 4. These specific designs would help to alleviate open questions, which are further discussed in Section 4. Please refer to the Appendix for more details.

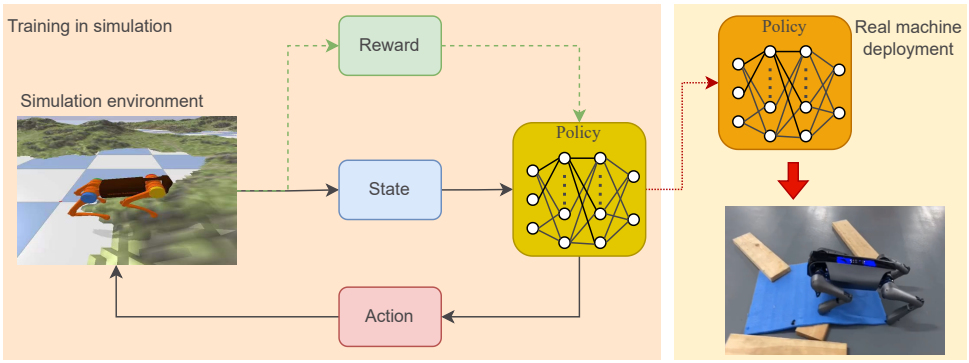


Figure 2. A common paradigm for DRL-based quadrupedal locomotion research. This paradigm is mainly divided into training and testing phases. The policy interacts with the simulated environment and collects data for iterative updates, and then the trained policy is deployed to the real robot.

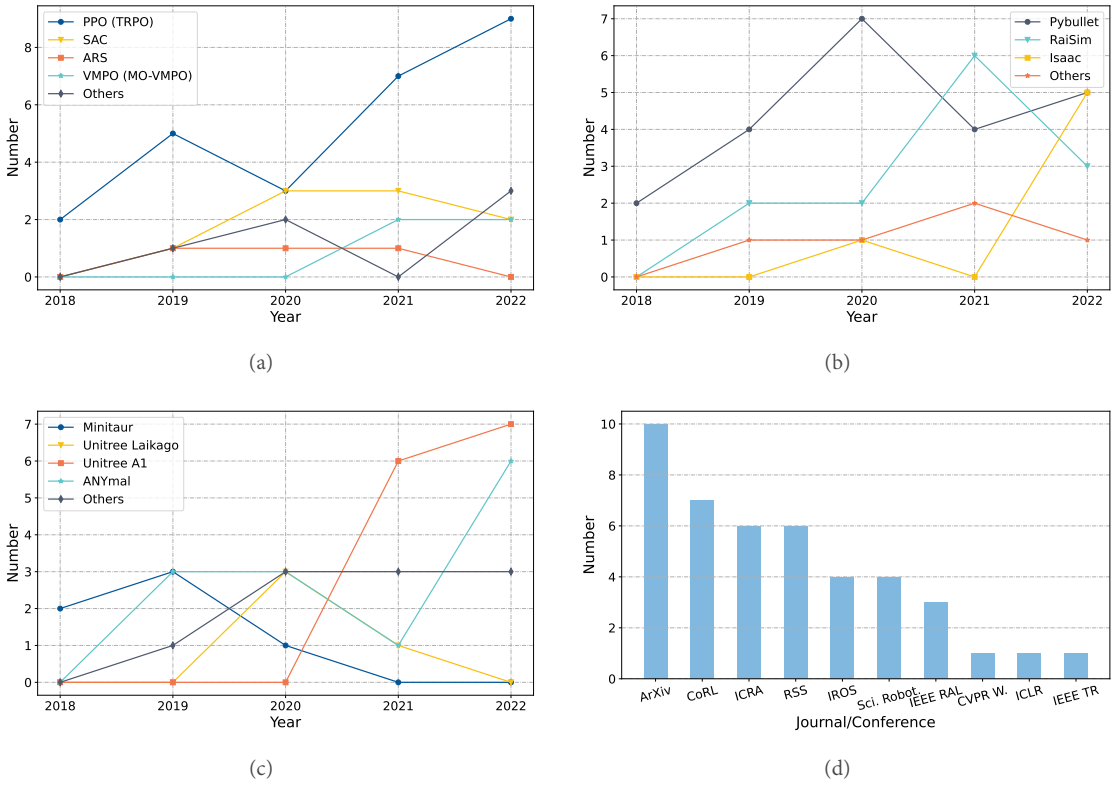


Figure 3. Several statistical results from important papers on quadrupedal locomotion research. A full summary of classification results of the most relevant publications is presented in Tables 1 and 2 in the Appendix. These papers were selected from journals and conferences (ArXiv, CoRL, ICRA, RSS, IROS, Science Robotics, ICLR, *etc.*) in recent years. (a-c) Trends in the usage times of several DRL-based algorithms, simulation platforms, and real robots. The x and y axes represent the year and the number used, respectively. (d) Number of papers accepted by the journal or conference. The x and y axes represent journals (or conferences) and the number of papers published, respectively.

3.1. DRL algorithm

Although many novel algorithms have been developed in the DRL community, most current quadrupedal locomotion controller designs still use model-free DRL algorithms, especially PPO and TRPO[20,21]. For a complex high-dimensional nonlinear system such as robots, stable control is the fundamental purpose. Most researchers choose the PPO (TRPO) algorithm for utilization in their research due to its simplicity, stability,

theoretical justification, and empirical performance[20–22].

Similar to on-policy algorithms, PPO (TRPO) has been criticized for its sample inefficiency; thus, more efficient model-free algorithms (ARS[23], SAC[24], V-MPO[25], *etc.*) are sometimes considered. Some researchers have also recently used advanced algorithms for more challenging tasks. For example, the multi-objective variant of the VMPO algorithm (MO-VMPO)[26] has been utilized to train a policy to track the planned trajectories[27]. Some researchers have introduced guided constrained policy optimization (GCPO) method for tracking base velocity commands while following defined constraints[28]. Moreover, for more efficient real-world fine-tuning and to avoid overestimation problems, REDQ, an off-policy algorithm[29], is used for real robots[30].

3.2. Simulator

The robot simulator should be able to simulate the dynamic physical laws of the robot itself more realistically and efficiently solve the collisions generated when the robot interacts with the environment. Over the past few years, the Pybullet[31] and RaiSim[32] simulation platforms have been the choice of most researchers. However, the current robotic simulators in academia are still relatively simple, and the precision is far less than that of simulators in games. For robots, directly realizing end-to-end decision making from perception to control is difficult without an accurate and realistic simulator. Common robotic simulators, such as Pybullet and RaiSim, can only solve control-level simulations, but they are stretched for real-world simulations. They have been developed to run on CPUs with reduced parallelism. On the other hand, while mujoco[33] is a popular DRL algorithm verification simulator, it is rarely used as a deployment and testing platform for real-world quadrupedal locomotion algorithms. A possible explanation is that the highly encapsulated mujoco simulator makes it difficult for researchers to develop it further.

Recently, NVIDIA released a new simulator, Isaac Gym[34], which simulates the environment with much higher accuracy than the aforementioned simulators, and can simulate and train directly on GPUs. This simulator is scalable and can simulate a large number of scenarios in parallel, so researchers can use DRL algorithms for large-scale training. It can also build large-scale realistic complex scenes, and its underlying PhysX engine can accurately and realistically model and simulate the motion of objects. Therefore, more researchers have begun to use Isaac Gym as the implementation and verification platform of DRL algorithm[35–38].

3.3. Hardware platform

In the early research stage, *Minitaur*[39] with only eight degrees of freedom was used to verify the feasibility of the DRL algorithm in simple experimental scenarios. To accomplish more complex tasks, robots (Unitree Laikago[1], Unitree A1[2], *ANYmal*[40], *etc.*) with more than 12 degrees of freedom are used by researchers. While the *ANYmal* series robots are well known for their high hardware costs, low-cost robots such as *Unitree A1* are a more prevailing choice among researchers. Lower-cost hardware platforms allow DRL algorithms to be more widely used. More recently, a wheel-legged quadruped robot[38] demonstrated skills learned from existing DRL controllers and trajectory optimization, such as ducking and walking, and new skills, such as switching between quadrupedal and humanoid configurations.

3.4. Publisher

Currently, DRL-based quadrupedal locomotion research is an emerging and promising field, and many papers have not been officially published. The published papers are mainly in journals or conferences related to the field of robotics, and there are four outstanding works[6–9] published on Science Robotics. It is worth noting that the field is actually an intersection of several fields, and some excellent studies have been published at

[1]https://www.unitree.com/products/laikago/

[2]https://www.unitree.com/products/a1/

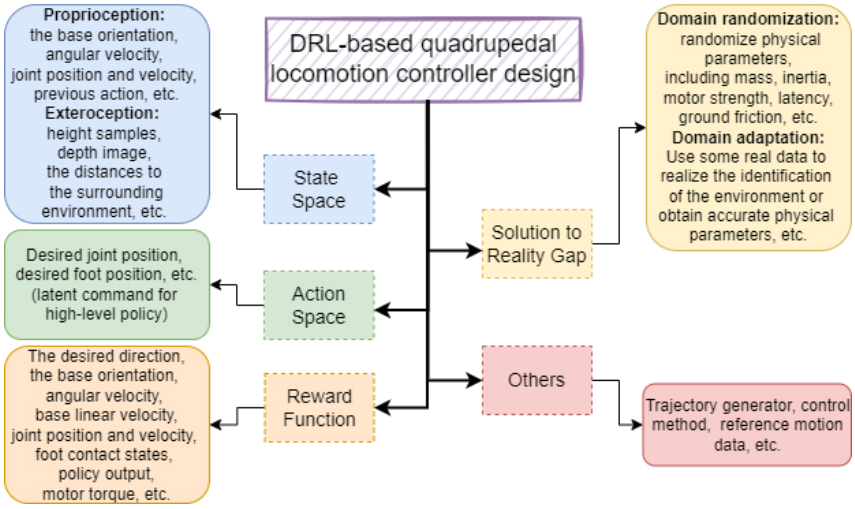


Figure 4. The key components of the DRL-based controller design from the classification result of the most relevant publications. Tables 1 and 2 in the Appendix provide a completed summary.

conferences in the machine learning field.

3.5. State, action, reward, and others

State, action, and reward are integral and important components for training controllers. The design of these components will directly affect the performance of the controller. However, there is no fully unified standard and method for the specific design.

For the design of state space, on the one hand, considering too few observations can lead to a partially observable controller. On the other hand, providing all available readings results in a brittle controller that is overfitted to the simulation environment. Both affect the performance of the controller in the real machine, so researchers can only make trade-offs based on practical problems. In current research works, for simple tasks (walking, turning on flat ground, *etc.*), proprioception alone (base orientation, angular velocity, joint position and velocity, *etc.*) is sufficient to solve the problem[10,39,41]. For more complex tasks (walking on uneven ground, climbing stairs or hillsides, avoiding obstacles, *etc.*), exteroception, such as visual information, needs to be introduced[8,13,42]. Adding additional sensors alleviates the partial observation issues to some extent.

Most researchers use the desired joint positions (residuals) as the action space and then calculate the torque through a PD controller to control the robot locomotion. Early studies[43] experimentally demonstrated that controllers with such action space can achieve better performance. However, recent studies also attempt to use lower-level control commands to obtain highly dynamic motion behavior to avoid the use of PD controllers and control torque directly[44]. Although the current DRL-based controllers have achieved outstanding performance[6–8], their stability is still not as good as the common control methods, such as MPC controllers[45]. The force–position hybrid control method adopted by MPC is worthy of reference and further research. Furthermore, in some studies based on hierarchical DRL, the latent commands serve as the action space of the high-level policy to guide the behavior of low-level policies[46,47].

In general, the design of the reward function is fairly laborious, especially for complex systems such as robots. Small changes in the reward function hyperparameters have the potential to have a large impact on the final performance of the controller. In order for the robot to complete more complex tasks, the reward function must be designed with sufficient detail[6–8,48]. Some specific factors include the desired direction, base orientation, angular velocity, base linear velocity, joint position and velocity, foot contact states, policy output, and motor

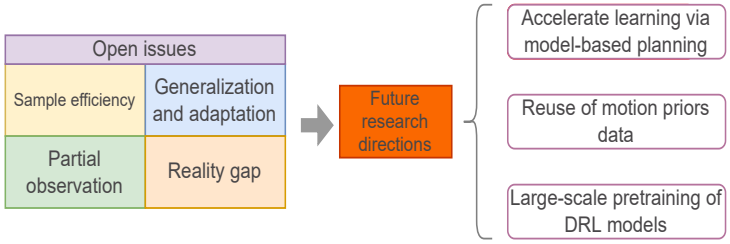


Figure 5. In the DRL-based real-world quadrupedal locomotion field, open problems mainly include sample efficiency, generalization and adaptation, partial observation, and reality gap. Future research directions are highlighted and pointed out around these open problems. Based on the current research states of quadrupedal locomotion, we expound the future research prospects from multiple perspectives. In particular, world models, skill data, and pre-trained models require significant attention, as these directions will play an integral role in realizing legged robot intelligence.

torque.

Many studies have also considered additional information, such as trajectory generators [46,49–51], control methods [52–54], motion data [10,12,55,56], *etc.* Trajectory generators and control methods mainly introduce prior knowledge in the action space, narrowing the search range of DRL control policies, which greatly improves the sample efficiency under a simple reward function. Motion data are often generated by other suboptimal controllers or assessed via public datasets. Through imitation learning based on the motion data, the robot can master behaviors and skills such as walking and turning. In both simulations and real-world deployment, the robot eventually manages to generate natural and agile movement patterns and completes the assigned tasks according to the external reward function.

3.6. Solution to reality gap

Under the current mainstream learning paradigm, the reality gap is an unavoidable problem that must be addressed. The domain randomization method is used by most researchers due to its simplicity and effectiveness. The difference between simulation and real environment is mainly reflected in physical parameters and sensors. Therefore, researchers mainly randomize physical parameters (mass, inertia, motor strength, latency, ground friction, *etc.*), add Gaussian noise to observations, and apply disturbing force, *etc.* [35,48,50,57,58]. However, domain randomization methods trade optimality for robustness, which can lead to conservative controllers [59]. Some studies have also used domain adaptation methods, that is, use real data to identify the environment [60,61] or obtain accurate physical parameters [62]. Furthermore, these methods can improve the generalization (adaptation) performance of robots in challenging environments. For more solutions to the reality gap, please refer to the relevant review paper [63].

4. OPEN PROBLEMS AND FUTURE PROSPECTS

In this section, we discuss the long-standing open questions and promising future research directions in the DRL-based quadrupedal locomotion field around these issues, as shown in Figure 5. Solutions to these open problems are described in Section 3.

4.1 Open problems

4.1.1. Sample efficiency

In many popular DRL algorithms, millions or billions of gradient descent steps are required to train policies that can accomplish the assigned task [64–66]. For real robotics tasks, therefore, such a learning process requires a significant number of interactions, which is infeasible in practical applications. In the face of increasingly complex robotic tasks, without improvement in the sample efficiency of algorithms, the number of training samples needed will only increase with model size and complexity. Furthermore, a sample-efficient DRL algo-

rithm can deal with sparse-reward tasks, which greatly reduces the difficulty of designing reward functions. It also alleviates the serious time burden for the researchers to tune the parameters of reward function.

4.1.2. Generalization and adaptation

Generalization is another fundamental problem of the DRL algorithm. Current algorithms perform well in single-task and static environments, but they struggle with multi-task and dynamically unstructured environments. That is, it is difficult for robots to acquire novel skills and quickly adapt to unseen environments or tasks. Generalization or adaptation to new scenarios remains a long-standing unsolved problem in the DRL community. In general, there are two broad categories of problems in robotics tasks: the observational generalization (adaptation) problem and the dynamic generalization (adaptation) problem. The former is a learning problem for robots considering high-dimensional state spaces, such as raw visual sensor observations. High-dimensional observations may incorporate redundant, task-irrelevant information that may impair the generalization ability of robot learning. Currently, there are many related studies published on physical manipulation[67–71] but only a few cutting-edge works on quadrupedal locomotion tasks[8,11,13]. The latter mainly takes into account the dynamic changes of the environment (e.g., robot body mass and ground friction coefficient)[72–74]. This causes the transition probability of the environment to change, i.e., the robot takes the same action in the same state, but it transitions to a different next state.

4.1.3. Partial observation

Simulators can significantly reduce the training difficulty of the DRL algorithms because we have access to the ground-truth state of the robots. However, due to the limitations of the onboard sensors of real robots, the policies are limited to partial observations that are often noisy and delayed. For example, it is difficult to accurately measure the root translation and body height of a legged robot. This problem is more pronounced when faced with locomotion or navigation tasks in complex and unstructured environments. Several approaches have been proposed to alleviate this problem, such as applying system identification[75], removing inaccessible states during training[39], adding more sensors[8,11,13], and learning to infer privileged information[7,76].

4.1.4. Reality gap

This problem is caused by differences between the simulation and real-world physics[16]. There are many sources of this discrepancy, including incorrect physical parameters, unmodeled dynamics, and random real-world environments. Furthermore, there is no general consensus on which of these sources plays the most important role. A straightforward approach is domain randomization, a class of methods that uses a wide range of environmental parameters and sensor noises to learn robust robot behaviors[39,77,78]. Since this method is simple and effective, most studies on quadrupedal locomotion have used it to alleviate the reality gap problem.

4.2 Future prospects

4.2.1. Accelerate learning via model-based planning

For sequential decision making problems, model-based planning is a powerful approach to improve sample efficiency and has achieved great success in applied domains such as game playing[79–81] and continuous control[82,83]. These methods, however, are both costly to plan over long horizons and struggle to obtain accurate world models. More recently, the strengths of model-free and model-based methods are combined to achieve superior sample efficiency and asymptotic performance on continuous control tasks[84], especially on fairly challenging, high-dimensional humanoid and dog tasks[85]. How to use model-based planning in DRL-based quadrupedal locomotion research is an issue worthy of further exploration.

4.2.2. Reuse of motion priors data

Current vanilla DRL algorithms have difficulty producing life-like natural behaviors for legged robots. Furthermore, reward functions capable of accomplishing complex tasks often require a tedious and labor-intensive tuning process. Robots also struggle to generalize or adapt to other environments or tasks. To alleviate this

problem to a certain extent, there have been recent DRL studies based on motion priors[86–90], which have been successfully applied to quadrupedal locomotion tasks[12,56,91]. However, the variety of motion priors in these studies is insufficient, and the robot's behavior is not agile and natural. This makes it difficult for robots to cope with complex and unstructured natural environments. Improving the diversity of motion priors is also an interesting direction in quadrupedal locomotion research. On the other hand, there is currently a lack of general real-world legged motion skills datasets and benchmarks, which would have significant value for DRL-based quadrupedal locomotion research. If many real-world data were available, we could study and verify offline RL[92] algorithms for quadrupedal locomotion. The main feature of offline RL algorithms is that the robot does not need to interact with the environment during the training phase, so we can bypass the notorious reality gap problem.

4.2.3. Large-scale pre-training of DRL models

The pre-training and fine-tuning paradigms for new tasks have emerged as simple yet effective solutions in supervised and self-supervised learning. Pre-trained DRL-based models enable robots to rapidly and efficiently acquire new skills and respond to non-stationary complex environments. Meta-learning methods seem to be a popular solution for improving the generalization (adaptation) performance of robots to new environments. However, current meta-reinforcement learning algorithms are limited to simple environments with narrow task distributions[93–96]. A recent study showed that multi-task pre-training with fine-tuning on new tasks performs as well as or better than meta-pre-training with meta test-time adaptation[97]. Research considering large-scale pre-trained models in quadrupedal locomotion research is still in its infancy and needs further exploration. Furthermore, this direction is inseparable from the motor skills dataset mentioned above, but it focuses more on large-scale pre-training of DRL-based models and online fine-tuning for downstream tasks.

5. CONCLUSIONS

In the past few years, there have been some breakthroughs in quadrupedal locomotion research. However, due to the limitations of algorithms and hardware, the behavior of robots is still not agile and intelligent. This review provides a comprehensive survey of several DRL algorithms in this field. We first introduce basic concepts and formulations, and then condense open problems in the literature. Subsequently, we sort out previous works and summarize the algorithm design and core components in detail, which includes DRL algorithms, simulators, hardware platforms, observation and action space design, reward function design, prior knowledge, solution of reality gap problems, *etc.* While this review considers as many factors as possible in systematically collating the relevant literature, there are still many imperceptible factors that may affect the performance of DRL-based control policies in real-world robotics tasks. Finally, we point out future research directions around open questions to drive important research forward.

DECLARATIONS

Authors' contributions

Made substantial contributions to conception and design of the study and performed data analysis and interpretation: Zhang H, Wang D

Performed data acquisition, as well as provided administrative, technical, and material support: He L

Availability of data and materials

Please refer to Table 1 and Table 2 in the appendix.

Financial support and sponsorship

This work was supported by the National Science and Technology Innovation 2030 - Major Project (Grant No. 2022ZD0208800), and NSFC General Program (Grant No. 62176215).

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Bledt G, Wensing PM, Ingersoll S, Kim S. Contact Model Fusion for Event-Based Locomotion in Unstructured Terrains. In: 2018 IEEE International Conference on Robotics and Automation (ICRA); 2018. pp. 4399–406. DOI
2. Hwangbo J, Bellicoso CD, Fankhauser P, Hutter M. Probabilistic foot contact estimation by fusing information from dynamics and differential/forward kinematics. In: 2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2016. pp. 3872–78. DOI
3. Camurri M, Fallon M, Bazeille S, et al. Probabilistic contact estimation and impact detection for state estimation of quadruped robots. *IEEE Robotics and Automation Letters* 2017;2:1023–30. DOI
4. Bloesch M, Gehring C, Fankhauser P, et al. State estimation for legged robots on unstable and slippery terrain. In: 2013 IEEE/RSJ International Conference on Intelligent Robots and Systems; 2013. pp. 6058–64. DOI
5. Sutton RS, Barto AG. Reinforcement learning: an introduction. *IEEE Trans Neural Netw* 2005;16:285–86. Available from: https://books.google.com.hk/books?hl=zh-CN&lr=&id=uWV0DwAAQBAJ&oi=fnd&pg=PR7&dq=Reinforcement+Learning:+An+Introduction&ots=mitFv1__l3&sig=tpgx07M0IomIGA4K13idTvhYFbo&redir_esc=y#v=onepage&q=Reinforcement%20Learning%3A%20An%20Introduction&f=false [Last accessed on 30 Aug 2022].
6. Hwangbo J, Lee J, Dosovitskiy A, et al. Learning agile and dynamic motor skills for legged robots. *Sci Robot* 2019;4. DOI
7. Lee J, Hwangbo J, Wellhausen L, Koltun V, Hutter M. Learning quadrupedal locomotion over challenging terrain. *Sci Robot* 2020;5. Available from: https://robotics.sciencemag.org/content/5/47/eabc5986 [last accessed on 30 Aug 2022].
8. Miki T, Lee J, Hwangbo J, et al. Learning robust perceptive locomotion for quadrupedal robots in the wild. *Sci Robot* 2022;7:eabk2822. Available from: https://www.science.org/doi/abs/10.1126/scirobotics.abk2822 [Last accessed on 30 Aug 2022].
9. Yang C, Yuan K, Zhu Q, Yu W, Li Z. Multi-expert learning of adaptive legged locomotion. *Sci Robot* 2020;5. Available from: https://robotics.sciencemag.org/content/5/49/eabb2174 [Last accessed on 30 Aug 2022].
10. Peng XB, Coumans E, Zhang T, et al. Learning agile robotic locomotion skills by imitating animals. In: Robotics: Science and Systems; 2020. DOI
11. Fu Z, Kumar A, Agarwal A, et al. Coupling vision and proprioception for navigation of legged robots. *arXiv preprint arXiv:211202094* 2021. Available from: https://openaccess.thecvf.com/content/CVPR2022/html/Fu_Coupling_Vision_and_Proprioception_for_Navigation_of_Legged_Robots_CVPR_2022_paper.html [Last accessed on 30 Aug 2022].
12. Bohez S, Tunyasuvunakool S, Brakel P, et al. Imitate and repurpose: learning reusable robot movement skills from human and animal behaviors. *arXiv preprint arXiv:220317138* 2022. DOI
13. Yang R, Zhang M, Hansen N, Xu H, Wang X. *Learning vision-guided quadrupedal locomotion end-to-end with cross-modal transformers*.
14. Zhang T, Mo H. Reinforcement learning for robot research: a comprehensive review and open issues. *Int J Advanc Robot Syst* 2021;18:17298814211007305. DOI
15. Ibarz J, Tan J, Finn C, et al. How to train your robot with deep reinforcement learning: lessons we have learned. *Int J Robot Res* 2021;40:698–721. DOI
16. Koos S, Mouret JB, Doncieux S. Crossing the reality gap in evolutionary robotics by promoting transferable controllers. In: Conference on Genetic and Evolutionary Computation. United States: ACM, publisher; 2010. pp. 119–26. DOI Available from: https://hal.archives-ouvertes.fr/hal-00633927.
17. Zhao W, Queralta JP, Westerlund T. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In: 2020 IEEE Symposium Series on Computational Intelligence (SSCI). IEEE; 2020. pp. 737–44. DOI
18. Yue J. Learning locomotion for legged robots based on reinforcement learning: a survey. In: 2020 International Conference on Electrical Engineering and Control Technologies (CEECT). IEEE; 2020. pp. 1–7. DOI

19. Sutton RS, Barto AG. *Introduction to reinforcement learning* 1998. Available from: https://login.cs.utexas.edu/sites/default/files/legacy_files/research/documents/1%20intro%20up%20to%20RL%3ATD.pdf [Last accessed on 30 Aug 2022].
20. Schulman J, Levine S, Abbeel P, Jordan M, Moritz P. Trust region policy optimization. In: International conference on machine learning. PMLR; 2015. pp. 1889–97. Available from: https://proceedings.mlr.press/v37/schulman15.html [Last accessed on 30 Aug 2022].
21. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. *CoRR* 2017;abs/1707.06347. Available from: http://arxiv.org/abs/1707.06347 [Last accessed on 30 Aug 2022].
22. Schulman J, Moritz P, Levine S, Jordan MI, Abbeel P. High-dimensional continuous control using generalized advantage estimation. In: Bengio Y, LeCun Y, editors. 4th International Conference on Learning Representations, ICLR 2016, San Juan, Puerto Rico, May 2-4, 2016, Conference Track Proceedings; 2016. Available from: http://arxiv.org/abs/1506.02438 [Last accessed on 30 Aug 2022].
23. Mania H, Guy A, Recht B. Simple random search provides a competitive approach to reinforcement learning. *CoRR* 2018;abs/1803.07055. Available from: http://arxiv.org/abs/1803.07055 [Last accessed on 30 Aug 2022].
24. Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: Dy JG, Krause A, editors. Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018. vol. 80 of Proceedings of Machine Learning Research. PMLR; 2018. pp. 1856–65. Avaialble from: http://proceedings.mlr.press/v80/haarnoja18b.html [Last accessed on 30 Aug 2022].
25. Song HF, Abdolmaleki A, Springenberg JT, et al. V-MPO: on-policy maximum a posteriori policy optimization for discrete and continuous control. OpenReview.net; 2020. Available from: https://openreview.net/forum?id=SylOlp4FvH [Last accessed on 30 Aug 2022].
26. Abdolmaleki A, Huang SH, Hasenclever L, et al. A distributional view on multi-objective policy optimization. In: Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event. vol. 119 of Proceedings of Machine Learning Research. PMLR; 2020. pp. 11–22. Avaialble from: http://proceedings.mlr.press/v119/abdolmaleki20a.html [Last accessed on 30 Aug 2022].
27. Brakel P, Bohez S, Hasenclever L, Heess N, Bousmalis K. Learning coordinated terrain-adaptive locomotion by imitating a centroidal dynamics planner. *CoRR* 2021;abs/2111.00262. Avaialble from: https://arxiv.org/abs/2111.00262 [Last accessed on 30 Aug 2022].
28. Gangapurwala S, Mitchell AL, Havoutis I. Guided constrained policy optimization for dynamic quadrupedal robot locomotion. *IEEE Robotics Autom Lett* 2020;5:3642–49. DOI
29. Chen X, Wang C, Zhou Z, Ross KW. Randomized ensembled double Q-learning: learning fast without a model. In: 9th International Conference on Learning Representations, ICLR 2021, Virtual Event, Austria, May 3-7, 2021. OpenReview.net; 2021. Avaialble from: https://openreview.net/forum?id=AY8zfZm0tDd [Last accessed on 30 Aug 2022].
30. Smith L, Kew JC, Peng XB, et al. Legged robots that keep on learning: fine-tuning locomotion policies in the real world. In: 2022 IEEE International Conference on Robotics and Automation (ICRA); 2022. pp. 1–7. DOI
31. Coumans E, Bai Y. PyBullet, a Python module for physics simulation for games, robotics and machine learning; 2016–2021. http://pybullet.org.
32. Hwangbo J, Lee J, Hutter M. Per-Contact Iteration Method for Solving Contact Dynamics. *IEEE Robotics Autom Lett* 2018;3:895–902. Avaialble from: https://doi.org/10.1109/LRA.2018.2792536 [Last accessed on 30 Aug 2022].
33. Todorov E, Erez T, Tassa Y. MuJoCo: a physics engine for model-based control. In: 2012 IEEE/RSJ International Conference on Intelligent Robots and Systems. IEEE; 2012. pp. 5026–33. DOI
34. Makoviychuk V, Wawrzyniak L, Guo Y, et al. Isaac gym: high performance GPU based physics simulation for robot learning. In: Vanschoren J, Yeung S, editors. Proceedings of the Neural Information Processing Systems Track on Datasets and Benchmarks 1, NeurIPS Datasets and Benchmarks 2021, December 2021, virtual; 2021. Avaialble from: https://datasets-benchmarks-proceedings.neurips.cc/paper/2021/hash/28dd2c7955ce926456240b2ff0100bde-Abstract-round2.html [Last accessed on 30 Aug 2022].
35. Rudin N, Hoeller D, Reist P, Hutter M. Learning to walk in minutes using massively parallel deep reinforcement learning. In: Conference on Robot Learning. PMLR; 2022. pp. 91–100. Avaialble from: https://proceedings.mlr.press/v164/rudin22a.html [Last accessed on 30 Aug 2022].
36. Margolis GB, Yang G, Paigwar K, Chen T, Agrawal P. Rapid locomotion via reinforcement learning. *arXiv preprint arXiv:220502824* 2022. DOI
37. Escontrela A, Peng XB, Yu W, et al. Adversarial motion priors make good substitutes for complex reward functions. *arXiv e-prints* 2022:arXiv:2203.15103. DOI
38. Vollenweider E, Bjelonic M, Klemm V, et al. Advanced skills through multiple adversarial motion priors in reinforcement learning. *arXiv e-prints* 2022:arXiv:2203.14912. DOI
39. Tan J, Zhang T, Coumans E, et al. Sim-to-real: learning agile locomotion for quadruped robots. In: Kress-Gazit H, Srinivasa SS, Howard T, Atanasov N, editors. Robotics: Science and Systems XIV, Carnegie Mellon University, Pittsburgh, Pennsylvania, USA, June 26-30, 2018; 2018. Avaialble from: http://www.roboticsproceedings.org/rss14/p10.html [Last accessed on 30 Aug 2022].
40. Hutter M, Gehring C, Jud D, et al. Anymal-a highly mobile and dynamic quadrupedal robot. In: 2016 IEEE/RSJ international conference on intelligent robots and systems (IROS). IEEE; 2016. pp. 38–44. DOI
41. Ha S, Xu P, Tan Z, Levine S, Tan J. *Learning to walk in the real world with minimal human effort* 2020;155:1110–20. Available from: https://proceedings.mlr.press/v155/ha21c.html [Lasta accessed on 30 Aug 2022].
42. Gangapurwala S, Geisert M, Orsolino R, Fallon M, Havoutis I. Rloc: terrain-aware legged locomotion using reinforcement learning and optimal control. *IEEE Trans Robot* 2022. DOI
43. Peng XB, van de Panne M. Learning locomotion skills using DeepRL: does the choice of action space matter? In: Teran J, Zheng C, Spencer SN, Thomaszewski B, Yin K, editors. Proceedings of the ACM SIGGRAPH / Eurographics Symposium on Computer Animation,

Los Angeles, CA, USA, July 28-30, 2017. Eurographics Association / ACM; 2017. pp. 12:1–2:13. Avaialble from: https://doi.org/10.1145/3099564.3099567 [Last accessed on 30 Aug 2022].
44. Chen S, Zhang B, Mueller MW, Rai A, Sreenath K. Learning torque control for quadrupedal locomotion. *CoRR* 2022;abs/2203.05194. Avaialble from: https://doi.org/10.48550/arXiv.2203.05194 [Last accessed on 30 Aug 2022].
45. Carlo JD, Wensing PM, Katz B, Bledt G, Kim S. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2018, Madrid, Spain, October 1-5, 2018. IEEE; 2018. pp. 1–9. Avaialble from: https://doi.org/10.1109/IROS.2018.8594448 [Last accessed on 30 Aug 2022].
46. Jain D, Iscen A, Caluwaerts K. Hierarchical reinforcement learning for quadruped locomotion. In: 2019 IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2019, Macau, SAR, China, November 3-8, 2019. IEEE; 2019. pp. 7551–57. Avaialble from: https://doi.org/10.1109/IROS40897.2019.8967913 [Last accessed on 30 Aug 2022].
47. Li T, Lambert NO, Calandra R, Meier F, Rai A. Learning generalizable locomotion skills with hierarchical reinforcement learning. In: 2020 IEEE International Conference on Robotics and Automation, ICRA 2020, Paris, France, May 31 - August 31, 2020. IEEE; 2020. pp. 413–19. Avaialble from: https://doi.org/10.1109/ICRA40945.2020.9196642 [Last accessed on 30 Aug 2022].
48. Lee J, Hwangbo J, Hutter M. Robust recovery controller for a quadrupedal robot using deep reinforcement learning. *CoRR* 2019;abs/1901.07517. Available from: http://arxiv.org/abs/1901.07517 [Lasta accessed on 30 Aug 2022].
49. Iscen A, Caluwaerts K, Tan J, et al. Policies modulating trajectory generators. In: 2nd Annual Conference on Robot Learning, CoRL 2018, Zürich, Switzerland, 29-31 October 2018, Proceedings. vol. 87 of Proceedings of Machine Learning Research. PMLR; 2018. pp. 916–26. Available from: http://proceedings.mlr.press/v87/iscen18a.html [last accessed on 30 Aug 2022].
50. Rahme M, Abraham I, Elwin ML, Murphey TD. Dynamics and domain randomized gait modulation with Bezier curves for sim-to-real legged locomotion. *CoRR* 2020;abs/2010.12070. Avaialble from: https://arxiv.org/abs/2010.12070 [Last accessed on 30 Aug 2022].
51. Zhang H, Wang J, Wu Z, Wang Y, Wang D. Terrain-aware risk-assessment-network-aided deep reinforcement learning for quadrupedal locomotion in tough terrain. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2021, Prague, Czech Republic, September 27 - Oct. 1, 2021. IEEE; 2021. pp. 4538–45. Avaialble from: https://doi.org/10.1109/IROS51168.2021.9636519 [Last accessed on 30 Aug 2022].
52. Yang Y, Zhang T, Coumans E, Tan J, Boots B. Fast and efficient locomotion via learned gait transitions. In: Faust A, Hsu D, Neumann G, editors. Conference on Robot Learning, 8-11 November 2021, London, UK. vol. 164 of Proceedings of Machine Learning Research. PMLR; 2021. pp. 773–83. Avaialble from: https://proceedings.mlr.press/v164/yang22d.html [Last accessed on 30 Aug 2022].
53. Gangapurwala S, Geisert M, Orsolino R, Fallon MF, Havoutis I. Real-time trajectory adaptation for quadrupedal locomotion using deep reinforcement learning. In: IEEE International Conference on Robotics and Automation, ICRA 2021, Xi'an, China, May 30 - June 5, 2021. IEEE; 2021. pp. 5973–79. Avaialble from: https://doi.org/10.1109/ICRA48506.2021.9561639 [Last accessed on 30 Aug 2022].
54. Yao Q, Wang J, Wang D, et al. Hierarchical terrain-aware control for quadrupedal locomotion by combining deep reinforcement learning and optimal control. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2021, Prague, Czech Republic, September 27 - Oct. 1, 2021. IEEE; 2021. pp. 4546–51. Avaialble from: https://doi.org/10.1109/IROS51168.2021.9636738 [Last accessed on 30 Aug 2022].
55. Singla A, Bhattacharya S, Dholakiya D, et al. Realizing learned quadruped locomotion behaviors through kinematic motion primitives. In: International Conference on Robotics and Automation, ICRA 2019, Montreal, QC, Canada, May 20-24, 2019. IEEE; 2019. pp. 7434–40. Avaialble from: https://doi.org/10.1109/ICRA.2019.8794179 [Last accessed on 30 Aug 2022].
56. Vollenweider E, Bjelonic M, Klemm V, et al. Advanced skills through multiple adversarial motion priors in reinforcement learning. *arXiv preprint arXiv:220314912* 2022. DOI
57. Li A, Wang Z, Wu J, Zhu Q. Efficient learning of control policies for robust quadruped bounding using pretrained neural networks. *arXiv preprint arXiv:201100446* 2020. DOI
58. Shao Y, Jin Y, Liu X, et al. Learning free gait transition for quadruped robots via phase-guided controller. *IEEE Robotics and Automation Letters* 2021;7:1230–37. DOI
59. Luo J, Hauser KK. Robust trajectory optimization under frictional contact with iterative learning. *Auton Robots* 2017;41:1447–61. Avaialble from: https://doi.org/10.1007/s10514-017-9629-x [Last accessed on 30 Aug 2022].
60. Kumar A, Fu Z, Pathak D, Malik J. RMA: rapid motor adaptation for legged robots. In: Proceedings of Robotics: Science and Systems. Virtual; 2021. DOI
61. Liu J, Zhang H, Wang D. DARA: dynamics-aware reward augmentation in offline reinforcement learning. *CoRR* 2022;abs/2203.06662. Avaialble from: https://doi.org/10.48550/arXiv.2203.06662 [Last accessed on 30 Aug 2022].
62. Shi H, Zhou B, Zeng H, et al. Reinforcement learning with evolutionary trajectory generator: A general approach for quadrupedal locomotion. *IEEE Robotics Autom Lett* 2022;7:3085–92. Avaialble from: https://doi.org/10.1109/LRA.2022.3145495 [Last accessed on 30 Aug 2022].
63. Zhao W, Queralta JP, Westerlund T. Sim-to-real transfer in deep reinforcement learning for robotics: a survey. In: 2020 IEEE Symposium Series on Computational Intelligence, SSCI 2020, Canberra, Australia, December 1-4, 2020. IEEE; 2020. pp. 737–44. Avaialble from: https://doi.org/10.1109/SSCI47803.2020.9308468 [Last accessed on 30 Aug 2022].
64. Mnih V, Kavukcuoglu K, Silver D, et al. Playing atari with deep reinforcement learning. *arXiv preprint arXiv:13125602* 2013. DOI
65. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. Proximal policy optimization algorithms. *arXiv preprint arXiv:170706347* 2017. DOI
66. Wang C, Yang T, Hao J, et al. ED2: an environment dynamics decomposition framework for world model construction. *CoRR* 2021;abs/2112.02817. Avaialble from: https://arxiv.org/abs/2112.02817 [Last accessed on 30 Aug 2022].

67. Kostrikov I, Yarats D, Fergus R. Image augmentation is all you need: regularizing deep reinforcement learning from pixels. *arXiv preprint arXiv:200413649* 2020. DOI
68. Yarats D, Fergus R, Lazaric A, Pinto L. Mastering visual continuous control: improved data-augmented reinforcement learning. *arXiv preprint arXiv:210709645* 2021. DOI
69. Ahmed O, Träuble F, Goyal A, et al. Causalworld: a robotic manipulation benchmark for causal structure and transfer learning. *arXiv preprint arXiv:201004296* 2020. DOI
70. Dittadi A, Träuble F, Wüthrich M, et al. The role of pretrained representations for the OOD generalization of RL agents. *arXiv preprint arXiv:210705686* 2021. DOI
71. Hsu K, Kim MJ, Rafailov R, Wu J, Finn C. Vision-based manipulators need to also see from their hands. *arXiv preprint arXiv:220312677* 2022. DOI
72. Eysenbach B, Asawa S, Chaudhari S, Levine S, Salakhutdinov R. Off-dynamics reinforcement learning: training for transfer with domain classifiers. *arXiv preprint arXiv:200613916* 2020. DOI
73. Liu J, Zhang H, Wang D. DARA: dynamics-aware reward augmentation in offline reinforcement learning. *arXiv preprint arXiv:220306662* 2022. DOI
74. Lee K, Seo Y, Lee S, Lee H, Shin J. Context-aware dynamics model for generalization in model-based reinforcement learning. In: Proceedings of the 37th International Conference on Machine Learning, ICML 2020, 13-18 July 2020, Virtual Event. vol. 119 of Proceedings of Machine Learning Research. PMLR; 2020. pp. 5757–66. Avaialble from: http://proceedings.mlr.press/v119/lee20g.html [Last accessed on 30 Aug 2022].
75. Yu W, Tan J, Liu CK, Turk G. Preparing for the unknown: learning a universal policy with online system identification. *arXiv preprint arXiv:170202453* 2017. DOI
76. Chen D, Zhou B, Koltun V, Krähenbühl P. Learning by cheating. In: Conference on Robot Learning. PMLR; 2020. pp. 66–75. Available from: http://proceedings.mlr.press/v100/chen20a.html [Last accessed on 30 Aug 2022].
77. Tobin J, Fong R, Ray A, et al. *Domain randomization for transferring deep neural networks from simulation to the real world* 2017. DOI
78. Peng XB, Andrychowicz M, Zaremba W, Abbeel P. *Sim-to-real transfer of robotic control with dynamics randomization* 2017. DOI
79. Kaiser L, Babaeizadeh M, Milos P, et al. Model-based reinforcement learning for atari. *arXiv preprint arXiv:190300374* 2019. DOI
80. Schrittwieser J, Antonoglou I, Hubert T, et al. Mastering atari, go, chess and shogi by planning with a learned model. *Nature* 2020;588:604–9. DOI
81. Ye W, Liu S, Kurutach T, Abbeel P, Gao Y. Mastering atari games with limited data. *Adv Neural Inform Proc Syst* 2021;34:25476–88. Available from: https://proceedings.neurips.cc/paper/2021/hash/d5eca8dc3820cad9fe56a3bafda65ca1-Abstract.html [Last accessed on 30 Aug 2022].
82. Chua K, Calandra R, McAllister R, Levine S. Deep reinforcement learning in a handful of trials using probabilistic dynamics models. *Adva neural inform proc syst* 2018;31. Available from: https://proceedings.neurips.cc/paper/2018/hash/3de568f8597b94bda53149c7d7f5958c-Abstract.html [Last accessed on 30 Aug 2022].
83. Janner M, Fu J, Zhang M, Levine S. When to trust your model: model-based policy optimization. *Adv Neural Inform Proc Syst* 2019;32. Available from: https://proceedings.neurips.cc/paper/2019/hash/5faf461eff3099671ad63c6f3f094f7f-Abstract.html [Last accessed on 30 Aug 2022].
84. Hansen N, Wang X, Su H. Temporal difference learning for model predictive control. *arXiv preprint arXiv:220304955* 2022. DOI
85. Tassa Y, Doron Y, Muldal A, et al. Deepmind control suite. *arXiv preprint arXiv:180100690* 2018. DOI
86. Peng XB, Ma Z, Abbeel P, Levine S, Kanazawa A. Amp: adversarial motion priors for stylized physics-based character control. *ACM Trans Graph (TOG)* 2021;40:1–20. DOI
87. Peng XB, Guo Y, Halper L, Levine S, Fidler S. ASE: large-scale reusable adversarial skill embeddings for physically simulated characters. *arXiv preprint arXiv:220501906* 2022. DOI
88. Merel J, Hasenclever L, Galashov A, et al. Neural probabilistic motor primitives for humanoid control. *arXiv preprint arXiv:181111711* 2018. DOI
89. Hasenclever L, Pardo F, Hadsell R, Heess N, Merel J. Comic: complementary task learning & mimicry for reusable skills. In: International Conference on Machine Learning. PMLR; 2020. pp. 4105–15. Available from: https://proceedings.mlr.press/v119/hasenclever20a.html [Last accessed on 30 Aug 2022].
90. Liu S, Lever G, Wang Z, et al. From motor control to team play in simulated humanoid football. *arXiv preprint arXiv:210512196* 2021. DOI
91. Escontrela A, Peng XB, Yu W, et al. Adversarial motion priors make good substitutes for complex reward functions. *arXiv preprint arXiv:220315103* 2022. DOI
92. Levine S, Kumar A, Tucker G, Fu J. Offline reinforcement learning: Tutorial, review, and perspectives on open problems. *arXiv preprint arXiv:200501643* 2020. DOI
93. Duan Y, Schulman J, Chen X, et al. Rl ^2: Fast reinforcement learning via slow reinforcement learning. *arXiv preprint arXiv:161102779* 2016. DOI
94. Nichol A, Achiam J, Schulman J. On first-order meta-learning algorithms. *arXiv preprint arXiv:180302999* 2018. DOI
95. Rakelly K, Zhou A, Finn C, Levine S, Quillen D. Efficient off-policy meta-reinforcement learning via probabilistic context variables. In: International Conference on Machine Learning. PMLR; 2019. pp. 5331–40. Available from: http://proceedings.mlr.press/v97/rakelly19a.html [Last accessed on 30 Aug 2022].
96. Seo Y, Lee K, James SL, Abbeel P. Reinforcement learning with action-free pre-training from videos. In: International Conference on

Machine Learning, ICML 2022, 17-23 July 2022, Baltimore, Maryland, USA. vol. 162 of Proceedings of Machine Learning Research. PMLR; 2022. pp. 19561–79. Avaialble from: https://proceedings.mlr.press/v162/seo22a.html [Last accessed on 30 Aug 2022].
97. Mandi Z, Abbeel P, James S. On the Effectiveness of Fine-tuning Versus Meta-reinforcement Learning. *arXiv preprint arXiv:220603271* 2022. DOI
98. Haarnoja T, Zhou A, Ha S, et al. *Learning to walk via deep reinforcement learning* 2019. DOI
99. Yang Y, Caluwaerts K, Iscen A, et al. Data efficient reinforcement learning for legged robots. In: Kaelbling LP, Kragic D, Sugiura K, editors. 3rd Annual Conference on Robot Learning, CoRL 2019, Osaka, Japan, October 30 - November 1, 2019, Proceedings. vol. 100 of Proceedings of Machine Learning Research. PMLR; 2019. pp. 1–10. Avaialble from: http://proceedings.mlr.press/v100/yang20a.html [Last accessed on 30 Aug 2022].
100. Tsounis V, Alge M, Lee J, Farshidian F, Hutter M. DeepGait: planning and control of quadrupedal gaits using deep reinforcement learning. *IEEE Robot Autom Lett* 2020;5:3699–706. DOI
101. Da X, Xie Z, Hoeller D, et al. Learning a contact-adaptive controller for robust, efficient legged locomotion. PMLR; 2020. Available from: https://proceedings.mlr.press/v155/da21a.html [Lasta accessed on 30 Aug 2022].
102. Liang J, Makoviychuk V, Handa A, et al. GPU-accelerated robotic simulation for distributed reinforcement learning. *CoRR* 2018;abs/1810.05762. Avaialble from: http://arxiv.org/abs/1810.05762 [Last accessed on 30 Aug 2022].
103. Escontrela A, Yu G, Xu P, Iscen A, Tan J. Zero-shot terrain generalization for visual locomotion policies. *CoRR* 2020;abs/2011.05513. Avaialble from: https://arxiv.org/abs/2011.05513 [Last accessed on 30 Aug 2022].
104. Jiang Y, Zhang T, Ho D, et al. *SimGAN: hybrid simulator identification for domain adaptation via adversarial reinforcement learning* 2021:2884–90. Available from: https://doi.org/10.1109/ICRA48506.2021.9561731 [last accessed on 30 Aug 2022].
105. Tan W, Fang X, Zhang W, et al. A hierarchical framework for quadruped locomotion based on reinforcement learning. In: IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS 2021, Prague, Czech Republic, September 27 - Oct. 1, 2021. IEEE; 2021. pp. 8462–68. Avaialble from: https://doi.org/10.1109/IROS51168.2021.9636757 [Last accessed on 30 Aug 2022].
106. Michel O. WebotsTM: professional mobile robot simulation. *CoRR* 2004;abs/cs/0412052. Avaialble from: http://arxiv.org/abs/cs/0412052 [Last accessed on 30 Aug 2022].
107. Fu Z, Kumar A, Malik J, Pathak D. Minimizing energy consumption leads to the emergence of gaits in legged robots. *CoRR* 2021;abs/2111.01674. Avaialble from: https://arxiv.org/abs/2111.01674 [Last accessed on 30 Aug 2022].
108. Kim S, Sorokin M, Lee J, Ha S. Human motion control of quadrupedal robots using deep reinforcement learning. *arXiv preprint arXiv:220413336* 2022. Avaialble from: http://www.roboticsproceedings.org/rss18/p021.pdf [Last accessed on 30 Aug 2022].
109. Bogdanovic M, Khadiv M, Righetti L. Model-free reinforcement learning for robust locomotion using trajectory optimization for exploration. *arXiv preprint arXiv:210706629* 2021. DOI
110. Fernbach P, Tonneau S, Stasse O, Carpentier J, Taïx M. C-CROC: continuous and convex resolution of centroidal dynamic trajectories for legged robots in multicontact scenarios. *IEEE Trans Robot* 2020;36:676–91. DOI
111. Zhang H, Starke S, Komura T, Saito J. Mode-adaptive neural networks for quadruped motion control. *ACM Trans Graph (TOG)* 2018;37:1–11. DOI
112. Feldman A, Goussev V, Sangole A, Levin M. Threshold position control and the principle of minimal interaction in motor actions. *Progr brain res* 2007 02;165:267–81. DOI
113. Winkler AW, Bellicoso CD, Hutter M, Buchli J. Gait and trajectory optimization for legged systems through phase-based end-effector parameterization. *IEEE Robot Autom Lett* 2018;3:1560–67. DOI
114. Liu H, Jia W, Bi L. Hopf oscillator based adaptive locomotion control for a bionic quadruped robot. *2017 IEEE Int Confer Mechatr Autom (ICMA)* 2017:949–54. DOI
115. Carlo JD, Wensing PM, Katz B, Bledt G, Kim S. Dynamic locomotion in the MIT cheetah 3 through convex model-predictive control. In: 2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS); 2018. pp. 1–9. DOI
116. Bellicoso D, Jenelten F, Fankhauser P, et al. Dynamic locomotion and whole-body control for quadrupedal robots. *2017 IEEE/RSJ Int Conf Intell Robots Sys (IROS)* 2017:3359–65. DOI
117. Sethian JA. Fast marching methods. *SIAM Rev* 1999;41:199–235. DOI
118. Ponton B, Khadiv M, Meduri A, Righetti L. Efficient multi-contact pattern generation with sequential convex approximations of the centroidal dynamics. *CoRR* 2020;abs/2010.01215. Avaialble from: https://arxiv.org/abs/2010.01215 [Last accessed on 30 Aug 2022].
119. Zhang T, McCarthy Z, Jow O, et al. Deep imitation learning for complex manipulation tasks from virtual reality teleoperation. *CoRR* 2017;abs/1710.04615. Avaialble from: http://arxiv.org/abs/1710.04615 [Last accessed on 30 Aug 2022].
120. Thor M, Kulvicius T, Manoonpong P. Generic neural locomotion control framework for legged robots. *IEEE Trans Neural Netw Learn Syst* 2021;32:4013–25. DOI

APPENDIX

Table 1. Classification of the most relevant publications in DRL-based quadrupedal locomotion research

Publication	Pub. Year	Description	Algorithm	State Space	Action Space	Reward Function	Simulator	Robot
Sim-to-Real: Learning Agile Locomotion For Quadruped Robots [39]	RSS 2018	A system to design agile locomotion by leveraging DRL algorithms.	PPO	Orientation (2-dim), Base Angular Velocities (2-dim), and Motor Angles (8-dim).	Desired Leg Pose.	Current and Previous Base Positions, Desired Running Direction, Motor Torques and Velocities.	Pybullet	Minitaur
Policies Modulating Trajectory Generators [49]	CoRL 2018	An architecture for learning behaviors by using PMTG that provides memory and prior knowledge.	PPO	Orientation, Base Angular Velocities, Desired Velocity (control input), and Phase of the TG.	Leg Swing Angles and Extensions, Frequency, Amplitude, Walking Height.	Speed Gap (Desired vs. Acutual).	Pybullet	Minitaur
Robust Recovery Controller for a Quadrupedal Robot using Deep Reinforcement Learning [48]	ArXiv 2019	A model-free DRL approach to control recovery maneuvers using a hierarchical controller.	TRPO + GAE [22]	Self-Righting: Gravity, Base Angular Velocity, Joint Positions, Velocities and History. Standing Up: Base Linear Velocity, State Space (Self-Righting). Locomotion: Command, Base Height, State Space (Standing Up).	Desired Joint Positions.	Angular and Linear Velocity, Height, Orientation, Torque, Power, Joint State, Body Impulse and Slippage, Foot Slippage and Clearance, Self Collision, and Action Gap.	RaiSim	ANYmal
Learning to Walk via Deep Reinforcement Learning [98]	RSS 2019	A sample-efficient Max. Entropy RL algorithm requiring minimal per-task tuning to learn neural network policies.	SAC	Motor Angles (8-dim), Orientation (2-dim), and Base Angular Velocities (2-dim).	Leg Swing Angles and Extensions.	Walking Distance, Joint Accelerations and Angles, and Base Roll Angle.	Pybullet	Minitaur
Data Efficient Reinforcement Learning for Legged Robots [99]	CoRL 2019	A model-based RL framework for learning locomotion from only 4.5 minutes of data collected on a quadruped robot.	MPC + CEM	Base Linear Velocity, Orientation (3-dim), and Motor Positions.	Leg Swing Angles and Extensions, and Phase Scales.	Speed Gap, and Base Orientation.	Pybullet	Minitaur
Hierarchical Reinforcement Learning for Quadruped Locomotion [46]	IROS 2019	A hierarchical framework to automatically decompose complex locomotion tasks.	ARS	High-Level: Base Position and Orientation. Low-Level: PMTG State, Orientation, Base Angular Velocities, and Latent Command.	High-Level: Command, Duration. Low-Level: Motor Position, PMTG Param.	Steering Angle, Moving Distance.	Pybullet	Minitaur
Realizing Learned Quadruped Locomotion Behaviors through Kinematic Motion Primitives (kMPs) [55]	ICRA 2019	kMPs is effective to learn quadrupedal walking using DRL, and realize these behaviors in Stoch.	PPO	Joint Angles, Velocities and Torques, and Orientation.	Leg End-Point Positions.	Position Change, Energy Consumption, and Motor Torques and Velocities.	Pybullet	Stoch
DeepGait: Planning and Control of Quadrupedal Gaits using Deep Reinforcement Learning [100]	ICRA 2019	A technique for training terrain-aware locomotion, which combines Model-Based Planning and RL.	PPO, TRPO, GAE	Planner: Terrain Elevation, Base State and Velocity, Joint Torques, Feet State, Phase Variables. Controller: Phase (Current&Next), Joint Angles and Velocities.	Planner: Candidate Phase. Controller: Desired Joint Positions.	Planner: Goal Distance and Orientation, Work, Stance Phases. Controller: Target and Slip, Joint Velocities and Torques, Base Velocities, Angular Deviation.	RaiSim	ANYmal

Learning agile and dynamic motor skills for legged robots [6]	Science Robotics 2019	A Sim2Real method leveraging fast, automated and cost-effective data generation schemes.	TRPO	Height, Base Velocities, Joint State (with History), Previous Action, Command, Gravity.	Desired Joint Positions.	Locomotion: Base Velocities, Cost, Torque, Joint Speed, Foot State, Direction, Fluency. Recovery: Torque, Joint Motion, HAA, HFE, KFE, Slip and Impulse, Internal Contact, Direction, Fluency.	RaiSim	ANYmal
Learning to Walk in the Real World with Minimal Human Effort [41]	CoRL 2020	A system for learning quadrupedal locomotion policies with Deep RL in the real world with minimal human effort.	SAC	Motor Angles (6-step), IMU Readings (6-step), and Previous Action (6-step).	Desired Joint Positions.	Base Position and Yaw, and Smoothness.	Pybullet	Minitaur
Dynamics and Domain Randomized Gait Modulation with Bezier Curves for Sim-to-Real Legged Locomotion [50]	ArXiv 2020	A quadrupedal Sim2Real framework utilizing offline RL with dynamics and domain randomized to allow traversing uneven terrain.	ARS [23]	Orientation (2-dim), Base Angular Velocities (3-dim) and Linear Accelerations (3-dim), and Foot Phase.	Foot Position Residuals.	Distance, Orientation (2-dim), Base Angular Velocities (3-dim).	Pybullet	Mini Spot
Guided Constrained Policy Optimization for Dynamic Quadrupedal Robot Locomotion [28]	IEEE Robotics Autom 2020	A CPPO-based RL framework for tracking velocity commands under constraints.	GCPO	Base Height and Velocities, Orientation, Joint State, Policy Output, Desired Base Velocity.	Desired Joint Positions.	Linear and Angular Velocity, Torque, Foot Acceleration and Slip, Smoothness, and Orientation.	RaiSim, PyBullet and MuJoCo.	ANYmal
Learning a Contact-Adaptive Controller for Robust, Efficient Legged Locomotion [101]	CoRL 2020	A hierarchical framework combining Model-Based Control and RL to synthesize robust quadrupedal controllers.	DQN	Pose (without linear positions and foot positions), Primitive (previously-used).	One-Hot Primitive Selection Vector (9-dim)	Torques, Base Linear Velocity, Desired Base Linear Velocity.	IsaacGym [102]	Unitree Laikago
Zero-Shot Terrain Generalization for Visual Locomotion Policies [103]	ArXiv 2020	A learning approach for terrain locomotion using exteroceptive inputs without ground-truth height maps.	PPO	Distance to Env., Orientation, Base Velocity, Joint Angles, Target Distance and Direction, Previous Action, Trajectory Generator Param.	Gait Frequency, Swing Height, Stride Length, Residual Action.	Euclidean Distance (Base to Target) and Timestep Duration.	Pybullet	Unitree Laikago
Learning Generalizable Locomotion Skills with Hierarchical Reinforcement Learning [47]	ICRA 2020	A sample-efficient and generalizable hierarchical framework for learning locomotion skills on real-world robots.	SAC	Phase (1-dim)	Desired Joint Positions.	Forward Distance and Orientation Deviation.	PyBullet	Daisy Hexapod
Learning Agile Robotic Locomotion Skills by Imitating Animals [10]	RSS 2020	A system enabling legged robot to learn agile locomotion skills by imitating real-world animals.	PPO	Pose (Orientation (3-dim) and Joint Rotations) and Action Sequence.	Joint Torques for Desired Positions.	Motion Gap (Current vs. Reference in Joint Velocities and State, End-Effector Positions, Base Pose and Velocity).	PyBullet	Unitree-Laikago

Learning Quadrupedal Locomotion over Challenging Terrain [7]	Science Robotics 2020	A novel Sim2Real solution incorporating proprioception showing remarkable zero-shot generalization.	TRPO	Goal Direction, Gravity, Base Velocities and Frequency, Joint States and Velocities, FTG Phases and Frequencies, Joint History, Terrain Normal Vector, Foot Height, Contact Forces and Target History, Contact States, Friction, External Force.	Leg Frequencies and Foot Position Residuals.	Linear and Angular Velocity, Base Motion Reward and Collision, Foot Clearance, Target Smoothness, and Torque.	RaiSim	ANYmal-B, ANYmal-C
Multi-expert learning (MEL) of adaptive legged locomotion [9]	Science Robotics 2020	A MEL architecture to generate adaptive skills from a group of expert skills.	SAC	Joint Position, Gravity, Base Velocities, Phase Vector, and Goal Position.	Expert: Desired Joint Positions. Gating: Variable Weights.	Base Pose, Height and Velocity, Regularisation (Torque, Velocity), Foot State, Body State, Reference Positions and Contacts, Goal Position.	PyBullet	Jueying[3]
Efficient Learning of Control Policies for Robust Quadruped Bounding using Pretrained Neural Networks [57]	ArXiv 2021	A training method for learning bounding gaits, which combines pre-training and DRL.	PPO2	Base Height, Gravity Direction, Base Angular Velocity and Linear Acceleration, and Joint Position and Angular Velocity.	Desired Joint Positions.	Base Velocity, Joint Torque and Velocity, Gait, Position uniformity, Torque uniformity, Smoothness, and Pitch Limitations.	RaiSim	Jueying-Mini robot
Learning Coordinated Terrain-Adaptive Locomotion by Imitating a Centroidal Dynamics Planner [27]	ArXiv 2021	A terrain adaptive controller obtained by training policies to reproduce trajectories planned by a non-linear solver.	V-MPO [25], MO-VMPO [26],	Image, State (Base, End-Effector, Joint, CoM), Velocities (Base, Joint), Orientation, Previous Action, Command.	Desired Joint Positions.	Joint Positions, Base Position, End-Effector Positions, Base Linear and Angular Velocities, and Quaternion.	Mujoco	ANYmal
Learning Free Gait Transition for Quadruped Robots via Phase-Guided Controller [58]	IEEE Robotics and Automation Letters 2021	A novel quadrupedal framework for training a control policy to locomote in various gaits.	PPO	Velocity Command, Sine and Cosine Values (4 phases), Joint Position and Velocity, Angular Velocities, Gravity.	Desired Joint Positions.	Joint Torque, Desired Velocity, Base Balance, and Foot Contact.	RaiSim	Black Panther robot
Fast and Efficient Locomotion via Learned Gait Transitions [52]	CoRL 2021	A hierarchical learning framework in which gait transitions emerge automatically with a reward of min. energy.	ARS [23]	Desired and Actual Base Linear Velocity	Desired Leg Frequency, Cutoff (Swing and Stance Phase), Phase Offset.	Torques, Base Velocities, and Survival.	Pybullet	Unitree A1
SimGAN: Hybrid Simulator Identification for Domain Adaptation via Adversarial RL [104]	ICRA 2021	A framework for domain adaptation by identifying a simulator to match the simulated trajectories to the target ones.	PPO	Orientation, Base Height, Base Linear Velocities and Joint Angles (12-dim).	Desired Joint Torques.	Base Forward Velocity, Joint Limit Count, and Torque.	Pybullet	Unitree Laikago
Hierarchical Terrain-Aware Control (HTC) for Quadrupedal Locomotion by Combining DRL and Optimal Control [54]	IROS 2021	A novel HTC framework leveraging DRL for the high-level and optimal control for the low-level.	SAC	Global Height Map, Motors Positions, Orientation, and Gait Phase.	Goal Swing, Base Height and Velocity, Orientation.	Desired Forward Velocity to Target and Orientation.	Pybullet	Unitree A1

[3]https://www.deeprobotics.cn/

A Hierarchical Framework for Quadruped Locomotion Based on RL [105]	IROS 2021	A well-performing quadruped robot system for learning locomotion in real-world terrains without pre-training.	SAC	Angle Error, Orientation, and Command.	Goal Position, Base Velocity, Pitch, and Leg Lift Max. Height.	Base Target Positions (with Previous), Orientation, and Survival.	Webots [106]	Yobogo
Terrain-Aware Risk-Assessment-Network-Aided (RAN) DRL for Quadrupedal Locomotion in Tough Terrain [51]	IROS 2021	A terrain-aware DRL-based controller integrating a RAN to guarantee the action stability.	SAC	Elevation Map, Goal Direction, Base Velocities, Joint State and Velocity, FTG Phases, Frequencies (Base and FTG), Joint History, Foot Targets and Contact Forces, Contact States, Env. Param.	Desired Joint Positions.	Linear and Angular Velocity, Base Motion and Collision, Foot Clearance, Target Smoothness, Torque, and Traversability Map.	Pybullet	Unitree A1
Real-Time Trajectory Adaptation for Quadrupedal Locomotion using DRL [53]	ICRA 2021	A policy using DRL to get noisy reference trajectory in order to generate a quadrupedal tracking system.	PPO	Robot Positions and Velocities, Reference Positions and Velocities, Corrected Trajectory Positions and Velocities, Height Maps.	Base State and Velocity Deviation, End-Effector State and Velocities Deviation.	Torque, Foot State, Smoothness, Orientation, Joint Motion, Trajectory Tracking, Goal.	RaiSim	ANYmal
Coupling Vision and Proprioception for navigation of Legged Robots [11]	CVPR workshop 2021	Incorporating vision and proprioception in navigation tasks of legged robots.	PPO	Proprioception, Command Velocities, Previous Action and Extrinsic Vector.	Desired Joint Positions.	Velocity Gap, Energy Consumption, Lateral Movement, and Hip Joints.	RaiSim	Unitree A1
Minimising Energy Consumption Leads to the Emergence of Gaits in Legged Robots [107]	CoRL 2021	Energy constraints leading to the emergence of natural locomotion, and the choice is related to the desired speed.	PPO	Joint Positions and Velocities, Orientation, Foot Contact, Previous Action.	Desired Joint Positions.	Linear and Angular Velocity, and Joint Torques and Velocities.	RaiSim	Unitree A1
RMA: Rapid Motor Adaptation for Legged Robots [60]	RSS 2021	RMA algorithm for real-time online adaptation problems in quadruped robots.	PPO	Joint Positions and Velocities, Orientation (2-dim), and Foot Contact Vector.	Desired Joint Positions.	Base Motion and Orientation, Work, Ground Impact, Smoothness, Joint Speed, Foot Slip.	RaiSim	Unitree A1
Human Motion Control of Quadrupedal Robots using DRL [108]	RSS 2022	A quadrupedal motion control system allowing human operation.	PPO	Sensor Data (with history), Actions History, and Reference Poses (with history).	Desired Joint Positions [43]	Joint Imitation, End-Effector, Base Sate, Deviation, and Acceleration.	RaiSim	Unitree A1
Learning Torque Control for Quadrupedal Locomotion [44]	ArXiv 2022	A quadrupedal torque control framework predicting high-frequency joint torques via RL.	PPO	Base Velocities, Gravity, Joint Position and Velocity, Command, Last Action.	Desired Torques.	Base Velocities and Height, Orientation, Joint Motion, Foot State, Knee Collision, Action Rate, Gaits, and Hips Reward.	Isaac, Pybullt	Unitree A1
Model-free RL for Robust Locomotion using Demonstrations from Trajectory Optimization [109]	ArXiv 2022	A RL approach to create robust policies deployable on real robots without additional training using a single optimised demonstration.	PPO	Joint Positions and Velocities, Orientation (Yall and Roll) and Angular Velocities (Yall and Roll).	Desired Joint Positions.	Joint Position, and Base Position, Quaternion and Angular Velocity.	Pybullet	Solo8
Legged Robots that Keep on Learning: Fine-Tuning Locomotion Policies in the Real World [30]	ICRA 2022	A robot RL system for fine-tuning real-world locomotion policies.	REDQ [29]	Orientation (3-step), Joint Angles (3-step), Actions (3-step), Future Target Poses.	Desired Joint Positions.	Joint States and Velocities, End-Effector State, Base Pose and Velocity.	Pybullet	Unitree A1

RLOC: Terrain-Aware Legged Locomotion using RL and Optimal Control [42]	IEEE Transactions on Robotics 2022	A unified model-based and data-driven approach for quadrupedal locomotion over uneven terrain.	SAC, TD3, GCPO [28]	Planning: Base State, Joint States, Goal Velocity, Elevation Map. Adaption: Base State, Feet Goal, Torques, Elevation Map. Recovery: Joint Position, Goal Positions and Velocity.	Planning: Coordinates. Adaption: Joint Torques. Recovery: Desired Joint Positions.	Planning: Base Velocities, Torque, Foot Slip, Stability. Adaption: State Deviation, Robot State. Recovery: State Space (Planning), Foot Motion (Foot, Joint), Smoothness.	RaiSim	ANYmal B, ANYmal C
Rapid Locomotion via Reinforcement Learning [36]	RSS 2022	A MIT Mini Cheetah controller achieving record agility.	PPO	Joint Angles and Velocities, Gravity, Previous Actions, Goal Velocity.	Desired Joint Positions.	Velocity Tracking, Base Pose, Self-Collision, Joint Limits, Torques, Action Rate, Airtime.	IsaacGym	MIT Mini Cheetah
Learning to Walk in Minutes Using Massively Parallel DRL [35]	CoRL 2022	A robotic training framework achieving fast policy generation via parallelism.	PPO	Base Velocities, Gravity, Joint Motion, Previous Actions, Terrain Measurements.	Desired Joint Positions.	Velocity Tracking, Joint Motion, Torques, Action Rate, Collisions, Feet Airtime.	IsaacGym	ANYmal B, ANYmal C, Unitree A1
Adversarial Motion Priors Make Good Substitutes for Complex Reward Functions [91]	ArXiv 2022	Substituting reward functions with stylish rewards learned from motion captures.	PPO	Joint Angles and Velocities, Orientations and Previous Actions.	Joint Torques for Desired Positions.	Linear and Angular Velocity Tracking, and Motion Prior Discrimination	Issac Gym	Unitree A1
Advanced Skills through Multiple Adversarial Motion Priors in RL [56]	ArXiv 2022	An adversarial motion prior-based RL approach to allow for multiple, discretely switchable styles.	PPO	Base State and Velocities, Gravity, Joint Positions and Velocity, and Wheel Positions.	Desired Joint Positions.	Linear and Angular Velocity Tracking, Pose, and Joint Velocity and Position.	Issac Gym	Quadruped Humanoid Transformer
Learning robust perceptive locomotion for quadrupedal robots in the wild [8]	Science Robotics 2022	A quadrupedal locomotion solution integrating exteroceptive and proprioceptive perception.	PPO	Command, Base Pose and Motion, Joint History, CPG Phase, Height Samples, Contact States, Friction, External Forces, Airtime.	Phase Offset, Joint Position Target.	Velocities, Body Motion, Foot Clearance, Collisions, Joint Motion and Constraint, Smoothness, Torque, Slip.	RaiSim	ANYmal-C
Imitate and Repurpose: Learning Reusable Robot Movement Skills From Human and Animal Behaviors [12]	ArXiv 2022	Learning reusable locomotion skills for real legged robots using prior knowledge of human and animal movement.	V-MPO, MO-VMPO,	Base States and Motion, Latent Command, Joint States and Velocities, Gravity, Goal Velocity and Position, Ball Position, End-Effector Position, Clip ID.	High-Level: Latent Command. Low-Level: Desired Joint Positions.	Imitation: Tracking CoM, Joint Velocities, End-Effector Positions, Body Quaternions, Current Draw. Walking: Tracking Velocity. Ball Dribbling: Tracking Ball Positions.	MuJoCo	ANYmal
Learning vision-guided quadrupedal locomotion end-to-end with cross-modal transformers [13]	ICLR 2022	An end-to-end RL method leveraging both proprioceptive states and visual observations for locomotion control.	PPO	Orientation, Joint Rotations, Previous Actions (3-step), and Image (4 dense depth).	Desired Joint Positions.	Distance, Motor Torques, Survival, and Collected Sphere Count.	Pybullet	Unitree A1
RL with Evolutionary Trajectory Generator: A General Approach for Quadrupedal Locomotion [62]	IEEE Robotics Autom 2022	A novel RL-based approach containing an evolutionary foot trajectory generator.	SAC	Orientation (3-dim), Joint Angles and Angular Velocities, Feet Contact Vector (4-dim), and Base Velocity.	Desired Joint Positions.	Base Position, Desired Direction, and Consumed Energy.	Pybullet	Unitree A1

Table 2. More information about publications (Supplement to Table 1)

Publication	Others	Solution to Reality Gap	Open-Source Package

Sim-to-Real: Learning Agile Locomotion For Quadruped Robots [39]	Open-loop Controller.	Improving Simulation Fidelity (Actuator Model, Latency), and Dynamics Randomization (Mass, Motor Strength, Inertia, Control Step, Latency, Battery, Friction, IMU bias and noise.	https://github.com/bulletphysics/bullet3/tree/master/examples/pybullet/gym/pybullet_envs/minitaur/envs
Policies Modulating Trajectory Generators [49]	Trajectory Generator.	Random Directional Virtual Forces.	/
Robust Recovery Controller for a Quadrupedal Robot using Deep Reinforcement Learning [48]	/	Randomized Physical Properties, Actuator Model, Additive Noise to the Observation.	/
Data Efficient Reinforcement Learning for Legged Robots [99]	Trajectory Generators.	/	/
Hierarchical Reinforcement Learning for Quadruped Locomotion [46]	PMTG [49]	/	/
Realizing Learned Quadruped Locomotion Behaviors through Kinematic Motion Primitives [55]	Recorded Data (Joint Angles and Orientation for 4800 steps).	/	/
DeepGait: Planning and Control of Quadrupedal Gaits using Deep Reinforcement Learning [100]	CROC [110]	/	/
Learning agile and dynamic motor skills for legged robots [6]	A controller generating foot trajectories to train the actuator model.	Actuator Model, Curriculum Training, Randomised Bodies (Size and Position), Random Command and Initial State.	/
Dynamics and Domain Randomized Gait Modulation with Bezier Curves for Sim-to-Real Legged Locomotion [50]	Open-loop Bezier curve Gait Generator.	Domain Randomization (Base Mass , Leg Link Masses, Foot Friction, XYZ Mesh Magnitude).	https://github.com/OpenQuadruped/spot_mini_mini
Guided Constrained Policy Optimization for Dynamic Quadrupedal Robot Locomotion [28]	/	Noisy Observations and Actions, and Domain Randomisation (Gravity, Actuator Torque Scaling, Link Mass and Size, Actuator Damping, and Step Time).	/
Learning a Contact-Adaptive Controller for Robust, Efficient Legged Locomotion [101]	A simple model-based method [45].	/	/
Zero-Shot Terrain Generalization for Visual Locomotion Policies [103]	PMTG	/	/
Learning Generalizable Locomotion Skills with Hierarchical Reinforcement Learning [47]	MPC (High-Level Planning), Sinusoidal Policy (TG, Low-Level Controller).	/	/

Learning Agile Robotic Locomotion Skills by Imitating Animals [10]	MoCap Clips [111].	Domain Randomization and Domain Adapation.	https://github.com/erwincoumans/motion_imitation
Learning Quadrupedal Locomotion over Challenging Terrain [7]	Foot Trajectory Generator.	Actuator Model, Randomized Physical Parameters, Teach-Student Training Set-up, and Automated Curriculum synthesizing Terrains.	/
Multi-expert learning of adaptive legged locomotion [9]	Reference Gait to provide Joint Position Reward and Foot Contact Reward.	Action Filtering and Smoothing Loss [112].	/
Efficient Learning of Control Policies for Robust Quadruped Bounding using Pretrained Neural Networks [57]	SLIP: Spring Linear Inverted Pendulum (Model-based Controller).	Domain randomization (Link Mass, Inertia and CoM, Initial Direction and Ground Friction and Restitution).	/
Learning Coordinated Terrain-Adaptive Locomotion by Imitating a Centroidal Dynamics Planner [27]	TOWR [113]	/	/
Learning Free Gait Transition for Quadruped Robots via Phase-Guided Controller [58]	Hopf Oscillator and manually designed functions [114].	Domain Randomization (External Force and Torque, Ground Friction and Height, Mass, Body Size, Noise of Joint Position and Velocity, Body Posture, and Angular Velocity).	https://github.com/ZJU-XMech/PhaseGuidedControl
Fast and Efficient Locomotion via Learned Gait Transitions [52]	Centroidal Dynamics Model [115].	/	/
SimGAN: Hybrid Simulator Identification for Domain Adaptation via Adversarial Reinforcement Learning [104]	/	Hybrid Simulator Identification.	/
Hierarchical Terrain-Aware Control for Quadrupedal Locomotion by Combining Deep Reinforcement Learning and Optimal Control [54]	Optimal Control.	Domain Randomization (Mass, Inertia, Motor Strength and Friction, Latency, Lateral Friction, Battery, Joint Friction, CoM position noise, External force, and Step Height and width.	/
A Hierarchical Framework for Quadruped Locomotion Based on Reinforcement Learning [105]	Trajectory Generator.	Domain Randomization (Leg Profile and Mass, Base Mass Distribution, Leg Inertia Matrix, and Communication Delay).	/
Terrain-Aware Risk-Assessment-Network-Aided Deep Reinforcement Learning for Quadrupedal Locomotion in Tough Terrain [51]	PMTG [49].	Domain Randomization (Mass, Inertia, Motor Strength and Friction, Latency, Lateral friction, Battery, Joint friction).	/

Real-Time Trajectory Adaptation for Quadrupedal Locomotion using Deep Reinforcement Learning [53]	TOWR [113], WBC [116].	Domain Randomization, Actuator Modelling, Shifting Initial Position, Changing Gravity, Actuator Torque Scaling, and Perturbing the Robot Base.	/
Coupling Vision and Proprioception for navigation of Legged Robots [11]	FMM [117], librealsense[4].	RMA-based Adaption Module [60].	/
Minimizing Energy Consumption Leads to the Emergence of Gaits in Legged Robots [107]	/	RMA-based Adaption Module.	/
RMA: Rapid Motor Adaptation for Legged Robots [60]	/	RMA-based Adaption Module.	/
Human Motion Control of Quadrupedal Robots using Deep Reinforcement Learning [108]	Human Motions.	Domain randomization (Link Mass, Ground Friction Coefficients and Slope, Proportional and Derivative Gain, and Communication Delay).	/
Learning Torque Control for Quadrupedal Locomotion [44]	/	Domain randomization (Base Linear and Angular Velocity, Projected Gravity, Joint Position and Velocity, Ground Friction and External Disturbances).	/
Model-free Reinforcement Learning for Robust Locomotion using Demonstrations from Trajectory Optimization [109]	Trajectory Optimization Algorithm [118].	/	/
Legged Robots that Keep on Learning: Fine-Tuning Locomotion Policies in the Real World [30]	MoCap Dog Recording [111] and Side-Step motion for A1 [10].	Real-World Fine-Tuning.	https://github.com/lauramsmith/fine-tuning-locomotion
RLOC: Terrain-Aware Legged Locomotion using Reinforcement Learning and Optimal Control [42]	Dynamic Gaits Controller.	Domain randomization (Gravity, Actuator Torque Scaling, Link Mass and Size, Actuator Damping), Perturbation on Robot Base, and Smoothing Filters for Elevation Map, Actuator Model.	/
Rapid Locomotion via Reinforcement Learning [36]	Curriculum Strategy.	Domain Randomization.	/
Learning to Walk in Minutes Using Massively Parallel Deep Reinforcement Learning [35]	/	Domain Randomization (Ground Friction and External Force), and Noisy Observations.	https://github.com/leggedrobotics/legged_gym

[4]https://github.com/IntelRealSense/librealsense

Adversarial Motion Priors Make Good Subsitutes for Complex Reward Functions [91]	German Shepherd Motion Dataset [119].	Domain Randomization (Friction, Base Mass, Velocity Perturbation, Motor Gain Multiplier).	https://github.com/Alescontrela/AMP_for_hardware
Advanced Skills through Multiple Adversarial Motion Priors in Reinforcement Learning [56]	Motion Data from another RL policy or an MPC controller.	Actuator Model, Domain Randomisation (Rough Terrin, Disturbances, External Force), Curriculum Training, and Joint-Velocity-Based Trajectory Termination.	/
Learning robust perceptive locomotion for quadrupedal robots in the wild [8]	Foot Trajectory Generator.	Actuator Model, Domain Randomisation (Robot Mass, Initial Joint Position and Velocity, Orientation, External Force, Friction Coefficient, Curriculum Learning, and Randomized Height Sampling.	/
Imitate and Repurpose: Learning Reusable Robot Movement Skills From Human and Animal Behaviors [12]	MoCap Dataset of dog walking and turning behaviors [111].	Domain Randomization (Body Mass, Centre of Mass, Joint Position and Reference, Joint Damping and Friction loss, Geom Friction, P Gain, and Torque Limit).	/
Learning Vision-Guided Quadrupedal Locomotion End-to-End with Cross-Modal Transformers [13]	/	Domain Randomization (KP, KD, Inertia, Lateral Friction, Mass, Motor Friction and Strength, and Sensor Latency), and Random Depth Input.	https://github.com/Mehooz/vision4leg
Reinforcement Learning with Evolutionary Trajectory Generator: A General Approach for Quadrupedal Locomotion [62]	Evolutionary Trajectory Generator [120].	Teacher-Student Learning Setting, Domain Adaptation,and Domain Randomization (Control Latency, Foot Friction, Base Mass, Leg Mass, and Motor Kp, Kd), and Noisy Sensor Input.	https://github.com/PaddlePaddle/PaddleRobotics/tree/main/QuadrupedalRobots/ETGRL

Peng *et al. Intell Robot* 2022;2(3):298-312
DOI: 10.20517/ir.2022.27

Intelligence & Robotics

Research Article

Open Access

Check for updates

Networked scheduling for decentralized load frequency control

Chen Peng[1], Hongchenyu Yang[1]

[1]School of Mechatronic Engineering and Automation, Shanghai University, Shanghai 200444, China.

Correspondence to: Prof. Chen Peng, School of Mechatronic Engineering and Automation, Shanghai University, No. 99, Shangda Road, Baoshan District, Shanghai 200444, China. E-mail: c.peng@shu.edu.cn

How to cite this article: Peng C, Yang H. Networked scheduling for decentralized load frequency control. *Intell Robot* 2022;2(3):298-312. http://dx.doi.org/10.20517/ir.2022.27.

Received: 14 Aug 2022 **First Decision:** 17 Aug 2022 **Revised:** 30 Aug 2022 **Accepted:** 1 Sep 2022 **Published:** 3 Sep 2022

Academic Editor: Simon X. Yang **Copy Editor:** Jia-Xin Zhang **Production Editor:** Jia-Xin Zhang

Abstract

This paper investigates the scheduling process for multi-area interconnected power systems under the shared but band-limited network and decentralized load frequency controllers. To cope with sub-area information and avoid node collision of large-scale power systems, round-robin and try-once-discard scheduling are used to schedule sampling data among different sub-grids. Different from existing decentralized load frequency control methods, this paper studies multi-packet transmission schemes and introduces scheduling protocols to deal with the multi-node collision. Considering the scheduling process and decentralized load frequency controllers, an impulsive power system closed-loop model is well established. Furthermore, sufficient stabilization criteria are derived to obtain decentralized H_∞ output feedback controller gains and scheduling protocol parameters. Under the designed decentralized output feedback controllers, the prescribed system performances are achieved. Finally, a three-area power system example is used to verify the effectiveness of the proposed scheduling method.

Keywords: Multi-area power systems, round-robin scheduling, try-once-discard scheduling, load frequency control

1. INTRODUCTION

With the vigorous development of power markets, the degree of interconnection among power grid regions is increasing. At present, power grids have developed into multi-area interconnected power systems[1,2]. Generally, there are two communication schemes to transmit data between neighborhood areas, that is, the dedicated channel and the open infrastructure. Since the latter has outstanding advantages over the former, such

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License (https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, sharing, adaptation, distribution and reproduction in any medium or format, for any purpose, even commercially, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

as higher flexibility and lower cost, it has been widely used in the communication of multi-area interconnected power systems[3–5].

An important indicator to measure power system operation quality is the fluctuation in frequency, and load frequency control (LFC) is widely used to maintain frequency interchanges at scheduled values and stifle frequency fluctuation caused by load disturbances[2,6,7]. Under the open communication infrastructure, the sensor in each area collects data information. Then, data is transmitted to the decentralized controller under the shared but band-limited network. The corresponding control commands are issued to actuators in each area respectively. However, due to the introduction of the network, the design and operation of LFC face some new challenges, such as node collision, data loss and network-induced delay[8–10].

Due to different locations of sub-region power grids and the wide distribution of system components, multi-channel transmission is inevitable in multi-area interconnected power systems, while most of the existing results[10–13] take the general assumption that sampled data is packaged into a single packet to transmit, which is not applied in large-scale multi-area interconnected power systems. On the other hand, the shared network has limited bandwidth, where the simultaneous transmission under multiple channels may cause node congestion. To solve this problem, scheduling protocols have been presented to decide which node to gain access to the communication network[15?].

Generally, multi-channel scheduling includes Round-Robin (RR) scheduling[16,17], try-once-discard (TOD) scheduling[18,19], and stochastic scheduling[20]. Under the RR scheduling protocol, each node is transmitted periodically with a fixed period whose value is the total number of transmission channels. Under the TOD scheduling, the sensor node with the largest scheduling error has access to the channel. Recently, a time-delay analysis method has been discussed to derive stability criteria for networked control systems (NCSs) which are scheduled by the above three communication protocols[14,21]. Besides, a hybrid system method has been employed to analyze NCSs with variable delays under the TOD communication protocol[19], where a partial exponential stabilization criterion has been derived. The H_∞ filtering of NCSs with multiple nodes has been investigated[22], in which TOD protocol is used to schedule sampled data.

This paper studies the H_∞ LFC for multi-area interconnected power systems with decentralized controllers under the shared but band-limited network. RR and TOD protocols are used to schedule the sampling information of different sub-grids, which could greatly improve communication efficiency. Through linear matrix inequality (LMI) technology and Lyapunov analysis methods, sufficient conditions that guarantee the prescribed H_∞ performance of the studied system are derived. Decentralized controller gains and protocol parameters are obtained simultaneously. The main contributions are summarized as follows.

(1) RR and TOD protocols are used to deal with the multi-node collision of large-scale power systems, which improve communication efficiency greatly. Compared with the existing LFC methods[10,23,24], the scheduling process under multi-area transmission schemes is investigated.

(2) An networked power system impulsive closed-loop model is well constructed, which covers the multi-channel scheduling, packet dropout, disturbance, and network-induced delays in a unified framework. Compared with the system without disturbance[19], this paper studies the anti-disturbance performance of the studied system. An H_∞ LFC method is presented to obtain decentralized controller gains and scheduling protocol parameters simultaneously.

1.1. Notations

Throughout this paper, $\mathbb{R}_n$ stands for the n-dimensional Euclidean space with vector norm $\|\cdot\|$. $diag\{\cdots\}$ and $col\{\cdots\}$ denote the block-diagonal matrix and block-column vector, respectively. The superscript T stands for

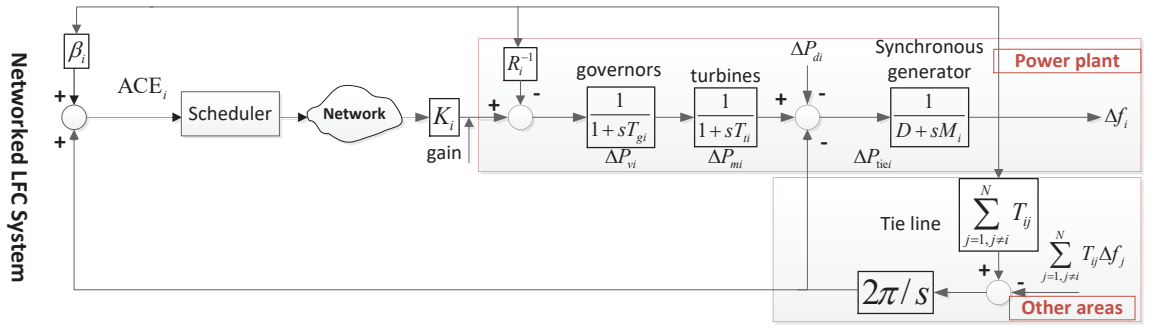


Figure 1. Multi-area decentralized power systems under scheduling protocols.

the transpose of a matrix or a vector. I is an identity matrix with an appropriate dimension. Matrix $X > 0 (X \geq 0)$ means that X is a positive definite (positive semi-definite) symmetric matrix. The symbol $*$ is the symmetric term in a matrix. $w[-h, 0]$ denotes the Banach type of absolutely continuous functions $\phi : [-h, 0] \rightarrow \mathbb{R}_n$ with $\dot{\phi} \in \mathbb{L}_2(-h, 0)$ with the norm $\| \phi \|_w = \max_{v \in [-h,0]} \| \phi(v) \| + [\int_{-h}^{0} \| \dot{\phi}(v) \|^2 ds]^{\frac{1}{2}}$.

2. PROBLEM FORMULATION

Considering single-packet size constraints, the dynamic model of multi-area interconnected power systems with decentralized load frequency controllers is constructed in this section. An impulsive system model under TOD or RR protocol is established.

2.1. Multi-area decentralized LFC model

Diagram of the multi-area decentralized LFC under scheduling protocols is shown in Figure 1, where data transmission from sensors to decentralize controllers is scheduled by RR or TOD scheduling protocol.

The system model is represented as[23,25]:

$$\begin{cases} \dot{x}(t) &= Ax(t) - Bu(t) + D\omega(t), \\ y(t) &= Cx(t), \end{cases} \tag{1}$$

where $x(t) \in \mathbb{R}_x$ is the state vector, $y(t) \in \mathbb{R}_{n_y}$ is the measurement and $\omega(t) \in \mathbb{R}_\omega$ is the disturbance, $u(t) \in \mathbb{R}_u$ is the control input vector. The multi-area power systems consist of generators, turbines, and governors, where $\Delta P_{vi}, \Delta P_{mi}, \Delta P_{di}$, and Δf_i denote the deviation of valve position, generator mechanical output, load, and the frequency of the ith sub-area, respectively. The area control error $ACE_i(s)$ is

$$ACE_i(s) = \beta_i \Delta f_i(s) + \Delta P_{tie-i}(s). \tag{2}$$

The state and measured output signals are

$$x_i^T(t) = [\Delta f_i(t), \Delta P_{tie-i}(t), \Delta P_{mi}, \Delta P_{vi}(t), \int ACE_i(t)]$$
$$\omega_i(t) = \Delta P_{di}(t), \; y_i^T(t) = [ACE_i(t), \int ACE_i(t)]$$

where $x(t) = [x_1^T(t), \cdots, x_n^T(t)]^T$, $y(t) = [y_1^T(t), \cdots, y_n^T(t)]^T \in \mathbb{R}_{n_y}, \omega(t) = [\omega_1^T(t), \cdots, \omega_n^T(t)]^T, A =$

$[A_{ij}]_{n\times n}, B = diag\{B_1, \cdots, B_n\}, C = [C_1^T, \cdots, C_N^T]^T, F = diag\{F_1, \cdots, F_n\}, C_i = [\underbrace{0, \cdots, 0}_{i-1}, \bar{C}_i, \underbrace{0, \cdots, 0}_{N-i}]$,

$$A_{ii} = \begin{bmatrix} -\frac{D_i}{M_i} & -\frac{1}{M_i} & \frac{1}{M_i} & 0 & 0 \\ 2\pi \sum\limits_{j=1, j\neq i}^{n} T_{ij} & 0 & 0 & 0 & 0 \\ 0 & 0 & -\frac{1}{T_{chi}} & \frac{1}{T_{chi}} & 0 \\ -\frac{1}{R_i T_{gi}} & 0 & 0 & -\frac{1}{T_{gi}} & 0 \\ \beta_i & 1 & 0 & 0 & 0 \end{bmatrix},$$

$$A_{ij} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ -2\pi T_{ij} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}, F_i = \begin{bmatrix} -\frac{1}{M_i} \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, B_i = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \frac{1}{T_{gi}} \\ 0 \end{bmatrix}, \bar{C}_i^T = \begin{bmatrix} \beta_i & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The multi-area power system includes N controlled areas, N decentralized controllers, which are connected via the network. Then, measurement signals are given by $y_i(t) = C_i x(t) \in \mathbb{R}_{n_i}, i = 1, \cdots, N, \sum\limits_{i=1}^{N} n_i = n_y$. Based on the phasor measurement unit, sampling instants of N sensors are synchronized in different areas of power systems. Denote sampling instant by s_q, satisfying that $0 = s_0 < s_1 < \cdots < s_q < \cdots, s_{q+1} - s_q \leq \varrho, \lim\limits_{q\to+\infty} s_q = +\infty$, where ϱ is the maximum allowable transfer interval.

In this paper, we take imperfect network conditions into account, such as data loss and network-induced delay. Denote the sequence after packet loss by $\{s_k\} \subseteq \{s_q\}$; that is, only at sampling instants s_k, the input of the controller can be updated. For $i = 1, \cdots, N$, an uncertain, time-varying delay $v_{s_k}^i \in [0, \tau_M^i]$ is assumed to occur, where τ_M^i is delay upper bound of the ith channel and $\max\limits_{i=1,\cdots,N}\{\tau_M^i\} = \tau_M$. Buffers are set to store and choose the largest communication delay τ_{s_k} of power system channels, i.e., $\tau_{s_k} = \max\limits_{i=1,\cdots,N}\{v_{s_k}^i\}$. Then, ZOH updating instant is $t_k = s_k + \tau_{s_k}$. The transmission delay is assumed to be bound, satisfying:

$$0 \leq \tau_{s_k} \leq \tau_M, \ t_{k+1} - t_k - \tau_{s_k} \leq \tau_M + \varrho = \bar{\tau}. \tag{3}$$

Let $\hat{y}(s_k) = [\hat{y}_1^T(s_k) \cdots \hat{y}_N^T(s_k)]^T \in \mathbb{R}_{n_y}$ denote the output signal transmitted to the scheduler. At instant s_k, only one node can be active. Let $\iota_k \in \{1, \cdots, N\}$ be the active output node at s_k, which is dependent on the scheduling rule. Then, we have

$$\hat{y}_i(s_k) = \begin{cases} y_i(s_k), i = \iota_k \\ \hat{y}_i(s_{k-1}), i \neq \iota_k \end{cases} \tag{4}$$

Consider the scheduling error between output $y_i(s_k)$ and the last available measurement $\hat{y}_i(s_{k-1})$:

$$e_i(t) = -y_i(s_k) + \hat{y}_i(s_{k-1}), \ e(t) = col\{e_1(t), \cdots, e_N(t)\}, \ t \in [t_k, t_{k+1}), \tag{5}$$

where $\hat{y}_i(s_{-1}) \triangleq 0, \ i = 1, \cdots, N$. In this paper, controllers and actuators in the ith area are event-driven. Let $L_i > 0, i = 1, \cdots, N$ be node weighting matrices. Under TOD scheduling,

$$\iota_k = \arg \max_{i\in\{1,\cdots,N\}} |\sqrt{L_i}(-y_i(s_k) + \hat{y}_i(s_{k-1}))|^2. \tag{6}$$

Under RR scheduling protocol, the active node ι_k is selected periodically:

$$\iota_k = \iota_{k+N}. \tag{7}$$

In the following, we will design the decentralized LFC law under the communication network. Similar to [23], a PI controller is used in this paper:

$$u_i(t) = -K_{Pi}ACE_i(t) - K_{Ii}\int ACE_i(t). \tag{8}$$

Therefore, dynamic model of the scheduled power systems under the decentralized LFC Equation (8) and imperfect network environments can be formalized as

$$\begin{cases} \dot{x}(t) = Ax(t) - BK\hat{y}(s_k) + F\omega(t) \\ y(t) = Cx(t), t \in [t_k, t_{k+1}), \end{cases} \tag{9}$$

where $u(t) = \sum_{i=1}^{N} K_i y_i(t) = Ky(t),\ K = [K_1, \cdots, K_N],\ K_i = [K_{Pi}, K_{Ii}]$.

2.2. Impulsive model and study objective

From Equation (5), one can obtain that

$$e_i(t_{k+1}) = \begin{cases} C_i x(s_k) - C_i x(s_{k+1}), i = \iota_k \\ C_i[x(s_k) - x(s_{k+1})] + e_i(t_k), i \neq \iota_k \end{cases} \tag{10}$$

Define an artificial delay $\tau(t) = t - s_k$, from which one arrives at

$$0 \leq \tau_{s_k} \leq \tau(t) \leq s_{k+1} - s_k + \tau_{s_{k+1}} \leq \tau_M,\ \dot{\tau}(t) = 1. \tag{11}$$

From Equation (5) and Equation (9), the impulsive power system model can be

$$\dot{x}(t) = Ax(t) - BKCx(t - \tau(t)) - \sum_{i=1, i\neq\iota_k}^{N} BK_iC_ie_i(t) + F\omega(t). \tag{12}$$

For system Equation (12), the initial condition of $x(t)$ on $[-\tau_M, 0]$ is supplemented as $x(t) = \phi(t), t \in [-\tau_M, 0]$, with $\phi(0) = x_0$, where $\phi(t)$ is a continuous function on $[-\tau_M, 0]$.

Using scheduling scheme Equation (6) and Equation (7), this paper is to design the decentralized controller Equation (8) such that system Equation (12) is exponentially stable with a prescribed H_∞ performance γ.

3. MAIN RESULTS

In the following, we first derive sufficient criteria under scheduling scheme Equation (6) and Equation (7) to ensure the exponential stability of system Equation (12) with a prescribed H_∞ performance. Then, criteria are proposed to design decentralized controllers under multi-channel transmission.

3.1. Stability analysis under the TOD scheduling scheme Equation (6)

Construct Lyapunov-Krasovskii functional candidate:

$$V(t) = \sum_{i=1}^{N} e_i^T(t)L_ie_i(t) + \Pi(t) + V_H, \tag{13}$$

where $O > 0, H_i > 0, T > 0, L_i > 0, \alpha > 0, S > 0, i = 1, \cdots, N, t \in [t_k, t_{k+1})$,

$$V_H = \sum_{i=1}^{N}\int_{s_k}^{t} \tau_M e^{2\alpha(s-t)} \left\|\sqrt{H_i}C_i\dot{x}(s)\right\|^2 ds,$$

$$\Pi(t) = \int_{t-\tau_M}^{t} e^{2\alpha(s-t)}x^T(s)Sx(s)ds + x^T(t)Ox(t) + \tau_M\int_{t-\tau_M}^{t}\int_{s}^{t} e^{2\alpha(s-t)}\dot{x}^T(v)T\dot{x}(v)dvds.$$

Remark 1 *V_H is introduced to cope with reset conditions, which is continuous on $[t_k, t_{k+1})$, and does not grow in the jumps when $t = t_{k+1}$ since*

$$-V_H(t_{k+1}^-) + V_H(t_{k+1}) \leq -\sum_{i=1}^{N} e^{-2\alpha\tau_M} \left\| \sqrt{H_i} C_i(-x(s_{k+1}) + x(s_k)) \right\|^2 \tag{14}$$

Theorem 1 *Under TOD scheduling scheme Equation (6), for given scalars $\tau_M, \alpha > 0, \gamma > 0$, impulsive system Equation (12) is exponentially stable with a prescribed H_∞ performance γ, if there exist appropriate dimensions Y and real matrices $O > 0, S > 0, T > 0, H_i > 0, J_i > 0, L_i > 0, i = 1, \cdots, N$ such that*

$$\begin{bmatrix} \Xi_{11} & * & * \\ \Xi_{21} & \Xi_{22} & * \\ \Xi_{31} & 0 & \Xi_{33} \end{bmatrix} < 0, \tag{15}$$

$$\begin{bmatrix} T & Y^T \\ Y & T \end{bmatrix} > 0, \tag{16}$$

$$\Omega_i = \begin{bmatrix} \Omega_{11} & \Omega_{12} \\ * & \Omega_{22} \end{bmatrix} < 0 \tag{17}$$

are feasible for $i, j = 1, \cdots, N$, where

$$\Xi_{11} = \begin{bmatrix} \Psi_{11} & \Psi_{12} & \Psi_{13} & \Psi_{14} & \Psi_{15} \\ * & \Psi_{22} & \Psi_{23} & 0 & 0 \\ * & * & \Psi_{33} & 0 & 0 \\ * & * & * & \Psi_{44} & 0 \\ * & * & * & * & \Psi_{55} \end{bmatrix}, \bar{\xi}_i = \begin{cases} [-PBK_2C_2, \cdots, -PBK_NC_N], \iota_k = 1 \\ [-PBK_1C_1, \cdots, -PBK_{N-1}C_{N-1}], \iota_k = N \\ [-PBK_1C_1, \cdots, -PBK_jC_{j|j\neq\iota_k}, \cdots, -PBK_NC_N], \iota_k \neq 1, N \end{cases}$$

$$\phi_i = \begin{cases} diag\{\psi_2, \cdots, \psi_N\}, \iota_k = 1 \\ diag\{\psi_1, \cdots, \psi_{N-1}\}, \iota_k = N \\ diag\{\psi_1, \cdots, \psi_{j|j\neq\iota_k}, \cdots, \psi_N\}, \iota_k \neq 1, N \end{cases}, \pi_i = \begin{cases} [BK_2C_2, \cdots, BK_NC_N], \iota_k = 1 \\ [BK_1C_1, \cdots, BK_{N-1}C_{N-1}], \iota_k = N \\ [BK_1C_1, \cdots, BK_jC_{j|j\neq\iota_k}, \cdots, BK_NC_N], \iota_k \neq 1, N. \end{cases}$$

$$\Psi_{11} = OA + A^TO + 2\alpha O + S + T, \Psi_{13} = -Y^T, \Psi_{14} = \chi_i, \Psi_{23} = Y^T - T, F_i = [A, BKC, 0, \pi_i, F],$$
$$\Psi_{22} = 2T - Y - Y^T, \Psi_{33} = T - S, \Psi_{44} = \chi_i, \Omega_{11} = -\tfrac{1-2\alpha\tau_M}{N-1} L_i + J_i, \Psi_{12} = -OBKC - T + Y^T,$$
$$\chi_i = [-OBK_1C_1, \cdots, -OBK_NC_N], \Omega_{22} = L_i - H_ie^{-2\alpha\tau_M}, \Xi_{22} = diag\{-T, -H_1, \cdots, -H_N\},$$
$$\Psi_{15} = OF, \Omega_{12} = L_i, \Xi_{21} = [\tau_M TF_i, \sqrt{\tau_M} H_1C_1F_1, \cdots, \sqrt{\tau_M} H_NC_NF_N], \Psi_{55} = -\gamma^2 I,$$
$$\psi_j = -\tfrac{1}{\tau_M} J_j + 2\alpha L_j, \Theta = \tau_M^2 T + \tau_M \sum_{i=1}^{N} C_i^T H_i C_i, \Xi_{31} = [C, 0, \cdots, 0].$$

Proof: Differentiating $V(t)$ along Equation (12) and applying Wirtinger-based integral inequality[26], we can obtain

$$\begin{aligned} &\dot{V}(t) + 2\alpha V(t) - \frac{1}{\tau_M} \sum_{i=1, i\neq\iota_k}^{N} |\sqrt{J_i} C_i e_i(t)|^2 - 2\alpha |\sqrt{L_{\iota_k}} e_{\iota_k}(t)|^2 + |y(t)|^2 - \gamma^2 |\omega(t)|^2 \leq -\frac{1}{\tau_M} \sum_{l=1, l\neq i}^{N} |\sqrt{J_l} C_l e_l(t)|^2 \\ &- \gamma^2 |\omega(t)|^2 - e^{-2\alpha\tau_M} x^T(t-\tau_M) S x(t-\tau_M) + |y(t)|^2 + \dot{x}^T(t) \Theta \dot{x}(t) - e^{-2\alpha\tau_M} \tau_M \int_{t-\tau_M}^{t} \dot{x}^T(s) T \dot{x}(s) ds \\ &+ 2\alpha \sum_{l=1, l\neq i}^{N} e_l^T(t) C_l L_l C_l e_l(t) + 2x^T(t) O \dot{x}(t) + x^T(t)(2\alpha O + S) x(t), \end{aligned} \tag{18}$$

where $\xi_i(t) = col\{x(t), x(t-\tau(t)), x(t-\tau_M), \bar{\xi}_i(t), \omega(t)\}$.

We use the reciprocally convex approach[27] to deal with the cross item in Equation (18). By using Schur's complement, one can get

$$\dot{V}(t) + 2\alpha V(t) - \frac{1}{\tau_M} \sum_{i=1, i\neq \iota_k}^{N} |\sqrt{J_i} C_i e_i(t)|^2 - 2\alpha |\sqrt{L_{\iota_k}} e_{\iota_k}(t)|^2 + |y(t)|^2 - \gamma^2 |\omega(t)|^2 < 0. \tag{19}$$

In the following, we will prove that

(i) with $\omega(t) = 0$, the impulsive system Equation (12) is exponentially stable.

Since $\frac{d}{dt} e^{2\alpha t} V(t) = e^{2\alpha t}(\dot{V}(t) + 2\alpha V(t))$ and $V(t)$ is continuous in $t \in [t_k, t_{k+1})$, the integration of Equation (19) yields

$$e^{2\alpha t} V(t) - e^{2\alpha t_k} V(t_k) \leq \int_{t_k}^{t} e^{2\alpha s} ds [2\alpha |\sqrt{L_{\iota_k}} e_{\iota_k}(t)|^2 + \frac{1}{\tau_M} \sum_{i=1, i\neq \iota_k}^{N} |\sqrt{J_i} C_i e_i(t)|^2]. \tag{20}$$

It follows from $\int_{t_k}^{t} e^{2\alpha(s-t)} ds \leq \tau_M$ that

$$V(t_{k+1}) \leq \Theta_k + e^{2\alpha(t_k - t_{k+1})} V(t_k), \tag{21}$$

where

$$\Theta_k = \sum_{j=1}^{N} [-V_H(t_{k+1}^-) + |\sqrt{L_{\iota_k}} e_{\iota_k}(t_{k+1})|^2 - |\sqrt{L_{\iota_k}} e_{\iota_k}(t_k)|^2 + V_H(t_{k+1})] + \sum_{i=1, i\neq \iota_k}^{N} |\sqrt{J_i} C_i e_i(t)|^2 + 2\alpha\tau_M |\sqrt{L_{\iota_k}} e_{\iota_k}(t_k)|^2. \tag{22}$$

Then taking Equation (17) and Equation (22) into account, denoting $\kappa_i = [e_i(t_k), C_i[-x(s_{k+1}) + x(s_k)]]^T$, we have

$$\Theta_k \leq -|(\sqrt{H_{\iota_k}} e^{-2\alpha\tau_M} - \sqrt{L_{\iota_k}}) C_{\iota_k} [x(s_k) - x(s_{k+1})]|^2 + \sum_{i=1, i\neq \iota_k}^{N} \kappa_i \Omega_i \kappa_i \leq 0. \tag{23}$$

Hence, one can conclude that

$$V(t) \leq \sum_{i=1, i\neq \iota_k}^{N} |\sqrt{J_i} e_i(t_k)|^2 + e^{2\alpha(t_0 - t)} V(t_0) + 2\alpha\tau_M |\sqrt{L_{\iota_k}} e_{\iota_k}(t_k)|^2. \tag{24}$$

From Equation (24), we have

$$V(t) \geq \lambda_{min}\{O\} |x(t)|^2 + \sum_{i=1}^{N} |\sqrt{L_i} e_i(t)|^2,$$

which implies that

$$\lambda_{min}\{O\} |x(t)|^2 \leq (2\alpha\tau_M - 1) |\sqrt{L_{\iota_k}} e_{\iota_k}(t)|^2 + \sum_{i=1, i\neq \iota_k}^{N} |\sqrt{J_i - L_i} e_i(t)|^2 + e^{2\alpha(t_0 - t)} V(t_0).$$

It follows from $\Omega_i < 0$ that for $t \geq t_0$

$$\lambda_{min}\{O\} |x(t)|^2 \leq e^{2\alpha(t_0 - t)} V(t_0). \tag{25}$$

For $t < t_0$, the system takes the form $\dot{x} = Ax(t) + F\omega(t)$, then its corresponding solution is given by

$$x(t) = e^{At} x(0) + \int_{0}^{t} e^{A(-s+t)} F\omega(s) ds. \tag{26}$$

Clearly, there exists the maximum of $|x(t)|$ for $t \in [t_0 - \tau_M, t_0]$. Thus, $V(t_0) \leq \mu \|x_{t_0}\|_W^2 + \sum_{i=1}^{N} e_i^T(t_0) L_i e_i(t_0)$ for some $\mu > 0$.

Therefore, the exponential stability of the system Equation (12) with $\omega(t) = 0$ is guaranteed.

Next, we will show that

(ii) under the zero initial condition, the inequality $\|y\|_{L_2} \leq \gamma \|\omega\|_{L_2}$ holds for any nonzero $\omega \in L_2[0, +\infty)$.

From Equation (19), we obtain that for $t \in [t_k, t_{k+1})$,

$$|y(t)|^2 - \gamma^2 |\omega(t)|^2 \leq -\dot{V}(t) + 2\alpha |\sqrt{L_{\iota_k}} e_{\iota_k}(t)|^2 + \frac{1}{\tau_M} \sum_{i=1, i \neq \iota_k}^{N} |\sqrt{J_i} C_i e_i(t)|^2. \tag{27}$$

Integrating Equation (27) on t from t_k to t_{k+1}^- yields

$$\int_{t_k}^{t_{k+1}^-} |y(s)|^2 - \gamma^2 |\omega(s)|^2 ds + V(t_{k+1}) \leq V(t_k) + \Theta_k \leq V(t_k). \tag{28}$$

Then, by summing Equation (28) on k from 0 to ρ, where $\rho \to +\infty$, we have

$$\int_{t_0}^{t_{\rho+1}^-} (|y(s)|^2 - \gamma^2 |\omega(s)|^2) ds \leq V(t_0). \tag{29}$$

Under the zero initial condition, $\int_0^{+\infty} y^T(s) y(s) ds \leq \int_0^{+\infty} \gamma^2 \omega^T(s) \omega(s) ds$. Therefore, one can derive that under the zero initial condition, $\|y\|_{L_2} \leq \gamma \|\omega\|_{L_2}$ for any nonzero $\omega \in L_2[0, +\infty)$. This completes the proof. □

Remark 2 *The feasibility of the linear martix inequalities has been sufficiently explained in*[23,28]*. Due to page limitations, this part is omitted here.*

3.2. Stability analysis under the RR scheduling scheme Eqution (7)

Construct the following Lyapunov-Krasovskii functional candidate:

$$V(t) = \Pi(t) + V_H + V_L, \ k \geq N-1, \ t \in [t_k, t_{k+1}), \tag{30}$$

where

$$V_H = \begin{cases} \sum_{i=1}^{N-1} \tau_M \int_{s_k}^{t} e^{2\alpha(s-t)} |\sqrt{H_i} C_i \dot{x}(s) ds, \ k \neq N-1, \\ \sum_{i=1}^{N-1} \tau_M \int_{s_0}^{t} e^{2\alpha(s-t)} |\sqrt{H_i} C_i \dot{x}(s) ds, \ k = N-1, \end{cases}$$

$$H_i = (N-1) L_i e^{2\alpha(N-1)\tau_M}, V_L = \sum_{i=1}^{N-1} \frac{t_{k+1} - t}{i \tau_M} |\sqrt{L_{\iota_{k-j}}} e_{\iota_{k-j}}(t)|^2.$$

Similar to Theorem 5.2 in[29], we establish the following result.

Theorem 2 *Under RR scheduling scheme Equation (7), given $\tau_M > 0$ and $\alpha > 0$, assume that there exist matrices $O > 0, S > 0, T > 0, L_i > 0, i = 1, \cdots, N$ and Y with appropriate dimensions such that Equation (15) and Equation (16) are feasible with $J_i = \frac{L_i}{N-1}$, where G_i is given by Equation (30). Then, system Equation (12) is exponentially stable with a prescribed H_∞ performance γ.*

Proof: The detailed derivation process can refer to[29] and the proof of Theorem 1, which is omitted here due to the limited pages. □

3.3. Controller Design under the TOD scheduling scheme Equation (6)

Theorem 3 *Under TOD scheduling scheme Equation (6), for given matrices A, B, C, F, and scalars $\tau_M, \alpha > 0, \gamma > 0, \zeta > 0$, system Equation (12) is exponentially stable with a prescribed H_∞ performance γ, if there exist real matrices $Z > 0, \tilde{S} > 0, \tilde{T} > 0, i, j = 1, \cdots, N, \tilde{J}_i > 0, H_i > 0, \tilde{L}_i > 0$ and appropriate dimensions $\tilde{Y}, \beth, \digamma$ such that*

$$\min tr\left\{\sum_{i=1}^{N} H_i h_i\right\} \tag{31}$$

$$s.t.\ Equation\ (17), Equation\ (32)^*, Equation\ (33), Equation\ (34), Equation\ (35)$$

$$\begin{bmatrix} \tilde{\Xi}_{11} & * & * \\ \tilde{\Xi}_{21} & \tilde{\Xi}_{22} & * \\ \tilde{\Xi}_{31} & 0 & \tilde{\Xi}_{33} \end{bmatrix} < 0, \tag{32}$$

$$\begin{bmatrix} \tilde{T} & \tilde{Y}^T \\ \tilde{Y} & \tilde{T} \end{bmatrix} > 0, \tag{33}$$

$$\begin{bmatrix} h_i & * \\ I & H_i \end{bmatrix} > 0, \tag{34}$$

$$\begin{bmatrix} -\zeta I & (\beth C - CZ)^T \\ \beth C - CZ & -I \end{bmatrix} < 0, \tag{35}$$

is solvable, and the controller gain is given by $K = \digamma \beth^{-1}$, where Equation (32)* is equivalent to Equation (32) by replacing $-H_i^{-1} - Z^T \tilde{T}_r^{-1} Z$ with $-h_i, \rho^2 \tilde{T}_r - 2\rho Z$, and

$$\tilde{\Xi}_{11} = \begin{bmatrix} \tilde{\Psi}_{11} & \tilde{\Psi}_{12} & \tilde{\Psi}_{13} & \tilde{\Psi}_{14} & \tilde{\Psi}_{15} \\ * & \tilde{\Psi}_{22} & \tilde{\Psi}_{23} & 0 & 0 \\ * & * & \tilde{\Psi}_{33} & 0 & 0 \\ * & * & * & \tilde{\Psi}_{44} & 0 \\ * & * & * & * & \tilde{\Psi}_{55} \end{bmatrix},$$

$$\tilde{\phi}_i = \begin{cases} diag\{\tilde{\psi}_2, \cdots, \tilde{\psi}_N\}, \iota_k = 1 \\ diag\{\tilde{\psi}_1, \cdots, \tilde{\psi}_{N-1}\}, \iota_k = N \\ diag\{\tilde{\psi}_1, \cdots, \tilde{\psi}_{j|j\neq\iota_k}, \cdots, \tilde{\psi}_N\}, \iota_k \neq 1, N \end{cases}$$

$$\tilde{\tilde{\xi}}_i = \begin{cases} [-BK_2C_2Z, \cdots, -BK_NC_NZ], \iota_k = 1 \\ [-BK_1C_1Z, \cdots, -BK_{N-1}C_{N-1}Z], \iota_k = N \\ [-BK_1C_1Z, \cdots, -BK_jC_{j|j\neq\iota_k}Z, \cdots, -BK_NC_NZ], \iota_k \neq 1, N \end{cases}$$

$$\tilde{\psi}_j = -\tfrac{1}{\tau_M}\tilde{J}_j + 2\alpha\tilde{L}_j, \tilde{\Xi}_{33} = -I, \tilde{\Psi}_{11} = AZ + ZA^T + 2\alpha Z + \tilde{S} + \tilde{T}, \tilde{\Psi}_{12} = -BKCZ - \tilde{T} + \tilde{Y}^T, \tilde{\Psi}_{22} = 2\tilde{T} - \tilde{Y} - \tilde{Y}^T, \tilde{T} = ZTZ,$$
$$\tilde{\Psi}_{23} = \tilde{Y}^T - \tilde{T}, \tilde{\chi}_i = [-BK_1C_1Z, \cdots, -BK_NC_NZ], \tilde{\Psi}_{33} = \tilde{T} - \tilde{S}, \tilde{\Psi}_{14} = \tilde{\chi}_i, \tilde{\Psi}_{44} = \tilde{\phi}_i, \tilde{\Psi}_{55} = -\gamma^2 I, \tilde{\Psi}_{13} = -\tilde{Y}^T, \tilde{S} = ZSZ,$$
$$\tilde{\Xi}_{31} = [CZ, 0, \cdots, 0], \tilde{\Xi}_{22} = diag\{-Z\tilde{T}^{-1}Z, -H_1^{-1}, \cdots, -H_N^{-1}\}, \tilde{\Xi}_{21}^T = [\tau_M F_i, \sqrt{\tau_M}C_1F_1, \cdots, \sqrt{\tau_M}C_NF_N], \tilde{Y} = ZYZ.$$

Proof: Let $Z = O^{-1}, \tilde{J}_i = ZJ_iZ, \tilde{L}_i = ZL_iZ, i = 1, \cdots, N$. Pre- and post-multiplying Equation (15) and Equation (16) with

$$diag\{Z, \cdots, Z, I, T^{-1}, H_1^{-1}, \cdots, H_N^{-1}, I, Z\}$$

Table 1. Configuration of three-area power systems

Parameter	$T_{ch}(s)$	$T_g(s)$	R	D	β	$M(s)$
*Area*1	0.30	0.37	0.05	1.0	$\frac{1}{R_1}+D_1$	10
*Area*2	0.17	0.40	0.05	1.5	$\frac{1}{R_2}+D_2$	10
*Area*3	0.20	0.35	0.05	1.8	$\frac{1}{R_3}+D_3$	12
$T_{12}=0.20, T_{13}=0.12, T_{23}=0.25(pu/rad)$						

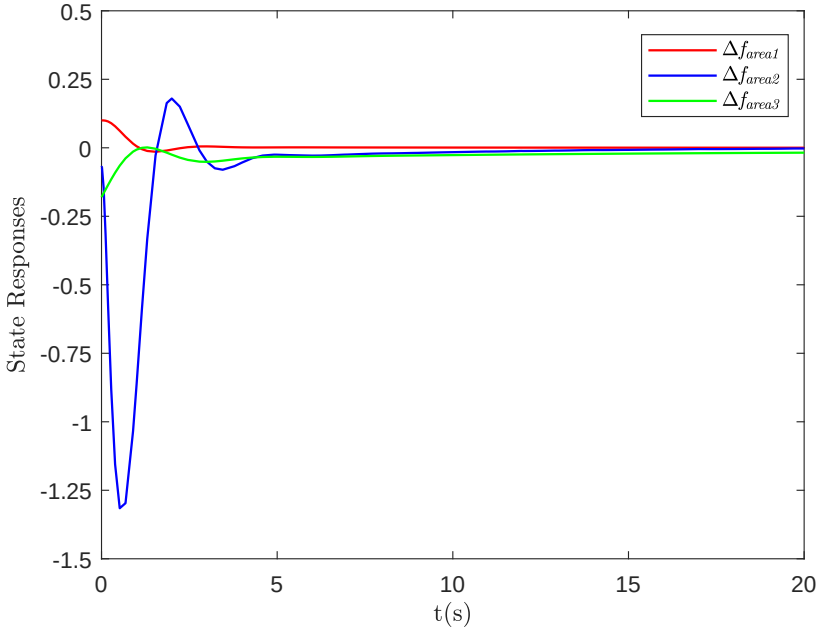


Figure 2. State responses of case 1.

and their transposes, respectively. For the nonlinear terms $-Z\tilde{T}_r^{-1}Z, r = 1,2$ and $-H_i^{-1}$, using inequalities $-Z\tilde{T}_r^{-1}Z \leq \rho^2\tilde{T}_r - 2\rho Z$ and the cone complementary linearization algorithm in [27], one can obtain *Equation* (32)*. By solving the minimization problem Equation (31), system Equation (12) is exponentially stable with a prescribed H_∞ performance γ and $K = F\Omega^{-1}$. □

3.4. Controller design under the RR scheduling scheme Equation (7)

Similar to Theorem 5.2 in [29], we establish Theorem 4.

Theorem 4 *Under the RR scheduling scheme Equation (7), for given matrices A, B, C, F, and scalars $\tau_M, \alpha > 0, \gamma > 0, \zeta > 0$, system Equation (12) is exponentially stable with a prescribed H_∞ performance γ, if there exist real matrices $Z > 0, \tilde{S} > 0, \tilde{T} > 0, \tilde{L}_i > 0, i = 1, \cdots, N$ and $\tilde{Y}, \Omega, F$ with appropriate dimensions such that Equation (31) is solvable with $J_i = \frac{L_i}{N-1}$, where H_i is given by Equation (30). The controller gain is given by $K = F\Omega^{-1}$.*

Proof: The detailed derivation process can refer to [29] and the proof of Theorem 3, which is omitted here. □

4. AN ILLUSTRATIVE EXAMPLE

In the following, we use a three-area power system [23,24,28] interconnected by the shared communication network to demonstrate the effectiveness of main results. Parameters are listed in Table 1.

Case 1: Stability of the studied system under TOD scheduling scheme Equation (6) and $\omega(t) = 0$.

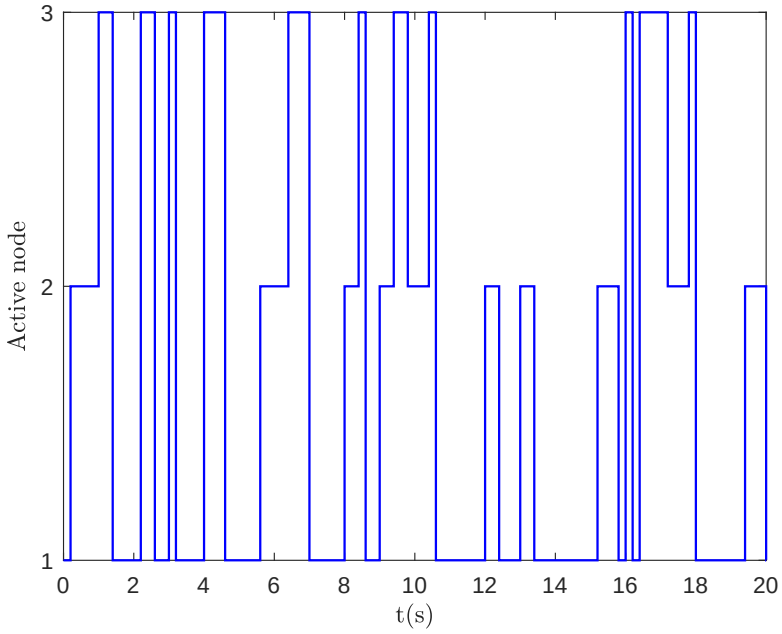


Figure 3. The switching behavior of active nodes of case 1.

Assume parameters are chosen as $\tau_M = 0.1s, \gamma = 10, \rho = 0.63$. We apply Theorem 3 which yields $K_1 = [0.2393, 0.0224]$, $K_2 = [0.2379, -0.1352]$, $K_3 = [0.2324, 0.1965]$.

Correspondingly, state responses and the switching behavior of active nodes are illustrated by Figure 2 and 3, respectively. Clearly, the designed dynamics output feedback controllers can stabilize the three-area power system under the TOD scheduling scheme Equation (6) in the absence of disturbance.

Case 2: Stability of the studied system under RR scheduling scheme Equation (7) and $\omega(t) = 0$.

To compare with the TOD protocol Equation (6), we use the same parameters as case 1. We apply Theorem 4 which yields $K_1 = [0.1876, 0.0184]$, $K_2 = [0.1899, -0.1596]$, $K_3 = [0.1824, 0.1498]$. Then, state responses and the switching behavior of active nodes are depicted in Figure 4 and 5. From Figure 2-5, one can obtain that the system can be stable under both scheduling protocols Equation (6), Equation (7). Compared with RR protocol, TOD protocol can achieve dynamic scheduling, which makes the control process more efficient.

Case 3: Stability of the studied system under TOD scheduling scheme Equation (6) and $\omega(t) \neq 0$.

According to Theorem 1-4, TOD protocol can degrade into RR protocol under certain conditions. In this case, we take the TOD protocol Equation (6) as an example to verify the anti-disturbance performance of the studied system. Under the disturbance $\omega(t)$ shown in Figure 6, we obtain controller gains $K_1 = [0.4005, -0.1702]$, $K_2 = [0.4839, -0.2094]$, $K_3 = [0.4914, -0.4009]$. State responses of the system are depicted in Figure 7. The designed decentralized controllers under the TOD protocol Equation (6) can ensure system stability with load disturbances.

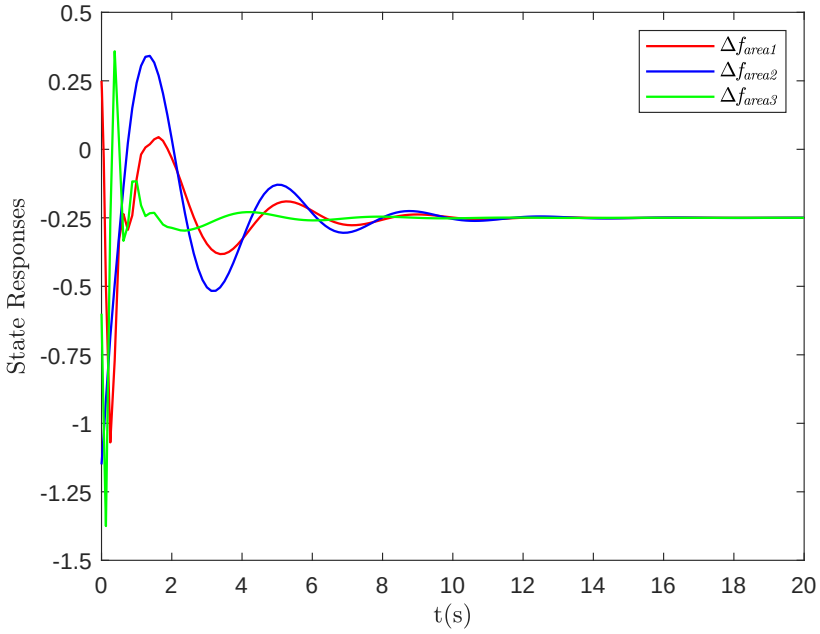


Figure 4. State responses of case 2.

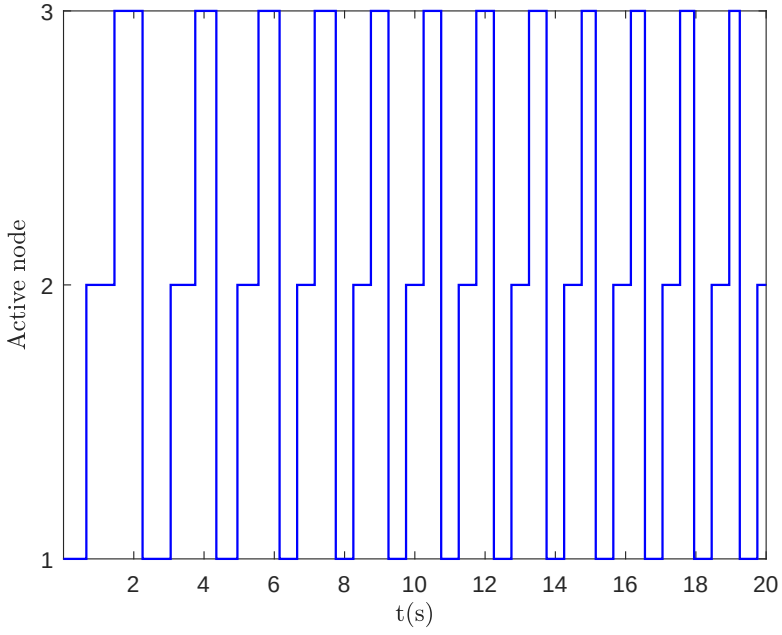


Figure 5. The switching behavior of active nodes of case 2.

5. CONCLUSIONS

This paper has designed H_∞ output feedback decentralized load frequency controllers for multi-area power systems under the shared but band-limited network. TOD and RR protocols have been used to improve communication efficiency. Since transmission priority has been well designed under scheduling protocols, a higher control performance has been obtained. Then, decentralized output feedback controllers with the prescribed

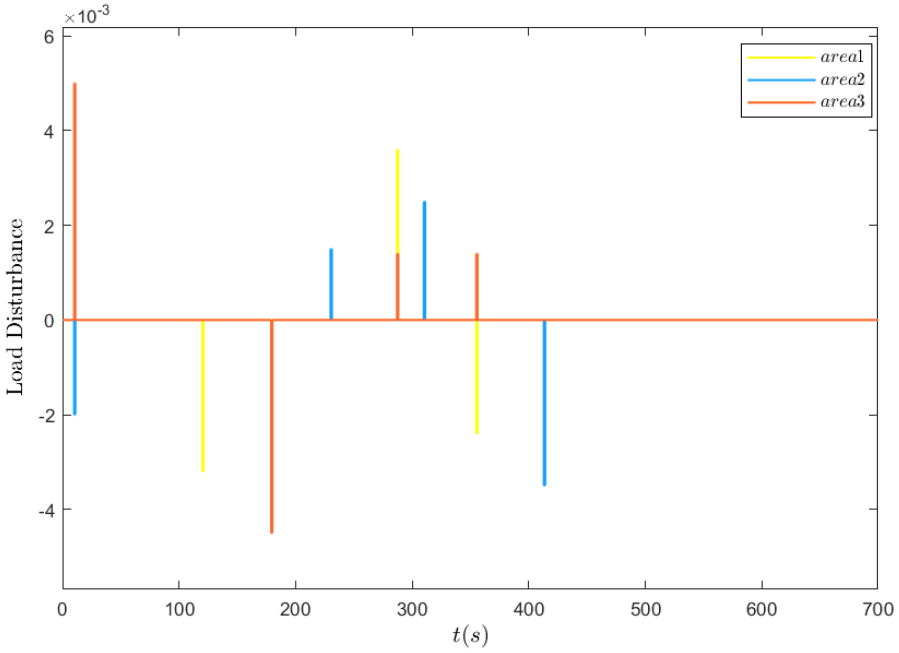


Figure 6. Disturbance of case 3.

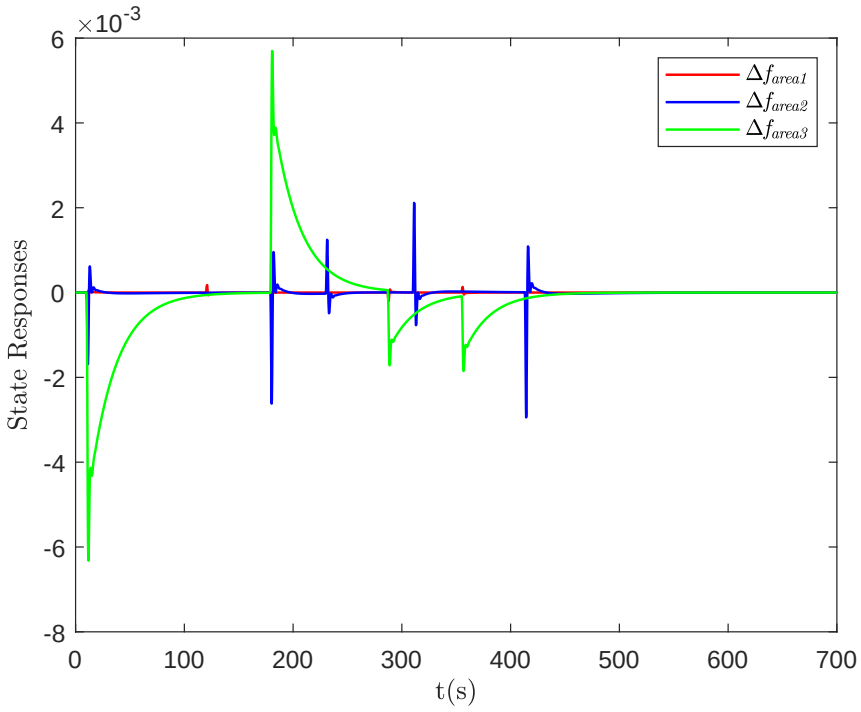


Figure 7. State responses of case 3.

performance have been designed in different areas. Finally, an example has verified the effectiveness of main results.

DECLARATIONS

Authors' contributions

All authors contributed equally.

Financial support and sponsorship

This work was supported in part by the National Natural Science Foundation of China under Grant 62173218, and the International Corporation Project of Shanghai Science and Technology Commission under Grant 21190780300.

Availability of data and materials

Not applicable.

Conflicts of interest

All authors declared that there are no conflicts of interest.

Ethical approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Copyright

© The Author(s) 2022.

REFERENCES

1. Mohandes B, El Moursi MS, Hatziargyriou N, El Khatib S. A review of power system flexibility with high penetration of renewables. *IEEE Trans Power Syst* 2019;34:3140–55. DOI
2. Ranjan M, Shankar R. A literature survey on load frequency control considering renewable energy integration in power system: Recent trends and future prospects. *J Energy Stor* 2022;45:103717. DOI
3. Peng C, Li F. A survey on recent advances in event-triggered communication and control. *Inform Sci* 2018;457:113–25. DOI
4. Zhou Q, Shahidehpour M, Paaso A, et al. Distributed control and communication strategies in networked microgrids. *IEEE Communic Surv & Tutori* 2020;22:2586–633. DOI
5. Shang Y. Group consensus of multi-agent systems in directed networks with noises and time delays. *Int J Syst Sci* 2015;46:2481–92. DOI
6. Alhelou HH, Hamedani-Golshan ME, Zamani R, Heydarian-Forushani E, Siano P. Challenges and opportunities of load frequency control in conventional, modern and future smart power systems: a comprehensive review. *Energies* 2018;11:2497. DOI
7. Yan Z, Xu Y. Data-driven load frequency control for stochastic power systems: a deep reinforcement learning method with continuous action search. *IEEE Trans Power Syst* 2018;34:1653–56. DOI
8. Peng C, Wu J, Tian E. Stochastic event-triggered H_∞ control for networked systems under denial of service attacks. *IEEE Trans Syst, Man, Cybern: Systems* 2021. DOI
9. Shangguan XC, He Y, Zhang CK, et al. Control performance standards-oriented event-triggered load frequency control for power systems under limited communication bandwidth. *IEEE Trans Contr Syst Technol* 2021;30:860–68. DOI
10. Tian E, Peng C. Memory-based event-triggering H_∞ load frequency control for power systems under deception attacks. *IEEE Trans Cybern* 2020;50:4610–18. DOI
11. Peng C, Sun H, Yang M, Wang YL. A survey on security communication and control for smart grids under malicious cyber attacks. *IEEE Trans Syst, Man, Cybern: Systems* 2019;49:1554–69. DOI
12. Wang D, Chen F, Meng B, Hu X, Wang J. Event-based secure H_∞ load frequency control for delayed power systems subject to deception attacks. *Appl Mathem Comput* 2021;394:125788. DOI
13. Wu J, Peng C, Yang H, Wang YL. Recent advances in event-triggered security control of networked systems: a survey. *Int J Syst Sci* 2022;0:1–20. DOI
14. Freirich D, Fridman E. Decentralized networked control of systems with local networks: a time-delay approach. *Automatica* 2016;69:201–9. DOI
15. Liu K, Fridman E, Hetel L, Richard JP. Sampled-data stabilization via round-robin scheduling: a direct Lyapunov-Krasovskii approach. *IFAC Proceed Vol* 2011;44:1459–64. DOI
16. Ding D, Wang Z, Han QL, Wei G. Neural-network-based output-feedback control under round-robin scheduling protocols. *IEEE trans*

cybern 2018;49:2372–84. DOI
17. Zou L, Wang Z, Han QL, Zhou D. Full information estimation for time-varying systems subject to Round-Robin scheduling: A recursive filter approach. *IEEE Trans Syst, Man, Cybern: Systems* 2019;51:1904–16. DOI
18. Zhang XM, Han QL, Ge X, et al. Networked control systems: a survey of trends and techniques. *IEEE/CAA J Autom Sinica* 2019;7:1–17. DOI
19. Liu K, Fridman E, Hetel L. Network-based control via a novel analysis of hybrid systems with time-varying delays. In: *2012 IEEE 51st IEEE Confer Decis Contr (CDC)*. IEEE; 2012. pp. 3886–91. DOI
20. Zhang J, Peng C, Xie X, Yue D. Output feedback stabilization of networked control systems under a stochastic scheduling protocol. *IEEE trans cybern* 2019;50:2851–60. DOI
21. Liu K, Fridman E, Johansson KH. Networked control with stochastic scheduling. *IEEE Trans Autom Contr* 2015;60:3071–76. DOI
22. Zhang J, Peng C. Networked H_∞ filtering under a weighted TOD protocol. *Automatica* 2019;107:333–41. DOI
23. Peng C, Zhang J, Yan H. Adaptive event-triggering H_∞ load frequency control for network-based power systems. *IEEE Trans Industr Electr* 2017;65:1685–94. DOI
24. Sun H, Peng C, Yue D, Wang YL, Zhang T. Resilient load frequency control of cyber-physical power systems under QoS-dependent event-triggered communication. *IEEE Trans Syst, Man, Cybernet: Systems* 2020;51:2113–22. DOI
25. Shang Y. Median-based resilient consensus over time-varying random networks. *IEEE Trans Circ Syst II: Express Briefs* 2021;69:1203–7. DOI
26. Seuret A, Gouaisbaut F. Wirtinger-based integral inequality: application to time-delay systems. *Automatica* 2013;49:2860–66. DOI
27. El Ghaoui L, Oustry F, AitRami M. A cone complementarity linearization algorithm for static output-feedback and related problems. *IEEE trans autom contr* 1997;42:1171–76. DOI
28. Peng C, Li J, Fei M. Resilient Event-Triggering H_∞ load frequency control for multi-area power systems with energy-limited DoS attacks. *IEEE Trans Power Syst* 2016;32:4110–18. DOI
29. Liu K, Fridman E, Xia Y. Networked control under communication constraints. Springer; 2020.

AUTHOR INSTRUCTIONS

1. Submission Overview

Before you decide to publish with *Intelligence & Robotics* (*IR*), please read the following items carefully and make sure that you are well aware of Editorial Policies and the following requirements.

1.1 Topic Suitability

The topic of the manuscript must fit the scope of the journal. Please refer to Aims and Scope for more information.

1.2 Open Access and Copyright

The journal adopts Gold Open Access publishing model and distributes content under the Creative Commons Attribution 4.0 International License. Copyright is retained by authors. Please make sure that you are well aware of these policies.

1.3 Publication Fees

Before December 31, 2024, there are no article processing charges for papers accepted for publication after peer review. OAE subsidizes and helps authors publish their manuscripts totally free. For more details, please refer to OAE Publication Fees.

1.4 Language Editing

All submissions are required to be presented clearly and cohesively in good English. Authors whose first language is not English are advised to have their manuscripts checked or edited by a native English speaker before submission to ensure the high quality of expression. A well-organized manuscript in good English would make the peer review even the whole Editorial handling more smoothly and efficiently.

If needed, authors are recommended to consider the language editing services provided by Charlesworth to ensure that the manuscript is written in correct scientific English before submission. Authors who publish with OAE journals enjoy a special discount for the services of Charlesworth via the following two ways.

Submit your manuscripts directly at http://www.charlesworthauthorservices.com/~OAE;

Open the link http://www.charlesworthauthorservices.com/, and enter Promotion Code “OAE” when you submit.

1.5 Work Funded by the National Institutes of Health

If an accepted manuscript was funded by National Institutes of Health (NIH), the author may inform editors of the NIH funding number. The editors are able to deposit the paper to the NIH Manuscript Submission System on behalf of the author.

2. Submission Preparation

2.1 Cover Letter

A cover letter is required to be submitted accompanying each manuscript. Here is a guideline of a cover letter for authors’ consideration:

List the highlights of the current manuscript and no more than 5 short sentences;

All authors have read the final manuscript, have approved the submission to the journal, and have accepted full responsibilities pertaining to the manuscript’s delivery and contents;

Clearly state that the manuscript is an original work on its own merit, that it has not been previously published in whole or in part, and that it is not being considered for publication elsewhere;

No materials are reproduced from another source (if there is material in your manuscript that has been reproduced from another source, please state whether you have obtained permission from the copyright holder to use them);

Conflicts of interest statement;

If the manuscript is contributed to a Special Issue, please also mention it in the cover letter;

If the manuscript was presented partly or entirely in a conference, the author should clearly state the background information of the event, including the conference name, time, and place in the cover letter.

2.2 Types of Manuscripts

There is no restriction on the length of manuscripts, number of figures, tables and references, provided that the manuscript is concise and comprehensive. The journal publishes Research Article, Review, Technical Note, etc. For more details about paper type, please refer to the following table.

Manuscript Type	Definition	Abstract	Keywords	Main Text Structure
Research Article	A Research Article is a seminal and insightful research study and showcases that often involves modern techniques or methodologies. Authors should justify that their work is of novel findings.	The abstract should state briefly the purpose of the research, the principal results and major conclusions. No more than 250 words.	3-8 keywords	The main content should include four sections: Introduction, Methods, Results and Discussion.
Review	A Review should be an authoritative, well balanced, and critical survey of recent progress in an attractive or a fundamental research field.	Unstructured abstract. No more than 250 words.	3-8 keywords	The main text may consist of several sections with unfixed section titles. We suggest that the author include an "Introduction" section at the beginning, several sections with unfixed titles in the middle part, and a "Conclusions" section at the end.
Technical Note	A Technical Note is a short article giving a brief description of a specific development, technique, or procedure, or it may describe a modification of an existing technique, procedure or device applied in research.	Unstructured abstract. No more than 250 words.	3-8 keywords	/
Editorial	An Editorial is a short article describing news about the journal or opinions of senior Editors or the publisher.	None required	None required	/
Commentary	A Commentary is to provide comments on a newly published article or an alternative viewpoint on a certain topic.	Unstructured abstract. No more than 250 words.	3-8 keywords	/
Perspective	A Perspective provides personal points of view on the state-of-the-art of a specific area of knowledge and its future prospects.	Unstructured abstract. No more than 250 words.	3-8 keywords	/

2.3 Manuscript Structure

2.3.1 Front Matter

2.3.1.1 Title

The title of the manuscript should be concise, specific and relevant, with no more than 16 words if possible.

2.3.1.2 Authors and Affiliations

Authors' full names should be listed. The initials of middle names can be provided. The affiliations and email addresses for all authors should be listed. At least one author should be designated as the corresponding author. In addition, corresponding authors are suggested to provide their Open Researcher and Contributor ID upon submission. Please note that any change to authorship is not allowed after manuscript acceptance. The authors' affiliations should be provided in this format: department, institution, city, postcode, country.

2.3.1.3 Abstract

The abstract should be a single paragraph with word limitation and specific structure requirements (for more details please refer to Types of Manuscripts). It usually describes the main objective(s) of the study, explains how the study was done, including any model organisms used, without methodological detail, and summarizes the most important results and their significance. The abstract must be an objective representation of the study: it is not allowed to contain results that are not presented and substantiated in the manuscript, or exaggerate the main conclusions. Citations should not be included in the abstract.

2.3.1.4 Graphical Abstract

The graphical abstract is essential as this can catch first view of your publication by readers. We recommend you submit an eye-catching figure. It should summarize the content of the article in a concise graphical form. It is recommended to use it because this can make online articles get more attention.

The graphical abstract should be submitted as a separate document in the online submission system. Please provide an image with a minimum of 531 × 1328 pixels (h × w) or proportionally more. The image should be readable at a size of 5 cm × 13 cm using a regular screen resolution of 96 dpi. Preferred file types: TIFF, PSD, AI, JPEG, and EPS files.

2.3.1.5 Keywords

Three to eight keywords should be provided, which are specific to the article, yet reasonably common within the subject discipline.

2.3.2 Main Text

Manuscripts of different types are structured with different sections of content. Please refer to Types of Manuscripts to make sure which sections should be included in the manuscripts.

2.3.2.1 Introduction

The introduction should contain background that puts the manuscript into context, allow readers to understand why the study is important, include a brief review of key literature, and conclude with a brief statement of the overall aim of the work and a comment about whether that aim was achieved. Relevant controversies or disagreements in the field should be introduced as well.

2.3.2.2 Methods

The methods should contain sufficient details to allow others to fully replicate the study. New methods and protocols should be described in detail while well-established methods can be briefly described or appropriately cited. Statistical terms, abbreviations, and all symbols used should be defined clearly. Protocol documents for clinical trials, observational studies, and other non-laboratory investigations may be uploaded as supplementary materials.

2.3.2.3 Results

This section contains the findings of the study. Results of statistical analysis should also be included either as text or as tables or figures if appropriate. Authors should emphasize and summarize only the most important observations. Data on all primary and secondary outcomes identified in the section Methods should also be provided. Extra or supplementary materials and technical details can be placed in supplementary documents.

2.3.2.4 Discussion

This section should discuss the implications of the findings in context of existing research and highlight limitations of the study. Future research directions may also be mentioned.

2.3.2.5 Conclusion

It should state clearly the main conclusions and include the explanation of their relevance or importance to the field.

2.3.3 Back Matter

2.3.3.1 Acknowledgments

Anyone who contributed towards the article but does not meet the criteria for authorship, including those who provided professional writing services or materials, should be acknowledged. Authors should obtain permission to acknowledge from all those mentioned in the Acknowledgments section. This section is not added if the author does not have anyone to acknowledge.

2.3.3.2 Authors' Contributions

Each author is expected to have made substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data, or the creation of new software used in the work, or have drafted the work or substantively revised it.

Please use Surname and Initial of Forename to refer to an author's contribution. For example: made substantial contributions to conception and design of the study and performed data analysis and interpretation: Salas H, Castaneda WV; performed data acquisition, as well as providing administrative, technical, and material support: Castillo N, Young V.

If an article is single-authored, please include "The author contributed solely to the article." in this section.

2.3.3.3 Availability of Data and Materials

In order to maintain the integrity, transparency and reproducibility of research records, authors should include this section in their manuscripts, detailing where the data supporting their findings can be found. Data can be deposited into data repositories or published as supplementary information in the journal. Authors who cannot share their data should state that the data will not be shared and explain it. If a manuscript does not involve such issues, please state "Not applicable." in this section.

2.3.3.4 Financial Support and Sponsorship

All sources of funding for the study reported should be declared. The role of the funding body in the experiment design, collection, analysis and interpretation of data, and writing of the manuscript should be declared. Any relevant grant numbers and the link of funder's website should be provided if any. If the study is not involved with this issue, state "None." in this section.

2.3.3.5 Conflicts of Interest

Authors must declare any potential conflicts of interest that may be perceived as inappropriately influencing the representation or interpretation of reported research results. If there are no conflicts of interest, please state "All authors declared that there are no conflicts of interest." in this section. Some authors may be bound by confidentiality agreements. In such cases, in place of itemized disclosures, we will require authors to state "All authors declared that they are bound by confidentiality agreements that prevent them from disclosing their conflicts of interest in this work.". If authors are unsure whether conflicts of interest exist, please refer to the "Conflicts of Interest" of *IR* Editorial Policies for a full explanation.

2.3.3.6 Ethical Approval and Consent to Participate

Research involving human subjects, human material or human data must be performed in accordance with the Declaration of Helsinki and approved by an appropriate ethics committee. An informed consent to participate in the study should also be obtained from participants, or their parents or legal guardians for children under 16. A statement detailing the name of the ethics committee (including the reference number where appropriate) and the informed consent obtained must appear in the manuscripts reporting such research.

Studies involving animals and cell lines must include a statement on ethical approval. More information is available at Editorial Policies.

If the manuscript does not involve such issue, please state "Not applicable." in this section.

2.3.3.7 Consent for Publication

Manuscripts containing individual details, images or videos, must obtain consent for publication from that person, or in the case of children, their parents or legal guardians. If the person has died, consent for publication must be obtained from the next of kin of the participant. Manuscripts must include a statement that written informed consent for publication was obtained. Authors do not have to submit such content accompanying the manuscript. However, these documents must be available if requested. If the manuscript does not involve this issue, state "Not applicable." in this section.

2.3.3.8 Copyright

Authors retain copyright of their works through a Creative Commons Attribution 4.0 International License that clearly states how readers can copy, distribute, and use their attributed research, free of charge. A declaration "© The Author(s) 2022." will be added to each article. Authors are required to sign License to Publish before formal publication.

2.3.3.9 References

References should be numbered in order of appearance at the end of manuscripts. In the text, reference numbers should be placed in square brackets and the corresponding references are cited thereafter. If the number of authors is less than or equal to six, we require to list all authors' names. If the number of authors is more than six, only the first three authors' names are required to be listed in the references, other authors' names should be omitted and replaced with "et al.". Abbreviations of the journals should be provided on the basis of Index Medicus. Information from manuscripts accepted but not published should be cited in the text as "Unpublished material" with written permission from the source.

References should be described as follows, depending on the types of works:

Types	**Examples**
Journal articles by individual authors	Weaver DL, Ashikaga T, Krag DN, et al. Effect of occult metastases on survival in node-negative breast cancer. *N Engl J Med* 2011;364:412-21. [PMID: 21247310 DOI: 10.1056/NEJMoa1008108]
Organization as author	Diabetes Prevention Program Research Group. Hypertension, insulin, and proinsulin in participants with impaired glucose tolerance. *Hypertension* 2002;40:679-86. [DOI: 10.1161/01.HYP.0000035706.28494.09]
Both personal authors and organization as author	Vallancien G, Emberton M, Harving N, van Moorselaar RJ; Alf-One Study Group. Sexual dysfunction in 1,274 European men suffering from lower urinary tract symptoms. *J Urol* 2003;169:2257-61. [PMID: 12771764 DOI: 10.1097/01.ju.0000067940.76090.73]
Journal articles not in English	Zhang X, Xiong H, Ji TY, Zhang YH, Wang Y. Case report of anti-N-methyl-D-aspartate receptor encephalitis in child. *J Appl Clin Pediatr* 2012;27:1903-7. (in Chinese)
Journal articles ahead of print	Odibo AO. Falling stillbirth and neonatal mortality rates in twin gestation: not a reason for complacency. BJOG 2018; Epub ahead of print [PMID: 30461178 DOI: 10.1111/1471-0528.15541]
Books	Sherlock S, Dooley J. Diseases of the liver and billiary system. 9th ed. Oxford: Blackwell Sci Pub; 1993. pp. 258-96.
Book chapters	Meltzer PS, Kallioniemi A, Trent JM. Chromosome alterations in human solid tumors. In: Vogelstein B, Kinzler KW, editors. The genetic basis of human cancer. New York: McGraw-Hill; 2002. pp. 93-113.
Online resource	FDA News Release. FDA approval brings first gene therapy to the United States. Available from: https://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm574058.htm. [Last accessed on 30 Oct 2017]

Conference proceedings	Harnden P, Joffe JK, Jones WG, Editors. Germ cell tumours V. Proceedings of the 5th Germ Cell Tumour Conference; 2001 Sep 13-15; Leeds, UK. New York: Springer; 2002.
Conference paper	Christensen S, Oppacher F. An analysis of Koza's computational effort statistic for genetic programming. In: Foster JA, Lutton E, Miller J, Ryan C, Tettamanzi AG, editors. Genetic programming. EuroGP 2002: Proceedings of the 5th European Conference on Genetic Programming; 2002 Apr 3-5; Kinsdale, Ireland. Berlin: Springer; 2002. pp. 182-91.
Unpublished material	Tian D, Araki H, Stahl E, Bergelson J, Kreitman M. Signature of balancing selection in Arabidopsis. *Proc Natl Acad Sci U S A*. Forthcoming 2002.

The journal also recommends that authors prepare references with a bibliography software package, such as EndNote to avoid typing mistakes and duplicated references.

2.3.3.10 Supplementary Materials

Additional data and information can be uploaded as Supplementary Materials to accompany the manuscripts. The supplementary materials will also be available to the referees as part of the peer-review process. Any file format is acceptable, such as data sheet (word, excel, csv, cdx, fasta, pdf or zip files), presentation (powerpoint, pdf or zip files), image (cdx, eps, jpeg, pdf, png or tiff), table (word, excel, csv or pdf), audio (mp3, wav or wma) or video (avi, divx, flv, mov, mp4, mpeg, mpg or wmv). All information should be clearly presented. Supplementary materials should be cited in the main text in numeric order (e.g., Supplementary Figure 1, Supplementary Figure 2, Supplementary Table 1, Supplementary Table 2, *etc*.). The style of supplementary figures or tables complies with the same requirements on figures or tables in main text. Videos and audios should be prepared in English, and limited to a size of 500 MB.

2.4 Manuscript Format

2.4.1 File Format

Manuscript files can be in DOC and DOCX formats and should not be locked or protected.
Manuscript prepared in LaTex must be collated into one ZIP folder (including all source files and images, so that the Editorial Office can recompile the submitted PDF).
When preparing manuscripts in different file formats, please use the corresponding Manuscript Templates.

2.4.2 Length

There are no restrictions on paper length, number of figures, or number of supporting documents. Authors are encouraged to present and discuss their findings concisely.

2.4.3 Language

Manuscripts must be written in English.

2.4.4 Multimedia Files

The journal supports manuscripts with multimedia files. The requirements are listed as follows:
Video or audio files are only acceptable in English. The presentation and introduction should be easy to understand. The frames should be clear, and the speech speed should be moderate;
A brief overview of the video or audio files should be given in the manuscript text;
The video or audio files should be limited to a size of up to 500 MB;
Please use professional software to produce high-quality video files, to facilitate acceptance and publication along with the submitted article. Upload the videos in mp4, wmv, or rm format (preferably mp4) and audio files in mp3 or wav format.

2.4.5 Figures

Figures should be cited in numeric order (e.g., Figure 1, Figure 2) and placed after the paragraph where it is first cited;
Figures can be submitted in format of TIFF, PSD, AI, EPS or JPEG, with resolution of 300-600 dpi;
Figure caption is placed under the Figure;
Diagrams with describing words (including, flow chart, coordinate diagram, bar chart, line chart, and scatter diagram, *etc*.) should be editable in word, excel or powerpoint format. Non-English information should be avoided;
Labels, numbers, letters, arrows, and symbols in figure should be clear, of uniform size, and contrast with the background;
Symbols, arrows, numbers, or letters used to identify parts of the illustrations must be identified and explained in the legend;
Internal scale (magnification) should be explained and the staining method in photomicrographs should be identified;
All non-standard abbreviations should be explained in the legend;
Permission for use of copyrighted materials from other sources, including re-published, adapted, modified, or partial figures and images from the internet, must be obtained. It is authors' responsibility to acquire the licenses, to follow any citation instruction requested by third-party rights holders, and cover any supplementary charges.

2.4.6 Tables

Tables should be cited in numeric order and placed after the paragraph where it is first cited;
The table caption should be placed above the table and labeled sequentially (e.g., Table 1, Table 2);
Tables should be provided in editable form like DOC or DOCX format (picture is not allowed);
Abbreviations and symbols used in table should be explained in footnote;
Explanatory matter should also be placed in footnotes;
Permission for use of copyrighted materials from other sources, including re-published, adapted, modified, or partial tables from the internet, must be obtained. It is authors' responsibility to acquire the licenses, to follow any citation instruction requested by third-party rights holders, and cover any supplementary charges.

2.4.7 Abbreviations

Abbreviations should be defined upon first appearance in the abstract, main text, and in figure or table captions and used consistently thereafter. Non-standard abbreviations are not allowed unless they appear at least three times in the text. Commonly-used abbreviations, such as DNA, RNA, ATP, *etc*., can be used directly without definition. Abbreviations in titles and keywords should be avoided, except for the ones which are widely used.

2.4.8 Italics

General italic words like *vs*., *et al*., *etc*., *in vivo*, *in vitro*; *t* test, *F* test, *U* test; related coefficient as *r*, sample number as *n*, and probability as *P*; names of genes; names of bacteria and biology species in Latin.

2.4.9 Units

SI Units should be used. Imperial, US customary and other units should be converted to SI units whenever possible. There is a space between the number and the unit (i.e., 23 mL). Hour, minute, second should be written as h, min, s.

2.4.10 Numbers

Numbers appearing at the beginning of sentences should be expressed in English. When there are two or more numbers in a paragraph, they should be expressed as Arabic numerals; when there is only one number in a paragraph, number < 10 should be expressed in English and number > 10 should be expressed as Arabic numerals. 12345678 should be written as 12,345,678.

2.4.11 Equations

Equations should be editable and not appear in a picture format. Authors are advised to use either the Microsoft Equation Editor or the MathType for display and inline equations.
Display equations should be numbered consecutively, using Arabic numbers in parentheses;
Inline equations should not be numbered, with the same/similar size font used for the main text.

2.4.12 Headings

In the main body of the paper, three different levels of headings may be used.
Level one headings: they should be in bold, and numbered using Arabic numbers, such as **1. INTRODUCTION**, and **2. METHODS**, with all letters capitalized;
Level two headings: they should be in bold and numbered after the level one heading, such as **2.1 Statistical analyses**, **2.2** ..., **2.3**..., *etc*., with the first letter capitalized;
Level three headings: they should be italicized, and numbered after the level two heading, such as *2.1.1 Data distributions*,and *2.1.2 outliers and linear regression*, with the first letter capitalized.

2.4.13 Text Layout

As the electronic submission will provide the basic material for typesetting, it is important to prepare papers in the general editorial style of the journal.
The font is Times New Roman;
The font size is 12pt;
Single column, 1.5× line spacing;
Insert one line break (one Return) before the heading and paragraph, if the heading and paragraph are adjacent, insert a line break before the heading only;
No special indentation;
Alignment is left end;
Insert consecutive line numbers;
For other details please refer to the Manuscript Templates.

2.5 Submission Link

Submit an article via https://oaemesas.com/login?JournalId=ir.

3. Publication Ethics Statement

OAE is a member of the Committee on Publication Ethics (COPE). We fully adhere to its Code of Conduct and to its Best Practice Guidelines.

The Editors of this journal enforce a rigorous peer-review process together with strict ethical policies and standards to guarantee to add high-quality scientific works to the field of scholarly publication. Unfortunately, cases of plagiarism, data falsification, image manipulation, inappropriate authorship credit, and the like, do arise. The Editors of *IR* take such publishing ethics issues very seriously and are trained to proceed in such cases with zero tolerance policy.

Authors wishing to publish their papers in *IR* must abide by the following:

The author(s) must disclose any possibility of a conflict of interest in the paper prior to submission;
The authors should declare that there is no academic misconduct in their manuscript in the cover letter;
Authors should accurately present their research findings and include an objective discussion of the significance of their findings;
Data and methods used in the research need to be presented in sufficient detail in the manuscript so that other researchers can replicate the work;
Authors should provide raw data if referees and the Editors of the journal request;
Simultaneous submission of manuscripts to more than one journal is not tolerated;
Republishing content that is not novel is not tolerated (for example, an English translation of a paper that is already published in another language will not be accepted);
The manuscript should not contain any information that has already been published. If you include already published figures or images, please get the necessary permission from the copyright holder to publish under the CC-BY license;
Plagiarism, data fabrication and image manipulation are not tolerated;
Plagiarism is not acceptable in OAE journals.

Plagiarism involves the inclusion of large sections of unaltered or minimally altered text from an existing source without appropriate and unambiguous attribution, and/or an attempt to misattribute original authorship regarding ideas or results, and copying text, images, or data from another source, even from your own publications, without giving credit to the source.

As to reusing the text that is copied from another source, it must be between quotation marks and the source must be cited. If a study's design or the manuscript's structure or language has been inspired by previous studies, these studies must be cited explicitly.

If plagiarism is detected during the peer-review process, the manuscript may be rejected. If plagiarism is detected after publication, we may publish a Correction or retract the paper.

Falsification is manipulating research materials, equipment, or processes, or changing or omitting data or results so that the findings are not accurately represented in the research record.

Image files must not be manipulated or adjusted in any way that could lead to misinterpretation of the information provided by the original image.

Irregular manipulation includes: introduction, enhancement, moving, or removing features from the original image; the grouping of images that should be presented separately, or modifying the contrast, brightness, or color balance to obscure, eliminate, or enhance some information.

If irregular image manipulation is identified and confirmed during the peer-review process, we may reject the manuscript. If irregular image manipulation is identified and confirmed after publication, we may publish a Correction or retract the paper.

OAE reserves the right to contact the authors' institution(s) to investigate possible publication misconduct if the Editors find conclusive evidence of misconduct before or after publication. OAE has a partnership with iThenticate, which is the most trusted similarity checker. It is used to analyze received manuscripts to avoid plagiarism to the greatest extent possible. When plagiarism becomes evident after publication, we will retract the original publication or require modifications, depending on the degree of plagiarism, context within the published article, and its impact on the overall integrity of the published study. Journal Editors will act under the relevant COPE guidelines.

4. Authorship

Authorship credit of *IR* should be solely based on substantial contributions to a published study, as specified in the following four criteria:

1. Substantial contributions to the conception or design of the work, or the acquisition, analysis, or interpretation of data for the work;
2. Drafting the work or revising it critically for important intellectual content;
3. Final approval of the version to be published;
4. Agreement to be accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

All those who meet these criteria should be identified as authors. Authors must specify their contributions in the section Authors' Contributions of their manuscripts. Contributors who do not meet all the four criteria (like only involved in acquisition of funding, general supervision of a research group, general administrative support, writing assistance, technical editing, language editing, proofreading, *etc*.) should be acknowledged in the section of Acknowledgement in the manuscript rather than being listed as authors.

If a large multiple-author group has conducted the work, the group ideally should decide who will be authors before the work starts and confirm authors before submission. All authors of the group named as authors must meet all the four criteria for authorship.

5. Reviewers Exclusions

You are welcome to exclude a limited number of researchers as potential Editors or reviewers of your manuscript. To ensure a fair and rigorous peer review process, we ask that you keep your exclusions to a maximum of three people. If you wish to exclude additional referees, please explain or justify your concerns—this information will be helpful for Editors when deciding whether to honor your request.

6. Editors and Journal Staff as Authors

Editorial independence is extremely important and OAE does not interfere with Editorial decisions. Editorial staff or Editors shall not be involved in processing their own academic work. Submissions authored by Editorial staff/Editors will be assigned to at least two independent outside reviewers. Decisions will be made by the Editor-in-Chief, including Special Issue papers. Journal staff are not involved in the processing of their own work submitted to any OAE journals.

7. Conflict of Interests

OAE journals require authors to declare any possible financial and/or non-financial conflicts of interest at the end of their manuscript and in the cover letter, as well as confirm this point when submitting their manuscript in the submission system. If no conflicts of interest exist, authors need to state "All authors declared that there are no conflicts of interest". We also recognize that some authors may be bound by confidentiality agreements, in which cases authors need to state "All authors declared that they are bound by confidentiality agreements that prevent them from disclosing their competing interests in this work".

8. Editorial Process

8.1. Pre-Check

New submissions are initially checked by the Managing Editor from the perspectives of originality, suitability, structure and formatting, conflicts of interest, background of authors, *etc*. Poorly prepared manuscripts may be rejected at this stage. If your manuscript does not meet one or more of these requirements, we will return it for further revisions.

Once your manuscript has passed the initial check, it will be assigned to the Assistant Editor, and then the Editor-in-Chief, or an Associate Editor in the case of a conflict of interest, will be notified of the submission and invited to review. Regarding Special Issue paper, after passing the initial check, the manuscript will be successively assigned to the Assistant Editor, and then to the Editor-in-Chief, or an Associate Editor in the case of conflict of interest for the Editor-in-Chief to review. The Editor-in-Chief, or the Associate Editor may reject manuscripts that they deem highly unlikely to pass peer review without further consultation. Once your manuscript has passed the Editorial assessment, the Associate Editor will start to organize peer-review.

All manuscripts submitted to *IR* are screened using CrossCheck powered by iThenticate to identify any plagiarized content. Your study must also meet all ethical requirements as outlined in our Editorial Policies. If the manuscript does not pass any of these checks, we may return it to you for further revisions or decline to consider your study for publication.

8.2. Peer Review

IR operates a single-blind review process, which means that reviewers know the names of authors, but the names of the reviewers are hidden from the authors. The scientific quality of the research described in the manuscript is assessed

by a minimum of two independent expert reviewers. The Editor-in-Chief is responsible for the final decision regarding acceptance or rejection of the manuscript.

All information contained in your manuscript and acquired during the review process will be held in the strictest confidence.

8.3. Decisions

Your research will be judged on scientific soundness only, not on its perceived impact as judged by Editors or referees. There are three possible decisions: Accept (your study satisfies all publication criteria), Invitation to Revise (more work is required to satisfy all criteria), and Reject (your study fails to satisfy key criteria and it is highly unlikely that further work can address its shortcomings). All of the following publication criteria must be fulfilled to enable your manuscript to be accepted for publication:

Originality
The study reports original research and conclusions.
Data availability
All data to support the conclusions either have been provided or are otherwise publicly available.
Statistics
All data have been analyzed through appropriate statistical tests and these are clearly defined.
Methods
The methods are described in sufficient detail to be replicated.
Citations
Previous work has been appropriately acknowledged.
Interpretation
The conclusions are a reasonable extension of the results.
Ethics
The study design, data presentation, and writing style comply with our Editorial Policies.

8.4. Revisions

Authors are required to submit the revised manuscript within one week if minor revision is recommended while two weeks if major revision recommended or one month if additional experiments are needed. If authors need more than one month to revise their manuscript, we usually require the authors to resubmit their paper. We request that a document of point-to-point response to all comments of reviewers and the Editor-in-Chief or the Associate Editor should be supplied along with the revised manuscript to allow quick assessment of your revised manuscript. This document should outline in detail how each of the comments was addressed in the revised manuscript or should provide a rebuttal to the criticism. Manuscripts may or may not be sent to reviewers after revision, dependent on whether the reviewer requested to see the revised version. Apart from in exceptional circumstances, *IR* only supports a round of major revision per manuscript.

9. Contact Us

Journal Contact

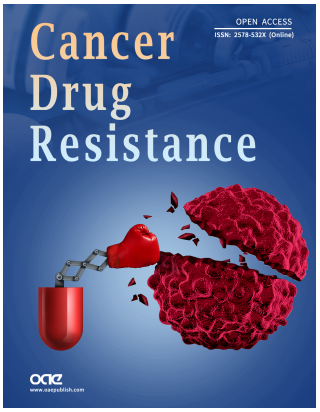
Intelligence & Robotics Editorial Office
Suite 1504, Plaza A, Xi'an National Digital Publishing Base,
No. 996 Tiangu 7th Road, Gaoxin District, Xi'an 710077, Shaanxi, China.

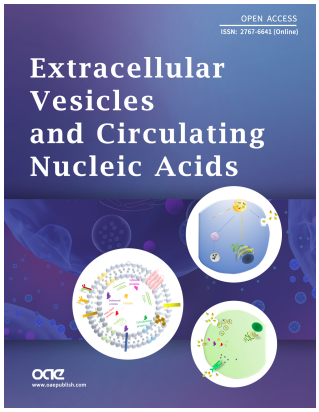
Managing Editor

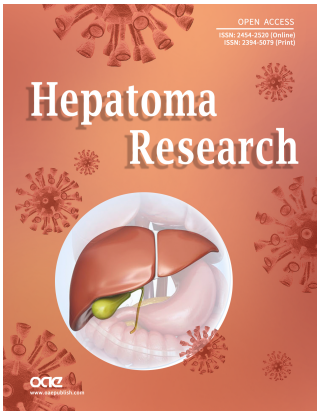
Lijun Jin
Email: editorial@intellrobot.com

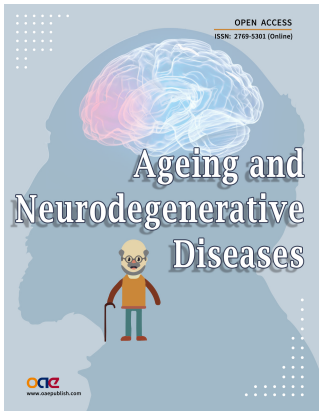
OAE Publishing Inc. (https:oaepublish.com) is a multidisciplinary open-access publishing company, founded in Los Angeles in 2015. Until now, OAE has been recognized by authoritative organizations in publishing industries,such as the *ORCID, COPE*, Scientific, Technical and Medical Publishers (STM), Crossref, and EASE.

As of June 2022, more than 1,200 outstanding scholars have joined OAE, who are from world-renowned universities and research institutions, including European Academy of Sciences, American Academy of Invention Sciences, Chinese Academy of Sciences, Royal Academy of Sciences of Belgium, British Academy of Medical Sciences, *etc*. There are more than 30 journals founded by OAE (https:oaepublish.com/about/journals), such as *Intelligence&Robotics*, *Journal of Materials Informatics, Complex Engineering Systems, Journal of Smart Environments and Green Computing*, and *Soft Science*, *etc.* Part of journals have been indexed by Scopus and CAS. We are currently working on database application including PubMed and ESCI. Up to June 2022, 3,154 articles have been published online, with 10,944,568 hits and 2,285,864 downloads. In the future, OAE Publishing Company will continue to found more quality journals with outstanding scholars, to promote the global academic development.









OAE Official Website