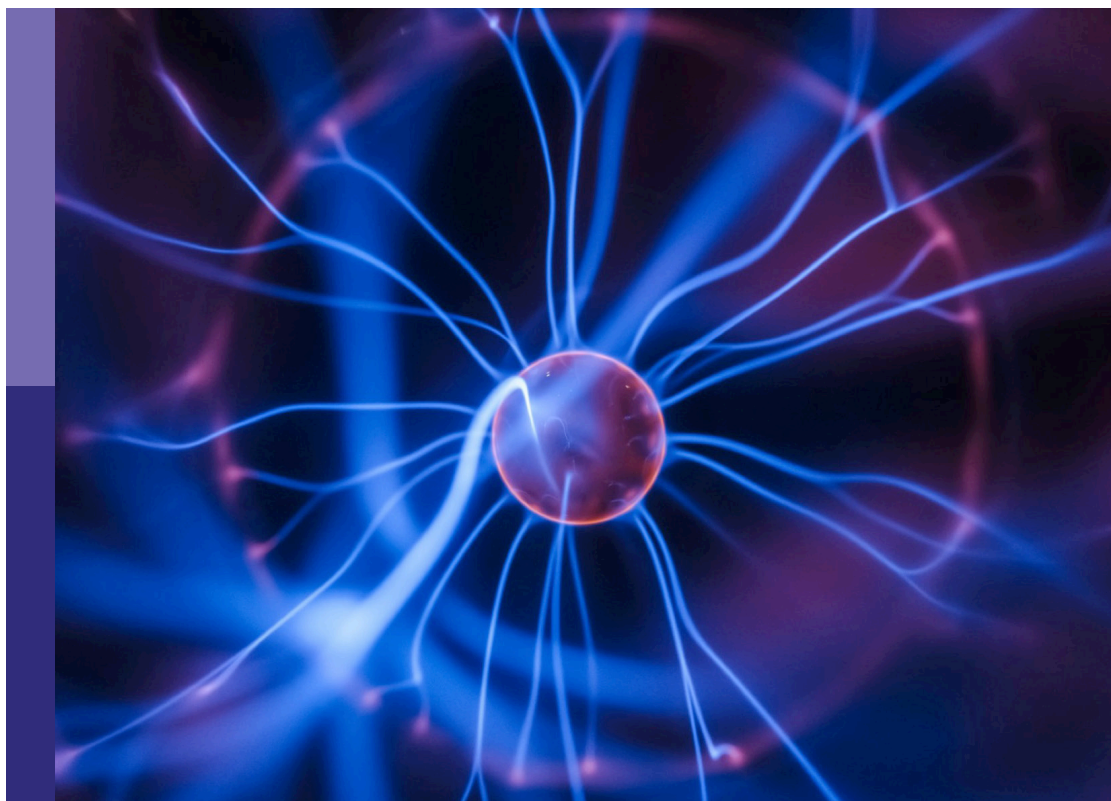# Real-world applications of game theory and optimization

Edited by
Dun Han, Jianbo Wang, Jianrong Wang and
Matjaž Perc

Published in
Frontiers in Physics

About Frontiers

Frontiers is more than just an open access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers journal series

The Frontiers journal series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the *Frontiers journal series* operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews. Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view. By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the *Frontiers journals series*: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area.

Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers editorial office: frontiersin.org/about/contact

Real-world applications of game theory and optimization

Topic editors

Dun Han — Jiangsu University, China

Jianbo Wang — Southwest Petroleum University, China

Jianrong Wang — Shanxi University, China

Matjaž Perc — University of Maribor, Slovenia

Table of contents

Editorial: Real-world applications of game theory and optimization

Dun Han[1]*, Jianrong Wang[2]*, Jianbo Wang[3]* and
Matjaž Perc[4,5,6,7]*

[1]School of Mathematical Sciences, Jiangsu University, Zhenjiang, China, [2]School of Mathematical
Sciences, Shanxi University, Taiyuan, China, [3]School of Computer Science and Software Engineering,
Southwest Petroleum University, Chengdu, China, [4]Faculty of Natural Sciences and Mathematics,
University of Maribor, Maribor, Slovenia, [5]Complexity Science Hub Vienna, Vienna, Austria, [6]Community
Healthcare Center Dr. Adolf Drolc Maribor, Maribor, Slovenia, [7]Department of Physics, Kyung Hee
University, Seoul, Republic of Korea

Editorial on the Research Topic
Real-world applications of game theory and optimization

Researching real-world applications of game theory and optimization is crucial due to its significant impact on decision-making, efficiency, and innovation across various sectors. Game theory provides valuable insights into strategic interactions among rational agents, aiding businesses in developing competitive strategies and assisting governments in policy design. Optimization techniques enhance operational efficiencies in industries such as logistics and manufacturing, leading to cost reductions and productivity gains. These fields are fundamental in technological advancements, particularly in artificial intelligence and machine learning, where they underpin many algorithms. Additionally, they play a vital role in addressing social and environmental challenges by optimizing resource management and promoting sustainable practices. In healthcare, they improve resource allocation and patient outcomes. By fostering interdisciplinary collaboration and theoretical advancements, research in these areas drives innovation and provides comprehensive solutions to complex problems. Ultimately, studying real-world applications of game theory and optimization not only enhances academic understanding but also delivers practical solutions that significantly benefit economic, technological, social, and environmental domains.

This Research Topic centers on the practical application of game theory and optimization methods to address complex challenges in real-world contexts. At its core, game theory provides a framework for analyzing strategic interactions among rational decision-makers, while optimization techniques seek the most favorable outcomes. These tools have proven to be powerful assets across a wide range of domains, from economics and computer science to social sciences and engineering. The goal of this Research Topic in Frontiers in Physics is to produce a comprehensive understanding of the real-world applications of game theory and optimization, highlighting their practical impact and potential for future use. It will provide valuable insights for professionals and researchers working in fields where these techniques can be applied and contribute to the body of knowledge in game theory and optimization.

Within this Research Topic, Wu et al. employ a game-theoretic approach to explore the dynamics of status transitions and the enhancement of employee performance in organizations. Zhang et al. utilize evolutionary game theory to examine the role transitions of employees in family businesses. To investigate the evolutionary game rules of strategic interactions between enterprises and employees during deviant

innovation, Zheng et al. construct a 2 × 2 asymmetric payoff matrix and use numerical simulations to demonstrate the influence of varying decision parameters and initial conditions on evolutionary outcomes. Sang et al. employ reinforcement learning algorithms to identify the optimal policy or equilibrium solution. Li et al. conduct an in-depth study on maritime area detection, proposing a cloud-edge cooperative-based scheme to address communication limitations by deploying edge computing nodes on the ship side of the gateway. Zhao and Yang establish a three-party evolutionary game model comprising an agricultural product data sharing platform, agricultural data providers, and agricultural data consumers. Zhou et al. propose a coupled disease-behavior model to describe the dynamic evolution of vaccination behavior during the spread of infectious diseases. Kan et al. explore a mixed updating strategy for vaccination decisions, where some individuals, termed intelligent agents, update their decisions based on a reinforcement learning strategy, while others, termed regular agents, use the Fermi function. Finally, Ma et al. investigate the decision-making behaviors of opinion leaders and netizens in the context of uncertain information dissemination, aiming to effectively manage online public opinion crises triggered by major sudden events.

Based on the contributions of these papers, it is evident that this Research Topic is highly valuable for understanding social systems. We hope that the theoretical models and practical applications presented in this research will inspire further exploration and development of real-world applications of game theory and optimization in social systems.

Author contributions

DH: Validation, Writing–original draft, Writing–review and editing. JrW: Validation, Writing–original draft, Writing–review and editing. JbW: Validation, Writing–original draft, Writing–review and editing. MP: Supervision, Validation, Writing–original draft, Writing–review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

Check for updates

Research on task offloading optimization strategies for vehicular networks based on game theory and deep reinforcement learning

Lei Wang[1], Wenjiang Zhou[2], Haitao Xu[1,3]*, Liang Li[1], Lei Cai[4] and Xianwei Zhou[1]

[1]School of Computer and Communication Engineering, University of Science and Technology Beijing, Beijing, China, [2]China Academy of Information and Communications Technology, Beijing, China, [3]Shunde Innovation School, University of Science and Technology Beijing, Foshan, China, [4]The North China Institute of Computing Technology, Beijing, China

With the continuous development of the 6G mobile network, computing-intensive and delay-sensitive onboard applications generate task data traffic more frequently. Particularly, when multiple intelligent agents are involved in tasks, limited computational resources cannot meet the new Quality of Service (QoS) requirements. To provide a satisfactory task offloading strategy, combining Multi-Access Edge Computing (MEC) with artificial intelligence has become a potential solution. In this context, we have proposed a task offloading decision mechanism (TODM) based on cooperative game and deep reinforcement learning (DRL). A joint optimization problem is presented to minimize both the overall task processing delay (OTPD) and overall task energy consumption (OTEC). The approach considers task vehicles (TaVs) and service vehicles (SeVs) as participants in a cooperative game, jointly devising offloading strategies to achieve resource optimization. Additionally, a proximate policy optimization (PPO) algorithm is designed to ensure robustness. Simulation experiments confirm the convergence of the proposed algorithm. Compared with benchmark algorithms, the presented scheme effectively reduces delay and energy consumption while ensuring task completion.

KEYWORDS

multi-access edge computing, cooperative game, task offloading, proximate policy optimization, deep reinforcement learning

1 Introduction

With the advent of the Internet of Things (IoT), many sensing devices have been deployed in networks. The data generated by these devices and related large-scale mobile applications are growing explosively [1]. In the context of IoT, the Internet of Vehicles (IoV) is a study hotspot. It uses IoV technology to provide services for vehicles through onboard processors [2,3]. However, task data also increase with a significant increase in the number of vehicles. The emergence of various computing-intensive tasks poses a significant challenge to the onboard computing capability of the vehicle itself Zhou et al. [4]. Multi-Access Edge Computing (MEC) is considered a feasible method to tackle this issue. MEC has significant advantages in addressing compute-intensive tasks in the IoV system. By moving
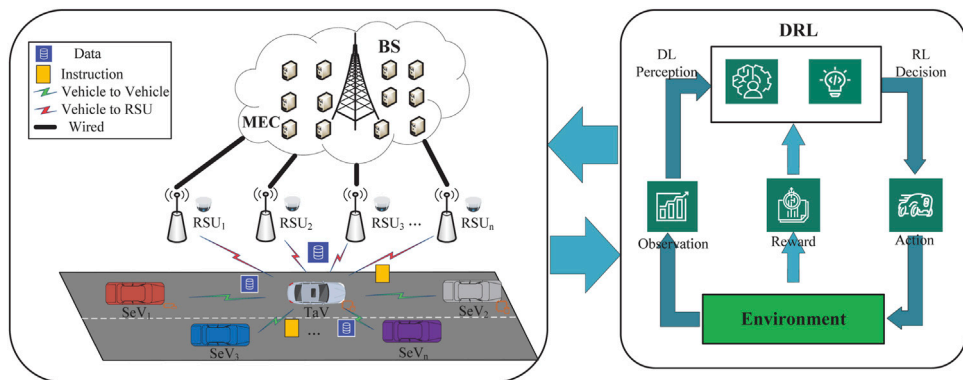
FIGURE 1
Intelligent architecture network of IoV.

TABLE 1 Parameter setting.

Parameter	Value
The transmission path loss index α	3.4
The transmission power of noise $\sigma_n^2/\sigma_{N+1}^2$	$10^{-13}w$
The wired transmission power from RSUs to MEC P'_{N+1}	2w
The height of RSUs H_{N+1}	20m
The distance from RSUs to the road center D_{N+1}	6m
The strength of vehicle coordinate transformation M_{tra}	1000 $cycles/bit$

computational and data processing functions to the network edge, it reduces task processing latency, enabling faster real-time decision-making, which is crucial for areas such as autonomous driving, traffic optimization, and intelligent traffic management. Moreover, MEC alleviates the burden on TaV, optimizes network load, and reduces energy consumption. MEC is considered a prevalent computing paradigm that has been widely studied to promote data processing efficiency, which can perform computation services closer to the data sources Porambage et al. [5].

Specifically, in the IoV system, tasks are offloaded to the service nodes (SNs) with computing power, and tasks are processed cooperatively to improve efficiency. The premise of task offloading is that jobs can be split into multiple subtasks and offloaded to SNs. Parked or moving vehicles, as idle resources, can provide specific computing and storage resources for task processing Sookhak et al. [6]. In addition to offloading tasks to the service vehicles (SeVs), task vehicles (TaVs) can also offload tasks to the MEC servers. The MEC servers coexist with the base station (BS) and connect the roadside units (RSUs) to provide services Xiao and Zhu [7]. In recent years, the issue of task offloading in the IoV system has received extensive research Zhou et al. [8], and task processing delay and energy consumption are essential indicators. It is challenging to minimize overall delay and energy consumption while completing the task Li et al. [9]. When the amount of task data is large, the task transmission delay is high, increasing the total task delay and energy consumption. To solve this problem, integrated

radar sensing and communication is a feasible solution. The integrated radar sensing and communication technology aims to reduce the task processing latency and energy consumption in the IoV, improving the efficiency and performance of task processing in IoV. By collecting data through radar sensing and sharing, instead of traditional data transmission, it reduces node waiting energy consumption and enhances the response speed of the IoV system. Its advantage lies in optimizing the overall performance of the IoV system, including perception of traffic data, improvement of communication quality, and increased accuracy of vehicle positioning, thereby enhancing the efficiency and safety of the entire IoV system. Game theory and optimization techniques provide technical support for it. In this study, a game theoretic approach was utilized to construct a game model, analyze the cooperative relationship between TaVs and SNs, and define the utility function for task offloading. This facilitated the development of an optimal task offloading decision strategy, encompassing task allocation and resource coordination. Optimization techniques were employed to achieve an optimal allocation of resources, including computing, storage, and communication resources, maximizing system utility while minimizing task processing delay and energy consumption.

Some scholars have conducted some studies on this issue. For example, in [10], a relatively practical IoV scenario was considered, and a matching game method was used to model the task allocation. The simulation results show that the input data transmission delay accounts for 73% of the total task processing time. In [11], the task is assigned to the MEC and the SeVs for processing. The results show that when the task size is 80 Mb, the input data transfer delay accounts for 50%. The delay in uploading data can significantly affect delay-sensitive applications. Therefore, several cars have an integrated radar system to sense the surrounding environmental data for local processing or assist connected vehicles in processing task data to ensure safe driving [12]. Furthermore, RSUs use radar to sense environmental data and use ecological data as input to reduce the transmission delay [13]. In summary, instruction transfer and environmental data sensing provide new possibilities for task offloading. For the issue of transmission delay, consider perceived environmental data and calculation instructions to reduce transmission delay [14].
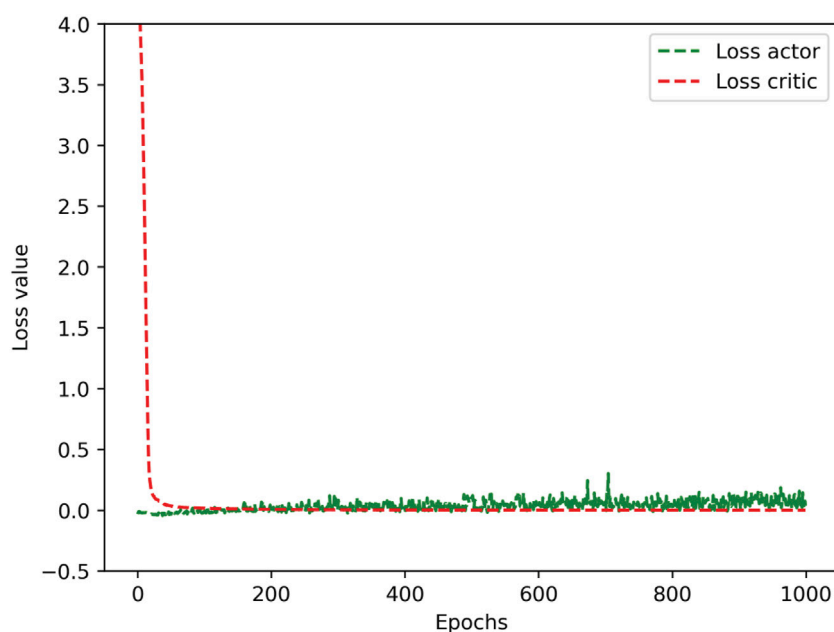
FIGURE 2
Training curve of PPOTR.

Energy is currently a major concern worldwide, and the increase in the number of IoV equipment will lead to increased energy demand and higher energy costs. Therefore, reducing energy consumption has become one of the issues that the IoV system needs to resolve [15]-[16]. To tackle this issue, Cesarano et al. designed a greedy heuristic algorithm to reduce the energy consumption of the task [17]. Some scholars applied minimizing of energy consumption and execution delay as the objective functions and reasonably selected the task offloading strategies [18]. In [19], the IoV system data transmission scheme adopts the deep Q-network (DQN) method to reduce transmission costs. Altogether, energy consumption is a key factor influencing the task offloading strategy. The focus of future study will be the proper selection of task offloading strategies to ensure delay and energy consumption.

For the aforementioned issues, many studies have adopted heuristic algorithms to solve them. For example, the author considers the reliability of task offloading in the IoV scenarios and uses heuristic algorithms to optimize the reliability [20]. Aiming at the issue of poor robustness of traditional heuristic algorithms for continuous state and action space in the IoV scenario [21], an offloading strategy-based method was studied to learn the optimal mapping from constant input state to discrete output and deal with continuous state space and action space scenarios. Although the aforementioned algorithm can solve the issue of the task offloading strategy in the IoV, the algorithm used has poor robustness in ensuring the reliability of data transmission [22].

DRL algorithms have significant advantages over heuristic algorithms. First, DRL algorithms can automatically learn and optimize decision strategies through large-scale data, eliminating the need for manual design of complex rules. Second, DRL

algorithms can handle high-dimensional and complex state and action spaces, making them suitable for solving complex real-world problems. Additionally, DRL algorithms have the ability to generalize learned knowledge to unseen environments, enabling more intelligent and flexible decision-making. Given the more significant potential and application value of policy-based deep reinforcement learning (DRL) [23], this paper discusses task offloading based on DRL. For sensitive applications with environmental data as input, we proposed a task offloading decision mechanism (TOMD) based on cooperative game and DRL. This paper is based on cooperative game theory, considered the overall task processing delay (OTPD) and overall task energy consumption (OTEC), and constructed a joint optimization issue. We transformed the joint optimization issue into a DRL issue and used the PPO algorithm to solve the issue. The main contributions are summarized as follows:

1. Considering dynamic wireless edge computing networks, a framework for joint task offloading is designed. On this basis, according to the wireless transmission requirements of SNs, combined with the game theory and communication function, a cooperative game and DRL-based TODM is proposed. The joint optimization issue is derived to minimize the delay and energy consumption.

2. DRL is more robust than the heuristic algorithm as it can make real-time online decisions. Therefore, combined with DRL, the designed joint optimization issues transformed into reinforcement learning (RL) issues. This paper develops an algorithm based on PPO to solve the aforementioned issues and theoretically analyze the algorithm's complexity.

3. Finally, we designed a simulation experiment to evaluate the algorithm's performance. The results show that the algorithm
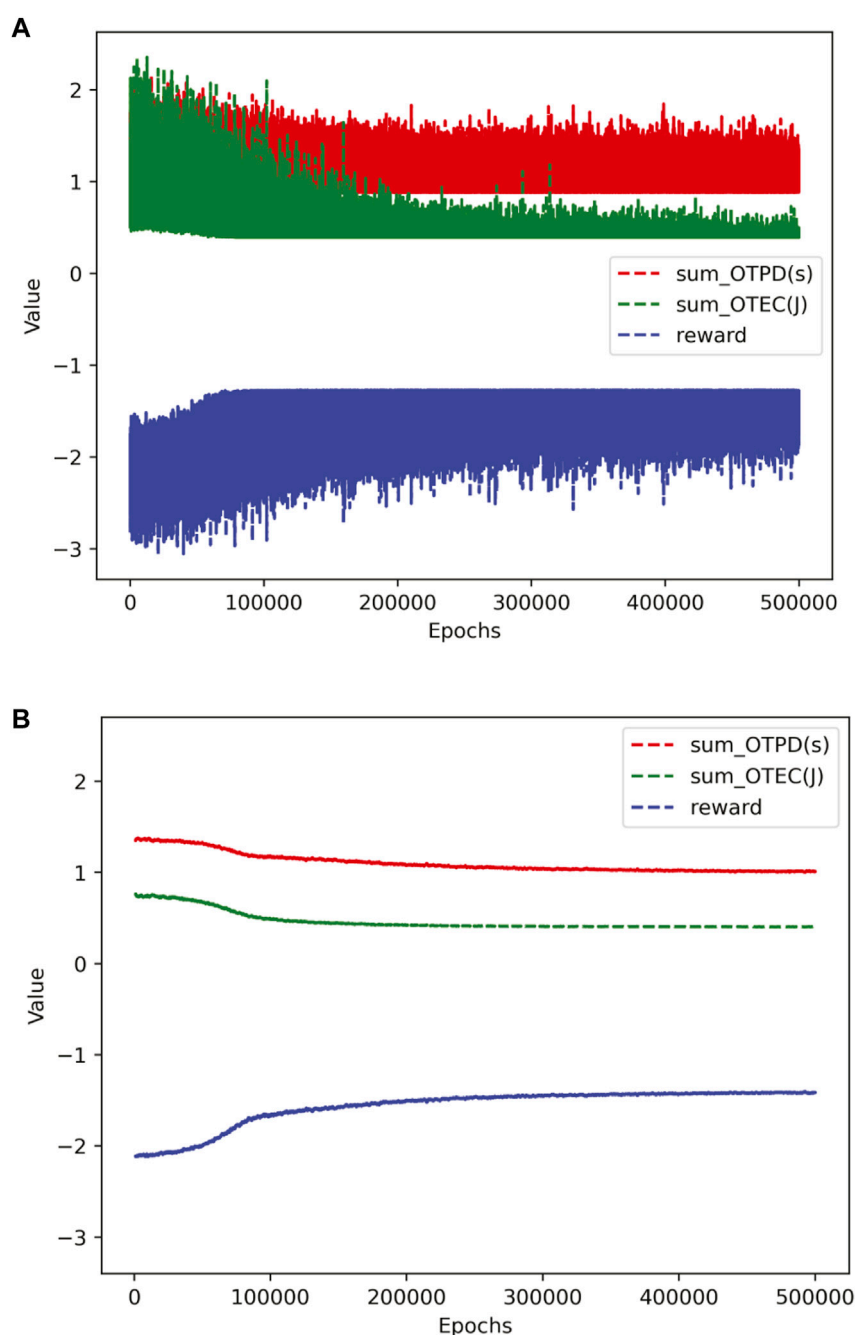
FIGURE 3
Convergence curve of PPOTR. **(A)** Unsmoothed convergence curve of PPOTR. **(B)** Smoothed convergence curve of PPOTR.

converges better than the soft actor-critic (SAC) algorithm, which can achieve the goal of a reasonable choice of the task offloading strategy. The proposed algorithm can reduce the task delay and energy consumption cost while improving the performance of the IoV system.

The remainder of this paper is arranged as follows: Section 2 presents relevant work. Section 3 presents the system model in detail, expounds on the task offloading mechanism TODM, and gives the issue formulation. Section 4 proposes a task offloading

algorithm based on DRL to solve the aforementioned issues. Section 5 proposes a simulation for evaluating the solution. Finally, Section 6 summarizes this paper.

2 Related works

This section summarizes the current study of the IoV, including the connected study of task offloading, radar sensing and communication, game theory, and DRL of the IoV edge intelligent system.
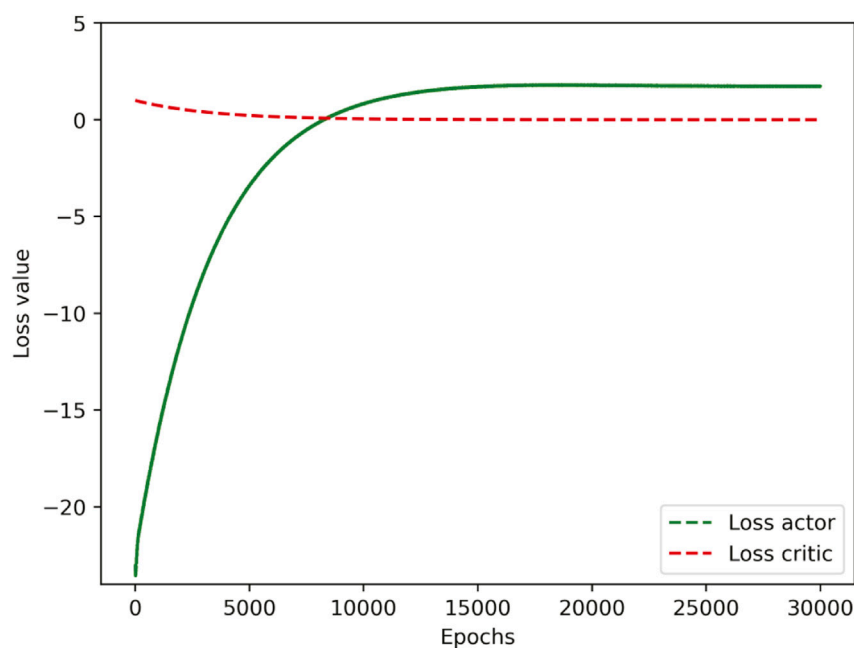
FIGURE 4
Training curve of SAC.

2.1 Task offloading of IoV

With the advent of the 6G era, mission data volume has experienced a blowout growth. With the intellectual development of the IoV intelligent system applications, the requirements for task data computation have also improved. Because the cloud is relatively far from users, traditional cloud computing has relatively high latency, which has become the focus of the task offloading strategy [24]. Researchers considered MEC as an effective technique to address the delay issue. Because the MEC servers are closer to users than the cloud, they can reduce the delay in task processing and enhance the user experience [25].

In light of MEC characteristics, it will be widely used in the future IoV system. In [26], the architecture of the vehicle network was defined according to the properties of the MEC, which can enhance the scalability of the network. In [27], an SDN-enabled network architecture assisted by the MEC was proposed to provide low-latency and high-reliability communication. In [28], the optimal task offloading issue in MEC was studied, which was transformed into two subproblems, task offloading and resource allocation, to minimize the delay.[29], considers an edge server and describes the computing and physical resource problems as optimization issues. In [30], a new offloading method was proposed to minimize transmission delay while improving resource utilization. In [31], a task offloading scheme fuzzy-task-offloading-and-resource-allocation (F-TORA) based on Takagi–Sugeno fuzzy neural network (T-S FNN) and game theory is designed.[32] proposes a UAV-assisted offloading strategy, which has been experimentally verified to reduce the delay by 30%.

2.2 Radar sensing and communication in the IoV

The integrated radar and communication design has great potential in cost-constrained scenarios. For example, by combining radar and communication functions, an IoV system can be designed to solve the issues of high latency and energy consumption. Some scholars have proposed a path estimation method to realize longitudinal and lateral vehicles followed only by radar and vehicle-to-vehicle (V2V) [33]. This paper introduces an intelligent real-time dual-functional radar–communication (iRDRC) system for autonomous vehicles (AVs) [34]. Obstacle detection is a very important part of the realization of intelligent vehicles. To avoid the problem that metal objects seriously block the millimeter wave, an active obstacle detection method based on a millimeter-wave radar base station is proposed [35]. The radar and communication integrated system (RCIS) can overcome the time-consuming problems of data format transfer and complex data fusion across multiple sensors in autonomous driving vehicles (ADVs) [36]. In summary, the integrated radar and communication design is a promising direction for future autonomous driving technology development.

2.3 Game theory in ToV

Game theory provides a framework for analyzing strategic interactions among rational decision-makers, while optimization techniques are designed to seek the most favorable outcomes. Some scholars have proposed a dependable content distribution framework that combines big data-based vehicle trajectory prediction with coalition game-based resource allocation in cooperative vehicular networks [37]. This paper proposes an
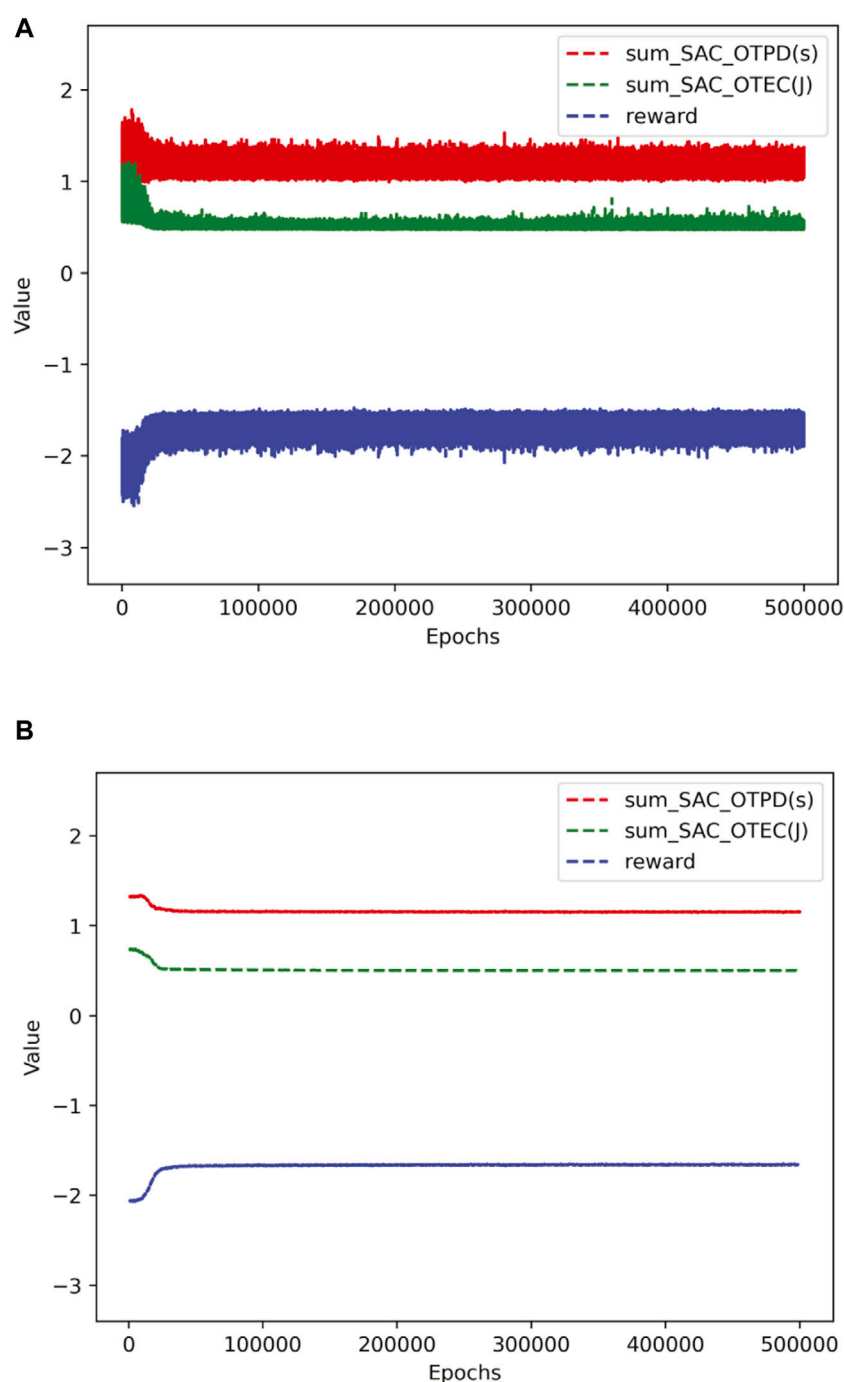
FIGURE 5
Convergence curve of SAC. **(A)** Unsmoothed convergence curve of SAC. **(B)** Smoothed convergence curve of SAC.

energy-efficient matching mechanism for resource allocation in device-to-device (D2D)-enabled cellular networks, which employs a game theoretical approach to formulate the interaction among end users and adopts the Gale–Shapley algorithm to achieve stable D2D matching [38]. Some scholars have proposed a novel game theoretical approach to encourage edge nodes to cooperatively provide caching services and reduce energy consumption [39]. In [40], the author has developed a two-player Stackelberg game-based opportunistic computation offloading scheme, which can

significantly shorten task completion delay. In conclusion, game theory holds significant and extensive application prospects within the realm of the IoV.

2.4 DRL methods for IoV

Regarding resource optimization for the IoV, DRL has strong sensing and decision-making capabilities compared to traditional
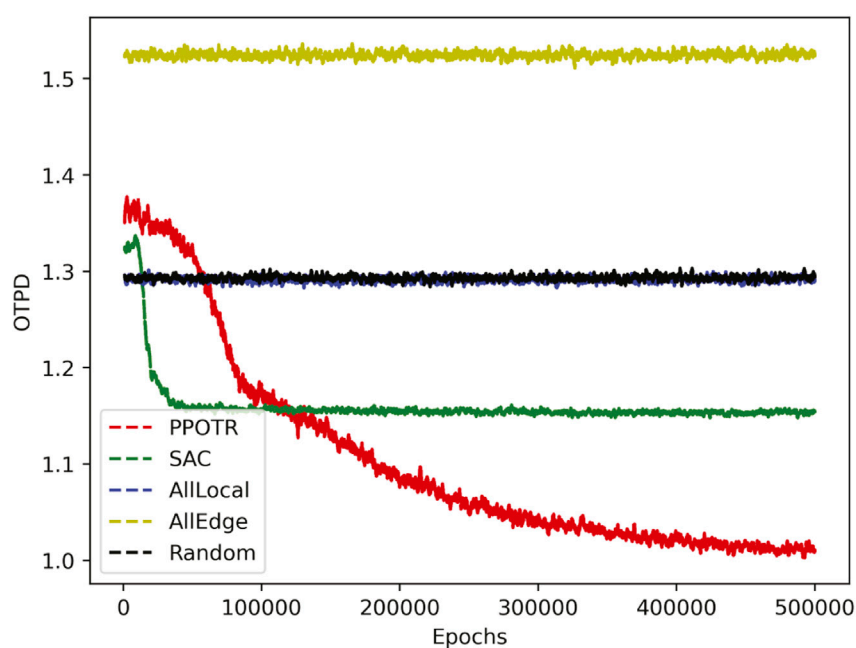
FIGURE 6
Total delay under different policies.
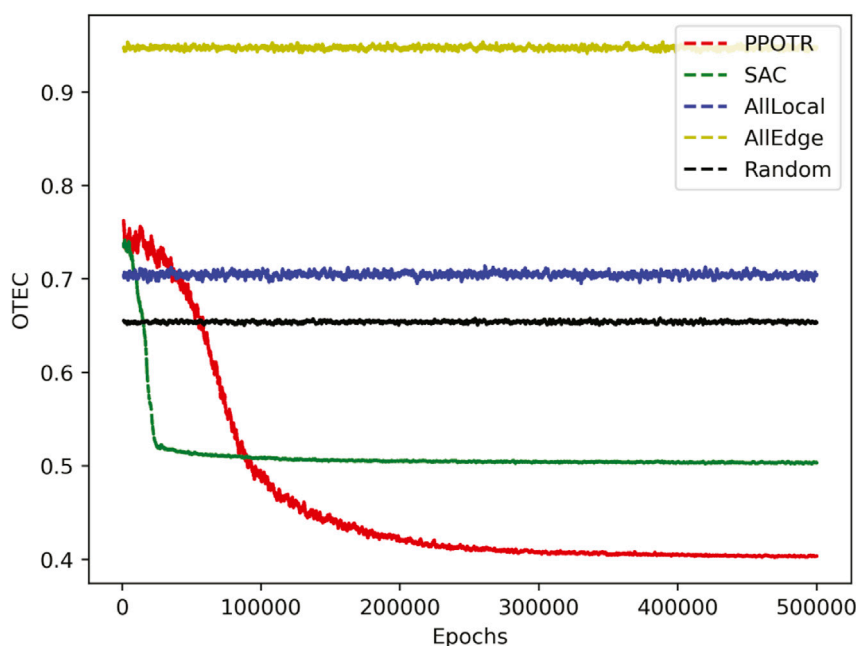
FIGURE 7
Total energy consumption under different policies.

heuristic algorithms and can analyze the long-term impact of current resource allocation on the system. Many scholars have applied DRL techniques to the study of the IoV. For example, in [41], DRL technology was used to transfer vehicle tasks to the edge server when facing the challenge of task delay. In [42], a UAV was placed in the vehicle network to assist resource allocation, and the deep deterministic policy gradient (DDPG) method was used to reduce the task delay. In [43], an online computation offloading strategy based on DQN was proposed, which takes the discrete channel gain as input to minimize energy consumption and delay
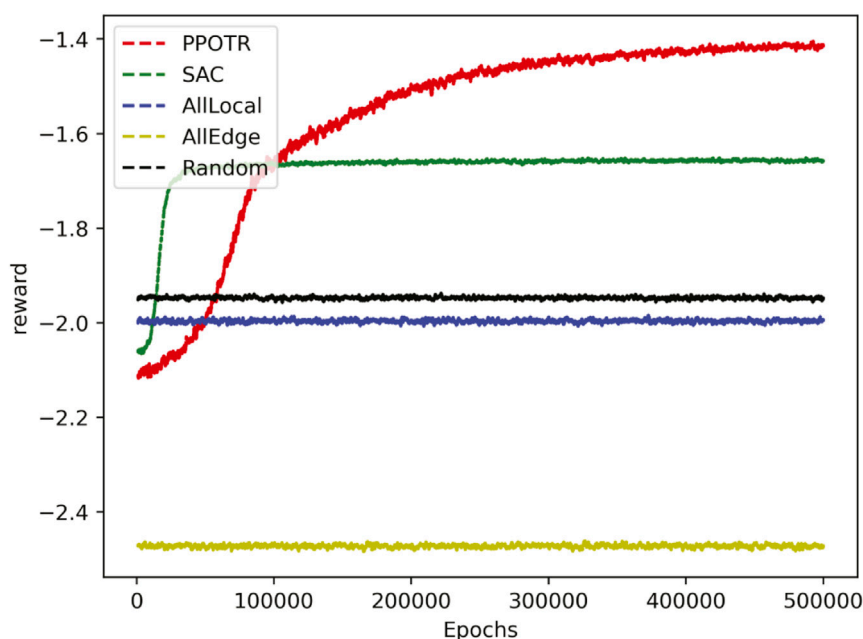
FIGURE 8
Reward under different policies.

and realize computation offloading and resource allocation. In [44], a hybrid scheduling mechanism to reduce computation was proposed for vehicle-to-vehicle communication in a specific area. In [45], the author proposed a priority-sensitive task offloading and resource allocation scheme in an IoV network to validate the feasibility of distributed reinforcement learning for task offloading in future IoV networks. In [46], the author proposed a multi-agent deep reinforcement learning (MA-DRL) algorithm for optimizing the task offloading decision strategy, while improving the offloading rate of the tasks and ensuring that a higher number of offloaded tasks are completed.

Given the preponderance of DRL techniques in the IoV system, two metrics are considered: delay and energy consumption. This paper aims to select the optimal task offloading strategy to save delay and resource costs. Consequently, we propose a framework for task offloading that uses a DRL-based algorithm to achieve optimal solutions in the network.

3 System description and problem formulation

In this section, Section 3.1 presents an edge computing network of the IoV. Section 3.2 presents an optimization issue.

3.1 System model

3.1.1 TODM mechanism based on the cooperative game

In light of the issue that the large amount of task data in the IoV leads to significant overall task delay and energy consumption, we build an intelligent system for the IoV by using the sensing capabilities of SNs. To achieve a practical and distributed solution, we realize that the task assignment problem in MEC architectures can also be formulated as a cooperative game. The cooperative game is applicable to the case of multi-node cooperation, where multiple agents work together to formulate resource allocation strategies to minimize overall delay and energy consumption. First, this paper defines the participants of the game, i.e., TaV, SeVs, and MEC. Second, it defines the strategies for task offloading, decomposing tasks into multiple subtasks assigned to different SNs, with the delay and energy consumption for nodes completing the task as the criteria for cooperative cost allocation. Finally, cooperative constraints are introduced to construct a cooperative game theory model.

The intelligent architecture network of the IoV is featured in Figure 1. A BS and MEC servers are deployed at the same location to improve MEC computing power and save costs. For RSUs reasonably deployed along the road, each RSU is equipped with storage resources and radars for real-time sensing of ambient data. The storage resources of RSUs support the storage of all sensed task data and are periodically cleared to maintain usability. RSUs are linked to the MEC via wired links. Each car is equipped with computing, storage resources, and radars. The TaV is linked to the SeVs and RSUs via wireless transmission. The communication between nodes adopts frequency division multiplexing (FDM) access technology, and the upload and feedback process adopts the time division duplexing (TDD) mode. This paper assumes that BS covers the entire IoV system, including all RSUs and vehicles. The coverage of RSUs is tangential to each other, and the TaV is always within the range of the nearest RSUs when processing the task. The task can be divided into several subtasks. Each subtask is independent and can be processed in parallel [47]. Considering the

impact of the delay and energy consumption on the offloading strategy, the MEC sends the offloading decision to the TaV and then offloads the task. In real-life scenarios, two-way roads are more practical. However, the study is still in its infancy. This paper only considers one-way lanes and ignores the car service in the opposite direction to the TaV in our model.

The delay and energy consumption are critical technical indicators in the TODM design, and this paper aims to minimize the OTPD and OTEC. The OTPD includes task description delay, offloading decision delay, offloading decision transmission delay, task upload delay, task processing delay, and task feedback delay. In this paper, each variable is represented by 64 bits, i.e., a double float. The task description size is a few kilobits, and the delay can be disregarded. The offloading decision comprises task allocation, transmission bandwidth, transmission power, and transmission policy. Compared with the amount of the task input data, the size of the offloading decision is small, so the offloading decision delay and transmission delay are disregarded. The amount of data after the task is completed is smaller than the amount of data input, and the task feedback delay can be disregarded. Thus, the OTPD consists of the task upload and computation delays. Similarly, the OTEC consists of the task upload energy consumption and the task computation energy consumption. It should be noted that if SNs perform other tasks, there will be waiting delays, and energy consumption is possible. In this paper, it is assumed that only one task needs to be processed, and the waiting delay and energy consumption are neglected. Multi-tasking will be considered in future study.

Due to the different perspectives of sensing environmental data, the TaV coordinate in the calculation instruction is used for coordinate transformation (CdT) preprocessing to eliminate differences [48]. The TaV has two ways to transmit the task: conventional data transmission (DaT) and instruction transmission (InT) with cooperative environment awareness. Different transfer methods offer new options for task offloading. The delay and energy consumption constraints affect the task upload mode, further affecting the task offloading strategy. Therefore, the transmission strategy can be chosen adaptively based on the objective function, traffic size, propagation capability, transmission delay, energy consumption, *etc.* Compared to the traditional offloading mechanism, TODM can potentially reduce energy consumption and transmission delay caused by the inputs. However, this mechanism incurs an additional cost to the overall IoV system, which is ignored in this paper.

3.1.2 Task model

The task of TaV is computationally intensive and delay-sensitive. The total task data are denoted by S_{DaT} and can be arbitrarily divided into infinitely many subtasks. The task ratio is denoted as $x_n (x_n \in [0,1]), n \in \mathcal{N} := \{0,1,2,\ldots,n,N+1\}$. H is used to denote the task, and the h-th subtask is denoted as h, $h \in \mathcal{N} := \{0,1,2,\ldots,n,N+1\}$, where $h \in H$. Some subtasks select local computations, while others select DaT or InT for SNs according to the task ratio. Task offloading to SNs can satisfy the delay and energy consumption constraints. V_n, $n \in \mathcal{N} := \{0,1,2,\ldots,n,N+1\}$, is used to denote the SNs; the

wireless bandwidth ratio is denoted as $b_n (b_n \in [0,1])$, $n \in \mathcal{N} := \{0,1,2,\ldots,n,N+1\}$; and the transmission power is denoted as $P_n (P_n \in [0.5w, 1.5w])$, $n \in \mathcal{N} := \{0,1,2,\ldots,n,N+1\}$. TaV is denoted as V_0, and the SeVs are denoted as V_n. The computational resources of both TaV and SeVs satisfy all computing tasks. RSUs and MEC are connected by wires, denoted by $V_N + 1$. The choice of the aforementioned three variables ensures the optimal task offloading strategy.

3.1.3 OTPD and OTEC of the TaV

When a subtask is selected to perform a local computation on V_0, the OTPD of the subtask is the local computation. The OTEC of the subtasks is the energy consumption computed locally and uploaded without energy consumption. The delay is denoted as T_0^{comput}, and the energy consumption is denoted as E_0^{comput}. The T_0^{comput} and E_0^{comput} are given as follows [49]:

$$T_0^{comput} = t_0^{comput}(x_0) = \frac{S_{DaT}x_0 M}{F_0}, \quad (1)$$

$$E_0^{comput} = e_0^{comput}(x_0) = K_0(f_0)^2 C_0, \quad (2)$$

$$C_0 = S_{DaT}x_0 M. \quad (3)$$

Here, M (in cycles/bit) is the task calculation strength, which refers to the computing resources required to input 1 bit of data. F_0 represents the CPU cycles of V_0. K_0 is the effective switching capacitance related to the chip structure in the car. f_0 is the computing capacity of the car itself. C_0 represents the number of CPU revolutions required for processing the subtasks h_0.

3.1.4 OTPD and OTEC of the SeVs

When a subtask is offloaded to the SeVs for processing, data or calculation instructions are transmitted wirelessly to the SeVs. For DaT, it is essential to consider the upload delay. For InT, the transmission delay is not considered, but it is essential to consider the CdT delay.

Uploading delay model: Based on comparing the delays between the two upload modes, the mode with the smaller delay is selected as the upload mode. The DaT and InT upload methods are considered, and an energy consumption model is built. The energy consumption corresponding to different upload methods is calculated. T_1^{upload} is used to denote the uploading delay for V_0 transmitting the task to V_n. The uploading rate from V_0 to V_n is given by

$$R_{V_0 \to V_n}^{DaT}(t) = B_{ToT}b_n \log_2\left(1 + \frac{P_n|\tilde{h}_n|^2 d_n^{-\alpha}(t)}{\sigma_n^2}\right). \quad (4)$$

Here, $R_{V_0 \to V_n}^{DaT}(t)$ denotes the upload rate in time t. B_{ToT} denotes the total bandwidth of the wireless transmission. b_n denotes the transmission bandwidth ratio. P_n denotes the transmission power. σ_n^2 denotes the noise power. $\tilde{h}_n$ denotes the channel fading coefficient from V_0 to V_n. $d_n(t)$ denotes the distance from V_0 to V_n in time t. $d_n^{-\alpha}(t)$ denotes the path loss from V_0 to V_n. α denotes the path loss index.

During the task data upload, the car's motion causes changes in $d_n(t)$. We assume that the coordinate of V_0 is 0, and $S_{DaT}x_n$ is G_n, where $G_n \neq 0$. The moving speeds of the cars are v_0 and v_n. The formula for calculating $d_n(t)$ is given by

$$d_n(t) = \sqrt{|G_n + (v_n - v_0)t|^2}, \quad G_n \neq 0, \quad n \in \mathcal{N}. \quad (5)$$

The cars are running on the expressway, and the maximum difference between their relative speeds does not exceed 30 km/h [50]. Take 10 ms as an example; $(v_n - v_0)t = 0.008 \ m$. The relative position changes are relatively small and do not affect the optimization results. This paper ignores the change in position. The calculation formula of $d_n(t)$ is given by

$$d_n(t) = |G_n| \quad G_n \neq 0 \quad n \in \mathcal{N}. \quad (6)$$

For DaT, $S_{DaT}x_n$ represents the amount of the task data allocated to V_n. $B_{ToT}b_n$ represents the transmission bandwidth from V_0 to V_n. The upload delay T_{1a}^{upload} is given by

$$T_{1a}^{upload} = t_{1a}^{DaT}(x_n, b_n, P_n) = \frac{S_{DaT}x_n}{R_{V_0 \to V_n}^{DaT}}$$

$$= \frac{S_{DaT}x_n}{B_{ToT}b_n \log_2\left(1 + \frac{P_n|\tilde{h}_n|^2|G_n|^{-\alpha}}{\sigma_n^2}\right)}. \quad (7)$$

For InT, this paper needs to consider the delay of CdT. Assume V_n stores the environmental data sensed by the radar and performs CdT immediately after receiving the calculation instruction. The delay of CdT depends on the amount of sensed data and the strength of the CdT calculation. T_{1b}^{upload} is used to denote the upload delay; T_{1b}^{upload} is given by

$$T_{1b}^{upload} = t_{1b}^{tra}(x_n) = \frac{S_{DaT}x_n M_{tra}}{F_n}, \quad (8)$$

where M_{tra} represents the computation intensity of CdT. F_n represents the CPU cycles of V_n.

Considering the TODM, DaT or InT with a lower delay is chosen as the upload method to minimize the task upload delay. The calculation formula of upload delay T_1^{upload} from V_0 to V_n is given by

$$T_1^{upload}(x_n, b_n, P_n) = \min\{T_{1a}^{upload}, T_{1b}^{upload}\}. \quad (9)$$

Uploading energy consumption model: Given the selected upload mode, the upload energy consumption model is built, and the upload energy consumption E_{1a}^{upload} and E_{1b}^{upload} are calculated. E_1^{upload} is used to denote the upload energy consumption; E_1^{upload} is given by [51]

$$E_1^{upload} = \begin{cases} E_{1a}^{upload} = e_{1a}^{upload}(x_n, b_n, P_n) = P_n T_{1a}^{upload} \\ \text{or} \\ E_{1b}^{upload} = e_{1b}^{upload}(x_n) = K_n(f_n)^2 C_{n1} \end{cases}, \quad (10)$$

$$C_{n1} = S_{DaT}x_n M_{tra}, \quad (11)$$

where K_n is the effective switching capacitor related to chip structure in cars. f_n is the calculation capacity of the car itself. C_{n1} represents the number of CPU revolutions required for processing the subtasks h_n.

Computing delay model: After the task is uploaded, the SeVs V_n start the parallel computation of the subtasks and obtain the computation delay. T_1^{comput} is used to denote the computing delay; T_1^{comput} is given by

$$T_1^{comput} = t_1^{comput}(x_n) = \frac{S_{DaT}x_n M}{F_n}. \quad (12)$$

Computing energy consumption model: The computing energy consumption model is designed according to the assigned task. C_{n2} is used to denote the number of CPU revolutions required for processing the subtasks h_n. E_1^{comput} is used to denote the computing energy consumption; E_1^{comput} is given by

$$E_1^{comput} = e_1^{comput}(x_n) = K_n(f_n)^2 C_{n2}, \quad (13)$$

$$C_{n2} = S_{DaT}x_n M. \quad (14)$$

3.1.5 OTPD and OTEC of MEC

When a subtask is offloaded to the MEC servers for processing, the upload delay includes both wireless and wired transmission delays. The upload energy consumption includes both wireless transmission energy consumption and wired transmission energy consumption.

Uploading delay model: The V_0 transmits the subtasks' data to RSUs via wireless transmission. RSUs transmit the subtasks' data to the MEC servers via wired transmission. The uploading rate from V_0 to V_{N+1} is given by

$$R_{V_0 \to V_{N+1}}^{DaT}(t) = B_{ToT}b_{N+1} \log_2\left(1 + \frac{P_{N+1}|\tilde{h}_{N+1}|^2 d_{N+1}^{-\alpha}(t)}{\sigma_{N+1}^2}\right), \quad (15)$$

where $R_{V_0 \to V_{N+1}}^{DaT}(t)$ denotes the upload rate in time t. b_{N+1} denotes the transmission bandwidth ratio. P_{N+1} denotes the transmission power. σ_{N+1}^2 denotes the noise power. $\tilde{h}_{N+1}$ denotes the channel fading coefficient from V_0 to V_{N+1}. $d_{N+1}(t)$ denotes the distance from V_0 to V_{N+1} in time t. $d_{N+1}^{-\alpha}(t)$ denotes the path loss from V_0 to V_{N+1}. During the upload of the task data, the movement of cars causes changes in $d_{N+1}(t)$. We assume that the coordinate of V_0 is 0, and V_{N+1} is G_{N+1}, where $G_{N+1} \neq 0$. The moving speed of the car is v_0; the calculation formula of $d_{N+1}(t)$ is given by

$$d_{N+1}(t) = \sqrt{|G_{N+1} - v_0 t|^2 + D_{N+1}^2 + H_{N+1}^2}, \quad G_{N+1} \neq 0, \quad N \in \mathcal{N}, \quad (16)$$

where D_{N+1} is the distance from RSUs to the centerline. H_{N+1} represents the height of RSUs. Take t = 20 ms as an example; when the speed of the car is 120 km/h, $v_0 t = 0.67$ m. The change in position is ignored compared with tens of meters. The calculation formula of $d_{N+1}(t)$ is given by

$$d_{N+1}(t) = \sqrt{|G_{N+1}|^2 + D_{N+1}^2 + H_{N+1}^2}, \quad G_{N+1} \neq 0, \quad N \in \mathcal{N}. \quad (17)$$

For DaT, $S_{DaT}x_{N+1}$ represents the amount of the task data allocated to V_{N+1}. $B_{ToT}b_{N+1}$ represents the transmission bandwidth from V_0 to V_{N+1}. The upload delay T_{2a}^{upload} is given by

$$T_{2a}^{upload} = t_{2a}^{DaT}(x_{N+1}, b_{N+1}, P_{N+1}) = \frac{S_{DaT}x_{N+1}}{R_{V_0 \to V_{N+1}}^{DaT}}$$

$$= \frac{S_{DaT}x_{N+1}}{B_{ToT}b_{N+1} \log_2\left(1 + \frac{P_{N+1}|\tilde{h}_{N+1}|^2 d_{N+1}^{-\alpha}(t)}{\sigma_{N+1}^2}\right)}. \quad (18)$$

After the task data are uploaded to RSUs, RSUs will transmit the data to MEC via wired transmission. T_{2R}^{upload} is used to denote the wired upload delay. R_{wired} is used to denote the wired transmission speed. The wired upload delay T_{2R}^{upload} is given by

$$T_{2R}^{upload}(x_{N+1}) = t_{2R}^{upload}(x_{N+1}) = \frac{S_{DaT}x_{N+1}}{R_{wired}}. \qquad (19)$$

Thus, let T_{2aR}^{upload} be the total upload delay, which is equal to the sum of T_{2a}^{upload} and T_{2R}^{upload}. T_{2aR}^{upload} is given by

$$T_{2aR}^{upload} = T_{2a}^{upload} + T_{2R}^{upload}. \qquad (20)$$

For InT, this paper assumes CdT is carried out immediately after the MEC servers receive the calculation instruction. Let T_{2b}^{upload} be the CdT delay. The calculation formula of T_{2b}^{upload} is given by

$$T_{2b}^{upload} = t_{2b}^{Rtra}(x_{N+1}) = \frac{S_{DaT}x_{N+1}M_{tra}}{F_{N+1}}, \qquad (21)$$

where F_{N+1} denotes the CPU cycles of the MEC. T_{2bR}^{upload} denotes the total upload delay, which is equal to the sum of T_{2b}^{upload} and T_{2R}^{upload}. The T_{2bR}^{upload} is given by

$$T_{2bR}^{upload} = T_{2b}^{upload} + T_{2R}^{upload}. \qquad (22)$$

Similarly, DaT or InT with a lower delay is chosen as the upload method. The formula for the upload delay from V_0 to V_{N+1} is given by

$$T_2^{upload}(x_{N+1}, b_{N+1}, P_{N+1}) = \min\{T_{2aR}^{upload}, T_{2bR}^{upload}\} \\ = \min\{T_{2a}^{upload}, T_{2b}^{upload}\} + T_{2R}^{upload}. \qquad (23)$$

Uploading energy consumption model: In light of the selected upload mode, build the upload energy consumption model and calculate the upload energy consumption E_{2a}^{upload} and E_{2b}^{upload}. Each transmission mode shall transmit data from RSUs to the MEC via wired mode, using E_{2R}^{upload} to denote the energy consumption of wired transmission. The E_{2R}^{upload} is given by

$$E_{2R}^{upload} = e_{2R}^{upload}(x_{N+1}) = P'_{N+1}T_{2R}^{upload}, \qquad (24)$$

where P'_{N+1} denotes the wired transmission power. E_2^{upload} denotes the total upload energy consumption. E_2^{upload} is given by

$$E_2^{upload} = \begin{cases} E_{2aR}^{upload} = E_{2a}^{upload} + E_{2R}^{upload} \\ or \\ E_{2bR}^{upload} = E_{2b}^{upload} + E_{2R}^{upload} \end{cases}, \qquad (25)$$

$$E_2^{upload} = \begin{cases} e_{2a}^{upload}(x_{N+1}, b_{N+1}, P_{N+1}) + e_{2R}^{upload}(x_{N+1}) \\ or \\ e_{2b}^{upload}(x_{N+1}) + e_{2R}^{upload}(x_{N+1}) \end{cases} \\ = \begin{cases} P_{N+1}T_{2a}^{upload} + P'_{N+1}T_{2R}^{upload} \\ or \\ K_{N+1}(f_{N+1})^2 C_{N+1}^1 + P'_{N+1}T_{2R}^{upload} \end{cases}, \qquad (26)$$

$$C_{N+1}^1 = S_{DaT}x_{N+1}M_{tra}, \qquad (27)$$

where K_{N+1} is the effective switching capacitor related to chip structure in the MEC. f_{N+1} is the calculation capacity of the server itself. C_{N+1}^1 represents the number of CPU revolutions required for processing the subtasks h_{N+1}.

Computing delay model: After the task is uploaded, the MEC servers start the parallel computation of the subtasks and obtain the computation delay. T_2^{comput} is used to denote the computing delay; T_2^{comput} is given by

$$T_2^{comput} = t_2^{comput}(x_{N+1}) = \frac{S_{DaT}x_{N+1}M}{F_{N+1}}. \qquad (28)$$

Computing energy consumption model: The computing energy consumption model is created according to the assigned task. Let C_{N+1}^2 be the number of CPU revolutions required for processing the subtasks h_{N+1}. Let E_2^{comput} be the computing energy consumption; E_2^{comput} is given by

$$E_2^{comput} = e_2^{comput}(x_{N+1}) = K_{N+1}(f_{N+1})^2 C_{N+1}^2, \qquad (29)$$

$$C_{N+1}^2 = S_{DaT}x_{N+1}M. \qquad (30)$$

3.2 Problem formulation

This paper aims to solve the issue of joint task offloading based on the edge computing network of IoV, that is, to minimize the task delay and energy consumption under the constraints of limited system resources. The payoff function is the weighted sum of task processing, energy consumption, and delay. The additional weight balances the effect of energy consumption and delay on the payoff function. T^{total} is used to denote the total delay, which is given by

$$T^{total} = T_0^{comput} + T_1^{upload} + T_1^{comput} + T_2^{upload} + T_2^{comput}. \qquad (31)$$

E^{total} is used to denote the total energy consumption, which is given by

$$E^{total} = E_0^{comput} + E_1^{upload} + E_1^{comput} + E_2^{upload} + E_2^{comput}. \qquad (32)$$

The payoff function S_R^{total} is expressed as

$$S_R^{total}(x, b, P) = \zeta T^{total}(x, b, P) + (1 - \zeta)E^{total}(x, b, P). \qquad (33)$$

The payoff function is transformed into the total objective function of the joint optimization issue. The optimization problem can be described as minimizing the delay and energy consumption under task allocation, transmission bandwidth allocation, and transmit power control constraints. Thus, the optimization issue can be formulated as

$$(P1): \quad \underset{x,b,P}{minimize} \quad S_R^{total}(x, b, P) := \underset{n=0,1,\ldots,N+1}{min}\left\{S_R^{total}(x_n, b_n, P_n)\right\}$$

$$s.t. \quad C1: \sum_{n=0}^{N+1} x_n = 1$$

$$C2: \sum_{n=1}^{n+1} b_n \leq 1$$

$$C3: 0 \leq x_n \leq 1, \quad n = 1, 2, \ldots\ldots, N+1$$

$$C4: 0 \leq b_n \leq 1, \quad n = 1, 2, \ldots\ldots, N+1$$

$$C5: 0.5 \leq P_n \leq 1.5, \quad n = 1, 2, \ldots\ldots, N+1$$

$$C6: 0 \leq \zeta \leq 1$$

$$(34)$$

In problem P1, constraint C1 represents the task allocation ratio, and the sum of the ratio is 1. C2 denotes the allocation ratio of wireless bandwidth. The sum of the wireless bandwidth allocation ratios is less than 1. C3 and C4 represent the value range of the task allocation ratio and wireless bandwidth ratio, respectively. C5 limits the transmit power of the uplink transmission rate. C6 represents the weight value.

4 DRL-based algorithm for task offloading

Section 4.1 presents DRL techniques and the Markov decision process (MDP). Section 4.2 proposes the conversion of the

optimization issues in the model into DRL issues. Section 4.3 proposes a PPO-based approach to address the task offloading issue.

4.1 DRL-based framework

4.1.1 DRL techniques

Deep learning (DL) has strong perception ability but lacks specific decision-making abilities; RL has decision-making abilities but does not address the solving of perception issues. DRL integrates DL's perception ability and RL's decision-making ability, which solves the perceptual decision issue of complex systems. DRL is an end-to-end sensing and control method with strong generality. Its learning process can be described as follows: (i) at each moment, the agent interacts with the environment to get a high-dimensional observation and specific state characteristics. (ii) The current state is mapped to the corresponding action through the strategy, and the value function of each action is evaluated. (iii) The environment gives feedback to the action to obtain the next observation object. The optimal policy is obtained by successive cycles of the aforementioned procedure.

4.1.2 Markov decision process

Almost all issues can be formulated as MDP in the formal description of RL environments. MDP refers to the decision-maker who periodically or continuously observes the stochastic dynamic system with Markov properties and makes decisions. It includes the environmental state, action, reward, state transition probability matrix, and discount factor. The process is given a state. The agent obtains the new state by performing actions based on the state transition probability matrix. Each strategy is rewarded for its implementation.

4.2 Problem transformation

The IoV scenario has continuous state and action space, which will increase the issue's complexity. So, it is a challenge to find the best task offloading strategy. Traditional optimization algorithms require significant iterations to achieve an approximate solution when solving such issues, which does not meet the requirements of time-varying systems. However, DRL algorithms can meet real-time decision-making requirements. Therefore, this paper adopted the DRL algorithm to solve the aforementioned issues. Get the optimal task offloading strategy through continuous interaction with the IoV environment.

Problem P1 is a complex issue with continuous real variables, which have strong coupling. Task allocation, transmission power, and transmission bandwidth are all continuous real variables. Therefore, P1 is a non-convex combined issue that cannot be solved directly through mathematical calculation. In light of this, this paper turns the optimization issue into a DRL issue and proposes adopting the DRL algorithm to solve the global optimization issue. Thus, the optimization issue (34) is established as follows:

$$(P2): \quad \underset{x,b,P}{minimize} \quad \tilde{S}_R^{total} \triangleq \mathbb{E}\left[\lim_{|T|\to\infty} \frac{1}{|T|} \sum_{t\in T} S_R^{total}\right], \quad (35)$$

$$s.t. \quad C1 - C6$$

where $\mathbb{E}(\cdot)$ represents the mathematical expectation.

In the IoV system, the cars are moving, and the vehicle status, edge server status, wireless transmission channel status, and RSU status are changing. The system needs to make different decisions to minimize delay and energy consumption and meet the reasonable allocation of resources. The transmission bandwidth and computing resources allocated by the IoV system to cars and RSUs are continuous values. Traditional DQN is mainly for discrete space. The DDPG is mainly for constant action space. The SAC and PPO can be applied to discrete and continuous spaces. Therefore, this paper designs a PPO-based method to find the optimal task offloading strategy. Next, the paper delves into the environmental state, action space, and reward function of Markov games.

4.2.1 Environment state

The environment state $S(t)$ reflects the impact of the channel condition information and agent behavior on the environment [52]. The state information includes the state of the cars, BS, and RSUs. The state of the car consists of the vehicle coordinates, transmission bandwidth, transmission power, and task allocation. The state of BS and RSUs includes the task size, transmission power, and transmission bandwidth. Each agent observes that the environment state is

$$\mathcal{S}(t) = \{U_n(t), N_n(t), R_n(t)\}, \quad (36)$$

where $U_n(t)$, $N_n(t)$, and $R_n(t)$ denote the status of vehicles, BS, and RSUs, respectively.

4.2.2 Action space

Although the computational complexity of DRL is relatively low in large-scale network scenarios, the spatial dimension changes as the number of agents increases. The high-dimensional space will make the system calculation difficult and affect the best decision. The algorithm's performance will suffer from dimension disaster due to the high-dimensional action and state space [53]. The agent takes actions according to the currently observed state to avoid the high computational complexity, that is, jointly optimize the task allocation, transmission bandwidth allocation, uplink power control, and offloading decision. Hence, the action is

$$a(t) = \{x_n(t), b_n(t), P_n(t)\}. \quad (37)$$

- x_n represents the task allocation policy.
- b_n represents the uplink transmission bandwidth.
- $P_n(t)$ represents the uplink transmission power.

The agent selects the offloading decision based on the present state. If the agent sets local computing for the task, the computing resources must meet the requirements. However, if the agent selects to calculate the task on the SeVs or MEC, the transmission bandwidth and computing resources must meet the needs.

4.2.3 Rewards

This paper should strictly follow constraints C1–C6 in the design of state space, action space, and reward function to optimize the task offloading strategies. The sum of the reward functions of nodes in all states is constant, and there is a competitive and cooperative relationship between nodes. Therefore, the paper sets the reward value as the opposite of the objective function. In optimization issue

P2, the agent maximizes the interests through action selection to affect the system's state. The reward function is

$$R_t = -S_R^{total}. \tag{38}$$

4.3 PPO-based algorithm framework

This paper proposes a PPO-based task offloading and resource allocation (PPOTR) algorithm to obtain stable performance in the actual changing network. The agent chooses the action to interact with the environment according to the policy, thus affecting the environment state and updating the environment parameters. Next, according to the new policy, the agent chooses actions to interact with the environment. Let $r_t(\theta)$ denote the action probability ratio of new and old strategies.

$$r_t(\theta) = \frac{\pi_\theta(a_t, s_t)}{\pi'_\theta(a_t, s_t)}. \tag{39}$$

When $r_t(\theta) > 1$, it indicates that the current strategy is more inclined to select the sampling action. Otherwise, it is not. The PPO algorithm improves the original policy gradient (PG) algorithm, and the formula of the new objective function is given by

$$L(\theta) = E_t \left[\frac{\pi_\theta(a_t, s_t)}{\pi'_\theta(a_t, s_t)} \cdot A_t \right]. \tag{40}$$

4.3.1 Training algorithm

PPO uses a new objective function to control the change in the strategy in each iteration, which is uncommon in other algorithms. The objective function is

$$L^{clip}(\theta) = E_t \left[\min(r_t(\theta))A_t, \, clip(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)A_t) \right], \tag{41}$$

where θ denotes the policy parameters. E_t denotes the empirical expectation of the time step. r_t denotes the probability ratio under the new and old strategies. A_t represents the estimated advantage. ε denotes the hyperparameter. The value is usually 0.1 or 0.2. $clip(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon)$ is given by

$$clip(r_t(\theta), 1 - \varepsilon, 1 + \varepsilon) = \begin{cases} 1 - \varepsilon, & if\, r_t(\theta) \le 1 - \varepsilon \\ 1 + \varepsilon, & if\, r_t(\theta) \ge 1 + \varepsilon \\ r_t(\theta), & otherwise \end{cases} . \tag{42}$$

4.3.2 Replay buffer

The static data in the DL differ from the data in the DRL, which is obtained according to machine learning. At each time step, the agent observes the current environment state and saves the state, action, reward, and prediction data $data_t = (s_t, a_t, r_t, s_{t+1})$ of the following environment state to the replay buffer [54]. In particular, in our model, the data of the training network will be aggregated after 1,000 time steps. We can see the data changes in the training process and avoid the correlation in the observation state sequence to reduce the update variance. Moreover, the data of each experiment can be used continuously in other weight updates to improve the efficiency of data use.

4.3.3 Algorithm steps

In the aforementioned architecture, the agent is the car, MEC is the policy decision center, and the SeVs and RSUs are the intermediaries of

perception information. The algorithm's input is the environment state information, and the output is the optimal offloading policy and target value. Algorithm 1 presents the pseudo-code.

```
Input: initial policy parameters and initial value
function parameters φ₀
1: for k = 0, 1, 2 ... do do
2:    Collect a set of trajectories Dₖ = {τᵢ} by running
      policy πₖ = π(θₖ) in the environment.
3:    Compute rewards-to-go Rₜ
4:    Compute advantage estimates, Aₜ (using any method
      of advantage estimation) based on the current
      value function V_θₖ.
5:    Update the policy by maximizing the PPO-clip
      objective, normally via stochastic gradient
      ascent with Adam.
      θₖ₊₁ = arg max_θ 1/|Dₖ|T Σ_{τ∈Dₖ} Σ_T^{t=0} m · (π_θ(aₜ|sₜ)/π'_θₖ(aₜ|sₜ) A^{π'_θₖ}(sₜ,aₜ), g(, A^{π'_θₖ}(sₜ,aₜ)))
6:    Fit value function by regression on the mean-squared
      error, normally via some gradient descent
      algorithm.
      φₖ₊₁ = arg min_φ 1/|Dₖ|T Σ_{T∈Dₖ} Σ_T^{t=0} (V_φ(sₜ) − Rₜ)²
7: end for
```

Algorithm 1. PPO-based algorithm for task offloading and resource allocation.

The following illustrates the steps of the proposed PPOTR algorithm. First, enter the initial policy parameter θ_0 and the value function parameter ϕ_0. Second, start iteration and collect a set of trajectories $D_k = \{\tau_i\}$ by running policy $\pi_k = \pi(\theta_k)$ in the environment. Then, in the fourth and fifth steps, calculate the reward value R_t and use the advantage estimation method based on the current value function V_{θ_k} to calculate the advantage estimation A_t. Then, in the sixth and seventh steps, update the strategy θ_{k+1} through PPO-clip objective function $L^{clip}(\theta)$ and the fit value function ϕ_{k+1} through mean square error regression. Finally, the algorithm iteration is ended.

Complexity analysis: The algorithm's main computational costs include the interaction with the environment, the action, and evaluation under the old and new strategies. In the process of interacting with the environment, the agent determines the input state according to the policy. Furthermore, the agent calculates the probability ratio of the action under the new and old policies through the transmission between the action network and the critic network. The time complexity of the training process interacting with the environment is given by [55]

$$C_{PPOTR}^1 = \mathcal{O}\left(\mathcal{N}_r \mathcal{N}_e d^{max} \sum_{x=0}^{X} \Omega_x^A \Omega_{x+1}^A \right). \tag{43}$$

The time complexity of policy updates is given by

$$C_{PPOTR}^2 = \mathcal{O}\left(\mathcal{N}_r \mathcal{N}_s \left(\sum_{x=0}^{X} \Omega_x^A \Omega_{x+1}^A + \sum_{y=0}^{Y} \Omega_y^C \Omega_{y+1}^C \right) \right), \tag{44}$$

where x, y are the quantities of full connection layers of the network, respectively. Ω_x^A represents the unit of the x-th actor network, and Ω_y^C represents the artificial neuron of the y-th critic network. Then, the total time complexity of Algorithm 1 is given by

$$C_{PPOTR} = \mathcal{O}\left(\mathcal{N}_r \left(\mathcal{N}_s + \mathcal{N}_e d^{\max} \right) \sum_{x=0}^{X} \Omega_x^A \Omega_{x+1}^A \right)$$

$$+ \mathcal{O}\left(\mathcal{N}_r \mathcal{N}_s \sum_{y=0}^{Y} \Omega_y^C \Omega_{y+1}^C \right). \tag{45}$$

The space complexity of the algorithm [56] is given by

$$C_{PPOTR}^{space} = \mathcal{O}\left(\mathcal{N}_r \left(\mathcal{N}_s + \mathcal{N}_e d^{\max} \right) \sum_{x=0}^{X} \Omega_x^A \Omega_{x+1}^A \right)$$

$$+ \mathcal{O}\left(\mathcal{N}_r \mathcal{N}_s \sum_{y=0}^{Y} \Omega_y^C \Omega_{y+1}^C \right) + \mathcal{O}(\mathcal{N}), \tag{46}$$

where $\mathcal{N}$ is the space complexity of the experience replay buffer in the algorithm.

5 Simulation results

In this section, a series of simulation experiments to verify the performance of the proposed algorithm have been proposed. The simulation results of different algorithms under the same network settings are given to compare the characteristics of different algorithms. This paper analyzes the convergence curves of delay, energy consumption, and objective function under the offloading strategy and shows that the algorithm is reasonable. This paper adopts four benchmark schemes from the perspective of convergence, and the effectiveness and efficiency of the algorithm are verified through the analysis of energy consumption and delay. The four schemes are as follows:

- Actor critical (AC) algorithm based on SAC [57]: Under the same environmental settings, this paper uses the AC algorithm with a soft update mechanism to solve the issue.
- Local computing policies for all the tasks (AllLocal): All the tasks are performed locally, and the local computing resources meet the requirements of task calculation. Calculate the corresponding delay, energy consumption, and objective function value.
- All-edge server-only execution policy (AllEdge): Offload all the tasks to the edge server, and the edge computing resources and transmission bandwidth meet the task's requirements.
- Random offloading policy (Random): Offload the task randomly and allocate resources randomly.

5.1 Simulation setup

This paper evaluates the performance of the proposed algorithm through multiple simulation experiments. We assume that four RSUs are set at the roadside, the coverage diameter of each RSU is 160 m, and the coverage is tangential to each other. For the sake of driving safety, we assume that there is one TaV and ten SeVs within the coverage of RSUs. The TaV and SeVs are always within the coverage of RSUs. Assume the input data S_{DaT} size is 25 Mb (one frame with a resolution of 1920*1080, 12 bits per pixel), and the total transmission bandwidth B_{ToT} is 100 MHz [58].

The calculation strength is 2,640 cycles/bit [59]. The CPU frequency of each car is randomly selected within the range of $0.3 \times 10^{12} \sim 0.6 \times 10^{12}$ cycles/s, and the CPU frequency of the MEC servers is randomly selected within the range of $1 {}^*10^{12} \sim 2{}^*10^{12}$ cycles/s [60]. The effective switching capacitor of the vehicle and MEC is 10^{-27}. The transmission rate R_{wired} for RSU wired transmission to the MEC is $100\ Gb/s$ [50]. The calculation capacity of cars and MEC is set to $1.4\ Gr/s$ and $2.8\ Gr/s$, respectively Song et al. [51].

This paper considers the effect of small-scale fading on transmission performance; the channel fading coefficients $|\tilde{h}_n|^2$ and $|\tilde{h}_{N+1}|^2$ are 1 [61]. The simulation experiment is completed in the environment of Pytorch 1.11.0 using Windows 10 system and Python 3.10 software. Other system parameters used in simulations are shown in Table 1.

5.2 Results

The learning curve of the PPOTR algorithm is shown in Figure 2, including the training losses of the action and criticism network. In the simulation experiment, if the value function of training loss does not tend to 0 for the action network, it proves that the whole action space has many places not explored, and there are still differences between the new and old action space. For the critic network, if the value function of training loss does not tend to 0, it proves that the critic network cannot perfectly predict the value of the state space. The simulations show that starting from the 50th training set, the value function fluctuates in a small range, and the gradient of the loss value decreases gradually. This indicates that the algorithm begins to converge and can quickly learn the optimal strategy.

Figures 3A, B show the convergence curves of the delay, energy consumption, and objective function of the PPOTR algorithm to solve the aforementioned issues. Figure 3A shows the unsmoothed curve of the training process, illustrating the total delay sum_OTPD in seconds, the total energy consumption sum_OTEC tin joules, and the reward value. Figure 3B is the convergence curve after smoothing in Figure 3A; the smooth curve is obtained by averaging the data under each training step with the previous 999 data. The simulations show that, although the curve fluctuates, the whole process tends to be flat and the algorithm converges.

The learning curve of the SAC algorithm is shown in Figure 4, including the training loss of the action and critic networks. It can be seen from the simulation results that the convergence speed of the loss function of the SAC algorithm is fast. Therefore, the algorithm can quickly learn the optimal strategy.

Figure 5A, B show the convergence curves of the delay, energy consumption, and objective function of the SAC algorithm to solve the aforementioned issues. Figure 5A shows the unsmoothed curve of the training process, illustrating the total delay sum_SAC_OTPD in seconds, the total energy consumption sum_SAC_OTEC in joules, and the reward value. Figure 5B is the convergence curve after smoothing in Figure 5A. The simulations show that the SAC algorithm converges quickly and can solve the aforementioned issues.

The aforementioned two algorithms can solve the issue in this paper. The SAC algorithm has a fast convergence rate because it scales the state characteristics before inputting data parameters into the model. There is no difference in orders of magnitude between

variables, which is conducive to optimizing the initial model. However, the convergence effect of the PPOTR algorithm is better because the algorithm is trained based on dynamic fitting data parameters. In this paper, the simulation is set to train once every 2,048 steps, so the convergence speed of the PPOTR algorithm is slow, but the convergence effect is good.

5.3 Performance comparison

In this part, this paper compares the PPOTR algorithm with the four benchmark algorithms in terms of delay, energy consumption, and reward value to verify the proposed algorithm's performance.

As shown in Figure 6, it represents the total delay of different policies under the same task data. Each scheme will converge to the optimal value with increased training times. Under the same computing task, the SAC algorithm converges faster, but the PPOTR algorithm converges better. When the task volume increases to 25 Mb, the proposed algorithm saves approximately 17.33%, 32.74%, 56.63%, and 32.63%, respectively, compared with the SAC algorithm, local computing, edge execution, and random computing of the time cost. This shows that the algorithm proposed in this paper can achieve better performance in terms of task processing delay.

Figure 7 illustrates the total energy consumption corresponding to different policies under the same task data. When the task data volume is 25 Mb, the energy consumption cost of edge execution calculation is about 2.3 times that of the PPOTR algorithm. The simulations show that the PPOTR algorithm saves approximately 25.79%, 77.53%, and 63.31%, respectively, compared with SAC, AllLocal, and Random of the energy consumption. The proposed algorithm achieves the lowest energy consumption cost.

Finally, this paper normalizes the delay and energy consumption and converts the objective function value into the reward value in DRL, as shown in Figure 8. The reward value is composed of delay and energy consumption, with these variables being highly coupled and interactive. Under the constraint conditions, the reward value is minimized to obtain the best task-unloading strategy. The reward values of the proposed algorithm in this paper were improved by 15%, 28%, 30%, and 44% compared to four baseline algorithms. Numerical comparative analysis provides strong evidence for the reliability of the algorithm and approach proposed in this paper. Compared with the four benchmark algorithms, the algorithm proposed in this paper is superior in terms of delay, energy consumption, and reward value. Therefore, this scheme can guarantee to minimize the energy consumption cost under the tolerable delay.

6 Conclusion

This paper investigates a joint optimization strategy for task offloading in the IoV edge computing network. In the IoV scenario, while considering the timeliness of task data and resource constraints, we constructed a model based on cooperative games and transformed it into a joint optimization issue. This paper models the optimization issue as a Markov game based on intelligent edge, game theory, communication, and DRL. The reward function is devised as the sum of delay and energy consumption. We adopted the PPO-based algorithm to solve the previously mentioned issue. Finally, the performance of the algorithm is verified using the simulation experiments. The numerical results show that, compared with SAC and other baseline schemes, this scheme can achieve stable convergence in the system environment and obtain the optimal reward value. This scheme minimizes the system cost and meets the development needs of the future IoV.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

LW: investigation, methodology, validation, writing–original draft, and writing–review and editing. WZ: writing–review and editing. HX: methodology, writing–original draft, and writing–review and editing. LL: writing–original draft. LC: writing–review and editing. XZ: writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Song C, Zhang M, Zhan Y, Wang D, Guan L, Liu W, et al. Hierarchical edge cloud enabling network slicing for 5g optical fronthaul. *J Opt Commun Networking* (2019) 11: B60–B70. doi:10.1364/JOCN.11.000B60

2. Liu B, Jia D, Wang J, Lu K, Wu L. Cloud-assisted safety message dissemination in vanet–cellular heterogeneous wireless network. *IEEE Syst J* (2017) 11:128–39. doi:10.1109/JSYST.2015.2451156

3. Ma Q, Liu Y-F, Huang J. Time and location aware mobile data pricing. *IEEE Trans Mobile Comput* (2016) 15:2599–613. doi:10.1109/TMC.2015.2503763

4. Zhou Y, Tian L, Liu L, Qi Y. Fog computing enabled future mobile communication networks: A convergence of communication and computing. *IEEE Commun Mag* (2019) 57:20–7. doi:10.1109/MCOM.2019.1800235

5. Porambage P, Okwuibe J, Liyanage M, Ylianttila M, Taleb T. Survey on multi-access edge computing for internet of things realization. *IEEE Commun Surv Tutorials* (2018) 20:2961–91. doi:10.1109/COMST.2018.2849509

6. Sookhak M, Yu FR, He Y, Talebian H, Sohrabi Safa N, Zhao N, et al. Fog vehicular computing: Augmentation of fog computing using vehicular cloud computing. *IEEE Vehicular Techn Mag* (2017) 12:55–64. doi:10.1109/MVT.2017.2667499

7. Xiao Y, Zhu C. Vehicular fog computing: Vision and challenges. In: 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops); March 13 2017; Atlanta, GA, USA (2017). p. 6–9.

8. Zhou C, Wu W, He H, Yang P, Lyu F, Cheng N, et al. Deep reinforcement learning for delay-oriented iot task scheduling in sagin. *IEEE Trans Wireless Commun* (2021) 20: 911–25. doi:10.1109/TWC.2020.3029143

9. Li L, Cheng Q, Xue K, Yang C, Han Z. Downlink transmit power control in ultra-dense uav network based on mean field game and deep reinforcement learning. *IEEE Trans Vehicular Techn* (2020) 69:15594–605. doi:10.1109/TVT.2020.3043851

10. Gu B, Chen Y, Liao H, Zhou Z, Zhang D. A distributed and context-aware task assignment mechanism for collaborative mobile edge computing. *Sensors* (2018) 18: 2423. doi:10.3390/s18082423

11. Zhou J, Wu F, Zhang K, Mao Y, Leng S. Joint optimization of offloading and resource allocation in vehicular networks with mobile edge computing. In: 2018 10th International Conference on Wireless Communications and Signal Processing (WCSP); October 18-20, 2018; Zhejiang, China (2018). p. 1–6.

12. Qi Y, Tian L, Zhou Y, Yuan J. Mobile edge computing-assisted admission control in vehicular networks: The convergence of communication and computation. *IEEE Vehicular Techn Mag* (2019) 14:37–44. doi:10.1109/MVT.2018.2883336

13. Zhang S, Chen J, Lyu F, Cheng N, Shi W, Shen X. Vehicular communication networks in the automated driving era. *IEEE Commun Mag* (2018) 56:26–32. doi:10.1109/MCOM.2018.1701171

14. Zhang K, Mao Y, Leng S, He Y, Zhang Y. Mobile-edge computing for vehicular networks: A promising network paradigm with predictive off-loading. *IEEE Vehicular Techn Mag* (2017) 12:36–44. doi:10.1109/MVT.2017.2668838

15. Cesarano L, Croce A, Martins LDC, Tarchi D, Juan AA. A real-time energy-saving mechanism in internet of vehicles systems. *IEEE Access* (2021) 9:157842–58. doi:10.1109/ACCESS.2021.3130125

16. Lixin L, Yan S, Cheng Q, Dawei W, Wensheng L, Wei C. Optimal trajectory and downlink power control for multi-type uav aerial base stations. *Chin J Aeronautics* (2021) 34:11–23. doi:10.1016/j.cja.2020.12.019

17. Zhao P, Tian H, Qin C, Nie G. Energy-saving offloading by jointly allocating radio and computational resources for mobile edge computing. *IEEE Access* (2017) 5: 11255–68. doi:10.1109/ACCESS.2017.2710056

18. Wang Y, Sheng M, Wang X, Wang L, Li J. Mobile-edge computing: Partial computation offloading using dynamic voltage scaling. *IEEE Trans Commun* (2016) 64: 1–4282. doi:10.1109/TCOMM.2016.2599530

19. Zhang K, Leng S, Peng X, Pan L, Maharjan S, Zhang Y. Artificial intelligence inspired transmission scheduling in cognitive vehicular communications and networks. *IEEE Internet Things J* (2019) 6:1987–97. doi:10.1109/JIOT.2018.2872013

20. Hou X, Ren Z, Wang J, Cheng W, Ren Y, Chen K-C, et al. Reliable computation offloading for edge-computing-enabled software-defined iov. *IEEE Internet Things J* (2020) 7:7097–111. doi:10.1109/JIOT.2020.2982292

21. Yang K, Shen C, Liu T. Deep reinforcement learning based wireless network optimization: A comparative study. In: IEEE INFOCOM 2020 - IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); 6-9 July 2020; Toronto, Ontario (2020). p. 1248–53.

22. Kong X, Duan G, Hou M, Shen G, Wang H, Yan X, et al. Deep reinforcement learning-based energy-efficient edge computing for internet of vehicles. *IEEE Trans Ind Inform* (2022) 18:6308–16. doi:10.1109/TII.2022.3155162

23. Li L, Cheng Q, Tang X, Bai T, Chen W, Ding Z, et al. Resource allocation for noma-mec systems in ultra-dense networks: A learning aided mean-field game approach. *IEEE Trans Wireless Commun* (2021) 20:1487–500. doi:10.1109/TWC.2020.3033843

24. Xia F, Ahmed AM, Yang LT, Ma J, Rodrigues JJ. Exploiting social relationship to enable efficient replica allocation in ad-hoc social networks. *IEEE Trans Parallel Distributed Syst* (2014) 25:3167–76. doi:10.1109/TPDS.2013.2295805

25. Abbas N, Zhang Y, Taherkordi A, Skeie T. Mobile edge computing: A survey. *IEEE Internet Things J* (2018) 5:450–65. doi:10.1109/JIOT.2017.2750180

26. Liu J, Wan J, Zeng B, Wang Q, Song H, Qiu M. A scalable and quick-response software defined vehicular network assisted by mobile edge computing. *IEEE Commun Mag* (2017) 55:94–100. doi:10.1109/MCOM.2017.1601150

27. Huang M, Liu W, Wang T, Liu A, Zhang S. A cloud–mec collaborative task offloading scheme with service orchestration. *IEEE Internet Things J* (2020) 7:5792–805. doi:10.1109/JIOT.2019.2952767

28. Chen M, Hao Y. Task offloading for mobile edge computing in software defined ultra-dense network. *IEEE J Selected Areas Commun* (2018) 36:587–97. doi:10.1109/JSAC.2018.2815360

29. Wang C, Yu FR, Liang C, Chen Q, Tang L. Joint computation offloading and interference management in wireless cellular networks with mobile edge computing. *IEEE Trans Vehicular Techn* (2017) 66:7432–45. doi:10.1109/TVT.2017.2672701

30. Xu X, Zhang X, Liu X, Jiang J, Qi L, Bhuiyan MZA. Adaptive computation offloading with edge for 5g-envisioned internet of connected vehicles. *IEEE Trans Intell Transportation Syst* (2021) 22:5213–22. doi:10.1109/TITS.2020.2982186

31. Xu X, Jiang Q, Zhang P, Cao X, Khosravi MR, Alex LT, et al. Game theory for distributed iov task offloading with fuzzy neural network in edge computing. *IEEE Trans Fuzzy Syst* (2022) 30:4593–604. doi:10.1109/TFUZZ.2022.3158000

32. Zou Y, Lin L, Zhang L. A task offloading strategy for compute-intensive scenarios in uav-assisted iov. In: 2022 IEEE 5th International Conference on Electronic Information and Communication Technology (ICEICT); 21-23 August 2022; Hefei, China. IEEE (2022). p. 427–31.

33. Wei S, Zou Y, Zhang X, Zhang T, Li X. An integrated longitudinal and lateral vehicle following control system with radar and vehicle-to-vehicle communication. *IEEE Trans Vehicular Techn* (2019) 68:1116–27. doi:10.1109/TVT.2018.2890418

34. Hieu NQ, Hoang DT, Luong NC, Niyato D. irdrc: An intelligent real-time dual-functional radar-communication system for automotive vehicles. *IEEE Wireless Commun Lett* (2020) 9:2140–3. doi:10.1109/LWC.2020.3014972

35. Zhao J, Lou C, Hao H. Intelligent vehicle communication and obstacle detection based on millimetre wave radar base station. In: 2021 International Wireless Communications and Mobile Computing (IWCMC); June 28 2021; Harbin, China. IEEE (2021). p. 1818–22.

36. Zhang Q, Li Z, Gao X, Feng Z. Performance evaluation of radar and communication integrated system for autonomous driving vehicles. In: IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS); 10-13 May 2021; New York City, NY, USA. IEEE (2021). p. 1–2.

37. Zhou Z, Yu H, Xu C, Zhang Y, Mumtaz S, Rodriguez J. Dependable content distribution in d2d-based cooperative vehicular networks: A big data-integrated coalition game approach. *IEEE Trans Intell Transportation Syst* (2018) 19:953–64. doi:10.1109/tits.2017.2771519

38. Zhou Z, Ota K, Dong M, Xu C. Energy-efficient matching for resource allocation in d2d enabled cellular networks. *IEEE Trans Vehicular Techn* (2016) 66:5256–68. doi:10.1109/tvt.2016.2615718

39. Xu Q, Su Z, Zheng Q, Luo M, Dong B. Secure content delivery with edge nodes to save caching resources for mobile users in green cities. *IEEE Trans Ind Inform* (2017) 14: 2550–9. doi:10.1109/tii.2017.2787201

40. Liwang M, Wang J, Gao Z, Du X, Guizani M. Game theory based opportunistic computation offloading in cloud-enabled iov. *Ieee Access* (2019) 7:32551–61. doi:10.1109/ACCESS.2019.2897617

41. Ye H, Li GY, Juang B-HF. Deep reinforcement learning based resource allocation for v2v communications. *IEEE Trans Vehicular Techn* (2019) 68:3163–73. doi:10.1109/TVT.2019.2897134

42. Peng H, Shen XS. Ddpg-based resource management for mec/uav-assisted vehicular networks. In: 2020 IEEE 92nd Vehicular Technology Conference (VTC2020-Fall); 4-7 October 2020; Victoria, Canada. IEEE (2020). p. 1–6.

43. Hu X, Wong K-K, Yang K. Wireless powered cooperation-assisted mobile edge computing. *IEEE Trans Wireless Commun* (2018) 17:2375–88. doi:10.1109/TWC.2018.2794345

44. Wei C-Y, Huang AC-S, Chen C-Y, Chen J-Y. Qos-aware hybrid scheduling for geographical zone-based resource allocation in cellular vehicle-to-vehicle communications. *IEEE Commun Lett* (2018) 22:610–3. doi:10.1109/LCOMM.2017.2784453

45. Hazarika B, Singh K, Biswas S, Li C-P. Drl-based resource allocation for computation offloading in iov networks. *IEEE Trans Ind Inform* (2022) 18:8027–38. doi:10.1109/TII.2022.3168292

46. Hazarika B, Singh K, Li C-P, Biswas S. Multi-agent drl-based computation offloading in multiple ris-aided iov networks. In: MILCOM 2022-2022 IEEE

Military Communications Conference (MILCOM); November-2 December 2022; Rockville, Maryland. IEEE (2022). p. 1–6.

47. Wang H, Li X, Ji H, Zhang H. Federated offloading scheme to minimize latency in mec-enabled vehicular networks. In: 2018 IEEE Globecom Workshops (GC Wkshps); 9-13 December 2018; Abu Dhabi, UAE (2018). p. 1–6.

48. Liu G, Wang W, Yuan J, Liu X, Feng Q. Elimination of accumulated error of 3d target location based on dual-view reconstruction. In: 2009 Second International Symposium on Electronic Commerce and Security. vol. 2; 22-24 May 2009; Nanchang City, China (2009). p. 121–4.

49. Sun Y, Guo X, Song J, Zhou S, Jiang Z, Liu X, et al. Adaptive learning-based task offloading for vehicular edge computing systems. *IEEE Trans Vehicular Techn* (2019) 68:3061–74. doi:10.1109/TVT.2019.2895593

50. Qi Y, Zhou Y, Liu Y-F, Liu L, Pan Z. Traffic-aware task offloading based on convergence of communication and sensing in vehicular edge computing. *IEEE Internet Things J* (2021) 8:17762–77. doi:10.1109/JIOT.2021.3083065

51. Song Z, Ma R, Xie Y. A collaborative task offloading strategy for mobile edge computing in internet of vehicles. In: 2021 IEEE 5th Advanced Information Technology, Electronic and Automation Control Conference (IAEAC). vol. 5; March 15-17, 2024; Chongqing, China (2021). p. 1379–84.

52. Pham Q-V, Mirjalili S, Kumar N, Alazab M, Hwang W-J. Whale optimization algorithm with applications to resource allocation in wireless networks. *IEEE Trans Vehicular Techn* (2020) 69:4285–97. doi:10.1109/TVT.2020.2973294

53. Wang Y, Fang W, Ding Y, Xiong N. Computation offloading optimization for uav-assisted mobile edge computing: a deep deterministic policy gradient approach. *Wireless Networks* (2021) 27:2991–3006. doi:10.1007/s11276-021-02632-z

54. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *nature* (2015) 518:529–33. doi:10.1038/nature14236

55. Chen Z, Yin B, Zhu H, Li Y, Tao M, Zhang W. Mobile communications, computing, and caching resources allocation for diverse services via multi-objective proximal policy optimization. *IEEE Trans Commun* (2022) 70:4498–512. doi:10.1109/TCOMM.2022.3173005

56. Qiu C, Hu Y, Chen Y, Zeng B. Deep deterministic policy gradient (ddpg)-based energy harvesting wireless communications. *IEEE Internet Things J* (2019) 6:8577–88. doi:10.1109/JIOT.2019.2921159

57. Haarnoja T, Zhou A, Abbeel P, Levine S. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor. In: International conference on machine learning (PMLR); 10-15 July 2018; Stockholm, Sweden (2018). p. 1861–70.

58. Choi Y-S, Shirani-Mehr H. Simultaneous transmission and reception: Algorithm, design and system level performance. *IEEE Trans Wireless Commun* (2013) 12: 5992–6010. doi:10.1109/TWC.2013.101713.121152

59. Kwak J, Kim Y, Lee J, Chong S. Dream: Dynamic resource and task allocation for energy minimization in mobile cloud systems. *IEEE J Selected Areas Commun* (2015) 33: 2510–23. doi:10.1109/JSAC.2015.2478718

60. Ye T, Lin X, Wu J, Li G, Li J. Toward dynamic computation offloading for data processing in vehicular fog based f-ran. In: 2019 IEEE Fourth International Conference on Data Science in Cyberspace (DSC); June 25 2019; Hangzhou, China. IEEE (2019). p. 196–201.

61. Liu Y, Wang S, Huang J, Yang F. A computation offloading algorithm based on game theory for vehicular edge networks. In: 2018 IEEE International Conference on Communications (ICC); May 20-24, 2018; Kansas City, MO, USA (2018). p. 1–6.

Exploring the dynamics of role transition of employees in family businesses through the evolutionary game theory

Linrong Zhang[1,2], Mengyun Wu[3]*, Shiyu Li[1], Ruiwen Liu[1] and Yuqing Zhu[1]

[1]School of Finance and Economics, Jiangsu University, Zhenjiang, China, [2]School of Economics and Management, Tongji University, Shanghai, China, [3]School of Accounting, Shanghai Lixin University of Accounting and Finance, Shanghai, China

This investigation delves into the dynamic optimization of the progression from outsider to insider status within the context of family businesses. Utilizing a dynamic game model that incorporates three agents—leaders, insiders, and outsiders—it conducts a rigorous examination of the optimal pathway for the status transition of outsider employees in family businesses. The paramount objective of this study is to generate theoretical insights that may inform the optimization of human resources management and thereby bolster the overall performance of family businesses. The key findings of the research are as follows: 1) Outsider employees necessitate support from both leaders and insiders for a successful elevation in ranks. A collaborative relationship with insiders significantly enhances their performance. 2) The harmonious functioning of the workplace demands concerted efforts from all parties—leaders, insiders, and outsiders. The upward mobility of outsiders is contingent upon synergistic cooperation amongst all stakeholders. 3) A myriad of factors such as potential costs, benefits, and favoritism heavily influence the degree to which leaders endorse the upward mobility of outsiders. 4) Encouraging outsiders to move up in the ranks can instigate a sense of urgency among insiders, serving as a deterrent against complacency. This urgency can act as a catalyst for insiders to enhance their performance and mitigate the perceived threat to their own status.

1 Introduction

Family businesses, representing a prevalent organizational form, contribute significantly to China's economic development by virtue of their quantity and substantial contribution to the Gross Domestic Product (GDP). Leadership, an essential organizational context factor, exhibits variations across different social and organizational settings, largely attributable to their unique cultural nuances. In this landscape, Differential Leadership emerges as a leadership style deeply rooted in the Chinese cultural milieu and social structure, wielding considerable influence over the survival and long-term growth of family businesses. This approach entails leaders categorizing employees into two groups, "insiders" and "outsiders", predicated on factors such as relational proximity, loyalty, and talent. This categorization often results in preferential treatment towards insiders in

management practices and resource allocation [1]. Research indicates that this insider-outsider dichotomy instigates an out-group favoritism among outsiders, fueling their desire to attain insider status [2]. Given the permeable boundaries between the two categories, outsiders are often motivated to strive for upward mobility, rendering the status transition a dynamic process. The transition from outsider to insider is a challenging journey, demanding consistent effort and a long-term commitment to surpass the established classification criteria. Outsiders typically employ an upward mobility strategy [3], exhibiting proactive attitudes in job performance and interpersonal interactions [4]. If sufficiently recognized and supported by leaders and insiders, their behavior can potentially enhance overall employee performance, mitigating conflicts and negative impacts [5]. Therefore, to facilitate the transition of outsiders to insider status, active collaboration between the two groups is encouraged. This dynamic optimization of upward mobility involves multiple agents and is a crucial aspect of maintaining equilibrium in the organizational structure.

The transition from outsider to insider status within an organization constitutes a multi-agent game, necessitating the reconciliation of divergent interests to achieve equilibrium. In this context, differential leadership assumes the pivotal role of decision-making and strategizing, acting as the organization's representative. Insiders, functioning as *de facto* leaders within their teams, establish substantial leadership relationships with outsiders [6,7]. The upward mobility of outsider employees to insider status mandates the collaborative involvement of leaders, insiders, and outsiders, thus establishing a dynamic interplay among them. To investigate this phenomenon, this study adopts dynamic game theory and identifies leaders, insider employees, and outsider employees as the three game agents. A tripartite dynamic game model is constructed to scrutinize the game relationships and strategic choices among these agents. The ultimate aim of this exploration is to elucidate the path optimization of outsider employees' transition to insider status. It was found that the mobility of external employees' status not only does not reduce

employee performance, but also can further improve employee performance based on the existing level, due to the efforts and input made by external employees, as well as the cooperation and interaction between them and internal employees. The insights gleaned from this study hold significant implications for the enhancement of human resources management strategies and the overall performance of family businesses.

2 The methods

In the context of this study, leaders, insider employees, and outsider employees are considered as participants in the transition from outsider to insider status. Each participant has a set of strategies: leaders can either encourage the status transition of outsiders (D_1) or refrain from doing so (D_2), with respective probabilities of x and $1-x$. Insider employees can accept the status transition of outsiders (E_1) or resist it (E_2), with respective probabilities of y and $1-y$. Outsider employees can actively engage in status transition (F_1) or participate passively (F_2), with respective probabilities of z and $1-z$. Here, $0 \le x$, y, $z \le 1$. The study presumes that the three agents are rational, each seeking to maximize their individual interests during the status transition process. Leaders aim to amplify team performance through member collaboration, enhancing leadership efficiency and personal benefits. Insider employees are primarily interested in improving their own performance, reaping benefits, and preserving their insider status. Outsider employees seek to boost their performance, gain benefits, and achieve upward mobility to become insiders.

Given the inherent incompleteness of information in practical game scenarios, the study assumes that despite having worked together for an extended period, leaders, insider employees, and outsider employees possess incomplete information about each other. As the *de facto* organizational representative, differential leadership assumes the role of decision-making. Insiders function as "colleague leaders," establishing factual leadership relations with
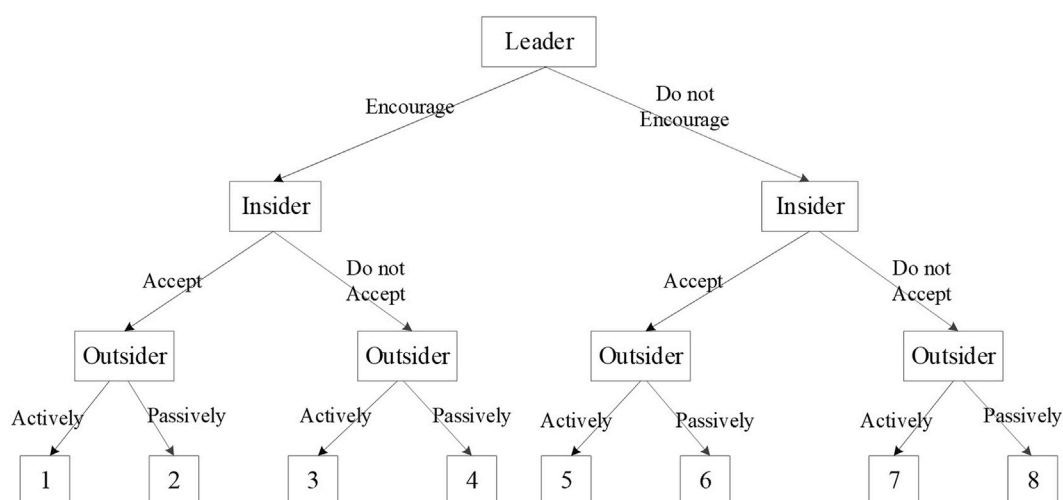

FIGURE 1
Three-agent game tree of leaders, outsiders, and insider employees.

TABLE 1 Payoff matrix for different strategy profile of leader, insider and outsider.

Strategy profile	Leader	Insider	Outsider
(Encourage, accept, actively)	$\triangle_{11} = \eta(V + \pi_1) - G$	$\triangle_{12} = S + \theta A_1 + G - B$	$\triangle_{13} = R + \zeta Q_1 - \lambda C_0$
(Encourage, accept, passively)	$\triangle_{21} = \eta V - G$	$\triangle_{22} = S + G$	$\triangle_{23} = R$
(Encourage, do not accept, actively)	$\triangle_{31} = \eta(V + \pi_1)$	$\triangle_{32} = S + \theta A_2$	$\triangle_{33} = R + \zeta Q_2 - \lambda D_0$
(Encourage, do not accept, passively)	$\triangle_{41} = \eta V$	$\triangle_{42} = S$	$\triangle_{43} = R$
(Do not encourage, accept, actively)	$\triangle_{51} = \eta(V + \pi_2)$	$\triangle_{52} = S + \theta A_1 - B$	$\triangle_{53} = R + \zeta Q_1 - \lambda C_0$
(Do not encourage, accept, passively)	$\triangle_{61} = \eta V$	$\triangle_{62} = S$	$\triangle_{63} = R$
(Do not encourage, do not accept, actively)	$\triangle_{71} = \eta(V + \pi_2)$	$\triangle_{72} = S + \theta A_2$	$\triangle_{73} = R + \zeta Q_2 - \lambda D_0$
(Do not encourage, do not accept, passively)	$\triangle_{81} = \eta V$	$\triangle_{82} = S$	$\triangle_{83} = R$

outsiders [6,7]. The study further presumes that the actions of the three parties occur sequentially rather than simultaneously. Subsequent actors can observe the actions of the preceding actors and infer the probabilities based on those actions. Thus, the study constructs a dynamic game model involving three parties with incomplete information. The game tree is illustrated in Figure 1, and the corresponding payoff parameters for the participants are delineated as follows:

(1) Leaders

In this study, V_1 denotes the leaders' payoff when they encourage outsider employees to transition, even though these outsiders choose not to participate in the transition but contribute to team performance. The conversion rate of team performance to personal benefits for leaders is represented by η, hence the leaders' payoff is ηV_1. Similarly, V_2 signifies the leaders' payoff when they do not encourage the outsiders' transition, and these employees still decide not to participate in the transition while contributing to team performance. Therefore, the payoff for leaders in this scenario is ηV_2. In both situations, we can assert that $V_1 = V_2 = V$. On the other hand, π_1 embodies the leaders' payoff when they encourage the transition of outsiders who actively participate in the process, thereby enhancing team performance. This additional performance improvement translates into supplementary benefits for leaders, denoted by $\eta\pi_1$. π_2 represents the leaders' payoff when they refrain from encouraging the transition of outsiders who, nonetheless, actively participate and contribute to the team's improved performance. In this case, the additional performance improvement converts into extra benefits for leaders, expressed as $\eta\pi_2$. Notably, $\pi_1 > \pi_2$. The term G denotes the additional cost incurred by leaders to persuade insider employees to support and collaborate with the policy of encouraging the outsiders' transition. This cost encompasses subsidies or appeasements directed towards insider employees, which consequently enhance their authorization level and psychological empowerment [8,9].

(2) Insider Employees

S_1 denotes the payoff for insider employees when they accept the status transition of outsider employees who choose not to participate in the transition. In this scenario, insiders receive their standard benefits. Similarly, S_2 signifies the payoff for insiders when they resist the status

transition of outsiders who also opt not to participate in the transition. In this instance, insiders also receive their typical benefits. In both situations, S_1 and S_2 can be equated as $S_1 = S_2 = S$. A_1 represents the payoff for insiders when they accept the transition of outsiders and cooperate with them by sharing resources. The active participation of outsiders in the transition results in an improvement in insiders' performance. This performance enhancement is converted into additional benefits for insiders at a rate of θ, thus yielding additional benefits of θA_1. In contrast, A_2 corresponds to the payoff for insiders when they do not accept the transition of outsiders, yet the outsiders actively participate in the transition and collaborate with insiders, thereby leading to enhanced performance. The additional benefits for insiders are θA_2, with the assumption that $A_1 > A_2$. Lastly, B encapsulates the potential cost incurred by insiders when they accept the outsiders' transition and cooperate with them. This cost may include resource sharing with outsiders, the investment of time and effort in collaboration, and the potential threat to insiders' status.

(3) Outsider Employees

In this framework, R designates the payoff for outsider employees when they abstain from participating in the status transition, thereby receiving their usual benefits. Q_1 signifies the payoff for outsider employees when insider employees accept the outsiders' status transition, and these outsiders actively engage in the transition, investing added time and effort to enhance performance. This additional performance improvement is converted into extra benefits for outsiders at a rate of ζ, culminating in benefits of ζQ_1. The performance improvement is positively correlated with the investment cost by outsider employees, which aligns with empirical research findings suggesting a positive association between work engagement and employee performance [10,11]. Thus, $Q_1 = f(C_1) = f(\lambda C_0)$. Q_2 represents the payoff for outsider employees when insider employees resist the status transition of outsiders, who still actively participate in the transition, thereby leading to performance enhancement. The additional benefits for outsiders are ζQ_2, and it is assumed that $Q_1 > Q_2$ and $Q_1 = Q_2 + Q_0$. Here, $Q_2 = f(C_2) = f(\lambda D_0)$. C_i encapsulates the cost and effort that outsider employees are willing to invest to actively participate in the status transition. Specifically, $C_1 = \lambda C_0$ and $C_2 = \lambda D_0$. The parameter $\lambda \in [0, 1]$ reflects the degree of outsider employees' out-group favoritism, varying from 0 to 1. The greater the degree of out-group favoritism,

the higher the cost that outsiders are willing to bear for achieving status mobility to insiders [12,13]. C_0 denotes the maximum cost that outsiders are willing to bear to participate in the transition when insiders accept the transition. Meanwhile, D_0 signifies the maximum cost that outsiders are willing to bear to participate in the transition when insiders do not accept the transition, with $C_0 \leqslant D_0$.

Table 1 presents the payoff matrices for the eight strategic options corresponding to the game tree illustrated in Figure 1.

3 Results

3.1 Model analysis and solution

Based on the game analysis described above, we can obtain the total expected payoff functions for the three major game agents: leaders, insider employees and outsider employees.

(1) The expected payoff for leaders is:

$$
\begin{aligned}
\sum_L = {} & xyz\Delta_{11} + xy(1-z)\Delta_{21} + x(1-y)z\Delta_{31} + x(1-y)(1-z)\Delta_{41} \\
& + (1-x)yz\Delta_{51} + (1-x)y(1-z)\Delta_{61} + (1-x)(1-y)z\Delta_{71} \\
& + (1-x)(1-y)(1-z)\Delta_{81}
\end{aligned}
\tag{1}
$$

(2) The expected payoff for insider employees is:

$$
\begin{aligned}
\sum_Z = {} & xyz\Delta_{12} + xy(1-z)\Delta_{22} + x(1-y)z\Delta_{32} + x(1-y)(1-z)\Delta_{42} \\
& + (1-x)yz\Delta_{52} + (1-x)y(1-z)\Delta_{62} + (1-x)(1-y)z\Delta_{72} \\
& + (1-x)(1-y)(1-z)\Delta_{82}
\end{aligned}
\tag{2}
$$

(3) The expected payoff for outsider employees is:

$$
\begin{aligned}
\sum_W = {} & xyz\Delta_{13} + xy(1-z)\Delta_{23} + x(1-y)z\Delta_{33} + x(1-y)(1-z)\Delta_{43} \\
& + (1-x)yz\Delta_{53} + (1-x)y(1-z)\Delta_{63} + (1-x)(1-y)z\Delta_{73} \\
& + (1-x)(1-y)(1-z)\Delta_{83}
\end{aligned}
\tag{3}
$$

Since dynamic games are sequential, backward induction is the fundamental method for solving equilibrium solutions in dynamic game models. Based on the previous assumptions, the sequence of the dynamic game is leaders, insider employees and outsider employees. Therefore, we first solve for the maximum expected payoff value of outsider employees, then substitute it into the expected payoff function of insider employees to obtain their maximum expected payoff value. Finally, we derive the maximum expected payoff value for leaders.

To determine the equilibrium solution for the maximum expected payoff of outsider employees, we first set the first derivative of Eq. 1 equal to zero:

$$
\begin{aligned}
\frac{d\sum_W}{dz} = {} & xy\Delta_{13} - xy\Delta_{23} + x(1-y)\Delta_{33} - x(1-y)\Delta_{43} + (1-x)y\Delta_{53} \\
& - (1-x)y\Delta_{63} + (1-x)(1-y)\Delta_{73} - (1-x)(1-y)\Delta_{83} = 0 \\
= {} & xy(\Delta_{13} - \Delta_{23}) + x(1-y)(\Delta_{33} - \Delta_{43}) + (1-x)y(\Delta_{53} - \Delta_{63}) \\
& + (1-x)(1-y)(\Delta_{73} - \Delta_{83}) = 0
\end{aligned}
\tag{4}
$$

$$
y = \frac{\lambda D_0 - \zeta(Q_1 - Q_0)}{\lambda(D_0 - C_0) + \zeta Q_0}
\tag{5}
$$

Equation 5 represents the probability of insider employees accepting the strategy of outsider employees' status movement to insiders, where their expected payoff is maximized.

To find the equilibrium solution for the maximum expected payoff of insider employees, we set the first derivative of Eq. 2 equal to zero:

$$
\begin{aligned}
\frac{d\sum_Z}{dy} = {} & xz\Delta_{12} + x(1-z)\Delta_{22} - xz\Delta_{32} - x(1-z)\Delta_{42} + (1-x)z\Delta_{52} \\
& + (1-x)(1-z)\Delta_{62} - (1-x)z\Delta_{72} - (1-x)(1-z)\Delta_{82} = 0 \\
= {} & xz(\Delta_{12} - \Delta_{32}) + x(1-z)(\Delta_{22} - \Delta_{42}) + (1-x)z(\Delta_{52} - \Delta_{72}) \\
& + (1-x)(1-z)(\Delta_{62} - \Delta_{82}) = 0
\end{aligned}
\tag{6}
$$

$$
x = \frac{[B - \theta(A_1 - A_2)]z}{G}
\tag{7}
$$

Equation 7 represents the probability of leaders encouraging outsider employees' status movement to insiders, where their expected payoff is maximized.

To find the equilibrium solution for the maximum expected payoff of leaders, we set the first derivative of Eq. 1 equal to zero:

$$
\begin{aligned}
\frac{d\sum_L}{dx} = {} & yz\Delta_{11} + y(1-z)\Delta_{21} + (1-y)z\Delta_{31} + (1-y)(1-z)\Delta_{41} \\
& - yz\Delta_{51} - y(1-z)\Delta_{61} - (1-y)z\Delta_{71} - (1-z)(1-y)\Delta_{81} = 0 \\
= {} & yz(\Delta_{11} - \Delta_{51}) + y(1-z)(\Delta_{21} - \Delta_{61}) + (1-y)z(\Delta_{31} - \Delta_{71}) \\
& + (1-y)(1-z)(\Delta_{41} - \Delta_{81}) = 0
\end{aligned}
\tag{8}
$$

$$
z = \frac{Gy}{\eta(\pi_1 - \pi_2)}
\tag{9}
$$

Equation 9 represents the probability of outsider employees actively and diligently participating in status movement, where their expected payoff is maximized.

By substituting Eq. 5 into Eq. 9, we obtain the probability of outsider employees actively participating in status movement under the condition of maximizing their payoff:

$$
z = \frac{G[\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_0]}
\tag{10}
$$

By substituting Eq. 10 into Eq. 7, we obtain the probability of leaders encouraging outsider employees' status movement under the condition of maximizing their payoff:

$$
x = \frac{[B - \theta(A_1 - A_2)][\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_0]}
\tag{11}
$$

Considering Eqs 7, 10 and 11, the equilibrium solution for the dynamic game is:

$$
(x^\star, y^\star, z^\star) = \left\{ \frac{[B - \theta(A_1 - A_2)][\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_O]}, \right.
$$
$$
\left. \frac{\lambda D_0 - \zeta(Q_1 - Q_0)}{\lambda(D_0 - C_0) + \zeta Q_O}, \frac{G[\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_O]} \right\}
\tag{12}
$$

3.2 Game equilibrium solution analysis

Under the equilibrium of the three-agent dynamic game, the probability of outsider employees actively participating in status movement to insiders is:

$$z^\star = \frac{G[\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_O]} \quad (13)$$

Firstly, to ensure $z^\star = \frac{G[\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_O]} > 0$, because $\pi_1 > \pi_2$, $\lambda(D_0 - C_0) + \zeta Q_O > 0$, so $\lambda D_0 - \zeta(Q_1 - Q_0) > 0$, $\lambda D_0 > \zeta(Q_1 - Q_0)$, $\lambda D_0 > \zeta Q_2$. This indicates that when insider employees do not accept the status movement of outsider employees, the cost λD_0 incurred by outsider employees to achieve status movement is greater than the benefits ζQ_2 they receive. The benefits are derived from the performance improvement resulting from the cost invested by outsider employees to achieve status movement, and for those outsiders with a higher degree of out-group favoritism λ, they are willing to invest more to achieve status movement. This further demonstrates that under differential leadership, the status movement of outsider employees requires acceptance and cooperation from insider employees.

From Eq. 10, it is evident that the probability z of outsider employees actively participating in status transition is directly proportional to G. This relationship implies that an increase in G corresponds to an increase in z. In practice, this suggests that in order to foster insider employees' support and collaboration with the strategy of encouraging outsider employees' status transition, leaders should extend additional subsidies and appeasements to insider employees. This action confers greater psychological authority to the insiders. This is because, in the highly personalized leadership atmosphere of China, leaders control and allocate resources within the organization. Insiders play the role of 'colleague leaders' within the organization or team, and they have the responsibility to support and execute the decisions of the leaders. If leaders provide larger subsidies and appeasement costs to insiders and increase their level of authorization, thereby reducing the concerns of insiders, insiders will be more proactive in responding to the policies of the leaders, accepting the status mobility of outsiders, and cooperating with them. This, in turn, reduces the costs and obstacles that outsiders face when striving for status mobility, making them more willing to invest in the long-term process of achieving their status mobility to insiders.

By solving the first derivative of z with respect to λ, we obtain $\frac{\partial z}{\partial \lambda} = \frac{\zeta G[D_0 Q_1 - C_0(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)(-\lambda C_0 + \lambda D_0 + \zeta Q_0)^2} > 0$. It can be observed that the probability z of outsider employees actively participating in status movement is a monotonically increasing function of λ. This means that as the degree of out-group favoritism λ of outsider employees, the probability of outsider employees actively participating in status movement also increases. This is consistent with reality, as individual attitudes determine behaviors. Therefore, only when outsider employees have a greater degree of out-group favoritism, they are more motivated to participate in the status movement.

Furthermore, by solving the first derivative of z with respect to C_0, we get $\frac{\partial z}{\partial C_0} = \frac{\lambda G[\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)(-\lambda C_0 + \lambda D_0 + \zeta Q_0)^2} > 0$. The probability of outsider employees actively participating in status movement is a monotonically increasing function of C_0. This indicates that

when insider employees accept the status movement of outsider employees, and thus reduce the obstacles faced by outsider employees in the status movement process, it increases the enthusiasm of outsider employees to participate in status movement. However, on the other hand, the realization of status movement for outsider employees is not an instantaneous process and requires significant time and long-term investment. Therefore, the higher the subjective willingness of outsider employees to invest and make efforts for status movement, the greater their likelihood of participating in status movement. Since employee performance is positively correlated with their effort invested in their work, outsider employees' willingness to invest more effort and cost for status movement leads to higher employee performance.

In the three-agent dynamic game equilibrium, the probability y of insider employees accepting the status movement of outsider employees is:

$$y^\star = \frac{\lambda D_0 - \zeta(Q_1 - Q_0)}{\lambda(D_0 - C_0) + \zeta Q_0} \quad (14)$$

Taking the derivative of y with respect to λ, we get: $\frac{\partial y}{\partial \lambda} = \frac{\zeta[D_0 Q_0 + (Q_1 - Q_0)(D_0 - C_0)]}{(-\lambda C_0 + \lambda D_0 + \zeta Q_0)^2} > 0$. This shows that the probability y of insider employees accepting the status movement of outsider employees is directly proportional to the degree of λ. In real life, when outsider employees have a stronger desire to become part of the insider employee group, the probability of the insider employees accepting the status movement of outsiders increases. This reflects the mutual influence between outsider employees, who are both the implementers of "colleague leader", and the decision-making of the "colleague leader".

Further, when we take the derivative of y with respect to C_0, we get: $\frac{\partial y}{\partial C_0} = \frac{\lambda[\lambda D_0 - \zeta(Q_1 - Q_0)]}{(-\lambda C_0 + \lambda D_0 + \zeta Q_0)^2} > 0$. It is evident that the probability y of insider employees accepting the status movement of outsider employees is a monotonically increasing function of C_0. This means that when insider employees accept the status movement of outsiders, the greater the cost and effort that outsider employees are willing to invest C_0, the higher the probability of insider employees accepting their movement. It because when insiders accept the status movement of outsiders, if outsider employees are willing to put in more effort and investment to their status movement, it not only improves their own performance but also demonstrates greater support for the organization, the team, the leader and especially their insider colleagues. It leads to more involvement in completing their work and actively assisting insider employees in overcoming work challenges. Such interactions between insider and outsider employees contribute to an increase in the performance of the insider employees, thereby increasing the likelihood of them accepting the status movement of outsider employees.

In order to ensure $x^\star = \frac{[B - \theta(A_1 - A_2)][\lambda D_0 - \zeta(Q_1 - Q_0)]}{\eta(\pi_1 - \pi_2)[\lambda(D_0 - C_0) + \zeta Q_O]} > 0$, it follows that $B - \theta(A_1 - A_2) > 0$, $B > \theta(A_1 - A_2)$. It implies that for insider employees, choosing to accept the status movement of outsider employees and cooperating with them may lead to an improvement in their performance level, but the net benefits derived from this improvement are not enough to compensate for the costs and threats they need to bear in accepting the status movement. Moreover, x is positively correlated with B, meaning that when insider employees have higher costs and face more threats in accepting the status

movement of outsider employees, they need stronger support from the leader to encourage the status movement. In this case, the leader needs to explicitly demonstrate their encouragement and support for the status movement of outsider employees, thereby exerting pressure on the insider employees. To encourage outsider employees' participation in status movement, the leader not only needs to apply pressure to the insiders to show their determination to encourage the movement of outsiders but also needs to provide them with subsidies and reassurance to alleviate their concerns and worries. Simultaneously, the leader's encouragement of outsider employees' status movement instills a sense of crisis and urgency to the insiders, keeping them vigilant throughout the long-term process of the status movement and maintaining and improving their performance.

From Eq. 11, we observe an inverse proportionality between x and A_1. This suggests that as insider employees accept the status transition of outsider employees and experience a substantial performance enhancement, the necessity for the leader to facilitate the status transition of outsider employees decreases. The rationale behind this is that the leader's facilitation of outsider employees' status transition, coupled with significant performance improvement and benefit increases for insider employees, results in minimal resistance to policy implementation from insiders. Consider that the status transition of outsider employees constitutes a long-term process. The sustained improvement in insider employees' performance not only yields additional benefits but also solidifies their status. This consolidation of status reduces insiders' perception of threat, thereby alleviating their concerns.

In this context, the derivative of x with respect to λ is $\frac{\partial x}{\partial \lambda} = \frac{\zeta [B_1 - \theta(A_1 - A_2)][Q_0 D_0 + (Q_1 - Q_0)(D_0 - C_0)]}{\eta(\pi_1 - \pi_2)(-\lambda C_0 + \lambda D_0 + \zeta Q_0)^2} > 0$, indicating that the probability x of the leader encouraging outsider employees' status movement is directly proportional to the degree of λ. In real-life, if outsider employees do not have out-group favoritism, the leader would be even less likely to provide opportunities for status movement. If outsider employees have a strong motivation to become insiders and if the leader perceives their desire, the probability of the leader encouraging outsider employees' status movement will increase.

4 Discussion

4.1 Conclusions

This study utilized a dynamic game methodology to establish an equilibrium among leaders, insider employees, and outsider employees. Findings underscore the positive impacts of effective status transition for outsider employees on the overall performance of the workforce. The research also highlighted the importance of synergy and alignment between leaders and employees for the successful status transition of outsiders. Variables such as the degree of outgroup favoritism, costs borne by employees, and leadership endorsement significantly affected the facilitation of status transition. The study offers valuable insights for future longitudinal inquiries and establishes a foundation for subsequent empirical research. Based on the dynamic game model

encompassing leaders, insider employees, and outsider employees, the study yielded key findings:

1) Effective status transition of outsider employees is contingent on the support from both leaders and insider employees. This is attributed to the cooperative efforts between outsider and insider employees during the status transition process, which can subsequently enhance the performance of both employee groups.
2) To achieve equilibrium among leaders, insider employees, and outsider employees, effective collaboration is essential. The status transition of outsider employees can only be facilitated when all stakeholders work in a synergistic manner.
3) The facilitation of status transition for outsider employees by leaders is influenced by several factors. These include potential costs and threats that insider employees may face when accepting the transition, net benefits gained by insiders from endorsing and cooperating with outsiders, and the degree of favoritism that outsiders have towards the insider group.
4) The role of crisis and urgency is crucial in this context. Leaders' facilitation of outsider employees' status transition creates a sense of crisis and urgency among insider employees, deterring complacency or opportunistic behavior. This sense of urgency motivates insider employees to enhance their performance and mitigate the perceived threat to their own status.

4.2 Practical significance

1) Maintain dynamic employee categorization. Leaders should maintain a dynamic employee categorization system and cultivate a fair and dynamic corporate culture. Emphasizing the advantages of differential leadership over time can have a long-term impact on improving employee performance. This approach can encourage outsider employees to accept differential treatment and find suitable justifications for this treatment. However, it is essential to ensure that accepting such treatment serves as a source of motivation for outsider employees.
2) Address the sense of crisis and urgency. Leaders should encourage outsider employees' status movement to instill a sense of crisis and urgency among insider employees. This will prevent insider employees from becoming complacent due to perceived preferential treatment from leaders. The encouragement of outsider employees' status movement, coupled with the motivation of outsider employees to improve their status, can further enhance employee performance and contribute to the sustainable development of the organization.
3) Promote synergy among leaders and employees. Leaders should maintain the dynamic categorization of insider and outsider employees, particularly in encouraging outsider employees' status movement. This can motivate outsider employees and also create a sense of crisis and urgency among insider employees. The long-term efforts and cooperation between outsider and insider employees during the status movement process can lead to mutual improvement in employee performance, thereby further enhancing overall employee performance.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

LZ: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. MW: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. SL: Investigation. RL: Writing–review and editing. YZ: Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Jiang D, Zhang W. Chinese differential leadership and subordinate effectiveness. *Indigenous Psychol Res Chin Societies* (2010) 33:109–77.

2. Zhang L. *The influence of family business differential leadership on employee performance and its mechanism*. China: Jiangsu University (2019).

3. Wu M, Zhang L, Imran M, Xu J, Yu R. Impact of differential leadership on innovative behavior of employees: a double-edged sword. *Soc Behav Personal* (2021) 49: 1–12. doi:10.2224/sbp.9746

4. Liu Y, Zhang Z, Zhao H, Liu L. The double-edged sword effects of differential leadership on deviant behavior. *Curr Psychol* (2022) 42:27888–900. doi:10.1007/s12144-022-03845-x

5. Liu N, Zhang H, Zhou J. Unraveling the effect of differential leadership on employee performance: evidence from China. *Front Psychol* (2023) 14:1081073. doi:10.3389/fpsyg.2023.1081073

6. Tao Y, Zhang J, Li L. Study on the influence of differential leadership on employees' prosocial organizational behavior. *China Ind Econ* (2016) 336:116–31.

7. Wang L. *Theoretical and empirical study on the effectiveness of differential leadership*. China: Northeast University of Finance and Economics (2013).

8. Mubarak F, Noor A. Effect of authentic leadership on employee creativity in project-based organizations with the mediating roles of work engagement and psychological empowerment. *Cogent Business Manag* (2018) 5:1. doi:10.1080/23311975.2018.1429348

9. Meyerson SL, Kline TJ. Psychological and environmental empowerment: antecedents and consequences. *Leadersh Organ Dev J* (2008) 29:444–60. doi:10.1108/01437730810887049

10. Kahn WA. Psychological conditions of personal engagement and disengagement at work. *Acad Manag J* (1990) 33:692–724. doi:10.5465/256287

11. Wang YF, Jin YX. The impact of proactive coping and work engagement on employee performance: the moderating effect of skills. *Sci Stud Sci Manag* (2016) 37: 158–71.

12. Donnenwerth GV, Foa UG. Effect of resource class on retaliation to injustice in interpersonal exchange. *J Personal Soc Psychol* (1974) 29:785–93. doi:10.1037/h0036201

13. Martin J, Harder JW. Bread and roses: justice and the distribution of financial and socioemotional rewards in organizations. *Soc Justice Res* (1994) 7:241–64. doi:10.1007/bf02334833

frontiers | Frontiers in Physics

Reward shaping using directed graph convolution neural networks for reinforcement learning and games

Jianghui Sang[1,2], Zaki Ahmad Khan[3], Hengfu Yin[1]* and Yupeng Wang[2]*

[1]Research Institute of Subtropical Forestry, Chinese Academy of Forestry, Hangzhou, China, [2]School of Computer Science and Engineering, Nanjing University of Science and Technology, Nanjing, China, [3]Department of Computer Science, University of Worcester, Worcester, United Kingdom

Game theory can employ reinforcement learning algorithms to identify the optimal policy or equilibrium solution. Potential-based reward shaping (PBRS) methods are prevalently used for accelerating reinforcement learning, ensuring the optimal policy remains consistent. Existing PBRS research performs message passing based on graph convolution neural networks (GCNs) to propagate information from rewarding states. However, in an irreversible time-series reinforcement learning problem, undirected graphs will not only mislead message-passing schemes but also lose a distinctive direction structure. In this paper, a novel approach called directed graph convolution neural networks for reward shaping φ_{DCN} has been proposed to tackle this problem. The key innovation of φ_{DCN} is the extension of spectral-based undirected graph convolution to directed graphs. Messages can be efficiently propagated by leveraging a directed graph Laplacian as a substitute for the state transition matrix. As a consequence, potential-based reward shaping can then be implemented by the propagated messages. The incorporation of temporal dependencies between states makes φ_{DCN} more suitable for real-world scenarios than existing potential-based reward shaping methods based on undirected graph convolutional networks. Preliminary experiments demonstrate that the proposed φ_{DCN} exhibits a substantial improvement compared to other competing algorithms on both Atari and MuJoCo benchmarks.

1 Introduction

Over the past few decades, game theory has utilized the concepts and methods of reinforcement learning (RL) to solve decision-making problems [1]; [2]; [3]. An RL problem can be seen as a game between individual decision-makers and the environment [4]; [5]. RL can be expressed as a Markov decision process (MDP) [6]; [7]. Through interaction between agents and the environment, RL is able to receive rewards and take actions to maximize those rewards. During the training of RL, agents often encounter situations where they cannot obtain rewards most of the time [8]; [9]. Providing rewards in a sparse environment makes learning difficult for agents [10]. It is important to note that RL has always been hampered by sparse rewards [11].

Reward shaping is a widely used technique to address the challenge of reward sparsity. The purpose of reward shaping is to guide agents in learning through providing artificially designed additional rewards. Nevertheless, artificially designed reward functions may result in agents learning non-optimal policies in certain situations. Therefore, potential-based reward shaping (PBRS) is proposed in literature [12]. In this manner, the optimal policy is maintained when the additional reward value can be expressed in the differential form of a potential function. Thus, PBRS effectively avoids the reward hacking problem while addressing the sparse reward issue. On the other hand, RL problems can be considered probabilistic inference problems in hidden Markov models, where forward–backward messages can be used for inference. According to existing research, the potential function is defined in the probabilistic inference view of RL. The probability of an optimal trajectory is usually defined as a potential function under a given state. Due to the complexity of computation, it is difficult to obtain the messages. As a consequence, reward shaping using graph convolution networks is developed since projections of functions on the eigenspace of the graph Laplacian produce smooth approximation with respect to the underlying state-space topology of the MDP.

Previous research relies on a traditional spectral-based GCN to leverage reward shaping [13], but often overlooks crucial temporal dependencies between states. Sami et al. [14] employed a recurrent neural network to record reward-shaping values at different times, but the issue of temporal dependencies is not addressed essentially. Their use of the undirected graph and symmetric Laplacian matrix resulted in messages being interfered with and discarded in different directions, which may lead to serious logical errors. From a macro perspective, there are indeed some tracks that are sequential and irreversible in the real world.

To tackle the aforementioned problem, we propose an approach termed directed graph convolution neural networks for reward shaping φ_{DCN}. Our approach extends spectral-based undirected graph convolution to a directed graph which is built on a state probability model of trajectory. The state transition matrix is approximated by a directed graph Laplacian in the process of reward shaping. Messages about states that are propagated on this directed graph Laplacian serve to learn potential functions. The incorporation of temporal dependencies between states makes φ_{DCN} more suitable for real-world scenarios than existing potential-based reward-shaping methods based on undirected graph convolutional networks. We have demonstrated that φ_{DCN} outperforms competitive baselines on both Atari [15] and [16] benchmarks.

The main contributions are summarized as follows.

- We implement reward shaping through the message-passing mechanism of directed graph neural networks for the first time, which is more in line with the logic of the real world.
- The stationary distribution of the classical directed graph the Markov chain builds on is not necessarily unique since the graph might not be necessarily irreducible and aperiodic. To counteract this, we added a PageRank-based teleportation back to each node.

- Experiments demonstrate that the performance of the proposed φ_{DCN} exceeds that of the baseline algorithm on the Atari and MuJoCo benchmark.

2 Related work

2.1 Reward shaping

Reward shaping is used to accelerate learning when the environment only provides sporadic, incomplete, or delayed rewards. Ng et al. [12] proposed an extended version called potential-based reward shaping. Its most acclaimed characteristic is its ability to ensure that the optimal policy remains unchanged, as supported [12]. On this basis, there are two development paths of potential-based reward shaping: a) potential-based advice [17]; [18] and b) dynamic potential-based reward shaping [19]. Potential-based advice adds state–action pairs into potential functions rather than individual states. It is possible to vary the potential energy function over time using the latter approach. [20] is a book about reward shaping, which provides a detailed summary of the methods, theories, and applications of reward shaping.

In addition to the potential-based reward-shaping methods, other outstanding research studies on reward shaping also embrace belief reward shaping, ethics shaping, and reward shaping via meta-learning. These works are different from ours as we utilize convolutional networks to leverage reward shaping. In this sense, we focus more on the message-passing mechanism in the network to learn the potential function.

2.2 Digraph convolution

Digraph convolution is a method for performing convolution operations on a directed graph, which aims to comprehensively analyze the topological structure of the graph and the feature information of nodes or edges. Compared to undirected graph convolutional networks, digraph convolutional networks have the advantage of better reflecting the directional relationships between nodes. [21] extended the graph convolutional kernel originally designed for an undirected graph to a directed graph by introducing a trainable binary gating mechanism. This enables the model to regulate information dissemination based on the directionality of edges. Moreover, Li et al. proposed a new neural network architecture that can directly process graph structured data [22], including graph convolution operations for directed graphs. [23] proposed a new graph convolution method called MixHop, which can simultaneously consider the information of all neighboring nodes and handle directed graphs.

Some GCNs are designed to adapt to directed graphs (digraphs) by looking for structural patterns and reformulating the graph [24]; [25]. In addition to their limitations, these methods rely on pre-defined structures and are not capable of handling complex structures. Similar to our approach, another approach [26] redefines the propagation scheme only for strongly connected digraphs. In contrast, our approach is universally applicable to digraphs, which is the most important difference.

FIGURE 1
Graphical model with the variable O_t.
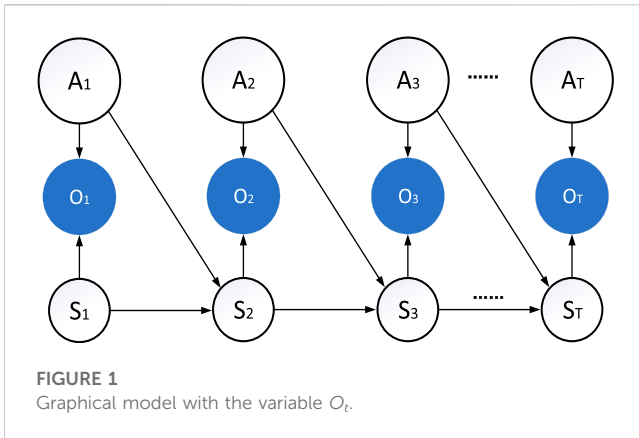
3 Background

3.1 Basic notions

As a mathematical expression, an MDP is represented by the tuple $\langle S, A, \gamma, r, P \rangle$, where S is the state space, A is the action space, r is the reward function, P is the transition probability matrix with $P(s' \mid s, a)$ giving the probability of transitioning to state s' when action a is taken at state s, and $\gamma \in (0, 1)$ is the discount factor. The state–action trajectory of a policy can be modeled by $\tau = (s_0, a_0, s_1, a_1 \ldots)$.

The policy π value function is defined as follows:

$$V_r^\pi(s) = \mathrm{E}_{\tau \sim \pi} \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s \right]. \tag{1}$$

The policy π action-value function is denoted as follows:

$$Q_r^\pi(s, a) = \mathrm{E}_{\tau \sim \pi} \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \mid s_0 = s, a_0 = a \right]. \tag{2}$$

The policy π expected discounted return is defined as follows:

$$J(\pi) = \mathrm{E}_{\tau \sim \pi} \left[\sum_{t=0}^{\infty} \gamma^t r(s_t, a_t) \right]. \tag{3}$$

Algorithms that learn from RL can determine the optimal policy $\pi^\star$:

$$\pi^\star = \mathrm{argmax}_\pi J(\pi). \tag{4}$$

Hence, given the initial state s_0 and its distribution $d(s_0)$, the gradient of the $J(\pi)$ over a parameterized policy π_θ can be expressed as

$$\frac{\partial J(\theta)}{\partial \theta} = \sum_s d(s; \theta) \sum_a \frac{\partial \pi_\theta(a \mid s)}{\partial \theta} Q_{\pi_\theta}(s, a),$$
$$d(s; \theta) = \sum_{s_0} d(s_0) \sum_{t=0}^{\infty} \gamma^t P^{\pi_\theta}(S_t = s \mid S_0 = s_0). \tag{5}$$

3.2 Control as inference

In existing studies, the RL problems have been translated directly into probabilistic inference problems [27]; [28]. So we use probability graph models to approximate RL. We then implement probabilistic inference through a message-passing mechanism. When the states are represented by nodes and the edges represent the transition probability between two different states, MDPs were considered probability graph models in previous RL research studies. As

shown in Figure 1, the RL structure approximates hidden Markov models (HMMs). Taking this into account, we introduce a binary variable $O = 0$ or 1 based on the action A_t and the state S_t. When $O_t = 1$, the state–action pair is optimal in time t.

In the probabilistic inference view of RL, the value function $V_\pi(S)$ can be approximately inferred through a message-passing mechanism. The forward–backward algorithm is an effective approach for performing inference in an HMM. A backward message is defined as $\beta(S_t, A_t) = P(O_{t:T} \mid S_t, A_t)$, and a forward message is defined as $\alpha(S_t, A_t) = P(O_{0:t-1} \mid S_t, A_t) P(S_t, A_t)$, where $O_{t:T}$ is the observation variable from time t to the end. Correspondingly, $O_{0:t-1}$ is the observation variable from the beginning to time $t - 1$. In the RL graph, given the current state S_t, the backward message reflects the probability that the current trajectory will lead to an optimal one in the future. This forward message indicates the probability of a past optimal trajectory for the current state S_t. By projecting maximum-entropy RL into probability space, the mapping function $f(\cdot)$ can be determined. $f(\cdot)$ maps rewards to a probability space by defining the distribution of this optimality variable as $P(O = 1 \mid S_t, A_t) = f(r(S_t, A_t))$.

As a result of recursion, the forward $\alpha(S_t, A_t)$ and backward $\beta(S_t, A_t)$ messages can be expressed as follows:

$$\alpha(S_t, A_t) = \sum_{S_{t-1}} \sum_{A_{t-1}} P(S_t \mid S_{t-1}, A_{t-1}) P(A_t) P(O_{t-1} \mid S_{t-1}, A_{t-1}) \alpha(S_{t-1}, A_{t-1}),$$
$$\beta(S_t, A_t) = \sum_{S_{t+1}} \sum_{A_{t+1}} P(S_{t+1} \mid S_t, A_t) P(A_{t+1}) P(O_t \mid S_t, A_t) \beta(S_{t+1}, A_{t+1}). \tag{6}$$

It should be noted that only the current state and reward are visible, not the action space. So the potential function of PBRS is designed in the state space. Once the actions are marginalized, we redefine the forward message $\alpha(S_t)$ and backward message $\beta(S_t)$ for learning potential functions.

According to [12], the optimal policy will remain unchanged after φ_{DCN} implements potential-based reward shaping. By replacing the original reward function $r(S_t, A_t)$ with a new reward function $R(S_t, A_t, S_{t+1})$, potential-based reward shaping can guarantee the optimal policy unchanged in RL:

$$R(S_t, A_t, S_{t+1}) = r(S_t, A_t) + F(S_t, S_{t+1}). \tag{7}$$

Here, $F(S_t, S_{t+1})$ is the shaping function calculated as follows:

$$F(S_t, S_{t+1}) = \gamma \Phi_{\alpha\beta}(S_{t+1}) - \Phi_{\alpha\beta}(S_t). \tag{8}$$

Literature reports have shown that the propagated messages can be used as potential functions [13]; [29]. Given the marginalized messages $\alpha(S_t)$ and $\beta(S_t)$, the potential function $\Phi(\cdot)$ is defined as

$$\Phi_{\alpha\beta}(S_t) = \alpha(S_t) \beta(S_t). \tag{9}$$

The potential function represents the probability of an entire trajectory being optimal. A high-return pathway is indicated by their likelihood.

4 Directed graph convolution neural networks for reward shaping

A DCN utilizes directed graph convolution to propagate messages $\alpha(S_t)$ and $\beta(S_t)$. Here, we illustrate digraph Laplacian,
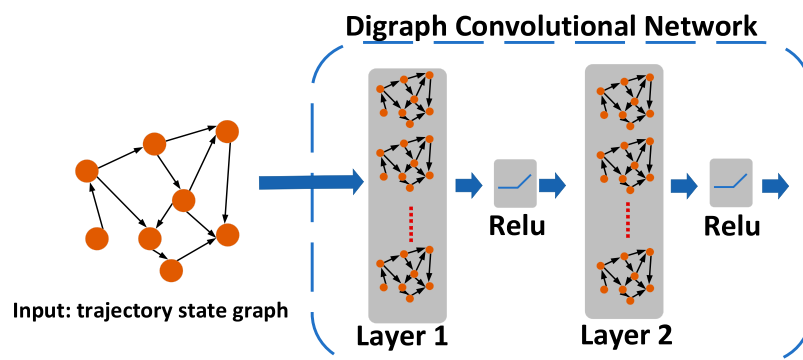
FIGURE 2
Network structure in φ_{DCN}. The trajectory in the MDP is defined as a graph, where each state is a node. The edge is the transition probability between two nodes. This probability directed graph is the input of φ_{DCN}.

TABLE 1 Configuration for experiments.

Hyperparameter	Value
Learning rate	2.5e-4
γ	0.99
Entropy coefficient	0.01
PPO steps	128
PPO clipping value	0.1
Mini batches	4
Processes	8
φ_{DCN}: μ	0.01
φ_{DCN}: λ	10
φ_{DCN}: δ	0.9

network structure, and loss function. In addition, we present a directed graph Laplacian. The structure of the directed graph convolutional neural network is shown in Figure 2.

4.1 Digraph Laplacian

The spectral-based graph convolution is first extended to the directed graph (digraph) by using the inherent relationship between the graph Laplacian and the stationary distribution of PageRank. Using the properties of Markov chains, we can solve the problem in digraphs using the internal relationship between graph Laplacian and PageRank. In a digraph $G = (V_d, E_d)$, a random walk on G is a Markov process with transition matrix $P_{rw} = D_d^{-1} A_d$, in which $D_d(i, i) = \sum_j A_d(i, j)$ is the diagonal degree matrix and A_d is the adjacency matrix. The stationary distribution of PageRank may not be unique if the transition matrix is not necessarily irreducible and aperiodic, especially when a graph contains isolated nodes in the periphery or can be formed into a bipartite graph. Irreducibility means that there exists a path between any two nodes in the network, while

aperiodicity means that the probability of returning to a node after a certain number of steps is not periodic. If a graph contains isolated nodes, then there is no path from those nodes to other nodes in the network, and the matrix is not irreducible. Similarly, if a graph can be formed into a bipartite graph, then there are no links between nodes in the same partition, which means that the matrix is not aperiodic.

Consequently, we slightly modify the random walk to PageRank which makes teleporting back to each node possible. In this way, the PageRank transition matrix P_{pr} is strictly irreducible and aperiodic, which is defined as $P_{pr} = (1 - \mu)P_{pr} + \frac{\mu}{n}1^{n \times n}$. It should be noted that the variable μ is small enough. Thus, according to the Perron–Frobenius theory [30], P_{pr} has a unique left eigenvector ξ_{pr} which is strictly positive with eigenvalue 1.

The row-vector ξ_{pr} corresponds to the stationary distribution of P_{pr}, and we have $\xi_{pr}(i) = \sum_{i, i \to j} \xi_{pr}(i) P_{pr}(i, j)$. According to the equation, the ξ_{pr} of node i is the sum of all incoming probabilities from node j to node i. Therefore, the ξ_{pr} and an undirected graph degree matrix D_u have similar properties. The digraph Laplacian using PageRank ϕ_{pr} in symmetric normalized format is defined as

$$L_{pr} = I - \frac{1}{2}\left(\Pi_{pr}^{\frac{1}{2}} P_{pr} \Pi_{pr}^{-\frac{1}{2}} + \Pi_{pr}^{-\frac{1}{2}} P_{pr}^T \Pi_{pr}^{\frac{1}{2}}\right), \qquad (10)$$

where we employ $\Pi_{pr} = \frac{1}{\|\xi_{pr}\|_1} Diag(\xi_{pr})$ to replace degree matrix D_u in an undirected graph. The definition is based on strongly connected digraphs, so it is not universally applicable. To deal with it, $\mu \to 0$ provides a generalized solution.

4.2 Loss function

One of the core characteristics of the GCN is that the message-passing mechanism is built on the graph Laplacian. Currently, PBRS is based on a traditional undirected GCN, where the undirected graph convolution is defined as $Z_u = \tilde{A}_u X W$. $\tilde{A}_u$ represents the normalized self-looped adjacency matrix of the undirected graph, and W represents the weight. The GCN and its variants require the undirected symmetric adjacency matrix A_u as the input. This not only aggregates features with incorrect weights but also discards
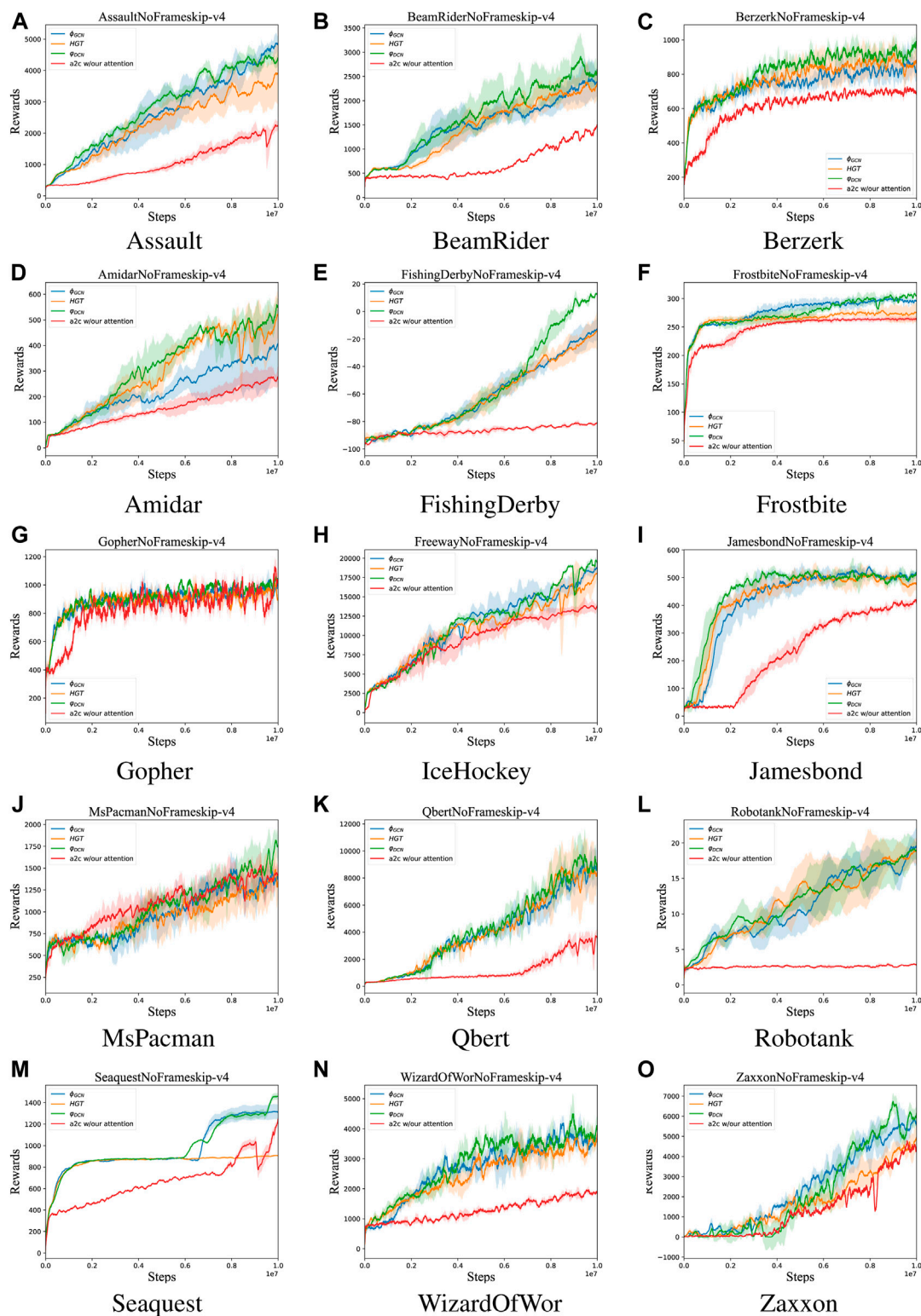
FIGURE 3
Comparison of rewards per episodes between a2c w/our attention, HGT, Φ_{GCN}, and φ_{DCN} on Atari. **(A)** Assault. **(B)** Beamrider. **(C)** Berzerk. **(D)** Amidar. **(E)** Fishing Derby. **(F)** Frostbite. **(G)** Gopher. **(H)**: IceHockey. **(I)** James Bond. **(J)** Ms. Pac-Man. **(K)** Q*bert. **(L)** Robo Tank. **(M)** SeaQuest. **(N)** Wizard Of Wor. **(O)** Zaxxon.

structures in different directions. In the MDP with temporal attributes, the undirected symmetric adjacency matrix A_u cannot be adopted. To approximate the transition matrix in the MDP, we

use the digraph Laplacian. Given a directed graph (digraph) $G = (V_d, E_d)$, the adjacency matrix is expressed as $A_d = \{0, 1\}^{n \times n}$, where the number of nodes is denoted by n. $X \in R^{n \times c}$ denotes the node features,

TABLE 2 Mean reward for 10 training processes on Atari. The better result is bolded.

	Alien	Amidar	Assault
ϕ_{GCN}	1,385.2	406.2	**4,845.1**
φ_{DCN}	**1,423.6**	**549.5**	4,403.4
	Beamrider	Berzerk	Breakout
ϕ_{GCN}	2,357.9	818.7	150.2
φ_{DCN}	**2,631.1**	**973.3**	**164.2**
	Demon Attack	Fishing Derby	Frostbite
ϕ_{GCN}	9,807.9	−13.1	296.3
φ_{DCN}	**12,448.6**	**12.5**	**305.2**
	Gopher	IceHockey	James Bond
ϕ_{GCN}	1,010.4	**−5.4**	511.2
φ_{DCN}	**1,045.8**	−5.7	**519.2**
	Ms. Pac-Man	Q*bert	Robo Tank
ϕ_{GCN}	1,422.3	**8,968.0**	**19.4**
φ_{DCN}	**1,754.4**	8,551.5	18.9
	SeaQuest	Wizard Of Wor	Zaxxon
ϕ_{GCN}	1,310.4	3,670.0	5,591.2
φ_{DCN}	**1,459.1**	**4,054.2**	**6,134.0**

with c being the number of features. Using Eq. 10 to define the digraph Laplacian, which is symmetric, we can then derive the digraph convolution definition as follows:

$$Z_d = \frac{1}{2}\left(\Pi_{pr}^{\frac{1}{2}}\tilde{P}_{pr}\Pi_{pr}^{-\frac{1}{2}} + \Pi_{pr}^{-\frac{1}{2}}\tilde{P}_{pr}^{T}\Pi_{pr}^{\frac{1}{2}}\right)XW, \qquad (11)$$

where $\tilde{P}_{pr}$ denotes a transition matrix with self-loops. Therefore, the message propagated by φ_{DCN} is as follows:

$$m_i = \text{ReLU}\left(W^T\sum_j\frac{1}{2}\left(\Pi_{pr}^{\frac{1}{2}}\tilde{P}_{pr}\Pi_{pr}^{-\frac{1}{2}} + \Pi_{pr}^{-\frac{1}{2}}\tilde{P}_{pr}^{T}\Pi_{pr}^{\frac{1}{2}}\right)_{ij}m_j\right), \qquad (12)$$

where the node j is a neighbor of node i and m_j is a message from node j.

Throughout φ_{DCN}, each state is represented by a node, while edges represent the transition probability between these states. Information about rewarding states is propagated through the message-passing mechanism of a directed graph convolutional network. In this paper, we propose a two-layer network as follows:

$$\varphi_{DCN} = \text{softmax}\left(\frac{1}{2}\left(\Pi_{pr}^{\frac{1}{2}}\tilde{P}_{pr}\Pi_{pr}^{-\frac{1}{2}} + \Pi_{pr}^{-\frac{1}{2}}\tilde{P}_{pr}^{T}\Pi_{pr}^{\frac{1}{2}}\right)\right.$$
$$\left.\text{ReLU}\left(\frac{1}{2}\left(\Pi_{pr}^{\frac{1}{2}}\tilde{P}_{pr}\Pi_{pr}^{-\frac{1}{2}} + \Pi_{pr}^{-\frac{1}{2}}\tilde{P}_{pr}^{T}\Pi_{pr}^{\frac{1}{2}}\right)XW^{(0)}\right)W^{(1)}\right). \qquad (13)$$

Then, we can express the loss function ℓ of φ_{DCN} as follows:

$$\ell = \ell_0 + \eta\ell_{prop}, \qquad (14)$$

where the parameter η is the weight assigned to the propagation loss ℓ_{prop}. Here, the supervised loss ℓ_0 is defined as the cross-entropy between the prediction result $\hat{Y}$ and the ground-truth label Y, denoted by the symbol $H(Y, \hat{Y})$. Y represents the probability that the path taken at the moment is the optimal trajectory. It is worth mentioning that $\hat{Y}$, found at the output of φ_{DCN}, is defined as a probability distribution $\varphi_{DCN}(S)$. According to the results of this study, we have calculated the supervised loss ℓ_0 as follows:

$$\ell_0 = H(P(O|S), \varphi_{DCN}(S)) = \sum_S P(O|S)\log(\varphi_{DCN}(S)). \qquad (15)$$

A propagation loss implemented as a recursive case is identified as ℓ_{prop} in Eq. 14. The recursive case of the message-passing mechanism can be implemented by the propagation loss ℓ_{prop} as follows:

$$\ell_{prop} = \sum_{vi,vj}\tilde{A}_{d_{v_i,v_j}}\left\|\varphi_{DCN}(X_{v_i}) - \varphi_{DCN}(X_{v_j})\right\|^2. \qquad (16)$$

4.3 Training

This paragraph describes the training process of φ_{DCN}. We propagate information about rewarding states through the message-passing mechanism of this directed graph convolution neural network. Then, the potential function $\Phi_{\alpha\beta}(\cdot)$ is learned on propagated messages $\alpha(S_t, A_t)$ and $\beta(S_t, A_t)$ (as in Eq. 9). Once the potential function $\Phi_{\alpha\beta}(\cdot)$ is learned, the new reward function $R(S_t, A_t, S_{t+1})$ is calculated to accelerate RL by replacing the original reward function $r(S_t, A_t)$.

In this case, the combined value function Q_{comb}^{π} of RL can be calculated using the sum of two value functions $Q_{comb}^{\pi}(s, a) = \delta Q^{\pi}(s, a) + (1 - \delta)Q_{\varphi}^{\pi}(s, a)$, where $Q^{\pi}(s, a) = \text{E}[\sum_t \gamma^t r(S_t, A_t)]$ is the original Q-value function and $Q_{\varphi}^{\pi}(s, a) = \text{E}[\sum_t \gamma^t r(S_t, A_t) + \gamma\varphi_{DCN}(S_{t+1}) - \varphi_{DCN}(S_t)]$ is the reward-shaped function. Two value functions can be weighted by the parameter δ. In this paper, we execute reward shaping φ_{DCN} on the underlying method PPO [31], which is a policy-based approach. The training process of φ_{DCN} is described in Algorithm 1.

```
1: Create an empty digraph G
2: for Episode = 0, 1, 2, ... do
3:     while t < T do
4:         Add transition (S_t, S_{t+1}) to digraph G
5:     end while
6:     if mod(Episode, N) then
7:         Train φ_DCN on the digraph G
8:     end if
9:     Q_comb^π = μQ^π + (1 − μ)Q_Φ^π
10:    Maximize E_π[∇ log π(A_t | S_t)Q_comb^π(S_t, A_t)]
11: end for
```

Algorithm 1. Directed graph convolution neural networks for reward shaping.

FIGURE 4
Delay rewards. Comparison of rewards per episode between HGT, PPO + RISE, Φ_{GCN}, and φ_{DCN} on MuJoCo. **(A)** HalfCheetah. **(B)** Hopper. **(C)** Walker2D. **(D)** Ant. **(E)** Swimmer. **(F)** Humanoid.
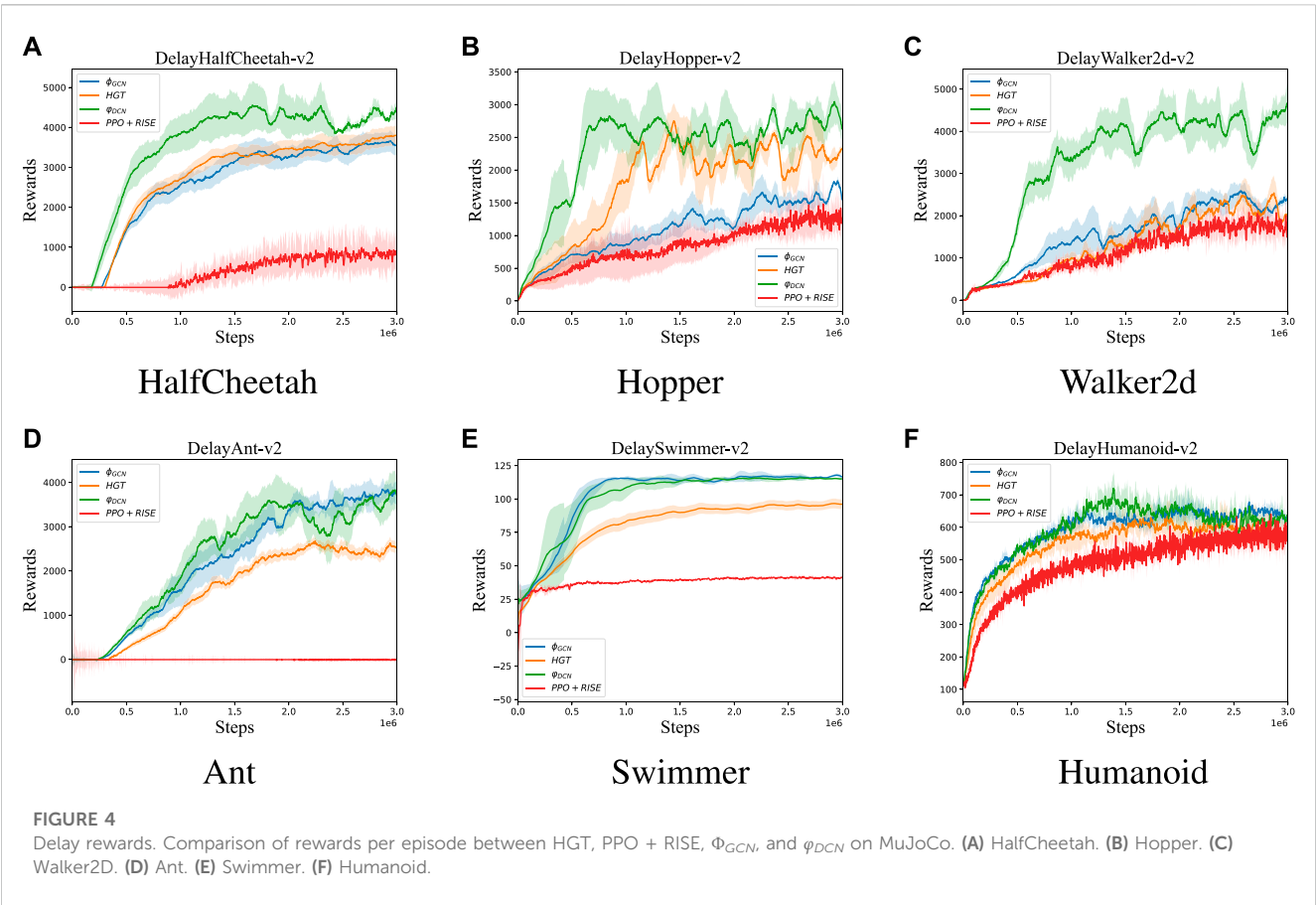
TABLE 3 Mean reward for 10 training processes on MuJoCo. The better result is bolded.

	HalfCheetah	Hopper	Walker2D
ϕ_{GCN}	3,543.9	1,550.4	2,413.5
φ_{DCN}	**4,496.8**	**2,630.6**	**4,620.9**
	Ant	Swimmer	Humanoid
ϕ_{GCN}	3,660.1	**116.6**	625.7
φ_{DCN}	**3,722.5**	115.0	**633.3**

5 Experiment

5.1 Experimental setup

In the experiment, the benchmarks from Atari and MuJoCo are adopted for evaluation.

Atari plays an important role in the field of RL [15]. Atari includes many classic games, such as Pac-Man, Space Invaders, Asteroids, and Pitfall. These games have become indelible symbols in the history of electronic games. Atari games are widely used as test platforms and benchmarks in the field of RL. Atari games have diversity, complexity, and challenges, covering various types of games, such as shooting, action, and adventure, at various difficulty levels. Therefore, they are used to evaluate the performance of the RL algorithm on complex tasks and help

researchers understand the advantages and limitations of the algorithm.

MuJoCo is a physics engine and simulator, which provides researchers and developers with an efficient and accurate physical simulation environment for training and evaluating RL algorithms [16]. MuJoCo provides a high-performance physics simulation engine that can simulate the dynamics and physical interactions of complex multi-joint robots and objects. This allows researchers to quickly test and verify RL algorithms in a simulation environment without the need for actual robots or real environments. MuJoCo supports various types of tasks and environmental settings, including robot control, object grasping, and movement [32]; [33,34]. It also allows users to customize the environment as needed to meet various research and application needs.

For the two categories of games, Atari and Mujoco, Table 1 shows the hyperparameters in the φ_{DCN}. The hardware components of our system include an RTX2070 GPU, CPU E5-2670V3, and 32 GB RAM.

We set ϕ_{GCN} as the baseline algorithm for comparison. In the literature [13], ϕ_{GCN} has been experimentally demonstrated to have better performance than others, such as the PPO [31], RND [35], ICM [36], and LIRPG [37]. For the purpose of comparing the proposed approach φ_{DCN} directly with other latest algorithms, PPO + RISE [38], a2c w/our attention [39], and HGT [29] are adopted as contenders. For a fair comparison, all competing algorithms use the default hyperparameters.

5.2 Experimental results on Atari

Due to its reactive and hard-exploration nature, the Atari benchmark is used for experiments. We repeat the experiment 10 times over ten million frames from each game in order to assess the applicability and effectiveness of the proposed φ_{DCN}.

In this experiment, we use potential-based reward shaping approaches HGT and ϕ_{GCN} as comparison methods, where HGT is an extended version of ϕ_{GCN}. We notice that HGT mines the logical correlations between states by enriching the propagated messages. In addition, we also compare a2c w/our attention as it is designed to improve exploration ability. However, a2c w/our attention does not guarantee the invariance of the optimal policy.

Figure 3 presents the mean rewards obtained from the 10 training processes using Atari tasks. In accordance with this, the proposed φ_{DCN} demonstrated good improvements in most games, including Ms Pac-Man, which displayed a 23% higher reward than the baseline ϕ_{GCN}. It is also observed that similar results are observed for Gopher, Demon Attack, and Amidar. Based on the results given in Table 2, the φ_{DCN} approach performs better than all other games in terms of improving learning performance by an average of 12.3%. It is concluded that reward shaping is enhanced by the message-passing mechanism in directed graph convolutional networks. A further analysis is conducted, with the results given in Table 2, which demonstrates empirically that the use of a directed graph Laplacian leads to performance improvement.

5.3 Experimental results on MuJoCo

In this experiment, we evaluate the performance of φ_{DCN} in continuous control tasks. We considered six standard environments in MuJoCo, namely, Ant, Humanoid, Hopper, Swimmer, Walker2D, and HalfCheetah. In order to increase the difficulty of the experiment, we used an environment with a delayed reward version that makes reward sparse. In this setting, agents only receive accumulated rewards for a period of time (20 steps), rather than receiving rewards for each step. Here, we choose ϕ_{GCN}, PPO + RISE, and HGT for comparison, which are considered strong state-of-the-art baselines.

According to Figure 4, when φ_{DCN} is trained on delayed reward environments, it is generally faster than baselines in all six MuJoCo environments. Particularly, φ_{DCN} achieves significant performance improvement in delayed reward environmental HalfCheetah, Hopper, and Walker2D. Although our approach φ_{DCN} performs at a similar level to ϕ_{GCN} which executes reward shaping through the message-passing mechanism on undirected graphs, we have surpassed the other two algorithms, namely, HGT and PPO + RISE. It is evidenced from Table 3 that our proposed approach has achieved much better rewards (37.1% higher) than the baseline ϕ_{GCN}. In this study, it is evident that the performance of the φ_{DCN} is improved in continuous control tasks. This suggests that the proposed approach

holds promise for accelerating RL in continuous control tasks when rewards are sparse.

5.4 Ablation analysis

An ablation analysis is conducted to determine how the directed graph Laplacian affects performance, as illustrated in Table 2 and Table 3. It should be noted that the bold one is the better one. The only difference between φ_{DCN} and ϕ_{GCN} is the graph Laplacian, where φ_{DCN} is the directed graph Laplacian and ϕ_{GCN} is the undirected graph Laplacian. According to this study, directed graph convolution networks have significantly improved performance in most of environments. Message passing with directional attributes can improve its performance. There is an improvement of 12.3% in the Atari experiment as compared to ϕ_{GCN}. The generalizability of φ_{DCN} is also demonstrated in several Atari tasks. Particularly in the continuous control tasks (MuJoCo), the performance is improved by an average of 37.1%.

6 Conclusion

Game theory utilizes the concepts and methods of RL to solve decision-making problems. However, the challenge of sparse rewards often exists in RL. Our proposed approach φ_{DCN} has been shown to be more effective in this issue as the message-passing mechanism of the directed graph Laplacian can be utilized to accelerate RL. In preliminary experiments conducted on Atari and MuJoCo, the proposed φ_{DCN} has demonstrated substantial improvement over conventional graph convolutional networks with an impressive increase of 12.3% and 37.1% compared to competing algorithms in terms of rewards per episode.

Despite this, there are still some shortcomings in certain aspects of φ_{DCN}, such as the high computational overhead of directed graph convolution operations. We are planning to conduct further research on this issue as the primary focus of our next project.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material; further inquiries can be directed to the corresponding authors.

Author contributions

JS: conceptualization, data curation, investigation, methodology, software, validation, and writing–original draft. ZA: writing–review and editing. HY: data curation, funding acquisition, project administration, resources, and writing–review and editing. YW: writing–review and editing.

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Song W, Sheng W, Li D, Wu C, Ma J. Modeling complex networks based on deep reinforcement learning. *Front Phys* (2022) 9:836. doi:10.3389/fphy.2021.822581

2. Grech L, Valentino G, Alves D, Hirlaender S. Application of reinforcement learning in the lhc tune feedback. *Front Phys* (2022) 10:929064. doi:10.3389/fphy.2022.929064

3. Guo H, Wang Z, Song Z, Yuan Y, Deng X, Li X. Effect of state transition triggered by reinforcement learning in evolutionary prisoner's dilemma game. *Neurocomputing* (2022) 511:187–97. doi:10.1016/j.neucom.2022.08.023

4. Ladosz P, Weng L, Kim M, Oh H. Exploration in deep reinforcement learning: a survey. *Inf Fusion* (2022) 85:1–22. doi:10.1016/j.inffus.2022.03.003

5. Han D, He Y. The reinforcement learning model with heterogeneous learning rate in activity-driven networks. *Int J Mod Phys C* (2023) 34:2350092. doi:10.1142/s0129183123500924

6. Han D, Li X. On evolutionary vaccination game in activity-driven networks. *IEEE Trans Comput Soc Syst* (2022) 10:142–52. doi:10.1109/tcss.2021.3137724

7. Meyn S. *Control systems and reinforcement learning*. Cambridge University Press (2022).

8. Zhu Y, Pang J-H, Tian F-B. Point-to-point navigation of a fish-like swimmer in a vortical flow with deep reinforcement learning. *Front Phys* (2022) 10:870273. doi:10.3389/fphy.2022.870273

9. Chen L, Wang C, Zeng C, Wang L, Liu H, Chen J. A novel method of heterogeneous combat network disintegration based on deep reinforcement learning. *Front Phys* (2022) 10:1021245. doi:10.3389/fphy.2022.1021245

10. Sang J, Wang Y. Graph convolution with topology refinement for automatic reinforcement learning. *Neurocomputing* (2023) 554:126621. doi:10.1016/j.neucom.2023.126621

11. Lu R, Jiang Z, Wu H, Ding Y, Wang D, Zhang H-T. Reward shaping-based actor–critic deep reinforcement learning for residential energy management. *IEEE Trans Ind Inform* (2022) 19:2662–73. doi:10.1109/tii.2022.3183802

12. Ng AY, Harada D, Russell S. Policy invariance under reward transformations: theory and application to reward shaping. Icml (*Citeseer*) (1999) 99:278–87.

13. Klissarov M, Precup D. *Reward propagation using graph convolutional networks* (2020). *arXiv preprint arXiv:2010.02474*.

14. Sami H, Bentahar J, Mourad A, Otrok H, Damiani E. Graph convolutional recurrent networks for reward shaping in reinforcement learning. *Inf Sci* (2022) 608:63–80. doi:10.1016/j.ins.2022.06.050

15. Bellemare MG, Naddaf Y, Veness J, Bowling M. The arcade learning environment: an evaluation platform for general agents. *J Artif Intelligence Res* (2013) 47:253–79. doi:10.1613/jair.3912

16. 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). p. 5026–33.

17. Harutyunyan A, Devlin S, Vrancx P, Nowé A Expressing arbitrary reward functions as potential-based advice. In: *Proceedings of the AAAI conference on artificial intelligence* (2015), 29.

18. Xiao B, Ramasubramanian B, Clark A, Hajishirzi H, Bushnell L, Poovendran R (2019). Potential-based advice for stochastic policy learning. In *2019 IEEE 58th conference on decision and control (CDC)* (IEEE), 1842–9.

19. Devlin SM, Kudenko D. Dynamic potential-based reward shaping. In: *Proceedings of the 11th international conference on autonomous agents and multiagent systems*. IFAAMAS (2012). 433–40.

20. Laud AD. *Theory and application of reward shaping in reinforcement learning*. University of Illinois at Urbana-Champaign (2004).

21. Tong Z, Liang Y, Sun C, Rosenblum DS, Lim A. *Directed graph convolutional network* (2020). *arXiv preprint arXiv:2004.13970*.

22. Li Y, Tarlow D, Brockschmidt M, Zemel R. *Gated graph sequence neural networks* (2015). *arXiv preprint arXiv:1511.05493*.

23. Li Q, Han Z, Wu X-M Deeper insights into graph convolutional networks for semi-supervised learning. In: *Proceedings of the AAAI conference on artificial intelligence* (2018), 32.

24. Kipf TN, Welling M. *Semi-supervised classification with graph convolutional networks* (2016). *arXiv preprint arXiv:1609.02907*.

25. Monti F, Otness K, Bronstein MM. Motifnet: a motif-based graph convolutional network for directed graphs. In *2018 IEEE data science workshop (DSW)*. (IEEE) (2018), 225–8.

26. Ma Y, Hao J, Yang Y, Li H, Jin J, Chen G. *Spectral-based graph convolutional network for directed graphs* (2019). *arXiv preprint arXiv:1907.08990*.

27. Ziebart BD, Maas AL, Bagnell JA, Dey AK. Maximum entropy inverse reinforcement learning. Aaai *(Chicago, IL, USA)* (2008) 8:1433–8.

28. Toussaint M, Storkey A. Probabilistic inference for solving discrete and continuous state markov decision processes. *Proc 23rd Int Conf Machine Learn* (2006) 945–52. doi:10.1145/1143844.1143963

29. Sang J, Wang Y, Ding W, Ahmadkhan Z, Xu L. *Reward shaping with hierarchical graph topology*. Pattern Recognition (2023), 109746.

30. Barker G, Schneider H. Algebraic perron-frobenius theory. *Linear Algebra its Appl* (1975) 11:219–33. doi:10.1016/0024-3795(75)90022-1

31. Schulman J, Wolski F, Dhariwal P, Radford A, Klimov O. *Proximal policy optimization algorithms* (2017). *arXiv preprint arXiv:1707.06347*.

32. Hu P, Chu X, Lv L, Zuo K, Ni T, Wang T, et al. An efficient and secure data collection scheme for predictive maintenance of vehicles. *Ad Hoc Networks* (2023) 146:103157. doi:10.1016/j.adhoc.2023.103157

33. Zhao R, Wang Y, Xiao G, Liu C, Hu P, Li H. A method of path planning for unmanned aerial vehicle based on the hybrid of selfish herd optimizer and particle swarm optimizer. *Appl Intelligence* (2022) 52:16775–98. doi:10.1007/s10489-021-02353-y

34. Zhao R, Wang Y, Xiao G, Liu C, Hu P, Li H. A selfish herd optimization algorithm based on the simplex method for clustering analysis. *The J Supercomputing* (2021) 77:8840–910. doi:10.1007/s11227-020-03597-0

35. Burda Y, Edwards H, Storkey A, Klimov O. *Exploration by random network distillation* (2018). *arXiv preprint arXiv:1810.12894*.

36. Pathak D, Agrawal P, Efros AA, Darrell T. Curiosity-driven exploration by self-supervised prediction. In: *International conference on machine learning*. PMLR (2017). p. 2778–87.

37. Zheng Z, Oh J, Singh S. On learning intrinsic rewards for policy gradient methods. *Adv Neural Inf Process Syst* (2018) 31.

38. Yuan M, Pun M-O, Wang D. Rényi state entropy maximization for exploration acceleration in reinforcement learning. *IEEE Trans Artif Intelligence* (2022) 4:1154–64. doi:10.1109/tai.2022.3185180

39. Wu H, Khetarpal K, Precup D. Self-supervised attention-aware reinforcement learning. *Proc AAAI Conf Artif Intelligence* (2021) 35:10311–9. doi:10.1609/aaai.v35i12.17235

Game-theoretic approach to understanding status transition dynamics and employee performance enhancement in organizations

Mengyun Wu[1,2], Yuqing Zhu[2]*, Ting Yang[2] and Yifan Xu[2]

[1]School of Accounting, Shanghai Lixin University of Accounting and Finance, Shanghai, China, [2]School of Finance and Economics, Jiangsu University, Zhenjiang, Jiangsu, China

To foster high-quality economic development, it is critical to not only enhance the "hard environment," such as infrastructure, but also to make significant strides in the "soft environment," such as the relationship between government and businesses. This study posits that the government, industry associations, and enterprises should collectively participate in fostering a "cordial and clean" government-business relationship. By resolving the equilibrium solution of the three-party game, it has been identified that achieving the goal of constructing a cordial and clean government-business relationship and aligning with the ambitions of enterprises necessitates policy guidance and a balance of interests among the government, industry associations, and enterprises. The research also contributes to the beneficial exploration of game theory, by constructing a network model from the perspective of public management and integrating it with the practice of local administrative reform. This integration is particularly relevant for industry associations, and their systematic analysis further enhances the practical applicability of the research.

1 Introduction

The government-business relationship is a timeless subject of discourse. It can be examined from a macroscopic perspective, where it pertains to the connection between the government and enterprises, or from a microscopic viewpoint, where it concerns the relationship between officials and entrepreneurs [1]. The proposition of a "cordial" and "clear" reformation of the existing government-business relationship provides innovative solutions to the challenges inherent in the current dynamics between enterprises and the government. It also outlines a direction for adjusting these existing relationships. Existing research on the government-business relationship primarily focuses on two levels: the macro and the micro. (1) Macro level: the relationship between the government and the market. Liberal economic theory, represented by thinkers such as Adam Smith, posits that the government's role should be limited to establishing and maintaining a strict law enforcement system, providing minimum public services like education, and minimally interfering with economic activities of enterprises. [2] further elaborates on this role from the perspective of the digital content industry, underscoring the government's various functions in the creation

of government-business relationships, including policy formulation and resource provision. Nationalistic economic theory, represented by [3], argues that a "visible hand" government can alleviate economic depression and unemployment through judicious and correct fiscal and monetary policy interventions. [4] also emphasizes that, considering the challenges to China's economic structure, the government should shift from a management focus to a service provider role, hence emphasizing the "cordial" aspect of the government-business relationship. (2) Micro level: the relationship between officials and entrepreneurs. The limited government theory, represented by [5], posits that the government is merely an agent entrusted by the people. It suggests that an alienation of the government-business relationship can occur, necessitating the establishment of a "limited government" subject to strict legal regulations. The distortion of the government-business relationship, evident in rent-seeking between the government and enterprises [6], underlines the importance of a "clear" relationship. On the other hand, public choice theory, represented by [7], views politics as essentially an exchange process. Here, the government, as a rational economic agent, acts as a 'shrewd buyer' providing paid services to enterprises. This results in enterprises receiving property protection and development opportunities, reflecting the "cordial" aspect of the government-business relationship. The shared interests and responsibilities between government and citizens reaffirm the importance of a "cordial" and "clear" relationship at the micro level.

In conclusion, a wealth of research on the construction, evolution, and governance of government-business relationships has been conducted by scholars both domestically and abroad [8]. Nonetheless, given that the concept of cordial and clean relationship between government and business is distinctively Chinese, most Chinese scholarly investigations have concentrated on its connotation or practical significance. A limited number of studies have attempted to develop network models within the realm of public management [9]. Furthermore, there is a scarcity of game-theoretic investigations that integrate with the practice of local administrative reform, particularly with regard to industry associations. These areas have not received systematic analysis and in-depth consideration. This study aims to redress this gap by embarking on a beneficial exploration of these less-explored aspects of government-business relations. The contributions of this paper are as follows: firstly, through the tripartite game model of government, industry association and enterprise, (1) it helps to promote the high-quality development of non-public economy; (2) It helps create a positive image of the government as friendly to the people and clean, and optimizes the business environment; (3) It helps to purify the political ecology and lead the social atmosphere. Secondly, the game model is used to simulate the influence of the cordial and clean relationship between government and business on the different decisions of the government, industry associations and enterprises, so as to provide practical support for narrowing the regional gap in the future and realizing the leapfrog development of the latecomer regions. Finally, the conclusions of this paper can enrich the relevant literature on the economic consequences of the pro-clean relationship between government and business and

regional economic development, and also provide valuable decision-making reference for realizing the strategic goal of coordinated development of regional economy and common prosperity of all people.

2 The methods

This study postulates that the government, industry associations, and enterprises are all participants in the construction of a cordial and clean relationship between government and business. For the government, there are two strategic alternatives: to construct "cordial" and "clean" policies that encourage enterprise behavior devoid of rent-seeking, or to abstain from constructing such policies. The probability of the government choosing to construct these policies is denoted by x, and the probability of refraining from construction is $1 - x$. In response to the government's policies, industry associations also have two strategic options: to supervise enterprises in promoting their actions, or to refrain from such supervision. The probability of choosing supervision is denoted by y, while the probability of not supervising is $1 - y$. In accordance with government policies, enterprises also have two strategic alternatives: proactive action or passive action. The probability of choosing a proactive action is symbolized by z, and the probability of choosing a passive action is $1 - z$. Simultaneously, this study assumes that the three participants in the dynamic game are rational economic actors, each pursuing the maximization of their self-interests in the process of constructing a "cordial" and "clean" policy for government-business relations. The government's interests primarily manifest in an increase in fiscal revenue and the maintenance of a positive governmental image. Trade associations' benefits are chiefly reflected in promoting economic development and enhancing government performance assessment. For businesses, the benefits mainly lie in increased operational income and profit. Given that the tripartite game involving the government, industry association, and enterprise is sequential, each subsequent actor can observe the behavior of the preceding actor and adjust their actions accordingly. Therefore, this study establishes a tripartite dynamic game model involving the government, industry association, and enterprise. The game tree is illustrated in Figure 1.

The corresponding benefit-cost parameters of participants in the dynamic game model are as: (1) The Government: B_1: Basic benefits for local governments (credibility, performance assessment and regional economic benefits, etc.); C_1: The cost of building a cordial and clean relationship between government and business; S_1: For the strong supervision of industry associations, the government subsidizes them; S_2: For the positive actions of enterprises, the government subsidizes them; R_1: For the positive actions of enterprises, the government gains benefits; R_2: For the strong regulation of industry associations, the government gains benefits. (2) Industry associations: B_2: Gains from participation in supervision by trade associations; C_2: Strong supervision costs for industry associations (construction labor costs, equipment costs and policy publicity, etc.); C_3: Weak regulatory costs for industry associations (policy advocacy); T_1: The additional benefits gained by industry associations for the establishment of cordial and clean relationship between
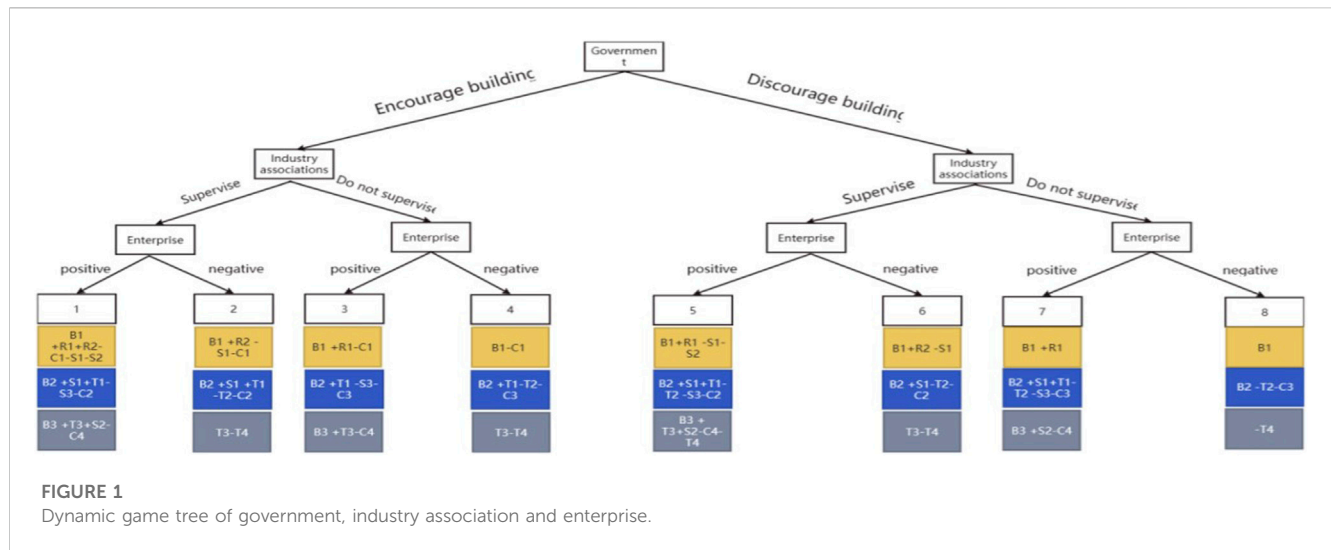
FIGURE 1
Dynamic game tree of government, industry association and enterprise.

TABLE 1 Income payment matrix under different strategy combinations of government, industry association and enterprise.

Id	Strategy combinations	Payoff matrix		
		The government	Industry associations	Enterprises
1	{construction, strong regulation, positive}	$B_1 + R_1 + R_2 - C_1 - S_1 - S_2$	$B_2 + S_1 + T_1 - S_3 - C_2$	$B_3 + T_3 + S_2 - C_4$
2	{construction, strong regulation, negative}	$B_1 + R_2 - S_1 - C_1$	$B_2 + S_1 + T_1 - T_2 - C_2$	$T_3 - T_4$
3	{construction, weak regulation, positive}	$B_1 + R_1 - C_1$	$B_2 + T_1 - S_3 - C_3$	$B_3 + T_3 - C_4$
4	{construction, weak regulation, negative}	$B_1 - C_1$	$B_2 + T_1 - T_2 - C_3$	$T_3 - T_4$
5	{no construction, strong regulation, positive}	$B_1 + R_1 - S_1 - S_2$	$B_2 + S_1 + T_1 - T_2 - S_3 - C_2$	$B_3 + T_3 + S_2 - C_4 - T_4$
6	{no construction, strong regulation, negative}	$B_1 + R_2 - S_1$	$B_2 + S_1 - T_2 - C_2$	$T_3 - T_4$
7	{no construction, weak regulation, positive}	$B_1 + R_1$	$B_2 + S_1 + T_1 - T_2 - S_3 - C_3$	$B_3 + S_2 - C_4$
8	{no construction, weak regulation, negative}	B_1	$B_2 - T_2 - C_3$	$-T_4$

government and business or the active actions of enterprises; T_2: The extra cost paid by industry associations for the government's failure to build a cordial and clean relationship between government and business or the passive actions of enterprises; S_3: For enterprises to actively act, industry associations to subsidize them. (3) Enterprises: B_3: The basic profit for the enterprise actively acts; C_4: The cost of doing something positive for the business; T_3: Build a cordial and clean relationship between government and business or strong supervision of industry associations, and the additional benefits obtained by enterprises; T_4: For the government does not build a cordial and clean relationship between government and business or the negative actions of enterprises, the additional losses caused by enterprises. Therefore, according to the assumptions of the model and the benefit-cost parameters mentioned above, there are eight strategy combinations among the three stakeholders of the government, industry associations and enterprises, which are 1 {construction, strong regulation, positive}, 2 {construction, strong regulation, negative}, 3 {construction, weak regulation, positive}, 4 {construction, weak regulation, negative}, 5 {no construction,

strong regulation, positive}, 6 {no construction, strong regulation, negative}, 7 {no construction, weak regulation, positive}, 8 {no construction, weak regulation, negative}, the payoff payment matrix of the three-game players under different strategy choices is shown in Table 1.

3 Results

3.1 Construction of dynamic equation

Based on the aforementioned game analysis, we can formulate the dynamic equation for the tripartite participants—government, industry association, and enterprise—as follows:

3.1.1 The government

(i) The expected return of constructing the cordial and clean relationship between government and business x

$$EX_1 = yz(B_1 + R_1 + R_2 - C_1 - S_1 - S_2) + y(1-z)(B_1 + R_2 - S_1 - C_1)$$
$$+ (1-y)z(B1 + R1 - C1) + (1-y)(1-z)(B_1 - C_1)$$
$$= -S_2yz + (R_2 - S_1)y + R_1z + B_1 - C_1$$

(ii) The expected return of $1 - x$ without constructing the "pro-clean" relationship

$$
\begin{aligned}
EX_2 &= yz(B_1 + R_1 - S_1 - S_2) + y(1-z)(B_1 + R_1 - S_1 - S_2) \\
&\quad + (1-y)z(B1 + R1) + (1-y)(1-z)(B_1) \\
&= -R_1yz + (R_1 - S_1 - S_2)y + R_1z + B_1
\end{aligned}
$$

(iii) Average expected return

$$
\begin{aligned}
EX &= xEX_1 + (1-x)EX_2 \\
&= x(EX_1 - EX_2) + EX_2 \\
&= x\big[(R_1 - R_2)yz + (R_2 - R_1 + S_2)y - C_1\big] - R_1yz \\
&\quad + (R_1 - S_1 - S_2)y + R_1z + B_1 \\
&= (R_1 - S_2)xyz + (R_2 - R_1 + S_2)xy - R1yz - C1x \\
&\quad + (R1 - S1 - S2)y + R_1z + B1
\end{aligned}
$$

(iv) Government replicates dynamic equations: $F(x) = \frac{dx}{dt} = x(EX_1 - EX)$

$$
\begin{aligned}
F(x) &= x(EX_1 - EX) \\
&= x(1-x)(EX_1 - EX_2) \\
&= x(1-x)\big[yz(R_2 - C_1) + y(1-z)(R_2 - R_1 + S_2 - C_1) \\
&\quad + (1-y)z(-C_1) + (1-y)(1-z)(-C_1)\big] \\
&= x(x-1)\big[C_1 + (R_1 - R_2 - S_2)y + (S_2 - R_1)yz\big]
\end{aligned}
$$

Let $F(x) = 0$, when $y = y^\star = \frac{C_1}{R_1z - S_2z + R_2 - R_1 + S_2}$, $F(x)$ is always 0. However, if $y \neq \frac{C_1}{R_1z - S_2z + R_2 - R_1 + S_2}$, we obtain that $x = 0$ and $x = 1$ are two possible equilibrium points for $F(x)$. According to the stability theorem of the replicated dynamic equation, When $F'(x) < 0$, this point is the stable strategy point of the evolutionary game. Take the derivative of F of x, $F'(x) = (1 - 2x)[C_1 + (R_1 - R_2 - S_2)y + (S_2 - R_1)yz]$. When $0 < y < \frac{C_1}{R_1z - S_2z + R_2 - R_1 + S_2}$, $F'(x)|_{x=0} < 0$, $F'(x)|_{x=1} > 0$, in this case, $x = 0$ is the stable strategy point of the evolutionary game. The government tends not to build a cordial and clean relationship between government and business. When $\frac{C_1}{R_1z - S_2z + R_2 - R_1 + S_2} < y < 1$, $F'(x)|_{x=0} > 0$, $F'(x)|_{x=1} < 0$, in this case, $x = 1$ is the stable strategy point of the evolutionary game. The government tends to build a cordial and clean relationship between government and business.

3.1.2 Industry associations

(i) Expected returns from implementing "strong regulation" y

$$
\begin{aligned}
EY_1 &= xz(B_2 + S_1 + T_1 - S_3 - C_2) + x(1-z) \\
&\quad \times (B_2 + S_1 + T_1 - T_2 - C_2) \\
&\quad + (1-x)z(B_2 + S_1 + T_1 - T_2 - S_3 - C_2) + (1-x)(1-z) \\
&\quad \times (B_2 + S_1 - T_2 - C_2) \\
&= (T_2 - T_1)xz + T_1x + (T_1 - S_3)z + B_2 + S_1 - T_2 - C_2
\end{aligned}
$$

(ii) Expected benefits of implementing "weak regulation" $1 - y$

$$
\begin{aligned}
EY_2 &= xz(B_2 + T_1 - S_3 - C_3) + x(1-z)(B_2 + T_1 - T_2 - C_3) \\
&\quad + (1-x)z(B_2 + S_1 + T_1 - T_2 - S_3 - C_3) + (1-x)(1-z) \\
&\quad \times (B_2 - T_2 - C_3) \\
&= (T_2 - T_1 - S_1)xz + T_1x + (S_1 - S_3 + T_1)z + B_2 + S_1 - T_2 \\
&\quad - C_3
\end{aligned}
$$

(iii) Average expected return

$$
\begin{aligned}
EY &= yEY_1 + (1-y)EY_2 \\
&= S_1xyz + (T_2 - T_1 - S_1)xz - S_1yz + T_1x + (S_1 + C_3 - C_2)y \\
&\quad + (S_1 - S_3 + T_1)z + B_2 - T_2 - C_3
\end{aligned}
$$

(iv) Government replicates dynamic equations: $F(y) = \frac{dy}{dt} = y(EY_1 - EY)$

$$
\begin{aligned}
F(y) &= y(EY_1 - EY) \\
&= y(1-y)(EY_1 - EY_2) \\
&= y(1-y)(-S_1z + S_1 - C_2 + C_3 + S_1xz)
\end{aligned}
$$

let $F(y) = 0$, when $z = z^\star = \frac{C_3 + S_1 - C_2}{S_1 - S_1X}$, $F(y)$ is always 0. However, if $z \neq \frac{C_3 + S_1 - C_2}{S_1 - S_1X}$, we obtain that $y = 0$ and $y = 1$ are two possible equilibrium points for $F(x)$. According to the stability theorem of the replicated dynamic equation, when $F'(y) < 0$, this point is the stable strategy point of the evolutionary game. Take the derivative of F of y, $F'(y) = (1 - 2y)(-S_1z + S_1 - C_2 + C_3 + S_1xz)$. When $0 < x < \frac{T_4y - B_3 - S_2 + C_4 - T_4}{S_2y + T_4y - S_2}$, $F'(z)|_{z=0} < 0$, $F'(z)|_{z=1} > 0$, $z = 0$ is the stable strategy point of the evolutionary game. That is, firms tend to act positively. When $0 < z < \frac{C_3 + S_1 - C_2}{S_1 - S_1X}$, $F'(y)|_{y=0} > 0$, $F'(y)|_{y=1} < 0$, $y = 1$ is the stable strategy point of the evolutionary game. That is, trade associations favor strong regulation. When $\frac{C_3 + S_1 - C_2}{S_1 - S_1X} < z < 1$, $F'(y)|_{y=0} < 0$, $F'(y)|_{y=1} > 0$, $y = 0$ is the stable strategy point of the evolutionary game. That is, trade associations favor weak regulation.

3.1.3 Enterprises

(i) The expected benefits of implementing "positive" z

$$
\begin{aligned}
EZ_1 &= xy(B_3 + T_3 + S_2 - C_4) + x(1-y)(B_3 + T_3 - C_4) \\
&\quad + (1-x)y(B_3 + T_3 + S_2 - C_4 - T_4) + (1-x)(1-y) \\
&\quad \times (B_3 + S_2 - C_4) \\
&= (S_2 + T_4 - T_3)xy + (T_3 - S_2)x + (T_3 - T_4)y + B_3 + S_2 - C_4
\end{aligned}
$$

(ii) Expected returns from implementing "negative" $1 - z$

$$
\begin{aligned}
EZ_2 &= xy(T_3 - T_4) + x(1-y)(T_3 - T_4) \\
&\quad + (1-x)y(T_3 - T_4) + (1-x)(1-y)(-T_4) \\
&= -T_3xy + T_3x + T_3y - T_4
\end{aligned}
$$

(iii) Average expected return

$$
\begin{aligned}
EZ &= zEZ_1 + (1-z)EZ_2 \\
&= (S_2 + T_4)xyz - T_3xy - S_2xz - T_4yz + T_3x + T_3y \\
&\quad + (B_3 + S_2 - C_4 + T_4)z - T_4
\end{aligned}
$$

(iv) Government replicates dynamic equations: $F(z) = \frac{dz}{dt} = z(EZ_1 - EZ)$

$$
\begin{aligned}
F(z) &= z(EZ_1 - EZ) \\
&= z(1-z)(EZ_1 - EZ_2) \\
&= z(1-z)\big[(S_2 + T_4)xy - S_2x - T_4y + B_3 - C_4 + T_4\big]
\end{aligned}
$$

let $F(z) = 0$, when $x = x^\star = \frac{T_4y - B_3 - S_2 + C_4 - T_4}{S_2y + T_4y - S_2}$, $F(z)$ is always 0. While $x \neq \frac{T_4y - B_3 - S_2 + C_4 - T_4}{S_2y + T_4y - S_2}$, we obtain that $z = 0$ and $z = 1$ are two possible equilibrium points for $F(z)$. According to the stability theorem of the replicated dynamic equation, when $F'(z) < 0$, this point is the stable strategy point of the evolutionary game. Take the derivative of F of z, $F'(z) = (1 - 2z)[(S_2 + T_4)xy - S_2x - T_4y + B_3 + S_2 - C_4 + T_4]$. When $0 < x < \frac{T_4y - B_3 - S_2 + C_4 - T_4}{S_2y + T_4y - S_2}$, $F'(z)|_{z=0} < 0$, $F'(z)|_{z=1} > 0$, $z = 0$ is the stable strategy point of the evolutionary game. That is, firms tend to act positively. When $\frac{T_4y - B_3 - S_2 + C_4 - T_4}{S_2y + T_4y - S_2} < x < 1$, $F'(z)|_{z=0} > 0$, $F'(z)|_{z=1} < 0$, in this case, $z = 1$ is the stable strategy point of the evolutionary game. That is, trade associations tend to behave negatively.

TABLE 2 Stability analysis of pure strategy equilibrium points.

Equalization point	λ_1	λ_2	λ_3	Stability condition
$(0,0,0)$	$-C_1$	$C_3 - C_2 + S_1$	$B_3 - C_4 + S_2 + T_4$	$C_3 < C_2 - S_1$, $B_3 < C_4 - S_2 - T_4$
$(1,0,0)$	C_1	$C_3 - C_2 + S_1$	$B_3 - C_4 + T_4$	Unstable point
$(0,1,0)$	$R_2 - R_1 - C_1 + S_2$	$C_2 - C_3 - S_1$	$B_3 - C_4 + S_2$	$R_2 < R_1 + C_1 - S_2$, $C_3 > C_2 - S_1$, $B_3 < C_4 - S_2$
$(0,0,1)$	$-C_1$	$C_3 - C_2$	$C_4 - B_3 - S_2 - T_4$	$C3 < C2$, $B_3 > C_4 - S_2 - T_4$
$(1,1,0)$	$C_1 + R_1 - R_2 - S_2$	$C_2 - C_3 - S_1$	$B_3 - C_4 + S_2 + T_4$	$C_1 < R_2 - R_1 + S_2$, $C_3 > C_2 - S_1$, $B_3 < C_4 - S_2 - T_4$
$(1,0,1)$	C_1	$C_3 - C_2 + S_1$	$C_4 - B_3 - T_4$	Unstable point
$(0,1,1)$	$R_2 - C_1$	$C_2 - C_3$	$C_4 - B_3 - S_2$	$C_1 > R_2$, $C_2 < C_3$, $C_4 < B_3 + S_2$
$(1,1,1)$	$C_1 - R_2$	$C_2 - C_3 - S_1$	$C_4 - B_3 - S_2 - T_4$	$C_1 < R_2$, $C_2 < C_3 + S_1$, $B_3 > C_4 - S_2 - T_4$

FIGURE 2
(A) System evolution path in the initial stage. **(B)** System evolution path during development stages. **(C)** System evolution path in the mature stage.
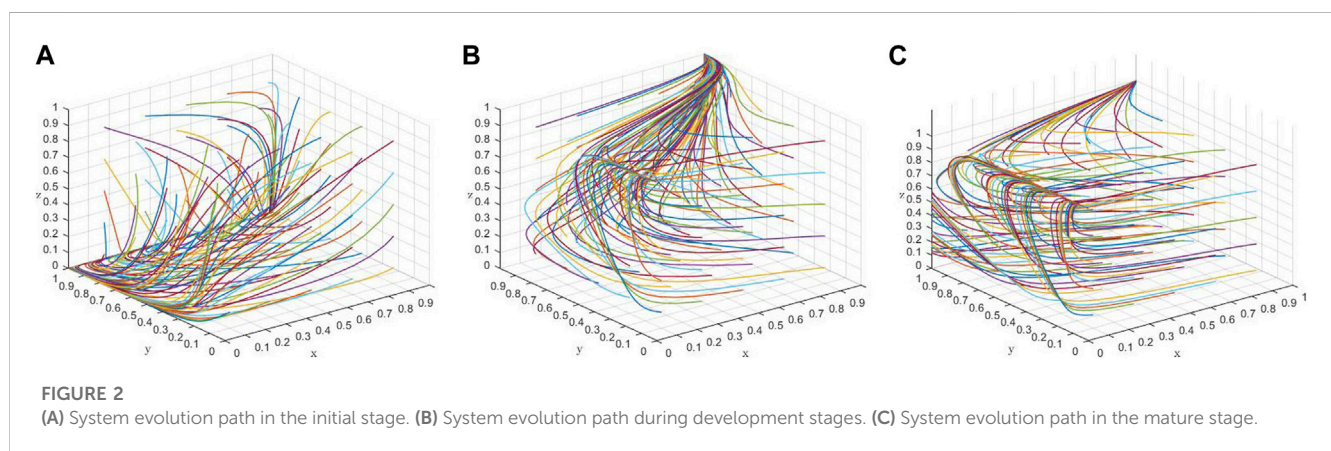
3.2 Strategy analysis

$$\begin{cases} F_x(x,y,z) = x(x-1)\left[C_1 + (R_1 - R_2 - S_2)y + (S_2 - R_1)yz\right] \\ F_y(x,y,z) = y(1-y)(-S_1z + S_1 - C_2 + C_3 + S_1xz) \\ F_z(x,y,z) = z(1-z)\left[(S_2 + T_4)xy - S_2x - T_4y + B_3 + S_2 - C_4 + T_4\right] \end{cases}$$

The Jacobian matrix of the system:

$$A = \begin{vmatrix} \dfrac{\partial F_x(x,y,z)}{\partial x} & \dfrac{\partial F_x(x,y,z)}{\partial y} & \dfrac{\partial F_x(x,y,z)}{\partial z} \\[2mm] \dfrac{\partial F_y(x,y,z)}{\partial x} & \dfrac{\partial F_y(x,y,z)}{\partial y} & \dfrac{\partial F_y(x,y,z)}{\partial z} \\[2mm] \dfrac{\partial F_z(x,y,z)}{\partial x} & \dfrac{\partial F_z(x,y,z)}{\partial y} & \dfrac{\partial F_z(x,y,z)}{\partial z} \end{vmatrix}$$

$$= \begin{vmatrix} \begin{array}{l} x(C_1 + R_1 y - R_2 y - S_2 y - R_1 yz + S_2 yz) \\ +(x-1)(C_1 + R_1 y - R_2 y - S_2 y \\ -R_1 yz + S_2 yz) \end{array} & x(1-x)(R_2 - R_1 + S_2 + R_1 z - S_2 z) & x(1-x)(R_1 y - S_2 y) \\ S_1 yz(1-y) & \begin{array}{l} -y(C_3 - C_2 + S_1 - S_1 z + S_1 xz) \\ -(y-1)(C_3 - C_2 + S_1 - S_1 z + S_1 xz) \end{array} & y(y-1)(S_1 - S_1 x) \\ z(1-z)(S_2 y - S_2 + T_4 y) & z(1-z)(S_2 x - T_4 + T_4 x) & \begin{array}{l} -z(B_3 - C_4 + S_2 + T_4 \\ -S_2 x - T_4 y + S_2 xy + T_4 xy) \\ -(z-1)(B_3 - C_4 + S_2 + T_4 \\ -S_2 x - T_4 y + S_2 xy + T_4 xy) \end{array} \end{vmatrix}$$

In this study, a total of thirteen equilibrium points have been derived. The stability of eight of these points, which represent pure strategy equilibria within the evolutionary system, is extensively analyzed. The difference between benefits and costs determines the choice of three subjects. It is more important to maintain the

relationship between government and business. According to the time sequence of the cordial and clean relationship between government and business, the evolution process is divided into three stages: the initial stage, the development stage and the mature stage which is shown in Table 2.

Initial stage (Figure 2A): The economic base determines the superstructure, and a good relationship between government and business plays an important role in promoting government, industry and enterprise [10]. In the early stage of the establishment of the cordial and clean relationship between government and business, most government departments lack the grasp of the policy and do not advocate the establishment of cordial and clean relationship between government and business; Industry associations do not have the case support to promote the success of government-entrepreneur cooperation, so they lack the consciousness to build a good relationship between government and business, and tend to choose the "weak regulation" strategy; Because enterprises fail to understand and grasp the essence of the new government-business relationship promptly and lack understanding of relevant policy publicity, they take negative actions. Therefore, this stage corresponds to the equilibrium point $A_1(0, 0, 0)$, We know from the table that it must be satisfied ① $C_2 < C_3 + S_1$: The cost of weak regulation is less than that of strong regulation; ② $B_3 < C_4 - S_2 - T_4$: The stable point is when the enterprise actively takes the condition that the profit is lower than the cost.

TABLE 3 The initial value of the parameter.

B_1	C_1	S_1	S_2	R_1	R_2	B_2	C_2	C_3	T_1	T_2	S_3	B_3	C_4	T_3	T_4
4	1.4	0.5	0.5	2	3	2	1.2	1.1	0.6	0.5	0.8	2	2.5	0.8	0.5

FIGURE 3
The impact of strategy proportion on system evolution results. **(A)** The initial choice probability of the three parties is 0.1 **(B)** The initial choice probability of the three parties is 0.5 **(C)** The initial choice probability of the three parties is 0.9.
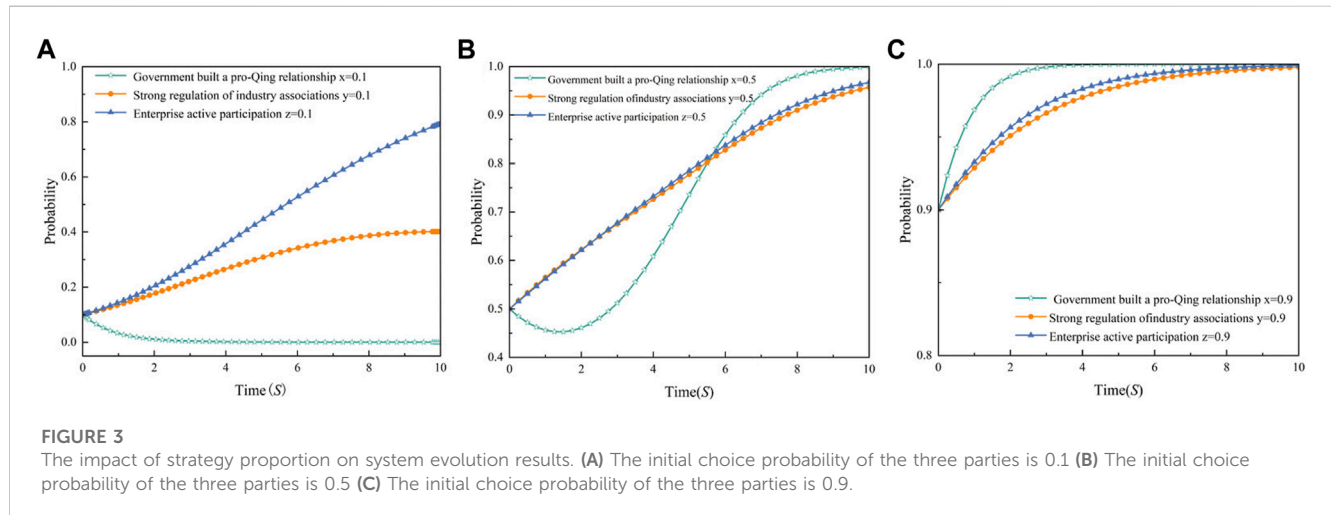
Development stage (Figure 2B): The government is inclined to actively build a new type of government-business relationship of "affinity, Qing and unity" and allocate more resources to enterprises and associations [11]. Industry associations began to improve the sound regulatory channels, motivated by government subsidies, keen to promote the connection between the government and entrepreneurs, and inclined to a "strong regulation" strategy; Entrepreneurs are gradually affected by the government's propaganda role, irregularly participate in the discussion, but the effect is still lower than expected, so they take negative actions. Therefore, this stage corresponds to the equilibrium point $A_5(1, 1, 0)$, We know from the table that it must be satisfied ① $C_1 < R_2 - R_1 + S_2$: The cost of government's construction of cordial and clean relationship between government and business is lower than the benefit of government's strong supervision of industry associations; ② $C_2 < C_3 + S_1$: The cost of strong supervision is greater than that of weak supervision; ③ $C_4 < B_3 + S_2 + T_4$: The stable point is when the enterprise actively takes the condition that the profit is lower than the cost. Because the evolution of the relationship between government and business has a historical stage, the traditional relationship between government and business is restricted by the social system in the development stage, which easily leads to the absence of government and market failure. The implementation of the policy of cordial and clean relationship between government and business through deepening the transformation of government functions, rationalizing the market order, so that the decisive role of the "invisible hand" and the strategic role of the "visible hand" are coordinated, laying the foundation for the return of healthy development of government-business relations. The government is highly motivated to build good relations and is willing to invest more energy and resources to build relations.

Maturity stage (Figure 2C): To realize the stable development of society, the government vigorously promotes the establishment of "pro-clean" government-business relationship [12]; Influenced by government policy welfare and entrepreneur donation behavior, industry associations can improve their own supervision and management ability, and tend to "strong supervision" strategy; Entrepreneurs are governed by relevant policies and profit maximization goals, and are willing to take positive actions. Therefore, this stage corresponds to the equilibrium point $A_8(1, 1, 1)$, We know from the table that it must be satisfied ① $C_1 < R_2$: The cost of constructing a cordial and clean relationship between government and business is lower than the benefit of strong supervision; ② $C_2 < C_3 + S_1$: The cost of strong supervision is less than that of weak supervision; ③ $C_4 < B_3 + S_2 + T_4$: The stable point is when the enterprise actively takes the condition that the profit is higher than the cost. The government pays more attention to sticking to the original intention. Take the initiative to be close to entrepreneurs in thought and emotion, have better platforms and support from industry associations, and actively guide entrepreneurs to communicate and exchange, so that they can get more sense of gain from the benign interaction with the government. In the mature stage, the market environment for fair competition is more optimized. The system and mechanism to prevent conflicts of interest will be more complete, and the norms of behavior for government and business exchanges are more perfect, so that the government can spend less energy to gain more trust.

3.3 Initial strategy simulation analysis

The influence of the initial strategy ratio on system evolution outcomes, given the initial parameter values, is explored in this research, as depicted in Table 3; Figure 3. The initial strategy selection ratios of the government, industry associations, and enterprises significantly affect the system's convergence speed [13]. A proximity between the initial strategy selection ratio and

the equilibrium point accelerates system convergence, highlighting the criticality of the initial strategy ratio in enabling the tripartite stakeholders to align in the direction of construction, robust supervision, and positivity. This suggests that industry associations should stimulate both the government and enterprises to actively engage in establishing a favorable pro-clean relationship. Concurrently, they should enhance their supervisory role through cooperative alliances, thus fostering a virtuous cycle within the system.

4 Discussion

Upon solving the game's equilibrium solution, it becomes evident that the realization of a "cordial" and "clean" government-business relationship necessitates policy guidance and interest balance between the government, industry associations, and enterprises. The following points elucidate this conclusion: 1. The extent of enterprise proactivity is directly proportional to the financial subsidies provided by the government to incentivize industry associations' supervision [14]. However, the costs incurred by enterprises due to active action, equipment transformation, and technological innovation exceed the opportunity loss from inaction. 2. Industry associations need to provide robust support and subsidies to enterprises to counterbalance profit reduction and cost increase [15]. Yet, they should avoid excessive investment in local enterprises to maintain a balanced relationship. 3. Enterprises require strong governmental support in response to profit reduction [16]. The government must grant sufficient subsidies to industry associations to overcome resistance and encourage policy compliance. Consequently, the government's macro-decisions guide industry associations, affecting the implementation of policies and subsequent enterprise behavior [17]. 4. Government governance should promote a rule of law, service orientation, and efficiency, providing high-quality public services [18]. The government's positive role in market system construction can be realized through strict power supervision and innovative service methods [19]. 5. Trade associations, based on market and enterprise, play a crucial role in market mechanism operation and industry interest realization [20]. Legal norms are necessary to ensure their healthy development, clarifying their legal status, power, responsibility, governance structure, and operation mechanism. Efficient undertaking of government service projects reflects the rule of law, specialization, and socialization of industry association governance [21].

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

MW: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. YZ: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. TY: Formal analysis, Writing—review & editing. YX: Conceptualization, Methodology, Writing–original draft, Writing–review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Ferasso M, Beliaeva T, Kraus S, Clauss T, Ribeiro-Soriano D. Circular economy business models: the state of research and avenues ahead. *Business Strategy Environ* (2020) 29:3006–24. doi:10.1002/bse.2554

2. Maisigova LA, Niyazbekova SU, Isayeva BK, Dzholdosheva TY. Features of relations between government authorities, business, and civil society in the digital economy. In: *Socio-economic systems: paradigms for the future*. Berlin, Germany: Springer (2021). p. 1385–91.

3. Keynes JM. *The general theory of employment, interest and money*. United Kingdom: Palgrave Macmillan (1936).

4. Moringiello JM. Dispossessing resident voice: municipal receiverships and the public trust. *U Mich JL Reform* (2019) 53:733. doi:10.36646/mjlr.53.4.dispossessing.moringiello

5. Locke J. *Two treatises of government*. United Kingdom: Awnsham Churchill (1689).

6. Yuxi H, Ahmad A, Talib ZM. The relationship between urban community collaborative governance and building resilience cities in zhengzhou city, henan province, China. *Int J Prof Business Rev* (2023) 8:e03561. doi:10.26668/businessreview/2023.v8i9.3561

7. Buchanan JM. The pricing of highway services. *Natl Tax J* (1952) 5:97–106. doi:10.1086/ntj41789931

8. Chen L, Tian Z. Current situation and prospect of the research on the relationship between government and business in the context of China. *Soft Sci* (2022) 36:125–31. doi:10.13956/j.ss.1001-8409.2022.04.17

9. Marques JC, Eberlein B. Grounding transnational business governance: a political-strategic perspective on government responses in the global south. *Regul governance* (2021) 15:1209–29. doi:10.1111/rego.12356

10. Zhao X, Yi C. Can the intervention of state-owned assets restrain the real enterprises from real to virtual? – on the regulating effect of pro-qing government business relationship. *Econ Manag* (2021) 43:61–74. doi:10.19616/j.cnki.bmj.2021.07.004

11. Jiang C, Jiang C, Zhen D. Can new political-business ties break the political resource curse of corporate innovation? *Soft Sci* (2021) 35:52–7. doi:10.13956/j.ss.1001-8409.2021.06.08

12. Liu N. Century-old construction of the relationship between government and enterprise under the leadership of the communist party of China: thought, practice and experience. *Financial Res* (2021) 47:4–18. doi:10.16538/j.cnki.jfe.20211016.201

13. Wei J, Zhao Q, Liu Y. New government-business relationship and robustness of corporate performance: evidence from listed companies. *J Manag Eng* (2021) 35:1–13. doi:10.13587/j.cnki.jieem.2021.04.001

14. Chen Y, Peng Z, Peng C, Xu W. Impact of new government–business relations on urban digital economy: empirical evidence from China. *Finance Res Lett* (2023) 58:104325. doi:10.1016/j.frl.2023.104325

15. Nugent JB. A new institutional perspective on business associations: filling a gap between firms and states in the dynamic analysis of richard day. *J Econ Behav Organ* (2023) 211:49–59. doi:10.1016/j.jebo.2023.04.015

16. Zhou J, Wang M. The role of government-industry-academia partnership in business incubation: evidence from new r&d institutions in China. *Tech Soc* (2023) 72:102194. doi:10.1016/j.techsoc.2022.102194

17. Wenqi D, Khurshid A, Rauf A, Calin AC. Government subsidies' influence on corporate social responsibility of private firms in a competitive environment. *J Innovation Knowledge* (2022) 7:100189. doi:10.1016/j.jik.2022.100189

18. Yu Y, Yin S, Zhang A. Clean energy-based rural low carbon transformation considering the supply and demand of new energy under government participation: a three-participators game model. *Energ Rep* (2022) 8:12011–25. doi:10.1016/j.egyr.2022.09.037

19. Mas JM, Gómez A. Social partners in the digital ecosystem: will business organizations, trade unions and government organizations survive the digital revolution? *Technol Forecast Soc Change* (2021) 162:120349. doi:10.1016/j.techfore.2020.120349

20. Tian Y, Wang Y, Xie X, Jiao J, Jiao H. The impact of business-government relations on firms' innovation: evidence from Chinese manufacturing industry. *Technol Forecast Soc Change* (2019) 143:1–8. doi:10.1016/j.techfore.2019.02.007

21. An H, Chen Y, Luo D, Zhang T. Political uncertainty and corporate investment: evidence from China. *J Corporate Finance* (2016) 36:174–89. doi:10.1016/j.jcorpfin.2015.11.003

Double-edged sword role of reinforcement learning based decision-makings on vaccination behavior

Jia-Qian Kan[1], Feng Zhang[2] and Hai-Feng Zhang[2]*

[1]School of Information and Network Engineering, Anhui Science and Technology University, Bengbu, China, [2]School of Mathematical Science, Anhui University, Hefei, China

Pre-emptive vaccination has been proven to be the most effective measure to control influenza outbreaks. However, when vaccination behavior is voluntary, individuals may face the vaccination dilemma owing to the two sides of vaccines. In view of this, many researchers began to use evolutionary game theory to model the vaccination decisions of individuals. Many existing models assume that individuals in networks use the Fermi function based strategy to update their vaccination decisions. As we know, human beings have strong learning capability and they may continuously search for the optimal strategy based on the surrounding environments. Hence, it is reasonable to use the reinforcement learning (RL) strategy to reflect the vaccination decisions of individuals. To this end, we here explore a mixed updating strategy for the vaccination decisions, specifically, some individuals called intelligent agents update their vaccination decisions based on the RL strategy, and the other individuals called regular agents update their decisions based on the Fermi function. We then investigate the impact of RL strategy on the vaccination behavior and the epidemic dynamics. Through extensive experiments, we find that the RL strategy plays a double-edged sword role: when the vaccination cost is not so high, more individuals are willing to choose vaccination if more individuals adopt the RL strategy, leading to the significant suppression of epidemics. On the contrary, when the vaccination cost is extremely high, the vaccination coverage is dramatically reduced, inducing the outbreak of the epidemic. We also analyze the underlying reasons for the double-edged sword role of the RL strategy.

KEYWORDS

epidemic spreading, vaccination game, reinforcement learning strategy, fermi function, double-edged sword role

1 Introduction

The spreading of large-scale epidemics, such as the Severe Acute Respiratory Syndrome (SARS), Avian influenza and Corona Virus Disease 2019 (COVID-19), not only seriously endangers human health, but also causes huge economic losses. Therefore, how to develop effective strategies to suppress the spreading of epidemics has always been an important issue. It has been proven that vaccination is the most successful intervention against the spread of epidemics, increasing life expectancy, and decreasing morbidity [1,2]. When considering the voluntary vaccination principle, an individual's vaccination decision may depend on the perceived risk of infection, cost of infection, cost of vaccination, and the vaccination behaviors of other individuals [3–5]. Thus, whether to take vaccination or not

represents a dilemma: vaccination protects not only those who are vaccinated but also their neighbors. In this case, many others in the community can also be benefited, so they have less incentive to be vaccinated. This scenario naturally leads to the "free-riding" problem commonly observed in public goods studies [6–8].

Inspired by these facts, researchers have investigated the impacts of vaccination behaviors on the epidemiological models within the game-theoretical framework [9–12]. For instance, Bauch *et al.* [9,10] analyzed the collective behavior of voluntary vaccination for various childhood diseases within a game-theoretic framework, and found that this voluntary strategy can not lead to the group-level optimum due to the risk perception pertaining to the vaccine and the effect of "herd immunity". The imitation dynamics inherent in the strategy-updating process was considered in the game-based vaccination model in Ref. [13], where the oscillations of vaccine uptake can emerge under some specific conditions, such as the change of disease prevalence or a high perceived risk of vaccine. Vardavas *et al.* [14] studied the effects of voluntary vaccination on the prevalence of influenza based on a minority game, and they demonstrated that severe epidemics could not be prevented unless vaccination programs offer incentives.

Since complex network provides an ideal and effective tool to describe the spreading of epidemics among populations, more works begun to study the voluntary vaccination behaviors within the network science framework [15–18]. For example, Perisic *et al.* studied the interplay of epidemic spreading dynamics and individual vaccination behavior on social contact networks. Compared to the homogeneously mixing model, they observed that increasing the neighborhood size of the contact network can eliminate the disease if individuals decide whether to vaccinate by accounting for infection risks from neighbors [19]. Mbah *et al.* considered the effects of both imitation behavior and contact heterogeneity on vaccination coverage and disease dynamics, and they found that the imitation behavior may impede the eradication of infectious diseases [20]. Fu *et al.* developed a network-based model to explore the effects of individual adaptation behavior and network structure on vaccination coverage as well as final epidemic size. Their findings indicate that the network structure can improve vaccination coverage when cost of vaccination is small; conversely, the network structure inhibits vaccination coverage when cost of vaccination is large [12]. Recently, a great deal of study has also focused on the impacts of various factors on individual vaccination behavior, such as perception [21], stubborn [22], social influence [23], different subsidy strategies [16,24,25], strategy conformity [15,26], anti-social behavior [27], hypergraph structure [28] and so on.

Given that individuals may have no complete information of the entire network and are not completely rational, the Fermi function based rule is often used to characterize the vaccination decision of individuals, i.e., the probability that individual *i* adopts individual *j*'s strategy is determined by their current payoff differences and the rationality level of individuals [12,16]. Nevertheless, the Fermi function based rule only considers the difference of the current payoffs, without fully considering the strong learning capability of human beings. In fact, individuals can continuously interact with the environment and then search for the best policy for themselves. Reinforcement learning (RL) is an aspect of machine learning where an agent tries to maximize the total amount of reward it receives

when interacting with a complex, uncertain environment, and it utilizes a Q-table to record and update the values for each state-action pair [29]. In practical scenarios, the number of state-action pairs is not fixed, so the deep reinforcement learning (DRL) was proposed to solve the problem [30]. As a pioneering work of DRL, the deep Q-network (DQN) algorithm is a representative method and has garnered widespread attention in recent years [31,32].

Motivated by the above considerations, in this work, we consider a mixed updating strategy for vaccination decision of individuals composed of Fermi function strategy and RL strategy, and then study the impact of such a mixed strategy on the vaccination behaviors and epidemic dynamics. Specifically, we divide individuals in networks into two categories: one group of individuals update their vaccination decisions based on Fermi function (referred to as regular agents), while the other group of individuals update their vaccination decisions based on RL strategy (referred to as intelligent agents). Since each individual's local information is flexible and dynamically changing, such as the number of neighbors, vaccinated neighbors or infected neighbors, we utilize DQN algorithm to update the decisions of intelligent agents. Experiments demonstrate that the RL strategy plays a double-edged sword role in vaccination behavior as well as epidemic dynamics. When the vaccination cost is not very high, a higher proportion of intelligent agents promotes vaccination coverage, leading to a dramatic reduction in the epidemic. However, when the vaccination cost is very high, the presence of intelligent agents can hinder the willingness to vaccinate, leading to an outbreak of the epidemic.

The rest of this paper is structured as follows. In Sec. 2, the descriptions of our model are introduced. In Sec. 3, main experimental results are presented and analyzed. Finally, the conclusions are summarized in Sec. 4.

2 Proposed model

We study the vaccination dynamics for the prevention of the flu-like disease, in which individuals in networks must make vaccination decision before the onset of each epidemic season. Due to the periodic outbreaks of flu-like diseases and the limited validity of the vaccines, individuals who receive vaccinations can only gain immunity to the disease during the current season. In this situation, we also model the vaccination dynamics as a two-stage iterative process [12]: the first stage is a public vaccination campaign, in which individuals determine whether to vaccinate or not based on the previous season's conditions. The second stage is the epidemic season stage, where vaccinated individuals cannot be infected, while unvaccinated individuals face a certain probability of being infected. In previous studies, individuals within social networks relied on the Fermi function to decide whether to vaccinate or not. In this work, we assume that some individuals update their vaccination decisions using the DQN method. As illustrated in Figure 1, the overall architecture of our proposed model can be subdivided into four steps: the initialization process, the decision-making process, the epidemic spreading process, and the payoff calculation process, where the last three steps repeatedly iterate until convergence or a given number of iterations. It should be noted that, the step 2 and step 4 correspond to the first stage of the two-stage iterative process,
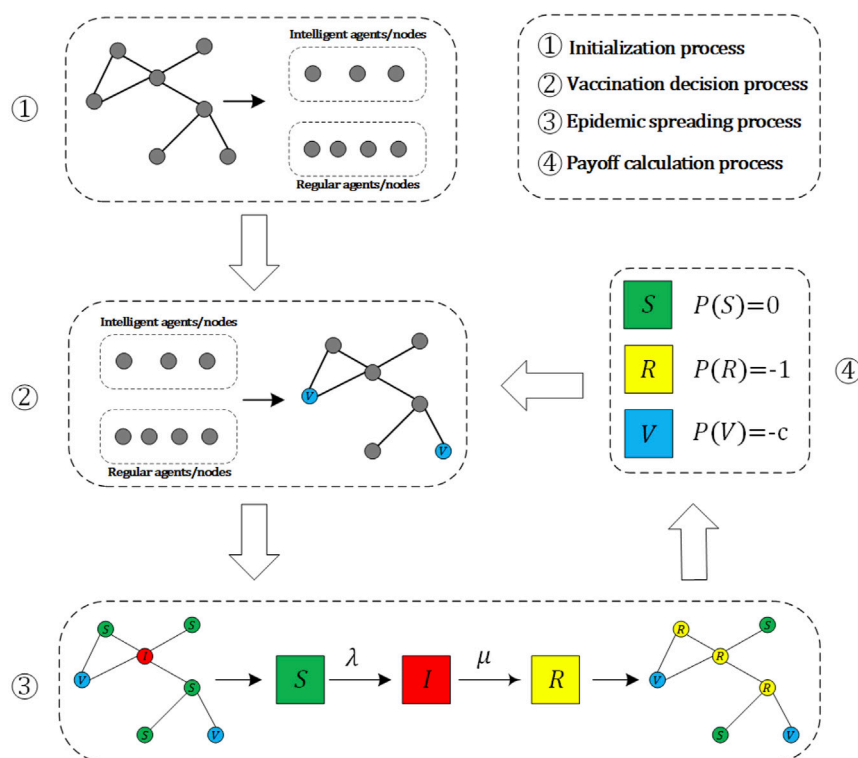
FIGURE 1
Overall architecture of the proposed model.

i.e., public vaccination campaign, where individuals make decisions based on the calculated payoffs. The step 3 is the epidemic season stage of the two-stage iterative process, after this step the payoff of each node can be calculated. Below, we provide detailed explanations for each of them.

2.1 Initialization process

In initialization phase, a proportion ρ of the total number of nodes in the network is randomly selected as the intelligent agents/nodes (i.e., updating vaccination decision based on DQN), and the other nodes are the regular agents/nodes (i.e., updating vaccination decision based on Fermi function). Once the categories of these nodes are established, they remain unchanged throughout the entire process. Meanwhile, we randomly select one-third of nodes to be vaccinated to begin the iterative process. After that, individuals need to use Fermi function or DQN method to decide whether to vaccinate or not based on the prior season's information, such as the vaccination status, epidemic infection situation, payoffs of individuals, and so forth.

2.2 Decision-making process

Since we consider a mixed updating strategy composed of the Fermi function and the DQN method, we will respectively introduce the details of them.

2.2.1 Fermi function based strategy

The regular nodes determine whether to vaccinate or not based on the Fermi function. In detail, for a regular node i, updates his/her vaccination decision by randomly choosing one of its immediate neighbors, say j, compares their costs, and adopts the strategy of j with the following probability [16]:

$$\pi_{i \to j} = \frac{1}{1 + e^{-\beta\left(P_j(t) - P_i(t)\right)}} \tag{1}$$

where $P_i(t)$ defined in Eq. 5 represents the payoff of individual i in the previous season, and β quantifies the uncertainty in the decision-making process [12]. In this work, we fix $\beta = 10$. Node i will adopt the strategy of neighbor j with a probability $\pi_{i \to j}$, and it will retain its own strategy with a probability $1 - \pi_{i \to j}$.

2.2.2 DQN based strategy

The intelligent nodes decide whether to vaccinate or not based on the DQN method. The specific steps are illustrated in Figure 2. We employ an ϵ − greedy strategy, which means randomly selecting an action with a probability of ϵ and choosing the action with the highest Q-value with a probability of $1 - \epsilon$.

The overall process of predicting the Q-values for actions can be divided into three steps. In the first step, all intelligent agents are encoded, and they can be categorized into three types: a) vaccinated and get immunity; b) unvaccinated and not being infected; c) unvaccinated and being infected. We encode these three categories as (1, 0, 1, 0, 0), (0, 1, 0, 0, 1), (0, 1, 0, 1, 0) respectively. The first two digits of the code represent vaccination

FIGURE 2
The whole process of Q-network. The encode module is to encode nodes in the network. The graph embedding module is responsible for representing both local and global information related to nodes through graph embeddings. The Q-values prediction module predicts the Q-values for an node's states and actions, allowing the selection of the action with the highest Q-value for the node's current state.

status, while the following three digits represent immunity, infection, and non-infection, respectively. It is important to emphasize that the encoding method for intelligent agents is not unique. Similar to Ref. [33], we also define a virtual node to represent the global information of the network. And the encoding with the largest number of individuals in three categories (i.e., a, b and c) is defined as the encoding of the virtual node, meanwhile, the neighborhood of the virtual node is the entire network.

In the second step, we utilize a Graph Neural Network (GNN) to generate their embedding representations for the encoded intelligent agents. The specific process is defined as follows [34]:

$$h_{N(v)}^{(l-1)} = \sum_{j \in N(v)} h_j^{(l-1)} \qquad (2)$$

and

$$h_v^{(l)} = ReLU\left(\left[W_1 \cdot h_v^{(l-1)}, W_2 \cdot h_{N(v)}^{(l-1)}\right]\right), \qquad (3)$$

where $h_{N(v)}^{(l-1)}$ represents the aggregated features of the neighbors of node v at the $(l-1)-$th convolutional layer, with $N(v)$ being the neighborhood set of node v. $ReLU$ represents the non-linear activation function. Eq. 2, 3 represent a single layer of graph convolution. During the convolution process, the virtual node aggregates information from its neighbors, its neighbors do not aggregate information from the virtual node.

In the last step, we need to predict the Q-values for the actions that intelligent agents may potentially undertake in a given state. Let $[z_i, z_g]$ be the state of intelligent node i, where z_i and z_g

represent the embeddings of node i and the virtual node, respectively. z_i and z_g also represent the local information of node i and global information, respectively. There are two situations in which node i may take action: taking vaccination or not. We encode it as $A = \{a_1, a_2\} = \{[1, 0], [0, 1]\}$. We then input the state and action into a Multilayer Perceptron (MLP) to predict the current state of node i and the Q-values of potential actions, namely,:

$$y = W_4^T \cdot ReLU\left(W_3^T \cdot \left[z_i, z_g, a_j\right]\right), \qquad (4)$$

where a_j ($j = 1$ or 2) represents the actions that intelligent node i may take, and W_i ($i = 1, 2, 3, 4$) in Eq. 3, 4 are the learnable parameters. The action with the highest Q-value prediction result among all possible actions is chosen. The training process and loss function of DQN are defined in Sub Section 2.5.

2.3 Epidemic spreading process

We use the Susceptibility-Infection-Recovery (SIR) model to simulate the epidemic spread process, with a transmission rate of λ and a recovery rate of μ [35]. In the beginning of each epidemic season, a small proportion of unvaccinated individuals are randomly selected as initial infection seeds I_0. Vaccinated individuals will not be infected in the upcoming season. The epidemic evolves until there are no more newly infected individuals.

FIGURE 3
Training process of DQN.
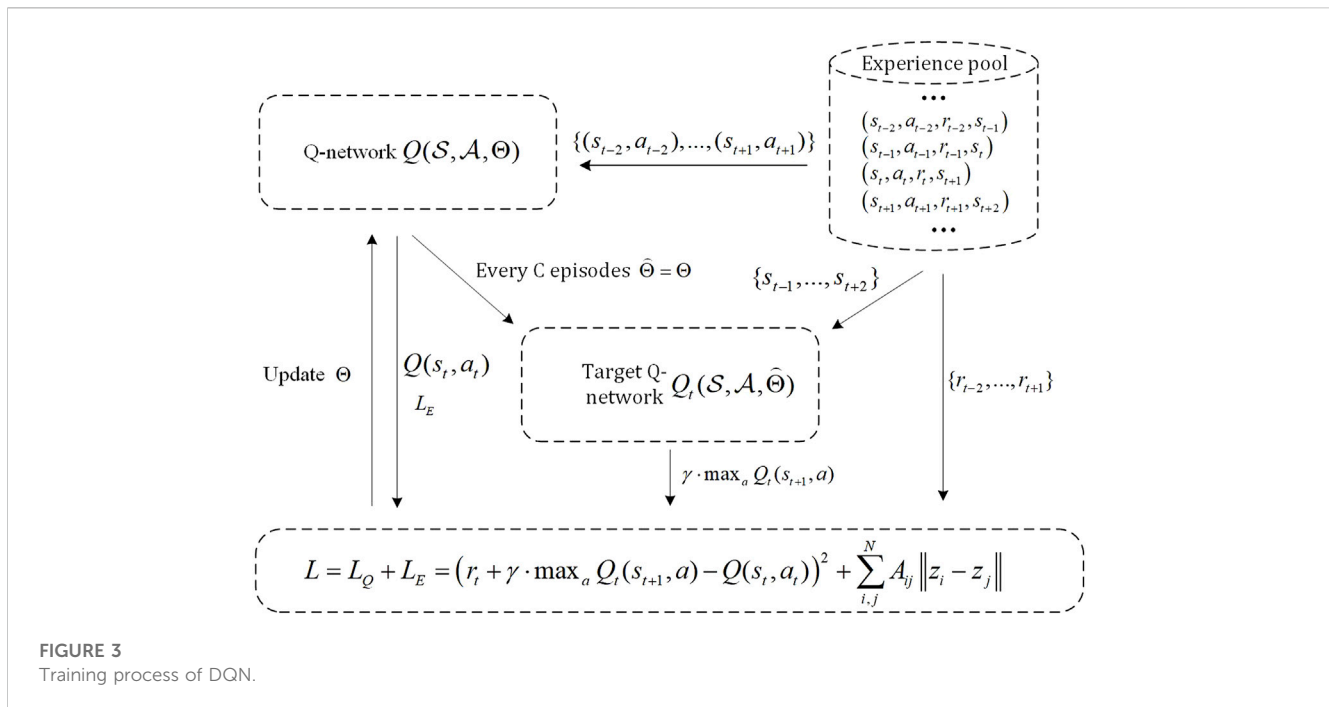
2.4 Payoff calculation process

When the epidemic season ends, it is necessary to calculate the payoffs of individuals in the previous season. Let C_V and C_I be the cost of vaccination and infection, respectively. Without loss of generality, one can set $c = C_V/C_I$ as the relative cost of vaccination with $0 < c < 1$. Namely, the cost of infection is 1. Further let $P_i(t)$ be the payoffs of node i in the t − th season, according to the costs of vaccination and infection, one has

$$P_i(t) = \begin{cases} -c, & \text{vaccination;} \\ -1, & \text{infected;} \\ 0, & \text{free} - \text{rider.} \end{cases} \tag{5}$$

2.5 DQN training process and loss function

Next, we will introduce the training process and loss function of DQN. The overall training process of DQN is depicted in Figure 3. Intelligent agents can obtain the current season's state s_t and the chosen action a_t based on the previous season's infection situation. Meanwhile, each intelligent agent can obtain its payoff according to their vaccination decision and whether to be infected or not, i.e., as defined in Eq. 5, which can be treated as the reward value r_t for (s_t, a_t). Similarly, based on the current season's infection situation, we can obtain the state s_{t+1} for the next season, and this process continues iteratively. We define (s_t, a_t, r_t, s_{t+1}) as an experience and store it in a fixed-size experience pool. We regard the experience of five seasons as an episode of DQN.

We need to define two identical models: one is the Q-network, and the other is the Target Q-network. After a fixed number of C episodes, the parameters of the Q-network are copied to the Target Q-network. To update the parameters of the Q-network, small batches of experiences from the experience pool are randomly

selected. We then need to compute the loss for each experience and then update the parameters of the Q-network through the Back-Propagation algorithm. The loss function consists of two components. The first component is the network embedding loss, which is defined as:

$$L_E = \sum_{i,j}^N A_{ij} \|z_i - z_j\| = 2tr(Z^T L Z), \tag{6}$$

where N and A_{ij} are the size and the adjacency matrix of the network, respectively. z_i and Z represent the embedding vector of node i and the matrix formed by embedding vectors of all nodes, respectively, and L is the Laplacian matrix. $tr(\cdot)$ denotes the trace of a matrix.

The second part is the Q-value prediction loss. For each experience, the objective of the Q-value prediction loss is to minimize the reward error between the predicted and actual values, which is described as:

$$L_Q = \left(r_t + \gamma \cdot \max_{a \in A} Q_t(s_{t+1}, a) - Q(s_t, a_t)\right)^2, \tag{7}$$

where the predicted reward value $Q(s_t, a_t)$ is predicted by the Q-network based on the state and action, and the current reward value r_t plus $\gamma \cdot \max_{a \in A} Q_t(s_{t+1}, a)$ from the Target Q-network is regarded as the actual reward value. γ represents the reward discount factor, which is used to balance the importance of future and current rewards. The overall loss for each experience is defined as a combination of L_E and L_Q with a balancing parameter α, i.e.,

$$L = L_Q + \alpha \cdot L_E. \tag{8}$$

The specific steps of the above process are outlined in Algorithm 1. The first line represents the initialization process for classifying individuals into intelligent or regular agents. Lines 3–14 depict how nodes with different categories decide whether to take vaccination based on various decision rules. Lines 15–16 simulate the epidemic

season, in which SIR model is used to model the spreading of epidemic, and line 17 indicates the payoff calculation process.

Input: The intelligent agents proportion ρ, the initial infection seed I_0 = 5, the season number 2000;

Output: The fraction of vaccination/infection/free-riders;

1: Number $N \cdot \rho$ of nodes are randomly selected as the intelligent agents, and the rest are the regular agents;

2: **for** t = 1 to 2000 **do**

3: **if** node is intelligent agents **then**

4:   **if** $t <$ = 1500 **then**

5:    Decide whether to vaccinate according to Q-network, and the parameter Θ of Q-network is updated;

6:   **else if** $t >$ 1500 **then**

7:    Decide whether to vaccinate according to Target Q-network;

8:   **end if**

9: **else if** node is regular agents **then**

10:   The decision rule is the Fermi function, as shown in Eq. (1);

11: **end if**

12: **if** $t\%50$ == 0 **then**

13:   The parameter of Target Q-network is updated as $\hat{\Theta} = \Theta$;

14: **end if**

15: The unvaccinated individuals are randomly selected as initial infection seeds I_0;

16: epidemic spreads via the SIR model until there are no new infected individuals;

17: The payoffs of individuals are calculated;

18: **end for**

19: Calculate the fraction of vaccination/infection/free-riders based on the Target Q-network.

Algorithm 1 Algorithm for the model.

3 Experiment

In this section, we investigate the impacts of different proportions of intelligent agents on the vaccination behaviors and the epidemic dynamics.

3.1 Experimental setup

Our experiments are employed on three types of networks: the Barabási-Albert (BA) network with $m = 3$ (number of edges with which a new node attaches to existing nodes) and $N = 2000$ [36], the Erdős-Rényi (ER) random network with average degree $\langle k \rangle = 6$ and $N = 2000$ [37], and a real-world Email network [38]. The GNN has 2 embedding layers with a dimensionality of 64 for the embeddings. As in Ref. [33,39], the reward discount factor γ, the size of the experience pool, and the ϵ for ϵ − greedy strategy are set as 0.99, 10000, and 0.05, respectively. We conduct a total of 2000 seasons,

with the initial 1500 seasons designated for model training, followed by the subsequent 500 seasons for testing. Meanwhile, the Q-network's parameters are copied to the Target Q-network for every fixed 50 seasons. In all experiments, without specification, the balancing parameter is $\alpha = 0.01$, the recovery rate is $\mu = 0.25$ and the initial infection seed $I_0 = 5$.

3.2 Experimental results

Figure 4 presents the heatmap results regarding ρ and c on the BA network, demonstrating their impacts on the fraction of vaccination (Figures 4A, C) and the fraction of infection (Figures 4B, D). Several observations can be concluded from Figure 4: Firstly, when the cost of vaccination c is not so high, such as $c < 0.6$ for $\lambda = 0.10$ (Figures 4A, B) and $c < 0.8$ for $\lambda = 0.18$ (Figures 4C, D), the fraction of vaccination increases with the value of ρ, leading to the reduction of the infection. In particular, when ρ is close to 1, almost all nodes take vaccination, giving rise to the complete extinction of epidemic; Secondly, when the cost of vaccination c is very high, such as $c = 0.9$, the opposite phenomenon happens, larger value of ρ induces lower level of vaccination, yielding higher level of infection. The result is also universal for different values of λ. Based on the two observations, one can conclude that the RL based strategy plays a double-edge sword role in the vaccination behavior and the epidemic dynamics. Thirdly, by comparing Figure 4A with Figure 4C, it is found that the fraction of vaccination for the case of $\lambda = 0.18$ is generally higher than that of $\lambda = 0.10$ when the values of ρ and c are not so large. That is to say, higher risk of infection encourages more individuals to take vaccination.

To validate the universality of our observations, we conduct experiments on the ER network (Figures 5A–C) and the Email network (Figure 5D–F) as well. To reflect the double-edge sword role of RL based strategy more clear, the fraction of vaccination, infection and free-riders as the function of ρ are shown in Figure 5. Similar to the results on the BA network, the double-edge sword role of RL based strategy can be observed in Figure 5. In other words, the existence of the intelligent agents has a beneficial impact on suppressing the spreading of epidemic when the cost of vaccination is not very high, whereas, it has a detrimental effect otherwise. As we know, taking vaccination is a better choice when the cost of vaccination is low, however, taking vaccination is almost unnecessary when the cost of vaccination is also equal to the cost of infection. Under different situations, intelligent agents prefer to select so-called "better choice" for themselves, therefore, the double-edge sword role of RL strategy is explainable.

To further elucidate the findings, we define $V(k)$ as the vaccination ratio among nodes with degree k, and let $P_V(k)$, $P_N(k)$ and $P_A(k)$ be the average payoffs of vaccinated nodes, unvaccinated nodes, and of all nodes with degree k, respectively. First, the experimental results for $\lambda = 0.18$ and $c = 0.1$ (low vaccination cost) on the BA network are presented in Figure 6. When $\rho = 0$ (i.e., without intelligent agents), Figure 6A indicates that the vaccination ratio $V(k)$ is proportional to the degree k. The reason is that the nodes with a higher degree are more susceptible to infection since they have a greater number of neighbors, thus nodes with higher degrees exhibit a greater inclination towards choosing vaccination. When $\rho = 0.5$, a notable increase in the vaccination ratio
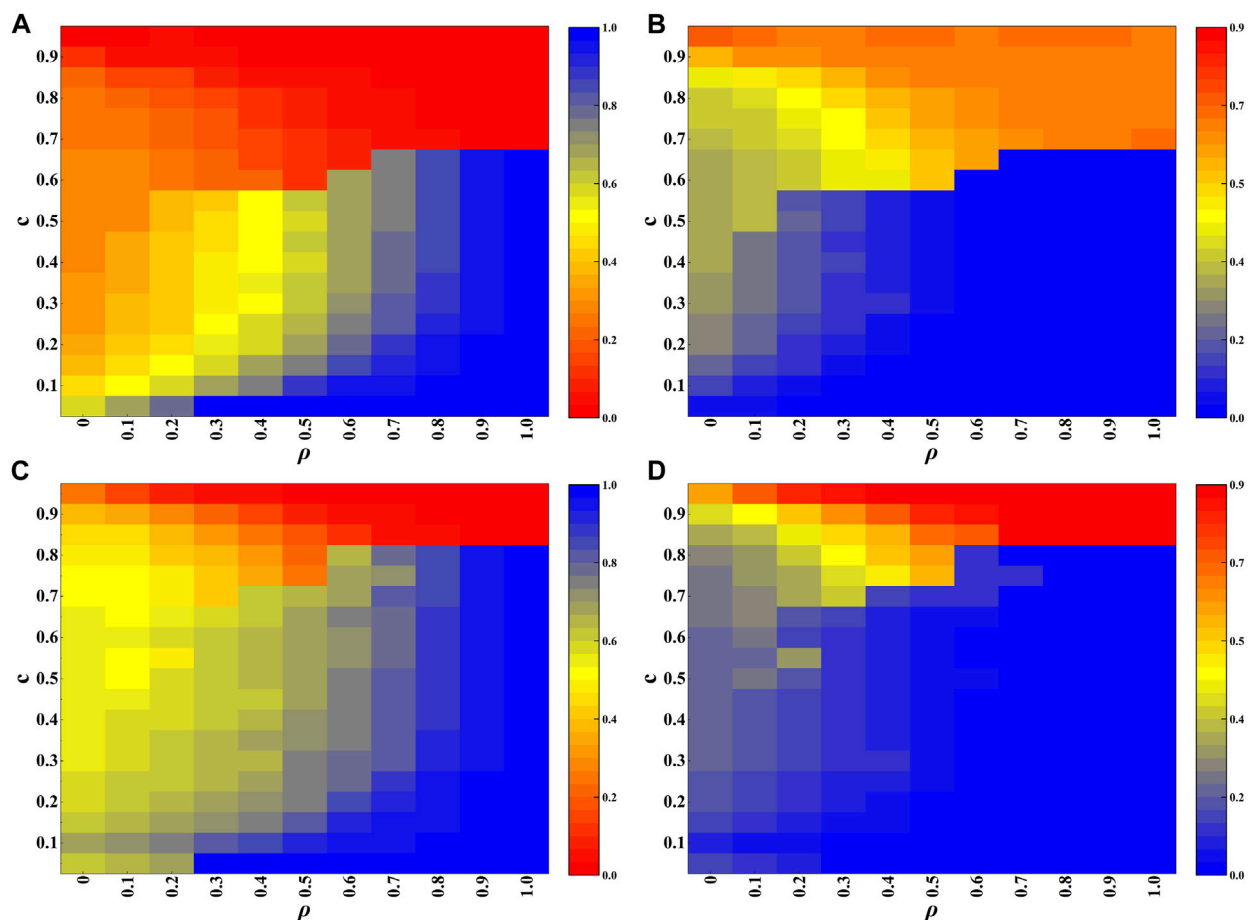
FIGURE 4
Experimental results on the BA network. **(A, B)** are heatmaps depicting the fractions of vaccination and infection, respectively, with transmission rate $\lambda = 0.10$. **(C, D)** are heatmaps depicting the fractions of vaccination and infection, respectively, with transmission rate $\lambda = 0.18$.

is observed compared to Figure 6A, especially for nodes with lower degrees (Figure 6B). Because the majority of nodes in the BA network are of low degree, it leads to a significant improvement in the vaccination coverage. When $\rho = 1.0$, it can be observed that all nodes have a high probability of vaccination, all exceeding 0.95 (Figure 6C). This results in a high overall vaccination coverage in the entire network. Figures 6D–F display the average payoffs of nodes with different degree k in various states. Overall, the average payoffs of unvaccinated nodes decreases with degree k, it is because nodes with lower degrees have a lower probability of being infected. Furthermore, the payoffs of vaccinated nodes significantly surpass unvaccinated nodes. As a result, individuals are more willing to get vaccination. This also explains why, in situations with a relatively low vaccination cost, the introduction of intelligent agents can encourage more nodes to take vaccination. Figures 6D–F also demonstrate that the average payoff of all nodes $P_A(k)$ increases with ρ, which indicates that the presence of intelligent agents also contributes to an overall increase in group benefits when c is relatively small.

Figure 7 displays the experimental results for $\lambda = 0.18$ and $c = 0.9$ (high cost of vaccination) in the BA network. When $\rho = 0$

(Figure 7A), the vaccination ratio $V(k)$ still increases with the degree k, however, owing to the higher cost of vaccination, its growth trend is slower than that shown in Figure 6A. In addition, the vaccination ratio $V(k)$ decreases dramatically as ρ increases to 0.5 (Figure 7B) and further to 1.0 (Figure 7B), especially for the case of $\rho = 1$, the values of $V(k)$ are almost equal to zero for different degree k. The observations imply that the presence of intelligent agents further lower the vaccination proportion when the vaccination cost is extremely high. The average payoffs of nodes with different degrees in different states are further illustrated in Figures 7D–F, they also imply that the average payoffs of unvaccinated nodes decrease with degree k. However, differing from the scenario with low vaccination cost, the average payoff of unvaccinated nodes with lower degrees, such as degree value is 3 or 4, is higher than that of vaccinated nodes. It is because the vaccination cost is extremely high (i.e., average payoff is very low), while the average payoff of unvaccinated nodes with lower degrees is not very small owing to the lower infection risk of them. One can also observe that the average payoff of all nodes $P_A(k)$ decreases with the value of ρ. This indicates that in scenarios with high vaccination cost, the
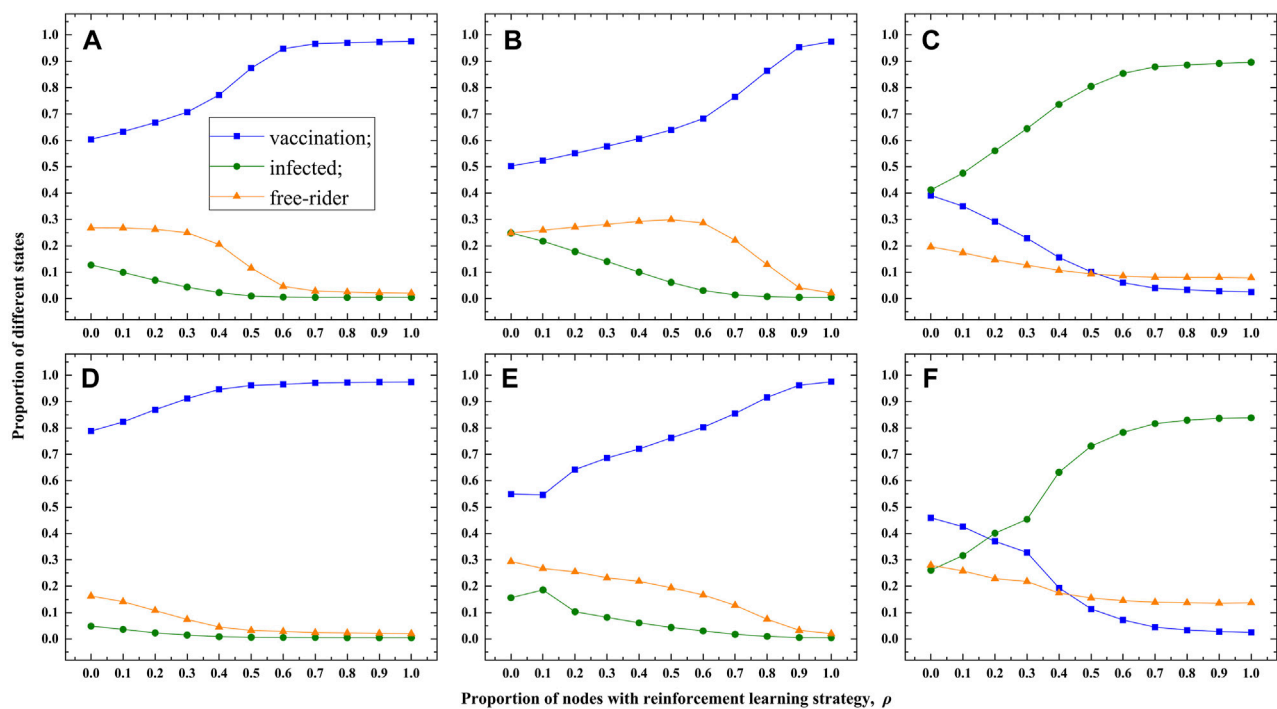
FIGURE 5
Experimental results on the ER and Email networks. **(A−C)** show the results for the ER network with vaccination cost set to *c* =0.1, 0.5, and 0.9, respectively. **(D−F)** display the results for the Email network with vaccination cost set to *c* = 0.1, 0.5, and 0.9, respectively. The transmission rate is *λ* = 0.18.
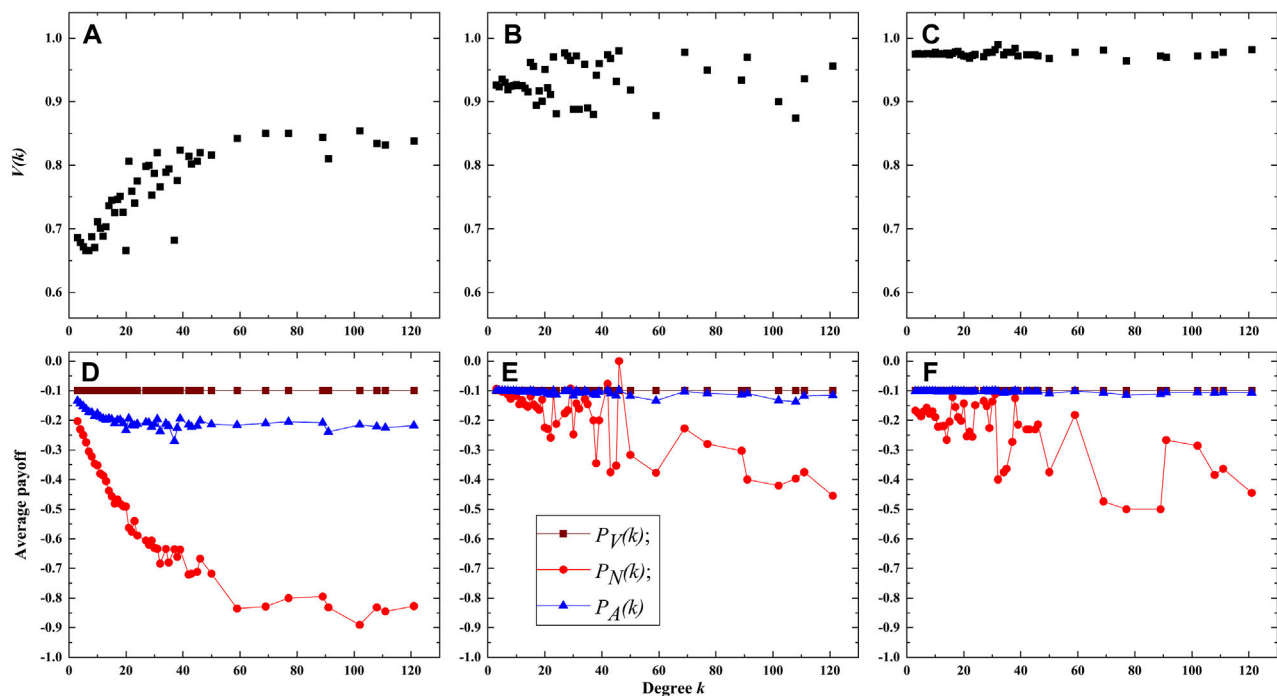
FIGURE 6
Vaccination ratio and payoffs of nodes with different degrees. **(A−C)** are the vaccination proportion *V(k)* for *ρ* = 0.0,0.5 and 1.0 respectively. **(D−F)** are the average payoffs of nodes in different states for *ρ* = 0.0, 0.5 and 1.0 respectively. The cost of vaccination is *c* = 0.1.
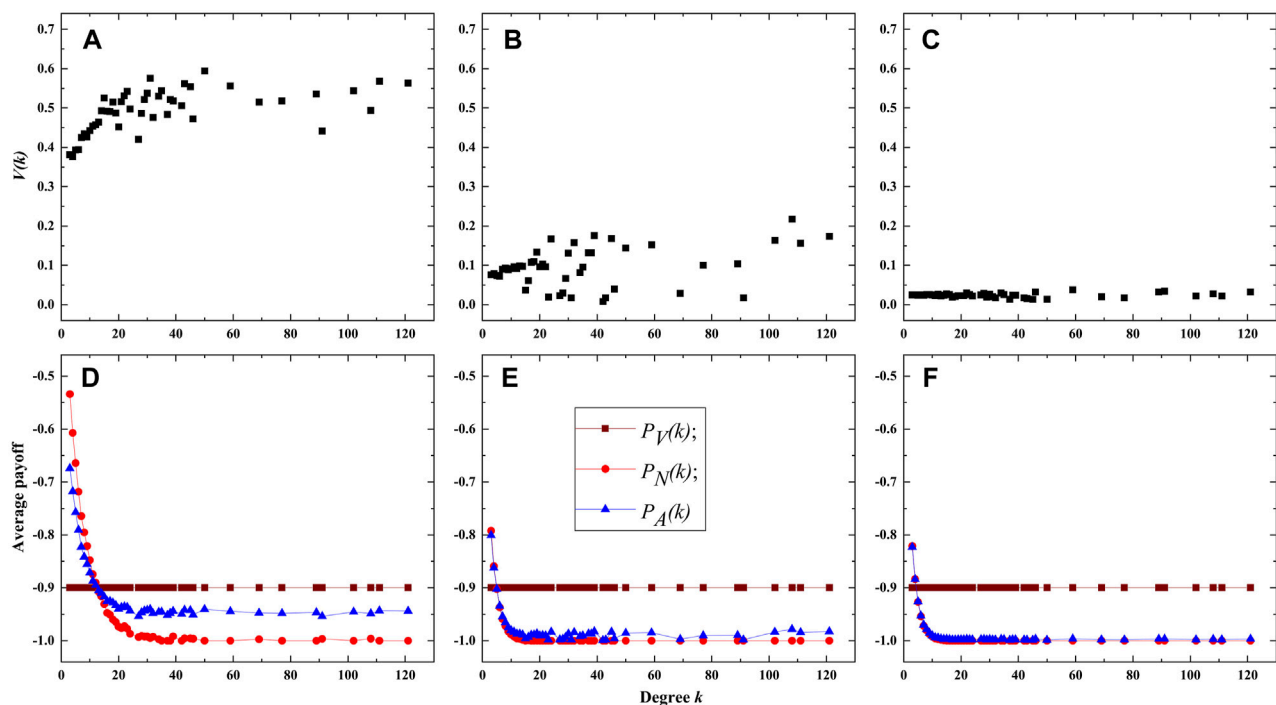
FIGURE 7
Vaccination ratio and payoffs of nodes with different degrees. **(A–C)** are the vaccination proportion $V(k)$ for ρ = 0.0,0.5 and 1.0 respectively. **(D–F)** are the average payoffs of nodes in different states for ρ = 0.0, 0.5 and 1.0 respectively. The cost of vaccination is c = 0.9.
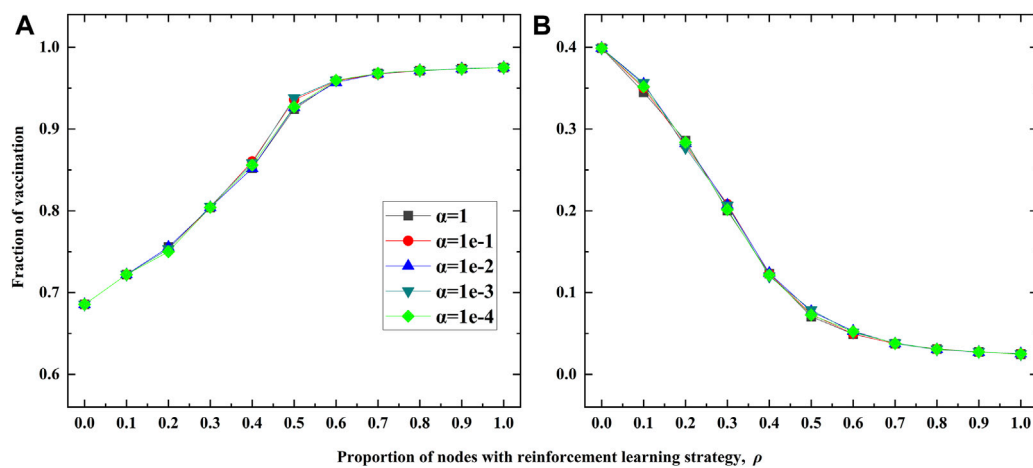
FIGURE 8
Sensitivity analysis of the balancing parameter α on the BA network. **(A)** is for c =0.1. **(B)** is for c = 0.9. Here the transmission rate is λ = 0.18.

presence of intelligent agents not only reduces the vaccination coverage but also leads to a decrease in overall group benefits.

Finally, we conduct the sensitivity analysis regarding the balancing parameter α, and the experimental results are shown in Figure 8. By varying the values of α from 1 to 1e-4, and one can observe that different values of α have minor impact on the fraction of vaccination, no matter c = 0.1 (Figure 8A) or c = 0.9 (Figure 8B). This indicates that the double-edged sword role of RL based strategy is robust to the value of α.

4 Conclusion

In this work, considering the strong learning capability of human beings, we introduced a mixed updating strategy for the vaccination decision of individuals. Specifically, we categorized individuals in the social networks into two groups: regular agents make vaccination decisions based on the Fermi function, primarily considering the difference in current payoffs, while intelligent agents' vaccination decisions are determined by the RL strategy, which relies on local

and global information. Since individuals' local information in the network is flexible and dynamic, we have further integrated the DQN algorithm into the RL strategy for intelligent agents. By varying the proportion of intelligent agents in networks, we found that under appropriate vaccination cost, increasing the proportion of intelligent agents can lead to a significant improvement of vaccination and an effective suppression of epidemic, also inducing an increase of the group benefits. Nevertheless, when the vaccination cost is extremely high, we observed an inverse relationship between the proportion of intelligent agents and vaccination coverage, which consequently leads to a decrease in the group benefits. That is to say, intelligent agents have a double-edged sword effect on vaccination behaviors and group benefits in pursuit of maximizing their own utilization. The findings enrich our understanding on the interplay of the human behavioral responses and epidemic spreading, and may also provide some insights for policymakers regarding the protection and control of epidemics.

There are a number of ways our methods can be extended in future work. For instance, we can consider more decision options for individuals, the incomplete effectiveness of vaccines, the subsidy of vaccines, the distinct structures of epidemic transmission and the vaccination decision updating process, and so on. In addition, we mainly focus on the repeated season model, namely, the vaccination decision should be repeatedly made before each epidemic season. In many situations, the decisions of individuals are often made before or during one emerging diseases. In this case, we should adjust our model to characterize the interplay of the vaccination behavior and the epidemic dynamics for the single season model [40].

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

J-QK: Validation, Writing–original draft, Writing–review and editing. FZ: Validation, Writing–original draft, Writing–review and editing. H-FZ: Conceptualization, Supervision, Writing–original draft, Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Stöhr K, Esveld M. Will vaccines be available for the next influenza pandemic? *Science* (2004) 306:2195–6. doi:10.1126/science.1108165

2. Wang Z, Bauch CT, Bhattacharyya S, d'Onofrio A, Manfredi P, Perc M, et al. Statistical physics of vaccination. *Phys Rep* (2016) 664:1–113. doi:10.1016/j.physrep. 2016.10.006

3. Yin Q, Wang Z, Xia C, Bauch CT. Impact of co-evolution of negative vaccine-related information, vaccination behavior and epidemic spreading in multilayer networks. *Commun Nonlinear Sci Numer Simulation* (2022) 109:106312. doi:10. 1016/j.cnsns.2022.106312

4. Wang W, Liu Q-H, Liang J, Hu Y, Zhou T. Coevolution spreading in complex networks. *Phys Rep* (2019) 820:1–51. doi:10.1016/j.physrep.2019.07.001

5. Kabir KA, Kuga K, Tanimoto J. The impact of information spreading on epidemic vaccination game dynamics in a heterogeneous complex network-a theoretical approach. *Chaos, Solitons & Fractals* (2020) 132:109548. doi:10.1016/j.chaos.2019.109548

6. Wu B, Fu F, Wang L. Imperfect vaccine aggravates the long-standing dilemma of voluntary vaccination. *PLoS One* (2011) 6:e20577. doi:10.1371/journal.pone.0020577

7. Chen X, Fu F. Imperfect vaccine and hysteresis. *Proc R Soc B* (2019) 286:20182406. doi:10.1098/rspb.2018.2406

8. Chang SL, Piraveenan M, Pattison P, Prokopenko M. Game theoretic modelling of infectious disease dynamics and intervention methods: a review. *J Biol Dyn* (2020) 14: 57–89. doi:10.1080/17513758.2020.1720322

9. Bauch CT, Galvani AP, Earn DJ. Group interest versus self-interest in smallpox vaccination policy. *Proc Natl Acad Sci* (2003) 100:10564–7. doi:10.1073/pnas. 1731324100

10. Bauch CT, Earn DJ. Vaccination and the theory of games. *Proc Natl Acad Sci* (2004) 101:13391–4. doi:10.1073/pnas.0403823101

11. Ndeffo Mbah ML, Liu J, Bauch CT, Tekel YI, Medlock J, Meyers LA, et al. The impact of imitation on vaccination behavior in social contact networks. *PLoS Comput Biol* (2012) 8:e1002469. doi:10.1371/journal.pcbi.1002469

12. Fu F, Rosenbloom DI, Wang L, Nowak MA. Imitation dynamics of vaccination behaviour on social networks. *Proc R Soc B: Biol Sci* (2011) 278:42–9. doi:10.1098/rspb. 2010.1107

13. Bauch CT. Imitation dynamics predict vaccinating behaviour. *Proc R Soc B: Biol Sci* (2005) 272:1669–75. doi:10.1098/rspb.2005.3153

14. Vardavas R, Breban R, Blower S. Can influenza epidemics be prevented by voluntary vaccination? *PLoS Comput Biol* (2007) 3:e85. doi:10.1371/journal.pcbi. 0030085

15. Wang X, Jia D, Gao S, Xia C, Li X, Wang Z. Vaccination behavior by coupling the epidemic spreading with the human decision under the game theory. *Appl Math Comput* (2020) 380:125232. doi:10.1016/j.amc.2020.125232

16. Zhang H-F, Shu P-P, Wang Z, Tang M, Small M. Preferential imitation can invalidate targeted subsidy policies on seasonal-influenza diseases. *Appl Math Comput* (2017) 294:332–42. doi:10.1016/j.amc.2016.08.057

17. Han D, Sun M. An evolutionary vaccination game in the modified activity driven network by considering the closeness. *Physica A: Stat Mech its Appl* (2016) 443:49–57. doi:10.1016/j.physa.2015.09.073

18. Han D, Wang X. Vaccination strategies and virulent mutation spread: a game theory study. *Chaos, Solitons & Fractals* (2023) 176:114106. doi:10.1016/j.chaos.2023. 114106

19. Perisic A, Bauch CT. Social contact networks and disease eradicability under voluntary vaccination. *PLoS Comput Biol* (2009) 5:e1000280. doi:10.1371/journal.pcbi. 1000280

20. Mbah MLN, Liu J, Bauch CT, Tekel YI, Medlock J, Meyers LA, et al. The impact of imitation on vaccination behavior in social contact networks. *PLoS Comput Biol* (2012) 8:e1002469. doi:10.1371/journal.pcbi.1002469

21. Feng X, Wu B, Wang L. Voluntary vaccination dilemma with evolving psychological perceptions. *J Theor Biol* (2018) 439:65–75. doi:10.1016/j.jtbi.2017.11.011

22. Fukuda E, Tanimoto J. Effects of stubborn decision-makers on vaccination and disease propagation in social networks. *Int J Automation Logistics* (2016) 2:78–92. doi:10.1504/ijal.2016.074909

23. Xia S, Liu J. A computational approach to characterizing the impact of social influence on individuals' vaccination decision making. *PLoS One* (2013) 8:e60373. doi:10.1371/journal.pone.0060373

24. Zhang H-F, Wu Z-X, Xu X-K, Small M, Wang L, Wang B-H. Impacts of subsidy policies on vaccination decisions in contact networks. *Phys Rev E* (2013) 88:012813. doi:10.1103/physreve.88.012813

25. Wang J, Zhang H, Jin X, Ma L, Chen Y, Wang C, et al. Subsidy policy with punishment mechanism can promote voluntary vaccination behaviors in structured populations. *Chaos, Solitons & Fractals* (2023) 174:113863. doi:10.1016/j.chaos.2023.113863

26. An T, Wang J, Zhou B, Jin X, Zhao J, Cui G. Impact of strategy conformity on vaccination behaviors. *Front Phys* (2022) 10:972457. doi:10.3389/fphy.2022.972457

27. Utsumi S, Arefin MR, Tatsukawa Y, Tanimoto J. How and to what extent does the anti-social behavior of violating self-quarantine measures increase the spread of disease? *Chaos, Solitons & Fractals* (2022) 159:112178. doi:10.1016/j.chaos.2022.112178

28. Nie Y, Su S, Lin T, Liu Y, Wang W. Voluntary vaccination on hypergraph. *Commun Nonlinear Sci Numer Simulation* (2023) 127:107594. doi:10.1016/j.cnsns.2023.107594

29. Kaelbling LP, Littman ML, Moore AW. Reinforcement learning: a survey. *J Artif Intelligence Res* (1996) 4:237–85. doi:10.1613/jair.301

30. Arulkumaran K, Deisenroth MP, Brundage M, Bharath AA. Deep reinforcement learning: a brief survey. *IEEE Signal Process. Mag* (2017) 34:26–38. doi:10.1109/msp.2017.2743240

31. Mnih V, Kavukcuoglu K, Silver D, Rusu AA, Veness J, Bellemare MG, et al. Human-level control through deep reinforcement learning. *Nature* (2015) 518:529–33. doi:10.1038/nature14236

32. Hafiz A (2022). A survey of deep q-networks used for reinforcement learning: state of the art. *Intell Commun Tech Virtual Mobile Networks: Proc ICICV* 2022, 393–402.

33. Fan C, Zeng L, Sun Y, Liu Y-Y. Finding key players in complex networks through deep reinforcement learning. *Nat Machine Intelligence* (2020) 2:317–24. doi:10.1038/s42256-020-0177-2

34. Wu Z, Pan S, Chen F, Long G, Zhang C, Philip SY. A comprehensive survey on graph neural networks. *IEEE Trans Neural Networks Learn Syst* (2020) 32:4–24. doi:10.1109/tnnls.2020.2978386

35. Pastor-Satorras R, Castellano C, Van Mieghem P, Vespignani A. Epidemic processes in complex networks. *Rev Mod Phys* (2015) 87:925–79. doi:10.1103/revmodphys.87.925

36. Barabási A-L, Albert R. Emergence of scaling in random networks. *Science* (1999) 286:509–12. doi:10.1126/science.286.5439.509

37. Erdös P, Rényi A. On the evolution of random graphs. *Publ Math Inst Hung Acad Sci* (1960) 5:43.

38. Rossi RA, Ahmed NK (2015). *The network data repository with interactive graph analytics and visualization*, West Lafayette: Purdue University

39. Yan D, Xie W, Zhang Y, He Q, Yang Y. Hypernetwork dismantling via deep reinforcement learning. *IEEE Trans Netw Sci Eng* (2022) 9:3302–15. doi:10.1109/tnse.2022.3174163

40. Tanimoto J. *Sociophysics approach to epidemics*, 23. Berlin, Germany: Springer (2021).

Transforming medical equipment management in digital public health: a decision-making model for medical equipment replacement

Luying Huang[1†], Wenqian Lv[2†], Qingming Huang[1,3], Haikang Zhang[1], Siyuan Jin[2], Tong Chen[2] and Bing Shen[1,4]*

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai, China, [2]Shanghai General Hospital, Shanghai, China, [3]School of Medical Imaging, Shanghai University of Medicine & Health Sciences, Shanghai, China, [4]Shanghai Tenth People's Hospital of Tongji University, Shanghai, China

Introduction: In the rapidly evolving field of digital public health, effective management of medical equipment is critical to maintaining high standards of healthcare service levels and operational efficiency. However, current decisions to replace large medical equipment are often based on subjective judgments rather than objective analyses and lack a standardized approach. This study proposes a multi-criteria decision-making model that aims to simplify and enhance the medical equipment replacement process.

Methods: The researchers developed a multi-criteria decision-making model specifically for the replacement of medical equipment. The model establishes a system of indicators for prioritizing and evaluating the replacement of large medical equipment, utilizing game theory to assign appropriate weights, which uniquely combines the weights of the COWA and PCA method. In addition, which uses the GRA method in combination with the TOPSIS method for a more comprehensive decision-making model.

Results: The study validates the model by using the MRI equipment of a tertiary hospital as an example. The results of the study show that the model is effective in prioritizing the most optimal updates to the equipment. Significantly, the model shown a higher level of differentiation compared to the GRA and TOPSIS methods alone.

Discussion: The present study shows that the multi-criteria decision-making model presented provides a powerful and accurate tool for optimizing decisions related to the replacement of large medical equipment. By solving the key challenges in this area as well as giving a solid basis for decision making, the model makes significant progress toward the field of management of medical equipment.

KEYWORDS

medical equipment, decision-making, MCDM, game theory, hospital management

1 Introduction

Large medical equipment is an essential material foundation for maintaining the normal operation of hospitals and improving their competitiveness (1, 2). With the iterative development of medical technology, hospitals should match the acquisition of medical equipment to the actual needs. In one survey, it was shown that nearly 60% of the total cost of a hospital project

involves hospital equipment (3). The Malaysian government invested about MYR27 million in healthcare facilities in 2018 by implementing a program of new and upgraded medical equipment purchases (4). According to the Chinese government, the total value of medical equipment in all hospitals rose from RMB320 billion to RMB629 billion from 2010 to 2015, thus medical equipment occupies an important investment in public hospitals (5).

In practice, however, some hospitals blindly pursue the advancement of equipment, leading to unbalanced resource allocation and waste of resources. The phenomenon of under-utilization and over-utilization of equipment occurs repeatedly, adding an invisible burden to patients, reducing the operational efficiency of hospitals, and neglecting the actual needs of hospital work use (6, 7). Hospital management decision makers are faced with the challenge of replacing medical equipment in an orderly manner, especially when it comes to old equipment, and need to prioritise the replacement of various medical equipment through assessment and quantitative tools for effective allocation of state funds and the healthy development of clinical departments in hospitals (8). Too little or too slow replacement of equipment can easily lead to stagnation in the development of the department, hindering the healthy development of the hospital and affecting the patient's experience.

Multiple Criteria Decision Making (MCDM) is a decision analysis method used to assist decision-makers in evaluating and selecting the best decision alternative among multiple decision criteria or standards (9). Evaluating major medical equipment replacement priorities is closely related to problem-solving using multicriteria decision-making (10). Due to the particular ambiguity and difficulty in defining indicators in solving multicriteria problems, MCDM calculates an overall score based on the weight of each criterion by quantifying the ranked quantitative criteria and provides effective decision-making on a more accurate basis (11). Common methods available include hierarchical analysis (AHP) (12), network analysis (ANP) (13), ideal solution similarity preference ranking (TOPSIS) (14), and data envelopment analysis (DEA). Presently, domestic and foreign scholars have thoroughly researched medical equipment replacement decisions. Mazloum Vajari S et al. (15) a proposed decision system that uses a hybrid SWOT-ANP-WASPAS approach provides solutions for medical equipment replacement programs. Ben Houria et al. (16) developed a multicriteria decision model based on AHP, TOPSIS and MILP methods to select the best maintenance strategy for the equipment by quantitatively ranking the different maintenance strategies of the

equipment according to their importance. Mora-García T et al. (17) using an assessment tool based on multicriteria decision analysis, 12 indicators were defined for technical and economic aspects, resulting in the Medical Equipment Replacement Priority Indicator (MERUPI), which provides supporting criteria for deciding which medical equipment should be replaced and for the purchase plan. Faisal M et al. (18) proposed an analytical hierarchy processes -group decision-making (AHP-GDM) model, which includes 11 quantitative and qualitative indicators as primary and secondary criteria to prioritize medical equipment replacement priorities.

The focus of this paper is to develop a comprehensive MCDM model to test the feasibility and superiority of the improved COWA-PCA and GRA-TOPSIS methods in evaluating the replacement priorities of large medical equipment based on the example of four MRI devices.

2 Constructing the evaluation indicator system

This article follows the principles of systematicity, operability, independence and measurability (19), and combines the demand characteristics of the hospital and the technical characteristics of the equipment to construct an evaluation system for the replacement priority of large medical equipment, so as to assists hospitals and related departments in managing and replacing medical equipment more effectively (20, 21), ensuring that the equipment is operated efficiently, and to reduce the operating costs.

First, relevant data involving the renewal of large medical equipment were studied, and the indicators related to the equipment were initially screened. Subsequently, experts engaged in medical equipment management and hospital management were consulted using the Delphi method (22, 23), and the initially selected evaluation indicators were refined and perfected from the 20 indicator datasets based on the criteria of high sensitivity and collectability of the indicators. Finally, the evaluation index system was constructed from four aspects, namely, operation guarantee, social benefits, technical indicators (24) and economic benefits, in order to comprehensively evaluate the performance, efficiency and effectiveness of the equipment in actual operation, and to provide a more comprehensive and objective basis for the decision-making of equipment renewal, as shown in Table 1.

TABLE 1 Large medical equipment renewal priority evaluation indicator system.

First-level indicator	Second-level indicator	Meaning of indicator	Property
Operational security(X_1)	Work Saturation(X_{11})	actual working load of the equipment	+
	Equipment Usage(X_{12})	actual working condition of the equipment	+
	Work Intensity(X_{13})	work density during the rated working time	+
Social benefits(X_2)	Average Patient Waiting Time (X_{21})	patient satisfaction with the examination	+
	Average Patient Interval Length (X_{22})	efficiency of the equipment examination	+
Technical specifications(X_3)	Life Index(X_{31})	current condition and performance of the equipment	+
	Failure Frequency(X_{32})	reliability of the equipment in actual operation	+
Economic benefits(X_4)	Cost–benefit Ratio(X_{41})	ratio of revenue generated to operating costs	−
	Payback Period(X_{42})	payback time of the equipment investment	−

2.1 Operational security (X_1)

Operational assurance indicators reflect the stability and reliability of the equipment during actual use. These indicators allow us to understand how the equipment operates and whether it can meet the existing workload and demands.

1) Operating saturation (X_{11}): higher operating saturation may lead To overworking of the equipment, thus affecting its stability and reliability.

$$X_{11} = \frac{(\tau + 2) \times \sum Z_i}{\eta_b} \qquad (1)$$

Where τ refers to the average inspection time per patient, Z refers to the total number of inspections and η_b the equipment's powered-on runtime, and 2 refers to the preparation time set aside.

2) Equipment utilization (X_{12}): a low equipment utilization rate may mean that equipment is sitting idle for a more extended period and resources are not being fully utilized.

$$X_{12} = \frac{\eta_w}{\eta_b} \qquad (2)$$

Where η_w refers to the total working time of the equipment, from the start of the first patient's inspection to the end of the last patient's inspection.

3) Work intensity (X_{13}): higher work intensities can lead to excessive wear and tear and equipment breakdowns

$$X_{13} = \frac{(\tau + 2) \times \sum Z_i}{\eta_r} \qquad (3)$$

Where η_r refers to the rated working time, generally 8 h daily.

2.2 Social benefits (X_2)

The social effectiveness indicators focus on patient satisfaction and quality of care. Understanding how well the equipment performs in meeting the needs of patients helps to evaluate the impact of the equipment on the overall reputation of the hospital and patient satisfaction. Societal benefits are important for decisions on equipment replacement.

1) Average patient waiting time (X_{21}): longer waiting times may affect patient satisfaction and, thus, the impact of equipment on patient service quality.

$$X_{21} = \frac{\sum (\tau_{bi} - \tau_{ci})}{\sum Z_i'} \qquad (4)$$

Where τ_{bi} refers to the patient's examination start time, τ_{ci} refers to the patient's examination end time, Z_i' and refers to appointment times.

2) Average patient interval length (X_{22}): longer examination intervals may mean that the equipment is insufficiently used, possibly due to poor appointment scheduling or operational delays resulting in longer waiting times for patients, which helps to understand the efficiency of the equipment.

$$X_{22} = \sum (\tau_{bi} - \tau_{e(i-1)}) \qquad (5)$$

Where $\tau_{e(i-1)}$ refers to the previous patient's end time.

2.3 Technical specifications (X_3)

Technical indicators focus on the technical performance and status of the equipment and can directly reflect the technical advancement and reliability. Technical indicators help to understand whether the equipment is at or near its expected service life and the failure rate of the equipment in actual operation.

1) Life index (X_3): equipment has a reference useful life of 6 years, and a higher life index means that the equipment is close to or Has reached its estimated useful life, which can help in planning for replacement or upgrading of the equipment and affect the efficiency of the equipment

$$X_{31} = \frac{\psi_a}{\psi_p} \qquad (6)$$

Where ψ_a refers to the current age of equipment, ψ_p refers to the Estimated valid lifetimes of equipment, generally taken to 6 years.

2) Failure frequency (X_{32}): a high number of failures may mean that the performance of the equipment decreases and helps to understand the stability and reliability of the equipment.

$$X_{32} = \sum \phi_b \qquad (7)$$

Where ϕ_b refers to several breakdown times.

2.4 Economic benefits (X_4)

The economic efficiency indicator looks at the equipment's cost-effectiveness and investment recovery. It reflects the value of the equipment on an economic level and helps to determine whether the economic performance of the equipment is in line with expectations.

1) Cost–benefit ratio (X_{41}): reflects the revenue generated To operating costs. A lower cost–benefit ratio may mean The equipment Is more expensive, reducing economic efficiency.

$$X_{41} = \frac{R_t}{C_t} \qquad (8)$$

Where R_t refers to total equipment inspection revenue C_t and equipment running costs, including staffing, maintenance, servicing, etc.

2) Payback period (X_{42}): reflects the payback time of the equipment investment. A more extended payback period may mean a lower rate of return on the equipment and the need to delay the replacement of the equipment, helping to understand

the economic value of the equipment and the benefits of the investment.

$$X_{42} = \frac{C_v}{R_t - C_t} \qquad (9)$$

Where C_v refers to the price when the equipment was purchased.

3 Materials and methods

3.1 Research framework

The research framework of this article is shown in Figure 1. In Part 1, the evaluation indicator system for prioritizing the replacement of large medical equipment is constructed by looking up the literature as well as consulting with experts (see Table 2).

In Part 2, COWA, PCA, and game theory methods (25) are combined to determine the weights of the indicators. The COWA method is mainly used to assign weights to different expert opinions to determine the key factors affecting the decision of equipment replacement. The PCA is used to calculate the objective weights of the indicators, and the raw indicators are transformed into principal

components. The weights are determined based on the contribution of each principal component to the variance, thus reducing subjective bias. Game theory is used to promote balance and co-operation between multiple stakeholders to develop an optimal weight allocation strategy.

In Part 3, the GRA-TOPSIS method was used to prioritise equipment replacement. The relative correlation between equipment is determined by gray relational analysis (GRA) and the distance of each piece of equipment relative to the ideal solution is calculated using the TOPSIS method to prioritise the replacement decision. By integrating these methods, a combination of expert opinion, reducing information loss and analysing weighting relationships was achieved and the validity of the method was tested by comparing multiple methods and the Kendall method, ultimately providing excellent guidance for the decision-making process surrounding the replacement of large medical equipment.

3.2 Combined weighted averaging(COWA)

Text for this sub-section. C-OWA (26) is a method for determining experts' weights. Based on the concept of Combined Ordered Weighted Average (OWA) and Compatibility Ranking (CO),
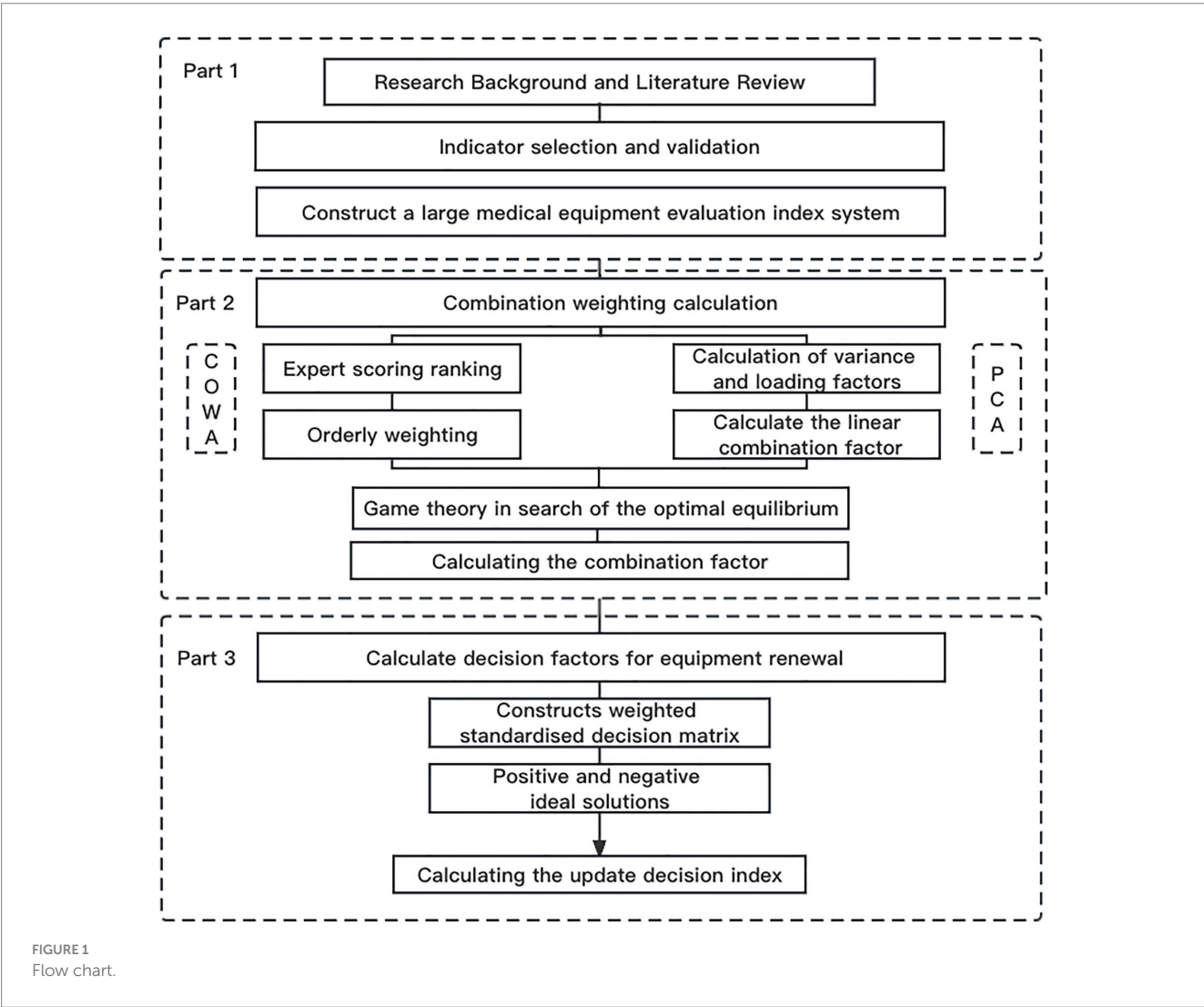

FIGURE 1
Flow chart.

TABLE 2 Results of principal component analysis weights.

Indicator	Components 1	Component 2	Component 3	Component 4	Score	Weights
Characteristic roots	4.013	1.652	1.186	0.924		
Explanation of variance	44.59%	18.36%	13.18%	10.27%		
X_{11}	0.443	0.158	0.297	0.218	0.333	0.124
X_{12}	0.402	0.246	0.325	0.321	0.347	0.129
X_{13}	0.392	0.195	0.123	0.335	0.302	0.113
X_{21}	0.226	0.576	0.062	0.057	0.255	0.095
X_{22}	0.358	0.151	0.229	0.348	0.293	0.109
X_{31}	0.388	0.151	0.367	0.397	0.335	0.125
X_{32}	0.097	0.612	0.336	0.046	0.237	0.088
X_{41}	0.381	0.099	0.432	0.328	0.322	0.120
X_{42}	0.065	0.339	0.549	0.588	0.259	0.096

the COWA operator first calculates the ranking compatibility of each expert on different evaluation indicators and then assigns weights to each expert based on the compatibility score. The method fully uses the experts' experience, eliminates the negative effects of individual extremes, improves the scientific nature of the indicator assignment and avoids the extremes of the experts' perceptions [27].

Step 1: by inviting n experts in the field of medical equipment to rate the importance of the indicators at each level (0 to 10 scale), the initial scoring data of the experts from the data set $\left(a_1, a_2, \cdots\cdots, a_j, \cdots\cdots, a_n\right)$ the scoring data in the data set are sorted from 0 to the smallest, and the result is $b_0 \geq b_1 \geq b_2 \geq \cdots \geq b_{n-1}$.

Step 2: using the combination number b_0 to determine the weights of the data, the weighting vector is applied to the decision data to obtain the absolute weights of the indicators ω_i :

$$\theta_{j+1} = \frac{C_{n-1}^j}{\sum_{k=0}^{n-1} C_{n-1}^k} = \frac{C_{n-1}^j}{2^{n-1}}, j = 0, 1, 2, \cdots, n-1 \quad (10)$$

$$\overline{\omega_i} = \sum_{j=0}^{n-1} ,_{j+1} b_j, i = 1, 2, \cdots, m \quad (11)$$

Where C_{n-1}^j is the combinatorial formula that calculates the number of methods to select j data from $n-1$ data, 2^{n-1} represents all possible combinations when selecting any quantity of data out of $n-1$, $,_{j+1}$ offers a normalized weight for each j.

Step 3: Calculate the relative weights of the indicators ω_i, which achieved by normalizing the absolute weights $\overline{E_i}$ such that their aggregate is equal to 1.

$$\omega_i = \frac{\overline{\omega_i}}{\sum_{i=1}^m w_i}, i = 1, 2, \cdots, m \quad (12)$$

Where $\sum_{i=1}^m w_i$ represents the summation of all absolute weights for all m indicators. This normalization ensures that the relative weights are proportional to the absolute weights, and their collective sum is 1, rendering them appropriate for scenarios where the relative significance between indicators is paramount.

3.3 Principal component analysis (PCA)

Principal Component Analysis (PCA) [28] is a method of multivariate statistical analysis in which multiple factors in an evaluation system are described by a few unrelated important variables, using linear equations to summarize and integrate all the factors so that they are used to reflect the variance at the higher level [29]. All linear combinations are a type of principal component, and the information reflected in the selected composite factors is interpreted to make the overall evaluation model more balanced.

Step 1: The evaluation indicators are normalized to obtain a judgment matrix Y_{ij}

$$Y_{ij} = \frac{x_{ij} - x_{\min}}{x_{\max} - x_{\min}}, (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m) \quad (13)$$

Where $x_{\max}$ and $x_{\min}$ are the highest and lowest values within a given unit.

Step 2: Calculate the correlation coefficient between the variables using the standardized data and calculate the correlation coefficient matrix $R = \left(r_{ij}\right)_{m \times n}$

$$r_{ij} = \frac{\sum_{k=1}^n y_{ki} y_{kj}}{n-1}, (i = 1, 2, \cdots, n; j = 1, 2, \cdots, m) \quad (14)$$

Step 3: Calculate the eigenvalues and eigenvectors to solve for the eigenvalues and eigenvectors of the correlation coefficient matrix. The characteristic equation for the correlation moment $|R - \lambda E| = 0$ calculates the eigenvalues $\lambda_1, \lambda_2 L, \lambda_n$, and the non-zero solution $|R - \lambda E| X = 0$ is the eigenvector $\propto_1, \propto_2 L, \propto_n$.

Step 4: Calculate the variance contribution and cumulative variance contribution of each principal component:

$$a_j = \frac{\lambda_j}{\sum_{j=1}^n \lambda_j}, b_p = \frac{\sum_{k=1}^p \lambda_k}{\sum_{k=1}^m \lambda_k} \quad (15)$$

Where a_j is the variance contribution, b_p is the cumulative variance contribution，p is the number of principal components.

Step 5: Calculate the composite score:

$$Z = \sum_{j=1}^{p} a_j P_j \tag{16}$$

Where P is the main component.

3.4 Game theoretical portfolio weights

In order to improve the objectivity, science and accuracy of the indicator assignment, game theory is introduced into the indicator weights (30). The game principle means that each player in the game decides which action to take, according to their interests and taking into account the possible impact of their decision-making behavior on the behavior of others, by considering the equilibrium between the mutually influencing behaviors, to achieve the goal of optimization of the subject's objectives in a state of compromise between the factors (31). Therefore, game theory is introduced to consider the COWA and the PCA methods as two sides of the game, seeking the optimal combination of weights that will bring both sides to equilibrium. At this balance level, the sum of the deviations between the optimal combination weights and the two is minimized.

Step 1: Using L methods to determine the weights of the n indicators, the set of indicator weights is expressed as $W_k = \left(W_{k_1}, W_{k_2}, \cdots, W_{k_n}\right)$ where $k = 1, 2, \cdots, L$ then the weight L vectors W_k combination weights w are

$$W = \sum_{k=1}^{L} \lambda_k W_k, k = 1, 2, \cdots, L \tag{17}$$

Where λ_k is the linear combination factor.

Step 2: Minimize the divergence $\Delta = \left(W - W_k\right)$ between W and W_k, and according to game theory principles, the corresponding optimization model is

$$\min \sum_{k=1}^{L} \left\| \lambda_k W_k^T - W_k \right\|_2, k = 1, 2, \cdots, L \tag{18}$$

Step 3: Uses two weighting methods, taking the value $L = 2$. According to the principle of differentiation, the set of linear equations for the optimal first-order derivative condition is obtained by substituting Eq:

$$\begin{bmatrix} W_1 W_1^T & W_1 W_2^T \\ W_2 W_1^T & W_2 W_2^T \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \end{bmatrix} = \begin{bmatrix} W_1 W_1^T \\ W_1 W_2^T \end{bmatrix} \tag{19}$$

Step 4: The combination coefficients λ_1 and λ_2 are obtained according to Eq. and normalized to λ_k^*, which in turn gives the game combination weights W^*:

$$W^* = \lambda_1 W_1^T + \lambda_2 W_2^T \tag{20}$$

3.5 Gray relational analysis and technique for order of preference by similarity to ideal solution (GRA-TOPSIS)

Text for this sub-section, the TOPSIS method (32) is suitable for applying large multi-factor systems, which avoids the subjectivity of data and describes the overall evaluation of multiple factors. GRA (33) can be used to judge an indicator's merits by the degree of similarity in geometric shape trends between factors. Each of the above two methods has its advantages (34). The combination of the weighted TOPSIS and GRA to construct a medical equipment replacement priority model makes the conclusions of the model calculation more consistent with practice and scientific.

Step 1: Construct a weighted normalized decision matrix S, calculate the Euclidean distance to the positive and negative ideal solutions for each evaluation object d_i^+, d_i^-:

$$S = \left(s_{ij}\right)_{m \times n} = \left(w_j \times z_{ij}\right)_{m \times n'} \tag{21}$$

$$d_i^+ = \sqrt{\sum_{j=1}^{n} \left(s_{ij} - s_j^+\right)^2},$$

$$d_i^- = \sqrt{\sum_{j=1}^{n} \left(s_{ij} - s_j^-\right)^2}, i = 1, 2, \cdots, m, j = 1, 2, \cdots, n \tag{22}$$

Where $s_j^+ = \max_j s_{ij}; s_j^- = \min_j s_{ij}.$

Step 2: Calculate the matrix of gray correlation coefficients between each solution and the positive and negative ideal solutions:

$$H^+ = \left\{h_{ij}^+\right\}_{m \times n'}, H^- = \left\{h_{ij}^-\right\}_{m \times n'} \tag{23}$$

$$h_{ij}^+ = \frac{\min\min_i \left|s_j^+ - s_{ij}\right| + \rho \max\max_i \left|s_j^+ - s_{ij}\right|}{\left|s_j^+ - s_{ij}\right| + \rho \max_i \max_j \left|s_j^+ - s_{ij}\right|} \tag{24}$$

$$h_{ij}^- = \frac{\min\min_i \left|s_j^- - s_{ij}\right| + \max\max_i \left|s_j^- - s_{ij}\right|}{\left|s_j^- - s_{ij}\right| + \rho \max_i \max_j \left|s_j^- - s_{ij}\right|} \tag{25}$$

Where ρ is the resolution factor and $\rho \in (0,1)$ is taken from $\rho = 0.5$.

Step 3: Calculate the gray correlation coefficients of each evaluation object and the positive and negative ideal solutions l_i^+, l_i^-, and dimensionless process the Euclidean distance d_i^+, d_i^- and the correlation l_i^+, l_i^-:

$$l_i^+ = \frac{1}{n}\sum_{j=1}^{n} W_j^* h_{ij}^+, l_i^- = \frac{1}{n}\sum_{j=1}^{n} W_j^* h_{ij}^- \qquad (26)$$

$$D_i^+ = \frac{d_i^+}{\max\limits_{i}}, D_i^- = \frac{d_i^-}{\max\limits_{i}} \qquad (27)$$

$$L_i^+ = \frac{l_i^+}{\max\limits_{i}}, L_i^- = \frac{l_i^-}{\max\limits_{i}}, i = 1, 2, 3, \cdots m \qquad (28)$$

Step 4: Combine the Euclidean distance D_i^+, D_i^- and the gray correlation coefficient L_i^+, L_i^- and calculate the replacement decision factor ξ_i:

$$T_i^+ = \alpha_1 D_i^- + \alpha_2 L_i^+, T_i^- = \alpha_1 D_i^+ + \alpha_2 L_i^- \qquad (29)$$

$$\xi_i = \frac{T_i^+}{T_i^+ + T_i^-}, i = 1, 2, 3, \cdots m \qquad (30)$$

Where α_1, α_2 reflects the decision maker's preference for location and shape, and $\alpha_1 + \alpha_2 = 1, \alpha_1, \alpha_2 \in [0, 1]$ α_1, α_2 values are empirically taken as 0.5 in general. The more significant the corresponding replacement decision factor ξ_i, the better the object; the smaller the corresponding replacement decision factor ξ_i, the worse the object.

4 Results

4.1 Case presentation and data sources

Text for this sub-section. Four magnetic resonance imaging (MRI) machines in the hospital are used alternately with old and new equipment. The equipment types are Signa HDx, MR750, MR750W and Prisma, located in different parts of the hospital.

In order to fully access the evaluation indicators for equipment replacement, multiple data sources are used to ensure the accuracy and completeness of the required information. Firstly, data collectors were installed on the large equipment to collect critical data such as hours of operation, start-up time and workload in real-time. These data collectors use advanced image recognition technology to accurately monitor the on/off status of the equipment. At the same time, the data collectors read information from the equipment's examination interface through a frequency divider and use image recognition technology to identify essential information such as patient numbers and examination sequences, enabling accurate calculation of the equipment's total working examination time.

In order to obtain more effective data, the real IoT collection data is matched with other information systems in the hospital as well as data integration. By interfacing with PACS, HIS, RIS, reservation system, ERP and other related systems, data acquisition of data fields required by indicators is achieved, to obtain metric fields such as patient appointment

days, average interval length, and revenue. Through in-depth analysis of patient payments and examination moments in these systems, key indicators such as patient appointment days, average interval length, and revenue were obtained. This process uses rigorous data cleaning and validation methods to ensure the reliability and accuracy of the data. In addition, a close working relationship is maintained with the equipment manufacturers to obtain data on failures data. Equipment manufacturers regularly provide information on equipment failure fills, which is carefully verified and collated to provide a reliable data source for assessing equipment failure count indicators.

In summary, the multiple data sources and rigorous data processing methods ensure that the indicator data used in assessing equipment replacement decisions are accurate and comprehensive. This provides a solid study database and helps make more scientific and rational equipment replacement decisions.

4.2 Correlation analysis

Text for this sub-section. Pearsons' method in SPSS is used to calculate correlations between indicators. It can reveal the strength and direction of linear relationships between variables, which indicators have strong positive or negative correlations with each other (35), to understand the interactions between indicators and provide a basis for the subsequent principal component analysis. The result is shown in Figure 2.

The evaluation indicators were subjected to Pearson correlation analysis to reveal their interrelationships. The analysis results show a high positive correlation between equipment utilization, work saturation and work intensity, indicating that these indicators may influence each other. The payback period was positively correlated with the average interval length, indicating that the average interval length also tends to increase when the payback period increases. In contrast, there is a negative correlation with equipment utilization, work saturation, average appointment length, work intensity, equipment life index and cost–benefit ratio, indicating that these indicators tend to decrease when the payback

FIGURE 2
Indicator correlation.
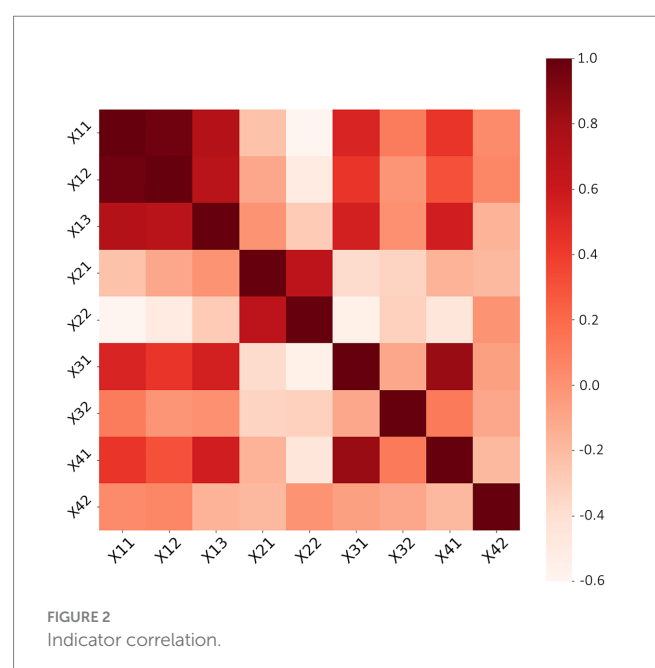
period increases. The average appointment length positively correlated with the number of breakdowns, suggesting that more equipment breakdowns may lead to increased appointment length.

4.3 Determination of weights

Text for this sub-section. To assess the importance of the indicators, eight experts in the field of medical device management were invited to score the indicators and weight them using the C-OWA method. The scores are shown in Figure 3.

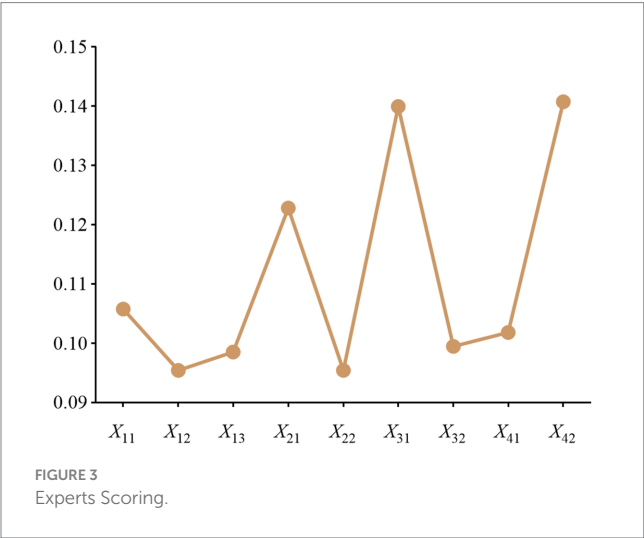Expert scores (0–10) for the evaluation indicators covered in this paper, with larger scores representing higher importance of the indicators.

FIGURE 3
Experts Scoring.

Take indicator X_{11} as an example, sort the scores of indicator X_{11} from the largest to the smallest to get (7–9), and calculate the weight vector by the formula to get: (0.07, 0.44, 1.31, 2.19, 2.19, 1.15, 0.38, 0.05). The results of the weight values are shown in Figure 4.

As shown in Figure 5, the gravimetric plot is a visual tool commonly used to present the results of principal component analysis, determining the number of principal components that should be retained. The graph shows the percentage of variance explained by each principal component and the cumulative percentage of variance explained. An inflection point can be found where the number of principal components retained explains most of the variance in the original data while avoiding the problem of overfitting due to retaining too many principal components. The curve begins to level off at the fourth eigenvalue point. The variance explained by these four principal components are 44.59, 18.36, 13.18 and 10.27% respectively, and the cumulative variance explained is 86.39%.

After calculating the PCA weights and COWA subjective weights, these weights were combined using a game-theoretic approach to obtain a combined weight value for each evaluation indicator. The comparison of the weights calculated by the three methods is shown in Figure 6, which takes into account both the variability of the evaluation indicators and the conflicting and variable nature of the indicator data, reflecting the objectivity of the data itself and helping to provide more targeted guidance to decision-makers, making the weight calculation results more reasonable.

4.4 Updating decision factor determination

Text for this sub-section. The Euclidean distances of the positive and negative ideal solutions s_j^+, s_j^- and the positive and negative ideal solutions d_i^+, d_i^- can be obtained by using Eq. The gray correlation matrix $H^+ = \left\{ h_{ij}^+ \right\}_{m \times n'}, H^- = \left\{ h_{ij}^- \right\}_{m \times n'}$ and the correlation degree

FIGURE 4
COWA weighting results.

FIGURE 5
Principal component feature results.

FIGURE 6
Comparison of weight values for different methods.
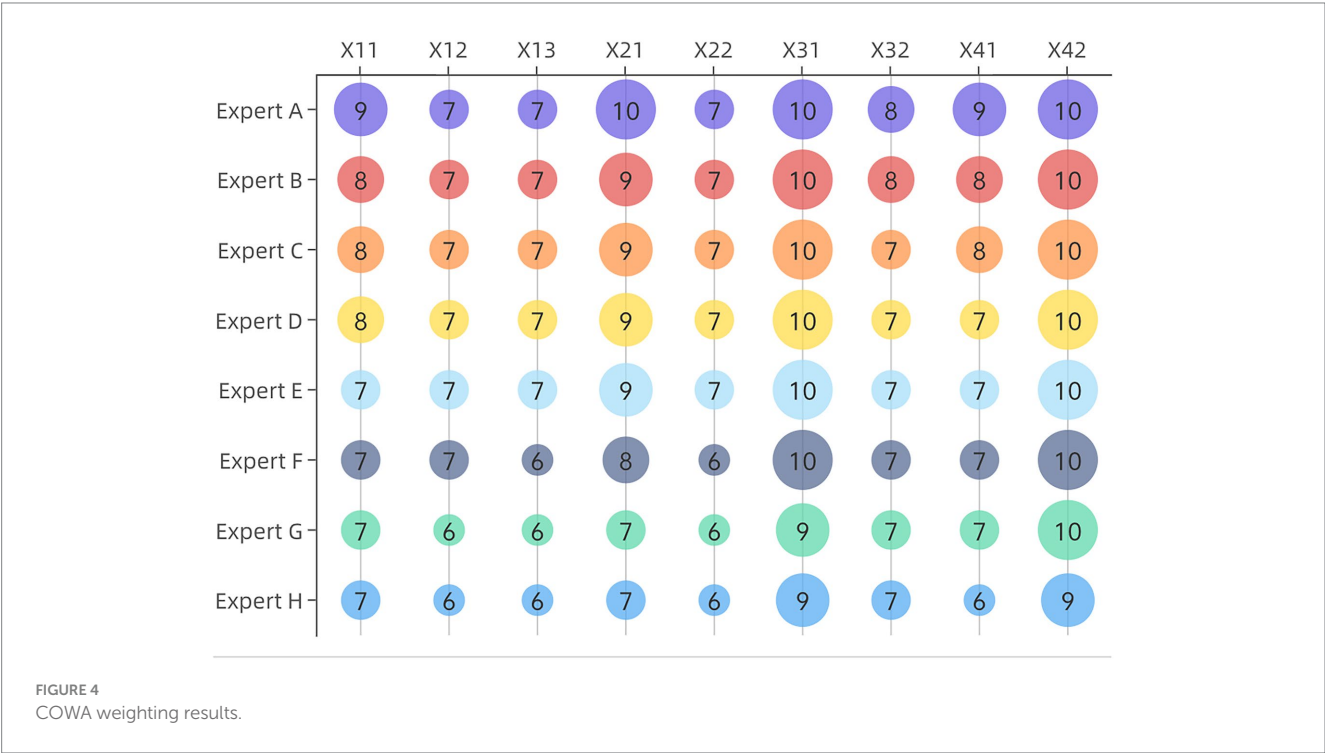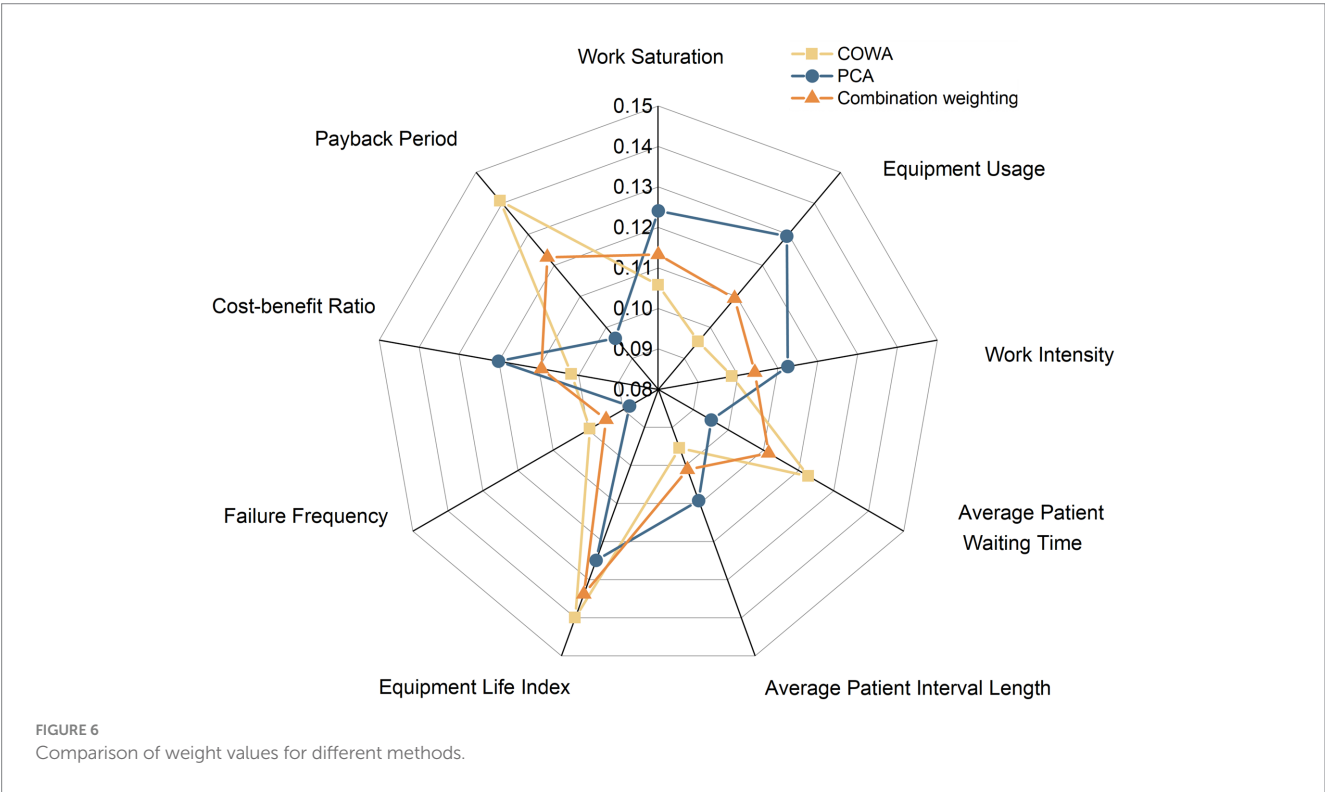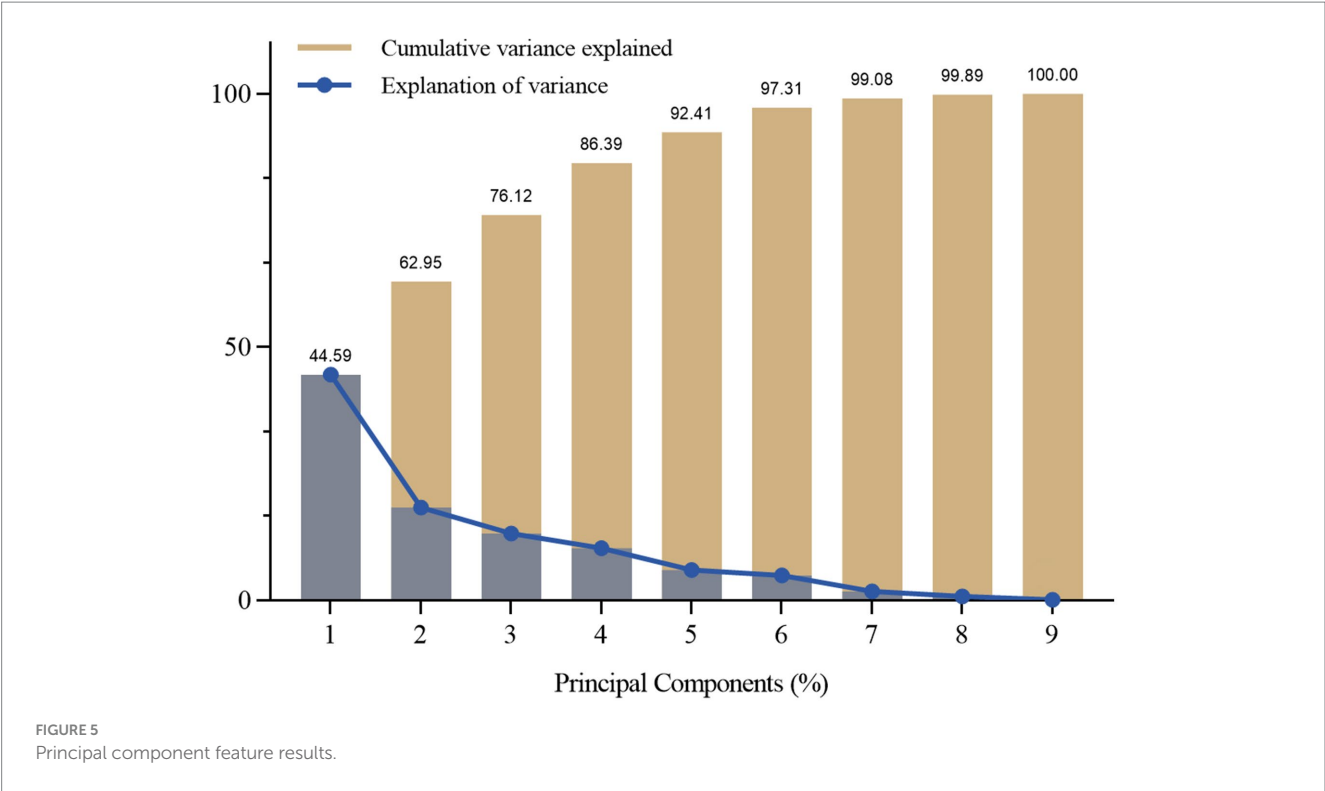
l_i^+, l_i^- are calculated for each device according to Eq. The correlations between the indicators are reflected in the three-dimensional space. Furthermore, the Euclidean distances D_i^+, D_i^- and the correlations L_i^+, L_i^- for each device are processed dimensionless according to (26)–(28). The results are shown in Table 3.

The more significant the replacement decision factor, the more the device needs replacement. The results of ranking the four MRI devices using the combined weight GRA-TOPSIS evaluation model are SignaHDx < MR750 < Prisma <

MR750W. The GRA-TOPSIS model calculates a uniform and reasonable distribution of the resulting values, shown in Figure 7.

The VIKOR method (36) is also applicable to multi-attribute decision problems and is able to take into account the complementarity and conflict between attributes when determining the weights of each attribute. This article compares the results of the calculations of several methods of the GRA method, TOPSIS method, GRA-TOPSIS method and VIKOR method were used to calculate the replacement priority of large medical equipment, respectively, and the data were standardized to take complete account of the closeness of individual indicators to the indicator series. The replacement decision results and ranking results are shown in Table 4.

The results in the table show that compared to the GRA and TOPSIS methods alone, the GRA-TOPSIS method has significantly improved in terms of correlation coefficients, thus better coping with the uncertainties and limitations in the evaluation process. The accuracy of the traditional GRA method may be limited when dealing with decision problems with complex quantitative attributes. In contrast, the accuracy of the TOPSIS method suffers when correlations exist between attributes. The VIKOR method is not applicable when the result values differ too much in the replacement decision evaluation process. This suggests that the GRA-TOPSIS method has higher accuracy and reliability in evaluating replacement priorities for large medical equipment. A comparison of the results obtained from the four methods is shown in Figure 8.

Kendall's test (37) can measure the correlation between multiple evaluation methods and assess their consistency in solving the problem of prioritizing the replacement of large medical equipment, which helps to reveal the similarities and differences between different methods in terms of evaluation results so that the best evaluation method suitable for the actual problem can be selected in a targeted manner. Kendall's test was used to analyze the consistency of the three methods, GRA, TOPSIS and GRA-TOPSIS, to better understand the strengths and weaknesses of the different methods and to provide a reference for subsequent research, which is shown in Table 5.

The table gives Kendall's W coefficient of 0.812, an X^2 value of 6.5 and a value of p of 0.039. Kendall's W coefficient is close to 1, indicating a high level of consistency in the replacement decision factors between the three methods. Also, as the value of p is less than 0.05, the overall data level shows significance. Therefore there is a significant correlation between the three methods regarding the replacement decision factor.

5 Discussion

This research constructs a large medical equipment renewal priority evaluation indicator system, covering four aspects: operational security, social benefits, technical indicators and economic benefits. This system is designed to provide a robust framework for medical equipment administrators and policymakers, thereby facilitating empirically informed renewal decisions. The employed methodology harnesses the combined strengths of the COWA and PCA methods, bridging both subjective judgment and objective data attributes. Through an amalgamation of subjective weights derived from COWA with objective weights from PCA, realized via game-theoretic reasoning, a comprehensive weighting system is established. This approach not only minimizes the potential for information loss, which is often inherent in isolated weighting schemes, but also augments error resilience and alignment, ensuring enhanced methodological precision and relevance.

TABLE 3 Results of principal component analysis weights.

Equipment Type	European distance		Gray correlation	
	Positive	Negative	Positive	Negative
Signa HDx	1.000	0.646	0.873	1.000
MR750	0.622	0.985	0.979	0.875
MR750W	0.410	1.000	1.000	0.818
Prisma	0.570	0.959	0.982	0.859

FIGURE 7
Updating decision results.
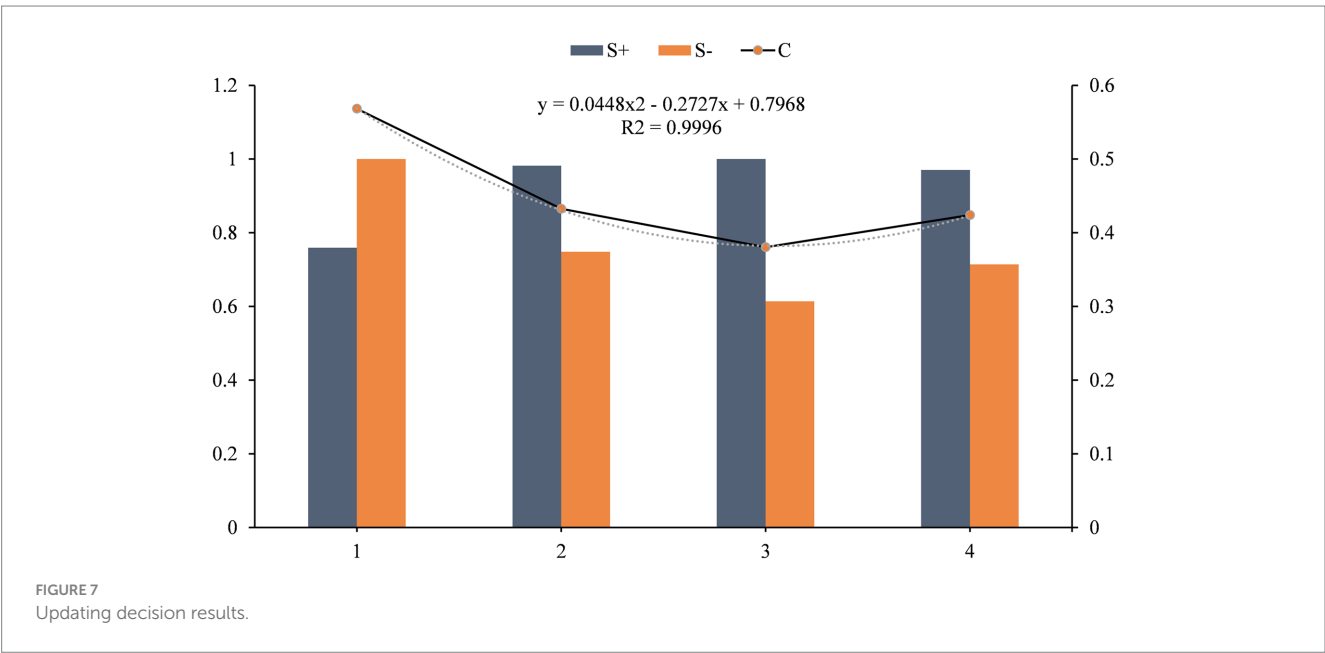
TABLE 4 Updating decision results.

Equipment type	GRA		TOPSIS		GRA-TOPSIS		VIKOR	
	Score	Sort	Score	Sort	Score	Sort	Score	Sort
Signa HDx	0.864	1	0.613	1	0.568	1	0.000	1
MR750	0.756	3	0.393	2	0.433	2	0.831	3
MR750W	0.707	4	0.296	4	0.380	4	0.807	2
Prisma	0.742	2	0.379	3	0.424	3	0.990	4

FIGURE 8
Sorting chart of replacement results.
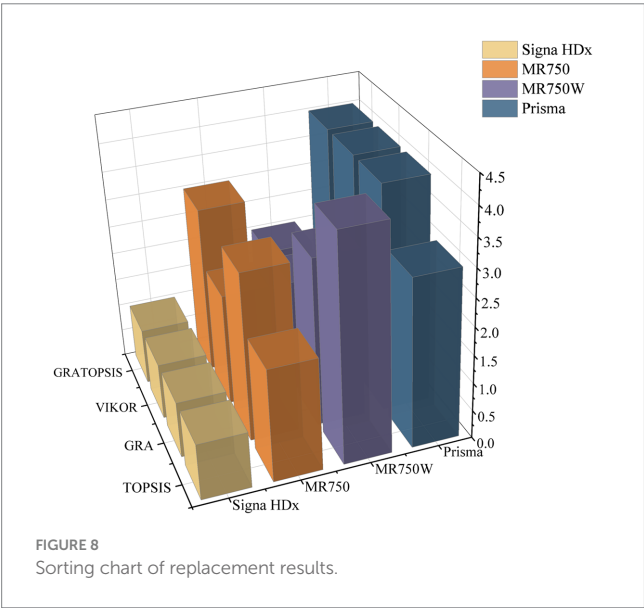
TABLE 5 Kendall's result.

Methods	Rank average	Median	Kendall's W	X^2	p
GRA	3	0.749			
TOPSIS	1.25	0.386	0.812	6.5	0.039
GRA-TOPSIS	1.75	0.428			

Building on this foundational methodology, the research introduces an integrative multicriteria decision model, combining the attributes of COWA-PCA and GRA-TOPSIS. The GRA method, adept at handling sparse and fragmented data sets, aligns seamlessly with the attributes of TOPSIS, which excels in analyzing multi-attribute decision matrices. The union of gray correlation (from GRA) with Euclidean distance (from TOPSIS) ensures a model capable of addressing both uncertainty and information incompleteness. This fusion not only circumvents the limitations of unilateral approaches but also amplifies the accuracy and robustness of evaluations regarding equipment renewal priorities.

Significant preparation and groundwork has been invested to ensure that the research has solid utility and enhances the full lifecycle management of large medical devices. The deployment of advanced IoT collectors, the development of interfaces with third-party systems, and rigorous database management have combined to create a robust dataset. Although the proposed model requires some computational power and computational cost, in terms of economic and social benefits, the model can better guide the management of medical

devices, thus saving costs and improving service quality for healthcare organizations, and bringing considerable economic dividends to hospitals, thus achieving greater economic benefits. The scalability and replicability of this project offer great prospects for subsequent data-centric exploration in this area.

For empirical validation, data from MRI equipment in a selected third-tier hospital was utilized. The results revealed that the GRA-TOPSIS method returned an R^2 value of 0.9667, attesting to its alignment with empirical realities. In addition, Kendall's test validates the robustness of GRA-TOPSIS with respect to the GRA and TOPSIS methods, emphasizing its usefulness as an effective tool for decision-making on equipment replacement in various industries.

6 Conclusion

This article constructs a comprehensive medical equipment replacement priority evaluation indicator system and proposes a comprehensive multicriteria decision model based on COWA-PCA and GRA-TOPSIS. In practical application, the medical equipment replacement priority evaluation model proposed in this paper can help hospital managers and policymakers to better understand and evaluate the need for equipment replacement so as to formulate more scientific and rational replacement strategies. Comparing and analyzing the equipment replacement priorities of different hospitals can provide a basis for resource allocation and policy formulation. In addition, the findings of this paper can also inspire equipment replacement decisions in other fields.

In future, research could further expand the study area and methodology to include a wider and more diverse range of data sources, diversifying the range of evaluation metrics to capture a broader range of operational realities and patient-centered outcomes. Continuous refinement and integration of evaluation methods is essential to improve the accuracy, robustness and generalisability of evaluation models. In addition, with the development of big data technology, artificial intelligence technology and other technologies (38, 39), in hopes of building an intelligent management platform for large medical equipments, ensuring a more scientific, rational and effective decision-making process driven by data. Through in-depth research and practical application, it is expected to provide more scientific, rational and effective support for medical equipment decision-making.

Data availability statement

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Jayawardena D. Hospital equipment management in district base hospitals in Kalutara district in Sri Lanka. *Biomed Stat Inform*. (2017) 2:18–21. doi: 10.11648/j.bsi.20170201.14

2. Liu G, Li N, Chen L, Yang Y, Zhang Y. Registered trials on artificial intelligence conducted in emergency department and intensive care unit: a cross-sectional study on ClinicalTrials.gov. *Front Med*. (2021) 8:265. doi: 10.3389/fmed.2021.634197

3. Patil PJ, Patil SP, Jaltade VG, Gupta SS. Departmental equipment maintenance system in government medical college. *Int Arch Integr Med*. (2015) 2:79–86.

4. Zamzam AH, Al-Ani AKI, Wahab AKA, Lai KW, Satapathy SC, Khalil A, et al. Prioritisation assessment and robust predictive system for medical equipment: a comprehensive strategic maintenance management. *Front Public Health*. (2021) 9:782203. doi: 10.3389/fpubh.2021.782203

5. Liu M, Qin X, Pan J. Does medical equipment expansion lead to more diagnostic services? Evidence from China's Sichuan Province. *Emerg Mark Financ Trade*. (2017) 53:1289–300. doi: 10.1080/1540496X.2016.1247689

6. Cao H, Zhang J, Liu Y. Exploration of medical equipment management development in public hospitals. *Zhongguo Yi Liao Qi Xie Za Zhi*. (2019) 43:65–8. doi: 10.3969/j.issn.1671-7104.2019.01.017 (In Chinese)

7. McDermott O, Antony J, Sony M, Looby E. A critical evaluation and measurement of organisational readiness and adoption for continuous improvement within a medical device manufacturer. *Intern J Manag Sci Engineering Manag*. (2023) 18:145–55. doi: 10.1080/17509653.2022.2073917

8. Vishwakarma A, Dangayach GS, Meena ML, Gupta S, Luthra S. Adoption of blockchain technology enabled healthcare sustainable supply chain to improve healthcare supply chain performance *Management of Environmental Quality*. An Intern J. (2023) 34:1111–28. doi: 10.1108/MEQ-02-2022-0025

9. Zavadskas EK, Turskis Z. Multiple criteria decision making (MCDM) methods in economics: an overview. *Technol Econ Dev Eco*. (2011) 17:397–427. doi: 10.3846/20294913.2011.593291

10. Ivlev I, Kneppo P, Bartak M. Multicriteria decision analysis: a multifaceted approach to medical equipment management. *Technol Eco Dev Eco*. (2014) 20:576–89. doi: 10.3846/20294913.2014.943333

11. Stević Ž, Pamučar D, Puška A, Chatterjee P. Sustainable supplier selection in healthcare industries using a new MCDM method: measurement of alternatives and ranking according to COmpromise solution (MARCOS). *Comput Ind Eng*. (2020) 140:106231. doi: 10.1016/j.cie.2019.106231

12. Hutagalung AO, Hasibuan S. Determining the priority of medical equipment maintenance with analytical hierarchy process. *Int J Online Biomed Eng*. (2019) 15:107. doi: 10.3991/ijoe.v15i10.10920

13. Yang JL, Tzeng G-H. An integrated MCDM technique combined with DEMATEL for a novel cluster-weighted with ANP method. *Expert Syst Appl*. (2011) 38:1417–24. doi: 10.1016/j.eswa.2010.07.048

14. Barrios MAO, De Felice F, Negrete KP, Romero BA, Arenas AY, Petrillo A. An AHP-topsis integrated model for selecting the most appropriate tomography equipment. *Int J Inf Tech Decis*. (2016) 15:861–85. doi: 10.1142/S021962201640006X

15. Mazloum Vajari S, Masoudi Asl I, Hajinabi K, Riahi L. Medical equipment replacement planning using the swot-anp-waspas hybrid approach. *J Healthc Manag*. (2019) 10:91–107.

16. Houria ZB, Masmoudi M, Hanbali AA, Khatrouch I, Masmoudi F. Quantitative techniques for medical equipment maintenance management. *Eur J Ind Eng*. (2016) 10:703–23. doi: 10.1504/EJIE.2016.081017

17. Mora-García T, Piña-Quintero F, Ortiz-Posadas M. Medical equipment replacement prioritization using multi-criteria decision analysis. *Prog Artif Intell Pattern Recognit*. (2018) 11047:271–9. doi: 10.1007/978-3-030-01132-1_31

18. Faisal M, Sharawi A. Prioritize medical equipment replacement using analytical hierarchy process. *IOSR J Electric Electron Eng*. (2015) 10:55–63. doi: 10.9790/1676-10325563

19. Mora-García TR, Hernández-López LA, Piña-Quintero MF, Pimentel-Aguilar AB, Ortiz-Posadas MR. Comparative analysis of two indicators of technical evaluation for the replacement of medical equipment In: . *VIII Latin American conference on biomedical engineering and XLII National Conference on biomedical engineering*. Cancún, México: Springer (2019). 1357–64.

20. Ji L. A comprehensive evaluation system of configuration feasibility for large medical equipment. *Chinese J Hospital Administration*. (2012) 12:47–50. (In Chinese).

21. Miri Lavassani K, Iyengar R, Movahedi B. Multi-tier analysis of the medical equipment supply chain network: empirical analysis and simulation of a major rupture. *BIJ*. (2023) 30:333–60. doi: 10.1108/BIJ-02-2021-0095

22. Ji-Zhuo P, Wei-Xuan W, Jun X, Ting-Ting Y, Zhang Z-C, Cheng-cheng L, et al. Evaluation of large-scale medical equipment in public hospitals in China from the perspective of value-based medicine: a Delphi study (2022). doi: 10.21203/rs.3.rs-2171870/v1,

23. Bhalaji RKA, Sankaranarayanan B, Alam ST, Ibne Hossain NU, Ali SM, Karuppiah K. A decision support model for evaluating risks in a collaborative supply chain of the medical equipment manufacturing industry. *Supply Chain Forum: An International Journal Taylor & Francis*. (2022) 23:227–51. doi: 10.1080/16258312.2021.1989268

24. Taghipour S, Banjevic D, Jardine AK. Prioritization of medical equipment for maintenance decisions. *J Oper Res Soc*. (2011) 62:1666–87. doi: 10.1057/jors.2010.106

25. Malmir, B, Dehghani, S, Jahantigh, FF, Najjartabar, M. (2016). "A new model for supply chain quality management of hospital medical equipment through game theory," in *Proceedings of the 6th international conference on information systems, logistics and supply chain, ILS*. 1–9.

26. Aggarwal M. Compensative weighted averaging aggregation operators. *Appl Soft Comput*. (2015) 28:368–78. doi: 10.1016/j.asoc.2014.09.049

27. Su G, Jia B, Wang P, Zhang R, Shen Z. Risk identification of coal spontaneous combustion based on COWA modified G1 combination weighting cloud model. *Sci Rep*. (2022) 12:2992. doi: 10.1038/s41598-022-06972-4

28. Pabalkar, V, Chanda, R, Sachin, J (2022). "Refurbished Medical Imaging Equipment through Technology," *2022 Interdisciplinary Research in Technology and Management (IRTM)*, Kolkata, India, 1–10.

29. Stević Ž, Miškić S, Vojinović D, Huskanović E, Stanković M, Pamučar D. Development of a model for evaluating the efficiency of transport companies: PCA–DEA–MCDM model. *Axioms*. (2022) 11:140. doi: 10.3390/axioms11030140

30. Daniele P, Sciacca D. A two-stage variational inequality formulation for a game theory network model for hospitalization in critic scenarios In: . Amorosi L, Dell'Olmo P, Lari I editors. *Optimization in Artificial Intelligence and Data Sciences*. Cham: Springer (2022). 14–7.

31. Peng J, Zhang J. Urban flooding risk assessment based on GIS-game theory combination weight: a case study of Zhengzhou City. *Int J Disast Risk Re*. (2022) 77:103080. doi: 10.1016/j.ijdrr.2022.103080

32. Chen Z, Shen D, Ren Y, Yu F, Yuan X. Airspace operation effectiveness evaluation based on q-rung Orthopair probabilistic hesitant fuzzy GRA and TOPSIS. *Symmetry*. (2022) 14:242. doi: 10.3390/sym14020242

33. Lu H, Zhao Y, Zhou X, Wei Z. Selection of agricultural machinery based on improved CRITIC-entropy weight and GRA-TOPSIS method. *PRO*. (2022) 10:266. doi: 10.3390/pr10020266

34. Chen W, Zeng S, Zhang E. Fermatean fuzzy IWP-TOPSIS-GRA multi-criteria group analysis and its application to healthcare waste treatment technology evaluation. *Sustainability*. (2023) 15:6056. doi: 10.3390/su15076056

35. Gupta AD, Rafi SM, Rajagopal BR, Milton T, Hymlin S. Comparative analysis of internet of things (IoT) in supporting the health care professionals towards smart health research using correlation analysis. *Bull Env Pharmacol Life Sci*. (2022) 1:701–8.

36. Mardani A, Zavadskas EK, Govindan K, Amat Senin A, Jusoh A. VIKOR technique: a systematic review of the state of the art literature on methodologies and applications. *Sustainability*. (2016) 8:37. doi: 10.3390/su8010037

37. Clarke GM. The genetic basis of developmental stability. IV. Individual and population asymmetry parameters. *Heredity*. (1998) 80:553–61. doi: 10.1046/j.1365-2540.1998.00326.x

38. Kumar A, Mani V, Jain V, Gupta H, Venkatesh VG. Managing healthcare supply chain through artificial intelligence (AI): a study of critical success factors. *Comput Ind Eng*. (2023) 175:108815. doi: 10.1016/j.cie.2022.108815

39. Wang L, Zhang Y, Wang D, Tong X, Liu T, Zhang S, et al. Artificial intelligence for COVID-19: a systematic review. *Front Med*. (2021) 8:1457. doi: 10.3389/fmed.2021.704256

Digital twin for multi-scenario emergency of railway passenger stations

Xiaoshu Wang[1], Wei Bai[2]*, Yuanqi Su[3], Guoyuan Yang[2], Chao Li[2], Xiaojun Lv[2], Kaibei Peng[2] and Jun Li[2]

[1]Postgraduate Department, China Academy of Railway Sciences, Beijing, China, [2]Institute of Computing Technology, China Academy of Railway Sciences Corporation Limited, Beijing, China, [3]School of Computer Science and Technology, Xi'an Jiaotong University, Xi'an, China

Emergency disposal is a critical aspect for railway stations to ensure safety. This requires the implementation of emergency plan simulations and cost-effective immersive drills. In the paper, we incorporate a set of disposals for events into emergency processes and model the personnel, supplies, and equipment to create multiple emergency scenarios. Additionally, we introduce a digital twin-based solution for multiple scenarios of emergencies. This solution completely restores the key components of a station in the information space and provides an immersive way for emergency disposals. The four-dimensional model used in the solution simulates and interacts with the station, and it is composed of the details of passenger station physical entity, multi-scenario emergency virtual entity, digital twin connection, and emergency twin service. The digital twin for versatile emergency events such as fire disaster, natural disaster (e.g., flood, earthquake), social security incident, and public health event have been constructed using the model. The solution was tested at Qinghe, a station during Beijing-Zhangjiakou high-speed railway in China. By utilizing the actual operation data of Qinghe, we validated the multi-scenario emergency drills, simulated the corresponding emergency disposal plans, and assessed the proposed solution from three aspects: the fidelity of the real-world simulation, coverage of the multiple emergency scenarios, and the user-friendliness. The evaluations indicate that the proposed solution attains *good* score and the *acceptable* level of system usability scale.

KEYWORDS

railway passenger stations, digital twin, emergency, emergency drill, disposal plan

1 Introduction

In China, the safety of passengers who board and alight at train stations is of utmost importance. Emergency events [1] such as fire disaster, natural disaster, public health event, and social security incident directly threaten the passengers' lives and property. According to regulations [2,3], station staff address these issues through emergency measures that comprise regular drills and validation of plans during emergencies. Existing emergency drill uses either tabletop exercises or live drills to enhance response capabilities, but they both have limitations. Tabletop exercises lack an immersive experience, while live drills waste resources and are limited in simulating multiple potential hazards. Additionally, management cannot ascertain the effectiveness of response plans during emergencies, nor assess their scientific validity. Recently, the

research [4] proposes to use IoT for Railway Emergency Detection and Response System. In China, currently the Intelligent Railway Passenger Station (IRPS) [5] system is used to enable real-time tracking of emergency processes and develop emergency drill plan for staff. Additionally, research has focused on railway 3-dimensional (3D) exercises, providing staff with single-scenario training [6,7]. To provide staff with a method to instantly validate emergency plans and offer immersive training in multiple disaster scenarios, we introduce digital twin technology. By replicating the real-world scenarios in a virtual environment and using actual data from station personnel, equipment, trains, passengers, and the environment, we enable the simulation of emergency processes and immersive training for various disaster events. The primary contributions of this paper can be summarized as follows.

- This study builds upon existing station emergency management protocols to meticulously construct a comprehensive outline of multi-emergency scenario, including fire, natural disaster (such as flood, earthquake), social security incident, and public health event. This outline provides a systematic approach to analyze and manage these situations.
- In analyzing emergency disposal process, this paper employs the Generalized Stochastic Petri Net (GSPN) [8] model. We have transformed each step of the response process into places, tokens, and transitions within the GSPN. The model calculates the duration of each transition through parallel, branching, and sequential structures. GSPN is then formulated as a Markov chain. It enables the study to identify weaknesses in the emergency disposal process, providing a basis for personnel adjustments.
- This research successfully integrates digital twin technology into emergency management. In this paper, we use a Digital Twin-Based model for Multi-Scenario Emergency (ED4D) to provide a detailed description of the process, which will furnish digital twin for process entities, facilitate interfaces to interact with reality, and offer services to manage emergency events. This model consists of the Passenger Station Physical Entity, Multi-Scenario Emergency Virtual Entity, Digital Twin Connections and Emergency Twin Services. By replicating the actual situation of the physical world in a virtual environment in real time, we allow the information exchange between both spaces. The characteristics facilitates the deduction of emergency processes and immersive drills for various disaster events.
- The ED4D model was validated at Qinghe Station, encompassing fire disasters, natural disaster (e.g., flood, earthquake), social security incidents, and public health events. By utilizing actual operational data from Qinghe Station (such as passenger numbers, staff, emergency resources, equipment status) through digital twin connections, we confirmed the effectiveness of the solution. This enables us to simulate emergency processes and conduct more efficient emergency drills.

2 Related works

2.1 Modeling techniques for emergency process

There are numerous tools proposed for modeling the common business process, such as the conditional directed graph [9], extended event-driven process Chain [10], Petri Net [8,11], and Integrated Computer Aided Manufacturing [12]. Among these methods, Petri Net and its extension have the strongest ability to explain the stochastic nature of events in business process. As a result, they are well-suited for representing random events that occur with a certain probability in business processes. Petri Net has gained extensive usage in the modeling of processes across diverse fields. In the context of emergency processes, they have found application in modeling emergency decision-making [13], scenario evolution of mass emergency' events [14], emergency response to flood [15], and emergency capability assessment of meteorological disasters [16]. In the field of railway processes, Petri Net has been employed in the modeling of customs clearance analysis of the China-Laos railway [17], emergency dispatching and cooperative disposal of high-speed railways [18], emergency evacuation of complex passenger flow in railway passenger stations [19], as well as safety assessment and risk control [20].

2.2 Digital twin technology

The concept of the digital twin, as referenced in [21], involves creating a replica of the physical world in a digital space. This allows for the exchange of data between the two realms. By connecting with real-world entities, it is able to gather real-time data on the current state of affairs, mimic real-life situations, simulate operational scenarios, make intelligent decisions, and relay this information back to the existing system for control interaction. This technology has found its way into various fields such as industry [22–24], culture [25,26], medicine [27–29], construction [30]; [31], and intelligent maritime transportation [32]. In the railway operation industry, digital twin technology can be leveraged to merge different data sources and models for monitoring [33], turnout management [34], diagnostics [35], and prognostics [36]. It has the advantages of acquiring the actual state, mirroring the real situation, simulating the operation scenario, making generation-assisted decisions, and realizing control interaction.

2.3 Emergency drill based on 3D technology

The advancement of computer technology has facilitated the integration of 3D technology in skill training programs, including automobile maintenance training [37], railway axle disc brake maintenance [38], tunnel fire emergency drills [39], ship fire emergency drills [40], railway workers' skill training [41], and train crew management [42]. The use of 3D technology enables the creation of a natural world in the information space that dynamically simulates disaster occurrences with negative outcomes, such as damages, losses, injuries, and fatalities. This

technology resolves the challenge of large resource consumption and enhances the sense of immersion.

2.4 Systems usability evaluation

User evaluation is a common approach to determine whether a software application is meeting its intended usage requirements. The system usability scale (SUS) [43] is a widely accepted questionnaire standard that consists of ten questions designed to assess usability. The after-scenario questionnaire [44] is composed of three questions that evaluate time, difficulty, and information support. The usability metric for user experience [45] is a rating system that involves four questions covering three dimensions: usefulness, efficiency, and satisfaction. The Happiness-Engagement-Adoption- Retention-Task Success (HEAR) framework [46] is an evaluation tool used for internal enterprise software applications, which assesses the time required to complete tasks and their level of difficulty.

3 The modeling of multi-scenario emergency

Multi-Scenario Emergency is a comprehensive framework that encompasses various types of emergency situations. A scenario, which is a delineation of emergency events, serves as the impetus for emergency response procedures and the mobilization of resources by personnel to execute a series of disposal actions aimed at mitigating and controlling the escalation of the situation. There are several types of emergencies in the station, categorized into four groups[1]: fire disaster, natural disaster, social security incident, and public health event. In the paper, we focus on five events for Multi-Scenario Emergency: fire, flood, earthquake, social security incident, and public health event. Although the fundamental process of emergency disposal remains constant for different events, the disposal personnel, resources and tasks differ. This paper provide a detailed outline of the disposal scenes for various emergency events in Section 3.1 and present a GSPN to describe the process of emergency disposal in Section 3.2, thereby defining a multi-scenario emergency disposal model.

3.1 Multiple emergency scenarios

Scenarios consist of a combination of events, equipment, personnel, resources, and individual actions. In accordance with the chronological progression of events, personnel execute a series of action measures to steer the events towards the intended outcome. In relation to transportation stations confronted with circumstances such as fire, flood, earthquake, social security incident, and public health event, we have undertaken an evaluation of emergency procedures, outlining personnel requirements, resource allocation, equipment utilization, and disposal methodologies. The specifics are presented as follows.

- **Fire disaster.** In the occurrence of a fire alarm signal, the passenger station is responsible for taking prompt measures to guarantee the security of its passengers. The primary action involves reporting the occurrence to the station manager, who will evaluate the extent of the incident and determine whether to initiate the emergency response plan. Pursuant to the plan, station employees must conduct various tasks, including consoling passengers, safely evacuating them, contacting local rescue services for aid, utilizing firefighting equipment, and extinguishing the fire. After the fire has been contained, the site must be restored, and a post-event summary should be created to evaluate the efficiency of the emergency response.

- **Natural disaster.** Upon receipt of an earthquake or severe weather alarm, the station manager initiates an assessment of the situation to determine the appropriate emergency response level. Subsequently, in compliance with the emergency plan's provisions, the station staff undertakes emergency tasks, including passenger reassurance, passenger evacuation, itinerary modifications, coordination with local rescue teams, acquisition of emergency resources, on-site rescue operations, and other related duties. Post-disaster, the site is restored to its previous state, and the emergency is concluded upon completion of the restoration process.

- **Social security incident.** Upon discovery of violent acts at the train stations, the staff promptly notifies and reports to the station manager. The manager evaluates the situation and implements the emergency response plan in accordance with established guidelines. The staff carries out emergency duties such as passenger evacuation, coordination with the police department, and on-site rescue operations. Once the offenders are apprehended, the site is restored to its original state. Ultimately, a post-incident analysis is performed.

- **Public health event.** When a case that is suspected to involve a contagious disease is reported, the station without delay informs the station manager for evaluation. Depending on the assessment, the station personnel are required to undertake several emergency tasks, which include, but are not limited to, providing reassurance to passengers, ensuring their safe evacuation, reaching out to the local disease control department, implementing on-site isolation measures, and more. Once all the individuals suspected of carrying the disease have been transported, the site must be restored to its original state, and a post-incident review must be carried out to determine the effectiveness of the response and identify areas for improvement.

To provide a succinct overview, the alarm circumstances exhibit a range of fluctuations in accordance with the emergency situations. In the course of these situations, personnel undertake assorted duties, necessitate a diverse array of external rescue measures, and obtain access to a variety of emergency resources. The specifics pertaining to these factors are concisely encapsulated in Table 1.

TABLE 1 Unique entity details for each scenario emergency of station.

| Class	Fire disaster	Natural disaster	Social security incident	Public health event
Aid workers	doctors, firemen	ministry of emergency management, doctors, firemen, policemen	policemen	Staffers of Government epidemic prevention department, doctors
equipment and facilities (*EQU*)	Fire extinguisher, fire hydrant, spray, blister, curtain door, smoke alarm	Pump and emergency light	blasting protection	Disinfection sprayers, isolation belts, protective articles[1]
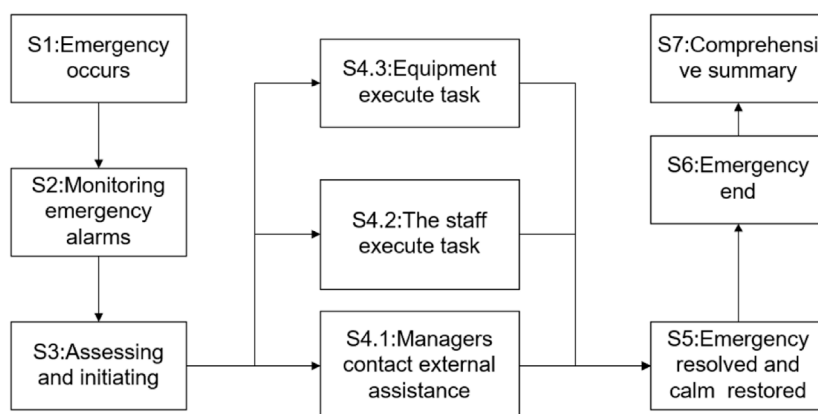
FIGURE 1
The common process of the emergency disposal.

3.2 The GSPN model for multi-scenario emergence processes

A strategy for managing various emergency situations is established as a common disposal protocol, as depicted in Figure 1. The procedure is initiated by monitoring emergency alarms and determining their seriousness level, as illustrated by S1, S2 and S3 in Figure 1. During an emergency, the station activates multiple staff members to participate in the disposal effort, as exemplified by S4.1, S4.2 and S4.3 in Figure 1. Once the situation has been resolved and the site has been restored, the emergency will be concluded, followed by a comprehensive summary of the emergency process, as illustrated by S5, S6 and S7 in Figure 1. The aforementioned process can be represented as a sequence S1 and S2 as shown in Figure 1, branching S3 and S4 as depicted in Figure 1, and parallel parts S4.1, S4.2 and S4.3 as illustrated in Figure 1. These components can be effectively portrayed by using a Petri Net [8,11].

Petri Net [8,11], a graphical mathematical modeling tool, is comprised of four essential components: place, transition, token, and directed arc. Place represents the current state of the system, transition signifies state changes, and token represents system resources. Directed arc, on the other hand, illustrates the relationship between Place and transition while also indicating the direction of token flow. However, Petri Net is not equipped to handle cases where the transition is a random delay, and the state space grows exponentially with the increase of the problem. The GSPN addresses this issue [18,47,48]. The introduction of time

values and random numbers in GSPN enhances Petri net, allowing for the modeling of features such as random delay, variable time, and randomness. Therefore, we adopt it to model the comprehensive emergency process. As the emergency process proceeds, tasks must be continuously adapted to site conditions, and each task may require varying amounts of time to complete.

The GSPN for emergencies disposal [15] is a directed graph described by eight elements:

$$GSPN^i = (P, T, F, W, M, \lambda, PIT, BT) \qquad (1)$$

Here, the raised index *i* is used to differentiate a variety of emergency occurrences and assumes values from a set of {1, 2, 3, 4, 5}. Specifically, 1 represents fire, 2 flood, 3 earthquake, 4 social security incident, and 5 public health event. The meaning of the eight constituents is outlined as follows.

- *P* denotes a finite collection of *places* that serve as a container for tokens. In this article, the term *Place* specifically refers to different sets of information, while *token* denotes specific pieces of information.
- *T* represents a finite assortment of *transitions* that bring about changes in emergency situations.
- *F* comprises the set of *directed arcs* that establish a connection between *P* and *T*.
- *W* is the collection of arc weights that specify the maximum number of tokens that can traverse each arc.
- *M* serves as a status flag, representing the quantity of tokens within each *P*, for each state.
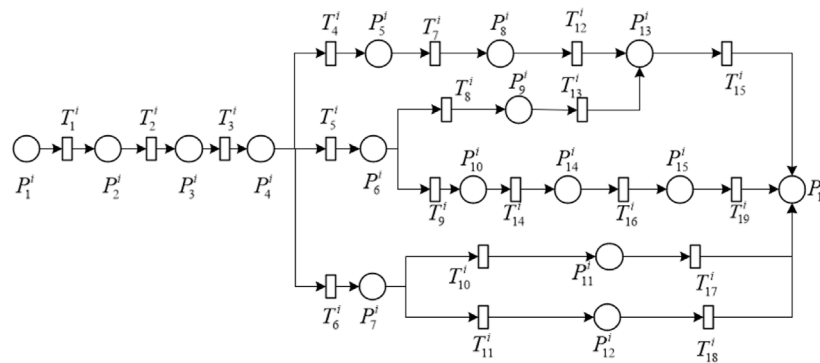
FIGURE 2
The GSPN of the emergency disposal.

- λ is associated with the *occurrence probability*. Upon fulfillment of specific conditions, the related transitions occur; the occurrence of the transition is independent of each other. λ then parametrizes the distribution function for the occurrence.
- *PIT* signifies the moment when the transition undergoes a change, represented by $\{Time_i\}$.
- *BT* keeps track of the time delay for each transition in *T*.

The entire process, beginning with monitoring and ending with the emergency summary, is modeled by GSPN, which encompasses all stages, as indicated in Figure 2. The GSPN model is made up of 16 place and 19 transitions, which effectively represent all crucial aspects of the emergency process within a restricted set. Further information regarding *P*, *T*, *PIT*, *BT* is available in Table 2.

The initial data set P_1 is made available at $Time0$, and the values of $Time_i$ are derived using Eq. 2. Through calculations, we can establish the corresponding point-in-time for every P_k, which in turn allows us to make an estimate of the total time required for the emergency process. As illustrated in Figure 2, the time for P_8 is determined by the leading P_5 as the start time, plus the duration needed to execute T_7. The start time of P_{13}, is determined as the maximum of two sums; one is the sum of the start time of P_8 plus the duration needed to execute T_{12}, and the other is the start time of P_9 plus the duration needed to execute T_{13}. The P_{13} can only start after the completion of both branches.

The transfer of the token from one *P* to another signifies a shift in the system's state. We set the flag for token changes in GSPN as denoted by vector set *M* in Eq. 3. The magnitude of vector M_i corresponds to the quantity of *P*. Each element in the vector M_i represents the quantity of tokens present in the *P* [49]. In the paper, the *M* is comprised of a set of 15 vectors, and each vector has 16 elements, as depicted in Eq. 3.

The model's initial sign is M_1, upon the occurrence of a token in P_1 that means the events occur alarm information, the transition T_1 is triggered with probability λ_1 and the token is transferred to P_2. The transition T_1 requires BT_1 time units and results in a change from M_1 to M_2, as depicted by the sequence shown in Figure 1 S1 to S2. In this moment, the token is in P_2 and the status is M_2 as Eq. 3, which means the event status exceeds the safety threshold. Similarly, the fourth status change is indicated by M_4, where P_4 possesses a token

and signifies the current state of emergency disposal plan determination.

$$\begin{cases} Time1 + BT_1 = Time2 \\ Time2 + BT_2 = Time3 \\ Time3 + BT_3 = Time4 \\ Time4 + BT_4 = Time5 \\ Time4 + BT_5 = Time6 \\ Time4 + BT_6 = Time7 \\ Time5 + BT_7 = Time8 \\ Time6 + BT_8 = Time9 \\ Time6 + BT_9 = Time10 \\ Time7 + BT_{10} = Time11 \\ Time7 + BT_{11} = Time12 \\ max(Time8 + BT_{12}, Time9 + BT_{13}) = Time13 \\ Time10 + BT_{14} = Time14 \\ Time14 + BT_{16} = Time15 \\ max(Time13 + BT_{15}, Time15 + BT_{19}, Time11 + BT_{17}, Time12 + BT_{18}) = Time16 \end{cases}$$

(2)

Subsequently, the token in P_4 split into three, as depicted by the branching shown in Figure 1 (S3 to S4), transitioning the status from M_4 to M_5. The transition T_4 describes one token changes from P_4 to P_5, which represents the assistance seeking from Aid workers. The transition T_5 describes other token changes from P_4 to P_6, which denote staff task reception. The transition T_6 describes another token changes from P_4 to P_7, which indicate station equipment task reception. M_5 corresponds to the fifth state change, with P_5, P_6, and P_7 containing one token each as Eq. 3, and the parallel shown in Figure 1 (S4.1, S4.2 and S4.3).

$$\begin{cases} M_1 = (1,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0) \\ M_2 = (0,1,0,0,0,0,0,0,0,0,0,0,0,0,0,0) \\ M_3 = (0,0,1,0,0,0,0,0,0,0,0,0,0,0,0,0) \\ M_4 = (0,0,0,1,0,0,0,0,0,0,0,0,0,0,0,0) \\ M_5 = (0,0,0,0,1,1,1,0,0,0,0,0,0,0,0,0) \\ M_6 = (0,0,0,0,0,1,1,1,0,0,0,0,0,0,0,0) \\ M_7 = (0,0,0,0,1,0,1,0,1,1,0,0,0,0,0,0) \\ M_8 = (0,0,0,0,1,1,0,0,0,0,1,1,0,0,0,0) \\ M_9 = (0,0,0,0,0,0,1,1,1,1,0,0,0,0,0,0) \\ M_{10} = (0,0,0,0,1,0,0,0,1,1,1,1,0,0,0,0) \\ M_{11} = (0,0,0,0,0,0,0,0,1,1,1,1,1,0,0,0) \\ M_{12} = (0,0,0,0,0,0,0,0,1,1,0,1,1,0,1,0,0) \\ M_{13} = (0,0,0,0,0,0,0,0,1,1,0,1,1,0,0,1,0) \\ M_{14} = (0,0,0,0,0,0,0,0,0,0,0,1,1,1,0,1,0) \\ M_{15} = (0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,1) \end{cases}$$

(3)

TABLE 2 P (Place), PIT, T (Transition) and delay meanings.

P(Place)	Meaning	PIT	T (Transition)	Meaning	Delay
P_1	Alarm information of the events occur	Time1	T_1	The event impact is increasing	BT_1
P_2	Alarm information of the event status exceeds the safety threshold	Time2	T_2	Disaster evaluation and emergency disposal are beginning	BT_2
P_3	Emergency information on disaster assessment and activation	Time3	T_3	The emergency response plan and problem are being generated	BT_3
P_4	Emergency solution information	Time4	T_4	Staff is seeking assistance from aid workers	BT_4
P_5	Information of Aid workers arrived station	Time5	T_5	Staff members are receiving the assigned tasks	BT_5
P_6	Information of station staff on site	Time6	T_6	The station equipment is receiving the assigned	BT_6
P_7	Information of Station equipment receives task feedback	Time7	T_7	Aid workers are developing rescue plans	BT_7
P_8[a]	Information of Aid workers feedback	Time8	T_8	Staff is guiding passengers with the evacuation process	BT_8
P_9	Feedback information of station staffers guided passengers evacuation	Time9	T_9	The staff is working on duty	BT_9
P_{10}	Feedback information of station staff execute task	Time10	T_{10}	The station equipment is executing commands	BT_{10}
P_{11}[b]	Information of station equipment finish tasks	Time11	T_{11}	The evacuation device is executing commands	BT_{11}
P_{12}	Information of evacuation devices finish tasks	Time12	T_{12}	Aid workers are implementing rescue plans	BT_{12}
P_{13}	Information of passenger safety	Time13	T_{13}	The passengers are completing the evacuation process and arriving at a safe location	BT_{13}
P_{14}	Feedback information of staff obtain emergency resource	Time14	T_{14}	The staff is acquiring emergency resources and executing tasks	BT_{14}
P_{15}	Feedback information of staff finish tasks	Time15	T_{15}	The staff is checking the safety status of the passengers	BT_{15}
P_{16}	Information of emergency disposal end	Time16	T_{16}	Staff is executing tasks by utilizing emergency resources	BT_{16}
-	-	-	T_{17}	The staff is inspecting the equipment to ensure completion of tasks	BT_{17}
-	-	-	T_{18}	The staff is inspecting the evacuation devices to ensure completion of tasks	BT_{18}
-	-	-	T_{19}	The staff is verifying the completion of field tasks	BT_{19}

[a]P_8 in GSPNa represents firemen; P_8 in GSPNb represents government; P_8 in GSPN3 represents government; P_8 in GSPN4 represents policemen; P_8 in GSPN5 represents doctors.
[b]P_{11} in GSPNa represents extinguisher, fire water monitor, loudspeaker, guide screen, and evacuation equipment; P_{11} in GSPNb represents pump, loudspeaker, guide screen and evacuation equipment; P_{11} in GSPN3 represents emergency light, loudspeaker, guide screen and evacuation equipment; P_{11} in GSPN4 represents explosion-proof blanket, loudspeaker, guide screen and evacuation equipment; P_{11} in GSPN5 represents isolation belt, Sterilizing instrument, loudspeaker, guide screen and evacuation equipment.

In GSPN, token transitions exhibit a certain degree of randomness. Given that the Markov chain is a type of stochastic process, GSPN can be effectively analyzed for system stability using Markov chain methods. GSPN having a finite number of places and tokens are equivalent to a one-dimensional continuous-time Markov chain model [49]. To gain a comprehensive grasp of the entire emergency process, we utilize Markov chain to analyze potential weaknesses. We can obtain Markov chain that are isomorphic to the simplified GSPN model, as shown in Figure 3. The Markov chain is obtained by using the probability occurrence λ_k of the triggered transitions as the directed edge. In branching scenarios, different states are generated due to varying branch duration. Branching starts from P_4, resulting in multiple states corresponding to M_5 and its subsequent transitions. During the transition from M_5 to M_6, M_7, and M_8, due to different branch times for P_5, P_6, and P_7, there is a probability λ_5 that P_5 has completed while P_6 and P_7 remains incomplete, as M_6. The same applies to the

FIGURE 3
The Markov chain is isomorphic to the GSPN model.
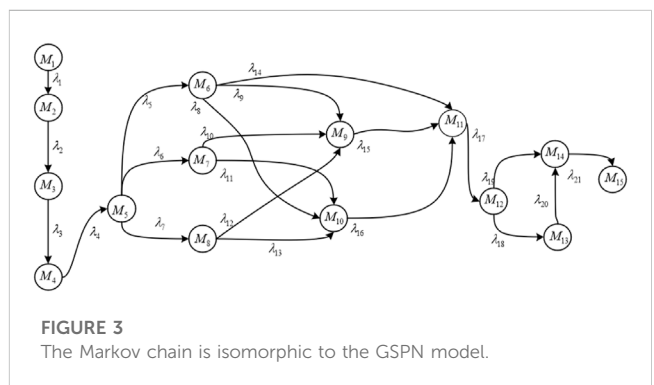
generation of M_7 and M_8. After the completion of P_5 and P_6, M_9 is generated, and after the completion of P_6 and P_7, M_{10} is generated. Therefore, from M_6, M_7, and M_8, we can reach both M_9 and M_{10}. M_9
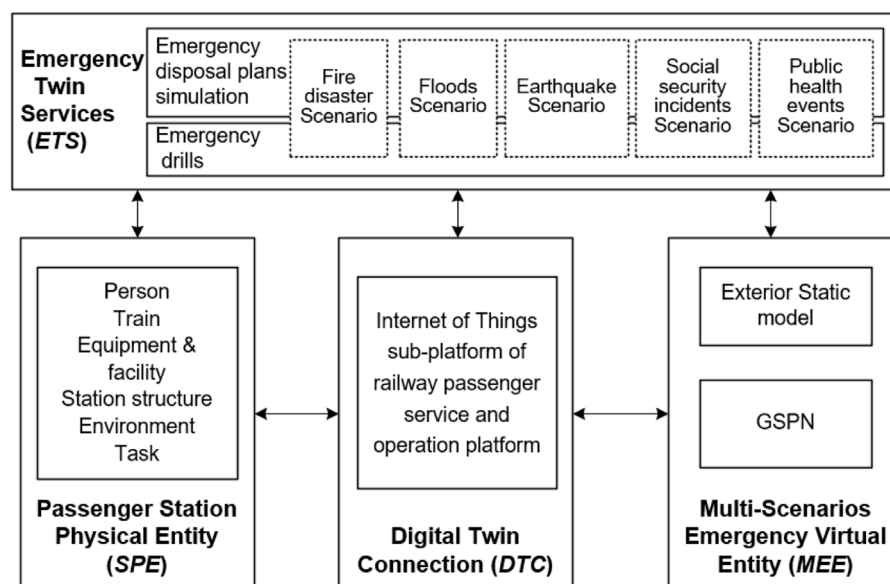
FIGURE 4
Four-dimensional Digital Twin model of Emergency Disposal for passenger station (ED4D).

corresponds to P_7, P_8, P_9 and P_{10}, M_{10} corresponds to P_9, P_{10}, P_{11} and P_{12}.

By utilizing Markov to calculate the average time on each state M_i of the GSPN, the determination of the duration of stay at different states is enabled [49]. Specifically, for the evacuation of passengers in group P_8 and P_9 to P_{13}, the duration BIT is closely linked to the number of passengers present, which poses a high risk of stagnation in state transition. P_{13} can lead to deadlock states in a Petri net, where the system is unable to progress and emergency procedures cannot be executed properly. By analyzing bottleneck P_{13} and promptly adjusting or supplementing them, we can optimize the entire emergency process and ensure smooth emergency handling.

4 Digital twin-based model for multi-scenario emergency

Utilizing the innovative digital twin technology as outlined in [21], we have developed a groundbreaking model consisting of four key components that effectively address emergency scenarios that may arise within passenger stations. This comprehensive model takes into account emergency events, personnel, equipment, resources, and individual actions, with designated personnel implementing a series of emergency response measures in a specific sequence. The 4-dimensional multi-scenario emergency model for passenger stations, known as *ED4D*, is formulated as a set of guidelines that are designed to effectively manage multiple potential emergency situations.

$$ED4D = \{SPE, MEE, DTC, ETS\} \qquad (4)$$

The set *ED4D*, as depicted in Eq. 4, has been established for multi-scenario emergency and disposal at passenger stations. This model consists of four crucial components, namely, the Passenger

Station Physical Entity (*SPE*), Multi-Scenario Emergency Virtual Entity (*MEE*), Digital Twin Connection (*DTC*), and Emergency Twin Service (*ETS*), which are clearly depicted in Figure 4. The following sections provide a detailed explanation of each component. The passenger station physical entity refers to a physical location designed to hold passengers as they board and alight. On the other hand, the multi-scenario emergency virtual entity is a virtual replica connected to its corresponding physical counterpart. The digital twin connection serves as an effective medium for exchanging data between the two realms, while the emergency twin service provides essential services to handle emergencies at the station.

4.1 Passenger station physical entity (*SPE*)

The Passenger Station Physical Entity is comprised of human (*HUM*), train (*TRA*), equipment and facility (*EQU*), station room (*ARC*), and environment (*ENV*), as follows:

$$SPE = \{HUM, TRA, EQU, ARC, ENV\} \qquad (5)$$

The precise details regarding every constituent element have been illustrated in Table 3. Nevertheless, the aid worker, equipment, and facility employed in an emergency situation may vary depending on the type of incident. For additional information of the incidents, please refer to Table1.

4.2 Multi-scenario emergency virtual entity (*MEE*)

In order to create a virtual representation of the passenger station, it is necessary to accurately replicate both the stationary

TABLE 3 Common physical entity details for all scenarios emergency of passenger station.

Class	Details
persons (HUM)	Staff, passengers, Aid workers
trains (TRA)	High-speed train, General-speed trains
equipment and facilities (EQU)	Loudspeaker, platform/Entry/check-in/Guiding screens, self-service ticket checking machine, real-name verification gate, security check instrument, security gate, camera, mobile emergency screen, evacuation indicator light, access control, automatic external defibrillator, help button
station rooms (ARC)	Station square, waiting hall, ticket hall, self-service, comprehensive service, Entrance/Exit/Transfer hall
environment (ENV)	Temperature, humidity, brightness, harmful gases

and moving attributes of physical objects within the digital realm. The multi-scenario emergency virtual entity comprises a Static model of the exterior, as well as a Dynamic Behavior Model that outlines the emergency response process based on the GSPN framework detailed in Section 3.2.

4.2.1 Exterior static model

To obtain a 3D representation and external image of a physical entity, on-site surveying is necessary to obtain geometry and texture information for trains, equipment, and facilities. One must retrieve the necessary geometric and texture information. In the case of a station structure, construction drawings are utilized for basic geometric information, while on-site photos are utilized for texture information. Personnel are differentiated based on their attire, with staffers further differentiated by gender.

4.2.2 Emergency process dynamic behavior model

In Section 3.2, the GSPN framework is utilized to create a multi-scenario emergency disposal process. The model is implemented in C# using object-oriented programming principles and involves ontology attributes and data related to personnel, trains, equipment, station houses, and environment. The emergency process comprises of different stages such as alarm, start of emergency, disposal, recovery, end, and summary. To simulate passenger arrival parameters, a log-normal distribution is used [50]. Staff are determined by the day's schedule, and train scheduling follows the charting plan for the day. Equipment operation is based on the equipment operation plan, and if an emergency disposal process encounters equipment failure, the Boolean value of the equipment status is changed accordingly.

4.3 Digital twin connection (DTC)

The physical entity of the passenger station and the Multi-Scenarios Emergency Virtual Entity both utilize the Internet of Things subsystem of the IRPS for integrated access and management, facilitating connectivity and interaction. In terms of person-to-person connection, staff members use Bluetooth positioning to obtain real-time location information and handheld terminals to send and receive information. The

platform and dispatching system-level interface are utilized to obtain information about train arrivals and departures. Common protocols are used to interactively control electronic equipment such as self-service ticket machines, real-name verification gates, speakers, and various types of screens. Additionally, for system-level interfaces like equipment and trains, we use the HTTP protocol for data acquisition, updating it at predetermined intervals. For real-time train arrival and departure information, the GET method is used, with an approximate cycle of 20 s. For train ticket availability, the POST method is adopted, with a frequency of about once per minute. For passenger and staff information, the POST protocol is utilized, while for equipment, the GET protocol is employed with a 1-min cycle. Fire extinguishers and disinfection sprayers are accessed by Radio Frequency Identification and maintained by staff. Real-time sensors such as smoke sensors and water level meters are used to obtain information about the state of the station environment.

4.4 Emergency twin service (ETS)

The digital twin emergence service offers a range of emergency scenario simulation and drill functions for various events. In terms of simulation, the service employs GSPN to automatically calculate the emergency disposal time and effectiveness based on a range of factors including the onset of the emergency, the station's trains, the number of passengers, the real-time location of staff, and the status of emergency resources and equipment. With regards to drills, the platform facilitates the setting of parameters such as the number of participating passengers, staff, and emergency resources to simulate the emergency disposal process. Both service functions provide real-time information on the status, video inspection, and maintenance of inspected equipment, such as fire extinguishers, fire hydrants, sprinklers, mobile emergency screens, and evacuation instructions.

5 A case study for Qinghe station

At Qinghe station, we conducted the validation of digital twin-based solution. The components of the station psychical entities include HUM (staff and passengers), TRA (high-speed trains), EQU, ARC (one underground level and three above-ground levels), and ENV. Utilizing 3D Max software, we captured on-site real scenes to generate the Exterior Static Model, employing these models to create a multi-scenario virtual entity of Qinghe Station. We utilized Unity3D [7] to develop GSPN processes to create the Emergency Process Dynamic Behavior model, enabling the implementation of all ETS functionalities. DTC was used to interact with reality and obtain real-time data from the Internet of Things subsystem of IRPS, including the train data, tasks, staff information, and more. The ETS can simulate the entire process of emergency disposal and display the number of personnel involved, the nature of resources required for disposal operations, and the duration of the evacuation process. The hardware environment used an Intel(R) Core(TM) i9-10900K CPU@3.7 GHz processor, a NVIDIA GeForce RTX 3080 Ti graphics card with 12 GB memory, 32 GB of DDR4 memory, and the Windows 10 operating system. The final size of this model is 3.8 GB.
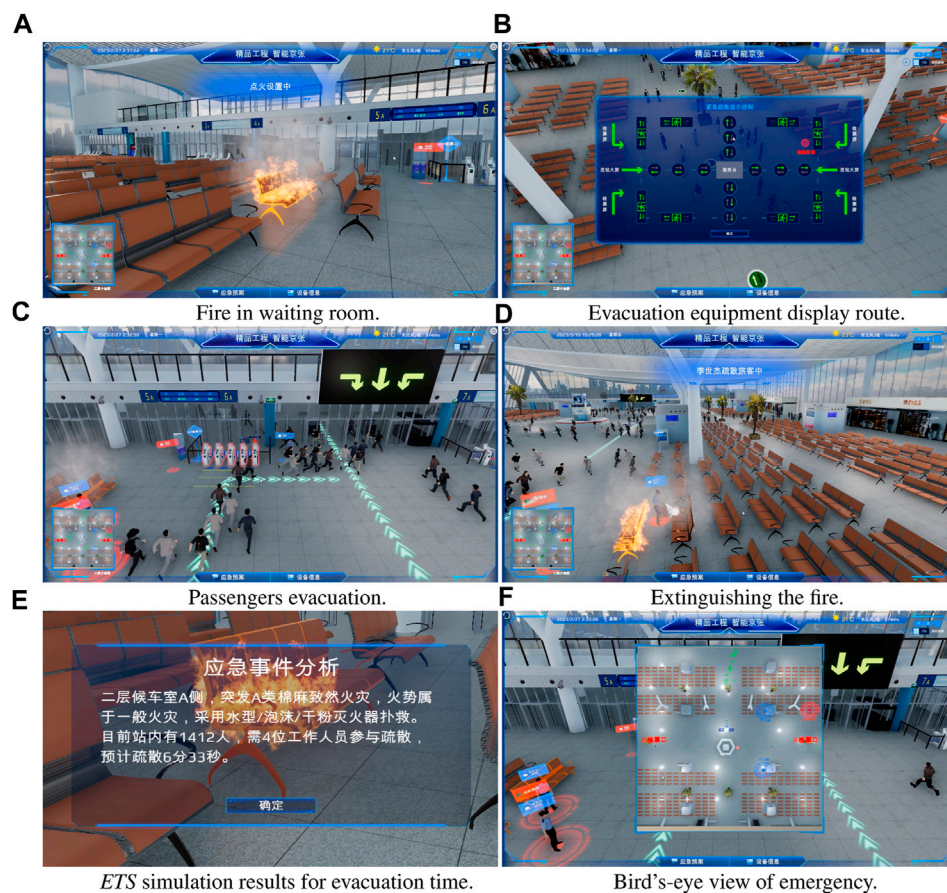
FIGURE 5
Fire scenario. **(A)** Fire in waiting room. **(B)** Evacuation equipment display route. **(C)** Passengers evacuation. **(D)** Extinguishing the fire. **(E)** ETS simulation results for evacuation time. **(F)** Bird's-eye view of emergency.

5.1 Fire

We have developed an extensive plan for a fire emergency, which covers the entire process from its initial outbreak to the conclusion. In the event of a fire, the *ETS* will share information about the burning chair (P_1 in GSPN) as indicated in Figure 5A. The rate of flame combustion may differ based on the material, resulting in varying values for the duration (BT_1). Once the burning surpasses the safety threshold (P_2), the station manager assesses the severity of the situation and activates emergency protocols before any information regarding the emergency is activated (P_3). Next, the staff executes the emergency disposal plan, with some of them seeking help from aid workers, such as firefighters (T_5). The *ETS* then gives information about staff present on the site (P_6), as well as the evacuation equipment executing the emergency disposal (P_7), which provides guidance on the prescribed evacuation route, as shown in Figure 5B. In addition to displaying the staff members responsible for organizing passenger evacuation (T_8), the *ETS* also provides instructions for a group of passengers (T_{13}) to follow and evacuation equipment to use in order to ensure a smooth evacuation process, as illustrated in Figure 5C. Additionally, during the disposal process, the *ETS* displays the staff responsible for extinguishing the fire (T_{14}) and those involved in restoring the scene after the fire has been

extinguished, as depicted in Figure 5D. Finally, the *ETS* delivers the concluding information (P_{16}), and a report on the emergency disposal is produced to document the resolution of the incident.

By simulating emergency disposal plans in the event of a fire scenario, *ETS* can improve the disposal process and achieve desired outcomes. These outcomes are illustrated in Figure 5E, which displays the analysis and results of the simulation of fire disaster emergency disposal plans. The station has 1,412 individuals, and four staff members are responsible for guiding passenger evacuation (T_8). The evacuation takes 6 min and 33 s to complete. Conducting emergency drills for *ETS* in fire scenarios allows staff to practice responding to simulated fire outbreaks in various locations, thereby enhancing their preparedness. It also provides an overview of the entire emergency disposal process, including resource location, disaster location, and passenger evacuation, among other elements, as shown in Figure 5F. The red circle marks the location of the accident, while the blue circle indicates the position of emergency resources.

5.2 Flood

For the scenario of flood, we set up a overall process covering the water-level rise, report, disposal, and conclusion. Initially, the
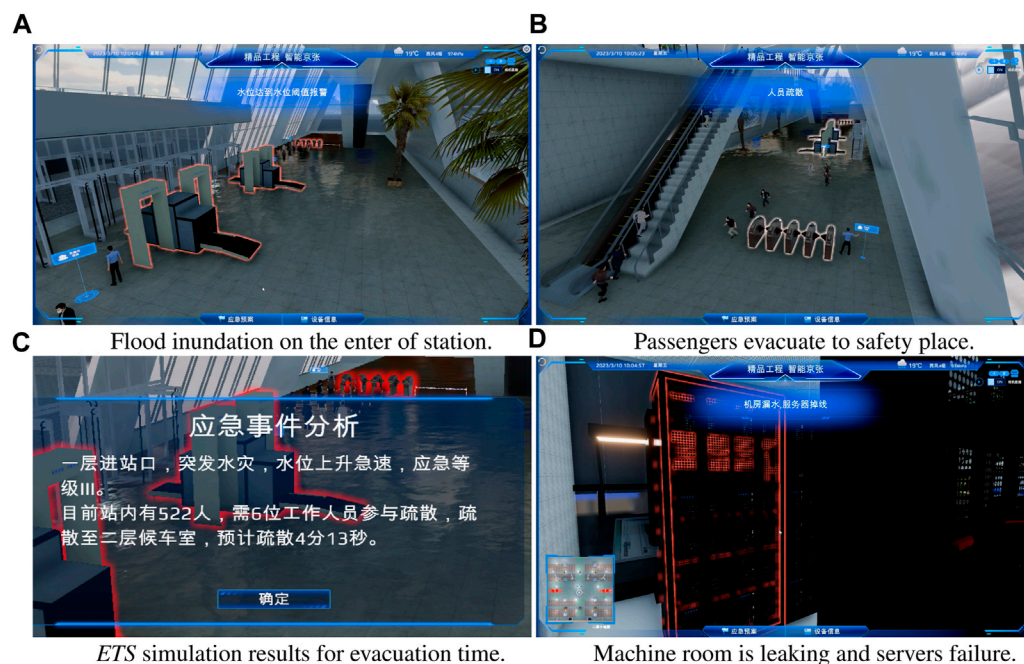
FIGURE 6
Flood scenario. **(A)** Flood inundation on the enter of station. **(B)** Passengers evacuate to safety place. **(C)** ETS simulation results for evacuation time. **(D)** Machine room is leaking and servers failure.

monitoring system provides an overview of the floodwater level entering the interior. As the water level continued to rise (T_1), the floodwater submerged the security check instruments, as depicted in Figure 6A (highlighted in red). The station manager evaluated the severity of the disaster and instructed the implementation of emergency procedures (T_2). At this point, the emergency level was determined to be level 3, and the ETS signaled the initiation of the emergency response (P_3). The emergency tasks were assigned to staff members at each location. These tasks included running to designated locations (P_6) and executing assigned duties (T_8) for some staff members, while equipment displayed emergency information (T_6 and T_{10}), and some staff members directed the evacuation of passengers (P_9).

The passengers adhered to the instructions given by the staff and utilized the evacuation equipment in order to carry out an organized evacuation process (T_{13}), as illustrated in Figure 6B. The staff then proceeded to open the pump (P_{14}) and restore the scene once the pumping was completed at (P_{15}). A report was compiled to document the conclusion of the incident, which is known as the emergency disposal report (P_{16}). Additionally, the emergency disposal plan simulation outlines the plans execution process and results. The flood simulation results, presented in Figure 6C, indicate that the disposal process involved the utilization of pumping pumps, the participation of 6 workers, and the evacuation of 522 individuals from the station. All these actions were successfully executed within a time frame of 4 min and 13 s. Furthermore, conducting emergency drills in flood scenario allows for setting the scale of water and simulating damage to various equipment, such as machine room leakage, as demonstrated in Figure 6D. This facilitates the smooth execution of the flood emergency drill.

5.3 Earthquake

For the scenario of earthquake, passengers experienced a sense of panic upon receiving an earthquake warning (P_1), as illustrated in Figure 7A. The staff on duty immediately responded to the warning and notified the station manager, who evaluated the situation and initiated the emergency protocol (T_3). The assigned staff members were responsible for implementing emergency procedures, including contacting the local earthquake rescue department (T_5), moving quickly to designated locations (T_6), and disseminating information about the emergency train stop (P_7), as demonstrated in Figure 7B. The staff successfully coordinated the evacuation of passengers (P_9), guiding them in an orderly manner using evacuation equipment and staff assistance (T_{13}), as shown in Figure 7C. The entire earthquake scenario was simulated by the solution, which provided valuable insights into the implementation process and the outcomes of the plans. The results of the earthquake simulation, as presented in Figure 7D, indicated that there were 1,478 individuals present in the station during the evacuation, involving the participation of 12 workers and lasting for a duration of 9 min and 46 s.

5.4 Social security incident

The social security situation is employed to tackle the disturbance created by individuals who present a threat to the station. An undisciplined person set fire to their carry-on luggage (T_1), as illustrated in Figure 8A. The responsible staff detected it and immediately notified the station manager to commence emergency disposal. The manager evaluated and initiated the emergency response
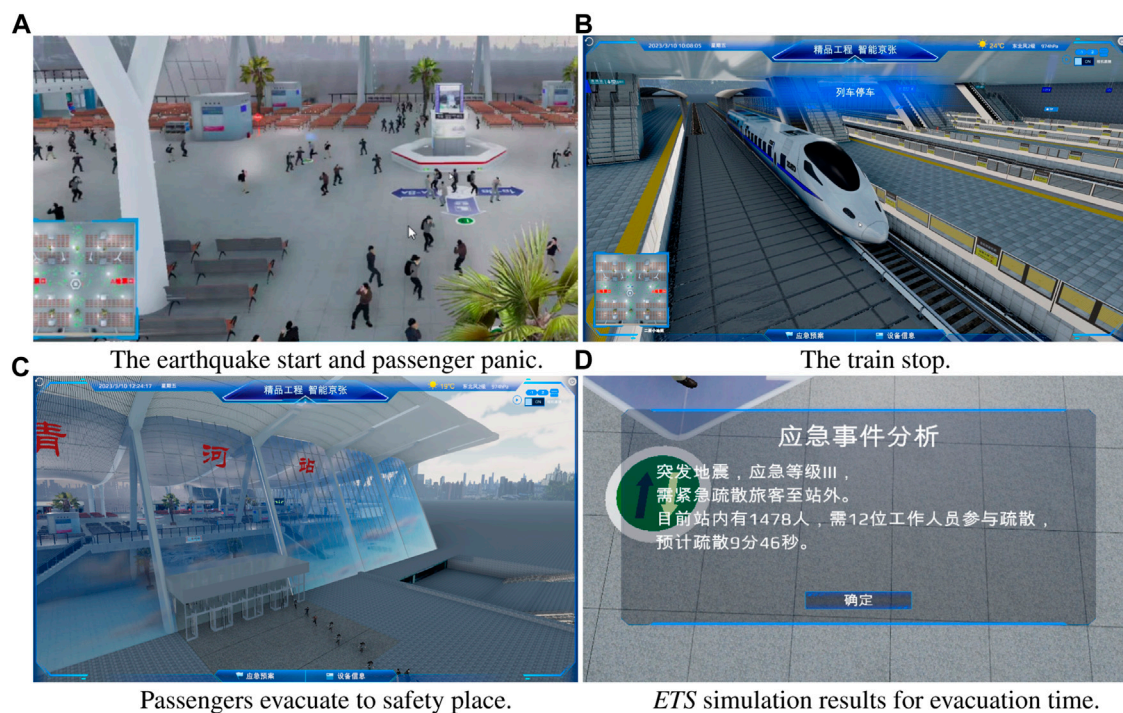
FIGURE 7
Earthquake scenario. **(A)** The earthquake start and passenger panic. **(B)** The train stop. **(C)** Passengers evacuate to safety place. **(D)** ETS simulation results for evacuation time.

(T_3). Subsequently, the station staff contacted the local public security department (P_5), and the public security personnel quickly arrived at the designated location (T_5), while the station staff equipped with mobile screens arrived at the assigned site (P_6) as depicted in Figure 8B.

Simultaneously, additional personnel obtained fire extinguishers to handle the fire (P_6), while a staff member organized the passengers' evacuation according to prescribed procedures (P_9). Public security personnel apprehended the culprits (T_{12}) and restored order to the scene, as shown in Figure 8C. The passengers ensured an orderly evacuation process by complying with staff instructions and the displayed evacuation equipment routes (T_{13}). Lastly, a report on emergency disposal was compiled to document the conclusion of the incident (P_{16}). The simulation process and results are provided by *ETS*. According to Figure 8D, there were 1,255 individuals in the station, requiring the involvement of 4 personnel for guidance, with the entire evacuation process taking 4 min and 51 s.

5.5 Public health event

In the event of a public health emergency at a station, the station personnel will receive notification of a passenger carrying a contagious disease. The alert message will appear on the *ETS* (P_1), prompting the station manager to initiate emergency protocols and release information to staff members (P_3). A staff member will then contact the local epidemic department (T_4), while another staff member locates the infected passenger at the designated location (P_6) as illustrated in Figure 9A, closing the entrance gate to prevent further spread of the disease (P_7) as

illustrated in Figure 9B. The epidemic prevention department will then rescue the passenger (P_8).

To prevent the spread of the disease, individuals who have had no contact with the infected passenger will be screened and evacuated according to staff guidance (T_{13}). The disposal staff will erect an isolation barrier (P_{14}) as illustrated in Figure 9C. Once the situation is fully under control, emergency protocols will be terminated, and an emergency disposal document will be created (P_{16}). The simulation results of public health event are presented in Figure 9D, showing that 933 people were in the station and that four workers were required for the evacuation task, which took 3 min and 55 s.

6 Usability evaluation

In order to validate the efficacy and efficiency of the *ED4D*, we conducted an assessment from three different perspectives: replication of real-world scenarios, simulation of emergency scenarios, and user-friendliness. The replication of real-world scenarios was utilized to gauge the degree of similarity between the virtual entity's and the physical entity's location, structure, quantity, and information data. The simulation of emergency scenarios was employed to assess the functional accuracy of the emergency process, with a focus on verifying whether the emergency treatment digital twin meet actual operational requirements and whether any critical steps are missing or inconsistent with reality. User-friendliness was measured by evaluating the ease of use and acceptability of the digital twin during practical use, including aspects such as ease of operation and the difficulty of task completion.
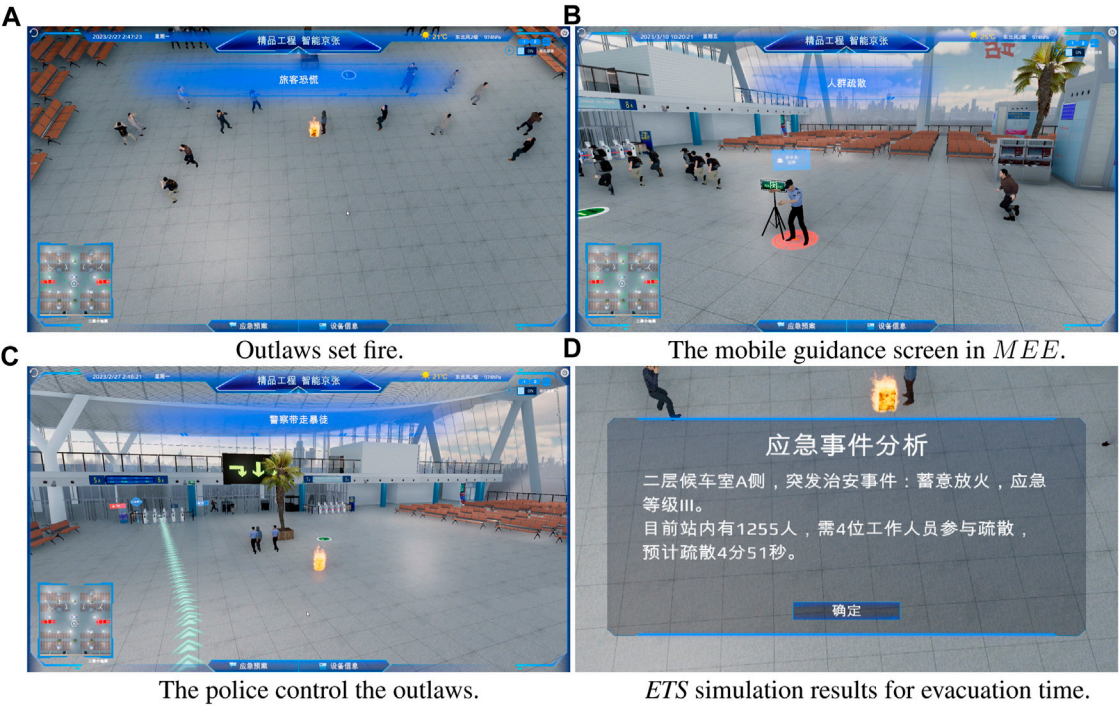
FIGURE 8
Social security incident scenario. **(A)** Outlaws set fire. **(B)** The mobile guidance screen in MEE. **(C)** The police control the outlaws. **(D)** ETS simulation results for evacuation time.
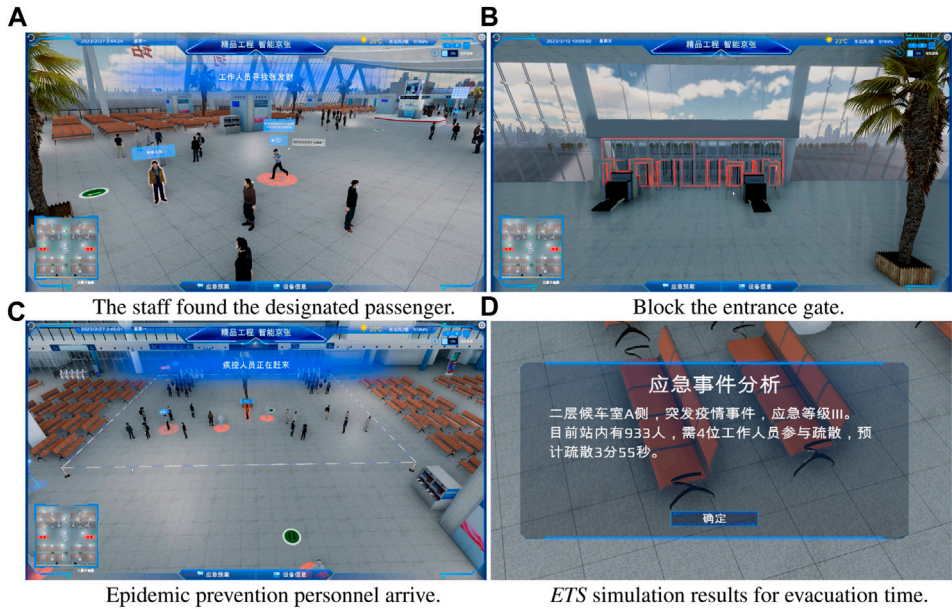
FIGURE 9
Public health event scenario. **(A)** The staff found the designated passenger. **(B)** Block the entrance gate. **(C)** Epidemic prevention personnel arrive. **(D)** ETS simulation results for evacuation time.

6.1 Questionnaire design

The questionnaire was specifically designed to cover the three evaluation aspects. The questionnaire was developed based on the usability metric from user experience framework [44], incorporating the core principles of HEART [45]. The scores were calculated using the SUS [43] based on the search questionnaire. Retaining happiness, engagement, adoption, and task success, the core idea of HEART discarded retention, which was not relevant to the application scenario, as it is intended as a tool in the workplace. After-Scenario Questionnaire [44] was utilized to assess usability, and the results were obtained through questionnaire research and subsequently analyzed.

The questionnaire consists of 12 inquiries, whose specifics are presented in Table 4. The inquiries for reproducing the real-world were designed in terms of station structure, equipment location, and data accuracy. For the simulation of the station emergency scenario, we asked questions about the correctness of the process in five scenarios by using the engagement and task success contents of the core idea of HEART [45]. The being user-friendliness concerns the happiness and adoption of system.

Each inquiry is graded on a scale of "1″" to "5″", wherein "1″" represents strong opposition, "2″" denotes opposition, "3″" represents neutrality, "4″" means agreement, and "5" signifies strong agreement. Questions 1 through 10 are intended to measure positive attitudes, while questions 11 and 12 are for negative attitudes. The SUS score (Brooke, 2013) is determined using Eq. 6.

$$\begin{cases} P_s = \sum (S_i - 1) & i = \{Q_1, Q_2, \ldots, Q_{10}\} \\ N_s = \sum (5 - S_i) & i = \{Q_{11}, Q_{12}\} \\ S = \dfrac{P_s + N_s}{c \cdot 12} \times 100 \end{cases} \quad (6)$$

Where P_s represents the cumulative sum of positive questions by an individual, N_s represents the cumulative sum of negative; S represent an individual's score, 12 represents the total number of questions, c signify a correction factor expressed as a percentage. The scores of each participant are tallied, with the positive question score being the original score minus 1. Meanwhile, the negative question score is obtained by subtracting the original score from 5. The scores for all 12 questions are added up and converted to a percentage scale. For the purpose of our analysis, we shall designate c as 4.

6.2 Experimental evaluations

In our evaluation, a group of 9 individuals participated, among which 6 had previous experience in passenger transportation, while 2 held managerial experience. Notably, 7 of them, accounting for 77.7%, possessed expertise in emergency disposal. The average age of the participants was calculated to be 34.22 years, with an average career length of 10.33 years. The evaluator conducted the trial through an online platform, utilizing a hardware configuration that consisted of a NVIDIA GeForce RTX 3080Ti graphics card.

In Figure 10, one can observe the distribution of scores for each inquiry, where the color red corresponds to a score of "1″", yellow of

"2″", blue of "3″", green of "4″", and purple of "5". For the first question, six individuals, accounting for 66.7%, provided a raw score of 5. For inquiries two to seven, eight individuals gave a raw score of 4 or greater. With respect to inquiries 8 and 10, seven individuals provided a raw score of 4 or greater. As for inquiry 9, all individuals provided a score of 4 or greater. For inquiry 11, which follows a "the smaller, the better" format, four individuals selected a score of 1. Concerning inquiry 12, five individuals selected a score of 2 or less. In regards to inquiries 4, 5, and 7, the present disposal process aligns with our simulations, underscoring the effectiveness of the *ED4D* model that has been developed. The short error bars illustrated in the graph suggest the data's reliability.

The data in Table 5 presents the mean and standard deviation for each question. The assessment of inquiry 1, 2, and 3 pertains to the precision of simulating real-world scenarios. The mean scores for each inquiry are 4.56, 4.33, and 4.22, respectively. Their aggregate mean score is 4.37, with a standard deviation of 0.01. The examination results validate that the exterior static model of *MEE* of *ED4D* agrees well with *SPE*.

The inquiries 4, 5, 6, 7, and 8 serve to assess the accuracy of the simulation of emergency scenarios. The assessment reveals an overall mean score of 3.91, with a standard deviation of 0.11. More specifically, the mean scores for each individual inquiry are 4.22, 4.11, 3.33, 4.00, and 3.89. The results of the evaluation demonstrate that scenarios related to fire disaster, flood, and social security incident accurately represent reality with a high degree of precision. With regards to inquiry 6, which pertains to the disposal process for earthquake emergency, users contend that the magnitude of the earthquake should factor into the process, with earthquake below level 3 not requiring consideration. The existing process aligns with the disposal process for earthquake between levels 3 and 6, but is not applicable for earthquake of level 8 or higher. Lastly, in relation to inquiry 8, the existing process for the disposal of public health event scenario is consistent with the prescribed disposal process for the period spanning 2020 to December 2022. However, it is no longer applicable to the disposal process for the new crown epidemic, as policy adjustments were made in 2023.

The inquiries 9 and 10 are utilized to assess the level of contentment. The overall average score obtained stands at 4.11, with a standard deviation of 0.12. The average rating for each question is established at 4.22 and 4.00 respectively. Inquiries 11 and 12 are employed to assess the level of acceptance. The overall average score obtained is 2.11 (note that a lower score indicates better performance), with a standard deviation of 0. The average rating for each question is established at 1.89 and 2.33 respectively. The evaluation results clearly indicate that the functioning of *ETS* is uncomplicated and user-friendly.

The distribution chart in Figure 11 illustrates the mean scores and number of survey participants in each age group for SUS scores. The evaluation revealed that the 21–30 age group obtained an average score of 77.08, while the 31–40 age group achieved the highest mean score of 82.29. In contrast, the 41–50 age group scored 68.75, and the 51–60 age group scored 58.33. The error bar is relatively short. Furthermore, the nine evaluators who participated in the survey received a mean score of 76.38 points, surpassing the minimum score of 74.1 points, which earned the system a rating of *acceptable* and the evaluators a grade of *good* [51,52].

TABLE 4 Questions.

Questions	Content	Class	HEART	Dimensions
Q1	The exterior Static model of Virtual Entity is consistent with the actual of Passenger Station physical entity building structure	station structure	-	Reproducing the real-world
Q2	The location of the equipment in the scene of Virtual Entity is consistent with the actual of Passenger Station physical entity	equipment location	-	Reproducing the real-world
Q3	The collected data is consistent with the actual of Passenger Station physical entity	data accuracy	-	Reproducing the real-world
Q4	With the *ED4D* model for passenger station, the business process is consistent with the disposal situation of fire disaster, and the function is correct	fire disaster	Task Success and Engagement	Simulating the station emergency scenario
Q5	With the *ED4D* model for passenger station, the business process is consistent with the disposal situation of flood natural disaster, and the function is correct	flood	Task Success and Engagement	Simulating the station emergency scenario
Q6	With the *ED4D* model for passenger station, the business process conforms to the disposal situation of earthquake disaster, and the function is correct	earthquake	Task Success and Engagement	Simulating the station emergency scenario
Q7	With the *ED4D* model for passenger station, the business process is in line with the disposal of social security and is functionally correct	Social security incident	Task Success and Engagement	Simulating the station emergency scenario
Q8	With the *ED4D* model of emergency disposal for passenger station, the business process conforms to the disposal of public health infectious disease, and the function is correct	Public health event	Task Success and Engagement	Simulating the station emergency scenario
Q9	This system is easy to use	-	Happiness	Being user-friendliness
Q10	This system is better integrated in terms of functionality	-	Happiness	Being user-friendliness
Q11	This system takes a lot of time to complete its tasks	-	Adoption	Being user-friendliness
Q12	This system requires additional knowledge to use	-	Adoption	Being user-friendliness

FIGURE 10
Stacked graph of percentage scores for each question.
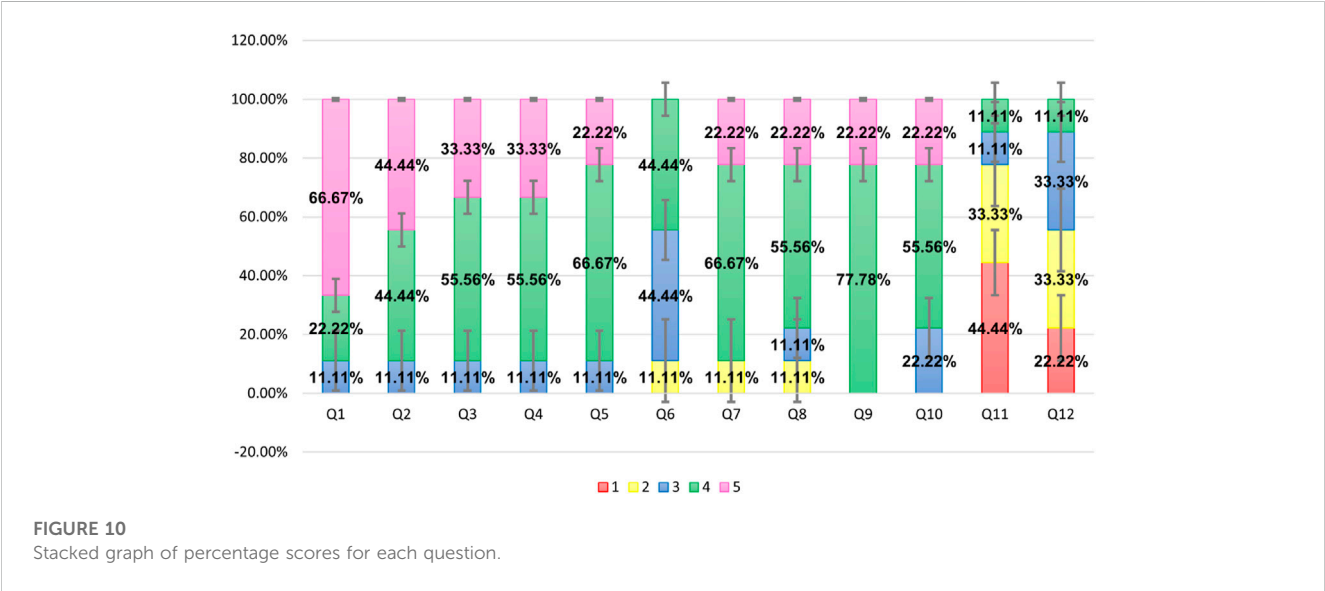
TABLE 5 Mean and standard deviation.

	Q1 ↑	Q2 ↑	Q3 ↑	Q4 ↑	Q5 ↑	Q6 ↑	Q7 ↑	Q8 ↑	Q9 ↑	Q10 ↑	Q11 ↓	Q12 ↓
Mean	4.56	4.33	4.22	4.22	4.11	3.33	4.00	3.89	4.22	4.00	1.89	2.33
Standard Deviation	0.66	0.66	0.64	0.64	0.60	0.81	0.83	0.89	0.46	0.70	0.98	0.98

FIGURE 11
Distribution of SUS scores for each age group.
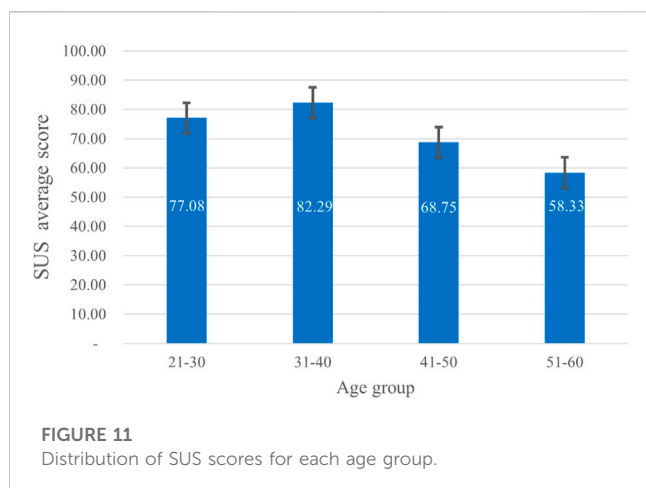
7 Conclusion

To facilitate the training of station personnel in emergency protocols, this research paper has implemented a digital-twin-based approach to create an *ED4D* model for emergency simulation. The model entails the extraction of various emergency scenarios, which are then digitized using GSPN. This process replicates the actual conditions of the physical world within a virtual environment and facilitates the exchange of information between the two spaces. The successful application of this model in constructing multiple emergency scenarios at Qinghe Station, including simulations of fire, natural disasters such as flood and earthquake, social security incident, and public health event, has enabled the effective simulation of emergency procedures and the execution of emergency drills with greater efficiency. In comparison to tabletop exercises and live drills, this solution offers an immersive experience while conserving resources, making it easily adaptable to various potential hazards.

Furthermore, we have evaluated the application in three key aspects: the replication of the real world, the simulation of the station emergency process, and user-friendliness. Based on the System Usability Scale (SUS) evaluation, the overall conclusion was graded as *good*, indicating that the application is suitable for emergency response in passenger stations. This model allows for the validation of emergency plans and the identification of vulnerabilities in the response process. It assists personnel in adjusting resource allocation and strengthening weak areas, thereby enhancing the efficiency of emergency response strategies. Additionally, it establishes a solid foundation for the digitalization and informatization of the emergency response modeling process. Furthermore, the *ED4D* model enables immersive drills for various disaster events, thereby improving the emergency response skills of staff members.

The developed *ED4D* model demonstrates significant applicability as it can encompass the entire process of various disasters, including pre-event preparation, mid-event management, and post-event recovery. Moreover, the design of the GSPN model allows for easy adaptation to other railway passenger stations, highlighting its wide-ranging applicability. Exterior Static models, such as personnel, trains, and equipment, can also be adjusted and

reused in different scenarios, further enhancing the practical value and potential for widespread adoption of this research. It is worth noting that the system developed in this study is currently only operational on computers. Moving forward, our plan is to utilize mixed reality devices to observe staff behavior, thereby achieving an emergency twin service that accurately reflects the composition of real multi-class passengers.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

XW: Methodology, Software, Validation, Writing–original draft, Writing–review and editing. WB: Software, Validation, Writing–original draft, Writing–review and editing. YS: Writing–original draft, Writing–review and editing. GY: Software, Validation, Writing–original draft. CL: Software, Validation, Writing–original draft. XL: Software, Validation, Writing–original draft. KP: Software, Writing–original draft. JL: Validation, Writing–original draft.

Conflict of interest

Authors WB, GY, CL, XL, KP, and JL were employed by Institute of Computing Technology, China Academy of Railway Sciences Corporation Limited.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Lei L, Fuzhang W. Railway incident and emergency decision-making research. *J Inst Disaster Prev* (2012) 14:58–63. doi:10.3969/j.issn.1673-8047.2012.03.012

2. Zhuguo Y. Analysis and optimization strategies of emergency disposal for high speed railway. *Railway Transport And Economy* (2022) 44:90–5. doi:10.16668/j.cnki.issn.1003-1421.2022.07.14

3. Xiaoqing C, Liming J, Yong Q. Research on railway emergency management. *J China Railway Soc* (2012) 34:7–13. doi:10.3969/j.issn.1001-8360.2012.03.002

4. Lipare S, Bhavathankar P. Railway emergency detection and response system using iot. In: 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT); July 1-3, 2020; Kharagpur, India (2020). p. 1–7.

5. Tianyun S, Chunjia Z. Overall design and evaluation of intelligent railway passenger station system. *Railway Comp Appl* (2018) 27:9–16.

6. Jianxi X, Zhao T, Xiaolin Y, Yinyu N, Zong M, Xihui W, et al. A vr-based the emergency rescue training system of railway accident. *Entertainment Comput* (2018) 27: 23–31. doi:10.1016/j.entcom.2018.03.002

7. Xingyu L. *Research on emergency desktop drill system based on unity3d* (2022). p. 3–13. Available at: https://cdmd.cnki.com.cn/Article/CDMD-10619-1022524942.htm.

8. Yuefeng Z, Jianmin H, Jianhua H. Modeling and analysis of emergency management workflow based on gspn. *Stat Decis* (2012) 4:44–7.

9. Peng W, Qiang H, Guozhu L. Xml oriented workflow modeling based on restricted digraph. *Nat Sci Edition* (2007) 28:240–3. doi:10.3969/j.issn.1672-6987.2007.03.015

10. Yuren S, Ying D. Comparison of eepc and petri net in modeling of business process improvement. *J Chongqing Univ Technology (Natural Sci)* (2013) 25:87–91. doi:10.3969/j.issn.1674-8425-B.2011.07.011

11. Molloy M. Performance analysis using stochastic petri nets. *IEEE Trans Comput* (1982) 31:913–7. doi:10.1109/tc.1982.1676110

12. Yi X, Renzhong T, Yaping M. Storage and retrieval of business process simulation models based on idef3. *Syst Eng Theor Pract* (2005) 25:69–75. doi:10.3321/j.issn:1000-6788.2005.12.011

13. Qing L, Wei L, Mengzhen G, Hongwei S. Modeling and analysis of emergency decision making based on logical probability game petri net. *Comp Sci* (2022) 49:294–301. doi:10.11896/jsjkx.210300224

14. Xunqing W, Yongjian L, Huali S. A scenario evolution model for mass emergencies based on stochastic petri nets. *Manag Rev* (2014) 26:53–62+116. doi:10. 14120/j.cnki.cn11-5057/f.2014.08.046

15. Chao Q, Lanjun L. Analysis of flood disaster case based on event chain and generalized stochastic petri nets. *J Wut (Information Manag Engineering)* (2017) 39:130–4. doi:10.3963/j.issn.2095-3852.2017.02.002

16. QiuQing L, XueLin W. Evaluation of meteorological disaster emergency capability based on fuzzy petri net. *J Henan Polytechnic Univ (Natural Science)* (2018) 37:32–7. doi:10.16186/j.cnki.1673-9787.2018.03.5

17. Runchao M, Xiangwei L, Yixuan C, Haiyan D. Research on customs clearance mode of China-laos railway based on petri-Markov chain. *China Transportation Rev* (2023) 02: 1–9. Available at: https://kns.cnki.net/kcms/detail/11.1197.u.20220512.0948.002.html.

18. Zhibo Z, Qiyuan P, Lingshuan Z, Denghui L, Bin L. Research on structural analysis thod of high speed raitway emergency dispatching cottaborative disposal process. *Railway Transport and Economy* (2021) 43:120–7. doi:10.16668/j.cnki.issn.1003—1421.2021.04.19

19. Zixian Y, Junhua C. A simulation study of emergency evacuation of railway passenger station under complex passenger flow condition. *Railway Transport and Economy* (2020) 42:66–71. doi:10.16668/j.cnki.issn.1003-1421.2020.11.12

20. Zelong L. In: thesis MS, editor. Beijing: Beijing Jiaotong University (2021). p. 26–43.*Research on passenger safety evaluation and risk control of high speed railway station based on fuzzy petri net and mop*

21. Fei T, Weiran L, Meng Z. Digital twins five-dimensional model and its application in ten fields. *Comp Integrated Manufacturing Syst* (2019) 25:1–18. doi:10.13196/j.cims.2019.01.001

22. Singh M, Srivastava R, Fuenmayor E, Kuts V, Qiao Y, Murray N, et al. Applications of digital twin across industries: a review. *Appl Sci* (2022) 12:5727. doi:10.3390/app12115727

23. Yicheng S, Yuqian L, Jinsong B, Fei T. Prognostics and health management via long short-term digital twins. *J Manufacturing Syst* (2023) 68:560–75. doi:10.1016/j.jmsy.2023.05.023

24. Baicun W, Huiying Z, Xingyu L, Geng Y, Pai Z, Ci S, et al. Human digital twin in the context of industry 5.0. *Robotics and Computer-Integrated Manufacturing* (2024) 85: 102626. doi:10.1016/j.rcim.2023.102626

25. Luther W, Baloian N, Biella D, Sacher D. Digital twins and enabling technologies in museums and cultural heritage: an overview. *Sensors* (2023) 23:1583. doi:10.3390/s23031583

26. Maria P, Valentina Y, Ekaterina R, Natalia K, Peter T, Tatiana D. Digital twin concept: healthcare, education, research. *J Pathol Inform* (2023) 14:100313. doi:10.1016/j.jpi.2023.100313

27. Falkowski P, Osiak T, Wilk J, Prokopiuk N, Leczkowski B, Pilat Z, et al. Study on the applicability of digital twins for home remote motor rehabilitation. *Sensors* (2023) 23:911. doi:10.3390/s23020911

28. Rajanikanth A, Swapna M, Kaladevi A, Rohith C, Sandhya M, Bhat CR. The novel emergency hospital services for patients using digital twins. *Microprocessors and Microsystems* (2023) 98:104794. doi:10.1016/j.micpro.2023.104794

29. Yilong H, Yinbo L, Yongkui L, Bin Y, Lingyan C. Digital twinning for smart hospital operations: framework and proof of concept. *Tech Soc* (2023) 74:102317. doi:10.1016/j.techsoc.2023.102317

30. Valerian Vanessa T, Joseph Handibry Mbatu T, Fonbeyin Henry A. Technologies for digital twin applications in construction. *Automation in Construction* (2023) 152: 104931. doi:10.1016/j.autcon.2023.104931

31. Victor Adetunji A, Robert Christian M, Yihai F. Digital twin technology for thermal comfort and energy efficiency in buildings: a state-of-the-art and future directions. *Energ Built Environ* (2023). doi:10.1016/j.enbenv.2023.05.004

32. Liu J, Li C, Bai J, Luo Y, Lv H, Lv Z. Security in iot-enabled digital twins of maritime transportation systems. *IEEE Trans Intell Transportation Syst* (2021) 1–9. doi:10.1109/tits.2021.3122566

33. Vladimir A, Alexander S, Andrey B, Aleksey R. Utilizing digital twin for maintaining safe working environment among railway track tamping brigade. *Transportation Res Proced* (2022) 61:600–8. XII International Conference on Transport Infrastructure: Territory Development and Sustainability (TITDS-XII). doi:10.1016/j.trpro.2022.01.097

34. Sakdirat K, Qiang L. Digital twin aided sustainability-based lifecycle management for railway turnout systems. *J Clean Prod* (2019) 228:1537–51. doi:10.1016/j.jclepro.2019.04.156

35. Shiyang Z, Alexander M, Yuxi X, Manfred G. A machine-learning-based surrogate modeling methodology for submodel integration in the holistic railway digital twin platform. *Proced CIRP* (2023) 119:345–50. doi:10.1016/j.procir.2023.02.141

36. Shiyang Z, Stefan D, Rebecca N, Rainer R, Ozan K, Martin K, et al. A conceptual model-based digital twins platform for holistic large-scale railway infrastructure systems. *Proced CIRP* (2022) 109:362–7. doi:10.1016/j.procir.2022.05.263

37. Baek S, Gil H, Kim Y. Vr based job training system using tangible interactions. *Sensors* (2021) 21:6794. doi:10.3390/s21206794

38. Kwon J, Lee I, Park H, Kim S. Design of augmented reality training content for railway vehicle maintenance focusing on the axle-mounted disc brake system. *Appl Sci* (2021) 11:9090. doi:10.3390/app11199090

39. Dunwen L, Haoran J, Jian Yinghua J, Fengkai Q, Hui Y. Construction and research of emergency training system for tunnel fire based on virtual reality technology. *J Saf Sci Tech* (2019) 15:131–7.

40. Bo Z, Jun S, Lei S, Hongji X, Cheng Y. Design and realization of ship virtual fire training system based on hmd. *J Syst Simulation* (2019) 31:43–52. doi:10.16182/j.issn1004731x.joss.17-0182

41. Ping Z, Ruijiang P, Ziding C. Research on safety quality improvement of railway staff based on 3d emergency drill. *China Saf Sci J* (2018) 28:138–42.

42. Pei S, Weifang D, Wei Y. Research on the construction of a practical training system for railroad crew emergency drills based on vr technology. *Sci Tech Vis* (2018) 05: 187–8.

43. Brooke J. Sus: a retrospective. *Usability Stud* (2013) 08:29–40. Available at: https://www.researchgate.net/publication/285811057_SUS_a_retrospective.

44. Lewis RJ. An after-scenario questionnaire for usability studies: psychometric evaluation over three trials. *ACM SIGCHI Bull* (1991) 23:79. doi:10.1145/126729.1056077

45. Finstad K. The usability metric for user experience. *Interacting Comput* (2010) 22: 323–7. doi:10.1016/j.intcom.2010.04.004

46. cleverism. *Complete guide to's heart framework for measuring the quality of ux* (2023).

47. Shenping H, Shengyuan L. Simulation modeling on emergency response process for lng-fueled vessel via gspn-mc method. *China Saf Sci J* (2021) 31:174–81. doi:10.16265/j.cnki.issn1003-3033.2021.10.024

48. Jianzhong C, Liangyou C. Extended event-process chain (eepc) and it's application in brp. *Syst Eng* (2000) 01:42–8.

49. Yanxiang H, Hua S. Rules and implementation of converting stochastic petri net model to Markov chain. *J Front Comp Sci Technolog* (2013) 07:55–62. doi:10.3778/j.issn.1673-9418.1209025

50. Jialin Y, Siyuan Z. Research on the calculation model of the maximum assembling passengers in large high-speed railway station. *J Railway Sci Eng* (2019) 16:34–41. doi:10.19713/j.cnki.43-1423/u.2019.01.005

51. Jiajun H, Ruihua X, Ling H, Zhaohong H. Simulation model of evacuation process in metro station under dynamic and uncertain environment. *J Transportation Syst Eng Inf Tech* (2011) 18:164–70. doi:10.16097/j.cnki.1009-6744.2018.02.025

52. Sauro J, Lewis J. When designing usability questionnaires, does it hurt to be positive? In: In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems; May 07-12, 2011; Vancouver, BC, Canada (2011). doi:10.1145/1978942.1979266

Evolutionary game analysis of stakeholders' decision-making behavior in agricultural data supply chain

Heyang Zhao[1]* and Jian Yang[2]

[1]School of Computer and Information Technology School of Big Data, Shanxi University, Taiyuan, China,
[2]School of Information, Shanxi University of Finance and Economics, Taiyuan, China

The significance of agricultural information sharing in fostering agricultural development cannot be overstated. This practice plays a pivotal role in disseminating cutting-edge agricultural technologies, cultivation methods, and pest control strategies, empowering farmers with valuable knowledge to enhance crop yield and quality. Moreover, it aligns with government objectives of resource sharing and addressing gaps, contributing to the advancement of agricultural modernization and the development of the industry chain. Despite its inherent benefits, the practical implementation of agricultural information sharing faces challenges. Stakeholders engaged in information sharing often prioritize individual benefits, potentially leading to a decline in agricultural information quality and the inefficient use of experimental resources. To confront this issue, the present research establishes a three-party evolutionary game model comprising an agricultural product data sharing platform, agricultural data providers, and agricultural data consumers. Leveraging dynamic system theory, the model analyzes the evolutionary stable strategies of stakeholders and investigates the critical factors influencing the strategic choices of these three parties. Experimental findings underscore the pivotal role of participants' initial strategies, regulatory intensity, reward and punishment mechanisms, and information feedback in shaping stakeholder decision-making behavior. Implementation of measures such as heightened scrutiny of information on the sharing platform and fostering consumer trust in data emerges as imperative for enhancing system stability. These actions are essential for constructing an efficient and reliable information-sharing ecosystem, thereby facilitating the sustainable development of modern agriculture.

KEYWORDS

agricultural data sharing, decision-making behavior, evolutionary game, influencing factors, stability analysis

1 Introduction

In contemporary society, the progress of agricultural information technology has emerged as a crucial area of development in the agricultural realm, spearheading the transformation of agricultural production, management, and decision-making processes. The continuous advancement and application of information technology have provided agriculture with novel prospects for development, allowing for more efficient, intelligent, and sustainable agricultural production [1]. By leveraging advanced sensing technology, the

Internet of Things (IoT), big data analysis, artificial intelligence (AI), and other approaches, agricultural informatization facilitates real-time monitoring and analysis of diverse agricultural data, encompassing factors such as soil conditions, weather patterns, crops, poultry, and livestock. This enables precise decision-making concerning key issues in agricultural production while simultaneously enhancing production efficiency and improving yield quality [2]. Nonetheless, the amount of data generated by a single institution or organization remains inherently limited. Consequently, establishing an efficient sharing mechanism between data providers and consumers has become an urgent priority. In a bid to addre ss the challenge of "data islands," agricultural product information sharing platforms have emerged. These platforms enable the collection, integration, sharing, and analysis of dispersed agricultural information sourced from individual companies or organizational systems, thereby effectively mitigating the issues posed by the inconvenient exchange of agricultural information.

Furthermore, with the ongoing digitalization and informatization transformation of the agricultural industry, there is a growing demand for diverse and reliable data sources, along with increased requirements for data quality control [3]. Consequently, establishing a data sharing platform with integrated data governance capabilities becomes paramount. Data sharing not only reduces redundancy in data processing but also lowers the overall cost of agricultural research, fostering advancements in agricultural scientific research [4].

However, the mere construction of an agricultural product data sharing platform does not guarantee high-quality data sharing. From an economic perspective, agricultural data sharing is an ongoing process involving multiple stakeholders, each driven by their own interests to maximize [5]. For instance, agricultural data providers may choose to provide incomplete or inadequate information to minimize costs and increase income. Conversely, data consumers may opt to verify the accuracy of the provided information to avoid losses resulting from erroneous data. This dynamic can lead to the inefficient use of scientific research resources and funding, ultimately hindering the progress of agricultural research endeavors.

Evolutionary game theory is a theory that combines game theory analysis with dynamic evolutionary process analysis, placing a strong emphasis on achieving dynamic equilibrium. This theory originally emerged from the field of biological evolution and has successfully explained certain phenomena observed in biological evolution processes [6]. In contemporary economics, evolutionary game theory is widely utilized to analyze the factors influencing the development of social habits, norms, institutions, or systems, and to explain the processes through which they form, yielding significant insights Liu et al. [7].

In the agricultural sector, due to pronounced information asymmetry within the agricultural supply chain and the challenges associated with monitoring the quality of agricultural products, certain stakeholders are motivated by personal interests, potentially leading to harm to the interests of others and resulting in conflicts of interest within the supply chain [8]. In the context of this three-party game, complex interactions arise among the agricultural product information sharing platform, information providers, and consumers.

As an information intermediary, the agricultural product information sharing platform needs to maintain a balance between information providers and consumers [9]. The decisions made by these parties directly affect the operation and reputation of the platform. Information providers can choose to provide data rigorously or perfunctorily, while consumers make choices between trusting or distrusting the data. These decisions directly impact the data quality and credibility of the platform. Therefore, there is a dynamic evolutionary process where the decisions of the platform, information providers, and consumers mutually influence each other. The audit strategy of the platform, data provision strategy of information providers, and consumer attitudes towards data trust are constantly evolving. This dynamic equilibrium process is influenced by government regulation and consumer feedback, thereby affecting decision-making behaviors of all parties involved.

In this study, assumptions are made regarding the decision-making behavior of stakeholders engaged in the sharing of agricultural data. Subsequently, a tripartite evolutionary game model is developed, encompassing an information sharing platform and involving information providers and consumers. The decision-making behavior of each party is then studied under various influencing factors, and system equilibrium points are explored through the construction of replicator dynamic equations. Finally, strategies of participants are dynamically adjusted to achieve an optimal balance of interests through numerical simulation experiments.

The remaining parts of this article are structured as follows. Section 2 provides a literature review on agricultural data sharing. Section 3 presents a tripartite evolutionary game model for agricultural data sharing. In section 4, using replicator dynamics equations, the evolution patterns and trends of stakeholders' interests are analyzed. The Jacobian matrix is employed to analyze the equilibrium points of the system and explore their stability. In Section 5, numerical simulation methods are used for analysis. Finally, relevant strategies and recommendations are summarized.

2 Related work

2.1 Agricultural data sharing

With the rapid development of information technology, the process of informatization in the agricultural field is accelerating. Agricultural informatization has not only changed the way of agricultural production, management, and decision-making but also has had a profound impact on the operation of the entire agricultural supply chain [10]. In order to better utilize agricultural data to promote agricultural production and research, the industry is studying agricultural data sharing models and how to achieve the optimal balance of interests between data providers, consumers, and sharing platforms.

[11] conducted a comprehensive review on the current status of open sharing of agricultural big data in China. Although the basic environment for data sharing has been continuously optimized, there are still several challenges hindering its progress. These challenges include constraints within the system, lack of standardized technology, and outdated laws and regulations.

Consequently, [12] proposes that challenges such as low data volume, poor readability, and insufficient update rates still exist in the open sharing of agricultural big data. In order to address these challenges and promote the sharing of agricultural data, various measures are suggested, including top-level design, technological advancements, content standards, sharing mechanisms, and legal regulations. These measures aim to accelerate the open sharing and integration of agricultural data, ultimately enhancing the governance capacity of modern agriculture. [13] based on analyzing the problems of existing agricultural scientific data sharing, take problem analysis as the guide and discuss the applicability of consortium chains and agricultural scientific data management from three dimensions: traceability, shareability, and security. The core of smart agriculture is big data, but the current big data management model still has certain risks. In response to this issue, [14] propose the use of blockchain technology to better achieve the sharing of agricultural data. They analyze the advantages of blockchain-based big data sharing and improve the regional independence and information independence of major agricultural information websites, as well as improve the sharing methods.

[15] summarizes the concepts of data, scientific data, and scientific data systems, and through the research and extraction of information system success models and other theories, constructs a sustainable model for the use of agricultural scientific data sharing platforms. [16] indicates that this is aimed at providing data support for improving agricultural scientific data sharing platforms and guiding practical activities in agricultural scientific data.

Agricultural data possesses distinct characteristics, including extensive geographical distribution, substantial data volume, intricate structure, diverse types, and challenging acquisition. Consequently, the utilization of big data technology for the acquisition, storage, processing, and analysis of agricultural big data has emerged as a significant challenge in agricultural informatization. Based on this, [17] have studied the application system architecture and platform construction of agricultural big data, which has guidance significance for the processing of various types of agricultural data; at the same time, it has practical significance for decision-making in the entire process of agriculture. In addition, [18] explores the necessity of improving the management and utilization of agricultural information resources and internet-based agricultural information resources from both theoretical and practical aspects. Corresponding strategies are proposed for effective integration and utilization.

2.2 Evolutionary game

Evolutionary game theory, integrating principles from game theory and genetic evolution, serves as a mathematical model to explore the dynamic evolution of individual behaviors within populations. This methodology extensively applies to decipher strategic interactions among participants, offering insights into the cooperative and competitive relationships that evolve over time. Researchers, through the lens of evolutionary game theory, simulate strategic decision-making processes among various stakeholders in supply chain management. This application enhances supply chain efficiency, stability, and resilience, providing decision-makers with valuable guidance amid market dynamics and competitive pressures.

Consequently, evolutionary game theory stands as a robust tool and framework for studying and optimizing supply chain management in scholarly research [19].

In an evolutionary game model with repeated interactions, [20] find that conditional cooperation is effective in promoting collaboration, especially under low-risk conditions. The risk threshold is crucial, as higher risks tend to enhance cooperation. Importantly, a timely transition from a defective to a cooperative strategy by conditional cooperators proves beneficial for sustaining high-level cooperation when the risk of failing to achieve collective goals exceeds a certain threshold.And by investigating coevolutionary dynamics between population state and reward systems, [21] find that interior stable coexistence state can emerge in our model.

While translating the original sentence into an academic style, I meticulously rephrased the text to augment conciseness, clarity, and the overall fluidity of the paragraph, demonstrating an assiduous commitment to linguistic refinement. The adjustments ensure that the content remains faithful to the original meaning but improves the quality of expression, [22] proposed a tripartite evolutionary game model to examine the interactions among crowd workers, crowd platforms, and task requesters in this study. Their focus was on analyzing the evolutionary stable strategies and trends exhibited by the different participants. The model considers the strategic decision-making of various stakeholders in the system and highlights the dynamic characteristics of the crowd working ecosystem. By examining the behaviors and choices of crowd workers, crowd platforms, and task requesters, the researchers acquire comprehensive insights into the emergence of long-term stability and patterns in this intricate environment. The analysis contributes to a more profound understanding of the interactions, adaptations, and evolution of these three key players over time.

Based on this, [23] used the evolution game theory to determine the optimal strategy energy system of multiple participants in the region and established a three-party game reciprocity model. The probability of strategy selection for the three participants under equilibrium and non-equilibrium strategies was analyzed. [24] addressed the insufficient stability research on three-party evolutionary games. They used a generalized three-dimensional dynamical system constructed by the replication dynamic equations to discuss the evolutionary trend of single-group strategies. Then, based on the Lyapunov stability theory, they analyzed the asymptotic stability of the system and, combined with the evolutionary trend of single-group strategies, discovered that all types of mixed strategy Nash equilibrium were saddle points. They also proved the zero eigenvalue non-ESS theorem and the ESS non-coexistence theorem, providing reference and inspiration for further research on evolutionary games.

In online knowledge communities, the monetization of knowledge has become a key focus of business models. [25] conducted an in-depth exploration of this topic. They established a three-party evolutionary game model to study different evolutionary stable strategies, evolutionary trends, and factors affecting these evolutionary paths. This research provides important decision support for the operation of online knowledge communities and proposes new approaches to solving conflicts of interest in knowledge sharing.

In the era of big data, the sharing of medical data has become an urgent need to improve the quality and efficiency of healthcare services. [26] developed a three-party evolutionary game model tailored to medical sharing scenarios, highlighting the significance of regulation
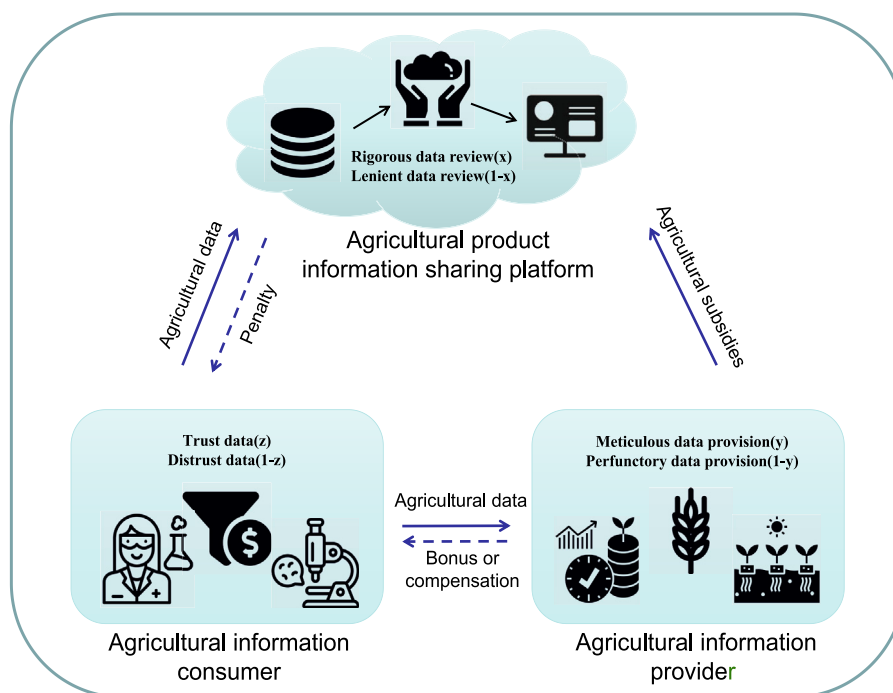
FIGURE 1
Agricultural big data sharing workflow.

and governance of data flow quality in addition to ensuring data security. By achieving stability within the medical data sharing system, this research provides significant insights and recommendations for enhancing the management of medical data sharing, providing valuable guidance for stakeholders' informed decision-making. Furthermore, this study also extends to mobile crowdsourcing, contributing to its efficiency and fairness.

[27] conducted an in-depth exploration of the behavioral strategies employed by crowdsourcing workers, crowdsourcing platforms, and task requesters. They achieved this by establishing a three-party evolutionary game model, enabling them to analyze the impact and constraints of these behavioral strategies on each other. This provided valuable insights into comprehending the dynamics of the mobile crowdsourcing market. Through simulation experiments, they verified the stability of equilibrium points within the three-party game system, emphasizing the interactions and mutual influences among different participants. The overarching objective was to promote the healthy development of the mobile crowdsourcing market and maximize social welfare.

The literature emphasizes the significant role of data sharing in modern agricultural informatization. These studies highlight challenges such as insufficient data volume, poor quality, and outdated regulations that impede its effectiveness. To overcome these challenges, studies suggest using big data technology and blockchain to enhance security and autonomy. Furthermore, the application of evolutionary game theory in domains like crowdsourcing and medical data sharing analyzes strategies and emphasizes stable approaches. This article aims to develop a multi-agent decision-making mechanism based on evolutionary game theory to optimize interactions between data providers, consumers, and platforms in order to improve governance and promote agricultural informatization.

3 Basic assumptions and model construction

3.1 Game relationship among three participants

As shown in Figure 1, in the agricultural product information sharing ecosystem, agricultural information providers provide data to the agricultural product information sharing platform. When providers are found to provide data with a lack of rigor, they will be fined by the platform. The rigor or lack thereof in providing data directly affects the data quality and stability of the ecosystem. At the same time, the agricultural product information sharing platform will review the collected information and then provide it to agricultural information consumers. If the data is reliable, agricultural information consumers will provide subsidies to the providers. However, if the data is unreliable, they will seek compensation from the platform. This complex ecosystem requires all parties to carefully balance their interests to ensure the sustainable development of information sharing.

3.2 Model assumptions

Assumption 1. The strategy of the agricultural product information sharing platform is represented as $\alpha = (\alpha_1, \alpha_2) =$ (rigorous data review, lenient data review). It selects α_1 with a probability of x and α_2 with a probability of $1 - x$, where x ranges from 0 to 1. The strategy space of the agricultural information provider is denoted as $\beta = (\beta_1, \beta_2) =$ (meticulous data provision, perfunctory data provision). It opts for β_1 with a probability of y and β_2 with a probability of $1 - y$, where y ranges

TABLE 1 Main symbols used in the paper.

Symbol	Description
C_a	Agricultural Platform Data Review Costs
C_b	The Cost of Accurate Agricultural Data Collection
C_d	The Cost of Distrust in Agricultural Data verification
P_a	Reputation Loss from Erroneous Agricultural Data
P_b	Reliable Agricultural Data's Reputation Gains
P_c	Wasted Resources from Superficial Agri Data Trust
R_a	Provider Benefits from Agricultural Data Sharing
R_b	Provider benefits from superficial agri data sharing
T	Platform Fines for Perfunctory Data in Audits
F	The Cost of Reacquiring Accurate Data Post-Audit
B	Consumer Claims for Mistrusted Agricultural Data
S	Provider Subsidies for Verified Accurate Data
H	Consumer Bonuses for Reliable Platform Data

from 0 to 1.The strategy space of the agricultural information consumer is expressed as $\gamma = (\gamma_1, \gamma_2)$ = (trust data, distrust data). It selects γ_1 with a probability of z and γ_2 with a probability of $1 - z$, where z ranges from 0 to 1.

Assumption 2. When the agricultural information provider chooses meticulous data provision, the profit is R_a, and the cost of using professional equipment to collect data rigorously is C_b. When the agricultural information provider chooses perfunctory data provision, the profit is R_b, where $R_a < R_b$.

Assumption 3. When the agricultural information sharing platform chooses rigorous data review, the required manpower and material costs are C_a. When the agricultural information sharing platform chooses lenient data review, the perfunctory data provision by the agricultural information provider leads to a decrease in the average quality of agricultural data. In this case, the loss caused by the reputation loss of the agricultural product information sharing platform is P_a. When the agricultural information sharing platform rigorously reviews the data and discovers that it has been provided perfunctorily, it imposes a fine of T on the agricultural information provider and incurs a cost of F to acquire accurate data. The reputation benefit obtained by the agricultural information sharing platform from having authentic and reliable data is P_b, and it also receives a bonus H from the agricultural information consumers.

Assumption 4. When the agricultural information consumer distrusts the data, the cost of verification is C_d. If perfunctory data provision is discovered, the consumer can file a claim of B against the agricultural information sharing platform. If the agricultural information provider is found to voluntarily provide meticulous data, they receive an agricultural subsidy S. When the agricultural information consumer trusts the data but it is provided perfunctorily, the waste of experimental resources resulting from incorrect data is P_c.

Based on the aforementioned assumptions, to enhance the clarity of the game model, Table 1 presents the relevant

behavioral parameters essential for the three-party evolutionary game model, along with their respective meanings.

3.3 Model construction

According to the above assumptions, the three-party game payoff matrix is established for the agricultural information sharing platform, the agricultural information provider, and the agricultural information consumer, as shown in Table 2.

4 Model analysis

4.1 Analysis of the stability of the strategy of agricultural product information sharing platform

The expected benefits and average expected benefits (U_{11}, U_{12}, U_1) for strict or lenient data review on the agricultural product information sharing platform are as follows:

$$\begin{cases} U_{11} &= -C_a + Fy - F + Hz + P_b - Ty + T \\ U_{12} &= -Byz + By + Bz - B + Hyz + P_a y \\ &\quad -P_a + P_b y \\ U_1 &= xU_{11} + (1-x)U_{12} \end{cases} \quad (1)$$

The equation for the replication dynamics of strategy selection in the agricultural product information sharing platform is:

$$\begin{aligned} F(x) &= x(U_{11} - U_1) \\ &= x(1-x)((B-H)yz - (B - F + P_a + P_b + T) \\ &\quad y - (B-H)z + B - C_a - F + P_a + P_b + T) \end{aligned} \quad (2)$$

The first-order derivative of x and the set $G(y)$ are respectively:

$$\begin{aligned} \frac{dF(x)}{dx} &= (2x-1)((B-H)yz - (B - F + P_a + P_b, +T) \\ &\quad y - (B-H)z + B - C_a - F + P_a + P_b + T) \end{aligned} \quad (3)$$

$$\begin{aligned} G(y) &= (B-H)yz - (B - F + P_a + P_b + T)y \\ &\quad -(B-H)z + B - C_a - F + P_a + P_b + T \end{aligned} \quad (4)$$

The agricultural information sharing platform must satisfy the following conditions to be in a stable state: $F(x) = 0$ and $\frac{dF(x)}{dx} < 0$. Since $\frac{dG(y)}{dy} < 0$, $G(y)$ is a decreasing function with respect to y. Therefore, when $y = ((B-H)z - B + C_a + F - P_a - P_b - T)/((B-H)z - B + F - P_a - P_b - T) = y*$, $G(y) = 0$, $\frac{dF(x)}{dx} = 0$, and the strategy cannot be determined. When $y < y*$, $G(y) > 0$ and $\frac{dF(x)}{dx}|_{x=0} < 0$, then $x = 0$ is the ESS. Conversely, when $y > y*$, $y = 1$ is the ESS. The phase diagram of strategy evolution for the agricultural information sharing platform is shown in Figure 2.

4.2 Stability analysis of agricultural information providers' strategies

The expected benefits and average expected benefits (U_{21}, U_{22}, U_2) for meticulous or perfunctory data provision by agricultural information providers are as follows:

TABLE 2 Tripartite game benefit matrix.

Data sharing platform	Data provider	Data consumer (trust data)	Data consumer (distrust data)
Rigorous data review	Meticulous data provision	$-C_a + P_b + H$	$-C_a + P_b$
		$R_a - C_b$	$R_a - C_b + S$
		$-H$	$-C_d - S$
	Perfunctory data provision	$-C_a + P_b + H + T - F$	$-C_a + P_b + T - F$
		$R_b - T$	$R_b - T$
		$-H$	$-C_d$
Lenient data review	Meticulous data provision	$P_b + H$	P_b
		$R_a - C_b$	$R_a - C_b + P_d + S$
		$-H$	$-C_d - S$
	Perfunctory data provision	$-P_a$	$-P_a - B$
		R_b	R_b
		$-P_c$	$-C_d + B$
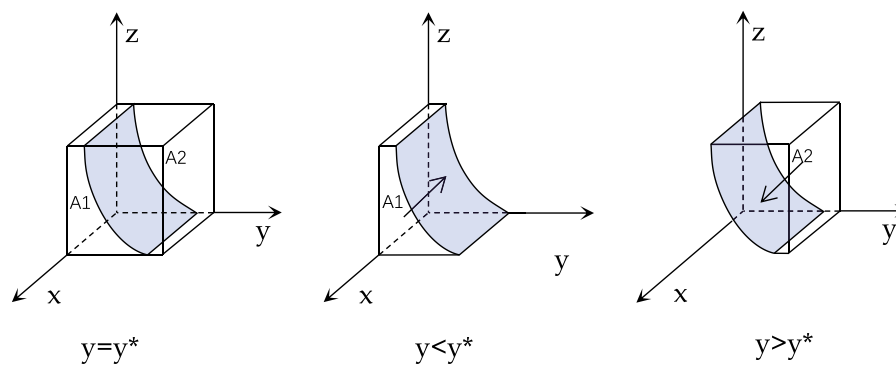
FIGURE 2
Phase diagram of the strategy evolution in agricultural information sharing platforms.

$$\begin{cases} U_{21} & = -C_b + R_a - Sz + S \\ U_{22} & = R_b - Tx \\ U_2 & = yU_{21} + (1 - y)U_{22} \end{cases} \quad (5)$$

The replication dynamics equation for strategy selection by agricultural information providers is:

$$\begin{aligned} F(y) &= y(U_{21} - U_2) \\ &= y(C_b y - C_b - R_a y + R_a + R_b y - R_b + Syz \\ &\quad - Sy - Sz + S - Txy + Tx) \\ &= y(1 - y)(-C_b + R_a - R_b + S + Tx - Sz) \end{aligned} \quad (6)$$

The first-order derivative of y and the set $J(z)$ are respectively:

$$\frac{dF(y)}{dy} = (2y - 1)(-C_b + R_a - R_b + S + Tx - Sz) \quad (7)$$

$$J(z) = -C_b + R_a - R_b + S + Tx - Sz \quad (8)$$

For agricultural information providers to be in a stable state, they must satisfy the conditions: $F(y) = 0$ and $dF(y)dy < 0$, when they provide data meticulously. Since $dF(z)dz < 0$, $J(z)$ is a decreasing function with respect to z. Therefore, when

$z = (-C_b + R_a - R_b + S + Tx)/S = z*$, $J(z) = 0$, $dF(y)dy = 0$, and the strategy cannot be determined. When $z < z*$, $J(z) > 0$, and $dF(y)dy|_y = 0 < 0$, then $y = 0$ is the ESS. Conversely, when $z > z*$, $y = 1$ is the ESS. The phase diagram of strategy evolution for agricultural information providers is shown in Figure 3.

4.3 Stability analysis of agricultural information consumer's strategies

The expected benefits and average expected benefits (U31, U32, U3) for agricultural information consumers' trust or distrust in data are as follows:

$$\begin{cases} U_{31} & = Hxy - Hx - Hy - P_c xy + P_c x \\ & \quad + P_c y - P_c \\ U_{32} & = Bxy - Bx - By + B - C_d - Sy \\ U_3 & = zU_{31} + (1 - z)U_{32} \end{cases} \quad (9)$$

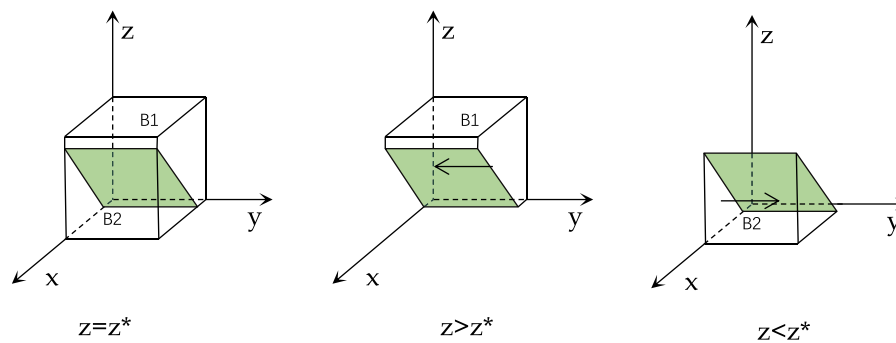The replication dynamic equation for strategy selection of agricultural information providers is:

FIGURE 3
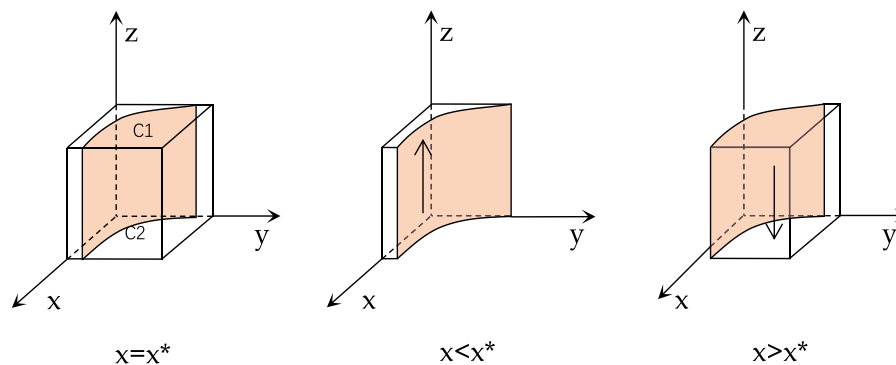Phase diagram of the strategy evolution for agricultural information providers.

FIGURE 4
Phase diagram of the strategy evolution for agricultural information consumers.

$$
\begin{aligned}
F(z) &= z(U_{31} - U_3) \\
&= z(-Hxy + Hx(y-1) + Hy(x-1) \\
&\quad -Pc(x-1)(y-1) + z(Hxy - Hx(y-1) \\
&\quad -yH(x-1) + Pc(x-1)(y-1)) \\
&\quad +(z-1)(C_d xy(x-1)(y-1) + y(C_d + S)(x-1) \\
&\quad +(B-C_d)(x-1)(y-1))) \\
&= z(1-z)(xy(C_d + S - H) + x(H - C_d) \\
&\quad +(H - C_d - S)y(x-1) \\
&\quad -(B - C_d + P_c)(x-1)(y-1))
\end{aligned}
\tag{10}
$$

The first-order derivative of z and the set $H(x)$ are respectively:

$$
\begin{aligned}
\frac{dF(z)}{dz} &= (2z-1)(xy(C_d + S - H) \\
&\quad +x(H - C_d)(y-1) + y(H - C_d - S)(x-1) \\
&\quad -(B - C_d + P_c)(x-1)(y-1))
\end{aligned}
\tag{11}
$$

$$
\begin{aligned}
H(x) &= xy(C_d + S - H) + x(H - C_d)(y-1) \\
&\quad +y(x-1)(H - C_d - S) \\
&\quad -(B - C_d + P_c)(x-1)(y-1)
\end{aligned}
\tag{12}
$$

For agricultural information consumers to trust data in a stable state, the following conditions must be met: $F(z) = 0$ and $\frac{dF(z)}{dz} < 0$. Since $\frac{dH(x)}{dx} > 0$, $H(x)$ is an increasing function with respect to x. Therefore, when $x = xy(C_d + S - H) + (H - C_d)x(y-1) + (H - C_d - S)y(x-1) - (B - C_d + P_c)(x-1)(y-1) = x*$, $H(x) = 0$, and $\frac{dF(z)}{dz}z = 0$,

the strategy cannot be determined. When $x < x*$, $H(x) < 0$, and $\frac{dF(z)}{dz}|_z = 1 < 0$, then $x = 1$ is the ESS. Conversely, when $x > x*$, $x = 0$ is the ESS. The phase diagram of the strategy evolution for agricultural information consumers is shown in Figure 4.

4.4 Stability analysis of the equilibrium points in the three-party evolutionary game system

By utilizing the joint replicator dynamics equations of the agricultural information provider and agricultural information consumer through the agricultural product information sharing platform, a three-party evolutionary game system can be obtained From the equation $F(x) = F(y) = F(z) = 0$, nine system equilibrium points can be derived: (0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1). However, these equilibrium points may not necessarily be the stable evolutionary strategies of the system. Therefore, it is necessary to further determine the stability of these nine points. Based on Friedman's research [28], we constructed the Jacobian matrix of the three-party game as follows:

$$
J = \begin{bmatrix} F_{11} & F_{12} & F_{13} \\ F_{21} & F_{22} & F_{23} \\ F_{31} & F_{32} & F_{33} \end{bmatrix}
\tag{13}
$$

TABLE 3 System equilibrium points and eigenvalues.

Equilibrium	λ_1	λ_2	λ_3
A (0, 0, 0)	$B - C_a - F + P_a + P_b + T$	$-C_b + R_a - R_b + S$	$-B + C_d - P_c$
B (0, 0, 1)	$-C_a - F + H + P_a + P_b + T$	$-C_b + R_a - R_b$	$B - C_d + P_c$
C (0, 1, 0)	$-C_a$	$C_b - R_a + R_b - S$	$C_d - H + S$
D (0, 1, 1)	$-C_a$	$C_b - R_a + R_b$	$-C_d + H - S$
E (1, 0, 0)	$-B + C_a + F - P_a - P_b - T$	$-C_b + R_a - R_b + S + T$	$C_d - H$
F (1, 0, 1)	$C_a + F - H - P_a - P_b - T$	$-C_b + R_a - R_b + T$	$-C_d + H$
G (1, 1, 0)	C_a	$C_b - R_a + R_b - S - T$	$C_d - H + S$
H (1, 1, 1)	C_a	$C_b - R_a + R_b - T$	$-C_d + H - S$

TABLE 4 Stability analysis of equilibrium points.

Equilibrium	Real part of symbol	Stable conditions	Stability
A (0, 0, 0)	(*, *, *)	①	ESS
B (0, 0, 1)	(*, −,*)	②	ESS
C (0, 1, 0)	(−,*, *)	③	ESS
D (0, 1, 1)	(−, +,*)	\	Unstable
E (1, 0, 0)	(*, *, *)	④	ESS
F (1, 0, 1)	(*, *, *)	⑤	ESS
G (1, 1, 0)	(+,*, *)	\	Unstable
H (1, 1, 1)	(+,*, *)	\	Unstable

①$B - C_a - F + P_a + P_b + T < 0; -C_b + R_a - R_b + S < 0; -B + C_d - P_c < 0.$
②$-C_a - F + H + P_a + P_b + T < 0; B - C_d + P_c < 0.$
③$C_b - R_a + R_b - S < 0; C_d - H + S < 0.$
④$-B + C_a + F - P_a - P_b - T < 0; -C_b + R_a - R_b + S + T < 0; C_d - H < 0.$
⑤$C_a + F - H - P_a - P_b - T < 0; -C_b + R_a - R_b + T < 0; -C_d + H < 0.$

Following the principles of evolutionary game theory, equilibrium points' nature can be determined by substituting them into the Jacobian matrix and analyzing the corresponding eigenvalues. An equilibrium point is classified as an ESS when all three eigenvalues are negative. A saddle point corresponds to the presence of one or two negative eigenvalues. Conversely, when all three eigenvalues are negative, it represents a see point. Table 3 provides a comprehensive overview of the eigenvalues associated with the eight equilibrium points obtained through their substitution into the Jacobian matrix.

Upon close scrutiny of Table 3, it is evident that the points (0,1,1), (1,1,0), and (1,1,1) each possess at least one positive eigenvalue, indicating their exclusion from the category of ESS. This highlights the infeasibility of depending solely on the supervision of the agricultural product information sharing platform. In contrast, the remaining five equilibrium points, as extensively examined in Table 4, may potentially acquire the status of ESS under specific conditions.

The comprehension of potential ESS necessitates an exploration of their characteristics and evolutionary trajectories. Within Table 4, the eigenvalues and evolutionary trends of these equilibrium points are presented, facilitating the identification of the conditions

contributing to their stability. As such, the elucidated potential ESS provide valuable managerial insights and act as reference points for informed decision-making in governing the agricultural product information sharing system.

In the subsequent analysis, we will delve into the unique attributes of each ESS point, shedding light on the intricate dynamics that shape the stability of the system and offering a more comprehensive understanding of their potential roles and implications within the agricultural information sharing ecosystem.

4.4.1 Quantitative analysis of point A (0, 0, 0)

When condition ① is met, A is the equilibrium point of a three-player game, and the strategy combination for the three players is {loose data auditing, superficial data provision, distrust of data}. λ_1, λ_2, and λ_3 need to be less than 0 to meet the stability condition. From λ_1, λ_2 and $\lambda_3 < 0$, we can conclude:

$$\begin{cases} -C_a - F + P_b + T < -B - P_a \\ R_a - C_b + S < R_b \\ C_d < P_c + B \end{cases} \qquad (14)$$

These three inequalities represent:

95

- The difference between the reputation income, fine income, and audit cost of the agricultural product information sharing platform and the cost of obtaining correct data again is smaller than the sum of compensation paid by agricultural information consumers and reputation losses.
- The difference between the income of rigorous data provision by agricultural information providers and the income from subsidies received from scientific research institutions is smaller than the benefits of providing data in a perfunctory manner.
- The verification cost of agricultural information consumers is smaller than the sum of the loss of wasted experimental resources and the income from claiming to the platform.

The agricultural product information sharing platform can provide more accurate and reliable data by strengthening the audit process and ensuring the accuracy and completeness of the data, in order to attract more agricultural information consumers and scientific research institutions to use the platform and increase the platform's reputation income.

4.4.2 Qualitative analysis of point B (0, 0, 1)

When condition is met, B is the equilibrium point of a three-player game, and the strategy combination for the three players is {loose data auditing, superficial data provision, trust in data}. Since $R_a < R_b$, we can infer that $\lambda_2 < 0$, and λ_1, λ_3 need to be less than 0 to meet the stability condition. From $\lambda_1, \lambda_3 < 0$, we can conclude:

$$\begin{cases} -C_a - F + H + P_b + T < -P_a \\ B + P_c < C_d \end{cases} \quad (15)$$

These three inequalities represent:

- The difference between the reputation income, fine income, institution bonus income, and audit cost of the agricultural product information sharing platform and the cost of obtaining correct data again is smaller than the reputation loss.
- The verification cost of agricultural information consumers is greater than the sum of the loss of wasted experimental resources and the income from claiming to the platform.

The agricultural product information sharing platform can take measures to help consumers conveniently verify data and provide relevant guidelines and tools. In addition, it can strengthen communication with agricultural information consumers and provide responsive customer support to reduce the likelihood of consumers wasting experimental resources and claiming income.

4.4.3 Quantitative analysis of point C (0, 1, 0)

When condition is met, C is the equilibrium point of a three-player game, and the strategy combination for the three players is {loose data auditing, rigorous data provision, distrust of data}. Since $C_a > 0$, we can infer that $\lambda_1 < 0$, and λ_2, λ_3 need to be less than 0 to meet the stability condition. From $\lambda_2, \lambda_3 < 0$, we can conclude:

$$\begin{cases} R_b < R_a - C_b + S \\ C_d - H < -S \end{cases} \quad (16)$$

These three inequalities represent:

- The profit from perfunctory provision of data by agricultural information providers is smaller than the difference between the profit from rigorous data provision, agricultural subsidy income, and the cost of data collection.
- The verification cost and bonus paid by agricultural information consumers are smaller than the agricultural subsidy paid.

To encourage agricultural information providers to provide data rigorously, the platform can consider offering more appealing incentive measures. This can include increasing the amount or frequency of agricultural subsidies, as well as providing other incentives to encourage providers to offer accurate and complete data, thereby improving the data quality of the platform.

4.4.4 Quantitative analysis of point E (1, 0, 0)

When condition is met, E is the equilibrium point of a three-player game, and the strategy combination for the three players is {strict data auditing, superficial data provision, distrust of data}. λ_1, λ_2 and λ_3 need to be less than 0 to meet the stability condition. From $\lambda_1, \lambda_2,$ and $\lambda_3 < 0$, we can conclude:

$$\begin{cases} C_a + F - P_a < P_b + B + T \\ R_a + S - C_b < R_b - T \\ C_d < H \end{cases} \quad (17)$$

These three inequalities represent:

- The difference between the cost of strict audit, cost of obtaining data again, and reputation loss, and the sum of reputation loss, compensation costs, and fine income, is smaller than reputation loss.
- The difference between the profit from careful selection, agricultural subsidy income, and the difference between the cost of collection, is smaller than the difference between the profit from perfunctory selection and the fine received.
- The cost borne by agricultural information consumers is smaller than the bonus provided to the platform.

To enhance consumers' trust and participation in the platform, the platform can consider providing a reasonable bonus mechanism to encourage consumers to actively participate in agricultural product traceability. This can include providing bonuses or discount coupons as feedback to stimulate consumer attention and willingness to purchase high-quality agricultural products. Agricultural consumers can also establish a comprehensive verification mechanism to reduce verification costs.

4.4.5 Quantitative analysis of point F (1, 0, 1)

When condition is met, F is the equilibrium point of a three-player game, and the strategy combination for the three players is {strict data auditing, superficial data provision, trust in data}. λ_1, λ_2 and λ_3 need to be less than 0 to meet the stability condition. From λ_1, $\lambda_2,$ and $\lambda_3 < 0$, we can conclude:

$$\begin{cases} C_a + F - P_a < P_b + H + T \\ R_a - C_b < R_b - T \\ H < C_d \end{cases} \quad (18)$$

These three inequalities represent:

TABLE 5 Numerical values of different variables and parameters in ESS.

ESS/Value	A (0,0,0)	B (0,0,1)	C (0,1,0)	E (1,0,0)	F (1,0,1)
C_a	80	80	80	80	80
C_b	50	50	10	50	50
C_d	80	80	30	10	80
P_a	20	20	20	20	40
P_b	20	20	20	20	40
P_c	70	20	70	70	70
R_a	100	100	100	100	100
R_b	150	150	110	150	150
T	10	10	10	30	30
F	30	30	30	30	30
B	20	50	20	50	50
S	50	50	30	50	50
H	20	20	70	20	20

- The difference between the cost of strict auditing, the cost of obtaining data again, and reputation loss is smaller than reputation loss, the bonus received from scientific research institutions, and fine income.
- The difference between the profit from careful selection and the cost of collection is smaller than the difference between the profit from perfunctory selection and the fine received.
- The cost borne by agricultural information consumers is greater than the bonus provided to the platform.

To reduce audit costs, the platform can utilize automation and intelligent technologies to process data, while establishing stable partnerships with data providers and collaborating with reliable data sources to reduce the cost of obtaining accurate data.

5 Numerical simulation and analysis

5.1 ESS comparative analysis

We conducted numerical simulation experiments to account for the complexity of relationships between multiple variables and parameters, which are difficult to measure accurately in practical situations. To ensure wide applicability and practicality of the obtained patterns, we established clear principles. These principles involve setting numerical values that accurately reflect the actual situation and minimizing changes in other variables when the ESS changes. This facilitates subsequent parameter sensitivity analysis. By adhering to these principles, our research effectively explores the evolutionary trends and patterns of three-party strategy equilibrium. See Table 5 for the initial numerical settings.

(1) (0,0,0) and (0,0,1) The distinction between the two ESS hinges on whether agricultural information consumers adopt a strategy of trusting the data. It is noteworthy that as this choice is made, P_c decreases while B increases. This suggests that augmenting the compensation cost for agricultural information consumers and curbing the wastage of experimental resources tend to steer agricultural information consumers towards embracing the data trust strategy. Figure 5. and Figure 6 illustrate the evolutionary trends within these two environmental systems. Notably, the x and y curves exhibit relatively minimal variations, while the z curve undergoes a transformation from 0 to 1. Initially, the z curve ascends and subsequently descends, indicating that research institutions' inclination to adopt the trust strategy is not particularly strong. Numerous numerical experiments have consistently demonstrated that further reductions in C_d can abbreviate the equilibrium time and sustain a consistent downward trend in the curve.

(2) (0,0,0) and (0,1,0) The distinction between the two ESS lies in whether the agricultural information provider opts for a meticulous data provisioning strategy. It is noticeable that, with this choice, C_b, C_d, S experience a decrease, while H registers an increase. This indicates that the agricultural information provider inclines towards a rigorous data provisioning strategy due to the reduced costs associated with collecting data rigorously, lower verification expenses for agricultural information consumers, a decrease in agricultural subsidies, and an increase in bonuses received by the platform from research institutions. Figure 5 and Figure 7 depict the evolutionary trends within these two environmental systems. Remarkably, the x curve exhibits relatively minimal variations, while the z curve displays a faster rate of decline. Meanwhile, the y curve undergoes a transformation from 0 to 1, albeit with a longer equilibrium time. This suggests a moderate level of enthusiasm from the agricultural information provider when it comes to rigorously providing data. Multiple numerical experiments have
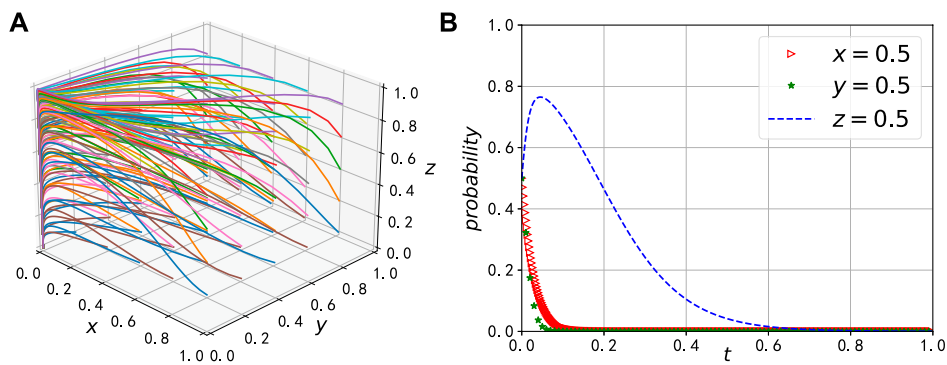
FIGURE 5
Evolutionary trajectory of *A* (0, 0, 0). **(A)** The 3D evolutionary chart, **(B)** Equilibrium state under the (corresponding point).
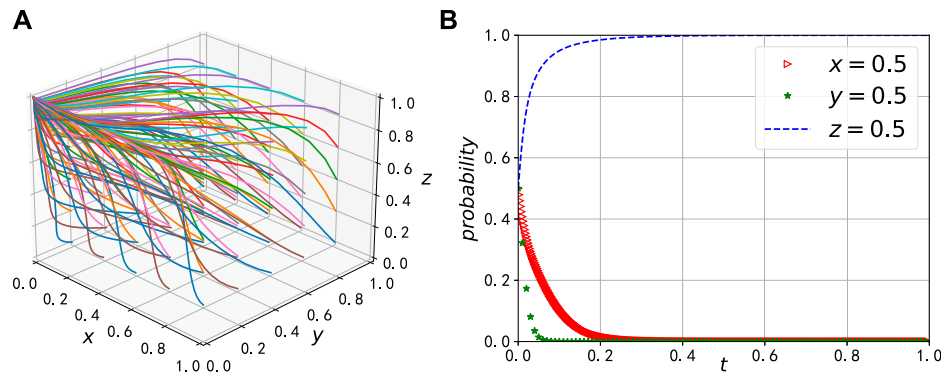

FIGURE 6
Evolutionary trajectory of *B* (0, 0, 1). **(A)** The 3D evolutionary chart, **(B)** Equilibrium state under the (corresponding point).
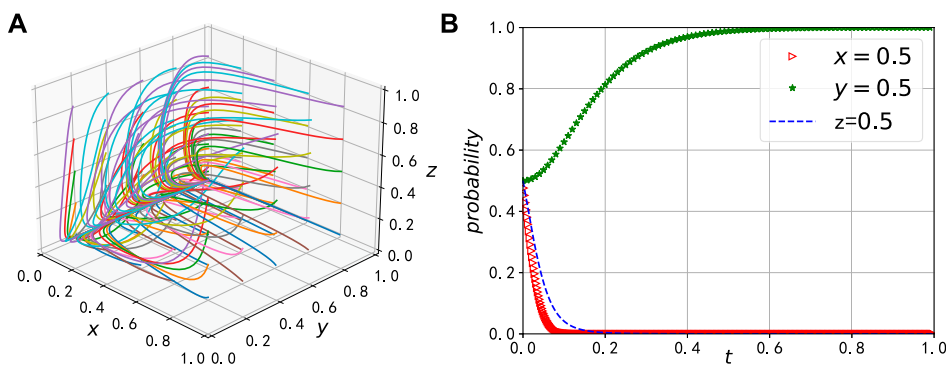

FIGURE 7
Evolutionary trajectory of *C* (0, 1, 0). **(A)** The 3D evolutionary chart, **(B)** Equilibrium state under the (corresponding point).

consistently demonstrated that further reductions in T have the potential to abbreviate the equilibrium time.

(3) (0,0,0) and (1,0,0) The distinction between the two ESS revolves around whether the agricultural product information sharing platform opts for a stringent data auditing strategy. It is evident that, with this choice, C_d experiences a decrease, while both *T* and *B* register an increase. This suggests that a reduction in the verification cost for agricultural information consumers, a decrease in penalties imposed by the platform on agricultural information providers, and a decline in compensation claims from research institutions to the platform collectively drive the agricultural product traceability platform toward a preference for the rigorous data auditing
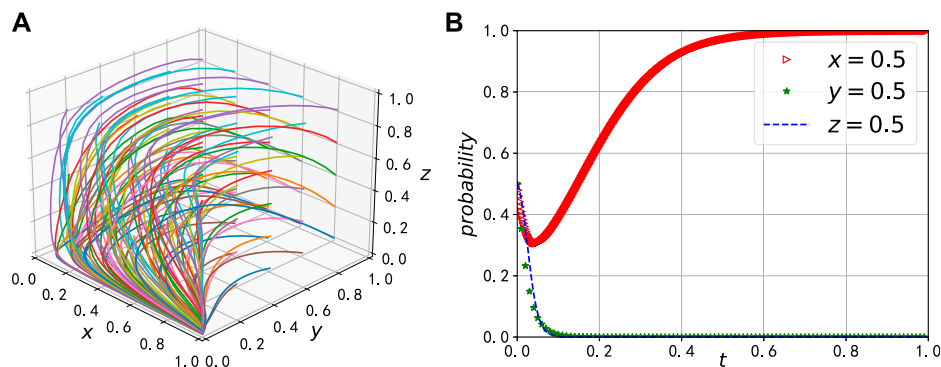
FIGURE 8
Evolutionary trajectory of E (1, 0, 0). **(A)** The 3D evolutionary chart, **(B)** Equilibrium state under the (corresponding point).

strategy. Figure 5 and Figure 8 portray the evolutionary trends within these two environmental systems. Notably, the y and z curves exhibit relatively minimal variations, while the x curve undergoes a transformation from 0 to 1. However, it initially decreases before increasing, resulting in a longer equilibrium time. This pattern indicates a reduced level of enthusiasm from the agricultural product traceability platform in conducting thorough data audits. Multiple numerical experiments have consistently shown that further reductions in P_a and P_b have the potential to abbreviate the equilibrium time.

(4) (0,0,1) and (1,0,1) The disparity between the two ESS centers on whether the agricultural product information sharing platform adopts a rigorous data auditing strategy. It becomes evident that, in making this choice, P_a, P_b, P_c, and T all experience an increase. This signifies an escalation in the reputation benefits and losses for the platform, the conservation of resources by research institutions, and the heightened penalties imposed by the platform on agricultural information providers. These factors collectively drive the agricultural product traceability platform toward a preference for the rigorous data auditing strategy. Figure 6 and Figure 9 depict the evolutionary trends within these two environmental systems. Remarkably, the y and z curves exhibit relatively minimal variations, while the x curve undergoes a transformation from 0 to 1. As discussed in part (3) of the analysis, further reductions in P_a and P_b have the potential to abbreviate the equilibrium time.

(5) (1,0,0) and (1,0,1) The distinction between the two ESS hinges on whether agricultural information consumers adopt a strategy of trusting the data. It can be observed that as this choice is made, C_d, P_a, and P_b all experience an increase. This signifies an escalation in the verification cost for research institutions and the reputation benefits and losses for the platform, thereby driving agricultural information consumers toward a preference for trusting the data strategy. Figure 8 and Figure 9 delineate the evolutionary trends within these two environmental systems. Notably, the x curve, as analyzed in part (3), and the y curve exhibit relatively minimal variations. In contrast, the z curve undergoes a transformation from 0 to 1 in a smooth and rapid manner, indicating an elevated level

of enthusiasm among agricultural information consumers to embrace the data trust strategy.

5.2 Initial probability analysis

In order to explore the impact of initial probabilities on the equilibrium of the three participants, the variables and parameters are kept constant.

In Figure 10, we set the initial probabilities of y and z to be both 0.5 and observe the evolution pattern of the system by changing the initial strategy of the agricultural product information sharing platform. As shown in Figure 10A, when the probability of strict data auditing increases from 0.2 to 0.8, the evolution speed of the agricultural information provider adopting a rigorous data provisioning strategy will increase, converging at a relatively stable rate. As shown in Figure 10B, the behavior of agricultural information consumers is less influenced, and their convergence rate to distrust is negatively correlated with z. This indicates that the auditing behavior of the agricultural product information sharing platform will influence the decision of whether agricultural information providers provide data rigorously or not, and strict auditing will compel them to provide more rigorous data.

In Figure 11, we initially set both the probabilities of x and z to 0.5 and observe the system's evolution pattern by varying the initial strategy of the agricultural information provider. As illustrated in Figure 11A, when the probability of rigorous data provisioning increases from 0.2 to 0.8, the behavior of the agricultural product information sharing platform is minimally affected, and its convergence rate to lenient data auditing remains relatively stable. In Figure 11B, the convergence rate of agricultural information consumers towards distrust is inversely correlated with the value of z. This implies that the data provisioning behavior of the agricultural information provider significantly influences the decision of whether agricultural information consumers choose to trust or not.

In Figure 12, we initialize the strategies of x and y at 0.5 and observe the system's evolution pattern by altering the initial strategy of the agricultural information consumer. As z gradually increases from 0.2 to 0.8, as depicted in Figure 12A, the behavior of the agricultural product information sharing platform is minimally
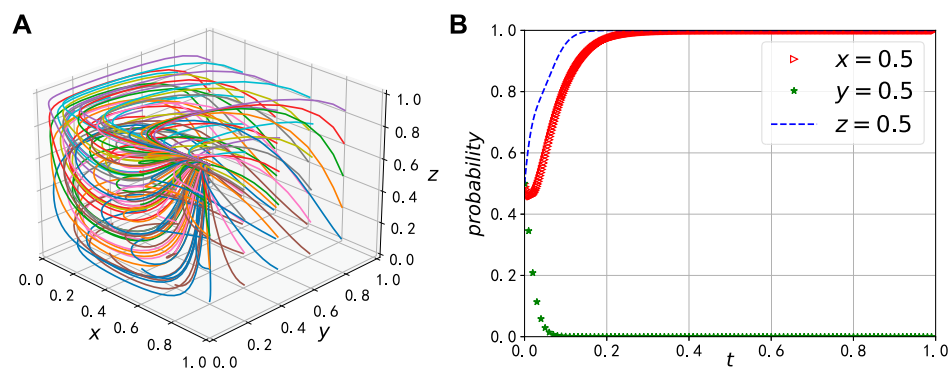
FIGURE 9
Evolutionary trajectory of F (1, 0, 1). **(A)** The 3D evolutionary chart, **(B)** Equilibrium state under the (corresponding point).

FIGURE 10
The impact of the initial probability of x on the evolutionary outcome and trajectory. **(A)** Effect on the evolution of y, **(B)** Effect on the evolution of z.

FIGURE 11
The impact of the initial probability of y on the evolutionary outcome and trajectory. **(A)** Effect on the evolution of x, **(B)** Effect on the evolution of z.

affected, and its convergence rate towards lenient data auditing remains nearly unchanged. As shown in Figure 12B, the evolutionary pace of the agricultural information provider, which adopts a rigorous data provisioning strategy, decelerates, and there are intermittent fluctuations towards the provision of perfunctory data. However, on the whole, it still leans towards the trend of providing well-researched data. This underscores the significant impact of strategy choices made by agricultural information consumers on the entire system. Since data represents a pivotal resource flowing throughout the entire sharing system, the reliability of the entire sharing system hinges on the decisions made by agricultural information consumers.
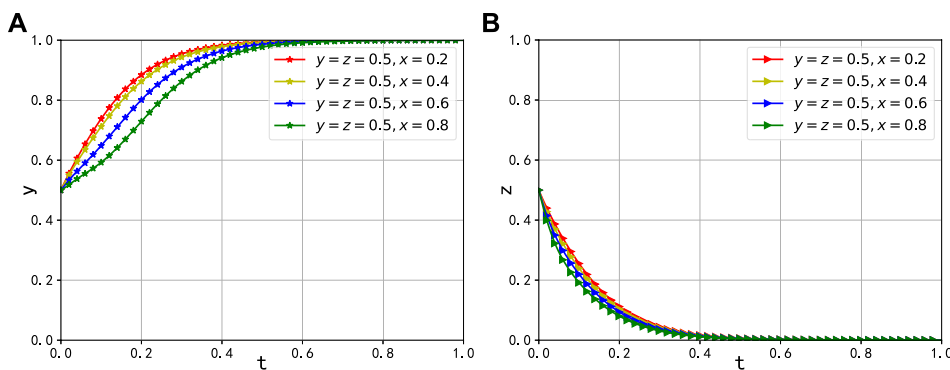
FIGURE 12
The impact of the initial probability of z on the evolutionary outcome and trajectory. **(A)** Effect on the evolution of x, **(B)** Effect on the evolution of y.
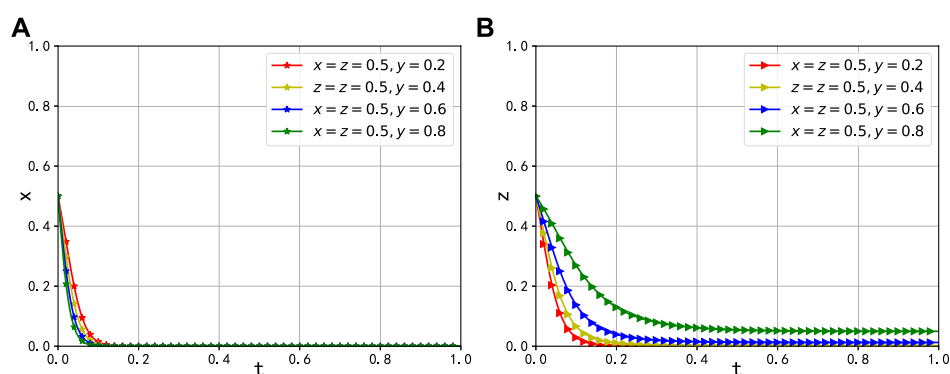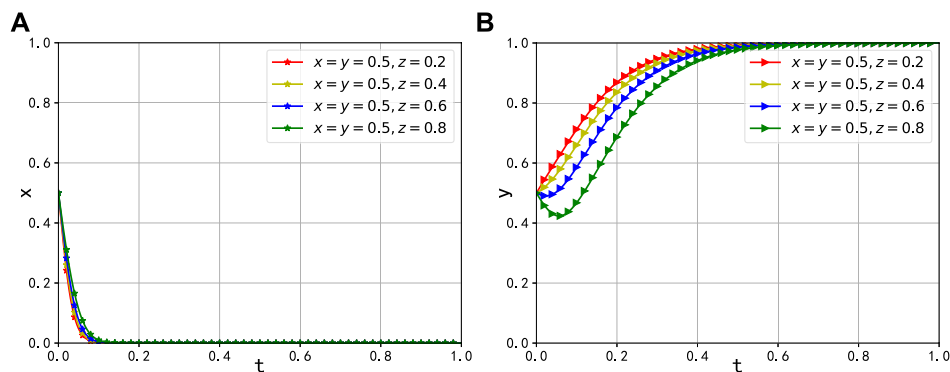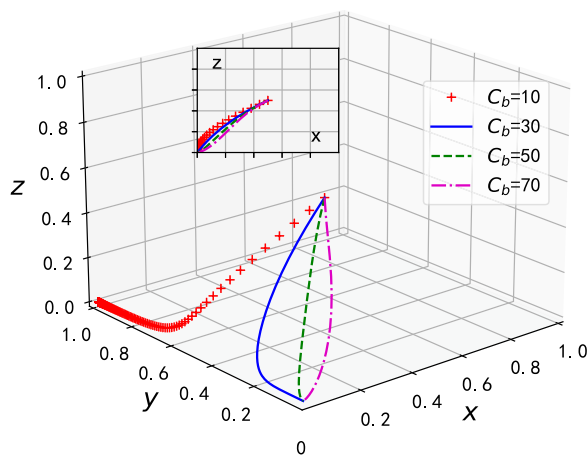

FIGURE 13
The impact of C_b on the evolutionary outcome and trajectory.
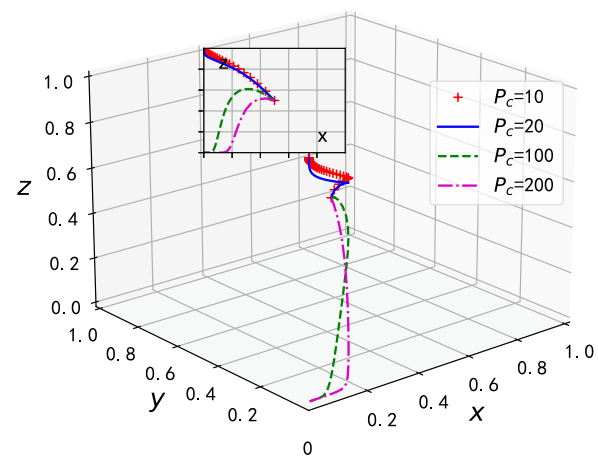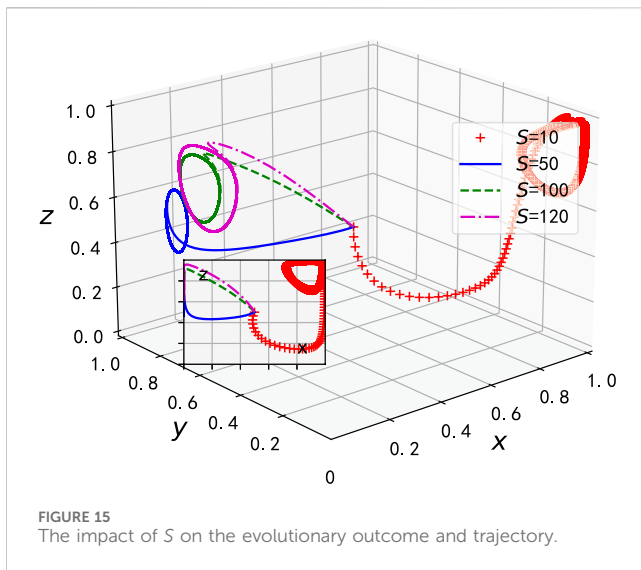

FIGURE 14
The impact of P_c on the evolutionary outcome and trajectory.

5.3 Partial parameter analysis

When the agricultural product information sharing platform chooses to relax data review standards, and scientific research institutions decide to distrust the data, agricultural information providers tend to reduce the cost of data collection rigor, as illustrated in Figure 13. With an initial value of C_b set at 30 and other parameters held constant, we examine the impact of reducing C_b to 10 and subsequently increasing it to 50 and 70, respectively. As C_b decreases to 10, the equilibrium point shifts from (0, 0, 0) to (0, 1, 0). However, when C_b increases to 50 and 70, there is no significant change in the equilibrium point. Furthermore, considering its implications in the $x − z$ dimension, as C_b further increases to 70, the efficiency of agricultural information providers in evolving towards rigorous data provision also improves. This suggests that with lower values of C_b, agricultural information providers typically adopt the strategy of immediately providing data rigorously. Conversely, when the cost of data collection rises due to the necessity of purchasing professional equipment or hiring agricultural experts, they are more inclined to opt for perfunctory data provision, thereby increasing their profits. Hence, it is plausible

to promote rigorous data provision by reducing the cost and price of professional equipment.

When the agricultural information sharing platform chooses to relax the review of data, and agricultural information providers choose to provide data perfunctorily, agricultural information consumers tend to reduce the waste of experimental resources caused by perfunctory low-quality data, as shown in Figure 14, we conducted a study on the impact of varying the initial value of P_c in an agricultural information consumer model. When the initial value of P_c is 100, we investigated its reduction to 20 and 10, as well as its increase to 200, while keeping the other parameters constant. When P_c decreases to 20 and 10, we observed a shift in the equilibrium point from (0, 0, 0) to (0, 0, 1). However, when P_c increases to 200, there is no notable change in the equilibrium point. Furthermore, the model predicts that as P_c further increases to 200, the efficiency of agricultural information consumers in evolving towards distrusting the data improves. This improvement is particularly evident when considering the x-z relationship. This indicates that for a smaller P_c, agricultural information consumers usually adopt the strategy of trusting the data immediately. However, when perfunctory low-quality data causes

FIGURE 15
The impact of *S* on the evolutionary outcome and trajectory.

FIGURE 16
The impact of *T* on the evolutionary outcome and trajectory.

errors in experimental results or incorrect influences on research ideas, leading to an overall waste of experimental resources and an increase in the cost of resource waste, they are more willing to choose to distrust the data and prefer to incur certain costs to verify the data in order to ensure the accuracy of scientific research experiments.

When agricultural information consumers confirm that the data is reliable and effective, they will provide agricultural subsidies to encourage agricultural information providers to continue growing crops and providing rigorous and reliable data in accordance with the original methods. If the bonuses received are generous enough, it will also incentivize agricultural information providers to continue practicing strict agriculture and provide reliable data. Hence, we initially set the value of *S* to 10 and gradually increased it to 50, 100, and 120 to observe its impact on decision-making behavior. As depicted in Figure 15, it becomes evident that the decision strategies of the three participants converge towards (0, 1, 1). Moreover, upon examining the $x - z$ projection graph, it is apparent that higher values of *S* lead to a decision bias favoring trust in data from scientific research institutions, while agricultural information providers also tend to offer data more rigorously. This suggests that as subsidies from scientific research institutions to agricultural information providers increase, the latter are incentivized to actively and rigorously provide data. Consequently, scientific research institutions also develop greater trust in the data due to the improved data provision from agricultural information providers, thus establishing a virtuous cycle.

When agricultural information providers choose to provide data perfunctorily, the agricultural product information sharing platform aims to mitigate the wastage of experimental resources resulting from subpar data, as illustrated in Figure 16. With an initial value of *T* set at 30, and keeping other parameters constant, we investigate the impact of reducing *T* to 10 and subsequently increasing it to 80 and 100.As *T* decreases to 10, there is no significant change in the equilibrium point. However, when *T* increases to 80 and 100, the equilibrium point transitions from (0, 0, 0) to (1, 0, 1). Furthermore, when considering its implications in the $x - z$ dimension, a decrease in *T* to 10 leads to an enhanced efficiency among agricultural information consumers in evolving towards a state of distrust in
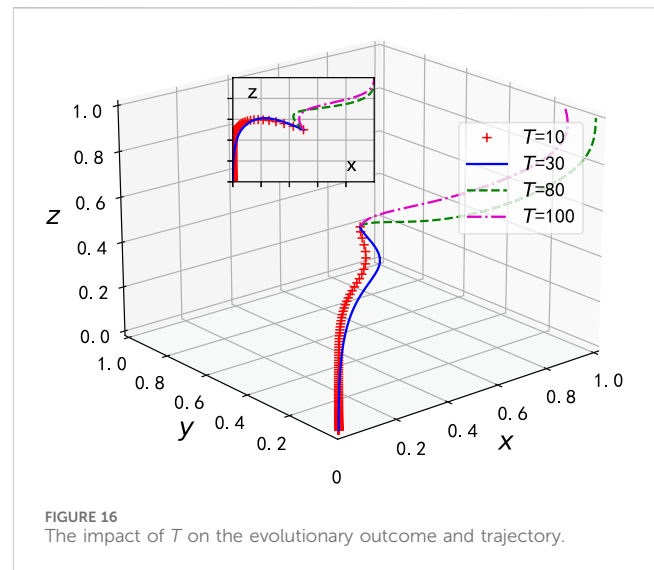
the data. This suggests that for smaller values of *T*, agricultural information consumers tend to adopt a strategy of distrusting the data. Nevertheless, when the severity of penalties for perfunctory data reaches a certain threshold, it indirectly promotes the rigor and reliability of the data. Consequently, agricultural information consumers become more inclined to trust the data.

6 Conclusion and future works

In the era of agricultural informatization, the reliability of data and the sharing of information are crucial for the sustainable development of the entire agricultural ecosystem. This paper delves into the interaction between agricultural information providers, agricultural information consumers, and information sharing platforms from the perspective of game theory. Specifically, we examine the system's evolutionary trends in light of whether agricultural information providers adopt a rigorous data-providing strategy and whether agricultural information consumers adopt a trusting data strategy. Based on the findings of this study, we propose some targeted recommendations.

1. Agricultural information sharing platform should establish a stringent review mechanism to ensure the reliability and accuracy of the agricultural data provided. This will enhance the trust of agricultural information consumers and promote information sharing and circulation. Simultaneously, implementing incentive measures is necessary to encourage agricultural information providers to provide reliable data. This will actively motivate them to collect and provide data, consequently improving the overall reliability of the system. Additionally, the information sharing platform should actively explore mutually beneficial cooperation mechanisms with other institutions to secure additional subsidies and policy support from the government.

2. Agricultural information providers should adopt a rigorous attitude to ensure that data collection methods are scientifically reliable, with regular data validation and

auditing. They should also evaluate strategies to improve data provision. This ensures that the provided data is reliable and accurate, which helps to enhance reputation and credibility, and enables access to more agricultural subsidies and bonuses, thereby increasing the benefits of data sharing.

3. It is important to choose a reliable platform to provide data for agricultural information consumers. This can promote information sharing and collaboration, reduce waste of experimental resources, and improve system efficiency. However, when using the data, consumers should remain cautious and verify its accuracy to ensure the reliability of scientific research experiments. By establishing a foundation of trust on the platform, consumers can better balance the relationship between information sharing and experiment reliability.

In future, we will further explore sustainable development strategies and methods for agricultural information sharing ecosystems. Here are some potential research agendas:

1. Model Refinement and Expansion: The model can be further improved by considering more real-world factors, such as market changes and policy adjustments, in order to enhance its accuracy and applicability. This will enable a more accurate analysis of decision-making behaviors and strategic choices of stakeholders in different scenarios.
2. Policy Formulation and Regulation: We will further study how to guide the development of agricultural product information sharing platforms through policy formulation and regulation. This includes incentive mechanisms, reward policies, and penalties for improper behavior.
3. Transaction Environment Construction: Adopting blockchain and mobile sensing technology [29] to build a secure and convenient agricultural data sharing trading platform. Additionally, this platform also provides data analysis, decision support, and agricultural management services to assist users in making better decisions and managing production using agricultural data.

These future research directions will contribute to deepening our understanding of the agricultural information sharing ecosystem and provide stronger support and guidance for promoting agricultural sustainable development.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The manuscript presents research on animals that do not require ethical approval for their study.

Author contributions

HZ: Writing–original draft, Writing–review and editing. JY: Formal Analysis, Funding acquisition, Supervision, Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Osinga SA, Paudel D, Mouzakitis SA, Athanasiadis IN. Big data in agriculture: between opportunity and solution. *Agric Syst* (2022) 195:103298. doi:10.1016/j.agsy.2021.103298

2. Qu D, Wang X, Kang C, Liu Y. Promoting agricultural and rural modernization through application of information and communication technologies in China. *Int J Agric Biol Eng* (2018) 11:1–4. doi:10.25165/j.ijabe.20181106.4228

3. Klerkx L, Jakku E, Labarthe P. A review of social science on digital agriculture, smart farming and agriculture 4.0: new contributions and a future research agenda. *NJAS-Wageningen J Life Sci* (2019) 90:1–16. doi:10.1016/j.njas.2019.100315

4. Kamilaris A, Kartakoullis A, Prenafeta-Boldú FX. A review on the practice of big data analysis in agriculture. *Comput Elect Agric* (2017) 143:23–37. doi:10.1016/j.compag.2017.09.037

5. Susha I, Rukanova B, Zuiderwijk A, Gil-Garcia JR, Hernandez MG. Achieving voluntary data sharing in cross sector partnerships: three partnership models. *Inf Organ* (2023) 33:100448. doi:10.1016/j.infoandorg.2023.100448

6. Mirjalili S, Mirjalili S. Genetic algorithm. Evolutionary Algorithms and neural networks: Theory and applications," in *Studies in computational intelligence*. Editors Mirjalili S. (Springer) (2019). p. 43–55.

7. Liu Y, Zuo J, Pan M, Ge Q, Chang R, Feng X, et al. The incentive mechanism and decision-making behavior in the green building supply market: a tripartite evolutionary game analysis. *Building Environ* (2022) 214:108903. doi:10.1016/j.buildenv.2022.108903

8. Herrera-Medina E, Riera Font A. A multiagent game theoretic simulation of public policy coordination through collaboration. *Sustainability* (2023) 15:11887. doi:10.3390/su151511887

9. Aaronson SA. *Could a global "wicked problems agency" incentivize data sharing?* United States: George Washington University (2023).

10. Shrivastava S, Marshall-Colon A. Big data in agriculture and their analyses. In: *Encyclopedia of food security and sustainability*. Amsterdam, Netherlands: Elsevier (2018). p. 233–7.

11. Guo T, Wang Y. Big data application issues in the agricultural modernization of China. *Ekoloji Dergisi* (2019).

12. Hu J, Zhang H, Irfan M. How does digital infrastructure construction affect low-carbon development? a multidimensional interpretation of evidence from China. *J Clean Prod* (2023) 396:136467. doi:10.1016/j.jclepro.2023.136467

13. Ngo QH, Kechadi T, Le-Khac N-A. Knowledge representation in digital agriculture: a step towards standardised model. *Comput Elect Agric* (2022) 199: 107127. doi:10.1016/j.compag.2022.107127

14. Zhu L, Li F. Agricultural data sharing and sustainable development of ecosystem based on block chain. *J Clean Prod* (2021) 315:127869. doi:10.1016/j.jclepro.2021. 127869

15. Liu W, Shao X-F, Wu C-H, Qiao P. A systematic literature review on applications of information and communication technologies and blockchain technologies for precision agriculture development. *J Clean Prod* (2021) 298:126763. doi:10.1016/j. jclepro.2021.126763

16. Hu J. Synergistic effect of pollution reduction and carbon emission mitigation in the digital economy. *J Environ Manag* (2023) 337:117755. doi:10.1016/j.jenvman.2023. 117755

17. Xie N, Zhang X, Sun W, Hao X. Research on big data technology-based agricultural information system. In: *International conference on computer information systems and industrial applications*. Netherlands: Atlantis Press (2015). p. 388–90.

18. Yu-xiao L, Min C, Xu-wen G, Lian-feng Z. Research on the utilization and the management of agriculture information resources on internet. In: 2011 IEEE 3rd International Conference on Communication Software and Networks (IEEE); Held 27-29 May 2011; Xi'an, China (2011). p. 542–4.

19. Zhu Z, Wang X, Liu L, Hua S. Green sensitivity in supply chain management: an evolutionary game theory approach. *Chaos, Solitons and Fractals* (2023) 173:113595. doi:10.1016/j.chaos.2023.113595

20. Hua S, Hui Z, Liu L. Evolution of conditional cooperation in collective-risk social dilemma with repeated group interactions. *Proc R Soc B* (2023) 290:20230949. doi:10. 1098/rspb.2023.0949

21. Hua S, Liu L. Coevolutionary dynamics of population and institutional rewards in public goods games. *Expert Syst Appl* (2024) 237:121579. doi:10.1016/j.eswa.2023. 121579

22. GuoHua Z, Wei W Study of the game model of e-commerce information sharing in an agricultural product supply chain based on fuzzy big data and lsgdm. *Technol Forecast Soc Change* (2021) 172:121017. doi:10.1016/j.techfore.2021.121017

23. Ren H, Huang H, Li Q, Wu Q, Yang Y. Operation optimization of multi-participants in a regional energy system based on evolutionary game theory. *Energ Rep* (2020) 6:1041–5. doi:10.1016/j.egyr.2020.11.079

24. Han R, Yang M. Profit distribution and stability analysis of joint distribution alliance based on tripartite evolutionary game theory under the background of green and low carbon. *Environ Sci Pollut Res* (2022) 29:59633–52. doi:10.1007/s11356-022-19712-y

25. Yang J, Yan X, Yang W. A tripartite evolutionary game analysis of online knowledge sharing community. *Wireless Commun Mobile Comput* (2022) 2022: 1–11. doi:10.1155/2022/4460034

26. Gao Y, Zhu Z, Yang J. An evolutionary game analysis of stakeholders' decision-making behavior in medical data sharing. *Mathematics* (2023) 11:2921. doi:10.3390/ math11132921

27. Hao H, Yang J, Wang J. A tripartite evolutionary game analysis of participant decision-making behavior in mobile crowdsourcing. *Mathematics* (2023) 11:1269. doi:10.3390/math11051269

28. Friedman D. Evolutionary games in economics. *Econometrica: J Econometric Soc* (1991) 59:637–66. doi:10.2307/2938222

29. Yang J, Di Zhang CH, Wang K. A spatial coverage-based participant recruitment scheme for mobile crowdsourcing. *Human-centric Comput Inf Sci* (2023) 13(20).

A study of production outsourcing strategies of dual oligopoly manufacturers based on quality investments

Yang Wang*, Tongbo Deng and Huibing Cheng

Guangzhou Railway Polytechnic, Guangzhou, Guangdong, China

Introduction: Self-production and outsourcing are two important production strategies for manufacturers, and the production capacity and investment capacity of manufacturers and suppliers play a decisive role in the quality of products. This study aims to analyze the manufacturer's best production strategy and the drivers of outsourcing in the context of quality.

Methods: This study constructs production outsourcing game models for duopoly manufacturers, examines the trade-offs between self-production and outsourcing when suppliers have the ability to invest in quality, explores the requirements and implications of outsourcing and compares the differences between "one-to-one" and "one-to-many" outsourcing structures.

Results: First, outsourcing can reduce the level of product quality and that a necessary condition for manufacturers' outsourcing is a strong advantage in the supplier's production costs. Second, duopoly manufacturers may face a prisoner's dilemma as a result of outsourcing. Finally, compared to two independent suppliers, outsourcing to a common supplier can increase the level of product quality by exploiting the centralization effect and increase the firm's profits when the market's competition intensity is low.

Conclusion: First, the production strategy balance of duopoly manufacturers is closely related to the outsourcing structure, production efficiency and investment efficiency. Second, duopoly manufacturers choose outsourcing may fall into the prisoner's dilemma, and outsourcing has the risk of quality reduction.

1 Introduction

The globalization of supply chains and the division of labor refinement have made production outsourcing more and more popular among manufactures. Self-production and outsourcing have become important production strategies for manufactures [1]. As consumers pay more attention to product quality, products with advanced processes tend to win consumers' favor. For this reason, manufactures are not only investing in internal production processes, but also actively investing in supplier capabilities [2], and even proactively seeking external suppliers with independent investment capabilities. This quality investment is essentially investing in innovating the production process to improve the quality of the product and increase the market share.

A leading example is Intel and AMD in the semiconductor industry, where the quality of the CPU products depends heavily on the performance of the chips. This requires a large injection of technology, and long technology refresh cycles can lead to market and consumer loss. According to a ChinaIRN.com study, Intel insisted on the route of integrated chip makers, building factories to produce chips on their own, thus occupying a dominant position in the field of chip manufacturing in the United States for nearly 50 years. AMD, on the other hand, outsourced its chip production to TSMC, a specialized foundry, and relied on TSMC's advanced technological innovations to gradually rise to a strong competitor in the semiconductor industry. H However, due to technical errors, Intel has lagged behind TSMC in recent years by a whole generation of technology updates, and product quality has been questioned. Forced by competition and technological pressure, Intel announced that it will launch its leading CPU products manufactured by TSMC in 2023[1].

The growing number of outsourcing cases in integrated electronics shows that the pursuit of higher levels of product quality has become an important driver for manufacturers to choose outsourcing. However, outsourcing makes manufactures' performance in competitive markets more dependent on supplier performance [3], and the impact of outsourcing structure and outsourced suppliers cannot be ignored. On the one hand, different manufacturers may choose the same supplier for outsourcing, which can raise concerns about product differentiation among manufacturers. For example, in response to Intel's outsourcing decision, CEO Gelsinger stated that outsourcing to TSMC would make it more difficult to differentiate its chips from AMD's, because they would be manufactured on the same process, potentially permanently preventing Intel from overtaking AMD. On the other hand, while stronger supplier quality investment capabilities can significantly increase the market demand for a product, higher product quality levels may prompt suppliers to increase the price they charge manufacturers for outsourcing. As a result, manufacturers have to face cost and quality tradeoffs given quality investments and production processes. For example, in 2018, Samsung, in order to compete with TSMC for chip orders from potential customers, not only made every effort to develop process technologies that could rival TSMC's, but even took the initiative to lower its foundry prices by 20%. However, the actual performance of Samsung is unsatisfactory due to customers' concerns about product process quality and qualification rate. According to SamMobile, until 2023, Samsung has made significant breakthroughs in volume production and yields of 3 nm chips, but competition with TSMC remains challenging.

Based on this fact, we will restrict our attention to the following issues. Can a manufacturer gain a competitive advantage through outsourcing when the supplier has the ability to invest on its own? For different outsourcing structures, under what market conditions do manufacturers outsourcing? How does production strategy affect the level of investment and profitability of manufacturers? How do duopoly manufacturers interact their strategies? To address these questions, this paper considers two outsourcing structures to construct dynamic game models of duopoly manufacturers, compares and analyzes the equilibrium results under different combinations of self-production and outsourcing. We also verify

the robustness of the conclusions with the arithmetic example analyses.

The theoretical contributions of this paper are:

(1) First, this paper can expand and enrich operations management research by extending outsourcing motivations from cost to investment. The quantitative approach fills the gap in explaining investment-based outsourcing.
(2) Second, this paper examines suppliers' production and investment capabilities to increase supply chain value. This is to compensate for the lower supplier dynamism in traditional production strategies.
(3) Finally, by examining different structures of outsourcing, this paper extends the framework for the study of supply chain competition and provides new insights into whether competition encourages or discourages investment.

The research in this paper contributes to the understanding of the important role of quality investment elements in outsourcing activities, with a view to informing the decision-making of manufacturing firms' production strategies.

This paper contributes to understanding the role of quality investment in outsourcing, with a view to informing production strategy decisions for manufacturers and suppliers.

2 Literature review

Our work belongs to the research field of production outsourcing. Earlier literatures explored the motivations, impacts, and decision-making frameworks of outsourcing from a strategic perspective, utilizing resource dependence theory, core competitiveness theory, and transaction cost theory [4–6].

From the perspective of operation management, firstly the impacts of outsourcing on supply chain members have been discussed. Gilbert et al. [7] argued that outsourcing mitigates downstream manufacturers' overinvestment in reducing production costs, thus mitigating mutually destructive cost competition. Kenyon et al. [8] found that production outsourcing has a negative impact on operational performance and that outsourcing significantly reduces the effectiveness of operational equipment and on-time delivery, and even negatively affects customer loyalty. Heydari et al. [9] argued that the use of flexible quantity contracts and outsourcing strategies under stochastic demand can share the risk of overstocking and overproduction, and that partial outsourcing can increase the profitability of the supply chain. Lee et al. [10] found that outsourced supplier development not only directly improves outsourcing performance but also indirectly improves outsourcing performance by facilitating contract and relationship governance.

Second, outsourcing forms and supplier selection have been investigated. Chen et al. [11] studied production strategies with demand uncertainty, quantity uncertainty and outsourcing information asymmetry for the supplier selection problem in outsourcing. Xiao and Qi [12] discussed the effect of product quality level on production outsourcing strategy and show that symmetric outsourcing tends to increase product line variability and that asymmetric outsourcing outperforms symmetric outsourcing when non-quality costs are low. Chen et al. [13] constructed the outsourcing relationship between an original equipment

manufacturer and a contract manufacturer when they compete in the same product market, identifying boundary conditions under which both parties can benefit from outsourcing. Niu et al. [14] investigated the dual horizontal outsourcing strategy of "production + sourcing" and its impact on the cost of investment in innovation from the perspective of the learning capability of manufacturers.

Quality investment has been a hot topic in operations management since the 1990s. Quality investment in the supply chain depends on both the manufacturer's or supplier's own development capability and the degree of cooperation among the members of the supply chain [15–17]. Karaer and Erhun [18] investigated how the decision of potential firms to enter a monopoly market is influenced by quality factors. Xie et al. [19] extended intra-chain competition to chain-chain competition to study the mechanism of quality improvement strategies in market segments. Xia et al. [20] analyzed the quality investment and pricing decisions of start-ups under financial constraints. In addition, there are a number of works dealing with cost sharing, investment incentives and supply chain coordination between manufacturers and suppliers in cooperative quality investment. Chen et al. [21] examined the importance of cooperative quality investment strategies, obtained analytical equilibrium by establishing a three-stage dynamic game framework, and proposed an investment share sharing agreement applicable to outsourced supply chains. Nagurney and Li [22] proposed a supply chain network quality management problem applicable to outsourced distribution with the objective of weighted minimization of production cost and reputation loss. Yang et al. [23] study the raw material quality control problem in production outsourcing, and design a procurement outsourcing contract to incentivize suppliers to select higher quality raw materials. Xiao et al. [24] studied the strategic outsourcing decisions of two competing manufacturers faced with improving the quality of key components, and the results show that asymmetric outsourcing is not a good solution to the problem. Xiao et al [24] study the strategic outsourcing decisions of two competing manufacturers faced with improving the quality of critical components, and the results show that the manufacturer that adopts an outsourcing strategy in asymmetric outsourcing can achieve a higher level of quality investment. Karaer et al. [25] investigated how manufacturers can use full control strategies and cost sharing mechanisms to develop the sustainable quality capabilities of their suppliers.

The literature has the following findings. Firstly, there are two different types of investment in outsourcing: process investment, which aims at reducing costs and gaining a price advantage, and quality investment, which aims at improving quality and gaining a non-price advantage. There are more studies on the former and fewer on the latter. Studies related to quality investment consider a single supply chain structure, ignoring the ability of suppliers to invest on their own, which is also inconsistent with the increasingly important role of suppliers in outsourcing. There is also the fact that the supply chain structure considered in quality investment is relatively homogenous, ignoring the ability of suppliers to invest on their own, which is also inconsistent with the increasingly important role of suppliers in outsourcing. Secondly, in reality, supply chain competition has long risen from inventory and price to higher levels such as quality and service, and it is more relevant to consider non-price competitive factors such as quality. Therefore, distinguishing from price competition, this study considers quality competition and aims to explore the non-price drivers in outsourcing using mathematical models.

3 The model

Consider two competing supply chains i ($i = 1, 2$), each consisting of a manufacturer (M_i) and a supplier (S_i) respectively. The duopoly manufacturers compete on output and quality sell products with certain substitutability and choose to either produce in-house or outsource to upstream suppliers, with both competing in terms of output and quality. We consider two outsourcing structures (shown in Figure 1): (a) a "one-to-one" structure, where a single supplier serves a single manufacturer; and (b) a "one-to-many" structure, where a common supplier serves two manufacturers.

Both outsourcing structures are reflected in practice [26–28]. As shown in the study by Feng et al. [29], in the computer industry, for example, Flextronics, as one of the largest contract manufacturers, also produces for Dell, Pratt & Whitney, etc., while another important contract manufacturer, Asahi Electric, produces for IBM and NEC. However, when Flextronics acquired Asahi Electric in 2007, manufacturers that used to compete with each other (e.g., HP and IBM) started to work with the same company.

3.1 Demand function

A linear inverse demand function is used to portray the quantity and quality competition of duopoly manufacturers [30, 31]: $p_i = a - q_i - bq_j + e_i$, $i, j = 1, 2$ and $i \neq j$. where p_i and q_i are the price and quantity of the manufacturer M_i, respectively, a denotes the basic demand in the market, e_i denotes the level of investment in quality, and b denotes the degree of product substitutability, with the product being fully independent at $b = 0$ and fully substitutable at $b = 1$.

3.2 Cost structure

Manufacturers and suppliers have certain technologies and resources to invest in product quality and promote market demand. In response to the law of diminishing marginal efficiency of investment, according to the literature [31, 32], the cost of quality investment $C(e)$ is a convex function of the form $C(e) = te^2/2$, where t is the marginal investment cost coefficient. Furthermore, referring to the literature [11, 14, 16], without loss of generality, it is assumed that the production costs of manufacturers and suppliers are c_m and c_s, and the quality investment efficiency is t_m and t_s, respectively.

Referring to the literature [29] for a convenient analysis, we make $t_s = 1$ and introduce the relative production cost advantage index $\delta = \frac{a - c_s}{a - c_m}$. Due to the investment burden, the supplier may no longer have an efficiency advantage, so we relax the range to $\delta > 0$ and the supplier has an absolute productivity advantage if and only if $\delta > 1$. In addition, to ensure the unique existence and non-negativity of the optimal solutions, it is assumed that $t_m > \frac{4}{(4 - b^2)^2}$ and $\frac{4bt_m(4 - b^2)}{8t_m(4 - b^2) - (8 - b^2)} < \delta < \frac{7 - b^2}{2b}$, and the upper and lower bounds of δ are notated as $\bar{\delta}$ and $\underline{\delta}$, respectively.

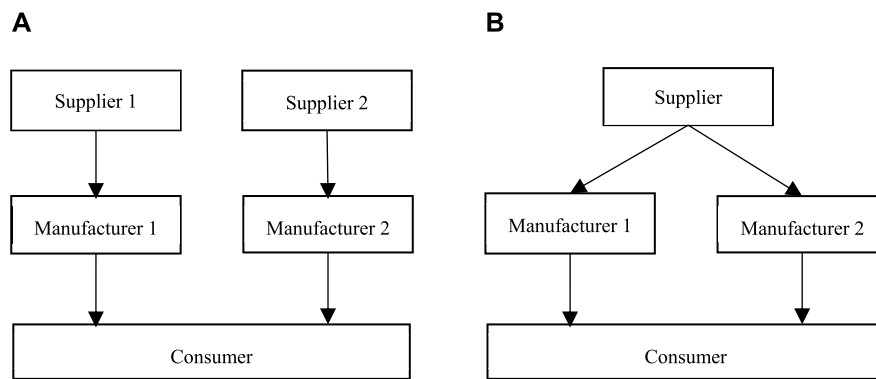Table 1 provides a summary of all the notations used in this paper.

FIGURE 1
Outsourcing structure. **(A)** "one-to-one" structure. **(B)** "one-to-many" structure.

TABLE 1 Summary of notation.

Notation	Explanation
a	Basic demand in the market
b	Degree of product substitutability
e_i	Level of product i's quality investment
w_i	Product i's wholesale price
c_s	Production costs of suppliers
c_m	Production costs of manufacturers
t_s	Quality investment efficiency of suppliers
t_m	Quality investment efficiency of manufacturers
δ	Relative production cost advantage index
p_i	Product i's retail price
q_i	Product i's demand quantity
π_{mi}	Manufacturer profit for product i
π_{si}	Supplier profit for product i

3.3 Game order

The sequence of the game considers two aspects: (a) Quality investment involves pre-preparation of technology, R&D, production lines, etc. Therefore, manufacturers must determine production outsourcing strategies in advance. (b) Quality investment determines the market positioning of a product, which is long-term and stable relative to pricing. Therefore, quality investment decisions come after outsourcing and before pricing. In addition, manufacturers can determine the competitor's production strategy by observing whether investing in a new production line, so they play a complete information dynamic game. The game sequence is as follows: firstly, the manufacturers determine the production strategy: self-production or outsourcing; secondly, the manufacturers and/or the suppliers announce the level of quality investment e_i; then the suppliers decide on the wholesale price w_i in case of outsourcing; and finally the manufacturers decide the product quantity q_i.

4 Analysis with two independent suppliers

In this section, we investigate the production strategies of duopoly manufacturers in a "one-to-one" structure, where each manufacturer has an independent supplier and chooses either self-production or outsourcing. Depending on manufacturers' choices, the market is characterized by four sub-game scenarios: (Self-production, Self-production), (Self-production, Outsourcing), (Outsourcing, Self-production), and (Outsourcing, Outsourcing), denoted by N and O for self-production and outsourcing, respectively.

4.1 Equilibrium solution

4.1.1 The NN scenario

In this scenario, both manufacturers choose the self-production strategy, i.e., the manufacturer autonomously decides to invest in product quality and pays for it. The manufacturers' profit functions are:

$$\pi_{mi}^{NN} = (p_i - c_m)q_i - \frac{1}{2}t_m e_i^2, i = 1, 2 \tag{1}$$

In the last stage of the game, the manufacturers decide their quantity decisions. According to the first-order conditional joint equations $\frac{\partial \pi_{m1}^{NN}}{\partial q_1} = 0$ and $\frac{\partial \pi_{m2}^{NN}}{\partial q_2} = 0$, their optimal quantity decisions are obtained as follows:

$$q_i = \frac{(2-b)(a-c_m) + 2e_i - be_j}{4 - b^2} \tag{2}$$

For the manufacturers' quality investment decisions, Eq. 2 is brought into Eq. 1 with the joint equations $\frac{\partial \pi_{m1}^{NN}}{\partial e_1} = 0$ and $\frac{\partial \pi_{m2}^{NN}}{\partial e_2} = 0$ to obtain the manufacturers' optimal levels of quality investment $e_i^{NN'}$. Table 2 lists all the optimal decisions of the manufacturers.

4.1.2 The NO/ON scenario

In these two scenarios, one manufacturer produces in-house and the other outsources. Without loss of generality, we take the NO scenario as an example, where manufacturer M_1 produces in-house and manufacturer M_2 outsources. Then manufacturer M_1 and

TABLE 2 Optimal decisions with independent suppliers.

	NN	ON/NO	OO
e_i^*	$\frac{2(a-c_s)}{t(2-b)(2+b)^2-2}$	$e_1^{ON'} = e_2^{NO*} = \frac{(a-c_s)[(8-b^2)\delta+4t(4-b^2)(b-2\delta)]}{(8-b^2)[t_s(8-b^2)-1]-8t(4-b^2)[2(4-b^2)t_s-1]}$	$\frac{2(8-b^2)(a-c_s)}{(4-b)^2(2+b)(4+b)t_s-2(8-b^2)}$
		$e_2^{ON*} = e_1^{NO*} = \frac{(a-c_s)(8-b^2)[1-(8-b^2-2b\delta)t_s]}{(8-b^2)[t_s(8-b^2)-1]-8t(4-b^2)[2(4-b^2)t_s-1]}$	
w_i^*	—	$w_1^{ON*} = w_2^{NO*} = (4-b^2)t_s e_1^{ON*} + c_s$	$\frac{(64-20b^2+b^4)t_s}{2(8-b^2)}e_i^{OOC*} + c_s$
q_i^*	$\frac{t}{2}e_i^{NN*}$	$q_1^{ON*} = q_2^{NO*} = 2t_s e_1^{ON*}$	$\frac{(16-b^2)t_s}{8-b^2}e_i^{OOC*}$
		$q_2^{ON*} = q_1^{NO*} = \frac{4t(4-b^2)}{8-b^2}e_2^{ON*}$	
π_{mi}^*	$\frac{t(4-b^2)^2-4}{t(4-b^2)^2}[q_i^{NN*}]^2$	$\pi_1^{ON*} = \pi_2^{NO*} = [q_1^{ON*}]^2$	$[q_i^{OOD*}]^2$
		$\pi_2^{ON*} = \pi_1^{NO*} = \frac{16t(4-b^2)^2-(8-b^2)^2}{16t(4-b^2)^2}[q_2^{ON*}]^2$	
π_{si}^*	0	$\pi_{s1}^{ON*} = \pi_{s2}^{NO*} = \frac{2t_s(4-b^2)-1}{4t_s}[q_1^{ON*}]^2$	$\frac{t_s(16-b^2)^2(4-b^2)-2(8-b^2)^2}{2t_s(16-b^2)^2}[q_i^{OOD*}]^2$

supplier S_2 decide independently the quality investment and pay investment cost. In addition, supplier S_2 also determines the outsourcing wholesale price w_2. The profit functions of the manufacturers and supplier are respectively:

$$\pi_{m1}^{NO} = (p_1 - c_m)q_1 - \frac{1}{2}t_m e_1^2 \quad (3)$$

$$\pi_{m2}^{NO} = (p_2 - w_2)q_2 \quad (4)$$

$$\pi_{s2}^{NO} = (w_2 - c_s)q_2 - \frac{1}{2}t_s e_2^2 \quad (5)$$

In the last stage of the game, the manufacturers decide their quantity decisions. According to the first-order conditional joint equations $\frac{\partial \pi_{m1}^{NN}}{\partial e_1} = 0$ and $\frac{\partial \pi_{m2}^{NN}}{\partial e_2} = 0$, their optimal quantity decisions are obtained as follows:

$$q_1 = \frac{a(2-b) - 2c_m + bw_2 + 2e_1 - be_2}{4-b^2} \quad (6)$$

$$q_2 = \frac{a(2-b) + bc_m - 2w_2 - be_1 + 2e_2}{4-b^2} \quad (7)$$

Next, we consider the supplier's wholesale price decision. By bringing Eqs 6, 7 into Eq. 5, we have the first-order condition $\frac{\partial \pi_{s2}^{NO}}{\partial w_2} = 0$, and the supplier's optimal wholesale price is obtained as:

$$w_2 = \frac{a(2-b) - b(c_m - e_1) - 2(c_s + e_2)}{4} \quad (8)$$

For the quality investment decisions, Eqs 6–8 are sequentially substituted back into Eqs 3–5. Then according to $\frac{\partial \pi_{m1}^{NO}}{\partial e_1} = 0$ and $\frac{\partial \pi_{s2}^{NO}}{\partial e_2} = 0$, the optimal levels of quality investment $e_i^{NO'}$ are obtained. The optimal decisions and profits for the manufacturers and the supplier are shown in Table 2.

4.1.3 The OO scenario

In this scenario, both manufacturers outsource, and the corresponding suppliers are responsible not only for manufacturing but also for quality investments. The profit functions of the manufacturers and suppliers are respectively:

$$\pi_{mi}^{OO} = (p_i - c_m)q_i \quad (9)$$

$$\pi_{si}^{OO} = (w_i - c_s)q_i - \frac{1}{2}t_s e_i^2 \quad (10)$$

In the last stage of the game, the manufacturers decide their quantity decisions. According to the first-order conditional joint equations $\frac{\partial \pi_{m1}^{OO}}{\partial q_1} = 0$ and $\frac{\partial \pi_{m2}^{OO}}{\partial q_2} = 0$, their optimal quantity decisions are obtained as follows:

$$q_i = \frac{a(2-b) - 2e_i + be_j + 2w_i - bw_j}{4-b^2} \quad (11)$$

Next, we bring Eq. 11 into Eq. 10, and obtain the first-order conditions $\frac{\partial \pi_{s1}^{OO}}{\partial w_1} = 0$ and $\frac{\partial \pi_{s2}^{OO}}{\partial w_2} = 0$. Then the suppliers' optimal wholesale price decisions are given as follows:

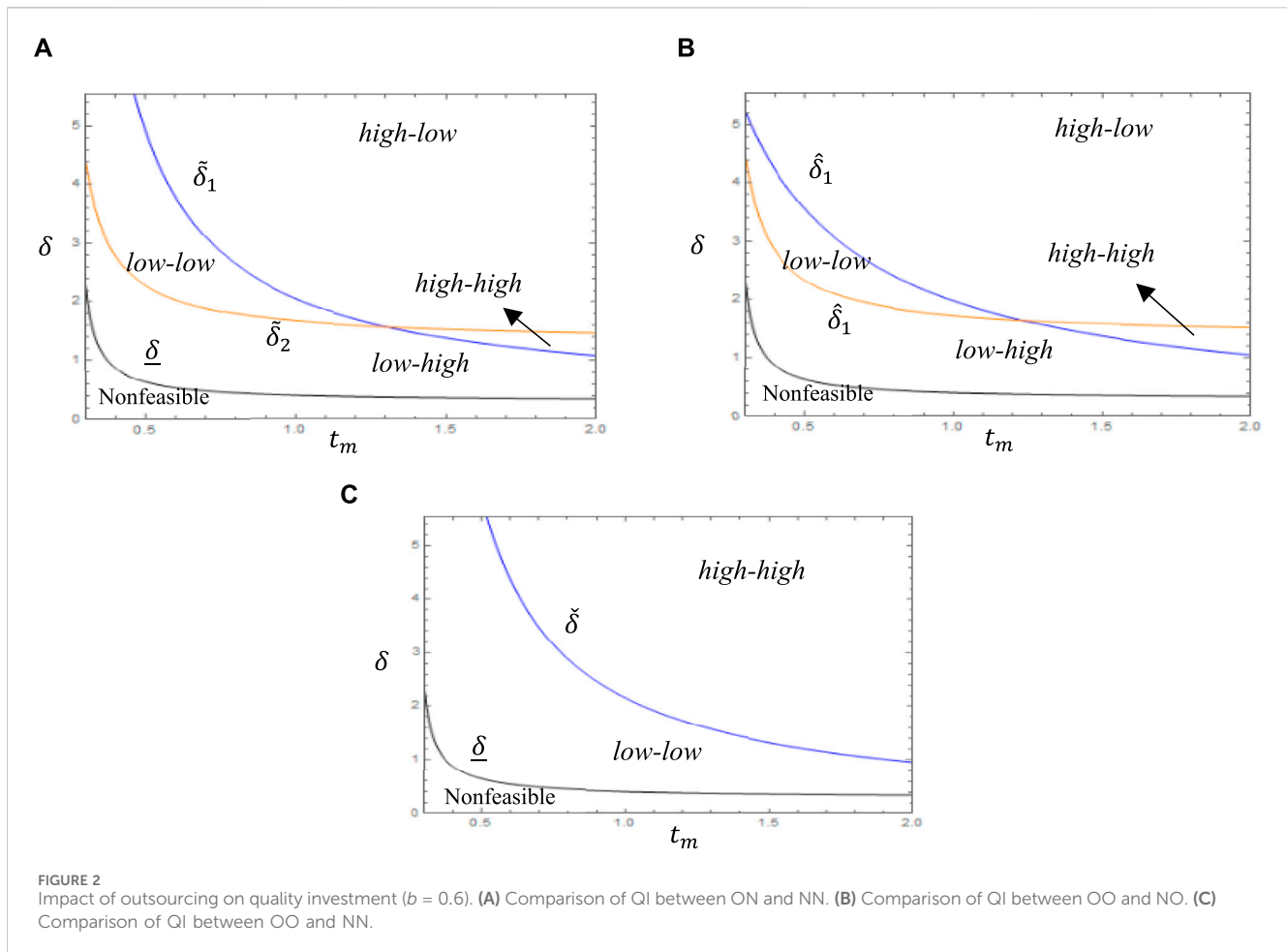$$w_i = \frac{a(2-b)(4+b) + 2(4+b)c_s + (8-b^2)e_i - 2be_j}{16-b^2} \quad (12)$$

For the quality investment decisions, Eqs 11, 12 are substituted back into Eq. 10. Then according to $\frac{\partial \pi_{si}^{OO}}{\partial e_1} = 0$ and $\frac{\partial \pi_{s2}^{OO}}{\partial e_2} = 0$, the optimal levels of quality investment $e_i^{OO'}$ are obtained. Substituting the above results into Eq. 9 can obtain the optimal solutions. The optimal decisions and profits for the manufacturers and suppliers are shown in Table 2.

Lemma 1. As δ decreases or t_m increases, the outsourced manufacturer has higher quality, more quantity and higher profit, while the self-produced manufacturer has lower quality, less output, and lower profit.

Lemma 1 shows that a decrease in the supplier's production cost increases the production advantage of outsourcing, while a decrease in the efficiency of the manufacturer's quality investment increases the investment advantage of outsourcing. Faced with either a production advantage or a quality advantage, the supplier will increase quality investment and reduce wholesale price, and the outsourced manufacturer will therefore increase quantity to realize higher profit. However, as a result of competition, the self-produced manufacturer will invest less in quality and produce less, leading to lower profit.

4.2 Comparative analysis of equilibrium

In this section, we compare the equilibrium results of the four sub-game models described above and analyze the impacts of different production strategies on quality investment and profit,

FIGURE 2
Impact of outsourcing on quality investment (b = 0.6). **(A)** Comparison of QI between ON and NN. **(B)** Comparison of QI between OO and NO. **(C)** Comparison of QI between OO and NN.

as well as the choice of the duopoly manufacturers between the two production strategies.

In order to analyze the effect of outsourcing on quality investment, we hold one manufacturer's production strategy constant and compare the quality investment of another manufacturer's self-production and outsourcing. The results are summarized in Theorem 1 and Figure 2.

Theorem 1. A comparison of the quality investment for the four scenarios is as follows:

Case 1. There exist two thresholds $\tilde{\delta}_1$ and $\tilde{\delta}_2$, such that
(i) $e_1^{ON*} > e_1^{NN*}$ if $\tilde{\delta}_1 < \delta < \bar{\delta}$; $e_1^{ON*} < e_1^{NN*}$ if $\underline{\delta} < \delta < \min\{\tilde{\delta}_1, \bar{\delta}\}$.
(ii) $e_2^{ON*} > e_2^{NN*}$ if $\underline{\delta} < \delta < \tilde{\delta}_2$; $e_2^{ON*} < e_2^{NN*}$ if $\tilde{\delta}_2 < \delta < \bar{\delta}$.
Case 2. There exist two thresholds $\hat{\delta}_1$ and $\hat{\delta}_2$, such that
(i) $e_1^{OO*} > e_1^{NO*}$ if $\hat{\delta}_1 < \delta < \bar{\delta}$; $e_1^{OO*} < e_1^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}_1$.
(ii) $e_2^{OO*} > e_2^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}_2$; $e_2^{OO*} < e_2^{NO*}$ if $\hat{\delta}_2 < \delta < \bar{\delta}$.
Case 3. There exist a threshold $\check{\delta}$, such that
$e_i^{OO'} > e_i^{NN'}$ if $\check{\delta} < \delta < \bar{\delta}$; $e_i^{OO'} < e_i^{NN'}$ if $\underline{\delta} < \delta < \min\{\check{\delta}, \bar{\delta}\}$.

Theorem 1 shows that outsourcing does not necessarily increase quality investment. Outsourcing has two effects on quality. On the one hand, the supplier with lower production cost or more efficient investment has an incentive to invest in higher quality; on the other hand, outsourcing makes the manufacturer more dependent on their supplier, and a decentralized channel structure reduces the manufacturer's

control over quality. Thus, outsourcing can improve quality only if the positive effects of low cost or high efficiency are higher than the negative effects of system decentralization. Specifically, given that one manufacturer self-produces, it holds the advantage of system centralization. If the other manufacturer's quality investment is more efficient, then outsourcing results in a higher loss of quality due to system decentralization. Combined with market competition, outsourcing reduces the manufacturer's quality. Conversely, given that one manufacturer outsources, then the other manufacturer's outsourcing exposes both parties to the same system decentralization. So even if the manufacturer's quality investment is more efficient, outsourcing will increase the manufacturer's quality as long as the supplier's relative cost is low. In addition, outsourcing may reduce the rival's quality investment if the supplier's relative cost is low, as market competition increases the pressure on rival to invest.

Next, We compare manufacturers' profits provided that only one party changes its production strategy, showing as Theorem 2 and Figure 3.

Theorem 2. A comparison of the profits for the four scenarios is as follows:

Case 1. There exist two thresholds $\tilde{\delta}_1^*$ and $\tilde{\delta}_2^*$, such that
(i) $\pi_{m1}^{ON*} > \pi_{m1}^{NN*}$ if $\tilde{\delta}_1^* < \delta < \bar{\delta}$; $\pi_{m1}^{ON*} < \pi_{m1}^{NN*}$ if $\underline{\delta} < \delta < \tilde{\delta}_1^*$.
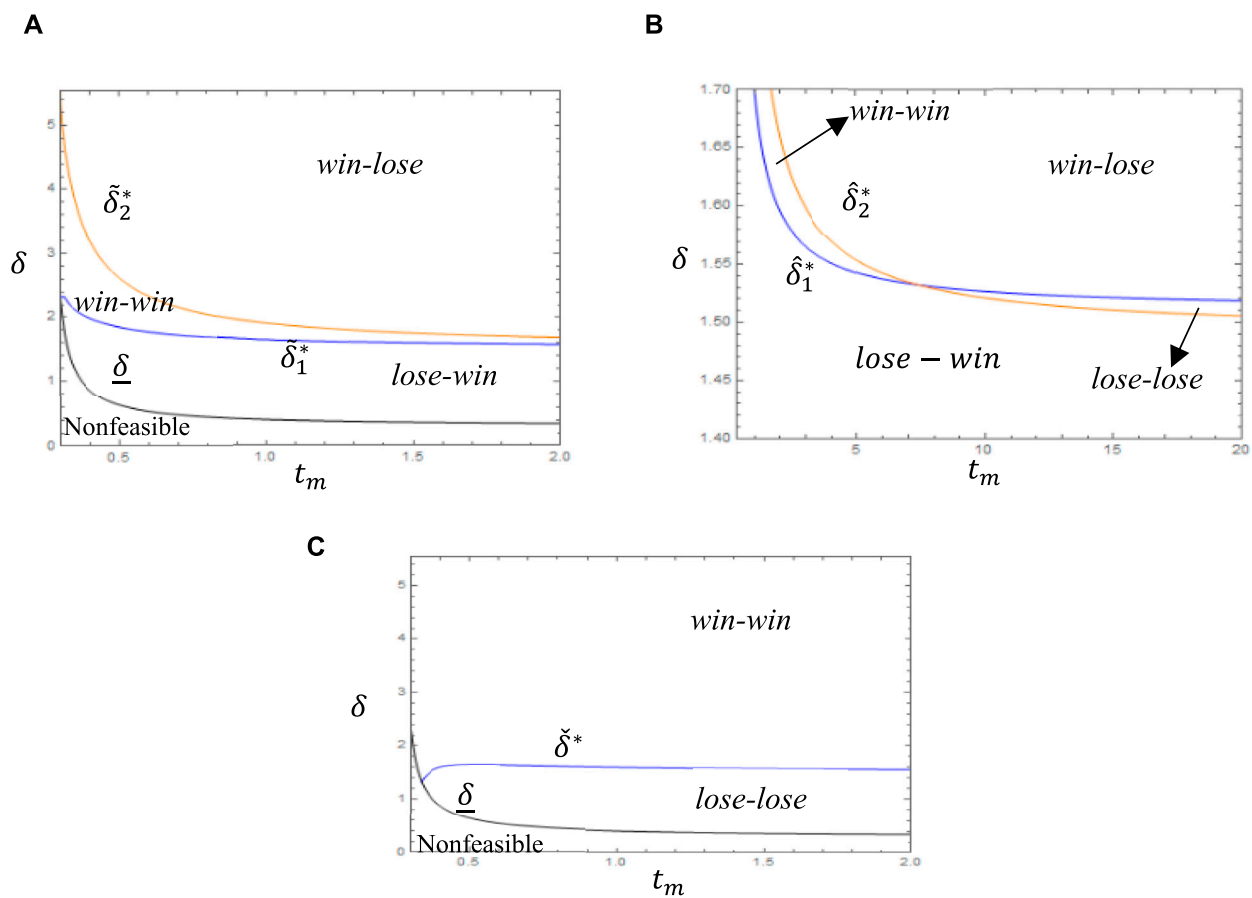
FIGURE 3
Impact of outsourcing on manufacturer profit ($b = 0.6$). **(A)** Comparison of profit between ON and NN. **(B)** Comparison of profit between OO and NO. **(C)** Comparison of profit between OO and NN.

(ii) $\pi_{m2}^{ON*} > \pi_{m2}^{NN*}$ if $\underline{\delta} < \delta < \tilde{\delta}_2^*$; $\pi_{m2}^{ON*} < \pi_{m2}^{NN*}$ if $\tilde{\delta}_2^* < \delta < \bar{\delta}$.

Case 2. There exist two thresholds $\hat{\delta}_1^*$ and $\hat{\delta}_2^*$, such that
(i) $\pi_{m1}^{OO*} > \pi_{m1}^{NO*}$ if $\hat{\delta}_1^* < \delta < \bar{\delta}$; $\pi_{m1}^{OO*} < \pi_{m1}^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}_1^*$.
(ii) $\pi_{m2}^{OO*} > \pi_{m2}^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}_2^*$; $\pi_{m2}^{OO*} < \pi_{m2}^{NO*}$ if $\hat{\delta}_2^* < \delta < \bar{\delta}$.

Case 3. There exist a threshold δ^*, such that
$\pi_{mi}^{OO*} > \pi_{mi}^{NN*}$ if $\max\left\{\underline{\delta}, \delta^*\right\} < \delta < \bar{\delta}$; $\pi_{mi}^{OO*} < \pi_{mi}^{NN*}$ if $\underline{\delta} < \delta < \delta^*$.
Among them, $\tilde{\delta}_1^* > 1$, $\hat{\delta}_1^* > 1$, $\delta^* > 1$, $\tilde{\delta}_2^* > \tilde{\delta}_1^*$.

Theorem 2 shows that a necessary condition for a manufacturer to outsource is that the supplier has an absolute cost advantage. When δ is small, outsourcing empties the manufacturer's investment cost but leads to double marginalization, and the loss of system efficiency dominates the decline in the manufacturer's profit. When δ is large enough, however, significant improvements in supplier productivity compensate for the loss of system efficiency, so outsourcing is more favorable to the manufacturer.

The results of Lemma 1 suggest that the supplier's lower production cost and higher investment efficiency have a boosting effect on the outsourced manufacturer and a dampening effect on the self-produced manufacturer, so whether outsourcing can achieve a win-win situation for the duopoly manufacturers depends not only on relative production cost and investment efficiency, but is also closely related to the rival's production strategy. Given that one manufacturer self-produces, outsourcing is win-win only if the supplier's relative cost of production is moderate, regardless of investment efficiency. Given

that one manufacturer outsources, outsourcing is win-win only if the manufacturer's quality investment is efficient and the supplier's relative cost of production is moderate ($\hat{\delta}_1^* < \hat{\delta}_2^*$).

Finally, we analyze the optimal production strategy of the duopoly manufacturers and solve the equilibrium in the first stage of the game based on the payment matrix (see Table 3).
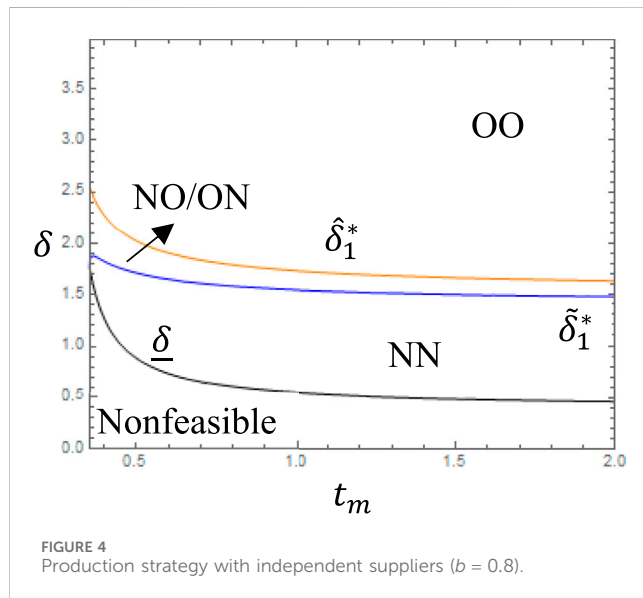
Due to the symmetry of the model, we can obtain the optimal production strategy of the duopoly manufacturers by comparing the positives and negatives of $\pi_{m1}^{NN} - \pi_{m1}^{ON}$ and $\pi_{m1}^{OO} - \pi_{m1}^{NO}$, respectively. It is summarized as Theorem 3 and Figure 4.

Theorem 3. The optimal production strategy of duopoly manufacturers with two independent suppliers is: (i) if $\underline{\delta} < \delta < \tilde{\delta}_1^*$, two manufacturers both choose self-production; (ii) if $\tilde{\delta}_1^* < \delta < \hat{\delta}_1^*$, one manufacturer choose self-production and the other manufacturer choose outsourcing; (iii) if $\hat{\delta}_1^* < \delta < \bar{\delta}$, two manufacturers both choose outsourcing.

Theorem 3 shows that the relative production cost of manufacturers and suppliers dominate the equilibrium of production strategy. The lower cost advantage makes manufacturers prefer self-production, and as the cost advantage increases, the attractiveness of outsourcing to manufacturers increases and one of them will favor outsourcing. The reason is that, for manufacturers, if the supplier's production cost is low, both

TABLE 3 Game payment matrix of duopoly manufacturers with independent suppliers.

Manufacturer M_1/Manufacturer M_2	Self-production (N)	Outsourcing (O)
Self-production (N)	$(\pi_{m1}^{NN}, \pi_{m2}^{NN})$	$(\pi_{m1}^{NO}, \pi_{m2}^{NO})$
Outsourcing (O)	$(\pi_{m1}^{ON}, \pi_{m2}^{ON})$	$(\pi_{m1}^{OO}, \pi_{m2}^{OO})$

FIGURE 4
Production strategy with independent suppliers ($b = 0.8$).

TABLE 4 Optimal decisions with a common supplier.

	OOC
$e^\star$	$\frac{k(a-c_s)}{2t_s(2+b)-k^2}$
w_i^*	$\frac{t_s(2+b)(a-c_s)}{2t_s(2+b)-k^2} + c_s$
q_i^*	$\frac{(a-c_s)t_s}{2t_s(2+b)-k^2}$
p_i^*	$\frac{t_s(3+b)(a-c_s)}{2t_s(2+b)-k^2} + c_s$
π_{mi}^*	$[q_i^{OO*}]^2$
π_s^*	$(a - c_s)q_i^{OO*}$

manufactures can enjoy the benefits of investment cost transfer and production cost saving through outsourcing. However, if the supplier's production cost is not low enough, both manufacturers outsource to make the product homogenization competition more intense, and the effect of outsourcing benefits are weakened, so the differentiated production strategy is more favorable to them.

Theorem 2 and Theorem 3 together show that outsourcing can be an equilibrium strategy for duopoly manufacturers under certain conditions. Outsourcing increases suppliers' profits, which can either increase or decrease manufacturers' profits. In short, only if outsourcing can increase the profits of the duopoly manufacturers at the same time, they will realize a win-win situation, otherwise outsourcing puts the duopoly manufacturers in a prisoner's dilemma.

5 Analysis with a common supplier

In this section, we study production strategy in a "one-to-many" structure, where the duopoly manufacturers choose to self-produce or outsource to a common supplier. There are four sub-game scenarios based on the manufacturer's choice: (Self-production, Self-production), (Self-production, Outsourcing), (Outsourcing, Self-production), and (Outsourcing, Outsourcing). Among them, the first three sub-game scenarios are the same as those in Section 4. To avoid confusion, we refer to (Self-production, Outsourcing) and (Outsourcing, Self-production) collectively as asymmetric outsourcing, outsourcing to two independent suppliers (OO scenario) as decentralized outsourcing, and outsourcing to a common supplier (OOC scenario) as centralized outsourcing.

5.1 Equilibrium solution

In this section, both manufacturers outsource to the same supplier. The supplier makes a one-time quality investment to provide equal quality to the duopoly manufacturers. Denoting this scenario by superscript OOC, the profit functions of the manufacturers and supplier are, respectively:

$$\pi_{mi}^{OOC} = (p_i - c_m)q_i \tag{13}$$

$$\pi_s^{OOC} = (w_i - c_s)q_i - \frac{1}{2}t_s e^2 \tag{14}$$

In the last stage of the game, the manufacturers decide their quantity decisions. According to the first-order conditional joint equations $\frac{\partial \pi_{m1}^{OOC}}{\partial q_1} = 0$ and $\frac{\partial \pi_{m2}^{OOC}}{\partial q_2} = 0$, their optimal quantity decisions are obtained as follows:

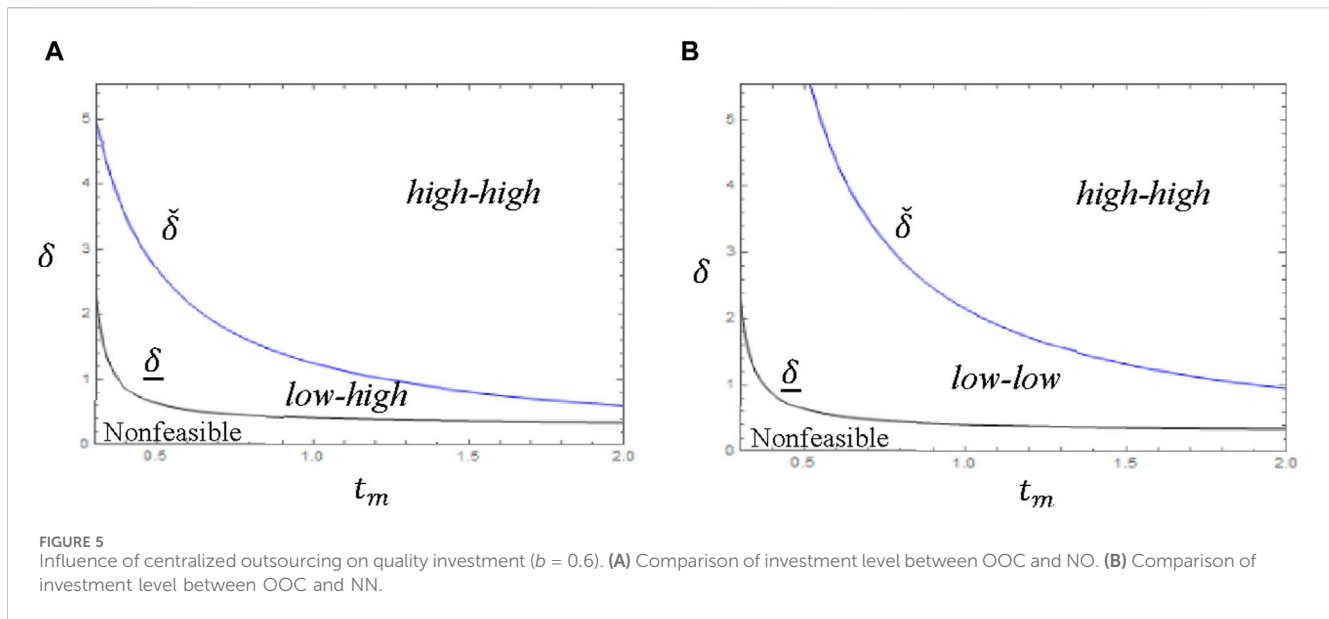$$q_i = \frac{(a+e)(2-b) - 2w_i + bw_j}{4 - b^2} \tag{15}$$

Next, we bring Eq. 15 into Eq. 14, and obtain the first-order condition $\frac{\partial \pi_s^{OOC}}{\partial w_i} = 0$. Then the suppliers' optimal wholesale price decisions are given as follows:

$$w_i = \frac{a + c_s + e}{2} \tag{16}$$

For the quality investment decisions, Eqs 15, 16 are substituted back into Eq. 14. Then according to $\frac{\partial \pi_s^{OOC}}{\partial e} = 0$, the optimal levels of quality investment $e_i^{OOC'}$ are obtained. Substituting the above results into Eq. 13 can obtain the optimal solutions. The optimal decisions and profits for the manufacturers and suppliers are shown in Table 4.

Lemma 2. (i) $e^{OOC*} > e_i^{OO*}$; (ii) $\pi_{mi}^{OOC*} > \pi_{mi}^{OO*}$ if $0 < b < b^\star$; $\pi_{mi}^{OOC} < \pi_{mi}^{OO*}$, if $b^\star < b < 1$; (iii) $\pi_{si}^{OOC*} > \pi_{si}^{OO*}$.

Lemma 2 compares the quality and profitability of two manufacturers when they both outsource to independent suppliers and common supplier. Obviously, regardless of the

FIGURE 5
Influence of centralized outsourcing on quality investment ($b = 0.6$). **(A)** Comparison of investment level between OOC and NO. **(B)** Comparison of investment level between OOC and NN.

production cost and the efficiency of quality investments, a common supplier uses centralized outsourcing to avoid duplication of investment and competition for quality, to increase the product quality with consolidation advantages, and to achieve a higher profit. However, the manufacturers may be suffer from higher product quality. This is because if products are more substitutable and more competitive, the prices increase for high-quality products may result in lower demand and thus lower profits for the duopoly manufacturers.

5.2 Comparative analysis of equilibrium

Next we analyze the impact of centralized outsourcing on product quality and profit, and compare the choices of the duopoly manufacturers between self-production, decentralized outsourcing and centralized outsourcing.

To analyze the effect of centralized outsourcing on the quality investment, we keep one manufacturer's production strategy unchanged and compare the quality investment under self-production and centralized outsourcing of another manufacturer, obtaining Theorem 4 and Figure 5.

Theorem 4. A comparison of the quality investment for the four scenarios is as follows:

Case 1. There exist a threshold $\hat{\delta}^C$, such that
(i) $e^{OOC*} > e_1^{NO*}$ if $\hat{\delta}^C < \delta < \bar{\delta}$; $e^{OOC*} < e_1^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}^C$.
(ii) $e^{OOC*} > e_2^{NO*}$.

Case 2. There exist a threshold $\underline{\delta}^C$, such that
$e^{OOC*} > e_i^{NN*}$ if $\max\left\{\underline{\delta}, \underline{\delta}^C\right\} < \delta < \bar{\delta}$; $e^{OOC*} < e_i^{NN*}$ if $\underline{\delta} < \delta < \min\left\{\overline{\delta}^C, \bar{\delta}\right\}$.

Theorem 4 shows that if one manufacturer outsources, then another manufacturer's choice of centralized outsourcing over self-production raises the competitor's quality investment. This is because centralized outsourcing gives the supplier more investment capital and avoids investment competition, and the consolidation effect and scale advantage of quality investment

promotes the outsourced product's quality. Similarly, the quality of the self-produced manufacturer can only improve from outsourcing if the supplier's relative production cost is below a certain threshold to compensate for the loss of investment due to double marginalization.

Next, We compare manufacturers' profits provided that only one party changes its production strategy, showing as Theorem 5 and Figure 6.

Theorem 5. A comparison of the profits for the four scenarios is as follows:

Case 1. There exist two thresholds $\hat{\delta}_1^{C*}$ and $\hat{\delta}_2^{C*}$, such that
(i) $\pi_1^{OOC*} > \pi_1^{NO*}$ if $\hat{\delta}_1^{C*} < \delta < \bar{\delta}$; $\pi_1^{OOC*} < \pi_1^{NO*}$ if $\underline{\delta} < \delta < \hat{\delta}_1^{C*}$.
(ii) $\pi_2^{OOC*} > \pi_2^{NO*}$ if $\underline{\delta} < \delta < \min\left\{\hat{\delta}_2^{C*}, \bar{\delta}\right\}$; $\overline{\delta}_2^{C*} < \delta < \bar{\delta}$ if $\pi_2^{OOC*} < \pi_2^{NO*}$.

Case 2. There exist a threshold $\underline{\delta}^{C*}$, such that
$\pi_i^{OOC*} > \pi_i^{NN*}$ when $\max\left\{\underline{\delta}, \underline{\delta}^{C*}\right\} < \delta < \bar{\delta}$; $\pi_i^{OOC*} < \pi_i^{NN*}$ when $\underline{\delta} < \delta < \overline{\delta}^{C*}$.
Where, $\hat{\delta}_1^{C*} > 1$, $\underline{\delta}^{C*} > 1$.

Theorem 5 shows that the self-produced manufacturer chooses centralized outsourcing only if the δ is sufficiently large. In addition, when the competitor outsource, if the degree of competition in the market is small, the product quality improvement brought by centralized outsourcing will unconditionally increase the outsourced manufacturer's profit; if the degree of competition in the market is large, the undifferentiated quality competition undermines the advantage of quality improvement. Therefore, centralized outsourcing can be a win-win situation only if the production cost advantage index is in the moderate range.

Finally, the optimal production strategy of the duopoly manufacturers is considered, and the equilibrium in the first stage of the game is solved based on the payment matrix (see Table 5), summarized in Theorem 6 and Figure 7.

Theorem 6. The optimal production strategy of duopoly manufacturers with a common supplier is: (i) if $\underline{\delta} < \delta < \min\left\{\tilde{\delta}_1^*, \hat{\delta}_1^{C*}\right\}$, two manufacturers both choose self-

FIGURE 6
Impact of outsourcing on manufacturer profit. **(A)** Profit comparison of OOC and NO ($b = 0.2$). **(B)** Profit comparison of OOC and NO ($b = 0.6$). **(C)** Profit comparison of OOC and NN ($b = 0.6$).
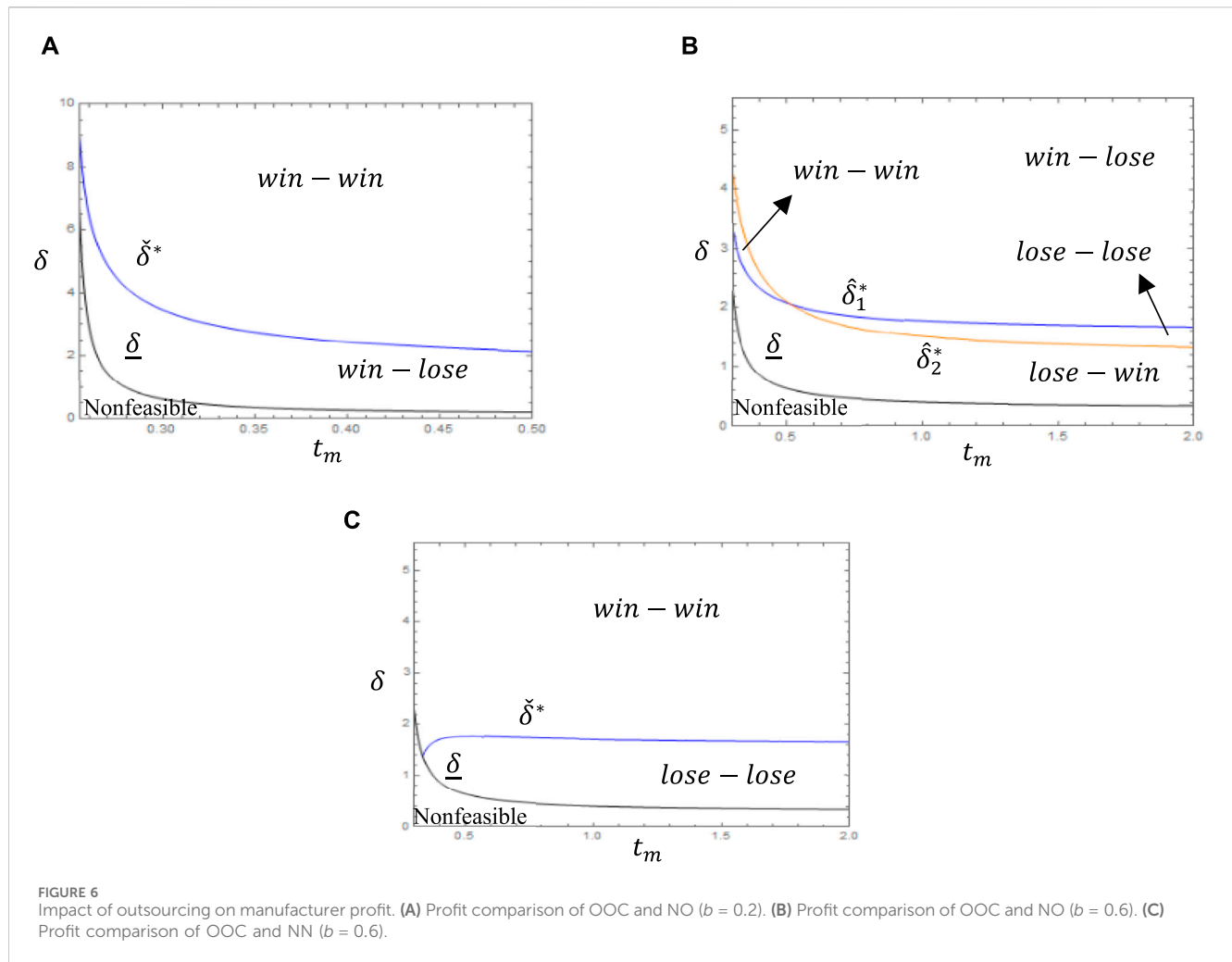
TABLE 5 Game payment matrix of duopoly manufacturers with a common supplier.

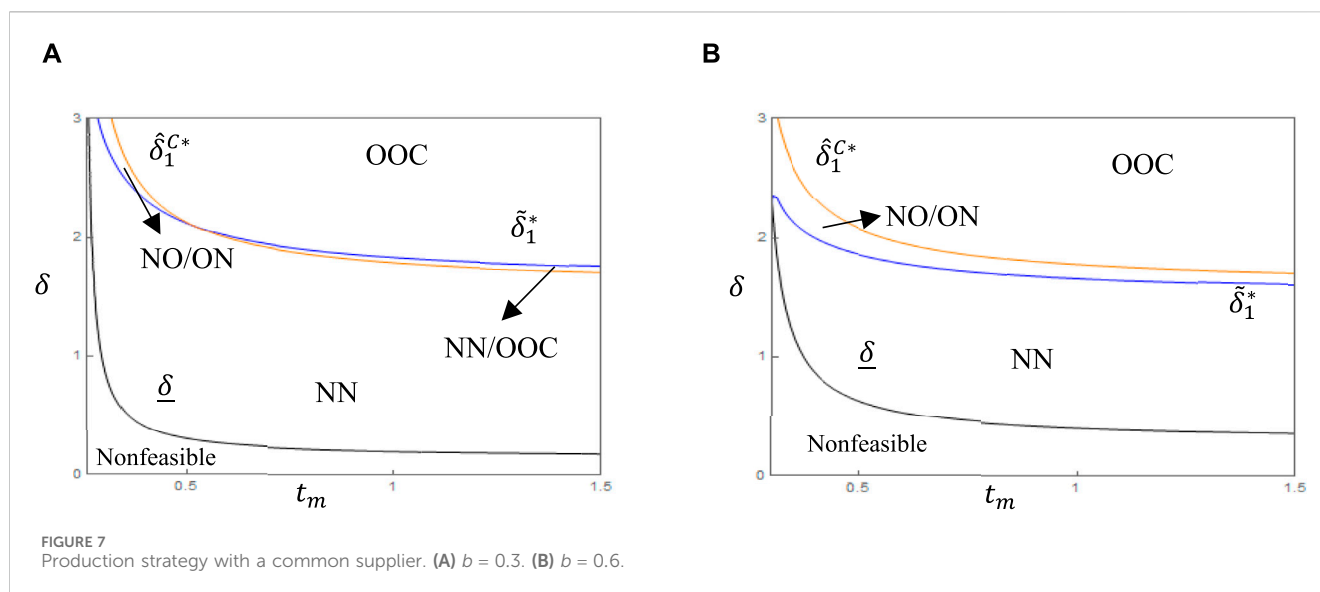| Manufacturer M_1/Manufacturer M_2	Self-manufacture (N)	Outsourcing (O)
Self-manufacture (N)	$(\pi_{m1}^{NN}, \pi_{m2}^{NN})$	$(\pi_{m1}^{NO}, \pi_{m2}^{NO})$
Outsourcing (O)	$(\pi_{m1}^{ON}, \pi_{m2}^{ON})$	$(\pi_{m1}^{OOC}, \pi_{m2}^{OOC})$

production; (ii) if $\max\left\{\tilde{\delta}_1^*, \hat{\delta}_1^{C*}\right\} < \delta < \bar{\delta}$, two manufacturers both choose outsourcing; (iii) if $\tilde{\delta}_1^* < \delta < \hat{\delta}_1^{C*}$, one manufacturer choose self-production and the other manufacturer choose outsourcing; (iv) if $\hat{\delta}_1^{C*} < \delta < \tilde{\delta}_1^*$, two manufacturers both choose self-production or outsourcing.

Theorem 6 and Theorem 3 show that a mixed-strategy Nash equilibrium may occur for two manufacturers when δ is moderate. Specifically, asymmetric outsourcing improves product quality variation, and centralized outsourcing or in-house manufacturing improves investment or production efficiency. When the market is not competitive and the manufacturer's investment in quality is inefficient (satisfying $\hat{\delta}_1^{C*} < \tilde{\delta}_1^*$), the incentive of the duopoly manufacturers to pursue efficiency gains dominates, and thus outsourcing on one side and self-production on the other side is no longer a stable equilibrium.

The above findings provide a new theoretical perspective on outsourcing strategies in the chip industry. As the mass production advantages of specialised chip foundries (e.g., TSMC) are exploited and advanced technological processes are improved, the tendency to outsource to the same foundries will gradually increase, and an oligopolistic competitive pattern will emerge in the foundry sector. With the aim of pursuing quality improvement, manufacturers' outsourcing also takes into account factors such as suppressing market competition and exploiting the advantages of intensification.

6 Conclusion

Quality investments are increasingly important in outsourcing. We investigate whether duopoly manufacturers will outsource to a supplier

FIGURE 7
Production strategy with a common supplier. **(A)** $b = 0.3$. **(B)** $b = 0.6$.

that has the ability to invest on its own, and compare the differences in the manufacturers' production strategies between two outsourcing structures: independent suppliers and common supplier.

The study has the following findings. First, two manufacturers choosing to outsource at the same time may be caught in a prisoner's dilemma. A necessary condition for outsourcing to be beneficial to the manufacturers is the supplier's strong production cost advantage, and even if the supplier's investment in quality is more efficient, outsourcing may reduce the quality of the product due to the risk of systemic decentralization. Secondly, compared to outsourcing to independent suppliers, duopoly manufacturers outsourcing to a common supplier can take advantage of the centralizing effect of outsourcing to improve quality and, when competition in the market is low, to improve profits. In addition, when the competitor outsource, the manufacturer's choice of a more productive independent supplier reduces the competitor's quality and its choice of a common supplier increases the competitor's quality. Finally, the equilibrium of production strategies of duopoly suppliers is closely related to the outsourcing structure, supplier production efficiency and investment efficiency. In the independent supplier structure, the supplier's productivity dominates the market equilibrium. If the supplier's productivity is low, then both manufacturer produce itself; if the supplier's productivity is moderate, then one manufacturer produces itself and the other outsources; if the supplier's productivity is high, then both manufacturers outsource. In a common supplier structure with moderate supplier production efficiency, lower market competition intensity and higher supplier quality investment efficiency may lead to a mixed strategy Nash equilibrium in which both manufacturers produce themselves or both outsource.

The findings of this research can provide some insights for foundry suppliers and manufacturing companies. When it comes to delivering high quality products to the market, the urgency for manufacturers to improve quality is at odds with their own inadequate level of quality improvement, and partnering with contract manufacturers who have the capacity to invest in quality is a viable solution to such problems. Therefore, how to fully exploit the capacity of suppliers to invest in the quality of their products to improve their competitive advantage is a key concern for manufacturing companies. Firstly, foundry suppliers can

increase the attractiveness of production outsourcing to manufacturing companies by improving production efficiency and investment efficiency in two ways, such as improving production processes to reduce production costs or improving production technology to increase investment efficiency. Foundry suppliers should actively improve their own strength, strengthen the centralization effect of outsourcing, and enhance the core advantages of outsourcing. Second, manufacturing enterprises choose to outsource does not necessarily enhance the competitive advantage, both sides choose to outsource instead may be due to product homogenization to intensify market competition, and then fall into the prisoner's dilemma. When competition in the market is weak, manufacturers should not rule out outsourcing to their competitors. In addition, they should also continue to improve their own capabilities to gain competitive advantage and, if necessary, reduce dependence on foundry suppliers. In conclusion, the choice of suppliers with independent investment capacity of outsourcing production can bring new opportunities to the product quality and profitability of manufacturing enterprises. Manufacturing enterprises should pay close attention to the foundry supplier production investment factors, but also to keep abreast of competitors' production decisions, pay attention to market competition and outsourcing structure changes, according to the strategic objectives of the enterprise in the choice of production strategy to balance the product quality and profitability.

The research in this paper has some limitations. First, due to the complexity of the parameters and the limitations of the analyses, this paper only considers the game models of two manufacturers and two suppliers with a small sample size. The limited case studies may affect the generality of the conclusions. In the future, other methods of analysis, such as simulations, may be considered to increase the size of the sample and to test the robustness of the results. Second, the model only addresses the issue of selecting production methods for two symmetric manufacturers or suppliers with comparable productivity and investment power. Future considerations may include asymmetric cases with disparities in power to better reflect real-world scenarios. Finally, the research only considers fixed production costs and ignores investment spillover effects. Including these factors may enrich the results. Future research could expand the cost structure and consider additional production and investment influences in the game model.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YW: Writing–original draft. TD: Formal Analysis, Writing–review and editing. HC: Visualization, Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Charles M, Ochieng SB. Strategic outsourcing and firm performance: a review of literature. *Int J Soc Sci Humanities Res (Ijsshr) ISSN 2959-7056 (O) 2959-7048 (P)* (2023) 1(1):20–9. doi:10.61108/ijsshr.v1i1.5

2. Mandal P, Jain T. Partial outsourcing from a rival: quality decision under product differentiation and information asymmetry. *Eur J Oper Res* (2021) 292(3):886–908. doi:10.1016/j.ejor.2020.11.018

3. Schaltegger S, Burritt R, Bai C, et al. Measuring and managing sustainability performance of supply chains. *Supply Chain Manag Int J* (2014) 19(3):275–91. doi:10.1108/scm-02-2014-0083

4. Schoenherr T. Outsourcing decisions in global supply chains: an exploratory multi-country survey. *Int J Prod Res* (2010) 48(2):343–78. doi:10.1080/00207540903174908

5. Dekkers R. Impact of strategic decision making for outsourcing on managing manufacturing. *Int J Operations Prod Manage* (2011) 31(9):935–65. doi:10.1108/01443571111165839

6. Gunasekaran A, Irani Z, Choy KL, Filippi L, Papadopoulos T. Performance measures and metrics in outsourcing decisions: a review for research and applications. *Int J Prod Econ* (2015) 161:153–66. doi:10.1016/j.ijpe.2014.12.021

7. Gilbert SM, Xia Y, Yu G. Strategic outsourcing for competing OEMs that face cost reduction opportunities. *IIE Trans* (2006) 38(11):903–15. doi:10.1080/07408170600854644

8. Kenyon GN, Meixell MJ, Westfall PH. Production outsourcing and operational performance: an empirical study using secondary data. *Int J Prod Econ* (2016) 171:336–49. doi:10.1016/j.ijpe.2015.09.017

9. Heydari J, Govindan K, Nasab HRE, Taleizadeh AA. Coordination by quantity flexibility contract in a two-echelon supply chain system: effect of outsourcing decisions. *Int J Prod Econ* (2020) 225:107586. doi:10.1016/j.ijpe.2019.107586

10. Lee G, Shin G, Hwang DW, Kuper P, Kang M. How manufacturers' long-term orientation toward suppliers influences outsourcing performance. *Ind Marketing Manage* (2018) 74:288–97. doi:10.1016/j.indmarman.2018.07.003

11. Chen X, Wang X, Xia Y. Production coopetition strategies for competing manufacturers that produce partially substitutable products. *Prod Operations Manage* (2019) 28(6):1446–64. doi:10.1111/poms.12998

12. Xiao T, Qi X. Product line design and outsourcing strategies in dyadic supply chains. *IEEE Trans Eng Manage* (2017) 64(3):316–26. doi:10.1109/tem.2017.2665702

13. Chen KS, Chung L, Chang TC. Developing a quality-based supplier selection model from the buying company perspective. *Qual Tech Quantitative Manage* (2021) 18(3):267–84. doi:10.1080/16843703.2020.1787307

14. Niu B, Zeng F, Chen L. "Production + procurement" outsourcing with competitive contract manufacturer's partial learning and supplier's price discrimination. *Int Trans Oper Res* (2021) 28(4):1917–51. doi:10.1111/itor.12904

15. Hsieh CC, Liu YT. Quality investment and inspection policy in a supplier–manufacturer supply chain. *Eur J Oper Res* (2010) 202(3):717–29. doi:10.1016/j.ejor.2009.06.013

16. Guan X, Liu B, Chen Y, Wang H. Inducing supply chain transparency through supplier encroachment. *Prod Operations Manage* (2020) 29(3):725–49. doi:10.1111/poms.13133

17. Chen Y, Karamemis G, Zhang J. A win–win strategy analysis for an original equipment manufacturer and a contract manufacturer in a competitive market. *Eur J Oper Res* (2021) 293(1):177–89. doi:10.1016/j.ejor.2020.12.016

18. Karaer Ö, Erhun F. Quality and entry deterrence. *Eur J Oper Res* (2015) 240(1):292–303. doi:10.1016/j.ejor.2014.07.016

19. Xie G, Wang S, Lai KK. Quality improvement in competing supply chains. *Int J Prod Econ* (2011) 134(1):262–70. doi:10.1016/j.ijpe.2011.07.007

20. Bing X, Jue G, Richard YK. Quality Investment Strategies of start-ups under capital constraints. *Syst Eng Theor Pract* (2019) 39(06):1373–84. doi:10.12011/1000-6788-2018-1303-12

21. Chen J, Liang L, Yang F. Cooperative quality investment in outsourcing. *Int J Prod Econ* (2015) 162:174–91. doi:10.1016/j.ijpe.2015.01.019

22. Nagurney A, Li D. A supply chain network game theory model with product differentiation, outsourcing of production and distribution, and quality and price competition. *Ann Operations Res* (2015) 226(1):479–503. doi:10.1007/s10479-014-1692-5

23. Yang FX, Zhang RQ, Zhu K. Should purchasing activities be outsourced along with production? *Eur J Oper Res* (2017) 257(2):468–82. doi:10.1016/j.ejor.2016.07.029

24. Xiao T, Xia Y, Zhang GP. Strategic outsourcing decisions for manufacturers competing on product quality. *IIE Trans* (2014) 46(4):313–29. doi:10.1080/0740817x.2012.761368

25. Karaer Ö, Kraft T, Yalçın P. Supplier development in a multi-tier supply chain. *IISE Trans* (2020) 52(4):464–77. doi:10.1080/24725854.2019.1659523

26. Heese HS, Kemahlıoğlu-Ziya E, Perdikaki O. Outsourcing under competition and scale economies: when to choose a competitor as a supplier. *Decis Sci* (2021) 52(5):1209–41. doi:10.1111/deci.12449

27. Pun H. The more the better? Optimal degree of supply-chain cooperation between competitors. *J Oper Res Soc* (2015) 66(12):2092–101. doi:10.1057/jors.2015.40

28. Li C. Supplier competition and cost reduction with endogenous information asymmetry. *Manufacturing Serv Operations Manage* (2020) 22(5):996–1010. doi:10.1287/msom.2019.0784

29. Feng Q, Lu LX. The strategic perils of low cost outsourcing. *Manage Sci* (2012) 58(6):1196–210. doi:10.1287/mnsc.1110.1481

30. Banker RD, Khosla I, Sinha KK. Quality and competition. *Manage Sci* (1998) 44(9):1179–92. doi:10.1287/mnsc.44.9.1179

31. Bing X, Jue G, Fung RYK. Quality Investment Strategies of start-ups under capital constraints. *Syst Eng Theor Pract* (2019) 39(06):1373–84.

Vaccination strategies in the disease–behavior evolution model

Lu Zhou[1], Jinying Dai[1], Bo Qu[2]* and Cong Li[1]*

[1]Adaptive Networks and Control Lab, Department of Electronic Engineering, School of Information Science and Technology, Fudan University, Shanghai, China, [2]Department of Computer Science, Guangdong University of Science and Technology, Dongguan, Guangdong, China

Previous studies on the co-evolving between vaccination strategies and epidemics mainly assumed that the vaccination strategies were made in the period between two spreading seasons. However, individual cognition during the spreading seasons might also alter the vaccination strategy and inversely influence the epidemic spreading. We propose a coupled disease–behavior model to describe the dynamic evolution of vaccination behavior during the spread of infectious diseases. The model integrates a susceptible–infected–vaccinated (SIV) model with the diffusion of vaccination behavior. We focus on the trade-off between perceptions of infection risk and the vaccination behaviors of neighbors, characterizing individual vaccination opinions. We introduce an opinion-critical value to map vaccination opinions into vaccination behavior. The vaccination coverage of the disease–behavior model is studied in network models and real-world networks. In addition, when societal costs are measured based on the degree of initial vaccinees, the cost of randomly selecting initial vaccinees is lower than selecting individuals with high or low degrees as vaccinees. Evaluating an individual's ability to transmit vaccination behavior based on the neighbor's number is inappropriate. We find that the impact of effective spreading rates on group vaccination is not one-sided and that reducing fear and highlighting the dangers of infectious diseases are crucial to increasing vaccination coverage.

KEYWORDS

vaccination behavior, infectious diseases, individual opinion, coupled model, social networks

1 Introduction

The outbreak of infectious diseases seriously endangers human health and social development [1, 2]. Intense research effort has been devoted to developing epidemic spreading models [3, 4]. In addition, previous studies [5–8] have also shown that the spread of infectious diseases interacts with individual behavior. Many studies focus on vaccination, which is considered one of the most successful and cost-effective health interventions [9, 10]. Individuals continuously adjust their attitude toward vaccination during the spread of infectious diseases. In-depth research on the dynamic evolution of individual vaccination behavior and its impact on the epidemic is of great significance for formulating more effective public health policies. Many studies have primarily constructed disease–behavior-coupled models from the perspective of information dissemination or economic costs [11–13]. The dissemination of disease-related information inhibits the spread of infectious diseases and contributes to the recovery of infected individuals [14]. [15] found that

vaccination coverage increases as people become sensitive to disease-related information, which increases the likelihood of herd immunity. Researchers have explored the dynamics of epidemic spreading in situations where vaccines are not fully effective. [16] studied both the case of a fixed immunity loss rate and an asymptotic total loss scenario based on the assumption of limited knowledge and temporary immunity. [17] pointed out that curbing the spread of negative information and improving vaccine effectiveness are effective means to prevent and control epidemics. Information-driven vaccination behavior significantly reduces the social cost of infection and facilitates the process of disease eradication [18] but ignores individual considerations of vaccine costs and vaccine spillover effects. The reason is that unvaccinated, self-interested individuals are dedicated to obtaining protection from other vaccinated individuals [19, 20]. For instance, [21] combined classical game theory with an epidemic model, revealing the "free-rider" behavior of self-interested individuals. [22, 23] found that the Nash equilibrium of vaccination game based on the vaccination cost could not form herd immunization due to conflict between herd and individual interests.

Psychological and behavioral experiments indicate that individual behavior tends to deviate from the rational criterion under the expected utility theory, exhibiting bounded rationality [24, 25]. For vaccination, even if a rational decision model predicts that vaccines will be accepted by individuals, in reality, low-cost and highly effective vaccines may still be rejected [26]. [10] proposed a two-stage vaccination game model that includes the disease spreading stage and the vaccination strategy update stage and illustrates the vaccine dilemma due to evolving psychological perceptions based on vaccine costs. Therefore, studying the impact of bounded rationality on individual vaccination decisions is crucial for disease control [27–29]. Prospect theory (PT) [30], which explains the decision-making process of individuals in the case of risk and uncertainty, not only captures the subjective perception of risk but also reveals the key role of bounded rationality in decision-making. PT contains two core concepts, namely, the weighting effect (WE), which describes an individual's subjective perception probability [31, 32], and the framing effect (FE), which indicates an individual's subjective evaluation of payoffs [33, 34]. In relevant studies, [35] developed an imperfect vaccination evolutionary game model, accounting for subjective perception and individual social differences. The results revealed that the epidemic threshold is significantly influenced by social differences in the epidemic spreading layer. [36] proposed an evolutionary vaccination game model in multiplex networks, incorporating an information-epidemic spreading process into vaccination dynamics. They found that the effect of information dissemination on vaccination decisions depends on vaccination costs, network topology, and the evolutionary stage of the system.

However, two-stage vaccination game models fail to capture the interactive dynamics between individual vaccination behavior and epidemic spreading. Moreover, traditional vaccination game models focus only on vaccination costs and payoffs, neglecting individual psychological cognition, i.e., perceptions of infection risk and vaccination behaviors of neighbors. In this work, we propose a disease–behavior-coupled model where individuals are exposed to the risk of infection and make vaccination decisions at each time step. In the context of free vaccines, we mainly focus on

psychological perceptions of influence risk. Here, an individual will have a vaccination behavior if their vaccination opinion is higher than a critical value. Given the general preference of individuals for reliable information sources, we assume that individuals make vaccination decisions based on local disease-related information. Each individual updates his/her vaccination opinion based on the weighted aggregation of the vaccination behaviors of neighbors and then adjusts his or her vaccination behavior at each time step. Specifically, the opinion weights depict the perceptions of infection risk, which are related to individual states. In addition, we study the dynamics of the coupled model and analyze the vaccination coverage of the proposed model in network models and real-world social networks.

The main contributions of our work are as follows: 1) we propose a coupled disease–behavior model to study the dynamic interactions between the spread of infectious diseases and the vaccination behaviors of individuals. Vaccination behavior is dominated by individuals' bounded rationality about infection risk, which is characterized by opinion weights. 2) The vaccination opinion in the co-evolution model could exhibit limited rationality. A relatively small infection fraction makes individuals underestimate the infection risks, while a large infection fraction leads individuals to "lie down" and be unwilling to defend themselves against infectious diseases. The phenomenon is verified with simulation results. 3) We find that vaccination coverage will reach convergence, which is strongly related to individual vaccination strategies. The opinion critical θ for an individual to get vaccinated has a decisive effect on the vaccination evolution game, leading to a clear phase transition in vaccination coverage *versus* the opinion critical. The link density of a network might influence vaccination coverage. Moreover, we analyze the performance of three vaccination strategies by administrators on the coupled model and find the advantage of the random-first vaccination strategy in promoting group vaccination. Compared with the random selection strategy, the high-degree individual priority vaccination strategy and the low-degree individual priority vaccination strategy exhibit low efficiency and high social costs. The findings in this work provide some clues for understanding the co-evolution of vaccination behavior and epidemic spread.

The remainder of this paper is arranged as follows: Section 2 illustrates the disease–behavior-coupled model in detail, which includes the dynamics of infectious disease spread and the evolution of vaccination behavior. The vaccination coverage of vaccination strategies by individuals and administrators is studied in network models and real-world networks in Section 3. Section 4 provides the conclusion.

2 Disease−behavior coupled model

2.1 Notations and preliminaries

We use nodes to represent the individuals in society and edges to indicate the interactions between members. A social interaction network could be characterized by an adjacency matrix $\mathbf{A} = [a_{ij}]_{N \times N}$. $a_{ij} = 1$ if there is a link between nodes i and j; otherwise, $a_{ij} = 0$. In this work, we focus on finite-size,
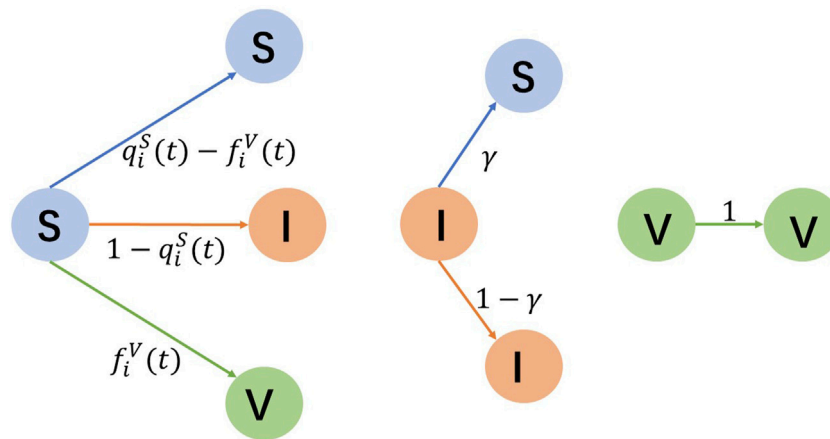
FIGURE 1
Transition of susceptible (S), infected (I), and vaccinated (V) states. S-individuals either remain susceptible or shift to be infected and vaccinated.
I-individuals have a probability γ of reverting to be susceptible. V-individuals stay vaccinated.
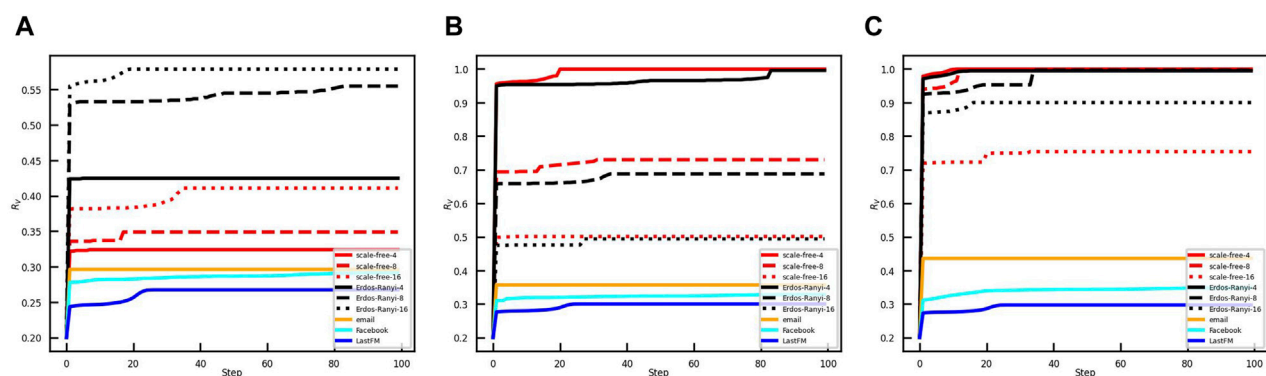

FIGURE 2
Convergence of vaccination coverage of the coupled model in the different networks. The ratio of initial vaccinees in the networks is 0.2. The red,
black, yellow, azure, and blue lines show vaccination coverage in scale-free networks, ER random networks, email network, Facebook network, and
LastFM network, respectively. The initial vaccinees are selected with **(A)** LFS, **(B)** SFS, and **(C)** RFS.

undirected, and unweighted networks [37]. The disease–behavior
coupled model is composed of two parts: the
susceptible–infected–vaccinated (SIV) spread model [38] and the
evolution rules of vaccination behavior.

2.2 Spread of infectious diseases

We first introduce the SIV model, where individuals have three
possible states: susceptible (S), infected (I), and vaccinated (V). The
diagram of the state transition is depicted in Figure 1. A susceptible
individual i would like to be vaccinated at time t with a probability
$f_i^V(t)$. Without the loss of generality, we here assume that the
vaccine is fully immune and long-term effective; in other words,
vaccinated individuals will not be infected. An infected individual
infects susceptible neighbors with an infection probability β and
cures with a recovery probability γ [39]. Let the symbols $p_i^S(t)$,
$p_i^I(t)$, and $p_i^V(t)$ denote the probabilities of being susceptible,
infected, and vaccinated for individual i at time t, respectively.

Then, there is the equation $p_i^S(t) + p_i^I(t) + p_i^V(t) = 1$. The
transition probability $q_i^S(t)$ of susceptible individual i not being
infected by neighbors is defined as follows:

$$q_i^S(t) = \prod_{j=1}^{N}\left(1 - a_{ij} \cdot p_j^I(t) \cdot \beta\right). \tag{1}$$

The continuous-time Markov approach can accurately
characterize the dynamics of infectious diseases [40]. However,
the state transition matrix is hardly available, especially for large-
scale networks [41]. Therefore, we use the microscopic Markov
chain approach [42, 43] to describe the probability of individual i
being susceptible, infected, and vaccinated at each moment as

$$\begin{cases} p_i^S(t+1) = p_i^I(t) \cdot \gamma + p_i^S(t) \cdot \left(q_i^S(t) - f_i^V(t)\right), \\ p_i^I(t+1) = p_i^I(t) \cdot (1 - \gamma) + p_i^S(t) \cdot \left(1 - q_i^S(t)\right), \\ p_i^V(t+1) = p_i^S(t) \cdot f_i^V(t) + p_i^V(t). \end{cases} \tag{2}$$

Notably, the dynamics will reach a steady state at the end of the
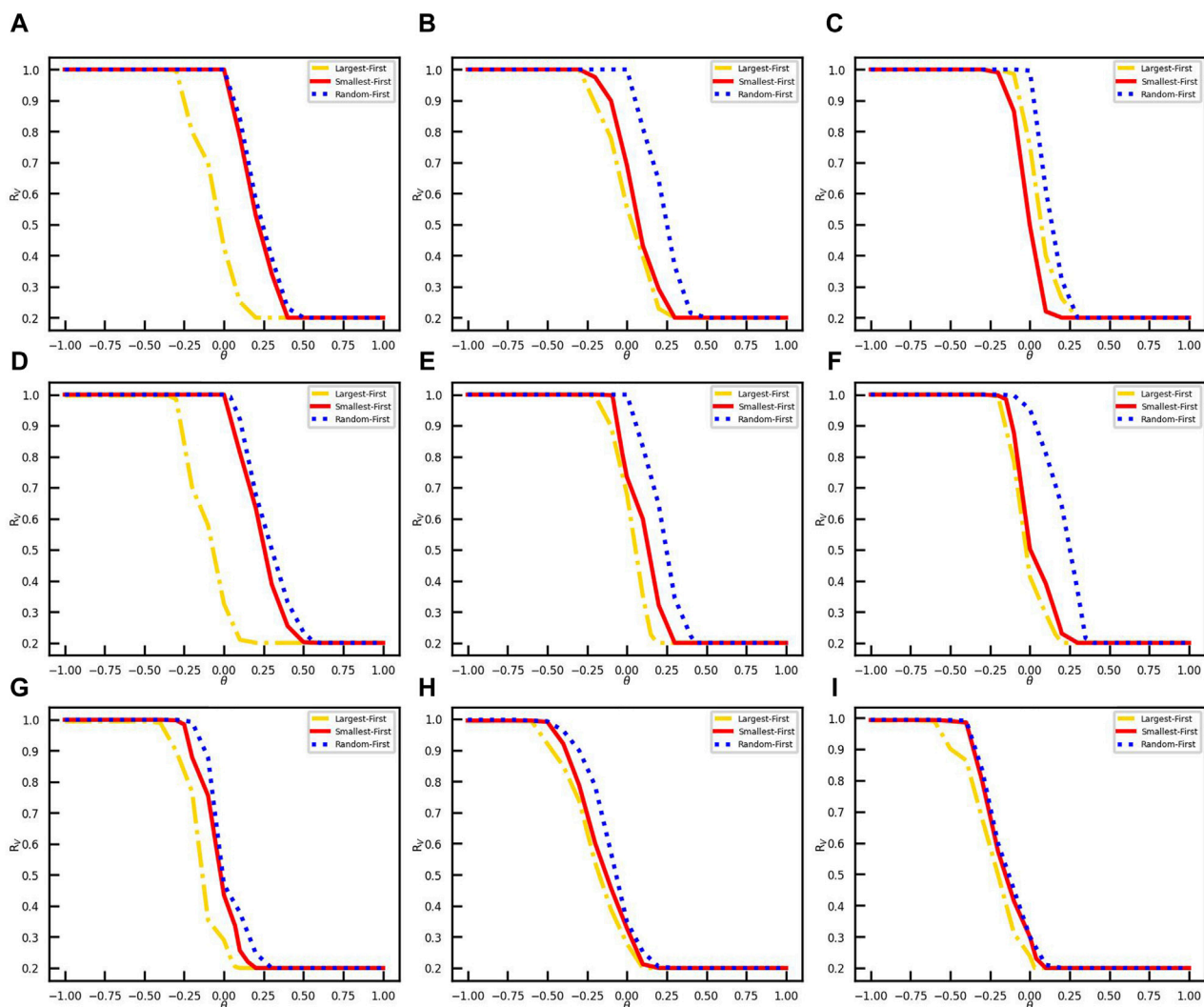spread process. Then, the probabilities of individual i being

FIGURE 3
Influence of opinion critical on vaccination coverage in different networks. The vaccination coverages R_V are obtained with three vaccination strategies, i. e., LFS, SFS, and RFS in yellow dot-line, red line, and blue dots. **(A)** Erdos−Ranyi-4 network. **(B)** Erdos−Ranyi-8 network. **(C)** Erdos−Ranyi-16 network. **(D)** Scale-free-4 network. **(E)** Scale-free-8 network. **(F)** Scale-free-16 network. **(G)** Email network. **(H)** Facebook network. **(I)** LastFM network.

susceptible, infected, and vaccinated no longer vary with time, namely, $p_i^S(t + 1) = p_i^S(t) = p_i^S$, $p_i^I(t + 1) = p_i^I(t) = p_i^I$, and $p_i^V(t + 1) = p_i^V(t) = p_i^V$, respectively. Naturally, the proportions of susceptible, infected, and vaccinated individuals in the population remain constant. We obtain $q_i^S(t + 1) = q_i^S(t) = q_i^S$ according to Eq. 1. Since the infection and recovery probabilities of infected individual are also constant, the transition probability of susceptible individual i to be vaccinated is constant, that is, $f_i^V(t + 1) = f_i^V(t) = f_i^V$. Then, we rewrite Eq. 2 as

$$\begin{cases} p_i^S = p_i^I \cdot \gamma + p_i^S \cdot (q_i^S - f_i^V), \\ p_i^I = p_i^I \cdot (1 - \gamma) + p_i^S \cdot (1 - q_i^S), \\ p_i^V = p_i^S \cdot f_i^V + p_i^V. \end{cases} \quad (3)$$

Based on the expression $p_i^S + p_i^I + p_i^V = 1$, we get Eq. 4 from Eq. 3

$$\begin{cases} p_i^S = \dfrac{\gamma \cdot (1 - p_i^V)}{1 + \gamma - q_i^S}, \\ p_i^I = \dfrac{(1 - q_i^S) \cdot (1 - p_i^V)}{1 + \gamma - q_i^S}. \end{cases} \quad (4)$$

Furthermore, there is $p_i^V = p_i^S \cdot f_i^V + p_i^V$ from Eq. 3. Therefore, the necessary condition for the steady state of the coupled model is $p_i^S \cdot f_i^V = 0$. When $p_i^S = 0$, we get $p_i^I = 0$ according to Equation 3. Therefore, there are only vaccinated individuals in the population, that is, $p_i^V = 1$. When $f_i^V = 0$ and $p_i^S \neq 0$, individuals exist in three states.

Let the symbols R_S, R_I, and R_V denote the proportions of susceptible, infected, and vaccinated individuals, respectively. Naturally, the three proportions satisfy $R_S + R_I + R_V = 1$ and are calculated using the Eq. 5:
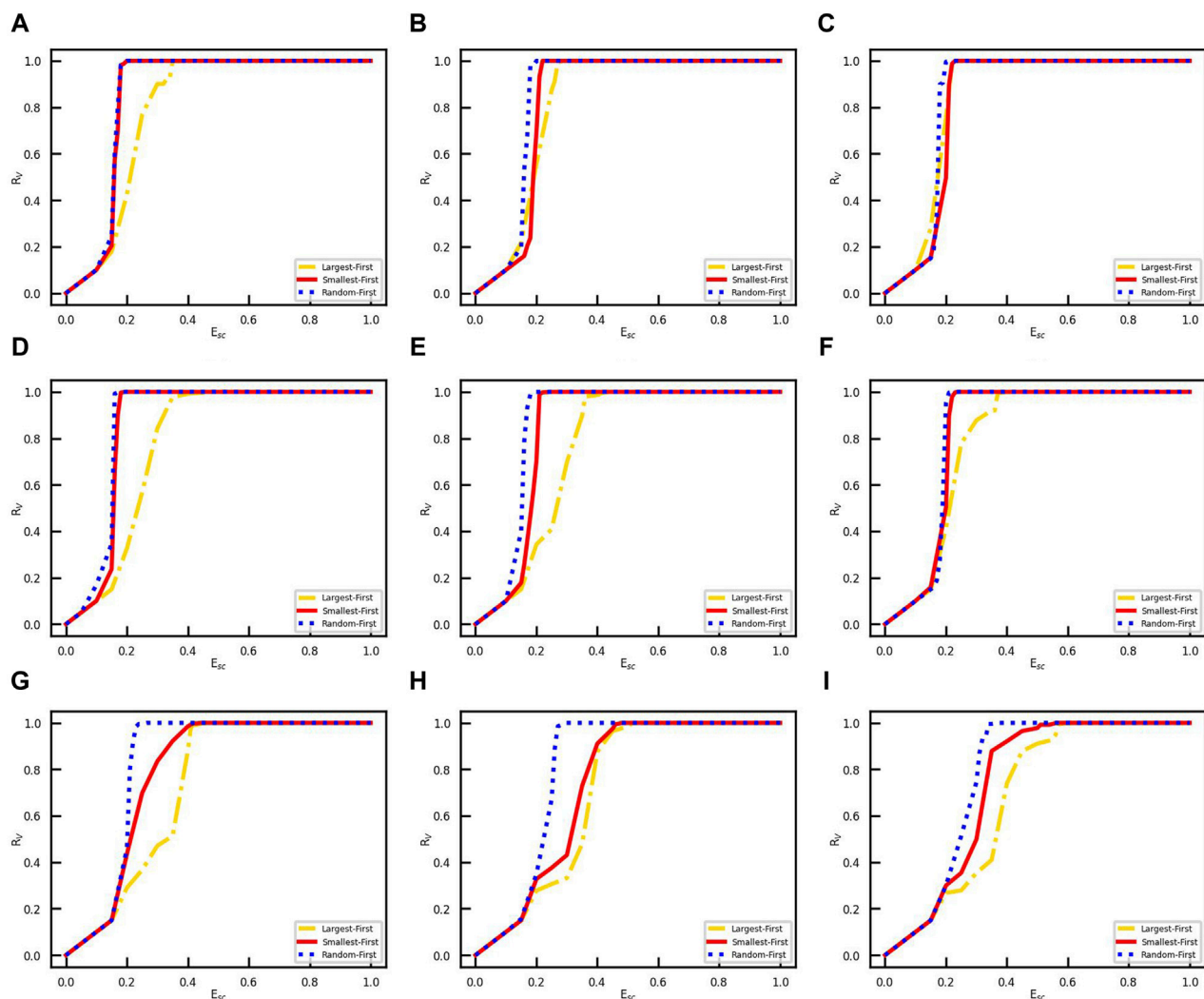
FIGURE 4
Vaccination coverage for different E_{sc} values. The vertical coordinate R_V denotes the rate of finally vaccinated nodes in the network, and the horizontal coordinate indicates the initially vaccinated degree. **(A)** Erdos−Ranyi-4 network. **(B)** Erdos−Ranyi-8 network. **(C)** Erdos−Ranyi-16 network. **(D)** Scale-free-4 network. **(E)** Scale-free-8 network. **(F)** Scale-free-16 network. **(G)** Email network. **(H)** Facebook network. **(I)** LastFM network.

$$
\begin{cases}
R_S = \dfrac{1}{N} \sum_{i=1}^{N} p_i^S, \\[2mm]
R_I = \dfrac{1}{N} \sum_{i=1}^{N} p_i^I, \\[2mm]
R_V = \dfrac{1}{N} \sum_{i=1}^{N} p_i^V.
\end{cases}
\tag{5}
$$

$$
Y_i(t) = Sgn\big(y_i(t) - \theta\big) = \begin{cases} 1, & y_i(t) \ge \theta, \\ -1, & \text{else}, \end{cases}
\tag{6}
$$

where the symbol θ indicates the opinion critical value for an individual to get vaccinated and $Sgn(.)$ is the sign function. When $Y_i(t) = 1$, individual i gets vaccinated at time t. Otherwise, i is out of vaccination. Notably, the increase in $y_i(t)$ indicates the enhancement of willingness for individual i to get vaccinated. The transition probability $f_i^V(t)$ of being vaccinated for susceptible individual i at time t equals to $y_i(t)$, when $y_i(t) \ge \theta$ at time t.

We use behavior vector $\mathbf{Y}(t) = [Y_1(t), Y_2(t), \dots, Y_N(t)]^T$ to represent behaviors of all individuals in the social network of size N at time t. Correspondingly, there is an opinion vector $\mathbf{y}(t) = [y_1(t), y_2(t), \dots, y_N(t)]^T$. The proportion of vaccinated individuals $R_V(t)$ at time t, named vaccination coverage, is computed as Eq. 7 that:

2.3 Evolution of vaccination behavior

The evolution process of vaccination behavior is the other important component of the coupled model. Let the notations $y_i(t)$ and $Y_i(t)$ denote the vaccination opinion and vaccination behavior of individual $i \in \{1, 2, \dots, N\}$ at time t, respectively. The relationship between $y_i(t)$ and $Y_i(t)$ is defined as Eq. 6 that

TABLE 1 Minimum social cost for the three strategies to reach vaccination coverage in networks with various typologies.

Network	Largest-first	Smallest-first	Random-first
Erdos–Ranyi-4	0.357	0.198	0.184
Erdos–Ranyi-8	0.293	0.217	0.191
Erdos–Ranyi-16	0.251	0.225	**0.209**
Scale-free-4	0.472	0.190	**0.179**
Scale-free-8	0.435	0.204	**0.189**
Scale-free-16	0.401	0.236	**0.215**
Email	0.451	0.440	**0.264**
Facebook	0.514	0.479	**0.295**
LastFM	0.609	0.561	**0.375**

We use red bold to highlight the results of the best options.

$$R_V(t) = \frac{\sum_{i=1}^{N} Y_i(t) + \sum_{i=1}^{N} |Y_i(t)|}{2N}. \quad (7)$$

A diagonal matrix $\Lambda(t) = diag(\lambda_{11}(t), \lambda_{22}(t), \ldots, \lambda_{NN}(t))$ is used to ensure that vaccinees cannot revert to being unvaccinated during the evolution process. If individual i is vaccinated, the element $\lambda_{ii}(t) = 0$; otherwise, $\lambda_{ii}(t) = 1$. In the disease–behavior-coupled model, individuals obtain vaccination opinions based on local information, namely, the perceptions of infection risk and vaccination behaviors of their neighbors. Here, when the number of infected neighbors increases, individuals perceive more disease risk and mitigate their willingness to receive vaccines. Meanwhile, when the number of infected nodes reaches a rather large number, individuals may have group psychology and give up vaccination. When the number of susceptible neighbors increases, individuals lessen their fear of disease and their inclination to receive vaccines. Individual vaccination behavior can be influenced by herd mentality [44], which is the tendency for people in a group to conform to the behavior of others in the group rather than acting as individuals. When the number of vaccinated neighbors increases, individuals are more likely to get vaccinated due to the influence of herd mentality. To simplify the complexity, we assume that individuals do not take into account their own opinion of the previous moment. We designed an opinion weight matrix $\mathbf{W}(t) = [w_{ij}(t)]_{N \times N}$ to characterize the perceptions of infection risk. The symbol $w_{ij}(t)$ indicates the influence weight of individual j on i at time t. The mathematical expression of opinion weight is shown as follows:

$$w_{ij}(t+1) = \begin{cases} \dfrac{1}{|N_i^I(t)|}, & j \in N_i^S(t) \text{ or } j \in N_i^V(t), \\ \dfrac{1}{|N_i^S(t)| + |N_i^V(t)|}, & j \in N_i^I(t), \\ 0, & \text{others}, \end{cases} \quad (8)$$

where the symbols $N_i^S(t)$, $N_i^V(t)$, and $N_i^I(t)$ denote the set of susceptible, vaccinated, and infected neighbors of individual i at time

t, respectively; and $|.|$ represents the cardinality of a set. Let the symbol N_i be the neighbors of individual i, thus $|N_i| = |N_i^S(t)| + |N_i^V(t)| + |N_i^S(t)|$. Particularly, for the extreme cases that all neighbors of individual i are infected or not infected, Eq. 8 is no longer applicable, and let the weight $w_{ij}(t) = 1$ if individuals i and j are neighbors.

Hence, we get the expression of individual vaccination opinion evolution equation as Eq. 9

$$y_i(t+1) = \lambda_{ii}(t) \cdot \sum_{j=1}^{n} w_{ij}(t) \cdot Y_j(t) + (1 - \lambda_{ii}(t)) \cdot Y_i(t), \quad (9)$$

and the corresponding matrix form is

$$\begin{aligned} \mathbf{y}(t+1) &= \Lambda(t) \cdot \mathbf{W}(t) \cdot \mathbf{Y}(t) + (\mathbf{E} - \Lambda(t)) \cdot \mathbf{Y}(t) \\ &= [\Lambda(t) \cdot \mathbf{W}(t) + (\mathbf{E} - \Lambda(t))] \cdot \mathbf{Y}(t) \end{aligned}. \quad (10)$$

If the individual i is vaccinated at time t, we get $y_i(t+1) = Y_i(t)$, which indicates that a vaccinated individual will maintain their current state. Otherwise, there is $y_i(t+1) = \sum_{j=1}^{n} w_{ij}(t) \cdot Y_j(t)$. The vaccination opinion $y_i(t+1)$ will be influenced by the vaccination behaviors of neighbors of individual i at time t. Moreover, the maximum number of possible neighbors for each node is $N-1$ in the connected network with N nodes. If the neighbor number of node i is $N-1$ and all neighbors are infected at time t, the minimum value of $y_i(t)$ is obtained from Eq. 10, and $y_i(t) = -N+1$. If the neighbor number of node i is $N-1$ and all neighbors are vaccinated at time t, the maximum value of $y_i(t)$ is $N-1$. Hence, the bounds of $y_i(t)$, $i \in \{1, 2, \ldots, N\}$ in the coupled model are between $-N+1$ and $N-1$.

From Eq. 10, the opinion vector $\mathbf{y}(t)$ is related to $\Lambda(0)$, $\mathbf{W}(t)$, and $\mathbf{Y}(0)$. The opinion weight matrix $\mathbf{W}(t)$ depends on the degrees of nodes in the network and the infection status of neighbors. The values of $\mathbf{Y}(0)$ and $\Lambda(0)$ depend on the initial vaccinees. Naturally, $\mathbf{y}(t)$ is determined by the network structure, initial vaccinees, and infection status of neighbors. Hence, changes in the value of θ reveal the role of network structure, initial vaccinees, and size of infected neighbors on vaccination behavior.

3 Evolutionary vaccination game in network models and real-world networks

We here study the vaccination coverage of the disease–behavior-coupled model in network models and real-world networks. Two types of vaccination strategies, namely, strategies by individuals and strategies by administrators, are studied. The individual vaccination strategy is based on the perceptions of infection risk and vaccination behaviors of neighbors, as introduced in Section 2. The vaccination strategies by administrators are used to select initial vaccinees. In this work, three vaccination strategies by administrators, namely, largest-first strategy (LFS), smallest-first strategy (SFS), and random-first strategy (RFS), were used. The LFS is used to select the initially vaccinated nodes based on the nodal degrees from large to small, and the SFS is the opposite. Naturally, RFS is used to randomly adopt the initial vaccinees. Notably, the network models are scale-free (SF) networks [45] and Erdos–Ranyi (ER) networks [46] with $N = 1,000$ and the
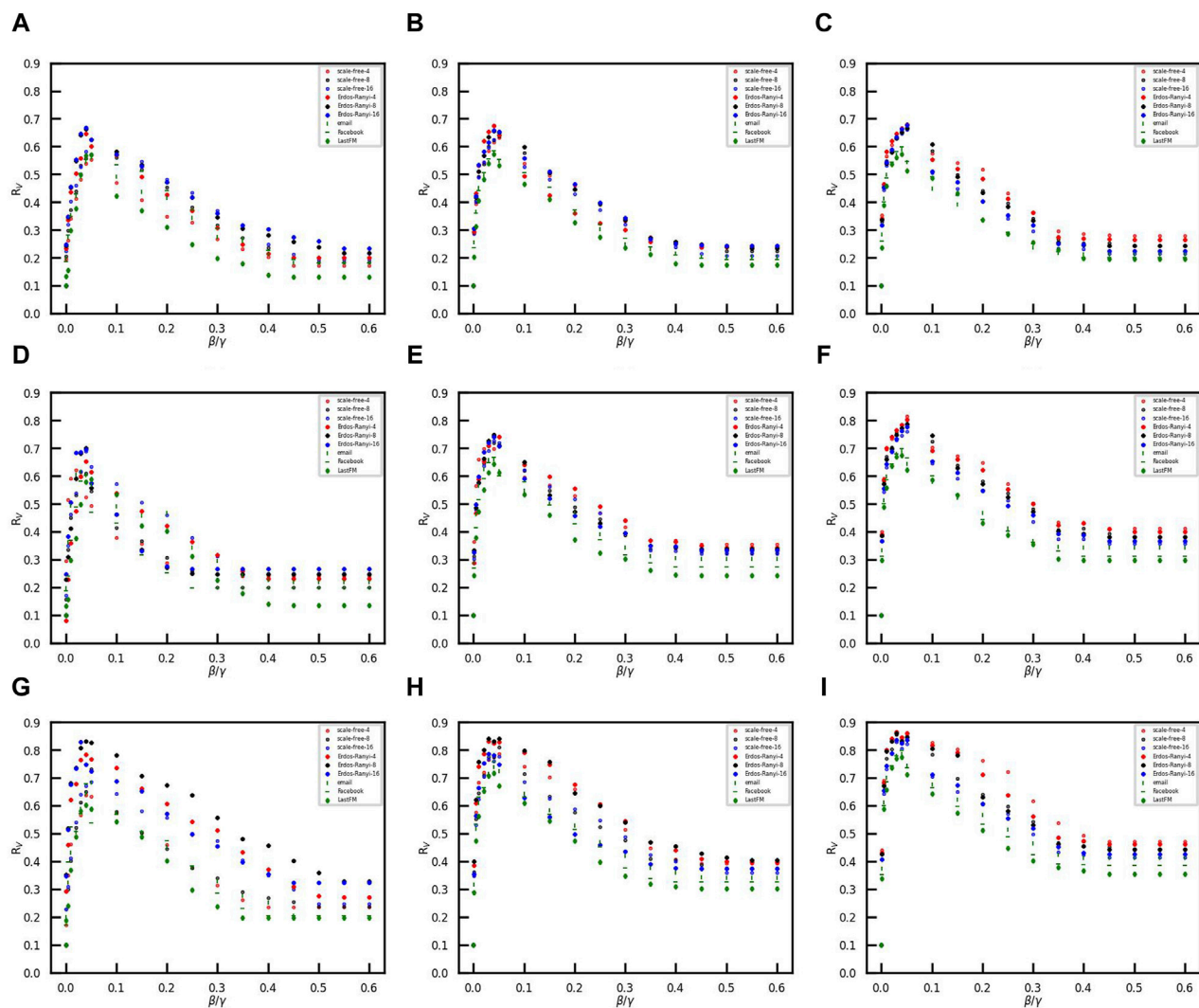
FIGURE 5
Vaccination coverage *versus* the effective spreading rate in the different networks. The ratio of initial vaccinees is 0.1, and the number of initial infected nodes is 50. **(A–C)** Opinion critical is 0.1..**(D–F)** Opinion critical is 0. **(G–I)** Opinion critical is −0.1. The initial vaccinees are selected with the largest-first strategy in **(A, D, and G)**, with the smallest-first strategy in **(B, E, and H)**, and with random-first strategy in **(C, F, and I)**.

average degrees $<k> = 4$, $<k> = 8$, and $<k> = 16$, respectively. The real social networks are email network [47], Facebook friendship network [48], and LastFM users' network [49]. The email network, generated using email data from a large European research institution, comprises 1,005 nodes and 25,571 edges. In the email network, users are represented as nodes and communication between them is represented as connected edges. The Facebook friendship network has been collected from survey participants using the Facebook application and consists of 4,039 nodes and 88,234 edges. The Facebook network is composed of users as nodes and friend relationships between them as edges. The LastFM users' network, consisting of 7,624 nodes and 27,806 edges, was collected from the public API in March 2020. Nodes are LastFM users from Asian countries, and edges are mutual follower relationships between them. In addition, each simulation result is the average of 50 times under the same parameters to avoid the accidentality of a single simulation.

3.1 Convergence of vaccination coverage of the coupled model

We first verify the convergence of the disease–behavior-coupled model in network models and the real-world social networks through Monte Carlo simulations. We attempt different combinations of all parameters, where the effective spreading rate β/γ ranges from 0.01 to 100, the opinion critical θ ranges from $-N + 1$ to $N - 1$, the number of initially infected nodes $N^I(0)$ ranges from 1 to N, and the number of initially vaccinated nodes $N^V(0)$ ranges from 1 to N. Without the loss of generality, the initially infected nodes are randomly selected since, in practice, infected individuals appear by chance.

We find that the proportion of vaccinated individuals R_V will converge to a constant, regardless of the effective spreading rate, opinion critical, and initially infected nodes. Hence, we give an illustration with the effective spreading rate $\beta/\gamma = 0.4$, the number of initially infected nodes $N^I(0) = 50$, and opinion critical $\theta = 0$. The

convergence of the model with three strategies, i.e., *LFS*, *SFS*, and *RFS*, are shown in Figure 2.

Vaccination coverage is influenced by the strategies by the administrator, network topology, and link density of the network. For the LFS, the vaccination coverage is the largest in the ER random networks compared to that in the SF networks and real-world networks. If the network topology and network size are given, the vaccination coverage is larger when the link density of the network is larger. However, the result is the opposite for SFS and RFS. Moreover, Figure 2 shows that vaccination coverage for LFS will converge to a smaller value than that for SFS and RFS.

3.2 Influence of opinion critical on vaccination coverage

The vaccination opinion critical θ in the coupled model represents the psychological threshold for vaccine acceptance. A higher opinion critical indicates that individuals are less inclined to prefer vaccines. We here study the role of opinion critical on vaccination coverage under the conditions of the effective spreading rate $\beta/\gamma = 0.4$, the number of initially infected nodes $N^I(0) = 50$, and the initial vaccinee ratio $R_V(0) = 0.2$. Figure 3 illustrates the variation of opinion critical θ from −1 to 1 in different networks. Under the different network structures and initial vaccination strategies, vaccination coverage decreases as the opinion critical increases from negative to positive values. Figures 3A–C show that in the Erdos–Ranyi networks, the SFS and LFS are both affected by the link density. The link density is higher; the psychological threshold for full vaccination coverage is lower with the SFS. However, with the LFS, θ for full vaccination coverage increases with the link density. Moreover, Figures 3D–F demonstrate that in the scale-free networks with the SFS, the psychological threshold for full vaccination coverage decreases with the increase in the link density. In contrast, the LFS exhibits a higher psychological threshold for full vaccination coverage as the link density of the network increases. The finding verifies the opposing results presented in Figures 2A,B.

3.3 Comparison of the three vaccination strategies

We further study the performances of the three vaccination strategies on the social cost in different networks. In complex network models such as Price's model and BA model, the node degree is used as an important indicator of the node's attractiveness to new nodes and ability to develop new links [50]. In this work, the social influence or status of a node is related to the degree $d_i = \sum_{j=1}^N a_{ij}$ of an individual i in the social networks. We assume that persuading a more influential person to get vaccinated will cost more socially. Hence, the social cost of initially vaccinated individuals is given as Eq. 11,

$$E_{sc} = \frac{\sum_{i \in N^V(0)} d_i}{\sum_{j=1}^N d_j}, \tag{11}$$

where $N^V(0)$ is the set of initially vaccinated nodes and E_{sc} is the social cost of strategy. It should be noted that the vaccine is free for

individuals, but there is a social cost to the government in promoting vaccination. We study the effect of the social cost of each strategy on vaccination coverage. We conduct experiments under the same initial conditions with the effective spreading rate $\beta/\gamma = 0.4$, the number of initially infected nodes $N^I(0) = 50$, and opinion critical $\theta = 0$. The vaccination coverage corresponding to each strategy at one social cost is the average of the results of 50 simulation experiments with the same parameters. We conducted the simulation experiment by setting the social cost E_{sc} values in 0.1, 0.02, 0.005, and 0.001 step sizes in turn. The approximate range of the minimum social cost is first determined in large steps, and then the step size of the experimental parameters is gradually reduced to determine an accurate minimum social cost. The relationship between social cost E_{sc} and vaccination coverage R_V is shown in Figure 4. We find that vaccination coverage will be reached when the social cost is greater than a threshold, which is listed in Table 1.

The results show that with the LFS, the minimum social cost E_{sc} for vaccination coverage decreases as the link density increases in the same network topology. However, the social cost threshold for the SFS and RFS gradually increases as the link density increases in the same network topology. In addition, we compare the performance of three strategies on the same network; surprisingly, RFS always needs the minimum social cost for vaccination coverage, both in network models and real-world networks. Then, the performance of the SFS is superior to that of the LFS.

3.4 Role of the effective spreading rate

We then explore the effect of the effective spreading rate β/γ on vaccination coverage under three vaccination strategies, i.e., LFS, SFS, and RFS. Figure 5 demonstrates that the effective spreading rate β/γ has a double-edged role in the diffusion of vaccination behavior when the social cost E_{sc} is less than the minimum social cost of the three strategies, as shown in Table 1.

When the effective spreading rate increases from small to large, vaccination coverage R_V first increases and then decreases. The phenomenon might be explained by the limited rationality of vaccination opinions. When the effective spreading rate is small, the number of infected individuals is also small. Therefore, susceptible individuals tend to ignore the risk of disease and refuse vaccination. However, when the effective spreading rate is large, the number of infected individuals is large. A large number of infected individuals leads to a tendency for susceptible individuals to coexist with the virus rather than resist disease transmission. The confidence of individuals in vaccines crumbles, and many individuals shift toward abandoning self-loathing due to the panic caused by rapid outbreaks of disease. Only if the effective spreading rate is moderate, individuals are not only aware of the risk of disease, but also inclined to accept the vaccine. As a result, the proportion of vaccinees is higher than that in the other two scenarios. Emphasizing the dangers of infectious diseases and reducing panic are both essential to increasing vaccination coverage in epidemic control. It is a primary concern for government policymakers to promote herd immunization when faced with the outbreak of infectious diseases.

4 Conclusion

In this work, we propose a coupled disease–behavior evolution model, providing a new perspective on the interactions between vaccination behavior and the spread of infectious diseases. We portray the mental choices of individuals facing disease risk and vaccination by the variable opinion weights, which capture the vaccination behaviors of neighbors. The vaccination strategy by individuals is based on their mental choices and exhibits limited rationality about infection risk. A large infection fraction may lead individuals to adopt negative strategies to resist infectious diseases, while a relatively small infection fraction makes individuals adopt positive strategies. A clear phase transition appears in the vaccination coverage compared to the opinion critical θ of an individual to be vaccinated. Meanwhile, the performance of three vaccination strategies, namely, LFS, SFS, and RFS, by the administrator is compared in this work. We find that with the three initial vaccination strategies, vaccination coverage, which is influenced by the link density of the network and network topology, always converges to a constant. The vaccination coverage of RFS and SFS is consistently higher than that of LFS. Persuading individuals with high influence to get vaccinated at the initial time is not optimal for promoting the diffusion of vaccination behavior. RFS has the best performance on both network models and real-world networks among the three strategies when studying the effect of the opinion critical and the social cost. In addition, the role of the effective spreading rate is not one-sided since the vaccination opinion exhibits limited rationality. Vaccination coverage R_V first increases and then decreases as the effective spreading rate increases from small to large. Controlling for outbreak information to make individuals perceive a "false and appropriate effective spreading rate" is an efficacious way to motivate individuals to be vaccinated.

The phenomena revealed by this work could provide a new perspective for guiding group vaccination opinions and improving vaccination coverage. The model introduced here also has some limitations and challenges. We did not account for the variability of opinion critical of different groups that belong to the same social network. Individuals between different groups have a greater difference in their opinion critical than individuals within a group. As an individual acquires information, his or her opinion critical may change. In addition, we focus on the neighbors' influence, but the influence of non-neighboring individuals or global information also merits further investigation.

Author contributions

LZ: data curation, formal analysis, methodology, software, writing–original draft, and writing–original draft. JD: writing–review and editing. BQ: formal analysis, methodology, writing–review and editing, and writing–review and editing. CL: conceptualization, funding acquisition, methodology, writing–review and editing, and writing–review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Zino L, Rizzo A, Porfiri M. On assessing control actions for epidemic models on temporal networks. *IEEE Control Syst Lett* (2020) 4:1–802. doi:10.1109/LCSYS.2020.2993104

2. Baker RE, Mahmud AS, Miller IF, Rajeev M, Rasambainarivo F, Rice BL, et al. Infectious disease in an era of global change. *Nat Rev Microbiol* (2022) 20:193–205. doi:10.1038/s41579-021-00639-z

3. Funk S, Salathe M, Jansen VAA. Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J R Soc Interf* (2010) 7:1247–56. doi:10.1098/rsif.2010.0142

4. Shang YL. Analytical solution for an in-host viral infection model with time-inhomogeneous rates. *Acta Physica Pol B* (2015) 46:1567. doi:10.5506/APhysPolB.46.1567

5. Glass RJ, Glass LM, Beyeler WE, Min HJ. Targeted social distancing designs for pandemic influenza. *Emerging Infect Dis* (2006) 12:1671–81. doi:10.3201/eid1211.060255

6. Mao L, Yang Y. Coupling infectious diseases, human preventive behavior, and networks–a conceptual framework for epidemic modeling. *Soc Sci Med* (2012) 74:167–75. doi:10.1016/j.socscimed.2011.10.012

7. Calabro GE, Carini E, Tognetto A, Giacchetta I, Bonanno E, Mariani M, et al. The value(s) of vaccination: building the scientific evidence according to a value-based healthcare approach. *Front Public Health* (2022) 10:786662. doi:10.3389/fpubh.2022.786662

8. Kuga K, Tanimoto J. Which is more effective for suppressing an infectious disease: imperfect vaccination or defense against contagion? *J Stat Mech Theor Exp* (2018) 2018:023407. doi:10.1088/1742-5468/aaac3c

9. Andre FE, Booy R, Bock HL, Clemens J, Datta SK, John TJ, et al. Vaccination greatly reduces disease, disability, death and inequity worldwide. *Bull World Health Organ* (2008) 86:140–6. doi:10.2471/blt.07.040089

10. Feng X, Wu B, Wang L. Voluntary vaccination dilemma with evolving psychological perceptions. *J Theor Biol* (2018) 439:65–75. doi:10.1016/j.jtbi.2017.11.011

11. Wang Z, Andrews MA, Wu ZX, Wang L, Bauch CT. Coupled disease–behavior dynamics on complex networks: a review. *Phys Life Rev* (2015) 15:1–29. doi:10.1016/j.plrev.2015.07.006

12. Wang W, Liu QH, Liang JH, Hu YQ, Zhou T. Coevolution spreading in complex networks. *Phys Rep* (2019) 820:1–51. doi:10.1016/j.physrep.2019.07.001

13. Shaham A, Chodick G, Shalev V, Yamin D. Personal and social patterns predict influenza vaccination decision. *BMC Public Health* (2020) 20:222–12. doi:10.1186/s12889-020-8327-3

14. Kabir KA, Kuga K, Tanimoto J. Analysis of sir epidemic model with information spreading of awareness. *Chaos, Solitons and Fractals* (2019) 119:118–25. doi:10.1016/j.chaos.2018.12.017

15. Xia S, Liu JM. A belief-based model for characterizing the spread of awareness and its impacts on individuals' vaccination decisions. *J R Soc Interf* (2014) 11:20140013. doi:10.1098/rsif.2014.0013

16. Shang YL. Immunization of networks with limited knowledge and temporary immunity. *Chaos: Interdiscip J Nonlinear Sci* (2021) 31:053117. doi:10.1063/5.0045445

17. Yin Q, Wang Z, Xia C, Bauch CT. Impact of co-evolution of negative vaccine-related information, vaccination behavior and epidemic spreading in multilayer networks. *Commun Nonlinear Sci Numer Simulation* (2022) 109:106312. doi:10.1016/j.cnsns.2022.106312

18. Buonomo B. Effects of information-dependent vaccination behavior on coronavirus outbreak: insights from a siri model. *Ricerche di Matematica* (2020) 69:483–99. doi:10.1007/s11587-020-00506-8

19. Xin Y, Gerberry D, Just W. Open-minded imitation can achieve near-optimal vaccination coverage. *J Math Biol* (2019) 79:1491–514. doi:10.1007/s00285-019-01401-z

20. Wang XY, Jia DY, Gao SP, Xia CY, Li XL, Wang Z. Vaccination behavior by coupling the epidemic spreading with the human decision under the game theory. *Appl Math Comput* (2020) 380:125232. doi:10.1016/j.amc.2020.125232

21. Bauch CT, Galvani AP, Earn DJD. Group interest versus self-interest in smallpox vaccination policy. *Proc Natl Acad Sci* (2003) 100:10564–7. doi:10.1073/pnas.1731324100

22. Li XJ, Li C, Li X. Vaccinating sis epidemics in networks with zero-determinant strategy. In: 2017 IEEE International Symposium on Circuits and Systems (ISCAS) (IEEE); May 28-31, 2017; Baltimore, MD, USA (2017). p. 2275–8.

23. Li XJ, Li C, Li X. Minimizing social cost of vaccinating network sis epidemics. *IEEE Trans Netw Sci Eng* (2017) 5:326–35. doi:10.1109/TNSE.2017.2766665

24. Simon HA. A behavioral model of rational choice. *Q J Econ* (1955) 69:99–118. doi:10.2307/1884852

25. Hota AR, Sundaram S. Game-theoretic vaccination against networked sis epidemics and impacts of human decision-making. *IEEE Trans Control Netw Syst* (2019) 6:1461–72. doi:10.1109/TCNS.2019.2897904

26. Oraby T, Bauch CT. Bounded rationality alters the dynamics of paediatric immunization acceptance. *Scientific Rep* (2015) 5:10724–12. doi:10.1038/srep10724

27. Ibuka Y, Chapman GB, Meyers LA, Li M, Galvani AP. The dynamics of risk perceptions and precautionary behavior in response to 2009 (h1n1) pandemic influenza. *BMC Infect Dis* (2010) 10:296–11. doi:10.1186/1471-2334-10-296

28. Dr?kowski D, Trepanowski R. Reactance and perceived disease severity as determinants of covid-19 vaccination intention: an application of the theory of planned behavior. *Psychol Health Med* (2021) 27:2171–8. doi:10.1080/13548506.2021.2014060

29. Wolff K. Covid-19 vaccination intentions: the theory of planned behavior, optimistic bias, and anticipated regret. *Front Psychol* (2021) 12:648289. doi:10.3389/fpsyg.2021.648289

30. Kahneman D, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* (1979) 47:263–91. doi:10.2307/1914185

31. Yang YX, Park LT, Mandayam NB, Seskar I, Glass A, Sinha N. Prospect pricing in cognitive radio networks. *IEEE Trans Cogn Commun Networking* (2015) 1:56–70. doi:10.1109/TCCN.2015.2488636

32. Jhala K, Natarajan B, Pahwa A. Prospect theory-based active consumer behavior under variable electricity pricing. *IEEE Trans Smart Grid* (2019) 10:2809–19. doi:10.1109/TSG.2018.2810819

33. Wang YP, Saad W, Sarwat AI, Hong CS. Reactive power compensation game under prospect-theoretic framing effects. *IEEE Trans Smart Grid* (2018) 9:4181–93. doi:10.1109/TSG.2017.2652846

34. Rahi GE, Etesami SR, Saad W, Mandayam NB, Poor HV. Managing price uncertainty in prosumer-centric energy trading: a prospect-theoretic stackelberg game approach. *IEEE Trans Smart Grid* (2019) 10:702–13. doi:10.1109/TSG.2017.2750706

35. Li C, Dai JY, Li X. A new species of *Entedon* Dalman (Hymenoptera, Eulophidae) and three newly recorded species from China. *IEEE Trans Comput Soc Syst* (2023) 1172:1–14. doi:10.3897/zookeys.1172.104676

36. Li XJ, Li C, Li X. The impact of information dissemination on vaccination in multiplex networks. *Sci China Inf Sci* (2022) 65:172202. doi:10.1007/s11432-020-3076-1

37. Proskurnikov AV, Tempo R. A tutorial on modeling and analysis of dynamic social networks. part i. *Annu Rev Control* (2017) 43:65–79. doi:10.1016/j.arcontrol.2017.03.002

38. Liu Q, Jiang DQ, Shi NZ, Hayat T, Alsaedi A. Nontrivial periodic solution of a stochastic non-autonomous sisv epidemic model. *Physica A: Stat Mech its Appl* (2016) 462:837–45. doi:10.1016/j.physa.2016.06.041

39. Mei WJ, Mohagheghi S, Zampieri S, Bullo F. On the dynamics of deterministic epidemic propagation over networks. *Annu Rev Control* (2017) 44:116–28. doi:10.1016/j.arcontrol.2017.09.002

40. Mieghem PV, Omic J, Kooij R. Virus spread in networks. *IEEE/ACM Trans Networking* (2009) 17:1–14. doi:10.1109/TNET.2008.925623

41. Wang W, Tang M, Stanley HE, Braunstein LA. Unification of theoretical approaches for epidemic spreading on complex networks. *Rep Prog Phys* (2017) 80:036603. doi:10.1088/1361-6633/aa5398

42. Granell C, Gómez S, Arenas A. Dynamical interplay between awareness and epidemic spreading in multiplex networks. *Phys Rev Lett* (2013) 111:128701. doi:10.1103/PhysRevLett.111.128701

43. de Arruda GF, Rodrigues FA, Moreno Y. Fundamentals of spreading processes in single and multilayer complex networks. *Phys Rep* (2018) 756:1–59. doi:10.1016/j.physrep.2018.06.007

44. Valle SYD, Hethcote HW, Hyman JM, Castillo-Chavez C. Effects of behavioral changes in a smallpox attack model. *Math Biosciences* (2005) 195:228–51. doi:10.1016/j.mbs.2005.03.006

45. Barabási AL, Albert R. Emergence of scaling in random networks. *Science* (1999) 286:509–12. doi:10.1126/science.286.5439.509

46. Erdös P, Rényi A. On random graphs i. *Publicationes Mathematicae Debrecen* (1959) 6:290–7. doi:10.5486/pmd.1959.6.3-4.12

47. Yin H, Benson AR, Leskovec J, Gleich DF. Local higher-order graph clustering. In: Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM); August 13-17, 2017; Halifax, NS (2017). p. 555–64.

48. McAuley J, Leskovec J. Learning to discover social circles in ego networks. In: Proceedings of the 25th International Conference on Neural Information Processing Systems (NIPS); December 3-8, 2012; Lake Tahoe (2012). p. 539–47.

49. Rozemberczki B, Sarkar R. Characteristic functions on graphs: birds of a feather, from statistical descriptors to parametric models. In: Proceedings of the 29th ACM International Conference on Information and Knowledge Management (ACM); October 19-23, 2020; Virtual Event, Ireland (2020). p. 1325–34.

50. Newman EJM. The structure and function of complex networks. *SIAM Rev* (2003) 45:167–256. doi:10.1137/S003614450342480

A game study on the impact of employees' deviant innovation behaviors on firms' organizational innovation performance

Han Zheng[1], Jie Lu[2]*, Yanxia Chen[3], Yingying Gu[4] and Zixin Zheng[4]

[1]School of Finance and Economics, Jiangsu University, Zhenjiang, Jiangsu, China, [2]School of Business Administration, Shanghai Lixin School of Accounting and Finance, Shanghai, China, [3]School of Management, Jiangsu University, Zhenjiang, Jiangsu, China, [4]School of Accounting, Shanghai Lixin School of Accounting and Finance, Shanghai, China

Organizational innovation performance is considered to be the key to maintaining competitiveness and achieving sustainable development in modern enterprises. Deviant innovation refers to that employees improve their working methods without the permission of the organization, break through the constraints of existing rules and achieve high performance. Additionally, Deviant innovation behavior can also stimulate the enthusiasm of other members of the organization. In order to study the evolutionary game rules of strategic interaction between enterprises and employees in the process of deviant innovation, this paper constructs a 2*2 asymmetric payoff matrix, and uses numerical simulation to show the influence of different values of decision parameters and changes of initial conditions on the evolutionary results. The research reveals that the interaction game between employees' deviant innovation behavior and enterprise enhancement of organizational innovation performance is a complex and significant topic. It is found that when enterprises actively improve organizational innovation performance and employees implement deviant innovation behavior, both sides engage in positive game interactions, maximizing enterprise benefits. However, when enterprises passively enhance organizational innovation performance but employees engage in deviant innovation behavior, it may lead to conflicts between the two parties. The research findings provide relevant strategies for employees to correctly implement deviant innovation behavior and for enterprises to enhance organizational innovation performance.

KEYWORDS

deviant innovation behavior, organizational innovation performance, evolutionary game, numerical experiment, undesirable locking state

1 Introduction

In modern enterprises, organizational innovation performance is considered to be the key to maintaining competitiveness and achieving sustainable development. However, innovation is no longer solely the responsibility of senior management but increasingly relies on employees at all levels within the organization. Deviant innovation refers to that employees improve their working methods without the permission of the organization, break through the constraints of existing rules and aim to improve the interests of the organization [1]. We find that employees' positive

deviant innovation behavior can prompt enterprises to better integrate and utilize internal and external resources, promote the landing of innovation culture, and accelerate the promotion of innovation projects, thus enhancing the innovation competitiveness of the enterprise and further optimizing the innovation performance of the organization [2].

Organizational innovation performance, as a comprehensive reflection of various innovation achievements of the enterprise, is expressed as the overall role and impact results brought to the enterprise from the generation of innovative ideas and the adoption of innovative measures to apply them to the actual production and operation activities, which itself is a multi-level concept [3]; [4]. This study believes that employees' deviant innovation behaviors can motivate other members of the organization, trigger chain reactions internally, and promote broader innovation, thus greatly contributing to the improvement of organizational innovation performance. When enterprises actively enhance organizational innovation performance and provide sufficient resources and support, while employees also engage in deviant innovation behavior, an ideal state is achieved. At this point, employees and enterprises jointly achieve optimal interaction effects, creating a win-win situation [5].

Evolutionary game theory is an interdisciplinary fusion theory that combines evolutionary biology and classical game theory [6]. It views individual strategies and behaviors as continually evolving through the transmission of genetic information and random mutations. In evolutionary game theory, individual behaviors are strategies, and the evolutionary process is a dynamic game process. Individuals adapt and adjust their strategies by considering the pros and cons of different strategies [7].

Currently, evolutionary game theory has become one of the primary mathematical tools for research in economics and finance [8]. In recent years, numerous scholars both domestically and internationally have conducted research in this field. Eid et al. [9] finds out the balanced distribution of post-disaster insurance plans purchased by households and sold by insurance companies and post-disaster relief implemented by government agencies. Feng and Hu [10] investigated the evolutionary game of financing empowerment mechanisms for the digital transformation of small and medium-sized enterprises, exploring the impacts of core enterprise support and government subsidies on the digital transformation financing of small and medium-sized enterprises.

In the practice of enterprises enhancing innovation performance and employees engaging in deviant innovative behavior, evolutionary game theory also provides a practical analytical framework. When employees are faced with different company innovation atmospheres and leadership decisions, their choices regarding deviant innovation exhibit characteristics of evolutionary game theory. Similarly, when enterprises decide on the level of investment in innovation, they consider various factors such as employee capabilities, operational objectives, and return on investment [11]. However, in situations of incomplete information, they often make decisions based on personal experience and habits, leading to interactive games between both parties. Therefore, to study the interaction between enterprises and employees, this paper constructs a dynamic game model composed of enterprises and employees.

TABLE 1 Game Matrix of employee deviant innovation and organizational performance.

	B1(q)	B2 (1-q)
A1(p)	$(\pi_1 - K + B, \pi_2 + X(K) + C)$	$(\pi_1 - K, \pi_2 + X(K))$
A2 (1-p)	$(\pi_1, \pi_2 - D)$	(π_1, π_2)

2 Model description

Due to the short-sighted behavior under the condition of limited rationality, in the process of innovation, whether the enterprise actively improves the organizational innovation performance and the employee's deviant innovation behavior is a dynamic process of continuous adjustment, similar to the characteristics in ecology. Enterprises, considering long-term benefits, may to some extent enhance organizational innovation performance [12]. However, their performance may be influenced by employees' deviant innovation behavior, which can manifest in two ways: 1) positively enhancing organizational innovation performance (A1) with a probability of p; 2) negatively enhancing organizational innovation performance (A2) with a probability of 1-p.

Similarly, assuming that the employees will initially engage in deviant innovation behaviors, and during the process from initiating deviant innovation to executing it, they will continuously adjust their behavior patterns. Therefore, the employees will also have two choices: 1) to carry out the deviant innovation behavior (B1), with the probability of q; 2) not to carry out the deviant innovation behavior (B2), with the probability of 1-q, the payoff matrix for the two sides of the game as shown in Table 1.

π_1 and π_2, respectively denote the basic profits of the company and employees themselves. K represents the investment made by the company to actively enhance organizational innovation performance, and X(K) represents the additional profit that can be gained due to the simultaneous enhancement of organizational innovation performance by the company and deviant innovative behavior by employees. $0 < X(K) < 1$, that is, the more active the enterprise is in terms of organizational innovation performance, the greater the benefits of employees will be.

The interpretation of the matrix is as follows: 1) When the company actively improves its organizational innovation performance and the employees conduct deviant innovation behaviors, the company's profit is represented as $\pi_1 - K + B$, and the employees' profit is represented as $\pi_2 + X(K) + C$. B is the revenue allocated from the "excess profit" gained by the employee from the deviant innovation, and C is the performance wage that the employee receives because he/she performs the innovative behavior and thus improves the organizational innovation performance of the firm. Where B > K, because in the long run, organizational innovation performance can improve the overall earnings of the enterprise, thus the profit obtained by the company through employees' deviant innovative behaviors should be greater than the investment made in actively enhancing organizational innovation performance. 2) When the firm actively improves organizational innovation performance but the employees do not engage in deviant innovation behavior, the firm's profit is $\pi_1 - K$, and the employees' profit is $\pi_2 + X(K)$. 3) When the firm negatively improves organizational innovation performance but employees

TABLE 2 Results of the analysis of the stability of the equilibrium point.

Equilibrium point	Det(J)	Tr	Results
O ($p = 0, q = 0$)	KD^+	$-(K + D)^-$	ESS
W ($p = 0, q = 1$)	$(B - K)D^+$	$(B - K) + D^+$	Precarious
U ($p = 1, q = 0$)	KC^+	$K + C^+$	Precarious
V ($p = 1, q = 1$)	$(B - K)C^+$	$K - (B + C)^-$	ESS
E ($p = \frac{D}{C+D}, q = \frac{K}{B}$)	$-\frac{DCK(B-K)}{B(C+D)}$	0	Saddle point

engage in deviant innovation behavior, the firm's profit is π_1, and the employee's profit is $\pi_2 - D$. D denotes the loss incurred by employees when they persist in deviate innovation while the company negatively improves organizational innovation performance. In this case, team members may exhibit lower acceptance of conflict, lower identification with organizational goals, and less trust among members. (4) When the company passively enhances organizational innovation performance and employees do not engage in deviant innovation behaviors, the firm's profit is π_1, and the employees' gain is π_2. If the employees continue not to carry out the deviant innovation, the company will gradually lose its competitiveness, which will lead to the continuous attrition. Therefore, both parties in the game can only maintain the original profit within a certain period.

According to the Malthusian dynamic equation [13], which states that the growth rate of a strategy is equal to its relative fitness, specifically as long as the fitness of individuals adopting the strategy is higher than the average fitness of the group, the strategy will grow over time. Therefore, the dynamical equation for the strategic interaction between organizational innovation performance and employees' deviant innovation over time is:

$$\begin{cases} \frac{d_p}{d_t} = p(1 - p)(qB - K) \\ \frac{d_q}{d_t} = q(1 - q)[(C + D)p - D] \end{cases} \tag{1}$$

According to equation set (1), we can study the evolution of organizational innovation performance and employees' strategic interaction behavior of deviant innovation. Let the Jacobian matrix of the system of equations (8.80) be J, then Eq. 2 can be derived.

$$J = \begin{matrix} (1 - 2p) & (qB - K)Bp(1 - p) \\ (C + D)q(1 - q) & (1 - 2q)[(C + D)p - D] \end{matrix} \tag{2}$$

Denoting the determinant of J as Det(J), and then we can discuss the stability of the system of Eq. 1.

Let $\frac{d_p}{d_t} = 0$ and $\frac{d_p}{d_t} = 0$, 5 equilibria of the evolutionary game can be obtained on the plane $M = \{(p, q) | 0 \leq p, q \leq 1\}$: O (0,0), W (0,1), U (1,0), V (1,1), and E ($\frac{D}{C+D}, \frac{E}{B}$), and their stability is shown in Table 2.

From Table 2, two of the five equilibrium points of the system are stable points, representing Evolutionarily Stable Strategies (ESS), corresponding to two game modes: 1) Mode (A1, B1) is that the employees engage in the deviant innovation behavior while the company positively enhances organizational innovation

FIGURE 1
The dynamic process of business-employee engagement.
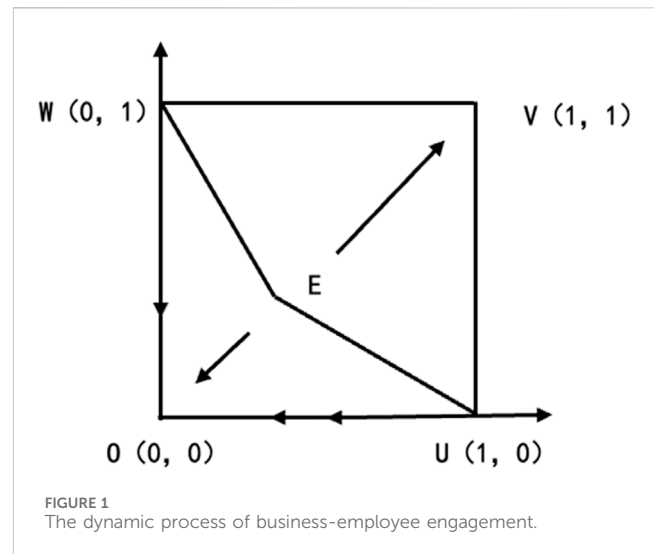
performance; 2) Mode (A2, B2) is that the employees do not engage in deviant innovation behavior while the enterprise passively enhances organizational innovation performance. Figure 1 illustrates the dynamic evolution process of organizational innovation performance and employees' deviant innovation behavior.

In Figure 1, the discount connected by two imbalance points (0,1), (1,0) and saddle point E ($\frac{D}{C+D}, \frac{K}{B}$) can be understood as the critical line where the system converges to different modes. If the initial position is located in the UEWO region, the system converges to the mode (A2, B2), which means the company negatively improves the organizational innovation performance, and the employees do not engage in deviant innovation behaviors. This is an undesirable locking state. Conversely, if the initial state is located in the UEWO region, the system converges to the mode (A1, B1), which means the company positively improves the organizational innovation performance, and the employees engage in deviant innovation behaviors. This is an ideal state. Both of these states are evolutionarily stable states, and participants adopting the alternative strategy will eventually vanish in evolution. Therefore, by adjusting parameters, the probability of the system reaching the ideal equilibrium state (A1, B1) can be increased.

When the parameters related to deviant innovation behavior and organizational innovation performance change, the speed of convergence of the model is also affected.

(1) The effect on the rate of convergence of the firm's benefit B from the employee's deviant innovation behavior and the employee's deviant innovation and hence performance pay C. At the saddle point, $\frac{\delta_q}{\delta_b} < 0$, with other parameters unchanged, the saddle point E ($\frac{D}{C+D}, \frac{K}{B}$) translates downward, the probability of convergence to mode (A1,B1) increases, and the probability of convergence to mode (A2, B2) decreases; conversely, the probability of convergence to mode (A1, B1) decreases, and the probability of convergence to mode (A2, B2) increases, as shown in Figure 2. At the saddle point, $\frac{\delta_p}{\delta_c} < 0$, the saddle point E ($\frac{D}{C+D}, \frac{K}{B}$) is translated to the left. The probability of convergence to mode (A1, B1) increases and the probability of convergence to mode (A2, B2) decreases; conversely, the probability of convergence to mode (A1, B1)
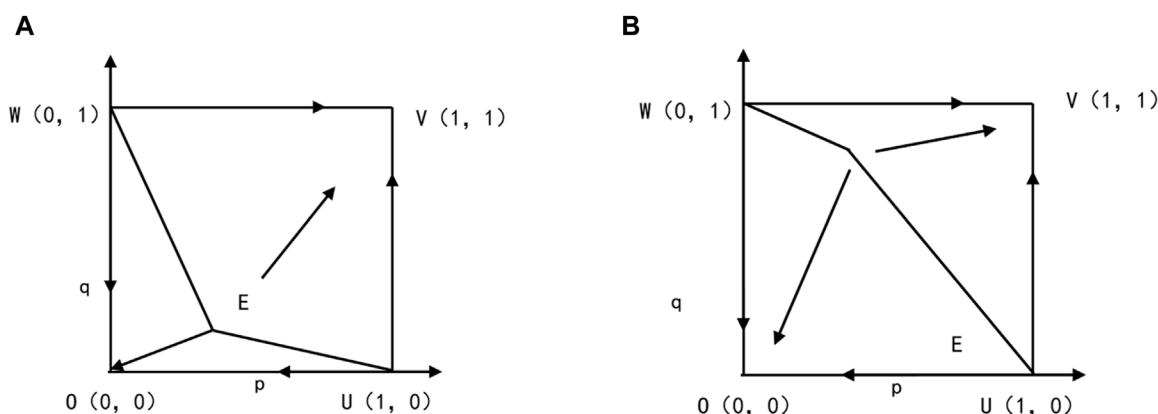
FIGURE 2
Effect of variation of parameters B, K on convergence rate. **(A)** When B increases and K decreases. **(B)** When B decreases and K increases.
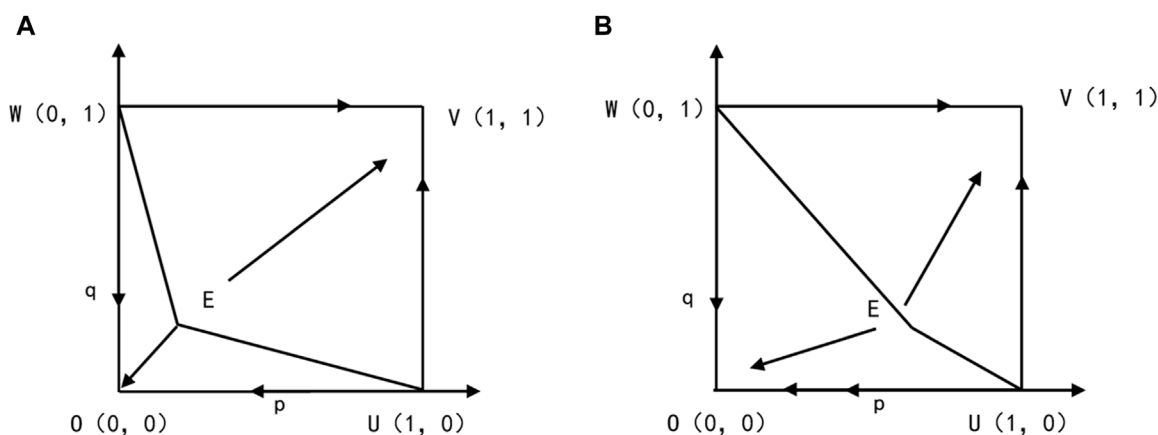

FIGURE 3
Effect of variation of parameters C, D on convergence rate. **(A)** When C increases and D decreases. **(B)** When C decreases and D increases.

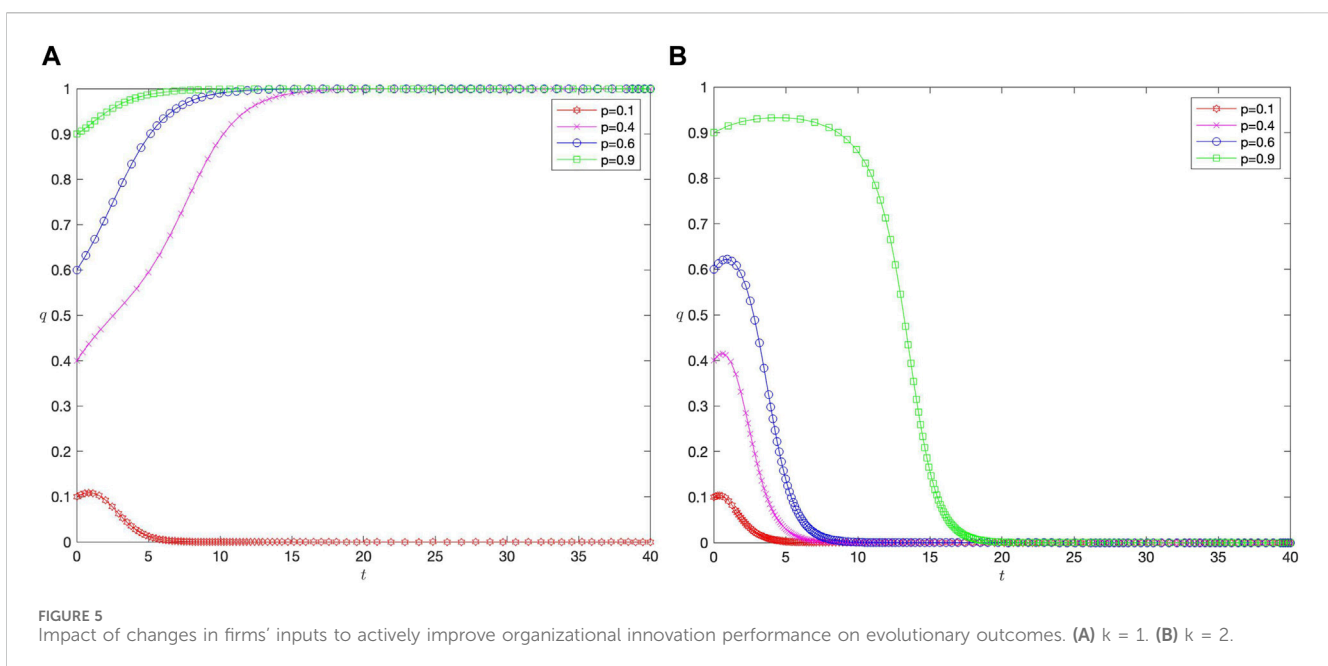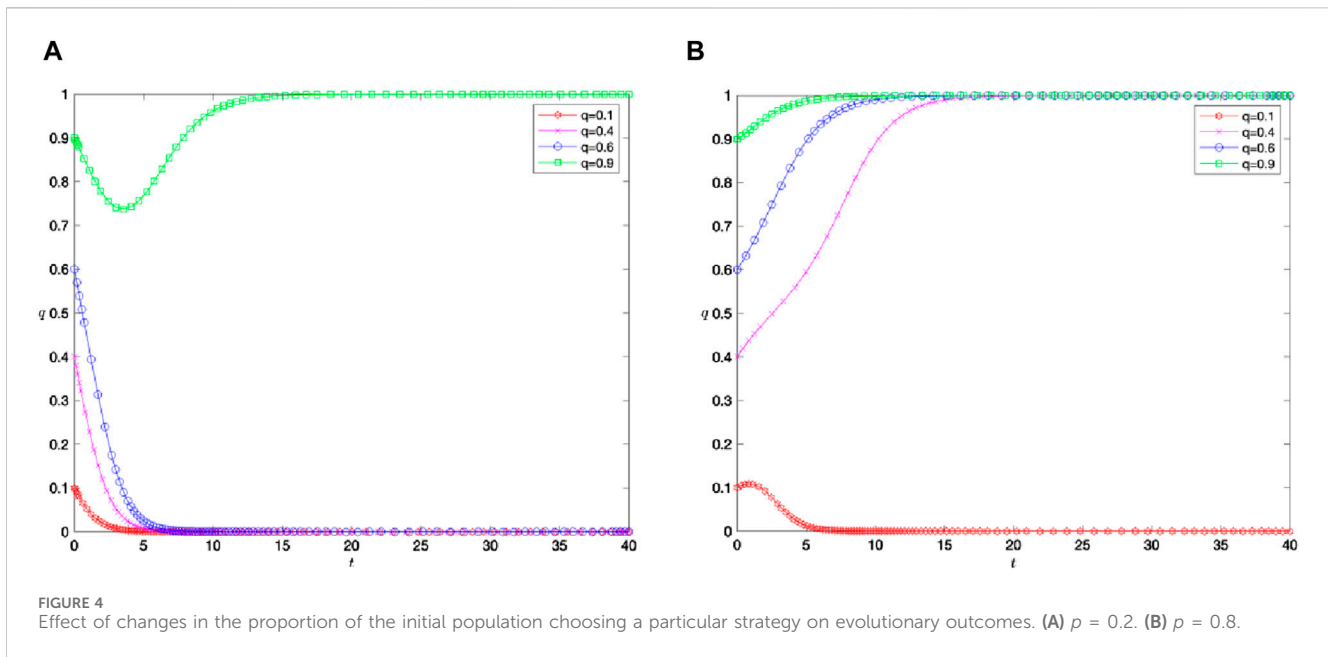decreases and the probability of convergence to mode (A2, B2) increases as shown in Figure 3.

(2) Similarly, the impact of the company's inputs K to improve the organizational innovation performance and the benefits X derived by the employees on the rate of convergence is shown in Figure 2. Because $X'(K) > 0$, indicating that within the domain of definition, a monotonic function possesses an inverse function, hence $K'(X) > 0$. This can be interpreted as follows: in an environment where employees actively strive to enhance organizational innovation performance, greater benefits can be obtained through the implementation of deviant innovation behaviors, thus promoting the survival and development of the enterprise. Consequently, the firm tends to invest more in this regard. X does not directly influence the convergence speed but affects it through the parameter K. In an environment where organizational innovation enhancement is lacking, the impact of loss D on the convergence speed when employees persist in implementing deviant innovation is depicted in Figure 3.

3 Simulation analysis

Next, by means of numerical experiments, we analyze the effects on the evolutionary results of the changes in parameters such as the proportion of the initial population choosing a certain strategy, the inputs carried out by enterprises to positively enhance the innovative performance of the organization, the additional benefits gained by the enterprises and the employees from the innovative behaviors, and the losses suffered by the employees who transgressed the innovation as a result of the negative enhancement of innovative performance of the organization.

(1) First of all, in this issue, we focus on the impact of the initial population proportion to choose the relevant strategy on the evolutionary outcome. The results of the numerical experiments are illustrated in Figure 4 below, where p and q respectively denote the initial proportions of firms choosing to actively enhance organizational innovation performance and the proportions of employees practicing deviant innovation behaviors. The horizontal axis represents time,

FIGURE 4
Effect of changes in the proportion of the initial population choosing a particular strategy on evolutionary outcomes. **(A)** *p* = 0.2. **(B)** *p* = 0.8.

FIGURE 5
Impact of changes in firms' inputs to actively improve organizational innovation performance on evolutionary outcomes. **(A)** k = 1. **(B)** k = 2.

and the vertical axis represents the proportion of employees who perform deviant innovative behaviors. The values of the parameters are K = 1, B = 2, C = 0.5, D = 1.

From the figure, it can be observed that the evolutionary paths of employees' deviant innovations and firms' organizational innovation behaviors depend largely on the choice of initial probabilities. The four different trajectories start from different initial probabilities and do not intersect or overlap before converging to the equilibrium state. In addition, the time required to converge to the equilibrium state varies in relation to the initial proportion of the population whose employees perform the deviant innovation behavior. The closer the initial proportion is

to the equilibrium state, the faster the convergence. Comparing the two sub-figures in Figure 2: we can see that the evolutionary outcome of employee behavior and the time of convergence are not only affected by the initial proportion of employees who choose deviant innovative behavior, but also by the initial proportion of the corporate population that actively improves the innovative performance of the organization. The closer the initial proportion is to the ideal state, the greater the possibility that the evolution of employee behavior will reach the ideal state.

(2) In order to understand the impact of firms' inputs to enhance organizational innovation performance on evolutionary outcomes, we conducted numerical experiments, the results
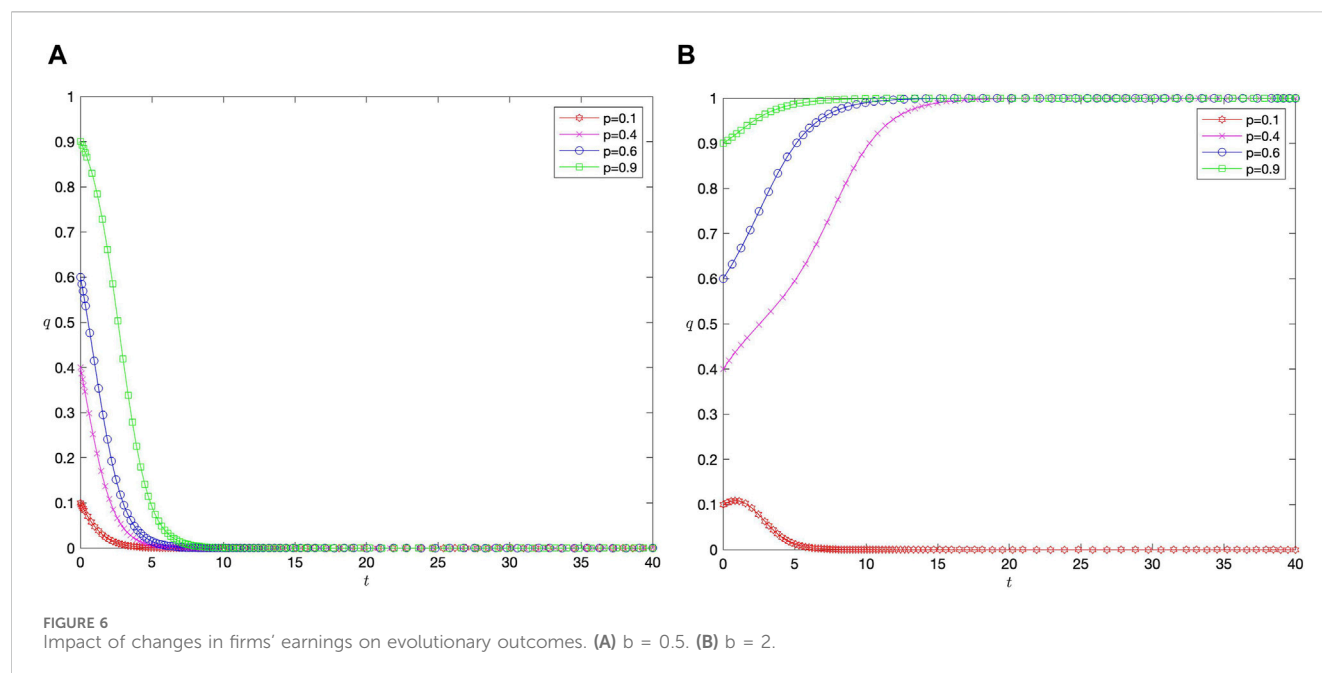
FIGURE 6
Impact of changes in firms' earnings on evolutionary outcomes. **(A)** b = 0.5. **(B)** b = 2.

of which are shown in Figure 5 below. In the figure, the horizontal axis represents time and the vertical axis represents the proportion of firms that are actively engaged in strategies to enhance organizational innovation performance. The proportion of employees choosing to perform deviant innovative behaviors is 0.8 (q = 0.8), and the other parameters are B = 2, C = 0.5, and D = 1.

Comparing the results in Figures 5A, B, we can observe that as the firm's investment in enhancing organizational innovation performance increases, the interaction behavior between the firm and its employees evolves to a completely different state, namely, mode (A2, B2). In this pattern, firms adopt negative strategies to enhance organizational innovation performance, while employees no longer perform deviant innovation behaviors, which is an undesirable locking state. However, ideally firms should actively enhance organizational innovation performance, so it is important to note that more investment is not always better, because it may fall into an undesirable locking state after exceeding a certain limit.

(3) Changes in firms' earnings have a significant impact on the evolutionary outcomes, and the results of the numerical experiments are shown in Figure 6. In the figure, the horizontal axis represents time, and the vertical axis represents the proportion of firms choosing to actively improve organizational innovation performance. The proportion of employees choosing to transgress innovation is 0.8, and the other parameters are K = 1, C = 0.5, and D = 1. By comparing the results in Figures 6A, B, it can be observed that the benefits that firms gain from the deviant innovation behaviors of their employees have a significant impact on the evolution of the interactive behaviors between firms and their employees to the desired state. When the proportion of employees engaging in deviant innovation is higher, the

firm's gain increases accordingly, which motivates the firm to be more active in improving organizational innovation performance, thus creating a win-win situation.

(4) The loss of deviant innovation behaviors suffered by employees has an impact on the evolutionary outcome in the case of firms taking negative measures to enhance organizational innovation performance, and the results of the numerical experiment are shown in Figure 7 below. In the figure, the horizontal axis represents time and the vertical axis represents the proportion of employees who choose deviant innovative behavior. The proportion of firms choosing to positively enhance organizational innovation performance is 0.8, and the other parameters are K = 1, B = 2, and C = 0.5. By comparing the results in Figures 7A, B, it can be observed that when the loss suffered by employees due to deviant innovation behaviors increases, it is more likely that the interaction behaviors between the firms and the employees will evolve into the (A2, B2) pattern. This is because if an employee's deviant innovative behavior may lead to an increase in the likelihood or degree of task conflict escalating into emotional conflict, then the psychological loss suffered by the employee will increase accordingly, which may lead the employee to choose a more conservative strategy, i.e., not to engage in deviant innovative behavior.

4 Game conclusion

The game analysis of the interaction between employees' deviant innovation and enterprises' organizational innovation performance provides us with profound insights into the complex relationship between employees' deviant innovation behaviors and enterprises' organizational innovation

FIGURE 7
Impact of changes in employee losses on evolutionary outcomes. **(A)** d = 0.5. **(B)** d = 3.5.
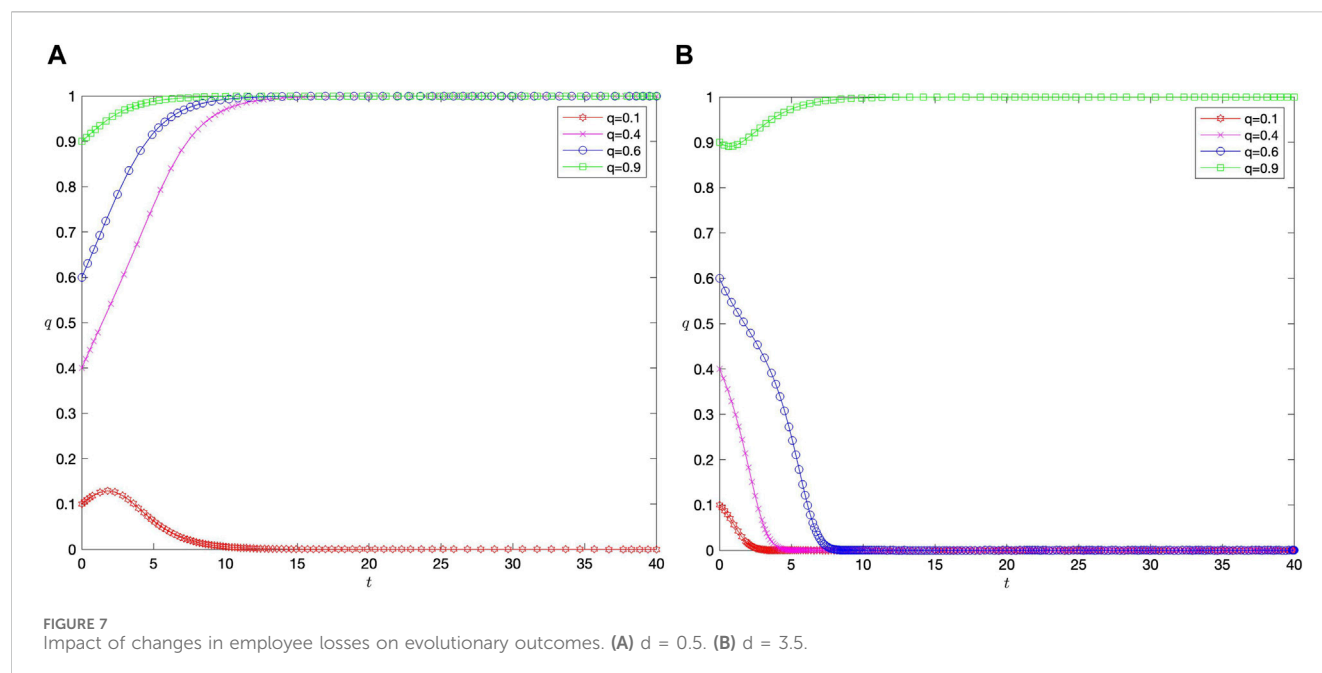
performance. It is initially found that when the enterprise actively improves organizational innovation performance and the employees implement deviant innovation behaviors, both sides of the game are positive, which maximizes the interests of the enterprise; when the enterprise actively improves organizational innovation performance but the employees do not implement deviant innovation behaviors, the enterprise's earnings are relatively reduced while the employee's earnings are affected by the environment to a certain extent; when the enterprise negatively improves organizational innovation performance but the employees implement deviant innovation behaviors, it may lead to the escalation of task conflict and the increase in the number of employees. When the enterprise negatively improves the organizational innovation performance but the employees carry out the deviant innovation behavior, it may lead to the possibility or degree of task conflict escalation into emotional conflict increases, then the employees in the psychological loss will be increased accordingly, which may lead to the employees to choose a more conservative strategy, that is to say, the enterprise negatively improves the organizational innovation performance, the employees do not carry out the deviant innovation behavior of the undesirable locking state. In order to maximize the interests of both enterprises and employees, this paper proposes relevant measures based on the game results.

From the enterprise perspective, companies need to manage this interaction wisely, as this interactive game is one of the key factors for organizational success. Firstly, it is essential to define organizational performance goals, requiring thoughtful and effective strategic planning to ensure rational resource allocation [14]. If companies fail to provide sufficient resources for organizational innovation performance at the outset, employees may feel a lack of support, resulting in the failure of innovation activities. Secondly, fostering an open, inclusive, and innovative cultural atmosphere, establishing efficient innovation teams, and

motivating employees' deviant innovation behavior are crucial [15]. Thirdly, scientific management is required to facilitate positive interactions among employees. When employees actively engage in deviant innovation, companies can leverage their employees' efforts to adjust their strategies to better respond to market demands. This includes enhancing products, services, or processes to meet evolving customer expectations [16]. Companies can thus adapt more rapidly to market dynamics, thereby enhancing organizational innovation performance.

From the employee perspective, proactive implementation of deviant innovation behavior by responsible employees is crucial [17]. Employees' deviant innovation behavior can inspire other members within the organization, meaning that when one employee demonstrates courage and innovation capability, other members may be encouraged to actively participate in innovation activities. This internal motivation can significantly contribute to the improvement of organizational innovation performance, thereby enhancing internal cohesion within the company. However, deviant innovation should be controlled within a reasonable range, as only moderate deviant innovation is suitable for the long-term development of the enterprise.

To sum up, the interactive game between employees' deviant innovation behavior and enterprises' enhancement of organizational innovation performance is a complex and important theme. Employees' positive deviant innovation behavior can stimulate the enthusiasm within the organization, and at the same time can help the enterprise to better cope with the market challenges. Organizations need to manage this interaction intelligently to ensure that resources are allocated appropriately to achieve optimal innovation performance. This interactive game is one of the key factors for organizational success and requires thoughtful and effective strategic planning. Only by running in the same direction, employees and organizations will work together to reach higher peaks of innovation.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

HZ: Conceptualization, Methodology, Writing–original draft, Writing–review and editing. JL: Data curation, Software, Validation, Writing–original draft, Writing–review and editing. YC: Data curation, Software, Validation, Writing–review and editing. YG: Writing–review and editing. ZZ: Writing–review and editing.

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The handling editor DH declared a shared affiliation with the authors HZ and YC at the time of review.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Zhang LY, Wu MY, Li SY, Liu R, Zhu Y. Exploring the dynamics of role transition of employees in family businesses through the evolutionary game theory. *Front Phys* (2023) 11:1295646. doi:10.3389/fphy.2023.1295646

2. Zhang Q, Zhang HF. A study on the impact of employee's constructive deviant behavior on their innovation performance based on the perspective of manager's meaning construction. *J Xidian Univ* (2023) 33(04):30–46.

3. Jiang Y. A review of research on employee transgressive innovation behavior and its prospects. *Sci Tech Management Res* (2018) 38(10):131–9. doi:10.3969/j.issn.1000-7695.2018.10.018

4. Ye CJ, Liu AL. Leaders' bottom-line mindfulness and employees' transgressive innovation-a moderated mediation mode. *Sci Tech Management Res* (2023) 43(12):176–82. doi:10.3969/j.issn.1000-7695.2023.12.020

5. Li YZ. The impact of transformational leadership on employee innovation behavior. The role of psychological empowerment and emotional commitment. *Res Management* (2018) 39(07):123–30. doi:10.19571/j.cnki.1000-2995.2018.07.015

6. Kai-Ineman DANIEL, Tversky A. Prospect theory: an analysis of decision under risk. *Econometrica* (1979) 47(2):363–91. doi:10.2307/1914185

7. Smith JM. Evolution and the theory of games. In: *Did Darwin get it right? Essays on games, sex and evolution*. Boston, MA: Springer US (1982). p. 202–15.

8. Amir R, Evstigneev VI, Schenk-Hoppé KR. Asset market games of survival: a synthesis of evolutionary and dynamic games. *Annals of Finance* (2013) 9(2):121–144. doi:10.1007/s10436-012-0210-5

9. Eid MS, El-Adaway IH, Coatney KT, Evolutionary stable strategy for postdisaster insurance: game theory approach. *J Management Eng* (2015) 31(6):04015005. doi:10.1061/(asce)me.1943-5479.0000357

10. Feng M, Hu YF. A cross-level study of job stressors on employees' breakthrough and progressive creativity. *J Management* (2021) 18(07):1012–21. doi:10.3969/j.issn.1672-884x.2021.07.007

11. Welbourne TM, Johnson DE, Erez A. The role-based performance scale: validity analysis of a theory-based measure. *Acad Manag J* (1998) 41(5):540–55. doi:10.5465/256941

12. Wang L, Schuetz CG, Cai D. Choosing response strategies in social media crisis communication: an evolutionary game theory perspective. *Inf Management* (2021) 58(6):103371. doi:10.1016/j.im.2020.103371

13. Friedman D. Evolutionary games in economics. *Econometrica* (1991) 59:637–66. doi:10.2307/2938222

14. Du PC, Fan MJ, Liu SY. The effect of job remodeling on employees' adaptive performance under the perspective of information processing theory. *J Capital Univ Econ Business* (2022) 24(02):101–12. doi:10.13504/j.cnki.issn1008-2700.2022.02.008

15. Roland Michel Déprez G, Battistelli A, Boudrias JS, Cangialosi N. Constructive deviance and proactive behaviors: two distinct approaches to change and innovation in the workplace. *Le travail humain* (2020) 83(3):235–67. doi:10.3917/th.833.0235

16. Wong SI, Kuvaas B. The empowerment expectation–perception gap: an examination of three alternative models. *Hum Resource Management J* (2018) 28(2):272–87. doi:10.1111/1748-8583.12177

17. Zhang H, Liu SP. Transformational leadership, employee responsibility and transgressive innovation behavior. *J Southwest Univ Polit Sci L* (2020) 22(02):140–51. doi:10.3969/j.issn.1008-4355.2020.02.12

Game theory based maritime area detection for cloud–edge collaboration satellite network

Yuan Li[1,2], Bingqian Wang[3], Yueqiang Xu[3] and Haitao Xu[3]*

[1]Department of Electronic Engineering, Tsinghua University, Beijing, China, [2]Beijing Spaceiot Technology Development Co., Ltd., Beijing, China, [3]University of Science and Technology Beijing, Beijing, China

Maritime area detection technology applies equipment such as high-orbit satellites, gateway ships and Unmanned Aerial Vehicles to detection. In this scenario, real-time uploading and analysis of maritime data is crucial. In the existing scenario, UAV data are gathered to the gateway ship and uploaded to the shore-based cloud via the high-orbit satellite, because the communication distance of the high-orbit satellite is far, and when the uploaded data volume is large or the access to the equipment increases, the propagation delay of the uploading of the data from the gateway ship to the satellite and the forwarding of the data from the satellite to the shore-based cloud is longer, and the processing delay of the shore-based cloud is increased, and the efficiency of the data transmission and communication will be affected as well. Aiming at the problem of increasing delay caused by communication limitations in maritime area detection, this paper proposes a maritime area detection scheme based on cloud-side collaboration. The scheme solves the problem of communication limitation from the following two aspects. First, the edge computing nodes are deployed on the ship side of the gateway, and the optimal offloading ratio is sought through game theory to offload a part of the tasks from the center cloud to the edge cloud for processing, which improves the efficiency of processing data and thus reduces the data transmission latency and data processing delay. Secondly, low-orbit (LEO) satellites are introduced to provide communication services, because low-orbit satellites have low orbital altitude and short propagation delay, which can transmit the data at the gateway ship to the shore-based cloud more quickly and improve the data transmission efficiency. Finally, it is also verified by designing experiments that the proposed scheme adopts the optimal offloading ratio and has a lower total delay than the original scheme, thus proving the effectiveness of the proposed scheme.

KEYWORDS

game theory, cloud edge collaboration, maritime regional exploration, low-orbit satellites, edge computing

1 Introduction

Timely mastery of the enemy's accurate real-time military intelligence is the key to the navy's military decision-making. UAV-based sea area detection technology has the advantages of wide detection area, high real-time intelligence acquisition, low detection cost, etc., which can comprehensively enhance the navy's ability to detect the sea area, and the technology has been widely used in sea area detection [1, 2].

The maritime area detection mission requires the shore-based cloud to send the maritime area target detection mission to the gateway ship through high-orbit satellite

communication, and the UAV-carrying gateway ship receives the mission instruction and then sends the UAV to detect the formulated area at sea [3, 4]. The UAV arrives at the designated position and altitude to carry out the detection and returns the detection data, including video, infrared images, radar detection signals, and data collected by the mounted sensor equipment, etc., and each UAV has a fixed collection frequency. Among them, ultra-short wave communication is used between the gateway ship and the UAV. The gateway ship summarizes the data collected by the drones, after which the data is transmitted back to the shore-based cloud center via satellite for data processing, and the powerful arithmetic capability of the cloud center is used to realize functions such as pattern recognition and trend prediction [5]. Through the above process, the UAV can quickly and efficiently complete the detection task, provide timely decision support for the command center, and assist the commander in quickly carrying out the subsequent task arrangement [6].

Currently, high-orbit satellite communication is used between the gateway ship and the shore-based cloud terminal [7]. However, as high-orbit satellites are far from the ground, they need to overcome a large spatial distance when transmitting data, resulting in long propagation delays for the uploading of data from the gateway ship to the satellite and for the forwarding of data from the satellite to the shore-based cloud. This may reduce the real-time nature of the data, which in turn affects the decision-making ability of the command center. Second, the current communication scheme needs to upload all the data collected by the UAVs pooled by the gateway ship via satellite to the shore-based cloud for processing. If the amount of data to be uploaded is large, it will increase the transmission delay of data uploading, which increases the pressure on the network bandwidth, making the utilization of limited network resources somewhat challenged, and the data may face the risk of being intercepted or tampered with in the transmission process. This may affect the decision-making efficiency of the command center and the security of the transmitted data. In addition, the efficiency of data transmission and communication may also be affected in the case of increased access to devices. In addition, when uploading data from gateway ships to satellites and forwarding data from satellites to shore-based clouds, the communication bandwidth of high-orbit satellites is relatively limited, and especially when processing a large amount of data, transmission bottlenecks may be encountered, which affects the speed and stability of data transmission. Based on the above, the use of high-orbit satellites for maritime area exploration missions, where a large amount of data is transmitted from the gateway ship via high-orbit satellites to the shore-based cloud for complex processing and analysis, affects the temporal complexity and spatial complexity of the data processed by the shore-based cloud and affects the effectiveness and validity of the data.

In this paper, we have conducted an in-depth study on the maritime area detection scheme and proposed a cloud-edge cooperative-based maritime area detection scheme, which solves the existing communication limitation problem from two aspects, including the deployment of edge computing nodes on the ship side of the gateway. In addition, to shorten the communication distance and reduce the latency, the new scheme also replaces the high-orbit satellites with low-orbit satellites. The main contribution of this paper:

1) First, a detailed description of the cloud-side collaboration-based maritime area detection scheme is presented, including two aspects to solve the problem of increased latency due to limited communication. The latency of each process is calculated from data acquisition to decision support based on data processing results.

2) Secondly, the performance of the proposed scheme is verified, and the latency is selected as the verification metric and the latency of each part of the service processing is calculated. The effectiveness of the improved scheme is verified by comparing the latency of the original scheme and the improved scheme.

The whole paper is organized as follows. The related work is summarized in Section 2. The system model is given in Section 3 and the experiment is given in Section 4. The conclusions are presented in section 5.

2 Related work

The maritime area detection task requires fast and efficient uploading of the data collected by the UAVs pooled on the ship side of the gateway to the shore-based cloud for data processing and target identification, to ensure that the whole process of maritime UAV detection can quickly and efficiently complete the detection task, to provide timely decision-making support for the command center, and to assist the commander in quickly carrying out the subsequent task arrangements. Therefore, the requirements for communication and computing capability are very high.

Aiming at the communication-computing resource collaboration problem, Ref. [8] proposed a computation offloading framework to offload computation or data to nearby devices. To reduce energy and time consumption, a multi-objective computation offloading method for cloud-edge on-dispersion in the Internet of Vehicles (IoV) is proposed to shorten the processing time of computation tasks by offloading them to edge computing devices [9]. An efficient computation offloading and resource allocation scheme in edge computing networks for the Internet of Vehicles is proposed to achieve low complexity and significantly improve system performance [10]. A collaborative computation offloading scheme is proposed to shorten the total service time for the problem of limited computation capability of underwater sensor nodes [11]. The authors in the paper [12] proposed a heterogeneous edge–cloud computing framework, in which a novel collaborative offloading scheme was designed to minimize the task execution latency. In Ref. [13], the authors formulated an optimization problem for edge–cloud computing networks, where the service placement and the power control of the base station were jointly considered, to enhance the resource efficiency.

For the cloud-network convergence scenario, considering that the general cloud center is far away from the end side and there is a latency problem. Therefore, it is necessary to offload the data and computational tasks of maritime area detection to the edge side for processing collaboratively to form a cloud-side collaboration model to reduce the latency. A cloud-edge collaboration approach is proposed to perform maintenance of critical vehicle components, which improves maintenance efficiency and safety [14]. A cloud-edge collaboration approach is proposed to improve the availability

FIGURE 1
Cloud-side collaboration-based maritime area detection scheme.
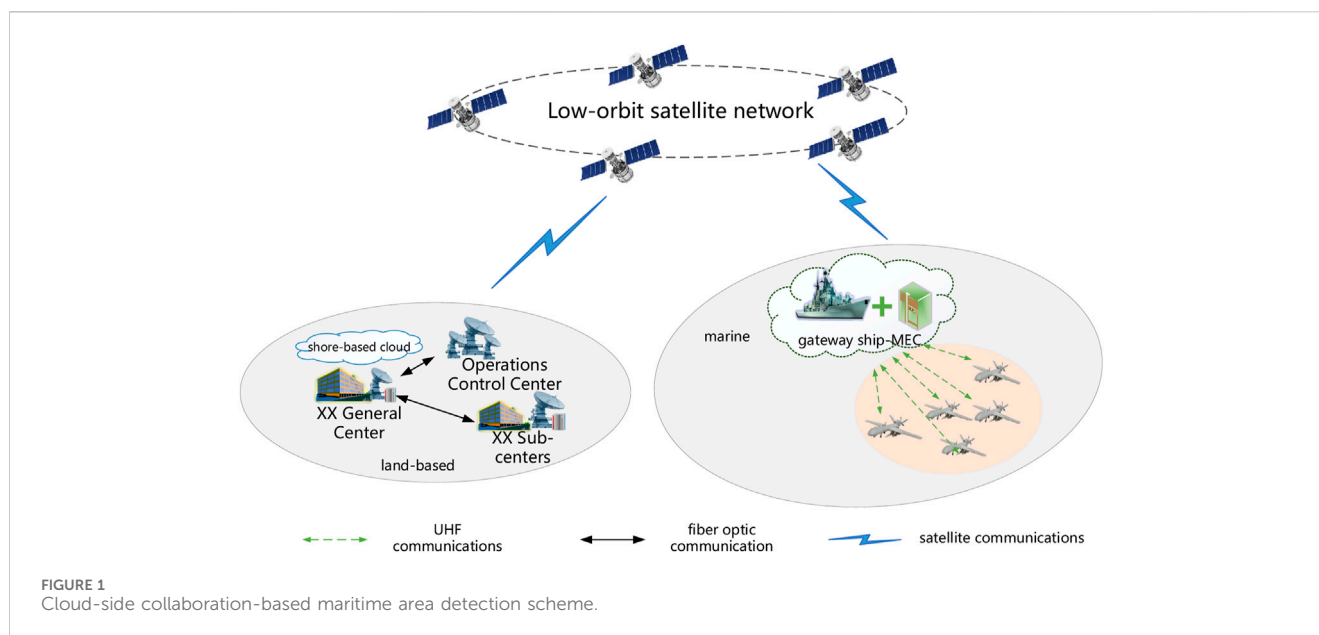
and effectiveness of safety supervision and inspection of power operations [15]. In MEC-enabled networks, Ref. [16] presented a two-level bargaining-based incentive mechanism for collaborative computing, aiming to maximize the offloading utilities of both edge cloud and central cloud. The authors in the paper [17] investigated a dynamic offloading scheme for an edge-cloud collaboration system and formulated an optimization problem to minimize the average cost under a series of constraints. In the paper [18], a collaborative cloud-edge computation offloading method was designed by considering both computing and transmission energy consumption.

3 System model

Aiming at the communication limitation problem faced by the current maritime area detection, this paper proposes a maritime area detection scheme based on cloud-side collaboration. The scheme solves the problem of increased latency caused by communication limitations from the following two aspects. First, the edge computing nodes need to be deployed on the gateway ship side, so that when the gateway ship collects the data collected by the UAV, it can process some of the data, and therefore reduce the amount of data transmission from the gateway ship to the cloud center as a way to alleviate the communication limitation problem. Secondly, the introduction of low-orbit satellites in this program, low-orbit satellites are close to the ground, which can greatly reduce the data propagation delay, and low-orbit satellites have small path loss and high data transmission rate, which can further alleviate the problem of communication limitation. The core idea of the program is to use arithmetic resources to make up for the problem of limited bandwidth resources. It should be noted that the data processing capacity of the gateway ship side is also limited, so this scheme also faces the same problem of limited arithmetic power. If the gateway ship processes the data completely locally, it may cause the local processing delay of the gateway ship side to increase due to the limited processing capacity of the gateway ship side, which

ultimately affects the processing delay of the whole task. If the gateway ship offloads the data completely to the cloud, it may cause a delay increase problem due to the limitation of network bandwidth, so it is necessary to offload part of the data measured by the gateway ship to the cloud for processing, which also indicates that the arithmetic power collaboration between the cloud center and the gateway ship is necessary.
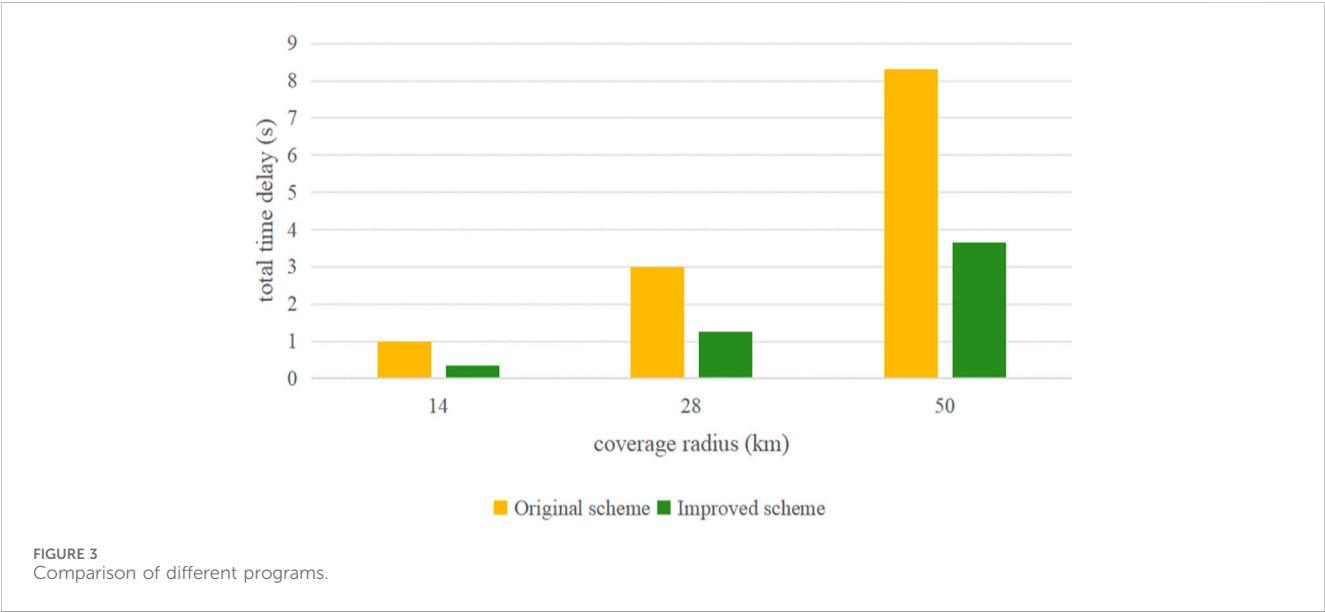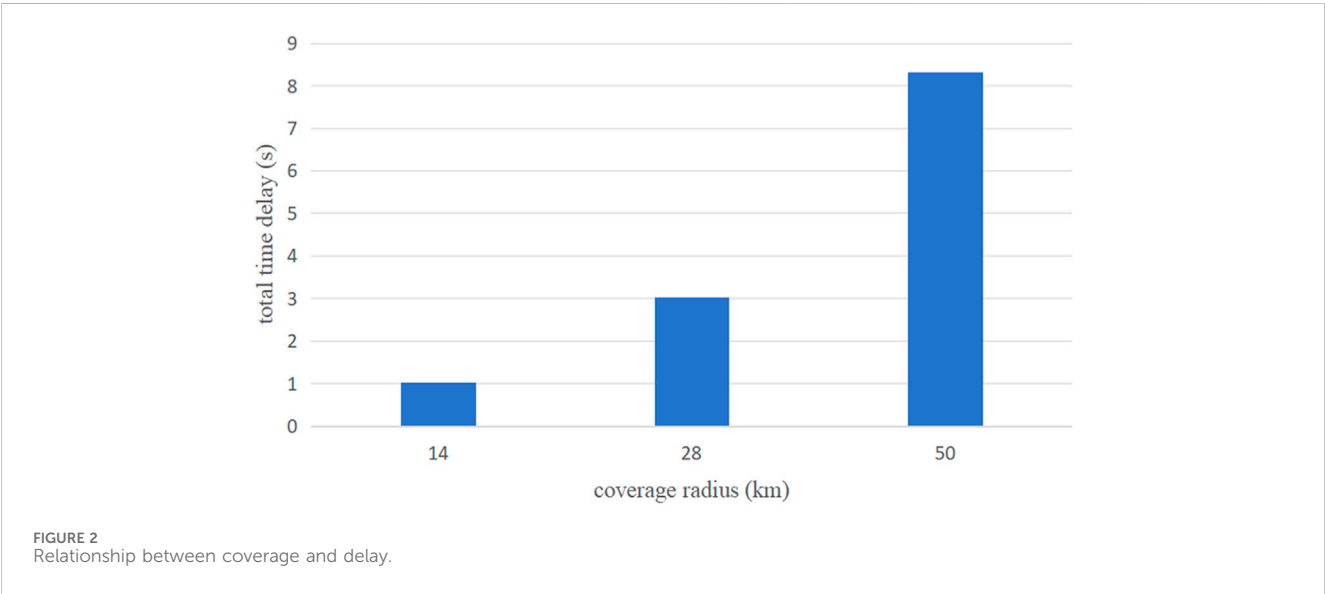
The scenario of the program is shown in Figure 1. In it, the UAV can be used as a terminal, whose task is to collect relevant monitoring data by utilizing various types of sensor devices it carries. The gateway ship side can be viewed as an edge cloud by deploying edge servers, and therefore can provide part of the data processing and storage capabilities. The LEO satellite is mainly used as a relay for data transmission to realize the data communication between the shore-based cloud and the gateway ship at sea. The cloud center is located at the shore, which can provide powerful data processing and storage capabilities. Therefore, the cloud-side collaboration in this scenario refers to the collaborative data processing between the gateway ship and the shore-based cloud center.
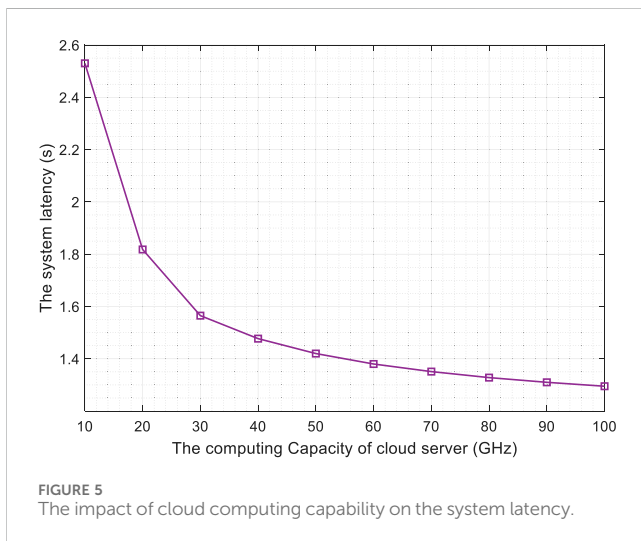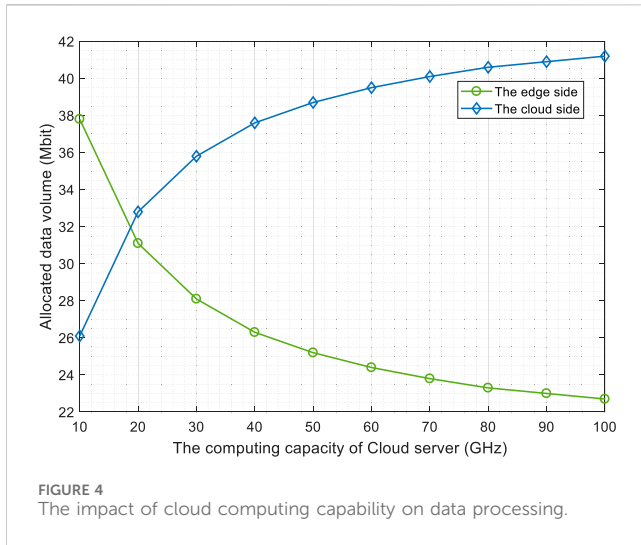
The following is the business processing flow of the maritime area detection scheme based on cloud-side collaboration, which mainly consists of six parts: 1) the gateway ship collects the data collected by each UAV through UHF communication; 2) the gateway ship calculates the optimal offloading scheme and divides the total collected data into two parts, and one part of the data completes the data processing on the local gateway ship side; 3) the gateway ship forwards the remaining data to the remote cloud center through the low orbit satellite network; 4) the ground cloud center processes the data sent by the gateway ship; 5) the gateway ship completes the data processing and sends the results to the ground cloud center through the low-orbit satellite network; 6) the cloud center synthesizes the results of the processing of the various parts and gives the final decision.

First, the gateway ship needs to collect data from the UAVs, and the amount of data collected by each UAV D_{uav} is constant. Let

TABLE 1 Parameters of simulation.

Parameters	Value
The data transmission rate from the UAV to the gateway ship	100 Mbps [23]
The required CPU cycles during the computing process	500 cycle/bit [24]
The computation capacity of the gateway ship	10 GHZ [25]
The computation capacity of the cloud center	100 GHZ [26]
The uplink transmission rate from the gateway ship to the satellite	80 Mbps [27]
The downlink transmission rate from the gateway ship to the cloud center	100 Mbps [28]
The number of UAVs	[1, 16]

FIGURE 2
Relationship between coverage and delay.

FIGURE 3
Comparison of different programs.

FIGURE 4
The impact of cloud computing capability on data processing.

FIGURE 5
The impact of cloud computing capability on the system latency.

$R_{wireless}^{uav}$ be the transmission rate between the UAV and gateway ship. Therefore, the transmission delay for the UAV to upload the data to the gateway ship is:

$$T_{uav} = \frac{D_{uav}}{R_{wireless}^{uav}} \tag{1}$$

Ultra-short wave communication is used between the UAV and the gateway ship, and its data propagation delay is negligible due to the relative proximity of the UAV and the gateway ship.

The gateway ship collects data from multiple UAVs, and N is the total number of UAVs, at which point the total amount of data pooled by the gateway ship is $D_{total} = N \times D_{uav}$. In the cloud-edge collaboration scheme, to improve the efficiency of processing data, both the gateway ship and the cloud center need to process a certain percentage of data, and the amount of data processed by the cloud center is recorded as D_{cloud} , and the amount of data processed by the edge of the gateway ship is D_{edge}. This offloading scheme is meaningful only if the cloud center and the gateway ship complete the data processing at the same moment as far as possible, and for this reason, it is proposed that the delay of

processing data at the edge of the gateway ship is equal to the total delay of offloading part of the data to the cloud for processing to have a higher efficiency. Namely:

$$\frac{D_{edge} \times \eta}{F_{edge}} = \frac{D_{cloud}}{R_{uplink}^{satellite}} + \frac{D_{cloud}}{R_{downlink}^{satellite}} + T_{progation}^{satellite} + \frac{D_{cloud} \times \eta}{F_{cloud}} \tag{2}$$

Where η is the number of cycles required for 1bit during data processing, and F_{edge} is the computing power of the gateway ship, and F_{cloud} is the computing power of the cloud center, and $R_{uplink}^{satellite}$ is the uplink data transfer rate from the gateway ship to the satellite, and $R_{downlink}^{satellite}$ is the downlink data transfer rate from the satellite to the center cloud. $T_{propagation}^{satellite}$ is the propagation delay of the data from the gateway ship to the gateway station using the LEO satellite.

Since LEO satellites are located at a distance of about D_{LEO} The propagation speed of the communication from the satellite is, then the propagation delay between the gateway ship and the gateway station $T_{propagation}^{satellite}$ is:

$$T_{propagation}^{satellite} = \frac{2 \times D_{LEO}}{V} \tag{3}$$

Simplify Equation 2 and define the unloading factor as C_{edge}, C_{edge} is the ratio of the amount of data processed at the edge of the gateway ship to the amount of data processed at the cloud center, and C_{edge} reflects the allocation of computational tasks between the gateway ship and the cloud center during data processing:

$$C_{edge} = \frac{D_{edge}}{D_{cloud}} = \left(\frac{1}{R_{uplink}^{satellite}} + \frac{1}{R_{downlink}^{satellite}} + \frac{\eta}{F_{cloud} \times 10^3} \right) \times \frac{F_{edge} \times 10^3}{\eta} \tag{4}$$

Where 10^3 is the unit matching factor. After the above derivation, the amount of data to be processed at the gateway ship is:

$$D_{edge} = \frac{C_{edge}}{1 + C_{edge}} \times D_{total} \tag{5}$$

Therefore, the data processing delay on the ship side of the gateway is:

$$T_{edge}^{computing} = \frac{D_{edge} \times \eta}{F_{edge}} \tag{6}$$

The gateway ship needs to send the data of size $D_{total} - D_{edge}$ to the cloud for processing. Considering the low-latency performance advantage of the low-orbit satellite network, the program adopts the low-orbit satellite as a relay for data transmission, and the low-orbit satellite adopts the data transmission method of Ka-band with a transmission rate of $R_{uplink}^{satellite}$ The data transmission delay between the gateway ship and the satellite is thus derived.

$$T_{uplink}^{satellite} = \frac{D_{total} - D_{edge}}{R_{uplink}^{satellite}} \tag{7}$$

The satellite needs to send the data to the cloud center via downlink with a transmission rate of $R_{downlink}^{satellite}$ The data transmission delay is:

$$T_{downlink}^{satellite} = \frac{D_{total} - D_{edge}}{R_{downlink}^{satellite}} \tag{8}$$
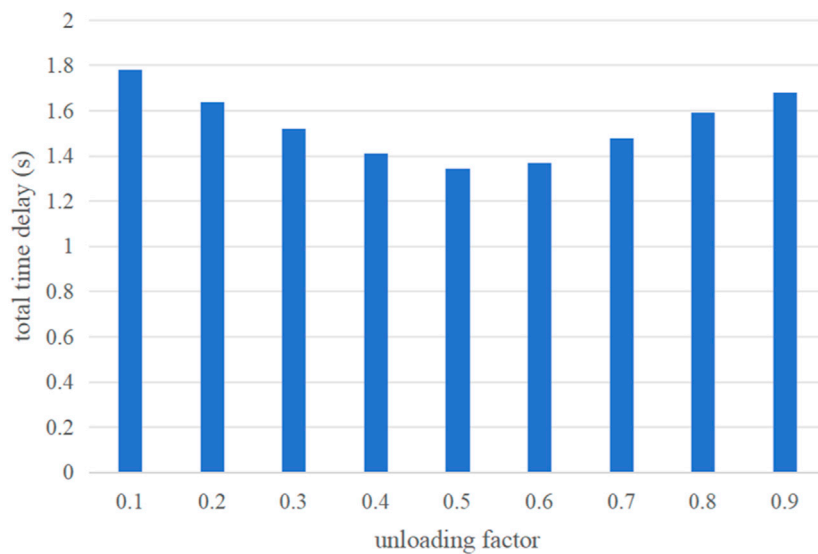
FIGURE 6
Performance comparison of schemes with different unloading ratios.

There is still a propagation delay in the above process, which has a value of approximately $T_{propagation}^{satellite}$. Based on the above analysis, the total delay of sending data from the gateway ship to the cloud can be derived and its value is:

$$T_{edge}^{offloading} = T_{uplink}^{satellite} + T_{downlink}^{satellite} + T_{propagation}^{satellite} \qquad (9)$$

The center cloud receives the data and starts processing the data with its data processing latency:

$$T_{cloud}^{computing} = \frac{\left(D_{total} - D_{edge}\right) \times \eta}{F_{cloud}} \qquad (10)$$

Note that after the gateway ship has processed some of the data, it needs to transmit the result to the cloud center, whose latency can be ignored because the amount of data to process the result is very small. Considering that the gateway ship and the center cloud can process the user's data in parallel, the total delay of their task processing is:

$$T_{total} = T_{uav} + \max\left(T_{edge}^{computing}, T_{edge}^{offloading} + T_{cloud}^{computing}\right) \qquad (11)$$

4 Experimental verification

In this section, the validity of the proposed scheme is verified. Firstly, the simulation parameters are set and then each validation process is described. Finally, the effectiveness of the proposed algorithm is verified by comparing the experimental results.

4.1 Analog setup

In this section, we will give some common simulation parameter settings. Specifically, the radius of the detection range of the UAV is about 14 km [19], take the radius of the detection range of 28 km as

an example, 4 UAVs are needed to perform the detection task, the data volume collected by the UAV is about 16 Mbit [20]. The distance from the ground to the low-orbit satellite is about 1,150 km [21], and the propagation speed of the satellite communication is 3×10^8 m/s [22]. The more detail of simulation parameters is given in Table 1.

4.2 Verification process

Based on the above description of the scheme, the performance of the proposed scheme needs to be validated, and here the delay is selected as the validation index, so it is necessary to calculate the delay of each part of the business processing. The validation process of the maritime area detection scheme based on cloud-edge collaboration is given below, which mainly includes the following steps:

Calculating the data collection latency: the communication between the UAV and the gateway ship is based on ultra-shortwave communication, and the communication rate is known. Each UAV has a fixed detection range, the video collected by the UAV is determined based on a combination of the collected image resolution (pixels), color depth (bits), and frame rate, and the amount of data collected by the UAV is also known. Based on the above information, the transmission delay for the UAV to upload the data to the gateway ship can be calculated. The delay is the ratio of the data volume to the transmission rate.

Calculating the optimal offloading ratio: In the cloud-edge collaboration scheme, in order to improve the efficiency of processing data, both the gateway ship and the cloud center need to process a certain proportion of the data, and this offloading ratio is optimal only if the cloud center and the gateway ship complete the data processing at the same moment as far as possible, for this reason, it is necessary to adjust the offloading ratio so that the latency of the data processing at the edge of the gateway ship is equal to the

total latency of offloading part of the data to the cloud for processing. To have a higher efficiency.

Calculation of data offloading delay: Gateway ships need to upload part of the data via LEO satellites to the shore-based cloud for processing. It should be noted that the gateway ship needs to take into account the delay of each part of the satellite in data offloading, including the uplink delay, downlink delay and propagation delay. Among them, the uplink delay is the delay between the gateway ship and the LEO satellite, which can be expressed as the ratio of the unloaded data volume to the uplink transmission delay; the downlink delay is the delay between the LEO satellite and the cloud center, which can be expressed as the ratio of the unloaded data size to the downlink transmission rate; the propagation delay is related to the distance between the LEO satellite and the gateway ship, and the delay can be expressed as the ratio of the distance to the speed of light. In summary, the offloading delay of the data can be derived, which is the sum of the above three parts of the delay.

Calculating the data processing delay: both the gateway ship and the cloud center need to process some of the data. In data processing, images need to be recognized, analyzed, etc. For this task type, it is necessary to determine the arithmetic resources required for each unit of the data volume of the task, therefore, after the size of each part of the processed data is known, its data processing delay can be expressed as the ratio of the total computational resource demand of the task to the capacity of the computational node. Where the total computational resource requirement is the product of the total data volume of the task and the arithmetic resources required per unit of data volume.

Calculate the total processing delay of the task: note that after the gateway ship has processed some of the data, it needs to transmit the result to the cloud center, and its delay can be ignored because the amount of data to process the result is very small. Considering that the gateway ship and the center cloud can process the user's data in parallel, the total delay of its task processing should be larger of the total data processing delay of the gateway ship side and the total data processing delay of the cloud center. Among them, the total data processing delay on the gateway ship side is mainly the delay of the gateway ship utilizing local arithmetic power for data processing. The total data processing delay of the cloud center includes the data offloading delay and the data processing delay of the cloud center.

4.3 Verification results

The UAV performs the maritime area detection mission by first using the sensors and equipment carried by the UAV itself to effectively and comprehensively detect the target area, then pooling the collected data through the gateway ship and uploading the data to the shore-based command post for data processing through satellite communications to obtain information about maritime activities for to provide decision support to the command center. To verify the effectiveness of the regional detection program based on cloud-side collaboration, the following objectives need to be verified.

1) **Coverage range and latency:** as the detection range increases, the processing latency of the data increases, the main reason for this is that a larger detection range needs to be accomplished by more than one UAV, and with the increase in the number of UAVs, the gateway ship-side collection of data from the UAVs increases and is sent to the shore-based cloud, and this process increases the processing latency of the data. To validate the above process, it is necessary to increase the coverage range and then calculate the total latency of the business processing, observe the relationship between the coverage range of the detection area and the latency, and expect to achieve the following results, as shown in Figure 2.

2) **Performance comparison of different schemes**: to verify the effectiveness of the proposed area detection scheme based on cloud-side collaboration, it is necessary to analyze the data processing latency under different schemes. For this purpose, different target area coverage radii need to be adjusted to calculate the total delay of business processing, and the following results are expected to be achieved, as shown in Figure 3. The original scheme refers to the gateway ship pooling the data collected by UAVs and then transmitting them to the cloud center for data processing via high-orbit satellites; the improved scheme refers to the scheme proposed in this paper, i.e., deploying an edge computing node on the side of the gateway ship, completing the data processing of a portion of the data collected by the gateway ship at the local gateway ship through the optimal offloading ratio, and the remaining data being forwarded to the cloud center through the low-orbit satellite network for data processing. Under the same coverage, the proposed scheme greatly reduces the data processing latency and meets the latency requirements for monitoring different ranges, thus saving valuable time for detection.

3) **Cloud computing capability and data processing:** we investigate the relationship between the performance of cloud servers and the amount of data processed by the servers. We assume that the computing power of edge nodes remains unchanged. With the improvement of cloud computing capabilities, the total processed data volume of the cloud will continue to increase, but the total processed data volume of edge nodes will continue to decrease, as shown in Figure 4. We can conclude that the computing power of nodes affects the proportion of data processing. This is because as cloud computing capabilities improve, the computing performance advantage of edge nodes will gradually decrease.

4) **Cloud Computing Capability and System Latency:** As shown in Figure 2, we analyze the impact of cloud computing capability on system latency. We fix the data processing capability of edge nodes. With the improvement of cloud computing capabilities, the data processing latency of the system will continue to decrease and tend toward balance, as shown in Figure 5. The main reason is that the latency of data processing is inversely proportional to the data processing capability, but the latency of data processing is also affected by network transmission performance. When the computing performance reaches a certain level, network transmission capacity will become the main factor affecting the total latency.

5) **Optimal task offloading ratio**: when designing the cloud-side collaboration scheme, the task processing ratio between the cloud center and the gateway ship is extremely critical, to

verify this conclusion, the detection range can be set to 28 km, the offloading ratio can be adjusted, and the service processing delay under different offloading ratios can be calculated and expected to reach the following results, as shown in Figure 6. From the figure, it can be seen that the proposed scheme can achieve optimal performance when the offloading ratio is about 0.5. If the tasks are all processed on the gateway ship or completely offloaded to the cloud for processing, the monitoring delay requirement may not be guaranteed, which also shows the necessity of cloud-side collaboration.

5 Conclusion

In this paper, we propose a cloud-edge collaboration-based maritime area detection scheme that addresses the current problem of increased latency due to limited communication from two aspects. One of the aspects is to deploy the edge computing nodes on the ship side of the gateway. On the other hand, low-orbit satellites can also be used instead of high-orbit satellites to shorten the communication distance. In this paper, the effectiveness of the maritime area detection scheme based on cloud-edge collaboration is also experimentally verified by using the time delay as a calculation index.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

YL: Writing–original draft, Writing–review and editing, Software, Supervision. BW: Formal Analysis, Writing–original draft. YX: Writing–original draft, Writing–review and editing, Conceptualization, Data curation, Formal Analysis, Methodology. HX: Funding acquisition, Project administration, Writing–original draft.

Funding

Conflict of interest

Author YL was employed by company Beijing Spaceiot Technology Development Co., Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

1. Yuan S, Li Y, Bao F, Xu H, Yang Y, Yan Q, et al. Marine environmental monitoring with unmanned vehicle platforms: present applications and future prospects. *Sci Total Environ* (2023) 858:159741. doi:10.1016/j.scitotenv.2022.159741

2. Wang Z, Yang T, Zhang H. Land contained sea area ship detection using spaceborne image. *Pattern Recognition Lett* (2020) 130:125–31. doi:10.1016/j.patrec.2019.01.015

3. Nie W, Han ZC, Zhou M, Xie LB, Jiang Q. UAV detection and identification based on WiFi signal and RF fingerprint. *IEEE Sensors J* (2021) 21(12):13540–50. doi:10.1109/jsen.2021.3068444

4. Dong Y, Ma Y, Li Y, Li Z. High-precision real-time UAV target recognition based on improved YOLOv4. *Comp Commun* (2023) 206:124–32. doi:10.1016/j.comcom.2023.04.019

5. Canello G, Mignardi S, Mikhaylov K, Buratti C, Hänninen T. Data collection from LoRaWAN sensor network by UAV gateway: design, empirical results and dataset. In: *2023 Joint European Conference on Networks and Communications & 6G Summit (EuCNC/6G Summit)*; July 2023; Gothenburg, Sweden. IEEE.

6. Pirmagomedov R, Kirichek R, Blinnikov M, Koucheryavy A. UAV-based gateways for wireless nanosensor networks deployed over large areas. *Comp Commun* (2019) 146: 55–62. doi:10.1016/j.comcom.2019.07.026

7. Zhou H, Wang X, Zhong S. A satellite orbit maneuver detection and robust multipath mitigation method for GPS coordinate time series. *Adv Space Res* (2023). doi:10.1016/j.asr.2023.08.006

8. Bajaj K, Sharma B, Singh R. Implementation analysis of IoT-based offloading frameworks on cloud/edge computing for sensor generated big data. *Complex Intell Syst* (2022) 8(5):3641–58. doi:10.1007/s40747-021-00434-6

9. Xu X, Gu R, Dai F, Qi L, Wan S. Multi-objective computation offloading for internet of vehicles in cloud-edge computing. *Wireless Networks* (2020) 26:1611–29. doi:10.1007/s11276-019-02127-y

10. Liu S, Tian J, Zhai C, Li T. Joint computation offloading and resource allocation in vehicular edge computing networks. *Digital Commun Networks* (2022) 9:1399–410. doi:10.1016/j.dcan.2022.12.002

11. Liu X, Du X, Zhang S, Han D. Cooperative computing offloading scheme via artificial neural networks for underwater sensor networks. *Appl Sci* (2023) 13(21):11886. doi:10.3390/app132111886

12. Hua W, Liu P, Huang L. Energy-efficient resource allocation for heterogeneous edge-cloud computing. *IEEE Internet Things J* (2023) 11:2808–18. doi:10.1109/jiot.2023.3293164

13. Han P, Liu Y, Zhang X, Guo L. Energy-efficient service placement based on equivalent bandwidth in cell zooming enabled mobile edge cloud networks. *IEEE Trans Vehicular Tech* (2022) 71(11):12275–90. doi:10.1109/tvt.2022.3194543

14. Shen B, Lin J, Jia Y. Data-driven urban rail vehicle critical parts maintenance system based on cloud-edge collaboration. *J Phys : Conf Ser* (2023) 2649(1):012057. doi:10.1088/1742-6596/2649/1/012057

15. Fei Z, Zhou H, Wang K. Cloud-edge collaboration based methods in improving the safety inspection efficiency on electric operation scenes. *J Phys : Conf Ser* (2023) 2584(1):012026. doi:10.1088/1742-6596/2584/1/012026

16. Chen G, Chen Y, Mai Z, Hao C, Yang M, Du L. Incentive-based distributed resource allocation for task offloading and collaborative computing in MEC-enabled networks. *IEEE Internet Things J* (2022) 10(10):9077–91. doi:10.1109/jiot.2022.3233026

17. Hu H, Wu D, Zhou F, Zhu X, Hu RQ, Zhu H. Intelligent resource allocation for edge-cloud collaborative networks: a hybrid DDPG-D3QN approach. *IEEE Trans Vehicular Tech* (2023) 72:10696–709. doi:10.1109/tvt.2023.3253905

18. Su Q, Zhang Q, Li W, Zhang X. Primal-dual-based computation offloading method for energy-aware cloud-edge collaboration. *IEEE Trans Mobile Comput* (2023) 1–15. doi:10.1109/tmc.2023.3237938

19. Li J, Zhang G, Zhang X, Zhang W. Integrating dynamic event-triggered and sensor-tolerant control: application to USV-UAVs cooperative formation system for maritime parallel search. *IEEE Trans Intell Transportation Syst* (2023) 1–13. doi:10.1109/tits.2023.3326271

20. Xu G, Yang Z, Liu H. Pursuit and evasion game between UVAs based on multi-agent reinforcement learning. In: 2019 Chinese Automation Congress (CAC); November 2019; Hangzhou, China. IEEE.

21. Yong D, Song J, Dayang W, Yang B, Huichen X, Jincheng Z. A stk-based constellation architecture implementation for 5g low-orbit satellites. In: *2022 IEEE 4th International Conference on Power, Intelligent Computing and Systems (ICPICS)*; September 2022; Shenyang, China. IEEE.

22. Qi Y, Meng W, Zeng C, He Q, Xu G, Wen X. Influence of Co-frequency interference on transmission performance in satellite communication. In: *2023 8th International Conference on Communication, Image and Signal Processing (CCISP)*; November 2023; Chengdu, China. IEEE.

23. Du P, Shi Y, Zeng Q, Zhang X. Joint trajectory design and transmit power control in NOMA-aided UAV communication systems. In: *2022 International Conference on Computing, Communication, Perception and Quantum Technology (CCPQT)*; August 2022; Xiamen, China. IEEE.

24. Farooq MO. Priority-based servicing of offloaded tasks in mobile edge computing. In: 2021 IEEE 7th World Forum on Internet of Things (WF-IoT); June 2021; New Orleans, LA, USA. IEEE.

25. Wang B, Tong F, Huang D. A joint computation offloading and resource allocation strategy for LEO satellite edge computing system. In: 2020 IEEE 20th International Conference on Communication Technology (ICCT); December 2020; Nanning, China. IEEE.

26. Tang Z, Zhou H, Ma T, Yu K, Shen XS. Leveraging LEO assisted cloud-edge collaboration for energy efficient computation offloading. In: *2021 IEEE Global Communications Conference (GLOBECOM)*; February 2021; Madrid, Spain. IEEE.

27. Hui LZ, Yang GQ. Uplink user power control for low-orbit satellite communication systems. In: 2018 IEEE 4th International Conference on Computer and Communications (ICCC); December 2018; Chengdu, China. IEEE.

28. Shi J, Hu J, Yue Y, Xue X, Liang W, Li Z. Outage probability for OTFS based downlink LEO satellite communication. *IEEE Trans Vehicular Tech* (2022) 71(3): 3355–60. doi:10.1109/tvt.2022.3144466

Analysis of regional carbon productivity differences and influencing factors—based on new green decomposition model

Min Fu[1,2], Ying Mei[1,3], Lixin Tian[1,2]* and Chao Zhang[1,3]

[1]Research Institute of Carbon Neutralization Development, School of Mathematical Sciences, Jiangsu University, Zhenjiang, China, [2]Jiangsu Province Engineering Research Center of Spatial Big Data, School of Mathematical Sciences, Nanjing Normal University, Nanjing, China, [3]Jiangsu Province Engineering Research Center of Industrial Carbon System Analysis, School of Mathematical Sciences, Jiangsu University, Zhenjiang, China

This paper introduces a new green decomposition model of carbon productivity that aims to further analyze the regional differences in carbon productivity and its interaction with regional industrial performance. We combine desired outputs and undesired outputs orientation, and multiple factor inputs to derive a new green decomposition theorem, establish a new green decomposition model of carbon productivity, and obtain nine effects of regional carbon productivity differences. Empirical analysis is conducted using input-output data from 29 provinces and 15 industries in China, comparing the differences in carbon productivity from both the provincial and industry perspectives and exploring the mechanism of action. This paper provides theoretical basis and empirical evidence for regional carbon productivity enhancement and economic and industrial optimization from the perspective of multi-factor inputs, as well as policy insights for regional low-carbon transition development.

1 Introduction

Since 2006, China has become the largest carbon-emitting country in the world. According to data released by the International Energy Agency (IEA), China's carbon dioxide emissions in 2021 were approximately 1.014 billion tons, accounting for around 27% of total global emissions, and its carbon dioxide emissions in 2022 are expected to remain at a similar level to those of 2021. China is faced with the dual constraints of economic growth and environmental protection [1], with underlying contradictions such as insufficient innovation capability, worsening environmental pollution, and regional development disparities [2]. Therefore, China has adopted carbon-emission reduction-oriented environmental policies, with a key focus on achieving the proper balance between environmental pollution and industrial performance [3–5]. Clearly, the adoption of carbon-emission reduction-oriented environmental policies requires a more comprehensive understanding of the interaction between carbon productivity and industrial performance, as well as a more in-depth search for the driving factors of carbon productivity.

However, traditional carbon productivity decomposition models often ignore undesirable outputs, making it difficult to efficiently measure production technology efficiency that includes undesirable outputs [6,7]. Additionally, these models overlook the diversity and complexity of factor inputs in the economic production process [8, 9], especially the critical role that knowledge inputs play in regional carbon productivity disparities [10, 11]. This limits our ability to thoroughly understand the interaction between regional carbon productivity and industrial performance, as well as to develop location-specific policies based on the driving factors and impact of carbon productivity. Against the backdrop of environmental policies aimed at reducing carbon emissions, how do the diversity and complexity of factor inputs and the duality of desirable and undesirable outputs affect regional carbon productivity disparities, and what impact do they have on regional industrial performance?

This paper addresses the aforementioned issues by constructing a new green decomposition model of carbon productivity. First, we refine energy input factors into renewable and non-renewable energy inputs, and draws from [2] and [12] to introduce knowledge stock and green knowledge stock inputs, defining a new green environmental production technology. Using data envelopment analysis (DEA) method, we establish a new linear programming model for green environmental production technology with constant returns to scale. Second, based on the traditional output-oriented Shephard distance function, we define a more practical new green Shephard distance function for desirable and undesirable outputs and their corresponding new green Farrell technical efficiency measures. Third, according to the definition of carbon productivity and input factors under the drive of green development, we define a new green decomposition theorem for carbon productivity and a new green decomposition theorem for carbon productivity embedded in distance function. By obtaining the decomposition equation for carbon productivity and constructing a new green decomposition model of carbon productivity, we have decomposed the regional differences of carbon productivity into 9 influencing factors: the green knowledge efficiency effect of renewable energy inputs, knowledge efficiency effect of non-renewable energy inputs, potential hybrid knowledge emission ratio effect, carbon performance index effect, carbon emissions structure effect, capital-hybrid energy substitution effect, labor-hybrid energy substitution effect, hybrid knowledge stock-hybrid energy substitution effect, and carbon factor effect. According to the definition of carbon productivity and input factors under green development drive, we define the new green decomposition theorem for carbon productivity and the embedded distance function green decomposition theorem for carbon productivity, obtaining the decomposition formula for carbon productivity. The new green decomposition model for carbon productivity is constructed, and the regional differences in carbon productivity are decomposed into 9 effects: green knowledge efficiency effect of renewable energy (*HEEE*), knowledge efficiency effect of non-renewable energy (*GEFE*), potential hybrid knowledge emission ratio effect (*PHGCRE*), carbon performance index effect (*CPIE*), carbon emission structure effect (*CESE*), capital - hybrid energy substitution effect (*KEFSE*), labor - hybrid energy substitution effect (*LEFSE*), hybrid knowledge stock - hybrid energy

substitution effect (*HGEFSE*), carbon factor effect (*CFE*). The empirical part of this paper selects input and output data from 15 industries in 29 provinces, municipalities, and autonomous regions in China in 2019. By using the new green decomposition model of carbon productivity and the expressions of each effect, we compare the differences in carbon productivity and explore the influencing factors behind the differences, enriching the research content at both the regional and industry levels. The research framework for the new green decomposition system of carbon productivity is shown in Figure 1.

The rest of this paper is described as follows; Section 2 provides a literature review on the research progress of carbon productivity methods and compares this paper with previous research; Section 3 constructs a new green decomposition model of carbon productivity, decomposing the differences in regional carbon productivity into 9 influencing factors; Section 4 presents an empirical analysis of the model, comparing and analyzing the differences and influential factors of carbon productivity from single and double dimensions; Section 5 is conclusion; Section 6 is policy recommendations.

2 Literature review

The concept of carbon productivity first appeared in the study by [13], which refers to the economic benefits generated by unit carbon emissions. In 2008, the report released by [14] elaborated on the significance of carbon productivity, indicating that any technology that successfully mitigates climate change must achieve two objectives: firstly, to stabilize greenhouse gas emissions, and secondly, to maintain high-quality and healthy economic development. Carbon productivity effectively combines both goals. Nowadays, carbon productivity has become one of the important indicators that measure the development of a low-carbon economy, attracting attention from policymakers and researchers in various countries and regions worldwide.

2.1 Research methods for carbon productivity

Regarding the research methods of carbon productivity, scholars and experts have done a lot of research. The main methods include Stochastic Frontier Analysis (SFA) [15] and Data Envelopment Analysis (DEA) [16]. The most important feature of SFA is that it considers the effect of stochastic factors on output, while it enables the study of panel data across time periods, resulting in more realistic research outcomes [17–19]. DEA methods are widely used in efficiency measurement research. [20] used a nonparametric approach to measure technical efficiency with constant returns to scale and proposed the environmental DEA technique. However, the traditional DEA model does not account for undesirable outputs, making it challenging to effectively measure the efficiency of production technology that includes them [7, 21]. Currently, there are three primary methods for dealing with undesirable outputs in efficiency analysis models. The first treats them as input variables for research, but this contradicts actual production processes [22, 23]. The second constructs a variable that

FIGURE 1
Framework for the new green decomposition system of carbon productivity.
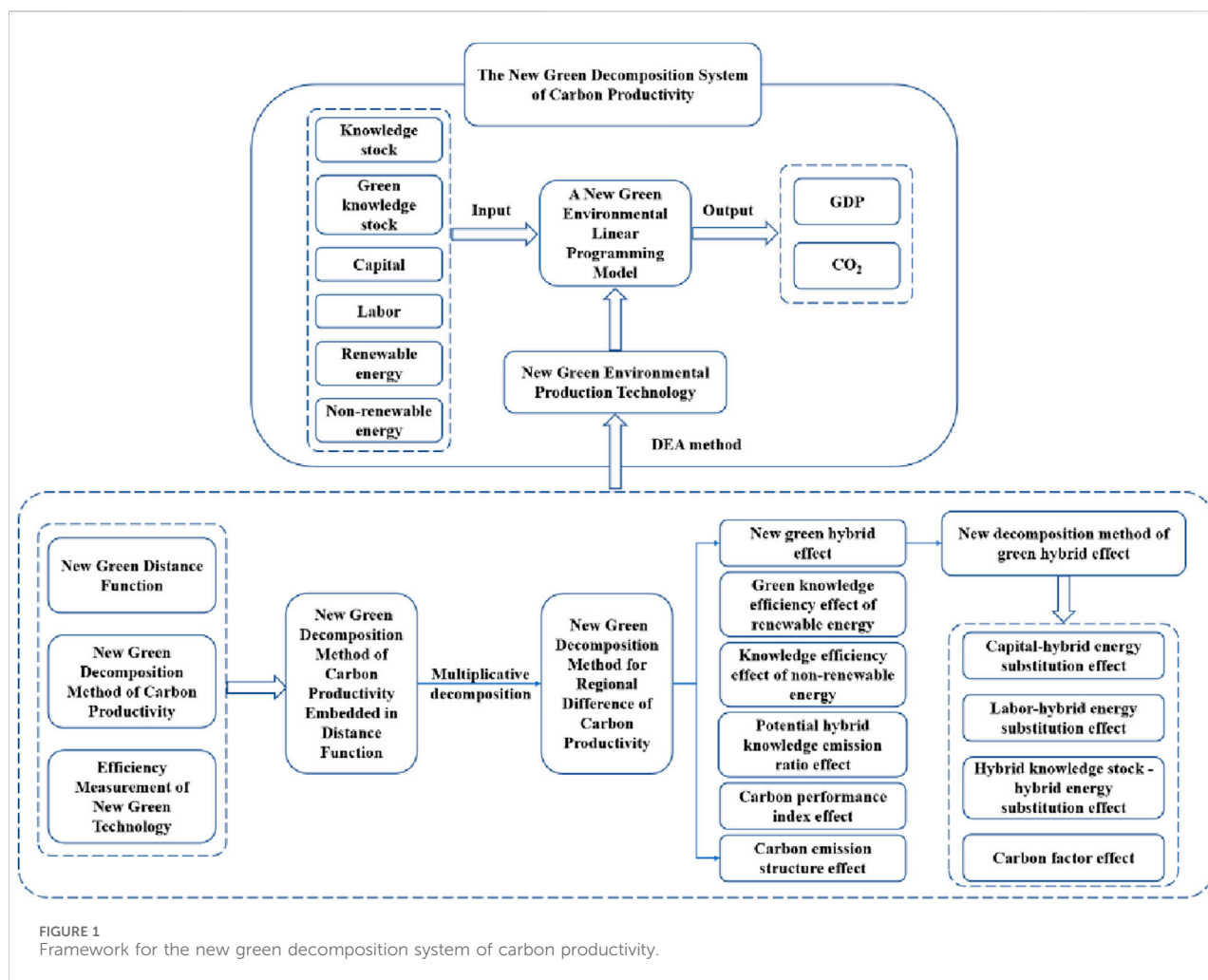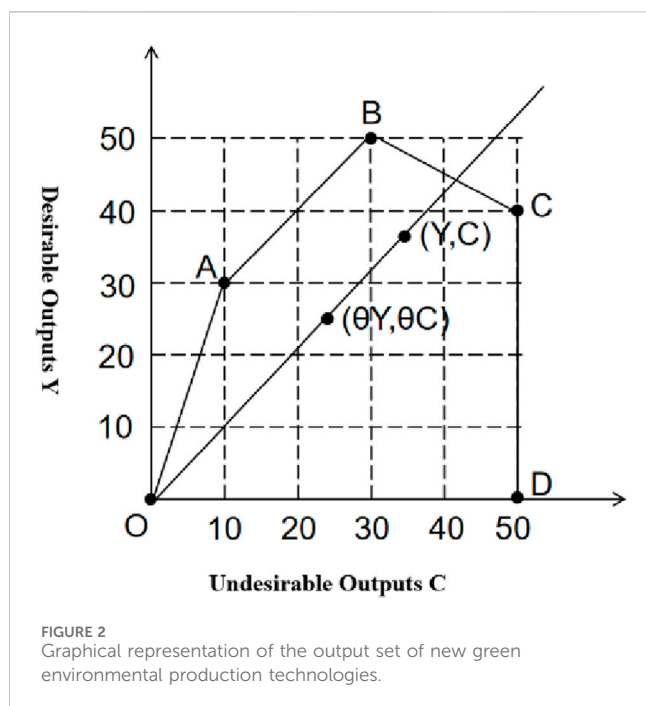
is negatively correlated with undesirable output, and then performs calculations within the traditional efficiency analysis model [24], but this method is limited to variable returns to scale cases. The third approach involves using environmental production technology to measure technical efficiency and incorporating undesirable outputs in the actual production process [25], rather than merely representing it through mathematical functions in abstract terms [5]. In addition to the above two traditional analysis methods, [26] have proposed a new approach to measuring carbon productivity in the generalized input-output model, which aims to construct more reasonable carbon productivity indicators. Thus, a theoretical framework for multiple research methods on carbon productivity has been preliminarily established.

2.2 Analysis methods for influencing factors of carbon productivity

The analysis of factors influencing carbon productivity usually uses econometric modeling and factor decomposition methods. Econometric models effectively combine economic theory and statistical methods to quantitatively analyze the research object [27, 28]. The factor decomposition methods are more widely used

because they can decompose the variation of the research object into the changes of multiple influencing factors. By comparing and analyzing the contribution rates of various influencing factors to the changes of the research object, it can offer more comprehensive insights. Based on the decomposition principle, factor decomposition methods mainly include Structural Decomposition Analysis (SDA) and Index Decomposition Analysis (IDA) methods [8, 29]. Compared to the SDA method, the IDA method is more widely applied at a cross-time and cross-regional level. The IDA method mainly includes Laspeyres index decomposition method and Divisia index decomposition method. The latter consists of Logarithmic Mean Divisia Index (LMDI) method and Arithmetic Mean Divisia Index (AMDI) method. LMDI is capable of handling residual and zero value problems better [30]. LMDI can be divided into LMDI-1 and LMDI-2, which use different weighting formulas. LMDI-1 is simpler, and thus, more widely applied in energy and environmental research [31]. In 1998, [32] introduced a fully decomposable LMDI method. In 2018, [9] used the LMDI method to decompose carbon emissions into six influencing factors and analyzed the top 23 countries in renewable energy use. [33] compared the contributions of different technology-driving factors related to CO_2 emission growth, using the LMDI method based on energy allocation analysis for different

FIGURE 2
Graphical representation of the output set of new green environmental production technologies.

time periods. The diversity of methods for analyzing the factors affecting carbon productivity provides a reference basis and analysis tool for the green decomposition of carbon productivity in this paper.

2.3 Research methods for regional differences in carbon productivity

The comparative research on regional differences of carbon productivity mainly involves two methods: the Theil index and decomposition analysis [34, 35]. The Theil index indicates the level of regional economic disparity, with a larger value indicating greater disparity. [36] effectively combined the decomposition analysis method and environmental production technology, proposing a new Production Decomposition Analysis (PDA) method using Shephard distance function and environmental DEA technology in production theory. [37] quantified the impact of carbon performance index, potential carbon factors, and industrial structure on regional differences in carbon emission intensity using a spatial PDA model based on the production theory framework. [38] used the spatial PDA method and data from various industrial sectors in Chinese provinces to study the regional differences in industrial carbon productivity as well as the driving factors at the provincial and sectoral levels. At the carbon trading and regional carbon productivity level, carbon trading can promote technological progress, factor accumulation, scale allocation and energy substitution effects [39]. [40] provided an in-depth explanation of energy saving and emission reduction in China's energy system in terms of carbon trading and new energy development mechanism, driving mechanism, evolutionary behavior and policy synergy. This provides us with reference for studying a green decomposition system for multiple sectors and regions.

Carbon productivity is a hot topic in the current field of energy and economic research. Scholars and experts mostly analyze the differences in carbon productivity by using existing methods, exploring the impact of different driving factors on carbon productivity. However, there is less research on the driving factors of carbon productivity that take into account the knowledge stock input needed in the production process and the green knowledge stock input, and most studies neglect the critical role of undesirable outputs, as well as the inherent interaction between regions and industries. This paper investigates multiple types of input factors in the economic production process, considering both desirable and undesirable outputs. It constructs a new green decomposition system of carbon productivity and obtains the results of carbon productivity decomposition in both single-dimensional and dual-dimensional dimensions. The comparative analysis of carbon productivity differences at the regional and industry levels from the perspective of inherent interaction provides targeted policy recommendations for low-carbon development in regions and industries.

3 New green decomposition model of carbon productivity

3.1 New green environmental production technology

Suppose there are N regions R_j ($j = 1, \ldots, N$) in the economic system, each of which contains M industries ($i = 1, \ldots, M$). Following the MR model in the spatial decomposition strategy, we take the average of all the regions under evaluation to obtain the reference region R_u as a benchmark for comparison. Therefore, our study includes a total of $N+1$ regions.

In economic production activities, a certain type and amount of production factors can generate economic benefits, or the desirable outputs (referring to GDP added value). However, at the same time, it can also produce some waste and pollutants, or the undesirable outputs (referring to CO_2 emissions). Based on the environmental production technology proposed by [20], this paper endogenizes the production factor input in economic production process in accordance with the requirements of green development [2, 41]. Building on the original energy input, energy input is divided into renewable energy input and non-renewable energy input, where the stock of knowledge input and green knowledge stock input are introduced. Definition 1 is given as follows: The set of new green environmental production technologies.

Definition 1: The set of new green environmental production technologies for industry i in economic production activities is:

$$T_i = \left\{ (K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i): \text{input}(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i), \right.$$
$$\left. \text{output}(Y_i, C_i) \right\} \tag{1}$$

Where, K_i represents the capital input for industry i, L_i represents the labor input for industry i, E_i represents the renewable energy input for industry i, F_i represents the non-renewable energy input for industry i, and the product of E_i and F_i represents the hybrid energy input for industry i. H_i represents the green knowledge stock input for industry i, G_i represents the

TABLE 1 Industry division.

Industrial division	Sub-industry	Symbol
Manufacturing industries	Food, tobacco and wine industry	S1
	Textile and garment industry	S2
	Stationery industry	S3
	Petroleum refineries	S4
	Pharmaceutical industry	S5
	Chemical industry	S6
	Black gold, gold processing and non-gold and metal products industry	S7
	General equipment industry	S8
	Special equipment industry	S9
	Transportation equipment industry	S10
	Electric appliance manufacturing	S11
	Communication equipment, office equipment and other manufacturing industries	S12
Electricity, heat, gas and water production and supply industries	Electricity and heat production and supply industry	S13
	Gas production and supply industry	S14
	Water production and supply industry	S15

knowledge stock input for industry i, and the product of H_i and G_i represents the hybrid knowledge stock input for industry i. Y_i represents the desirable outputs for industry i, referring to the industry added value, while C_i represents the undesirable outputs for industry i, referring to the amount of CO_2 emissions.

According to [42], we present Property 1 of the set of new green environmental production technologies.

Property 1: states that the set of new green environmental production technologies exhibits the following properties:

(1) The production technology demonstrates output neutrality;
(2) The set of new green environmental production technologies is a bounded and closed set;
(3) The input factors have free disposability;
(4) The desirable outputs have free disposability;
(5) The undesirable outputs have weak disposability;
(6) The desirable and undesirable outputs exhibit zero jointness.

Note 1: Explanation of the 6 properties in Property 1:

(1) In the economic production process, the production technology exhibits output neutrality, which means that when the inputs are not all zero, the output can be zero, that is, $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, \ 0, 0) \in T_i$, indicating the decision-making unit may not engage in production activities;
(2) In Eq. 1, for any input vector $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$, the subscript i represents the industry, of which there are only finitely many. A finite quantity of inputs can only produce a finite quantity of outputs. Therefore, the set of new green environmental production technologies in production theory is a closed and bounded set;

(3) Each input factor in the production process has free disposability. The free disposability of capital input is expressed as: if $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$ and $K_i \leq K_i'$, then $(K_i', L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$. The free disposability of labor input is expressed as: if $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$ and $L_i \leq L_i'$, then $(K_i, L_i', E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$. Similar expressions can be derived for disposal of other input factors;
(4) The desirable outputs have free disposability, which means that if $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$ and $Y_i' \leq Y_i$, then $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_{i_i}', C_i) \in T_i$;
(5) The undesirable outputs exhibit weak disposability, which means that if $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$ and $0 \leq \theta \leq 1$, then $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, \theta Y_i, \theta C_i) \in T_i$. This implies that undesirable outputs cannot be reduced separately during the production process and can only be reduced proportionally with desirable outputs;
(6) The "zero jointness" of the desirable and undesirable outputs indicates that when $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$, if there exists $C_i = 0$, then $Y_i = 0$. This means that obtaining desirable outputs during the production process will necessarily generate undesirable outputs, and only by completely terminating the entire production process can undesirable outputs be eliminated. The "zero jointness" is also known as by-productivity, which refers to the feature in the production process where undesirable outputs are considered as a by-product of desirable outputs, and C and Y will be produced simultaneously.

The new green environmental production technology can be modeled and described through an output set. Definition 2 is

TABLE 2 Descriptive statistics of input-output data of 29 provinces in China.

Variable	Unit	Max	Min	Mean	Median	Std
K	1 hundred million yuan	27682.68	200.46	8104.78	5568.93	7449.03
L	10 thousand people	2020.16	37.47	457.12	228.56	470.08
F	10^{16} Joule	702.00	25.79	208.78	162.86	160.02
E	10^{16} Joule	81.40	1.84	21.97	17.74	17.41
G	Piece	517942	2756	81170	53020	111107
H	Piece	65794	603	14231	9226	17018
Y	1 hundred million yuan	37014.80	583.73	9626.34	7407.29	9289.71
C	10 thousand tons	80819.58	3070.10	29498.90	21440.85	22638.97
P	Yuan/kg	11.81	0.49	3.26	3.16	2.70

provided as follows: the output set of new green environmental production technology.

Definition 2: In economic production activities, the output set of new green environmental production technology for a given industry is defined as:

$$P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i) = \{(Y_i, C_i): \ (K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i\}$$ (2)

Where $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$ represents the input factor combination in industry i, (Y_i, C_i) represents the output factor combination in industry i, and T_i represents the set of new green environmental production technologies in industry i as defined in Definition 1.

Based on the properties of the set of new green environmental production technologies, we present Property 2 of the output set of new green environmental production technologies:

Property 2: The output set of new green environmental production technologies satisfies the following properties:

(1) $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i) \in T_i$ if and only if $(Y_i, C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$;
(2) Production technologies exhibit output neutrality;
(3) The output set of new green environmental production technologies is a bounded closed set;
(4) Input factors have free disposability;
(5) Desirable outputs have free disposability;
(6) Undesirable outputs have weak disposability;
(7) Desirable outputs and undesirable outputs exhibit "zero association".

Note 2: Explanation of the 7 properties in Property 2:

(1) When an input-output factor combination $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i)$ belongs to the set of new green environmental production technologies T_i, that is, given the input vector $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$, the output set of the new green environmental production technologies $P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$ as defined in Definition 2 is

composed of all the feasible output vectors (Y_i, C_i). The converse also holds true.

(2) The output neutrality of production technologies is demonstrated by the fact that for any input vector $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$, then $\{0, 0, 0, 0\} \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$.

(3) In Eq. 2, for any given input vector $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$, the subscript i denotes industry, which is of a finite number. A finite number of inputs can only generate a finite number of outputs, ensuring the closure of the output vector. Clearly, the output set of new green environmental production technologies in the production theory is a closed bounded set.

(4) Each input factor in the production process has free disposability. The free disposability of capital input is manifested by the fact that if $K_i \leq K_i'$, then $P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i) \subseteq P_i(K_i', L_i, E_i \cdot F_i, H_i \cdot G_i)$. The free disposability of labor input is manifested by the fact that if $L_i \leq L_i'$, then $P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i) \subseteq P_i(K_i', L_i', E_i \cdot F_i, H_i \cdot G_i)$. Similarly, the free disposability of other input factors can be derived.

(5) The free disposability of undesirable outputs is manifested by the fact that if $(Y_i, C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$ and $Y_i' \leq Y_i$, then $(Y_i', C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$.

(6) The weak disposability of undesirable outputs is manifested by the fact that if $(Y_i, C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$ and $0 \leq \theta \leq 1$, then $(\theta Y_i, \theta C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$.

(7) The "zero association" of desirable and undesirable outputs is manifested by the fact that when $(Y_i, C_i) \in P_i(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i)$ and $C_i = 0$, we must have $Y_i = 0$.

Based on Definition 2 and Property 2, we have drawn Figure 2 to represent the output set of new green environmental production technologies.

Note 3: Explanation on Figure 2:

(1) In Figure 2, the horizontal axis represents non-desirable output, and the vertical axis represents desirable output. The slope from the origin O to any point in the production set represents the ratio of desirable output Y to

TABLE 3 Descriptive statistics of input-output data of 15 industries in China.

Variable	Unit	Max	Min	Mean	Median	Std
K	1 hundred million yuan	40128.20	2366.28	15669.23	14720.29	10075.83
L	10 thousand people	2238.89	103.76	883.77	858.39	599.72
F	10^{16} Joule	3105.82	13.35	403.64	83.76	802.86
E	10^{16} Joule	301.83	1.82	42.47	10.71	79.30
G	Piece	510926	8175	156929	109049	136528
H	Piece	264948	85	27514	1821	72349
Y	1 hundred million yuan	53561.78	900.46	18610.93	14135.69	14069.42
C	10 thousand tons	531655.36	46.58	57031.20	2401.12	147985.20
P	Yuan/kg	357.20	0.33	3.26	41.45	103.25

non-desirable output C, which indicates the carbon productivity of the corresponding output combination.

(2) For example, assuming that A, B, and C are three different decision-making units that use the same amount of input factors to produce different output factor combinations (C, Y), which are (10,30), (30,50), and (50,40), respectively. They are represented by the three vertices A, B, and C in the Figure 2, and the region OABCD represents the production set. Based on the specific values of output factor combinations from A, B, and C, we can obtain the carbon productivities of the three decision-making units $P_A = Y_A/C_A = 30/10 = 3$, $P_B = Y_B/C_B = 50/30 = 5/3$, and $P_C = Y_C/C_C = 40/50 = 4/5$, respectively. Clearly, $P_A > P_B > P_C$, and the slopes of the three lines OA, OB, and OC decrease in turn in Figure 2.

(3) According to the sixth property of Property 2, any output factor combination (Y, C) in the production set, (θY, θC) after the reduction according to the same ratio θ still belongs to the output set.

(4) According to the seventh property of Property 2, the only intersection of the horizontal and vertical axes with the production set is at the origin O, which indicates that non-desirable output C and desirable output Y are "zero correlated" and can only equal 0 at the same time. Non-desirable output C is the by-product of desirable output Y.

New green environmental production technologies effectively combine six inputs with two outputs. We will now use DEA method to construct a new green environmental linear programming model based on Definition 1 under the assumption of constant returns to scale.

Model 1: A new green environmental linear programming model for industry i in economic production activities under the assumption of constant returns to scale.

$$
\begin{aligned}
T_i = \Big\{ &\big(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\big): \\
&\sum_j \lambda_j K_{ij} \le K_i, \\
&\sum_j \lambda_j L_{ij} \le L_i, \\
&\sum_j \lambda_j \big(E_{ij} \cdot F_{ij}\big) \le (E_i \cdot F_i), \\
&\sum_j \lambda_j \big(H_{ij} \cdot G_{ij}\big) \le (H_i \cdot G_i), \\
&\sum_j \lambda_j Y_{ij} \ge Y_i, \\
&\sum_j \lambda_j C_{ij} = C_i, \\
&\lambda_j \ge 0, j = 1, 2, \cdots, N + 1 \Big\},
\end{aligned}
\tag{3}
$$

Where, in Eq. 3, $(K_i, L_i, E_i \cdot F_i, H_i \cdot G_i, Y_i, C_i)$ is the input-output vector for industry i, $(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ is the input-output vector for industry i in region j, and λ_j represents the intensity variable.

Note 4: Explanation of the constraints in Model 1:

(1) Inequalities are used to constrain each input in the production process. The right-hand side of the constraint represents the actual input used by the producer, while the left-hand side represents the input used by a theoretically efficient producer

TABLE 4 Input-output related data of reference region.

Symbol	K	L	F	E	G	H	Y	C	P
S1	833.96	38.49	4.49	0.54	3760	66	1083.10	154.98	69.88
S2	450.22	29.60	3.38	0.37	2581	45	487.44	55.34	88.08
S3	300.60	19.84	2.89	0.39	2268	63	333.21	82.80	40.24
S4	96.34	6.06	15.77	1.52	908	247	427.43	676.77	6.32
S5	229.57	16.56	0.93	0.11	2963	8	258.81	15.22	170.03
S6	836.17	39.93	23.93	2.80	7562	205	1064.08	445.52	23.88
S7	1383.73	65.46	107.10	10.41	11696	4127	1846.96	9259.23	1.99
S8	507.60	24.43	2.75	0.30	4651	98	429.15	111.88	38.36
S9	475.76	20.36	1.91	0.17	5062	31	320.46	77.32	41.45
S10	624.60	41.90	2.19	0.25	9247	53	840.82	56.14	149.78
S11	521.31	32.50	1.56	0.20	7957	26	657.12	30.95	212.31
S12	768.19	77.20	2.96	0.42	17618	36	1164.52	32.60	357.20
S13	786.48	33.45	35.62	4.16	3494	9136	608.16	18332.94	0.33
S14	81.60	3.58	2.84	0.27	282	86	74.05	165.60	4.47
S15	208.63	7.75	0.46	0.06	1122	2	31.05	1.61	193.30
Aggregate	8104.78	457.12	208.78	21.97	81170	14231	9626.34	29498.90	3.26

at the optimal production frontier. The inequality indicates that the input used by a theoretically efficient producer must be less than or equal to the input used by the actual producer, which reflects property 1 (3), that input factors are freely disposable. In this paper, constraints on the input of knowledge stock and green knowledge stock are added to the new green environmental linear programming model to further determine the feasible range. The product of renewable energy input and non-renewable energy input is treated as an input factor, and the product of knowledge stock input and green knowledge stock input is treated as an input factor for ease of calculation in subsequent steps.

(2) Inequalities are used to constrain the desirable outputs Y, where the right-hand side of the constraint represents the actual desirable outputs received by the producer, and the left-hand side represents the desirable outputs that a theoretically efficient producer would receive at the optimal production frontier. The inequality indicates that the desirable outputs received by a theoretically efficient producer must be greater than or equal to the actual desirable outputs received by the producer, which reflects Property 1 (4), that desirable outputs are freely disposable.

(3) Equalities are used to constrain the undesirable outputs, which reflects Property 1 (5), that undesirable outputs have weak disposability. The desirable outputs and undesirable outputs can only be reduced proportionally.

(4) λ_j represents the intensity variable, which indicates the weight assigned to each observation when constructing the production possibility boundary. λ_j in Model 1 is unconstrained and hence the new green environmental linear programming model is constructed under the assumption of constant returns to scale.

3.2 New green distance function and technical efficiency measurement

Drawing on the research of [37], a new definition is presented for the Shephard distance function for undesirable outputs orientation, incorporating input factors under the framework of green development and in accordance with Definition 1, based on the traditional output-oriented Shephard distance function.

Definition 3: The new green Shephard distance function for undesirable outputs orientation is defined as:

$$D_{io}^{C}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$$
$$= \sup\left\{\beta_{io}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \frac{C_{ij}}{\beta_{io}}\right) \in T_i\right\}, \quad (4)$$

Where, in Eq. 4, the subscript O indicates the study area.

Property 3: The new green Shephard distance function for undesirable outputs orientation, as defined in Definition 3, has the following properties:

(1) $D_{io}^{C}\left(K_{ij}, L_{ij}, \quad E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \geq 1 \Leftrightarrow (K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) \in T_i$;

(2) $D_{io}^{C}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, \quad H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) = 1 \Leftrightarrow (K_{ij}, L_{ij}, \quad E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ at the production technical frontier;

(3) $D_{io}^{C}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$ possesses first-degree homogeneity with respect to undesirable outputs: when α is a positive scalar, $D_{io}^{C}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \alpha \cdot C_{ij}\right) = \alpha D_{io}^{C}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$.
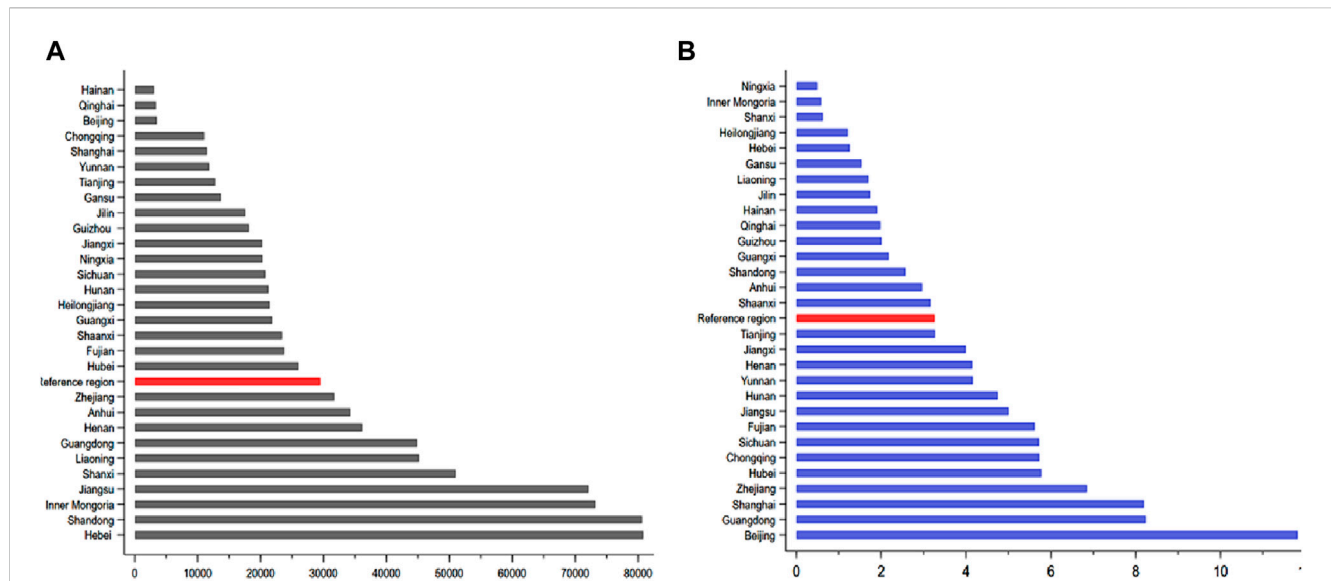
FIGURE 3
Carbon emissions and carbon productivity of 29 provinces and reference regions in 2019. **(A)** CO2 emissions (10^4 tons), **(B)** Carbon productivity (CNY/Kg).

Proof:

(1) When a combination of input-output factors $(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ belongs to the set of new green production technology T_i for undesirable outputs orientation, the new green Shephard distance function is used to measure the technical efficiency of each decision-making unit in terms of carbon emissions, as represented by $D_{io}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) = C_{io}/C_{io}^*$ which is the ratio of actual observed undesirable outputs to potential ideal undesirable outputs. Clearly, the actual observed undesirable outputs are greater than or equal to the potential ideal undesirable outputs. Therefore, $D_{io}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) \geq 1$, the opposite is also true;

(2) When a combination of input-output factors $(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ lies on the production technology frontier, the actual observed undesirable outputs are exactly equal to the potential ideal undesirable outputs, that is, $C_{io} = C_{io}^*$ Therefore, $D_{io}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) = C_{io}/C_{io}^* = 1$, and the opposite is also true;

(3) According to the formula of the new green Shephard distance function for undesirable outputs orientation:

$$D_{ij}^C\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \alpha \cdot C_{ij}\right)$$

$$= sup\left\{\beta_{ij}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \frac{\alpha \cdot C_{ij}}{\beta_{ij}}\right) \in T_i\right\}$$

$$= sup\left\{\alpha \cdot \frac{\beta_{ij}}{\alpha}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \frac{\alpha \cdot C_{ij}}{\beta_{ij}}\right) \in T_i\right\}$$

$$= sup\left\{\alpha \cdot \frac{\beta_{ij}}{\alpha}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \frac{C_{ij}}{\beta_{ij}/\alpha}\right) \in T_i\right\}$$

Based on Definition 3 and Model 1, the definition of the new green Farrell technical efficiency measure for undesirable outputs orientation is presented in this paper.

Definition 4: The new green Farrell undesirable outputs measure of production technical efficiency in Eq. 5 is defined as:

$$\left(D_{io}^C\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)\right)^{-1} = min\, \beta_{io},$$

$$s.t. \begin{cases} \sum_j \lambda_j K_{ij} \leq K_{io}, \\ \sum_j \lambda_j L_{ij} \leq L_{io}, \\ \sum_j \lambda_j \left(E_{ij} \cdot F_{ij}\right) \leq \left(E_{io} \cdot F_{io}\right), \\ \sum_j \lambda_j \left(H_{ij} \cdot G_{ij}\right) \leq \left(H_{io} \cdot G_{io}\right), \\ \sum_j \lambda_j Y_{ij} \geq Y_{io}, \\ \sum_j \lambda_j C_{ij} = \beta_{io} C_{io}, \\ \lambda_j \geq 0, j = 1, 2, \cdots, N + 1. \end{cases}$$

$$(5)$$

Note 5: The new green Farrell undesirable outputs measure used to measure production technical efficiency represents the proportion that needs to shrink to the minimum possible value of undesirable outputs, while keeping the desirable outputs constant, given a constant input factor. According to the meaning of the new green Shephard distance function for undesirable outputs orientation, the new green Farrell undesirable outputs measure of production technical efficiency is the reciprocal of the distance function defined in Definition 3, that is, $D_{io}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) \geq 1$. Therefore, β_{io} is no greater than 1. When $\beta_{io} = 1$, that is, the new green Shephard distance function for undesirable outputs orientation is 1, the new green Farrell undesirable outputs measure of production technical efficiency is 1, and the system is in the optimal production state.
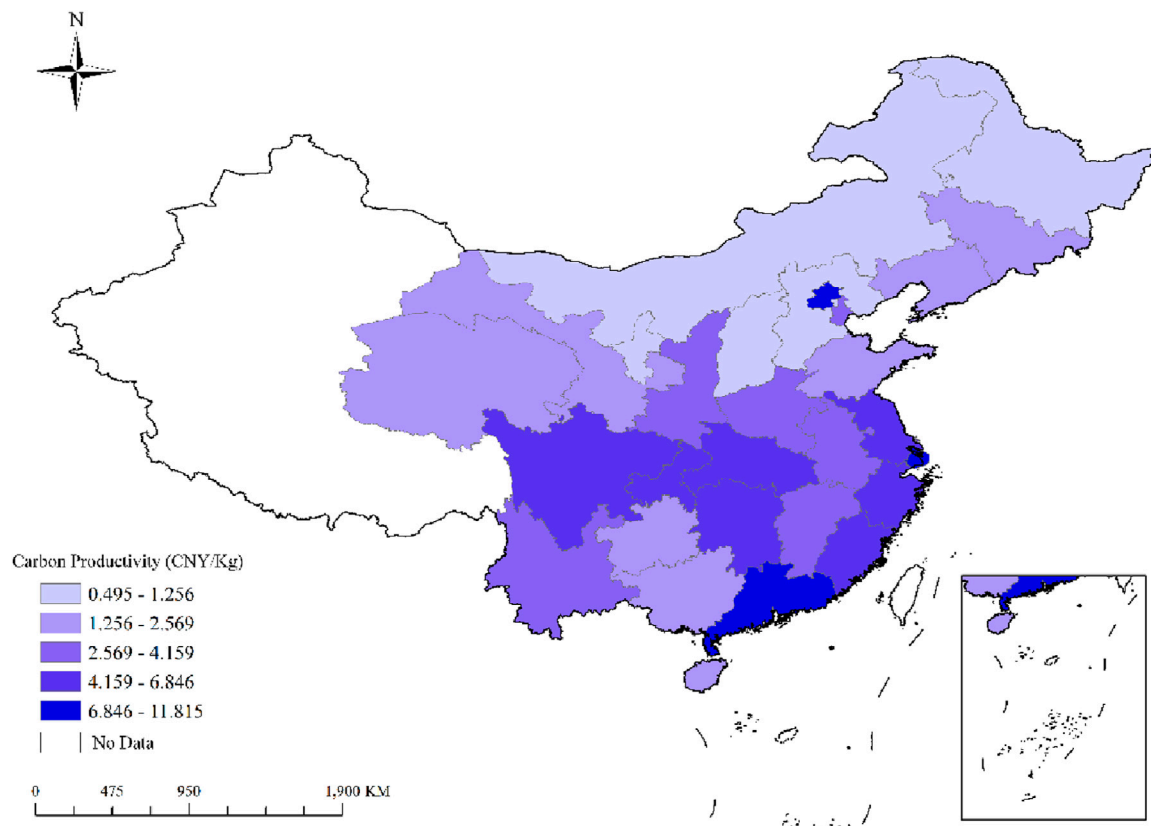
FIGURE 4
Spatial distribution of China's carbon productivity in 2019.

Definition 5: The new green Shephard distance function for desirable outputs orientation is in Eq. 6:

$$
\begin{aligned}
&D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \\
&= \inf\left\{\theta_{io}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, \frac{Y_{ij}}{\theta_{io}}, C_{ij}\right) \in T_i\right\},
\end{aligned}
\tag{6}
$$

Property 4: The new green Shephard distance function for desirable outputs orientation defined in Definition 5 has the following properties:

(1) $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, \ H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \le 1 \Leftrightarrow \ \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \in T_i$;

(2) $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, \ \ H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) = 1 \Leftrightarrow \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$ at the production technical frontier;

(3) $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$ is positively homogeneous of degree +1 with respect to desirable outputs: when α is a positive scalar, $D_{io}^{Y}(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, \alpha \cdot C_{ij}) = \alpha D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)$.

Proof:

(1) When a set of input-output factor combinations $(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ belongs to the new green environmental production technology set T_i, the new green Shephard distance function for desirable outputs orientation is used

to measure the technical efficiency of each decision unit in producing desirable outputs, represented as $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) = Y_{io}/Y^{\star}$, the ratio of observed desirable outputs to potential ideal desirable outputs. Obviously, observed desirable outputs are less than or equal to potential ideal desirable outputs, therefore $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \le 1$ and *vice versa*.

(2) When a set of input-output factor combinations $(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$ is at the production technical frontier, the observed desirable outputs are exactly equal to the potential ideal desirable outputs, $Y_{io} = Y_{io}^{\star}$. Therefore, $D_{io}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) = Y_{io}/Y_{io}^{\star} = 1$ and *vice versa*.

(3) According to the formula for the new green Shephard distance function for desirable outputs orientation:

$$
\begin{aligned}
&D_{ij}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \\
&= \inf\left\{\theta_{ij}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, \frac{\alpha \cdot Y_{ij}}{\theta_{ij}}, C_{ij}\right) \in T_i\right\} \\
&= \inf\left\{\alpha \cdot \frac{\theta_{ij}}{\alpha}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, \frac{Y_{ij}}{\theta_{ij}/\alpha}, C_{ij}\right) \in T_i\right\} \\
&= \alpha \cdot \inf\left\{\frac{\theta_{ij}}{\alpha}: \left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, \frac{Y_{ij}}{\theta_{ij}/\alpha}, C_{ij}\right) \in T_i\right\} \\
&= \alpha \cdot D_{ij}^{Y}\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right).
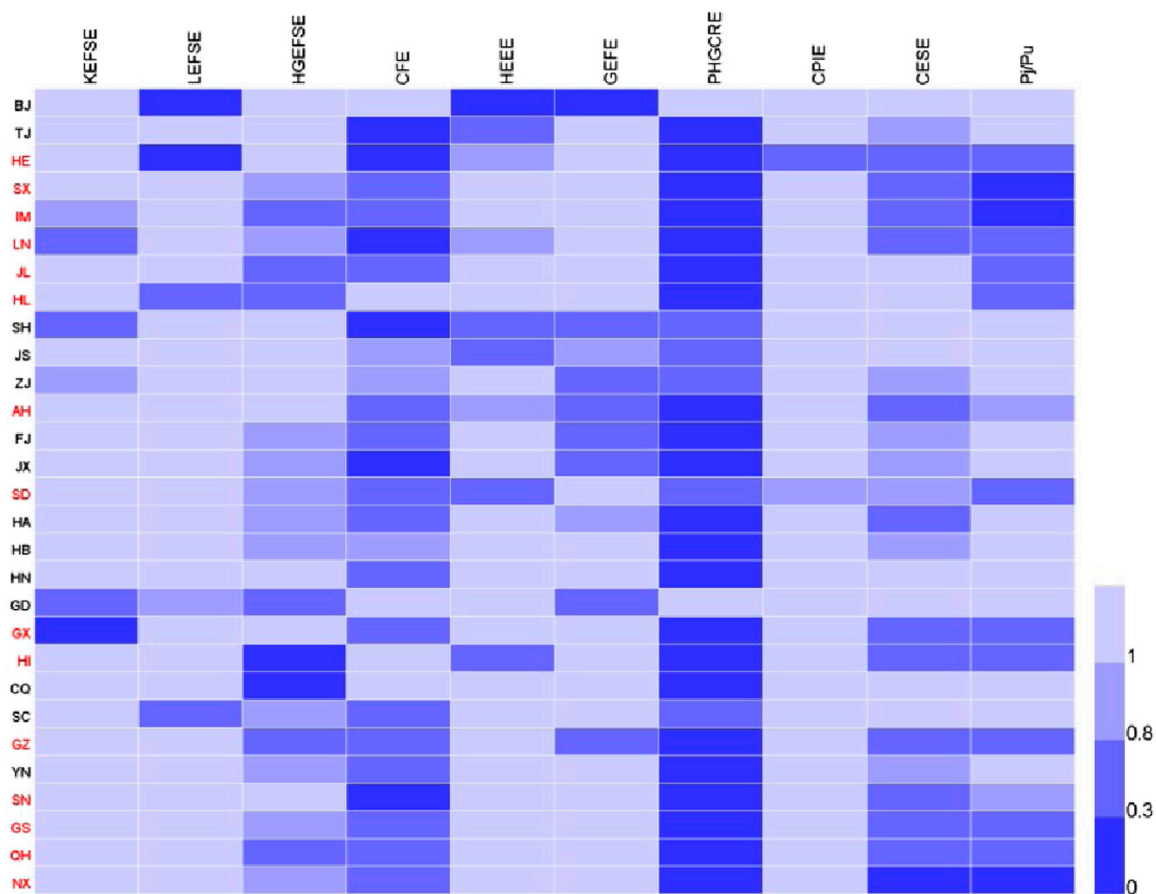\end{aligned}
$$

FIGURE 5
Decomposition results of provincial carbon productivity differences.

According to Definition 5 and Model 1, we present Definition 6: new green Farrell technical efficiency measure for desirable outputs orientation.

Definition 6: The new green Farrell measure of production technical efficiency for desirable outputs orientation is in Eq. 7 defined as follows:

$$\left(D_{io}^Y\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right)\right)^{-1} = \max \theta_{io},$$

$$s.t.\begin{cases} \sum_j \lambda_j K_{ij} \le K_{io}, \\ \sum_j \lambda_j L_{ij} \le L_{io}, \\ \sum_j \lambda_j\left(E_{ij} \cdot F_{ij}\right) \le \left(E_{io} \cdot F_{io}\right), \\ \sum_j \lambda_j\left(H_{ij} \cdot G_{ij}\right) \le \left(H_{io} \cdot G_{io}\right), \\ \sum_j \lambda_j Y_{ij} \ge \theta_{io} Y_{io}, \\ \sum_j \lambda_j C_{ij} = C_{io}, \\ \lambda_j \ge 0, j = 1, 2, \cdots, N + 1. \end{cases} \quad (7)$$

Note 6 The new green Farrell measure for desirable outputs orientation used to measure production technical efficiency represents the expansion ratio required to achieve the maximum possible value of desirable outputs, under the condition that undesirable outputs remain constant and given input factors remain unchanged. Based on the definition of the new green Shephard distance function for desirable outputs orientation, it can be seen that the new green Farrell measure of production technical efficiency for desirable outputs orientation is the reciprocal of the distance function defined in Definition 5, $D_{io}^Y\left(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}\right) \le 1$. Therefore, θ_{io} is not less than 1, and when $\theta_{io} = 1$, the new green Shephard distance function for desirable outputs orientation is 1, the new green Farrell measure of production technical efficiency for desirable outputs orientation is also 1, which represents the optimal production state.

3.3 The new green carbon productivity decomposition theorem

According to [13] definition of carbon productivity and considering the input factors in the context of green development, this paper derives a new green decomposition theorem for carbon productivity in the context of green development.

Theorem 1: (New Green Decomposition Theorem for Carbon Productivity): The carbon productivity in region j can be decomposed into new green components as follows:

$$P_j = \sum_{i=1}^{M} EFE_{ij} \cdot HEE_{ij} \cdot GEF_{ij} \cdot HGCR_{ij} \cdot CES_{ij}, \qquad (8)$$

Where, in Eq. 8 $EFE_{ij} = \frac{Y_{ij}}{E_{ij} \cdot F_{ij}}$ represents the hybrid energy utilization efficiency of industry i in region j, which is defined as GDP added value by unit of hybrid energy input; $HEE_{ij} = E_{ij}/H_{ij}$ represents the green knowledge efficiency of renewable energy input, which is defined as the consumption of renewable energy per unit of green knowledge stock input; $GEF_{ij} = F_{ij}/G_{ij}$ represents the knowledge efficiency of non-renewable energy input, which is defined as the consumption of non-renewable energy per unit of knowledge stock input; $HGCR_{ij} = H_{ij} \cdot G_{ij}/C_{ij}$ represents the hybrid knowledge emission ratio, which is defined as the consumption of hybrid knowledge stock per unit of carbon emission spatial resources; $CES_{ij} = C_{ij}/C_j$ represents the carbon emission structure, which is defined as the proportion of carbon emissions from industry i in region j to the total carbon emissions from all industries in region j.

Proof: Carbon productivity in region j is equal to the ratio of the added value of GDP in region j to CO_2 emissions, and the added value of GDP in region j is written in the form of the sum of the added value of GDP in each industry in region j. Proof of Theorem 1 in Eq. 9:

$$\begin{aligned} P_j &= \frac{Y_j}{C_j} = \sum_{i=1}^{M} \frac{Y_{ij}}{C_j} \\ &= \sum_{i=1}^{M} \frac{Y_{ij}}{E_{ij} \cdot F_{ij}} \cdot \frac{E_{ij} \cdot F_{ij}}{H_{ij} \cdot G_{ij}} \cdot \frac{H_{ij} \cdot G_{ij}}{C_{ij}} \cdot \frac{C_{ij}}{C_j} \\ &= \sum_{i=1}^{M} \frac{Y_{ij}}{E_{ij} \cdot F_{ij}} \cdot \frac{E_{ij}}{H_{ij}} \cdot \frac{F_{ij}}{G_{ij}} \cdot \frac{H_{ij} \cdot G_{ij}}{C_{ij}} \cdot \frac{C_{ij}}{C_j} \\ &= \sum_{i=1}^{M} EFE_{ij} \cdot HEE_{ij} \cdot GEF_{ij} \cdot HGCR_{ij} \cdot CES_{ij}. \end{aligned} \qquad (9)$$

Note 7: Theorem 1 decomposes carbon productivity of any region into the above-mentioned five indicators by using a new green decomposition approach. This paper further introduces three decomposed indicators, namely, green knowledge efficiency of renewable energy input, knowledge efficiency of non-renewable energy input, and hybrid knowledge emission ratio, into the green-decomposed results of carbon productivity. The mechanism for the driving factors of carbon productivity is analyzed from two perspectives, that is, input-output ratio and input-input ratio.

Based on Theorem 1, Definitions 3 and Definitions 5, a new green decomposition theorem of carbon productivity with a Shepherd distance function directed to both desirable outputs and undesirable outputs is derived.

Theorem 2: (New green-decomposition theorem of carbon productivity with embedded distance function) Carbon productivity of region j can be decomposed into a new green-decomposition with an embedded distance function in Eq. 10 as follows:

$$P_j = \sum_{i=1}^{M} MIX_{ij} \cdot HEE_{ij} \cdot GEF_{ij} \cdot PHGCR_{ij} \cdot CPI_{ij} \cdot CES_{ij}, \qquad (10)$$

Where, $MIX_{ij} = \frac{Y_{ij}/(E_{ij} \cdot F_{ij})}{D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})}$ represents the new green hybrid indicator of industry i in region j; $PHGCR_{ij} = (H_{ij} \cdot G_{ij})D_{ij}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})/C_{ij}$ represents the potential hybrid knowledge emission ratio of industry i in region j; $CPI_{ij} = \frac{D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})}{D_{ij}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})}$ represents the carbon efficiency index of industry i in region j; while the meanings of HEF_{ij}, GEF_{ij} and CES_{ij} remain the same as in Theorem 1.

Proof: Theorem 2 is proved by embedding the new green Shepherd distance function directed to both desirable outputs and undesirable outputs on the basis of the proof of Theorem 1 in Eq. 11.

$$\begin{aligned} P_j &= \sum_{i=1}^{M} \frac{Y_{ij}}{E_{ij} \cdot F_{ij}} \cdot \frac{E_{ij}}{H_{ij}} \cdot \frac{F_{ij}}{G_{ij}} \cdot \frac{H_{ij} \cdot G_{ij}}{C_{ij}} \cdot \frac{C_{ij}}{C_j} \\ &= \sum_{i=1}^{M} \frac{Y_{ij} / D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})}{E_{ij} \cdot F_{ij}} \cdot \frac{E_{ij}}{H_{ij}} \cdot \frac{F_{ij}}{G_{ij}} \\ &\qquad \cdot \frac{H_{ij} \cdot G_{ij}}{C_{ij} / D_{ij}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})} \\ &\qquad \cdot \frac{D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})}{D_{ij}^C(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})} \cdot \frac{C_{ij}}{C_j} \\ &= \sum_{i=1}^{M} MIX_{ij} \cdot HEE_{ij} \cdot GEF_{ij} \cdot PHGCR_{ij} \cdot CPI_{ij} \cdot CES_{ij}. \end{aligned}$$
$$(11)$$

Note 8: Theorem 2 further decomposes carbon productivity of any region into the above-mentioned six indicators by embedding a new green Shepherd distance function directed to both desirable outputs and undesirable outputs on the basis of Theorem 1. This enables a more in-depth study of the driving factors of carbon productivity from the perspectives of both actual observed state and potential ideal state. The following are explanations for the three newly added indicators among the six decomposed indicators: new green hybrid indicator, potential hybrid knowledge emission ratio, and carbon efficiency index.

(1) In the new green hybrid indicator, the numerator is changed from Y_{ij} in the original hybrid energy utilization efficiency indicator to $Y_{ij}/D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$, which means that desirable outputs are expanded to potential ideal state under the corresponding undesirable outputs-directed new green Shepherd distance function. The denominator remains the hybrid energy input, so the whole fraction represents the potential hybrid energy utilization efficiency, indicating the potential level of hybrid energy utilization efficiency when the technical performance of desirable outputs reaches the ideal state.

According to [43] study, the new green Shephard distance function with outputs oriented desirable outputs have homogeneous of degree +1, that is, $D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, \alpha \cdot Y_{ij}, C_{ij}) = \alpha \cdot D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$, and with inputs oriented desirable outputs have homogeneous of degree -1, that is, $D_{ij}^Y(\beta \cdot K_{ij}, \beta \cdot L_{ij}, \beta \cdot E_{ij} \cdot F_{ij}, \beta \cdot H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij}) = \beta^{-1} \cdot D_{ij}^Y(K_{ij}, L_{ij}, E_{ij} \cdot F_{ij}, H_{ij} \cdot G_{ij}, Y_{ij}, C_{ij})$, where
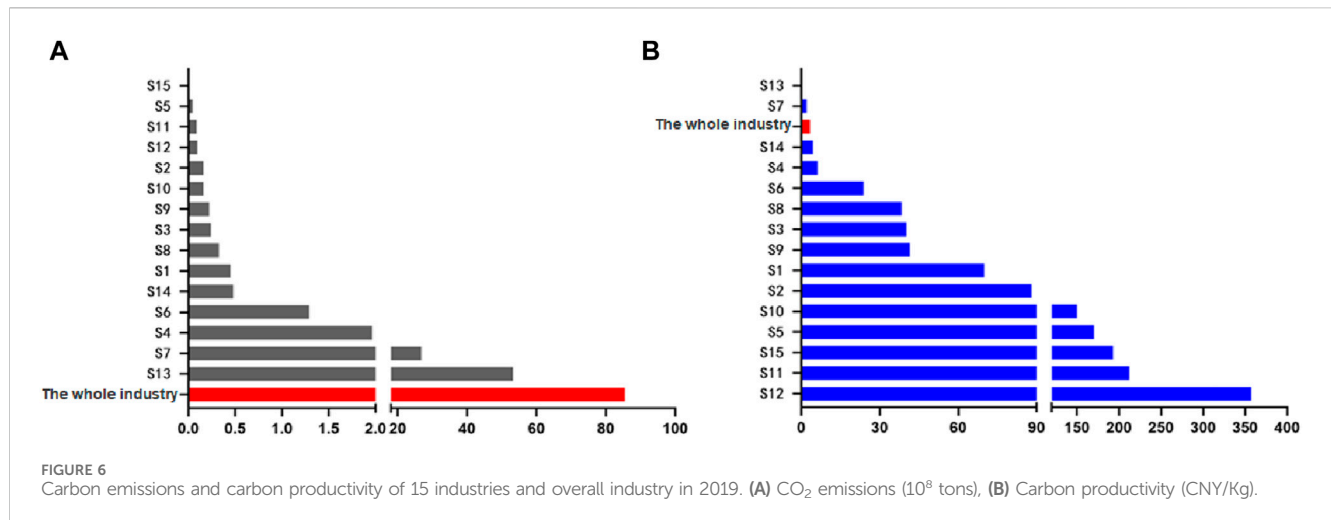
FIGURE 6
Carbon emissions and carbon productivity of 15 industries and overall industry in 2019. **(A)** CO_2 emissions (10^8 tons), **(B)** Carbon productivity (CNY/Kg).

α and β are positive scalars. Hence, the new green hybrid index can be formulated as:

$$\frac{Y_{ij}\big/D_{ij}^Y\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)}{E_{ij}\cdot F_{ij}}$$

$$= \left[\frac{1}{Y_{ij}}\cdot\left(\frac{1}{E_{ij}\cdot F_{ij}}\right)^{-1}\cdot D_{ij}^Y\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)\right]^{-1}$$

$$= \left[\frac{1}{Y_{ij}}\cdot D_{ij}^Y\left(k_{ij},l_{ij},1,h_{ij}\cdot g_{ij},Y_{ij},c_{ij}\right)\right]^{-1}$$

$$= \left(D_{ij}^Y\left(k_{ij},l_{ij},1,h_{ij}\cdot g_{ij},1,c_{ij}\right)\right)^{-1},$$

Where, $k_{ij}=\frac{K_{ij}}{E_{ij}\cdot F_{ij}}$ represents the capital - hybrid energy ratio (*KEF*); $l_{ij}=\frac{L_{ij}}{E_{ij}\cdot F_{ij}}$ represents the labor - hybrid energy ratio (*LEF*); $h_{ij}\cdot g_{ij}=\frac{H_{ij}\cdot G_{ij}}{E_{ij}\cdot F_{ij}}$ represents the green knowledge stock input ratio (*HGEF*), and $c_{ij}=\frac{C_{ij}}{E_{ij}\cdot F_{ij}}$ represents the carbon factor (*CF*), which denotes the carbon conversion rate of the hybrid energy input. The calculation process highlights the superiority of representing the hybrid energy input as a product of renewable energy input E and non-renewable energy input F and that of representing the hybrid knowledge stock input as a product of green knowledge stock input H and total knowledge stock input G.

(2) In the potential hybrid knowledge emission ratio, the denominator has been changed from C_{ij}, which was the hybrid knowledge emission ratio for the original index, to $\frac{C_{ij}}{D_{ij}^C\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)}$, which represents the undesirable outputs based on the corresponding undesirable outputs-oriented new green Shephard distance function. This indicates a reduction of the potential output to an ideal state of undesirable outputs. The numerator remains as the hybrid knowledge stock input. The entire fraction represents the potential level of hybrid knowledge emission ratio when the undesirable outputs technical efficiency reaches its ideal state.

(3) The carbon performance index is expressed as the ratio of the new green Shephard distance function oriented towards desirable outputs and the one oriented towards undesirable outputs. According to the meanings of the two distance

functions, $D_{ij}^Y\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)=Y_{ij}/Y_{ij}^*$ and $D_{ij}^C\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)=C_{ij}/C_{ij}^*$. By substituting these values into the expression for the carbon performance index, we obtain $CPI_{ij}=\frac{D_{ij}^Y\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)}{D_{ij}^C\left(K_{ij},L_{ij},E_{ij}\cdot F_{ij},H_{ij}\cdot G_{ij},Y_{ij},C_{ij}\right)}=\frac{Y_{ij}/Y_{ij}^*}{C_{ij}/C_{ij}^*}=\frac{Y_{ij}/C_{ij}}{Y_{ij}^*/C_{ij}^*}=\frac{P_{ij}}{P_{ij}^*}$, which represents the ratio of the actual observed carbon productivity to the potential ideal state carbon productivity. The potential ideal state of carbon productivity is greater than or equal to the actual observed carbon productivity. Therefore, the range of the carbon performance index *CPI* is between 0 and 1, where *CPI* approaching 1 indicates higher carbon productivity, which means that the actual observed carbon productivity is closer to the potential ideal state carbon productivity. When *CPI* equals 1, it represents the potential ideal state. This is the optimal level of carbon emission performance.

Based on Theorem 2 and Note 8(1), the decomposition of carbon productivity in the j region can be obtained as:

$$P_j = \sum_{i=1}^M \left(D_{ij}^Y\left(k_{ij},l_{ij},1,h_{ij}\cdot g_{ij},1,c_{ij}\right)\right)^{-1}\cdot HEE_{ij}\cdot GEF_{ij}$$
$$\cdot PHGCR_{ij}\cdot CPI_{ij}\cdot CES_{ij} \qquad (12)$$

The decomposition results of carbon productivity in the j region have included three indicators: the green knowledge efficiency of renewable energy input, the knowledge efficiency of non-renewable energy input, and the potential hybrid knowledge emission ratio. This paper explores the influential factors of carbon productivity from the perspectives of input quantity and potential ideal states. To compare the differences of carbon productivity among different regions, two methods can be used: additive decomposition and multiplicative decomposition. In this paper, we adopt the multiplicative decomposition method to derive a new green decomposition theorem for regional differences in carbon productivity.

Theorem 3: (A new green decomposition theorem for regional differences of carbon productivity) The carbon productivity ratio
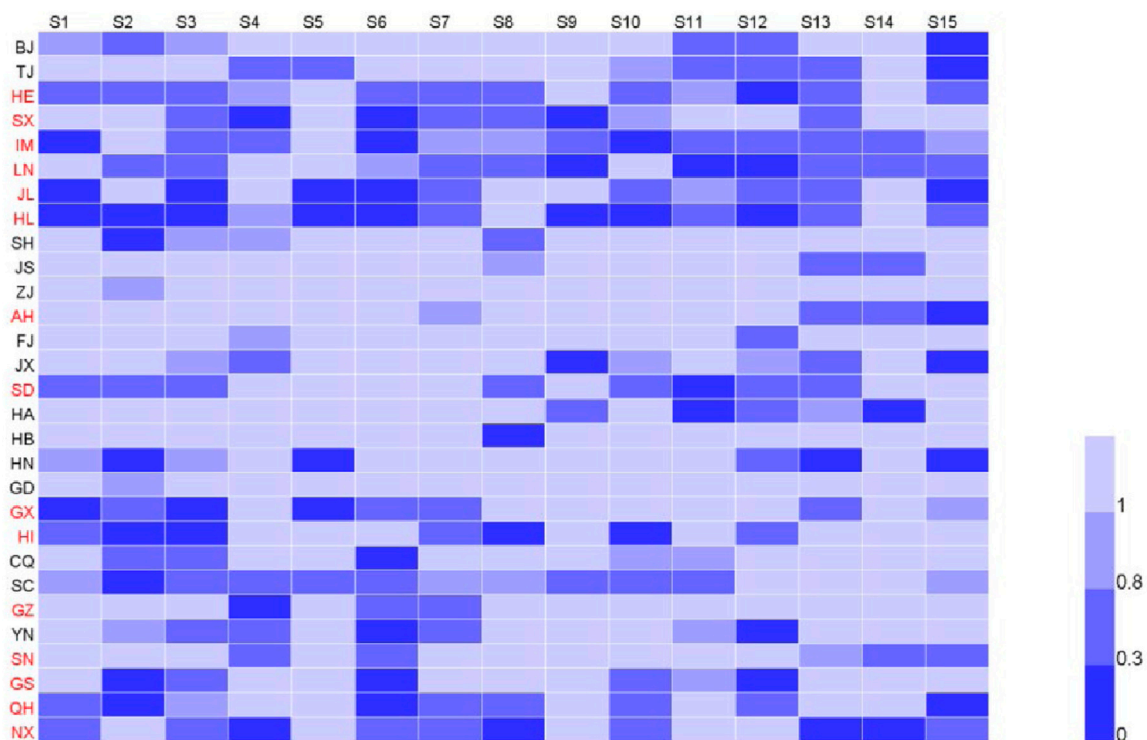
FIGURE 7
Industrial decomposition results of provincial carbon productivity differences.

between any given region j and a reference region can be decomposed into a new green decomposition as follows:

$$\frac{P_j}{P_u} = A_{MIX}^{j,u} \cdot A_{HEE}^{j,u} \cdot A_{GEF}^{j,u} \cdot A_{PHGCR}^{j,u} \cdot A_{CPI}^{j,u} \cdot A_{CES}^{j,u}, \quad (13)$$

Where, $A_{MIX}^{j,u}$ represents the new green hybrid effect (MIX), which includes four types of differences: the capital-hybrid energy ratio, labor - hybrid energy ratio, hybrid knowledge stock - hybrid energy ratio, and carbon factor, between the two regions. $A_{HEE}^{j,u}$ represents the green knowledge efficiency effect of renewable energy input ($HEEE$); $A_{GEF}^{j,u}$ represents the knowledge efficiency effect of non-renewable energy input ($GEFE$); $A_{PHGCRE}^{j,u}$ represents the potential hybrid - knowledge emission ratio effect ($PHGCRE$); $A_{CPI}^{j,u}$ represents the carbon performance index effect ($CPIE$); and $A_{CES}^{j,u}$ represents the carbon emission structure effect ($CESE$).

Proof: According to Eq. 12, dividing the carbon productivity decomposition equation in Eq. 14 of two regions yields:

$$\frac{P_j}{P_u} = \frac{\sum_{i=1}^{M} \left(D_{ij}^{Y}\left(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij}\right) \right)^{-1} \cdot HEE_{ij}}{\sum_{i=1}^{M} \left(D_{iu}^{Y}\left(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu}\right) \right)^{-1} \cdot HEE_{iu} \cdot GEF_{iu}} \quad (14)$$
$$= A_{MIX}^{j,u} \cdot A_{HEE}^{j,u} \cdot A_{GEF}^{j,u} \cdot A_{PHGCR}^{j,u} \cdot A_{CPI}^{j,u} \cdot A_{CES}^{j,u}.$$

Note 9: Theorem 3 uses the new green decomposition to analyze the regional differences in carbon productivity, decomposing the ratio of carbon productivity between two regions into six effects, which correspond to the six decomposition indicators in Eq. 12 of

carbon productivity. Furthermore, we have added the green knowledge efficiency effect of renewable energy inputs, the knowledge efficiency effect of non-renewable energy inputs, and the potential hybrid - knowledge emission ratio effect to investigate the influencing factors of regional differences in carbon productivity from a new perspective. In this paper, carbon productivity is represented as the ratio of industry added value and CO_2 emissions, thus, it is an intensity variable. We use a multiplication decomposition method to study the differences in carbon productivity between regions. Based on the research of [31], we select the LMDI-1 decomposition weight formula to obtain the six effects in Eqs 15–20 in Theorem 3.

$$A_{MIX}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{D_{iu}^{Y}\left(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu}\right)}{D_{ij}^{Y}\left(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij}\right)} \right), \quad (15)$$

$$A_{HEE}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{HEE_{ij}}{HEE_{iu}} \right), \quad (16)$$

$$A_{GEF}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{GEF_{ij}}{GEF_{iu}} \right), \quad (17)$$

$$A_{PHGCR}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{PHGCR_{ij}}{PHGCR_{iu}} \right), \quad (18)$$

$$A_{CPI}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{CPI_{ij}}{CPI_{iu}} \right), \quad (19)$$

$$A_{CES}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{CES_{ij}}{CES_{iu}} \right), \quad (20)$$

Where, $w_i^{j,u} = \frac{L(Y_{ij}/C_{ij}, Y_{iu}/C_{iu})}{L(P_j, P_u)}$ represents the weight function, and

$$L(a, b) = \begin{cases} \dfrac{a - b}{\ln a - \ln b}, a \neq b \\ a, a = b \end{cases}$$

represents the logarithmic averaging function.

Regarding the new green hybrid effect, $A_{MIX}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}\right)$, there are some parts that can be further decomposed. By separating the four indicators of the capital-hybrid energy ratio, labor-hybrid energy ratio, hybrid knowledge stock-hybrid energy ratio, and carbon factor between the two regions, we derived a decomposition theorem for the new green hybrid effect.

Theorem 4: (New green hybrid effect decomposition theorem): The new green hybrid effect can be decomposed into the following components:

$$A_{MIX}^{j,u} = A_{KEF}^{j,u} \cdot A_{LEF}^{j,u} \cdot A_{HGEF}^{j,u} \cdot A_{CF}^{j,u}, \tag{21}$$

Where, $A_{KEF}^{j,u}$ represents the capital-hybrid energy substitution effect (*KEFSE*), $A_{LEF}^{j,u}$ represents the labor-hybrid energy substitution effect (*LEFSE*), $A_{HGEF}^{j,u}$ represents the hybrid knowledge stock-hybrid energy substitution effect (*HGEFSE*), and $A_{CF}^{j,u}$ represents the carbon factor effect (*CFE*).

Proof: Referring to the "one-factor-at-a-time" principle of the Laspeyres-linked approach in [43] study, we keep all other factors constant and change only one factor at a time to analyze its driving effect during the research period. For the logarithmic part, $\frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{iu}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}$, of the new green hybrid effect, we use this method by sequentially changing k, l, $h \cdot g$, and c in the distance function ratio in different orders, where $h \cdot g$ is treated as a whole for changes. According to permutation and combination, there are $A^4 = 4! = 24$ decomposition forms in total.

We sequentially alter the values of k, l, $h \cdot g$, and c in sequence in the equation, and the results of the decomposition are as follows:

$$\frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}$$

$$= \frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})} \cdot \frac{D_{i}^{Y}(k_{ij}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}$$

$$\cdot \frac{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{iu})} \cdot \frac{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}. \tag{22}$$

The four terms on the right-hand side of Eq. 22 respectively reflect the influences of *KEF*, *LEF*, *HGEF*, and *CF*, and similarly, we can obtain 23 other decomposition results. [44] proposed using a geometric mean with equal weights to obtain a complete decomposition of the distance function ratio for all possible decompositions, but the large number of factors can make the calculation process cumbersome. [45] introduced the polar coordinate method, which approximates the complete decomposition by averaging a group of mirrored decomposition possibilities and captures the influences of each factor from the

opposite direction. Drawing on the research of [37], we utilize the polar coordinate method in this paper. The decomposition results of the mirrored sequence c, $h \cdot g$, l, and k in Eq. 22 are presented in Eq. 23 as follow:

$$\frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}$$

$$= \frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{ij})} \cdot \frac{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{ij})}{D_{i}^{Y}(k_{iu}, l_{i}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}$$

$$\cdot \frac{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}{D_{i}^{Y}(k_{iu}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})} \cdot \frac{D_{i}^{Y}(k_{iu}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}. \tag{23}$$

The four terms on the right-hand side of Eq. 22 respectively reflect the influences of *CF*, *HGEF*, *LEF*, and *KEF*. By averaging the group of mirrored decomposition possibilities, we obtain the final decomposition of the distance function ratio as follows:

$$\frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}$$

$$= \left[\frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})} \cdot \frac{D_{i}^{Y}(k_{iu}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}\right]^{\frac{1}{2}}$$

$$\cdot \left[\frac{D_{i}^{Y}(k_{ij}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})} \cdot \frac{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}{D_{i}^{Y}(k_{iu}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}\right]^{\frac{1}{2}}$$

$$\cdot \left[\frac{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{iu})} \cdot \frac{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{ij})}{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}\right]^{\frac{1}{2}}$$

$$\cdot \left[\frac{D_{i}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})} \cdot \frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{i}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{ij})}\right]^{\frac{1}{2}}$$

$$= KEF_i^{j,u} \cdot LEF_i^{j,u} \cdot HGEF_i^{j,u} \cdot CF_i^{j,u}. \tag{24}$$

Substituting Eq. 24 into Eq. 15 yields Eq. 25:

$$A_{MIX}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln \frac{D_{iu}^{Y}(k_{iu}, l_{iu}, 1, h_{iu} \cdot g_{iu}, 1, c_{iu})}{D_{ij}^{Y}(k_{ij}, l_{ij}, 1, h_{ij} \cdot g_{ij}, 1, c_{ij})}\right)$$

$$= \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln\left(KEF_i^{j,u} \cdot LEF_i^{j,u} \cdot HGEF_i^{j,u} \cdot CF_i^{j,u}\right)\right)$$

$$= \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln KEF_i^{j,u}\right) \cdot \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln LEF_i^{j,u}\right)$$

$$\cdot \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln HGEF_i^{j,u}\right) \cdot \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln CF_i^{j,u}\right)$$

$$= A_{KEF}^{j,u} \cdot A_{LEF}^{j,u} \cdot A_{HGEF}^{j,u} \cdot A_{CF}^{j,u}. \tag{25}$$

Note 10: Theorem 4 decomposes the new green hybrid effect into four effects. In this paper, we adopt a hybrid energy input of the product of renewable energy input and non-renewable energy input, which adds the mutual substitution effect between hybrid knowledge stock input and hybrid energy input. We consider using knowledge stock input to replace energy input in the production process, which optimizes the input factor ratio from a new perspective.
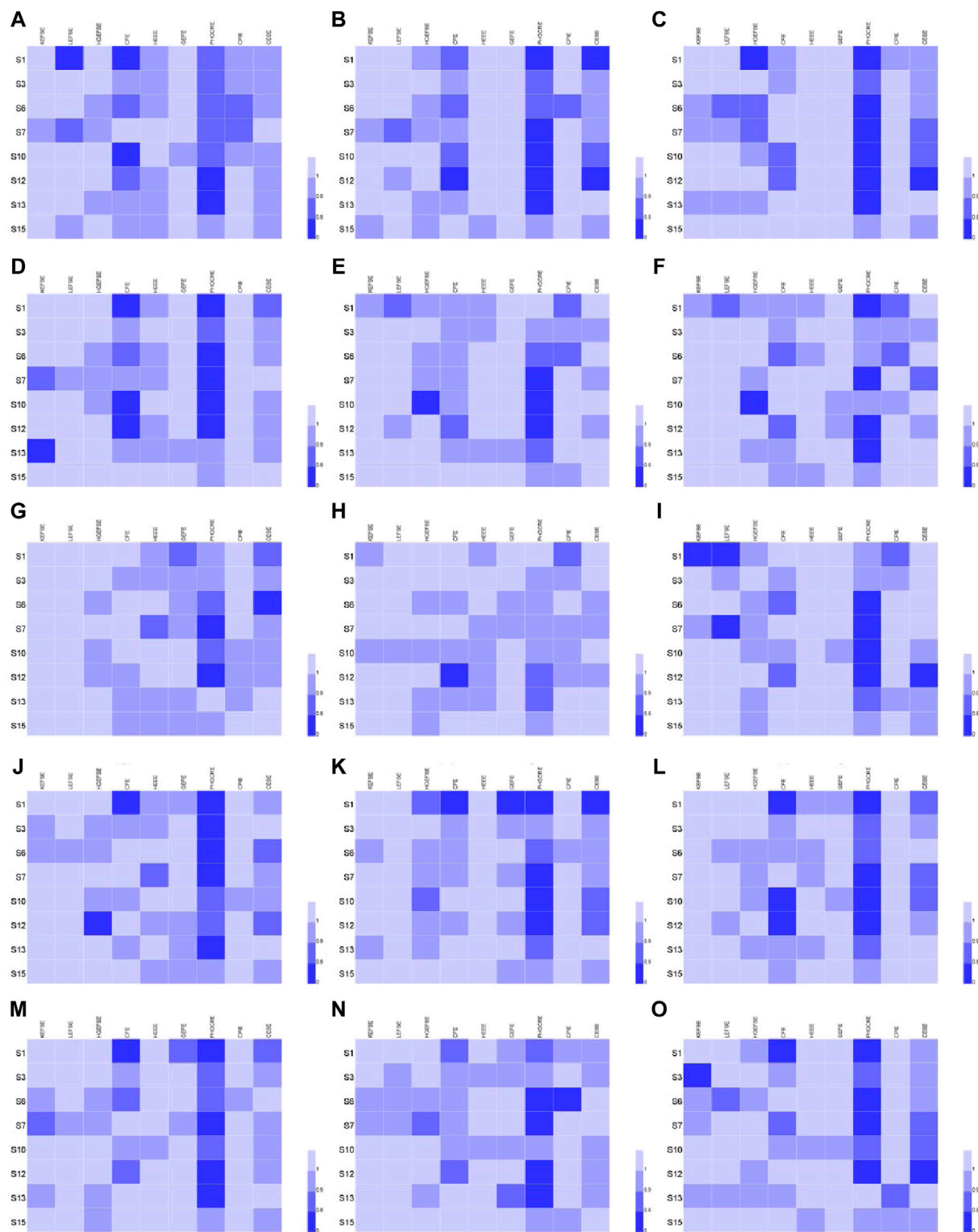
FIGURE 8
Decomposition results of carbon productivity differences in 8 key industries in 15 key provinces. **(A)** Hebei, **(B)** Shanxi, **(C)** Inner Mongolia, **(D)** Liaoning, **(E)** Jilin, **(F)** Heilongjiang, **(G)** Anhui, **(H)** Shandong, **(I)** Guangxi, **(J)** Hainan, **(K)** Guizhou, **(L)** Shaanxi, **(M)** Gansu, **(N)** Qinghai, **(O)** Ningxia.

According to Theorem 3 and Theorem 4, substituting Eq. 21 into Eq. 13 yields the decomposition formula for carbon productivity regional differences:

$$\frac{P_j}{P_u} = A_{KEF}^{j,u} \cdot A_{LEF}^{j,u} \cdot A_{HGEF}^{j,u} \cdot A_{CF}^{j,u} \cdot A_{HEE}^{j,u} \cdot A_{GEF}^{j,u} \cdot A_{PHGCR}^{j,u} \cdot A_{CPI}^{j,u} \cdot A_{CES}^{j,u}$$

$$(26)$$

We decompose the carbon productivity ratio between region *j* and reference region *u* into the above 9 effects. If the decomposition value of any effect is greater than 1, it indicates that the effect has enlarged the differences in carbon productivity between region *j* and reference region *u*; otherwise, it has reduced the differences.

4 Empirical analysis

4.1 Data sources and description

Sample data were selected for the provinces and cities of China in 2019, excluding Hong Kong, Macao, Taiwan, Xinjiang, and Tibet due to missing data. The study objects were divided according to China's seven major geographical regions (see Table A1), and the reference region was obtained by taking the average of the 29 provinces and autonomous regions, representing the average level of socio-economic activities. Based on China's national economic industry classification, 38 sub-industries in manufacturing and electricity, heat, gas, and water production and supply were classified and summarized into 15 industries (see Table 1).

The new green decomposition model of carbon productivity involves a total of 9 variables, namely, capital input *K*, labor input *L*, renewable energy input *E*, non-renewable energy input *F*, knowledge stock input *G*, green knowledge stock input *H*, industrial added value *Y*, carbon dioxide emissions *C*, and carbon productivity *P*. This paper collected and compiled 6 input and 2 output data for 15 industries in 29 provinces in 2019. The capital input was obtained from fixed asset investment data from the "China Investment Statistical Yearbook" in billions of yuan. The labor input was obtained from employment population data from the "China Population and Employment Statistical Yearbook" in ten thousand people. The non-renewable energy input used non-renewable energy heat generation data from the CEADs China carbon accounting database. Referring to relevant research by [46], the usage of different types of energy in provincial energy inventories was converted into unified standard coal based on the conversion factors in the "China Energy Statistical Yearbook" for each type of energy, then converted into heat value, with units in joules of 10'16. Renewable energy input used renewable energy heat generation data from the National Energy Administration and "China Energy Statistical Yearbook", converted to heat value based on the conversion of each province's renewable energy power consumption to unified standard coal, with units in joules of 10'16. The knowledge stock input used ordinary patent data from the State Intellectual Property Office and provincial statistical yearbooks, with units in pieces. The green knowledge stock input used green patent data from provincial green patent panel data, with units in pieces. The desirable outputs are used u industrial added value data from the National Bureau of Statistics, "China Industrial Statistical Yearbook", and provincial statistical yearbooks in billions of yuan. The undesirable outputs are used carbon dioxide emissions data [46] from the CEADs China carbon accounting database's provincial emission inventory, with units in ten thousand tons. Carbon productivity was calculated based on the *P*=*Y/C* formula with units in yuan/kg. A few missing data were obtained using interpolation.

Table 2 summarizes the input and output data for the 29 provinces studied in this paper. There are significant differences between provinces. Table 3 summarizes the input and output data for the 15 industries studied in this paper. There are also significant differences between different industries. Looking at the mean column of Table 2 and Table 3, the mean column data in Table 2 is the input-output data of the reference region, and the mean column in Table 3 is the average data of the 15 industries. The mean carbon productivity of all provinces is the same as the mean carbon productivity of all industries (3.26 yuan/kg), because the mean carbon productivity is not the mean of carbon productivity, but the ratio of industrial added value means to carbon dioxide emissions mean, which matches the provincial and industrial data.

Table 4 summarizes the input and output data for various industries in the reference region. The carbon productivity of the overall industry is not the sum of carbon productivity of each industry, but rather the ratio of the overall industry GDP added value to the overall industry carbon dioxide emissions. The overall industry carbon productivity for the reference region, located in the bottom right corner of the table, is 3.26 yuan/kg, which further confirms the consistency between provincial and industrial data.

4.2 One-dimensional comparative analysis

4.2.1 Comparative analysis at provincial level

From a provincial perspective, Figure 3 depicts the carbon dioxide emissions and carbon productivity for the 29 provinces and the reference region studied in this paper. As shown in Figure 3A, the top 3 provinces with the highest carbon dioxide emissions are Hebei (808.20 million tons), Shandong (806.06 million tons), and Inner Mongolia (732.32 million tons), while the top 3 provinces with the lowest carbon dioxide emissions are Hainan (30.70 million tons), Qinghai (33.52 million tons), and Beijing (35.15 million tons). The carbon dioxide emissions for the reference region (294.99 million tons) are marked in red and ranked 11th among the 29 provinces. As shown in Figure 3B, the top 3 provinces with the highest carbon productivity are Beijing (11.81 yuan/kg), Guangdong (8.24 yuan/kg), and Shanghai (8.19 yuan/kg), while the top 3 provinces with the lowest carbon productivity are Ningxia (0.49 yuan/kg), Inner Mongolia (0.58 yuan/kg), and Shanxi (0.63 yuan/kg). The carbon productivity for the reference region (3.26 yuan/kg) is marked in red and ranked 15th among the 29 provinces. By combining Figure 3, it can be found that the provincial ranking of carbon dioxide emissions and carbon productivity varies significantly due to the inter-provincial differences in industrial added value.

Figure 4 shows the spatial distribution of provincial-level carbon productivity in China, with different shades of color indicating the numerical value of carbon productivity for each province. The white color in the map represents provinces that have not been counted due to data deficiency. From the map, it can be clearly seen that the level of carbon productivity in China is generally higher in the southeast coastal provinces, followed by central provinces, and lower in the southwest and northwest regions. Looking at China's seven major geographical regions, for the North China region, Beijing has the highest carbon productivity in the country, while Tianjin's carbon productivity is close to the reference region and therefore

at the average level. The carbon productivity of the other provinces in the region is relatively low, especially Shanxi and Inner Mongolia. The overall level of carbon productivity in the three provinces in the Northeast region is poor, all lower than the reference region. In the East China region, except for Anhui and Shandong which are below the average level, the carbon productivity of the other provinces is relatively high, especially Shanghai, Jiangsu, and Zhejiang, which are among the top provinces in the country. The carbon productivity of all three provinces in the Central China region is higher than the average level. In the South China region, Guangdong has a high level of carbon productivity, while Guangxi and Hainan have lower levels than the average. In the southwest region, except for Guizhou, the carbon productivity of the other three provinces is higher than the average level. In the northwest region, the overall level of carbon productivity in the four provinces is relatively low, all lower than the average level.

Using the new green decomposition model of carbon productivity, which takes into account carbon emissions, we can apply Theorem 3 and Theorem 4 to compare the carbon productivity of 29 provinces with that of the reference regions. Eq. 26 can be used to decompose the regional differences in carbon productivity:

$$\frac{P_j}{P_u} = A_{KEF}^{j,u} \cdot A_{LEF}^{j,u} \cdot A_{HGEF}^{j,u} \cdot A_{CF}^{j,u} \cdot A_{HEE}^{j,u} \cdot A_{GEF}^{j,u} \cdot A_{PHGCR}^{j,u} \cdot A_{CPI}^{j,u} \cdot A_{CES}^{j,u},$$

By applying the calculation formula for each specific effect, taking the capital-hybrid energy substitution effect KEF as an example, we obtain:

$$A_{KEF}^{ij,iu} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln KEF_i^{j,u}\right),$$

The remaining 8 types of effects can be obtained similarly. By taking j from 1 to 29, we can obtain the effect decomposition results of carbon productivity differences between each province and the reference region, and present them as a heat map. Figure 5 shows the ratio of industry-wide carbon productivity of 29 provinces in this paper to that of the reference region, as well as the effect decomposition results.

The y-axis of Figure 5 represents the 29 provinces, and the first nine columns of the x-axis represent the results of the nine decomposition effects. The values of each decomposition effect are represented using different shades of color, with lighter colors indicating larger values of the specific decomposition effect. The lightest color indicates that the decomposition effect is greater than 1, which means it has a positive impact on carbon productivity. On the other hand, darker colors indicate a smaller value of the specific decomposition effect, which has a negative impact on carbon productivity. The last column shows the ratio of the industry-wide carbon productivity of each province to that of the reference region. Provinces whose industry-wide carbon productivity is lower than the reference region are marked in red on the y-axis, indicating that they have significant potential for improvement.

From the first nine columns, we can see that the $KEFSE$, $LEFSE$, and $CPIE$ have a positive influence on carbon productivity for the majority of provinces, indicating that these provinces have a reasonable input factor structure and carbon emission

performance levels. In contrast, the $HGEFSE$ has a negative impact on the carbon productivity of most provinces, except for the 10 provinces of Beijing, Tianjin, Hebei, Shanghai, Jiangsu, Zhejiang, Anhui, Hunan, Guangxi, and Shaanxi. Therefore, when it comes to improving the carbon productivity of these 19 provinces, the $HGEFSE$ has greater potential compared with the $KEFSE$, $LEFSE$, and $CPIE$. It is a driving factor that requires special attention to improve carbon productivity by increasing the substitution of hybrid knowledge stock for hybrid energy in input factors. The CFE and $PHGCRE$ have negative impacts on carbon productivity of most provinces, indicating the need to optimize energy consumption structure and improve the potential ideal hybrid knowledge emission ratio level in future development. The $HEEE$ and $GEFE$ also have different promotion and inhibition effects on carbon productivity for different provinces. These two effects are important factors that promote further improvement of carbon productivity for provinces with a higher level of carbon productivity. By increasing scientific research investment, promoting technological innovation, encouraging patent research and development, and actively developing green low-carbon technologies, carbon productivity can be enhanced to a higher level. The $CESE$ is another important factor that leads to significant differences in carbon productivity among provinces. Except for the provinces of Beijing, Jilin, Heilongjiang, Shanghai, Jiangsu, Hunan, Guangdong, Chongqing, and Sichuan, there is no significant improvement in the carbon productivity of the remaining 20 provinces, indicating that there is great potential to improve the level of carbon productivity by improving industrial sector structure. The $KEFSE$ and $LEFSE$ are the main driving forces for the development of carbon productivity in most provinces, while the substitution of hybrid knowledge stock for hybrid energy and the potential ideal hybrid knowledge emission ratio still needs to be improved. In particular, the $HEEE$ and $GEFE$ are key factors for further improving carbon productivity in economically developed provinces.

4.2.2 Industry-level comparative analysis

From an industry-level perspective, Figure 6 shows the carbon emissions and carbon productivity of the 15 industries studied in this paper, as well as the overall industry.

From Figure 6A, we can see that the top three industries with the highest carbon emissions are the power and heat production and supply industry (S13) with 5.317 billion tons, followed by the black gold, non-ferrous metal, and metal products industry (S7) with 2.685 billion tons, and the petroleum processing industry (S4) with 0.196 billion tons. All of these industries are high-energy consumption industries. The three industries with the lowest carbon emissions are the water production and supply industry (S15) with 0.0047 billion tons, followed by the medical industry (S5) with 0.0441 billion tons, and the electrical equipment manufacturing industry (S11) with 0.0898 billion tons. The industry-wide carbon emissions of all industries combined (85.55 billion tons) are marked in red. From Figure 6B, we can see that the top three industries with the highest carbon productivity are the communication equipment, office equipment, and other manufacturing industry (S12) with 357.20 yuan/kg, followed by the electrical equipment manufacturing industry (S11) with 212.31 yuan/kg, and the water production and supply industry (S15) with 193.30 yuan/kg. The

three industries with the lowest carbon productivity are the power and heat production and supply industry (S13) with 0.33 yuan/kg, followed by the black gold, non-ferrous metal, and metal products industry (S7) with 1.99 yuan/kg, and the gas production and supply industry (S14) with 4.47 yuan/kg. The industry-wide carbon productivity (3.26 yuan/kg) is marked in red, which is obtained by comparing the industry-wide GDP added value with the industry-wide carbon emissions. This once again confirms the consistency between provincial data and industry data. By combining Figure 6, we can conclude that high-energy consumption industries tend to have higher carbon emissions but lower carbon productivity. Conversely, industries with lower carbon emissions tend to have higher carbon productivity. For example, the power and heat production and supply industry (S13) have the highest carbon emissions and lowest carbon productivity. While the power industry has been committed to reducing energy consumption, its high reliance on fossil fuels results in high carbon emissions, leaving significant challenges to improve carbon productivity.

We divide the carbon productivity of each industry in the 29 provinces by the carbon productivity of the corresponding industry in the reference region (P_{ij}/P_{iu}). By taking j from 1 to 29 and i from 1 to 15, we obtain the industry decomposition results of the differences in carbon productivity between each province and the reference region and present them in a heat map. Figure 7 depicts the ratio of carbon productivity between the 29 provinces and the reference region, as well as the industry decomposition results. The vertical axis represents the 29 provinces, while the horizontal axis represents the 15 industries. Observing the vertical aspect of Figure 7, we can see that the industries with lower carbon productivity than the reference region are mainly concentrated in 15 provinces, namely, Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Anhui, Shandong, Guangxi, Hainan, Guizhou, Shaanxi, Gansu, Qinghai, and Ningxia. We mark these 15 provinces in red on the vertical axis. Combining this with the provinces marked in red on the vertical axis in Figure 4, we find that the provinces with overall lower carbon productivity and specific industries with lower carbon productivity than the reference region are completely consistent. Therefore, we identify these 15 provinces as potential areas for improving carbon productivity. Compared to provinces with higher carbon productivity, these provinces are of greater concern and worthy of further study.

4.3 Dual dimensional comparative analysis

Based on the comparative analysis of single-dimension province and single-dimension industry, and the comprehensive effect decomposition results and industry decomposition results of each province, we found that there are 15 provinces with carbon productivity levels lower than the reference region. These provinces are Hebei, Shanxi, Inner Mongolia, Liaoning, Jilin, Heilongjiang, Anhui, Shandong, Guangxi, Hainan, Guizhou, Shaanxi, Gansu, Qinghai, and Ningxia. The 8 weak industries that cause the carbon productivity of these 15 provinces to be lower than the average level are: food, tobacco and liquor industry (S1), culture and education supplies industry (S3),

chemical industry (S6), black and non-ferrous metal smelting and processing industry (S7), transportation equipment industry (S10), communication equipment, office equipment and other manufacturing industry (S12), electric power, heat production and supply industry (S13), and water production and supply industry (S15).

By combining the differences and influencing factors of carbon productivity at the regional and industry levels, and considering the industry heterogeneity within regions, we obtained the effect decomposition of specific industry-level carbon productivity differences in different regions. In this paper, we conducted an effect decomposition of the carbon productivity differences of these 15 key provinces and 8 key industries from the two dimensions of province and industry, and investigated the factors influencing the differences. Using the new green decomposition model of carbon productivity and applying the decomposition Eq. 26 for regional differences, we obtained the decomposition formula for the carbon productivity differences with provincial and industrial subscripts.

$$\frac{P_{ij}}{P_{iu}} = A_{KEF}^{ij,iu} \cdot A_{LEF}^{ij,iu} \cdot A_{HGEF}^{ij,iu} \cdot A_{CF}^{ij,iu} \cdot A_{HEE}^{ij,iu} \cdot A_{GEF}^{ij,iu} \cdot A_{PHGCR}^{ij,iu} \cdot A_{CPI}^{ij,iu} \cdot A_{CES}^{ij,iu},$$

This refers to the decomposition effect of the carbon productivity difference between the j region industry and the reference region industry, which is the product of the decomposition effects of the carbon productivity difference between each industry in the j region and the reference region. This paper uses the capital-hybrid energy substitution effect KEF as an example.

$$A_{KEF}^{j,u} = \exp\left(\sum_{i=1}^{M} w_i^{j,u} \ln KEF_i^{j,u}\right) = \prod_{i=1}^{M} \exp\left(w_i^{j,u} \ln KEF_i^{j,u}\right)$$
$$= \prod_{i=1}^{M} A_{KEF}^{ij,iu},$$

Furthermore, we obtain the capital-hybrid energy substitution effect KEF of the carbon productivity difference between the i industry in the j region and the i industry in the reference region:

$$A_{KEF}^{ij,iu} = \exp\left(w_i^{j,u} \ln KEF_i^{j,u}\right),$$

The remaining eight effects are obtained similarly. Taking j from 1 to 15 and i from 1 to 8, we obtain the decomposition results of the effects of carbon productivity differences between each province and industry and the reference region industry, which can be presented in heat maps. Figure 8 contains 15 heat maps, each describing the decomposition results of the effects on eight key industries of the relative carbon productivity laggard provinces. Overall, the *KEFSE* and *LEFSE* have higher decomposition results in most provinces and industries, indicating a positive role in promoting the development of carbon productivity. In contrast, the *PHGCRE* has lower decomposition results in most provinces and industries, indicating a negative impact on carbon productivity and significant potential for improvement. This is consistent with the provincial decomposition results presented in Figure 5. The *PHGCRE* indicates, that is, the consumption of hybrid knowledge resources per unit of carbon emission space, at the point when

undesirable outputs reach the ideal state, corresponding to measures aimed at increasing hybrid knowledge stock input and reducing carbon emissions.

Figure 8A shows the decomposition results of the effects on eight key industries in Hebei province. It can be observed that there are significant variations in the decomposition results of various effects across different industries. Overall, the *HGEFSE* and *GEFE* have a positive impact on the development of carbon productivity in most industries. On the other hand, the *CFE*, *HEEE*, *CPIE*, and *CESE* have a negative impact on carbon productivity in most industries. Thus, it is important to focus on increasing the decomposition values of these effects. As an industrial province, Hebei's economic development has long relied on heavy industry. From the figure, it can be observed that the chemical industry (S6) needs to improve the *HGEFSE*, *CFE*, *HEEE*, *CPIE*, and *CESE*.

Figure 8B, C show the decomposition results of the effects on eight key industries in Shanxi and Inner Mongolia, respectively. The results in these two regions are similar. From the figures, it can be observed that the *HEEE*, *GEFE*, and *CPIE* have a positive impact on carbon productivity in most industries. On the other hand, the *HGEFSE*, *CFE*, and *CESE* still have significant potential for improvement. As an economically underdeveloped region, the carbon productivity of many industries in Shanxi province is below the average level. For example, the *LEFSE*, *CFE*, and *CESE* have a negative impact on carbon productivity in the manufacturing industry of communication equipment, office equipment, and other industries (S12). Although Inner Mongolia is also an economically underdeveloped region, it has abundant natural resources that can be reasonably developed and utilized to reduce carbon emissions and improve carbon productivity.

Figure 8D–F show the decomposition results of the effects on eight key industries in the three provinces of Northeast China. Among them, the *HGEFSE*, *GEFE*, and *CPIE* in Liaoning have a positive impact on carbon productivity in most industries. On the other hand, the *CFE*, *HEEE*, and *CESE* have great potential for improvement. The decomposition results of Jilin and Heilongjiang are relatively similar. The *HEEE*, *GEFE*, and *CESE* have a positive impact on carbon productivity in most industries; however, the *HGEFSE* and *CFE* need to be improved. For example, the black and non-ferrous metal, and metal products industry (S7) in Liaoning needs to adopt policies such as capital replacement of hybrid energy, labor replacement of hybrid energy, and hybrid knowledge stock replacement of hybrid energy, increase scientific and technological inputs, and encourage the research and application of green patents. The transportation equipment industry (S10) in Jilin needs to vigorously promote the replacement of hybrid knowledge stock for hybrid energy, regulate the structure of production factors, improve technical efficiency, and the potential emission ratio of hybrid knowledge. The chemical industry (S6) in Heilongjiang needs to raise the level of technology and carbon emission efficiency, advocate the development and application of green patents, and vigorously develop high-tech emerging industries.

Figure 8G shows the decomposition results of the effects on eight key industries in Anhui province. The *HGEFSE* and *CPIE* have a positive impact on carbon productivity in most industries (S1, S3, S6, S7, S13, S15), while the *HEEE*, *GEFE*, and *CESE* have great potential for improvement. Anhui province is a relatively backward province in the Yangtze River Economic Zone, with great potential

for economic development. For instance, the black and non-ferrous metal, and metal products industry (S7) should focus on increasing scientific and technological inputs, promoting the development and application of patents, especially green patents, improving technical efficiency, balancing industrial structure, and applying green and clean energy and low-carbon technologies in the production process to effectively improve the level of carbon productivity.

Figure 8H shows the decomposition results of the effects on eight key industries in Shandong province. The *GEFE* and *CESE* have a positive impact on carbon productivity in most industries, while there is still room for improvement for *HEEE* and *CPIE*. Shandong province is an important energy base in China, with abundant coal and oil resources. However, this advantage also leads to an imbalanced energy consumption structure, with a high proportion of fossil energy consumption and resulting in large amounts of carbon emissions and low levels of carbon productivity. Therefore, Shandong province urgently needs to improve its energy consumption structure, reduce the use of traditional fossil fuels, and vigorously develop clean energy sources such as wind, solar, and hydropower to reduce carbon emissions.

Figure 8I shows the decomposition results of the effects on eight key industries in Guangxi province. The *HEEE*, *GEFE*, and *CPIE* have a positive impact on carbon productivity in most industries, while there is still room for improvement for *HGEFSE* and *CFE*. Guangxi province is relatively backward in the southern region of China, and for specific industries such as the food, tobacco, and liquor industry (S1), which is mainly focused on the sugar industry, it is necessary to focus on improving the capital substitution of hybrid energy, labor substitution of hybrid energy, and hybrid knowledge stock substitution of hybrid energy. Adjusting the proportions of various production factors inputs, further improving technical efficiency, and carbon performance levels, increasing the economic benefits of the industry, and thus further enhancing the level of carbon productivity.

Figure 8J shows the decomposition results of the effects on eight key industries in Hainan province. The *CPIE* has a positive impact on carbon productivity in most industries, while there is still room for improvement for *HEEE* and *CESE*. Hainan has a favorable geographical location, rich forestry and marine resources, and a strong ability to absorb and store carbon dioxide. Hainan has diverse energy structure and enormous potential for developing renewable energy sources such as solar and hydropower generation. Additionally, Hainan has abundant geothermal and tidal energy resources, which should be fully utilized to balance the energy consumption structure, reduce carbon emissions, and promote the absorption and storage of carbon emissions. By combining both carbon sources and carbon sinks, Hainan can lower carbon dioxide concentration in the atmosphere and enhance its carbon productivity.

Figure 8K shows the decomposition results of the effects on eight key industries in Guizhou province. The *HEEE* and *CPIE* have a positive impact on carbon productivity in most industries, while there is still room for improvement for *HGEFSE*, *CFE*, *GEFE* and *CESE*. Guizhou province is relatively underdeveloped in the southwestern region of China, and the carbon productivity levels of most industries need to be improved. For example, in the electric and thermal power generation and supply industry (S13), which is a

key area for transmitting electricity from western to eastern China, there is a need to increase capital substitution of hybrid energy and hybrid knowledge stock substitution of hybrid energy, and optimize the input structure of production factors while improving the potential emission rate of hybrid knowledge. By doing so, carbon productivity can be enhanced comprehensively. Another key industry in Guizhou is the food, tobacco, and liquor industry (S1), for which there is a need to increase hybrid knowledge stock substitution of hybrid energy, adjust the energy consumption structure, encourage the development of green patents, reduce carbon emissions, improve economic efficiency, and gradually enhance the level of carbon productivity.

Figure 8L–O show the decomposition results of the effects on eight key industries in four northwestern provinces of China. The results in Shaanxi, Gansu, Qinghai, and Ningxia are quite similar, the *HEEE*, *GEFE*, and *CPIE* have a positive impact on carbon productivity in most industries, while the *CFE* and *CESE* still have room for improvement. These northwestern provinces have sparse populations, arid climates, and lack of water resources, while possessing abundant desert and wind-blown sand resources. Therefore, it is necessary to increase scientific and technological investment and make use of existing natural resources to vigorously develop solar power stations and ecological photovoltaic power stations. At the same time, renewable energy sources such as wind power and biogas should be reasonably developed and utilized as new forms of clean energy in order to jointly reduce CO_2 emissions and enhance the level of carbon productivity.

5 Conclusion

This paper explores the input elements in the economic production process under the driving force of green development, where energy input is refined into renewable energy input and non-renewable energy input, and knowledge stock input and green knowledge stock input are introduced. By using the data envelopment analysis (DEA) method and combining traditional environmental production technology with green input elements, a new linear programming model with a constant return to scale is constructed. Based on the traditional output-oriented Shephard distance function, a new green Shephard distance function and corresponding green Farrell technical efficiency measures for desirable outputs orientation and undesirable outputs orientation are also constructed. In addition, a series of new green decomposition methods are derived based on the definition of carbon productivity, and a new green decomposition model of carbon productivity is established. The carbon productivity differences between two regions are decomposed into nine effect types, resulting in the formation of a new green decomposition system of carbon productivity. Using input-output data from 15 industries in 29 Chinese provinces in 2019, this paper investigates the regional differences in carbon productivity and the factors that influence these differences from both a single-dimensional and a two-dimensional perspective.

(1) Conclusion of single-dimensional provincial level comparison analysis: Out of the 29 provinces studied in this paper, 14 provinces have higher carbon productivity than the

reference region, mainly including Beijing, Tianjin, most of Eastern China, Central China, and parts of Southwest China, while the remaining 15 provinces have lower carbon productivity than the reference region, mainly including some parts of North China, Northeast China, South China and Northwest China. The regions with relatively backward carbon productivity mostly have rich natural resources, but poor resource utilization efficiency and over-reliance on heavy industry, especially industrial sectors that have been transferred from the eastern regions to the western regions. Therefore, it is necessary to optimize the industrial structure in these regions to gradually improve their carbon productivity.

(2) Conclusion of single-dimensional industry level comparison analysis: Among the 15 industries studied in this paper, high-energy-consuming industries such as the electric power generation and supply industry (S13), the processing of black and non-ferrous metals and other minerals industry (S7), and the petroleum processing industry (S4) have high levels of CO_2 emissions and relatively low carbon productivity. In contrast, industries with low levels of CO_2 emissions such as the water production and supply industry (S15), the pharmaceutical industry (S5), and the electrical machinery manufacturing industry (S11) have relatively higher levels of carbon productivity. Most high-energy-consuming industries use fossil energy sources, which leads to excessive CO_2 emissions and low carbon productivity levels. Therefore, new energy sources need to be continuously developed and utilized, and the proportion of traditional fossil energy use needs to be reduced at its core to minimize carbon emissions, enhance energy utilization efficiency, and improve carbon productivity.

(3) Conclusion of two-dimensional comparison analysis: Among the eight key industries and 15 provinces studied in this paper, the decomposition results for the capital - hybrid energy substitution effect (*KEFSE*) and labor - hybrid energy substitution effect (*LEFSE*) are high. This indicates that capital and labor substitution are the main driving forces behind the development of carbon productivity in these 15 key provinces. However, the decomposition results for the potential hybrid knowledge emission ratio effect (*PHGCRE*) are comparatively low, indicating that there is significant room for improvement in the potential level of hybrid knowledge emission rate under ideal conditions in these 15 key provinces.

6 Policy suggestions

China is the country with the highest carbon dioxide emissions in the world. To achieve low-carbon development, we must simultaneously meet the requirements of increasing industrial added value and reducing carbon emissions, so as to fundamentally and effectively improve carbon productivity. Firstly, based on the input and output data of 15 industries in 29 provinces and the new green decomposition model of carbon productivity, the influencing factors of regional carbon productivity differences are quantitatively studied. Secondly, through the decomposition of provincial effects and provincial industries,

15 key provinces and 8 key industries that affect the improvement of China's overall carbon productivity will be identified, and potential areas for the development of carbon productivity will be identified. Based on the decomposition results and the actual situation, the following policy recommendations are given:

(1) At the single-dimension provincial level, the carbon factor effect *CFE* and the potential hybrid knowledge emission ratio effect *PHGCRE* have negative effects on the carbon productivity of most provinces. Therefore, priority should be given to improving these two effects, optimizing the energy consumption structure in time, improving the level of hybrid knowledge emission ratio, and implementing differentiated energy policies according to the characteristics of natural resources and industrial structure in different regions. Hybrid knowledge stock - hybrid energy substitution effect *HGEFSE* has a negative impact on the carbon productivity of the rest of the provinces, except for the positive effect on the carbon productivity of nearly one-third of the provinces. Therefore, it is necessary to continuously improve the substitution of hybrid knowledge stock for hybrid energy and adjust the structure of input factors, so as to improve the overall carbon productivity. The carbon emission structure effect *CESE* has no significant improvement in the carbon productivity of the remaining 20 provinces, except for its positive effects on the nine provinces of Beijing, Jilin, Heilongjiang, Shanghai, Jiangsu, Hunan, Guangdong, Chongqing and Sichuan. Therefore, we should give priority to the development of low-carbon industries, get rid of the high dependence on energy-intensive industries and heavy industries, learn from provinces with advanced technology and experience in this field, and adjust the industrial structure, so as to gradually improve carbon productivity.

(2) At the single-dimension industry level, stationery industry (S3), communication equipment, office equipment and other manufacturing industries (S12) and power and heat production and supply industry (S13) have negative effects on the carbon productivity of most provinces, and the carbon productivity of these three industries needs to be greatly improved. From the perspective of the seven major regions, North China needs to focus on the development of the stationery industry (S3), petroleum processing industry (S4), transportation equipment industry (S10), electrical equipment manufacturing industry (S11), communication equipment, office equipment and other manufacturing industries (S12), power and heat production and supply industry (S13) and water production and supply industry (S15). Northeast China needs to improve weak industries such as stationery industry (S3), chemical industry (S6), metal products industry (S7), electrical appliance manufacturing industry (S11), communication equipment, office equipment and other manufacturing industry (S12), power and heat production and supply industry (S13) and water production and supply industry (S15). Communication equipment, office equipment and other manufacturing industries (S12) and power and heat production and supply industries (S13) in East and Central China need to be further improved. South China needs to focus on

improving the relatively backward industries such as food, tobacco and wine industry (S1), textile and clothing industry (S2), stationery industry (S3) and metal products industry (S7). Southwest China needs to focus on textile and clothing industry (S2), stationery industry (S3), petroleum processing industry (S4), chemical industry (S6), metal products industry (S7) and electrical manufacturing industry (S11). The northwest region needs to vigorously develop the stationery industry (S3), chemical industry (S6), transportation equipment industry (S10) and water production and supply industry (S15).

(3) At the provincial level and industry level, priority should be given to improving the carbon productivity of 8 key industries in 15 key provinces. On the whole, it is necessary to improve potential hybrid knowledge emission ratio effect *PHGCRE*. The government should strengthen investment in science, technology and education, train high-tech talents, and promote the research, development and application of patents. The government should introduce low-carbon technologies in various industries to reduce carbon emissions, so as to vigorously develop low-carbon industries, balance the structure of low-carbon industries and energy-intensive industries, and achieve balanced development. For the eight key industries, each province should adjust the input proportion of various production factors, optimize the energy structure, and reduce carbon emissions under the condition of ensuring stable and healthy economic development, so as to reduce the gap in carbon productivity between provinces. Provinces with close geographical locations should implement assistance policies to improve carbon productivity.

(4) In terms of specific industries in specific provinces, the chemical industry in Hebei should focus on the development of hybrid knowledge stock to replace hybrid energy, adjust the proportion of energy use, increase investment in scientific and technological research and development, and improve technical efficiency and carbon emission performance. Shanxi Province should focus on the development of labor substitution for hybrid energy in communication equipment, office equipment and other manufacturing industries, reduce the use of fossil energy, and reduce carbon emissions. Inner Mongolia should make full and efficient use of natural resources and improve its ability to absorb and store carbon dioxide. The three northeastern provinces as a whole need to develop advanced production technologies and improve traditional production methods. They should vigorously develop capital to replace hybrid energy, labor to replace hybrid energy and hybrid knowledge stock to replace hybrid energy, adjust the input structure of production factors, and increase the research, development and use of green patents. Anhui should increase scientific and technological investment in black gold, gold processing and non-gold and metal products industries, promote scientific and technological innovation, especially the use of clean energy, to reduce carbon dioxide emissions. Shandong Province should focus on adjusting the energy consumption structure in the province and improve the level of carbon productivity fundamentally. Guangxi Province

needs to focus on improving the substitution of capital, labor force and hybrid knowledge stock for hybrid energy, and improve the proportion of production factors input, so as to improve the level of technology and carbon emission performance. Hainan Province should make full use of the advantages of natural resources, actively develop renewable energy, reduce carbon emissions, and absorb carbon dioxide that has been generated. The combination of the two will further improve carbon productivity. Guizhou needs to increase the substitution of capital and hybrid knowledge stock for hybrid energy, optimize the input structure of production factors, and comprehensively improve carbon productivity. The four provinces in northwest China need to focus on developing new energy, such as solar power, wind power, and ecological photovoltaic power, to reduce fossil energy consumption and reduce carbon emissions.

Author contributions

MF: Conceptualization, Methodology, Project administration, Supervision, Writing–original draft. YM: Data curation, Methodology, Investigation, Writing–original draft. LT: Funding acquisition, Methodology, Resources, Supervision, Writing–review and editing. CZ: Data curation, Software, Visualization, Writing–review and editing.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Dong F, Yu B, Hadachin T, Dai Y, Wang Y, Zhang S, et al. Drivers of carbon emission intensity change in China. *Resour Conservation Recycling* (2018) 129:187–201. doi:10.1016/j.resconrec.2017.10.035

2. Jin P, Peng C, Song M. Macroeconomic uncertainty, high-level innovation, and urban green development performance in China. *China Econ Rev* (2019) 55:1–18. doi:10.1016/j.chieco.2019.02.008

3. Wang Y, Shen N. Environmental regulation and environmental productivity: the case of China. *Renew Sustain Energ Rev* (2016) 62:758–66. doi:10.1016/j.rser.2016.05.048

4. Jin J, Ma XY, Gao YS. China's carbon emission productivity and its development tendency from an environmental protection perspective. *Nat Environ Pollut Tech* (2017) 16(4):1293–301.

5. Chen YF, Ma LH, Zhu ZT. The environmental-adjusted energy efficiency of China's construction industry: a three-stage undesirable SBM-DEA model. *Environ Sci Pollut Res* (2021) 28(41):58442–55. doi:10.1007/s11356-021-14728-2

6. Scheel H. Undesirable outputs in efficiency valuations. *Eur J Oper Res* (2001) 132(2):400–10. doi:10.1016/s0377-2217(00)00160-0

7. Sueyoshi T, Yuan Y, Goto M. A literature study for DEA applied to energy and environment. *Energ Econ* (2017) 62:104–24. doi:10.1016/j.eneco.2016.11.006

8. Plank B, Eisenmenger N, Schaffartzik A, Wiedenhofer D. International trade drives global resource use: a structural decomposition analysis of raw material consumption from 1990–2010. *Environ Sci Tech* (2018) 52(7):4190–8. doi:10.1021/acs.est.7b06133

9. Moutinho V, Madaleno M, Inglesi-Lotz R, Dogan E. Factors affecting CO2 emissions in top countries on renewable energies: a LMDI decomposition application. *Renew Sustain Energ Rev* (2018) 90:605–22. doi:10.1016/j.rser.2018.02.009

10. Isaksen A, Trippl M. Innovation in space: the mosaic of regional innovation patterns. *Oxford Rev Econ Policy* (2017) 33(1):122–40. doi:10.1093/oxrep/grw035

11. Crescenzi R, Rodríguez-Pose A. The geography of innovation in China and India. *Int J Urban Reg Res* (2017) 41(6):1010–27. doi:10.1111/1468-2427.12554

12. Wan B, Tian L. Health-education-disaster green low-carbon endogenous economic growth model and its new accompanying effects. *J Clean Prod* (2022) 359:131923. doi:10.1016/j.jclepro.2022.131923

13. Kaya Y, Yokobori K. *Environment, energy, and economy: strategies for sustainability*. Tokyo: United Nations University Press (1997).

14. McKinsey Global Institute. *The carbon productivity challenge: curbing climate change and sustaining economic growth*. New York, NY: McKinsey and Company, McKinsey Global Institute (2008).

15. Aigner DJ, Lovell CAK, Schmidt P. Formulation and estimation of stochastic frontier production function models. *J Econom* (1977) 6(1):21–37. doi:10.1016/0304-4076(77)90052-5

16. Charnes A, Cooper WW, Rhodes E. Measuring the efficiency of decision making units. *Eur J Oper Res* (1978) 2(6):429–44. doi:10.1016/0377-2217(78)90138-8

17. Kumbhakar SC, Lovell CAK. *Stochastic frontier analysis*. New York: Cambridge University Press (2003).

18. Li J, Li P. Analysis on urban production efficiency and its influencing factors in the three provinces of NortheastNortheast China—based on the study of stochastic frontier analysis method with three factors input. *Econ Horizon* (2018) 35(1):14–21. doi:10.15931/j.cnki.1006-1096.20171129.013

19. Fang Y, Ma R. Study on the cultural trade potential and influencing factors of countries along the belt and road based on the stochastic frontier gravity model. *World Econ Res* (2018) 1:112–21. doi:10.13516/j.cnki.wes.2018.01.009

20. Färe R, Grosskopf S. Modeling undesirable factors in efficiency evaluation: comment. *Eur J Oper Res* (2004) 157(1):242–5. doi:10.1016/s0377-2217(03)00191-7

21. Chen S Environmental pollution emissions, regional productivity growth and ecological economic development in China. *China Econ Rev* (2015) 35:171–82. doi:10.1016/j.chieco.2014.08.005

22. Seiford LM, Zhu J. Modeling undesirable factors in efficiency evaluation. *Eur J Oper Res* (2002) 142(1):16–20. doi:10.1016/s0377-2217(01)00293-4

23. Gomes EG, Lins MPE Modelling undesirable outputs with zero sum gains data envelopment analysis models. *J Oper Res Soc* (2008) 59(5):616–23. doi:10.1057/palgrave.jors.2602384

24. Hua Z, Bian Y, Liang L. Eco-efficiency analysis of paper mills along the Huai River: an extended DEA approach. *Omega* (2007) 35(5):578–87. doi:10.1016/j.omega.2005.11.001

25. Chen F, Zhao T, Wang J. The evaluation of energy–environmental efficiency of China's industrial sector: based on Super-SBM model. *Clean Tech Environ Policy* (2019) 21(7):1397–414. doi:10.1007/s10098-019-01713-0

26. Gurgul H, Lach Ł. Linkages-based indicators of production-source sectoral eco-efficiency with application to Polish data. *J Clean Prod* (2021) 279:123545. doi:10.1016/j.jclepro.2020.123545

27. Pan XF, Pan XY, Ming Y, Zhang J. The effect of regional mitigation of carbon dioxide emission on energy efficiency in China, based on a spatial econometrics approach. *Carbon Manage* (2018) 9(6):665–76. doi:10.1080/17583004.2018.1537514

28. Li SJ, Wang SJ. Examining the effects of socioeconomic development on China's carbon productivity: a panel data analysis. *Sci Total Environ* (2019) 659:681–90. doi:10.1016/j.scitotenv.2018.12.409

29. Wang ZG, Su B, Xie R, Long H. China's aggregate embodied CO_2 emission intensity from 2007 to 2012: a multi-region multiplicative structural decomposition analysis. *Energ Econ* (2020) 85:104568. doi:10.1016/j.eneco.2019.104568

30. Ang BW, Choi KH. Decomposition of aggregate energy and gas emission intensities for industry: a refined Divisia index method. *Energ J* (1997) 18(3):59–73. doi:10.5547/issn0195-6574-ej-vol18-no3-3

31. Ang BW LMDI decomposition approach: a guide for implementation. *Energy Policy* (2015) 86:233–8. doi:10.1016/j.enpol.2015.07.007

32. Ang BW, Zhang FQ, Choi KH. Factorizing changes in energy and environmental indicators through decomposition. *Energy* (1998) 23(6):489–95. doi:10.1016/s0360-5442(98)00016-4

33. Chong CH, Tan WX, Ting ZJ, Liu P, Li Z. The driving factors of energy-related CO_2 emission growth in Malaysia: the LMDI decomposition method based on energy allocation analysis. *Renew Sustain Energ Rev* (2019) 115:109356. doi:10.1016/j.rser.2019.109356

34. Zhang C, Cai WH, Yu TS, Hu SJ, Liu HY, He P. Regional economic growth and carbon productivity: an analysis based on convergence and decoupling index. *China Ind Econ* (2013) 5:18–30. doi:10.1186/1999-3110-54-18

35. Cheng Y, Sun YX, Wang XY. Study on the influence and countermeasures of global technological innovation on carbon productivity. *China Popul Resour Environ* (2019) 29(9):30–40. doi:10.12062/cpre.20190625

36. Zhou P, Ang BW. Decomposition of aggregate CO_2 emissions: a production-theoretical approach. *Energ Econ* (2008) 30(3):1054–67. doi:10.1016/j.eneco.2007.10.005

37. Wang H, Zhou P. Multi-country comparisons of CO_2 emission intensity: the production-theoretical decomposition analysis approach. *Energ Econ* (2018) 74:310–20. doi:10.1016/j.eneco.2018.05.038

38. Yang H, Lu ZN, Shi XP, Mensah IA, Luo Y, Chen W. Multi-region and multi-sector comparisons and analysis of industrial carbon productivity in China. *J Clean Prod* (2021) 279:123623. doi:10.1016/j.jclepro.2020.123623

39. Liu B, Ding CJ, Hu J, Su Y, Qin C. Carbon trading and regional carbon productivity. *J Clean Prod* (2023) 420:138395. doi:10.1016/j.jclepro.2023.138395

40. Fang GC. *The path optimization and policy synergy of new energy development under carbon trading-driving*. Beijing, China: China Financial and Economic Publishing House (2023).

41. Wan B, Tian L, Zhang W, Zhang G. Environmental effects of behavior growth under green development. *Environ Develop Sustainability* (2022) 25:10821–55. doi:10.1007/s10668-022-02508-y

42. Färe R, Grosskopf S, Noh DW, Weber W. Characteristics of a polluting technology: theory and practice. *J Econom* (2005) 126(2):469–92. doi:10.1016/j.jeconom.2004.05.010

43. Wang C. Sources of energy productivity growth and its distribution dynamics in China. *Resource Energ Econ* (2011) 33(1):279–92. doi:10.1016/j.reseneeco.2010.06.005

44. Zhang XP, Zhang J, Tan QL. Decomposing the change of CO_2 emissions: a joint production theoretical approach. *Energy policy* (2013) 58:329–36. doi:10.1016/j.enpol.2013.03.034

45. Dietzenbacher E, Hoen AR, Los B. Labor productivity in Western Europe 1975–1985: an intercountry, interindustry analysis. *J Reg Sci* (2000) 40(3):425–52. doi:10.1111/0022-4146.00182

46. Guan YR, Shan YL, Huang Q, Chen H, Wang D, Hubacek K. Assessment to China's recent emission pattern shifts. *Earth's Future* (2021) 9(11):e2021EF002241. doi:10.1029/2021ef002241

Appendix A

TABLE A1 Seven geographical divisions and symbols of China.

Region	Name and symbol
North China	Beijing (BJ), Tianjing (TJ), Hebei (HE), Shanxi (SX), Inner Mongoria (IM)
Northeast China	Liaoning (LN), Jilin (JL), Heilongjiang (HL)
East China	Shanghai (SH), Jiangsu (JS), Zhejiang (ZJ), Anhui (AH), Fujian (FJ), Jiangxi (JX), Shandong (SD)
Central China	Henan (HA), Hubei (HB), Hunan (HN)
South China	Guangdong (GD), Guangxi (GX), Hainan (HI)
Southwest China	Chongqing (CQ), Sichuan (SC), Guizhou (GZ), Yunnan (YN)
Northwest China	Shaanxi (SN), Gansu (GS), Qinghai (QH), Ningxia (NX)
Reference region	Reference region (U)

Research on multi-layer network topology optimization strategy for railway internet of things based on game theory benefits

Fang Wang[1†], Kaixuan Su[2†], Bo Liang[3]*, Jian Yao[3] and Wei Bai[3]

[1]Department of Mathematics, Xinzhou Normal University, Xinzhou, China, [2]School of Mathematics and Statistics, Shanxi University, Taiyuan, China, [3]Institute of Computing Technology, China Academy of Railway Sciences Company Ltd., Beijing, China

In the railway system environment, the interconnection of a vast array of intelligent sensing devices has brought about revolutionary changes in the management and monitoring of railway transportation. However, this also poses challenges to the communication service quality within the railway Internet of Things (IoT). Through collective intelligence and collaboration, the nodes within the railway IoT can not only share data and information but also work synergistically to enhance the overall intelligence level and improve decision-making quality of the network. Therefore, this paper proposes a reconnection mechanism based on the computation of node game-theoretic benefits and optimizes this process with the concept of swarm intelligence collaboration. Initially, the game-theoretic benefit values of the nodes in the railway IoT network are calculated. Subsequently, based on the weight priority of the edges, the two edges with the larger weights are selected, and connections are established between nodes with similar game-theoretic benefit values to enhance the network's robustness. This approach enables rapid networking and efficient communication transmission within the railway IoT, providing robust assurance for the safe and stable operation of the railway.

1 Introduction

With the rapid development of new generation information [1] and communication technologies (5G), the scale of Internet of Things (IoT) applications in the railway environment has increased dramatically [2]. The rapid expansion of IoT applications, while enhancing network service quality, also brings greater risks of network paralysis [3]. Device nodes in network applications are susceptible to failure due to various factors, such as natural disaster activities, node malfunctions, and malicious attacks. Moreover, since IoT applications are interconnected, the paralysis of a single network can easily lead to a chain reaction causing the collapse of the entire railway IoT system [4]. Therefore, in the face of partial device node failures, how to improve the quality of service (QoS) of complex IoT applications and their robustness against network attacks [5], and to maximize the maintenance of network topology communication capabilities, has become a bottleneck in the development of large-scale railway IoT applications [6].

When the topological structure of IoT applications is attacked, measuring and optimizing the robustness and reliability of the network topology is an important means to minimize the risk of network failures [7]. In practical dynamic issues such as the propagation and immunity of complex networks [8], and link control [9], different topological structures exhibit different levels of robustness in the face of deliberate attacks or random failures. Therefore, altering the network's topological structure is of significant importance for enhancing the robustness of complex networks.

The robustness of complex networks measures the network's ability to respond to external disturbances, such as deliberate attacks or random failures [10]. Networks that maintain their functionality under such changes (by removing some nodes or edges) are considered more robust than those that do not. Current research on improving the robustness of complex networks is mainly divided into two categories: increasing the connections within the network and reconnecting the network's edges. Many infrastructures have capacity limitations, such as transmission lines of power plants and the number of flights in aviation systems, making it impractical to increase the number of connections in a network. A substantial amount of research has confirmed that the reconnection mechanism [11] is a simple and effective method often used to adjust the network's topological structure to enhance its robustness. For the topological structure of complex dynamic networks, it is generally described by a coupling matrix, and most of the existing literature discusses the case of constant coefficient coupling matrices. In actual complex dynamic networks, due to the influence of external environments, the topological structure cannot remain constant. Therefore, it is necessary to introduce the concept of time-varying topological structures in complex dynamic networks. Some scholars have begun to consider this issue. For example, Refs. [12,13] discuss the time-varying coupling strength in complex dynamic networks, using adaptive laws to identify unknown parameters, achieving adaptive synchronization of complex dynamic networks; Ref. [9] discusses the control of complex networked supply chains with multiple time-delay couplings and time-varying topological structures, thereby enhancing the network's ability to resist collaborative attacks; and how the network's performance in the dynamics process feeds back and affects the evolution of the topological structure.

Currently, IoT topology optimization methods [12] primarily use the largest connected subgraph as a robustness metric to quantify and optimize network topologies. Since the IoT topology optimization problem is an NP-hard problem [14], to seek approximate optimal solutions, most IoT topology optimization methods are based on heuristic algorithms [15,16]. Meanwhile, in complex networks, the scale-free network model more closely resembles the structure of real-world networks and performs well in resisting random network attacks (where each device node fails or leaves the network with equal probability). However, it is prone to network paralysis when facing malicious attacks (where important nodes fail first). A multitude of researchers have proposed efficient strategies to enhance the stability and resilience of IoT topological structures. Rong et al [14] first classified the edges in the network as effective, ineffective, and flexible edges, and then proposed a heuristic optimization algorithm based on edge classification (EC), designed to enhance the robustness of scale-free (SF)

networks against malicious attacks (MA). Qiu et al. [13] introduced the ROSE strategy, an enhancement method for wireless sensors in scale-free networks, which identifies and protects key nodes in the network and optimizes the connection structure to fend off malicious attacks. Zhao et al studied [17] a specific type of network attack under the knapsack constraint—the maximum vertex cover attack, which aims to maximize the number of links associated with removed nodes under a limited budget. Game theory, which describes the micro-interactions of network nodes, is naturally a powerful tool to guide the adjustment of topological structures, and applying cooperative game theory to the optimization of IoT topologies will also face greater opportunities and challenges.

Based on this, to address the failure of some critical equipment nodes and the chain reaction collapse effect of network attacks, which affect the service quality of railway system communication applications, this paper proposes a reconnection mechanism calculated based on the measurement of game-theoretic benefits between nodes. The first chapter introduces the essential knowledge of complex networks and model construction, and on this basis, constructs a cooperative game strategy, including interaction strategies, transformation rules, and equilibrium balance analysis. The second chapter proposes the initialization of the model and the reconnection edge mechanism. The third chapter verifies the optimized network robustness based on the game value reconnection edge mechanism using Monte Carlo experiments in complex networks.

2 Construction of multi-layer topological network model in the railway internet of things environment

2.1 Network representation

The network is modeled as a weighted graph $G = (V, E)$, where $V = \{e_1, e_2, e_3 ..., e_n\}$ represents the set of nodes in the network and $E = \left\{(e_i, e_j) \mid e_i, e_j \in V\right\}$ represents the set of edges. Let $N = |V|$ denote the number of nodes and $M = |E|$ denote the number of edges. Each edge e_{ij} connects nodes e_i and e_j, and has a weight w_{ij}, which reflects the importance of the connection between the nodes.

This paper only considers simple connected undirected graphs without duplicate edges or self-loops, meaning that there is at most one edge between any pair of nodes in the network, each edge has two distinct endpoints, and there is at least one path between any two nodes. The neighborhood of a node v_i in the network G is denoted as $N_{v_i} = \{(u, v_i) \in E, u \in V\}$, representing the set of nodes adjacent to node v_i in the network G. To avoid confusion, it is simply noted as N_i. The degree $d_{v_i} = |N_i|$ of a node v_i, denoted as d_i, represents the number of edges associated with node v_i in the network G.

2.2 Game strategy

Swarm intelligence collaboration is a large-scale sensing and computing model based on collective intelligence, emphasizing the completion of complex sensing and computing tasks through the

TABLE 1 Rules of cooperative evolution in network structure.

Color	Previous round strategy	Current round strategy
Blue	Cooperate	Cooperate
Red	Defect	Defect
Green	Defect	Cooperate
Yellow	Cooperate	Defect

participatory and collaborative nature of the group. This model typically involves a large number of individuals who are interconnected through communication technologies such as the Internet, collectively engaging in a particular task or goal. The core of swarm intelligence collaboration lies in how to effectively organize and motivate these participants, as well as how to process and analyze the vast amounts of data they generate.

By integrating swarm intelligence collaboration with cooperative game theory, a more efficient and fair system can be created. In this system, participants (who can be humans or intelligent devices) are motivated to cooperate and jointly complete tasks or solve problems. Cooperative game theory provides the tools and methods to design such incentive mechanisms.

In the cooperative evolution within network structures, it is assumed that each individual engages in a round of public goods games with its neighbors at each clock cycle, where the neighbors are the subjects of the individual's game, related to the social network structure that hosts the gaming group. The model includes three types of topological structures for the gaming group: regular lattice networks, BA scale-free networks, and WS small-world networks. In each round of the game, there is no central leader to control; individuals have two choices: to cooperate or to defect. The specific strategy an individual chooses is related to a series of factors, which is also what the experiment will investigate. The strategies adopted will have an impact on neighboring subjects; adopting a cooperative strategy generates positive externalities, from which other subjects can benefit. These externalities reflect the social ability of subjects to interact with their adjacent subjects. After multiple rounds of the game, whether a stable evolutionary outcome can be achieved is also influenced by the introduction of incentive mechanisms in the experiment, which aim to shift the original evolutionary outcome from defection to cooperation. Each subject in the model is described by four elements:

1) The strategy for this round is either cooperation or defection;
2) The strategy from the previous round was either cooperation or defection;
3) The score for this round;
4) To facilitate the observation of results during the simulation process, subjects are also assigned different colors for distinction. The specific rules are as shown in Table 1.

2.3 Interaction and strategy transition rules

The most important part of evolutionary game theory is the interaction between individuals (interact) and the selection stage of the next round of strategy. Assuming all individuals are

homogeneous and have the same set of strategies. When an individual adopts a cooperative strategy, the net benefit it receives is the cooperative benefit R minus the cooperation cost c. For the convenience of studying the problem, let the benefit R be the total number of all neighboring nodes plus one (which is itself), and the cooperation cost is 1, so the net benefit for each node is its total number of neighboring nodes. When a node chooses a defection strategy, the defector will gain a certain benefit, and due to the cooperation of neighboring nodes, the defector will free-ride and enjoy the positive external benefits from the cooperating neighbors, so its benefit is the defection incentive multiplied by the total number of cooperators among its neighboring nodes.

Therefore, for each individual, the number of cooperators among its neighboring nodes is counted, and if the individual is a cooperator, its score is the total number of cooperators among its neighbors; if the individual is a defector, its score is the product of the defection benefit and the total number of cooperating neighbors. In this paper, we mainly adopt the strategy of choosing the one with the highest score among the neighboring nodes as the strategy for the next round, characterizing the feature of learning from the strong in evolutionary games.

2.4 Strategy space and equilibrium analysis

The incentive rules here are refined into rewards (R) and punishments (P). The incentive policy will either reward cooperators as the sole strategy or punish defectors as the sole strategy. Individuals will react to external stimuli, that is, they will consider the rewards or punishments given by the central authority before making decisions. The model compares three scenarios: no policy, reward policy, and punishment policy, to determine which one achieves the optimal outcome (the optimal outcome is determined by the number of cooperators), in order to identify the best solution. When a reward mechanism is introduced into the model, corresponding to rewards for cooperators in reality, a reward (reward) is introduced, and the utility function of cooperators increases, while the utility function of individuals adopting a defection strategy remains unchanged. When a punishment mechanism is introduced into the model, corresponding to punishments for defectors in reality, a punishment (punishment) is introduced, and the utility function of defectors will decrease by a value P, while the utility function of individuals adopting a cooperation strategy remains unchanged.

The individual payoff is characterized as follows:

TABLE 2 Experimental parameters and their value ranges.

Network structure parameters	Parameter meaning	Random network	BA scale-free network	WS small-world network
Num-nodes	Total Number of Network Nodes, Initial	500	1,000	1,500
Initial-cooperation	Cooperation Probability	0–1	0–1	0–1
Defection-award	Defection Reward	B	B	B
Reward	Reward	R	R	R
Punishment	Punishment	P	P	P

$$p_{xy} = p_x p_y R + (1 - p_x) p_y T + p_x (1 - p_y) S + (1 - p_x)(1 - p_y) P \quad (1)$$

The average revenue is as follows:

$$r_x = \sum_{y=1}^{|\Omega_x|} p_{xy} / |\Omega_x| \quad (2)$$

where $|\Omega_x|$ is the number of neighbors of node x.

The expected outcome is related to the number of initial cooperators, the size of the group, the magnitude of the defection payoff, and the network's topological structure. Subsequently, targeted improvements in parameter settings, as well as the magnitude of imposed incentives, expected rewards, and punishments, will also affect the final evolutionary outcome.

2.5 Paramter settings

The previous section introduced the algorithms of game theory; now we will begin to set initial values for the relevant parameters of network initialization, representing their meanings and ranges.

1) Num—nodes: It represents the total number of nodes in the network, which is also the scale of the gaming group, and its value range varies slightly depending on the network;
2) Initial—cooperation: It represents the initial probability of cooperation, with a value range between 0% to 100%;
3) Defection—award: It represents the defection payoff, with a value range between 0 to 10;
4) In models with incentive mechanisms, there are also parameters for reward and punishment.

Table 2 displays the experimental parameters and their value ranges.

3 Topology reconstruction of railway internet of things based on cooperative benefits

3.1 Network state initialization

Initializing the network requires setting the network's state information vector to the topological environmental state. The following description details how to convert the IoT topology structure into the corresponding environmental state vector, as shown in Figure 1. It is essential to reduce the storage space of the topology structure for large-scale topologies. First, the topology structure is converted into an adjacency matrix, an operation frequently used in other IoT applications. In this paper, by analyzing the network topology structure and optimization action operations, it is found that only network nodes within each other's communication range can have a connection relationship. Therefore, information from other nodes within the node's communication range is of significant value. As shown in Figure 1, node a's upper triangular relationship is [1, 0, 0, 0], but since a is not within the communication range of c and d, the only effective position is b with [1]; similarly, node b has no connection with node d and node e, so its effective position is c with [1]; node c's effective position is d with [1]; node d's effective position is a with [1], and node d has no effective position because it only considers the upper triangular matrix. In summary, finally, the topology is linked in the order of the adjacency matrix numbers to form an environmental vector of [0, 1, 1, 1].

3.2 Initialization of network edge weights

In the analysis of complex networks, the degree of a node plays a fundamental and crucial role. BA scale-free networks [17] typically exhibit scale-free characteristics, with a degree distribution that follows a power-law rule, indicating that most nodes in the network have only a few connections, while a small fraction of nodes have a large number of connections. In the context of cooperative and competitive relationships between nodes, the most important factor is the expected benefit of cooperation with neighbors. According to the clustering effect reflected in Ref. [18], it is known that the more neighbors a node has, the easier it is to form a cluster, so nodes with a higher degree may often generate better benefits over time. This paper assigns weights to each edge in the network by calculating the absolute difference in the expected benefits of node cooperation, thereby characterizing the similarity between nodes, and using the magnitude of edge weights to measure whether nodes have the conditions for reconnection, as shown in the following formula:

$$w_{ij} = |r_i - r_j| \quad (3)$$
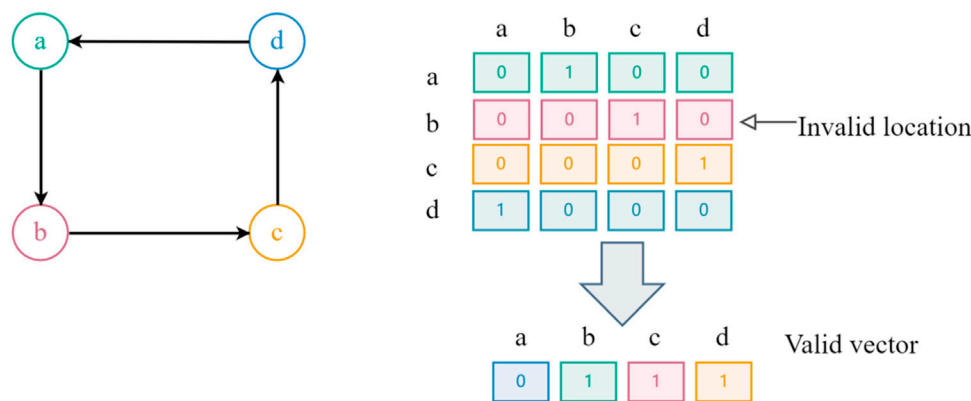
FIGURE 1
The state information vector of the network is transformed into the state of the topological environment.

where r_i represents the degree value of node v_i, and w_{ij} represents the weight of edge (v_i, v_j). The larger the weight value of the edge, the greater the difference between the two endpoints of that edge.

3.3 Edge selection strategy

To enhance the stability of the network, the strategy for selecting edges to swap is to disconnect edges with a large difference in expected benefits, while nodes with similar benefit values are more likely to connect with each other. The algorithm proposed in this paper assigns weights to each edge using Formula 3, and when selecting edges for reconnection, the decision is based on the weight of the edges. The probability of choosing two edges is directly proportional to the weight of the edges, following the rule as stated.

$$\Pi_{ij} = \frac{w_{ij}}{\sum\limits_{(u,v)\in E} w_{uv}} \tag{4}$$

According to Formula 4, edges with larger weights have a higher probability of being selected, and these edges typically connect two nodes with a large deviation in benefit values. Therefore, the algorithm can quickly identify the edges with larger differences and then select the two edges with the highest edge weights (requiring at least four nodes and two edges for disconnection and reconnection), and reconnect their endpoints. This can eliminate the benefit gap and rapidly enhance the stability of the network.

3.4 Node homophily

Homophily can analyze and understand the connection preferences between nodes, which is the phenomenon that similar nodes tend to connect with each other. Connection preferences also stem from the attributes of nodes, such as routing frequency, bandwidth, climatic environment, antenna technology, etc. This can help designers optimize network structures and improve network efficiency and performance. According to Formula 5, prioritize the reconnection operation of

four distinct nodes connected by two edges with larger weights. According to the benefit values, the four nodes are sorted in non-decreasing order, i.e., $r_1 \geq r_2 \geq r_3 \geq r_4$. Then, there are three ways for the reconnection mechanism to connect these four nodes: (1) $(v_{r_{(1)}}, v_{r_{(2)}})$ and $(v_{r_{(3)}}, v_{r_{(4)}})$; (2) $(v_{r_{(1)}}, v_{r_{(3)}})$ and $(v_{r_{(2)}}, v_{r_{(4)}})$; (3) $(v_{r_{(1)}}, v_{r_{(4)}})$ and $(v_{r_{(2)}}, v_{r_{(3)}})$.Therefore, this paper determines the optimal connection method between nodes by comparing the changes in the homophily coefficient of the three edge connection methods.

The homophily coefficient is an indicator that measures the similarity or difference in connections between nodes in a network. Choosing the reconnection method that results in the smallest change in the homophily coefficient can be considered optimal. The homophily coefficient reveals the preference for connections between nodes in the network. When nodes with higher connectivity in the network tend to connect with other nodes of similar connectivity, the network is referred to as a homophilous network. As shown in Formula 6.

$$\rho = 1 - \frac{\sum\limits_{(i,j)\in E} \left(r_i - r_j \right)^2}{\sum\limits_{i=1}^{N} r_i^3 - \frac{1}{2M}\left(\sum\limits_{i=1}^{N} r_i^2 \right)^2} \tag{5}$$

where r_i represents the expected benefit value of node v_i , N represents the number of nodes in the network, and M represents the number of edges in the network.

Since the reconstruction mechanism in this paper does not change the degree values of the nodes during the reconnection process. Therefore, $\sum\limits_{i=1}^{N} r_i^3 - \frac{1}{2M}(\sum\limits_{i=1}^{N} r_i^2)^2$ is a constant during the calculation process. The network's homophily coefficient is only related to the square of the difference in degree values at the two endpoints of each edge. Consequently, the difference in the homophily coefficient for the three edge connection methods is as shown in Formulas 6–8.

$$\rho_{(1)} - \rho_{(2)} = \frac{2\left(r_{(2)} - r_{(3)} \right)\left(r_{(1)} - r_{(4)} \right)}{\sum\limits_{i=1}^{N} r_i^3 - \frac{1}{2M}\left(\sum\limits_{i=1}^{N} r_i^2 \right)^2} \geq 0 \tag{6}$$

$$\rho_{(2)} - \rho_{(3)} = \frac{2\left(r_{(1)} - r_{(2)}\right)\left(r_{(3)} - r_{(4)}\right)}{\sum\limits_{i=1}^{N} r_i^3 - \frac{1}{2M}\left(\sum\limits_{i=1}^{N} r_i^2\right)^2} \geq 0 \qquad (7)$$

$$\rho_{(1)} - \rho_{(3)} = \frac{2\left(r_{(2)} - r_{(4)}\right)\left(r_{(1)} - r_{(3)}\right)}{\sum\limits_{i=1}^{N} r_i^3 - \frac{1}{2M}\left(\sum\limits_{i=1}^{N} r_i^2\right)^2} \geq 0 \qquad (8)$$

From Formulas 6–8, it can be seen that $\rho_{(1)} \geq \rho_{(2)} \geq \rho_{(3)}$, which means that the first reconnection method will cause the greatest increase in homophily. Therefore, this paper performs the reconnection operation based on the first method of edge connection, which is to sort the four nodes selected by edge weight in non-increasing order according to their degree values, creating two edges between nodes with similar game-theoretic benefits. Similarly, during the reconnection process, the network is not allowed to have duplicate edges and self-loops, and the network must remain connected. Additionally, this paper uses the collective cooperation rate to judge the cooperative performance of the network after reconnection; if the network cooperation increases, the reconnection operation is maintained, otherwise, the reconnection operation is undone.

4 Experimental simulation

4.1 Reconstruction network statistical indicator test

The experiment is conducted using the python, where the network is modeled as a weighted graph and evolves according to the method by Nowak and May in reference [19]. This method places "agents" in a two-dimensional spatial array to explore the changing spatial patterns of nodes. In this paper's experiment, the two-dimensional grid space is replaced with a more three-dimensional BA scale-free network. In the model, when a reward system is implemented to encourage cooperative behavior, the utility function of cooperators is enhanced, manifested as an additional reward, while the utility function of individuals adopting a defection strategy remains unchanged. Similarly, when the model adopts punitive measures to sanction defection, the utility function of defectors will suffer a loss, specifically reflected in the deduction of a fixed amount P as a penalty. In this case, the utility function of individuals choosing a cooperative strategy remains unchanged. The specific value is set to $R = 1, T = b, S = 0, P = 0$. At the beginning of the experiment, an initial probability of cooperation is randomly assigned to each node in the network graph. The evolutionary game strategy introduced in Chapter 1 is then used for calculations, with the defection payoff set to b = 1.2 and the time step MCS = 5,000. Subsequently, the network begins to evolve. After the network reaches a preliminary state of stability, reconnection operations are performed based on the average payoff of each node using the algorithm from Figure 2, and the reconnected network undergoes degree distribution testing and natural connectivity testing.

The degree distribution in a communication network describes the probability distribution of the number of connections (degree) of various nodes within the network, such as routers, switches, and terminal devices. This distribution influences the efficiency of communication within the network.

Through the simulation analysis of the optimization of complex network topological structures, the following basic conclusions can be drawn:

1) The simulation analysis has verified the effectiveness and feasibility of the complex network topology optimization model and the solution model. As the topological structure continues to be optimized, the natural connectivity value continuously increases, indicating that the network's resilience to destruction is constantly strengthening.

2) By comparing the results of Figures 3A, B, it can be observed that the degree distribution of the network has changed after optimization, where nodes with high degree values and those with low degree values have gradually decreased, while nodes with a moderate degree (around 74) have significantly increased. This results in the entire network's cooperation rate being in an optimal state, possibly because a very high degree value can lead to traffic congestion, which is not conducive to collective interests; too low a degree value can also affect the decrease in selfish benefits of individual entities. Therefore, the reconnection edge strategy in this paper serves as the best way to balance the interests of individuals and the collective.

3) Analyzing the relationship between natural connectivity and the number of evolutionary generations, as well as the relationship between fitness values and the number of evolutionary generations, as depicted in Figures 4, 5, reveals a common pattern. It is evident that both of these metrics increase with the increase in the number of evolutionary generations. This suggests that nodes with higher degrees are more likely to form new connections with other nodes that also have high degrees.

4.2 Robustness analysis

The purpose of complex network topology optimization is to enhance the network's resilience. The analysis mentioned above establishes a topology optimization model using natural connectivity as a measure of network resilience, significantly improving the network's resistance to destruction. Here, two types of strategies are adopted: random (node and edge removal) attacks and malicious (node and edge removal) attacks. These are implemented before and after the optimization of complex network topologies to analyze the changes in network resilience, as shown in Figures 6–9.

Through simulation analysis, it is known that complex network topologies before and after optimization exhibit the following characteristics when facing different attack strategies:

1) In the random node removal attack strategy (Figure 6), the optimized network's resistance to attacks is higher than the resistance before optimization, mainly reflected in the slower decline in network resilience as the proportion of node removal increases. Analyzing the network's topological structure, using the average degree as a dividing line, the proportion of nodes with low degrees in the optimized network increases, while the proportion of nodes with high degrees decreases. Therefore, under the random attack
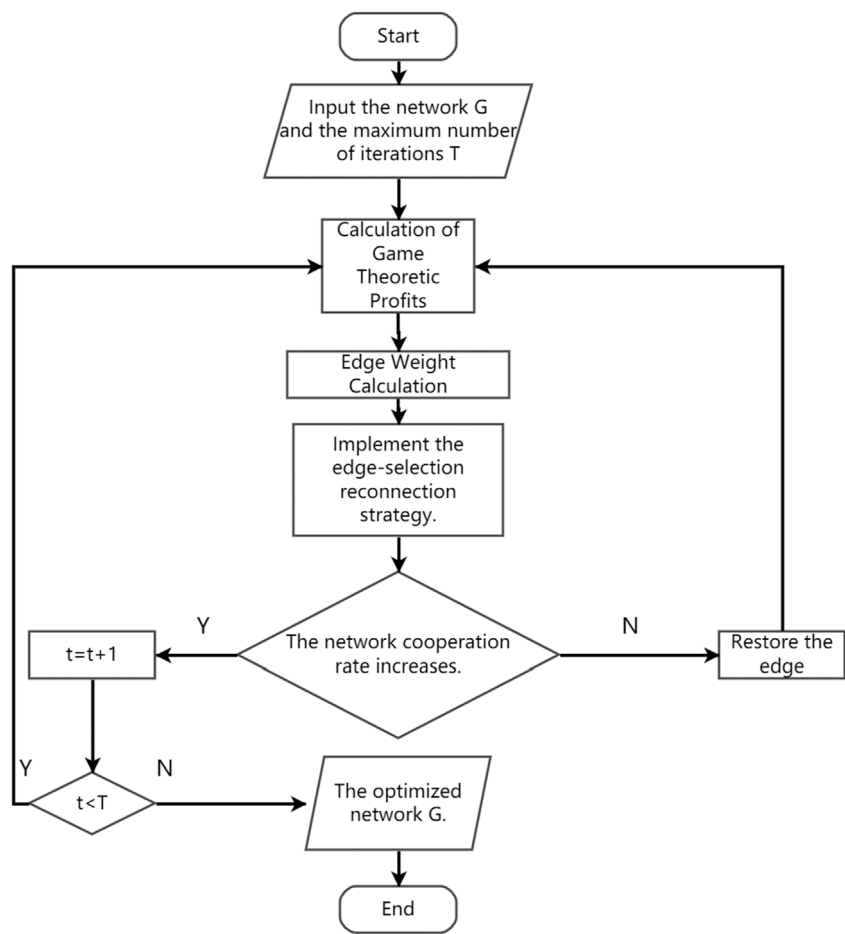
FIGURE 2
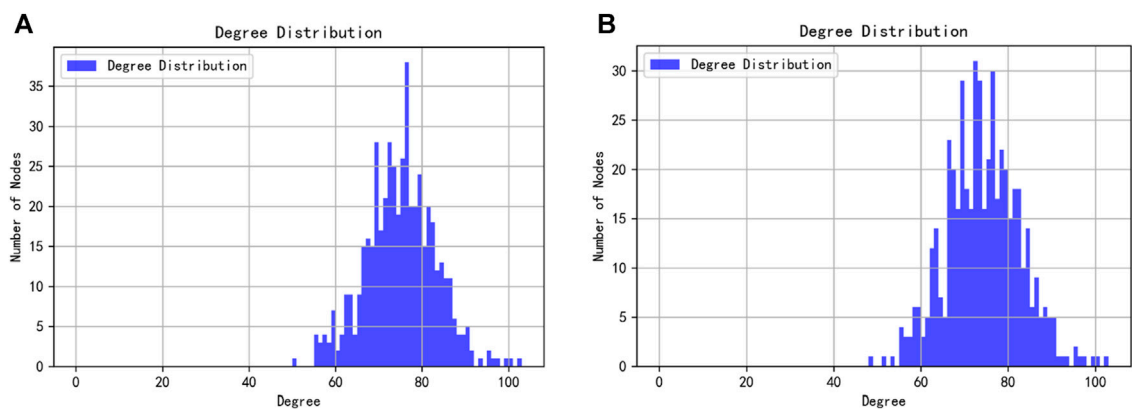Demonstration framework for optimization algorithms based on game reconnection mechanism.

FIGURE 3
Compared Degree Distribution between no optimized network **(A)** and optimized network **(B)**.

strategy, the average probability of low-degree nodes being attacked in the optimized network is higher than that of high-degree nodes, and the impact of low-degree nodes on network resilience is also smaller, thus demonstrating stronger resistance to attacks.

2) In the random edge removal attack strategy (Figure 7), the network's resistance to attacks is similar before and after optimization, and it is only when the edge removal ratio exceeds approximately 0.6 that the optimized network's resistance to attacks shows a slight improvement.
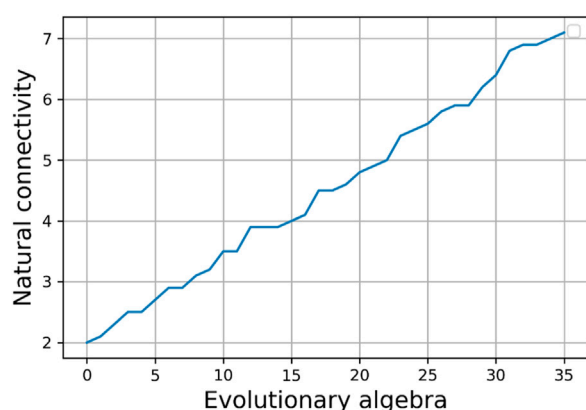
FIGURE 4
The relationship between natural connectivity and evolutionary generations.
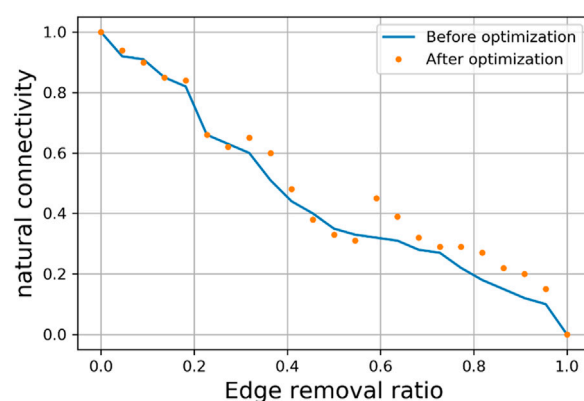
FIGURE 7
Comparison of random edge removal attack and initial connectivity.
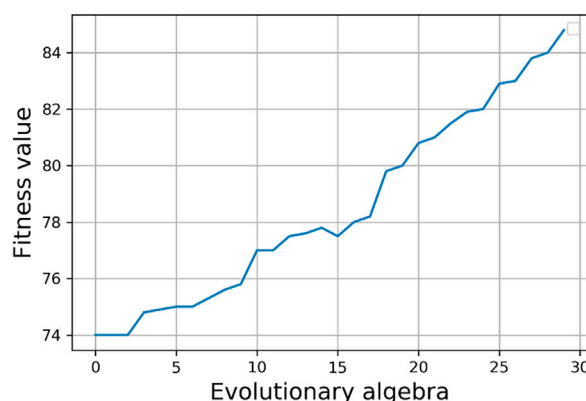
FIGURE 5
The relationship between fitness value and evolutionary generations.
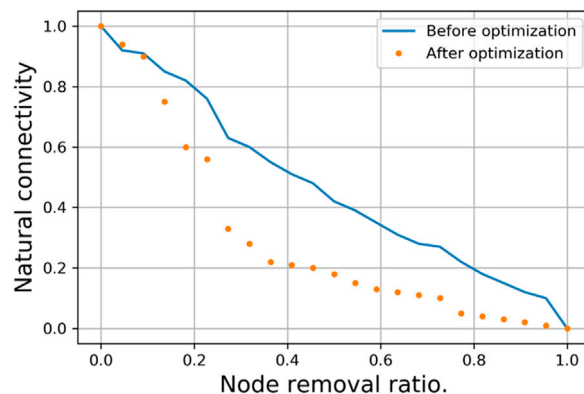
FIGURE 8
Comparison of degree-based node removal attack and initial connectivity.
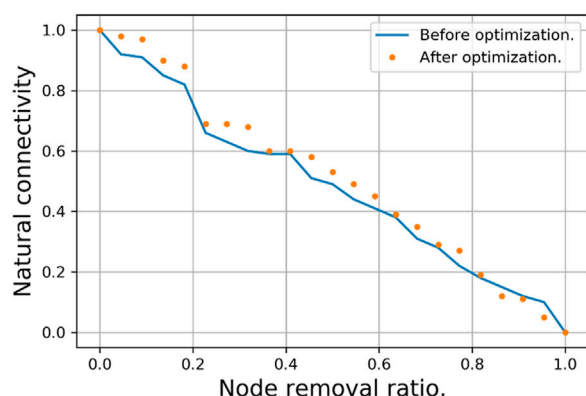
FIGURE 6
Comparison of random node removal attack and initial connectivity.

3) In the degree-based node removal attack strategy (see Figure 8), where nodes are removed in descending order of degree, the optimized network exhibits poorer resilience against this attack method. The removal of a small number (about 10%) of high-degree nodes leads to a significant decrease (about 40%) in network resilience, almost paralyzing the network; whereas the network before optimization demonstrates relatively stronger resilience against such attacks.

4) In the betweenness-based edge removal attack strategy (see Figure 9), where edges are removed in order of their betweenness centrality, the optimized network shows relatively stronger resilience against this attack method. The removal of a minority (about 20%) of edges with high betweenness has a limited impact on network resilience. The analysis attributes this to the tendency of low-degree nodes in the optimized network to connect with high-degree nodes, which results in the edges connecting high-degree nodes also having relatively high betweenness, thus the edges have stronger redundancy.

FIGURE 9
Comparison of betweenness-based edge removal attack and initial connectivity.
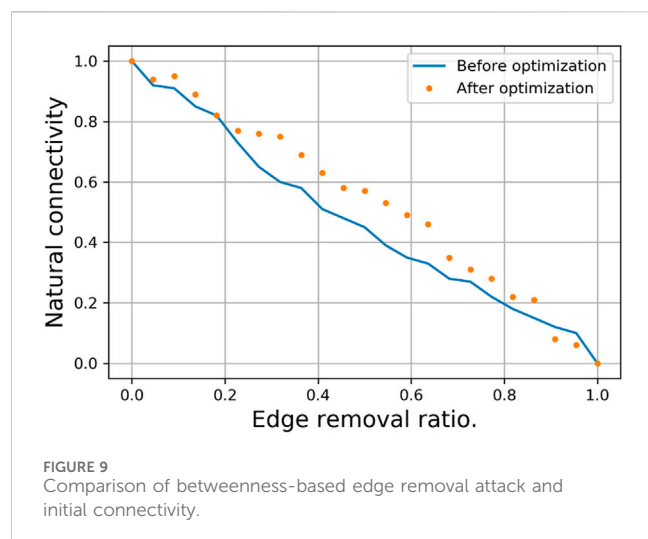
Conversely, for the network structure before optimization, the situation is the opposite.

This paper considers attack strategies, which belong to the dynamic optimization of network topological structures. The optimized network has enhanced resilience against node removal attacks.

5 Conclusion

This paper introduces a rewiring strategy based on node game payoff, which assigns initial cooperation rates to nodes through cooperative game play, and adjusts cooperation benefits during dynamic interactions. It introduces reward and punishment mechanisms to achieve average payoff. In this process, each network node not only makes decisions based on its own payoff but also considers the collective intelligence synergy of the entire network to achieve overall optimization. Building on this, the paper uses the average payoff between nodes for rewiring and evaluates the rewiring effect with the collective cooperation rate. The experimental results show that the optimization algorithm proposed in this paper performs better in terms of the proportion of nodes in natural connectivity and robustness indicators when subjected to attacks such as random node attacks, random edge attacks, degree centrality attacks, and betweenness centrality attacks on the multi-layer network topology of the Internet of Things (IoT) in the railway sector. It is also found that the network exhibits poor resistance to attacks when facing degree-based node removal attacks. Therefore, in practical applications, it is necessary to choose different network topological structures based on the specific real threats faced. In conclusion, the research findings of this paper provide a reference for achieving rapid networking, efficient transmission, and secure and stable operation of communication in the IoT for the railway sector.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

FW: Formal Analysis, Methodology, Software, Validation, Writing–original draft, Writing–review and editing. KS: Investigation, Methodology, Software, Validation, Writing–original draft, Writing–review and editing. BL: Conceptualization, Funding acquisition, Investigation, Writing–review and editing. JY: Conceptualization, Supervision, Validation, Writing–review and editing. WB: Software, Supervision, Validation, Writing–original draft, Writing–review and editing.

Funding

Conflict of interest

Authors BL, JY, and WB were employed by China Academy of Railway Sciences Company Ltd.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Skouby KE, Lynggaard P. Smart home and smart city solutions enabled by 5G, IoT, AAI and CoT services. In: *International conference on contemporary computing and informatics (IC3I)* (2014). p. 874–8.

2. Varma PS, Anand V. Intelligent scanning period dilation based Wi-Fi fingerprinting for energy efficient indoor positioning in IoT applications. *J supercomputing* (2023) 79:7736–61. doi:10.1007/s11227-022-04980-9

3. Yang Y, Zhou W, Zhao SR. Survey of IoT security research: threats, detection and defense. *J Commun* (2021) 42(8):188–205. doi:10.11959/j.issn.1000-436x.2021124

4. Nie Y, Li J, Liu G, Zhou P. Cascading failure-based reliability assessment for post-seismic performance of highway bridge network. *Reliability Eng Syst Saf* (2023) 238: 109457. doi:10.1016/j.ress.2023.109457

5. Hilal AM, Alohali MA, Al-Wesabi FN, Nemri N, Alyamani HJ, Gupta D. Enhancing quality of experience in mobile edge computing using deep learning based data offloading and cyberattack detection technique. *Cluster Comput* (2023) 26:59–70. doi:10.1007/s10586-021-03401-5

6. Stiawan D, Wahyudi D, Septian TW. The development of an Internet of Things (IoT) network traffic dataset with simulated attack data. *J Internet Tech* (2023) 24: 345–56. doi:10.53106/160792642023032402013

7. Dilek GF, Salman FS. Assessing the reliability and the expected performance of a network under disaster risk. *OR Spectr quantitative approaches Manag* (2011) 33(3): 499–523. doi:10.1007/s00291-011-0250-7

8. Bródka P, Musial K, Jankowski J. Interacting spreading processes in multilayer networks: a Systematic Review. *IEEE Access* (2020) 8:10316–41. doi:10.1109/access.2020.2965547

9. Jalili M, Orouskhani Y, Asgari M, Alipourfard N, Perc M. Link prediction in multiplex online social networks. *R Soc Open Sci* (2017) 4(2):160863. doi:10.1098/rsos.160863

10. Albert R, Jeong H, Barabasi A. Error and attack tolerance of complex networks. *Nature* (2000) 406(6794):378–82. doi:10.1038/35019019

11. Mcdonald DC *Optimally reconnecting weighted graphs against an edge-destroying adversary* (2021).

12. Qiu T, Liu J, Si W, Wu DO. Robustness optimization scheme with multi population co-evolution for scale-free wireless sensor networks. *IEEE/ACM Trans Networking* (2019) 27:1028–42. doi:10.1109/tnet.2019.2907243

13. Zhao T, Si W, Li W, Zomaya AY. Optimizing the maximum vertex coverage attacks under knapsack constraint. *IEEE/ACM Trans Networking* (2021) 29(3): 1088–104. doi:10.1109/tnet.2021.3056450

14. Qiu T, Zhao A, Xia F, Si W, Wu DO. ROSE: robustness strategy for scale-free wireless sensor networks. *IEEE/ACM Trans Networking* (2017) 25:2944–59. doi:10. 1109/tnet.2017.2713530

15. Mateev V, Marinova I. Modified chromosome pooling genetic algorithm for resource allocation optimization. *AIP Conf Proc* (2023) 2939(1). doi:10.1063/5. 0178510

16. Rong L, Liu J. A heuristic algorithm for enhancing the robustness of scale-free networks based on edge classification. *Physica A: Stat Mech its Appl* (2018) 503:503–15. doi:10.1016/j.physa.2018.02.173

17. Vázquez A, Moreno Y. Resilience to damage of graphs with degree correlations. *Phys Rev E* (2003) 67(1):015101. doi:10.1103/physreve.67.015101

18. Wang Y, Wang BH. Evolution of cooperation on spatial network with limited resource. *PLoS ONE* (2015) 10(8):e0136295. doi:10.1371/journal.pone.0136295

19. Nowak MA, May RM. Evolutionary games and spatial chaos. *Nature* (1992) 359: 826–9. doi:10.1038/359826a0

Dynamic evolutionary analysis of opinion leaders' and netizens' uncertain information dissemination behavior considering random interference

Lin Ma[1], Bowen Li[2]* and Junyao Wang[3]

[1]School of Arts, University of Science and Technology Liaoning, Anshan City, China, [2]School of Business Administration, University of Science and Technology Liaoning, Anshan City, China, [3]School of Electronic and Information Engineering, University of Science and Technology Liaoning, Anshan City, China

This paper investigates the decision-making behaviors of opinion leaders and netizens in the context of uncertain information dissemination with the aim of effectively managing online public opinion crises triggered by major sudden events. The decision-making behaviors of opinion leaders are categorized into positive and negative guidance, while those of netizens are classified into acceptance and nonacceptance. Using an evolutionary game model, this study introduces random factors to examine their influence on the decision-making processes of both groups. A stochastic evolutionary game model is constructed to analyze the behaviors of opinion leaders and netizens in the context of uncertain information dissemination. The evolutionary stability strategies and stochastic evolutionary processes of the model are analyzed based on the theory of Itô stochastic differential equations. The impacts of key variables such as random disturbances, the degree of psychological identification of netizens with opinion leaders, and the intensity of government penalties for those spreading negative information are examined through numerical simulations. The findings indicate that opinion leaders evolve to make stable strategies more rapidly than netizens do; random disturbances slow the evolution of stable strategies for both groups but do not alter their strategic choices; a higher degree of psychological identification increases the likelihood of netizens adopting the views of opinion leaders; and as punitive measures intensify, both opinion leaders and netizens are inclined to choose strategies of positive guidance and acceptance. The results of this study offer theoretical insights and decision-making guidance for future government strategies for managing similar online collective behaviors.

KEYWORDS

stochastic evolution game, opinion leaders, psychological identification, random disturbance, netizens

1 Introduction

The emergence of major emergencies can cause great damage to the lives and property of governments and the public. For example, the 6.2 magnitude earthquake that struck Gansu Province, China, on 18 December 2023 and the 7.4 magnitude earthquake that struck off the west coast of Honshu, Japan, on 1 January 2024 resulted in a considerable loss of life

and extensive property damage. The harm caused by major emergencies is not only limited to the impact of the event itself but also to the various secondary impacts it causes, which can also hinder the smooth operation of society. Sometimes, the damage caused by secondary impacts may even exceed the damage caused by the major emergency itself. With the continuous iteration of internet technologies, the internet is gradually emerging as the predominant conduit through which the general public obtains information. Individuals within society can conveniently access platforms such as Facebook, X (Twitter), TikTok, Weibo, and WeChat through computers or mobile devices at any time and from any location to acquire information of interest and engage in real-time communication and interaction with others. Therefore, when a major emergency occurs in a certain place, netizens can obtain all kinds of related information quickly andonduct information exchange around it. However, because major emergencies are characterized by randomness, wide coverage, uncertain duration and a large degree of harm, relevant departments are unable to publish all the information related to major emergencies in a timely and effective manner, resulting in a large amount of unsubstantiated and uncertain information on social platforms. If opinion leaders choose to guide netizens negatively for certain purposes, netizens will make incorrect judgments after adopting the information, thus spreading panic and even triggering online public opinion crises resulting in mass incidents. This will not only cause great obstacles to relevant government departments in dealing with major emergencies, but also have a negative impact on the prosperity and stability of society. Therefore, analyzing the decision-making behaviors of opinion leaders and netizens in the dissemination of uncertain information to identify strategic choices that can contribute to the smooth operation of society are highly important for preventing and controlling online public opinion crises triggered by major emergencies.

Based on the summary and combination of previous studies, to analyze the decision-making behaviors of opinion leaders and netizens in uncertain information dissemination after major emergencies occur, this paper combines stochastic analysis theory with an evolutionary game model. First, we take opinion leaders and netizens as research objects and construct an evolutionary game model between them. Second, considering that random factors from the internal or external environment affect decision-making behavior, Gaussian white noise is introduced on the basis of the evolutionary game model to construct a stochastic evolutionary game model between opinion leaders and netizens. Again, the stochastic evolutionary game model is solved, the conditions when the model reaches a stable state are analyzed, and the numerical solution of the equilibrium solution of the model is solved using stochastic Taylor expansion. Finally, the model is numerically simulated using MATLAB 2017b to analyze the effects of different variables on the model evolution process.

The main contributions of the research reported in this paper are as follows: 1) Considering that after the occurrence of major emergencies, opinion leaders and netizens are the main actors in the dissemination of uncertain information in social platforms. Therefore, we study the decision-making behavior of opinion leaders and netizens. 2) When constructing the model, we took into account that netizens may not fully trust the statements released by opinion leaders. Therefore, we included netizens' psychological

identification with opinion leaders in the model. 3) Compared to other scholars' research on the decision-making behaviors of different groups in the process of uncertain information dissemination, this paper incorporates random disturbance factors into the model construction. The random disturbance factors not only includes the bounded rationality of opinion leaders and netizens but also reflects the complexity of the real world and the impact of random interference factors on their decision-making process. This paper introduces Gaussian white noise into the construction of the evolutionary game model of opinion leaders and netizens. In this way, this paper depicts the evolution of decision-making behaviors of opinion leaders and netizens in the process of uncertain information dissemination more realistically. The research in this paper provides the corresponding theoretical basis and decision-making reference for the future exploration of the decision-making behavior of opinion leaders and netizens in uncertain information dissemination in the context of major emergencies.

2 Related work

The embryonic form of evolutionary game theory was initially applied by biologists in research pertaining to the species evolution of animals and plants. It was not until 1973 that the evolutionary stability strategy proposed by Smith et al. [1] marked the formalization of evolutionary game theory. Evolutionary game theory is widely used in sociology, management, cybernetics, biology and other disciplines because it does not require game subjects to be completely rational, and the information between game subjects is not fully disclosed. Wölfl et al. [2] used the evolutionary game model to study the evolution of cancer. Li et al. [3] analyzed the decision-making behaviors of the government, online media and netizens in the process of disinformation dissemination by constructing a three-party evolutionary game model. Wang et al. [4] analyzed the decision-making behaviors of workers, platforms and requesters in spatial crowdsourcing based on an evolutionary game model. Wu et al. [5] studied the behavior of enterprises facing energy transition in the carbon trading market. Shi et al. [6] studied decision-making behaviors among the government, service providers and elderly people in a smart aging system. Sun et al. [7] studied the decision-making behavior of civil aviation and high-speed rail under a carbon trading price. Shi et al. [8] studied the behavior of the government, automotive suppliers and logistics companies in the context of green transition. As Perc et al. [9] and Tanimoto et al. [10] explored whether the emergence of random disturbances will affect the payoffs of the game system. Different game subjects in a real environment are affected by the uncertainties they experience or by the external environment, which results in great uncertainty in decision-making behavior among game subjects. Therefore, some scholars have combined stochastic analysis theory with evolutionary game theory, introduced Gaussian white noise into the construction of models, and constructed stochastic evolutionary game models. For example, Mo et al. [11] suggested that random factors interfere with the decision-making behavior of game subjects and constructed a stochastic evolutionary game model to study the decision-making behavior of different game subjects in the online car market. Kang

et al. [12] constructed a stochastic evolutionary game model among multinational corporations, international dealers and the government. Xie et al. [13] contructed the model to study the behavior of market regulators and risky units in electricity markets. Du et al. [14] constructed an evolutionary game model between e-commerce firms and banks without considering the interference of stochastic factors and a stochastic evolutionary game model considering the interference of stochastic factors. Scatà et al. [15] used the stochastic evolutionary game model to investigate the cooperation mechanism between humans.

Scholars in various countries have carried out related research from various perspectives against the background of major emergencies. Among them, some scholars have studied the impact of major emergencies. Mos et al. [16] believe that major emergencies can cause enormous financial risks and economic losses. Cheng et al. [17] analyzed the oil price data of the month when major emergencies occurred from 2009 to 2020 and found that major emergencies in major economies around the world would seriously affect international oil prices. Some other scholars, studied the behavior of different types of actors after major emergencies. For example, Jia et al. [18] constructed a stochastic evolutionary game model based on the Moran process to study the effects of different game actors on prevention and control behavior in public health emergencies. Using game theory, Özkaya et al. [19] investigated the effect of self-isolation on the spread of COVID-19. Salarpour et al. [20] used COVID-19 as a background for a game-theoretic study of the behavior of countries in the supply of medical supplies at different stages. Wang et al. [21] studied the dispatching behavior of relief supplies after a major emergency. Meanwhile, some scholars conducted research about online public opinion caused by major emergencies. Wei [22] used the theory of heat conduction to study the propagation behavior of online public opinion after major emergencies. Chen et al. [23] argue that the topics that are widely discussed over time will generate new sub topics which will combine with existing public opinion to form multidimensional new public opinion. Lu et al. [24] argue that netizens pay limited attention to different events occurring at the same time and do not pay attention to all the information. Apuke et al. [25] argued that factors such as altruism, quality of information, and netizens' thirst for information determine whether false information will be spread on social media. Chew et al. [26] argued that relevant government departments can utilize online social media platforms to disseminate truthful information related to major emergencies, thereby reducing public panic. Yu et al. [27] argued that when there is a major emergency, spreading false information related to the event will increase public panic. Zhang et al. [28] argued that in major emergencies, the identity and social influence of opinion leaders have a positive effect on the efficiency of crisis information dissemination. Guan et al. [29] argued that information released by opinion leaders with a certain social status is more likely to be adopted by netizens, and in most cases, the gender factor does not affect netizens' decision-making behavior. Alvarez-Galvez [30] argues that even if a message is not adopted by the majority of netizens, they will choose to adopt the message when a central opinion leader supports the message. Parsegov et al. [31] argue that netizens are able to discuss topics of interest in social networks and form clusters of different types of networks.

In past research on different actors after major emergencies occurred, few scholars have studied decision-making behavior in the process of uncertain information dissemination. Based on previous research, this paper combines stochastic analysis theory with evolutionary game models to analyze the decision-making behavior of opinion leaders and netizens in uncertain information dissemination after major emergencies occur. Considering that random factors from internal or external environments can affect the decision-making behavior of both parties, we construct a stochastic evolutionary game model between 2 research objects, opinion leaders and netizens, by introducing Gaussian white noise. This study provides theoretical basis and decision-making reference for exploring the decision-making behavior of opinion leaders and netizens in uncertain information dissemination under the background of major emergencies in the future.

3 Construction of the evolutionary game model

The workflow comprises four key steps: 1) Defining game subjects and their decision-making behaviors; 2) Constructing the model; 3) Computing equilibrium solutions; 4) Analyzing evolutionary stability strategies. Organizational diagram of the current study as shown in Figure 1.

3.1 Basic assumptions and notation

In the era of new media, the public can use social platforms to obtain information triggered by major emergencies worldwide, but due to the diverse forms, large quantities and rich content of network information, it is impossible for the public to grasp all the information. Therefore, the public can watch the audio and video of video bloggers, live broadcasts of anchors, and information organized by self-media to obtain rapid access to relevant information. This paper defines video bloggers, anchors, self-media outlets and other people who have a certain degree of discourse power and are able to influence and shape the opinions of others through their own speech or behavior as opinion leaders. The members of the public who can use online social media platforms to obtain information are defined as netizens. Based on this, this paper defines opinion leaders and netizens as the main actors in the process of disseminating uncertain information after major emergencies. Both of them are limited rational participants with learning ability; in the case of incomplete information, they cannot judge whether the choices they make can maximize the benefits the first time, but due to their learning ability, they can make choices toward the strategy of maximizing the benefits in the process of learning continuously.

After major emergencies occur, opinion leaders need to integrate and sort out the relevant information related to major emergencies and propose their own views for netizens to adopt to gain attention. Since opinion leaders have the right to speak, which leads to their remarks being able to guide netizens' thoughts, the gaming strategy of opinion leaders can include positive guidance (by investigating the relevant events and releasing the real information about the

TABLE 1 Gain matrix of the two-party game between opinion leaders and netizens.

		Internet user	
		Acceptance	Nonacceptance
Opinion leader	Positive guidance	$(-C_1 +I_1, \lambda M_1 +I_3 -(1-\lambda) C_4 -C_3)$	$(-C_1 -L_1, -C_2)$
	Negative guidance	$(I_1 +I_2 -\pi L_2, -\lambda M_2 +I_4 -C_3 -(1-\lambda) C_4)$	$(-\pi L_2 -L_1, -C_2)$

events) or negative guidance (by rumor mongering, releasing unconfirmed and false information, stirring up the netizens' emotions, and other means to make profits). The netizens' game strategies include adopting or not adopting the opinions published by opinion leaders. Upon combinations of the above game strategies, the resultant scenarios encompass: (positive guidance, adoption), (negative guidance, adoption), (positive guidance, nonadoption), and (negative guidance, nonadoption).

This paper is based on the literature [32]. The following assumptions are made for this evolutionary game model:

(1) The probability that an opinion leader chooses posotive guidance or negative one is x and 1-x; the probability that a netizen chooses to adopt or not is y and 1-year, respectively.

(2) Neitziens' psychological identification with opinion leaders is $\lambda(0 < \lambda < 1)$. In this paper, it is argued that netizens' psychological identity toward opinion leaders is a kind of trust relationship formed after netizens pay attention to the remarks of opinion leaders for a long time; when the identity degree is zero, it means that netizens have no trust in opinion leaders; and when the identity degree is one, it means that netizens have complete trust in opinion leaders.

(3) When the opinion leader provides positive guidance, they will investigate and collect evidence on a series of uncertain information triggered by major emergencies, the cost of investigation will be C1, and the public speech from the opinion leader will be M_1; if the netizens adopt the speech from the opinion leader's positive guidance, then the opinion leader will receive the benefit of I_1. When the opinion leader leads negatively, since he or she publishes false information and does not verify the uncertain information, he or she does not pay investigation costs, and the published speech is M_2. However, since negative guidance is more capable of stirring up netizens' emotions, leading to netizens' panic or increasing conflict between netizens and the government, there is a probability of π for the opinion leader to be punished by the government, and the punishment is L_2. If the netizen adopts the speech published by the opinion leader under negative guidance, the opinion leader receives the benefit of I_2. As long as negatively guided speech is not adopted, the opinion leader loses L_1 due to a decrease in his or her popularity.

(4) When netizens agree with the opinion leader's speech, they will pay the cost of time and energy C_3 and the cost of judging the relevant speech (1-λ) C4. Meanwhile, they will also have a certain sense of satisfaction and participation, I_3 and I_4, respectively, and $I_4 > I_3$ because the speech in negative steering was more appealing. If netizens adopt positively guided speech, they will have a positive gain λM_1; if netizens adopt negatively guided speech, they will have a

negative gain λM_2 because the speech is not true but false and harmful information (recorded as $-\lambda M_2$ in the modeling operation). When netizens do not adopt any remarks of opinion leaders, they will experience panic (C_2) because they are in a state of information vacuum.

The above parameters C_1, C_2, C_3, C_4, I_1, I_2, I_3, I_4, L_1, L_2, M_1, M_2, π, λ are all greater than 0, where $0 \le x \le 1$, $0 \le y \le 1$, $0 \le \pi \le 1$, $0 < \lambda < 1$, $I_4 > I_3$.

3.2 Benefits matrix analysis

Based on the assumptions above, we can examine that when netizens adopt the positive guidance of opinion leaders, the total profit of the opinion leader is composed of their own investigation cost C_1 and the profit obtained from netizens' adoption I_1, the gain of opinion leaders is $-C_1 + I_1$; the total profit of netizens is composed of their own time and effort cost C_3, the cost of judging the statements released by opinion leaders $(1-\lambda) C_4$ and the profit from obtaining true information from opinion leaders, as well as the satisfaction and sense of participation gained from participating in the discussion of uncertain information $\lambda M_1 +I_3 -(1-\lambda) C_4 -C_3$. In this way, when netizens adopt the negative guidance of opinion leaders, the total profit of the opinion leader is $I_1 +I_2 -\pi L_2$ and the total profit of netizens is $-\lambda M_2 +I_4 -C_3 -(1-\lambda) C_4$. When netizens do not adopt the positive guidance of opinion leaders, the gain of opinion leaders is $-C_1 - L_1$, and the gain of netizens is $-C_2$. When netizens do not adopt the negative guidance of opinion leaders, the gain of opinion leaders is $-\pi L_2 - L_1$, and the gain of the netizen is $-C_2$. The gain matrix is shown in Table 1.

Combining the benefit matrix in Table 1, let the expected benefit of opinion leaders choosing positive guidance be U_{11}, the expected benefit of opinion leaders choosing negative guidance be U_{12}, and the average benefit of opinion leaders be $\overline{U_1}$; then, U_{11}, U_{12}, and $\overline{U_1}$, respectively:

$$U_{11} = y(-C_1 + I_1) + (1 - y)(-C_1 - L_1) = y(I_1 + L_1) - (C_1 + L_1) \tag{1}$$

$$U_{12} = y(I_1 + I_2 - \pi L_2) + (1 - y)(-\pi L_2 - L_1)$$
$$= y(I_1 + I_2 + L_1) - (\pi L_2 + L_1) \tag{2}$$

$$\overline{U_1} = xU_{11} + (1 - x)U_{12}$$
$$= -xyI_2 + x(\pi L_2 - C_1) + y(I_1 + I_2 + L_1) - (\pi L_2 + L_1) \tag{3}$$

Based on Eqs 1–3, it can be obtained that, the replication dynamic equation for opinion leaders' choice of positive guidance strategy is:

$$F(x) = \frac{dx}{dt} = x(U_{11} - \overline{U_1}) = x(1-x)(U_{11} - U_{12})$$

$$= x(1-x)(-yI_2 + \pi L_2 - C_1) \tag{4}$$

Similarly, let the expected benefit of the netizen's choice to adopt be U_{21}, the expected benefit of the netizen's choice not to adopt be U_{22}, and the average benefit of the netizen be $\overline{U_2}$; then, U_{21}, U_{22}, and $\overline{U_2}$, respectively:

$$U_{21} = x[\lambda M_1 + I_3 - (1-\lambda)C_4 - C_3]$$
$$+ (1-x)[-\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3] \tag{5}$$

$$U_{22} = x(-C_2) + (1-x)(-C_2) = -C_2 \tag{6}$$

$$\overline{U_2} = yU_{21} + (1-y)U_{22}$$
$$= xy[\lambda M_1 + I_3 - (1-\lambda)C_4 - C_3]$$
$$+ (1-x)y[-\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3] - (1-y)C_2$$
$$= xy[\lambda M_1 + I_3 + \lambda M_2 - I_4]$$
$$+ y[-\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2] - C_2 \tag{7}$$

Based on Eqs 5–7, it can be obtained that, the replication dynamic equation for netizens' choice of adoption strategy is:

$$F(y) = \frac{dy}{dt} = y(U_{21} - \overline{U_2}) = y(1-y)(U_{21} - U_{22})$$
$$= y(1-y)[x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4$$
$$- (1-\lambda)C_4 - C_3 + C_2] \tag{8}$$

4 Construction of a stochastic evolutionary game model

4.1 Introduction of Gaussian white noise

Due to limitations, the above Eq. 4 considers only the expected and overall benefits of opinion leaders adopting positive guidance strategies and does not take into account the difference between the benefits of opinion leaders adopting positive guidance and the benefits of adopting negative guidance; therefore, the above Eq. 4 adjusted to:

$$\frac{dx}{dt} = x(-yI_2 + \pi L_2 - C_1) \tag{9}$$

The above Eq. 8 considers only the expected and overall benefits of adopting an adoption strategy by netizens and does not take into account the difference between the benefits of adopting and the benefits of not adopting by netizens; therefore, the above Eq. 8 adjusted to:

$$\frac{dy}{dt} = y[x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2] \tag{10}$$

Considering that in the real world, opinion leaders and netizens are influenced by uncertain factors caused by their own or external environment, making their decision-making behavior highly uncertain. Therefore, to better understand the influence of the random perturbation term on the decision-making behavior of the game subjects, this section combines stochastic analysis

theory and introduces Gaussian white noise into the replication dynamic equation of the evolutionary game between opinion leaders and netizens Eqs 9, 10 in the game, i.e.,

$$dx(t) = x(t)(-yI_2 + \pi L_2 - C_1)dt + \sigma x(t)dw(t) \tag{11}$$

$$dy(t) = y(t)\begin{bmatrix} x(\lambda M_1 + \lambda M_2 + I_3 - I_4) \\ -\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2 \end{bmatrix}dt + \sigma y(t)dw(t) \tag{12}$$

where $w(t)$ is a one-dimensional standard Brownian motion; $dw(t)$ is Gaussian white noise; when $t > 0$, the step size $h > 0$, and its increment $\Delta w(t) = w\pi(t+h) - w(t)$ obeys the normal distribution $N(0, \sqrt{h})$; σ refers to the intensity of the random perturbation; and when σ is 0, the model is not affected by random disturbing factors.

The above equation is a one-dimensional Itô stochastic differential equation containing a Gaussian random disturbance term, which represents the replicated dynamic equations for opinion leaders and netizens subjected to random perturbations.

Compared with the evolutionary game model between opinion leaders and netizens constructed in literature [32], the stochastic evolutionary game model between opinion leaders and netizens constructed in this paper not only takes into account the impact of known factors on both parties when making behavioral choices but also considers the influence of uncertainties caused by themselves or the external environment. After introducing Gaussian white noise, the stochastic evolutionary game model differs from the ordinary evolutionary game model in the calculation of equilibrium solutions. The ordinary evolutionary game is mathematically expressed based on ordinary differential equations, while the stochastic evolutionary game model is mathematically expressed based on stochastic differential equations.

4.2 Existence and stability analysis of equilibrium solutions

Analyze the existence and stability of equilibrium solutions for stochastic evolutionary game models separately. Firstly, evaluate the existence of equilibrium solutions. Lemma 1 below is a sufficient condition for the solutions of stochastic differential equations to satisfy existence and uniqueness.

Lemma 1. There is a stochastic process $x = \{x(t), t \geq 0\}$ satisfying the Itô differential equation:

$$dx(t) = f(t, x(t))dt + g(t, x(t))dw(t), \forall 0 \leq t \leq T \tag{13}$$

If $f(t, x), g(t, x): [0, T] \times R \to R$ satisfies all of the following conditions:

(1) The measurability condition, i.e., $f(t, x), g(t, x)$ is binary measurable, $|f(t, x)|^{\frac{1}{2}}, |g(t, x)| \in L^2_{T \times R}$

(2) The Lipschitz condition, i.e., the existence of a constant H such that $|f(t, x) - f(t, y)| + |g(t, x) - g(t, y)| \leq H|x - y|$, where $\forall 0 \leq t \leq T, x, y \in R$;

(3) Linear growth bounded condition: There exists a positive constant P such that $|f(t, x)| + |g(t, x)| \leq P(1 + |x|)$, where $\forall 0 \leq t \leq T, x \in R$;

(4) Initial conditions: $x(t_0)$ on F_{t_0} is measurable and $Ex^2(t_0) < \infty$.

Then, there exists a unique process $x = \{x(t), t \geq 0\}$ satisfying the Itô differential Eq. 13, and x(t) is adaptive; with respect to F_{t_0} being measurable, then $Ex^2(t_0) < \infty$ and $\forall 0 \leq t \leq T$.

Proposition 1. A stochastic differential equation has a unique solution under $x, y \in [0, 1]$, $t \in [0, T]$.

Proof:

(1) Measurability condition

Rewrite the equations as $dx(t) = f_i(x, y)dt + g_i(x, y)dw(t)$, $i = 1, 2$, Among them,

$$\begin{cases} f_1(x, y) = x(-yI_2 + \pi L_2 - C_1) \\ g_1(x, y) = \sigma x \\ f_2(x, y) = y[x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2] \\ g_2(x, y) = \sigma y \end{cases}.$$

Clearly, $f_1(t, x)$, $g_1(t, x)$, $f_2(t, x)$ and $g_2(t, x)$ are continuous at $[0, 1] \times [0, 1]$, and condition (1) holds.

(2) Lipschitz condition

Due to $\forall x, x^* \in [0, 1]$:

$$\left| f_1(x, y) - f_1(x^*, y) \right| = \left| -yI_2 + \pi L_2 - C_1 \right| \left| x - x^* \right| \leq \left| \pi L_2 - C_1 \right| \left| x - x^* \right| \tag{14}$$

$$\left| g_1(x, y) - g_1(x^*, y) \right| = \left| \sigma x - \sigma x^* \right| \leq \left| \sigma \right| \left| x - x^* \right| \tag{15}$$

where $\left| \pi L_2 - C_1 \right|$ and $\left| \sigma \right|$ are both positive constants. Assume $H = \max\{\left| \pi L_2 - C_1 \right|, \left| \sigma \right|\}$: Based on Eqs 14, 15, it can be obtained that, $\max\{\left| f_1(x, y) - f_1(x^*, y) \right|, \left| g_1(x, y) - g_1(x^*, y) \right|\} \leq H \left| x - x^* \right|$, therefore, when $y \in [0, 1]$, the Eq. 11 satisfies the local Lipschitz condition. Similarly, Eq. 12 also satisfies the local Lipschitz condition. Thus, condition (2) holds.

(3) Linear growth bounded condition

For any $x \in [0, 1]$:

$$\begin{aligned} \left| f_1(x, y) \right|^2 &= \left| (-yI_2 + \pi L_2 - C_1)x \right|^2 \leq \left| \pi L_2 - C_1 \right|^2 \left| x \right|^2 \\ &\leq \left| \pi L_2 - C_1 \right|^2 (1 + \left| x \right|^2) \end{aligned} \tag{16}$$

$$\left| g_1(x, y) \right|^2 = \left| \sigma x \right|^2 \leq \left| \sigma \right|^2 (1 + \left| x \right|^2) \tag{17}$$

where $\left| \pi L_2 - C_1 \right|^2$ and $\left| \sigma \right|^2$ are both positive constants, such that $P = \max\{\left| \pi L_2 - C_1 \right|^2, \left| \sigma \right|^2\}$, then: Based on Eqs 16, 17, it can be obtained that, $\max\{\left| f_1(x, y) \right|, \left| g_1(x, y) \right|\} \leq P(1 + \left| x \right|^2)$. Therefore, the equation satisfies the linear growth bounded condition. Similarly, the Eq. 12 also satisfies the linear growth bounded condition. Thus, condition (3) holds.

(4) Initial conditions

The initial condition (4) clearly holds.

Thus, Proposition 1 is proved.

From Proposition 1, we can see that there is an equilibrium solution in the stochastic evolution game model of opinion leaders and netizens; i.e., under the premise of no Gaussian white noise interference, opinion leaders and netizens can reach a consensus with each other at the beginning and choose a strategy that meets

their own interests. However, in reality, with the passage of time and changes in the internal or external environment, opinion leaders and netizens may be affected by random interference, which changes their decision-making behavior. For example, after a period of time after a major emergency, as relevant investigations are conducted, an increasing number of details are reported, and netizens are able to obtain more information, which makes netizens quickly choose whether to adopt or not to adopt the comments released by opinion leaders. This random factor plays an important role in the original decision-making behavior, and the random factor plays a crucial role in the selection of the final stabilization strategies of opinion leaders and netizens. Therefore, fully considering random factors is more conducive to understanding the decision-making behaviors of opinion leaders and netizens in real society.

Next, the stability of the equilibrium solution is analyzed based on Lemma 2 below:

Lemma 2. Let the stochastic differential equation $x = \{x(t), t \geq 0\}$ satisfy the solution of the Itô differential equation initial value problem:

$$dx(t) = f(t, x(t))dt + g(t, x(t))dw(t), \forall t \geq 0, x(t_0) = x_0 \tag{18}$$

There exist continuous differentiable functions $V(t, x)$ with positive constants c_1 and c_2 such that $c_1 \left| x \right|^p \leq V(t, x) \leq c_2 \left| x \right|^p$.

(1) If there exists a positive constant γ such that $LV(t, x) \leq -\gamma V(t, x), t \geq 0$, then the Eq. 18 of the zero solution of the pth order moment index is stable and holds $E \left| x(t, x_0) \right|^p \leq (\frac{c_2}{c_1}) \left| x_0 \right|^p e^{-\gamma t}, t \geq 0$.

(2) If there exists a positive constant γ such that $LV(t, x) \geq \gamma V(t, x), t \geq 0$, then the Eq. 18 of the zero solution of the p-order moment index is unstable and holds $E \left| x(t, x_0) \right|^p \geq (\frac{c_2}{c_1}) \left| x_0 \right|^p e^{-\gamma t}, t \geq 0$.

Among others, $LV(t, x) = V_t(t, x) + V_x(t, x)f(t, x) + \frac{1}{2}g^2(t, x)V_{xx}(t, x)$.

Proposition 2. For the Eq. 11, when $c_1 = c_2 = 1$, $p = 1, \gamma = 1, V(t, x) = x^2$, $LV(t, x) = 2x^2(-yI_2 + \pi L_2 - C_1) + \sigma^2 x^2$.

(1) The exponential stability condition for the zero solution moment of the stochastic evolution equation satisfying the opinion leader is DL_1.

When $y \geq \frac{-1 - \sigma^2 - 2\pi L_2 + 2C_1}{-2I_2}$ and $-\pi L_2 + C_1 + I_2 \geq \frac{1 + \sigma^2}{2}$, DL_1.

(2) The zero-solution moment-exponential instability condition for a stochastic evolutionary equation satisfying the opinion leader is DL_2.

When $y \leq \frac{1 - \sigma^2 + 2C_1 - 2\pi L_2}{-2I_2}$ and $\pi L_2 - C_1 \geq \frac{1 - \sigma^2}{2}$, DL_2.

Proof: For Eq. 11, when $c_1 = c_2 = 1, p = 1, \gamma = 1$, $V(t, x) = x^2$, $LV(t, x) = 2x^2(-yI_2 + \pi L_2 - C_1) + \sigma^2 x^2$.

(1) When the zero-solution moment exponent is stabilized, the equation is required to satisfy $2x^2(-yI_2 + \pi L_2 - C_1) + \sigma^2 x^2 \leq -x^2$. Because $x \in [0, 1]$ and $-I_2 < 0$, $2(-yI_2 + \pi L_2 - C_1) + \sigma^2 \leq -1$; that is, $y \geq \frac{-1 - \sigma^2 - 2\pi L_2 + 2C_1}{-2I_2}$; and because $y \in [0, 1]$, $\frac{-1 - \sigma^2 - 2\pi L_2 + 2C_1}{-2I_2} \leq 1$, the solution is $-\pi L_2 + C_1 + I_2 \geq \frac{1 + \sigma^2}{2}$. Based on this, we can obtain Eq. 11 The exponential stabilization condition for the zero-solution moment of Eq. $y \geq \frac{-1 - \sigma^2 - 2\pi L_2 + 2C_1}{-2I_2}$ and $-\pi L_2 + C_1 + I_2 \geq \frac{1 + \sigma^2}{2}$.

(2) When the zero-solution moment exponent is unstable, the equation is required to satisfy $2x^2(-yI_2 + \pi L_2 - C_1) + \sigma^2 x^2 \geq x^2$. Because $x(t) \in [0, 1]$ and $-I_2 < 0$, $2(-yI_2 + \pi L_2 - C_1) + \sigma^2 \geq 1$; that is, $y \leq \frac{1 - \sigma^2 + 2C_1 - 2\pi L_2}{-2I_2}$; and because $y \in [0, 1]$, $\frac{1 - \sigma^2 + 2C_1 - 2\pi L_2}{-2I_2} \geq 0$; thus, the solution is $\pi L_2 - C_1 \geq \frac{1 - \sigma^2}{2}$. Based on this, we can obtain Eq. 11 The zero-solution-moment exponential instability condition of Eq. $y \leq \frac{1 - \sigma^2 + 2C_1 - 2\pi L_2}{-2I_2}$ and $\pi L_2 - C_1 \geq \frac{1 - \sigma^2}{2}$.

Proposition 2 suggests that the stabilizing strategy of opinion leaders will be influenced by their own factors as well as by random factors. When condition DL_1 is satisfied, opinion leaders will eventually choose the strategy of negative guidance; i.e., over time, no matter what the initial state is, opinion leaders will eventually reach a stable state under the strategy of negative guidance after continuously adjusting their decision-making behavior. When condition DL_2 is met, the opinion leader will eventually choose the strategy of positive guidance; that is, over time, regardless of the initial state, the opinion leader will eventually reach a stable state under the strategy of positive guidance after constantly adjusting his or her decision-making behavior.

From condition DL_1 and condition DL_2 of Proposition 2, it is clear that opinion leaders are more inclined to choose a strategy that is in their own interest. That is, opinion leaders adopt the strategy of negative steering when the expected benefit of choosing negative steering is greater than the expected benefit of choosing positive steering; conversely, they adopt the strategy of positive steering. This finding is consistent with the actual situation because the decision-making behavior of opinion leaders is influenced by factors such as penalties from regulatory agencies and the traffic generated by the attention of netizens, and which strategy is adopted is determined by the expected benefit of that strategy. However, random factors from internal or external sources can also have an impact on the final decision-making behavior of opinion leaders. For example, to smear the public image of China, hostile forces force opinion leaders to make negatively guided choices by means of coercion and enticement when major emergencies occur. In addition, to help the government reduce the impact of major emergencies as soon as possible, opinion leaders always choose to guide them positively, regardless of their own interests. All of these random factors may affect the decision-making behavior of opinion leaders, causing the final stabilization strategy to fluctuate or change.

Proposition 3. For Eq. 12, when $c_1 = c_2 = 1$, $p = 1$, $\gamma = 1$, $V(t, y) = y^2$, $LV(t, y) = 2y^2 [x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 y^2$.

(1) The exponential stability condition for the zero-solution moment of the stochastic evolutionary equation that satisfies the netizen is $DW_1 \cup DW_2$.

DW_1 when $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \leq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $\lambda M_2 - I_4 + (1 - \lambda)C_4 + C_3 - C_2 \geq \frac{1}{2}(1 + \sigma^2)$.

DW_2 when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \geq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $(1 - \lambda)C_4 + C_3 - C_2 - \lambda M_1 - I_3 \geq \frac{1}{2}(1 + \sigma^2)$.

(2) The zero-solution moment-exponential instability condition for the stochastic evolutionary equations that satisfy the netizens is $DW_3 \cup DW_4$.

DW_3 when $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \geq \frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $(\lambda M_1 + I_3) - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$.

DW_4 when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \leq \frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $-\lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$.

Proof: For Eq. 12, when $c_1 = c_2 = 1$, $p = 1$, $\gamma = 1$, $V(t, y) = y^2$, $LV(t, y) = 2y^2 [x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 y^2$.

(1) When the zero-solution moment exponent is stabilized, the equation is required to satisfy $2y^2 [x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 y^2 \leq -y^2$. Since $y \in [0, 1]$, $2[x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 \leq -1$.

In the case when $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \leq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$, since $x \in [0, 1]$, $\frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)} \geq 0$, and $\lambda M_2 - I_4 + (1 - \lambda)C_4 + C_3 - C_2 \geq \frac{1}{2}(1 + \sigma^2)$ is solved.

In the case when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \geq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$, since $x \in [0, 1]$, $\frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)} \leq 1$, and $(1 - \lambda)C_4 + C_3 - C_2 - \lambda M_1 - I_3 \geq \frac{1}{2}(1 + \sigma^2)$ is solved.

Based on this, one can obtain Eq. 12 The exponential stability condition for the zero solution moment is:

When $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \leq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $\lambda M_2 - I_4 + (1 - \lambda)C_4 + C_3 - C_2 \geq \frac{1}{2}(1 + \sigma^2)$; alternatively, when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \geq \frac{2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2 - (1 + \sigma^2)}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $(1 - \lambda)C_4 + C_3 - C_2 - \lambda M_1 - I_3 \geq \frac{1}{2}(1 + \sigma^2)$.

(2) When Eq. 12 The zero solution moment exponent is unstable, the equation needs to satisfy $2y^2 [x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 y^2 \geq y^2$. Since $y \in [0, 1]$, $2[x(\lambda M_1 + \lambda M_2 + I_3 - I_4) - \lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2] + \sigma^2 \geq 1$.

In the case when $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \geq \frac{(1-\sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$, because $x \in [0, 1]$, $\frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)} \leq 1$, and $(\lambda M_1 + I_3) - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$ is solved.

In the case when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \leq \frac{(1-\sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$, because $x \in [0, 1]$, $\frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)} \geq 0$, and $-\lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$ is solved.

Based on this, one can obtain Eq. 12 The exponential instability condition for the zero solution moment is:

When $\lambda M_1 + \lambda M_2 + I_3 - I_4 > 0$, $x \geq \frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $(\lambda M_1 + I_3) - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$; alternatively, when $\lambda M_1 + \lambda M_2 + I_3 - I_4 < 0$, $x \leq \frac{(1 - \sigma^2) + 2\lambda M_2 - 2I_4 + 2(1-\lambda)C_4 + 2C_3 - 2C_2}{2(\lambda M_1 + \lambda M_2 + I_3 - I_4)}$ and $-\lambda M_2 + I_4 - (1 - \lambda)C_4 - C_3 + C_2 \geq \frac{1}{2}(1 - \sigma^2)$.

Proposition 3 suggests that netizens' stabilizing strategies are influenced by their own factors as well as by random factors. When $DW_1 \cup DW_2$ is satisfied, netizens will eventually choose the strategy of nonadoption; i.e., over time, regardless of the initial state, after continuously adjusting their own decision-making behavior, they will eventually reach a stable state under the strategy of nonadoption. When $DW_3 \cup DW_4$ is met, netizens will eventually choose the strategy of adoption; i.e., over time, regardless of the initial state, netizens will eventually reach a stable state under the

strategy of adoption after continuously adjusting their decision-making behavior.

From the condition and the condition of Proposition 3, it is clear that netizens are more inclined to choose a strategy that is in their own interest. That is, when the expected benefit of not adopting is greater than the expected benefit of adopting, netizens will choose to not adopt; in contrast, they will adopt the strategy of adopting. Because netizens' decision-making behavior is influenced by the degree to which they pay attention to major emergencies and by their psychological recognition of opinion leaders, which strategy to adopt is determined by the expected benefit of the strategy. Moreover, random factors from internal or external sources may also have an impact on netizens' final decision-making behavior. For example, at the early stage of major emergencies, the investigation of major emergencies by relevant departments has just begun, and specific real information has not yet been released in time; in these cases, netizens can only pay attention to major emergencies through the remarks released by opinion leaders. In this case, netizens can only pay attention to the comments of opinion leaders. Or for other reasons, the attention of netizens may be drawn to other events and they may give up paying attention to major emergencies. All of these random factors may affect the decision-making behavior of netizens, which may lead to fluctuations or shifts in the final stabilization strategy.

Based on Proposition 3 in Proposition 4, the final evolutionary stability strategies of the random evolutionary game between opinion leaders and netizens can be described by the following four scenarios:

(1) When $DL_1 \cap (DW_1 \cup DW_2)$ is satisfied, there exists a unique evolutionary stable strategy, ESS(0,0), for the system; i.e., the opinion leaders and netizens will ultimately choose the strategy of negative guidance or nonadoption over time. In this case, although the opinion leader chooses to negatively guide and does not provide netizens with true information related to major emergencies, netizens ultimately choose not to adopt the statements posted by the opinion leader. This prevents netizens from being swayed by the negative guidance released by opinion leaders, limits the spread of panic, and reduces the burden on the relevant authorities to maintain the smooth operation of society.

(2) When $DL_1 \cap (DW_3 \cup DW_4)$ is satisfied, there exists a unique evolutionary stable strategy, ESS (0,1), for the system; i.e., the opinion leaders and netizens will ultimately choose the strategy of negative guidance or adoption over time. In this case, as netizens adopt the negative guidance s released by opinion leaders, some of them will reprocess and spread the negative statements, which is very likely to exacerbate the spread of panic. If not channeled in a timely manner, it is very easy for major emergencies to cause secondary impacts and may even evolve into severe network mass incidents.

(3) When $DL_2 \cap (DW_1 \cup DW_2)$ is satisfied, there exists a unique evolutionary stable strategy, ESS (1,0), for the system; i.e., opinion leaders and netizens will ultimately choose the strategy of positive guidance and nonadoption over time. In this case, although the opinion leaders choose to

positively guide the released speech is not adopted by netizens, their decision-making behavior does not cause obstacles to relevant government departments in the process of dealing with major emergencies. It also slows the spread of panic and reduces the burden of relevant departments in maintaining the smooth operation of society from a side perspective.

(4) When $DL_2 \cap (DW_3 \cup DW_4)$ is satisfied, there exists a unique evolutionary stable strategy, ESS (1,1), for the system; i.e., opinion leaders and netizens will eventually choose the strategy of positive guidance and adoption over time. In this case, the opinion leader chooses to guide netizens positively, and netizens adopt the statements posted by the opinion leader. Both of these decision-making behaviors greatly reduce the pressure on relevant departments when dealing with major emergencies and the burden of maintaining smooth operation.

Among the above four scenarios, the strategy most conducive to the society stablity is (positive guidance, adoption). In this scenario, opinion leaders and netizens will cooperate together. Opinion leaders publish positive guidance statements, and netizens choose to believe the truthful information released by opinion leaders. Both parties work together to clarify uncertain information, which can greatly reduce the adverse effects caused by the dissemination of uncertain information. The strategy least conducive to the stable operation of society is (negative guidance, adoption). In this situation, opinion leaders, for some purpose, publish negative guidance statements on social platforms, inciting panic among netizens who believe in negative guidance statements, which will directly exacerbate the adverse effects caused by the dissemination of uncertain information.

4.3 Equilibrium solution

Since Eqs 11, 12 is a nonlinear Itô stochastic differential equation, it is impossible to find its analytical solution directly, so it needs to be solved numerically; in the next part of this paper, we use the stochastic Taylor expansion and Itô's formula to solve Eqs 11, 12 In the next part of this paper, the stochastic Taylor expansion and Itô's formula are applied to numerically solve Eq.

For the following Itô stochastic differential equation:

$$dx(t) = f(t, x(t))dt + g(t, x(t))dw(t) \tag{19}$$

where $t \in [t_0, T], x(t_0) = x_0, x_0 \in R$ and $w(t)$ are standard Brownian motions obeying a normal distribution $N(0, t)$ and $dw(t)$ obeys a normal distribution $N(0, \Delta t)$.

Let the step size $h = \frac{T - t_0}{N}, t_n = t_0 + nh$, where N is the number of samples, t_0 is the initial time point, t_n is the nth sample time point, and $n \in \{0, 1, 2, \ldots, N\}$. The stochastic Taylor expansion of Eq. 19 the stochastic Taylor expansion of Eq:

$$x(t_{n+1}) = x(t_n) + P_0 f(x(t_n)) + P_1 g(x(t_n)) \\ + P_{11} K^1 g(x(t_n)) + P_{00} K^0 f(x(t_n)) + R \tag{20}$$

where R is the remainder term of the expansion, $P_0 = h, P_1 = \Delta w_n, P_{11} = \frac{1}{2}[(\Delta w_n)^2 - h], P_{00} = \frac{h^2}{2}, K^1 = \frac{\partial g(x)}{\partial x}, K^0 = \frac{\partial f(x)}{\partial x} + \frac{\partial^2 g(x)}{2\partial x^2}$.
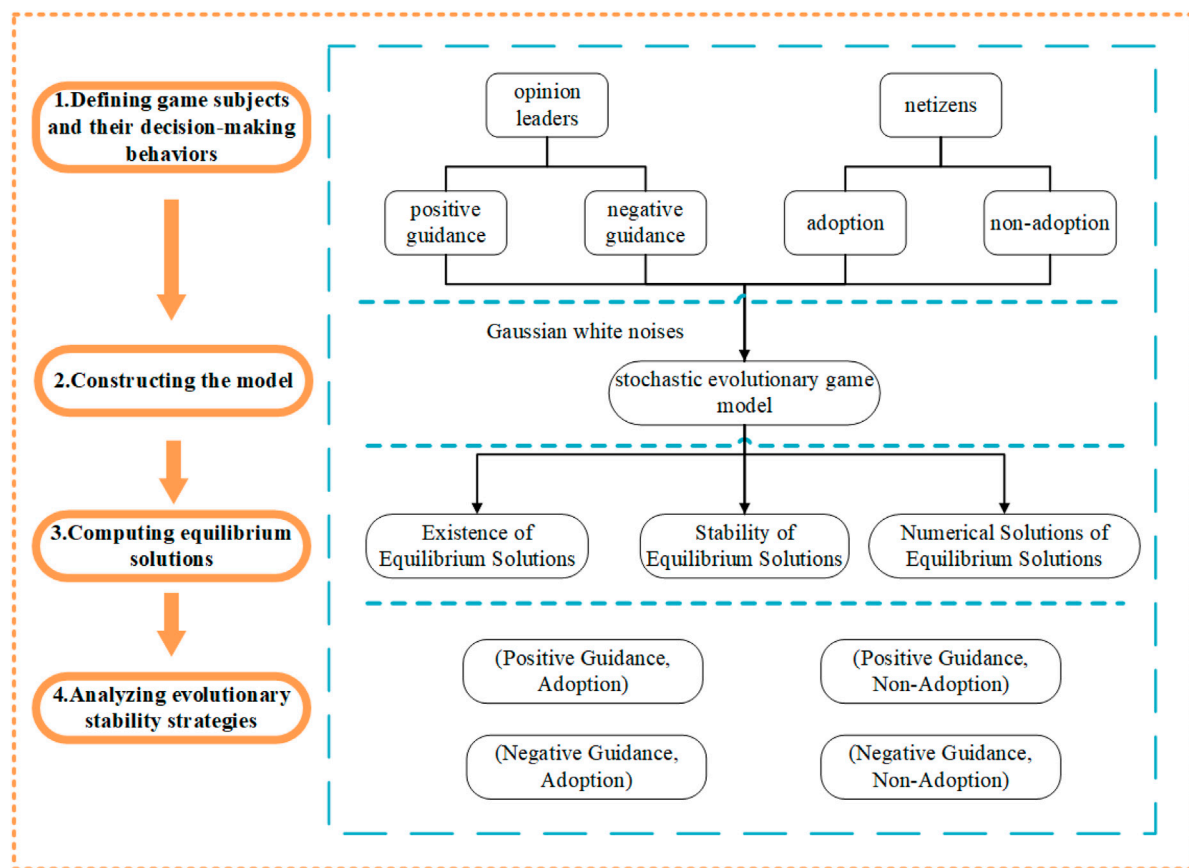
FIGURE 1
Organizational diagram of the current study.

Based on the Formula 20, the random Taylor expansions for Eqs 11, 12 are:

$$x(t_{n+1}) = x(t_n) + hx(t_n)\left[-yI_2 + \pi L_2 - C_1\right] + \Delta w_n \sigma x(t_n)$$
$$+ \frac{1}{2}\left[(\Delta w_n)^2 - h\right]\sigma^2 x(t_n) + \frac{h^2}{2}\left[-yI_2 + \pi L_2 - C_1\right]^2 x(t_n) + R_1$$

(21)

$$y(t_{n+1}) = y(t_n) + hy(t_n)\begin{bmatrix} x(\lambda M_1 + \lambda M_2 + I_3 - I_4) \\ -\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2 \end{bmatrix}$$
$$+ \Delta w_n \sigma y(t_n) + \frac{1}{2}\left[(\Delta w_n)^2 - h\right]\sigma^2 y(t_n)$$
$$+ \frac{h^2}{2}\begin{bmatrix} x(\lambda M_1 + \lambda M_2 + I_3 - I_4) \\ -\lambda M_2 + I_4 - (1-\lambda)C_4 - C_3 + C_2 \end{bmatrix}^2 y(t_n) + R_2$$

(22)

where R_1 and R_2 are the remainder of the expansion of Eqs 21, 22.

In real-world applications, the model can be simulated numerically using Euler's method or Milstein's method, which involves partial term interception of the stochastic Taylor expansion and subsequent numerical solution. In this paper, the numerical solution is based on Milstein's method, taking Eq. 20 For example, the intercept format of the Milstein method is shown in Eq. 23:

$$x(t_{n+1}) = x(t_n) + hf(x(t_n)) + \Delta w_n g(x(t_n))$$
$$+ \frac{1}{2}\left[(\Delta w_n)^2 - h\right]g'(x(t_n))g(x(t_n))$$

(23)

5 Discussion

This paper investigates the decision-making behavior of opinion leaders and netizens on online social media platforms in the process of uncertain information dissemination after a major emergency. A two-party stochastic evolutionary game model of opinion leaders and netizens is constructed considering random interference factors. The evolutionary stability strategy of the model and the stochastic evolution process are solved and analyzed, and the effects of different variables on the decision-making behavior of the two parties are discussed.

To more accurately reflect the influence of the interference strength of random factors, the strength of the regulatory body on the negative guidance of opinion leaders and the psychological identity of opinion leaders toward the decision-making behavioral choices and stochastic evolutionary process of opinion leaders and netizens should be considered. In the following, MATLAB 2017b software is used to numerically simulate the behavioral evolution
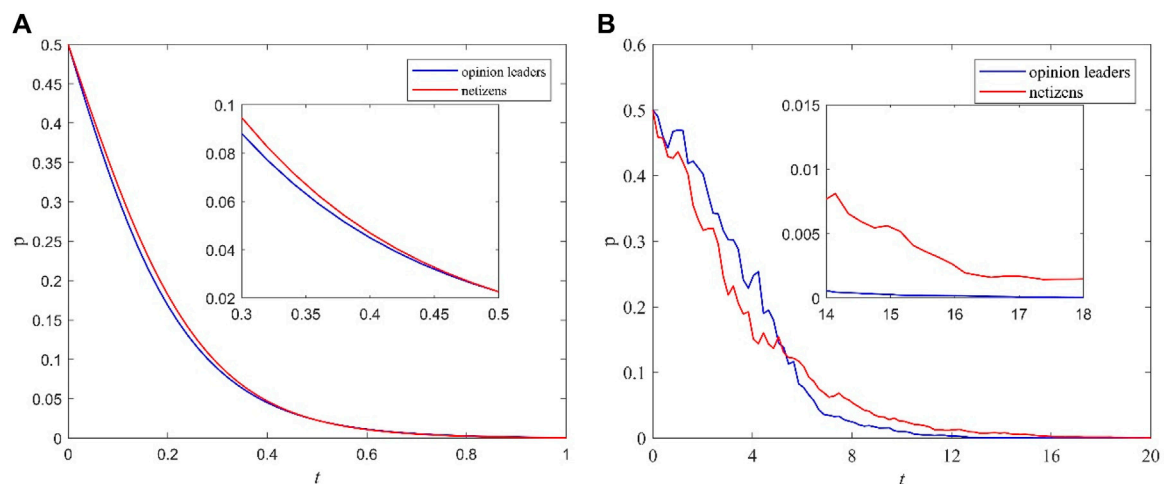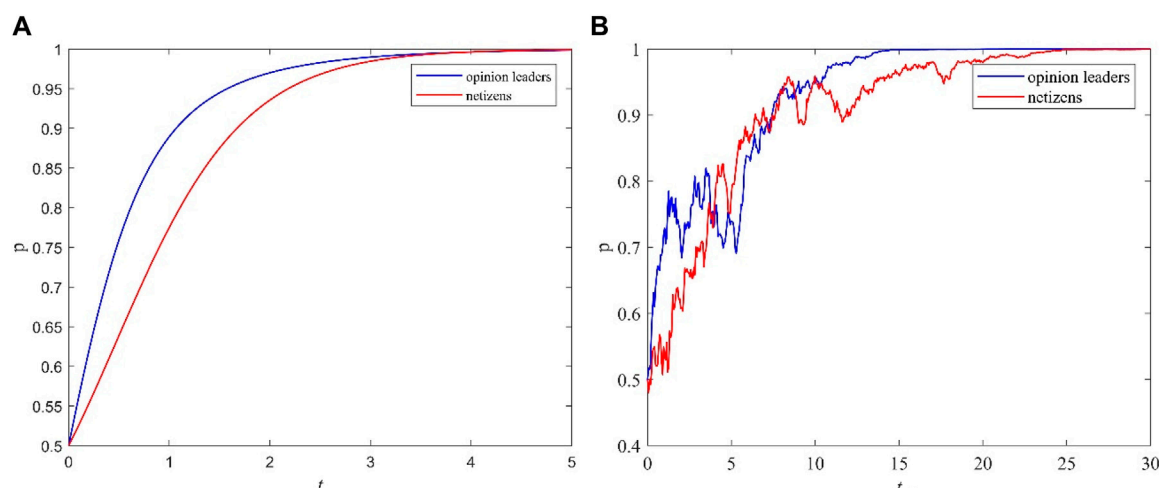
FIGURE 2
Stochastic evolutionary trend of opinion leaders and netizens' choices (negative guidance, nonadoption). **(A)** $\sigma = 0$ **(B)** $\sigma = 1$.

process of the above two game subjects and verify the theoretical analysis above.

5.1 Analysis of the stochastic evolutionary process of opinion leaders and netizens

To better analyze the influence of random interference factors on opinion leaders and netizens in making decision-making behavioral choices, i.e., to analyze the evolutionary stabilization process of the game system after the introduction of Gaussian white noise. The other parameters in the model are randomly assigned, and the evolutionary processes in the absence of random interference factors ($\sigma = 0$) and the presence of random interference factors ($\sigma = 1$) are compared and analyzed.

(1) Analysis of the stochastic evolutionary process of opinion leaders' and netizens' choices (negative guidance, nonadoption).

The relevant parameters of the model are assigned as follows: $I_1 = 10, I_2 = 3, I_3 = 2, I_4 = 4, L_1 = 8, L_2 = 10, C_1 = 9, C_2 = 6, C_3 = 10, C_4 = 10, M_1 = 3, M_2 = 3, \lambda = 0.35$, and $\pi = 0.2$. The initial values of the opinion leaders' choice of the positive steering strategy and the netizens' choice of the adoption strategy are set to x (0) = 0.5 and y (0) = 0.5, respectively, and the simulation step size is h = 0.01. The random evolutionary trend of opinion leaders and netizens can be obtained as shown in Figure 2 below. 0.01. Under the premise of ensuring that the values of the remaining parameters remain unchanged, the stochastic evolution trend graphs of opinion leaders and netizens at $\sigma = 0$ and $\sigma = 1$ can be obtained, as shown in Figure 2 below.

The blue line in Figure 2 represents the probability curve of opinion leaders choosing the positive guidance strategy, and the red line represents the probability curve of netizens choosing the adoption strategy. Figure 2 shows that after a period of evolution, opinion leaders and netizens gradually converge to 0. That is, no matter what the initial values of opinion leaders choosing the positive guidance strategy and netizens choosing the adoption strategy are, opinion leaders and netizens will ultimately choose

the strategy of negative guidance or nonadoption over time, and opinion leaders tend to converge to 0 at a slightly faster rate than netizens. This is because, after major emergencies, opinion leaders can obtain more information from multiple sources and are able to adjust more quickly after gaming behavior has begun. This is because after a major emergency occurs, opinion leaders can obtain more information from multiple sources and make faster adjustments after the game behavior begins. When the benefits of negative guidance strategies are discovered due to positive guidance, opinion leaders can quickly choose negative guidance strategies.

By comparing Figures 2A, B, it can be found that at $\sigma = 0$, opinion leaders and netizens evolve to a stable strategy faster because they are not interfered with by random factors in the decision-making process and need to combine each other's expected benefits to make a strategy choice. Therefore, in an idealized situation without random interference, the two parties are able to quickly make choices that are in their own interest.

(2) Analysis of the stochastic evolutionary process of opinion leaders' and netizens' choices (positive guidance, adoption).

The relevant parameters of the model are assigned as follows: $I_1 = 10, I_2 = 3, I_3 = 2, I_4 = 4, L_1 = 8, L_2 = 20, C_1 = 6, C_2 = 6, C_3 = 6, C_4 = 4, M_1 = 3, M_2 = 3, \lambda = 0.5$, and $\pi = 0.5$. The initial values of the opinion leaders' choice of the positive steering strategy and the netizens' choice of the adoption strategy are set as x (0) = 0.5 and y (0) = 0.5, and the simulation step size is h = 0.01. The random evolutionary trends of opinion leaders and netizens can be obtained as shown in Figure 3. 0.01. Under the premise of ensuring that the values of the remaining parameters remain unchanged, the stochastic evolution trend graphs of opinion leaders and netizens at $\sigma = 0$ and $\sigma = 1$ can be obtained, as shown in Figure 3 below.

The blue line in Figure 3 represents the probability curve of opinion leaders choosing the positive guidance strategy, and the red line represents the probability curve of netizens choosing the adoption strategy. Figure 3 shows that after a period of evolution, opinion leaders and netizens gradually converge to 1. That is, no matter what the initial values are for opinion leaders to choose the positive guidance strategy and for netizens to choose the adoption

FIGURE 3
Stochastic evolutionary trend of opinion leaders' and netizens' choices (positive guidance, adoption). **(A)** $\sigma = 0$ **(B)** $\sigma = 1$.

strategy, opinion leaders and netizens will ultimately choose the strategy of positive guidance adoption over time, and the speed at which opinion leaders converge to 1 is slightly faster than that of netizens.

By comparing Figures 3A, B, it can be observed that the speed at which opinion leaders and netizens evolve to a stable strategy is relatively slow. This is because both are subject to interference from internal or external random factors during the decision-making process, resulting in oscillations during the evolution to a stable state. Therefore, in a nonidealized situation, opinion leaders and netizens will be partially limited by random factors in their decision-making, and they will slow their decision-making speed when thinking about the pros and cons of random factors.

5.2 The effect of psychological identity on stochastic evolutionary processes

To analyze the influence of netizens' psychological identity toward opinion leaders on the stochastic evolution process, other parameters in the model are randomly assigned, and the evolution process is simulated. The influence of changes in psychological identity on the decision-making behavior of opinion leaders and netizens should be observed.

The relevant parameters of the model were assigned as $I_1 = 10$, $I_2 = 3, I_3 = 2, I_4 = 4, L_1 = 8, L_2 = 20, C_1 = 6, C_2 = 6, C_3 = 6, C_4 = 4, M_1 = 3, M_2 = 3, \pi = 0.5$, and $\sigma = 1$. Setting the initial value of the opinion leader's choice of the positive steering strategy, the netizen's choice of the adoption strategy is set to $x(0) = 0.5, y(0) = 0.5$, and the simulation step size h = 0.01. Under the premise of ensuring that the values of the remaining parameters remain unchanged, we change the value of λ and set $\lambda = 0.1, 0.5$, and 0.8. We can construct a stochastic evolution trend graph of opinion leaders and netizens, as shown in Figure 4 below.

The blue line in Figure 4A represents the probability curve of opinion leaders choosing the positive guidance strategy when the psychological identity $\lambda = 0.1$, the red line represents the probability

curve when the psychological identity $\lambda = 0.5$, and the green line represents the probability curve when the psychological identity $\lambda = 0.8$. The three lines in Figure 4B are the probability curves of netizens choosing the adoption strategy. Figure 4A shows that the netizens' psychological identity λ of the opinion leader does not affect the final strategy choice of the opinion leader; however, the smaller the value of psychological identity λ is, the faster the opinion leader equals 1. This is because, in reality, the less value a netizen places on the opinions of the opinion leader, the less likely it is that the opinion leader's published speech will be adopted by the netizens. To increase their influence among netizens, opinion leaders need to make strategic choices as early as possible to attract netizens' attention. The more content an opinion leader publishes, the more netizens will pay attention to it, and the more likely it is to increase netizens' psychological identification with it.

Figure 4B shows that as netizens' recognition of opinion leaders gradually increases, they gradually convert from converging to 0 to converging to 1; that is, netizens convert from not adopting the remarks released by opinion leaders to adopting the remarks released by opinion leaders. This is because the more netizens trust opinion leaders, the easier it is for them to believe their statements. However, the psychological recognition of opinion leaders by netizens is a double-edged sword. Due to their limited understanding, it is difficult to accurately distinguish whether the statements made by opinion leaders are true. This leads to the fact that when opinion leaders provide positive guidance, they can largely maintain the stable operation of society, but when they provide negative guidance, it is easy to exacerbate the spread of panic and greatly increase the difficulty of government work.

5.3 Effect of penalty strength on stochastic evolutionary processes

To analyze the influence of the government's punishment strength on those who propagate and spread negative information on the stochastic evolution process, the other
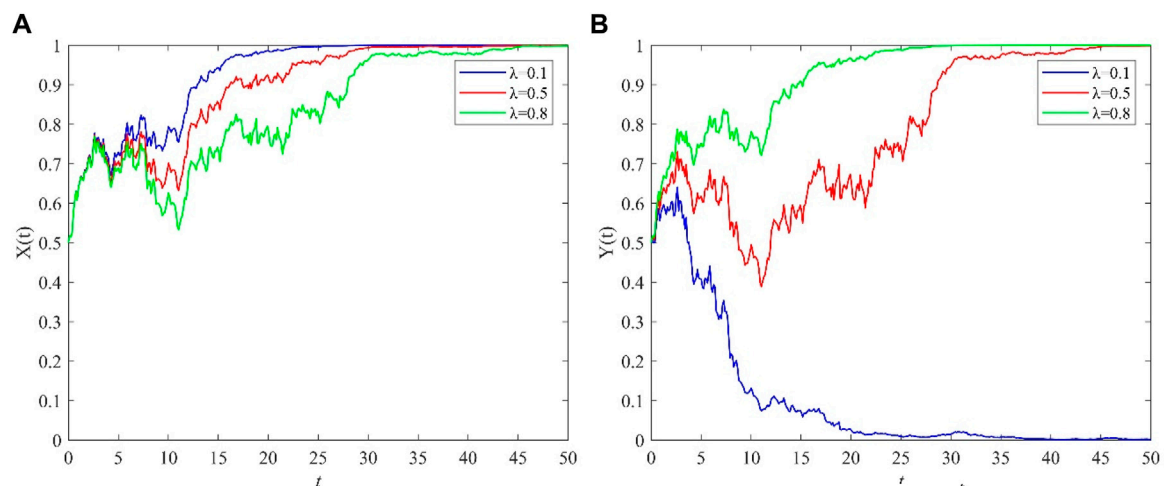
FIGURE 4
The effect of psychological identity λ on the evolutionary strategies of opinion leaders and netizens. **(A)** opinion leaders **(B)** netizens.
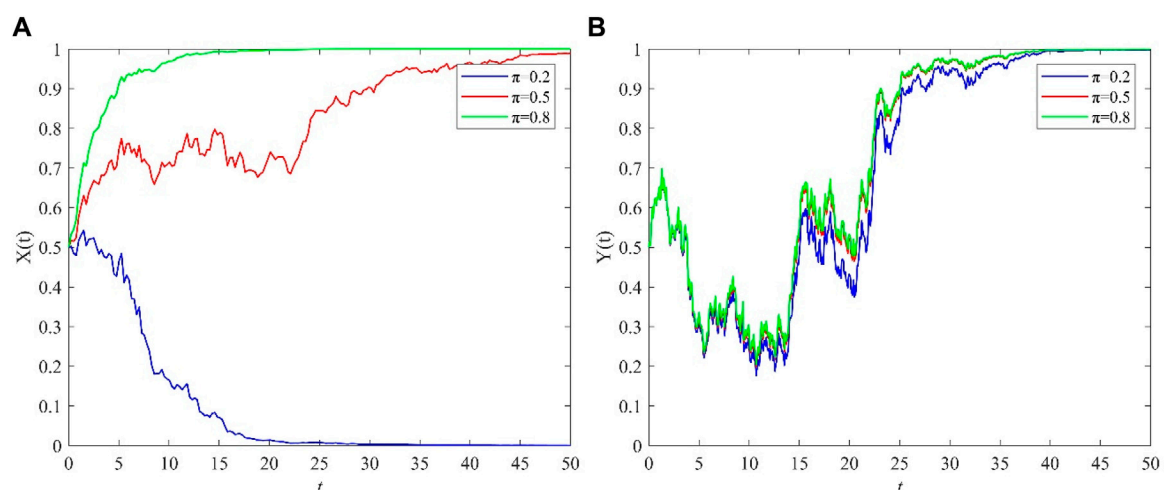
FIGURE 5
Impact of punishment probability π on the evolutionary strategies of opinion leaders and netizens. **(A)** opinion leaders **(B)** netizens.

parameters in the model are randomly assigned values, and the evolution process is simulated. The influence of changes in punishment strength on the decision-making behavior of opinion leaders and netizens should be observed.

Since both π and L_2 in the model can affect the punishment strength, for the convenience of analysis, only the government's punishment probability π for those who propagate and spread negative information is selected for analysis. The relevant parameters of the model are assigned as follows: $I_1 = 10$, $I_2 = 3$, $I_3 = 2$, $I_4 = 4$, $L_1 = 8$, $L_2 = 20$, $C_1 = 6$, $C_2 = 6$, $C_3 = 6$, $C_4 = 4$, $M_1 = 3$, $M_2 = 3$, $\lambda = 0.35$, and $\sigma = 1$. The initial values of the opinion leaders' choice of the positive guidance strategy and the netizens' choice of the adoption strategy are set as $x(0) = 0.5$ and $y(0) = 0.5$, respectively, and the simulation step size is set as h = 0.01. Under the premise of ensuring that the values of the remaining parameters remain unchanged, changing the value of π and setting

$\pi = 0.2, 0.5,$ and 0.8, we can obtain the stochastic evolution trend graph of opinion leaders and netizens, which is shown in Figure 5 below.

The blue line in Figure 5A represents the probability curve of opinion leaders choosing the positive guidance strategy when the penalty probability *is* $\pi = 0.2$, the red line represents the probability curve when the penalty probability *is* $\pi = 0.5$, and the green line represents the probability curve when the penalty probability *is* $\pi = 0.8$. The three lines in Figure 5B are the probability curves of netizens choosing the adoption strategy. From Figure 5A, it can be seen that when the government's punishment for propagandizing and spreading negative information gradually increases, the opinion leader gradually changes from 0 to 1, i.e., the opinion leader converts from negative guidance to positive guidance. This is because when the government's punishment gradually exceeds the range that opinion leaders can bear, to avoid being punished,

such as through banning, blocking or administrative punishment, opinion leaders will switch from negative guidance to positive guidance. Therefore, when there are major emergencies, the government should increase punishment for those who publicize and disseminate negative information to encourage opinion leaders to choose a positive guidance strategy.

Figure 5B shows that the government's punishment strength for those who publicize and spread negative information does not affect netizens' final strategy choice; however, the stronger the punishment is, the faster the netizens tend to 1. This is because, in netizens' thinking, the more the government punishes opinion leaders for spreading negative statements, the more correct the statements spread by the possible opinion leaders will be; therefore, the netizens will adopt the strategy more quickly.

5.4 Limitations and future prospects

This study has certain limitations: 1) this paper assumes that the ability of opinion leaders to disseminate statements is the same and does not distinguish the ability of opinion leaders to disseminate statements, but in reality, the statements released by different opinion leaders are all different; 2) the vague relationship between opinion leaders and netizens has not been clearly distinguished; In the real world, opinion leaders can be natural persons or institutional accounts; 3)This paper uses MATLAB for numerical simulation of the model without combining real-world data.

In future work, we will conduct separate research on opinion leaders with different dissemination capabilities based on different network structures [33, 34] by combining complex networks with stochastic evolutionary game models. Furthermore, we will refine the differences between opinion leaders and netizens, incorporate more detailed decision-making behaviors, and include more accurate assumptions in the construction of the model. Additionally, we will try to combine real-world data and use real cases to simulate the model.

6 Conclusion

This paper takes the decision-making behavior of opinion leaders and netizens in the process of uncertain information dissemination on online social media platforms after major emergencies as the research background. In the real world, opinion leaders and netizens are subject to random interference from their own internal factors or external environmental factors in the decision-making process. Based on the traditional evolutionary game model, Gaussian white noise is introduced to construct a stochastic evolutionary game model based on uncertain information dissemination behavior between opinion leaders and netizens. Using the theory of Itô stochastic differential equations and stochastic Taylor expansion, the evolutionary stability strategy and stochastic evolution process of the model are analyzed, and the numerical solution of the equilibrium solution of the model is found. Finally, through numerical simulation software, we analyze the influence of the interference intensity of random factors, the punishment strength of regulatory agencies when opinion leaders spread false

or negative information, and the opinions of netizens toward opinion leaders on the stochastic evolution process. The results of this paper can be summarized as follows:

(1) Random disturbances will not change the strategic choices of opinion leaders or netizens during the process of stochastic evolution to a stable state. However, due to the emergence of random interference, opinion leaders and netizens will think about the advantages and disadvantages of random factors in the process of decision-making, which leads to the slowing of their progression to a stable state and the occurrence of oscillations in the process of evolution. After a major emergency occurs, opinion leaders can obtain more information from multiple sources and make faster adjustments after the game behavior begins. The speed at which opinion leaders reach a stable state is faster than the speed at which netizens reach a stable state.

(2) A change in the netizens' psychological identity λ will not affect the final strategy choice of opinion leaders, but it will affect the speed at which opinion leaders reach a stable state; the smaller the value of λ is, the faster the opinion leaders reach a stable state. A change in psychological identity λ has a greater impact on netizens and even changes their final strategy choice. Under the premise that the other variables remain unchanged, when the value of λ gradually increases, the netizens' final stable state will be converted from nonadoption to adoption of the speech released by opinion leaders.

(3) Changes in the government's punishment of those who publicize and disseminate negative information will not affect netizens' final strategic choices but will affect the speed at which netizens reach a stable state; the greater the punishment is, the more quickly netizens reach a stable state. A change in punishment intensity has a greater impact on opinion leaders and even changes their final strategy choice. Under the premise that other variables remain unchanged, when the punishment intensity gradually increases, the final stable state of opinion leaders will be converted from negative guidance to positive guidance.

Therefore, when major emergencies occur again, government regulators can encourage opinion leaders to make positive decisions by increasing punishment for those who publicize and disseminate negative information, accelerating the investigation of major emergencies and releasing the real situation in a timely manner, and improving the channels for opinion leaders to obtain information. Opinion leaders can attract the attention of netizens by increasing their psychological recognition and improving and enriching the content of the real information they publish to encourage netizens to adopt positive statements.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary material, further inquiries can be directed to the corresponding author.

Author contributions

Funding

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Smith JM, Price GR. The logic of animal conflict. *Nature* (1973) 246(5427):15–8. doi:10.1038/246015a0

2. WöLFL B, Te Rietmole H, Salvioli M, Kaznatcheev A, Thuijsman F, Brown JS, et al. The contribution of evolutionary game theory to understanding and treating cancer. *Dynamic Games Appl* (2022) 12(2):313–42. doi:10.1007/s13235-021-00397-w

3. Li XY, Li QZ, Du YJ, Fan Y, Chen X, Shen F, et al. A novel tripartite evolutionary game model for misinformation propagation in social networks. *Security Commun Networks* (2022) 2022:1–13. doi:10.1155/2022/1136144

4. Wang H, Liu W, Liu AF, Wang T, Song H, Zhang S. SQCS: a sustainable quality control system for spatial crowdsourcing via three-party evolutionary game: theory and practice. *Expert Syst Appl* (2024) 238:122132. doi:10.1016/j.eswa.2023.122132

5. Wu ZQ, Yang C, Zheng RJ. An analytical model for enterprise energy behaviors considering carbon trading based on evolutionary game. *J Clean Prod* (2024) 434: 139840. doi:10.1016/j.jclepro.2023.139840

6. Shi QN, Yang SM, Wang N, Zhang SE, Wang Y, Wu B, et al. An evolutionary game-based simulation study of a multi-agent governance system for smart senior care services in China. *Bmc Geriatr* (2023) 23(1):871. doi:10.1186/s12877-023-04521-w

7. Sun B, Xu ZH, Wei M. Evolutionary game model of civil aviation and high-speed rail interaction strategies based on the passenger ticket and carbon trading prices. *J Adv Transportation* (2023) 2023:1–16. doi:10.1155/2023/7675900

8. Shi WQ, Hu QD, Zhou YM. Evolutionary game analysis of vehicle procurement in the courier industry from the perspective of green supply chain. *Int J Ind Eng Computations* (2024) 15(1):223–34. doi:10.5267/j.ijiec.2023.10.002

9. Perc M. Coherence resonance in a spatial prisoner's dilemma game. *New J Phsyics* (2006) 8(3):22. doi:10.1088/1367-2630/8/2/022

10. Tanimoto J. Promotion of cooperation by payoff noise in a 2x2 game. Physical review E, Statistical, nonlinear. *soft matter Phys* (2007) 76(4 Pt 1):041130. doi:10.1103/physreve.76.041130

11. Mo D, Chen X, Zhu Z, Liu C, Xie N. A stochastic evolutionary dynamic game model for analyzing the ride-sourcing market with limited platform reputation. *Transportmetrica B-Transport Dyn* (2023) 11(1). doi:10.1080/21680566.2023.2248399

12. Kang K, Bai L, Zhang J. A tripartite stochastic evolutionary game model of complex technological products in a transnational supply chain. *Comput Ind Eng* (2023) 186:109690. doi:10.1016/j.cie.2023.109690

13. Xie JD, Guan BW, Yao Y, Li R, Shi Q. Market power risk prevention mechanism of China's electricity spot market based on stochastic evolutionary game dynamics. *Front Energ Res* (2023) 11. doi:10.3389/fenrg.2023.1270681

14. Du J, Li JJ, Li JX, Li W. Competition-cooperation mechanism of online supply chain finance based on a stochastic evolutionary game. *Oper Res* (2023) 23(3):55. doi:10.1007/s12351-023-00792-8

15. Scata M, Di Stefano A, La Corte A, Liò P, Catania E, Guardo E, et al. Combining evolutionary game theory and network theory to analyze human cooperation patterns. *Chaos Solitons and Fractals* (2016) 91:17–24. doi:10.1016/j.chaos.2016.04.018

16. Mo TC, Xie C, Li KL, Ouyang Y, Zeng Z. Transmission effect of extreme risks in China?s financial sectors at major emergencies: empirical study based on the GPD-CAViaR and TVP-SV- VAR approach. *Electron Res Archive* (2022) 30(12):4657–73. doi:10.3934/era.2022236

17. Cheng A, Chen TH, Jiang GG, Han X. Can major public health emergencies affect changes in international oil prices? *Int J Environ Res Public Health* (2021) 18(24):12955. doi:10.3390/ijerph182412955

18. Jia FJ, Wang DD, Li LS. The stochastic evolutionary game analysis of public prevention and control strategies in public health emergencies. *Kybernetes* (2023) 52(6): 2205–24. doi:10.1108/k-10-2021-0988

19. Özkaya M, Izgi B. Effects of the quarantine on the individuals' risk of Covid-19 infection: game theoretical approach. *Alexandria Eng J* (2021) 60(4):4157–65. doi:10.1016/j.aej.2021.02.021

20. Salarpour M, Nagurney A. A multicountry, multicommodity stochastic game theory network model of competition for medical supplies inspired by the Covid-19 pandemic. *Int J Prod Econ* (2021) 236:108074. doi:10.1016/j.ijpe.2021.108074

21. Wang HC, Ma XY. Research on multiobjective location of urban emergency logistics under major emergencies. *Math Probl Eng* (2021) 2021:1–12. doi:10.1155/2021/5577797

22. Wei Y. Network public opinion propagation control model of major emergencies based on heat conduction theory. *Wireless Commun Mobile Comput* (2022) 2022:1–14. doi:10.1155/2022/1476231

23. Chen TG, Yin XH, Yang JJ, Cong G, Li G. Modeling multi-dimensional public opinion process based on complex network dynamics model in the context of derived topics. *Axioms* (2021) 10(4):270. doi:10.3390/axioms10040270

24. Lu P, Chen DH, Zhang G, Ding J. Online attention dynamics: the triangle framework of theory, big data and simulations. *Expert Syst Appl* (2023) 233:120900. doi:10.1016/j.eswa.2023.120900

25. Apuke OD, Omar B. Fake news and COVID-19: modelling the predictors of fake news sharing among social media users. *Telematics Inform* (2021) 56:101475. doi:10.1016/j.tele.2020.101475

26. Chew C, Eysenbach G. Pandemics in the age of twitter: content analysis of tweets during the 2009 H1N1 outbreak. *Plos One* (2010) 5(11):e14118. doi:10.1371/journal.pone.0014118

27. Yu LA, Li L, Tang L, Dai W, Hanachi C. A multi-agent-based online opinion dissemination model for China's crisis information release policy during hazardous chemical leakage emergencies into rivers. *Online Inf Rev* (2017) 41(4):537–57. doi:10.1108/oir-04-2015-0126

28. Zhang L, Wang X, Wang J, Yang P, Zhou P, Liao G. A study on predicting crisis information dissemination in epidemic-level public health events. *J Saf Sci Resilience* (2023) 4(3):253–61. doi:10.1016/j.jnlssr.2023.02.003

29. Guan WQ, Gao HY, Yang MM, Li Y, Ma H, Qian W, et al. Analyzing user behavior of the micro-blogging website Sina Weibo during hot social events. *Physica a-Statistical Mech Its Appl* (2014) 395:340–51. doi:10.1016/j.physa.2013.09.059

30. Alvarez-Galvez J. Network models of minority opinion spreading: using agent-based modeling to study possible scenarios of social contagion. *Soc Sci Comp Rev* (2016) 34(5):567–81. doi:10.1177/0894439315605607

31. Parsegov SE, Proskurnikov AV, Tempo R, Friedkin NE. Novel multidimensional models of opinion dynamics in social networks. *Ieee Trans Automatic Control* (2017) 62(5):2270–85. doi:10.1109/tac.2016.2613905

32. Li B, Li H, Sun Q, Lv R, Zhao J. Evolutionary game analysis of the dissemination of false information by multiple parties after major emergencies. *Complexity* (2022) 2022: 1–14. doi:10.1155/2022/3527674

33. Li H-J, Feng Y, Xia C, Cao J. Overlapping graph clustering in attributed networks via generalized cluster potential game. *ACM Trans Knowledge Discov Data* (2023) 18(1): 1–26. doi:10.1145/3597436

34. Li H, Cao H, Feng Y, Li X, Pei J. Optimization of graph clustering inspired by dynamic belief systems. *IEEE Trans Knowledge* (2023)(01) 1–14. doi:10.1109/tkde. 2023.3274547

![frontiers logo] Frontiers in Physics

Game-theory based truck platoon avoidance modes selection near the highway off-ramp in mixed traffic environment

Yi Li[1]*, Lan Wang[1], Zhaoze Xuan[1] and Wenzhe Shen[2]

[1]Logistics Research Center, Shanghai Maritime University, Shanghai, China, [2]Lianyungang New Harbour Terminal Co., Ltd., Lianyungang, Jiangsu, China

Introduction: The rise of autonomous vehicles has brought about a transformative shift in transportation, witnessing the coexistence of human-driven and autonomous vehicles on highways in the United States, Europe, and China. This coexistence poses challenges to traffic operations, particularly in intricate scenarios like highway ramps. The interaction between autonomous truck platoons, displaying heightened maneuverability, and human-driven vehicles has emerged as a critical concern. Consequently, this research aims to propose and investigate three avoidance modes (overall, gap and cross) employed by truck platoons, evaluating their comprehensive impact on human-driven vehicles.

Methods: Multiple scenarios are simulated utilizing the Simulation of Urban Mobility (SUMO) software, collecting data on three distinctive avoidance modes concerning Travel Time (TT) and Time to Collision (TTC). Employing principles of game theory, a comprehensive assessment is undertaken to evaluate the traffic efficiency and safety of each mode. Comparative analyses against a no-avoidance baseline are conducted, offering a holistic evaluation of each mode's applicability across diverse scenarios.

Results: The findings highlight the commendable performance of gap mode and overall mode in enhancing traffic efficiency, while cross mode excels in fortifying traffic safety. Overall, the gap mode emerges as the optimal choice among the three.

Discussion: This study introduces a game-theoretic approach to managing human-machine mixed traffic flow, establishing a foundational framework for theoretical research in decision-making for emerging mixed traffic environments. It considers safety and efficiency perspectives across different types of traffic entities. The insights gained contribute to the evolving discourse on the integration of autonomous vehicles into existing traffic systems, addressing the intricate challenges posed by the coexistence of various vehicle types on highways.

KEYWORDS

truck platoon, avoidance mode, mixed traffic flow, off-ramp area, game theory

1 Introduction

With the development of logistics industry and the intensification of market competition, truck transportation plays a crucial role in cargo transportation. However, truck transportation also faces many challenges, such as traffic congestion, environmental pollution, and energy consumption. To address these issues, researchers have begun to

consider using autonomous truck platoons to reduce transportation costs, improve transportation efficiency, and reduce environmental impacts. In practical scenarios, an excessively long truck platoon needs to avoid other cars in specific areas, so how to choose an avoidance mode is currently one of the research topics. Game theory-based methods can provide effective decision support for autonomous truck platoon system, which helps to choose the best avoidance mode.

2 Literature review

The emergence of autonomous driving technology has given rise to various new technology, equipment, and theoretical approaches related to autonomous driving. Based on such development, truck platoon as a new freight transportation mode has demonstrated significant advantages in terms of reducing energy consumption [1, 2], operating costs [3, 4], and improving road capacity [5, 6]. Truck platoons are two or more trucks in a fleet that are connected using vehicle-to-vehicle communication and autonomous driving support systems. When these trucks are connected to each other on certain sections (e.g., on a highway), they automatically maintain a close-headway platoon [7]. This means that these trucks must be equipped with Level 1 or Level 2 and even higher autonomous driving systems [8]. However, the current state of autonomous driving technology is insufficient to support full truck platoon driving in the entire road network. Instead, truck platoon is applicable only in specific road scenarios. Moreover, due to the spatial characteristics of truck platoons, they are more suitable in the wider road scenarios [9]. Thus, it is crucial to explore the safety aspects and assess the impact on the traffic environment when implementing truck platoons in specific scenarios.

Previous research has investigated various aspects of truck platooning in specific road environments. Zhao focused on optimizing overall traffic efficiency by examining the appropriate length of truck platoons in off-ramp regions [10]. Chandra explored how different highway routes affect the accessibility of truck platoons [11]. Faber studied the impact of different truck platoon characteristics on traffic safety and efficiency in the context of highway merging scenarios [12]. It can be seen that the highways are recognized as suitable environments for truck platoon driving, particularly during merging and diverging scenarios where human-driven vehicles (HVs) frequently interact with truck platoons. Scholars have conducted studies on truck platooning during highway merging situations [13–15]. However, fewer studies have been conducted for the diversion scenario. Especially, how can truck platoons meet the needs of HVs on the adjacent lanes waiting for getting off the highway when passing through the highway diversion area. Tabibi suggested that high traffic volumes make it challenging for the vehicles to find gaps to cross the truck platoon [16].

Recent research has focused on the avoidance strategies of the connected autonomous truck platoons. For example, Wu applied the intra-platoon field theory in the feasible gap model [17]. Shi proposed a cooperative connected autonomous vehicle (CAV) lane-changing model to build more CAV platoon in mixed traffic conditions [18]. Wang designed a collaborative lane change control algorithm based on model predictive control algorithm, which could reduce the influence of lane change behavior on traffic flow [19].

The above studies showed that CAV platoon could improve the lane capacity and decrease the average time headway in a great deal [20, 21]. However, in a mixed traffic flow scenario (HV + CAV), the benefits of platooning can be influenced by various factors, including traffic signals, highway ramps, vehicle types, and human errors. Few of these factors were considered or discussed in past studies. Therefore, further studies are needed to explore these issues.

This study focuses on the highway off-ramp scenario to analyze the impact of truck platoon with different avoidance modes on the traffic flow when surrounded with human-driven vehicles. The objective of this study is to provide an effective method to assess the impact of different truck platoon avoidance modes on traffic flow, which helps to improve traffic efficiency and safety. It is assumed that all the vehicles in the simulation experiment are in the state of vehicle networking. The lane-changing information of human-driven vehicles can be obtained directly by the truck platoons. Besides, the truck platoons travel on the right lane that is compliant with the traffic rules in China. Simulation experiments based on SUMO were conducted in this study. Meanwhile, collision risk factor and traffic efficiency are also quantified to evaluate the applicability of each avoidance mode.

The rest of the study is organized as follows. Section 2 analyzes the lane-changing scenarios of HVs near the off-ramp areas and proposes three truck platoon avoidance modes. Section 3 describes the design of the simulation experiments. Section 4 proposes to evaluate and analyze the performance of each avoidance mode. Section 5 discusses the impact mechanisms of different avoidance modes. Section 6 summarizes the whole work and its contributions and limitations.

3 Truck platoon avoidance mode

The technical difficulty and complexity of the truck platoon model is higher than that of the car platoon model. The flexibility of car platoons is stronger, and the control strategies can be applied on all lanes. However, truck platoons can only operate on designated lanes and have lower traffic rights than cars or car platoons on special road sections. Therefore, the avoidance of truck platoons is one of its characteristics, on the other hand its driving mode is also constrained. Therefore, in this study, a two-lane highway off-ramp scenario was designed, with the assumption that the truck platoon travels on the second lane (the right one). As the truck platoon approached the off-ramp, evasive maneuvers became necessary to satisfy the lane change requirements of HVs, as shown in Figure 1.

Based on the above scenarios, three avoidance modes of the truck platoon are proposed as follows.

(1) Overall mode: All trucks in the platoon will change lanes to allow the green HVs to access the off-ramp. Once the HVs have entered the ramp, the platoon will transition back to the right lane.

(2) Gap mode: The platoon will disperse to create additional space for the green HVs to perform their lane changes. After the HVs have completed the off-ramp maneuver, the sub-platoons will merge to form the original long platoon configuration.
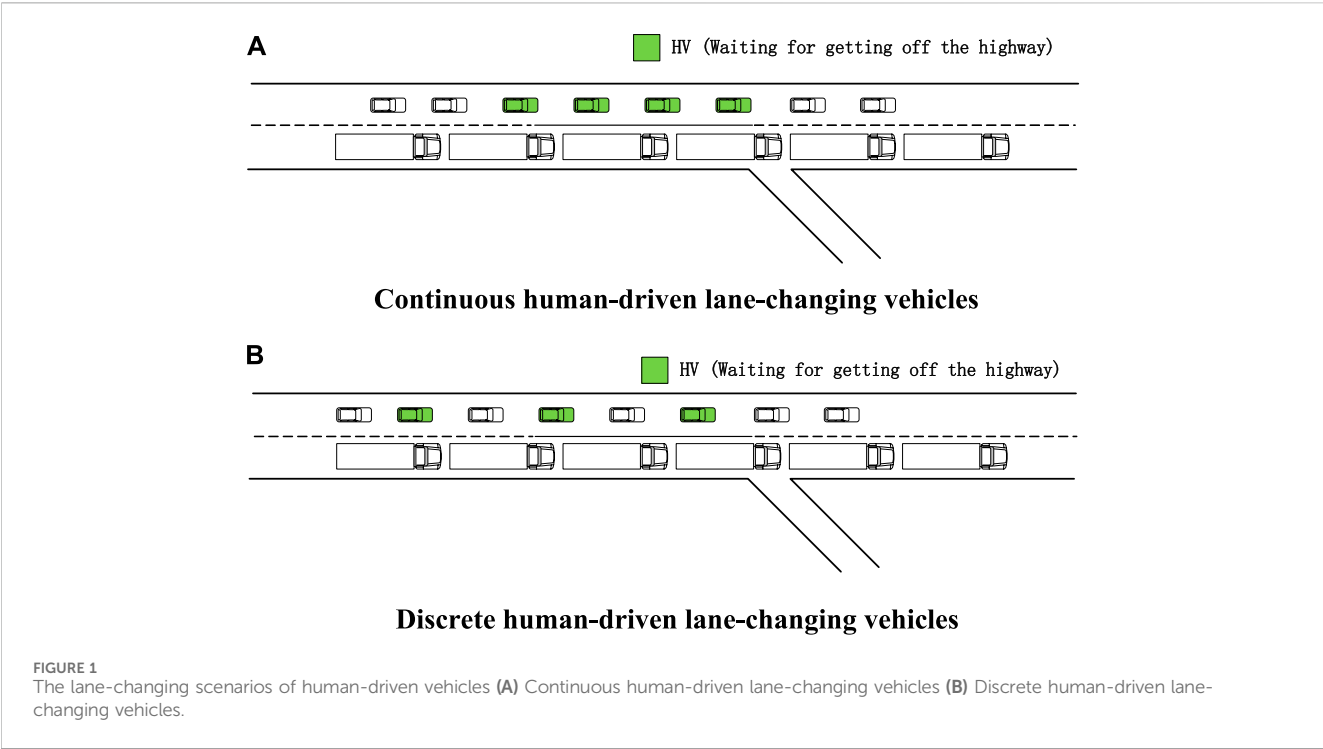
FIGURE 1
The lane-changing scenarios of human-driven vehicles **(A)** Continuous human-driven lane-changing vehicles **(B)** Discrete human-driven lane-changing vehicles.

TABLE 1 Avoidance modes.

Modes	Illustration	
	Continuous off-ramp scenario	Discrete off-ramp scenario
Overall Mode		
Gap Mode		
Cross Mode		

(3) Cross mode: In this mode, the platoon will dissolve into three sub-platoons. The lead platoon continues straight ahead. The center platoon will change lanes for the green HVs. The rear platoon will continue straight ahead and keep a safe distance from the last green HVs.

In the overall mode, the whole truck platoon will change lanes and return to the original lane after the HV vehicles goes off ramp. In the gap mode, the truck platoon stays on the same lane, then the platoon will split into two sub-platoons to provide additional space for off-ramp vehicles, and will re-merge after the HV vehicles go off ramp. In the cross mode, the truck platoon will split into three sub-platoons, then the

middle sub-platoon changes the lane to give way to the HV vehicles, and will return to the original lane after the HV vehicles go off ramp.

The illustration of the three modes is shown in Table 1.

4 Simulation experiment

4.1 Simulation scenario and settings

In this study, a highway simulation scenario was constructed based on SUMO. This highway included two lanes and an off-ramp area, as shown in Figure 2.

FIGURE 2
SUMO-based simulation scenarios.

FIGURE 3
The distribution of lane change locations near the off-ramp.
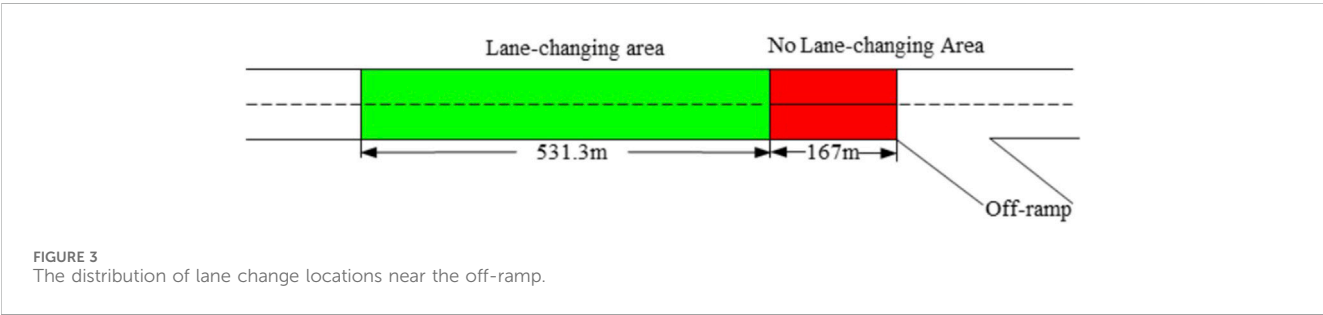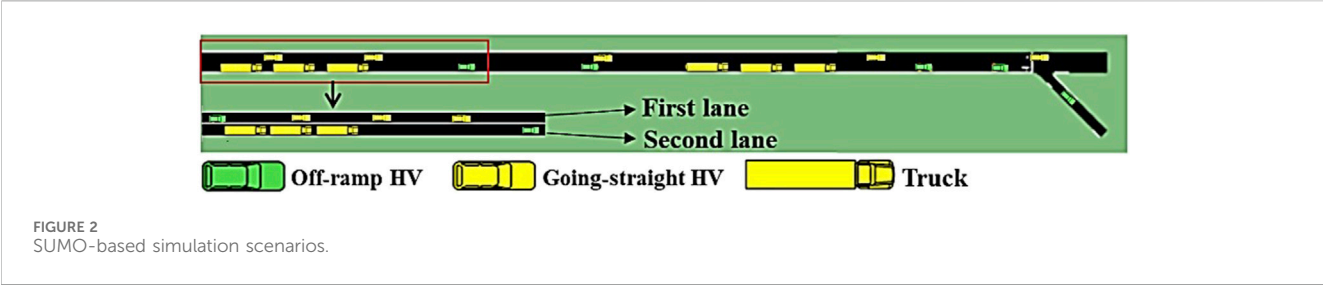
TABLE 2 Key parameters of the simulation experiment.

No.	Parameter	Value
1	Length of scenario	2000 m
2	Length of HV	5 m
3	Length of truck	15 m
4	Initial speed of HV	100 km/h
5	Initial speed of truck	80 km/h
6	Maximum acceleration	2.6 m/s²
7	Maximum decelerate	−4.5 m/s²
8	Initial lane of HV	Random
9	Initial lane of truck platoon	Second lane
10	Length of lane-changing area	531.3 m
11	Time headway between trucks	0.1 s

TABLE 3 Three kinds of traffic flow volume.

No.	Flow volume type	Range
1	Low flow	1800–2,300 pcu/h
2	Medium flow	2,400–2,900 pcu/h
3	High flow	3,000–3,500 pcu/h

the TraCI module, parameters such as the platoon status and vehicle trajectory in the platoon could be obtained in real time. At the same time, the vehicle status could be modified and controlled in the simulation scenario. The key parameters mentioned above are shown in Table 2.

In this study, the traffic flow volume of HVs around the truck platoon was regarded as an important factor in the mode choice [22]. Therefore, five ratio levels (K) of the off-ramp vehicles were simulated in SUMO, including $K = 10\%$, $K = 20\%$, $K = 30\%$, $K = 40\%$, $K = 50\%$. K can be calculated as Eq. 1.

$$K = \frac{ORV_s}{GV_s + ORV_s} \tag{1}$$

where, ORVs is the number of off-ramp vehicles in the scenario. GVs is the number of going-straight vehicles in the scenario.

Each mode was simulated in the three kinds of traffic flow volume, as shown in Table 3. The three kinds of traffic flow volume is based on the service level 1 to 4 in "Technical Standard of Highway Engineering (JTG B01-2014)" of China.

4.2 Car-following models

In this study, the truck platoon was assumed to be equipped with V2V communication and automatic cruise control (ACC) system.

In our study, 40 drivers participated in simulation experiments in Silab simulator. The lane-changing position of each driver was recorded to obtain the lane change area. The lane-changing positions of HVs on the upstream of highway off-ramps were recorded. Based on this, this study determined the HVs' lane-changing area for HVs upstream of the highway off-ramps as shown in Figure 3.

The speed limit of all vehicles was 120 km/h. The HVs entered the simulation lane randomly. The off-ramp vehicles must complete the lane-changing behavior before entering the no lane-changing area. In the continuous HVs scene, the four continuous off-ramp vehicles kept driving with a safe time of headway like other HVs until they reach the lane-changing area. Then they could change lane freely.

The trucks could keep driving by platoon and disband to yield by the TraCI module (Traffic Control Interface) in SUMO. Through

Therefore, different following models should be adopted for the truck platoon and HVs.

4.2.1 Car-following model
4.2.1.1 Human-driven vehicle

IDM, as a basic model to study the dynamic changes of traffic, is mainly modeled based on speed, inter-vehicle distance and the speed difference between the rear vehicle and the front vehicle [23, 24]. Therefore, this study uses IDM as the car-following model for human-driven vehicles in the simulation scenarios.

The acceleration function in IDM is shown as Eqs 2, 3

$$\dot{v} = a\left[1 - \left(\frac{v}{v_0}\right)^\delta - \left(\frac{S'(v, \Delta v)}{S}\right)^2\right] \quad (2)$$

$$s'(v, \Delta v) = s_0 + Tv + \frac{v\Delta v}{2\sqrt{ab}} \quad (3)$$

where, $\dot{v}$ is the vehicle acceleration, a is the maximum acceleration, v is the actual speed, v_0 is the desired speed, δ is the free acceleration component, Δv is the velocity difference from the leading vehicle, S is the distance difference from the leading vehicle, S' is desired following distance, s_0 is jam distance, T is safe-time headway, b is the desired deceleration rate.

4.2.1.2 Truck platoon

The CACC car-following model based previous studies is used to describe the driving behavior between trucks [25–27]. Corresponding equations are listed in Eqs 4–6.

$$v_{i,k} = v_{i,k-1} + k_p e_{i,k-1} + k_d \frac{d(e_{i,k-1} - e_{i,k-2})}{dt} \quad (4)$$

$$e_{i,k} = p_{i-1,k-1} - p_{i,k-1} - L - t_{des}v_{i,k-1} - d_0 \quad (5)$$

$$d_0 = \begin{cases} 0 & v_{i,k} \geq 10m/s \\ -0.125v & v_{i,k} \leq 10m/s \end{cases} \quad (6)$$

where, $v_{i,k}$ is the speed of vehicle i at time k, $e_{i,k}$ is the difference between actual space headway and desired space headway of vehicle i at time k, $p_{i,k}$ is the position of vehicle i at time k, L is the length of the vehicle, t_{des} is the desired headway, d_0 is the spacing margin, k_p and k_d are the control gains.

4.2.2 Lane-changing model

In the SUMO, the vehicle needs to determine whether the gap on the target lane satisfies the minimum safe gap before executing the lane-changing command [28]. If it is satisfied, the lane-changing behavior would be executed, otherwise it needs to wait for a safer gap before changing lanes.

The safe gap of vehicles is calculated as follows.

$$S_{sg} = d_{sg}^f + d_{sg}^b + d_1 + d_2 + d_3 \quad (8)$$

$$d_{sg}^f(t) = v_e(t) \times T_e + d_{bg}^e(t) - d_{bg}^l(t) \quad (9)$$

$$d_{sg}^b(t) = v_f(t) \times T_f + d_{bg}^f(t) - d_{bg}^l(t)$$
$$= \begin{cases} v_s(t) \times n_s(t) - v_r^s(t) \times n_s(t)\frac{n_s(t)+1}{2}\Big)T_s \\ v_r^s(t) = bs_{max} \\ n_s(t) = \left\lceil \frac{v_s(t)}{v_r^s(t)} \right\rceil \end{cases} \quad (10)$$

where, d_{sg}^f is the safe front gap of vehicles, d_1 is the minimum parking gap for following vehicles, d_2 is the ego vehicle's length, d_3 is the minimum parking gap for ego vehicles, $d_{sg}^f(t)$ and $d_{sg}^b(t)$ are the safe gap difference from the leading vehicle and the back vehicle at time t, $v_e(t)$ and $v_f(t)$ are the speed of ego vehicle and following vehicle at time t, T_e and T_f are the headway of ego vehicle and following vehicle, T is the time step.

5 Results and analysis

5.1 Evaluation indicators

In the simulation experiments, the average value of the travel time (AVG TT) and the minimum value of time to collision (minTTC) of all HVs were recorded and calculated by Eqs 11, 12. AVG TT shows the overall travel efficiency of all vehicle in the simulation scenarios. MinTTC shows the worst safety situation influenced by the avoidance modes. Therefore, the best avoidance mode needs to decrease the negative impact on efficiency and increase the safety of the whole off-ramp area.

$$TT = \frac{l}{v} \quad (11)$$

$$minTTC = min\{minTTC_1, minTTC_2, \ldots, minTTC_n\} \quad (12)$$

where, l is the distance between the HV and ramp, v is the vehicle speed,

The relative value (RV) of TT and minTTC between the avoidance mode and no-avoidance mode were calculated by Eqs 13, 14.

$$RV_{TT_i} = AVG(TT_{m_i}) - AVG(TT_{m_0}) \quad (13)$$

$$RV_{TTC_i} = minTTC_{m_i} - minTTC_{m_0} \quad (14)$$

where, m_i is the avoidance mode i, m_0 is the no-avoidance mode, and AVG is the abbreviation of average.

5.2 RV_{TT} of three modes under different traffic flow

In Figure 4, it can be observed that the increase of the traffic flow volume generally leads to a noticeable upward trend in the average travel time (AVG TT) of HVs across various ORVs ratio scenarios. To illustrate the impact of three avoidance modes on GVs, ORVs and overall HVs, the detailed analysis is presented as follows.

5.2.1 Impact on the RV_{TT} of GVs

As shown in Figure 4, among the 15 scenarios (5 × 3 = 15, five ratio levels and three traffic flow volume), the gap mode performed best in 13 scenarios and the overall mode performed best in two scenarios in the discrete flow scenario. In these scenarios, these two modes made the HVs' travel time shorter than the no-avoidance mode (keeping going-straight without any avoidance behavior). As shown in Figure 5, the gap mode performed best in 13 scenarios in the continuous flow scenario. Therefore, the gap mode is applicable in most scenarios. The overall mode is only suitable for a few
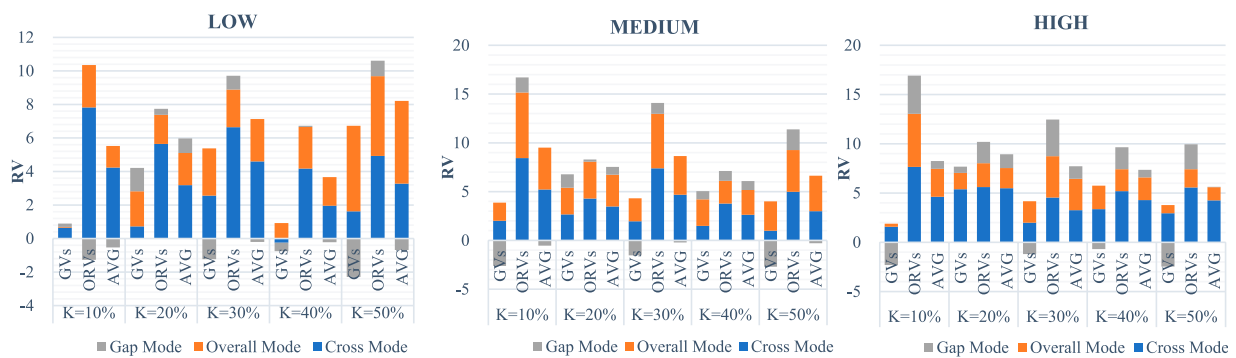
FIGURE 4
RV_{TT} of discrete flow under three kinds of traffic volume.
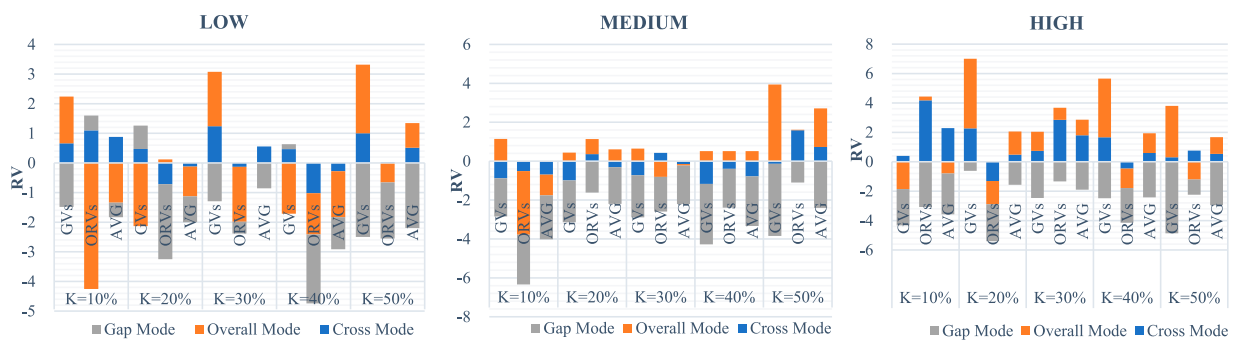
FIGURE 5
RV_{TT} of continuous flow under three kinds of traffic volume.

scenarios. The cross mode did not help to decrease the GVs' travel time, but increased it in almost all scenarios.

5.2.2 Impact on the RV_{TT} of ORVs

As shown in Figure 5, the gap mode performed best in 13 scenarios. It was found that the gap mode was still benefit for the majority of the ORVs in the discrete flow scenario. The overall mode had negative influence on the ORVs' travel time (the minimum travel time of ORVs changed from 2s to 4s). However, in the continuous flow scenario, no-avoidance showed good performance in 14 scenarios. This means that the avoidance behavior of truck platoon is not useful to reduce the ORVs' travel time.

5.2.3 Impact on the RV_{TT} of all HVs near the off-ramp area

The AVG TT represented the overall impact of different modes on all nearby HVs.

As shown in Figure 4, the gap mode performed well in 12 scenarios and the overall mode performs well in three scenarios in the discrete flow scenario. As for the continuous flow scenario, the gap mode performed well in 13 scenarios and the overall mode performs well for the HV in the three scenarios.

The above results show that in the discrete flow scenario, adopting a certain avoidance mode (overall mode or gap mode)

is useful to improve the traffic efficiency near the off-ramp. The gap mode obviously showed the best performance among the three modes to reduce the HVs' AVG TT in the continuous scenario.

In the discrete flow, when the traffic flow was between 1800–2,300 pcu/h, the overall mode and gap mode had different performance in decreasing AVG TT of HVs with the increase of K. However, in the medium flow (2,400–2,900 pcu/h) or high flow (3,000–3,500 pcu/h), only the gap mode could decrease the AVG TT of HVs. Therefore, the gap mode was more suitable for medium and high traffic flow volume scenarios. It can be concluded that the cross mode had a poor performance in reducing the AVG TT of HVs. On the contrary, the gap mode was the best choice in improving the driving efficiency in the mixed traffic flow near the off-ramp.

5.3 Game-theoretic analysis of safety and efficiency

In this study, the relative value of AVG TT is used to measure the driving efficiency. The small AVG TT value can result in more efficient avoidance mode. The relative value of minTTC is used to measure the driving safety. The large minTTC value means a safer avoidance mode. Therefore, a perfect avoidance mode can reach the
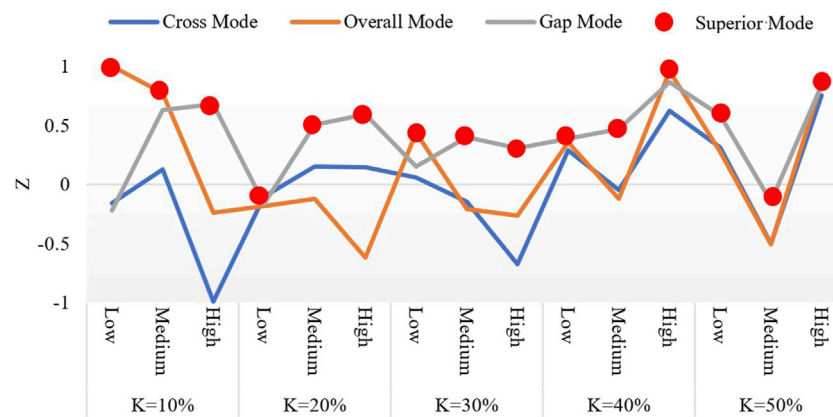
FIGURE 6
Z-value of each avoidance mode in the discrete flow.
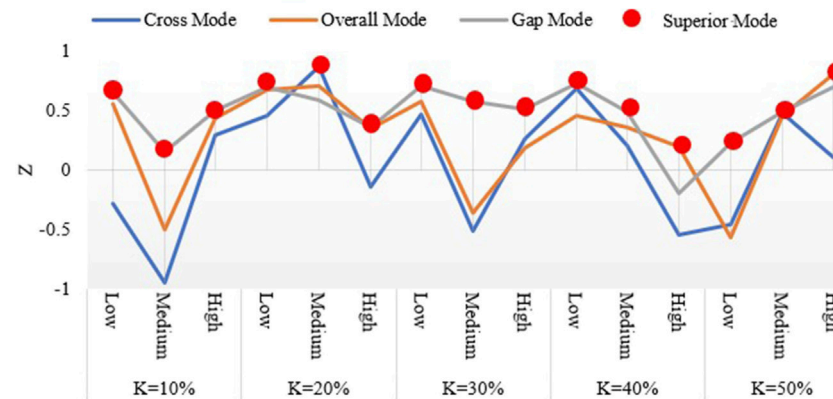
FIGURE 7
Z-value of each avoidance mode in the continuous flow.

relative maximum value (maxZ) of two aspects. The calculation model is Eq. 15.

$$maxZ = \pm \sqrt{\left(E_{q,k,m_i} - E_{q,k,m_0}\right)^2 \pm \left(S_{q,k,m_i} - S_{q,k,m_0}\right)^2} \quad (15)$$

where, E is the maximum relative value of average travel time in GVs and ORVs, S is the minimum relative value of minTTC in lane-changing area, q is the rate of flow volume, which includes low, medium and high, K is the ratio levels of the off-ramp HVs, which includes 10%, 20%, 30%, 40% and 50%, m_i is the avoidance mode i, and m_0 is the no-avoidance mode.

To get the solution of the above equation, the Euclidean Distance based single-objective linear programming method is applied to calculate the maximum Z-value. The mode with large Z-value is safer and more efficient in the specific scenario. When Z > 0, the avoidance mode is more superior than no-avoidance mode in terms of safety and efficiency. Otherwise, when Z < 0, the performance of the avoidance mode is not as good as no-avoidance mode. When Z = 0, it means that the avoidance mode and no-avoidance have the same effect, so it is not necessary to take avoidance behavior for the ORVs.

Based on this, the Z-value of each avoidance mode in the discrete flow and continuous flow scenarios are obtained and shown in Figure 6 and Figure 7.

The above two figures show that gap mode (the grey line) had the highest Z-value among the 15 discrete flow and 15 continuous flow scenarios. The gap mode was suitable for 70% off-ramp scenarios. The overall mode was suitable for 23.3% off-ramp scenarios, and the cross mode was suitable for only 6.7% off-ramp scenarios.

In the cases of Z < 0, three avoidance mode had different limitations among the 30 situations.

• Cross mode

The Z-value of this mode were below 0 in nearly half of the 30 scenarios. It performed better under the high ORVs ratio in discrete traffic flow. In the continuous flow, it was the best choice under the 20% ORVs ratio in the medium traffic flow. However, in the rest circumstances, its performance was not stable and was inferior to the other two modes.

TABLE 4 The game analysis results of three modes in all scenarios (discrete).

Superior avoidance mode	Low	Medium	High	Best mode
K = 10%	cross	overall, gap, cross	gap	overall
K = 20%	None	gap, cross	gap, cross	gap
K = 30%	overall, gap, cross	gap	gap,	gap
K = 40%	gap, overall, cross	gap	overall, gap, cross	gap
K = 50%	gap, cross, overall	None	overall, gap, cross	gap
Best Mode	overall	gap	gap	gap

Note: ▮ is the cross mode, ▮ is the overall mode, ▮ is the gap mode.

TABLE 5 The game analysis results of three modes in all scenarios (continuous).

Superior avoidance mode	Low	Medium	High	Best mode
K = 10%	gap, overall	gap	gap, overall, cross	gap
K = 20%	gap, overall, cross	cross, overall, gap	gap, overall	gap
K = 30%	gap, overall, cross	gap	gap, cross, overall	
K = 40%	gap, cross, overall	gap, overall, cross	overall	overall
K = 50%	gap	gap, cross, overall	gap, cross, overall	gap
Best Mode	gap	gap	overall	

Note: ▮ is the cross mode, ▮ is the overall mode, ▮ is the gap mode.

- Overall mode

The overall mode performed badly in the high discrete traffic low under low ORVs ratio (K = 10%, 20%, 30%) and the medium discrete traffic low under high ORVs ratio (K = 30%, 40%, 50%). In the continuous flow, this mode is not an option for the medium traffic flow under low ORVs ratio (K = 10% and K = 30%) and the low traffic flow under high ORVs ratio (K = 50%).

- Gap mode

It had no benefit for the low traffic flow under low ORVs ratio (K = 10% and K = 20%) and the medium/high traffic flow under high ORVs ratio (K = 50%/K = 40%).

Therefore, the gap mode was the best choice for the truck platoon to improve the safety and efficiency for the mixed traffic flow in the most scenarios. The other two modes were preferred in some specific scenarios. Table 4 and Table 5 show the game analysis results of all scenarios.

In low traffic volume and low-level ORVs ratio, the overall mode is optimal for discrete off-ramp vehicles. For other cases at low ORVs ratio, the gap mode and cross mode are better choice. In medium traffic flow, the cross mode and gap mode are optimal for continuous off-ramp vehicles. In high traffic volume, the overall mode is optimal if the ORV ratio is K = 50% or 40% but it does not perform well in medium and low traffic flow for continuous off-

ramp vehicles. For other cases at high flow rate, the gap mode is optimal.

6 Discussion

The simulation results show that in the two-lane highway off-ramp area, the AVG TT of HVs increases with the traffic flow volume under mixed traffic flow. This is because when the volume is low, HVs can drive freely at the ideal speed or designed speed. It is easy for them to slow down and change lanes near the ramp, which can result in small influence on the other vehicles or truck platoon. With the increase of traffic flow, the traffic density near the off-ramp will also increase. In this situation, once the ORVs decelerate, several GVs following the ORVs need to decelerate, which leads to a decrease in the average speed and travel time of HVs. This phenomenon indicates that the off-ramp demand and state are the basis to evaluate the need and performance of truck platoon avoidance mode. It can be seen that the above results have some similarity with the previous studies. For example, Shen [29] mentioned that in order to improve the traffic efficiency of urban expressways, V2V technology was needed near the expressway off-ramp and downstream intersections as collaborative control areas. This indicates that the status of the off-ramp is an important scenario when evaluating the demand and performance of the truck platoon avoidance modes. The traffic volume and off-ramp

vehicle ratio were considered in past studies [20, 30], in which high traffic volume or ORV ratio would result in the platoon-single vehicle congestion, so short platoon modes were preferred. This is consistent with the results in Tables 4, 5 that gap mode is more suitable in high traffic volume and ORV ratio scenarios, and overall mode is the best mode in low traffic volume and ORV ratio scenarios. On the other hand, this research also has some difference with previous studies. Some studies focused on the operation characteristics and safety effects [31]. The lateral and longitudinal steady control is the key factors to guarantee the safety of truck platoons and other vehicles near the off-ramp. However, this study does not focus on the micro-aspect of single truck operation, but applies best avoidance mode to enhance the overall safety and efficiency of off-ramp area, which can be applied in most two-lane highway off-ramp scenarios. Different from Liu et al.'s study [20] and Wang et al.'s study [32], three avoidance modes were proposed and analysed in the off-ramp simulation experiments, other than the single sub-platoon control mode or merging scenarios. This provides a macro-decision strategy for platoon management system. According to the simulation and game analysis results, the main findings are discussed as follows.

6.1 The efficiency gain of different avoidance modes

The results show that there are obvious differences in efficiency gain among three avoidance modes. In some continuous flow scenarios, if the truck platoon takes no-avoidance action, the ORVs' AVG TT of most scenarios are smaller than the three modes. The reason is that the continuous lane-changing demand requires a large gap in the truck platoon to ensure the safe off-ramp behavior. Therefore, when truck platoon slows down or changes lanes to yield for a high rate of continuous ORVs, the following ORVs and ORVs on the adjacent lane have to slow down as well, which results in the increase of ORVs' AVG TT. To overcome this problem, the gap mode is a solution to provide ORVs with enough space to change lanes before the lane-changing area. It makes the truck platoon to split and leave enough space for ORVs to change lanes. This can not only reduce the waiting time for the ORVs, but also release the left lane for the GVs to drive freely. Therefore, this mode greatly reduces the traffic delay and improves the HVs and ORVs' driving efficiency.

The poor performance of the cross mode is due to the low gap acceptance between the sub-platoon. HVs have to follow the sub-platoon for a few seconds to seek for a lane-changing opportunity. On the other hand, before the ORVs change lanes, they need to check the traffic state on the target lane. If there are already several ORVs in the "cross gap" (the gap between the first and the third sub-platoon), the following ORVs have to wait until the target lane has enough space. These two aspects are the main reasons for the large AVG TT result when the truck platoon taking the cross mode.

Considering the lane-changing characteristics of truck platoon, the overall mode yields for the ORVs by changing the whole platoon to the left lane. This requires no additional control of sub-platoon. However, the overall mode has some limitations in some situations.

For example, when the HVs travel on the left lane near the off-ramp, the truck platoon cannot directly change lanes. The container trucks in truck platoon must change lanes one by one according to the nearby traffic conditions. Sometimes, the following trucks have to look for another appropriate lane-changing opportunity after the leading truck changes lanes. This will reduce the average speed of HVs on the left lane, which results in the decrease of the HVs' driving efficiency.

6.2 The safety gain of different avoidance modes

When the ORVs reach the off-ramp area, they need enough space on the right lane to meet their lane-changing demand. In the no-avoidance state, the truck platoon occupies 100 m on the right lane, so that the ORVs have two choices to get off the highway: a) overtaking the whole truck platoon before reaching the end line of the lane-changing area. b) keeping a distance with the truck platoon, and waiting for the truck platoon passing the off-ramp exit, then finding the appropriate opportunity to change lanes. If the ORVs need to get off the ramp quickly, they must sacrifice the safety to decrease the relative distance and to increase the speed. Therefore, the most unsafe status usually appears in this period.

If the traffic flow volume increases constantly and reaches a high flow volume, the traffic density will increase. Therefore, when HVs are closer to the off-ramp and still do not change to the right lane, they will develop a strong intention to change lanes. If truck platoon disbands enough space, ORVs will adjust the speed between the two sub-platoons. Otherwise, there will be a risk of rear-end accident. In the simulation results, this phenomenon is reflected with low AVG min TTC of HVs.

Generally, the avoidance of truck platoon near the highway off-ramp can increase traffic safety. The gap mode provides great space for ORVs to meet their lane-changing requirement. However, the gap mode will cause the rear sub-platoon to decelerate, which will increase the conflict probability between the rear sub-platoon and the following HVs. In addition, if the leading platoon does not accelerate enough, the ORVs cannot overtake the platoon before the lane-changing area. In this case, they must wait on the left lane at the off-ramp exit, which results in traffic congestion (choice b). The overall mode allows the ORVs to change lanes freely by clearing the right lane, which reduces the conflicts between HVs and other vehicles. However, during the lane-changing process of each container truck, the conflict probability between truck platoon and HVs climbs. The HVs following or travelling in parallel with truck platoon must decelerate to the same speed as the platoon to change lane safely. Therefore, how to balance the truck lane-changing behavior and traffic efficiency is an important issue of this avoidance mode. The cross mode combines the advantages and disadvantages of the other two modes. It includes the lane change and disband behavior of platoons, so it provides proper gap for ORVs to change lanes. Compared with the other two modes, the minTTC of cross mode is larger. This is because the ORVs in this mode usually change lanes at a relative slower speed, which sacrifices the travel efficiency but improves safety.

6.3 Summary

A comprehensive evaluation of different avoidance modes in mixed traffic flow necessitates an examination from two key perspectives: travel efficiency and safety. The findings indicate that the gap mode consistently demonstrates superior performance across various scenarios, while the overall mode exhibits commendable efficiency in several instances. Conversely, the cross mode proves applicable in only a limited number of scenarios. Comparative analysis against a no-avoidance state reveals that all three modes contribute to enhancing the efficiency and safety of both human-driven vehicles (HVs) and truck platoons to a certain extent.

The results of game analysis highlight the gap mode as the optimal choice, displaying the most favourable impact on travel efficiency and safety, especially with increasing traffic volume and off-ramp vehicle ratio. Although not always the primary choice in specific situations, the gap mode remains a viable option, particularly for static truck platoon control systems. In cases where individual trucks within the platoon can be monitored and controlled, the avoidance mode should be dynamically tailored to match the traffic environment, as outlined in Table 4 and Table 5.

7 Conclusion

This study introduces three avoidance modes for truck platoons in the off-ramp area of a two-lane highway within a human-machine mixed traffic flow context. Employing game theory to analyze the safety and efficiency of each avoidance mode, the study concludes that the gap mode proves most effective in enhancing travel safety and efficiency. By segmenting the long platoon into multiple sub-platoons, the gap mode generates ample spaces for multiple off-ramp vehicles.

However, in scenarios characterized by high traffic volume and consecutive high-speed vehicles waiting to exit the highway, the overall mode emerges as a preferable option over the gap mode. In the overall mode, all trucks shift lanes to the left, significantly reducing travel time for nearby vehicles. Additionally, the overall mode demonstrates adaptability to various Off-Ramp Vehicle (ORV) ratio levels in low traffic volume, ensuring greater safety for high-speed human-driven vehicles compared to other avoidance modes.

Despite these findings, this study still has some limitations. Firstly, all vehicles in the scenarios are assumed to be connected in the Intelligent Transportation Systems environment, but information transmission delay in avoidance mode is not analyzed. Secondly, the study only focuses on the differences between the three modes without considering the position, speed, and length of sub-platoons for each mode. Thirdly, the evaluation indicators for HVs do not include other factors, such as queue length and deceleration rate. Consequently, future research needs to conduct field tests to investigate the safety, traffic efficiency, and response delay of diverse avoidance modes in the Vehicle-to-Everything (V2X) environment. Detailed micro-macro evaluation of different modes would be preferred to obtain more accurate avoidance strategies.

Data availability statement

The raw data supporting the conclusion of this article will be made available by the authors, without undue reservation.

Author contributions

YL: Conceptualization, Methodology, Writing–original draft, Funding acquisition. LW: Writing–original draft, Writing–review and editing. ZX: Formal Analysis, Writing–original draft, Writing–review and editing. WS: Data curation, Software, Writing–original draft.

Conflict of interest

Author WS is employed by Lianyungang New Harbour Terminal Co., Ltd., Jiangsu, China.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Tsugawa S, Jeschke S, Shladovers SE. A review of truck platooning projects for energy savings. *IEEE Trans Intell Vehicles* (2016) 1:68–77. doi:10.1109/tiv.2016.2577499

2. Alam A, Besselink B, Turri V, Mårtensson J, Johansson KH. Heavy-duty vehicle platooning for sustainable freight transportation: a cooperative method to enhance safety and efficiency. *IEEE Control Syst Mag* (2015) 35(6):34–56. doi:10.1109/MCS.2015.2471046

3. Bösch PM, Becker F, Becker H, Axhausen KW. Cost-based analysis of autonomous mobility services. *Transport Policy* (2018) 64:76–91. doi:10.1016/j.tranpol.2017.09.005

4. Wadud Z, MacKenzie D, Leiby P. Help or hindrance? The travel, energy and carbon impacts of highly automated vehicles. *Transportation Res A: Pol Pract* (2016) 86:1–18. doi:10.1016/j.tra.2015.12.001

5. Motamedidehkordi N, Margreiter M, Benz T. Effects of connected highly automated vehicles on the propagation of congested patterns on freeways. In: Proceedings of the 95th Annual Meeting of the Transportation Research Board; Washington DC, United States (2016). p. 16–1802.

6. Ramezani H, Shladover SE, Lu XY, Altan OD. Micro-simulation of truck platooning with cooperative adaptive cruise control: model development and a case study. *Transportation Res Rec* (2018) 2672(19):55–65. doi:10.1177/0361198118793257

7. Bouchery Y, Hezarkhani B, Stauffer G. Coalition formation and cost sharing for truck platooning. *Transportation Res B: Methodological* (2022) 165:15–34. doi:10.1016/j.trb.2022.08.007

8. Calvert SC, Schakel WJ, van Arem B. Evaluation and modelling of the traffic flow effects of truck platooning. *Transportation Res C: Emerging Tech* (2019) 105:1–22. doi:10.1016/j.trc.2019.05.019

9. Ali A, Garcia G, Martinet P. The flatbed platoon towing model for safe and dense platooning on highways. *IEEE Intell Transportation Syst Mag* (2015) 7(1):58–68. doi:10.1109/mits.2014.2328670

10. Zhao C, Li L, Li J, Li K, Li Z. The impact of truck platoons on the traffic dynamics around off-ramp regions. *IEEE Access* (2021) 9:57010–9. doi:10.1109/access.2021.3072070

11. Chandra S, Thai T. Analyzing freight truck platoon accessibility with route deviations. *Sustainability* (2022) 14(4):2130–17. doi:10.3390/su14042130

12. Faber T, Sharma S, Snelder M, Klunder G, Tavasszy L, van Lint H. Evaluating traffic efficiency and safety by varying truck platoon characteristics in a critical traffic situation. *Transportation Res Rec* (2020) 2674(10):525–47. doi:10.1177/0361198120935443

13. Buisson C, Keyvan-Ekbatani M, Wagner P. Impede autonomous vehicles merging at on-ramps. In: Proceedings of the Transportation Research Board 97th Annual Meeting; Washington DC, United States (2018). p. 18–00938.

14. Duret A, Wang M, Leclercq L. Truck platooning strategy near merge: heuristic-based solution and optimality conditions. In: Proceedings of the Transportation Research Board 97th Annual Meeting; Washington DC, United States (2018). p. 18–00902.

15. Scarinci R, Heydecker B. Control concepts for facilitating motorway on-ramp merging using intelligent vehicles. *Transport Rev* (2014) 34(6):775–97. doi:10.1080/01441647.2014.983210

16. Tabibi M. Automated truck traffic: design and Control of automated truck traffic at motorway ramps (2014). Available from: http://resolver.tudelft.nl/uuid:d1009e78-be81-4140-8930-5028ffe455c0 (Accessed January 15, 2024).

17. Wu R, Li L, Lu W, Rui Y, Ran B. Improving the safe operation of platoon lane changing for connected automated vehicles: a novel field-driven approach. *Appl Sci* (2021) 11(16):7287–17. doi:10.3390/app11167287

18. Shi Y, He Q, Huang Z. Capacity analysis and cooperative lane changing for connected and automated vehicles: entropy-based assessment method. *Transportation Res Rec* (2019) 2673(8):485–98. doi:10.1177/0361198119843474

19. Wang D, Hu M, Wang Y, Wang J, Qin H, Bian Y. Model predictive control-based cooperative lane change strategy for improving traffic flow. *Adv Mech Eng* (2016) 8(2):168781401663299–17. doi:10.1177/1687814016632992

20. Liu Y, Wang X, Li L, Cheng S, Chen Z. A novel lane change decision-making model of autonomous vehicle based on support vector machine. *IEEE Access* (2019) 7:26543–50. doi:10.1109/access.2019.2900416

21. Jing S, Hui F, Zhao X, Rios-Torres J, Khattak AJ. Cooperative game approach to optimal merging sequence and on-ramp merging control of connected and automated vehicles. *IEEE Trans Intell Transportation Syst* (2019) 20(11):4234–44. doi:10.1109/tits.2019.2925871

22. Zhang C, Wang B, Yang S, Zhang M, Gong Q, Zhang H. The driving risk analysis and evaluation in rightward zone of expressway reconstruction and extension engineering. *J Adv Transportation* (2020) 2020:1–13. doi:10.1155/2020/8943463

23. Zhou M, Qu X, Jin S. On the impact of cooperative autonomous vehicles in improving freeway merging: a modified intelligent driver model-based approach. *IEEE Trans Intell Transportation Syst* (2017) 18(6):1–7. doi:10.1109/tits.2016.2606492

24. Saifuzzaman M, Zheng Z. Incorporating human-factors in car-following models: a review of recent developments and research needs. *Transportation Res Part C: Emerging Tech* (2014) 48:379–403. doi:10.1016/j.trc.2014.09.008

25. Milanés V, Shladover SE. Modeling cooperative and autonomous adaptive cruise control dynamic responses using experimental data. *Transportation Res Part C: Emerging Tech* (2014) 48:285–300. doi:10.1016/j.trc.2014.09.001

26. Xiao L, Wang M, Van Arem B. Realistic car-following models for microscopic simulation of adaptive and cooperative adaptive cruise control vehicles. *Transportation Res Rec* (2017) 2623:1–9. doi:10.3141/2623-01

27. Xiao L, Wang M, Schakel W, van Arem B. Unravelling effects of cooperative adaptive cruise control deactivation on traffic flow characteristics at merging bottlenecks. *Transportation Res Part C: Emerging Tech* (2018) 96:380–97. doi:10.1016/j.trc.2018.10.008

28. Erdmann J. SUMO's lane-changing model. In: *Modeling mobility with open data* (2015). p. 105–23.

29. Shen H. Research on control strategies and methods of entrances and exits of urban expressway-Take Beijing Expressway as an example. *Road Traffic Sci Tech* (2018) 13–9.

30. Pang M, Huang J. Cooperative control of highway on-ramp with connected and automated vehicles as platoons based on improved variable time headway. *J Transportation Eng A: Syst* (2022) 148(7):04022034. doi:10.1061/jtepbs.0000683

31. Liu Z, Ma Q, Shao H, Fang R, Lin T. Study on the exit characteristics and safety impact of single lane exit ramp of expressway. *Highway* (2023) 333–42.

32. Wang M, van Maarseveen S, Happee R, Tool O, van Arem B. Benefits and risks of truck platooning on freeway operations near entrance ramp. *Transportation Res Rec* (2019) 2673(8):588–602. doi:10.1177/0361198119842821

Frontiers in
Physics

Investigates complex questions in physics to understand the nature of the physical world

Addresses the biggest questions in physics, from macro to micro, and from theoretical to experimental and applied physics.

Discover the latest Research Topics

See more →

Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact

frontiers

Frontiers in
Physics