

The Influence of AI Algorithm-aided Decision on the Diffusion of Industrial Chain Integration Innovation Decision

Li Dan¹ and Jiao Runyu²

¹School of Business Administration, Liaoning University of Technology, Xingcheng 125105,
Liaoning Province CHINA;
Email : 2106116882@qq.com

²School of Business Administration, Liaoning University of Technology, Xingcheng 125105,
Liaoning Province CHINA;
Email : 13589416990@163.com

ABSTRACT

Algorithm-assisted decision-making is a crucial approach to enhance decision-making efficiency, scientific rigor, and objectivity. This study constructs a dual-layer coupled network model integrating evolutionary game theory and deep reinforcement learning to investigate the impact of algorithmic intervention on innovation decision diffusion. Simulation results demonstrate: (1) Embedding algorithm-assisted decision-making significantly accelerates innovation diffusion speed and stability, yet exhibits a dual-threshold effect in algorithm weight allocation. Additionally, diffusion dynamics are influenced by upstream-downstream integration innovation ratios, exploration rates, network scale, and initial diffusion proportions. (2) When initial algorithm weights approach thresholds, decision diffusion demonstrates strong adaptability to integration ratios, strategy update sensitivity, and initial diffusion proportions, whereas high exploration rates and undersized networks markedly suppress diffusion velocity. Low algorithm weights hinder integration innovation between upstream and downstream enterprises in driving diffusion. (3) When initial algorithm weights exceed thresholds, innovation diffusion becomes constrained, yet efficiency can be partially restored by increasing initial integration ratios and network scale. These findings provide theoretical foundations for optimizing enterprise innovation decision processes and guiding governmental policies on algorithmic governance and industrial chain coordination.

Keywords: Artificial Intelligence, Algorithm-Assisted Decision-Making, Integration Innovation, Decision Diffusion, Complex Network.

Mathematics Subject Classification: 68T07, 91A25, 91A35, 91A44
Computing Classification System: I.2.5

1. INTRODUCTION

Under the backdrop of accelerated economic globalization and technological innovation, integration innovation across industrial chains has emerged as a pivotal strategy to enhance corporate competitiveness and drive industrial upgrading. Through industrial chain integration innovation, resource sharing and complementary advantages can be achieved, accelerating the Research and Development (R&D) of new technologies and products⁰, effectively addressing industrial chain bottlenecks, and advancing technological supplementation and synergistic strengthening. However,

challenges such as heterogeneity among decision-making entities⁰, information asymmetry⁰, and others persist in the diffusion of these decisions, necessitating novel decision-making paradigms to transcend the limitations of traditional game-theoretic analyses.

The groundbreaking advancements in artificial intelligence and big data technologies have provided enterprises with new tools for decision-making. According to the *Survey Report on Manufacturing Enterprises' Integration of DeepSeek and Other Large Model Applications* released on April 10, 2025, among 159 manufacturing enterprises surveyed, 78.7% had already adopted or planned to adopt large models like DeepSeek, indicating a growing reliance on algorithmic assistance in critical decisions⁰. Techniques such as data mining, random forests, and reinforcement learning—particularly the latter, which enables autonomous learning from historical actions⁰—are increasingly utilized. Specifically, reinforcement learning optimizes decision-making by determining optimal actions in specific states to maximize benefits⁰. Consequently, the introduction of algorithms may reshape enterprises' innovation decision-making processes and strategic choices, thereby influencing the diffusion of integration innovation decisions across industrial chains. On one hand, algorithms enhance decision-making efficiency through real-time information sharing and rapid data processing, improving scientific rigor and objectivity⁰⁰, thus fostering collaborative diffusion of innovation decisions⁰. On the other hand, the "black-box" nature of algorithms may introduce decision biases⁰ and exacerbate free-riding behaviors within innovation networks.

Building on existing research, this study constructs a coupled network model integrating evolutionary game theory and deep reinforcement learning to investigate the impact of algorithm-assisted decision-making on the diffusion of integration innovation decisions under varying scenarios. The research motivations are threefold: First, how does algorithm-assisted decision-making alter the evolutionary trajectory of innovation decision diffusion? Second, does the influence of algorithmic intervention on integration innovation diffusion vary depending on assigned algorithm weights? Third, while prior studies predominantly focus on algorithmic risks, ethics, and governance^[0-0]—examining consequences in contexts such as price competition⁰ and R&D rivalry⁰—the role of algorithmic decision-making in industrial chain innovation diffusion remains underexplored. For China, where AI adoption is vigorously promoted, it is imperative to identify potential innovation-inhibiting effects under algorithmic empowerment.

This paper employs simulation analysis within an industrial chain integration innovation context to examine the potential outcomes of reinforcement learning-assisted decision-making, addressing the above questions. The findings aim to elucidate the role of algorithmic assistance in corporate innovation decisions, offering guidance for enterprises to optimize decision-making processes and improve innovation efficiency, while providing policymakers with insights to formulate industrial chain integration promotion strategies.

2. LITERATURE REVIEW

Integration innovation, as a cross-organizational collaborative innovation paradigm driven by industrial chain division of labor among upstream, midstream, and downstream enterprises, as well as between

large firms and SMEs⁰⁰, fundamentally strengthens the technological and economic linkages among enterprises through "vertical integration" and "horizontal integration." This enhances resource allocation efficiency, innovation efficiency, and generates synergistic effects that amplify overall innovation performance⁰.

Specifically, vertical integration manifests as innovation collaborations among enterprises across different industrial chain segments to boost overall competitiveness. High product interdependency fosters strong interest alignment, facilitating the formation of innovation communities and vertical integration mechanisms characterized by transmission, learning, and safeguarding processes⁰⁰. Horizontal integration, conversely, focuses on competitive enterprises at the same industrial chain level adopting alliances, joint production, or similar approaches to mitigate excessive competition, reduce innovation costs, share risks, and accelerate product development⁰, thereby establishing diffusion-, competition-, and learning-driven innovation mechanisms⁰. Existing studies on promoting integration innovation predominantly explore digital transformation⁰, technology spillovers⁰, smart manufacturing⁰, and policy tools⁰, emphasizing how technological applications optimize resource integration or incentivize collaboration. However, there remains a lack of in-depth analysis on dynamic optimization mechanisms within the decision-making process itself. Thus, while prior research has extensively examined the connotation, collaborative mechanisms, and influencing factors of industrial chain integration innovation, systematic studies on providing decision-makers with toolkits to enhance integration innovation decision efficiency and accelerate decision diffusion remain scarce.

Advances in artificial intelligence, particularly deep reinforcement learning (DRL), offer new pathways for optimizing decisions in dynamically uncertain environments through agent-environment interactions⁰. The evolution of DRL—from Mnih et al.'s (2015) Deep Q-Network (DQN)⁰ to Silver et al.'s (2017) AlphaZero framework for autonomous strategy generation⁰—marks its expansion from closed environments to open, complex systems. In corporate decision-making, DRL's transformative role manifests in two dimensions: R&D innovation and industrial chain coordination. On one hand, DRL can simulate inter-firm R&D competition⁰ and design mechanisms for exploratory innovation⁰; on the other, it significantly enhances industrial chains' capacity to navigate market uncertainties⁰. Concurrently, DRL has catalyzed the rise of "human-machine collaboration" paradigms, such as the Soft Actor-Critic algorithm balancing safety and efficiency in industrial robotics⁰, productivity improvements through human-machine synergy⁰, and optimization models for supply chain innovation investments⁰. However, deepening human-machine collaboration faces critical challenges: dynamic control allocation in autonomous vehicles⁰, reconciling algorithmic efficiency with expert experience in transportation scheduling⁰, and defining algorithmic boundaries in judicial decision-making⁰, highlighting the necessity of algorithmic governance and human autonomy. Collectively, existing studies reveal the potential of algorithms to optimize decisions in contexts like R&D competition and supply chain coordination, yet few address innovation decision diffusion while balancing algorithmic assistance and human agency.

Complex network theory provides a structural-behavioral framework for analyzing innovation diffusion across industrial chains. The small-world network model proposed by Watts and Strogatz (1998), characterized by high clustering coefficients and short average path lengths⁰, has been widely

adopted to capture the dual processes of "local infiltration" and "global leap" in innovation diffusion⁰. In traditional industrial chains, enterprises within clusters form localized innovation pools through strong ties⁰, while hub enterprises bridge cross-module connections to facilitate technology diffusion⁰. These structural properties make small-world networks ideal for simulating innovation diffusion⁰. However, most studies focus on single-layer homogeneous networks⁰, neglecting the evolution of enterprise decision networks into dual-layer coupled systems—comprising an "enterprise layer" and an "algorithmic agent layer"—under algorithmic intervention⁰. The enterprise layer retains the small-world network's localized collaboration features, while the agent layer dynamically reconfigures task-oriented networks through algorithmic coordination. Algorithmic agents not only act as information conduits but also reshape latent inter-firm connections via predictive recommendations⁰, transcending geographic and organizational constraints on innovation diffusion.

Based on this literature, this paper's marginal contributions are threefold: (1) Prior studies predominantly focus on single-layer homogeneous networks, inadequately addressing human-machine collaborative decision-making scenarios post-algorithmic intervention. By integrating evolutionary game theory and DRL to construct an "enterprise-agent" coupled network, this study combines algorithmic dynamic optimization with complex network structural properties, offering new theoretical insights into algorithm-empowered innovation diffusion. (2) Moving beyond the limitations of existing literature on technological empowerment and policy incentives, this work focuses on dynamic optimization of decision processes, exploring how algorithm-assisted decisions influence integration innovation diffusion, thereby broadening research perspectives. (3) By addressing the dynamic balance between algorithmic assistance and human agency, this study reveals the interactive effects of algorithm weights, upstream-downstream integration ratios, strategy update sensitivity, and network scale on innovation diffusion, providing a theoretical basis for optimizing corporate innovation workflows and guiding government algorithmic governance policies.

3. RESEARCH HYPOTHESES AND MODEL SPECIFICATION

3.1. Problem Description

The strategic choices of focal enterprises in industrial chains can be viewed as outcomes of multi-party game interactions, driven jointly by "competitive pressures" and "algorithmic advisory." On one hand, focal enterprises' integration innovation decisions exhibit a dual-driven feature of vertical collaboration and horizontal game dynamics, influenced by both the innovation incentives of upstream/downstream partners and comparisons with neighboring competitors. Enterprises weigh the benefits of integration innovation against traditional operations, with decisions to adopt integration innovation propagating through social learning mechanisms as early adopters gain advantages. On the other hand, algorithm-assisted decision-making mitigates subjective biases and information asymmetry, providing enterprises with more intelligent and reliable decision support.

Based on this theoretical framework, we first assume each enterprise has two pure strategies: Strategy A (integration innovation) and Strategy B (traditional operations). The decisions of focal enterprises determine the diffusion extent of integration innovation behaviors. Each enterprise interacts closely with upstream/downstream partners. Focal enterprises decide whether to adopt

integration innovation by comparing economic outcomes through either collaborative decision-making with algorithmic agents or game-based interactions with neighboring peers. Let n_A and n_B denote the number of enterprises adopting Strategy A and B, respectively, with $n_A + n_B = N$ (total enterprises). Let I represent the total number of upstream/downstream partners, and $i = I/(n_B + n_A)$ denote the average number of partners per enterprise. Define p as the proportion of integration innovation among upstream/downstream partners, and T as the initial proportion of integration innovation adopters in the network. To quantify decision diffusion, we use $\theta = n_A/(n_B + n_A)$ to represent the adoption rate of integration innovation and $\beta = n_B/(n_B + n_A)$ for traditional operations. The dynamic evolution of decision adoption is thus captured by changes in θ and β .

3.2. Model Specification

The proposed industrial chain integration innovation diffusion model comprises three components: (1) focal enterprise game-theoretic model, (2) network structure and evolutionary rules, and (3) algorithmic agent decision-making.

Focal Enterprise Game-Theoretic Model:

Basic Assumptions: Enterprises engage in strategic games based on predefined payoff functions.

Payoff Matrix: Constructed to quantify economic outcomes of Strategy A and B under varying conditions.

Network Structure and Evolutionary Rules:

An NW small-world network model simulates the game relationships among focal enterprises within the industrial chain.

Strategy Update Rules: Enterprises adjust strategies probabilistically based on payoff comparisons and social learning.

Algorithmic Agent Decision-Making:

A reinforcement learning-based agent is trained to optimize collaborative decisions.

Training Framework: Includes state space design, reward mechanisms, and policy iteration processes to prepare for simulation.

3.3. Focal Enterprise Game-Theoretic Model

Based on practical considerations, the following assumptions are proposed:

H1: Integration innovation among upstream, midstream, and downstream enterprises can effectively share innovation risks, reduce costs, shorten cycles, and enhance efficiency, albeit requiring increased R&D investments and financial support.

Integration innovation adheres to the principles of shared benefits and shared risks⁰, emphasizing efficient knowledge flow and collaborative resource utilization. By integrating and sharing production resources, it maximizes the complementary effects of homogeneous and heterogeneous resources, improving resource allocation efficiency. Additionally, integration innovation across upstream, midstream, and downstream enterprises generates synergistic effects exceeding the sum of individual contributions "1+1>2"⁰ through division of labor, collaboration, and competition, significantly lowering innovation costs.

Thus, we assume the innovation cost under traditional operations for a focal enterprise is C , comprising R&D expenditure a and other related expenses e , i.e., $C = a + e$. For integration innovation, the cost becomes $(1-b)(C + \Delta c)$, where Δc represents the incremental innovation investment required for integration, and b is the cost-efficiency coefficient of integration innovation. Each upstream/downstream partner generates a revenue f for the focal enterprise. When both the focal enterprise and its partners adopt integration innovation, the synergistic revenue⁰ is modeled as $F = i(1 - e^{-ktp})f$

H2: Traditional operations by upstream/downstream partners impose constraints on the focal enterprise's technology and product development, limiting its growth potential and innovation outcomes. Accordingly, traditional operations by partners impose an additional integration innovation cost Δg on the focal enterprise.

H3: When partners adopt integration innovation while the focal enterprise maintains traditional operations, the focal enterprise gains free-riding benefits due to knowledge spillovers but loses competitiveness, resulting in partial disengagement from partners. We assume the free-riding benefit is H . If n neighboring enterprises adopt integration innovation, the focal enterprise m loses a proportion ξ of its upstream/downstream partners.

H4: Competitive relationships exist between focal enterprise m and neighboring enterprise n . Divergent strategies affect competitiveness and market share.

If neighbor n adopts integration innovation while focal enterprise m retains traditional operations, m 's market share decreases by d_1 . Conversely, if m adopts integration innovation and n does not, m 's market share increases by d_2 .

Based on H1–H4, the game payoff matrix for interactions between enterprises is constructed as follows (taking focal enterprise m as an example):

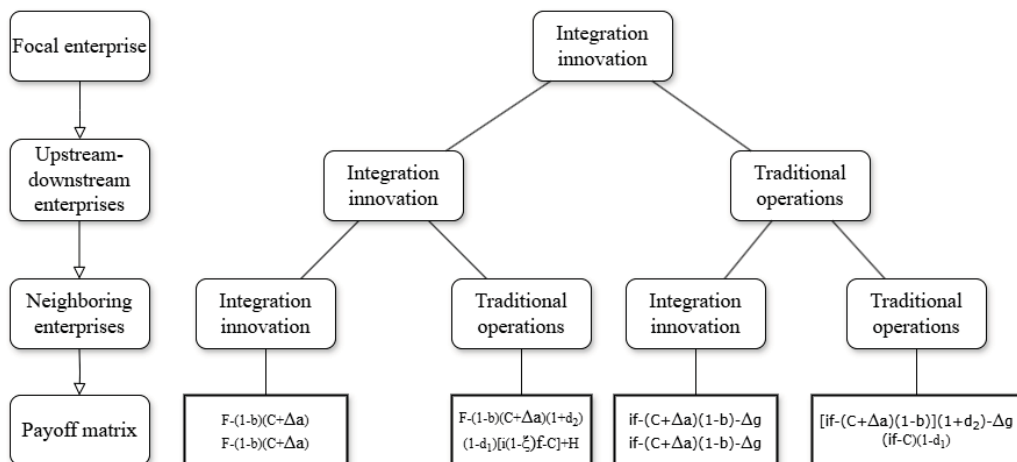


Figure 1. Payoff Matrix for Integration Innovation of Focal Enterprise m

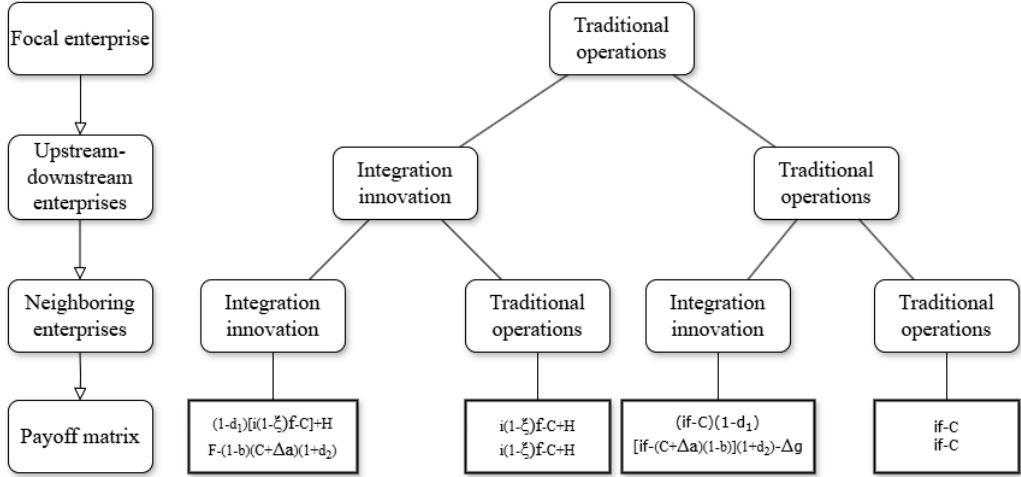


Figure 2. Payoff Matrix for Traditional Operations of Focal Enterprise m

3.4. Network Structure

3.4.1. Basic Framework

We construct a dual-layer coupled "enterprise-agent" network with a total of NN nodes, where each node in the two layers corresponds one-to-one, i.e., every enterprise is embedded with algorithm-assisted decision-making. Both layers adopt small-world network topologies. Existing studies commonly employ Watts-Strogatz (WS) or Newman-Watts (NW) small-world networks⁰. The WS model disrupts original connections by rewiring edges⁰, while the NW model enhances connectivity through random edge additions without breaking existing ties⁰. Given the need to preserve original relationships in simulating the evolutionary game dynamics of integration innovation diffusion, the NW small-world network is selected.

3.4.2. Interlayer Connectivity

Three interlayer coupling modes are common: assortative, disassortative, and random connections⁰. Competitive relationships among enterprises often correlate with finite market resources, human capital, and infrastructure⁰. Large enterprises with intense competition typically possess stronger computational resources (e.g., high-performance clusters, cloud platforms, big data capabilities) to train advanced algorithmic agents, and vice versa. To simplify the model while aligning with research objectives, we adopt assortative coupling: enterprises with higher degrees in the enterprise layer connect to agents with higher degrees in the agent layer, and vice versa. Nodes with equal degrees are randomly paired.

3.4.3. Enterprise Layer Network

The enterprise layer network $G_1 = (V, E)$ is defined as follows:

$V = \{V_1, V_2, \dots, V_N\}$: Node set representing enterprises.

E : Edge set describing competitive-cooperative relationships, represented by the adjacency matrix:

$$E = \begin{bmatrix} e_{11} & e_{12} & \cdots & e_{1N} \\ e_{21} & e_{22} & \cdots & e_{2N} \\ \vdots & \cdots & \ddots & \vdots \\ e_{N1} & e_{N2} & \cdots & e_{NN} \end{bmatrix} \quad (1)$$

Where e_{mn} represents the relationship between nodes m and n , reflecting their competitive-cooperative dynamics:

$e_{mn} = 1$: Direct competitive-cooperative relationship exists between nodes m and n .

$e_{mn} = 0$: No direct competitive-cooperative relationship.

Constraints: Self-loops (e_{mn}) and duplicate edges between the same node pair are prohibited.

For a focal enterprise m :

Initial neighbor count: t .

Edge addition probability: η .

Proportion of neighbors adopting integration innovation: $\frac{t_n^{RT}}{t_n}$.

The expected payoff for m adopting integration innovation (Strategy A) is:

$$U_A = \left(\frac{t_n^{RT}}{t_n}, 1 - \frac{t_n^{RT}}{t_n} \right) \times \begin{pmatrix} F - (1-b)(C + \Delta c) & F - (1-b)(C + \Delta c)(1+d) \\ if - (C + \Delta c)(1-b) - \Delta g & [if - (C + \Delta c)(1-b)](1+d) - \Delta g \end{pmatrix} \begin{pmatrix} p \\ 1-p \end{pmatrix} \quad (2)$$

The expected payoff for traditional operations (Strategy B) is:

$$U_B = \left(\frac{t_n^{RT}}{t_n}, 1 - \frac{t_n^{RT}}{t_n} \right) \begin{pmatrix} (1-d)[i(1-\xi)f - C] + H & i(1-\xi)f - C + H \\ (if - C)(1-d) & if - C \end{pmatrix} \begin{pmatrix} p \\ 1-p \end{pmatrix} \quad (3)$$

3.4.4. Agent Layer Network

The agent layer network $G_2 = (V', E')$ is defined as:

$V' = \{V'_1, V'_2, \dots, V'_N\}$: Node set representing algorithmic agents.

E' : Edge set describing information interactions, represented by the adjacency matrix:

$$E' = \begin{bmatrix} e'_{11} & e'_{12} & \cdots & e'_{1N} \\ e'_{21} & e'_{22} & \cdots & e'_{2N} \\ \vdots & \cdots & \ddots & \vdots \\ e'_{N1} & e'_{N2} & \cdots & e'_{NN} \end{bmatrix} \quad (4)$$

Where e'_{mn} indicates an information interaction between agents m' and n' , and $e'_{mn} = 0$ otherwise.

3.4.5. Enterprise-Agent Coupled Network

A coupled network is a complex system comprising multiple network layers, each with distinct connection patterns that interact through predefined rules⁰.

Building upon the previously constructed enterprise layer network G_1 and agent layer network G_2 , this study establishes an enterprise-agent coupled network. During construction, nodes in the enterprise layer are connected to corresponding nodes in the agent layer via interlayer assortative connections—specifically, enterprises with higher degrees in G_1 are linked to agents with higher degrees in G_2 , and vice versa. Nodes with identical degrees are randomly paired.

The coupled network $G = (Q, R, S, Z)$ integrates the enterprise and agent layers:

Node Set: $Q = \{q_1, q_2, \dots, q_{2N}\}$.

Edge Set: $R = \{\langle q_m, q_n | q_m, q_n \in Q, 1 \leq m, n \leq 2N \rangle\}, R \subseteq Q \times Q$.

Coupling Relationships: $S = S_1 \times S_2 = \{(s_1, s_2) | s_m \in S_m, 1 \leq m \leq 2\}$, with S_1 for enterprise relationships and S_2 for agent relationships.

Mapping: $Z: R \rightarrow S$, defining the correspondence between edges and coupling relationships.

This framework enables systematic analysis of how algorithmic agents and enterprise interactions jointly shape innovation diffusion dynamics.

3.5. Strategy Update

Under the dual-layer coupled small-world network framework, this study constructs a strategy diffusion system driven by interactions between the agent layer and enterprise layer. The enterprise layer nodes represent physical enterprises, while the agent layer serves as an algorithmic decision-making hub, where nodes generate strategy recommendations through Proximal Policy Optimization (PPO) training in a virtual environment. Interlayer coupling of heterogeneous nodes is achieved via cross-network connections, resulting in cross-layer synergistic evolution:

Enterprise Layer: Nodes acquire strategy information from peers through competitive-cooperative networks while receiving decision recommendations from corresponding agent nodes.

Agent Layer: Nodes provide decision support to enterprises based on algorithmic training in the virtual environment.

To investigate the impact of algorithm-assisted decision-making, two experimental setups are defined:

1. Control Group (No Algorithm Assistance)

Enterprises rely solely on payoff comparisons with same-layer neighbors for strategy updates. Each enterprise adopts the Fermi rule⁰⁰:

In each round, focal enterprise m randomly selects a neighbor n and compares payoffs.

If neighbor n 's payoff exceeds m 's, m probabilistically updates its strategy in the next round. This simulates traditional industrial chains without algorithmic decision support.

2. Experimental Group (Algorithm Assistance)

Decision mechanisms embody human-machine collaboration:

Focal enterprises integrate strategy recommendations from agents with observed payoff performances of neighbors. Enterprises autonomously adjust the integration weights of these two information sources:

If agent strategies demonstrate high payoff consistency in the virtual environment, algorithmic recommendations gain higher weight. Conversely, imitation of real-world neighbors is prioritized.

This design preserves the bounded rationality of organizational decision-making while allowing flexibility for algorithm-driven strategy optimization.

3.5.1. Imitation Probability Function

In the control group, focal enterprises update strategies solely by comparing current payoffs with neighboring enterprises. The Fermi strategy update probability is defined as:

$$W_1 = \frac{1}{1 + e^{-\varphi(U_n - U_m)}} \quad (5)$$

where U_m represents the payoff of the focal enterprise, U_n denotes the neighbor's payoff, and φ reflects the decision-maker's sensitivity to external payoff differences. As φ increases, strategy changes are triggered only when the neighbor's payoff significantly exceeds that of the focal enterprise, effectively lowering the threshold for decision updates. Conversely, smaller φ values imply greater susceptibility to neighbor payoffs, increasing the likelihood of strategy shifts.

In the experimental group, focal enterprises integrate algorithmic strategy recommendations with neighbor payoff observations. The modified Fermi probability is:

$$W_2 = \frac{1}{1 + e^{-\varphi \Delta U}} \quad (6)$$

with the composite payoff difference:

$$\Delta U = \omega_{agent}(U_{agent} - U_m) + \omega_{neighbor}(U_n - U_m) \quad (7)$$

where ΔU is the integrated payoff difference, ω_{agent} is the algorithm recommendation weight, $\omega_{neighbor} = 1 - \omega_{agent}$ is the neighbor imitation weight, and U_{agent} is the agent-predicted payoff.

Dynamic Weight Adjustment Rule:

If the correlation coefficient ρ between agent-predicted and actual payoffs satisfies $\rho > 0.7$, $\rho \in [-1, 1]$:

$$\omega_{agent} \leftarrow \min(\omega_{agent} + 0.1, 0.9) \quad (8)$$

If $\rho < 0.3$:

$$\omega_{agent} \leftarrow \max(\omega_{agent} - 0.1, 0.1) \quad (9)$$

3.6. Observation Metric Design

To comprehensively measure the scope, speed, and stability of decision diffusion, we design the following metrics based on prior research⁰⁰:

Decision Adoption Rate: $\theta = \frac{n_A}{n_R + n_A}$, which directly compares the convergence values of adoption curves.

Innovation Diffusion Speed: The time required for θ to reach 80%, evaluating efficiency advantages across scenarios.

Decision Adoption Rate Variance: Lower variance indicates greater system-wide strategy uniformity and diffusion stability

4. Algorithm Selection and Model Specification

4.1. Algorithm Selection

Current mainstream deep reinforcement learning (DRL) algorithms include Asynchronous Advantage Actor-Critic (A2C), Proximal Policy Optimization (PPO), and Deep Q-Network (DQN)⁰. While DRL algorithms generally deliver impressive results across tasks, specific algorithms exhibit strengths suited to distinct environmental challenges. DQN excels in environments requiring rapid learning and deployment⁰, whereas PPO and A2C may dominate in scenarios demanding deep exploration and comprehensive strategic development⁰⁰. This study adopts PPO for agent training based on the following rationale:

1. Practical Relevance:

Enterprises rarely disclose their proprietary algorithms. However, PPO's widespread adoption in computational science—a field closely aligned with industrial applications—suggests its potential integration into existing or emerging enterprise decision systems.

2. Technical Advantages:

Policy-Based Optimization: As a policy-based DRL algorithm, PPO addresses limitations of prior methods, such as high variance and instability during training⁰.

Robust Policy Updates: PPO's advanced policy optimization capabilities excel in complex state spaces, ensuring stable and efficient policy updates⁰ while balancing simplicity, performance, and implement ability⁰.

Exploration-Exploitation Balance: The discount factor in PPO mitigates trade-offs between exploring novel strategies and exploiting known rewarding behaviors. Compared to A2C, PPO's stable update mechanism is less sensitive to discount factor variations⁰.

Limitations: PPO may exhibit longer execution times and slower convergence compared to alternatives.

4.2. Model Specification

4.2.1. Environment Design

The integration innovation decision problem is formalized as a Markov Decision Process (MDP). Key elements—states, actions, and rewards—are designed as follows (see Figure 3 for the reinforcement learning interaction framework):

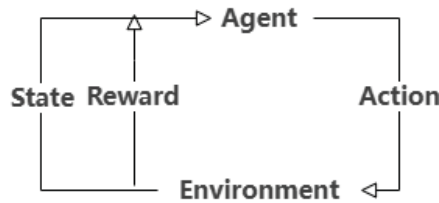


Figure 3 Reinforcement Learning Interaction Process

Goal: The agent aims to learn a policy that maximizes the focal enterprise's cumulative rewards

States: Knowledge spillover, Market competition intensity, Enterprise capability, Upstream/downstream innovation incentives

Reward Mechanism:

+1: Focal enterprise's profit increases or remains positive.

+50: Profit increases/remains positive for ≥ 10 consecutive timesteps.

+100 (with episode termination): Profit increases/remains positive for ≥ 10 consecutive timesteps.

-1: Profit decreases or turns negative.

-100 (with episode termination): Profit decreases/remains negative for ≥ 10 consecutive timesteps.

Actions:

The agent selects one action from three independent groups:

Internal R&D or External Procurement

Inter-enterprise Collaboration or Non-collaboration

Industry-Academia Collaboration or Non-collaboration

A total of $2 \times 2 \times 2 = 8$ action combinations exist, including:

Internal R&D + Inter-enterprise Collaboration + Industry-Academia Collaboration

Internal R&D + Inter-enterprise Collaboration + Non-collaboration

Internal R&D + Inter-enterprise Non-collaboration + Industry-Academia Collaboration

Internal R&D + Inter-enterprise Non-collaboration + Industry-Academia Non-collaboration

External Procurement + Inter-enterprise Collaboration + Industry-Academia Collaboration

External Procurement + Inter-enterprise Collaboration + Industry-Academia Non-collaboration

External Procurement + Inter-enterprise Non-collaboration + Industry-Academia Collaboration

External Procurement + Inter-enterprise Non-collaboration + Industry-Academia Non-collaboration

Following Zhang and Li's approach⁰, selecting Internal R&D, Inter-enterprise Collaboration, and Industry-Academia Collaboration simultaneously is defined as the integration innovation strategy.

4.2.2. Parameter Configuration

The core of the PPO algorithm lies in balancing exploration (acquiring new knowledge) and exploitation (leveraging existing knowledge). This study employs the widely used ϵ -greedy

exploration model. With probability ε , the agent selects actions uniformly at random. With probability $1-\varepsilon$, it chooses the current optimal action.

To adaptively reduce exploration over time, the exploration rate ε decays as:

$$\varepsilon(t) = \varepsilon(1 - \xi)^t \quad (10)$$

Practical Implementation:

A minimum exploration rate $\underline{\varepsilon}$ is enforced to ensure persistent exploratory behavior.

When $\varepsilon < \underline{\varepsilon}$, the agent uses $\underline{\varepsilon}$ as the exploration probability.

State Space:

A four-dimensional continuous observation vector (state=4) captures dynamic interactions: Knowledge spillover, Market competition intensity, Enterprise capability, Upstream/downstream innovation incentives

Action Space:

Discrete action space with action=[2, 2, 2], reflecting binary choices across three dimensions.

Learning Rates:

Actor network: Lr actor=0.0003 (slow updates for stable policy optimization).

Critic network: Lr critic=0.001 (fast convergence for value estimation).

Discount Factor:

$\gamma=0.99$ ensures long-term reward consideration ($\gamma^{200} \approx 0.134$), aligning with PPO's stability in strategic planning.

Policy Optimization Constraint:

PPO-Clip parameter $\varepsilon=0.2$, limiting policy update ratios to $[1-\varepsilon, 1+\varepsilon]$ for trust-region stability.

Table 1: PPO Algorithm Parameter Settings

Category	Parameter Name	Description	Default Value
Environment Parameters	state	State Space Dimension	4
	action	Action Space Dimension	[2, 2, 2]
Network Optimization	Lr actor	Actor Learning Rate	0.0003
	Lr critic	Critic Learning Rate	0.001
Core RL Parameters	γ	Discount Factor	0.99
	ε	Clip Range	0.2
	epochs	Policy Update Iterations per Batch	10
	lam	GAE Parameter	0.95
	entropy	Entropy Coefficient	0.01
Regularization & Stabilization	Max grad norm	Gradient Clipping Threshold	0.5
	value	Value Function Loss Weight	0.5
Training Process	episodes	Total Training Episodes	200
	Max steps	Max Steps per Episode	200
	Update interval	Policy Update Interval	40
	Eval interval	Evaluation Interval	100
	Log interval	Logging Interval	10

4.2.3. Training Process

The PPO algorithm is implemented in Python to train agents for optimizing decision strategies in the customized integration innovation environment. The training protocol is as follows:

Environment Initialization:

A custom integration innovation environment is created with an observation space dimension of 4 and an action space comprising three two-dimensional discrete actions.

Random seeds are set to ensure experiment reproducibility.

Real-time state normalization is applied to stabilize input data distribution.

Training Configuration:

Episodes: 200 episodes, each with a maximum of 200 timesteps.

Policy Updates: Triggered every 40 timesteps using Generalized Advantage Estimation to compute advantage functions.

Optimization: Policy networks are updated via PPO's clip mechanism, while value networks are adjusted to minimize prediction errors against actual returns.

Monitoring and Evaluation:

Training statistics (e.g., rewards, losses) are logged per episode.

Periodic Evaluation: Every 100 episodes, the current policy undergoes exploration-free evaluation to compute average rewards.

Model Checkpointing: If the average reward exceeds historical best performance, model parameters (policy/value network weights, normalization statistics) are saved.

4.2.4. Agent Training Outcome

Post-training, the best-performing model is loaded for final evaluation:

Generalization Metrics: Average reward and standard deviation across multiple test episodes quantify agent robustness and stability.

Action Probability Analysis:

Integration Innovation Strategy (1,1,1): 87.5% adoption probability.

Internal R&D + Non-collaboration + Industry-Academia Collaboration (1,0,1): 12.5% probability.

These results provide foundational insights for algorithm-guided innovation decision diffusion.

5. Simulation Analysis

5.1. Case Study and Parameter Configuration

To enhance the interpretability and reliability of simulation results, parameter settings are grounded in real-world cases and policy frameworks. This study references the Xuzhou Economic Development Zone's construction machinery industrial chain from the *Typical Models of Integration Innovation Between Large Enterprises and SMEs Case Collection* published by the China Academy of Information and Communications Technology.

Case Background:

Industrial Scale: Xuzhou Economic Development Zone hosts China's largest construction machinery R&D and manufacturing cluster, with 71 "Specialized, Refined, Distinctive, Innovative" enterprises, 53

provincial engineering technology research centers, 187 high-tech enterprises, 6 listed companies, Ranked 25th among 230 national-level economic development zones (2024).

Innovation Output: 1,186 new invention patents granted in 2024 (8.1% YoY growth), totaling 5,654 valid patents. Full-chain collaboration spanning "raw materials → components → host equipment → aftermarket."

XCMG Group Leadership:

As the industrial chain leader, XCMG Group established an integration innovation consortium with upstream/downstream partners (e.g., AVIC Liyuan, Xingcheng Special Steel, Shengbang Hydraulics), achieving: 5 consecutive years of revenue exceeding 100 billion CNY. 9,600+ valid patents, 2300+ invention patents. 30% supply chain cost reduction and 20% efficiency gains via smart logistics platforms. 10% production cost reduction through unified procurement and capacity coordination.

(1) Enterprise Parameter Settings

R&D Investment:

2024 estimated integration innovation R&D investment: 7.373 billion CNY

Traditional operation cost (C): 7.373B CNY.

Incremental integration innovation costs (Δc):

$$\Delta c = 33.73 + 20.19 + 0.39 + 2.92 + 5.89 = 6.312 \text{ B CNY}$$

Cost-Efficiency Coefficient (b):

$$b = \frac{(881.83 - 866.40)/866.40}{(938.17 - 928.48)/928.48} = 0.67$$

Synergistic Revenue (F):

Technology Synergy (F_1):

$$F_1 = 11.6 \times 4 \times 2.37 = 10.996 \text{ B CNY}$$

Market Synergy (F_2):

$$F_2 = 880(16\% - 8.87\%)/1 + 8.87\% = 5.763 \text{ B CNY}$$

Cost Synergy (F_3):

$$F_3 = 16.06 \times 4 \times 6\% = 0.385 \text{ B CNY}$$

Total Synergy:

$$F = 109.96 + 57.63 + 3.85 = 171.44 \text{ B CNY}$$

Free-Riding Benefit (H):

$$H = 285.16 \times 6.7\% = 19.11 \text{ B CNY}$$

Competitiveness Loss Parameters:

Following existing literature⁹, the proportion of upstream-downstream collaborative partners lost by the focal enterprise due to a decline in competitiveness, denoted as ξ , is set to 0.2.

Market share reduction: $d = 30.81\% \times 4.11\% = 1.27\%$

Additional integration cost (Δg): 0.385B CNY.

(2) Network Structure Parameters

Initial Neighbor Connections: $t=12$ (based on 13 Chinese enterprises in "2024 Top 50 Global Construction Machinery Manufacturers").

Edge Addition Probability: $\eta=0.1^0$.

Network Nodes: $N=264$ (71 SRDI + 187 high-tech + 6 listed enterprises).

Table 2 Initial Parameter Settings

Parameter	θ	i	T	C	Δc	b	f	F	p	Δg	H	ξ	d	η	t
Assignment	0.1	8.08	0.2	77.04	63.12	0.67	9.61	171.44	0.2	3.85	19.11	0.2	1.27%	0.1	12

Based on the above game - theoretic model and complex network setup, and by incorporating existing relevant studies, this analysis of decision - making diffusion in industrial innovation integration utilizes Python software. To more intuitively show the diffusion of innovation - integration decisions, simulations are run for different scenarios. The proportion of innovation - integration - adopting firms among all firms after diffusion ends is used as a measure of the diffusion degree of such decisions in the industry. Given that exploration rates and strategy - updating sensitivity can lead to randomness in game results, each game round is simulated 20 times. The average of the simulation results is taken to reduce random errors and enhance result robustness.

5.2. The Impact of Algorithm - Assisted Decision - Making on the Diffusion of Innovation - Integration Decisions

Figure 5 illustrates the temporal evolution of the integration innovation adoption rate (θ) under scenarios with and without algorithm-assisted decision-making. After 300 evolutionary rounds, the experimental group (enterprise-agent collaborative evolution) exhibits a significantly higher θ than the control group (traditional enterprise network evolution), demonstrating that embedded algorithm-assisted decision-making substantially promotes the diffusion of integration innovation decisions across industrial chains. Analysis of θ dynamics reveals fluctuations in both groups during the initial diffusion phase. These fluctuations arise from the inclusion of an exploration rate (ϵ) to balance the exploration of new strategies and exploitation of known strategies during agent learning, while ensuring comparability between the two networks by applying the same exploration rate to the control group⁰⁰. This design simulates the bounded rationality of enterprise decisions (e.g., non-fully rational responses to policy changes or market volatility) and mimics exogenous shocks. In later stages, the experimental group's θ gradually converges to 1, entering a steady-state phase, whereas the control group's θ fluctuates around 0.7 with persistently high variance.

As shown in Figures 6 and 7, the experimental group achieves $\theta=0.8$ at Round 33, with variance decreasing from 0.2 to 0, while the control group's θ remains near 0.7 with consistently high variance. These results confirm that algorithm-assisted decision-making not only accelerates the diffusion of integration innovation decisions but also significantly enhances diffusion stability.

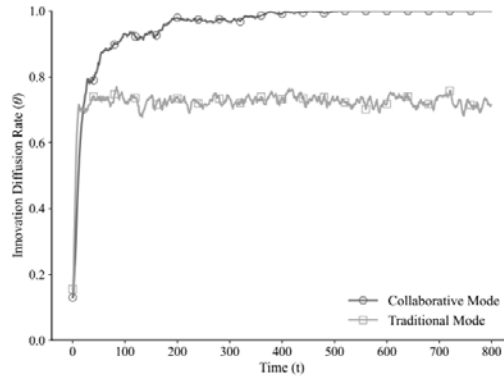


Figure4 shows the impact of algorithm - assisted decision - making on the diffusion of innovation - integration decisions in industrial chains.

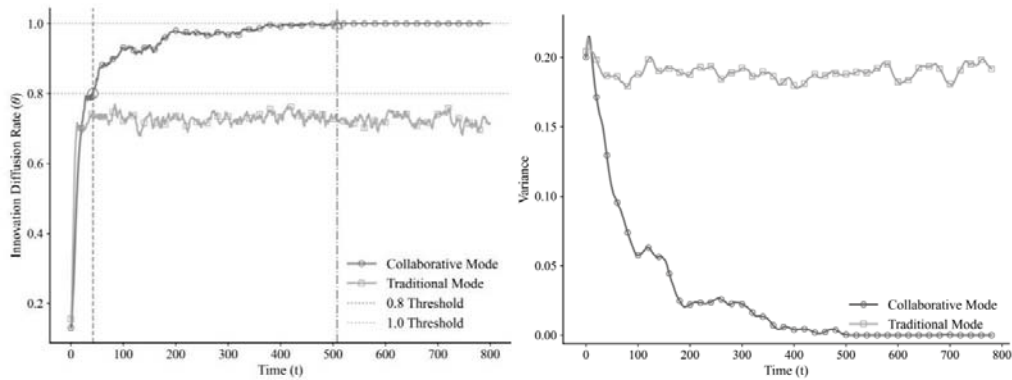


Figure5 analyzes innovation diffusion speed.

Figure 6 examines decision - making stability.

5.2.1. Impact of Initial Algorithm Weights on Integration Innovation Decision Diffusion

As established earlier, algorithm-assisted decision-making significantly promotes the diffusion of integration innovation decisions. However, the initial algorithm weight (ω_{agent}), a critical parameter influencing human-machine collaboration, requires careful balancing: excessively high weights may lead to algorithmic dominance, suppressing organizational autonomy, while overly low weights underutilize algorithmic optimization capabilities. Thus, identifying an optimal weight threshold is essential.

Figure 8 illustrates the impact of varying initial algorithm weights ($\omega_{agent}=0.2,0.5,0.8$) on diffusion dynamics, with other parameters fixed. Key findings include: the adoption rate (θ) converges to 1.0 by Round 300, achieving optimal diffusion and converges to 0.8, demonstrating weaker algorithmic facilitation compared to the traditional network.

Interpretation: In reinforcement learning, algorithms learn rules from historical data, but reliance on past data to predict future behaviors can introduce biases. Additionally, feedback loops—where algorithm-generated data is fed back into training—may amplify errors. To mitigate these risks, an

initial weight of $\omega_{agent} = 0.5$ balances algorithmic guidance and human oversight. Excessive weights risk innovation-inhibiting effects by over-relying on potentially flawed algorithmic recommendations. Practical Implication: Algorithmic assistance should remain subordinate to human decision-making, preserving human experts' authority to review and correct critical decisions.

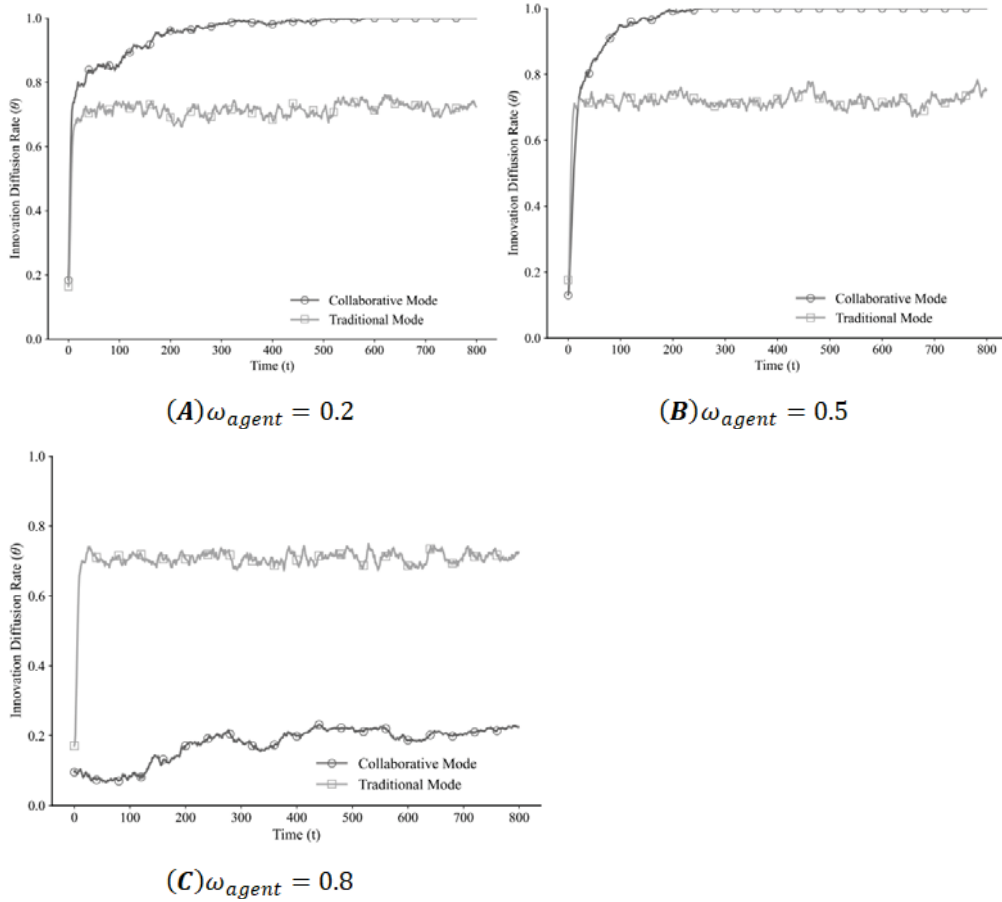


Figure 7 Impact of Initial Algorithm Weight (ω_{agent}) on Integration Innovation Decision Diffusion

5.2.2. Interactive Effects of Initial Algorithm Weights with Other Factors

(1) Upstream-Downstream Integration Innovation Ratio

As shown in Figure 9, the impact of varying upstream-downstream integration innovation ratios (p) on decision diffusion under different initial algorithm weights reveals distinct patterns:

$\omega_{agent} = 0.5$: Optimal diffusion occurs regardless of p , with integration innovation decisions achieving rapid and stable full diffusion ($\theta=1$) across the industrial chain.

$\omega_{agent} = 0.2$: Innovation diffusion levels exhibit a significant negative correlation with increasing p , indicating that low algorithm weights fail to counteract declining collaboration incentives.

$\omega_{agent} = 0.8$: Diffusion stagnates near initial levels irrespective of p , reflecting algorithmic overdominance that suppresses adaptive coordination.

These results demonstrate a dual-threshold effect in embedded algorithm-assisted decision-making. Deviations from the moderate weight range disrupt upstream-downstream synergies:

Underweight Algorithms : Excessive reliance on human experiential judgments limits the algorithm's global optimization potential.

Overweight Algorithms: Algorithmic rigidity stifles decision-makers' innovation agency, hindering adaptive responses to collaboration dynamics.

Conclusion: Balancing algorithm weights within an intermediate range is critical to maintaining effective human-machine collaboration and ensuring stable innovation diffusion.

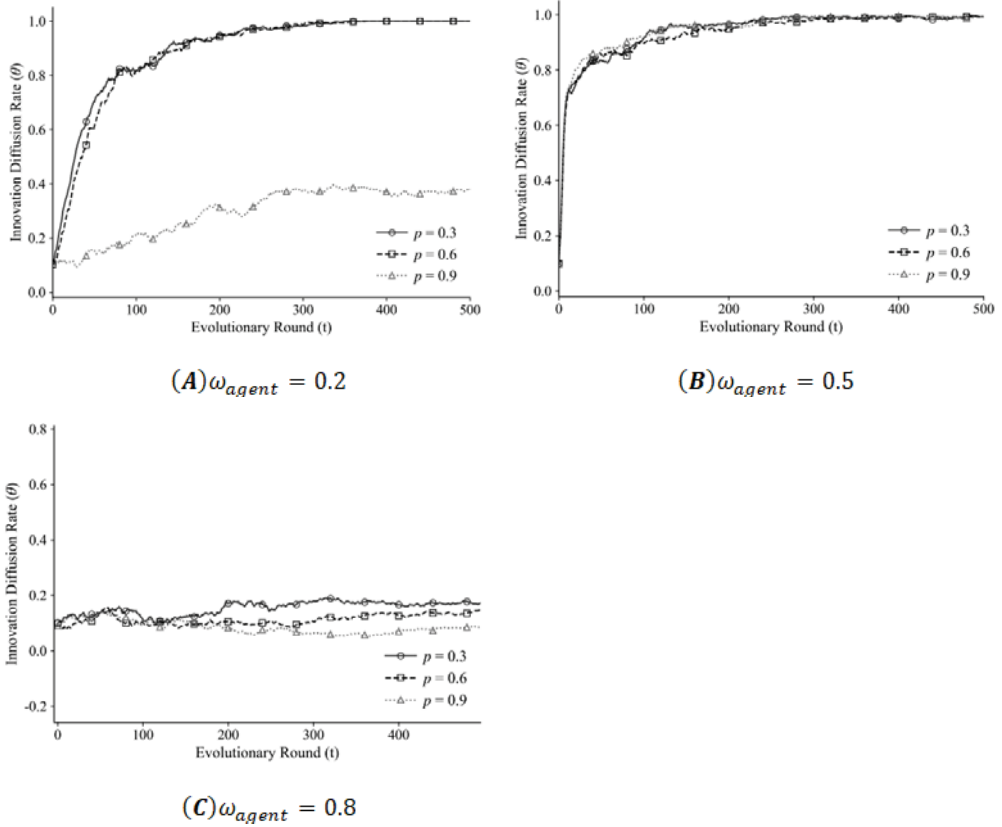


Figure 8. illustrates the impact of upstream-downstream integration innovation ratio (p) under varying ω_{agent}

(2) Impact of Strategy Update Sensitivity (ϕ)

Strategy update sensitivity regulates agents' responsiveness to payoff differences, directly shaping the dynamic evolution of collective decision-making. As shown in Figure 10, variations in ϕ under different algorithm weights ($\omega_{agent} = 0.2, 0.5, 0.8$) exhibit distinct effects on integration innovation diffusion:

For $\omega_{agent} = 0.2$ and 0.5 : Full diffusion ($\theta=1$) is achieved rapidly and stably across all ϕ values.

For $\omega_{agent} = 0.8$: Diffusion stagnates regardless of ϕ , indicating algorithmic overdominance hinders adaptive strategy shifts.

Despite full convergence at moderate weights, diffusion speed varies with ϕ :

$\phi=2.4$: Fastest diffusion (optimal balance between responsiveness and stability).

$\phi = 3.6$: Moderate speed (slightly delayed due to oversensitivity to noise).

$\phi = 1.2$: Slowest diffusion (requires larger payoff gaps to trigger updates).

Practical Implications:

Moderate ϕ (e.g., 2.4): Suitable for scenarios requiring gradual diffusion, such as technology adoption.

High ϕ (e.g., 3.6): Risks overreacting to stochastic fluctuations (e.g., panic selling in financial markets).

Low ϕ (e.g., 1.2): Aligns with conservative decision-making in traditional industries, where large payoff disparities are needed to drive changes.

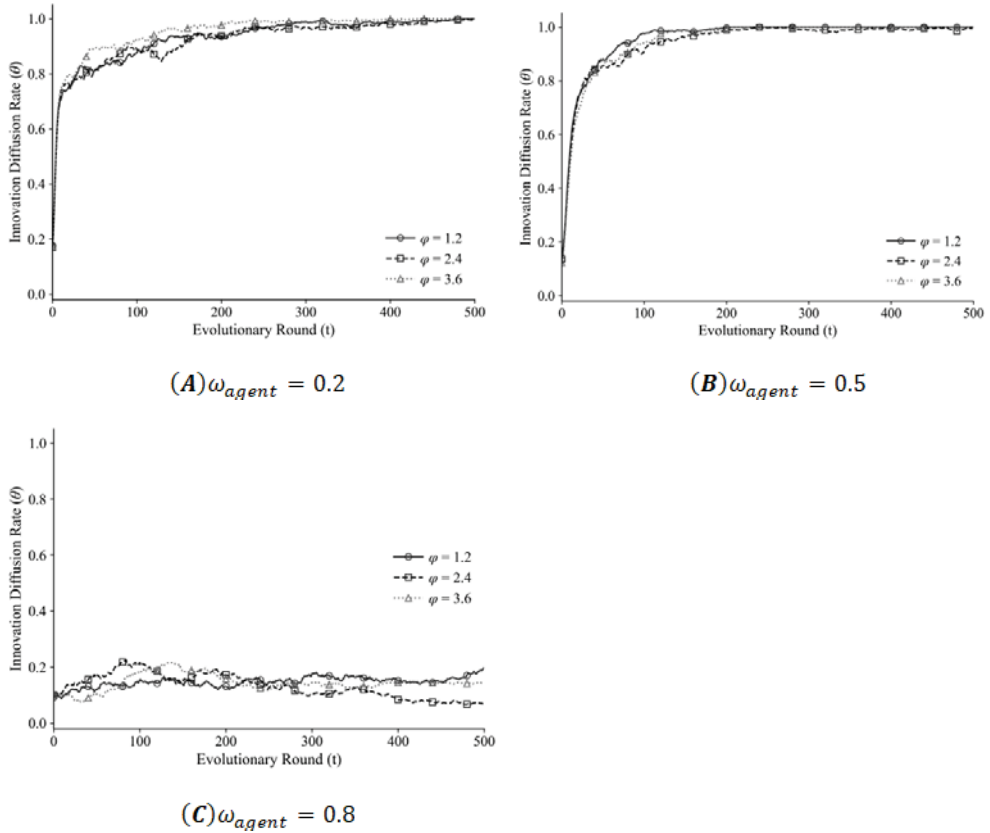


Figure 9. illustrates the impact of strategy update sensitivity (ϕ) under varying ω_{agent}

(3) Impact of Initial Integration Innovation Diffusion Proportion

To examine how variations in the initial integration innovation diffusion proportion (set to 0.2, 0.4, 0.6) interact with algorithm weights (ω_{agent}), simulations are conducted as shown in Figure 11. Key findings include:

Moderate-to-Low Algorithm Weights: Innovation diffusion outperforms scenarios with high weights.
Initial Proportion Effects: Higher initial proportions accelerate diffusion speed. Even under high algorithm weights that generally suppress diffusion, increased initial proportions still enhance innovation adoption. This is attributed to the formation of a high-trust, high-synergy innovation environment at elevated initial proportions, where diffusion and learning mechanisms propagate innovation behaviors across the industrial chain.

Practical Implications: In innovation-intensive sectors with high trust and dense collaborative networks—such as advanced materials, new energy, AI, blockchain, or clusters like Zhong guan cun—algorithm-assisted decision-making exhibits stronger catalytic effects due to inherent synergy and rapid knowledge transfer.

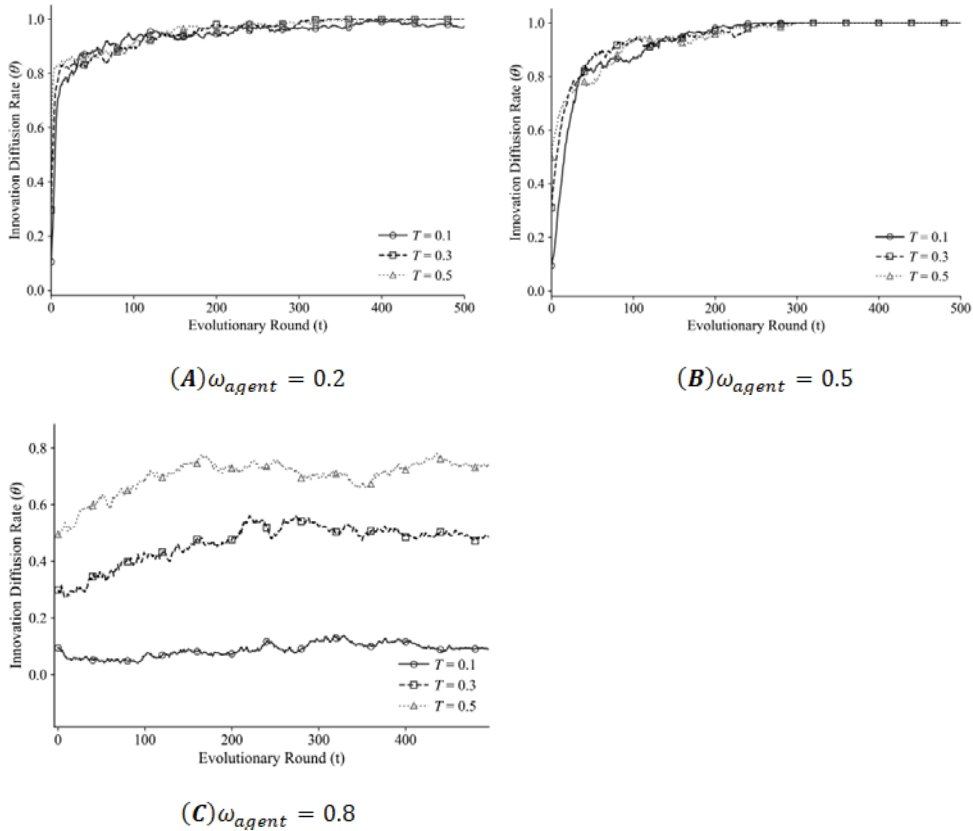


Figure 10. illustrates the impact of initial integration innovation proportions under varying ω_{agent}

(4) Impact of Network Scale (N)

Social network theory posits that enterprises, as economic entities embedded in relational networks, engage in competition and learning through network connections. Network characteristics shape enterprises' decision-making behaviors and influence the diffusion of innovation decisions. Current integration innovation practices vary significantly in the number of competing enterprises across industries. According to CSMAR data, by 2025, the artificial intelligence sector comprises 168,000 enterprises, nonferrous metal smelting and rolling processing has 9,946 enterprises, and the chemical fiber manufacturing industry has only 2,422 enterprises. To investigate the impact of network scale on integration innovation diffusion, simulations are conducted by adjusting N within small-world networks of varying sizes while maintaining other parameters.

A comparison of results from Figures 11(a)(b)(c) reveals that when algorithm weights are at moderate-to-low levels, network scale exerts minimal differential effects on innovation diffusion. Among these, $\omega_{agent} = 0.5$ achieves the fastest diffusion speed, likely because human-machine collaboration balances algorithmic global optimization capabilities with organizational autonomy. However, when the algorithm weight increases to 0.8, the moderating effect of network scale becomes significantly stronger: The larger the network scale is, the more conducive it is to accommodate the diffusion of integration innovation decisions. This may be because overly small networks cause enterprises to be overly influenced by local neighborhood interactions.

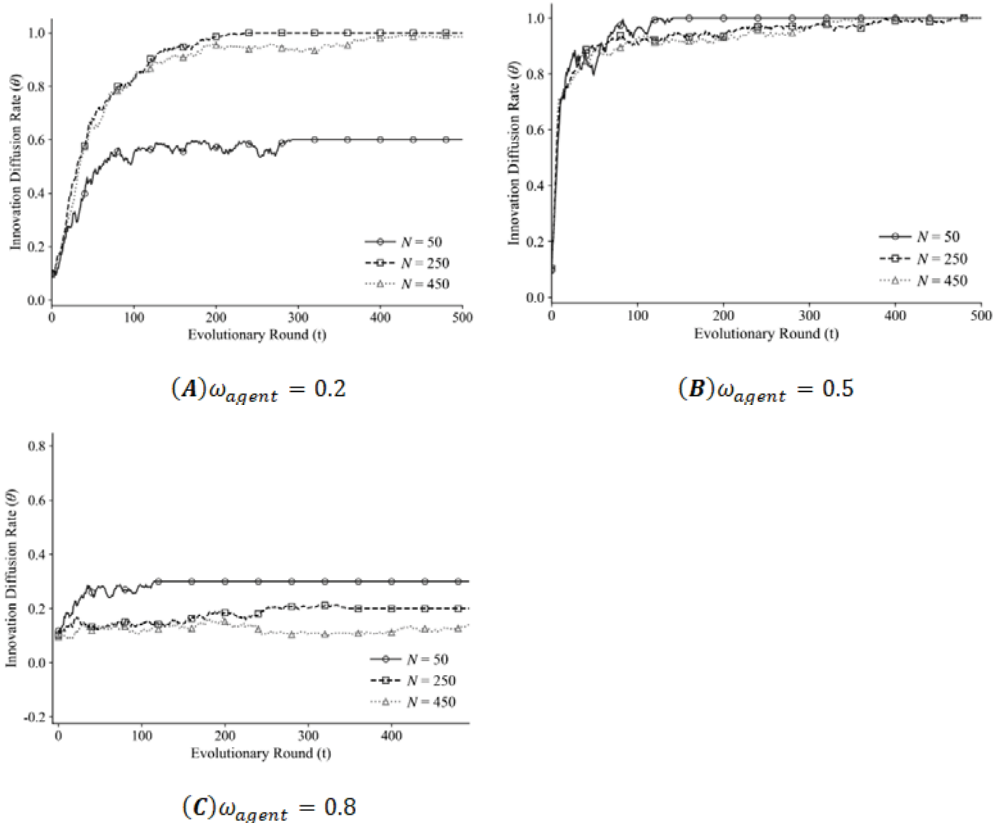


Figure 11. Impact of network scale (N) on integration innovation diffusion under varying ω_{agent}

6. Conclusions and Implications

This study constructs a dual-layer coupled "enterprise-agent" network model, integrating deep reinforcement learning with evolutionary game theory to simulate the impact of algorithm-assisted decision-making on integrated innovation diffusion. The core findings and their implications are synthesized as follows:

First, algorithmic empowerment exhibits a pronounced "dual-threshold" effect on innovation diffusion efficiency. While algorithm-assisted decision-making significantly enhances diffusion stability, its effectiveness is non-linear. Within the high-efficiency synergy zone ($\omega_{agent} \in [0.2, 0.5]$), the system achieves rapid convergence, with the adoption rate (θ) reaching a global optimum of approximately 80%. However, once the weight surpasses the suppression threshold (0.8), the system lapses into innovation stagnation due to algorithmic over-dominance. This confirms that excessive reliance on algorithms can paradoxically stifle the autonomous exploratory drive of firms. Second, algorithmic weight serves as a systemic "buffer" that modulates environmental and structural constraints. Under moderate weighting, the system maintains robust efficiency despite variations in integration ratios or network scales. Crucially, algorithmic intervention facilitates a "decoupling" of decision quality from initial endowments, even manifesting a late-mover advantage where lower initial diffusion proportions stimulate stronger learning momentum. Furthermore, while network expansion (N) inherently introduces "coordination friction," an optimized ω_{agent} effectively mitigates this scale-induced latency, proving the compensatory role of human-AI collaboration in complex industrial chains.

Based on the quantitative evidence of the "dual-threshold" effect and the systemic robustness analysis, this study offers the following strategic implications for policymakers and industrial stakeholders:

First, implement an adaptive, tiered algorithm embedding strategy based on the "Synergy Corridor." Decision-makers should move beyond binary "adoption vs. rejection" logic and instead focus on calibrated algorithmic weighting. This study identifies a high-efficiency synergy zone (representative benchmark $\omega_{agent} \in [0.2, 0.5]$) where human intuition and algorithmic precision generate the strongest complementary effects. For high-tech clusters with mature collaboration, initial weights should be anchored toward the upper bound to maximize diffusion velocity. Conversely, in traditional manufacturing chains where tacit knowledge is paramount, a more conservative initial embedding is recommended, coupled with a dynamic adjustment mechanism that scales with the algorithm's prediction accuracy and the firms' trust levels. Second, establish a "Guardrail Mechanism" to prevent algorithmic over-dominance and rigidity. The identification of a suppression threshold provides a critical quantitative heuristic for risk management. To prevent the "low-level trap" where excessive algorithmic reliance stifles organizational exploration, regulators should mandate transparency in the parameterization of intelligent agents. Governance should shift from "black-box" oversight to parameter-based regulation, specifically monitoring micro-behavioral indicators such as the exploration rate and strategy sensitivity. By maintaining these parameters within stable bounds, firms can ensure that algorithmic assistance serves as a catalyst for innovation rather than a constraint on autonomous strategic diversity. Third, leverage the "Decoupling Effect" to foster inclusive innovation

ecosystems. A pivotal finding of this research is that optimized algorithmic intervention can decouple innovation quality from a firm's initial resource endowments, creating a "late-mover advantage." Policymakers should utilize this mechanism to support small-to-medium enterprises (SMEs) or emerging industrial chains that suffer from low initial integration. By providing subsidized access to algorithm-assisted decision-making platforms, governments can help these "late-movers" overcome traditional innovation barriers. Furthermore, for large-scale networks prone to coordination friction, targeted support should focus on reducing information sparsity, allowing the human-AI collaborative framework to maintain high adoption rates even under increased structural complexity.

This study acknowledges certain limitations. The current model does not incorporate stochastic external shocks (e.g., disruptive technological breakthroughs or sudden geopolitical shifts) or potential ethical risks like algorithmic bias in multi-agent settings. Future research will explore the co-evolutionary dynamics of multi-agent systems under external volatility to further advance the theoretical framework of algorithm-enabled innovation ecosystems.

7. REFERENCES

- Akossou, A.Y.J., 2005, *Impact de la structure des données sur les prédictions en régression linéaire multiple*. PhD Thesis, Fac. Univ. Sci. Agron., Gembloux, Belgium, 215 p.
- Bendel, R.B., Afifi, A.A., 1977, Comparison of stopping rules in forward stepwise regression. *J. Amer. Stat. Assoc.* **72**, 46-53.
- Copas, R.D., 1983, Regression, prediction and shrinkage. *J. R. Stat. Soc.* **B 45**, 311-354.
- Dempster, A.P., Schatzoff, M., Wermuth N., 1977, A simulation study of alternatives to ordinary least squares. *J. Amer. Stat. Assoc.* **72**, 77-106.
- Meg, B. C., 1988, Determining the optimum number of predictors for linear prediction equation. *Amer. Météo. Soc.* **116**, 1623-1640.
- Miller, A.J., 1990, *Subset selection in regression*. Monographs on statistics and applied probability 40. Chapman and Hall.
- Palm, R., De Bast, A., Lahlou, M., 1991, Comparaison des modèles agrométéorologiques de type statistique empirique construits à partir de différents ensembles de variables météorologiques. *Bull. Rech. Agron. Gembloux* **26**, 71-89.
- Roecker, E.B., 1991, Prediction error and its estimation for subset-selected models. *Technometrics* **33**, 459-468.
- WANG X W, LI Y, CHEN N X. Digital transformation, supply chain spillover, and corporate innovation integration [J/OL]. *East China Economic Management*, 2025,39(04):43-55. DOI:10.19629/j.cnki.34-1014/f.240906015.
- LI M, XIA S Y. Heterogeneous executive teams, industry background, and a meta-analysis of corporate innovation performance [J]. *Journal of Anhui University (Philosophy and Social Sciences Edition)*, 2022,46(04):147-156. DOI:10.13796/j.cnki.1001-5019.2022.04.017.
- LI W J, ZHENG M M. Fundamental innovation or strategic innovation? —The impact of macro industrial policies on micro corporate innovation [J]. *Economic Research*, 2016,51(04):60-73.

XIONG X, CHEN R X, ZHANG W, et al. Research on the architecture of digital intelligence decision-making theater for financial scenarios [J]. *Frontiers of Engineering Management and Technology*, 2023,42(04):35-42.

XU L, LI Z, GUO X L. Artificial intelligence algorithmic decision-making and corporate R & D "conspiracy" [J]. *China Soft Science*, 2024,(06):214-224.

Li Y. Deep reinforcement learning: An overview[J]. *arXiv preprint arXiv:1701.07274*, 2017.

ZHANG Y L, LI L L, DING Z B. Artificial intelligence decision-making in organizational management: Review and prospects [J]. *Foreign Economics and Management*, 2024,46(10):18-38. DOI:10.16538/j.cnki.fem.20240712.101.

ZHANG Z X, HUA Z S, XIE X Y. Research status and future directions of human-machine collaboration in the digital intelligence era [J]. *Journal of Management Engineering*, 2024,38(01):1-13. DOI:10.13587/j.cnki.jieem.2024.01.001.

WANG Y, WANG C L, QIU G D. Group intelligent learning-based decision-making: The evolution of decision-making paradigms empowered by "Big Data + AI" [J]. *China Soft Science*, 2024,(12):35-50.

GU W, LIU Y J, LU S F, et al. Artificial intelligence-driven decision-making in management: Applications, perceptions, and biases [J/OL]. *Chinese Journal of Management Science*, 1-14 [2025-03-20]. <https://doi.org/10.16381/j.cnki.issn1003-207x.2023.1896>.

SUN L W, LI S S. Attribution and governance of risks in the industrial innovation ecosystem of artificial intelligence technologies [J]. *Science and Technology Progress and Countermeasures*, 2021,38(17):69-78.

ZHU Y N, LU Y X. Practices and challenges of AI ethical governance in China ——A perspective of three overlapping phases [J]. *China Forum on Science and Technology*, 2025,(01):148-153+178. DOI:10.13580/j.cnki.fstc.2025.01.021.

LI Z, LUO X. Restructuring human resources in the smart revolution: Challenges, impact-s, and governance of DeepSeek and Manus generative AI in the HR market [J/OL]. *Journal of Chongqing University (Social Sciences Edition)*, 1-13 [2025-03-20]. <http://kns.cnki.net/kcms/detail/50.1023.C.20250311.1349.007.html>.

ZHOU W. Antitrust regulation of personalized pricing algorithms in the AI era [J]. *Wuhan University Journal (Philosophy and Social Sciences Edition)*, 2021,74(01):108-120. DOI:10.14086/j.cnki.wujss.2021.01.010.

LEI H N. Competitive risks, regulatory logic, and antitrust responses of generative AI [J]. *China Forum on Science and Technology*, 2024,(07):64-74. DOI:10.13580/j.cnki.fstc.2024.07.003.

ZHOU Y L, ZENG L. Emergence, connotation, and extension of integrated innovation [J/OL]. *Scientific Studies*, 1-20 [2025-04-15]. <https://doi.org/10.16192/j.cnki.1003-2053.20240816.004>.

CHEN J, YANG Y J, LIU C. Exploration of China's technological innovation paradigm by 2035: Integrated innovation [J]. *China Forum on Science and Technology*, 2020,(10):7-10. DOI:10.13580/j.cnki.fstc.2020.10.005.

CHEN J, YANG Y J, LIU C. Theoretical connotations and practical exploration of integrated innovation [J]. *China Society for Management Science*, 2022. DOI:10.19345/j.cxkj.1671-0037.2020.02.001.

Sturgeon, T. J. (2002). Modular production networks: A new American model of industrial organization. *Industrial and Corporate Change*. <https://doi.org/10.1093/icc/11.3.451>

Masucci, Monica, Stefano Brusoni, and Carmelo Cennamo. "Removing bottlenecks in business ecosystems: The strategic role of outbound open innovation." *Research Policy* 49.1 (2020): 103823. <https://doi.org/10.1016/j.respol.2019.103823>

Gnyawali, D. R., & Park, B. J. (2011). Coopetition between giants: Collaboration with competitors for technological innovation. *Research Policy*. <https://doi.org/10.1016/j.respol.2011.01.009>

LI X J. Digital economy driving innovation and development of large, medium and small enterprises: Mechanisms, models and paths [J]. *Contemporary Finance and Economics*, 2023,(04):3-14. DOI:10.13676/j.cnki.cn36-1030/f.2023.04.012.

LI X X, ZHANG S H. How does digital transformation affect corporate innovation integration? [J]. *Economic Management*, 2023,45(04):41-63. DOI:10.19616/j.cnki.bmj.2023.04.003.

ZENG J L, ZHOU J. Does intelligent manufacturing promote corporate innovation integration? [J]. *Journal of Capital University of Economics and Business*, 2024,26(06):79-93. DOI:10.13504/j.cnki.issn1008-2700.2024.06.006.

CHEN B P, WU X L. The impact of government-enterprise value co-creation on innovation integration in tech-based enterprises ——A case study based on innovation and entrepreneurship competitions [J]. *Science and Technology Management Research*, 2023,43(24):171-179.

SUTTON R S, BARTO A G. Reinforcement learning: An introduction [M]. [s.l.]: MIT Press, 2018. <https://doi.org/10.1137/21N975254>

Mnih, V., et al. (2015). Human-level control through deep reinforcement learning. *Nature*, 518(7540), 529-533. <https://doi.org/10.1038/nature14236>

Silver, D., Schrittwieser, J., Simonyan, K., et al. (2017). Mastering the game of Go without human knowledge. *Nature*, 550(7676), 354-359. <https://doi.org/10.1038/nature24270>

Bharathi B, Shareefa P, Maheshwari P U, et al. Exploring the Possibilities: Reinforcement Learning and AI Innovation [J]. *Int. J. Adv. Res. Sci., Commun. Technol*, 2023: 356-362. DOI : 10.48175/IJARSCT-8837

Batsis, Alex, and Spyridon Samothrakis. "Contextual reinforcement learning for supply chain management." *Expert Systems with Applications* 249 (2024): 123541. <https://doi.org/10.1016/j.eswa.2024.123541>

Haarnoja T, Zhou A, Abbeel P, et al. Soft actor-critic: Off-policy maximum entropy deep reinforcement learning with a stochastic actor [C]//International conference on machine learning. Pmlr, 2018: 1861-1870.

LIU W L, LI X J. Industrial robots and the increase in domestic value-added export ratios of enterprises——An empirical study based on human-machine collaboration [J]. *Management World*, 2024,40(3):45-60. DOI:10.14116/j.nkes.2024.12.009.

MI L Y, SHANG C Y. Research on the coordination mechanism of fairness concern and innovation input in closed-loop supply chains with algorithm embedding [J]. *Systems Engineering Theory and Practice*, 2022,42(11):2987-3000.

YANG J R, CHU D F, LU L P, et al. A review of research on human-machine shared control in intelligent vehicles [J]. *Journal of Mechanical Engineering*, 2022,58(18):31-55.

LIU T, YOU H L. Optimization of transportation scheduling based on human-machine collaborative decision-making [J]. *Transportation Systems Engineering and Information*, 2024,24(02):136-148. DOI:10.16097/j.cnki.1009-6744.2024.02.014.

WANG Y W, GAO P. Development concerns and improvement paths of judicial applications of AI algorithms [J]. *Academic Exchange*, 2023,(11):69-83.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684), 440–442. <https://doi.org/10.1038/30918>

- Cowan, R., & Jonard, N. (2004). Network structure and the diffusion of knowledge. *Journal of Economic Dynamics and Control*, 28(8), 1557–1575. <https://doi.org/10.1016/j.jedc.2003.04.002>
- Giuliani, E. (2007). The selective nature of knowledge networks in clusters: Evidence from the wine industry. *Journal of Economic Geography*, 7(2), 139–168. <https://doi.org/10.1093/jeg/lbl014>
- Ernst, D. (2002). Global production networks and the changing geography of innovation systems: Implications for developing countries. *Economics of Innovation and New Technology*, 11(6), 497–523. <https://doi.org/10.1080/10438590214341>
- Funk, R. J., & Owen-Smith, J. (2017). A dynamic network measure of technological change. *Management Science*, 63(3), 791–817. <https://doi.org/10.1287/mnsc.2015.2366>
- HE Y B, LIN X H. Research on the diffusion of disruptive technologies based on complex network evolutionary games [J]. *Soft Science*, 2024,38(06):28-36. DOI:10.13956/j.ss.1001-8409.2024.06.05.
- Rahwan,I., Cebrian, M., Obradovich, N. *et al.* Machine behaviour. *Nature* 568, 477–486 (2019). <https://doi.org/10.1038/s41586-019-1138-y>
- Centola, D.(2018). How behavior spreads: The science of complex contagions. Princeton University Press.
- SUN Q, TIAN B F. The impact of integrated innovation on embedding in digital global value chains—An analysis based on new productive forces [J]. *Business Research*, 2024,(05):10-18. DOI:10.13902/j.cnki.syyj.2024.05.003.
- Bertello A,De Bernardi P,Santoro G,et al.Unveiling the microfoundations of multiplex boundary work for collaborative innovation[J].*Journal of Business Research*, 2022,139:1424-1434. <https://doi.org/10.1016/j.jbusres.2021.10.039>
- ZHANG H Y, WAN D F, YUAN A F, et al. Analysis of corporate R & D behavior based on innovation spillover and uncertainty [J]. *Journal of Management Engineering*, 2005,(01):60-64.
- LIU T, CHEN Z, CHEN X R. An overview of complex network theory and its applications [J]. *Systems Engineering*, 2005,(06):1-7.
- LIU F N, HU H Q, QIN X Y. Research on the diffusion of resource sharing behaviors in entrepreneurial teams based on complex network evolutionary games [J]. *Soft Science*, 2025,39(01):54-60+84. DOI:10.13956/j.ss.1001-8409.2025.01.07.
- SUN Y W, WEI Y P. Research on knowledge diffusion in high-tech enterprise alliances ——A perspective of small-world networks [J]. *Journal of Management Science*, 2011,14(12):17-26.
- YU K, RONG L L, GUO W G, et al. Research on public opinion propagation models based on online and offline networks [J]. *Management Review*, 2015,27(08):200-212. DOI:10.14120/j.cnki.cn11-5057/f.2015.08.018.
- XIE F J, CUI W T, WU X P. Modeling and structural analysis of competition networks in the express delivery industry [J]. *Systems Engineering*, 2017,35(07):101-106.
- LIU J, GUO Y F. Analysis of public opinion evolution mechanisms based on coupled networks [J]. *Intelligence Magazine*, 2025,44(02):136-143+182.
- Traulsen A, Nowak M A, Pacheco J M. Stochastic dynamics of invasion and fixation[J]. *Physical Review E*, 2006, 74(1): 011909. DOI: <https://doi.org/10.1103/PhysRevE.74.011909>
- SZABÓ G, TOKE C. Evolutionary prisoner's dilemma on a square lattice [J]. *Physical Review E*, 1998, 58:69-73. DOI: <https://doi.org/10.1103/PhysRevE.58.69>
- XUE X F, FANG X Y, WANG Y W, et al. Do rumors end with the wise or the truth? ——A simulation study on rumor control based on a double-layer coupled network [J]. *Intelligence Science*, 2024,42(06):176-185. DOI:10.13833/j.issn.1007-7634.2024.06.019.

SUN B, SU X, XU X F. Research on technology diffusion based on new technology adoption decision-making games [J]. Industrial Technology Economy, 2019,38(06):10-19.

Mousavi S S, Schukat M, Howley E. Deep reinforcement learning: an overview[C]//Proceedings of SAI Intelligent Systems Conference (IntelliSys) 2016: Volume 2. Springer International Publishing, 2018: 426-440.

Tan F, Yan P, Guan X. Deep reinforcement learning: From Q-learning to deep Q-learning[C]//Neural Information Processing: 24th International Conference, ICONIP 2017, Guangzhou, China, November 14–18, 2017, Proceedings, Part IV 24. Springer International Publishing, 2017: 475-483.

Schulman J, Wolski F, Dhariwal P, et al. Proximal policy optimization algorithms[J]. arXiv preprint arXiv:1707.06347, 2017. <https://doi.org/10.48550/arXiv.1707.06347>

Huang S, Kanervisto A, Raffin A, et al. A2C is a special case of PPO[J]. arXiv preprint arXiv:2205.09123, 2022. <https://doi.org/10.48550/arXiv.2205.09123>

A. d. Rio, D. Jimenez and J. Serrano, "Comparative Analysis of A3C and PPO Algorithms in Reinforcement Learning: A Survey on General Environments," in IEEE Access, vol. 12, pp. 146795-146806, 2024 DOI: 10.1109/ACCESS.2024.3472473

Yu, Chao, et al. "The surprising effectiveness of ppo in cooperative multi-agent games." Advances in neural information processing systems 35 (2022): 24611-24624.

De La Fuente, Neil, and Daniel A. Vidal Guerra. "A comparative study of deep reinforcement learning models: Dqn vs ppo vs a2c." arxiv preprint arxiv:2407.14151 (2024). <https://doi.org/10.48550/arXiv.2407.14151>

Del Rio, Alberto, David Jimenez, and Javier Serrano. "Comparative Analysis of A3C and PPO Algorithms in Reinforcement Learning: A Survey on General Environments." IEEE Access (2024). DOI: 10.1109/ACCESS.2024.3472473

ZHANG S H, LI X X. The impact of digital transformation on enterprise innovation integration——An analysis based on the boundary perspective [J]. Soft Science, 2023,37(12):86-92. DOI:10.13956/j.ss.1001-8409.2023.12.12.

LI D, MEI X D, YI X T. The impact of industrial chain linkage effects on the diffusion of digital decision-making in small and medium-sized manufacturing enterprises ——Based on complex network game models [J]. Operations and Management, 2024,33(11):218-225.

Friedman D. Evolutionary games in economics[J]. Econometrica: journal of the econometric society, 1991: 637-666. <https://doi.org/10.2307/2938222>

Nowak M A. Evolutionary dynamics: exploring the equations of life[M]. Harvard university press, 2006.