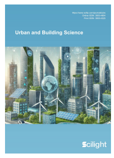
Urban and Building Science
https://www.sciltp.com/journals/ubs



Review

A Conceptual Framework for Performance Evaluation and Classification of Multi-Agent Hybrid Game Problems Based on Reinforcement Learning: Impact Factors and Evaluation Metrics

Lihua Zhang [1,*], Xiaolei Yuan [2,*], Wenqi Jia [3] and Angelica Elhokayem [4]

[1] Tsinghua Shenzhen International Graduate School, Tsinghua University, Nanshan, Shenzhen 518071, China
[2] Department of Mechanical Engineering, Aalto university, FI-00076 Aalto, Finland
[3] Department of Computer Science and Engineering, University of Texas at Arlington, Arlington, TX 76019, USA
[4] Faculty of Engineering, Free University of Bozen-Bolzano, 39100 Bolzano-Bozen, Italy
* Correspondence: lihuazhang@sz.tsinghua.edu.cn (L.Z.); xiaolei.yuan@aalto.fi (X.Y.)

How To Cite: Zhang, L.; Yuan, X.; Jia, W.; et al. A Conceptual Framework for Performance Evaluation and Classification of Multi-Agent Hybrid Game Problems Based on Reinforcement Learning: Impact Factors and Evaluation Metrics. *Urban and Building Science* **2025**, *1*(1), 4. https://doi.org/10.53941/ubs.2025.100004

Received: 13 July 2025
Revised: 4 August 2025
Accepted: 18 August 2025
Published: 25 August 2025

Abstract: Multi-agent reinforcement learning (MARL) method for hybrid game is characterized by the coexistence of cooperativeness and competitiveness. This method has become more applicable for their complexity and wide applications in fields of architectural optimization, urban resource scheduling, and intelligent environmental control. However, the absence of a systematic framework for MARL algorithms' assessment limits practical deployment in such fields. To address this concern, this paper intends to conduct a critical literature review to delve into the decisive factors of the scenarios and further develop conceptual framework for the subdivision classification, algorithm design and evaluation of hybrid games. The core research goal is to expound the differentiated classification system of hybrid game and the key determinants that significantly affect the performance of multi-agent reinforcement learning algorithms. Therefore, this paper proposes an evaluation conceptual framework for classification of hybrid games based on objectives and application scenarios, providing a reference for developers to select and design appropriate MARL algorithms and evaluate their performance according to different types of cooperative games.

Keywords: multi-agent systems; hybrid game problems; reinforcement learning; architectural optimization

1. Introduction

Multi-agent reinforcement learning (MARL) has emerged as a pivotal field, enabling systems and artificial intelligence to address complex decision-making scenarios where multiple intelligent entities interact within shared environments [1,2]. Among the diverse landscapes of MARL, hybrid game problems have gained significant attention. They are characterized by the simultaneous presence of cooperative and competitive dynamics. This growing interest is largely driven by their relevance to real-world applications, including architectural optimization, urban resource scheduling, and intelligent environmental control.

These problems pose unique challenges in such fileds, as they require agents to balance collaborative strategies with competitive behaviors to achieve optimal performance. Despite the growing interests in practical applications of the MARL algorithms, there still is a shortage of a applied conceptual framework for evaluating and classifying multi-agent hybrid game problems, which has hindered the standardization and deployment of MARL algorithms in the real world [3,4].

1.1. Research on Classification Frameworks in Multi-Agent Reinforcement Learning

Multi-agent reinforcement learning (MARL) methods have witnessed substantial progress in addressing fundamental challenges of multi-agent systems. Du et al. established a comprehensive structural taxonomy of MARL algorithms, with particular emphasis on centralized training with decentralized execution (CTDE) frameworks such as Multi-Agent Deep Deterministic Policy Gradient (MADDPG) and Deep Multi-Agent Reinforcement Learning QMIX for cooperative scenarios [1]. Wong et al. proposed interdisciplinary methodologies to tackle critical issues including non-stationarity and partial observability [2]. Nguyen et al. conducted a systematic analysis of technical solutions, encompassing experience replay mechanisms and deep recurrent neural networks for practical implementations [3]. Gronauer and Diepold investigated emergent behavioral patterns and communication protocols in hybrid environments [4].

Oroojlooy and Hajinezhad developed a systematic classification of cooperative MARL approaches into five distinct algorithmic categories, elucidating the pivotal role of CTDE in agent coordination [1]. Their comprehensive review underscored successful applications in domains ranging from traffic management to robotic systems, while simultaneously identifying persistent limitations in scalability and evaluation methodologies. Promising research directions include the development of hybrid games and hierarchical MARL paradigms, with particular emphasis on the need for advancements in off-policy learning, safety-constrained systems, and heterogeneous MARL configurations.

However, the current literature exhibits a notable deficiency in the systematic classification of mixed cooperative-competitive scenarios, rigorous analysis of key influencing factors, and a unified conceptual framework for algorithmic performance evaluation.

1.2. Research on Classification Frameworks in Multi-Agent Systems

The previous section introduced the relevant review work in multi-agent reinforcement learning, and this section presents the related work in multi-agent systems. Li et al. conducted a rigorous examination of consensus algorithms and control strategies within MAS frameworks, meticulously analyzing how dynamic models and network architectures influence convergence properties [5]. Their systematic review encompassed critical consensus paradigms including leader-follower dynamics, finite-time convergence, and cluster consensus mechanisms. The authors particularly elucidated the dichotomy between robust and adaptive control approaches for disturbance mitigation, while underscoring practical implementations such as formation control and wireless sensor networks as pivotal application domains. Their analysis revealed the pressing need for comprehensive evaluation frameworks to assess consensus algorithms across diverse network topologies and dynamic environments.

Zhang et al. conducted an exhaustive survey on the dual challenges of physical safety and cybersecurity in Multi-Agent Systems (MAS), covering fault estimation, detection, diagnosis, fault-tolerant control, cyberattack identification, and secure control mechanisms [4]. Their taxonomy categorized physical threats (e.g., actuator and sensor faults) and cyber threats (e.g., denial-of-service and deception attacks), with in-depth analyses of both homogeneous and heterogeneous MAS configurations. The study highlighted the effectiveness of model-based approaches (including sliding mode observers and unknown input observers) combined with adaptive control strategies in mitigating system vulnerabilities, while emphasizing the benefits of distributed frameworks for enhancing robustness. It identified critical limitations in prior research: existing studies either focused narrowly on access control and reputation mechanisms or restricted fault-tolerant solutions to topology reconfiguration, whereas this work pioneered an integrated physical-cyber security perspective.

However, these otherwise comprehensive investigations remain limited in their treatment of MARL approaches, representing significance in the current MAS research landscape [6–9]. The absence of systematic analysis regarding reinforcement learning paradigms in multi-agent contexts suggests an important direction for future scholarly inquiry.

1.3. Research on Classification Frameworks in Reinforcement Learning

The above two sections respectively introduced the related work in multi-agent reinforcement learning and multi-agent systems, and this section presents the related work in reinforcement learning.

Shakya et al. systematically synthesized RL's theoretical underpinnings, algorithmic architectures, and application scenarios through comprehensive bibliometric analysis [10]. Their work not only elucidated the applicability of model-free RL and function-approximated DRL in stochastic environments but also pioneered a taxonomy that systematizes model-based and multi-agent extensions. Notably, their methodological framework provides a referential foundation for emerging researchers.

From a methodological perspective, Luo's team made groundbreaking progress in model-based reinforcement learning (MBRL) [11]. Their work innovatively integrated Lipschitz-constrained error correction mechanisms with hybrid architectures combining dynamic programming and gradient propagation. This methodology addresses the general challenges inherent in practical deployments while maintaining computational tractability.

Empirical investigations by Vázquez's group yielded critical insights for demand-response systems [12]. Their empirical analysis revealed that while existing pricing mechanisms facilitate load shifting, they exhibit limited efficacy in aggregate demand reduction—an observation underscoring the imperative of integrating behavioral modeling and self-learning mechanisms to enhance grid flexibility. Crucially, this work underscores the necessity for adaptive learning paradigms capable of synthesizing human behavioral patterns.

Alongside the application of professional scenarios in the urban construction field, these studies offer comprehensive overviews of reinforcement learning methodologies. However, they collectively exhibit a notable lack of rigorous analysis on multi-agent systems. The existing literature would stand to gain from more systematic investigations into distributed decision-making frameworks and emergent behaviors within multi-agent reinforcement learning environments.

1.4. Current Studies in Industrial Applications

The previous section introduced the relevant work in the filed of theoretical optimization, and this section presents the related work in the field of industrial applications. Noaeen et al. investigated a holistic reinforcement learning (RL) application per urban traffic signal control, examining how RL algorithms achieve optimal signal timing from intersection to intersection to improve traffic flow efficiency [13]. Their investigation encompassed critical dimensions compared multi-agent with single-agent system architectures, implemented deep reinforcement learning methodologies, and the formulation of state representations, action spaces, and reward functions tailored to traffic management scenarios. Notably, the authors identified significant gaps in the real-world implementation of RL-based control systems, emphasizing the necessity for further research into large-scale network coordination and adaptive strategies for complex environmental factors such as emergency vehicle prioritization and weather-induced traffic variations.

Similarly, Zhang et al. analyzed the transformative deep learning's potential in urban water resource management, demonstrating how advanced techniques—long short-term memory (LSTM) models, including convolutional neural networks (CNNs), and generative adversarial networks (GANs)—can significantly enhance operational efficiency in key areas such as water demand prediction, pipeline leak detection, sewer infrastructure assessment, and flood risk modeling [14]. However, their review revealed a persistent disconnecting between theoretical advancements and practical deployment, with most studies relying on simulated or limited pilot datasets. The authors underscored the urgent need to address fundamental challenges including data privacy concerns, model interpretability, and seamless integration with digital twin technologies to facilitate real-world adoption.

While these studies provide valuable insights into industrial applications, they collectively highlight a critical research gap: the absence of a unified conceptual framework for evaluating multi-agent systems in practical settings.

1.5. Original Contribution

These prior studies have significantly advanced our understanding of MARL in hybrid games, but they primarily focus on algorithmic challenges and solution methodologies rather than establishing a structured framework for problem classification and performance evaluation. Existing studies have not provided a systematic approach for decomposing hybrid game problems according to the interaction between cooperative and competitive dynamics. As a result, the development of targeted evaluation metrics that reflect the unique characteristics of these environments remains limited.

This paper aims to bridge these gaps by proposing a conceptual framework for the performance evaluation and classification of multi-agent hybrid game problems. The core originality is mentioned as below: (1) A Differentiated Classification System. This study develops a comprehensive taxonomy of hybrid game problems based on the degree of cooperation, information structure, and reward dynamics, providing a clear framework for problem categorization; (2) Impact Factor Analysis. This study identifies and analyzes the key determinants that significantly affect MARL algorithm performance in hybrid games, including cooperation-competition balance, observability levels, and agent heterogeneity; (3) Metrics Framework for Assessment. The research proposes a set of standardized evaluation metrics that integrate both cooperative and competitive performance indicators, enabling systematic assessment of algorithms across diverse hybrid game scenarios; (4) Practical Guidance. The framework offers developers a reference for selecting and designing appropriate MARL algorithms, as well as evaluating their performance based on the classified problem types and identified impact factors.

By addressing these gaps, this study aims to advance the field of MARL for hybrid games by providing a structured foundation for problem analysis, algorithm design, and performance evaluation, ultimately facilitating broader adoption and practical deployment in real-world applications.

1.6. Paper Arrangement

This paper comprises five main sections, each covering critical aspects of multi-agent hybrid game problems and their evaluation frameworks. Section 2 classifies such games based on cooperation-competition dynamics and illustrates with typical experimental scenarios. Section 3 details reinforcement learning algorithms—including fundamental, value-based, policy-based, and hybrid methods—along with their strengths and limitations. Section 4 presents comprehensive metrics for evaluating algorithm performance, cooperation-competition dynamics, and feasibility, while discussing influencing factors and robust evaluation methodologies. Section 5 concludes by summarizing key findings, challenges, and future research directions in the field.

2. Classification of Multi-Agent Hybrid Games

Multi-agent hybrid games refer to complex interactive scenarios in which multiple agents within a multi-agent system exhibit both cooperative and competitive behaviors [15–17]. Such problems are widely observed in real-world settings, such as collaboration and competitions among enterprises in a market economy or multi-robot teams working together to accomplish complex tasks. This section presents a systematic classification of multi-agent hybrid games across multiple dimensions and introduces typical experimental environments commonly used in current research.

2.1. Literature Search of Multi-Agent Game

To systematically analyze the research progress in the field of multi-agent systems, delineate its developmental trajectory, and explore its applications in urban and architectural domains, we conducted a comprehensive literature search in the Web of Science database. The search employed the following key terms: “Multi-Agent”, “Agent-based building”, “Agent-based cities”, “Building Design”, “Built Environment”, “Smart City”, “Traffic City”, “Traffic Control”, “Urban Smart Technology”, and “Urban City”.

Figure 1 presents a quantitative analysis of publication trends in multi-agent systems research. Figure 2 specifically examines the application of multi-agent systems in urban and architectural contexts through keyword co-occurrence analysis. It should be noted that the data for 2025 in both figures include only the first six months of the year. As illustrated, since 2018, multi-agent reinforcement learning has garnered increasing attention, demonstrating remarkable potential for addressing complex industrial tasks, which is reflected in the substantial growth of related publications. The data reveal that since 2000, multi-agent reinforcement learning has gained progressively more research interest in these domains. Notably, the fields of agent-based building design and traffic control systems have emerged as particularly active areas of investigation, accounting for the majority of publications in this interdisciplinary research space.

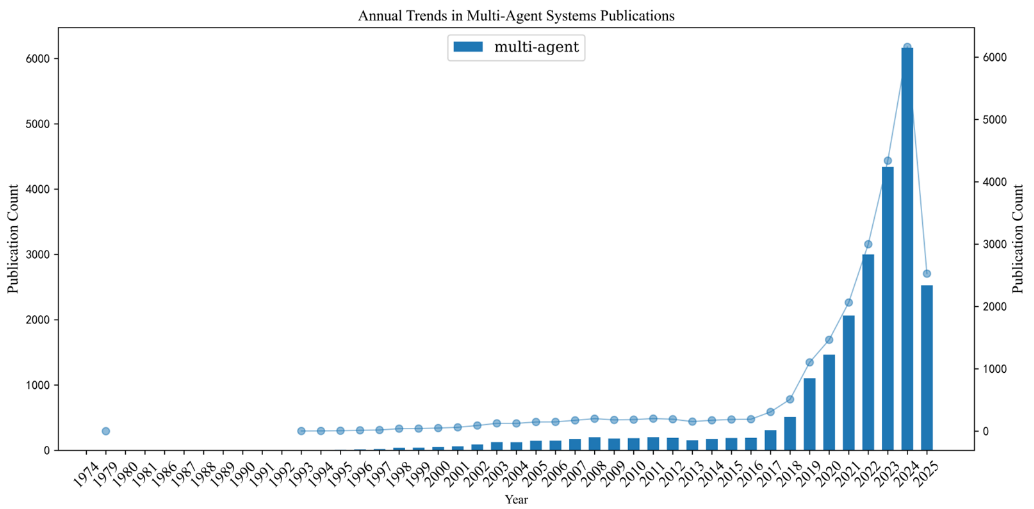


Figure 1. Annual Trends in Multi-Agent Systems Publications.

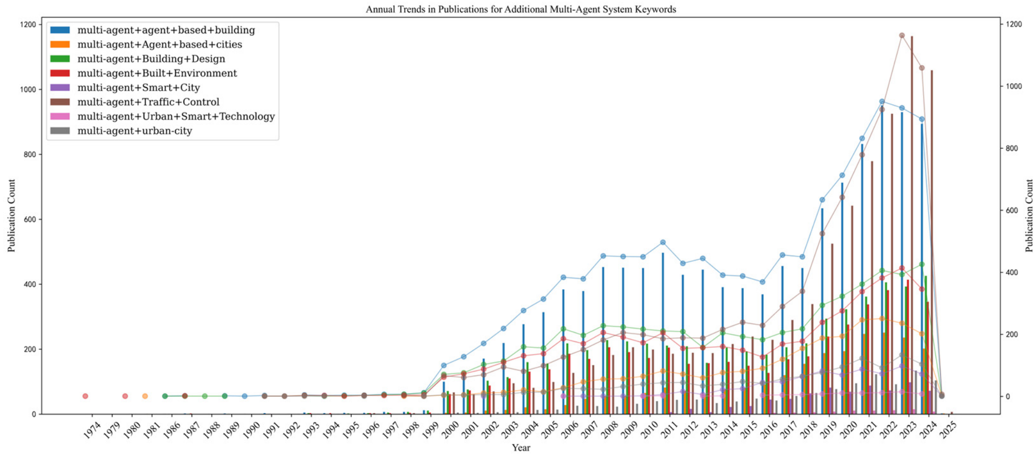


Figure 2. Annual Trends in Publications for Additional Multi-Agent System Keywords.

2.2. Basic Classification of Multi-Agent Game

Multi-agent game can be categorized based on the nature of agent interactions, information availability, and reward structures [18]. This classification clarifies the characteristics of different types of games and provide a foundation for selecting appropriate reinforcement learning methods [19]. According to the interactive nature among agents, multi-agent games can be divided into the three categories: fully cooperative games, fully competitive games, hybrid games.

In fully cooperative games, all agents share identical objectives and collaborate to maximize a reward. For example, in a task where a team of robots cooperatively transports a heavy object, successful completion requires coordinated actions among agents, with no underlying competition. Fully cooperative games emphasize team coordination, high levels of information sharing, and unified objective functions, and are solved using centralized control or cooperative strategies [20].

In fully competitive games, agents pursue strictly opposing objectives, where the gain of one agent directly corresponds to the loss of another. A typical example is a zero-sum game, such as chess, in which the victory of one player inherently entails the defeat of the other. These games emphasize adversarial strategies, with agents often operating under partial information and seeking solutions such as Nash equilibria through competitive learning approaches [21].

Hybrid games, by contrast, represent the most complex and realistic class of multi-agent systems. In these settings, agents simultaneously exhibit cooperative and competitive behaviors. For example, companies in a marketplace may compete for market share while forming temporary price alliances to maximize collective profit. Such environments demand flexible and adaptive strategies, as agents must continuously balance cooperation and competition [22].

2.3. Subcategories of Hybrid Games

Hybrid game problems constitute an important component in multi-agent systems, holding significant application potential in fields such as multiplayer games, urban architectural design, and smart city design. However, due to the inherent complexity of the problems themselves, along with issues like the curse of dimensionality and credit assignment, it is difficult for algorithms to converge to Nash equilibrium. Researchers have proposed different research algorithms for specific problems [23]. To assist researchers in selecting appropriate practical algorithms for their own tasks, this section classifies hybrid game problems. Based on the degree of competition and cooperation as well as profit distribution, hybrid game problems are subdivided into: coordination games, cooperative-competitive games, adversarial games with cooperative elements, zero-sum games, general-sum games, and team games. A summary of the classification of hybrid game problems is provided in Table 1 and Figure 3.

In coordination games, agents align their actions to achieve a common objective yet may face competition in the process of reaching that goal [24,25]. For example, two teams may need to collaborate on a joint project while competing for limited resources. These games emphasize communication and negotiation mechanisms and require the design of effective incentive schemes to promote cooperation.

In cooperative-competitive games, agents both collaborate to achieve shared objectives and compete to optimize their individual performance, making them prevalent in multi-agent reinforcement learning, collaborative

https://doi.org/10.53941/ubs.2025.100004

control, and decision-making. A prime example is Public Goods Games [26], where agents either contribute to or extract from a common resource, yielding outcomes that blend cooperation and competition.

Table 1. Subcategories of Hybrid Games.

Hybrid Game Type	Core Characteristics	Representative Experimental Scenarios
Coordination Games	Agents must align actions to achieve a common objective while managing competition for resources. Emphasizes communication, negotiation, and incentive design.	✓ Multi-agent traffic signal coordination (Traffic Control) in smart cities (Smart City) [29] ✓ Agent-based Building energy management systems with cooperative-competitive device interactions[30]
Cooperative-Competitive Game	Participants must collaborate to achieve shared objectives and compete to optimize individual performance. It involves balancing collective interests and individual gains, with challenges such as externalities and free-riding needing to be addressed.	✓ Urban green space (Outdoor Environment [31–33] ✓ Shared infrastructure (e.g., EV charging stations) allocation in Agent-based Cities [34–36] ✓ Distributed energy trading markets (e.g., solar power exchange) in Smart Cities [30,37,38]
Adversarial Games with Cooperative Elements	Primarily competitive in nature, yet necessitates localized cooperation (such as conflict mitigation). It emphasizes the balance between adversarial strategies and defensive mechanisms.	✓ Multi-agent disaster response coordination vs. resource competition in Urban Planning [39,40] ✓ Cybersecurity (e.g., hacker vs. defender) in Agent-based Buildings [41]
Zero-Sum Game	The interests of the participants are completely opposed, with the sum of payoffs always being zero. There is no room for cooperation, and the gain of one party directly equals the loss of the other. The focus is on formulating optimal adversarial strategies to maximize one's own gains while minimizing the opponent's.	✓ Chess competitions (e.g., chess, go), where the victory of one player means the defeat of the other [42] ✓ Futures market trading between long and short positions, where the profit of long positions comes from the loss of short positions and vice versa [43]
General-Sum Game	The sum of payoffs of participants is not fixed, which can be positive, negative, or zero. There are both competitive and cooperative possibilities. The key is to find a balance between cooperation and competition to maximize the overall interests while considering individual reasonable distribution.	✓ International trade negotiations between countries, where trade wars (competitive) lead to negative-sum results, while free trade agreements (cooperative) bring positive-sum benefits [44] ✓ Joint venture cooperation between enterprises, where they compete for market share in some respects but cooperate in technology research and development to achieve mutual benefit [45]
Team Game	All participants have a common goal, and individual payoffs are closely linked to the team's overall results. There is no competition among individuals, and the focus is on establishing efficient cooperation mechanisms and information sharing channels to give full play to the team's synergy.	✓ Scientific research teams tackling key technical problems, where members with different expertise cooperate to complete experiments and achieve research breakthroughs [46] ✓ Sports teams in collective events (e.g., football, basketball), where players coordinate their actions to win the game through teamwork [47–51]

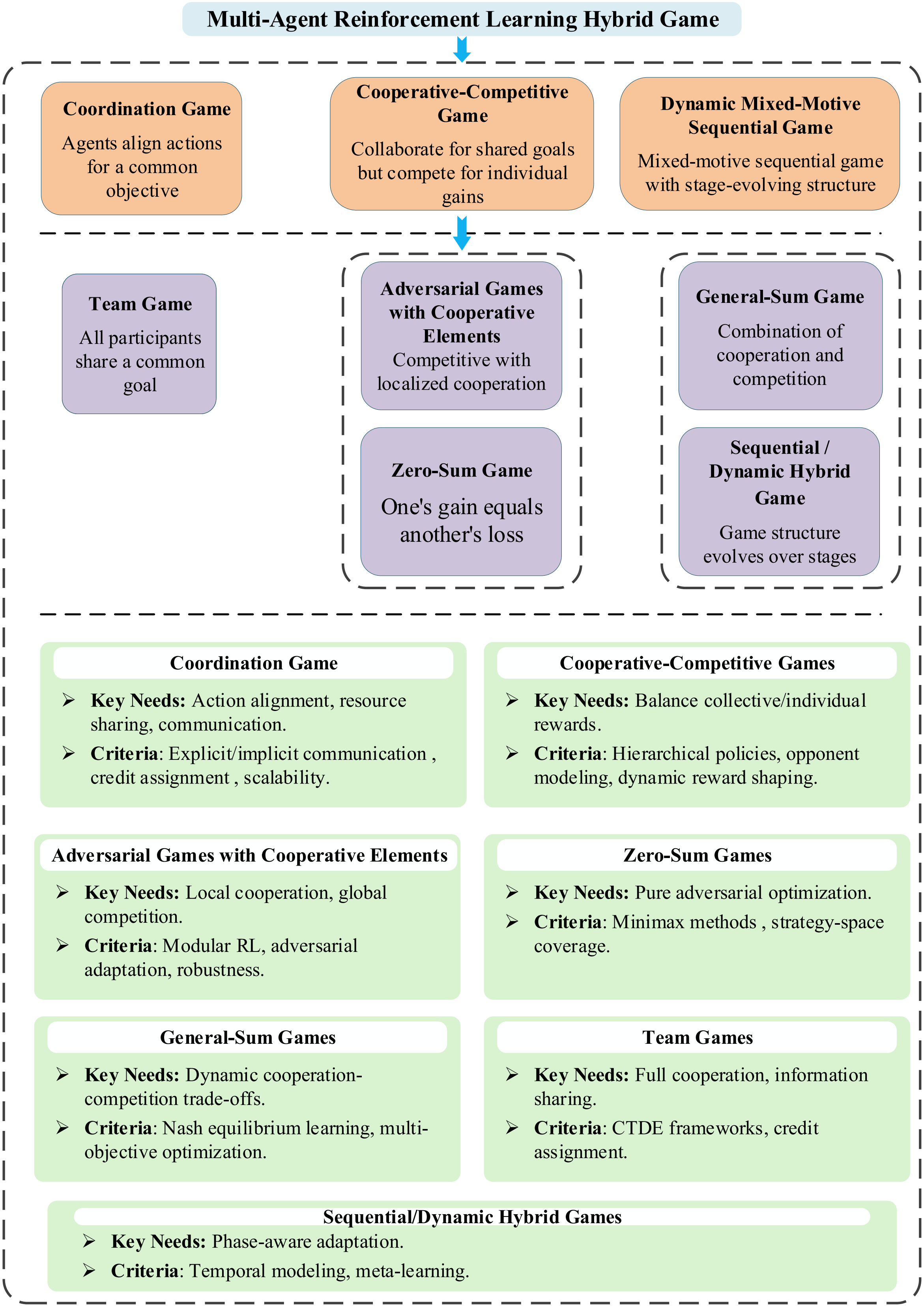


Figure 3. A Taxonomy of Hybrid Multi-Agent Games and Corresponding Algorithm Selection Criteria. This framework categorizes hybrid multi-agent games by their cooperation-competition balance (**up**) and maps them to algorithm selection criteria (**down**).

The management of fisheries provides a clear illustration. Each fisher has an incentive to harvest as much as possible; however, excessive harvesting risks depleting the shared resource. This scenario highlights critical challenges such as externalities and free-riding, which necessitate mechanisms to prevent overexploitation. Another important example is market games [27]. These models describe interactions between buyers and sellers as a blend of competition and cooperation. For instance, firms in a competitive market may compete for market share while simultaneously collaborating to achieve economies of scale.

Additionally, adversarial games with cooperative elements are primarily competitive in nature but incorporate elements of cooperation. A typical example is military confrontation, where opposing forces engage in adversarial actions yet may still need to coordinate on specific operational or humanitarian matters. Such games emphasize adversarial strategies and defense mechanisms, often requiring careful consideration of the intensity and scope of confrontation. Another example lies in constrained settings: Sai et al. proposed a mul-ti-agent framework that integrates game theory with deep reinforcement learning [28]. By constructing a "po-lice-robber" asymmetric game model and designing targeted reward functions, they achieved the collaborative learning of deception and counter-deception strategies in adversarial environments. Experimental validation demonstrated that this framework can significantly reduce the rate of resource misallocation and improve in-terception efficiency, providing a scalable learning paradigm for adversarial multi-agent systems and bridging the theoretical integration of classical game theory and deep reinforcement learning.

A zero-sum game is a type of completely adversarial game. In this kind of game, the gain of one party necessarily implies the loss of the other party, and the sum of the payoffs of all participants is always zero. This determines that there is no possibility of cooperation among the participants. The goals of the two sides are completely opposed, and the success of one side is built on the failure of the other side. For example, in a chess game, when one side captures the opponent's pieces and occupies favorable positions, it aims to weaken the opponent's strength and increase its own winning probability. Eventually, when one side wins, the other side lose, and the interests of the two sides show a strictly inverse relationship. In the futures trading market, the game between long and short positions also belongs to a zero-sum game. The profit of the long side is exactly the loss of the short side, and vice versa. The core of this type of game lies in the antagonism of strategies and the strict opposition of payoffs. Participants need to focus on formulating the optimal counter-strategy, maximize their own payoffs by predicting the opponent's actions, and minimize the opponent's payoffs at the same time, which requires extremely high strategic analysis and prediction capabilities of the participants.

A general-sum game is a form of game with an unfixed sum of payoffs. The sum of the payoffs of the participants can be positive, negative, or zero. This makes there exist both competitive and potential cooperative spaces among the participants. Different from the complete antagonism of the zero-sum game, the interests of the participants in a general-sum game are not completely opposed. Through cooperation, it is possible to achieve a common increase in the payoffs of both sides, while excessive competition may lead to the damage of the interests of both sides. For instance, in international trade, the trade game between two countries belongs to a general-sum game. If both sides adopt trade protectionism and impose tariffs on each other, it may lead to a decline in the trade volume of both sides and damage to their economic interests, with a negative sum of payoffs. However, if the two sides carry out free trade cooperation and reduce trade barriers, they can achieve complementary advantages and promote economic growth, making the sum of payoffs of both sides positive. The key to a general-sum game lies in finding the balance between cooperation and competition. Participants must avoid inefficient competitive scenarios and instead seek to maximize overall benefits through mechanisms such as negotiation and agreement formation, while ensuring a reasonable allocation of individual interests. This requires in-depth investigation of strategic interactions and payoff distribution mechanisms among the participants.

A team game takes cooperation among participants as the core. All participants have a common goal, and their payoffs depend on the overall action results of the team. The individual payoff is closely related to the overall payoff of the team, and there is no competitive relationship among individuals. In this kind of game, participants need to work together, coordinate their actions, and give full play to the advantages of each member to achieve the optimal realization of the team goal. For example, a scientific research team is carrying out a major scientific research project [52]. Team members are responsible for different research links. Some are responsible for theoretical research, some for experimental operations, and some for data analysis, etc.

Only when all members cooperate efficiently and coordinate with each other can the project proceed smoothly and finally achieve scientific research results. And the payoffs (such as honors, bonuses, etc.) of each member are closely related to the success of the project. The project team in an enterprise also falls within the category of a team game. To ensure timely and high-quality completion of project tasks, team members must engage in close communication, effectively divide responsibilities, and collaborate efficiently. They must also work together to address challenges that arise during project implementation [53].

The focus of a team game lies in establishing an efficient cooperation mechanism and information sharing channel to ensure that team members can clarify their own responsibilities and give full play to the synergy effect. At the same time, it is necessary to design a reasonable team incentive mechanism to stimulate the enthusiasm and creativity of each member and improve the overall performance of the team.

2.4. Experimental Scenarios in Current Research

The previous section introduced the classification of hybrid game problems, and this section presents the applications of multi-agent reinforcement learning methods in real-world hybrid game problems [54–56].

In path planning, multiple agents are required to find paths from their respective starting points to destinations while potentially avoiding conflicts and sharing resources, which is common in fields such as robotic navigation and traffic management. In such tasks, agents compete for limited spatial resources while also needing to cooperate in certain situations to avoid collisions. In pursuit-evasion games, a group of pursuers attempts to capture an evader, requiring coordination among pursuers, and this class of problems serves as a classical benchmark for studying multi-agent cooperation. The pursuers must work together to surround and capture the evader, while the evader learns to counter their strategies. In resource allocation, multiple agents compete for limited resources and must balance individual and collective interests, with such problems prevalent in domains such as cloud computing and energy management where agents seek cooperative opportunities within competitive environments. In robotic control, particularly in manufacturing or logistics environments, multiple robots must collaborate while also competing for limited resources or goals, and these tasks emphasize the coordination and cooperation among robots, accounting for physical constraints and task dependencies.

In energy management, the allocation of energy resources in smart buildings is often modeled as a hybrid game; for example, one study proposed an Uncertainty-Aware Minority Game-based Energy Management System (UAMG-EMS) for precise energy allocation between hybrid renewable sources and the main grid in smart buildings. In this system, multiple agents are deployed within the building and account for two types of uncertainty: (i) random noise from energy meters/sensors, and (ii) unpredictable working behaviors on the load side, and these agents allocate limited solar energy resources using a modified minority game framework. In cybersecurity, the protection of cyber-physical systems (CPS) involves hybrid game formulations, as one study proposed a hybrid game-theoretic and reinforcement learning-based method for CPS security, modeling defense strategies at two levels: the strategic and battlefield levels. The strategic layer is formulated as an extensive-form game with incomplete information, where human administrators and malware authors decide on defense and attack strategies, respectively; at the battlefield layer, strategies are implemented by learning agents that derive optimal policies in real time, with the resulting outcomes influencing the utility at the higher level, aiming to reach a Nash equilibrium that favors the defender.

In power grid control, voltage regulation in active distribution networks with high renewable penetration involves hybrid games, and one study proposed a two-timescale voltage regulation method for active distribution networks considering heterogeneous devices. The problem was first formulated to minimize power losses across the network while coordinating various mixed devices; given the complexity, it was reformulated as a bi-level Markov game, and a hierarchical multi-agent attention-based deep reinforcement learning algorithm was developed to solve the problem. Specifically, the upper-level Markov game was solved using a discrete Multi-Actor Attention Critic (MAAC) algorithm, while the lower-level game was solved using a continuous MAAC algorithm, with two-timescale coordination between upper- and lower-level agents achieved via reward-sharing during training, and simulation results demonstrated the proposed method's effectiveness, robustness, and scalability in voltage regulation.

In urban traffic signal control, coordinated traffic signal control in urban road networks is another instance of hybrid games, and one study introduced a distributed signal control method based on game theory and multi-agent reinforcement learning. Intersections were modeled as players in a distributed game framework, and Nash equilibrium solutions were obtained using a mixed-strategy game formulation; within a multi-agent reinforcement learning framework, Nash Q-learning was used to update the payoff matrices and learn optimal control policies, with simulations showing that under low traffic demand, the proposed method significantly reduced average delays and waiting times, and under high traffic demand, it outperformed single-agent control, effectively avoiding congestion and reducing network-wide delays.

These classifications and experimental scenarios illustrate the diversity and complexity of multi-agent hybrid games [57]. Such problems incorporate both cooperative and competitive elements, requiring agents to make trade-offs and derive optimal strategies during decision-making. In the following section, we delve into the technical approaches and algorithmic frameworks for addressing hybrid game problems using reinforcement learning.

3. Reinforcement Learning Algorithms for Multi-Agent Hybrid Games

Reinforcement Learning (RL), as a powerful learning paradigm, offers novel approaches and methods for addressing multi-agent hybrid game problems. Through trial-and-error interactions with the environment, agents can autonomously discover optimal strategies without relying on expert domain knowledge [58]. This section systematically introduces the application of RL in multi-agent hybrid games, covering fundamental RL algorithms, value-based methods, policy-based methods, and hybrid approaches.

3.1. Fundamental Reinforcement Learning Algorithms

Basic RL algorithms lay the foundation for solving multi-agent hybrid game problems. The following introduces two fundamental value-based RL algorithms.

Q-Learning is a value-based reinforcement learning algorithm that finds the optimal policy by learning a value function over state-action pairs [59]. In Q-Learning, an agent estimates the Q-value of each state-action pair, representing the expected cumulative reward obtained by taking a given action in a given state and subsequently following the optimal policy. The update rule for Q-Learning is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha \left[r + \gamma \max_{a'} Q\,(s',a') - Q(s,a)\right], \tag{1}$$

where α is the learning rate, γ is the discount factor, r is the immediate reward, s is the current state, a is the current action, and s' is the next state.

In multi-agent hybrid games, Q-Learning faces significant challenges, primarily *non-stationarity* and *policy overgeneralization*. Non-stationarity refers to the dynamic nature of an agent's reward function, which changes as other agents adapt their policies. Policy overgeneralization occurs when an agent incorrectly applies learning experiences from one state-action pair to similar but distinct situations. To address these challenges, researchers have proposed various improvements, such as SARSA, hybrid Q-learning/SARSA algorithms, and other stabilized learning frameworks.

SARSA (State-Action-Reward-State-Action) is another value-based RL algorithm similar to Q-Learning, but it updates the Q-value using the action actually taken by the agent under its current policy, rather than the greedy action [60]. Its update rule is:

$$Q(s,a) \leftarrow Q(s,a) + \alpha[r + \gamma Q(s',a') - Q(s,a)]. \tag{2}$$

SARSA's primary advantage lies in its on-policy nature—it learns directly from interactions using the current policy without requiring a model of the environment. In multi-agent hybrid games, SARSA often exhibits greater stability compared to Q-Learning, due to its more conservative update mechanism, which avoids overestimation of action values.

3.1.1. Value-Based Methods

Value-based methods are among the most prominent approaches for solving multi-agent hybrid game problems. Representative algorithms include Value-Decomposition Networks (VDN) and QMIX. Value-Decomposition Networks (VDN**)** assumes that the global team reward can be decomposed into the sum of individual agents' rewards [61]. Under this assumption, each agent learns its own Q-value function independently, and the global Q-value is obtained by summing all individual Q-values. The architecture of VDN is relatively simple and computationally efficient, making it suitable for solving basic cooperative tasks. VDN offers fast training and ease of implementation, with simple neural architectures and few parameters. It typically uses a single shared value network, resulting in a lightweight model. However, its expressiveness is limited, making it less capable of handling complex cooperative tasks.

Monotonic Value Function Factorization for QMIX [62] extends value decomposition by learning a more flexible mixing function to combine agents' Q-values. Unlike VDN, QMIX does not require that the global reward be a linear sum of individual rewards. Instead, it uses a mixing network conditioned on the global state to learn how to combine individual agent Q-values. This enables more expressive coordination among agents. QMIX achieves relatively fast training with improved final performance, moderate parameter size, and intermediate architectural complexity. It is suitable for large-scale systems and more complex cooperative tasks.

3.1.2. Policy-Based Methods

Policy-based methods directly mappings from states to actions and include approaches such as MADDPG [63] and Multi-Agent Twin Delayed Deep Deterministic Policy Gradient (MATD3) [64].

Multi-Agent Deep Deterministic Policy Gradient (MADDPG) adopts the actor-critic architecture, where each agent has an actor network (for policy learning) and a critic network (for evaluating policy quality). The critic learns the joint Q-function over all agents' actions, while the actor updates its policy by maximizing the critic's output.

MADDPG is effective in continuous action spaces and heterogeneous agent settings. However, it suffers from moderate training speed, high parameter complexity, and intricate network structure.

Multi-Agent Twin Delayed DDPG (MATD3) improves upon MADDPG by introducing two critic networks and applying delayed policy updates. This delay stabilizes learning by ensuring that the actor is updated less frequently than the critics, reducing the variance and mitigating instability caused by rapid policy shifts.

MATD3 demonstrates relatively fast training, superior final performance, and strong robustness in continuous and complex environments. Despite having high parameter complexity and network overhead, it is well-suited to challenging real-world tasks.

3.2. Hybrid Methods

Hybrid methods, by integrating the strengths of diverse learning strategies or algorithms, offer effective solutions to the complexities inherent in multi-agent hybrid games. A prominent paradigm within this domain is Centralized Training with Decentralized Execution (CTDE), which enables agents to share information centrally during training while acting autonomously during execution. This framework strikes a balance between collaborative learning and independent decision-making, making it suitable for multi-agent hybrid game scenarios [1].

Hybrid methods integrate diverse learning strategies or algorithms to solve the complexities of multi-agent hybrid games. Centralized Training with Decentralized Execution (CTDE) a key paradigm. It enables agents to share information centrally during training but act autonomously during execution, balancing collaborative learning and independent decision-making [1]. To handle multi-agent systems with mixed action spaces, researchers extended Proximal Policy Optimization (PPO) using CTDE, combined with Markov games. As standalone PPO struggles with mixed action spaces, CTDE allows agents to learn from shared info during training and act independently later. Incorporating a Beta distribution and tailored output modes improved policy expression. Simulations showed this CTDE-enhanced PPO outperformed standalone PPO [65].

Hybrid Independent Learner methods integrate multiple strategies to mitigate non-stationarity and overgeneralization. In dynamic environments, independent learners encounter specific challenges: an emphasis on immediate rewards can result in short-sighted behavior, whereas reliance on average rewards may slow the learning process. The hybrid approach combines optimistic and average-reward strategies. Agents first maximize the observed rewards, then switch to average-reward to ensure long-term stability. It updates Q-values with a hybrid rule, improving convergence and stability [66]. Genetic Programming (GP) combined with multi-agent Reinforcement Learning (RL) merges gradient-free and gradient-based methods. Traditional gradient-based RL is prone to local optima, while GP excels at global exploration but lacks fine-grained refinement [18]. In complex battlefield games, integrating them optimized policy parameters, improving win rates and stability compared to traditional methods.

Hierarchical Multi-Agent Reinforcement Learning (HRL) decomposes tasks to address dimensionality issues [67]. In 3D underwater UUV scenarios, a hierarchical framework separated high-level strategic planning and low-level maneuvering control. A constrained Apollonius model managed UUV interactions, and asynchronous training with curriculum learning handled complexity. Experiments showed it improved credit assignment and generalization. Hybrid Game-Theoretic Methods combine game models for large-scale multi-agent systems. Single models can’t capture LS-MAS complexity [68]. A distributed method integrated mean-field, Stackelberg, and cooperative games. A hybrid actor-critic algorithm managed leader and follower strategies. Simulations showed it achieved stable formation control with lower complexity.

3.3 Domain-Knowledge-Enhanced Reinforcement Learning for Hybrid Games

Addressing the intertwined challenges of non-stationarity, high dimensionality, and partial observability in hybrid multi-agent games requires reinforcement learning (RL) methods that are informed by domain expertise. Among the most effective approaches are game-theoretic reward shaping [69,70]. and Policy Space Response Oracle (PSRO)-based techniques [71–73], both of which integrate structural knowledge of the interaction environment to accelerate convergence, improve generalization, and promote stable equilibria. Reward shaping incorporates task-specific insights to align agent incentives with socially optimal outcomes, mitigating the risk of local, self-serving strategies that undermine global performance In hybrid games where cooperative and competitive elements coexist, such alignment is crucial. For example, shared goal representation can be achieved by embedding global team objectives within a deterministic finite automaton (DFA) and dynamically adapting rewards according to the frequency of successful transitions in historical trajectories, as demonstrated in “Intruders and Guards” scenarios where this mechanism fosters emergent role specialization [70]. Similarly, responsibility-aware reward decomposition in dynamic job shop scheduling partitions global objectives—such as minimizing total tardiness—into local performance responsibilities, distributing penalties proportionally across upstream and downstream machines when delays occur, thereby encouraging system-wide optimization [69]. Reward functions can also be made sensitive to critical states by amplifying penalties for near-miss failures or high-stakes misdirections, focusing learning on pivotal decisions with disproportionate impact on long-term success.

PSRO-based methods complement reward shaping by enabling meta-strategy learning and equilibrium computation in complex, multi-agent environments, particularly those exhibiting non-transitive dynamics or adaptive adversaries. Through iterative expansion of strategy pools and optimization of best responses, PSRO constructs increasingly robust policy sets. Incorporating diversity constraints, such as Determinantal Point Processes, ensures that the strategy pool spans the action space, avoiding premature convergence to suboptimal equilibria; the CPPU algorithm for Integrated Sensing and Communication systems exemplifies this approach, iteratively refining policies until exploitability approaches zero [72]. In non-transitive games, the Unified Diversity Measure (UDM) enhances PSRO's adaptability by preserving a balanced repertoire of offensive, defensive, and mixed strategies, thereby mitigating overfitting and stabilizing equilibria. In adversarial contexts such as network deception, PSRO can be coupled with deep Q-networks to model attack–defense interactions as zero-sum differential games [71]. Alternating updates to attacker and defender networks, followed by saddle-point equilibrium computation, yield deception and counter-deception strategies that converge more rapidly than conventional PSRO—achieving, for example, a 95.6% reduction in convergence time in cloud-native security scenarios [73].

In summary, domain-knowledge-enhanced RL methods operationalize structured priors—through both game-theoretic reward shaping and PSRO-based equilibrium learning—to reconcile individual and collective objectives, foster robust and adaptive meta-strategies, and accelerate convergence in non-stationary, high-complexity hybrid games. These techniques collectively strengthen performance stability and generalization in dynamic, competitive–cooperative environments.

4. Evaluation Metrics and Key Influencing Factors

In multi-agent mixed games, evaluating the performance of reinforcement learning (RL)-based methods is a complex yet critical task. It requires a comprehensive assessment from multiple dimensions to accurately reflect the algorithm's effectiveness in practical applications. This section systematically discusses the evaluation metrics for RL methods in multi-agent mixed games, analyzes key factors influencing algorithmic performance, and explores how to scientifically and comprehensively assess the efficacy of these approaches.

4.1. Individual Performance Metrics

Individual performance metrics focus on the behaviors and outcomes of agents within the mixed game setting [38]. This subsection reviews several metrics for evaluating the performance of MARL algorithms, namely Cumulative Reward, Discounted Cumulative Reward, Task Success Rate, and Task Completion Time. These metrics directly reflect the learning effectiveness and adaptability of agents.

Cumulative reward is the most direct metric for evaluating an agent's performance in an environment [74]. It represents the total reward accumulated by the agent over one or more episodes. Typically, cumulative reward increases as training progresses. During evaluation, the average cumulative reward across multiple episodes is calculated to assess long-term performance. For example, in a robotic navigation task, an agent receives positive rewards for reaching target locations and negative rewards for colliding with obstacles. The cumulative reward thus reflects the agent's overall task performance. A higher cumulative reward indicates better performance.

To account for the diminishing value of future rewards compared to immediate ones, a discount factor, discounted cumulative reward, is applied. Discounted cumulative reward emphasizes the agent's ability to balance short-term and long-term returns. It is calculated as:

$$G_t = \sum_{k=0}^{\infty} \gamma^k r_{t+k+1}, \tag{3}$$

Where G_t is the discounted return at time step t, γ is the discount factor ($0 \leq \gamma \leq 1$), and r_{t+k+1} is the reward received at time $t + k + 1$.

For tasks with defined goals, the success rate is a critical metric. It denotes the proportion of episodes in which the agent successfully completes the task. For instance, in a game environment, it is the number of successful level completions divided by the total number of games played; in a robotic grasping task, it is the number of successful grasps divided by the total number of attempts. In addition to success rate, the time required to complete a task is also essential, particularly in real-time scenarios such as autonomous driving or logistics. The agent's ability to complete tasks within a shorter time frame is an important indicator of its performance [74].

4.2. Collective Performance Metrics

Collective performance metrics assess the overall behavior of multi-agent systems, reflecting the quality of cooperation and the level of social welfare. This subsection reviews three metrics: Social Optimality, Fairness, and Stability. Social optimality refers to the sum of cumulative rewards obtained by all agents. It evaluates the overall system performance; a higher social optimum indicates a better-performing system as a whole.

Fairness pertains to the distribution of rewards or costs among agents. More equitable distributions are desirable. Fairness can be quantified using various measures such as the Gini coefficient or max-min fairness. Stability evaluates the convergence of policies during training and their robustness to environmental changes. More stable policies are more reliable. In multi-agent systems, stability is particularly important, as interactions among agents may lead to policy oscillation or divergence.

4.3. Cooperation and Competition Metrics

This subsection reviews three specific metrics for hybrid games: the degree of cooperation, the degree of competition, and the Nash equilibrium [75].

The metric, degree of cooperation, quantifies how well agents cooperate with each other. A high degree of cooperation indicates close coordination in pursuing shared goals. In cooperative tasks, higher cooperation is generally preferred [67]. The metric, the degree of competition, reflects the intensity of competition among agents. A higher degree of competition suggests stronger pursuit of individual goals. In competitive tasks, increased competition is desirable. The metric, Nash equilibrium, assesses whether the agents' strategies converge to a Nash equilibrium, a state in which no agent can unilaterally change its strategy to improve its own reward. As a fundamental concept in game theory, a Nash equilibrium represents a stable strategic configuration [5].

4.4. Implementation-Specific Metrics

Implementation-specific metrics evaluate the feasibility and efficiency of reinforcement learning algorithms in real-world applications. These metrics are critical for the practical deployment and operational performance of the algorithms. This subsection reviews three metrics: training efficiency, communication efficiency, privacy preservation.

Training efficiency refers to the amount of training data or iterations required for the algorithm to reach a desired performance level. Higher training efficiency indicates faster learning. This is important in real-world scenarios, where prolonged training can increase computational costs and delay deployment. Communication efficiency measures the volume of information exchanged among agents. Lower communication requirements are generally desirable, especially in environments where communication costs are high. In multi-agent systems, communication overhead can be a major concern, particularly in distributed settings. Privacy preservation reflects the algorithm's ability to protect agents' private information during the learning process. Higher levels of privacy are often preferable, especially in applications involving sensitive data. In certain scenarios, agents may need to safeguard their policies or internal states, requiring the design of privacy-preserving reinforcement learning algorithms.

4.5. Key Factors Influencing Algorithm Performance

This subsection reviews five key factors that influence the performance of RL algorithms in multi-agent mixed games, including algorithm design, environmental complexity, parameter settings, communication mechanisms, and reward design.

The factor of algorithm design has a direct impact on its performance. Different frameworks—such as value-based methods (e.g., VDN, QMIX), policy-based methods (e.g., MADDPG, MATD3), and hybrid approaches—exhibit different strengths and limitations [76,77]. For example, VDN offers fast training but limited performance; QMIX improves both training speed and performance; MADDPG is well-suited for heterogeneous agents and continuous action spaces; and MATD3 performs well in complex environments with continuous action spaces.

The factor of environmental complexity affects algorithm performance. Factors such as dynamics, stochasticity, and the nature of agent interactions influence both the applicability and effectiveness of RL methods. For instance, dynamic environments require greater adaptability; stochastic environments necessitate extensive exploration; and intricate agent interactions may demand more expressive policy representations. Algorithmic performance is also sensitive to hyperparameter configurations. Parameters such as learning rate, discount factor, and the exploration–exploitation trade-off must be carefully tuned based on the specific task [8]. For example, an excessively high learning rate may cause instability, while a very low rate may slow convergence; a large discount factor could emphasize future rewards, whereas a small one may lead to short-sighted behavior.

In multi-agent systems, the mechanism by which agents communicate also influences performance [7]. The frequency, content, and modality of communication affect coordination and policy optimization. Full information sharing may introduce privacy concerns and communication overhead, while limited or no sharing may result in poor coordination. Reward design is a crucial engineering challenge in reinforcement learning. The reward function should encourage desirable behaviors while avoiding unintended incentives. In mixed games, reward design becomes even more important, as it must balance individual interests with collective welfare.

4.6. Evaluation Methodologies

Evaluating reinforcement learning methods in multi-agent mixed games requires a multi-faceted approach. This subsection reviews five Commonly used evaluation methodologies: baseline comparison, ablation studies, sensitivity analysis, real-world validation, visualization analysis.

The proposed method is compared against existing baseline algorithms to assess its strengths and weaknesses. Baselines may include classical algorithms such as Q-learning [78] and SARSA [79], as well as multi-agent-specific methods like MADDPG and QMIX. Ablation studies assess the contribution of individual components by removing or modifying them. For instance, one can evaluate the impact of using mixed strategies or centralized training with decentralized execution [80,81].

Sensitivity analysis examines how variations in parameter settings affect performance, providing insight into the robustness of the algorithm. For example, the influence of learning rate and discount factor changes can be investigated. Algorithms are tested in real-world environments to verify their practical effectiveness. Applications may include robotic control systems, traffic management, and other real-time decision-making scenarios [82]. Visualization techniques help intuitively evaluate the behavior and policies of agents. For example, visualizing agent trajectories or policy distributions can provide insight into the rationality and effectiveness of the learned strategies.

4.7. Evaluation Methodologies

Based on the core characteristics of each hybrid game type, particularly the levels of Goal Coupling and Resource Competition defined in Table 1, we identify and map the most critical evaluation metrics for each category. This prioritized mapping guides researchers in selecting the most relevant indicators for assessing agent performance in specific hybrid game settings. The core metrics for each type are summarized in Table 2 below.

Table 2. Mapping of Core Evaluation Metrics to Hybrid Game Types.

Hybrid Game Type	Core Evaluation Metrics	Rationale (Linked to Core Characteristics)
Coordination Games	Social Welfare (Sum Reward), Task Success Rate, Degree of Cooperation, Communication Efficiency.	High goal coupling requires measuring collective success (Social Welfare, Success Rate) and the quality of coordination (Degree of Cooperation, Comm. Eff.). Low resource competition minimizes the need for intense competitive metrics.
Cooperative-Competitive Game	Social Welfare, Fairness (e.g., Gini), Degree of Cooperation, Degree of Competition, Stability.	Balancing collective gains (Social Welfare) with individual equity (Fairness) is paramount. Both cooperative (Cooperation Degree) and competitive (Competition Degree) behaviors need quantification. Stability is key as the balance can be delicate.
Adversarial Games with Cooperative Elements	Individual Cumulative Reward, Nash Convergence, Degree of Competition, Task Success Rate (for cooperative subtasks).	High resource competition focuses evaluation on individual outcomes (Individual Reward, Nash Conv.) and competitive intensity (Competition Degree). Medium goal coupling necessitates measuring success in cooperative elements (Success Rate on specific cooperative tasks).
Zero-Sum Game	Individual Cumulative Reward, Nash Convergence, Stability (against opponent adaptation)	Extreme competition requires assessing individual strategic dominance (Individual Reward, Nash Conv.) and the robustness of that dominance (Stability). No goal coupling makes collective metrics irrelevant.
General-Sum Game	Social Welfare, Fairness, Nash Convergence, Degree of Cooperation, Degree of Competition.	Variable coupling/competition demands a flexible set. Core metrics include overall efficiency (Social Welfare), distribution (Fairness), strategic stability (Nash Conv.), and the balance of collaborative vs. antagonistic behaviors (Coop./Comp. Degree).

Table 2. *Cont.*

Hybrid Game Type	Core Evaluation Metrics	Rationale (Linked to Core Characteristics)
Team Game	Social Welfare, Task Success Rate, Degree of Cooperation, Communication Efficiency, Training Efficiency.	Extreme goal coupling makes collective success (Social Welfare, Success Rate) and seamless coordination (Cooperation Degree, Comm. Eff.) paramount. Training Efficiency is often crucial for complex team tasks. No resource competition minimizes competitive metrics.
Sequential/Dynamic Hybrid Game	Metrics relevant to each phase (see above), Adaptability (measured via Stability across phases or Task Success Rate under transitions), Training Efficiency.	Dynamic nature requires applying the core metrics relevant to each phase (e.g., Coordination metrics in collaborative phases, Competitive metrics in competitive phases). Crucially, Adaptability to changing game structures and Training Efficiency to handle complexity become core concerns.

4.8. Application-Driven Guidance for Metric Selection

The scenarios in Section 2.4 demonstrate the diversity of real-world hybrid game problems. While the classification in Section 2.3 structures the understanding of mixed interactions, and Section 4.7 links core metrics to these categories, final metric selection must be grounded in the application's operational demands. Each domain imposes unique constraints, prioritizes distinct objectives, and carries different societal implications, making certain metrics more critical than others.

In navigation domains such as path planning and robotic guidance (e.g., warehouse robots, UAVs), safety—often quantified via collision avoidance—is paramount, followed by efficiency, path optimality, robustness to dynamic conditions, and resource utilization. In pursuit–evasion settings (e.g., security drones, benchmark simulations), primary metrics are task success (capture or escape rates), team coordination efficiency, evader survival, and robustness against adaptive strategies.

Resource allocation problems (e.g., cloud computing, microgrid energy sharing, intelligent building EMS) require balancing efficiency with fairness, assessed through system utility, proportionality, envy-freeness, and stability. In CPS security, defense success, false positive/negative rates, latency, computational overhead, and adversarial robustness dominate. For power grid control, maintaining voltage stability within safe limits is non-negotiable, with efficiency, control smoothness, and robustness as secondary concerns. In urban traffic signal control, network-wide delay, throughput, fairness across routes, robustness to demand variation, and environmental impacts are key.

Across applications, safety and stability take precedence in physical systems, efficiency–fairness trade-offs define resource-sharing contexts, and robustness is universally essential for reliable performance under uncertainty. This application-driven prioritization ensures evaluation aligns with practical goals, constraints, and societal impacts.

4.9. Challenges and Future Directions in Performance Evaluation

Despite advancements in multi-agent reinforcement learning (MARL) and the development of evaluation metrics for hybrid game, several critical challenges remain unresolved, leaving substantial room for improvement in the field. This section examines current gaps in performance evaluation frameworks and outlines potential future research directions to address these limitations.

4.9.1. Current Gaps in Performance Evaluation

The lack of unified evaluation frameworks stands as a significant hurdle, while existing studies have proposed diverse evaluation metrics for multi-agent systems, no comprehensive, standardized framework yet integrates individual and collective performance metrics, cooperation-competition dynamics, and environmental factors. This fragmentation impedes systematic comparisons of algorithms across diverse hybrid game scenarios, as metrics emphasize disjoint aspects of system behavior. Compounding this issue, current evaluation methodologies predominantly rely on controlled or simulated environments, failing to fully capture the complexity and unpredictability of real-world applications. Critical factors such as partial observability, communication latency, and external disturbances are frequently oversimplified or omitted in benchmark tasks, creating potential mismatches between theoretical performance and practical deployment outcomes.

Moreover, many existing evaluation metrics and algorithms are tailored to small-scale or homogeneous multi-agent systems. As applications expand into large-scale urban or industrial scenarios, evaluation frameworks must accommodate heterogeneity, high-dimensional state spaces, and distributed decision-making—areas where current metrics often break down or become computationally intractable.

4.9.2. Future Research Directions

Future research should expanded real-world validation of MARL algorithms in complex, dynamic environments are also essential, which involves developing benchmark tasks that embed realistic constraints—such as sensor noise, intermittent communication, and human-agent interaction—and establishing standardized testbeds (e.g., physical robot swarms) to facilitate cross-study comparisons. Collaborative academia-industry initiatives can accelerate this by sharing real-world datasets and deployment challenges.

Future work should also design scalable and generalizable metrics for large-scale systems, leveraging techniques like graph-based representations (to model agent interactions) or hierarchical evaluation, and ensure these metrics generalize across scenarios. Evaluations should integrate tools to enhance interpretability and transparency, such as attention mechanisms (identifying which agent inputs drive decisions) or counterfactual analysis to enable stakeholders to trust and debug algorithms, especially in regulated domains. Research should focus on developing adaptive evaluation metrics that capture context-dependent trade-offs between cooperation and competition, which could involve dynamic reward functions that adjust based on environmental cues or game-theoretic metrics to quantify strategic alignment.

5. Conclusions

As a frontier in artificial intelligence, the multi-agent mixed game problem has garnered growing attention due to its inherent complexity and high practical significance, with applications spanning architectural optimization design, urban resource scheduling, and intelligent indoor environmental control. This study establishes a systematic analytical framework that examines the problem from three core dimensions: problem categorization, reinforcement learning approaches, and performance evaluation.

A conceptual framework for evaluating the performance of multi-agent reinforcement learning (MARL) algorithms in mixed game scenarios is proposed, featuring a multi-dimensional indicator system that integrates individual performance, collective performance, cooperation-competition dynamics, and implementation feasibility. This framework enables comprehensive assessment of algorithmic effectiveness through metrics such as cumulative reward, social optimality, and Nash equilibrium attainment rate, thereby providing a quantitative foundation for algorithm design and comparative analysis.

Key findings reveal that the intertwined cooperation and competition in multi-agent mixed games necessitate dynamic balancing of these two dimensions in algorithm design. Moreover, algorithm selection must align with problem characteristics: for example, value-based methods demonstrate superiority in cooperative tasks, while policy-based methods are more suitable for scenarios involving continuous action spaces and heterogeneous agents. The proposed multi-dimensional, multi-level performance evaluation framework is proven essential for assessing both the effectiveness and practical applicability of MARL algorithms.

Funding

This work was partially supported by the grant entitled “Non-cooperative Heterogeneous Multi-agent Brain Learning and Game Theory” (Grant No. 2021ZD0201404), which supported the contributions of L.Z. and X.Y. and A.E. W.J. did not receive any support from this grant.

Acknowledgments

We authors acknowledge the contribution and support from Tsinghua Shenzhen International Graduate School, Tsinghua University, China; Department of Mechanical Engineering, Aalto University, Espoo, Finland; Department of Computer Science and Engineering, University of Texas at Arlington, TX, USA; Faculty of Engineering, Free University of Bozen-Bolzano, Bolzan, Italy.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Hu, K.; Li, M.; Song, Z.; et al. A review of research on reinforcement learning algorithms for multi-agents. *Neurocomputing* **2024**, *599*, 128068.
2. Nguyen, T.T.; Nguyen, N.D.; Nahavandi, S. Deep reinforcement learning for multiagent systems: A review of challenges, solutions, and applications. *IEEE Trans. Cybern.* **2020**, *50*, 3826–3839.
3. Du, W.; Ding, S. A survey on multi-agent deep reinforcement learning: From the perspective of challenges and applications. *Artif. Intell. Rev.* **2021**, *54*, 3215–3238.
4. Zhang, K.; Yang, Z.; Başar, T. Multi-agent reinforcement learning: A selective overview of theories and algorithms. In *Handbook of Reinforcement Learning and Control*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 321–384.
5. Li, H.; Huang, W.; Duan, Z.; et al. A survey on algorithms for Nash equilibria in finite normal-form games. *Comput. Sci. Rev.* **2024**, *51*, 100613.
6. Li, Z.; Wu, W.R.; Guo, Y.L.; et al. Embodied Multi-Agent Systems: A Review. *IEEE-CAA J. Autom.* **2025**, *12*, 1095–1116. https://doi.org/10.1109/Jas.2025.125552.
7. Dou, Y.H.; Xing, G.S.; Ma, A.H.; et al. A review of event-triggered consensus control in multi-agent systems. *J. Control Decis.* **2025**, *12*, 1–23. https://doi.org/10.1080/23307706.2024.2388551.
8. Wang, Y.M.; Lin, H.; Lam, J.; et al. Differentially private consensus and distributed optimization in multi-agent systems: A review. *Neurocomputing* **2024**, *597*, 127986. https://doi.org/10.1016/j.neucom.2024.127986.
9. Amirkhani, A.; Barshooi, A.H. Consensus in multi-agent systems: A review. *Artif. Intell. Rev.* **2022**, *55*, 3897–3935. https://doi.org/10.1007/s10462-021-10097-x.
10. Shakya, A.K.; Pillai, G.; Chakrabarty, S. Reinforcement learning algorithms: A brief survey. *Expert Syst. Appl.* **2023**, *231*, 120495.
11. Luo, F.M.; Xu, T.; Lai, H.; et al. A survey on model-based reinforcement learning. *Sci. China Inf. Sci.* **2024**, *67*, 121101. https://doi.org/10.1007/s11432-022-3696-5.
12. Vázquez-Canteli, J.R.; Nagy, Z. Reinforcement learning for demand response: A review of algorithms and modeling techniques. *Appl. Energy* **2019**, *235*, 1072–1089. https://doi.org/10.1016/j.apenergy.2018.11.002.
13. Noaeen, M.; Naik, A.; Goodman, L.; et al. Reinforcement learning in urban network traffic signal control: A systematic literature review. *Expert Syst. Appl.* **2022**, *199*, 116830. https://doi.org/10.1016/j.eswa.2022.116830.
14. Fu, G.T.; Jin, Y.W.; Sun, S.A.; et al. The role of deep learning in urban water management: A critical review. *Water Res.* **2022**, *223*, 118973. https://doi.org/10.1016/j.watres.2022.118973.
15. Dey, S.; Xu, H.; Fadali, M.S. Adaptive distributed formation control for multi-group large-scale multi-agent systems: A hybrid game approach. *IFAC-Pap.* **2023**, *56*, 5482–5487.
16. Khoury, J.; Nassar, M. A hybrid game theory and reinforcement learning approach for cyber-physical systems security. In Proceedings of the NOMS 2020 IEEE/IFIP Network Operations and Management Symposium, Budapest, Hungary, 20–24 April 2020.
17. Gan, W.; Qiao, L. Many-Versus-Many UUV Attack-Defense Game in 3D Scenarios Using Hierarchical Multi-Agent Reinforcement Learning. *IEEE Internet Things J.* **2025**, *12*, 23479–23494.
18. Fitch, N.; Clancy, D. Genetic Programming+ Multi-Agent Reinforcement Learning: Hybrid Approaches for Decision Processes. In Proceedings of the 2022 IEEE Aerospace Conference (AERO), Big Sky, MT, USA, 5–12 March 2022.
19. Qiu, S.; Li, Z.; Pang, Z.; et al. Multi-agent optimal control for central chiller plants using reinforcement learning and game theory. *Systems* **2023**, *11*, 136.
20. Yehoshua, R.; Amato, C. Hybrid Independent Learning in Cooperative Markov Games. In Proceedings of the Distributed Artificial Intelligence: Second International Conference, DAI 2020, Nanjing, China, 24–27 October 2020.
21. Jiang, T.Y.; Ju, P.; Lin, Z.J.; et al. Competitive Incentive Mechanism for Multi-Agents in Demand Response via a Hierarchical Game Considering Joint Uncertainties. *IEEE Trans. Smart Grid* **2025**, *16*, 2913–2925. https://doi.org/10.1109/Tsg.2025.3555590.
22. Azimi, Z.; Afshar, A. Hybrid game-theoretic security assessment of cyber-physical power systems using partial-information multi-agent reinforcement learning. *Sustain. Energy Grids* **2025**, *43*, 101727. https://doi.org/10.1016/j.segan.2025.101727.
23. Marc, F.; Degirmenciyan-Cartault, I.; El Fallah-Seghrouchni, A. An integral cycle for building feasible multi-agent plans. In Proceedings of the IEEE/WIC/ACM International Conference on Intelligent Agent Technology, Beijing, China, 24 September 2004.
24. Zhou, L.Q.; Zheng, Y.S.; Zhao, Q.; et al. Game-based coordination control of multi-agent systems. *Syst. Control Lett.* **2022**, *169*, 105376. https://doi.org/10.1016/j.sysconle.2022.105376.
25. Zhang, J.; Wang, G.; Yue, S.H.; et al. Multi-agent system application in accordance with game theory in bi-directional coordination network model. *J. Syst. Eng. Electron.* **2020**, *31*, 279–289. https://doi.org/10.23919/Jsee.2020.000006.

26. Fley, B.; Florian, M. Trust and the economy of symbolic goods: A contribution to the scalability of open multi-agent systems. *Lect. Notes Artif. Int.* **2005**, *3413*, 176–198.
27. Sajid, A.H.; Iftikhar, M.Z.; Kazmi, S.A.A.; et al. Multi-agent system for optimized energy management in multi-smart buildings via deregulated market system. *Energy Rep.* **2025**, *14*, 455–472. https://doi.org/10.1016/j.egyr.2025.06.031.
28. Mareddy, S.K.R.; Maity, D. Learning Deceptive Strategies in Adversarial Settings: A Two-Player Game with Asymmetric Information. *Appl. Sci.* **2025**, *15*, 7805.
29. Zhang, H.P.; Wang, Z.W.; Yu, J.L.; et al. Learning Simultaneous and Sequential Decisions in Multi-Agent Systems With Application to Traffic Signal Control. *IEEE Trans. Intell. Transp. Syst.* **2025**, *26*, 8257–8267. https://doi.org/10.1109/Tits.2025.3560712.
30. Guan, Q.Q.; Zhang, H.X.; Yang, J.; et al. Design of energy management strategy for electro-hydraulic hybrid vehicles based on multi-agent systems. *Energy Sources Part A* **2025**, *47*, 2512996. https://doi.org/10.1080/15567036.2025.2512996.
31. Liu, S.; Feng, Y.; Ren, N. Distributed Urban Road Network Signal Coordination Control Method Based on Game Theory and Multi-Agent Reinforcement Learning. In Proceedings of the 2024 12th International Conference on Information Systems and Computing Technology (ISCTech), Xi'an, China, 8–11 November 2024.
32. Laudan, J.; Heinrich, P.; Nagel, K. High-Performance Simulations for Urban Planning: Implementing Parallel Distributed Multi-Agent Systems in MATSim. In Proceedings of the 2024 23rd International Symposium on Parallel and Distributed Computing (ISPDC), Chur, Switzerland, 8–10 July 2024. https://doi.org/10.1109/Ispdc62236.2024.10705395.
33. Caprioli, C. The integration of multi-agent system and multicriteria analysis for developing participatory planning alternatives in urban contexts. *Environ. Impact Assess. Rev.* **2025**, *113*, 107855. https://doi.org/10.1016/j.eiar.2025.107855.
34. do Nascimento, L.V.; de Oliveira, J.P.M. A multi-agent architecture for context sources integration in smart cities. *Future Gener. Comput.Syst.* **2025**, *172*, 107862. https://doi.org/10.1016/j.future.2025.107862.
35. Kalyuzhnaya, A.; Mityagin, S.; Lutsenko, E.; et al. LLM Agents for Smart City Management: Enhancing Decision Support Through Multi-Agent AI Systems. *Smart Cities* **2025**, *8*, 19. https://doi.org/10.3390/smartcities8010019.
36. Suanpang, P.; Jamjuntr, P. Optimizing Electric Vehicle Charging Recommendation in Smart Cities: A Multi-Agent Reinforcement Learning Approach. *World Electr. Veh. J.* **2024**, *15*, 67. https://doi.org/10.3390/wevj15020067.
37. Lin, Z.; Xiuying, X.; Guiqing, Z.; et al. Building energy saving design based on multi-agent system. In Proceedings of the 2010 5th IEEE conference on industrial electronics and applications, Taichung, Taiwan, 15–17 June 2010.
38. Qiu, D.; Wang, J.; Dong, Z.; et al. Mean-field multi-agent reinforcement learning for peer-to-peer multi-energy trading. *IEEE Trans. Power Syst.* **2022**, *38*, 4853–4866.
39. Mancy, H.; Ghannam, N.E.; Abozeid, A.; et al. Decentralized multi-agent federated and reinforcement learning for smart water management and disaster response. *Alex Eng. J.* **2025**, *126*, 8–29. https://doi.org/10.1016/j.aej.2025.04.033.
40. Zhang, K.; Bian, F.L. A design and application of a multi-agent system and GIS for simulation of expansion in urban planning. In Proceedings of the 2nd IEEE International Conference on Advanced Computer Control (Icacc 2010), Shenyang, China, 27–29 March 2010.
41. Kong, G.H.; Chen, F.C.; Yang, X.H.; et al. Optimal Deception Asset Deployment in Cybersecurity: A Nash Q-Learning Approach in Multi-Agent Stochastic Games. *Appl. Sci.* **2024**, *14*, 357. https://doi.org/10.3390/app14010357.
42. Haase, H.; Glake, D.; Clemen, T. Multi-Agent Imitation Learning for Agent Typification: A Proof-of-Concept for Markov Games with Chess. In Proceedings of the 2024 Annual Modeling and Simulation Conference, Annsim, Washington, DC, USA, 20–23 May 2024. https://doi.org/10.23919/Annsim61499.2024.10732442.
43. Hernes, M.; Korczak, J.; Krol, D.; et al. Multi-agent platform to support trading decisions in the FOREX market. *Appl. Intell.* **2024**, *54*, 11690–11708. https://doi.org/10.1007/s10489-024-05770-x.
44. Thiel, D.; Hovelaque, V.; Pham, D.N. A multi-agent model for optimizing supermarkets location in emerging countries. In Proceedings of the 2012 IEEE 13th International Symposium on Computational Intelligence and Informatics (CINTI), Budapest, Hungary, 20–22 November 2012.
45. Çevikarslan, S. Research Joint Ventures in an R&D Driven Market with Evolving Consumer Preferences: An Evolutionary Multi-agent Based Modeling Approach. In *Conference of the European Social Simulation Association*; Springer Nature: Cham, Switzerland, 2024. https://doi.org/10.1007/978-3-031-57785-7_16.
46. Agah, A. Robot teams, human workgroups and animal sociobiology: A review of research on natural and artificial multi-agent autonomous systems. *Adv. Robot.* **1996**, *10*, 523–545.
47. Yeung, C.L.; Bunker, R.; Fujii, K. Unveiling Multi-Agent Strategies: A Data-Driven Approach for Extracting and Evaluating Team Tactics from Football Event and Freeze-Frame Data. *J. Robot. Mechatron.* **2024**, *36*, 603–617. https://doi.org/10.20965/jrm.2024.p0603.
48. Song, Y.; Jiang, H.; Tian, Z.; et al. An Empirical Study on Google Research Football Multi-agent Scenarios. *Mach. Intell. Res.* **2024**, *21*, 549–570. https://doi.org/10.1007/s11633-023-1426-8.

49. Ribeiro, A.F.A.; Lopes, A.C.C.; Ribeiro, T.A.; et al. Probability-Based Strategy for a Football Multi-Agent Autonomous Robot System. *Robotics* **2024**, *13*, 5. https://doi.org/10.3390/robotics13010005.
50. Smit, A.; Engelbrecht, H.A.; Brink, W.; et al. Scaling multi-agent reinforcement learning to full 11 versus 11 simulated robotic football. *Auton. Agents Multi-Agent Syst.* **2023**, *37*, 20. https://doi.org/10.1007/s10458-023-09603-y.
51. Gu, C.Y.; De Silva, V.; Artaud, C.; et al. Embedding Contextual Information through Reward Shaping in Multi-Agent Learning: A Case Study from Google Football. In Proceedings of the 2023 Ieee 13th International Conference on Pattern Recognition Systems, Guayaquil, Ecuador, 4–7 July 2023. https://doi.org/10.1109/Icprs58416.2023.10179030.
52. Yu, J.C.; Hu, B.M. Multi-agent Simulation of Knowledge Innovation Activities of University Research Team. In Proceedings of the 2009 International Conference on Industrial and Information, Haikou, China, 24–25 April 2009. https://doi.org/10.1109/Iis.2009.72.
53. Ozsoyeller, D. TAP: Distributed team assignment in heterogeneous multi-agent systems. *Future Gener. Comput. Syst.* **2026**, *174*, 107925. https://doi.org/10.1016/j.future.2025.107925.
54. Qiao, B.; Liu, K.; Guy, C. Multi-agent building control in shared environment. In Proceedings of the International Conference on Enterprise Information Systems, Funchal, Portugal, 12–16 June 2007.
55. Kofinas, P.; Dounis, A.; Korkidis, P. Fuzzy Reinforcement Learning Multi-agent System for Comfort and Energy Management in Buildings. In Proceedings of the Sixth International Congress on Information and Communication Technology: ICICT 2021, London, UK, 25–26 February 2021.
56. Li, X.; Zhang, P.; Gu, Q.; et al. Multi-agent Hybrid Architecture Design for Naval Warfare Game. In Proceedings of the International Conference on Autonomous Unmanned Systems, Shenyang, China, 19–21 September 2023.
57. Panisson, A.R.; Farias, G.P. A Multi-level Semantics Formalism for Multi-Agent Microservices. In Proceedings of the Brazilian Conference on Intelligent Systems, Belém do Pará, Brazil, 17–21 November 2024.
58. Rodriguez-Soto, M.; Lopez-Sanchez, M.; Rodriguez-Aguilar, J.A. Multi-objective reinforcement learning for designing ethical multi-agent environments. *Neural Comput. Appl.* **2023**. https://doi.org/10.1007/s00521-023-08898-y.
59. Yang, Y.; Gao, Y.; Ding, Z.; et al. Advancements in Q-learning meta-heuristic optimization algorithms: A survey. *Wiley Interdiscip. Rev. Data Min. Knowl. Discov.* **2024**, *14*, e1548.
60. Ferdous, J.; Murshed, M.; Meneguette, R.I.; et al. SARSA RL for Edge Connectivity Management in Vehicular Edge Networks. In Proceedings of the 2024 IEEE 13th International Conference on Cloud Networking (CloudNet), Rio de Janeiro, Brazil, 27–29 November 2024.
61. Cai, W.; Chen, H.; Chen, H.; et al. Optimal Defense Strategy for Multi-agents Using Value Decomposition Networks. In Proceedings of the International Conference on Intelligent Computing, Tianjin, China, 5–8 August 2024.
62. Zhang, M.; Tong, W.; Zhu, G.; et al. SQIX: QMIX algorithm activated by general softmax operator for cooperative multiagent reinforcement learning. *IEEE Trans. Syst. Man Cybern. Syst.* **2024**, *54*, 6550–6560.
63. Jiang, C.; Lin, Z.; Liu, C.; et al. MADDPG-Based Active Distribution Network Dynamic Reconfiguration with Renewable Energy. *Prot. Control Mod. Power Syst.* **2024**, *9*, 143–155.
64. Xing, X.; Zhou, Z.; Li, Y.; et al. Multi-UAV adaptive cooperative formation trajectory planning based on an improved MATD3 algorithm of deep reinforcement learning. *IEEE Trans. Veh. Technol.* **2024**, *73*, 12484–12499.
65. Ivison, H.; Wang, Y.; Liu, J.; et al. Unpacking dpo and ppo: Disentangling best practices for learning from preference feedback. *Adv. Neural Inf. Process. Syst.* **2024**, *37*, 36602–36633.
66. Tang, C.; Peng, T.; Xie, X.; et al. 3D path planning of unmanned ground vehicles based on improved DDQN. *J. Supercomput.* **2025**, *81*, 1–31.
67. Cao, J.; Dong, L.; Yuan, X.; et al. Hierarchical multi-agent reinforcement learning for cooperative tasks with sparse rewards in continuous domain. *Neural Comput. Appl.* **2024**, *36*, 273–287.
68. Zhao, R.; Liu, Y.; Wang, X. Graph limit and exponential consensus for the large-scale multi-agent system with delay. *Commun. Nonlinear Sci. Numer. Simul.* **2025**, *149*, 108919.
69. Liu, R.; Piplani, R.; Toro, C. A deep multi-agent reinforcement learning approach to solve dynamic job shop scheduling problem. *Comput. Oper. Res.* **2023**, *159*, 106294.
70. Velasquez, A.; Bissey, B.; Barak, L.; et al. Multi-agent tree search with dynamic reward shaping. In Proceedings of the International Conference on Automated Planning and Scheduling, Virtual, 13–24 June 2022; pp. 652–661.
71. Liu, Z.; Yu, C.; Yang, Y.; et al. A unified diversity measure for multiagent reinforcement learning. *Adv. Neural Inf. Process. Syst.* **2022**, *35*, 10339–10352.
72. Zhu, X.; Liu, J.; Zhang, T.; et al. CPPU: Policy Space Diversity for Informative Path Planning and GAI-enabled Updating CSI in ISAC. *IEEE Trans. Cogn. Commun. Netw.* **2024**, *11*, 777–790.
73. He, W.; Tan, J.; Guo, Y.; et al. A deep reinforcement learning-based deception asset selection algorithm in differential games. *IEEE Trans. Inf. Forensics Secur.* **2024**, *19*, 8353–8368.

74. Abouelazm, A.; Michel, J.; Zöllner, J.M. A review of reward functions for reinforcement learning in the context of autonomous driving. In Proceedings of the 2024 IEEE Intelligent Vehicles Symposium (IV), Jeju Island, Republic of Korea, 2–5 June 2024; pp. 156–163.
75. Feng, Z.; Hu, G.Q.; Dong, X.W.; et al. Adaptively Distributed Nash Equilibrium Seeking of Noncooperative Games for Uncertain Heterogeneous Linear Multi-Agent Systems. *IEEE Trans. Netw. Sci. Eng.* **2023**, *10*, 3871–3882. https://doi.org/10.1109/Tnse.2023.3275326.
76. Sunehag, P.; Lever, G.; Gruslys, A.; et al. Value-Decomposition Networks For Cooperative Multi-Agent Learning Based On Team Reward. *arXiv* **2018**, arXiv:1706.05296.
77. Rashid, T.; Samvelyan, M.; de Witt, C.S.; et al. QMIX: Monotonic Value Function Factorisation for Deep Multi-Agent Reinforcement Learning. *arXiv* **2018**, arXiv:1803.11485.
78. Yang, R.Q.; Yu, L.; Li, Z.; et al. Rethinking Offline Reinforcement Learning for Sequential Recommendation from A Pair-Wise Q-Learning Perspective. In Proceedings of the 2024 International Joint Conference on Neural Networks (IJCNN), Yokohama, Japan, 30 June 2024–5 July 2024. https://doi.org/10.1109/Ijcnn60899.2024.10650400.
79. Suh, J.; Tanaka, T. SARSA(0) Reinforcement Learning over Fully Homomorphic Encryption. In Proceedings of the 2021 SICE International Symposium on Control Systems (SICE ISCS), Tokyo, Japan, 2–4 March 2021.
80. Michailidis, P.; Michailidis, I.; Kosmatopoulos, E. Review and Evaluation of Multi-Agent Control Applications for Energy Management in Buildings. *Energies* **2024**, *17*, 4835. https://doi.org/10.3390/en17194835.
81. Olorunfemi, B.O.; Nwulu, N. Multi-agent system implementation in demand response: A literature review and bibliometric evaluation. *Aims Energy* **2023**, *11*, 1179–1210. https://doi.org/10.3934/energy.2023054.
82. Saldaña, D.; Ovalle, D.; Montoya, A. A Multi-Agent Model to Control Robotic Sensor Networks. In Proceedings of the 2012 7th Colombian Computing Congress (CCC), Medellin, Colombia, 1–5 October 2012.