

TOPICAL REVIEW

Recent Developments of Game Theory and Reinforcement Learning Approaches: A Systematic Review

GARIMA JAIN¹, ARUN KUMAR^{1,2}, AND SHAHID AHMAD BHAT³¹Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh 474005, India²Thapar Institute of Engineering and Technology, Patiala, Punjab 147005, India³LUT Business School, LUT University, 53851 Lappeenranta, Finland

Corresponding author: Shahid Ahmad Bhat (shahid.bhat@lut.fi)

ABSTRACT In the ever-changing world of decision-making, when game theory and reinforcement learning(RL) come together, they create a fascinating combination that shows a new way to solve complex problems in many fields. The combination of game theory and RL is a powerful convergence that opens up a hopeful new frontier for dealing with complex decision-making problems in many different fields. Research on the convergence of game theory and RL has shown to be beneficial, providing essential insights into challenging decision-making issues in various disciplines. This study investigates the recent developments of game theory and RL approaches through a systematic review and highlights the significance of game theory in boosting reinforcement algorithms and increasing the interaction of autonomous vehicles, safeguarding edge caching, and more. It offers a thorough account of the developments at the confluence of game theory and RL. The reviewed papers mainly focus on broad themes and address three important research questions: the impact of game theory on multi-agent reinforcement learning (MARL), the significant contributions of game theory to RL, and the significant impact areas. Following the methodology, search outcomes, and study areas is a discussion on game theory-related terminology, followed by study findings. The review's conclusions offer ideas for further study and open research questions. The importance of game theory in advancing MARL, the potential of game theory in promoting RL strategies, and the opportunities for combining game theory and RL in cutting-edge fields like mobile edge caching and cyber-physical systems(CPS) are all emphasized in the conclusion. This review article advances our knowledge of the theoretical underpinnings and real-world applications of game theory and RL, laying the groundwork for future improvements in decision-making techniques and algorithms.

INDEX TERMS Game theory, reinforcement learning, multi-agent reinforcement learning, decision-making, autonomous vehicles, edge caching, cyber-physical systems.

I. INTRODUCTION

Game theory is a mathematical framework for studying situations in which two or more persons (or “players”) must make decisions that influence each other’s outcomes [1]. In game theory, players are supposed to be rational and self-interested, meaning they would attempt to make the decision that will benefit them the most given the actions of the other players. It gives a method for analyzing the

strategic interactions between players and predicting their likely results.

Reinforcement learning(RL), a branch of machine learning, is concerned with how an agent can learn to take subsequent actions in a setting to maximize cumulative rewards [2]. RL agents interact with the environment discretely. At each time step, the agent observes the environment, follows its policy, and is quantitatively rewarded by the environment. The agent seeks a strategy that maximizes the expected cumulative long-term benefit. In the RL framework, the agent learns via trial and error. Value functions or action-value

The associate editor coordinating the review of this manuscript and approving it for publication was Jiachen Yang¹.

functions, help the agent estimate future rewards for specific states or state-action combinations. These value functions help the agent evaluate and refine the policy.

Due to its capacity to resolve challenging decision-making issues in various fields, such as robotics, natural language processing, and game playing, RL has recently gained popularity. Games may be solved using RL by treating the game environment as an RL problem and training the model with strategies for playing the game. The field of game research has significantly benefited from RL in designing intelligent agents with strong gaming skills. In contrast, the study of strategic decision-making in settings where the results depend on the actions of numerous players is called game theory [2]. Game theory is now applied to investigate a wide range of phenomena, from voting behavior to market competitiveness, and it provides a framework for examining strategic interactions between participants.

The commonality between game theory and RL is their emphasis on decision-making in uncertain, complicated contexts. RL can be used to describe and solve games by treating them as RL problems, where the objective is to identify an optimal policy that maximizes the reward. In contrast, game theory provides a framework for evaluating strategic interactions between players and can assist in forecasting the equilibrium results of these interactions [3]. Combining RL and game theory has been used to create more intelligent agents capable of playing games at a high level, as evidenced by AlphaGo's victory over human Go champions [4]. Game theory can also be used to analyze the behavior of RL agents in social problems when individual and collective goals conflict [5]. Combining RL and game theory can increase the understanding of strategic decision-making in complex and uncertain environments and create more intelligent agents in various contexts.

This paper aims to provide an inclusive insight and a comprehensive overview of the works done at the intersection of game theory and RL, the role of game theory in bettering the reinforcement algorithms, contributions of game theory towards improving the interaction of unmanned vehicles, the role of game theory in securing edge caching and furthermore. The analysis was done by conducting a systematic literature review to address the following research questions(RQ):

RQ1: What are the advancements made in RL via game theory?

RQ2: How game theory has improved Multi-Agent Reinforcement Learning(MARL)?

RQ3: Identifying the major areas of contribution of game theory in RL.

The remaining paper is structured as follows: Section II explains the methodology followed while conducting the review. Section III describes the patterns observed in the search results and the research fields that came forward. Section IV defines some basic terms related to RL and game theory to familiarise the reader with the research topic. Section V summarizes the complete paper readings, bringing

TABLE 1. Critical analysis.

Author[Ref]	MARL	NE	Cooperative Games	Non-Cooperative Games	Zero-sum Game	Markov Decision Process
Yang and Wang [6]	✓	✗	✗	✗	✓	✗
Yunlong and Kai [7]	✗	✓	✗	✗	✗	✗
Gao [8]	✗	✗	✓	✓	✗	✗
Zhang et al. [9]	✗	✗	✗	✗	✗	✗
Moll and Kunczik [10]	✓	✗	✗	✗	✗	✗
Sharma and Gopal [11]	✓	✗	✗	✗	✗	✓
Our Survey	✓	✓	✓	✓	✓	✓

forth the results and future scopes in the field. Section VI identifies the potential areas of future research, followed by the research gaps and future research opportunities in Section VII. We wind up the paper with conclusions and future work in section VIII.

II. CRITICAL ANALYSIS

This paper stands out from other review papers as it compares the various areas in which game theory and RL have proven beneficial on the computer science front. The paper provides a significant headstart to future researchers who are new to the field and want a collective overview of the current state-of-the-art methods, advancements, and future scopes. Yang et al. made a similar effort, giving a detailed overview of MARL and game theory and reviewing the developments of that time [6]. Yunlong and Kai, in their paper, discuss solution algorithms from three aspects: independent RL, RL with Nash equilibrium(NE), and beyond ash equilibrium [7]. In his paper, Gao discusses the intersection of RL and game theory, limiting game theory to cooperative and non-cooperative games [8]. Other papers restrict the discussion to robotics and game theory [9] or finding common ground between the two fields [10], [11]. Our paper firmly sets a base by explaining the fundamentals of the two fields with a thorough literature review, providing future researchers with a solid start for their work.

III. PRELIMINARIES: UNDERSTANDING GAME THEORY AND RL

This section gives definitions of the terms very commonly used while referring to Game Theory and RL.

A. GAME THEORY

Game is a description of a strategic interaction that also contains the interests of the participants and the limitations on the activities they are permitted to do, but it does not identify the actual actions that the players take [12].

Players are the individuals taking part in the game [12].

Strategy is a comprehensive course of action that a player follows throughout the game, given some prior knowledge

about the game and the other player's strategy [13]. A strategy is often defined as a plan of action intended to accomplish a specific goal [14].

A pure strategy for player i is a deterministic plan of action [14]. The set of all pure strategies for player i is denoted S_i . A profile of pure strategies $s = (s_1, s_2, \dots, s_n)$, $s_i \in S_i$ for all $i = 1, 2, \dots, n$, describes a particular combination of pure strategies chosen by all n players in the game. The word 'pure' indicates the idea of following a particular action plan.

Payoff is the consequences received by the players in response to their actions [15]. Payoffs can be rewards, utility, or any other measurable benefit that the players feel due to their actions.

A payoff function $u : X \rightarrow R$ represents the preference relation \succeq if for any pair $x, y \in X$, $u(x) \geq u(y)$ if and only if $x \succeq y$ [14]. The preference relation \succeq translates that the payoff function $u : X \rightarrow R$, assigns to every outcome in X a real number, iff the function assigns a comparatively higher value.

A rational player chooses the action that gives him the highest possible payoff from the possible set of actions at his disposal [14]. Hence by maximizing his payoff function over his set of alternative actions, a rational player will choose his optimal decision.

NE is an action profile a^* with the property that no player i can do better by choosing an action different from a_i^* , given that every other player j adheres to a_j^* [1].

In the idealized setting in which the players in any given play of the game are drawn randomly from a collection of populations, NE corresponds to a steady state. If, whenever the game is played, the action profile is the same NE a^* , then no player has a reason to choose any action different from her component of a^* ; there is no pressure on the action profile to change. Expressed differently, NE embodies a stable "social norm": if everyone else adheres to it, no individual wishes to deviate from it.

NE is a solution concept in game theory that describes a set of strategies, one for each player, such that no player can improve their payoff by unilaterally changing their strategy [16].

The pure-strategy profile $s^* = (s_1^*, s_2^*, \dots, s_n^*) \in S$ is a NE if s_i^* is a best response to s_{-i}^* , for all $i \in N$, that is,

$$v_i(s_i^*, s_{-i}^*) \geq v_i(s_i', s_{-i}^*) \quad (1)$$

for all $s_i' \in S_i$ and all $i \in N$ [14].

Cooperative Games By using the term cooperative, we mean to imply that the players have complete freedom of communication and complete information on the structure of the game [17]. Furthermore, there should be the possibility of making enforced agreements, binding either one or both players to a certain agreement or policy. It is assumed that either player may secure a commitment (enforced policy contract) upon himself if he so desires. But each player is supposed not to have any commitments upon himself before entering into the negotiation involved in this game, or at least none relevant to the situation.

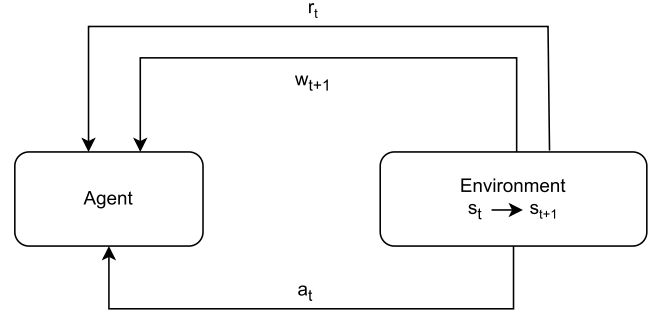


FIGURE 1. Possible results of actions.

Non-cooperative Games focus on simulating how agents behave in a defined process as they try to maximize their utility [18]. In Non-cooperative game theory (NCGT), agents make decisions independently, and the analysis is based on a thorough description of the actions and data that each agent has access to. It emphasizes individual decision-making and strategic behavior without assuming any formal coordination among the agents.

Stackelberg Games A Stackelberg game is a two-player extensive game with perfect information in which a "leader" chooses an action from a set A_1 and a "follower", informed of the leader's choice, chooses an action from a set A_2 [12]. The solution usually applied to such games in economics is that of subgame perfect equilibrium (though this terminology is not always used). Some (but not all) subgame perfect equilibria of a Stackelberg game correspond to solutions of the maximization problem

$$\max_{(a_1, a_2) \in A_1 \times A_2} u_1(a_1, a_2) \text{ subject to } a_2 \in \arg \max_{a_2' \in A_2} u_2(a_1, a_2') \quad (2)$$

[12]

where u_i is a payoff function that represents the player i 's preferences. If the set A_i of actions of each player i is compact and the payoff functions u_i are continuous then this maximization problem has a solution.

B. REINFORCEMENT LEARNING

State presents the current situation of the model under consideration. It consists of a set of variables describing the environment.

Action is the decision made by the agent in response to the current state [19]. It is similar to deciding upon a choice from an available set of options. At each time step t , the agent has to take an action $a_t \in A$. As illustrated in Fig 1 it follows three consequences: (i) the agent obtains a reward $r_t \in R$, (ii) the state transitions to $s_{t+1} \in S$, and (iii) the agent obtains an observation $w_{t+1} \in \Omega$.

Reward is a positive response given by the environment to the agent against the action taken [19]. $R: S \times A \times S \rightarrow R$ is the reward function, where R is a continuous set of possible rewards in a range $R_{max} \in R^+$ (e.g., $[0, R_{max}]$) [19]. At each time step t , the probability of moving to s_{t+1} is given by the

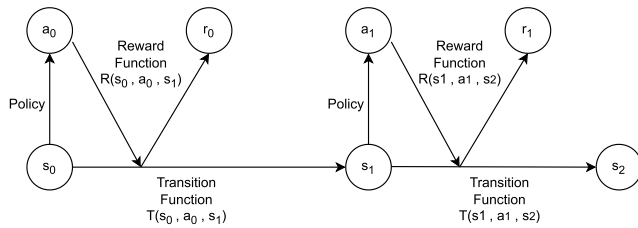


FIGURE 2. At each step, the agent takes an action that changes its state and gives a reward.

state transition function $T(s_t, a_t, s_{t+1})$ and the reward is given by a bounded reward function $R(s_t, a_t, s_{t+1}) \in \mathbb{R}$. Refer to Fig 2

Policy governs an agent's choice of activities. Policies can be categorized as either stationary or non-stationary based on this criterion. When an agent tries to maximize cumulative rewards across a finite number of future time steps, a non-stationary policy is advantageous since it relies on the time steps [20].

Value Function is an estimated measure of how beneficial it is for the agent to be in a particular state [21]. The value of a state s under policy π , denoted $V^\pi(s)$ is the expected return when starting in s and following π thereafter.

Markov Decision Process (MDP) is an environment, modeled as a set of states and actions that are performed to control the system's state [21]. The goal is to control the system such that the performance is maximized. MDPs have modeled stochastic planning, learning robot control, and game-playing challenges.

IV. METHODOLOGY

In this study, we use the Systematic Literature Review (SLR) approach, often viewed as a benchmark to review existing literature and methodology. Fig 3 represents the SLR framework and the review process as a whole. The review process spreads over three phases. Phase I includes finalizing the keyword combination and carrying out the search over various databases available. Phase II witnesses the segregation and sorting of the acquired papers. Finally, in phase III, we conduct complete paper readings of the selected articles and analysis to meet the study objectives. The following subsections give a brief description of the entire review process.

A. KEYWORD COMBINATION

Acquiring a worthy collection of research articles is a crucial part of the SLR review methodology. "Reinforcement Learning" and "Game Theory" are the principally emphasized keywords for this paper. The combination of the keywords and operators used for the paper is TITLE (reinforcement learning) AND TITLE (game theory) to restrict the search of the keywords only to the title in Google Scholar, Web of Science, and Scopus. The AND operator ensures that the search results include the terms "reinforcement learning" and "game theory" in the title.

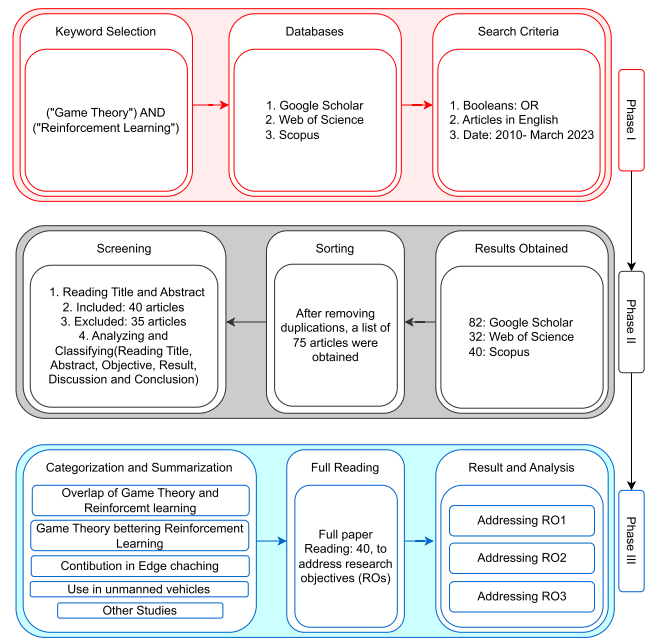


FIGURE 3. Systematic flow of the conducted literature review.

B. DATABASES

Selecting credible sources is essential when doing a literature review to guarantee that the contents reviewed are of sufficient quality and relevance. This paper uses a holistic approach by searching three prestigious databases: Google Scholar, Web of Science, and Scopus. Each of these databases has its strengths and areas of focus, and together, they provide for an extensive literature assessment. We aim to achieve a comprehensive and diverse collection of sources by incorporating these three databases into our literature review process.

C. ARTICLE SCREENING AND SORTING: INCLUSION AND EXCLUSION CRITERIA

Formerly, we searched all three databases for the keyword combination ("Game Theory") AND ("Reinforcement Learning") inclusive of all the work done since 2010. The search results were as follows: 82 results from Google Scholar, 32 from Web of Science, and 40 from Scopus. Since we found only a few documents on the research topic from before 2010 in the databases Fig 4, we decided against including them in our review paper. Sorting and removing the duplications left us with 75 research papers. Moving forward with screening, we selected papers based on analysis, including reading the titles, abstract, results, and discussion. We included a total of 40 articles after the screening process. Inclusion criteria included (i) Bettering RL algorithms using game theory and (ii) Advancements in computer science. When conducting a literature review, choosing reliable sources is essential to ensure that all the literature reviewed is relevant. This paper uses three databases: Google Scholar, Web of Science, and Scopus.

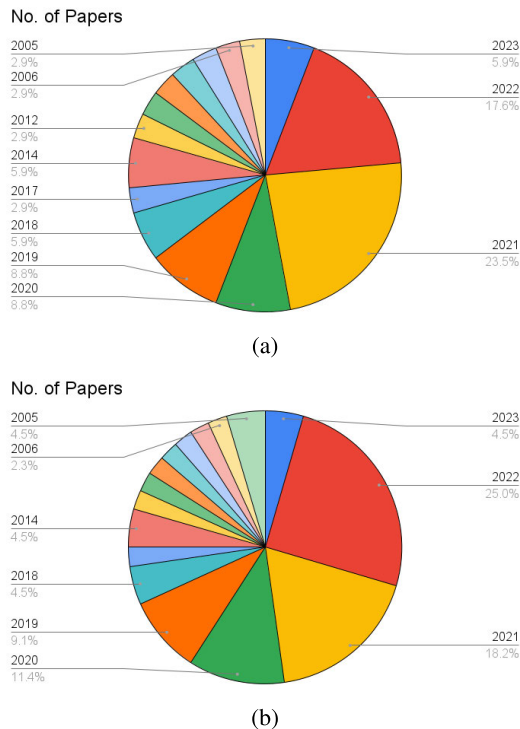


FIGURE 4. Distribution of research papers under the keyword search TITLE (reinforcement learning) AND TITLE (game theory) over years: Fig 4a Search from Web of Science Fig 4b Search from Scopus.

D. CRITERIA OF CATEGORIZATION OR CLASSIFICATION OF REVIEWED PAPERS

Screening and reading the shortlisted papers brings forward a pattern. Hence, the papers could be broadly divided into the following categories: papers discussing advancement in RL through game theory, advances in MARL using game theory, enhancement in edge caching, enhancements in unmanned vehicles and handling the vehicular traffic and modeling Cyber-Physical Human Systems(CPHS).

V. CLASSIFICATION OF SEARCH RESULTS

This section accounts for the various fields in which game theory and RL have cumulatively enhanced the scenario. Fig 5 presents the many areas influenced by game theory and RL. Out of the many fields, we selected a few that piqued our interest. We mainly picked up fields in Computer Science and Engineering like enhancement of CPHS, Traffic management and control (vehicular and network), intrusion detection in cyber security, computations in multi-agent systems, and mobile edge-caching. The most commonly seen game theory concepts are zero-sum games, NE, cooperative and non-cooperative games, and payoff optimization.

Table 2 shows the distribution of the referred papers among the various publishers.

VI. LITERATURE REVIEW

We comprehensively reviewed various papers on the intersection of game theory and RL. It gave an insight into how different game theory algorithms are used to solve the

TABLE 2. Major publishers.

Publisher	Count
ARC	1
arXiv	2
Elsevier	4
IEEE	15
Inderscience	1
IOP Publishing	1
MDPI	2
Reno ProQuest Dissertations Publishing	1
SAE	1
SPIE	1
Wiley	1

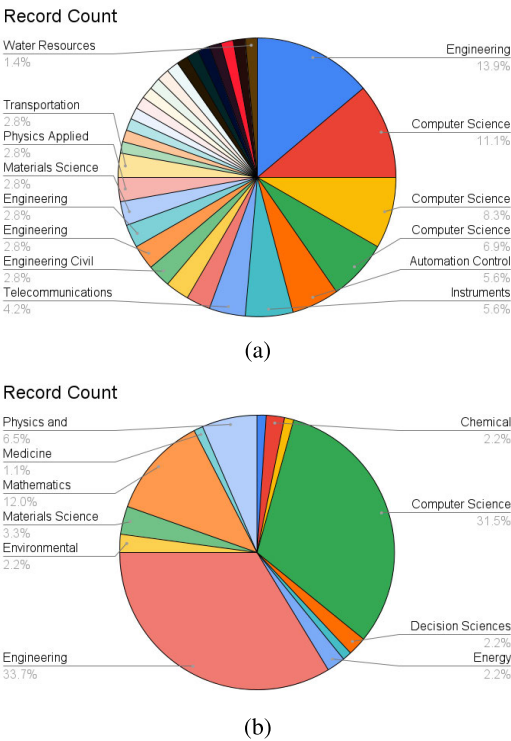


FIGURE 5. Distribution of research papers in various fields: Fig 5a Search from Web of Science Fig 5b Search from Scopus.

problem. We identified some major fields in which we were able to categorize the papers.

CPHS is a significant field of interest for researchers. Jin et al. address the optimal policy selection problem of attackers and sensors in cyber-physical systems (CPSs) under denial of service (DoS) attacks. The sensor-attacker game is described as a two-player zero-sum game. RL algorithm is designed for dependable and unreliable channels. The simulation results of both scenarios show that the proposed RL method can swiftly converge to both sides' NE [22]. Schelble et al. explores how recent RL methods and game theory scenarios affect human-machine team cooperation levels. Three RL algorithms, namely Vanilla Policy Gradient, Proximal Policy Optimisation(PPO), and Deep Q-Network(DQN) and two game theory situations, namely Hawk Dove and Prisoners

dilemma, were tested in a large-scale experiment. Results show that each RL model interacts differently with humans, with Deep-Q promoting greater cooperation [23]. Khoury and Nassar suggest two complimentary cybersecurity risk assessment methods for comprehensive CPS network evaluation. First, they present a game-theoretical model for Industrial Control System (ICS) cybersecurity using Monte Carlo simulations to evaluate payoffs for variable randomness, strategies, budget expenditure, and look-ahead. Secondly, they improved the CPS security framework by combining game theory and MARL ideas at strategic and battlefield levels [24]. Albaba and Yildiz propose a computationally feasible approach to model multiple humans as decision-makers, simultaneously instead of determining the decision dynamics of the intelligent agent of interest [25]. Table 3 presents our literature review for papers where game theory and RL were used to enhance CPHS.

Edge caching and edge computing became another field of interest for researchers. Wang et al. combine Multi-Access Edge Computing (MAEC) with artificial intelligence (AI) to provide a potential task-offloading strategy. They propose a task-offloading decision mechanism (TODM) based on cooperative games and Deep Reinforcement Learning (DRL) [26]. Wang et al. propose game theory-based distributed edge computing server task scheduling. The suggested method balances mobile device-server link quality and computing resource allocation while selecting edge computing servers. They create an IndRNN-LSTM time series prediction method that accurately forecasts link quality [27]. Li et al. explore the offloading decision problem in a software-defined networking-driven MAEC system with numerous users and servers. They optimize server selection, offload data size, and compute service price to maximize MAEC server profit. According to them, the problem in an MAEC context is to find the best strategy for dynamic and stochastic end-users. MAEC server selection is addressed via PPO reinforcement learning. A two-step optimization process determines the offloading data size and computing service pricing [28]. Liang et al. study the MEC's interference-aware multi-user computation offloading. A multi-user computation offloading game model analyses NE for this problem. They create the computation offloading approach using NE and RL to reduce system overhead. We gain a DRL algorithm to avoid dimensional crises by adding neural networks to Nash-Q-learning [29]. Xu et al. suggest a secure edge caching strategy for Mobile Social Networks (MSN) content providers and mobile consumers. The Stackelberg game represents the interactions between content providers and edge caching devices. Q-learning determines the best payment method for content providers and security strategies for edge caching devices [30]. Table 4 presents the literature review of papers representing the enhancement of edge caching methods using game theory and RL algorithms.

Ways to enhance the automation of unmanned vehicles and traffic control became another widely researched field.

In their paper, Xue et al. discuss a mechanism to control Multiple Unmanned Vehicles (MUV) in stable formation using game theory and RL, where they create a MUV model and convert formation tracking into an optimization problem [31]. Fu et al. optimize power distribution and 3D deployment of multiple Ultra-dense unidentified aerial vehicle base stations (UAV-BSs). They model the power allocation problem as a non-cooperative game with a pricing mechanism miming UAV-BSs-served user interactions for efficient interference management. UAV-BS power distribution and 3D deployment are turned into Markov decision problems, which use price-based proximal policy optimization (3PO) to find the best system throughput policy [32]. Cipolina-Kun solves the coalition formation problem and proposes a socially optimal rider-car allocation mechanism. An equitable cost-sharing method is proposed to encourage ridesharing [33]. Li et al. offer a data-driven background vehicle (BV) behavior model for virtual autonomous vehicle testing. The technique represents the merging vehicle decision process as a regular MDP. Deep maximum entropy-inverse reinforcement learning and the game matrix are used to find the reward function for BV behavior. BV is simulated using a deep Q-network technique based on the reward function [34]. According to Yin et al., air warfare is competitive, and the opponent needs to be more explicit, making it hard to choose the best approach. The DRL-game theory algorithm solves the problem of existing methods failing to solve NE strategy in highly competitive environments [35]. Using game theory and RL, Duan et al. propose an automatic drive model for multi-agent autonomous driving in challenging traffic conditions. By extending the game description language, this model enables strategic reasoning and negotiation in traffic scenarios. The automatic drive model based on game theory and RL is proposed and applied to multi-agent cooperative driving to enable strategic reasoning with negotiation in traffic scenarios by extending the game description language and proposing the constrained multi-agent deep deterministic policies gradient (CMADDPG) algorithm [36]. Zwillinger et al. aims to quickly broadcast all Mobile Ad-hoc Network (MANET) messages and reduce their number. RL and game theory are compared for agent choice optimization. The RL framework treats MANET nodes as RL agents that must learn message timing and recipients. The game theory framework game tree comprises nodes for message knowledge and connectivity information and decision branches for neighbor messaging [37]. Zhan et al. demonstrate how kinematics, game theory, and RL-based optimization can address the conflict of interest between lane-changing and target lane cars in fully autonomous driving. The reward function is created using the kinematic method to calculate the payoff value for the two vehicles' strategy combinations [38]. Guo developed an intelligent traffic signal control model utilizing a game-theoretical framework, treating inbound links as players and signal light status as the decision. This model adopted the baseline-constant technique, ensuring consistent

and periodic red and green light intervals. He created an RL system for traffic light control and single-agent RL to address the challenges of learning and computation in game theory-based methods. Increased traffic density in the prior decomposition approach leads to longer execution times. Thus, a DQN approach optimized computational time, albeit with reduced efficiency [39]. Liang et al. study how an IDS recognizes environmental change and adapts to different environments. The Vehicle Ad-hoc Network(VANET)-based IDS GaDQN-IDS uses Bayesian Game theory and DQN. They call IDS-attacker interactions a dynamic intrusion detection game in which the IDS changes the accuracy-efficiency tradeoff or is retrained as its detection capability drops [40]. Albaba and Yildiz describe a stochastic modeling method using DRL and game theory for simultaneous decision-making in multi-agent traffic scenarios, which combines level-k reasoning with DQN reinforcement learning [41]. Abdoos' article creates intersection multi-agent traffic signal controllers. A two-mode agent architecture with independent and cooperative methods is offered for traffic congestion control. Individual Q-learning agents regulate each junction. Game theory determines how agents might collaborate in cooperative mode to regulate traffic signals at multiple crossings [42]. Yildiz et al. aim for Next Generation Air Traffic System capacity and safety. Their approach models human goals and predicts actions using RL instead of explicitly modeling them [43]. Table 5 presents the literature review in the field discussed above.

Extending single-agent RL to MARL has long been an active subject. In his paper, Arend et al. introduce us to MLPro. In the first subtopic, MLPro-RL for reinforcement learning, the abstract ML model becomes a landscape of model-free or model-based single-/multi-agents, and the surroundings become environments. In the second subtopic, package MLPro-GT for cooperative game theory, large parts of MLPro-RL are taken up, transferred to the established terminology, and further specialized [44]. In an automated warehouse scenario, Ho et al. consider 5G service provisioning using swarm robotics managed by an industrial controller that offers routing and task instructions over the 5G network. They employ coordinated multipoint(CoMP) and ultra-reliable low-latency communication(URLLC) beamforming to manage robots in an automated warehouse for goods storage while following planned reference tracks [45]. Teymoori and Boukerche examine multi-user compute offloading to a single-cell edge server in a dynamic setting. They employ game theory to characterize the compute offloading choice process as a stochastic game to limit user interference on wireless channels. Next, they use a payoff-based MARL to verify the suggested game model's NE exists [46]. Farabaksh discusses the conflict between profit-seeking people's individual advantage and the resource's persistence-based collective well-being. Game theory and RL have been used to represent rational agent decisions in these systems. Research shows that human learning models can significantly

impact common-pool resource systems and generate sustainable results among self-interested agents [47]. Zhou discusses the challenges faced by multi-agent systems and introduces mean field game theory to generate a decentralized optimal control framework named the Actor-critic-mass (ACM) [48]. Liu et al. offer game theory and RL-based radar network time scheduling for inverse synthetic aperture radar imaging with targets in various radar beams. The optimization problem's game behavior is analyzed, and a time-scheduling game is created to find the strategy [49]. Adams et al. use DQN for policy learning and NE for action selection to solve implicit coordination problems in Multi-agent Deep Reinforcement Learning(MADRL), including non-stationarity and exponential state-action space growth. Joint action selection is based on Nash Q-values proxy payoffs and mutual optimum replies. Coordination is implicit since multiple/no Nash equilibria are resolved deterministically [50]. Morrison studies the problem of automated object manipulation and becomes the first to use a game theory-based strategy for dual-arm manipulation. Each arm learns independent policies and is trained using the Twin Delayed Deep Deterministic Policy Gradients(TD3) algorithm utilizing two neural networks [51]. Tang and Dong propose a MAXQ method to improve the speed and efficiency of task exploration in multi-robot systems [52]. Table 6 presents our literature review on MARL enhancement.

We came across some papers that did not fit into any one of these categories, yet they were an essential part of the literature review. Zhu et al. use RL and game theory to find the optimal strategy for the 2-player simultaneous Pig game, including the optimal strategy against a specific independent strategy and the NE mixed strategy. They offer a new Stackelberg value iteration for multi-agent (SVIMA) RL to find a near-optimal approach. The best strategy for the 3-player simultaneous game under the independent strategy setting is then developed, followed by the extension of the n-player optimal strategy [53]. Yunxiu and Kai stress recognizing other players' ambitions and considering player deception.

By abstracting deception features, the general deceptive behavior model is used to design a behavior plan that best matches the deceiver's past behavior data using inverse reinforcement learning (IRL) [54]. Yifei and Lakshminarayanan plan to configure RL agents for MARL control systems to show their efficacy in managing multiloop processes with significant interconnections. They offer an RL agent configuration with a feedback controller and decoupler functions in a control loop. After deploying two agents, they create a MARL system that learns to regulate a two-input, two-output system with strong interactions. In systems with weak to moderate loop interactions, the performance of the MARL system is only a little affected by the reward function setup. MARL with mixed strategies outperforms pure cooperation in systems with strong loop interactions [55]. Liu et al. construct joint optimization as a mixed integer nonlinear programming (MINLP) issue, which can also be a complex

TABLE 3. Literature review on CPS.

Author	Research Focus	Methodology	Proposed Algorithm	Results Findings
Jin et al. (2022) [22]	Optimal policy selection in DoS attack	Interaction between attacker and sensor formulated as a zero-sum game	RL algorithms for Reliable and Unreliable channel	Algorithms quickly converge to NE
Schellble et al. (2021) [23]	Human-machine cooperation levels	Various RL and game theory scenarios are examined		social and task framing in human-artificial intelligence systems is necessary
Khoury and Nassar (2020) [24]	Cyber-security	Two complementary risk assessment approaches are proposed		Defender can learn optimal policies to counter viruses
Albaba and Yildiz (2019) [25]	CPS outcomes	RL is employed to predict the time-extended interaction dynamics		Enhanced performance compared to traditional algorithms

TABLE 4. Literature review on MEC.

Author	Research Focus	Methodology	Proposed Algorithm	Results Findings
Wang et al. (2023) [26]	Task Offloading	MEC with AI as a task-offloading strategy	TODM	Stable convergence attained
Wang et al. (2022) [27]	Scheduling server tasks	Game theory-based distributed edge computing server task scheduling	Ind-LSTM	Outperforms other time series prediction algorithm
Li et al. (2021) [28]	Computation offloading policy	A two-step optimization issue determines offloading data size and survive pricing	PPO RL framework	NE achieved relatively faster
Liang et al. (2020) [29]	Multiuser computation offloading	A computation offloading algorithm called Nash-Q-learning is proposed to minimize system overhead.	Computation offloading algorithm	System overhead reduced considerably
Xu et al. (2020) [30]	Edge-buffering mechanism	Stackelberg game models content provider-edge caching device interactions	Q-Learning	Edge caching device able to provide high-quality secure services

TABLE 5. Literature review on automation control.

Author	Research Focus	Methodology	Proposed Algorithm	Results Findings
Xue et al.(2023) [31]	Manage stable formation of MUVs	Formation tracking is made into an optimization problem	Real-time Leader Switching algorithm	Tests with simulation tools prove effective
Fu et al. (2023) [32]	Optimizing power allocation	The problem made into a Pricing factor optimization	3PO algorithm	Considerably better than base algorithms
Cipolina-Kun (2023) [33]	Sustainable Mobility Service	<ul style="list-style-type: none"> Coalition Formation problem solved Optimal rider-car allocation mechanism is proposed 	Centralized approach	Large number of users pose a challenge
Li et al. (2023) [34]	Safety methods for autonomous vehicles	Vehicle decision making formulated as MDP	Deep Maximum Entropy Inverse RL	Accurate replication of virtual-vehicle engagement behavior
Yin et al. (2022) [35]	Uncertainty of air combat environment	1v1 air combat modeled as two-player zero-sum Markov game	DQN	Algorithm shows convergence
Duan et al. (2022) [36]	Autonomous driving for traffic safety	An extension of actor-critic strategy gradient method	C-MADDPG	The algorithm works in competitive and cooperative environments
Zwillinger et al. (2022) [37]	Route optimization	<ul style="list-style-type: none"> Game tree is created Solved using MCTS version 	Monte-Carlo Tree Search	Able to identify optimal paths
Zhan et al. (2022) [38]	Vehicle lane changing	<ul style="list-style-type: none"> Payout function formed using kinematics Game payoff matrix determines the strategy 		Cars select best strategy combination
Guo (2022) [39]	Increasing Algorithm complexity with traffic density	A novel semi-cooperative Nash Q-learning approach	Request-respond mechanism	Increasing lane changing efficiency
Liang et al. (2021) [40]	Self-adaptive IDS for VANETs	A dynamic game describes IDS-attacker interactions	GaDQN-IDS	IDS can retrain itself and detect attacks well
Albaba and Yildiz (2021) [41]	Addressing human-human and human-automation interactions	Game theoretical driver models are compared to real human driving behavior patterns	Level-k reasoning	GT policies perform 52% better than a benchmark policy
Abdoos (2020) [42]	Traffic congestion problem	Game theory dynamically controls traffic signals	Q-learning cooperative games	The agent must cooperate to reduce average traffic delay
Yildiz et al. (2014) [43]	Pilot automation	RL predicts actions based on modeled goals	RL pilot model	Introducing self-automation proves beneficial

multiple resource allocation (MRA) optimization problem with distinct allocation constraints. They employ DRL to forecast future rewards of actions based on user data from the networks and present single-layer MRA methods based on DQN and deep deterministic policy gradient (DDPG) for downlink wireless broadcasts. We also present a two-layer iterative solution to tackle the NP-hard MRA problem using the data-driven DQN and NCGT model, which can increase

communication performance [56]. Zheng et al. claim that the hierarchical actor-critic interaction of actor-critic RL algorithms calls for a game-theoretic interpretation. Theory-wise, they develop a policy gradient theorem for the refined update and ensure local convergence of the Stackelberg actor-critic algorithms to local equilibrium [57]. Salgado and Clempner use NCGT to describe agent interaction and RL to introduce environmental inputs. Markov chains

represent emotions as states with probabilities. They describe agent interaction using NCGT and solve the game using a novel two-step method. An RL mechanism to provide inputs to the environment is also developed [58]. Zazo et al. analyze dynamic potential games, which can be solved by a single multivariate optimum control problem. Four different problems are solved using analysis and numerical methods namely Energy demand control in smart-grid networks, network flow optimization with limited relay capacity and battery life, uplink multiple access communication for battery optimization, and two optimal scheduling games with time-varying channels [59]. In his paper, Marden introduces a new framework, state-based potential games, introducing an underlying state space to potential games. This state space increases system designers' flexibility to coordinate group behavior and overcome restrictions [60]. Arthurs and Birnbaum ML technique generates a function Q that approximates the expected utility of action and state s , given the rules of a two-player adversarial game. They train the Q -function neural network using RL using selfplay data [61]. These papers are briefed in Table 7.

VII. POPULAR GAME THEORY APPROACHES IN RESEARCH

The literature review brought forward some game theory algorithms that proved the most effective when combined with RL. These included Cooperative, Non-Cooperative, and Stackelberg games, which we already discussed in Section 2.2. We identified some other frequently encountered game theory approaches.

A. POTENTIAL GAMES

In exact potential games, the potential function shifts by the same amount whenever a single player's utility shifts, owing to a divergence in strategy [62].

The game G is an exact potential game if and only if a potential function $F(S) : S \mapsto \mathbb{R}$ exists such that $\forall i \in N$

$$U_i(T_i, S_{-i}) - U_i(S_i, S_{-i}) = F_i(T_i, S_{-i}) - F_i(S_i, S_{-i}) \\ \forall S_i, T_i \in S_i, \forall S_{-i} \in S_{-i} \quad (3)$$

[63]

B. FICTITIOUS SELF PLAY

The concept of fictitious play pertains to a dynamic iterative procedure when participants in a recurrent game modify their strategy in accordance with their perceptions of the opponents' prior actions. This dynamic method is commonly employed within the framework of non-cooperative games [64].

C. STOCHASTIC GAMES

In a stochastic game the play proceeds by steps from position to position, according to transition probabilities controlled jointly by the two players [65].

D. MARKOV PROCESS

A group of Markov processes, denoted as $M_{x,y}$, is provided with two indices. Each instance of a Markov process, representing different states or histories, is associated with a payment from player B to player A. Player A selects the value of x in order to optimize the reward, whereas player B selects the value of y to minimize it [66].

E. MEAN-FIELD GAME THEORY

The mean-field game (MFG) theory focuses on examining the presence of Nash equilibria in games that involve a substantial number of agents that are modeled as controlled stochastic dynamical systems and whose individual contributions are considered to be asymptotically unimportant. This is accomplished by leveraging the connection between the finite population problem and its related infinite limit counterpart [67].

F. ZERO-SUM GAME

A two-player game is zero-sum if the sum of the payoffs to the two players at each terminal node is zero [68].

G. BAYESIAN GAME

Bayesian games are a category of games within the realm of game theory that expands upon the conventional notion of strategic games by incorporating elements of uncertainty and incomplete knowledge. In the context of Bayesian games, participants possess individual beliefs or private information regarding the attributes of the game. These views are quantitatively expressed using probability distributions, which serve as mappings from preferences to distributions across possible actions [69].

VIII. IDENTIFYING POTENTIAL AREAS

A. MULTI-AGENT REINFORCEMENT LEARNING APPROACH

The merging of game theory and RL has greatly improved MARL methods. Game theoretical ideas have improved MARL algorithms' capacity to handle complex problems across various fields.

In Radar networking, an equilibrium-based MARL approach is proposed to find the optimal strategy profile so that scheduling can be done to capture high-resolution images [49]. Integration of game theory and RL to propose an algorithm that combines Q-learning and NE for coordination in multi-agent RL results in better convergence than traditional NE. [50]. MARL strategy for automated object manipulation, referring to the case of using the two arms of a Bexter robot, has enabled the coordination of both arms to execute tasks collaboratively [51]. Each arm is treated as an independent agent, and coordination is achieved using a game-theory based distributed coordination strategy. Computation Offloading is also explored through game theory [46]. Formulating the offloading process as

TABLE 6. Literature review on MARL.

Author	Research Focus	Methodology	Proposed Algorithm	Results Findings
Arend et al. (2022) [44]	Designing of MLPro 1.0	An abstract adaptive ML model interacts with specific scenarios and is trained and tuned on it	MLPro-RL MLPro-GT	Accelerate the development of new subtopics
Ho et al. (2022) [45]	5G service provisioning	A time-varying joint CoMP and 5G URLLC design problem is formulated	CoMP clustering with PPO	Outperforms baseline algorithm
Teymoori and Boukerche (2022) [46]	Computation Offloading	Users' computation offloading decision process is formulated as a stochastic game model		Comparison with conventional methods proves efficiency of model
Farahbakhsh (2021) [47]	Shared resource optimization	Decisions of rational agents modeled from a game theoretic and RL approach		Human learning models greatly affect common-pool-resource systems
Zhou (2021) [48]	MAS optimal control problem	Mean field game theory is introduced to generate a decentralized optimal control framework	ACM	Target Quality-of-Service achieved
Liu et al (2020) [49]	Radar-network time scheduling	Time scheduling game formulated and analyzed	Equilibrium based MARL	Algorithm attains NE and optimizes time scheduling
Adams et al. (2020) [50]	Implicit coordination in MADRL	DQN combined with NE		Simulation testing yields suboptimal results
Morrison (2020) [51]	Automated object manipulation	Individual agents trained using standard TD3 algorithm	TD3 algorithm	
Tang and Dong (2019) [52]	Task handling in fuzzy models	MAXQ is used to adjust the parameters of the collaborative rules	HRL based MAXQ	Higher convergence speed

TABLE 7. Literature review in miscellaneous areas.

Author	Area of Focus	Research Focus	Methodology	Proposed Algorithm	Results Findings
Zhu et al. (2023) [53]	Dice game Pig	Strategy selection	Strategy selection analytically and using framework	SVIMA	
Yunxiu and Kai (2023) [54]	RTS games	Role recognition	goal recognition for deceptive agent to abstract deceptive features		Identifies deceptive behavior
Yifei and Lakshminarayanan (2023) [55]	Control Theory	Controlling multiloop processes	Impact of various reward functions investigated	A3C algorithm	Under pure cooperation, MARL with mixed strategy performs better
Liu et al. (2022) [56]	Network Optimization	Difficulties of charging wireless stations	<ul style="list-style-type: none"> Joint optimization formulated as a MINLP problem NP-hard MRA problem 	Two-layer MRA algorithm	Two-layer MRA performs better than single layer
Zheng et al. (2022) [57]	Game Theory framework	Actor-critic based RL	Meta-framework for Stackelberg actor-critic algorithms is proposed	Stackelberg actor-critic algorithm	Algorithm seldom outperforms actor-critic algorithm
Salgado and Clempner (2018) [58]	Emotional Perception	Measuring emotional state	NCGT approach for representing the interaction between agents.	<ul style="list-style-type: none"> NCGT approach RL algorithm 	
Zazo et al. (2016) [59]	Dynamic Games	Potential Dynamic Potential Games	Single multivariate optimal control problem		Theoretical outcomes naturally follow static potential game analysis
Marden (2012) [60]	Game theory frameworks	State based potential games	State-based potential games as used in cooperative control problems		Underlying design fits with state games
Arthurs and Birnbaum [61]	General game playing	Two player adversarial games	The Q-function is implemented as a neural network		The model can plan multiple moves in advance

a Stochastic game and proposing a payoff-based MARL approach has proven effective.

B. AUTOMOBILE CONTROL

The analysis and solution of problems in automobile control, such as intrusion detection in vehicular ad hoc networks (VANETs) and the incorporation of new technology, can be significantly aided by game theory. Game-theoretic models identify the best tradeoffs between accuracy and efficiency in IDS to improve network security in-vehicle environments.

The papers discuss several topics related to the interaction between game theory and RL in several domains, including traffic modeling, intrusion detection in VANETs, and automation of air traffic systems. Approaches like continuous policy spaces, Gaussian Processes, and hierarchical decision-making frameworks are introduced to improve the accuracy and efficiency of driver models and traffic congestion control to improve the modeling of human-driver interactions and traffic signal control by applying game theoretical concepts [70]. The studies also

highlighted the significance of agent collaboration and cooperation in multi-agent systems for efficient traffic signal control and congestion reduction [42]. Game theory and RL approaches made Independent and collaborative decision-making processes possible, which also improved traffic flow and decreased delays. A Bayesian game theory and deep Q-learning network-based IDS that adapts to changing environments and optimizes the tradeoff between accuracy and efficiency was also proposed as a solution to the problems that IDS in VANETs confront [39], [40]. The papers also examine incorporating new technology into air traffic control systems and highlight the viability of behavior modeling with RL based on goals [43]. This method enables evaluating and validating automation scenarios while considering the intricate relationships between various system decision-makers.

C. MOBILE EDGE CACHING

MARL techniques benefit greatly from game theory in several areas, including computation offloading and mobile edge computing. Game theory and MARL integration enable

effective and reasonable offloading decisions in dynamic situations with numerous users and limited resources. Nash equilibria can be proven to exist by framing offloading decision processes as stochastic game models, which enables the creation of distributed computing offloading algorithms based on payoff-based MARL strategies. The optimization of offloading rules, choice of Mobile Edge Caching(MEC) servers, and estimation of offloading data amount and pricing are further made possible by applying deep RL and game theory. The new algorithms have performed better than existing methods through simulations, reducing system overhead, profit maximization for MEC servers, and secure caching services in MSNs. The efforts of these studies demonstrate the value of game theory in refining MARL methods and providing reliable, practical solutions for decision-making and coordination in multi-agent systems [28], [29], [30], [46].

D. CYBER-PHYSICAL SYSTEMS

Game theory and RL are employed in CPSs to address complex issues. Game theory and RL applications are applied to forecast the results of CPSs involving many human interactions [25]. This method enables a computationally viable solution that captures the dynamics of interactions by permitting the simultaneous modeling of several human decision-makers. Formulating The interaction between the sensor and attacker as a two-player zero-sum game, deriving the NE strategies, and proposing an RL algorithm that dynamically adjusts the strategy helps to address the optimal policy selection problem in CPSs under a DoS attacks [22]. The algorithm considers variables like security state estimation and channel dependability to assess the effectiveness of offensive and defensive methods.

IX. CURRENT GAPS AND FUTURE RESEARCH OPPORTUNITIES

Based on our detailed literature survey, we identified some research gaps. A detailed discussion about the research gaps in game theory is as follows:

A. REAL-WORLD ENVIRONMENT

Creating machine learning algorithms that can replicate and represent more human-like surroundings in game theory settings is one exciting area for the future. At the moment, game theory frequently uses simplified suppositions and abstractions, which might not fully convey the complexity of actual circumstances. Traffic and automation control provides scope for betterment in this area where a more practical human-like environment setting can enhance the results [40], [42]. Machine learning methods like deep learning and generative models pave the way for developing more dynamic and realistic environments that better reflect human behavior. A practical environment modeling can benefit human-AI cooperation by efficiently capturing nuances of human actions, preferences, and decision-making processes [23].

B. INTEGRATING OVERLOOKED DETAILS

Extending the research to previously overlooked aspects can lead to further refined results. In [28], researchers can extend the model to include transmission power regulation and communication interference as a component of the joint optimization problem, to achieve a more thorough knowledge of the system dynamics. Liang et al., in [29], must create an efficient compute offloading mechanism to ensure optimum resource utilization, taking into account the mobility of mobile users (MUs). Additionally, enhancing the lane-changing models by considering continuous velocity and space can result in more accurate and realistic representations [39]. It is essential to concentrate on speed difficulties and the classification and identification of fatal scenarios to solve safety issues [36]. Investigating confined control or state space inside the already-existing algorithms can improve their performance. The science of game theory and machine learning can progress and offer more thorough answers by including these ignored subtleties.

C. ALGORITHM REFINEMENT

Sophisticated methods to overcome constraints improve previous approaches. In [25], one computationally expensive solution is to update agents' assumptions using Bayesian inference and build non-predefined policies to solve level-k thinking's static assumptions. In [49], dynamic multi-agent optimization can be reduced via cooperative approaches. RL techniques extended to multi-sensor single-attacker and multi-sensor multi-attacker security issues can bring forward exciting results. Model expansion under various attacks helps understand system flaws. Constrained control, state space, and deeper neural networks can improve existing algorithms [22], [48]. Future studies could examine compliance weight, learning timescales, and naive agents [47]. Studying cooperative games using MARL and formulating new algorithms can be a significant breakthrough [53]. MARL and mixed-strategy equilibria RL methods can help solve complicated issues. These algorithm modifications can help game theory and machine learning solve problems efficiently.

X. DISCUSSION AND CONCLUSION

The intersection of game theory and RL is a valuable and productive field of research. Due to the fusion of these two sciences, significant progress has been made in understanding and solving complex decision-making issues across various areas. Game theory has enabled the handling of multi-agent settings. Game theory has been essential in developing more intelligent and effective strategies in MARL, auto control, mobile edge caching, CPSs, human-AI cooperation, Internet-of-Things, and real-time strategies.

There are several intriguing directions for future investigation. We gain a deeper understanding of these strategies' theoretical foundations and real-world applications by exploring the improvements in RL made possible by game theory. It allows us to handle even more complicated decision-

making situations, enhance the overall effectiveness of RL agents, and formulate new algorithms and approaches.

Game theory's contribution to advancing MARL warrants additional study. It is possible to improve the coordination, collaboration, and decision-making capacities of MARL systems by investigating fresh game-theoretic concepts and strategies, resulting in more efficient and effective outcomes in various fields, including traffic control, network security, and automation. Integrating game theory and RL in cutting-edge areas like MEC and CPSs offers fascinating study opportunities. Further research into these topics may lead to the creation of dependable and scalable solutions for resource allocation, decision-making regarding offloading, and security augmentations in dynamic and complex contexts.

ACKNOWLEDGMENT

The authors wish to thank the anonymous reviewers for their valuable suggestions.

REFERENCES

- [1] M. J. Osborne, *An Introduction to Game Theory*, vol. 3, no. 3. New York, NY, USA: Oxford Univ. Press, 2004.
- [2] R. S. Sutton and A. G. Barto, *Reinforcement Learning: An Introduction*. Cambridge, MA, USA: MIT Press, 2018.
- [3] D. Fudenberg and J. Tirole, *Game Theory*. Cambridge, MA, USA: MIT Press, 1991.
- [4] D. Silver, A. Huang, C. J. Maddison, A. Guez, L. Sifre, G. van den Driessche, J. Schrittwieser, I. Antonoglou, V. Panneershelvam, M. Lanctot, S. Dieleman, D. Grewe, J. Nham, N. Kalchbrenner, I. Sutskever, T. Lillicrap, M. Leach, K. Kavukcuoglu, T. Graepel, and D. Hassabis, "Mastering the game of go with deep neural networks and tree search," *Nature*, vol. 529, no. 7587, pp. 484–489, Jan. 2016.
- [5] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, U.K.: Cambridge Univ. Press, 2008.
- [6] Y. Yang and J. Wang, "An overview of multi-agent reinforcement learning from game theoretical perspective," 2020, *arXiv:2011.00583*.
- [7] Y. Lu and K. Yan, "Algorithms in multi-agent systems: A holistic perspective from reinforcement learning and game theory," 2020, *arXiv:2001.06487*.
- [8] Z. Gao, "Understanding the future of deep reinforcement learning from the perspective of game theory," *J. Phys., Conf. Ser.*, vol. 1453, no. 1, Jan. 2020, Art. no. 012076.
- [9] C. Zhang, R. T. Ma, B. Gu, N. Kamiyama, W. Li, and Y. Tanaka, "Future research trends by integrating deep reinforcement learning, robotics and game theory," in *Proc. IEICE Conf. Arch.* Tokyo, Japan: Institute of Electronics, Information and Communication Engineers, 2019.
- [10] M. Moll and L. Kunczik, "Two perspectives on playing games: Reinforcement learning vs game theory," in *Proc. Int. Conf. Artif. Intell. (ICAI), World Congr. Comput. Sci.*, 2019, pp. 60–61.
- [11] R. Sharma and M. Gopal, "Synergizing reinforcement learning and game theory—A new direction for control," *Appl. Soft Comput.*, vol. 10, no. 3, pp. 675–688, Jun. 2010.
- [12] M. J. Osborne and A. Rubinstein, *A Course in Game Theory*. Cambridge, MA, USA: MIT Press, 1994.
- [13] A. K. Dixit and B. J. Nalebuff, *Thinking Strategically: The Competitive Edge in Business, Politics, and Everyday Life*. New York, NY, USA: W. W. Norton, 1993.
- [14] S. Tadelis, *Game Theory: An Introduction*. Princeton, NJ, USA: Princeton Univ. Press, 2013.
- [15] R. B. Myerson, *Game Theory: Analysis of Conflict*. Cambridge, MA, USA: Harvard Univ. Press, 1991.
- [16] J. Nash Jr., "Non-cooperative games," in *Essays Game Theory*. Cheltenham, U.K.: Edward Elgar, 1996, pp. 22–33.
- [17] J. Nash Jr., "Two-person cooperative games," *Econometrica*, vol. 21, no. 1, pp. 128–140, Jan. 1953.
- [18] O. Chatain, *Cooperative and Non-Cooperative Game Theory*. Philadelphia, PA, USA: Univ. Pennsylvania, 2014.
- [19] V. François-Lavet, P. Henderson, R. Islam, M. G. Bellemare, and J. Pineau, "An introduction to deep reinforcement learning," *Found. Trends Mach. Learn.*, vol. 11, nos. 3–4, pp. 219–354, Mar. 2019.
- [20] D. Bertsekas, *Dynamic Programming and Optimal Control*, vol. 1. Nashua, NH, USA: Athena Scientific, 2012.
- [21] M. A. Wiering and M. Van Otterlo, "Reinforcement learning," *Adaptation, Learn., Optim.*, vol. 12, no. 3, p. 729, 2012.
- [22] Z. Jin, S. Zhang, Y. Hu, Y. Zhang, and C. Sun, "Security state estimation for cyber-physical systems against DoS attacks via reinforcement learning and game theory," *Actuators*, vol. 11, no. 7, p. 192, Jul. 2022.
- [23] B. Schelble, C. Flathmann, L.-B. Canonico, and N. Mcneese, "Understanding human-AI cooperation through game-theory and reinforcement learning models," in *Proc. Annu. Hawaii Int. Conf. Syst. Sci.*, 2021.
- [24] J. Khoury and M. Nassar, "A hybrid game theory and reinforcement learning approach for cyber-physical systems security," in *Proc. IEEE/IFIP Netw. Oper. Manage. Symp. (NOMS)*, Apr. 2020, pp. 1–9.
- [25] B. M. Albaba and Y. Yildiz, "Modeling cyber-physical human systems via an interplay between reinforcement learning and game theory," *Annu. Rev. Control*, vol. 48, pp. 1–21, Jan. 2019.
- [26] L. Wang, W. Zhou, H. Xu, L. Li, L. Cai, and X. Zhou, "Research on task offloading optimization strategies for vehicular networks based on game theory and deep reinforcement learning," *Frontiers Phys.*, vol. 11, Oct. 2023.
- [27] T. Wang, B. Lu, W. Wang, W. Wei, X. Yuan, and J. Li, "Reinforcement learning-based optimization for mobile edge computing scheduling game," *IEEE Trans. Emerg. Topics Comput. Intell.*, vol. 7, no. 1, pp. 55–64, Feb. 2023.
- [28] S. Li, X. Hu, and Y. Du, "Deep reinforcement learning and game theory for computation offloading in dynamic edge computing markets," *IEEE Access*, vol. 9, pp. 121456–121466, 2021.
- [29] S. Liang, H. Wan, T. Qin, J. Li, and W. Chen, "Multi-user computation offloading for mobile edge computing: A deep reinforcement learning and game theory approach," in *Proc. IEEE 20th Int. Conf. Commun. Technol. (ICCT)*, Oct. 2020, pp. 1534–1539.
- [30] Q. Xu, Z. Su, and R. Lu, "Game theory and reinforcement learning based secure edge caching in mobile social networks," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 3415–3429, 2020.
- [31] L. Xue, B. Ma, J. Liu, C. Mu, and D. C. Wunsch, "Extended Kalman filter based resilient formation tracking control of multiple unmanned vehicles via game-theoretical reinforcement learning," *IEEE Trans. Intell. Vehicles*, vol. 8, no. 3, pp. 2307–2318, Mar. 2023.
- [32] S. Fu, X. Feng, A. Sultana, and L. Zhao, "Joint power allocation and 3D deployment for UAV-BSS: A game theory based deep reinforcement learning approach," *IEEE Trans. Wireless Commun.*, 2023.
- [33] L. Cipolina-Kun, "Enhancing smart, sustainable mobility with game theory and multi-agent reinforcement learning," in *Proc. Int. Conf. Auto. Agents Multiagent Syst.*, Jun. 2023, pp. 3035–3037.
- [34] W. Li, F. Qiu, L. Li, Y. Zhang, and K. Wang, "Simulation of vehicle interaction behavior in merging scenarios: A deep maximum entropy-inverse reinforcement learning method combined with game theory," *IEEE Trans. Intell. Vehicles*, 2024.
- [35] S. Yin, Y. Kang, Y. Zhao, and J. Xue, "Air combat maneuver decision based on deep reinforcement learning and game theory," in *Proc. 41st Chin. Control Conf. (CCC)*, Jul. 2022, pp. 6939–6943.
- [36] W. Duan, Z. Tang, W. Liu, and H. Zhou, "Autonomous driving planning and decision making based on game theory and reinforcement learning," *Expert Syst.*, vol. 40, no. 5, p. e13191, Jun. 2023.
- [37] D. Zwilling, J. Sierchio, M. Gerken, and E. Clark, "Distributing data throughout a MANET in a communications denied environment: Reinforcement learning and game theory approaches," *Proc. SPIE*, vol. 12113, pp. 123–136, Jun. 2022.
- [38] Z. Zhan, G. Lu, K. Wang, J. Li, and J. Wang, "Game theory and reinforcement learning based smart lane change strategies," *SAE Tech. Paper 2022-01-0221*, 2022.
- [39] J. Guo, "Multi-agent traffic control using game theory and reinforcement learning," *Tech. Rep.*, 2022.
- [40] J. Liang, M. Ma, and X. Tan, "GaDQN-IDS: A novel self-adaptive IDS for VANETs based on Bayesian game theory and deep reinforcement learning," *IEEE Trans. Intell. Transp. Syst.*, vol. 23, no. 8, pp. 12724–12737, Aug. 2022.

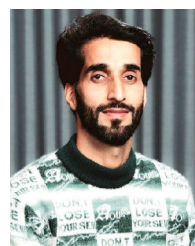
- [41] B. M. Albaba and Y. Yildiz, "Driver modeling through deep reinforcement learning and behavioral game theory," *IEEE Trans. Control Syst. Technol.*, vol. 30, no. 2, pp. 885–892, Mar. 2022.
- [42] M. Abdoos, "A cooperative multiagent system for traffic signal control using game theory and reinforcement learning," *IEEE Intell. Transp. Syst. Mag.*, vol. 13, no. 4, pp. 6–16, Winter 2021.
- [43] Y. Yildiz, A. Agogino, and G. Brat, "Predicting pilot behavior in medium-scale scenarios using game theory and reinforcement learning," *J. Guid., Control, Dyn.*, vol. 37, no. 4, pp. 1335–1343, Jul. 2014.
- [44] D. Arend, S. Yuwono, M. R. Diprasetya, and A. Schwung, "MLPro 1.0—Standardized reinforcement learning and game theory in Python," *Mach. Learn. Appl.*, vol. 9, Sep. 2022, Art. no. 100341.
- [45] T. M. Ho, K.-K. Nguyen, and M. Cheriet, "Converging game theory and reinforcement learning for Industrial Internet-of-Things," *IEEE Trans. Netw. Service Manage.*, 2022.
- [46] P. Teymoori and A. Boukerche, "Dynamic multi-user computation offloading for mobile edge computing using game theory and deep reinforcement learning," in *Proc. IEEE Int. Conf. Commun. (ICC)*, May 2022, pp. 1930–1935.
- [47] I. Farahbakhsh, "Modeling human-coupled common pool resource systems with techniques in evolutionary game theory and reinforcement learning," M.S. thesis, Univ. Waterloo, Waterloo, ON, Canada, 2021.
- [48] Z. Zhou, "Large-scale multi-agent decision-making using mean field game theory and reinforcement learning," Ph.D. dissertation, Univ. Nevada, Reno, NV, USA, 2021.
- [49] X.-W. Liu, Q. Zhang, Y. Luo, X. Lu, and C. Dong, "Radar network time scheduling for multi-target ISAR task with game theory and multiagent reinforcement learning," *IEEE Sensors J.*, vol. 21, no. 4, pp. 4462–4473, Feb. 2021.
- [50] G. Adams, S. J. Padmanabhan, and S. Shekhar, "Resolving implicit coordination in multi-agent deep reinforcement learning with deep Q-networks & game theory," 2020, *arXiv:2012.09136*.
- [51] C. M. Morrison, "Game theory based distributed coordination with multi-agent reinforcement learning," in *Proc. SOARS*, 2020.
- [52] H. Tang and C. Dong, "Hierarchical reinforcement learning based on multi-agent cooperation game theory," *Int. J. Wireless Mobile Comput.*, vol. 16, no. 4, pp. 369–376, 2019.
- [53] T. Zhu, M. Ma, L. Chen, and Z. Liu, "Optimal strategy of the simultaneous dice game pig for multiplayer: When reinforcement learning meets game theory," *Sci. Rep.*, vol. 13, no. 1, p. 8142, May 2023.
- [54] Y. Zeng and K. Xu, "Recognition and interfere deceptive behavior based on inverse reinforcement learning and game theory," *J. Syst. Eng. Electron.*, vol. 34, no. 2, pp. 270–288, Apr. 2023.
- [55] Y. Yifei and S. Lakshminarayanan, "Multi-agent reinforcement learning for process control: Exploring the intersection between fields of reinforcement learning, control theory, and game theory," *Can. J. Chem. Eng.*, vol. 101, no. 11, pp. 6227–6239, Nov. 2023.
- [56] J. Liu, C.-H.-R. Lin, Y.-C. Hu, and P. K. Donta, "Joint beamforming, power allocation, and splitting control for SWIPT-enabled IoT networks with deep reinforcement learning and game theory," *Sensors*, vol. 22, no. 6, p. 2328, Mar. 2022.
- [57] L. Zheng, T. Fiez, Z. Alumbaugh, B. Chasnov, and L. J. Ratliff, "Stackelberg actor-critic: Game-theoretic reinforcement learning algorithms," in *Proc. AAAI Conf. Artif. Intell.*, 2022, vol. 36, no. 8, pp. 9217–9224.
- [58] M. Salgado and J. B. Clempner, "Measuring the emotional state among interacting agents: A game theory approach using reinforcement learning," *Expert Syst. Appl.*, vol. 97, pp. 266–275, May 2018.
- [59] S. Zazo, S. V. Macua, M. Sánchez-Fernández, and J. Zazo, "Dynamic potential games with constraints: Fundamentals and applications in communications," *IEEE Trans. Signal Process.*, vol. 64, no. 14, pp. 3806–3821, Jul. 2016.
- [60] J. R. Marden, "State based potential games," *Automatica*, vol. 48, no. 12, pp. 3075–3088, Dec. 2012.
- [61] N. Arthurs and S. Birnbaum, *Deep Reinforcement Learning for General Game Playing* (Theory and Reinforcement), 2016.
- [62] D. Monderer and L. S. Shapley, "Potential games," *Games Econ. Behav.*, vol. 14, no. 1, pp. 124–143, May 1996.
- [63] Q. D. Lă, Y. H. Chew, B.-H. Soong, Q. D. Lă, Y. H. Chew, and B.-H. Soong, "An introduction to game theory," in *Potential Game Theory: Applications in Radio Resource Allocation*. Cham, Switzerland: Springer, 2016, pp. 3–22.
- [64] J. Heinrich, M. Lanctot, and D. Silver, "Fictitious self-play in extensive-form games," in *Proc. Int. Conf. Mach. Learn.*, 2015, pp. 805–813.
- [65] L. S. Shapley, "Stochastic games," *Proc. Nat. Acad. Sci. USA*, vol. 39, no. 10, pp. 1095–1100, Oct. 1953.
- [66] L. E. Zachrisson, *Markov Games*, vol. 52. Princeton, NJ, USA: Princeton Univ. Press, 1964.
- [67] P. E. Caines, "Mean field games," in *Encyclopedia of Systems and Control*. London, U.K.: Springer, 2021, pp. 1197–1202.
- [68] D. Koller and N. Megiddo, "The complexity of two-person zero-sum games in extensive form," *Games Econ. Behav.*, vol. 4, no. 4, pp. 528–552, Oct. 1992.
- [69] J. C. Ely and W. H. Sandholm, "Evolution in Bayesian games I: Theory," *Games Econ. Behav.*, vol. 53, no. 1, pp. 83–109, Oct. 2005.
- [70] C. O. Yaldiz and Y. Yildiz, "Driver modeling using continuous reasoning levels: A game theoretical approach," in *Proc. IEEE 61st Conf. Decis. Control (CDC)*, Dec. 2022, pp. 5068–5073.



machine learning-based solutions for real-world problems, showcasing her commitment to practical applications of her skills.



Quality Teacher with the Technical Education Quality Improvement Program (TEQIP)–III by the World Bank and the Government of India, for three years. His research interests include machine learning, computational biology, and sentiment analysis.



also an Assistant Professor with the Department of Mathematics, Maharishi Markandeshwar University, India. He has published many reputed research reports in high impact journals, such as *Engineering Applications of Artificial Intelligence*, *Applied Soft Computing*, *Resources Policy*, and *Expert Systems with Applications*. His current research interests include the development of intelligent food supply chain sustainability model based on the Internet of Things (IoT) and artificial intelligence (AI) in the wake of COVID-19 pandemic, optimization techniques, progressive feature extraction of biological sequences for machine learning, fuzzy mathematical model for tumor growth pattern using generalized hukuhara derivative and its numerical analysis, logistic growth model of tumor population under fuzzy environment and its numerical analysis, and the development of novel multi-attribute decision making (MADM) methods under fuzzy environment and its extensions. Entitled some methods for prioritization of alternatives of MADM problems in fuzzy environment and its extensions.

• • •