



Breaking the traditional: a survey of algorithmic mechanism design applied to economic and complex environments

Qian Chen¹ · Xuan Wang^{1,2,3} · Zoe Lin Jiang^{1,2,3} · Yulin Wu^{1,3} · Huale Li¹ · Lei Cui^{1,2} · Xiaozhen Sun¹

Received: 28 April 2022 / Accepted: 2 May 2023 / Published online: 20 May 2023
© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2023

Abstract

The mechanism design theory can be applied not only in the economy but also in many fields, such as politics and military affairs, which has important practical and strategic significance for countries in the period of system innovation and transformation. As Nobel Laureate Paul said, the complexity of the real economy makes it difficult for “Unorganized Markets” to ensure supply-demand balance and the efficient allocation of resources. When traditional economic theory cannot explain and calculate the complex scenes of reality, we require a high-performance computing solution based on traditional theory to evaluate the mechanisms, meanwhile, get better social welfare. The mechanism design theory is undoubtedly the best option. Different from other existing works, which are based on the theoretical exploration of optimal solutions or single perspective analysis of scenarios, this paper focuses on the more real and complex markets. It explores to discover the common difficulties and feasible solutions for the applications. Firstly, we review the history of traditional mechanism design and algorithm mechanism design. Subsequently, we present the main challenges in designing the actual data-driven market mechanisms, including the inherent challenges in the mechanism design theory, the challenges brought by new markets and the common challenges faced by both. In addition, we also comb and discuss theoretical support and computer-aided methods in detail. This paper guides cross-disciplinary researchers who wish to explore the resource allocation problem in real markets for the first time and offers a different perspective for researchers struggling to solve complex social problems. Finally, we discuss and propose new ideas and look to the future.

Keywords Algorithmic mechanism design · Incentive mechanism · Online advertising · Privacy preserving

1 Introduction

Mechanism design is one of the research areas in economics and game theory that uses the engineering methods to design economic mechanisms or incentives to achieve desired goals. It is also called “Reverse Game Theory” because it starts at the end of the game and then goes backwards. Mechanism design explains the problem that the objective function is given when the mechanism is unknown. Therefore, the design problem is contrary to traditional economic theory, which usually focuses on analyzing the performance of the given mechanism. It has a wide range of applications, from market, auctions, voting to inter-domain routing, sponsored search auctions and other resource allocation scenarios.

Algorithmic game theory studies the interaction between rational and selfish agents. It is an emerging field combining algorithms and game theory, and its achievements are related to other fields, including networks and artificial intelligence. Nisan et al. [111] published the representative book of algorithmic game theory in 2017, which summarized two essential branches of this field: game strategy and algorithmic mechanism design. Algorithm game theory studies the calculation model in the real market from the game theory perspective and uses economics and calculation theory. It offers the possibility to break through the traditional mechanism design theory and various unrealistic assumptions, such as infinite computational power and ex-ante constrained assumption conditions. Figure 1 shows the branches and relationship of algorithmic game theory, and we will use mechanism design without annotation to denote algorithmic mechanism design in the following.

Extended author information available on the last page of the article

Both efficient game strategy solving and rational game mechanism design are essential and are two different aspects of the same thing. Obtaining the Nash equilibrium in the game strategy solution is optimal from the player's perspective. Moreover, the design of a rational game mechanism is to create systems or rules that satisfy consensual goals, that guide agents in the system to make strategic actions, and retain private information relevant to the decision. The research on mechanism design is not descriptive or evaluative but the process of opening the black box of the research object to the public.

The algorithmic game theory of strategy solving is more familiar to researchers in computing, which searches for the optimal strategy under a fixed mechanism. The methods of finding equilibrium are constantly being improved, from the DeepBlue victory over chess master Kasparov, the AlphaGo [139], AlphaGo Zero [140] defeat of Go world champion Li Shishi and Ke Jie, to the Cepheus [18], DeepStack [104], Libratus [16] and Pluribus [19] successive conquests of the challenges in two-player limit and no-limit, multiplayer no-limit Texas Hold'em games. The state-of-the-art researches also include the Awareness that defeated a team of professional players from the Glory of Kings and exploring correlated equilibrium [23, 48, 176, 177], Stackelberg equilibrium [10, 81, 89, 182] and other different kinds of equilibrium solutions. As "reverse game theory," the mechanism design explores the mechanism, while agents use rational strategy to maximize global social welfare. In chess, for example, by modifying the appropriate rules, it is possible to improve the global goal and the optimal equilibrium that agents can achieve.

The most common applications of algorithmic mechanism design explored in this paper include auctions, advertising, and other economics scenarios. How to break through the limitation of this traditional concept, or whether mechanism design can be applied in other complex environments, is what this paper is going to study and looking for.

As the Royal Swedish Science Council said in the award ceremony, mechanism design theory provides a coherent framework for analyzing this great variety of institutions, or "allocation mechanisms," with a focus on the problems associated with incentives and private information. Mechanism design theory allows researchers to systematically analyze and compare a broad variety of institutions under less stringent assumptions. It allows economists and other social scientists to analyze the performance of institutions relative to the theoretical optimum. The Nobel Prize in Economics has been awarded four times for research in the area of mechanism design: Vickery and Morris for their work [98, 154] on asymmetric information in 1996, Hertz, Maskin and Myerson for their pioneering work [65, 90, 105] on mechanism design in 2007, Roth and

Shapley for their essential contributions [55, 128] to matching markets in 2012. Paul Milgrom and Robert Wilson for their refinement and invention [116, 160] of auction theory in 2020. In short, mechanism design, like game strategy solving, has provided important insights into economic, social and even political-military issues, especially regarding its significant impact on policies and institutions.

The modern form of game theory existed as early as the 1940s and 1950s, and outstanding researchers such as John von Neumann and Morgenstein (Theory of Games and Economic Behavior, 1944 [156]), Nash (Non-cooperative Games, 1950 [107]; Two-person cooperative games, 1953 [108]), and even has earlier explorations such as Zermelo's theorem, where Zermelo's thesis was published in German in 1913 and translated into English by Ulrich Schwalbe and Paul Walker in 1997. Although game theory occasionally overlapped with computer science for many years, most of the early research in game theory was carried out by economists. Indeed, game theory is now the primary analytical framework for microeconomic theory, as evidenced by its presence in economics textbooks (e.g., Microeconomic Theory, 1995 [91]) and many Nobel Prizes in the economic sciences. In recent years, scholars have made some progress in discussing the single-item restriction of the famous Myerson theorem in game theory, studying one additive buyer with two items [61, 62, 88] and one-unit demand buyer with two items [117, 118, 148]. The latest is Yao's theoretical proof for two items with multiple bidders [169]. The multi-item mechanism has encountered a bottleneck in theory, but this has not stopped researchers from moving forward to combine neural networks, reinforcement learning and other computer-aided algorithms with exploring [31, 39, 120, 123], even adaptive mechanisms are gradually becoming possible. This is undoubtedly an essential breakthrough in the design of traditional mechanisms, but there are still many challenges ahead that we need to explore.

There are at least two main reasons why game theory became a major research topic for computer scientists.

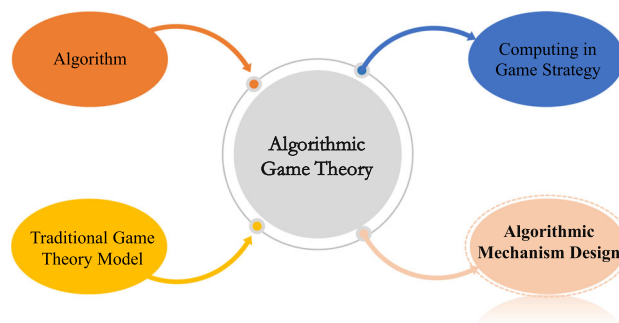


Fig. 1 Branches of algorithmic game theory

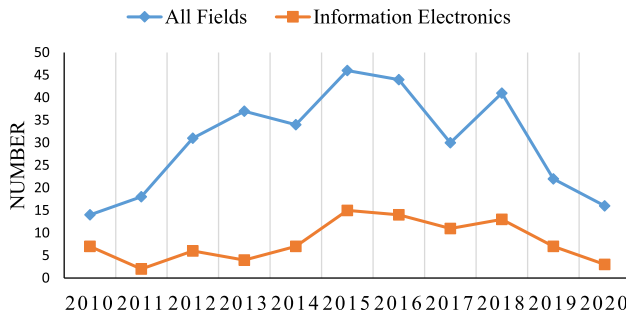


Fig. 2 Number of research projects in different fields of mechanism design

First, economists became interested in problems that seriously impede practical applications because of their computational properties. Thus, they started to approach computer science early on in the study of combinatorial [5] auction problems. Secondly, with the rise of distributed computing, especially the continuous development of big data and artificial intelligence, the traditional economic scene is no longer static. Internet-based advertising auctions have also become dynamic and variable at any time. The evidence based on the assumption of traditional mechanism design theory is no longer strong enough. In recent years, with the rapid development of machine learning technologies such as deep learning, for example, in 2012, Jeffrey Hinton uses the AlexNet model [74] to win the championship and become an iconic node. Subsequent improved and representative models have grown exponentially, including ELMO, GPT, BERT, XLM [30] and other models applied to different fields.

Corresponding to the above two waves, Figs. 2, 3, 4 and 5 show the number and trend of research projects¹ and academic papers² with “mechanism design” or “incentive mechanism.” In addition to the development of traditional mechanism design in around 2012, due to the emergence of new opportunities, the research projects of mechanism design have substantially increased. The USA, Canada and China, as the countries with the most significant investment in mechanism design, are far ahead of other countries. For example, the National Science Foundation of the USA funded the university with \$499,972.00 [54] to explore new directions of social choice and mechanism design.

The numbers of academic papers that shown in Fig. 4 are based on the database of web of science, which are calculated by excluding the research directions with low correlation (such as biology and chemistry). We can see that around 2015, there was a new increase in the number of academic papers, which proves the emergence of new

opportunities. In addition to the traditional essential conferences and journals in the field of artificial intelligence, three major important conferences are developed in the direction of game theory methods and applications in computing. There are the International Symposium on Algorithmic Game Theory (SAGT), the ACM Conference on Economics and Computation (ACM EC) and The Conference on Web and Internet Economics (WINE). It provides a good platform for the development of this field.

Game strategy solving and mechanism design are all interdisciplinary, which pose considerable challenges to computer science and economics researchers, and have their own perspectives and emphases. Compared with traditional game theory, algorithmic game theory faces more challenges but has achieved remarkable results.

For static game strategy solution scenario with complete information, it is usually modeled as game trees, which is used for searching to obtain the best strategy. We often reduce the space of the gaming tree by the techniques such as pruning. For large-scale game trees, supervised learning [139] or self-game [140] methods are needed to evaluate the state, so as to conduct accurate and efficient heuristic search.

As for the incomplete information game, solving the Nash equilibrium becomes more difficult because of the larger scale of the game tree, so it is necessary first to perform game abstraction [130] and then execute the iterative algorithm based on CFR, residual solution and so on. The current state-of-the-artwork has been able to beat the world’s top humans at two-player no-limit Texas game.

Typical applications of the mechanism design include how sellers design prices to sell under asymmetric information between buyers and sellers in a buy-sell transaction scenario (e. g. online advertising). Auction theory has undergone significant changes since the early studies of mechanism design theory. Although the early “impractical” models proved their worth in guiding the actual auction design, some of these were not entirely correct. As Paul Milgrom, the winner of Nobel Prize in Economics in 2020, summed up in his paper [95], simulation and calculation methods may become increasingly important in the future. However, understanding this problem and the proposed solutions still need theoretical support. It is time for the old and new methods to work together. Artificial intelligence-based mechanism design methods are a hot topic in current research at the intersection of artificial intelligence and economics. The research mainly includes optimizing mechanism design through reinforcement learning, modeling players through machine learning and behavioral economics, modeling players through experimental economics and verifying and evaluating with mechanism design.

¹ China Knowledge Centre for Engineering Sciences and Technology.

² Web of Science.

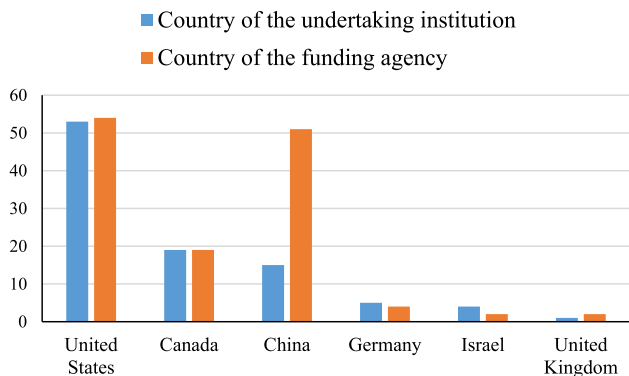


Fig. 3 Undertaking and funding about mechanism design research projects

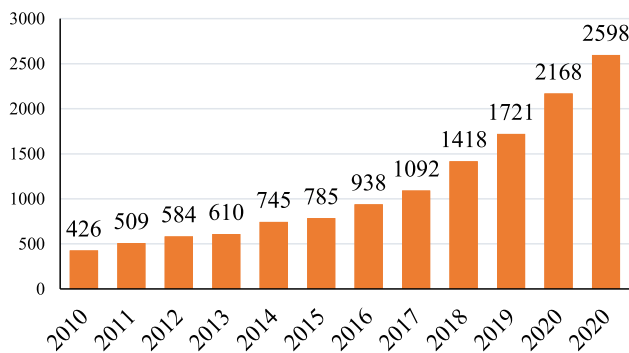


Fig. 4 Number of academic papers on mechanism design



Fig. 5 Word cloud statistics

Based on the algorithm game theory, the researchers in their respective disciplines have conducted in-depth research on game strategy and mechanism design, but they lack systematic summaries and comments. Our team has summarized a review of Nash Equilibrium Strategy Solving [78] in 2020. However, there is a lack of comprehensive papers that can describe the challenges in this interdisciplinary field of algorithm mechanism design. This paper attempts to summarize this field, try to give new research points and give new inspiration to researchers. On

the one hand, static mechanism design theory needs to be optimized in the context of dynamically variable data in more complex scenarios. On the other hand, with the widespread use of artificial intelligence for game strategy solving, combining computer-aided modeling approaches is necessary to counter the enhanced game strategy.

In addition, apart from the outstanding performance in economic policy and market systems, whether mechanism design is suitable for other more complicated social interaction problems, such as data sharing and privacy-preserving, and classification of scarce resources. Moreover, how it can be broken through the traditional theory of mechanism design and traditional research areas is also the problem we want to discuss.

The main difference between this paper and other work is that we stand at the intersection of multiple disciplines, hoping to help researchers quickly understand the hot spots, feasible solutions and challenges in this intersection. This is the author’s original goal. We hope beginners do not need to spend too much time looking for fragmented information and finding the point of convergence between the two disciplines. It is also hoped that experts can quickly determine the complex problems that can be solved in the next stage.

The rest of the paper is as follows. We introduce the history and comparison of mechanism design in Sect. 2. In Sect. 3, the main problems and challenges faced by the mechanism design at this stage are shown. And then at Sects. 4 and 5, we systematically introduce and explore the traditional mechanism design theory and computer-aided incentive mechanism methods. Finally, we discuss the application and prospect of algorithmic mechanism design in the economic and complex environment from three aspects, the optimal mechanism of internet online advertising (traditional economics scenarios), the incentive mechanism for data sharing and privacy preserving (complex environment) and the allocation of other scarce resources (abstraction of mechanism design issues) in Sect. 6. And we make a conclusion of the paper in Sect. 7.

2 Related work

2.1 History of game theory and mechanism design

John von Neumann’s theory of games and economic behavior [156], co-authored with Oskar Morgenstern in 1944, is the pioneering work in game theory. The authors think that game theory is the only mathematical method suitable for studying economic problems. They proved the existence of zero-sum game solutions for two or more agents and gave the solutions of cooperative games.

Schelling's "The Strategy of Conflict" [131] made an in-depth analysis of the mutually distrustful relationship and the critical role of credible commitment in conflict or negotiation, although it was non-technical, while the basis of axiomatic cooperation theory is "Value" put forward by Shapley [134] in 1953 and the "Bargaining Solution" put forward by Nash in 1950. Among them, Shapley Value distributes the interests of alliance agents, which reflects the contribution of each alliance member to the overall goal of the alliance, avoids egalitarianism in distribution, and is more reasonable and fair than any other distribution methods based only on the value of input, the efficiency of resource allocation and the combination of both, and also reflects the process of the competing among the agents. Nash's bargaining solution is based on the insight that the value of personal utility can be assigned to distinguish incremental linear transformations.

Mechanism design theory not only points out the dilemmas but, more importantly, it provides a way out of the dilemmas under specific circumstances: how to design mechanisms or rules so that micro-actors can genuinely show their personal preferences and the behaviors determined by personal preferences and economic mechanisms can finally ensure the realization of social goals. At present, mechanism design theory has entered the core of mainstream economics. It is widely used in monopoly pricing, optimal taxation, contract theory, principal-agent theory, auction theory and other fields. Many practical and theoretical problems, such as the formulation of laws and regulations, optimal tax design, administrative management, democratic election and social system design, can all be attributed to the design of the mechanism.

The pioneering work on mechanism design theory originated from Hurwitz's studies in 1960 and 1972. It solves the general problem of whether and how to design an economic mechanism under decentralized decision-making conditions such as free choice, voluntary exchange and incomplete information so that the interests of the participants in economic activities are consistent with the goals announced by the designer for any given economic or social purpose. That is, by designing the specific form of the game mechanism, the interaction of the strategies chosen by the participants under the self-interested behavior can satisfy their respective constraints [68, 126, 147, 159] and make the configuration results consistent with the expected goal.

According to Adam Smith's assumption, the invisible hand of the market can efficiently allocate resources under ideal conditions. However, the real world is not so perfect, and there will always be various constraints that hinder the market from giving full play. In other words, the market is always prone to failures. Under the circumstances of incomplete competition, incomplete information, public

goods and increasing returns to scale, the market mechanism cannot automatically realize the effective allocation of resources. Take incomplete information as an example [4, 6, 144]. We are facing a society with incomplete information. No one can fully grasp the personal information of others. Since all personal information cannot be fully available to one person, people will always want to decentralize their decisions. However, information about personal preferences and existing production technologies is distributed among many participants, such as federated learning [170, 175], who may hide their accurate information and use private information to maximize their benefits. From the point of view of the whole society, this will lead to a loss of resource allocation efficiency.

As mentioned in Sect. 1, mechanism design is a branch of game theory. Both for game strategy solving and mechanism design, we usually classify them into "cooperative games" and "non-cooperative games," depending on whether the players are allied or not. Therefore, they correspond to different game models and are suitable for all environments. However, the environment in real life is much more complex. Sometimes, the information on both sides of the game is not entirely observable and need for data privacy preserving and other security issues. The last three sections of this paper, three aspects are studied in detail with examples from the traditional economic fields, more complex social environments and resource allocation for the abstract of mechanism design, respectively.

2.2 Comparison of traditional and algorithmic mechanism design

There is a complete theoretical system in traditional mechanism design. Compared with algorithmic mechanism design, it has a longer history and commonly uses auction, contract and bargaining theories, while the difference between auction theory with the latter theories is that it is widely used in non-cooperative games. Traditional theory of mechanism design is the cornerstone of mechanism design system. Whether it is the application of theoretical proofs to complex environments or, conversely, the enrichment of mechanism design theory through artificial intelligence, there is no doubt about its importance.

The typical application of mechanism design includes valuable items, reverse bidding of easily corroded items, spectrum auctions and other economic fields. The mechanism design method based on artificial intelligence is a hot topic of interdisciplinary research in computer and economics, which broadens the practical application scenarios for mechanism design. Almost all resource allocation problems [47, 50, 113, 122, 135] can be solved by mechanism design. The primary research includes optimizing the mechanism design through reinforcement learning,

modeling agents through machine learning and behavioral economics, and verifying and evaluating the mechanism through experimental economics.

Moreover, algorithmic mechanism design, as a new field of theoretical computer science, focuses on and solves the fundamental problem of real-time high-frequency behavior in networks and resource allocation under realistic complexity.

The difference between traditional mechanism design and game theory lies in the following aspects. First, it has different application fields, including more complex social environments such as big data and non-traditional auction scenarios. Secondly, from the point of view of specific optimization problems to the application of modeling, the search for optimal solutions and the study of upper and lower limits of solvable optimization, they are different in application and quantitative methods. Third, computability differs because the traditional mechanism design approach is ideal and cannot be solved by a computer in polynomial time (NP-hard problem). In contrast, the algorithmic mechanism design takes computability as a constraint that must be considered for algorithm implementation.

3 Challenges in mechanism design

As Victor and Myerson applied the theory of economic mechanism design to auction theory, they solved many complex resource allocation problems. Scholars in various fields have summarized and extended them from different perspectives, including auction mechanism, pricing mechanism, contract mechanism, monopoly pricing mechanism, principal-agent mechanism, and optimal tax mechanism.

Starting from the basic connotation of the traditional auction mechanism, some scholars have made in-depth discussions on the theoretical models of the single-stage VCG mechanism, the first-price and price-increase combined auction mechanism, auction processes and other mechanical issues. Some scholars also analyzed and compared the optimal mechanism design of homogeneous multi-item auctions [1, 3] under the conditions of unit demand or linear demand. Analyze the selection of auction methods in the process of state-owned assets [3, 87] separation, for example.

With the reform of market mechanisms, many scholars have gradually begun to pay attention to and extend the design theory of economic mechanisms. For example, considering the characteristics of the electricity system and the strategic behavior of electricity providers triggered by incomplete information problems [158, 184], which affects the security and economy of the electricity market. Some scholars applied mechanism design theory, supply function, Cuomo model, Stackelberg model to explore the

strategic behavior of electricity suppliers and incentive bidding mechanism and other issues.

This seems to be very successful, but there are many hidden problems with the traditional mechanism design. When putting theoretical models into practice, there are enormous challenges. For example, the classic Victor auction could not be used in the spectrum auction organized by the FCC in 2012. For example, the classic Victor auction could not be applied in the spectrum auction organized by the FCC in 2012 and faced with the dilemma that the optimization method could not be calculated on modern computers [94]. Spectrum auctions contain tens of millions of variables and 2.7 million constraints, which makes their calculations very complicated and beyond the scope of computer solutions.

Described in terms of computational disciplines, it is a typical graph coloring problem with huge constraints and is an NP-hard problem that cannot be solved optimally. However, there is no need to be discouraged. In the real world, many problems are NP-hard, but suboptimal solutions are also good options. Many scholars have explored clock auctions [96], simulation studies [110], the FUEL bid language [14] and other methods to tame this auction challenge.

There are still many difficulties to be solved in exploring the design of algorithmic mechanisms in complex environments such as big data and the Internet. For example, we need to consider the mechanism's computational complexity, time complexity and information security to design an efficient pricing mechanism and maximize the social welfare.

In this paper, starting from the main difficulties faced by traditional mechanism design, algorithm mechanism design and their combination, the current problems are divided into three categories, as shown in Fig. 6. In addition, Sect. 4 introduces the basic theory of traditional mechanism design in detail. Section 5 introduces the algorithm of computer-aided incentive mechanism in detail. And, Sect. 6 introduces the application of the combination in the economic and complex social environment. Furthermore, other application scenarios are given under the abstract goal of mechanism design to address resource allocation.

Through the exploration and research of the above problems, it is helpful to promote the optimization and application of algorithm mechanism design for more complicated and changeable application scenarios, to maximize the expected goals, such as social welfare, to achieve consistency between social goals and goals of game strategy.

3.1 Strategic behavior of rational users

With the development of machine learning and deep learning technology, the algorithm of game strategy solving is also improving. Moreover, for dynamic application scenarios, the game strategy solving ability of real-time processing of new scenes and data has been significantly improved.

Accordingly, the design of the game mechanism also needs to be improved at the algorithm level to keep up with the continuous improvement of the performance of the game strategy solving algorithm. The studies to explore this problem have recently been published in Nature and Science journals. Deepmind has developed a human-centered democratic AI mechanism design algorithm [73] based on reinforcement learning to validate AI that satisfies human values based on an online investment game. Salesforce has developed a system called The AI Economist [183], which introduces reinforcement learning to tax policy design for the first time through the two-level RL model. In addition, the framework was used to design policies to deal with the COVID-19 epidemic based on historical data and achieved good results.

Nevertheless, this is still in its infancy and needs to be explored in depth by researchers, such as how to address the robustness of two-layer reinforcement learning better and how to explain the mechanisms designed by deep learning.

3.2 Diversity of user requirements

The demand of users will become more complex because of the limited budget, incomplete information and bounded rationality of the markets. Traditional approaches of mechanism design, such as the VCG mechanism, cannot deal well with problems such as the allocation of remuneration with budget constraints. Therefore, new methods are needed to design mechanisms to meet budget

constraints. Similarly, in the case of incomplete information, the complexity of calculating the maximum benefit is also an inevitable difficulty in designing the mechanism. In addition, for the participants with bounded rationality, the assumptions in the traditional theory will fail, and new situations need to be explored.

Earlier, Che [25] studied the performance of the first-price and second-price auctions when the participants had economic constraints and proved that the existence of constraints makes the final results of the mechanism inconsistent. Its constraints include marginal expenditure costs increase, opportunity costs increase and agent's moral hazard. It follows that budget constraints have a significant influence on mechanism design.

Recent years, new budget constraint problems have been studied, including multi-unit actions [36], multi-objective [12], incomplete information [121] and approximate solution [141]. Paul [41] demonstrated incentive compatibility, individual rationality and Pareto optimality of heterogeneous goods and budget constraints auction, which are essential contributions. He also applied statistical machine learning techniques to design payment rules for the first time by replacing incentive compatibility with the minimization of expected ex-post regret [40]. Based on the proposed of RegretNet, follow-up explorations [42, 52] are constantly emerging.

In the emerging data market [37], data are an information commodity with a fixed initial investment cost, while the marginal cost can be neglected. This new cost feature makes it no longer impossible for traditional strategies to be used in pricing [79]. The price of data depends on the requirements and valuation of consumers. However, in the data market, service providers do not know the consumer's requirements precisely, nor do they know the corresponding valuation. For new data products, the service provider does not even know the approximate market demand distribution function, forcing him to make real-time data pricing decisions in an incomplete information environment.

In the process of the maximum profit from data transactions, service providers need to consider the cost structures of specific data and online pricing decisions under incomplete information, which undoubtedly increases the complexity of pricing mechanisms in emerging markets.

Due to the late emergence of algorithmic game theory, there is less research on diverse user requirements in emerging complex markets, including limited budget, incomplete information and bounded rationality. In particular, it is lesser in the research of incomplete information mechanisms since it is commonly used for strategy solving, communications or network security. Therefore, it is urgent to explore mechanism design problems arising from users' requirements in more realistic scenes.

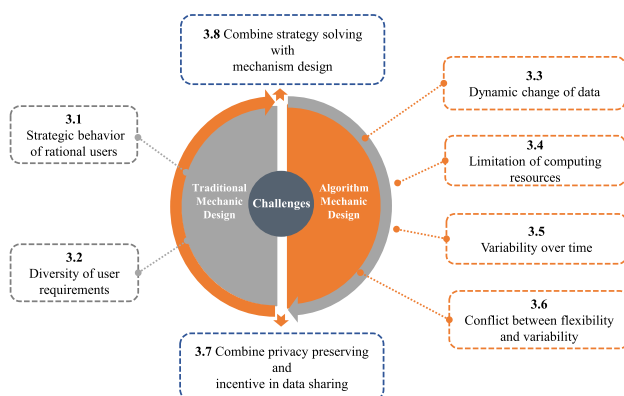


Fig. 6 Challenges in mechanism design

3.3 Dynamic change of data

Auction theory and mechanism design provide a brand-new way of thinking to solve various application problems. Unfortunately, solutions to static problems usually cannot directly translate to dynamic scenarios. For example, consider a standard second-price auction for a single item with private value. Buyers must bid for the goods and then distribute them to the bidder with the highest bid. The winning bidder pays the next highest bid. As Vickrey [154] first demonstrated, buyers bidding on actual value is a (weakly) dominant strategy, and, as a result, items are efficiently allocated. However, the second-price auction does not need to maintain its characteristics [13] in a dynamic environment.

Consider a more general situation when the buyers are uncertain about the arrival times and willingness to pay competitors. The second-price sealed bidding auction usually cannot produce effective results, even if bidders are forward-looking and entirely rational. He proved that the second price sealed bidding auctions usually cannot produce effective results, even if the bidders are forward-looking and completely rational. Therefore, to get the desired results in a dynamic environment [58], we must go beyond the most suitable tools for static environments.

In recent years the learning-based approaches have had better results in the field, such as no-regret learning [23], reinforcement learning [136], and the AI Economist [183]. Many new explorations of dynamic data have emerged from the classic multi-armed bandit (MAB) problem, including data privacy [149], auction-based mechanisms with strategic arms [56] and truthful mechanisms [9]. Significantly, the methods based on deep learning and machine learning developed late in this field. How to ensure the theoretical and increase their robustness and stability in the presence of dynamic data needs to be explored in the future. It is also a new track for computer-aided multi-item or multi-agents-based gaming methods.

3.4 Limitation of computing resources

In computational complexity theory, polynomial time means that the computational time of a problem is not greater than a polynomial multiple of the problem size n . Previous work only considered the benefit maximization problem in some specific scenarios to find the optimal solution, for example, choosing a specific benefit function or considering a simple network topology.

However, solving optimal social welfare is usually an NP-Hard problem, and it is challenging to get in polynomial time. From the point of view of algorithm mechanism design, to make the model better describe practical

application problems, we need to consider the computational complexity and limited computational resources in solving social welfare problems.

The research on how prices can be used to guide the allocation of resources when optimization is practically impossible is a new frontier. The backpack and graph coloring problems are typical NP-hard problems. However, from the point of view of mechanism design, there are still feasible schemes to reduce the time complexity. The spectrum auction theory introduced at the beginning of this section is a complex application of the graph coloring problem, which was solved by Paul Milgrom using the idea of clock auctions, for which he won the 2020 Nobel Prize in Economics. This is a challenging problem, but it is not insurmountable. Researchers [66, 82] are still exploring heuristic methods to solve the complicated situation of dynamic spectrum mechanisms.

3.5 Variability over time

In the changing scenes, the value of many tasks will be discounted over time, making the traditional pricing mechanisms impossible. Therefore, unlike traditional performance indicators (such as delay and throughput), AoI is usually used to consider the freshness of information.

There are, however, two challenges in designing such mechanisms. The priorities of real-time computing tasks are crucial to effective resource allocation but usually are private information to the users. Second, as time passes, the value of tasks can be discounted, making traditional pricing mechanisms unfeasible. Asnat [60] predicted the impact of product price elasticity based on the log–log demand model with the GBM regression algorithm without historical price elasticity information. Yang [167] explored the dynamic pricing scheme of perishable fresh agricultural products based on reinforcement learning, which provided a new direction for studying the design of value mechanisms of goods that changed with time.

There may be many variables in the mechanism that will change over time, not only the price of the item itself but also the change of equilibrium point, marginal costs and the difficulty of obtaining data. The features involved in the design of multi-attribute auctions are becoming increasingly numerous, and the changes in the mechanisms caused by this are a subject worthy of in-depth study.

3.6 Conflict between flexibility and variability

This challenge mainly comes from designing data pricing mechanisms, balancing flexibility and economic stability. The existing data pricing strategies are relatively rigid. The service providers set fixed prices for the whole dataset or several data subsets. This simple data pricing mechanism

not only forces service providers to guess the interest points of consumers to provide a subset of data that meets the market needs, but also forces consumers to buy things beyond their actual requirements. In the data market, we should adopt more sophisticated data-selling methods.

Although the flexible data pricing strategy can provide convenience for both parties to data transactions, it can also bring some traps [53], such as possible arbitrage. Specifically, data consumers can infer the results of expensive queries from a combination of low-priced queries. In particular, in perceptual data markets, the complex correlation of data provides a breeding ground for arbitrage.

A robust and reasonable data pricing mechanism should provide flexible data interfaces and satisfy the robustness, such as anti-arbitrage. Considering the complex data markets, it is a challenging research problem to design the data pricing strategy that considers flexibility and economic robustness.

Schneider [132] performed many computational experiments to analyze different mechanisms and explore robustness regarding revenue, efficiency and convergence rate. The fundamental theoretical framework of combining flexibility and robustness is not only the subject to be explored in the field of AI [49] but also the challenge to be explored in the field of mechanism design.

3.7 Combine privacy preserving and incentive in data sharing

In all kinds of complicated social scenes, data are a necessary part of game strategy solving and mechanism design. Designing incentive mechanisms to motivate participants to balance data sharing and privacy-preserving, and to provide more data for designing better mechanisms, is a problem that conflicts with and promotes each other.

The interaction between mechanism and cryptography is worthy of mention. Although both consider the problem of controlled information, they are mainly different. On the one hand, they are divergent in that mechanism design tries to enforce the disclosure of information, while cryptography tries to private the data. Besides, they have traditionally represented different models of paranoia. Although game theory assumes that all agents' expected utility is balanced, but the idea of cryptography is more straightforward. It considers the harsh environment as much as possible.

However, recent work has begun to bridge these gaps [45]. Multi-party computation (MPC) as a security model is gradually becoming possible by applying its decentralized ideas [138] to the field of mechanism design, where elliptic-curve cryptography [93] can also be applied as an aid to security mechanisms.

However, the traditional encryption methods need to be verified by the security of the mechanism design scheme, whether the decentralization of MPC is consistent with the decentralization of the correlated equilibrium in the game, and whether the fully homomorphic of quantum-resistant encryption can satisfy the high security and acceptable time complexity of the mechanism. All these can be explored in depth. The application of federated learning combines privacy-preserving and dynamic data. As one of the complex scenarios for mechanism design, it is worth studying in detail in Sect. 6.

3.8 Combine strategy solving with mechanism design

Poker is a crucial application platform for researching various strategy algorithms and is often used as a testing ground for the research of mathematics, economy, and game theory. Texas Hold'em is a game with incomplete information, while the modeling for this game can generate almost unlimited computational complexity. The latest advances in poker research for solving gaming strategies can already beat most human players.

On the other hand, auctions, as the originated field of mechanism design, can hardly leave the support of distributive rules, such as money, to maximize social welfare. Meanwhile, the rules of board games such as poker also seem difficult to change. It is a complicated and never-before-researched area to combine the fundamental theorems of mechanism design with scenarios of game strategy solving to optimize their inherently established mechanisms.

Perhaps we can start with Bang! or Legend of the Three Kingdoms, these multi-participant award-winning card games, and re-optimize their game mechanics. As a simple abstraction of military warfare, this type of card game has its special strategic significance. It may be a novel idea to design a dynamic mechanism to achieve the highest degree of game experience resulting from the voting of each participant. In addition to the democratic AI recently discussed by the Deepmind [73] team, we may be able to develop a more peaceful AI by exploring the mechanics of the Three Kingdoms game.

4 Theoretical support for mechanism design

4.1 Measurement indicators for mechanism design

That there are three basic measurement indicators for evaluating an economic mechanism, the efficient allocation of resources, effective utilization of information and

incentive compatibility. The verification of effective allocation usually uses the Pareto optimal standard, the effective utilization of information requires the mechanism to operate at the lowest information cost as possible, and incentive compatibility requires the consistency of individual and collective rationality. Therefore, the problem convert to what kind of economic system can satisfy these three requirements at the same time. How to design the mechanism to satisfy or approach these three requirements infinitely. The incentive is the core of economics, because it ultimately determines the efficiency of resource allocation. In addition, the central issue of incentive design is that the seller must consider two constraints (the incentive compatibility and the participation) when designing the mechanisms to maximize the social welfare.

We depict the main equilibria in game theory and its intensity relationships in Fig. 7. Nash equilibrium is the equilibrium state with the relatively weakest strength in game theory. No participant can improve their benefits by only adjusting their strategy in Nash equilibrium. Pareto optimal equilibrium is slightly better than Nash equilibrium. No combination of strategies can individually increase a participant’s benefits without harming others in Pareto optimal equilibrium. The dominant strategy equilibrium is the most powerful concept among all equilibriums. In dominant strategy equilibrium, any participant can maximize their benefits by using their dominant strategy regardless of the combination strategy of other participants.

4.1.1 Nash equilibrium

Nash equilibrium is the fundamental concept in non-cooperative game. The agent/player will choose the solid strategy no matter the opponents is called the dominant strategy. If both agents’ strategies are dominant, then the combination is defined as the Nash equilibrium.

Definition 1 Pure Strategy Nash Equilibrium

A strategy profile $s \in S$ is in Nash equilibrium if s_i is the best action for player i given that other players are playing s_{-i} .

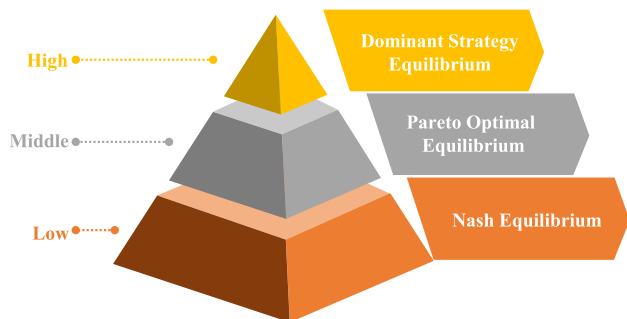


Fig. 7 Equilibrium intensity in game theory

$$u_i(s_i, s_{-i}) \geq u_i(s'_i, s_{-i})$$

Furthermore, Nash proved that the Nash equilibrium game has a limited number of players and strategies when they independently choose the strategy using a probability distribution.

4.1.2 Pareto optimal equilibrium

Definition 1 Pareto Optimal Equilibrium

In a strategy game, a strategy s_i is a Pareto optimal equilibrium for player i when and only when for that player i satisfies $u_i(s_i) \geq u_i(s'_i)$, there is no other player j whose strategy satisfies $u_j(s_i) \geq u_j(s'_i)$.

Pareto optimal resource allocation not always be the Nash equilibrium of the complete information static game. The difference with Nash equilibrium is that the latter considers individual interests, while Pareto optimal equilibrium is usually considered to maximizing the general welfare. Pareto optimality is used to measure the effective allocation of resources. It is called Pareto improvement or Pareto optimization when the change from one state to another makes at least one player better without making anyone worse.

4.1.3 Dominant strategy equilibrium

An important concept in game theory is dominant strategy equilibrium, which is defined as follows.

Definition 1 Dominant Strategy Equilibrium

In a strategy game, a strategy s_i is the dominant strategy of player i if and only if for that player any other strategy $s'_i \neq s_i$ and the set S_i of possible strategies of other players, the following equation holds.

$$u_i(s_i, S_i) \geq u_i(s'_i, S_i)$$

Intuitively, the dominant strategy enables the player always to maximize its revenue, regardless of which of the set of strategies the other players choose. Thus, a rational and selfish player will choose the dominant strategy if it exists. Putting the above definition another way, a strategy s_i is a dominant strategy for player i when and only if that strategy can dominate any other strategy of player i .

4.1.4 Informational efficiency

Informational efficiency is the question of how much information is required for an incentive mechanism to achieve a given social goal, i.e., the cost of operating the

mechanism. The excellent mechanism requires less knowledge about consumers, producers, and other participants, i.e., a lower information cost. The information is obviously expensive, so the designer must keep the dimensionality of information space lower.

4.1.5 Individual rationality (IR) and incentive compatibility (IC)

The other two basic concepts in mechanism design are individual rationality and incentive compatibility. They are defined as follows.

Definition 1 Individual Rationality (IR)

The payoff of player i who participates in the game is not less than zero. Here, we assume the payoff is zero for the player who does not participate.

$$u_i(s_i) \geq 0$$

The individual rationality constraint, also known as the participation constraint, means that if a rational agent is interested in the mechanism designed by the designer (principal), the expected utility that the agent receives under that mechanism must be no less than the maximum expected utility he did not participate the game. The maximum expected utility an agent can obtain outside the game is called the agent's reservation utility.

It is also called opportunity cost due to the agent losing the other opportunities when they participate. However, the participation constraint sometimes does not need to consider the opportunity cost. For example, suppose a resident does not have the freedom to immigrate. In this case, the government does not need to consider participation constraints when formulating tax policy. That is, all country residents must unconditionally abide by the tax law.

Definition 2 Incentive Compatibility (IC)

When the strategies chosen by the participants are their real preferences, they can obtain the most favorable revenue. Then this mechanism is called incentive compatibility (IC).

The concept of dominant strategy in the previous subsection is the basis for the incentive compatibility property. A mechanism that achieves a dominant strategy equilibrium is also referred to as having incentive compatibility.

The relevant direct-revelation mechanism here refers to a particular class of mechanisms in which the participants' strategies are restricted to those that decide to reveal information based on their personal preferences. In a direct-revelation auction, the incentive-compatible nature means that no player has the incentive to manipulate his

private information strategically when truthful disclosure of personal information is the dominant strategy for each player and maximizes player gain. A direct-revelation mechanism satisfies incentive compatibility when it guarantees that each participant maximizes their payoff when revealing their real preferences directly.

4.1.6 Strategy-proofness

When a direct revealing mechanism satisfies both individual rationality (IR) and incentive compatibility (IC), it is called the strategy-proof mechanism.

The advantage of strategy-proof mechanism is that it enables each participant to the advantage of strategy-proof mechanism is that it enables each participant to make an optimal global decision quickly based only on individual local information without considering others' strategies. And it allows the distributed network to converge to the optimal state in one step. Based on incentive compatibility and individual rationality, the strategy-proof mechanism is introduced.

4.2 Traditional mechanism design theory

4.2.1 Auction theory

Game models of auctions started in the 1960s with the seminal papers in by William Vickrey [154, 155] and Reichert [114]. Robert Wilson is also a significant contributor to the research of auction theory and has written some fundamental studies on predatory pricing, price wars and the role of reputation in other competitions. With his student Paul Milgrom, he designed the spectrum license auction for Pacific Bell, which was adopted by the Federal Communications Commission (FCC), which is an actual application of auction theory. The two scholars were recently awarded the 2020 Nobel Prize in Economics for improving auction theory and new forms of auctions.

The wave of auction theory and market design [5, 7] is from the FCC spectrum auction and the combinatorial auction problem. The problem studies how to sell and allocate items to maximize the overall benefits when sellers have several items (e.g., spectrum) for sale and when buyers do not simply equal the valuation of multiple items to the sum of their valuations of individual items.

Market design, which focuses on the application of mechanism design, has existed in microeconomics for decades. However, it was not until recent years that the tremendous success of the automatic kidney transplant system [2, 8, 33, 46] in the USA aroused widespread concern in the computing field. From the point of view of calculation, the algorithm implementation of this mechanism is a non-deterministic polynomial completeness

(NPC) problem. Designing the practical mechanism is an exciting problem in computer science. The wave is still developing, and much research work is expected to emerge in the coming years.

The classic approach to solving the auction mechanism design problem is the Vickrey–Clerke–Grove (VCG) mechanism, a generic term for a class of direct display mechanisms in environments where actors have linear-like preferences, which is the only class of mechanisms that simultaneously satisfy validity, individual rationality, and incentive compatibility. However, it has been proved that this mechanism does not have polynomial time realization [14], so the algorithm needs to be improved. Representative works to design fast search algorithms to realize VCG mechanism and an effective approximate VCG that is close to the optimal global benefit. Shen [137] made new progress in coupon research by optimizing the VCG auction theory and applying it to online advertising. The development of the industrial internet accompanies the combination of auction theory and computer technology. Many e-commerce companies based on combined auctions have appeared, such as trading dynamics, combined and so on. Before defining the VCG mechanism, we need to define auction theory’s allocation and payment rules and the important Myerson lemma.

By using easy-to-understand and straightforward auction examples, researchers can better understand the ideas and theories of mechanism design and further explore and improve the mechanism design theory. As shown in Fig. 8, the typical auction mechanism can be divided into three stage.

First of all, the principal needs to design a mechanism. Here, the mechanism is the game rule, according to which each agent sends a signal (e.g., buyer’s offer), and the signal decides the configuration result (e.g., who gets the auction item and what price to pay); second, the agent

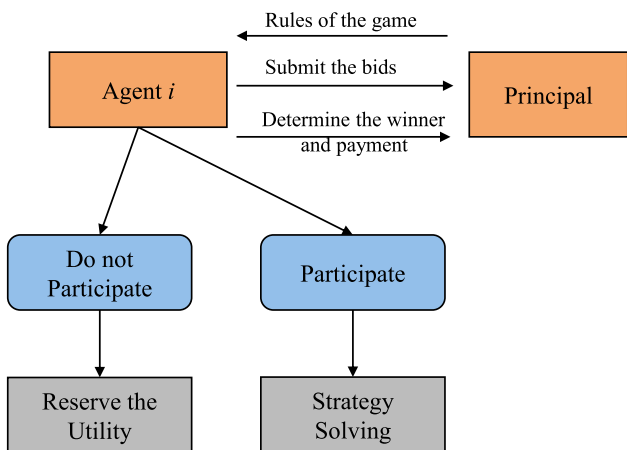


Fig. 8 Typical auction mechanism and game strategy solving process

chooses to accept or not simultaneously. If the agent does not select the accept, he will get the exogenous appointment effect. Finally, the accepted agent will participate in this mechanism. It can be seen that the game decision and mechanism design are complementary and antagonistic.

(1) Allocation and Payment Rules

Usually, the seller decides the auction rules during the auction process, and he must make the rules only when he has limited knowledge of the bidder’s willingness to pay. The main auction types can be divided into open outcry and sealed bids.

In open outcry auctions, bidders bid openly, and all bidders can observe the bidding price. This auction may be the most suitable for people’s understanding of auctions. There are two open auctions, only one of which may have a “crazy” bid. The ascending auction (English auction) is the type that best accords with the popular image of open bidding. It is also the alternative name for the standard auction held by British auction houses (such as Christie or Sotheby).

Another type of open outcry auction is the reduced-price auction (Dutch auction). Contrary to an English auction, the auctioneer starts with a very high price and then calls for lower and lower prices until a bidder accepts the price and bids, thus winning the auction. For speed reasons, Dutch-style flower auctions, as with other auctions of produce or perishable goods, use a “clock” to count down each bid until someone “stops the clock” and takes the item. In many cases, the auction clock reveals a lot of information about the current items for sale and the price reductions.

The second-price auction is the sealed bid, where bids are made in secret, bidders cannot observe the information of others, and in many cases, only the winner’s bid is announced. In this type of sealed-bid auction, as in Dutch auctions, bidders have only one chance to bid (only the highest price is relevant to the bid outcome). A sealed auction does not need an auctioneer, only a supervisor to announce the bid and decide the winner.

In a sealed auction, the price paid by the winning bidder is determined in one of two ways. (1) Under the first-price sealed bid auction, the highest bidder wins the item and pays the price it bids. (2) Under the auction with the second price, the highest bidder wins the item but only pays the second highest bid.

For example, in the sealed-bid auction, the mechanism designer needs to make two decisions: one is who gets the auction item, and the other is to set how much each person needs to pay. These two decisions are called the allocation and payment rules, respectively. The specific steps are shown below.

Definition 3 Allocation and Payment Rules

1. Collect the bids $\mathbf{b} = (b_1, \dots, b_n)$ for all the agents. We refer to the vector \mathbf{b} as the bid vector.

2. Allocation Rule: Choose a feasible allocation $x(\mathbf{b}) \in X \subseteq R^n$, which is a function of the bid vector.

3. Payment Rule: Choose a payment $p(\mathbf{b})$, which is also a function of the bid vector.

Finally, bidders can evaluate the value of auction items in many ways. The most crucial distinction in this auction environment is between the values of the items. In the common or objective value auction, the value of the auction item is the same for all bidders, but everyone only knows the inaccurate estimation of their value. Bidders have some knowledge of its possible distribution, but each bidder must evaluate it before bidding. In the common value auction, every bidder should know that the other bidders have a rough understanding of the value of the auction. He should try to infer this information from the bidding behavior of competitors. In the private or subjective value auction, each bidder decides the value of the auction item to them. In this case, each bidder will evaluate the same auction item differently.

(2) Myerson lemma

In this subsection, we introduce a crucial sufficient necessary condition for designing a unidimensional auction mechanism (i.e., the auction participants have only unidimensional private information) to satisfy the strategy-proofness.

According to Myerson’s theorem, the auction mechanism satisfies strategy-proofness, when and only when the winner selection rule (or resource allocation rule) is monotonic. Before introducing Myerson’s theorem, let us introduce enforceable allocation rules and monotone allocation rules.

Definition 4 Enforceable Allocation Rules

For an allocation rule x in a univariate setting, have a payment rule p such that the direct display mechanism (x, p) is dominant strategy incentive compatible (DISC). Then, this allocation rule x is said to be enforceable. That is, those allocation rules that can be extended into DISC mechanisms are enforceable.

Definition 5 Monotone Allocation Rules

An allocation rule x is monotonic if, for each agent i and the bid vector b_{-i} of all other agents, the allocation function $x_i(z, b_{-i})$ for agent i is a monotonic non-decreasing of i ’s bid function z . That is, under a monotonic allocation rule, higher bids will win more items.

Definition 6 Myerson’s Lemma

For a univariate environment:

1. An assignment rule x is enforceable when and only when it is monotonic.

2. If x is monotonic, then there is a unique payment rule such that the direct display mechanism (x, p) of the sealed auction mechanism is dominant strategy incentive compatible (DSIC).

3. The payment rules of (x, p) have exact expressions.

(3) VCG Mechanism

The full name of the VCG mechanism is the Vickery–Clarke–Groves mechanism [154, 155], which comes from the names of three scholars, so VCG is used as its abbreviation. It provides the conclusion that “a DSIC mechanism that maximizes welfare is enforceable” for multiple private parameter scenarios.

Designing such a mechanism is very difficult, and we take a two-step approach. First, assuming that the mechanism is DSIC, and everyone bids according to the facts, and design an allocation rule to maximize the welfare $\sum_{i=1}^n x_i v_i$. Second, design a payment rule after obtaining the allocation rule, and the guarantee the mechanism is DSIC. Myerson’s Lemma proves that for a monotonic allocation rule, there is a unique and computable payment rule that satisfies DSIC. But that is for univariate and may not apply to multivariate, even if we cannot define what monotonic means in a multivariate setting. Therefore, we have to find a different way of thinking.

The VCG mechanism takes advantage of the externality induced by agent i to set prices. An externality is the welfare loss to others caused by the presence of i , i.e., the sum of the social welfare of others when i is not present minus the sum of the social welfare of others after i is present (excluding himself).

Definition 7 VCG Mechanism

The allocation and payment rules of the VCG mechanism satisfy the following two equations, respectively.

$$x(\mathbf{b}) = \arg \max_{\omega \in \Omega} \sum_{i=1}^n b_i(\omega)$$

$$p_i(\mathbf{b}) = \left(\max_{\omega \in \Omega} \sum_{j \neq i} b_j(\omega) \right) - \sum_{j \neq i} b_j(\omega^*)$$

An alternative formula representation of the payment rules in the VCG mechanism is given below.

$$p_i(\mathbf{b}) = b_i(\omega^*) - \left[\sum_{j=1}^n b_j(\omega^*) - \max_{\omega \in \Omega} \sum_{j \neq i} b_j(\omega) \right]$$

4.2.2 Contract theory

The contract theory in economics is critical. It started with the classic paper “The Nature of the firm” by Coase [28],

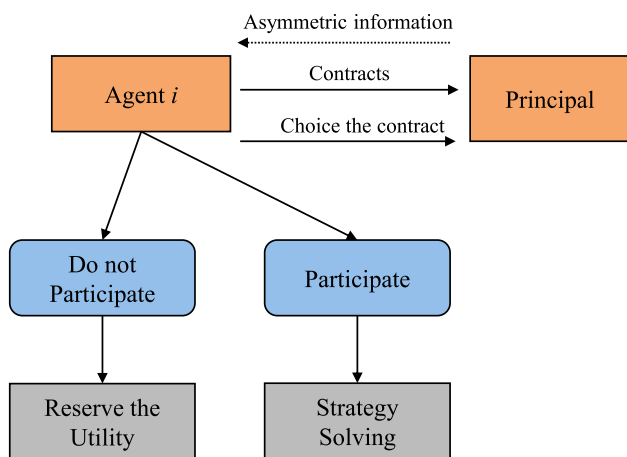


Fig. 9 Typical contract mechanism design and game strategy solving process

winner of the 1991 Nobel Prize in Economics in 1937. Coase points out that the longer the contract duration regarding the supply of goods or services, the more impossible and inappropriate it is for the buyer to specify what the other agents should do.

Contract theory is one of the applications of game theory, mainly used to study the economic behavior of transactions between different contractors. Contract theory is one of the branches of economics that has developed rapidly in recent 30 years. It can solve the problem of information asymmetry in transactions between different contractors, such as incentive theory, complete contract theory and incomplete contract theory [179]. The assumed constraints and models designed in contract theory can simplify the tedious transactions between different contractors and get relevant theoretical opinions. Contract theory [28, 119] is also a promising and widely adopted theoretical tool for dealing with private information problems. In 2016, economists Oliver Hart and Bengt Holmström were awarded the Nobel Prize in Economics for their contributions to contract theory.

Contract theory is widely used in economics to solve the problem of information asymmetry between employers and employees, sellers and buyers by entering into contracts [15]. The asymmetric information problem includes the hidden information problem and the hidden action problem. In the field of communication, contract theory has also been applied to set incentives and solve maximization problems. For example, in the literature [178], it is considered that transmitting the data to other users will bring different fees. The authors propose a contract-theoretic method to design the pricing mechanism and create an appropriate match between users who need and are willing to relay data.

The typical process of contract mechanism and game strategy solution is shown in Fig. 9. Unlike auction theory, contract theory needs part of agent information, widely used in cooperative games. Through some theorems and examples, this paper deepens the understanding of the contract theory. The motivation for considering cooperative games lies in the following example: In many games, the Nash equilibrium income is not optimal compared to other unbalanced results. In the following, we refer to a modified version of the Prisoner’s dilemma problem, which has the following payoff matrix.

In the game above, the only equilibrium is (y_1, y_2) , and the payoff from this equilibrium is $(1, 1)$. However, the table shows that the non-equilibrium outcome (x_1, x_2) produces a higher payoff $(2, 2)$. At this point, the participants may wish to transform the game so that its equilibrium set includes better outcomes, and there are several ways to achieve this transform. The participants express their agreement in the form of a contract. The participants commit to coordinating their behavior. The participants play a repeated game.

We build this transformation on the basis of the first method, in which the participants who sign the contract must act according to the established strategy, which is called correlated strategy.

Participant i is willing to sign a contract and thus choose the correlated strategy α only if $u_i(a) \geq v_i$. This is called the individual rationality or participation constraint of participant i . This leads to the following definition.

Definition 8 Individual Rational Correlated Strategy

Given a correlated strategy $\alpha \in \Delta(S_1 \times \dots \times S_n)$ for all participants in N , if the following equation is satisfied.

$$u_i(a) \geq v_i, \forall i \in N$$

Then the correlated strategy is individually rational.

Contract theory converts a game with fewer consensual equilibria into more consensual equilibria. In the game shown in Table 1, let two participants sign the following contract in Table 2.

1. If both agents sign the contract, then agent 1 chooses strategy x_1 and agent 2 chooses x_2 .
2. If only agent 1 signs the contract, then agent 1 will choose y_1 .
3. If only agent 2 signs the contract, then agent 2 will choose y_2 .

The act of signing a contract by participant i is called a_i . Now we can expand the set of strategies of each of the two agents as $S_i = \{x_i, y_i, a_i\}$. The payoff matrix for the transformed game is shown below.

At this point, the transformed game has a new equilibrium (a_1, a_2) , which has a payoff of $(2, 2)$. Although the strategy (y_1, y_2) also constitutes an equilibrium, it is not a

Table 1 Payoff matrix

		Agent 2	
		x_2	y_2
Agent 1	x_1	2, 2	0, 6
	y_1	6, 0	1, 1

Bold represents their equilibrium solutions

Table 2 Payoff matrix for the transformed game

		Agent 2		
		x_2	y_2	a_2
Agent 1	x_1	2, 2	0, 6	0, 6
	y_1	6, 0	1, 1	1, 1
	a_1	6, 0	1, 1	2, 2

Bold represents their equilibrium solutions

dominant strategy equilibrium. From this, we give the following definition.

Definition 9 Contract Theory

Consider the vector $\tau = (\tau_C)_{C \subseteq N}$, note here that $\tau \in \times_{C \subseteq N} (\Delta(\times_{i \in C} S_i))$.

The vector τ of correlated equilibria of all possible coalitions is a contract.

The correlated strategy τ_C for alliance $C \subseteq N$ is that the correlated strategy τ_C will be implemented by the members in alliance C if it is composed of those participants who signed the contract.

4.2.3 Bargaining problems

Apart from auction and contract theories, many market transactions and coordination can also be modeled by bargaining theory. Bargaining theory, as a branch of game theory [103, 108, 109], is a nonzero-sum game theory developed with the continuous improvement of game theory. Agents can solve the problem of profit distribution through consultation.

Nash proposed the Nash bargaining solution to the bargaining problem [109] as early as 1950. The general model for problem-solving is established in this paper and given an axiomatic solution method, which paves the way for the future study of axiomatic solutions to bargaining problems.

The two-player bargaining problem has been applied to many important situations, including management of labor contract and trade union negotiation in labor arbitration. Negotiation is between two countries or groups of countries in international relations, such as reduction of nuclear weapons, military cooperation, and counter-terrorism strategies; in the duopoly market game, two competing enterprises negotiate the output to maximize the total

income; bargaining between buyers and sellers in bilateral trade; moreover, in the supply chain cooperation, buyers and sellers negotiate mutually beneficial contracts to promote long-term relationships between the two parties. The solve dispute between companies and individuals in property rights disputes. The following abstraction of various problems defines the two-person bargaining problem and its Nash bargaining solution.

Definition 10 Two-Person Bargaining Problem

The question consists of three elements:

1. Agent 1 and Agent 2;
2. Feasible distribution set S (containing the case of negotiations failed d);

The allocation of two bargains is generally denoted by $s = (s_1, s_2), s \in S$, where s_1 and s_2 represent the distribution of two agents, respectively.

3. Utility allocation set U .

The allocation of utility is generally denoted by $u = (u_1, u_2), u \in U$, where u_1 and u_2 represent the expected utility of agents, where $u_i : S \rightarrow R$ is the is the real-valued function from the feasible distribution set S to the set of real numbers R .

For a bargaining model to be meaningful, the following equation must be satisfied.

$$\forall s \in S, u_1(s) \geq u_1(d), u_2(s) \geq u_2(d)$$

$$\exists s \in S, u_1(s)u_1(d), u_2(s)u_2(d)$$

All $s \in S$ is assignable and there exists at least one $s \in S$ such that the expected utility of the agents is greater than the negotiations failed point. Denote the problem as $B = (S, d; u_1, u_2)$. For $\forall s \in S$, participants can obtain a pair of utility values $(u_1(s), u_2(s))$, which is called the utility configuration of problem B . The set is $U(B) = \{(u_1(s), u_2(s)) : s \in S\}$.

There are many solutions for the bargaining problem [71], such as Nash solution, Kalai-Smorodinsky solution [70, 71], Egalitarian solution [69] and so on. In recent years, many researchers have made corresponding extensions and applications of the Nash bargaining problem, including the distribution of returns [150] in investment which is an alternative to the game associated with solving the Nash bargaining problem. Bargaining theory has been used to improve the traditional K-means algorithm [124] in machine learning and energy efficiency optimization for networks [99]. Since the Nash solution is the most representative, we show here a simple form of it.

Definition 11 Nash Bargaining Solution

For any two-person bargaining problem $B = (S, d; u_1, u_2)$, the following solution set needs to be determined:

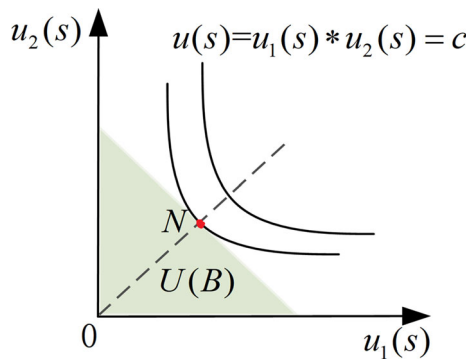


Fig. 10 Nash solution of bargaining problems

$$\sigma^N(B) = \{s \in \arg \max [u_1(s) - u_1(d)] * [u_2(s) - u_2(d)]\}$$

Without considering the specific bargaining process, now assume that there is a referee whose utility preferences are represented as undifferentiated curves in Fig. 10: $\mu(s) = \mu_1(s) * \mu_2(s)$. The tangent point N of its curve to the set of utility configurations $U(B)$ is the Nash solution of the bargaining problem. A formulaic definition of the Nash solution is given above.

5 Computer-aided approach to incentive mechanism

Although the mechanism design theoretical is very important, in the face of increasingly complicated big data Internet scenarios, computing is needed to assist the mechanism design. Internet advertising and search auctions, as the most widely used scenes in mechanism design, are naturally common research fields in algorithmic game theory. We take this as the entry point for research, and some representative papers of Internet advertising based on reinforcement learning are listed in Table 3.

Research based on the combination of mechanism design theory and machine learning algorithms is becoming hotter and hotter. As the earliest application of the latest technology, the Internet advertising field has given us cutting-edge technology. For example, the latest paper [180] combines the Actor-Critic model with the traditional Generalized Second-Price auction. It proposes a new Deep GSP Auctions model for Internet advertising.

This section explores several algorithms with a high degree of integration. Specific application scenarios, including Internet online advertising, data sharing and privacy preservation in complex environments, will be discussed in Sect. 6.

Machine learning is applied to game strategy solving, and computer-aided algorithms are essential in optimizing

algorithm mechanism design. The idea of reinforcement learning was developed at the beginning of the twenty-first century, and its core concept was perfected by Richard S. Sutton of the university of alberta. This view stems from behaviorism in psychology. That is, in the process of long-term interaction with the environment, agents can optimize their behavior by trial and error or searching for a memory.

In the reinforcement learning process, the algorithm needs the best response strategy to ensure that an agent can get the maximum expected return in the constantly changing dynamic environment [139]. One of the strategies consists of a series of continuous actions, which correspond to the agent's response to the environment in the corresponding environment state. Figure 11 represents the environment exchange process of a single agent.

Many tasks can correspond to reinforcement learning ideas. For example, an information search task can be regarded as a sequential interaction between an RL agent (the system) and a user (environments), in which the agent can constantly update the policy based on real-time feedback from the environment during the interactions until the system converges to the optimal policy that generates an object that best matches the user's dynamic preferences. Second, the RL framework's goal is to maximize users' cumulative long-term return. Therefore, the agent can identify objects with small immediate returns but essential to long-term returns. Because of the advantages of reinforcement learning, people are very interested in developing a technology based on reinforcement learning.

Mainstream reinforcement learning algorithms do not require state prediction. They do not consider how actions affect the environment, so they need little prior knowledge. Theoretically, it is an effective way to solve the problem of autonomous learning in complex and changeable environments. However, the complexity of the reinforcement learning algorithm increases exponentially with the increases of state action space, so it is difficult to break through the limitation of high-dimensional action space. The following are the representative structural modeling, value function-based and policy gradient-based reinforcement learning methods.

Theoretically, it is an effective method to solve the problem of autonomous learning in a complex and changeable environment. However, the complexity of reinforcement learning algorithms grows exponentially with the growth of state-action space, so it is difficult to break the limits of high-dimensional action space. The representative structural modeling schemes based on the value function and the policy gradient reinforcement learning are below.

Table 3 Internet advertising research based on reinforcement learning

	Year	Academic paper	Methods	Main contribution
[181]	2021	DEAR: Deep Reinforcement Learning for Online Advertising Impression in Recommender Systems	DQN, Rec/Ads Trade-Off	Consider the negative impact of advertising on the recommended user experience to maximize returns in the long run
[180]	2021	Optimizing Multiple Performance Metrics with Deep GSP Auctions for E-commerce Advertising	Actor-Critic, Generalized Second-Price	New rating functions are designed using deep learning within the well-known GSP auction framework and applied in e-commerce advertising
[17]	2021	Reinforcement Learning of Sequential Price Mechanisms	Partially Observable, MDP	Combining RL with sequential price mechanisms, the method can learn optimal or near-optimal mechanisms
[26]	2019	Pricing Average Price Advertising Options When Underlying Spot Market Prices Are Discontinuous	Jump-Diffusion Stochastic Process, Monte Carlo	The new Advertising options pricing framework is proposed, and a display pricing formula is derived
[162]	2018	A Multi-Agent Reinforcement Learning Method for Impression Allocation in Online Display Advertising	Multi-Agent RL, MDP, Policy Optimization	Formulate the impression allocation problem as an auction problem in which each contract can submit virtual bids for individual impressions
[161]	2018	Budget-Constrained Bidding by Model-free Reinforcement Learning in Display Advertising	Model-Free RL, MDP	Solving the optimization problem with an innovative reward function design method
[38]	2019	Infer Your Enemies and Know Yourself, Learning in Real-Time Bidding with Partially Observable Opponents	Actor-Critic, Deep Attentive Survival Analysis	Generalized multi-agents framework for inferring missing data, developed mean field equilibrium analysis for second price auctions
[97]	2021	Real-time bidding campaigns optimization using user profile settings	Dynamic Programming	Complementary strategies to increase profitability were considered
[129]	2021	A Unified Optimization Framework for Auction and Guaranteed Delivery in Online Advertising	Column Generation Strategy	Combining pay-per-click auctions and guaranteed delivery with the different marketing objectives of advertisers results
[112]	2016	A Combinatorial-Bandit Algorithm for the Online Joint Bid/Budget Optimization of Pay-per-Click Advertising Campaigns	Gaussian Processes, Bayesian Bandit	Modeling Display Advertisement as a MAB problem, a novel dynamic contextual MAB approach is proposed
[163]	2018	Budget-Constrained Bidding by Model-free Reinforcement Learning in Display Advertising	Model-Free RL, MDP	An algorithm for online joint bid/budget optimization of multi-channel campaigns was investigated in paid point-of-sale advertising
[106]	2018	Data efficient hierarchical reinforcement learning	Layered RL, Model-Free RL	Be used with a moderate number of interaction samples

5.1 Markov decision process

While Markov decision processes are fundamental to understanding reinforcement learning, they are necessary for understanding problem abstraction. Parkes [115] designed an online VCG mechanism design based on the MDP method, which implements optimal strategies in a truth-revealing Bayes-Nash equilibrium to maximize the total long-term value of the system. It can be seen that the MDP method is also one of the critical methods of computer-aided incentive mechanisms.

The Markov decision process can be represented as a quintuple (S, A, T, γ, R) . Among them, S is the state space; A is the behavior space, $T : S \times A \times S \rightarrow [0, 1]$ is the transfer function; $\gamma \in [0, 1]$ is the discount factor; $R : S \times A \times S \rightarrow R$ is the reward function. The transfer

function represents the probability distribution of transfer to the next state, given the current state and behavior. For $\forall a \in A, \forall s \in S$:

$$\sum_{s' \in S} T(s, a, s') = 1$$

In the above equation, s' denotes the possible states at the next moment. The payoff function represents the payoff obtained in the next state, given the current behavior and state. The MDP has the following Markov property: the agent's next state and payoff depend only on the agent's current state and behavior. The agent strategy $\pi : S \rightarrow A$ is defined as the probability distribution of the agent's behavior in a given state. Agent strategy $\pi(s, a)$ should satisfy the following equation.

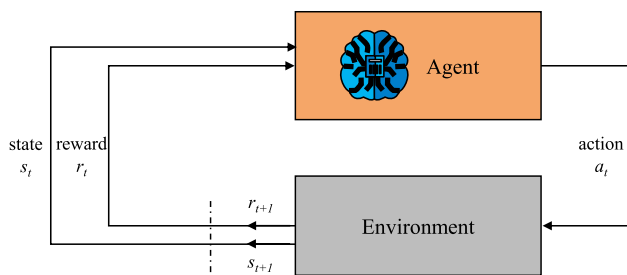


Fig. 11 Environment exchange process of single agent

$$\sum_{a \in A} \pi(s, a) = 1, \forall s \in S$$

In any MDP model, there exists a deterministic optimal strategy for the agent, where $\pi^*(s, a) \in \{0, 1\}$. The goal of an agent in an MDP is to maximize the expected long-term payoff. To evaluate an agent strategy, the following state-value function is required. When an agent starts in state s and subsequently executes strategy π , the value of state s under that strategy is defined as the expected payoff. As a result, the state-value function is:

$$V^\pi(s) = E_\pi \left\{ \sum_{k=0}^T \gamma^k r_{k+t+1} \mid s_k = s \right\}$$

In the above equation, T is the final moment; t is the current moment; r_{k+t+1} is the direct return obtained at the moment $k + t + 1$; and $\gamma \in [0, 1]$ is the discount factor.

Based on MDP, many optimization improvement algorithms exist, such as the optimal strategy-solving problem for non-complete information games that can be modeled using the Partially Observable Markov Decision Process (POMDP). POMDP is an extension of the Markov Decision Process. It can be transformed into an MDP model and solved using reinforcement learning-related algorithms [139].

5.2 Deep Q network

Early research on deep reinforcement learning algorithms focused on value function-based. Moreover, Q-learning is the milestone of the value function algorithm. Mnih et al. [100] proposed Deep Q-Network (DQN) algorithm by

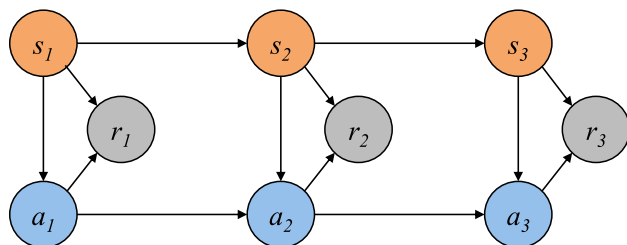


Fig. 12 Example of MDP

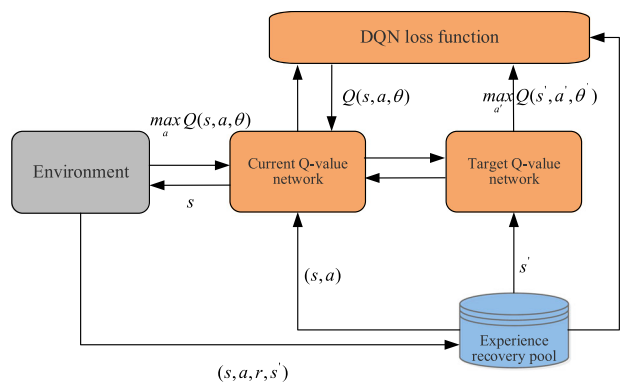


Fig. 13 DQN model structure

combining a convolutional neural network with a Q-learning algorithm in traditional reinforcement learning. Furthermore, Double DQN [151] and DRQN [63] are the improved algorithms based on it which as shown in Fig. 12.

The DQN algorithm [64] uses a deep neural network to approximate the $Q(s, a)$ function. Because the goal of reinforcement learning Q -value is dynamic, and deep learning usually needs a fixed training goal, directly using Q -value as the training goal will lead to excessive oscillation in the training process. It is difficult for training to converge. DQN, therefore, proposes an isomorphic Q -value network for fitting the agent’s state action-value function and a target network for representing the optimization goal at the current stage. DQN updates the Q -value network in real-time according to the agent’s action state and updates the parameters of the target network every fixed time based on the current Q -value network, whose model architecture is shown in Fig. 13 below.

The loss function of the DQN algorithm is shown in the following equation, where $r + \gamma \max_{a'} Q(s', a', \theta')$ is the target Q value and $Q(s, a, \theta)$ is the predicted actual Q value.

$$L(\theta) = E \left[(r + \gamma \max_{a'} Q(s', a', \theta') - Q(s, a, \theta))^2 \right]$$

5.3 Actor-critic model

Policy gradients-based deep reinforcement learning methods can directly fit the agent’s policy space and use neural networks to predict the optimal action of the agent in the current state. This section focuses on the Actor-Critic algorithm, the reinforcement learning framework that combines the value function and the strategy gradient method and combines the advantages of estimating the state value function and directly optimizing the strategy space.

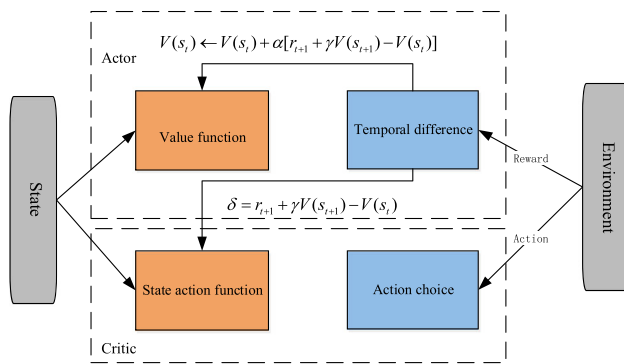


Fig. 14 Actor-critic structure

The Actor [145] determines the action strategy of the agent in the environment. Critics judge the current strategy by estimating the function of the state value of the action strategy. After combining with deep learning, which can represent actors and critics, respectively, the actor module is updated by using a gradient rising method. The evaluation module estimates the state value function by calculating the time differential error. Figure 14 shows the structure of the model.

In addition, the A2C and A3C [101] algorithms are improvements of the AC algorithm that optimize the algorithm's performance by introducing the concept of dominance function and multiple agents learning asynchronously, respectively.

6 Economic and complex environment applications

There is a solid theoretical foundation for studying economic problems in the design of the algorithm. It is a good starting point to discuss and optimize the game theory and mechanism design of data-driven algorithms from the perspective of economic problems. As a typical internet application of economics, the optimal mechanism design of advertising provides a testing ground for researching cutting-edge theories and algorithms. Figure 17 shows the significant contributions made by researchers in the Internet advertising environment. It lists the representative literature, including their methods and significant contributions, which are realized by combining reinforcement learning and other algorithms.

We list the research points in online advertising applications on the Internet from a global perspective and discuss the current state of research in several of these branches in detail in Sect. 6.1. In addition, we explore the research on incentives in complex environments concerning data sharing and privacy-preserving in Sect. 6.2. In

addition, in Sect. 6.3 we give the current state of research on other scarce resource fee allocation scenarios.

6.1 Optimal mechanism design of internet online advertising

Pricewaterhouse Coopers (Pw) made a financial report on internet advertising in 2021. It shows that since 2010, the compound annual growth rate (CAGR) of total Internet advertising revenue has been 18% (due to the impact of the epidemic, which is slightly lower than the 19% in 2010–2019 but compared with other industries Internet advertising is an undoubted hot spot. industry). Mobile CAGR was 66.5% over the same period (down from 73.8% in 2010–2019), while desktop growth remained modest at 4.6%. Figures 15 and 16 are from the IAB/PwC Internet Ad Revenue Report, FY 2020. In the figure, in addition to the growing revenue of Internet advertising, we also found three main popular areas of Internet advertising: search, display and digital video, which account for about 92.4%. Below we, respectively, research these three popular areas of Internet advertising.

6.1.1 Real-time bidding

Internet advertising is an essential economic environment in the current stage of the big data scenario. It is a typical application scenario of scarce resource allocation, i.e., how to allocate the resources of advertising space to the right users. However, it needs to be calculated in real-time by extensive data calculations to maximize social welfare, including but not limited to maximizing revenue, increasing click-through rate (CTR) and return on investment (ROI). A recent study on this aspect by ByteDive [168] also considered another point that cannot be ignored, i.e., the possible negative impact of advertising on user experience, through an improved DQR model that considers maximizing revenue in the long run.

Therefore, online display advertisements, search advertisements and digital videos cannot be separated from real-time bidding. RTB is one of the most important developments of advertising distribution mechanisms in the past decade [21, 35, 57, 68, 68, 112, 126, 133, 146, 147, 159, 161, 166, 171], because it is widely used by major online advertising platforms, including but not limited to Google, Baidu [147], Facebook and Amazon. In RTB for display advertising, a user visiting a web page immediately triggers an auction held by an ad exchange. The auction's winner is given advertising space and pays the price to the publisher.

6.1.2 Online display advertising

Considering the importance of online display advertising, this subsection summarizes the existing literature research methods from the perspective of every participant in the display advertising ecosystem. Figure 17 presents a comprehensive view of the online display advertising ecosystem, including advertisers' real-time bidding decisions regarding selling advertisers, allocating exposure and other scenarios. Problems faced by intermediaries in matching advertisers and publishers. Challenges such as the pricing of publishers (publishers of the system suppliers) and the design of auction mechanisms based on different perspectives, the researcher can clearly understand and explore display advertising research in the Internet scenario.

6.1.3 Online search advertising

Among all search advertising auctions, generalized second-price auctions have been widely studied [20, 27, 44, 75, 127, 152]. A significant part of this work focuses on the nature of social welfare in equilibrium to measure the price of anarchy [22, 59, 77, 85]. Another important part focuses on income guarantees [43, 86, 153], aiming at explaining the income guarantees of generalized second price auctions under different equilibrium concepts.

The mechanism design requires computer-aided algorithms to approximate the best incentive mechanism in data-driven real-world scenes. The literature [81] focuses on the revenue maximization problem in the search advertising market. They assume that the historical data satisfy the Markov model to understand the buyers' behavior in the search advertising market. The literature [102] investigates the revenue maximization problem in an online pricing auction scenario from an experimental perspective. The literature [180] combined generalized

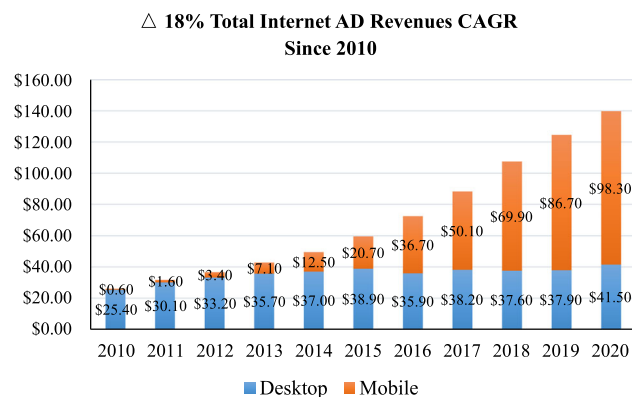


Fig. 15 Desktop versus mobile full-year internet AD revenues (2010–2020)

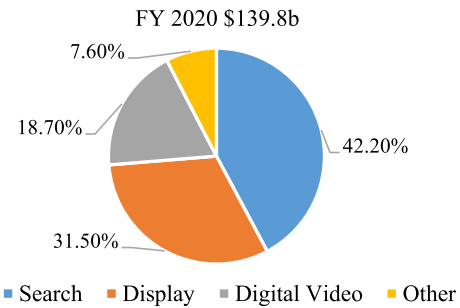


Fig. 16 2020 advertising format by share

second-price auction with the Actor-Critic algorithm to propose a new scoring function, and the experiments achieved better results.

The literature [147] presents a representative framework for designing an online advertising search mechanism, which includes Markov Bidder Model, which determines how bidders adjust their bids based on KPI feedback. In this mechanism, the bidders interact with the seller's behavior and receive KPI as feedback. The literature translates the revenue-maximizing household problem in designing the sponsored search auction mechanism into a Markov decision process in reinforcement learning $(N, S, R, G, REV(s, r), \gamma)$. The number of daily queries per keyword in the search engine is assumed to be constant. Therefore, KPI is entirely determined by the bid distribution s and reserved price curve r .

6.2 Incentives for data sharing and privacy preserving

In complex social scenes, data are a vital link. Designing incentive mechanisms to motivate participants to balance data sharing and privacy-preserving and to provide more data for designing better mechanisms is a hot topic. The profit of data transactions can be defined as the difference between the revenue of data transactions and the expenditure during data collection.

Data are an emerging electronic product with unique cost characteristics: fixed initial and negligible marginal costs. Therefore, the revenue that data can generate in the market is determined by consumers' demand, not the cost of data collection. Based on the prior knowledge of market demand, data service providers can express the expected revenue from data as a function of market demand. However, as the market demand is usually complex and diverse, it is challenging to conduct an in-depth mathematical analysis of the income and profit of data transactions. Even if we can accurately calculate the optimal expected return, the profit maximization problem can be transformed into the cost minimization problem, which is an NP-hard problem. The optimal solution cannot be obtained in

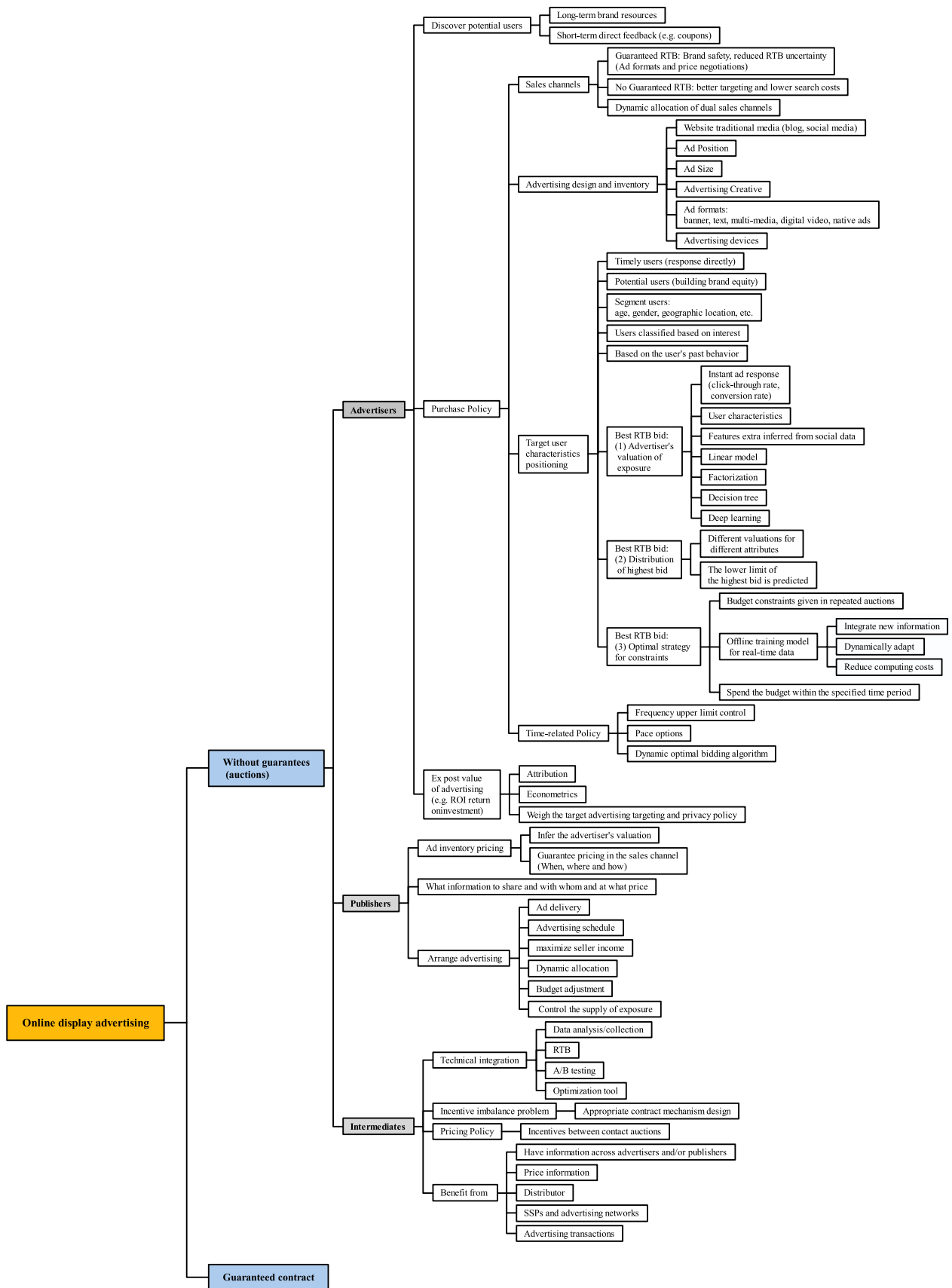


Fig. 17 A comprehensive view of online display advertising research

polynomial time. Therefore, to realize the optimal profit of data transactions, we must consider the complicated market demand and simultaneously reduce the computational complexity of solving the optimal solution.

Besides, the interaction between mechanism design and cryptography is worth mentioning. Traditional mechanism design assumes that the set of participants is fixed and the mechanism (market owner) knows it in advance. However, market owners can only have direct contact with a few participants. Therefore, owners usually need expensive promotions to recruit more participants to get the expected results, such as social welfare or income maximization. On the other hand, various countries and organizations are strengthening data privacy protection. The EU General Data Protection Regulations (GDPR), enacted in 2018, opened a new era of privacy-preservation. In December 2019, Twitter announced a new privacy center to update its privacy policy. Mozilla has announced plans to allow users to delete their data from Mozilla's servers starting January 1, 2020.

The incentive mechanism of design game theory has been widely studied in other fields, such as crowdsourcing sensing and edge computing [97, 162]. Furthermore, as a feasible data privacy solution, federated learning has recently been applied in more and more fields, and the framework of federated learning is shown in Fig. 18. However, quantifying the value of each client's training data is a challenge in federated learning. In addition, because of the complexity of federated learning algorithms, it is not easy to model the final learning performance of the federated learning system. Therefore, it is challenging to model the utility function of each participant (parameter

server and client) in federal learning, which makes existing incentive design work unable to be directly applied. This subsection focuses on federal learning and discusses the application of incentive mechanisms in data sharing and privacy-preserving.

By applying encryption algorithms with different security levels, such as semi-homomorphic encryption, secure multiparty computing and other technologies, the federated learning framework can be secure. However, participating in federated learning will lead to computing resources of client. Designing an appropriate incentive mechanism to make students actively participate in the online training of federated learning is another problem that needs to be solved urgently. The Shapley value is the main criterion to measure the contribution and profit distribution [142, 157]. The Stackelberg game is also another [51, 157, 174]. The Stackelberg game is also another better solution. In addition, blockchain and other technologies, the Stackelberg game is also another [83], are occasionally used as auxiliary technologies in mechanical design. The following section explores applying the two most important theories in microeconomics, namely, auction theory and contract theory, to federal learning.

6.2.1 Auction theory-based federated learning incentive mechanism

Auction theory is applied to many fields, and its underlying theory is described in Sect. 4. This subsection explores in depth the current work and future research points in the popular area of auction theory-based federal learning incentive mechanism design as an example.

The literature [29, 32, 67, 76, 84, 172, 173] represents the incentive mechanism of federal learning based on auction theory. The literature [173] is based on the classic literature [24] in microeconomics, which is a classic literature in the field of multi-dimensional auctions in 1993. In the procurement of weapons by the Ministry of National Defense, in addition to the bidding price, it is also necessary to consider various performance indexes of weapons. Similarly, model training for federal learning considers multi-dimensional input variables, such as customer data, calculation and communication, and bid price. The literature [173] elaborated on multi-dimensional motivation and completes the theoretical proof. Both proposed a light-weight server counting and scoring method, achieving better results than traditional FL, including higher accuracy and faster loss convergence. The literature [173] is a representative paper of auction theory in federal learning. It perfectly combines auction rationality in economics with the user's demand in federal learning, expands and applies it, and achieves good experimental results. The mobile edge computing (MEC) problem has also been developed

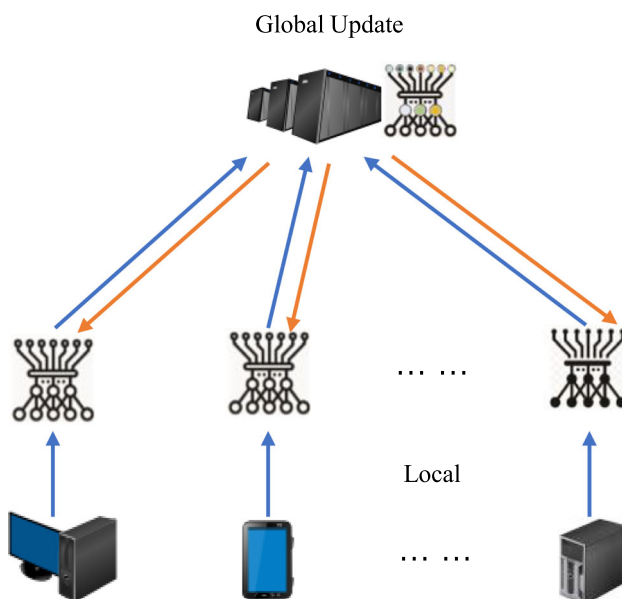


Fig. 18 Structure diagram of federal learning

Table 4 The comparison of representative auction theory of federal learning incentives

	Individual rationality	Incentive compatibility	Pareto efficiency	Budget balance
[173]	✓	✓	✓	×
[29]	✓	✓	×	✓
[67]	✓	✓	×	×
[76]	✓	×	×	×
[32]	✓	×	×	×
[84]	×	×	×	×

Table 5 The comparison of representative auction theory of federal learning incentives

	Individual rationality	Incentive compatibility	Pareto efficiency	Budget balance
[34]	✓	✓	×	×
[80]	✓	✓	×	×
[72]	✓	✓	×	×
[165]	✓	✓	×	×
[143]	✓	✓	×	✓

in the literature [84]. In this method, cluster federated learning is used to offset the influence of non-independent identical distribution (non-IID) and local data size imbalance, and the energy heterogeneity of the system is considered. It is proved that this method can converge to the approximate optimal solution in Table 4.

Table 5 compares the existing literature in proving four typical mechanism design judgment criteria. In the literature [29], Cong proposed a VCG-based incentive mechanism for supply side FL: Fair-VCG (FVCG). Vickrey-Clerke-Groves (VCG), a classical mechanism in auction theory, can incentivize participants to report their training costs truthfully. The model maximizes social surplus and minimizes federal inequity. The literature uses neural networks to solve function optimization problems and, under some reasonable assumptions, provides theoretical proofs for FVCG with Dominant Incentive Compatibility, Social Surplus Maximization, Individual Rationality and Weak Budget, the theoretical proof of Budget Balance. Moreover, this is the first theoretical proof of Budget Balance.

The literature [67] studies federal learning in wireless communication scenarios, and two auction mechanisms are designed to maximize the total utility of the federal learning market, i.e., a formulaic representation of social welfare. In addition, the earth mover's distance (EMD) is used as a metric for metric data. GNN networks are also combined with deep reinforcement learning methods to implement automatic auction mechanisms and are

theoretically shown to achieve strategy-proof, individually rational and truthful. The literature [76] studies federal learning for Cellular Wireless Networks using a stochastic auction framework for incentive design, guaranteeing truthfulness, individual rationality and efficiency.

The literature [32] is notable for adding contribution measurement to the auction theory, the proposed Learning Quality Maximization (LQM) problem and the proposed FAIR system that integrates quality estimation, quality perception and federated learning. The proposed FAIR system integrates quality estimation, quality perception and federated learning aggregation. It is also proved theoretically that the mechanism proposed in the paper is truthful, individually rational and computationally efficient.

Although not utilizing auction theory, the literature [172] also involves the concept of contribution-aware and considers the heterogeneity of the distribution data. It designs a new loss function and proves that it is convex and has optimization preferences. The table shows that Individual Rationality is what most of the literature can achieve, followed by Incentive Compatibility. The economics theorem tells us that to satisfy the above two conditions. The incentive mechanism is Strategy-Proof. Furthermore, Pareto Efficiency as an intuitive measure of efficient resource allocation, few papers can make theoretical proofs, except for the literature [173].

The Budget Balance is rarely mentioned in the current literature as an essential part to be considered in realizing

the federal learning scenarios. Although the mechanism design problem can achieve approximate optimal or locally optimal solutions under certain conditions, it would be more relevant to realistic scenarios if Budget Balance, optimality under conditionally constrained or limited budget situations, is considered. This will also be the focus of research in the future.

6.2.2 Contract theory-based federated learning incentive mechanism

Contract theory is different from auction theory. It needs the distribution information of some participants as the basis for contract-making and is a representative economic concept in cooperative games. The design of contracts enables application scenarios to achieve a better balance of benefits, and its basic theory is described in Sect. 4. This section discusses the current work and future research points in the hot field of federal learning incentives design based on contract theory.

There are similarities in the research scenario between literature [34] and literature [173] in the previous subsection, both considering multi-dimensional information for federal learning problems. Unlike [173], which is based on the auction theory of reverse auctions without knowing any information about each participant; literature [34] is based on the cooperative contract theory. Due to the essential difference between the two underlying theories, the literature [34] makes different decisions concerning the degree of knowledge of the client's private information, classified into four cases (i.e., complete information, weakly incomplete information, strongly incomplete information and non-independent identically distributed data), respectively. The complexity of contract design increases with the incompleteness of information. In addition to the proof of optimal contract design for the solid incomplete information scenario in non-IID data, the literature completes the complete theoretical proof of the other works. It verifies the correctness of the theory with experiments.

The literature [143, 165] considers the role of contract theory from another perspective, which explores how to reduce the information leakage of participants in federated learning. The literature [143] makes the participants' private information private through differentially private encryption techniques and proposes the DPFL framework (DESIGN MULTIDIMENSIONAL CONTRACT) in the context of this information asymmetry. Simulation results verify the effectiveness of this incentive mechanism. The literature [143] proposes a Pain-FL framework based on

contract-theoretic customization of privacy levels, which can be analyzed for optimal contracts under the full-information and non-full-information models, respectively.

Compared with the theoretical improvements, how to apply the existing basic mechanism design theory to the Internet scene reasonably is also an urgent problem to be solved. At present, the design of incentive mechanism for federal learning is still in the initial stage. How to design a better incentive mechanism, meet different security levels and make the mechanism more in line with the actual application needs, etc., all need to be explored constantly.

6.3 Other scarce resources allocation

In addition to optimal mechanism design in economic-oriented environments such as online advertising on the Internet and research in data sharing and privacy protection in complex environments, algorithmic mechanism design has been applied to a broader range of fields and more complex scenarios and analysis in the allocation of complex scarce resources are discussed below.

In allocating resources, we often use utility monetary payments to achieve design goals. Matching is another widely studied field, including the transaction cycle algorithm in housing allocation [125]. The gale-Shapley algorithm in stable marriage is an effective solution [55]. Barter has become an essential aspect of e-commerce and multi-intelligence systems. Among various forms of barter, the exchange of living organs is the primary form. In kidney transplantation [2, 8], the donor cannot charge the patient by monetary payment. Without utility transfer tools, researchers focus on achieving the best or most stable match. In the last decade, organ exchange has become an in-depth research topic for the AI and EC communities, with several researchers starting in-depth studies with kidney exchange scenarios [2, 8, 33, 46]. In the USA, kidney exchange transplantation accounts for about 10% of all living donor kidney transplantation. Recently, this idea has been investigated in several organs, including the kidneys and the liver.

Social choice is a large area of research, and many scenarios can be explored as applications, such as voting, equity policy, tax policy, and so on. It also has the social choice theory worthy of further study. Voting is an essential application in the mechanism design of social choice [11], which does not transfer money. With the development of the Internet, electronic voting mechanisms [92] have emerged, providing us with new research directions. With the development of social networks, more and

more voting is conducted online by inviting participants on social media, which brings many security [164] and fairness problems. Therefore, it is another research direction to study and improve the mechanism design of data-driven voting algorithms.

Big data and the internet bring convenience to people's lives. However, the applications that need to be solved are becoming more and more complex, such as common carpooling, bike sharing, and optimized parking spots that need mechanism design. It can also be delegated to more profound and broader social fields, including auditing, government, military, and many other fields. The purpose of carpooling is to reduce traffic congestion, vehicle trips, vehicle emissions, and other environmental advantages. This concept has been widely accepted all over the world, which not only provides convenient and economical service for drivers and passengers.

We can see that the rational allocation of limited resources is a complex issue covering all aspects of daily life. There are different problems to be solved in different application fields. Designing the best mechanism is a problem we must discuss from a new height and angle.

7 Conclusion

The algorithmic mechanism design method is based on the traditional mechanism design, supplemented by machine learning or other algorithms. Able to cope with complex environmental problems and abstract resource allocation problems, the traditional mechanism design theory and computer-aided mechanism design approaches are compared and combined, and the main problems and challenges in the current mechanism design are proposed in this paper. In the face of these challenges, we discuss three applications representing the economy and complex environment: optimal mechanism design for online internet advertising, incentive mechanism for data sharing, privacy-preserving and other scarce resource allocation. At the same time, we have pointed out possible future research directions in different fields and from which point to start overcoming new challenges.

Funding This research was funded by Basic Research Project of Shenzhen, China (No. JCYJ20180507183624136, JCYJ20200109113405927), National Natural Science Foundation of China (No. 61872109, 61902093, 62176164, 62203134), Science and Technology Project of Guangzhou (No. 2020A1515010652), Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies (No. 2022B1212010005), Key Fields R &D Project of

Guangdong Province (No. 2020B0101380001) and PINGAN-HITsz Intelligence Finance Research Center.

Availability of data and materials Part of the datasets generated during or analyzed during the current study are available in the Web of Science (<https://www.webofscience.com/>) and China Knowledge Centre for Engineering Sciences and Technology (<https://www.ckcest.cn/index>) repository. Part of the datasets that support the findings of this study are available from the first or corresponding author, Qian Chen or Wang Xuan, upon reasonable request.

Declarations

Conflict of interest We have no conflict of interest to declare.

References

1. Abhishek V, Hosanagar K (2013) Optimal bidding in multi-item multislot sponsored search auctions. *Operat Res* 61(4):855–873
2. Abraham DJ, Blum A, Sandholm T (2007) Clearing algorithms for barter exchange markets: enabling nationwide kidney exchanges. In: *Proceedings of the 8th ACM conference on electronic commerce*, pp 295–304
3. Adomavicius G, Gupta A, Yang M (2018) Designing real-time feedback for bidders in homogeneous-item continuous combinatorial auctions. *MIS Quarterly* 43(3):19–89
4. Allen B (1997) Implementation theory with incomplete information. *Cooperation: game-theoretic approaches*, pp 115–126
5. Ausubel LM, Milgrom PR (2006) The lovely but lonely Vickrey auction. *Combinatorial Auctions* 17:22–26
6. Ausubel LM, Cramton P, Deneckere RJ (2002) Bargaining with incomplete information. *Handbook of game theory with economic applications*, vol 3, pp 1897–1945
7. Ausubel LM, Cramton P, Milgrom PR (2016) The clock-proxy auction: a practical combinatorial auction design. *Handbook of spectrum auction design*, vol 1, pp 120–140
8. Awasthi P, Sandholm T (2009) Online stochastic optimization in the large: application to kidney exchange. *IJCAI*, pp 405–411
9. Babaioff M, Sharma Y, Slivkins A (2009) Characterizing truthful multi-armed bandit mechanisms. In: *Proceedings of the 10th ACM conference on electronic commerce*, pp 79–88
10. Bai Y, Jin C, Wang H et al (2021) Sample-efficient learning of Stackelberg equilibria in general-sum games. *Adv Neural Inf Process Syst* 34:25,799–25,811
11. Barberà S (2011) Strategyproof social choice. In: *Handbook of social choice and welfare*, vol 2, pp 731–831
12. Benoit JP, Krishna V (2001) Multiple-object auctions with budget constrained bidders. *Rev Econ Stud* 68(1):155–179
13. Bergemann D, Välimäki J (2019) Dynamic mechanism design: an introduction. *J Econ Lit* 57(2):235–74
14. Bichler M, Milgrom P, Schwarz G (2022) Taming the communication and computation complexity of combinatorial auctions: the fuel bid language. *Manag Sci* 69(4):2217–2238
15. Bolton P, Dewatripont M (2004) *Contract theory*. MIT press
16. Bowling M, Burch N, Johanson M et al (2015) Heads-up limit hold'em poker is solved. *Science* 347(6218):145–149
17. Brero G, Eden A, Gerstgrasser M et al (2021) Reinforcement learning of sequential price mechanisms. *AAAI Press*, pp 5219–5227

18. Brown N, Sandholm T (2018) Superhuman ai for heads-up no-limit poker: Libratus beats top professionals. *Science* 359(6374):418–424
19. Brown N, Sandholm T (2019) Superhuman ai for multiplayer poker. *Science* 365(6456):885–890
20. Bu TM, Deng X, Qi Q (2012) Multi-bidding strategy in sponsored search auctions. *J Comb Optim* 23(3):356–372
21. Cai H, Ren K, Zhang W, et al (2017) Real-time bidding by reinforcement learning in display advertising. In: Proceedings of the tenth ACM international conference on web search and data mining, pp 661–670
22. Caragiannis I, Kaklamanis C, Kanellopoulos P et al (2011) On the efficiency of equilibria in generalized second price auctions. In: Proceedings of the 12th ACM conference on Electronic commerce, pp 81–90
23. Celli A, Marchesi A, Farina G et al (2020) No-regret learning dynamics for extensive-form correlated equilibrium. *Adv Neural Inf Process Syst* 33:7722–7732
24. Che YK (1993) Design competition through multidimensional auctions. *The RAND J Econ*. pp 668–680
25. Che YK, Gale I (1998) Standard auctions with financially constrained bidders. *Rev Econ Stud* 65(1):1–21
26. Chen B, Kankanhalli MS (2019) Pricing average price advertising options when underlying spot market prices are discontinuous. *IEEE Trans Knowl Data Eng* 31(9):1765–1778
27. Chen X, Hu G, Lu P et al (2011) On the approximation ratio of k-lookahead auction. In: International workshop on internet and network economics, Springer, pp 61–71
28. Coase RH (1937) The nature of the firm. *Economica* 4(16):386–405
29. Cong M, Yu H, Weng X et al (2020) A VCG-based fair incentive mechanism for federated learning. *arXiv preprint arXiv:2008.06680*
30. Conneau A, Lample G (2019) Cross-lingual language model pretraining. In: Advances in neural information processing systems, pp 7057–7067
31. Curry MJ, Lyi U, Goldstein T et al (2022) Learning revenue-maximizing auctions with differentiable matching. In: International conference on artificial intelligence and statistics, PMLR, pp 6062–6073
32. Deng Y, Lyu F, Ren J et al (2021) Fair: quality-aware federated learning with precise user incentive and model aggregation. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10
33. Dickerson J, Sandholm T (2014) Multi-organ exchange: the whole is greater than the sum of its parts. In: Proceedings of the AAAI conference on artificial intelligence, pp 1412–1418
34. Ding N, Fang Z, Huang J (2020) Optimal contract design for efficient federated learning with multi-dimensional private information. *IEEE J Sel Areas Commun* 39(1):186–200
35. Ding W, Qin T, Zhang X et al (2013) Multi-armed bandit with budget constraint and variable costs. In: AAAI, pp 232–238
36. Dobzinski S, Lavi R, Nisan N (2012) Multi-unit auctions with budget limits. *Games Econ Behav* 74(2):486–503
37. Du J, Jiang C, Han Z et al (2017) Contract mechanism and performance analysis for data transaction in mobile social networks. *IEEE Trans Netw Sci Eng* 6(2):103–115
38. Du M, Cowen-Rivers AI, Wen Y et al (2019) Infer your enemies and know yourself, learning in real-time bidding with partially observable opponents. *arXiv preprint arXiv:1902.11212*
39. Duan Z, Tang J, Yin Y et al (2022) A context-integrated transformer-based neural network for auction design. *ICML*, pp 5609–5626
40. Dütting P, Fischer F, Jirapinyo P et al (2015a) Payment rules through discriminant-based classifiers. *ACM Trans Econ Comput* 3(1):1–41
41. Dütting P, Henzinger M, Starnberger M (2015) Auctions for heterogeneous items and budget limits. *ACM Trans Econ Comput (TEAC)* 4(1):1–17
42. Dütting P, Feng Z, Narasimhan H, et al (2019) Optimal auctions through deep learning. In: International conference on machine learning. PMLR, pp 1706–1715
43. Edelman B, Schwarz M (2010) Optimal auction design and equilibrium selection in sponsored search auctions. *Am Econ Rev* 100(2):597–602
44. Edelman B, Ostrovsky M, Schwarz M (2007) Internet advertising and the generalized second-price auction: Selling billions of dollars worth of keywords. *Am Econ Rev* 97(1):242–259
45. Elkind E, Lipmaa H (2004) Interleaving cryptography and mechanism design: the case of online auctions. In: *Financial Cryptography*, pp 117–131
46. Ergin H, Sönmez T, Ünver MU (2015) Lung exchange. Department of Economics, Boston College
47. Fard MV, Sahafi A, Rahmani AM et al (2020) Resource allocation mechanisms in cloud computing: a systematic literature review. *IET Softw* 14(6):638–653
48. Farina G, Bianchi T, Sandholm T (2020) Coarse correlation in extensive-form games. In: Proceedings of the AAAI Conference on Artificial Intelligence, pp 1934–1941
49. Fazlyab M, Morari M, Pappas GJ (2020) Safety verification and robustness analysis of neural networks via quadratic constraints and semidefinite programming. *IEEE Trans Autom Control* 67:1–15
50. Feldman M, Lai K, Zhang L (2005) A price-anticipating resource allocation mechanism for distributed shared clusters. In: Proceedings of the 6th ACM conference on Electronic commerce, pp 127–136
51. Feng S, Niyato D, Wang P et al (2019) Joint service pricing and cooperative relay communication for federated learning. In: 2019 international conference on internet of things (iThings) and IEEE green computing and communications (GreenCom) and IEEE cyber. Physical and Social Computing (CPSCom) and IEEE Smart Data (SmartData). IEEE, pp 815–820
52. Feng Z, Narasimhan H, Parkes DC (2018) Deep learning for revenue-optimal auctions with budgets. In: Proceedings of the 17th international conference on autonomous agents and multiagent systems, pp 354–362
53. Fernholz D, Karatzas I (2010) On optimal arbitrage. *Ann Appl Probab* 20(4):1179–1204
54. Foundation NS (2015) New directions in computational social choice and mechanism design. https://www.nsf.gov/awardsearch/showAward?AWD_ID=1527434
55. Gale D, Shapley LS (1962) College admissions and the stability of marriage. *Am Math Mon* 69(1):9–15
56. Gao G, Huang H, Xiao M et al (2021) Auction-based combinatorial multi-armed bandit mechanisms with strategic arms. In: IEEE INFOCOM 2021-IEEE conference on computer communications. IEEE, pp 1–10
57. Gasparini M, Nuara A, Trovò F et al (2018) Targeting optimization for internet advertising by learning from logged bandit feedback. *IEEE*, pp 1–8
58. Gershkov A, Moldovanu B (2014) *Dynamic allocation and pricing: a mechanism design approach*. MIT Press
59. Gomes R, Sweeney K (2014) Bayes-nash equilibria of the generalized second-price auction. *Games Econ Behav* 86:421–437

60. Greenstein-Messica A, Rokach L (2020) Machine learning and operation research based method for promotion optimization of products with no price elasticity history. *Electron Comm Res Appl* 40(100):914
61. Haghpanah N, Hartline J (2015) Reverse mechanism design. In: *Proceedings of the sixteenth ACM conference on economics and computation*, pp 757–758
62. Haghpanah N, Hartline J (2021) When is pure bundling optimal? *Rev Econ Stud* 88(3):1127–1156
63. Hausknecht M, Stone P (2015) Deep recurrent q-learning for partially observable mdps. In: *2015 AAAI fall symposium series*
64. Hosu I, Rebedea T (2016) Playing atari games with deep reinforcement learning and human checkpoint replay. *CoRR abs/1607.05077*. <https://arxiv.org/abs/1607.05077>
65. Hurwicz L (1960) Optimality and informational efficiency in resource allocation processes. In: Arrow KJ, Karlin S, Suppes P (eds) *Mathematical methods in the social sciences*. Stanford University Press, pp 27–46
66. Jia J, Zhang Q, Zhang Q, et al (2009) Revenue generation for truthful spectrum auction in dynamic spectrum access. In: *Proceedings of the tenth ACM international symposium on mobile ad hoc networking and computing*, pp 3–12
67. Jiao Y, Wang P, Niyato D et al (2020) Toward an automated auction framework for wireless federated learning services market. *IEEE Trans Mob Comput* 20(10):3034–3048
68. Jin J, Song C, Li H et al (2018) Real-time bidding with multi-agent reinforcement learning in display advertising. In: *The 27th ACM international conference on information and knowledge management*, pp 2193–2201
69. Kalai E (1977) Nonsymmetric Nash solutions and replications of 2-person bargaining. *Int J Game Theory* 6(3):129–133
70. Kalai E, Smorodinsky M (1975a) On a game theoretic notion of complexity for compact convex sets. *Proc Am Math Soc* 49(2):416–420
71. Kalai E, Smorodinsky M (1975b) Other solutions to Nash's bargaining problem. *Econom J Econ Soc* 71(3):513–518
72. Kang J, Xiong Z, Niyato D et al (2019) Incentive mechanism for reliable federated learning: a joint optimization approach to combining reputation and contract theory. *IEEE Internet Things J* 6(6):10700–10714
73. Koster R, Balaguer J, Tacchetti A et al (2022) Human-centered mechanism design with democratic AI. *Nat Hum Behav* 6(9):1320–1320
74. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. pp 1106–1114
75. Lahaie S, Pennock DM, Saberi A et al (2007) Sponsored search auctions. In: *Algorithmic game theory*, vol 1, pp 699–716
76. Le THT, Tran NH, Tun YK et al (2020) Auction based incentive design for efficient federated learning in cellular wireless networks. In: *2020 IEEE wireless communications and networking conference (WCNC)*. IEEE, pp 1–6
77. Leme RP, Tardos E (2010) Pure and bayes-Nash price of anarchy for generalized second price auction. In: *2010 IEEE 51st annual symposium on foundations of computer science*. IEEE, pp 735–744
78. Li H, Wang X, Jia F et al (2021) A survey of Nash equilibrium strategy solving based on CFR. *Arch Comput Methods Eng* 28(4):2749–2760
79. Li J, Zhao L (2020) Game analysis of hydropower bidding under incomplete information. *J Coast Res* 104(SI):398–401
80. Lim WYB, Xiong Z, Miao C et al (2020) Hierarchical incentive mechanism design for federated machine learning in mobile networks. *IEEE Internet Things J* 7(10):9575–9588
81. Ling CK, Brown N (2021) Safe search for Stackelberg equilibria in extensive-form games. In: *Proceedings of the AAAI conference on artificial intelligence*, pp 5541–5548
82. Liu Y, Xu G, Tan X (2009) A novel spectrum allocation mechanism based on graph coloring and bidding theory. In: *2009 International conference on computational intelligence and natural computing*. IEEE, pp 155–158
83. Liu Y, Ai Z, Sun S et al (2020) Fedcoin: A peer-to-peer payment system for federated learning. In: *Federated learning: privacy and incentive*, pp 125–138
84. Lu R, Zhang W, Li Q et al (2021) Auction based clustered federated learning in mobile edge computing system. *arXiv preprint arXiv:2103.07150*
85. Lucier B, Leme RP (2011) GSP auctions with correlated types. In: *Proceedings of the 12th ACM conference on electronic commerce*, pp 71–80
86. Lucier B, Paes Leme R, Tardos É (2012) On revenue in the generalized second price auction. In: *Proceedings of the 21st international conference on World Wide Web*, pp 361–370
87. Mafita (2019) Implementation of auctions through the internet for state-owned assets of the state assets service office and Serang auctions based on the principle of legal certainty. *Nurani Hukum* 2:25
88. Manelli AM, Vincent DR (2006) Bundling as an optimal selling mechanism for a multiple-good monopolist. *J Econ Theory* 127(1):1–35
89. Martín-Herrán G, Rubio SJ (2021) On coincidence of feedback and global Stackelberg equilibria in a class of differential games. *Eur J Oper Res* 293(2):761–772
90. Maskin E (1999) Nash equilibrium and welfare optimality. *Rev Econ Stud* 66(1):22–38
91. Mas-Colell A, Whinston M, Green J (1995) *microeconomic theory*. Oxford student edn. Oxford University Press. <https://books.google.com/books?id=sQGDQgAACAAJ>
92. McCarthy A, Smyth B, Quaglia EA (2014) Hawk and Aucitas: e-auction schemes from the Helios and Civitas e-voting schemes. In: *International conference on financial cryptography and data security*. Springer, pp 51–63
93. Milani S, Chatzigiannakis I (2021) Design, analysis, and experimental evaluation of a new secure rejoin mechanism for Lorawan using elliptic-curve cryptography. *J Sens Actuator Netw* 10(2):36
94. Milgrom P (2017) *Discovering prices: auction design in markets with complex constraints*. Columbia University Press
95. Milgrom P (2021) Auction research evolving: theorems and market designs. *Am Econ Rev* 111(5):1383–1405
96. Milgrom P, Segal I (2020) Clock auctions and radio spectrum reallocation. *J Polit Econ* 128(1):1–31
97. Miralles-Pechuán L, Qureshi MA, Namee BM (2021) Real-time bidding campaigns optimization using user profile settings. *Electron Commer Res* 23:1297–1322
98. Mirrlees JA (1971) An exploration in the theory of optimum income taxation. *Rev Econ Stud* 38(2):175–208
99. Mkiramweni ME, Yang C (2018) Energy efficiency optimization for wireless unmanned aerial vehicle communication networks: a bargaining game approach. In: *2018 10th international conference on wireless communications and signal processing (WCSP)*. IEEE, pp 1–6

100. Mnih V, Kavukcuoglu K, Silver D et al (2015) Human-level control through deep reinforcement learning. *Nature* 518(7540):529–533
101. Mnih V, Badia AP, Mirza M et al (2016) Asynchronous methods for deep reinforcement learning. In: *International conference on machine learning*, PMLR, pp 1928–1937
102. Mohri M, Munoz A (2014) Optimal regret minimization in posted-price auctions with strategic buyers. In: *Advances in neural information processing systems*, NIPS, pp 1871–1879
103. Moldovanu B (1990) Competition and bargaining in games and markets. PhD thesis, University of Bonn, Germany. <https://d-nb.info/911190341>
104. Moravčík M, Schmid M, Burch N et al (2017) Deepstack: expert-level artificial intelligence in heads-up no-limit poker. *Science* 356(6337):508–513
105. Myerson RB (1979) Incentive compatibility and the bargaining problem. *Econometr J Econ Soc* 47(1):61–73
106. Nachum O, Gu SS, Lee H et al (2018) Data-efficient hierarchical reinforcement learning. In: *Advances in neural information processing systems*, NIPS, pp 3307–3317
107. Nash J (1951) Non-cooperative games. *Ann Math* 54:286–295
108. Nash J (1953) Two-person cooperative games. *Econom J Econom Soc* 21:128–140
109. Nash Jr JF (1950) The bargaining problem. *Econom J Econom Soc* 18:155–162
110. Newman N, Leyton-Brown K, Milgrom P et al (2020) Incentive auction design alternatives: a simulation study. In: *Proceedings of the 21st ACM conference on economics and computation*, pp 603–604
111. Nisan N, Roughgarden T, Tardos E et al (2007) *Algorithmic game theory*. Cambridge University Press. <https://books.google.com/books?id=f4SFngEACAAJ>
112. Nuara A, Trovo F, Gatti N et al (2018) A combinatorial-bandit algorithm for the online joint bid/budget optimization of pay-per-click advertising campaigns. In: *Thirty-second AAAI conference on artificial intelligence*
113. Nurcahyani I, Lee JW (2021) Role of machine learning in resource allocation strategy over vehicular networks: a survey. *Sensors* 21(19):6542
114. Ortega-Reichert A (1968) Models for competitive bidding under uncertainty. Stanford University
115. Parkes DC, Singh S (2003) An MDP-based approach to online mechanism design. In: *Thrun S, Saul L, Schölkopf B (eds) Advances in neural information processing systems*, vol 16. MIT Press
116. Paul M (2000) Putting auction theory to work: the simultaneous ascending auction. *J Polit Econ* 108(2):245–272
117. Pavlov G (2011) Optimal mechanism for selling two goods. *BE J Theor Econ* 11(1):1–35
118. Pavlov G (2020) Selling two units of a customizable good. Tech rep, Working paper, University of Western Ontario
119. Peleg B, Sudhölter P (2007) *Introduction to the theory of cooperative games*, vol 34. Springer
120. Peri N, Curry M, Dooley S et al (2021) Preferencenet: encoding human preferences in auction design with deep learning. *Adv Neural Inf Process Syst* 34:17532–17542
121. Pitchik C (2009) Budget-constrained sequential auctions with incomplete information. *Games Econ Behav* 66(2):928–949
122. Qureshi MB, Dehnavi MM, Min-Allah N et al (2014) Survey on grid resource allocation mechanisms. *J Grid Comput* 12(2):399–441
123. Rahme J, Jelassi S, Bruna J et al (2021) A permutation-equivariant neural network architecture for auction design. In: *Proceedings of the AAAI conference on artificial intelligence*, pp 5664–5672
124. Rezaee MJ, Eshkevari M, Saberi M et al (2021) GBK-means clustering algorithm: an improvement to the k-means algorithm based on the bargaining game. *Knowl-Based Syst* 213(106):672
125. Rodrigues TK, Suto K, Nishiyama H et al (2019) Machine learning meets computation and communication control in evolving edge and cloud: challenges and future perspective. *IEEE Commun Surv Tutor* 22(1):38–67
126. Rohde D, Bonner S, Dunlop T et al (2018) Recogym: a reinforcement learning environment for the problem of product recommendation in online advertising. *CoRR abs/1808.00720*. <https://arxiv.org/abs/1808.00720>
127. Ronen A (2001) On approximating optimal auctions. In: *Proceedings of the 3rd ACM conference on electronic commerce*, pp 11–17
128. Roth AE (1985) The college admissions problem is not equivalent to the marriage problem. *J Econ Theor* 36(2):277–288
129. Salomatin K, Liu TY, Yang Y (2012) A unified optimization framework for auction and guaranteed delivery in online advertising. In: *Proceedings of the 21st ACM international conference on information and knowledge management*, pp 2005–2009
130. Sandholm T (2015) Solving imperfect-information games. *Science* 347(6218):122–123
131. Schelling T (1960) *The strategy of conflict*. Mass. <https://books.google.com/books?id=BeG2swEACAAJ>
132. Schneider S, Shabalin P, Bichler M (2010) On the robustness of non-linear personalized price combinatorial auctions. *Eur J Oper Res* 206(1):248–259
133. Schwartz EM, Bradlow ET, Fader PS (2017) Customer acquisition via display advertising using multi-armed bandit experiments. *Mark Sci* 36(4):500–522
134. Shapley LS (1953) Stochastic games. *Proc Natl Acad Sci* 39(10):1095–1100
135. Sheikholeslami F, Jafari Navimipour N (2018) Auction-based resource allocation mechanisms in the cloud environments: a review of the literature and reflection on future challenges. *Concurr Comput Pract Exp* 30(16):e4456
136. Shen W, Peng B, Liu H et al (2020) Reinforcement mechanism design: with applications to dynamic pricing in sponsored search auctions. In: *Proceedings of the AAAI conference on artificial intelligence*, pp 2236–2243
137. Shen W, Tang P, Wang X et al (2021) Coupon design in advertising systems. AAAI Press, pp 5717–5725
138. Shi E, Chung H, Wu K (2022) What can cryptography do for decentralized mechanism design. *arXiv preprint arXiv:2209.14462*
139. Silver D, Huang A, Maddison CJ et al (2016) Mastering the game of go with deep neural networks and tree search. *Nature* 529(7587):484–489
140. Silver D, Schrittwieser J, Simonyan K et al (2017) Mastering the game of go without human knowledge. *Nature* 550(7676):354–359
141. Singer Y (2010) Budget feasible mechanisms. In: *2010 IEEE 51st annual symposium on foundations of computer science*. IEEE, pp 765–774
142. Song T, Tong Y, Wei S (2019) Profit allocation for federated learning. In: *2019 IEEE international conference on big data (big data)*. IEEE, pp 2577–2586
143. Sun P, Che H, Wang Z et al (2021) Pain-FL: personalized privacy-preserving incentive for federated learning. *IEEE J Sel Areas Commun* 39(12):3805–3820



144. Sun Z, Liu Y, Wang J et al (2021b) Applications of game theory in vehicular networks: a survey. *IEEE Commun Surv Tutor*
145. Sutton RS, McAllester DA, Singh SP et al (1999) Policy gradient methods for reinforcement learning with function approximation. The MIT Press, pp 1057–1063
146. Tang L, Rosales R, Singh A et al (2013) Automatic ad format selection via contextual bandits. In: *Proceedings of the 22nd ACM international conference on Information & Knowledge Management*, pp 1587–1594
147. Tang P (2017) Reinforcement mechanism design. *IJCAI*, pp 5146–5150
148. Thirumulanathan D, Sundaresan R, Narahari Y (2017) On optimal mechanisms in the two-item single-buyer unit-demand setting. *CoRR abs/1705.01821*. <https://arxiv.org/abs/1705.01821>
149. Tossou AC, Dimitrakakis C (2016) Algorithms for differentially private multi-armed bandits. In: *Thirtieth AAAI conference on artificial intelligence*
150. Valenčík R, Čermík O et al (2021) The inequality in society and a multipoint extension of Nash bargaining problem. *Econ Bus J* 15(1):221–232
151. van Hasselt H, Guez A, Silver D (2016) Deep reinforcement learning with double q-learning. *AAAI Press*, pp 2094–2100
152. Varian HR (2007) Position auctions. *Int J Ind Organ* 25(6):1163–1178
153. Varian HR (2009) Online ad auctions. *Am Econ Rev* 99(2):430–34
154. Vickrey W (1961) Counterspeculation, auctions, and competitive sealed tenders. *J Financ* 16(1):8–37
155. Vickrey W (1962) Auctions and bidding games. In: *Recent advances in game theory*, Princeton University conference, pp 15–29
156. Von Neumann J, Morgenstern O (2020) *Theory of games and economic behavior*. Sidney Bond. <https://books.google.com/books?id=nQ1NzQEACAAJ>
157. Wang G, Dang CX, Zhou Z (2019) Measure contribution of participants in federated learning. In: *2019 IEEE international conference on big data (big data)*. IEEE, pp 2597–2604
158. Wang M, Song Y, Sui B et al (2022) Comparative study of pricing mechanisms and settlement methods in electricity spot energy market based on multi-agent simulation. *Energy Rep* 8:1172–1182
159. Wang Y, Yin D, Jie L et al (2018) Optimizing whole-page presentation for web search. *ACM Trans Web* 12(3):1:91–19:25
160. Wilson RB (1969) Communications to the editor—competitive bidding with disparate information. *Manag Sci* 15(7):446–452
161. Wu D, Chen X, Yang X et al (2018) Budget constrained bidding by model-free reinforcement learning in display advertising. In: *Proceedings on 27th ACM international conference on information and knowledge management*, pp 1443–1451
162. Wu D, Chen C, Yang X, et al (2018a) A multi-agent reinforcement learning method for impression allocation in online display advertising. *CoRR abs/1809.03152*. <https://arxiv.org/abs/1809.03152>
163. Wu D, Chen X, Yang X et al (2018b) Budget constrained bidding by model-free reinforcement learning in display advertising. pp 1443–1451
164. Wu HT, Yang CY (2018) A blockchain-based network security mechanism for voting systems. In: *2018 1st international cognitive cities conference (IC3)*. IEEE, pp 227–230
165. Wu M, Ye D, Ding J et al (2021) Incentivizing differentially private federated learning: a multidimensional contract approach. *IEEE Internet Things J* 8(13):10,639–10,651
166. Xu M, Qin T, Liu T (2013) Estimation bias in multi-armed bandit algorithms for search advertising. *Adv Neural Inf Process Syst* 26:2400–2408
167. Yang C, Feng Y, Whinston A (2022) Dynamic pricing and information disclosure for fresh produce: an artificial intelligence approach. *Prod Oper Manag* 31(1):155–171
168. Yang H, Lu Q (2016) Dynamic contextual multi arm bandits in display advertisement. *IEEE Computer Society*, pp 1305–1310
169. Yao ACC (2017) Dominant-strategy versus Bayesian multi-item auctions: maximum revenue determination and comparison. In: *Proceedings of the 2017 ACM conference on economics and computation*, pp 3–20
170. Ye D, Zhang M, Sutanto D (2013) Self-adaptation-based dynamic coalition formation in a distributed agent network: a mechanism and a brief survey. *IEEE Trans Parallel Distrib Syst* 24(5):1042–1051
171. Yuan S, Wang J, van der Meer M (2013) Adaptive keywords extraction with contextual bandits for advertising on parked domains. *CoRR abs/1307.3573*. <https://arxiv.org/abs/1307.3573>
172. Zeng H, Zhou T, Guo Y et al (2021) Fedcav: contribution-aware model aggregation on distributed heterogeneous data in federated learning. In: *50th international conference on parallel processing*, pp 1–10
173. Zeng R, Zhang S, Wang J et al (2020) Fmore: an incentive scheme of multi-dimensional auction for federated learning in MEC. In: *2020 IEEE 40th international conference on distributed computing systems (ICDCS)*. IEEE, pp 278–288
174. Zhan Y, Li P, Qu Z et al (2020) A learning-based incentive mechanism for federated learning. *IEEE Internet Things J* 7(7):6360–6368
175. Zhan Y, Zhang J, Hong Z et al (2021) A survey of incentive mechanism design for federated learning. *IEEE Trans Emerg Top Comput* 10(2):1035–1044
176. Zhang BH, Farina G, Celli A et al (2022a) Optimal correlated equilibria in general-sum extensive-form games: fixed-parameter algorithms, hardness, and two-sided column-generation. In: *Proceedings of the 23rd ACM conference on economics and computation*, pp 1119–1120
177. Zhang BH, Farina G, Sandholm T (2022b) Team belief DAG form: a concise representation for team-correlated game-theoretic decision making. *arXiv preprint arXiv:2202.00789*
178. Zhang T, Wang H, He J (2016) An incentive mechanism under hidden-action for device-to-device content sharing. In: *2016 IEEE 13th international conference on signal processing (ICSP)*. IEEE, pp 1288–1292
179. Zhang Y, Song L, Saad W et al (2015) Contract-based incentive mechanisms for device-to-device communications in cellular networks. *IEEE J Sel Areas Commun* 33(10):2144–2155
180. Zhang Z, Liu X, Zheng Z et al (2021) Optimizing multiple performance metrics with deep GSP auctions for e-commerce advertising. *ACM*, pp 993–1001
181. Zhao X, Gu C, Zhang H et al (2021) DEAR: deep reinforcement learning for online advertising impression in recommender systems. *AAAI Press*, pp 750–758
182. Zheng L, Fiez T, Alumbaugh Z et al (2022a) Stackelberg actor-critic: game-theoretic reinforcement learning algorithms. In: *Proceedings of the AAAI conference on artificial intelligence*, pp 9217–9224
183. Zheng S, Trott A, Srinivasa S et al (2022) The AI economist: optimal economic policy design via two-level deep reinforcement learning. *Sci Adv* 8(18):eabk2607

184. Zou X (2009) Double-sided auction mechanism design in electricity based on maximizing social welfare. *Energy Policy* 37(11):4231–4239

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Qian Chen¹  · Xuan Wang^{1,2,3}  · Zoe Lin Jiang^{1,2,3} · Yulin Wu^{1,3} · Huale Li¹ · Lei Cui^{1,2} · Xiaozhen Sun¹

✉ Xuan Wang
wangxuan@cs.hitsz.edu.cn

Qian Chen
qianchen@stu.hit.edu.cn

Zoe Lin Jiang
zoeljiang@hit.edu.cn

Yulin Wu
yulinwu@cs.hitsz.edu.cn

Huale Li
lihuale@stu.hit.edu.cn

Lei Cui
cuilei@cs.hitsz.edu.cn

Xiaozhen Sun
sxzhit2018@163.com

¹ Computer Science and Technology, Harbin Institute of Technology Shenzhen, Shenzhen 518000, Guangdong, China

² Peng Cheng Laboratory, Shenzhen 518000, Guangdong, China

³ Guangdong Provincial Key Laboratory of Novel Security Intelligence Technologies, Shenzhen 518000, Guangdong, China