An Attack-Defense Game-Based Reinforcement Learning Privacy-Preserving Method Against Inference Attack in Double Auction Market

Donghe Li, *Member, IEEE*, Chunlin Hu, Qingyu Yang, *Senior Member, IEEE*, Yuhao Ma, Feiye Zhang, and Dou An

***Abstract*— Auction mechanism, as a fair and efficient resource allocation method, has been widely used in varieties trading scenarios, such as advertising, crowdsensoring and spectrum. However, in addition to obtaining higher profits and satisfaction, the privacy concerns have attracted researchers' attention. In this paper, we mainly study the privacy preserving issue in the double auction market against the indirect inference attack. Most of the existing works apply differential privacy theory to defend against the inference attack, but there exists two problems. First, 'indistinguishability' of differential privacy (DP) cannot prevent the disclosure of continuous valuations in the auction market. Second, the privacy-utility trade-off (PUT) in differential privacy deployment has not been resolved. To this end, we proposed an attack-defense game-based reinforcement learning privacy-preserving method to provide practically privacy protection in double auction. First, the auctioneer acts as defender, adds noise to the bidders' valuations, and then acts as adversary to launch inference attack. After that the auctioneer uses the attack results and auction results as a reference to guide the next deployment. The above process can be regarded as a Markov Decision Process (MDP). The state is the valuations of each bidders under the current steps. The action is the noise added to each bidders. The reward is composed of privacy, utility and training speed, in which attack success rate and social welfare are taken as measures of privacy and utility, a delay penalty term is used to reduce the training time. Utilizing the deep deterministic policy gradient (DDPG) algorithm, we establish an actor-critic network to solve the problem of MDP. Finally, we conducted extensive evaluations to verify the performance of our proposed method. The results show that compared with other existing DP-based double auction privacy preserving mechanisms, our method can achieve better results in both privacy and utility. We can reduce the attack success rate from nearly 100% to less than 20%, and the utility deviation is less than 5%.**

***Note to Practitioners*—Privacy protection in trading markets, such as advertising, crowdsensing, and spectrum, is crucial. Traditional approaches like differential privacy have been unable to entirely guard sensitive data against inference attacks. To address this, we introduce a novel privacy-preserving mechanism for double auction markets. Our approach employs an attack-defense game model, where noise is added to bidders' valuations and then used to launch an inference attack. This process allows for the evaluation of the noise's effectiveness and iteratively refines the privacy protection method. Transformed into a reinforcement learning model and optimized through a DDPG network, our mechanism reduces computational complexity. It has been shown to significantly diminish the success rate of inference attacks, while maintaining a minimal utility deviation. Practitioners in auction-based markets can leverage our approach to enhance privacy protection without negatively impacting market performance. By integrating our mechanism into their operations, auctioneers can foster a safer and more efficient trading environment.**

***Index Terms*— Double auction, inference attack, differential privacy, privacy-utility trade-off, reinforcement learning.**

I. Introduction

With the rapid development of sensing and information technology, big data from power systems, transportation systems, financial systems, etc., has significantly enhanced the performance of prediction, control, decision-making, and other functions [1], [2]. However, the increasing granularity of data not only improves system performance but also poses a serious threat to safety and privacy [3]. In the realm of network security, scholars have employed differential equations and deep learning technologies to detect network attacks. Song et al. proposed a switching event-triggered state estimation method based on reaction-diffusion neural networks to counter DoS attacks [4]. Similarly, Zhang et al. discussed a hybrid-driven fuzzy secure filtering approach for network attacks in [5], which, along with other network security technologies, forms a robust foundation for privacy protection in the face of cyber threats. Concurrently, privacy protection technologies have garnered substantial research attention, leading to the proposal of various effective methods to tackle different types of privacy threats, such as encryption [6], differential privacy

Received 23 January 2024; revised 21 July 2024 and 29 September 2024; accepted 3 November 2024. This article was recommended for publication by Associate Editor H.-J. Kim and Editor J. Li upon evaluation of the reviewers' comments. This work was supported in part by the National Natural Science Foundation of China under Grant 62203350 and Grant 62173268, in part by the Key Program of the National Natural Science Foundation of China under Grant 61833015, in part by the Key Research and Development Program of Shaanxi under Grant 2022GY-033, and in part by the Industrial Field Project—Key Industrial Innovation Chain (Group) of Shaanxi Province under Grant 2022ZDLGY06-02. *(Corresponding author: Qingyu Yang.)*

Donghe Li and Dou An are with the School of Automation Science and Engineering and the MOE Key Laboratory for Intelligent Networks and Network Security, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: lidonghe2020@xjtu.edu.cn; douan2017@xjtu.edu.cn).

Chunlin Hu, Yuhao Ma, and Feiye Zhang are with the School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: hucl0918@stu.xjtu.edu.cn; mayuhao@stu.xjtu.edu.cn; zhangfy19970324@stu.xjtu.edu.cn).

Qingyu Yang is with the SKLMSE Laboratory and the School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China (e-mail: yangqingyu@mail.xjtu.edu.cn).

Color versions of one or more figures in this article are available at https://doi.org/10.1109/TASE.2024.3496869.

Digital Object Identifier 10.1109/TASE.2024.3496869

[7], and anonymity [8]. These advancements underscore the importance of safeguarding privacy while leveraging the full potential of data-driven technologies.

Differential privacy stands as a crucial tool in the realm of privacy protection technologies. By incorporating a measure of noise into the data or process, it ensures a theoretical 'indistinguishability' that safeguards privacy. To be specific, differential privacy was first proposed by Dwork in 2008 [9], it was originally proposed to deal with the problem of privacy leakage in data query. It ensures that the public output results will not change significantly due to the change of a certain input by adding random noise, that is, the adversary cannot infer which private inputs (neighboring databases) produced the result. Due to the advantages of lightweight calculation, theoretically provable, simple deployment, and strong adaptability, differential privacy has been widely used in various scenarios to provide theoretical privacy protection, such as data querying, data publishing, resource dispatching, and so on.

Resource allocation/trading markets have been promoted recently, such as energy auction market [10], spectrum market [11], and crowdsensing market [12]. Similarly, the privacy protection of participants in the trading market is also the focus of current scholars, and the scholars are devoted to the study of the privacy preserving auction mechanisms [13], [14]. Nonetheless, there are still several problems in the existing researches on the application of differential privacy in the auction market [13], [15], [16], [17]. First of all, differential privacy is designed to resist inference attacks, and most of the current DP-based auction mechanisms do not consider the attack mechanism when they are designed. For example, in [18], the author designed a privacy protection bidding strategy to solve the problem of bidding privacy leakage in crowdsourcing auction market and proves it in detail. But it lacks the detailed mathematical modeling of the attack method, or even the introduction of the attack model. This makes it impossible to reflect the necessity of adding differential privacy in the auction market, let alone evaluate the effectiveness of the proposed mechanism (except for a series of theorem proofs). With the deepening of research, scholars have gradually become aware of this issue, and some of the latest research authors have also begun to explore inference attacks against the auction market. In [19], the author provided examples of privacy inference attacks in the combinatorial auction market, highlighting the necessity of differential privacy protection. However, it still does not model privacy attacks, so in the performance evaluation section, there is still a lack of persuasiveness about the effectiveness of privacy protection algorithms. Secondly, most of the studies on DP-based auction mechanism have only discussed its privacy performance theoretically (i.e., ε-differential privacy), but there are few researches have discussed the privacy performance practical. Thirdly, most studies have not given the method of selecting differential privacy parameter ε explicitly, which brings great difficulties to the practice of differential privacy in auction market. To sum up, the current research on DP-based auction mechanism is more on the theoretical level. From the perspective of practice, the necessity, usability and deployment methods of differential privacy in auction mechanism need to be further studied.

According to our previous work published in 2022 IEEE T-ASE [20], we found that the current DP mechanism is unable to get a good effect in resisting the inference attacks in practice. Meanwhile, the noise is the most effective method against such attack. Therefore, we think that we should jump out of the constraint of differential privacy concept and find a directed noise deployment mechanism, so as to protect the privacy in practice, while minimizing the impact on auction performance as much as possible. However, considering that the noise will affect the performance of the auction market, how to locate the position and size of the noise added has become a problem that needs to be solved. Actually, this can be always considered as a Privacy-Utility Trade-off (PUT) noise optimal deployment problem [21]. Because this optimization problem is continuous in the time series, and has a very large computational complexity, the traditional solution methods can not solve it. At present, the commonly used method is to use the reinforcement learning method to learn the optimal strategy by constantly interacting with the environment through agent [22], [23]. This method has been applied to time series-based data sharing [21], data aggregation [24], data publishing [25], and even smart grid load hiding [26]. Nonetheless, there still remains several challenges in designing such a reinforcement learning-based noise deployment method in double auction market. First of all, the existing reinforcement learning-based PUT works focus on the time series data, so it is naturally a Markov Decision Process (MDP). Although the auction market is also a continuous process, the time interval between each round and each round is far and the correlation is limited, so it is unrealistic to build it into a MDP. Secondly, the existing works focus on data privacy protection, and has not introduced such a complex function as the auction mechanism. Therefore, the measurement of privacy and utility in existing work is based on Mean Square Error (MSE), mutual information and other theoretical values. These measurement methods have limited effect in dealing with such a complex process as the auction market. To this end, it is necessary to design a practical and effective reinforcement learning-based privacy preserving method in double auction market.

In this paper, we introduce a practical method for privacy preservation against inference attacks in a double auction market, which is grounded in the principles of reinforcement learning. It is proved to resist the attack in practice with minimal impact on auction performance (social welfare). **Note that although our research is aimed at the typical Mcafee mechanism in double auction market, the idea of designing practical privacy protection method against privacy inference attacks from the perspective of offense and defense can be applied to most of the current scenarios where differential privacy is applied.** This work continues our previous research published in 2022 IEEE T-ASE [20], where we designed a Bayesian privacy inference attack for the auction market from the attacker's point of view. In our paper, we focused on the protection perspective and studied the limitations of differential privacy on practical applications. We formalize the privacy protection problem against privacy inference attacks in the auction market as a privacy-utility trade-off (PUT) problem and design a novel reinforcement learning privacy protection method, thereby achieving a truly practical privacy protection method. This paper and previous work have comprehensively studied the threat of privacy inference attacks in the auction market from the perspectives of attackers and defenders, respectively. Specifically, the contributions of this paper are listed below:

- ***Attack-defense privacy preserving game model in double auction market:*** To tackle the inadequacies of differential privacy in mitigating inference attacks within the double auction market, we propose a novel model based on an attack-defense privacy preserving game. Specifically, after collecting all the bids from the bidders, the auctioneer will add noise to the bids to act as a defender, and judge its effect by executing the McAfee mechanism and Bayesian inference attack method to act as an attacker. By repeatedly executing this operation, the auctioneer aims to find a set of noise deployment methods to get better utility loss as well as privacy performance.
- ***Reinforcement learning-based privacy preserving method:*** Since the two objectives of the game model need to execute other two algorithms, namely McAfee transaction mechanism and Bayesian privacy inference attack method, it is difficult to establish a standard mathematical optimization model. To solve the game model, we formalize the above attack-defense game model into a Markov Decision Process (MDP) to facilitate the solution. Specifically, the state is the observation of the overall valuations under the current steps. The action is the noise added to each bidder, and the added noise can only be selected from 0, $+\gamma$, $-\gamma$ (action iteration noise step). The reward consists of three parts. The Bayesian attack mechanism is used to attack the current valuations, and the obtained attack success rate is taken as the privacy reward. While using the McAfee mechanism to make decisions on current valuations, the difference between the current social welfare and the original social welfare is taken as the utility reward. We also introduced a delayed penalty item to reduce training time. Finally, we deploy a DDPG algorithm to solve this problem.
- ***Extensive evaluations:*** We conducted extensive evaluations to verify the effectiveness of our proposed attack-defense game-based reinforcement learning privacy-preserving method. First, in terms of reinforcement learning network training, our mechanism can achieve convergence only after 800 iterations. Secondly, in terms of the trade-off between utility and privacy, our mechanism can reduce the attack success rate from nearly 100% to less than 20%, and make the utility deviation less than 5%. Finally, compared with the existing differential privacy protection mechanism, our mechanism can achieve the optimal privacy protection effect, while ensuring the minimum deviation of utility.

The rest of this paper is organized as follows. In Section II, the related work is proposed. In Section III, we briefly review the technical preliminaries and introduce the notations in this paper. In Section IV, we proposed a game model of defense and offence in auction market. In Section V, we introduce the practical reinforcement learning-based privacy preserving method. In Section VI, a comprehensive evaluation is conducted. Section VII, the discussion and future work are given. Finally, we conclude this paper in Section VIII.

II. Related Work

Privacy protection theory have evolved in recent years, and it mainly includes two kinds of privacy protection methods, which are cryptography-based methods and perturbation-based methods. Cryptographic-based methods protect the privacy by encrypting the sensitive data directly, and it mainly includes homomorphic encryption [6], multi-party secure computing [27], block chain [28] and so on. While perturbation-based methods protect the privacy by adding some noise to the sensitive data, and it mainly includes differential privacy [29], [30], k-anonymity [31], l-diversity [32], and so on.

Differential privacy was originally used to protect privacy information in simple scenarios such as data query and data publication [30], [33]. With the in-depth research, it is now widely used in location privacy protection [34], machine learning [35], [36], auction market [37] and other scenarios. For example, Ye et al. [35] ensured that the machine learning trainer cannot threaten personal privacy while ensuring the availability of the dataset by adding some noise to the training data, and it is proved to satisfy the ε differential privacy. Andres et al. [34] proposed a differential privacy algorithm for Geo-indistinguishability, which brings binary differential privacy to the field of location privacy protection. Zhang and Zhu [36] proposed a privacy preserving method based on differential privacy for distributed machine learning architecture. In fact, the essence of the above-mentioned research on location information and machine learning is still to protect the privacy of the database.

Recently, differential privacy has gradually been used to protect privacy in more complex algorithms, such as auction mechanisms [18], [38]. Initially, scholars directly introduced differential privacy into the auction market, utilizing the "indistinguishable" feature provided by differential privacy in databases to ensure the security of certain sensitive information in the auction market. For example, Li et al. [13] designed a DP-based double auction mechanism in energy trading market. The mechanism is proved to satisfy ε-differential privacy, individual rationality and incentive compatibility. Aiming at the IoT-based data trading market, Zhang and Zhong [38] proposed a DP-based auction mechanism using an exponential mechanism, and it can ensure that the transaction process is not affected by inference attacks. The above studies have verified the effectiveness of the proposed mechanism through theoretical proof, indicating that in practical situations such as auctions, this verification method is not convincing. Gradually, scholars have begun to test the effectiveness of differential privacy through privacy inference attacks [19]. Especially in the preliminary work of this paper [20], we proposed a privacy inference attack method targeting bilateral auction mechanisms, providing a practical evaluation for privacy protection auction mechanisms. However, in the field of differential privacy protection auction mechanisms, to truly implement practical, it is not only necessary to implement reliable privacy protection functions, but also to reduce the impact on data availability and auction results. In the auction market, the choice of differential privacy parameter ε will significantly affect the auction results, but few researchers have studied the selection method in detail. Therefore, how to choose the appropriate level of differential privacy is a key issue in the practical design of differential privacy protection auction mechanisms, and it is also the main problem that this paper aims to solve.

Scholars define differential privacy as a privacy-utility trade-off (PUT) problem, which contradicts privacy protection and data availability. In the field of database privacy protection, scholars often use optimization methods or reinforcement learning methods to solve the PUT problem. For example, Ecenaz Erdemir et al. [21] and Jiang et al. [25] proposed the time series-based mobility data noising methods based on reinforcement learning, which use mutual information to represent privacy. And the experiments show that they can provide effective privacy protection performance. Aiming at the PUT problem in data query scenario, Zhang et al. [24] proposed a reinforcement learning-based privacy preserving method. In the smart metering system, Sun et al. [39] proposed a reinforcement learning-based Electric Vehicle-assisted battery load hiding mechanism. This method formalizes the charging and discharging behavior of electric vehicles into a Markov Decision Process(MDP). With the goal of realizing the trade-off between cost and privacy, it is solved through a model-free Q-learning algorithm. In the auction market, how to build Markov Decision Process (MDP), measure utility and privacy, and reduce training time are three urgent problems to be solved. The aforementioned works address the PUT problem of differential privacy in static data storage. In contrast, the auction market is a dynamic process, presenting new challenges in building Markov Decision Processes and developing privacy and utility reward indicators.

Therefore, in response to the privacy inference attack threat faced by the auction market and the shortcomings of commonly used differential privacy in practical aspects, a new practical privacy protection method based on reinforcement learning is studied in this paper.

III. Preliminaries

In this section, we will introduce the technical knowledge of our paper. We will first introduce the system model and threaten model of the double auction market. Then introduce the most basic double auction algorithm McAfee. Finally we will review the Bayesian-based inference attack model against McAfee.

A. Model of Double Auction Market and Notations

1) System Model: The typical double auction market always contains multiple sellers, multiple buyers and an auctioneer. The sellers and buyers are allowed to submit their bids to the auctioneer freely, and the bidding information always include valuations and demand/supply volume. After receiving the bidding information, the auctioneer is responsible for making the decision about winners, payments and trading volume. Note that, in order to prevent collusion between participants and protect bidding privacy, so as to achieve a fair transaction, the double auction market is seal-bid market. That is to say, bidding information of both buyers and sellers is private information and not disclosed to anyone except auctioneer. The auction results, which include the winners, payment and trading volume, are public information. And the results can be viewed by anyone to ensure that there is no backroom. The decision making rule of the auctioneer will be the McAfee mechanism, and it can be proved to satisfy the properties of Incentive Compatibility (IC), Individual Rationality (IR), and weak budget balance.

2) Threaten Model: In most existing researches on privacy preserving double auction mechanism designing, the bidders' valuations are considered as the sensitive information that needs to be protected [13], [40], [41]. The reason behind this is that the valuations represent the bidders' willingness of the goods. Once this information is leaked, it may lead to targeted bidding and threaten the fairness of the market. On the other hand, in some particular markets, this information may also reflect the personal information of bidders. In this paper, we make a reasonable assumption that auctioneer is trusted, so that adversaries cannot invade into the auctioneer to steal the privacy information. Instead of that, we only consider indirect privacy inference attack threats [13], [20], [41]. The adversary can be an external attacker or a participant, and he is familiar with the auction process and mechanism. By collecting the results of multiple round of auctions, the adversary would compare them and reverse privacy valuations of specific bidders.

3) Notations: The notations involved in our paper are shown in Table I. Specifically, the double auction market happens in a slotted time period, and saved as $\mathrm{T} = [1, 2, \ldots, \mathrm{t}, ..]$. Then, we denote the buyers and sellers sets as B and C respectively. The buyers and sellers would submit the valuations to bid for the commodity. Note that, in our paper, the commodity is a single unit item, so the participant's valuation is the willing price for that single unit of the commodity. We denote the valuations of buyer i and seller j at time slot t as $b_{t,i}$ and $c_{t,j}$ respectively. Regarding to the auction results, eb_t and es_t are used to denote the valid price at time slot t. For buyer i and seller j, $R_{t,i}$ and $R_{t,j}$ denote their winning status, while 0 represents they do not win the bid, and 1 represents they win the bid. Furthermore, when the buyer i and seller j win the bid, the actual payments are denoted as $pb_{t,i}$ and $pc_{t,j}$.

B. Basic Double Auction Mechanism

McAfee mechanism is the most commonly used auction mechanism in the double auction market. The advantage of McAfee mechanism is that it can maximize the social welfare of participants while ensuring economic properties. Specifically, incentive compatibility ensures the fairness of the market, individual rationality ensures the non-negative utility of the participants, and weak budget balance ensures the profits of the auctioneer. In our paper, all of our research has focused on McAfee double auction mechanism. Since this

TABLE I
NOTATIONS

Symbols	Descriptions
T	Sets of time slots
B, C	Sets of buyers and sellers
n, m	Number of buyers and sellers
$b_{t,i}, c_{t,j}$	Valuation of buyer i and seller j
eb_t, es_t	Valid price at time slot t
$R_{t,i}, R_{t,j}$	Winning status
$pb_{t,i}, pc_{t,j}$	Actual payment
K	Sets of iteration rounds
ε_i^k	The added noise of buyer i in k^{th} round iteration
η_j^k	The added noise of seller j in k^{th} round iteration
γ	Action iteration noise step
r_p^k	Privacy reward
r_u^k	Utility reward
λ	Trade-off parameter
κ	Delay penalty factor
K_{max}	Maximum number of iterations

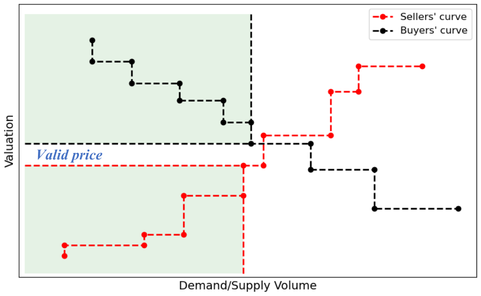


Fig. 1. Auction process.

mechanism has been proposed for many years and is widely used in the double auction market, we will intuitively describe its process. After collecting bidding information from all participants, the auctioneer first ranks all buyers in ascending order of their valuations and all sellers in descending order of their valuations. Then the vertical axis denotes the valuation, and the horizontal axis denotes the demand/supply volume (always 1 in our paper). The auctioneer will plot the sorted participants as shown in Fig. 1. Then the auctioneer will find the intersection of these two lines. The valuation of the first buyer and seller before the intersection (on the left) will be regarded as the valid price. All the buyers (sellers) whose valuations are larger (smaller) than the valid price will be selected as the winners of this round of auction. The actual payment of these bidders is the related valid price.

C. Bayesian-Based Inference Attack

In our previous work [20], we have proposed a Bayesian-based inference attack method against the McAfee double auction mechanism. In this section, due to the limitation of the space, we will make a brief review of the attack process.

First of all, the adversary will participant in multiple rounds of auctions (two rounds of auctions for example) to get some prior knowledge, including the valid price (eb_1, es_1, eb_2, es_2), winning status of all the bidders ($R_{t,i}$, $R_{t,j}$, $i \in B$, $j \in C$), adversary's own auction information (valuations, bidding results). Then the adversary will preprocess the prior knowledge. Specifically, the adversary will select the bidders whose winning status changed in the two rounds of auctions as attack target. And then determine the valuation candidates of the attack target i according to the valid price. After that, the adversary will calculate the condition probability (prior) $P(R_{1,i}, c_o)$, $P(R2, i, c_o)$ of each candidate c_o. That is, calculate the probability that the winning status is $R_{1,i}$ and $R_{2,i}$ when the valuation expectation of target user i is c_o. Then, the adversary will calculate the conditional probability (posterior) corresponding to all the num_c candidates:

$$P(c_o|R_i^1, R_{2,i}) = \frac{P(R_i^1, R_{2,i}|c_o)}{\sum_{j=1}^{num_c} P(R_i^1, R_{2,i}|c_j)}. \quad (1)$$

Finally, the valuation candidate c_o with the highest conditional probability (posterior) will be selected as the attack valuation. The specific technologies and proofs can be viewed in detail in [20].

D. Important Definitions of Differential Privacy

First of all, we will give a brief introduction of the differential privacy.

Definition 1 (Neighboring Database): Two databases which only differ in one value are denoted as neighboring databases. For example, databases $D_1 = \{d_1, d_2, \ldots, d_m, \ldots, d_n\}$ and $D_2 = \{d_1, d_2, \ldots, d'_m, \ldots, d_n\}$ can be treated as neighboring databases.

Definition 2 (Differential Privacy [9]): Consider a random mapping function $\mathcal{M}$ from D to R, if it holds the following constrain, it satisfy the differential privacy.

$$Pr[\mathcal{M}(D_1) \in R] \leq \exp(\varepsilon) Pr[\mathcal{M}(D_2) \in R], \quad (2)$$

where ε is the privacy parameter which determine the privacy protection efforts.

There are two main ways to implement differential privacy according to whether the random algorithm is discrete or continuous. The Laplacian/Gaussian Mechanism [42] ensures privacy by adding random noise to the continuous algorithm, while the Exponential Mechanism [43] ensures privacy by adding random process to the selection process of the discrete algorithm. Due to the space limitation, it will not be introduced in detail here.

IV. Attack-Defense Privacy Preserving Game Model in Double Auction Market

In this section, we will introduce why differential privacy does not work in the double auction market (McAfee mechanism) by showing how it works in its original scenario.

A. Why Differential Privacy Cannot Protect Privacy in Auction Market

Differential privacy is proposed to ensure the privacy of individual information in a database, and it would ensure that the overall availability of the database will not be affected. To be specific, we take a medical database as an example, consider a medical database of 100 people, in which 20 COV-19 patients were queried. When Tim is added to the database, the number of COV-19 queries becomes 21. Then without invading into the database, outsiders (adversary) will capture such private information as Tim having COV-19. Generally speaking, this process can be viewed as a mapping process. The input is all the information of the database, and the output is the number of people who get a certain disease in the database. The adversary performs this mapping several times and compares the results so as to infer the privacy of a certain data in the database without directly obtaining any data from the database. This can be regard as an inference attack. Differential privacy is to add a small noise to the input of mapping, so that the adversary can not confirm whether its attack result is credible. In particular, when Tim is added to the database, the number of COV-19 queries becomes 21, the adversary can not distinguish whether Tim has COV-19 or not (regard as two neighboring databases). Meanwhile, since the noise added is very small, a query result of 21 or 20 has little impact on overall data availability.

Because differential privacy can protect individual privacy without affecting the overall data availability, and the deployment difficulty is very low compared with encryption mechanism, it has been widely promoted to various scenarios, such as auction market. However, with the in-depth study of differential privacy, we find that its applicability in some scenarios is questionable, such as auction market, location privacy protection and so on. In this paper, we will take the auction market as an example to discuss why differential privacy is not effective in resisting such privacy inference attacks. Meanwhile, we will design a practical privacy protection method for the auction market that can really resist such inference attacks.

Regarding to the auction market, the auction mechanism can be viewed as a mapping process, where the input is the bidders's bidding information (valuations) and the output is the auction results. Similarly, the input is the sensitive information, and the output is a public information. The adversary's goal is to prevent adversary from inferring valuation from results, and this is also a process that protects individual privacy without compromising overall availability. However, there are some problems that apply differential privacy theory in this scenario.

First of all, different from data query, the mapping relationship of auction mechanism is very complex. So, from the adversary's point of view, it is impossible to infer the private input from the public output so easily like the data query scenario. **Whether such inference attacks exist, or whether such attacks can threaten all inputs, is questionable.** However, almost all current differential privacy-based auction mechanisms consciously ignore the construction of attack models. As a result, the existing researches only stays at the theoretical level, and even fails to verify its effectiveness in resisting inference attacks. According to our previous research, privacy inference attacks similar to database scenarios do exist in the auction market, but it can only threaten part of the input [20]. Secondly, because the current research on differential privacy has ignored the in-depth study of inference attack mechanism, almost all differential privacy deployment methods uniformly add noise to all input private information. According to our previous research, **privacy inference attacks in the auction market will not threaten all bidders' valuations, and the indiscriminate differential privacy deployment method will reduce the overall data availability.** Thirdly, applying differential privacy mechanisms (whether Gaussian or Exponential mechanism) in this scenario can indeed prove that the auction process satisfies $\Delta\varepsilon$-differential privacy. That is, the adversary cannot distinguish which neighboring databases produced the auction result. However, the sensitivity of the mapping process $\mathcal{M}$ needs to be limited to achieve a good differential privacy performance theoretically. Therefore, the difference of neighboring databases would not be very large. Specifically, an adversary may not be able to distinguish a bidder's valuation as 2 or 2.1. But in this case, we can say that the adversary has successfully stolen the privacy of bidder. The reason is that the difference between real and inferred valuation is too small, and either will threat the bidder's privacy. **Therefore, the 'indistinguishability' achieved by differential privacy may not be enough to really protect the inferred privacy threat.**

In conclusion, considering the existence of privacy inference attacks in the auction market and the aforementioned problems with differential privacy, which lead to unreliable privacy protection, it is urgent to investigate practical privacy-preserving mechanisms that can defend against privacy inference attacks as an alternative to differential privacy.

B. Design Rational

As introduced before, we know that the inference attack is to infer the private input (disease or valuation) through the public output (query result or auction result) and the adversary's understanding of the algorithm process (query algorithm or auction algorithm). Therefore, this kind of attack does not threaten the process of system storage and transmission. Even if encryption is adopted, this kind of attack can not be resisted. Obviously, the most effective way to resist this attack is to add noise when executing the algorithm, thus reducing the correspondence between output and input. However, according to the above analysis, we find that there are two important problems in the application of differential privacy to the auction market. One is that the inference attack in the auction market is not aimed at all inputs, and the indiscriminate deployment of differential privacy will reduce the availability of the overall data. Second, the "indistinguishable" property that differential privacy can guarantee is of little significance in protecting evaluation in the auction market. Therefore, we need to propose a novel privacy protection method based on noise to resist the threat of privacy inference in the auction market.

According to the inference attack mechanism, we find that there is a strong correlation between the inference attack threaten and the input data (valuation) in the auction market. Specifically, adversary launches inference privacy attacks by repeatedly entering the auction market, modifying their own valuations, and comparing the differences between auction results. So that, we can conclude that for those participants whose valuation are relatively extreme (large or small), no matter how the adversary modifies his own valuation, the winning results of these participants will not change (they will

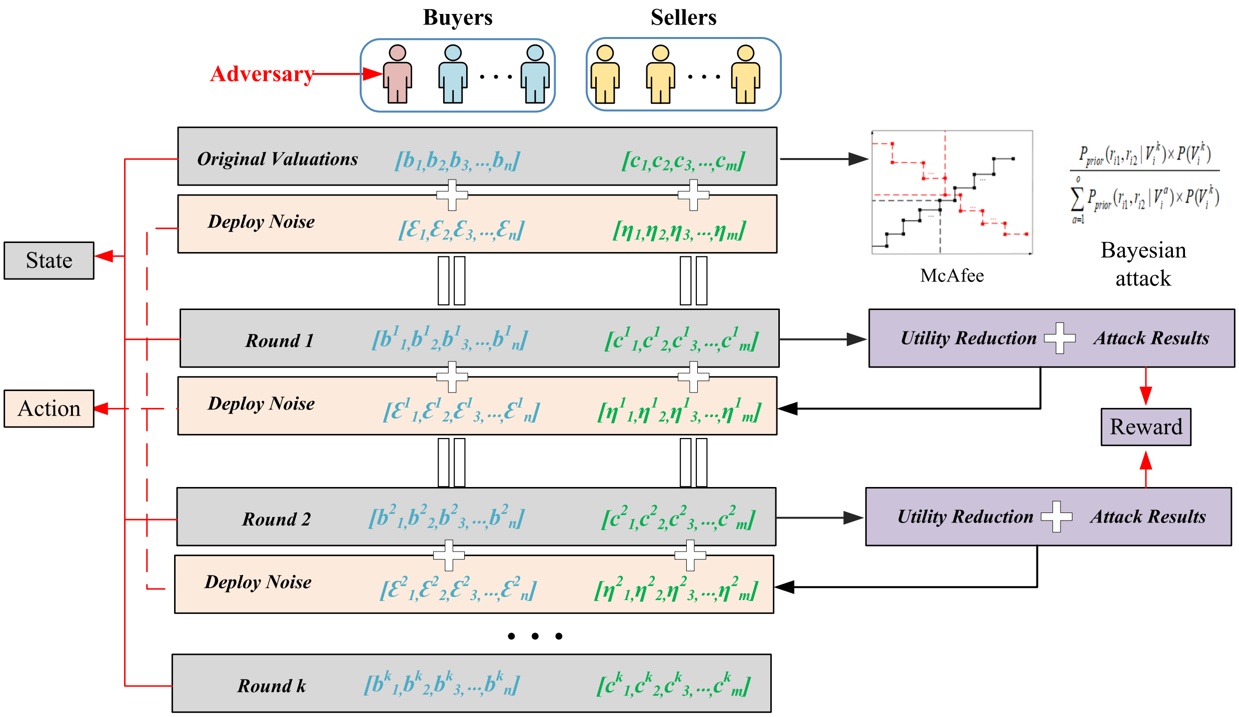

Buyers
Sellers
Adversary
Original Valuations
$[b_1,b_2,b_3,...,b_n]$
$[c_1,c_2,c_3,...,c_m]$
Deploy Noise
$[\varepsilon_1,\varepsilon_2,\varepsilon_3,...,\varepsilon_n]$
$[\eta_1,\eta_2,\eta_3,...,\eta_m]$
McAfee
$\frac{P_{prior}(r_{i1},r_{i2}|V_i^k)\times P(V_i^k)}{\sum_{a=1}^{o} P_{prior}(r_{i1},r_{i2}|V_i^a)\times P(V_i^k)}$
Bayesian attack
State
Action
Round 1
$[b^1{}_1,b^1{}_2,b^1{}_3,...,b^1{}_n]$
$[c^1{}_1,c^1{}_2,c^1{}_3,...,c^1{}_m]$
Deploy Noise
$[\varepsilon^1{}_1,\varepsilon^1{}_2,\varepsilon^1{}_3,...,\varepsilon^1{}_n]$
$[\eta^1{}_1,\eta^1{}_2,\eta^1{}_3,...,\eta^1{}_m]$
Utility Reduction
Attack Results
Reward
Round 2
$[b^2{}_1,b^2{}_2,b^2{}_3,...,b^2{}_n]$
$[c^2{}_1,c^2{}_2,c^2{}_3,...,c^2{}_m]$
Utility Reduction
Attack Results
Deploy Noise
$[\varepsilon^2{}_1,\varepsilon^2{}_2,\varepsilon^2{}_3,...,\varepsilon^2{}_n]$
$[\eta^2{}_1,\eta^2{}_2,\eta^2{}_3,...,\eta^2{}_m]$
Round k
$[b^k{}_1,b^k{}_2,b^k{}_3,...,b^k{}_n]$
$[c^k{}_1,c^k{}_2,c^k{}_3,...,c^k{}_m]$

Fig. 2. Game model of attack and defense in double auction market.

still win or lose with their extremely large and small valuations). The vulnerable participants are those whose valuation is on the edge of winning or losing. Once the adversary's valuation changes, the vulnerable participants' winning status will also change accordingly, and their valuation information will be stolen as a result. Therefore we can conclude that, once a set of valuations is determined, the privacy threats it contains are determined too. However, from the perspective of privacy protection deployment, due to the strong coupling relationship between valuations, it is difficult for us to obtain a specific optimal privacy protection strategy (i.e. noise adding strategy for each bidder) from the theoretical level, so as to protect privacy in practical. Furthermore, to genuinely achieve a practical privacy protection mechanism in double auction markets, it is crucial to consider the trade-off between privacy protection effectiveness and data availability. In other words, enhancing privacy protection effectiveness cannot be pursued solely by indiscriminately increasing the magnitude of noise, as this would significantly decrease the economic indicators of the auction market. Therefore, the two essential challenges in designing a truly practical privacy protection mechanism are to ensure its resilience against privacy inference attacks and, concurrently, to minimize the impact of the introduced noise on the original functionality of the auction market.

To this end, this paper adopts a novel perspective to investigate privacy concerns in auction markets, treating privacy protection as a game model to ensure the resilience of privacy protection mechanisms against privacy inference attacks, in which defenders deploy differential privacy noise while attackers launch privacy inference attacks, and the game model aims to get better utility loss as well as privacy performance (PUT). Since the two objectives of the game model need to execute other two algorithms, namely McAfee transaction mechanism and Bayesian privacy inference attack method, it is difficult to establish a standard mathematical optimization model. Additionally, it leverages reinforcement learning to autonomously explore optimal noise generation rules, aiming to balance the impact of noise on utility as much as possible. On the premise of existing inference attack model, we are going to design a game model of attack and defense to continuously attack and defend a group of valuation input, so as to find a noise deployment method to determine the location and size of noise for this group of valuation input from the perspective of practice. Similarly, the goal of noise deployment method is consistent with that of differential privacy, that is, to improve privacy protection performance as much as possible and reduce the impact on data availability as much as possible, which can be regarded as Privacy-Utility Trade-off (PUT) problem. Meanwhile, this optimization problem needs to locate the position where the noise is applied, and to determine the size of each noise, so the dimension of the solution is large. Considering that reinforcement learning can solve such problems simply after training by continuously learning the experience of deployment noise. Therefore, we consider adopting the method of reinforcement learning to solve this game model. This innovative game approach addresses non-intrusive inference attack problems, providing assistance not only in enhancing privacy protection in auction markets but also offering insights for defending against inference threats in other domains.

C. Game Model of Attack and Defense in Double Auction Market

In this section, we will introduce the game model of attack and defense in double auction market which is shown in Fig. 2.

In general, we consider a specific noise deployment progress for a group of bids containing n buyers and m sellers to protect privacy. The game model is run round by round, and we save the rounds in set $K = [1, 2, 3, \ldots, k, ..]$. In which, we choose a buyer as the adversary, and continue to attack this group of valuations by modifying his/her valuation in this process. Note that, when a group of bids is determined, the sensitive valuations will be determined. And the sensitive valuations are not depending on the adversary, so we can choose any bidders as the adversary. Here we set the adversary as a fixed buyer. First of all, we collect the original valuations set as:

$$\xi(0) = [b_1, b_2, b_3, \ldots, b_n, c_1, c_2, \ldots, c_m] \tag{3}$$

where b_i, c_j represent the valuations of buyer i and seller j.

The McAfee mechanism is executed on it to obtain the original utility. Meanwhile, buyer 1 is used as the adversary to launch inference attack and calculate the success rate of attack on this group of data. These two results are used as the evaluation benchmark for the subsequent noise deployment strategy. Then, we will randomly generate a set of noise deployment strategies as:

$$N^0 = [\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_n, \eta_1, \eta_2, \ldots, \eta_m], \tag{4}$$

where ε_i, η_j represent the noise added to buyer i and seller j.

And add them to the original valuations set $\xi(0)$ to obtain the first round valuations set as:

$$\xi(1) = [b_1^1, b_2^1, b_3^1, \ldots, b_n^1, c_1^1, c_2^1, \ldots, c_m^1] = \xi(0) + N^0, \tag{5}$$

where b_i^1, c_j^1 represent the valuations of buyer i and seller j at round 1.

Then the McAfee mechanism and inference attack method are executed on the valuations set ξ_1. Then we compare the utility and attack success rate of this round with the original utility and attack success rate, so as to determine the next round of noise deployment strategy N^1. Eventually, the process repeats until we find the set of valuations that minimizes the attack rate and minimizes utility reduction.

In this process, we simulate the game process of attack and defense, so as to find a defense method that can effectively resist inference attacks (reduce the attack success rate as much as possible), and meanwhile reduce the impact on the auction process. This dynamic process obviously conforms to Markov property, so it is most suitable to use reinforcement learning method to solve it.

V. Reinforcement Learning-Based Privacy Preserving Method in Double Auction Market

In this section, we will introduce the reinforcement learning-based privacy preserving method in double auction market. First, we will formalize the game model as a Markov Decision Process (MDP), and then we will introduce the reinforcement learning solution and training method based on DDPG.

A. MDP for Game Model

Regarding to the game model of attack and defense in double auction market which is shown in Fig. 2, the auctioneer is regarded as the agent, and the auctioneer is responsible for determining the noise of each bidders. Specifically, the auctioneer is able to observe the states s_k (valuations set) of the environment repeatedly, and determine the action a_k (noise deployment strategy). Then the state transfers to the next state s_{k+1} according to the action a_k. Finally, a reward which reflects the action a_k's utility and privacy is used to instruct the next action. In this paper, the MDP model of the game model of attack and defense in double auction market is defined as three elements $\{S, A, R\}$, and we will introduced them in detail.

1) Valuations-Based State Space: We use $S = [s_1, s_2, \ldots, s_k]$ to denote the system state space. Specifically, the state at round k is the observation of the valuations set:

$$s_k = [b_1^k, b_2^k, b_3^k, \ldots, b_n^k, c_1^k, c_2^k, \ldots, c_m^k], \tag{6}$$

where b_i^k, s_j^k represent the valuations of buyer i and seller j at round k.

2) Noise Deployment-Based Action Space: $A = [a_1, a_2, \ldots, a_k]$ represents the action space. After obtaining the observation of valuations set at each time slot, the agent (auction) will determine the noise of each bidders, which can be formalized as the action:

$$a_k = [a_1^k, a_2^k, \ldots, a_n^k, a_{n+1}^k, a_{n+2}^k, \ldots, a_{n+m}^k], \tag{7}$$

where a_i^k represent the noise added to buyer i (or seller $i-n$, when $i > n$) at round k.

In our paper, we need to find a specific noise deployment method for a specific set of valuations. Therefore, the noise to be added to each bid should be a certain value, rather than white noise subject to some distribution. To this end, we will approach the optimal noise deployment method step by step through action. Then, the values of action at round k can be expressed as:

$$a_i^k = \begin{cases} -\gamma, \\ 0, \\ \gamma. \end{cases} \tag{8}$$

where γ represents the action iteration noise step. When the value is large, the training speed is fast, but the final result may be further optimized. In Section VI, we will analyze its sensitivity.

3) Utility and Privacy Measurement: There exists two main targets in our game model of attack and defense in double auction market. The one is to minimize the impact on the economic attributes of the auction market. The other is to increase the privacy protection effect as much as possible. Therefore, we need to consider how to quantify these two indicators to provide a basis for the MDP model.

First regarding to the utility target, we intend to use the difference of the total utility (social welfare) of the auction market participants to represent it:

$$\Delta U^k = |u^k(s^k) - u^0(s^0)| \tag{9}$$

where ΔU^k represents the utility indicator at round k, and it is calculated by the difference between the total utility $u^k(s^k)$ of the market at round k and the total utility $u^0(s^0)$ of the original market. Furthermore, the total utility can be expressed as:

$$u^k(s^k) = \sum_{i=1}^{n}(b_i^k - pb_i^k) + \sum_{j=1}^{m}(ps_j^k - c_j^k) \tag{10}$$

where b_i^k, c_j^k represents the valuations of buyer i and seller j, and pb_i^k, ps_j^k represents the actual payment/reimbursement of buyer i and seller j.

Second regarding to the privacy target, we intend to use the attack success rate to represent it. Specifically, the agent will launch an inference attack as the adversary (buyer 1 in our paper) at round k to obtain the attack success rate. As introduced in [20], the Attack Success Rate (ASR) is denoted as:

$$P^k = \frac{Num_s^t}{Num_t^t} \tag{11}$$

where P^k represents the attack success rate at time slot t, Num_s^t represents the number of attack success bidders at time slot t, and Num_s^t represents the number of attack target at time slot t. Note that, since the valuation in the auction market is a continuous value, we believe that the inference results are successful within a range. Therefore, in our paper, the attack success bidder o should follow:

$$\hat{b}_o^k \in [0.9 \cdot b_o^k, 1.1 \cdot b_o^k] \tag{12}$$

$$\hat{c}_o^k \in [0.9 \cdot c_o^k, 1.1 \cdot c_o^k] \tag{13}$$

where $\hat{b}_o^k, \hat{c}_o^k$ represent the inferred valuation of buyer/seller o at round k.

4) Trade-off of Reward: The reward can be mainly divided into three parts: (i) utility reward, (ii) privacy reward, (iii) delay penalty item.

First, we normalized the utility difference indicator ΔU^k to obtain the utility reward r_u^k:

$$r_u^k = -\frac{\Delta U^k}{u^0(s^0)} \tag{14}$$

where r_u^k represents the utility difference, so we need to obtain a smaller value of r_u^k to guarantee a smaller utility difference. Note that, although the above formula cannot guarantee that r_u^k is strictly less than 1, when the value is greater than 1, it indicates that the difference between the utility at round $k =$ and the original utility is too large, which will not be allowed. So it is reasonable to assign a value greater than 1 to prohibit this behavior.

Second, considering that the attack success rate is within the range of [0, 1], then the privacy reward r_p^k can be expressed as:

$$r_p^k = 1 - P^k \tag{15}$$

Third, considering that the auction market is continuous, the auction decision at a moment needs to be completed as soon as possible, which requires that the number of iterations of reinforcement learning be as small as possible. Therefore, here we define a delay penalty reward to constrain the number of iterations:

$$r_d^k = \frac{k\kappa}{K_{max}} \tag{16}$$

where κ represents the delay penalty factor, K_{max} represents the set maximum number of iterations.

Then the total reward can be expressed as:

$$r^k = \lambda r_p^k + (1-\lambda) r_u^k - r_d^k, \tag{17}$$

where λ represent the trade-off parameter. The agent can adjust the parameter size to decide whether to focus more on privacy or utility.

B. Deep Deterministic Policy Gradient

Since the double auction market contains multiple bidders, each bidder has different noise adding strategies, resulting in a large action space. Traditional DQN algorithms mainly focus on discrete actions and are not suitable for reinforcement learning models with large action spaces. The biggest advantage of DDPG over other deep neural networks is its ability to learn more efficiently on successive actions. Submiliary with DQN, DDPG also uses a replay buffer and two neural networks with the same structure but different parameter update frequency, which can effectively improve the learning efficiency, reduce the correlation between parameters and increase the convergence speed. At the same time, as an off-policy algorithm, it absorbs the advantages of Actor-Critic and policy gradient. Policy gradient is different from updating the network based on the reward value. This method directly selects the action, this is consistent with the setting of adding noise to the data in our model. Therefore, DDPG is more suitable for offensive and defensive models of electrical energy trading than other reinforcement learning algorithms.

The proposed DDPG algorithm is shown in Algorithm 1, which mainly includes two parts. The first part is the interaction between the environment and the agent. This part defines the specific parameters of the environment, actions and rewards. The second part is the update of Actor and critic network, this part defines the replay buffer and update formula.

The interaction process has been mentioned in the MDP introduction. Rejudging to the network update process, it contains four networks, i.e., actor network, critic network, target actor network and target critic network. The critic network evaluates the behavior of the agent which uses the loss function to update the parameter θ_c of the Critic network. The TD-error of the critic network can be expressed as:

$$TD_{error} = r_i + \gamma_{ac} Q'(s_{i+1}, \pi'(s_{i+1})|\theta_c') - Q(s_i, a_i|\theta_c) \tag{18}$$

where r_i represents reward, γ_{ac} represents learning rate. And then, we use the Mean Square Error (MSE) of TD-error to represent the loss function of the critical network:

$$L = \frac{1}{\omega}(TD_{error})^2 \tag{19}$$

Actor network update adopts the deterministic gradient descent method, and the negative value of the generated Q value is used as the loss. If the gradient is large, it means that the parameters of the actor network should be updated in this direction. In this paper, we update the actor network with multiple pairs of Δ-samples:

$$\Delta = \frac{1}{\omega} grad(s, a|\theta_c)|_{s=s_i, a=\pi(s_i)} \times grad(s|\theta_a)|_{s_i} \tag{20}$$

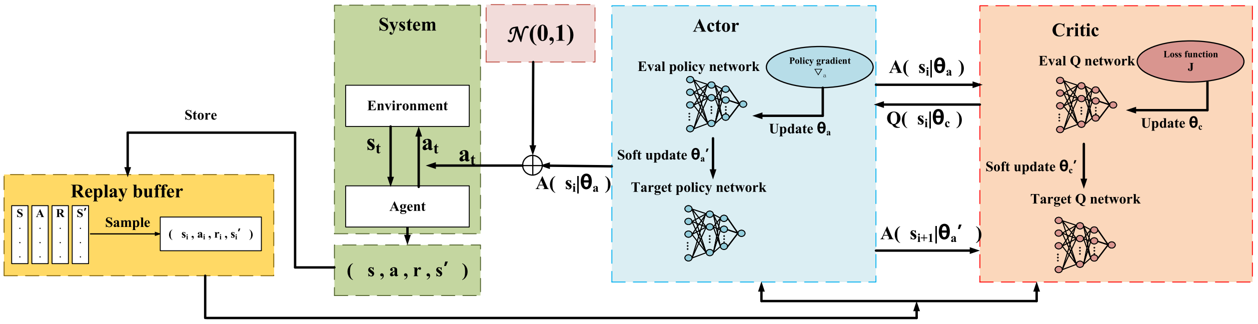


Fig. 3. Network parameter update.

Algorithm 1 DDPG Algorithm

```
Input: Original bid Information $\xi(k_0)$, actor and critic network
       parameters $\theta_a$, $\theta_c$, discount factor, learning rate $\gamma_{ac}$, $\eta_{ac}$,
       maximum round limit $K_{max}$, reward decay factor $\kappa$
Output: Optimal policy $\pi(s|a)$
Initialize environment $E$
Initialize round set $k = 0$
Initialize replay buffer $\Omega$
for $k = 0$ to $K_{max}$ do
    Get current observation value $o_k$, state value $s_k$
    Select action $a_k = \pi(o_k, s_k) + \Psi$, the current action is
    chosen
    based on strategy and exploration noise.
    $r_k = -\lambda r_u^k - (1-\lambda) r_p^k - r_d^k$,
    The reward function is mainly composed of three parts,
    privacy, utility, and the reward decay function
    Update environment,$s_k \leftarrow s_{k+1}$
    for $\Omega < \Omega_{max}$ do
        Store transition $(s_k, a_k, r_k, s_{k+1})$ in $\Omega$
    end
end
Sample a random minibatch of $\omega$ transitions from $\Omega$
Set $TD_{error} = r_i + \gamma_{ac} Q'(s_{i+1}, \pi'(s_{i+1})|\theta_c') - Q(s_i, a_i|\theta_c)$
for $k < K_{max}$ do
    for $i < k$ do
        $L = \frac{1}{\omega}(TD_{error})^2$
        $\Delta = \frac{1}{\omega} grad(s, a|\theta_c)|_{s=s_i, a=\pi(s_i)} \times grad(s|\theta_a)|_{s_i}$
    end
    Update critic by minimizing $L$
    Update actor by using sampled policy gradient $\Delta$
    Asynchronous update the target networks:
    $\theta_a' = k\theta_a + (1-k)\theta_a'$
    $\theta_c' = k\theta_c + (1-k)\theta_c'$
end
```

Note that, there is the soft update between the main and target networks. The algorithm does not copy all the parameters of the main network to the target, but updates parameters with a small step. This method is also called soft update. The passing of network parameters can be described as:

$$\theta' = k\theta + (1-k)\theta' \tag{21}$$

The updates of Actor and Critic are shown in Figure 3.

VI. Evaluation

In this section, evaluations are conducted to verify the effectiveness of our proposed attack-defense game-based reinforcement learning privacy-preserving method.

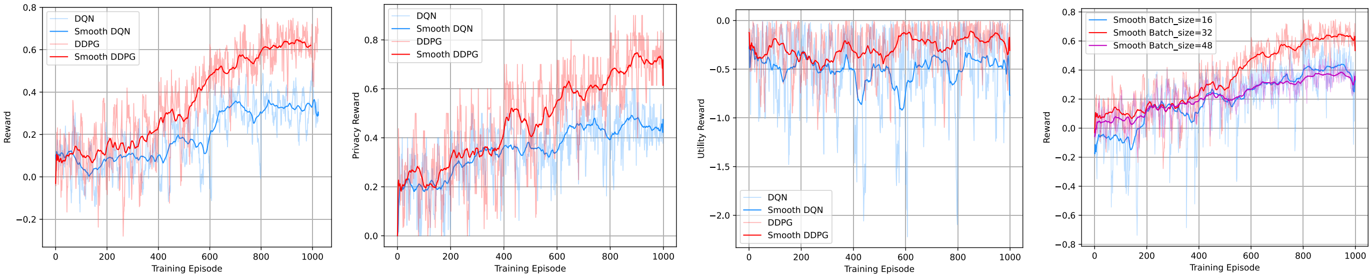
A. Experimental Settings

We consider a double auction market which consists of 10 buyers and 10 sellers. The valuations of them follow the normal distribution: $N(v_i, 1)$, where v_i represents the expected valuation of the participant i. While the expected valuations of each participants also follows a uniformly distributed from 1 to 10. The attack accuracy is set as 0.5, which means when the difference between the inferred valuation and the expected valuation is less than 0.25, this attack will be regarded as a successful attack. Without specific instructions, the trade-off parameter λ is 0.75 and the action iteration noise step γ is 1, and the batch size is 32. Regarding to the reinforcement training process, the learning rates of Actor and Critic networks in DDPG are 0.01, 0.02 respectively. Soft replacement parameter is 0.01 and the reward decay is 0.9 for both networks. Both reinforcement learning networks were explored with e-greedy, e-greedy parameter was 0.9, while considering the practicality of the model, we limited the number of steps to 1024. All experiments in this Section were configured to complete the simulation with Python software on a desktop server with two 2.10 GHz Intel(R) Xeon(R) Gold 6230 CPUs, 64 GB RAM, and two NVIDIA GEFORCE RTX 2080 Ti.

B. Evaluation of Convergence

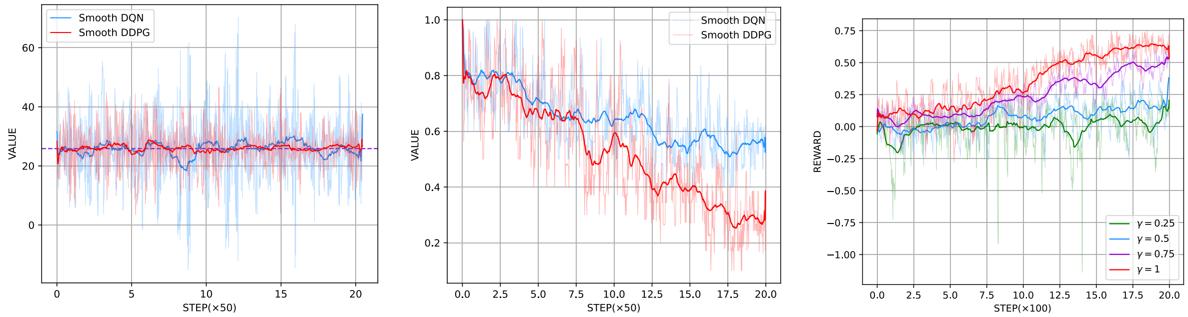
In the following, we will discuss the performance of the reinforcement learning network.

We first compare the reward update process of our method with typical DQN in Fig. 4-(a) to (c). According to Fig. 4-(a), it can be seen that the reward increase rate of the DDPG algorithm is significantly better than that of DQN, including convergence speed and maximum reward. Under the same reward function setting, with a limited number of 1000 steps, the reward value of DDPG can reach 0.7, and the reward value of DQN can only reach 0.3. Regarding to Fig. 4-(b), we can see that the privacy reward follows the same trend with the total reward. According to utility reward, it is a penalty item. As can be seen from Fig. 4-(c), the jitter of DQN is obviously more severe, and the effect is not as good as that of DDPG. Fig. 4-(d) indicates the influence of batch size on convergence speed and reward size. It can be seen that when the batch size is 32, both the convergence rate and the final value of the reward are optimal. So this paper selects the batch size as 32.



(a) (b) (c) (d)

Fig. 4. (a) to (c) Reward comparison of DDPG and DQN, (d) Batch sizes *v.s.* reward.



(a) (b) (c)

Fig. 5. (a) Utility, (b) Privacy, (c) Action iteration noise step.

Fig. 5-(a) and (b) show the real values of utility (social welfare) and privacy (attack success rate) after each iteration. According to these two figures, we can see that the trend of utility and privacy real value is consistent with that of reward. The index of attack success rate can be reduced to less than 20%, and the utility basically remains consistent with the original social welfare. Finally, fig. 5-(c) represents the impact of the action iteration noise step on the reward. As we analyzed before, smaller step size will slow down the convergence speed. And we also add delay penalty reward, so the smaller size cannot obtain a larger reward, when the number of iterations is long.

C. Trade-Off

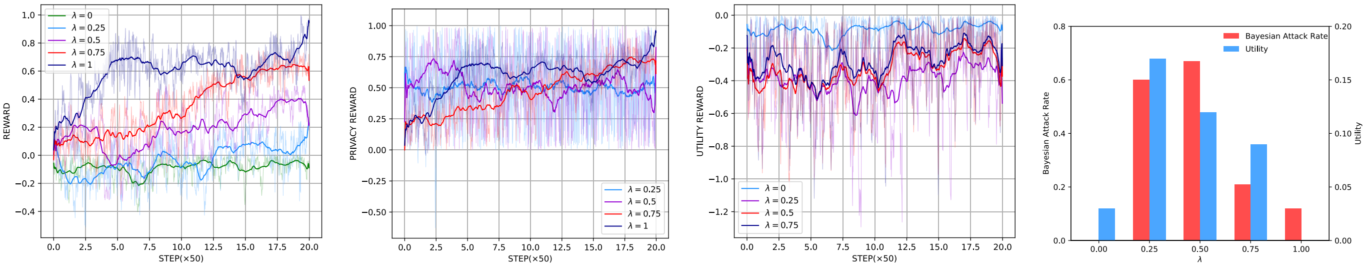
Next, we will discuss the impact of the trade-off parameter λ on utility and privacy.

Fig. 6-(a) shows the total reward under different trade-off parameter λ from 0 to 1. We can see that when λ is large, the overall reward is large. The reason is that privacy reward is a positive reward, while utility reward is actually a penalty item, which must be negative. Therefore, the smaller λ is, the larger the percentage of privacy is, and the overall reward will be higher. This is why when λ is equal to 1, the reward is maximum. However, the overall reward does not necessarily represent the actual privacy protection, and the transaction performance. Further, we explored the relationship between λ with privacy reward and utility reward. From Fig 6-(b) and (c), we can see that when $\lambda = 0.75$, the privacy reward is almost similar to that when $\lambda = 1$, and the utility reward is the best expect for $\lambda = 0$. Finally, Fig 6-(d) shows the utility deviation and the attack success rate of privacy attack average value of the last 50 iterations. It can also be seen that when λ equals 0.75, the proposed method has the best trade-off between utility and privacy. Therefore, λ is set as 0.75 in this paper.

D. Comparison With Other Privacy Preserving Methods

Finally, we will compare the utility and privacy performance of our proposed method with the existing typical privacy preserving methods. We choose several common differential privacy methods as a comparison to demonstrate the advantages of the proposed methods in terms of PUT, including: (i) Original method: solve the auction determining problem by typical McAfee mechanism. (ii) Gaussian-based Differential Privacy (Gaussian mechanism $\varepsilon = 1$ or 0.5, [15]): McAfee mechanism with Gaussian noise-based differential privacy ($\varepsilon = 1$ or 0.5) mechanism. (iii) Exponential-based Differential Privacy (Exponential mechanism $\varepsilon = 0.5$, [13]): McAfee mechanism with exponential mechanism. (Note that typical exponential mechanism focus on one side auction market, here, the double auction is divided into multiple one side auctions, so as to avoid the problem of very low efficiency when used directly, resulting in no comparability). (iv) Individual differential privacy ($\varepsilon = 0.5$) [20]: Exponential mechanism (ε) is adopted to make decisions on the bidding of target bidders.

First, Table. II shows a numerical example under differential privacy preserving methods, including Differential Privacy ($\varepsilon = 0.5$), Individual Differential Privacy ($\varepsilon = 0.5$) and our method. It is worth mentioning that, different adversaries may bring different attack results. The attack results shown in Table. II are obtained by assuming the seller 2 is the adversary. Furthermore, we obtain a mean Attack Success Rate (ASR) by assuming different adversaries to verify the privacy disclose threat of this group of valuations. From this table, we can



(a)

(b)

(c)

(d)

Fig. 6. (a),(b),(c) Reward, Privacy Reward, Utility Reward *v.s.* trade-off parameter λ, (d) Utility and Privacy *v.s.* trade-off parameter λ.

TABLE II
A NUMERICAL EXAMPLE

		Original			Gaussian DP [15]			Individual DP [20]			Our Approach		
		Valuation	Utility	Attack	Valuation	Utility	Attack	Valuation	Utility	Attack	Valuation	Utility	Attack
Seller	1	4.99	0		4.94	0		4.99	0		4.89	0	
	2	3.26	0.24		3.52	0.79		3.26	0.24		3.25	0.24	
	3	5.06	0		4.05	0	Y(5)	5.06	0		5.06	0	
	4	2.73	0.77		2.21	1.32		2.73	0.77		2.73	0.77	
	5	1.96	1.54		1.77	2.09		1.96	1.54		1.86	1.54	
	6	3.17	0.33		3.10	0.88		3.17	0.33		3.10	0.33	
	7	3.50	0		3.59	0.55		3.50	0	N(5)	3.50	0	N(4.5)
	8	5.22	0		4.92	0		5.22	0		5.22	0	
	9	4.21	0	Y(4.5)	4.04	-0.16	N(3)	4.07	0	Y(4)	4.00	0	N(6)
	10	6.73	0		7.92	0		6.73	0		6.73	0	
Buyer	1	2.98	0		3.54	0	Y(3)	2.98	0		2.98	0	
	2	5.98	0	Y(6)	6.56	1.89		5.48	0	N(3)	5.91	0	N(4)
	3	6.63	0.65		6.28	2.54		5.63	1.15		6.63	0.72	
	4	3.72	0		4.22	-0.37		3.72	0	N(3)	3.72	0	
	5	7.13	1.15		7.77	3.04		7.13	1.65		7.25	1.22	
	6	4.005	0	Y(4)	4.09	0	Y(4)	3.51	0		3.9	0	
	7	2.75	0		3.06	0		2.75	0		2.75	0	
	8	6.49	0.51		6.87	2.4		6.49	1.01		7.15	0.58	N(7)
	9	2.25	0		2.73	0		2.25	0		2.13	0	
	10	7.65	1.67		7.63	3.56		7.65	2.17		7.07	1.74	
Summary		Utility: 6.86, Average ASR: 100%, Seller/buyer valid price: 3.50, 5.98			Utility: 18.53, Average ASR: 68.7%, Seller/buyer valid price: 4.05, 4.09			Utility: 8.86, Average ASR: 41.5%, Seller/buyer valid price: 3.50, 5.48			Utility: 7.14, Average ASR: 18.6%, Seller/buyer valid price: 3.55, 5.91		
Note that: N indicates the target that has not been successfully attacked, Y indicates the target that has been successfully attacked.													

clearly see the mechanisms and effects of several privacy protections.

Regarding to the typical Gaussian DP method, we find that the protection strategy of this method is to add differential privacy noise indiscriminately to the valuations of all bidders. In terms of auction results, the valuations of each bidders and the valid price have greatly changed. Therefore, the winning status of many bidders has changed. For example, some bidders who cannot win have won at this time, and it will seriously affects the fairness of the auction market. Meanwhile, we also find that the personal utility of some bidders is less than 0, which also makes the market no longer satisfy the property of individual rationality. Overall, the totaly utility has improved to 270% of the original utility (from 6.86 to 18.35). Such impact on auction performance will significantly reduce the enthusiasm of bidders to participate in the market. In terms of privacy protection performance, the ASR of DP method is reduced to 68.7%. However, according to our in-depth research, we found that the DP noise may introduce more targets, which will lead to more bidders' privacy being stolen than the original scenario. Therefore, the performance of traditional Gaussian DP method in privacy protection is still poor.

Regarding to the exponential-based DP method, it is a mechanism based on probability winner selection, unlike other noised-based mechanisms that add noise to each participant's bid, so its privacy protection process cannot be demonstrated in Table. II. While from Table. III, we can still see the

TABLE III
UTILITY AND PRIVACY COMPARISON

Method	ASR	Utility
Original	100%	–
Gaussian DP(ε=1)[15]	84.4%	24%
Gaussian DP(ε=0.5)[15]	44.2%	41.9%
Exponential DP (ε=0.5)[13]	40.5%	58.6%
Individual DP(ε=0.5) [20]	32.5%	9.5%
Our approach	17.2%	3.9%

performance of the exponential-based DP method in terms of privacy and utility still has a significant gap compared to the algorithm proposed in this paper.

Regarding to the individual DP method, its privacy protection strategy is to add differential privacy noise to the bidders that act as valid prices in the original market. We have to admit that this has no impact on the winning status of bidders in the auction market. However, due to the significant change of the valid price, the total utility market also has 29% deviation (from 6.86 to 8.86). In terms of privacy protection, the ASR is further reduced to 41.5%. Meanwhile, we found that the individual DP method will protect the attack target bidders as much as possible. However, due to the deep coupling between valuations, it is possible to introduce new targets. In general, individual DP is proposed on the premise of understanding the attack mechanism, which can provide more targeted privacy protection than typical DP. However, because the added noise is random, or we cannot give a targeted noise adding strategy, the final utility and privacy performance is still uncontrollable.

Regarding to our approach, we can see that the strategy of autonomous learning through reinforcement learning is similar to the individual difference method, which is to add noise to sensitive bidders. However, the difference with individual DP is that it reduces the size of noise as much as possible, and it chooses to add a certain amount of noise to some bidders who are not target bidders. From the auction results, the winners have hardly changed, and the valid price has not changed much, and all of the bidders satisfy the property of individual rationality. So the impact on utility is very small, with only 4% deviation (from 6.86 to 7.14). From the perspective of privacy protection effect, our method can indeed further reduce the ASR to 18.6%. The reason is that the reinforcement learning method can automatically learn the coupling relationship between valuation data, and avoid introducing other attack targets while adding noise to protect them. Therefore, our method successfully learned the noise deployment strategy and optimized it on the basis of individual DP, thus achieving better results.

Furthermore, Table. III shows the average value of utility and privacy under multiple experiments. It further verifies our conclusions above. Our approach can indeed reduce the ASR to within 20% and control the utility deviation to within 5%.

In general, our mechanism provides the best privacy protection performance while ensuring that the auction results remain unchanged. Therefore, we can say that our mechanism realizes the trade-off between utility and privacy.

VII. Discussion and Future

The differential privacy protection method was originally designed to protect the privacy of data sets and make data indistinguishable from each other. It has been gradually applied to other scenarios, among which the electronic trading market is a very typical example. Due to the competitive relationship between buyers and sellers and the need to quickly solve the transaction results in the electronic trading market, differential privacy, a lightweight privacy protection method, is indeed more suitable for this scenario. At present, a lot of results have been achieved in this area, and most of the work has proved that the proposed mechanism can theoretically provide the required privacy protection performance (ε-differential privacy). However, according to our preliminary research, compared with data sets, electronic trading markets need stronger privacy protection to provide stricter non-discrimination in order to protect privacy. For example, the buyer's bid is 2.1 \$ or 2.2 \$, even if they cannot be distinguished, but it has fully exposed the buyer's private information, you can make profits from it. Therefore, this paper innovatively incorporates the Attack process into the design of defense mechanism. Defenders deploy differential privacy noise while attackers launch privacy inference attacks. An attack-defense privacy preserving game model is designed. This paper tries to find a privacy protection method that can really resist privacy inference attacks. Finally, it is proved that this method is better than the traditional method in Privacy protection performance and preventing Utility reduction, and solves the trade-off problem of privacy and utility. We believe that this idea can also be extended to other related privacy protection fields to provide more accurate privacy protection performance.

As mentioned above, this paper focuses on designing a new attack-defense game-based privacy protection framework and model, and to solve this game model, this paper adopts reinforcement learning which is more effective for this kind of problem. However, for the game optimization model, in fact, there are still a lot of mature and efficient methods in the field, here is a brief discussion, which is also the future research direction of this work. Firstly, traditional methods for solving optimization problems include applied mathematics and computation, heuristic algorithm [44], [45], etc. These algorithms are characterized by high stability, that is to say, no matter how complex the model is, the strategy solved is relatively optimal, if not optimal. However, there are some limitations in the solving speed of complex models and the ability of model reuse. Specially in our work, since the two objectives of the game model need to execute other two algorithms, namely McAfee transaction mechanism and Bayesian privacy inference attack method, it is difficult to establish a standard mathematical optimization model, so it is hard to use the traditional applied mathematics and computation method. Therefore, artificial intelligence algorithms represented by reinforcement learning have also become a mainstream method to solve such problems, such as reinforcement learning, online reinforcement learning, offline reinforcement learning, Continuous-Time Markov Jump Linear Systems, which can be used in dynamic systems through online reinforcement learning. Strategy learning and optimization for multi-agents can achieve their goals in this randomly changing environment. Compared with traditional optimization methods or reinforcement learning methods, it can better solve the optimal strategy in the abrupt environment of this complex game. However, the training speed of reinforcement learning is slow, and the effect is not guaranteed, which needs further research in the future. In short, at present, there are still many parts worth studying for the optimization and solution of the game architecture model of the offensive and defensive game proposed in this paper, which is also the future research direction of this paper.

VIII. Conclusion

In this paper, we studied the privacy preserving issue in the double auction market. First, we discussed that differential privacy has two problems in resisting indirect privacy inference attacks in auction market. The 'indistinguishability' property is not enough to resist the inference attack, and the privacy-utility trade-off (PUT) problem remains unsolved. To this end, we proposed an attack-defense game model in double auction market. The auctioneer acted as both the defender and the adversary, and found the optimal noise deployment strategy for a group of bids by constantly playing games with himself. This process is formalized as a Markov Decision Process (MDP). Specifically, the noise added to each bidders constitute the action space, and the valuations of each round constitutes the state space. While the reward consists of two parts: one is the utility privacy formed by the utility deviation, and the other is the privacy reward formed by the success rate of privacy attacks. At the same time, a penalty term related to the iteration time is added to speed up the iteration. We construct a DDPG network to solve the MDP. The simulation results show that compared with the existing privacy protection methods, our method is better both in the performance of utility and privacy, and achieves the trade-off of utility and privacy.

References

[1] P. McDaniel and S. McLaughlin, "Security and privacy challenges in the smart grid," *IEEE Secur. Privacy*, vol. 7, no. 3, pp. 75–77, May 2009.

[2] J. Lin, W. Yu, N. Zhang, X. Yang, H. Zhang, and W. Zhao, "A survey on Internet of Things: Architecture, enabling technologies, security and privacy, and applications," *IEEE Internet Things J.*, vol. 4, no. 5, pp. 1125–1142, Oct. 2017.

[3] M. Binjubeir, A. A. Ahmed, M. A. B. Ismail, A. S. Sadiq, and M. K. Khan, "Comprehensive survey on big data privacy protection," *IEEE Access*, vol. 8, pp. 20067–20079, 2020.

[4] X. Song, N. Wu, S. Song, and V. Stojanovic, "Switching-like event-triggered state estimation for reaction–diffusion neural networks against DoS attacks," *Neural Process. Lett.*, vol. 55, no. 7, pp. 8997–9018, Dec. 2023.

[5] Z. Zhang, X. Song, X. Sun, and V. Stojanovic, "Hybrid-driven-based fuzzy secure filtering for nonlinear parabolic partial differential equation systems with cyber attacks," *Int. J. Adapt. Control Signal Process.*, vol. 37, no. 2, pp. 380–398, Feb. 2023.

[6] J. H. Cheon and J. Kim, "A hybrid scheme of public-key encryption and somewhat homomorphic encryption," *IEEE Trans. Inf. Forensics Security*, vol. 10, no. 5, pp. 1052–1063, May 2015.

[7] J. Soria-Comas, J. Domingo-Ferrer, D. Sánchez, and D. Megías, "Individual differential privacy: A utility-preserving formulation of differential privacy guarantees," *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 6, pp. 1418–1429, Jun. 2017.

[8] B. Gedik and L. Liu, "Protecting location privacy with personalized K-anonymity: Architecture and algorithms," *IEEE Trans. Mobile Comput.*, vol. 7, no. 1, pp. 1–18, Jan. 2008.

[9] C. Dwork, "Differential privacy: A survey of results," in *Proc. Int. Conf. Theory Appl. Models Comput. (TAMC)*. Cham, Switzerland: Springer, 2008, pp. 1–19.

[10] T. AlSkaif, J. L. Crespo-Vazquez, M. Sekuloski, G. van Leeuwen, and J. P. S. Catalão, "Blockchain-based fully peer-to-peer energy trading strategies for residential energy systems," *IEEE Trans. Ind. Informat.*, vol. 18, no. 1, pp. 231–241, Jan. 2022.

[11] Y. Cui, L. Yang, R. Li, and X. Xu, "Online double auction for wireless spectrum allocation with general conflict graph," *IEEE Trans. Veh. Technol.*, vol. 71, no. 11, pp. 12222–12234, Nov. 2022.

[12] Z. Cai, Z. Duan, and W. Li, "Exploiting multi-dimensional task diversity in distributed auctions for mobile crowdsensing," *IEEE Trans. Mobile Comput.*, vol. 20, no. 8, pp. 2576–2591, Aug. 2021.

[13] D. Li, Q. Yang, W. Yu, D. An, Y. Zhang, and W. Zhao, "Towards differential privacy-based online double auction for smart grid," *IEEE Trans. Inf. Forensics Security*, vol. 15, pp. 971–986, 2019.

[14] M. Clark and K. Psounis, "Optimizing primary user privacy in spectrum sharing systems," *IEEE/ACM Trans. Netw.*, vol. 28, no. 2, pp. 533–546, Apr. 2020.

[15] Z. Chen, T. Ni, H. Zhong, S. Zhang, and J. Cui, "Differentially private double spectrum auction with approximate social welfare maximization," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 11, pp. 2805–2818, Nov. 2019.

[16] M. U. Hassan, M. H. Rehmani, and J. Chen, "DEAL: Differentially private auction for blockchain-based microgrids energy trading," *IEEE Trans. Services Comput.*, vol. 13, no. 2, pp. 263–275, Feb. 2019.

[17] G. Gao, M. Xiao, J. Wu, S. Zhang, L. Huang, and G. Xiao, "DPDT: A differentially private crowd-sensed data trading mechanism," *IEEE Internet Things J.*, vol. 7, no. 1, pp. 751–762, Jan. 2020.

[18] Y. Xu, M. Xiao, A. Liu, and J. Wu, "Edge resource prediction and auction for distributed spatial crowdsourcing with differential privacy," *IEEE Internet Things J.*, vol. 9, no. 17, pp. 15554–15569, Sep. 2022.

[19] T. Ni, Z. Chen, L. Chen, S. Zhang, Y. Xu, and H. Zhong, "Differentially private combinatorial cloud auction," *IEEE Trans. Cloud Comput.*, vol. 11, no. 1, pp. 412–425, Jan. 2023.

[20] D. Li, Q. Yang, C. Li, D. An, and Y. Shi, "Bayesian-based inference attack method and individual differential privacy-based auction mechanism for double auction market," *IEEE Trans. Autom. Sci. Eng.*, vol. 20, no. 2, pp. 950–968, Apr. 2023.

[21] E. Erdemir, P. L. Dragotti, and D. Gündüz, "Privacy-aware time-series data sharing with deep reinforcement learning," *IEEE Trans. Inf. Forensics Security*, vol. 16, pp. 389–401, 2021.

[22] A. D. Martinez, J. Del Ser, E. Osaba, and F. Herrera, "Adaptive multifactorial evolutionary optimization for multitask reinforcement learning," *IEEE Trans. Evol. Comput.*, vol. 26, no. 2, pp. 233–247, Apr. 2021.

[23] L. Huang, C. Liu, and Z. Dong, "Deep reinforcement learning based collaborative optimization of communication resource and route for UAV cluster," in *Proc. IEEE Int. Conf. Unmanned Syst. (ICUS)*, Oct. 2021, pp. 69–73.

[24] W. Zhang, B. Jiang, M. Li, and X. Lin, "Privacy-preserving aggregate mobility data release: An information-theoretic deep reinforcement learning approach," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 849–864, 2022.

[25] Y. Jiang, K. Zhang, Y. Qian, and L. Zhou, "Reinforcement-learning-based query optimization in differentially private IoT data publishing," *IEEE Internet Things J.*, vol. 8, no. 14, pp. 11163–11176, Jul. 2021.

[26] M. Shateri, F. Messina, P. Piantanida, and F. Labeau, "Privacy-cost management in smart meters with mutual-information-based reinforcement learning," *IEEE Internet Things J.*, vol. 9, no. 22, pp. 22389–22398, Nov. 2022.

[27] M. Keller, "MP-SPDZ: A versatile framework for multi-party computation," in *Proc. ACM SIGSAC Conf. Comput. Commun. Secur.*, 2020, pp. 1575–1590.

[28] U. Bodkhe et al., "Blockchain for Industry 4.0: A comprehensive review," *IEEE Access*, vol. 8, pp. 79764–79800, 2020.

[29] L. Zhao et al., "InPrivate digging: Enabling tree-based distributed data mining with differential privacy," in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2018, pp. 2087–2095.

[30] Y. Cao, M. Yoshikawa, Y. Xiao, and L. Xiong, “Quantifying differential privacy in continuous data release under temporal correlations,” *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 7, pp. 1281–1295, Jul. 2019.

[31] K. Wang, W. Zhao, J. Cui, Y. Cui, and J. Hu, “A K-anonymous clustering algorithm based on the analytic hierarchy process,” *J. Vis. Commun. Image Represent.*, vol. 59, pp. 76–83, Feb. 2019.

[32] A. Machanavajjhala, D. Kifer, J. Gehrke, and M. Venkitasubramaniam, “*L*-diversity: Privacy beyond *k*-anonymity,” *ACM Trans. Knowl. Discovery Data*, vol. 1, no. 1, pp. 1–52, 2007.

[33] J. Wang, S. Liu, and Y. Li, “A review of differential privacy in individual data release,” *Int. J. Distrib. Sensor Netw.*, vol. 11, no. 10, 2015, Art. no. 259682.

[34] M. Andrés, N. Bordenabe, K. Chatzikokolakis, and C. Palamidessi, “Geo-indistinguishability: Differential privacy for location-based systems,” in *Proc. ACM Conf. Comput. Commun. Secur. (CCS)*, Nov. 2013, pp. 901–914.

[35] D. Ye, S. Shen, T. Zhu, B. Liu, and W. Zhou, “One parameter defense—Defending against data inference attacks via differential privacy,” *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 1466–1480, 2022.

[36] T. Zhang and Q. Zhu, “Dynamic differential privacy for ADMM-based distributed classification learning,” *IEEE Trans. Inf. Forensics Security*, vol. 12, no. 1, pp. 172–187, Jan. 2017.

[37] F. Hu, B. Chen, J. Wang, M. Li, P. Li, and M. Pan, “MastDP: Matching based double auction mechanism for spectrum trading with differential privacy,” in *Proc. IEEE Global Commun. Conf. (GLOBECOM)*, Dec. 2019, pp. 1–6.

[38] J. Zhang and C. Zhong, “Differential privacy-based double auction for data market in blockchain-enhanced Internet of Things,” *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–15, Jun. 2022.

[39] Y. Sun, L. Lampe, and V. W. S. Wong, “EV-assisted battery load hiding: A Markov decision process approach,” in *Proc. IEEE Int. Conf. Smart Grid Commun. (SmartGridComm)*, Nov. 2016, pp. 160–166.

[40] Q. Xiang, L. Kong, X. Liu, J. Xu, and W. Wang, “Auc2Reserve: A differentially private auction for electric vehicle fast charging reservation,” in *Proc. IEEE 22nd Int. Conf. Embedded Real-Time Comput. Syst. Appl. (RTCSA)*, Aug. 2016, pp. 85–94.

[41] R. Zhu and K. G. Shin, “Differentially private and strategy-proof spectrum auction with approximate revenue maximization,” in *Proc. IEEE Conf. Comput. Commun. (INFOCOM)*, Apr. 2015, pp. 918–926.

[42] F. Liu, “Generalized Gaussian mechanism for differential privacy,” *IEEE Trans. Knowl. Data Eng.*, vol. 31, no. 4, pp. 747–756, Apr. 2019.

[43] F. McSherry and K. Talwar, “Mechanism design via differential privacy,” in *Proc. 48th Annu. IEEE Symp. Found. Comput. Sci. (FOCS)*, Oct. 2007, pp. 94–103.

[44] S.-C. Huang, M.-K. Jiau, and K.-H. Chong, “A heuristic multi-objective optimization algorithm for solving the carpool services problem featuring high-occupancy-vehicle itineraries,” *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 8, pp. 2663–2674, Aug. 2018.

[45] S. M. Zandavi, V. Y. Y. Chung, and A. Anaissi, “Stochastic dual simplex algorithm: A novel heuristic optimization algorithm,” *IEEE Trans. Cybern.*, vol. 51, no. 5, pp. 2725–2734, May 2021.

Donghe Li (Member, IEEE) received the B.S. degree in automation and the Ph.D. degree in control science and engineering from Xi’an Jiaotong University, Xi’an, China, in 2015 and 2020, respectively. He is currently an Associate Professor with the School of Automation Science and Engineering, Faculty of Electronic and Information Engineering, Xi’an Jiaotong University. His current research interests include cyber-physical systems, auction mechanisms, energy trading markets, reinforcement learning, and privacy preservation.

Chunlin Hu received the B.S. degree in automation from Jilin University, Changchun, China, in 2022. He is currently pursuing the M.S. degree with the Faculty of Electronics and Information Engineering, Xi’an Jiaotong University, Xi’an, China. His current research interests include smart grid energy trading and attack and privacy security.

Qingyu Yang (Senior Member, IEEE) received the B.S. and M.S. degrees in mechatronics engineering and the Ph.D. degree in control science and engineering from Xi’an Jiaotong University, Xi’an, China, in 1996, 1999, and 2003, respectively. He is currently a Professor with the Faculty of Electronics and Information Engineering, Xi’an Jiaotong University, where he is also with the State Key Laboratory for Manufacturing System Engineering. His current research interests include cyber-physical systems, power grid security and privacy, control and diagnosis of mechatronic systems, and intelligent control of industrial processes.

Yuhao Ma received the B.S. degree in automation from North China Electric Power University, Baoding, China, in 2019. He is currently pursuing the M.S. degree with the Faculty of Electronics and Information Engineering, Xi’an Jiaotong University, Xi’an, China. His current research interests include differential privacy and generative adversarial networks.

Feiye Zhang received the B.S. degree in electronic science and technology from Xi’an Jiaotong University, Xi’an, China, in 2019, where he is currently pursuing the Ph.D. degree with the Department of Automation Science and Technology, School of Electronics and Information Engineering. His current research interests include multi-agent systems, reinforcement learning, and auction mechanisms design for smart grids.

Dou An received the Ph.D. degree in control science and engineering from Xi’an Jiaotong University, Xi’an, China, in 2017. He is currently an Associate Professor with the Department of Automation Science and Engineering, Faculty of Electronics and Information Engineering, Xi’an Jiaotong University. His current research interests include cyber-physical systems, the IoT security and privacy, and incentive mechanism design for smart grids and the IoT.