A Comprehensive Survey on Auction Mechanism Design for Cloud/Edge Resource Management and Pricing

NAFISEH SHARGHIVAND[1], **FARNAZ DERAKHSHAN**[1], **AND NAZLI SIASI**[2], **(Member, IEEE)**

[1]Department of Computer Engineering, University of Tabriz, Tabriz 5166616471, Iran
[2]Department of Physics, Computer Science and Engineering, Christopher Newport University, Newport News, VA 23606, USA

Corresponding author: Nafiseh Sharghivand (n.sharghivand@tabrizu.ac.ir)

ABSTRACT Driven by the vision of cloud and edge computing as the 5th utility, a major challenge is the efficient resource management and pricing taking into account incentives and preferences of cloud/edge providers, users, and the system. In recent years, auction mechanism design has gained wide attention as a tool for addressing this challenge. Here existing surveys lack detailed studies in this area, hence this work presents a comprehensive survey on the state-of-the-art auction-based mechanisms in the field of cloud/edge computing. First, the problem statements and background are presented for cloud/edge computing and auction theory. This is followed by a collaboration model on the benefits of auction mechanism designs to cloud/edge computing. Next, various auction mechanisms in cloud computing are reviewed based upon the following aspects: (1) direction of bids—forward, reverse, and two-sided auctions, (2) the heterogeneity of resources—combinatorial auctions, (3) the number and types of adopted attributes in winner determination—multi-attribute auctions, (4) the setting of auction—offline, online, or sequential auctions, and (5) the interoperability among cloud providers—inter-clouds. For each aspect, existing studies are summarized along with their saliencies and drawbacks. Moreover, existing auction mechanisms in the context of edge computing are also presented. Our comprehensive survey shows that although significant progress has been achieved in this field, there still exist key challenges that need to be investigated, as discussed here to provide future research directions.

INDEX TERMS Auction mechanism design, cloud computing, edge computing, resource management, resource pricing.

I. INTRODUCTION

In the past decade, cloud computing has created a new provisioning paradigm for network operators, organizations, and users by shifting from on-premise hardware towards shared, accessible computing resource pools. However, cloud services are usually accompanied by long propagation delays making them inappropriate for time-sensitive applications. This has consequently given rise to the concept of edge computing which aims to bring computing resources to the edge of the network, i.e., enabling time-sensitive internet of things (IoT) applications.

Cloud/edge infrastructures usually incur high maintenance and operational costs [1]–[6], which requires efficient resource management mechanisms that are deployed in two

The associate editor coordinating the review of this manuscript and approving it for publication was Asad Waqar Malik.

tiers. Physical resources such as CPU, RAM, storage, and bandwidth are efficiently provisioned in the form of different Virtual Machines (VMs) in the first tier. Meanwhile, the provisioned VMs are proficiently allocated to users in the second tier.

Tackling the cloud/edge resource management problem is very challenging due to several reasons such as the large scale of the market, the heterogeneity of the resources, the hierarchy of providers, the interdependency between users' requests, the dynamicity and unpredictability of user demands, the interoperability between providers, the wide range of preferences and incentives of users and providers, etc.

Another significant challenge that is highly affiliated with resource management and often addressed concurrently is resource pricing. In a conventional cloud/edge environment, the providers require satisfactory monetary rewards to

share their resources. Meanwhile, the users may have different budget preferences regarding their requirements. Therefore, a trade exists between the users requesting computing resources with budget limitations and the providers offering computing resources with monetization goals. In this respect, the cloud/edge resource management and pricing problems need to be addressed jointly in order to improve system efficiency and incentivize price-driven self-interested users and providers to use the system.

This problem has subsequently attracted a lot of attention both in academia and in industry. Multiple efforts in literature proposed centralized, or semi-centralized solutions for the problem of resource provisioning and VM allocation and placement in cloud/edge computing [7]–[10]. These solutions mostly leverage optimization techniques which require global knowledge of the market, while assuming providers and users as non-decision makers. However in real world, cloud/edge providers and users are actually rational, self-interested entities, exercising their partial or complete autonomy to achieve their objectives [11].

In order to avoid the aforementioned limitations and to model the strategic behavior of providers and users, another venue of literature have been dedicated to developing game-theoretic solutions [12]–[16]. Although, game theory provides a suitable tool for analyzing the strategies, behaviors, and interactions of players in a multi-agent environment [17], it is not capable of enforcing any desired structures to the system. Hence, auction mechanism design [18] has gained wide attention in order to develop variant solutions which are able to employ the incentives of rational agents to induce the desired behavior. Furthermore, they are capable of providing an efficient framework for the allocation and pricing of commodities via a process called bidding [19]. Finally, they only require a limited global knowledge as they are decentralized in nature.

All the aforementioned advantages have resulted in proposing various auction-based solutions in the recent years to address the problem of cloud/edge resource management and pricing. Especially that the cloud/edge environments can be simply modeled as auction markets, where providers and users can be assumed as the sellers and buyers in the market respectively, and the cloud/edge services can be assumed as the commodities of the market. Therefore, the possible interactions between the providers and the users can be modeled in the auction mechanisms properly.

In this survey, we aim to provide a comprehensive review of the state-of-the-art auction-based solutions in the literature for the problem of cloud/edge resource management and pricing. Furthermore, we provide a concrete discussion on the open challenges of the field. We highlight our contributions in more details in the following section.

A. MOTIVATION AND CONTRIBUTION
Currently, many survey papers are available in the field of cloud/edge computing, focusing on different areas such as, security [20]–[26], energy efficiency [27]–[30], elastic-

ity [31], network virtualization and software defined networking [32], [33], interconnected clouds (inter-clouds) [34], resource management [35], [36], resource scheduling [37], [38], pricing [39]–[42], and other topics [43]–[46].

There are also several other surveys which have discussed auction mechanisms in cloud and edge computing including [41], [47]–[49]. In [47], auction-based cloud resource allocation mechanisms have been surveyed and classified into four categories including one-sided, double-sided, combinatorial, and other types of auction-based mechanisms. However, compared to our work, their study covers a much smaller body of existing literature. They do not address other types of auction mechanisms such as online auctions, sequential auctions, and multi-attribute auctions. Also, they do not discuss the proposed auction mechanisms in the context of inter-clouds and edge computing. In [48] a very concise literature review in the context of cloud computing has been provided, lacking many of the previous studies. Also, the authors do not make any conclusions over existing studies, nor provide any guidelines for future researches. In [41] only the proposed double auction mechanisms in the context of cloud computing have been reviewed and thus the survey does not include other types of auctions or edge computing environments. Finally, the survey on auction mechanisms for edge computing in [49] lacks some key models in cloud computing. However, we believe that due to the many similarities that the problem of resource management and pricing in cloud and edge computing share, having a comprehensive knowledge on existing studies in the area of cloud computing can act as a building block for what can be done for edge computing in the future. So, in contrast to their work, we have also surveyed the existing proposed auction models for cloud computing to provide a comprehensive overview of what is done previously, and what can be done in the future in this field. Moreover, the authors provide different comparison and classification models than our work.

Therefore, to the best of our knowledge, there is no comprehensive survey in the literature discussing different proposed auction mechanisms to deal with resource management and pricing in cloud/edge computing. This motivates us to fill this gap with a comprehensive literature review over stat-of-the-art auction-based mechanisms for cloud/edge resource management and pricing. Overall, the main contributions of our work are as follows:

- We discuss why the auction-based mechanisms are considered as promising solutions for cloud/edge resource management and pricing.
- We provide a comprehensive survey of auction-based mechanism for cloud resource management and pricing, presenting a novel categorization over them based on the direction of bids (single-sided and two-sided auctions), the number and types of attributes used in winner determination (multi-attribute auctions), the setting of auction (online, offline, or sequential auctions), the heterogeneity of cloud resources (combinatorial

auctions), and the interoperability among clouds (inter-clouds).

- In the course of each of the above aspects, we discuss the advantages and disadvantages of existing proposed models.
- We review the existing studies in auction mechanism design for edge computing and discuss their pros and cons.
- We discuss remaining issues in the field of cloud/edge resource management and pricing which can be addressed using auction mechanism design, and we provide directions for future works.

B. THE STRUCTURE OF THE PAPER

The rest of the paper is organized as following. In Sections II and III, we present some fundamental concepts about cloud and edge computing respectively. In Section IV we first provide the basic concepts of auction mechanism design and then discuss how designing efficient auction mechanisms can lead towards market-oriented cloud/edge computing. Through Sections V to VIII, we review different proposed auction models for cloud markets, including single-sided and two-sided auctions, combinatorial auctions, multi-attribute auctions, and sequential and online auctions respectively. Then, we review the proposed auction models for inter-clouds in Section IX. In Section X, we review the proposed auction models for edge resource management and pricing. Finally, we present remaining challenges in the field, and provide directions for future works in section XI, and conclude our survey in Section XII.

II. CLOUD COMPUTING

Cloud computing is defined by National Institute of Standards and Technology (NIST) as a model for enabling ubiquitous, convenient, on-demand network access (usually via Internet) to a shared pool of configurable computing resources [50]. Cloud computing can be categorized based on its service model or deployment model.

The service model typically categorizes cloud computing into three types of cloud services offered at three levels [50]. Software as a Service (SaaS) is provided at the highest level offering varieties of software applications to cloud users. Platform as a Service (PaaS) is provided at the middle level provisioning a computing platform to users, where they can develop and run their applications. Infrastructure as a Service (IaaS) is provided at the lowest level provisioning different fundamental computing resources such as processing, storage, and bandwidth to users. Figure 1a illustrates different cloud computing services offered at different levels. However, over the years, many functions have been transformed into new services (e.g., ML as a service) and are represented by Anything as a Service (XaaS).

On the other hand, the deployment model categorizes cloud computing into four sub-models based on its computing infrastructure [50]. A public cloud is owned by an organization selling cloud services to the general public over a public network (usually via Internet). A private cloud gives an exclusive access to an organization or cloud users for using infrastructure and computational resources. A community cloud serves a group of cloud users who share similar concerns such as mission objectives, security, privacy, and compliance policy. A hybrid cloud employs a combination of two or more distinct clouds (private or public). Figure 1b shows different cloud deployments models.

Another important concept in the context of cloud computing is called Quality of Service (QoS) which is used as a measurement to denote the levels of performance, reliability, and availability offered by cloud services [51], [52]. Obviously, cloud users always prefer to receive higher QoS, but clearly this comes at the expense of higher operational costs for providers. Hence, from the providers' perspective, it is not always economical to satisfy all user expectations. In this respect, an agreement needs to be achieved between cloud users and providers on the different aspects of delivered services which is referred as service level agreement (SLA). An SLA will then serve as the foundation for the expected level of service between the user and the provider [53].
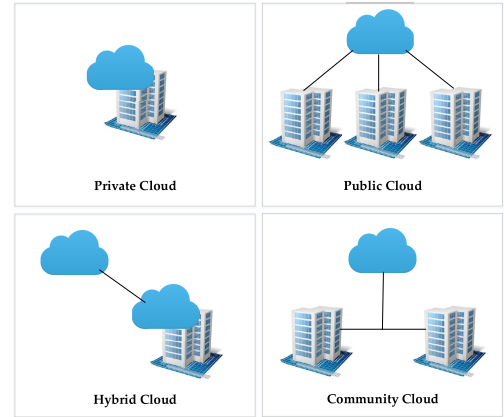
III. EDGE COMPUTING

It is projected that the exponential growth in the number of smart connected (IoT) devices such as smart phones, wearable devices, autonomous vehicles/drones, etc, will reach 75 billion by 2025 [54]. However, due to the limited computational and energy resources of IoT devices, and the data-intensiveness and energy-hungriness of the developed applications, the core services and processing are usually performed on cloud severs. Although leveraging services of distant cloud servers improves the computational capabilities of IoT devices, it results in high latency and mobility-related issues [3, 4]. In fact, for many time-sensitive applications such as online gaming, speech recognition, virtual reality, and real-time traffic monitoring that require fast processing and quick response time, offloading computations to a far centralized cloud data center is considered infeasible.

In this respect, edge computing have been presented in order to address the aforementioned challenges by bringing computing resources to the vicinity of IoT devices. This way, instead of transmitting a large amount of raw data to a distant cloud, the data is processed in micro data centers, placed at the edge of the network to ensure real-time ultra-low-latency processing. These small-sized data centers which are usually referred as "Cloudlets", can thereby mitigate the overload of IoT devices by accepting their offloaded computational tasks.

Other computing paradigms similar to edge computing have also been proposed in the literature, such as Cloudlets, fog computing, and mobile edge computing (MEC) [55]. The concept of MEC was firstly proposed by the European Telecommunications Standard Institute (ETSI) in 2014, where mobile users can utilize the computing services within the Radio Access Network (RAN) [56]. The concept of Fog Computing has been proposed by Cisco, which extends the cloud computing paradigm to the edge of the network,
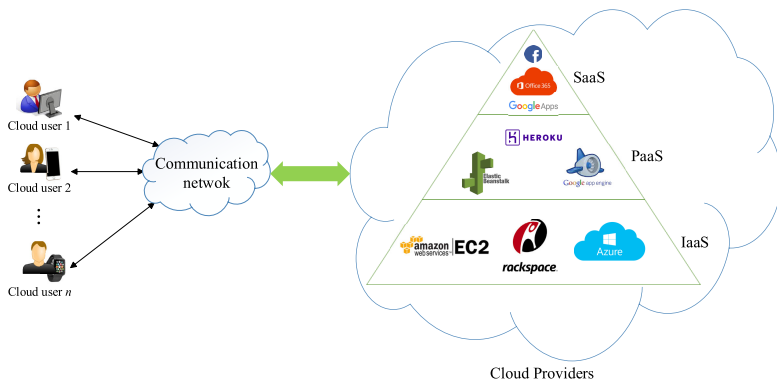
(a) Cloud computing service models

(b) Cloud computing deployment models

FIGURE 1. Cloud computing classification based on service model and deployment model.
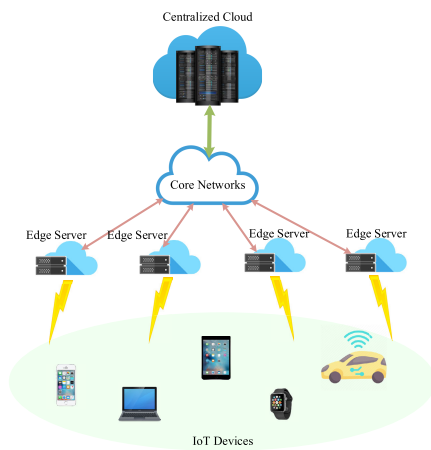

FIGURE 2. The basic architecture of edge computing.

thus allowing data computation, storage, and networking services at reduced latencies [55], [57], [58]. Further, Fog Computing has also been identified as a suitable platform for network function virtualization (VNF) for delay-sensitive applications [59]–[61]. Finally the concept of Cloudlets has been introduced by Satyanarayanan *et al.* [62] which aim at providing computing resources available within the local network.

However, all the three areas are overlapping as they are all characterized by high bandwidth, ultra-low latency, and real-time access to computing services which is achieved by a decentralized model for server distribution. Therefore, the proposed technologies in each case are often applicable to the other two cases [55]. Figure 2 shows the basic architecture of edge computing.

IV. AUCTION MECHANISM DESIGN

In algorithmic game theory, mechanism design provides a suitable platform to attain a social choice in the presence of selfish agents within a competitive environment [63], satisfying desired properties such as high revenue [64]. An auction mechanism on the other hand provides a suitable scheme for allocation and pricing of commodities through a process known as bidding [19]. In this section, we discuss some of the fundamental concepts about auction mechanism design. Furthermore, we argue why auction mechanisms can be considered as promising solution for market-oriented cloud/edge computing.

A. TERMINOLOGIES AND CLASSIFICATION OF AUCTION MECHANISMS

Here, we provide the common terminologies and classification of auction mechanisms. Foremost the commonly used terminologies in auction literature are described as follows:

- ***Bidder*** is the one who submits a request or ask in the auction to acquire or sell commodities, respectively.
- ***Auctioneer*** is in charge of managing and directing the auction process.
- ***Commodity*** is the object being traded in the auction.
- ***Valuation*** is used for evaluating the commodities in the market.

Auctions may be characterized in different ways theoretically and practically [65], but they are commonly classified based on the following features:

- ***single-sided vs two-sided auctions:*** In a single-sided (forward and reverse) auction, only one side (buyers or sellers) can place bids. However, in a two-sided auction both parties can place their bids.
- ***uni-attribute vs multi-attribute:*** In a uni-attribute auction, the allocations are determined only based on one attribute (e.g., bid price). However, in a multi-attribute auction, other attributes (e.g., QoS) are considered as well.
- ***open-cry vs sealed-bid:*** In an open-cry auction each bidder calls out her bid, while in a sealed-bid auction each bid is submitted confidentially.

- *single-item vs multi-item (combinatorial):* In a single-item auction only one type of item is traded in the market, while in a combinatorial auction multiple types of items are traded.
- *single-unit vs multi-unit:* In a single-unit auction only one unit of each item can be traded, whereas in a multi-unit auction multiple units can be traded.

An auction mechanism can be further classified based on the time intervals that the market is cleared:

- *Offline auction:* The market is cleared once all the bids are collected.
- *Sequential auction:* The buyers can submit their bids multiple times and the auctioneer clears the market periodically on specified time frames.
- *Online auction:* The buyers come and go on the fly without any prior notification, and the auctioneer clears the market in an online fashion as soon as a new request is submitted or a new resource becomes available.

B. DESIRED PROPERTIES IN AUCTION MECHANISM DESIGN

The main goal of the auction mechanism design is to construct the rules of an auction such that in the equilibrium of selfish agents' behavior, a desired objective is obtained [66], [67]. Therefore, in an efficiently designed auction mechanism, the desired objectives of an auction will be simply achieved while selfish agents follow their own incentives and preferences. In the following, we describe some of the desired properties in auction mechanism design:

- *Incentive Compatibility (IC)/ Truthfulness (T)/ Strategy Proofness (SP):* Ensures that truthfulness would be within the best interests of participants, and thus they cannot benefit from lying [63].
- *Individual Rationality (IR):* Ensures that each participant can increase his/her utility by using the mechanism [63].
- *Budget Balance (BB):* Implies that the market neither accumulates surplus, nor runs in deficit [63].
- *Allocative Efficiency (AE):* Implies a socially efficient distribution of the resources, as the utility across all participants (social welfare) is maximized [63]. Moreover, if the welfare loss converges to zero as the maximum social welfare approaches infinity, then the mechanism would be called Asymptotically Efficient (AsE) [68].
- *Computational Tractability (CT):* Implies that the auction mechanism clears the market by finding allocation and pricing of commodities within a tractable computation time.

In many cases, some desired properties (e.g., AE) need to be relaxed in order to achieve more preferable ones (e.g., IC). This is because of the impossibility of attaining all properties of IC, IR, BB, and AE by an auction mechanism at the same time [69]. For example, the celebrated Vickrey–Clarke–Groves (VCG) mechanism [70]–[72] is a type of sealed-bid auction in which the items are assigned in a socially optimal manner. This mechanism satisfies IC, IR, and AE, but it is not BB. However, when the classic VCG mechanism is directly extended to other auction variations, such as online auctions, it may loose some of its desirable properties like truthfulness [18]. Therefore, auction mechanisms should be carefully engineered based on the existing challenges in each domain to satisfy required properties. It is worth noting that mechanism designers can seek to optimize several other desirable properties such as following:

- *Revenue Maximization (RM):* Implies to the selfish perspective of the mechanism designer which is to maximize the obtained revenue. This goal turns out to be quite different from the goal pursued by AE [63].
- *Fairness (F):* Implies that the outcome of the mechanism would be fair to all participants. Providing the exact definition of fairness is somehow tricky, but in many cases it is defined as the case where the fairest outcome for a mechanism is the one that makes the least-happy agent the happiest [63].

C. AUCTIONS FOR MARKET-ORIENTED CLOUD/EDGE COMPUTING

Currently, the problems of resource management and pricing are two major challenges in the field of cloud/edge computing. These two problems are assumed tightly coupled as they can directly affect the outcome of each other. For example, the over-charging/under-charging of the resources can result in the under-utilization/over-utilization of the resources and subsequently cause profit loss. As another example, inefficient pricing of the resources does not allow for a suitable distribution of workload over time in order to achieve load balancing. Therefore, the problem of cloud/edge resource management needs to be addressed in conjunction with the pricing problem, which is referred as market-oriented cloud/edge computing [73], [74].

However, the joint problem of cloud/edge resource management and pricing is considered very challenging due to the large scale of cloud market, the heterogeneity of resources, the interdependencies between the requests of the users, the variability and unpredictability of the demands, the possibility of interoperability between the providers, and the vast range of the preferences and incentives of the users and providers. These challenges have given rise to the variety of proposed solutions.

The proposed conventional solutions for the aforementioned challenges are mostly centralized and require global and complete knowledge of the environment. Moreover, they often assume providers and users as non-strategic actors having no interactions with each other. In other words, they do not consider the impacts of the user and provider behaviors on the obtained results for resource allocation and pricing. However, in a real cloud/edge environment, the users and providers act as selfish agents with conflicting goals. For example, on one hand the providers need enough motivations in the form of monetary rewards in order to share their resources. On the

other hand, the users need efficient mechanisms that can guarantee providing their required resources with least price and best QoS. Hence, they will always seek for the best strategies that can optimize their own utilities, e.g. through lying about their true preferences. This way, they can indirectly manipulate the system and impact the overall resource allocation and pricing. Therefore, methods such as machine learning based techniques [75], [76] alone cannot be sufficient and new solutions are needed to be devised in order to address the selfish behaviors of the users and providers and their impact on the cloud/edge resource allocation and pricing [77], [78].

In this regard, mechanism design as a subfield of economics and game theory has attracted a great deal of attention in the recent years, which mainly revolves around studying how to employ the incentives of rational agents to induce the desired behavior for the system. In other words, a well-designed mechanism can give enough incentives to rational agents to act truthfully in order to achieve the desired goals for the system. Furthermore, auction mechanisms have been extensively used in the literature attributed to their decentralized nature, suitability for distributed systems, and their minimal need for global knowledge. This is while a cloud/edge market can be simply modeled as an auction market, where providers and users are considered as the sellers and buyers of the market, respectively. Note here that cloud/edge services can also be assumed as the commodities being traded in the market. Therefore, an auction-based mechanism can be considered as a promising solution to provide a market-driven cloud/edge system.

Figure 3 illustrates some of the advantages of using auction mechanism design for cloud/edge resource management and pricing. Moreover, Table 1 summarizes the mapping of a cloud/edge system to an auction market.

However, there exists several challenges in developing a well-designed auction mechanism for cloud/edge computing. This includes following:

- The auction mechanism should be IC in order to be resistant to the strategic behavior of users and providers so that it can prevent any manipulation of the market.
- The auction mechanism should be IR to incentivize users and providers participation in the auction.
- The auction mechanism should be BB in order to prevent any loss for the broker.
- The auction mechanism should be AE in order to maximize the social welfare.
- The auction mechanism should be multi-attribute in order to consider other important attributes such as QoS metrics beside the pricing of resources.
- The auction mechanism should be combinatorial to offer various resources.
- The auction mechanism should be online so that it can provide real-time responses to service requirements.
- The auction mechanism should be able to model different types of interactions among cloud/edge actors, such

as the interoperability between providers in the form of inter-clouds.
- The auction mechanism should be fair in order to increase satisfaction among the users and the providers.

In this regard, multiple studies have been performed developing different auction-based solutions, each addressing a subset of the above challenges. In this survey, we aim to provide a comprehensive survey of these studies. In this regard, since the number of existing literature in cloud computing is much larger than edge computing, we first review the auction-based solutions in cloud computing from different perspectives. We then go over the proposed auction-based approaches in edge computing. We finally summarize the remaining open challenges in the field.

V. SINGLE-SIDED AND TWO-SIDED CLOUD AUCTIONS

Depending on the direction of offers, auctions can be categorized into two types: single-sided auctions, which are further categorized into forward and reverse auctions, and two-sided auctions, which are also referred as double auctions. In a forward auction, the competition is on the user side, where multiple users can bid for their required resources offered by a single cloud provider (Figure 4a). In a reverse auction, the competition is on the cloud side, where multiple providers can submit asking prices for leasing their resources (Figure 4b). Finally, in a double auction, the competition is on both sides as both users and cloud providers can submit their bid and ask prices, respectively (Figure 4c).

Forward auction models are considered as the most common auction types. For instance, Amazon EC2's Spot market is the best known forward auction based model in real world in which users with bid prices higher than the spot price can acquire their requested resources [79]. However, Amazon spot market has several deficiencies; First, requests can be submitted only in the form of individual VM instances and not a set of different VM types. Second, all winning users are charged equally, under-exploiting the payment potential of users. Third, dishonest users in Amazon may abuse the system by bidding high prices while being charged at lower spot prices [80], resulting in the dissatisfaction of honest users. Finally, according to [81], Amazon's spot price is not likely to be market driven, rather it seems to be generated randomly within a narrow interval close to the hidden reserve price. However, due to the efficiency and simplicity of forward auction based models, they have also been broadly employed in the literature. Some of the major contributions in this field are as following.

Wang *et al.* [80] formulate the cloud service pricing problem as a multi-unit combinatorial auction problem and then propose a suite of computationally efficient and truthful auction mechanisms to solve it. In their proposed solutions, they resort to greedy allocation mechanisms to find an efficient allocation of resources and then employ a critical-value based payment scheme to ensure truthfulness. One of the main contributions of their work is to address the case of
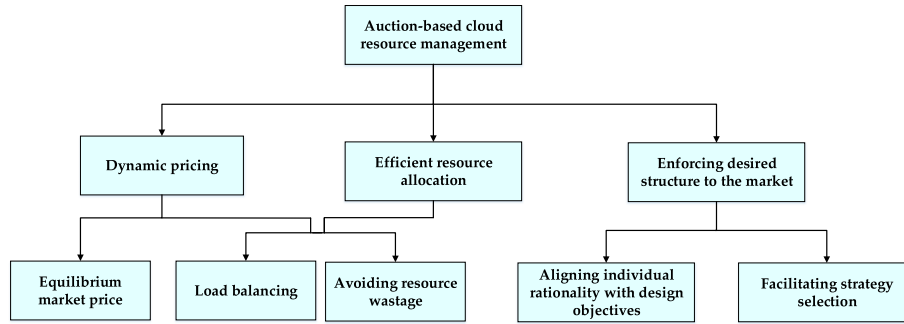
FIGURE 3. Advantages of using auction-based approaches for cloud resource management and pricing.

TABLE 1. Mapping of a cloud system to an auction market.

Auction terminology	Matching entity in cloud
Bidder	Forward auction: Cloud users submit their requests
	Reverse auction: Cloud providers submit their asks
	Two-sided auction: Both cloud users and providers submit their requests and asks respectively
Auctioneer	Forward auction: A trusted third-party or cloud provider
	Reverse auction: A trusted third-party or cloud user
	Two-sided auction: A trusted third-party
Commodity	Cloud services
Valuation	Uni-attribute auction: price
	Multi-attribute auction: price and other attributes such as QoS

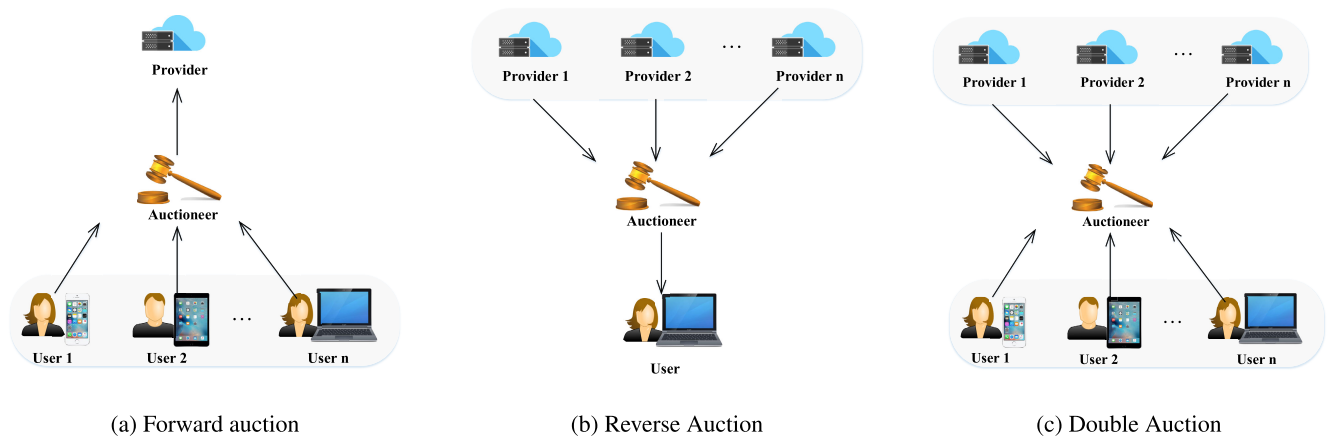(a) Forward auction   (b) Reverse Auction   (c) Double Auction

FIGURE 4. Different cloud markets considering the direction of bids.

collusive bidders by proposing a collusion-resistant cloud resource auction mechanism which can effectively defend against small-sized collusion. They show that the proposed algorithms can increase the revenue of providers, especially when allocating relatively limited resources to a potentially large number of users.

In [82] false-name bids cheating pattern is addressed by proposing a new false-name-proof auction mechanism (FAITH) for VM instance allocation. FAITH not only resists the false-name bid cheating, it also provides truthfulness and polynomial execution time. It first computes the payments for each bidder and then determines the auction winners. Moreover, it supports diverse request formats by

users and allows the auctioneer to configure the quantity of cloud resources for auction to maximize the revenue. FAITH can achieve high auction efficiency in terms of social welfare and revenue.

Huu and Tham [83] propose three truthful combinatorial auction mechanisms, namely exhaustive search algorithm (ESA), linear relaxation based randomized algorithm (LRRA) and green greedy algorithm (GGA). The two first algorithms aim to maximize the total bid values of users minus their energy cost, while the third one is based on the energy consumption of users' resource bundles. They also present a payment scheme corresponding to each algorithm. According to their obtained results, GGA can significantly

reduce energy consumption while increasing the revenue for providers.

Table 2 at the end of this section includes the main features and the pros and cons of the aforementioned works.

As mentioned before, another type of single-sided auctions are reverse auctions (also called procurement auctions [84]). Reverse auctions are a widely used model for procuring resources as they can reduce procurement costs, improve market efficiency [85], and avoid undesirable consequences like nepotism and political ties [86]. Smeltzer and Carr *et al.* [87] also outline several other advantages of using reverse auctions, as well as several required conditions for the efficiency of reverse auction based market. In a reverse auction based cloud market, the cloud user is the one who creates a call for proposal, and thus, the competition is undergoing among providers. We review some of the major studies in this area in the following.

Bonacquisto *et al.* [88] propose a procurement auction mechanism to allocate providers' spare computing resources. They also devise an adaptive bidding strategy to guide providers in the choice of the right actions through the auction. Finally, they employ an overbooking mechanism to overcome the problem of resource under-utilization.

Tanaka and Murakami [89] propose a dynamic programming based algorithm for service selection and VCG payment calculation. They also provide an extended model to reduce computation complexity by recording iterative calculations in the service selection process. This way, they avoid doing similar calculations repetitively which can increase the overall execution time of algorithm.

Prasad and Rao [86] propose three reverse auction based mechanisms namely cloud-dominant strategy incentive compatible (C-DSIC), cloud-Bayesian incentive compatible (C-BIC), and cloud optimal (C-OPT). C-DSIC is a low-bid Vickrey auction and in the case that all providers use the same probability distribution of price and QoS, C-DSIC is preferred over the other two mechanisms. In C-BIC, each provider contributes a participation fee which is used to pay other providers. This way, C-BIC ensures to be budget balanced. C-BIC is suitable for government-sponsored procurement auctions as it leads to less procurement cost for users in comparison with C-DSIC. In both C-DSIC and C-BIC mechanisms, the provider with the lowest cost per unit QoS is declared the winner. C-OPT employs different winner determination and payment rules from C-DSIC and C-BIC. In C-OPT, the cloud vendor with the least virtual cost is declared the winner. Virtual cost is defined as a function of cost and QoS. Then, the payment is computed based on the quoted cost and the expectation of the allocation. All the three proposed mechanisms reduce the resource procurement cost with the increase in number of providers.

Table 2 illustrates the major features, advantages, and disadvantages of the above works.

Single-sided auctions are only well-suited for markets with a limited number of buyers or sellers, and thus they loose their efficiency as the market size grows [68]. As a result,

two-sided auctions have drawn attention among researchers in order to establish a healthier and more efficient cloud market [90]. This is while the nature of cloud market also reveals a two-way trading model. On one side the cloud users compete to obtain their required resources at optimal cost, and on the other side cloud providers compete to serve more users in order to increase their revenue. Moreover, a two-sided auction can prevent monopolies and provide an unbiased, fair cloud market, by considering the preferences of both sides in the outcome of the auction. In this respect, various double auction mechanisms have been proposed to address the problem of cloud resource management and pricing. We review some of the major studies in this area.

Kumar *et al.* [41] propose a double auction model for resource allocation and pricing in cloud computing which is truthful for both user and provider sides. The proposed resource allocation method performs partial allocation similar to cloud spot market. Also, their proposed model does not consider the case of offering different types of VMs in the auction, i.e. combinatorial auctions.

Wang *et al.* [91] propose a fitness-enabled resource allocation model based on the continuous double auction, which guarantees the fitness of allocated cloud resources to cloud services. They also propose a dynamic asking/bidding strategy so that providers and users can autonomously achieve trade-offs between prices, profits, and the QoS. They show that the idea of fitness improves economic efficiency and system performance.

Wang *et al.* [92] propose a combinatorial double auction which enables task partitioning among multiple providers. They also propose a price formation mechanism using a back propagation neural network based price prediction algorithm and a price matching algorithm. A reputation system is also employed to exclude dishonest participants from the market. They use the improved paddy field algorithm to solve the winner determination problem. Although their simulations show that the proposed mechanism can encourage participants to behave truthfully, it has not been theoretically proved.

Kumar *et al.* [93] propose a truthful combinatorial double auction mechanism using linear programming based padded method. For payment, truthful and novel schemes are designed for both customers and providers. The user payments are calculated using critical value schemes, whereas the provider payments are calculated using marginal values of resources offered by various providers.

Lu *et al.* [94] claim that in practice users rarely know the exact number of VMs that meet their task requirements. Hence, they propose a truthful double auction mechanism, including a matching process as well as a VM allocation and pricing scheme. In the matching process, a cost-aware resource algorithm based on Lyapunov optimization techniques is designed to find the number of VMs that meets users' task requirements. For the VM allocation and pricing, the idea of second-price auction is applied to determine the final price and the number of provisioned VMs in the auction.

TABLE 2. Single-sided (forward and reverse) and two-sided auction mechanisms in cloud computing.

Model	Major features	Advantages	Disadvantages
Forward auctions			
[80]	• Proposes a suite of computationally efficient multi-unit combinatorial auction-based mechanisms • Considers collusion among users	• Guarantees truthfulness • Tolerates a collusion group of size t with probability at least p	• Limits the requests of users to one unit per instance in the proposed collusion-resistant auction mechanism • Does not guarantee IR
[82]	• Addresses a new cheating pattern, named false-name bids • Makes allocation decisions based on payment decisions	• Ensures IC and false-name-proof • Supports diverse request formats by users • Provides high auction efficiency in terms of social welfare and revenue • Allows for dynamic provisioning of resources	• Lacks IR and BB guarantee
[83]	• Proposes three different combinatorial auction mechanisms, while addressing green cloud computing	• Ensures truthfulness • Reduces energy consumption of data centers	• Lacks IR and BB guarantee
Reverse auctions			
[88]	• Proposes a procurement auction mechanism to allocate providers' spare resources • Devises an adaptive bidding strategy for providers	• Employs resource overbooking mechanism to improve resource utilization	• Lacks IC, IR, and BB guarantee • Lacks any guarantees for the social welfare nor the obtained utility
[89]	• Considers offering SaaS cloud services • Considers the minimum quality requested by the user for the composite service in the winner determination	• Ensures IC • Provides a quasi-polynomial time while giving an exact solution	• Does not guarantee IR • Does not provide a clear definition of the quality paarmeter which has been used in the proposed model
[86]	• Proposes three reverse auction based mechanisms, namely C-DSIC, C-BIC, and C-OPT • Selects the provider with lowest cost per unit QoS as the winner in the C-DSIC and C-BIC , and the provider with the lowest virtual cost as the winner in the C-OPT	• C-DSIC ensures IC, IR, and AE • C-BIC ensures Bayesian IC, BB and AE • C-OPT ensures Bayesian IC, IR, and optimality • Considers QoS guarantees in winner determination	• Is only suitable for scenarios with single resource requirement at a time • Limits the providers in C-DSIC and C-BIC to have symmetric price distributions • C-DSIC is prone to bidder collusion and is not BB • C-BIC is not BB and makes the losing providers to suffer a loss
Two-sided auctions			
[41]	• Considers a cloud environment with homogeneous VMs • Allows for multi-unit VM allocation	• Ensures the properties of IC, IR, BB, and AsE	• May result in the partial fulfilment of requested bundles • Lacks different VM types • May terminate allocated instances at any time due to price variations
[91]	• Employs the idea of continuous double auction to propose a novel fitness-enabled auction	• Guarantees the fitness of performance traits between cloud resources (sellers) and cloud services (buyers)	• Lacks IC, IR, and BB guarantee • Presents a centralized system architecture which may put the auctioneer node in face of bottleneck problem
[92]	• Proposes a combinatorial double auction based on a bio-inspired algorithm, which enables task partitioning among multiple providers • Proposes a price formation mechanism using a back propagation neural network	• Devises a novel reputation system to suppress dishonest participants • Considers several types of services	• Lacks guarantee IC, IR, and BB guarantee • Lacks any guarantee for economic efficiency
[93]	• Proposes a combinatorial double auction mechanism based on the padding method in which a virtual user is added to the system	• Ensures IC, IR, BB, and AsE • Allows each provider to serve multiple users at the same time	• Eliminates several possible allocations due to the adding of virtual padding user, which decreases the allocative efficiency
[94]	• Proposes a double auction mechanism to bridge the task requirements of users and the resources (VMs) of providers • Designs a cost-aware resource provisioning algorithm for determining the number of VMs that meets users' task requirements	• Ensures IC, IR, and BB • Improves user satisfaction by allocating a suitable number of VMs to each user based on his task requirements	• Lacks guarantees for the obtained social welfare
[95]	• Proposes a secure and fair double auction framework for cloud virtual machines • Employs a three-party computation model to provide security	• Achieves bid-privacy protection and trade fairness	• Does not guarantee IC, IR, and BB

Cheng *et al.* [95] propose a secure double Auction framework which can guarantee both bid privacy and trade fairness. In this respect, they propose secure 3-party computation based protocols that support secure comparison and secure sorting. As for fairness, they present a fair and efficient trading scheme based on smart contracts to incentivize each auction participant to follow the auction protocols.

Table 2 includes the major studies in this area, highlighting their advantages and disadvantages.

Summary and Discussion. This section discusses different auction models based on the direction of bids in the cloud auction market, i.e. forward, reverse, and two-sided auctions. Currently, forward auctions are the most common approach in the literature, which have been adopted in the industry such as in Amazon Spot Market [79] as well. In contrast, reverse auctions are not used very often. This is maybe because in the cloud environments the number of users is very larger compared to the number of clouds. Therefore, holding a distinct auction for each individual user does not seem as an effective approach. In this respect, two-sided auctions show a great potential for future cloud markets as they consider the incentives and preferences of both sides, creating a healthier cloud market as a result. However, most of the existing two-sided auctions demonstrate a centralized architecture which can be considered as a potential risk for single point of failure. Because, if that one existing broker fails for any reason, the whole system will become unavailable. Hence, this issue needs to be further investigated in the literature.

VI. COMBINATORIAL CLOUD AUCTIONS

Given the inherent heterogeneity in real world computing tasks, cloud users often demand for bundles of heterogeneous VM instances to satisfy their requirements. For instance, a social game application consisting of a front-end web server layer, a load balancing layer, and a back-end data storage layer, is better served by a bundle of different VMs intended for communication-intensive, computation-intensive, and storage-intensive tasks, respectively [96]. Furthermore, in many scenarios, the value that a user derives from a cloud service highly depends on what other services she can obtain. Therefore, the capability of users to bid on combinations of cloud services in such scenarios can increase the efficiency of the market. As a result, combinatorial auctions reveal more consistency with real world scenarios as they can enable expressive bids for requesting bundles of different VM types. They can be considered as a drastic departure from existing homogeneous auction designs which assume VMs as type-oblivious commodities, i.e., VMs that are essentially of the same type, or interchangeable up to a multiplicative factor.

Obviously, bringing the combinatorial nature of cloud into auction design adds to the complexity of winner and payment determination problem due to the higher dimensions of the problem. In such setting, the celebrated VCG mechanism is also inapplicable due to computational complexity [70]. As a result, many new combinatorial auction based approaches have been proposed in the recent years which provide computational efficiency, as well as satisfying other desirable properties such as truthfulness.

Chichin *et al.* [97] propose a combinatorial greedy allocation mechanism based on the sorted order of allocation candidates to solve the winner determination problem in a two-sided market. They further employ critical-value scheme to compute user prices, and a mixed surplus-distribution rule that relies on direct and proportional-value payment to derive provider payments. They assume that each requested VM bundle must be fulfilled by a single provider. This however can result in an inefficient resource management in some cases. Because, even though a single provider may be able to serve the user request, the user's task requirements may be fulfilled more efficiently through task partitioning between multiple providers. Also, there may be cases that no single provider is capable of providing required resources all by itself. Thus, in these cases there would be no choice other than task partitioning between several providers.

Yang *et al.* [98] propose a combinatorial auction mechanism in a single provider scenario. aiming to maximize the cloud providers' revenue, and achieve envy-freeness and truthfulness. The proposed mechanism combines two general ideas of consensus estimate and RevenueExtraction to respectively achieve an approximate optimal target revenue without market manipulation, and to determine the winners and equally share the generated target revenue among them.

Tafsiri and Yousefi [99] formulate the combinatorial double auction-based market for the cloud VM allocation to users as an integer linear programming, aiming to maximize the total profit of users and providers. Then, they employ a heuristic resource allocation algorithm to overcome the time complexity of their proposed model.

Singhal and Singhal [100] propose a feedback-based combinatorial double auction mechanism in which genuine providers with good feedback are prioritized over in-genuine providers with bad feedback from users. The proposed model also penalizes market spoilers who aim to overcome the market by manipulating prices so that their competitors loose their interest in the market.

However, beside the efficient solutions for VM allocation similar to above studies, cloud providers also need to devise efficient approaches for the provisioning of different computing resources, such as CPU, RAM and Disk, into various VM types. Currently, most of the studies employ Static Resource Provisioning (SRP) approaches in which VMs are assembled before the auction starts. They usually employ simple heuristics or historical demand patterns as a guide for an efficient provisioning of resources (Figure 5a). However, estimation approaches are often prone to errors, even when implemented at their best cases. Thereby, a more efficient approach is Dynamic Resource Provisioning (DRP), which assembles the VM instances dynamically during the auction, tailored to the real-time user demands (Figure 5b). As a result, DRP approaches can improve the resource utilization, the revenue of providers, and the social welfare of the entire
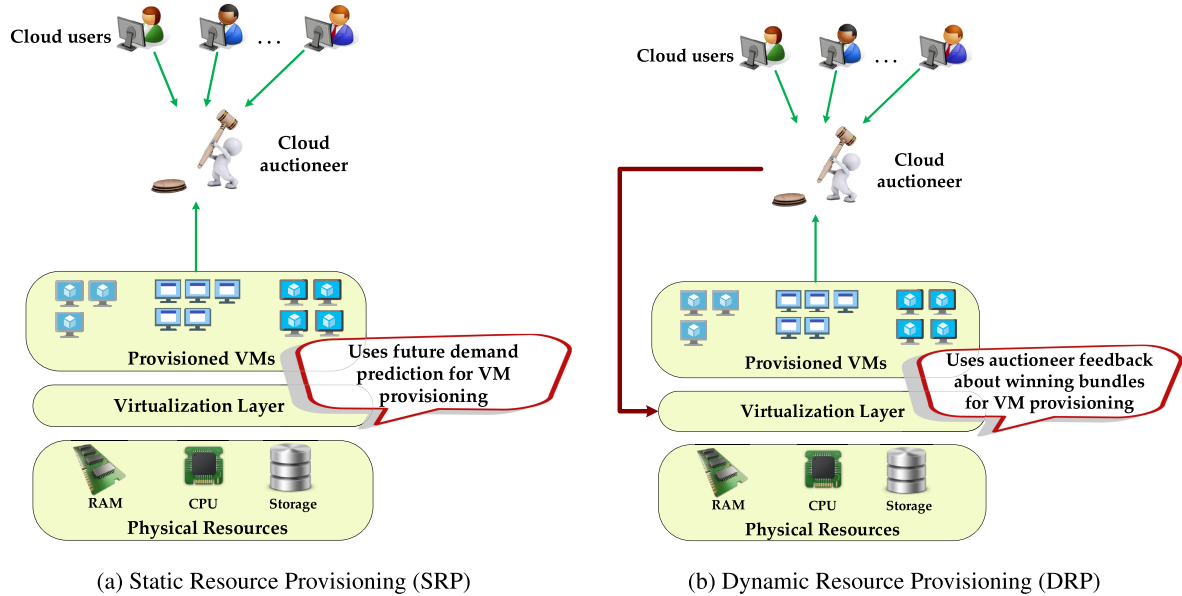
(a) Static Resource Provisioning (SRP)

(b) Dynamic Resource Provisioning (DRP)

FIGURE 5. Different VM provisioning approaches.

(a) Static Resource Assembling (SRA)
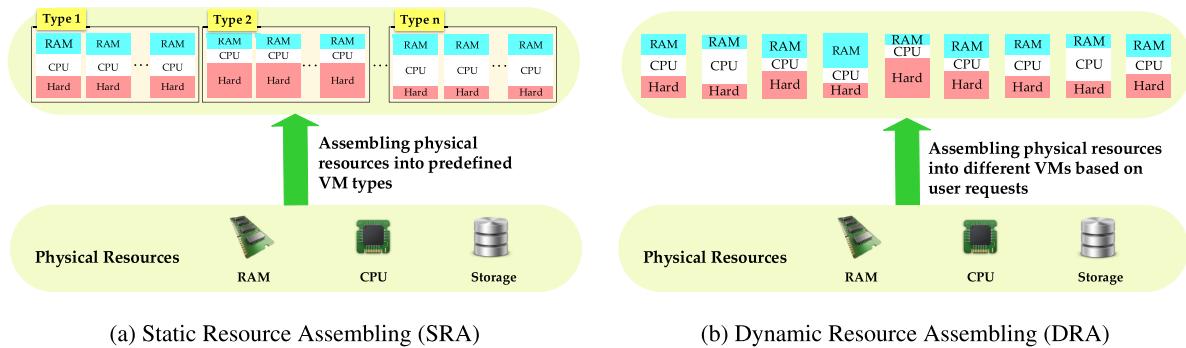
(b) Dynamic Resource Assembling (DRA)

FIGURE 6. Different VM assembling approaches.

cloud system. There are several studies in the literature that have adopted DRP approaches.

Mashayekhy *et al.* [101] propose a strategy-proof polynomial time approximation scheme (PTAS) mechanism to solve the VM provisioning and allocation problem in the presence of a single cloud provider. They show that the proposed approximation mechanism is by far the strongest approximation result that can be achieved for this problem, unless $P = NP$.

Nejad *et al.* [102] propose a family of truthful greedy mechanisms for VM provisioning and allocation such that the cloud provider provisions VMs based on the requests of the winning users and determines their payments.

Mashayekhy *et al.* [103] consider the physical machine (PM) resource management problem in the presence of multiple PMs and multiple types of resources in an auction-based setting. They propose a strategy-proof greedy mechanism called G-PMRM. In order to guarantee strategy-proofness of

G-PMRM, the winner determination algorithm is designed such that it determines loser-independent allocations on each PM. This can make the mechanism robust against users who tried to manipulate the system by changing the allocations of other users.

Zhang *et al.* [104] first propose a cooperative primal dual approximation algorithm and then employ it as a building block to design a randomized combinatorial auction for dynamic resource provisioning. The proposed mechanism is computationally efficient, truthful in expectation, and achieves the same approximation ratio in social welfare as the cooperative algorithm does.

Li *et al.* [105] considers time-varying user demands and propose a truthful online auction mechanism for maximizing the profit of the cloud provider. The proposed mechanism consists of a price-based allocation rule and a payment rule. The allocation rule determines the number of VMs of each type that should be allocated to each user. The payment rule

determines how much the cloud provider should charge each user based on a marginal price function for each type of VMs.

Although the aforementioned studies have taken major steps towards efficient resource management, effective solutions are also required for defining VM types at the lowest level. In this regard, two different approaches are usually adopted. In the first approach, similar to the above studies, the computing resources are assembled into predefined VM types (Figure 6a). This approach which we refer as Static Resource Assembling (SRA), is a very common technique both in the industry and academia. However, it enforces users to choose their required resources from a limited list of predefined VM types with specified configurations. These VM types are actually defined by the providers beforehand. In contrast, in the second approach, the VM types are defined based on user demands (Figure 6b). In other words, this approach which we call Dynamic Resource Assembling (DRA), assembles VMs based on the computing requirements of users. Currently, a few cloud companies employ DRA strategies in order to create VM instances based on user-specified demands, such as [106]. The research community has also proposed several combinatorial auctions based on DRA such as the following.

Zhang *et al.* [78] propose an efficient randomized auction mechanism based on a novel application of smoothed analysis and randomized reduction for the provisioning and pricing of customized VM instances in a geo-distributed cloud.

In [107] a more realistic case of online VM auction design is considered, where cloud users can bid for customized VMs with desired duration of occupation. Furthermore, truthful and polynomial time auctions are proposed to achieve social welfare and/or the provider's profit maximization with acceptable competitive ratios.

Zhang *et al.*s [108] propose an online auction mechanism, where they consider both user evaluation and cost in their model. In the proposed mechanism, users can submit multiple requirements at any time, but only one requirement is satisfied.

Table 3 illustrates the major proposed combinatorial auctions, including those ones that have employed DRP and DRA strategies.

Summary and Discussion. This section discusses proposed combinatorial auction models in cloud computing which allows for offering and requesting a variety of cloud resources. It also discusses different resource provisioning techniques, i.e. Static Resource Provisioning (SRP) and Dynamic Resource Provisioning (DRP), and explains how DRP strategies improve resource utilization, the revenue of providers, and the total social welfare compared to SRP techniques. Moreover, another line of research in combinatorial auctions is the dynamic assembling of resources into different VM types based on users' requirements. This approach can highly improve user satisfaction, as the VMs are exactly assembled based on the user requirements. However, the number of existing studies in this area is quite rare and thereby more research is needed in this venue.

VII. MULTI-ATTRIBUTE CLOUD AUCTIONS

As the ever-growing cloud market approaches perfect competition, it enables cloud users to choose their required services from a variety of cloud providers. As a result, the perceived QoS is becoming an important differentiator, other than price, between cloud service providers [109], [110]. Therefore, if a cloud provider compromises its offered QoS substantially or does not commit to the quoted QoS, the unsatisfied cloud users may refuse to accept the service and decide to choose another cloud provider. In contrast, meeting or exceeding QoS expectations of users will improve the cloud provider's reputation, and will subsequently enhance its resource utilization and obtained revenue.

However, designing multi-attribute auction mechanisms pursuing other desired properties such as truthfulness, individual rationality, and so on, is very challenging due to the dependence of winner selection on many attributes. Moreover, a good understanding of the types of attributes and their measuring and comparison approaches is required. For example, QoS attributes can be divided into two categories of quantitative and qualitative attributes. Qualitative attributes are those metrics which cannot be quantified and are mostly inferred by user experiences. Whereas, quantitative attributes are those which can be measured using software and hardware monitoring tools and can be declared in numeric values [111]. For instance, provider's ethicality and security are qualitative attributes, whereas service response time, availability, and reliability are quantitative attributes.

Hence, qualitative attributes are first required to be converted to a numeric value in order to be measurable and comparable. The Analytic Hierarchy Process (AHP) [112] is a common technique for quantifying QoS attributes which has been also used by [86] and [113]. Moreover, inspired from [114], [115], the authors in [86] and [113] have employed the Simple Additive Weighting (SAW) approach for scaling and normalizing of QoS attributes, and then for comparison of them. However, the SAW method only supports quantitative QoS attributes and thus all attributes need to be first quantified, similar to what has been done in [86] and [113].

In general, two different approaches have been adopted towards integrating QoS attributes into the field of auction mechanism design in cloud computing. In the first approach, it is assumed that selfishness is limited to the pricing aspect, and thereby the cloud providers act truthfully in technical matters. Conversely, in the second approach, it is assumed that the providers may misreport their QoS in order to achieve more benefit.

In [86] and [113], the rationality of providers is limited to pricing aspects. In [86], three different reverse auction based mechanisms are proposed, each satisfying different properties. All the three proposed mechanisms can reduce the resource procurement cost as the number of providers grows. The authors extend their work in [113] by considering

TABLE 3. Combinatorial auction mechanisms in cloud computing.

Model	Major features	Advantages	Disadvantages
		Simple combinatorial auctions	
[97]	• Proposes a combinatorial greedy allocation mechanism based on a two-sided cloud market • Assumes that the whole requested VM bundle must be delivered by a single cloud provider	• Ensures IR and BB • Ensures truthfulness for the users	• Lacks guarantee of truthfulness for the providers • Never allows for task partitioning between multiple providers
[98]	• Employs the idea of the consensus auction and cost sharing mechanism • Addresses envy-freeness in a single provider scenario	• Ensures IC and IR • Achieves an approximate optimal revenue for the cloud provider	• Lacks BB guarantee • Employs single sale price strategy which under exploits the payment potential of users
[99]	• Proposes a double auction mechanism which aims at maximizing the total profit of users and providers	• Ensures truthfulness	• Exposed to single point of failure due to employing a single broker • lacks IR and BB guarantee
[100]	• Proposes a feedback-based combinatorial double auction mechanism • Prioritizes genuine providers with good feedback over in-genuine providers with bad feedback	• Prevents market spoilers to manipulate prices by penalizing them	• Does not ensure BB and IR • Does not ensure truthfulness completely
		Combinatorial auctions with DRP	
[101]	• Proposes a strategy-proof polynomial-time approximation Scheme • Considers the heterogeneity of the system and user requests	• Guarantees a near optimal solution • Ensures truthfulness	• The running time can be exponential in $1/\epsilon$, where ϵ denotes the approximation error • Lacks IR guarantee
[102]	• Proposes a family of truthful greedy mechanisms • Employs a forward auction based mechanism	• Ensures truthfulness • Ensures approximate efficiency	• Lacks IR guarantee • Runs the proposed mechanisms periodically, making them unsuitable for online responding to user requests
[103]	• Considers the physical machine resource management problem in a forward auction-based setting • Addresses green cloud computing strategies	• Ensures IC and IR • Ensures an approximation ratio • Satisfies the loser-independent property • Improves energy savings and reduce costs through the efficient placement of VMs over physical machines	• Lacks BB guarantee
[104]	• Proposes a randomized single-round auction, employing the cooperative approximation algorithm as a building block • Adopts the XOR bidding language, in which a user can win at most one bid even if it submits multiple bids	• Ensures truthfulness • Allows for multiple bid submission by users	• Lacks IR and BB guarantee
[105]	• Proposes a novel online auction mechanism aiming to maximize the profit of the cloud provider	• Ensures IC and IR • Considers online VM auction design • Allows for time-varying user demands	• Lacks BB guarantee • It Lacks multiple provider scenarios
		Combinatorial auctions with DRA	
[78]	• Proposes a randomized auction mechanism for the provisioning and pricing of customized VM instances	• Ensure truthfulness • Achieves $(1 - \epsilon)$-optimal social welfare, where ϵ can be arbitrarily close to 0	• Does not guarantee IR and BB
[107]	• Proposes online auctions for social welfare maximization or profit maximization • Considers the resource composition variation of each VM over its duration	• Ensures IC and IR • Considers online VM auction design • Enables each VM to consume a different amount of each type of resource over time	• Does not guarantee BB • Does not include multiple provider scenarios
[108]	• Proposes an online auction mechanism based on user evaluation and cost • Adopts an XOR bidding language	• Ensures truthfulness • Improves the social welfare and revenue for resource providers by allowing users to submit multiple requirements • Allows users to submit their requests at any time	• Does not guarantee IR and BB • Does not consider multiple provider scenarios

different types of resources. They propose combinatorial auctions branch on bids focusing on time complexity rather than incentive compatibility.

However, as we mentioned, there are a few other studies that have considered the strategic behavior of providers towards their QoS commitments. Therefore, in order to

prevent any manipulations of the market by malicious providers, they have designed reputation-, priority-, and penalty-based mechanisms to decrease untruthful behaviors by cloud providers in terms of QoS. This way, a more healthy competition is encouraged among the providers.

In this respect, Baranwal and Vidyarthi [116] propose a multi-attribute combinatorial double auction for winner determination considering price and non non-price attributes for winner determination. In the proposed mechanism, a penalty is imposed on the providers which do not offer the agreed upon services. Although this approach can decrease the utility of dishonest providers, it is not guaranteed that the imposed penalty could be sufficient in order to prevent any untruthful behavior. In other words, the proposed solution only decreases the probability of untruthful behavior, but it does not provide any guarantees. Another advantage of the proposed work is that it offers compensation to the user who cannot receive the agreed upon QoS. The authors also employ a similar strategy to prevent QoS commitment violation by providers in [117], where their proposed mechanism is based on a reverse auction model. The first work does not ensure IC and IR properties, whereas the latter one does.

Table 4 illustrates major proposed multi-attribute auctions in the literature.

Summary and Discussion. This section discusses multi-attribute auctions for the problem of cloud resource management and pricing. In these auctions, other attributes in addition to price, such as QoS attributes, are considered in the winner determination function. The multi-attribute auctions may limit the strategic behavior of cloud providers to the pricing aspect or generalize it to a more realistic case, where providers may also violate their QoS commitments in order to gain more profit. In this respect, some studies have designed reputation-, priority-, and penalty-based mechanisms to prevent market manipulations by those providers which may lie about their QoS guarantees. However, these proposed approaches do not ensure truthfulness completely. In other words, the aforementioned methods mainly focus on reducing the probability of untruthful behaviors by cloud providers in terms of the QoS commitments. Hence, more studies are required to design multi-attribute auction mechanisms which can guarantee truthfulness in all attribute dimensions.

VIII. ONLINE AND SEQUENTIAL CLOUD AUCTIONS

Auctions may vary depending on how winners and payments are determined over time. Here an auction can be designed in either offline (single-round auctions), sequential, or online setting.

In an offline auction mechanism, it is assumed that all users submit their bids and then a one-time decision is made to clear the market (Figure 7a). So far, different offline auction mechanisms have been proposed in the literature, such as [80], [101], [102], [104], [118]–[120], to address the problem of cloud resource management and pricing. However, in a real world cloud market, the cloud servers constantly fluctuate between busy and idle states while the users come and go

on the fly. Moreover, many users may lose their interest in obtaining their required resources over time, whereas an offline auction ignores such temporal correlation in decision making [104]. Such deficiencies in offline auction mechanisms have thereby motivated researches to propose different sequential and online auction mechanisms.

In a sequential auction, users submit their bids multiple times over time in the auction. If a user does not win in the current round of the auction, then the user can submit another bid in the subsequent round (Figure 7b). One simple approach to design sequential auctions is to hold a series of single-round auctions over time. However, single-round auctions usually lose their truthfulness when extended to sequential auctions [18]. This problem resulted in the development of multiple approximation methods.

Wang *et al.* [121] propose dynamic auctions where a sequence of auctions are periodically carried out to accommodate demands. Also, a near optimal capacity allocation scheme is presented to determine the number of instances that are auctioned off in each period.

Zaman and Grosu [122] propose an auction mechanism for dynamic VM provisioning and allocation, where the provider runs an auction periodically (e.g., once an hour) to provision and allocate the VM instances such that its profit is maximized. The users bid for obtaining their requested bundles for one unit of time and if they require a bundle for more than that, they have to bid again in the next round. Thus, this mechanism is not appropriate for uninterruptible jobs as users may not be able o acquire the required bundle in the next round.

Inspired by the group-buying websites, Sun *et al.* [123] propose a combinatorial double auction in which the allocation and payment schemes are designed with respect to group-buying idea. In this respect, cloud providers are allowed to submit group-buying rules along with their bids. In the proposed mechanism, the central auctioneer is responsible to form buying groups such that the total utility is maximized. Similar to the above works, the proposed auction mechanism takes an iterative manner, i.e. the auction is run in sequential rounds.

Table 5 includes the major sequential auctions which are proposed for the problem of cloud resource management and pricing. However, sequential auctions also suffer from two main shortcomings. First, they do not provide real-time allocation and payment outcomes. Second, they are not suitable for user applications which require uninterruptible computing resources. Because in the case that users need their requested bundles for more than one unit of time, they may have to bid on each round. However, there is no guarantee that they can win in the upcoming rounds as well. Therefore, possible interruptions seem irresistible using some type of sequential auctions.

As a result, various online auction mechanisms have been proposed to address the above shortcomings, in which the allocation and payment decisions are made in real-time without any prior knowledge about the future. Indeed, each transaction takes place as soon as a new user request

TABLE 4. Multi-attribute auction mechanisms in cloud computing.

Model	Major features	Advantages	Disadvantages
[117]	• Proposes a multi-attribute combinatorial reverse auction • Considers the strategic behavior of providers towards their QoS commitments • Penalizes cheating providers and decreases their reputation, meanwhile it offers compensation to the affected users	• Ensures IC, IR, BB, and an approximation ratio • Provides fairness in the auction • Decreases untruthful behavior of providers towards their QoS commitments • Offers compensation to the users who do not receive the agreed upon QoS	• Does not guarantee the truthfulness of providers towards their QoS commitments
[116]	• Proposes a fair multi-attribute combinatorial double auction • Considers the strategic behavior of providers towards their QoS commitments • Penalizes cheating providers and decreases their reputation, meanwhile it offers compensation to the affected users	• Ensures BB • Provides fairness in the auction • Decreases untruthful behavior of providers towards their QoS commitments • Offers compensation to the users who do not receive the agreed upon QoS	• Does not guarantee IC and IR • Does not guarantee the truthfulness of providers towards their QoS commitments
[86]	• Proposes three reverse auction based mechanisms, namely C-DSIC, C-BIC, and C-OPT • Selects the provider with lowest cost per unit QoS as the winner in the C-DSIC and C-BIC , and the provider with the lowest virtual cost as the winner in the C-OPT	• C-DSIC ensures IC, IR, and AE • C-BIC ensures Bayesian IC, BB and AE • C-OPT ensures Bayesian IC, IR, and optimality	• Is only suitable for scenarios with single resource requirement at a time • Limits the providers in C-DSIC and C-BIC to have symmetric price distributions • C-DSIC is prone to bidder collusion and is not BB • C-BIC is not BB and makes the losing providers to suffer a loss • Limits the rationality of providers to pricing aspects
[113]	• Uses Combinatorial Auctions Branch on Bids • Employs a reverse auction model • Is an extension of [86] to a combinatorial model	• Provides a highly scalable winner determination solution	• Does not guarantee IC, IR, and BB • Limits the rationality of providers to pricing aspects

arrives or some cloud resources become available. Figure 7c shows that in an online auction users can submit their bids any time they want. Therefore, online auctions can better capture the existing dynamism in the cloud market compared to offline and sequential auction mechanisms. Because, firstly they reflect the fluctuation of supply-demand in real-time. Secondly, they allow cloud users to ask for resources whenever they need them, while their requests are processed immediately. Thirdly, they do not need to have any information or make any predictions about future.

However, the lack of future information in an online auction mechanism design brings a key challenge in pursuing truthfulness. For example, the direct extension of the classic VCG mechanism to an online auction breaks its truthfulness property [18]. Since, even in the presence of unlimited computational resources, the optimal allocation for the future cannot be calculated. As a result, various online auction mechanisms have been proposed in the recent years to overcome the above challenges.

Mashayekhy *et al.* [124] propose an online mechanism for VM provisioning, allocation, and pricing in clouds, considering several types of resources. The auction is invoked as soon as a user places a request or some of the allocated resources become available. Once a user wins in the auction, the mechanism allocates his required VM instances to him for the entire requested period.

Shi *et al.* [125] propose an online combinatorial auction which also models dynamic provisioning of heterogeneous VM types. In the proposed mechanism, first, a primal-dual algorithm is designed to decompose the online social welfare optimization problem into a series of one-round optimizations. Then, a randomized subframework is adopted that applies primal-dual optimization to translate a centralized cooperative social welfare approximation algorithm into an auction mechanism. Finally, a greedy primal-dual algorithm is applied that approximates the one-round social welfare optimization.

In [126], an online auction mechanism in a cloud spot market is proposed considering multi-unit of a single VM type. The mechanism is envy-free and has a high probability of being truthful. Along with the proposed auction, a new scheme is adopted for dynamically calculating reserve prices based on data center power usage effectiveness and electricity costs.

Zhang *et al.* [127] first design a novel bidding language to express users' heterogeneous requirement on their desired allocation time, application type, and valuation. Then, they propose a truthful online auction mechanism on top of the designed bidding language. Their proposed mechanism consists of two components. First, a payment-function-based payment rule which is uniquely determined by the allocation result and the request submission time. Second, an allocation rule that aims to maximize bidders' utility.
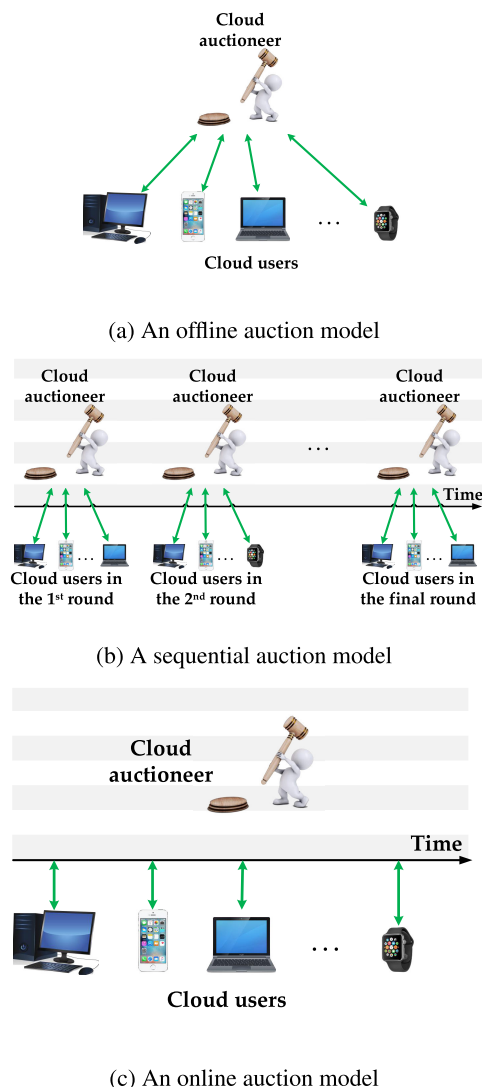
Cloud
auctioneer

Cloud users

(a) An offline auction model

Cloud auctioneer   Cloud auctioneer   Cloud auctioneer

...

Time

Cloud users in
the 1st round

Cloud users in
the 2nd round

Cloud users in
the final round

(b) A sequential auction model

Cloud
auctioneer

Time

Cloud users

(c) An online auction model

FIGURE 7. Different auction clearance settings.

Patel *et al.* [128] propose an online double auction for balancing the multi-objective trade-offs between energy, revenue, and performance in IaaS clouds. In this respect, they employ a weighted bipartite matching based algorithm for winner determination, and a VCG driven algorithm for payment determination.

Table 5 consists of the major features, and pros and cons of the above studies.

Summary and Discussion. This section discusses different proposed auction models in cloud computing based on market clearance intervals, which are offline, online, and sequential auctions. Offline auctions are simpler to design, but they make a one-time decision assuming that all the information is available beforehand which is unrealistic. Sequential auctions clear the market periodically at specific time intervals, yet, similarly they are not capable of making real-time decisions. Moreover, some of the proposed sequential auctions allocate

the resources to winning users only for one interval of time, and thus do not consider the time preferences of users. This clearly makes sequential auctions inappropriate for uninterruptible user applications. As a result, online auctions are preferred over offline and sequential auctions in general. Because, online auctions are in more consistency with real cloud markets in which the users can come and go on the fly over time. Moreover, they are able to respond user requests instantaneously. They also allocate VM instances to users as long as they require them. In other words, they avoid undesirable service interruption, unless the users ask for VM termination.

IX. AUCTION-BASED APPROACHES FOR INTER-CLOUDS

With the rapid growth in the amount and variety of computing resources required by data-intensive applications, it is expected that in the near future the cloud providers will not be able to fulfill many user demands individually anymore. However, rejecting the user requests due to the insufficiency of available resources clearly leads to both profit and reputation loss for the provider. Moreover, using a single cloud as the only service provider may lead to issues such as vendor lock-in and single point of failure [129].

Such emerging challenges have thereby arisen the idea of reshaping the business structure of cloud providers from single clouds to a network of interconnected clouds, termed as inter-clouds [130]. In other words, the expected future for the cloud market is more likely to be a market where cloud services are provided by several clouds, and standardization and interoperability between providers are widespread [34], [131].

In a typical inter-cloud, the disparate cloud providers will be thereby able to dynamically partner with each other to create a seemingly infinite pool of computing resources. Moreover, an inter-cloud can democratize the supply side of market through allowing small and medium-sized cloud providers to enter a fair competition. Beside to the all aforementioned advantages, the interoperability between clouds in an inter-cloud brings many other advantages such as, optimizing market profits, load balancing, preventing legal issues, avoiding vendor lock-in, improving availability, enabling disaster recovery, lowering access latency, decreasing operational costs, handling a burst of incoming jobs, enabling better application resilience, providing better SLA, and partitioning extra-large tasks between multiple providers which cannot be accommodated by a single provider [34], [132].

So far, different types of inter-clouds have been introduced in the literature including federated clouds, hybrid clouds, cloud-of-clouds, sky computing, aggregated clouds, multi-tier clouds, etc [133]. Although, the taxonomy is not well defined yet, there is a general agreement that there are generally two types of delivery models in inter-clouds: federated clouds and multi-clouds [133]; where the main difference is defined by the degree of collaborations between the involved Clouds. In particular, a cloud federation is formed as a result of cooperation among cloud providers to increase their own

TABLE 5. Sequential and online auction mechanisms in cloud computing.

Model	Major features	Advantages	Disadvantages
Sequential auctions			
[121]	• Proposes dynamic auctions to periodically auction off VMs to accommodate user demands over time • Considers a capacity allocation scheme to determine the amount of instances to be auctioned off in each period	• Ensures IC and AsE • Allows the provider to reserve some available capacity for future high-bid requests • Does not terminate instances against users' will	• Considers only one type of VM • Does not guarantee IR and BB
[122]	• Proposes a combinatorial forward auction based on dynamic resource provisioning • Calculates the profit of the provider based on the revenue, the running cost of allocated VMs, and the cost of idle resources	• Ensures truthfulness • Improves resource utilization through dynamic provisioning	• Is not appropriate for uninterruptible tasks as the resources are taken back at the end of each round • Does not consider multiple provider scenarios • Does not guarantee IR and BB
[123]	• Proposes a combinatorial double auction, employing the idea of group-buying discounts offered by providers	• Ensures IR, IC, and BB • Enables group-buying discounts • Enable providers to benefit from demand aggregation	• Exposed to single point of failure due to the existence of a single auctioneer
Online auctions			
[124]	• Proposes an online mechanism which is invoked as soon as a user places a request or some allocated resources become available	• Ensures IC and IR • Considers multiple types of resources	• Does not guarantee BB • Does not include multiple provider scenarios
[125]	• Proposes an online combinatorial auction mechanism which allows for dynamic resource provisioning	• Ensures truthfulness • Guarantees an approximation ratio	• Does not guarantee IR and BB • The system runs in a time-slotted fashion, which tends toward a sequential model
[126]	• Proposes an online auction mechanism in a cloud spot market, maximizing the cloud provider's profit • Dynamically calculates reserve prices based on data center power usage effectiveness and electricity costs • Operates under conditions similar to the EC2 spot market	• Provides envy-freeness through determining equal sale prices for all winning bidders • Ensures truthfulness with high probability	• Considers only one type of VM • Does not guarantee the complete fulfilment of requests • Even though the proposed auction is called online, it runs in sequential rounds
[127]	• Proposes a novel bidding language to translate user-specific heterogeneous demands into regulated forms • Proposes a forward auction mechanism on top of the proposed bidding language	• Ensures truthfulness • Allows users to determine their desired allocation time and application type beside valuation	• Does not guarantee IR and BB • Does not consider multiple provider scenarios
[128]	• Proposes an online double auction mechanism aiming at revenue maximization, energy cost minimization, and performance maintenance • Consists of a weighted bipartite matching based algorithm for winner determination and a VCG driven algorithm for payment calculation	• Ensures IC, IR, and computational efficiency • Considers server maintenance times for an efficient resource allocation • Improves energy efficiency	• Does not guarantee BB

profit and the satisfaction of cloud users [134], [135]. In other words, in a cloud federation the cloud providers voluntarily choose to interconnect with each other to share resources. However, the existing collaboration is not realized by any user as each user will interact with only a single cloud and will have a transparent access to cloud resources. In contrast, in a typical Multi-Cloud, no volunteer interconnection or sharing of resources exist between clouds. Thus, the user or a trusted third-party would be responsible to acquire required resources from several independent clouds [133]. It should be mentioned that a very common type of multi-cloud is called hybrid cloud, which is often used in the case of cloud bursting. Figure 8 illustrates inter-clouds including a cloud federation,

a multi-cloud, and a hybrid cloud. It also demonstrates how users can access different cloud environments.

Although the interoperability among cloud providers brings many advantages, managing such interactions are believed to be complex as the resources in inter-clouds are distributed and controlled by different clouds [136]. In this respect, auction mechanisms have been employed in the literature for the problem of cloud resource management and pricing in inter-clouds due to their simplicity, versatility, and economic efficiency. In the following, we review some of the major studies in the literature.

Lee *et al.* [137] propose a real-time group auction system based on a combinatorial double auction in the cloud
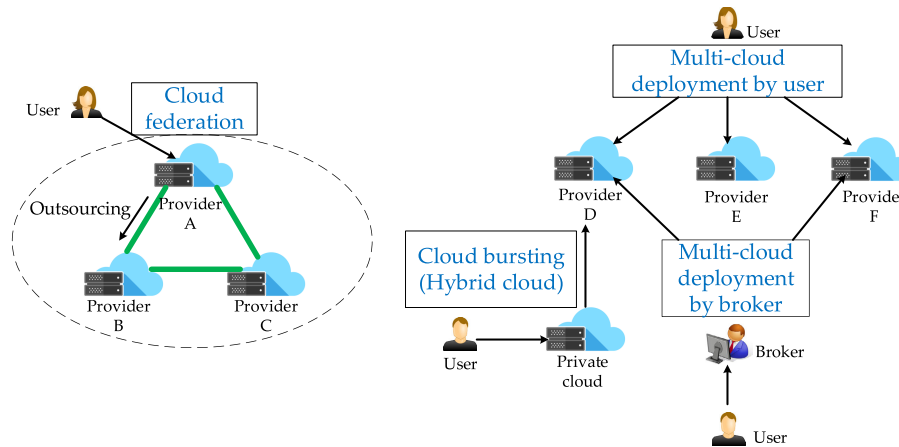
FIGURE 8. An inter-cloud including a cloud federation, a multi-cloud, and a hybrid cloud.

instance market. In the proposed mechanism, multiple users are allowed to form user coalition in order to achieve discounted prices, while multiple providers can form a coalition in order to improve resource utilization by attracting more users. The proposed mechanism guides participants' through their cooperation decisions, i.e., when to or not to cooperate.

Li *et al.* [138] address the problem of VM trading and scheduling in a cloud federation. In this respect, they propose a double auction mechanism combined with a dynamic VM trading and scheduling algorithm. The dynamic algorithm is responsible for determining the best VM valuation in the auction, optimally scheduling stochastic job arrivals with different resource requirements and SLAs onto different data centers, and judiciously turning on and off servers based on the current operational costs. However, they do not address the undergoing competition on the user side.

Mihailescu and Teo [139] propose a reverse auction-based mechanism considering a centralized auctioneer. Hence, the auction is prone to single point of failure. Unfortunately, the steps of the proposed auction mechanism are not very well defined, and its properties such as IC and IR are not theoretically proved.

Hosseinalipour and Dai [140] study both the interactions between the customers and cloud managers, and the interactions between the cloud managers and the cloud providers. For the interactions between customers and cloud managers, the options-based sequential auctions is adopted. For the interactions between cloud managers and providers, a novel model for resource gathering is proposed which consists of two parallel markets: flat-price market and auction-based market.

Table 6 summarizes the above studies, highlighting their main characteristics and pros and cons.

Summary and Discussion. According to what have been discussed in this section, the future cloud market is more likely to be a network of interconnected clouds to provide a seemingly infinite pool of computing resources. The network of interconnected clouds is usually referred as inter-clouds which can be further categorized into federated clouds and multi-clouds. However, inter-clouds research is still in its early stages and requires investigation in different aspects, such as resource management and pricing, security, privacy, collusion, energy efficiency, etc. Moreover, the effect of undergoing competition between users on the inter-cloud VM trading and the relation between front-end charging strategies and inter-cloud trading strategies require further studies.

X. AUCTION-BASED EDGE RESOURCE MANAGEMENT AND PRICING

As mentioned earlier in Section III, we are facing an exponential growth in the number of different IoT devices in the recent years. Nonetheless, the energy and resource constraints of such devices limit their capability to support many power-hungry, resource-intensive, and time-sensitive applications. Moreover, offloading such applications to conventional cloud infrastructures is not usually applicable due to existing long distances between the cloud data centers and the end users, imposing long communication latencies.

Along these lines, the concept of Edge Computing (EC) has recently emerged as a key technology that processes data generated from IoT devices at the edge of network [141], [142]. The geographically distributed computing resources at the edge of the network are usually referred as Cloudlets [62], which are actually trusted, resource-rich, Internet-connected computer or a cluster of computers, accessible via a high-speed WLAN. The close physical proximity of Cloudlets can thus be exploited to enable smoother interactions with the low one-hop communication latency. This is while sufficient number of Cloudlets should be deployed at the edge in order to satisfy the resource requirements of the users and preserve their quality of experience. In this regard, single or multiple Cloudlets may be assigned to each user depending on her requirements [77], [143].

TABLE 6. Auction mechanisms in inter-clouds.

Model	Major features	Advantages	Disadvantages
[137]	• Employs a combinatorial double auction mechanism • Allows multiple users to form a coalition in order to get discount • Allows multiple providers to form a coalition to share their resources • Divides the total payment of each user coalition among its users • Distributes the obtained revenue by each provider coalition among its providers	• Ensures IR, BB, and AE • Improves resource efficiency and monetary benefits compared to a conventional individual auction	• Does not guarantee IC • Although the auction is claimed to be real-time, it considers a concept of time slot
[138]	• Proposes an online double auction mechanism for VM trading in a federation of clouds • Aims to maximize the profit of each individual cloud as well as the social welfare of the federation	• Ensures IC, IR, and BB • Achieves near optimal solution • Achieves AsE under homogeneous cloud settings • Considers the SLA requirements of user jobs and the operational costs of servers	• Does not consider front-end job pricing and competition for customers among the clouds • Does not address the relation between front-end charging strategies and inter-cloud trading strategies • Runs in a time-slotted fashion and thus does not provide real-time decisions
[139]	• Proposes a dynamic pricing scheme for multiple types of shared resources in federated clouds • Employs a reverse auction-based mechanism	• Ensures IC, IR, and BB	• Does not discuss cloud resource allocation and pricing to cloud customers • Is prone to single point of failure
[140]	• Considers cloud of clouds networks consisting of heterogeneous cloud servers and customers with different demands • Proposes a two-stage auction framework to capture both the interactions among customers and cloud managers, and the interactions between cloud managers and providers	• Avoids the delayed entrance issue • Addresses both the exiting competition among cloud users to obtain resources from cloud providers, and the existing competition among cloud managers to rent resources from other providers	• Ensures truthfulness only in the first stage • Does not consider the optimization of any parameters of interest such as the social welfare

Beside to the edge resource management issues, resource pricing should also be addressed in parallel. In fact, pricing strategies need to be tailored in conjunction with edge resource management, with respect to the preferences of users and Cloudlets. Because, in a real edge environment, IoT devices often have different valuations toward Cloudlets according to their spatial locations, or offered resources and QoS. On the other hand, the Cloudlets need to be rewarded which can work as a motivation to share their resources. Hence, there exists a trade between the mobile users requesting the services on one hand, and the Cloudlets providing the services on the other hand.

However, most existing works only focus on task offloading strategies ignoring the motivations of edge Cloudlets to offer services. This is while we deal with resource-poor computing facilities in EC, and not big data centers. Therefore, a main challenge in EC is the resource poverty at the edge which intensifies the undergoing competition among users for acquiring their required resources within their proximity. Therefore, similar to cloud computing, edge resource allocation and economics must be addressed jointly in order to improve system efficiency and motivate price-driven self-interested users and Cloudlets.

Unfortunately, the existing auction based resource management and pricing mechanisms designed for cloud systems are not directly applicable to the edge computing systems.

It is because most of the proposed auction-based solutions for the cloud computing do not consider many features that are specific to edge computing such as resource-poverty and geographical distribution. Moreover, most of them do not consider the QoS requirements of the users which are of very high importance in edge computing.

In this respect, novel auction based mechanisms have been proposed in the recent years to address the problem of edge resource management and pricing. These studies have approached the aforementioned problem from different perspectives. We will review these studies and discuss their advantages and disadvantages in this section. As discussed earlier in Section III, different models of edge computing have been introduced in the literature including Cloudlets, fog computing, and MEC [55]. Here the proposed auction mechanism for any model can be simply extended for the other ones, i.e., attributed to the major similarities that exist between them.

In this respect, we reviewed the existing studies in this area based on the hierarchy levels of the deployed edge infrastructure and categorized them into two major groups. The first group of studies consider a flat infrastructure where computing resources are distributed at the very edge of the network. However, the second group includes those works which have assumed that computing resources are distributed at different levels of the network. We first review the studies which

have considered a flat architecture, including our previous studies.

In [77], [144] we propose two-sided auction mechanisms to match edge resources to IoT devices regarding the QoS requirements of the users and the QoS guarantees of the providers. In [144], we assume that one type of VM is offered by the providers and present two QoS-aware double auction based solutions namely QMECS and M-QMECS. Our proposed auctions not only provide efficient resource allocation solutions to maximize obtained social welfare, they also ensure that the matching providers will be able to provide the required QoS by users. M-QMECS is the modified version of QMECS which can augment social welfare and avoid zero budget surplus compared to QMECS. We then extend our work in [77] considering multiple types of VMs, while preserving all the desired properties such as IC, IR, BB, and AsE. We also improve the computational efficiency of the proposed mechanism by grouping the users with similar QoS requirements in the matching process.

Kayal and Liebeherr [143] address distributed service placement in fog computing by employing an iterative combinatorial auction approach. They aim to find an efficient placement such that the energy consumption at fog nodes as well as communication costs of applications are reduced.

Zavodovski *et al.* [145] proposes a decentralized double auction mechanism which is inspired by a distributed trust model that smart contracts can offer. Along this, they can provide a secure distributed sealed bids auction framework without relying on a centralized auctioneer.

Abdulsalam and Hossain [146] propose a multi-attribute auction mechanism that considers QoS attributes as well as pricing. The proposed mechanism consists of two stages. In the first stage, the mechanism uses the bids to qualify providers based on the broker's reservation quality and determines the cost value of different attributes. In the second stage, the maximum score is shared by the broker in order to allow the providers to increase their bids.

Luong *et al.* [147] propose a forward auction mechanism based on deep learning for the problem of resource allocation in fog computing which is specifically developed to support blockchain applications. They assume a single provider and multiple miners in their setting, where the provider deploys fog nodes across the blockchain network to provide nearby computing resources to the miners.

Sun *et al.* [148] model the two-sided interaction between MEC servers and industrial IoT (IIoT) mobile devices (MDs) under the constraints of limited computing resource of edge servers and locality between edge servers and IIoT MDs. In this respect, they propose two double auction mechanisms, namely a breakeven-based double auction (BDA) and a more efficient dynamic pricing based double auction (DPDA). DPDA can achieve more system efficiency when the behavior of edge servers can be regulated by the provider; otherwise, BDA is preferred in order to maintain a healthier market.

Zhang *et al.* [149] propose a multi-round-sealed sequential combinatorial auction mechanism in a MEC environment

with multiple providers and users. However, the proposed mechanism is not robust as it does not ensure truthfulness.

Zhang *et al.* [150] consider an ad-hoc Cloudlet and assume there are multiple MDs and IoT devices, and a dispatcher in the system. The dispatcher and MDs play the roles of the auctioneer and bidders, respectively. In this regard, they propose an online rewards-optimal auction to optimize the long-term sum-of-rewards for processing offloaded tasks, which is based on Lyapunov optimization and VCG auction.

Jin *et al.* [151] propose a two-sided auction mechanism to coordinate resource allocation and pricing between users and Cloudlets. They further extend their proposed auction mechanism to achieve higher system efficiency at the expense of losing truthfulness for the user side.

Wang *et al.* [152] propose a profit maximization multi-round auction considering a trusted third party as an auctioneer. In their proposed model, first the users submit their resource requirement to the auctioneer and then the providers submit their bids accordingly to it. Next, the auctioneer performs winner determination and pricing based on the proposed auction mechanism.

Mashhadia *et al.* [153] propose an auction mechanism based on deep neural networks in order to maximize the profit of the edge servers, while satisfying the task processing delay and energy consumption constraints of MDs. In his respect, they design new penalty functions for the task computing delay and energy consumption constraints to guide the neural networks though feasible solutions.

Peng *et al.* [154] consider the case of vehicular fog computing and propose a multi-attribute based double auction to make vehicles/fog nodes willing to buy/sell resources. The proposed auction mechanism satisfies computational efficiency, individual rationality, budget balance, and truthfulness. However, the proposed auction mechanism is offline and thus does not allow client vehicles and vehicular fog nodes to dynamically join the auction in real-time.

Table 7 summarizes the main features of the above studies. However, as we mentioned earlier in this section, another group of the proposed auction mechanisms in edge computing have assumed a hierarchical architecture of computing resources at the edge. In the following, we review this group of studies.

Kiani and Ansari [155] present the Cloudlet tier with three hierarchical levels by introducing the concept of field, shallow, and deep Cloudlets based on the principle of LTE-Advanced backhaul network. They consider a two time scale mechanism for the allocation of computing and communication resources as they argue that the traffic loads of the users change much faster than the economics of computing resources. In particular, computing resources are offered in an auction-based profit maximization manner in longer time frames, while the communication resources are allocated in shorter time slots such that the experienced delay by the users are minimized.

Gao *et al.* [156] consider a set of geo-distributed edge cloud nodes and a centralized cloud in their proposed model.

TABLE 7. Auction mechanisms in edge computing with flat infrastructures.

Model	Major features	Advantages	Disadvantages
[144]	• Proposes novel two-sided matching solutions, namely QMECS and M-QMECS, for edge services considering QoS requirements and guarantees • M-QMECS augments social welfare and avoids zero budget surplus compared to QMECS	• Ensures IC, IR, BB, and AsE • Improves the quality of experience of the users by guaranteeing to satisfy their QoS requirements	• Considers one type of VM • Lacks truthfulness in terms of QoS
[77]	• Extends [144] to a combinatorial model where users/Cloudlets can request/offer a variety of VMs • Devises a grouping approach for the users to reduce the computation	• Ensures IC, IR, BB, and AsE • Improves the quality of experience of the users by guaranteeing to satisfy their QoS requirements	• Lacks truthfulness in terms of QoS
[143]	• Proposes a fully distributed combinatorial auction mechanism for micro-service placement in fog computing	• Reduces energy consumption of fog nodes and communication costs of applications • Preserves the privacy of both fog nodes and applications	• Lacks IC and IR guarantee • Lacks satisfying QoS requirements
[145]	• Proposes a decentralized double auction mechanism based on a distributed ledger trust model	• Ensures IC, IR, and BB • Ensure client's privacy and security requirements	• A heuristic approach without leading to AE • Lacks satisfying QoS requirements
[146]	• Proposes a two-stage multi-attribute auction mechanism for edge resource allocation and pricing	• Ensures IC and IR • Considers QoS attributes beside pricing to provide more efficient resource procurement	• Lacks BB guarantee • Considers only one type of service • Lacks satisfying QoS requirements
[147]	• Proposes an optimal auction based on deep learning for the fog resource allocation in the blockchain network	• Scalable with the increase of number of miners	• Does not ensure IC, IR, and BB • Lacks any guarantees for the obtained social welfare • Absence of multiple provider scenarios
[148]	• Addresses the joint problem of network economics and resource allocation in MEC for IIoT • Proposes two double auction mechanisms, namely BDA and DPDA	• Both mechanisms ensure IR and BB • BDA ensures IC for both sides, whereas DPDA ensures IC only for IIoT mobile devices	• DPDA does not ensure truthfulness for the sellers • Limits incoming computing tasks to indivisible atomic tasks
[149]	• Proposes a multi-round sealed bid sequential combinatorial auction to allocate wireless and computational edge resources to multiple users	• Ensures IR and BB	• It does not ensure truthfulness • Considers only one type of computational resource • May result in late resource allocation due to the multi-round nature of the auction
[150]	• Investigates computation offloading in an edge system consisting of energy harvesting-enabled mobile devices and a dispatcher • Proposes an online rewards-optimal auction, based on a reverse auction model	• Ensures IC and AsE • Does not require a prior knowledge of the task arrivals or wireless channel statistics • Addresses green computing issues	• Does not ensure IR and BB • Assumes that each mobile device can win at most one computation task at each time slot which is not efficient
[151]	• Proposes two-sided auction mechanisms to coordinate the resource allocation and pricing between mobile devices and Cloudlets • Improves the proposed mechanism in terms of system efficiency by involving randomness and bidding uncertainty	• Ensures IC, IR, and BB	• Assumes that all provided resources are homogeneous • Limits each edge provider to serve only one request • Making individual bids for each provider may be exhaustive for the users
[152]	• Proposes an online multi-round reverse auction mechanism for the resource trading between edge clouds and mobile devices • Aims at maximizing the profit of providers	• Ensures IC and IR • Considers the mobility of users	• Does not ensure BB • Does not provide any guarantees for the achieved profit
[153]	• Proposes a deep-learning auction mechanism for edge resource allocation, satisfying the energy and delay constraints of users	• Ensures IC and IR • Considers the delay and energy constraints of mobile devices	• Does not ensure BB • Making individual bids for each edge server may be exhaustive for the users
[154]	• Proposes a multi-attribute double auction mechanism for a vehicular fog computing environment	• Ensures IC, IR and BB • Considers both price and non-price attributes to provide an efficient resource allocation	• Lacks support of real-time solutions (offline mechanism) • Does not consider the mobility of users

TABLE 8. Auction mechanisms in edge computing with hierarchical infrastructure.

Model	Major features	Advantages	Disadvantages
[155]	• Considers the Cloudlet tier in three hierarchical levels • Proposes a two time scale mechanism to allocate the computing and communication resources to the mobile users • Aims at maximizing the provider profit during longer time frames, while improving QoS for the users in shorter time slots	• Considers communication resources as well as computation resources	• Does not guarantee IC, IR, and BB • Postpones user requests until the beginning of each time frame for running the auction
[156]	• Considers several geo-distributed edge cloud nodes and one centralized cloud • Proposes an auction mechanism to solve the problem of VM allocation to mobile users with deadline-sensitive tasks	• Ensures IC and IR • Ensures an approximation ratio • Addresses the deadline requirements of the users	• Does not guarantee BB • Is not suitable for uninterruptible tasks as users have to bid on each auction cycle • Does not consider other QoS requirements beside time latency • Considers only one type of VM
[157]	• Considers a two-level computing infrastructure including cloud and edge resources • Proposes an envy-free forward auction mechanism for the allocation and pricing of edge/cloud resources to maximize the obtained social welfare • Combines features from both position and combinatorial auctions	• Ensures IR and envy-freeness • Considers heterogeneous user demands	• Does not ensure IC and BB • Does not provide any guarantees for the obtained social welfare
[158]	• Addresses the problem of provisioning heavily stateful LLAs over a hierarchical fog infrastructure • Proposes a decentralised pricing mechanism for on-path, on-demand LLA provisioning with respect to the availability of Cloudlets' computing resources and the users' expected QoS gain	• Ensures truthfulness	• Does not guarantee IR and BB • Does not consider the mobility of users • Does not consider combinatorial requirements of the users
[159]	• Proposes a multi-attribute combinatorial reverse auction for a three-layered fog-integrated cloud architecture • The number of resources are acquired based on the customer requirements and not the offered bundles by the provider • Categorizes pricing schemes into local fog resources, remote fog resources, and cloud resources	• Ensures IC, IR and BB • Ensures egalitarian social welfare • Considers both price and non-price attributes to provide an efficient resource allocation • Provides robustness by imposing penalty to providers with QoS violations and providing compensations for the effected customers	• Does not consider the price preferences of the customers

Then, they present an auction mechanism to solve the problem of VM resource allocation to mobile users with deadline-sensitive tasks. Their proposed mechanism consists of a greedy winning bid selection algorithm and a critical payment based pricing algorithm.

Bahreini *et al.* [157] consider a two-level edge computing system in which servers with different capacities are located in the cloud or at the edge of the network. Then, they propose an auction mechanism edge/cloud resource allocation and pricing. Their proposed mechanism combines features from both position and combinatorial auctions.

Tasiopoulos *et al.* [158] address the problem of provisioning heavily stateful Low Latency Applications (LLAs) over the fog infrastructure. They assume that computation resources are available at the edge and middle-tier of the network in the form of Cloudlets, and also at distant clouds. In this respect, they propose FogSpot, which is a pricing mechanism for on-path on-request LLA provisioning. In this mechanism, as LLA requests pass the Cloudlets towards the cloud, if the price of a market is less than the estimated QoS gain a LLA has, an available VM is allocated to it by the

corresponding Cloudlet; otherwise, the request is rejected and continues its journey towards the cloud.

Finally, Aggarwal *et al.* [159] propose a multi-attribute combinatorial reverse auction, assuming local Fog, remote Fog, and Cloud in the system. In their proposed mechanism, response time, data source mobility requirements, and Fog resource limitations are considered together for resource allocation. Another novelty of their work is that instead of allocating all resources of the bundle, only the required resources are allocated to the customer with the bundle discount.

Table 8 summarizes the aforementioned studies.

Summary and Discussion. This section discusses the necessity of edge resource management and pricing. It also argues why proposed auction mechanisms for cloud computing cannot be directly applied in the context of edge computing as they do not consider many features specific to those of edge computing environments. It then provides a comprehensive review over the proposed auction mechanisms, categorizing them into two groups based on their adopted architecture, i.e., flat and hierarchical architectures.

One of the major drawbacks of the existing studies is that most of them do not consider the mobility of users in their model. This is however in contrast to real world situations in which the majority of edge users include mobile devices that can move arbitrarily in the environment. Therefore, more studies are required in order to address the mobility of users in edge computing environments and its effect on the quality of experience of users.

XI. OPEN CHALLENGES

As discussed in Section IV-C, auction mechanism design is a promising solution for market-based cloud/edge computing. Along this, auction based market is adopted by pioneers in the cloud industry, such as Amazon EC2's Spot market [79].

However, existing employed auction mechanisms still suffer from various drawbacks. For example, in the case of Amazon's Spot market, firstly it does not allow users to request different VM types at the same time. Secondly, it charges winning users at equal prices and thus under-exploits the payment potential of users. Thirdly, it is vulnerable to untruthful behavior by users [80]. Finally, Amazon's spot price is not likely to be market driven [81].

As we discussed throughout the paper, different auction based mechanisms have been proposed in the literature to address the problem of cloud/edge resource management and pricing from different perspectives to make them applicable in real world situations. For example, two-sided auction mechanisms are able to model the undergoing competition on both sides which is more in line with real world scenarios. Because, on one hand the users compete to obtain cheapest computing resources with highest QoS, and on the other hand the providers compete to gain more profit by attracting more users. Moreover, combinatorial auction methods allow users to demand multiple types of resources, which arise due to the variation in the real world applications that require different resource types as well. Further, online auction methods enable real-time solutions that are vital for edge computing, i.e., to provide low-latency computing services. Another mechanism is the multi-attribute auction that features additional attributes besides pricing, which allows it to provide an enhanced efficiency in the resource allocation process. Finally, different auction mechanisms based upon flat and hierarchical architectures have been proposed for edge computing environments, taking into account the various features associated with edge computing.

However, despite the significant advances in the literature, there are still multiple challenges that need to be addressed. Hence this section highlights some of these key challenges that require further investigation regarding the real world scenarios.

A. DYNAMICITY

An interesting venue of research for designing auction-based resource management and pricing mechanisms in cloud/edge computing is considering the dynamicity in the requirements of the users and the availability of resources. Since, in a real world scenario most of the users demonstrate a dynamic model of requirement for the computing resources. In other words, the computing requirements of users are usually subject to change over time with regard to their computing tasks. Meanwhile, their QoS requirements or pricing desires may also change over different periods of time.

Cloud/edge providers show a very similar behavior as well. Indeed, the availability of their resources, QoS guarantees, and pricing preferences may vary over time. Hence, novel auction mechanisms need to be designed allowing the dynamicity in the preferences of the users and providers.

B. FALSE-NAME BIDDING

Many of the proposed auction mechanisms in the literature have considered incentive compatibility as a critical requirement in their work. However, just a few works have addressed false-name bidding which is another cheating pattern that can appear in the system.

False-name bidding happens when a bidder can increase her utility by submitting multiple bids under multiple fictitious names, which can lead to untruthful bidding and unfair scarcity [160]. It has previously emerged in various auctions running on Internet such as [161], [162]. Therefore, the same cheating behavior is inherent to breed in cloud/edge auctions. Hence, false-name-proof mechanisms should be designed to avoid false-name manipulations in cloud/edge environments under different scenarios.

In a false-name-proof mechanism, bidders are encouraged to bid truthfully under a single identifier, as the profit gain via false-name bidding is impossible. In other words, a false-name-proof auction guarantees that the price of obtaining a bundle of VM instances in a cloud auction would be smaller than or equal to the sum of prices of obtaining these instances separately by bidding under multiple fictitious names.

C. COMMUNICATION-COST AWARE AUCTION MECHANISMS

Multiple studies in the literature have considered the case of delivering the required resources of a single user by several providers. However, they do not account for the extra communication costs which can arise due to the placement of different VMs of a single requested bundle over distinct physical machines.

This problem becomes apparent when different physical machines are located at very distant geographical locations from each other or the VMs placed on different physical machines exchange high amounts of data with each other. Also, there may be cases that the physical machines are close but the communication links are unable to transmit large amount of data for some reason such as data congestion. Hence, the deployed VMs on these servers will experience communication delay despite physical proximity. Therefore, more studies are required to design new auction mechanisms considering the communication costs in the proposed allocation model.

D. ENERGY-AWARE AUCTION MECHANISMS

One of the major issues in cloud/edge industry is to reduce energy consumption. Since, cloud data centers generally consume enormous amount of electrical energy, contributing to high operational costs and carbon footprints to the environment.

Although energy consumption is a major dimension being considered by the cloud providers while setting their prices, efficient approaches are also needed to be devised during resource allocation process to reduce energy consumption and its imposed extra costs. For example, a very important factor in reducing energy consumption is to provide load balancing. Thereby, clearly addressing the efficient placement of VMs over different servers during auction mechanism design can lead to load balancing and subsequently provide energy efficiency.

Unfortunately, existing auction mechanisms have rarely considered energy efficiency in their proposed models. Therefore, an interesting venue for future researches can be designing of energy-aware auction mechanisms.

E. COLLUSION

In a typical cloud/edge environment, efficient mechanisms are needed to be designed to manage the interactions between different users and providers. Such mechanisms should also be able to prevent any misbehaviour by the rational users and providers who may intentionally collaborate to manipulate the system towards their own interests. In other words, different cloud/edge users and providers may collude with each other in order to increase their benefit. For example, consider the case of VCG-based pricing in which multiple users may collude so that the winning user acquires his required resources with a lower price. Hence, collusion-resistant mechanisms need to be designed to avoid such misbehavior.

F. SECURITY

Multiple security issues are likely to occur in an auction-based cloud/edge market, e.g. bid packet loss or broker exposure to a DoS attack. Such problems can thereby lead to user dissatisfaction due to the service unavailability.

Some aspects of security can be considered as one of QoS dimensions offered by the providers and thus implemented by means of multi-attribute auctions as discussed in Section VII. However, other solutions need to be devised for problems like DoS attack to a broker in order to avoid the failure of whole system or auction process. Hence, each security issue needs to be addressed properly in order to provide user satisfaction.

Unfortunately, existing auction mechanisms lack such security considerations in their proposed models and thereby there exist a research gap in this area. In this respect, efficient auction mechanisms need to be designed considering different security aspects to fill this gap.

G. DECENTRALIZED AUCTIONS

Different auction mechanisms may face scalability issues as the number of user requests or the variety of resource types increases. Moreover, auction mechanisms which are run by a single broker are exposed to the single point of failure. In this respect, decentralized auctions can overcome such problems by making allocation and payment decisions in a distributed manner.

In particular, the decentralized nature of edge computing infrastructure suggests even a much more suitable platform for developing distributed solutions compared to the centralized structure of classical cloud computing. Clearly, this will be accompanied by other challenges such as data sharing, privacy issues, etc. Hence, much more studies are required in this venue.

H. VARIETY IN THE SERVICES

Cloud/edge providers can provide a variety of different services such as IaaS, PaaS, SaaS, and lately XaaS. However, most of the previous studies have only considered the case of IaaS. Therefore, there is a need for new auction mechanisms which can allow for offering a variety of services by the providers.

XII. CONCLUSION

This paper has presented a comprehensive survey on auction based mechanisms for cloud/edge resource management and pricing. First, we have provided a review on fundamental concepts of cloud computing, edge computing and auction mechanism design. Then we have clarified how auction mechanism design can be used as a promising solution for cloud/edge resource management and pricing. Afterwards, we have provided detailed reviews, comparisons and discussions of the proposed auction mechanisms. Different two-sided, online, multi-attribute, and combinatorial auctions have been proposed in the literature to address different real world demands. Also, various auction mechanisms have been presented specifically for inter-clouds and edge computing environments. However, despite the achieved advances that have also led to the adoption of auction based mechanisms in the cloud industry, several challenges are still remaining which require further investigation. We have finalized our paper by presenting these open challenges in order to shed lights for future works.

REFERENCES

[1] H. Yuan, J. Bi, and M. Zhou, "Spatial task scheduling for cost minimization in distributed green cloud data centers," *IEEE Trans. Autom. Sci. Eng.*, vol. 16, no. 2, pp. 729–740, Apr. 2018.
[2] H. Rong, H. Zhang, S. Xiao, C. Li, and C. Hu, "Optimizing energy consumption for data centers," *Renew. Sustain. Energy Rev.*, vol. 58, pp. 674–691, May 2016.
[3] A. L. D. Salvo, F. Agostinho, C. M. Almeida, and B. F. Giannetti, "Can cloud computing be labeled as 'green'? Insights under an environmental accounting perspective," *Renew. Sustain. Energy Rev.*, vol. 69, pp. 514–526, Mar. 2017.
[4] T. Ouyang, Z. Zhou, and X. Chen, "Follow me at the edge: Mobility-aware dynamic service placement for mobile edge computing," *IEEE J. Sel. Areas Commun.*, vol. 36, no. 10, pp. 2333–2345, Oct. 2018.

[5] C. Jiang, Y. Qiu, H. Gao, T. Fan, K. Li, and J. Wan, "An edge computing platform for intelligent operational monitoring in internet data centers," *IEEE Access*, vol. 7, pp. 133375–133387, 2019.

[6] X. Sun and N. Ansari, "EdgeIoT: Mobile edge computing for the Internet of Things," *IEEE Commun. Mag.*, vol. 54, no. 12, pp. 22–29, Dec. 2016.

[7] S. Singh and I. Chana, "QoS-aware autonomic resource management in cloud computing: A systematic review," *ACM Comput. Surveys*, vol. 48, no. 3, pp. 1–46, Feb. 2015.

[8] G. Portaluri, D. Adami, A. Gabbrielli, S. Giordano, and M. Pagano, "Power consumption-aware virtual machine placement in cloud data center," *IEEE Trans. Green Commun. Netw.*, vol. 1, no. 4, pp. 541–550, Dec. 2017.

[9] W. Wang, Y. Zhao, M. Tornatore, A. Gupta, J. Zhang, and B. Mukherjee, "Virtual machine placement and workload assignment for mobile edge computing," in *Proc. IEEE 6th Int. Conf. Cloud Netw. (CloudNet)*, Sep. 2017, pp. 1–6.

[10] Z. Tao, Q. Xia, Z. Hao, C. Li, L. Ma, S. Yi, and Q. Li, "A survey of virtual machine management in edge computing," *Proc. IEEE*, vol. 107, no. 8, pp. 1482–1499, Aug. 2019.

[11] M. Mihailescu and Y. M. Teo, "Dynamic resource pricing on federated clouds," in *Proc. 10th IEEE/ACM Int. Conf. Cluster, Cloud Grid Comput. (CCGrid)*, May 2010, pp. 513–517.

[12] G. Wei, A. V. Vasilakos, Y. Zheng, and N. Xiong, "A game-theoretic method of fair resource allocation for cloud computing services," *J. Supercomput.*, vol. 54, no. 2, pp. 252–269, Nov. 2010.

[13] F. Teng and F. Magoulès, "A new game theoretical resource allocation algorithm for cloud computing," in *Advances in Grid and Pervasive Computing*, P. Bellavista, R.-S. Chang, H.-C. Chao, S.-F. Lin, and P. M. A. Sloot, Eds. Berlin, Germany: Springer, 2010, pp. 321–330.

[14] K. M. Sim, "Agent-based interactions and economic encounters in an intelligent InterCloud," *IEEE Trans. Cloud Computing*, vol. 3, no. 3, pp. 358–371, Jul. 2015.

[15] K. M. Sim, "Agent-based cloud computing," *IEEE Trans. Serv. Comput.*, vol. 5, no. 4, pp. 564–577, Oct. 2011.

[16] W. Wei, X. Fan, H. Song, X. Fan, and J. Yang, "Imperfect information dynamic stackelberg game based resource allocation using hidden Markov for cloud computing," *IEEE Trans. Serv. Comput.*, vol. 11, no. 1, pp. 78–89, Jan. 2018.

[17] R. B. Myerson, *Game Theory*. Cambridge, MA, USA: Harvard Univ. Press, 2013.

[18] N. Nisan, T. Roughgarden, E. Tardos, and V. V. Vazirani, *Algorithmic Game Theory*. Cambridge, U.K.: Cambridge Univ. Press, 2007.

[19] R. P. McAfee and J. McMillan, "Auctions and bidding," *J. Econ. Literature*, vol. 25, no. 2, pp. 699–738, Apr. 1987.

[20] S. Subashini and V. Kavitha, "A survey on security issues in service delivery models of cloud computing," *J. Netw. Comput. Appl.*, vol. 34, no. 1, pp. 1–11, 2011.

[21] C. Rong, S. T. Nguyen, and M. G. Jaatun, "Beyond lightning: A survey on security challenges in cloud computing," *Comput. Electr. Eng.*, vol. 39, no. 1, pp. 47–54, Jan. 2013.

[22] C. Modi, D. Patel, B. Borisaniya, A. Patel, and M. Rajarajan, "A survey on security issues and solutions at different layers of cloud computing," *J. Supercomput.*, vol. 63, pp. 561–592, Oct. 2013.

[23] Q. Yan, F. R. Yu, Q. Gong, and J. Li, "Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: A survey, some research issues, and challenges," *IEEE Commun. Surveys Tuts.*, vol. 18, no. 1, pp. 602–622, Jan. 2015.

[24] S. Singh, Y.-S. Jeong, and J. H. Park, "A survey on cloud computing security: Issues, threats, and solutions," *J. Netw. Comput. Appl.*, vol. 75, pp. 200–222, Nov. 2016.

[25] R. Roman, J. Lopez, and M. Mambo, "Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges," *Future Gener. Comput. Syst.*, vol. 78, pp. 680–698, Jan. 2018.

[26] J. Zhang, B. Chen, Y. Zhao, X. Cheng, and F. Hu, "Data security and privacy-preserving in edge computing paradigm: Survey and open issues," *IEEE Access*, vol. 6, pp. 18209–18237, 2018.

[27] T. Mastelic, A. Oleksiak, H. Claussen, I. Brandic, J.-M. Pierson, and A. V. Vasilakos, "Cloud computing: Survey on energy efficiency," *ACM Comput. Surv.*, vol. 47, no. 2, p. 33, 2015.

[28] A. Beloglazov, R. Buyya, Y. C. Lee, and A. Zomaya, "A taxonomy and survey of energy-efficient data centers and cloud computing systems," in *Advances in Computers*, vol. 82. Amsterdam, The Netherlands: Elsevier, 2011, pp. 47–111.

[29] A. Hameed, A. Khoshkbarforoushha, R. Ranjan, P. P. Jayaraman, J. Kolodziej, P. Balaji, S. Zeadally, Q. M. Malluhi, N. Tziritas, A. Vishnu, and S. U. Khan, "A survey and taxonomy on energy efficient resource allocation techniques for cloud computing systems," *Computing*, vol. 98, no. 7, pp. 751–774, Jul. 2016.

[30] C. Jiang, T. Fan, H. Gao, W. Shi, L. Liu, C. Cérin, and J. Wan, "Energy aware edge computing: A survey," *Comput. Commun.*, vol. 151, pp. 556–580, Feb. 2020.

[31] G. Galante and L. C. E. de Bona, "A survey on cloud computing elasticity," in *Proc. IEEE 5th Int. Conf. Utility Cloud Comput.* Washington, DC, USA: IEEE Computer Society, Nov. 2012, pp. 263–270.

[32] R. Jain and S. Paul, "Network virtualization and software defined networking for cloud computing: A survey," *IEEE Commun. Mag.*, vol. 51, no. 11, pp. 24–31, Nov. 2013.

[33] A. C. Baktir, A. Ozgovde, and C. Ersoy, "How can edge computing benefit from software-defined networking: A survey, use cases, and future directions," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2359–2391, 4th Quart., 2017.

[34] A. N. Toosi, R. N. Calheiros, and R. Buyya, "Interconnected cloud computing environments: Challenges, taxonomy, and survey," *ACM Comput. Surveys*, vol. 47, no. 1, pp. 1–47, Jul. 2014.

[35] S. S. Manvi and G. K. Shyam, "Resource management for infrastructure as a service (IaaS) in cloud computing: A survey," *J. Netw. Comput. Appl.*, vol. 41, pp. 424–440, May 2014.

[36] C.-H. Hong and B. Varghese, "Resource management in fog/edge computing: A survey on architectures, infrastructure, and algorithms," *ACM Comput. Surv.*, vol. 52, no. 5, pp. 1–37, Oct. 2019.

[37] Z.-H. Zhan, X.-F. Liu, Y.-J. Gong, J. Zhang, H. S.-H. Chung, and Y. Li, "Cloud computing resource scheduling and a survey of its evolutionary approaches," *ACM Comput. Surveys*, vol. 47, p. 63, Jul. 2015.

[38] S. Singh and I. Chana, "A survey on resource scheduling in cloud computing: Issues and challenges," *J. Grid Comput.*, vol. 14, no. 2, pp. 217–264, Jun. 2016.

[39] C. Wu, R. Buyya, and K. Ramamohanarao, "Cloud pricing models: Taxonomy, survey, and interdisciplinary challenges," *ACM Comput. Surveys*, vol. 52, no. 6, pp. 1–36, Jan. 2020.

[40] D. Kumar, G. Baranwal, Z. Raza, and D. P. Vidyarthi, "A survey on spot pricing in cloud computing," *J. Netw. Syst. Manage.*, vol. 26, no. 4, pp. 809–856, Oct. 2018.

[41] D. Kumar, G. Baranwal, Z. Raza, and D. P. Vidyarthi, "A systematic study of double auction mechanisms in cloud computing," *J. Syst. Softw.*, vol. 125, pp. 234–255, Mar. 2017.

[42] N. C. Luong, P. Wang, D. Niyato, Y. Wen, and Z. Han, "Resource management in cloud networking using economic analysis and pricing models: A survey," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 2, pp. 954–1001, 2nd Quart., 2017.

[43] P. Mach and Z. Becvar, "Mobile edge computing: A survey on architecture and computation offloading," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 3, pp. 1628–1656, 3rd Quart., 2017.

[44] Q.-V. Pham, F. Fang, V. N. Ha, M. J. Piran, M. Le, L. B. Le, W.-J. Hwang, and Z. Ding, "A survey of multi-access edge computing in 5G and beyond: Fundamentals, technology integration, and state-of-the-art," *IEEE Access*, vol. 8, pp. 116974–117017, 2020.

[45] S. Wang, J. Xu, N. Zhang, and Y. Liu, "A survey on service migration in mobile edge computing," *IEEE Access*, vol. 6, pp. 23511–23528, 2018.

[46] W. Yu, F. Liang, X. He, W. G. Hatcher, C. Lu, J. Lin, and X. Yang, "A survey on the edge computing for the Internet of Things," *IEEE Access*, vol. 6, pp. 6900–6919, 2018.

[47] F. Sheikholeslami and N. J. Navimipour, "Auction-based resource allocation mechanisms in the cloud environments: A review of the literature and reflection on future challenges," *Concurrency Comput., Pract. Exper.*, vol. 30, no. 16, p. e4456, Aug. 2018.

[48] N. Shinde and P. S. Kiran, "A survey of cloud auction mechanisms & decision making in cloud market to achieve highest resource & cost efficiency," in *Proc. Int. Conf. Autom. Control Dyn. Optim. Techn. (ICACDOT)*, Sep. 2016, pp. 1158–1162.

[49] H. Qiu, K. Zhu, N. Cong Luong, C. Yi, D. Niyato, and D. I. Kim, "Applications of auction and mechanism design in edge computing: A survey," 2021, *arXiv:2105.03559*. [Online]. Available: http://arxiv.org/abs/2105.03559

[50] P. Mell and T. Grance, "The NIST definition of cloud computing," Nat. Inst. Standards Technol., Gaithersburg, MD, USA, Tech. Rep. SP 800-145, 2011.

[51] D. Ardagna, G. Casale, M. Ciavotta, J. F. Pérez, and W. Wang, "Quality-of-service in cloud computing: Modeling techniques and their applications," *J. Internet Services Appl.*, vol. 5, no. 1, pp. 1–17, Dec. 2014.

[52] M. Jelassi, C. Ghazel, and L. A. Saïdane, "A survey on quality of service in cloud computing," in *Proc. 3rd Int. Conf. Frontiers Signal Process. (ICFSP)*, Sep. 2017, pp. 63–67.

[53] P. Patel, A. H. Ranabahu, and A. P. Sheth, "Service level agreement in cloud computing," Wright State Univ., Dayton, OH, USA, Tech. Rep., 2009.

[54] *The Number of IoT Connected Devices*. Accessed: Sep. 1, 2021. [Online]. Available: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

[55] W. Z. Khan, E. Ahmed, S. Hakak, I. Yaqoob, and A. Ahmed, "Edge computing: A survey," *Future Gener. Comput. Syst.*, vol. 97, pp. 219–235, Aug. 2019.

[56] M. Patel, B. Naughton, C. Chan, N. Sprecher, S. Abeta, and A. Neal, "Mobile-edge computing introductory technical white paper," Mobile-Edge Comput. (MEC) Ind. Initiative, London, U.K., White Paper, 2014, pp. 854–864, vol. 29.

[57] *Cisco Fog Computing Solutions: Unleash the Power of the Internet of Things*. Accessed: Sep. 1, 2021. [Online]. Available: https://www.cisco.com/c/dam/en_us/solutions/trends/iot/docs/computing-solutions.pdf

[58] Y. Mao, C. You, J. Zhang, K. Huang, and K. B. Letaief, "A survey on mobile edge computing: The communication perspective," *IEEE Commun. Surveys Tuts.*, vol. 19, no. 4, pp. 2322–2358, 4th Quart., 2017.

[59] N. Siasi, M. Jasim, and N. Ghani, "Service function chain mapping in fog networks," *IEEE Commun. Lett.*, vol. 25, no. 1, pp. 99–102, Jan. 2020.

[60] N. Siasi, M. Jasim, A. Aldalbahi, and N. Ghani, "Delay-aware SFC provisioning in hybrid fog-cloud computing architectures," *IEEE Access*, vol. 8, pp. 167383–167396, 2020.

[61] N. Siasi, M. Jasim, A. Aldalbahi, and N. Ghani, "Deep learning for service function chain provisioning in fog computing," *IEEE Access*, vol. 8, pp. 167665–167683, 2020.

[62] M. Satyanarayanan, P. Bahl, R. Caceres, and N. Davies, "The case for VM-based cloudlets in mobile computing," *IEEE Pervas. Comput.*, vol. 8, no. 4, pp. 14–23, Oct. 2009.

[63] Y. Shoham and K. Leyton-Brown, *Multiagent Systems: Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge, U.K.: Cambridge Univ. Press, 2008.

[64] N. Nisan and A. Ronen, "Algorithmic mechanism design," *Games Econ. Behav.*, vol. 35, nos. 1–2, pp. 166–196, 2001.

[65] S. Parsons, J. A. Rodriguez-Aguilar, and M. Klein, "Auctions and bidding: A guide for computer scientists," *ACM Comput. Surveys*, vol. 43, no. 2, p. 10, 2011.

[66] J. D. Hartline and B. Lucier, "Bayesian algorithmic mechanism design," in *Proc. 42nd ACM Symp. Theory Comput. (STOC)*, 2010, pp. 301–310.

[67] A. Mas-Colell, M. D. Whinston, and J. R. Green, *Microeconomic Theory*, vol. 1. New York, NY, USA: Oxford Univ. Press, 1995.

[68] L. Y. Chu, "Truthful bundle/multiunit double auctions," *Manage. Sci.*, vol. 55, no. 7, pp. 1184–1198, Jul. 2009.

[69] R. B. Myerson and M. A. Satterthwaite, "Efficient mechanisms for bilateral trading," *J. Econ. Theory*, vol. 29, no. 2, pp. 265–281, Apr. 1983.

[70] W. Vickrey, "Counterspeculation, auctions, and competitive sealed tenders," *J. Finance*, vol. 16, no. 1, pp. 8–37, 1961.

[71] E. H. Clarke, "Multipart pricing of public goods," *Public Choice*, vol. 11, no. 1, pp. 17–33, Sep. 1971.

[72] T. Groves, "Incentives in teams," *Econometrica, J. Econ. Soc.*, vol. 41, pp. 617–631, Jul. 1973.

[73] R. Buyya, C. S. Yeo, S. Venugopal, J. Broberg, and I. Brandic, "Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility," *Future Generat. Comput. Syst.*, vol. 25, no. 6, pp. 599–616, 2009.

[74] R. Buyya, C. S. Yeo, and S. Venugopal, "Market-oriented cloud computing: Vision, hype, and reality for delivering IT services as computing utilities," in *Proc. 10th IEEE Int. Conf. High Perform. Comput. Commun.*, Sep. 2008, pp. 5–13.

[75] J. Ren, L. Pang, and Y. Cheng, "Dynamic pricing scheme for IaaS cloud platform based on load balancing: A Q-learning approach," in *Proc. 8th IEEE Int. Conf. Softw. Eng. Service Sci. (ICSESS)*, Nov. 2017, pp. 806–810.

[76] M. Khodak, L. Zheng, A. S. Lan, C. Joe-Wong, and M. Chiang, "Learning cloud dynamics to optimize spot instance bidding strategies," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, Apr. 2018, pp. 2762–2770.

[77] N. Sharghivand, F. Derakhshan, L. Mashayekhy, and L. M. Khanli, "An edge computing matching framework with guaranteed quality of service," *IEEE Trans. Cloud Comput.*, early access, Jun. 29, 2020, doi: 10.1109/TCC.2020.3005539.

[78] X. Zhang, C. Wu, Z. Li, and F. C. Lau, "A truthful $(1 − \epsilon)$-optimal mechanism for on-demand cloud resource provisioning," *IEEE Trans. Cloud Comput.*, vol. 8, no. 3, pp. 735–748, Sep. 2020.

[79] *Amazon EC2 Spot Instances*. [Online]. Available: https://aws.amazon.com/ec2/spot/pricing/

[80] Q. Wang, K. Ren, and X. Meng, "When cloud meets eBay: Towards effective pricing for cloud computing," in *Proc. IEEE INFOCOM*, Mar. 2012, pp. 936–944.

[81] O. A. Ben-Yehuda, M. Ben-Yehuda, A. Schuster, and D. Tsafrir, "Deconstructing Amazon EC2 spot instance pricing," *ACM Trans. Econ. Comput.*, vol. 1, no. 3, pp. 16:1–16:20, Sep. 2013.

[82] Q. Wang, B. Ye, B. Tang, S. Guo, and S. Lu, "EBay in the clouds: False-name-proof auctions for cloud resource allocation," in *Proc. IEEE 35th Int. Conf. Distrib. Comput. Syst.*, Jun. 2015, pp. 153–162.

[83] T. T. Huu and C.-K. Tham, "An auction-based resource allocation model for green cloud computing," in *Proc. IEEE Int. Conf. Cloud Eng. (IC E)*, Mar. 2013, pp. 269–278.

[84] P. Klemperer, "Auction theory: A guide to the literature," *J. Econ. Surv.*, vol. 13, no. 3, pp. 227–286, 1999.

[85] G. Manoochehri and C. Lindsy, "Reverse auctions: Benefits, challenges, and best practices," *California J. Oper. Manage.*, vol. 6, no. 1, pp. 123–130, 2008.

[86] A. S. Prasad and S. Rao, "A mechanism design approach to resource procurement in cloud computing," *IEEE Trans. Comput.*, vol. 63, no. 1, pp. 17–30, Jan. 2013.

[87] L. R. Smeltzer and A. Carr, "Reverse auctions in industrial marketing and buying," *Bus. Horizons*, vol. 45, no. 2, pp. 47–52, Mar. 2002.

[88] P. Bonacquisto, G. D. Modica, G. Petralia, and O. Tomarchio, "A procurement auction market to trade residual cloud computing capacity," *IEEE Trans. Cloud Comput.*, vol. 3, no. 3, pp. 345–357, Jul. 2014.

[89] M. Tanaka and Y. Murakami, "Strategy-proof pricing for cloud service composition," *IEEE Trans. Cloud Comput.*, vol. 4, no. 3, pp. 363–375, Jul. 2014.

[90] A. Bestavros and O. Krieger, "Toward an open cloud marketplace: Vision and first steps," *IEEE Internet Comput.*, vol. 18, no. 1, pp. 72–77, Jan. 2014.

[91] H. Wang, Z. Kang, and L. Wang, "Performance-aware cloud resource allocation via fitness-enabled auction," *IEEE Trans. Parallel Distrib. Syst.*, vol. 27, no. 4, pp. 1160–1173, Apr. 2015.

[92] X. Wang, X. Wang, H. Che, K. Li, M. Huang, and C. Gao, "An intelligent economic approach for dynamic resource allocation in cloud services," *IEEE Trans. Cloud Comput.*, vol. 3, no. 3, pp. 275–289, Apr. 2015.

[93] D. Kumar, G. Baranwal, Z. Raza, and D. P. Vidyarthi, "A truthful combinatorial double auction-based marketplace mechanism for cloud computing," *J. Syst. Softw.*, vol. 140, pp. 91–108, Jun. 2018.

[94] L. Lu, J. Yu, Y. Zhu, and M. Li, "A double auction mechanism to bridge users' task requirements and Providers' resources in two-sided cloud markets," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 4, pp. 720–733, Apr. 2018.

[95] K. Cheng, W. Tong, L. Zheng, J. Fu, X. Mu, and Y. Shen, "A secure and fair double auction framework for cloud virtual machines," *IEEE Access*, vol. 9, pp. 87982–87994, 2021.

[96] H. Fu, Z. Li, C. Wu, and X. Chu, "Core-selecting auctions for dynamically allocating heterogeneous VMs in cloud computing," in *Proc. IEEE 7th Int. Conf. Cloud Comput.*, Jun. 2014, pp. 152–159.

[97] S. Chichin, Q. B. Vo, and R. Kowalczyk, "Towards efficient and truthful market mechanisms for double-sided cloud markets," *IEEE Trans. Services Comput.*, vol. 10, no. 1, pp. 37–51, Jul. 2016.

[98] B. Yang, Z. Li, S. Jiang, and K. Li, "Envy-free auction mechanism for VM pricing and allocation in clouds," *Future Gener. Comput. Syst.*, vol. 86, pp. 680–693, Sep. 2018.

[99] S. A. Tafsiri and S. Yousefi, "Combinatorial double auction-based resource allocation mechanism in cloud computing market," *J. Syst. Softw.*, vol. 137, pp. 322–334, Mar. 2018.

[100] R. Singhal and A. Singhal, "A feedback-based combinatorial fair economical double auction resource allocation model for cloud computing," *Future Gener. Comput. Syst.*, vol. 115, pp. 780–797, Feb. 2021.

[101] L. Mashayekhy, M. M. Nejad, and D. Grosu, "A PTAS mechanism for provisioning and allocation of heterogeneous cloud resources," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 9, pp. 2386–2399, Sep. 2014.

[102] M. M. Nejad, L. Mashayekhy, and D. Grosu, "Truthful greedy mechanisms for dynamic virtual machine provisioning and allocation in clouds," *IEEE Trans. Parallel Distrib. Syst.*, vol. 26, no. 2, pp. 594–603, Feb. 2014.

[103] L. Mashayekhy, M. M. Nejad, and D. Grosu, "Physical machine resource management in clouds: A mechanism design approach," *IEEE Trans. Cloud Comput.*, vol. 3, no. 3, pp. 247–260, Jul. 2014.

[104] L. Zhang, Z. Li, and C. Wu, "Dynamic resource provisioning in cloud computing: A randomized auction approach," in *Proc. IEEE Conf. Comput. Commun. (IEEE INFOCOM)*, Apr. 2014, pp. 433–441.

[105] J. Li, Y. Zhu, J. Yu, C. Long, G. Xue, and S. Qian, "Online auction for IaaS clouds: Towards elastic user demands and weighted heterogeneous VMs," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 9, pp. 2075–2089, Sep. 2018.

[106] *CloudSigma*. Accessed: Sep. 1, 2021. [Online]. Available: https://www.cloudsigma.com

[107] X. Zhang, Z. Huang, C. Wu, Z. Li, and F. C. M. Lau, "Online auctions in IaaS clouds: Welfare and profit maximization with server costs," *IEEE/ACM Trans. Netw.*, vol. 25, no. 2, pp. 1034–1047, Apr. 2016.

[108] J. Zhang, N. Xie, X. Zhang, and W. Li, "An online auction mechanism for cloud computing resource allocation and pricing based on user evaluation and cost," *Future Gener. Comput. Syst.*, vol. 89, pp. 286–299, Dec. 2018.

[109] T. Hobfeld, R. Schatz, M. Varela, and C. Timmerer, "Challenges of QoE management for cloud applications," *IEEE Commun. Mag.*, vol. 50, no. 4, pp. 28–36, Apr. 2012.

[110] A. Abdelmaboud, D. N. A. Jawawi, I. Ghani, A. Elsafi, and B. Kitchenham, "Quality of service approaches in cloud computing: A systematic mapping study," *J. Syst. Softw.*, vol. 101, pp. 159–179, Mar. 2015.

[111] S. Garg, S. Versteeg, and R. Buyya, "A framework for ranking of cloud computing services," *Future Gener. Comput. Syst.*, vol. 29, no. 4, pp. 1012–1023, 2013.

[112] T. L. Saaty, *The Analytic Hierarchy Process: Planning, Priority Setting, Resource Allocation*. New York, NY, USA: McGraw-Hill, 1980.

[113] G. V. Prasad, A. S. Prasad, and S. Rao, "A combinatorial auction mechanism for multiple resource procurement in cloud computing," *IEEE Trans. Cloud Comput.*, vol. 6, no. 4, pp. 904–914, Oct. 2016.

[114] L. Zeng, B. Benatallah, A. H. H. Ngu, M. Dumas, J. Kalagnanam, and H. Chang, "QoS-aware middleware for web services composition," *IEEE Trans. Softw. Eng.*, vol. 30, no. 5, pp. 311–327, May 2004.

[115] M. Mohabey, Y. Narahari, S. Mallick, P. Suresh, and S. V. Subrahmanya, "A combinatorial procurement auction for QoS-aware web services composition," in *Proc. IEEE Int. Conf. Autom. Sci. Eng.*, Sep. 2007, pp. 716–721.

[116] G. Baranwal and D. P. Vidyarthi, "A fair multi-attribute combinatorial double auction model for resource allocation in cloud computing," *J. Syst. Softw.*, vol. 108, pp. 60–76, Oct. 2015.

[117] G. Baranwal and D. P. Vidyarthi, "A truthful and fair multi-attribute combinatorial reverse auction for resource procurement in cloud computing," *IEEE Trans. Services Comput.*, vol. 12, no. 6, pp. 851–864, Nov. 2019.

[118] U. Lampe, M. Siebenhaar, A. Papageorgiou, D. Schuller, and R. Steinmetz, "Maximizing cloud provider profit from equilibrium price auctions," in *Proc. IEEE 5th Int. Conf. Cloud Comput.*, Jun. 2012, pp. 83–90.

[119] S. Zaman and D. Grosu, "Combinatorial auction-based allocation of virtual machine instances in clouds," *J. Parallel Distrib. Comput.*, vol. 73, no. 4, pp. 495–508, Apr. 2013.

[120] L. Mashayekhy, M. M. Nejad, and D. Grosu, "A truthful approximation mechanism for autonomic virtual machine provisioning and allocation in clouds," in *Proc. ACM Cloud Autonomic Comput. Conf. (CAC)*, 2013, pp. 1–10.

[121] W. Wang, B. Liang, and B. Li, "Revenue maximization with dynamic auctions in IaaS cloud markets," in *Proc. IEEE/ACM 21st Int. Symp. Qual. Service (IWQoS)*, Jun. 2013, pp. 1–6.

[122] S. Zaman and D. Grosu, "A combinatorial auction-based mechanism for dynamic VM provisioning and allocation in clouds," *IEEE Trans. Cloud Comput.*, vol. 1, no. 2, pp. 129–141, Jul./Dec. 2013.

[123] Z. Sun, Z. Zhu, L. Chen, H. Xu, and L. Huang, "A combinatorial double auction mechanism for cloud resource group-buying," in *Proc. IEEE 33rd Int. Perform. Comput. Commun. Conf. (IPCCC)*, Dec. 2014, pp. 1–8.

[124] L. Mashayekhy, M. M. Nejad, D. Grosu, and A. V. Vasilakos, "An online mechanism for resource allocation and pricing in clouds," *IEEE Trans. Comput.*, vol. 65, no. 4, pp. 1172–1184, Apr. 2015.

[125] W. Shi, L. Zhang, C. Wu, Z. Li, and F. C. M. Lau, "An online auction framework for dynamic resource provisioning in cloud computing," *IEEE/ACM Trans. Netw.*, vol. 24, no. 4, pp. 2060–2073, Aug. 2016.

[126] A. N. Toosi, K. Vanmechelen, F. Khodadadi, and R. Buyya, "An auction mechanism for cloud spot markets," *ACM Trans. Auto. Adapt. Syst.*, vol. 11, no. 1, pp. 1–33, Apr. 2016.

[127] H. Zhang, H. Jiang, B. Li, F. Liu, A. V. Vasilakos, and J. Liu, "A framework for truthful online auctions in cloud computing with heterogeneous user demands," *IEEE Trans. Comput.*, vol. 65, no. 3, pp. 805–818, Mar. 2015.

[128] Y. S. Patel, Z. Malwi, A. Nighojkar, and R. Misra, "Truthful online double auction based dynamic resource provisioning for multi-objective trade-offs in IaaS clouds," *Cluster Comput.*, vol. 24, pp. 1–25, Jan. 2021.

[129] D. Hilley, "Cloud computing: A taxonomy of platform and infrastructure-level offerings," Georgia Inst. Technol., Atlanta, GA, USA, Tech. Rep., 2009.

[130] D. Bernstein, E. Ludvigson, K. Sankar, S. Diamond, and M. Morrow, "Blueprint for the intercloud–protocols and formats for cloud computing interoperability," in *Proc. 4th Int. Conf. Internet Web Appl. Services*, 2009, pp. 328–336.

[131] B. Rochwerger, D. Breitgand, A. Epstein, D. Hadas, I. Loy, K. Nagin, J. Tordsson, C. Ragusa, M. Villari, S. Clayman, E. Levy, A. Maraschini, P. Massonet, H. Mu, and G. Tofetti, "Reservoir—When one cloud is not enough," *Computer*, vol. 44, no. 3, pp. 44–51, Mar. 2011.

[132] M. Grozev and R. Buyya, "Inter-Cloud architectures and application brokering: Taxonomy and survey," *Softw., Pract. Exper.*, vol. 44, no. 3, pp. 369–390, Mar. 2014.

[133] D. Petcu, "Multi-cloud: Expectations and current approaches," in *Proc. Int. Workshop Multi-Cloud Appl. Federated Clouds (MultiCloud)*, 2013, pp. 1–6.

[134] T. Kurze, M. Klems, D. Bermbach, A. Lenk, S. Tai, and M. Kunze, "Cloud federation," in *Proc. Cloud Comput.*, Sep. 2011, pp. 32–38.

[135] R. Buyya, R. Ranjan, and R. N. Calheiros, "InterCloud: Utility-oriented federation of cloud computing environments for scaling of application services," in *Proc. Int. Conf. Algorithms Archit. Parallel Process.* Berlin, Germany: Springer, 2010, pp. 13–31.

[136] K. M. Sim, "Agent-based approaches for intelligent intercloud resource allocation," *IEEE Trans. Cloud Comput.*, vol. 7, no. 2, pp. 442–455, Apr. 2019.

[137] C. Lee, P. Wang, and D. Niyato, "A real-time group auction system for efficient allocation of cloud internet applications," *IEEE Trans. Services Comput.*, vol. 8, no. 2, pp. 251–268, Mar./Apr. 2013.

[138] H. Li, C. Wu, Z. Li, and F. C. M. Lau, "Virtual machine trading in a federation of clouds: Individual profit and social welfare maximization," *IEEE/ACM Trans. Netw.*, vol. 24, no. 3, pp. 1827–1840, Sep. 2015.

[139] M. Mihailescu and Y. M. Teo, "Strategy-proof dynamic resource pricing of multiple resource types on federated clouds," in *Proc. Int. Conf. Algorithms Archit. Parallel Process.* Berlin, Germany: Springer, 2010, pp. 337–350.

[140] S. Hosseinalipour and H. Dai, "A two-stage auction mechanism for cloud resource allocation," *IEEE Trans. Cloud Comput.*, vol. 9, no. 3, pp. 881–895, Jul. 2021.

[141] P. Garcia Lopez, A. Montresor, D. Epema, A. Datta, T. Higashino, A. Iamnitchi, M. Barcellos, P. Felber, and E. Riviere, "Edge-centric computing: Vision and challenges," *ACM SIGCOMM Comput. Commun. Rev.*, vol. 45, no. 5, pp. 37–42, 2015.

[142] W. Shi, J. Cao, Q. Zhang, Y. Li, and L. Xu, "Edge computing: Vision and challenges," *IEEE Internet Things J.*, vol. 3, no. 5, pp. 637–646, Oct. 2016.

[143] P. Kayal and J. Liebeherr, "Distributed service placement in fog computing: An iterative combinatorial auction approach," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2019, pp. 2145–2156.

[144] N. Sharghivand, F. Derakhshan, and L. Mashayekhy, "QoS-aware matching of edge computing services to Internet of Things," in *Proc. IEEE 37th Int. Perform. Comput. Commun. Conf. (IPCCC)*, Nov. 2018, pp. 1–8.

[145] A. Zavodovski, S. Bayhan, N. Mohan, P. Zhou, W. Wong, and J. Kangasharju, "DeCloud: Truthful decentralized double auction for edge clouds," in *Proc. IEEE 39th Int. Conf. Distrib. Comput. Syst. (ICDCS)*, Jul. 2019, pp. 2157–2167.

[146] Y. Abdulsalam and M. S. Hossain, "COVID-19 networking demand: An auction-based mechanism for automated selection of edge computing services," *IEEE Trans. Netw. Sci. Eng.*, early access, Sep. 24, 2020, doi: 10.1109/TNSE.2020.3026637.

[147] N. C. Luong, Y. Jiao, P. Wang, D. Niyato, D. I. Kim, and Z. Han, "A machine-learning-based auction for resource trading in fog computing," *IEEE Commun. Mag.*, vol. 58, no. 3, pp. 82–88, Mar. 2020.

[148] W. Sun, J. Liu, Y. Yue, and H. Zhang, "Double auction-based resource allocation for mobile edge computing in industrial Internet of Things," *IEEE Trans. Ind. Informat.*, vol. 14, no. 10, pp. 4692–4701, Oct. 2018.

[149] H. Zhang, F. Guo, H. Ji, and C. Zhu, "Combinational auction-based service provider selection in mobile edge computing networks," *IEEE Access*, vol. 5, pp. 13455–13464, 2017.

[150] D. Zhang, L. Tan, J. Ren, M. K. Awad, S. Zhang, Y. Zhang, and P.-J. Wan, "Near-optimal and truthful online auction for computation offloading in green edge-computing systems," *IEEE Trans. Mobile Comput.*, vol. 19, no. 4, pp. 880–893, Apr. 2020.

[151] A.-L. Jin, W. Song, P. Wang, D. Niyato, and P. Ju, "Auction mechanisms toward efficient resource sharing for cloudlets in mobile cloud computing," *IEEE Trans. Services Comput.*, vol. 9, no. 6, pp. 895–909, Nov./Dec. 2015.

[152] Q. Wang, S. Guo, J. Liu, C. Pan, and L. Yang, "Profit maximization incentive mechanism for resource providers in mobile edge computing," *IEEE Trans. Services Comput.*, early access, Jun. 24, 2020, doi: 10.1109/TSC.2019.2924002.

[153] F. Mashhadi, S. A. S. Monroy, A. Bozorgchenani, and D. Tarchi, "Optimal auction for delay and energy constrained task offloading in mobile edge computing," *Comput. Netw.*, vol. 183, Dec. 2020, Art. no. 107527.

[154] X. Peng, K. Ota, and M. Dong, "Multiattribute-based double auction toward resource allocation in vehicular fog computing," *IEEE Internet Things J.*, vol. 7, no. 4, pp. 3094–3103, Apr. 2020.

[155] A. Kiani and N. Ansari, "Toward hierarchical mobile edge computing: An auction-based profit maximization approach," *IEEE Internet Things J.*, vol. 4, no. 6, pp. 2082–2091, Dec. 2017.

[156] G. Gao, M. Xiao, J. Wu, H. Huang, S. Wang, and G. Chen, "Auction-based VM allocation for deadline-sensitive tasks in distributed edge cloud," *IEEE Trans. Services Comput.*, early access, Mar. 4, 2019, doi: 10.1109/TSC.2019.2902549.

[157] T. Bahreini, H. Badri, and D. Grosu, "An envy-free auction mechanism for resource allocation in edge computing systems," in *Proc. IEEE/ACM Symp. Edge Comput. (SEC)*, Oct. 2018, pp. 313–322.

[158] A. Tasiopoulos, O. Ascigil, I. Psaras, S. Toumpis, and G. Pavlou, "FogSpot: Spot pricing for application provisioning in edge/fog computing," *IEEE Trans. Services Comput.*, early access, Jan. 24, 2019, doi: 10.1109/TSC.2019.2895037.

[159] A. Aggarwal, N. Kumar, D. P. Vidyarthi, and R. Buyya, "Fog-integrated cloud architecture enabled multi-attribute combinatorial reverse auctioning framework," *Simul. Model. Pract. Theory*, vol. 109, May 2021, Art. no. 102307.

[160] Y. Zhang, C. Lee, D. Niyato, and P. Wang, "Auction approaches for resource allocation in wireless systems: A survey," *IEEE Commun. Surveys Tuts.*, vol. 15, no. 3, pp. 1020–1041, Nov. 2012.

[161] M. Yokoo, Y. Sakurai, and S. Matsubara, "The effect of false-name bids in combinatorial auctions: New fraud in internet auctions," *Games Econ. Behav.*, vol. 46, no. 1, pp. 174–188, Jan. 2004.

[162] M. Yokoo, Y. Sakurai, and S. Matsubara, "The effect of false-name declarations in mechanism design: Towards collective decision making on the internet," in *Proc. 20th IEEE Int. Conf. Distrib. Comput. Syst.*, Apr. 2000, pp. 146–153.

NAFISEH SHARGHIVAND received the B.Sc., M.Sc., and Ph.D. degrees in computer engineering from the University of Tabriz, Tabriz, Iran. She is currently a Lecturer with the Department of Computer Engineering, University of Tabriz. Her research interests include cloud and edge computing, the Internet of Things, game theory, mechanism design, machine learning, and multiagent systems.

FARNAZ DERAKHSHAN received the Ph.D. degree in computer engineering from the University of Liverpool, in 2008. She is currently an Assistant Professor with the Department of Computer Engineering, University of Tabriz. Her research interests include multiagent systems, game theory, and the Internet of Things.

NAZLI SIASI (Member, IEEE) received the M.S. and Ph.D. degrees in electrical engineering from the University of South Florida, FL, USA. She is currently an Assistant Professor with Christopher Newport University, Newport News, VA, USA. Her research interests include virtualization, and cloud and edge computing.

• • •