Future Generation Computer Systems 89 (2018) 286–299

Contents lists available at ScienceDirect

Future Generation Computer Systems

journal homepage: www.elsevier.com/locate/fgcs

An online auction mechanism for cloud computing resource allocation and pricing based on user evaluation and cost

Jixian Zhang [a], Ning Xie [a], Xuejie Zhang [a], Weidong Li [b,*]

[a] School of Information Science and Engineering, Yunnan University, Kunming, Yunnan, 650500, China
[b] School of Mathematics and Statistics, Yunnan University, Kunming, Yunnan, 650500, China

HIGHLIGHTS

- We propose a new auction-based mathematical model.
- We propose a truthful online auction mechanism based on user evaluation and cost.
- We propose a fast heuristic algorithm to solve resource allocation problems.

ARTICLE INFO

Article history:
Received 7 August 2017
Received in revised form 30 October 2017
Accepted 20 June 2018
Available online 3 July 2018

Keywords:
Cloud computing
User evaluation
Truthful online auction mechanism
Multi-requirements
Resource allocation
Dynamic pricing

ABSTRACT

Auction-based virtual resource allocation is a major problem for cloud computing. However, previous studies have typically equated virtual resources with ordinary items and designs corresponding to specific auction mechanisms, which cannot truly reflect the characteristics of the virtual resource. This inability to reflect the virtual resource characteristics arises because the users have only the right to use resources without ownership, the use of virtual resources has significant impacts on the social welfare and revenue of the resource provider, and the virtual resource usage will generate operating costs. Considering the above issues, this work proposes a truthful online auction mechanism based on user evaluation and cost and applies it to the allocation and pricing of cloud computing virtual resources. Based on a pay-as-you-go model, users can submit multiple requirements at any time, but only one requirement can be satisfied, which is known as a multi-requirement, single-minded scenario. We prove that the resource providers can obtain increased social welfare and guarantee that the mechanism is truthful. For the resource allocation problem, we propose a heuristic algorithm to obtain the allocation result quickly and to maximize the social welfare of the cloud resource provider. The payment algorithm accounts for the interests of both the users and resource provider based on the dichotomy. We experimentally analyze the solution in terms of social welfare, execution time, resource utilization and users served.

1. Introduction

In 2016, the global cloud computing market reached 240 billion US dollars, with many enterprises and individuals submitting computing and storage tasks to the cloud. Presently, the auction-based mechanism has become the primary method of cloud resource allocation. In the auction-based mechanism, resource providers are willing to add redundant resources to the market using a less fixed price manner, such as Amazon EC2 [1]. The advantages of this mechanism are that it can improve the revenue and resource utilization of resource providers and meet more user requirements. The introduction of the auction-based mechanism breaks the traditional resource allocation system by replacing the original fixed price model with user bidding as the dominant allocation of resources.

However, the auction-based resource allocation mechanism used in existing research has several limitations. First, the auction mechanism does not account for the use of virtual resources and their costs, which impact the results of resource allocation. Second, the majority of current auction mechanisms are offline or static. Allocation and payment procedures start after all of the user requirements are known, and the users must wait for a long period of time. Third, the current auction mechanism allows each user to submit only one requirement, which makes it difficult for the user to determine whether his requirement can be satisfied. Finally, the lack of consideration of the user's evaluation is not conducive to the final payment price calculation.

* Corresponding author.
E-mail address: weidong@ynu.edu.cn (W. Li).

https://doi.org/10.1016/j.future.2018.06.034
0167-739X/

Considering the above limitations, the auction-based cloud computing resource allocation and pricing mechanism is faced with the following three challenges:

(1) Mechanism design: The auction mechanism design should consider the incentives for both cloud providers and users with the goal of maximizing the social welfare for the cloud providers. A mechanism design should ideally be reasonable, truthful (incentive compatibility) and online. Reasonable means the social welfare-maximized theoretical model should be as accurate and practical as possible. The existing research equates virtual resources with the ordinary item and design of the auction mechanism, which cannot truly reflect the characteristics of the virtual resource. First, the users have only the right to use resources, no ownership, and the use of virtual resources will affect the social welfare and revenue of resource providers. Second, the virtual resource usage will generate operating costs, and thus, these factors should be considered in the mechanism design. Truthful means the auction mechanism should prevent users from submitting untruthful bids to acquire increased revenue. Several previous studies have considered the difference between feasible solutions and the optimal solution while ignoring truthfulness, which makes the auction mechanism lose some meaning. Online means that the mechanism should enable the dynamic allocation of resources. Previous studies regarding mechanism design have largely been static (offline), which means that there exists a wait in obtaining all users' requirements before solving the allocation and payment. However, in practice, when a user arrives and submits a resource requirement that is unpredictable, it is not possible to obtain all user requests in advance. A user who participates in the auction wants to use the resource in a pay-as-you-go manner; thus, the online auction mechanism design will be more suitable for real environments. Therefore, the design of a reasonable, truthful, and online auction mechanism is a significant challenge.

(2) Resource allocation: All users' requirements cannot be satisfied in the auction process. Therefore, the main goal of the allocation is to allocate limited resources to the users in such a way that resource providers obtain the greatest social welfare and revenue. A single-type resource allocation problem has been proven to be NP-hard and thus cannot be solved in polynomial time. In online multi-type resources allocation problems, different resources (e.g., CPU, memory, and bandwidth) must be taken into account, and the arrival and execution times of user requirements must be considered. Thus, online multi-type resource allocation problems are strongly NP-hard.

(3) Payment: In many current auction rules, the allocation winner users do not need to pay the bid price to the cloud provider. The payment price must consider both the cloud provider and users. Nejad M noted that if the payment price was based on the critical value [2], then it would be possible to ensure that the auction mechanism is truthful. The Vickrey–Clarke–Groves (VCG) mechanism [3] provides an ideal method to obtain the critical value, but it does not apply to online auctions. The users' evaluation should be considered in the final payment price algorithm. The users' evaluation allows the users who have a good reputation to pay less for the allocation of resources and can also punish users who have a poor reputation by making them pay more. However, no existing research accounts for the users' evaluation.

1.1. Our contribution

The online auction mechanism allows prices to be determined by real-time supply and demand conditions while also maximizing the revenue to cloud providers. For the above problems, this paper proposes an online auction mechanism based on user evaluation and cost, which is applied to cloud computing virtual resource allocation and pricing. The mechanism will make the social welfare maximization theoretical model more accurate and reasonable. It is characterized by the user satisfying multiple requirements in online auctions and ensuring that the users' bids are truthful. We can prove that the cloud provider can obtain increased social welfare under this mechanism compared to the previous auction mechanism. Then, we propose the online virtual resource allocation and payment (OVRAP) algorithm to solve online multi-requirement, multi-type resource allocation problems. Finally, we introduce the users' evaluation into the payment price algorithm by considering the interests of cloud providers and users. The integer programming and weighted VCG algorithms are used to solve the static optimal solution. In the experiment, the optimal solution is compared to the results obtained using the OVRAP algorithm in terms of execution time, social welfare, resource utilization and served users. The experimental results demonstrate that the OVRAP algorithm can obtain feasible solutions close to the optimal solution while satisfying the truthfulness property.

1.2. Related works

The existing research on cloud resource allocation is mainly focused on auction mechanism design, resource allocation algorithms, payment algorithms and cloud frameworks.

Auction mechanism design. The current research on auction mechanisms is mainly based on the premise of a static, single-type resource with a single requirement. The goal of the auction mechanism is to maximize the social welfare for resource providers such that their revenue is maximized. One of the most important features of an auction mechanism is truthful responses (strategy proof or incentive compatibility) [4], That is, the users cannot obtain a higher revenue in the auction through untruthful behavior. In the literature, Mashayekhy and Nejad et al. [2,4] proposed a static, truthful auction mechanism and provided an approximate algorithm for multi-resource allocation; however, under this mechanism, the user can submit only one requirement in the auction. Sandholm et al. [5] proposed a heuristic algorithm and validated it for combinatorial auctions. Wu et al. [6] transformed a combination auction into a winner decision problem (WDP) and used the clique-based exact method to obtain the exact solution. Lai et al. [7] obtained the optimal solution of a combined auction by utilizing a monotone branch-and-bound search method. Mashayekhy et al. [8,9] used the approximate algorithm OVMAP to solve the problem of online resource allocation and payments. Shi et al. [10] transformed the online auction problem into a continuous static auction problem for a period of time. They used the primal–dual algorithm to obtain the approximate resource allocation solution, where the algorithm was proven to have a competitive ratio of $e + \frac{1}{e-1}$. Zhang et al. [11,12] proposed an incentive-compatible online cloud auction mechanism (COCA), which can ensure truthfulness with heterogeneous and online user demand, Moreover, theoretical analysis shows that the worst-case performance of COCA can be well bounded. The main problems with the current auction mechanism design are as follows. First, virtual cloud computing resources cannot be completely equivalent to ordinary items because the use of virtual resources will affect the social welfare and revenue of resource providers and the virtual resource usage generates operating costs. These two factors should be considered in the mechanism design. Second, the static auction mechanism must know all user requirements in advance and cannot satisfy the users who want to use resources by a pay-as-you-go model. Third, existing online auction mechanisms can submit only one requirement to the users, which is not conducive to maximizing social welfare for resource providers.

Resource allocation. Virtual resource allocation problems in cloud computing are similar to multidimensional knapsack problems (MKPs), which are strongly NP-hard [13]. These problems