A Blockchain Platform of Crowdsensing for Cloud Re-allocation

Hong Jiang[1,a], Hui Chen[2,b,*],Yuxin Qiu[3,c], Xin Ma[3,d]and Zhe Zhang[4,e]

{jiang.hong@lpunetwork.edu.ph[a], Correspondence: 2533507448@qq.com[b,*] q13850392637@126.com[c], 2222902019@cnu.edu.cn[d], 1995asd@gmail.com[e]}

Claro M. Recto Academy of Advanced Studies, Lyceum of the Philippines University, Manila 1002 , Philippines[1]
School of Economics and Management,Anyang University,Henan,453599,China[2]
School of Management, Capital Normal University, Beijing 100089, China[3]
School of Management, Henan University of Technology, Henan 450001, China[4]

Abstract: Cloud computing is a technology facilitating broad access to diverse computing services, predominantly reliant on centralized mechanisms for resource allocation. This study introduces a decentralized platform, harnessing crowdsensing and machine learning, for cloud provisioning and pricing. It involves a blockchain-based trading platform, enabling sellers (primary users) to auction their cloud resources to buyers (secondary users) while considering buyer reputations. To incentivize crowd sensors in gathering and sharing information about available cloud resources, an incentive mechanism is implemented. The pricing is estimated through a supervised machine learning algorithm, specifically linear regression, incorporating critical values and the Vickrey-Clarke-Groves (VCG) algorithm. Results indicate that supervised linear regression is a superior approach for enhancing overall utilization. This research presents a robust methodology for integrating cloud computing and machine learning in practical pricing decisions.

Keywords: Cloud Computing, Blockchain, Crowdsensing, Incentive Mechanism, Supervised Linear Regression

1. Introduction

With the rapid development of cloud computing technology, it has become increasingly important as a technology that facilitates broad access to a wide range of computing services. The centralised resource allocation mechanism of cloud computing is a key factor that plays an important role in its commercial applications [1], which provides users with efficient and flexible computing resource management. However, in order to further drive innovation in the field of cloud computing, this study introduces an emerging technology based on a decentralised platform that aims to provide more intelligent cloud services and pricing by integrating crowdsensing and machine learning.

With the development of cloud computing, research on combining applications with blockchain and introducing machine learning algorithms to provide smart pricing is becoming increasingly popular. In the early days of the rise of cloud computing, researchers mainly studied the advantages and limitations of centralised resource allocation mechanisms [2]. Along with the rise of blockchain technology, researchers began to think about how to apply it in cloud computing,

MSIEID 2023, December 08-10, Guangzhou, People's Republic of China

DOI 10.4108/eai.8-12-2023.2344773

especially the potential applications in secure transactions [3,4]. Crowd sensing is an important channel for participants to access data and services, and incentives are a means of achieving crowd sensing, which addresses the utility maximisation problem faced by each of the providers and participants[5].In order to motivate and attract participants to cloud computing, more and more people have begun to study crowd-sensing and incentives [6,7],Incentive mechanism based on reputation theory has also been extensively studied [8].With the rapid development of machine learning, its application in the field of cloud computing has attracted wide attention, especially the application of supervised linear regression algorithm in cloud computing[9], such as the classification and analysis of crowdsourced information,pricing, etc.[10]. On the other hand, in order to achieve computational efficiency, bid authenticity and competition fairness, auction pricing model is often applied to cloud resource auctions [11,12].

This study mainly relies on a blockchain-based trading platform that connects sellers and buyers by integrating cloud computing and machine learning, enabling sellers (primary users) to auction their cloud resources to buyers (secondary users) while considering the reputation of the buyers. It also motivates crowd sensors to collect and share cloud resource information through incentives to further improve resource utilisation. In terms of pricing, pricing is estimated through supervised machine learning algorithms, especially linear regression, combined with critical value and Vickery-Clark-Groves (VCG) algorithms. With this research, we present a powerful approach for integrating cloud computing and machine learning in real-world pricing decisions, bringing a higher level of intelligence and efficiency to the field of cloud services.

This paper is structured as follows: Section 2 depicts the decentralized cloud trading platform. Section 3 proposes a machine learning mechanism for cloud pricing: the supervised linear regression. Section 4 illustrates the pricing algorithm. Section 5 addresses and verifies the procedures for training and testing the supervised linear regression. Section 6 is the conclusion.

2. The Blockchain-based Cloud Trading Platform

2.1 Incentive Mechanism in Crowdsensing

Crowdsensing is an interactive and participatory sensing network formed through devices such as smartphones, laptops and wearable devices, with participants publishing information. It enables data collection, information analysis and resource sharing, and is an important way to access environmental data and services. Incentives Mechanism participants to engage in sensing tasks by designing methods that encourage the provision of high quality and reliable information. The cloud server solves the problem of information authenticity by recruiting the lowest cost sensors whose incentives should motivate high quality data, which can be solved by using reputation values. The server sets the reputation value according to the difficulty of the task to ensure that the task is completed and the smart contract automatically releases the reward to protect the sensor's rights and interests.

2.2 The Procedure of Cloud Sensing

In this study, the blockchain platform is an alliance system consisting of primary user (seller PU), secondary user (buyer SU), cloud server S, and cloud sensor CS. The perception task is processed by S, which sets a credit value according to the task difficulty to ensure that the task is completed. The blockchain-based cloud sensing system has a physical layer, a transport layer and an application layer. The physical layer includes CS and PU, the transport layer connects the physical layer and the application layer, and the blockchain network is located at the transport layer. Transactions can be added to the block after verifying the original sensing data. The application layer consists of SU and S, where SU looks for the idle PU and S handles the authenticated SU requests.Its operation process is as follows:(1) Multiple SUS send a request to S, and the winner pays S's service fee and PU access fee.(2) The cloud server determines the cloud-aware task according to the request and reputation value. The higher the reputation value, the more difficult the task and the more reward.(3)S issues tasks with smart contracts, specifying mission credit and rewards for each SU.(4)CS provides quotes for different tasks, and CS with high reputation can perform more tasks.(5)S selects a certain number of CS to complete the task, broadcasts the results and records them on the blockchain.(6)CS completes the processing and transmits the unprocessed data to the miner, who processes the data and records it on the block-chain.(7)S rewards CS, SU can use cloud services. Smart contracts automatically reward CS and adjust reputation values after verification. When the task is complete, the transaction is closed.

3. The Supervised Linear Regression Algorithms

Aiming to distribute the sensing rewards fairly, the rewards to each CS are based on the reputation value; a CS will obtain more rewards if the sensor works more. This will effectively stimulate CS's willingness to join in cloud sensing.

3.1 A Hedonic Regression Model of Cloud Pricing

The pricing metrics include the reputation, the cloud features, and the time of the same item from different customers' bids. The proposed hedonic regression model is shown in equation (1):

$$P_{it}= f\,(R_{it}, C_{it}, T_t), \tag{1}$$

Where, P_{it} is price, R_{it} is the reputation of a certain cloud service, C_{it} is cloud feature attributes, T_{it} is time trend, i represents a certain cloud service, t represents a specific time.

3.2 The Allocation Model of Cloud

This paper is designed based on the auction pricing mechanism.Auction design consists of two components: resource allocation and price estimation, and a reliable auction needs to fulfil both truthfulness and accuracy requirements. Truthfulness means that users cannot benefit from false bids, and accuracy requires that the allocation strategy is the optimal solution or at least very close to the optimal solution [13]. In cloud computing, resource allocation is an NP problem. If possible, algorithms will be used to obtain the optimal solution; otherwise, approximate or heuristic algorithms will be used as a viable solution to the NP-hard problem [14]. The next sections describe and explain cloud reallocation and pricing in detail.

Assume that there are m users in the set U. U = {1, 2, …, m}, user $i \in U$. User i proposes resources requests $K_i^{k_r}$. k_r is a certain cloud resource, r = 1, 2, …, n. User i's resource request $K_i^{k_r} = (k_r, R_r^i)$, R_r^i is the reputation value of user i on resource k_r.

Definition 1. Monotonicity: If the request submitted by a user (B_r^i) can be allocated, any other request ($B_r^{i'}$) from the same user will be allocated on the condition of $B_r^{i'} > B_r^i$. This is the monotonicity of resource allocation.

Definition 2. Critical Value: If the request submitted by a user is allocated, there exists a critical value (CV_r). If the user bid $B_r^i > CV_r$, the request can be satisfied; otherwise, it cannot be satisfied.

Lemma 1. If the distribution of resources in the auction mechanism is consistent with the principle of monotonicity, and the ultimate price adheres to the critical value, then the mechanism can be considered truthful.

The VCG auction mechanism is realistic because it relies on the optimal allocation solution. However, the final price obtained using VCG cannot be computed in polynomial time. In this study, supervised learning classification and regression are used to design cloud resource allocation. The basic principle is to select some requests from all user requests and estimate the optimal allocation and price. By fitting the optimal policy, the model is trained to be applied to all users to predict the resource allocation. In this section, we design resource allocation based on linear regression algorithm (LN) and construct price algorithm based on critical value theory.

3.3 The LN Algorithm of Cloud Allocation

In the auction design, the hypothesis function ($h_\theta(k^i)$)(equation (2)) is constructed according to the SU's request for different resources.

$$h_\theta(k^i) = \theta_0 + \theta_1 k_1^i + \theta_2 k_2^i + \cdots + \theta_n k_n^i + \theta_{n+1}\sqrt{k_1^i} + \theta_{n+2}\sqrt{k_2^i} + \cdots + \theta_{2n}\sqrt{k_n^i} \quad (2)$$

The goal of the supervised linear regression is to find the rules that SU wins, it is $\theta_{LN} = (\theta_0 \theta_1 \dots \theta_{2n}) \in \mathbb{R}^{2n+1}$. The optimal strategy and the final price of each winning SU will be calculated. p_F^i is the final price from SU_i. If SU_i wins, $p_F^i > 0$, otherwise, $p_F^i = 0$.

According to the optimal solution and all SUs' requests, the matrix of SU request is $K = [k^1, k^2, \dots, k^m]^T$. The vector of the optimal allocation is $X = (x^1, x^2, \dots, x^m)^T$. The vector the SUs' bidding is $B = (b^1, b^2, \dots, b^m)^T$. The vector of the final price is $P = (p^1, p^2, \dots, p^m)^T$. We have F (θ),see equation(3),

$$F(\theta) = \frac{1}{2m}\left[\sum_{i=0}^{m} x^i \left(h_\theta(k^i) - p^i\right)^2 + \lambda \sum_{i=0}^{m} \theta_j^2\right] \quad (3)$$

To get Min $F(\theta)$, θ can be solved by the normal equation based on the above function. Hence,see equation(4),

$$\theta = (K^T K - \lambda L)^{-1} K^T P$$
$$L = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix} \tag{4}$$

Then, we use the Sigmoid Function to estimate whether an SU's bid wins or not.

$$P_i^W = \frac{1}{1+e^{-(b^i - h_\theta(k^i))}},\ W^i \in (0,1) \tag{5}$$

P_i^W (see equation(5)) is the probability of an SU winning the bid. If an SU wants to win the bid, $b^i > h_\theta(k^i)$. Hence, $W^i \geq 0.5$ means that an SU has a strong chance of winning the resources being bided for.

4. The Pricing Algorithms

4.1 The Modified VCG Pricing Algorithm

A reliable auction mechanism will ensure that the price paid by buyers is optimal. The final price proposed in this study is based on the critical value. The estimated winning set W^i is used to calculate the final price. Specifically, p_F^{i+} is the maximum and p_F^{i-} is the minimum. When $p_F^{i+} - p_F^{i-} > \delta$, the final price that an SU needs to pay is p_F^{i+}.

4.2 Proof of Truthfulness

Lemma 2. The supervised linear regression is monotonicity.

Proof. Suppose P_L^W is the probability that the last SU can be satisfied for the expected request. Based on the estimation function (Last Function), we have equation(6)

$$P_i^W = \frac{1}{1+e^{-(b^i - h_\theta(k^i))}} > P_L^W \tag{6}$$

So,

$$b^i > h_\theta(k^i) - \ln[(1 - P_L^W)/P_L^W] \tag{7}$$

This indicates that an SU will win the bidding if the SU's bid satisfies the above function. If an SU's bid is greater than b^i(see equation(7)), the SU can be guaranteed to obtain the requested cloud resources. Definition 1 presents the identical tone of monotonicity.

Lemma 3. The final price algorithm satisfies the theory of critical value.

Proof. When $b^i > p_F^{i+}$, an SU will win the bid; when $b^i < p_F^{i-}$, an SU will lose the bid. There exists a critical value (Definition 2). Thus, the final price that the SU needs to pay is p_F^{i+}, if $p_F^{i+} - p_F^{i-} > \delta$.

Theorem 1. The cloud allocation design proposed is truthfulness.

Proof. According to Lemma 1, the supervised linear regression algorithm satisfies resource allocation monotonicity.Also, the final price algorithm satisfies the theory of critical value.Therefore, the proposed cloud allocation design is truthfulness.

5. The Procedure of Training and Testing

5.1 Resource Allocation Algorithm Training

In the training tests, SU requests are simulated using the DAS-2 open source dataset. Experimental platform: Intel Core i7 6500U CPU, 16GB memory, 1TB DDR storage. The experimental conditions include (1) simulating SU requests using CPU, memory, and storage information; (2) randomly generating integers from 1 to 100 to simulate bids with preset resource reputation values; (3) solving the optimal allocation using IBM CPLEX; (4) solving the optimal price to pay based on the VCG mechanism; and (5) programming to implement the algorithm. The test selected 5000 records as SU requests and generated corresponding bids. We calculated the resource density of each SU (d_i),see equation(8),

$$d_i = \frac{b_i}{\sqrt{\sum_{r=1}^{n}(\frac{1}{c_r} * k_r^i)}}, \forall i = 1,2, \ldots, m \tag{8}$$

Based on resource density, SU requests are sorted in descending order to form total samples. The system samples 500 samples at a time, with a total of 20 sample sets, of which 17 are training sets and 3 are cross-validation sets. The predictive models from the training set are substituted into the validation set in the test and the best model is selected to estimate all SU requests. The models were evaluated using Prediction Accuracy (PA) and Prediction Error (PE) with PA + PE = 1. PA is defined as the number of SUs with the same feasible and optimal solution divided by the total number of SUs,it's calculation is shown in equation(9),

$$P_i^W = \begin{cases} 1, & P_i^W \geq V_L \\ 0, & P_i^W < V_L \end{cases}$$
$$PA = \frac{1}{m} \sum_{i=1}^{m} (P_i^W = x_i) \tag{9}$$

V_L represents the predicted value of the last allocated SU, P_i^Wrepresents the probability of a SU winning in the resource reallocation. The greater the value P_i^W, the higher the probability that a SU wins the bid. To solve θ in LN, the coefficient λ needs to be adjusted appropriately to ensure higher prediction accuracy in the cross-validation set. Figure 1 shows the change in the prediction error rate with λ when fitting the model in the training set to the cross-validation set. For LN, when λ=3, the prediction error rate is the smallest at 1.4% (Figure 1(a)). Similarly, for LG, when λ= {30, 40}, the prediction error rate is the smallest at 3.2%, (Figure 1(b)).

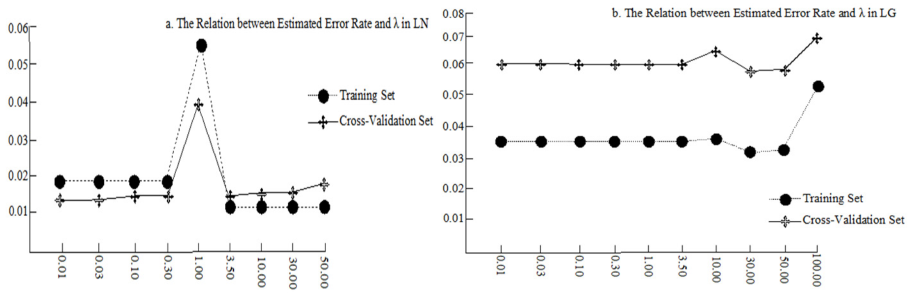

Figure 1. Comparison of Estimated Error Rate between LN and LG.

5.2 Comparison of training Time among the Three Algorithms

To verify the effect of the LN algorithm proposed in this paper, it is compared with the Algorithm of Logistic Regression (LG) and the Algorithm of Support Vector Machine (SVM) .The LG algorithm does not compute the final price, which is an important factor of the hypothesis function ($h_\theta(k^i)$) that is constructed according to the SU's request for different resources. $\theta_{LG} = (\theta_0\theta_1 \ldots \theta_{n+1}) \in \mathbb{R}^{n+2}$, the conditions are shown in equation(10):

$$f_\theta(k^i) = \theta_0 + \theta_1 k_1^i + \theta_2 k_2^i + \cdots + \theta_n k_n^i + \theta_{n+1}(b^i)^2$$

$$g\,(z) = \frac{1}{1+e^{-z}} \tag{10}$$

$$h_\theta(k^i) = g(f_\theta(k^i))$$

The function F (θ) is shown in equation(11),

$$F(\theta) = \frac{1}{m}\sum_{i=0}^{m}[-x^i \lg\left(h_\theta(k^i)\right) - (1-x^i)\lg\,(1-h_\theta(k^i))] + \frac{\lambda}{2m}\sum_{i=0}^{n}\theta_j^2 \tag{11}$$

Then, the estimated winning probability is equation (12),

$$P_i^W = \frac{1}{1+e^{-(h_\theta(k^i))}}, W^i \in (0,1) \tag{12}$$

The SVM algorithm was also used for comparison, see Equation (13),

$$\min C\sum_{i=1}^{m}[x^i COST_1(\theta^T f^i) + (1-x^i) * COST_0(\theta^T f^i)] + \frac{1}{2}\sum_{j=1}^{m}\theta_j^2$$

$$COST_1(x) = \begin{cases} -\dfrac{11}{16}x + \dfrac{11}{16}, & x < 1 \\ 0, & x \geq 1 \end{cases} \tag{13}$$

$$\mathrm{COST}_0(x) = \begin{cases} \frac{11}{16}x + \frac{11}{16} & , x > -1 \\ 0, & x \le -1 \end{cases}$$

$$x^i = \begin{cases} 1, & \theta^T f^i \ge 0 \\ 0, & \theta^T f^i < 0 \end{cases}$$

$$f^i = (1 f_1^i f_2^i \dots f_m^i)$$

$$f_j^i = \exp\left(-\frac{\left|R^i - R^j\right|^2}{2\sigma^2}\right)$$

$$\forall j = 1, 2, \dots, m$$

When using SVM to predict resource reallocation, C and σ are important parameters. C indicates the accuracy of the estimation boundary, and σ indicates the range of influence of each value. The advantage of SVM is that when a small sample training set is used for training, it can also obtain good estimation accuracy,

$$\theta = (\theta_0 \theta_1 \dots \theta_m) \tag{14}$$

According to SU requests, the estimation function is shown in equation(15),

$$P_i^W = \theta_0 + \theta_1 f_1^i + \theta_2 f_2^i + \dots + \theta_m f_m^i,\ P_i^W \in \mathbb{R} \tag{15}$$

We compared the training times among the three algorithms (Figure 2). When the training set size was the same, the SVM's training time was the longest, and the LG's time was the shortest. The LN's speed was in the middle. Among the three algorithms, the LN had the smallest error rate. The main reason is that its cost function has the characteristics of the optimal payment price. Compared to LG and SVM, an extra factor needs to be considered in LN to give a higher prediction accuracy. Overall, the proposed LN was found to be qualified and can be implemented in the decentralized auction design.

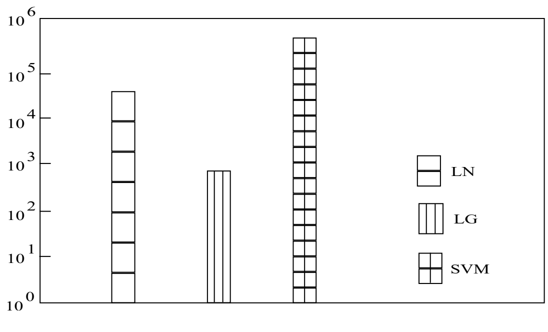

Figure 2. Comparison of Training Time among the Three Algorithms

5.3 Analysis of Resource Allocation Forecast Results

After obtaining the optimal prediction models of the three algorithms, 5 test sets were randomly generated, for instance, 1000, 2000, 3000, 4000, and 5000 SU requests. The social welfare obtained by the two algorithms (LG and SVM), based on supervised learning, was lower than the proposed optimal strategy, but was very close (Figure 3). This shows that the optimal allocation solution has specific pattern. It can be classified by a supervised learning algorithm and can be fitted by a regression model.

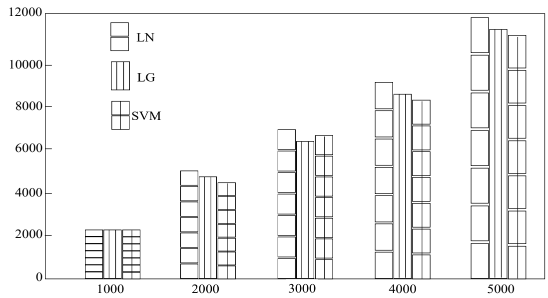

Figure 3. Comparison of Social Welfare

Figure 4 shows the prediction accuracy of different algorithms relative to the optimal allocation. The accuracy can reflect the fairness of an important indicator of the algorithm in resource reallocation. All predictions based on supervised learning algorithms had a high accuracy rate (above 95%). Among them, the accuracy of the LN algorithm was above 97%, and the proposed optimal strategy had the highest accuracy among all of the testing sets.

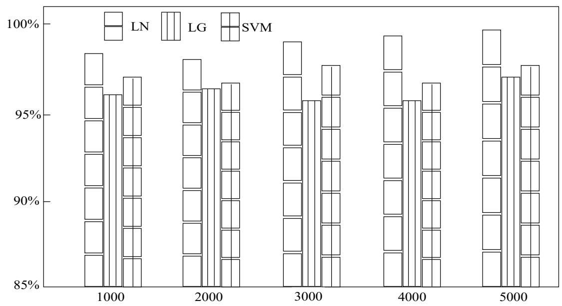

Figure 4. Comparison of Prediction Accuracy

Figure 5 shows three resource utilizations in different algorithms, given the resource capacities of the CPU, the RAM, and the storage. Based on the results, LN and LG had similar good performances, but the optimal strategy performed the best among the three resource utilizations -- in detail, CPU (100%), RAM (60%), and storage (100%).

The supervised linear regression (LN) performed very well in the test. Consistent with the previous theoretical study, the hypothesis function of the linear regression had the characteristics of the optimal payment. Also, the linear regression had the extra dimension of variable in the calculation. Hence, the accuracy of the prediction was higher.

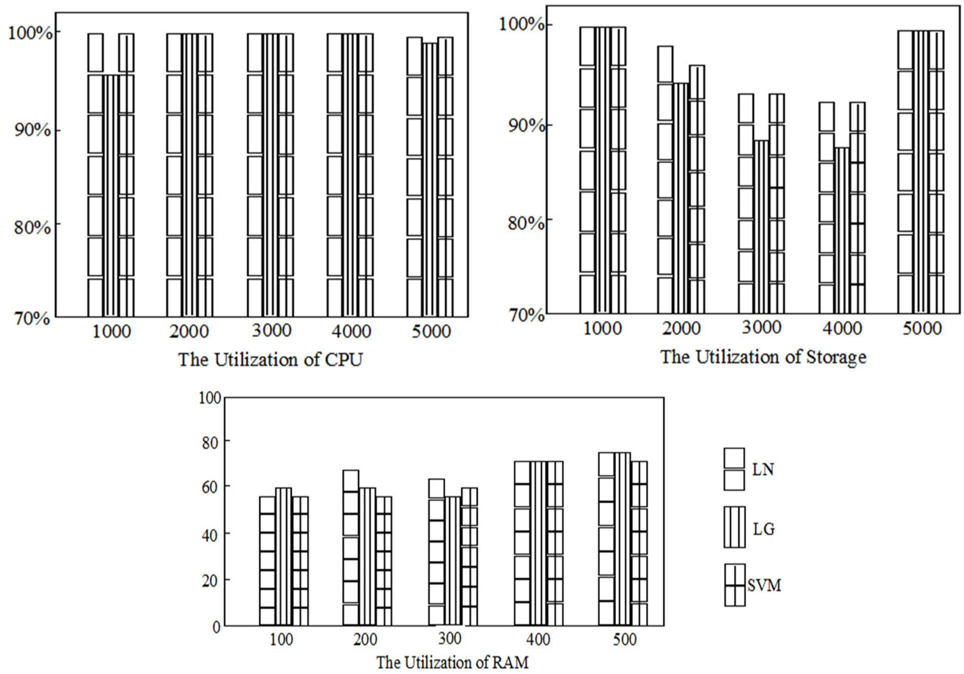

Figure 5. Comparison of Resources Utilization

6. Conclusion

Key factors in the study of cloud computing resource allocation and pricing include user competition and reputation value. Future research should explore different metrics, taking into account NP difficulty and computational complexity. This study uses reputation value as a pricing metric but stresses the potential importance of security and network externalities. To improve algorithmic accuracy, more parameters are introduced, although this may increase complexity. Blockchain-based transaction mechanisms encounter challenges such as high costs, limited data storage, inefficient communication, and platform selection. Our research concentrates on reallocating cloud resources to secondary users via a blockchain-based crowdsensing system, which delivers adaptable offers. By exploring the decentralised crowdsensing process, reputation incentives, and implementing supervised linear regression algorithms, we have effectively converted the issue of redistributing and pricing cloud resources into a problem of training and classification. Utilising the supervised linear algorithm's parameters has enabled us to provide optimal auction design strategies and guarantee the precise selection of successful users. Overall, our proposed model demonstrates strong performance with regard to accuracy, realism, social welfare, and resource utilization, and presents promising avenues for further research in the domain of cloud computing.

Reference

[1]Vakilinia, Shahin, Mustafa Mehmet Ali, and Dongyu Qiu. "Modeling of the resource allocation in cloud computing centers." *Computer Networks* 91 (2015): 453-470.

[2]Vinothina, V. V., R. Sridaran, and Padmavathi Ganapathi. "A survey on resource allocation strategies in cloud computing." *International Journal of Advanced Computer Science and Applications* 3.6 (2012).

[3]Alouffi, Bader, et al. "A systematic literature review on cloud computing security: threats and mitigation strategies." *IEEE Access* 9 (2021): 57792-57807.

[4]Achar, Sandesh. "Cloud Computing Security for Multi-Cloud Service Providers: Controls and Techniques in our Modern Threat Landscape." *International Journal of Computer and Systems Engineering* 16.9 (2022): 379-384.

[5]Ganti, R.K.; Ye, F.; Lei, H. Mobile Crowdsensing: Current State and Future Challenges. *Ieee Commun Mag* 2011, *49*, 32-39.

[6]Zhou, Bowen, Satish Narayana Srirama, and Rajkumar Buyya. "An auction-based incentive mechanism for heterogeneous mobile clouds." *Journal of Systems and Software* 152 (2019): 151-164.

[7]Zhan, Yufeng, et al. "A survey of incentive mechanism design for federated learning." *IEEE Transactions on Emerging Topics in Computing* 10.2 (2021): 1035-1044.

[8]Li Jie, Wang Xingwei, and Liu Rui."User Reputation-Based Participatory Incentive Mechanism in Social and Community Intelligence Systems." *Computer Science and Exploration* 9.12 (2015): 1471.

[9]Zhang, Hanlin, et al. "Machine learning on cloud with blockchain: a secure, verifiable and fair approach to outsource the linear regression." *IEEE Transactions on Network Science and Engineering* 9.6 (2021): 3956-3967.

[10]Khandelwal, Veena, Anand Kishore Chaturvedi, and Chandra Prakash Gupta. "Amazon EC2 spot price prediction using regression random forests." *IEEE Transactions on Cloud Computing* 8.1 (2017): 59-72.

[11]Kang, Kai, et al. "Auction-based cloud service allocation and sharing for logistics product service system." *Journal of Cleaner Production* 278 (2021): 123881.

[12]Funaro, Liran, Orna Agmon Ben-Yehuda, and Assaf Schuster. "Efficient Multi-Resource, Multi-Unit VCG Auction." *Economics of Grids, Clouds, Systems, and Services: 16th International Conference, GECON 2019, Leeds, UK, September 17–19, 2019, Proceedings 16*. Springer International Publishing, 2019.

[13]Milgrom, Paul R., and Robert J. Weber. "A theory of auctions and competitive bidding." *Econometrica: Journal of the Econometric Society* (1982): 1089-1122.

[14] McAfee, R. Preston, and John McMillan. "Auctions and bidding." *Journal of economic literature* 25.2 (1987): 699-738.