

Research on the Efficiency of Cloud Resource Allocation and Operation Based on DEA Method

Li Wenjuan¹, Wu Jiyi^{2,3,*}, Zhang Qifei³ and Zhao Haili¹

¹*Qianjiang College, Hangzhou Normal University, Hangzhou China;*

²*Key Lab of E-Business, Hangzhou Normal University, Hangzhou, China;*

³*School of Computer Science and Technology, Zhejiang University, Hangzhou China*

**Corresponding author: CloudLab@aliyun.com*

Abstract

In order to compensate for the current lack of assessing cloud resource allocation and business operation from the economical perspective, this paper introduced the Data Envelopment Analysis (DEA) method which is one of the most famous methods in economics into this field. Based on it, it proposed several new methods to evaluate the efficiency of cloud resource allocation and cloud providers' business operation status through the way of case study.

Keywords: *Cloud platform, resource allocation, business analysis, efficiency evaluation, DEA method*

1. Introduction

Cloud computing is a new method of sharing IT resources by using the integration of a large number of decentralized or distributed resources. Clouds are able to provide users with low-cost on-demand services with an Internet based architecture and virtualized background. Undoubtedly, Cloud is promised with good prospects of further development and wide deployment [1, 2]. Since Cloud computing technology is a comprehensive re-development of traditional technologies such as Parallel computing, Grid Computing and Distributed Computing, or be more precisely, the commercial implementations of these technologies, thus it has attracted both the attention in academia and industry. In the last five years, cloud and cloud-related researches were emerging. However, from the existing research fruits, we can see that it lack the enough concern about the economic issues involved in cloud systems and seldom were researches trying to apply economics approaches to solve these problems. While as a business model using Internet to provide services, Cloud should ensure the economic interests of the service providers and also should promise the economic efficiency of cloud platforms' operation. Unfortunately, till now, very few researchers focused on the efficient allocation of cloud resources, the service performance of cloud provider's, the healthy and orderly development of cloud economic and other related issues from the perspective of economics.

In response to this phenomenon, this paper introduced the famous DEA (Data Envelopment Analysis) method into cloud computing area. And based on DEA method, it studied the effectiveness of cloud resource allocation process, and also put forward a new method for analyzing the operational benefits of cloud providers. The innovations of this paper include: (1) it firstly introduced DEA method to cloud computing realm; (2) it proposed the method for the operations analysis of cloud providers.

This paper is organized as follows: section 2 briefly introduced the main DEA methods and models, section 3 proposed an evaluation model for cloud resource allocation based on DEA, cloud business operations evaluation model based on DEA was discussed in section 4 and the last section is the conclusion.

2. DEA Method Overview [3-7]

2.1. Brief Introduction of DEA

Data envelopment analysis (DEA) method is an efficient evaluation method proposed by the famous American operations researchers COOPER and CHARNES in 1978 and it has become one of the most common and important analytical tool in the fields of Management, Systems Engineering, Evaluation and decision analysis after nearly 40 years' development. Currently, the main application areas of DEA method include: technical and economic management, resource optimization, logistics and supply chain management, risk assessment, asset restructuring, industry structure analysis, etc. Today, the researches in DEA remained strong momentum of development.

DEA method is a method for evaluating the efficiency of multiple decision units with multi-input/multi-output conditions. By using the ratio of output to input, it evaluates the economic efficiency of the measure units (or decision-making unit) and also through the comparison of a specific unit with a group of similar units providing the same services, evaluates the efficiency and performance of the specific unit. DEA method can help enterprises' managers compare different strategies, easier do decision making and in addition, find ways to improve self-efficiency through comparing with other good or bad units.

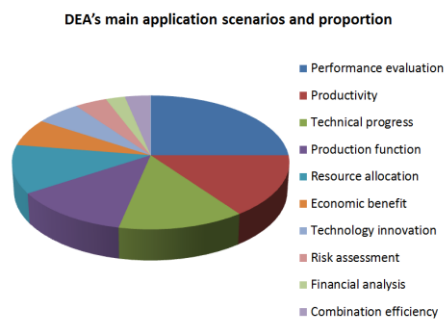


Figure 1. The Main Application Fields of DEA

When evaluation using DEA method, the general steps include:

(a) Stage one. The main tasks in this stage include: determining evaluation objectives and various factors, determining the weight of different factors and their relationship.

(b) Stage two. The main tasks in this stage include: the establishment of the index system, selection of DMUs, collect and collate relevant data and select the appropriate DEA model for the further calculations.

(c) Stage three. The main tasks in this stage include: analyzing the calculation results, distinguishing valid and invalid units, finding out the reasons for invalid units through comparison and thus proposing effective means for improvement, study and analyze the rationality of the results.

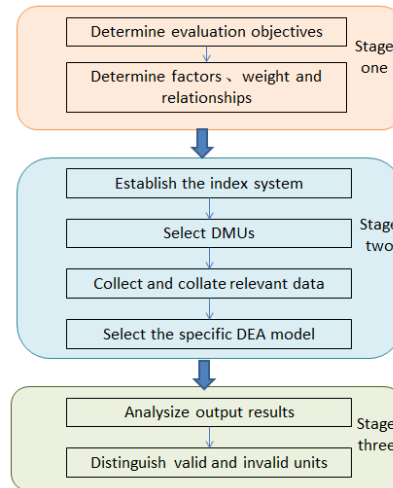


Figure 2. The General Steps of DEA Evaluation

2.2. The Main DEA Evaluation Models

The most basic DEA model is C^2R model put forward by CHARNES in the context of solving relative efficiency and this method, for the first time, extended the definition of engineering efficiency from single-input single-output to multi-input multi-output areas. Later on, researchers proposed a series of new DEA models. In general, there are mainly four DEA models to measure the effectiveness of decision units including C^2R model, BC^2 model to evaluate the relative efficiency of production technology, FG model meeting the requirement of non-increasing returns, and the last ST model meeting the requirement of non-decreasing returns. Since the latter three models can all be regarded as the promotion of C^2R model, thus we will only discuss the mechanism of C^2R model briefly.

It is assumed that there were n decision units DMUs (each DMU has m kinds of input and s types of output). We used $x(m*n)$ as input matrix (where x_{ij} represents the input amount of the j^{th} DMU to the i^{th} input species and $x_{ij}>0$), $y(s*n)$ as output matrix (where y_{rj} represents the output amount of the j^{th} DMU to the r^{th} output species and $y_{rj}>0$), Vector v (m -dimensional) represents the input weights (where v_i is the weight of the i^{th} input), Vector u (s -dimensions) represents the output weights (where u_r is the weight of the r^{th} output). Then the efficiency evaluation index of the j^{th} DMU ($1 \leq j \leq n$) can be expressed as below formula 1:

$$h_j = \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \quad (1)$$

The specific C^2R model can be constructed as follows if the efficiency of the j^{th} DMU taken into consideration under the condition that the efficiency index of all the other DMUs were less than 1.

$$(C^2R) \begin{cases} \max \frac{u^T y_j}{v^T X_j}, \\ s.t. \frac{u^T y_i}{v^T X_i} \leq 1, \quad i = 1, 2, \dots, n, \\ v \geq 0, \\ u \geq 0. \end{cases} \quad (2)$$

Since the order of DMUs did no influences to the final results, for simplicity, say $(x_j, y_j) = (x_0, y_0)$. Using Charnes-Cooper transformation, the above model can be converted into an equivalent linear programming problem.

let

$$t = \frac{1}{v^T X_0}, \quad \omega = tv \quad , \quad \mu = tu \quad (3)$$

The equivalent linear programming model is as below.

$$(P_{C2R}) \begin{cases} \max \mu^T y_0 = V_p, \\ s.t. \quad \omega^T X_i - \mu^T y_i \geq 0, \quad i = 1, 2, \dots, n, \\ \omega^T X_0 = 1, \\ \omega \geq 0, \mu \geq 0. \end{cases} \quad (4)$$

If the optimal solution of the above linear programming (P_{C2R}) meets $VP=1$, the j_0^{th} DMU is called weak DEA efficient. However, if it meets

$$\omega^0 > 0, \quad \mu^0 > 0, \quad V_p = \mu^{0T} y_0 + \delta_1 \mu_0^0 = 1 \quad (5)$$

It is called DEA efficient.

The dual programming of (P_{C2R}) is

$$(D) \begin{cases} \min \theta = V_D, \\ s.t. \quad \sum_{j=1}^n x_j \lambda_j \leq \theta X_0, \\ \sum_{j=1}^n y_j \lambda_j \geq y_0, \\ \lambda_j \geq 0, \quad j = 1, 2, \dots, n. \end{cases} \quad (6)$$

And introducing slack variable s^- and remaining variable s^+ to the above dual programming (D) , (D) can be changed to (D_{C2R}) .

$$(D_{C2R}) \begin{cases} \min \theta = V_D, \\ s.t. \quad \sum_{j=1}^n x_j \lambda_j + S^- = \theta X_0, \\ \sum_{j=1}^n y_j \lambda_j - S^+ = y_0, \\ \lambda_j \geq 0, \quad j = 1, 2, \dots, n, \\ S^- \geq 0, S^+ \geq 0. \end{cases} \quad (7)$$

Similarly, if the optimal result of (D_{C2R}) is equal to 1, the j_0 DMU is called weak DEA efficient; if at the same time, for each of the optimal result $\lambda^0 = (\lambda_1^0, \dots, \lambda_n^0)^T, s^{-0}, s^{+0}, \theta^0$, existing the condition that $s^{-0} = 0, s^{+0} = 0$, it is called DEA efficient.

3. DEA Based Cloud Resource Allocation Evaluation Model

Cloud services are generally the instant combination of virtual resources [8]. Since there are several solutions to complete the same task, in order to provide the most appropriate services while making maximize profit for providers, it is necessary to assess and evaluate the corresponding several combination solutions quickly and accurately. This paper introduced DEA method to evaluate the resource combination scheme according to the costs and benefits of related resource. The following two cases show how the specific evaluation process works.

3.1. Case Study of Cluster Efficiency Evaluation

Suppose the completion of the task T exist four cluster configuration programs in a provider, the input and output data were shown in Table 1 below:

Table 1. Resource Allocation Schemes

Evaluation parameters	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Energy cost (\$)	200	180	150	100
Depreciation (\$)	10	7	8	5
User evaluation (%)	88%	80%	75%	85%
Income (\$)	1000	840	900	600

For Cluster 1, the linear programming (P_{C2R}^2) is as follows

$$(P_{C2R}) \begin{cases} \max 0.88\mu_1 + 1000\mu_2 = V_p, \\ 200\omega_1 + 10\omega_2 - 0.88\mu_1 - 1000\mu_2 \geq 0, \\ 180\omega_1 + 7\omega_2 - 0.8\mu_1 - 840\mu_2 \geq 0, \\ 150\omega_1 + 8\omega_2 - 0.75\mu_1 - 900\mu_2 \geq 0, \\ 100\omega_1 + 5\omega_2 - 0.85\mu_1 - 600\mu_2 \geq 0, \\ 200\omega_1 + 10\omega_2 = 1, \\ \omega_1 \geq 0, \omega_2 \geq 0, \mu_1 \geq 0, \mu_2 \geq 0. \end{cases} \quad (8)$$

An optimal solution of the above (P_{C2R}) is $\mu_1=0, \mu_2=0.1, \omega_1=0, \omega_2=0.00083$ and the optimal objective function value is 0.83333. So the Cluster 1 is DEA invalid.

For Cluster 2, the linear programmin (P_{C2R}) is as follows

$$(P_{C2R}) \left\{ \begin{array}{l} \max 0.8\mu_1 + 840\mu_2 = V_p, \\ 200\omega_1 + 10\omega_2 - 0.88\mu_1 - 1000\mu_2 \geq 0, \\ 180\omega_1 + 7\omega_2 - 0.8\mu_1 - 840\mu_2 \geq 0, \\ s.t. \quad 150\omega_1 + 8\omega_2 - 0.75\mu_1 - 900\mu_2 \geq 0, \\ 100\omega_1 + 5\omega_2 - 0.85\mu_1 - 600\mu_2 \geq 0, \\ 180\omega_1 + 7\omega_2 = 1, \\ \omega_1 \geq 0, \omega_2 \geq 0, \mu_1 \geq 0, \mu_2 \geq 0. \end{array} \right. \quad (9)$$

Similarly, we could obtain one of the optimal solutions of Cluster 2 is $\mu_1=0$, $\mu_2=0.14286$, $\omega_1=0$, $\omega_2=0.00119$ and the optimal objective function value is 1. So we came to the conclusion that Cluster 2 is DEA valid.

3.2. Searching For the Best Ratio of Different Resources

Cloud provides users all kinds of services including storage, computation and network, etc. However, in most cases, tasks could not be simply classified into a certain service type, but the need for the integration of different types of services. For simple services, it is easy to calculate the efficiency according to the input to output ratio. While for integration ones, it seems a lot more difficult. This section discuss how to configure the services in order to obtain the relative highest profit using DEA model and thus judge the service preference of the resource combinations.

3.2.1. Integrated DEA Model Based On Engineering Efficiency Concept: Assume v and u are weighting coefficients ($v \in E^m, u \in E^s$), after transform, (D_{C2R}) can be changed to (P) model below.

$$(P) \left\{ \begin{array}{l} \max \frac{u^T y_0 + \delta_1 u_0}{v^T x_0} = V_p, \\ s.t. \quad \frac{u^T y_j + \delta_1 u_0}{v^T x_j} \leq 1, \quad j = 1, 2, \dots, n, \\ v \geq 0, u \geq 0, \delta_1 \delta_2 (-1)^{\delta_3} u_0 \geq 0. \end{array} \right. \quad (10)$$

(P) is an integrated DEA model If the best solutions u^0, v^0, u_0^0 meets the requirement of $u^0 > 0, v^0 > 0, V_p=1$, (P) is determined to be DEA valid.

3.2.2. Case Study of The Best Ratio of Combining Service Resources: Below a case study will help understand how to evaluate the best service ratio using the above integrated DEA model to gain the best efficiency. Assuming only three types of cloud services (computation, storage and network) under consideration, completing one unit task can earn 100 yuan, while the cost for different service type was: computation 1, storage 1.2, and network 0.8. Table 2 shows the amount of tasks that can be completed by different resource combinations within a unit of time.

Table 2. The Capability of Different Resource Combinations

Service type	Combination1	Combination 2	Combination 3
computation (task/sec)	0.40	0.30	0.50
storage (task/sec)	0.20	0.30	0.15
network (task/sec)	0.30	0.30	0.20

The Integrated DEA model (P) can solve the problem of how to arrange different service type in order to gain the highest income per unit time, thus determine the comprehensive service position of the resource portfolio.

Firstly, Combination 1 was considered. Using x, y, z to represent the ratio of different services and (P) model applied, we got the following linear programming.

$$(P) \left\{ \begin{array}{l} \max \frac{0.40x + 0.20y + 0.30z}{x + 1.2y + 0.8z} = V_p, \\ s.t. \frac{0.40x + 0.20y + 0.30z}{x + 1.2y + 0.8z} \leq 1, \\ \frac{0.30x + 0.30y + 0.30z}{x + 1.2y + 0.8z} \leq 1, \\ \frac{0.50x + 0.15y + 0.20z}{x + 1.2y + 0.8z} \leq 1 \\ x + y + z = 1, x \geq 0, y \geq 0, z \geq 0. \end{array} \right. \quad (11)$$

The optimal solution is $x=1, y=0, z=0$ which means providers will gain the highest revenue if the resource combination 1 is positioned to be computation service.

Similarly, the linear programming for the combination 2 is as follows.

$$(P) \left\{ \begin{array}{l} \max \frac{0.30x + 0.30y + 0.30z}{x + 1.2y + 0.8z} = V_p, \\ s.t. \frac{0.40x + 0.20y + 0.30z}{x + 1.2y + 0.8z} \leq 1, \\ \frac{0.30x + 0.30y + 0.30z}{x + 1.2y + 0.8z} \leq 1, \\ \frac{0.50x + 0.15y + 0.20z}{x + 1.2y + 0.8z} \leq 1 \\ x + y + z = 1, x \geq 0, y \geq 0, z \geq 0. \end{array} \right. \quad (12)$$

The best solution is $x=0, y=0, z=1$ from which it is judged that combination 2 should be better defined network service for the highest income. Moreover, combination 3 is better to be computation if the maximum benefits were required.

4. DEA Based Evaluation Model for the Operation of Cloud Enterprises

4.1. The Operating Data of Some Listed Cloud Companies in the Mainland

As a business platform, it is necessary for the effective assessment of the economic benefits in cloud, of which to evaluate the operations of cloud enterprises is a very important aspect that can discover the loopholes and improve the future competitiveness of the corresponding enterprises. This paper got the operation data of some mainland cloud companies listed in the 2013 cloud computing industry research report [9,10] available on the "Securities Star" site which is shown in Table 3.

Table 3. The Operation Data of Some Listed Cloud Enterprises In The Mainland (2014-03-31)

Company name	Total assets (Million)	Total operating costs (Million)	Main business income (Million)	Net profit (Million)
ZTE (000063)	10258177.4	1869994.5	1905274	62218.2
Neusoft Group (600718)	896113.23	157816.5	161419.87	2760.1
Wave software (600756)	140762.35	20492.4	19929.99	677.2
Dr. Peng (600804)	1254829.61	145111.36	165170.55	17168.84
Yonyou Software (600588)	684442.31	65091.38	47921.33	-9085.39
Focus Technology (002315)	207866.5	11411.21	12678.24	1784.8
TOYOU Technology (300302)	51891.35	3325.36	3086.3	-196.41
IFLYTek (002230)	432581.0	57222.07	26193.89	5089.63

4.2. Analysis of the Operation Status of the Above Cloud Enterprises

Using the data in table 3 and constructing formula with C^2R model, the DEA evaluation results are shown in table 4.

Table 4. DEA Evaluation Results

Company name	ZTE	Neusoft Group	Wave Software	Dr. Peng	Yonyou Software	Focus Technology	TOYOU Technology	IFLYTek
result	1	0.99678	0.908	1	0.60907	0.74094	0.66251	0.62988

From the results, we could see that currently the most DEA efficiency enterprises were ZTE and Dr. Peng and that the rest ones (in descending order) are: Neusoft Group, Wave software, Focus Technology, TOYOU Technology, IFLYTEK and Yonyou software.

5. Conclusion

This paper introduced Data Envelopment Analysis method(DEA) into cloud and proposed several novel methods for the evaluation of cloud resource allocation and cloud business operations. Case studies show that our methods were able to cope with multi-input and multi-output conditions in the efficiency evaluation situation of cloud economics related problems and were simple and effective. Our future work include: to realize the above models in our experimental platform, to test the efficiency of the models and also to perfect the models.

Acknowledgements

This work is supported by the National Natural Science Foundation of China (NO.61402144), the Natural Science Fund of Zhejiang Province (NO.LQ12G02016).

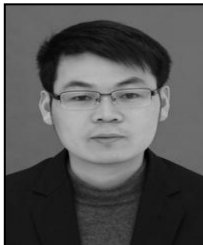
References

- [1] P. Liu, "Cloud Computing (second edition), Publishing House of Electronics Industry, Beijing", (2011).
- [2] "Chinese Cloud Forum", <http://bbs.chinacloud.cn>.
- [3] Z. X. Ma, "Data Envelopment Analysis Models and Methods (first edition), Science Press, Beijing (2010).
- [4] Q. L. Wei, "Data Envelopment Analysis models to evaluate relative efficiency, DEA and network DEA", China Renmin University Press, Beijing (2012).
- [5] Q. L. Wei and G. LU, "Applications of DEA methods and models—Data Envelopment Analysis", System Engineering- Theory & Practice, vol. 3, no. 5, (1989), pp. 67-75.
- [6] D. Du, Q. H. Pang and Y. WU, "Modern Comprehensive Evaluation methods and cases selection", (second edition), Tsinghua University Press, Beijing (2008).
- [7] Z. X. Ma, "Method of Data Envelopment Analysis Based on the Theory of Partially Ordered Sets", System Engineering- Theory & Practice, vol. 23, no. 4, (2003), pp. 11-17.
- [8] J. Y. Wu, L. D. Ping and X. Z. Pan, "Cloud computing, Concept and Platform, Telecommunications Science", vol. 12, (2009), pp. 23-30.
- [9] "Chinese Research Reports Network. 2013 Chinese Cloud Computing Industry Research and Analysis Report", <http://wenku.baidu.com/link?url=pEd67NUJTT4Y9UclB7vbhErJSPWYKj1XbNBnbXTEPkxxmmrMRUHp8vP20Bsk4CjSSCKdHo8Vf7EE9DscdS9uwz6A4y9SEZhKRrhJGieh4gO>, (2013).
- [10] "Stock Star Network", <http://www.stockstar.com/>
- [11] J. Y. Wu, W. Li, H. Juan, J. P. Zhang, J. Lin and D. R. Chen, "Key techniques for Mobile Internet a survey", SCIENTIA SINICA Information, vol. 45, no. 1, (2015), pp. 45-69.
- [12] J. Zhang and J-Y. Wu, "Provable Secure Efficient Arbitrated Quantum Signature Scheme", Journal of Beijing University of Posts and Telecommunications, vol. 36, no. 2, (2013), pp. 113-116.

Authors



Li WenJuan, she was born in 1978. She received ph. D. degree in 2012 from Zhejiang Unoversity, Hangzhou China, in computer science. She is now a lecture in Hangzhou Normal University. Her research interests include cloud computing security and trust.



Wu Jiyi, he was born in 1979. He received Ph. D. degree in 2011 from Zhejiang University, Hangzhou China, in computer science. He is now a researcher at E-Service Research Center, Zhejiang University, and the director of Key Lab of E-Business, Hangzhou Normal University. His research interest includes services computing, trust and reputation. He has published more than 30 journal articles.

