Soft Comput (2017) 21:7221–7234
DOI 10.1007/s00500-016-2267-y

METHODOLOGIES AND APPLICATION

Evaluating the efficiency of cloud services using modified data envelopment analysis and modified super-efficiency data envelopment analysis

Chandrashekar Jatoth[1,2] · G. R. Gangadharan[2] · Ugo Fiore[3]

Published online: 11 July 2016

Abstract Several cloud services with comparable functionality are now available to customers at different prices and performance levels. Often, there may be trade-offs among different functional and non-functional requirements fulfilled by different cloud providers. Hence, it is difficult to evaluate the relative performances of the cloud services and their ranking based on various quality of service attributes. In this paper, we propose a modified data envelopment analysis and a modified super-efficiency data envelopment analysis for evaluating the cloud services and their efficiencies considering user preferences. We compare these methods of cloud service selection based on sensitivity analysis, adequacy to changes in DMUs, adequacy to support decision making and modeling of uncertainty. The comparison helps customers to choose a cloud service that is most suitable to their requirements and also creates a healthy competition among the cloud service providers.

Communicated by V. Loia.

✉ G. R. Gangadharan
grgangadharan@idrbt.ac.in

Chandrashekar Jatoth
chandrashekar.jatoth@gmail.com

Ugo Fiore
ufiore@unina.it

1 School of Computer and Information Sciences, University of Hyderabad, Hyderabad, India

2 Institute for Development and Research in Banking Technology (IDRBT), Hyderabad, India

3 Department of Molecular Medicine and Medical Biotechnology, Federico II University, Naples, Italy

Keywords Cloud computing · Data envelopment analysis · Multi-criteria decision making · Analytic hierarchy process · Analytic network process

1 Introduction

Cloud computing is a network-based model whereby shared resources, software and information are provided to computers and other devices on-demand, like a public utility. A user can access technology-enabled services from the cloud through Internet without owning the technology infrastructure that supports them (Buyya et al. 2010). The ideology behind development of the cloud computing is that information processing could be done more efficiently and centrally on large firms of computing and storage systems accessible via the Internet. The cloud offers enormous benefit to businesses by providing various services at reduced cost (pay for what you use). Businesses no longer need to spend large amounts of capital on buying expensive application software or sophisticated hardware.

Several cloud services with comparable functionality are available to customers at different prices and performance levels. Often, there may be trade-offs between different functional and non-functional requirements fulfilled by different cloud providers. This makes it difficult to evaluate the cloud services (Garg et al. 2013; Shivakumar et al. 2013; Zheng et al. 2013; Chandrashekar et al. 2016). Thus, a methodology is required to evaluate the relative performance of the cloud services based on quality of service (QoS) attributes specified by a user. This kind of evaluation will provide a solution to customers to select a best cloud service that can satisfy their needs. Moreover, this enables a provider to know the areas in which improvement is required so as to meet the customer demands and be proficient in the market.

Multi-criteria decision making (MCDM) is widely used in the literature, capable of selecting, comparing and ranking of different attributes of multiple alternatives. Analytic hierarchy process (AHP) is a subjective method to analyze the qualitative criteria for generating weight of decision criteria (Saaty 1988). Data envelopment analysis (DEA) (Charnes et al. 1978; Banker et al. 1984) is a linear programming methodology based on multi-criteria decision making.

Several integrated DEA and AHP methods (Shang and Sueyoshi 1995; Yang and Kuo 2003; Ertay and Ruan 2005; Sinuany-Stern et al. 2000; Lin et al. 2011; Mohajeri and Amin 2010; Ertay et al. 2006; Azadeh et al. 2011; Ahmad et al. 2006; Wen et al. 2015; Chou 2010) are used in various real-world applications. While AHP is an effective method for decision making, it does not consider the uncertainty of human decision making in pairwise comparison evaluation. To overcome the uncertainty, fuzzy AHP (Buckley 1985; Buckley et al. 2001) allows the decision makers to use fuzzy ranking instead of exact ranking. As AHP does not include dependency and feedback for decision making in criteria weights, analytic network process (ANP) is introduced (Saaty 1996). ANP derives a super matrix to handle the interdependence among weights of criteria for all alternatives. Several integrated DEA and ANP (Kuo and Lin 2012; Lin 2010; Ramanathan 2006) methods are used in various real-world applications. DEA models evaluate the efficiency of an observation with respect to a reference set comprising of all sample observations. As super-efficiency DEA (SDEA) excludes each observation from its own reference set, the resulting efficiency score could exceed one (Banker and Chang 2006).

With the increasing popularity of cloud computing, a lot of research has been done to compare the cloud services for different types of applications based on attributes including security, accountability, assurance, performance, cost.

Silas et al. (2012) present a methodology for service selection of middle-ware services in cloud computing environment by using ELECTRE based on multi-criteria decision-making model. The parameters considered for influencing the process of service selection in Silas et al. (2012) include flexibility, time, service cost, scalability, trust and capability. A framework to find reliable cloud service providers using a recommendation from reliable sources is proposed by Bedi et al. (2012). Shivakumar et al. (2013) proposed a fuzzy multi-attribute decision-making method for ranking cloud services. Kwon and Seo (2013) proposed a fuzzy AHP method to evaluate and select the cloud service selection. Yan et al. (2012) proposed a systematic framework on top of a hybrid cloud management platform for enterprises to automatically recommend and select cloud services according to business requirements, company policies and standards, and the specifications of cloud offering. Esposito et al. (2015) developed a novel method for cloud service selection. In their model, the fuzzy sets theory is used to set the preferences of the users and service selection is resolved with Dempster-Shafer theory and game theoretic approach. Zheng et al. (2013) proposed a QoS ranking prediction framework for cloud services by taking advantage of the past service usage experiences of other consumers, and two personalized QoS ranking prediction approaches were proposed to predict the QoS rankings directly. Li et al. (2010) proposed a framework CloudCmp to compare the performance of various cloud services. CloudCmp systematically compares the performance and cost of cloud providers along dimensions that matter to customers. Garg et al. (2013) introduced a framework for ranking cloud services using analytical hierarchy process which evaluates the cloud services based on different applications depending on the QoS requirement. Xu et al. (2015) proposed a nonparametric DEA method for evaluating cloud services based on the values of price/hour, virtual core, compute units, memory and disk. In our proposed method, we present a modified DEA and a modified SDEA with AHP/ANP that includes user preferences for evaluating cloud services based on QoS attributes (CPU performance, disk I/O consistency, disk performance and memory performance).

As there are several cloud service providers offering various cloud services with different quality of services in today's market, the obvious question is to select a particular cloud service to deploy our applications that work with robustness and help to improve business (Menzel et al. 2013). Despite the huge amount of research illustrating the use of AHP, ANP, DEA and SDEA in various applications, to the best of our knowledge, this is the first study applying DEA and SDEA integrated with AHP/ANP for comparing various real-world cloud services provided by Amazon, HP, Azure, Rackspace, Google, Century Link, City-Cloud, Linode, GoGrid, Softlayer and Joyent.

The salient contributions of this study are: (i) a modified DEA method integrated with AHP/ANP to rank the cloud services based on their QoS attributes. (ii) A modified super-efficiency DEA method integrated with AHP/ANP to rank the cloud services based on their QoS attributes. (iii) A comprehensive comparison on sensitivity analysis, adequacy to changes in DMUs, adequacy to support decision making and modeling uncertainty of the proposed methods to determine efficiency of cloud services.

The remainder of this paper is organized as follows. Section 2 describes the concepts of DEA, SDEA, AHP and ANP. In Sect. 3, we present our modified DEA and modified SDEA integrated with AHP and ANP. Section 4 describes our data collection methodology and data set description. We describe performance evaluations of cloud services using our proposed DEA and SDEA methods in Sect. 5. A comparative analysis of our proposals is illustrated in Sect. 6 followed by concluding remarks in Sect. 7.

2 Preliminaries

This section presents the concepts of data envelopment analysis (DEA), super-efficiency data envelopment analysis (SDEA), analytical hierarchical process (AHP) and analytical network process (ANP).

2.1 Data envelopment analysis (DEA)

DEA is a multi-criteria decision-making model based on linear programming methodology. The "Envelopment" in DEA stems from of the way observations are enveloped in order to identify the "Pareto frontier" that is used to evaluate the relative performance of all peer entities (Banker et al. 1984). It was introduced as a "mathematical programming model applied to observational data that provides a new way of obtaining empirical estimates of relation" (Charnes et al. 1978).

In the DEA model, decision-making units (DMUs) represent the operations or processes which convert multiple inputs to multiple outputs. The model maximizes the efficiency of a DMU subject to constraints relative to the best DMUs. The efficiency of all DMUs is, then, less than or equal to 1. Through performance evaluation, the model determines the "Efficiency Frontier" that represents the relative best practices. Inefficient strategies can be improved with the suggested directions (Cooper et al. 2011). The basic assumption is that if a DMU is inefficient, then a virtual DMU obtained via a combination of other efficient DMUs can either result in greater output for the same level of inputs, use fewer inputs to produce the same level of outputs, or some combination of both. There are two approaches to determine an efficiency frontier: input oriented and output oriented (Cooper et al. 2011).

Input-oriented DEA The DEA model is input oriented when inputs are minimized while keeping outputs fixed at their current value. The linear programming formulation of the input-oriented DEA model is given as follows. Suppose there are N DMUs, of which DMUo is the one under evaluation. Let x_{ij} represent the amount of the jth (out of P) input used by DMUi and let y_{ik} denote the amount of the kth (out of Q) output produced by DMUi. The efficiency score of DMUo is given by

$$\theta^* = \min_{\theta,\mathbf{w}} \theta \tag{1}$$

subject to

$$\sum_{j=1}^{N} w_j x_{ij} \leq \theta x_{io} \quad i = 1, \ldots, P \tag{2}$$

$$\sum_{j=1}^{N} w_j y_{kj} \geq y_{ko} \quad k = 1, \ldots, Q \tag{3}$$

$$w_j \geq 0 \quad j = 1, \ldots, N \tag{4}$$

where $\mathbf{w}$ is a weight vector (w_j represents the contribution of the jth DMUs to the virtual DMU). If $\theta^* = 1$, then the current input level cannot be reduced indicating that DMUo is on the "efficient frontier." Otherwise, if $\theta^* < 1$, then DMUo is inefficient and there is room for increasing efficiency by reducing input with output staying at the same level.

Output-oriented DEA In the output-oriented DEA model, outputs are maximized while keeping input levels constant. The corresponding linear program is the same as in the input-oriented DEA, Eqs. (1)–(4), with the modifications that in Eq. (1), maximization is substituted with minimization, and that the inefficiency score θ', the parameter corresponding to θ, appears in the constraint on outputs as a multiplier for y_{ko} and it is no longer present in the constraint for inputs, yielding the following modified constraints in place of Eqs. (2) and (3):

$$\sum_{j=1}^{N} w'_j x_{ij} \leq x_{io} \quad i = 1, \ldots, P \tag{5}$$

$$\sum_{j=1}^{N} w'_j y_{kj} \geq \theta' y_{ko} \quad k = 1, \ldots, Q \tag{6}$$

The two formulations are equivalent with the position

$$\theta = 1/\theta' \quad w = w'/\theta' \tag{7}$$

There is also a possibility of improvement in a DMU that is located at the same level as of another DMU of the efficient frontier. It can be done by varying the proportions in which inputs are utilized. Defining the (nonnegative) input excesses and output shortfalls as

$$s_i^- = \theta^* x_{io} - \sum_{j=1}^{N} w_j x_{ij} \qquad i = 1, \ldots, P \quad \text{(input excess)} \tag{8}$$

$$s_k^+ = \sum_{j=1}^{N} w_j y_{kj} - y_{ko} \qquad k = 1, \ldots, Q \quad \text{(output shortfall)} \tag{9}$$

and using them as slack variables in the following linear program:

$$\max \omega = \sum_{i=1}^{P} s_i^- + \sum_{r=1}^{Q} s_k^+ \tag{10}$$

subject to

$$s_i^- = \theta^* x_{io} - \sum_{j=1}^{N} w_j x_{ij} \qquad i = 1, \ldots, P \tag{11}$$

$$s_k^+ = \sum_{j=1}^{N} w_j y_{kj} - y_{ko} \qquad k = 1, \ldots, Q \tag{12}$$

$$w_j \geq 0, \quad s_j^- \geq 0, \quad s_j^+ \geq 0 \tag{13}$$

the optimal slacks (max-slack solution) can be computed. Now the notions of weak and strong efficiency can be given. A DMU for which $\theta^* = 1$ is said to be weakly efficient. A DMU is strongly efficient (Pareto–Koopmans efficient) if it is weakly efficient and all the optimal slacks are zero. With a strongly efficient DMU, any improvement on an input or output will produce a worsening of some other input or output. This is important from a service provider point of view, because, for a given set of user-specific performance attributes, it is not only necessary to determine which one is performing better, but also to know the specific areas where efficiency is lacking and improvement is needed.

2.2 Super-efficiency data envelopment analysis (SDEA)

Since the early 1980s, DEA has been used as an alternative method of classification for evaluating the relative efficiency of independent homogeneous units which use the same inputs to produce the same outputs (Tone 2002). However, DEA inconveniently leaves room for ties between units with relative efficiency equal to 100 %. SDEA is used for ranking the performance of efficient DMUs (Andersen and Petersen 1993). DEA evaluates the efficiency of a unit relative to a reference set comprising of all units (including itself), whereas SDEA excludes each unit from its own reference set. It is, therefore, possible to obtain efficiency scores that exceed 1.

Input-oriented SDEA The linear program for the input-oriented model used for super-efficiency DEA is the same as the one in Eqs. (1)–(4), with the difference that DMUo is excluded from the summations in Eqs. (2) and in (3), which become:

$$\sum_{j=1, j\neq o}^{N} w_j x_{ij} \leq \theta x_{io} \quad i = 1, \ldots, P \tag{14}$$

$$\sum_{j=1, j\neq o}^{N} w_j y_{kj} \geq y_{ko} \quad k = 1, \ldots, Q \tag{15}$$

This will allow a super-efficiency score greater than one, enabling distinguishability between efficient observations. Super-efficiency measures can be calculated for both inefficient and efficient observations, but in the case of inefficient observations, the values of the efficiency measure do not change, while efficient observations may obtain higher values.

Output-oriented SDEA The output-oriented version of SDEA is the same as the corresponding DEA version, with the same modifications as above, i.e., DMUo is being excluded from the summations in the constraints, i.e., Eqs. (5) and (6).

2.3 Analytical hierarchical process (AHP)

AHP is a multi-criteria decision-making method to allow a decision makers to compute a ratio scale from preferences and model a complex problem in a hierarchical structure. This structure consists of goal, criteria (factors), sub-criteria (sub-factors) and alternatives (Saaty 1980). In AHP, criteria are evaluated at top level and alternatives are evaluated at bottom level for each criterion. Each level and sub-level are evaluated separately by the decision makers. The decision makers should determine the weights of all criteria in order to do pairwise comparison among them. It helps the decision makers to incorporate a group agreement using questionnaire for comparing each element and geometric mean to form a final solution (Vaidya and Kumar 2006). The procedure of AHP is as follows (Saaty 1988):

(i) Model the problem structure by breaking down the decision problem (goal, criteria, sub-criteria and alternatives)
(ii) Establish the priority of input data, by pairwise comparison matrix on each element of the hierarchical structure.
(iii) Develop the pairwise comparison matrix at each level of hierarchy (the pairwise comparison determined according to the scale from 1 to 9).
(iv) Calculate vector of weights by using eigenvector procedure.
(v) Calculate the consistency ratio (CR) to check the consistency of the judgment. If CR < 0.1, then the pairwise comparison is consistent and acceptable.
(vi) Aggregate of the relative weights of decision elements and form a set of preferences for alternatives.

The consistency ratio (CR) of the pairwise comparison matrix A is defined as follows:

$$\mathrm{CR} = \frac{\mathrm{CI}}{\mathrm{RI}} \tag{16}$$

$$\mathrm{CI} = \frac{\lambda_{\max} - n}{n - 1} \tag{17}$$

where CI is the consistency index, n is the order of the pairwise comparison matrix A, and $\lambda_{\max}$ is its maximum eigenvalue, while the random index RI is the average CI value for random matrices.

2.4 Analytical network process (ANP)

Analytic network process (ANP) (Saaty 2006, 1996) aims to surmount the limitations of AHP. AHP does not allow the elements of hierarchical model to have dependence and feedback between each criterion and alternatives. ANP allows interdependence interaction between involved elements that can be criteria and alternatives. ANP is not limited by the independent assumptions between the criteria and the alternatives of the decisions, or simply among the criteria or among the alternatives themselves. However, the importance of criteria and alternatives are determined. ANP shows the interdependencies with feedback in the structure which has two-way arrows and connected cycles of its clusters like a network instead of level hierarchy in AHP (Kuo et al. 2015). This approach has the potential for handling interdependent relationships among the criteria weights by procure composite weights through the development of a "supermatrix". The supermatrix concept is similar to the Markov chain process (Saaty 1996) where respective prime weights are forming a supermatrix from their eigenvectors. The supermatrix consists of control hierarchy and network hierarchy. The control hierarchy is an interaction between the criteria and sub-criteria, and it also includes the goal of the problem, criteria and sub-criteria, each criterion and sub-criteria are independent at the same level. The weight of criteria is evaluated with AHP in the control hierarchy. The network hierarchy indicates the relationship between elements and clusters, and criteria interaction (Kuo et al. 2015).

3 Modeling modified DEA and modified SDEA for cloud services

MCDM techniques are predominantly empowered to structure the problem systematically and distinctly. Based on this aspect, the decision makers have the prospect to certainly explore and scale the problem in obedience with their requirement. The main objective of this paper is to select the best cloud service with respect to the user preferences. Hence, we use AHP and ANP to determine the priorities or weights of different criteria of QoS attributes of cloud services. Then, we integrate AHP/ANP with DEA and SDEA for selecting the best cloud service.

3.1 Modified data envelopment analysis

We propose a DEA model by considering a cloud service as a decision-making unit (DMU) and price per hour as input and performance values (CPU performance, disk I/O consistency, disk performance, memory performance) as output parameters. The DEA model locates cloud services that lie on the efficient frontier. This indicates that these cloud services are performing relatively high. However, there is still a chance of improvement for these cloud services if the slack variable is nonzero. The DEA model gives the projected value, for improving the performance of a service which does not lie on the efficient frontier. To give preference to the QoS attributes for measuring the performance of the cloud services, we use the input-oriented model.

We modify the input-oriented model to evaluate the preferred efficiency by assigning weights to the output parameters obtained by using AHP. The DEA discrimination ability reduces if more numbers of DMUs are identified as efficient units because the sum of the number of inputs and outputs is large as compared to the total number of DMUs in the sample (Andersen and Petersen 1993; Azadeh et al. 2008; Seiford and Zhu 1999). Although DEA differentiates between efficient and inefficient DMUs, assigning a unique rank to them is not possible in DEA. While ranking the DMUs using DEA model, various DMUs achieve an efficiency score of 1. In such cases, ranking of efficient DMUs is a pivot challenge faced by the decision maker. In our modified DEA, the weights are assigned to the input variables by using AHP/ANP method. The modified objective function is given as follows:

$$\theta^* = \min_{\theta, \mathbf{w}} \theta \tag{18}$$

subject to

$$\sum_{j=1}^{N} w_j x_{ij} \leq \theta x_{io} \quad i = 1, \ldots, P \tag{19}$$

$$\sum_{j=1}^{N} w_j y'_{kj} \geq y'_{ko} \quad k = 1, \ldots, Q \tag{20}$$

$$w_j \geq 0 \qquad j = 1, \ldots, N \tag{21}$$

where y'_{ko} is the modified kth output of DMUo. Other notations have their usual meanings. Similarly, we modify the input-oriented model by assigning weights using ANP to the output parameter to calculate the preferred efficiency of DEA model.

A pseudo-code to determine the most preferable efficiency of cloud services among the available services, with the modified DEA model integrated with AHP/ANP, is shown in Algorithm 1.

3.2 Modified super-efficiency data envelopment analysis

We modify the super-efficiency DEA model for cloud services to find the preferred efficiency of cloud services using

Algorithm 1: Pseudo-code representation of DEA integrated with AHP/ANP

Input: m cloud services with n QoS attributes.
Price attribute is an input for a cloud service in a DEA model.
Output: cloud service DMUs relative efficiency scores.
1. Initialize each cloud service as a DMU.
2. Determine the relative performance of cloud services using the modified input-oriented DEA model.
3. Evaluate the preferred efficiency by assigning QoS weights using AHP/ANP method.
4. Calculate the technical efficiency score of DMU by using equations 18–21.
5. IF $\theta^* = 1$
then
DMU is on the efficient frontier
else
DMU is inefficient.
6. Improve the efficient frontier of DMU by calculating slack by using equations 8–10.
7. Rank the cloud services based on relative efficiency score.

Algorithm 2: Pseudo-code representation of SDEA integrated with AHP/ANP

Input: m cloud services with n QoS attributes.
Price attribute is an input for a cloud service in a SDEA model.
Output: cloud service DMUs relative efficiency scores.
1. Initialize each cloud service as a DMU.
2. Calculate the relative performance of cloud services using the modified input-oriented SDEA model.
3. Determine the preferred efficiency by assigning QoS weights using AHP/ANP method.
4. Calculate the technical efficiency score of DMU using equations 22–24.
5. IF $\theta^* = 1$
then
DMU is on the efficient frontier
else
DMU is inefficient.
6. Improve the efficient frontier of DMU by calculating slack by using equations 8–10.
7. Rank the cloud services based on relative efficiency score.

AHP (as done in DEA). The modified objective function is described as follows:

$$\theta^* = \min_{\theta, \mathbf{w}} \theta \tag{22}$$

subject to

$$\sum_{j=1, j\neq E}^{N} w_j x_{i\,j} \leq \theta x_{i\,o} \qquad i = 1, 2, \ldots, P; \tag{23}$$

$$\sum_{j=1, j\neq E}^{N} w_j y_{k\,j}^{T} \geq y_{k\,o}^{T} \qquad k = 1, 2, \ldots, Q;$$

$$w_j \geq 0, \qquad j = 1, 2, \ldots, N; \tag{24}$$

where the notations are having their usual meaning as defined in the modified DEA model. Similarly, we modify the input-oriented model of SDEA by assigning weights using ANP. A pseudo-code to determine the most preferable efficiency of cloud services among the available services, with the modified SDEA model integrated with AHP/ANP is shown in Algorithm 2.

4 Data collection methodology and data set description

We have considered 11 real-world cloud service providers including Amazon, HP, Azure, Rackspace, Google, Century Link, City-Cloud, Linode, GoGrid, Softlayer and Joyent (not in any order). Each service provider provides cloud services based on the number of virtual cores. The services with 2 virtual cores are specified as large cloud services, services with 4 virtual cores are specified as extra-large cloud services, services with 8 virtual cores are specified as 2x-extra-large cloud services, and services with 16 virtual cores are specified as 3x-extra-large cloud service. The dataset for the analysis is illustrated in Table 1. The cloud service providers are coded as C_1, C_2, ... and C_{11} (not in any order), and the services provided by each are further coded as S_1, S_2, S_3, and so on. For each service, the specified data for price per hour (doller), virtual core and memory (GB) are collected from the respective cloud service providers and the values for CPU performance, disk I/O consistency, disk performance and memory performance are obtained from cloudharmony.com. The list of QoS attributes and their description is presented in Table 2.

During data collection, consistency check is performed to identify data that are out of range, are logically inconsistent or have extreme values. Inconsistent data for any service provider are inadmissible, and we either corrected it if possible otherwise we did not consider the service provider for the analysis.

One may argue that the different quantitative QoS attributes of cloud service providers considered in this study are rather limited. However, collecting the real-world data set regarding quantitative QoS attributes of cloud service providers was extremely challenging. Our collected dataset could be used by several researchers further use for their research purposes.

5 Performance evaluation

This section describes the selection of cloud services using our modified DEA and modified SDEA for the given dataset.

Table 1 Collected dataset

Providers	Services	Virtual cores	Memory (in GB)	Price/h	CPU performance	Disk I/O consistency	Disk performance	Memory performance
C_1	$C1S1$	1	1.6	4.4	0.92	74	27.08	64
	$C1S2$	4	7.5	0.17	63.44	66	56.82	91
	$C1S3$	4	15	0.36	5.15	66	109	91
	$C1S4$	4	15	0.29	126.66	92.89	110.33	104
	$C1S5$	8	30	0.59	231.77	53.28	67.22	104
C_2	$C2S1$	2	7.5	0.16	65.72	114.44	97.38	144.86
	$C2S2$	4	15	0.42	111.95	119.63	100.5	131.81
C_3	$C3S1$	8	30	0.63	244.36	77.46	73.44	125.59
	$C3S2$	2	4	0.12	77.49	23.43	40.23	80.67
	$C3S3$	4	15	0.45	152.96	29.07	42.47	90.83
	$C3S4$	8	30	0.9	249.72	35.35	55.07	83.92
C_4	$C4S1$	1	2	0.06	4.66	62.08	56.08	78.51
	$C4S2$	2	4	0.12	5.45	78.56	109.2	84.2
	$C4S3$	4	8	0.24	5.53	72.92	110.78	76.04
	$C4S4$	8	16	0.42	6.19	74.53	130.84	90.37
	$C4S5$	16	32	0.9	6.82	71.45	144.79	100.87
C_5	$C5S1$	2	3.5	0.12	44.53	67.87	83.73	53.27
	$C5S2$	4	7	0.24	82.2	67.97	78.49	61.8
	$C5S3$	8	14	0.48	140.89	96	133	70.91
C_6	$C6S1$	2	2	0.12	54.96	68.41	62.46	63.04
	$C6S2$	4	4	0.24	41.85	70.29	63.1	63.44
	$C6S3$	8	8	0.48	82.84	72.56	78.12	63.48
	$C6S4$	8	14.5	0.9	81.21	100.15	78.66	70.09
C_7	$C7S1$	4	2	0.36	5.11	41.07	72.45	95.74
	$C7S2$	8	16	1.05	58.42	31.22	68.45	78.15
C_8	$C8S1$	4	16	0.56	17.34	43.02	141.23	51.71
	$C8S2$	8	64	1.65	37.05	36.15	102.74	132.87
	$C8S3$	16	128	2.52	71.11	39.66	99.15	135.88
C_9	$C9S1$	2	4	0.11	23.43	89.31	173.49	89.84
	$C9S2$	4	8	0.21	42.05	59.63	174.5	97.16
	$C9S3$	8	16	0.34	75.89	64.64	174.12	100.14
C_{10}	$C10S1$	1	1	0.06	4.42	108.19	89.12	79.72
	$C10S2$	2	2	0.12	4.87	133.61	82.01	83.74
	$C10S3$	4	4	0.24	9.28	161.14	86.15	96.12
	$C10S4$	8	8	0.48	23.2	130.68	122.62	89.74
C_{11}	$C11S1$	1	1	0.278	5.48	34.47	89.12	84.12
	$C11S2$	2	2	0.056	5.38	47.17	84.56	81.23
	$C11S3$	8	8	0.222	5.33	95.38	78.15	82.16
	$C11S4$	16	16	0.444	6.3	106.06	82.14	78.38

5.1 Cloud service selection using modified DEA and modified SDEA with AHP

Our experimental evaluation is based on the input-oriented DEA that addresses the problem "by how much can input parameter (price per hour) be proportionally be decreased without changing the output parameter (performance values)." The input parameter considered for the analysis is price per hour charged by the cloud service provider for a cloud service. We employed four output parameters; however, any number of parameters could be included for both outputs and inputs. The output parameters include CPU performance, disk I/O consistency, disk performance and memory performance.

Table 2 List of QoS attributes and their description

S. no.	QoS attributes name	Description
1	Price per hour	The cost of a virtual machine per hour
2	CPU performance	The number of jobs that a computer can execute in a given amount of time (the processing and orchestration of all applications as integer and floating point operations per second)
3	Disk I/O consistency	The average time required for disk I/O operations to remain consistent (measured in I/O operations per second)
4	Disk performance	The number of operations performed on a disk in a certain amount of time (measured in I/O operations per second)
5	Memory performance	The relationship between speed and latency

Table 3 Scales for comparison matrix of criteria

Intensity of Importance	Definition
1	Equal importance
3	Moderate importance of one over another
5	Essential or strong importance
7	Very strong importance
9	Extreme importance
2, 4, 6, 8	Intermediate value between the two adjacent judgments

By using modified DEA and modified SDEA, we evaluate the preferred efficiency by assigning weights to the QoS attributes obtained by using AHP. The scale that we use ranges from 1 to 9 and is presented in Table 3. The pairwise comparison was made with the help of domain experts.

Table 4 shows the relative importance among QoS attributes. For example, the relative importance of CPU performance is 5 times as that of disk I/O consistency and the relative importance of disk performance is 0.25 (1/4) times as that of memory performance, and its reciprocal is in lower triangular matrix, i.e., memory performance is 4 times as important parameter as disk performance.

After normalizing the resultant matrix and averaging the value, we get the weights as in Table 5. To check the consistency of the calculated weights, we obtain consistency ratio (CR) as 0.040291. The consistency ratio tells how inconsistent the matrix is, and the result is acceptable if CR ≤ 0.1. So our matrix is consistent and weights are valid. These weights are used to evaluate the preferred values of the QoS attributes which is done by multiplying these fractions to the sum of output parameters.

Table 5 Weights for QoS attributes

QoS attributes	Weights (%)
CPU performance	48.66
Disk I/O consistency	8.48
Disk performance	12.13
Memory performance	30.73

The results of modified DEA with AHP and modified SDEA with AHP in comparison with DEA and SDEA are shown in Figs. 1 and 2, respectively. The results in blue color (DEA/SDEA) indicate the efficiency score of DMUs with equal preference for all QoS attributes, and the results in red color (modified DEA/modified SDEA) indicate the efficiency score of DMUs with priority weights obtained by using AHP.

From Figs. 1 and 2, we observe that C11S2 is performing relatively better among other cloud services with virtual core 2. The service C11S2 provides better performance on the preferred QoS attribute at a reasonable service charge; however, it has virtual core 2. So if a user wishes for a better service with higher virtual core (for example, a service with virtual core 4), then C1S2 is efficient on relative scale. Similarly, among services with virtual core 8, C3S3 is better than other services. It can be seen that the service provider C11 is relatively performing best among the cloud service providers.

Table 4 Relative importance among QoS attributes

	CPU performance	Disk I/O consistency	Disk performance	Memory performance
CPU performance	1	5	4	2
Disk I/O consistency	0.2	1	0.5	0.33
Disk performance	0.25	2	1	0.25
Memory performance	0.5	3	4	1

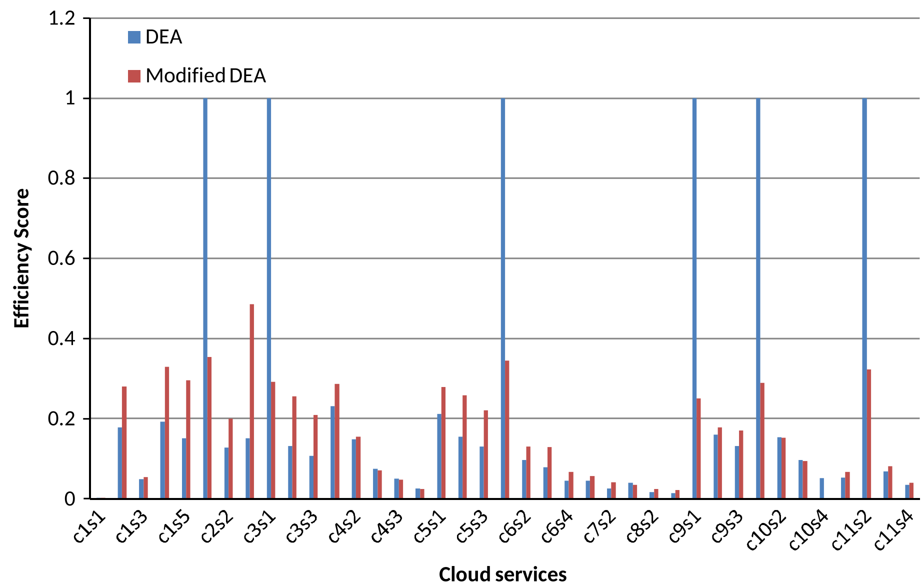

Fig. 1 Relative efficiency score of cloud service using DEA and modified DEA with AHP

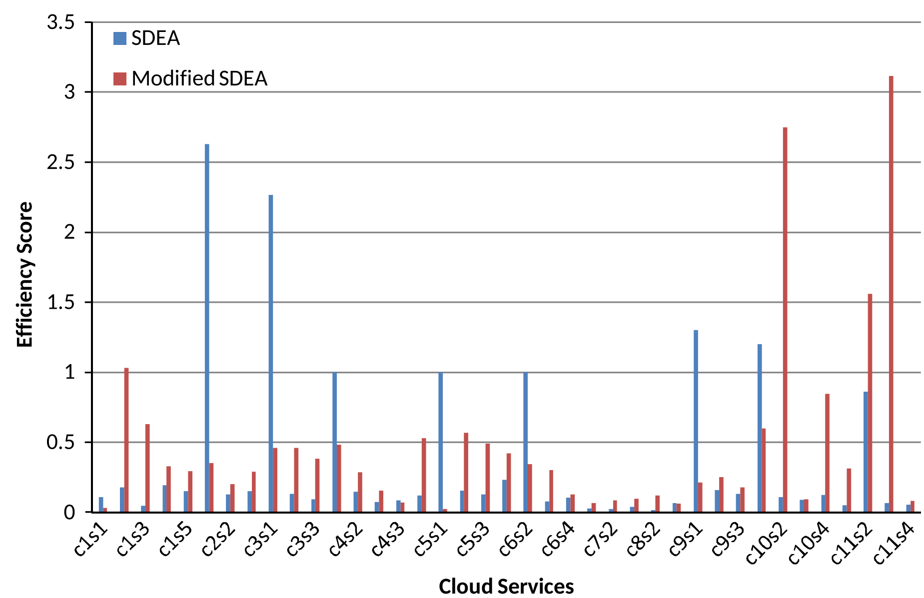

Fig. 2 Relative efficiency score of cloud service using SDEA and modified SDEA with AHP

Table 6 Relative importance among QoS attributes

	CPU performance	Disk I/O consistency	Disk performance	Memory performance
CPU performance	1	0.25	0.33	0.50
Disk I/O consistency	4	1	0.20	0.14
Disk performance	3	5	1	0.25
Memory performance	2	7	4	1

5.2 Cloud service selection using modified DEA and modified SDEA with ANP

By using modified DEA and modified SDEA, we evaluate the preferred efficiency by assigning weights to the QoS attributes obtained using ANP. The relative preference of cloud service selection criteria was evaluated by ANP using ANP super decision 1.6.0 software. The scale that we use ranges from 1 to 9 is presented in Table 3. The pairwise comparison was made with the help of domain experts. The weight of each criterion is described in Table 6. Table 7 shows the relative importance among

Table 7 Weights for QoS attributes

QoS attributes	Weights (%)
CPU performance	27.69
Disk I/O consistency	10.64
Disk performance	19.80
Memory performance	41.88

QoS attributes. The consistency ratio values are less than 1.0.

The results of modified DEA and modified SDEA integrated with ANP in comparison with DEA and SDEA are shown in Figs. 3 and 4. The results in blue colors (DEA/SDEA) indicates that the efficiency score of DMUs equal priorities for all QoS attributes, and the results in red colors (modified DEA/modified SDEA) indicate the efficiency score of DMUs with priority weights obtained by using ANP.

From Figs. 3 and 4, we observe that C2S1, C2S2, C3S2, C4S1, C5S2, C11S2, C10S2, C9S1 and C1S5 are performing relatively better than other cloud services. The service C3S2 provides better performance on the preferred QoS attributes at a reasonable service charge; however, it has virtual core 4. So if a user wishes for a better service with higher virtual core (for example, a service with virtual core 8), then C1S5 is efficient on the relative scale. Similarly, among services with virtual core 4 and 2, C3S2 and C2S1 are better than other services. It can be seen that the service provider C3 is relatively performing superior than other cloud service providers.

6 A comparative analysis of modified DEA and modified SDEA

This section presents a comparison of proposed two methods, based on the following factors: sensitivity analysis, adequacy changes in services, adequacy to support decision making and modeling of uncertainty.

6.1 Sensitivity analysis

Sensitivity analysis determines the robustness of the proposed models. To determine the robustness of a model, we gradually change the priorities and find the final weight of each QoS attribute. We moderately change the weights of one QoS attribute and keeping all other weights the same. Hence, we find the impact of final decisions. Thus, the performance of each QoS attribute and its results are analyzed and applied to modified DEA and modified SDEA. The sensitivity analysis for CPU performance and disk I/O consistency attributes against 39 cloud services of modified DEA with AHP are shown in Figs. 5 and 6. In Figs. 5 and 6, the x-axis represents the cloud services and y-axis represents the efficiency scores of DMUs. Hence, each plot corresponds to a specific weight for changing the weights of a selected QoS attribute (CPU Performance in Fig. 5 and disk I/O consistency in Fig. 6). In order to determine the trend, the priority of all DMUs is calculated for each of the different weights for the selected QoS attribute. From Fig. 5, we observe that C3S1, C2S3, C6S1, C1S4 and C2S1 give the best efficiency scores for different weights of CPU performance. We also observe that there is no significant changes in DMUs efficiency and ranking if different weights are given. Similarly, in Fig. 6, we perceive that C2S3, C6S1, C2S1, C1S4 and C5S1 give the best efficiency scores for different weights of disk I/O consistency. We also perceive that there is no significant changes in DMUs efficiency and ranking if different weights are given. Similarly, we calculate the sensitivity analysis of disk performance and memory performance. As the weight of attribute increases, the priority of the DMU also increases. However, the decision does not change. Thus, we conclude that DMU C2 with high virtual core is the best cloud service provider for all values.

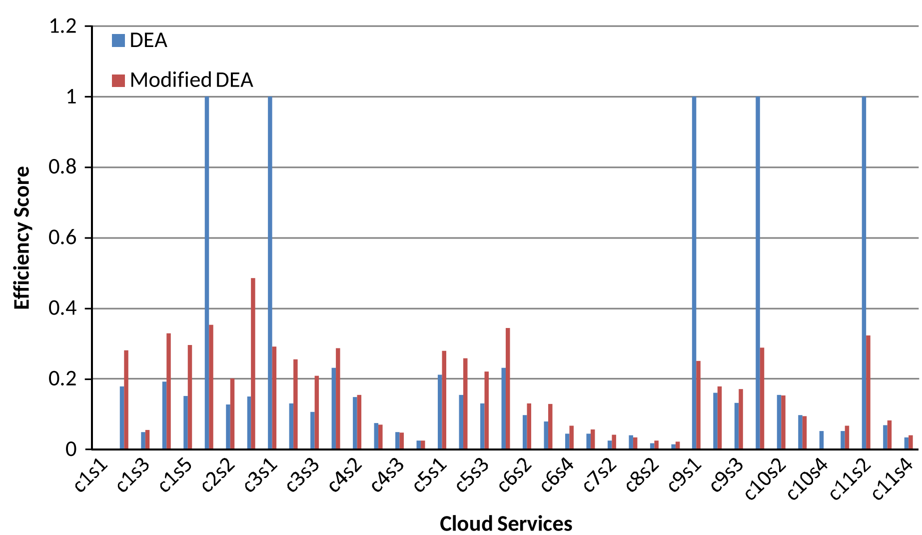

Fig. 3 Relative efficiency score of cloud service using DEA and modified DEA with ANP

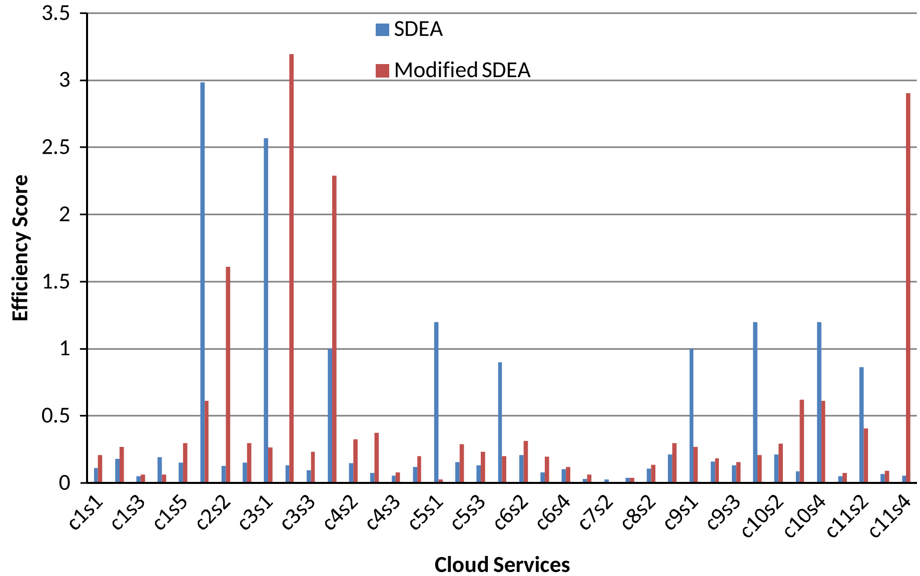

Fig. 4 Relative efficiency score of cloud service using SDEA and modified SDEA with ANP

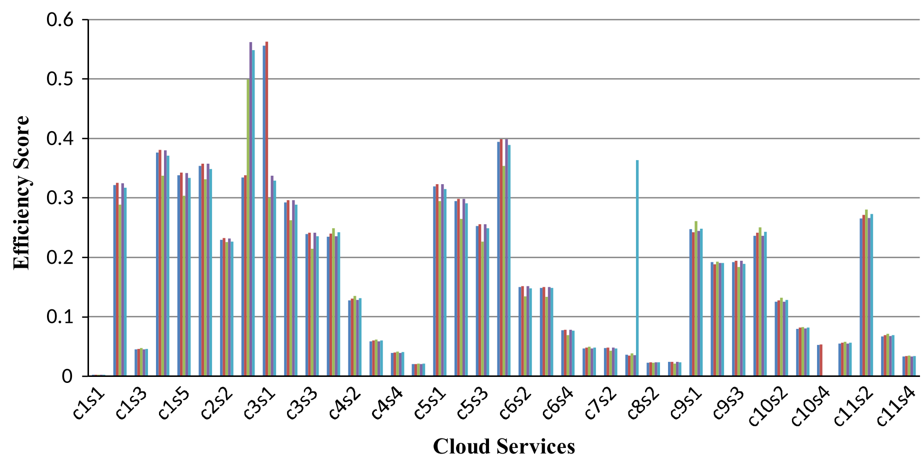

Fig. 5 Sensitivity analysis of modified DEA with AHP for CPU performance

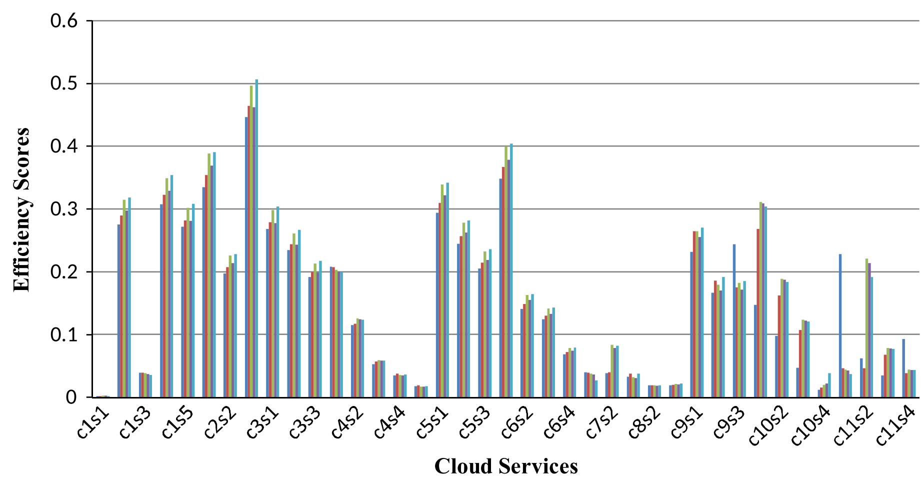

Fig. 6 Sensitivity analysis of modified DEA with AHP for disk I/O consistency

The sensitivity analysis for CPU performance and disk I/O consistency attributes against 39 DMUs of the modified SDEA with ANP is shown in Figs. 7 and 8. In Figs. 7 and 8, the x-axis represents the cloud services and the y-axis represents the efficiency scores of DMUs. Hence, each plot corresponds to a specific weight for changing the weights of a selected QoS attribute (CPU performance in Fig. 7 and disk I/O consistency in Fig. 8). In order to determine the trend, the priority of all DMUs is calculated for each of the different weights for the selected QoS attributes. From Fig. 7, we

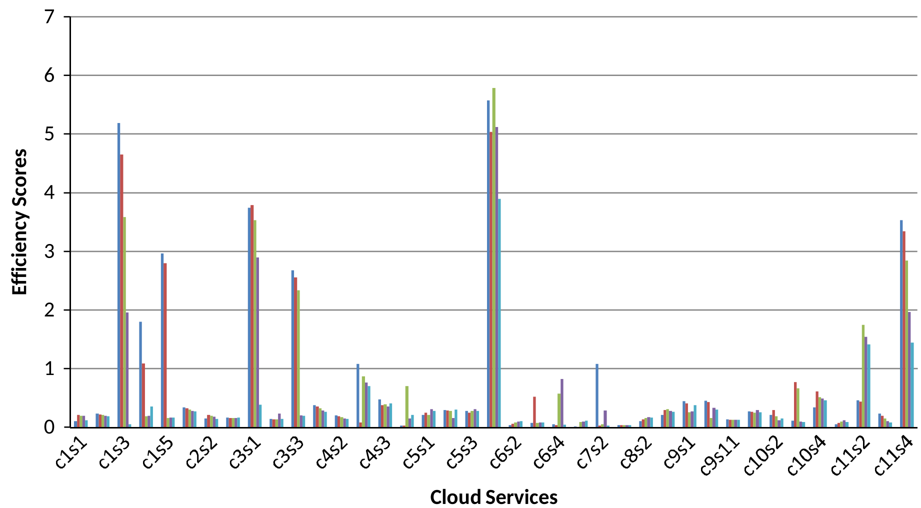

Fig. 7 Sensitivity analysis of modified SDEA with ANP for CPU performance

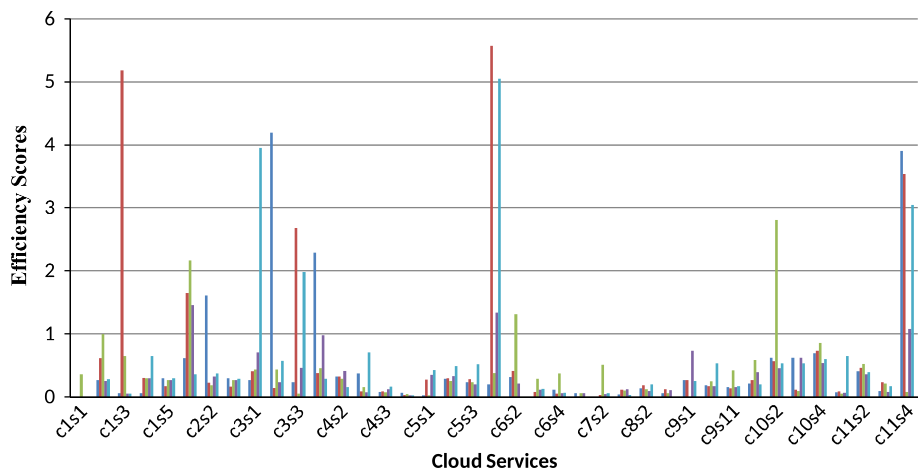

Fig. 8 Sensitivity analysis of modified SDEA with ANP for disk I/O consistency

observe that C6S1, C1S3, C3S1, C11S4 and C1S5 give the best efficiency scores for different weights of CPU performance. We also observe that there is no significant changes in the efficiency of DMUs and ranking if different weights are given. Similarly, in Fig. 6, we perceive that C6S1, C3S1, C1S3, C11S4 and C3S3 services give the best efficiency scores for different weights of disk I/O consistency. We also perceive that there is no significant changes in DMUs efficiency and ranking if different weights are given. Similarly, we calculate the sensitivity analysis of disk performance and memory performance. As the weight of attribute increases, the priority of the DMU also increases. However, the decision does not change. Thus, we conclude that DMU C6 with high virtual core is the best cloud service provider for all values.

6.2 Adequacy to changes in services

In order to rank the 39 DMUs, modified DEA and modified SDEA integrated with AHP/ANP with equal weights for all QoS attributes are used. The resultant ranking are as follows:

- $C2S3$, $C1S5$, $C1S4$, $C1S2$, $C6S1$, $C11S2$, $C3S1$, $C5S1$, $C10S1$ *and* $C4S1$ (using modified DEA with AHP).
- $C2S3$, $C1S5$, $C1S4$, $C3S1$, $C1S2$, $C5S2$, $C2S1$, $C6S1$, $C11S2$, $C5S1$, $C10S1$ *and* $C4S1$ (using modified DEA with ANP).
- $C4S4$, $C11S3$, $C10S4$, $C5S3$, $C3S3$, $C1S4$, $C1S2$, $C1S3$, $C5S2$, $C3S2$, $C6S2$, $C10S2$ *and* $C11S2$ (using modified SDEA with AHP).
- $C11S4$, $C10S4$, $C1S5$, $C3S2$, $C2S2$, $C10S3$, $C4S3$, $C6S2$, $C2S1$, $C11S2$, $C3S1$ *and* $C2S1$ (using modified SDEA with ANP).

These ranks are given according to the descending order of virtual core 16 to 2. To test these ranks of modified DEA and modified SDEA, we added an additional cloud service with equal priority weights to the existing cloud service. In several cases, we observed that the results have not shown any significant changes in the efficiency of cloud services.

Similarly, the same sequence of tests are applied to all models and no significant change is observed. The order of priority endures the same in all test cases and having the same ranking as the equal priority.

6.3 Adequacy to support decision making

Modified DEA and modified SDEA models allow aggregation of judgment of more than one decision makers. In the case of modified DEA and modified SDEA, the aggregation of different judgment is made by AHP and ANP methods. AHP does not consider uncertainty of human decision making in pairwise comparison and does not consider dependence and feedback for decision making in criteria weights. Similarly, traditional DEA does not deal with uncertainty and other limitations, whereas SDEA deals. Due to the limitation of AHP and DEA, SDEA with ANP emerges as the most preferred method for evaluating the relative efficiency of cloud services.

6.4 Modeling of uncertainty

In our proposed models, we utilize ANP to deal with an intrinsic lack of clarity of data regarding the efficiencies of cloud services. In both models, AHP is the main resource for quantifying vagueness. Due to the vagueness of judgments of quantitative variables and uncertainty of human decision making in pairwise comparison, ANP model is selected. In modified DEA and modified SDEA, we used pairwise comparison by means of comparative linguistic variables. Tables 4 and 6 present the judgment of cloud services and weights of QoS attributes.

7 Conclusions

There are many cloud service providers delivering several services with different prices and performance levels. It has now become a challenge for the customers to select a cloud service which will best satisfy their needs. To select the best service, customers need to have a methodology to identify and measure key performance criteria according to their requirements. This paper helps in selecting the most suitable cloud service among various other cloud services delivered by various cloud service providers based on user-specific QoS requirements. Although many frameworks exist to evaluate the performance of cloud services, this paper proposed a modified DEA model and a modified SDEA model for evaluating cloud services, considering the preferences of users.

Further, we performed sensitivity analysis, adequacy to changes in services, adequacy to support decision making and modeling uncertainty of the proposed methods. From these analyses, our modified SDEA seems as the best method and the DMU C3 seems as the best cloud service provider among the available services providers. In our future work, we plan to work on evolutionary algorithms for evaluating cloud services.

Acknowledgements We thank Saurabh Kumar (IIT, Kanpur, India) and Akshay Jaiswal (IIT-BHU, Varanasi, India) for their help in implementing parts of DEA and SDEA (during their internships at IDRBT) in this work.

Compliance with ethical standards

Conflict of interest Chandrashekar Jatoth declares that he has no conflict of interest. G. R. Gangadharan declares that he has no conflict of interest. Ugo Fiore declares that he has no conflict of interest.

Ethical approval This article does not contain any studies with human participants or animals performed by any of the authors.

References

Ahmad N, Berg D, Simons GR (2006) The integration of analytical hierarchy process and data envelopment analysis in a multi-criteria decision-making problem. Int J Inf Technol Decis Mak 5(2):263–276

Andersen P, Petersen NC (1993) A procedure for ranking efficient units in data envelopment analysis. Manag Sci 39(10):1261–1264

Azadeh A, Ghaderi S, Izadbakhsh H (2008) Integration of dea and ahp with computer simulation for railway system improvement and optimization. Appl Math Comput 195(2):775–785

Azadeh A, Ghaderi S, Mirjalili M, Moghaddam M (2011) Integration of analytic hierarchy process and data envelopment analysis for assessment and optimization of personnel productivity in a large industrial bank. Expert Syst Appl 38(5):5212–5225

Banker RD, Chang H (2006) The super-efficiency procedure for outlier identification, not for ranking efficient units. Eur J Oper Res 175(2):1311–1320

Banker RD, Charnes A, Cooper WW (1984) Some models for estimating technical and scale inefficiencies in data envelopment analysis. Manag Sci 30(9):1078–1092

Bedi P, Kaur H, Gupta B (2012) Trustworthy service provider selection in cloud computing environment. In: Proceedings of the 2012 international conference on communication systems and network technologies, IEEE, pp 714–719

Buckley JJ (1985) Fuzzy hierarchical analysis. Fuzzy Sets Syst 17(3):233–247

Buckley JJ, Feuring T, Hayashi Y (2001) Fuzzy hierarchical analysis revisited. Eur J Oper Res 129(1):48–64

Buyya R, Broberg J, Goscinski AM (2010) Cloud computing: principles and paradigms. Wiley, New Jersey

Chandrashekar J, Gangadharan GR, Buyya R (2016) Computational intelligence based qos-aware web service composition: a systematic literature review. IEEE Trans Serv Comput. doi:10.1109/TSC.2015.2473840

Charnes A, Cooper WW, Rhodes E (1978) Measuring the efficiency of decision making units. Eur J Oper Res 2(6):429–444

Chou CC (2010) An integrated quantitative and qualitative FMCDM model for location choices. Soft Comput 14(7):757–771

Cooper WW, Seiford LM, Zhu J (2011) Handbook on data envelopment analysis, vol 164. Springer Science & Business Media, New York

Ertay T, Ruan D (2005) Data envelopment analysis based decision model for optimal operator allocation in cms. Eur J Oper Res 164(3):800–810

Ertay T, Ruan D, Tuzkaya UR (2006) Integrating data envelopment analysis and analytic hierarchy for the facility layout design in manufacturing systems. Inf Sci 176(3):237–262

Esposito C, Ficco M, Palmieri F, Castiglione A (2015) Smart cloud storage service selection based on fuzzy logic, theory of evidence and game theory. IEEE Trans Comput. doi:10.1109/TC.2015.2389952

Garg SK, Versteeg S, Buyya R (2013) A framework for ranking of cloud computing services. Future Gener Comput Syst 29(4):1012–1023

Kuo R, Hsu C, Chen Y (2015) Integration of fuzzy anp and fuzzy TOPSIS for evaluating carbon performance of suppliers. Int J Environ Sci Technol 12(12):3863–3876

Kuo RJ, Lin Y (2012) Supplier selection using analytic network process and data envelopment analysis. Int J Prod Res 50(11):2852–2863

Kwon HK, Seo KK (2013) A decision-making model to choose a cloud service using fuzzy ahp. Adv Sci Technol Lett 35:93–96

Li A, Yang X, Kandula S, Zhang M (2010) Cloudcmp: comparing public cloud providers. In: Proceedings of the 10th ACM SIGCOMM conference on internet measurement, ACM, pp 1–14

Lin HT (2010) Personnel selection using analytic network process and fuzzy data envelopment analysis approaches. Comput Ind Eng 59(4):937–944

Lin MI, Lee YD, Ho TN (2011) Applying integrated DEA/AHP to evaluate the economic performance of local governments in China. Eur J Oper Res 209(2):129–140

Menzel M, Schönherr M, Tai S (2013) (mc2) 2: criteria, requirements and a software prototype for cloud infrastructure decisions. Softw Pract Exp 43(11):1283–1297

Mohajeri N, Amin GR (2010) Railway station site selection using analytical hierarchy process and data envelopment analysis. Comput Ind Eng 59(1):107–114

Ramanathan R (2006) Data envelopment analysis for weight derivation and aggregation in the analytic hierarchy process. Comput Oper Res 33(5):1289–1307

Saaty T (1980) Fundamentals of decision making and priority theory with analytical hierarchical process, vol 6. RWS Publications, University of Pittsburgh, Pittusburgh

Saaty TL (1988) What is the analytic hierarchy process?. Springer, Berlin

Saaty TL (1996) Analytical network process. RWS Publications, Pittsburgh

Saaty TL (2006) The analytic network process. Springer, Berlin

Seiford LM, Zhu J (1999) An investigation of returns to scale in data envelopment analysis. Omega 27(1):1–11

Shang J, Sueyoshi T (1995) A unified framework for the selection of a flexible manufacturing system. Eur J Oper Res 85(2):297–315

Shivakumar U, Ravi V, Gangadharan GR (2013) Ranking cloud services using fuzzy multi-attribute decision making. In: Proceedings of the IEEE international conference on fuzzy systems, pp 1–8

Silas S, Rajsingh EB, Ezra K (2012) Efficient service selection middleware using ELECTRE methodology for cloud environments. Inf Technol J 11(7):868–875

Sinuany-Stern Z, Mehrez A, Hadad Y (2000) An AHP/DEA methodology for ranking decision making units. Int Trans Oper Res 7(2):109–124

Tone K (2002) A slacks-based measure of super-efficiency in data envelopment analysis. Eur J Oper Res 143(3):32–41

Vaidya OS, Kumar S (2006) Analytic hierarchy process: an overview of applications. Eur J Oper Res 169(1):1–29

Wen M, Qin Z, Kang R, Yang Y (2015) Sensitivity and stability analysis of the additive model in uncertain data envelopment analysis. Soft Comput 19(7):1987–1996

Xu C, Ma Y, Wang X (2015) A non-parametric data envelopment analysis approach for cloud services evaluation. In: Proceedings of the service-oriented computing-ICSOC 2014 workshops, Springer, pp 250–255

Yan S, Chen C, Zhao G, Lee BS (2012) Cloud service recommendation and selection for enterprises. In: Proceedings of the 8th international conference on network and service management and workshop on systems virtualization management, IEEE, pp 430–434

Yang T, Kuo C (2003) A hierarchical AHP/DEA methodology for the facilities layout design problem. Eur J Oper Res 147(1):128–s136

Zheng Z, Wu X, Zhang Y, Lyu MR, Wang J (2013) Qos ranking prediction for cloud services. IEEE Trans Parallel Distrib Syst 24(6):1213–1222