

Cloud Service Selection Based on the Aggregation of User Feedback and Quantitative Performance Assessment

Lie Qu

*Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
lie.qu@mq.edu.au*

Yan Wang

*Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
yan.wang@mq.edu.au*

Mehmet A. Orgun

*Department of Computing
Macquarie University
Sydney, NSW 2109, Australia
mehmet.orgun@mq.edu.au*

Abstract—Cloud computing has been attracting huge attention in recent years. More and more individuals and organizations have been moving their work into cloud environments because of its flexibility and low-cost. Due to the emergence of a variety of cloud service providers, selecting the most suitable cloud service becomes increasingly important for potential cloud users. In prior studies, the selection and comparison of cloud services usually focus on objective performance analysis based on cloud monitoring and benchmark testing without considering the viewpoints of cloud users who are indeed consuming cloud services. This causes a problem that some vital aspects which concern cloud consumers (e.g., privacy and cloud providers' reputation) are not taken into account in cloud service selection.

In this paper, we propose a novel model of cloud service selection by aggregating the information from both the feedback from cloud users and objective performance analysis from a trusted third party. Based on this model, we first propose a framework which supports our cloud service selection approach. Then after classifying subjective assessment and objective assessment, we present a novel cloud service selection approach to aggregate all subjective assessments and objective assessments through a fuzzy simple additive weighting system. In addition, to reduce the bias caused by unreasonable feedback from unprofessional or malicious cloud users, a method is proposed for filtering the feedback from such users. After processing, the aggregated result can quantitatively reflect the overall quality of a cloud service. Finally, a case study is presented to illustrate the advantages of our approach.

Keywords—Cloud computing; Service selection; Subjective and objective attributes; Fuzzy simple additive weighting system;

I. INTRODUCTION

In recent years, cloud computing has become the most popular paradigm for storage and service solutions. Due to many advantages of cloud services, especially the pay-as-you-go model [1], companies or organizations can significantly reduce their costs by outsourcing their IT hardware and software to a variety of cloud providers. Comparing to traditional IT service models, there is no need for cloud users to pay a large amount of money for expensive IT equipment before setting up their own work, or to worry about these large upfront costs which could be devalued and eliminated in the future [8]. Cloud computing is service-oriented [19]. Cloud users without IT background can also consume the well-defined services for their own needs, and only pay for what they have consumed and promised to consume. Furthermore, the property of elasticity in cloud computing guarantees that the IT resources can be dynamically dis-

tributed to every cloud user near real-time [17], even though these users may be working in peak loads.

Because of these advantages, more and more individuals and organizations have started to consume cloud services. However, it should be noted that the widespread use of cloud computing is also accompanied by some problems. One big problem is how to evaluate the performance of a cloud service since it is quite necessary for potential cloud consumers to know the quality of services they will consume and pay for. To this end, many cloud performance monitoring, testing and comparison approaches [5] [11] [20] have been proposed. The usual way for cloud service evaluation is to compare the performance differences between similar cloud services. Such a comparison is usually based on the results of a predesigned set of benchmark tools [11] [10]. As cloud services are highly virtualized, the benchmark tools for traditional computation performance measurement can be appropriately applied in cloud environments. By combining these benchmark tools according to cloud features, many metrics can be quantitatively measured (e.g., the speed of CPU, memory read/write and storage, service response time and throughput).

Nevertheless, the benchmark testing results usually may not reflect the real performance of a cloud service for ordinary cloud users. This is because the testing environment is usually not the same as that of ordinary users' daily work, and a variety of real tasks currently executed in a cloud may not be perfectly simulated by a limited number of tests. In addition, these benchmark tests are usually spot-check tests. It is hard to carry out continuous testing because such tests might lead to the cost no less than that of consuming a real cloud service. Furthermore, some crucial but qualitative aspects of a cloud service can hardly be tested through such objective and quantitative measurement. For example, considering a company working for processing a large amount of sensitive customer data, the security and privacy of the data have a crucial impact on the company's survival. If the company plans to move their work into cloud in order to reduce the optional costs, it must choose a cloud provider which has very good reputation on data security and privacy. In addition, as the company is not a professional IT company, good and comprehensive after-sales services are highly desired. Moreover, due to the sensitivity of the company's data, a variety of encryption approaches are frequently applied in daily work. Hence, the speed of data encryption and decryption is a big concern for the company.

In this example, in addition to the typical performance of a cloud service (e.g., CPU, memory, response time and costs) that can be quantitatively tested by common benchmark testing, the company needs to carefully consider the data privacy and security and the quality of after-sales services, which can be very hard to quantify. In addition, the benchmark tests of the cryptographic calculation speed of a cloud service may need to be specifically designed according to the company's requirement. All the issues mentioned above may cause great difficulty for the company in cloud service selection.

In this paper, different from all the existing models of cloud service selection, we propose a novel model based on the aggregation of the feedback from cloud consumers and the objective performance measurement from a trusted third party's testing. After a discussion of some related work in Section 2, we propose the framework of our cloud service selection model in Section 3. Then, in Section 4, we present the details of our cloud service selection model that evaluates the performance of cloud services by aggregating all the subjective assessments and the objective assessments through a fuzzy simple additive weighting system [3]. It should be noted that cloud users' subjective assessment may be biased and inaccurate since they are usually not professional IT staff or even some of them are malicious users. Hence, in our model, cloud users' subjective assessments and the third party's objective assessment are compared, so that unreasonable subjective assessments can be filtered out before aggregation. This makes our approach more accurate and effective. After that, a case study is presented in Section 5 to illustrate the advantages of our model. We conclude this paper and discuss future work in Section 6.

II. RELATED WORK

In [7], Goscinski *et al.* present a detailed description of the process of cloud service publication, discovery and selection. Through a higher abstraction of clouds based on dynamical and characteristic attributes, they propose a general framework which is shown to be sound and feasible by a proof of concept. However, there is no concrete selection approach proposed in this work.

In the literature, the problem of cloud service selection can usually be solved by approaches of cloud service comparison based on objective performance analysis. In [11], Li *et al.* propose a systematic comparator, called *CloudCmp*. *CloudCmp* can be applied to compare three aspects of the performance and cost of a cloud (i.e., elastic computing, persistent storage, and intra-cloud and wide area networking). These comparisons are realized by a set of standard benchmark tools, whose results display the objective assessment of a cloud. In [2], Binning *et al.* consider the diversity of cloud services. They propose a benchmarking approach for cloud services based on the metrics of scalability, cost, peak-load and fault tolerance. Another discussion of cloud benchmark testing is presented by Lenk *et al.* [10]. In their work, they point out that the performance indicators provided by cloud providers may not be enough to judge the real performance of a virtual machine, and propose a new performance measurement method which considers the types of services executed on a virtual machine for Infrastructure-as-a-Service clouds. Recently, some third-

party organizations (e.g., CloudHarmony¹) have started to offer cloud monitoring and benchmarking services. When compared to the performance indicators provided by cloud providers, such third party testing may be more credible due to no direct profits involved. All the approaches mentioned above focus on analyzing the objective aspects of a cloud service based on quantitative testing. No subjective aspect is taken into account to reflect the overall performance of a cloud service.

In traditional e-commerce or e-service environments, service selection usually depends on the reputation-based trust evaluation of services. Comparing to early trust evaluation approaches based on computing a single trust value for a service, Li *et al.* [12] [14] propose several trust vector based evaluation approaches, where a trust vector is calculated to reflect both the current trustworthiness of a service and its trust trend. Such trust values or vectors are all evaluated from ratings which represent the subjective assessment of services given by service consumers [13]. In order to consider subjective aspects of a cloud service, Rehman *et al.* [20] propose a simple framework for monitoring cloud performance based on user feedback, in which the performance of a cloud service is monitored and predicted by users' feedback. Their approach only considers cloud users' subjective assessment. There is no mechanism to check the reliability of users' feedback. And the objective assessment of a cloud service is not considered in their framework.

Another solution for cloud service selection problem is to model the problem as a multi-criteria decision-making (MCDM) problem [9] which can be typically solved by Analytic Hierarchy Process (AHP) [18]. In [6], Godse *et al.* focus on the selection of Software-as-a-Service clouds based on AHP. Five factors (i.e., functionality, architecture, usability, vendor reputation and cost) are considered in their approach. It should be noted that all these factors except cost can hardly be quantified by an objective measure. Thus, their approach is still mainly based on subjective assessment. Another AHP-based cloud comparison approach is proposed by Garg *et al.* [5]. In their work, they attempt to standardize the performance attributes for cloud comparison. However, the standardization for some attributes (e.g., sustainability and transparency) is too simple to reflect the complex situations of cloud services in the real world.

III. THE PROPOSED FRAMEWORK

In this section, we present our framework of cloud service selection based on both the feedback from cloud users and the objective performance benchmark testing from a trusted third party. Figure 1 illustrates our framework, which consists of four components, namely, (1) *cloud selection service*, (2) *benchmark testing service*, (3) *user feedback management service*, and (4) *assessment aggregation service*, where *cloud selection service* is in the higher layer of our framework, and others are in the lower layer.

A. Cloud Selection Service

The *cloud selection service* is in charge of accepting and preliminarily processing the requests for cloud service selection from potential cloud consumers. In addition, it issues requests for the services from the lower layer components.

¹cloudharmony.com

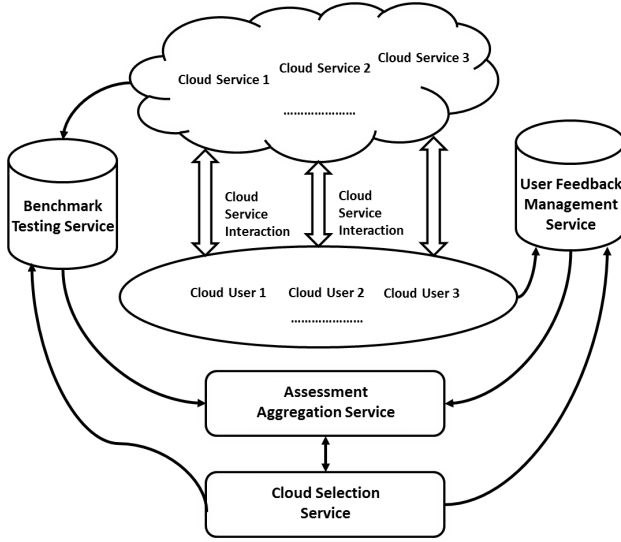


Figure 1. Our Proposed Framework for Cloud Service Selection

When a potential cloud consumer submits a request for selecting the most suitable cloud service, the *cloud selection service* firstly chooses the cloud services which can meet all the objective and quantitative requirements (e.g., the type of services, the specification of virtual machines and costs) of the potential user from a candidate list of cloud services. Then, according to the user's further requirements, it sends requests to the *benchmark testing service* and the *user feedback management service* for accessing the related records of alternative clouds. These records are then sent to the *assessment aggregation service*. By aggregating these records through our proposed model (see details in Section 4), the *assessment aggregation service* returns the final score of each alternative cloud to the *cloud selection service*. All these scores are shown to the potential user for cloud selection.

B. Benchmark Testing Service

The *benchmark testing service* is provided by a trusted third party which designs a variety of testing scenarios for the common performance aspects of a cloud service (e.g., availability, elasticity, service response time, and cost per task) by standard benchmark suites. In addition, some specific tests can be designed and run according to potential cloud consumers' needs, such as testing the speed of cryptographic calculations. The common performance tests can be executed in the spot-check form over time for a long period. The specific tests can run in a continuous but short period according to potential users' requirements. Each tested performance aspect of a cloud service can be considered as an *objective attribute* of the cloud service. All these objective attributes are expressed in quantified forms (e.g., 90% for availability, 200ms for response time or 47.5 benchmark scores for CPU performance). All the attribute values are recorded and maintained for the requests from the *cloud selection service*.

C. User Feedback Management Service

The *user feedback management service* is used to collect and manage the feedback from the users who are consuming

cloud services. For each performance aspect of a cloud service, a user gives his/her subjective assessment according to his/her intuition. Each aspect that users assess can be considered as the *subjective attribute* of the cloud service. These subjective attributes are expressed by linguistic variables (e.g., "good", "fair" and "poor").

Note that some subjective attribute and some objective attribute can represent the same performance aspect of a cloud service. For example, the response time of a cloud service can be accurately calculated by benchmark testing under different circumstances. By analyzing the testing results, an objective assessment of the service response time can be given. Meanwhile, a user of this cloud can also give subjective assessment of response time by feeling how long the cloud responds to his/her requests in daily work. Figure 2 introduces an example to explain the relationship between these attributes. In our framework, such attributes (e.g., response time which belongs to both subjective attributes and objective attributes) are named as *associated attributes*. In Figure 2, for a cloud service, we assume there are s subjective attributes, o objective attributes and u associated attributes for a cloud service ($u \leq s, u \leq o$), where *privacy*, *after-sales services*, *availability* and *response time* are its *subjective attributes* extracted from users' feedback. On the other hand, it also has *availability*, *response time* and *cryptographic calculation* as its *objective attributes* extracted from a third party's benchmark testing. Thus, *availability* and *response time* are considered as the *associated attributes*. The *associated attributes* from subjective assessment are called *subjective associated attributes*, and those from objective assessment are called *objective associated attributes*. Therefore, each *subjective associated attribute* corresponds to an *objective associated attribute*.

Furthermore, in some cases, the users' subjective assessment of a cloud service may not be absolutely subjective. For those users with IT background, some simple testing tools or status checking commands (e.g., Command *xentop* for Xen Hypervisor²) can be used to help them make an assessment. In such a situation, we consider that the assessment still belongs to subjective assessment since the results from such cloud users' testing are usually incomprehensive and inaccurate without scientific and statistical analysis.

D. Assessment Aggregation Service

The *assessment aggregation service* is in charge of aggregating the values of the subjective attributes from the *user feedback management service* and the objective attributes from the *benchmark testing service*, and computing the final score for each alternative cloud service according to the importance weights that are set by a potential cloud user in the form of linguistic variables. In the example introduced in Section 1, suppose that the company is quite concerned about the security and privacy issues of user data and hopes to receive excellent after-sales service. Therefore, the importance weights for *security* and *privacy* as the attribute of a cloud service can be set as "very high", and the weight for *after-sales service* can be set as "high". Likewise, all the other subjective attributes and objective attributes are given such importance weights. By using these weights, the potential cloud user (i.e., the company) can also determine

²xen.org

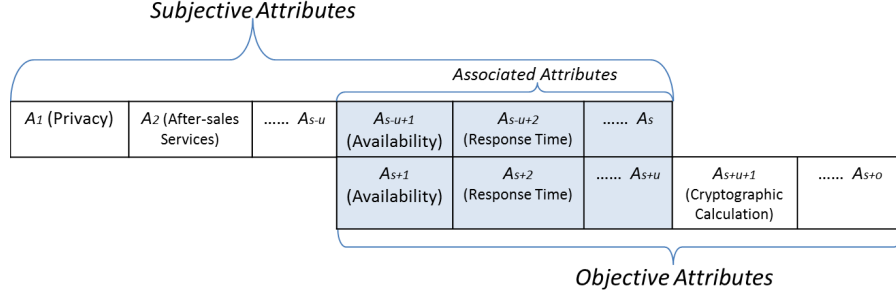


Figure 2. The Relationship of Subjective Attributes and the Objective Attributes

whether to put more trust on subjective assessment or objective assessment, so that the final score based on aggregating all these attributes can comprehensively reflect the various needs of the potential cloud user (i.e., the company). In Section 4, we introduce the details of this aggregation process.

IV. THE CLOUD SERVICE SELECTION MODEL

In this section, we propose a new cloud service selection model which is based on comparing and aggregating the subjective attributes and the objective attributes of a cloud service. In our model, for the sake of simplicity, we assume there are sufficient users offering their subjective assessments, and there is only one third party offering objective assessment for a cloud service. That is, there is only one set of values for objective attributes and there are many sets of values for subjective attributes for one cloud service. In addition, the third party is assumed to be a trusted third party which offers honest results of the objective performance testing. The trustworthiness of the third party and the situation with multiple third parties will be considered in our future work. Hence, in our model, we take objective assessment as a benchmark to evaluate the accuracy of subjective assessments, since objective assessment is usually more accurate to reflect the real performance of a cloud service than subjective feeling.

In order to compare and aggregate subjective attributes and objective attributes, the values of these attributes should be normalized. Here, we apply a fuzzy simple additive weighting system (FSAWS) proposed by Chou *et al.* [3] to convert all the attribute values into ratings. Firstly, after gathering both the values of the subjective attributes and the objective attributes of an alternative cloud service, the values of the subjective attributes are converted into ratings through a mapping from linguistic variables to fuzzy numbers. Secondly, the values of the objective attributes are also converted into the ratings by comparing the values of the same objective attributes in all the alternative cloud services. Thirdly, as introduced above, for an alternative cloud service there are many sets of values for the subjective attributes from different users' feedback and only one set of values for the objective attributes from the third party's testing. We compute the Euclidean distance [4] between the set of values for the objective associated attributes and each set of the values for the subjective associated attributes through the corresponding normalized ratings. The set of values for the subjective attributes with the Euclidean distance over a threshold (e.g., 80% of the maximum distance) will be considered unreasonable and eliminated from the

following aggregation of all the attributes. After that, the importance weight for each attribute is computed according to the potential cloud user's preference. Finally, a score is calculated for each alternative cloud service. The higher the score, the better the overall performance of an alternative cloud service.

In real world situations, subjective assessment for cloud services and the importance weight for each performance attribute are usually represented in the form of linguistic variables (e.g., "good" and "bad"). In order to deal with the inherent uncertainty of human languages, we apply a fuzzy simple additive weighting system in our model. Through this system, linguistic variables can be represented by fuzzy numbers for their fuzziness. And quantitative terms can also be represented in fuzzy number form. Hence, by using this system, our model can effectively normalize and aggregate all different types of subjective attributes and objective attributes in real world situations. Before presenting the details of our model, some basic knowledge of FSAWS will be introduced.

A. Fuzzy Simple Additive Weighting System

The fuzzy simple additive weighting system [3] is originally proposed for solving the facility location selection problem which is a multiple attributes decision-making problem under homo/heterogeneous group decision-making environments. In this system, the decision-makers are grouped together to determine the most suitable facility location. The evaluation of an alternative location depends on a variety of subjective attributes and objective attributes. The importance weight and the rating value for each attribute are represented by trapezoidal fuzzy numbers. An aggregated fuzzy score is calculated for each alternative location through some fuzzy set operations. Then, each fuzzy score is defuzzified to a crisp number as the final score of each alternative location.

To explain how this system works, some basic knowledge of fuzzy numbers is introduced in [3].

1. *Trapezoidal Fuzzy Number:* $\tilde{A} = (a, b, c, d)$ is a fuzzy set on \mathbb{R} , where $a < b < c < d$ are real numbers. If

$$\mu_{\tilde{A}}(x) = \begin{cases} \frac{x-a}{b-a}, & a \leq x \leq b \\ 1, & b \leq x \leq c \\ \frac{x-d}{c-d}, & c \leq x \leq d \\ 0, & \text{otherwise,} \end{cases}$$

then $\tilde{A} = (a, b, c, d)$ is called a trapezoidal fuzzy number, where $\mu_{\tilde{A}}(x)$ is its membership function. The most probable

value of the evaluation data is represented in the interval $[b, c]$. The intervals $[a, b]$ and $[c, d]$ show the fuzziness of the evaluation data. For example, “good” can be represented by $(5, 7, 7, 10)$, and “very good” can be represented by $(7, 10, 10, 10)$. In addition, some quantitative terms can also be represented by trapezoidal fuzzy numbers. For example, “equal to 50” can be represented by $(50, 50, 50, 50)$, and “approximately equal to 700” can be represented by $(690, 700, 700, 710)$.

2. Operations of Trapezoidal Fuzzy Numbers: Given two trapezoidal fuzzy numbers $\tilde{A} = (a, b, c, d)$ and $\tilde{B} = (e, f, g, h)$, and a real number k , some operations are defined as follows:

Addition:

$$\tilde{A} \oplus \tilde{B} = (a + e, b + f, c + g, d + h), a \geq 0, e \geq 0.$$

Multiplication:

$$\tilde{A} \otimes \tilde{B} = (ae, bf, cg, dh), a \geq 0, e \geq 0;$$

$$k \otimes \tilde{A} = (ka, kb, kc, kd), a \geq 0, k \geq 0;$$

$$\tilde{A} \otimes k = (ka, kb, kc, kd), a \geq 0, k \geq 0.$$

Division:

$$\tilde{A} / \tilde{B} = (\frac{a}{h}, \frac{b}{g}, \frac{c}{f}, \frac{d}{e}), a \geq 0, e > 0;$$

$$k / \tilde{A} = (\frac{k}{d}, \frac{k}{c}, \frac{k}{b}, \frac{k}{a}), a > 0, k \geq 0;$$

$$\tilde{A} / k = (\frac{a}{k}, \frac{b}{k}, \frac{c}{k}, \frac{d}{k}), a \geq 0, k > 0.$$

Commutative operations:

$$\tilde{A} \oplus \tilde{B} = \tilde{B} \oplus \tilde{A}, a \geq 0, e \geq 0;$$

$$\tilde{A} \otimes \tilde{B} = \tilde{B} \otimes \tilde{A}, k \otimes \tilde{A} = \tilde{A} \otimes k, k \geq 0, a \geq 0, e \geq 0.$$

3. Defuzzification: Taking a trapezoidal fuzzy number $\tilde{A} = (a, b, c, d)$ as input, the defuzzified output is a crisp number defined by computing the signed distance of \tilde{A} :

$$d(\tilde{A}) = \frac{1}{4}(a + b + c + d). \quad (1)$$

B. Our Proposed Approach

In this section, we present the details of our cloud service selection model. We modify the fuzzy simple additive weighting system [3] in order to fit for our targeted problem. In addition, in our model, unreasonable subjective assessments can be filtered out before the aggregation of all the assessments. Therefore, the aggregated score of each alternative cloud service can more accurately reflect the overall performance of a cloud service with less noise.

Assume that a potential cloud user submits its request to the *cloud selection service* for finding the most suitable cloud service meeting all the user’s requirements. After the preliminary selection according to the functional requirements, suppose that there are m clouds left as the alternative clouds denoted by $\{C_j\}$, where $j = 1, \dots, m$. The final score of each alternative cloud is computed based on s subjective attributes extracted from cloud users’ feedback and o objective attributes extracted from the benchmark testing of a trusted third party, where there are u ($u \leq s, u \leq o$) associated attributes. Thus, all the subjective attributes and the objective attributes are denoted as $\{A_i\}$, where $i = 1, 2, \dots, s + o$. $\{A_i\}$ ($i = 1, \dots, s$) denotes the subjective attributes, where $\{A_i\}$ ($i = s - u + 1, \dots, s$) denotes the subjective associated attributes; $\{A_i\}$ ($i = s + 1, \dots, s + o$) denotes the objective attributes, where $\{A_i\}$ ($i = s + 1, \dots, s + u$) denotes the objective associated

Table I
DECISION-MAKERS OF AN ALTERNATIVE CLOUD C_j

Attributes	1	2	...	$s + o$
Decision-makers				
1	A_{1j1}	A_{2j1}	...	$A_{(s+o)j1}$
2	A_{1j2}	A_{2j2}	...	$A_{(s+o)j2}$
...
n	A_{1jn}	A_{2jn}	...	$A_{(s+o)jn}$

attributes. The corresponding objective associated attribute of the subjective associated attribute A_i is A_{i+u} for each $i = s - u + 1, \dots, s$.

For an alternative cloud C_j , the *user feedback management service* returns n feedbacks given by n cloud users. Each feedback includes s values in the form of linguistic variables corresponding to the s subjective attributes. The feedbacks are denoted as $\{F_{jk}\}$, where $k = 1, \dots, n$. On the other hand, the *benchmark testing service* returns one benchmark testing report for each alternative cloud. Each report includes o values in the form of quantitative terms corresponding to the o objective attributes. The reports are denoted as $\{T_j\}$, where $j = 1, \dots, m$. m is the number of all the alternative cloud services.

It should be clarified that, for an alternative cloud, there are n feedbacks and only one testing report. That is, all the n feedbacks correspond to only one testing report since we assume there is only one trusted third party in our model. Here, for an alternative cloud, we connect the set of values for the o objective attributes from the only one testing report to the set of values for the s subjective attributes from each feedback. The final score of an alternative cloud is computed based on aggregating all the sets of values of the $s + o$ attributes. We define such a set of values of the $s + o$ attributes as a decision-maker, denoted as DM_{jk} ($k = 1, \dots, n$), for an alternative cloud C_j . Thus, the value of i th attribute of the j th alternative cloud from the k th decision-maker can be denoted as A_{ijk} .

Table I illustrates the decision-makers of an alternative cloud C_j . There are n decision-makers for C_j . n is the number of user feedbacks for C_j . Note that, in Table I, for each $i = s + 1, \dots, s + o$, all the A_{ijk} are the same for different decision-makers DM_{jk} since they are all the values of the objective attributes from the only one testing report T_j . That is, for C_j , $A_{ij1} = A_{ij2} = \dots = A_{ijn}$ for each $i = s + 1, \dots, s + o$.

The detailed procedure of our approach is shown below:

Step 1 (Converting the values of subjective attributes into ratings): As we have introduced before, in our model, all the subjective attributes are represented in the form of linguistic variables. By using a mapping from linguistic variables to fuzzy numbers, the linguistic values for the subjective attributes (e.g., “good” and “fair”) can be converted into fuzzy numbers. Here, we use the mapping illustrated in Table II from [3], which is frequently employed in research of the multi-criteria decision-making problem for real world situations, such as [16] and [15]. Each fuzzy number in Table II represents the fuzzy rating corresponding to the linguistic variable, denoted as \tilde{r} . In addition, a crisp rating corresponding to each linguistic variable is computed by defuzzifying its fuzzy rating with the signed distance (Equation (1)), which is denoted as r . Hence, for a decision-maker DM_{jk} of an alternative cloud C_j , \tilde{r}_{ijk} and r_{ijk} denote the fuzzy rating and the crisp rating for A_{ijk} respectively,

where $i = 1, \dots, s$.

Step 2 (Converting the values of objective attributes into ratings): As we have introduced in Section 4.1, the values in the form of quantitative terms can also be represented by fuzzy numbers. For any decision-maker DM_{jk} of an alternative cloud C_j , the values of the objective attributes $\{A_{ijk}\}$ ($i = s+1, \dots, s+o$) are represented in the form of fuzzy numbers. Then, these fuzzy numbers are converted into fuzzy ratings by comparing the values of the same objective attribute in all the alternative clouds. Let $\tilde{x}_{ijk} = (a_{ijk}, b_{ijk}, c_{ijk}, d_{ijk})$ ($i = s+1, \dots, s+o$) be the fuzzy numbers of the objective attribute values of the cloud C_j for any DM_{jk} . The fuzzy rating of each objective attribute value is computed as follows:

$$\tilde{r}_{ijk} = (\tilde{x}_{ijk} / \max_j(d_{ijk})) \otimes 100, \text{ where} \quad (2)$$

$$i = s+1, \dots, s+o, j = 1, \dots, m, k = 1, \dots, n.$$

$$\tilde{r}_{ijk} = (\min_j(a_{ijk}) / \tilde{x}_{ijk}) \otimes 100, \text{ where} \quad (3)$$

$$i = s+1, \dots, s+o, j = 1, \dots, m, k = 1, \dots, n.$$

Equation (2) is for the situation where the larger objective attribute value is the better (e.g., benchmark score for CPU). And Equation (3) is for the situation where the smaller objective attribute value is the better (e.g., response time).

Step 3 (Filtering unreasonable subjective assessments): So far, all the values in Table I have been converted into the form of fuzzy ratings. For each decision-maker DM_{jk} ($k = 1, \dots, n$) of an alternative cloud C_j , the Euclidean distance between the ratings of the corresponding subjective associated attributes and the objective associated attributes is computed as follows:

$$ED_{jk} = \sqrt{\sum_{i=s-u+1}^s (d(\tilde{r}_{ijk}) - d(\tilde{r}_{(i+u)jk}))^2} \quad (4)$$

$$= \sqrt{\sum_{i=s-u+1}^s (r_{ijk} - r_{(i+u)jk})^2}.$$

ED_{jk} represents the aggregated difference between the crisp ratings of the corresponding subjective associated attributes and the objective associated attributes of the decision-maker DM_{jk} for the alternative cloud C_j . In our model, we take the objective assessment as the benchmark to filter out unreasonable subjective assessments. If the distance exceeds a threshold (e.g., 80% of the maximum Euclidean distance), the decision-maker offering such values of the subjective attributes will be removed from the list of the decision-makers for an alternative cloud.

Step 4 (Computing the importance weight for each attribute): According to the potential cloud user's requirement, an importance weight in the form of linguistic variables is given to each subjective or objective attribute, so that the user can determine the importance of each attribute. In addition, the user can also determine how much to trust the subjective or objective assessment through these importance weights. A fuzzy weight in the form of fuzzy numbers is given to each attribute, which is denoted as \tilde{W}_i , where $i = 1, \dots, s+o$. Table III [3] illustrates the mapping from linguistic variables to fuzzy weights. This mapping is also

Table II
MAPPING FROM LINGUISTIC VARIABLES TO FUZZY RATINGS AND CRISP RATINGS [3]

Linguistic Variables	Fuzzy Ratings (\tilde{r})	Crisp Ratings (r)
Very poor (VP)	(0, 0, 0, 20)	5
Between very poor and poor (B.VP&P)	(0, 0, 20, 40)	15
Poor (P)	(0, 20, 20, 40)	20
Between poor and fair (B.P&F)	(0, 20, 50, 70)	35
Fair (F)	(30, 50, 50, 70)	50
Between fair and good (B.F&G)	(30, 50, 80, 100)	65
Good (G)	(60, 80, 80, 100)	80
Between good and very good (B.G&VG)	(60, 80, 100, 100)	85
Very good (VG)	(80, 100, 100, 100)	95

Table III
MAPPING FROM LINGUISTIC VARIABLES TO FUZZY WEIGHTS [3]

Linguistic Variables	Fuzzy Weights
Very low (VL)	(0, 0, 0, 3)
Low (L)	(0, 3, 3, 5)
Medium (M)	(2, 5, 5, 8)
High (H)	(5, 7, 7, 10)
Very High (VH)	(7, 10, 10, 10)

frequently used in prior studies for real world situations [15] [16]. Then, the fuzzy weights are defuzzified by computing their signed distances (Equation (1)). The crisp weight of the attribute A_i is denoted as W_i which is computed as follows:

$$W_i = \frac{d(\tilde{W}_i)}{\sum_{i=1}^{s+o} d(\tilde{W}_i)}, \text{ where } i = 1, \dots, s+o. \quad (5)$$

Step 5 (Aggregating all attributes): Assume there are n' decision-makers left for the cloud C_j after Step 3. A matrix \tilde{M}_j based on the fuzzy ratings of each attribute from different decision-makers is constructed as follows:

$$\tilde{M}_j = \begin{bmatrix} \tilde{r}_{1j1} & \tilde{r}_{2j1} & \dots & \tilde{r}_{(s+o)j1} \\ \tilde{r}_{1j2} & \tilde{r}_{2j2} & \dots & \tilde{r}_{(s+o)j2} \\ \dots & \dots & \dots & \dots \\ \tilde{r}_{1jn'} & \tilde{r}_{2jn'} & \dots & \tilde{r}_{(s+o)jn'} \end{bmatrix}.$$

According to the crisp weight of each attribute, the fuzzy scores of the cloud C_j from every decision-maker DM_{jk} ($k = 1, \dots, n'$) are computed as follows:

$$\tilde{S}_j = \tilde{M}_j \otimes \begin{bmatrix} W_1 \\ W_2 \\ \dots \\ W_{s+o} \end{bmatrix} = \begin{bmatrix} \tilde{f}_{j1} \\ \tilde{f}_{j2} \\ \dots \\ \tilde{f}_{jn'} \end{bmatrix}, \quad (6)$$

where \tilde{f}_{jk} is the fuzzy score of C_j from DM_{jk} ($k = 1, \dots, n'$). Here, the operation \otimes is generalized to matrices in the standard way of matrix multiplication. Then, the final score of C_j is computed as follows:

$$\bar{S}_j = \frac{1}{n'} \left(\sum_{k=1}^{n'} d(\tilde{f}_{jk}) \right). \quad (7)$$

Finally, according to the final scores, all the alternative cloud services are ranked for the selection of the potential cloud user.

V. A CASE STUDY

In this section, a case study over the example introduced in Section 1 is presented to illustrate the effectiveness of our proposed model for cloud service selection.

Assume that the company which plans to consume a cloud service requests to select the most suitable cloud services

Table IV
THE FUZZY VALUES OF THE OBJECTIVE ATTRIBUTES FOR EACH
ALTERNATIVE CLOUD

Alter. Clouds	Obj. Attr.	A_5 (%)	A_6 (ms)	A_7 (scores)
C_1		98.8, 99.2, 99.2, 99.8	50, 56, 56, 60	136.4, 136.4, 136.4, 136.4
C_2		99.2, 99.8, 99.8, 100	75, 83, 83, 90	102.5, 102.5, 102.5, 102.5
C_3		97.9, 98.5, 98.5, 99.1	180, 200, 200, 220	62.6, 62.6, 62.6, 62.6
C_4		98.2, 98.8, 98.8, 99.4	160, 173, 173, 190	69.8, 69.8, 69.8, 69.8

based on seven attributes, including four subjective attributes and three objective attributes. The four subjective attributes are *privacy* (A_1), *after-sales service* (A_2), *availability* (A_3) and *service response time* (A_4); and the three objective attributes are *availability* (A_5), *service response time* (A_6) and *cryptographic calculation speed* (A_7), where *availability* and *service response time* are the associated attributes.

Step 1: Assume that there are four alternative cloud services left after preliminary selection. For each alternative cloud, there are ten feedbacks from the *user feedback management service* to assess the four subjective attributes in the form of linguistic variables. According to Table II, all the linguistic variables are mapped into the fuzzy numbers which represent the fuzzy ratings for the subjective attributes.

Step 2: Every alternative cloud service has a performance report for the three objective attributes from the *benchmark testing service*. The value of the attribute *availability* is the average percentage of the times of successfully accessing a cloud service over the last twelve months. The value of the attribute *service response time* is the average of the service response time of every spot-check test over the last twelve months. The fuzziness of these values depends on the minimum value and the maximum value for each attribute recorded in the last twelve months. The speed of cryptographic calculation is represented by a benchmark score from the short-term specific tests according to the company's requirement. The higher the score, the faster the speed. Table IV shows the fuzzy values of the three objective attributes for the four alternative clouds. And according to Equations (2)&(3), all these fuzzy values are converted into fuzzy ratings. Table V shows the fuzzy ratings corresponding to Table IV.

Step 3: Combining the fuzzy ratings for all the subjective attributes and the objective attributes, the decision-makers are listed for an alternative cloud. Due to space limitations, we only list the decision-makers for C_1 in Table VI. The Euclidean distance between the corresponding associated attribute pairs are computed for each decision-maker with Equation (4). For DM_4 and DM_8 , we can see that the ratings of the subjective associated attributes A_3 and A_4 are much lower than the ratings of the corresponding objective associated attributes A_5 and A_6 . By Equation (4), the Euclidean distances of DM_4 and DM_8 are 120.02 and 113.11 respectively, which exceed the threshold of 107.48 (i.e., 80% of the maximum distance of 134.35). Thus, DM_4 and DM_8 are considered unreasonable and filtered out of the decision-maker list.

Step 4: According to the potential cloud user (i.e., the company)'s requirement, the importance weight for each attribute is given in the form of linguistic variables. According to Table III, all the weights are mapped into the fuzzy weights which are then converted into the crisp weights by Equation (5). All the weights are shown in Table VII.

Table V
THE FUZZY RATINGS OF THE OBJECTIVE ATTRIBUTES FOR EACH
ALTERNATIVE CLOUD

Alter. Clouds	Obj. Attr.	A_5 (\bar{r}_5)	A_6 (\bar{r}_6)	A_7 (\bar{r}_7)
C_1		98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
C_2		99.2, 99.8, 99.8, 100	55.56, 60.24, 60.24, 66.67	75.15, 75.15, 75.15, 75.15
C_3		97.9, 98.5, 98.5, 99.1	22.73, 25, 25, 27.78	45.89, 45.89, 45.89, 45.89
C_4		98.2, 98.8, 98.8, 99.4	26.32, 28.90, 28.90, 31.25	51.17, 51.17, 51.17, 51.17

Step 5: According to Table VI and Table VII, the fuzzy scores of C_1 based on the remaining eight decision-makers are computed by Equation (6). Finally, by Equation (7), the final score of C_1 is 74.3.

Results: Table VIII shows the scores of the four alternative clouds. Row 2 presents the scores based on our model considering both the subjective attributes and the objective attributes. Row 3 and Row 4 present the scores of considering only the subjective attributes and only the objective attributes respectively according to the corresponding importance weights. The ranking computed based on our model is C_1, C_2, C_3, C_4 . The ranking computed based on the subjective assessment only is C_2, C_1, C_3, C_4 . And the ranking computed based on the objective assessment only is C_1, C_2, C_4, C_3 .

From Table VIII, we can see that, although the subjective assessment from cloud users places C_1 as the second best, C_1 is the best cloud service according to our model since our model considers both its subjective assessment and objective assessment. It should be noted that the gap of the scores between C_1 and C_2 based on our model is smaller than the gap of the scores between them based on considering only their objective assessments. That is because our model considers subjective assessment for some important aspects of a cloud service, which can hardly be measured by quantitative testing. Thus, in this case, C_1 is the best cloud service by comprehensively considering all the concerned performance aspects of a cloud service according to the potential cloud user (i.e., the company)'s requirements.

Based on the analysis in the above case study, our proposed model of cloud service selection has the following advantages:

(1) By considering subjective assessment from cloud consumers, our model takes into account some vital but qualitative performance aspects in the selection process of a cloud service as well as quantitative performance aspects.

(2) Our model considers the situation in the real world, where cloud users' subjective assessments are fuzzy in linguistic form as well as the importance weight for each performance attribute. Thus, our model can effectively deal with the uncertainty of human languages in cloud service selection.

(3) According to different concerns of potential cloud users for different subjective attributes and objective attributes, our model presents an overall performance score for a cloud service by aggregating all subjective assessments and objective assessments with less noise from unreasonable subjective assessments.

VI. CONCLUSION

In this paper, we have proposed a novel model of cloud service selection by aggregating subjective assessment from cloud consumers and objective performance assessment from

Table VI
THE DECISION-MAKERS OF THE ALTERNATIVE CLOUD C_1

Attributes	A_1	A_2	A_3	A_4	A_5	A_6	A_7
Decision-makers							
1	B.P&F	VP	B.G&VG	B.G&VG	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
2	B.P&F	G	P	F	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
3	G	VG	B.F&G	G	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
4	B.G&VG	B.F&G	B.VP&P	VP	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
5	G	G	B.F&G	B.P&F	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
6	P	B.VP&P	G	G	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
7	B.VP&P	F	B.F&G	B.G&VG	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
8	F	VG	B.VP&P	B.VP&P	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
9	B.F&G	B.F&G	B.G&VG	B.P&F	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100
10	B.P&F	B.F&G	B.F&G	B.P&F	98.8, 99.2, 99.2, 99.8	83.33, 89.29, 89.29, 100	100, 100, 100, 100

Table VII
THE WEIGHTS OF EACH ATTRIBUTE

Attributes	Weights	Linguistic Variable	Fuzzy Weights	Crisp Weights
A_1		Very High	7,10,10,10	0.2022
A_2		High	5,7,7,10	0.1585
A_3		Low	0,3,3,5	0.0601
A_4		Medium	2,5,5,8	0.1093
A_5		Medium	2,5,5,8	0.1093
A_6		High	5,7,7,10	0.1585
A_7		Very High	7,10,10,10	0.2022

Table VIII
THE SCORES OF EACH ALTERNATIVE CLOUD

Approaches	C_1	C_2	C_3	C_4
Sub.&Obj.	74.3	67.8	51.5	50.6
Sub.	54.4	60.5	51.8	47.0
Obj.	96.6	76.0	51.1	54.7

a trusted third party. In order to consider real world situations, we apply a fuzzy simple additive weighting system to normalize and aggregate all different types of subjective attributes and objective attributes of a cloud service, so that some specific performance aspects of a cloud service can also be taken into account according to potential cloud users' requirements. In addition, our model can identify and filter unreasonable subjective assessments. This makes the results delivered by our model more accurate and effective with less noise. Finally, a case study demonstrates the advantages of our proposed model in cloud service selection.

For future work, we plan to apply different fuzzy weight systems and study the effect of different system factors to our model (e.g., defuzzification methods, sensitivity of membership functions and mappings from linguistic variables to fuzzy numbers). In addition, we plan to improve our model for the situation with multiple third parties which offer different testing results for cloud services. And the trustworthiness of every third party will be evaluated and considered in our improved model.

REFERENCES

- [1] M. Armbrust, A. Fox, R. Griffith, A. D. Joseph, R. H. Katz, A. Konwinski, G. Lee, D. A. Patterson, A. Rabkin, I. Stoica, and M. Zaharia. A view of cloud computing. *Commun. ACM*, 53(4):50–58, 2010.
- [2] C. Binnig, D. Kossmann, T. Kraska, and S. Loesing. How is the weather tomorrow?: towards a benchmark for the cloud. In *ACM SIGMOD Conference*, 2009.
- [3] S.-Y. Chou, Y.-H. Chang, and C.-Y. Shen. A fuzzy simple additive weighting system under group decision-making for facility location selection with objective/subjective attributes. *European Journal of Operational Research*, 189(1):132–145, 2008.
- [4] M. M. Deza and E. Deza. *Encyclopedia of Distances*. Springer, 2009.
- [5] S. K. Garg, S. Versteeg, and R. Buyya. SMICloud: A framework for comparing and ranking cloud services. In *Utility and Cloud Computing*, pages 210–218, 2011.
- [6] M. Godse and S. Mulik. An approach for selecting Software-as-a-Service (SaaS) product. In *IEEE CLOUD*, pages 155–158, 2009.
- [7] A. Goscinski and M. Brock. Toward dynamic and attribute based publication, discovery and selection for cloud computing. *Future Generation Comp. Syst.*, 26(7):947–970, 2010.
- [8] M. Y. Hajjat, X. Sun, Y.-W. E. Sung, D. A. Maltz, S. G. Rao, K. Sripanidkulchai, and M. Tawarmalani. Cloudward bound: planning for beneficial migration of enterprise applications to the cloud. In *ACM SIGCOMM Conference*, pages 243–254, 2010.
- [9] H. K. Jain, M. Tanniru, and B. Fazlollahi. MCDM approach for generating and evaluating alternatives in requirement analysis. *Information Systems Research*, 2(3):223–239, 1991.
- [10] A. Lenk, M. Menzel, J. Lipsky, S. Tai, and P. Offermann. What are you paying for? performance benchmarking for Infrastructure-as-a-Service offerings. In *IEEE CLOUD*, pages 484–491, 2011.
- [11] A. Li, X. Yang, S. Kandula, and M. Zhang. CloudCmp: comparing public cloud providers. In *Internet Measurement Conference*, pages 1–14, 2010.
- [12] L. Li and Y. Wang. A trust vector approach to service-oriented applications. In *IEEE International Conference on Web Services (ICWS)*, pages 270–277, 2008.
- [13] L. Li and Y. Wang. Subjective trust inference in composite services. In *the Twenty-Fourth AAAI Conference on Artificial Intelligence (AAAI)*, 2010.
- [14] L. Li, Y. Wang, and V. Varadharajan. Fuzzy regression based trust prediction in service-oriented applications. In *the 7th International Conference on Autonomic and Trusted Computing (ATC)*, pages 221–235, 2009.
- [15] G.-S. Liang. Fuzzy MCDM based on ideal and anti-ideal concepts. *European Journal of Operational Research*, 112(3):682–691, 1999.
- [16] G.-S. Liang and M.-J. J. Wang. Fuzzy multi-criteria decision-making method for facility site selection. *International Journal of Production Research*, 29(11):2313–2330, 1991.
- [17] M. Mao and M. Humphrey. A performance study on the VM startup time in the cloud. In *IEEE CLOUD*, pages 423–430, 2012.
- [18] K. Muralidhar, R. Santhanam, and R. L. Wilson. Using the analytic hierarchy process for information system project selection. *Information & Management*, 18(2):87–95, 1990.
- [19] J. Namjoshi and A. Gupte. Service oriented architecture for cloud based travel reservation Software as a Service. In *IEEE CLOUD*, pages 147–150, 2009.
- [20] Z. ur Rehman, O. K. Hussain, S. Parvin, and F. K. Hussain. A framework for user feedback based cloud service monitoring. In *International Conference on Complex, Intelligent and Software Intensive System*, pages 257–262, 2012.