CLOUD SERVICE BROKER: SELECTION OF PROVIDERS USING DTRFV EVALUATION

1ANUPRIYA KONERU, 2SREELATHA M

1Research Scholar, Acharya Nagarjuna University, Department of Computer Science and Engineering, Guntur, India
2Professor, RVR & JC College of Engineering, Department of CSE, Guntur, India
E-mail: 1smartykonuru@gmail.com, 2lathamoturi@rediffmail.com

ABSTRACT

Multicloud environment possess new challenges to the Cloud Service Requesters. Cloud Service Requesters are not aware of the quality of service offered by Cloud Service Providers. To select the suitable provider for a service the requester meets the Cloud Service Broker. Broker finds the best suitable provider for every request. Broker identifies the previous service ratings to evaluate the Reputation Factor Value of provider. In this paper, we present Dynamic True Reputation Factor Value Evaluation algorithm to find the best providers. By using this algorithm Broker can identify the unfair ratings and reduce the effect of unfair ratings in the calculation of True Reputation Factor Value of provider. Results are presented to appraise the success of the proposed model.

Keywords: Multicloud, Cloud service requester, Broker, Reputation, Provider

1. INTRODUCTION

A Multi cloud environment consists of huge number of cloud service providers to serve the needs of different organizations which are scaled up. It is really a great opportunity to the Cloud Service Requesters for choosing the Cloud service Providers in this kind of environment. In this place, the Cloud Service Requesters can take different services from different Cloud Service Providers. If a particular service is available with multiple Cloud Service Providers then the selection of a single Cloud Service Provider is really a difficult task to the Cloud Service Requester. So, in the industry most of the organizations are contacting a mediator to select the best suitable Cloud Service Provider. The responsibility of the mediator is to choose the suitable Cloud Service Provider on behalf of Cloud Service Requester. In the IT Industry, the mediator is termed as a CLOUD SERVICE BROKER. The Cloud Service Providers and Cloud Service Requesters are interested to communicate through Cloud Service Broker for simplifying their work.

In the Multi Cloud Environment, the cloud service providers and cloud service requesters register with the Broker. Cloud Service Broker maintains all the services offered by each registered Cloud Service Provider and accepts the request from the Cloud service Requesters with all the required specifications. Based on the specifications of each Cloud Service Requester, Cloud Service Broker selects that particular Cloud Service Provider. The advantage of using Cloud Service Broker is to get the maximum services at minimum cost from the best suitable Cloud Service Provider to the Cloud Service Requester. This advantage really attracts the entire IT Industry towards Cloud Service Brokering. There are Number of Cloud Service Brokers in the real environment which are running profitably. In this environment, the Cloud Service Requesters have to pay the amount to the Cloud Service Broker when it takes suggestion from the Broker. The Cloud Service Broker use different mechanisms to map the requirements specified by the CSRs to the services offered by CSPs. One of the mechanism is to use different filters to select the Cloud Service Provider like cost, availability, security, QoS etc. The Cloud Service Requester specifies the cost, level of availability, level of security and also the level of quality in the requirement specification form. The Broker apply these values to the filters and find out the suitable Cloud Service Provider among the registered Providers. Another mechanism is by applying a service mapping using mining algorithm. There are different approaches to select.
This paper is mainly concentrating on the evaluation of fair and unfair ratings by Cloud Service Broker (CSB) when the Cloud Service Requester rated the Cloud Service Provider (CSP). After identifying the unfair ratings CSB reduce the weightage of those ratings to reduce the effect of unfair ratings during the CSP reputation evaluation. This could be done by using Dynamic True Reputation Factor Value Evaluation Algorithm.

This paper consists of introduction in the first section, problem statement in the second section, survey in the third section, methodology is explained in the fourth section, experimental work in the fifth section, Results and Discussions along with two case studies in the sixth section, and concluded with implications of research in the seventh section.

2. PROBLEM STATEMENT

In the previous work [15], Reputation Factor Values (RFVs) are calculated based on the response form submitted by the Cloud Service Requester. But in the real environment the most of the Cloud Service Requesters are no longer interested to fill the lengthy response form. Some of the Cloud Service Requesters interested to fill response form when they were not satisfied by that Cloud Service Provider Service. To reduce the complexity of filling lengthy response forms, in this paper consider star ratings which are very popular in most of the e-commerce applications. In this work the complete Cloud Service Provider selection depends on the feedback given by the used CSR. If the Cloud Service Requester is not fairly rated the CSP then that impacts the total Broker system. So, it shows the necessity of evaluating the fair and unfair ratings given by the CSRs.

3. LITERATURE SURVEY

3.1 Survey On Multicloud Adoption

To improve the cloud availability most of the organizations migrating from single to multicloud environment. Many of the researchers worked on this to postulate the effectiveness of this migration. In [1] this paper, author specified a novel scheme to store data among multiple clouds without using any encryption scheme. It works on the basis of log range interpolation algorithm. When they compared this with existing methods, it has given good performance through availability, confidentiality and lock-in issues. In the paper [2], author raised an issue regarding the complexity in deploying multicloud applications and proposed CYCLONE architecture. This method merged number of existing solutions to provide multicloud application for research organizations. In [3], author proposed a model called multicloud searchable encryption which enables the individual providers to collaborate to search in their own data. The process is carried out at three levels: key generation, keyword identification, and obfuscation. In this the multicloud architecture searches over obfuscated keyword. In [4], author explained the way how this multicloud environment provides data privacy and availability. As per this, data is portioned into slices and each provider does not have all the data slices to see the content. This way of utilizing multicloud really improves the privacy of data. Along with privacy author introduced data duplication to improve the availability also. Many researchers [5, 6, 7], worked towards this at different applications.

3.2 Survey on Broker

In [8], author specified the significance of Broker, role of Broker, and the services provided by the Broker. In this work the Brokers are classified based on the service they provide towards the user and the provider. In [9], author specified the Broker services. As per this work, Broker reserves the services from different providers and lease their services to the users at cost benefit. So, this model motivates the organizations to take services from providers only through Broker to get at profitable cost. In [10] this work, Broker guides the users to get volume discount pricing offered by providers using cost efficient online scheduling algorithm.

3.3 Survey on Trust Model

Trust is an important factor which is needed in between the providers and users. The factors that impact the trust are scalability, availability, security, and usability. In [11] this paper, author proposed a fuzzy based trust model which evaluates the performance of providers. To describe this feature IaaS based e-learning system is considered. The main important steps in this model are defining the fuzzy sets to all possible factors and also for the output which is the trust value of the provider, fuzzification process, designing the rules based on the neural network, use of Sugeno’s method, and
defuzzification for generating a single output value. In [12] this paper, author proposed a method for CSR to identify trustworthy cloud providers. It deals with different kinds of threads against the system and methods to handle these threads. [13], proposed a method called RLM for reputation evaluation. This reputation value is mainly depending on the feedbacks. To identify the motivated feedbacks author used EM algorithm. [14], describes the threads that compromises the reliability of the trust system. In the previous work, RFV [15] is evaluated for every CSP and selected the CSP which has high RFV. But in RFV evaluation, the CSB is not evaluating the trust worthiness of the ratings. It effects the future business of CSP.

4. METHODOLOGY

Cloud Service Broker is evaluating the performance of Cloud Service Provider using Cloud Service Requesters response. If the response given by the CSR is not fair and motivated by any of the other sources then the entire Broker system will be collapsed and the future business of that Cloud Service Provider will be get affected. So, the Cloud Service Broker should maintain a mechanism to detect the unfair and fair raters. This mechanism is termed as Dynamic True Reputation Factor Value Evaluation (DTRFVE).

The Cloud Service Broker architecture consists of a local repository with list of registered Cloud Service Providers and list of Cloud Service Requesters. This repository is maintained by the Broker by assigning the local unique identity numbers to all the providers and requesters. The purpose of these IDs is to recognize the individual providers and requesters and also to hide their original identity which may interm overcome the Sybil attacks. The internal components of the Cloud Service Broker are discussed in [15].

4.1 Dynamic True Reputation Factor Value Evaluation

The Dynamic Trust Reputation Factor Value Evaluation algorithm takes list of Cloud Service Providers and their star ratings as an Input. All the ratings are normalized using min-max normalization [16] method. Normalized values are stored for further process in the dataset which is named as Norm_Data. Norm_Data is sorted in ascending order using an order ( ) function and stored as Order_Data. Find the first quartile [17] and third quartile of the Order_Data and store these values in Quartile_1, Quartile_3. These values are used to calculate the Inter Quartile Range (IQR).

\[ \text{IQR} = \text{Quartile}_3 - \text{Quartile}_1 \]

Find the upper and lower boundary values using IQR.
\[ \text{Lower \_Quartile} = \text{Quartile}_1 - 0.5*\text{IQR} \]
\[ \text{Upper \_Quartile} = \text{Quartile}_3 + 0.5*\text{IQR} \]

Now, every rating against each Cloud Service Provider in Order_Data is considered and compared their range.
If the rating falls below the Lower_Qartile or above the Upper_Qartile then consider it is as unfair and assign the very low weightage score to its rating.
If the rating falls above the Lower_Qartile and below the Quartile_1 then consider it is as fair and assign the average weightage score to its rating.
If the rating falls below the Upper_Qartile and above the Quartile_3 then consider it is as fair and assign the high weightage score to its rating.
If the rating falls in between the Quartile_1 and the Quartile_3 then consider it is as fair and assign the high weightage score to its rating.

After assigning weightage scores to their ratings find the mean of all these weighted ratings for every provider in the data set.

\[ \text{TRFV} = \frac{\sum W_s \times \text{Rating}_i}{n} \]

Calculate the True Reputation Factor Value of Cloud service Provider by using the above equation. In that Ws indicates the Weightage Score, Rating is the normalized rating given by the Cloud Service Requester, n is the number of Ratings. This process gives the True Reputation Factor Value of every Cloud Service Provider.

When the CSR gives the feedback immediately the CSB find out the Divergence Value of that rating.

\[ \text{Divergence} = \text{Normalized (New\_rating)} - \text{True RFV of CSP} \]

If the Divergence is less than a Threshold value, then that rating will be added to the Data set.

Divergence <= 0.5 then True Rating
Divergence > 0.5 then False Rating

Otherwise, CSB allots minimum rating to the CSP by specifying that as the CSR rating as unfair.
4.2 Algorithm

Import the data
Replace non-numeric cells with NaN
Create output variable
Create dataset array
Allocate imported array to column variable names
Clear temporary variables
Take size of data in to Row_Count and Coloumn_Count
Loop 1: Repeat 1 to Coloumn_Count
  Loop 2: Repeat 1 to Row_Count
  Norm_Data = Normalization (Data)
  Order_Data = Sort (Norm_Data )
  Quartile_1 = quantile(Order_Data,0.25)
  Quartile_3 = quantile(Order_Data,0.75)
  IQR = Quartile_3 - Quartile_1
  Lower_Quartile = Quartile_1 – 0.5*IQR
  Upper_Quartile = Quartile_3 + 0.5*IQR
  Unfair [] = Order_Data ((Order_Data < Lower_Quartile) | (Order_Data > Upper_Quartile))
  Weighted_Score = 0.2
  Unfair [] = Unfair []* Weighted_Score
  Fair[] = Order_Data (Order_Data < Quartile_1 & Order_Data > Lower_Quartile) & (Order_Data > Quartile_3 & Order_Data < Upper_Quartile)
  Weighted_Score = 0.5
  Fair [] = Fair []* Weighted_Score
  Fairest[] = Order_Data (Order_Data <= Quartile_1 & Order_Data >= Quartile_3)
  Weighted_Score = 0.8
  Fairest [] = Fairest []*Weighted_Score
End Loop 2
True RFV= Average of Unfair [ ], Fair [ ], and Fairest [ ]
End Loop 1

5. EXPERIMENTAL WORK

In this section, we investigate the effectiveness of our proposed method DTRFVE. Initially this is simulated using Java. We compare the performance of DTRFVE with RFV [15]. This investigation is done by applying our algorithm in the simulated cloud application. The following figure 1 shows the flow of work. It consists of the data collection, cleaning the data, Normalization of data and apply the Dynamic True Reputation Factor Evaluation Algorithm.

5.1 Data Collection

Dataset named Cloud Service Providers is used in experiment presented here. Table 1 shows the list of Cloud Service Providers considered in the Dataset. This data is collected from www.g2crowd.com. This dataset consists of ten Cloud Service Providers in the Market.

<table>
<thead>
<tr>
<th>IDs of the CSP in the Dataset</th>
<th>Name of the Cloud Service Provider</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>AmazonVPC</td>
</tr>
<tr>
<td>2</td>
<td>Digital Ocean</td>
</tr>
<tr>
<td>3</td>
<td>Amazon S3</td>
</tr>
<tr>
<td>4</td>
<td>Softlayer Virtual Servers</td>
</tr>
<tr>
<td>5</td>
<td>Linode</td>
</tr>
<tr>
<td>6</td>
<td>Cloud stack</td>
</tr>
<tr>
<td>7</td>
<td>Microsoft Azure</td>
</tr>
<tr>
<td>8</td>
<td>Compute Engine</td>
</tr>
<tr>
<td>9</td>
<td>Open Stack</td>
</tr>
<tr>
<td>10</td>
<td>Rackspace Managed Cloud</td>
</tr>
</tbody>
</table>

The initial star ratings of the Cloud Service Provider are shown in figure 2. In the dataset, the ratings of each provider varies based on the number of Cloud Service Requesters who rated the Providers after taking the service.
In figure 3 the X-axis shows the Ids of Cloud Service Providers and the Y-axis shows number of ratings for each of the Cloud Service Provider.

5.2 Cleaning

Cleaning is the process making the data ready for the task. In this dataset all the Cloud Service Providers are not having the same number of Cloud service Requesters. The first row consists of the Providers names. So, to make it ready for the task we apply replacement of non-numeric elements with NaN.

5.3 Normalization

The ratings in the Dataset are 1-5 star ratings. These are normalized in the range of 0-1. When we multiplied these range of ratings with the weighted score in some cases where rating is 0 the product will become 0. To avoid this situation, we are normalizing the ratings in between 0.1 - 0.9. Here we used min-max Normalization [16] method.

5.4 Sorting

The ratings which are normalized in the range of 0.1 - 0.9 are need to be sorted in ascending order to apply the Dynamic Reputation Factor Value Evaluation Algorithm.

5.5 Dynamic True Reputation Factor Value Evaluation

Apply the DTRFV algorithm on the data set to find the TRFV of ten Cloud Service Providers in the data set.

6. RESULTS & DISCUSSIONS

Cloud Service Broker received a request from Cloud Service Requester for IaaS. Broker search for the Providers who offer IaaS in its local repository. The result of this search has given ten providers who offer this service. Now Broker has to select one best Provider among ten. Broker applied the Dynamic True Reputation Factor Value Evaluation Algorithm on the Cloud Service Provider Data set which has those ten providers with ratings.
CSP Ids. Now the Broker decides that CSP ID 3 as a suitable provider to serve the present request.

The above figure shows the Reputation Factor Values of ten Cloud Service Providers. As per the previous model [15] the Broker selected the CSP ID 4 as a service provider to serve the request. In this method, equal priority has been given to each rating. The presence of unfair ratings effects the complete Broker model and leads to effect the business of the Cloud Service Providers. This is the main limitation of RFV.

As per the above observation, the number of ratings effects the True Reputation Factor Value. If the Number of ratings are more that Cloud Service Provider is getting maximum TRFV. Then it creates load imbalance problem. To solve it, we consider only the past ‘n’ transactions. This ‘n’ value is decided by the Cloud Service Broker. For example the CSB decided n=7, then past seven ratings of each provider is considered and apply the Dynamic True Reputation Factor Value Evaluation Algorithm. The accuracy of this algorithm is helpful to the Cloud Service Broker to improve Brokering Service among the Cloud Service Requester and Cloud Service Provider. The effect of number of ratings is shown in figure 6 when we consider Cloud Provider Dataset.

6.1 Case Study I

To understand the usage of Dynamic Reputation Factor Value Evaluation Algorithm we consider a simple situation where there are 10 normalized ratings for a Cloud Service Provider1 and 15 normalized ratings for a Cloud Service Provider2. Their ratings are as follows.

Rating [ ] = [0.8 0.6 0.5 0.5 0.4 0.4 0.3]
As per our algorithm the sorted array of Ratings Rating [ ] = [0.3 0.4 0.4 0.5 0.5 0.6 0.8]
First quartile range Q1 = 0.4
Third quartile range Q3= 0.6
IQR = 0.2
Lower Limit= 0.3
Upper Limit =0.7
Consider all ratings which falls under the range of above Upper Limit and below Lower Limit. Those are unfair ratings. In this case only one ratings falls under this category. Weighted Score is 0.2.

[313x526]0.8*0.2 = 0.16  (1)
Consider all ratings which falls under the range of Q3 and Upper Limit or Q1 and Lower Limit. Those are fair ratings. In this case only one ratings falls under this category. Weighted Score is 0.5.

[313x526]0.3*0.5 = 0.15  (2)
Consider all ratings which falls under the range of Q1 and Q3. Those are also fair ratings. In this case five ratings fall under this category. Weighted Score is 0.8.

[313x526]0.6*0.8 = 0.48  (3)
0.5*0.8 = 0.40  (4)
0.5*0.8 = 0.40  (5)
0.4*0.8 = 0.32  (6)
0.4*0.8 = 0.32  (7)

TRFV of the provider = (Eq1+ Eq2+ Eq3+ Eq4+ Eq5+ Eq6+ Eq7)/n
TRFV = (0.16+0.15+0.48+0.40+0.40+0.32+0.32)/7
= 0.31
Like this the Cloud Service Broker evaluates the True RFV of all Cloud Service Providers.
Whenever the Cloud Service Broker receives a request for a Cloud Service Provider from the Cloud Service Requester immediately it search in its local repository for its service match. If there are multiple providers who could offer the same required services then CSB choose the one which has maximum TRFV. That Cloud Service Provider will serve the request. After taking the service from the Provider the Cloud Service Requester rate the service by giving a Star rating. This is not directly included in the local repository of the CSB. The CSB finds the Divergence value of this rating by subtracting the normalized value of this rating from the TRFV of that CSP. If the Divergence is less than the 0.5 then consider the rating as valid and insert it in the repository. Whenever it is needed CSB finds the TRFV of all the required providers. For example the new normalized rating is 0.6. Then Divergence = 0.3 which is less than 0.5. So, this rating is added to the dataset because it is Valid.

6.2 Case Study II

Consider two Cloud Service Providers and assigned Identity Numbers are as CSP 1 and CSP 2. Ratings are as follows:

CSP 1: 5 3 2 1 3 5 2 1 3 3
CSP 2: 1 5 3 2 4 2 4 3 3 2 3 4 2 5 5

The CSP 1 has 11 Ratings whereas CSP 2 has 17 Ratings. These are normalized to 0.1-1. After applying Min- Max Normalization:

CSP 1: 1 0.5 0.3 0.1 0.5 1 1 0.3 0.1 0.5 0.5
CSP 2: 0.1 1 0.5 0.3 0.7 0.3 0.7 0.7 0.5 0.5 0.5 0.3 0.5 0.7 0.3 1 1

The sorted list of ratings are

CSP 1: 0.1 0.1 0.3 0.3 0.3 0.5 0.5 0.5 1 1 0.3 0.1 0.5 0.5
CSP 2: 0.1 0.3 0.3 0.3 0.3 0.5 0.5 0.5 0.5 0.5 0.5 0.7 0.7 0.7 1 1 1

The RFV values of CSP 1 and CSP 2 are

RFV (CSP1) = 0.549
RFV (CSP2) = 0.60

Apply Dynamic True Reputation Factor Value Evaluation Algorithm on CSP 1 and CSP 2

TRFV (CSP 1) = 0.416
TRFV (CSP 2) = 0.344

By considering RFV values the Broker would select the CSP 2 which has high RFV value. But when the Broker uses the TRFV of Cloud Service Providers, selects CSP 1.

Table 2 shows the RFV and TRFV of CSP 1 and CSP 2. In this table, initially we consider 11 ratings of CSP 1 and 17 ratings of CSP 2.

Table 2: Number Of Ratings For Each CSP

<table>
<thead>
<tr>
<th>CSP ID</th>
<th>No of Ratings</th>
<th>RFV</th>
<th>TRFV</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>11</td>
<td>0.549</td>
<td>0.416</td>
</tr>
<tr>
<td>2</td>
<td>17</td>
<td>0.6</td>
<td>0.344</td>
</tr>
<tr>
<td>1</td>
<td>5</td>
<td>0.48</td>
<td>0.25</td>
</tr>
<tr>
<td>2</td>
<td>5</td>
<td>0.52</td>
<td>0.33</td>
</tr>
</tbody>
</table>

The above figure shows the RFV and TRFV of CSP 1 and CSP 2. When we consider only first 5 ratings of CSP1 and 17 ratings of CSP2 then TRFV (CSP1) is 0.25 and TRFV (CSP2) is 0.344. Then Broker selects CSP2 which has high TRFV value. Figure 8 shows these values. Based on the above result we could say if the number of ratings are more that CSPs TRFV value would be high. To overcome this problem we consider equal number of past ‘n’ ratings from each provider.

Figure 7: RFV and TRFV of CSPs

The above figure shows the RFV and TRFV of CSP 1 and CSP2. When we consider only first 5 ratings of CSP1 and 17 ratings of CSP2 then TRFV (CSP1) is 0.25 and TRFV (CSP2) is 0.344. Then Broker selects CSP2 which has high TRFV value. Figure 8 shows these values. Based on the above result we could say if the number of ratings are more that CSPs TRFV value would be high. To overcome this problem we consider equal number of past ‘n’ ratings from each provider.
In this case consider n value as 5.

CSP 1: 5 3 2 1 3

CSP 2: 1 5 3 2 4

TRFV (CSP 1) = 0.25

TRFV (CSP 2) = 0.33

The Broker selects CSP 2. Figure 9 shows the RFV and TRFV of CSP 1, CSP 2 by considering past five ratings in both the providers.

We identified, however, that ratings given by the Cloud Service Requester indicates the degree of their satisfaction with the service they received which really implies to that service provider also. In future we plan to improve the Broker services by taking necessary actions on the Providers who got poor TRFV and any SLA violations.

7. CONCLUSIONS

This paper defines the effect of false reputation in the Cloud Service Broker system where the Cloud Service Requesters depend on the decision of Cloud Service Broker in the Multicloud environment. To identify the false ratings, we proposed an algorithm called Dynamic True Reputation Factor Value Evaluation. This algorithm helps to reduce the effect of unfair ratings by assigning low weightage to those ratings and encourage the fair ratings by giving high weightage in the TRFV evaluation process. We applied this algorithm on the Cloud Service Provider rating data set to show the efficiency of this algorithm. We compared these results with RFV and identified the effectiveness of DTRFV evaluation algorithm by reducing the effect of unfair ratings over RFV.

REFRENCES:


