

Evaluation of Parameters Importance in Cloud Service Selection Using Rough Sets

Yongwen Liu, Moez Esseghir, Leila Merghem Boulahia

Technology University of Troyes, Troyes, France

Email: yongwen.liu@utt.fr, moez.esseghir@utt.fr, leila.merghem_boulahia@utt.fr

Received 12 February 2016; accepted 27 March 2016; published 30 March 2016

Copyright © 2016 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

With the rapid development of the cloud computing technology, it has matured enough for a lot of individuals and organizations to move their work into the cloud. Correspondingly, a variety of cloud services are emerging. It is a key issue to assess the cloud services in order to help the cloud users select the most suitable cloud service and the cloud providers offer this service with the highest quality. The criteria parameters defining the cloud services are complex which lead to cloud service deviation. In this paper, we propose an assessment method of parameters importance in cloud services using rough set theory. The method can effectively compute the importance of cloud services parameters and sort them. On the one hand, the calculation can be used as the credible reference when users choose their appropriate cloud services. On the other hand, it can help cloud service providers to meet user requirements and enhance the user experience. The simulation results show the effectiveness of the method and its relevance in the cloud context.

Keywords

Cloud Computing, Cloud Service Selection, Rough Set Theory

1. Introduction

For several years, cloud computing has been influencing the IT landscape and becomes an important economic factor [1] due to its mode of operation that is the pay-as-you-go to provide service. Since cloud computing is a minimal barrier to entry and economic scaling, there are a lot of prospective clients to move their business on it. In this context, many small and large cloud service providers emerge every day. However, not all of them are the first-hand owners of a cloud infrastructure. This means that for those smaller cloud service providers, they are only partnered with a bigger provider which owns the infrastructure. Normally, this is not a big problem, even though they are all connected to a bigger infrastructure provider, when it goes down, all “agentmiddleman”

go down with it. Since cloud service providers have their specific service model, therefore, it is difficult for users to compare the cloud services offered by the different providers. Consequently, the cloud users face a challenge to select an appropriate provider taking into account their specific requirements.

Some cloud users take into consideration only their subjective preference parameters of the assessment criteria, while ignoring the importance of objective assessment parameters obtained from other customers who have the same service requirements when they are selecting the cloud services. Most cloud users cannot find an appropriate cloud service matching their individual requirements when they are using a given cloud service for the first. In fact, as they are not sure that the performance and quality of the selected service are good, they choose on the basis of their subjective judgment to the adapted decision parameters. Furthermore, when cloud users try to give an overall assessment for a cloud service, it is also not objective that the parameter weights of cloud service are generated usually subjective experience or experts scoring. This affects the cloud users' choice of a suitable cloud service.

For all the issues mentioned above, we can obtain the importance rating of attributes and rank them through the rough set theory; thereby we determine the objective weight of the assessment indexes of cloud services. Our proposal not only can guide cloud users, facing a lot of choices of cloud services, concerning assessment indexes they should focus more, but also helps cloud providers to improve the performance and quality of the cloud services with the emphasis to attract more cloud users to make themselves have a predominance in future competition of IT industry.

2. Related Works

With the development of cloud computing technology, the cloud service is becoming a mature concept concerning the delivery of software services, infrastructure services and platform services. Many techniques have been proposed by researchers from academia and industry for cloud services publication, interface definition and service discovery. Cloud service techniques (e.g., virtualization technique) have greatly accelerated the adoption and deployment of cloud services.

At the same time, more and more cloud service providers are offering all kinds of cloud services. For users, it is uneasy to make decision about the services meeting their requirements. To allow customers to evaluate cloud offerings and rank them based on their ability to meet the user's QoS (Quality of Service) requirements, Garg, S.K. *et al.* proposed a framework and a mechanism that measure the quality and prioritize Cloud services [2]. In this framework, the authors presented a rank cloud services mechanism using AHP (Analytic Hierarchy Process) [3] for solving problems related to MCMD (Multiple-criteria Decision-making). AHP is a widespread service ranking method. It is a structured technique for organizing the cloud service information and analyzing complex decisions. The analytic network process (ANP) [4] can provide a solution to problems that cannot be structured hierarchically, and is considered as an extension of AHP. An AHP-based SaaS services selection method is introduced in literature [5] to score and rank services. The researchers construct an AHP hierarchy to represent SaaS service attributes. Although the use of AHP can improve the objective rating based on selection attributes, however, the importance of the service attributes is judged by aggregating user preferences and the opinions of experts, so the result of services ranking is more subjective. On the basis of AHP hierarchy, N. Boussoualim [6] proposed an approach to calculate the weights of the various attributes of choice parameters and score the different products in a SaaS selection to help users to make decision. Since weights of various factors are assigned according to the user preferences, therefore, this method is also limited by the subjective judgment. Karim *et al.* [7] defined an AHP hierarchy of a cloud service weighting model, in which a mechanism (a set of rules to perform the mapping process) is explored to map the users' QoS requirements of cloud services to the right QoS specifications of SaaS. Nie G. h. *et al.* [8] proposed a cloud service evaluation index system to guide users in the choice of cloud services. These works have some common features, such as the proposed models are based on AHP, the initial importance of the parameters based on subjective judgment and so on.

Unlike AHP, other approaches for cloud service selection are proposed. Han S.M. *et al.* [9] presented a cloud service selection framework in the cloud market to help users select the better services. This cloud service recommendation system is based on a utility function to quantify the preferences of a decision maker. In [10], authors described a framework for reputation-aware software service selection and rating. It aims to rate SaaS services while reducing the time and risk of the selection and utilization of software services. The proposed selection mechanism aids service users to select services based on quality, cost and reputation. Saripalli *et al.* [11]

discussed Multiple Attribute Decision Methodology to rank alternatives in a decision problem in cloud service adoption. In this work, the authors analyzed the possible decision problems the service users might encounter. The Simple Additive Weighting (SAW) method is used to rank the service candidates based on the rating values generated.

In mentioned above works, the researchers proposed various ranking approaches for cloud services selection. To rank the cloud services, it is necessary to evaluate the importance of the parameters given in cloud services selection. As the weight for each parameter acquired by conducting experts' opinions or user preferences in above works, as a result, certain recommended cloud services are not always the best to meet users' requirements.

Different from research emphasis of the above works, our study focuses on the parameters importance evaluation to guide users in cloud services selection. To get a rational evaluation result for each cloud service parameter, we use the rough set theory to carry out our work. In [12], the author proposed an approach for mining significant factors affecting the adoption of SaaS using the rough set theory. Although we are using the same theory in a similar context, our work makes a further study. The method we proposed not only can explore the significant factors but also can rank and weight these parameters in cloud services selection.

3. Evaluation Parameters of Cloud Service

With the rapid development of cloud computing, more and more cloud service providers join cloud market. Businesses and consumers have more choices because a large number of industry application solutions emerge. The global market scale for cloud services is increasing. Cloud computing providers carry on the business on a unified platform by building cloud resource pool for resource sharing, resource centralization, service network, billing and demand elasticity, to achieve cloud business structure on a scale. From a marketing perspective, the main types of cloud services are cloud hosting services, object storage services, cloud database services, cloud engine services, block storage services, cloud caching services, on-line application services, load balancing services and cloud distribution services. From another perspective, cloud services include IaaS, PaaS and SaaS. Moreover, cloud can be divided into public cloud, private cloud and hybrid cloud (on deployment).

The core business is various from different cloud service providers. For example, Amazon's business is more interested in the platform and software (PaaS and SaaS), which are public cloud services. However, IBM has a wider range of business, and its hardware and platforms are more advanced; IaaS, PaaS, SaaS and other aspects of the business are involved, it is favored in building private and hybrid clouds. Therefore, it is difficult for the user to define what cloud service providers are the best on the basis of a certain point. There are some configuration parameters for every type of cloud services to assess their performance. For example, the number of CPU, the size of memory, the space of storage, operate system and so on, these parameters determine the performance of cloud hosting services. When users are choosing one type cloud service, there are many alternative cloud service providers. When users make choices they need some parameters to evaluate cloud service providers' comprehensive ability, such as the capacity for innovation, the service capability, product technologies, the solutions, brand influence. Usual evaluation parameters of cloud service and cloud services providers as follows.

- **Cloud service availability:** availability is the proportion of time a system in functioning condition. For cloud service availability, it can be defined as the capacity of an IT system to provide continuous service delivery. We give an example to understand what exactly it means. Let's take a 99.9% SLA (service-level agreement), in practice, this means that in any given month (assuming a 30-day month), the service can only be unavailable for about 4 minutes and a few seconds, or only about 50 minutes per year. It includes connectivity, reliability, delay, data leakage and the tenant's business does not meet expectations or entirely suspended caused by any accident on IaaS, PaaS and SaaS. As cloud services mature, cloud service availability becomes as important as price or other factors in choosing the right service provider.
- **Cloud service scalability:** scalability is a broad concept. It appears in a wide range of applications. For cloud service, simply speak, scalability is the ability of the whole system to sustain increasing workloads by making use of additional resources. It is about how to deal with the large-scale business and attract more users. It is not directly related to how well the actual resource demands are matched by the provisioned resources at any point in time, even if there is more than a single point of failure. However, scalability of cloud service composition needs to meet the requirement for expanding users and technology upgrade.
- **Cloud service elasticity:** elasticity has become a key metric of cloud service. The term elasticity is used in the

naming of specific cloud products or service. It is an ability of a system to adapt to change in workloads and resource demands. Users expect to obtain the best service with the cheapest way. As we all know, cloud services provide multi-service contracts depending on the different hierarchical levels of users' needs. This dynamic proposition allows the users selecting the suitable options according to their needs and the amount of the resource they used. Therefore, users use the service quite flexibly with defined rights at any moment to save money. Usually, the term elasticity is one of the keywords for promoting the development of cloud service [13].

- Cloud service security: Cloud service concerns a number of security issues [14], such as software platform security and infrastructure security via the cloud. Cloud service providers must ensure their clients' data and applications are protected, while users can through authentication enhance their application security. Cloud service providers often store many users' data on the same server to save costs, conserve resources and maintain efficiency. As a result, there is a chance that user's private data can be viewed by other users without taking effective measures. Moreover, the precautionary measures to prevent Internet from hacking and virus damage. Therefore, cloud service security is an important index when evaluating the quality of the service.
- Capacity of innovation: innovation is described in terms of changes in what a company offers the product or service upgrade and the ways it creates and delivers those offerings (process improvement) [15]. Innovation is the soul of enterprise progress, the core of economic competition. An enterprise's ability to innovate is a key to its success. When most competitors within an industry have acquired the same level of competence in areas of management, such as marketing operations, human resources and strategy, they need to look for some innovations, such as incentive, resource investment and enterprise's self-fulfillment as a key factor for significant competitive advantages
- Total Cost of Ownership: Total Cost of Ownership (TCO) is an analysis technology to uncover all the lifetime costs that follow from owning certain kinds of assets. TCO provides a cost basis for determining the total economic value of an investment when incorporated in any financial benefit analysis [16]. TCO analysis attempts to uncover both the obvious costs and the "hidden" costs of ownership. Obvious costs in TCO are the costs involved during planning and vendor selection, such as purchase cost and the actual price paid. "Hidden" costs include acquisition costs, upgrade costs, security costs and so on. TCO is a scientific, rational economic evaluation index for firms.
- Service capability: Service capability is the degree of capability in a service system to provide services and is commonly defined as the maximum output rate of the system. Compared with the manufacturing industry, service capability of IT enterprises stresses the technology and skills to meet the needs of customers with high quality serving products [17] [18]. Enhancing of service capability can improve competitive advantages.
- Solution: For some problems (such as deficiency, demands, shortage) that already occurred or can be predicted in an enterprise, solution is a specific plan or proposal that can be effectively implemented. An excellent solution offers a series of conclusion: Why it happens? Whether it occurs again or not? Does it lead to other problems? How to avoid related problems? What experiences are accumulated from the solution? [19] As well as in some fields, solution should meet customers' demands to achieve the expected effects.
- Brand influence: Brand influence refers to the ability of opening up market and gaining the benefits with the brand [20] [21]. It has been an important element for customers to choose their cloud service providers.

4. Rough Set Theory

Rough set theory proposed by Pawlark in [22] is a mathematical approach to uncertain knowledge. Rough set theory has been applied in many interesting areas. The rough set approach is of fundamental importance to artificial intelligence and cognitive sciences, especially in the fields of machine learning, knowledge acquisition, knowledge discovery, decision analysis, expert systems, inductive reasoning and pattern recognition [23]. The main advantage of rough set theory in the process of knowledge analysis is based on dataset rather than subjective judgment.

Definition 1 [22] [24] [25]. Let $T = (U, A, V, f)$ be an information system, where $U = \{X_1, X_2, \dots, X_n\}$ is the finite set of objects; $A = C \cup D$ is the set of attributes, C is a conditional attributes set, D is the decision attribute set; $V = \bigcup_{\alpha} V_{\alpha}$ where V_{α} is the set of values of attributes $\alpha \in A$. f is an information function and denotes the map of $U \times A \rightarrow V$, which means a value to each attribute for each object.

Definition 2 [22] [24] [25]. Given an information system $T = (U, A, V, f)$, $A = C \cup D$. The expression $Pos_C(D)$, called a positive region of the partition U/D with respect to condition attributes C , is the set of all elements of U that can be uniquely classified to blocks of the partition U/D , by means of C . U/D indicates elementary concepts of information system T about decision attribute set D . For $\alpha \in C$, we have:

- a) If $Pos_{C-\{\alpha\}}(D) = Pos_C(D)$, then α is an unnecessary attribute of C ;
- b) If $Pos_{C-\{\alpha\}}(D) \neq Pos_C(D)$, then α is a necessary attribute of C .

Definition 3 [22] [24] [25]. Given an information system $T = (U, A, V, f)$, $A = C \cup D$. Attribute importance of the decision information system can be tested by the classification ability for T when removing an attribute $\alpha \in C$ from condition attribute set C , the significance of the attribute α is defined by [22] as in:

$$Sig(\alpha) = \frac{|\text{card}(Pos_C(D))| - |\text{card}(Pos_{C-\{\alpha\}}(D))|}{|U|} \quad (1)$$

Card presents the set cardinality of the attributes. Sig_α represents the dependence of decision attribute D relative to condition attribute α , and which reflects the classification discrimination ability of the attribute α . The larger value of Sig_α , the more stronger of dependency relationships between condition attribute α and decision attribute D , and the more discriminative the attribute α is.

5. The Cloud Service Selection Method with Preference Information

Cloud users usually give the subjective weight to different parameters of the cloud service based on personal preference when they are choosing the cloud service, thus resulting into non practical choices. Therefore, in this section we introduce an approach to rank the importance of the cloud service indexes and provide the objective weight about different parameters based on the rough set theory.

5.1. The Objective Ranking of Attributes Approach Based on Rough Set Theory

Rough set theory analysis is based on upper and lower approximations space. The lower approximation of the set can describe the precise knowledge in an information system, which is called positive region and is defined by Definition 2. If the lower approximation will not be changed when an attribute is deleted, then the attribute is unnecessary and can be reduced. Otherwise, the attribute is called core attribute, which is necessary. In other words, the Definition 2 can distinguish the core attributes and unnecessary attributes while ignoring the effect of the relatively necessary attributes. For all relatively necessary attributes, we can rank them in an information system according to the significance values of different attributes. The significance of an attribute defined by Definition 3 can reflect the variety of the lower approximation space when the attribute is deleted.

Since cloud service is characterized by various parameters, such as availability or scalability, elasticity and so on, it is difficult to define selection criteria valid for different customer needs. For this problem, we give a cloud service selection method using rough set theory to help user make decision.

We get the users' subjective preferences information through interacting with users. If some users provide incomplete information, we can adopt data complete mode translating the incomplete information into complete one. The method of getting user's preferences information is shown in **Figure 1**.

First, we obtain the preference values of parameters of cloud services. Then, we compute the preference weight of various parameters. The user preference levels are shown in **Table 1**. To facilitate computations and storage in the database, we assign the preference levels with numerical values. * means that users do not provide personal preferences, which are null.

We construct an information system based on a large preference datasets collected from users of certain cloud service providers (google, Alibaba et al). **Table 2** is an assessment and requirement system of users about the cloud services. U represents the cloud services set, $U = \{s_1, s_2, \dots, s_m\}$; Condition attributes set represents the assessment parameters of cloud services, $C = \{\text{availability, scalability, reliability, credit, \dots, loads}\}$, that is $C = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$; decision attribute set is satisfied with the cloud service or not, $D = \{\text{Yes, No}\}$, that is, $\{1, 0\}$, where, * represents incomplete information.

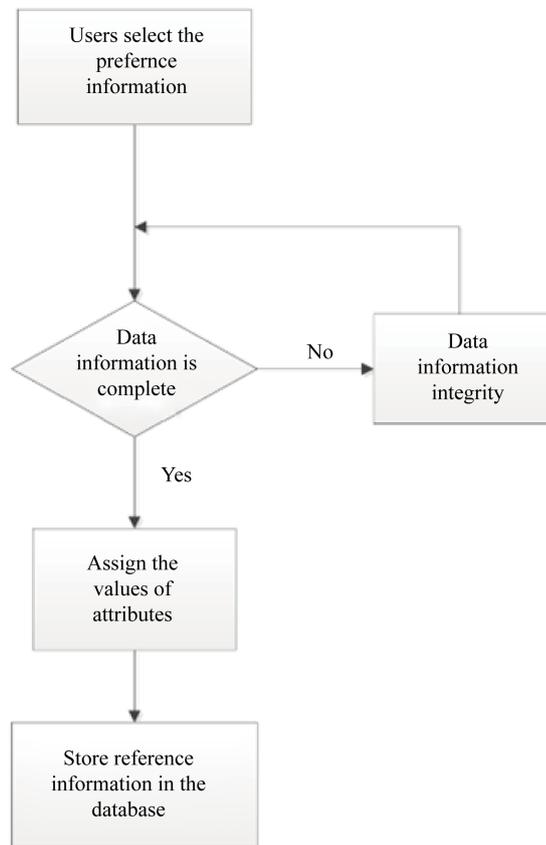


Figure 1. Getting the preference information.

Table 1. The preference levels of users.

Very important	Important	Not important	No selection
2	1	0	*

Table 2. User preferences and assessment for cloud service.

	α_1	α_2	α_3	α_4	...	α_n	d
S_1	1	*	2	1	...	0	0
S_2	0	0	1	2	...	1	0
S_3	1	2	1	0	...	*	1
S_4	2	1	0	*	...	1	1
S_5	1	2	0	0	...	0	0
S_6	0	0	2	1	...	0	1
S_7	1	*	1	1	...	1	1
...	0
S_m	*	1	0	2	...	1	1

To obtain the parameters importance of cloud service, the ranking of attributes algorithm as follows:

Algorithm 1: The ranking attributes of cloud services

Input:
The information system of cloud services;

Output:
The attributes ranking of the cloud service;

1: Input the information table of cloud services;

2: Set $C = \{\alpha_1, \alpha_2, \dots, \alpha_n\}$

3: Compute all partition U/D with respect to condition attributes C ;

4: Set $i = 1$;

5: if $i \leq \text{number of the attributes}$;
 then
 Compute all partition U/D with respect to condition attributes $c = \{C - \alpha_i\}$;

$i++$;

6: Compute all the significance of the condition attributes with respect to decision attribute D

$$Sig(\alpha) = \frac{|\text{card}(Pos_c(D))| - |\text{card}(Pos_{c-\alpha_i}(D))|}{|U|}$$

7: Rank the attributes of cloud services.

5.2. Application of the Objective Ranking of Attributes Approach in Cloud Service Selection

Choosing the cloud services is a multiple attributes decision making problem, and the key is to determine the weight of parameters. There are several ways to determine the weight of indicators, on general, which fall into two categories: subjective and objective assignment methods. The subjective assignment method is assigning weight based on subjective information of decision-making. It is arbitrary with poor accuracy and reliability of decision-making. In the objective assignment method, each parameter is evaluated with the actual data. In cloud service selection system, the importance of attributes is different. The objective weight of attributes can be defined as in (2):

$$W_\alpha = \frac{Sig_\alpha(\alpha)}{\sum_{c \in C} Sig_c(c)} \tag{2}$$

The comprehensive weight with regard to parameters can be defined as in (3):

$$I(w) = \beta W_o(w) + (1 - \beta) W_{so}(w), 0 \leq \beta \leq 1 \tag{3}$$

where, β which is called weight coefficient reflects cloud user preference for subjective and objective weights of parameter when they make decisions in cloud services selection. $W_o(w)$ and $W_{so}(w)$ respectively represents the weight of parameters of cloud services with objective dataset and subjective dataset. Smaller value of β indicates that users value more their subjective preference. Conversely, higher value of β users emphasizes the objective importance of parameters. Specially, if $\beta = 0$, users judging the parameters' importance of cloud services totally depend on their subjective awareness; if $\beta = 1$, users completely rely on the objective weight.

An application is illustrated in determining the comprehensive weights of cloud service parameters based on the rough set theory. Obtaining the comprehensive weight of each parameter includes two parts. The first part is acquiring the weight of the parameters based on the subjective data which comes from the cloud user preferences. The second part is acquiring the objective weight based on the data without subjective information of decision-maker. The application model of the objective ranking of attributes in cloud service selection system is shown in **Figure 2**.

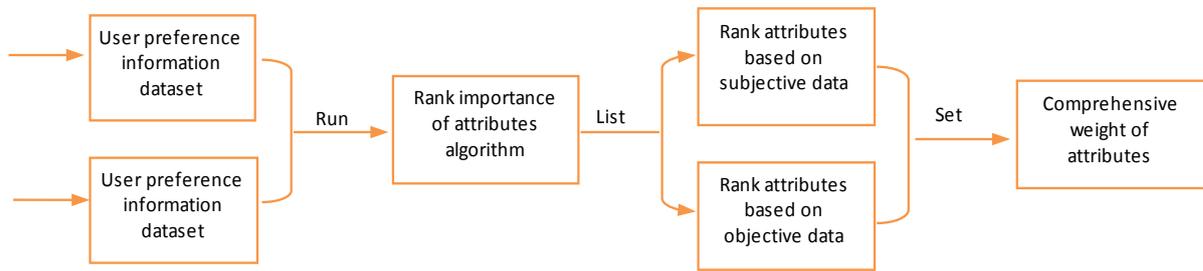


Figure 2. Application model of the objective ranking of attributes.

5.3. Application of Attributes Ranking Approach in Cloud Service Selection

There are corresponding indexes designed to evaluate a system or a service. When cloud service providers launch a service product to consumers, they should provide quality of services and they hope to get the feedback from consumers early to improve their products, at the same time, the evaluation indexes of the services to be design accordingly. For cloud service users, when they choose a cloud service, they will consider some factors to obtain the suitable service, such as cloud service availability, cloud service elasticity, brand of service etc. As we know, in economical market, the cost control and the pursuit of efficiency are the primary goals of each company management. The reason cloud users choose moving their business to cloud computing center is because this is a good way to save capital and improve efficiency compare to their traditional development model. However, in practice, cloud users should balance the weight of factors used to evaluate cloud service.

Here we demonstrate an instance to use rough set theory to rank the factors of cloud service providers because the overall strength of cloud service provider is important for cloud users to choose the suitable cloud service. The real data in Table 3 is the list of cloud service providers according to their all-round capacity in 2014. The cloud service providers operate in China. The data is published in the journal of China Internet Weekly [26]. In Table 3, the factors *CI* (capacity for innovation), *SC* (service capability), *PT* (product technologies), *S* (solution), *TCO* (total cost of ownership) and *BI* (Brand influence) are the evaluation factors of cloud service providers. The factor *CS* (comprehensive score) is the assessment result of the cloud service providers.

In rough set theory, every cloud service provider is represented as a research object, and the factors as its attributes. Among them, the factor *CS* is decision attribute, while others are condition attributes. Simply, columns of Table 3 are attributes and rows are objects, whereas entries of the table are attribute values. Thus, each row of the table can be seen as information about specific cloud service provider. Our research purpose is to rank the weight of the factors to assess the comprehensive strength of cloud service providers.

We abstract randomly a cloud service provider from Table 3 to explain what it is the purpose we study, for example, *Amazon*. We can see from Table 3 that cloud service provider is characterized by the following attribute-value set $(CI,9),(SC,9),(PT,9),(S,9),(TCO,5),(BI,9) \rightarrow (CS,8.8)$, which form the information about the cloud service provider.

In order to decide the weight of factors of cloud service providers to assess their comprehensive strength, we can get the attributes rank and weight values of Table 3 by the ranking of attributes algorithm we proposed which are shown in Table 4. It shows that the factor *S* is very important than other factors when the given parameters are used for evaluating cloud service providers. The weights of the factor *TCO* and *BI* are the smallest ones. They are not the key factors. According to the result of ranking factors, we able to reduce flexibly the evaluation factors.

5.4. An Example of Application of the Objective Ranking of Attributes Approach in Cloud Service Selection

We give an example to explain how to apply our model with personal preference. Table 5 and Table 6 are two information systems respectively based on the user preference dataset and the third-party objective dataset. To distinguish cloud service elements of subjective dataset and objective dataset, we use $s_j(j = 1, 2, \dots, 9)$ and $e_k(k = 1, 2, \dots, 20)$ to represent respectively the cloud service elements in Table 5 and Table 6. Attribute $\alpha_i(i = 1, 2, 3, 4)$ represents various parameters of cloud services. The value of attribute d is used to show the different decision results per cloud service. They are shown as follows.

Table 3. The scores of cloud service providers.

Rank	Manufacture	CS	CI	SC	PT	S	TCO	BI
1	IBM	8.9	10	9	9	9	4	10
2	Amazon	8.8	9	9	9	9	5	9
3	HP	8.7	10	8	9	9	6	9
4	Cisco	8.7	9	9	8.5	9	4.5	9
5	Saleforce	8.7	9	9	9	8.5	5	9.5
6	Dell	8.6	8.5	9.8	8.5	8.5	8.5	8.5
7	Huawei	8.6	9	8	8.5	9	8	9
8	Oracle	8.5	9	8.5	8.5	9	7	8
9	Microsoft	8.5	8	8.5	8.5	9	5	9
10	Google	8.5	8	10	8	9	8	7
11	Intel	8.4	8.5	8.5	8.5	9	7	8
12	EMC	8.3	9	8.5	9	9	5	8.5
13	SAP	8.2	8	8.5	8.5	8.5	7.5	8.5
14	H3C	8.2	8	8.5	9	8.5	5	8.5
15	ZTE	8.2	8	8.5	8.5	8	5	8.5
16	Alibaba	8.1	8	8.5	8.5	8	5	8
17	Fujitsu	8.0	8	8.5	8	8	5	8
18	Neusoft	8.0	8	8	8.5	8	5	8
19	Packspace	7.8	8	7	8	8.5	7	7
20	Teradata	7.8	8	8	7.5	8	7	6
21	NEC	7.6	8	7.5	8	7.5	5	8
22	Tencent	7.6	7	8	8	7.5	6	7.5
23	Citrix	7.6	7	8	7.5	7.5	7	8
24	Lenovo	7.6	8	8.5	7.5	7	4.5	9
25	Joyent	7.3	9	8	8	6	6	8
26	Inspur	7.2	7.5	7	7.5	7.5	4	8
27	NetApp	7.2	7	8	7	7	7	6
28	Vmware	7.2	7	8	7	7	7	6
29	Akamai	7.2	7	8	6	7	8	8
30	Sugon	7.1	6	8	7	7	7.5	6
31	JNPR	7.1	8	7	7.5	7	4	7.5
32	Xtools	7.1	7	7.5	7	7	6	6.5
33	SNDA	7.1	7	7	8	7	4	7
34	Jingdong	7.1	7	7	7.5	7	6	7
35	Infor	6.9	7	7.5	7	6.5	6	7
36	Symantec	6.9	7	8	7.5	6	4	7.5
37	FastTrek	6.9	7	7.5	7	6.5	5	7
38	ChinaTelecom	6.9	7	7	7.5	6.5	5	7.5
39	800APP	6.8	7.5	7	7	6.5	4	7.5
40	DigitalChina	6.8	7	7.5	7.5	6	4	7.5
41	Netsuite	6.7	7.5	7	6	7	4	7.5
42	UFIDA	6.6	7	5	7	7.5	6	7
43	PowerLeader	6.6	6.5	6	6.5	7	7	7
44	Juniper	6.6	7	7	6.5	7	7	6
45	Ruijie	6.6	6	7	6.5	6.5	7	6
46	Kingdee	6.6	6.5	7	7.5	6	4	7.5
47	Vianet	6.6	7	7	6.5	6	7	7.5
48	Ucloud	6.6	7	7	7	6	4	8
49	PedHat	6.5	7	7	6	6	7	7.5
50	Unicom	6.4	6	7	7	6	4.5	7

Table 4. The ranking and weight of attributes.

Ranking	Weight					
	<i>CI</i>	<i>SC</i>	<i>PT</i>	<i>S</i>	<i>TCO</i>	<i>BI</i>
$S \succ SC \succ PT \succ CI \succ TCO = BI$	0.1	0.25	0.2	0.35	0.05	0.05

Table 5. Users' preference information dataset.

	α_1	α_2	α_3	α_4	d
s_1	2	0	1	1	0
s_2	0	1	1	1	1
s_3	1	0	1	1	1
s_4	1	2	0	0	1
s_5	2	1	1	1	0
s_6	2	2	0	1	1
s_7	1	0	0	1	1
s_8	0	1	1	0	0
s_9	0	1	0	1	1

Table 6. Third party objective dataset.

	α_1	α_2	α_3	α_4	d
e_1	0	1	1	1	1
e_2	2	0	0	1	1
e_3	0	1	1	2	0
e_4	1	1	1	0	1
e_5	1	0	1	0	0
e_6	1	1	0	0	1
e_7	1	1	1	2	0
e_8	2	1	0	2	1
e_9	0	1	0	1	1
e_{10}	2	1	0	0	1
e_{11}	2	2	0	1	1
e_{12}	0	1	1	1	1
e_{13}	0	2	0	1	0
e_{14}	1	0	1	0	0
e_{15}	0	1	0	1	1
e_{16}	1	1	0	1	0
e_{17}	0	0	2	1	1
e_{18}	2	1	0	1	0
e_{19}	0	1	2	2	1
e_{20}	0	2	0	0	1

We can get the attributes rank, significance and weight values of **Table 5** and **Table 6** by Definition 2, 3 and Equation (2), or we get the result integrating **Algorithm 1** and Equation (2). The results are shown in **Table 7**. According to Equation (3), we can obtain the attributes ranking of cloud services with different values of weight coefficient β shown in **Table 8**.

6. Experiments Result and Analysis

The experiment has two goals. The first one aims for sorting the parameters of cloud services according to their significance to guide the new cloud service users to make decision. The second one aims to prove the method is effective in the application of the cloud services selection with preference information. Due to lack of the related standard test platform of users' preference and the standard test datasets, here we adopt data sets (download from the UCI [27]) as the training samples to carry out. Besides that, the original datasets are pre-processed to be easily used for calculating and program designing.

Table 9 shows the basic information of the data sets. Programming code is by Java language. It is executed sequentially on a processor Intel Core2 Duo CPUs x64. The main function of the algorithm is to give the importance order of the attributes. We can get the comprehensive weights of attributes according to the result of ranking and significance of attributes. We can get the ranking attributes by setting the different values of weight coefficient β . Thus we compare to the services matching rate successfully. The experiment regards the objective datasets as the benchmark for analysis to draw graphic. Services matching is used to describe the intention of the selection of cloud users for cloud services providers. We can get the result shown in **Figure 3**.

Table 7. The ranking, significance and weight of attributes.

	Significance				Ranking	Weight			
	α_1	α_2	α_3	α_4		α_1	α_2	α_3	α_4
Dataset in Table 5	0.444	0	0	0.222	$\alpha_1 \succ \alpha_4 \succ \alpha_2 = \alpha_3$	0.67	0	0	0.33
Dataset in Table 6	0.3	0.45	0.1	0.6	$\alpha_4 \succ \alpha_2 \succ \alpha_1 \succ \alpha_3$	0.2069	0.3103	0.0689	0.4138

Table 8. Ranking for attributes selection.

The value of weight coefficient	Ranking
Subjective dataset	
$\beta = 0$	$\alpha_1 \succ \alpha_4 \succ \alpha_2 = \alpha_3$
Objective dataset	
$\beta = 1$	$\alpha_4 \succ \alpha_2 \succ \alpha_1 \succ \alpha_3$
Comprehensive dataset	
$\beta = 0.1$	$\alpha_1 \succ \alpha_4 \succ \alpha_2 \succ \alpha_3$
$\beta = 0.3$	$\alpha_1 \succ \alpha_4 \succ \alpha_2 \succ \alpha_3$
$\beta = 0.5$	$\alpha_1 \succ \alpha_4 \succ \alpha_2 \succ \alpha_3$
$\beta = 0.7$	$\alpha_4 \succ \alpha_1 \succ \alpha_2 \succ \alpha_3$
$\beta = 0.9$	$\alpha_4 \succ \alpha_2 \succ \alpha_1 \succ \alpha_3$

Table 9. Basic information of test datasets.

Datasets	1	2	3	4	5
Number of Attributes	5	5	7	5	7
Number of Objects	24	150	287	625	1727

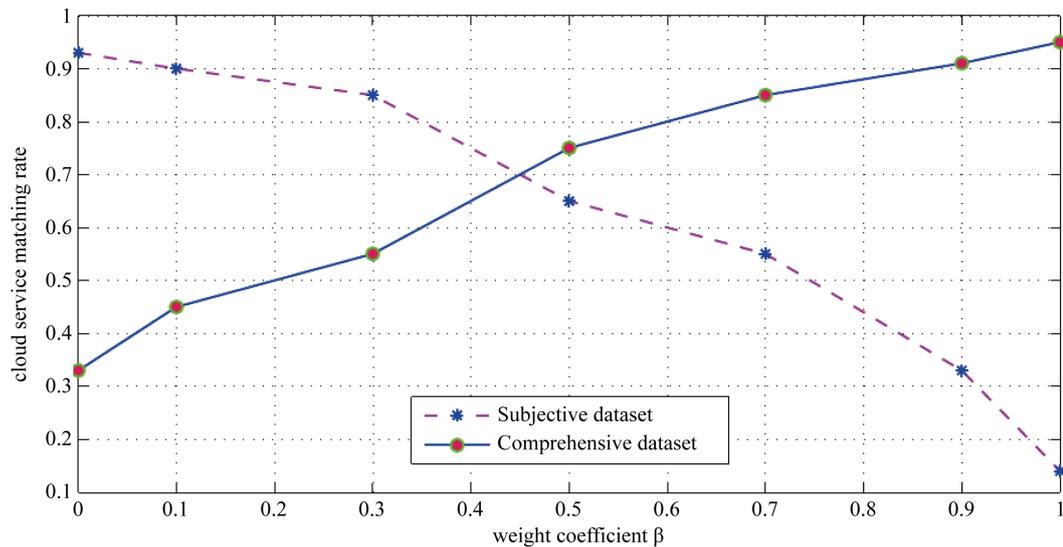


Figure 3. Cloud services match-making with various value of β .

It can be seen from **Figure 3** that with weight coefficient β greater, users' subjective preference becomes important, and the service match-making rate decreases; rather, combining the subjective data and objective data, the cloud service match-making rate increases.

The users with the different subjective preference of the attribute weight use the random data to get the subjective service matching rate. As mentioned above, we use the rough set methods to get the objective weight of the attribute, integrating the objective and subjective weight to get the comprehensive matching rate of the service. Here, we set weight coefficient β is 0.1, 0.3, 0.5, 0.7 and 0.9 separately. The results are shown in **Figure 4**.

We can see in **Figure 4**, when the datasets have less service objects, the comprehensive selection or subjective selection has high service matching rate successfully. With the data increases, the comprehensive weight matching rate increases, whereas the cloud service match-making rate decreases based on the subjective preference information.

In [12], the author proposed an analytical framework to explore the significant factors affecting the adoption of SaaS for enterprise users using rough set theory. The main contribution is to mine the important factors. Although our work is similar it in context, but our study goes to one step further, mining the significant factors in assessing cloud service providers (shown in **Table 3**), for example. There are six factors (CI , SC , PT , S , TCO , BI) in the information system of cloud service provider. It can mine four factors (CI , SC , PT , S) which are the important influence factors for evaluating the cloud service providers using the approach in [12]. Beyond that, we can't get the additional information about the result. However, in our study, we not only can know which factor is the important evaluation index of cloud service provider assessment but also rank them according to their weight, as the result shown in **Table 6**. Further, we can define a threshold to select evaluation factors at a stretch based on the result to design the evaluation system. In **Table 6**, we suppose that, for some reason, we need to reduce the number of evaluation factors from 6 to 4. The method in [12] and ours both are effective. That is, the factors TCO and BI would be removed because their influence is smaller than others for evaluating cloud service providers. And if, we need to reduce the number of evaluation factors from 6 to 3, first, we remove the two factors (TCO , BI), after that, we don't know which factor would be removed among the other four factors (CI , SC , PT , S , TCO , BI) based on the approach in [12], because there is no more information to guide us to do further. Therefore, the method proposed in [12] is failed in this case. However, in our work, beside removing the two factors (TCO , BI), we can judge easily to remove the factor (CI), because its weight is lower than the other factors', or according to the rank of factors importance shown in **Table 6**.

7. Conclusion

To provide a guide choosing the appropriate cloud services for cloud users, we present the rank-making of the parameters importance in cloud services selection and propose a attribute ranking method based on the rough

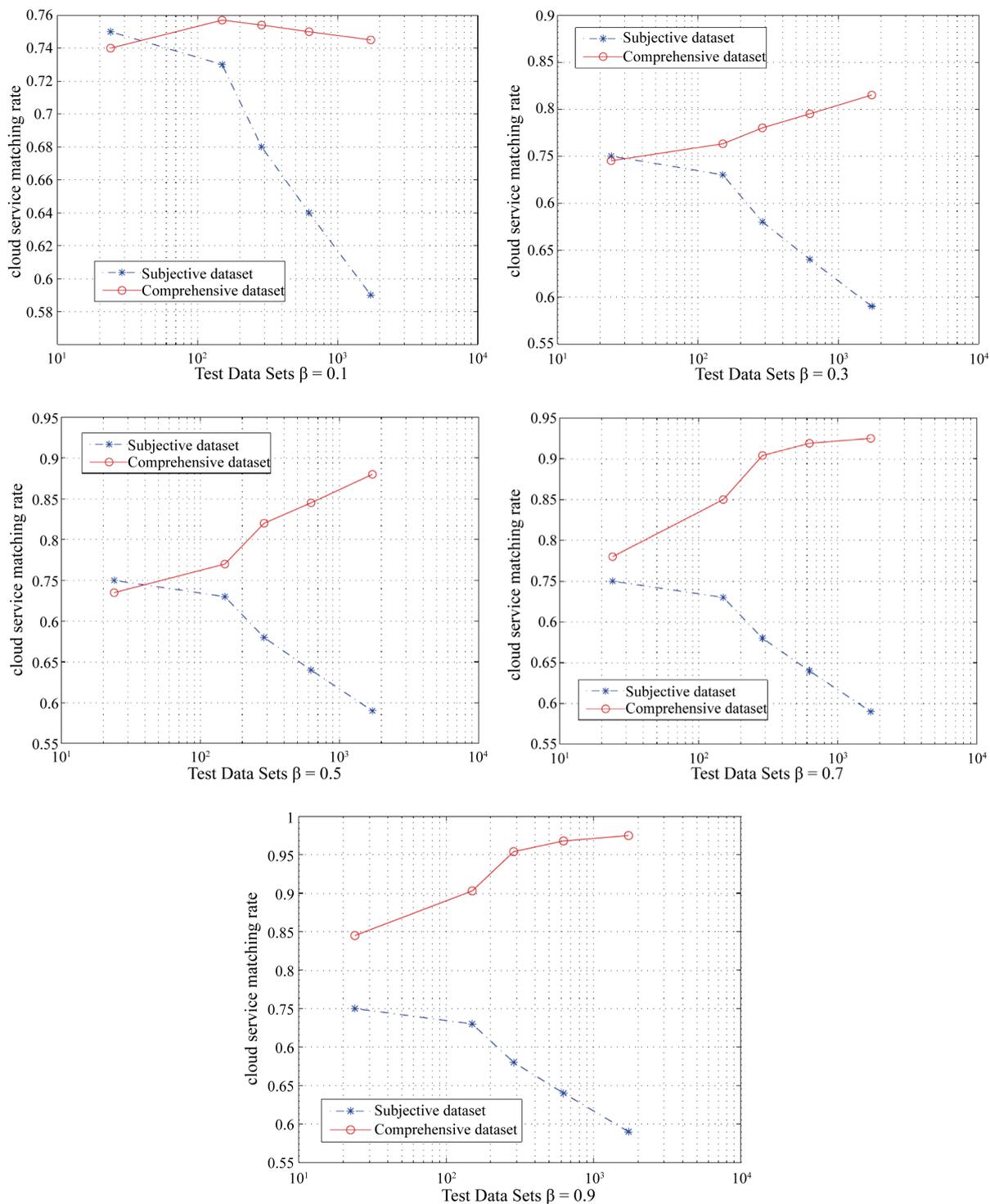


Figure 4. Cloud service match-making with varies datasets.

set theory. It can explore the significant factors affecting the adoption of cloud services for users. At the same time, it can help the cloud service providers to specifically improve their quality of services to win more customers. We use rough set theory in the design of the algorithm to rank the parameters of cloud services. Then we can get the different weights of attributes of cloud services from subjective dataset and objective dataset. Our experimental results show that our approach is effective in services matching. Our future work will focus on op-

timizing the cloud services selection with more complex preferences.

References

- [1] Repschlager, J., Wind, S., Zarnekow, R. and Turowski, K. (2011) Developing a Cloud Provider Selection Model. *Enterprise Modeling and Information Systems Architecture*, 163-175.
- [2] Garg, S.K., Versteeg, S. and Buyya, R. (2011) Smicloud: A Framework for Comparing and Ranking Cloud Services. 2011 4th IEEE International Conference on Utility and Cloud Computing (UCC), Melbourne, December 2011, 210-218. <http://dx.doi.org/10.1109/ucc.2011.36>
- [3] Buyukyazlcl, M. and Sucu, M. (2003) The Analytic Hierarchy and Analytic Network Processes. *CRITERION*, **1**, C1.
- [4] Saaty, T.L. (1990) How to Make a Decision: The Analytic Hierarchy Process. *European Journal of Operational Research*, **48**, 9-26. [http://dx.doi.org/10.1016/0377-2217\(90\)90057-1](http://dx.doi.org/10.1016/0377-2217(90)90057-1)
- [5] Godse, M. and Mulik, S. (2009) An Approach for Selecting Software-as-a-Service (SaaS) Product., *CLOUD'09 IEEE International Conference on Cloud Computing*, Bangalore, 21-25 September 2009, 155-158. <http://dx.doi.org/10.1109/CLOUD.2009.74>
- [6] Boussoalim, N. and Aklouf, Y. (2014) An Approach Based on User Preferences for Selecting SaaS Product. 2014 International Conference on Multimedia Computing and Systems (ICMCS), Marrakech, 14-16 April 2014, 1182-1188. <http://dx.doi.org/10.1109/icmcs.2014.6911278>
- [7] Karim, R., Chen, D. and Miri, A. (2013) An End-to-End Qos Mapping Approach for Cloud Service Selection. *Proceedings of the IEEE 9th World Congress on Services (SERVICES)*, Santa Clara Marriott, 27 June-2 July 2013, 341-348. <http://dx.doi.org/10.1109/services.2013.71>
- [8] Nie, G.H., She, Q.P. and Chen, D.L. (2011) Evaluation Index System of Cloud Service and the Purchase Decision-Making Process Based on AHP. *Proceedings of the 2011 International Conference on Informatics, Cybernetics, and Computer Engineering (ICCE2011)*, Melbourne, 19-20 November 2011, 345-352.
- [9] Han, S.M., Hassan, M.M., Yoon, C.W. and Huh, E.N. (2009) Efficient Service Recommendation System for Cloud Computing Market. *Proceedings of the 2nd International Conference on Interaction Sciences: Information Technology, Culture and Human*, Seoul, 24-26 November 2009, 839-845. <http://dx.doi.org/10.1145/1655925.1656078>
- [10] Limam, N. and Boutaba, R. (2010) Assessing Software Service Quality and Trustworthiness at Selection Time. *IEEE Transactions on Software Engineering*, **36**, 559-574. <http://dx.doi.org/10.1109/TSE.2010.2>
- [11] Saripalli, P. and Pingali, G. (2011) MADMAC: Multiple Attribute Decision Methodology for Adoption of Clouds. *IEEE International Conference on Cloud Computing (CLOUD)*, Washington DC, 4-9 July 2011, 316-323. <http://dx.doi.org/10.1109/CLOUD.2011.61>
- [12] Wu, W.-W. (2010) Mining Significant Factors Affecting the Adoption of SaaS Using the Rough Set Approach. *The Journal of Systems and Software*, **84**, 435-441. <http://dx.doi.org/10.1016/j.jss.2010.11.890>
- [13] Copil, G., Trihinas, D., Truong, H.-L., et al. (2014) ADVISE—A Framework for Evaluating Cloud Service Elasticity Behavior. In: Franch, X., Ghose, A.K., Lewis, G.A., Bhiri, S., Eds., *Service-Oriented Computing*, Springer Berlin Heidelberg, 275-290. http://dx.doi.org/10.1007/978-3-662-45391-9_19
- [14] Balduzzi, M., Zaddach, J., Balzarotti, D., Kirda, E. and Loureiro, S. (2012) A Security Analysis of Amazon's Elastic Compute Cloud Service. *Proceedings of the 27th Annual ACM Symposium on Applied Computing*, Riva del Garda, 26-30 March 2012, 1427-1434. <http://dx.doi.org/10.1145/2245276.2232005>
- [15] Cropley, D.H., Cropley, A.J., Chiera, B.A. and Kaufman, J.C. (2013) Diagnosing Organizational Innovation: Measuring the Capacity for Innovation. *Creativity Research Journal*, **25**, 388-396. <http://dx.doi.org/10.1080/10400419.2013.843330>
- [16] Martens, B., Walterbusch, M. and Teuteberg, F. (2012) Costing of Cloud Computing Services: A Total Cost of Ownership Approach. 45th Hawaii International Conference on System Science, Maui, 4-7 January 2012, 1563-1572.
- [17] Kulvatunyou, B., Lee, Y., Ivezic, N., et al. (2015) A Framework to Canonicalize Manufacturing Service Capability Models. *Computers and Industrial Engineering*, **83**, 39-60. <http://dx.doi.org/10.1016/j.cie.2015.01.027>
- [18] Carroll, N., Helfert, M. and Lynn, T. (2014) Towards the Development of a Cloud Service Capability Assessment Framework. In: Mahmood, Z., Ed., *Continued Rise of the Cloud*, Springer, London, 289-336. http://dx.doi.org/10.1007/978-1-4471-6452-4_12
- [19] Christensen, C. and Raynor, M. (2013) *The Innovator's Solution: Creating and Sustaining Successful Growth*. Harvard Business Review Press, Boston.
- [20] Lipsman, A., Mudd, G., Rich, M. and Bruich, S. (2012) The Power of "Like": How Brands Reach (and Influence) Fans through Social-Media Marketing. *Journal of Advertising Research*, **52**, 40-52.

<http://dx.doi.org/10.2501/jar-52-1-040-052>

- [21] Pham, M.T., Geuens, M. and De Pelsmacker, P. (2013) The Influence of Ad-Evoked Feelings on Brand Evaluations: Empirical Generalizations from Consumer Responses to More than 1000 TV Commercials. *International Journal of Research in Marketing*, **30**, 383-394. <http://dx.doi.org/10.1016/j.ijresmar.2013.04.004>
- [22] Pawlak, Z. (1982) Rough Sets. *International Journal of Computer & Information Sciences*, **11**, 341-356. <http://dx.doi.org/10.1007/BF01001956>
- [23] Liu, Y.W., Esseghir, M. and Boulahia, L.M. (2014) Cloud Service Selection Based on Rough Set Theory. *International Conference and Workshop on the Network of the Future (NOF)*, Paris, 3-5 December 2014, 1-6.
- [24] Skowron, A., Komorowski, J., Pawlak, Z. and Polkowski, L. (2002) Rough Sets Perspective on Data and Knowledge. In: Klösgen, W. and Zytkow, J.M., Eds., *Handbook of Data Mining and Knowledge Discovery*, Oxford University Press, New York, 134-149.
- [25] Rissino, S. and Lambert-Torres, G. (2009) Rough Set Theory—Fundamental Concepts, Principals, Data Extraction and Applications. In: Ponce, J. and Karahoca, A., Eds., *Data Mining and Knowledge Discovery in Real Life Applications*, I-Tech, 438-462.
- [26] Zhao, Y.X. (2014) 2014 Cloud Service Provider Charts. *China Internet Weekly*, **24**, 62-63.
- [27] UCI Machine Learning Repository: Data Sets. <https://archive.ics.uci.edu/ml/datasets.html>