International Journal of u- and e- Service, Science and Technology
Vol.8, No.2 (2015), pp.185-196
http://dx.doi.org/10.14257/ijunesst.2015.8.2.18

Collaborative Optimization of Profit and Cost of Hybrid Cloud Storage Service

Chao-ping SUN [1,2], Jing ZHAO [1] and Su-mei Ruan [2, 3]

[1.] *School of Management, Hefei University of Technology, Hefei 230009, PR. China*
[2.] *Key Laboratory of Process Optimization and Intelligent Decision-making, Ministry of Education, Hefei 230009, PR. China*
[3.] *School of Business, Anhui University of Finance and Economics, Bengbu 233041, PR. China*
Corresponding author's mail: tedsuncp@163.com

Abstract

Aiming at hybrid cloud storage service (HCSS), this paper conducts research on collaborative optimization of corresponding profits and costs. Profit function of HCSS considers flexibility, extendibility, decrease in investment, improvement on access speed, data sharing, disaster recovery, and green IT, while cost function of HCSS considers data transmission, data storage, and facilities of hardware and software. Based on profit and cost functions of HCSS, a goal programming model is constructed to collaboratively optimize profits and costs of HCSS. The model is validated by making a comparative analysis in two representative contexts.

Keywords: *HCSS, cloud service profit, cloud service cost, collaborative optimization*

1. Introduction

Promoted by IT giants like Google, Amazon, IBM *etc.*, [1], cloud computing, integrating computing resources, storage resources with software resources and forming a huge virtual pool of resources via virtualization technology, is likely to become basic necessities like water, electricity, gas, and telephone in the future [2], providing the required information service for users [3].

Cloud storage service, as one of the typical cloud computing applications, provides enterprises with an economical, feasible, flexible and convenient storage mode [4]. Cloud storage service falls into public cloud storage service（PuCSS）, private cloud storage service（PrCSS）, and HCSS in terms of technology. PuCSS is characterized by being highly efficient, reliable, and low cost, provided by a third party; PrCSS is characterized by being highly safe, controllable, and reliable, provided by enterprises themselves; HCSS strikes a balance between enterprise's business and storage cost, having the characteristics of both PuCSS and PrCSS.

As to the application to different types of enterprises, HCSS can not only meet enterprise's dynamic, elastic, and changeable storage demands but also eliminate their concerns about the security of confidential data, thus promoting the popularization of cloud storage service [5]. Cost and safety are two major considerations in the configuration of public and private cloud, and next comes profit.

This paper intends to study HCSS from the aspect of profits and costs. The following seven factors like flexibility, extendibility, decrease in investment, improvement on access speed, data sharing, disaster recovery, and green IT are considered in the service

ISSN: 2005-4246 IJUNESST

profit function, while the following three factors like data transmission, data storage, and facilities of hardware and software are considered in the cost function. The goal programming model is constructed based on the profit function and cost function to make a collaborative optimization of profits and costs in HCSS to achieve their best allocation.

With the development of information technology, the college informationization has become an important indicator to measure the level of scientific research. However, there also exist problems with heavy investment in hardware facilities, low utilization of resources, and low information sharing [6]. HCSS can pose opportunities to college informationization via the virtual storage technology, thus reducing data storage and management cost, realizing the dynamic extension of storage capacity, improving information sharing and resources utilization.

The information technology can also find applications in the modern China's manufacturing industry [7]. The modern manufacturing industry aims at developing the service manufacturing, posing an urgent demand for the integrated manufacturing service platform construction for group enterprises and enterprise alliances. HCSS provides a feasible scheme for the construction of integrated manufacturing service platform. In other words, HCSS can provide enterprise groups and enterprise alliances with a stable storage service platform of high performance, thus improving the storage efficiency and resources utilization.

Therefore, this paper elaborates the application and effectiveness of the collaborative optimization of profit and cost of HCSS and makes a contrasting analysis by taking a certain university research institute and an automobile manufacturing enterprise as an example.

2. Constructing Profit Function and Cost Function of HCSS

2.1. Constructing Profit Function of HCSS

2.1.1. Influencing Factors of HCSS Profit

HCSS profits are mainly the long-term returns and the cost reduction compared with the traditional storage [8]. These returns, far exceeding the immediate cost reduction, are gained by the long-term application of HCSS. Specifically, these returns can be gained from the following 7 factors.

(1) Flexibility, is the maximum returns that CSS can bring to enterprises by preparing the computing resources rapidly, shortening the preparation time, enabling the business departments to launch new products rapidly to be competitive in the market.

(2) Extendibility, means that CSS can provide enterprise with storage resources and related services as required. HCSS can provide dynamic storage service for enterprises according to their different demands, bringing about implicit profits.

(3) Decrease in investment

HCSS users can reduce their maintenance cost of hardware facilities considerably, improve their resources utilization, and lower their energy consumptions, thereby reducing their investment.

(4) Improvement on access speed

HCSS can reduce single machine's access load effectively through the load balancing strategy and reading-writing isolated strategy according to vast data in the enterprise informationization, thereby improving data access speed in storage services and contributing to enterprises' long-term development.

(5) Data sharing

Realizing data sharing at a maximum level is an important goal of enterprise informatization. HCSS can increase enterprise's profits through the establishment of highly efficient, reliable data sharing mechanism, updating it in a real-time and dynamic way, improving the data efficiency, reducing the management cost of enterprise data.

(6) Disaster recovery

HCSS can ensure the integrity and consistency of enterprises' data through multiple backup mechanisms, improving the security of enterprises' data, bringing about implicit profits to enterprises.

(7) Green IT

HCSS based on green IT [9] can reduce energy consumption, reduce the IT cost, and improve IT resources utilization, thereby bringing about long-term profits to sustain enterprises' development.

2.1.2. HCSS Profit Function

Considering the long-term profits brought about by HCSS, we construct the service profit function considering the above 7 factors.

As to PuCSS, we regard uploading data size (GB), downloading data size (GB), data storage size (GB) as dependent variables, denoted as x_1, x_2, x_3, satisfying $x_1 > 0$, $x_2 > 0$, $x_3 >$. As to private cloud, we regard unit price of purchasing facilities as the dependent variable, denoted as x_j ($j = 4, ..., J$), st. $x_j > 0$, where $J-3$ denotes the number of types of purchase facilities.

Taking full account of PuCSS and PrCSS, We construct profit function of HCSS based on the above 7 factors, denoted as $f(x_1, \ldots, x_J)$. Suppose variables $x_1, \ldots, x_J$ are independent of each other, profit of each factor can be represented as the multi-function of variables $x_1, \ldots, x_J$, denoted as $u_i = \sum_{j=1}^{J} \lambda_{ij} g_{ij}(x_j)$ ($i = 1, \ldots, 7$), satisfying $\sum_{j=1}^{J} \lambda_{ij} = 1$, where $g_{ij}(x_j)$ denotes the profit of factor i on the variable x_j; λ_I represents the relative weight of $g_{ij}(x_j)$. Further, assuming that the above seven factors are independent of each other, We can obtain the total profit function $f(x_1, \ldots, x_J) = \sum_{i=1}^{7} \beta_i u_i$ based on the constructed u_i ($i = 1, \ldots, 7$), satisfying $\sum_{i=1}^{7} \beta_i = 1$, where β_i ($i = 1, \ldots, 7$)denotes the relative weight of u_i. It is apparent that determining $g_{ij}(x_j)$, λ_{ij}, and β_i is fundamental to construct $f(x_1, \ldots, x_J)$. Especially, determining $g_{ij}(x_j)$ is of particular importance.

Properties of function $g_{ij}(x_j)$ are summarized by interviewing some network engineers experienced in network implementation, as shown below.

Property 1 Suppose x_j ($j = 1, ..., J$) denotes uploading data size, downloading data size, data storage size, and unit price of purchase facilities $J-3$, profit function $g_{ij}(x_j)$ of factor i ($i = 1, \ldots, 7$) satisfies:

(1) being continuously differentiable at the interval $(0, +\infty)$;

(2) the first-order derivative being greater than 0 at the interval $(0, +\infty)$;

(3) the first-order derivative being monotone decreasing at the interval $(0, +\infty)$;

(4) being not convergent at the interval $(0, +\infty)$.

For different applications, the value of x_j ($j = 1, ..., J$) is different. To ensure $g_{ij}(x_j)$ ($j = 1, ..., J$) can find applications on different occasions, condition (1) in property 1 must be satisfied.

When the value *of* x_j ($j = 1, ..., J$) increases, profits keep increasing, so condition (2) of Property 1 needs to be satisfied.

When the value of x_j ($j = 1, ..., J$) increases, the generated marginal profit follows the law of diminishing returns. Therefore, condition (3) in Property 1 needs to be satisfied, namely the second-order derivative of $g_{ij}(x_j)$ is continuous, and less than 0.

To ensure the applicability of $g_{ij}(x_j)$ (j = 1, ..., J) on different occasions, the value of $g_{ij}(x_j)$ should be differentiated in the given range of x_j (j = 1, ..., J). Therefore, condition (4) in Property 1 needs satisfying.

In general, functions satisfying Property 1 can be used to build $g_{ij}(x_j)$. We set the constructed u_i (i = 1, …,7) the synthesis of $g_{ij}(x_j)$ in the uniform dimension as the normalized function considering different values of x_j (j = 1, ..., J) in the practical application. The power function $(x_j)^a$ $(0 < a < 1)$ and logarithmic function $\log_a(x_j)$ $(a > 1)$ are two typical elementary functions satisfying Property 1 and normalized $g_{ij}(x_j)$ can be constructed according to the given range of x_j (j = 1, ..., J). Obviously, it is key to determine the parameter a so as to determine a normalized $g_{ij}(x_j)$ by taking account of the practical application. Given the range $[x_j^-, x_j^+]$ of x_j (j = 1, ..., J), corresponding parameter a_{ij} of different factors i (i = 1, …, 7) should be different, reflecting different influences of variable x_j (j = 1, ..., J) on different factors. The higher the influencing degree is, the smaller value of a_{ij} is, and vice versa.

Though the value of logarithmic function $\log_a(x_j)$ $(a > 1)$ is related to that of parameter a, the normalized $g_{ij}(x_j)$ based on its construction is independent of the value of parameter a.

Proposition1. Let $g_{ij}(x_j)=\dfrac{\log_{a_{ij}}(x_j)-\log_{a_{ij}}(x_j^-)}{\log_{a_{ij}}(x_j^+)-\log_{a_{ij}}(x_j^-)}$, $x_j \in (x_j^-, x_j^+)$, then $\forall\, i_1, i_2 \in \{1,...,7\}$, $g_{i_1 j}(x_j)=g_{i_2 j}(x_j)$.

Proof According to $g_{ij}(x_j)=\dfrac{\log_{a_{ij}}(x_j)-\log_{a_{ij}}(x_j^-)}{\log_{a_{ij}}(x_j^+)-\log_{a_{ij}}(x_j^-)}$, then

$$\frac{g_{i_1 j}(x_j)}{g_{i_2 j}(x_j)}=\frac{\log_{a_{i_1 j}}(x_j)-\log_{a_{i_1 j}}(x_j^-)}{\log_{a_{i_1 j}}(x_j^+)-\log_{a_{i_1 j}}(x_j^-)} * \frac{\log_{a_{i_2 j}}(x_j^+)-\log_{a_{i_2 j}}(x_j^-)}{\log_{a_{i_2 j}}(x_j)-\log_{a_{i_2 j}}(x_j^-)}$$

$$=\frac{\log_{a_{i_1 j}}(x_j/x_j^-)/\log_{a_{i_1 j}}(x_j^+/x_j^-)}{\log_{a_{i_2 j}}(x_j/x_j^-)/\log_{a_{i_2 j}}(x_j^+/x_j^-)}$$

$$=\frac{\log_{a_{i_1 j}}(x_j/x_j^-)}{\log_{a_{i_2 j}}(x_j/x_j^-)}\cdot\frac{\log_{a_{i_2 j}}(x_j^+/x_j^-)}{\log_{a_{i_1 j}}(x_j^+/x_j^-)}$$

$$=\log_{a_{i_1 j}}(a_{i_2 j})\cdot\log_{a_{i_2 j}}(a_{i_1 j})=1.$$

So, $g_{i_1 j}(x_j)=g_{i_2 j}(x_j)$.

According to Proposition 1, the normalized $g_{ij}(x_j)$ based on of the construction of the logarithmic function $\log_a(x_j)$ $(a > 1)$ makes no difference to each factor. Therefore, it is difficult to reflect the different influence of the variable x_j (j = 1, ..., J) on each factor by selecting different a_{ij}.

Considering using the power function $(x_j)^a$ $(0 < a < 1)$ to construct regularized $g_{ij}(x_j)$.

Proposition 2. If $g_{ij}(x_j)=\dfrac{(x_j)^{a_{ij}}-(x_j^-)^{a_{ij}}}{(x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}}}$, $x_j \in (x_j^-, x_j^+)$, $g_{ij}(x_j)$ is a monotone decreasing function of aij.

Proof To resolve the first-order partial derivatives of $g_{ij}(x_j)$ by taking a_{ij} as the variable, we can get

$$\frac{\partial g_{ij}(x_j)}{\partial a_{ij}}=\frac{((x_j)^{a_{ij}}\ln x_j-(x_j^-)^{a_{ij}}\ln x_j^-)((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})}{((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})^2}$$

$$-\frac{((x_j^+)^{a_{ij}}\ln x_j^+-(x_j^-)^{a_{ij}}\ln x_j^-)((x_j)^{a_{ij}}-(x_j^-)^{a_{ij}})}{((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})^2}.$$

Let

$h(x_j)=((x_j)^{a_{ij}}\ln x_j-(x_j^-)^{a_{ij}}\ln x_j^-)((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})$

$-((x_j^+)^{a_{ij}}\ln x_j^+-(x_j^-)^{a_{ij}}\ln x_j^-)((x_j)^{a_{ij}}-(x_j^-)^{a_{ij}})$.

Apparently, $h(x_j)$ is continuous at the interval $[x_j^-,\ x_j^+]$ and differentiable at the interval $(x_j^-,\ x_j^+)$, satisfying $h(x_j^-)=h(x_j^+)=0$, according to the Rolle theorem, there exists at least one point $\xi\in(x_j^-,\ x_j^+)$, satisfying $h'(\xi)=0$ξ. Again

$h'(x_j)=(a_{ij}(x_j)^{a_{ij}-1}\ln x_j+(x_j)^{a_{ij}-1})((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})$

$-a_{ij}(x_j)^{a_{ij}-1}((x_j^+)^{a_{ij}}\ln x_j^+-(x_j^-)^{a_{ij}}\ln x_j^-)$

$=(x_j)^{a_{ij}-1}((a_{ij}\ln x_j+1)((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})-a_{ij}((x_j^+)^{a_{ij}}\ln x_j^+-(x_j^-)^{a_{ij}}\ln x_j^-))$,

It can be inferred from $h'(\xi)=0$ that

$$(a_{ij}\ln\xi+1)((x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}})-a_{ij}((x_j^+)^{a_{ij}}\ln x_j^+-(x_j^-)^{a_{ij}}\ln x_j^-)=0.$$

Because ln (XJ) is an increasing function, it can be inferred that when $x_j\in(x_j^-,\xi)$, $h'(x_j)<0$, when $x_j\in(\xi,\ x_j^+)$, $h'(x_j)>0$.

So, when $x_j\in(x_j^-,\ x_j^+)$, $h(x_j)<0$, namely $\frac{\partial g_{ij}(x_j)}{\partial a_{ij}}<0$, showing that $g_{ij}(x_j)$ is a monotone decreasing function of aij.

When $x_j=x_j^-$, $g_{ij}(x_j)\equiv0$, when $x_j=x_j^+$, $g_{ij}(x_j)\equiv1$, having nothing to do with the value of aij.

According to Proposition 2, it is known that when $x_j\in(x_j^-,\ x_j^+)$, the value of $g_{ij}(x_j)$ $=\frac{(x_j)^{a_{ij}}-(x_j^-)^{a_{ij}}}{(x_j^+)^{a_{ij}}-(x_j^-)^{a_{ij}}}$ is associated with a_{ij}. Therefore, it is a reasonable normalized payoff function.

Given the influencing weight of x_j (j = 1, ..., J) to different factors, namely θ_{ij} (i = 1, …, 7), satisfying $\sum_{i=1}^{7}\theta_{ij}=1$.

Let $\theta_{i^*j}=max\{\theta_{ij}, i=1,\dots,7\}$. Without loss of generality, we define a^* ($0<a^*<1$) as the ideal minimum in the practical application. Assign $a_{i^*j}=a^*$ (j = 2, ..., J), and $a_{ij}=a^*\cdot\theta_{i^*j}/\theta_{ij}$ $(i\neq i^*)$.

Profit function $g_{ij}(x_j)$ (i = 1, …, 7, j = 1, ..., J) can be determined.

To avoid the negative influence of subjective cognition deviation on each factor profit and overall service profit, we determine λ_{ij} (i = 1, …, 7, j = 1, ..., J) and β_i (i = 1, …, 7) according to $g_{ij}(x_j)$by following fair and optimistic principle. Namely, $\lambda_{ij}=g_{ij}(x_j)/\sum_{j=1}^{J}g_{ij}(x_j)$, j = 1, ..., J, $\beta_i=u_i/\sum_{i=1}^{7}u_i$, i = 1, …, 7.

Where λ_{ij} (j = 1, ..., J) represents profit weights of public and private cloud service variables on each factor.

Having got λ_{ij} (i = 1, …, 7, j = 1, ..., J) and β_i (i = 1, …, 7), we can obtain standardized profit function of mixing cloud storage services $f(x_1, \ldots, x_J)$, and the profit ranges from 0 to 1.

2.2. Constructing Cost function of HCSS

The cost of HCSS consists of public cloud service cost and private cloud service cost.

2.2.1. PuCSS Cost

PuCSS consists of data uploading service and data downloading service [10]. service cost (c_1) includes data transmission service (c_{11}), data storage service (c_{12}) [11], namely $c_1=c_{11}+c_{12}$.

Data transmission cost refers to the cost of migrating data to public cloud storage platform, determined by uploading data size, downloading data size, unit price of uploading data (p_1), and unit price of downloading data (p_2), so $c_{11}=p_1 \cdot x_1+p_2 \cdot x_2$.

Data storage costs are determined by storage time (t), data storage size and unit data storage cost within the unit time (p_3), so $c_{12}=p_3 \cdot x_3 \cdot t$.

To sum up, $c_1=c_{11}+c_{12}=p_1 \cdot x_1+p_2 \cdot x_2+p_3 \cdot x_3 \cdot t$.

2.2.2. PrCSS Cost

Enterprise's deployment of PrCSS platform mainly consists of infrastructure construction and virtual construction. Accordingly, PrCSS cost includes infrastructure cost and virtualization cost. This paper unifies the above two costs into the soft and hard infrastructure cost.

Suppose deploying PrCSS needs purchasing J-3 facilities, and unit price of each facility is fixed at x_j; the number is n_j(j = 4, ..., J), then the PrCSS cost is

$c_2=\sum_{j=4}^{J} n_j \cdot x_j$.

Integrating PuCSS cost with PrCSS cost, total service cost is $c=c_1+c_2=p_1*x_1+p_2*x_2+p_3*x_3*t+\sum_{j=4}^{J} n_j \cdot x_j$. Like HCSS profit function, we standardize the cost function, namely $\bar{c}=(c-c^-)/(c^+-c^-)$.

Where $c^- = p_1 \cdot x_1^- + p_2 \cdot x_2^- + p_3 \cdot x_3^- \cdot t + \sum_{j=4}^{J} n_j \cdot x_j^-$,

$c^+ = p_1 \cdot x_1^+ + p_2 \cdot x_2^+ + p_3 \cdot x_3^+ \cdot t + \sum_{j=4}^{J} n_j \cdot x_j^+$.

When $\bar{c}$ is 0, the cost is the minimum; when $\bar{c}$ is 1, the cost is the maximum.

3. Collaborative Optimization Model of Profits and Costs of HCSS

Goal programming model is constructed based on the above profit and cost functions of HCSS to collaboratively optimize profits and costs of HCSS, achieving their best balance of profit and cost, realizing the best configuration of PuCSS with PrCSS. Specific models are as follows:

$$\text{MIN } \delta_1+\delta_2 \quad (1)$$

$$\text{s.t.} \quad \sum_{i=1}^{7} \beta_i u_i + \delta_1 = 1, \quad (2)$$

$$\bar{c} - \delta_2 = 0, \quad (3)$$

$$\beta_i = u_i / \sum_{i=1}^{7} u_i, \quad i = 1, \ldots, 7, \quad (4)$$

$$u_i = \sum_{j=1}^{J} \lambda_{ij} g_{ij}(x_j), \quad i = 1, \ldots, 7, \quad (5)$$

$$\lambda_{ij}= g_{ij}(x_j) / \sum_{j=1}^{J} g_{ij}(x_j)\ ,\ \ j = 1, ..., J, \tag{6}$$

$$g_{ij}(x_j)= \frac{(x_j)^{a_{ij}} - (x_j^-)^{a_{ij}}}{(x_j^+)^{a_{ij}} - (x_j^-)^{a_{ij}}}\ ,\ \ i = 1, \ldots, 7,\ \ j = 1, ..., J, \tag{7}$$

$$a_{ij}= a^* \cdot \theta_{i^* j} / \theta_{ij}\ ,\ \ i = 1, \ldots, 7,\ \ j = 1, ..., J, \tag{8}$$

$$\overline{c} = (c - c^-) \big/ (c^+ - c^-)\ , \tag{9}$$

$$c=p_1 * x_1+p_2 * x_2+p_3 * x_3 * t+ \sum_{j=4}^{J} n_j \cdot x_j \tag{10}$$

$$c^- = p_1 \cdot x_1^- + p_2 \cdot x_2^- + p_3 \cdot x_3^- \cdot t + \sum_{j=4}^{J} n_j \cdot x_j^- \tag{11}$$

$$c^+ = p_1 \cdot x_1^+ + p_2 \cdot x_2^+ + p_3 \cdot x_3^+ \cdot t + \sum_{j=4}^{J} n_j \cdot x_j^+ \tag{12}$$

$$x_j^- \le x_j \le x_j^+\ ,\ \ j = 1, ..., J, \tag{13}$$

$$0 \le \delta_1 \le 1,\ \ 0 \le \delta_2 \le 1 \tag{14}$$

In the above model, θ_{ij} ($i = 1, \ldots, 7$) and a^* ($0 < a^* < 1$) are engineered by enterprise's decision makers, $\theta_{i^* j}$ $=max\{\theta_{ij}, i = 1, \ldots, 7\}$. p_1, p_2, p_3 are specified by the cloud storage service providers, while n_j ($j = 4, ..., J$) is specified by enterprise's decision makers.

If the enterprise has clarified its total budget (P) in HCSS, the above model needs considering the corresponding budget constraint: $p_1 \cdot x_1+p_2 \cdot x_2+p_3 \cdot x_3 \cdot t+ \sum_{j=4}^{J} n_j \cdot x_j \le P$. case I

4. Cases Analysis

4.1. Case I

A certain research institute accumulates a large number of data with its research deepening, scattered in different servers. Due to the further study of scientific research and accumulation, managers upgrade and transform its data management platform by using HCSS, considering characteristics of research data and research demands.

As for public cloud storage, Google is chosen to be the service provider. Consulting Website of Chinese Internet Network Information Center, it can be seen that unit price of Google GAE (RMB¥) is p1=0.68, p2=0.816, p3=1.02 in sequence (by month) in survey of the market price. It can be obtained that uploading data size (GB), downloading data size (GB), storage data size (GB) are at the interval[260,400], [230,420], [400,780] respectively in terms of interval range by making analysis of the existing data of the Institute. Data storage time to be signed lasts a year, namely t=12.

As for private cloud storage, it can be found that deploying the system needs 2 servers, a set of virtualization software by integrating resources of the Institute. It can be determined that unit price of the server is at the interval (RMB¥) [11000, 30000] and unit price of virtualization software is at the interval (RMB¥) [1000, 6000] by making an analysis of the data storage and management needs of the Institute.

The total budget for HCSS is fixed at 50,000 RMB¥, considering the research data storage, management situation and research funds of the Institute. Meanwhile, a^* is specified as 0.1 by managers.

Managers determine the influencing weights of each variable on each factor by using the method in References [12], as shown in Table 1.

Table 1. Influencing Weights of Each Variable on Factors in Case I

	Uploaded data size (x_1)	Downloaded data size	Storage data size (x_3)	server unit price(x_4)	Unit price of virtual

		(x_2)			software (x_5)
Flexibility (u_1)	0.3	0.3	0.25	0.2	0.25
Scalability (u_2)	0.1	0.1	0.2	0.1	0.15
Cut in input (u_3)	0.15	0.15	0.2	0.3	0.2
Improvement on access speed (u_4)	0.15	0.15	0.15	0.15	0.15
Data sharing (u_5)	0.15	0.15	0.1	0.1	0.15
Disaster recovery (u_6)	0.05	0.05	0.05	0.1	0.05
Green IT(u_7)	0.1	0.1	0.05	0.05	0.05

Based on the given information, we construct the goal programming model considering budget constraints, finding the solution using the fmincon function in the Matlab environment. The best regularized total profit is 0.7828; the regularized total cost is 0.2105; variables (x_1, x_2, x_3, x_4, x_5) = (400, 420, 500.94, 14795.23, 2007.3); PuCSS fee is 6744.7 RMB¥; PrCSS fee is 31599.25 RMB¥; the total service cost is 38343.95 RMB¥. It is obvious that PuCSS costs account for 17.59% of the total cost of storage service. In general, suppose the proportion of PuCSS fees to total storage service fees in the optimized results is r_1, then $r_1=(p_1*x_1+p_2*x_2+p_3*x_3*t)/(p_1*x_1+p_2*x_2+p_3*x_3*t+\sum_{j=4}^{J} n_j \cdot x_j)$, $0 \le r_1 \le 1$.

For total budget ranges from 30,000 RMB¥ to 60,000 RMB¥ and the optimization step is 5,000 RMB¥, the optimization results are listed in Table 2.

Table 2. The Optimization Results for Different Budgets in Case I

Total budgets	Total profits	Total costs	r_1	x_1	x_2	x_3	x_4	x_5	Total fees
30000	0.0669	0.0363	17.53%	260	230	400	11869.76	1000	30000
35000	0.2458	0.1407	15.03%	260	230	400	13995.81	1747.9	35000
40000	0.7896	0.245	15.6%	400	420	459.47	15744.82	2267.51	39995.81
45000	0.7833	0.2098	17.52%	400	420	498.07	14796.65	2006.16	38310.56
50000	0.7828	0.2105	17.59%	400	420	500.94	14795.23	2007.3	38343.95
55000	0.7833	0.2098	17.52%	400	420	498.07	14796.63	2006.22	38310.63
60000	0.7825	0.2111	17.65%	400	420	503.27	14795.62	2007.62	38373.54

It can be easily seen from Table 2 that when the total budget is small, the total costs and total profits are also small, and most of the variables taking the lower limit is conducive to PrCSS to some extent, the total storage cost equals to the total budget. With the increase in the total budget, the relevant variables of PuCSS taking the upper limit, is conducive to the PuCSS, total profit, total cost, r_1, and the total cost of storage service tend to be stable.

4.2. Case II

A certain automobile manufacturing enterprise has accumulated a large amount of data in its design and production, and the data is distributed in the server of various departments, leading to difficulties in data management and sharing. In order to overcome the difficulties, improve the data management, the enterprise's managers consider managing data by using HCSS.

As for public cloud storage, Google is chosen to be the service provider. Consulting Website of Chinese Internet Network Information Center, it can be seen that unit price of Amazon EC2 (RMB¥) is p_1=0.68， p_2=1.292， p_3=0.748 in sequence (by month) in survey of the market price. Amazon EC2 provides 160G of free storage space, and the exceeding part is calculated at p_3=0.748. It can be obtained that uploading data size (GB), downloading data size (GB), storage data size (GB) are at the interval[3200,4500]，[3300,4700]，[4000,7000] respectively in terms of interval range by making analysis of the existing data of the Institute. Data storage time to be signed lasts one year, namely t=12.

On analysis of existing enterprise data, we obtain the upload data size enterprises (GB), download the data size (GB), data storage size (GB) interval were [32004500], [33004700], [40007000], data storage time plan signed for a year, namely t=12.

It is found that the system deployment needs 5 servers, and a set of virtualization software by integrating existing enterprise resources for private cloud storage,. It can be determined that unit price of the server is at the interval of [70000, 150000], and the unit price of virtualization software is at the interval of [5000, 20000] by analyzing enterprise's data storage and management needs.

Considering data storage and liquidity in enterprise design and production, managers determine the total budget for HCSS is 600,000 RMB¥. Meantime, a^* is specified as 0.1. The influencing weights of each variable on each factor are determined by using the method in References [12], as shown in Table 3.

Table 3. Influencing Weights of Each Variable on Factors in Case II

	uploading data size (x_1)	downloading data size (x_2)	storage data size(x_3)	Unit price of server (x_4)	Unit price of visualization software(x_5)
flexibility (u_1)	0.15	0.15	0.2	0.2	0.2
extendibility (u_2)	0.25	0.25	0.15	0.1	0.15
decrease in investment (u_3)	0.15	0.15	0.2	0.2	0.2
improvement on access speed (u_4)	0.15	0.15	0.15	0.15	0.1
data sharing (u_5)	0.15	0.15	0.1	0.1	0.15
disaster recovery (u_6)	0.1	0.1	0.1	0.15	0.15
green IT(u_7)	0.05	0.05	0.1	0.1	0.05

Based on above information, we can construct the goal programming model according to the budget constraints, using the fmincon function in the Matlab environment to find the solutions: the best regularized total profit is 0.5734; the regularized total cost is 0.459, variables(x_1, x_2, x_3, x_4, x_5) = (3200, 3300, 4000, 109440.88, 11888.18); PuCSS fee is 40907.44 RMB¥; PrCSS fee is 559092.56 RMB¥; the total storage service fee is 600,000 RMB¥. Obviously, the proportion of PuCSS costs to the total cost of storage service is 6.82%.

Suppose the total budget ranges between 600,000 RMB¥ and 900,000 RMB¥, and the optimization step is 50,000 RMB¥, and the optimized results are listed in table 4.

Table 4. The Optimization Results for Different Budgets in Case II

Total budget	Total profit	Total cost	r_1	x_1	x_2	x_3	x_4	x_5	Total fees
600000	0.5734	0.459	6.82%	3200	3300	4000	109440.88	11888.18	600000
650000	0.687	0.5715	6.29%	3200	3300	4000	119056.26	13811.25	650000
700000	0.7917	0.6839	5.84%	3200	3300	4000	128671.65	15734.33	700000
750000	0.8892	0.7964	5.45%	3200	3300	4000	138287.03	17657.41	750000
800000	0.5632	0.5	9.01%	3850	4000	5500	110000	12500	618217.84
850000	0.5327	0.5	9.01%	3850	4000	5500	110000	12500	618217.84
900000	0.5327	0.5	9.01%	3850	4000	5500	110000	12500	618217.84

As can be inferred from Table 4, total profits and total costs are on the gradual rise with the increase in total budgets. Related variables of the PuCSS taking the lower limit, is conducive to PrCSS to some extent, and the total storage cost is roughly the same as the total budget. When the total budget increases to 800000 RMB¥, total profit, total cost, r_1, and total fees of storage service are almost stable. Compared with relatively small total budget, the stable total cost and total profit are on the decline due to the fact that stable r_1 is on the increase. Namely, related variables of PuCSS and PrCSS can achieve the best balance.

Comparing Table 2 with Table 4, it can be seen that r_1 in case I is significantly higher than that in case II, which coincides with the application and demand of the above two cases. As for University Research Institute, the large number of scientific data is non confidential, and only a small part of unpublished data is confidential. Therefore, the requirement of the server performance is not strict, preferring PuCSS to PrCSS. However, the large number of data related to design, production in the Automobile Manufacturing Enterprise needs keeping secret. Besides, the requirement for data management and sharing is strict. Therefore, its requirement for the server performance is stricter, preferring PrCSS. Accordingly, the total profit is higher and the total cost is low in case I in the steady state, while total profit and total cost in case II is balanced.

It is worthy of note that costs of PrCSS are much higher than those of PuCSS in terms of absolute costs from the above two cases. What accounts for this is that PrCSS can bring about long-term profits, while PuCSS can only lead to temporary profits.

5. Conclusions

In conclusion, the goal programming model collaboratively optimizing HCSS profits and costs is constructed in this paper, to achieve the best configuration of public and cloud storage service. HCSS profit function considers 7 factors, while HCSS cost function contains 3 factors. The constructed collaborative optimization model is validated by making a comparative analysis of a University Research Institute and an Automobile Manufacturing Enterprise.

Acknowledgements

The research is supported by Anhui provincial Natural Science Foundation Program (1408085MG136) ; the Fundamental Research Funds for the Central Universities （2013HGBZ0599）; Humanities and Social Science Foundation Program of Ministry of Education (14YJA630051); China National Nature Science Foundation (71403001).

References

[1]. K. Chen and W. M. Zheng, “Cloud Computing: System Instances and Current Research”, Journal of Software（in Chinese, vol. 20, no. 5, (**2009)**, pp. 1337-1348.

[2]. R. Buyya, Yeo, C. S, S. Venugopal, J. Broberg, and I. Brandic, “Cloud computing and emerging IT platforms: Vision, hype, and reality for delivering computing as the 5th utility”, Future Generation Computer Systems, vol. 25, no. 6, (**2009)**, pp. 599-616.

[3]. D. G. Feng, M. Zhang, Zhang Y and Xu Z, “Study on cloud computing security”, Journal of Software (in Chinese), vol. 22, no. 1, (**2011)**, pp. 71-83.

[4]. W. Zen, W. Zhao and M. Shang, “Research on Cloud Computing and Cloud Storage Ecosystem”, Journal of Computer Research and Development (in Chinese), vol. 48, (**2011)**, pp. 234-239.

[5]. J. Gao, Z. Wu and S. Sun, “Research on the Application of Cloud Storage in Digital Preservation”, New Technology of Library and Information Service (in Chinese), vol. 6, **(2010),** pp. 1-6.

[6]. Li Cuixia, “Study on library storage at colleges and universities based on hybrid cloud”, Research on Library Science, (in Chinese), vol. 3, **(2011),** pp. 68-71.

[7]. S. Wang, L. Guo, L. Kang, Q. Li, W. Song and G. Chen, “Cloud manufacturing application mode and program analysis”, Computer Integrated Manufacturing System (in Chinese), vol. 18, no. 7, **(2012),** pp.

1637-1643.
[8]. C. M. Subhas and M. Arka, "Identification of a company's suitability for the adoption of cloud computing and modeling its corresponding Return on Investment", Mathematical and Computer Modeling, vol. 53, no. 3-4, **(2011),** pp. 504-521.
[9]. K. Ye, Z. Wu, X. Jiang and H. Qinming, "Power Management of Virtualized Cloud Computing Platform", Chinese Journal of Computers (in Chinese), vol. 35, no. 6, **(2012),** pp. 1262-1285.
[10]. Q. Lin, X. Gui, D. Shi and X. Wang, "Study of the Secure Access Strategy of Cloud Storages', Journal of Computer Research and Development (in Chinese), S1, **(2011),** pp. 240-243.
[11]. T. Hong-Linh and D. Schahram, "Composable cost estimation and monitoring for computational applications in cloud computing environments", Procedia Computer Science, vol. 1, no. 1, **(2010),** pp. 2175-2184.
[12]. A. I. Ölçer and A. Y. Odabaşsi, "A new fuzzy multiple attributive group decision making methodology and its application to propulsion/manoeuvring system selection problem", European Journal of Operational Research, vol. 166, **(2005),** pp. 93-114.

Authors

Chao-ping SUN

PhD, Master Tutor, associate professor at Hefei University of Technology, Anhui Provincial Distinguished Teacher
Email: tedsuncp@163.com

Jing ZHAO

Master of Management, Hefei University of Technology
Email: 41789357@qq.com

Su-Mei RUAN

PhD, Master Tutor, associate professor at Anhui University of Finance and Economics
Email: 2413814264@qq.com