International Journal of Computers and Applications, Vol. 29, No. 2, 2007

ON PROFIT DENSITY BASED GREEDY ALGORITHM FOR A RESOURCE ALLOCATION PROBLEM IN WEB SERVICES

J. Sum*,*** J. Wu,** and C.-S. Leung***

Abstract

Allocating limited computational resources to different clients is always a challenging problem to a web service provider (WSP). Profit density based greedy knapsack algorithm is one simple approach that can ensure near-optimal profit. However, profit gain is sometimes not the only factor concerned in making important management decisions. Other factors, such as the number of clients that a WSP can serve and the number of un-used resources that remain, are also important. By assuming that (a) the pricing curves of the buyer are all identical and their marginal utility (i.e., $\Delta Price/\Delta Size$) is decreasing, (b) the resource is divisible, (c) the resource quantity each client requests follows uniform distribution $U[0,1]$ and (d) the available resource is constrained by $\bar{k}$; equations for the expected number of clients who can get the resource, denoted by $\langle b\rangle$, and the expected quantity of resource being allocated, denoted by $\langle s\rangle$, are derived analytically. By observing the numerical plots of $\langle b\rangle$ and $\langle s\rangle$ against the number of clients n, it is found that $\langle b\rangle \approx n$ for $n \leq 2\bar{k}$ and $\langle b\rangle \approx (-1+\sqrt{1+8n\bar{k}})/2$ for $n \geq 2\bar{k}$. Comparing with another simple selling mechanism, we call it first-come-first-serve, it is found that resource allocation via greedy algorithm might not always be the best choice as far as the number of units being sold and the number of clients being served are concerned.

Key Words

Knapsack problem, order statistics, profit density greedy algorithm, sum of random variables, uniform distribution

1. Introduction

Extended from the ideas of software reuse and component based development, web service is a new paradigm and possibly a new direction for system development. A web services provider (WSP) makes application components available on the web. System developers can thus integrate those components (URLs) together to develop an application system. Certainly, the usage of these remote resources is usually not free. Allocating limited computational resources to clients to maximize the profit is one, but not the only, issue that every WSP needs to consider.

To solve this problem, one can apply an off-line allocation method. Let us consider a simple but normally not quite realistic situation. For clarification, Table 1 summarizes the notations appearing in the paper. Consider a WSP that has 20 servers available to support the service and 8 clients are willing to pay for their services:

Client i	B1	B2	B3	B4	B5	B6	B7	B8
Server k_i	2	4	5	1	3	4	2	5
Price p_i	10	30	35	6	15	18	12	35

As resource is limited, the WSP has to select the most profitable clients and sign the service contracts. Obviously, this problem is essentially the 0/1 knapsack problem [1, 2] (or recently it is called multi-units combinatorial auction problem [3]) that can be formulated by the following constraint optimization problem:

$$\begin{aligned}
\text{Maximize} \quad & 10s_1 + 30s_2 + 35s_3 + 6s_4 + 15s_5 + 18s_6 \\
& + 12s_7 + 35s_8 \\
\text{Subject to} \quad & 2s_1 + 4s_2 + 5s_3 + s_4 + 3s_5 + 4s_6 + 2s_7 \\
& + 5s_8 \leq 20 \\
& s_i \in \{0,1\}\ \forall\, i = 1,\ldots,8
\end{aligned}$$

For this simple problem, the WSP can profit 133 by allocating all 20 servers to B2, B3, B4, B5, B7 and B8. However, it is known that solving this constraint optimization problem will be intractable if the number of clients is large. Profit density based greedy algorithm is a near-optimal alternative [2], which profits 128, by allocating 19 servers

* Department of Information Management, Chung Shan Medical University, Taichung, Taiwan, ROC; e-mail: pfsum@csmu.edu.tw.

** Department of Computer Science and Engineering, Florida Atlantic University, Boca Raton, FL 33431, USA; e-mail: jie@cse.fau.edu.

*** Department of Electronic Engineering, City University of Hong Kong, Kowloon Tong, KLN, Hong Kong; e-mail: eeleungc@cityu.edu.hk.

Recommended by Prof. Xiao Chen
(paper no. 202-1911)

Table 1
Notations. [Here $k_i \sim U[0,1]$ means k_i is a random variable following uniform distribution.

Notation	Description
n	Number of customers
$\overline{k}$	Resource available
(p_i, k_i)	Offer that the i-th customer gives p_i – price; $k_i \sim U[0,1]$ – quantity
i	Index of the customers according to their arrival sequence i.e., offer given from the 1st customer is earlier than the 2nd customer, and so on (Used in FCFS analysis)
$i{:}n$	Index of the customers according to their profit densities* i.e., $p_{i:n}/k_{i:n} > p_{j:n}/k_{j:n}$ if $i<j$ (Used in greedy algorithm analysis)
$S_r(w,n)$	$\sum_{i=1}^{w} k_i$ – sum of the quantities of the first w customers (Used in FCFS analysis)
$S(w,n)$	$\sum_{i=1}^{w} k_{i:n}$ – sum of the quantities of the first w customers whose orders are sorted according to profit density (Used in greedy algorithm analysis)
b	Number of customers being served
$\langle b\rangle$	Expected number of customers being served
s	Quantity of resource being allocated
$\langle s\rangle$	Expected quantity of resource being allocated

* Suppose there are three customers, their offers are $(3,0.2)$, $(7,0.8)$ and $(5,0.4)$. (p_i,k_i) for $i=1,2,3$ will be $(3,0.2)$, $(7,0.8)$ and $(5,0.4)$, respectively. $(p_{i:3},k_{i:3})$ for $i=1,2,3$ will be $(3,0.2)$, $(5,0.4)$ and $(7,0.8)$. Then $S_r(2,3)=1.0$ and $S(2,3)=0.6$.

to B1, B2, B3, B4, B7 and B8. One server remains in the stock.

To trade off the computational complexity, another even simpler mechanism called first-come-first-serve (FCFS) – allocating resource to the client whenever the number of servers is available – can be used instead. In terms of profit gain, it is clear that profit density greedy algorithm is a better allocation method as it will ensure near-optimal profit for the number of customers is large. However, profit gain is sometimes not the only measure a company would like to know. Other factors, such as the number of clients it serves and the number of unused resources remaining, are also important for making management decision.

In this paper, we assume that $\overline{k} \gg 1$ units of resource are available. Let the expected number of clients who can get the resource be $\langle b\rangle$, and the expected quantity of resource being allocated be $\langle s\rangle$, the purpose of the paper is to find out their relationships in terms of n and $\overline{k}$. The essential technique being used is a formula derived by Weisberg [4] for a linear combination of order statistics and a formula derived by Feller (p. 27 of [5]) for the sum of uniformly random variables. The next section will describe the basic assumptions on p_i and k_i. The profit density greedy algorithm and the FCFS mechanism will be presented. The expected number of customers $\langle b\rangle$ and the expected number of product being sold $\langle s\rangle$ for the mechanisms will be derived in Section 3. A discussion comparing greedy algorithm against FCFS method will be presented in Section 4. Then the conclusion will be presented in Section 5.

2. Greedy Algorithm and FCFS

Without loss of generality, we assume that k_i is a random variable from $U(0,1)$. Next, we assume that the pricing function is marginal utility decreasing [6]. That is to say, a client would like to have a larger discount for a larger purchase. Mathematically, (i) $p'(k_i) \geq p'(k_j)$, $\forall\, 0 \leq k_i \leq k_j \leq 1$ and (ii) $p(0) \geq 0$ and $p'(0) > 1$, where $p'(k_i)$ is the first derivative of the function $p(k)$ at k_i. Two exemplar functions satisfying the assumption are $p(k) = \alpha k + \beta$ and $p(k) = \alpha \log(1+k) + \beta$, where $k \in [0,1]$, α and β are non-negative constant values. It should be noted that the pricing function is a deterministic function depending solely on the quantity of resource requested. The following lemma will be used for latter analysis.

Lemma 1. For any non-negative real-valued function $f(x)$ that satisfies $f'(x) \geq f'(y) \geq 0$ for all $0 \leq x \leq y$ and $f(0) \geq 0$, then the following conditions hold: (i) $\frac{f(x)}{x} \geq f'(x)$; (ii) $\frac{f(x)}{x} \geq \frac{f(y)}{y}$, for all $0 \leq x \leq y$.

Proof: The proof of the first inequality is straightforward. As $f(x) = f(0) + \int_0^x f'(u)du \geq f(0) + xf'(x)$, $f(x)/x \geq f'(x)$. Using the fact that $f(y) = f(x) + \int_x^y f'(u)du$, then dividing both sides by y and the condition, $f'(x) \geq f'(y)$, the following inequality can be obtained:

$$\frac{f(y)}{y} \leq \frac{f(x)}{x} + \frac{(y-x)}{y}\left[f'(x) - \frac{f(x)}{x}\right]$$

As $\frac{f(x)}{x} \geq f'(x)$, for all $y \geq x \geq 0$, $\frac{f(x)}{x} \geq \frac{f(y)}{y}$ and the proof is completed. **Q.E.D.** □

2.1 Greedy Algorithm

Suppose there are n clients whose prices and quantities are $p_1, \ldots, p_n$ and $k_1, \ldots, k_n$, respectively. We call (p_i, k_i) for all $i = 1, 2, \ldots, n$ the offers the clients give. Once all the offers have been collected, the WSP can apply the algorithm below to determine the allocation:

1: WAITFOR (p_i, k_i), $i = 1, \ldots, n$;
2: SORT $\{\frac{p_i}{k_i}\}$ s.t. $\frac{p_{i:n}}{k_{i:n}} \geq \frac{p_{j:n}}{k_{j:n}} \quad \forall\, i \leq j$;

3: SET $C = \overline{k}$;
4: SET $P = 0$;
5: SET $j = 1$;
6: WHILE($C - k_{j:n} > 0$ and $j \leq n$)
   $C = C - k_{j:n}$;
   $P = P + p_{j:n}$;
   $j = j + 1$;
  END

First, their offers are ranked in descending order with respect to their profit density, i.e.,

$$\frac{p_{1:n}}{k_{1:n}} \geq \frac{p_{2:n}}{k_{2:n}} \geq \cdots \geq \frac{p_{n:n}}{k_{n:n}} \tag{1}$$

Then, we allocate the resource to the first $1{:}n$, $2{:}n$, ..., $b{:}n$ clients, such that:

$$\sum_{i=1}^{b} k_{i:n} \leq \overline{k} \quad \sum_{i=1}^{b+1} k_{i:n} > \overline{k} \tag{2}$$

In accordance with the condition (1), the condition $\frac{p_{i:n}}{k_{i:n}} \geq \frac{p_{j:n}}{k_{j:n}}$ implies that $k_{1:n} \leq k_{2:n} \leq \cdots \leq k_{n:n}$ whenever price p is a function of k and its marginal utility is decreasing. So, $\frac{p_{1:n}}{k_{1:n}} \geq \frac{p_{2:n}}{k_{2:n}} \geq \cdots \geq \frac{p_{n:n}}{k_{n:n}}$ implies that $k_{1:n} \leq k_{2:n} \leq \cdots \leq k_{n:n}$ and their offers can be ranked in accordance with k_i. Again, the $\overline{k}$ units are allocated to the first b bidders according to conditions in (2).

2.2 First-Come-First-Serve

FCFS method is similar to selling products in a flea market. Once a customer has walked in and given an offer, the seller will check with the stock. The customer gets the product as long as there is available stock. One advantage of this FCFS method apart from its simplicity is that the customers do not have to wait. Besides, the seller has no need to anticipate the number n.

In web service provision, the WSP simply denies the service request whenever the available resource is not large enough to support the service. The FCFS method can be described by the following algorithm:

1: SET $C = \overline{k}$;
2: SET $P = 0$;
3: SET $j = 1$;
4: WHILE($C - k_j > 0 \quad$ and $\quad j \leq n$)
   $C = C - k_j$;
   $P = P + p_j$;
   $j = j + 1$;
  END

Here, the index $i = 1, 2, \ldots,$ represent the sequence of the offers made by the clients. In other words, the index $i = 1, 2, 3, \ldots,$ indicates their timing of visit. The i-th client makes an offer earlier than the j-th client if $i < j$. In this method, the WSP has no need to wait until all the offers have been collected. The decision is simply made by investigating the number of resources remaining. The resource is allocated to the first $1, 2, \ldots, b$ clients:

$$\sum_{i=1}^{b} k_i \leq \overline{k} \quad \sum_{i=1}^{b+1} k_i > \overline{k} \tag{3}$$

If the resource remaining is larger than the quantity requested by the walk-in client, the resource will be allocated accordingly.

3. Analysis

For the sake of analysis, let $S(w,n) = \sum_{i=1}^{w} k_{i:n}$ be the sum of units being sold to the $\{1{:}n\}$, $\{2{:}n\}$, ..., $\{b{:}n\}$ customers based on the profit density greedy algorithm. Similarly, we let $S_r(w,n) = \sum_{i=1}^{w} k_i$ be the sum of units being sold to the $1, 2, \ldots, b$ customers based on the FCFS method.

3.1 Greedy Algorithm

As k_i is a random variable drawn from uniform distribution for all $i = 1, 2, \ldots, n$, $k_{i:n}$ (after being sorted by profit density) is also a random variable drawn from uniform distribution for all $i = 1, 2, \ldots, n$:

$$S(w,n) \;=\; \sum_{i=1}^{n} d_i \; k_{i:n} \tag{4}$$

$$d_i \;=\; \begin{cases} 1 \;\forall\, i = 1, \ldots, w \\ 0 \;\forall\, i = w+1, \ldots, n \end{cases} \tag{5}$$

The cumulative probability distribution $Pr\{S(w,n) \leq \overline{k}\}$ can be evaluated by a formula derived by Weisberg in [4] (see Appendix A),

$$Pr\{S(w,n) \leq \overline{k}\} \;=\; 1 - \sum_{j=1}^{r} \frac{(c_j - \overline{k})^n}{c_j \prod_{j \neq i}(c_j - c_i)} \tag{6}$$

$$c_i \;=\; \begin{cases} w - i + 1 \;\forall\, i = 1, \ldots, w \\ 0 \qquad\qquad \forall\, i = w+1, \ldots, n \end{cases} \tag{7}$$

Unfortunately, this formula (as well as another formula from Feller [5]) cannot be reduced to a simple close form. To obtain the solution, one needs to do it numerically.

For the case that exactly w customers are allocated with resources, it is equivalent to the case:

$$\{S(w,n) \leq \overline{k} \quad \text{and} \quad S(w+1,n) > \overline{k}\}$$

Consider the following events,

$$\begin{aligned} E_1 &= \{S(w,n) \leq \overline{k} \quad \text{and} \quad S(w+1,n) \leq \overline{k}\} \\ E_2 &= \{S(w,n) \leq \overline{k} \quad \text{and} \quad S(w+1,n) > \overline{k}\} \\ E_3 &= \{S(w,n) > \overline{k} \quad \text{and} \quad S(w+1,n) \leq \overline{k}\} \\ E_4 &= \{S(w,n) > \overline{k} \quad \text{and} \quad S(w+1,n) > \overline{k}\} \end{aligned}$$

and the facts that (i) $Pr\{E_1\} + Pr\{E_2\} + Pr\{E_3\} + Pr\{E_4\} = 1$ and (ii) $E_3 = \phi$ the empty set, the probabilities for the events can readily be determined as follows:

$$
\begin{aligned}
Pr\{E_1\} &= Pr\{S(w+1,n) \le \overline{k}\} \\
Pr\{E_2\} &= Pr\{S(w,n) \le \overline{k}\} - Pr\{S(w+1,n) \le \overline{k}\} \\
Pr\{E_3\} &= 0 \\
Pr\{E_4\} &= 1 - Pr\{S(w,n) \le \overline{k}\}
\end{aligned}
$$

3.1.1 Number of Clients $\langle b \rangle$ Being Allocated with Resources

The probability of exactly w clients being allocated with resources can be determined as follows:

$$
Pr\{w \text{ clients}\} = \begin{cases} Pr\{S(w,n) \le \overline{k}\} & \\ -Pr\{S(w+1,n) \le \overline{k} & \text{if } w < n\} \\ Pr\{S(n,n) \le \overline{k}\} & \text{if } w = n \end{cases} \tag{8}
$$

This equation applies for all $w \ge \overline{k}$ and the evaluation of the $Pr\{S(w,n) \le \overline{k}\}$ can be based on (6). Thus, the expected number of clients being allocated with resources, $\langle b \rangle$, can be determined by the following formula:

$$
\begin{aligned}
\langle b \rangle &= \sum_{w=1}^{n-1} w(Pr\{S(w,n) \le \overline{k}\} - Pr\{S(w+1,n) \le \overline{k}\}) \\
&\quad + nPr\{S(n,n) \le \overline{k}\}
\end{aligned} \tag{9}
$$

As $Pr\{b = w\} = 1$ for all $w \le \overline{k}$, the summation can be started with $w = \overline{k}$:

$$
\begin{aligned}
\langle b \rangle &= \sum_{w=\overline{k}}^{n-1} w(Pr\{S(w,n) \le \overline{k}\} - Pr\{S(w+1,n) \le \overline{k}\}) \\
&\quad + nPr\{S(n,n) \le \overline{k}\}
\end{aligned} \tag{10}
$$

It is a function dependent on n and $\overline{k}$. Once n and $\overline{k}$ are known, $\langle b \rangle$ can be evaluated numerically. Fig. 1 illustrates the case when $\overline{k} = 20$. We have also plotted the curve for the cases when $\overline{k}$ equals to 30 and 40, respectively. All of them show the same shape. It can be observed that for $n \le 2\overline{k}$ and $n \ge 2\overline{k}$, $\langle b \rangle$ can be approximated as follows:

$$
\langle b \rangle \approx \begin{cases} n & \text{if } n \le 2\overline{k} \\ \dfrac{-1+\sqrt{1+8n\overline{k}}}{2} & \text{if } n \ge 2\overline{k} \end{cases} \tag{11}
$$

The approximations are shown by dotted line and the dot-solid line, respectively, in Fig. 1. A derivation for the case when $n \gg \overline{k}$ can be found in Appendix C.

3.1.2 Quantity of Resource $\langle S \rangle$ Being Allocated

The expected quantity of resource $\langle S \rangle$ being allocated can thus be evaluated by using a similar argument. First, let us consider the event (exactly w clients will be allocated with resources and x quantities of resource will be allocated), i.e., $\{S(w,n) \le x$ and $S(w+1,n) \ge \overline{k}\}$. Obviously, $\overline{k} - 1 \le x \le \overline{k}$. As $\{S(w,n) \le x\}$ equals

Figure 1. The expected number of clients being allocated with resource against the number of customers for $\overline{k} = 20$ is shown by solid line with circles. The dotted line corresponds to $\langle b \rangle = n$ and the dot-solid line corresponds to $\langle b \rangle = \frac{-1+\sqrt{1+8n\overline{k}}}{2}$.

$$
\begin{aligned}
&\{S(w,n) \le x \quad \text{and} \quad S(w+1,n) \le \overline{k}\} \bigcup \{S(w,n) \\
&\quad \le x \quad \text{and} \quad S(w+1,n) \ge \overline{k}\}
\end{aligned}
$$

and the first event is equivalent to $\{S(w+1,n) \le \overline{k}\}$, it is readily shown that:

$$
\begin{aligned}
&Pr\{S(w,n) \le x \quad \text{and} \quad \text{exactly } w \text{ clients being allocated}\} \\
&\quad = Pr\{S(w,n) \le x \quad \text{and} \quad S(w+1,n) \ge \overline{k}\} \\
&\quad = Pr\{S(w,n) \le x\} - Pr\{S(w+1,n) \le \overline{k}\}
\end{aligned} \tag{12}
$$

for all $x \in \{y | Pr\{S(w,n) \le y\} - Pr\{S(w+1,n) \le \overline{k}\} \ge 0\}$. Let

$$
\begin{aligned}
h(x|w,n,\overline{k}) &= Pr\{S(w,n) \\
&\qquad = x| \text{ exactly } w \text{ clients being allocated}\}
\end{aligned}
$$

It can thus be evaluated as follows:

$$
h(x|w,n,\overline{k}) = \frac{d}{dx}\left\{ \frac{\begin{array}{c} Pr\{S(w,n) \le x\} \\ -Pr\{S(w+1,n) \le \overline{k}\} \end{array}}{\begin{array}{c} Pr\{S(w,n) \le \overline{k}\} \\ -Pr\{S(w+1,n) \le \overline{k}\} \end{array}} \right\} \quad \text{if } w < n \tag{13}
$$

$$
h(x|w,n,\overline{k}) = \frac{d}{dx}\left\{ \frac{Pr\{S(n,n) \le x\}}{Pr\{S(n,n) \le \overline{k}\}} \right\} \quad \text{if } w = n \tag{14}
$$

for all $x \in \{y | Pr\{S(w,n) \le y\} - Pr\{S(w+1,n) \le \overline{k}\} \ge 0\}$. The expected quantity of resource being allocated $\langle S \rangle$ can thus be written as follows:

$$\langle S\rangle = \sum_{w=\overline{k}}^{n-1}\int_{x_w}^{\overline{k}} x\, d(Pr\{S(w,n)\leq x\} - Pr\{S(w+1,n) \leq \overline{k}\}) + \int_0^{\overline{k}} x\, dPr\{S(n,n)\leq x\} \quad (15)$$

for all $n \geq \overline{k}$ and x_w satisfies the condition:

$$Pr\{S(w,n)\leq x_w\} = Pr\{S(w+1,n)\leq \overline{k}\}$$

Fig. 2 shows the case when $\overline{k}$ equals to 20.

Figure 2. The expected quantity of resource being allocated $\langle S\rangle$ against the number of customers n for $\overline{k} = 20$.

For large n, an approximated equation for the expected quantity of resource being allocated can be derived. Considering the residue, $R(n,\langle b\rangle,\overline{k}) = \overline{k} - \sum_{i=1}^{\langle b\rangle}\frac{i}{n}$ satisfies the following inequality: $0 \leq R(n,\langle b\rangle,\overline{k}) \leq (\langle b\rangle + 1)/n$ and supposing that this residue is uniformly distributed on $[0,(\langle b\rangle+1)/n]$. The expected residue $\langle R\rangle$ can be written as follows: $\langle R\rangle = (\langle b\rangle + 1)/2n$. Substituting the approximation for $\langle b\rangle$ in (11), the approximation of the expected quantity of resource being allocated can be written as follows:

$$\langle S\rangle \approx \overline{k}\left(\frac{\sqrt{1+8n\overline{k}}-3}{\sqrt{1+8n\overline{k}}-1}\right) \quad (16)$$

for $n \gg \overline{k}$. Reader can also refer to Appendix C for a derivation of the above equation.

3.2 First-Come-First-Serve

For the case that the resource is allocated in an FCFS basis, we consider the following equation: $S_r(w,n) = \sum_{i=1}^{w} k_i$, for all $\overline{k} \leq w \leq n$. By replacing $S(w,n)$ by $S_r(w,n)$, we can use the same argument used for greedy method to derive the equations for the expected number of clients being allocated with resource $\langle b_r\rangle$ and the expected quantity of resource $\langle S_r\rangle$ being allocated.

3.2.1 Number of Clients $\langle b_r\rangle$ Being Allocated with Resource

The expected number of clients being allocated with resources $\langle b_r\rangle$ can be determined by the following formula:

$$\langle b_r\rangle = \sum_{w=\overline{k}}^{n-1} w\{Pr\{S_r(w,n)\leq \overline{k}\} - Pr\{S_r(w+1,n) \leq \overline{k}\}\} + nPr\{S_r(n,n)\leq \overline{k}\} \quad (17)$$

The expression for $Pr\{S_r(w,n)\leq x\}$ will be from Feller formula [5]:

$$Pr\{S_r(w,n)\leq x\} = \frac{1}{w!}\sum_{i=0}^{w}(-1)^i C_i^w (x-i)_+^w \quad (18)$$

where

$$x_+ = \frac{x+|x|}{2} \quad \text{and} \quad C_i^w = \frac{w!}{i!(w-i)!}$$

It should be noted that $Pr\{S_r(w,n)\leq x\}$ is independent of n. Fig. 3 shows the expected number of clients being allocated with resources against number of customers n for the case that $\overline{k} = 20$.

3.2.2 Quantity of Resource $\langle S_r\rangle$ Being Allocated

Using the same technique as for $\langle S\rangle$, the expected number of units being sold $\langle S_r\rangle$ can be determined by the following equation:

Figure 3. The expected number of clients being allocated $\langle b_r\rangle$ with resource against the number of customers n for $\overline{k} = 20$.

$$\langle S_r \rangle = \sum_{w=\overline{k}}^{n-1} \int_{x_w}^{\overline{k}} x\, d\{Pr\{S_r(w,n) \le \overline{k}\} - Pr\{S_r(w+1,n) \le \overline{k}\}\} + \int_0^{\overline{k}} x\, dPr\{S_r(n,n) \le x\} \tag{19}$$

for all $n \ge \overline{k}$. x_w satisfies the condition:

$$Pr\{S_r(w,n) \le x_w\} = Pr\{S_r(w+1,n) \le \overline{k}\}$$

Fig. 4 shows the case when $\overline{k} = 20$. It should be noted that the expected quantity being allocated by FCFS method is slightly larger than the expected quantity being allocated by profit density based greedy algorithm (Fig. 5).

Figure 4. The expected quantity of resource being allocated $\langle S_r \rangle$ against the number of customers n for $\overline{k} = 20$.

Figure 5. Comparison between FCFS and the greedy method in terms of the expected quantity of resource being allocated $\langle S \rangle$ (solid line with circles) and $\langle S_r \rangle$ (solid line with dots) for $\overline{k} = 20$.

Table 2
Greedy Algorithm versus FCFS

(a)	Greedy Algorithm	FCFS
$n \le 2\overline{k}$	$\langle b \rangle \approx n$ $\langle S \rangle \approx n/2$	$\langle b_r \rangle \approx n$ $\langle S_r \rangle \approx n/2$
$n \gg 2\overline{k}$	$\langle b \rangle = (-1 + \sqrt{1 + 8n\overline{k}})/2$ $\langle S \rangle \approx \overline{k}$	$\langle b_r \rangle \approx \overline{k}$ $\langle S_r \rangle \approx \overline{k}$
(b)	Greedy Algorithm	FCFS
$n \le 2N/M$	$\langle b \rangle \approx n$ $\langle S \rangle \approx n/2$	$\langle b_r \rangle \approx n$ $\langle S_r \rangle \approx n/2$
$n \gg 2N/M$	$\langle b \rangle = (-1 + \sqrt{1 + 8nN/M})/2$ $\langle S \rangle \approx N$	$\langle b_r \rangle \approx N/M$ $\langle S_r \rangle \approx N$

(a) Summary on the expected number of customers and the expected number of units being sold for both auction and the FCFS. (b) N and M (>1) correspond to the total number of units for sale and the $\max\{k_i\}$.

4. Greedy Algorithm versus FCFS

The results obtained in this section are summarized in Table 2. Without loss of generality, the results obtained in this paper can be extended for the case when k_i is uniformly distributed on the range $[0, M]$:

$$\begin{aligned} &\text{Maximize} && \textstyle\sum_{i=1}^{n} p_i s_i \\ &\text{Subject to} && \textstyle\sum_{i=1}^{n} k_i s_i \le N \\ & && s_i \in \{0,1\}\ \forall i = 1, \ldots, n \end{aligned}$$

Here N is the total number of resource available.

By comparing the number of clients being allocated with resources, it is found that there is no difference between the greedy algorithm or the FCFS when the number of customers is less than $2\overline{k}$. When $n > 2\overline{k}$, the greedy algorithm can allocate resource to more clients than the FCFS method. Obviously, the resources are allocated to those clients whose requested quantities are comparatively small. On the other hand, by comparing the expected quantity of resources being allocated, the FCFS can allocate more resources than the greedy algorithm, Fig. 5, irrespective to the number of customers n.

Suppose only the client who can get resources will have to pay and the service charge is defined as $P_0 + P_1 k_i, P_0, P_1 > 0$. The constant price P_0 can be interpreted as a premier that every client has to pay and P_1 can be interpreted as the unit resource price. The expected profit the WSP can gain by using the greedy algorithm and the FCFS, respectively, can be written as follows:

$$G = P_0 \langle b \rangle + P_1 \langle S \rangle \tag{20}$$
$$G_r = P_0 \langle b_r \rangle + P_1 \langle S_r \rangle \tag{21}$$

With reference to the numerical results (Fig. 5) obtained for $n = 50$ (i.e., $n = 2.5\overline{k}$), it is clear that the difference between $\langle S\rangle$ and $\langle S_r\rangle$ is about $0.013\times\overline{k}$ and the difference between $\langle b\rangle$ and $\langle b_r\rangle$ is about $0.2\times\overline{k}$. Therefore, the difference between G and G_r can be expressed as follows:

$$G - G_r = 0.2P_0\overline{k} - 0.013P_1\overline{k}$$

Accordingly, $G > G_r$ if $P_0/P_1 > 0.013/0.2$. The profit gain by using the greedy algorithm will be more than using the FCFS. If $P_0/P_1 < 0.013/0.2$, profit gain by using the FCFS will be more.

Of course, this comparison is only valid if the service charge model is linear. For other service charge models, a conclusion cannot easily be achieved. Numerical analysis will be needed.

5. Conclusion

In this paper, we have analyzed two properties of the profit density based greedy algorithm for a resource allocation problem in web service. The allocation problem is essentially a well-known knapsack problem. In terms of profit making, greedy algorithm can ensure a near-optimal solution. However, profit making is sometimes not the only consideration in management decision. Other properties such as the number of clients being allocated with resources and the quantity of resource being allocated are also important. In this regard, we have given a numerical analysis on these properties with respect to the profit density based greedy algorithm and the FCFS method.

The major tools that we used in the analysis are (i) the application of a formula derived by Weisberg in [4] for a linear combination of order statistics to analyze the greedy algorithm and (ii) the application of a formula derived by Feller in [5] for sum of uniform random variables to analyze the FCFS method. In accordance with the numerical results obtained, it is found that both the profit density based greedy algorithm and the FCFS method have very similar properties when the number of customers is not large, i.e., $n \le 2\overline{k}$. If $n = 2.5\overline{k}$, greedy algorithm has an advantage in letting more clients have resources allocated.

We have not concluded which algorithm is the best algorithm in resource allocation as resource allocation is itself a complicated problem, in particular when other management decisions are concerned. What we have presented here is simply additional remarks on profit density based greedy algorithm.

Appendix A Weisberg Formula: Linear Combination of Order Statistics

To analyze the expected $\sum_{i=1}^{b} k_{i:n}$, we apply the formula derived by Weisberg [4] for linear combination of order statistics. Let

$$S(n) = d_1U_{1:n} + \cdots + d_nU_{n:n} \tag{22}$$

where $U_{i:n}$ is the i-th order statistic drawn from $U[0,1]$ and d_is are real numbers. The probability for event $\{S(n) \le x\}$ is given by the following formula:

$$Pr\{S(n) \le x\} = 1 - \sum_{j=1}^{r} \frac{(c_j - x)^n}{c_j \prod_{j\neq i}(c_j - c_i)} \tag{23}$$

where c_is are given as follows:

$$c_{n+1} = 0 \quad c_k = c_{k+1} + d_k \tag{24}$$

$Pr\{S(n) < x\}$ is defined for all $0 \le x \le d_1 + d_2 + \cdots + d_n$ and r is the largest integer such that $x \le c_r$.

Illustrative example. Suppose $S(n) = U_{1:n} + U_{2:n} + U_{3:n}$. $d_1 = d_2 = d_3 = 1$ and $d_i = 0$ for all $i = 4, \ldots, n$. Then all the c_is will be given as follows:

$$c_1 = 3 \quad c_2 = 2 \quad c_3 = 1 \quad c_4 = 0 \quad \ldots \quad c_{n+1} = 0$$

The cdf can be written as the following equations:

$$\begin{array}{ll}
\forall\, x \ge 3 & Pr\{S(n) \le x\} = 1 \\
\forall\, 2 \le x < 3 & Pr\{S(n) \le x\} = 1 - \frac{(3-x)^n}{3\prod_{j\neq i}(3-c_i)} \\
\forall\, 1 \le x < 2 & Pr\{S(n) \le x\} = 1 - \frac{(3-x)^n}{3\prod_{j\neq i}(3-c_i)} \\
 & \qquad - \frac{(2-x)^n}{2\prod_{j\neq i}(2-c_i)} \\
\forall\, 0 \le x < 1 & Pr\{S(n) \le x\} = 1 - \frac{(3-x)^n}{3\prod_{j\neq i}(3-c_i)} \\
 & \qquad - \frac{(2-x)^n}{2\prod_{j\neq i}(2-c_i)} \\
 & \qquad - \frac{(1-x)^n}{1\prod_{j\neq i}(1-c_i)} \\
\forall\, x < 0 & Pr\{S(n) \le x\} = 0
\end{array}$$

Appendix B Feller Formula: Sum of n Uniform Random Variables

To analyze the expected L of the case that the products are sold in FCFS basis, we need the following formulae derived by Feller (p. 27 of [5]). Let $S_r(n)$ be the sum of uniform random variables defined as follows:

$$S_r(n) = U_1 + U_2 + \cdots + U_n \tag{25}$$

Noted that U_is are not ordered. For $n = 1, 2, \ldots$, and $0 \le x \le n$,

$$Pr\{S_r(n) \le x\} = \frac{1}{n!}\sum_{v=0}^{n}(-1)^v C_v^n (x-v)_+^n \tag{26}$$

where

$$x_+ = \frac{x + |x|}{2} \quad \text{and} \quad C_v^n = \frac{n!}{v!(n-v)!}$$

Note that for a point x between $(k-1)$ and k only k terms of the sum are different from zero.

Illustrative example. Let $S_r(3) = U_1 + U_2 + U_3$ and U_is are not in order, the cdf can be written as the following equations:

$$
\begin{array}{ll}
\forall\, x \geq 3 & Pr\{S_r(3) \leq x\} = 1 \\
\forall\, 2 \leq x < 3 & Pr\{S_r(3) \leq x\} = \frac{C_0^3 x^3 - C_1^3 (x-1)^3 + C_2^3 (x-2)^3}{3!} \\
\forall\, 1 \leq x < 2 & Pr\{S_r(3) \leq x\} = \frac{C_0^3 x^3 - C_1^3 (x-1)^3}{3!} \\
\forall\, 0 \leq x < 1 & Pr\{S_r(3) \leq x\} = \frac{C_0^3 x^3}{3!} \\
\forall\, x < 0 & Pr\{S_r(3) \leq x\} = 0
\end{array}
$$

Appendix C Derivation of (11) and (16)

For n is large and $n \gg \overline{k}$, we assume that the k_is are distributed evenly in $[0,1]$. Without loss of generality, we further assume $k_i < k_j$ if $i < j$. So, the values of k_is can be written as:

$$k_i = \frac{i}{n}$$

for all $i = 1, 2, \ldots, n$. As greedy allocation implies, the resource will be allocated to k_1, k_2 and so on until further allocation is not possible. Let $\langle r \rangle$ be the last one who can be allocated with resource. It turns out that

$$\sum_{i=1}^{\langle b \rangle} k_i \approx \overline{k}$$

It is equivalent to that

$$
\begin{aligned}
\sum_{i=1}^{\langle b \rangle} \frac{i}{n} &\approx \overline{k} \\
\frac{\langle b \rangle^2 + \langle b \rangle}{2n} &\approx \overline{k}
\end{aligned}
$$

The solution of $\langle b \rangle$ is thus approximately equal to $\frac{-1+\sqrt{1+8n\overline{k}}}{2}$. As $\langle R \rangle = (\langle b \rangle + 1)/2n$,

$$
\begin{aligned}
\langle S \rangle &\approx \overline{k} - \frac{\langle b \rangle + 1}{2n} \\
&= \overline{k} - \frac{1 + \sqrt{1+8n\overline{k}}}{4n} \\
&= \overline{k}\left(1 - \frac{1 + \sqrt{1+8n\overline{k}}}{4n\overline{k}}\right) \\
&= \overline{k}\left(1 - \frac{8n\overline{k}}{4n\overline{k}(\sqrt{1+8n\overline{k}} - 1)}\right) \\
&= \overline{k}\left(\frac{\sqrt{1+8n\overline{k}} - 3}{\sqrt{1+8n\overline{k}} - 1}\right)
\end{aligned}
$$

Whenever n is large, $\langle S \rangle \to \overline{k}$.

Acknowledgement

The authors would like to thank the reviewers for giving valuable comments on the earlier version of this paper.

References

[1] S. Martello & P. Toth, *Knapsack problems: Algorithms and computer implementations* (John Wiley & Sons, 1990).

[2] C.H. Papadamitrious & K. Steiglitz, *Combinatorial optimization: Algorithms and complexity* (Dover, 1998).

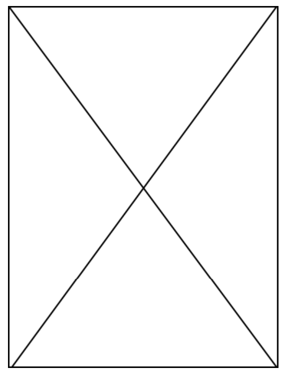
[3] E.S. Maskin & J.G. Riley, Optimal multi-unit auctions, in F. Hahn (ed.), *The economics of missing markets, information, and games* (Oxford: Oxford University Press, 1989).

[4] H. Weisberg, The distribution of linear combinations of order statistics from the uniform distribution, *Annals of Mathematical Statistics, 42*, 1971, 704–709.

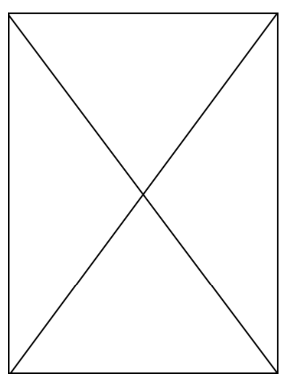
[5] W. Feller, *Probability theory: Volume II* (Wiley, 1971).

[6] B. Lehmann, D. Lehmann, & N. Nisan, *Combinatorial auctions with decreasing marginal utilities, EC'01* (Tampa, FL, 2001).

Biographies

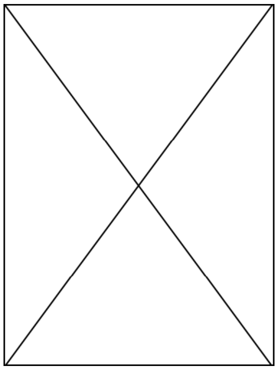


John Sum received the B.Eng. in Electronic Engineering from Hong Kong Polytechnic University in 1992. Then, he received his M.Phil. and Ph.D. in Computer Science and Engineering from Chinese University of Hong Kong in 1995 and 1998, respectively. Before joining the Chung Shan Medical University (Taiwan), John was teaching in several universities in Hong Kong, including the Hong Kong Baptist University, the Open University of Hong Kong and the Hong Kong Polytechnic University. Currently, he is an Assistant Professor of the Department of Information Management in the Chung Shan Medical University, Taichung ROC. His current research interests include neural computation, mobile sensor networks and scale-free network. John Sum is a Senior Member of IEEE and an Associate Editor of the *International Journal of Computers and Applications.*



Jie Wu is a Distinguished Professor at the Department of Computer Science and Engineering, Florida Atlantic University. He has published over 350 papers in various journals and conference proceedings. His research interests are in the areas of wireless networks and mobile computing, routing protocols, fault-tolerant computing and interconnection networks. Dr. Wu was on the Editorial Board of IEEE Transactions on Parallel and Distributed Systems and was a co-guest-editor of IEEE Computer and *Journal of Parallel and Distributed Computing.* He served as the program co-chair for MASS 2004, program vice-chair for ICDCS 2001 and program vice-chair for ICPP 2000. He was also general co-chair for MASS 2006 and is general vice-chair for IPDPS 2007. He is the author of the text "*Distributed system design*" published by the CRC press. He was also the recipient of the 1996–1997 and 2001–2002 Researcher of the Year Award at Florida Atlantic University.

Dr. Wu has served as an IEEE Computer Society Distinguished Visitor and is the Chairman of IEEE Technical Committee on Distributed Processing (TCDP). He is a Member of ACM and a Senior Member of IEEE.



Chi-Sing Leung received his B.Sci. in Electronics from the Chinese University of Hong Kong in 1989. He received his M.Phil. degree in Information Engineering and Ph.D. in Computer Science from the same university in 1991 and 1995, respectively. He is currently an Associate Professor in the Department of Electronic Engineering, City University of Hong Kong. His research interests include neural computing, data mining and computer graphics. He has published over 60 international journal papers and 4 book chapters. He received the IEEE Transactions on Multimedia 2005 Prize Paper Award for the paper titled, "The plenoptic illumination function" in *IEEE Transactions on Multimedia*, Vol. 4, Issue 3, September 2002.