On Unbounded Zero-One Knapsack with Discrete-Sized Objects

John Sum[1], Jie Wu[2] and Kevin Ho[3]
[1]Department of Information Management, Chung Shan Medical University
Taichung 402, Taiwan
[2]Department of Computer Science and Engineering, Florida Atlantic University
Boca Raton, Florida 33431, USA
[3]Department of Computer Science and Communication Engineering
Providence University, Sha Lu, Taiwan

Abstract

The problem being discussed in this paper is a special case of the unbounded knapsack problem :

$$\max \quad z_n(M) = \frac{1}{n}\sum_{i=1}^{n} p_i x_i$$
$$\text{s.t.} \quad \sum_{i=1}^{n} c_i x_i \leq \beta_0 n$$
$$x_i \in \{0,1\} \ \forall i = 1, \ldots, n;$$

where p_i's are uniformly distributed random variables in $[0,1]$, c_is are discrete random variables distributed uniformly in $\{1/M, 2/M, \ldots, (M-1)/M, 1\}$ and $z_n(M)$ is the optimal objective function value. Assuming that M is large, $z_n(M)$ approximately equals to $\sqrt{\frac{2\beta_0}{3}}(1 - 0.3062(\sqrt{\beta_0}M)^{-1})$. An application of this formula is exemplified by an auction problem.

1 Introduction

The unbounded knapsack problem can be stated as the following combinatorial optimization problem :

$$\max \quad z_n = \frac{1}{n}\sum_{i=1}^{n} p_i x_i$$
$$\text{s.t.} \quad \sum_{i=1}^{n} c_i x_i \leq \beta_0 n \tag{1}$$
$$x_i \in \{0,1\} \ \forall i = 1, \ldots, n.$$

By making assumptions on the statistical distribution for p_is and c_is as $f(p,c)$ and using Central Limit Theorem[1], it has been shown in [1, 3, 4] that the optimal z_n (when $n \to \infty$) is given by

$$\int_0^{1/m^*} \int_{m^*c}^{1} p f(p,c) \, dp \, dc \tag{2}$$

where m^* is obtained by the following equation :

$$\int_0^{1/m^*} \int_{m^*c}^{1} c f(p,c) dp \, dc = \beta_0, \tag{3}$$

for $n \to \infty$ and $\beta_0 \leq 0.5$.

---

[1]The essential idea is due to the relation that

$$\lim_{n \to \infty} \frac{1}{n}\sum_{i=1}^{n} c_i x_i = \int_0^{1/m^*} \int_{m^*c}^{1} c f(p,c) dp \, dc,$$

where

$$x_i(p_i, c_i) = \begin{cases} 1 & \text{if } p_i \geq m^* c_i \\ 0 & \text{otherwise.} \end{cases}$$

In this short paper, we extend the work in [4] and [3] simply by considering c_is are discrete random variables uniformly distributed in $\{1/M, 2/M, \ldots, (M-1)/M, 1\}$. In the next section, the motivation of the study and its relation with the multi-unit combinatorial auction will be presented. The approximated profit gain for the unbounded zero-one knapsack with discrete-sized objects will be derived in Section 3. An application of the approximated profit gain in a notebook auction problem will be illustrated in Section 4. Section 5 will discuss certain issues regarding the situation where the object size follows Beta distribution and the price function is a deterministic function. Finally, the conclusion of the paper will be presented in Section 6.

2 Background

Multi-units combinatorial auction (CA) has recently been researched within the e-commerce community for selling products [2], and extended to the solve problems in web services resource allocation [6, 7]. This auction mechanism is especially designed for selling a collection of products or services. In contrast with the conventional English auction in which products can only be auction off one item or one unit at a time, CA allows bidder to bid for a subset of collection.

2.1 A problem in using CA

In theory, Multi-units CA is an efficient mechanism since it can ensure the profit of *a collection of products* being sold is maximum. It is because the allocation mechanism for multi-units CA is simply a combinatorial optimization problem. However, there are at least two reasons leading CA not yet commonly used in the auction industry.

The first reason is due to the fact that *the expected profit being gained by using CA cannot be pre-computed in advance*. Without the estimation on the profit return, seller can hardly compare from their anticipated profit the efficient of different selling mechanisms. Eventually, seller usually prefers other selling mechanism instead of CA. Nevertheless, without the estimation on the profit return, auctioneer is unlikely to determine the management fee, by commission for instance, for organizing such an auction.

The second reason is due to a logistic consideration. Since it is more efficient if products are packed and delivered in a size of multiple units at a time, restricting the bid size on discrete levels, say $\{20, 40, \cdots, 1000\}$, might reduce the total delivery cost and hence increase the net profit. Therefore, a seller would like to know *the expected profit if the bid sizes are restricted to discrete levels and hence decide the allowable bid sizes for the bidders*. However, there is not much information being provided for such anticipation.

2.2 Profit gain for continuous c_i's [3, 4]

One useful result is from [3, 4] regarding the size of knapsack is proportion to the number of items and the number of items is large. In this regard, the profits being gained by using either dynamic programming or profit-density based greedy algorithm will be the same. And the expected profit can thus be *asymptotically* expressed in terms of the joint probability distribution of the object size (c) and object price (p), $f(p, c)$, [3, 4].

Still, the expected profit gain can only be expressed analytically, Equation (3) and Equation (2). Obtaining the exact values for m^* and z_n, one needs to solve Equation (3) and Equation (2) numerically.

One special case is when both p_is and c_i's are continuous random variables distributed uniformly in $[0, 1]$, i.e. $f(p, c) = 1 \forall (p, c) \in [0, 1]^2$, the optimal solution (expected profit) $\lim_{n \to \infty} \hat{z}_n$ can be written as follows [3] :

$$m^* = \sqrt{\frac{1}{6\beta_0}} \tag{4}$$

$$\lim_{n \to \infty} \hat{z}_n = \sqrt{\frac{2\beta_0}{3}} \tag{5}$$

for $\beta_0 \in (0, 0.5]$.

2.3 Practical situation

The usefulness of the above result relies on the assumption that the size of the object can be in any precision, i.e. product is divisible. As it is commonly known that many consumer products, like notebook and air cargo, are not divisible. Even a product can be divisible, it might be more cost effective (due to logistic consideration) if only discrete sizes are allowed.

Let us have a simple example. Suppose an auctioneer has 1 000 000 units of notebook computers for auction off. The auctioneer anticipates that 2 000 bidders will come. Moreover, the auctioneer also expects that their bid sizes are uniformly random in $\{1, 2, 1\,000\}$ and the bid price is random variables uniformly in $\{100, 200, \cdots, 100\,000\}$. Accordingly, the average profit gain is estimated as follows :

$$
\begin{aligned}
2\,000 \times 100\,000 \times \hat{z}_{2000} &\approx 2\,000 \times 100\,000 \times \sqrt{\frac{2(0.5)}{3}} \\
&= 115.47 \times 10^6.
\end{aligned}
$$

Since the quantized levels in both bid size and bid price are small, 0.001 and 0.001 respectively, the approximation is close to the actual value. However, it will be inaccurate if the quantized level is not small.

3 Average profit gain : Discrete c_i's

Let M be the maximum size that a bidder will bid. The actual inequality constraint will be expressed as follows :

$$
\frac{1}{n} \sum_{i=1}^{n} c_i M x_i \leq \beta_0 M. \tag{6}
$$

Here $c_i M \in \{1, 2, \ldots, M\}$, is the actual bid size for the ith bidder. Let $\lim_{n \to \infty} \hat{z}_n(M)$ be the optimal average profit gain,

$$
\lim_{n \to \infty} \hat{z}_n(M) \leq \lim_{n \to \infty} \hat{z}_n.
$$

Equality holds when $M \to \infty$. Besides, let us define the percentage error, $E(M)$, as follows :

$$
E(M) = \frac{\lim_{n \to \infty} \hat{z}_n - \lim_{n \to \infty} \hat{z}_n(M)}{\lim_{n \to \infty} \hat{z}_n}. \tag{7}
$$

We now consider the case that the probability distribution for the bid size c, $P(c)$, is uniformly distributed in the set $\{1/M, 2/M, \ldots, 1\}$, that is,

$$
\begin{aligned}
P(c) &= \sum_{k=1}^{M} \frac{1}{M} \delta(c, k/M) \tag{8} \\
\delta(c, k/M) &= \begin{cases} 1 & \text{if } c = k/M \\ 0 & \text{otherwise.} \end{cases}
\end{aligned}
$$

In practice, we assume that each bidder whose original bid size is in the range $[j/M, (j+1)/M]$ ($j = 0, 1, \ldots, M-1$) will bid up to $(j+1)/M$. Therefore, the equality constraint in Equation (3) can be written as follows :

$$
\sum_{i=1}^{k^*} \left(1 - i\frac{m^*}{M}\right) \left(\frac{i}{M^2}\right) = \beta_0 \tag{9}
$$

where m^* is the slope of the straight line for the decision boundary, see Figure 1, and m^* and k^* can be related by the following inequality.

$$
\frac{k^*}{M} < \frac{1}{m^*} < \frac{k^*+1}{M}.
$$

Solving Equation (9), it is able to obtain

$$
\frac{k^*(k^*+1)}{2} - \frac{m^*}{M}\frac{k^*(k^*+1)(2k^*+1)}{6} = \beta_0 M^2. \tag{10}
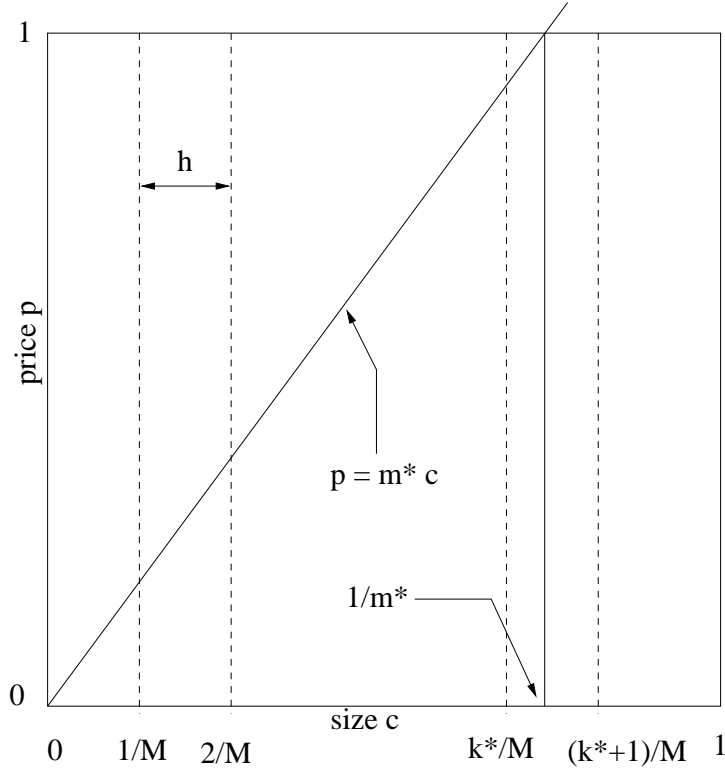$$

Figure 1: The straight line $p = m^*c$ is the decision boundary for product allocation. All the bids (p, c) above this line will be allocated.

Since the number of bidders n and the number of quantized intervals M are large, it can be assumed that $k^* \gg 1$ and $m^*/M \approx 1/k^*$. Hence,

$$k^* \approx \sqrt{6\beta_0}\, M. \tag{11}$$

The average profit gain $\lim_{n\to\infty} \hat{z}_n(M)$ can be written as follows :

$$
\begin{aligned}
\lim_{n\to\infty} \hat{z}_n(M) &= \sum_{i=1}^{k^*} \frac{1}{M} \int_{i/k^*}^{1} p\, dp \\
&= \frac{1}{2M}k^* - \frac{1}{6M}k^* - \frac{3}{12M} - \frac{1}{12k^*M} \\
&\approx \frac{1}{3M}k^* - \frac{1}{4M}.
\end{aligned}
$$

Since $k^* \approx \sqrt{6\beta_0}\, M$,

$$
\begin{aligned}
\lim_{n\to\infty} \hat{z}_n(M) &\approx \sqrt{\frac{2\beta_0}{3}} - \frac{1}{4M}; \\
E(M) &\approx \frac{1}{4M}\sqrt{\frac{3}{2\beta_0}}.
\end{aligned}
$$

Equation (7) can thus be approximated by the following equation.

$$E(M) \approx \sqrt{\frac{3}{32}}\frac{h}{\sqrt{\beta_0}} = 0.3062\frac{1}{\sqrt{\beta_0}M}. \tag{12}$$

Figure 2 shows the cases when β_0 equals to 0.2 and 0.5 respectively. M is taken from 40 up since the approximation does not hold for small M. It is found that the percentage error is less than 1% even when the number of intervals is only 70.
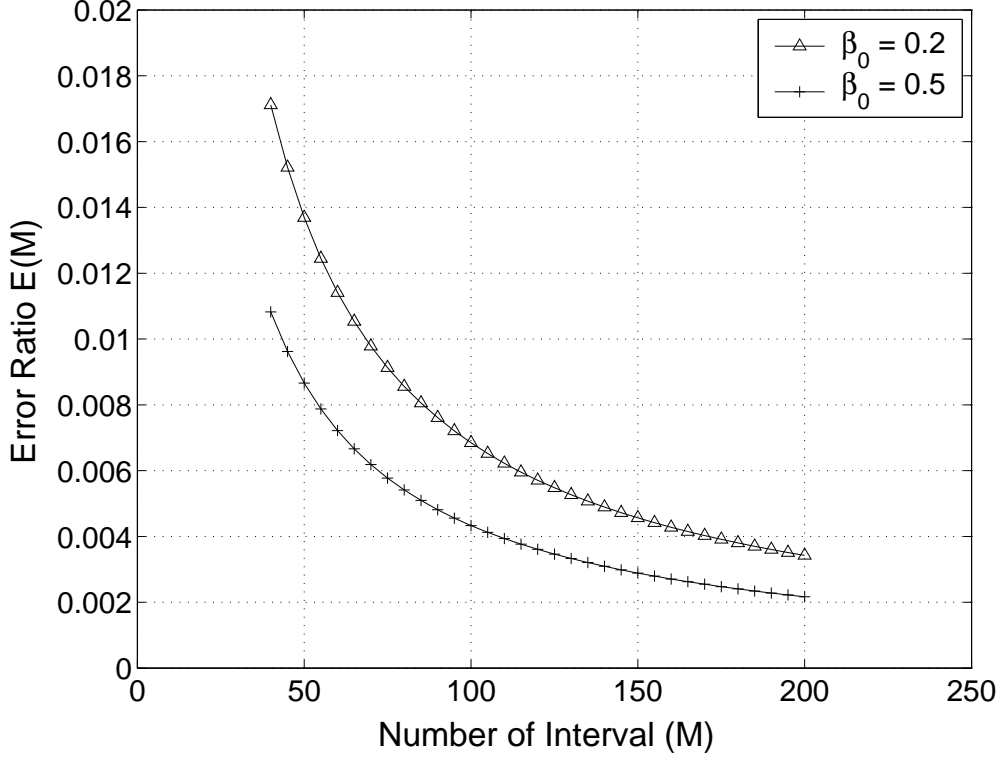
4

Figure 2: The error ratio $E(M)$ against the number of interval M.

4 Profit gain for notebook auction

In the illustrative example mentioned before, we roughly estimated the average profit gain, $\lim_{n\to\infty} \lim_{M\to\infty} \hat{z}(M)$, as $\sqrt{\frac{2(0.5)}{3}}$. In accordance with Equation (12), the percentage error can thus be obtained.

$$E(M) = 0.3062 \frac{1}{\sqrt{0.5}\ 1000} = 4.33 \times 10^{-4}$$

which is less than 0.05%. A better estimation for the average profit gain can also be evaluated as follows :

$$\begin{aligned} \lim_{n\to\infty} \hat{z}(M = 1000) &= \left(1 - 4.33 \times 10^{-4}\right) \sqrt{\frac{2(0.5)}{3}} \\ &= 0.5771 \end{aligned}$$

Using Equation (12), we can determine the maximum integer of M such that the percentage error is less than certain threshold η.

$$M \approx 0.3062 \frac{1}{\sqrt{\beta_0}\ \eta}.$$

For $\eta = 0.01$ and $\beta_0 = 0.2$, it can readily be shown that $M \approx 69$. For $\eta = 0.01$ and $\beta_0 = 0.5$, $M \approx 44$. For M equals to 100 and β_0 is in $[0.2, 0.5]$, the percentage error is less than 1%.

Therefore, the profit gain of auction off 1 000 000 notebook computers by using multi-units CA can be estimated. Anticipating that there are 2000 bidders and their bid sizes are uniformly random in $[1, 1\,000]$. If the bid size is restricted to be $\{20, 40, \cdots, 1000\}$, the expected profit gain will be $0.99 \times 115.47 \times 10^6$, i.e. 114.3×10^6. If the bid size is restricted to $\{50, 100, \cdots, 1000\}$, the profit will be $0.978 \times 115.47 \times 10^6$, i.e. 113×10^6.
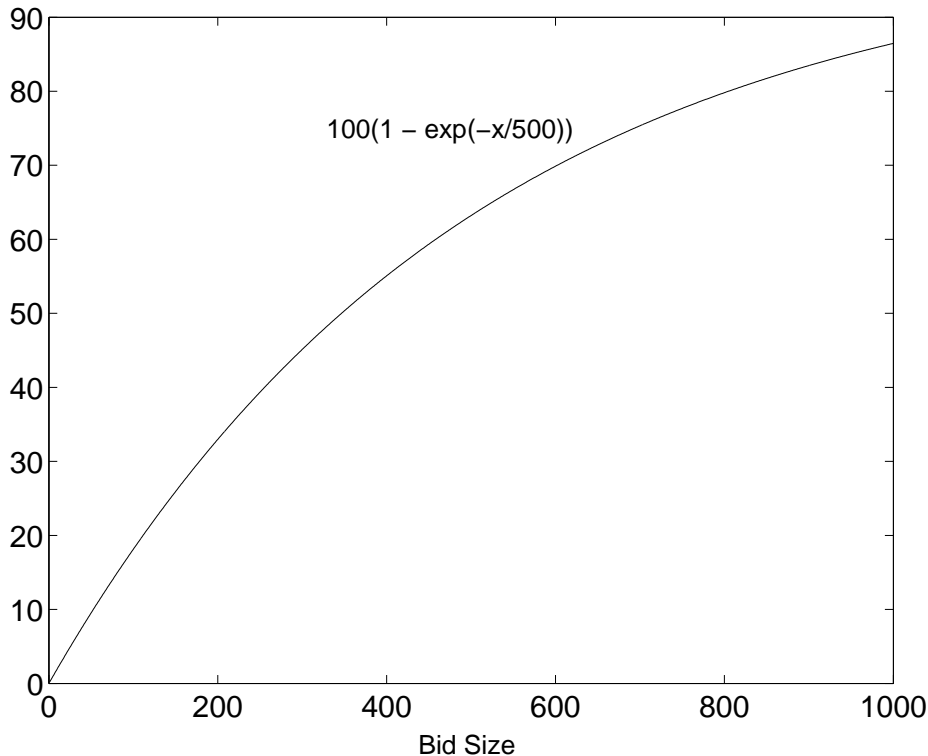
Figure 3: Deterministic bid price.

5 Discussion

Interesting enough, based on simulation result, we have found that the profit error can be rather small for some special distributions for bid size and bid price. For example, we have attempted the situation when the bid price is a deterministic function (Figure 3) :

$$p(c) = 100 \times (1 - \exp(-c/500)),$$

for $c \in [1, 1000]$. This mimics the volume discount behavior, i.e. the decreasing marginal utility behavior [2]. The bid size is modeled as a Beta distribution defined as follows :

$$Prob(c) = \frac{c^a(1000 - c)^b}{\int_0^{1000} c^a(1000 - c)^b dc},$$

where the parameter (a, b) are set to three cases: $(5, 5)$, $(5, 15)$ and $(15, 5)$, Figure 4. We assume that there are 10000 bidders and the range for c is $[1, 1000]$. β_0 is set to be 0.2. It is found that no matter the discrete level M is set to $10, 20$ and 50, the profit error is still less than 1%. If we assume the distribution is uniform and the percentage error is estimated by using Equation (12), the profit error are 6.8%, 3.4% and 1.4% respectively. Therefore, without much information about the price distribution and the object size distribution, Equation (12) can be used to determine the value M.

6 Conclusion

While combinatorial auction has been claimed to be an efficient algorithm for selling products, many sellers are still reluctant to adopt it as an alternative selling mechanism. In this paper, we have discussed a number of practical issues regarding to the application of this new type of auction and the profit gain estimation is one of them. Extended from the work done by Lueker [3], the average profit gain for the unbounded
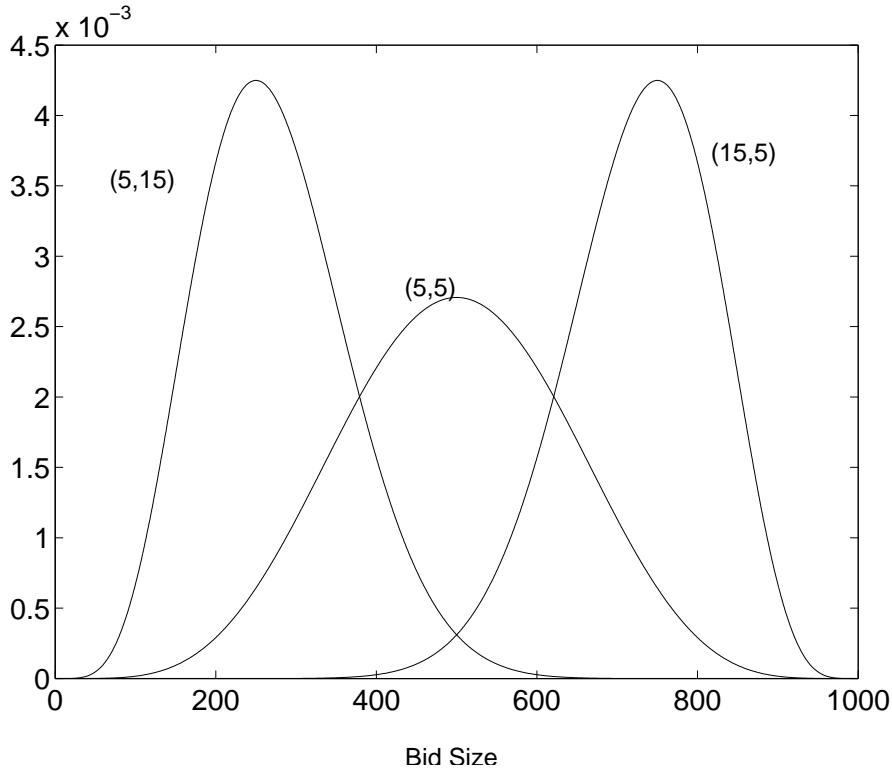
Figure 4: Beta distribution for bid sizes.

knapsack problem with discrete-sized objects has been obtained. It has been shown that the average discrete-sized profit gain can be expressed in term of β_0 and M as $\sqrt{\frac{2\beta_0}{3}}(1 - 0.3062(\sqrt{\beta_0}M)^{-1})$. Since multi-units combinatorial auction is essentially the same as a knapsack problem, we applied the formula to determine the number of allowable bid sizes for a notebook auction problem. Noted that the bid price and the bid size distributions are normally not uniform, a situation where the bid price is a deterministic function of the bid size and the bid size follows Beta distribution has been investigated. Though, no close formula has been obtained to estimate the profit gain, the factor $\sqrt{\frac{2\beta_0}{3}}(1 - 0.3062(\sqrt{\beta_0}M)^{-1})$ can still be used as an upper error bound. For the bid size and bid price following other distributions, numerical algorithm would be needed for obtaining the expected profit and/or the profit error term. Research along this line is valuable for future work.

References

[1] Lai T-C, Worst-case analysis of greedy algorithms for the unbounded knapsack, subset-sum and partition problems, *Operations Research Letters*, Vol.14:215-220, 1993.

[2] Lehmann B., D. Lehmann and N. Nisan, Combinatorial auctions with decreasing marginal utilities, EC'01, Tampa, Florida, 2001.

[3] Lueker G.S., Average-case analysis of off-line and on-line knapsack problems, *Journal of Algorithms*, Vol.29, 277-305, 1998.

[4] Marchetti-Spaccamela A. and C. Vercellis, Stochastic on-line knapsack problems, *Mathematical Programming*, Vol.68:73-104, 1995.

[5] Maskin E.S. and J.G. Riley, Optimal multi-unit auctions, *The Economics of Missing Markets, Information, and Games*, Oxford University Press, 1989.

[6] John Sum *et al*, On Profit Density Based Greedy Algorithm for a Resource Allocation Problem in Web Services, accepted for publication in *International Journal of Computers and Applications*.

[7] John Sum, Kevin Ho, G. Young, Pricing Web Services, in Y.-C. Chung and J.E. Moreira (Eds.) *Advances in Grid & Pervasive Computing*, p.147-156, 2006. Spring-Verlag Berlin Heidelberg.