数理解析研究所講究録
1489巻 2006年 50-56

オンラインオークション型資源配分問題

東北大学大学院 情報科学研究科　原田 薫明, 瀧本 英二, 丸岡 章
Graduate School of Information Sciences,
Tohoku University

概 要

本論文では，次々とやってくる入札者に商品を売るにあたり値段をオンライン的に決定するオークショニアの利益を最大化するという，オンラインオークション問題を取り上げる．このオンラインオークション問題は，デジタル商品の価格決定を想定しており，Bar-Yossef らによって初めて設計された [BHW02]．この問題に対する結果として，オンライン資源配分問題へと帰着することにより，ヘッジアルゴリズムを適用した結果 [BKRW03] や，仮想的なスコアを利用する HG アルゴリズム [BH05] などが知られる．本論文では，Bar-Yossef らのオークション問題から帰着した資源配分問題は，通常の資源配分問題とは異なる特徴を持つ問題であることを指摘し，Vovk の統合戦略 [V98] に基づいてオークション問題に対するオークション型統合アルゴリズムを新しく導出した．

1 はじめに

オンラインオークション問題とは，次々とやってくる入札者に商品をひとつずつ売るにあたり，一回の取引における商品の値段をオンライン的に決定するオークショニアを考え，オークショニアの利益を最大化するという問題である．この問題は，品質劣化のない複製を作るのが容易なデジタル商品におけるオークションを想定して，Bar-Yossef らによって初めて設計された [BHW02] ものである．

これから，本論文で扱う Bar-Yossef らのオンラインオークション問題について，t 回目の入札者が訪れたときのやり取りを説明する．まず，オークショニアは過去の入札額 $m_1, \ldots, m_{t-1}$ を参考に商品の値段 r_t を決定する．また，オークショニアは，商品の最低落札価格を 1 としたとき，入札者の入札額の最高額が高々 h であることを予め知っているものとする．なお，入札者も最低落札価格以上の入札をするものとする．すなわち，任意の時刻 t について $1 \leq m_t \leq h$ である．よって，オークショニアが選ぶ商品の値段も $1 \leq r_t \leq h$ である．

オークショニアが商品の値段 r_t を提示した後，入札者は自分の入札額 m_t を告げる．オークショニアのつけた値段 r_t が入札者の入札額 m_t 以下ならば，入札者は商品を購入していく．よって，オークショニアの利得は商品の値段が入札額以下のとき ($r_t \leq m_t$ のとき) のみ発生し，このときのオークショニアが受け取る利得を $g_{A,t}$ であらわせば，$g_{A,t} = r_t$ である．一方，入札者の入札額を商品の値段が上回った場合 ($r_t > m_t$ のとき) は商品は売れず，オークショニアは利得を受け取ることができない ($g_{A,t} = 0$)．もちろんこの問題では，オークショニアは t 回目の入札を見てからそのときの商品の値段 r_t を変更することを許されない．そして，入札者も正直な入札額を告げるとし，オークショニアが提示した値段 r_t を見て入札額を調整することはないと仮定する．

この問題の形式的なプロトコルは次のようになっている．

Bar-Yossef のオンラインオークション問題　各入札者 $t = 1, 2, \ldots$ について，

1. 学習者は過去の入札履歴を参考に，値段 $r_t \in [1, h]$ を決定する．
2. 学習者に入札金額 m_t が明かされる．
3. 学習者は $r_t \leq m_t$ のとき利得 $g_{A,t} = r_t$ を得ることができるが，そうでなければ利得を得られない ($g_{A,t} = 0$)．

この問題の目標は，時刻 T までの累積利得 $G_{A,T} = \sum_{t=1}^{T} g_{A,t}$ を，“最適な価格で売り続けたときの利得” に匹敵させることである．これから最適な価格で売り続けたときの利得を最大利得と呼び，オンラインオークション問題における最大利得を定義する．

今，時刻 T までの入札額の履歴を $(m_1, \ldots, m_T)$ とする．このとき，この入札の履歴を入札額の大きい順に並べ替え，$(m_{\pi(1)}, \ldots, m_{\pi(T)})$ とあらわすことにする．ここで，$\pi : \{1, 2, \ldots, T\} \mapsto \{1, 2, \ldots, T\}$ は入札額を大きい順に並べ替える置換関数とする．今，オンラインオークション問題の特性として，入札者の入札額以下の値段を提示した場合は必ず商品は売れるので，T 人の入札者全員に対して商品の値段を $m_{\pi(k)}$ と提示した場合は k 回売れることになる．したがって，オンラインオークション問題における最大利得は次のように定義される．

定義 1 (最大利得) T 人の入札価格を大きい順に $(m_{\pi(1)},\ldots,m_{\pi(T)})$ とあらわしたとき，最大利得 OPT は

$$\mathrm{OPT} = \max_{1\le k\le T} k m_{\pi(k)}$$

で与えられる．また，この最大利得を達成することのできる値段を m_{OPT} とあらわす．

2 オンライン資源配分問題への帰着

オンラインオークション問題はオンライン資源配分問題へ還元できる．ここではまず，オンライン資源配分問題について説明する．オンライン資源配分問題は，毎時刻予測を出力する N 人のエキスパート，それらを統合して自分の予測を行うアルゴリズム (学習者)，そして，結果を表示する環境の間で行なわれる繰り返しゲームとして形式化される．

今，P を予測の集合，Y を結果の集合とし，予測の良さを評価するための利得関数 $\lambda : P\times Y \mapsto [0,\infty]$ が与えられているとする．このとき，オンライン資源配分問題では，各時刻 $t=1,2,\ldots$ について次のようなやり取りが行なわれる．

1. 各エキスパート $i\in\{1,\ldots,N\}$ は予測 $x_{i,t}\in P$ を行なう．
2. 学習者は 1. に基づき，自分の予測 $r_t\in P$ を行なう．
3. 環境は結果 $y_t\in Y$ を公開する．
4. 学習者は利得 $\lambda(y_t,p_t)$ を，各エキスパート i は利得 $\lambda(y_t,x_{i,t})$ を受け取る．

このような一連の試行を時刻 T まで行なったとき，学習者 A の累積利得は $G_{A,T}=\sum_{t=1}^{T}\lambda(y_t,p_t)$ で与えられ，同様に，エキスパート i の累積利得は $G_{i,T}=\sum_{t=1}^{T}\lambda(y_t,x_{i,t})$ で与えられる．資源配分問題における目標は，アルゴリズムの累積利得を，最良のエキスパートの累積利得 $\max_{1\le i\le N} G_{i,T}$ に匹敵させることである．

さて，オンラインオークション問題は次のようにオンライン資源配分問題へと帰着される．まず，学習者，各エキスパートの予測値を離散的な値とするため，N 個の値段の集合 R を $R=\{b(1),b(2),\ldots,b(N)\}$ と定義する．ここで，一般性を失わず，$h\ge b(1)\ge b(2)\ge\cdots\ge b(N)=1$ とする．このとき，予測の集合 P は R 上の N 次元確率ベクトル空間とする．各エキスパート i は，予測として常に同じ値段を支持するとする．すなわち，エキスパート i の予測は $x_{i,t}(i)=x_i(i)=1$，$x_{i,t}(j\neq i)=x_i(j\neq i)=0$ からなるベクトル $\boldsymbol{x}_i\in P$ とする．また，学習者は時刻 t の値段 r_t を決定するときは，確率分布 $\boldsymbol{p}_t\in P$ に従いエキスパートを選ぶことによって間接的に行なわれるとする．結果の集合 Y は N 次元空間で，$\boldsymbol{y}_t\in Y$ の第 i 成分は $y_t(i)\in\{0,b(i)\}$ である．ところで，オークション問題では学習者の利得は $\boldsymbol{p}_t$ に従う確率的なものであるから，学習者のの利得の期待値で評価が行なわれる．よって，利得関数は確率分布 p_t による期待値として内積により $\lambda(y_t,p_t)=\sum_{i=1}^{N}\boldsymbol{p}_t\cdot\boldsymbol{y}_t$ と定義される．また，期待累積利得を $EG_{A,T}=\sum_{t=1}^{T}\lambda(y_t,p_t)$ とする．

ところで，値段の設定法もアルゴリズムの性能に大きく影響する．値段の決め方のひとつの方法として，値段を等比数列的に定める手法が知られている [BKRW03]，[BH05]．今，$\rho>1$，$N=\lfloor\log_\rho h\rfloor+1$ とし，各値段を $b(i)=\rho^{N-i},(i=1,\ldots,N)$ とする．すなわち，値段の集合を $R=\{\rho^{N-1}(\le h),\ldots,\rho^2,\rho,1\}$ とする．このように R を設定することで，ある i^* で $b(i^*)\le m_{\mathrm{OPT}}\le b(i^*-1)=\rho b(i^*)$ となり，この値段 $b(i^*)$ を支持するエキスパート i^* の累積利得は，

$$G_{i^*}\ge\frac{\mathrm{OPT}}{\rho}$$

が保証される．

このようにしてオンライン資源配分問題として還元されたオンラインオークション問題に，ヘッジアルゴリズムを適用した場合，アルゴリズムの期待累積利得は，

$$EG_{\mathrm{Hedge},T}\ge\frac{\ln\alpha}{\alpha-1}\left(\frac{\mathrm{OPT}}{\rho}-\frac{\rho^{\lfloor\log_\rho h\rfloor+1}}{\ln\alpha}\ln\left(\lfloor\log_\rho h\rfloor+1\right)\right)$$

という下界となる [BKRW03]．

さらに，このヘッジアルゴリズムにおける下界の，付加損失項の h に対する影響を改善したとしてHGアルゴリズムが示されている [BH05]．このアルゴリズムは学習者の判断にランダム性を与えるために，ノイズベクトルとして仮想的な利得 (Hallucinated-Gain)$\boldsymbol{h}$ を扱う．これは，エキスパート i が推薦している値段を r_i とするとき，$\delta \in (0,1)$ として $(1-\delta)^k\delta$ の確率で $h_i = kr_i$ となるように設定される．そして，時刻 t までのエキスパート i の累積利得を $g_{i,t}$ とあらわすと，アルゴリズムはスコア s_i と呼ばれる実際の累積利得と仮想的な利得の和 $s_i = g_{i,t} + h_i$ に基づき，最も高いスコアを示すエキスパートの値段を商品に設定する．

このようなシンプルなアイディアに基づいているのにもかかわらず，HGアルゴリズムの期待累積利得の下界は，$\nu(\rho) = \lfloor \log_\rho 2 \rfloor + 1$ として，

$$EG_{\mathrm{HG},T} \geq (1-\delta)\left(\frac{\mathrm{OPT}}{\rho} - 2h\left(\frac{2}{\delta}\ln\nu(\rho) + \frac{\nu(\rho)}{\delta^2}(1-\delta)^{\nu(\rho)} + 1\right)\right).$$

となる．これは，ヘッジアルゴリズムと比較すると，h に対する性能は，付加損失項の $\ln\ln h$ の影響分だけ優れているといえる．

このように還元されたBar-Yossefのオンラインオークション問題であるが，結果の与えられ方に何の仮定もおかない通常のオンライン資源配分問題と若干モデルが異なっている．このオンラインオークション問題は，時刻 t の入札者の入札額を m_t としたとき，入札額以下の値段 $(b(i) \leq m_t)$ を支持するエキスパートは皆，それぞれ $b(i)$ の報酬を得ることができるという問題設計がなされている．これは，エキスパートへの報酬の与えられ方が独立である通常の資源配分問題とは異なっている．よって本論文では，このような性質を持つ問題をオークション型資源配分問題と呼ぶ．この性質を考慮したアルゴリズムは，通常のオンライン資源配分問題に対するアルゴリズムを素直に適用するよりも，優れた性能評価が与えられることがわかった．なお，HGアルゴリズムもオークション型資源配分問題の特徴を考慮したアルゴリズムのひとつといえる．本論文では，オークション型資源配分問題にVovkの統合戦略 [V98] を適用し，オークション型の問題における新しいアルゴリズムを導出する．

3 オークション型統合アルゴリズム

3.1 Vovkの統合戦略

Vovkの統合戦略 (Aggregating Strategy) は，幅広い問題に適用できるメタ戦略としてVovkによって与えられた．この戦略の特徴は，次にどんな結果が観測されようとも，エキスパートの予想との相対損失がある一定の値以下となるよう予測を行うことにある．この戦略に基づくアルゴリズムの累積損失上界は，ある緩やかな条件の下では最適であることが示されている [V98]．これをオンライン資源配分問題に適用すると，ヘッジアルゴリズムの累積損失上界よりも小さな累積損失上界が保証される．Vovkの統合戦略に基づくアルゴリズムをVovkの統合アルゴリズム (Aggregating Algorithm, AA) と呼ぶ．

Vovkの統合戦略では，各時刻 t において，各エキスパート i に重み $v_{i,t}$ を次のように割り当てる．

$$v_{i,t} = \frac{v_{i,1}\alpha^{G_{i,t-1}}}{\sum_{j=1}^{N} v_{j,1}\alpha^{G_{j,t-1}}} \tag{1}$$

ここで，$G_{i,t-1} = \lambda(y_t, x_{i,t-1})$ は，エキスパート i のそれまでの累積利得で，$alpha > 1$ は学習定数と呼ばれるパラメータである．$v_{i,1}$ は各エキスパート i に対する初期重みで，エキスパート i の信頼度に応じて任意につけることができるが，通常は一様な重み $(1/N)$ を割り当てる．

AAは，エキスパートの予測 $\boldsymbol{x}_t = (x_{1,t}, \ldots, x_{N,t})$ が与えられたとき，

$$\forall y \in Y,\ \ \lambda(y, p_t) \geq c(\alpha)\log_\alpha \sum_{i=1}^{N} v_{i,t}\alpha^{\lambda(y, x_{i,t})} \tag{2}$$

を満たす p_t を出力する．ここで，$c(\alpha)$ は予測ゲームと α にのみ依存する実数で，任意のエキスパートの重みと予測値に対し，式 (2) を満たす p_t を保証する最小の実数として定義される．

Vovkの統合戦略の累積利得下界は $c(\alpha)$ を用いて次の定理のようにあらわされる．

定理 1 ([V98]) 任意のゲーム終了時刻 T, 任意の結果系列 $y_1, \ldots, y_T$ に対して, 初期重みを一様 $v_{i,1} = 1/N$ とすると, *Vovk* の統合戦略に基づくアルゴリズムの累積利得下界は次のようになる.

$$G_{AA,T} \geq c(\alpha)\left(\max_{1\leq i\leq N} G_{i,T} - \frac{\ln N}{\ln \alpha}\right)$$

3.2 オークション型統合アルゴリズム

本節では, Vovk の統合戦略に基づいたオンラインオークション問題のアルゴリズムを導く. 式 (2) より, 時刻 t における AA の予測 $\boldsymbol{p}_t$ は, エキスパート i の重みを $v_{i,t}$, 予測を $\boldsymbol{x}_i$ とすると,

$$\boldsymbol{p}_t = \arg\sup_{\boldsymbol{p}\in P}\inf_{y\in Y} \frac{\lambda(y,\boldsymbol{p})}{\log_\alpha \sum_{i=1}^N v_{i,t}\alpha^{\lambda(y,\boldsymbol{x}_i)}} \tag{3}$$

である. また,

$$g(y) = \log_\alpha \sum_{i=1}^N v_{i,t}\alpha^{\lambda(y,\boldsymbol{x}_i)}$$

を仮予測と呼ぶ.

今, $\rho > 1$, $N = \lfloor \log_\rho h \rfloor + 1$ として, $b(i) = \rho^{N-i}, (i = 1, \ldots, N)$ としたとき, 結果空間 Y が

$$Y = \left\{ \begin{array}{c} (b(1), b(2), \ldots, b(N-1), b(N)) \\ (0, b(2), \ldots, b(N-1), b(N)) \\ \vdots \\ (0, 0, \ldots, 0, b(N)) \end{array} \right\}$$

の N 通りであるということに注意して式 (3) を解くことによって, オンラインオークション問題における予測が得られる. Vovk の統合戦略に基づくオンラインオークション問題におけるアルゴリズムを, これからはオークション型統合アルゴリズム (AuAA) と呼ぶことにする.

定理 2 $b(1) \geq \cdots b(N) > 0$ とする. 任意の $\boldsymbol{v}_t$ が与えられたとき,

$$p_t(j) = \frac{1}{b(j)}\log_\alpha\left(1 + \frac{(\alpha^{b(j)}-1)v_t(j)}{B(j)}\right) \Big/ Normalizer \tag{4}$$

$$B(j) = 1 + \textstyle\sum_{k=j+1}^N (\alpha^{b(k)} - 1)v_t(k)$$

で与えられるベクトル $\boldsymbol{p}_t$ は, $\sup_{\boldsymbol{p}} \inf_y \frac{\lambda(y,\boldsymbol{p})}{g(y)}$ の上限を達成する.

このように, オンラインオークション問題における予測ベクトルは複雑な調整をしている. ここで, $i < j$ について $B(i) > B(j)$ は明らかなので, この予測は大きな値段 $b(i)$ を支持するエキスパートほど, 重み付き平均 $v_t(i)$ を小さくして非線形変換していると解釈することができる. これは直感的に, オンラインオークション問題では, ある値段 $b(i)$ で売れるときは値段 $b(j) < b(i)$ でも必ず売れるので, 値段 $b(j)$ への重みの割り当てがその分増やされるためだと考えられる.

3.3 アルゴリズムの性能評価

次に, 前節で得られたオークション型統合アルゴリズムの性能評価を与える. アルゴリズムの評価は, 定理 1 における $c(\alpha)$ の導出である. なお, オークション型統合アルゴリズムの評価においては, 値段情報 $\boldsymbol{b} = (b(1), b(2), \ldots, b(N))$ を明示するため, $c(\alpha)$ の代わりに $c(\alpha, \boldsymbol{b})$ と表記する. さて, 定理 2 によって得られる $\boldsymbol{p}_t$ を予測ベクトルとして用いるとき, $c(\alpha, \boldsymbol{b})$ は

$$\begin{aligned} c(\alpha,\boldsymbol{b}) &= \inf_{\boldsymbol{q}} \frac{1}{\sum_{j=1}^N \frac{1}{b(j)}(d_j - d_{j+1})} \\ d_j &= \log_\alpha\left(1 + \sum_{k=j}^N (\alpha^{b(k)} - 1)q(k)\right) \end{aligned} \tag{5}$$

で与えられる．ここで，この式 (5) も右辺の分母に注目する．今，

$$f(\boldsymbol{q}) = -\sum_{j=1}^{N} \frac{1}{b(j)}(d_j - d_{j+1})$$

とすると，式 (5) を与える $\boldsymbol{q}$ は，$\sum_{j=1}^{N} q(j) = 1$，全ての j について $q(j) \geq 0$ と言う制約条件の下，目的関数 $f(\boldsymbol{q})$ を最小化する凸計画問題を解くことにより求められる．ここで，目的関数および定義域が凸であるとき，凸計画問題の最適解となる $\boldsymbol{q}^*$ についての必要十分条件として，Kursh-Kuhn-Kucker(KKT) 条件が知られている．これを適用すると，求めるべき $c(\alpha, \boldsymbol{b})$ を与える $\boldsymbol{q}^*$ は，ある t と $s(j)$ $(j = 1, \ldots, N)$ について次の条件を満たす $\boldsymbol{q}$ である．

$$\frac{\partial}{\partial q(j)} f(\boldsymbol{q}) - s(j) + t = 0, \quad j = 1, \ldots, N \tag{6}$$

$$\sum_{j=1}^{N} q(j) - 1 = 0, \tag{7}$$

$$s(j) \cdot (-q(j)) = 0, \quad j = 1, \ldots, N \tag{8}$$

$$-q(j) \leq 0, \quad j = 1, \ldots, N \tag{9}$$

$$s(j) \geq 0, \quad j = 1, \ldots, N \tag{10}$$

ここで，最初の条件 (6) は，ラグランジュ関数 $\nabla_{\boldsymbol{q}} L(\boldsymbol{q}, \boldsymbol{t}, \boldsymbol{s})$ から導かれる．

この KKT 条件を満たすような $\boldsymbol{q}$ を求め，$c(\alpha, \boldsymbol{b})$ に代入することで，次の定理が導かれる．

定理 3 $b(1) > b(2) > \cdots > b(N) > 0$ とし，便宜的に $b(0) = \infty$ とする．このとき，

$$r_j = \left(\frac{1}{b(j)} - \frac{1}{b(j-1)} \right) b(N), \tag{11}$$

$$s_j = \left(\frac{1}{\alpha^{b(j)} - 1} - \frac{1}{\alpha^{b(j-1)} - 1} \right) (\alpha^{b(N)} - 1) \tag{12}$$

とすると，

$$c(\alpha, \boldsymbol{b}) = \frac{b(N) \ln \alpha}{D(\boldsymbol{r} || \boldsymbol{s}) + b(N) \ln \alpha}. \tag{13}$$

ただし，$D(\boldsymbol{r}||\boldsymbol{s})$ は *KL-ダイバージェンス (Kullback-Leibler divergence)* である．

証明: まず，次のような関数 $F(x, y)$ を定義する．

$$F(x, y) = \frac{\frac{1}{x} - \frac{1}{y}}{\frac{1}{\alpha^x - 1} - \frac{1}{\alpha^y - 1}}.$$

この関数 $F(x, y)$ は，任意の $a < b < c$ について $F(a, b) < F(b, c)$ である．今，$b(1) > b(2) > \cdots > b(N) > 0$ であり，便宜的に $b(0) = \infty$ である．このとき，$j = 1, \ldots, N$ について，

$$q(j) = \frac{F(b(j), b(j-1)) - F(b(j+1), b(j))}{t(\alpha^{b(j)} - 1)}, \tag{14}$$

$$s(j) = 0, \tag{15}$$

$$t = \frac{\frac{1}{b(N)}}{1 + \frac{1}{\alpha^{b(N)} - 1}} \tag{16}$$

は，KKT 条件 (式 (6)〜式 (10)) を満たす．ただし，便宜的に $F(b(N+1), b(N)) = t$ とした．

さて，$d_j = \log_\alpha \left(1 + \sum_{k=j}^{N} (\alpha^{b(k)} - 1) q_t(k) \right)$ および式 (14) より，

$$d_j = \log_\alpha F(b(j), b(j-1)) - \log_\alpha t.$$

よって，式(5)より，

$$c(\alpha, \boldsymbol{b}) = \frac{\ln \alpha}{\sum_{j=1}^{N} \left(\frac{1}{b(j)} - \frac{1}{b(j-1)} \right) \ln F(b(j), b(j-1)) - \frac{1}{b(N)} \ln t}.$$

ここで，$F(x,y) = \frac{\frac{1}{x} - \frac{1}{y}}{\frac{1}{\alpha^x - 1} - \frac{1}{\alpha^y - 1}}$ より，r_j および s_j を式(11)，(12)のようにおくと，$\sum_{j=1}^{N} r_j = 1$，$\sum_{j=1}^{N} s_j = 1$ である．さらに，

$$F(b(j), b(j-1)) = \frac{r_k}{s_k} \frac{\alpha^{b(N)} - 1}{b(N)}$$

であり，

$$t = \frac{\frac{1}{b(N)}}{1 + \frac{1}{\alpha^{b(N)} - 1}} = \frac{\alpha^{b(N)} - 1}{b(N)\alpha^{b(N)}}$$

だから，

$$\begin{aligned} c(\alpha, \boldsymbol{b}) &= \frac{\ln \alpha}{\sum_{j=1}^{N} \frac{r_j}{b(N)} \ln \frac{r_k}{s_k} \frac{\alpha^{b(N)} - 1}{b(N)} - \frac{1}{b(N)} \ln \frac{\alpha^{b(N)} - 1}{b(N)\alpha^{b(N)}}} \\ &= \frac{\ln \alpha}{\sum_{j=1}^{N} \frac{r_j}{b(N)} \ln \frac{r_k}{s_k} - \frac{1}{b(N)} \ln \frac{1}{\alpha^{b(N)}}}. \end{aligned}$$

□

さて，従来の手法と比較するため，$N = \lfloor \log_\rho h \rfloor + 1$ とし，値段の集合 R を $R = \{\rho^{N-1}, \ldots, \rho^2, \rho, 1\}$ とする．すなわち，各値段は $b(i) = \rho^{N-i}$，$(i = 1, \ldots, N)$ であらわされる．これを $c(\alpha, \boldsymbol{b})$ に代入すれば良いのだが，残念ながら，このオークション型統合アルゴリズムの $c(\alpha, \boldsymbol{b})$ を従来のアルゴリズムと直接比較できるような平易な形にあらわすのは困難である．そのため，数値計算により比較を行う．

3.4 数値計算による比較

オークション型統合アルゴリズムの累積利得の下界をヘッジアルゴリズム及びHGアルゴリズムの累積利得の下界と比較するため，数値計算によりそれぞれの下界を求める．今，$N = \lfloor \log_\rho h \rfloor + 1$，$\boldsymbol{b} = (\rho^{N-1}, \ldots, \rho^2, \rho, 1)$ としたとき，それぞれのアルゴリズムの評価式は，

$$EG_{\mathrm{Hedge},T} \geq \frac{\ln \alpha}{\alpha - 1} \frac{\mathrm{OPT}}{\rho} - \frac{\rho^{\lfloor \log_\rho h \rfloor + 1}}{\alpha - 1} \ln \left(\lfloor \log_\rho h \rfloor + 1 \right), \tag{17}$$

$$EG_{\mathrm{HG},T} \geq (1-\delta)\frac{\mathrm{OPT}}{\rho} - 2h(1-\delta)\left(\frac{2}{\delta} \ln \nu(\rho) + \frac{\nu(\rho)}{\delta^2}(1-\delta)^{\nu(\rho)} + 1 \right), \tag{18}$$

$$EG_{\mathrm{AuAA},T} \geq c(\alpha, \boldsymbol{b})\frac{\mathrm{OPT}}{\rho} - \frac{c(\alpha, \boldsymbol{b})}{\ln \alpha} \ln \left(\lfloor \log_\rho h \rfloor + 1 \right) \tag{19}$$

である．ただし，$\nu(\rho) = \lfloor \log_\rho 2 \rfloor + 1$ とした．ここで，$\frac{\mathrm{OPT}}{\rho}$ の係数を r に統一した場合のそれぞれのアルゴリズムの付加損失項の大きさを，様々な h において比較する．すなわち，ある h に対して $\frac{\ln \alpha}{\alpha - 1} = 1 - \delta = c(\alpha', \boldsymbol{b}) = r$ となるためのパラメータ α, δ, α' の値をそれぞれ求め，各アルゴリズムの付加損失項にそれぞれ代入し，得られた値を比較する．

図1，に示したグラフは，ヘッジアルゴリズム，HGアルゴリズムおよびオークション型統合アルゴリズムによる付加損失項の h の増加に対する相対的な増加量を，値段の刻みを $\rho = 1.001$ について描いたものである．今回行なった数値計算の範囲では，HGアルゴリズムの付加損失項に対するオークション型統合アルゴリズムの付加損失項は極めて小さいため，それぞれの図の左のグラフに3つのアルゴリズムの付加損失項の大小関係を比較するためのグラフを，右のグラフにはオークション型統合アルゴリズムの付加損失項の傾向を見るためにHGアルゴリズムの付加損失項のグラフを除いたグラフを掲載した．

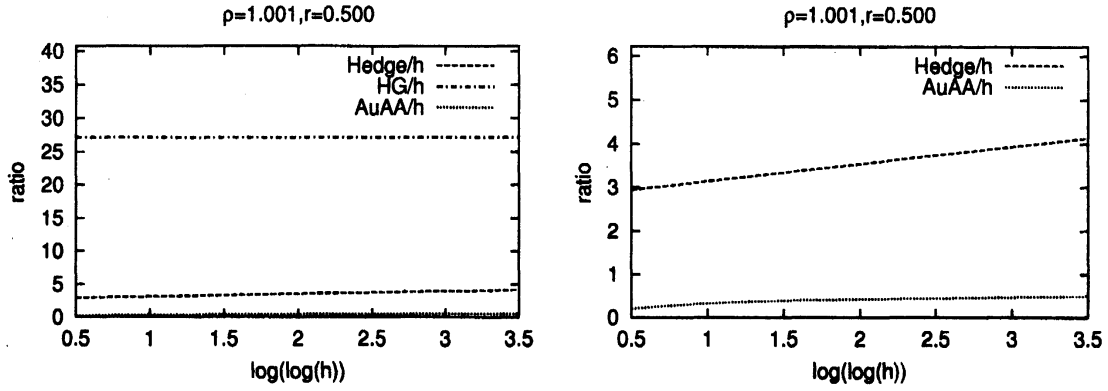

図 1: $\rho = 1.001$ のとき

今回数値計算した範囲では，競合比 r を固定した場合のオークション型統合アルゴリズムの付加損失項は，他のアルゴリズムと比較して小さいものとなっていることが確認できる．

しかし，残念ながら，式 (19) に基づく数値計算の結果は，ヘッジアルゴリズムと同様に，h の増加に伴い付加損失項が増加していることが確認できる．現在確認できる傾向が $\ln \ln h \in [0.5, 3.5]$ と狭い範囲であるため，このまま h が増加していったときに (19) の評価式におけるオークション型統合アルゴリズムの付加損失項が常に増加し続けるかどうかはわからない．しかし，仮にこのまま増加していくのならば，巨大な h では，HG アルゴリズムの付加損失項を上回ってしまうことを否定できない．その原因として考えられることは，本論文では詳しく言及しなかったが，初期重み $v_{i,1}$ の設定法にある．初期重みの調節については今後の課題である．

4 まとめ

本論文では，オンラインオークション問題がエキスパートへの報酬の与えられ方に依存関係がある特殊な資源配分問題であることを言及した．オークション型の問題では，エキスパートが受け取る報酬がそれぞれ独立に決まる通常のオンライン資源配分問題と比較し，考慮すべき結果空間が狭い．そこで，本論文では，オークション型の問題に適合するように Vovk の統合戦略を適用し，新たにオークション型統合アルゴリズムを導いた．

オークション型問題に特化したオークション型統合アルゴリズムの累積利得の下界は，従来のアルゴリズムによる累積損失の下界よりも，かなり良いという結果が数値計算により確認された．ただし，今回確認した範囲においては，オークション型統合アルゴリズムの累積利得の下界の付加損失項の比は，h の増加に伴って増加していた．これを改善するため，今後は初期重みの調整を考慮した評価を行う必要がある．

参考文献

[BKRW03] A. Blum, V. Kumar, A. Rudra, and F. Wu. Online Learning in Online Auctions. *In Proc. 14th Symposium on Discrete Algorithms. ACM/SIAM*, 2003.

[BH05] A. Blum and J. D. Hartline. Near-Optimal Online Auctions. *In Proc. SODA 2005, pages 1156-1163*, 2005.

[BHW02] Z. Bar-Yossef, K. Hildrum and F. Wu. Incentive-Compatible Online Auctions for Digital Goods. *In Proc. SODA 2002, pages 964-970*, 2002.

[V98] V. Vovk. A game of prediction with expert advice. *JCSS, 56(2):153-173*, 1998.