

WEB 広告における実時間オークションのシミュレーションモデル

○鍋田武頼 山本学 吉川厚 寺野隆雄 (東京工業大学)

Study on The Simulation Model of A Real-Time Auction in WEB Advertising Frame

*T. Nabeta, G. Yamamoto, A. Yoshikawa, and T. Terano (Tokyo Institute of Technology)

Abstract— Real-Time Bidding (RTB) is such that goods are traded by way of auction on a real-time basis and ad spot has been done in this way in recent years. RTB is thus a new auction style for bidders called Demand-Side Platform (DSP) because the number of trades and bidding styles are different from those in conventional auctions. This study explores an effective and less costly bidding strategy to make more successful bids in RTB by agent-based approach. In the simulation model, auctions of ad spots occur in accordance with time slots in a day and the participants bid based on their strategies. The computational experiments show that staying out is of particular importance.

Key Words: Real-Time Bidding, Auction, Simulation

1 序論

1.1 目的

近年、WEB 広告の市場は金額ベースで拡大し続けており、それに伴い様々な形で広告枠の取引がされている。その中で、WEB 広告の一種であるディスプレイ広告の取引形態である Real-Time Bidding (RTB) に注目が集まっている。RTB とは、ユーザーが広告掲載面のある WEB ページにアクセスし、広告枠が発生すると、その枠について広告主側でオークションを行い、配信する広告を決める方式のことを指す¹⁾。

本研究では RTB を模したオークションのモデル化に取り組む。RTB をモデル化するにあたり本研究では、Agent-Based Modeling (ABM) を用いる。ABM とは、「エージェント」と呼ぶ内部状態と意思決定力を備えた複数の主体についてモデル化する手法である²⁾。ABM は、複雑な社会システムの特徴をシミュレーションによる分析によって理解することを目的とする³⁾。RTB におけるオークションは、入札者である Demand-Side Platform (DSP) という複数の主体による意思決定、すなわち入札戦略が絡みあいながら動いている複雑なシステムになっており、ABM の手法はモデル化にあたり、非常に有用といえる。

本研究では RTB 型のオークション取引をモデル化し、そのモデルの中で入札者はどのような入札戦略をとれば、より少ない金額で多くの商品を落札していけるのか考察する。

1.2 本稿の構成

本稿では、2 章で RTB について説明をする。3 章ではオークションの関連研究について述べる。4 章では、実際の RTB におけるオークションの流れについて説明をする。5 章では、本研究で提案する RTB の擬似オークション取引モデルについての説明をする。6 章では、提案したモデルにおける入札者 (DSP) の行動モデルについて述べる。7 章では、提案したモデル及び行動モデルを基に行ったシミュレーション結果と、その結果からの入札戦略についての考察を述べる。最後に 8 章では、本研究のまとめと今後の課題について述

べる。

2 Real-Time Bidding について

2.1 背景

RTB では 2 つのプレイヤーが存在する。1 つ目は Demand-Side Platform (DSP) である。DSP は広告主の利益向上のための広告枠の買い付け判断を行うプラットフォームである。2 つ目は Supply-Side Platform (SSP) である。SSP は媒体主の広告収益を最大化させるためのプラットフォームである。本研究では DSP に注目する。

ここで RTB の市場全体の枠組みを Fig. 1 に示す。

ある WEB サイト 1 は SSP 1 と、WEB サイト 2 は SSP 2、WEB サイト 3 は SSP 3 と契約している。そして SSP 1 は DSP 1、2 と取引しており、SSP 2 は DSP 1、3、SSP 3 は DSP 2、3 と取引している。RTB の市場は SSP ごとに生まれる。Fig. 1 では、SSP 1 が持つ RTB 市場は SSP 1、DSP 1、DSP 2 との間で生まれる。SSP 2、3 についても同様である。

ここで DSP から見た時の RTB 市場におけるオンライン広告出稿までのプロセスを Fig. 2 に示す。A さん (Cookie=123) がコスメサイトにアクセスしたとする。その際、A さんがアクセスしたコスメサイトでは広告枠が発生する。広告枠が発生すると、そのコスメサイ

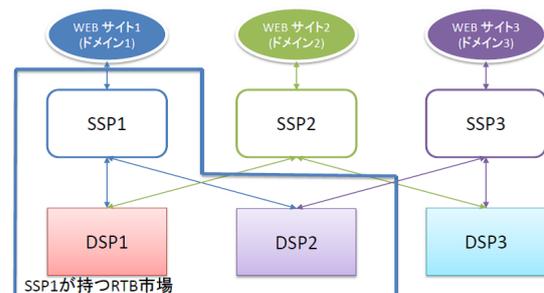


Fig. 1: RTB の市場全体の枠組み

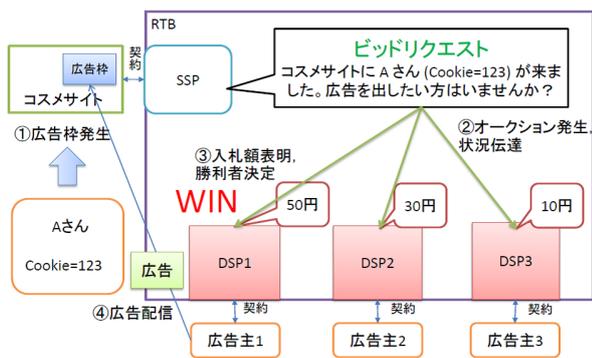


Fig. 2: RTB の流れ

トと契約をしている SSP が DSP に対して、ビッドリクエストを行う。ビッドリクエストとは、オークションにおける商品の出品である。ビッドリクエストは、「A さん (Cookie=123) がただ今コスメサイトにアクセスしました。広告を出したい方はいませんか?」といった形で行われる。そのリクエストを参考に、各 DSP がオークションに参加するか否か、参加するのであれば A さんに出稿したい広告を選び、入札価格を決め、SSP にレスポンスをする。各 DSP の入札の中から最高の値をつけた DSP がオークションの勝者となり、広告出稿の権利を得る。権利を得た DSP が出稿したい広告を A さんに配信する。なお、RTB ではセカンドプライスオークションが主流であり、オークションの勝者となった DSP が実際に支払う金額は入札価格の中で二番目に高い額である。

RTB は両プレイヤーにメリットが有る方式である。DSP から見ると、ターゲットとしたユーザーのみに広告を配信することにより、広告効果の向上が図れる。SSP から見ると、広告単価の高い広告主を優先することが可能になり、媒体の利益向上につながる。このような背景もあり、RTB の市場規模は拡大し続けている。

そういった中で DSP の入札戦略に関する研究⁴⁾⁵⁾⁶⁾や、ユーザーの行動を予測するモデルの研究⁷⁾など、RTB に関する研究も多くなされてきている。

2.2 問題点

ここで一般的なオークションの例としてヤフオク!⁸⁾を取り上げ、入札者にとって一般的なオークションと RTB におけるオークションとの間にどのような違いがあるのか、比較したものを Table. 1 に示す。

Table 1: ヤフオク! と RTB におけるオークションとの比較

	ヤフオク!	RTB
1 日の落札数	数点	数万点
形式	競り上げ	一発勝負
制約金額	一商品ごと	複数商品

まず、1 つ目が 1 日の落札数である。一般的なユーザーの場合、ヤフオク! などの一般的なオークションで落札するのは 1 日数点だと考えられる。RTB の場合、1 日数万件落札していく必要がある。2 つ目が形式である。ヤフオク! は入札額を入札者同士で競り上げていく形式である。RTB では、1 商品に対して入札額を 1 回しか示せない一発勝負である。3 つ目が制約金額で

ある。ヤフオク! などの一般的なオークションであれば、一商品ごとに制約金額を決めることが多いと考えられる。RTB では、入札者である DSP は広告主からオークションで使える金額 (予算) や、目標の広告枠獲得数が与えられている。すなわち複数商品で制約金額を決める。

上記に挙げた点が一般的なオークションと比較した時の RTB におけるオークションの特徴であり、これらの特徴が絡み合ったオークションは、他にはない。RTB は近年誕生した新しい形態であり、その性質自体が理解されていないという問題点がある。そこで、本研究では RTB の特徴を持つオークションをモデル化しシミュレーションを行うことで、RTB のオークション取引における入札戦略の理解が深まることを期待している。

3 オークションの関連研究

RTB はオークション形式で広告枠の取引がされている。本章では既存のオークション理論の関連研究について述べる。

オークションとは、売り手が提供する財に対して、買い手が提示する入札額をもとに落札者と支払額を定める手続きのことである⁹⁾。オークションには様々な形式が存在する。入札者同士で入札額の競り上げを行い、入札者が番号札を上げて、会場全体に入札額を知らせる形式である“イギリス式オークション”や競売人が入札価格を下げていき、最初に購入希望を出した買い手に商品を販売する形式である“オランダ式オークション”などである¹⁰⁾。RTB でのオークションは、一発勝負でオークションの勝者が決定し、セカンドプライスオークションが主流である。既存のオークション形式と RTB でのオークションを比較した時、一番近い形式なのが 2 位価格封印入札オークションである。2 位価格封印入札オークションとは、自分の入札額を他の入札者に知られないように提出し、最も高い額で入札した者が 2 番目に高い入札額を支払い落札する形式である。この方式はアメリカの経済学者ウィリアム・ヴィックリーによって考案された。2 位価格封印入札オークションでは理論上、入札者の評価額を入札すること (真実表明入札) が利益を最大化することにつながるとされている¹¹⁾。理想的に言えば、各 DSP が自分の評価額を入札することが最適戦略といえるのかもしれないが、DSP は広告主から様々な評価の仕方をされている。例えば広告を出稿できた“数”であったり、広告の平均落札“額”である。これらの背景も有り、入札者は最適な戦略を見つけるのが難しい。

続いてオークションの戦略についての関連研究について述べる。売り手の行動解析に関する研究¹²⁾や、数理的なモデルからの入札戦略に関する研究¹³⁾など、オークション戦略に関する研究はいくつかされている。その中で戦略を学習していくようなモデルの研究は大竹ら¹⁴⁾を除いて、あまり行われていない。大竹らはイギリス式オークションにおいて入札戦略を強化学習する入札アルゴリズムの提案をした。強化学習とは、試行錯誤しながら最適な行動を学習していくことを指す。本研究では、大竹らの戦略モデルを、DSP の行動モデルに適用する。

Table 2: サンプルデータ

day	hour	advertised	domain	brng	res	imp	click	ttlbid	ttlwin
201304XX	0	1	sample1.com	0	10	0	0	100	0
201304XX	23	3	sample2.co.jp	1000	100	10	1	100000	10000

4 実際の RTB におけるオークションについて

本章では、RTB におけるオークションのモデル化のために、実際の RTB におけるオークションについて説明していく。

ここでサンプルデータについて説明する。このデータは実際に DSP のサービスを運営する企業と共同で作成したサンプルデータである。仮定している対象期間は、2013 年 4 月の連続した 7 日間である。データ対象広告主数は 3 社とした。対象ドメインは 1 週間で 20 万件以上入札するドメインを想定し、合計のレコード数は約 52 万件とした。データのサンプルを Table. 2 に示す。

Table. 2 の項目について説明する。day は日付を表す。hour は時間帯を表す。0 は 0 時 00 分～0 時 59 分を、1 は 1 時 00 分から 1 時 59 分を表す。advertised は広告主 ID を表し、ここでは 3 種類ある。domain は広告が配信される媒体サイト (ドメイン) を表す。なお、サンプルデータのためドメインの表記は“sample...”とした。brng は入札価格帯を表す。0 は 0.000～0.049 円の価格帯、50 は 0.050～0.099 円までの価格帯を表す。1000 は 1.000 円以上の入札価格である。res はこの条件に対する入札件数を表す。例えば、201304xx の 0 時間帯では、“sample1.com”に 0 の価格帯で 10 回入札していることを表す。imp はインプレッション件数、すなわちこの条件で入札を行った中で勝利した件数を表す。click はクリック数を表す。落札した広告枠の中でユーザーがクリックした想定回数である。ttlbid はこの条件で入札した合計の“入札”額であり、ttlwin は合計の“落札”額である。このサンプルデータを参考にし、モデルのパラメータを設定していく。サンプルデータは、時間帯及びドメインごとに広告枠の発生数を変化させるなど、実際の RTB の特徴を取り入れた。

次に RTB でのオークションの流れについて説明していく。RTB でのオークションは主に 3 つのフェーズに分けられる。

1. 広告枠発生 (ビッドリクエスト発生)
2. DSP 入札
3. オークション (DSP の入札額から出稿広告及び落札額を決定)

以下、それぞれのフェーズについて説明していく。

4.1 広告枠の発生について

あるユーザーが広告掲載面のある WEB サイト (ドメイン) にアクセスすると広告枠が発生する。広告枠が発生すると同時にそのドメインと契約している SSP に広告枠発生情報が伝わり、その SSP と契約している DSP にビッドリクエストが送られる。ビッドリクエストには cookie ベースで分かるユーザー情報 (例:

よく見ているサイトやアクセス端末の情報) やドメインの情報などが入っている。

4.2 DSP の入札について

実際の DSP の入札では SSP から送られてくるビッドリクエストから広告枠のユーザー情報やドメインなどの情報を受け取り、その情報を参考に、DSP ごとに広告枠の価値を評価し、どの広告を表示させるかを決め、入札額を決定する。それらの情報を SSP にレスポンスする事で、入札が終了する。以下の Fig. 3 に実際の入札の流れの例をフローチャートで示す。

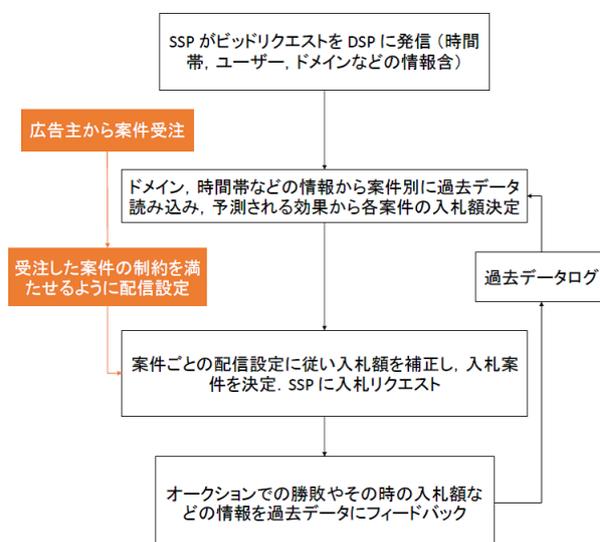


Fig. 3: 実際の DSP の入札までの流れ

Fig. 3 のオレンジで示した枠がオークション外で行われるものであり、黒で示した枠がオークション内で行われるものである。まずオークションが行われる前に、DSP は広告主から案件受注する。案件を受注したら、既に受注している案件との兼ね合いも考慮し、それぞれの案件の制約を満たせるように配信設定を行う。そして実際のオークションに入る。DSP は SSP からビッドリクエストを受信する。ビッドリクエストには時間帯やユーザー、ドメインなどの情報が含まれている。DSP はドメイン、時間帯などの情報から案件別に過去データ読み込み、予測される効果から各案件の入札額決定する。次に先に説明した案件ごとの配信設定に従い入札額を補正し、入札案件を決定する。そしてそれらの情報を SSP に入札リクエストとして発信する。最後にその入札リクエストによるオークションでの勝敗や、その時の入札額などの情報を過去データログにフィードバックする。以上が DSP の入札までの流れの一例である。

4.3 オークションについて

実際の RTB におけるオークションは以下の流れで進む。

1. SSP が DSP から入札額のリクエストを受け取りと勝者 (最高額入札者) を決定する。
2. 勝者の DSP は二番目に高い入札額を勝者は支払う

オークションでは競り上げは行われずに、一発で勝敗が決まる。すなわち、1つの広告枠に対して1回しか入札が許されていない。

5 オークションモデルの設計

本モデルは DSP から見た RTB の特徴を持った擬似オークションである。はじめに大まかなモデルの概要を説明する。

まず1日を1単位として考えることとし、ユーザーがアクセスする可能性のあるドメイン数は10とした。そして、1日1,218,983件広告枠が発生するとした。これは、サンプルデータより1日あたりに入札してる件数の平均値である。DSPはビッドリクエストから時間とドメインの情報のみ受け取りとした。最後に DSP は3社の広告主と取引しているものとし、以下“案件 A”、“案件 B”、“案件 C”と呼ぶ。

ここで RTB の市場全体を考えるためには SSP の要素も考慮する必要がある。しかし、本研究で注目するのは DSP から見た時のオークション“取引”である。DSP から見た時、ビッドリクエストがどの SSP からきたリクエストだろうが、入札するのに差異はない。これより本モデルにおいて SSP の要素は除外した。

そして、実際の RTB でのオークションでは時間帯及びドメイン別に広告枠の発生数に違いがあることから、各時間帯の広告枠の発生件数、ドメイン別の広告枠の発生確率を設定した。まず、先に述べたサンプルデータから各時間帯の広告枠の発生件数を Fig. 4 のように設定した。

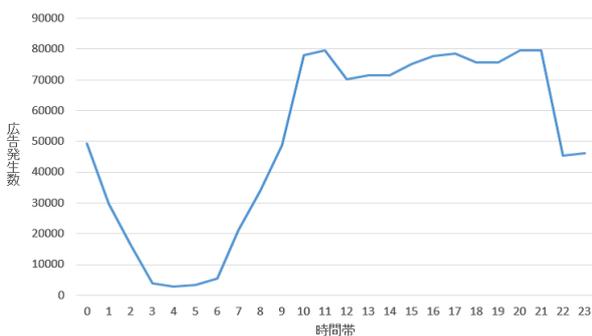


Fig. 4: 時間帯別の広告発生数

早朝の時間帯は広告発生数は少なく、昼から夕方にかけて多くなっている。

次に、時間帯別のドメインにおける広告枠の発生確率を Table. 3 のように設定した。例えば、0時帯にビッドリクエストがあった場合、そのビッドリクエストのドメインが“sample1.com”である確率は23.2%であるとした。

Table 3: 時間帯・ドメイン別の広告枠発生確率例

時間帯	ドメイン	発生確率 (%)
0	sample1.com	23.2
0	sample2.co.jp	24.4
...
11	sample4.com	17.4
11	sample5.co.jp	7.08
11	sample6.com	6.00
...
23	sample9.com	5.09
23	sample10.com	2.11

6 DSP の行動モデルの設計

本モデルにおいて DSP は共通の行動ルール、及び個別のビッドポリシーを持っているものとした。DSP は入札に際しそれぞれに従うものとする。まず、共通の行動ルールを Fig. 5 に示す。

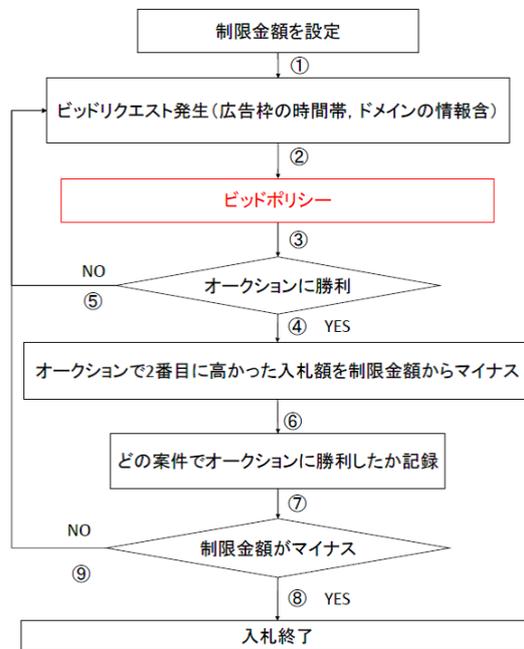


Fig. 5: 共通の行動ルール

まず制限金額を設定する。そしてビッドリクエストが発生する。このビッドリクエストには広告枠の時間帯やドメインの情報が含まれる。次の“ビッドポリシー”の部分で、正直モデル及び確率モデルに依存しており、そこで各 DSP の入札案件及び入札額が決まるものとした。オークションに勝利した場合、オークションで2番目に高かった入札額を制限金額からマイナスする。オークションで負けてしまった場合、次のビッドリクエストまで待機となる。それから、どの案件でオークションに勝利したか記録する。最後に制限金額がマイナスであるなら入札終了とし、プラスであるなら次のビッドリクエストに待機するものとした。

ここでビッドポリシーの概要について説明する。ビッドポリシーは計5つある。1つ目のビッドポリシーは、案件に対する入札額に対して補正を行った後、それぞれ

案件の中で最高額であったものを入札するものであり、以下“正直モデル”と呼ぶ。2つ目のビッドポリシーは、入札する案件を確率的に決めるものであり、以下“確率モデル”と呼ぶ。3つ目のビッドポリシーは、著者ら¹⁵⁾の研究で用いられたモデルを、本モデルに適用したものであり、以下“競り上げモデル”と呼ぶ。4つ目のビッドポリシーは、大竹ら¹⁴⁾の研究で用いられたモデルを、本モデルに適用したものである。以下“つり上げモデル”と呼ぶ。5つ目の行動ルールは、“正直モデル”、“競り上げモデル”、“つり上げモデル”及び“降りる”の、計4つの行動に対して強化学習を適用するものとした。以下“学習モデル”と呼ぶ。

以下、この5つのモデルについて詳細に説明していく。

6.1 正直モデル

まず正直モデルについて説明していく。以下の Fig. 6 に正直モデルによる入札の流れをフローチャートで示す。

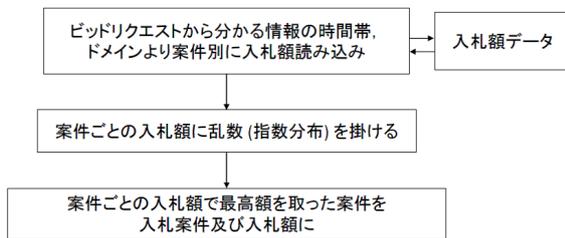


Fig. 6: 正直モデル

まずビッドリクエストから分かる情報の時間帯、ドメインから、サンプルデータよりあらかじめ設定した入札額データを読み込む。この入札額データとは案件ごとの時間帯、ドメイン別の平均入札額を記したものである。データ例を以下の Table. 4 に示す。

Table 4: 案件・時間帯別の平均入札額

案件	時間帯	ドメイン	平均入札額
A	0	sample1.com	0.0967
B	0	sample1.com	0.168
...
A	11	sample4.com	0.000
B	11	sample4.com	0.0802
C	11	sample4.com	0.0294
...
B	23	sample10.com	0.462
C	23	sample10.com	0.000

なお、平均入札額が0.000のところは、その条件の際に入札していないことを意味している。この入札額を読み込んだ後に、乱数を掛ける。

乱数は以下の数式で表される指数分布に依存しているものとした。指数分布の確率密度関数は以下の (1) 式で表される。

$$f(x; \lambda) = \lambda \exp(-\lambda x) \quad (1)$$

これより分布関数 $g(\tau)$ は、

$$\begin{aligned} g(\tau) &= \int_0^\tau \lambda \exp(-\lambda \tau) d\tau \\ &= 1 - \exp(-\lambda \tau) \end{aligned} \quad (2)$$

逆関数法により、指数分布に従う乱数 τ は、

$$\tau = -\frac{1}{\lambda} \log(1 - g(\tau)) \quad (3)$$

で表される。

そして案件ごとの入札額を決定し、その中で最高額であった案件を入札案件及び入札額に決定する。

以上が正直モデルである。

6.2 確率モデル

次に確率モデルについて説明していく。以下の Fig. 7 に確率モデルによる入札案件及び入札額決定までの流れをフローチャートで示す。

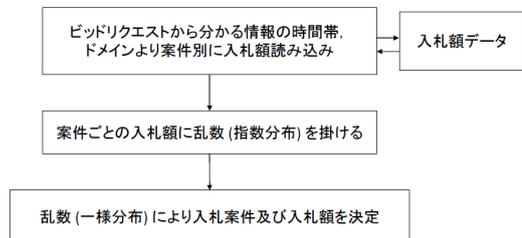


Fig. 7: 確率モデル

ビッドリクエストから分かる情報の時間帯、ドメインから、あらかじめ設定されている入札額を読み込み、入札額に乱数を掛ける点までは正直モデルと同一である。最後の案件を選ぶ点であるが、正直モデルでは案件の中で最高額なものを入札案件としていた。しかし、確率モデルは一様分布による確率で入札案件を決めるものとした。

以上が確率モデルである。

6.3 競り上げモデル

本節では競り上げモデルについて説明していく。このモデルは、著者ら¹⁵⁾の研究で用いられたモデル参考にした。著者らのモデルでは逐次型の封印入札オークションにおいて、1回目のオークションのみランダムに入札額を決め、以降は1回前のオークションに勝利したら入札額を下げ、敗北したら入札額を上げるというものである。

Fig. 8 に競り上げモデルによる入札案件及び入札額決定までの流れをフローチャートで示す。競り上げモデルでは、ある時間帯の1回目は0.001~0.150円の額をランダムに補正值に設定する。補正值を入札額とし、入札額が最大な案件を入札案件とする。勝利した場合補正值を0.010下げ、敗北した場合補正值を0.005上げる。ただし、入札額が0.300を上回ったら、当初の状態に戻る。例えば、入札額0.250で勝利した場合、次の入札額は0.240とし、敗北したら場合、次の入札額は0.255とする。

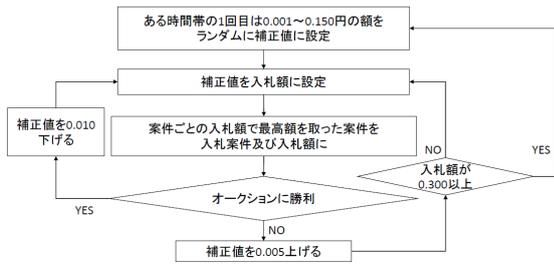


Fig. 8: 競り上げモデル

以上が競り上げモデルである。

6.4 つり上げモデル

本節ではつり上げモデルについて説明していく。このモデルは大竹ら¹⁴⁾の研究で用いられたモデルを、本モデルに適用したものである。大竹らはイギリス式オークションにおいて、入札者は以下の3つの行動を取るものとした。

1. 単位入札額だけせり上げる (競り上げ)
2. 自分の評価額との差を半分にする (つり上げ)
3. オークションから降りる

上記の“つり上げ”を、本モデルに適用した。

以下の Fig. 9 につり上げモデルによる入札案件及び入札額決定までの流れをフローチャートで示す。

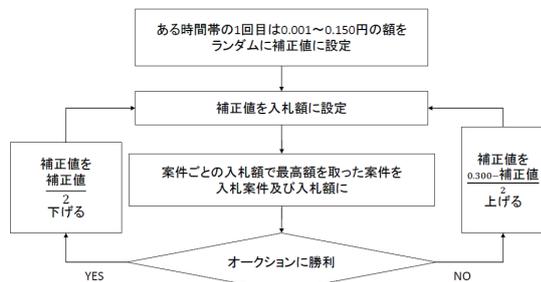


Fig. 9: つり上げモデル

つり上げモデルでは、ある時間帯の1回目は0.001～0.150円の額をランダムに補正值に設定する。補正值を入札額とし、入札額が最大な案件を入札案件とする。勝利した場合補正值を“ $\frac{\text{補正值}}{2}$ ”下げ、敗北したら補正值を“ $\frac{0.300 - \text{補正值}}{2}$ ”上げるものとする。例えば、入札額0.250で勝利した場合、次の入札額0.125とし、敗北した場合、次の入札額は0.275にする

以上がつり上げモデルである。

6.5 学習モデル

本節では学習モデルについて説明していく。学習モデルでは強化学習の代表的手法であるQ学習を参考にする。Q学習ではどの状況でどの行動をとることがどれくらい望ましいかを示すQ値を用いる。状態 s_{t+1} で報酬 r を得るごとに、次式に従って前状態 s_t におけるQ値を更新する。ここでQ値の更新式は下式で表される。

$$Q(s_t, a) \leftarrow Q(s_t, a) + \alpha[r_{t+1} + \gamma \max_p Q(s_{t+1}, p) - Q(s_t, a)]$$

ここで $\alpha(0 \leq \alpha \leq 1)$ は学習率と呼ばれる。

今回のモデルでは、初期状態はビッドリクエストが来た時であり、終端状態は落札したか否かの時である。すなわち1回しか状態は遷移しないため、 γ の項より右は使用しないものとした。よって、本モデルで用いるQ値の更新式は下式で表される。

$$Q(s_t, a) \leftarrow Q(s_t, a) + \alpha r_{t+1} \quad (4)$$

こうして得られたQ値をもとに、意思決定は ϵ -greedy選択によって行動を決定する。これは各状態において、確率 $1 - \epsilon$ でQ値を最大とする行動 a を選択し、確率 ϵ で行動をランダムに選択する方法である。

本モデルにおいて、状態とはビッドリクエスト情報(時間帯, ドメイン)及び落札したか否かで表されるものとする。すなわち、

$$s_t = [\text{hour}][\text{domain}][\text{rakusatsu}]$$

で表されるものとする。ここで $hour$ は広告枠の発生時間帯、 $domain$ は広告枠のドメイン、そして $rakusatsu$ は落札したか否か、とした。

次に行動である。行動 a は、正直モデル、競り上げモデル、つり上げモデル、そして降りるのいずれかを用いるものとした。降りるというのは、あるビッドリクエストに対して、入札しないことを意味する。

最後に報酬は以下のとおり2種類設定した。

1. オークションで勝利し、その時の落札額が0.200以下の場合 r_1
2. オークションで勝利し、その時の落札額が0.200以上の場合 r_2

以上が学習モデルである。

7 実験と考察

はじめに共通の実験条件を述べる。1日の制約金額は27,500円とした。これはサンプルデータの1日に使用した平均金額を参考にしている。そして、それぞれの案件ごとにビッドする時間帯が決まっているものとする。Table 5に案件ごとのビッドする時間帯を示す。

Table 5: 案件別のビッドする時間帯

案件	ビッドする時間帯
A	0～2, 7～23
B	0～23
C	10～21

今回は以下の4種類の実験を行った。

1. N のDSPが正直モデルで入札
2. $\frac{N}{2}$ のDSPが正直モデルに、残りの $\frac{N}{2}$ のDSPが確率モデルに従い入札

- $\frac{N}{3}$ の DSP が正直モデルに, $\frac{N}{3}$ の DSP が競り上げモデルに, 残りの $\frac{N}{3}$ の DSP がつり上げモデルに従い入札
- $\frac{N-1}{3}$ の DSP が正直モデルに, $\frac{N-1}{3}$ の DSP が競り上げモデルに, $\frac{N-1}{3}$ の DSP がつり上げモデルに, 残り 1 つの DSP が学習モデルに従い入札

そして (3) 式で表される乱数 τ を発生させるときに必要な λ は, $\lambda = 3$ とした. 実験はそれぞれ 10 試行 (10 日分) ずつ行った. 以下, それぞれの実験について詳細に述べる.

7.1 実験 1

まず N の DSP が正直モデルで入札する場合でシミュレーションを行った. DSP の数 N は 12 と 24 の 2 パターン行った.

7.1.1 実験 1 の結果

実験結果として, 1 広告枠あたりの平均落札額と時間帯別の 1 広告枠あたりの平均落札額の推移を, それぞれ Table. 6, Fig. 10 に示す.

Table 6: 実験 1 の 1 広告枠あたりの平均落札額

DSP 数	平均落札額	標準偏差
12	0.168	5.13×10^{-4}
24	0.216	8.34×10^{-4}

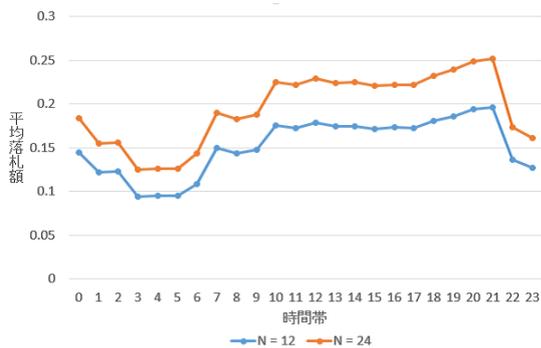


Fig. 10: 実験 1 における時間帯別の平均落札額推移

7.1.2 実験 1 の考察

Table. 6, Fig. 10 から DSP の数が多くなれば多くなるほど, 平均落札額が高くなっていることが読み取れる. これは, DSP が多くなれば多くなるほど競争が激しくなるためであると考えられる. ここから DSP と SSP の契約による戦略に関して考えられる事がある. 例えば, DSP は SSP と契約する際, 多くの DSP と契約している SSP よりも 少ない数の DSP と契約している SSP と契約したほうが良いし, 逆に SSP は多くの DSP と契約を結んだ方が, 広告枠は高い額で売れやすくなる. この部分に関して, RTB の市場全体をモデル化することにより SSP と DSP との関係についても考察することが, 今後の研究課題である.

7.2 実験 2

次に $\frac{N}{2}$ の DSP が正直モデルに, 残りの $\frac{N}{2}$ の DSP が確率モデルに従い入札する場合でシミュレーションを行った. N は 12 に固定した.

7.2.1 実験 2 の結果

実験結果として, 各モデルの DSP 1 社あたりの 1 広告枠あたりの平均落札額及び平均落札数, そして各モデルの DSP 1 社あたりの平均勝率を, それぞれ Table. 7, Table. 8, Fig. 11 に示す.

Table 7: 実験 2 の 1 広告枠あたりの平均落札額

モデル	平均落札額	標準偏差
正直 確率	0.143	3.73×10^{-4}
	0.143	4.66×10^{-4}

Table 8: 実験 2 の DSP 1 社あたりの平均落札数

モデル	平均落札数	標準偏差
正直 確率	149,000	355
	54,100	204

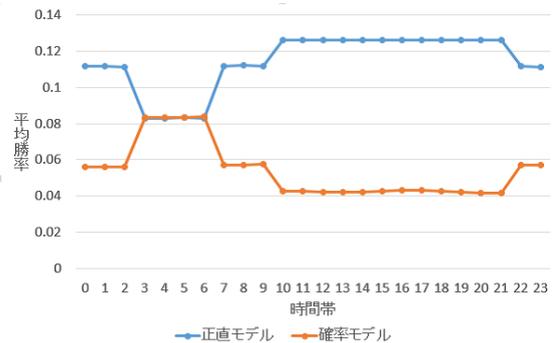


Fig. 11: 実験 2 における各モデルの時間帯別平均勝率

7.2.2 実験 2 の考察

まず, Table. 7 から正直モデルと確率モデルとでは 1 広告枠あたりの落札額は変わらないことが読み取れる. 確率モデルは案件をランダムに選ぶものであるが, 落札額自体は入札額が最大な案件を選ぶ正直モデルと変わらないことが分かった. 続いて, Table. 8 から正直モデルの方が確率モデルよりも, より多くの広告枠を落札できることがわかった. そして, Fig. 11 からは正直モデルと確率モデルとで勝率の差に 3 つのパターンが有ることが分かる. 1 つ目は 3~6 時帯, 2 つ目は 0~2, 7~9, 22~23 時帯, 3 つ目は 10~21 時帯である. これは表 5 に示した案件ごとのビッドする時間帯に依存していることが分かる. 1 つ目は 1 案件, 2 つ目は 2 案件, 3 つ目は 3 案件ビッドするものがある. ここから, ビッドする案件が増えるごとに確率モデルの勝率が下がっていくと考えられる. 正直モデルと確率モデルは入札案件を選ぶ点のみが違っていたが, これらの結果より, 入札案件が多くなるほど正直モデルの方が確率モデルよりも多くオークションに勝てるモデルと言える.

7.3 実験 3

次に $\frac{N}{3}$ の DSP が正直モデルに, $\frac{N}{3}$ の DSP が競り上げモデルに, 残りの $\frac{N}{3}$ の DSP がつり上げモデルに従い入札する場合でシミュレーションを行った. N は 12 に固定した.

7.3.1 実験 3 の結果

実験結果として, 各モデルの DSP 1 社あたりの 1 広告枠あたりの平均落札額及び平均落札数, そして各モデ

ルの DSP 1 社あたりの時間帯別平均落札額及び平均勝率を、それぞれ Table. 9, Table. 10, Fig. 12, Fig. 13 に示す。

Table 9: 実験 3 の 1 広告枠あたりの平均落札額

モデル	平均落札額	標準偏差
正直	0.188	6.43×10^{-4}
競り上げ	0.243	4.80×10^{-3}
つり上げ	0.279	5.09×10^{-4}

Table 10: 実験 3 の DSP 1 社あたりの平均落札数

モデル	平均落札数	標準偏差
正直	9.33×10^4	362
競り上げ	1.13×10^5	2280
つり上げ	9.85×10^4	17.9

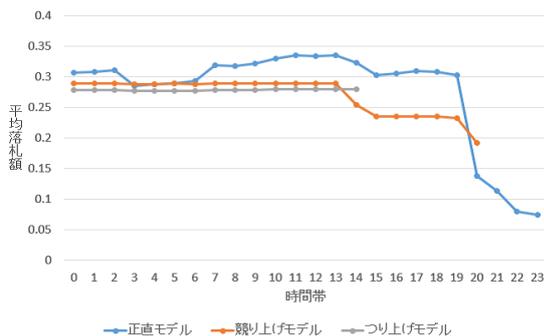


Fig. 12: 実験 3 における各モデルの時間帯別平均落札額

7.3.2 実験 3 の考察

まず Table. 9 から平均落札額の低い順に正直モデル, 競り上げモデル, つり上げモデルであることが分かる。この理由は、Fig. 12, Fig. 13 から分かる。つり上げモデルは 14 時間帯の時点で制限金額を使い果たしてしまっており、15 時間帯以降は入札できていない。すなわち、15 時間帯以降は多くとも 8 社の DSP のみでオークションが行われているということである。同様に競り上げモデルも 20 時間帯で制限金額を使い果たしてしまっており、21 時間帯以降は入札できていない。これより 21 時間帯以降は多くとも 4 社でオークションが行われていることになる。実験 1 で示したように、DSP の数が多くなればなるほど平均落札額は上がるので、時間帯が遅くなるにつれて平均落札額が下がっていることが、Fig. 12 から分かる。これより、遅い時間帯まで残っているモデルの方が平均落札額は低くなるのが分かる。

次に Table. 10 から平均落札数の多い順に競り上げモデル, つり上げモデル, 確率モデルであることが分かる。この理由も、Fig. 12, Fig. 13 から分かる。Fig. 13 から朝の時間帯はつり上げモデル, 昼頃は競り上げモデル, そして夜の時間帯は正直モデルの勝率が高いことが分かる。前述のとおり、つり上げモデル及び競り上げモデルは、平均落札額が高い時間帯に多くの金額を使ってしまい、遅い時間帯まで持たなかった。競り上げモデルは、つり上げモデルより平均落札額が低い時間帯に勝率が高かったため、落札数がつり上げモデルよりも多くなっていると考えられる。

これらより、一概にどの戦略が良いとは言えない。

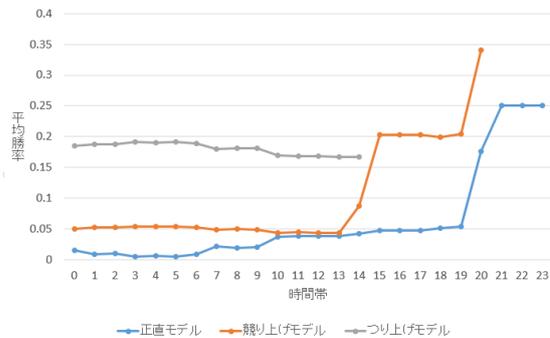


Fig. 13: 実験 3 における各モデルの時間帯別平均勝率

7.4 実験 4

最後に $\frac{N-1}{3}$ の DSP が正直モデルに、 $\frac{N-1}{3}$ の DSP が競り上げモデルに、 $\frac{N-1}{3}$ の DSP がつり上げモデルに、残り 1 つの DSP が学習モデルに従い入札する場合でシミュレーションを行った。(4) 式で用いるパラメータの値は $\alpha = 0.80$, ϵ -greedy 選択で用いるパラメータの値は $\epsilon = 0.30$ とした。そして報酬は、 $r_1 = 100$ と固定し、 r_2 は -1.00, -50.0 の 2 パターン行った。

7.4.1 実験 4 の結果

実験結果として、まず 3 パターンの r_2 における、各モデルの DSP 1 社あたりの 1 広告枠あたりの平均落札額及び平均落札数を、それぞれ Table. 11, Table. 12 に、各モデルの DSP 1 社あたりの時間帯別平均落札額及び平均勝率を、それぞれ Fig. 14, Fig. 15, Fig. 16, Fig. 17 に示す。

Table 11: 実験 4 の 1 広告枠あたりの平均落札額

r_2	モデル	平均落札額	標準偏差
-1	正直	0.230	1.18×10^{-3}
	競り上げ	0.246	4.12×10^{-3}
	つり上げ	0.281	3.14×10^{-5}
	学習	0.268	1.17×10^{-3}
-50	正直	0.247	3.60×10^{-3}
	競り上げ	0.245	3.42×10^{-3}
	つり上げ	0.281	3.71×10^{-5}
	学習	0.177	2.96×10^{-3}

Table 12: 実験 4 の DSP 1 社あたりの平均落札数

r_2	モデル	平均落札数	標準偏差
-1	正直	6.91×10^4	353
	競り上げ	1.12×10^5	1.92×10^3
	つり上げ	9.80×10^4	11.0
	学習	1.02×10^5	449
-50	正直	6.44×10^4	1.10×10^3
	競り上げ	1.12×10^5	1.60×10^3
	つり上げ	9.80×10^4	12.9
	学習	1.20×10^5	4.13×10^3

7.4.2 実験 4 の考察

Table. 11 から報酬 r_2 を -50.0 にした際、学習モデルの平均落札額が下がっていることが分かる。そして Table. 12 から報酬 r_2 を -50.0 にした際、学習モデルの平均落札数が多くなっていることが分かる。この点をさらに詳細に調べるため、Fig. 14, Fig. 15, Fig. 16, Fig. 17 を考察する。

Fig. 14, Fig. 15 より平均落札額の推移は、 r_2 が -1.00, -50.0 との間であまり変化は見られなかった。し

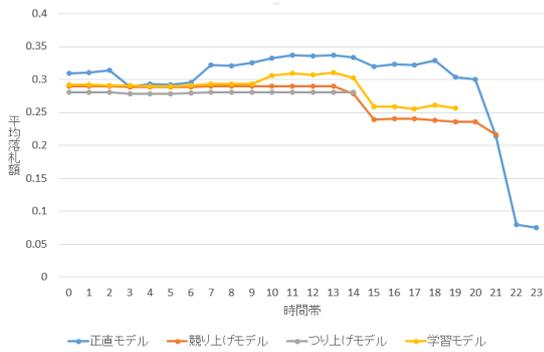


Fig. 14: 実験 4 における各モデルの時間帯別平均落札額 ($r_2 = -1.00$)

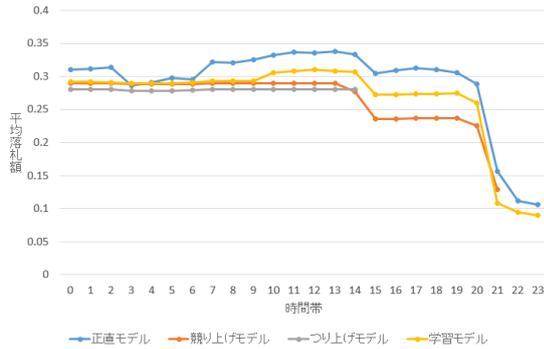


Fig. 15: 実験 4 における各モデルの時間帯別平均落札額 ($r_2 = -50.0$)

しかし、Fig. 16, Fig. 17 の平均勝率をしてみると $r_2 = -50.0$ における学習モデルの勝率が、 $r_2 = -1.00$ の時と比較して、特に 15~18 時の間で抑えられている事が読み取れる。これは $r_2 = -50.0$ としたため、落札額が 0.200 を超えると Q 値に大きく影響し、入札額が抑えられているからだと考えられる。そのため、 $r_2 = -50.0$ の際の学習エージェントは“降りる”の選択肢を多く使用していると考えられる。実際 $r_2 = -50.0$ の際の学習エージェントの Q 値を検証した結果、20 時間帯まで“降りる”の以外の選択肢の Q 値がマイナスになっていた。ここから、朝 (平均落札額が高い) の時間帯には勝率を抑え、夜 (平均落札額が低い) の時間帯に勝率を上げていく戦略が有用といえる。ただし、学習モデルを増やすと、同様の戦略の DSP が出てくることも考えられ、この点を今後検証していく必要がある。

8 結論

8.1 まとめ

本研究では、WEB 広告におけるオークション取引である、RTB を模したオークションのモデル化に取り組んだ。モデル化のモチベーションは、入札戦略に関して考察することであった。

モデル化にあたり、本研究では RTB のオークションにおいて入札者の立場である DSP に注目した。DSP の立場から見ること、特にオークション取引に限定したモデルとした。モデル化のために、既存のオークションとの比較や、実際に DSP のサービスを運営している企業へのヒアリングを行った。

モデルでは、実時間を考慮した広告枠の発生数や、ユーザーがアクセスしたドメインによる広告枠の発生数を設定した。そして、SSP から送られてくるビッド

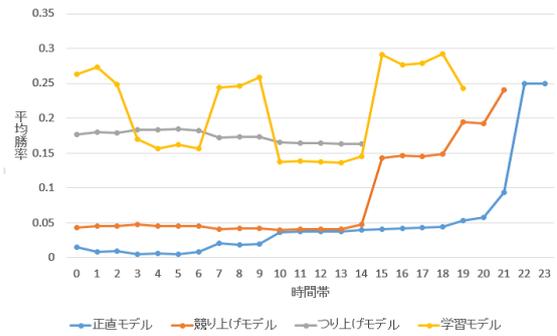


Fig. 16: 実験 4 における各モデルの時間帯別平均勝率 ($r_2 = -1.00$)

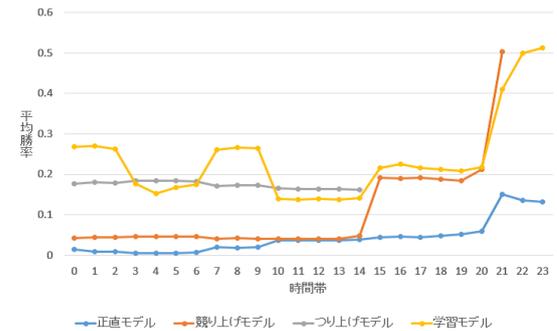


Fig. 17: 実験 4 における各モデルの時間帯別平均勝率 ($r_2 = -50.0$)

リクエストには時間帯、及びドメインの情報が含まれているものとした。DSP はそれらの情報から入札額及び入札案件を決定できるものとした。

さらに、DSP の行動モデルを作成し、その中でビッドポリシーを 5 つ作成した。1 つ目が過去のデータから入札額を決定し、入札額が最高のもを入札案件に決定する”正直モデル”、2 つ目が過去データから入札額を決定するが、入札案件は確率的に決める”確率モデル”である。3 つ目が勝利したら入札額を“0.010”下げ、敗北したら入札額を“0.005”を上げていく”競り上げモデル”、4 つ目が勝利したら入札額を“ $\frac{\text{入札額}}{2}$ ”下げ、敗北したら入札額を“ $\frac{0.300 - \text{入札額}}{2}$ ”上げていく”つり上げモデル”である。5 つ目が“正直モデル”、“競り上げモデル”、“つり上げモデル”及び“降りる”の、計 4 つの行動に対して強化学習を行い、時間帯及びドメイン別に最適な戦略を探しだそうとする“学習モデル”である。以上をオークションモデルに組み込み、取引の様子を分析した。

実験は 4 つ行った。実験 1 では N の DSP が正直モデルで入札する場合でシミュレーションを行った。DSP の数 N は 12 社と 24 社の 2 パターン行った。結果として DSP の数が増えるほど 1 広告枠あたりの落札額が増加する傾向があることが分かった。ここから、今後の課題として、RTB の市場全体をモデル化し、DSP と SSP との関係の理解も深めることが挙げられた。

実験 2 では $\frac{N}{2}$ の DSP が正直モデルに、残りの $\frac{N}{2}$ の DSP が確率モデルに従い入札する場合でシミュレーションを行った。 N は 12 社に固定した。結果として、正直モデルと確率モデルとでは 1 広告枠あたりの落札額は変わらず、オークションの勝率は正直モデルの方が高くなることがわかった。ここから正直モデルの方

が確率モデルよりも多くオークションに勝てるモデルと言えた。

実験3では $\frac{N}{3}$ の DSP が正直モデルに、 $\frac{N}{3}$ の DSP が競り上げモデルに、残りの $\frac{N}{3}$ の DSP がつり上げモデルに従い入札する場合でシミュレーションを行った。 N は 12 社に固定した。結果として、平均落札額の低い順に正直モデル、競り上げモデル、つり上げモデルとなった。この理由は、競り上げモデルやつり上げモデルは早い時間帯に制限金額を使い果たしてしまったため、結果として遅い時間帯の落札額が下がった。それ故、最後まで残った正直モデルが平均落札額で一番優秀という結果になった。しかし平均落札数では、多い順に競り上げモデル、つり上げモデル、確率モデルとなった。ここから、一概にどの戦略が良いとは言えない結果となった。

実験4では $\frac{N-1}{3}$ の DSP が正直モデルに、 $\frac{N-1}{3}$ の DSP が競り上げモデルに、 $\frac{N-1}{3}$ の DSP がつり上げモデルに、残り1つの DSP が学習モデルに従い入札する場合でシミュレーションを行った。 N は 13 社に固定した。結果として、報酬 r_2 が -50.0 にした際、学習モデルの平均落札額が下がっていることが分かった。この理由は、学習モデルの勝率が早い時間帯に抑えられている事だった。学習モデルは“降りる”選択肢を多く使い、朝の時間帯には勝率を抑え、夜の時間帯に勝率を上げていく戦略をとり、これは有用であった。ただし、同様の戦略の DSP が出てくることも考えられ、この点を検証していくことが今後の課題として挙げられた。

8.2 今後の展望

本研究では、RTB 型のオークション取引モデルの提案を行った。提案したモデルにおいて DSP の戦略を学習させた時、朝の時間帯は降りて夜の時間帯に集中して入札する戦略に、効果があることが分かった。ただし、他の DSP も同様な戦略をとった時に関して、検証していく事が今後の課題として挙げられた。

今後、SSP から送られてくるビッドリクエストに、ユーザーの情報を載せ、より現実の RTB に近いモデルを実現することや、DSP の戦略をより詳細に設定しモデルに組み込んでいくことで、さらに多くの知見が得られることに期待している。

最終的に SSP の要素もモデルに組み込むことで、RTB の市場全体のモデル化、さらには本モデル用いて入札戦略を策定していくサービスの実用化もできるのではないかと期待する。

謝辞

本研究を行うにあたり、Real-Time Bidding に関しての知見を頂きました、ソネット・メディア・ネットワークス株式会社様に深く感謝致します。

参考文献

- 1) 横山, 菅原, 榎田: DSP/RTB オーディエンスターゲティング入門 ビッグデータ時代を実現する「枠」から「人」への広告革命, インプレス R&D (2012)
- 2) 寺野: なぜ社会システム分析にエージェント・ベース・モデリングが必要か (ミニ特集 社会デザイン), 横幹, 4.2, 56/62 (2010)
- 3) アクセルロッド; 寺野: 対立と協調の科学, 64, ダイヤモンド社 (2003)
- 4) 海津研: 広告オークションにおける入札額変動の分析に基づく入札戦略の提案, (2012)
- 5) Ghosh, Rubinstein, Vassilvitskii, Zinkevich: Adaptive bidding for display advertising, Proceedings of the 18th international conference on World wide web, ACM, 251/160 (2009)
- 6) Cui, Zhang, Li, Mao: Bid landscape forecasting in online ad exchange marketplace, Proceedings of the 17th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 265/273 (2011)
- 7) Attenberg, Pandey, Suel: Modeling and predicting user behavior in sponsored search, Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining, ACM, 1067/1076 (2009)
- 8) ヤフオク!, <http://auctions.yahoo.co.jp/>
- 9) 坂井豊貴; 藤中裕二; 若山琢磨: メカニズムデザイン, ミネルヴァ書房 (2008)
- 10) Steiglitz: Snipers, Shills, & Sharks: eBay and Human Behavior, Princeton Univ Pr (2007)
- 11) Vickrey: Counterspeculation, auctions, and competitive sealed tenders, The Journal of finance, 16.1, 8/37, (1961)
- 12) 荒木, 松原, 櫻井: 1V-7 競り上げオークションと固定価格販売が混在する電子商取引市場における売り手の行動の解析 (マルチエージェント (1), 学生セッション, 人工知能と認知科学, 情報処理学会創立 50 周年記念), 全国大会講演論文集, 2, (2010)
- 13) 服部, 横尾, 櫻井, 新谷: 逐次型オークションの入札戦略決定手法: 準線形効用と予算制約の導入, 電子情報通信学会論文誌, DI 情報・システム I-情報処理, 85.10, 974/984 (2002)
- 14) 大竹, 西田: 強化学習を用いた交渉戦略学習アルゴリズム, 電子情報通信学会技術研究報告, AI 人工知能と知識処理, 101.210, 37/44 (2001)
- 15) 鍋田, 山本, 吉川, 寺野: RTB における競売参加者の戦略に関する研究, SIG-DOCMAS (2013)