

ディスプレイ広告における広告掲示系列の改善

Improving the Order of Posting Advertisements in Display Advertising

齋藤 陽介

Yosuke SAITO

松原 繁夫

Shigeo MATSUBARA

京都大学 情報学研究科 社会情報学専攻

Department of Social Informatics, Kyoto University

This paper proposes a new method for determining the order of posting advertisements to increase the conversion probability. Display advertisements have been traded by auction and attaining the high conversion rate within the pre-determined budget is required. However, previous studies have not paid a lot of attention to the sequence of different advertisements. This problem can be represented as Markov decision process but there are two problems, modeling the situation where the number of advertisements is enormous and dealing with the dynamic environment. To overcome these problems, we reduced the number of states of the model formulated as MDP using a method of clustering, and improved the order of posting advertisements increasing a conversion probability using a method of reinforcement learning. Finally, we showed that the developed method outperforms the simple method by experiments.

1. はじめに

ディスプレイ広告とは、Web 広告の形式の一種で、Web ページの一部として埋め込まれて表示される、画像や Flash、動画などによる広告を指す。その一部は、バナー広告と呼ばれることもある。これまで、Google AdWords に代表される検索連動型広告に関心が移っていたが、広告枠をオークションで売買する AdExchange の出現などにより、動画広告を含めてディスプレイ広告が再び注目を集めている [Goldfarb 11]。広告に関する技術はアドテクノロジーと呼ばれ、ビジネスという側面だけでなく技術面でも、人工知能を始め情報技術の新たな適用分野として注目を集めている [Ghosh 09, Goldstein 11, Balseiro 11]。

ディスプレイ広告に関しては、広告エコシステムとして、広告枠の売り手であるメディア、広告枠の買い手である広告主、広告枠の取引市場の運営者と様々なプレイヤーが存在している。本稿では、広告主、あるいは、広告主を支援する DSP (Demand Side Platform) と呼ばれるプレイヤーの立場に立って議論を行う。広告主の目的は、一定の予算の中で、最大のコンバージョンが得られるように、広告を掲示することである。ここでコンバージョンとは、閲覧者がディスプレイ広告をクリックして、広告主の web ページを訪れ、商品を購入する、資料を請求するといった広告主の目的が達成されることである。

広告枠取引市場である AdExchange でどのようなオークション方式が採られているようとも、広告主にとっては、広告掲示の価値を正確に評価することが必要である。過小に評価すれば、オークションに勝つことができなくなり、過大に評価すれば、オークションには勝つが、得られる効果に比べて出費が大きくなり、全体としての効率が低下する。しかし、広告掲示の価値を正確に評価することは簡単ではない。例えば、一人の閲覧者に同じ広告を何回も掲示すると、段々とコンバージョン率が低下することが知られている。

この掲示頻度の問題に対して、Jordan らは、現行の広告解析における、閲覧者の広告クリック率が、それまでにその広告を見た回数とは独立という仮定を問題点として指摘し、マルコフ決定過程 (Markov Decision Process, MDP) を用いて閲覧

者の行動をモデル化する方法を示している [Jordan 11]。ある商品の広告を閲覧者が見たときに、商品に興味を持ち、購入を検討し、他の類似品と比較し、実際に購入に至るといった購買の漏斗 (purchasing funnel) を考慮に含めようとの意図である。つまり、掲示回数が増えるにつれ、単調に割り引いて入札するのではなく、閲覧者の反応をより精密にモデル化しようとの試みである。Jordan らは MDP として定式化し、社会的効用最大化を目的として、いつ広告を掲示するかを決定する広告最適配置問題を議論している。

しかし、Jordan らは 1 つの広告の掲示回数の問題を扱うのみで、異なる広告に関する掲示系列の問題を議論していない。例えば、紅茶の広告を見た後で、ケーキの広告を見れば、コンバージョン率が上がるが、一方、ダイエットサブリの広告を見た後に、ケーキの広告を見れば、コンバージョン率が下がるかもしれない。ここでは、紅茶とケーキが一緒に買われやすいといった点への着目ではなく、紅茶の広告の次にケーキの広告といった順序に着目している点に注意されたい。

さて、このような異なる広告間の関係を考慮して掲示系列を決定しようとする場合、Jordan らのモデルの単純な拡張としては扱えない。そこには、以下の 2 つの解決すべき課題が存在する。

- 異なる広告間の関係を扱おうとすれば、扱う広告数が膨大となり、MDP として表現するのが困難になる。

Jordan らのモデルは個々の広告に着目しており、ある広告主とその他競合者の広告の二つに分けて、MDP として定式化している。しかし、本研究では広告全体を対象として広告掲示系列の改善を目的としているため、Jordan らと同じように MDP として定式化すると状態数が膨大となり、計算が非常に困難になる。

- 閲覧者の選好が経時的に変化する動的状況を適切に取り扱えない。

同じ商品であっても、クリエイティブと呼ばれる表示用画像が変われば、閲覧者の反応も変化と言われており、また、閲覧者自身の嗜好も変化するかもしれない。しかし、Jordan らのモデルでは、このような変化が生じた

ことが想定されておらず、この種の動的状況を適切に扱えない。

本研究では、上記2つの課題を解決し、高いコンバージョン率の達成が可能となるような広告掲示系列を得ることを目的とする。AdExchangeでのオークションでは組み合わせ入札は認められていない。そのため、掲示系列によって得られる期待報酬が計算できても、それを直接入札値として用いることはできないが、広告主に対する入札額決定の支援に使えると我々は考えている。

2. 広告掲示系列決定問題の定式化

本章では、まず、複数広告間の広告系列決定問題のMDPとしての定式化を示し、つぎに、前章で述べた1つ目の課題、すなわち、MDPモデルの中で膨大な数の広告をどう扱うかという課題を解決する方法について述べる。

2.1 基本モデル

本稿では、広告の数は多いかもしれないが、その数は有限であると考えて、離散時間の有限MDPを考える。広告 ad_k ($k = 1, \dots, n$) があるとし、閲覧者は各時刻に1つの広告を見たとする。また、閲覧者が現在見ているページのリンクをクリックしたことにより、つぎに表示されるページ $path_l$ ($l = 1, \dots, m$) があるとする。さて、本来広告のクリックとコンバージョンは区別して扱われるべきであるが、本稿では、議論を簡単にするため、クリックをコンバージョンとして扱う。つまり、広告のクリックが生じれば、それが広告主に何らかの利得を与えると仮定する。このとき、広告掲示系列の問題は、マルコフ決定過程 $\{X_t\}$ として以下のように表現される。

- 状態 $S = \{(add_1, path_1), (add_1, path_2), \dots, (add_n, path_m)\} \cup \{converted, drop_out\}$

状態は、現時点で掲示されている広告と、次に表示されるページの組、あるいは、コンバージョンが達成された状態 $converted$ 、閲覧者が当該メディアの閲覧を止めた状態 $drop_out$ のいずれかで表現される。

- 行動 $A(s_i) = \{ad_1, ad_2, \dots, ad_n\}$

状態 s_i で取り得る行動の集合には、 s_i に含まれる $path$ で許容されている広告掲示行動が含まれる。これは、例えば、これから掲示されるページが子ども向けのページである場合、酒類の広告掲示行動を選択肢から取り除かれることになる。

- 遷移確率 $p(s_j | s_i, a)$: 状態 s_i で、行動 a を実行したときに状態 s_j に遷移する確率を表す。

ここでは、閲覧者があるページを見ているとき、どのリンクをクリックするかの確率についても、得られていると仮定する。

- 報酬は、状態 $converted$ に到達したときに、1与えられ、それ以外のときには、報酬が得られないと仮定する。本来、広告主により、コンバージョンの価値は異なるが、議論を簡単にするため、一律に1であるとする。

遷移確率の中でも、状態 $converted$ に遷移する確率を、コンバージョン率 λ_{ad_i, ad_j} として表記する。これは、前回広告 ad_i が掲示されていた場合に、広告 ad_j を掲示する行動を取った場合に、コンバージョンに成功する確率を表す。

上記のモデルでは、直前の掲示広告と、現時点で掲示される広告によってコンバージョンされるかどうか決定すると考えた。これとは別の可能性について検討しよう。1つの可能性として、コンバージョンは現在掲示されている広告のみに影響され、過去に何が掲示されたかに影響を受けないという場合が考えられる。しかし、広告掲示頻度による影響は議論されており [Bronnenberg 98]、また、少なくとも、同じ広告が掲示され続けられれば、クリック率が低下することは経験的に知られている。よって、過去の掲示広告にまったく影響されないという考えは棄却できる。ただし、異なる広告間で、実際にどの程度の影響が生じるかは、実験等により確認する必要がある。

一方、別の可能性として、コンバージョンは直前の掲示広告だけではなく、さらに遡っての掲示広告にも影響されるという場合が考えられる。この場合は、例えば過去の k 回の掲示系列と、これから掲示されるページの組を状態と定義し直すことで、対処し得る。ただし、状態数の爆発はより深刻な問題となる。

さて、コンバージョンが直前の広告掲示と現時点での広告掲示のみに影響されると仮定しても、単純にMDPとして表現しようとするならば、広告数が増えた場合に、状態数が膨大となり、計算機での処理が困難になる。この問題を解決するため、我々は、クラスタリング手法を適用することを提案する。つまり、いくつかの類似した広告を1つの広告として扱うことにより、MDPの状態数を減らそうという考え方である。

2.2 クラスタリングによる表記量の削減

クラスタリングとは代表的なデータ解析手法のひとつであり、クラスタと呼ばれる部分集合にデータ集合を分割する方法である。クラスタリングには様々な手法が存在するが、本研究では実装の容易さの観点から、K-means法を採用した。ただし、広告間の近接度合いを直接計測するのは難しいため、分類対象である広告を特徴語のベクトルとして表現する。広告文からその広告の特徴ベクトル表現を得る方法として、本研究ではtf-idf法を用いた。

特徴語抽出からクラスタリングまでの流れを以下に示す。

1. 広告のタイトルと説明から特徴的な単語を抽出する。例えばタイトルと説明が以下のようなであれば、特徴語は「旅行、楽天、トラベル、予約」となる。

タイトル：旅行に行くなら楽天トラベル

説明：250万件のクチコミから比べて選べる！旅行の予約は楽天トラベル

2. tf-idf法を用いて各広告の特徴ベクトルを求める。
3. K-means法を用いて与えられた広告の特徴ベクトル集合を、いくつかのクラスタに分類する。

このように、複数の類似した広告をひとつの広告として扱うことにより、広告数が膨大な場合でも、状態数をある一定数に抑ええたままで、MDPとして表現することが可能となる。

3. 強化学習による動的環境への対処

1. で述べたように、閲覧者の選好が経時的に変化することが考えられる。広告主としては、このような動的環境に対処する必要がある。この問題に対して、本稿では強化学習手法の適用を提案する。

静的環境を仮定すれば、コンバージョン率を改善する広告掲示系列は、強化学習を用いなくとも求められる。しかし、実際はクリエイティブや閲覧者の嗜好の変化によって、広告への反応が変化すると考えられる。ここで、静的環境を仮定して得ら

れた広告掲系列では、この変化に対応できない。しかし、強化学習を用いて、つねに学習を続けることにすれば、その変化に追従できると期待できる。

強化学習とは、試行錯誤を通じて環境に適応する学習制御の枠組みである [Kaelbling 96]。学習主体であるエージェントの目的は利得を最大化する政策の獲得である。本研究ではいくつかある強化学習手法のうち、アルゴリズムの簡潔さを優先し Q 学習を用いた。Q 学習ではすべての行動を十分な回数選択すれば、行動選択方法（探査戦略）に依存せず、Q 値は最適な値に収束することが知られている。

4. 仮想環境の構築

本章では、クラスタリングと強化学習を用いた広告掲系列改善方法を評価するための実験方法について説明する。ただし、実際の広告配信を操作することは難しいため、仮想的な環境を構築し、その中で評価することとした。まず、環境構築のための行ったデータ取得方法について述べ、それから Q 学習を用いた広告掲系列改善方法の詳細について述べる。

4.1 データ取得方法

提案方法評価のためにはコンバージョン率 $\lambda_{i,j}$ のデータが必要である。しかし、実データを得ることは難しいため、以下の実験から得られたデータで代替することにした。

1. 被験者に Yahoo! ブログ (<http://blogs.yahoo.co.jp/>) 内のページに掲示されている 5 つのテキストの広告について、タイトルとその説明のまとめを記録させる。
2. 被験者には掲示された広告の中で、興味を持った広告があれば、それをクリックするよう指示し、どの広告をクリックしたかを記録させる。
3. ページを更新し、ステップ 1 と 2 を 10 回繰り返して行わせる。

上記の一連の操作から得られた広告集合を一つのデータとした。実験は 16 人の大学生に行ってもらい、合計で 16 セットのデータを収集した。1 試行当たり、5 個 \times 10 回でのべ 50 個の広告が掲示され、これが 16 セットあるため、計 800 個の広告が掲示されたことになる。ただし、重複して掲示される広告もあるため、広告の種類としては、354 個であった。800 個の広告中、クリックされたのは 60 個であった。得られたデータのうち、8 セットのデータを学習用に用い、残りの 8 セットを評価用に用いた。

さて、コンバージョンは、実際には、系列の影響を強く受けるものもあれば、影響が小さいものも含まれると考えられる。そのため、コンバージョン率 $\lambda_{i,j}$ を以下のように与えた。

$$\lambda_{ad_i, ad_j} = 1 - (1 - \lambda_{ad_j})(1 - \lambda'_{ad_i, ad_j})$$

ここで、 λ_{ad_j} と λ'_{ad_i, ad_j} は以下の式で表される。

$$\lambda_{ad_j} = \frac{\text{広告 } ad_j \text{ をクリックした総数}}{\text{広告 } ad_j \text{ を掲示した総数}}$$

$$\lambda'_{ad_i, ad_j} = \frac{\text{広告 } ad_i \text{ から広告 } ad_j \text{ に遷移して、広告 } ad_j \text{ をクリックした総数}}{\text{広告 } ad_i \text{ から広告 } ad_j \text{ に遷移した総数}}$$

問題の定式化での議論では、1 ページに広告が 1 つしか掲示されないと仮定していたが、実験では 1 ページに 5 つの広告が掲示されている。このため、5 つの広告を別々に扱い、前後のページの広告の組み合わせすべてが 1 回ずつ発生したものと仮定して、計算した。

表 1: 各クラスタを表す特徴語

クラスタ	広告数	特徴語
1	30	ウェディング, 結婚, 婚活, ペット, プリーダー
2	42	髪ドック, 育毛, 白髪
3	10	腹巻, 靴下, レッグウォーマー
4	35	カード, ローン, 銀行, 融資
5	69	ラーメン, チョコレート, 楽天, 旅行, 観光
6	25	紅茶, ワイン, コーヒー
7	43	住宅, 家, 物件
8	2	音楽, 療法, ストレス
9	53	英会話, TOEIC, 英語, 英検
10	45	パソコン, PC, 家具, オンライン

表 2: 学習結果

状態	最適行動
初期状態	クラスタ 3
クラスタ 1	クラスタ 3
クラスタ 2	クラスタ 6
クラスタ 3	クラスタ 6
クラスタ 4	クラスタ 3
クラスタ 5	クラスタ 3
クラスタ 6	クラスタ 3
クラスタ 7	クラスタ 3
クラスタ 8	クラスタ 6
クラスタ 9	クラスタ 3
クラスタ 10	クラスタ 3

また、得られた広告集合に対し、2.2 節で述べたクラスタリング手法を適用した。本実験ではクラスタ数を 10 と設定した。各クラスタを表す特徴語を表 1 に示す。

4.2 Q 学習の適用

MDP として定式化した閲覧者のモデルに Q 学習を適用し、広告掲系列を改善することを行う。学習用環境で、行動選択回数を 1,000,000 回として学習を行った。状態 *drop.out* への遷移確率は 0 と設定した。また、行動選択方法は ϵ -greedy 選択方法を用い、 $\epsilon = 0.1$ とした。

さて、Q 学習は、一方で、学習率や割引率などのパラメータ変化に敏感であると指摘されており、事例によってパラメータの最適値は異なると考えられる。本研究では議論を簡単にするために、学習率 α と割引率 γ を 0.1, 0.2, ..., 0.9 の中から決めることとした。一般的な強化学習では強化信号をできるだけ大きくするために、 γ を 1 に近い固定値にすることが多い [Dean 94]。そこで本研究では $\gamma = 0.9$ とした。 α は 0.1 から 0.9 まで変化させて、コンバージョン率（得られたコンバージョン総数を 1,000,000 回の行動選択回数で除算した確率）が最大のときの α を採用することとした。

10 回学習を行ってコンバージョン率の平均を求めた結果、 $\alpha = 0.1$ のとき最もコンバージョン率が高いことが分かった。そこで本研究では学習率 $\alpha = 0.1$ 、割引率 $\gamma = 0.9$ を採用し、このときの学習結果を用いることにした。ただし、学習率 $\alpha = 0.1$ 、割引率 $\gamma = 0.9$ として何度も学習を行ったところ、いくつかの状態の最適行動が一意に定まらなかった。そのため、学習を 10 回行い、その中で Q 値の平均が最も高い行動を最適行動とした。

表 3: 得られたコンバージョン率

手法	設定 1	設定 2
提案方法	0.2604	0.5103
比較方法	0.2702	0.2671

このようにして得た、学習結果を表 2 に示す。最適行動としてクラス 3 が多く選ばれているが、これはクラス 3 の状態でクラス 6 を掲示するときのコンバージョン率が高いためと考えられる。

5. 評価

本章では提案方式による広告掲示系列の改善の程度を評価する。まず、評価方法について説明し、つぎにその結果を述べ、考察を行う。

5.1 評価方法

評価用データを学習用データに加え、4.1 節と同様に環境を構築し、この環境下でどの程度広告掲示系列が改善されるかを評価した。比較対象として広告単体のコンバージョン率が最大の広告を掲示する方法を考える。すなわち、学習用環境における広告単体のコンバージョン率 λ_j が最大となる広告を繰り返し掲示する。学習用環境ではクラス 10 のコンバージョン率が最大であったため、クラス 10 に含まれる広告を繰り返し掲示することになる。本来であれば、予算面で各広告の掲示回数に制限があるが、今回の実験では無視した。一方、提案方法では、表 2 に従って広告を掲示した。

評価は以下の 2 つの設定に対して行った。

設定 1: 初期状態をランダムに設定し、次に掲示する広告のみを選択する。

設定 2: 初期状態をランダムに設定し、次に掲示する広告と、さらにその次に掲示する広告を選択する。

評価尺度としては、コンバージョン数を広告掲示回数で割った系列全体のコンバージョン率を用いた。

5.2 評価結果

評価結果を表 3 に示す。表中のコンバージョン率はそれぞれ評価用の環境における 10 回の試行の平均を示している。表 3 を見ると、設定 1 では、提案方法と比較方法とあまり差がないが、設定 2 では、提案方法が比較方法の結果を大きく改善できていることがわかる。これは、提案方法は目先のコンバージョンだけでなく、以降のコンバージョンも考慮しているためと考えられる。今回の実験では直前の広告のみ考慮したが、より長い系列を考慮するようにすれば、さらに先のコンバージョンも考慮した系列となるため、より高いコンバージョン率を与える広告掲示系列を得られると期待できる。

6. おわりに

本研究では、コンバージョン率を改善するディスプレイ広告掲示系列の決定法を提案した。ディスプレイ広告はオークションによる取引が始まっており、広告主からは、予算の範囲内で高いコンバージョン率を得ることが求められている。しかし、先行研究では、異なる広告の掲示系列に関して十分な注意が払われてこなかった。この問題はマルコフ決定過程としてモデル

化することができるが、その適用には 2 つの課題が存在する。1 つは、膨大な数の広告をどう扱うかという点であり、もう 1 つは閲覧者の変化など動的な環境にどう対処するかという点である。この課題に対し、我々はクラスタリングと強化学習を導入することで解決を目指した。初期的な実験結果ではあるが、単純な掲示広告決定法と比べて、提案方法では、高いコンバージョン率が得られることを確認した。

今後の課題には、掲示系列の変更によって、閲覧者がどの程度影響を受けるのか、より実際の環境で評価を行うことや、適切なクラス数数の発見、また、得られた系列からオークションでの入札戦略を生成することなどが含まれる。

謝辞

本研究は、日本学術振興会科学研究費基盤研究 (B) (22300052, 平成 22 年度 ~ 24 年度) の補助を受けた。

参考文献

- [Balseiro 11] Balseiro, S., Feldman, J., Mirrokni, V., and Muthukrishnan, S.: Yield optimization of display advertising with ad exchange, in *Proceedings of the 12th ACM conference on Electronic commerce*, EC '11, pp. 27–28, New York, NY, USA (2011), ACM
- [Bronnenberg 98] Bronnenberg, B. J.: Advertising Frequency Decisions in a Discrete Markov Process under Budget Constraint, *Journal of Marketing Research*, Vol. 35, No. 3, pp. 399–406 (1998)
- [Dean 94] Dean, T., Basye, K., and Shewchuk, J.: Reinforcement Learning for Planning and Control, in Minton, S. ed., *Machine Learning Methods for Planning*, chapter 1, pp. 1–20, Morgan Kaufmann Publishers Inc. (1994)
- [Ghosh 09] Ghosh, A., Rubinstein, B. I., Vassilvitskii, S., and Zinkevich, M.: Adaptive bidding for display advertising, in *Proceedings of the 18th international conference on World wide web*, WWW '09, pp. 251–260, New York, NY, USA (2009), ACM
- [Goldfarb 11] Goldfarb, A. and Tucker, C.: Online Display Advertising: Targeting and Obtrusiveness, *Marketing Science*, Vol. 30, No. 3, pp. 389–404 (2011)
- [Goldstein 11] Goldstein, D. G., McAfee, R. P., and Suri, S.: The effects of exposure time on memory of display advertisements, in *Proceedings of the 12th ACM conference on Electronic commerce*, EC '11, pp. 49–58, New York, NY, USA (2011), ACM
- [Jordan 11] Jordan, P., Mahdian, M., Vassilvitskii, S., and Vee, E.: The multiple attribution problem in pay-per-conversion advertising, in *Proceedings of the 4th international conference on Algorithmic game theory*, SAGT'11, pp. 31–43, Berlin, Heidelberg (2011), Springer-Verlag
- [Kaelbling 96] Kaelbling, L. P., Littman, M. L., and Moore, A. W.: Reinforcement learning: a survey, *Journal of Artificial Intelligence Research*, Vol. 4, pp. 237–285 (1996)