2H1-5

大規模リワード広告システムにおける 行動履歴と広告属性を利用したコンバージョン予測モデルの構築

Conversion Prediction in Large-Scale Reward Advertising System Based on User Access History and Properties of Ads

宮西　一徳 *1　　高野　雅典 *2　　吉田　岳彦 *3*†
Kazunori Miyanishi　Masanori Takano　Takehiko Yoshida

*1株式会社サイバーエージェント　アドテクスタジオ
AdTech Studio, CyberAgent Inc.

*2株式会社サイバーエージェント　技術本部　　*3ヤフー株式会社
Technical Department, CyberAgent Inc.　Yahoo! Japan Corporation

Serving ads that are relevant to user interests, or “re-marketing”, usually helps to improve the conversion rate of Internet Advertising systems. User interests are often extracted from user behaviors in the past, such as browsing history. In this research, we propose a conversion prediction method to find out ads that are relevant to a user for re-marketing. The proposed method combines a recommendation algorithm to recommend appropriate ads to users and a classification algorithm to reveal the ads that are likely to become a conversion. We apply the proposed method to a real-world reward advertising system and show that the method can significantly improve the conversion rate.

1. はじめに

近年，インターネット広告費が増加傾向にある．株式会社電通の発表では，2014 年のインターネット広告全体の広告費（媒体費＋広告制作費）は，前年比 12.1%増の 1 兆 519 億円である [電通 15].

インターネット広告は従来のメディア広告と比べ，特定のユーザにターゲティングすることが容易である．そのため，ユーザの好みに合った広告をレコメンドすることによって，広告効果を高めることが可能となる．事実，インターネット広告費のなかでも，データプラットフォームとアドテクノロジーにより広告配信を最適する運用型広告の媒体費は前年比 23.9%増で 5,106 億円と大きく伸びている（同 [電通 15] 参照）.

本論文で対象とするリワード広告とは，ユーザがアプリインストールや会員登録等の成果地点に達した（コンバージョン（以下 CV）した）段階で，仮想コインやポイント等のインセンティブが付与される成果報酬型の広告である．リワード広告におけるユーザ，メディア，広告主の関係を図 1 に示す．ユーザがメディアサイトやアプリ内で広告をクリックすると，広告主サイトやアプリダウンロードサイトへ遷移する．そこで商品購入や会員登録，アプリインストールなど広告によって設定された異なる成果地点に達する（＝コンバージョンする）と，リワード広告プラットフォームに成果が発生したことが通知される．この段階で，広告主に広告費を請求し，メディアへ掲載費が支払われる．さらに，ユーザに対してはメディア内で使用できるポイント等が付与される．リワード広告の領域でも上記同様に，ユーザの興味・関心に基づいた広告配信を行うことによって広告効果を向上させる（コンバージョン率（以下 CVR）を高める）仕組みが求められている．

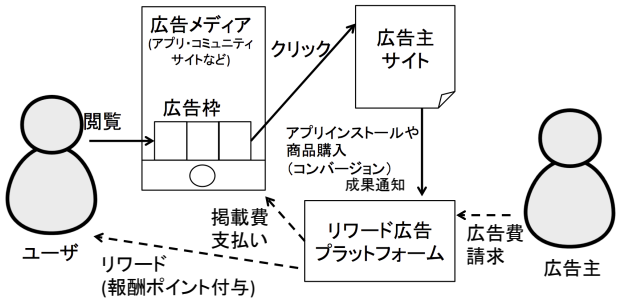

図 1: リワード広告の概要

そこで，本論文では，ユーザの行動履歴と広告属性に基づく CV 予測モデルを構築し，レコメンド広告枠において A/B テストを行い，提案モデルの効果を検証した．

本稿は以降，第 2. 章でレコメンド手法とクリック・コンバージョン (CV) 予測の関連研究を紹介し，既存研究でまだ解決されていない課題を説明する．第 3. 章で非負値行列因子分解 (NMF) による協調フィルタリングとロジスティック回帰の組み合わせでコンバージョン予測手法を提案する．第 4. 章で提案手法の評価実験とその結果について述べ，第 5. 章でまとめと今後の課題について説明する．

2. 関連研究

ユーザの行動履歴に基づく主要なレコメンド手法として協調フィルタリングがある [Goldberg 92]．応用事例として，購買履歴に基づいた商品のレコメンド [Schafer 99] や閲覧履歴から映画レコメンド [Koren 09a] に関する研究などがある．

大量のユーザとアイテム（広告）の組み合わせの中で CV するケースは少数に限られるため，単純な協調フィルタリングではスパース性の問題が発生する．これに対する手法として，行列因子分解 (Matrix Factorization(MF)) や非負値行列因子分解 (Non-Negtive Matrix Factorization(NMF)) などがある．これらは，高次元なデータの情報をできるだけ保持した状態で次元を削減する手法である [Lee 99]．リワード広告では，ポイントを獲得できるまでの時間や成果地点の種類（アプリインス

連絡先: 宮西　一徳，株式会社サイバーエージェント　アドテクスタジオ，Email: miyanishi_kazunori@cyberagent.co.jp
*†　株式会社 AMoAd 在籍時本研究を行った

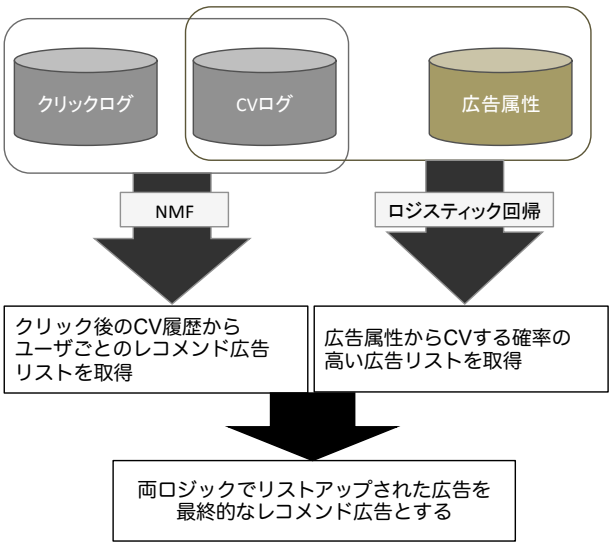

図 2: 提案手法の概要図

トールや会員登録．有料・無料など）といった広告属性の違いが CVR へ影響を及ぼすと考えられるため，ユーザベースでのレコメンドだけでなく広告の属性情報を予測モデルに組み込むべきである．

一方，属性情報に基づく予測・分類モデルとして，ロジスティック回帰を用いた研究がある [田頭 14, Rosales 12]．単純に属性情報のみで予測した場合，ユーザの興味・関心の違いを考慮しないため，全ユーザに対して同一のレコメンド結果となってしまう．本論文では，ユーザの興味・関心と広告属性の両面から，各ユーザごとに CV する確率が高い広告を予測するモデルとして，NMF による協調フィルタリングとロジスティック回帰を組み合わせたモデルを提案する．つまり，本手法では，ユーザの行動履歴による広告レコメンドとユーザや広告の属性による分類手法を組み合わせてコンバージョンを予測する．

3. 提案手法

本論文では，ユーザの行動履歴と広告属性を両方利用した CV 予測モデルを提案する．提案モデルの概要を図 2 に示す．まず，行動履歴 (CV ログ) からユーザ × 広告の CV 可能性を表す関連行列を作成し，その行列に非負値行列因子分解（NMF）を適用し，協調フィルタリングで，CV 可能性が未知の（ユーザ，広告）ペアに対して，CV 可能性のレコメンド値を求める．次に，ユーザと広告属性を利用して，ロジスティック回帰で CV 確率を予測する．上記の 2 つの指標を組み合わせて，あるユーザに対して，CV し易さで整列された広告リストを出力する．以下，提案手法の各段階を詳細に説明する．

3.1 協調フィルタリングによる CV 予測

本手法では，まず，ユーザの行動履歴を利用するために，ユーザ u が広告 a におけるコンバージョン予測問題をユーザ u に対する広告レコメンド問題として扱う．行動履歴から次の行動（商品を購入するかどうか）を予測する問題はよくレコメンドの問題として扱われている [Goldberg 92, Schafer 99, Koren 09a]．

具体的には，下記のように，クリックと CV のログから，一定回数以上クリックした履歴のあるユーザのみを対象として，式 1 のように，ユーザ × 広告ペアにおける CV の可能性を表す行列 $\mathbf{R}$ を作成する．

$$\mathbf{R} = \begin{array}{l} \\ \text{ユーザ} \downarrow \end{array} \overrightarrow{\text{広告}} \begin{pmatrix} r_{1,1} & r_{1,2} & \cdots & r_{1,n} \\ r_{2,1} & r_{2,2} & \cdots & r_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ r_{m,1} & r_{m,2} & \cdots & r_{m,n} \end{pmatrix} \quad (1)$$

行列 $\mathbf{R}$ の各要素 r_{ij} は下記の式 2 のように，CV したユーザ，広告ペアに対応するとき，値を 1，クリックのみで CV しなかったペアに対応する場合は 0 とする．

$$r_{ij} = \begin{cases} 1 & \text{ユーザ } i \text{ が広告 } j \text{ で CV した} \\ 0 & \text{ユーザ } i \text{ が広告 } j \text{ で CV しなかった} \end{cases} \quad (2)$$

行列 $\mathbf{R}$ を作成するときに，全広告に渡す合計クリックがある一定回数以上のユーザのみを対象にした理由はユーザが誤クリックなどのようなノイズを除去したいからである．ユーザ × 広告単位で考えると，クリック率 (CTR) やコンバージョン率 (CVR) は普通非常に小さい (1%程度) ので，上記の行列は疎行列である (クリックが発生していないユーザ × 広告に対応する行列の要素では，値が未知である)．そこで，この未知要素を予測するために，レコメンドアルゴリズムでよく用いられる非負値行列因子分解 (NMF [Lee 99]) 手法を適用し，協調フィルタリングを使い，各ユーザ × 広告に対してレコメンド値 (予測値) を計算する．NMF では，行列 $\mathbf{R}$ を下記のような 2 つの非負値行列 $\mathbf{W}^T$, $\mathbf{H}$ の積に近似する

$$\mathbf{R} \approx \mathbf{W^T H} = \hat{\mathbf{R}} \quad (3)$$

$\mathbf{W}$ は $k \times m$, $\mathbf{H}$ は $k \times n$ ($k \ll m, n$) 行列である．これにより，ユーザ u_i の特徴量は $\mathbf{W}_i$ ($\mathbf{W}$ の i 番目の列) に相当し，広告 j は $\mathbf{H}_j$ ($\mathbf{H}$ の j 番目の列) に相当する．$k \ll m, n$ なので，ユーザと広告の重要な特徴しか取られるように次元圧縮がされている．次元圧縮した後，元々の行列 $\mathbf{R}$ の要素 r_{ij} 近似値は $\hat{r}_{ij} = \mathbf{W}_i{}^T \mathbf{H}_j$ になるので，未知要素はこの値が予測値として入れることが出来る．

本研究では大規模の行列を扱うため，実装は Spark の MLlib ライブラリを利用している [MLlib 14, Koren 09b]．MLlib における行列 $\mathbf{R}$ の近似 (式 3) は下記の式を最小化する:

$$(r_{ij} - \hat{r}_{ij})^2 + \lambda(\|\mathbf{W}_i\| + \|\mathbf{H}_j\|) \quad (4)$$

上記の式の λ は $\mathbf{W}$, $\mathbf{H}$ の要素の値の影響を調整する係数である．式 4 は alternating least squares (ALS) という反復法で最小化出来る [Koren 09b]．

3.2 ロジスティック回帰による CV 予測

前節で説明したレコメンド手法は行動履歴を利用出来ているが，広告の属性（例えば，カテゴリなど）やユーザの属性（性別，年齢など）を利用するためには，複雑な行列分解アルゴリズムに組み込む必要であり，その影響が明示に調査しにくい．そこで，本手法では，その組み合わせを行列分解段階で行わず，CV 予測を別途で分類問題として扱う．

本システムでは，ある広告に対して関心を持ちクリックしたユーザに対して，類似のコンバージョン可能性が高い広告を配信したい．そのため，ユーザ情報とそのユーザがクリックした広告の情報に基づいて CVR を予測し，CVR の高い広告をク

リックされた広告と関連付けて配信する．広告のコンバージョン率を予測する問題では，既存研究でロジスティック回帰がよく用いられる [Rosales 12, Lee 12]．具体的には，下記の式 5 で表される確率を予測する

$$p(y|u,a) = \frac{1}{1+\exp\{-(b_0 + b_1 x_1 + b_2 x_2 ...)\}} \quad (5)$$

ここで，目的変数 y は過去の履歴で CV した場合を 1（正例）, しなかった場合を 0（負例）とする．u はユーザ属性，a は広告属性を表す．x_i は u, a の属性値，b_i は係数である．

各事例はユーザ属性と広告属性により特徴付ける．広告掲載メディア（サイトやアプリケーション）内で課金するアクティブなユーザほど広告にも関心を持つと考えられるため，ユーザ属性として前月のメディア内での課金額を使用する．広告の中には，アプリインストールでポイント獲得できるものや，サービスへの会員登録の申請を行ってから数週間程度の審査期間の後にポイント獲得できるものなど，ユーザにとってハードルの高さが大きく異なるものがある．このような特徴を CV 予測モデルに反映させるために，広告属性として広告カテゴリや課金種別などを使用する．使用した属性は以下の通りである．

- ユーザ属性
  性別，年齢，前月の課金額
- 広告属性
  広告カテゴリ（月額案件，インストール案件など），課金種別（有料，無料），その他属性（新着案件，ゲーム，グルメなど 45 種類）

3.3 NMF とロジスティック回帰の組み合わせによる CV 予測

NMF によるレコメンドでは，各ユーザのクリック，CV 履歴を使用してユーザの興味・関心に基づいた広告レコメンドを行うため，広告自体の特徴を捉えたレコメンドは難しい．一方，ロジスティック回帰による CV 予測では，ユーザ属性と広告属性の組み合わせから予測するため，ユーザの興味・関心に基づいたレコメンドができない．

提案手法では，両手法を組み合わせることによって，ユーザの興味・関心とユーザ・広告の属性を捉えたレコメンドを可能にする．ユーザごとに各広告のレコメンドスコア S を式 6 に示す通り，NMF で推薦する広告の予測評価値とロジスティック回帰での発生確率の和で定義する．

$$S(u,a) = r_{u,a} + p(1|u,a) \quad (6)$$

各ユーザに対して，スコア $S(u,a)$ の上位一定数の広告をレコメンドする．流れとしては，図 2 に示した通りである．

4. 実験

NMF とロジスティック回帰を組み合わせることによる有効性を検証するため，小規模のデータで予備実験を行った．その後，提案手法によるレコメンド配信を A/B テストで評価した．結果を表 1 に示す．

4.1 予備実験

NMF のみの場合，ロジスティック回帰のみの場合と両者を組み合わせる提案手法との精度比較を行う．1 日分のクリックログと CV ログを使用し，5-fold cross validation により精度の評価を行った．評価結果を表 1 に示す．ここで再現率，適合率は以下の通りである．

表 1: 各ケースの精度比較

手法	再現率	適合率	F 値
NMF	0.68	0.60	0.64
ロジスティック回帰	0.73	0.63	0.68
提案手法	0.54	**0.70**	0.61

$$\text{再現率} = \frac{\text{正しく } CV \text{ すると予測できた事例数}}{\text{実際に } CV \text{ した事例数}}$$

$$\text{適合率} = \frac{\text{正しく } CV \text{ すると予測できた事例数}}{CV \text{ すると予測した事例数}}$$

$$F \text{ 値} = \frac{2 \times \text{再現率} \times \text{適合率}}{\text{再現率} + \text{適合率}}$$

結果から，NMF とロジスティック回帰を組み合わせることで再現率, F 値ともに低下しているが，適合率は向上している．レコメンド配信では，CV すると予測した広告のうち限られた広告枠数分だけが配信に使われるため，ユーザが CV しそうな広告を漏れなくリストアップするよりも，CV の確率が特に高い広告のみを提示することが求められる．従って，適合率が高い提案手法はレコメンド配信に適している．

4.2 A/B テストによる評価

実際のレコメンド配信での A/B テストにより提案手法の有効性を評価した．A/B テストとは，同一の広告枠において異なる 2 種類の広告配信を行い，広告効果を比較するテストである．

A/B テストを行った広告枠は，ユーザが関心を持った広告をクリックした際に遷移する広告詳細ページに表示されるレコメンド広告枠である．この枠は，初めにユーザがクリックした広告（メイン広告）1 つに対して最大 6 つの広告をレコメンドする．提案手法の比較対象とするベースライン手法は，メイン広告と同じ広告カテゴリに含まれる広告の中で獲得ポイント数が近いもの（メイン広告で獲得できるポイントの ±25%）をクリック数等に基づく人気順で配信する．

ベースライン手法と提案手法は，ユーザの初回来訪時に無作為に選択して配信し，2 回目以降の来訪時には初回に選択されたベースライン手法か提案手法のどちらかの方法で配信を行う．この配信方式では，ユーザがランダムにどちらかの手法の配信対象になるため，同一のユーザに両方の手法で配信されることはない．提案手法では，ユーザの過去のクリックや CV の履歴を使用して予測を行うため，新規ユーザに対してはレコメンドをすることができない．この場合には，ベースライン手法と同様の配信方法によって広告を表示する．レコメンドの評価指標としては，CVR を用いる．

4.2.1 データセットと実験設定

過去 2ヶ月分のクリックログと CV ログを使用し，2015 年 2 月 20 日から 2 月 22 日の 3 日間 A/B テストを行った．NMF による手法では，2ヶ月間の間に 3 クリック以上したユーザを対象とし，ユーザ数 137,822, アイテム数（広告数）が 1,466 の行列を作成した．ユーザが一度 CV した広告は，同一ユーザに再度配信することがないため，NMF で生成した行列を基にして各ユーザに対して自分自身が CV していない広告をレコメンドする．ロジスティック回帰では，ユーザと広告の各組み合わせを 1 事例とする．NMF と同様に過去 2ヶ月間のクリックログと CV ログに基づいて，ユーザと広告の組み合わせで

CV した事例を正例，CV しなかった事例を負例として学習してモデルを生成する．NMF とロジスティック回帰を組み合わせた結果，ユーザごとにランク上位の広告を一定数レコメンドする．今回の A/B テストでは，式 6 で表すスコア $S(u, a)$ の上位 40 件をレコメンドするようにし，この中からランダムで選択した広告を 6 つのレコメンド広告枠にて配信する．配信時に配信不可となる広告もあり得るため，枠数に対して多めの最大 40 件をレコメンドする設定とした．

4.2.2 実験結果

ベースライン手法と提案手法の比較結果を表 2 に示す．

表 2: A/B テスト結果

手法	表示回数	クリック数	CV 数	CVR
ベースライン手法	45,738	1,730	533	0.308
提案手法	45,501	926	315	**0.340**

提案手法により CVR が 3.2%向上していることが分かる．さらに，提案手法では新規ユーザに対してレコメンドできないためベースライン手法と同様の配信に切り替えるケースがある．その内訳として，提案手法でレコメンドできないためベースライン手法と同等の配信を行った場合（ケース (1)）と，提案手法によりレコメンドできた場合（ケース (2)）に分けた結果を表 3 に示す．

表 3: 提案手法の内訳

ケース	表示回数	クリック数	CV 数	CVR
(1)	16,164	512	180	0.352
(2)	29,337	414	135	0.326

全ユーザの 64.5%程度に対しては提案手法で表示することができている．(1), (2) どちらのケースにおいてもベースライン手法の場合に比べて CVR が高くなっている．(1) はベースライン手法と同様の配信方法であるが，表 2 でのベースライン手法に比べて CVR が 4.4%高くなっている点から，提案手法を用いることでベースライン手法では CVR が低くなるユーザ層に対して CVR を改善できたと考えられる．

本論文では CVR の向上を目的としているが，表示回数に対するクリック数を比較すると，提案手法はベースライン手法に比べて 54%程度と低い値になっている．結果として，表示回数ベースでは広告効果が低くなってしまっているため，今後の課題として，表示回数に対するクリック数を改善することが挙げられる．

5. 結論

本論文では，非負値行列因子分解とロジスティック回帰を組み合わせたコンバージョン予測モデルを提案した．予備実験により組み合わせることで適合率が向上することを確認し，実際の配信で A/B テストを行うことでコンバージョン率が向上することを確認した．今後の課題としては，今回の A/B テストでは NMF とロジスティック回帰の組み合わせ結果のランク上位 40 件の中からランダムに選択した広告を配信したが，ランキングに従い上位の広告から順番に配信することが考えられる．これにより，提案手法のモデルにより広告効果が高いと判定した広告から確実に配信することが可能となる．また，表示回数に対するクリック数の改善や新規ユーザに対するレコメンド手法の確立も今後の課題として考えられる．

参考文献

[Goldberg 92] Goldberg, D., Nichols, D. A., Oki, B. M., and Terry, D. B.: Using Collaborative Filtering to Weave an Information Tapestry, *Communications of the ACM*, Vol. 35, No. 12, pp. 61–70 (1992)

[Koren 09a] Koren, Y.: Collaborative Filtering with Temporal Dynamics, in *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2009, Paris, France*, pp. 447–456 (2009)

[Koren 09b] Koren, Y., Bell, R. M., and Volinsky, C.: Matrix Factorization Techniques for Recommender Systems, *IEEE Computer*, Vol. 42, No. 8, pp. 30–37 (2009)

[Lee 99] Lee, D. D. and Seung, H. S.: Learning the parts of objects by non-negative matrix factorization., *Nature*, Vol. 401, No. 6755, pp. 788–791 (1999)

[Lee 12] Lee, K.-c., Orten, B., Dasdan, A., and Li, W.: Estimating Conversion Rate in Display Advertising from Past Erformance Data, in *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD '12, pp. 768–776, New York, NY, USA (2012)

[MLlib 14] MLlib, S.: Collaborative Filtering – Spark MLlib Documentation, http://spark.apache.org/docs/1.2.1/mllib-collaborative-filtering.html (2014)

[Rosales 12] Rosales, R., Cheng, H., and Manavoglu, E.: Post-click Conversion Modeling and Analysis for Non-guaranteed Delivery Display Advertising, in *Proceedings of the Fifth ACM International Conference on Web Search and Data Mining*, WSDM '12, pp. 293–302, New York, NY, USA (2012)

[Schafer 99] Schafer, J. B., Konstan, J. A., and Riedl, J.: Recommender Systems in E-Commerce, in *Proceedings of the First ACM Conference on Electronic Commerce. EC99*, pp. 158–166 (1999)

[田頭 14] 田頭 幸浩, 小野 真吾, 田島 玲：オンライン広告における CVR 予測モデルの素性評価, in *The 6th Forum on Data Engineering and Information ManagementDEIMs Forum 2014* (2014)

[電通 15] 電通 株式会社電通：「2014 年 日本の広告費」は 6 兆 1,522 億円、前年比 102.9 %, http://www.dentsu.co.jp/news/release/2015/0224-003977.html (2015)