

クロスドメインの情報推薦に関するサーベイ

深澤佑介 太田順

東京大学 人工物工学研究センター

yusuke.fukazawa@gmail.com

概要 本稿ではクロスドメインの情報推薦(クロスレコメンド)に関するサーベイ結果を報告する。クロスレコメンドの定義を試み、クロスレコメンドの研究課題について述べる。課題解決に向けた関連研究を協調フィルタリング、コンテンツフィルタリング、ハイブリッドのアプローチの3種類の方法に分類しそれぞれ概観する。

キーワード クロスドメイン、レコメンド、情報推薦

1 はじめに

人々のニーズは様々である。単一のサービスではニーズの一部しか捕捉できない。人々の様々なニーズを捉え、個々人に適したコンテンツを配信するため、様々なドメインを横断するレコメンド・広告が提供されている。たとえば、Google AdSense に代表されるコンテンツ連動広告、Amazon.com・楽天に代表される物販広告、Yahoo・RevenueScience に代表される行動ターゲティング広告がある。Google AdSense では、業態、業種、ターゲットユーザも異なる様々なドメインのサービスの広告が混在している。内容に基づくレコメンド方式により、様々なドメインの中から現在閲覧しているページと関連するコンテンツを出力している。Amazon.com では書籍、音楽、家電など様々なドメインの物販サービスを提供している。協調フィルタリングによりドメイン間で横断的なレコメンドを実現している。行動ターゲティングとはユーザの履歴情報を AdNetwork 上で一元管理しておき、AdNetwork に参加している様々なサイトでパーソナライズされた広告・コンテンツを配信する。このように、クロスドメインでのレコメンド(以下クロスレコメンド)はビジネス領域においても非常に重要な課題である。本稿ではクロスレコメンドに関する最先端の研究を紹介する。

2 クロスレコメンドの定義

クロスレコメンドとは、目的のサービスとは異なるサービスで蓄積された履歴を用いて目的のサービス上でレコメンドを行う手法である。図1に従来のレコメンドとクロスレコメンドの違いを示す。従来のレコメンドは、ユーザの履歴対象のサービスとレコメンド対象のサービスは一致している。一方、クロスレコメンドでは、前者と、後者は異なる。以下、履歴対象のサービスをソースドメイン、レコメンド

対象のサービスをターゲットドメインと呼ぶ。ドメインの定義については文献により異なる。ビジネス的な観点ではサービス単位がドメインとなるのが自然である。Li ら[1][2]の研究では、書籍、映画など全く異なるドメイン間のレコメンドーションを行っている。一方、Berkovsky ら[3]は、同一ユーザの複数ドメインにわたる履歴データの入手が困難であることを理由として、代替手段として映画の複数ジャンル(SF、ロマンス)を複数ドメインとして扱っている。

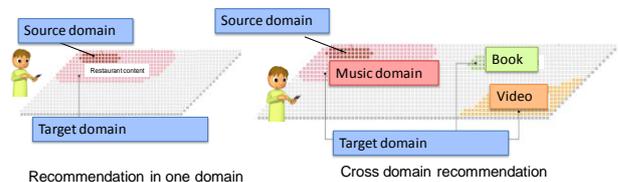


図1 従来レコメンドとクロスレコメンドの違い

3 クロスレコメンドの研究課題

クロスレコメンドでは、レコメンドの主要な手法である、協調フィルタリング方式およびコンテンツフィルタリング方式の二つの手法の適用が可能である。協調フィルタリングでは、あるユーザに対してソースドメインでの購入傾向の近いユーザを探し、当該ユーザがターゲットドメインで購入したコンテンツを推薦することが可能になる。一方、コンテンツフィルタリングでは、ソースドメインで構築したユーザプロフィールをターゲットドメインでの情報推薦に利用する。しかしながらクロスドメインに特有の問題として以下の課題がある。

- 協調フィルタリングの課題: 一般に複数のドメインにまたがってアイテムの評価履歴を持つユーザは少ない。そのため、既存の協調フィルタをそのまま適用すると履歴の集中している各ドメイン内のレコメンドに偏る傾向がある。

- コンテンツフィルタリングの課題: コンテンツ

フィルタリングでは、ユーザおよびコンテンツを特徴量で表現(プロファイリング)する必要があるが、異なるドメイン間で共通する特徴量が少ない場合、レコメンドの精度が著しく低下する恐れがある。

4 クロスレコメンドの動向

4.1 協調フィルタリングベース

Li ら[1][2]は、ソースドメインターゲットドメイン間のユーザの嗜好の共通性に関して以下の仮説をおくことで上記の課題を解決している。具体的には、ソースドメインの特定の Item 集合に対して評価の高い User 集合は、ターゲットドメインの特定の Item 集合に対しても評価を高くつけやすい、逆もまた然り、という仮説を設定している。まず、ソースドメインにおける User-Item マトリクスにおいて、評価の高い領域、低い領域が集中するように行および列を並び替える。次に、ターゲットドメインの User-Item マトリクスの評価の高い領域、低い領域が、ソースドメインの並び替えられたマトリクスを近似するように行と列を並び替える。そして、ターゲットドメインの User-Item マトリクスの未知の部分についてそれが評価の高い領域にある場合評価を高くし、評価が低い領域にある場合評価を低くする。

4.2 コンテンツフィルタリングベース

Loizou ら[4]は、複数のドメインにまたがる知識を Wikipedia を用いて獲得し、共通する特徴量の少なさをカバーしている。著者ら[5]は、名詞と動詞から構成されるタスクのオントロジを構築し、それを複数ドメインの共通の特徴として用いる手法を提案している。

4.3 ハイブリッドアプローチ

クロスレコメンドでは、上記の双方の課題を解決する手段としてハイブリッドのアプローチが最も多く提案されている。Berkovsky ら[3]は特徴が共通するコンテンツ集合はドメインが異なったとしても似たような評価を得る、という仮説のもと、以下のハイブリッドアプローチを提案している。1) すべてのコンテンツに対して特徴量ベースでクラスタリングを実施し、特徴が類似するコンテンツ集合を複数作成する。2) 分割後のコンテンツ集合とそれに対するユーザの評点をベースに User-Item マトリクスを構築。各クラスター内で協調フィルタリングを実施し、レコメンドを出力する。Berkovsky らの評価結果は大変興味深い。以下の3つの手法を MAE (二乗誤差) で比較している。**Standard**:分割なしで全サービスを横断的に協調フィルタリングを実施、**Local**:ターゲットドメインでの User-item 行列だけを利用して協調フィルタリングを実施、**Remote-average** (提

案手法) : **Local** の評価+**Local** のコンテンツと類似する他の User-item マトリクスの協調フィルタリングの結果を平均化して出力。その結果、提案手法が最も精度がよかったが、**Standard** と **Local** では、**Standard** のほうが多くの履歴を考慮しているにも関わらず、**Local** のほうが精度が高いという結果であった。これは、履歴が多くなるからといていたずらにクロスレコメンドを実施するのではなく、ドメインにあった履歴がある場合はその履歴を利用したレコメンドのほうが有効である、ということを示唆している。

中辻ら[6]は、協調フィルタリングの上記の課題をハイブリッド方式で解決している。既存の共通の購入アイテムの類似性に加え、1) ソーシャルネットワークのつながり、2) タクソノミーによるユーザの興味表現、類似度計算を導入している。Mustafa ら[7]は、デモグラベースレコメンド、協調フィルタリングとコンテンツフィルタリングを独立に実装しているが、現在のユーザに関するデータからの3種類のレコメンドの中で最も精度の高いクロスレコメンドが可能な方法をルールベースで選ぶ機構を導入している。

5 今後のクロスドメインの課題

今後、クロスドメインのレコメンドでは以下の課題が重要になると考えられる。様々なオンラインサービスの増加に伴い有益なクロスレコメンドが可能なドメインとそうでないドメインの選別が必要になる。また、サービスの個人最適化がさらに進むことを鑑み、ユーザの実世界行動、ソーシャルネットワーク上の行動等を利用したクロスレコメンドの可能性の検証が必要になると考えられる。

参考文献

- [1] B. Li, Q. Yang and X. Xue: Can Movies and Books Collaborate? Cross-Domain Collaborative Filtering for Sparsity Reduction, IJCAI, 2009.
- [2] B. Li, Q. Yang, and X. Xue: Transfer learning for collaborative filtering via a rating-matrix generative model, ICML, 2009.
- [3] S. Berkovsky, T. Kuflik and F. Ricci: Cross-Domain Mediation in Collaborative Filtering, User Modeling, 2007.
- [4] A. Loizou: How to recommend music to book lovers: Enabling the provision of recommendations from multiple domains, PhD thesis, University of Southampton, 2009.
- [5] Y. Fukazawa, J. Ota: User-centered Profile Representation for Recommendations across Multiple Content Domains, International Journal of Knowledge-Based Intelligent Engineering Systems, Vol. 15, No. 1, pp.1-14, 2011.
- [6] 中辻 真, 藤原 靖宏, 内山 俊郎: ユーザグラフ上のランダムウォークに基づくクロスドメイン推薦, 人工知能学会論文誌, Vol. 27, No .5, 2012.
- [7] A. Mustafa, B. Aysenur: CrosSing Framework: A Dynamic Infrastructure to Develop Knowledge-Based Recommenders in Cross Domains, International Conference on Web Information Systems and Technologies (WEBIST'10), 2010.