

# アクセスログを用いたクロスドメイン環境における情報推薦

堤田 恭太<sup>1,a)</sup> 中辻 真<sup>1</sup> 内山 俊郎<sup>1</sup> 戸田 浩之<sup>1</sup> 内山 匡<sup>1</sup>

**概要:** ポータルサイトや EC サイトでは様々な異なる種類のコンテンツが扱われており、アクセスログは複数のドメインを横断した形で存在している。こうしたログを活用することで、事業者は集客力のあるサービスドメインの利用者を他のドメインへ誘導して新規利用者の増加等を見込むことができ、利用者は通常訪れないドメインからも興味のあるアイテムを発見できるようになると期待される。しかし、こうしたクロスドメインの推薦には、(1) ユーザが普段利用しないドメインには利用履歴がないため推薦が難しい、(2) 複数ドメイン間でログの規模が異なる不均衡状態があり、ログの多いドメインのアイテムが提示されやすい、(3) データの疎性が促進されて推薦精度が低下しやすい、という問題がある。そこで本研究は、データの疎性に強い Random Walk with Restart による推薦手法をベースとして、アクセスログを用いて構築するグラフ構造をドメイン間の規模の不均衡の問題に対処して複数ドメインに対応させる手法を提案する。評価実験では、大規模なアクセスログの実データを用いて state-of-the-art な情報推薦の手法を含む様々なアプローチと比較し、提案法の有効性を示した。

## Cross-domain Recommendation Using Web Access Logs

**Abstract:** We propose a graph-based recommendation method that deals with multiple types of contents domain by using large scale web access logs. Recently, some portal sites and content providers have begun to offer various kinds of contents, however, few studies have been done on the recommendation across multiple domains. At the same time, it is much difficult that users find the interesting items from unknown markets by themselves. Therefore, there is a great need for the research on cross-domain recommendation. In this paper, we present a novel random-walk based method: an extension of random walk with restarts (RWR), which is a robust approach to data sparseness. Considering the imbalance in service scale, we propose a graph construction methodology for RWR to realize cross-domain recommendation. After that, we describe experiments on large scale real access logs and our results show that proposed method outperforms state-of-the-art collaborative filtering methods.

### 1. はじめに

ユーザが興味を持つアイテムを推定して利用者に提示する推薦システムは、サービスの利便性や売上の向上を目的として、Amazon.com<sup>\*1</sup>等の様々な商用の web サイトで利用されている [28]。推薦の手法についても幅広く研究がなされており、特に、協調フィルタリング (CF) [11], [14], [22], [23] によるアプローチは、被推薦ユーザの興味を利用傾向の似たユーザの情報を用いて推定して推薦を行う方法としてよく知られている。

また最近では、複数サービスをまたがったクロスドメ

イン環境での推薦が注目を集めている [4], [17], [18], [19], [20], [26]。クロスドメイン環境の代表例としては、複数サービスを提供するポータルサイトや、スマートフォンを利用して複数のサービスを使う環境が挙げられる。このようなクロスドメイン環境を対象とした研究としては、大きく分けて二つの種類がある。一つは、データの疎性に対処する目的で異なるドメインから得られる統計量を用い、推薦システムの評点の予測を高精度に行うことを目指した研究 [18], [19], [20] である。もう一つは、特定のドメインのユーザの利用履歴から、他のドメインでのアイテムの好みを予測する研究 [4], [26] である。前者は例えば、音楽の好みが利用履歴等から得られているユーザに対して、おすすめ音楽をより高精度に推定するために、ユーザの映画や書籍の好みを利用する研究であり、ユーザが既に利用しているドメインでの推薦精度を改善する研究である。後者は

<sup>1</sup> 日本電信電話株式会社 NTT サービスエボリューション研究所  
NTT Service Evolution Laboratories, NTT Corporation

<sup>a)</sup> tsutsumida.kyota@lab.ntt.co.jp

<sup>\*1</sup> <http://www.amazon.com/>

例えば、音楽の好みを利用履歴等から得られているユーザに対して、おすすめの映画や書籍を推薦することを旨とした研究であり、ユーザが既に利用しているドメインの情報を元に、異なるドメインのアイテムを推薦するための研究である。

特に後者については、事業者、ユーザの両者にとってメリットは大きいと考えられる。事業者側のメリットとしては、既存サービスの利用者に新たなサービスのアイテムを推薦する事で新サービスの立ち上げ、ビジネスの拡大に貢献することが考えられる。また、エンドユーザにとってみれば、これまで利用していないサービスからもアイテムの推薦を受ける事ができ、思いがけない発見につながる事が期待される。

そこで、本研究ではポータルサイト等複数のサービスを提供する環境におけるクロスドメイン推薦に取り組む。ポータルサイトにおけるアイテムの推薦方法としては、ポータルサイトの TOP ページ等において数件をタイトルとともに提示する形が一般的と考えられるため、推薦結果上位の推薦精度を向上させる事を旨とする。

クロスドメイン環境の推薦では、先行研究と実サービスのアクセスログを調査したところ、次の3つの問題があることが分かった。

#### (1) ユーザが未利用ドメインでの推薦

ユーザが普段利用しないドメインには、そのユーザの利用履歴がないため推薦がより難しい [10], [24].

#### (2) 複数ドメイン間でログの規模が異なる不均衡状態

特に、扱うドメインが多数になる場合、推薦システムはユーザによりログの多いドメインのアイテムを提示しやすく、結果として推薦精度が低下することがある。

#### (3) データ疎性の促進

多くのユーザは多数あるドメインの中から特定の2, 3つを利用し、その他のドメインはほとんど利用しないことが多いため、結果としてクロスドメイン環境のアイテムとユーザのログがより疎になる。また、その結果として推薦精度が低下する傾向がある。

そこでアプローチとして、データの疎性に強い Random Walk with Restart (RWR) [7], [13], [16] による推薦手法をベースとして、アクセスログを用いて構築するグラフ構造を複数ドメインに対応させることを考える。具体的には、ユーザとアイテムの関係を表す重み付き有向グラフにおいて、サービスの規模を考慮して規模の大きいサービス内のエッジの重みを軽くすることで、複数ドメイン間のログの不均衡の問題に対処する。

評価実験では、state-of-the-art な手法を含む様々なアプローチについて、性質の異なる複数のデータにおいて推薦精度を検証し、提案法の有効性を確認した。特に、アクセスログを用いたクロスドメイン環境での検証に、ポータル

サイト\*2の大規模な実データを用いた。また、評価のためのテストデータには、3000名を超えるモニタユーザによる人手の5段階評価のデータを用いた。

また、前述のような利用目的では top-k を考慮した評価が必要なため、各手法の推薦精度の評価には、IR や情報推薦の分野においてアイテムのランキングの精度評価に用いられる normalized Discounted Cumulative Gain (nDCG) [12] を用いた。

本研究の貢献をまとめると以下ようになる。

- (1) アクセスログを用いたクロスドメイン環境において高精度に推薦を実現する手法を提案した。
- (2) 定量評価実験において、性質の異なる複数のデータセットを用いて、state-of-the-art な推薦手法を含む種々のアプローチについて実験を行い、提案法の有効性を検証した。

本稿の構成は以下の通りである。2章では、クロスドメイン推薦およびグラフベースの推薦手法についての関連研究について述べる。3章では、提案法について述べる。4章では、検証用のデータセット、定量評価に用いた推薦精度の評価指標、比較手法、実験結果といった評価実験について述べる。5章では、本研究のまとめを述べる。

## 2. 関連研究

本章では、2.1 節でクロスドメイン環境での推薦に関する関連研究について述べ、2.2 節でグラフを用いた推薦手法の関連研究について述べる。

### 2.1 クロスドメイン推薦

複数のドメインやメディアを扱うクロスドメイン環境での推薦システムが研究されている [4], [17], [18], [19], [20], [26], [27]. クロスドメイン推薦には、大きく分けて2つの種類がある。一つは、データの疎性に対処する目的で異なるドメインから得られる統計量を用い、推薦システムの評点の予測精度を改善する研究 [18], [19], [20] である。もう一つは、特定のドメインのユーザの利用履歴から、他のドメインでのアイテムの好みを予測する研究 [4], [26], [27] である。前者は例えば、音楽の好みを利用履歴等から得られているユーザに対して、おすすめの音楽をより高精度に推定するために、ユーザの映画や書籍の好みを利用する研究であり、ユーザが既に利用しているドメインでの推薦精度を改善する研究である。後者は例えば、音楽の好みを利用履歴等から得られているユーザに対して、おすすめの映画や書籍を推薦することを旨とした研究であり、ユーザが既に利用しているドメインの情報を元に、異なるドメインのアイテムを推薦するための研究である。本研究は主に後者の研究に取り組む。

\*2本研究では、コンテンツ系サービスのドメインを対象とした

柳原ら [27] は、書籍や音楽等の様々なメディア（ドメイン）の利用履歴が、個人の携帯端末によって紐付けられる状態を想定し、クロスメディア型の推薦として提案がなされている。実験では、MovieLens のジャンルの違いをメディアの違いとみなして、ジャンルの組合せについて共通するユーザの数と推薦精度の関係について詳細に調査した。Cremonesi ら [4] は、ドメイン間での共通するユーザの割合等によって、一般的な協調フィルタリングや特異値分解 (SVD) による手法等がどのようなパフォーマンスを示すかを検証している。堤田ら [26] は、ドメイン間で共通するユーザやメタデータが全く存在しない場合に、blog 上でのアイテムの共起関係を用いてクロスドメインの推薦を実現する方法を提案した。

## 2.2 グラフベースの推薦手法

グラフを用いた情報推薦では、まず、ユーザやアイテムなどをノード、それらの関係の有無をエッジで表し、関係の強さをエッジの重みとして表す重み付き有向グラフを構築する。次に、そのグラフ上を被推薦ユーザを起点として Random Walk または、RWR を用いてアイテムへの関連度を表す到達確率を計算し、それに基づいてアイテムをランキングする。グラフ上を値が次々に伝播し、次第に減衰することに特徴があり [7]、特に、疎なデータを用いた推薦のタスクにおいて有効なことが知られている。

Yildirim ら [25] は、アイテムをノードとし、アイテムの類似度をエッジの重みとして持つグラフ上で被推薦ユーザが持つアイテムを起点として Random Walk し、被推薦ユーザへの推薦の予測値を計算することで、グラフの sparsity に起因する推薦精度の低下を抑え、従来のメモリベースの推薦システムよりも推薦精度を改善した。Gori ら [7] は、アイテム間の関連度を計算し、アイテムのみをノードとするグラフを構築し、そのグラフ上で RWR を適用することでアイテムをランキングする ItemRank と呼ばれる手法を提案し、他のグラフベースの推薦手法と比べて高い精度となることを示した。Konstas ら [13] は、ソーシャルネットワーク上のユーザ間の友人関係やユーザによってアイテムに付与されたソーシャルタグ等を扱うためのグラフ構築手法を提案した。彼らはユーザ、アイテム、タグの任意の組合せについて、図.1 のような関連性を表す 9 つの部分行列を構築し、それらを結合したグラフを用いることで推薦精度を改善した。また、ItemRank で用いられたアイテムグラフと比べて、グラフを疎行列や隣接リスト等を用いて表現する事で、消費メモリ量が抑えられるというメリットがある。この性質によって、アイテム数が膨大な場合にも手法が適用可能となっている。

これらの手法はデータの疎性に強いという利点があるが、複数のドメインを想定したものでないため、例えば、他よりも著しくログの規模が大きいドメインが存在するな

	User	Item	Tag
User	UU	UI	UT
Item	IU	0	IT
Tag	TU	TI	0

図 1 Konstas ら [13] による SNS を表す隣接行列。U, I, T はそれぞれユーザ、アイテム、タグを表し、例えば、UI はユーザ-アイテム行列を表す。0 は関係性の定義されていない領域を表し、0 行列となる。

表 1 本研究で用いる変数の一覧

変数	説明
$UI$	ユーザ-アイテム行列
$IU$	アイテム-ユーザ行列
$U^x$	ドメイン $x$ のユーザ集合
$I^x$	ドメイン $x$ のアイテム集合
$D^{x,y}$	ドメイン $x$ のユーザとドメイン $y$ のアイテムの関係を表す部分行列
$A$	グラフを表す隣接行列
$S$	扱うドメインの数
$N$	グラフを表す隣接行列の次元数
$K$	ノード $j$ から出るエッジの集合
$\alpha$	RWR のリスタートパラメータ
$a$	被推薦ユーザ
$p$	起点ノードからの到達確率を表す $1 \times N$ のベクトル
$q$	$a$ を表す値を 1、それ以外に 0 を持つ $1 \times N$ のベクトル

ど、ログの不均衡がある場合には推薦精度が低下することがあると考えられる。

## 3. 提案法:

### Cross-Domain RWR (CDRWR)

本章では、提案法となる Cross-domain RWR (CDRWR) について述べる。3.1 節では、提案法で用いるドメインの不均衡の問題に対処したグラフを表す隣接行列の構築手順について述べる。3.2 節では、隣接行列を用いてアイテムをランキングする推薦手法であり、データの疎性に強い特徴がある RWR について述べる。

#### 3.1 グラフの構築手順

本節では、提案法が扱うグラフの構築手順について述べる。

まず、扱うドメインのすべての組合せについて、多くの

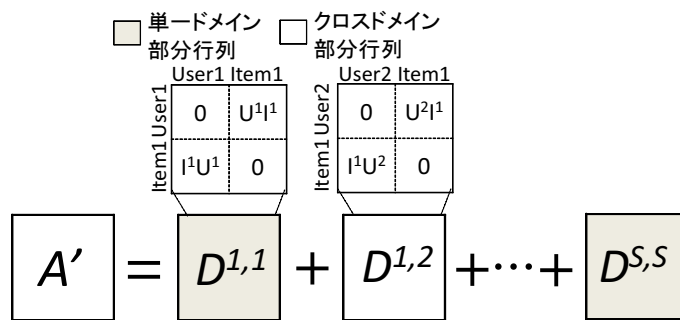


図 2 提案法で用いる隣接行列の、正規化前の部分行列の和を求める処理の概念図。影付きの領域は単一ドメインの部分行列、他の領域はクロスドメインの部分行列を表す。

小さく疎な部分行列  $D$  で表現されるグラフ構造を構築する。ドメイン  $x$  のユーザとドメイン  $y$  のアイテムの関係を表す部分行列を  $D^{x,y}$  とする。  $x$  と  $y$  が同じドメインを表すとき、一般的な単一ドメインでのグラフの隣接行列表現と同じになる。

部分行列  $D^{x,y}$  におけるユーザとアイテムの関係は双方向の有向グラフで表され、ユーザからアイテムへのエッジを表すユーザ-アイテム行列  $U^x I^y$  と、アイテムからユーザへのエッジを表すアイテム-ユーザ行列  $I^y U^x$  とから構成される。ユーザ-アイテム行列の要素  $U^x I^y_{u,i}$  は、ドメイン  $x$  のユーザを  $u$ 、ドメイン  $y$  のアイテムを  $i$  とするとき、  $u$  が  $i$  を利用履歴に持つ場合に 1、持たない場合に 0 とする。

$$D^{x,y} = \begin{pmatrix} 0 & U^x I^y \\ I^y U^x & 0 \end{pmatrix}. \quad (1)$$

さらに、それぞれの部分行列は列の値の合計が 1 になるように正規化する。正規化後の行列を  $D^{x,y}$  とおくと、あるドメイン  $x$  のあるノード  $j$  からドメイン  $y$  のあるノード  $k$  へのエッジの重みは以下の式で表される。

$$D^{x,y}_{j,k} = D^{x,y}_{j,k} / \sum_{k \in K} D^{x,y}_{j,k} \quad (2)$$

ここで、推薦システムが扱うドメインの数を  $S$  とするとき、結果として  $S \times S$  個の部分行列  $D$  を得る。

前述のように、 $D$  の正規化処理をドメインの組合せ毎に個別に行うことで、ドメイン間でデータサイズが極端に異なるといったドメインの不均衡による影響を減らすことができる。また別の視点では、ユーザが何アイテムも利用するサービスのドメインと、そうでないドメインとの 1 クリック（アクセスログの場合）の重要度に差を付けていると考えることができる。例えば、Q&A ドメインではユーザは多くのページを閲覧する傾向にあるが、EC のドメインで商品を買うのは 1 つだけ、といった状況を想定している。この例の場合、個別の正規化処理により、1 ユーザから多くのアイテムに向けてエッジが出やすい Q&A ドメインでは、エッジの重みが他のドメインに対して相対的に軽くなる効果が得られる。

最後に、全ての部分行列  $D$  について、同一のユーザ、同一のアイテムは同じ行列のインデックスを表す様に配置し、式 3 のように部分行列の和を求め、

$$A' = \sum_n \sum_m D^{n,m}. \quad (3)$$

その行列を  $A'$  とする。その行列  $A'$  に  $D$  と同様にして列の和が 1 になるように式 4 の正規化処理を施し、提案法の RWR の計算で用いる一つの隣接行列  $A$  を得る (図 2)。

$$A_{j,k} = A'_{j,k} / \sum_k A'_{j,k}. \quad (4)$$

### 3.2 Random Walk with Restart (RWR)

提案手法では、前節で示した行列で表現されるグラフ構造に対して RWR を適用することで、推薦アイテムを特定する。そこで本節では、RWR[6], [9], [15] について述べる。

RWR は、Personalized PageRank (PPR) [21] としても知られる、グラフ上の起点となるノードからの各ノードの関連度を計算するアルゴリズムである。通常の Random Walk と異なり、各ステップにおいてリスタートパラメータと呼ばれる一定の確率  $\alpha$  に従って起点ノード  $a$  に値が付与される点に特徴がある。起点ノード  $a$  を被推薦ユーザとすることで、個人化を施したアイテムのランキングが可能になり、上位にランクされたアイテムを推薦アイテムとするアルゴリズムである。

RWR をベースにした推薦システムは、以下の二つの仮説に基づいて推薦するアイテムを決定する。

- (1) 被推薦ユーザ  $a$  とグラフ上での距離が近いアイテムは、同じアイテムを持つユーザが持つアイテム、といった協調フィルタリングによる推薦の性質を持ち、ユーザがより興味を持ちやすいと考えられる (個人化の側面)
- (2) PageRank の様に、グラフ上で多くのノードからエッジが張られているアイテムは、多くのユーザによって支持されており、被推薦ユーザ  $a$  も興味を持ちやすいと考えられる (人気度の側面)

リスタートパラメータと呼ばれる定数  $\alpha \in [0, 1]$  を用いて、

RWR はこれらの 2 つの側面の影響度合いを制御する。α の値が大きければ、モデルはより強く個人化されたものと考えることができ、逆に小さければ、PageRank の様な人気度の側面が大きいモデルになる。

被推薦ユーザを表す起点ノードを  $a$ 、起点ノード  $a$  から各ノードの到達確率を表す  $1 \times N$  のベクトルを  $\mathbf{p}$ 、前述のグラフを隣接行列として表したものを  $\mathbf{A}$ 、 $a$  に対応する要素を 1、それ以外を 0 とする  $1 \times N$  のベクトルを  $\mathbf{q}$ 、RWR のハイパーパラメータである定数  $\alpha$  として、式 5 を収束するまで繰り返し更新することにより、起点ノードから全てのノードへの関連度を表す到達確率を計算する。最終的に、アイテムを表すノードについて、 $a$  からの関連度に基づいてアイテムをランキングする。

$$\mathbf{p} = (1 - \alpha)\mathbf{A}\mathbf{p} + \alpha\mathbf{q} \quad (5)$$

## 4. 評価実験

本章では、実験で用いたデータセット、評価尺度、比較手法、実験結果について述べる。

### 4.1 データセット

評価実験では、アクセスログを用いたクロスドメイン環境での推薦について、次の 4 つの疑問を明らかにするためのデータセットを用意した (表 2, 表 3, 表 4)。

- 疑問 1 : 既存の推薦システムとの比較  
どのような推薦システムのアプローチが、サービスの規模が異なる不均衡な複数ドメインを扱い、データが疎なアクセスログを用いたクロスドメイン推薦に適しているのか。この検証は、本研究のメインタスクである CrossDomain データセットの結果について、手法間の比較を行うことで明らかになる。
- 疑問 2 : アクセスログのデータの疎性による影響  
推薦システムの学習に用いるデータが疎なアクセスログの場合には、どういったアプローチの推薦システムが適するのか。この検証は、後述の MovieLens データセットと Q&A データセットの結果を比較することによって明らかになる。
- 疑問 3 : クロスドメインのデータの不均衡による影響  
ドメイン数が単一の場合と複数の場合とで、データの不均衡の影響は実験結果に影響するのか。この検証は、Q&A データセットと CrossDomain データセットとの結果を比較することによって明らかになる。
- 疑問 4 : クロスドメインのドメインの利用経験による影響  
クロスドメイン環境での推薦において、ユーザの各ドメインでの利用経験による影響はあるか。例えば、ユーザが学習データ収集期間内に利用していたドメインと、未利用のドメインとで推薦システムの精度に

表 2 実験に用いたデータセットの種類

データセット	学習データ	ドメイン
MovieLens	密な評点	単一
Q&A	疎なアクセスログ	単一
CrossDomain	疎なアクセスログ	複数

表 3 Q&A および CrossDomain データセットにおけるテストデータ内のタスク量の内訳。ユーザ 3029 名が最大 40 の評価を付与しているが、出現頻度による閾値のため量が少量となっている。

テストデータセット	全体	利用	未利用
Q&A	39,381	14,599	-
CrossDomain	94,236	28,942	5,985

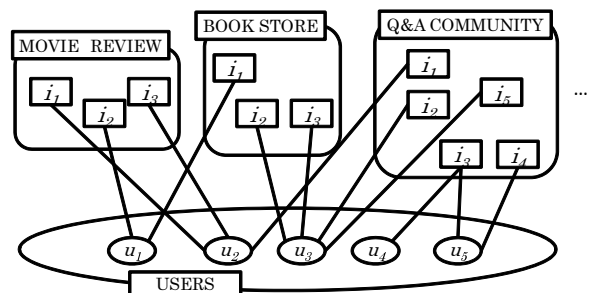


図 3 複数のドメインを横断したポータルサイトのアクセスログのイメージ図

差があるか。この検証のため、テストデータセットをユーザの利用経験を元に 3 つのタスクに分割した (表 3)。これらのタスク間での推薦精度の比較を行うことで明らかになる。

#### 4.1.1 MovieLens (密な評点/単一ドメイン)

MovieLens データセット<sup>\*3</sup> は、推薦システムの定量評価に広く用いられているデータセットである。このデータセットには、943 人のユーザによる人手の明示的な 5 段階の評点が 1682 の映画に対して、100,000 件付与されている。本実験では、学習データが密な、単一ドメインのデータセットとして用いる。

#### 4.1.2 CrossDomain (疎なアクセスログ/クロスドメイン)

CrossDomain データセットは、本研究のメインタスクとなるポータルサイトのアクセスログを用いたクロスドメイン推薦の評価用データセットである。多くのユーザは複数のサービスドメインを利用しており、ドメイン間に共通するユーザを媒介して、ユーザが未知のドメインにおいても推薦が可能である (図 3)。

ポータルサイトのアクセスログの中から Q&A ドメイン、映画評判ドメイン、音楽ドメインなど、ポータルサイト内の 35 のサービスドメインを対象として抜き出したものを用意し、推薦システムの学習データセットとした。ユーザ・アイテム行列の密度は、MovieLens データセットが 0.057

<sup>\*3</sup> <http://www.grouplens.org/node/73>

表 4 各データセットごとのユニークユーザ数, ユニークアイテム数, 評点/アクセスログ数, およびユーザ-アイテム行列の密度.

データセット	ユニークユーザ	ユニークアイテム	評点/アクセスログ	行列の密度
MovieLens (学習データ)	943	1,682	90,570	0.057
MovieLens (テストデータ)	943	1,129	9,430	-
Q&A (学習データ)	543,323	362,581	1,072,538	5.44e-6
Q&A (テストデータ)	1,881	11,840	39,381	-
CrossDomain (学習データ)	543,323	559,291	39,956,136	1.32e-4
CrossDomain (テストデータ)	3,029	49,318	94,236	-

なのに対して 1.32e-4 とかなり疎なデータとなっている。テストデータについては、3029 名のモニタユーザが一定期間の学習データ収集期間の後、ユーザ毎に最大 40 件のポータルサイト内コンテンツ (Q&A ページ等) を提示して、興味があるかどうかを 5 段階で評価したデータをこのデータセットのテストデータとして用いた。

また、比較手法のデータ規模の制約のため、実験ではアクセスログ中の出現頻度が 30 件以上のユーザとアイテムを対象として用い、ユニークユーザ数約 54 万人、ユニークアイテム数約 56 万件となった (表 4)。

さらに、CrossDomain データセットのテストデータは、ユーザのドメイン利用経験による推薦システムの推薦精度の違いを検証する目的で、以下の 3 つのデータセットに分かれている。それぞれの規模は表 3 の通りである。

#### (1) 利用ドメイン (利用)

推薦システムの学習に用いられたトレーニングデータ中に、ユーザの利用があったサービスのアイテムについて、推薦精度を測るテストデータセット。例えば、Q&A サービスと映画の評判サービスを利用しているユーザに対し、好む他の Q&A 記事もしくは映画評判記事を推薦する場合の精度を検証する。

#### (2) 未利用ドメイン (未利用)

推薦システムの学習に用いられたトレーニングデータ中に、ユーザの利用がなかったサービスのアイテムについて、推薦精度を測るテストデータセット。例えば、Q&A サービスを利用したことがないユーザに対して、他の映画の評判を扱ったサービスの利用傾向からそのユーザが他の Q&A 記事よりも映画関連の Q&A 記事を好むだろうと推測して提示できるかを検証する。

#### (3) データセット全体 (全体)

上記 2 つに加えて、ユーザの利用経験を考慮せずに提示されたアイテムを含むテストデータセット。

#### 4.1.3 Q&A (疎なアクセスログ/単一ドメイン)

前述の CrossDomain データセットの中から、Q&A ドメインの利用履歴のみを抜き出したデータセットである (表 4)。テストデータについても、同様に Q&A ドメインに関するものだけを抜き出して用い、このデータセットのテストデータとして用いた。テストデータセットの内訳は表 3 の通りである。

MovieLens データセットの結果との比較によって、疎なアクセスログを用いた場合の推薦について検証が可能となり、また、CrossDomain データセットの結果との比較によって、複数ドメインと単一ドメインの推薦の違いを検証することができる。

## 4.2 推薦精度の評価尺度

一般に、推薦精度の評価尺度は大きく二種類に分けることができる。一つは、ユーザがアイテムに対して明示的に付ける評点を予測する問題であり、正解データとなる実際に付けられた評点と予測した評点との予測誤差を計測する問題である [3]。この場合、平均絶対誤差 (MAE) や平均二乗誤差 (RMSE) がよく利用される。もう一つは、ユーザが興味を持つであろうアイテムをより上位にランキングする問題である。正解がユーザの興味の有無などの 2 値分類で与えられる場合には、top-k を提示した場合に正解を含む割合を表す Precision@k や、平均適合率 (AP) が用いられ、ユーザによる 5 段階評価等で与えられる場合には nDCG[12] が用いられる。

本研究では、前述の通り top-k を考慮した評価が必要であり、ユーザによる 5 段階評価のある正解データのうち、評点の高いアイテムをいかに上位にランクできるかを評価するために nDCG を用いた。

## 4.3 比較手法

本節では、比較実験を行った提案法を含む 7 つの推薦手法について述べる。各手法は、予測した評点や確率の値に基づいてテストデータのアイテムの評点を並べ換えたリストを返し、nDCG を用いて推薦精度を比較する。

(1) **CF (Pearson)** は、アイテム・ベースの協調フィルタリング [23] による推薦手法である。今回の実験では、アイテム間の類似度としてピアソン係数を用い、類似度が正になる組合せのみを用いた。

(2) **LDA** は、LDA[2] による推薦手法である。モデルの学習にはギブスサンプリング [8] を用い、また、トピック数は 10, 50, 100, 150 として実験し、最も精度の高くなる 100 を用いた。

(3) **PureSVD**[4] は、ユーザ-アイテムの評点や利用履歴を表す行列を特異値分解 (SVD) することに基づく

表 5 MovieLens データセットにおける nDCG の値

methods	nDCG
CF(Pearson)	0.927
LDA	0.918
PureSVD	0.950
SVD++	<b>0.951</b>
PageRank	0.910
RWR/CDRWR (提案法)	0.920

state-of-the-art な推薦手法の一つである [1], [4], [5]. ユーザ-アイテム行列において利用履歴や評点のないものは 0 とし [4], SVD の rank は, 10, 50, 100, 150 の場合について実験して, 最も精度の高くなる 100 を用いた.

- (4) **SVD++**[14] は, PureSVD をより発展させた SVD に基づく state-of-the-art な推薦手法である. ユーザが付けた評点や, アイテムに付けられた評点の情報を暗黙的なフィードバックとして用いることで, PureSVD に対してさらに精度を向上させている. SVD の rank は, 10, 50, 100, 150 の場合について実験し, 最も精度の高くなる 100 を用いた.
- (5) **PageRank**[21] は, ユーザ毎の個人化を考慮せずに, PageRank の値に基づいてアイテムをランキングする手法である. RWR を用いる推薦手法は 3.2 節で述べた通り, PageRank によるアイテムの人気度の側面と個人化の側面の両方を考慮するが, その人気度の側面のみによる推薦精度を検討するために用意した. PageRank のスケールパラメータは, RWR と CDRWR (提案法) に合わせて 0.1 を用いた.
- (6) **RWR**[13] は, Konstas らによって提案された部分行列の結合によってグラフを表す隣接行列を構築し, そのグラフ上で RWR を適用してアイテムをランキングする手法である. なお, RWR の個人化の度合いを表すパラメータ  $\alpha$  は, 0.1 毎に 0.1 から 0.9 の場合について実験し, 最も精度の高くなる  $\alpha = 0.1$  を用いた.  $\alpha$  の値が小さいため, 個人化の側面よりもより人気度の側面を重視する方法であると考えることができる.
- (7) **CDRWR (提案法)** は, 3 章で述べた複数のサービスの存在を考慮した提案手法である. RWR と同様に, パラメータは  $\alpha = 0.1$  を設定した.  $\alpha$  の値が小さいため, 個人化の側面よりもより人気度の側面を重視する方法であると考えることができる. また, 単一ドメインのデータセットにおける実験では, Konstas らの RWR と同様の手法となるため, 実験結果の表では RWR/CDRWR (提案法) というように併記した.

#### 4.4 実験結果

各比較手法の, MovieLens データセットにおける結果を表 5, Q&A データセットにおける結果を表 6, CrossDomain

表 6 単一ドメインの Q&A データセットにおける nDCG の値

methods	利用	全体	マクロ平均
CF (Pearson)	0.974	0.938	0.956
LDA	<b>0.976</b>	0.938	0.957
PureSVD	0.950	0.942	0.946
SVD++	0.950	0.942	0.946
PageRank	0.975	0.940	0.958
RWR / CDRWR (提案法)	<b>0.976</b>	<b>0.943</b>	<b>0.960</b>

表 7 CrossDomain データセットにおける nDCG の値

methods	利用	未利用	全体	マクロ平均
CF (Pearson)	0.913	0.955	0.895	0.921
LDA	0.914	0.956	0.890	0.920
PureSVD	0.915	0.957	0.897	0.923
SVD++	0.915	0.957	0.898	0.923
PageRank	0.914	0.956	0.894	0.921
RWR	0.921	0.961	0.912	0.931
CDRWR (提案法)	<b>0.925</b>	<b>0.962</b>	<b>0.914</b>	<b>0.934</b>

データセットにおける結果を表 7 に示す. 表中の太字は各データセットにおいて最も高い推薦精度となったものの値を表している.

これら 3 つの実験結果の表から分かることは, 次の通りである.

- (1) CrossDomain データセットにおける結果 (表 7) から, 提案法が各手法の中で最高精度となり, アクセスログを用いたクロスドメイン推薦に適していることが分かった. また, RWR と CDRWR との比較により, ドメイン間のデータ量の不均衡の問題に対処することで推薦精度が改善することが分かった (4.1 節の疑問 1).
- (2) SVD を用いた手法 (PureSVD, SVD++) が, 従来の知見通り MovieLens の実験では最高精度となった (表 5) が, データの疎性が強いアクセスログを用いた実験では十分な精度が得られなかった (表 6, 表 7). これにより, データが疎なアクセスログの場合は, SVD 系の手法より, LDA やグラフベースの手法が優れることが分かった (4.1 節の疑問 2).
- (3) 単一ドメインの Q&A データセットにおいては, RWR と CDRWR は個人化を行わない PageRank とに近い推薦精度だった (表 6) が, CrossDomain の実験では個人化によって大きく精度を改善することが分かった (表 7). (4.1 節の疑問 3).
- (4) ユーザの利用経験による違いに関しては, CrossDomain データセットにおいて RWR と CDRWR は他の手法と比べ, 利用, 未利用の値に対して全体の結果が良くなっている (4.1 節の疑問 4). これは, 実験結果の簡単な観察では, RWR と提案法と比べて他の手法が利用ドメインのアイテムを未利用ドメインのアイテムより上位にランクする傾向があったことに起因して

いるように見受けられたが、こういったドメインの組合せにおいて起きているのか等のより詳細な分析が必要である。

## 5. まとめ

本研究では、データの疎性に強い Random Walk with Restart による推薦手法をベースとして、アクセスログを用いて構築するグラフ構造を複数ドメインに対応させる手法を提案した。評価実験では、大規模なアクセスログの実データを用いて state-of-the-art な情報推薦の手法を含む様々なアプローチと比較し、特に、クロスドメイン環境での推薦精度において提案法が有効であることを示した。

## 参考文献

- [1] Barbieri, N. and Manco, G.: An analysis of probabilistic methods for top-n recommendation in collaborative filtering, *Machine Learning and Knowledge Discovery in Databases*, pp. 172–187 (2011).
- [2] Blei, D. M., Ng, A. Y. and Jordan, M. I.: Latent Dirichlet Allocation, *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022 (2003).
- [3] Chen, W., Chu, J., Luan, J., Bai, H., Wang, Y. and Chang, E.: Collaborative filtering for orkut communities: discovery of user latent behavior, *Proceedings of the 18th international conference on World wide web*, ACM, pp. 681–690 (2009).
- [4] Cremonesi, P., Tripodi, A. and Turrin, R.: Cross-domain recommender systems, *Data Mining Workshops (ICDMW), 2011 IEEE 11th International Conference on*, IEEE, pp. 496–503 (2011).
- [5] Cremonesi, P., Koren, Y. and Turrin, R.: Performance of recommender algorithms on top-n recommendation tasks, *Proceedings of the fourth ACM conference on Recommender systems*, RecSys '10, New York, NY, USA, ACM, pp. 39–46 (2010).
- [6] Fujiwara, Y., Nakatsuji, M., Onizuka, M. and Kitsuregawa, M.: Fast and exact top-k search for random walk with restart, *Proceedings of the VLDB Endowment*, Vol. 5, No. 5, pp. 442–453 (2012).
- [7] Gori, M. and Pucci, A.: ItemRank: A Random-Walk Based Scoring Algorithm for Recommender Engines, *Proc. IJCAI '07*, pp. 2766–2771 (2007).
- [8] Griffiths, T. and Steyvers, M.: Finding scientific topics, *Proceedings of the National Academy of Sciences of the United States of America*, Vol. 101, No. Suppl 1, p. 5228 (2004).
- [9] Haveliwala, T. H.: Topic-sensitive PageRank, *Proceedings of the 11th international conference on World Wide Web*, WWW '02, New York, NY, USA, ACM, pp. 517–526 (2002).
- [10] Herlocker, J. L., Konstan, J. A., Terveen, L. G. and Riedl, J. T.: Evaluating collaborative filtering recommender systems, *ACM Trans. Inf. Syst.*, Vol. 22, No. 1, pp. 5–53 (2004).
- [11] Hofmann, T.: Collaborative filtering via gaussian probabilistic latent semantic analysis, *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in information retrieval*, ACM, pp. 259–266 (2003).
- [12] Järvelin, K. and Kekäläinen, J.: Cumulated gain-based evaluation of IR techniques, *ACM Trans. Inf. Syst.*, Vol. 20, No. 4, pp. 422–446 (2002).
- [13] Konstantas, I., Stathopoulos, V. and Jose, J. M.: On social networks and collaborative recommendation, *Proc. SIGIR '09*, pp. 195–202 (2009).
- [14] Koren, Y.: Factorization meets the neighborhood: a multifaceted collaborative filtering model, KDD '08, New York, NY, USA, ACM, pp. 426–434 (2008).
- [15] Langville, A. and Meyer, C.: Deeper inside pagerank, *Internet Mathematics*, Vol. 1, No. 3, pp. 335–380 (2004).
- [16] Lee, S., Song, S.-i., Kahng, M., Lee, D. and Lee, S.-g.: Random walk based entity ranking on graph for multidimensional recommendation, *Proceedings of the fifth ACM conference on Recommender systems*, RecSys '11, New York, NY, USA, ACM, pp. 93–100 (2011).
- [17] Li, B.: Cross-Domain Collaborative Filtering: A Brief Survey, *Proceedings of the 2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, ICTAI '11, Washington, DC, USA, IEEE Computer Society, pp. 1085–1086 (2011).
- [18] Li, B., Yang, Q. and Xue, X.: Can movies and books collaborate?: cross-domain collaborative filtering for sparsity reduction, *Proc. IJCAI '09*, pp. 2052–2057 (2009).
- [19] Li, B., Yang, Q. and Xue, X.: Transfer learning for collaborative filtering via a rating-matrix generative model, *Proc. ICML '09*, pp. 617–624 (2009).
- [20] Nakatsuji, M., Fujiwara, Y., Tanaka, A., Uchiyama, T. and Ishida, T.: Recommendations Over Domain Specific User Graphs, *Proc. ECAI '10* (2010).
- [21] Page, L., Brin, S., Motwani, R. and Winograd, T.: The PageRank citation ranking: Bringing order to the web, Technical report, Stanford Digital Library Technologies Project (1998).
- [22] Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews, *Proc. CSCW '94*, pp. 175–186 (1994).
- [23] Sarwar, B., Karypis, G., Konstan, J. and Reidl, J.: Item-based collaborative filtering recommendation algorithms, *Proc. WWW '01*, pp. 285–295 (2001).
- [24] Schein, A. I., Popescul, A., Ungar, L. H. and Pennock, D. M.: Generate models for coldstart recommendations, *ACM SIGIR 2001 Workshop on Recommender Systems* (2001).
- [25] Yildirim, H. and Krishnamoorthy, M. S.: A random walk method for alleviating the sparsity problem in collaborative filtering, *Proc. RecSys '08*, pp. 131–138 (2008).
- [26] 堤田恭太, 中辻 真, 内山俊郎, 藤村 考: ソーシャルメディア上のアイテム共起を用いたサービス横断推薦, *DEIM Forum 2011* (2011).
- [27] 柳原 正, 帆足啓一郎, 小野智弘: クロスメディア型レコメンデーションの提案と評価, *日本データベース学会論文誌*, Vol. 8, No. 2, pp. 13–18 (2009-09).
- [28] 神島敏弘: 推薦システムのアルゴリズム (1)~(3), *人工知能学会誌*, Vol. 22, No. 6 ~ Vol. 23, No. 2 (2007-2008).