

## 近傍ユーザを用いたコールドユーザへの多目的推薦手法の提案

## A Neighbor-based Multi-objective Recommendation Approach for Cold-start Users

藤後 英哲<sup>†</sup>      渡邊 一生<sup>†</sup>  
Eitetsu Togo      Kazuo Watanabe

## 1. はじめに

近年、インターネット上の情報量が急速に増加する中で、ユーザが自身の好みに合った情報を効率的に取得するための推薦システムの重要性が高まっている。ユーザの満足度を高める手法の一つとして、ユーザの興味関心、推薦リストの多様性など、複数の目的を同時に考慮する多目的推薦が注目を集めている。

しかし、一般的に多目的推薦が十分に機能するためには、ユーザの過去の行動履歴や評価データが豊富に必要となる。そのため、新規ユーザや行動履歴がほとんどない「コールドユーザ」に対しては、これら複数の目的をバランス良く最適化することは困難である。

本研究では、この課題に焦点を当て、コールドユーザに対して効果的な多目的推薦を実現するためのアプローチを提案した。

## 2. 関連研究

## 2.1 多目的推薦における評価指標

多目的推薦は、多様化するユーザニーズに対応し、ユーザの満足度を高めるために精度以外の指標を考慮する手法として近年広く研究されている。特に、精度以外にも Diversity, Novelty, Coverage, Serendipity など、多岐にわたる評価指標が活用されている。ただし、同一の名称であっても利用目的や解釈の違いによって、異なる定義がされている。本研究では多目的推薦の代表的な指標として Diversity と Novelty に着目する。

**Diversity (多様性)** 本研究では Diversity を、ユーザが評価したアイテムリスト  $R$  内で得たアイテムベクトル間の散らばりを測定する指標として扱う[1]。この指標は、評価したアイテムがどれだけ異なる特徴を持つアイテムを含んでいるかを定量化するものであり、次の式で計算される：

$$\text{Diversity}(R) = 1 - \frac{\sum_{i,j \in R} \text{cosineSimilarity}(\mathbf{q}_i, \mathbf{q}_j)}{|R|(|R| - 1)/2}$$

ここで、 $\mathbf{q}_i$  および  $\mathbf{q}_j$  はそれぞれアイテム  $i$  および  $j$  のベクトル表現を示し、 $|R|$  はリスト  $R$  内のアイテム数を表す。

**Novelty (新規性)** 本研究では Novelty を、ユーザが評価したアイテムが他のユーザにとってどの程度馴染みのないものかを測定する指標として扱う[2]。この指標は、アイテム

$i$  の全ユーザ集合  $U$  における普及度に基づき、次の式で計算される：

$$p(i) = \frac{|u \in U; r_{ui} \neq \emptyset|}{|U|}$$

$$\text{Novelty}(R) = \frac{\sum_{i \in R} -\log_2 p(i)}{|R|}$$

ここで、 $U$  は全ユーザの集合、 $|U|$  は全ユーザの数を示し、 $r_{ui}$  はユーザ  $u$  がアイテム  $i$  の評価を行ったかどうかを示す値である。また、 $p(i)$  はアイテム  $i$  を評価したユーザの割合を表す。

## 2.2 多目的推薦の手法

コールドユーザへの多目的推薦に関する従来研究の多くは、コンテンツベースのアプローチを採用しており、協調フィルタリングを利用した研究は比較的少ない。これは、協調フィルタリングが一般的にユーザの行動履歴や評価履歴に大きく依存し、履歴が乏しいコールドユーザに対して Diversity や Novelty を効果的に推定することが困難なためである。

協調フィルタリングに基づく多目的推薦のアプローチとしては、主に次の2つの手法が挙げられる。

**リランキング手法** 初期の推薦リストを作成した後、Diversity や Novelty などの指標を考慮して順序を再構成する手法である[3]。この手法では、ユーザごとに多様性や新規性の重視度を調整することが可能なことが多い。しかし、ユーザの行動履歴が十分に存在することが前提となることが多く、コールドユーザへの適用は困難である。

**多目的性を考慮したモデル学習手法** 学習段階において、精度に加え Diversity や Novelty の指標も考慮することで、多目的な最適化を図る手法である[4]。この手法は、推薦リスト全体にわたって一貫した特性を持たせやすいという利点があり、また、推薦リストの生成時に追加の処理を必要としないため、推論時の計算コストを抑えられる。一方で、複数の目的を同時に含めることによる学習の不安定化や個々のユーザに個別に多様性や新規性を反映させることが難しいという問題点がある。

本研究では、既存の推薦モデルとの統合の容易性や実装上の柔軟性といった観点から、リランキングによる多目的推薦手法に着目する。リランキングによる多目的推薦では、コールドユーザにおける多様性や新規性の算出の際、スコアが偏りやすく、信頼性の高い値を算出することが難しい。

そこで本研究では、このような既存手法の課題を解決するために、十分にレコード数のあるユーザの情報を活用す

<sup>†</sup> 株式会社日本総合研究所 The Japan Research Institute, Limited

ることでコールドユーザへの多目的推薦の精度向上を図る方法を提案する。

### 3. 提案手法

コールドユーザへの多目的推薦が難しいという課題を解決するため、本研究では、ユーザの少ない行動履歴のみで推薦するのではなく、十分な行動履歴を持つ他のユーザ群から嗜好が類似したユーザ（以下、近傍ユーザ）を活用する方法を提案する。

提案手法の概要を図 1 に示す。まず、データセットをもとに学習した推薦システムを用い、コールドユーザの推薦リストを作成する。次に、コールドユーザの行動履歴をもとに、十分に行動履歴のあるユーザから近傍ユーザを取得する。その後、取得した近傍ユーザについて多様性や新規性を算出し、算出した値に基づいて推薦リストをリランキングする。

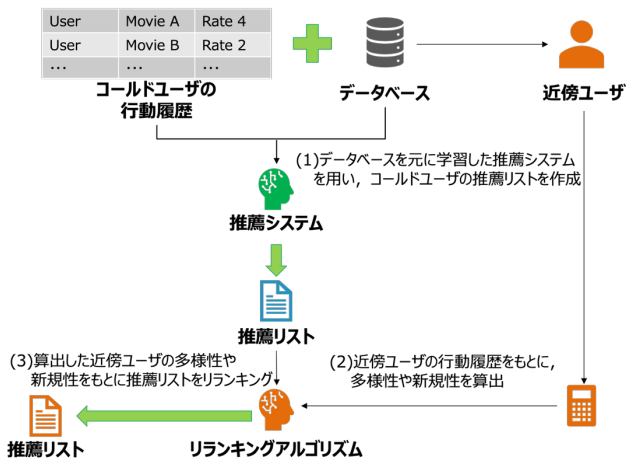


図 1 提案手法の概要

従来手法では、コールドユーザ自身の限られたデータから多様性や新規性を直接算出して推薦を行うが、本研究では、データが十分に蓄積されたユーザ群の情報を代理として活用することで、より精度の高い多目的推薦を目指す。

そのため、本研究では以下の 3 つのアプローチを比較する。

**既存手法:** コールドユーザについて、少ないレコード数のまま多様性や新規性を算出し、それらをコールドユーザの推定値として用いて多目的推薦を行う方法。

**提案手法:** まず、十分にレコード数のあるユーザの中から、推薦対象となるコールドユーザの近傍ユーザを  $k$  件取得する。その後、そのユーザの多様性や新規性の平均値をコールドユーザの推定値として算出し、その値に基づいて多目的推薦を行う方法。

**比較手法:** まず、ランダムに  $k$  件のユーザを取得する。その後、そのユーザの多様性や新規性の平均値をコールドユーザの推定値として算出し、その値に基づいて多目的推薦を行う方法。

なお、本研究では  $k$  を 5 件に統一した。

## 4. 実験

本実験の目的は、コールドユーザに対する多様性と新規性の推定において、3 つのアプローチを比較検討することである。

### 4.1 実験の流れ

本実験では、重複を避けてランダムに 1,200 人のユーザ  $u$  を選出し、各ユーザに対して以下の手順を実施した。

#### 1. ユーザ $u$ をランダムに選択

ただし、ユーザ  $u$  はアイテム数が 20 件以上のユーザとする。

#### 2. コールドユーザ $u'$ を作成

ユーザ  $u$  の最初の  $n$  件 ( $n \in \{2, 3, \dots, 20\}$ ) を抽出し、ユーザ  $u'$  のレコードとする。

#### 3. 学習データセットの作成およびモデルの学習

データセットにユーザ  $u'$  を追加し、モデルを学習する。その後、学習したモデルからユーザベクトル  $q_u$  およびアイテムベクトル  $q_i$  を取得する。

#### 4. 3 つのアプローチを比較

ユーザ  $u'$  に対して、以下の 3 つのアプローチの精度を比較する。

1. ユーザ  $u'$  のまま算出した評価指標と、ユーザ  $u$  の評価指標の絶対誤差の平均。（既存手法）
2. ユーザ  $u'$  の最近傍ユーザ  $v$  を  $k$  件取得し、その  $k$  件のユーザ  $v$  の評価指標の平均値とユーザ  $u$  の評価指標の絶対誤差の平均。（提案手法）
3. ランダムにユーザ  $v$  を  $k$  件取得し、その  $k$  件のユーザ  $v$  の評価指標の平均値とユーザ  $u$  の評価指標の絶対誤差の平均。（比較手法）

ユーザ  $v$  を取得する際、以下の 2 つの処理を実施した。

- ユーザ  $v$  は、レコード数が 20 件以上あるユーザから取得した。これは、提案手法が十分にレコード数のあるユーザの中から近傍ユーザを取得することを前提としているためである。
- ユーザ  $v$  を取得する際、ユーザ  $u$  を取得しないよう処理をした。

### 4.2 データセット

本研究では、MovieLens 1M[6]、Netflix Prize[7]、Amazon-Reviews(Books)[8] の 3 つのデータセットを使用した。全てのデータセットで総レコード数が概ね 100 万件になるように、Netflix Prize および Amazon-Reviews(Books)についてはそれぞれランダムに 5,000 ユーザ、350,000 ユーザをサンプリングした。表 1 は各データセットの概要である。

表 1 データセットの概要

データセット	件数	ユーザ数	アイテム数
MovieLens	1,000,209	6,040	3,706
Netflix	1,025,250	5,000	17,751
Amazon	1,010,896	350,000	564,301

なお、各データセットに含まれるユーザの行動履歴は全て学習に利用した。また本研究では協調フィルタリングに基づく手法を採用しているため、ユーザやアイテムの属性

情報は利用しておらず、ユーザがアイテムを評価したかどうかのみを使用している。

### 4.3 使用したモデル

本研究では、SVD[9]を使用した。SVD は行列分解に基づく手法で、各ユーザおよびアイテムを低次元の潜在的特徴空間に写像し、それらの特徴ベクトルを用いて評価値を予測する。具体的には、ユーザ $u$ とアイテム $i$ の評価値 $\hat{r}_{ui}$ は次の式で表される：

$$\hat{r}_{ui} = \mu + b_u + b_i + \mathbf{q}_i^T \mathbf{p}_u$$

ここで、 $\mu$ は全ユーザ・アイテムの評価値の平均を示し、 $b_u$ と $b_i$ はそれぞれユーザ $u$ とアイテム $i$ の評価値におけるバイアスを表す。また、 $\mathbf{p}_u$ と $\mathbf{q}_i^T$ は、それぞれユーザ $u$ とアイテム $i$ に対応する潜在特徴ベクトルである。

## 5. 結果

### 5.1 Diversity

まず、MovieLens 1Mにおける Diversityの結果を図 2 に示す。横軸はコールドユーザにおける行動履歴の件数 ( $n$ ) である。比較手法と提案手法を比較したところ、提案手法は常に高い精度で多様性を算出することができた。また、既存手法との比較では、 $n$ が 9 件以下の場合において提案手法の方が高い精度を示した。

次に、Netflix Prizeにおける Diversityの結果を図 4 に示す。ここでも比較手法と提案手法を比較すると、提案手法が常に高精度であった。さらに、 $n$ が 10 件以下では既存手法に対しても提案手法が高い精度を示した。

続いて、Amazon-Reviews(Books)における Diversityの結果を図 6 に示す。このデータセットにおいても、比較手法と比較し、提案手法は常に高い精度であった。既存手法との比較において、 $n$ が 9 件以下の範囲で提案手法が高い精度であった。

### 5.2 Novelty

MovieLens 1Mにおける Noveltyの結果を図 3 に示す。比較手法と提案手法の性能はほぼ同程度であったが、既存手法と提案手法を比較すると、 $n$ が 6 件以下において提案手法が高い精度を示した。

続いて Netflix Prizeの結果を図 5 に示す。Noveltyにおいても Diversityと同様に、提案手法が常に比較手法より高い精度を示した。また、 $n$ が 3 件以下の場合には既存手法よりも提案手法の方が優れていた。

最後に Amazon-Reviews(Books)における Noveltyの結果を図 7 に示す。このデータセットにおいても、比較手法と比較し、Diversityと同様に提案手法は常に高い精度を示した。さらに、 $n$ が 5 件以下の場合において、既存手法よりも提案手法の精度が高かった。

## 6. 考察

本研究では、コールドユーザに対する多目的推薦において、十分な行動履歴を持つ近傍ユーザの情報を活用するア

プローチを提案した。実験結果から、次の点が明らかになった。

まず、提案手法はユーザの行動履歴が限られた状態において、従来手法よりも高精度に多様性や新規性を算出することができた。これは、コールドユーザ自身の行動履歴だけでは多様性や新規性の信頼性のある推定が困難であり、近傍ユーザの情報を用いることで、より正確に評価できるためだと考えられる。特に、行動履歴の少ない初期段階のユーザに対して、近傍ユーザの情報を利用することの有効性が示された。

また、ユーザの行動履歴が非常に少ない場合では、既存手法が比較手法よりも精度が低いことがあった。これは、データが極めて少ない状況では、コールドユーザの行動履歴から多様性や新規性を推定することが困難であり、むしろ全体的な平均を用いる方が良いことを示唆している。

さらに、提案手法とランダムに選択したユーザの情報を用いた比較手法との比較においても、提案手法は比較的一貫して高い精度を示した。これは、ユーザの嗜好の類似性を考慮して近傍ユーザを選定することが、精度向上に重要であることを示している。

以上より、提案手法はコールドユーザの多目的推薦において有効なアプローチであることが示された。本研究では、多様性や新規性を高精度に算出する段階にとどまっている。そのため、実際に推薦リストの作成まで行い、その推薦精度についても比較評価することを今後の課題とする。

### 参考文献

- [1] Smyth, B., McClave, P., "Similarity vs. Diversity", In Aha, D. W., Watson, I. (Eds.), Case-Based Reasoning Research and Development, pp. 347-361 (2001).
- [2] Zhou, T., Kuscsik, Z., Liu, J.-G., Medo, M., Wakeling, J. R., and Zhang, Y.-C., "Solving the apparent diversity-accuracy dilemma of recommender systems", Proceedings of the National Academy of Sciences, Vol. 107, No. 10, pp. 4511-4515 (2010).
- [3] Abdollahpour, H., Burke, R., and Mobasher, B., "Managing popularity bias in recommender systems with personalized re-ranking", In Proceedings of the 32nd International Florida AI Research Society Conference (FLAIRS' 19), pp. 413-418 (2019).
- [4] Qin, L., and Zhu, X., "Promoting diversity in recommendation by entropy regularizer", In Proceedings of the Twenty-Third International Joint Conference on Artificial Intelligence (IJCAI' 13), pp. 2698-2704 (2013).
- [5] Harper, F. M., and Konstan, J. A., "The MovieLens Datasets: History and Context", ACM Transactions on Interactive Intelligent Systems, Vol. 5, No. 4 (2015).
- [6] Bennett, J., and Lanning, S., "The Netflix Prize", Proceedings of the KDD Cup Workshop 2007, pp. 3-6 (2007).
- [7] McAuley, J., Pandey, R., and Leskovec, J., "Inferring Networks of Substitutable and Complementary Products", Proceedings of the 21st ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD ' 15), pp. 785-794 (2015).
- [8] Koren, Y., Bell, R., and Volinsky, C., "Matrix Factorization Techniques for Recommender Systems", Computer, Vol. 42, No. 8, pp. 30-37 (2009).

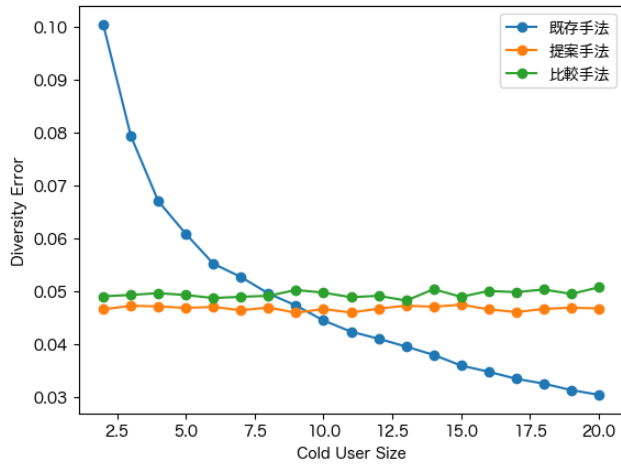


図 2 MovieLens 1M における Diversity 推定誤差

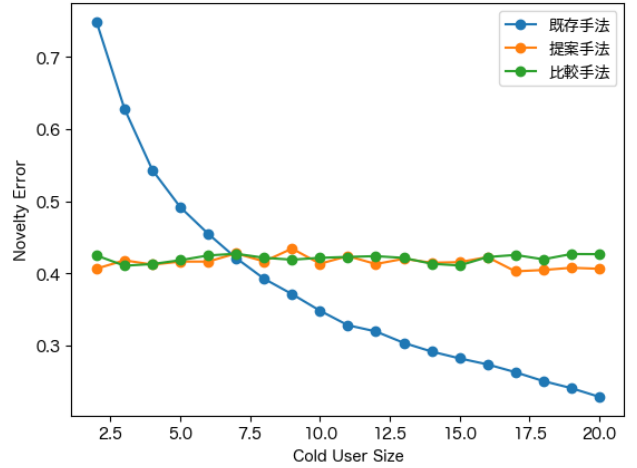


図 3 MovieLens 1M における Novelty 推定誤差

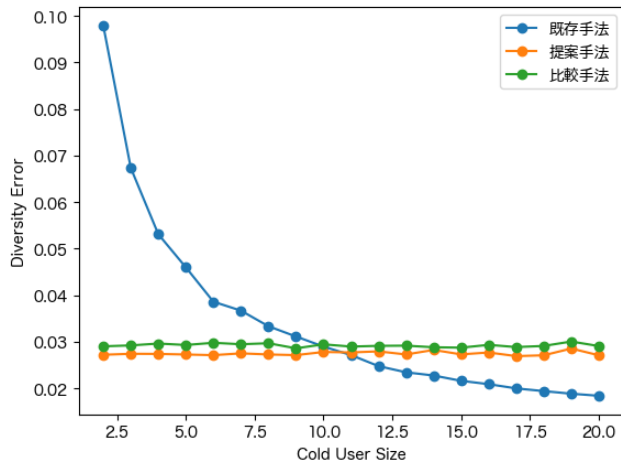


図 4 Netflix Prize における Diversity 推定誤差

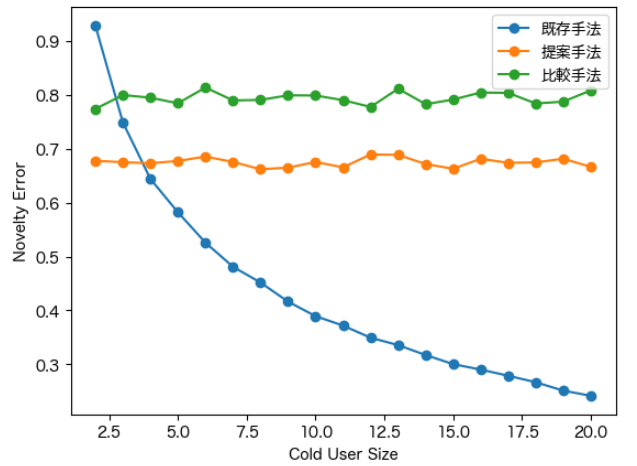


図 5 Netflix Prize における Novelty 推定誤差

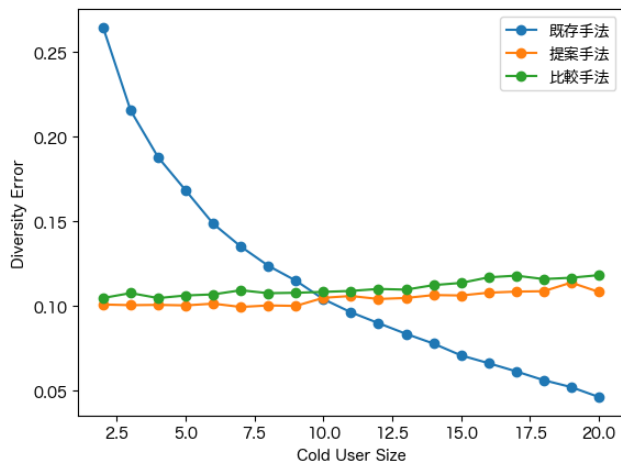


図 6 Amazon-Reviews(Books) における Diversity 推定誤差

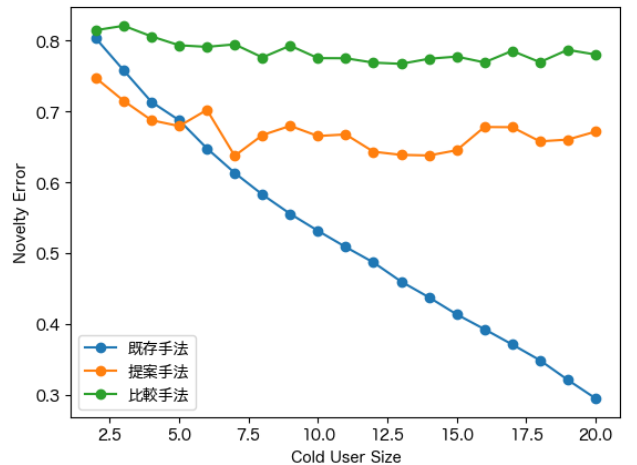


図 7 Amazon-Reviews(Books) における Novelty 推定誤差