

特定アイテムへのフィードバックを重視した 協調フィルタリング手法の提案

Proposal of Collaborative Filtering Focusing on Feedback on Specific Items

○野崎 海斗
○ Nozaki Kaito

柴田 祐樹
Shibata Hiroki

高間 康史
Takama Yasufumi

東京都立大学大学院システムデザイン研究科
Graduate School of System Design, Tokyo Metropolitan University

Abstract: This paper proposes a method for matrix factorization-based collaborative filtering (CF) that focuses on feedback to specific items. Although a large amount of user information is usually required for accurate recommendation, privacy concern makes it difficult to collect user information. As a solution, we focus on applying CF to a small amount of feedback on a few items by developing a method for determining the items for which user feedback can be efficiently obtained (Probing Items), and a method for learning a recommendation system focusing on the probing items. This paper focuses on the latter and proposes a method for learning a matrix factorization-based CF, in which probing items are considered at every epoch while other items are randomly dropped. Experimental results show the effectiveness of the proposed method.

1 はじめに

本稿では、行列分解ベース協調フィルタリングにおいて、特定アイテムへのフィードバックを重視する手法を提案する。近年、インターネット利用者の増加に伴い、大量の情報がインターネット上に流通するようになった。その結果、ユーザが効率よく必要な情報にアクセスするために情報推薦システムの需要が高まってきている。情報推薦手法には大きく分けて内容ベースフィルタリングと協調フィルタリングがある [1]。内容ベースフィルタリングでは、アイテムの属性やユーザの嗜好データを用いて推薦アイテムを決定する。一方で、協調フィルタリングはアイテムの属性を用いずに、ユーザの評価情報に基づいて推薦アイテムを決定する手法であり、アイテム属性などの情報が不要であるなどの利点があるため、多数の推薦手法が提案されている。

一般に、精度の高い推薦にはユーザの情報が大量に必要になるが、最近ではプライバシー保護の観点からユーザ情報の取得に対する目が厳しくなっている。そこで本稿では、特定の少数アイテムへのフィードバックを重視した協調フィルタリングの学習アプローチを検討する。特定の少数アイテムに着目することにより、ユーザから取得するフィードバック情報を最小限に抑えられるため、ユーザのプライバシーに配慮した情報推薦システムの実現が期待できる。

提案するアプローチにおいて、ユーザのフィードバックを効率よく得られるアイテムをプロービングアイテム

と定義する。プロービングアイテムの選択手法、およびそれに対するフィードバックを重視した推薦モデルの学習手法を検討する必要があるが、本稿は後者を対象とする。提案する学習手法では行列分解ベース協調フィルタリングの学習時にプロービングアイテムは全エポックにて考慮するのに対し、それ以外のアイテムはエポック毎にランダムに選択・考慮することで、プロービングアイテムを重視した学習を行う。

人気アイテムを重視した場合とランダムに選択したアイテムを重視した場合を比較する評価実験を行い、提案手法の有効性について考察する。

2 関連研究

2.1 協調フィルタリング

情報推薦手法は内容ベースフィルタリングと協調フィルタリングに大別される。内容ベースフィルタリングではアイテムの属性やユーザの嗜好データに基づいて、属性値の近さから推薦アイテムを決定する。しかし、音楽や映像といった人の感性に左右されるアイテムに対して属性値を設定するのは難しく、内容ベースフィルタリングにおける推薦の精度を高めることは困難になっている。そのため、アイテムの属性を用いずにユーザの評価情報に基づいて推薦アイテムを決定する協調フィルタリングが注目されている [2]。

現在、協調フィルタリングはインターネットショッピングをはじめとする多くのオンラインシステムに実装されている。協調フィルタリングの基本的な考え方として

「口コミによる推薦のモデル化」がある [3]. 対象ユーザのアイテムへの評価履歴から嗜好が類似しているユーザを発見し、その類似ユーザが好むアイテムを対象ユーザに推薦アイテムとして提示する.

推薦対象アイテムに関する知識を必要としないことや、セレンディピティの高い推薦が可能である点が協調フィルタリングの利点とされている. しかし一方で、ユーザ数が少ないことで起こる「cold-start 問題」や、ユーザ数に対してアイテム数が膨大な場合にユーザとアイテムの関係性が希薄になってしまう「sparsity 問題」が指摘されている.

協調フィルタリングはメモリベース法とモデルベース法に分類することができる [4]. メモリベース法は、推薦の際にユーザの嗜好データや評価情報をそのまま利用する手法であり、代表的な手法として GroupLens が挙げられる [5]. モデルベース法はあらかじめモデルを構築しておき、推薦の際にそのモデルを用いて推薦を行う手法であり、代表的な手法として Matrix Factorization (MF) が挙げられる [6].

2.2 Matrix Factorization (MF)

MF は行列分解をすることで評価値行列の次元圧縮を行う. スパースな行列に対しても有効であることから、従来の協調フィルタリングで指摘されていた「sparsity 問題」を解決する手法として注目されている.

一般的に MF はユーザ数を M 、アイテム数を N として、 $k < M, N$ を満たす k に対して、 $M \times N$ の評価値行列 R を $M \times k$ の行列 P と $N \times k$ の行列 Q に分解する. ここで、 P, Q の各行が、ユーザ、アイテムの因子ベクトルにそれぞれ対応する. ユーザ i の因子ベクトルを p_i 、アイテム j の因子ベクトルを q_j とすると、ユーザ i のアイテム j に対する予測評価値 \hat{r}_{ij} は式 (1) で表される.

$$\hat{r}_{ij} = p_i \times q_j \quad (1)$$

また、実際の評価値 r_{ij} と予測評価値 \hat{r}_{ij} の二乗誤差の総和を目的関数 L とすると、 L は式 (2) で表される.

$$L = \sum_{i,j} (r_{ij} - \hat{r}_{ij})^2 \quad (2)$$

式 (2) において目的関数 L を最小化するような P, Q を近似的に求めることで、その積 PQ^T から予測評価値行列を得ることができる.

3 提案手法

3.1 提案手法概要

本稿では、プロービングアイテムを重視した学習手法を提案する. 提案手法では、一つ以上のアイテムがプロービングアイテムとして与えられることを想定する.

プロービングアイテムは協調フィルタリングの学習時に全エポックにて考慮することで、それらに対する予測評価値が実際の評価値と近づくように学習する. しかし、少数のプロービングアイテムのみで学習を行うことは困難なため、他のアイテムを一定数考慮する必要がある. そのため、プロービングアイテムの他にエポック毎にランダムでアイテムを一定数選出し、学習に考慮する. 全エポックで学習に利用されるプロービングアイテムは、他のアイテムよりも学習結果に影響を与えることが期待されるため、プロービングアイテムを重視したモデルが得られることが期待できる.

3.2 Biased-Matrix Factorization (B-MF)

協調フィルタリングの学習には Biased-Matrix Factorization (B-MF) を用いる [6]. B-MF では 2.2 節で述べた基本的な MF にユーザやアイテムに関するバイアスを導入したものである. B-MF では、予測評価値 \hat{r}_{ij} は式 (3) で求められる. ここで、 μ は全評価値の平均、 a_i, b_j はそれぞれ i 番目のユーザ、 j 番目のアイテムのバイアスを表す.

$$\hat{r}_{ij} = p_i \times q_j + \mu + a_i + b_j \quad (3)$$

バイアス項を用いることで、高評価を付けられやすいアイテムや高評価を付けやすいユーザといった、ユーザやアイテムの傾向を表現することができ、評価の偏りを考慮することができる.

提案手法で用いる目的関数 L を式 (4) に示す.

$$L = \sum_i \left\{ \sum_j e_{ij}^2 * w_j \right\} + \lambda (\|P\|^2 + \|Q\|^2 + \|A\|^2 + \|B\|^2) \quad (4)$$

$$e_{ij} = r_{ij} - \hat{r}_{ij} \quad (5)$$

式 (4) において、 λ は正則化係数を表しており、 λ 以降の項を正則化項と呼ぶ. 正則化項を用いることで局所解に陥ることによる過学習のリスクを軽減するはたらかが期待できる. また、 w_j はアイテムに対する重みを表しており、アイテム j を学習に考慮する場合 $w_j = 1$ 、学習に考慮しない場合 $w_j = 0$ とする. ただし、プロービングアイテムは学習において全エポックにて考慮するため、常に $w_j = 1$ とする. $A \in R^M, B \in R^N$ はそれぞれユーザバイアス、アイテムバイアスである.

学習において各値の更新には式 (6), (7), (8), (9) で表す更新式を用いる.

$$p_i \leftarrow p_i + \gamma (e_{ij} w_j q_j - \lambda p_i) \quad (6)$$

$$q_j \leftarrow q_j + \gamma(e_{ij}w_jp_i - \lambda q_j) \quad (7)$$

$$a_i \leftarrow a_i + \gamma(e_{ij}w_j - \lambda a_i) \quad (8)$$

$$b_j \leftarrow b_j + \gamma(e_{ij}w_j - \lambda b_j) \quad (9)$$

ここで、 γ は学習率を表しており、エポック毎に学習するスケールを調整する役割を持つ。エポック毎に各値を確率的勾配降下法を用いて更新していき L を最小化するような P, Q, A, B を求める。

4 評価実験

4.1 実験概要

本節では、データセットとして Movielens^{*1}を用いて提案手法を適用し、人気アイテムを重視した場合とランダムに選択したアイテムを重視した場合で予測精度や多様性、人気バイアスといった観点で違いがみられるか比較を行い、提案手法の有効性について考察する。

Movielens は映画に関するデータセットで 610 人のユーザによる 9,742 本の映画に対する評価 100,836 件が含まれている。ユーザは映画に対して、0.5 から 5.0 まで 0.5 刻みで評価をしており、最低でも 20 件の評価を行っている。このデータセットを訓練データ 60%、テストデータ 20%、パラメータ調整用 20% として分割して実験を行った。

評価実験では、「人気度の高いアイテム 100 件」と「ランダムに選択したアイテム 100 件」をそれぞれプロービングアイテムとみなしてモデルを構築した。人気度として各アイテムに対する評価件数を採用し、評価件数が多いアイテムを人気アイテムとしている。多くの人に知られている人気アイテムはフィードバックを得るのにふさわしいと考え、プロービングアイテムの選択基準として採用した。また、ランダムに選択した場合と比較することで、プロービングアイテムの選択基準が与える影響についても考察する。

エポック毎にランダムに選択し学習に考慮するアイテム数については 100 件、1,000 件、プロービングアイテム以外の全アイテム (All) の 3 条件についてそれぞれ実験を行った。その他、学習率 γ や正則化係数 λ 、因子ベクトルのサイズ (因子数) k 、エポック数については表 1 に示す値を用いた。

4.2 実験結果

提案手法を適用し、プロービングアイテムを重視したモデルによる予測精度の算出を行った。本稿では予測精

表 1 評価実験で設定したパラメータ

パラメータ	設定値
学習率 γ	0.02
正則化係数 λ	0.02
因子数 k	100
エポック数	30

度として平均絶対値誤差 (MAE) を用いた。MAE は実際の評価値と予測評価値の差を表しており、テストデータを R_{test} とすると、MAE は式 (10) で求められる。

$$MAE = \frac{1}{|R_{test}|} \sum_{i,j \in R_{test}} |r_{ij} - \hat{r}_{ij}| \quad (10)$$

プロービングアイテムを「人気度の高いアイテム 100 件」としたモデルによる予測精度を表 2 に、「ランダムに選択したアイテム 100 件」としたモデルによる予測精度を表 3 に示す。

表 2, 3 において $\#other_item$ はプロービングアイテム以外で学習に用いるアイテム数、 $corr_num$, $corr_rate$ はユーザが評価したアイテムの中にプロービングアイテムが含まれる個数あるいは割合と MAE の Pearson 相関を表している。

MAE に注目すると、 $\#other_item$ が 1,000 件、All の場合は表 2, 3 でほとんど差が見られないが、100 件のとき「人気度の高いアイテム 100 件」をプロービングアイテムとしたモデルの方が MAE が小さいことがわかる。 $\#other_item$ が大きい方が MAE が小さくなる傾向にあるが、これは学習に用いるデータ数が増加するためであり、妥当な結果といえる。 $\#other_item$ が 100 のときに表 2 の方が MAE の増加が抑えられていることから、人気アイテムの方がランダムに選択するよりもプロービングアイテムとして適している可能性があるといえる。

表 2 人気度の高いアイテム 100 件をプロービングアイテムとしたモデルの実験結果

$\#other_item$	100	1,000	ALL
MAE	0.896	0.769	0.754
$corr_num$	-0.142	-0.135	-0.140
$corr_rate$	-0.153	-0.114	-0.097

表 3 ランダムに選択したアイテム 100 件をプロービングアイテムとしたモデルの実験結果

$\#other_item$	100	1,000	ALL
MAE	0.922	0.772	0.754
$corr_num$	-0.091	-0.070	-0.077
$corr_rate$	-0.067	0.020	-0.014

^{*1} <https://grouplens.org/datasets/movielens/>

表 4 推薦アイテムの多様性 (単位: 種類)

#other_item	100	1,000	All
人気度	98	109	942
ランダム	16	29	942

表 5 推薦アイテムの平均人気度

#other_item	100	1,000	All
人気度	98.1	92.2	29.9
ランダム	30.0	30.1	29.9

$corr_num$, $corr_rate$ に関しては, 負の相関が観測されれば, プロービングアイテムを多く評価しているユーザほど MAE が小さくなることを意味する. 表 2, 3 より, どの実験結果でも相関は見られなかったが, 表 2 の方が表 3 よりも負で絶対値が大きくなっている. さらに, $corr_rate$ では #other_item が小さくなるほど負で絶対値が大きくなる傾向がみられる. 以上の結果から, 人気度でプロービングアイテムを選択した場合には, 学習においてプロービングアイテムの割合を多くすることで, プロービングアイテムを重視してユーザモデルが構築されていると考える.

次に, プロービングアイテムを重視したモデルについて多様性の評価を行った. 各ユーザに対して予測評価値の高いアイテム 10 件を推薦した場合, 何種類のアイテムが推薦アイテムとして登場したかカウントすることで, ユーザに推薦されるアイテムの多様性を評価する.

「人気度の高いアイテム 100 件」, 「ランダムで選択したアイテム 100 件」をそれぞれプロービングアイテムとして学習したモデルに対する多様性の評価結果を表 4 に示す. 表 4 より, #other_item が 100 件, 1,000 件のとき, 推薦アイテムとして全アイテム 9,742 件のうち 1% 程度しか選択されていないことから, 推薦アイテムの多様性は低いといえる. これは, #other_item が小さいほどプロービングアイテムによる学習への影響が大きくなるため, 学習において偏りが表れたためだと考えられるが, 人気度で選択した場合の方がランダムに選択した場合よりも多様性の低下が抑制されていることも確認できる.

最後に人気バイアスによる影響について分析を行った. 多様性の評価と同様の基準で推薦アイテムを決定し, その平均人気度を比較する.

「人気度の高いアイテム 100 件」, 「ランダムで選択したアイテム 100 件」をそれぞれプロービングアイテムとしたモデルに対する人気バイアスの評価結果を表 5 に示す. 表 5 より, 人気度でプロービングアイテムを選択した場合の方が, 推薦アイテムの平均人気度が高くなって

いることから, 人気バイアスの影響を強く受けていると考えられる. #other_item が小さいほどプロービングアイテムによる学習の影響が大きくなるため, バイアスが強いことも確認できる.

5 まとめ

本稿では, 行列分解ベース協調フィルタリングの学習においてプロービングアイテムを重視する学習手法を提案した. Movielens を用いた評価実験では, 人気度の高いアイテムをプロービングアイテムとした場合に, MAE の増加が抑えられることや, プロービングアイテムの割合を高めることでそれを重視したモデルが構築されていることを示した. 今後はプロービングアイテムの決定方法について検討する予定である. 具体的には, 少量でも MAE の増加が抑制され, モデルへの影響が強く出るようなアイテムの選択方法について検討する. また, 人気バイアスは人気アイテムをプロービングアイテムとして利用したことが原因と考えられることから, 人気バイアスが発生しないようなアイテムの選択方法についても検討する.

参考文献

- [1] 神島 敏弘, 推薦システムのアルゴリズム (2), 人工知能学会誌, Vol. 23, No. 1, pp. 89-103, 2008.
- [2] 角 康之, 協調フィルタリング, 日本ファジィ学会誌, Vol. 12, No. 6, pp. 783, 2000.
- [3] 岡本 一志, 藤井 流華, 協調フィルタリング入門, 知能と情報, Vol. 31, No. 1, pp. 5-9, 2019.
- [4] J.S. Breese, David Heckerman, C. Kadie, "Empirical Analysis of Predictive Algorithms for Collaborative Filtering," UAI'98, pp. 43-52, 1998.
- [5] P. Resnick, N. Iacovou, M. Suchak, P. Bergstrom, J. Riedl, "GroupLens: An Open Architecture for Collaborative Filtering of Netnews," CSCW'94, pp. 175-186, 1994.
- [6] Y. Koren, R. Bell, C. Volinsky, "Matrix Factorization Techniques for Recommender Systems," IEEE Computer, Vol. 42, No. 8, pp. 30-37, 2009.

連絡先

野崎 海斗

東京都立大学大学院システムデザイン研究科

E-mail : nozaki-kaito@ed.tmu.ac.jp