

エージェントモデルによるユーザ間と アイテム間協調フィルタリングの性能評価

山下晃弘^{†1} 川村秀憲^{†1,†2,†3}
飯塚博幸^{†4} 大内東^{†1,†3}

Web 上の情報量が拡大する中、推薦システムが注目されている。これまでに、ユーザの行動履歴を入力として推薦の精度を最大化する手法について多くの研究が報告されてきた。しかし、推薦結果がその後のユーザ行動に影響することも考慮すると、推薦システムとユーザは相互に影響を及ぼしあう関係であるにも関わらず、そのようなダイナミクスはこれまでほとんど議論されてこなかった。そこで本研究では、推薦システムとユーザの相互作用という観点に着目し、推薦システムの利用者を購入格付けエージェントとして全体をモデル化することで推薦アルゴリズムの特長分析を行った。本稿では、推薦アルゴリズムの原点であるユーザ間協調フィルタリングと、現在一般的に利用されているアイテム間協調フィルタリングを取り上げその結果について報告する。

Performance Evaluation of User-based and Item-based Collaborative Filtering by Multi-agent Modeling

AKIHIRO YAMASHITA,^{†1} HIDENORI KAWAMURA,^{†1,†2,†3}
HIROYUKI IIZUKA^{†4} and AZUMA OHUCHI^{†1,†3}

Information overload problem has expanded further more. A recommender system provides personalized recommendation for each user to reduce the problem. Many techniques to maximize the accuracy of recommendation based on browse history or purchase of certain products have been already proposed. Moreover, the recommendation result influences the user's behavior. Basically, the recommender system and the users are interactive, however, this dynamics had not been researched. In this research, we modeled the user as a purchase-rating agent, and analyzed the performance of the recommendation algorithm considering the system-user interaction. In this paper, we report a result of the analysis of user-based and item-based collaborative filtering.

1. はじめに

ウェブ上における情報量が拡大する中で、推薦システムが注目され、多くのオンラインストアで導入されている¹⁾。推薦システムは、ユーザの興味に合わせた情報提示により利便性を向上させるだけでなく、サイト運営側は宣伝効果による収益拡大が期待できるため、双方にメリットのあるシステムである。中でも

Amazon.com は推薦システムを効果的に利用した典型的な例である。

推薦システムは協調フィルタリング (Collaborative Filtering 以下 CF) の基本概念を始めて示した GroupLens プロジェクト²⁾ に端を発しこれまで多くの改良、分析、評価がなされてきた³⁾⁵⁾。推薦システムは、ユーザの行動履歴から推薦アイテムを決定し、推薦結果はユーザの行動に影響を与える。つまり推薦システムとユーザは相互に影響を及ぼしあう関係であり、そのインタラクションの中でユーザ全体の満足度を向上させる必要がある。しかしこれまでの研究では、ある瞬間において推薦精度を最大化することに目的が置かれ、推薦結果がユーザ行動に影響し次の推薦結果に影響するというようなダイナミクスは考慮されていない。

そこで本研究では、推薦システムとユーザの相互作用が推薦結果に与える影響についてマルチエージェントモデルを用いて解析を行った。本稿では、最も基本的

^{†1} 北海道大学 大学院情報科学研究科
Graduate School of Information Science and Technology, Hokkaido University
^{†2} 日本学術振興会, JSPS
JSPS, Japan Society for the Promotion of Science
^{†3} 科学技術振興機構, CREST
CREST, Japan Science and Technology Agency
^{†4} 公立はこだて未来大学
Department of Media Architecture, Futur University-Hakodate

なアルゴリズムであるユーザ間 CF (User-based CF) と、多くの利用実績を持つアイテム間 CF (Item-based CF) についての分析結果を報告する。

2. 関連研究

CF は、メモリベース方式とモデルベース方式に分類される⁶⁾。モデルベース手法では、予め一般化したモデルを使用して推薦を行うが、メモリベース方式ではユーザから得た評価がすぐに次の推薦結果に影響する。本稿は相互作用の効果検証を目的としており、より互いの影響が強いメモリベースの CF を扱う。

推薦システムの評価は難しいと言われている³⁾。これまで様々な評価指標が提案されているが、最も一般的な指標は予測精度であり、多くの場合、実際の Web サイトから何らかのアイテムに対するユーザの格付けデータを用いて検証している⁷⁾。この場合、データの一部を用いて推薦アルゴリズムを学習させ、残りの部分を推測させることで検証を行う。

しかしこの方法では、推薦システムとユーザ間の相互作用は観察できず、実際に実データから観察するには多くのコストと時間が必要となる。また、複数の推薦アルゴリズムについて同条件下で検証することが難しい。そこで本研究では、エージェントベースで環境を構築し、アルゴリズムの特性を分析した。

3. モデル

3.1 モデルの構成要素

本研究で構築したモデルは、購入格付けエージェント (以下エージェント) とアイテム、推薦システムの 3 つの要素から構成される。

推薦システムはエージェント集合 $U = \{i | i = 1, 2, \dots, n\}$ の各エージェントに対して、アイテム集合 $C = \{j | j = 1, 2, \dots, m\}$ から適切なアイテムを推薦する。各エージェントの嗜好 (preference) は嗜好ベクトル $\mathbf{p}_i = (p_{i1}, p_{i2}, \dots, p_{iN_f})$ として表現する。同様に、各アイテムの特徴 (feature) は特徴ベクトル $\mathbf{v}_j = (v_{j1}, v_{j2}, \dots, v_{jN_f})$ とする。 \mathbf{p}_i 及び \mathbf{v}_j の次元は共に $N_f = 5$ として実験を行った。

エージェント i がアイテム j を推薦されたとき、エージェントはその満足度を表す効用 $s_{i,j}$ と、そこから求まる格付け $r_{i,j}$ を計算する。格付けは推薦システムに入力される評価であり、ユーザ-アイテムの行列として表現できる。本稿ではこの行列を格付け行列と呼ぶ。

3.2 嗜好ベクトルと特徴ベクトル

嗜好ベクトルの分布により、エージェント全体の嗜好の分布 (トレンドと呼ぶ) が設定される。本稿ではその典型的な例として、次の 3 つの分布を想定し、実験を行った。

- 一様分布 (トレンドが殆んど存在しない場合)
- 多変量正規乱数 (多数が似た嗜好を持つ場合)

- 二つの頂点を持つ分布 (全体の嗜好が 2 つのグループに分かれる場合)

多変量正規分布に基づく乱数は、各要素間で独立な p 変量正規乱数 $\mathbf{z} = (z_1, z_2, \dots, z_p)$ と、分散協分散行列 Σ をコレスキー分解 $\Sigma = LL^T$ することで得られる行列 L の積で求まる。本稿では、 Σ として対角要素が全て v で、非対角要素が全て 0 の行列を、 Σ_v と定義しこれを用いる。従って、 v が小さくなるほど強いトレンドを持つ。また、2 つの頂点を持つ分布は、予め平均ベクトル μ が異なり分散共分散行列 Σ が等しい 2 つの多変量正規分布を作成し、等確率でどちらかの分布に従うこととした。本実験では、2 つの平均ベクトルとして、要素が全て 0.5 及び -0.5 のベクトルとし、分散共分散行列は $\Sigma_{0.04}$ を使用した。

一方特徴ベクトルは、様々な特徴を持つアイテムが満遍なく分布している状況を想定し、一様分布に基づく乱数で生成した。

3.3 効用と格付け

本稿における効用 $s_{i,j}$ の定義を以下に示す。

$$\begin{aligned} f_{utility}(\mathbf{p}_i, \mathbf{v}_j) &= \exp(-\alpha \|\mathbf{p}_i - \mathbf{v}_j\|) \\ &= \exp\left(-\alpha \sqrt{\sum_{k=1}^{N_f} (p_{ik} - v_{jk})^2}\right) \end{aligned}$$

従って、 \mathbf{p}_i と \mathbf{v}_j に距離がある場合効用は 0 に近くなり、距離が近づくほどその効用は指数関数的に増加する。ここで α は予備実験により $\alpha = 0.5$ とした。

一方、格付けは商用 Web サイトなどでも良く用いられる 5 段階による格付けを使用した。従って、格付けを求める関数 $f_{rating}(s_{i,j})$ は効用を 5 段階に分ける 4 つの閾値を持った関数である。各閾値は、嗜好ベクトルを一様分布で与えた場合に、5 段階の格付けが同じ回数行われるよう予め予備実験により決定した。従って、嗜好ベクトルの分布を変えれば 5 段階の格付けの割合は変化する。

3.4 推薦アルゴリズム

本稿では、4 つの推薦アルゴリズムで実験を行った。**ランダム推薦** ランダム推薦は、他の推薦手法に対する基準であり、格付けの推定値に一様乱数を使用する手法である。

人気推薦 人気推薦は、アイテム毎の格付けの平均値が最も高いアイテムを推薦する。従って、格付けの平均値が最も高いアイテムをまだ格付けしていない全てのエージェントは同じアイテムを推薦される。

ユーザ間 CF ユーザ間 CF は Resnick らによってその基本アルゴリズムが提案された²⁾。本稿ではアルゴリズムの詳細は割愛する。

アイテム間 CF アイテム間 CF は、アイテム間の類似性から推薦するアイテムを求める手法である⁸⁾。本稿ではアルゴリズムの詳細は割愛する。

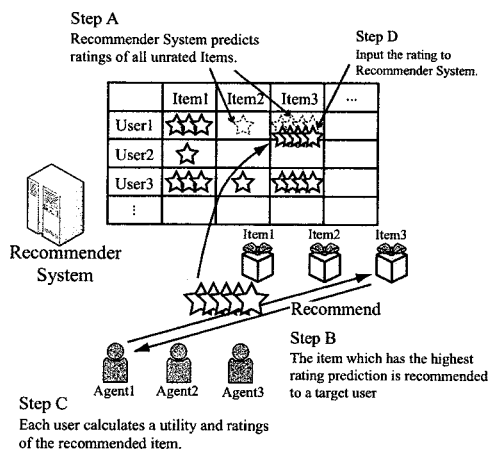


図1 シミュレーションの流れ

4. 実験

4.1 シミュレーションの流れ

本実験は初期設定の後、次の4ステップを1シミュレーションステップとして繰り返す(図1)。

初期設定: パラメータとして、 n, m, \mathbf{p}_i の分布、推薦アルゴリズムを決定する。

ステップA: 推薦システムは、収集済みの格付けから、未入力の格付けを推定する。初期状態では収集済みの格付けが無いので、推薦アルゴリズムによらずランダム推定になる。

ステップB: エージェント毎に格付けの推定値が最大となるアイテムを1つ推薦する。

ステップC: エージェントは推薦された商品が無条件に購入するものとし、その効用及び格付けを計算する。つまり、ステップ t において全エージェントが t 個の格付けをする。ここで、効用の平均値(以下 utility average)はステップ t におけるエージェント全体の満足度と捉えられる。実験結果では utility average の推移を観察し考察する。

ステップD: 各エージェントは推薦され購入した商品の格付けを推薦システムに入力する。その後ステップAに戻り繰り返す。

4.2 実験1

本稿の実験結果は全て10回の試行の平均値である。実験結果を説明するために、本稿で使用した2つの指標について述べる。一つ目は、推薦効果が現れるまでのステップ数(立ち上がりステップ)である。本稿では、ランダム推薦に対して average utility が0.5高くなったステップと定義する。二つ目は最も推薦効果が高いときの満足度(満足度のピーク値)である。本稿

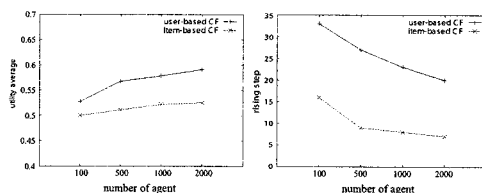


図2 エージェント数を変化させたときの満足度のピーク値(左)及び立ち上がりステップ(右)

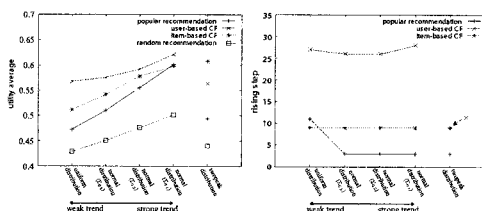


図3 嗜好ベクトルの分布を変化させたときの満足度のピーク値(左)及び立ち上がりステップ(右)

においては、立ち上がりステップ $t = r$ から $t = r + 30$ までの全エージェントの効用の平均値と定義する。

図2は、 $n = 100, 500, 1000, 2000, m = 500, \mathbf{p}_i$ を一様分布で与えたときの結果である。左に満足度のピーク値、右に立ち上がりステップを示す。横軸は n である。また、図3は $n = 500, m = 500$ において、 \mathbf{p}_i の分布を変更して実験した結果である。

図2から、両CFにおいて、 n の増加に伴って満足度のピーク値が上昇し立ち上がりステップが短縮しているが、アイテム間CFの立ち上がりステップのほうが短いことが分かる。しかし、図2(左)及び図3(左)から、満足度のピーク値はユーザ間CFの方が良いことが分かる。この原因を調べるために、両CFの推薦の特徴について分析を行った。

その結果、ユーザ間CFでは全てのアイテムがほぼ偏り無く推薦されているのに対し、アイテム間CFでは何度も推薦されるものと全く推薦されないものがあることが明らかとなった。つまり、アイテム間CFはその時点での人気アイテムを積極的に推薦する手法であり、ユーザ間CFの方がより広い視野でエージェントの嗜好に合わせたアイテムを選択していると言える。

この現象は類似度計算の仕組みに起因していると推測される。アイテム間CFでは、何度も格付け(推薦)されるアイテムはアイテム間類似度が計算可能であり、再び推薦される機会を得やすいが、逆に格付けされないアイテムは推薦されにくい。一方で本実験では必ず全てのエージェントが同じ回数格付けを行うため、全てのユーザ間類似度が計算可能である。そのため、ユーザ間CFの方が全てのアイテムについて偏り無く推薦される結果になる。

また図3(左)から、トレンドが強まるに従って満足

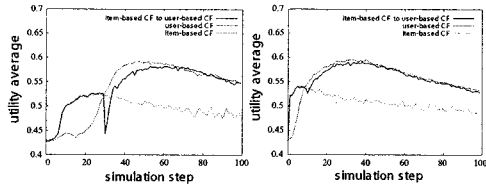


図4 アイテム間CFからユーザ間CFに切り替えた時のエージェント満足度の推移

度が向上することが分かるが、ランダム推薦も向上していることに注意が必要である。本モデルでは、トレンドが強くなるにつれて、嗜好ベクトルは空間における中心に集まりやすいため、全体の満足度が増加する。

4.3 実験 2

次に、予め一定の格付けが入力されている状況から新規エージェントに対する推薦の特性について分析する。まず、 $n = 100, 500, 1000, 2000, m = 500, \mathbf{p}_i$ を一様分布で生成するものとし、予め半数のエージェントに対して100個のアイテムをランダム推薦する。その状況で、残りの半分のエージェントに対してユーザ間CFとアイテム間CFで推薦を行った。

その結果、アイテム間CFは2つ目の推薦から効果を得られるのに対し、ユーザ間CFは人数が増加しても推薦効果を得るのに5つ程度の格付けが必要であった。この原因はアイテム間類似度は予め計算可能であるのに対して、新規エージェントについてのユーザ間類似度は計算不可能であることが原因として挙げられる。一方推薦効果のピーク値はユーザ間CFの方が良い結果となった。

4.4 実験 3

実験1、実験2より、ユーザ間CFのほうが推薦効果のピーク値が高く、アイテム間CFの方が立ち上がりステップが早いことが明らかになった。そこで、最初はアイテム間CFで推薦し、推薦効果が逆転する地点でユーザ間CFに切り替えることを考える。予め得た結果より実験1の設定では $t = 30$ 付近で逆転し、実験2の設定では $t = 10$ で逆転することが明らかとなっている。

図4(左)は、全く格付けが無い状態から $t = 30$ で切り替えた場合の図であり、また図4(右)は、実験2の設定において、 $t = 10$ で同様に切り替えた場合の結果である。どちらの図も、切り替えなかった場合の結果のグラフもプロットしている。図4(左)によると、切り替えた瞬間に満足度が落ち込み、すぐに元よりも高い値まで増加している。切り替える位置を変更した場合、 $t = 50$ より後であれば満足度の落ち込みが少なく押さえられることを確認している。

一方で、図4(右)では切り替えた瞬間の満足度の落ち込みは僅かである。即ち、推薦システムの運用を考慮すると、すでに多数のユーザが使用している状況

において新規ユーザに対しては、10回程度までの推薦はアイテム間CFを使用し、それ以降の推薦はユーザ間CFによる推薦に切り替えたほうが効果的であることが示された。

5. 結 論

本稿では、推薦システムとユーザ間の相互作用に着目し、ユーザを購入格付けエージェントとしてモデル化することでユーザ間CF及びアイテム間CFの特徴分析を行った。その結果、ユーザ間CFによる推薦とアイテム間CFによる推薦には、その特性に明確な違いがあることが分かった。また、新規ユーザに対して推薦を行う場合、アイテム間CFの方がより早く効果的な推薦を実現できる一方で、ユーザ間CFの方が満足度のピーク値が高く、途中で切り替える手法が有効であることが明らかとなった。

より現実に近いシミュレーション結果を得るためには、実データによるモデルの拡張と検証が必要であるが、本稿で示した結果は、特定のユーザやアイテムに依存しない一般的な特性として新たな知見を獲得したものである。

参 考 文 献

- 1) J. B. Schafer, et al.: E-Commerce recommendation applications. *Data Mining and Knowledge Discovery*, Vol.5, pp.115-153 (2001).
- 2) P. Resnick, et al.: GroupLens: An Open Architecture for Collaborative Filtering of Netnews In *CSCW '94: Conference on Computer Supported Cooperative Work*, ACM, pp. 175-186 (1994).
- 3) J.L. Herlocker, et al.:Evaluating Collaborative Filtering Recommender Systems, *ACM Transactions on Information Systems*, Vol.22, No.1, pp.5-53 (2004).
- 4) G. Adomavicius et al.: Toward the Next Generation of Recommender Systems, *IEEE Transactions on Knowledge and Data Engineering*, Vol.17, No.6, pp.734-749 (2005).
- 5) 土方嘉徳：嗜好抽出と情報推薦技術，*情報処理*，Vol.48, No.9, pp.957-965 (2007).
- 6) B. Sarwar, et al.: Analysis of Recommendation Algorithms for E-Commerce, *Proceedings of the 2nd ACM conference on Electronic commerce*, p.158-167, October 17-20, 2000, Minneapolis, Minnesota, United States (2000).
- 7) B. Sarwar, et al.: Item-based Collaborative Filtering Recommendation Algorithms, *Proc. of WWW '01*, pp.285-295 (2001).