

情報推薦における評価履歴及び 相対的評価に基づいた調整型評価手法の検討

山内 一騎† 當間 愛晃‡
赤嶺有平‡ 山田孝治‡ 遠藤聡志‡

琉球大学大学院理工学研究科情報工学専攻† 琉球大学工学部情報工学科‡

1 はじめに

情報推薦において、推薦システムがユーザに有益な情報を提供することは最も重要な課題である。ユーザの嗜好に適した推薦結果を提示することが推薦システムの有用性を決定づけるという観点から、情報検索の分野で確立された評価指標を適用して、予測の正確さ (accuracy) が評価されることが多かった。しかし、推薦システムの有用性は、ユーザの現在の嗜好に適した推薦結果を提示することに加え、ユーザに未知の情報 (novelty) や多様な情報 (diversity)、意外な情報 (serendipity) などを提供することによる新たな嗜好を提供することにも大きく関係している。その中でも、近年では、意外性が着目され、これを評価する指標 [1] や推薦手法 [2] が提案されてきた。しかし、仮に意外性を完全に評価できる指標が定義されたとしても、それが全ユーザにとっての最適な評価値であるという可能性は考えられず、さらに同ユーザにおいても時間などの違いによって変化する可能性が大いに考えられる。本稿では、これらのように変動する評価値に対応するため、パラメータによる調整型評価値を提案する。本評価値では、協調フィルタリング (CF) などといった従来の推薦結果に基づいた相対的な評価を導入すると共に評価履歴を加え、それらをパラメータにより調整する。各パラメータと従来の推薦手法の関係性を分析した上で本評価値における妥当性を検討した。

2 提案手法

本稿では、評価履歴及び相対評価に基づいた調整型評価手法を提案する。この評価手法は、先に述べたように、ユーザによる最適な評価値の違いや同ユーザにおいても時間による最適な評価値の違いに対応するためにパラメータを用意することで評価値を調整可能にした。

ユーザが求める情報とは、novelty や diversity、serendipity だけでなく、多種多様なものが存在する。これらの種類は serendipity のように完全に定義されていないものだけでなく、そもそも認識すらされていないものも考えられ、ユーザ自身、求めている情報を明確に把握することは非常に難しい。そこで本稿では、ユーザが求めない情報を推薦するシステム (不満足な推薦システム) を基準にし、評価される推薦システムとの差異を比較することにより、ユーザが求める情報を推薦するシステム (満足な推薦システム) を評価できるようにした (式 (1))。

$$D(\text{Rank}_{i,j}) = \frac{P(\text{Rank}_{i,j})}{\text{base}(\text{Rank}_{i,j})} \quad (1)$$

$\text{Rank}_{i,j}$ はユーザ i にコンテンツ j が推薦される順位であり、基準となる不満足な推薦システムによる順位を $\text{base}(\text{Rank}_{i,j})$ 、評価される推薦システムによる順位を $P(\text{Rank}_{i,j})$ とする。

不満足な推薦システムによって上位にソートされる (推薦される) コンテンツとは、ユーザが求めない情報であり、逆に下位にソートされる (推薦されない) コンテンツは、ユーザが求める情報である。そこで、 $D(\text{Rank}_{i,j})$ では、不満足な推薦システムによって上位にソートされるコンテンツ (ユーザが求めない情報) が評価される推薦システムで下位にソートされるほど点数が高く、逆も同様に点数が高くなるようにしている。よって、 $D(\text{Rank}_{i,j})$ が高いほど、満足な推薦システムといえる。この評価値はユーザによって求める情報が異なったり、時間によって変化しようとも、この式 (1) のようにユーザにとって不満足な推薦システムを基準にすることにより、その逆である満足な推薦システムとして評価できるようになる。さらに、これを正規化し、パラメータ ($p1$) を付与した (式 (2))。なお、 n はユーザ総数、 m はコンテンツ総数である。

$$p1 \sum_i^n \sum_j^m \frac{D(\text{Rank}_{i,j}) - \min(D(\text{Rank}_i))}{\max(D(\text{Rank}_i)) - \min(D(\text{Rank}_i))} \quad (2)$$

また、これに付随して、どんなに novelty や diversity、

An examination of an adjustable evaluation by user's rating and relative evaluation on Information Recommendation
Kazuki YAMAUCHI†, Naruaki TOMA‡, Yuhei AKAMINE‡, Koji YAMADA‡, Satoshi ENDO‡
Graduate School of Eng. and Sci. Univ. of the Ryukyus†, Faculty of Engineering, Univ. of the Ryukyus‡

serendipity な情報が推薦されたとしてもその情報自体の評価 (accuracy) が低ければ、ユーザに最適でないという観点から、式 (3) を用意し、式 (2) に足しあわせた (式 (4))。この観点とは逆に、コンテンツの評価値 (accuracy) よりも novelty や diversity、serendipity といった情報に価値を置くユーザも十分考えられるのでパラメータ (p2) を付与した。なお、 $Rating_{i,j}$ はコンテンツ j に対するユーザ i の評価値である。

$$p2 \sum_i^n \sum_j^m \frac{Rating_{i,j} - \min(Rating_i)}{\max(Rating_i) - \min(Rating_i)} \quad (3)$$

$$\text{式 (2) + 式 (3)} \quad (4)$$

3 実験

本実験ではデータセットとして、MovieLens を用いた。ユーザ数は 943、映画数は 1682、評価数は 100,000 であり、その中からランダムで選んだ 9 割を学習データ、残り 1 割をテストデータとした。各パラメータにおける従来の推薦手法の関連性を分析するために、本提案手法のベースとなる推薦手法にベイズ推定、コンテンツベース CF(CCF) を用いて、既存の推薦手法であるユーザベース CF(UCF) を評価した。なお、本提案手法で評価されるコンテンツは推薦システムに推薦される上位 3 つのコンテンツである。

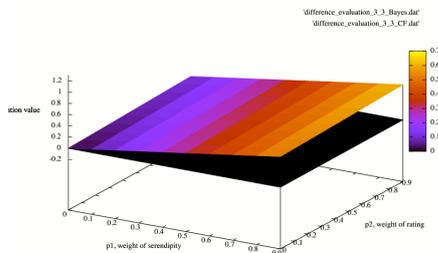


図 1: ベイズ推定をベースにした UCF

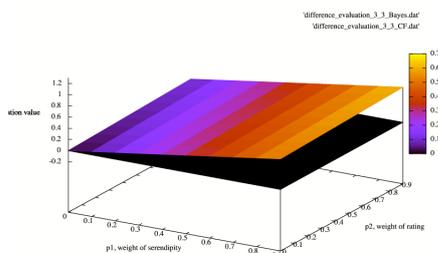


図 2: CCF をベースにした UCF

図 1 及び、図 2 の X 軸は式 (2) のパラメータ (p1) の値であり、Y 軸は式 (3) のパラメータ (p2) の値である。

Z 軸は式 (4) の値である。図 1 はベイズ推定をベースにして評価した UCF である。パラメータ 1(p1) に比例して Z 軸が高くなっていることがわかる。これは、ベイズ推定に比べ UCF の方が式 (2) の得点が高く、その分だけベイズ推定と UCF の推薦結果に違いがあるということである。また、パラメータ 2(p2) も Z 軸に比例して高くなることから、UCF の方が式 (3) の得点が多く、ベイズ推定よりもコンテンツの評価 (accuracy) が高い推薦ということである。図 2 は CCF をベースにして評価した UCF である。この図から MICHAEL ら [1] が調査した結果と同じく CCF より UCF の方が優れていることが確認できる。また、全体的に見ると、図 1 と同じ傾向だが、Z 軸が約 6 倍低い。これもまた、MICHAEL ら [1] が調査した結果と同じく、両者の推薦結果はさほど変わらないということが確認できる。このことから、ベイズ推定と UCF の違いと CCF と UCF の違いを数値化できた。

4 おわりに

本稿では、パラメータによる調整型評価値を提案し、各パラメータと従来の推薦手法の関連性を分析した上で本評価値における妥当性を検討した。今後の課題として、今回は MovieLens によるデータセットで実験を行い、実験結果の偏りを防ぐため全ユーザで評価し、各パラメータによる実験結果を示したが、本提案手法はユーザそれぞれに合ったパラメータを用いることを前提としているため、実際にユーザにアンケートを取り、実験結果から得た数値の違いと満足度の違いと関連性を確認することがあげられる。

参考文献

- [1] 村上知子, 森紘一郎, 折原良平, “推薦結果の意外性を評価する指標の提案”, The 21st Annual Conference of the Japanese Society for Artificial Intelligence, 2C5-2, 2007
- [2] Oku, et al., “User Evaluation of Fusion-based Recommender System for Serendipity-oriented Recommender System,” Proc. of the Workshop on Recommendation Utility Evaluation: Beyond RMSE, at the 6th ACM International Conf. on RecSys 2012 (Web), pp.(6 pages), Dublin, Ireland, 2012.
- [3] MICHAEL J.PAZZANI, “A Framework for Collaborative, Content-Based and Demographic Filtering”, Artificial Intelligence Review 13: 393-408, 1999.