

## 嗜好の個人差と状況依存性を考慮した映画推薦システムの評価方法の検討と基礎評価

小野 智弘<sup>†</sup> 本村 陽一<sup>‡</sup> 麻生 英樹<sup>‡</sup>

株式会社 KDDI 研究所<sup>†</sup>

独立行政法人 産業技術総合研究所<sup>‡</sup>

近年の情報やコンテンツの爆発的な増大およびユーザーニーズの多様化に伴い、ユーザが欲する情報の選択を支援する推薦システムへの要求が益々高まっている。ユーザの嗜好は例えば映画鑑賞の場合、家族と観るか恋人と観るか、元気であるか沈んでいるか等、状況や気分に応じて変化するため、ユーザ毎の差異に加えて同一ユーザ内でも状況に応じた差異を考慮する必要がある。筆者らはこれらの問題を解決するためにユーザの履歴やプロフィール等の情報と気分や場所などの状況の依存関係をベイジアンネットによりモデル化し、個人差と状況に応じて最適なコンテンツを推薦する方式およびシステムの研究を行っている。推薦システムの有効性の評価については推薦精度が高いことが必ずしもユーザ満足度に繋がらない等困難であり、これまでに統一的な評価方法は確立されていない。本稿では、推薦システムの評価方法について検討するとともに実装した映画推薦システムを用いた基礎的な評価結果について述べる。

## Evaluation of Movie Recommendation System considering both users' personality and situation.

Chihiro Ono<sup>†</sup>, Youichi Motomura<sup>‡</sup> Hideki Asoh<sup>†</sup>

KDDI R&D Laboratories Inc.<sup>†</sup>

National Institute of Advanced Industrial Science and Technology<sup>‡</sup>

With the flood of various information and contents through the Internet, the need for recommendation systems that assist users in finding the information they desire is increasing. For realizing timely movie recommendation, both the user's personality and his situation should be considered at the same time as user preference may change according to his situation such as mood and location etc. We have proposed a recommendation system which achieves context-aware personalized recommendation based on Bayesian network by making use of various information related to user profiles, histories, situations, and content attributes. So far there is no perfect methodology for evaluating usefulness of the recommendation systems as the accuracy is not necessarily connected with the user satisfaction. In this paper, we discuss the evaluation methodology for a movie recommendation system and show the result of basic experiments.

### 1. はじめに

近年の情報やコンテンツの増大およびユーザーニーズの多様化を受け、ユーザが欲する情報の選択を支援する推薦システムへの要求が益々高まっており、アマゾン社の書籍の推薦や TSUTAYA online のレンタル・販売 DVD の推薦等効果を挙

げている例も報告されている[1][2].

これまでにユーザ個別の嗜好に基づく推薦を行う枠組みとして、ユーザの購買やコンテンツ評価履歴からユーザの嗜好傾向を探り、嗜好傾向の類似した他人が好むコンテンツを推薦する協調フィルタリング方式や、コンテンツそのものや属

性群からユーザの嗜好構造を反映する特徴の抽出を行い、ユーザの好きなコンテンツ群の特徴と類似した特徴を持つコンテンツを推薦するコンテンツベース（属性ベース）の方式が提案されている[3]。これらは一定の効果があるものの、前者は大量の履歴データが必要なこと、後者は適切な特徴抽出が困難といった欠点がある。このため、両者を組み合わせた方式として、両者の結果を重み付けして結合した方式[4][5]や、アスペクトモデルと呼ばれる潜在クラスを用いた統計的混合分布モデルで結合した方式[6][7][8][9]が検討されている。

さらに、たとえば映画の推薦を対象とする場合、家族と観るか恋人と観るか、元気であるか沈んでいるか等、状況や気分に応じてユーザの嗜好が変化するが、これまでの推薦方式にはユーザがどんな状況でも同じ結果になってしまうという問題がある。ユーザの個性と状況を同時に考慮した推薦を行うためには、ユーザの履歴やプロフィール等の情報と気分や場所などの状況の情報をユーザの個性や状況に応じて重み付け等を変更しながら統合する必要があるが、データ間には複雑な依存関係があり、既存方式の単純な組み合わせでは実現は困難である。

近年、確率変数間の複雑な依存関係をグラフ構造によって表現する手法として、ベイジアンネットワーク[10]が注目されており、ソフトウェアの機能推薦[11]やプリンタの故障診断[12]等で利用されている。そこで筆者らは上記のような問題を解決するために、ユーザの履歴やプロフィール等の情報に加えて気分や場所などの状況も含めた複雑な依存関係をベイジアンネットによりモデル化し、個人差と状況に応じて最適な映画コンテンツを推薦する方式を検討・提案してきた[13, 15]。

本稿では、提案システムの有効性の評価について検討する。推薦システムやアルゴリズムの有効

性をどのように評価すればよいかという問題については、推薦性能が対象領域やデータの性質によって大きく変化する、推薦精度が高いことが必ずしもユーザ満足度に繋がらない等、考慮すべき点が多く、評価方法が確立されているとはいえない。そこで、本稿では、まず、推薦システムの評価方法のあり方について、既存の研究[4~8, 16]をもとに整理する。さらに、我々のシステムに対して基礎的評価を行った結果についても報告する。

以下、まず2章では映画推薦タスクの性質について検討する。3章ではベイジアンネットワークを用いた映画推薦システムの概要について述べる。4章では推薦システムの評価を考える上での重要な観点について検討し、5章では実施した基礎評価結果について述べる。6章はまとめである。

## 2. 映画推薦タスクの性質

Herlocker[16]でも述べられているように推薦システムの評価をするためには、まず、その推薦システムが行うタスクの性質を明確化することが不可欠である。タスクが明確化されて初めてそれにあつた評価用のデータセットや評価尺度の選択を行えるようになる。以下では、筆者らの推薦システムのタスクの性質について、[16]に倣って、6つの観点から検討する

- 1) 推薦システムのタスク：ユーザがタイムリーに見たいと思われる映画を推薦する。本研究では特にタイムリー性に対する要求が強い携帯電話上での映画推薦を対象とし、ユーザの嗜好の個人差と状況依存性を同時に考慮した推薦を行うことを目標とする。
- 2) 推薦リスト：出力する推薦リスト数は表示デバイスや応用の種類により決定される。PC画面では10~20個程度、画面サイズの小さい携帯電話は3~5個程度であることが多い。
- 3) 推薦対象とのコンテンツ集合：映画館でみる場合とレンタルやDVDで見る場合で異なる。

- (ア) 映画館：推薦対象となる上映中のコンテンツ数は、2005年2月15日時点で旧作上映や単館上映も含めて約200本である（全国規模の上映は42本）
- (イ) レンタルやDVD販売：在庫数は店舗によるが数千～1万本規模となる。  
また推薦に利用するコンテンツ情報は、コンテンツに付随するジャンルや製作年、映画の解説文等が利用可能である。
- 4) ユーザ集合：サービス提供サイトの会員と非会員の場合で性質が異なる
- (ア) 会員：ユーザ数はそれほど大きくないが、会員登録時等にユーザ属性を細かく取得できる
- (イ) 非会員：ユーザ数は大きいがユーザ属性は取得できない。ただし、Cookie ID等で対応づけることにより継続的に履歴を取得可能なことも多い。また、推薦要求を受けた際にユーザに簡単な質問によりユーザ情報を取得したあとに推薦すること等も考えられる
- 本研究では会員登録情報または推薦要求時の質問により、ユーザ情報を得られることを想定する。
- 5) ユーザの嗜好評価の性質と粒度：明示的評価と暗示的評価の2種類がある。
- (ア) 明示的評価：ユーザがアンケートや推薦結果に対するフィードバックにより各コンテンツに対する評価を入力する。通常は5段階で評価されていることが多い。
- (イ) 暗示的評価：映画情報サイトにおけるコンテンツページへのアクセスやDVDの購買をコンテンツへの評価とみなす。  
本研究では明示的評価を想定する。
- 6) 嗜好評価の密度：  
コンテンツの密度 = 評価値数 / (コンテンツ数

\*ユーザ数) とすると、明示的評価では密度は通常あまり高くなく、ユーザやコンテンツによりかなりばらつきがある。ミネソタ大のMovieLensプロジェクトで公開しているデータセットはユーザ数が6040、コンテンツ数が3900、密度が4.2%程度であった。

### 3. 推薦システムの概要

筆者らはこれまでにユーザの履歴やプロフィール等の情報と気分や場所などの状況の依存関係をベイジアンネットによりモデル化し、個人差と状況に応じた最適なコンテンツを推薦する推薦システムを提案してきた。本節ではこの概要を述べる。

#### 3.1. データ

2004年12月（第1回目）と2005年8月（第2回目）と2回に分けてWEBアンケート調査によりモデル化のためのデータを取得した。

第1回目はユーザ属性、コンテンツ属性、コンテンツ評価履歴に関するデータを取得した。概要は以下の通り。

- 1) 被験者数：1616人
- 2) リクルート条件：提示した100コンテンツ中20コンテンツ以上視聴経験のある被験者
- 3) 調査項目（各ユーザ）：
  - (ア) デモグラフィック・ライフスタイル属性：年齢、性別、職業等27項目
  - (イ) 映画視聴に関する態度属性：映画コンテンツの鑑賞頻度、映画選択時の重視項目(7項目)、ジャンル選好度(12項目)、映画を見る主要目的（感動したい等7項目）
  - (ウ) ユーザのコンテンツ評価履歴：視聴経験のある20コンテンツ（ランダムに抽出）について、評価{良い・悪い}、その時の気分（感動した等7項目）
- 4) コンテンツ属性（100コンテンツ分）：
  - (ア) メタ属性：ジャンル、収録時間等
  - (イ) 特徴属性：映画紹介文からテキストマ

イニングで抽出したキーワード群

第2回目は状況属性に関するデータを取得した。概要は以下の通り。

- 1) 被験者数：1088人
- 2) リクルート条件：提示したコンテンツのうち15件以上が見たことの有る被験者
- 3) 調査項目（各ユーザ）：以下の項目を自由記述で取得
  - (ア) どこ（場所）
  - (イ) 誰と何人
  - (ウ) どんな映画どんな時

### 3.2. モデル構築の概要

図1に示す評価構造の仮説に基づき、ユーザ情報、コンテンツ情報、ユーザの評価情報、状況情報間の依存関係をベイジアンネットワークでモデル化した。この評価構造は、ユーザのコンテンツ評価が、コンテンツに対する感じ方によって決まり、感じ方はユーザ属性、状況属性、コンテンツ属性によって決まることを意味している。

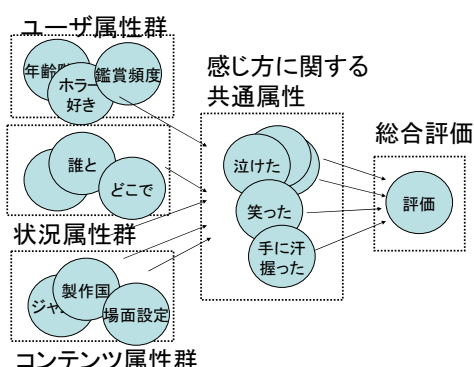


図1 評価構造の仮説

モデル構築ではベイジアンネットワーク構築ツール BayoNet [14] を利用した。詳細な手順は文献 [15] を参照されたい。構築したモデルは、ユーザ属性群 (22 ノード) とコンテンツ属性群 (6 ノード)、状況属性群 (3 ノード)、共通属性 (7 ノード)、総合評価の合計 39 ノードとなった。

### 3.3. 推薦システムの実装

構築したユーザ嗜好モデルを用いて映画推薦システムのプロトタイプを構築した。

ユーザに最適なコンテンツを発見するためには、推薦要求があったユーザに対する全コンテンツの予測評価を求め、予測評価値の高いコンテンツから順に薦めることが考えられるが、候補となるコンテンツ数と同じ回数分の推論が必要となり、計算時間の観点から現実的ではない。そこで推薦要求時の推論回数を抑えるために、以下の手法を提案した。まず、各コンテンツに対応する共通属性の事後確率を平均的なユーザを想定して（つまり、ユーザ属性値を設定せずに）あらかじめ計算してコンテンツ特徴ベクトルとして保存しておく。そして、ユーザからの推薦要求時には、ユーザ属性と状況をエビデンスとして共通属性の事後確率を求め、保存済みのコンテンツ特徴ベクトルとのマッチングを行い、近いコンテンツから順に推薦を行う。

推薦システムの処理の流れを図2に示す。

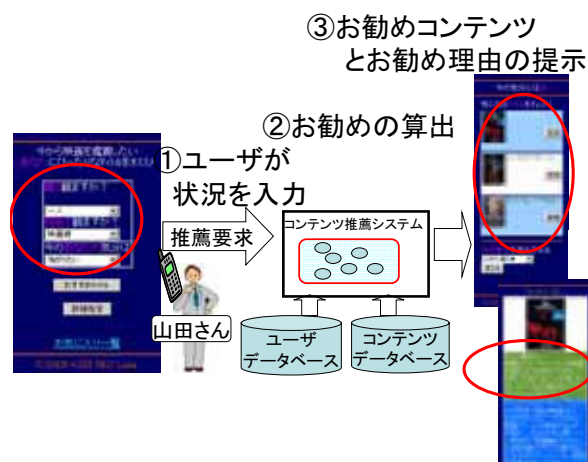


図2 コンテンツ推薦システムの概要

- 1) ユーザから携帯電話またはPCWEB経由で状況に関する情報とともにコンテンツ推薦要求を受け付ける。
- 2) データベースから登録済みのユーザ属性情報を取得し、状況を利用して特徴ベクトル (= 共通属性の値) を算出する。
- 3) ユーザの特徴ベクトル値とあらかじめ計算済みのコンテンツ群の特徴ベクトルとの類

似性を判定し、類似度の高いコンテンツを推薦する。

- 4) 推薦したコンテンツに対するフィードバックをユーザから受付け、予測精度の向上のための学習に用いる。

#### 4. 推薦システムの評価

##### 4.1. 推薦システムの評価方法

映画推薦システムの評価には、以下の5つの重要な観点がある。

- 1) 推薦精度：システムが推薦した作品をユーザが実際に好むかどうかという観点である。尺度としては予測誤差やランキング精度が用いられる。予測誤差はシステムの予測評価値と正解値との平均絶対誤差 (MAE) や平均2乗誤差 (MSE) をとったもの、ランキング精度はユーザのランキングと推薦システムのランキングの相関や順位相関をとったものが用いられる。ただし、ユーザが評価済みのデータに対する誤り率は、そもそもユーザが評価をつけたものは関心があるものだという偏りがあるため正確とは限らない[16]等の指摘もある。
- 2) 新規性と思わぬ発見：どれくらいあたりまえでないコンテンツを提示できるかという観点で、新規性 (Novelty) と思わぬ発見 (Serendipity)がある[16]。両者の違いは、新規性は（自分で見つけられたかもしれないが、たまたま）気づかなかったコンテンツを見つけてくれることを指し、Serendipity とは、推薦してくれなければ自力では思いもよらなかったコンテンツを教えてくれることを指す。この観点は推薦システムがユーザにとって魅力的なものになるために重要であるが、評価が難しいため、研究があまり進んでいない。
- 3) 満足度：ユーザの推薦システム全体に対するユーザの主観的な満足度で、アンケート調査

等によるオフライン評価と、その場でのフィードバックによる評価がある。

- 4) 有用性：推薦機能の付与による商業的な効果の観点で、サイトアクセス数、サイト訪問率等の尺度で測定する。新規コンテンツやユーザ追加の容易性等も評価対象となりうる。

- 5) 性能：推薦システムの応答速度や対応できる同時接続ユーザ数やコンテンツ数等である。

その他、カバー率（ユーザが興味をもつコンテンツのうち、どれくらいの割合を推薦できているか）、信頼性（システムの推薦結果がどれくらい信頼できるか）等の尺度もある。

##### 4.2. 本研究特有の考慮点

本研究特有の考慮点として、以下の3点があげられる。

- 1) 精度向上率：提案する推薦システムはあらかじめ収集したデータを用いてオフラインで構築した初期モデルを利用するとともに、運用中にフィードバックにより動的にモデル精練を行う。このため、ユーザに関する履歴を蓄積することによる推薦精度の向上度合いという観点が考えられる。

- 2) 予測値の捉え方：ベイジアンネットによるユーザ嗜好モデルでは、ユーザによるコンテンツ評価値の事後確率が出力される。従って、予測値として、評価の各値を事後確率に基づき加算した期待値やMAP値（最大事後確率を持つ値）を用いることが考えられる。

- 3) 推論回数と推薦精度：提案するシステムで行っている要求時の推論回数削減のためのマッチング処理の結果は、コンテンツとユーザを与えた際の予測評価値とは必ずしも一致しないため、許容できる推論回数と推薦精度とのトレードオフを考慮する必要がある。

##### 5. 推薦システムの基礎評価

評価の手始めとして、3.1 節で述べたアンケートデータから生成した初期モデルを用いて、4.2 節の3)に述べた観点、すなわち、推論回数削減の

ためのマッチングによるコンテンツランキングとコンテンツとユーザを与えた際の予測評価値によるコンテンツランキングとの差の検証を行った。なお、検証にはモデル構築に利用していないデータを用いた。具体的には 2005 年 12 月～2006 年 2 月に実施した映画視聴に関する WEB 調査で取得した 787 人分の評価データを用いた。

具体的にはアンケート結果に現れたコンテンツから抽出した 20 コンテンツを対象として、ユーザ毎に、ユーザ属性とコンテンツ属性を入力して求めた予測評価値によるランキングと、下記の 3 通りの条件でのマッチング方式によるランキングとを算出し、二つのランキングのスピアマンの順位相関を求めた。なお、コンテンツ特徴ベクトルとユーザ属性・状況から計算されるベクトルとの間の近さの計算にはユークリッド距離を用いた。

試行	対象ユーザ	マッチングに用いた属性	順位相関の中央値
試行 1	全員	共通 7 属性	0.257
試行 2	全員	3 つの共通属性	0.495
試行 3	注 1)	3 つの共通属性	0.539

注 1) 3 つの共通属性に高い評価をしたユーザ群

表 1 予測評価値によるランキングとの順位相関

試行 1, 2, 3 それぞれの順位相関の全ユーザについての中央値を表 1 に示す。なお、試行 2, 3 の 3 つの共通属性とは、総合評価の期待値との相関が高く互いに独立な“感動した”, “笑った”, “手に汗握った”を採用した。この結果、試行 1 よりも 2, 3 が高い順位相関を示していることから、ユーザ (セグメント) 毎に適切な共通属性の組を用いることが効果的であるという示唆を得た。今後はさらに細かい分析を進め、モデルを精練していく予定である。

## 6. おわりに

本稿では、ベイジアンネットによるユーザの個人差と状況を考慮したユーザモデルおよび推薦

システムの評価方針の検討ならびに基礎評価結果を報告した。今後は 4 章で述べた新たな観点も含めて評価をさらに進める予定である。最後に日頃ご指導頂く KDDI 研究所秋葉重幸代表取締役所長に深く感謝致します。

## 参考文献

- [1] <http://www.amazon.co.jp>
- [2] <http://www.tsutayaonline.co.jp>
- [3] P. Resnic and H.R. Varian, “Recommender systems,” Communications of the ACM, 40, No.3, 56-58, 1997
- [4] M. Balabanovic and Y. Shoham, “Fab: Content-Based, Collaborative Recommendation,” Communications of the ACM, 40, No.3, pp. 66-72, 1997
- [5] C. Basu, H. Hirsh, and W. W. Cohen, “Recommendation as classification: Using social and content-based information in recommendation,” Proc. 15<sup>th</sup> National Conference on Artificial Intelligence, pages 714--720, Madison, WI, July 1998
- [6] Hofmann, T., Probabilistic latent semantic analysis, In Proceedings of the Fifteenth Conference on Uncertainty in Artificial Intelligence, pp. 289-296, 1999
- [7] Hofmann, T. and Puzicha, J.: Latent class models for collaborative filtering, In Proceedings of the Sixteenth International Joint Conference on Artificial Intelligence, pages .688-693, 1999
- [8] A. Popescul, L. H. Ungar, D. M. Pennock, and S. Lawrence, “Probabilistic models for unified collaborative and content-based recommendation in sparse-data environments,” Proc. 17<sup>th</sup> Conference on Uncertainty in Artificial Intelligence, 2001
- [9] 小野, 本村, 麻生, “確率モデルによる映画推薦方式の検討”, 人工知能学会研究会資料, SIG-FPAI-A502-02, pp7-12, 2005
- [10] 本村陽一, “ベイジアンネットワーク”, 学習システムの理論と実現 (渡辺澄夫著), 4 章, 森北出版, 2005.
- [11] E. Horvitz, “Principles of Mixed-Initiative User Interfaces,” Proc. ACM SIGCHI Conference on Human Factors in Computing Systems, 1999.
- [12] F.V. Jensen et.al., “The SACS methodology for troubleshooting complex systems,” Artificial Intelligence for Engineering Design, Analysis and Manufacturing(AIEDAM), Vol. 15, pp. 321–333, 2001.
- [13] 小野, 本村, 麻生, “ベイジアンネットによる映画コンテンツ推薦方式の検討”, 信学技法, NC2004—66, pp.55-60, 2004
- [14] 本村, “ベイジアンネットソフトウェア BayoNet” 計測と制御, vol.42, No.8, pp.693-694, 2003
- [15] 小野, 本村, 麻生, “嗜好の個人差と状況依存性を考慮した映画推薦方式の検討 情報処理学会 DPS 研究会研究報告 DPS-125, PP79-84, 2005
- [16] J, Herlocker et al, “Evaluating Collaborative Filtering Recommender Systems”, ACM Transactions on Information Systems, Vol.22, No.1, pp5-53, 2004