

## 7 音楽推薦システム



吉井 和佳      後藤 真孝

産業技術総合研究所

近年 iTunes Music Store や Last.fm<sup>☆1</sup> などの数百万曲を取り扱うオンライン音楽配信サービスが多数のユーザを獲得している。(財)デジタルコンテンツ協会の報告によると、2007年の国内デジタルコンテンツ市場規模は2兆6,947億円であり、このうち音楽コンテンツが占める割合は28.5%の7,677億円であった。近年は映像コンテンツの伸びが著しいが、依然として音楽コンテンツは重要な位置を占めている。この種のサービスにおいて注目すべき現象として、**ロングテール効果**がある。図-1に示すように、全楽曲を人気度順に並べると、楽曲の再生回数(購入回数)は指数的に減少していくが、ほとんど最後までゼロにはならない<sup>☆2</sup>。このように、どの曲も誰かに必要とされているのであるが、ニューリリースや週間ランキングを参考にして人気が高い楽曲を聴くだけでは、好みに合った楽曲を発見できる範囲は限定されてしまう。

こうした状況を受けて、各々のユーザに適切な楽曲との出会いを提供するため、音楽推薦システムの重要性は増している。本稿では、音楽推薦システムに関して、音楽特有の事情を取り上げながら解説を行う。音楽に限定しない推薦システム全般の優れた解説としては、神畠の一連の記事<sup>1)~3)</sup>がある。以降、まず、音楽推薦システムを支える主要な技術について解説し、音楽に特有な推薦システムを紹介する。次に、音楽推薦システムの性能を検証するための実験用データの収集・解析方法および評価尺度について説明する。最後に、音楽推薦システムの研究を行っていく上での今後の指針について述べる。

### 主要な音楽推薦技術

音楽推薦の一般的な目的は、あるユーザの過去の行動

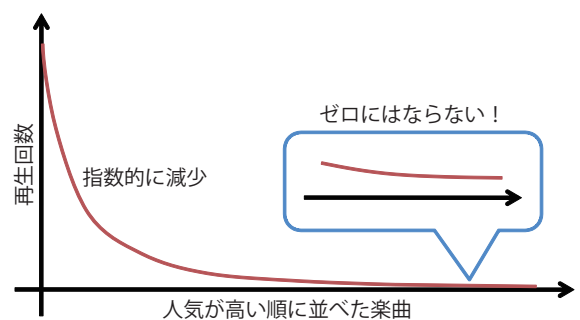


図-1 ロングテール効果

履歴(楽曲の五段階評価や再生回数など)に基づいて、そのユーザの好みに合う楽曲をおすすめ順にいくつか提示することである。推薦時に利用するデータとして、他のユーザの行動履歴を利用するものを「協調フィルタリング」、楽曲の音楽内容を考慮するものを「内容に基づくフィルタリング」、それら両者を考慮するものを「ハイブリッド型フィルタリング」と呼ぶ。

#### ■ 協調フィルタリング

協調フィルタリングは、ほとんどの音楽配信サービスで採用されている最も身近な推薦技術である。基本的には、あるユーザに対し、好みの似た他のユーザが好む楽曲を推薦する。例として、Amazon.co.jpの「この商品を買った人はこんな商品も買っています」などが挙げられる。協調フィルタリングはユーザの行動履歴さえあれば映画でも本でもアイテムの種類を問わずに適用できるため、非常に多くの研究がなされてきた。

このアプローチはユーザの好みに合う楽曲を精度よく推薦できるものの、有益であるとは必ずしも言えない。ある知名度の高いアーティストの楽曲を好むユーザに対しては、同じアーティストの楽曲ばかり推薦されることが多い。したがって、新たな楽曲との出会いを求めるユーザにとって良い推薦とはならない。このように、ユー

☆1 <http://www.lastfm.jp>

☆2 音楽に限らず、店舗面積などの物理的制約を受けないオンライン販売で扱われるさまざまなアイテムに当てはまる。

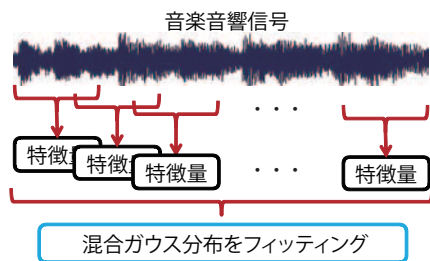


図-2 混合ガウス分布による特徴量のモデル化

ザの行動履歴が十分に蓄積されていない楽曲は推薦されにくいことを**コールドスタート問題**と呼ぶ。

### ■ 内容に基づくフィルタリング

上記問題を克服するため、音楽情報処理分野では、内容に基づくフィルタリングが盛んに研究されている。その多くは、あるユーザに対して、そのユーザが好む楽曲と音楽内容が類似した楽曲を推薦する方式を採用している。しかし、ユーザの音楽的嗜好を形成する要因である文化的背景や市場の人気度を反映できないという問題が指摘されている。

利用する音楽内容は、音楽音響信号に基づくものとメタデータに基づくものとに大別できる。前者は、音楽音響信号を自動解析して、各楽曲の内容を計算機が扱える形式のデータ(**特徴量**と呼ぶ)で表現したものである。後者は、楽曲の音楽要素に対するアノテーションや、不特定多数のユーザが楽曲に自由に付与したソーシャルタグ(ジャンル名や印象語などが多い)などである。音楽配信サービス Pandora<sup>☆3</sup>では、専門家が数百種類の属性をアノテーションしたデータを利用している。

内容に基づくフィルタリングでは、ある尺度に基づいて全楽曲間の特徴量の類似度を計算することが基本となる。音楽音響信号の解析は、短時間フレームを始端から終端まで少しずつシフトさせながら行い、各楽曲につき特徴量ベクトルをフレーム数個得るのが一般的である。その後、全フレームにわたる平均と分散を求めれば(ガウス分布で近似する)、楽曲を1つの特徴量ベクトルで表現できる。類似度の尺度としては、コサイン距離やピアソンの相関係数がよく利用される。また、図-2に示すように、全フレームの特徴量ベクトル群に対して、混合ガウス分布(GMM)をフィッティングさせることで、特徴量分布をより精密に表現する方法もある。GMM間の距離尺度としては、Earth Movers Distance(EMD)がよく利用される。これは、あるGMMの分布をもう片方のGMMに変形させるためのコストを、最小輸送量

問題として定式化したものである。このように、フレーム順序を考慮しない内容表現を**Bag-of-Features**モデルと呼ぶ。これは、文書の一般的な内容表現である**Bag-of-Words**モデルとのアナロジーである。

### ■ ハイブリッド型フィルタリング

前述した2つのアプローチの欠点を解消するため、近年はハイブリッド型フィルタリングに注目が集まっている。Yoshiiらは、Probabilistic Latent Semantic Indexing(pLSI)に基づくハイブリッド型推薦手法を提案している<sup>4)</sup>。pLSIは元々自然言語処理分野で、文書のトピックを潜在変数として表現し、文書の生成過程を確率的にモデル化するために考案された。具体的には、ある文書における複数のトピックの混合比と、文書中の各単語がどのトピックから生成されたのかを推定する。音楽推薦の場合は、文書をユーザ、単語を楽曲とみなすことで、トピックを介したユーザと楽曲の関連の強さを推定でき、さらに特徴量との関連を考慮できるように拡張できる。ほかにも、Tiemannらが、協調フィルタリングの結果と内容に基づくフィルタリングの結果とをアンサンブル学習法を用いて統合する手法を提案している<sup>5)</sup>。

## 音楽に特有な推薦システム

ここでは、音楽を対象とするが故の特徴的な推薦形態として、プレイリストの自動生成と状況・目標に基づく音楽推薦について解説する。

### ■ プレイリストの自動生成

プレイリストとは、何らかの意図をもって楽曲を並べたものであり、順番に再生して鑑賞することが想定されている。多くのユーザが類似した曲調の楽曲を集めたもの、楽曲間のスムーズな曲調の遷移を考慮したものなどさまざまなプレイリストを作成している。音楽に限定しない推薦システムの研究では、この種のトピックは扱われていない。本や映画などのリストを作成しても、1つ1つの鑑賞時間が非常に長くなりがちで、全体を通して何度も鑑賞することは現実的ではないからである。一方、音楽はアルバムという形態が存在することから分かるように、楽曲だけでなくその順序が大きな意味を持つ。しかし、通常の音楽推薦システムはおすすめ順に楽曲をランキングするだけなので、必ずしも適切なプレイリストが得られるとは限らない。

プレイリストの自動生成手法の多くは、類似楽曲検索に基づくものである。一般的なシステム利用法として、ユーザはまず「種」となる楽曲(**seed song**)を指定する。次にシステムは、データベース中から種となる楽曲

☆3 <http://www.pandora.com/mgp.shtml>

と音楽内容が類似した楽曲をいくつか選択し、全体として曲調がそろったプレイリストを生成する。

一方、楽曲の順序を明示的に考慮した手法も提案されている。Aucouturierらは、プレイリスト作成問題をユーザが指定した条件を満たす制約充足問題として定式化している<sup>6)</sup>。ユーザが指定する条件には、1つの楽曲に対するもの(例:あるタイトルの楽曲を含んでほしい)、2つの楽曲の関係に関するもの(例:これらの曲は連続してほしい)、プレイリスト全体に関するもの(例:テンポが徐々に速くなってほしい)などがある。このような複雑な条件のもとで最適なプレイリストを求める問題はNP困難であるが、ローカルサーチを用いると効率的に局所最適解が求められる。Flexerらは、プレイリストの開始曲と終端曲を指定すると、開始から終端まで曲調がスムーズに遷移するように途中の楽曲を自動選択する手法を提案している<sup>7)</sup>。Baccigalupoらは、既存のプレイリストの中から種となる楽曲と関連性が高いプレイリストを選択し、それらを統合することで最終的なプレイリストを生成する手法を提案している<sup>8)</sup>。Pohleらは、**図-3**に示すように、音響的特徴量の類似度から楽曲間の距離を計算し、NP困難である巡回セールスマン問題を近似的に解くことで楽曲順を決定している<sup>9)</sup>。

### ■ 状況や目標に基づく音楽推薦

単純にユーザが好む楽曲というより、ユーザを取り巻く状況によく合致する楽曲を提示する音楽推薦システムも存在する(context-aware recommendation)。小田川らは、ユーザが聴きたい楽曲の印象語(明るい・かなしいなど)を入力することで、そのときの気分にあった楽曲を推薦する技術を提案している<sup>10)</sup>。推薦された楽曲をユーザが最後まで聴いた場合は推薦が適切であった、途中でユーザがしっくりこないと入力した場合は推薦が不適切であったとしてユーザの好みを強化学習することで、推薦精度が向上する。この技術を応用し、時間帯・天気・走行状況に応じて最適な選曲が行えるカーナビがパイオニアから2004年に発売されている。

また、音楽の心理的効果を利用して、ユーザの目標とする活動を支援するための音楽推薦システムも提案されている(purpose-aware recommendation)。Oliverらは、ユーザのエクササイズをサポートする音楽推薦システムを携帯端末上に実装した<sup>11)</sup>。ジョギング中は携帯端末がユーザの心拍数を常時監視しており、目標となる心拍数との差を補償するように楽曲を推薦する。目標より負荷が低すぎるとアップテンポの楽曲を、負荷が高すぎるとスローテンポの曲を推薦し、適切な負荷が維持できるようにユーザをサポートする。

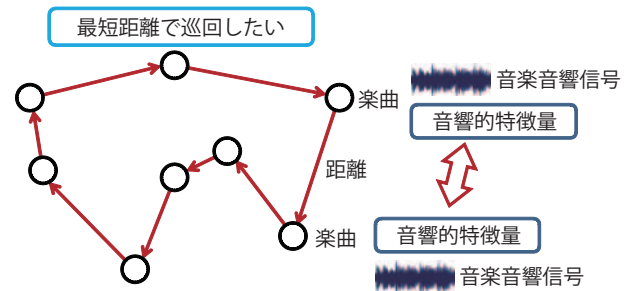


図-3 プレイリスト生成のための巡回セールスマン問題

## 音楽推薦システムの検証

音楽推薦システムを開発する際には、実際に動作させてみてその性能を検証することは重要である。ここでは、推薦に利用するデータの収集・解析方法と音楽推薦システムの評価尺度について述べる。

### ■ ユーザの行動履歴の収集

ユーザの行動履歴には明示的なものと暗黙的なものの2種類が存在する。前者は、Amazon.co.jpの五段階評価で見られるように、ユーザが楽曲評価を直接表明するタイプであり、「好き」だけでなく「嫌い」という評価を入力できることが多い。後者は、iTunesで見られるように、ユーザの購入履歴や再生回数など、自動的に収集されるタイプであり、直接ユーザの評価を反映しているわけではない。たとえば、試しに買ったアイテムが好みではないことはしばしばある。一般に、暗黙的な評価の方が(サービスを提供している企業は)大量に入手できる利点がある。推薦アルゴリズムを設計する際には、このような性質を意識することが重要である。

ユーザの行動履歴を入手する方法の1つは、Web APIを利用することである。たとえば、Amazon Web Services<sup>☆4</sup>を用いれば、タイトル、アーティスト名、レーベル、発売日などの書誌情報はもちろん、カテゴリやカスタマーレビューなどを取得できる。レビューには、ユーザの五段階評価のスコアとコメントが含まれる。すべてのユーザに固有のIDが割り振られているので、どのユーザがどの楽曲に評価を行ったのかを識別でき、明示型の行動履歴データとして利用できる。また、Last.fmが公開しているAPI<sup>☆5</sup>では、書誌情報だけでなく、それらにユーザが付与したタグや人気度、類似したアイテムなどを取得することができる。ユーザがこの1週間でのどのような楽曲を何回再生しているかの情報も取得可能

☆4 Amazon.co.jp以外の企業から提供されている情報(商品の説明など)はAPIでは入手できないので、直接HTMLの解析を行う必要がある。

☆5 <http://www.lastfm.jp/api/>

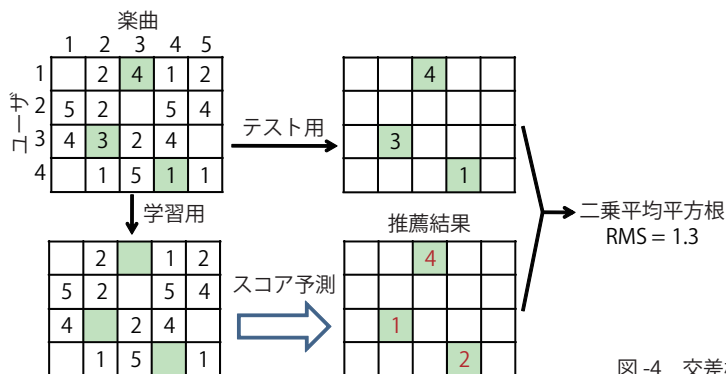


図-4 交差検証法に基づく推薦精度の算出

であり、暗黙型の行動履歴データとして利用できる。

### ■ 音楽音響信号の解析

音楽音響信号の解析結果を推薦に利用したい場合は、Web上で公開されているソフトウェアの利用を検討するとよい。最も有名なものに Marsyas<sup>☆6</sup>があり、信号波形のゼロ交差率・スペクトル特徴量（重心・フラックス・ロールオフ）・メル周波数ケプストラム係数などを抽出できる。これらは楽曲の自動ジャンル分類をはじめ、さまざまな応用に利用される汎用性の高い特徴量である。また、CLAM<sup>☆7</sup>は、特徴量の抽出を行うだけでなく、楽曲にアノテーションを行ったり修正したりできる環境を提供する。ほかに、商用の数値計算システム MATLAB もよく利用される。MATLAB上で動作する MIRtoolbox<sup>☆8</sup>では、音楽情報処理分野でよく利用される信号処理アルゴリズムが数多く実装されている。

### ■ 音楽推薦システムの評価尺度

推薦システムの最も基本的な評価軸は推薦精度であり、交差検証法（cross validation）で測定するのが一般的である。具体的には、行動履歴データを学習用データとテスト用データに分け、学習用データに基づいて推薦を行った結果をテスト用データと比較する。明示型の行動履歴データを用いてユーザーが評価を行っていないアイテムのスコアを予測する場合は、図-4に示すように、実際のスコアと予測スコアとの二乗平均平方根（RMS: Root Mean Square）に基づく評価尺度が利用できる。一方、アイテムを直接ランキングする場合は、スピアマンの順位相関係数や期待値効用<sup>12)</sup>などに基づく評価尺度が利用できる。

推薦精度の評価尺度として、適合率と再現率もよく用いられる。交差検証法において、各ユーザーにN個のアイテムを推薦する場合を考える。推薦されたN個の中

で、テスト用データ中でユーザーが好むと評価していたものがK個含まれていたとすると、適合率はK/Nとなる。また、テストデータ中でユーザーが好むと評価していたアイテムが全部でM個であったとすると、再現率はK/Mとなる。これは、検索システムの評価尺度に由来している。一般に検索システムにおいては、クエリに対する検索結果として「返されるべき」アイテムが想定されている。そのため、検索結果の評価として、適合率だけでなく再現率も考慮する（返されるべきアイテムを上位に多く検索できるほど高精度と判定する）のは理にかなっている。一方、推薦システムにおいては好みに合うアイテムを「発見する」ことが重視され、「網羅的に探す」ことはあまり重要ではないので、適合率に比重が置かれる。

ここで、交差検証法では推薦システムの真の性能を測定しているわけではないことに注意すべきである。二乗平均平方根に基づく評価尺度では、誤差が計算できるのは実際にユーザーが評価を行ったアイテムに限定され、真に未評価のアイテムに対するスコア予測性能は測れない。ランキングや適合率・再現率に基づく評価尺度では、テスト用データ中でユーザーが好きであると評価したアイテムを上位に多く推薦するほうが、より高い推薦精度であると判定されやすい。しかし、そのような既知のアイテムの推薦は有益とは言えない。

近年は、推薦精度の向上だけを目標にするのではなく、推薦精度以外の評価軸を重視する研究が増加しつつある。その背景として、数値的な推薦精度の改善がほとんど頭打ちになってきたことや、推薦精度がある一定以上に達すると、システムを利用するユーザーはほとんどその違いに気づかなくなるという報告がある。推薦システムに求められる特性の1つに、思いがけない発見を可能にする能力（serendipity）がある。音楽推薦の場合は特にこの性質が重要であるので、推薦される楽曲がどの程度多様なアーティストを含んでいるかの評価尺度を提案し、推薦精度との両立を目指している研究もある<sup>4)</sup>。また、現実的なデータサイズ(数百万曲・数百万人オーダ)に対応できるようにするため、推薦にかかる計算コストを評価

☆6 <http://marsyas.sness.net/>

☆7 <http://clam-project.org/index.html>

☆8 <http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox/>

尺度とすることも重要である。さらに、悪意を持つユーザやボットが一部のアイテムを異常に貶めたり、賞賛したりするなどの攻撃をしかけた場合に、推薦結果を乱さないような頑健性も必要である。

## 音楽推薦システムの今後

最後に、音楽推薦システム研究で今後取り組むべき課題を述べ、研究の指針としたい。

### ■ 推薦システム分野との技術統合

音楽情報処理分野では、ハイブリッド型フィルタリングなどで機械学習に基づく音楽推薦手法がいくつか提案されているものの、導入はあまり進んでいない。一方、音楽に限定しない汎用的な推薦システム分野では、機械学習手法を推薦に応用するという観点で多くの研究がなされている。この理由は、音楽推薦が音楽検索の延長であるとの認識が一般的であるため、と考えられる。2000年に初開催された音楽情報処理分野におけるトップレベルの国際会議である ISMIR<sup>☆9</sup>は、当初は音楽情報検索 (MIR: Music Information Retrieval) をメイントピックとして扱っていたが、その後発表されるトピックは多様化してきた。この流れを受けて、内容に基づく検索技術を応用して内容に基づくフィルタリングが盛んに研究されている。今後、推薦システム分野と音楽情報処理分野の技術統合が進み、より洗練されたシステムが開発されることを期待している。

### ■ 音楽内容表現の洗練化

音楽推薦・検索システムでは、スペクトルレベルの低次特徴量が利用されることがほとんどである。この種のアプローチは曲調の類似した楽曲を推薦・検索する目的では有用であるが、リズム構造やコード進行などの高次の音楽内容を考慮することはできない。今後は音楽音響信号に対して高次の音楽内容を自動的に認識する技術だけでなく、それらを効果的に特徴量化する技術についての研究が必要である。また、Bag-of-Features モデルとは異なり、時系列の長いスパンの動的変化をとらえられるような特徴量を設計することが、より本質的に音楽内容を表現することにつながると思われる。

<sup>☆9</sup> <http://www.ismir.net>

### ■ 生きたデータベースの取り扱い

音楽推薦システムを実用化するためには、巨大な楽曲データベースを対象とし、データ変化 (新曲のリリースやユーザの入れ替わり) に対して迅速に適応できなければならない。多くの場合、システムの開発や検証は、静的な小規模データを対象とした実験室環境で行われており、時間的・計算量的コストに制約はない。そのため、実用化を困難にしている。今後は、推薦システムを設計する段階から実際のサービス運用時に必要な要件を列挙し、それを満たすように技術開発していくことが重要である。

#### 参考文献

- 1) 神島敏弘: 推薦システムのアルゴリズム (1), 人工知能学会誌, Vol.22, No.6, pp.826-837 (2007).
- 2) 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).
- 3) 神島敏弘: 推薦システムのアルゴリズム (3), 人工知能学会誌, Vol.23, No.2, pp.248-263 (2008).
- 4) Yoshii, K., Goto, M., Komatani, K., Ogata, T. and Okuno, H. G.: An Efficient Hybrid Music Recommender System Using an Incrementally Trainable Probabilistic Generative Model, IEEE Transactions on Audio, Speech, and Language Processing, Vol.16, No.2, pp.435-447 (2008).
- 5) Tiemann, M., Pauws, S. and Vignoli, F.: Ensemble Learning for Hybrid Music Recommendation, Proc. of ISMIR, pp.179-180 (2007).
- 6) Aucouturier, J.-J. and Pachet, F.: Scaling Up Music Playlist Generation, Proc. of ICME, pp.105-108 (2002).
- 7) Flexer, A., Schnitzer, D., Gasser, M. and Widmer, G.: Playlist Generation Using Start and End Songs, Proc. of ISMIR, pp.173-178 (2005).
- 8) Baccigalupo, C. and Plaza, E.: Case-Based Sequential Ordering of Songs for Playlist Recommendation, Lecture Notes in Computer Science, Springer, Vol.4106, pp.286-300 (2006).
- 9) Pohle, T., Pampalk, E. and Widmer, G.: Generating Similarity-based Playlists Using Traveling Salesman Algorithms, Proc. of DAFx, pp.220-225 (2005).
- 10) 小田川智, 児玉泰輝, 菟山真一, 鈴木康悟, 松下文雄, 塩田岳彦: 楽曲レコメンドシステム, 電子情報通信学会技術研究報告, 応用音響, Vol.106, No.160, pp.49-53 (2006).
- 11) Oliver, N. and Kreger-Stickles, L.: PAPA: Physiology and Purpose-Aware Automatic Playlist Generation, Proc. of ISMIR, pp.250-253 (2006).
- 12) Breese, J., Heckerman, D. and Kadie, C.: Empirical Analysis of Predictive Algorithms for Collaborative Filtering, Proc. of UAI, pp.43-52 (1998).

(平成 21 年 7 月 3 日受付)

#### 吉井 和佳 (正会員)

産業技術総合研究所研究員, 2008年京都大学情報学研究科博士後期課程修了, 博士 (情報学), 音楽推薦や音楽と画像・テキストとの対応付けの研究に従事, 電子情報通信学会, IEEE 各会員。

#### 後藤 真孝 (正会員)

産業技術総合研究所情報技術研究部門メディアインタラクション研究グループ長, 1998年早大大学院博士後期課程修了, 博士 (工学), 現在, 統計数理研究所客員教授, 筑波大学大学院准教授 (連携大学院), IPA 未踏ユース PM を兼任。