

# ユーザ状況と音楽特徴量に基づく音楽推薦の一手法 MALL

宇野愛 伊藤貴之

お茶の水女子大学大学院

## MALL: A music recommendation technique based on user situations and musical features

Ai Uno Takayuki Itoh

Graduate School of Humanities and Sciences, Ochanomizu University

{ai\_u, itot} @ itolab.is.ocha.ac.jp

### 概要

聞きたい曲を選ぶ際に、メタデータから選ぶだけでなく、状況や環境（例えば、天気や季節、時間帯、場所など）に合った曲を選びたいときがある。これらの状況や環境の多くはライフログとして記録可能である。一方、我々は日常的に膨大な楽曲を所持していることが多く、状況にあった楽曲をその都度選ぶのは困難である。そこで近年、ライフログから推定される個人の選曲傾向に従って楽曲を自動推薦する手法が多く研究されている。その多くは個人の行動と音楽のメタ情報を結びつけるものである。それに対して本報告では、楽曲再生時のライフログ情報と音楽特徴量を照合し、その相関ルールを導出することで楽曲を推薦する MALL (Music Adviser with Life Log) を提案する。本報告では MALL に関する検証実験結果を報告し、MALL が導出した相関ルールの妥当性について検証する。

## 1 はじめに

ポータブル音楽プレーヤー上での選曲画面には、一般的にメタデータ（タイトル、アーティスト名、ジャンルなど）が表示される。一方で我々が 236 人の大学生にアンケート調査した結果によると、特定の状況や心境には特定の音楽を聴きたいという要求をもつ人は非常に多い。例えば回答者の 71% は「特定の季節には特定の音楽が合う」と感じ、回答者の 62% は「特定の時間帯には特定の音楽が合う」と感じたことがあると回答している。他にも半分以上の回答者が、「特定の場所や天気には特定の音楽が合う」と感じたことがあると回答している。このような状況や心境はライフログとして記録が可能である。そして我々は、ライフログベースの音楽推薦システムがあれば、このような要求を満足させることが可能であると考えられる。このような音楽推薦システムとして発表されている研究の多くは、ユーザの行動と楽曲のメタ情報に関する確率を導出している。一方で本手法の開発

に先立って我々は、「静かな場所では静かな楽曲を推薦する」「スポーツの前には元気な曲を推薦する」といった直感的な楽曲推薦こそ、ライフログとの相関ルール導出において重要であると考えられる。また、特定の音楽ジャンルにこだわることなく一定の特徴を有する楽曲を推薦することが、推薦結果における意外性、またユーザにとっての新たな音楽との出会いにつながる可能性を高めると考える。以上のことから楽曲のメタ情報に加えて、楽曲自体が有する特徴量をユーザ行動に紐付ける形での楽曲推薦が有効であると考えられる。

本報告では音楽リスナーの状況と音楽特徴量の関係に基づく音楽推薦の一手法 MALL (Music Adviser with Life Log) を提案する。本手法では音楽リスナーの状況（例として時間帯、季節、天気、イベントなど）を記録したライフログと、そのときリスナーが「状況に合っている」と判断した楽曲の音楽特徴量との関係を、数値属性相関ルールによって導き、その結果に沿って楽曲を推

薦する。本手法の確立のために我々は、Android 上のアプリとして MALL の専用音楽プレイヤーを開発している。このアプリには、リスナーが「状況に合っている」と判断した時に押すボタンを備えている。そしてリスナーがこのボタンを押すと、アプリはその時点での状況と楽曲名をログファイルに記録する。MALL はこのログファイルを解析し、数値属性相関ルールを定期的に導出することで、音楽推薦のためのルールを構築する。

本報告では MALL の処理手順に加えて、1ヶ月間のユーザ実験結果を紹介することで、MALL の有効性について検証する。なお本報告の骨子は既に我々自身による国際会議論文でも発表されている [12] が、本報告ではその内容に対して、被験者実験結果を追加し、MALL の設計方針や実験結果に関する議論を深めている。

## 2 関連研究

他の情報推薦と同様に、音楽推薦の研究も黎明期には協調フィルタリング [1, 9] やメタデータ解析 [7] などに基づく手法が発表されてきたが、近年になって提案手法にも関連する音楽特徴量ベースの手法 [8] やユーザの状態に基づく手法 [4] が数多く発表されている。特にユーザのコンテキストに基づいた音楽推薦手法 [2, 5, 6, 11, 13] は近年さかんに発表されているが、その大半はコンテキストとメタデータの共起性に注目している。MALL は前章でも述べた通り、メタデータではなく音楽特徴量に着目することで「静かな曲」「元気な曲」といった抽象的かつ直感的な楽曲選択を目指しており、その実現のために数値属性相関ルールを導入している。

## 3 提案手法 MALL の処理手順

図 1 は提案手法 MALL のユースケースシナリオを示したものである。本研究で開発するポータブル音楽プレイヤー上のアプリが、リスナーの聞いた曲とその時の状況をログファイルに記録する。MALL ではそれらの音楽特徴量とユーザ状況に関する相関ルールを導出し、そのルールに該当する楽曲群を集めてプレイリストを生成する。このプレイリストを提供することで、個々のリスナーが各々の状況に合うと感じる楽曲群を推薦することができる。

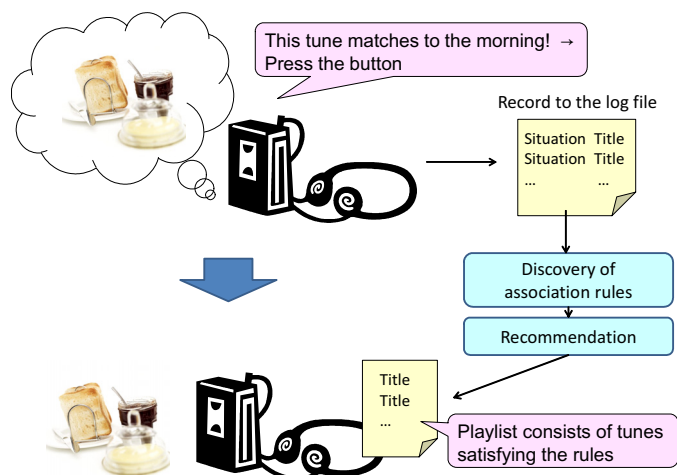


図 1: MALL のユースケースシナリオ。(上) ライフログの記録。(下) 音楽推薦結果としてのプレイリストの提供。

### 3.1 音楽特徴量

現時点での我々の実装では、MIRtoolbox<sup>1</sup> という音楽情報抽出ソフトウェアを用いて、以下の音楽特徴量を各楽曲に対して算出している。

RMS energy は音響エネルギーの二乗平均平方根を示すもので、J-POP やロックなどでは高く、アコースティックな曲やクラシック音楽などでは低い傾向にある。

Tempo は音響振幅のピークや和音の変更頻度などから算出可能である。

Brightness は 1500Hz 以上 (主に倍音) の周波数領域が占めるエネルギー比率である。倍音の豊かな楽器 (バイオリン等の擦弦楽器、金属製の打楽器など) を含む編成であるか否か、などの観点から楽曲を特徴づけられる。

Roughness は不協和音が占めるエネルギー比率である。伝統的な音楽やポップスなどでは不協和音比が小さく、モダンな音楽では不協和音比が高い傾向にある。

Mode は長和音と短和音が占める時間比率である。

### 3.2 ライフログの記録

我々は MALL 専用のポータブル音楽プレイヤーを Android アプリとして開発した。このアプリには、リスナーが「この曲は現在の状況や環境に合っている」と感じたら押すボタンを備えている。図 2 がそのスナップショットである。下部に 4 つ並んでいるボタンの一番右が、状

<sup>1</sup><http://www.jyu.fi/hum/laitokset/musiikki/en/research/coe/materials/mirtoolbox>

況や環境に合っていると感じたら押すボタンである。このボタンを押すとアプリは、再生中の楽曲のタイトルと、その時の状況（現在の実装では天気、曜日、時間帯）をログファイルに記録する。

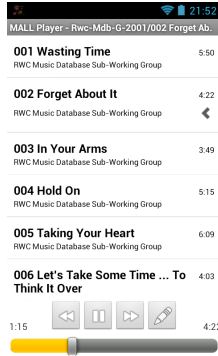


図 2: MALL 専用の音楽プレイヤーのスナップショット。

### 3.3 数値属性相関ルールマイニング

本報告ではログファイルに記録された「ライフログ条件」を  $A$ 、音楽特徴量の数値範囲を  $B$  とし、相関ルール ( $A \rightarrow B$ ) を求めることによって、楽曲推薦の条件を特定する。相関ルールマイニングは一般的に、以下の「支持度」と「確信度」の両者を同時に満たすルールを発見する問題に相当する。

支持度  $P(A, B)$  は  $A$  と  $B$  が同時に成立する確率を示す。

確信度  $P(B|A)$  は  $A$  が成立している条件下において  $B$  も成立する確率を示す。

現時点での我々の実装では、天気、季節、曜日、時間帯に関するライフログ条件をあらかじめ何種類か設定してある。そしてその各々に対して、音楽特徴量の数値範囲の論理積として表現される音楽特徴量条件とのルールを設定する。そして各々のルールに対して支持度と確信度を算出し、その両者があらかじめ指定された閾値を上回るルールを抽出することで、ライフログ条件と音楽特徴量の相関ルールとする。以上を定式化すると以下のようになる。

$$A \rightarrow B_1[a_{1min}, a_{1max}] \& \dots B_n[a_{nmin}, a_{nmax}] \quad (1)$$

ここで  $A$  はライフログ条件であり、 $B_i[a_{imin}, a_{imax}]$  は  $i$  番目の音楽特徴量の数値範囲である。上述のようなルールは一般的に数値属性相関ルールと呼ばれており、多くのマイニング手法 [3, 10] が既に知られている。

### 3.4 音楽推薦

MALL は前節の方法で抽出されたルールに合致する楽曲をリスナーに推薦する。ここで非常に多くの楽曲が用意されているとき、推薦される楽曲数も膨大になる可能性がある。現時点での我々の実装では、個々の楽曲が有する  $n$  次元の音楽特徴量ベクタ ( $a_1, \dots, a_n$ ) と、相関ルールによって抽出された数値範囲の中心点との距離を算出し、その距離が小さい順に音楽推薦の優先順を定義している。また現時点での我々の実装では、M3U ファイル形式で楽曲群を出力することで、Android アプリのプレイリストとして楽曲を推薦する。

## 4 実行例

我々は Android 音楽プレイヤーを被験者（20 代の女子大生）にもたせ、収録曲を聴いてその時の状況に合っていると感じたら所定のボタンを押してもらい、という実験を行った。本実験では RWC Music Database<sup>2</sup> に収録されている 313 曲を用いた。ただし本実験では収録曲数が限られている関係で、3 個以上の音楽特徴量の論理積として相関ルールを抽出するとそれに該当する楽曲数が少なくなりすぎる、という問題が発生した。そのため本実験では、相関ルールは任意の 2 個の音楽特徴量の論理積として抽出した。

結果として本実験では、被験者のうち 3 人に対して相関ルールを導出できた。被験者 A には「夜」「水曜日」「晴天」の 3 つのライフログ条件に対して、被験者 B には「朝」「夜」「晴天」「曇天」の 4 つのライフログ条件に対して、被験者 C には「朝」「昼」「夕方」「夜」「晴天」「月曜日」「金曜日」「土曜日」の 8 つのライフログ条件に対して、それぞれ相関ルールを導出できた。表 1 および表 2 に、被験者 A, B に対して導出された相関ルールの一覧を示す。被験者 C の相関ルールは大量に導出されたため割愛する。

これらの相関ルールに沿って推薦された楽曲を 3 人の被験者に聴いてもらい、その 5 段階評価と感想を尋ねた。被験者 A の回答は以下のとおりである。

- 夜：評価 5 静かな曲が多く、夜には非常に合うと思った。
- 水曜：評価 4 確かに自分の好きな曲が多かった。しかし、なぜそれが水曜なのかは自分にも説明できない。

<sup>2</sup><http://staff.aist.go.jp/m.goto/RWC-MDB/>

- 晴天：評価 4 元気が出る曲が多く、晴天に合うと思った。

被験者 B の回答は以下のとおりである。

- 朝：評価 5 元気が出る曲が多く、朝には非常に合うと思った。
- 夜：評価 4 スローテンポまたは暗めの曲が多く、夜に合うと思った。
- 晴天：評価 3 元気が出る曲もあれば、心地よい曲もあり、暗い曲もあり、選曲に一貫性がなかった。
- 曇天：評価 4 穏やかな曲が多く、理解できる選曲であった。

被験者 C の回答は以下のとおりである。

- 朝：評価 4 朝に合う元気が出るような明るい曲、またはゆっくりと静かで落ち着いた曲が多かった。
- 昼：評価 4 昼に合う元気が出るような明るい曲が多かった。
- 夕方：評価 5 夕方に合う静かで悲しい曲、落ち着いたイメージの曲が多かった。
- 夜：評価 5 夜に合う静かで落ち着いた曲や、音の小さく起伏の少ない曲が多かった。
- 晴天：評価 4 晴れに合う長調の明るい元気が出るような曲が多かった。
- 月曜日：評価 4 週の初めに合う落ち着いた曲や長調で元気が出るような曲が多かった。
- 金曜日：評価 5 平日の最後に合う静かで落ち着いた曲、悲壮感漂うが切なさが顕著ではない曲が多かった。
- 土曜日：評価 5 土曜に合う長調でのんびりとした、もしくは明るく元気が出るような曲が多かった。

この結果から、MALL による楽曲推薦結果は概ね好評と考えられる。特に被験者 3 名の回答から「元気な曲」「落ち着いた曲」といった抽象的な印象を意味する単語が頻出された点に注目されたい。このような抽象的な印象によって紐付けられた楽曲を特定するためには、アーティスト名や音楽ジャンル名などのメタ情報よりも、音楽特徴量を手がかりにするほうが有利である。例えば被験者 B の「朝」に対する回答に「元気な曲」とい

う単語があるが、実際に表 2 に示した被験者 B の「朝」に関する相関ルールには、RMSEnergy が高いポップな曲や、Brightness が高い華やかなサウンドの曲が導出されている。このことから、相関ルールの導出に音楽特徴量を用いることの妥当性が示唆される。

一方で、被験者 B の「晴天」に対する推薦結果だけはあまり良好でなかった。ここで被験者 B の「晴天」に対して MALL が導出した相関ルールは「RMSEnergy が低く Roughness が低い」「Roughness が中程度で Brightness が低い」の 2 つであった。前者のルールからはアコースティックでかつ不協和音が少ない音楽が推薦される可能性が高く、後者のルールからは不協和音が中程度で倍音が少ない編成の音楽が推薦される可能性が高い。このような一貫性のなさが評価の低さにつながっていることが示唆されるため、今後の課題として同一のライフログ条件から複数の相関ルールを導出した際の組み合わせについて議論したい。

被験者には以下の質問にも回答してもらった。

Q1: 天気、曜日、時間帯の他にどのようなライフログ条件があるとよいか。

全ての被験者が「気分・感情」をあげた。1 人の被験者からは「場所」という意見もあがった。

Q2: Android アプリにどんな機能を追加してほしいか。

再生曲のブックマーク、ボタン以外のインタフェースによる「状況に合う曲」の記録、といった意見があがった。

## 5 まとめ・今後の課題

本報告では、ライフログ条件と音楽特徴量の相関ルール導出によって楽曲を推薦する一手法 MALL を提案し、その実験結果を示した。MALL によって推薦された楽曲を被験者に評価させたところ、「元気な曲が多い」「落ち着いた曲が多い」といった抽象的な印象をもって推薦結果を高く評価したコメントが多く得られた。このことから、音楽特徴量に関する相関ルールが楽曲推薦において一定の妥当性を有することが示唆される。一方で、今後の課題として以下の点が考えられる。

まず 3.1 節に示した以外の音楽特徴量を試してみたい。現時点での実装では、我々自身の先行研究での実験結果にもとづいて音楽特徴量を選択しただけであり、本報告で示した音楽特徴量が MALL の目的に合致しているかは検証されていない。本来ならもっと多くの音楽特徴量に対して相関ルールを導出する実験を試すか、あ

るいは深層学習などを利用して各音楽特徴量の妥当性を導くなどの形での検証が望ましいであろうと考える。また現時点での実装では、特にメロディやリズムパターンに関連する音楽特徴を十分に考慮できていない。近年では音源分離技術も発達しており、メロディやリズムパターンだけを単体で分析することも徐々に可能になっているので、これらの技術と組み合わせることで音楽特徴量の充実を図りたい。

関連ルールマイニングにおいて、支持度と確信度の閾値自動設定は重要な問題である。我々の経験上、支持度と確信度を全てのユーザ・全てのライフログ条件において同一値にしておくことが適切な楽曲推薦を導くとは限らないようである。これらの値を適切に自動設定することは、実用上の観点から不可欠であると考えている。

楽曲の特徴が1曲の中で大きく変化するとき、どの時点の特徴を採用するか、という問題もある。ポップス曲の場合にはサビ検出手法などを導入することで一貫性のある選曲が可能になるが、クラシック音楽などのように多彩な展開を有する楽曲の場合には簡単ではない。また同様な問題は例えば、クラシック音楽の組曲のように複数の楽曲をあわせて1曲と扱うような楽曲の推薦や、また1曲単位ではなく1アルバム単位、1プレイリスト単位での楽曲の推薦、などを視野に入れた時にも議論が必要となる。

「状況に合う曲」の意思表示入力をボタン以外の方法で、という意見もある。我々の過去の議論の中でも例えば、「途中で早送りして最後まで聴かなかった曲が状況に合わない曲で、それ以外は全て状況に合う曲」というインターフェースにしたほうが適切に楽曲を推薦できるのでは、という意見が交わされたことがある。しかし我々はこのような考え方には悲観的である。我々はMALLにおける関連ルールマイニングにおいて、リスナーが能動的に「状況に合う」と意思表示した楽曲のみから関連ルールを導出することが、その満足度において非常に重要であると考えている。一方で、ボタンを押すという動作だけのためにAndroidアプリを開くという操作は煩雑である。対策としてインターフェース面での工夫、例えばケーブル中にボタンを搭載したヘッドフォンを採用する、などを検討したい。

ライフログ条件の追加実装については既に多くの人から意見をいただいている。その中でも特に、以下によるライフログ条件の実装が有効であると考えている。

場所: ポータブル機器中のGPSによって取得する緯度・経度の情報から「海」「山」といった大きな地理的意味

を抽出可能である。またAndroid機器が接続したWi-FiのIPアドレスから、その建物で経営される業種などを抽出可能である。

行動: ポータブル機器中の加速度センサで記録した数値から「走行中」「起立中」といった情報を抽出可能である。

テキスト情報: メールやSNSなどの記述内容に対して自然言語処理技術を適用することにより、その日のイベントや、気分・感情などを抽出可能である。

以上について議論した後、さらに多くの楽曲、さらに多くの被験者を交えてユーザ評価を実施したい。今回の実験でも実際には4人以上の被験者に協力をお願いしていたが、Android機器の故障などでログファイルが壊れるなどのトラブルがあり、結果的には3人の被験者からのみ関連ルールを抽出できたという経緯があったので、その反省も踏まえて再実験を実施したい。さらに、ランダムな選曲結果との比較実験、音楽特徴量以外の情報(メタデータなど)に基づく音楽推薦結果との比較実験、などについて検討したい。

## 参考文献

- [1] M. Anderson, M. Ball, H. Boley, S. Greene, N. Howse, D. Lemire, S. McGrath, RACOFI: A Rule-Aplying Collaborative Filtering System, *Workshop on Collaboration Agents: Autonomous Agents for Collaborative Environments*, 2003.
- [2] L. Baltrunas, M. Kaminskas, B. Ludwig, O. Moling, F. Ricci, A. Aydin, K.-H. Luke, R. Schwaiger, InCar-Music: Context-Aware Music Recommendations in a Car, *E-Commerce and Web Technologies*, LNBP 85, 89-100, 2011.
- [3] T. Fukuda, Y. Morimoto, S. Morishita, T. Tokuyama, Mining Optimized Association Rules for Numeric Attributes, *ACM Symposium on Principles of Database Systems*, 182-191, 1996.
- [4] K. Kaji, K. Hirata, K. Nagao, A Music Recommendation System Based on Annotations about Listeners' Preferences and Situations, *International Conference on Automated Production of Cross Media Content for Multi-Channel Distribution*, 231-234, 2005.
- [5] J.-S. Lee, J.-C. Lee, Context Awareness by Case-based Reasoning in a Music Recommendation System, *Ubiquitous Computing System*, 45-58, 2007.
- [6] H.-S. Park, J.-O. Yoo, S.-B. Cho, A Context-Aware Music Recommendation System Using Fuzzy Bayesian Networks with Utility Theory, *Fuzzy Systems and Knowledge Discovery*, 970-979, 2006.

- [7] S. Pauws B. Eggen, Realization and User Evaluation of an Automatic Playlist Generator , *Journal of New Music Research*, 32(2), 179-192(14), 2003.
- [8] Y. Saito, T. Itoh, MusiCube: A Visual Music Recommendation System featuring Interactive Evolutionary Computing, *Visual Information Communication - Information Symposium (VINCI'11)*, 2011.
- [9] U. Shardanand, P. Maes, Social Information Filtering: Algorithms for Automating "Word of Mouth", *ACM SIGCHI Conference on Human Factors in Computing Systems*, 210-217, 1995.
- [10] R. Srikant, R. Agrawal, Mining Quantitative Association Rules in Large Relational Tables, *ACM SIGMOD International Conference on Management of Data*, 1-12, 1996.
- [11] J.-H. Su, H.-H. Yeh, P. S. Yu, V. S. Tseng, Music Recommendation Using Content and Context Information Mining, *IEEE Intelligent Systems*, 25(1), 16-26, 2010.
- [12] A. Uno, T. Itoh, MALL: A Life Log Based Music Recommendation System and Portable Music Player, *ACM Symposium on Applied Computing, Multimedia Visualization Track*, 939-944, 2014.
- [13] X. Wang, D. Rosenblum, Y. Wang, Context-aware Mobile Music Recommendation for Daily Activities, *ACM International Conference on Multimedia*, 99-108, 2012.

表 1: 被験者 A について抽出された関連ルール .

| ライフログ条件 | 音楽特徴量   |
|---------|---|
| 夜       | RMSEnergy [0.04, 0.05]<br>Roughness [5.0, 55.0]     |
| 夜       | RMSEnergy [0.03, 0.06]<br>Rolloff [1000.0, 2000.0]  |
| 夜       | Rolloff [1500.0, 2000.0]<br>Brightness [0.16, 0.21] |
| 夜       | Rolloff [1000.0, 2000.0]<br>Roughness [5.0, 55.0]   |
| 水曜日     | RMSEnergy [0.03, 0.06]<br>Roughness [5.0, 55.0]     |
| 晴天      | RMSEnergy [0.04, 0.07]<br>Roughness [55.0, 105.0]   |

表 2: 被験者 B について抽出された関連ルール .

| ライフログ条件 | 音楽特徴量   |
|---------|---|
| 朝       | RMSEnergy [0.03, 0.05]<br>Roughness [5.0, 55.0]     |
| 朝       | Tempo [100.0, 115.0]<br>Brightness [0.430, 0.480]   |
| 朝       | RMSEnergy [0.100, 0.130]<br>Tempo [105.0, 110.0]    |
| 夜       | RMSEnergy [0.05, 0.07]<br>Rolloff [2000.0, 3000.0]  |
| 夜       | RMSEnergy [0.05, 0.07]<br>[Mode -0.10, -0.06]       |
| 夜       | RMSEnergy [0.03, 0.06]<br>Tempo [115.0, 135.0]      |
| 夜       | RMSEnergy [0.06, 0.08]<br>Roughness [105.0, 155.0]  |
| 夜       | Tempo [115.0, 135.0]<br>Roughness [5.0, 55.0]       |
| 夜       | Rolloff [1000.0, 2000.0]<br>Brightness [0.13, 0.18] |
| 夜       | Rolloff [1500.0, 2500.0]<br>Roughness [5.0, 55.0]   |
| 晴天      | RMSEnergy [0.03, 0.05]<br>Roughness [5.0, 55.0]     |
| 晴天      | Rolloff [1500.0, 2000.0]<br>Brightness [0.16, 0.21] |
| 曇天      | RMSEnergy [0.08, 0.11]<br>Roughness [155.0, 255.0]  |
| 曇天      | RMSEnergy [0.10, 0.13]<br>Tempo [105.0, 115.0]      |