

ユーザの行動に基づいた楽曲推薦システムの提案

東海菜摘^{†1} 湯川高志^{†1}

概要: ユーザ個人の位置情報や楽曲再生記録といったライフログを利用し、ユーザ個人の行動に適した楽曲推薦システムの実現を目的とする。ユーザの所在位置の変化や時間帯から行動の種別が推定できるとともに、聴きたい楽曲が行動に影響される点に着目し、スマートフォンのセンサ情報のログからの行動推定機能と、楽曲再生記録に基づき生成する種プレイリストを利用した推薦機能を持った楽曲推薦システムを提案した。提案システムを実装し、実験的に評価して、その有効性を検証した。

キーワード: 楽曲推薦 ライフログ 行動推定 トピックモデル

A Music Recommendation System based on Users' Behavior

NATSUMI TOKAI^{†1} TAKASHI YUKAWA^{†1}

Abstract: The present paper proposes a new music recommendation system which employs a user's lifelog including location information, playback records, and so on. The system estimates classes of the user's behavior using a sensor log in his/her smartphone and captures playback records to construct base play lists. Then, it recommends songs based on the class of behavior associated with the base play list and category of the songs. The system was evaluated experimentally to verify its usefulness.

Keywords: Music Recommendation, Lifelog, Behavior Estimation, Topic Model

1. はじめに

1.1 背景

近年、MP3 プレイヤーやスマートフォンなどの小型端末の普及や保存容量の増大により、音楽をデジタルデータとして保有するユーザが増加している。その楽曲数の中から移動、家での余暇、外出中など行動によって大きく変化するユーザ自身の聴きたい音楽に合わせて選曲することはユーザの負荷が大きいと考えられる。

一方、スマートフォンをはじめとする小型端末は、現在は常時インターネットに接続可能で、高性能なセンサが多く搭載されるようになった。それにより、個人が自身のインターネット内外での日常の活動記録、ライフログを手軽に残すことが可能となった。これらの情報は日々の活動にとともに蓄積されていく。この蓄積された情報を解析することでユーザの嗜好や興味、習慣などを推定することが可能となった。

1.2 目的

MP3 プレイヤーやスマートフォンなどの小型端末の高性能化により、楽曲選択時のユーザの負荷が大きくなる一方でライフログを活用した行動推定が可能となった。そこで、位置情報を利用した行動推定が可能であること、ユーザの聴きたい楽曲が行動に影響を受けることに着目し、ユーザの行動に基づく楽曲推薦システムの実現を目的とする。

2. 背景知識

2.1 楽曲推薦を行う既存サービス

現在様々なサービスに推薦システムが導入されている。音楽特有の推薦システムとしてプレイリストを用いた推薦がある[1]。現在の一般的な音楽プレイヤーでは、アーティストやアルバムなどを選択して複数の曲を順に再生することができる。その中でプレイリストを作る場合、作成者が曲調や歌詞から判断した印象やルールによって作成されると考えられる。既存のサービスでもプレイリストの利用が進んでいる。Google music のインスタントミックスと呼ばれるサービスはその一つである。これはユーザの Google music 内の楽曲に関する購入や視聴などのログに基づきプレイリストが生成される。

2.2 ライフログを用いた楽曲推薦に関する既存研究

ライフログを用いた楽曲推薦に関する研究として宇野ら[2]の研究について述べる。宇野らは、ユーザは聴きたい曲を選ぶ際、曲名や歌手名などのメタデータから選ぶだけでなく、その日の出来事や天気、時間といった状況や環境にあった選曲をしたときがあると考え、これらの状況や環境をライフログとして記録し、同時に楽曲再生ログも記録し、このログにしたがって楽曲を自動推薦する Android 用のポータブル音楽プレイヤーアプリケーションを提案している。

宇野らの提案したアプリケーションでは、音楽を鑑賞している人がその曲が現在の状況に適していると感じた場合

^{†1} 長岡技術科学大学
Nagaoka University of Technology

に、所定のボタンを押し、ライフログ情報と再生楽曲を記録する。記録される情報は、ボタンを押した日付、曜日、時間、曲名、歌手名となる。位置情報や天気などを後から編集で付加することが可能である。このアプリケーションでは、楽曲・ライフログの記録がともに原則としてユーザによる操作を必要とする。

2.3 行動推定に関する既存研究

スマートフォンなど小型端末に搭載された GPS から得られる位置情報を用いた行動推定の研究として宮崎ら[3]の研究について述べる。

宮崎らは次の3種類のデータを入力として利用している。

- スマートフォンから取得するデータ
 ユーザの位置情報やスマートフォンの操作ログ
- ユーザが登録するデータ (ユーザ情報)
 自宅や職場などユーザにとって特別な意味を持つ位置と名称 (有意エリア), 起床や就寝, 就業時間などユーザ自身の生活サイクルに関する時間情報 (生活時間帯情報)
- 行動繊維テーブル (行動モデル)
 学生や主婦, 会社員などのユーザの特性に合わせて, 取りうる一般的な行動の遷移関係のモデル

これらの入力データから推定できる結果を宮崎らは図1のように分類した。横軸は、推定の対象となるデータの量である。蓄積度が最も低い単一データであれば、ある日時とそのときにいた地点の経度緯度情報のセットを指す。縦軸は、推定された結果の抽象度である。抽象度が最も低い行動要素はひとつの行動を構成する各要素を意味する。在宅中に音楽を聴くという行動の行動要素は、自宅にいて、滞在中であること、スマートフォンで音楽再生アプリが起動していることなどが考えられる。

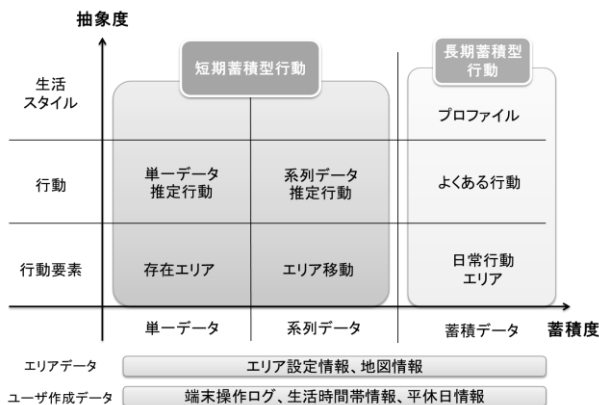


図1 行動分類マトリクス

3. ユーザの行動に基づいた楽曲推薦システムの提案

本章では、2.3 節で述べたようなスマートフォンを用いた行動推定の原理を利用し、ユーザに記録操作の負荷をかけることのない、行動に基づく楽曲推薦システムの提案をする。

3.1 提案システムの概要

提案するシステムの構成図を図2に示す。

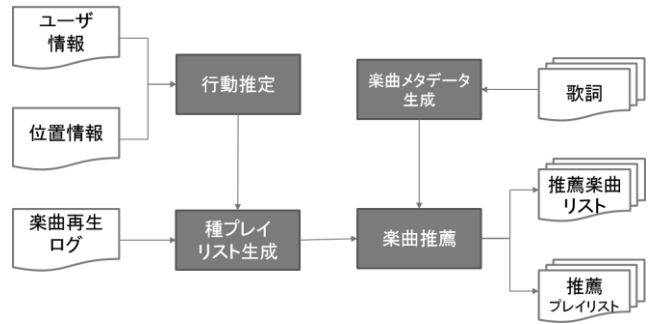


図2 提案するシステムの構成図

各ユーザの以下の情報を用いる。

- 位置情報
- 楽曲再生ログ
- 自宅の住所 (ユーザ情報)

また、音楽再生端末内の全楽曲の歌詞をあらかじめ用意する。本システムは以下のように動作する。

1. 位置情報とユーザ情報を入力とし、行動推定部が行動を推定する。
2. 推定結果と楽曲再生ログから種プレイリスト生成部が行動ごとに種プレイリストを生成する。
3. 楽曲メタデータ生成部が歌詞から楽曲メタデータを生成する。
4. 種プレイリストと楽曲メタデータから楽曲推薦部が推薦曲と推薦プレイリストを出力する。

3.2 位置情報を用いた行動推定手法

まず、位置情報とユーザ情報、行動モデルを用意する。位置情報は Android を用いて記録した Google ロケーション履歴を用いる。ユーザ情報として有意エリアの住所を登録し、住所を経度緯度に変換し、図3のようにその座標を中心とした $120\sqrt{2}$ m 四方を有意エリアとする。行動モデルは図4のようなモデルとする。行動を推定する際に推定結果は、在宅から外出、および、外出から在宅へは直接遷移せず、必ず移動を介して遷移する。

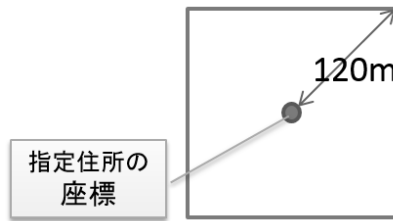


図3 有意エリア



図4 行動モデル

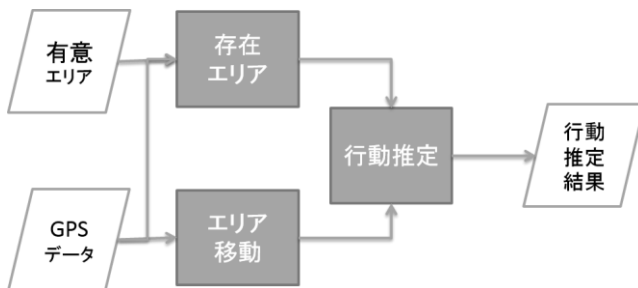


図5 行動推定手法モデル

存在エリア判定では位置情報の各点において各有意エリア内に存在するかを判定する。このとき、前後の存在エリア判定結果は考慮しない。その時点でどの有意エリアにいるのみを判定する。エリア移動判定では位置情報の任意の一点を基準に過去の位置情報の空間的な重複を判定する。5分以上重複する場合は滞在、重複しない場合は移動と判定する。このとき、前後のエリア移動情報は考慮しない。

行動推定は、存在エリアとエリア移動の判定結果を元に行動モデルに基づき、以下の手順で在宅・移動・外出のいずれかの行動を推定する。

- エリア移動情報が移動の場合、
→行動推定結果は移動となる。
- エリア移動情報が滞在の場合、
 - ・ひとつ前の行動推定結果が移動もしくは在宅かつ、存在エリアが自宅の場合、
→行動推定結果は在宅となる。
 - ・ひとつ前の行動推定結果が移動もしくは外出かつ、存在エリアが自宅以外の場合、
→行動推定結果は外出となる。

3.3 種プレイリストの生成

まず、位置情報記の録時と同期間の楽曲再生ログを取得する。今回、楽曲再生ログは Last.fm の Android アプリで記録し、事前処理として記録されたものから歌手名、楽曲名、再生日時を抽出する。これを楽曲再生ログと呼ぶ。

種プレイリストは、3.2 節で述べた行動推定手法で得られる行動推定結果と、楽曲再生ログを参照し、作成する。それぞれの行動の間に再生されている曲の再生回数をカウントし、2 回以上再生された楽曲をリストにする。種プレイリストは行動ごとに生成されるため、在宅、移動、外出の各行動に対する種プレイリストが存在する。

3.4 楽曲メタデータの生成

楽曲メタデータは端末内にある全楽曲の楽曲名、歌手名、作詞家、作曲家、歌詞のトピックについての情報である。これらの情報を得るために、次のような処理を行う。

1. 端末内の全楽曲の歌詞情報を収集する。
2. 楽曲ごとに、歌詞情報の中に含まれる歌手名・作詞家・作曲家の情報と歌詞を分ける。
3. 2 で得た歌詞を形態素解析し、日本語は動詞・名詞・形容詞のみ、英語は頻出単語など指定した単語を除いた単語を抽出する。これを歌詞コーパスと呼ぶ。今回、形態素解析には MeCab を用いた。
4. LDA を用いて、歌詞コーパスから歌詞ごとのトピック分布を得る。今回、LDA には gensim を用いた。
5. 4 で得られた歌詞ごとのトピック分布から歌詞ごとに帰属度の高い3つのトピックを歌詞のトピックとする。

以上が楽曲メタデータを得るための処理となる。

3.5 楽曲の推薦

楽曲推薦は以下の手順で行う。

1. ベクトル生成
 2. 類似度計算
 3. 推薦プレイリスト生成
- この手順を在宅・移動・外出の行動ごとに行う。

(1) ベクトル生成

行動別楽曲情報ベクトルは種プレイリストに含まれる歌手、作詞家、作曲家についてのベクトルである。楽曲メタデータに含まれる歌手、作詞家、作曲家の総数の次元のベクトルで、要素は種プレイリスト内の出現頻度とする。行動別歌詞ベクトルは種プレイリスト内の楽曲の歌詞のトピックについてのベクトルである。トピック次元のベクトルで、要素は種プレイリスト内の出現頻度に帰属度の重みをかけたものとする。行動別推薦対象ベクトルは、楽

曲メタデータにある全楽曲の中で、ある行動の種プレイリストに含まれないすべての楽曲をその行動の推薦対象楽曲とし、楽曲情報ベクトル、歌詞ベクトルと同様のベクトルを曲ごとに求めたものとなる。また、すべてのベクトルは大きさが1となるように正規化する。

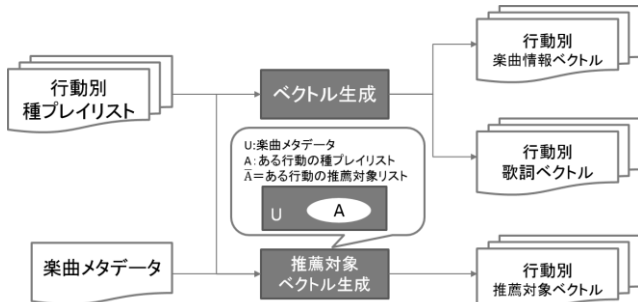


図6 ベクトル生成処理

(2) 類似度計算

推薦対象ベクトル内の各曲のベクトルに対して、行動別楽曲情報ベクトル、行動別歌詞ベクトルについて、それぞれコサイン類似度を計算する。

(3) 推薦プレイリスト生成

類似度の高い順に推薦対象ベクトルをソートする。このとき、歌詞ベクトルの類似度を優先する。そして、類似度の高い10曲を推薦曲とし、この10曲を種プレイリストに追加し、推薦プレイリストとする。推薦曲、推薦プレイリストともに行動ごとに生成される。

4. 提案システムの評価

提案した楽曲推薦システムを実装し実験的に評価した。実験の詳細と評価結果を述べる。

4.1 提案システム評価実験概要

構築したシステムについて評価実験を行った。その実験の詳細について述べる。

(1) 実験条件

被験者数：4名

- 入力データ

位置情報・楽曲再生ログ：3週間分

ユーザ情報：各被験者の自宅住所

楽曲リスト：各被験者が自身の保有楽曲から100曲以上選択したリスト

いずれのデータも被験者により異なる。

- 位置情報・楽曲再生ログ取得環境

OS：Android

以下のアプリが扱えるバージョンとする

使用アプリ：Last.fm, Rocket Player

使用楽曲：楽曲リスト内の全楽曲

- パラメータ検証

歌詞のトピック分類時のトピック数：10/13/15を検討
 再生回数の反映：行動別楽曲情報ベクトルに再生回数を反映

このときトピック数は10とした。

(2) 評価方法

システムによる推薦結果を被験者に見せ、結果についてアンケート評価を行った。

- 質問内容

適合率

各推薦曲に対して、行動に適しているかをYes/Noで回答する。

プレイリスト評価

各推薦プレイリスト内の曲がどのくらい行動に適しているかを4段階で評価し、全く適していない場合は1、すべての曲が適している場合は4とする。

4.2 提案システム評価結果

提案した楽曲推薦システムについて評価実験を行った。今回の評価実験においては、おおむね良い結果が得られた。4名分の結果を平均したものを次に示す。

表1 推薦プレイリスト評価結果

		在宅	外出	移動	
プレイリスト評価	トピック数	10	2.5	2.75	3
		13	3	3.25	3.25
		15	3	3.25	3.25
	再生回数反映	3.25	3	3.25	

推薦プレイリスト評価結果はトピック数を10とした場合を除き、同程度の評価結果となった。これは、今回推薦されたプレイリストに対する評価がトピック数を10とした場合を除き同程度であるといえる。

表2 推薦曲の行動適合率

		在宅	外出	移動	
適合率[%]	トピック数	10	75	80	75
		13	90	92.5	97.5
		15	90	87.5	95
	再生回数反映	87.5	82.5	95	

推薦曲の行動適合率はトピック数を13とした場合が最も良い評価結果となった。楽曲の行動適合率は、提案システムを繰り返し利用した際にプレイリストがどの程度よくなっていくかを表すものである。今回の結果からはトピッ

ク数を13とした場合に、プレイリストが最も良くなっていると考えられる。

5. まとめ

今回、位置情報を利用した行動推定が可能であること、ユーザの聴きたい楽曲が行動に影響を受けることに着目し、ユーザの行動に基づく楽曲推薦システムの実現を目的にシステムの提案と構築、および評価を行った。評価実験の結果からユーザから満足を得られる十分な結果が得られた。今後は、行動の種類についてと行動推定における生活時間帯情報の反映の影響を検討する必要がある。

まず、行動の種類について述べる。今回の行動の分類は、誰もが必ず取りうる行動であるため、今回の各行動から細分化していくことで、行動の種類を増やしていくことが望ましい。例えば、外出行動を細分化すると、買い物や出社・在学などになる。細分化することで、より細かい推薦が可能となる。しかし、細分化する際には、どこまで細分化することが、ユーザにとって選択の補助となりうるのかの検討が必要となる。

次に、行動推定における生活時間帯情報の反映の影響について述べる。今回は、被験者にわかりやすい生活サイクルで活動しているものが少なかったため、要素としては検討していない。しかし、生活サイクルがわかりやすい場合は、生活時間帯情報を行動推定に反映させることで、推薦結果の改善に影響すると考えられる。生活時間帯情報を反映することで、行動推定結果がより正確になり、現在種プレイリストから漏れている楽曲再生ログを拾うことが可能となる。種プレイリストが増えるため推定の精度も向上すると考えられる。

謝辞 提案システムを実現する際に使用した、genismの作成に関わったRadimRehurek,およびMeCabの作成に関わった京都大学情報工学研究科-日本電信電話株式会社コミュニケーション科学基礎研究所ユニットプロジェクトに関わられた皆様に謹んで感謝の意を表す。

参考文献

- [1] 吉井和佳, 後藤真孝, 音楽推薦システム, 情報処理学会誌, Vol.50, No.9 (2009)
- [2] 宇野愛, 伊藤貴之, MALL: ライフログに基づく推薦機能を備えたポータブル音楽プレイヤー, WISS (2012)
- [3] 宮崎雄一郎, 山田直治, 住谷哲夫, 磯田佳徳, ユーザの行動に合わせたサービス実現のための行動推定技術の開発, NTT DOCOMO テクニカル・ジャーナル, VOL.17, NO.3, PP55-61 (2009)