

分野横断の架け橋となる情報推薦



分野横断の架け橋となる情報推薦

静岡大学情報学部
行動情報学科
奥瀬 雄哉

Bachelor Thesis Series of Yamamoto Laboratory
Faculty of Informatics,
Shizuoka University
Copyright © 2021 Yuya Okuse

目次

第 1 章　はじめに	5
第 2 章　関連研究	7
2.1　グラフ構造を考慮した情報推薦	7
2.2　意外性のある情報推薦	7
2.3　未知分野の知識習得	8
第 3 章　提案手法	9
3.1　提案アルゴリズムの概要	9
3.2　定式化	10
3.3　音楽推薦への適用例	11
第 4 章　実験	15
4.1　実験協力者	15
4.2　音楽リスト	15
4.3　実験データ	16
4.4　実験システム	16
4.5　実験手順	17
4.6　評価・測定指標	18
第 5 章　結果	19
第 6 章　考察	21
6.1　結果の解釈	21
6.2　実験設計の改善	21
6.3　提案手法の改善点	22

6.4 提案手法の応用性 23

第7章 おわりに 24

参考文献 25

図目次

3.1 グラフ構造	12
4.1 分野選択画面	17
4.2 音楽再生画面	17

表目次

3.1 提案アルゴリズムの出力例	14
5.1 各音楽リストにおけるアンケート Q2 の評価値 (上段:平均値, 下段:標準 偏差)	20

第1章

はじめに

新たな価値を創出するアプローチとして、異なる分野が連携する動きが増えている。医療では情報化が進み、電子カルテの普及やAIを活用した診断支援システムの開発が実施されている [4][2]。文部科学省は、大学での異分野融合を目的としたオープン・イノベーションの推進を支援している [3]。異分野融合の重要性は高く、自身の専門分野だけではなく、専門分野外の知見を取り入れることの重要性が今後ますます高まると考えられる。

しかし、専門外の分野に入り込むことは簡単なことではない。例えば、医療とAIの分野は大きく異なり、医学分野には馴染みのない考え方やアプローチがAI分野には存在する。そのため、医学分野の人が医学への応用を意図してAI技術を学ぼうとしても、分野間に関連性がないため、AI技術をどのように医療に応用できるのかを考えることは容易ではない。結果的に、医学分野の人がAI技術について学ぶことは難しいと考えられる。

また、新しい分野に入り込むことの難しさは、「学習」だけではなく「他分野に興味を持ちたい場合」についても同様であると考えられる。例えば、普段はクラシック音楽しか聞かない人が新しくロック音楽に興味を持ちたいとする。クラシックとロックではジャンルが大きく異なるため、興味を保ちながらロック音楽を聴き続けることは容易ではない。

本稿では、興味を持てるようになりたい分野や詳しくなりたい分野があるときに、既に興味のある分野や詳しい分野から対象分野に徐々になじんでいくための情報推薦手法を提案する。本稿では、分野横断の情報推薦の題材として「音楽」を取り上げる。本稿では、まず音声データから各音楽の波形をもとに、音楽間の類似度を求める。その後、その類似度から音楽間の関係性を示すグラフ構造を構築する。その上で、興味のある音楽分野から興味を持てるようになりたい音楽分野に段階的に遷移していくよう、グラフ構造を考慮して鑑賞する音楽を段階的に推薦する。

未知分野の知識の習得を支援する方法としては、当該分野におけるユーザのスキルや知

識レベルに応じた情報推薦が考えられる [4]. また, ある分野について全く知識がない場合は, その分野の入門書籍で学習を開始するというアプローチも考えられる. この種のアプローチは, 学習者が未知分野に関連する知識を持っている場合や, ゼロからの学習に対して高いモチベーションを持っている場合に有効であると考えらえる. しかし, 自分の興味のある分野や専門分野とは全く異なる分野の知識を学ぶような場合は, 高いモチベーションを保ちながら学習を続けることが困難であることは想像に難くない. また, Crounch らは, 新しい分野の学習において, 既存の知識と関連づけて学ぶことがプラスの影響を与えると報告している [5]. しかし, 自分の専門分野や興味のある分野の知識と専門外の分野を関連づけて学習を進めることは容易ではない. 提案手法を用いることで, 興味のある分野や詳しい分野の知識や経験を踏まえつつ, 興味を保ちながら専門外の分野の学習をスムーズに行うことが可能になる.

第2章

関連研究

2.1 グラフ構造を考慮した情報推薦

X.Wang らは、グラフ構造からユーザとアイテムの依存関係を利用し、ユーザの好みを推論することができる Knowledge-aware Path Recurrent Network というモデルを提案している [6]. H. Wang らは、グラフ構造から、レコメンドシステムのために、ユーザのパーソナライズされた潜在的な興味を学習することができる Knowledge Graph Convolutional Networks というフレームワークを提案している [7]. Lei らは、ユーザからのフィードバックに応じてグラフ構造を変更することで、動的にユーザの嗜好を把握する手法を提案している [8].

これらの関連研究では、グラフ構造からユーザの嗜好を把握する手法を提案している。本研究では、グラフ構造からユーザの嗜好を考慮し、ユーザの興味のある分野から興味を持てるようになりたい分野へと分野横断できる推薦経路を発見し、情報推薦を行う。

2.2 意外性のある情報推薦

意外性のある情報推薦の研究として、セレンディピティを誘発する情報推薦の研究がある。この情報推薦方法は、ユーザにとって新規性や意外性のある情報を推薦するものである。

Iaquinta らは、ユーザにとって新規性や意外性のある情報を推薦するために、セレンディピティ指向のレコメンドシステムを設計している [9]. Pardos らは、大学生を対象に、新規性や意外性がありつつも、学生の興味に関する性のある授業を提示するレコメンドシステムを提案している [10]. Kotkov らは、セレンディピティの8つの定義に着目し、ユー

ザの嗜好や満足度にどのような影響があるかを調査した。その結果、セレンディピティはユーザの嗜好を広げるが、満足度を低下させる場合があることを明らかにした [4].

意外性のある情報推薦では、ユーザの専門外の情報を推薦し、ユーザにとって新規性や意外性のある情報を推薦することを目指している。しかし、Kotkov らの研究結果より、意外性を狙った情報推薦はユーザの満足度を下げる可能性もある。本研究では、ユーザの専門外の情報のうち、ユーザが興味を持てるようになりたい分野、詳しくなりたい分野の情報に限定し、推薦を行う。

2.3 未知分野の知識習得

未知分野の知識を習得するための研究はいくつか行われている。Umemoto らは、未知分野の知識の習得を支援する方法として、当該分野におけるユーザのスキルや知識レベルに応じた情報推薦手法についての研究を行っている [4]. Zhao らは、オンライン学習システムにおいて、学習者の属性を組み込むことができるフレームワークを提案している。このフレームワークにより、個人の知識を追跡し、学習者の習熟度を予測することができ、各学習者に対して最適な学習内容を提案することが可能となる [12]. これらの研究では、ユーザのスキルレベルや習熟度に応じて、推薦する情報を決定している。また Crounch らは、新しい分野の学習において、既存の知識と関連づけて学ぶことがプラスの影響を与えると報告している [5].

この研究結果を踏まえて、本研究では、ユーザの既存の知識と関連づけて、未知分野の学習を行うための情報推薦を目指す。

第3章

提案手法

本研究では、ユーザが興味のある分野（もしくは詳しい分野）と興味を持てるようになりたい分野（もしくは詳しくなりたい分野）を入力することで、興味のある分野から興味を持つようになりたい分野へスムーズに横断できるような情報を段階的に推薦するアルゴリズムを提案する。

本章では、まず提案アルゴリズムの概要を述べる。その後、情報推薦の適用ドメインとして音楽を取り上げ、提案アルゴリズムの具体的な適用方法について述べる。

3.1 提案アルゴリズムの概要

前提知識や興味が持ちにくい分野の情報を得ようとする場合、唐突にその分野の知識を学ぼうとするよりは、関心のある分野や馴染みのある分野の知識と対応づけながら、徐々に対象分野の情報を習得していくほうが、心理的抵抗や習得コストが少ないと考えられる。そこで、本稿では興味のある分野から興味を持つようになりたい分野へスムーズに横断できる情報推薦のアルゴリズムを提案する上で、以下の2つの仮定を置く。

仮定1 未知の情報であっても、その情報が前提知識や関心のある分野に属するものであれば、そうでない分野の情報よりも習得にかかるコストが小さい

仮定2 情報 x に前提知識や関心がある場合、 x に類似する情報 y の習得にかかるコストは、 x に類似していない情報 z の習得にかかるコストよりも小さい

提案手法は、推薦対象となる情報および既知の情報の関係性をグラフ構造で表現する。その上で、上記仮定に基づき既知の情報となるべく類似しつつ、かつ前提知識や関心のある

分野に属する情報を徐々に取得していき、最終的に興味を持てるようになりたい分野の情報を獲得できるようなグラフ上の経路を発見・推薦する。

3.2 定式化

本稿では、ユーザが興味のある分野（もしくは詳しい分野）をソース分野、興味を持つようになりたい分野（もしくは詳しくなりたい分野）をターゲット分野と呼ぶことにする。

今、情報推薦の対象となる情報の集合を I とする。推薦対象情報 $i \in I$ は、1つ以上の分野に属することが分かっており、 i が分野 c に属することを $c \in Cat(i)$ と表すことにする。

提案手法は推薦対象情報の集合 I をノード、推薦対象情報間の関係性を表す頂点対集合 E をエッジと見なし、後述する修正版ダイクストラアルゴリズムを用いて、ソース分野からターゲット分野に至る最短経路の分析・提示を行う。提案手法は入力として、ソース分野の集合 $L = \{l_1, l_2, \dots, l_k\}$ 、ターゲット分野 t 、ユーザが獲得済みの情報の集合 $I_u (I_u \subset I)$ を受け取る。提案手法は入力をもとに、順序付き情報系列 $O = \{(i_1, i_2, \dots, i_{m-1}, i_m) | m \geq 2, i_1 \in I_u, \forall k \geq 1, i_k \in \overline{I_u} \cap I, L \cap C(i_1) \neq \emptyset, t \in C(i_m)\}$ を出力する。以下、修正版ダイクストラアルゴリズムについて述べる。

ソース分野の集合は、集合 I_u から暗黙的に取得する。集合 I_u のそれぞれの情報が属する分野を明らかにする。その後、各分野の出現回数を求め、多い順に上位 k 件をソース分野とする。また、情報 i_n が属するソース分野の数を $u_n (0 \leq u_n \leq k)$ とする。グラフ構造において、集合 I の要素がノードとなる。またノード i_n は、 u_n の情報を保持する。次に、ある情報 i_m と i_n の類似度の大きさを $s_{mn} (0 \leq s_{mn} \leq 1)$ とする。 s_{mn} の大きさが 0 より大きいとき、それらを類似する情報とみなし、ノード間にエッジが張られる。情報 i_m と i_n 間に張られるエッジを e_{mn} とし、エッジの集合を $E = \{e_{12}, e_{34}, \dots, e_{mn}\}$ とする。

本手法では、エッジの重みを、類似度とそのエッジの遷移先ノードの u_n の値から導く。 s_{mn} がとりうる値の範囲と u_n がとりうる値の範囲が異なる。そこで、min-max スケーリングにより、 u_n のとりうる値の範囲をスケーリングする。スケーリング後の u_n の値を z_n とすると、 z_n の値は以下の式で求めることができる。

$$z_n = \frac{u_n - u_{min}}{u_{max} - u_{min}} \quad (3.1)$$

ここで、 u_{max} は u_n がとりうる値の最大値、 u_{min} は u_n がとりうる値の最小値である。

以上のスケーリングにより、類似度 $s_{(m,n)}$ とノードが属するソース分野の数 z_n がとりうる値は 0 以上 1 以下の範囲で一致する。 s_{mn} と z_n から、エッジ e_{mn} の重み w_{mn} の値は以下の式で求まる。

$$w_{mn} = \alpha s_{mn} + (1 - \alpha) z_n \quad (3.2)$$

パラメータとして $\alpha (0 \leq \alpha \leq 1)$ を設定し、最適なエッジの重みを求める。ここで、エッジの重みは、類似度とノードが属するソース分野の数を足し合わせた値となっている。そのため、重みが大きいエッジの遷移先ノードは、1つ前のノードと類似性が高く、かつ多くのソース分野に属したノードということとなり、重みの大きいエッジが推薦するべき経路であるといえる。本手法では、最終的にダイクストラアルゴリズムにより、グラフ構造から最短経路を発見する。そのため、 w_{mn} の逆数をとったものを最終的なエッジの重み g_{mn} とする。

$$g_{mn} = 1/w_{mn} \quad (3.3)$$

これまでの式から、作成されるグラフ構造を図 3.1 に示す。

提案アルゴリズムではスタートをソース分野の情報群、ゴールをターゲット分野の情報群とする。式 3.3 の値をエッジの重みとして構築したグラフから、ダイクストラアルゴリズムにより、最短経路を求める。それにより、前提知識や関心のある分野の情報となるべく類似し、かつそのような情報と類似性のある情報を段階的に推薦する。

3.3 音楽推薦への適用例

本研究では、音楽に関するデータセットを用いて、提案アルゴリズムによる情報推薦を行う。

3.3.1 類似度の定義

本研究では「ISMIR04 Genre Identification task dataset」^{*1}というデータセットを用いる。本データには、729曲の音楽データがあり、全ての音楽に対するmp3ファイルがある。また、「classical, electronic, jazz-blues, metal-punk, rock-pop, world」の6種類の分野があり、それぞれの音楽はどれか1つの分野に属している。本研究では、各音楽の

^{*1} <https://zenodo.org/record/1302992#.YBJ8dnczZhH>

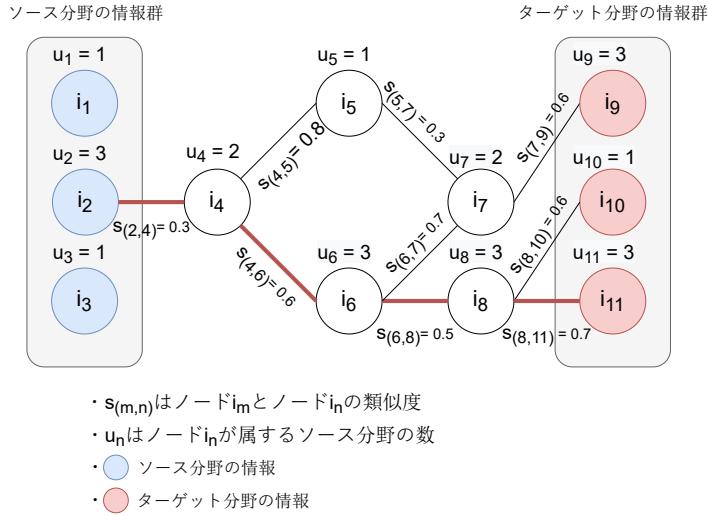


図 3.1 グラフ構造

波形を比較することで、音楽間の類似度を求めた。はじめに、mp3 ファイルからスペクトル特徴量のメル周波数ケプストラム係数 (MFCC)[[13](#)] を時間フレーム数分だけ抽出する。波形を比較する際、音楽によってフレーム数が異なる。そのため、系列長が異なるデータ同士の比較が可能な動的時間伸縮法 [[13](#)] を用いることで、フレーム数が異なる音楽同士の比較を行った。以上により、各音楽の波形を比較することで、音楽間の類似度を算出した。

3.3.2 グラフ構築

本研究では、節 [3.3.1](#) で算出した音楽間の類似度からグラフ構造を構築した。各ノードは音楽を表す。ある音楽間の類似度が 0 より大きい場合は、その音楽間に類似性があるとみなし、そのノード間にエッジが張られる。また、各ノードは、その音楽が属する分野の情報を保持する。

3.3.3 音楽推薦における提案アルゴリズムの出力例

「ISMIR04 Genre Identification task dataset」のデータに対して、提案アルゴリズムを適応させた際の出力例を以下に示す。表 [3.1](#) はソース分野を jazz&blues、ターゲット分野を world とし、提案アルゴリズムを実行した際の出力例である。音楽 ID は、データセットに用意されているユニークな ID である。リストの行は、提案アルゴリズムが出力

した推薦経路の音楽 ID と、その推薦経路での式 B.3 の値の累計を表す。出力されるリストは、累計コストが小さい順にソートされている。

表 3.1 提案アルゴリズムの出力例

音楽 ID	音楽 ID	音楽 ID	累計コスト
artist_73_album_1_track_6	artist_128_album_1_track_2		1.07
artist_73_album_1_track_6	artist_127_album_4_track_2		1.15
artist_73_album_1_track_6	artist_75_album_2_track_1	artist_127_album_3_track_1	1.27
.	.	.	.
.	.	.	.
.	.	.	.
artist_73_album_1_track_6	artist_75_album_2_track_1	artist_111_album_1_track_1	1.74
artist_73_album_1_track_6	artist_110_album_1_track_1		1.84

第 4 章

実験

本章では、提案アルゴリズムが推薦した情報がユーザの分野横断を可能にするのかを検証するための実験を行った。実験では、比較手法と提案手法から提示された音楽リストの音楽を聴き、ターゲット分野の音楽を好きになることができるかを調べることで、提案手法の有効性を評価した。

4.1 実験協力者

静岡大学浜松キャンパスの学部生および大学院生に協力を依頼した。計 24 名（男性 18 名、女性 6 人）の実験協力者のデータを分析に利用した。

4.2 音楽リスト

本実験では、入力されたソース分野及びターゲット分野から、ターゲット分野の音楽を好きになれるような音楽リストを 3 種類用意した。

SimilarityList 提案アルゴリズムにおいて、式 3.2 の α の値を 1 として算出されたリスト

SimilaritySourceList 提案アルゴリズムにおいて、式 3.2 の α の値を 0.5 として算出されたリスト

TypicalList データセットにおけるデータの典型性を考慮して算出されたリスト

SimilarityList は、式 3.2 の α の値を 1 とし、提案アルゴリズムが音楽間の類似度の

みから算出した推薦経路の音楽リストである。つまり、**SimilarityList** は、実験協力者のソース分野からターゲット分野へと音楽が遷移する際、リスト内の x 番目の音楽は、 $x - 1$ 番目の音楽と最も類似度の高い音楽となる。

SimilaritySourceList は、式 B.2 の α の値を 0.5 とし、提案アルゴリズムが音楽間の類似度と、音楽が属するソース分野の数から算出した推薦経路の音楽リストである。つまり、**SimilaritySourceList** は、実験協力者のソース分野からターゲット分野へと音楽が遷移する際、リスト内の x 番目の音楽は、 $x - 1$ 番目の音楽と類似度が高く、またソース分野に属する音楽となる。

TypicalList は、VisualRank[14] を参考に、データセットの典型性を考慮して算出されたリストである。本実験では、VisualRank アルゴリズムにおける画像データを音楽データと読みかえ、データ間の類似度を MFCC から計算することで、各音楽データの典型性を求めた。

4.3 実験データ

本実験では、あるソース分野及びターゲット分野の入力の組みに対して、上記の 3 種類の音楽リストデータを用意した。提案アルゴリズムは、ソース分野及びターゲット分野を入力とし、スタートをソース分野に属する音楽、ゴールをターゲット分野に属する音楽とする。そのため、ソース分野とターゲット分野の入力に対して、提案アルゴリズムは複数の推薦経路を算出する。そこで、実験の際は、提案アルゴリズムが算出した推薦経路のうち、式 B.3 のエッジの重みの累計が最も小さい経路を、その入力された分野間の推薦経路とし、実験に用いる音楽リストとした。したがって、全てのソース分野・ターゲット分野の組み合わせについて、**SimilarityList**, **SimilaritySourceList** のデータを 1 つずつ用意した。また、各分野に関して、典型的な曲上位 3 件のデータを **TypicalList** として用意した。

4.4 実験システム

本研究では、ユーザ実験のためのシステム（以下、実験システム）を構築した。実験システムは、入力としてソース分野とターゲット分野を受け取ることで、その分野に応じて、節 4.2 の音楽リスト 3 種類を表示するものである。実験協力者は、分野選択画面において、ソース分野とターゲット分野を選択する。図 4.1 が実際の分野選択画面である。音楽再生画面では、選択された分野に応じた音楽リストの曲を再生できるようになっている。

図 4.2 が実際の音楽再生画面である。1 つの音楽再生画面には、1 つのリストの音楽再生 UI が表示される。そのため、実験協力者は合計で 3 つの音楽再生画面を使用する。

学籍番号

興味のある音楽カテゴリ
[ロック&ポップ]

興味を持てるようになりたい音楽カテゴリ
[ロック&ポップ]

「興味のある音楽カテゴリ」と「興味を持てるようになりたい音楽カテゴリ」は、必ず異なるものを選んでください。また後のアンケートで使用しますので、選択した2つのカテゴリを忘れないようお願いします。

選択

図 4.1 分野選択画面



図 4.2 音楽再生画面

4.5 実験手順

本実験では、節 4.2 の 3 つの音楽リスト内の音楽を聴いてもらい、それぞれの音楽についてのアンケートに回答をしてもらった。実験は、下記の構成で行った。

- (1) 実験の概要説明
- (2) ソース分野及びターゲット分野の入力
- (3) 音楽リストの音楽の再生・評価
- (4) 音楽リストについてのアンケートへの回答

はじめに、実験協力者に、「ロック&ポップ、クラシック、エレクトロニック、ジャズ、メタル&パンク、ワールド」の中から、ソース分野とターゲット分野を選択してもらった。次に、入力された2つの分野に基づいて、実験システムの音声再生画面にTypicalListの音楽が表示される。実験協力者には、TypicalListの音楽を上から順番に再生してもらった。実験協力者は1つの音楽を全て聴く必要はなく、早送りおよび巻き戻しを行い、音楽を適切に評価できたと判断したタイミングで、その音楽の評価をしてもらい、次の音楽の再生へと移ってもらった。各音楽の評価値は、事後アンケートにて必要となる。そのため、実験協力者は、TypicalList内のそれぞれの音楽に対して評価をしたタイミングで、5段階評価の値をメモしてもらった。TypicalList内の全ての音楽の評価が完了したら、完了ボタンを押してもらい、TypicalListに関する事後アンケートに回答してもらった。アンケート回答後、上記の手順(2), (3)をSimilarityList, SimilaritySourceListについても同様に行ってもらった。

4.6 評価・測定指標

実験協力者が分野横断をできたかどうかを評価するためにするため、事後アンケートにおいて以下の2つの質問にそれぞれのリストについて5段階評価で回答をもらった。

- Q1 音楽リスト内のそれぞれの音楽に対する評価をしてください。
- Q2 今回の音楽リストの順に音楽を聴くことによって、あなたが選択した「興味を持てるようになりたい分野」の曲に興味を持つことはできましたか？

第 5 章

結果

本章では、計 24 名の実験協力者からのアンケート分析結果について述べる。

提示されたリスト内の音楽を順番に聴くことで、実験協力者がターゲット分野に属する音楽に興味を持つことができたかを調べるために、アンケートの Q2 の評価値を調査した。表 5-1 は各音楽リストにおけるアンケート Q2 の評価値についてまとめたものである。それぞれのリストにおける評価値の平均値は TypicalList が 3.33 と最も大きく、続いて SimilaritySourceList の平均値が 3.29、SimilarityList の平均値が 3.25 という結果であった。ここで、これら 3 群間に差があるかどうかを調べるために、検定を行った。はじめに、それぞれのリストのデータが正規分布しているかどうかを調べるために、シャピロ・ウィルク検定を行った。その結果、SimilarityList のデータは正規分布しているが、TypicalList と SimilaritySourceList は正規分布していなかった。そのため、ノンパラメトリックな検定法である Friedman 検定を行った。その結果、統計的有意差は確認されなかった ($p = 0.82$)。

表 5.1 各音楽リストにおけるアンケート Q2 の評価値（上段：平均値、下段：標準偏差）

	TypicalList	SimilarityList	SimilaritySourceList	p 値
音楽リストの評価値	3.33 (0.91)	3.25 (1.07)	3.29 (1.23)	0.82

第 6 章

考察

本章では、5 章の実験結果をもとに考察を行う。本研究の目的は、興味を持てるようになりたい分野や詳しくなりたい分野があるとき、既に興味がある分野や詳しい分野から対象分野に徐々になじんでいくための情報推薦経路を算出することである。

6.1 結果の解釈

本研究では、仮定 1、仮定 2 からアルゴリズムを提案した。比較手法および提案手法から算出された 3 つのリスト間において、統計的有意差を確認することができなかった。そのため、提案アルゴリズムにより算出された音楽リストを聴くことがユーザの分野横断を可能にすると立証することはできなかった。

6.2 実験設計の改善

実験結果から、提案手法の有効性は確認できなかった。その理由の 1 つとして実験設計があげられる。

1 つ目の問題点は、評価指標の設定である。本実験では、アンケート Q2 において、実験協力者にターゲット分野に属する音楽に興味を持つことができたかどうかについて評価を行ってもらった。しかし、この評価指標では、実験協力者がターゲット分野に属する音楽自体に興味を持ったのか、ターゲット分野自体に興味を持ったのかを判断することはできない。そのため、実験では「音楽リストの音楽を順番に聴くことで、ターゲット分野自体に興味を持つことができたかどうか」をアンケートの質問として設定するべきであったと考える。だが、この評価指標であっても、最終的にターゲット分野に属する音楽をもと

にして，ターゲット分野自体の評価をすることになる。この場合，最後に聴いた音楽そのものが評価に大きく影響すると考える。最後の音楽が実験協力者好みであれば，ターゲット分野自体にも興味を持つことができたと評価をし，逆に，最後の音楽が好みでない場合，ターゲット分野自体に興味を持つことができないと評価する可能性がある。したがって，聴いた音楽に関わらず，ターゲット分野に興味を持つことができたかを評価できる指標を設定する必要がある。

また，本実験では，各音楽リストについて，音楽を順番に聴くことでターゲット分野の音楽に興味を持つことができたかを評価してもらった。しかし，この評価方法では，実験協力者の分野横断のためにどのリストが有効であったかを評価することが難しいと考えられる。そのため，どのリストが分野横断をするにあたり有効であったかを評価するために，実験協力者に全てのリストを比較してもらい，どのリストが最も有効であったかを評価してもらう必要があった。また，本実験では，実験協力者に音楽リスト内の音楽同士のつながりを評価してもらう質問項目を設定しなかった。本研究の目的は，ソース分野からターゲット分野へと徐々に馴染んでいくための情報推薦を行うことである。最終的にターゲット分野に興味を持つことができたを評価することに加え，提案アルゴリズムが output した推薦経路の音楽が，ソース分野からターゲット分野へスマーズに遷移しているかどうかを評価してもらう必要がある。

2つ目の問題点は，実験に用いるリスト内の音楽の数が少なかったことである。本実験で用いたリスト内の音楽は，提案アルゴリズムが output した経路上の音楽である。提案アルゴリズムは最終的に最短経路問題に落とし込み経路出力を行った。そのため，実験に用いた音楽リストの音楽は多くても 3 曲であった。音楽リストの長さが短い場合，ソース分野からターゲット分野までに聴く音楽が少ないため，ユーザが分野横断することは難しいと考えられる。しかし，音楽リストの長さが長すぎても，ターゲット分野の音楽までに聴く音楽が多くなり，実際の利用を考えるとユーザが分野横断をするまでのコストが大きくなってしまう。したがって，実際にユーザが分野横断をすることができるような長さの音楽リストを実験で用いる必要がある。

6.3 提案手法の改善点

6.3.1 推薦経路決定時に考慮する分野

本研究では，推薦経路を決定する際に，遷移先ノードの属するソース分野の数を考慮した。ソース分野からターゲット分野へと分野横断をする際，推薦される情報がソース分野

に属している方が、ユーザにとって習得しやすい情報であると考えたからである。しかし、学習をしていくにつれて、ソース分野との関連性に比べ、ターゲット分野との関連性の方が重要になると考える。例えば、医学分野の人が新しく機械学習を学ぶ際に、初めは医学と関連した情報の方が習得しやすいが、学習を進めていくにつれて、医学との関連性よりも、機械学習とより関連性のある情報の方が、分野横断に適していると考える。そのため、各ノードの推薦経路上での位置に応じて、遷移先ノードを決定する際に考慮する分野を動的に変化させることが必要であると考える。

6.3.2 習得コスト

本研究では、推薦経路を最短経路問題に落とし込み算出した。そのため、提案手法が推薦する経路上の情報の数は3つほどであった。しかし、実際にユーザが本推薦システムを利用するケースを考えると、推薦される情報の数が少ない経路が、必ずしもユーザの分野横断に適しているというわけではない。例えば、医学分野の人が新しく機械学習について学びたいとした時、「できるだけ少ない工数で学習を行いたいユーザ」と「工数が多くても良いので段階的に学びたいユーザ」がいると考えられる。そのため、推薦される情報の数をユーザのニーズに合わせて動的に変化させることにより、ユーザに適した情報推薦を行うことができると考える。

6.4 提案手法の応用性

本稿では、提案手法の適用例として「音楽」と取り上げた。提案手法は音楽以外にも適応させることが可能である。提案アルゴリズムは、情報集合において、情報間での類似度及び、それぞれの情報が属する分野が決められているものに対して、適用可能である。例えば、「学習」を適用例とし、ソース分野の学習内容からターゲット分野の学習内容へと類似する内容を推薦することも可能となる。本稿では、分野横断をする対象を「音楽」とし、提案手法の評価実験を行った。しかし、音楽では、実験協力者の評価から実験協力者自身が分野横断できたのかどうかを評価するのは難しいと考えられる。[\[5\]](#), [\[12\]](#) のように「学習」を適用例とし、習熟度などの客観的に数値として測れる事柄で実験を行うことで、提案手法を適切に評価することができると考えた。

第7章

おわりに

本稿では、興味を持てるようになりたい分野や詳しくなりたい分野があるとき、既に興味がある分野や詳しい分野から対象分野に徐々になじんでいくための情報推薦手法を提案した。

提案手法を評価するために、ユーザ実験を行い、比較手法と提案手法から算出される音楽リストを聴くことで、提案手法による情報推薦により、ユーザが分野横断をすることができるかを検証した。ユーザがターゲット分野に属する音楽に興味を持つことができるようになったかを調査するための評価値には、どのリスト間にも、統計的有意差は見られなかった。そのため、提案手法による情報推薦がユーザの分野横断を可能にすることを立証することができなかった。本稿の実験設計では、本研究の目的を達成できたかどうかを測るための十分な指標を設定することができていなかったと考えられる。これより、提案手法を適切に評価するためにも、実験設計を改善する必要がある。

提案手法は、「音楽」に限らず、様々なテーマに適用可能である。提案手法による情報推薦により、人々が興味を持ちたいと思っている分野に簡単に興味持てるようになることを期待する。

参考文献

- [1] 総務省. 消費者視点に基づく新製品・サービス提供のための異業種連携, 2018. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h30/html/nd124220.html>.
- [2] 総務省. 医療・介護・健康分野におけるict利活用の推進, 2019. <https://www.soumu.go.jp/johotsusintokei/whitepaper/ja/r01/html/nd246220.html>.
- [3] 文部科学省. 文教・科学技術施策の動向と展開, 2015. https://www.mext.go.jp/b_menu/hakusho/html/hpab201601/1375335_013.pdf.
- [4] Kazutoshi Umemoto, Tova Milo, and Masaru Kitsuregawa. Toward recommendation for upskilling: Modeling skill improvement and item difficulty in action sequences. In *2020 IEEE 36th International Conference on Data Engineering (ICDE)*, pp. 169–180. IEEE, 2020.
- [5] Catherine H Crouch, Panchompoo Wisittanawat, Ming Cai, and K Ann Renninger. Life science students’ attitudes, interest, and performance in introductory physics for life sciences: An exploratory study. In *Physical Review Physics Education Research*, pp. 010111–1–010111–14. APS, 2018.
- [6] Xiang Wang, Dingxian Wang, Canran Xu, Xiangnan He, Yixin Cao, and Tat-Seng Chua. Explainable reasoning over knowledge graphs for recommendation. In *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 33, pp. 5329–5336, 2019.
- [7] Hongwei Wang, Miao Zhao, Xing Xie, Wenjie Li, and Minyi Guo. Knowledge graph convolutional networks for recommender systems. In *The world wide web conference*, pp. 3307–3313, 2019.
- [8] Wenqiang Lei, Gangyi Zhang, Xiangnan He, Yisong Miao, Xiang Wang, Liang Chen, and Tat-Seng Chua. Interactive path reasoning on graph for conversa-

- tional recommendation. In *Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*, pp. 2073–2083, 2020.
- [9] Leo Iaquinta, Marco De Gemmis, Pasquale Lops, Giovanni Semeraro, Michele Filannino, and Piero Molino. Introducing serendipity in a content-based recommender system. In *2008 Eighth International Conference on Hybrid Intelligent Systems*, pp. 168–173. IEEE, 2008.
 - [10] Zachary A Pardos and Weijie Jiang. Designing for serendipity in a university course recommendation system. In *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, pp. 350–359, 2020.
 - [11] Denis Kotkov, Joseph A Konstan, Qian Zhao, and Jari Veijalainen. Investigating serendipity in recommender systems based on real user feedback. In *Proceedings of the 33rd Annual ACM Symposium on Applied Computing*, pp. 1341–1350, 2018.
 - [12] Jinjin Zhao, Shreyansh Bhatt, Candace Thille, Dawn Zimmaro, and Neelesh Gattani. Interpretable personalized knowledge tracing and next learning activity recommendation. In *Proceedings of the Seventh ACM Conference on Learning@ Scale*, pp. 325–328, 2020.
 - [13] Lindasalwa Muda, Mumtaj Begam, and Irraivan Elamvazuthi. Voice recognition algorithms using mel frequency cepstral coefficient (mfcc) and dynamic time warping (dtw) techniques. *arXiv preprint arXiv:1003.4083*, 2010.
 - [14] Yushi Jing and Shumeet Baluja. Visualrank: Applying pagerank to large-scale image search. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 30, No. 11, pp. 1877–1890, 2008.

謝辞

本研究の遂行ならびに論文の作成にあたり，ご指導を賜りました静岡大学情報学部講師山本祐輔先生に謹んで深謝の意を表します。本論文をまとめるにあたり副査として有益なご助言とご教示を賜りました静岡大学情報学部准教授福田直樹先生に心より感謝の意を表します。本研究の遂行にご協力いただいた，静岡大学情報学部行動情報学科山本研究室の皆様に感謝致します。

2021年3月 奥瀬雄哉