



Title	深層学習と知識グラフ解析に基づくハイブリッド推薦システムに関する研究
Author(s)	櫻井, 慶悟
Degree Grantor	北海道大学
Degree Name	博士(情報科学)
Dissertation Number	甲第16673号
Issue Date	2025-09-25
DOI	https://doi.org/10.14943/doctoral.k16673
Doc URL	https://hdl.handle.net/2115/98283
Type	doctoral thesis
File Information	Keigo_Sakurai.pdf, 全文



博士論文

深層学習と知識グラフ解析に基づく
ハイブリッド推薦システムに関する研究



北海道大学 大学院情報科学院
情報科学専攻

櫻井 慶悟

2025 年 9 月

目次

第1章 序論	5
1.1 本研究の背景	5
1.2 本研究の目的	7
1.3 論文の構成	8
第2章 関連研究	10
2.1 はじめに	10
2.2 協調フィルタリングに基づく推薦手法に関する研究	10
2.3 コンテンツベースフィルタリングに基づく推薦手法に関する研究	12
2.4 ハイブリッド推薦に関する研究	14
2.4.1 特徴統合に基づく手法	14
2.4.2 知識グラフに基づく手法	15
2.5 まとめ	16
第3章 知識グラフと強化学習に基づくプレイリスト生成	19
3.1 はじめに	19
3.2 Preliminary	20
3.3 提案手法	21
3.3.1 マルコフ決定過程の定式化	23
3.3.2 知識グラフを用いた楽曲特徴の抽出	24
3.3.3 状態表現の設定	25
3.3.4 報酬関数の設定	26
3.3.5 最適化	30
3.4 実験	31
3.4.1 実験設定	32
3.4.2 結果と考察	35
3.5 結論	37
第4章 音響特徴を導入した知識グラフを用いた楽曲推薦	38
4.1 はじめに	38
4.2 提案手法	39

4.2.1	知識グラフの構築	41
4.2.2	知識グラフ上でのエージェント学習	41
4.2.3	方策に基づく楽曲の推薦	45
4.3	実験	47
4.3.1	実験設定	47
4.3.2	実験結果	49
4.4	結論	52
第 5 章	LLM の出力を用いた知識グラフに基づくアイテム推薦	53
5.1	はじめに	53
5.2	提案手法	54
5.2.1	LLM によるユーザの将来的嗜好に関する直感の獲得	56
5.2.2	LLM の直感と強化学習に基づく知識グラフ推論による推薦 . .	57
5.3	実験	59
5.3.1	実験設定	59
5.3.2	全体性能比較 (RQ1)	61
5.3.3	LLM の種類とプロンプトの影響分析 (RQ2)	62
5.3.4	推薦戦略の影響分析 (RQ3)	63
5.4	結論	66
第 6 章	結論	67
	謝辞	69
	参考文献	70
	著者の研究業績	79

図目次

2.1	リサーチマップと本研究の位置付け.	18
3.1	提案するプレイリスト生成手法の概要図.	22
4.1	提案する楽曲推薦手法の概要図.	40
4.2	報酬 R_t の計算方法の概要.	43
4.3	知識グラフ $\mathcal{G}$ におけるビームサーチアルゴリズムの概要.	46
4.4	PM および CMs 1-5 における推薦性能の比較.	51
5.1	LIK R の概要図.	55
5.2	LLM の直感に基づく報酬と知識グラフ埋め込みに基づく報酬の影響.	65

表目次

3.1	PM ベースのモデルと報酬関数 R_t における各パラメータ値.	34
3.2	nDCG@ k および Hit Rate@ k ($k = 1, 5, 10$) の結果. 各評価指標における 最大値を太字で示す.	36
3.3	M_{aco} , M_{pop} , M_{new} の結果. 太字は最大値, 下線は2番目に高い値を示 す.	37
4.1	推薦可能ユーザ率 [%].	49
5.1	各データセットの統計量.	60
5.2	MovieLens-1M および Lastfm-1M における推薦性能の比較. 最も高い値 を太字, 2番目に高い値を <u>下線付き</u> で示す.	62
5.3	使用する LLM の違いによる LIKR の推薦性能の比較.	63
5.4	プロンプトの違い (時系列意識の有無) による LIKR の推薦性能の比 較.	63

第1章

序論

1.1 本研究の背景

インターネットとデジタル技術の急速な発展により、人々は膨大な量の情報やコンテンツに日常的にアクセス可能になった [1,2]. 音楽、映像、書籍、ニュース記事、商品レビューなど、ユーザが接触可能な情報やアイテムの種類と量はかつてないほど増加している。これにより、ユーザが目的に適したコンテンツを探索し、選択する際には多大な労力を要する。この現象は「情報過多 (information overload)」と呼ばれ、現代社会における重要課題である [3-5]. 情報過多の状況下では、ユーザは有用な情報に辿り着く前に大量の無関係な情報に直面し、探索の効率性や意思決定の質が低下する可能性がある。このような情報過多の問題に対処可能な技術として、推薦システムが広く活用されている [6]. 推薦システムは、ユーザの過去の行動履歴やアイテムの属性情報をもとに、個々のユーザにとって有益または興味深いと推定されるアイテムを自動的に提示するシステムである [7,8]. 特に、オンラインショッピングサイトや音楽ストリーミングサービス等、ユーザとのインタラクションが多く発生するサービスにおいては、推薦システムがユーザエクスペリエンスの向上やサービスの継続利用において中核的な役割を果たしている。

推薦システムにおける代表的なアプローチには、協調フィルタリングとコンテンツベースフィルタリングの二種類が存在する [9]. 協調フィルタリングは、ユーザの視聴履歴や購入履歴、評価スコアなどを分析し、類似した嗜好を持つ他のユーザを特定することで、それらのユーザの履歴内のアイテムを推薦対象として提示する手法である [10-12]. この手法はユーザの明示的な嗜好モデルを必要とせず、

ユーザ間の集合知を活用することで高精度なアイテムの推薦が可能である．一方で，十分な履歴が存在しない新規ユーザに対しては，推薦精度が低下するという「コールドスタート問題 (cold start problem)」が顕在化する [13,14]．コンテンツベースフィルタリングは，ユーザの履歴内に存在するアイテムに類似した属性を持つ他のアイテムを推薦するアプローチである [15,16]．具体的に，ジャンル，作成者，キーワード，時間帯などのメタ情報を活用することで個人の嗜好に合致した推薦を行う．アイテムの属性情報を用いるため，他ユーザの行動履歴に依存することなく推薦することができ，コールドスタート問題を回避可能である．しかしながら，ユーザの嗜好は必ずしも明示的な属性のみで説明しきれるものではなく，潜在的・多面的な関心やその変化を捉えることが困難であるため，十分な推薦精度が得られないケースも少なくない．

このように，協調フィルタリングとコンテンツベースフィルタリングは，それぞれに特長と限界を持ち，推薦システムの根幹技術として広く利用されているものの，単独で用いた場合には情報の制約やデータの偏りにより，ユーザに最適な推薦を常に実現することは難しい．こうした背景のもと，両者の利点を統合することでそれぞれの欠点を補完し合う「ハイブリッド推薦 (hybrid recommendation)」が注目されている．ハイブリッド推薦は，協調フィルタリングによるユーザ間の嗜好類似性と，コンテンツベースの属性情報に基づく個別嗜好モデルを組み合わせることで，推薦精度の向上やコールドスタート問題の緩和を図る手法である [17–19]．近年では，ユーザ・アイテム・属性といった異なる情報源を統合的に扱うための有効な表現として，「知識グラフ (knowledge graph)」の活用が進められている．知識グラフは，ノードとしてユーザやアイテム，属性，カテゴリなどを表現し，それらの関係性をエッジで接続することで構造化された情報ネットワークを構成する [20–22]．これにより，単一の手法による推薦と比較して，より複雑で高次の関係に基づく推薦が実現される．

機械学習技術の発展により，知識グラフからの情報抽出や推薦モデルの構築は飛躍的に発展したが，それでもなお実用上の課題は残されている．第一の課題は，ユーザの目的や状況に応じた柔軟な推薦方針の制御が困難であるという点である．従来手法では知識グラフを用いた推薦において，ユーザの目的や状況といった文脈情報を明示的に入力・反映する仕組みが限定的であり，それに応じた嗜好の変動

を柔軟に推薦結果へ反映することが困難であった．例えば音楽ドメインでは，同一ユーザにおいても，時間帯や状況に応じて好まれる楽曲が異なるが，従来手法ではこれらに起因する明示的な意図を推薦に反映することが難しい．第二の課題は，コールドスタート状況における意味的推論能力の限界が存在する点である．実データに基づいて構築された知識グラフは，推薦に有効な情報を十分に含んでいない可能性があるため，意味的關係がスパースになりやすい．このスパース性により，潜在的に関連したアイテム間の關係が限定されるため，依然としてユーザの履歴が乏しいコールドスタート状況では，嗜好に合致するアイテムとの意味的關係を把握・抽出することが難しく，適切な推薦を行うことが困難となる．

以上のように，情報過多の現代社会において推薦システムは重要な役割を担っているが，その精度と柔軟性を高めるには，ユーザの目的や状況に応じた推薦方針の動的制御と，コールドスタート環境下での高い意味推論能力の両立が求められている．

1.2 本研究の目的

本研究は，ハイブリッド推薦における「ユーザの目的や状況に応じた柔軟な推薦方針の制御の困難さ」と「コールドスタート状況における意味的推論能力の限界」に対処することを目的とする．これを実現するために，深層学習と知識グラフ解析を基盤とし，ユーザの明示的な意図や潜在的な嗜好を考慮しながら，柔軟かつ高精度な推薦を実現するハイブリッド推薦システムを提案する．

まず，1つ目の課題であるユーザの目的や状況に応じた推薦の柔軟な制御に対応するために，本研究では特に音楽ドメインにおける実用性の高い推薦タスクである「プレイリスト生成」に着目する．音楽の推薦においては，状況や意図に応じて異なる選好を持つことが多い．本研究では，こうしたユーザの意図を推薦方針に明示的に反映するため，強化学習（Reinforcement Learning）の枠組みを導入し，報酬関数の設計によって推薦方針を制御可能なプレイリスト生成手法を提案する．具体的には，知識グラフを活用して楽曲間の意味的關係を特徴量として抽出し，それらを基盤とした推薦環境を構築する．その上で，ユーザの入力する意図に応じて重みを調整可能な報酬関数を設計し，推薦方針を動的に変更できるようにすることで，文脈に応じた柔軟な推薦を実現する．本手法により，従来困難であった「状

況や目的に応じた推薦の制御」が可能となる。

次に、2つ目の課題であるコールドスタート状況における意味的推論の限界に対処するため、本研究では知識グラフの構造的限界を補完する2つの技術的アプローチを提案する。第一のアプローチは、音響特徴に基づく知識グラフの拡張である。従来の知識グラフは、主に明示的なメタ情報（例：ジャンル、アーティスト、リリース年など）に基づいて構築されているが、音楽の「聴感上の類似性」や「雰囲気」はこのような属性だけでは捉えきれない。本研究では、楽曲に内在する音響的特徴を分析し、それらを知識グラフ上に新たな関係として追加する。これにより、音響的に類似した楽曲間の意味的つながりが強化され、履歴情報が少ないユーザに対しても、嗜好に合致する推薦が可能となる。特に、履歴が存在しない新規ユーザや新規アイテムにおいても、音響的特徴の共通性を手がかりにした推薦が期待できる。第二のアプローチは、大規模言語モデル（Large Language Models, LLM）を用いた知識グラフ推論の補完である。LLMは、自然言語理解と知識推論において近年顕著な性能向上を示しており、ユーザの履歴や記述的情報をもとに高次の意味を抽出する能力を有する。本研究では、ユーザの行動履歴やプロフィールを入力としてLLMを活用し、ユーザの嗜好や興味に関する抽象的かつ潜在的な特徴を推定する。さらに、LLMから得られた情報を知識グラフ上の推論に組み込むことで、既存の構造にない潜在的な関係性の補完を図る。これにより、知識グラフのスパース性によって生じる推論の限界を緩和し、履歴の乏しいユーザに対しても意味的に妥当な推薦を提供することが可能となる。

以上の3つの技術的提案、すなわち「報酬関数設計による意図に応じた推薦方針の制御」「音響特徴による知識グラフの拡張」「大規模言語モデルによる意味推論の補完」を統合することで、本研究では従来手法の限界を克服し、ユーザの多様な状況や嗜好に対応可能なハイブリッド推薦システムの実現を目指す。

1.3 論文の構成

本論文は6章からなり、以下に各章の概要を述べる。

- 第1章では、研究の背景および目的について述べた。
- 第2章では、協調フィルタリング、コンテンツベースフィルタリング、そして

それらを統合するハイブリッド推薦手法に関する従来手法について説明し、現状の課題と本論文の解決手法について明確にする。

- 第3章では、知識グラフを用いた強化学習に基づく楽曲プレイリストの生成手法を提案し、ユーザの明示的な意図に応じたプレイリストを生成可能であることを確認する。
- 第4章では、音響特徴を導入した知識グラフを用いた楽曲推薦手法を提案し、コールドスタート問題が緩和されることを確認する。
- 第5章では、LLMの出力を用いた知識グラフに基づくアイテム推薦手法を提案し、実験により、その有効性を確認する。
- 第6章では、本論文のまとめと今後の課題について述べる。

第2章

関連研究

2.1 はじめに

本章では、本論文の関連研究について整理することで、関連研究の中での本論文の位置づけおよび意義を明確にする。従来の推薦システムを実現する手法は、協調フィルタリングとコンテンツベースフィルタリングに大別される。さらに、第1章で説明したこれらの手法の課題を解決するための手法として、ハイブリッド推薦手法が提案されている。本章では、協調フィルタリング手法、コンテンツベースフィルタリング手法およびハイブリッド推薦手法に関する従来研究について述べ、本研究の位置づけを明確にする。以降、2.2節で協調フィルタリングに基づく推薦手法に関する従来研究について説明する。次に、2.3節で、コンテンツベースフィルタリングに基づく推薦手法に関する従来研究について説明する。また、2.4節で、ハイブリッド推薦手法に関する従来研究について説明する。最後に、2.5節で本章のまとめとする。

2.2 協調フィルタリングに基づく推薦手法に関する研究

協調フィルタリングは、ユーザの視聴履歴や購入履歴、評価スコアなどを分析し、類似した嗜好を持つ他のユーザを特定することで、それらのユーザの履歴内のアイテムを推薦対象として提示する手法である。協調フィルタリングに関する従来研究について以降にまとめる。

ItemKNN [23]

文献 [23] では、協調フィルタリングの初期の代表的手法である Item-based K-

Nearest Neighbor (ItemKNN) アルゴリズムを提案している。ItemKNNでは、ユーザの行動履歴から各アイテム間の類似度を計算し、対象ユーザが高く評価したアイテムに似たアイテムを推薦する。ItemKNNは、計算効率とスケーラビリティに優れることから、実用的な場面で広く採用されている。

MF [24]

文献 [24] では、行列分解 (Matrix Factorization; MF) によるモデルベース協調フィルタリングが提案された。ユーザ-アイテムの評価行列を潜在因子空間に埋め込み、それらの内積によりスコア予測を行うこの手法である。この研究は、協調フィルタリングにおける潜在特徴モデルの先駆けとなっている。

MFBPR [25]

文献 [25] では、Bayesian Personalized Ranking (BPR) 損失と行列分解に基づく推薦モデルである MFBPR が提案されている。MFBPR は、「ユーザが暗黙的に評価したアイテムの方が、未観測のアイテムより好む確率が高い」という仮定のもと、ユーザの埋め込みと暗黙的に評価したアイテムの埋め込みが類似するように学習を行うことにより、高精度な推薦を実現している。

NGCF [26]

文献 [26] では、Neural Graph Collaborative Filtering (NGCF) が提案された。NGCF は GCN に基づき、ユーザのアイテムに対するインタラクションを明示的に考慮することで、表現力の高い埋め込み特徴量を算出可能とし、高精度な推薦を実現している。

LightGCN [27]

文献 [27] では、Light Graph Convolutional Network (LightGCN) が提案された。LightGCN は、文献 [26] で提案された NGCF に加えて、グラフ上の埋め込み特徴量を平準化し、構造を単純化することにより、計算量の大幅な削減および推薦精度の向上を実現している。

BUIR [28]

文献 [28] では、One-Class Collaborative Filtering (OCCF) 問題において、負例のサンプリングを用いずにユーザ・アイテム表現を学習する Bootstrapping User and

Item Representations (BUIR) が提案されている。BUIR は 2 つのエンコーダ（オンライン・ターゲット）を用いたブートストラップ学習と自己教師ありデータ拡張により、データスパース性と負例不確実性の課題を克服し、高精度な推薦性能を実現している。

EGLN [29]

文献 [29] では、Enhanced Graph Learning Network (EGLN) が提案されており、ユーザ・アイテムの埋め込みとグラフ構造を相互に強化することで、より適応的かつ精緻な推薦を実現している。本手法では mutual information 最大化と local-global 一貫性最適化により、暗黙フィードバックの不完全性を克服し、信頼性の高い推薦グラフの構築を可能としている。

GTN [30]

文献 [30] では、ユーザ・アイテム間の相互作用の信頼性に着目し、適応的な伝播を行う Graph Trend Filtering Network (GTN) が提案された。GTN は、信頼性の低いクリックやノイズの多いフィードバックを抑制し、グラフ構造上の過度な平滑化を回避することで、ロバストかつ高精度な推薦性能を達成している。

協調フィルタリングはこのように、ユーザ間の集合知を活用することで高精度なアイテムの推薦が可能である。しかしながら、十分な履歴を持たない新規ユーザに対しては推薦が困難であるため、このようなコールドスタート問題の解決が急務となっている。

2.3 コンテンツベースフィルタリングに基づく推薦手法に関する研究

コンテンツベースフィルタリング (Content-Based Filtering) は、ユーザが過去に好んだアイテムの属性情報を基に、類似する特徴を持つ他のアイテムを推薦する手法である。ユーザの評価履歴からその嗜好を反映したプロファイルを作成し、推薦候補のアイテムとユーザプロファイルとの類似度に基づいて推薦を行う。コンテンツベースフィルタリングに関する従来研究について以降にまとめる。

TF-IDF ベース [31]

文献 [31] では、Web ページ推薦のための Fab システムが提案されている。このシステムでは、TF-IDF に基づく特徴抽出を用いてユーザとアイテムをベクトルで表現し、コサイン類似度により推薦を行うことで、個別の嗜好に対応可能な動的推薦を実現している。

ナイーブベイズ [32]

文献 [32] では、ニュース記事推薦のために Naive Bayes 分類器を用いた手法が提案された。ユーザが以前に好んだ記事の語彙的特徴を学習し、そのパターンに一致する新しい記事を推薦することで、教師あり学習に基づくコンテンツベース推薦の有効性が示されている。

タグベース [33]

文献 [33] では、HTML タグや構造情報を利用して Web ページ推薦を行うコンテンツベースフィルタリング手法が提案された。この手法は、Web ページの構造的特徴を活用することで、テキストだけでは捉えきれないコンテンツの特性を反映させている。

KNN 分類ベース [34]

文献 [34] では、k-近傍法 (KNN) を用いてユーザのプロファイルとアイテムの特徴ベクトルの類似性を評価し、分類ベースで推薦を行う手法が提案されている。この方法は単純であるが、適切な特徴設計によって高い精度を発揮することが報告されている。

CNN ベース [35]

文献 [35] では、畳み込みニューラルネットワーク (CNN) を用いてアイテムの画像・テキスト情報を特徴ベクトルに変換し、ユーザプロファイルと比較する手法が提案されている。この深層学習に基づくアプローチにより、非構造的コンテンツを含む推薦でも高い精度が実現されている。

コンテンツベースフィルタリングは特に新規アイテムに対するコールドスタート問題に強いという利点がある。しかしながら、ユーザが過去に評価したアイテムに似たものばかりが推薦される過特化 (overspecialization) 問題や、協調フィルタリングと比較して一般的に推薦精度が劣るという課題が存在する。

2.4 ハイブリッド推薦に関する研究

協調フィルタリングとコンテンツベースフィルタリングの課題を解決するための手法として両者を組み合わせたハイブリッド推薦手法が提案されている。ハイブリッド推薦手法は特徴統合に基づく手法と知識グラフに基づく手法が提案されている。以下、2.4.1にて特徴統合に基づく手法について説明し、2.4.2にて知識グラフに基づく手法について説明する。

2.4.1 特徴統合に基づく手法

特徴統合に基づく手法は、ユーザ、アイテムおよびアイテムの属性などを特徴量に変換し、それらの特徴量を統合的に活用して推薦を行う手法である。以下に特徴統合に基づく手法の従来研究をまとめる。

FM [36] 文献 [36] では、Factorization Machines (FM) が提案された。FM は、特徴量間のすべてのペアワイズな相互作用を潜在因子ベクトルの内積で表現することにより、高次元かつスパースな特徴空間でも高精度な予測が可能である。また、線形時間での学習が可能であり、様々な推薦タスクに柔軟に適用可能な汎用モデルとなっている。

LightFM [37] 文献 [37] では、ユーザおよびアイテムのメタデータを用いて埋め込みを学習する LightFM が提案された。本手法は、メタデータの線形結合によってユーザ・アイテムの表現を構成し、コールドスタート条件下でも高精度な推薦が可能である。また、コンテンツ情報と協調情報の双方を統合するハイブリッド手法である。

DeepFM [38] 文献 [38] では、Factorization Machines と Deep Neural Network を統合した DeepFM が提案された。DeepFM は、共通の入力から低次および高次の特徴相互作用を同時に学習可能であり、手動の特徴設計を不要とするエンドツーエンドの CTR 予測モデルである。

NFC [39] 文献 [39] では、コールドスタートアイテム推薦に向けた Neural Feature Combiner (NFC) が提案された。NFC は、アイテムの内容特徴を低次元のハイブリッド埋め込み空間にマッピングし、既存アイテムとの類似度を学習する深

層ニューラルネットワークであり，協調情報と内容情報を非線形に統合することで高精度な推薦を実現している．

これらの特徴統合に基づく手法は，多様な情報源を活用することにより高い推薦性能を実現可能である．しかしながら，これらの手法はアイテム間の意味的關係を直接捕捉することが困難である．

2.4.2 知識グラフに基づく手法

知識グラフに基づく手法は，知識グラフ内のパスを推論することにより推薦アイテムを決定する方法である．以下に知識グラフに基づく手法の従来研究をまとめる．

CKE [40] 文献 [40] では，Collaborative Knowledge Base Embedding (CKE) が提案された．CKE は，構造・テキスト・画像といった異種情報を知識ベースから抽出し，これらを協調フィルタリングと統合的に学習することで，表現力豊かな推薦モデルを実現している．

KGAT [41] 文献 [41] では，Knowledge Graph Attention Network (KGAT) が提案された．KGAT は，知識グラフとユーザ-アイテムグラフを統合し，高階関係を明示的にモデリングするための注意機構を導入することで，高精度かつ解釈可能な推薦を実現している．

CAFE [42] 文献 [42] では，Coarse-to-Fine な Neural Symbolic Reasoning を行うモデルである CAFE が提案された．本手法では，ユーザの履歴から行動パターンを抽出し，それをガイドとした知識グラフ上のパス推論により，推薦と説明性の両立を図っている．

UCPR [43] 文献 [43] では，User-Centric Path Reasoning (UCPR) が提案された．UCPR は，ユーザの需要ポートフォリオを多視点的に捉える構造を導入し，知識グラフ上の動的パス推論を通じて多様で説明性の高い推薦を可能にしている．

PGPR [44] 文献 [44] では，Policy-Guided Path Reasoning (PGPR) が提案された．本手法は強化学習を用いて知識グラフ上でパス探索を行い，推薦理由となる推論経路を明示的に提供できる点が特徴である．

RippleNet [45] 文献 [45] では, RippleNet が提案された. RippleNet は, ユーザ履歴から知識グラフ上に「波紋 (Ripple)」のように興味を伝播させることで, ユーザの潜在的な関心を探査し, 高精度な推薦と説明性を実現している.

CPER [46] 文献 [46] では, Counterfactual Path-based Explainable Recommendation (CPER) が提案された. CPER は, パス表現とトポロジ構造の両面から反実仮想推論を行い, 従来の不安定なアテンション重みの代替として, より信頼性の高い推薦根拠となる説明パスを抽出可能とする.

知識グラフに基づく手法は, 知識グラフを通じて意味的關係に基づく推薦が可能であることから特徴統合に基づく手法の課題を解決することが可能である. 一方で, これらの手法は, ユーザの目的や状況に応じた柔軟な推薦方針の制御が困難であるという課題が存在する. 知識グラフに基づく手法では, ユーザの目的や状況といった文脈情報を明示的に入力・反映する仕組みが限定的であり, それに応じた嗜好の変動を柔軟に推薦結果へ反映することが困難である. 例えば音楽ドメインでは, 同一ユーザにおいても, 時間帯や状況に応じて好まれる楽曲が異なるが, 従来手法ではこれらに起因する明示的な意図を推薦に反映することが難しい. さらに, コールドスタート状況における意味的推論能力の限界という課題も存在する. 実データに基づいて構築された知識グラフは, 推薦に有効な情報を十分に含んでいない可能性があるため, 意味的關係がスパースになりやすい. このスパース性により, 潜在的に関連したアイテム間の関係が限定されるため, 依然としてユーザの履歴が乏しいコールドスタート状況では, 嗜好に合致するアイテムとの意味的關係を把握・抽出することが難しく, 適切な推薦を行うことが困難となる. 本研究では, これらの2つの課題を解決するための手法の提案を行う.

2.5 まとめ

本章では, 推薦システムに関する代表的なアプローチである協調フィルタリング, コンテンツベースフィルタリング, およびそれらを統合したハイブリッド推薦に関する先行研究を概観した. リサーチマップと本研究の位置付けを図 2.1 に示す. 既存手法は推薦精度や解釈性の向上に寄与してきたが, ユーザの意図や状況を反映した柔軟な推薦制御, ならびにコールドスタート状況における意味的推論には

限界が存在することが明らかとなった．特に，知識グラフを用いた推薦では，高次な関係性のモデリングが可能である一方，推薦方針の制御性や関係のスパース性に起因する課題が依然として残されている．次章では，これらの課題を踏まえ，ユーザの明示的な意図を柔軟に反映可能な知識グラフ強化学習に基づくプレイリスト推薦手法を提案する．さらに第4章，第5章では，知識の補完によってコールドスタート問題を緩和するための手法について述べ，本研究の貢献の具体的な技術的枠組みを明らかにする．

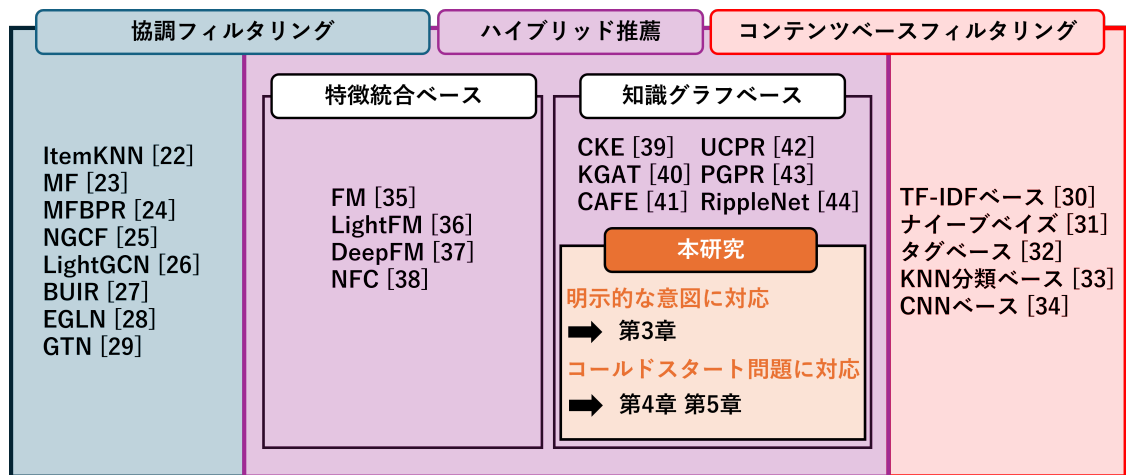


図 2.1: リサーチマップと本研究の位置付け.

第3章

知識グラフと強化学習に基づくプレイリスト生成

3.1 はじめに

本章では，知識グラフおよび強化学習に基づく制御可能なプレイリスト生成手法を提案する．提案手法は，対象ユーザの聴取履歴と明示的な入力に基づき，文脈的な意図や嗜好の変化に応じたプレイリスト生成を可能とすることを目的とする．従来のプレイリスト生成手法は，主にユーザの短期的な嗜好に基づいた単一楽曲の推薦に特化しており，リスト全体としての整合性や長期的な満足度には十分に対応できていない．さらに，聴取履歴のみに依存する方策の学習は，ユーザの嗜好探索が不十分となり，フィルターバブルの形成につながる懸念もある．これらの課題を解決するためには，スパースな履歴情報を補完し，ユーザの現在および将来的な意図に沿った推薦を可能とする柔軟な手法が求められる．提案手法では，まず，ユーザの過去の聴取履歴から短期・長期の嗜好を抽出し，さらに現在の関心や将来的な嗜好の変化を捉えるために，過去・現在・将来に関する嗜好を表す3種類の状態表現を導入する．これらの状態表現は，知識グラフを用いた楽曲間の意味的な関係に基づいて定義され，ユーザの嗜好を多面的に捉えることが可能である．次に，ユーザの明示的な入力（例：音響類似度，新規性，人気度などの重視）に応じて設計される報酬関数に基づき，強化学習によりプレイリスト生成方策を最適化する．具体的に，REINFORCE アルゴリズムを用いることにより，生成されたプレイリストと聴取履歴との楽曲特徴の類似度や，ユーザの意図を反映する誘

導的な評価指標に基づいた方策の学習を行う。以上により、提案手法は、ユーザのスパースな履歴情報と知識グラフ上の意味的関係を統合的に活用することで、短期的な嗜好予測の精度を維持しつつ、ユーザの現在の関心や将来的な意図を柔軟に反映したプレイリスト生成を実現する。

本章の構成は以下の通りである。3.2節では、プレイリスト生成タスクに関する基礎的事項を整理する。3.3節では、提案手法の詳細な設計とその各構成要素を説明する。3.4節では、提案手法の有効性を検証するための実験を行い、3.5節において、本章のまとめを述べる。

3.2 Preliminary

本節では、プレイリスト生成タスクを明確化するために、強化学習の概念を含む予備知識について述べる。はじめに、プレイリスト生成のプロセスを明確にするために、本タスクにおける記法を導入する。さらに、本タスクの定義および強化学習の中核概念であるマルコフ決定過程 (MDP) について説明する。

記法

プレイリスト生成のシナリオは、対象ユーザの聴取履歴およびその明示的な入力から構成される。ユーザ集合を $\mathcal{U}$ 、楽曲集合を $\mathcal{M}$ とする。各対象ユーザ u_v ($v = 1, 2, \dots, N_u$; N_u は対象ユーザ数) に対し、聴取履歴を $l_{1:N^{u_v}}^{u_v} = l_1^{u_v} \rightarrow l_2^{u_v} \rightarrow \dots \rightarrow l_{N^{u_v}}^{u_v}$ と表す (N^{u_v} は対象ユーザ u_v の聴取履歴に含まれる楽曲数)。特に、 $l_1^{u_v}$ は対象ユーザ u_v が最初に聴取した楽曲を表す。提案手法では、対象ユーザの聴取履歴に加えて、報酬関数のパラメータとして対象ユーザの明示的な入力 γ を利用する。さらに、各レコードは2つのエンティティとその関係からなる三つ組で構成される知識グラフ $\mathcal{G}$ をプレイリスト生成で用いる。ここで、楽曲集合 $\mathcal{M}$ はこの知識グラフ上に配置可能である。知識グラフに基づいて、楽曲に関連する知識情報 (例: 楽曲のアーティスト, アーティストのジャンル) を取得することが可能である。

タスク定義

プレイリスト生成には強化学習を用いる。上述の記法に基づき、モデルはプレイリストを生成し、それを対象ユーザに推薦する。提案手法の主な利点は、

対象ユーザの聴取履歴に基づいて嗜好を予測し、ユーザの明示的な入力に基づいて新しいタイプ(ジャンル, アーティスト, 音響特徴等)の楽曲への誘導を可能にする点にある.

マルコフ決定過程 (MDP)

MDPは五つ組 $\langle S, \mathcal{A}, T, R, \pi \rangle$ によって表される:

状態 S

S は状態の集合を表し, 各 $s \in S$ は環境におけるエージェントの情報状態を表す.

行動 $\mathcal{A}$

$\mathcal{A}$ は行動の集合を表し, 各 $a \in \mathcal{A}$ はエージェントが環境に対して取り得る行動を示す.

遷移関数 T

T は行動および現在の状態に基づいて状態を更新する遷移関数であり, すなわち $s_{t+1} = T(s_t, a_t)$ で表される.

報酬関数 R

R は報酬関数であり, 例えばエージェントが状態 s_t において行動 a_t をとったときに即時報酬 $R(s_t, a_t)$ を与える.

方策 π

π はエージェントの行動方策を表す. 一般に, これは可能な行動に対する確率分布としてモデル化される.

以上の五つの定義に基づき, 学習を繰り返すことによりエージェントの最適方策の獲得を目指す.

3.3 提案手法

本節では, 提案する制御可能なプレイリスト生成手法について詳述する. 提案手法の概要を図3.1に示す. 提案手法は, 対象ユーザの聴取履歴および明示的な入力に基づき, 強化学習アルゴリズムを用いてプレイリストを生成する.

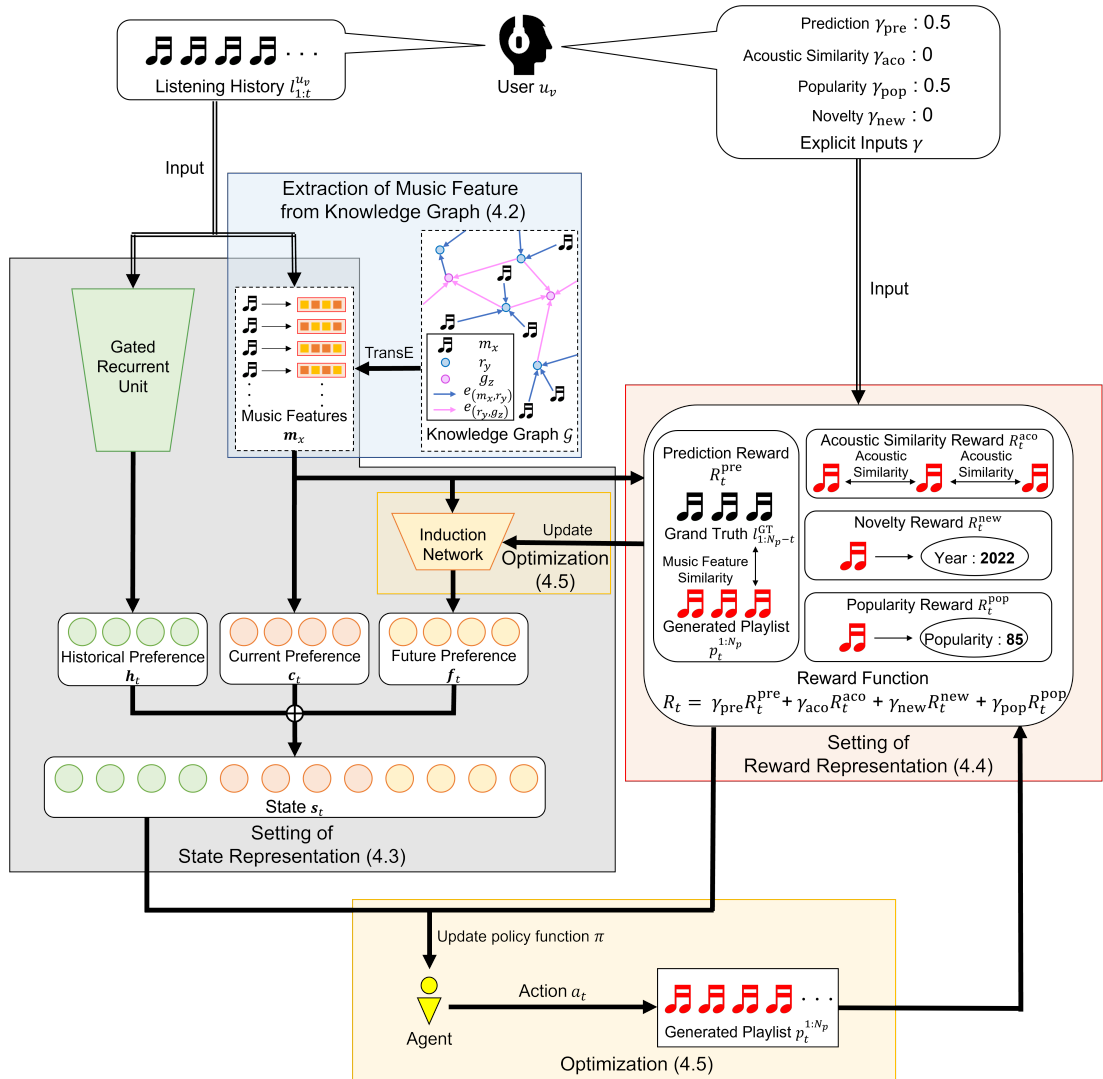


図 3.1: 提案するプレイリスト生成手法の概要図.

3.3.1 マルコフ決定過程の定式化

まず、最適な方策を獲得し洗練されたプレイリストを生成するために、MDPを用いる。MDPにおいては、エージェントは各タイムステップにおいて環境と相互作用し、状態を観測する。本タスクでは、環境の状態は対象ユーザの聴取履歴と、データベース中の楽曲の高次関係を記述する知識グラフの情報から構成される。タイムステップ t における状態は以下のように定義される：

$$s_t = [l_{1:t}^{u_v}, \mathcal{G}], \quad (3.1)$$

ここで、 $l_{1:t}^{u_v}$ ($1 \leq t \leq N^{u_v}$) は聴取履歴 $l_{1:N^{u_v}}^{u_v}$ における t 番目までの楽曲列、 $\mathcal{G}$ は知識グラフを表す。初期状態 s_0 は以下のように定義される：

$$s_0 = [\emptyset, \mathcal{G}]. \quad (3.2)$$

続いて、エージェントは行動 $a_t \in \mathcal{A}$ をとり、楽曲集合 $\mathcal{M}$ から楽曲を選択して、タイムステップ t における対象ユーザ u_v に対してプレイリスト $p_t^{1:N_p}$ (N_p はプレイリスト中の楽曲数) を生成する。特に、 p_t^1 はタイムステップ t に生成されたプレイリストの最初の楽曲を示す。エージェントの行動は、方策 $\pi(a_t|s_t)$ に基づいて決定される。本モデルでは、 $\pi(a_t|s_t)$ は各楽曲の確率を出力する softmax 関数として定義される。この方策は以下のように定義される：

$$\pi(a_t|s_t) = \frac{\exp\{p_t W s_t\}}{\sum_{m' \in \mathcal{M}} \exp\{m' W s_t\}}, \quad (3.3)$$

$$p_t = \sum_{i=1}^{N_p} \frac{p_t^i}{N_p}, \quad (3.4)$$

ここで、 m' および p_t^i はそれぞれ楽曲 m' および p_t^i に対応する楽曲特徴を表し、これらは後述する 3.3.2 項で得られるものである。 W は方策関数のパラメータ、 s_t は状態 s_t の埋め込み特徴を表す。

行動に応じて、エージェントは報酬 r_{t+1} を獲得する。この報酬は、対象ユーザの聴取履歴および明示的な入力に基づいて設計されている。さらに、次の状態への遷移を表す遷移関数 $T(s_t, a_t)$ は以下のように定義される：

$$T(s_t, a_t) = T([l_{1:t}^{u_v}, \mathcal{G}], p_t^{1:N_p}) \quad (3.5)$$

$$= s_{t+1}. \quad (3.6)$$

新たな状態 s_{t+1} は $[l_{1:t+1}^u, \mathcal{G}]$ として定義され、それに対応する埋め込み特徴 s_{t+1} が与えられる。以上の MDP の定式化に基づき、聴取履歴および明示的入力から方策を徐々に更新しつつ、洗練されたプレイリストを生成する。

3.3.2 知識グラフを用いた楽曲特徴の抽出

本項では、知識グラフを用いて楽曲特徴を算出する方法について述べる。これは後述される音響特徴とは異なる概念であることに注意されたい。知識グラフはノード間の高次の関係を表現可能な有向グラフである。提案手法では、楽曲およびそのメタデータを用いて知識グラフを構築する。ノード集合 $\mathcal{N}$ およびエッジ集合 $\mathcal{E}$ を有する知識グラフ $\mathcal{G}$ は、以下のように定義される： $\mathcal{G} = \{(n_{\text{head}}, e, n_{\text{tail}}) \mid n_{\text{head}}, n_{\text{tail}} \in \mathcal{N}, e \in \mathcal{E}\}$ 。ここで、 n_{head} および n_{tail} はそれぞれヘッドノードとテールノードを表し、 e は両ノードを接続するエッジを表す。

提案手法では、これらのノードは以下の3種類から構成される：楽曲 m_x ($x = 1, 2, \dots, N_m$; N_m は楽曲数)、アーティスト r_y ($y = 1, 2, \dots, N_r$; N_r はアーティスト数)、およびジャンル g_z ($z = 1, 2, \dots, N_g$; N_g はジャンル数) である。

さらに、以下の2種類の有向エッジを定義する：楽曲 m_x を制作したアーティスト r_y に接続するエッジ $e_{(r_y, m_x)}$ 、およびアーティスト r_y が属するジャンル g_z に接続するエッジ $e_{(m_x, g_z)}$ である。このような有向グラフを用いることで、異種エンティティ間の階層的な関係を把握することが可能となる。

加えて、各ノードおよびエッジの埋め込み特徴を算出するために、代表的な知識グラフ埋め込み手法である TransE [47] を用いる。TransE では、ノードおよびエッジの特徴が以下の関係を満たすように学習される：

$$n_{\text{head}} + e = n_{\text{tail}}, \quad (3.7)$$

ここで、 n_{head} 、 e 、および n_{tail} はそれぞれ n_{head} 、 e 、 n_{tail} の埋め込み特徴を表す。

本研究では、楽曲 m_x 、アーティスト r_y 、ジャンル g_z 、および対応するエッジ $e_{(r_y, m_x)}$ 、 $e_{(m_x, g_z)}$ の埋め込み特徴をそれぞれ m_x 、 r_y 、 g_z 、 $e_{(r_y, m_x)}$ 、 $e_{(m_x, g_z)}$ と定義する。特に、 m_x を各楽曲に対応する楽曲特徴と見なす。このようにして、楽曲とそのメタデータとの高次関係を反映した楽曲特徴を獲得することが可能となる。

3.3.3 状態表現の設定

適切な状態表現を設計することは、強化学習アルゴリズムにおける重要な要素である。提案手法では、知識グラフからの知識情報を取り込んで良好な状態表現を学習するために、過去嗜好に基づく表現、現在嗜好に基づく表現、将来嗜好に基づく表現の3種類の状態表現を設定する。これにより、知識グラフから得られる豊富な情報を活用し、ユーザの嗜好を理解する強化学習アルゴリズムを構築可能となる。

過去嗜好に基づく状態表現

まず、対象ユーザの聴取履歴をエンコードするために、標準的な再帰型ニューラルネットワーク（RNN）モデルである Gated Recurrent Unit（GRU）[48]を用いる。この表現の目的は、ユーザの過去の聴取嗜好の特徴を捉えることであり、知識グラフから得られる情報は使用しない。過去嗜好に基づく状態表現 $\mathbf{h}_t$ は以下のように定義される：

$$\mathbf{h}_t = \text{GRU}(\mathbf{h}_{t-1}, \mathbf{l}_{1:t}^{\mu_v}; \Phi_{\text{gru}}) \quad (3.8)$$

ここで、 Φ_{gru} は GRU に関連する全パラメータを表す。

現在嗜好に基づく状態表現

次に、対象ユーザの聴取履歴および 3.3.2 項で算出された楽曲特徴を用いる。この表現の目的は、知識グラフの豊富な情報からユーザが好む楽曲内容の特徴を抽出することである。先行研究[49]に基づき、対象ユーザが聴取した楽曲の知識グラフ特徴を平均プーリングにより集約する。現在嗜好に基づく状態表現は以下のように定義される：

$$\mathbf{c}_t = \sum_{j=1}^t \frac{\mathbf{l}_j^{\mu_v}}{t}, \quad (3.9)$$

ここで、 $\mathbf{l}_j^{\mu_v}$ は $\mathbf{l}_j^{\mu_v}$ に対応する楽曲特徴を表す。なお、この表現は時間的な情報や注意機構を考慮していないが、この単純な表現と比較して有意な性能向上が得られないため、それらは導入していない。

将来嗜好に基づく状態表現

さらに、将来的な時点における対象ユーザの関心の変化を捉えるために、将来嗜好を導入する。提案手法では、現在嗜好に基づく表現を入力として、将来の嗜好を予測する誘導ネットワーク（Induction Network）を構築し、洗練されたプレイリストを生成する鍵とする。本モデルにおける誘導ネットワークは、多層パーセプトロン（MLP）により構成される。現在嗜好に基づく表現 c_t を入力とし、将来嗜好に基づく状態表現を予測する。将来嗜好に基づく状態表現は以下のように定義される：

$$f_t = \text{IDN}(c_t; \Phi_{\text{idn}}), \quad (3.10)$$

ここで、 $\text{IDN}(\cdot)$ は MLP に基づいて構築された誘導ネットワークを表し、 Φ_{idn} はそのパラメータである。対象ユーザの楽曲内容に対する嗜好は、連続したタイムステップ間で大きく変動しないと仮定している。したがって、本モデルでは現在の嗜好に基づいて将来の嗜好を予測することを目的とする。知識グラフから将来嗜好を学習することは、合理的かつ有益な探索を実現する上で強化学習アルゴリズムにとって重要である。MLP を用いることで、誘導ネットワークが対象ユーザの嗜好の発展過程を把握することが期待される。

最終的な状態表現

以上の議論に基づき、最終的な状態表現を定義する。状態 s_t に対して、その表現 s_t は 3 種類の状態表現を結合することで構成される：

$$s_t = h_t \oplus c_t \oplus f_t, \quad (3.11)$$

ここで、 $\oplus$ はベクトルの連結を示す。本モデルは、対象ユーザの過去・現在・将来の嗜好を捉えることにより、対象ユーザが好むと予測される楽曲を含むプレイリストの生成を可能とする。

3.3.4 報酬関数の設定

本項では、報酬関数の設定について述べる。提案手法における主な目的は、対象ユーザの聴取履歴に基づいて嗜好を予測すること、および対象ユーザの明示的

な入力に基づいて新しいタイプの楽曲への誘導を行うことである．これらの目的を同時に達成するために，2種類の報酬を導入する．1つ目は，生成されたプレイリストと正解となる聴取履歴との類似度に基づく予測報酬である．2つ目は，音響特徴の類似性，年代，人気度に対する対象ユーザの特定の嗜好を反映する誘導報酬である．対象ユーザが明示的な入力によりこれら2つの報酬のバランスを制御可能にすることで，個々のユーザの嗜好にパーソナライズされたプレイリスト生成を実現する．以下に，各報酬の内容を説明する．

予測報酬

予測報酬では，生成されたプレイリストの楽曲特徴の質を評価することを目的とする．正解となる聴取履歴（Ground Truth: GT）と生成されたプレイリストとの完全一致は，楽曲を正確に予測する上で有効な手法である．しかし，GTとなる聴取履歴はデータベース全体に対して疎であるため，完全一致に基づく学習は困難である．この疎な問題を回避し，効率的な学習を促進するための合理的な解決策として，3.3.2項で得られた楽曲特徴の類似性を利用する方法が挙げられる．楽曲特徴の類似性は疎ではなく，また楽曲のメタデータに関する高次関係を表現可能であるため，近似的な一致を実現可能である．そこで，GTの聴取履歴と生成されたプレイリストの楽曲特徴間のコサイン類似度を予測報酬として導入する．これにより，疎な問題を解消しつつ，効率的な学習を可能とする．

GTとなる聴取履歴および生成されたプレイリストをそれぞれ $I_{1:N_p-t}^{\text{GT}}$, $p_t^{1:N_p}$ とし，単純平均法により各楽曲特徴を集約することで， c_t^{GT} および $p_t^{1:N_p}$ と定義する．このとき，予測報酬 R_t^{pre} は以下のように定義される：

$$R_t^{\text{pre}} = \frac{c_t^{\text{GT}} \cdot p_t^{1:N_p \top}}{\|c_t^{\text{GT}}\| \cdot \|p_t^{1:N_p \top}\|}. \quad (3.12)$$

予測報酬 R_t^{pre} は，楽曲の予測性能に大きな影響を与えるため，将来状態 f_t を更新する誘導ネットワークの学習にも用いられる．誘導ネットワークの学習については，3.3.5項で詳述する．

誘導報酬

誘導報酬は、対象ユーザの特有の嗜好を満たす新しいタイプの楽曲への誘導を目的として設計されている。この報酬は、音響類似度、人気度、新規性の3つの指標から構成される。対象ユーザの特有の嗜好は、時間の経過や文脈の変化に伴って変動すると想定され、聴取履歴だけでは完全に把握することが困難である。このような時間的な嗜好の変化を捉えるために、ユーザから明示的な入力として報酬関数のパラメータを受け取り、各報酬のバランスに反映させる。

音響類似度報酬

一般に、楽曲間の滑らかな遷移をもつプレイリストは、ユーザの満足度を高める上で効果的であるとされている [50,51]。提案手法では、楽曲間の音響特徴の類似性に基づく報酬を設計し、スムーズなプレイリストを好むユーザに対応する。具体的には、生成されたプレイリスト中の i 番目と $i+1$ 番目の楽曲の音響特徴ベクトルのコサイン類似度に基づき、音響類似度報酬 R_t^{aco} を以下のように定義する：

$$R_t^{\text{aco}} = \frac{1}{N_p} \sum_{i=0}^{N_p-1} \frac{\mathbf{p}_{t,\text{aco}}^i \cdot \mathbf{p}_{t,\text{aco}}^{i+1 \top}}{\|\mathbf{p}_{t,\text{aco}}^i\| \cdot \|\mathbf{p}_{t,\text{aco}}^{i+1 \top}\|}, \quad (3.13)$$

ここで、 $\mathbf{p}_{t,\text{aco}}^i$ は i 番目の楽曲 p_i^i の音響特徴を表す。

人気度報酬

音楽に不慣れなユーザや慎重な傾向を持つユーザは、楽曲の人気度に基づいて聴取することが多いと想定される。実際、音楽ストリーミングサービスの多くは、ユーザが自身の嗜好を明確に把握していない場合でも、人気楽曲を推薦することで支持を得ている。このように、人気楽曲はユーザの注目を集める可能性が高い。したがって、人気楽曲を好むユーザに対応するために、Spotify API (<https://developer.spotify.com/documentation/web-api/>, アクセス日：2022年5月12日) から取得される人気度の値を用いて人気度報酬 R_t^{pop} を以下のように定義する：

$$R_t^{\text{pop}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{\text{Popularity}(p_i^i)}{\text{Popularity}(p_{\text{highest}})}, \quad (3.14)$$

ここで、 $\text{Popularity}(\cdot)$ は Spotify API から取得される楽曲の人気度を返す関数、 p_{highest} はデータベース中で最も人気度の高い楽曲を表す。

新規性報酬

音楽に詳しいユーザの中には、新しい楽曲に注目する傾向を持つ者も多い。特に Shih ら [52] は、プレイリスト内における楽曲の新規性を考慮することの重要性を指摘している。提案手法では、プレイリストにより多くの新しい楽曲を含めることを目的として、楽曲のリリース年に基づいた報酬を設計する。新規性報酬 R_t^{new} は以下のように定義される：

$$R_t^{\text{new}} = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{\text{Year}(p_t^i) - \text{Year}(m_{\text{oldest}})}{\text{Year}(m_{\text{newest}}) - \text{Year}(m_{\text{oldest}})}, \quad (3.15)$$

ここで、 $\text{Year}(\cdot)$ は楽曲のリリース年を返す関数、 m_{oldest} および m_{newest} はデータベース内の最古および最新の楽曲をそれぞれ表す。

最終的な報酬関数

以上の議論に基づき、最終的な報酬関数を定義する。対象ユーザからの明示的な入力 γ に基づき、各報酬の影響度のバランスを制御可能であることは、提案手法の重要な特徴の1つである。明示的な入力 γ は、各報酬に対する重み係数から構成され、以下のように表される：

$$\gamma = \{\gamma_{\text{pre}}, \gamma_{\text{aco}}, \gamma_{\text{pop}}, \gamma_{\text{new}}\}, \quad (3.16)$$

ここで、 γ_{pre} , γ_{aco} , γ_{pop} , γ_{new} はそれぞれ R_t^{pre} , R_t^{aco} , R_t^{pop} , R_t^{new} に対応する係数パラメータを表す。

ただし、上記の報酬は分布が異なるため、単純な線形結合によって等価に比較することはできない。各報酬を等価に比較可能とするために、経験分布関数 [53] を導入する。経験分布関数は統計学における分布関数の1つであり、各データ点に対して $1/n$ ずつ増加するステップ関数として定義される。これを用いることで、異なる分散や外れ値に対して頑健な報酬分布 R_t^{pre} , R_t^{aco} , R_t^{pop} , R_t^{new} を構成する。

最終的に、これらの正規化された報酬分布に基づき、最終報酬関数を以下のよう
に定義する：

$$R_t = \gamma_{\text{pre}} R_t^{\text{pre}} + \gamma_{\text{aco}} R_t^{\text{aco}} + \gamma_{\text{pop}} R_t^{\text{pop}} + \gamma_{\text{new}} R_t^{\text{new}}. \quad (3.17)$$

3.3.5 最適化

本項では、提案手法における最適化手順について説明する．定義された状態，行動，報酬に基づいて最適な戦略を獲得するために，方策勾配法の標準的アルゴリズムである REINFORCE [54] を用いてマルコフ決定過程 (MDP) の学習を行った．さらに，楽曲予測性能の向上を目的として，将来状態を獲得する誘導ネットワークも同時に学習する．以下に詳細を示す．

方策勾配による学習

プレイリスト生成タスクにおいては，全対象ユーザに対して期待累積報酬 $J(\Theta)$ を最大化する確率の方策 π を獲得することを目的とする． $J(\Theta)$ の勾配は以下のよう
に表される：

$$\nabla J(\Theta) = E_{\pi} \left[\sum_{\mathcal{U}} \sum_{t'=t}^T \gamma_{\pi}^{t'-t} R_{t'} \frac{\nabla \pi(a_t | s_t; \Theta)}{\pi(a_t | s_t; \Theta)} \right], \quad (3.18)$$

ここで， $E_{\pi}[\cdot]$ はエージェントが方策 π に従って行動する際の期待値， γ_{π} は割引率， Θ は学習対象となる全パラメータ， T は1エピソードにおける終端タイムステップを表す．提案手法では，REINFORCE [54] に基づき，方策関数のパラメータを学習する．

エージェントは各タイムステップにおいて D 個のプレイリストを生成する．具体的には，各状態 s_t に対して，方策関数（式 (3)）に従い， N_p 個の楽曲からなる D 個のプレイリスト $p^{1:N_p}(d)$ ($1 \leq d \leq D$) をサンプリングする．生成されたプレイリスト $p^{1:N_p}(d)$ に対する学習更新式は以下のとおりである：

$$\nabla \Theta = \sum_{t'=t}^T \gamma_{\pi}^{t'-t} R_{t'} \frac{\nabla \pi(p_t^{1:N_p}(d) | s_t; \Theta)}{\pi(p_t^{1:N_p}(d) | s_t; \Theta)}. \quad (3.19)$$

より良い累積報酬を推定するために，上記のプロセスを D 回繰り返すことで学習を行う．

誘導ネットワークの学習

対象ユーザの嗜好を予測するうえで提案手法の鍵となる構成要素は、式 (9) に示した誘導ネットワークである。このニューラルネットワークの学習において、単純な解法としては平均二乗誤差 (MSE) などの回帰損失を適用する方法が考えられる。しかし、GT となる聴取履歴が疎であるため、単純な回帰損失では誘導ネットワークを効率的に学習させることが難しい。そこで、先行研究 [49] に着想を得て、ペアワイズランキング戦略を用いて誘導ネットワークを学習する。

生成されたプレイリストに対して、式 (9) に従い、楽曲特徴に基づく将来嗜好表現 $f_t^{(1)}, f_t^{(2)}, \dots, f_t^{(D)}$ を取得可能である。誘導ネットワークの学習目標は、より正確な将来嗜好表現を獲得することである。この目標を達成するために、ペアワイズ比較による追加制約を導入する。具体的には、 $f_t^{(d)}$ と $f_t^{(d')}$ が与えられたとき、報酬に基づいて D 個のプレイリストの優先順位を決定し、それに従って誘導ネットワークに対して制約を加える。すなわち、 $R_{\text{pre}}(l_{t+1:N_p}^{\mu}, p_t^{1:N_p}(d)) > R_{\text{pre}}(l_{t+1:N_p}^{\mu}, p_t^{1:N_p}(d'))$ が成立する場合には、 $\text{MLP}(f_t^{(d)}) \succ \text{MLP}(f_t^{(d')})$ となるよう学習を行う。

最終的に、学習済みの誘導ネットワークおよび方策関数に基づいて、対象ユーザに対して N_p 個の楽曲からなるプレイリストを生成する。テストフェーズにおいては、学習済みの誘導ネットワークにより、対象ユーザの聴取履歴および明示的な入力に基づく状態と報酬を獲得する。そのうえで、学習済みの方策関数が、対象ユーザに対するプレイリストを出力する。

3.4 実験

本節では、提案手法の有効性を確認するために実施した実験について述べる。本実験では、従来のプレイリスト生成および音楽推薦手法と比較し、提案手法の有効性を以下の2つの観点から評価する：(1) 対象ユーザの聴取履歴に基づいて嗜好を予測可能であるか、(2) 対象ユーザの明示的な入力に基づいて新しいタイプの楽曲に適切に誘導可能であるか。以降、3.4.1 項に実験設定の詳細、3.4.2 項に結果と考察をそれぞれ示す。

3.4.1 実験設定

実験では、Spotify Million Playlist Dataset [55] に含まれる 1006 名のユーザの聴取履歴と、Spotify API から取得した楽曲のメタデータを使用した。このデータセットには、57,880 曲の楽曲、1006 名のユーザ、14,973 人のアーティスト、および 2,517 種類のジャンルが含まれている。各アーティストは少なくとも 1 つのジャンルに属している。TransE によって得られる楽曲特徴ベクトル $\mathbf{m}_i$ の次元数は 50 に設定した。楽曲の音響特徴としては、Spotify API が提供する danceability, energy, speechiness, acousticness, valence, instrumentalness, liveness, tempo の 8 次元ベクトルを用いた。楽曲の人気度は 0 から 100 の範囲で与えられており、本データセットでは最も人気の高い楽曲 p_{highest} の人気度は 90、最も古い楽曲 m_{oldest} のリリース年は 1935 年、最も新しい楽曲 m_{newest} のリリース年は 2022 年である。学習データおよびテストデータは、各プレイリストに含まれる楽曲のうち、先頭 90% を学習データ、残り 10% をテストデータとした。生成されるプレイリストの長さ（楽曲数）は 10 に設定した。

また、提案手法のモデルの学習には、学習率 0.001、バッチサイズ 1,024 の設定で REINFORCE アルゴリズムを適用した。学習はエピソード単位で行い、検証データにおける予測報酬 R_t^{pre} の移動平均が 10 エピソード連続で増加しない場合に早期終了とした。また、学習の最大エピソード数は 1000 と設定した。誘導ネットワークの学習においても同様に、ペアワイズランキング損失が 10 エピソード連続で改善しない場合に学習を終了した。テストフェーズでは、学習済みの誘導ネットワークと方策関数を用いて、対象ユーザの聴取履歴および明示的入力 γ に基づいて状態 s_t を構築し、方策 π に従って 1 つのプレイリスト（10 曲）を生成した。生成されたプレイリストに対して、各評価指標を算出し、モデル性能の定量的評価を行った。

提案手法（PM）の性能を検証するため、以下の 4 つのプレイリスト生成手法と比較を行った。これらは基本的な RNN モデル、グラフ探索ベースの手法、および知識グラフ・強化学習ベースの state-of-the-art 推薦モデルである。

CM1:

GRU [48] に基づき、対象ユーザの聴取履歴のみを用いて学習された手法。

CM2 [51]:

楽曲の音響類似度に基づいて構築されたグラフの探索によりプレイリストを生成する手法.

CM3 [56]:

対象ユーザの聴取履歴と楽曲の音響特徴を用いた知識グラフを構築し、深層強化学習により推薦を行う音楽推薦手法.

CM4 [44]:

強化学習による知識グラフ推論を活用し、推薦の説明可能性を重視したアイテム推薦手法.

プレイリスト生成は、音楽推薦の特殊なケースと見なせる. したがって、知識グラフ・強化学習ベースの state-of-the-art の音楽推薦モデルである [51], およびその前身であるアイテム推薦モデル [44] を比較手法として採用した. なお, CM4 は連続する楽曲の音響特徴が類似するようなプレイリストを生成するが, 対象ユーザの聴取履歴に基づく予測性能は重視していない点に留意されたい.

PM と CM の性能を正当に比較するためには, 大規模なオンライン評価を実施し, 対象ユーザ自身がプレイリストの品質を評価することが理想的である. しかし, このようなオンライン評価は複雑なインフラを必要とするため, 本研究の範囲を超える. 代替手段として, [52] における既存手法と同様に, 実ユーザに基づいた標準的なオフライン評価を実施した.

本評価では, 様々な背景や嗜好を持つユーザを想定し, 報酬関数 R_t のパラメータを変更した7種類の PM ベースのモデルを用いた. それらのモデル名と各報酬パラメータの値を表 3.1 に示す. 例えば, PM-A は音響類似度を重視するユーザ, PM-PN は新規性と人気度を重視するユーザ, PM-ALL は音響類似度, 新規性, 人気度のすべてを重視するユーザを想定している.

なお, パラメータ γ はユーザが任意に設定可能であるため, 実際の運用においてはユーザ自身が任意の γ 値を指定することで個別のプレイリスト生成が可能である. また, γ_{pre} を 0 に設定すると誘導ネットワークが R_{pre} の空値を受け取ってしまい, 状態更新が不可能になるため, 常に 0 より大きい値に設定している.

表 3.1: PM ベースのモデルと報酬関数 R_t における各パラメータ値.

モデル名	γ_{pre}	γ_{aco}	γ_{pop}	γ_{new}
PM-A	0.5	0.5	0	0
PM-P	0.5	0	0.5	0
PM-N	0.5	0	0	0.5
PM-AP	0.33	0.33	0.33	0
PM-AN	0.33	0.33	0	0.33
PM-PN	0.33	0	0.33	0.33
PM-ALL	0.25	0.25	0.25	0.25

評価指標として, (1) の観点 (嗜好予測性能) に対しては, プレイリスト生成の予測性能に広く用いられる指標である $\text{nDCG}@k$ および $\text{Hit Rate}@k$ ($k = 1, 5, 10$) を用いた [57, 58]. 一方, (2) の観点 (誘導性能) を客観的に評価するため, それぞれの報酬 R'_{aco} , R'_{pop} , R'_{new} に対応する以下の 3 つの評価指標を用いた.

音響類似度指標 M_{aco}

生成されたプレイリスト内の連続する楽曲間の音響特徴の類似性を測定するため, 以下のように定義する:

$$M_{\text{aco}} = \sum_{\mathcal{U}} \frac{1}{N_{u_v}^p} \sum_{i=0}^{N_{u_v}^p - 1} \frac{\mathbf{p}_{u_v, \text{aco}}^i \cdot \mathbf{p}_{u_v, \text{aco}}^{i+1}{}^{\top}}{\|\mathbf{p}_{u_v, \text{aco}}^i\| \cdot \|\mathbf{p}_{u_v, \text{aco}}^{i+1}\|}, \quad (3.20)$$

ここで, $N_{u_v}^p$ はユーザ u_v に対して生成されたプレイリストの楽曲数, $\mathbf{p}_{u_v, \text{aco}}^i$ は i 番目の楽曲の音響特徴ベクトルである.

人気度指標 M_{pop}

生成されたプレイリストに人気楽曲が多く含まれているかを評価するため, 楽曲の人気度の平均値を以下のように定義する:

$$M_{\text{pop}} = \sum_{\mathcal{U}} \frac{1}{N_{u_v}^p} \sum_{i=1}^{N_{u_v}^p} \text{Popularity}(p_{u_v}^i), \quad (3.21)$$

ここで, $p_{u_v}^i$ はユーザ u_v のプレイリスト内の i 番目の楽曲を表す.

新規性指標 M_{new}

生成されたプレイリストに新しい楽曲が含まれているかを評価するため、リリース年の平均を用いて以下のように定義する：

$$M_{\text{new}} = \sum_{\mathcal{U}} \frac{1}{N_{u_v}^p} \sum_{i=1}^{N_{u_v}^p} \text{Year}(p_{u_v}^i). \quad (3.22)$$

3.4.2 結果と考察

本項では、実験結果を提示し、その考察を行う。まず、PM ベースのモデルおよび各 CM における nDCG@ k および Hit Rate@ k の値を表 3.2 に示す。すべての k の値において、PM ベースのすべてのモデルが、すべての CM よりも高い nDCG@ k および Hit Rate@ k を達成していることが確認される。

PM は CM1 を上回っており、これは短期的な嗜好のみならず長期的な嗜好まで分析可能となる強化学習の導入が効果的であることを示している。CM3 および CM4 は、対象ユーザの聴取履歴を系列としてではなく知識グラフの構成要素として用いているが、PM では聴取履歴を系列として GRU および強化学習により解析し、短期・長期の嗜好を共に捉えることが可能である。CM2 の予測性能が低いのは、楽曲の音響特徴の類似性のみを考慮しているためである。

次に、PM ベースのモデル同士の比較を行う。nDCG@1, nDCG@10, Hit Rate@ k ($k=1, 5, 10$) において最も高い値を示したのは PM-P であり、nDCG@5 において最高値を示したのは PM-N であった。これらのモデルの予測重みパラメータ γ_{pre} は 0.5 と、他のモデルよりも高く設定されている。これは、知識グラフから得られる楽曲特徴の類似性が予測性能の向上に寄与していることを示唆する。以上の結果から、PM は強化学習と知識グラフを用いることで、対象ユーザの嗜好を正確に予測可能であることが確認された。

表 3.2: nDCG@ k および Hit Rate@ k ($k = 1, 5, 10$) の結果. 各評価指標における最大値を太字で示す.

	nDCG@ k [$\times 10^{-2}$]			Hit Rate@ k [%]		
	$k = 1$	$k = 5$	$k = 10$	$k = 1$	$k = 5$	$k = 10$
PM-A	18.2	24.7	26.2	18.2	30.9	37.9
PM-P	19.3	25.6	27.9	19.3	33.4	40.4
PM-N	18.2	26.1	26.8	18.2	32.9	39.6
PM-AP	17.5	23.5	24.8	17.5	29.8	36.8
PM-AN	17.0	24.9	24.5	17.0	27.2	34.2
PM-PN	18.7	22.1	24.6	18.7	30.5	38.1
PM-ALL	18.0	23.0	24.1	18.0	29.0	36.0
CM1	10.2	13.9	17.2	10.2	21.8	28.8
CM2	0.00	0.52	0.40	0.00	0.43	1.03
CM3	3.42	5.30	6.22	3.42	6.71	11.5
CM4	3.59	4.30	5.94	3.59	7.05	12.0

表 3.3 に, PM ベースのモデルおよび各 CM における M_{aco} , M_{pop} , M_{new} の結果を示す. M_{pop} および M_{new} においては, それぞれ PM-P および PM-N が最高値, PM-AP および PM-PN が次点の値を示している. これらの結果は, 人気かつ新しい楽曲を求める対象ユーザに対して, それに対応するプレイリストが生成可能であることを示している.

M_{aco} において最高値を示したのは CM2 であった. これは, CM2 が連続する楽曲の音響類似度の最大化のみに特化した手法であるため, 自明な結果といえる. 一方, CM2 は予測性能が極めて低く, M_{pop} および M_{new} の制御も不可能である. M_{aco} の第 2 位および第 3 位のモデルは, それぞれ PM-A および PM-AN であり, これは音響類似度を重視するユーザの嗜好にも対応可能であることを示している.

表 3.3: M_{aco} , M_{pop} , M_{new} の結果. 太字は最大値, 下線は2番目に高い値を示す.

	M_{aco}	M_{pop}	M_{new}
PM-A	<u>0.90</u>	19.0	2001.9
PM-P	0.57	52.8	2012.4
PM-N	0.64	37.6	2013.4
PM-AP	0.81	<u>50.6</u>	2006.6
PM-AN	0.86	27.2	2011.4
PM-PN	0.65	48.0	<u>2012.8</u>
PM-ALL	0.73	42.0	2009.8
CM1	0.56	41.2	2005.4
CM2	0.97	19.3	2004.2
CM3	0.65	38.6	2006.6
CM4	0.60	40.7	2007.0

3.5 結論

本章では, 知識グラフおよび強化学習に基づく制御可能なプレイリスト生成手法を提案した. 提案手法は, 対象ユーザの聴取履歴に基づく嗜好予測に加え, 明示的な入力に基づき新たなタイプの楽曲への誘導を実現可能である. 実験により, 提案手法は既存のプレイリスト生成手法よりも高い性能を示すことが確認された.

第4章

音響特徴を導入した知識グラフを用いた楽曲推薦

4.1 はじめに

本章では、音響特徴を導入した知識グラフと深層強化学習に基づく楽曲推薦手法を提案する。知識グラフを活用した推薦手法は、ユーザとアイテム、およびそれらの属性間の関係を構造的に捉えられる点から、高次の意味的推論を通じた説明可能な推薦を可能とする有力なアプローチである。しかしながら、推薦精度はユーザの履歴情報の密度に大きく依存するため、履歴の少ないユーザに対しては依然としてコールドスタート問題が生じ、意味的なつながりを十分に活用した推薦が困難であるという課題が残されている。本章ではこの課題に対処するため、楽曲が持つ音響特徴に基づいて類似性の高い楽曲同士を結ぶエッジを知識グラフに追加することで、ユーザ履歴が不十分な場合でも意味的なつながりを補完できるグラフ構造を実現する。さらに、この高密度化された知識グラフを探索空間とみなし、深層強化学習により最適な推薦経路を学習することで、コールドスタート環境下でもユーザ嗜好に合致した楽曲の推薦を可能とする手法を提案する。提案手法では、TransEに基づく知識グラフ埋め込みを用いて推薦経路の特徴をベクトル空間上に表現し、その類似性に基づく明示的な報酬設計により、推薦理由の説明性と推薦精度の両立を図っている。また、推薦フェーズにおいては、学習済みの方策に基づくビームサーチを用いて、効率的かつ妥当性の高い推薦経路を選択する。

本章の構成は以下の通りである．4.2節では，提案手法の詳細について説明する．4.3節では，提案手法の有効性を検証するための実験を行い，4.4節において，本章のまとめを述べる．

4.2 提案手法

本節では提案手法の詳細について説明する．図4.1に，提案手法の概要図を示す．本手法は，知識グラフの構築，知識グラフ上でのエージェントの学習，および学習済み方策に基づく楽曲推薦の3つのフェーズから構成される．

第1フェーズでは，ユーザの聴取履歴および楽曲のアーティスト情報に加え，楽曲の音響特徴を用いて知識グラフを構築する．第2フェーズでは，構築した知識グラフを対象に，深層強化学習アルゴリズムを用いてエージェントの学習を行う．第3フェーズでは，獲得されたエージェント方策とビームサーチアルゴリズムに基づき，推薦する楽曲を決定する．

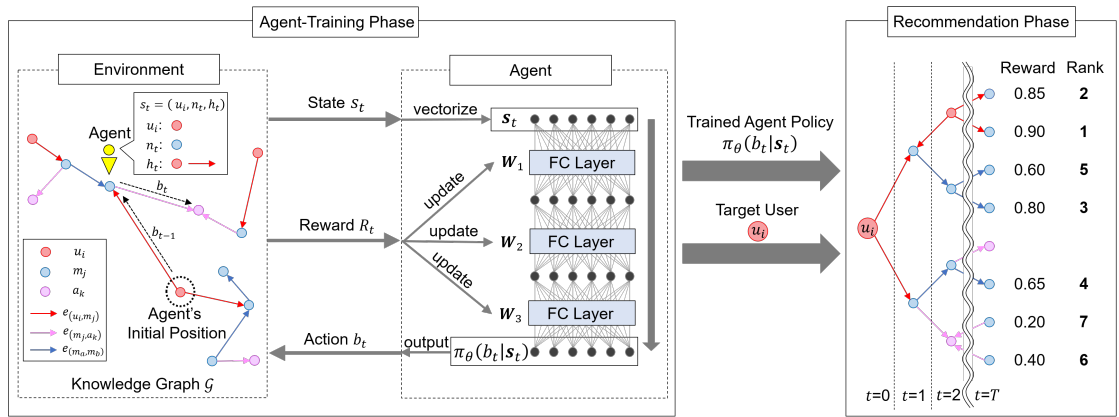


図 4.1: 提案する楽曲推薦手法の概要図.

4.2.1 知識グラフの構築

まず、提案手法に新たに導入される知識グラフ $\mathcal{G}$ を構築する．知識グラフにおけるノードは、ユーザ、アーティスト、および楽曲に対応し、エッジは聴取履歴、アーティスト情報、および音響特徴の類似度に基づいて定義される．ノード集合 $\mathcal{N}$ およびエッジ集合 $\mathcal{E}$ から構成される知識グラフ $\mathcal{G}$ は、以下のように定義される：

$$\mathcal{G} = \{(n_{\text{head}}, e, n_{\text{tail}}) \mid n_{\text{head}}, n_{\text{tail}} \in \mathcal{N}, e \in \mathcal{E}\},$$

ここで、 n_{head} は始点ノード、 n_{tail} は終点ノード、 e は両者を結ぶエッジである．

本手法では、ノードはユーザ u_i ($i = 1, 2, \dots, N_u$; N_u はユーザ数)、楽曲 m_j ($j = 1, 2, \dots, N_m$; N_m は楽曲数)、アーティスト a_k ($k = 1, 2, \dots, N_a$; N_a はアーティスト数) から構成される．次に、ユーザ u_i が過去に聴取した楽曲 m_j を接続する有向エッジ $e_{(u_i, m_j)}$ を定義する．さらに、アーティスト a_k が制作した楽曲 m_j を接続する有向エッジ $e_{(m_j, a_k)}$ を定義する．このように、有向グラフを用いることで、異種エンティティ間の階層的関係性を理解することが可能となる．例えば、ユーザ u_i が楽曲 m_j を聴取したという関係は、グラフ上で「 $u_i \xrightarrow{e_{(u_i, m_j)}} m_j$ 」と表現され、推薦における説明性の根拠として機能する．このような有向グラフによる階層構造の記述は、説明可能な推薦に関する多くの先行研究 [41, 44, 59] でも広く採用されている．

さらに、本手法では、先述の2種のエッジに加え、楽曲間の音響特徴に基づくエッジを新たに定義する．具体的には、各楽曲 m_j に対して M 次元の音響特徴ベクトル f_j を抽出し、任意の2曲間のコサイン類似度 $w_{j_1 j_2}$ があらかじめ定めた閾値 δ より大きい場合に、 m_{j_1} と m_{j_2} の間にエッジ $e_{(m_{j_1}, m_{j_2})}$ を定義する．このようにして、ユーザの聴取履歴、アーティスト情報、および音響特徴に基づいて知識グラフを構築する．音響特徴に基づくエッジを導入することで、知識グラフの密度を高めるとともに、ユーザ嗜好に関する情報不足を補完することができる．

4.2.2 知識グラフ上でのエージェント学習

エージェント学習フェーズの概要を図4.1左側に示す．本フェーズでは、前項で構築した知識グラフ $\mathcal{G}$ 上で、深層強化学習アルゴリズムに基づき、対象ユーザノードから楽曲ノードへの探索を行うエージェントの学習を行う．

深層強化学習においては、環境・状態・行動・報酬関数を定義することで、報酬の最大化を目的とした方策の学習を実現する。楽曲推薦タスクにおける各要素の定義は以下のとおりである：

環境. 環境は前項で構築した知識グラフ $\mathcal{G}$ である。

状態. エージェントが t ステップ ($t = 0, 1, \dots, T$; T は探索の終了ステップ数) 遷移した時点の状態は、 $s_t = (u_i, n_t, h_t)$ と定義される。ここで、 u_i は初期位置にあるユーザノード、 n_t はステップ t においてエージェントが到達したノード、 h_t はエージェントがステップ t 以前に通過したすべてのノードおよびエッジの列（すなわち $h_t = (n_0, e_{(n_0, n_1)}, \dots, n_{t-1}, e_{(n_{t-1}, n_t)})$ ）を表す。

行動. 行動 b_t は、ステップ t においてノード n_t から次に遷移するノードの選択である。候補ノードは n_t に接続するすべてのノードのうち、 t ステップ以前に通過していないノードとする。

報酬. 従来の知識グラフおよび強化学習に基づく推薦手法では、設計意図や妥当性が明確でない「ソフトな」報酬関数が用いられていた。本手法では、「推薦理由としての適切性」を明確に表現することを目的とし、知識グラフ埋め込み法の加法的合成性および埋め込み特徴間のコサイン類似度を活用して報酬を定義する。

まず、報酬定義の前処理として、前項で構築した知識グラフ $\mathcal{G}$ に対して、知識グラフ埋め込み法によりベクトル化を行う。図 4.2 に、報酬の計算方法の概要を示す。ここでは、代表的な埋め込み手法である TransE [47] に着想を得て、知識グラフ $\mathcal{G}$ をベクトル空間に写像する。トリプレット $(n_{\text{head}}, e, n_{\text{tail}})$ に対して、 $\mathbf{n}_{\text{head}} + \mathbf{e} = \mathbf{n}_{\text{tail}}$ が成立するように、ノードおよびエッジのベクトル表現を学習する。

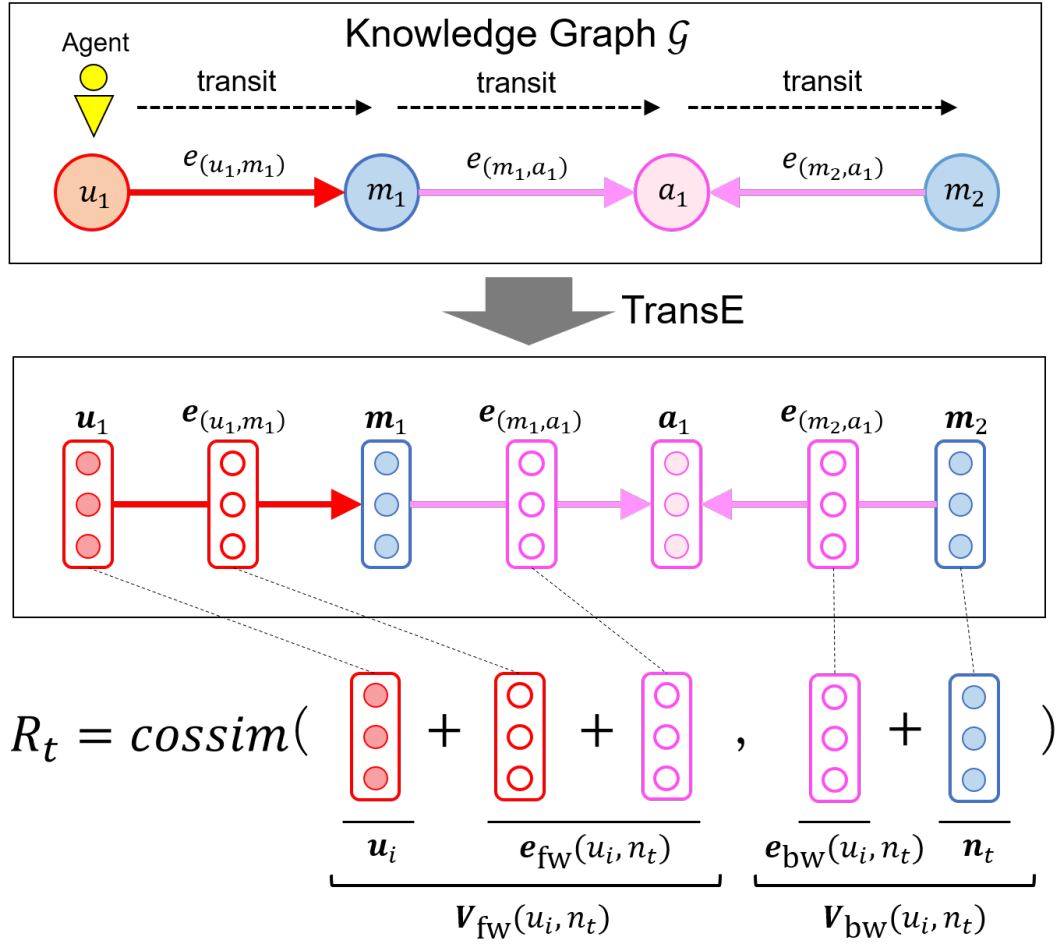


図 4.2: 報酬 R_t の計算方法の概要.

ここで、初期ユーザノード u_i からステップ t で到達したノード n_t までの探索パスに含まれるエッジの集合を $\mathcal{E}_{u_i, n_t}$ とし、その順方向および逆方向の合成ベクトル表現をそれぞれ $\mathbf{e}_{\text{fw}}(u_i, n_t)$, $\mathbf{e}_{\text{bw}}(u_i, n_t)$ と定義する．このとき、報酬 R_t は以下のように定義される：

$$R_t = \begin{cases} \frac{\mathbf{V}_{\text{fw}}(u_i, n_t) \cdot \mathbf{V}_{\text{bw}}(u_i, n_t)}{\|\mathbf{V}_{\text{fw}}(u_i, n_t)\| \|\mathbf{V}_{\text{bw}}(u_i, n_t)\|}, & \text{if } t = T \text{ かつ } n_t \in \mathcal{M} \\ 0, & \text{その他の場合} \end{cases} \quad (4.1)$$

$$\mathbf{V}_{\text{fw}}(u_i, n_t) = \mathbf{u}_i + \mathbf{e}_{\text{fw}}(u_i, n_t), \quad (4.2)$$

$$\mathbf{V}_{\text{bw}}(u_i, n_t) = \mathbf{n}_t + \mathbf{e}_{\text{bw}}(u_i, n_t), \quad (4.3)$$

ここで、 $\mathcal{M}$ は楽曲ノード集合であり、 $\mathbf{u}_i$ および $\mathbf{n}_t$ はそれぞれ u_i および n_t のベクトル表現である．

推薦の妥当性は、探索経路の順方向ベクトル $\mathbf{e}_{\text{fw}}$ と逆方向ベクトル $\mathbf{e}_{\text{bw}}$ のコサイン類似度により表現できると仮定している．類似度が高いほど推薦が自然で妥当であることを意味し、低いほど不自然・不適切であると解釈される．TransE による埋め込みベクトルのコサイン類似度を報酬として用いる手法は、知る限りでは先行研究に存在しない．

最適化. 状態ベクトル $\mathbf{s}_t$ を入力とし、行動 b_t を出力するエージェントの方策関数 $\pi_\theta(b_t | \mathbf{s}_t)$ は、以下のように定義される：

$$\pi_\theta(b_t | \mathbf{s}_t) = F_{\text{drop}}(\sigma(F_{\text{drop}}(\sigma(\mathbf{s}_t \mathbf{W}_1)) \mathbf{W}_2)) \mathbf{W}_3, \quad (4.4)$$

$$\theta = (\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3), \quad (4.5)$$

ここで、 $\mathbf{W}_1, \mathbf{W}_2, \mathbf{W}_3$ は方策のパラメータ、 $\mathbf{s}_t$ は TransE による u_i, n_t, h_t の埋め込みベクトルの連結である． $\sigma(\cdot)$ は非線形活性化関数である指数線形単位 (ELU), $F_{\text{drop}}(\cdot)$ は Dropout [60] を表す．

最適化の目的は、方策関数のパラメータ集合 θ を最適化し、以下の目的関数 $J(\theta)$ を最大化することである：

$$J(\theta) = E_{\pi_\theta}[\pi_\theta(b_t | \mathbf{s}_t) G(\mathbf{s}_t)], \quad (4.6)$$

$$G(\mathbf{s}_t) = \sum_{t'=t}^{T-1} \gamma^{t'-t} R_{t'+1}, \quad (4.7)$$

ここで、 $E[\cdot]$ は期待値、 γ は報酬の割引率である。

目的関数 $J(\theta)$ の最適化には、REINFORCE アルゴリズム [54] を用いる。以下の方策勾配を最小化することでパラメータ θ を学習する：

$$\nabla_{\theta} J(\theta) = E_{\pi_{\theta}}[\nabla_{\theta} \log \pi_{\theta}(b_t | s_t) G(s_t)]. \quad (4.8)$$

このように、目的関数 $J(\theta)$ の最大化解を直接求めることは困難であるが、REINFORCE アルゴリズムにより解析的に最適化を進めることが可能である。

4.2.3 方策に基づく楽曲の推薦

推薦フェーズの概要を図 4.1 右側に示す。前項で学習された方策 $\pi_{\theta}(b_t | s_t)$ とビームサーチアルゴリズムに基づき、対象ユーザに推薦すべき楽曲およびその経路を決定する。

図 4.3 に、知識グラフ上でのビームサーチアルゴリズムの概要を示す。ビームサーチは、各ステップにおいてスコアの高い上位 K ノードを保持しながら探索を進めるヒューリスティックな探索アルゴリズムである [61]。

まず、対象ユーザ u_i とステップ数 T を定める。次に、対象ユーザノード u_i を出発点として、知識グラフ $\mathcal{G}$ 上をビームサーチにより探索する。この探索では、各時刻において、学習済みの方策 $\pi_{\theta}(b_t | s_t)$ から出力される遷移確率に基づいて、上位 K 個のノードを選択し、エージェントを進める。

最終的に、エージェントがステップ T で到達した候補楽曲ノードとの間の経路の報酬値に基づき、対象ユーザに対する推薦楽曲を高報酬順に提示する。

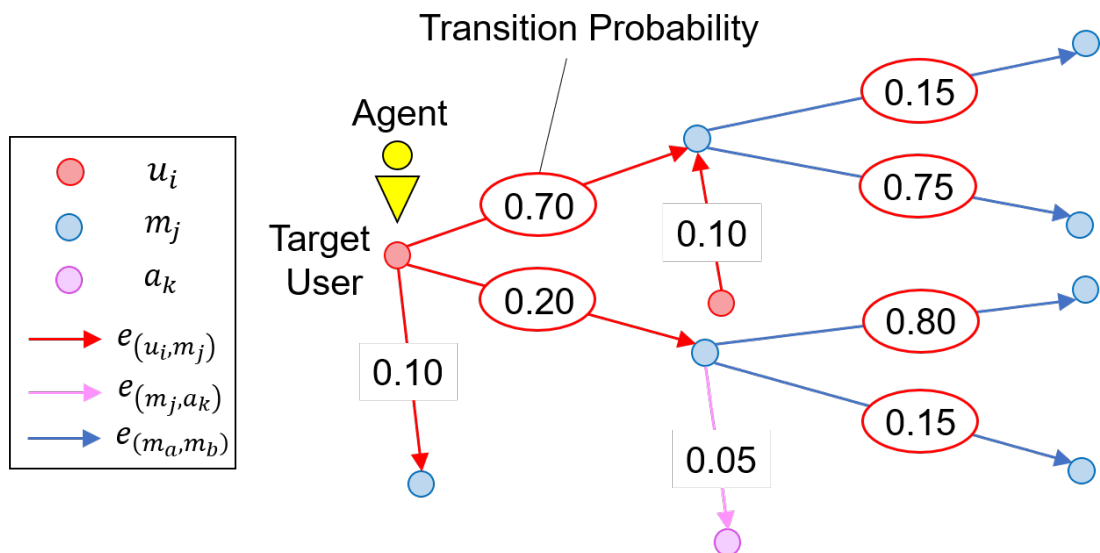


図 4.3: 知識グラフ G におけるビームサーチアルゴリズムの概要.

4.3 実験

本節では、提案手法の有効性を確認するために、既存の代表的な推薦手法および最先端の推薦手法と比較を行う。実験設定を4.3.1項に、実験結果を4.3.2項に記載する。

4.3.1 実験設定

本実験では、The Music Listening Histories Dataset [62] に収録された 1,536 名分のリスニング履歴を使用した。各ユーザについて過去5年間のリスニング履歴を取得し、再生回数の多い上位 50 曲を「そのユーザが聴いている楽曲」と定義した。各楽曲については、MusicBrainz¹からアーティスト情報を取得した。さらに、Spotify API²より、danceability, energy, speechiness, acousticness, valence, instrumentality の 6 次元の音響特徴を取得した。先行研究 [63] に従い、これらの 6 特徴量を用いることで楽曲の多面的な特徴を表現できる。

二値情報である “mode” は使用せず、代わりに float 値で歌声の有無を考慮できる “instrumentality” を採用した。これにより、歌詞を重視するユーザや、楽器の構成に注目するユーザなど、多様な嗜好に対応することができる。たとえば、“speechiness” は発話の多さを示し、“acousticness” および “instrumentality” はそれぞれアコースティック性とボーカルの有無を示す。また、“danceability” は楽曲のダンス適性を、“valence” および “energy” は楽曲のポジティブ／ネガティブさを表す。これらを統合することで、ユーザの気分や音響的嗜好を反映した推薦が可能となる。コサイン類似度の閾値 δ は 0.999 に実験的に設定した。結果として、38,526 曲、1,500 ユーザ、6,975 アーティストを対象としたデータセットを構築した。

エージェント学習フェーズでは、ステップ数 T を 4、報酬割引率 γ を 0.99 に設定した。推薦フェーズにおけるビームサーチのサンプリングサイズ K は 4 とした。また、REINFORCE アルゴリズムに基づき、ユーザごとにエピソードを反復生成しながら方策関数の最適化を行った。1 ユーザあたり 10 エピソードを生成し、各ステップにおける状態・行動・報酬をもとに勾配を計算した。学習率は 1×10^{-4} に固定し、最

¹<https://musicbrainz.org/>

²<https://developer.spotify.com/documentation/web-api/>

適化にはAdamを用いた。学習の最大エポック数は100とし、バリデーションセット上の累積報酬が10エポック連続で改善しなかった場合に早期終了を適用した。

テストフェーズでは、学習済みの方策 $\pi_{\theta}(b_t | s_t)$ を固定し、テストデータに含まれる未知履歴ユーザに対して、知識グラフ $\mathcal{G}$ 上でのビームサーチ探索（サンプリングサイズ $K=4$ ）を行った。このとき、各ユーザに対して最大ステップ数 $T=4$ の範囲で推薦候補を生成し、報酬値に基づくスコア順で推薦リストを決定した。

まず、コールドスタート問題に対する音響特徴に基づくエッジ導入の有効性を、推薦可能ユーザ率により評価した。グラフベースの推薦手法は、楽曲に対応するエッジが少ない場合、推薦困難となるコールドスタート問題が発生しやすい。本手法では、この課題に対処するために、音響特徴に基づくエッジを導入している。

具体的には、ユーザの再生履歴のうち、下位 α ($=90, 80, 70, 60, 50$) [%] に属する楽曲に対応するエッジを削除し、10曲以上推薦されたユーザの割合を「推薦可能ユーザ率」と定義した。すなわち、 α が大きくなるほど、コールドスタートの程度が高くなる。この評価指標は、[64]などでも用いられている一般的な手法である。本比較では、提案手法(PM)と、音響特徴を除いたPM-AF、さらにアーティスト情報も除いたPM-AF-AIを比較した。

加えて、提案手法の推薦性能についても評価を行った。評価指標として、Precision, Recall, nDCG, Hit Rateを用いた。各評価では、ユーザの聴取履歴のうち α ($=90, 80, 70, 60, 50$) [%] のエッジを削除した知識グラフを用いて推薦を行い、削除された楽曲が推薦されるかを評価した。比較対象は以下の5手法である：

CM1 [41]

知識グラフとRNNに基づく最先端の推薦手法 (KGAT)。

CM2 [65]

知識グラフとTransE埋め込みを用いた協調フィルタリング (CFKG)。

CM3 [25]

暗黙的フィードバックに基づく行列分解手法 (BPR)。

CM4 [40]

知識ベースからアイテムの意味的表現を学習する協調知識ベース埋め込み (CKE)。

表 4.1: 推薦可能ユーザ率 [%].

	$\alpha=90$	$\alpha=80$	$\alpha=70$	$\alpha=60$	$\alpha=50$
PM	18.5	64.5	90.1	99.9	99.9
PM - AF	12.0	62.3	81.6	88.0	96.0
PM - AF - AI	10.4	43.0	66.9	76.4	88.4

CM5 [66]

音響特徴からセマンティックタグを予測するハイブリッド楽曲推薦モデル.

なお, [59] の手法は全パスの列挙が必要であり, 本研究の推薦タスクには適さないため, より新しい [41] を採用した. また, 提案手法との公平な比較のため, 各比較手法においても音響特徴を用いている.

4.3.2 実験結果

表 4.1 に推薦可能ユーザ率の結果を示す. すべての α において, PM の値が PM - AF および PM - AF - AI を上回っており, 音響特徴を導入することで, 楽曲ノードへの接続エッジ数が増加し, 推薦可能ユーザ数が増加している. これにより, 音響特徴ベースのエッジがコールドスタート問題の緩和に有効であることが確認できる.

図 4.4 に, Precision, Recall, nDCG, Hit Rate の各評価指標における結果を示す. PM はすべての α において, 既存のすべての比較手法 (最先端手法含む) よりも Precision, Recall, Hit Rate が高い結果を示している. 特に $\alpha=70$ において各指標が最大となっており, これは知識グラフが適度に疎であるときに, 音響特徴が推薦に効果的に働くことを示唆している.

一方, nDCG においては, $\alpha=90$ のときに PM が高い値を示した. nDCG は推薦順位を評価する指標であり, 上位に正解楽曲が含まれるほど高い値を取る. エッジが大きく削除される $\alpha=90$ では候補が多くなるが, 音響特徴ベースのエッジを活用することで, 上位への推薦が可能となったと考えられる.

ただし, $\alpha=60$ および 70 において, PM の nDCG は CM3 および CM4 より低く, さらに $\alpha=50$ では CM1, CM3, CM4, CM5 よりも低くなった. これは, 知識グラフ探索による正解楽曲の推薦には優れるが, ランキング最適化には改善の余地があることを示している. また, $\alpha=50$ では, PM の Precision, Recall, Hit Rate が CM3, CM4, CM5 に劣る場合があった. これは, 音響特徴エッジがユーザの聴取履歴ベースの探

索を妨げている可能性があるが，そもそも α が小さい設定はコールドスタート問題ではないため，特段不自然ではない．今後は，コールドスタート／非コールドスタートの両問題に対応可能なハイブリッドモデルの構築を目指す．

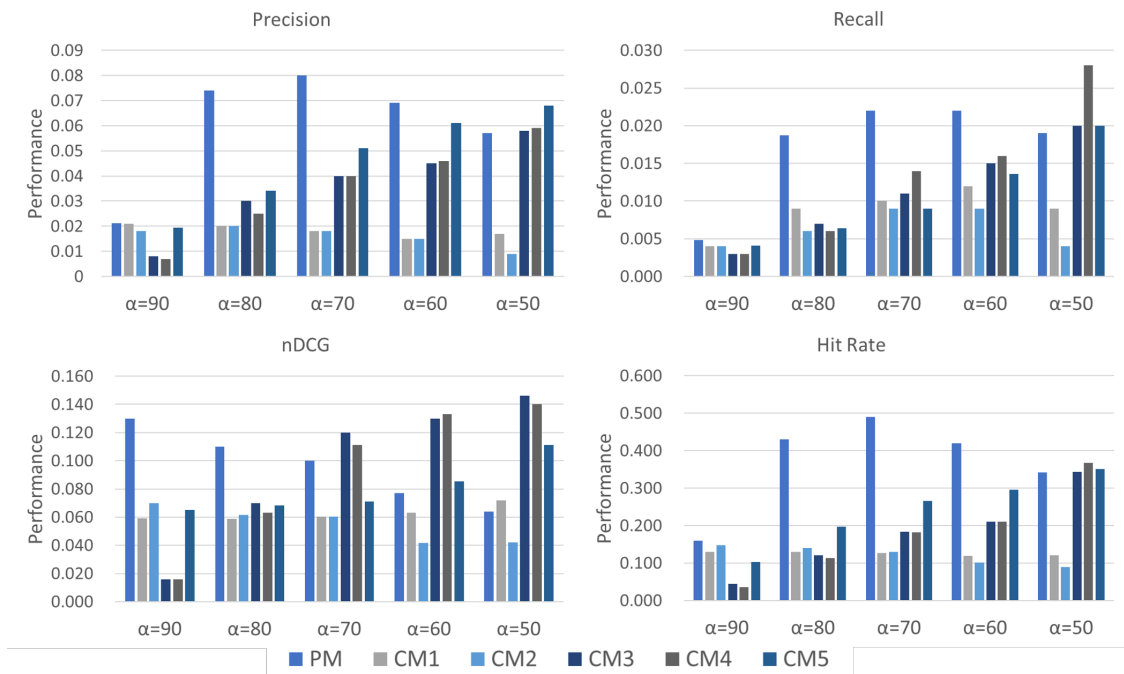


図 4.4: PM および CMs 1-5 における推薦性能の比較.

4.4 結論

本章では，知識グラフと強化学習を組み合わせた新たな楽曲推薦手法を提案した．提案手法は，楽曲の音響特徴を知識グラフに導入することで，コールドスタート問題を緩和しつつ，高精度な推薦を実現可能である．実験結果から，提案手法が従来手法を上回る性能を有することが確認された．

第5章

LLMの出力を用いた知識グラフに基づくアイテム推薦

5.1 はじめに

本章では、LLMの出力を用いた深層強化学習と知識グラフに基づくシーケンシャルアイテム推薦手法について説明する。近年、知識グラフを活用した推薦システムは、ユーザとアイテム、およびその属性間の構造的関係を利用することで、高次の意味的推論や推薦の説明可能性を実現してきた。しかしながら、従来の知識グラフベース推薦は、時系列的な文脈変化への対応が困難であり、特にユーザ履歴が乏しいコールドスタート環境では推薦精度が著しく低下するという課題がある。一方で、LLMは、事前学習によって豊富な知識と推論能力を獲得しており、少ない情報からでもユーザの嗜好や意図を推定できる点で、コールドスタート状況における有望な手がかりとなる。本研究では、このLLMの直感的推論能力を知識グラフ推薦に統合し、推薦方策の学習と探索の指針として活用することで、履歴の少ない状況においても高精度な推薦を実現する手法を提案する。具体的には、LLMによりユーザの直近の行動履歴をプロンプトとして入力し、その出力から推薦に有効なアイテム属性（例：ジャンル、アーティスト、年代）を抽出する。そのうえで、この出力を報酬設計に組み込んだ深層強化学習を用いて、知識グラフ上の推薦経路を学習する。LLMが示唆するユーザ意図と、知識グラフ埋め込みに基づく意味的類似性を統合的に考慮した報酬関数により、推薦経路の探索方針を柔軟かつ文脈依存的に制御することが可能となる。

本章の構成は以下の通りである．5.2節では，提案手法の詳細について説明する．5.3節では，提案手法の有効性を検証するための実験を行い，5.4節において，本章のまとめを述べる．

5.2 提案手法

本節では，提案手法である LIKR の詳細を説明する．コールドスタート設定とは，ユーザのインタラクション履歴が少量である状況を指す．コールドスタートシーケンシャル推薦タスクでは，限られたインタラクションデータしか利用できない状況下で，ユーザの次のインタラクション対象を予測することが求められる．LIK R はこの課題に対して，以下の2フェーズで対応する：(1) LLM によるユーザの将来的な嗜好に関する直感 (intuition) の獲得，および (2) 獲得された直感と強化学習に基づく知識グラフ推論による推薦である．LIK R フレームワークの全体像を図 5.1 に示す．

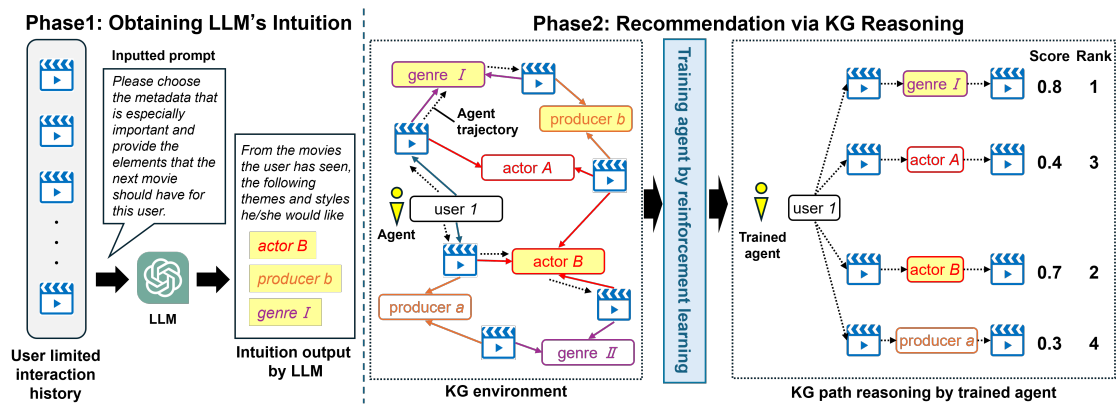


図 5.1: LIKR の概要図.

5.2.1 LLMによるユーザの将来的嗜好に関する直感の獲得

第一フェーズでは、ユーザの将来的な嗜好に関するLLMの直感を取得する。この直感は、後述する知識グラフ推論の戦略決定（5.2.2節）に影響を与える重要な要素となる。先行研究[67]に従い、LLMをreasonerとして扱い、時系列意識（temporal awareness）を考慮したプロンプトを用いて将来の嗜好を予測する。

具体的には、ユーザ u のインタラクション履歴を時系列順に並べた $\mathcal{H}_u = \{i_1^u, i_2^u, \dots, i_{N_i^u}^u\}$ と、推薦対象ドメイン d におけるメタデータ種別の集合 $\mathcal{Y}^d = \{y_1^d, y_2^d, \dots, y_{N_y^d}^d\}$ を入力としたプロンプトをLLMに与える。ここで、 i_j^u はユーザ u がインタラクトしたアイテム、 N_i^u はその総数を示す。また、 y_j^d はドメイン d に含まれるメタデータの種別を表し、 N_y^d はその総数である。

たとえば、映画ドメイン（ $d = \text{movie}$ ）におけるプロンプトは以下のように設定される：

“I am thinking of recommending the next movie for a user to watch. The user has watched the following movies in this order, with the more recently watched movies appearing at the end: $i_1^u, i_2^u, \dots, i_{N_i^u}^u$.

Based on the metadata types such as $y_1^{\text{movie}}, y_2^{\text{movie}}, \dots, y_{N_y^{\text{movie}}}^{\text{movie}}$, please choose the metadata that is especially important, and provide the elements that the next movie should have for this user.

Please format the output as follows:

y^{movie} : element”

下線部分は、時間的文脈を強調するための記述である。このようなプロンプトをLLMに入力することで、ユーザ u にとって次に重要と予測されるメタデータ要素の集合 $\mathcal{E}_u$ が出力される。この $\mathcal{E}_u$ は、上記プロンプトにおける“element”に対応するものである。

LLMはTransformerアーキテクチャ、特に自己注意機構に基づいて設計されており、インプット系列全体に同時に注意を向けることが可能である。この構造により、ユーザのインタラクション履歴と各メタデータ種別の双方を時間的な流れを加味しつつ効果的に解釈することができる。ここで得られるLLMの直感は、単なるパターン認識を超え、事前学習された知識とトレンドに基づいてユーザ行動を文脈化する能力に裏付けられている。

本アプローチの利点は、その適応性と拡張性にある。LLMは、特定のドメインにおけるファインチューニングを必要とせず、事前学習された知識を活用することで、様々なドメインにおけるインタラクション系列やメタデータ種別を柔軟に解釈できる。

得られたメタデータ要素集合 $\mathcal{E}_u$ は、ユーザの将来的嗜好を反映したものであり、後続の知識グラフ推論戦略を洗練させるための基盤情報として機能する。

5.2.2 LLMの直感と強化学習に基づく知識グラフ推論による推薦

第二フェーズでは、5.2.1節で得られたLLMの直感を用いて、知識グラフ推論を実行し、推薦対象アイテムを決定する。

まず、データセットから知識グラフ $\mathcal{G} = (\mathcal{V}, \mathcal{R})$ を構築する。ここで、 $\mathcal{V}$ はエンティティの集合、 $\mathcal{R}$ はそれらの間の関係を表す。エンティティには、ユーザ、アイテム、およびアイテムの各種メタデータが含まれる。

構築した知識グラフの意味的情報を活用し、関係の多様性を捉えるため、知識グラフの埋め込みを行う。知識グラフ埋め込みにはTransE [47]を用いる。TransEでは、エンティティと関係を同一のベクトル空間に埋め込み、関係 r によってエンティティ e_1 から e_2 へのベクトル変換が実現されるよう学習される。具体的には、以下の式を満たすようにベクトルが学習される：

$$\mathbf{e}_1 + \mathbf{r} \approx \mathbf{e}_2, \quad (5.1)$$

ここで、 $\mathbf{e}_1$ および $\mathbf{e}_2$ はエンティティ e_1 と e_2 のベクトル表現、 $\mathbf{r}$ は関係 r のベクトル表現である。TransEは構造が単純で計算効率に優れており、推薦システムにおける大規模なグラフデータにおいて、多対多関係を適切に捉えることが示されている [44, 68]。

LLMによる直感と知識グラフ埋め込みベースの手法を組み合わせた推薦戦略を実現するため、強化学習による知識グラフ推論を導入する。強化学習により、複数の指標を同時に考慮した推薦戦略の最適化が可能となる。具体的には、[44]の枠組みに従い、本推薦タスクをマルコフ決定過程 (MDP) として定式化する。以下にMDPの構成要素を示す：

- **状態 (State)** : 各時刻 t ($t = 0, 1, \dots, T$; T は探索のステップ数) において, 状態は $s_t = (u, e_t, h_t)$ と定義される. ここで, $u \in \mathcal{U}$ はユーザエンティティ, e_t は t 時点で到達しているエンティティ, h_t は過去に通過したエンティティと関係の履歴を表す. 初期状態では, エージェントはユーザノード上に位置し, $t = 0$ において $e_0 = u$ が成り立つため, 初期状態は $s_0 = (u, u, \emptyset)$ と定義される. 最終的に到達したアイテムノードが推薦対象となる.
- **行動 (Action)** : 各時刻 t において, エージェントは e_t から遷移可能な出次数エッジから行動 a_t を選択し, 新たな状態 s_{t+1} に遷移する. 多様な経路探索を促進するため, 行動空間 $A(s_t)$ には, すでに訪問したエンティティおよび関係は含まれないよう設計する.
- **報酬 (Reward)** : e_0 から e_k までの経路に基づき, 報酬 R_t を定義する. 報酬は LLM の直感に基づく項と, 知識グラフ埋め込みに基づく項から構成される.

- **LLM 直感に基づく報酬 R_t^{LLM}** : エージェントがユーザの直感集合 $\mathcal{E}_u$ に含まれるノードに到達した場合に報酬を与える:

$$R_t^{\text{LLM}} = \begin{cases} 1, & \text{if } e_0 = u \text{ and } e_t \in \mathcal{E}_u \\ 0, & \text{otherwise.} \end{cases} \quad (5.2)$$

- **知識グラフ埋め込みに基づく報酬 R_t^{KG}** : 経路上のエンティティ間の内積を用いて, 以下のように定義される:

$$R_t^{\text{KG}} = \sum_{i=0}^t \langle \mathbf{e}_i, \mathbf{e}_{i+1} \rangle, \quad (5.3)$$

ここで, $\mathbf{e}_i$ はエンティティ e_i のベクトル表現を表す.

全体の報酬 R_t は, 2つの項を重み付きで加算することで定義される:

$$R_t = R_t^{\text{LLM}} + \alpha R_t^{\text{KG}}, \quad (5.4)$$

ここで, α は両者の寄与を調整するハイパーパラメータである.

上記の MDP 設定に基づき, 方策ネットワークの最適化には REINFORCE アルゴリズム [69] を用いる. 学習済みの方策ネットワークをもとに, エージェントは知識グラフ推論により推薦アイテムを決定する.

ユーザノード $u \in \mathcal{U}$ を起点として、現在の状態 s_t からの有効な行動（エッジ）は学習済み方策に基づいて評価される．推論効率を高めるため、各時刻 t において方策スコアの高い上位 z 個の行動を選択し、この処理を T ステップにわたって再帰的に実行する．生成された全経路を記録し、最終的な Top- N 推薦を構築する．

5.3 実験

本節では、実世界の知識グラフベースの推薦データセットを用いた実験により、以下の研究課題（RQ）に対する検証を行う．

RQ1: コールドスタートのシーケンシャル推薦シナリオにおいて、LIK R は従来手法と比較してどの程度の性能を示すか．

RQ2: LLM の種類およびプロンプトの設計は、LIK R の推薦性能にどのような影響を与えるか．

RQ3: 報酬関数に含まれる異なる推薦戦略は、LIK R の推薦性能にどのように影響するか．

5.3.1 実験設定

データセット: 本研究では、知識グラフを用いた推薦において代表的な2つの実世界データセット MovieLens-1M [70] および Lastfm-1M [71] を用いた．各データセットの統計量を表 5.1 に示す．Lastfm-1M は [71] により提供されており、知識グラフ推薦の公平な比較のために Lastfm-1B [72] から整備されたものである．

本手法では、LLM からの出力を得るためにアイテムタイトルのテキストと、LLM と知識グラフとを接続するためにメタデータのテキストが必要となる．そのため、アイテムおよびメタデータが ID のみで構成されるデータセットは除外した．インタラクション時刻に基づくホールドアウト戦略に従い、データを古い順に 60%, 20%, 20% で訓練・検証・テストに分割した．

推論フェーズでは、学習済みの方策 $\pi_\theta(b_t | s_t)$ を固定し、テストデータに含まれる未知履歴ユーザに対して、知識グラフ $\mathcal{G}$ 上でのビームサーチ探索（サンプリングサイズ $K=4$ ）を行った．このとき、各ユーザに対して最大ステップ数 $T=4$ の範囲で推薦候補を生成し、報酬値に基づくスコア順で推薦リストを決定した．

表 5.1: 各データセットの統計量.

		MovieLens-1M	Lastfm-1M
Interactions	Users	6,040	4,817
	Items	2,984	12,492
	Interactions	932,295	1,091,275
	Density	0.05	0.01
知識グラフ	Entities (Types)	13,804 (12)	17,492 (5)
	Relations (Types)	193,089 (11)	219,089 (4)
	Sparsity	0.0060	0.0035

さらに、コールドスタートシーケンシャル推薦を模擬するために、訓練データ中のユーザー-アイテムインタラクションを制限し、過去 10 件のみを保持した（先行研究 [56, 73–75] に準拠）。これにより、ユーザの直近の関心や嗜好を予測する能力を評価可能とした。

比較手法: 本研究では、シーケンシャル推薦モデルおよび知識グラフベース推薦モデルを比較対象とした。シーケンシャル推薦モデルとしては、GRU4Rec [76] および SASRec [74] を採用した。前者は GRU [48] を用いたセッションベース推薦、後者は Self-Attention 機構により次アイテム予測を行う手法である。実験 [74, 76] では、検証データをテスト時に使用しているが、知識グラフベース手法は訓練データのみで推論を行うため、本研究でも後者の評価方法に従い、テスト時には検証データを入力しない設定とした。

知識グラフベース手法として、強化学習に基づく経路推論を行う PGPR [44], CAFE [42], およびユーザの中心性を考慮した UCPR [43] を採用した。なお、LLM を用いた推薦モデル（例：LLMRank [67]）は、候補アイテム数が限定される前提で設計されており、本研究のように大規模候補集合を扱う推薦シナリオとは異なるため、比較から除外した。同様に、LLM から知識グラフの構築を行う CSRec [77] のような手法は公式実装がないため、本研究の範囲外とした。

シーケンシャル推薦モデルは RecBole [78] ライブラリを用いて実装し、知識グラフモデルは [71] の実装に準拠した。すべての手法について、論文で効果が報告されているハイパーパラメータを中心にグリッドサーチにより最適値を探索した。

評価指標: トップ k 推薦精度の評価指標として、Recall@ k および nDCG@ k ($k=20, 40$) を使用した。

LLM への入力には、各データセットで提供されているメタデータを用いたが、

MovieLens-1M の “wikipage” および “category”, Lastfm-1M の “genre” は ID の照合困難性から除外した。ユーザによる推薦理由の理解を考慮し、最大パス長は 3 に設定した。TransE による知識グラフの埋め込み次元は 100, 行動空間の最大サイズは 400, アクションドロップアウトは 0.5 とした。強化学習の割引率は 0.99, 報酬関数のバランスをとるパラメータ α は {0, 0.25, 0.5, 0.75, 1} から選択した。すべてのデータセットにおいて、学習は Adam 最適化器 (学習率 0.001, バッチサイズ 64) を用い、1 ユーザあたり 10 エピソードを生成し、各ステップにおける状態・行動・報酬をもとに勾配を計算した。学習の最大エポック数は 100 とし、バリデーションセット上の累積報酬が 10 エポック連続で改善しなかった場合に早期終了を適用した。ビームサーチのサンプリングサイズは、MovieLens-1M ではそれぞれ 1, 2, 4, Lastfm-1M では 1, 5, 10 とした。

5.3.2 全体性能比較 (RQ1)

RQ1 に対する検証として、提案手法 LIKR とシーケンシャル推薦モデルおよび知識グラフベース推薦モデルとの推薦性能の比較を行った。その結果を表 5.2 に示す。この実験では、2024 年 10 月時点で最新の GPT 系モデルである GPT-4o-preview を用いた。

まず、シーケンシャル推薦モデルとの比較において、LIK R はすべての評価指標およびすべてのデータセットにおいて最も高い性能を示した。GRU4Rec および SASRec は、セッションベース推薦のために設計されたモデルであり、コールドスタートシーケンシャル推薦のシナリオにおいても一定の性能を示した。特に、知識グラフベース手法と比較して競合または優れた性能を示すことが確認された。一方、nDCG の値が相対的に低い理由としては、これらのモデルが訓練・検証データを併用した単一アイテム予測を目的として設計されていることが挙げられる。一方、LIK R は LLM の時系列的理解能力を活かすことにより、コールドスタート状況においても高精度な推薦を実現できることが確認された。

次に、知識グラフベース推薦手法との比較において、LIK R は Lastfm-1M の Metrics@40 を除くすべての指標において、従来手法を上回る性能を示した。特に、強化学習による経路推論を行う PGPR と比較して、LIK R は全指標において上回る結果を示し、LLM の直感 (intuition) を強化学習に統合するアプローチの有効性が示

表 5.2: MovieLens-1M および Lastfm-1M における推薦性能の比較．最も高い値を太字，2 番目に高い値を下線付きで示す．

	MovieLens-1M				Lastfm-1M			
	Metrics@20 [%]		Metrics@40 [%]		Metrics@20 [%]		Metrics@40 [%]	
	Recall	nDCG	Recall	nDCG	Recall	nDCG	Recall	nDCG
GRU4Rec	<u>4.77</u>	4.73	9.02	6.30	1.22	2.65	2.16	2.67
SASRec	4.63	4.47	8.72	6.04	0.77	1.28	1.27	1.36
PGPR	4.63	<u>18.53</u>	<u>9.32</u>	23.12	1.27	<u>12.73</u>	2.29	14.30
CAFE	4.14	15.35	7.27	22.86	1.19	10.10	2.20	15.05
UCPR	2.64	14.47	7.31	<u>23.28</u>	<u>1.38</u>	12.34	3.32	18.72
LIK R (GPT-o1-preview)	4.83	19.14	9.46	23.80	1.42	12.93	<u>2.83</u>	<u>16.32</u>

された．ただし，Lastfm-1M の Metrics@40 において，UCPR が最も高い値を示したことから，LIK R は時系列を考慮することで短期的な嗜好予測に優れる一方で，長期的な嗜好予測においては限界がある可能性が示唆される．

以上より，LIK R はシーケンシャル推薦モデル・知識グラフ推薦モデルの両カテゴリにおける代表的手法を上回る性能を示し，コールドスタート状況下においても有効な推薦が可能であることが確認された．

5.3.3 LLM の種類とプロンプトの影響分析 (RQ2)

RQ2 に対する検証として，LIK R において用いる LLM の種類およびプロンプト設計が推薦性能に与える影響を分析した．

まず，LLM の種類による影響を検証するために，以下の 6 種類の代表的な LLM を用いた：Gemini1.5-flash, Claude3.5-sonnet, GPT-3.5-turbo, GPT-4-turbo, GPT-4, および GPT-4o-preview である．その結果を表 5.3 に示す．両データセットにおいて，GPT-4o-preview を用いた LIK R が最も高い推薦性能を示した．GPT-4o-preview はプロンプトの内容を精緻に読み取り，推論過程を反映した出力を生成できるため，インタラクション履歴および時系列情報を入力することで，ユーザの将来的な嗜好をより高精度に予測できると考えられる．GPT 系モデルにおいては，新しい世代のモデルほど推薦性能が向上する傾向が見られたが，GPT-3.5-turbo と GPT-4 の間では顕著な差は確認されなかった．一方で，Gemini1.5-flash はプロンプトに指定された出力形式を十分に理解できず，嗜好予測を拒否する出力が他のモデルと比べて多く見られたため，性能が相対的に低下したと考えられる．

次に，プロンプトの設計が推薦性能に与える影響を検証するために，時系列意

表 5.3: 使用する LLM の違いによる LIKR の推薦性能の比較.

	MovieLens-1M		Lastfm-1M	
	Metrics@20 [%]		Metrics@20 [%]	
	Recall	nDCG	Recall	nDCG
LIKR (Gemini1.5-flash)	4.61	18.90	1.13	11.41
LIKR (Claude3.5-sonnet)	4.76	19.16	1.30	<u>12.79</u>
LIKR (GPT-3.5-turbo)	4.73	19.01	1.29	12.33
LIKR (GPT-4-turbo)	4.76	18.92	1.32	12.33
LIKR (GPT-4)	<u>4.78</u>	19.05	<u>1.36</u>	11.56
LIKR (GPT-o1-preview)	4.83	<u>19.14</u>	1.42	12.93

表 5.4: プロンプトの違い（時系列意識の有無）による LIKR の推薦性能の比較.

手法	MovieLens-1M		Lastfm-1M	
	Metrics@20 [%]		Metrics@20 [%]	
	Recall	nDCG	Recall	nDCG
LIKR (GPT-4-turbo：時系列意識なし)	4.62	18.78	1.20	12.30
LIKR (GPT-4-turbo)	4.76	18.92	1.32	12.33
LIKR (GPT-4：時系列意識なし)	4.66	18.97	1.28	11.89
LIKR (GPT-4)	4.78	19.05	1.36	11.56
LIKR (GPT-o1-preview：時系列意識なし)	4.65	18.60	1.36	12.65
LIKR (GPT-o1-preview)	4.83	19.14	1.42	12.93

識を考慮するプロンプトと考慮しないプロンプトの比較を行った．時系列を考慮しないプロンプトでは，5.2.1 節で下線部として示された文言を省略した．その結果を表 5.4 に示す．GPT-4 における nDCG@20 を除くすべての指標において，時系列を考慮したプロンプトの方が高い推薦性能を示した．特に GPT-4o-preview はプロンプトの構造に敏感であり，時系列情報の有無により出力結果が大きく変化することが確認された．また，時系列を考慮しない一部の LIKR モデルでは，シーケンシャル推薦モデルよりも推薦性能が劣る場合も見られた．

以上より，LLM の選択とプロンプトの設計は，提案手法 LIKR の効果に対して顕著な影響を与えることが明らかとなった．

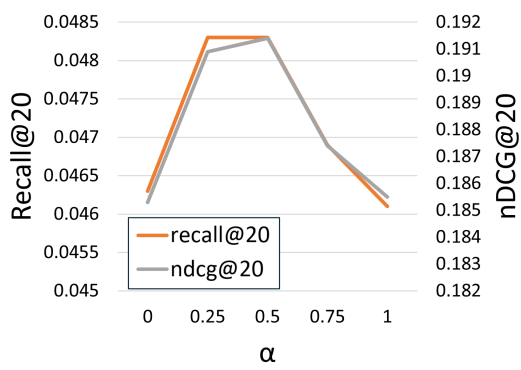
5.3.4 推薦戦略の影響分析 (RQ3)

RQ3 に対する検証として，報酬関数に含まれる推薦戦略の影響を評価するため，報酬関数中のハイパーパラメータ α を変更した際の推薦性能を比較した．その結果を Fig. 5.2 に示す．

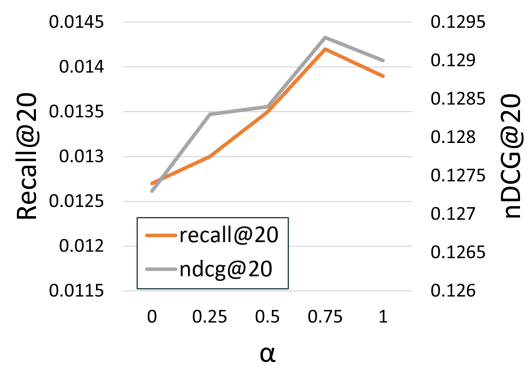
まず Fig. 5.2 (a) に示される MovieLens-1M の結果を見ると、 $\alpha = 0.25$ から 0.5 の範囲で最も高い推薦性能が得られていることが分かる。一方で Fig. 5.2 (b) に示される Lastfm-1M では、 $\alpha = 0.75$ のときに最大性能が得られている。これは、映画と音楽というドメインの違いにより、LLM の直感 (intuition) に依拠する度合いの最適値が異なることを示唆している。

また、 $\alpha = 0$ または $\alpha = 1$ とした場合、両データセットにおいて推薦性能が低下する傾向が見られた。これは、知識グラフ埋め込みに基づく情報と LLM の直感の両方を適切に組み合わせて用いることが、推薦性能の向上に不可欠であることを意味している。

以上の結果から、LLM ベースの直感的知識と知識グラフベースの構造的知識の両方が、コールドスタート推薦において相補的な役割を果たしており、報酬関数における 2 種類の指標はいずれも推薦性能に対して重要な影響を持つことが確認された。



(a) Reward impact in MovieLens-1M



(b) Reward impact in Lastfm-1M

図 5.2: LLM の直感に基づく報酬と知識グラフ埋め込みに基づく報酬の影響.

5.4 結論

本研究では、LLMの直感を活用した強化学習および知識グラフに基づく推薦モデルLIKRを提案した。提案手法は、LLMを推論器として位置づけ、知識グラフ上の直感的な探索戦略を出力させることで、コールドスタートなシーケンシャル推薦において推薦性能の向上を実現した。2つの実世界データセットを用いた従来の推薦モデルとの比較実験を通じてLIKRの有効性を示した。本手法におけるLLMの利用戦略は単純かつ柔軟であり、従来のLLMベース推薦システムが抱えるスケーラビリティの課題を克服可能である。また、さまざまな推薦モデルにおける報酬関数や損失関数への応用も期待される。

しかしながら、LLMと強化学習の利用に関して、LIKRにはいくつかの課題が存在する。

LLMの課題：LIKRでは、アイテムタイトルを入力としてLLMの推薦リテラシーを活用している。そのため、アイテムタイトルが利用できないデータセットに対して適用が困難である。さらに、LLMが十分な知識を持たないニッチなアイテムがユーザ履歴に含まれる場合、ユーザ嗜好の正確な把握が困難となる可能性がある。

強化学習の課題：LIKRでは、古典的な方策勾配法であるREINFORCEを用いているが、累積報酬の分散が大きくなることにより、学習の安定性が損なわれる恐れがある。また、方策が急激に変化することで、過度な探索が生じ、学習時間が長期化する可能性もある。

今後の課題として、ニッチなアイテムに対応するためのLLMのファインチューニングに加え、より安定した学習を実現するために、Proximal Policy Optimization (PPO) や Deep Deterministic Policy Gradient (DDPG) などの強化学習手法の導入を検討する予定である。

第6章

結論

本論文では、深層学習と知識グラフ解析に基づくハイブリッド推薦システムを提案し、ユーザの目的や状況に応じた柔軟な推薦制御の実現、およびコールドスタート状況における意味的推論の限界という2つの主要な課題に対処することを目的とした。

まず、ユーザの明示的な意図に基づく柔軟な推薦制御の実現に向けて、音楽ドメインにおけるプレイリスト生成を対象とした知識グラフと強化学習に基づく推薦手法を提案した。本手法では、知識グラフに基づく意味的關係を特徴量として用い、ユーザの目的に応じて調整可能な報酬関数を設計することで、推薦方針の柔軟な制御を可能とした。次に、コールドスタート問題への対応として、知識グラフを外部情報によって補完する2つの手法を提案した。1つ目は、楽曲の音響特徴を知識グラフに反映させ、音響的類似性に基づく新たな関係を構築することで、ユーザ履歴が乏しい状況でも音響的嗜好に基づく推薦を実現する手法である。2つ目は、LLMによる高次推論を活用し、ユーザ履歴から抽出された情報をもとに意味的に関連するアイテムを推定し、知識グラフ上での探索と推薦に活用する手法である。これにより、従来手法では困難であったスパースな関係に基づく推薦の精度向上を実現した。本論文で提案した各手法について、音楽ドメインおよび一般的な推薦タスクに対する実データを用いた評価実験を行い、既存手法と比較して高い推薦精度を達成することを確認した。さらに、提案手法はいずれも説明可能性を保持しており、推薦の透明性と信頼性の向上にも寄与することが示された。

今後の研究方針としては、以下の点が挙げられる。第一に、ユーザの意図の反映およびコールドスタート問題の解決の両方に対処可能な推薦システムの構築であ

る．本論文では，ユーザの意図に基づく柔軟な推薦方針の制御とコールドスタート問題の解決をそれぞれ独立に解決したため，それぞれの手法において未着手の課題が残存している．したがって，今後の研究では，本論文のアプローチを統合し，これらの課題を同時に解決可能な手法を構築する予定である．第二に，提案手法の他ドメインへの汎化性能や拡張可能性についての追加検証である．特に，第3章および第4章では音楽ドメインに注目した手法を提案したため，提案手法の他ドメインへの汎化性能に関しては未検証である．したがって，今後の研究では，動画やSNS等のドメインにおける提案手法の拡張可能性に関して検証を行う予定である．

謝辞

本研究は、著者が北海道大学および北海道大学大学院に在学した2019年10月から2025年9月までの約6年間にわたって行ったものである。

本研究に関して、研究遂行のみならず、終始御指導および御鞭撻を頂きました長谷山美紀教授および小川貴弘教授に心より感謝申し上げます。加えて、多くの国内・国外学会への参加、論文執筆、および教育活動等、様々な有益な機会を頂けたこと、深く御礼申し上げます。本論文をまとめるにあたり、御助言いただき、また、副査をお引き受けいただいた北海道大学大学院情報科学研究院 情報メディア環境学研究室 土橋宜典教授に感謝申し上げます。また、本研究の遂行において、多大なる御助力を賜りました北海道大学大学院情報科学研究院メディアダイナミクス研究室 藤後廉准教授に心より御礼申し上げます。研究活動のみならず公私に渡り様々な御助言いただけたこと、ご多忙の中においても真剣に対応していただきましたこと、感謝申し上げます。本研究室のゼミにおいて有益なフィードバックを提供してくださった北海道大学大学院情報科学研究院メディア創生学研究室 前田圭介准教授、北海道大学総合IR本部 斉藤直輝助教、北海道大学 数理・データサイエンス教育研究センター 李広 特任助教に感謝申し上げます。加えて、本研究の遂行にあたり音楽情報学の観点より御助言いただいた国立研究開発法人産業技術総合研究所 後藤真孝 上級主席研究員をはじめとした人間情報インタラクション研究部門メディアインタラクション研究グループの皆様および経済学の観点より御助言いただいた大阪公立大学大学院経営学研究科 中川慧教授に感謝申し上げます。

在学中、多くの御協力を賜りました北海道大学大学院情報科学院メディアダイナミクス研究室の先輩、同輩ならびに後輩学生の皆様に感謝申し上げます。皆様と共に高め合うことで、私は約6年間研究を成し遂げることができました。

最後に、日ごろから私の大学院生活を応援してくださった友人たち、そして越後魚沼の地から私の決断した道を暖かく見守り続けてくださった家族に感謝の意を表し、謝辞とさせていただきます。

参考文献

- [1] C. Matt, T. Hess, and A. Benlian, “Digital transformation strategies,” *Business and Information Systems Engineering*, vol. 57, pp. 339–343, 2015.
- [2] D. Wang, S. Deng, and G. Xu, “Sequence-based context-aware music recommendation,” *Information Retrieval Journal*, vol. 21, no. 2, pp. 230–252, 2018.
- [3] A. Borchers, J. Herlocker, J. Konstan, and J. Reidl, “Ganging up on information overload,” *Computer*, vol. 31, no. 4, pp. 106–108, 1998.
- [4] D. Bawden and L. Robinson, “The dark side of information: Overload, anxiety and other paradoxes and pathologies,” *Journal of Information Science*, vol. 35, no. 2, pp. 180–191, 2009.
- [5] G. Häubl and V. Trifts, “Consumer decision making in online shopping environments: The effects of interactive decision aids,” *Marketing Science*, vol. 19, no. 1, pp. 4–21, 2000.
- [6] M. Montaner, B. López, and J. L. De La Rosa, “A taxonomy of recommender agents on the internet,” *Artificial Intelligence Review*, vol. 19, no. 4, pp. 285–330, 2003.
- [7] J. Lu, D. Wu, M. Mao, W. Wang, and G. Zhang, “Recommender system application developments: A survey,” *Decision Support Systems*, vol. 74, pp. 12–32, 2015.
- [8] S. Zhang, L. Yao, A. Sun, and Y. Tay, “Deep learning based recommender system: A survey and new perspectives,” *ACM Computing Surveys*, vol. 52, no. 1, pp. 1–38, 2019.
- [9] R. Glauber and A. Loula, “Collaborative filtering vs. content-based filtering: Differences and similarities,” *arXiv preprint arXiv:1912.08932*, 2019.
- [10] X. Su and T. M. Khoshgoftaar, “A survey of collaborative filtering techniques,” *Advances in Artificial Intelligence*, vol. 2009, no. 1, pp. 421425, 2009.

- [11] Y. Koren, S. Rendle, and R. Bell, “Advances in collaborative filtering,” *Recommender Systems Handbook*, pp. 91–142, 2021.
- [12] J. L. Herlocker, J. A. Konstan, and J. Riedl, “Explaining collaborative filtering recommendations,” in *Proceedings of the ACM Conference on Computer Supported Cooperative Work*, 2000, pp. 241–250.
- [13] J. Gope and S. K. Jain, “A survey on solving cold start problem in recommender systems,” in *Proceedings of the International Conference on Computing, Communication and Automation*, 2017, pp. 133–138.
- [14] A. I. Schein, A. Popescul, L. H. Ungar, and D. M. Pennock, “Methods and metrics for cold-start recommendations,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2002, pp. 253–260.
- [15] R. Van Meteren and M. Van Someren, “Using content-based filtering for recommendation,” in *Proceedings of the Machine Learning in the New Information Age: MLnet*, 2000, vol. 30, pp. 47–56.
- [16] J. Son and S. B. Kim, “Content-based filtering for recommendation systems using multiattribute networks,” *Expert Systems with Applications*, vol. 89, pp. 404–412, 2017.
- [17] P. B. Thorat, R. M. Goudar, and S. Barve, “Survey on collaborative filtering, content-based filtering and hybrid recommendation system,” *International Journal of Computer Applications*, vol. 110, no. 4, 2015.
- [18] G. Geetha, M. Safa, C. Fancy, and D. Saranya, “A hybrid approach using collaborative filtering and content based filtering for recommender system,” *Journal of Physics: Conference Series*, vol. 1000, pp. 012101, 2018.
- [19] E. Çano and M. Morisio, “Hybrid recommender systems: A systematic literature review,” *Intelligent Data Analysis*, vol. 21, no. 6, pp. 1487–1524, 2017.
- [20] J.-C. Zhang, A. M. Zain, K.-Q. Zhou, X. Chen, and R.-M. Zhang, “A review of recommender systems based on knowledge graph embedding,” *Expert Systems with Applications*, p. 123876, 2024.

- [21] J. Chicaiza and P. Valdiviezo-Diaz, “A comprehensive survey of knowledge graph-based recommender systems: Technologies, development, and contributions,” *Information*, vol. 12, no. 6, pp. 232, 2021.
- [22] Y. Cao, X. Wang, X. He, Z. Hu, and T.-S. Chua, “Unifying knowledge graph learning and recommendation: Towards a better understanding of user preferences,” in *Proceedings of the World Wide Web Conference*, 2019, pp. 151–161.
- [23] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, “Item-based collaborative filtering recommendation algorithms,” in *Proceedings of the International Conference on World Wide Web*, 2001, pp. 285–295.
- [24] Y. Koren, R. Bell, and C. Volinsky, “Matrix factorization techniques for recommender systems,” *Computer*, vol. 42, no. 8, pp. 30–37, 2009.
- [25] S. Rendle, C. Freudenthaler, Z. Gantner, and L. Schmidt-Thieme, “BPR: Bayesian personalized ranking from implicit feedback,” in *Proceedings of the Conference on Uncertainty in Artificial Intelligence*, 2009, pp. 452–461.
- [26] X. Wang, X. He, M. Wang, F. Feng, and T.-S. Chua, “Neural graph collaborative filtering,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019, pp. 165–174.
- [27] X. He, K. Deng, X. Wang, Y. Li, Y. Zhang, and M. Wang, “LightGCN: Simplifying and powering graph convolution network for recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020, pp. 639–648.
- [28] D. Lee, S. Kang, H. Ju, C. Park, and H. Yu, “Bootstrapping user and item representations for one-class collaborative filtering,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 317–326.
- [29] Y. Yang, L. Wu, R. Hong, K. Zhang, and M. Wang, “Enhanced graph learning for collaborative filtering via mutual information maximization,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 71–80.

- [30] W. Fan, X. Liu, W. Jin, X. Zhao, J. Tang, and Q. Li, “Graph trend filtering networks for recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2022, pp. 112–121.
- [31] M. Balabanović and Y. Shoham, “Fab: Content-based, collaborative recommendation,” *Communications of the ACM*, vol. 40, no. 3, pp. 66–72, 1997.
- [32] R. J. Mooney and L. Roy, “Content-based book recommending using learning for text categorization,” in *Proceedings of the ACM Conference on Digital Libraries*, 2000, pp. 195–204.
- [33] E. P. Olvera and D. Godoy, “Evaluating term weighting schemes for content-based tag recommendation in social tagging systems,” *IEEE Latin America Transactions*, vol. 10, no. 4, pp. 1973–1980, 2012.
- [34] W. Qu, K.-S. Song, Y.-F. Zhang, S. Feng, D.-L. Wang, and G. Yu, “A novel approach based on multi-view content analysis and semi-supervised enrichment for movie recommendation,” *Journal of Computer Science and Technology*, vol. 28, no. 5, pp. 776–787, 2013.
- [35] A. Da ’ u, N. Salim, and R. Idris, “An adaptive deep learning method for item recommendation system,” *Knowledge-Based Systems*, vol. 213, pp. 106681, 2021.
- [36] S. Rendle, “Factorization machines,” in *Proceedings of the IEEE International Conference on Data Mining*, 2010, pp. 995–1000.
- [37] M. Kula, “Metadata embeddings for user and item cold-start recommendations,” *arXiv preprint arXiv:1507.08439*, 2015.
- [38] H. Guo, R. Tang, Y. Ye, Z. Li, and X. He, “DeepFM: A factorization-machine based neural network for ctr prediction,” in *Proceedings of the International Joint Conference on Artificial Intelligence*, 2017, pp. 1725–1731.
- [39] C. Bernardis and P. Cremonesi, “NFC: A deep and hybrid item-based model for item cold-start recommendation,” *User Modeling and User-Adapted Interaction*, pp. 1–34, 2022.

- [40] F. Zhang, N. J. Yuan, D. Lian, X. Xie, and W.-Y. Ma, “Collaborative knowledge base embedding for recommender systems,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2016, pp. 353–362.
- [41] X. Wang, X. He, Y. Cao, M. Liu, and T.-S. Chua, “KGAT: Knowledge graph attention network for recommendation,” in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, 2019, pp. 950–958.
- [42] Y. Xian, Z. Fu, H. Zhao, Y. Ge, X. Chen, Q. Huang, S. Geng, Z. Qin, G. De Melo, S. Muthukrishnan, et al., “CAFE: Coarse-to-fine neural symbolic reasoning for explainable recommendation,” in *Proceedings of the ACM International Conference on Information and Knowledge Management*, 2020, pp. 1645–1654.
- [43] C.-Y. Tai, L.-Y. Huang, C.-K. Huang, and L.-W. Ku, “User-centric path reasoning towards explainable recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2021, pp. 879–889.
- [44] Y. Xian, Z. Fu, S. Muthukrishnan, G. De Melo, and Y. Zhang, “Reinforcement knowledge graph reasoning for explainable recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2019, pp. 285–294.
- [45] H. Wang, F. Zhang, J. Wang, M. Zhao, W. Li, X. Xie, and M. Guo, “RippleNet: Propagating user preferences on the knowledge graph for recommender systems,” in *Proceedings of the ACM International Conference on Information and Knowledge Management*, 2018, pp. 417–426.
- [46] Y. Li, X. Sun, H. Chen, S. Zhang, Y. Yang, and G. Xu, “Attention is not the only choice: Counterfactual reasoning for path-based explainable recommendation,” *IEEE Transactions on Knowledge and Data Engineering*, vol. 36, no. 9, pp. 4458–4471, 2024.
- [47] A. Bordes, N. Usunier, A. Garcia-Duran, J. Weston, and O. Yakhnenko, “Translating embeddings for modeling multi-relational data,” *Advances in Neural Information Processing Systems*, vol. 26, pp. 2787–2795, 2013.

- [48] K. Cho, B. Van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, “Learning phrase representations using rnn encoder-decoder for statistical machine translation,” *arXiv preprint arXiv:1406.1078*, 2014.
- [49] P. Wang, Y. Fan, L. Xia, W. X. Zhao, S. Niu, and J. Huang, “KERL: A knowledge-guided reinforcement learning model for sequential recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2020, pp. 209–218.
- [50] S. Ben-Elazar, G. Lavee, N. Koenigstein, O. Barkan, H. Berezin, U. Paquet, and T. Zaccai, “Groove radio: A bayesian hierarchical model for personalized playlist generation,” in *Proceedings of the ACM International Conference on Web Search and Data Mining*, 2017, pp. 445–453.
- [51] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Music playlist generation based on graph exploration using reinforcement learning,” in *Proceedings of the IEEE Global Conference on Life Sciences and Technologies*, 2021, pp. 53–54.
- [52] S.-Y. Shih and H.-Y. Chi, “Automatic, personalized, and flexible playlist generation using reinforcement learning,” *arXiv preprint arXiv:1809.04214*, 2018.
- [53] B. W. Turnbull, “The empirical distribution function with arbitrarily grouped, censored and truncated data,” *Journal of the Royal Statistical Society: Series B (Methodological)*, vol. 38, no. 3, pp. 290–295, 1976.
- [54] R. S. Sutton and A. G. Barto, *Reinforcement learning: An introduction*, MIT press, 2018.
- [55] C.-W. Chen, P. Lamere, M. Schedl, and H. Zamani, “Recsys challenge 2018: Automatic music playlist continuation,” in *Proceedings of the ACM Conference on Recommender Systems*, 2018, pp. 527–528.
- [56] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Deep reinforcement learning-based music recommendation with knowledge graph using acoustic features,” *ITE Transactions on Media Technology and Applications*, vol. 10, no. 1, pp. 8–17, 2022.

- [57] A. Ferraro, Y. Kim, S. Lee, B. Kim, N. Jo, S. Lim, S. Lim, J. Jang, S. Kim, X. Serra, et al., “Melon playlist dataset: A public dataset for audio-based playlist generation and music tagging,” in *Proceedings of the IEEE International Conference on Acoustics, Speech and Signal Processing*, 2021, pp. 536–540.
- [58] D. Jannach, L. Lerche, and I. Kamehkhosh, “Beyond “hitting the hits” generating coherent music playlist continuations with the right tracks,” in *Proceedings of the ACM Conference on Recommender Systems*, 2015, pp. 187–194.
- [59] X. Wang, D. Wang, C. Xu, X. He, Y. Cao, and T.-S. Chua, “Explainable reasoning over knowledge graphs for recommendation,” in *Proceedings of the AAAI Conference on Artificial Intelligence*, 2019, vol. 33, pp. 5329–5336.
- [60] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, “Dropout: A simple way to prevent neural networks from overfitting,” *The Journal of Machine Learning Research*, vol. 15, no. 1, pp. 1929–1958, 2014.
- [61] S. K. Nair, L. S. Thakur, and K.-W. Wen, “Near optimal solutions for product line design and selection: Beam search heuristics,” *Management Science*, vol. 41, no. 5, pp. 767–785, 1995.
- [62] G. Vigliensoni and I. Fujinaga, “The music listening histories dataset,” in *Proceedings of the International Society for Music Information Retrieval Conference*, 2017, pp. 96–102.
- [63] M. Bakhshizadeh, A. Moeini, M. Latifi, and M. T. Mahmoudi, “Automated mood based music playlist generation by clustering the audio features,” in *Proceedings of the International Conference on Computer and Knowledge Engineering*, 2019, pp. 231–237.
- [64] L. Tong, “Personal recommendation based on community partition of bipartite network,” in *Proceedings of the International Conference on Cloud Computing and Big Data*, 2015, pp. 336–341.
- [65] Q. Ai, V. Azizi, X. Chen, and Y. Zhang, “Learning heterogeneous knowledge base embeddings for explainable recommendation,” *Algorithms*, vol. 11, no. 9, pp. 137, 2018.

- [66] D. Liang, M. Zhan, and D. P. Ellis, “Content-aware collaborative music recommendation using pre-trained neural networks,” in *Proceedings of the International Society for Music Information Retrieval Conference*, 2015, pp. 295–301.
- [67] Y. Hou, J. Zhang, Z. Lin, H. Lu, R. Xie, J. McAuley, and W. X. Zhao, “Large language models are zero-shot rankers for recommender systems,” in *Proceedings of the European Conference on Information Retrieval*, 2024, pp. 364–381.
- [68] C. Yang, R. Zheng, X. Chen, and H. Wang, “Content recommendation with two-level transe predictors and interaction-aware embedding enhancement: An information seeking behavior perspective,” *Information Processing and Management*, vol. 60, no. 4, pp. 103402, 2023.
- [69] R. J. Williams, “Simple statistical gradient-following algorithms for connectionist reinforcement learning,” *Machine Learning*, vol. 8, pp. 229–256, 1992.
- [70] F. M. Harper and J. A. Konstan, “The MovieLens datasets: History and context,” *ACM Transactions on Interactive Intelligent Systems*, vol. 5, no. 4, pp. 1–19, 2015.
- [71] G. Balloccu, L. Boratto, C. Cancedda, G. Fenu, and M. Marras, “Knowledge is power, understanding is impact: Utility and beyond goals, explanation quality, and fairness in path reasoning recommendation,” in *Proceedings of the European Conference on Information Retrieval*, 2023, pp. 3–19.
- [72] M. Schedl, “The LFM-1b dataset for music retrieval and recommendation,” in *Proceedings of the ACM on International Conference on Multimedia Retrieval*, 2016, pp. 103–110.
- [73] F. Cai, Z. Pan, C. Song, and X. Zhang, “Exploring latent connections in graph neural networks for session-based recommendation,” *Information Retrieval Journal*, vol. 25, no. 3, pp. 329–363, 2022.
- [74] W.-C. Kang and J. McAuley, “Self-attentive sequential recommendation,” in *Proceedings of the IEEE International Conference on Data Mining*, 2018, pp. 197–206.
- [75] Z. Pan, F. Cai, Y. Ling, and M. d. Rijke, “Rethinking item importance in session-based recommendation,” in *Proceedings of the International ACM SIGIR Conference on Research*

and Development in Information Retrieval, 2020, pp. 1837–1840.

- [76] B. Hidasi, A. Karatzoglou, L. Baltrunas, and D. Tikk, “Session-based recommendations with recurrent neural networks,” in *Proceedings of the International Conference on Learning Representations*, 2015, pp. 1–10.
- [77] S. Yang, W. Ma, P. Sun, M. Zhang, Q. Ai, Y. Liu, and M. Cai, “Common sense enhanced knowledge-based recommendation with large language model,” *arXiv preprint arXiv:2403.18325*, 2024.
- [78] L. Xu, Z. Tian, G. Zhang, J. Zhang, L. Wang, B. Zheng, Y. Li, J. Tang, Z. Zhang, Y. Hou, X. Pan, W. X. Zhao, X. Chen, and J. Wen, “Towards a more user-friendly and easy-to-use uenchmark library for recommender systems,” in *Proceedings of the International ACM SIGIR Conference on Research and Development in Information Retrieval*, 2023, pp. 2837–2847.

著者の研究業績

(A) 学会誌

- [A-1] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Deep reinforcement learning-based music recommendation with knowledge graph using acoustic features,” *ITE Transactions on Media Technology and Applications* (2023 IF: 0.5), vol. 10, no. 1, pp. 8–17, 2022.
- [A-2] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Controllable music playlist generation based on knowledge graph and reinforcement learning,” *Sensors* (2023 IF: 3.4), vol. 22, no. 10: 3277, 2022.
- [A-3] 櫻井慶悟, 前田圭介, 藤後廉, 小川貴弘, 長谷山美紀, “地下鉄トンネル点検時の一人称視点映像を用いた Vision Transformer に基づく変状検出,” 土木学会 AI・データサイエンス論文集, vol. 3, no. J2, pp. 470–478, 2022.
- [A-4] 櫻井慶悟, 前田圭介, 藤後廉, 小川貴弘, 長谷山美紀, “地下鉄トンネル点検技術者の新規変状の発見率向上に向けた一人称視点映像からの変状検出,” 土木学会 AI・データサイエンス論文集, vol. 4, no. 3, pp. 393–401, 2023.

(B) 国際学会

- [B-1] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Music playlist generation based on reinforcement learning using acoustic feature map,” in *Proceedings of the IEEE Global Conference on Consumer Electronics (IEEE GCCE)*, pp. 942–943, 2020.
- [B-2] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Music playlist generation based on graph exploration using reinforcement learning,” in *Proceedings of the IEEE Global Conference on Life Sciences and Technologies (IEEE LifeTech)*, pp. 53–54, 2021.
- [B-3] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “User background information-aware music recommendation with reinforcement learning-based knowledge graph exploration,”

- in *Proceedings of the IEEE International Conference on Consumer Electronics-Taiwan (IEEE ICCE-TW)*, pp. 1–2, 2021.
- [B-4] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Listener recommendation for artist based on knowledge graph and reinforcement learning,” in *Proceedings of the IEEE Global Conference on Consumer Electronics (IEEE GCCE)*, pp. 202–203, 2021.
- [B-5] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Explainable artist recommendation based on reinforcement knowledge graph exploration,” in *Proceedings of the International Workshop on Advanced Image Technology (IWAIT)*, pp. 91–96, 2022.
- [B-6] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Knowledge-guided sequential recommendation with reinforcement learning using empirical distribution function,” in *Proceedings of the IEEE Global Conference on Life Sciences and Technologies (IEEE LifeTech)*, pp. 187–188, 2022.
- [B-7] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Bundle detection based on graph convolutional network considering categorical hierarchical relationship,” in *Proceedings of the IEEE Global Conference on Consumer Electronics (IEEE GCCE)*, pp. 827–828, 2023.
- [B-8] Y. Moroto, R. Yanagi, N. Ogawa, K. Kamikawa, K. Sakurai, R. Togo, K. Maeda, T. Ogawa, and M. Haseyama, “Personalized content recommender system via non-verbal interaction using face mesh and facial expression,” in *Proceedings of the ACM International Conference on Multimedia (ACM MM)*, pp. 9399–9401, 2023.
- [B-9] J. Zhu, K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Discriminator-enhanced music generation based on multitrack music transformer,” in *Proceedings of the IEEE International Conference on Consumer Electronics-Taiwan (IEEE ICCE-TW)*, pp. 363–364, 2024.
- [B-10] J. Zhu, K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “Structured polyphonic music generation with diffusion transformer,” in *Proceedings of the IEEE Global Conference on Consumer Electronics (IEEE GCCE)*, pp. 8–9, 2024.

[B-11] J. Zhu, K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “MMT-BERT: Chord-aware symbolic music generation based on multitrack music transformer and MusicBERT,” in *Proceedings of the International Society for Music Information Retrieval Conference (ISMIR)*, pp. 470–477, 2024.

[B-12] K. Sakurai, R. Togo, T. Ogawa, and M. Haseyama, “LLM is knowledge graph reasoner: LLM’s intuition-aware knowledge graph reasoning for cold-start sequential recommendation,” in *Proceedings of the European Conference on Information Retrieval (ECIR)*, pp. 263–278, 2025.

(C) 技術報告

[C-1] 櫻井慶悟, 藤後廉, 小川貴弘, 長谷山美紀, “ユーザの嗜好を考慮した強化学習と知識グラフに基づく楽曲プレイリスト生成に関する検討,” 映像情報メディア学会技術報告, vol. 46, no. 6, pp. 109–112, 2022.

[C-2] 櫻井慶悟, 藤後廉, 小川貴弘, 長谷山美紀, “Graph masked autoencoder を用いた知識グラフ推論に基づく説明可能性のある推薦に関する検討,” 2024 年度人工知能学会全国大会, pp. 1–4, 2024.

[C-3] 櫻井慶悟, 岡村洋希, 藤後廉, 前田圭介, 小川貴弘, 長谷山美紀, “[特別講演] 画像生成および補完推薦技術を活用したインテリアコーディネート支援に関する検討,” 映像情報メディア学会技術報告, vol. 49, no. 4, pp. 408–412, 2025.

[C-4] 櫻井慶悟, 小川貴弘, 長谷山美紀, 阿南晏樹, 中川慧 “効用を考慮した行列分解に基づく株式推薦,” 2025 年度人工知能学会全国大会, pp. 1–4, 2025.

(D) 講演

[D-1] 櫻井慶悟, 藤後廉, 小川貴弘, 長谷山美紀, “自己組織化マップを用いた強化学習に基づく楽曲プレイリスト生成に関する検討,” 令和 2 年度 電気・情報関係学会北海道支部連合大会 講演論文集, pp. 100–101, 2020.

[D-2] 櫻井慶悟, 藤後廉, 小川貴弘, 長谷山美紀, “メタデータを考慮した楽曲グラフ探索に基づく強化学習を用いた楽曲プレイリスト生成に関する検討,” 令和 3 年度 電気・情報関係学会北海道支部連合大会 講演論文集, pp. 108–109, 2021.

[D-3] K. Sakurai, ““Listener recommendation for artists and effective method with high explainability,” 1st Hokkaido Young Professionals Workshop, Muroran Institute of Technology, 2021.

(E) 受賞

[E-1] The 1st Hokkaido Young Professionals Workshop Best Student Presentation Award (2021 年 10 月)

[E-2] 令和 3 年度 電気・情報関係学会北海道支部連合大会 若手優秀論文発表賞 (2021 年 11 月)

[E-3] 2022 International Workshop on Advanced Image Technology Best Paper Award (2022 年 1 月)

[E-4] 映像情報メディア学会 優秀研究発表賞 (2022 年 12 月)

[E-5] 電子情報通信学会北海道支部 学生奨励賞 (2023 年 3 月)

[E-6] 北海道大学大学院情報科学院 学院長賞 (2023 年 3 月)

[E-7] 土木学会 AI・データサイエンス論文賞 (2023 年 12 月)

[E-8] IEEE International Conference on Consumer Electronics – Taiwan Best Paper Award 1st Place (2024 年 7 月)

[E-9] IEEE Global Conference on Consumer Electronics Presentation Award (2024 年 10 月)

[E-10] 人工知能学会 2025 年度全国大会優秀賞 (2025 年 7 月)

(F) 奨学金・助成金

[F-1] 日本証券奨学財団 奨学金 (2021 年 4 月～2022 年 3 月)

[F-2] 似鳥国際奨学財団 北海道大学みらい IT 人財奨学金 (2022 年 4 月～2023 年 3 月)

[F-3] 日本学生支援機構 第一種奨学金 特に優れた業績による奨学金返還全額免除 (2021 年 4 月～2023 年 3 月)

[F-4] NEC C&C 財団 研究員助成 (2023 年 4 月～2024 年 3 月)

[F-5] 日本学術振興会 特別研究員 DC1 (2023 年 4 月～2025 年 9 月)

[F-6] S k y 大浦 ICT 奨学財団 (2024 年 4 月～2025 年 9 月)

(G) メディア紹介

[G-1] “知識グラフと強化学習を組み合わせた説明可能な新しいアーティスト推薦システム,” AI-SCHOLAR, 2022/08/18. (K. Sakurai et al., IWAIT2022 に関して)

[G-2] “「ChatGPT」など大規模言語モデルの仕組みと土木領域での可能性,” BUILT-ITmedia, 2023/06/20. (櫻井慶悟 他, 土木学会 AI・データサイエンス論文集に関して)

[G-3] “土木における AI 活用の現状と将来,” けんせつ Plaza, 2024/07/17. (櫻井慶悟 他, 土木学会 AI・データサイエンス論文集に関して)