

修士論文概要書

Master's Thesis Summary

Date of submission: 01/11/2022 (MM/DD/YYYY)

専攻名 (専門分野) Department	経営システム工学専攻	氏名 Name	楊 冠宇 Yang Guanyu	指導 教員 Advisor	後藤 正幸 印 Seal
研究指導名 Research guidance	情報数理応用研究	学籍番号 Student ID number	CD 5221C044-5		
研究題目 Title	知識グラフと強化学習に基づく説明可能性を考慮した効率的な推薦アルゴリズム An Efficient Recommendation Algorithm Considering Explainability Based on Knowledge Graph and Reinforcement Learning				

1. 研究背景と目的

近年、機械学習モデルを用いたユーザの行動分析および商品推薦システムの研究が行われ、顧客の商品購買支援や売上向上の面で一定の成果を上げている。一方で、分析者はブラックボックスのモデルよりも説明可能なモデルを好む傾向がある。これは、モデル内部を明らかにすることで施策立案の一助としたり、推薦理由の提示による満足度向上が見込めるためである。そのうちのひとつとして、ユーザ・アイテム・補助情報の間の関係性をモデル化する知識グラフ [1] を用いたモデルの有効性が示されている。

知識グラフとは、複数種類のノード (エンティティ) をエッジ (リレーション) で結んだ異種知識ネットワークである。知識グラフを推薦システムに活用することで、ユーザとアイテムの購買等の相互作用に関する情報だけでなく、その他の補助情報を考慮した推薦を行うことが可能となる。さらに強化学習のフレームワークを用いて知識グラフを学習し、推薦に至るまでの関係を解釈する手法として PGPR (Policy-Guided Path Reasoning) [2] が提案されている。PGPR では、アイテムとユーザとの間に知識グラフ上における有効な経路を見つけることにより、各推薦結果に対する意思決定プロセスを解釈可能である。

しかし、一般に知識グラフの学習には膨大な量のデータが必要となり、さらに推定が必要なパラメータも多く、計算コストが非常に高いという課題がある。そこで本研究では有効な経路を効率的に探索可能な、分析者の事前知識を活用したヒューリスティックな経路探索アルゴリズムを提案する。これにより、知識グラフを用いた PGPR による推薦の精度を向上させると共に、計算量を削減することを本研究の目的とする。最後に実際のデータに提案手法を適用し、アイテムの推薦精度と計算コスト、および解釈性の観点から評価を行う。

2. 準備

2.1. 知識グラフ

知識グラフとは、エンティティとリレーションを用いて、アイテムの属性情報や、アイテム間の関係性などを表現した有向グラフである。エンティティは主に、ユーザ、アイ

テム、カテゴリーなどの実際の物体や概念を表しており、リレーションはそれらの概念を繋がる関係性を表している。知識グラフは、先頭のエンティティ h 、リレーション r 、末尾のエンティティ t の3つの組み合わせからなる「トリプレット」 (h, r, t) の集合として表現される。

2.2. 知識グラフを利用した推薦システム

知識グラフは、異なるエンティティの関係性に関する豊富な情報をキャプチャできるため、推薦精度と共に説明可能性を高めることが期待される。例えば知識グラフに映画、俳優、監督、およびその間のリレーションに関する情報が含まれている場合、ユーザの過去の視聴履歴、好み、およびその他の関連する要因を基に、これらの情報を推薦理由としてユーザに映画を提示することが考えられる。しかし、知識グラフ自体は単に様々な種類の情報の集合体であり、適切なアイテムを推薦する機能は持たない。そこで、グラフ埋め込み手法によって知識グラフを表現し、それによって元のアイテムやユーザ表現の意味情報を拡張する手法が研究されている。その代表的手法である TransE[3] は、先頭エンティティのベクトルとリレーションのベクトルの和が、末尾エンティティのベクトルに近づくようにそれぞれの埋め込み表現を学習する手法である。すなわち、トリプレット (h, r, t) について、 $h + r \approx t$ を満たすようにベクトル表現を学習する。その他、TransH[3] や、TransR[3]、TransD[3] など、様々な拡張手法が提案されている。しかし、2つのエンティティの間には多段階なリレーションからなる複合リレーションが存在し、さらに、このようなリレーションが複数存在する場合もある。従来の埋め込み表現に基づく推薦手法では、こういった多段階な関係をモデル化することが難しく、推薦理由を説明することは困難である。

そこで Yu ら [4] は、ユーザとアイテム間の相互関係を meta-path によって表現した知識グラフの経路を活用した推薦システムを提案し、多段階な関係をモデル化した。例えば、映画の推薦タスクでは、ユーザとアイテムの間にある関係性を「user - item - actor - item」または「user - user - item」などの meta-path として、それぞれ異なるパター

ンと定義してモデル化する。モデル化された meta-path (知識グラフ上の経路) を辿ることにより、ユーザへアイテムを推薦する理由を示すことが可能になる。しかし、知識グラフの経路を活用した推薦システムでは meta-path を事前に定義する必要があり、自動的に推薦経路を探索することが困難である。また、知識グラフに含まれる全ての経路の中から有効な経路を抽出する必要があり、大量のデータが蓄積されているようなサービスへの適用は計算コストの面で現実的ではない。

2.3. 強化学習

近年、推薦システムにおいて、強化学習 [5] を援用する研究が盛んに行われている [2]。強化学習は、幅広い分野において、エージェントが環境と相互作用しながら、絶え間ない試行錯誤により逐次決定問題に対する最適戦略を学習する機械学習手法である。近年、強化学習は様々な分野で活用されており、例えばゲームやロボット、自然言語処理の分野において深層学習やマルコフ決定過程 [6] (MDP) を用いた強化学習が実用化されている。MDP を用いた強化学習では、エージェントは時間をかけて環境と相互作用するものと仮定している。エージェントと環境との相互作用を、状態・動作・遷移確率・報酬・割引率という5つの要素に注目して逐次学習する。ここで、エージェントの目標は、各状態における長期報酬の期待値を最大化することである。

2.4. PGPR(Policy-Guided Path Reasoning)

2.4.1. 概要

Xian ら [2] は 2.2 節で示した問題を解決するため、PGPR を提案している。PGPR は有効な推薦経路 (推薦理由) の発見を目的として、強化学習のアイデアを援用し、効率的な推薦経路を探索・抽出し、解釈性の向上と計算コストの削減を実現する手法である。具体的には、知識グラフ上で有効な推薦経路を探索する問題を、MDP に基づいて定式化している。

まず、「ターゲットユーザ」、「現在のステップで探索したエンティティ」、「現在ステップの前の全ての探索履歴」からなる要素の集合を「状態」と定義する。また、現在のステップまで探索したエンティティから、リレーションを通して次のエンティティに遷移する行為を「行動」と定義する。これらの行動が実行される確率は、ニューラルネットワークで表現することができる。また、報酬関数を定義し、報酬の期待値の合計を最大化する推薦経路を MDP により学習する。さらに剪定などのアプローチを用いることにより、必要性の低い経路を探索するコストの削減も可能となる。実際に PGPR を用いることで、知識グラフ上で有効な推薦経路を見つけることが可能となり、解釈性のある説明可能な推薦システムの構築が可能となることが示されている。

2.4.2. 定式化

PGPR は、知識グラフ G 、ユーザ u 、経路数の上限 K 、推薦アイテム数 N を入力し、総数 N 個の推薦アイテム $i_n \in \mathcal{I}$ と各アイテム i_n と関連している推薦経路 $p_k(u, i_n)$ を出力するモデルである ($n = 1, 2, \dots, N$)。PGPR では、式 (1) のように、 k 個のリレーションで結ばれた $k+1$ 個のエンティティのシーケンス (多段階な関係性) を「経路」と定義する。 e_i と r_i は、それぞれ i 番目のリレーションとエンティティを表す。

$$p_k(e_0, e_k) = \left\{ e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k \right\} \quad (1)$$

なお、一方向の経路は式 (2) で表される。

$$\vec{p}_k(e_0, e_k) = \left\{ e_0 \xrightarrow{r_1} e_1 \xrightarrow{r_2} \dots \xrightarrow{r_k} e_k \right\} \quad (2)$$

式 (2) で表される一方向の経路のリレーションの集合 $\tilde{r}_k = \{r_1, \dots, r_k\}$ を k -hop 経路と呼ぶ。また、経路の探索中、多様なシナリオをより包括的にとらえるため、前のエンティティへの振り返りの発生をある程度許容するものとする。さらに、フォワード関係とバックワード関係の両方を持つ特殊なケースを考慮し、経路の一部の関係を反転させた一次逆転 k -hop 経路パターン $\tilde{r}_{k,j} = \{r_1, \dots, r_j, r_{j+1}, \dots, r_k \mid j \in [0, k]\}$ を導入する。具体的には、以下のような経路を表す。

$$e_0 \xrightarrow{r_1} \dots \xrightarrow{r_j} e_j \xleftarrow{r_{j+1}} e_{j+1} \xleftarrow{r_{j+2}} \dots \xleftarrow{r_k} e_k \quad (3)$$

しかし、知識グラフにおいて学習されるデータ量は膨大であるため、PGPR では経路探索の際、候補である行動集合からランダムに1つの行先を決め、次のエンティティに遷移する形で計算が行われており、計算効率が悪いという問題がある。

次に、PGPR で用いる MDP の定式化について説明する。ユーザ u の時刻 t の状態を $s_t = (u, e_t, h_t)$ で表すこととする。ここで、 h_t は時刻 t までの探索履歴、 e_t 時刻 t において到達したエンティティである。時刻 t の行動候補の集合を $\mathcal{A}_t = \{(r, e) \mid (e_t, r, e) \in G, e \notin \{e_0, \dots, e_{t-1}\}\}$ と定義する。ここで、状態 $s_t = (u, e_t, h_t)$ と行動 $a_t = (r_{t+1}, e_{t+1})$ が与えられたときの、次の状態 s_{t+1} に遷移する確率を $\mathbb{P}[s_{t+1} \mid s_t, a_t]$ で表す。また、最終的な (時刻 t における) 報酬関数は式 (4) で定義される。

$$R_T = \begin{cases} \max \left(0, \frac{f(u, e_T)}{\max_{i \in \mathcal{I}} f(u, i)} \right), & \text{if } e_T \in \mathcal{I} \\ 0, & \text{otherwise} \end{cases} \quad (4)$$

ただし、 $f(\cdot)$ は経路の観点から二つのエンティティの類似度を計算する評価関数である。

$$f(e_0, e_k \mid \tilde{r}_{k,j}) = \left\langle e_0 + \sum_{s=1}^j r_s, e_k + \sum_{s=j+1}^k r_s \right\rangle + b_{e_k} \quad (5)$$

ただし、 b_{e_k} はバイアスであり、 e, r はそれぞれエンティティ e とリレーション r に対応した d 次元のベクトル、 $\langle a, b \rangle$ はベクトル a と b の内積を表す。式 (6) を最大化するような方策 π を求め、最適な経路を明らかにする。ここで、 $\gamma \in [0, 1]$ は割引率であり、空集合 \emptyset は「 t_0 における探索履歴はまだ何もない」ということを表している。

$$J = \mathbb{E}_\pi \left[\sum_{t=0}^{T-1} \gamma^t R_{t+1} \mid s_0 = (u, u, \emptyset) \right] \quad (6)$$

$$f((r, e) \mid u) = f\left(u, e \left[\tilde{r}_{k_e, j} \right] \right) \quad (7)$$

ここで、経路探索の行き先が膨大であるため、PGPR はコスト削減のため、式 (7) の行動剪定関数を用い、行動集合において剪定を行っている。 k_e は $\{u, e\}$ の「有効な」 k -hop 経路パターンになりうる最小の k 、すなわち最短経路である。剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ を式 (8) で表す。

$$\tilde{\mathcal{A}}_{u,t} = \{(r, e) \mid \text{rank}(f((r, e) \mid u)) \leq \alpha, (r, e) \in \mathcal{A}_t\} \quad (8)$$

ここで、 α は剪定下限を決めるパラメータである。剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ は、行動剪定関数によって算出した値の大きさ順の上位 α 位の行動のみを含む集合である。また、図 1 のように、ニューラルネットワークを用いて方策 π を推定する。

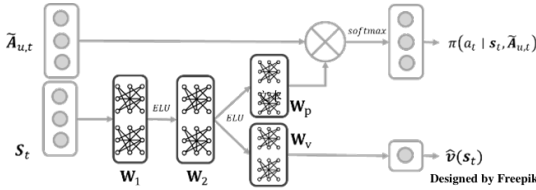


図 1: 方策 π を推定するニューラルネットワークの構造のイメージ

ここで入力には状態 s_t および剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ であり、状態価値関数 \hat{v} の結果 $\hat{v}(s_t)$ と各行動 a を行う確率 (方策) $\pi(a_t \mid s_t, \tilde{\mathcal{A}}_{u,t})$ を計算する。 $\hat{v}(s_t)$ は状態 s_t の結果を表し、剪定が行われない時との比較に使われている。 $\pi(a_t \mid s_t, \tilde{\mathcal{A}}_{u,t})$ は剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ の事前情報のもとで、剪定が行われた時の結果として使われている。実際に、剪定することによる精度の向上が示されている。

3. 提案手法：E-PGPR

PGPR では計算量を減らすために行動空間の剪定アルゴリズムが導入されているが、剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ からランダムに一個の行先を選択し、次のエンティティに遷移することで探索を行っており、多くの計算コストが必要になるという課題がある。そこで本手法では、「逆数型重み配分」を利用して遷移候補の各リレーションに重みを付けることを考える。そして、PGPR の探索方法を拡張し、より効率的な推薦経路探索を実現する E-PGPR (Efficient PGPR) を提案する。

逆数型重み配分とは、分析者の事前知識をベースとして事前に付与された重要度に基づき、重要度の順位の数によって、重み付けする手法である。なお、重みは総和が 1 になるように正規化を行う。具体的には、以下のプロセスにより、遷移先の候補となる (r_1, r_2, \dots, r_l) という l 個のリレーションに重みを付与する。

1. 事前に規定された重要度の順番により、リレーション (r_1, r_2, \dots, r_l) をソートする。
2. ソートされた新しい系列を $(r^*_1, r^*_2, \dots, r^*_n)$ とする。
3. $(r^*_1, p^*_2, \dots, r^*_l)$ に対応した重みの総和が 1 に等しくなるように正規化された降順の重みの系列 $(p^*_1, p^*_2, \dots, p^*_l)$ を作成する。
4. 3. で設定した重みの大きさに応じて、次の遷移先を選択し、次のエンティティに遷移する。

ここで、 l 個のリレーション (r_1, r_2, \dots, r_l) に対し、順位 $(1, 2, \dots, l)$ にそれぞれとある定数 c を加え、逆数を取ることで、式 (9) のようなベクトルが得られる。

$$(p^*_1, p^*_2, \dots, p^*_l) = \left(\frac{1}{1+c}, \frac{1}{2+c}, \dots, \frac{1}{l+c} \right) \quad (9)$$

次に、各要素の和が 1 になるように式 (9) を正規化し、新に得られたベクトル $(p^*_1, p^*_2, \dots, p^*_l)$ を重みベクトルとして用いる。ここで、 $c \geq 0$ は重みベクトルにおける各要素の分散が大きくなりすぎないようにするパラメータ均質因子である。 c によって、ベクトルの要素のばらつきを小さくすることが可能となり、一番大きい重みと一番小さい重みの差が大きくなりすぎないように調整することができる。

E-PGPR は、以上の手順で経路探索の各ステップにおいて、剪定された行動集合 $\tilde{\mathcal{A}}_{u,t}$ の行き先の候補である各リレーションに重み $p^*_1, p^*_2, \dots, p^*_l$ を付与し、重みの大きさによって次のステップの探索を行っている。

4. 検証実験

4.1. 実験条件

提案手法の有効性を検証するために、オープンソースのデータセット Amazon datasets[7] における、化粧品・美容品などのビューティーアイテムに関わる購買履歴データである Amazon beauty を用いて実験を行う。Amazon beauty のデータ概要を表 1 に示す。

表 1: エンティティの種類とその数

エンティティ	エンティティの数
ユーザ	22,363
カテゴリ	248
単語	22,563
メーカー	2,077
アイテム	12,101

また、対象データにおいてリレーションは 8 種類ある。その中で、提案手法においては、購買関係は閲覧やメーカー

情報より重要であるという考えのもと、それぞれ表2のように重要度の順位を事前に人工的に付与する。ここで、「所属する」、「生産される」、「言葉で表現される」はそれぞれアイテムとカテゴリ、メーカー、単語との関係性を表現したリレーションである。また、検証実験における各パラメータは表3のように設定した。

表 2: 含まれるリレーションの種類と設定した重要度

リレーション	重要度
購入する	1st
ともに購入される	2nd
ともに購入する	3rd
ともに閲覧される	4th
所属する	5th
生産される	6th
言葉で表現される	7th
言及する	8th

表 3: 各パラメータの設定値

パラメータ	設定値
最大エポック数	50
バッチサイズ	32
学習率	1×10^{-4}
最大行動数	250
最大経路長さ	3
均質因子	3

4.2. 実験結果

提案手法の計算コストに関する有効性を確認するために、従来と提案が収束するまでの速さ（エポック数）を確認する比較実験を行う。学習の終了条件は、「エポック数が10より大きい」かつ「損失の変化率が閾値 5×10^{-3} より小さい」こととした。初期値を変更し行った10回の実験結果の平均を表4に示す。また、表4において、*は提案手法E-PGPRと従来手法PGPRの結果を比較したt検定の結果、1%の有意差があることを示す。

表 4: 収束するまでの平均エポック数と平均精度

手法	Epoch	NDCG	Recall	HR	Prec
PGPR	18.7	5.213	7.928	13.791	1.624
E-PGPR	10.3*	5.480*	8.395*	14.465*	1.728*

結果より、提案手法の方が収束が速く、かつ精度が向上していることがわかる。

5. 考察

実験結果より、逆数型重み配分を利用し、経路探索の各ステップにおいてそれぞれのリレーションに重みを付けを行う提案手法の方が、ランダム探索を行う従来手法より素早く学習の損失関数が収束することが明らかになった。実応用のためには、計算量や予算の制約から、より低いコストで高精度なモデルを使用することが望ましい場合が多いため、E-PGPRは実務で有効であると考えられる。

また、あるユーザからあるアイテムまでの推薦経路を示した一例を図2に示す。図2に示した事例の場合、ユーザAに育毛飲み薬が推薦される理由は、「ユーザAが過去に



図 2: 抽出した推薦経路の例

育毛シャンプー A を購入したことがあり、その育毛シャンプー A と B、さらにある育毛剤が共に購入される傾向にあり、その育毛シャンプー B と育毛剤が育毛飲み薬と共に購入されるため」であることがわかる。提案手法の推薦システムでは以上の理由に基づき、育毛飲み薬をユーザ A に推薦する。このように、得られた経路を辿ることで、推薦理由を説明することが可能である。

6. まとめと今後の課題

本研究では、知識グラフと強化学習に基づく推薦のための経路探索手法 E-PGPR を提案した。提案手法では、逆数型重み配分を用いることにより、精度が向上すると共に、計算時間コストが軽減されることを確認した。また、実例を用いて、推薦結果が解釈可能であることを確認した。

一方で、PGPR や E-PGPR では、推薦に対するフィードバック（推薦されたアイテムの理由について、どう感じたか）の情報が得られたとしても、モデルの学習に活用することが難しい。今後の課題として、フィードバックの情報を活用することで、より継続的に推薦システムの性能を改善することなどが挙げられる。

参考文献

- [1] X. Wang, X. He, et al. Kgat: Knowledge graph attention network for recommendation. *In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining*, pp. 950–958, 2019.
- [2] Y. Xian, Z. Fu, et al. Reinforcement Knowledge Graph Reasoning for Explainable Recommendation. *In Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 285–294, 2019.
- [3] Q. Wang, Z. Mao, et al. Knowledge Graph Embedding: A Survey of Approaches and Applications. *IEEE Transactions on Knowledge and Data Engineering*, Vol. 29, No. 12, pp. 2724–2743, 2017.
- [4] X. Yu, X. Ren, et al. Personalized Entity Recommendation: A Heterogeneous Information Network Approach. *In Proceedings of the 7th ACM International Conference on Web Search and Data Mining*, pp. 283–292, 2014.
- [5] Y. Li. Deep reinforcement learning: An overview. *arXiv preprint arXiv:1701.07274*, 2017.
- [6] R. Bellman. A markovian decision process. *Indiana University Mathematics Journal*, Vol. 6, No. 4, pp. 679–684, 1957.
- [7] J. McAuley, C. Targett, Q. Shi, et al. Image-based recommendations on styles and substitutes. *In Proceedings of the 38th international ACM SIGIR conference on research and development in information retrieval*, pp. 43–52, 2015.