

補助行動データと明示的な選好データを活用した求人推薦手法

齋藤 陽介[†] 杉山 一成^{††}

[†] 京都大学大学院情報学研究科 〒 606-8501 京都府京都市左京区吉田本町

^{††} 京都大学大学院情報学研究科 〒 606-8501 京都府京都市左京区吉田本町

E-mail: [†]saito@db.soc.i.kyoto-u.ac.jp, ^{††}kaz.sugiyama@i.kyoto-u.ac.jp

あらまし 求人情報のオンライン化が進み、その情報を推薦する重要性が増している。転職サービスの推薦タスクでは、ユーザの求人に対する応募を目的として設定することが多い。その際、一般的に応募以外の求人情報の閲覧や求人のお気に入り登録などの補助行動データと、ユーザの希望条件などの明示的な選好データを利用できる。しかし、いずれかのデータを利用し推薦精度の向上を目指した研究は行われていたものの、両方を活用して、どのように組み合わせるかに取り組んだ研究は、これまでに行われていなかった。そこで、本研究ではオンライン転職サービス上の閲覧やお気に入りといった、応募以外の補助行動データと、ユーザの明示的な選好データの両方を活用した推薦手法を提案する。実際のオンライン転職サービスから得られたデータを使った実験の結果、提案手法は最新の手法よりも高い推薦精度を達成することができた。

キーワード 推薦システム, 求人推薦, 補助行動データ, グラフ表現

1 はじめに

近年、様々なサービスがインターネット上で利用できるようになり、多くの転職活動もオンライン転職サービスを介して行われるようになってきている。こうしたオンライン上の求人市場の拡がりとともに、これを対象とした様々な研究が行われている [12], [14], [18]。求人推薦はこうした研究の1つであり、求職者（ユーザ）の転職サービスにおける行動データや職務経歴を基に、関連性の高い求人を提供することを目的としている。オンライン上に掲載される求人の数は膨大であり、従来の検索機能などを用いて自身に適した求人を探すことは多くの時間を要するため、近年、求人推薦の重要性が増している。Eコマースや映画などの他のドメインにおける推薦と比較して、求人推薦の特徴はいくつか挙げられるが、本研究では (1) 複数種類の行動の存在およびその活用可能性と、(2) ユーザ自身の情報を利用できる点の2点に着目する。(1) について、オンライン転職サービスにおける推薦タスクでは、ユーザ-求人インタラクション行列の値として、ユーザの求人に対する応募を利用することが多い（すなわち、求人に応募があれば1、そうでなければ0）。このとき、応募以外にもユーザは求人の閲覧や、求人のお気に入りの登録などといった、応募に至るまでの補助的な行動を実行できる。これらは補助行動データと呼ばれ、求人市場以外のEコマースなどの分野では、推薦精度の向上に利用されている [10]。これまで補助行動データを活用した求人推薦の研究はなかったが、転職サービス上における応募が、Eコマースの購買イベントや、映画における視聴イベントなどと比較して発生しにくいことを考えると、補助行動データの活用により求人推薦の精度を大きく改善できる可能性がある。(2) について、オンライン転職サービスでは一般的に、ユーザは仕事に対

する自身の希望条件を入力することが多く、こうした明示的な選好情報を推薦で活用できる。

これまでの研究では、補助行動データとユーザの明示的な選好データの両方を同時に活用した求人推薦の手法はなかった。そこで、本研究ではユーザの求人に対する、補助行動データを含む暗黙的フィードバックのデータと、ユーザの仕事に対する希望条件等の選好データの両方を活用した推薦手法を提案する。提案手法は2種類のデータから、それぞれグラフの表現学習手法を用いてユーザと求人の潜在表現を獲得し、それらを統合した潜在表現を距離学習の方法を用いてユーザと適合した求人の距離が近くなるように学習を行う。行動データの学習モジュールでは、個々の行動の潜在表現を得る代わりに、複数の行動の組み合わせの潜在表現を獲得することで、ユーザの求人に対する行動のきめ細かい表現を獲得する。提案手法の有効性を検証するために、国内で実際に運営されているオンライン転職サービスのデータを利用して、評価実験を行った。その結果、提案手法は最新の推薦手法よりも高い推薦精度を達成することができた。

2 関連研究

提案手法は次の3つの研究分野と関連がある。1つは求人推薦、2つめは補助行動データを用いた推薦手法、そして3つめはグラフ表現に基づいた推薦手法である。本章では、それぞれの分野におけるいくつかの関連研究を紹介し、本研究との差異について述べる。

2.1 求人推薦

求人推薦はユーザのオンライン転職サービスにおける活動や職務経歴書を基に、彼らに対し関連性の高い求人を提供するこ

とを目的としている [17], [19]. 映画や音楽といった他の分野の推薦タスクと異なり, 求人推薦はユーザの職務経歴書や希望条件などのユーザ自身の情報を利用できる点が特徴である. これは求人市場が企業とユーザをマッチングさせるための市場であり, 企業側も採用に値するユーザかどうか判断するための情報が必要だからである. 求人推薦に関しては, これまで多くの研究が様々なアプローチの推薦手法を提案してきた. Dave ら [5] は, ユーザの職務経歴情報から仕事の遷移ネットワーク・仕事とスキルの関係ネットワーク・スキルの共起ネットワーク, の3つのネットワークを作成し, 求人推薦とスキル提案タスクに活用している. Lacic ら [15] は, 変分オートエンコーダを利用した求人推薦手法を提案し, ユーザの行動履歴と求人職務内容に基づいた推薦を行っている. Lacic らは精度での評価だけでなく, 推薦のカバレッジについても評価を行っている. 本研究は, これまでの研究であまり活用されてこなかった, 閲覧やお気に入りなどの補助行動データを活用する点が特徴の1つである.

2.2 補助行動データを用いた推薦手法

補助行動データを用いた推薦 (Multi Behavior Recommendation) は, 近年多くの研究者の注目を集めている. これはユーザのアイテムに対するターゲット行動の確率を予測するために, ターゲット行動とは別の種類の行動データを活用する推薦手法である. Gao ら [6] は行動タイプ別に Neural Collaborative Filtering [9] のユニットを用意し, マルチタスク学習 [3] を行っている. Chen ら [4] は, Graph Convolutional Network を用いて, ユーザの複数種類の行動データ間の関係を表す表現学習を行い, かつサンプリングを行わない最適化手法により高い推薦精度を達成している. また, Xia ら [26] はユーザーとアイテム間の複数種類の行動グラフに加え, アイテム間の関係グラフを用いて推薦を行っている. 本研究では補助行動データに加え, ユーザの明示的な選好データを利用している点が既存の補助行動データを用いた推薦手法にはない特徴である.

2.3 グラフ表現に基づいた推薦手法

グラフ表現に基づいた推薦手法 (Graph-based Recommendation) は, Graph Neural Network [21] や TransE [1] などの, グラフの表現学習を行う手法を推薦タスクに応用した手法である. グラフとして利用するデータの観点から見ると, (1) ユーザとアイテムのインタラクションデータのグラフを利用する手法と, (2) インタラクションデータの他に, アイテムの side information を利用する手法の2つに大きく分けられる. (1) について, Wang ら [23] は Graph Neural Network [21] を用いてユーザとアイテムの直接の関係だけでなく, 高次の繋がりを捉えることで推薦精度を向上させている. また, He ら [8] はインタラクションデータをエンティティ間の意味的關係を保持するグラフである知識グラフに変換し, アイテムを知識グラフ上のエンティティ, ユーザを個々のリレーションとみなし学習を行う逐次推薦の手法を提案している. ユーザを表現するリレーションはユーザによって観測されたアイテムの系列から学習さ

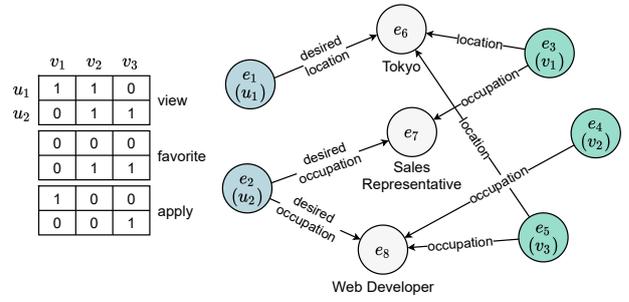


図1 入力データの例. 左図はユーザ u と求人 v の行動データ行列の入力例を表す. 右図はグラフデータの入力例を表す.

れる. (2) について, Zhang ら [28] は, アイテムが持つ情報を知識グラフとして表現し, それをアイテムのテキストデータと画像データとともに利用する推薦手法を提案している. Cao ら [2] は知識グラフの未完全な性質を考慮し, TransH [25] を用いて, 知識グラフの完成タスクとアイテム推薦タスクを同時に学習する手法を提案している. 同時学習の結果, 2つのタスクがシナジーを生み, より良い性能を示すとしている. 本研究と既存のグラフ表現に基づいた推薦手法との差異として, 次の2点を挙げる事ができる. 1つはグラフデータの種別であり, 本研究ではユーザの明示的な選好情報から構築されるグラフデータを用いる. もう1つは, 本研究ではユーザの選好を表すグラフに加え, 補助行動データも活用し, 両者を組み合わせることで更なる推薦精度向上を試みている点である.

3 提案手法

本章では, 補助行動データとユーザの明示的な選好データを活用した提案手法である Job Recommendation Model with Multi Behaviors and Explicit Preferences (JME) について述べる. はじめに問題定義について説明し, 次に提案手法の概要を説明する. 提案手法はいくつかのモジュールから構成されているため, それぞれのモジュールの詳細について説明を行い, 最後に提案手法の学習方法と推論方法について述べる.

3.1 問題定義

ユーザと求人集合をそれぞれ U, V で表し, 個々のユーザと求人をそれぞれ u と v で表す. 本研究で扱う求人推薦タスクは, ユーザ集合 U と求人集合 V に関するデータを入力として受け取り, ユーザ u の求人 v に対する応募の尤度 $\hat{y}_{u,v}$ を推定することである. $\hat{y}_{u,v}$ はユーザ u が将来的に求人 v に対し応募する確率の程度を表す. 推薦時はユーザ u が未観測の求人に対し $\hat{y}_{u,v}$ が計算され, $\hat{y}_{u,v}$ の降順に求人が推薦される. 入力となるユーザ u と求人 v に関するデータは (1) ユーザの求人に対するインタラクションデータと, (2) ユーザの選好データの2種類ある. (1) について, 行動タイプの集合を W とすると, ユーザと求人インタラクションデータはテンソル $\mathbf{X} = [x_{u,v,w}]_{|U| \times |V| \times |W|} \in \{0, 1\}$ で表現できる. $x_{u,v,w}$ は行動タイプ w のもとで, ユーザ u の求人 v に対するインタラクションがあったかどうかを表す. 本研究では次の3つの行動タ

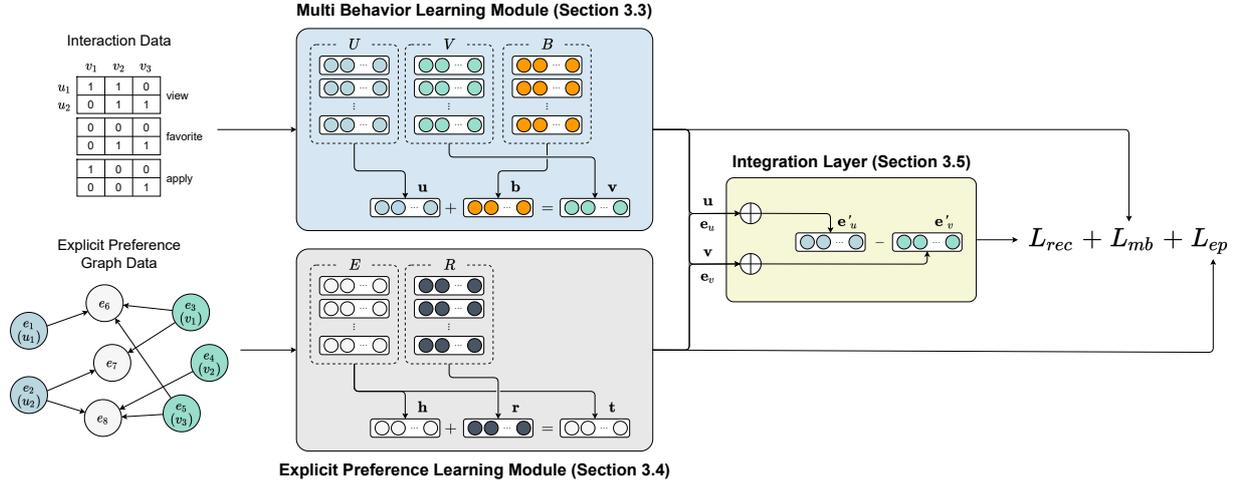


図2 提案手法の概要.

イブを扱う: *view*(閲覧), *favorite*(お気に入り登録), *apply*(応募). (2) について, これはユーザーがオンライン転職サービス上で明示的に設定する, 求人に対する希望条件などのデータである. 本研究ではグラフとして表現し利用する. グラフのノード集合 E にはユーザーと求人のノード以外にカテゴリデータが含まれるため, $|E| > |U| + |V|$ となる. カテゴリデータは, たとえば「営業」という職種や「東京都」という勤務地に関するものなどがあり, ユーザーと求人を媒介する役割を担っている. また, 以下において, ユーザーの選好データから構築されるグラフに関して, リレーション集合を R , トリプル集合を $S = \{(h, r, t) | h, t \in E, r \in R\}$ で表す. 図1は問題の入力例である.

3.2 提案手法の概要

提案手法の概要を図2に示す. 本提案手法は, (1) Multi Behavior Learning, (2) Explicit Preference Learning の2つのモジュールと, (1)と(2)のそれぞれから得られるユーザーと求人の潜在表現を統合する, (3) 統合レイヤから構成される. (1) Multi Behavior Learning モジュール (3.3 節) では, インタラクションデータのテンソルから, ユーザーと求人の表現学習を行う. (2) Explicit Preference Learning モジュール (3.4 節) は, ユーザーの選好データを表すグラフから, ユーザーと求人の表現学習を行う. (3) 統合レイヤ (3.5 節) では2つのモジュールから得られたユーザーと求人の潜在表現から, ユーザーの求人に対する応募尤度の計算を行う. 次節以降, 提案手法を構成する各コンポーネントについて説明する.

3.3 Multi Behavior Learning モジュール

本モジュールではユーザーの求人に対するインタラクションデータのテンソルから, ユーザーと求人それぞれの潜在ベクトルを獲得する. それぞれの潜在表現を得るために, 本モジュールでは入力データをグラフとして変換し, 知識グラフの埋め込み手法の一つである TransE [1] を用いて表現学習を行う. 行動データから潜在表現を得る方法としては, Matrix Factorization をはじめ様々な手法が存在するが, 本研究では補助行動データを

含む複数の行動データとの関係を表現するために TransE を採用する. TransE では, 知識グラフ中のリレーション r で繋がれた2つのエンティティ h, t が, ユークリッド空間上で $h + r \approx t$ となるように学習を行う. 本モジュールでは, ユーザーと求人を知識グラフのエンティティとして利用し, インタラクションをリレーションとして利用する. インタラクションをリレーションとして利用する際に, インタラクションデータには複数の行動タイプのデータが存在するため, ナイーブな方法として行動タイプ毎にリレーションとして定義する方法が考えられる. しかし, 本研究では行動タイプの組み合わせ毎にリレーションを割り当てることで, ユーザーの行動を, よりきめ細やかに表現したベクトルが獲得できるのではないかと考えた. 例えば, (1) ユーザー u が求人 v に対し閲覧と応募を行っている場合と, (2) ユーザー u が求人 v に対し閲覧, お気に入り, 応募を行っている場合では, (2) の方が, u の選好に v はより適合していると考えられる. この場合, $|W|$ 種類の行動タイプが存在する場合, リレーションは $2^{|W|} - 1$ 種類となる.

行動タイプの組み合わせ集合を B とすると (ただし, $|B| = 2^{|W|} - 1$), インタラクションデータのテンソル \mathbf{X} は $X' = \{(u, v, b) | u \in U, v \in V, b \in B\}$ に変換できる. このとき, 本モジュールにおける損失関数は式 (1) のように定義できる.

$$L_{mb} = \sum_{\substack{(u,v,b) \\ \in \mathbf{X}'}} \sum_{\substack{(u,v',b) \\ \notin \mathbf{X}'}} [m_{mb} + d(\mathbf{u} + \mathbf{b}, \mathbf{v}) - d(\mathbf{u} + \mathbf{b}, \mathbf{v}')]_+, \quad (1)$$

$\mathbf{u}, \mathbf{v}, \mathbf{b}$ はそれぞれ u, v, b の潜在ベクトルである. v' はユーザーが未観測の求人を表す. $m_{mb} > 0$ はマージンを表すスカラのパラメータである. $d(\mathbf{x}, \mathbf{y})$ は x と y の距離関数を表し, 本研究ではユークリッド距離を用いる. また, $[x]_+ = \max(0, x)$ である.

3.4 Explicit Preference Learning モジュール

本モジュールはユーザーの選好データをグラフで表現し, ユーザーと求人それぞれの潜在ベクトルを獲得することを目的とする. 潜在ベクトルを獲得する方法としては, 任意のグラフ表現学習

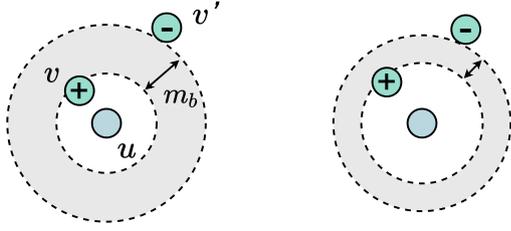


図3 Behavior Aware マージン. 強い選好の適合求人とネガティブサンプリングされた求人との距離は遠くなる(左図). 弱い選好の適合求人とネガティブサンプリングされた求人との距離は近くなる(右図).

の手法が利用可能だが, 本研究では Multi Behavior Learning モジュールと同様 TransE を採用した. 学習時の損失関数 L_{ep} は式 (2) で表される.

$$L_{ep} = \sum_{(h,r,t) \in S} \sum_{(h',r,t') \in S'(h,r,t)} [m_{ep} + d(\mathbf{h} + \mathbf{r}, \mathbf{t}) - d(\mathbf{h}' + \mathbf{r}, \mathbf{t}')]_+, \quad (2)$$

$\mathbf{h}, \mathbf{r}, \mathbf{t}$ はそれぞれ h, r, t の潜在ベクトルである. $S'(h,r,t)$ は学習データ中のトリプル (h, r, t) に対しネガティブサンプリングを行ったデータで, 式 (3) で表される. $m_{ep} > 0$ はマージンを表すスカラのパラメータである.

$$S'(h,r,t) = \{(h', r, t) | h' \in E\} \cup \{(h, r, t') | t' \in E\}. \quad (3)$$

3.5 統合レイヤ

本レイヤでは 3.3 節と 3.4 節で述べた 2 つのモジュールから得られるユーザと求人の潜在ベクトルを統合し, 推薦タスクで用いる最終的な潜在ベクトルを獲得する. 最終的なユーザ, 求人の潜在ベクトルをそれぞれ $\mathbf{e}'_u, \mathbf{e}'_v$ とする. Multi Behavior Learning モジュールで得られたユーザ, 求人の潜在ベクトルをそれぞれ \mathbf{u}, \mathbf{v} とし, Explicit Preference Learning モジュールで得られたユーザ, 求人の潜在ベクトルをそれぞれ $\mathbf{e}_u, \mathbf{e}_v$ とすると, $\mathbf{e}'_u, \mathbf{e}'_v$ は, 式 (4) で表される.

$$\mathbf{e}'_u = \mathbf{u} + \mathbf{e}_u, \mathbf{e}'_v = \mathbf{v} + \mathbf{e}_v. \quad (4)$$

統合レイヤにおける損失関数は式 (5) で表される.

$$L_{rec} = \sum_{(u,v,b) \in \mathbf{X}'} \sum_{(u,v',b) \notin \mathbf{X}'} [m_b + d(\mathbf{e}'_u, \mathbf{e}'_v) - d(\mathbf{e}'_u, \mathbf{e}'_{v'})]_+, \quad (5)$$

ここで, m_b は行動タイプの組み合わせ b における Behavior Aware マージンで, 式 (6) で表される.

$$m_b = 1.5 - \sigma(\|\mathbf{b}\|_2), \quad (6)$$

$\sigma(x)$ はシグモイド関数である. 式 (5) より, ユーザの選好に近い求人は, ユークリッド空間上でユーザとの距離が小さくなるように学習される. よって, Multi Behavior Learning モジュールにおける行動タイプの組み合わせベクトルも, より強い選好を表すベクトルはそのノルムが小さくなると考えられる. そこで, そのノルムをマージンの計算に利用し, 選好の強さに

表1 インタラクションデータの統計情報.

| ユーザ | 求人 | インタラクション | | |
|-------|-------|----------|-------|-------|
| | | 閲覧 | お気に入り | 応募 |
| 1,996 | 3,444 | 24,542 | 5,918 | 4,112 |

表2 明示的な選好情報から作成したグラフの統計情報.

| エンティティ | | | リレーション | トリプル |
|--------|-------|-----|--------|--------|
| ユーザ | 求人 | その他 | | |
| 1,996 | 3,444 | 546 | 10 | 45,529 |

応じたマージンの大きさにすることで, 強い選好の求人はユーザとの距離がより近くなるような学習を実現している. 図3 は Behavior Aware マージンを図示したものである.

3.6 学習・推論方法

提案手法の損失は式 (7) で表される. 学習時はマルチタスク学習 [3] を用いて, 式 (7) の損失を最小化するように各モジュールのパラメータを確率的勾配降下法などで更新していく.

$$L = L_{rec} + L_{mb} + L_{ep}. \quad (7)$$

推論時は式 (8) の計算により, ユーザ u の求人 v に対する応募尤度 $\hat{y}_{u,v}$ を計算する.

$$\hat{y}_{u,v} = -d(\mathbf{e}'_u, \mathbf{e}'_v). \quad (8)$$

ユーザの選好に近い求人はユーザとの距離が小さくなる一方で, 推薦では応募尤度の高い求人から推薦していくため $d(\mathbf{e}'_u, \mathbf{e}'_v)$ に対し -1 を乗算している.

4 評価実験

本章では, はじめに評価実験で用いるデータセットについて説明し, 次に評価方法について説明する. その後, 実験結果とその考察を示す. 提案手法は複数のモジュールから成り, また複数種類のデータを用いているため, それぞれについての ablation test の結果を最後に示す.

4.1 データセット

本研究では国内で実際に運営されている転職サービス^(注1)で得られたデータを用いて評価実験を行った. はじめに, 2020年1月1日から2021年7月31日の期間にインタラクションのあったユーザと求人のデータを収集し, 次に (1) 1度も応募のないユーザと求人を除き, (2) インタラクションの総数が5件未満のユーザと求人を除く前処理を行った. 表1に, インタラクションデータの統計情報を示す.

また, インタラクションデータに含まれるユーザ, 求人について, 転職サービスに登録されているカテゴリデータを基にグラフを作成した. 表2に, グラフの統計情報を示す. その他のエンティティがカテゴリデータであり, 例えば, Webエンジニアや営業といった職種, 東京都や大阪府といった勤務地データを表す. リレーションはユーザの希望条件や求人の属性を表す.

(注1): <https://glit.io>

表 3 ベースラインとの比較実験結果. “**”は、提案手法と KGIN(下線付き)の差が、 $p < 0.01$ の水準で統計的に有意であることを示す.

| | | MRR | nDCG@5 | nDCG@10 | nDCG@20 |
|--------------------------------------|---------------|-----------------|-----------------|-----------------|-----------------|
| General Recommendation | MF-BPR | 0.2511 | 0.2475 | 0.2836 | 0.3090 |
| Multi Behavior Recommendation | KHGT | 0.1415 | 0.1229 | 0.1593 | 0.2014 |
| | MB-GMN | 0.2577 | 0.2502 | 0.2901 | 0.3333 |
| | GHCF | 0.3112 | 0.3305 | 0.3573 | 0.3706 |
| Graph-based Recommendation | CKE | 0.2981 | 0.3041 | 0.3327 | 0.3571 |
| | KGIN | <u>0.4415</u> | <u>0.4607</u> | <u>0.5031</u> | <u>0.5344</u> |
| Proposed Model | JME | 0.5133** | 0.5385** | 0.5793** | 0.6025** |

例えば、ユーザ u が職種エンジニアを希望している、求人 v の勤務地は東京都、などを表す。本研究で用いるデータセットのサイズは比較的小さく、データの提供元である転職サービスも大きな市場シェアを持つものではないが、IT 業界に限らない、非常に幅広い種類の求人とユーザを扱っている。よって、本研究のデータセットは提案手法の汎化性能を検証するのに適したデータセットと考えられる。

4.2 評価

4.2.1 評価方法

時系列を考慮してデータセットを学習用、バリデーション用、テスト用の 3 つに分割した。具体的には、ターゲット行動である応募データについては次のように分割した。

- 3 つ以上の応募データがあるユーザ (344 ユーザ): 最後の応募をテストデータとし、その 1 つ前の応募をバリデーションデータに使用。残りの応募を学習データに使用。
- 2 つの応募データがあるユーザ (371 ユーザ): 最後の応募をテストデータかバリデーションデータにランダムに使用。残りの応募を学習データに使用。
- 1 つの応募データがあるユーザ (1,281 ユーザ): 応募を学習データに使用。

補助行動データはすべて学習データとして利用した。[9] の評価方法に倣い、ユーザ毎にテストデータ中の適合求人 1 件と、ランダムに選択したユーザが未観測の求人 99 件を合わせた 100 件の求人を応募尤度の降順に並べ、mean reciprocal rank (MRR) [22] と normalized discounted cumulative gain (nDCG) [11] を用いてベースライン手法との比較を行った。

4.2.2 ベースライン

2 章で述べたように、本研究は推薦手法における複数の分野と関連があるため、それぞれの分野からベースラインを採用する。

General Recommendation: 補助行動データとユーザの嗜好を表すグラフの両方を使用しないベースライン。

- **MF-BPR** [20]: Bayesian Personalized Ranking を用いて、ランキング指標を直接最適化した Matrix Factorization.
- Multi Behavior Recommendation:** 補助行動データを利用するベースライン。
 - **KHGT** [26]: 階層的グラフ Transformer ネットワークを用いて、行動タイプ別の重要度の違いを考慮した推薦手法。
 - **MB-GMN** [27]: グラフメタネットワークを用いて、行

動タイプ間の相関を考慮した推薦手法。

- **GHCF** [4]: ユーザ-アイテムネットワークにおける離れたペアの相互作用を探索するために、グラフ畳み込みネットワークを活用した推薦手法。

Graph-based Recommendation: インタラクションデータに加え、アイテム間の知識グラフを利用するベースライン。本実験ではユーザの選好情報から作成したグラフから、ユーザのエンティティを除いたグラフをアイテムの知識グラフとして利用する。

- **CKE** [28]: 協調フィルタリングに知識グラフ、テキストデータ、画像データの埋め込み表現を組み込んだ手法。なお、本研究における実験では公平性のために知識グラフの埋め込み表現のみを使用した。
- **KGIN** [24]: インタラクションの意図を、知識グラフのリレーションのアテンションとしてモデル化し、意図を探索する推薦手法。

なお、補助行動データを考慮しないベースライン手法では、学習データの多寡で不公平が生じないように、補助行動データをターゲット行動データとして扱い学習を行った [16]。

4.2.3 パラメータ設定

提案手法では、Adam optimizer [13] を用いて最適化を行い、バッチサイズは 64 で固定した。スカラ値のマージンパラメータである式 (1) の m_{mb} と式 (2) の m_{ep} はそれぞれ 1.0 とした。また、潜在ベクトルの次元数と学習率をそれぞれ $\{8, 16, 32, 64\}$ と $\{5e^{-2}, 1e^{-2}, 5e^{-3}, 1e^{-3}, 5e^{-4}, 1e^{-4}\}$ の中からグリッドサーチを用いて探索し、最も推薦精度の高い組み合わせを採用した。提案手法の潜在ベクトルの初期化は Xavier initialization [7] を用いた。

ベースラインでは、それぞれの論文で提示されている値でパラメータを初期化し、潜在ベクトルの次元数と学習率を提案手法と同じ候補の中からグリッドサーチで探索した。

4.3 実験結果

実験は、バリデーション用とテスト用のデータセットの分割方法をランダムに変更したデータセットを 5 つ用意して行った。表 3 に、5 回の実験結果の平均値を示す。

まず、表 3 から、提案手法の推薦精度が、他の最新のベースライン手法に比べて、いずれの評価指標においても上回っていることが分かる。この推薦精度の差が統計的に有意であることを確認するため、提案手法と最良のベースライン手法 (KGIN)

表 4 Ablation test の結果.

| | | MRR | nDCG@5 | nDCG@10 | nDCG@20 |
|----------------|--------------------|---------------|---------------|---------------|---------------|
| Data Ablation | w/o View | 0.2699 | 0.2648 | 0.3077 | 0.3493 |
| | w/o Favorite | 0.5107 | 0.5369 | 0.5768 | 0.6012 |
| | w/o View, Favorite | 0.0803 | 0.0650 | 0.0772 | 0.0951 |
| Model Ablation | w/o EPL Module | 0.3364 | 0.3414 | 0.3732 | 0.3992 |
| Full Model | JME | 0.5133 | 0.5385 | 0.5793 | 0.6025 |

で得られた結果について、両側 t 検定を行った。その結果、いずれの指標においても $p < 0.01$ の水準で統計的に有意な差があることが示された。これは補助行動データとユーザの明示的選好を表すグラフの両方を同時に活用したことによる効果であると考えられる。

次に、ターゲット行動データのみを利用する一般的な推薦手法である MF-BPR と比べ、その他の最新のベースライン手法が高い推薦精度を示していることが分かる。これは、補助行動データ、または side information を利用することにより推薦精度が向上するという先行研究の内容を確認できる結果となった。KHGT のみ推薦精度の低下が見受けられるものの、これは使用したデータセットの規模が、KHGT が実験を行ったデータセットより小さいため、うまく学習できなかった可能性があることによるものと考えている。

最後に、Multi Behavior Recommendation の手法と Graph-based Recommendation の手法を比較する。今回の実験では、Graph-based Recommendation の手法である KGIN が、提案手法を除くと最も良い結果となった。Multi Behavior Recommendation の手法の中で最良の GHCF と比べ、KGIN はより高い推薦精度を示しているが、これは本実験のタスクにおいて、ユーザの選好データの有効性が高かったためであると考えられる。よって、データセットやタスクの設定によって、両方の手法の優劣は変わるものと考えられる。

4.4 Ablation Test

補助行動データと Explicit Preference Learning モジュールの推薦精度に及ぼす効果を確認するため、データまたはモジュールを除いた場合の ablation test を次の 4 つのパターンについて行った。

Data Ablation

- **w/o View:** 行動データから閲覧データを除いたパターン。
- **w/o Favorite:** 行動データからお気に入りデータを除いたパターン。
- **w/o View, Favorite:** 行動データから閲覧とお気に入りデータを除いたパターン。つまり、応募データのみを用いたパターン。

Model Ablation

- **w/o EPL Module:** Explicit Preference Learning モジュールを除いたパターン。統合レイヤで用いるユーザと求人 の潜在表現として、Multi Behavior Learning モジュールで得られたものをそのまま用いる。

表 4 に、実験結果を示す。まず、最も推薦精度が良かったのは提案手法の Full Model パターンであった。続いて、お気に入りデータを用いない w/o Favorite のパターンの推薦精度が高く、Full Model のパターンとの差はわずかであった。僅差になった理由として、お気に入りデータは推薦精度向上に役立つ半面、ノイズとして働く可能性もあるのではないかと考えている。

次に、他の Data Ablation の結果を見ると、大きく推薦精度を下げていることが分かる。これは閲覧データを利用しないことにより、学習データのサイズが大幅に減少してしまったことが原因と考えられる。したがって、ターゲット行動データのみではなく補助行動データを用いることは推薦精度向上に寄与すると考えられる。

最後に、Model Ablation の結果について、こちらも大きな推薦精度の低下が見受けられる。よって、ユーザの選好データを反映したグラフの表現学習を行うことは、推薦精度の向上に役立つものと考えられる。

5 議 論

本章では、JME をより深く理解するために JME の実験結果をさらに分析する。まず、JME が学習した潜在表現について考察する。次に、具体的なテストケースを採り上げ、JME が適合データを高いランクで推薦できる場合に、テストデータにどのような特徴があるかを考察する。

5.1 表現学習

はじめに、提案手法の学習で得られた行動タイプの組み合わせの潜在表現について考察する。式 (5) に示すように、ユーザの選好に近い求人は、ユーザとのユークリッド距離が小さくなるように学習される。そのため、Multi Behavior Learning モジュールにおける強い選好を表す行動タイプの組み合わせは、そのベクトルのノルムが他の組み合わせよりも小さくなると考えられる。図 4 に、行動タイプの組み合わせのノルムを示す。図中の v, f, a は行動タイプを表しており、それぞれ *view* (閲覧), *favorite* (お気に入り), *apply* (応募) を表す。2 文字以上からなる文字列は複数の行動タイプを含む組み合わせであり、例えば、 vf は閲覧とお気に入りの組み合わせを表す。 a と fa はデータセットの中には存在しない組み合わせであったため、初期値が保たれている。最も強いユーザの選好を表すと考えられる vfa のノルムが最も小さく学習されており、予想通りの結果となっている。一方で、最もノルムの大きかった組み合わせは f であった。これは、お気に入り単体だけでは、ユーザがその

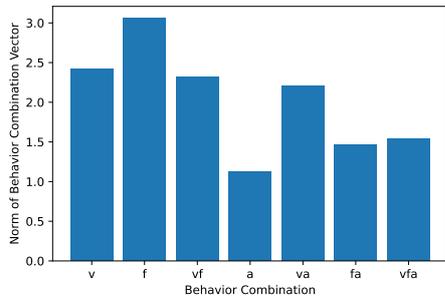


図4 行動タイプの組み合わせベクトルのノルム.

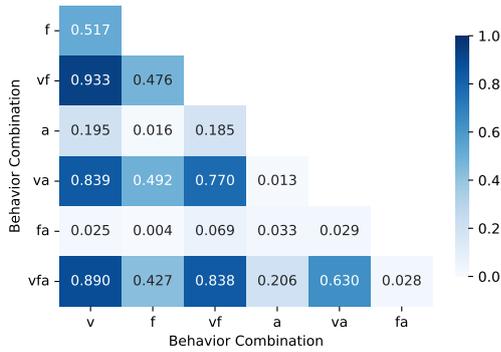


図5 行動タイプの組み合わせベクトル間のコサイン類似度.

求人に興味があると結論づけるのは早急であるという考えを示唆しており、4.4節の考察を裏付ける結果となっている。

次に、図5に、行動タイプの組み合わせベクトル間のコサイン類似度を示す。色の濃いブロックが類似度が高いことを示している。この図を見ると、同じ行動タイプを含む組み合わせ、例えばvのベクトルとvfのベクトルなど、の類似度が高いことが分かる。同じ行動タイプを含むベクトルが類似していることは自然であると考えられることから、提案手法は行動タイプの組み合わせの潜在表現を、うまく学習できていることが分かる。

5.2 詳細分析

JMEをさらに理解するため、本節では具体的なテスト例を分析する。JMEの推薦の成功例としてユーザ1845、失敗例としてユーザ323を採り上げる。

ユーザ1845のテストでは、JMEは適合求人(3358)を1位で推薦することができた。一方、グラフ表現に基づいた推薦手法の中で最も高い推薦精度を示したKGINは適合求人を63位、補助行動データを用いた推薦手法の中で最も高い推薦精度を示したGHCFは60位での推薦となった。図6はユーザ1845とインタラクションのあった求人(3347, 3367)、適合求人(3358)の関係を示す。ユーザ1845は希望の業種(Hotel and Travel Industry)を明示しており、適合求人はその業種に属している。JMEはインタラクションのあった求人と適合求人との間には表れない、ユーザの明示的選好を考慮できたため、適合求人を1位で推薦できたと考えられる。

次に、ユーザ323のテストについて、JMEは適合求人(1178)を44位で、KGINは53位、GHCFは56位で推薦する結果となった。図7はユーザ323がインタラクションした求人(49,

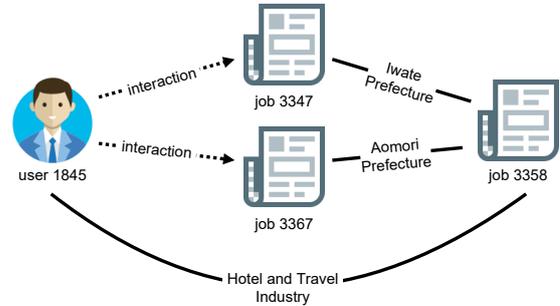


図6 JMEが推薦に成功したテスト例.

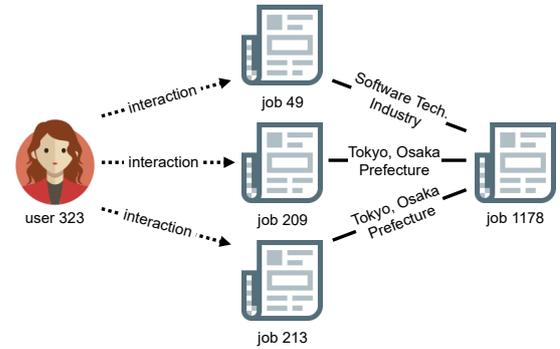


図7 JMEが推薦に失敗したテスト例.

209, 213)と適合求人(1178)がそれらの求人と属性情報を共有していることを示す。しかし、成功例とは異なり、ユーザ323は求人1178と属性情報を共有していない。この共通属性の欠落と、インタラクションのあった求人と適合求人にも共通する属性(Software Tech. Industry, Tokyo Prefecture, Osaka Prefecture)に紐づく求人数の多さから、ユーザ323に適合求人を上位で推薦することが難しかったと考えられる。

上記2つの例より、JMEはユーザが適合求人と職種や勤務地など、何らかの属性を共有している場合に、特に効果的に推薦可能と考えられる。

6 まとめ

本研究では、求人推薦における補助行動データと、ユーザの明示的な選好データを表現するグラフの両方を活用した推薦手法を提案した。提案手法では知識グラフの埋め込み手法を用いて、それぞれのデータからユーザと求人の表現学習を行う。さらに、両方のデータから得られたユーザと求人の潜在表現から、距離学習の学習方法に基づき、ユーザの求人に対する選好を学習する。実際の転職サービスから得られたデータを使った実験の結果、提案手法は最新手法よりも高い推薦精度を達成することができた。1つの利点として提案手法はテキストデータを利用しないため、あらゆる言語の転職サービスに適用可能である。

今後の研究課題として、推薦の説明性を向上させることを計画している。本研究では明示的に活用できていないが、ユーザの選好データの表現学習時に得られる、ユーザと求人以外のカテゴリデータのエンティティや、リレーションの潜在表現を用いることで、ユーザがより納得できる推薦を生成できるのではないかと考えている。

文 献

- [1] Antoine Bordes, Nicolas Usunier, Alberto García-Durán, Jason Weston, and Oksana Yakhnenko. Translating Embeddings for Modeling Multi-relational Data. In *Proceedings of the 27th Annual Conference on Neural Information Processing Systems (NIPS '13)*, pp. 2787–2795, 2013.
- [2] Yixin Cao, Xiang Wang, Xiangnan He, Zikun Hu, and Tat-Seng Chua. Unifying Knowledge Graph Learning and Recommendation: Towards a Better Understanding of User Preferences. In *Proceedings of the 28th World Wide Web Conference (WWW '19)*, pp. 151–161, 2019.
- [3] Rich Caruana. Multitask learning. *Machine Learning*, Vol. 28, No. 1, pp. 41–75, 1997.
- [4] Chong Chen, Weizhi Ma, Min Zhang, Zhaowei Wang, Xiquiang He, Chenyang Wang, Yiqun Liu, and Shaoping Ma. Graph Heterogeneous Multi-Relational Recommendation. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-21)*, pp. 3958–3966, 2021.
- [5] Vachik S Dave, Baichuan Zhang, Mohammad Al Hasan, Khalifeh AlJadda, and Mohammed Korayem. A Combined Representation Learning Approach for Better Job and Skill Recommendation. In *Proceedings of the 27th ACM International Conference on Information and Knowledge Management (CIKM '18)*, pp. 1997–2005, 2018.
- [6] Chen Gao, Xiangnan He, Danhua Gan, Xiangning Chen, Fuli Feng, Yong Li, Tat-Seng Chua, Lina Yao, Yang Song, and Depeng Jin. Learning to Recommend with Multiple Cascading Behaviors. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Vol. 33, No. 6, pp. 2588–2601, 2021.
- [7] Xavier Glorot and Yoshua Bengio. Understanding the difficulty of training deep feedforward neural networks. In *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics (AISTATS '10)*, Proceedings of Machine Learning Research, pp. 249–256, 2010.
- [8] Ruining He, Wang-Cheng Kang, and Julian McAuley. Translation-based Recommendation. In *Proceedings of the 11th ACM Conference on Recommender Systems (RecSys '17)*, pp. 161–169, 2017.
- [9] Xiangnan He, Lizi Liao, Hanwang Zhang, Liqiang Nie, Xia Hu, and Tat-Seng Chua. Neural Collaborative Filtering. In *Proceedings of the 26th International Conference on World Wide Web (WWW '17)*, pp. 173–182, 2017.
- [10] Chao Huang. Recent advances in heterogeneous relation learning for recommendation. In *Proceedings of the 30th International Joint Conference on Artificial Intelligence (IJCAI-21)*, pp. 4442–4449, 2021.
- [11] Kalervo Järvelin and Jaana Kekäläinen. Cumulated gain-based evaluation of IR techniques. *ACM Transactions on Information Systems (TOIS)*, Vol. 20, No. 4, pp. 422–446, 2002.
- [12] Miao Jiang, Yi Fang, Huangming Xie, Jike Chong, and Meng Meng. User click prediction for personalized job recommendation. *World Wide Web*, Vol. 22, No. 1, pp. 325–345, 2019.
- [13] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. Proceedings of the 3rd International Conference on Learning Representations (ICLR '15), 2015.
- [14] Nagaraj Kota, Venkatesh Duppada, Ashvini Jindal, and Mohit Wadhwa. Understanding job seeker funnel for search and discovery personalization. In *Proceedings of the 30th ACM International Conference on Information & Knowledge Management (CIKM '21)*, pp. 3888–3897, 2021.
- [15] Emanuel Lacic, Markus Reiter-Haas, Dominik Kowald, Manoj Reddy Daredy, Junghoo Cho, and Elisabeth Lex. Using Autoencoders for Session-based Job Recommendations. *User Modeling and User-Adapted Interaction (UMUAI)*, Vol. 30, No. 4, pp. 617–658, 2020.
- [16] Qiang Liu, Shu Wu, and Liang Wang. Multi-Behavioral sequential prediction with recurrent Log-Bilinear model. *IEEE Transactions on Knowledge and Data Engineering (TKDE)*, Vol. 29, No. 6, pp. 1254–1267, 2017.
- [17] J Malinowski, T Keim, O Wendt, and T Weitzel. Matching People and Jobs: A Bilateral Recommendation Approach. In *Proceedings of the 39th Annual Hawaii International Conference on System Sciences (HICSS '06)*, pp. 137c–137c, 2006.
- [18] Qingxin Meng, Hengshu Zhu, Keli Xiao, Le Zhang, and Hui Xiong. A hierarchical Career-Path-Aware neural network for job mobility prediction. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining (KDD '19)*, pp. 14–24, 2019.
- [19] Ioannis K. Paparrizos, Berkant Barla Cambazoglu, and Aristides Gionis. Machine Learned Job Recommendation. In *Proceedings of the 5th ACM Conference on Recommender Systems (RecSys '11)*, pp. 325–328, 2011.
- [20] Steffen Rendle, Christoph Freudenthaler, Zeno Gantner, and Lars Schmidt-Thieme. BPR: Bayesian Personalized Ranking from Implicit Feedback. In *Proceedings of the 25th Conference on Uncertainty in Artificial Intelligence, (UAI '09)*, pp. 452–461, 2009.
- [21] Franco Scarselli, Marco Gori, Ah Chung Tsoi, Markus Hagenbuchner, and Gabriele Monfardini. The graph neural network model. *IEEE Transactions on Neural Networks*, Vol. 20, No. 1, pp. 61–80, 2009.
- [22] Ellen M Voorhees and Others. The TREC-8 question answering track report. In *Trec*, Vol. 99, pp. 77–82, 1999.
- [23] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural graph collaborative filtering. In *Proceedings of the 42th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '19)*, pp. 165–174, 2019.
- [24] Xiang Wang, Tinglin Huang, Dingxian Wang, Yancheng Yuan, Zhenguang Liu, Xiangnan He, and Tat-Seng Chua. Learning intents behind interactions with knowledge graph for recommendation. In *Proceedings of the Web Conference 2021 (WWW '21)*, pp. 878–887, 2021.
- [25] Zhen Wang, Jianwen Zhang, Jianlin Feng, and Zheng Chen. Knowledge graph embedding by translating on hyperplanes. In *Proceedings of the 28th AAAI Conference on Artificial Intelligence (AAAI-14)*, Vol. 28, pp. 1112–1119, 2014.
- [26] Lianghao Xia, Chao Huang, Yong Xu, Peng Dai, Xiyue Zhang, Hongsheng Yang, Jian Pei, and Liefeng Bo. Knowledge-Enhanced Hierarchical Graph Transformer Network for Multi-Behavior Recommendation. In *Proceedings of the 35th AAAI Conference on Artificial Intelligence (AAAI-21)*, pp. 4486–4493, 2021.
- [27] Lianghao Xia, Yong Xu, Chao Huang, Peng Dai, and Liefeng Bo. Graph meta network for Multi-Behavior recommendation. In *Proceedings of the 44th International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR '21)*, pp. 757–766, 2021.
- [28] Fuzheng Zhang, Nicholas Jing Yuan, Defu Lian, Xing Xie, and Wei-Ying Ma. Collaborative Knowledge Base Embedding for Recommender Systems. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*, pp. 353–362, 2016.