

# Human-in-the-loop 潜在空間学習による 空間型インタラクティブ文献管理システム

渡邊 真悟<sup>†</sup> 伊藤 寛祥<sup>††</sup> 松原 正樹<sup>††</sup> 森嶋 厚行<sup>††</sup>

<sup>†</sup> 筑波大学 知識情報・図書館学類 〒 305-8550 茨城県つくば市春日 1-2

<sup>††</sup> 筑波大学 図書館情報メディア系 〒 305-8550 茨城県つくば市春日 1-2

E-mail: †shingo.watanabe.2021b@mlab.info, ††{ito,masaki,mori}@slis.tsukuba.ac.jp

**あらまし** 昨今、研究者が管理しなければならない研究論文やドキュメントの数は膨大なものになっている。管理する文書が増え続ける中で、研究者が一貫した基準に則って文書を管理し続けるということは、各個人の基準を明示的に表現する方法が明らかになっていないため困難である。本論文では、潜在空間中に配置された各文書の位置に対して、ユーザがフィードバックを与えることによる Human-in-the-Loop 潜在空間学習を適用して、ユーザの文書管理の基準を得る文書管理システムを提案する。このようなシステムを設計する上での重要な課題は、文書間の関係を表現するグラフ表現をどのように学習するかである。提案するフレームワークでは、meta-path を文書管理の基本的な基準とし、meta-path によって構築される文書グラフの表現を学習する。実験では、ユーザからのフィードバックを考慮に入れ、ユーザの文献管理の基準を推定することで、ユーザが求める位置に正しく文献を配置できるかを検証した。実験結果により、提案するフレームワークは既存手法などと比較して、ユーザが求める位置に正しく文献を配置できていることを示した。

**キーワード** Human-in-the-loop, Graph Auto-Encoders, グラフ埋め込み, meta-path, 異種混合ネットワーク

## 1 はじめに

現在、研究者が管理しなければならない研究論文やドキュメントの数は膨大なものになっている。Michael らの研究 [1] では、2014 年以降は減少傾向にあるものの、近年では、毎年 700 万件以上の新しい学術論文が発表されており、そのうち、5 件以上文献を引用している論文は毎年 180 万件を超えていることが示されている。研究者が電子文献を管理する場合、主にオンラインの文献管理ソフトを使用することが考えられ、どの程度オンラインの文献管理ソフトが使用されているか [2] や、文書管理システムの機能性 [3] に関する研究がこれまでに行われてきた。[2] では、トリノ大学の学生や研究者を対象に、どの程度オンラインの文献管理ソフトを使用しているかが調査された。調査の結果、文献管理ソフト上に 200 以上の文書を保存している研究者の割合が半分を超えていることが示されている。

文献を研究者の基準に従って適切に分類し、管理することは研究者が関連研究の発見や、その分野の研究動向などを把握するなどの点で重要であるが、管理する文書が増え続ける中で、研究者が、単なる文献検索を超えて、一貫した基準に則って文書を管理し続けることは困難である。したがって、このような文献管理を支援するシステムがあれば望ましいと考えられるが、そこで大きな問題となるのが、各個人の文書管理の基準をどのように取得し、表現するかが明らかになっていないことである。

以上の問題を解決するために、本論文ではユーザと対話を繰り返すことでユーザの文書の管理基準を取得するインタラクティブな文献管理システムの提案を行う。より具体的には、潜

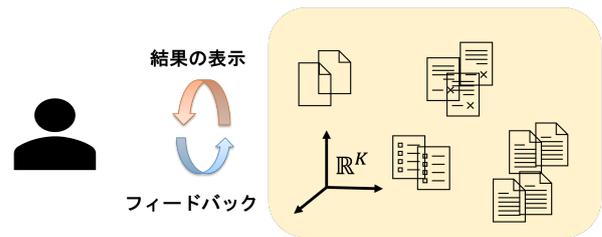


図 1: 空間型インタラクティブ文献管理システムのイメージ: ユーザは、潜在空間中に配置された文書の位置に対しフィードバックを与える。システムは、与えられたフィードバックを基に、ユーザの文書管理の基準を推定し、より良い潜在空間を表示する。

在空間中に配置された各文書の位置に対して、ユーザがフィードバックを与えることによる Human-in-the-loop 潜在空間学習を適用することで、ユーザの文書の管理基準を推定する。本研究では、管理の対象とする文献を異種混合ネットワークとしてモデル化し、ユーザの文献の管理基準を meta-path として表現する。ここで、異種混合ネットワークとは複数の種類のノードやエッジをもつネットワーク構造であり、たとえば文献情報の異種混合ネットワークであれば、文書が持つ属性として、著者や出版年、文書中に出現する単語などが挙げられる。また、meta-path とは異種混合ネットワーク中におけるスキーマレベルのパスのことである。文献情報の異種混合ネットワークであれば、共著関係の文書であることや、同じ年に発表された文書などといったような意味情報を含んだ関係を表現できる。本研究では meta-path が表現する、文献間の意味情報を利用する

ことで、各個人の文献管理基準を表現する。

提案手法の有効性を検討するために、未知の文書が与えられたときにその文書を潜在空間中でユーザが求める位置に正しく配置できるかどうかを実験した。被験者実験では、ある文書について、その文書を潜在空間中において配置されて欲しい位置に動かしてもらった。被験者には文献が持つ情報として毎回、文書のタイトル、著者、発表された会議、発表された年を与えた。実験の評価では、潜在空間中で被験者が求める文書の配置と、提案手法による文書の配置を比較し、被験者が潜在空間中で近くに配置した文書を正しく推定できているかどうかを検証した。評価の結果、提案手法は設定したベースライン手法と比較して、未知の文書を被験者が求める位置に正しく配置できていることを示した。

本論文の貢献は次の通りである。(1) 研究者の文書管理の基準として meta-path が有効である可能性を示した。(2) 研究者の評価基準を自動的に推定する Human-in-the-loop 潜在空間学習のフレームワークを提案した。(3) 複数回にわたる被験者実験によって提案手法の有効性を示した。

本論文の構成は以下である。2章で関連研究と本論文の位置付けについて説明を行い、3章では本論文に必要な予備知識に関する説明を行う。4章では本論文で扱う問題設定についての定義を記述する。5章では本論文で提案する Human-in-the-loop 潜在空間学習の方法について説明する。6章において被験者実験とその結果を示し、7章で実験の結果に対する考察を行う。最後に8章にて本論文をまとめる。

## 2 関連研究

本節では、本研究の関連研究として、文献管理の方法に関する研究、推薦システムに関する研究、アクティブラーニングに関する研究、そしてネットワークの表現学習に関する研究についてそれぞれ述べる。

### 2.1 文献管理の方法に関する研究

関連研究の整理や研究動向の把握のための研究者を支援するようなツールは広く利用されており、このようなツールを対象に行われている研究がある。Francesse [2] は、学生や研究者がどのように参考文献を管理しているかについて、トリノ大学を対象に調査を行なった。調査の結果、研究者がオンライン上で電子文献を管理する際、最も利用されている文献管理ソフトは、EndNote [4] であり、回答者の 49% が利用している、次点で BibTex が 11%、Mendley [5] が 9% で続くという結果になった。EndNote や Mendley などの文献管理ソフトは、プロジェクトや研究分野によって分けられたフォルダ内に文書を保存したり、それぞれの文書にタグをつけることによって、文書を管理するという方法をとっている。一方で、本論文で提案する文献管理システムは、ユーザの文書の管理基準を推定し、文書を空間上にマッピングするという点で既存の文献管理システムとは異なる。本論文で提案するように、文書を空間上にマッピングする利点としては、既存の文献管理システムがとっているよ

うなフォルダ分けによる管理や文書にタグを付与する必要がなくなり、自分が管理している文書集合の全体像を把握することが可能になる点が挙げられる。

### 2.2 推薦システムに関する研究

eコマースの普及に伴って、多くのeコマースプラットフォームでは、内容ベースの推薦や協調フィルタリングによる推薦が広く利用されている。ユーザやアイテム、クエリなどが属性として挙げられるような異種混合ネットワークを用いて、ユーザに新しいアイテムやクエリを推薦するような研究は、近年さまざまな行われている。たとえば、Yu ら [6] が提案するフレームワークでは、まず、モデルがユーザの嗜好に合わせていくつかのクエリを生成する。そしてモデルは、ユーザがクエリをクリックして閲覧するなどの行動をフィードバックとして受け取ることにより良いアイテムの推薦を可能にした。実際のeコマースプラットフォーム上で行われた A/B テストでは、提案したフレームワークを利用した場合に閲覧数の向上が確認された。既存のフレームワークでは、嗜好が似ているユーザに、彼らが好むであろうアイテムを推薦できるようにユーザの嗜好を推定するが、本論文で提案するフレームワークではユーザ 1 人ごとにユーザの管理基準を推定する。

### 2.3 アクティブラーニングに関する研究

近年では、機械学習モデルと人間の対話によるインタラクティブな推薦システム [7] に関する研究が盛んに行われている。ユーザの興味関心を取得する方法として、アクティブラーニングの手法をとっている研究 [8] もある。アクティブラーニングでキーとなっている考え方は、たとえば分類問題であれば、分類するのに不確実なものを選択し問い合わせることである。このように、アクティブラーニングはモデルの性能を向上させるために、性能を向上させられると考えられるようなデータをオラクルに問い合わせることによる半教師あり学習を行う。本研究では、ユーザが潜在空間中に配置された文書の位置にフィードバックを与えるが、このフィードバックをオラクルとして潜在空間学習を行う。

### 2.4 ネットワークの表現学習に関する研究

近年、ネットワークの構造を分析するための研究として、ネットワークの埋め込み手法が注目を集めている。ネットワークの埋め込み手法として有名なものに、ランダムウォークを用いてノードの分散表現を獲得する Deepwalk [9] や Node2vec [10] が、直接エッジが貼られていないようなノード同士でも、同じようなノードの特徴が見られればそれを考慮して埋め込む LINE [11] などが挙げられる。また、Kipf と Welling が提案した GCN [12] は、ネットワークの構造を加味しながらノードの潜在ベクトルを学習する。さらに、この GCN を用いてオートエンコーダ [13] にネットワークデータを当てはめた Graph Autoencoders (GAE) や、オートエンコーダに確率分布を仮定することによる変分オートエンコーダをネットワークデータに当てはめた、Variational Graph Autoencoders (VGAE) を

用いることでネットワークの表現学習をする手法 [14] などが存在する。GAE や VGAE は 2 層のグラフ畳み込みネットワークを持ち、エンコーダ・デコーダのアルゴリズムによって隣接行列を再構成する。本論文では、文書同士の関係をもとに文書を空間に埋め込む方法として、GAE を用いる。

### 3 予備知識

本節では、本論文で提案するフレームワークで用いられる、いくつかの前提知識について記述する。

**異種混合ネットワーク**：我々の実世界に普遍的に存在するソーシャルネットワークや文献情報のネットワークは、異種混合ネットワーク [15] [16] というもので構成されている。異種混合ネットワークは複数種類の属性のノードと、複数種類の関係を表現するエッジから構成されており、有向グラフ  $G(\mathcal{V}, \mathcal{E})$  において、ノードから属性に写像する関数  $\tau: \mathcal{V} \rightarrow \mathcal{A}$  と、エッジから属性に写像する関数  $\rho: \mathcal{E} \rightarrow \mathcal{R}$  をもつ。加えて、 $\mathcal{A}$  と  $\mathcal{R}$  はそれぞれ、ノードの属性とエッジの関係を表すラベルの集合だが、これらは  $|\mathcal{A}| + |\mathcal{R}| > 2$  という関係を満たす、たとえば、文献情報ネットワークでは、文書 (P)、著者 (A)、文書中に登場する用語 (T) そして、文書が発表された年 (Y) などの属性が考えられる。そして、これらの属性をもとに、文書と著者の間の共著関係といった、意味情報を含んだ関係を表現することが可能である。

**meta-path**：異種混合ネットワークにおいて、異なる属性のノードを結ぶパスを meta-path と呼ぶ。meta-path は、ノードの属性  $A_1 \in \mathcal{A}$  と  $A_{l+1} \in \mathcal{A}$  の間の二項関係  $R = R_1 \circ R_2 \circ \dots \circ R_l$  を用いて、 $A_1 \xrightarrow{R_1} A_2 \xrightarrow{R_2} \dots \xrightarrow{R_l} A_{l+1}$  で表される。ただし  $\circ$  は二項関係を合成する演算子である。本論文では文書同士の関係を対象とするため、meta-path の始点と終点の両方が文書 (P) であるものを考える。たとえば、「文書 (P) – 著者 (A) – 文書 (P)」という meta-path は、共著の論文であるという関係を表している。

### 4 問題定義

本節では、本研究で対象とする問題の定義について記述する。本論文で用いる主な記号とその定義は表 1 に示す。本研究では入力として、重みなし無向グラフ  $G$  の隣接行列  $\mathbf{A}$  と各ノードの特徴量  $\mathbf{X}_D$  が与えられるとする。ここで、 $\mathbf{A}$  はある meta-path による隣接行列である。また、ユーザの文書管理の基準を推定するために、ユーザとのインタラクションの過程も入力として与える。本研究の目的は、未知の文書集合  $\mathcal{Q}$  が与えられたときに、ユーザのインタラクションの結果をもとに  $\mathcal{Q}$  が潜在空間中でユーザが求めるような位置に配置できるかを扱う。したがって、ユーザからのフィードバックによって重み付けされた隣接行列  $\tilde{\mathbf{A}}$  の潜在変数  $Z_{\mathcal{Q}}$  を出力する。

### 5 提案手法

本節では、本論文で提案する Human-in-the-loop 潜在空間

表 1: 主な記号と定義

記号	定義
$d \in \mathcal{D}$	文書の集合
$p \in \mathcal{P}$	meta-path の集合
$v \in \mathcal{V}$	ノードの集合
$e \in \mathcal{E}$	エッジの集合
$G(\mathcal{D}, \mathcal{E})$	無向グラフ
$A \in \mathcal{A}$	ノードの属性
$R \in \mathcal{R}$	エッジの属性
$L$	埋め込み次元数
$\gamma_p \in \Gamma_P$	ある meta-path $p$ の寄与率
$w_p \in \mathcal{W}_P$	ある meta-path $p$ の重み
$\mathbf{A}^p \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	ある meta-path $p$ についての隣接行列
$\tilde{\mathbf{A}} \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	$W$ で重み付けされた隣接行列
$\tilde{\mathbf{A}}_{d_i, d_j}$	$\tilde{\mathbf{A}}$ の $(i, j)$ 要素
$\mathbf{X} \in \mathbb{R}^{\mathcal{D} \times \mathcal{D}}$	ノードの特徴量
$\mathbf{z} \in \mathbb{R}^L$	埋め込みベクトル
$\mathbf{Z} \in \mathbb{R}^{\mathcal{D} \times L}$	$\mathbf{z}$ の行列
$(z, \hat{z}) \in \hat{\mathbf{Z}}$	ユーザがフィードバックを与えたノード
$\theta$	デコーダのパラメータ
$\phi$	エンコーダのパラメータ
$q \in \mathcal{Q}$	未知の文書
$z_q \in \mathcal{Z}_{\mathcal{Q}}$	未知の文書の潜在変数

学習のためのモデルに関して記述する。提案手法は次の 3 つの段階からなる。

- (1) GAE による各文書の潜在ベクトル  $\mathbf{z}_d$  の学習 (5.1)
- (2) ユーザからのフィードバックを元に、各 meta-path の重み付けによる隣接行列の再計算 (5.2)
- (3) meta-path の重み付けによる、ユーザの文献管理基準の推定 (5.3)

#### 5.1 GAE による各文書の潜在ベクトル $\mathbf{z}_d$ の学習

ここでは、ユーザの文書管理基準によって重み付けされた隣接行列  $\tilde{\mathbf{A}}$  を用いて、以下に続く GAE に基づいて、各文書の潜在空間中で埋め込みベクトルを獲得する。なお、ユーザの文書管理基準の推定方法については、5.3 にて説明する。

**エンコーダの構成** エンコーダは 2 層の GCN [12] によって構成され、各文書の潜在ベクトルに関する提案分布  $q_{\phi}(\tilde{\mathbf{Z}}_D | \mathbf{X}_D, \tilde{\mathbf{A}})$  に基づき、以下の式で潜在表現  $\mathbf{z}_d$  を推定することによって  $\phi$  が計算される。また、本提案手法では、各ノードの特徴量  $\mathbf{X}_D$  は各文書の文書番号のみを与える。これは、たとえば特徴量として各文書のタイトルを Bag-of-Words 化したのもを与えた場合、ネットワークの構造に基づいて特徴量を畳み込む際に、ネットワークの隣接関係と各特徴量を考慮に入れながら畳み込まれるが、提案手法では meta-path ベースの隣接関係のみを参照し畳み込むためである。

ここで、

$$\mathbf{Z} = GCN_{\phi}(\mathbf{X}_D, \tilde{\mathbf{A}}) \quad (1)$$

であり,  $GCN_\phi(\mathbf{X}_D, \tilde{\mathbf{A}})$  は,

$$GCN_\phi(\mathbf{X}_D, \tilde{\mathbf{A}}) = \mathring{A}ReLU(\mathring{A}\mathbf{X}_D\mathbf{W}_0)\mathbf{W}_1 \quad (2)$$

で表される. ここで,  $\phi = \{\mathbf{W}_0, \mathbf{W}_1\}$  であり,  $GCN$  のパラメータ集合である. ただし,  $\mathbf{W}_0$  は 1 層目の重み,  $\mathbf{W}_1$  は 2 層目の重みである. また,  $\mathring{A}$  は,

$$\mathring{A} = \mathbf{D}^{-\frac{1}{2}}\tilde{\mathbf{A}}\mathbf{D}^{-\frac{1}{2}} \quad (3)$$

で表される.

**デコーダの構成** デコーダでは各文書の潜在ベクトルに基づくエッジの生成確率  $p_\theta(\tilde{\mathbf{A}}|\mathbf{Z}_D)$  を計算することで隣接行列  $\tilde{\mathbf{A}}$  を再構成する. ただし,

$$p_\theta(\tilde{\mathbf{A}}|\mathbf{Z}_D) = \prod_{d_i \in \mathcal{D}} \prod_{d_j \in \mathcal{D}} p_\theta(\tilde{\mathbf{A}}_{d_i, d_j} | z_{d_i}, z_{d_j}) \quad (4)$$

である. 本研究で提案する生成モデルでは, 各潜在表現のユークリッド距離によってデコーダを構成する. これはユーザがシステムにフィードバックを与える際に, 文書同士の距離でフィードバックを与えることから, より近くに配置された文書同士のエッジの生成確率を大きくすることを狙いとしている.

$$p_\theta(\tilde{\mathbf{A}}_{d_i, d_j} | z_{d_i}, z_{d_j}) = \sigma\left(\frac{a}{\|z_{d_i} - z_{d_j}\|_2} + b\right) \quad (5)$$

ここで,  $\sigma(\cdot)$  はシグモイド関数である.  $\theta = \{a, b\}$  はパラメータ集合である.

**目的関数** 本研究の目的関数は以下のクロスエントロピー誤差関数であり, これを最大化するようにパラメータ  $\phi, \theta$  を学習する.

$$\phi, \theta = \arg \max_{\phi, \theta} \log p_\theta(\tilde{\mathbf{A}} | GCN_\phi(\mathbf{X}_D, \tilde{\mathbf{A}})) \quad (6)$$

$$= \arg \max_{\phi, \theta} \sum_{d_i \in \mathcal{D}} \sum_{d_j \in \mathcal{D}} \log p_\theta(\tilde{\mathbf{A}}_{d_i, d_j} | z_{d_i}, z_{d_j}) \quad (7)$$

## 5.2 インタラクションによる隣接行列の再構成

提案手法では, 潜在空間中に配置された文書の位置に対し, ユーザがフィードバックを与えることによる Human-in-the-loop 潜在空間学習によって, 隣接行列を再構成し, ユーザがどの meta-path を重要だと考えているかを推定する. より具体的には, 各 meta-path による隣接行列について, フィードバックが与えられたノードに対し, そのノードが移動する前に周辺にいた上位  $k$  件のノードとはエッジが貼られていないとみなす. さらに, 移動した後にそのノードの周辺に上位  $k$  件のノードとエッジが貼られるとみなして, 隣接行列を再構成する. 再構成後の隣接行列は, 以下の Algorithm 1 で表される. ただし,  $\mathcal{N}_\epsilon(\hat{\mathbf{Z}})$  は, ユーザがフィードバックを与えたノードにおいて, フィードバックを与えた後の状態の潜在空間中での近傍を表す. 加えて,  $\mathcal{N}_\epsilon(\mathbf{Z})$  は, ユーザがフィードバックを与えたノードにおいて, フィードバックを与えた前の状態の潜在空間中での近傍を表す. より具体的に図 2 を用いて説明する. ユーザが赤い

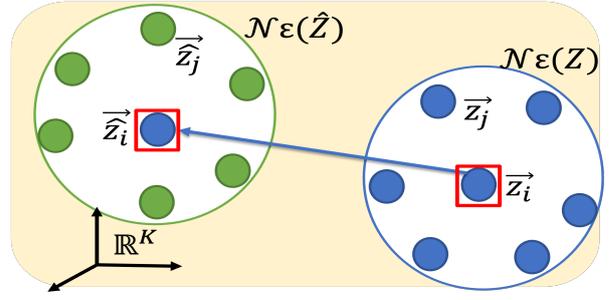


図 2: ユーザが与えるフィードバックのイメージ: 赤い枠で囲まれたノードについて潜在空間中の位置対し, 図中矢印の方向に向かってフィードバックを与えた場合,  $\mathcal{N}_\epsilon(\mathbf{Z})$  は, 図 2 の大きな青い円で囲まれたノードの潜在表現の集合を表す. 同様に,  $\mathcal{N}_\epsilon(\hat{\mathbf{Z}})$  は, 図 2 において, ユーザがフィードバックを与えた先にある大きな緑色の円で囲まれたノードの潜在表現の集合を表す.

枠で囲まれたノードについて, 潜在空間中の位置対し, 図中矢印の方向に向かってフィードバックを与えた場合,  $\mathcal{N}_\epsilon(\mathbf{Z})$  は, 図 2 の大きな青い円で囲まれたノードの潜在表現の集合を表す. 同様に,  $\mathcal{N}_\epsilon(\hat{\mathbf{Z}})$  は, 図 2 において, ユーザがフィードバックを与えた先にある大きな緑色の円で囲まれたノードの潜在表現の集合を表す.

### Algorithm 1 インタラクションによる隣接行列の再構成

**Input:**  $\hat{\mathbf{Z}}$

**Output:**  $\{\mathbf{A}^p | p \in \mathcal{P}\}$

```

1: for  $p \in \mathcal{P}$  do
2:   for  $(z, \hat{z}) \in \hat{\mathbf{Z}}$  do
3:     for  $\hat{z}_j \in \mathcal{N}_\epsilon(\hat{z}_i)$  do
4:       if  $\mathbf{A}_{d_i, d_j}^p \geq 1$  then
5:          $\mathbf{A}_{d_i, d_j}^p \leftarrow 0$ 
6:       end if
7:     end for
8:     for  $\hat{z}_j \in \hat{\mathcal{N}}_\epsilon(\hat{z}_i)$  do
9:        $\mathbf{A}_{d_i, d_j}^p + 1$ 
10:    end for
11:  end for
12: end for

```

## 5.3 meta-path の重み付けによる, ユーザの文献管理基準の推定

ここでは, ユーザの文献管理の基準を推定することで, 各 meta-path に対し重み付けをする. meta-path の重みを計算するために, まず meta-path  $p$  の寄与率として対数尤度を計算する.

$$\gamma_p = \sum_{d_i \in \mathcal{D}} \sum_{d_j \in \mathcal{D}} \mathbf{A}_{d_i, d_j}^p \log p_\theta(\mathbf{A}_{d_i, d_j} | z_{d_i}, z_{d_j}) \quad (8)$$

その後, 全ての meta-path の寄与率の合計に対するそれぞれの meta-path の寄与率の比を, その meta-path の重みとして

算出する.

$$w_p = \frac{\gamma_p}{\sum_{q \in P} \gamma_q} \quad (9)$$

最後に、各 meta-path ベースの隣接行列  $A^p$  にそれぞれの meta-path の重み  $w$  をかけ、それを足し合わせたものを重み付けされた隣接行列  $\tilde{A}$  とする.

$$\tilde{A} = \sum_{p \in P} w_p A^p \quad (10)$$

## 6 実験

本節では、データ収集フェイズによって得られたデータを用いて、未知の文書に対し、本提案手法がどれだけ正確に被験者が求める位置に配置できるかを検証する.

### 6.1 実験設計

**予測精度の評価実験のためのデータ収集フェイズ**：評価実験のためのデータ収集フェイズでは、図3のようなインターフェイスを用いてある文書について、その文書を潜在空間中において配置されて欲しい位置に動かしてもらおうとい、そのログを取った。表示される文書は、それぞれの被験者がこれまでに参考文献として目を通したことのあるもので、被験者には表示される文書について、毎回、文書のタイトル、著者、発表された会議、発表された年を、それぞれ文書がもつ情報として与えた。文書の数については、被験者1人あたり50件を操作してもらった。

**予測精度の評価実験フェイズ**：評価実験フェイズでは、データ収集フェイズで獲得したデータを利用して、提案手法の有効性を検討した。具体的には、被験者が求める潜在空間の最後の状態を対象に、そのデータを5分割し、10件ずつ学習データを増やしたときにテストデータの精度がどのように変化するかを検証した。

### 6.2 評価指標

評価には Liang らの研究 [17] にて用いられている Recall@k と nDCG@k を用いる。なお、今回の実験の評価では k=6 を行った。

評価指標である Recall@k は次の式で表される。

$$Recall@k = \frac{|U \cap P_k|}{|U|} \quad (11)$$

ここで、 $U$  はユーザが実際に配置した潜在空間中で、テストデータに対しユークリッド距離が小さい上位  $k$  件の文書の集合、 $P_k$  はモデルが予測した潜在ベクトルのうち、テストデータに対しユークリッド距離が小さい上位  $k$  件の集合である。

また nDCG@k は、DCG@k の値を、最も理想的な、すなわちモデルの予測が全て正しかったと仮定したときの DCG@k の値で割ったものである。すなわち、

$$nDCG@k = \frac{DCG@k}{IdealDCG@k} \quad (12)$$

ただし、DCG@k は

$$DCG@k = \frac{2^{rel_i} - 1}{\log_2(i + 1)} \quad (13)$$

で表される。評価で用いる Recall@k, nDCG@k はどちらも値が大きいほど精度が良いことを示している。

### 6.3 ベースライン手法

本論文の提案手法と比較したベースライン手法には次に示すものを用いた。

- **VGAE**： VGAE は Kipf ら [14] が提案したネットワークの表現学習の手法である。評価実験では、4節の(5)式で構成されていたデコーダを通常の GAE と同じように各潜在表現同士の内積：

$$p_{\theta}(\mathbf{A}_{d_i, d_j} | z_{d_i}, z_{d_j}) = \sigma(z_{d_i} \cdot z_{d_j}) \quad (14)$$

で構成し、本提案手法を当てはめた。ただし、 $\sigma(\cdot)$  はシグモイド関数である。

また、用いる meta-path の数を変更することで、予測の精度に変化が見られるかを調べた。用いた meta-path は以下の通りである。また、各 meta-path の意味については以下の表2に示す。

- **ALL+others(PAP, PTP, PYP, PVP, others)**： PAP, PTP, PYP, PVP に加えて、これらの meta-path では抽出できなかった関係 (others) を隣接行列に取り入れたもの。

- **ALL(PAP, PTP, PYP, PVP)**： PAP, PTP, PYP, PVP によるもの。

- **PAP Only**： PAP のみで隣接行列を構成したもの。
- **PTP Only**： PTP のみで隣接行列を構成したもの。
- **PYP Only**： PYP のみで隣接行列を構成したもの。

表 2: 各 meta-path の意味

meta-path	意味
PAP	共著経験のある著者同士の論文
PTP	同じ単語について言及している論文
PYP	同じ年に発表された論文
PVP	同じ会議で発表された論文

### 6.4 実験結果

実験結果は次の図 4a~図 5b に示す。それぞれの図の赤い線は、入力として与える隣接行列が (6.3) で挙げた **ALL+others** によって構成されたものである。また、図 4a と図 5a の青い線は、(6.3) で挙げた **VGAE** を提案手法に当てはめ、入力として **ALL+others** で構成される隣接行列を与えたものである。また、図 4b と図 5b の緑、桃、紫色の線は提案手法において、入力として与える meta-path の種類を限定したものである。

図 4a と図 4b は、フィードバックの回数を変化させたときの recall@k の値の変化を、図 5a と図 5b はフィードバックの回数を変化させたときの nDCG@k の値の変化を表している。図を見ると、提案手法は、今回比較した手法の全てと比べて、フィードバックの回数を増やしたことで精度が向上することが

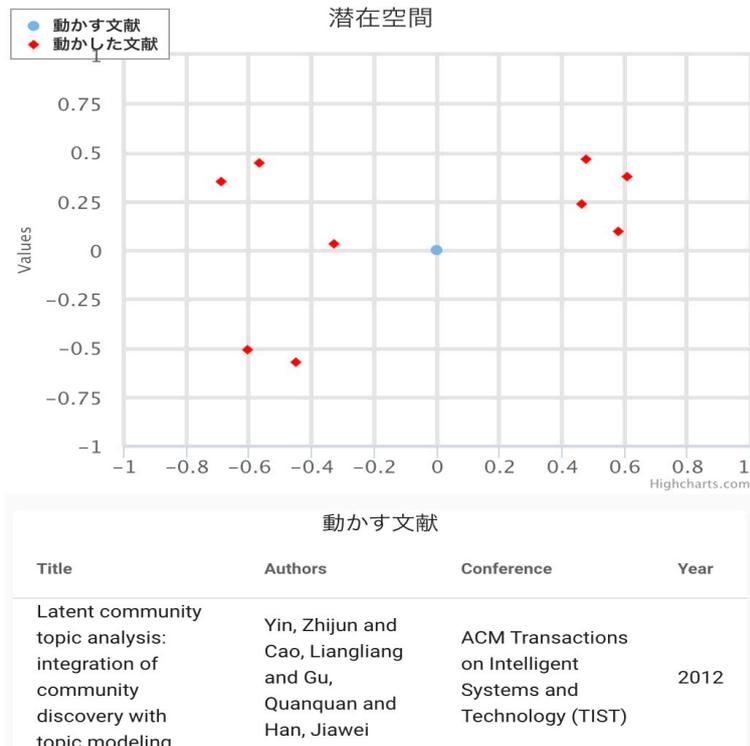


図 3: 被験者実験で使用したインターフェイス：毎回潜在空間中の原点の位置に置かれる点を、被験者が好みの位置に配置する。被験者には、文献が持つ情報として、文書のタイトル、著者、会議名、発表された年を表示した。

確認された。

また、図 6a と図 6b には、実際の潜在空間の様子を示す。図 6a はある被験者が求めた潜在空間の最終状態を表す。図 6b は提案手法による潜在空間の状態である。図中の紫色の点が未知の文書であり、図 6a の桃色の点が、紫色の点に対してユーザが近くに配置した 6 つの文書である。図 6b は、紫色の点が未知の文書として与えられたときのモデルの予測結果である。プログラムの都合上、未知の文書以外の潜在表現についても再計算されてしまうため、潜在空間の形そのものは変わってしまっているものの、紫色の点と桃色の点の相対的な位置関係は比較的近いものになっている、また、文書同士の相対的な位置関係が重要であるため、図 6a と図 6b 中の横軸、縦軸の数値の大きさ自体に意味はない。

## 7 議論

本節では上記の実験結果に対する議論と考察を行う。

### 7.1 提案手法と既存手法を比較した場合

ここでは、提案手法と、VGAE を使用した場合での実験結果に対する考察を行う。図 4a、図 5b を見ると、フィードバックを 40 回与えた段階の recall@k、nDCG@k の値はどちらも提案手法の方が大きいという結果になった。提案手法と既存手法の大きな違いは、生成モデルであるデコーダを提案手法では (5) 式で構成したのに対し、VGAE は (14) 式によって構成している。これによって、提案手法は潜在空間中で距離に近いノードほどエッジの生成確率が大きくなり、VGAE では 2 つのノード

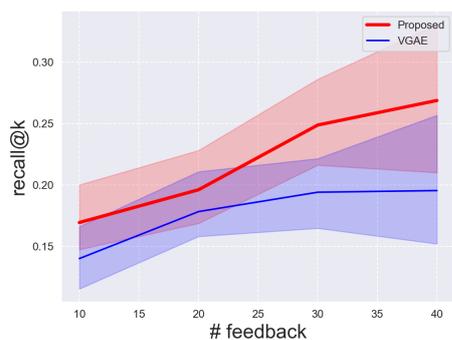
ドで作られる潜在空間中の角度が大きいものほどエッジの生成確率が大きくなる、ユーザは潜在空間中で、文書同士の角度ではなくユークリッド距離に基づいてフィードバックを与えることから、文書同士のユークリッド距離によってエッジの生成確率を計算する提案手法の方が精度が良くなったと言える。

### 7.2 提案手法と使用する meta-path の種類を限定した場合

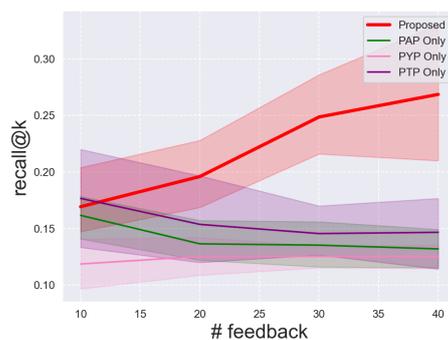
ここでは、提案手法と、使用する meta-path の種類を限定した場合での実験結果に対する考察を行う。図 4b を見ると、提案手法 (図の赤の線) はフィードバックの回数が増えると recall@k の値が大きくなっているのに対し、使用する meta-path の種類を限定した場合 (図の緑・桃・紫の線) ではフィードバックの回数が増えても recall@k の値に大きな変化はない、もしくは値が減少するという結果となった。また、図 5b から同様に、フィードバックの回数を増やしたことで、使用する meta-path の種類を限定した場合と比べて、提案手法は nDCG@k の値が増加していく様子がわかる。このことから、研究者の文献管理の基準は単一のものではなく、複数の meta-path を用いてユーザの文献管理の基準を推定できる提案手法が有効であったことが確認できる。

### 7.3 今後の展望

最後に実験結果に基づいて、課題と今後の展望について記述する。図 4a~図 5b の赤い線は、使用した meta-path が ALL+others の場合である。今回の実験では、基本的にはこの赤い線が最も良い精度を実現していた。このようになった要因の 1 つとして、多くの meta-path を使用したということが

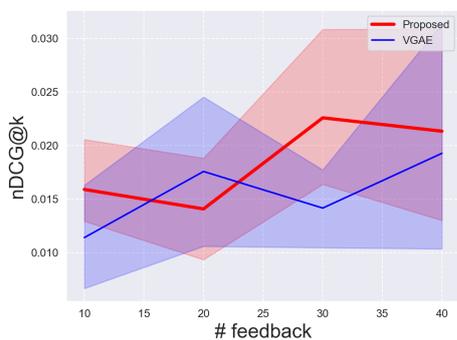


(a) 提案手法と VGAE の recall@k の値

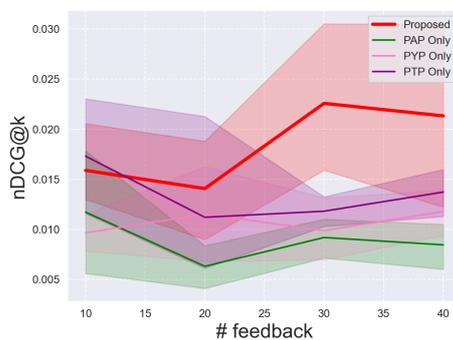


(b) 提案手法と、meta-path の種類を限定したときの recall@k の値

図 4: 横軸はユーザからのフィードバックの回数, 縦軸は recall@k の値を表す. 設定したベースライン手法と比較して, 提案手法 (赤) はフィードバックの回数を増やすと recall@k の値が大きくなっている.

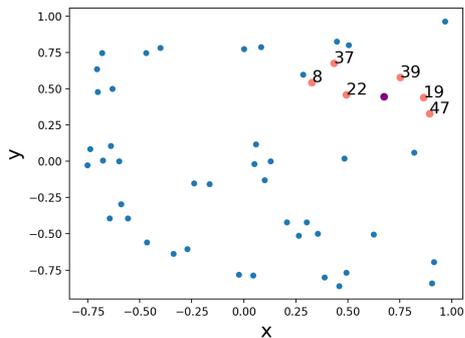


(a) 提案手法と VGAE の ndcg@k の値

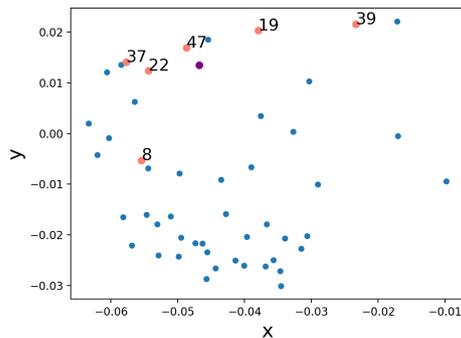


(b) 提案手法と、meta-path の種類を限定したときの ndcg@k の値

図 5: 横軸はユーザからのフィードバックの回数, 縦軸は nDCG@k の値を表す. 平均値で見ると, 提案手法 (赤) は設定したベースライン手法よりも大きな nDCG@k の値を獲得している.



(a) ある被験者が求める潜在空間



(b) モデルが出力した潜在空間

図 6: 未知の文書 (紫) が与えられたとき, 図 6a はある被験者が求める文書配置を, 図 6b は提案手法が予測した文書配置を表す. また, 桃色の点は紫色の点に対して被験者が近くに配置した 6 つの文書である. 潜在空間は再計算されるため, 形自体は異なるが, 紫色の点と桃色の点の相対的な位置関係は, 図 6a と図 6b で, 比較的似ていると言える.

挙げられるが, 今回の実験では, 今回用意した meta-path だけでは抽出できなかった関係 (others) も含めて隣接行列を構成してしまっている. したがって, 今後の展望の 1 つ目として, 複雑な meta-path を導入することが挙げられる. 今回の実験では, 「文書 (P) - 著者 (A) - 文書 (P)」のような最小構成

の meta-path を使用した. 今後は「文書 (P) - 著者 (A) - 文書 (P) - 単語 (T) - 文書 (P) - 著者 (A) - 文書 (P)」のように複雑な meta-path を導入することで, 今回の実験で others とされてしまったような関係を抽出できるようにする. そして, 文書同士の関係性の表現力を高めることによって, ユー

ザの文書の管理基準をより良い精度で推定できるようにすることが必要である。

また、本実験において、多くの場合ではフィードバックの回数が少ない段階から設定したベースライン手法よりも高い精度を実現していたが、たとえば図 5a では、インタラクションの回数が増えた段階から提案手法の精度がベースラインを追い抜くという結果になった。したがって、今後の展望の 2 つ目として、少ないフィードバックの回数で効率よくモデルが学習できるような手法の考案が挙げられる。これを実現するためには、どのような文書に対してユーザがフィードバックを与えればモデルが効率よく学習できるのかについて、考察が必要である。今後は、ユーザに問い合わせる文書の順番を変更して被験者実験を行うなどして、ユーザが操作する回数が少ない段階から、より高い精度になるようなフレームワークの考案を計画している。また、今回の実験では、Recall@k や nDCG@k の値には、まだまだ精度が向上する余地が残っている。今後は、上に記したようなフレームワークの考案の計画と同時に、高い精度を実現するための仕組みも必要になってくると考えられる。

## 8 ま と め

本研究では、Human-in-the-loop 潜在空間学習を適用することでユーザの文献管理の基準を獲得し、その基準に基づいて文書を低次元空間に埋め込む手法を提案した。提案手法ではまず、ユーザの文書の管理基準を推定し、各 meta-path による隣接行列に重みをつける、次に、Graph Autoencoders (GAE) を用いて各文書の潜在表現  $z$  の学習を行う。そして、ユーザから与えられたフィードバックをもとに各 meta-path による隣接行列を作り直し、再度隣接行列の重み付けを行い潜在表現  $z$  の学習をすることで、より良い潜在空間を作る。また、ユーザは潜在空間中において、文書同士の距離に基づいてフィードバックを与えるため、(5) 式のように、文書同士のユークリッド距離によってデコーダを構成した。

提案手法の有効性を検討するために、未知の文書を被験者が求める位置にどれだけ正確に配置できるかを評価する実験を行った。実験の結果、提案手法はベースラインの手法と比較して、未知の文書をユーザが求める位置に配置できていた。加えて、複数の meta-path を用いて実験した場合と、使用する meta-path の種類を限定して実験した場合で比較すると、複数の meta-path を用いた場合の方が精度が高く、ユーザが様々な基準で文書を管理する場合でも、提案手法が有効であることが明らかになった。

今後の展望として、(1) 複雑な meta-path の導入 (2) 少ないフィードバックの回数で効率よくモデルが学習できるような手法の考案の 2 つが挙げられる。この 2 つを実現することで、より良い Human-in-the-loop 潜在空間学習のフレームワークの構築が行える。

## 謝 辞

本研究の一部は JST CREST (# JPMJCR16E3), JSPS 科

研費 (JP21H03552) および株式会社熊谷組の支援による。

## 文 献

- [1] Michael Fire and Carlos Guestrin. Over-optimization of academic publishing metrics: Observing goodhart's law in action. *GigaScience*, Vol. 8, , 06 2019.
- [2] Enrico Francese. Usage of reference management software at the university of torino. *JLIS.it*, Vol. 4, pp. 145–174, 2013.
- [3] Bo-Christer Björk. Electronic document management in construction – research issue and results. *Electronic Journal of Information Technology in Construction, ITcon, 2003*, Vol. 8, pp. 101-113, Vol. 8, , 05 2003.
- [4] Clarivate. Endnote. <https://endnote.com>. (Accessed on 12/30/2021).
- [5] ELSEVIER. Mendekey. <https://www.mendeley.com/>. (Accessed on 12/30/2021).
- [6] Yu Zhu, Yu Gong, Qingwen Liu, Yingcai Ma, Wenwu Ou, Junxiong Zhu, Beidou Wang, Ziyu Guan, and Deng Cai. Query-based interactive recommendation by meta-path and adapted attention-gru. *CIKM '19*, p. 2585–2593, New York, NY, USA, 2019. Association for Computing Machinery.
- [7] Michael Jugovac and Dietmar Jannach. Interacting with recommenders—overview and research directions. *ACM Trans. Interact. Intell. Syst.*, Vol. 7, No. 3, sep 2017.
- [8] Neil Rubens, Mehdi Elahi, Masashi Sugiyama, and Dain Kaplan. *Active Learning in Recommender Systems*, pp. 809–846. Springer US, Boston, MA, 2015.
- [9] Bryan Perozzi, Rami Al-Rfou, and Steven Skiena. Deepwalk: Online learning of social representations. p. 701–710, 2014.
- [10] Aditya Grover and Jure Leskovec. node2vec: Scalable feature learning for networks, 2016.
- [11] Jian Tang, Meng Qu, Mingzhe Wang, Ming Zhang, Jun Yan, and Qiaozhu Mei. Line: Large-scale information network embedding. *Proceedings of the 24th International Conference on World Wide Web*, May 2015.
- [12] Thomas N. Kipf and Max Welling. Semi-supervised classification with graph convolutional networks. In *International Conference on Learning Representations (ICLR)*, 2017.
- [13] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. In *2nd International Conference on Learning Representations, ICLR 2014, Banff, AB, Canada, April 14-16, 2014, Conference Track Proceedings*, 2014.
- [14] Thomas N Kipf and Max Welling. Variational graph auto-encoders. *NIPS Workshop on Bayesian Deep Learning*, 2016.
- [15] Yizhou Sun and Jiawei Han. Meta-path-based search and mining in heterogeneous information networks. *Tsinghua Science and Technology*, Vol. 18, No. 4, pp. 329–338, 2013.
- [16] Yuxiao Dong, Nitesh V. Chawla, and Ananthram Swami. Metapath2vec: Scalable representation learning for heterogeneous networks. In *Proceedings of the 23rd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '17*, p. 135–144, New York, NY, USA, 2017. Association for Computing Machinery.
- [17] Dawen Liang, Rahul G. Krishnan, Matthew D. Hoffman, and Tony Jebara. Variational autoencoders for collaborative filtering, 2018.