

# 量子モデルを利用した推薦システムの提案

和山 弘<sup>†</sup> 杉山 一成<sup>†</sup>

<sup>†</sup> 京都大学大学院情報学研究科 〒606-8501 京都市左京区吉田本町

E-mail: <sup>†</sup>wayama.hiroshi.84h@st.kyoto-u.ac.jp, <sup>††</sup>kaz.sugiyama@i.kyoto-u.ac.jp

**あらまし** クリックなどのユーザからの暗黙的なフィードバックを用いた推薦システムにおいては、ユーザーとアイテムの間にインタラクションがあれば1を、なければ0を付け、隣接行列に基づいてモデルが作成されることが多い。ここで、隣接行列とは、ユーザーとアイテムを無向二部グラフとする対称行列である。しかしながら、隣接行列はすべて1か0かの二値で表現され、グラフの結びつきの強さを表現することが出来ないという問題がある。その一方で、連続時間の量子ウォークの概念を用いると、グラフ上の情報の伝播をシュレディンガー方程式に従う形で表現することが可能である。このようなモデルでは、古典的な隣接行列を用いた情報の伝播とは振る舞いが大きく異なる事が予想される。そこで本研究では、平均混合行列で量子ウォークを表現し、隣接行列に量子モデルの考えを応用する事により精度の高い推薦システムを提案する。

**キーワード** 推薦システム, グラフ, 量子情報

## 1 はじめに

推薦システムで利用される手法の一つに協調フィルタリング [1], [2] がある。協調フィルタリングはユーザーの購入やクリックなどの暗黙的なフィードバックを元に生成されるユーザーとアイテムの嗜好行列を元に推薦を行う手法である。協調フィルタリングは、さらに (1) モデルベースの方法と (2) メモリベースの方法に分けられる。(1) モデルベースの方法は、行列分解 [3] などにより、高次元のデータを低次元のデータとして表現し、特定のユーザーと親和性の高い商品を抽出する方法である。ユーザーとアイテムを学習可能なベクトルとして表現し、ユーザーとアイテムのインタラクションを二つのベクトルの内積で表現する。(2) メモリベースの方法は、同じような行動履歴を持つユーザーは同じような嗜好を持つという仮定の元に、コサイン類似度などを利用しユーザーの近傍探索を行い、似たようなユーザーを抽出する手法である。

ユーザーやアイテムの隣接情報を利用する手法として、GCN [4], NGCF [5], LightGCN [6] などのグラフ畳み込みネットワーク (GCN) が提案され、協調フィルタリングを利用した最先端の手法として盛んに研究されている。ユーザーとアイテムの隣接行列を利用して、ユーザー、アイテムそれぞれの初期ベクトルを隣接行列に従い伝播させるとことでそれぞれのベクトルを改善する手法である。LightGCN [6] では、従来のグラフ畳み込みネットワークを大幅に簡素化し、計算量を大幅に削減しても精度の良い推薦システムが提供可能である事が示されている。

最近では量子アルゴリズムを利用した推薦システムが提案されている。[7] では、嗜好行列が低ランク近似可能という前提条件の元、あるユーザー  $i$  に対する価値が高いアイテムを  $O(\text{poly}(k) \text{polylog}(mn))$  の計算量で嗜好行列の特異ベクトルをサンプリングを行う手法が提案されている。 $m$  がユーザー数、

$n$  がアイテム数、 $k$  が  $m \times n$  の嗜好行列のランクである。従来のコンピュータを利用した場合、特異値分解は  $O(\text{poly}(mn))$  の速度が必要なため、指数関数的な高速化を実現している。一方、低ランク近似を仮定すれば、従来のコンピュータを利用しても同程度の計算速度で特異ベクトルがサンプリング可能である事が示された [8]。さらに、Bai ら [9] は、量子空間グラフ畳み込みニューラルネットワークというグラフ構造を持つ量子アルゴリズムを提案した。この研究では、これまで提案された畳み込み層のグラフ表現と、量子空間グラフ畳み込み層を統合してグラフを分類する深層学習のアーキテクチャを提案している。その中で、量子版ランダムウォークである量子ウォークが提案されている [10]。また、グラフ分類タスクを量子回路を用いて実行する研究も行われ、Zhang ら [11] は、量子グラフ畳み込みニューラルネットワークを提案している。

このように、近年グラフ構造や量子アルゴリズムなどを利用した推薦システムが提案され、活発に研究されている。

## 2 関連研究

### 2.1 グラフ畳み込みネットワーク

本節では、グラフ畳み込みネットワークを利用した最先端のモデルである LightGCN [6] を中心に説明を行う。

#### 2.1.1 嗜好行列と隣接行列

協調フィルタリングではユーザーがアイテムに対してクリックや購入などの暗黙的なフィードバックを受けた場合は1、その他は0とする嗜好行列 (user-item interaction matrix) が利用される。嗜好行列は以下のように定義される。

$$\mathbf{R}_{u,i} = \begin{cases} 1, & \text{if } (u, i) \text{ interaction is observed} \\ 0, & \text{otherwise} \end{cases}$$

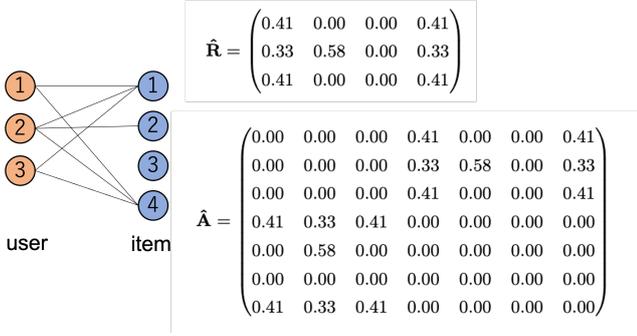


図1 2部グラフと正規化嗜好行列と隣接行列の例. ユーザーがアイテムに対してクリックや購入などのフィードバックを行うとグラフ上で頂点同士が辺で結ばれる.  $\hat{\mathbf{R}}$  と  $\hat{\mathbf{A}}$  は正規化された嗜好行列と隣接行列である.

### 2.1.2 隣接行列

隣接行列は嗜好行列を用いて以下のように定義される.

$$\mathbf{A} = \begin{pmatrix} \mathbf{0} & \mathbf{R} \\ \mathbf{R}^T & \mathbf{0} \end{pmatrix}$$

行列の大きさはユーザー数を  $m$ , アイテム数を  $n$  とすると,  $(m+n) \times (m+n)$  である. グラフ畳み込みネットワークではこの形の隣接行列では計算が収束しないことが知られており, 正規化が行われる. 各ユーザーの interaction 数の合計を対角成分に持つ正方行列  $\mathbf{D}_U = \text{Diag}(\mathbf{R} \cdot \mathbf{1})$  と, 各アイテムの interaction 数の合計を対角成分に持つ正方行列  $\mathbf{D}_I = \text{Diag}(\mathbf{1}^T \cdot \mathbf{R})$  を利用して, 正規化嗜好行列は以下のように定義される.

$$\hat{\mathbf{R}} = \mathbf{D}_U^{-\frac{1}{2}} \mathbf{R} \mathbf{D}_I^{-\frac{1}{2}}$$

この正規化嗜好行列を利用して正規化隣接行列が定義される.

$$\hat{\mathbf{A}} = \begin{pmatrix} \mathbf{0} & \hat{\mathbf{R}} \\ \hat{\mathbf{R}}^T & \mathbf{0} \end{pmatrix} \quad (1)$$

図1に, ユーザーとアイテムの2部グラフから, 正規化隣接行列を作成する図を示す.

### 2.1.3 LightGCN

LightGCN [6] では, グラフの接点を持つ情報を正規化嗜好行列に従って伝播させている. 最初にユーザーとアイテムをベクトルで表現し, それに対して隣接行列をかけることで情報の伝播を実現している. ユーザー  $u_k$  の潜在ベクトルを  $\mathbf{e}_{u_k}$ , アイテム  $j$  の潜在ベクトルを  $\mathbf{e}_{i_j}$  とすると,  $\mathbf{E}^{(0)}$  を以下のように定義できる.

$$\mathbf{E}^{(0)} = \begin{pmatrix} \mathbf{e}_{u_1} \\ \vdots \\ \mathbf{e}_{u_m} \\ \mathbf{e}_{i_1} \\ \vdots \\ \mathbf{e}_{i_n} \end{pmatrix}$$

この初期状態  $\mathbf{E}^{(0)}$  は, 正規化隣接行列に従って, グラフ上で隣接する頂点に伝播される.

$$\mathbf{E}^{(n+1)} = \hat{\mathbf{A}} \mathbf{E}^{(n)}$$

## 2.2 量子ウォーク

本節では, 量子ウォークの簡単な説明を行う. 量子ウォークはランダムウォークの量子版として研究されている. 量子ウォークは物理学で微視的な粒子の状態を記述するシュレディンガー方程式  $i\hbar \frac{d}{dt} |\psi\rangle = H|\psi\rangle$  に従い時間発展する. ハミルトニアン  $H$  が時間に依存しない場合,  $|\psi(t)\rangle = e^{\frac{H}{i\hbar}t} |\psi(0)\rangle$  と表現することができる.  $|\psi(0)\rangle$  は系の初期状態である. 一般に, ハミルトニアン  $H$  エルミート行列でありユニタリ行列ではないが,  $U = e^{\frac{H}{i\hbar}t}$  はユニタリ行列となり, この行列に従って系は時間発展する.

[10] では, 平均混合行列を用いて連続時間量子ウォークを表し, グラフ上の各頂点における時間平均的な量子状態の変化を論じている. ここでは, [10] に従い, 量子ウォークの定式化を行い, 簡単に説明する.

### 2.3 連続時間量子ウォーク

正規化隣接行列  $\hat{\mathbf{A}}$  をハミルトニアン  $H$  として考え, 以下の行列を定義する.

$$\hat{\mathbf{T}}(t) := \exp(it\hat{\mathbf{A}}) \quad (2)$$

$\hat{\mathbf{T}}(t)$  と  $\hat{\mathbf{T}}(t)$  の随伴行列 ( $\hat{\mathbf{T}}^\dagger(t) = \hat{\mathbf{T}}(-t)$ ) の要素積を計算した行列を  $\hat{\mathbf{M}}(t)$  と定義する.

$$\hat{\mathbf{M}}(t) := \hat{\mathbf{T}}(t) \circ \hat{\mathbf{T}}(-t) \quad (3)$$

$\circ$  は要素毎の積 (アダマール積) を表す.  $\hat{\mathbf{T}}(t)$  がユニタリ行列なので,  $\hat{\mathbf{M}}(t)$  の各行, 各列における和は常に1になる. よって,  $\hat{\mathbf{M}}(t)$  は確率行列として捉える事も可能である. そのように考えると,  $\hat{\mathbf{M}}(1)$  は, 量子状態  $|\psi(t)\rangle$  から  $|\psi(t+1)\rangle$  へと遷移させる遷移行列として定義する事ができる.

#### 2.3.1 スペクトル分解

$\hat{\mathbf{M}}(t)$  を具体的に計算するために  $\hat{\mathbf{A}}$  のスペクトル分解を利用する.  $\hat{\mathbf{A}}$  はエルミート行列であり, スペクトル分解が可能である. すなわち以下のように書くことができる.

$$\hat{\mathbf{A}} = \sum_r \lambda_r |r\rangle \langle r| \quad (4)$$

$|r\rangle$  は  $\hat{\mathbf{A}}$  の固有ベクトルあり,  $\lambda_r$  は  $|r\rangle$  に対する固有値である.  $\hat{\mathbf{A}}$  はエルミート行列なので, 各固有ベクトルは直交し,  $\langle i|j\rangle = \delta_{ij}$  が成立する.

式(4)を式(2)に代入して計算を進めると,  $\hat{\mathbf{T}}(t)$  を簡単に表現することができる. ここで,  $\langle i|j\rangle = \delta_{ij}$  を利用した.

$$\hat{\mathbf{T}}(t) = \sum_r \exp(i\lambda_r t) |r\rangle \langle r| \quad (5)$$

### 2.3.2 量子ウォークによる隣接行列

式 (5) を式 (3) に代入すると、遷移行列  $\hat{M}$  を具体的に計算することができる。ここで、 $\hat{M}$  の  $t=1$  の行列を量子ウォークによる隣接行列と考え、 $\hat{A}_Q$  と表現する。

$$\hat{A}_Q = \sum_r |r\rangle\langle r|^2 + 2 \sum_{r<s} \cos(\lambda_r - \lambda_s) |r\rangle\langle s| \circ |r\rangle\langle s| \quad (6)$$

ここで、 $|r\rangle\langle r|^2$  は  $|r\rangle\langle r| \circ |r\rangle\langle r|$  である。

### 2.4 量子ウォークの利点

古典的なランダムウォークはグラフ上の隣接する頂点をマルコフ過程に沿って時間発展するモデルである。量子ウォークはグラフ上の頂点を遷移することは同じだが、ユニタリ行列に沿って時間発展する。量子ウォークによる実装については、例えば、[12] で提案されている。また、通常グラフ上の畳み込みは、グラフの高周波成分がカットされるローパスフィルタの性質を持つが、量子ウォークはラプラシアン行列による高周波成分のフィルタリングを防ぐ効果があると考えられる。このように、量子ウォークの状態ベクトルは複素数を取り、古典的なランダムウォークとは異なる挙動を示すことが期待される。

## 3 提案手法

古典的な正規化隣接行列 (式 (1)) と量子ウォークに基づく量子化隣接行列 (式 (4)) の二つの隣接行列を利用した二つのモデルを提案する。

### 3.1 モデル 1

モデル 1 は単純に隣接行列として、式 (4) を利用するモデルである。初期化されたユーザーとアイテムの潜在ベクトルを  $\mathbf{E}^{(n+1)} = \hat{A}_Q \mathbf{E}^{(n)}$  という漸化式に沿って情報を伝播させる。

例えば、ユーザーの初期潜在ベクトルを  $\mathbf{e}_u^{(0)}$  と、アイテムの初期潜在ベクトルを  $\mathbf{e}_i^{(0)}$  とすると、次のレイヤー 1 におけるベクトル  $\mathbf{e}_u^{(1)}$  と  $\mathbf{e}_i^{(1)}$  は、

$$\begin{pmatrix} \mathbf{e}_u^{(1)} \\ \mathbf{e}_i^{(1)} \end{pmatrix} = \hat{A}_Q \begin{pmatrix} \mathbf{e}_u^{(0)} \\ \mathbf{e}_i^{(0)} \end{pmatrix}$$

と表現できる。このようにして次々と伝播される情報を最後には平均して、最終的なユーザーベクトル  $\mathbf{e}_u$ 、アイテムベクトル  $\mathbf{e}_i$  とする。

$$\mathbf{e}_u = \frac{1}{L+1} \sum_{k=0}^L \mathbf{e}_u^{(k)}; \quad \mathbf{e}_i = \frac{1}{L+1} \sum_{k=0}^L \mathbf{e}_i^{(k)}$$

推薦に必要なスコアはユーザーとアイテムの二つのベクトルの内積によって計算する。

$$\hat{r}_{ui} = \mathbf{e}_u^T \mathbf{e}_i \quad (7)$$

### 3.2 モデル 2

モデル 2 では従来の隣接行列 (式 (1)) と量子化隣接行列 (式

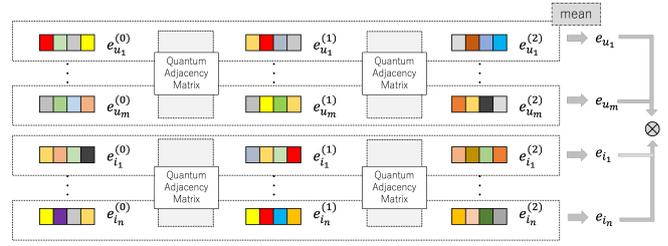


図 2 レイヤー数を 1 にしたモデル 2 のアーキテクチャ。

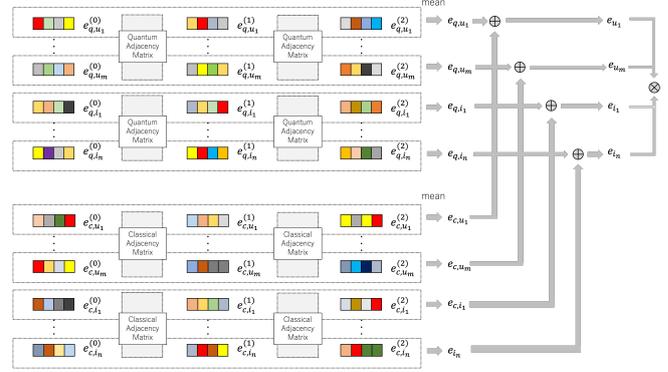


図 3 レイヤー数を 2 にしたモデル 2 のアーキテクチャ。

(4)) の両方を利用したハイブリッド形式のモデルである。

ユーザー、アイテムの初期潜在ベクトルとして、従来の隣接行列に作用させる潜在ベクトル  $\mathbf{e}_{c,u}^{(0)}$ 、 $\mathbf{e}_{c,i}^{(0)}$ 、量子化隣接行列に作用される潜在ベクトル  $\mathbf{e}_{q,u}^{(0)}$ 、 $\mathbf{e}_{q,i}^{(0)}$  をそれぞれ用意する。モデル 1 と同様に、これらの潜在ベクトルに対して、従来の隣接行列と量子化隣接行列をそれぞれ作用させる。

$$\begin{pmatrix} \mathbf{e}_{c,u}^{(1)} \\ \mathbf{e}_{c,i}^{(1)} \end{pmatrix} = \hat{A} \begin{pmatrix} \mathbf{e}_{c,u}^{(0)} \\ \mathbf{e}_{c,i}^{(0)} \end{pmatrix}$$

$$\begin{pmatrix} \mathbf{e}_{q,u}^{(1)} \\ \mathbf{e}_{q,i}^{(1)} \end{pmatrix} = \hat{A}_Q \begin{pmatrix} \mathbf{e}_{q,u}^{(0)} \\ \mathbf{e}_{q,i}^{(0)} \end{pmatrix}$$

最終的なユーザーとアイテムの潜在ベクトルは、それぞれの潜在ベクトルの和の平均として計算する。

$$\mathbf{e}_u = \frac{1}{L+1} \sum_{k=0}^L (\mathbf{e}_{c,u}^{(k)} + \mathbf{e}_{q,u}^{(k)}); \quad \mathbf{e}_i = \frac{1}{L+1} \sum_{k=0}^L (\mathbf{e}_{c,i}^{(k)} + \mathbf{e}_{q,i}^{(k)})$$

推薦スコアはモデル 1 と同様に式 (7) を用いて計算する。

### 3.3 最適化

提案するモデルの最適化パラメータは初期潜在ベクトルだけである。損失関数として、グラフ畳み込みネットワークで良く利用される BPR Loss (*Bayesian Personalized Ranking loss*) [13] を利用する。ユーザー数を  $m$ 、 $\mathcal{N}_u$  をユーザー  $i$  とインタラクションがあるアイテムの集合とすると、以下のように定義される。

$$L = - \sum_{u=1}^m \sum_{i \in \mathcal{N}_u} \sum_{j \notin \mathcal{N}_u} \ln \sigma(\hat{r}_{ui} - \hat{r}_{uj}) + \lambda \left\| \mathbf{E}^{(0)} \right\|^2$$

$\lambda$  は過学習を防ぐための  $L_2$  正則化係数である。

## 4 実験

### 4.1 データセット

今回提案したモデルを比較するために, [5] や [6] でも利用されている, MovieLens 1M, Gowalla, Yelp, Amazon-Books の4つのデータセットを利用した。いずれのデータセットも一般的に公開されており, 誰でも利用する事が可能である。MovieLens はユーザーの映画のレビュー結果をまとめたデータセットである。Gowalla はユーザーがチェックインすることで場所を共有する事ができるソーシャルネットワークサービスから作成されたデータセットである。Yelp はバーやレストランなどの飲食店を中心としたお店の評価がまとめられたデータセットである。Amazon-Books はその名の通りアマゾンの本を対象としたデータセットである。

いずれの4つのデータセットも推薦システムを評価する際に, よく利用される。これらのデータセットのユーザー数やアイテム数, インタラクション数, スパース性などは様々である。本実験では計算時間を有限に抑えるためにも, 各ユーザー, アイテムに対してインタラクション数に対して下限と上限を設けてフィルタリングを行った。

また, 今回提案するモデルの考察を深めるために, MovieLens-1Mを除く三つのデータセットについて, interactionの少ない疎なデータセットも作成した。すなわち, 今回利用するデータセットはMovieLens-1Mについてはinteractionの多いユーザーとアイテムを抽出した密なデータセット, 他の三つのデータセットについては, 密なデータセットに加え, interactionの少ないユーザーとアイテムを抽出した疎なデータセットも利用する。

表1と表2に, 密なデータセットのフィルタリングの条件と, 疎なデータセットのフィルタリングの条件を, それぞれ示す。

フィルタリングを行った後の統計情報は表3と表4の通りである。密なデータセットと疎なデータセットではinteractionにおおよそ一桁程度の差がある。

### 4.2 評価指標

評価指標には推薦システムの評価でしばしば利用される Recall(再現率), normalized discounted cumulative gain (nDCG) [14] を利用する。Recallは正解データの中でどれだけの割合を推薦できたかを表す指標で, 上位10件のアイテムを対象とする。nDCGはDCG(Discounted Cumulative Gain)を正規化することで得られる指標であり, ランキング上位のアイテムを正解できると高い評価を得ることができるという性質を持つ。nDCGも上位10件のアイテムを対象とする。

### 4.3 計算量

式(6)の計算量について議論する。表現を簡素化するために,

表1 Filtering conditions for dense dataset.

Dataset	Lower	Upper
MovieLens-1M	10	$\infty$
Gowalla	20	$\infty$
Yelp	35	$\infty$
Amazon-Books	32	$\infty$

表2 Filtering conditions for sparse dataset.

Dataset	Lower	Upper
Gowalla	3	25
Yelp	3	25
Amazon-Books	3	20

表3 Statistic of the dense datasets.

Dataset	#Users	#Items	#Interactions	Density
MovieLens-1M	6,041	3,261	998,539	0.005069
Gowalla	5,993	5,640	281,412	0.008326
Yelp	9,024	7,789	657,692	0.009357
Amazon-Books	8,465	9,107	608,034	0.007887

表4 Statistic of the sparse datasets.

Dataset	#Users	#Items	#Interactions	Density
Gowalla	3,518	3,412	13,483	0.00112
Yelp	5,482	5,202	11,665	0.00041
Amazon-Books	5,574	5,587	21,885	0.00070

式(6)を以下のように分解する。

$$\hat{\mathbf{A}}_{\mathbf{Q}} = Q_0 + Q_1 \quad (8)$$

$$Q_0 = \sum_r |r\rangle\langle r|^{\circ 2} \quad (9)$$

$$Q_1 = 2 \sum_{r < s} \cos(\lambda_r - \lambda_s) |r\rangle\langle s| \circ |r\rangle\langle s| \quad (10)$$

#### 4.3.1 $Q_0$ の計算量

$Q_0$  を正確に計算するには,  $|r\rangle\langle r|$  の計算に  $O(N^2)$  で, さらにその和を計算する必要があるため,  $O(N^3)$  の計算量が必要になる。

#### 4.3.2 $Q_1$ の計算量

$Q_1$  を正確に計算するには,  $|r\rangle\langle s|$  の計算に加えて,  $\sum_{r < s}$  の  $O(N^2)$  の計算量が必要のため, 合計で  $O(N^4)$  の計算量が必要になる。

以上の議論により,  $Q_0 + Q_1$  を正確に計算するには  $O(N^3) + O(N^4) \sim O(N^4)$  という膨大な計算量が必要になる。今回, 我々が用意したデータセットでは現実的な時間で計算が終わらない事が分かった。

表 5 実験結果. 上の表が密なデータセット 1, 下の表が疎なデータセット 2 に関する結果である. Ours-Model1 はモデル 1, Ours-Model2 はモデル 2 を示す. (†) はスコアの最大値である.

Dataset	MovieLens-1m		Gowalla		Yelp		Amazon-Books	
Method	Recall@10	nDCG@10	Recall@10	nDCG@10	Recall@10	nDCG@10	Recall@10	nDCG@10
BPR	0.1671	0.2589	0.1684	0.1425	0.0808	0.081	0.0872	0.0828
NGCF	0.1670	0.2558	0.1575	0.1363	0.0786	0.0774	0.0837	0.0788
LightGCN	0.1744	0.2688(†)	0.1711	0.1494(†)	0.0913(†)	0.0914(†)	0.0958	0.0927
Ours-Model1	0.1692	0.2619	0.1641	0.1424	0.0853	0.0847	0.0907	0.0860
Ours-Model2	0.1746(†)	0.2688(†)	0.1713(†)	0.1477	0.0869	0.0865	0.0984(†)	0.0943(†)

Dataset	Gowalla		Yelp		Amazon-Books	
Method	Recall@10	nDCG@10	Recall@10	nDCG@10	Recall@10	nDCG@10
BPR	0.0994	0.0669	0.4370	0.4313	0.3470	0.2709
NGCF	0.1201	0.0711	0.4473(†)	0.4393	0.3824	0.2831
LightGCN	0.1257	0.0782(†)	0.4433	0.4397(†)	0.3845	0.2927
Ours-Model1	0.068	0.0425	0.2806	0.2047	0.3903	0.2995(†)
Ours-Model2	0.1290(†)	0.0778	0.4444	0.4387	0.3914(†)	0.2989

#### 4.3.3 $Q_1$ サンプリング

$Q_1$  を正確に計算する事が不可能なので, 式 (10) の係数  $\cos(\lambda_r - \lambda_s)$  に着目し, この大きさに応じて確率的にサンプリングする. 今回の実験ではサンプリング数は 100,000 とした.

$$Q_1 \sim p(r, s) = \cos^2(\lambda_r - \lambda_s) \quad (11)$$

#### 4.4 実験結果

表 5 に, 実験結果を示す. 我々が提案する手法とこれまで提案されてきた BPR [13], NGCF [5], LightGCN [6] の手法で得られた結果を比較する. モデル 1 とモデル 2 にはハイパーパラメータがあるが, 最も結果が良いパラメータセットの結果を載せている.

なお, 従来の手法の評価には, 様々な推薦システムの評価を統一したインターフェースで利用できる RecBole [15] を利用した.

モデル 2 の一部のデータセットにおいて, 最良のベースラインである LightGCN を上回る結果も得ることができた. 従来の隣接行列に加え, 量子的な隣接行列の効果により, 隠れたグラフ間の隣接関係を表現することでできた事が原因だと考えられる. ただ, 上回った量も高々 5% 程度であり, 統計的に LightGCN よりも優れたモデルとは言いがたい.

##### 4.4.1 モデル 1 とモデル 2 の比較

モデル 1 とモデル 2 を比較すると, モデル 2 の方が良い結果が得られた. 原因としては, モデル 1 では従来の隣接行列を全く考慮していないため, グラフ間の情報の伝達が十分ではなかった事が考えられる.

#### 4.5 考察

今回提案した我々のモデルについて, 以下の 3 つの観点から (Research Questions, RQs) 考察を行う.

##### 4.5.1 [RQ1] $Q_0$ と $Q_1$

式 (8) において, 第一項である  $Q_0$  と第二項である  $Q_1$  が結果にどのような影響を与えるか確認する. 評価指標は Recall@10 で, Gowalla と Amazon-Books の密なデータセットを利用する. 表 6 に結果を示す.

Gowalla と Amazon-Books においてどちらも  $Q_1$  が無い方が良い Recall を得ることができた.  $Q_1$  はあくまでも 100,000 個をサンプリングしている事も原因と考えられる.

##### 4.5.2 [RQ2] 対角成分の影響

式 (6) を計算すると, 対角成分が非常に大きな値を持つ事が分かった. 対角成分が大きいと, グラフ上の情報の伝播において自分自身の要素が最大になる事を意味し, 接続されているグラフの多様な成分を伝播させるという目的に沿わない事が予想される. よって, RQ2 では対角成分を 0 にしてその後正規化処理をした行列を作成し, 結果がどのように変化するか調べた.

こちらも評価指標は Recall@10 で, データセットは Gowalla と Amazon-Books の密なデータセットを利用する.

対角成分を 0 にしてその分他の成分を強くしたことで, グラフ上で接続されている他の頂点から情報が伝播され良い結果に繋がったと考える事ができる.

##### 4.5.3 [RQ3] $\hat{\mathbf{A}}_Q$ のフィルタリング

式 (6) における  $Q_0$  と  $Q_1$  は従来の隣接行列と比較して 0 の成分が圧倒的に少ない. これは量子化隣接行列の特徴であり, 従来の隣接行列にはないグラフ間の情報伝播を可能にするものである. しかしながら,  $\hat{\mathbf{A}}_Q$  が密になりすぎると, 他の重要な

表 6 RQ1 の結果

	Q0	Q1	Gowalla	Amazon-Books
Ours-Model1	TRUE	TRUE	0.1491	0.0885
Ours-Model1	TRUE	FALSE	0.1641	0.0892
Ours-Model2	TRUE	TRUE	0.1697	0.0943
Ours-Model2	TRUE	FALSE	0.1705	0.0955

表 7 RQ2 の結果

	Diag set to 0	Gowalla	Amazon-Books
Ours-Model1	FALSE	0.1406	0.0776
Ours-Model1	TRUE	0.1493	0.0845
Ours-Model2	FALSE	0.1697	0.0942
Ours-Model2	TRUE	0.1713	0.0961

表 8 RQ3 の結果

	$\hat{\mathbf{A}}_Q$ Filter	Gowalla	Amazon-Books
Ours-Model1	TRUE	0.1406	0.0776
Ours-Model1	FALSE	0.1491	0.0885
Ours-Model2	TRUE	0.1697	0.0942
Ours-Model2	FALSE	0.1697	0.0943

成分の接続の強度が弱まる可能性が考えられる。

よって RQ3 として、 $\hat{\mathbf{A}}_Q$  の大きい方からフィルタリングを行い、0 の成分の割合を従来の隣接行列と同等になるように行いその影響を調べた。

結果はフィルタリングしない方が高い Recall が得られた。本来値がある成分を 0 にフィルタリングする事で、量子化隣接行列特有の成分が無効化され、結果が悪くなると考えられる。

#### 4.6 多様性, 意外性の考察

推薦システムでは、上記で議論した Recall や Precision などの評価指標の他に、多様性や意外性 (セレンディピティ) も求められる。推薦システムがユーザーが欲しい情報をタイミング良くユーザーに提供できたとしても、同じようなアイテムを推薦し続けるとユーザーに飽きが生じる。現代の推薦システムでは、式 (7) で定義されるスコアが上位のアイテムは高い Recall や Precision を持つ一方で、スコアが下位のアイテムでは高いセレンディピティを持つアイテムをユーザーに推薦することが求められている。

##### 4.6.1 定義

ユーザー  $u$  にとって、式 (7) が  $s$  番目に高いアイテムのベクトルを  $\mathbf{e}_{i,u(s)}$  と表記する。また、 $s$  番目から  $t$  番目までのアイテムのベクトルの平均を以下の様に表記する。

$$\bar{\mathbf{e}}_{i,u(s,t)} = \frac{1}{t-s+1} \sum_{r=s}^t \mathbf{e}_{i,u(r)} \quad (12)$$

本研究では、多様性を定量的に考察するために、あらかじめ各ユーザーのアイテムの順位の範囲を定め、その順位の範囲にあるアイテムのベクトルの平均を計算し、各ベクトルとその平

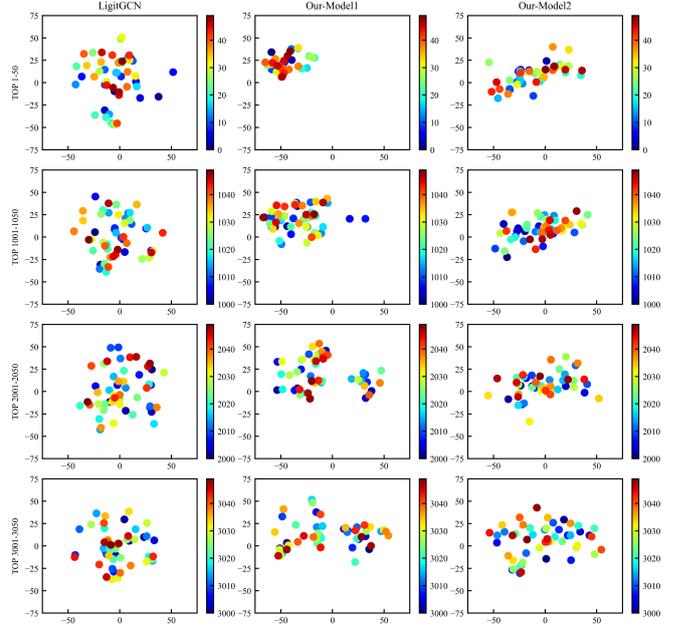


図 4 t-SNE [16] を利用して、 $\mathbf{e}_{i,u(s)}$  を 2 次元にプロットした散布図。データセットは Yelp であり、ある一人のユーザーをランダムにサンプリングした。

均のベクトルのユークリッド距離の平均を求め、さらにその平均を全ユーザーに対して計算し平均をとった  $d_i$  を以下の様に定義する。

$$d_i(s,t) = \frac{1}{m(t-s+1)} \sum_{u=1}^m \sum_{r=s}^t \|\mathbf{e}_{i,u(s)} - \bar{\mathbf{e}}_{i,u(s,t)}\| \quad (13)$$

例えば、 $d_i(25,29)$  は、あるユーザーに対するアイテムのランキング順위가 25 位から 29 位のアイテムのベクトルと、その 5 つのベクトルの平均のユークリッド距離の平均を求め、さらに、すべてのユーザーに対して同様の計算を行い、平均をとる事を意味する。

$d_i(s,t)$  が大きければ、ランキング順위가  $s$  から  $t$  までのアイテムの集合がベクトル空間上で広い範囲に存在することを示し、多様性のある推薦システムを構築できたことを意味する。

以下では主に本研究にて提案する Our-Model1 と Our-Model2 について考察を行い、LightGCN [6] を比較対象とする。

##### 4.6.2 定性的考察

最初に、広いランキング範囲を対象にした結果を図 4 に示す。図 4 は、ランダムに選んだある一人のユーザーに対して、 $1 \leq s \leq 50$ ,  $1001 \leq s \leq 1050$ ,  $2001 \leq s \leq 2050$ ,  $3001 \leq s \leq 3050$  における  $\mathbf{e}_{i,u(s)}$  をそれぞれその距離関係を可能な限り保存した状態で 2 次元にプロットした図である。それぞれのベクトルは 64 次元であるが、それを t-SNE [16] を利用して、2 次元に圧縮した。利用したデータセットは Yelp である。

横軸には、比較対象として選択した LightGCN [6] と Our-Model1, Our-Model2 の各モデルが並び、縦軸には、サンプルの順位  $s$  ( $1 \leq s \leq 50$ ,  $1001 \leq s \leq 1050$ ,  $2001 \leq s \leq 2050$ ,  $3001 \leq s \leq 3050$ ) を取り、順位に対するユークリッド距離

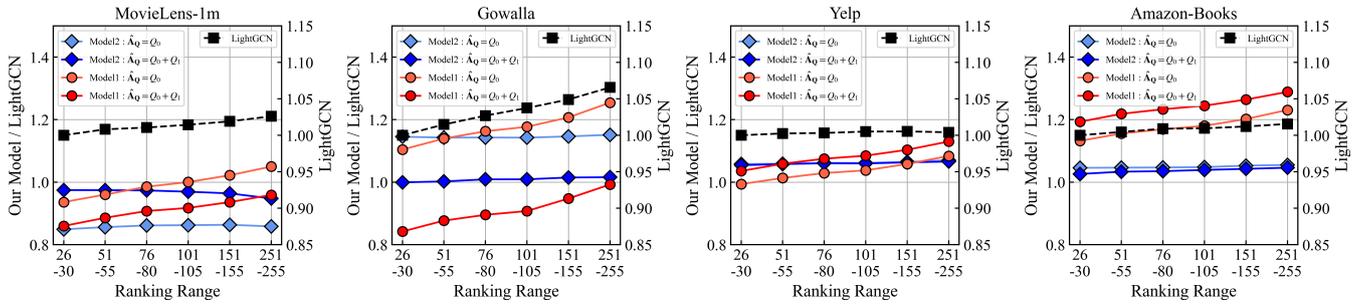


図5 ランキング順位が下位のアイテムについて  $d_i(s, t)$  の値をプロットした図。左軸が Our-Model1 と Our-Model2 の  $d_i(s, t)$  を LightGCN [6] の  $d_i(s, t)$  を基準に計算した値、右軸が LightGCN [6] の  $d_i(1, 5)$  の値を基準に計算した値。Our-Model1 では、LightGCN [6] と比較してもアイテムのベクトル空間上で広い範囲に存在していることがわかる。

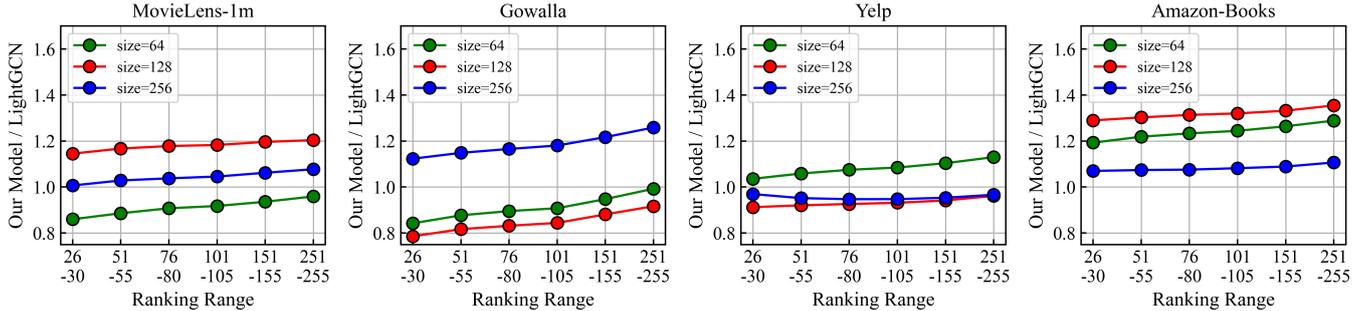


図6 Our-Model1 の隣接行列が  $\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q}_0 + \mathbf{Q}_1$  の場合の  $d_i(s, t)$  のアイテムのベクトルの次元数依存性。各プロットはベクトルの次元数が 64, 128, 256 の場合を示す。Yelp のベクトルの次元数が 256 のデータを除き、すべてランキング順位が下がるにつれて増加する。

の依存性をプロットした。Our-Model1 では、LightGCN や Our-Model2 と比較して、順位が下がるにつれてベクトル間の距離が広がっている事がわかる。

#### 4.6.3 定量的考察

次に粒度を細かくし、推薦システムにおいて、多様性を必要とする比較的低いランキング順位におけるアイテムについて定量的な考察を行う。

具体的には順位  $s$  が、 $26 \leq s \leq 30$ ,  $51 \leq s \leq 55$ ,  $76 \leq s \leq 80$ ,  $101 \leq s \leq 105$ ,  $151 \leq s \leq 155$ ,  $251 \leq s \leq 255$  のアイテムを対象とする。これらの順位について、定義である式 (13) に基づき  $d_i(26, 30)$ ,  $d_i(76, 80)$ ,  $d_i(101, 105)$ ,  $d_i(151, 155)$ ,  $d_i(251, 255)$  を計算する。

図5は LightGCN [6] と比較して、ランキング順位が下位のアイテムについて  $d_i(s, t)$  の値をプロットした図である。左軸は、Our-Model1 と Our-Model2 に対して、隣接行列が  $\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q}_0 + \mathbf{Q}_1$  と  $\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q}_0$  の  $d_i(s, t)$  を LightGCN [6] の  $d_i(s, t)$  で規格化した値を示す。右軸は、比較対象である LightGCN [6] の変化をプロットするために、LightGCN [6] の  $d_i(1, 5)$  で各  $d_i(s, t)$  を規格化した値である。右軸である LightGCN [6] の  $d_i(s, t)$  は、データセットが Gowalla の場合、ランキングの順位が下がるにつれて、大きくなっていることがわかる。また、他のデータセットについては、ほぼ一定の値となっている。また、その  $d_i(s, t)$  で規格化した Our-Model1 の  $d_i(s, t)$  についても、ランキング順位が下がるに毎に大きくなっていることがわかる。ただし、Our-Model2 については、ほぼ一定の値をと

り、 $d_i(s, t)$  に変化がないことがわかる。

図6は、Our-Model1 の隣接行列が  $\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q}_0 + \mathbf{Q}_1$  の場合の  $d_i(s, t)$  のアイテムのベクトルの次元数に対する依存性を示すグラフである。本研究で基本となる次元数は 64 であるが、その倍である 128、さらにその倍である 256 について実験を行った。データセットが Yelp の次元サイズが 256 のデータを除き、すべての次元数についてランキング順位が下がるにつれて上昇することが確認できる。

#### 4.6.4 多様性に関する結果

図5と図6の結果から、Our-Model1 においては、LightGCN [6] と比較して、ランキング順位が下がるにつれて、ベクトル空間上での距離が広くなり、多様性がある推薦システムを構築する事ができた。Our-Model2 については、ランキング順位が下がるにつれて、距離は広くなるが、LightGCN [6] と比較するとほぼ同程度であり、優位性があるとは認められなかった。

量子化隣接行列 (式 (4)) は、従来利用されてきた正規化隣接行列 (式 (1)) と比較して、二部グラフ上で直接接続されていないアイテム-ユーザーの成分を持つ。この成分が LightGCN [6] と比較して多様性や意外性などを持つ事が期待される。今回の Our-Model1 では、最適化されたアイテムはベクトル空間上で広い範囲に分布し、その予想を定量的に裏付ける結果となった。

## 5 結 論

### 5.1 本研究のまとめ

本研究の結論は以下の通りである。

(1) 推薦システムに量子ウォークに基づく量子化隣接行列を利用するモデルを二つ提案した。

(2) ランキング順位が上位のアイテムについては, LightGCN [6] と同等の Recall や Precision どの評価指標を得ることができた。

(3) ランキング順位が下位のアイテムについては, 比較対象である LightGCN [6] と比較して, 多様性のある推薦システムを構築する事ができた。

### 5.2 本研究の課題

#### 5.2.1 Our-Model1 と Our-Model2 の差異

本研究の課題の一つに, Our-Model1 と Our-Model2 の差異の考察が不十分という点があげられる。例えば, 図 5 で, Our-Model1 では, ランキング順位が下がるに応じて, ユークリッド距離も広がっているが, Our-Model2 ではそのような傾向は見られなかった。この差異については, さらに考察を深める必要がある。

#### 5.2.2 計 算 量

$Q_1$  を求める際, 式 (11) に従ってサンプリングを行うが, あくまでもサンプリングを行っているため, 量子化隣接行列を正確に計算できていないという課題がある。また, 量子化隣接行列 (式 (4)) は, 正規化隣接行列 (式 (1)) と比較して密な行列になっているため, Our-Model1 や Our-Model2 において, ニューラルネットワークの計算を行う際, 疎行列として定義することができず, 計算時に余計なメモリや計算量が必要という課題もある。

### 5.3 今後の展望

今後は量子ウォークに関するモデルの理解を深め, さらなるモデル改善に取り組み, 精度がより高く, かつ多様性や意外性のある量子モデルを利用した推薦システムの研究を進める予定である。

## 文 献

- [1] J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. *Collaborative Filtering Recommender Systems*, p. 291–324. Springer-Verlag, Berlin, Heidelberg, 2007.
- [2] Badrul Sarwar, George Karypis, Joseph Konstan, and John Riedl. Item-Based Collaborative Filtering Recommendation Algorithms. In *Proceedings of the 10th International Conference on World Wide Web (WWW 2001)*, p. 285–295.
- [3] Yehuda Koren. Factorization Meets the Neighborhood: A Multifaceted Collaborative Filtering Model. In *Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD 2008)*, p. 426–434, New York, NY, USA.
- [4] Thomas N. Kipf and Max Welling. Semi-Supervised Classification with Graph Convolutional Networks. In *Proceedings of the 5th International Conference on Learning Representations (ICLR 2017)*.
- [5] Xiang Wang, Xiangnan He, Meng Wang, Fuli Feng, and Tat-Seng Chua. Neural Graph Collaborative Filtering. In *Proceedings of the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2019)*, p. 165–174.
- [6] Xiangnan He, Kuan Deng, Xiang Wang, Yan Li, YongDong Zhang, and Meng Wang. LightGCN: Simplifying and Powering Graph Convolution Network for Recommendation. In *Proceedings of the 43rd International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2020)*, p. 639–648.
- [7] Iordanis Kerenidis and Anupam Prakash. Quantum Recommendation Systems. In Christos H. Papadimitriou, editor, *8th Innovations in Theoretical Computer Science Conference, ITCS 2017, January 9–11, 2017, Berkeley, CA, USA*, Vol. 67, pp. 49:1–49:21.
- [8] Ewin Tang. A Quantum-Inspired Classical Algorithm for Recommendation Systems. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing (STOC 2019)*, p. 217–228.
- [9] Lu Bai, Yuhang Jiao, Luca Rossi, Lixin Cui, Jian Cheng, and Edwin R. Hancock. Graph Convolutional Neural Networks based on Quantum Vertex Saliency. *arXiv preprint arXiv:1809.01090*, 2019.
- [10] Chris Godsil. Average mixing of continuous quantum walks. *Journal of Combinatorial Theory, Series A*, Vol. 120, No. 7, pp. 1649–1662, 2013.
- [11] Jin Zheng, Qing Gao, and Yanxuan Lv. Quantum Graph Convolutional Neural Networks. *arXiv preprint arXiv:2107.03257*, 2021.
- [12] B. L. Douglas and J. B. Wang. Efficient quantum circuit implementation of quantum walks. *arXiv preprint arXiv:0706.0304*, 2009.
- [13] 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 2009)*, p. 452–461.
- [14] Kalervo Järvelin and Jaana Kekäläinen. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Trans. Inf. Syst.*, Vol. 20, No. 4, p. 422–446, oct 2002.
- [15] Wayne Xin Zhao, Shanlei Mu, Yupeng Hou, Zihan Lin, Kaiyuan Li, Yushuo Chen, Yujie Lu, Hui Wang, Changxin Tian, Xingyu Pan, Yingqian Min, Zhichao Feng, Xinyan Fan, Xu Chen, Pengfei Wang, Wendi Ji, Yaliang Li, Xiaoling Wang, and Ji-Rong Wen. RecBole: Towards a Unified, Comprehensive and Efficient Framework for Recommendation Algorithms. *arXiv preprint arXiv:2011.01731*, 2020.
- [16] Laurens van der Maaten and Geoffrey Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research (JMLR)*, Vol. 9, No. 86, pp. 2579–2605, 2008.