日本データベース学会和文論文誌
一般論文 Vol. 21-J, Article No. 3, 2023年 3月

量子モデルを利用した推薦システムの提案

和山弘[1]　杉山一成[2]

クリックなど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2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ことが予想される．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ことにより精度の高い推薦システムを提案する．

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 [10]， [9]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 [7]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アイテム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 [6]，NGCF [13]，LightGCN [3]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 [3]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ことが示されている．

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． [5]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$O(\mathrm{poly}(k)\ \mathrm{polylog}(mn))$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$O(\mathrm{poly}(mn))$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ことが示された [11]．さらに，Baiら [1]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 [2]．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 [15]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 [3]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．

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

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$\mathbf{D_U} = \mathrm{Diag}(\mathbf{R} \cdot \mathbf{1})$と，各アイテムのinteraction数の合計を対角成分に持つ正方行列 $\mathbf{D_I} = \mathrm{Diag}\left(\mathbf{1}^T \cdot \mathbf{R}\right)$を利用して，正規化嗜好行列は以下のように定義される．

$$\hat{\mathbf{R}} = \mathbf{D_U}^{-\frac{1}{2}} \mathbf{R} \mathbf{D_I}^{-\frac{1}{2}}$$

[1] 学生会員　京都大学大学院情報学研究科
wayama.hiroshi.84h@st.kyoto-u.ac.jp
[2] 正会員　京都大学大学院情報学研究科
kaz.sugiyama@i.kyoto-u.ac.jp

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．

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

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 [3]では，グラフのノードが持つ情報を正規化嗜好行列に従って伝播させる．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作用させることで情報の伝播を実現している．ユーザー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$H$は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．

[2]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，[2]に従い，量子ウォークの定式化を行う．

2.3　連続時間量子ウォーク

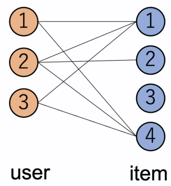
正規化隣接行列$\hat{\mathbf{A}}$をハミルトニアンHとして考え，以下の行列を定義する．

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

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

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

$\circ$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ことも可能である．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ことができる．



$$\hat{\mathbf{R}} = \begin{pmatrix} 0.41 & 0.00 & 0.00 & 0.41 \\ 0.33 & 0.58 & 0.00 & 0.33 \\ 0.41 & 0.00 & 0.00 & 0.41 \end{pmatrix}$$

$$\hat{\mathbf{A}} = \begin{pmatrix} 0.00 & 0.00 & 0.00 & 0.41 & 0.00 & 0.00 & 0.41 \\ 0.00 & 0.00 & 0.00 & 0.33 & 0.58 & 0.00 & 0.33 \\ 0.00 & 0.00 & 0.00 & 0.41 & 0.00 & 0.00 & 0.41 \\ 0.41 & 0.33 & 0.41 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.58 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 & 0.00 \\ 0.41 & 0.33 & 0.41 & 0.00 & 0.00 & 0.00 & 0.00 \end{pmatrix}$$

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．

2.3.1　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．

$$\hat{\mathbf{A}} = \sum_r \lambda_r |r\rangle\langle r| \tag{4}$$

$|r\rangle$は$\hat{\mathbf{A}}$の固有ベクトルあり，λ_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．

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

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$\hat{\mathbf{M}}$を具体的に計算することができる．本研究では，$\hat{\mathbf{M}}$の$t = 1$の行列を量子ウォークによる隣接行列と考え，$\hat{\mathbf{A}}_{\mathbf{Q}}$と表現する．

$$\hat{\mathbf{A}}_{\mathbf{Q}} = \sum_r |r\rangle\langle r|^{\circ 2} + 2\sum_{r<s} \cos(\lambda_r - \lambda_s)\,|r\rangle\langle s| \circ |r\rangle\langle s| \tag{6}$$

ここで，$|r\rangle\langle r|^{\circ 2}$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．また，式(1)を利用したグラフ上の畳み込みは高周波成分をカットするローパスフィルタの性質を持つ [14]．式(6)の第一項の最大固有値は1であり，式(1)と比較して効率的に高周波成分をカットすることが可能である．グラフ上の高周波成分は，隣接するノード間の値の差が大きいということを意味し，推薦システムにおいては悪い影響があると考えられる．このように，量子ウォークによる隣接行列(式(6))は従来の隣接行列(式(1))とは異なる性質を持ち，推薦システム全体に対しても良い影響を与えることが期待される．

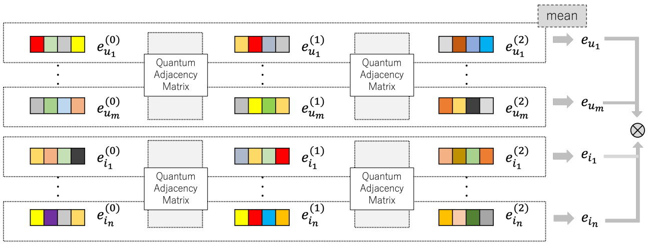


図2　レイヤー数を1にしたモデル2のアーキテクチャ．

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(6))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(6)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$\mathbf{E}^{(n+1)} = \hat{\mathbf{A}}_{\mathbf{Q}}\mathbf{E}^{(n)}$という漸化式にしがた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，

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

と表現できる．ここで，$\mathbf{e}_u^{(0)}$と$\mathbf{e}_i^{(0)}$はすべてのユーザーとすべてのアイテムのベクトルを並べた行列を表す．

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

このようにして次々と伝播される情報を最後に平均して，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．

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

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(6))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．

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

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

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．

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

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よく利用されるBPR Loss (*Bayesian Personalized Ranking loss*) [8]を利用する．ユーザー数をm，$\mathcal{N}_u$をユーザーu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．

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

λ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，[13]や[3]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，誰でも利用することが可能である．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こと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，アイテムの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．

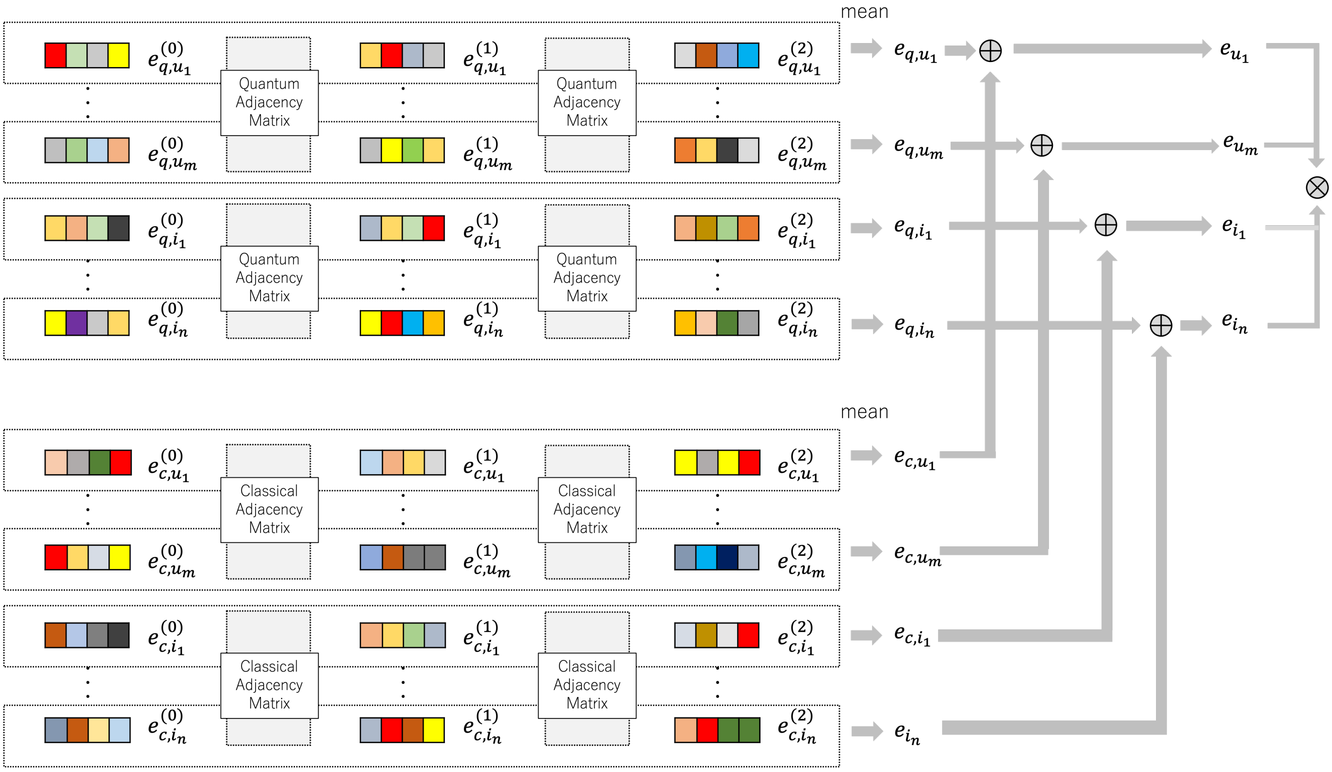


図3　レイヤー数を2にしたモデル2のアーキテクチャ.

表1　Filtering conditions for dense dataset.

Dataset	Lower	Upper
MovieLens-1M	10	∞
Gowalla	20	∞
Yelp	35	∞
Amazon-Books	32	∞

表2　Filtering conditions for sparse dataset.

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

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におよそ一桁程度の差がある.

表3　Statistis 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　Statistis 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

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) [4]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，式(6)を以下のように分解する.

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

$$\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q_0} + \mathbf{Q_1} \tag{8}$$

$$\mathbf{Q_0} = \sum_r |r\rangle\langle r|^{\circ 2} \tag{9}$$

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

4.3.1 $\mathbf{Q_0}$の計算量

$\mathbf{Q_0}$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ので，$O(N^3)$の計算量が必要になる．

4.3.2 $\mathbf{Q_1}$の計算量

$\mathbf{Q_1}$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$O(N^2)$の計算量が必要なため，合計で$O(N^4)$の計算量が必要になる．

以上の議論により，$\mathbf{Q_0} + \mathbf{Q_1}$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ことが分かった．

4.3.3 $\mathbf{Q_1}$サンプリング

$\mathbf{Q_1}$を正確に計算することが不可能なので，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．

$$\mathbf{Q_1} \sim p(r,s) = \cos^2(\lambda_r - \lambda_s) \tag{11}$$

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 [8]，NGCF [13]，LightGCN [3]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示している．

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きたこと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ことが考えられる．

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] $\mathbf{Q_0}$と$\mathbf{Q_1}$

式(8)において，第一項である$\mathbf{Q_0}$と第二項である$\mathbf{Q_1}$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$\mathbf{Q_1}$がない方が良いRecallを得ることができた．$\mathbf{Q_1}$はあくまでも100,000個をサンプリングしていること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ことが分かった．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ことを意味し，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ことが予想される．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．

表6 RQ1の結果

	Q_0	Q_1	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

	$\hat{\mathbf{A}}_\mathbf{Q}$ 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

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ことができる．

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

式(6)における$\mathbf{Q_0}$と$\mathbf{Q_1}$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，$\hat{\mathbf{A}}_\mathbf{Q}$ が密になりすぎると，他の重要な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，$\hat{\mathbf{A}}_\mathbf{Q}$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ことで，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．最新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ように表記する．

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

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に対して計算し平均をとったd_iを以下のように定義する．

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

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位が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，平均をとることを意味する．

$d_i(s,t)$が大きければ，ランキング順位が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．

以下においては，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 [3]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，$1 \le s \le 50$，$1001 \le s \le 1050$，$2001 \le s \le 2050$，$3001 \le s \le 3050$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 [12]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 [3]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（$1 \le s \le 50$，$1001 \le s \le 1050$，$2001 \le s \le 2050$，$3001 \le s \le 3050$）を取り，順位に対するユークリッド距離の依存性をプロットした．Our-Molde1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ことがわかる．

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$26 \le s \le 30$，$51 \le s \le 55$，$76 \le s \le 80$，$101 \le s \le 105$，$151 \le s \le 155$，$251 \le s \le 255$ 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．

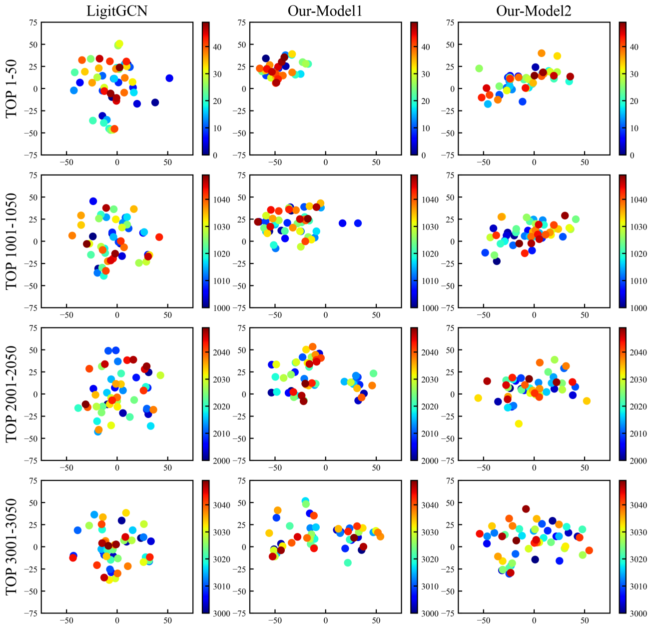


図4 t-SNE [12]を利用して、$\mathbf{e}_{i,u(s)}$を2次元にプロットした散布図．データセットはYelpであり，ある一人のユーザーをランダムにサンプリングした.

図5はLightGCN [3]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 [3]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 [3]の変化をプロットするために，LightGCN [3]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 [3]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数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 [3]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ことができた．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 [3]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(6))は，従来利用されてきた正規化隣接行列(式(1))と比較して，2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 [3]と比較して多様性や意外性などを持つことが期待される．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 [3]と同等のRecallやPrecisionなどの評価指標を得た.
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 [3]と比較して，多様性のある推薦システムを構築した.

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

$\mathbf{Q_1}$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(6))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.

5.2.3 連続時間量子ウォーク

本研究では連続時間量子ウォークの概念を取り入れた際，遷移行列$\hat{\mathbf{M}}$において，$t = 1$を仮定し隣接行列を計算した．この仮定が妥当なものなのか，もしくは，$\mathbf{Q_1}$の時間平均が0であることに着目し，式(6)の第一項である$\mathbf{Q_0}$のみを考慮すれば十分なのか考察を深める必要がある.

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，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.

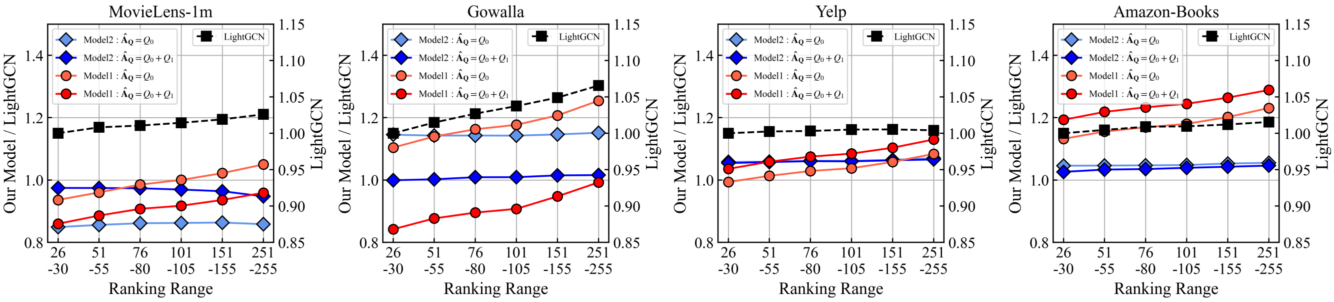


図5 ランキング順位が下位のアイテムについて$d_i(s,t)$の値をプロットした図．左軸がOur-Model1とOur-Model2の$d_i(s,t)$をLightGCN [3]の$d_i(s,t)$を基準に計算した値，右軸がLightGCN [3]の$d_i(1,5)$の値を基準に計算した値．Our-Model1では，LightGCN [3]と比較してもアイテムのベクトル空間上で広い範囲に存在していることがわかる．

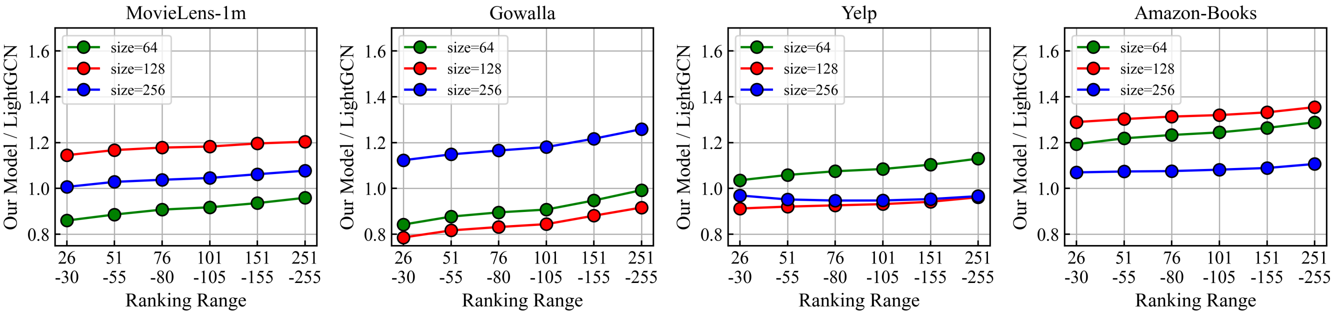


図6 Our-Model1の隣接行列が$\hat{\mathbf{A}}_{\mathbf{Q}} = \mathbf{Q_0} + \mathbf{Q_1}$の場合の$d_i(s,t)$のアイテムのベクトルの次元数依存性．各プロットはベクトルの次元数が64, 128, 256の場合を示す．Yelpのベクトルの次元数が256のデータを除き，すべてランキング順位が下がるにつれて増加する．

参考文献

[1] 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.

[2] Chris Godsil. Average mixing of continuous quantum walks. *Journal of Combinatorial Theory, Series A*, 120(7):1649–1662, 2013.

[3] 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)*, page 639–648.

[4] Kalervo Järvelin and Jaana Kekäläinen. Cumulated Gain-Based Evaluation of IR Techniques. *ACM Trans. Inf. Syst.*, 20(4):422–446, oct 2002.

[5] 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*, volume 67, pages 49:1–49:21.

[6] 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)*.

[7] 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)*, page 426–434, New York, NY, USA.

[8] 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)*, page 452–461.

[9] 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)*, page 285–295.

[10] J. Ben Schafer, Dan Frankowski, Jon Herlocker, and Shilad Sen. *Collaborative Filtering Recommender Systems*, page 291–324. Springer-Verlag, Berlin, Heidelberg, 2007.

[11] 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)*, page 217–228.

[12] Laurens van der Maaten and Geoffrey Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research (JMLR)*, 9(86):2579–2605, 2008.

[13] 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)*, page 165–174.

[14] Felix Wu, Amauri H. Souza Jr., Tianyi Zhang, Christopher Fifty, Tao Yu, and Kilian Q. Weinberger. Simplifying Graph Convolutional Networks. In Kamalika Chaudhuri and Ruslan Salakhutdinov, editors, *Proceedings of the 36th International Conference on Machine Learning (ICML 2019)*, pages 6861–6871.

[15] Jin Zheng, Qing Gao, and Yanxuan Lv. Quantum Graph Convolutional Neural Networks. In *Proceedings of the 40th Chinese Control Conference (CCC '21)*, pages 6335–6340, 2021.