|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| テーマ番号 | 1EP070 | |  | | |
| プロジェクト テーマ | 和文 | 図書推薦システムにおける書籍共起ネットワークの構築ならびに更新手法の開発 | | 指導教員 | 元木 光雄 准教授 |
| 英文 | Construction of book co-occurrence network in book recommendation system and development of update method | |
| プロジェクト メンバー | | 4EP4-24　熊ノ郷 脩 (Shu Kumanogo) | | | |

Currently, most Japanese libraries do not provide the book recommendation system. Collaborative filtering, which is the mainstream of recommendation systems, uses personal action history. However, it is difficult to use personal lending history in libraries from the viewpoint of privacy protection. Therefore, it is difficult to build a book recommendation system using collaborative filtering. Based on this background, Motoki Lab has been studying book recommendation systems based on the idea that books with the same user's loan period overlap are related. We have calculated book usage statistics and created a book co-occurrence network when books are returned. However, it does not consider the case where there are multiple same books. Furthermore, there is a problem that a book co-occurrence network including books borrowed until a certain period cannot be created. In this research, we aim to calculate book usage statistics in consideration of the case where there are multiple identical books, and to construct a book co-occurrence network including books borrowed until a certain period.

recommendation system, library loan records, co-occurrence network

# まえがき

図書館の蔵書検索システムは，検索ボックスに検索キーワードを入力すると，検索キーワードと完全一致もしくは部分一致した書籍のリストを表示する．しかし，検索キーワードを含まないが，関連のある書籍が見つからないことがある．そこで，図書館に図書推薦システムを提供することができれば，より書籍を借りやすくなることが期待される．

推薦システムの主流である協調フィルタリングは，利用者の購入履歴や閲覧履歴などに基づいて各利用者の嗜好を推測し，利用者に適した推薦を行っている．しかし，図書館に図書推薦システムを提供するにあたって，争点となるのは，利用者個人の貸返データが必要となることである．貸返データには，書籍のタイトルだけでなく，利用者のIDや身分など，個人を特定することができる情報も含まれている．そのため，個人の思想や信条を示す読書傾向の情報となり得る．日本図書館協会の「図書館の自由に関する宣言」には，「読者が何を読むかはその人のプライバシーに属することであり，図書館は，利用者の読書事実を外部に漏らさない．」とある[1]．また，貸返データを図書館に保存することを認めないという意見も一部存在し，書籍を返却したと同時に貸返データを削除しているところもある．以上の理由から，協調フィルタリングは，図書館で用いることは難しい．

元木研究室では，同一利用者の貸出期間が重なる書籍同士には関連があると考え，図書推薦システムの研究を行ってきた．具体的には，利用者の個人情報を使用せず，同一利用者が同時期に借りた書籍に関係を定義することで，図書推薦システムの開発[2][3][4]を行ってきた．これまでに開発した手法では，書籍を返却した際に、書籍の利用統計を求め、書籍共起ネットワークの作成を行った．

しかし，実際の図書館での運用を考慮すると，同じ書籍が複数存在する場合があり，従来の手法を用いて，正確な書籍の利用統計を求めることが難しい．また，従来の手法では，書籍を返却した際に書籍の利用統計の計算と書籍共起ネットワークの作成を行っている．しかし，実際の図書館での運用を考慮すると，ある期間ごとに書籍共起ネットワークを作成するほうが，現実的である．

そこで，本研究では，同じ書籍が複数存在する場合を考慮した書籍の利用統計の計算と，ある期間内で返却処理がされていない書籍を含めた図書推薦システムの開発を目的とする．

# これまでの手法の概略と問題点

これまでの手法では，実際の図書館での運用を考慮して，貸出と返却に応じた処理を行っている．具体的には，貸出と返却の2通りに場合分けをした処理を行う．

貸出処理時は，借りた日時を利用者番号と書籍番号，蔵書番号ごとに貸出中データベースに記録する．実際の図書館での貸出中データベースは，書籍の保管状況を図書館側で常に把握するためにある．

返却処理時は，利用者が返却日時まで借りている各書籍らの貸出期間と書籍同士の重なる貸出期間の合計を算出する．算出した結果をグラフデータベースに記録する．その後，貸出中データベースから返却処理をした書籍の情報を削除する．

これまでの手法の問題点として2つある．1つ目は，同じ書籍が複数存在する場合を考慮した書籍利用統計の計算が出来ていないことである．そのため，これまでの手法では，正確な書籍利用統計を求められていない．

2つ目は，ある期間ごとに書籍共起ネットワークの更新が出来ていないことである．実際の図書館での運用を考慮した際に，ある期間ごとに書籍共起ネットワークを更新することが必要ある．

# 提案手法

## 書籍の関連を正確にするための月末処理

ある期間ごとに書籍共起ネットワークの更新を行う手法として，ある期間まで借りられたままの書籍を含めた書籍利用統計の計算を行う．具体的には，ある期間内の貸返処理が終わり，その後，貸出中データベースに書籍がある場合，貸出中データベースの書籍で書籍利用統計を求める．貸出中データベースにある書籍を含めた利用統計の計算を終えると，貸出中データベースにある書籍の借りた日時をある期間に更新する．その結果，ある期間ごとに書籍共起ネットワークの更新を行うことができる．

## 同じ書籍が複数ある場合の関連度の計算方法

同じ書籍が複数ある場合を考慮した関連度の計算方法は，算出した書籍利用統計と書籍の存在期間から書籍同士のを計算する．は，

，

で求められる．ただし，は書籍の貸出期間の合計を書籍の存在期間の合計で割った値(生起頻度)，は，書籍と書籍の重なる貸出期間の合計から書籍と書籍が同時に借りられる存在期間で割った値(共起頻度)で求められる．

# 実験による評価

表1に貸返処理の例を示す．対象期間を2019/4/1 07:00 ~ 2019/4/1 12:00の5時間としている．

1. 書籍の貸返処理の例
2. Examples of book lending data

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **利用者**  **番号** | **日付** | **時刻** | **書籍名** | **蔵書**  **番号** | **貸返区分** |
| 1111111 | 2019/4/1 | 7:00 | A | 10000 | 貸出 |
| 1111111 | 2019/4/1 | 8:30 | A | 10000 | 返却 |
| 2222222 | 2019/4/1 | 9:00 | B | 20000 | 貸出 |
| 2222222 | 2019/4/1 | 10:00 | A | 10001 | 貸出 |
| 2222222 | 2019/4/1 | 11:00 | B | 20000 | 返却 |

## 書籍の関連を正確にするための月末処理

表2に，表1を用いた従来手法の書籍利用統計を示す．

1. 従来手法を用いた書籍利用統計
2. Book usage statistics using conventional method

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **書籍** | **貸出期間** | **書籍1** | **書籍2** | **貸出期間** |
| A | 2:30:00 | A | B | 1:00:00 |
| B | 2:00:00 |  |  |  |

表3に，表1を用いた今回の手法の書籍利用統計を示す．

1. 今回の手法を用いた書籍利用統計
2. Book usage statistics using this method

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **書籍** | **貸出期間** | **書籍1** | **書籍2** | **貸出期間** |
| A | 3:30:00 | A | B | 1:00:00 |
| B | 2:00:00 |  |  |  |

## 同じ書籍が複数ある場合の関連度の計算方法

従来の計算方法を用いた結果を表4に示す．

1. 従来手法を用いた計算結果
2. Calculation results using the conventional method

|  |  |  |  |
| --- | --- | --- | --- |
| **生起頻度** | | **共起頻度** | |
|  | 0.5 |  | 0.2 |
|  | 0.4 |  |  |

今回の計算方法を用いた結果を表5に示す．

1. 今回の手法を用いた計算結果
2. Calculation results using this method

|  |  |  |  |
| --- | --- | --- | --- |
| **生起頻度** | | **共起頻度** | |
|  | 0.35 |  | 0.2 |
|  | 0.4 |  |  |

# むすび

本研究では，これまでの研究手法の問題点を解決し，実際の図書館に対応した実用的な推薦システムの開発を行った．

個人情報を使用せず，書籍の利用統計と書籍の情報のみを用いて推薦システムの開発を行った．また，実際の図書館に対応できるように，同じ書籍が複数存在する場合を考慮した利用統計の計算とある期間ごとに書籍共起ネットワークの更新を行う手法を試みた．その結果，従来の手法よりも精度の高い推薦を行うことができた．

1. 社団法人日本図書館協会，「図書館の自由に関する宣言」，1979．
2. 鳥山希美 : 「貸出データに基づくレコメンドシステムの研究」，金沢工業大学 工学部 情報工学科 PDⅢプロジェクトレポート，2016．
3. 小泉良太 : 「レコメンドシステムにおける図書同士の関連性の重み付けの研究」，金沢工業大学 工学部 情報工学科 PDⅢプロジェクトレポート，2017．
4. 川内憲翔 : 「データ研磨による精度の高い図書レコメンドシステムの提案」，金沢工業大学 工学部 情報工学科 PDⅢプロジェクトレポート，2018．