

パーソナライズド空間型インタラクティブ文献推薦システムの開発

新延 空也

膨大な文献を参照、保管することや、文献を発見することは研究活動において不可欠である。重要であるが故、文献管理するためのソフトウェアが存在する。これはタグやディレクトリなどを用い、ユーザが入力したメタデータから文献を単一基準から分類するものが大半である。

Watanabe らは 2 次元空間上に文献を配置するインタフェースを持ち、ユーザの感覚に基づいて文献間の遠近を決めることで、新たに文献を追加した際に、既にある文献から分類基準を推定し、自動で空間に配置、分類する文献管理システムを提案している。この研究では、空間の近い位置にある文献集合を 1 つのクラスタと見なし、クラスタ内の文献に共通する分類基準を認識し、新しい文献がシステムに登録された際に、空間中のどの位置にユーザが配置するかを推定している。

情報過多な社会において検索が困難であることから、ショッピングサイトなどではユーザの嗜好からアイテムを推薦するシステムが開発されている。一方で学術情報の流通量は増加しており、研究者が重要な文献を発見することは困難になってきている。これに対応し、近年では論文の推薦システムは多数報告されている一方、Watanabe らの研究は各論文の重要性を計算し、推薦する機能を持たない。

本研究では、Watanabe らが提案したシステムに対して、登録した文献に関連する重要度の高い文献を推薦する機能を追加することを検討する。より具体的には、Watanabe らの研究に基づき、論文の関係をラベル付きグラフ構造としてモデル化し、各クラスタに属する論文集合を入力とし、グラフにおけるノードランキングのアルゴリズムである Personalized PageRank を推薦対象の文献集合に対して適用することで、ユーザに対して推薦する論文を決定する。

しかし、Personalized PageRank による重要度の高い文献と、想定ユーザにとって重要な文献とが合致するか不明である。そこで、データセットと提案手法に基づくシステムを構築し、実験を行った。まず、ユーザ A が管理している文献を入力用、検証用のグループに分離する。次に、Crossref の文献と、複数人のユーザの入力用グループとからデータベースを構築する。最後に、ある 1 ユーザが管理する文献クラスタを推薦システムを入力した際に、出力される文献が検証用グループとどれほど一致しているかを調べた。

本論文で示した提案手法による推薦結果として $Recall@5 = 0.21$ が得られた。この値が有用なものであるかは今後ユーザ実験を行い、正確な値を算出することで検証する必要がある。

(指導教員 伊藤 寛祥)