

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

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。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る。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では、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。

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、文献が持つ情報として、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、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うということを行なった。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、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。

(指導教員 伊藤 寛祥)