

# シンガポール国立大学・Web Information Retrieval and Natural Language Processing Group (WING) における 電子図書館、情報検索、自然言語処理に関する研究

**講演者 杉山一成氏** (シンガポール国立大学 Senior Researcher)

本講演では、シンガポール国立大学・Web Information Retrieval and Natural Language Processing Group における、電子図書館、情報検索、自然言語処理に関する主要な研究について発表する。

最初に、電子図書館に関する研究として、学術論文の推薦システムに関する研究について発表する。我々の提案する推薦システムでは、協調フィルタリングを利用して、“potential citation papers” を同定する。また、論文の異なる章はそれぞれ異なる重要度があるので、論文を効果的に表現するためには、どの章を重視すればよいのかを検討する。学術論文のデータセットを用いた実験の結果、提案手法は、最新のベースラインとなる手法と比較して、推薦精度が大きく改善されたことを示す。

次に、情報検索に関する研究、特に、携帯アプリの推薦システムに関する研究について発表する。我々の提案手法は、携帯アプリのツイッターアカウントを利用し、そのフォロワーのIDを用いる。まず、ある携帯アプリに興味があるツイッターユーザのIDを含む疑似文書を作成し、LDAを適用して潜在的なグループを作成する。次に、推薦対象ユーザをこれらの潜在的なグループに写像する。携帯アプリと潜在的なグループの関係を用いて、対象ユーザが推薦候補となる携帯アプリを好む確率を推定する。我々の手法は、ツイッターからの情報を統合することによって、cold start となる携帯アプリ推薦システムの難しさを克服することを示す。

また、自然言語処理の研究の一つとして、概要的要約に関する研究について発表する。まず、複数文書集合において、最も重要な文書を同定する。次に、その最も重要な文書における文を、その他の文書における文に割り当て、類似した文のクラスタを生成する。また、各クラスタにおける文から、単語のグラフ構造を利用し、K個の最短経路を生成する。さらに、最終的な要約の情報内容と読みやすさを最大化するために、新たに提案する整数線形計画法を利用して、すべてのクラスタから生成される最短経路の集合から、文を選択する。データセット、DUC2004、ならびに2005を用いた実験の結果、提案手法は、情報の有用性と読みやすさがあり、ベースラインと最新の抽出的要約手法を上回ることを示す。

最後に、健康に関するフォーラムの推薦システムに関する我々の最新の研究について、簡単に紹介する。

日時

2017年7月27日 (木)

16:00 - 17:00

会場

筑波大学 筑波キャンパス

春日エリア 情報メディアユニオン

3階 共同研究会議室1

**参加費無料、参加申し込み不要。**

学生、教員、一般、学内外を問わず、  
どなたでもご自由にご参加下さい。

主催: 筑波大学知的コミュニティ基盤研究センター  
<http://www.kc.tsukuba.ac.jp/index.html>

お問い合わせ:

メール [kc-office@ml.cc.tsukuba.ac.jp](mailto:kc-office@ml.cc.tsukuba.ac.jp)

電話 029-859-1524 (学内: 内線81524)

