

システム情報工学研究科修士論文概要

年 度	平成 24 年度	学位名	修士(工学)
専 攻	知能機能システム	専攻	著者氏名 鈴木 浩子
指導教員氏名 宇津呂 武仁			
論文題目 トピックモデルにおけるトピック集約手法の研究			
論文概要 <p>本論文では文書集合に対してトピックモデルを推定し、さらに推定されたトピックを集約することにより、トピック、および文書集合を俯瞰する研究を行った。はじめに、特定のキーワードで検索された文書集合に対して、トピックモデルの推定により話題分布を推定することで、文書集合の俯瞰を実現する。しかし、文書集合から直接推定されたトピックは、トピックの一つ一つが比較的専門性の高い話題の文書の集まりに対応しており、文書集合全体にわたる分野や話題の分布を俯瞰することが難しい。そこで、本論文では、収集された文書集合から直接推定されたトピックモデルに対して、Wikipedia を知識源として用いる「分野トピックモデル」を提案する。Wikipedia を知識源とする分野トピックモデルの推定時には、まず、収集された文書集合から Wikipedia エントリのタイトルが抽出される。そして、各タイトルのエントリ本文を Wikipedia から収集し、収集された Wikipedia エントリ本文の集合を対象としてトピックモデルを推定し、得られたトピックモデルを分野トピックモデルとする。この分野トピックモデルを、従来型の「Wikipedia を介さないトピックモデル」の各トピックとあわせて相補的に用いることにより、文書集合の効率的閲覧が促進される。</p> <p>本論文の枠組みにおいては、まず、文書集合に対して「Wikipedia を介さないトピックモデル」を推定する。次に、Wikipedia を知識源とする分野トピックモデルの推定を行う。その後、文書へと Wikipedia エントリを付与し、付与された Wikipedia エントリと強い関わりを持つ分野トピックモデルのトピックを分野ファセットとして文書へと付与する。そして、分野ファセットと、「Wikipedia を介さないトピックモデル」のトピックとの対応をとり、「分野トピックモデル」と「Wikipedia を介さないトピックモデル」によって、階層的に文書集合の俯瞰が実現できることを示す。</p>			
審査日	平成 25 年 1 月 30 日		
審査員	(大学名 職名)	(学位)	(氏名)
主査	筑波大学 教授	博士(工学)	宇津呂 武人
副査	筑波大学 准教授	博士(工学)	古賀 弘樹
副査	筑波大学 准教授	博士(工学)	矢野 博明