

氏名（本籍）	ふくだ さとし 福田 悟志（鳥取県）
学位の種類	博士（情報科学）
学位記番号	乙第11号
学位授与年月日	平成28年9月26日
学位授与の要件	広島市立大学大学院学則第36条第3項及び広島市立大学学位規程第3条第3項の規定による
学位論文題目	技術文書を統合した動向分析システムの自動構築
論文審査委員	主査 教授 竹澤 寿幸 副査 教授 高濱 徹行 副査 教授 日浦 慎作 副査 准教授 難波 英嗣

## 論文内容の要旨

産業と関連性が高い分野の研究者にとって、論文や特許などの技術文書を検索・分析することは、その分野の動向を知るうえで重要である。本論文では、このような作業を支援するため、論文と特許から技術動向に関する情報を自動的に抽出する手法を提案する。また、要素技術とその効果を用いた応用例を示す。

論文や特許からの技術動向の抽出にあたって、本論文では、特定分野において使用された基礎的な技術（要素技術）とそれを用いて得られた知見（効果）に着目する。要素技術とその効果の変遷を知ることは、その分野における技術動向のあらましを把握する重要な情報となる。そこで本論文では、様々な研究分野における要素技術とその効果に関する表現を自動的に抽出するための手法を提案した。

一般に、技術文書において、「を用いた」や「を具備する」といった表現の直前には、要素技術を表す用語が出現する。また、「が可能になる」や「ができる」の直前には、効果を表す表現が出現する可能性が高い。さらに、効果表現は、「精度（属性）が向上（属性値）」など、属性や属性値になりやすい用語で構成されている場合が多い。そこで本論文では、このような手掛かり語を用いた機械学習による自動解析手法を提案した。さらに、機械学習において、単独のドメインコーパスから得られる情報量には限界があることに着目し、ドメイン適応による学習量の増加を図り、さらなる解析精度の向上を試みた。そして、NTCIR-8 特許マイニングタスクで提供されたデータを用いて実験を行った結果、論文解析では再現率 0.254、精度 0.469 および特許解析では再現率 0.441、精度 0.537 が得られた。これらの結果は、特許マイニングタスクの formal run で提示されたシステムの結果よりも優れていることが分かった。

要素技術と効果に関する情報を利用したシステムとして、特定の分類体系に基づく研究領域全般を横断した学術論文の自動分類システムを構築した。すべての研究領域を網羅した自動分類は、網羅的かつ効率的な文書検索を実現する。このようなタスクは文書分類として扱われて

おり、機械学習に基づいた手法が多く提案されている。本論文では、機械学習に用いる手掛かり語として、要素技術と効果に関する表現を用いた。これらの情報は、特定分野の特徴を表す重要な手掛かりになる。なお、使用する分類体系として、すべての研究分野を網羅して構築された科学研究助成事業データベースの分類体系を用いた。そして、この分類体系における「分野・分科・細目表」を論文の分類対象とした時、実験からそれぞれ平均 0.853、0.712、0.615 の分類精度が得られた。これらの値は、要素技術とその効果に関する表現を用いない場合より高いことから、文書分類における手掛かり語として要素技術とその効果を用いることの有効性が示された。

論文や特許中に含まれる要素技術とその効果を利用したもう一つのシステムとして、技術動向分析システムを構築した。ある技術分野において、「どのような要素技術がいつ頃から使われており、どのような効果が得られているのか」という情報を網羅的に収集し整理することは、その分野の技術動向を概観するのに必要不可欠である。しかし、このような動向調査には多大な時間と労力を要する。そこで、技術文書から技術動向に関する情報を自動的に抽出し、可視化するための方法を提案した。まず、特定のキーワードにより検索された論文と特許から要素技術と効果を抽出する。次に、左から順番に、要素技術、論文または特許の著作年、効果に関する情報を列挙する。これにより、特定のキーワードを中心とした要素技術の変遷を提示することができる。そして、膨大な関連論文や特許を人手で分析することなく、技術の将来性や研究の方向性を効率的に判断できることを示した。

本論文では、日本語論文と特許を対象としたが、今後は、英語や中国語など様々な言語の論文および特許も対象とすることで、海外との技術動向を比較したシステムの構築などを検討している。また、論文や特許のような技術文書だけでなく、ニュース記事や SNS といったメディア情報や評価報告書、決算短信などの文書も分析対象とし、技術的側面と経済的・社会的側面の両方から技術を評価する動向分析システムの構築も検討している。

## 論文審査の結果の要旨

平成28年8月19日13:00から14:15まで博士学位論文発表会（公聴会）を開催した。申請者が論文内容について説明を行い、その後、論文内容に関する質疑応答および議論を行った。発表会終了後の14:15から14:30まで審査委員会を開催し、論文の可否に関する審議を行った。

産業と関連性の高い分野の研究者にとって、論文や特許などの技術文書を検索・分析することは、その分野の動向を知るうえで重要である。このような作業を支援するため、論文と特許から技術動向に関する情報を自動的に抽出する手法を提案している。要素技術と効果に関する手掛かり語を用いた機械学習を提案し、ドメイン適応により学習量の増加を図り、高い性能を達成した点が優れている。特許や論文はそれぞれの分類体系で分類されるが、時代の変化に合わせて粒度や細目が更新される。すべての研究領域を網羅した自動分類は従来から機械学習に基づく多くの手法が提案されている。機械学習の手掛かり語に要素技術と効果に関するものを追加することで従来手法より高い分類精度が得られることを示した点も優れている。これらの成果をもとに技術動向分析システムを構築している。

学位論文の概要は次の通りである。1章で研究の背景と目的を述べ、2章で関連研究を概観している。3章で特許と技術文書から要素技術と効果を自動抽出する手法を提案し、ドメイン適応の効果を実験により示している。4章で要素技術と効果を用いた学術論文の自動分類を提案し、その効果を実験により示している。これらの成果をもとに構築した技術動向分析システムを5章で述べ、6章が結論である。

本論文にまとめられている主な成果は、国内ジャーナル論文（フルペーパー）2編として公表されている。関連する内容が2編の査読付き国際会議等で発表されている。

発表会（公聴会）では、申請者から博士論文研究の内容が的確に説明され、質疑応答も適切であった。聴講者や審査委員との間では、申請者が構築したCiNii Miningシステムも含めた議論が活発に行われた。