

平成 28 年 11 月 29 日

豊橋技術科学大学長 殿

情報・知能工学 専攻
学位審査委員会
委員長 増山 繁

論文審査及び最終試験の結果報告

このことについて、学位審査会を実施し、下記の結果を得ましたので報告いたします。

学位申請者	Md Zia Ullah		学籍番号	第 115309 号
申請学位	博士（工学）	専攻名	大学院工学研究科博士後期課程 情報・知能工学専攻	
博士学位論文名	Bipartite Graph Based Ranking Methods for Subtopic Mining and Genetic Disease Prediction (サブトピック・マイニングと遺伝病予測のための二部グラフに基づくランキング手法)			
論文審査の期間	平成28年 10 月 27 日 ~ 平成 28 年 11 月 30 日			
公開審査会の日	平成 28 年 11 月 18 日	最終試験の実施日	平成 28 年 11 月 18 日	
論文審査の結果*	合格		最終試験の結果*	合格
審査委員会(学位規程第6条)				
学位申請者にかかる博士学位論文について、論文審査、公開審査会及び最終試験を行い、別紙論文内容の要旨及び審査結果の要旨のとおり確認したので、学位審査委員会に報告します。				
委員長	増山 繁			
委員	梅村 恭司		青野 雅樹	
		印		印
		印		印

※論文審査の結果及び最終試験の結果は「合格」又は「不合格」の評語で記入すること。

論文内容の要旨

本論文は、二部グラフ (bipartite graph) というデータ表現を用いて、2つの実用的で挑戦的な課題に取り組んだ手法を論じたものである。ひとつは、情報検索で必須な検索質問 (クエリ) の意図を知的に汲み取る手法である。すなわち、情報検索においては、クエリから一義的に検索意図を解釈することは一般的には困難であり、多義的な解釈がされる場合が多い。そこで、多義性 (サブトピック性) に応じて、検索意図をマイニングし、それを検索結果のランキングに反映する手法が要求される。このため、検索候補となるサブトピックと素性 (features) の間に二部グラフを構成することで、サブトピックの検索におけるランキングが向上できる仕組みを述べている。もうひとつは、インターネット上に公開されている遺伝子情報と、患者に観察される症状、ならびに病気の3つ組の情報を基に、遺伝子 (gene) と病気 (disease) で二部グラフを作成し、症状 (phenotype) と遺伝子 (gene) で別の二部グラフを作成しておき、これら2セットの二部グラフに重みや情報伝播を適切に定義することで、症状から (遺伝子因子で決まる) 病気を予測する手法である。この他、上記2つの手法において、ランキング手法だけでなく各種の新規な素性を提案している。

論文の構成は第1章で、対象とする問題の全体像を述べ、第2章で関連研究を述べ、第3章では二部グラフを始め、論文で用いている数学的な道具の定義を行っている。第4章でサブトピック・マイニングを述べ、同時に、検索意図が多義にわたる場合の多様性 (diversification) を創出するために必要な素性の提案を行っている。第5章では、遺伝子情報を介して、症状からの遺伝的要因で決まる病気の予測手法を論じている。第6章では、本論文を総括し、今後の展望について述べている。

審査結果の要旨

二部グラフ表現は、リンクのある大規模データ処理に適したデータ表現として、Webマイニング、特に情報検索において、Webサイトの品質を数値化する手法として用いられてきた。本論文では、二部グラフを、情報検索分野での研究の最前線テーマのひとつである、ビッグデータを仮定した情報検索のフロントエンドである検索意図マイニング (intent mining)、ならびに検索結果の多様化 (diversification) に適用する画期的な手法を提案している。すなわち、一般に曖昧な検索質問 (クエリ) が与えられ、それに関わる複数の関連するサブトピックとの関連度を数値化するために二部グラフを導入し、さらに、「情報多様化のための新規度 (novelty)」という新たな尺度を導入することで、検索結果の多様化に成功している。その結果は、国際コンテストである NTCIR-10 (INTENT-2 タスク) ならびに NTCIR-12 (IMINE-2 タスク) の2セットのデータでの実験ですべての性能尺度において参加したチームのどの手法よりも高精度な多様化性能を発揮することが確認されている。また、この二部グラフを巧妙に活用して、病気と遺伝子の間の関係を二部グラフで表現し、一方、症状と遺伝子の間の関係を二部グラフで表現し、これら2セットの二部グラフで、相関の度合いをグラフのエッジの重みに用い、遺伝子情報部分を経由することで、遺伝的要因に起因する病気の予測を、症状だけから推定する画期的な手法を提案し、従来手法である Phenomizer システムに比べて予測精度が大幅に向上することを示している。このように、情報検索ならびにデータマイニングの次世代技術を開拓したことと、これらの分野の発展に大きく貢献した点が高く評価できる。

以上により、本論文は博士 (工学) の学位論文に相当するものと判定した。

(各要旨は1ページ以上可)