

信頼される AI の基盤技術  
2021 年度採択研究者

2021 年度 年次報告書
------------------

佐々木勇和

大阪大学 大学院情報科学研究科  
助教

グラフデータの説明可能なバイアスに関する基盤技術の創出

## § 1. 研究成果の概要

本研究はグラフデータに含まれる差別的なバイアスを削除することで、公平な機械学習モデルの構築を目指す。2021年度は、グラフデータ内に含まれるバイアスの定式化を目標とし、経路に基づく相関ルールマイニングのコンセプトを具体化した。さらに、経路探索の高速化のための索引技術の開発、および知識グラフ内のバイアス調査を実施した。

まず、経路に基づく相関ルールマイニングにおいては、(1) 経路問合せ技術、および(2) グラフ相関ルールマイニングの設計を解決する必要がある。本研究では、経路問合せにおいて、一般的に利用される結合操作、クリーネ閉包、および属性値指定の3つの操作を対象とした。条件に適合する経路のうち、始点を共有する経路を共起するパターンとして、相関ルールマイニングで一般的である支持度、信頼度、およびリフトを指標として計算可能とした。提案技術により、知識グラフにおける人と出生地の偏り、ソーシャルネットワークにおける人間関係の偏りを発見できる。さらに、高速化のために支持度以下の出現頻度の経路パターンを効率的に枝刈りするアルゴリズムおよび並列処理アルゴリズムを開発した。枝刈り無しアルゴリズムと比較して、少なくとも6倍程度の高速化を実現した。

次に、経路問合せのひとつである連言経路問合せ(Conjunctive path query)を対象とした索引技術を開発した。この索引は連言経路問合せに特化させ、高速化を実現した。この索引技術により、連言経路問合せを最大で数千倍の高速化を達成した。

最後に、知識グラフを対象としたバイアスの調査を実施した。今回は、YAGO, Nell, DBpedia, および Wikidata という代表的な4つの知識グラフを対象とし、人物に対する内容が最も充実している Wikidata を詳細に調査した。例えば、性別、国籍、宗教などのセンシティブな属性と職業の共起数を計測し、大きなバイアスがあることを発見した。