

主専攻実験 A 最終レポート

岡部 純弥

2022 年 7 月 22 日

概要

本課題では、Google の検索アルゴリズムとして非常に有名な PageRank アルゴリズム [1] の理論を理解し、これを用いた計算機実験を行った。実際に、航空ネットワークに関するダミーデータを作成し、PageRank アルゴリズムを用いて、各空港の重要度を計算した。さらに計算結果から、航空ネットワークが中央集権的である^{*1}べきか否かを考察した。

1 はじめに

1.1 PageRank とは

PageRank とは、Google の検索システムで用いられているアルゴリズムである。PageRank では、Web ページ間のハイパーリンク関係を用いて、各ページの重要度を計算する。これは、**良い Web ページは別の良い Web ページからリンクされている**という考え方をもとに実現されている。Facts about Google and Competition によると、

PageRank works by counting the number and quality of links to a page to determine a rough estimate of how important the website is. The underlying assumption is that more important websites are likely to receive more links from other websites

と確かに記載されている。またこの考え方は、論文

^{*1} 一般的に、特定のいくつかの空港に路線が集中するようなネットワークをハブアンドスポークシステムという。一方で非中央集権的な、各空港間に直行便が運行しているシステムをポイントトゥポイントシステムという。

の引用/被引用数ネットワークや共著ネットワークと非常に似ている。つまり良質な論文は、別な良質な論文からリンクされているという考え方である。実際にこの分野に関する研究として、Ma et al.[2], Ding et al.[3] などが挙げられる。

1.2 応用先

PageRank は、Web サイトの重要度付けの他にも、(ソーシャル) ネットワーク分析、物理学、化学、生物学など多数の応用先がある。ソーシャルネットワーク分析の事例としては Bahmani et al.[4] などが挙げられる。また、PageRank の応用に関する総説論文としては Gleich[5], Berkhin[6] が著名である。

1.3 アルゴリズム

ここでは、基本的な PageRank のアルゴリズムを紹介する。

PageRank のより詳細な理論に関しては、Page et al.[1], Bianchini et al.[7], Langville, Meyer[8] などを参照されたい。

1.4 変数の定義

1.5 最適化問題の定式化

2 計算機実験

3 考察

4 まとめ

参考文献

- [1] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. The pagerank citation ranking: Bringing order to the web. Technical Report 1999-66, Stanford InfoLab, November 1999. Previous number = SIDL-

WP-1999-0120.

- [2] Nan Ma, Jiancheng Guan, and Yi Zhao. Bringing pagerank to the citation analysis. *Information Processing & Management*, 44(2):800–810, 2008.
- [3] Ying Ding, Erjia Yan, Arthur Frazho, and James Caverlee. Pagerank for ranking authors in co-citation networks. *Journal of the American Society for Information Science and Technology*, 60(11):2229–2243, 2009.
- [4] Bahman Bahmani, Abdur Chowdhury, and Ashish Goel. Fast incremental and personalized pagerank. *arXiv preprint arXiv:1006.2880*, 2010.
- [5] David F Gleich. Pagerank beyond the web. *siam REVIEW*, 57(3):321–363, 2015.
- [6] Pavel Berkhin. A survey on pagerank computing. *Internet mathematics*, 2(1):73–120, 2005.
- [7] Monica Bianchini, Marco Gori, and Franco Scarselli. Inside pagerank. *ACM Transactions on Internet Technology (TOIT)*, 5(1):92–128, 2005.
- [8] Amy N Langville and Carl D Meyer. Deeper inside pagerank. *Internet Mathematics*, 1(3):335–380, 2004.