

## 文献における参照関係の可視化

北川 晴香<sup>†</sup> 宮村 (中村) 浩子<sup>††</sup> 古谷 雅理<sup>†††</sup> 斎藤 隆文<sup>††</sup>

東京農工大学 工学部情報コミュニケーション工学科<sup>†</sup>

東京農工大学 大学院生物システム応用科学府<sup>††</sup>

東京農工大学 大学院工学府<sup>†††</sup>

### 1. はじめに

ある分野の研究を調査するうえで、その分野の文献の参照関係を調べることは重要である。ある文献が参照している文献を順に辿っていくことで多くの研究の基礎となる文献を知ることができ、その分野で注目度の高い文献を知ることができる。

これまででは文献から必要な文献を探す際には Web 上で検索エンジンを用いたり、文献の参照関係から調べたりしていた。しかし、膨大な数の文献の中から必要な文献を探し出すことは困難である。そこで、効率的に検索するための方法として、文献情報によって選択した文献データの中で多くの文献に参照されている文献や、多くの文献を参照している文献などの代表的な文献に注目することが有効であると考えられる。

そこで本研究では多数の文献の参照関係を高速かつ手軽に、そして整然と描画する可視化手法を提案する。また、参照関係に着目したフィルタリングを用い、ある文献と近い関係にある文献のみを提示する選択的表示を提案する。この手法によって必要な文献の参照関係に着目した情報を対話的に表示し、観察することが可能になる。本提案手法を用いることで、ある分野の研究をこれから始める人や、新しく注目する文献を必要とする人など、ユーザの目的に応じた文献の検索が可能となる。

## 2. 文献の参照関係の可視化

### 2.1 データ形式

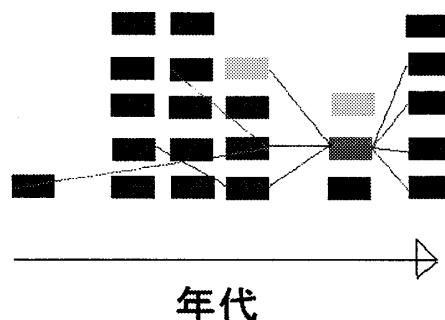
各文献について、文献 ID、それが参照している文献の ID、それを参照している文献の ID、文献名、著者名、掲載雑誌名または学会名、発表

年、文献の詳細情報の URL、所属団体のドメイン名、概要を文献データとして使用する。今回は Google Scholar [1] の検索サービスを利用して実験用の文献データを作成した。

### 2.2 可視化

可視化は図 1 のように行う。ノードは文献を表している。各ノードの色はその文献が参照された回数を擬似カラーで示しており、赤に近いほど参照された回数が多く、青に近い色ほど参照された回数が少ないことを表している (図 2)。ノードは左から年代順に配置する。データの中には発表年が明記されていないものもあるが、参照されている文献の年代から年代を推定し、おおまかに年代順に並ぶようにする。年情報をもとにして配置しているため、リンクは基本的に右側にある文献が左側にある文献を参照していることになるが、中には年代の古いものが新しいものを参照していたり、同じ年内で参照していたりするものもある。そのようなリンクは赤色で表示することによって特異性を発見しやすくする。また、同じ年内での参照関係はノードに対するリンクの表示方法によって読み取ることができる。参照している文献のリンクはノードの右側に、参照されている文献のリンクはノードの左側に接続されている。なお、画面の簡略化のために参照関係以外の情報は省略する。

図 1 可視化概要



Visualization of reference relations in documents

Haruka KITAGAWA<sup>†</sup>, Hiroko Nakamura MIYAMURA<sup>††</sup>, Tadasuke HURUYA<sup>†††</sup>, Takahumi SAITO<sup>††</sup>

<sup>†</sup>Department of Computer, Information and Communication Sciences, Tokyo University of Agriculture and Technology

<sup>††</sup>Graduate School of Bio-Applications and Systems Engineering, Tokyo University of Agriculture and Technology

<sup>†††</sup>Graduate School of Engineering, Tokyo University of Agriculture and Technology

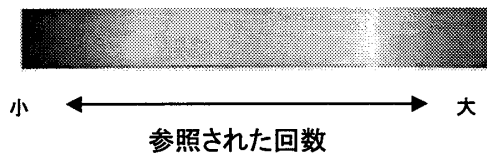


図2 擬似カラー表示

### 2.3 可視化結果

図3は1990年に発表された文献A(青丸で示す)に関連した53の文献に対して可視化を適用した結果である。データを取得する際には1990年から1999年間で参照された回数の多いもののみを選択した。この図において赤色のノードは6個あるため、それらのノードが重要であると考えられる。ノードが赤いにも関わらずリンクが繋がっていないものはデータを取得しなかったものであると考えられる。そのためその文献について知りたいときにはさらにその文献から繋がるデータを取得することによってより詳しい可視化結果を得ることができる。また、各文献についての詳細情報が知りたいときには、その文献をマウスで選択することによって文献名、ID、発表年、仮に設定した発表年が画面下部にテキストで表示され、それ以外の情報については別のウィンドウに表示した。

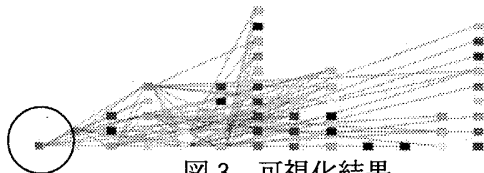


図3 可視化結果

文献の参照関係の詳細を知る際には、注目している2001年に発表された文献B(赤丸のついている文献)を選択することによってそれに繋がっているリンクとその先のノードを表示した(図4)。ここでリンクが繋がっていないノードについては、全体の形状を維持しながら必要な情報を強調したいので明度を下げて表示することとした。この手法によって接続関係と個々の文献の情報を同時に見せ、双方の情報を選択して表示できる。

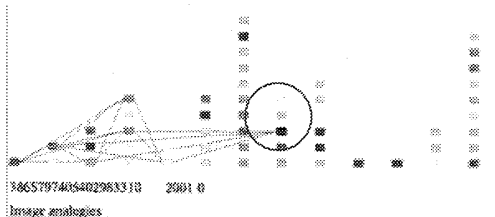


図4 参照関係の詳細の表示

この図から文献Aと文献Bは直接の参照関係はないが、2つの文献は共通して3つの同じ文献からのリンクが繋がっているため、関連性の強い研究であることが読み取れる。

図5は図3において文献数を1000程度まで増やしたものである。全体的にノードとリンクの数が増えた分見た目が乱雑になってしまっている。しかし、個々の文献の情報(ノードの色)から文献を選択して、その文献の接続情報を提示することで、目的の文献への参照関係を把握することができる。

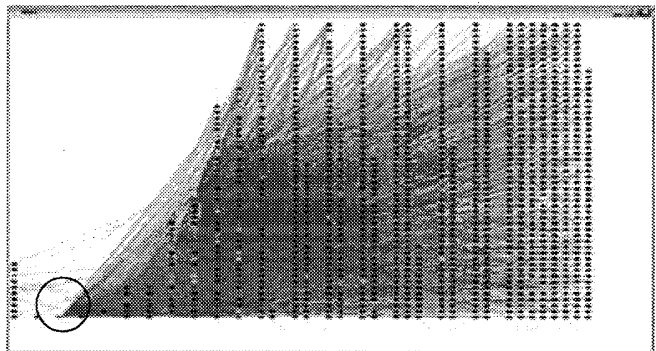


図5 多くの文献における可視化

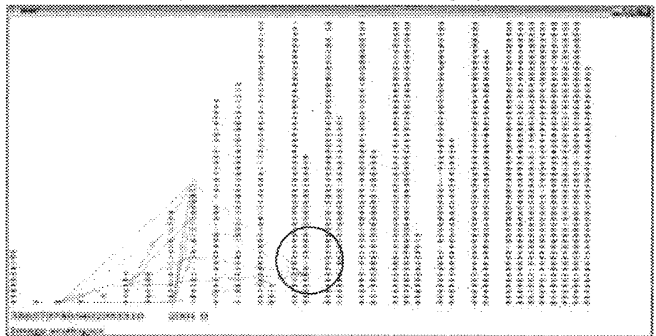


図6 図5における参照関係の詳細の表示

### 3. おわりに

本稿では文献の参照関係の可視化手法を提案した。この可視化手法によってある文献から直接の参照関係はなくとも、関連性が深いと考えられる文献を発見することが可能である。

今後の課題として、現在では雑多になってしまっている全体の概観をより見やすくするために、リンク情報の集約や、文献の内容の関連性を考慮した可視化手法を行う。また、インタフェースや機能面での充実化により、より必要な情報を取得しやすいものを目指す。

### 参考文献

- [1] Google Scholar, <http://scholar.google.com/>