

第6回 SPARC Japan セミナー 2008

「IF を越えて—さらなる研究評価の在り方を考える」

アイゲンファクター (Eigenfactor) : 学術情報のランキングおよびマッピング

ジェヴィン・ウェスト (米ワシントン大学 生物学部)

要旨

これまで数十年にわたり、引用数とインパクトファクター (Impact Factor) が学術雑誌を評価する主要な指標でした。これらの指標は理解しやすいという利点があるものの、引用のネットワークが本来持っている有益な情報の多くが捨てられています。そこでわれわれは、アイゲンファクターというアルゴリズムを提案しました (<http://www.eigenfactor.org/methods.htm>)。この指標は、その雑誌が何回引用されているかだけでなく、どこで引用されているかを考慮に入れています。このことはグーグルがウェブサイトをランキングするのに類似していますが、この指標はウェブサイトの代わりに雑誌を、ハイパーリンクの代わりに引用数をランキングしています。さらに書誌学データにこのアプローチを用いると、学術コミュニケーションについて時間を追ってマッピングすることが可能になります。これはある雑誌を科学界全体の中で位置づけるツールとして役立つ可能性があります。もちろんわれわれは、各論文を実際に読むことが、いろいろなデータや統計に優ることを理解しています。しかしながら、時間と予算の制約がますます厳しくなるにつれて、図書館員、雑誌出版者、編集者、および研究者にとって、学術雑誌に対する定量的な評価指標の必要性が当然ながら高まりつつあります。このような背景において、アイゲンファクターは正しい方向への第一歩であると考えています。



Jevin West (ジェヴィン・ウェスト)

米ワシントン大学、生物学部

ジェヴィン・ウェスト氏は、ワシントン大学生物学部の ARCS (Achievement Rewards for College Scientists) フェローです。主な研究は、数理生物学です。ウェスト氏は、ネットワーク理論、およびその理論を生物学における広範囲な問題 (抗生物質に対する抵抗や病気の進行など) に適用することに特に興味を持っています。ウェスト氏の生物学に対する定量的なアプローチは、氏を計量書誌学的な分野など他の分野のさまざまな研究にも導きました。

ウェスト氏は、eigenfactor.org を開発した代表メンバーです。eigenfactor.org は、図書館員、雑誌編集者、研究者、出版者に対して、増え続ける科学文献を評価するためのツールやリソースを無料で提供するウェブサイトです。ウェスト氏は、これまでアイゲンファクターについてハーバード大学、スタンフォード大学、ESF (欧州科学財団)、SLA (米国専門図書館協会)、ALA (米国図書館協会)、CSE (科学編集者評議会)、NAS (全米科学アカデミー) などで講演を行なってきました。

はじめに

学術雑誌の評価は、1927年にカリフォルニア州南部の小さな大学であるポモナカレッジで始まりました。当時の図書館員は、今では更に深刻になっていますが、文献が増え続けるのに予算が削減されるといふ難しい課題に直面していました。このような状況から、「図書館員はどの雑誌を購入すべきか」という問題に注目が集まりました。

1927年にポモナカレッジの2人の化学者が、引用数を計算するという当時では新しい手法を見出しました。しかし、単純に引用数を計算することが雑誌のランキングを行う最良の方法とは限りません。そこで、1955年にユージン・ガーフィールド氏が、購入すべき雑誌を決定するためのより良い手法であるインパクトファクターを考案しました。現在はこのインパクトファクターが、雇用の決定、雑誌購読の決定、昇進と終身在職権審査、広告の出稿先、英国の研究評価法に基づく研究費補助の決定、および大学のランキングなど、学術的な分野における各種の意思決定に活用されています。

結果として、当初は科学者のために考案されたインパクトファクターは、科学者だけでなく、雑誌の編集者や出版者、および大学や研究機関の管理者が偏重するツールになりました。言い換えると、犬が尻尾を振り回すのではなく、尻尾に犬が振り回されるような事態になりました。この結果、最近はこの問題の是正に関するいくつかの論文が出版されています。われわれを研究に駆り立てる課題は、「どうすれば学術文献をより良く評価できるか」です。

文献評価の新しい方法の研究

文献評価のために最も良い方法はすべての文献を読むことですが、そのための時間がある人はまずいません。そこで、時間や経費を削減するだけでなく、インパクトファクターの乱用を抑えるためにも、新しい指標が必要とされています。

われわれは、「分厚い科学論文の価値がどれだけあるのか」といった質問に答えたいと考えています。「生物学者はどのくらいの頻度で経済学の

論文を引用するか」、「最も価値が高いのはどの出版者のパッケージか」といった質問にも答えることができたらと思います。最後の質問は、図書館員がどの雑誌を購入すべきかの意思決定で直面する問題です。

雑誌評価の新しい指標を開発するにあたって、われわれは次のような基準を設定しました：

- その計量書誌学的指標は無料で入手出来なければなりません。また、オープンかつ自由にコピー出来なければなりません。
- それは、昨日までは存在せず、今日出現したものを含めて、多くのデータベースを統合出来なければなりません。
- その指標は分野を越えても通用しなければなりません。インパクトファクターの弱点の一つは、分野が異なると引用文化が違うために通用しないという点です。例えば、数学の論文における引用数が通常 20 以下であるのに対して、バイオケミストリーでは 60 もあります。
- その指標は不正行為を防止できるしっかりしたものでなければなりません。
- その指標は、情報理論とネットワーク科学の最新のアルゴリズムと方法を用いて、最大限の情報を提供しなければなりません。

われわれはアイゲンファクターがこれらの基準を満たす第一歩にあると考えています。

図1は、インパクトファクターとアイゲンファクターの主な相違点を示したものです。

Information Comparison

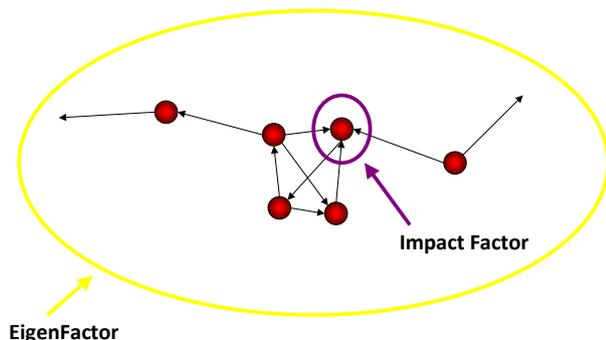


図1: インパクトファクターと
アイゲンファクターの相違点

100年近くの間、われわれはインパクトファクターの基礎となっている引用数を数えてきました。これに対して、アイゲンファクターはネットワーク全体を考慮し、ネットワークの構造に着目しています。こうして、この指標は間接的な作用を取り込んでいます。言い換えると、インパクトファクターは単に入ってくる矢印だけを考慮していて、その矢印がどこから来ているかに留意しないため、ネットワーク内の重要な情報を無視しています。実際、アイゲンファクターがネットワーク全体を考慮しているのに対して、インパクトファクターはネットワークの小さな部分だけに着目しています。

アイゲンファクターは何を表しているのでしょうか？

アイゲンファクターのアルゴリズムの背景には、ボナチッチ氏が1972年に提唱した、固有ベクトル中心性(Eigenvector centrality)という概念があります。固有ベクトル中心性の応用例は、グーグルのページランクのアルゴリズムです。アイゲンファクターとページランクのアルゴリズムは非常に良く似ていますが、アイゲンファクターの場合は、ウェブサイトのランキングの代わりに雑誌をランクづけ、ウェブサイト間のハイパーリンクの代わりに雑誌による引用数を用いています。アイゲンファクターは実際に何を表している

のでしょうか。例えば、研究者が図書館に入室し、ランダムに雑誌を選んでからランダムに引用を選び、そこから次の雑誌にたどり着くというように、このプロセスを次々に繰り返すと、研究者はそれを無限に続ける可能性があります。このプロセスは数学的にモデル化することができ、それが、基本的にアイゲンファクターが行っていることです。

このプロセスを無限に繰り返した後、「研究者はどこまで長い時間を費やしたか」という疑問が生じます。これを評価することで、アイゲンファクターは全体としての価値を計量しています。図書館員は通常この価値に特に興味を持っています。なぜなら図書館員は、研究者が図書館内をランダムに歩き回ることをのみを許された場合に、特定の雑誌にどれだけ時間を費やすかという点に関心があるからです。

一方、研究者は特定の雑誌の1論文あたりの価値に興味を持っているので、われわれはアイゲンファクターに加えてその他の指標も提供しています。原則的に、アイゲンファクターを論文の数で割ったものが論文影響値(Article Influence)です。論文影響値はインパクトファクターとの比較がより容易な指標です。しかしながら、評価する分野によっては二つの指標に相違が生じます。例えば、数学や経済学などの分野では、論文影響値の方がインパクトファクターよりも大きくなります。一方、医学などの分野では2つの指標間には高い相関があります。

アイゲンファクターの特徴

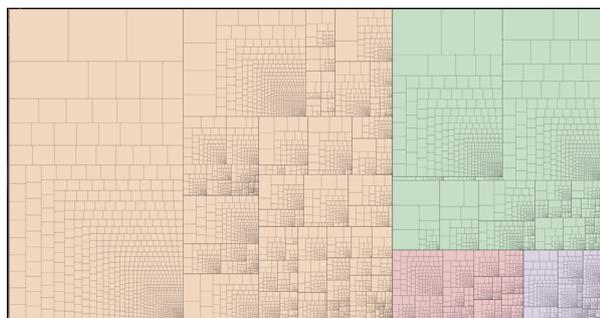
アイゲンファクター指標には次のような特徴があります。優れた雑誌からの引用は価値が高く、この点はグーグルがウェブページをランク付けするのと似ているのですが、レビュー誌以外からの引用はさらに価値が高くなります。数学や経済学のような「質素な」分野からの引用は、インパクトファクターでは低く評価されますが、アイゲンファクターでは比較的高く評価されます。

この指標を用いると、従来評価されなかった雑誌の評価が可能になります。われわれはこれを灰

色文献 (grey literature) と呼びます。従来はランキングされなかった学術雑誌が 10 万以上存在していましたが、アイゲンファクターを用いるとこれらもランキングすることができます。

アイゲンファクターには加法性があり、これは非常に重要な特徴です。このため、この指標の利点として、各雑誌のアイゲンファクターの値を足すだけでその雑誌のグループ全体のおよその価値が求められます。下図中の各四角は、それぞれ特定の雑誌を表します。図の下方には経済学分野があり、一番上には、予想通り、Nature や Science などの雑誌が位置しています。

Eigenfactor is additive



アイゲンファクターに加法性があるため、このコスト効率分析ツールを用いて雑誌のパッケージを評価することができます。図書館員たちは、引用に基づいた価値だけでなく、コストに基づいた価値にも興味を持っています。この点を評価するために、われわれはアイゲンファクターをコストで割っています。

われわれのウェブサイトでは、研究者は分野または個別の雑誌名をクリックすることで検索できて、データの動向、価格情報、収益状況、その雑誌がどのグループに属するか、われわれの指標に基づいた科学全体に占める割合などの様々な情報を入手することが出来ます。

また、論文影響値とアイゲンファクター値という 2 つの異なる評価基準があり、雑誌の価値に関する質問の内容に応じて、どちらかを用いて検索することができます。雑誌の総合的な価値を求めたいときは、アイゲンファクターを、また、論文ごとの価値を求めたいときは、論文影響値を用い

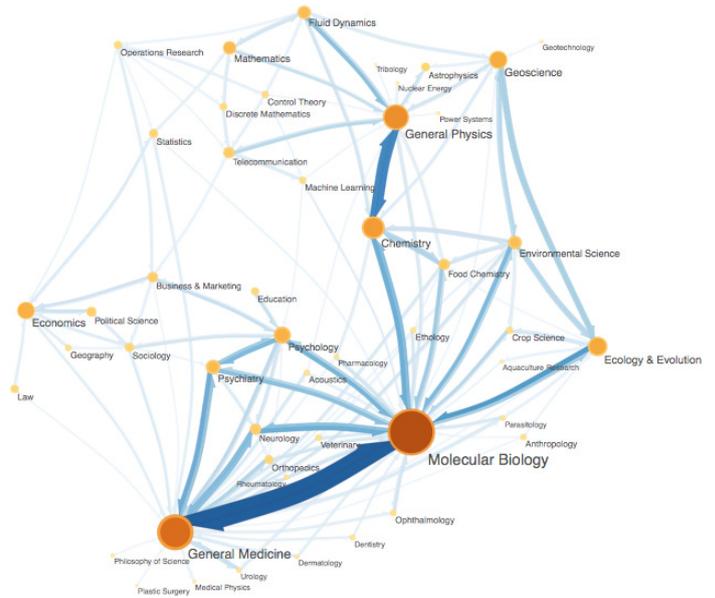
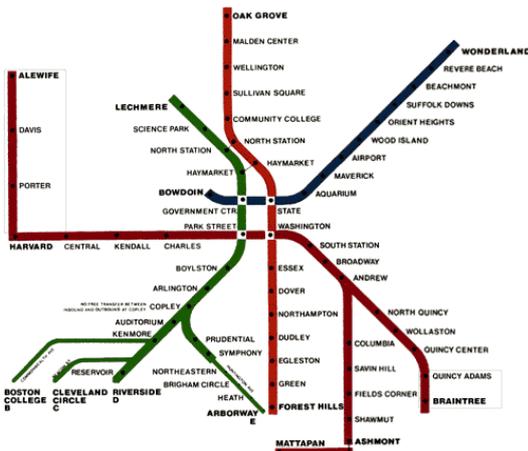
ます。たとえばコスト効率に関しては、ある雑誌のグループがもたらすと思われるアイゲンファクター値の合計やそれを購入するために必要なコストを示すツールがあります。これらのツールは図書館員たちの間で大いに普及しつつあります。

別のツールとして役立つのが、Eigenfactor.org のバブルチャートです。バブルチャートは、たとえば雑誌のブランド力がとても重要になってくることを示しています。例として何社かの有名な雑誌社がある分野の雑誌の買収に着手しました。

(ウェブのバブルチャートをご覧ください)。雑誌の中には、論文の数を増やすことなくその価値を高めているものもあり、それは編集者や出版者が望むことです。一般的に、ある雑誌がその価値を高めるためには、論文の数を増やします。しかし、論文数を増やすことなく価値を高めているとしたら非常に成功しているといえます。以上、説明したツールや情報はわれわれのウェブサイトでご覧可能です。

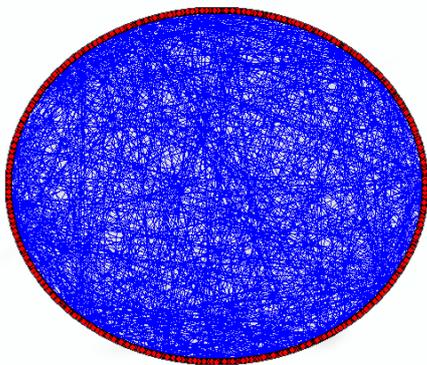
マッピングの利用

われわれのもう一つの研究は科学のマッピングに関するもので、過去 100 年間に科学で生じたことをマップ化するためにネットワーク科学のツールを用いることに興味を持っています。生物学者は酵母のネットワークのようなネットワークに大いに馴染みがあります。実際、生物学者はそのようなデータを溢れるほど持っており、この事情は計量書誌学分野の文献と類似しています。しかしながら、われわれはモデルあるいは写真をマップに変換することができます。たとえば、グーグルで表示されるボストンの街のグラフィックは詳しく過ぎるために、ボストン空港からハーバード大学までの経路を探すには不向きです。地下鉄マップのように情報を圧縮したマップ情報の方がこの目的に適しています。



写真の代わりにマップを作成すると同様に、われわれは引用ネットワークについてもマップを作成しました。

Citation Network



$$L(M) = q \circ H(Q) + \sum_{i=1}^m p_i \circ H(P^i).$$

われわれは、情報理論のアルゴリズムを用いて、初めて科学のマップを作成しました。

このマップは情報理論的アルゴリズムを用いて作成されたもので、円の大きさは、例えば科学者が分子生物学の分野でランダムに歩き回って文献を探した場合に費やす時間を表します。エッジは、一般医学から分子生物学への引用の数を表します。円の境界線は特定の分野から外へ出た引用の数を表します。通常、引用は各分野内に留まりますが、時としてその分野の外に出ることもあります。

現在、引用データに関する多数のマップを作成していますが、その主な目的はこれらのマップを評価ツールとして用いることであり、また科学の特定分野で起こっていることを、時間を追って観察することです。われわれのウェブサイトでは、科学のマップを使ってナビゲイトすることができます。最終的には、個々の論文にまで案内したいと考えています。矢印は流れの方向を表し、例えば生態・進化学のような特定分野では分析化学に向かう矢印はありますが、逆向きの矢印は存在しません。このようにして、マップを用いると、「それぞれの分野がどのように関連しているか」「アイデアの流れはどこからきてどこへ行くのか」といった疑問について調査することができます。われわれのマップは、このような疑問に適したツールなのです。また、各分野にはクリックすることでより詳しい情報が得られる雑誌もあり

ます。通常のマップのように、マップを詳細化することも簡略化することも可能です。

結論

アイゲンファクター・プロジェクトの趣旨は、ネットワーク科学や生物学で開発されたツールが、他のデータセットに対しても同様に適用できる、ということです。これらのデータセットは情報の宝庫ですから、図書館員、出版者、編集者、研究者、そして学術にかかわるすべての人に役立つように、その情報を抽出するアルゴリズムやビジュアルツールを開発することがわれわれの挑戦すべき課題です。アイゲンファクター・プロジェクトの主な目標の一つは、学術文献を横断的に検索するためのより良い方法を開発するためにこれらのツールを活用することであり、これにより研究者が読むべき文献を探し当てることができるようにすることです。

謝辞

共同研究者である、カール・バーグストローム、テド・バーグストローム、ベン・アルシュース、モーリッツ・ステファナー、およびマーチン・ロスヴァルに深く感謝します。