

Title	科学論文を引用することは特許の影響力を増大させるか
Author(s)	富澤, 宏之
Citation	年次学術大会講演要旨集, 25: 499-501
Issue Date	2010-10-09
Type	Conference Paper
Text version	publisher
URL	http://hdl.handle.net/10119/9346
Rights	本著作物は研究・技術計画学会の許可のもとに掲載するものです。This material is posted here with permission of the Japan Society for Science Policy and Research Management.
Description	一般講演要旨

科学論文を引用することは特許の影響力を増大させるか

○富澤宏之（文科省・科学技術政策研）

1. はじめに

科学研究がイノベーションに及ぼす影響の解明は、科学技術政策研究や進化経済学における重要課題である。この課題へのアプローチの一つとして、特許における科学論文の引用データから作成されたサイエンスリンケージ指標（特許一件当たりの科学論文の引用件数）の分析が試みられてきた。この指標は、漠然とした意味では、特許と論文のリンケージの大きさを示していると考えられる。しかし、そもそも、特許が科学論文を引用することの意味は十分に解明されておらず、したがって、この指標が何を示すのか、必ずしも明確ではない。

このような現状を踏まえ、本研究では「特許が科学論文を引用することは、特許の影響力を増大させるのか」という疑問について、実証的に検討する。このアプローチは、特許が科学論文を引用することの意味を内在的論理から追及するものではないが、その意味を、外在的論理から明らかにしようとする試みであると言える。なお、このようなアプローチでの先行研究は多少、存在するが、十分な研究が行われているとは言い難い。

2. 特許による科学論文引用の概要

本研究では、1995年から2005年までの11年間に米国特許商標庁に登録された特許のデータセット（ipIQ社製）を用いた。このデータセットには、各特許のフロント・ページで引用された文献の情報が収録されている。これらの文献は、特許の内容の説明や、既存の科学技術との関係を示すために、特許の出願者や特許審査官が示した文献である。これらの引用文献の多くは先行特許であるが、非特許文献も含まれている場合がある。非特許文献には、科学論文、単行本、会議資料、新聞記事、各種雑誌記事、広告やカタログ、あるいはwebサイトなど、様々な文献が含まれている。

このデータにより、特許による科学文献の引用という現象が、どの程度の規模で起きているのかがわかる。1995～2005年に米国特許商標庁が発行した特許は1,591,996件であるが、そのうち非特許文献を引用している特許は585,184件であり、全体の36.8%を占めている。また、非特許文献は4,863,560件であるので、単純に計算すれば、非特許文献を引用している特許1件あたり平均8.3件の非特許文献を引用していることになる。

非特許文献4,863,549件のうち、科学論文と見なすことができるものは2,737,174件である。その科学論文を引用した特許の数は299,295件であり、全登録特許の18.8%である。このように、特許が科学論文を引用するという現象は、全ての特許で起きているわけではない。このことは、科学論文と特許の関係を考察する際の前提として重要である。

3. 科学論文の引用と特許の被引用度の関係

以上のようなデータに基づいて、本論文の主題である疑問に答えるために、「科学論文を多く引用する特許は、影響力の強い特許となる場合が多い」あるいは「影響力の強い特許は、一般的な特許に比べて科学論文を多く引用している」という仮説を設定して、その検証を試みる。

影響力の強い特許を同定する方法として、特許の被引用度を用いた。これは、科学論文の引用と同様に、他の特許に頻繁に引用された特許は影響力がある、という考えに基づいている。ただし、ある特許が先行特許を引用する理由は、先行特許の技術的内容との相違ないし関係を示すためであることが多く、先行特許の影響力と直接的に関係があるとは限らない。それでも、頻繁に引用される特許は、そうでない特許に比べると、統計的な意味で（すなわち確率的に）影響力が高い、と考えることは十分に妥当であろう。

ここで用いた特許の被引用度のデータは、2009年9月時点の観測値である。つまり、1995～2005年に登録された特許について、それらの特許が2009年9月までの期間に引用された回数が基本的なデータである。ただし、このような被引用回数は、技術領域によって大きく異なり、また、発行されてからの経過年の長さによっても異なる。そこで、被引用回数をそのまま用いるのではなく、特許の技術分野と発行年に関して標準化した値を用いた、具体的には、技術分野Aにおける発行年yの特許の被引用回数を、技術分野Aで発行年yの特許全体における被引用回数の期待値（平均値に等しい）で除して標準化した。技術分野の分類は、米国特許についての信頼性の高い分野分類であるとされる米国特許分類（約750分野）を用いた。

まず、「影響力の強い特許は、一般的な特許に比べて科学論文を多く引用している」という仮説を検討する。そのために、前述のように算出した特許の被引用度指数の階級（上位1%、3%、10%、・・・）ごとに、サイエンスリンケージ（特許1件あたりの科学論文引用件数）を計算し、図1に示した。

図1 特許の被引用頻度階級別のサイエンスリンケージ

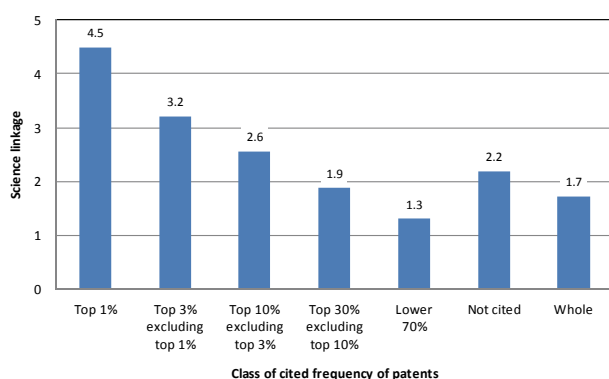


図1によると、米国特許全体のサイエンスリンケージが1.7であるのに対し、被引用度上位1%特許のサイエンスリンケージは4.5（前者の2.6倍）であり、被引用度上位3%特許（被引用度上位1%特許を除く）の値も3.2と高く、「被引用度の高い特許は、一般的な特許に比べて科学論文をよく引用していた」ことが示された。この結果は、「他の特許によく引用される特許は、平均的な特許に比べて、科学知識との結びつきが強い」と解釈できる。ただし、こ

のような関係は、被引用度上位1%や3%というかなり限られた特許についてのみ成り立っており、上位3%以下の特許では明確な関係があるとは言えない。

このように、特許の被引用度とサイエンスリンケージの間の統計的關係は明確ではあるものの、限定的な関係である理由として、次の二点が考えられる。第一に、特許の被引用度が高くなる要因として、科学論文の引用以外の要因もあることは明らかである。つまり、科学知識の活用とはあまり関係なくとも、技術的に革新的であり大きな影響力をもつ特許が存在することは当然である。第二に、既に述べたように、被引用度は特許の影響力の精度の高い指標ではないため、上位1%や3%のように被引用度が十分に大きい場合に初めて、特許の影響力が大きいことを示しているためであると考えられる。

図1を解釈する際には、前節で述べたように、特許が科学論文を引用するという現象は全ての特許で起きているわけではない、ということを念頭に置く必要がある。そのことを明確にして、図1の解釈を深めるために、図1のサイエンスリンケージ指標に替えて特許の参考文献の引用パターンを用い、特許の被引用度との関係を図2に示した。

図2 特許の被引用頻度階級別の引用パターン

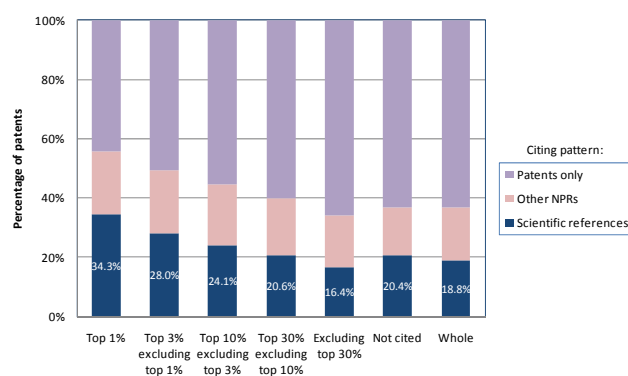
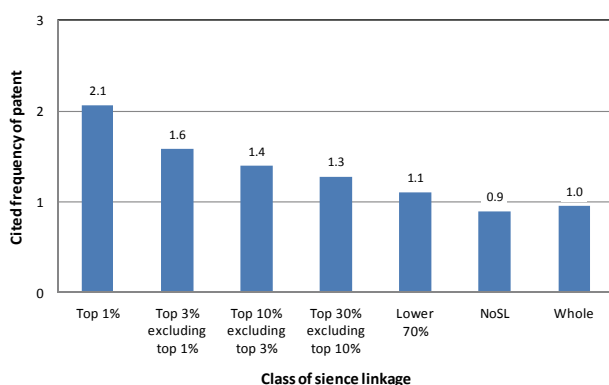


図2によると、被引用度上位1%の特許では、その34.5%が科学論文を引用している。この割合は、米国特許全体についての値（18.8%）の1.8倍であり、やはり、前述の仮説を支持する結果である。ただし、被引用度上位1%の特許の大部分が科学論文を引用していたわけではないので、科学論文の引用が影響力の強い特許を生み出す要因だとしても、部分的なものであることが改めて示された。

次に、「科学論文を多く引用する特許は、影響力の強い特許となる場合が多い」という仮説について検討する。これは、先に検討した仮説と同じ二つの変数の関係を扱っているが、「科学論文の引用」を説明変数とし、「特許の影響力」を被説明変数としており、自然な因果関係により近い表現となっている。もちろん、これだけのデータによって因果関係を立証できるわけではなく、ここでは統計的な相関を調べているに過ぎない。

図3は、サイエンスリンケージの階級ごとの特許被引用度を示している。ちょうど、図1の説明変数と被説明変数を入れ替えた形であり、説明変数であるサイエンスリンケージは、値の大きい方から上位1%, 3%, 10%, …という階級に区分し、一方、被説明変数として特許の被引用度指数を表示している。図3によると、特許全体の被引用度指数が1.0であるのに対し、サイエンスリンケージが上位1%の特許の被引用度指数は2.1と高く、サイエンスリンケージの大きさは、特許の被引用度にある程度影響を及ぼしているようであり、設定した仮説を支持する結果となっている。

図3 サイエンスリンケージ階級別の特許被引用度指数

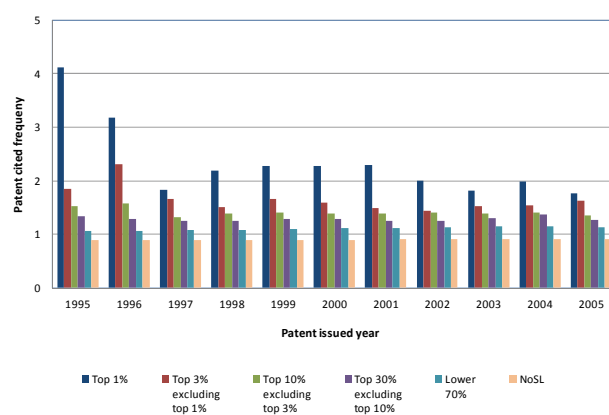


しかし、特許の被引用度は時間がたつにつれて変化する値であり、図3に示した関係については、年別に見る必要がある。そこで、図3に示した変数を、特許の発行年別に分けて、図4に示した。

この図によると、特許の発行年が古い場合には、図3に示された関係が強く成り立っており、一方、発行年の新しい特許では、その関係は弱い。これは、「サイエンスリンケージの大きさは、特許の被引用度を上昇させるような影響を及ぼす。ただし、その

影響が現れるまでは何年か要する」ということを意味していると考えられる。また、被引用度の大きさを正しく観測するためには、ある程度の時間が必要である、という引用の基本的性質の反映であるとも考えられる。このような時間の影響は、常識的な実感に合っており、ここでのデータ分析の結果にリアリティを与えている。

図4 特許登録年およびサイエンスリンケージ階級別の特許被引用度



6. まとめ

本研究では、特許が科学論文を引用することは、特許の影響力を増大させるか、という問題について、米国特許のデータを用いて検討した。データ分析によって示されたことは、「他の特許によく引用される特許は、平均的な特許に比べて、科学論文をよく引用する傾向がある」および「科学論文をよく引用する特許は、被引用度の大きい特許となる傾向がある」ということである。ただし、これらの関係は、特許の被引用度と科学論文引用件数（サイエンスリンケージ）のそれぞれが上位1%や3%といった極めて高い値の時にのみ成り立っていた。また、その関係自体も統計的な関係であり、確率的にそのような傾向があるというに過ぎない。

以上の結果は、「特許と科学論文との結びつきは、影響力の強い特許を生み出す要因であるが、部分的な要因である」ということを示している。このことは、部分的とはいえ、科学が技術の発展に影響を及ぼしていることを定量的に示した点で、一定の意義があると言う事ができるであろう。