

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 24 日現在

機関番号：13102

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730140

研究課題名(和文) 自然言語処理を利用したパテントポートフォリオのスコアリング

研究課題名(英文) Patent Portfolio scoring based on natural language processing

研究代表者

野中 尋史 (Nonaka, Hirofumi)

長岡技術科学大学・工学(系)研究科(研究院)・講師

研究者番号：70544724

交付決定額(研究期間全体)：(直接経費) 3,000,000円

研究成果の概要(和文)：本研究ではパテントポートフォリオを統合的に評価するためのスコアについての検討を行った。具体的には引用ベースと地域ベースのスコアの開発を行った。引用ベーススコアについてはPagerankをベースとした指標を開発した。財務指標などとの相関による評価を行い妥当性を確認した。一方、地域性評価スコアは空間的自己相関をベースとしたスコアを開発した。産業クラスターの発展状況などとの比較により妥当性を確認した。

研究成果の概要(英文)：We developed patent portfolio score based on citation and location. Citation score is based on Pagerank. On the other hand, location score is based on spatial auto-correlation.

研究分野：知財情報処理，自然言語処理

キーワード：空間的自己相関 Pagerank Moran's I 産業クラスター パテントポートフォリオ評価

1. 研究開始当初の背景

パテントポートフォリオとは、出願人が戦略的に出願・取得を行った基本特許とそれを支える周辺特許から構成される特許群を指す。強固なパテントポートフォリオを保有する企業ほど他社からの侵害を防止し、独占力を有するためきわめて高い付加価値を生み出す。このため、パテントポートフォリオの価値を定量的に評価し、それに基づき財務指標等の予測を行うことは、出願人にとっては、自社の経営戦略の妥当性評価につながり、また、競合相手の評価も行えるなど経営の意思決定に重要な役割を果たす。また、金融機関・投資家にとっては、投資先の技術力評価、投資の意思決定に役立つ。

このような背景のもと特許価値を評価するためのモデルの研究が経営学において盛んにおこなわれている。この中では、リアルオプションと呼ばれる金融工学におけるオプションの評価モデルを現物資産評価に応用した手法により知財価値を評価することが提唱されている[1]。ただし、この手法では、将来のキャッシュフローや金利など分析者が設定せねばならないパラメータがあり、分析者によってまちまちの結果が得られる可能性が大きい。もっともらしいパラメータを設定するためには知財の内容を厳密に検討する必要がある。一方、PatVM[2]と呼ばれるパテントポートフォリオ評価手法を提唱している。この手法では、知財の評価項目(権利の強さなど)を設定し、それぞれの項目を専門家が特許文書をはじめとする知財の内容に基づき評価していくものである。この手法でも、分析者の主観は排除できず、また、厳格に行うためには多大な労力を要する。このため、より客観的で簡便に特許文書など知財の内容を評価できる手法の開発が求められている。内容の評価を簡便、かつ、定量的に行うためには、文書を解析する技術である自然言語処理の手法が必要となる。

一方、自然言語処理の分野において、特許文書をはじめとする知財文書を対象とした研究は機械翻訳や特許検索に関するものが主流となっており、特許文書の内容評価に踏み込んだ研究は少ない。この中では、梅村ら[3]のように、特許庁の査定を通過するかどうかについて特許文書の内容から機械学習を利用して判定する手法や、佐藤ら[4]のように引用情報に基づき発明協会の賞をとるような重要な特許であるかどうかの判定を行った研究がある。しかしながら、評価を行う上でもっとも重要となる特許文書の内容と経営指標との因果関係について分析したのではない。このため、文書の内容と経営指標の関係性を分析する手法の開発が求められていた。また、これまでの手法は単体特許の内容のみを評価するものであり、特許群であるパテントポートフォリオを評価するものではなかった。そこで、本申請研究では、地域性や引用件数に着目してパテントポ

リオを定量的に評価し、経営指標を予測するモデルを構築する。

2. 研究の目的

パテントポートフォリオを統合的に評価するため引用や地域性の観点から評価するスコアを開発する。

3. 研究の方法

引用に基づく評価はWebサイトの重要性評価に使用されるHITSやPagerankアルゴリズムをベースとしたスコアの開発を行う。また、出願人のパテントポートフォリオの評価を行うにあたっては、地域性を加味することも重要となる。産業クラスターなどの地域内の技術集積ネットワークが企業の研究開発に重要な効果をもたらすことから、地域性にそったテーマのパテントポートフォリオは価値が高いと考えられる。本研究では、空間的自己相関モデルにより検定を行うことでその地域的偏在性を見出し、地域の技術開発の潮流と合致しているか判定を行えるベースとなるスコアを開発を行う。

これらのスコアについては統計モデルを利用して財務指標・株価との関係性を調査し妥当性を評価する。

4. 研究成果

まず、平成25年度はリンク評価手法であるHITS(Authority値を使用)を用いた引用ベースパテントポートフォリオスコアリング手法を検討した。妥当性を評価するため株価との相関分析を行った。なお、HITS authority値の計算に用いた特許は1996年から2006年までに公開された公開特許公報3,875,711件である。株価は1996年から2012年までのデータを使用し、東証一部上場企業のうち、対象期間に社名や商号を変更して行っていない企業を調査対象とした。相関分析においては決定係数を用い、モデルから予測される各年の株価と実際の株価のズレを評価した。決定係数は観測値の分散とモデルによる値の分散の比で定義され、式(3)によって定義される。

決定係数は0から1までの値を取り、1に近いほどモデルにより観測値の変動が説明できることを表す。決定係数から企業を3つのグループに分け、A:モデルに強く従っている、B:モデルに従っている、C:ほとんど従っていないに分類し、業種や企業規模別のモデル評価を行った。

Table 1. Classification by R^2 (決定係数による分類)

A	B	C
$0.5 \leq R^2 \leq 1$	$0.3 \leq R^2 \leq 0.5$	$0 \leq R^2 \leq 0.3$

モデルの回帰分析の結果を、業種毎に分類した結果を表2に示す。表中の値は、各企業に対し最もよい(大きい) Lag をその企業

の属する表1のグループとした時のその業界の企業数に対する比である(小数点以下四捨五入)。企業を分類する業種としては証券コード協議会が個別の上場企業に定める業種コードを利用した。

Table 2. Industry-classified result(業種別評価結果)

業種	社数	A	B	C
医薬品	13	0.10	0	0.90
電気ガス	8	0.24	0.63	0.12
食料品	25	0.55	0.14	0.32
サービス	2	0.50	0	0.50
陸運	5	1.00	0	0
パルプ紙	5	0.80	0	0.20
ガラス	15	0.50	0.17	0.33
繊維	13	1.00	0	0
鉄鋼	20	0.89	0.06	0.06
精密機器	8	0.50	0.12	0.38
非鉄金属	11	0.45	0.09	0.45
輸送機器	25	0.57	0.24	0.19
機械	40	0.72	0.08	0.21
化学	54	0.45	0.22	0.33
鉱業	1	1.00	0	0
建設	39	0.91	0.09	0
情報通信	3	0	1.00	0
水産農林	2	0.5	0	0.5
ゴム	6	0.83	0	0.17
石油石炭	4	1.00	0	0
電気機器	62	0.41	0.14	0.44
金属製	10	0.60	0.2	0.2

品				
その他	18	0.87	0.07	0.07
倉庫運輸	1	1.00	0	0
卸売業	9	0.86	0	0.14

Table 3. Size-classified result(規模別評価結果)

規模	A	B	C
TOPIX30	0	0.44	0.56
TOPIX70	0.20	0.35	0.45
TOPIX400	0.50	0.15	0.36
TOPIX Small	0.73	0.10	0.18

結果から、規模が小さい企業ほど HITS を指標とする特許の質が、その企業の株価に影響を与えることが分かった。これは、規模が小さい企業ほど技術力により付加価値を生み出していることが影響していると考えられる。また、このことから HITS が中堅以下の企業の技術ポートフォリオを測る上での指標として有効であることも示すことができた。実用上、技術力が収益に直結するような企業の評価において、特許ポートフォリオ評価手法が重要となることから、本手法の有用性は大きいと考える。一方、前節の表2から示唆されるように、業種別でも差異が見られた。傾向として B to B の事業を展開している鉄鋼、機械などの業種はモデルに従いやすい傾向にあるといえるが、B to C の事業を展開している電機(大企業が多いことも影響)などはあまりモデルに従っていない。これは、B to C の事業の場合は B to B の事業と比較して、広告宣伝力など技術面以外の要素がより企業業績に反映されるためと推測される。

一方、地域に着目したスコアについては空間的自己相関をベースに検討を行った。具体的にはグローバル、および、ローカルな空間的自己相関による特許出願の空間的偏在性の定量的評価法の提案と評価を提案した。さらに、国家戦略において重要な分野であり、他の技術分野と比較して地域分散傾向を示したバイオ分野に焦点をあててスコアの妥当性の分析を行った。

Local Moran's I は(2)式で表され、個別地域ごとに算出される。当該地域の Local Moran's I が正の大きな値をとるとき、自身の値が周囲と似ていて類似性が高い傾向を示す(局所的な正の空間的自己相関)。また、当該地域の Local Moran's I が負の大きな値をとるとき、近隣地域と非常に異なるよう

な傾向を示す（局所的な負の空間的自己相関）。さらに、当該地域の Local Moran's I が 0 に近い値をとるとき、周囲のデータとの間の関連性が低いを示す。

$$I_i = \frac{n(x_i - \bar{x})^2 \sum_j^N w_{ij}(x_j - \bar{x})}{\sum_i^N \sum_j^N w_{ij} \sum_i^N (x_i - \bar{x})^2} \quad (2)$$

ここで、 $I = \sum_i^N I_i / (\sum_i^N \sum_j^N w_{ij})$

であり、各地域の Local Moran's I の総和は、Global Moran's I と比例関係が成立する。その結果、市町村のような狭い行政区域での横断、もしくは、局在型でのクラスター整備がなされていること、三大都市圏以外の地方都市レベルでも市町村をまたぐ横断型クラスターが形成されていることを示唆していること、大学医学部などの学術機関が中核となっている可能性があること、すでにクラスターが形成される地域は固定化しつつあり深化傾向にあることなどが示唆された。これらは実際のクラスター構想とリンクしていることが分かりスコアが地域性をよく表していることが分かった。

<引用文献>

- [1] 榊原茂樹，リアルオプションと特許権の評価，国民経済雑誌，188(5)，pp.1-14，2003
 [2] 広瀬義州他，特許権価値評価モデル Patent VM，東洋経済新報社，2005

5. 主な発表論文等

（研究代表者、研究分担者及び連携研究者には下線）

〔雑誌論文〕(計 5 件)

- [1]木村誠，野中尋史，平岡透，太田貴久，増山繁，引用ネットワークを利用した企業の特許ポートフォリオ評価と株価変動の関係分析，電子情報通信学会論文誌 D，J97-D，No.3，pp.700-702,2014.
 [2]太田貴久，野中尋史，平岡透，増山繁，自然言語処理を用いた特許拒絶理由の統計的分析と審査プロセス解析への応用，日本知財学会誌，第 10 巻，第 3 号，pp.71-88，2014.
 [3]Hirofumi Nonaka, Daiki Kubo, Makoto Kimura, Toru Hiraoka, Takahisa Ota, Shigeru Masuyama, Correlation Analysis between Financial Data and Patent Score based on Reference Network Algorithm HITS, Proc of IEEE International Technology Management Conference(ITMC) Chicago, USA, 2014.
 [4]Hirofumi Nonaka, Seiya Kawano, Toru Hiraoka, Takahisa Ota, Shigeru

Masuyama, Evaluating Industrial Cluster by using Spatial Auto Correlation of Patent Applications, Proc. of International Conference on Advanced Informatics: Concepts Theory and Applications (ICAICTA2014), Bandung, Indonesia pp.188-191, 2014.

[5]Hirofumi Nonaka, Seiya Kawano, Kenjiro Anai, Toru Hiraoka, Takahisa Ota, Shigeru Masuyama, Assessment Method for Technology Development of an Industrial Cluster by Using Spatial Auto-correlation of Patent Applications, Proc. of 43th International Conference on Computers and Industrial Engineering (CIE43), Hong Kong, China, 2013.

〔学会発表〕(計 9 件)

- [1]川嶋和希，野中尋史，小林暁雄，太田貴久，増山繁，ソフトウェア使用許諾書を対象とした重要条項の抽出，人工知能学会第 28 回全国大会，2014.
 [2]久保大輝，野中尋史，特許引用ネットワークの分野補正 HITS スコアと株価・財務指標との相関分析，日本経営工学会九州支部第 40 回学士・修士論文発表会，2014. 最優秀賞
 [3]河野誠也，野中尋史，特許出願の局所空間的自己相関と技術分野分布に基づくエントロピーを利用した地域イノベーション力の評価手法の検討，日本経営工学会九州支部第 40 回学士・修士論文発表会，2014.
 [4]太田貴久，野中尋史，平岡透，増山繁，自然言語処理技術に基づく拒絶理由の自動推定とその分析，第 11 回日本知財学会年次学術研究発表会，2013.
 [5]河野誠也，野中尋史，穴井賢次郎，平岡透，太田貴久，増山繁，空間的自己相関を利用した特許文書の解析による技術集積度指標の検討，第 11 回日本知財学会年次学術研究発表会，2013.
 [6]野中尋史，久保大輝，木村誠，平岡透，太田貴久，増山繁，自然言語処理を利用した特許スコアリングと株価・財務指標の関係性分析，第 11 回日本知財学会年次学術研究発表会，2013.
 [7]木村誠，野中尋史，引用ネットワークを利用した特許文書の重要性評価の研究，日本経営工学会九州支部第 39 回学士・修士論文発表会，2013.
 [8]穴井賢次郎，野中尋史，空間統計学を利用した特許出願の地域偏在性に関する研究，日本経営工学会九州支部第 39 回学士・修士論文発表会，2013. 最優秀賞
 [9]原田綾花，太田貴久，小林暁雄，増山繁，野中尋史，酒井浩之，特許文書からの発明に関する特徴的技術とその効果の抽出，言語処理学会第 19 回年次大会発表論文集，2013.

〔図書〕(計 1 件)

〔産業財産権〕
出願状況（計 件）

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況（計 件）

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

6. 研究組織

(1) 研究代表者

野中 尋史 (NONAKA, Hirofumi)
長岡技術科学大学・工学研究科・講師
研究者番号： 70544724

(2) 研究分担者

()

研究者番号：

(3) 連携研究者

()

研究者番号：