

# ネットワーク分析を用いた計算機利用技術分野の変遷

○原田泰輔 吉川厚 寺野隆雄 (東京工業大学)

## Complex network approach to the trend of Computer utilization technology Field

\*T.Harada, A. Yoshikawa, and T. Terano (Tokyo Institute of Technology)

**Abstract**—In this research, we analyze the transition of research fields and scholar's research trend based on more than 10 million JST (Japanese Science and Technology agency) database and KAKENHI (Grants-in-Aid for Scientific Research) data. The contribution of this research is to investigate how researchers become the contributors in their research fields and won the research fund by applying the methodology of complex network. We analyze the trend of relation between researchers' network and research fund in 10 year's time span using the academic discipline data. Also, the pattern analysis has been applied to classify the grow process of researchers who has won large amount of fund based on their affiliation and field information.

**キーワード:** ネットワーク分析, 共著者ネットワーク, 競争的研究資金

### 1 はじめに

研究者が研究活動を行う上で研究資金は欠かせないものであり, 科学技術の発展は研究資金によって支えられてきた. 研究資金の多寡によっては, 研究規模, 研究プロセスが制限される. そのため, 研究者にとって研究資金の獲得は, 重要な課題となっており, 研究費を得やすい課題名についての分析といった研究もなされている<sup>1)</sup>.

研究資金は, 大学や研究施設の校費や企業からの融資といった非競争的資金と, 研究者からの応募に対して評価, 採用され資金を配分する, 競争的資金とに分けられる. その中でも, 特に大きな役割を果たしているのが, 科学研究費助成事業(以下, 科研費)である. 科研費は, 『人文・社会科学から自然科学まで全ての分野にわたり, 基礎から応用までのあらゆる「学術研究」(研究者の自由な発想に基づく研究)を格段に発展させること』を目的とする「競争的研究資金」であり, ピア・レビューによる審査を経て, 助成されるものである. 平成26年度の競争的資金制度の予算額は, 全体で4,162億円, そのうち科研費額は2,276億円で, 科研費額が全体の約55%を占めており, 科研費は多くの研究者にとって欠かせないものである.

一方で, 研究活動は多くの場合, 研究者個人だけで

はなく, 大学の研究室や研究グループといった集団で行われている. また, 研究領域によっては, 他の研究分野の研究者と共同研究という形で研究が行われており, より効果的な共同研究を行うための取り組みが行われている<sup>2)3)</sup>. これらのことから, より良い研究成果をあげる上で, 他の研究者との関係が如何に重要であるかがわかる. また, 研究者間の関係の重要性は, 科研費がピア・レビューによる審査で決定されていることから明らかである.

こうした背景から, 研究者間において有意な位置を占めている研究者は, 研究費の獲得においても有意であると考えることができる. そこで, 本研究では, 研究者ネットワークでの優位性と研究費との関係を明らかにすることを目的とし, 研究者ネットワークでの中心性と科研費の取得率について分析を行う. さらに, 研究者の動向を分析し, 共著者ネットワーク上で優位な位置を占める研究者の特徴の抽出を行う.

本論文の構成は以下の通りである. 2章で関連研究の紹介をし, 3章では, 使用データについての説明を行い, 4章で科研費と共著者ネットワークの関係についての分析を行う. 5章で共著者ネットワークを用いた研究者のクラスタリングを行い, 6章で研究者クラスタの特徴について分析を行う. 最後に

7章で全体のまとめを述べる。

## 2 関連研究

この章では、本研究に関連する研究について述べる。

### 2.1 研究者の科研費動向に関する研究

西澤らは、学術賞受賞者の科研費取得遍歴を分析することにより、大型研究の採択に至るまでのパターン、分野の違いによる特徴的な採択パターンについて分析を行っている<sup>4)</sup>。また、柿沼らは、研究者のキャリアパスを年齢により4つのステージに分け、各ステージにおける研究者の科研費額、学問分野間の科研費額の隔たりについての分析を行っている<sup>5)</sup>。しかし、これらは、科研費取得における共同研究者の要因や研究領域内でのポジションには着目していない。

### 2.2 研究者ネットワーク

研究者ネットワークに関する研究は数多く行われている<sup>6)-10)</sup>。篠田は、ある研究領域における第一人者や貢献者を、論文の共著関係から見つけ出し、それらの研究者の系譜の作成を行っている<sup>8)</sup>。

しかし、これらの研究では、その領域の第一人者や貢献者を見つけて出すことに主眼を置いており、研究の成長過程に着目はしていない。

## 3 使用データ

ここでは、本研究で使用したJST学術文献データと科学研究費助成事業データベース(KAKEN)についての説明を行う。

### 3.1 JST学術文献データ

本研究では、科学技術振興機構の提供する、JST科学技術文献データをもとに、2003年から2013年までの『計算機利用技術分野』を対象として分析を行なった。JST科学技術文献データには、2003年から2013年までの論文タイトル、著者、論文の分野、発行年

月日、発行国、著者の所属機関の情報が含まれている。JST科学技術文献データにおける計算機利用技術分野には、Table1に示す分野が含まれている。計算機利用技術分野の論文数、著者数、ネットワーク密度、連結成分の数の変化はそれぞれの期間についてTable2のようである。また、2003年時の計算機利用技術分野の共著者ネットワークがFig. 1である。

本研究では、著者の氏名表記の揺れを抑える為、JST科学技術文献データの人名名寄せデータを用いてネットワークの作成を行なった。人名名寄せデータは、各著者に固有のIDを付与したものである。

Table1: 計算機技術利用分野に含まれる分野

計算機利用技術分野			
計算機利用技術一般	数値計算	記号処理	図形・画像処理
図形・画像処理[*81-'92]	図形・画像処理一般	医用画像処理	音声処理
自然語処理	パターン認識	人工知能	CAI
CAD, CAM	計算機シミュレーション	医用情報処理[*81-'92]	医用情報処理
事務・経営情報処理[*81-'92]	その他の計算機利用技術	情報処理	情報処理一般
事務・経営情報処理	分子・遺伝情報処理	その他の情報処理	

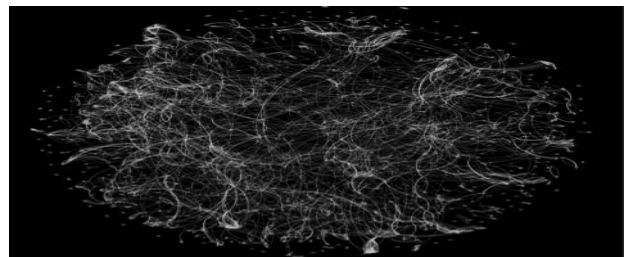


Fig.1: 計算機技術利用分野の共著者ネットワーク

Table2: 計算機技術利用分野の共著ネットワーク推移

	2003-2006	2004-2007	2005-2008	2006-2009	2007-2010	2008-2011	2009-2012	2010-2013
ノード数	69629	77151	82724	87841	90257	90121	89536	82421
エッジ数	192203	224639	250099	271661	278845	273459	266320	221437
ネットワーク密度	7.93E-05	7.55E-05	7.31E-05	7.04E-05	6.85E-05	6.73E-05	6.64E-05	6.52E-05
連結成分数	8023	8703	9207	9497	9816	10123	10383	10018
最大連結成分の占める割合	45.74%	45.89%	46.25%	46.40%	46.64%	45.58%	43.75%	40.86%

### 3.2 科学研究費助成事業データベース(KAKEN)

研究者の科研費取得状況を把握するため、科学研究費助成事業データベース(KAKEN)より科研費情報の取得を行った。KAKENには、当初採択時のデータ(採択課題)、研究成果の概要(研究実施状況報告書、研究実績報告書、研究成果報告書概要)、研究成果報告書及び自己評価報告書が収録されている。本研究では、JST学術文献データより抽出した研究

者名に対し、2009年から2012年までの採択の有無と研究種目をKAKENより取得した。

## 4 科研費と研究者ネットワーク

この章では、共著者ネットワークの作成手法、科研費と共著者ネットワークとの中心性指標による順位の関係についての分析を行う。

### 4.1 共著者ネットワークの作成

共著者ネットワークの作成は下記の手順で行った。

- (1) 各論文をその発行年をもとに 第1期 (2003-2006), 第2期 (2004-2007), 第3期 (2005-2008), 第4期 (2006-2009), 第5期 (2007-2010), 第6期 (2008-2011), 第7期 (2009-2012), 第8期 (2010-2013) の8つの期間に分け、各期間での論文の著者データから、共著者間全てにリンクを張る。
- (2) (1)をもとに、複数回の共起を考慮しない、重みなし無向グラフを作成する。

### 4.2 中心性指標による研究者の順位付け

4.1で作成した、各期間での共著者ネットワークを用い、次数中心性、媒介中心性、近接中心性の値を全著者に対し求め、その値により各中心性での順位付けを行った。用いた中心性指標の詳細が下記である。

#### ・次数中心性

次数中心性は、ネットワーク内でより多くのリンクを持つノードを高く評価する指標であり、ネットワーク内のノードとどの程度リンクを張っているかを示す指標である。グラフの隣接行列を  $a_{ij}$  とすると、次のように定式化される。

$$C_d(i) = \sum_{j=i}^n a_{ij} \quad (1)$$

ここで、 $n$  はグラフに含まれるノード数である。

#### ・媒介中心性

媒介中心性は、あるノードが他のノード間の最短経路上に位置する程度を示す指標であり、他のノード同士をつなぐ働きをするノードを示す。

$$C_b(i) = \sum_{i \neq j \neq k} \frac{g_{jk}(i)}{g_{jk}} \quad (2)$$

ここで、 $g_{jk}$  は頂点  $j$  と頂点  $k$  の間の最短経路のうち頂点  $i$  を通るものである。

#### ・近接中心性

近接中心性は、他のノードへの最短距離の総和の逆数であり、ネットワーク内の全てのノードとどれくらい近いのかを示す指標であり、次のように定式化される。

$$C_c(i) = \frac{1}{\sum_{j=1}^n d_{ij}} \quad (3)$$

ここで、 $\sum_{j=1}^n d_{ij}$  は頂点  $i$  から他の頂点への最短距離の合計である。

### 4.3 中心性順位と科研費

ここでは、科研費と共著者ネットワークの中心性指標による順位との関係について分析を行った。

第7期で各中心性 Top2000 位以内の研究者について、科研費(基板 B 以上)の取得状況の調査を行った。ここでは、各中心性指標での順位を 100 名ずつ 20 区間に区切り、各順位区間の研究者の科研費の取得率を求めた。その結果が、Fig.2 である。また、グラフ内の赤線は、1000 人の無作為に抽出した研究者の科研費の取得割合である。

Fig.2 から、各中心性順位の上位者は、科研費を取得している割合が高いことがわかる。各中心性順位と科研費の取得率との相関係数をまとめたものが、Table 3 である。Table 3 から、いずれの中心性指標でも科研費の取得割合と相関があることがわかる。

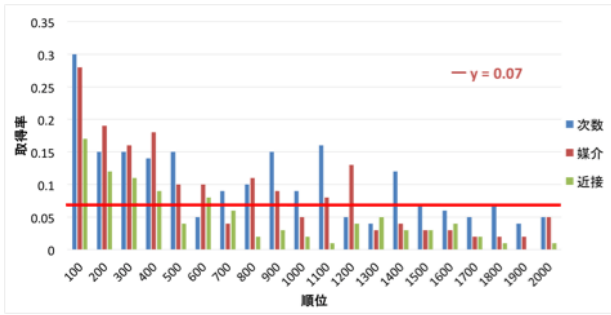


Fig.2: 中心性での順位と科研費取得率

Table3: 中心性指標での順位と科研費取得率の相関

	次数順位-科研費取得率	媒介順位-科研費取得率	近接順位-科研費取得率
標本数	20	20	20
自由度	18	18	18
r 値	-0.71	-0.82	-0.80
p 値	4.4E-06	8.1E-06	1.9E-05

Table4: 過去期間での媒介順位と第7期での科研費取得割合の相関

	第1期	第2期	第3期	第4期	第5期	第6期
標本数	20	20	20	20	20	20
自由度	18	18	18	18	18	18
r 値	-0.76	-0.82	-0.80	-0.75	-0.89	-0.77
p 値	8.3E-05	8.6E-06	2.1E-05	1.4E-04	1.8E-07	7.5E-05

次に、第七期で科研費を取得している研究者の過去の期間での媒介中心性順位との相関係数をまとめたものが Table 4 である。Table4 から、過去の順位と、その後の期間での科研費の取得状況に相関があることがわかる。

これらの結果から、中心性指標での順位と科研費の取得率との間に相関関係があることがわかる。

#### 4.4.4 章のまとめ

4 章では、共著者ネットワークの中心性指標での順位と科研費との関係について分析を行った。その結果、t 期において中心性順位の高い研究者は、t+1 期で科研費を取得できる割合が高いことが確認できた。

## 5 研究者のクラスタリング

ここでは、共著者ネットワーク分析の中心性指標による研究者の順位を用いた研究者のクラスタリングを手法、研究者のクラスタリング結果を記す。

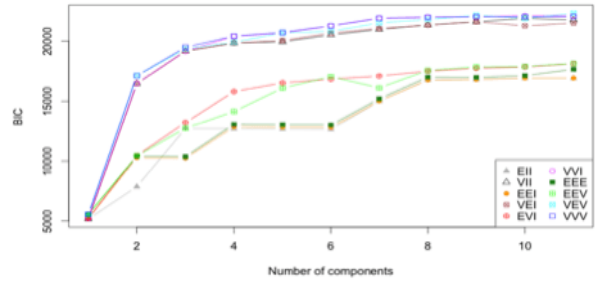


Fig.3: クラスタ数と BIC 値

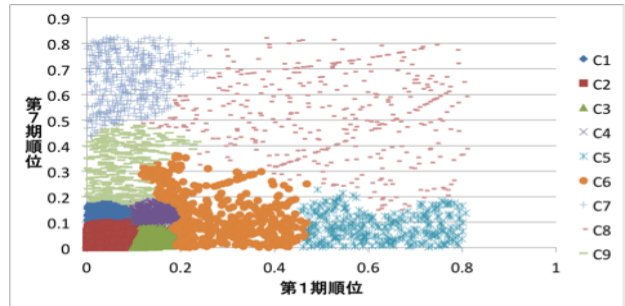


Fig.4: 第1期と第7期の順位を用いた研究者のクラスタリング

### 5.1 研究者のクラスタリング手法とクラスタ数の決定

第1期間と第七期の中心性指標での順位を用い、2003年から2012年の7期間で存在する、5658人を対象とし研究者のクラスタリングを行う。混合分布モデルを適用し、BIC値(ベイズ情報基準)から、混合分布モデルとクラスタ数の決定を行った。混合分布モデルによるクラスタリングとBIC値についての詳細は以下の通りである。

混合分布モデルによるクラスタリングは、i 番目のクラスタについて、パラメータ  $\theta_i$  をもつ対象の確率(密度)分布  $f_i(x|\theta_i)$  と、クラスタの混合比  $\forall \alpha_i > 0, \sum_i \alpha_i = 1$  で Table される次の混合分布を用いる手法である<sup>11)</sup>。

$$P_r[x|\theta_1, \dots, \theta_k] = \sum_i \alpha_i f_i(x|\theta_i) \quad (4)$$

### 5.2 研究者クラスタリングの結果

クラス多数別のBIC値がFig.3 である。Fig.3 より、クラスタ数を9とした場合が最も高いBIC値が得

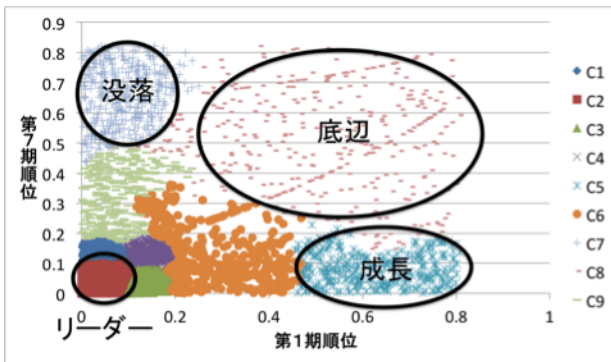


Fig.5: 分析対象とするクラスタ

Table5: 各クラスタの人数と全体に占める割合

	リーダー	成長	没落	底辺
人数	2174	319	420	380
全体に占める割合	38.4%	5.6%	7.4%	6.7%

られることがわかる。クラスタ数を9とした、混合分布モデルでのクラスタリング結果がFig.4である。

Fig.4から、第1期、第7期の両方で高順位の研究者が多く、逆に両期間で共に低順位となっている研究者は少ないことがわかる。

本研究では、中心性順位の高い研究者の特徴抽出を行うため、Fig.5の4つのクラスタに着目し、分析、比較を行った。各クラスタに属する研究者数は下記のTable5であった。

[リーダー]の研究者は、全期間で高い順位を維持しており、この領域の中心的役割を果たしている研究者の集団である。逆に、[底辺]の研究者は、低い順位で推移しており、この領域の研究に貢献しているとは言えない。[成長]の研究者は、第1期では順位が低かったものの、第2期では、高順位となっており、成長を遂げた研究者である。また、[没落]の研究者は、高順位から低順位へ推移した集団である。

## 6 クラスタの特徴抽出

ここでは、5章での研究者のクラスタリング結果を用い、各クラスタの特徴の抽出を行う。

本研究では、『媒介中心性順位の推移』について分析を行う。また、共著者ネットワーク上での位置

に影響の大きい『所属機関』、『研究分野』についても分析を行う。これは、『所属機関』では、研究成果の結果として所属機関の変更が起こり、所属機関の変更時には共著者の変化が起こるためである。また、『研究分野』については、研究分野の変更は共著者を変更することであるからである。

### 6.1各クラスタの媒介中心性順位の推移

媒介中心性の推移の分析について、分析手法と結果を記す。

#### 6.1.1手法

第1期～第7期の全7期間を対象とし、前後期間での順位の差を調査した。前後期間での順位差のうち、最も順位が上昇したものを最大上げ幅、最も順位の低下したものを最大下げ幅とし、クラスタ間での比較を行った。また、各クラスタでの第1期での順位と第7期での順位差のについても比較を行った。

#### 6.1.2結果

各クラスタの最大上げ幅と最大下げ幅の平均をまとめたものがFig.6である。また、ここでは最上げ幅と最大下げ幅について、クラスタによる分散分析と多重比較を行った。その結果がTable6からTable9である。

この結果より、[リーダー]では、順位の変動は小さく、安定していることがわかる。一方で、[成長]では、徐々に順位が上昇していくのではなく、ある期間に急激に順位の上昇が起こっていることがわかる。[没落]でも、ある期間に急激な順位の下降が起こっていることがわかる。また、[底辺]では、急激な順位の変動はあるものの、最終的には始めとほぼ同じ順位となっている。検定の結果、全クラスタ間で有意差が確認できた。

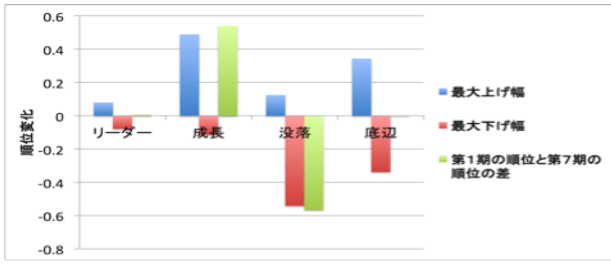


Fig.6: 各クラスターの最大上げ幅と最大下げ幅

Table6: 最大上げ幅の分散分析

	自由度	平方和	平均平方和	F 値	P 値
クラスター	3	61.88	20.626	896.7	<2E-16
誤差	3292	75.72	0.023		

Table7: 最大上げ幅の多重比較

	リーダー	成長	没落
成長	<2.0E-16	-	-
没落	2.0E-07	<2.0E-16	-
底辺	<2.0E-16	<2.0E-16	<2.0E-16

Table8: 最大下げ幅の分散分析

	自由度	平方和	平均平方和	F 値	P 値
クラスター	3	87.76	29.252	1342	<2E-16
誤差	3292	71.75	0.022		

Table9: 最大下げ幅の多重比較

	リーダー	成長	没落
成長	0.00031	-	-
没落	<2.0E-16	<2.0E-16	-
底辺	<2.0E-16	<2.0E-16	<2.0E-16

## 6.2 所属情報を用いた特徴抽出

所属情報を用いた分析について、分析手法と結果を記す。

### 6.2.1 手法

所属情報を用いた分析では、所属機関の変更の有無と、所属機関変更前後での媒介中心性順位の変動についてクラスター間での比較を行った。

所属変更前と所属変更後の順位差の取り方については、Fig. 7のように期間を定め、差の測定を行った。

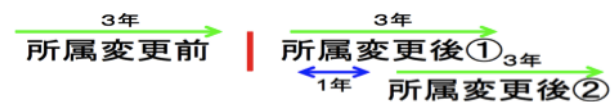


Fig.7: 所属変更期間の取り方の例

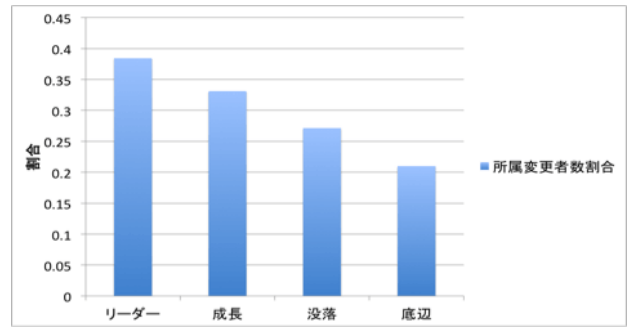


Fig.8: 所属変更者の割合

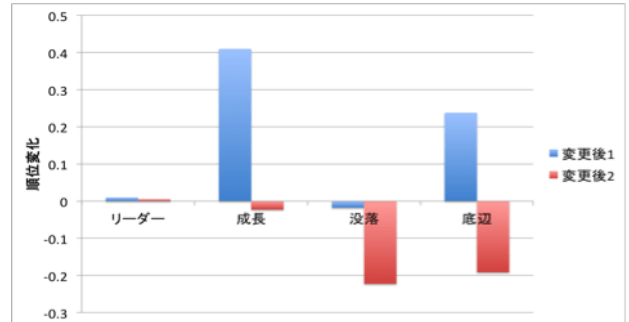


Fig.9: 所属変更と順位変化

Table10: 変更後1での分散分析

	自由度	平方和	平均平方和	F 値	P 値
クラスター	3	3.515	1.1717	34.23	<2E-16
誤差	265	9.070	0.0342		

Table11: 変更後1の多重比較

	リーダー	成長	没落
成長	2.7E-16	-	-
没落	1.000	5.3E-14	-
底辺	2.1E-05	0.039	3.4E-05

Table12: 変更後2の分散分析

	自由度	平方和	平均平方和	F 値	P 値
クラスター	3	2.04	0.68	28.64	4.9E-16
誤差	265	6.30	0.024		

Table13: 変更後2の多重比較

	リーダー	成長	没落
成長	1.0000	-	-
没落	2.4E-14	3.7E-05	-
底辺	9.6E-06	0.0085	1.0000

### 6.2.2 結果

各クラスターでの所属変更者割合がFig.8 である。[リーダー]が最も変更者割合が高く、[成長],[没



落], [底辺]の順で変更者割合が高いことがわかる. 最も高い[リーダー]と最も低い[底辺]では, 変更者割合に約2倍の差が見られる. 次に, 各クラスターの所属変更前後での媒介中心性の順位の差がFig. 9 である. また, 分散分析, 多重比較の結果がTable10からTable13である. [リーダー], [没落]では, 変更後1では順位の変動がほぼないことがわかる. 一方, [成長],

[底辺]では, 所属変更後に大きな順位の上昇があることがわかる. また, 所属変更2では, [リーダー], [成長]には大きな変化はみられず, [没落], [底辺]は順位が大きく下がっている. また, 分散分析, 多重比較の結果がTable10からTable13である. 分散分析, 多重比較の結果, 所属変更後1では, [リーダー]と[没落]間のみ有意差はなかった. 所属変更後2では, [リーダー]と[成長]間, [没落]と[底辺]間には有意差がないという結果であった.

### 6.3 研究分野情報からの特徴抽出

研究分野情報を用いた分析について, 分析手法と結果を記す.

#### 6.3.1 手法

はじめに, 計算技術利用分野に含まれる小分類に対しクラスタリングを行う. JST 学術文献データでは, 各論文に対し, 一つまたは複数の小分野分類が付与されている. ここでは, 各論文に付与された小分野分類情報を用い, 小分野の共起回数によるクラスタリングを行う. これは, 計算技術利用分野には, 23の研究分野が含まれているが, 各研究者の研究分野の動向と分析する上で, 研究者の動向を把握し易くするためである.

クラスタリング手法としては, 階層的クラスタリングを用い, ウォード法により, 距離の計算を行う. ウォード法では,  $x_i$ を分類対象,  $C_i$ を分割されたクラスター,  $c_i$ を $C_i$ のセントロイドとしたとき, クラス

ター  $C_1$  と  $C_2$  の距離関数  $D(C_1, C_2)$  を下記により式で求める.

$$D(C_1, C_2) = E(C_1 \cap C_2) - E(C_1) - E(C_2) \quad (5)$$

ただし

$$E(C_i) = \sum_{x \in C_i} (D(x, C_i))^2 \quad (6)$$

次に, 研究者が最も多く論文を提出している分野をその研究者の主分野とし, 次点の分野を副分野1, 副分野2, ...と決め, 各研究者の主野, 副分野の比率について, クラスター間の比較を行う. 主分野, 副

	分野1	分野2	分野3	分野4
研究者A	10 第一分野	5 第三分野	8 第二分野	1
研究者B	2 第三分野	0	3 第二分野	5 第一分野

Fig.10: 主分野, 副分野の決定方法

分野の決定方法についての具体例が, Fig.10 である.

Fig.10 の例では, 枠内の数値がその分野での研究者の論文である. この場合, 研究者 A の主分野は分野1, 研究者 B の主分野は分野4となる.

また, 研究者の研究パターンとして, 3つのパターンを想定し, 各パターンの割合と, 各パターンにおける主分野の変更者の割合についても比較を行う. 想定するパターンは, Table14の3つである.

Table14: 研究者の研究分野のパターン

A	常に一つの分野の研究のみを行っている研究者
B	常に複数分野の研究を行っている研究者
C	一つの分野の研究のみを行う期間と複数分野の研究を行う期間が混在する研究者である

#### 6.3.2 結果

研究分野のクラスタリング結果がFig. 11, Table15 である.

Fig.11 から, 計算機技術利用分野は大きく分けて5つにわけられることがわかる. しかし, クラスター1に関しては, 本研究で分析対象としている期間では一度も付与されていない分野であるため, クラスター2から5の4分野として以降の分析を行った.

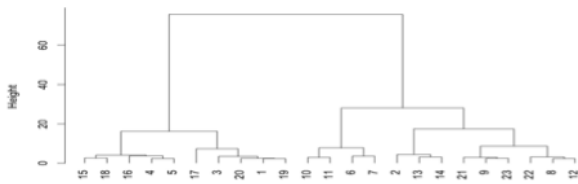


Fig.11: 研究分野のデンドログラム

Table15: 分野クラスタの内訳

クラスター	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23
1	15	18	16	4	5	17	3	1	19	11	10	6	7	2	13	14	21	9	23	22	8	12	

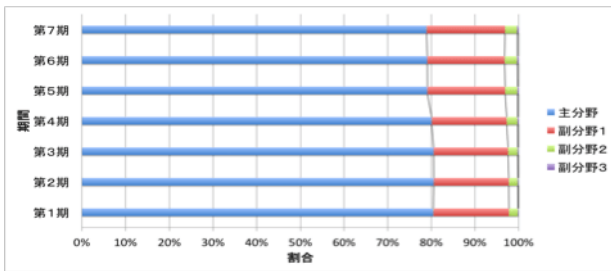


Fig.12: リーダーの分野推移

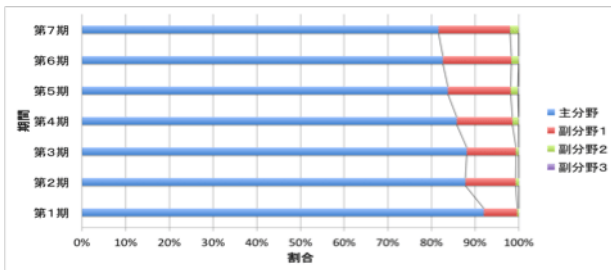


Fig.13: 成長の分野推移

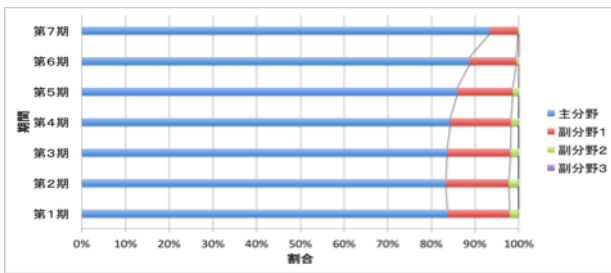


Fig.14: 没落の分野推移

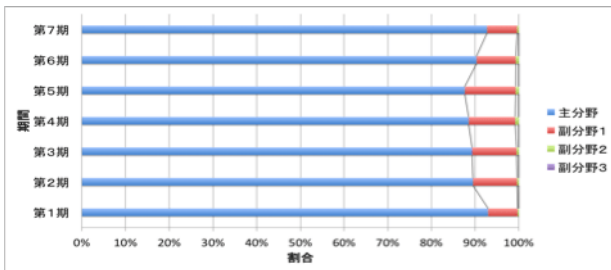


Fig.15: 底辺の分野推移

各期間での研究者の主分野, 副分野 1, 副分野 2, 副分野 3 の割合が Fig. 12 から Fig. 15 である.

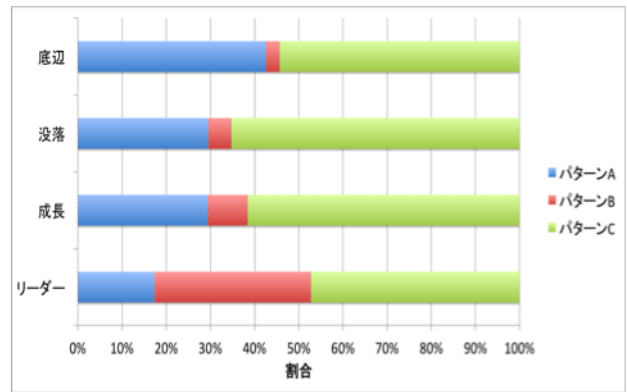


Fig.16: 各クラスタにおけるパターンの割合

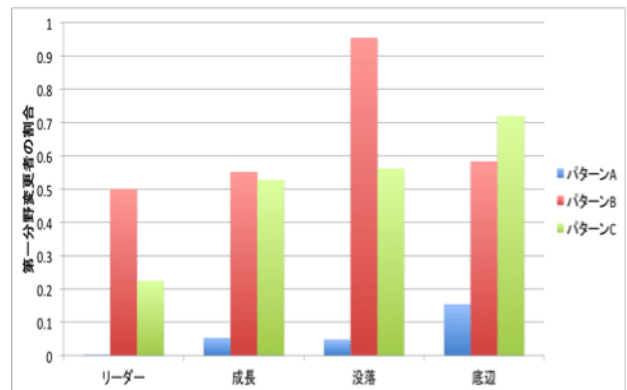


Fig.17: 各パターンでの主分野変更者割合

Fig.12-15 より, [リーダー]では全期間で主分野が約 80%, 副分野 1 が約 20%となっている. [成長]では第 1 期では主分野の割合が 90%以上あるが, 期間が進むにつれ副分野 1 の割合が大きくなり, 第 7 期では, 主分野と副分野 1 の割合が[リーダー]と同程度となっている. 一方, [没落]では, 第 1 期では, 主分野の割合が 80%程度であるが, 第 7 期では, 90%以上となっている. また, [底辺]では, どの期間も主分野の割合が 90%前後で推移している.

次に, 研究者の研究パターンとして想定した 3 パターンの割合をクラスタ別にまとめたものが Fig.16 である. Fig.16 から, 全体では, パターン C ののが最も多いことがわかる. また, [リーダー]ではパターン A ののが最も少なく, パターン B の研究者が多い. [底辺]は, パターン A が他のクラスタに比べ多いことがわかる.

各パターンの研究者のうち, 主分野を変更した者



の割合が Fig.17 である。[リーダー]では、パターン A で主分野を変更したものはほぼないことがわかる。また、パターン C での変更者の割合が他のクラスタに比べ小さい。[没落]では、パターン B での変更者の割合が大きく、9割以上が分野の変更を行っている。[底辺]では、パターン A での変更者の割合が他のクラスタに比べ大きいことがわかる。

#### 6.46 章のまとめ

6章では、各研究者クラスタに対し、媒介中心性の順位推移、所属、分野の分析を行った。

順位推移に対する分析では、クラスタ間に最大上げ幅と最大下げ幅の比較を行い、クラスタ間での差が確認できた。[リーダー]クラスタは、順位の変動が小さく、安定している。[成長]、[没落]のクラスタでは、ある期間に順位の急激な変化が起こっていることがわかった。また、[底辺]では、ある期間に順位の上昇があるが、同様に、急激な下降もあり、最終的な順位は変わらないことが明らかとなった。

所属に対する分析では、所属変更前後での順位変化の比較を行い、所属変更前後での順位の変化にクラスタ間での差が確認できた。[リーダー]では所属変更による順位変化はほぼ見られなかった。[成長]、[底辺]では、所属変更後1で順位の上昇が確認できた。また、[没落]、[底辺]では、所属変更2で順位の上昇が見られた。

分野に対する分析では、分野の偏り、3つの研究パターンによる比較を行い、各クラスタ間の差が確認できた。[リーダー]では継続して副分野1を持っている研究者が多く、また、[リーダー]の一つの分野のみの研究のみを行う期間と複数分野の研究を行う期間が混在する研究者は、主分野の変更者が少ないことがわかった。[没落]では、継続して副分野1を持っている研究者の多くが主分野を変更していることがわかった。また、[底辺]では、他のクラスタに

比べ、一つの分野のみの研究のみを行う研究者が主分野の変更を行っていることが確認できた。

## 7 まとめ

本研究では、JST 学術文献データと科研費情報を使用し、計算機利用技術分野を対象に、共著者ネットワーク上での中心性と、科研費の取得率との関係について分析を行い、その関係を明らかにした。さらに、共著者ネットワーク上優位な位置に存在する研究者の特徴について、中心性指標での順位、所属、研究分野情報を用いた分析を行い、特徴を抽出した。得られた研究者の特徴は以下である。

### a) リーダー

- ・順位の上昇が急激な変動がほとんどない
- ・所属変更が多く、初速変更時の変化が小さい
- ・分野変更者が少ない

### b) 成長

- ・短期間に急激な順位の上昇がある。
- ・所属変更直後に順位の上昇がある。

### c) 没落

- ・短期間に急激な順位の上昇がある。
- ・所属変更後の期間で順位の上昇がある。
- ・主分野変更者が多い。特に、一分野の研究のみを行っている研究者について、他のクラスタとの差が大きい。

### d) 底辺

- ・ある程度大きな順位の上昇があるが、同程度の下降もある。
- ・所属変更直後に順位の上昇があるが、その後の期間で順位の上昇がある。

最後に、今後の展望を述べる。本研究では、共著者ネットワーク、研究分野、所属、科研費の情報を活用し、計算機技術利用分野の研究者を対象とした分析を行った。そのため、今後さらに広い分析が可

能である。いかにそれらを列挙する。

#### 1) 対象範囲の拡張

本分析では、計算機技術利用分野のみを対象に分析を行ったため、今後、分野や国など幅広い領域を対象にした分析が期待できる。

#### 2) 研究成果の考慮

本分析では、単リンクのみの無向グラフとして共著者ネットワークの作成を行い、分析を行った。論文の被引用数による重み付きリンクや、引用関係でのネットワークでの分析が今後考えられる。

#### 3) リンク予測による

本分析では新規の共著者による分析では、クラスター間に有意差が見られなかったが、研究者ネットワーク上で将来重要となる研究者との共著関係についての分析が考えられる<sup>12)</sup>。

#### 謝辞

本研究を進めるに当たり、JST学術文献データとして「2013年度データサイエンス・アドベンチャー杯」で提供していただきましたJST関係者の皆様に深く感謝し、お礼申し上げます。

#### 参考文献

- 1) 西澤正己, 孫媛: キーワード分析による科研費におけるゲノムおよびナノテクノロジー関連研究の動向調査, 情報知識学会誌, **17-2**, 117/122 (2007)
- 2) 浅田洋平, et al.: Web からの研究者ネットワーク抽出の大規模化, 人工知能学会論文誌, **20-6**, 370/378 (2005)
- 3) 松尾豊, et al.: 効果的な共同研究を支援するための

Web からの研究者ネットワーク抽出: 情報処理学会, 2005 年情報学シンポジウム (2005)

- 4) 西澤正己, et al.: 科学研究費データベースによる採択と研究成果のパターン分析, 情報知識学会誌, **16-2**, 1/6 (2006)
- 5) 柿沼澄男, et al.: 科学研究費補助金による研究助成の効果に関する調査—研究キャリアステージにおける科学研究費補助金取得パターンの分析—, 情報知識学会誌, **17-2**, 111/116 (2007)
- 6) 篠田孝祐: 日本における人工知能研究の系譜, 人工知能学会誌, **26-6** (2011)
- 7) Newman, Mark EJ: Coauthorship networks and patterns of scientific collaboration., Proceedings of the National Academy of Sciences 101. suppl 1, 5200/5205 (2004)
- 8) 内藤理, et al.: 日本におけるロボット工学の研究者ネットワークの分析, 日本ロボット学会誌, **30-6**, 629/638 (2012)
- 9) 安田雪, et al.: 人工知能学会におけるネットワーク構造と変化, 人工知能学会全国大会論文集, 1F2-1 (2006)
- 10) 安田雪, 松尾豊: 人工知能学会における研究者ネットワークの分析, 第 19 回人工知能学会全国大会 (JSAI 2005), 2A3-02 (2005)
- 11) 神島敏弘: データマイニング分野のクラスタリング手法 (1), 人工知能学会誌 **18-1** (2003).
- 12) 宮西大樹, et al.: リンク予測を基にした時系列ネットワーク中でのオブジェクトランキング, 人工知能学会論文誌, **27-3**, 223/234 (2012)