

科研費における応募細目の変遷による 細目間の関係抽出とその予測

山下 長義^{†1} 沼尾 正行^{†2} 市瀬 龍太郎^{†3}

本論文では学術の動向を把握するために科研費における応募細目の変遷を用いて研究領域ネットワークを抽出する手法と、基盤研究・若手研究と挑戦的萌芽研究との細目別重複応募状況によってその経年変化を予測する手法を提案する。その結果、科研費の応募細目の変遷による研究領域ネットワークによって学術の動向の現状を把握し、挑戦的萌芽研究との重複応募研究者数によって高い適合率で研究者の移動が増加する細目対を予測でき提案手法の有効性を示すことができた。

Relation Extraction of Research Fields Using Proposal Transitions of Research Fields in Research Grants and Its Prediction

NAGAYOSHI YAMASHITA,^{†1} MASAYUKI NUMAO^{†2}
and RYUTARO ICHISE^{†3}

This paper proposes methods for comprehending the trend of science using research-fields proposal transitions networks and predicting the change of the network using pairs of research fields proposed by the same researchers to both “Scientific Research” and “Challenging Exploratory Research”. The experimental results showed these networks enable to comprehend trends of science and changes of these networks are forecasted by research-fields proposed by the same researchers to two types of grant with high precision.

1. はじめに

数学や哲学などの古典的な学問から環境学やゲノム科学などの幅広い分野をカバーする学問が存在している現在、学術の動向を把握することは重要になりつつある。科学における知の発見は、既存の研究の新たな組み合わせから生み出されることが多いからである。しかし、学問が細分化されているため、学術の動向を把握することは難しいのが現状である。そこで、本論文では学術の動向を把握するために科学研究費補助金（以下では科研費とする）における応募細目の変遷によって細目対の関係を抽出し、これら細目対の関係から研究領域ネットワークを抽出する手法を提案する。科研費では研究者自らの研究計画の内容から審査を希望する細目を選択しなければならず、審査を希望する細目が前回と異なるのは前回の応募時から研究分野が変わったり同じ研究内容のまま関連する他の細目に応募したりしたためであり、ある時期に特定の細目対において応募する細目を変更する研究者が増加すれば、その時期に研究のトレンドが変化したと考えられる。よって、応募細目の変遷は学術の動向を示していると考えられる。

さらに自らの関連細目に応募している研究者が次にどの分野の研究を行なおうとしているかなどの学術の動向を予測できれば、将来の研究方針の決定に役立つことが期待される。そこで、基盤研究・若手研究と挑戦的萌芽研究（萌芽研究を含む）との重複応募状況によってその経年変化を予測する手法を提案する。挑戦的萌芽研究は「独創的な発想に基づく、挑戦的で高い目標設定を掲げた芽生え期の研究」を行う研究種目であり、基盤研究（S）（A）（B）と若手研究（A）との重複応募が可能で、基盤・若手研究により助成を受けながら挑戦的萌芽研究にも応募することができるため、応募する細目が基盤研究・若手研究と挑戦的萌芽研究とで異なる場合、基盤研究・若手研究で応募した研究分野をもとに、挑戦的萌芽で応募した研究分野で「独創的な発想に基づく、挑戦的で高い目標設定を掲げた芽生え期の研究」を行っていると考えられる。よって、基盤研究・若手研究と挑戦的萌芽研究の重複応募状況は応募細目の遷移の先行指標となり、これによって研究領域ネットワークの経年変化を予測できると考えられる。

^{†1} 日本学術振興会 研究事業部

Research Program Department, Japan Society for the Promotion of Science

^{†2} 大阪大学 産業科学研究所

The Institute of Scientific and Industrial Research, Osaka University

^{†3} 国立情報学研究所 情報学プリンシプル研究系

Principles of Informatics Research Division, National Institute of Informatics

本論文の構成を以下に示す。2章では関連研究について述べ、3章では科研費について述べ、4章では応募細目の変遷による細目対の関係を抽出する手法を説明し、5章では挑戦的萌芽研究との重複応募状況による応募細目の変遷の予測手法について、6章ではまとめとこれからの課題について説明する。

2. 関連研究

これまで学術の動向を調査する研究のほとんどは、論文における情報を用いたものであった。代表的なものに論文の共著関係を用いた研究¹⁾、論文の文献情報を用いた研究²⁾や文献キーワードを利用した研究³⁾などがある。しかし、論文のデータは様式が一様ではなく、さまざまな分野の論文を揃えることが難しいといった問題がある。学術に関するデータは論文の他に科研費の応募データがある。科研費の応募データの様式は一様で、人文学・社会科学から医学まで幅広い分野の研究者からの応募があり、多くの研究者が毎年継続して応募しているため年度ごとの変化を調査するのに適している。また、それぞれの応募課題に対して分野、分科、細目などの研究領域やキーワード、研究者の所属機関などの属性が付加されている。科研費などのグラントの応募採択情報を用いた研究は、科研費における研究代表者と研究分担者の関係によって研究領域ネットワークを抽出しその経年変化の予測する研究⁴⁾や、NIH（アメリカ国立衛生研究所）の採択課題のタイトルやアブストラクトから抽出した関係を可視化しインタラクティブに検索できるシステムを提案している論文⁵⁾がある。

佐藤らは研究代表者と研究分担者の関係による研究領域ネットワークの分析を行ったが、本論文では応募細目の変遷による研究領域ネットワークを構築する手法とその経年変化を基盤研究・若手研究と挑戦的萌芽研究との細目別重複応募状況によって予測する手法を提案する。

3. 科研費

文部科学省および日本学術振興会によって公募、審査、交付される科研費は、日本国の政府から提供される助成金の中でもっとも一般的なものである。人文・社会科学から自然科学までの全ての分野にわたり、基礎から応用まであらゆる「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「競争的資金」であり、ピア・レビュー（専門分野に近い複数の研究者による審査）により、豊かな社会発展の基盤となる独創的・先駆的な研究に対する助成を行うものである⁶⁾。

科研費の応募データは、応募課題の研究領域を表す細目やキーワード、応募研究者の所属

や研究分担者、研究種目などの様々な属性が付加されている。応募課題の研究領域は、応募に際して研究計画の内容に照らし、審査希望分野を示す分類表である「系・分野・分科・細目表」から適切な細目を応募研究者によって選定されるとともに、最も関連が深いと思われるキーワードを「キーワード一覧」より選定される。主な研究種目には基盤研究（S）・（A）・（B）・（C）、挑戦的萌芽研究、若手研究（A）・（B）があり、基盤研究は独創的・先駆的な研究を助成の対象とし、挑戦的萌芽研究は独創的な発想に基づく挑戦的で高い目標設定を掲げた芽生え期の研究を助成対象とし、若手研究は39歳以下の研究者が1人で行う研究を対象に助成を行っている。ただし、挑戦的萌芽研究は基盤研究（S）・（A）・（B）、若手研究（A）と重複して助成を受けることが可能である⁶⁾。本論文では、平成15年度から平成22年度までの基盤研究（海外学術調査を除く）・若手研究・挑戦的萌芽研究における応募課題の情報を用いる。

4. 応募細目の変遷による細目対の関係の抽出

科研費では研究者自らの研究計画の内容から審査を希望する細目を選択しなければならない。審査を希望する細目が前回と異なるのは、前回の応募時から研究分野が変わったり同じ研究内容のまま関連する他の細目に応募したりしたためであり、ある時期に特定の細目対において応募する細目を変更する研究者が増加すれば、その時期に研究のトレンドが変化したと考えられる。よって応募細目の変遷は学術の動向を示していると考えられる。そこで学術動向の現状を把握するために、応募細目の変遷（細目対における研究者の移動）による細目対の関係と研究領域ネットワークを抽出する手法を提案する。4.1節にて細目対の関係と研究領域ネットワークの抽出法を説明し、4.2節にて研究領域ネットワークによって細目全体の関係による評価を行ない、移動研究者数の増分が大きい細目対を取り上ることで細目別の評価を行う。

4.1 細目対の関係と研究領域ネットワークの抽出法

平成 $n-1$ 年度に細目 X に応募した研究者集合を $R_x(n-1)$ 、平成 n 年度に細目 Y に応募した研究者集合を $R_y(n)$ として、平成 n 年度における細目 X から細目 Y への移動研究者数 $M_{xy}(n)$ を以下の式により定義する。

$$M_{xy}(n) = |R_x(n-1) \cap R_y(n)| \quad (1)$$

ただし同一課題の継続分を除外するために、平成 n 年度における研究者集合は新規応募から抽出し、平成 $n-1$ 年度における研究者集合は新規と継続課題から抽出する。たとえば、平成22年度におけるメディア情報学・データベースAから知能情報学への移動研究者数は、

平成 21 年度にメディア情報学・データベース A へ応募し、平成 22 年度には知能情報学へ応募した研究者数となる。

細目対における移動研究者の関係により形成される研究領域ネットワークは、ノードを各細目とし、 α 人以上の研究者の移動があればリンクでノード間をつなぐことによって形成される。ノードの大きさは 8 年間の応募件数の平均に比例している (図 1)。本論文では α を 3 とし、情報学の細目と情報学の細目と研究者の移動がある細目により研究領域ネットワークを形成する。

4.2 評価

細目全体の関係による評価を行うために平成 16 年度から平成 22 年度まで 1 年度ごとに研究領域ネットワークを抽出した。このようにして得られた研究領域ネットワークから、メディア情報学・データベース A では入次数が年々増加して特に知覚情報処理・知能ロボティクス A から研究者の転入が増加傾向にあり、知能情報学では平成 21 年度に医用システムへの移動研究者数や言語学からの移動研究者数が増加し、平成 22 年度には社会システム工学・安全システム A からの移動研究者数が増加していることがわかる。また教育学では平成 21 年度までメディア情報学・データベース A からの移動研究者数が増加傾向にあり、リハビリテーション科学・福祉科学 B では当初メディア情報学・データベース B、知覚情報処理・知能ロボティクス A との間で研究者の移動があったが、平成 21 年度には知覚情報処理・知能ロボティクス B からの移動研究者が増加していることがわかる。

次に細目別に評価を行うために移動研究者数の 1 年あたりの増分が大きい細目対を取り上げる。移動研究者数の 1 年度あたりの増分は、基準年度から 3 年間の移動研究者数を最小自乗法により近似した直線の傾きとする。平成 18 年度からの計算機システム・ネットワーク B からソフトウェアへと、平成 17 年度からの認知科学から実験心理学への移動研究者数の 1 年度あたりの増分がともに 6.5 と最も高いことがわかる。また、情報学の細目と他分野の細目との関係では、知覚情報処理・知能ロボティクス B からリハビリテーション科学・福祉工学 B への移動研究者数と、感性情報学・ソフトウェア A から生活科学一般 B への移動研究者数が平成 19 年度から増加し、またメディア情報学・データベース A からリハビリテーション科学・福祉工学 B への移動研究者数が平成 20 年度から増加傾向にあることがわかる。

5. 重複応募状況による応募細目の変遷の予測

科研費の研究種目には「独創的な発想に基づく、挑戦的で高い目標設定を掲げた芽生え期

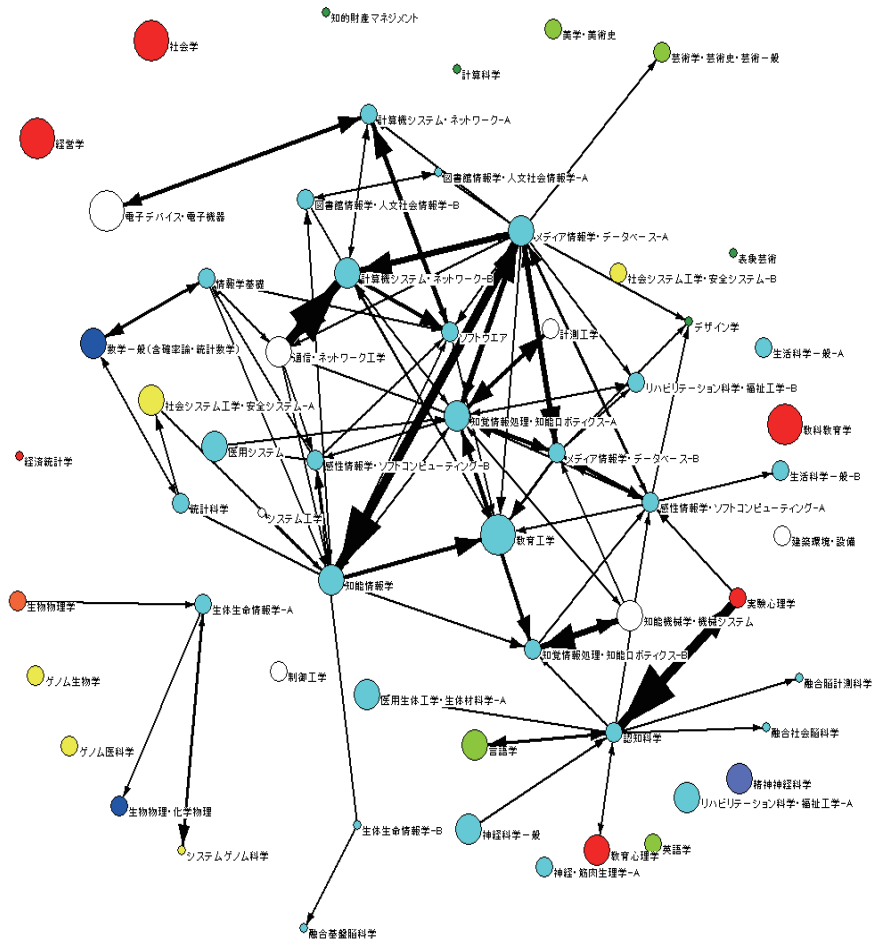


図 1 研究者の移動による研究領域ネットワーク (平成 22 年度)
Fig.1 A research field network by researchers transition (2010)

の研究」を行う挑戦的萌芽研究がある。挑戦的萌芽研究は、基盤研究 (S) (A) (B) と若手研究 (A) との重複応募が可能で、基盤・若手研究により助成を受けながら挑戦的萌芽研究にも応募することができる。したがって、応募する細目が基盤研究・若手研究と挑戦的萌芽研究とで異なる場合、基盤研究・若手研究で応募した研究分野をもとに、挑戦的萌芽で応募した研究分野で「独創的な発想に基づく、挑戦的で高い目標設定を掲げた芽生え期の研究」を行っているため、基盤研究・若手研究と挑戦的萌芽研究との重複応募状況は応募細目の変遷の年次推移の先行指標となると考えられる。また応募細目の変遷による研究領域ネットワークにおいて、共通の隣接ノード (共参照ノード) を多く持っているほど2つのノード間にリンクが現れやすいと考えられる。そこで、挑戦的萌芽研究への重複応募状況と共参照ノード数によって応募細目の変遷の経年変化を予測する手法を提案する。5.1 節では重複応募研究者数と共参照ノード数を求める方法を説明し、5.2 節で移動研究者数が増加する細目対を予測し、いくつかの細目対において重複応募研究者数と共参照ノード数が移動研究者数の先行指標となっているかを調査する。

5.1 重複応募研究者数と共参照ノード数

平成 n 年度において基盤研究の細目 X に応募した研究者集合を $R_x(n)$ 、挑戦的萌芽研究の細目 Y に応募した研究者集合を $CH_y(n)$ とすると、平成 n 年度に基盤・若手研究において細目 X に応募し挑戦的萌芽研究において細目 Y に応募した重複応募研究者数 $CH_{xy}(n)$ は、以下の式により求まる。

$$CH_{xy}(n) = |R_x(n) \cap CH_y(n)| \quad (2)$$

ただし $R_x(n)$ 、 $CH_y(n)$ とともに新規・継続課題を用いる。たとえば、平成 22 年度における知能情報学からメディア情報学・データベース A への重複応募研究者数は、平成 22 年度において基盤・若手研究の知能情報学に応募し、挑戦的萌芽研究のメディア情報学・データベース A へ応募している研究者数となる。さらに、累積重複応募数 $CHC_{xy}(n)$ を以下の式により求める。

$$CHC_{xy}(n) = CH_{xy}(n) + CH_{xy}(n-1) + CH_{xy}(n-2) \quad (3)$$

また平成 n 年度におけるノード x とノード y に対する共参照ノード数 $CO_{xy}(n)$ は、 $\Gamma_x(n)$ を n 年度にノード x を参照しているノード集合とすると以下の式により求まる。

$$CO_{xy}(n) = |\Gamma_x(n) \cap \Gamma_y(n)| \quad (4)$$

5.2 評価実験

情報学の細目から研究領域ネットワーク上に存在するすべての細目において移動研究者数が増加する細目対を予測し (5.2.1 節)、重複応募研究者数と共参照ノード数が移動研究者

表 1 平成 22 年度において累積重複応募数と隣接ノード数の合計が 8 以上の細目対

Table 1 Pairs of research fields with 8 or more total of the cumulative number of proposals on duplication and the number of nodes linking to both of a pair of nodes in 2010

基盤・若手研究での応募細目	挑戦的萌芽研究での応募細目
知覚情報処理・知能ロボティクス A	知覚情報処理・知能ロボティクス B
知覚情報処理・知能ロボティクス A	メディア情報学・データベース B
知覚情報処理・知能ロボティクス B	知覚情報処理・知能ロボティクス A

数の先行指標となっているかを調査するために、いくつかの細目対において移動研究者数、重複応募研究者数、共参照ノード数の年次推移による評価を行う (5.2.2 節)。

5.2.1 移動研究者数が増加する細目対の予測

細目対 (x, y) における移動研究者数 $M_{xy}(n)$ の 1 年あたりの増分が 1 以上の細目対を正例、1 未満の細目対を負例とし、これらクラスラベルと累積重複応募数と共参照ノード数との関係を調査した。ただし、移動研究者数がすでに 10 以上の細目対は、多数の研究者がすでに細目対を移動しているため除外した。平成 17 年度と平成 20 年度において評価を行った結果、累積重複応募数が 7 人以上の細目対は 9 対、適合率は 0.67 であり (図 2)、累積重複応募数と共参照ノード数の合計が 8 以上の細目対は 14 対、適合率は 0.71 であり提案手法が有効であることが示された (図 3)。

平成 22 年度における累積重複応募数と共参照ノード数の合計が 8 以上の細目対を表 1 に示す。平成 17 年度と平成 20 年度でこの条件を満たす細目対の移動研究者数は 71% の確率で増加していたため、これらの細目対でも将来研究者の移動が活発になることが予想される。

5.2.2 移動研究者数、重複応募研究者数、共参照ノード数の年次推移による評価

知覚情報処理・知能ロボティクス A (x_1) から知能情報学 (y_1) への関係 (図 4) と、メディア情報学・データベース A (x_2) から知能情報学 (y_2) への関係 (図 5) を移動研究者数 ($M_{xy}(n)$)、重複応募研究者数 ($CH_{xy}(n)$)、共参照ノード数 ($CO_{xy}(n)$) によって示す。ただし、 $(x, y) = (x_1, y_1), (x_2, y_2)$ 、 $15 \leq n \leq 22$ とする。

知覚情報処理・知能ロボティクス A から知能情報学への移動研究者数は平成 19 年度と平成 22 年度において極大となり、重複応募研究者数は平成 15 年度から平成 18 年度で 2 人以上と高い値を示し、共参照ノード数は平成 20 年度と平成 21 年度において極大となっている (図 4)。また、知覚情報処理・知能ロボティクス A のキーワードにおいて平成 19 年度に 3 名が画像情報処理を選択、2 名がセンサー融合・統合を選択、1 名がパターン認識を選択している一方、平成 22 年度では 5 名がパターン認識を選択、2 名が画像情報処理を選択

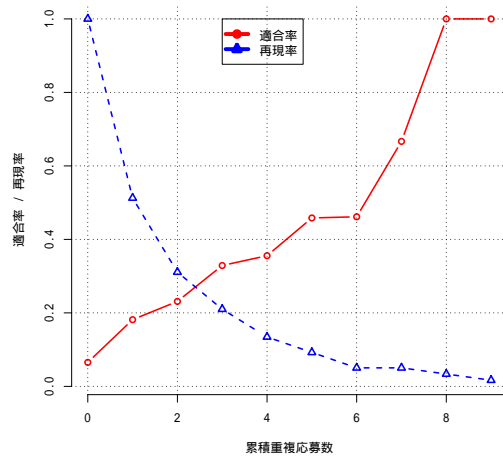


図 2 累積重複応募数と適合率・再現率

Fig. 2 The cumulative number of proposals on duplication and precision, recall

した。このように、知覚情報処理・知能ロボティクス A から知能情報学へ移動している研究者の中で、知覚情報処理・知能ロボティクス A のキーワードにおいてパターン認識を選択した研究者の占める割合が近年増加傾向にあることがわかる。

またメディア情報学・データベース A から知能情報学への移動研究者数は平成 19 年度と平成 22 年度において極大となり、重複応募研究者数は平成 21 年度、共参照ノード数は平成 18 年度と平成 21 年度において極大となっている (図 5)。上記の 2 つの例において移動研究者における極大となる年の前年に重複応募研究者数または共参照ノード数が極大となり、重複応募研究者数と共参照ノード数が移動研究者数の先行指標となっていることから提案手法が有効であることが示された。

6. まとめとこれからの課題

研究者の応募細目の変遷を調査することによって科学研究の動向を調査する手法と、基盤研究・若手研究と挑戦的萌芽研究における重複応募状況によってその経年変化を予測する手法を提案した。その結果、科研費の応募細目の変遷による研究領域ネットワークによって学

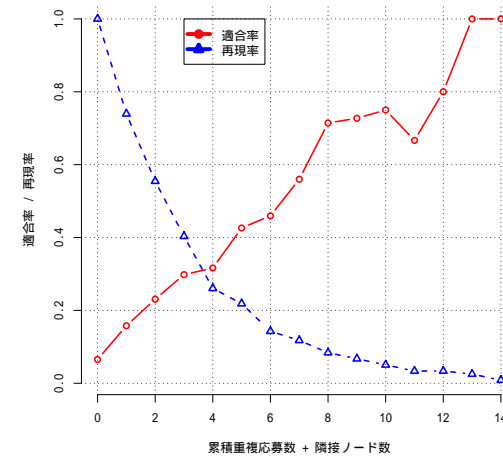


図 3 共参照ノード数と累積重複応募数の合計と適合率・再現率

Fig. 3 A total of the cumulative number of proposals on duplication and the number of nodes linking to both of a pair of nodes and precision, recall

術の動向の現状を把握することが可能であることを示し、挑戦的萌芽研究の重複応募状況によって高い適合率で研究者の移動が増加する細目対を予測することによって提案手法の有効性を示した。

これからの課題は、本論文で用いたデータに加え研究代表者と研究分担者の関係を属性として用いて、決定木等の学習手法によって細目対の研究者の移動の予測を行うことである。また、細目対を移動する研究者数が急激に増加しているなど大きな変化が起こっている 2 つの細目において、研究代表者と研究分担者による研究者ネットワークの経年変化を調査することがある。

謝辞 本研究を進めるにあたり貴重なご意見をいただいた日本学術振興会 学術システム研究センターの研究員の先生方に感謝いたします。

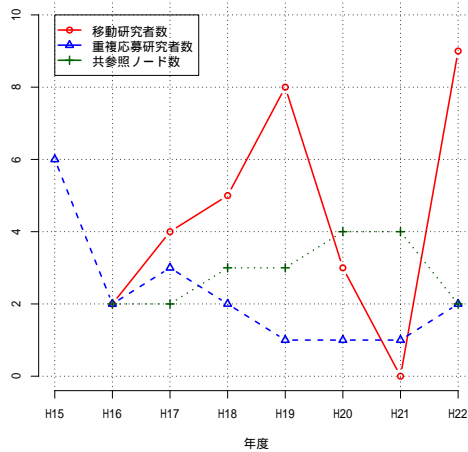


図 4 知覚情報処理・知能ロボティクス A から知能情報学への移動研究者数、重複応募研究者数と隣接ノード数の年次推移

Fig. 4 Secular change of the number of moving researchers, the number of proposals on duplication and the number of nodes linking to both of a pair of nodes from Perception information processing/Intelligent robotic A to Intelligent informatics

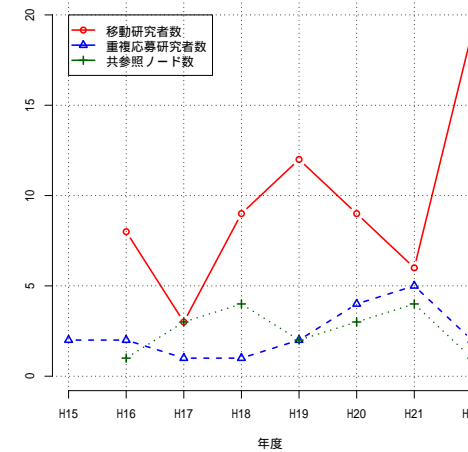


図 5 メディア情報学・データベース A から知能情報学への移動研究者数、重複応募研究者数と隣接ノード数の年次推移

Fig. 5 Secular change of the number of moving researchers, the number of proposals on duplication and the number of nodes linking to both of a pair of nodes from Media informatics/Database A to Intelligent informatics

参考文献

- 1) Newman, M. E.J.: Coauthorship Networks and Patterns of Scientific Collaboration, *Proceedings of the National Academy of Sciences*, pp.5200–5205 (2004).
- 2) 阪 彩香, 伊神正貫, 桑原輝隆: サイエンスマップ 2008 -論文データベース分析 (2003年から2008年)による注目される研究領域の動向調査, *NISTEP REPORT No.110* (2010).
- 3) 奥岡晋大, 片上大輔, 新田克己: 研究分野ネットワークによる学術分野動向の解析と評価, 人工知能学会知識ベースシステム研究会資料, Vol.84, pp.21–26 (2009).
- 4) 佐藤和宏, 市瀬龍太郎, 栗原 聡, 相澤彰子, 沼尾正行: 科学研究費申請データの解析, 第25回ファジィシステムシンポジウム (2009).
- 5) II, B. W.H., Talley, E.M., Burns, G.A., Newman, D. and LaRowe, G.: The NIH Visual Browser: An Interactive Visualization of Biomedical Research, *13th International Conference Information Visualisation* (2009).

- 6) 独立行政法人日本学術振興会: 平成 23 年度科学研究費補助金公募要領 (特別推進研究, 基盤研究, 挑戦的萌芽研究, 若手研究 (A・B)) (2010).