

# ネットワーク構造分析を用いた投資信託の分析

松井藤五郎<sup>1\*</sup> 武田祐太<sup>2</sup> 犬塚信博<sup>2</sup>

<sup>1</sup> 中部大学 <sup>2</sup> 名古屋工業大学

**Abstract:** 本論文では、投資信託の組入比率上位 10 銘柄に基づいて、投資信託と投資先企業を頂点とした 2 部グラフを作成し、これをネットワーク構造分析の技術を用いて分析する方法を提案する。提案手法では、まず、Web ページのリンク関係を表すネットワーク構造から重要度が高い Web ページを抽出するために開発された HITS (Hypertext Induced Topic Selection) アルゴリズムを用いて投資信託と投資先企業間の投資関係を表すネットワーク構造を分析し、独自性が高い投資信託を抽出する。その後、投資信託を隣接ベクトルで表現し、ユークリッド距離に基づく  $k$  平均法を用いてクラスタリングを行うことによって投資信託を  $k$  個のグループに分ける。  $k = 5$  として分析した結果、J-REIT、日経 225、TOPIX、小型アクティブ、大型アクティブに分けることができた。

## 1 はじめに

投資信託は、一般の投資家から集めた資金をまとめて複数の株式や債券などに投資する金融商品である。投資信託には、販売会社、運用会社、受託銀行が関わっている。販売会社は、一般の投資家に対して投資信託を販売する。運用会社は、販売会社が集めた資金の投資先を決定する。受託銀行は、運用会社の決定に基づいて株式等の売買を行い、資産を管理する。

投資信託の銘柄は数百にも及び、一般の投資家はどの投資信託を選択すれば良いのか選ぶのが難しい。投資信託の多くには、「グロース」「バリュー」「高配当」「小型株」「不動産」など、運用方針を示すような名称が付けられている。また、投資信託には目論見書と呼ばれる費用や運用方針等を記した説明書の交付が義務づけられており、目論見書には一般社団法人投資信託協会が定めた運用対象に基づく商品分類と属性区分が記載されている。しかしながら、実際には設立時に付けられた名称や商品分類とは実体が異なる運用を行っている投資信託もあると考えられる。

投資信託が実際にどのような運用を行っているかを分析するには、実際の投資先を調べればよい。そこで、本論文では、投資信託の投資先情報に基づいて投資信託をいくつかのグループに分ける方法を提案する。具体的には、投資信託の組入比率上位 10 銘柄に基づいて、投資信託と投資先企業を頂点、投資関係を辺とした 2 部グラフを作成し、これをネットワーク構造分析の技術を用い

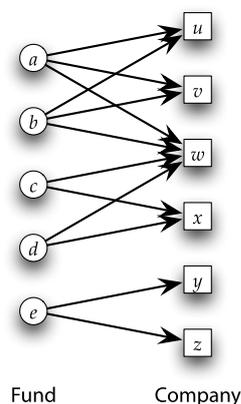


図 1 投資信託と投資先企業の 2 部グラフ

て分析する方法を提案する。

## 2 投資信託と投資先企業間のネットワーク構造

投資信託と投資先企業を頂点、投資関係を辺として表すと、2 部グラフ (同じグループに属する頂点間に辺がない 2 つのグループに頂点を分けることができるグラフ) となる。例を図 1 に示す。左側の頂点  $a$  から頂点  $e$  が投資信託、右側の頂点  $u$  から頂点  $z$  が投資先企業を表す。頂点  $a$  から頂点  $u$  への矢印は、投資信託  $a$  が上場企業  $u$  に投資していることを示している。

このように、投資信託と投資先企業の間をネットワークとして表すことによって、投資信託や投資先企業を特徴付けることができる。この例においては、まず、投資信託  $e$  は、他の投資信託とは投資先が異なっており、

\* 連絡先: TohgorohMatsui@tohgoroh.jp, <http://とうごろう.jp>

独自性が高いと考えられる。投資信託  $a$  と投資信託  $b$  は投資先企業が  $u, v, w$  と共通していることから、良く似た投資信託であると考えられる。同様に、投資信託  $c$  と投資信託  $d$  は、投資先企業が  $w, x$  と共通していることから、良く似た投資信託であると考えられる。また、企業側に着目すると、企業  $w$  は、多くの企業から投資されていることから、投資先として重要な企業であることがわかる。

### 3 分析手法

本論文では、投資信託を表す頂点集合をクラスタリングする手法を提案する。クラスタリングは、同じ特徴をもつデータが同じグループに属するようにデータ集合を分割する技術であり、データ・マイニングの手法の一つとしてよく用いられている。

ところが、図 1 の例における投資信託  $e$  のように、他のデータと共通する特徴を持たないデータが含まれていると、うまくグループ分けすることができない。そこで、本論文では、まずはじめに投資信託  $e$  のような独自性が高い投資信託を抽出し、取り除くことを提案する。

本論文では、独自性が高い投資信託を抽出するのに、HITS (Hypertext Induced Topic Selection) アルゴリズム [1, 2] を用いる。HITS アルゴリズムは、Web ページのリンク関係を表す 2 部グラフから重要度が高い Web ページを抽出するために開発されたものである。HITS アルゴリズムを投資信託と投資先企業間の投資関係を表す 2 部グラフに適用すると、よく投資されている企業や、よく投資されている企業に投資している投資信託を抽出することができる。本研究では、これを利用して、他の投資信託があまり投資していない企業に多く投資している投資信託を、独自性が高い投資信託として抽出する。

Algorithm 1 に本研究で用いた HITS アルゴリズムを示す。投資関係  $(f, c) \in I$  ( $f \in F, c \in C$ ) は、投資信託  $f$  が企業  $c$  に投資していることを表す。ここで、 $F$  は投資信託の集合、 $C$  は投資先企業の集合を表す。良く投資されている企業は、そのオーソリティー度  $a(c)$  が高くなり、よく投資されている企業に投資している投資信託は、そのハブ度  $h(f)$  が高くなる。そこで、このアルゴリズムを用いてすべての投資信託  $f \in F$  についてハブ度  $h(f)$  を求め、ハブ度が小さいものを独自性が高い投資信託として抽出する。

次に、投資信託を隣接ベクトルで表現し、ユークリッド距離に基づく  $k$  平均法 [3, 4] を用いてクラスタリングを行うことによって投資信託を  $k$  個のグループに分け

---

#### Algorithm 1 投資信託と投資先企業間の投資関係ネットワークに対する HITS アルゴリズム

---

入力: 投資信託集合  $F$ , 投資先企業集合  $C$ , 投資関係集合  $I$ , ステップ数  $n$   
 出力: ハブ度  $h$ , オーソリティー度  $a$   
**For all**  $f \in F$  **do**  $h(f) \leftarrow 1$   
**For all**  $c \in C$  **do**  $a(c) \leftarrow 1$   
**loop** (以下を  $n$  回繰り返す)  
     **For all**  $c \in C$  **do**  $a(c) \leftarrow \sum_{(f,c) \in I} h(f)$   
      $N \leftarrow \sqrt{\sum_{c \in C} a(c)^2}$   
     **For all**  $c \in C$  **do**  $a(c) \leftarrow \frac{a(c)}{N}$   
     **For all**  $f \in F$  **do**  $h(f) \leftarrow \sum_{(f,c) \in I} a(c)$   
      $N \leftarrow \sqrt{\sum_{f \in F} h(f)^2}$   
     **For all**  $f \in F$  **do**  $h(f) \leftarrow \frac{h(f)}{N}$   
**end loop**

---



---

#### Algorithm 2 投資信託集合に対する $k$ 平均法

---

入力: 投資信託集合  $F$ , クラスタ数  $k$   
 出力: クラスタ集合  $Y$   
**for all**  $f \in F$  **do**  
      $f$  を  $y_1, \dots, y_k \in Y$  のいずれかにランダムに割り当てる  
**end for**  
**repeat**  
     **for all**  $y_i \in Y$  **do**  
          $\mu_i \leftarrow y_i$  に属する投資信託の隣接ベクトルの平均  
     **end for**  
     **for all**  $f \in F$  **do**  
          $\mu_1, \dots, \mu_k$  とのユークリッド距離を求め、 $f$  を最も近いクラスタに割り当て直す  
     **end for**  
**until** クラスタの割り当てが変化しない

---

る。投資信託  $f$  の隣接ベクトルの要素  $x_i$  は、次のように定義される。

$$x_i = \begin{cases} 1 & (\text{投資信託 } f \text{ が企業 } c_i \text{ に投資しているとき}) \\ 0 & (\text{そうでないとき}) \end{cases}$$

隣接ベクトルの次元数  $m$  は、投資先企業の総数である。この隣接ベクトルに対して、 $m$  次元空間におけるユークリッド距離を用いて  $k$  平均法によるクラスタリングを行う。Algorithm 2 に  $k$  平均法のアルゴリズムを示す。  $k$  平均法により、投資先が似ている投資信託が同じクラスタに属するように投資信託集合  $F$  が  $k$  個のクラスタに分割される。

## 4 分析結果

### 4.1 分析データ

Yahoo!ファイナンスで公開されているモーニングスター社の投資信託情報<sup>\*1</sup>から、投資対象が国内株式型となっている投資信託 691 銘柄について、各投資信託の組み入れトップ 10 銘柄を取得した。これらの投資信託の組み入れトップ 10 銘柄に含まれていた企業は 769 社あった。これらの情報は、2012 年 10 月 17 日の時点で公開されていたものである。

### 4.2 HITS アルゴリズムによる独自性が高い投資信託の分析結果

まずはじめに、このデータに対して  $n = 1$  として HITS アルゴリズムを適用した。その結果、ハブ度は、最も小さいもので 10、最も高いもので 2,060 であった。ハブ度を小さい方から調べたところ、最初に大きなギャップが生じていたのが 250 近辺であったため、ハブ度が 250 より小さい 130 銘柄を独自性が高い投資信託として抽出した。表 1, 2, 3 に、それぞれ、独自性が高い（ハブ度が小さい）投資信託、ハブ度が閾値近辺の投資信託、独自性が低い（ハブ度が高い）投資信託のリストを示す。名称が長いものについては名称の一部を省略した。また、『』内は投資信託の愛称である。

HITS アルゴリズムを用いて分析した結果、低位株や小型株を投資対象としている投資信託がハブ度が低い、すなわち、独自性が高いものとして抽出された。今回は組入比率上位 10 銘柄を対象として分析するため、 $n = 1$  のときのハブ度の理論的最小  $h_{\min}$  は、ひとつの投資信託だけからしか投資されていない最小のオーソリティー度  $a_{\min} = 1$  の企業だけに投資しているとき  $h_{\min} = 10$  である。つまり、最もハブ度が低かった投資信託「低位株オープン」は、他のどの投資信託とも組入比率上位 10 銘柄が異なっていた。

また、ハブ度が最も高かった（独自性が最も低かった）投資信託は、すべて TOPIX（東証株価指数）をベンチマークとする TOPIX 連動型インデックス・ファンドであった。

### 4.3 $k$ 平均法によるクラスタリングの結果

独自性が高い投資信託 130 銘柄を除いた 561 銘柄に対し、 $k$  平均法を用いてクラスタリングを行った。いく

表 1 独自性が高い投資信託。

投資信託 $f$	$h(f)$
低位株オープン	10
シンプレクスジャパンバリューアップ F	12
日本低位株ファンド	16
野村 日本真小型株投信	19
SBI 中小型割安成長株ファンド『jrevive』	22
野村 小型株オープン	23
スパークス M&S ジャパンファンド『華咲く中小型』	24
(グローバルラップ) 日本小型株式 F	24
BNP パリバ 日本成長株ファンド『鞍馬天狗』	25
大和住銀 日本小型株ファンド	26
スパークス ジャパンスモールキャップ F『ライジングサン』	26
スーパー店頭株兆 (きざし) ポートフォリオ	29
日興ターゲットジャパンファンド	35
YMR みずほ 日本株 F (ヤングカンパニー)	36
TA 中部経済圏株式ファンド『ゲンキ中部』	40
東京海上 スーパーセレクト日本株ファンド『日本の匠』	40
YMR みずほ 日本株 F (スモールキャップ)	40
三菱 UFJ 技術のチカラ	41
アパディーン 日本小型株ファンド	42
黒田アクティブジャパン	42
黒田アクティブアルファ	42
いちよし公開ベンチャーファンド	43
J-Stock アクティブオープン	44
JF ジャパンディスカバーファンド	45
JF ジャパンテクノロジーファンド	45

表 2 独自性が低い投資信託。

投資信託 $f$	$h(f)$
三菱 UFJ DC 国内株式インデックスファンド	2,060
三菱 UFJ TOPIX ファンド	2,060
三菱 UFJ (DC) TOPIX ファンド	2,060
インデックスファンド TSP	2,060
東京海上セレクション日本株 TOPIX	2,060
日立 国内株式インデックスファンド	2,060
『ゆうせい DC TOPIX インデックス』	2,060
年金積立インデックスファンド TOPIX『DCTOPIX』	2,060
日本株式インデックスファンド	2,060
三菱 UFJ DC トピックスオープン	2,060
三菱 UFJ トピックスオープン	2,060
DC 日本株式インデックスファンド A	2,060
DC 日本株式インデックスファンド	2,060
ファンドマネジャー (国内株式)	2,060
MHAM 日本株式インデックスファンド (FW)	2,060
野村 日本株インデックス (一任口座)	2,060
三菱 UFJ トピックスインデックスオープン	2,060
トピックスオープン	2,060
MHAM TOPIX オープン	2,060
しんきん トピックスオープン	2,060
三菱 UFJ トピックスオープン DC	2,060
ニュートピックスインデックス	2,060
インデックスコレクション (国内株式)	2,060
野村 インデックス FTOPIX『Funds-i TOPIX』	2,060
eMAXIS TOPIX インデックス	2,060

<sup>\*1</sup> <http://www.morningstar.co.jp/webasp/yahoo/>

表3 閾値付近の投資信託.

投資信託 $f$	$h(f)$
京都滋賀インデックスファンド『京(みやこ)ファンド』	210
ダイワニッポン応援 F-京都の志士達-	210
DIAM ストラテジック J-REIT ファンド	226
(業種別インデックス SF) O 公益	227
インバスコ ジャパングロスファンド	229
インバスコ 日本株式グロスファンド	229
MHAM 日本成長株ファンド (DC 年金)	229
神奈川応援ファンド	233
インバスコ 日本中小型株オープン	233
海外成長日本株式オープン『グローバル J』	235
日本株式アクティブオープン	237
DIAM 成長株オープン『出世株』	238
新世代成長株ファンド『ダイワ大輔』	239
日興 キャッシュリッチファンド	246
DIAM J-REIT オープン『オーナーズインカム』	275
DIAM J-REIT アクティブファンド『ハッピーオーナー』	275
DIAM J-REIT オープン『オーナーズインカム 2M』	275
BR アクティブジャパン	280
ジャパンアグレッシブグロス	280
アジアターゲットエクイティオープン	281
(セクター 10) H 商業サービス等	297
スパークス新国際優良日本株ファンド『厳選投資』	303
アムンディターゲットジャパンファンド	303
中小型株式オープン『投資満々』	319
彩の国応援ファンド	342

表4 各クラスターに属する投資信託の数

番号	名称	投信数
1	J-REIT	53
2	日経 225	59
3	TOPIX	72
4	小型アクティブ	183
5	大型アクティブ	194
	計	561

つかの  $k$  を試した結果,  $k = 5$  のときに比較的良好な結果が得られた. 各クラスターに属する投資信託の数を表 4 に示す. 本論文では, それぞれのクラスターに属する投資信託の名称を参考にして, (1) J-REIT, (2) 日経 225, (3) TOPIX, (4) 小型アクティブ, (5) 大型アクティブという名称を付けた. 表 5-7 に, それぞれに属する投資信託のリストの一部として, ハブ度が小さいもの 5 銘柄とハブ度が小さいもの 5 銘柄を示す.

独自性が高い投資信託を除いた 561 銘柄とその投資先企業間の投資関係ネットワークを図 2 に示す.  $k$  平均法によるクラスタリングの結果によって, クラスターごとに色分けされている.

#### 4.4 投資先企業の分析

最後に,  $k$  平均法によって求められたクラスターごとに, どのような企業に投資しているかを調べた. 図 3-5 に, それぞれのクラスターに属する投資信託の投資先

表5 クラスター 1 (J-REIT) に属する投資信託

投資信託 $f$	$h(f)$
Jリートアクティブファンド(1年決算)	459
ダイワ J-REIT オープン	459
インデックスファンド J リート	459
MHAM J-REIT インデックス F (DC)	459
新光 J-REIT オープン (DC 年金)	459
J-REIT リサーチ (毎月決算型)	373
三菱 UFJ J リートオープン (3 カ月決算型)	359
DIAM J-REIT オープン『オーナーズインカム』	275
DIAM J-REIT アクティブファンド『ハッピーオーナー』	275
DIAM J-REIT オープン『オーナーズインカム 2M』	275

表6 クラスター 2 (日経 225) に属する投資信託

投資信託 $f$	$h(f)$
国際 ジャパンオープン	1302
エスピー日本株オープン 225	1161
野村 インデックス F 日経 225『Funds-i 日経 225』	1161
野村 日経 225 インデックス F (確定拠出年金)	1137
三井住友 DC 年金日本株式 225 ファンド	1137
野村 日本ブランド株投資 (リアル) 年 2 回	543
野村 日本ブランド株投資 (資源国通貨) 年 2 回	543
野村 日本ブランド株投資 (アジア通貨) 毎月	543
野村 日本ブランド株投資 (アジア通貨) 年 2 回	543
野村 日本ブランド株投資 (資源国通貨) 毎月	543

表7 クラスター 3 (TOPIX) に属する投資信託

投資信託 $f$	$h(f)$
三菱 UFJ DC 国内株式インデックスファンド	2060
三菱 UFJ TOPIX ファンド	2060
三菱 UFJ (DC) TOPIX ファンド	2060
インデックスファンド TSP	2060
東京海上セレクション日本株 TOPIX	2060
大和住銀 日本グロス株ファンド『海のくに』	1584
好配当日本株式オープン『好配当ニッポン』	1507
MHAM 日本株式	1480
MHAM 日本株式 (DC 年金)	1480
DIAM 国内株オープン『自由演技』	1436

表8 クラスター 4 (小型アクティブ) に属する投資信託

投資信託 $f$	$h(f)$
YMR みずほ 日本株 F (エクセレント)	1353
日本技術評価オープン『スーパーテクノロジー』	1214
(日本株セレクトオープン) 日本株グロス『日本新世紀』	1211
アパディーンジャパンオープン『メガトレンド』	1210
YMR みずほ 日本株 F (インカム)	1201
アムンディターゲットジャパンファンド	303
(セクター 10) H 商業サービス等	297
アジアターゲットエクイティオープン	281
BR アクティブジャパン	280
ジャパンアグレッシブグロス	280

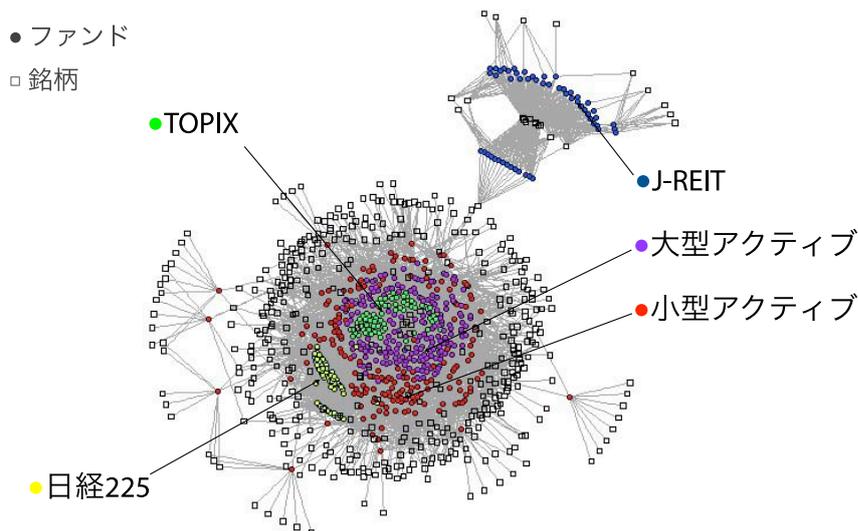


図2 投資信託と投資先企業の投資関係ネットワーク

表9 クラスタ5 (大型アクティブ) に属する投資信託

投資信託 $f$	$h(f)$
DC ファンダメンタルグロス F	1947
DC ファンダメンタルバリュ F	1947
YMR みずほ 日本株 F (ラージキャップ)	1934
ダイワバリュアアップファンド	1921
日興 アクティブバリュ F	1883
日本実力株ファンド『リアルエコノミー (DC 年金)』	858
ダイワ 金融新時代ファンド	803
(レインボー F) ファイナンシャル情報株 F	791
システムオープン	790
株ちょファンド日本 毎月『カブチョコファンド』	755

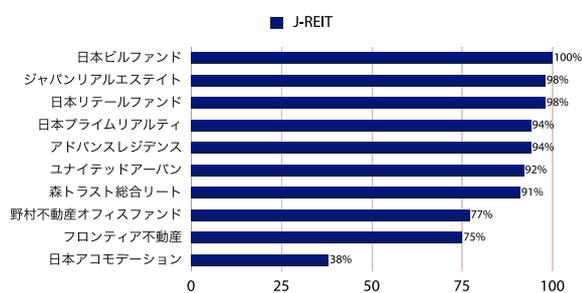


図3 クラスタ1 (J-REIT) に属する投資信託の投資先上位10銘柄とその被投資割合

図4 クラスタ2 (日経225) に属する投資信託の投資先上位10銘柄とその被投資割合

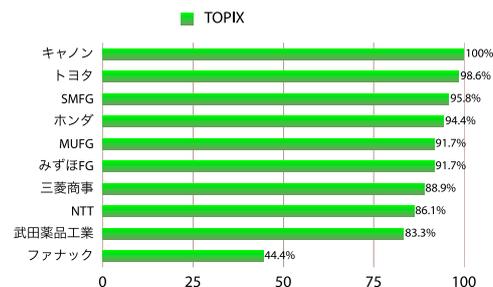


図5 クラスタ3 (TOPIX) に属する投資信託の投資先上位10銘柄とその被投資割合

上位10銘柄とその被投資割合を示す。三菱東京UFJフィナンシャルグループのように名称が長いものは略称を用いた。また、NTT (日本電信電話) のように正式名称よりもブランド名がわかりやすいものはブランド名を

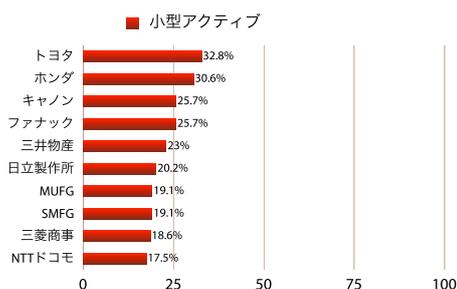


図6 クラスタ4 (小型アクティブ) に属する投資信託の投資先上位10銘柄とその被投資割合

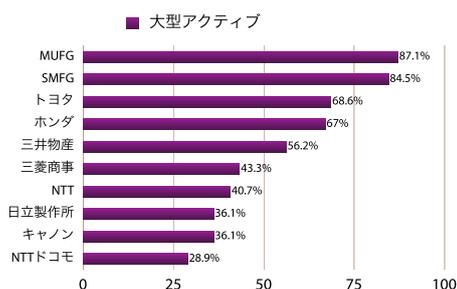


図7 クラスタ5 (大型アクティブ) に属する投資信託の投資先上位10銘柄とその被投資割合

用いた。

クラスタ1 (J-REIT)、クラスタ2 (日経225)、クラスタ3 (TOPIX) においては、投資先上位銘柄の被投資割合が高いことから、それぞれのクラスタに属する投資信託が非常に良く似ていることがわかる。また、クラスタ1 (J-REIT) における日本ビルファンド (日本ビルファンド投資法人)、クラスタ2 (日経225) におけるファナック、クラスタ3 (TOPIX) におけるキャノンには、そのクラスタに属するすべての投資信託が投資していることから、これらの企業がそのクラスタに属する投資信託にとって非常に重要な投資先であることがわかる。

一方で、クラスタ4 (小型アクティブ) とクラスタ5 (大型アクティブ) においては、投資先上位銘柄の被投資割合がそれほど高くないことから、これらのクラスタに属する投資信託はそれほど似ていない、すなわち、投資先のバラツキが大きいことがわかる。

## 5 考察とまとめ

本論文では、投資信託の組み入れトップ10銘柄に基づいて、投資信託と投資先企業を頂点とした2部グラフを作成し、これをネットワーク構造分析の技術を用いて分析する方法を提案した。提案手法では、まずHITSアルゴリズムを用いて投資信託と投資先企業間の投資関係を表すネットワーク構造を分析し、独自性が高い投資信託を抽出する。その後、 $k$ 平均法を用いてクラスタリングを行うことによって投資信託を $k$ 個のグループに分ける。 $k=5$ として分析した結果、J-REIT、日経225、TOPIX、小型アクティブ、大型アクティブに分けることができた。

提案手法を用いることによって、複数の投資信託を購入したときにそれらの投資先が似ていてリスク分散になっていないということを回避することが期待できる。HITSアルゴリズムで抽出した独自性が高いもの及び $k$ 平均法で得られた各クラスタから等しく投資信託を選択することによって投資先の銘柄及び業界が分散されることを確認し、文献[5]で報告した。

本論文では分析対象を組み入れトップ10銘柄としたが、これを全銘柄に拡張することによってより詳細な分析が期待できる。また、本論文で提案したネットワーク構造分析を用いた手法は、投資信託だけでなく、上場企業とその大株主間の投資関係や個人投資家とその投資先の投資関係など、様々な投資関係に対しても応用が期待できる。

## 参考文献

- [1] Jon M. Kleinberg. Authoritative sources in a hyperlinked environment. *Journal of the ACM (JACM)*, Vol. 46, No. 5, pp. 604–632, 1999.
- [2] Amy N. Langville and Carl D. Meyer. *Google's PageRank and Beyond: The Science of Search Engine Rankings*. Princeton University Press, 2006. 岩野和生, 黒川利明, 黒川洋 訳. Google PageRank の数理—最強検索エンジンのランキング手法を求めて—. 共立出版, 2009.
- [3] J. MacQueen. Some methods for classification and analysis of multivariate observations. In *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability*, pp. 281–297, 1967.
- [4] Toby Segaran. *Programming Collective Intelligence: Building Smart Web 2.0 Applications*. O'Reilly Media, 2007. 當山仁健, 鴨澤真夫 訳. 集合知プログラミング. オライリー・ジャパン, 2008.
- [5] 武田祐太, 松井藤五郎, 犬塚信博. 投資信託と銘柄のネットワーク構造に基づいた投資信託の分類. 第75回情報処理学会全国大会, 5M-9, 2013.