

# ネットワーク分析 2

林 大地

同志社大学 統計ファイナンス研究室 M1

2019 年 12 月 9 日

- 多くの企業などは職階性や指揮系統に見られるようにピラミッド型の階層的な組織構造を持っている。
- 企業内の組織図にかかわらず、仲間集団におけるリーダーとその配下のような人間関係においても階層的構造が形成されている場合もある。
- グラフ構造でいうと、最も上位にある 1 つの頂点から他の下位の頂点へと有向辺の連なる out-tree 構造を最も典型的な階層構造を持つネットワークとした。

## out-tree 構造なグラフ

$n$  個の頂点と  $(n-1)$  本の有向辺からなり、ある単一の頂点から他の全ての頂点へと到達可能な有向グラフ.

- その中で、Krackhardt はこのような out-tree 構造の必要十分条件として 4 つの性質を考え、この性質に違反している程度から、ある有向グラフが階層的構造を持っているかどうかを評価している. その 4 つの指標は次の通りである.

- 連結性とは、その名の通り対象となる有向グラフにおいて、どの程度繋がっているかを表している。具体的な定義は次の通りである。

### Definition 1 (連結性)

連結性とは  $V$  を辺の向きを無視しても相互に到達不可能な頂点の数, その有向グラフに含まれる頂点数を  $n$  とすると

$$\text{connectedness} = 1 - \frac{V}{\frac{n(n-1)}{2}} = 1 - \frac{2V}{n(n-1)}$$

で算出される。

- out-tree 構造において、到達可能な 2 つの頂点間の関係は上位のポジションから下位のポジションへと常に一方的である。
- そこで、もし 2 つの頂点間にお互いに到達可能な関係があるならその頂点間では上下関係はなく、階層性に違反していることになる。
- よって、少なくとも一方から他方へ到達可能な頂点の組みのうち、互いに行き来できない頂点の組みの比率によって階層性を定義することができる。

## Definition 2 (階層性)

$V$  は互いに到達可能な頂点数,  $MaxV$  は少なくとも一方から他方へ到達可能な頂点の数とする。すると、階層性を以下で定義する。

$$\text{hierarchy} = 1 - \frac{V}{MaxV}$$

- out-tree 構造は上位の頂点から下位の頂点へと必ず 1 つの経路を通して到達できるという意味で**効率的**であると言える.
- そこで, もし, 上位の頂点  $i$  から下位の頂点  $j$  への経路が 2 つ以上ある場合は冗長であるので, そのネットワークは効率的であるとは言えない.
- この効率性を定量化したのが次の効率性といった概念である.

- $n$  個の頂点と  $m$  個の有効辺からなる弱連結 (互いに行き来できることを認めると任意の頂点間を行き来できるグラフ構造のこと) の有向グラフを考える.
- もし, このグラフが out-tree 構造なら, そこに含まれる有向辺の数  $m$  は  $n$  より 1 少なく  $m = n - 1$  になる.
- よって, 必ずしも out-tree 構造でない実際の有向グラフに含まれる冗長な有向辺の数は  $m - (n - 1)$  と表すことができる.
- $m$  の理論的な最大値は  $n(n - 1)$  であるので, 冗長な有向辺の数の理論的最大値は  $n(n - 1) - (n - 1)$  である.

- これらを用いて、効率性は以下のように定義できる.
- ここで、 $V$  は  $n - 1$  を超える冗長な有向辺の数、 $\text{Max}V$  は冗長な有向辺の理論的最大値である.

### Definition 3 (効率性)

ここで、 $V$  は  $n - 1$  を超える冗長な有向辺の数、 $\text{Max}V$  は冗長な有向辺の理論的最大値である. このとき、効率性は以下のように定義される.

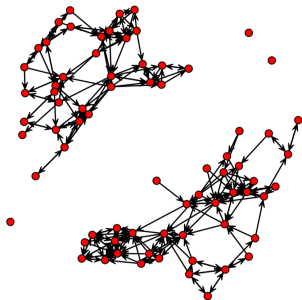
$$\begin{aligned}\text{efficiency} &= 1 - \frac{V}{\text{Max}V} \\ &= 1 - \frac{m - (n - 1)}{n(n - 1) - (n - 1)} \\ &= 1 - \frac{m - (n - 1)}{(n - 1)^2}\end{aligned}$$



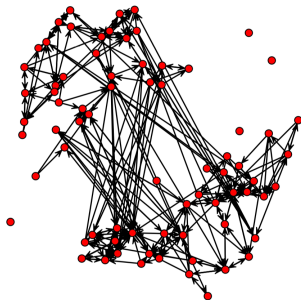
# 実際の分析例

- ここでは、実データを用いてネットワーク構造の指標を算出してみる。今回は Coleman データを用いる。
  - ▶ Coleman による 73 人の男子高校生の友人ネットワークである。
  - ▶ 1957 年の秋、1958 年の春の 2 回、よく付き合いのある相手をあげてもらった方向のあるデータである。

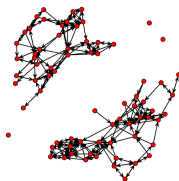
Fall, 1957



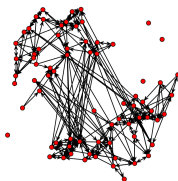
Spring, 1958



Fall, 1957



Spring, 1958



- ネットワーク構造を可視化してみると、友人間のネットワーク構造が成長していることがわかる。
  - 今回は、これを定量的に評価する。
- 
- 考察対象となる指標は以下の通りである。
  - 密度・推移性・相互性・連結性・階層性などなど...
  - とりあえず、算出した結果を表にまとめたものを次に示します。

Table: 指標のまとめ

指標	1957 秋	1958 春
密度	0.04623	0.05003
推移性	0.4627	0.3803
相互性	0.3425	0.3019
相互性の数	62	61
連結性	0.45433	0.89269
階層性	0.7001	0.8029
効率性	0.92456	0.95782
最小上界性	0.6660	0.33494

- 以下、この表を基に考察を行う。

- まずは以下の指標について考察する.

指標	1957 秋	1958 春
密度	0.04623	0.05003
推移性	0.4627	0.3803
相互性	0.3425	0.3019
相互性の数	62	61

- まずは、密度を見てみよう.
  - ▶ 密度は、対象となるグラフがどの程度、完全2部グラフ(任意の頂点同士が繋がっている)になっているかの比率を表していることを思い出そう.
  - ▶ その点では、少しばかりネットワーク構造は成長していると言える。(仲良くなったのかな?)

指標	1957 秋	1958 春
推移性	0.4627	0.3803
相互性	0.3425	0.3019
相互性の数	62	61

- 次に、相互性についてみる。
  - ▶ 相互性は簡単にいうと、対象の有向グラフにおける各ノード間の両思いの割合である。
- 相互性の値が下がっているのので、仲良しであった人の友情にヒビがはいた事を示唆している。
  - ▶ 秋から春にかけて何かあったんでしょうね...

指標	1957 秋	1958 春
推移性	0.4627	0.3803

- 推移性について見てみよう。
  - ▶ 推移性は、自分の友達の友達はまた自分の友達になっている割合を示す指標でした。
- 推移性は減少しているので、秋先は友人の友人だからとりあえず友人と思い接していたが、よくよく付き合ってみると、こいつとは友達にはなれないと判断した人が一定数いたという事を示唆しています。(残酷な世界)

指標	1957 秋	1958 春
連結性	0.45433	0.89269

- 特筆すべきは、連結性の値の上昇であろう。
- これは時間が経つにつれて、友人間のネットワークが成長している事を示しています。
  - ▶ つまり、クラスの皆が様々な人と関わりを持つようになったという事を示唆しています (これは、グラフからも見て取れましたね。)
  - ▶ クラスの担任の先生からすると微笑ましい限りですね。
- 実際に、当初は手探りだった人も学祭などを通して仲良くなるとういったことがあると思います。
- このような現象を実際に定量化できることはネットワーク分析の魅力の一つだと思います。

## 考察・その2(階層性・効率性)

指標	1957 秋	1958 春
階層性	0.7001	0.8029
効率性	0.92456	0.95782

- 次に階層性をみる.
  - ▶ 階層性は、対象となる有向グラフにおいて、どのくらい縦社会表すようなピラミッド構造になっているかを示す指標でした。(簡単にいうのが難しい...)
- この値の上昇は、友人間のネットワークが発展して生徒間に**力関係が生じている**事を表しています。(残酷な世界)
  - ▶ ドラえもんていうと、ジャイアンみたいに一方的にのび太や他のクラスメイトを友人と思って、こき使うような人物がいる事を示唆しています。(のび太はジャイアンを友人だと思っていると信じたいですが...)



指標	1957 秋	1958 春
階層性	0.7001	0.8029
効率性	0.92456	0.95782

- 効率性は、誰か 1 人に伝達事項を伝えと、そのネットワークの他の人に少ない人を介して (効率よく) 伝達できる事を示唆しています。
- これは、階層性の上昇からも、友人間ネットワークのヒエラルキーの頂点に立つ人物がいる事を示唆しています。

- この2回でネットワーク分析の基本的な評価指標を紹介した。
- 実際のデータを用いてネットワーク分析を行い、各指標が示す事について考察した。
  - ▶ 友人間ネットワークだと、微笑ましい分析結果も得られるが、**残酷な分析結果**を得られることの方が今回は多かった。
- 次回は、また大事な指標の一つである中心性指標について取り上げるつもりである。