社会ネットワーク分析（SNA: Social Network Analysis）

URL: https://www.kkaneko.jp/ai/st/sna.pptx

金子邦彦

社会ネットワーク分析の基本概念、グラフによる関係性の表現、指標を用いたノードの重要性評価

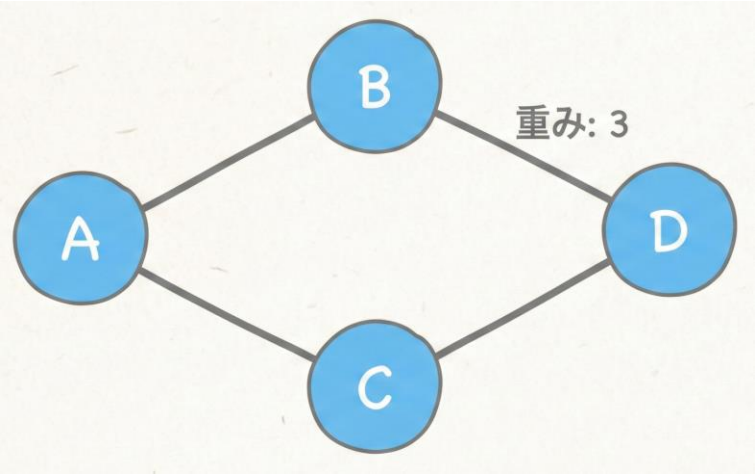
【学習内容の構成】

1. **グラフ表現**：ネットワークをノードとエッジで表現し、有向・無向グラフや重み付けで関係の方向性・強度を扱う
2. **中心性指標**：次数中心性（接続数）、媒介中心性（仲介者としての重要性）、近接中心性（到達しやすさ）、PageRank（リンク元の重要性を考慮）による評価
3. **Python実装**：NetworkXライブラリを用いた有向グラフの構築と各指標の算出

- 前提：Pythonプログラミングの基礎
- 意義：個々の属性ではなく関係性からネットワーク構造を読み解く視点の獲得

社会ネットワーク分析（SNA）

社会ネットワーク分析（SNA）は**関係性の分析手法**である



図解は Nano Banana Pro を用いて作成

グラフ表現

ネットワークを**ノード**（node）と**エッジ**（edge）で表現

- **ノード**：**人物、組織など**の表現
- **エッジ**：**ノード間の関係**の表現

エッジへの**重み**（weight）付与が可能

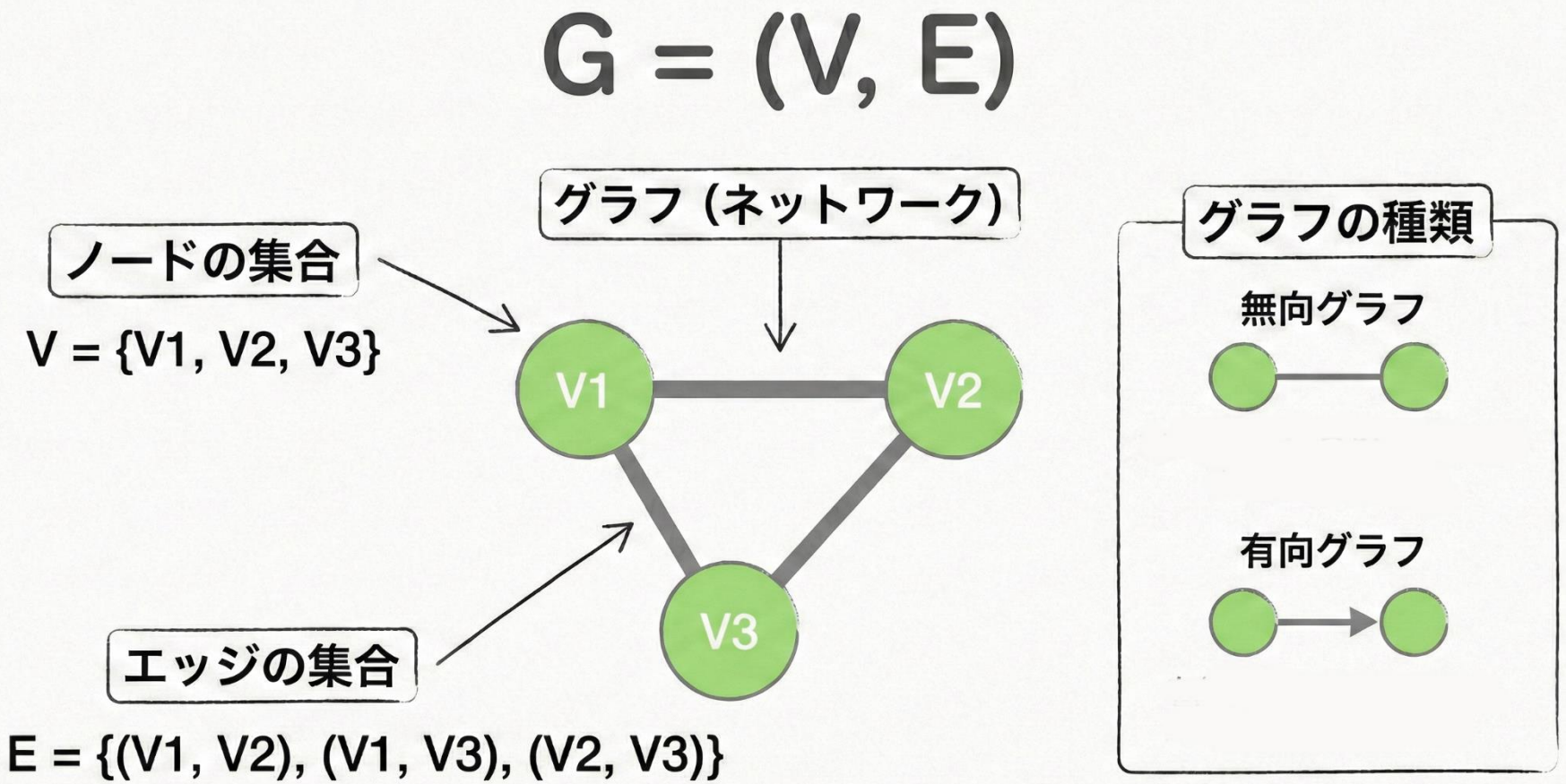
従来手法との比較

観点	統計手法	社会ネットワーク分析
分析対象	個々の属性	関係性
問い	個体の特性は何か ※	個々の要素が、どのようにつながっているか

※ 相関分析など関係性を扱う統計手法も存在

ネットワークのグラフ表現

ネットワークの**グラフ** G = (V, E) による表現



図解は Nano Banana Pro を用いて作成

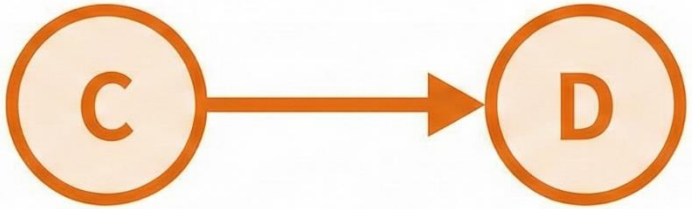
有向グラフと無向グラフ

無向グラフ

エッジに方向性がない

適用例: 友人関係（双方向の関係）

有向グラフ



エッジに方向性がある

適用例: 引用関係、影響関係

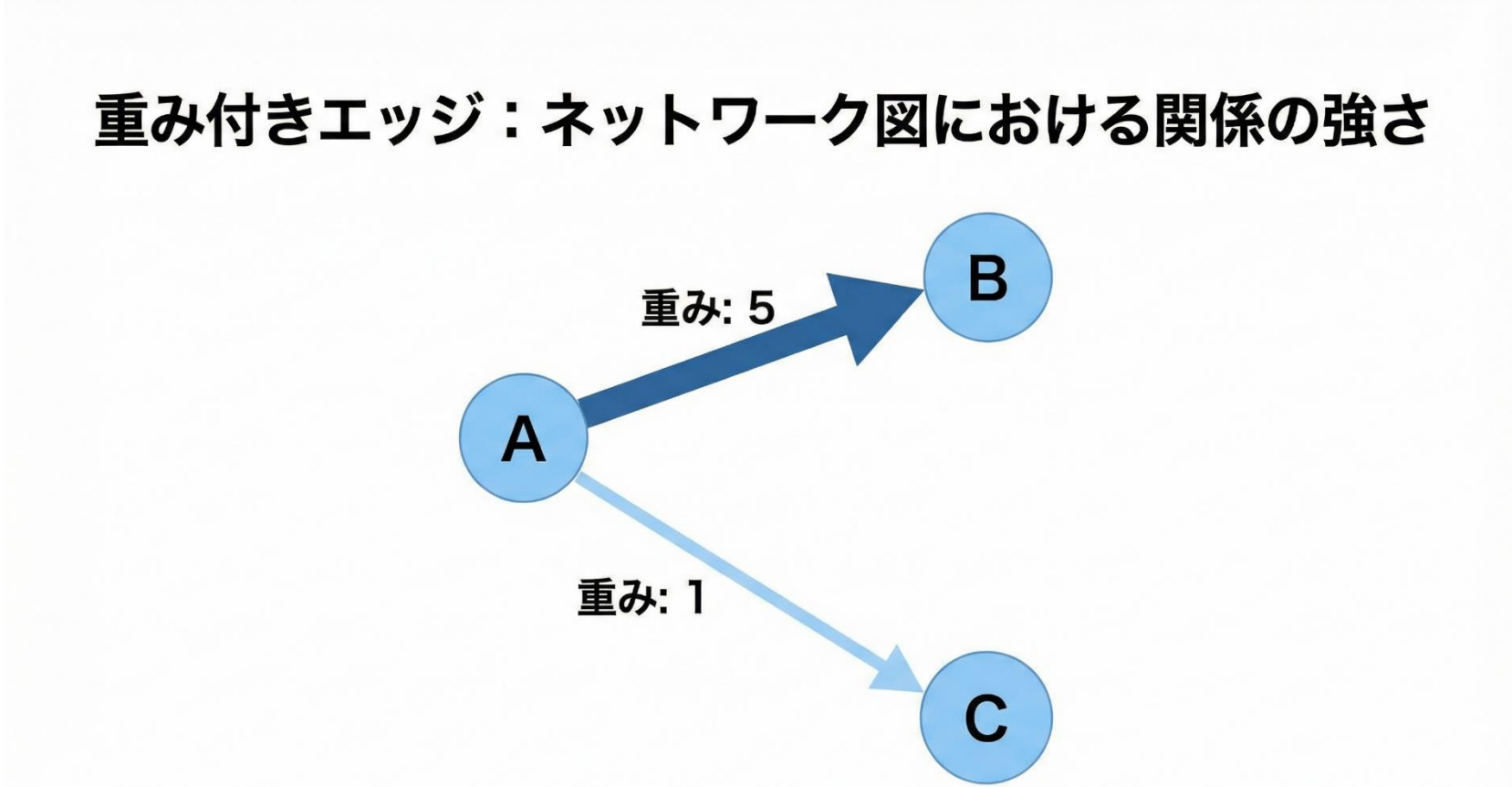
本資料では有向グラフを扱う

図解は Nano Banana Pro を用いて作成

重み付きエッジ

エッジへの**重み付与**が可能

- **重み：関係の強さや頻度の表現**
- 例：引用回数、影響の強さ



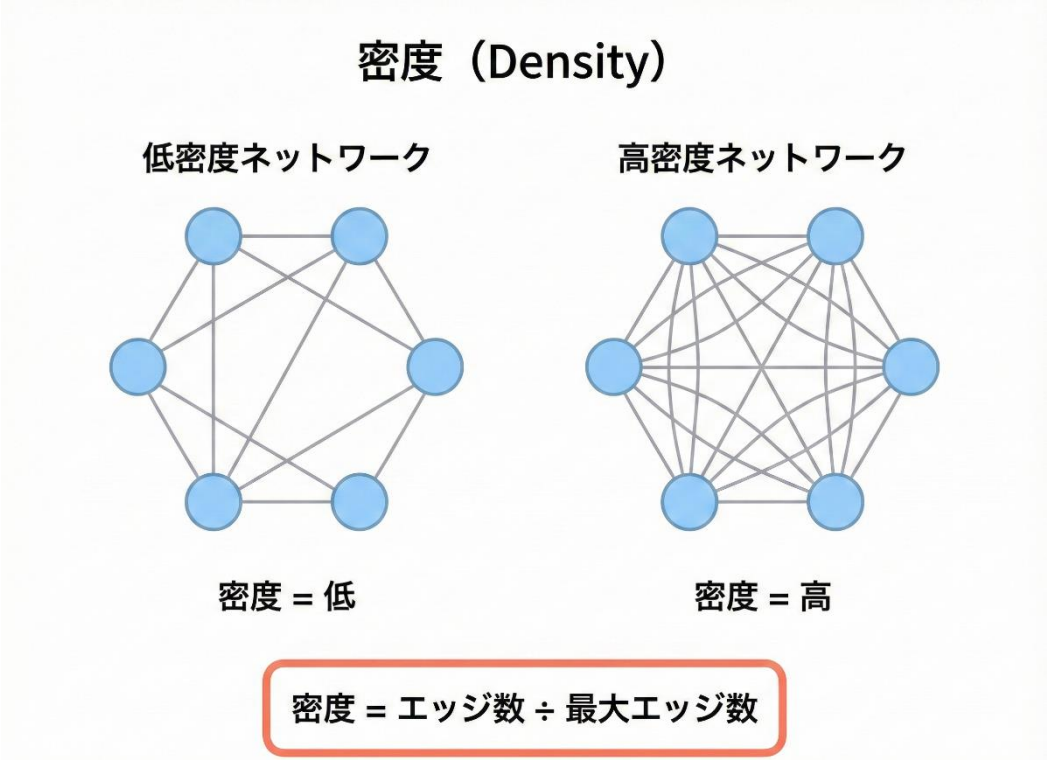
図解は Nano Banana Pro を用いて作成

主な分析手法

	分析対象	重みの考慮
密度	ネットワーク全体	なし
次数中心性（入次数・出次数）	ノード	なし
媒介中心性	ノード	**あり**
近接中心性	ノード	**あり**
PageRank	ノード	**あり**

密度（Density）

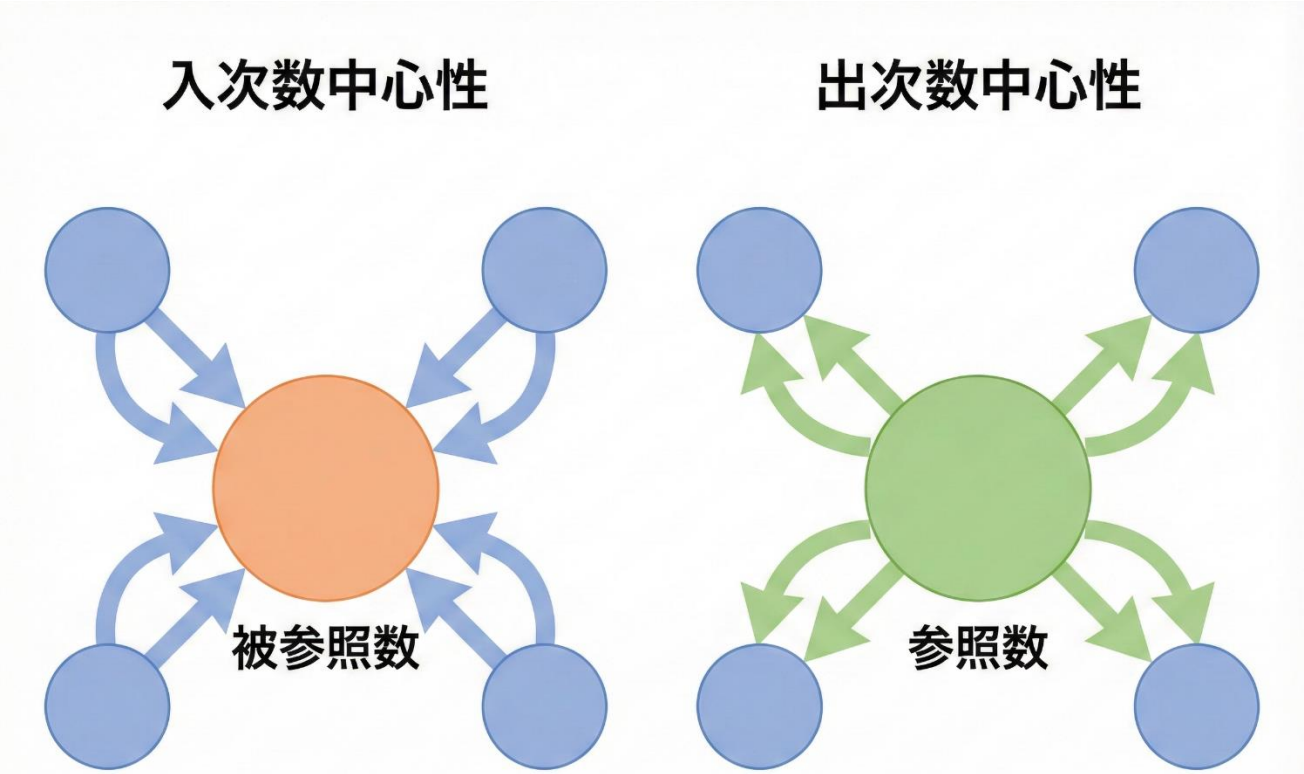
- **実際のエッジ数を最大エッジ数で割った値**
- **ネットワーク全体**の**結合度**の指標
- **有向グラフ**ではA→BとB→Aを**別エッジ**として算出
- 計算式：D = E / (N × (N-1))
  - N：ノード数
  - E：エッジ数
- 密度：重みは非考慮



図解は Nano Banana Pro を用いて作成

次数中心性（Degree Centrality）

- **ノードに接続するエッジの数**で重要性を評価
- **有向グラフ**では**入次数**と**出次数**を区別
- 重み：非考慮

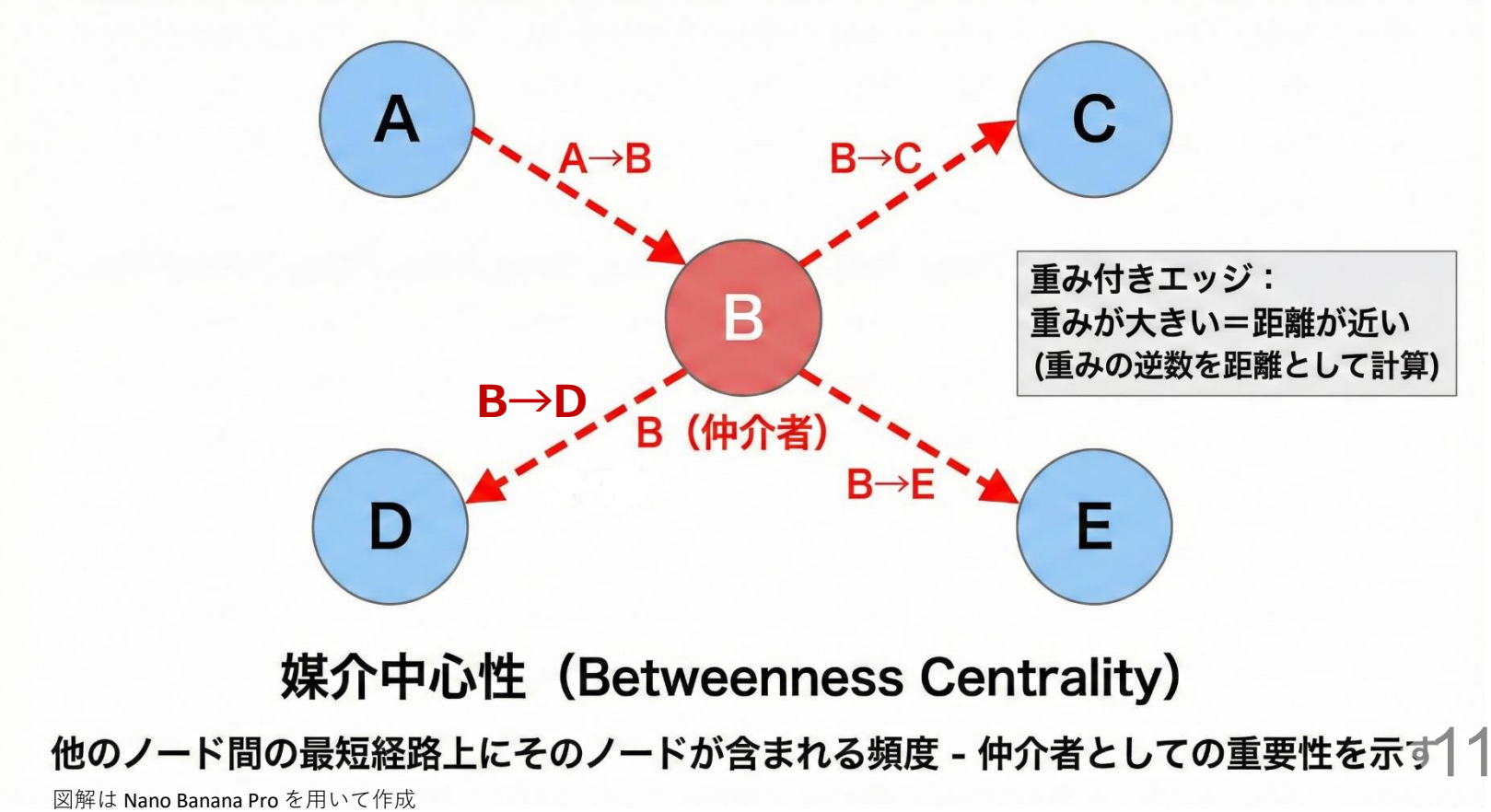


そのノードに入るエッジの数　そのノードから出るエッジの数

図解は Nano Banana Pro を用いて作成

媒介中心性（Betweenness Centrality）

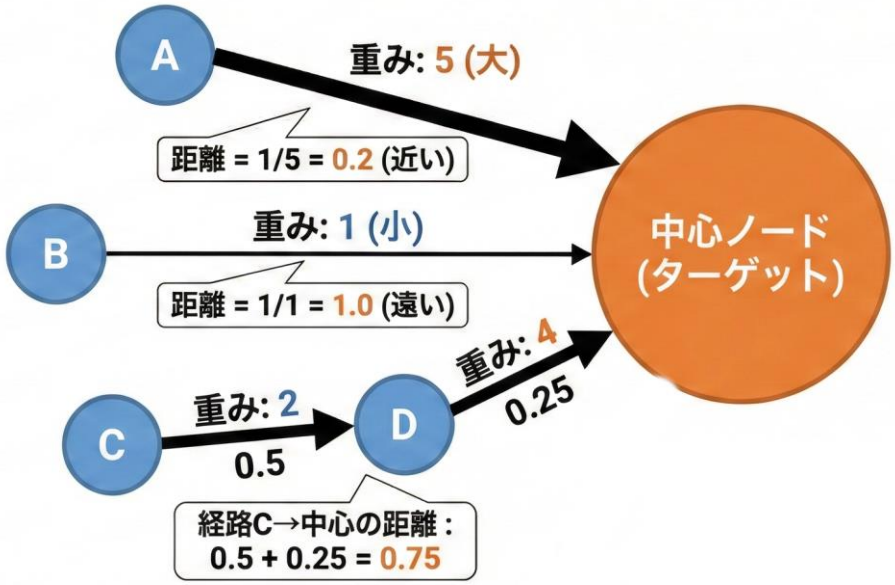
- **他のノード間の最短経路上にそのノードが含まれる頻度**
- 重み付きエッジの場合：**重みの逆数を距離として計算**



図解は Nano Banana Pro を用いて作成

近接中心性（Closeness Centrality）

- **他の全ノード**から**そのノード**への**平均経路長の逆数**
- **他のノード**からの**到達しやすさ**を示す
- 入ってくるエッジがないノード：**値は0**
- 重み付きエッジの場合：**重みの逆数を距離として計算**（重みが大きいほど距離が近い）

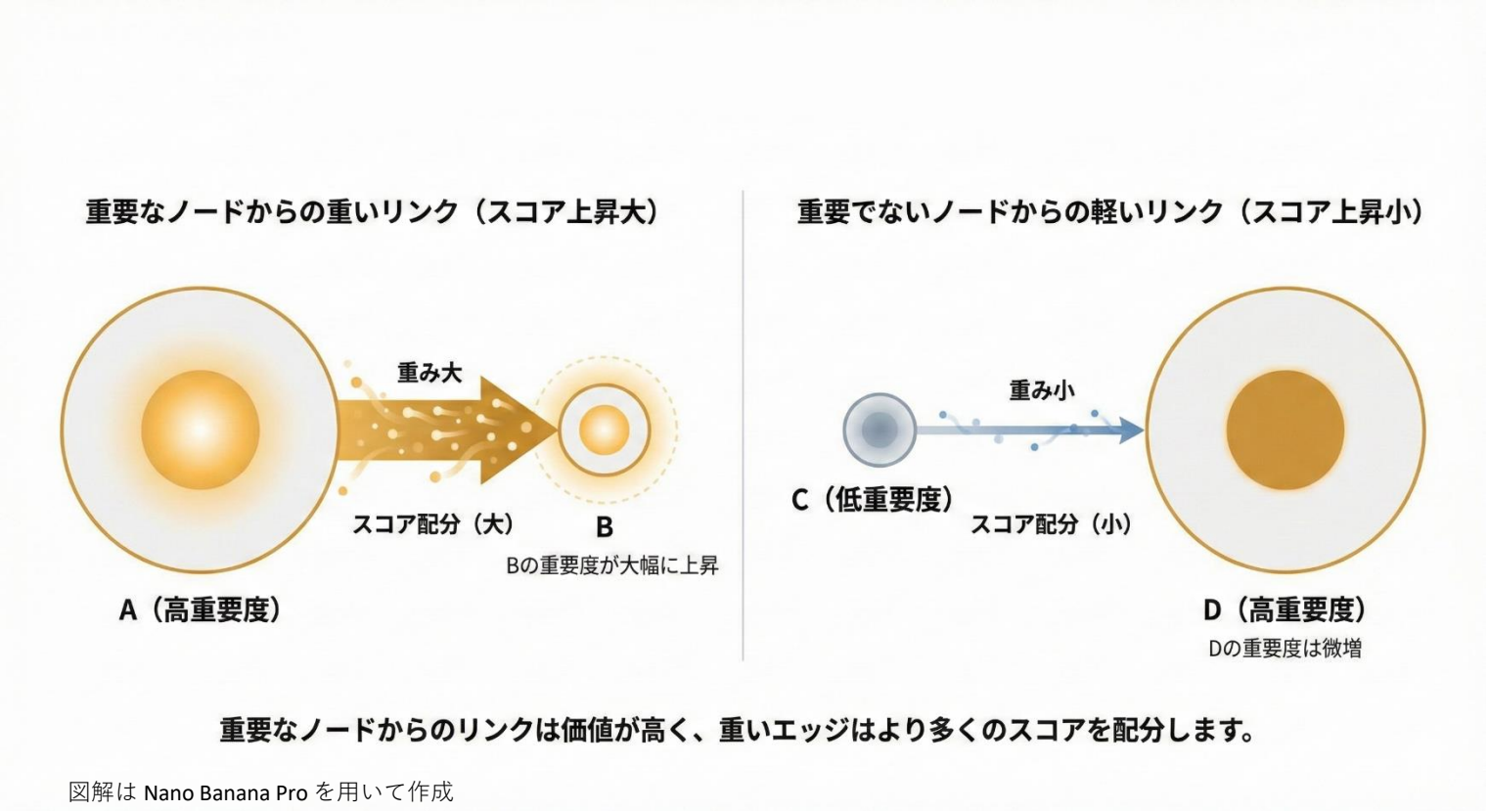


※ 近接中心性は 2種類。
この資料では「**入方向**」のものを説明
出方向：引用先への到達しやすさ（参照の効率性）
入方向：引用元からの到達しやすさ（被引用の容易さ）

図解は Nano Banana Pro を用いて作成

PageRank

- **重要なノードからリンクされているノード**ほど高い値
- **入次数**だけでなく**リンク元の重要性**も考慮
- 重み付きエッジの場合：重みに応じたスコア配分



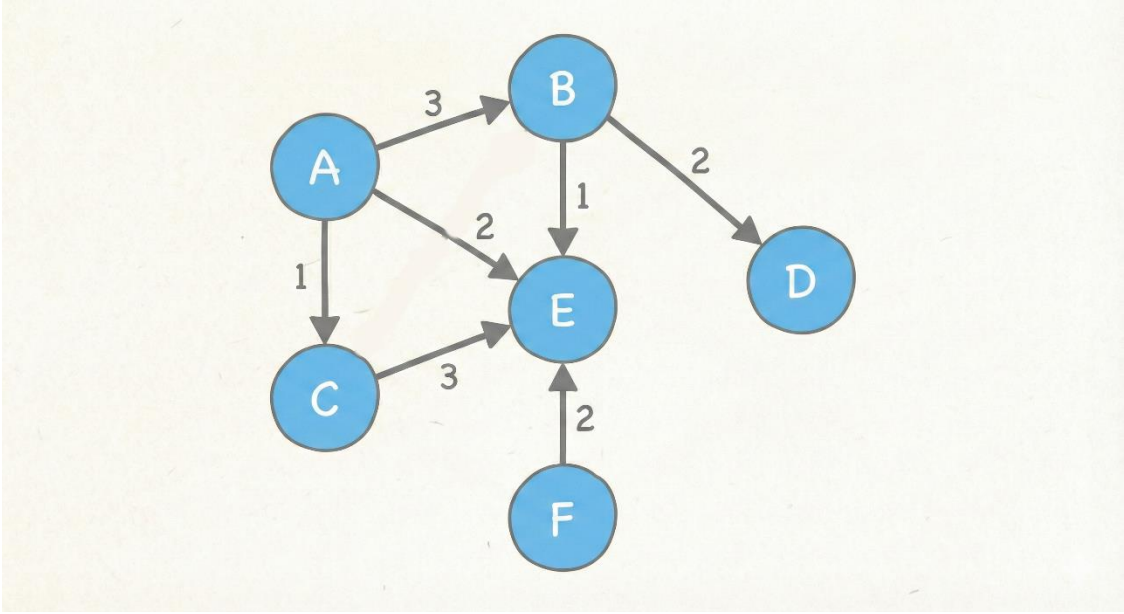
図解は Nano Banana Pro を用いて作成

サンプルネットワーク

論文の引用ネットワークを例として使用。

論文Aが論文Bを引用する場合、**A→Bのエッジ**を作成。**重み**は引用の**重要度**を表現。

エッジ	重み
A → B	3
A → C	1
A → E	2
B → D	2
B → E	1
C → E	3
F → E	2



図解は Nano Banana Pro を用いて作成

Python による社会ネットワーク分析

```
import networkx as nx

# 有向グラフの構築（重み付き）
G = nx.DiGraph()
edges = [
    ('A', 'B', 3), ('A', 'C', 1), ('A', 'E', 2),
    ('B', 'D', 2), ('B', 'E', 1),
    ('C', 'E', 3), ('F', 'E', 2)
]
G.add_weighted_edges_from(edges)

# 重みの逆数を距離として設定
for u, v, data in G.edges(data=True):
    data['distance'] = 1 / data['weight']

# 各指標の計算
in_degree_cent = nx.in_degree_centrality(G)
out_degree_cent = nx.out_degree_centrality(G)
betweenness_cent = nx.betweenness_centrality(G, weight='distance')
closeness_cent = nx.closeness_centrality(G, distance='distance')
pagerank = nx.pagerank(G, weight='weight')

# 結果の表示（小数点以下3桁）
print("入次数中心性:", {k: f"{v:.3f}" for k, v in in_degree_cent.items()})
print("出次数中心性:", {k: f"{v:.3f}" for k, v in out_degree_cent.items()})
print("媒介中心性:", {k: f"{v:.3f}" for k, v in betweenness_cent.items()})
print("近接中心性:", {k: f"{v:.3f}" for k, v in closeness_cent.items()})
print("PageRank:", {k: f"{v:.3f}" for k, v in pagerank.items()})
```

```
>>> # 結果の表示（小数点以下3桁）
>>> print("入次数中心性:", {k: f"{v:.3f}" for k, v in in_degree_cent.items()})
入次数中心性: {'A': '0.000', 'B': '0.200', 'C': '0.200', 'E': '0.800', 'D': '0.200', 'F': '0.000'}
>>> print("出次数中心性:", {k: f"{v:.3f}" for k, v in out_degree_cent.items()})
出次数中心性: {'A': '0.600', 'B': '0.400', 'C': '0.200', 'E': '0.000', 'D': '0.000', 'F': '0.200'}
>>> print("媒介中心性:", {k: f"{v:.3f}" for k, v in betweenness_cent.items()})
媒介中心性: {'A': '0.000', 'B': '0.050', 'C': '0.000', 'E': '0.000', 'D': '0.000', 'F': '0.000'}
>>> print("近接中心性:", {k: f"{v:.3f}" for k, v in closeness_cent.items()})
近接中心性: {'A': '0.000', 'B': '0.600', 'C': '0.200', 'E': '1.371', 'D': '0.600', 'F': '0.000'}
>>> print("PageRank:", {k: f"{v:.3f}" for k, v in pagerank.items()})
PageRank: {'A': '0.101', 'B': '0.144', 'C': '0.116', 'E': '0.355', 'D': '0.183', 'F': '0.101'}
>>>
```

社会ネットワーク分析の結果

実行結果

ノード	入次数中心性	出次数中心性	媒介中心性	近接中心性	PageRank
A	0.000	**0.600**	0.000	0.000	0.101
B	0.200	0.400	**0.050**	0.600	0.144
C	0.200	0.200	0.000	0.200	0.116
D	0.200	0.000	0.000	0.600	0.183
E	**0.800**	0.000	0.000	**1.371**	**0.355**
F	0.000	0.200	0.000	0.000	0.101

ネットワーク全体：**ノード数 6、エッジ数 7、密度 0.233**

結果の解釈

入次数中心性と出次数中心性

- **Eの入次数中心性**が**0.800と最も高い**（4つの論文から引用）
- AとFは0.000（引用されていない）
- **出次数中心性**は**Aが0.600と最も高い**（3つの論文を引用）
- DとEは0.000（引用していない）

媒介中心性

- **Bのみが0.050**を示す（**A→D経路の仲介**）
- 他のノードは0.000

結果の解釈

近接中心性

- **Eが1.371と最も高い（4つのノードから引用され、到達されやすい）**
- BとDが0.600（それぞれ引用元から短い距離）
- AとFは0.000（引用されておらず、他から到達できない）

PageRank

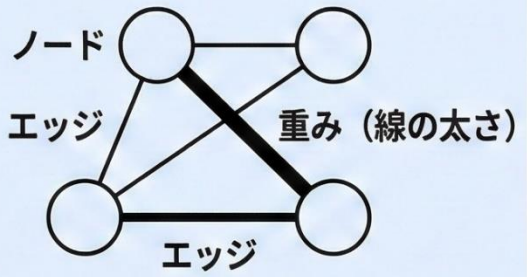
- **Eが0.355と最も高い（4つのノードからリンク）**
- Dが0.183（Bからのみリンクされるが、Bの重要度が反映）

密度

- **0.233**は、最大エッジ数**30本（6×5）**のうち**7本のエッジ**が存在することを示す

全体まとめ — 社会ネットワーク分析

基本概念：グラフで関係を表す



- ノード（○）→ 人物や組織
- エッジ（線）→ 関係
- 重み（線の太さ）→ 関係の強さ

グラフの種類：2種類のグラフ

双方向
例：友人

一方向
例：引用

分析手法：何を測るか

全体の指標
- 密度：つながりの多さ

ノードの重要性（中心性）
- 次数：接続の数
- 媒介：仲介者の度合い
- 近接：到達しやすさ
- PageRank：重要なノードからのリンクを重視

重みの扱い：重みを考慮するか

考慮しない	密度、次数中心性
考慮する	媒介中心性、近接中心性、PageRank

従来手法との違い：統計手法との違い

	統計手法	SNA
対象	個々の属性	関係性
問い	特性は何か	どうつながるか

図解は Nano Banana Pro を用いて作成