

高頻度経済データの経済物理学解析

プロジェクト責任者
高安美佐子:東京工業大学

報告者:高安秀樹:ソニーCSL

プロジェクトの目的

これまで科学的な観点から解析されたことがない膨大で詳細な経済データを、経済物理学の解析手法に基づいて解析する。これによって、高頻度経済データを超並列計算機で解析するための手法を確立し、超並列計算機の新たな利用方法の有用性を実証する。

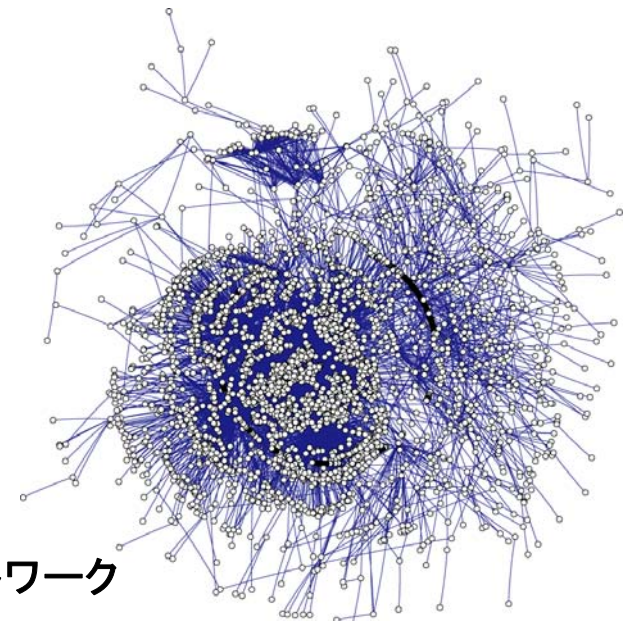
高安美佐子・鮫島伸二郎・渡邊隼史:東京工業大学

大西立顕・紺野友彦:東京大学

池田裕一:日立総合計画研究所

渡邊國彦:地球シミュレータセンター

高安秀樹:ソニーCSL



銀行業のネットワーク

長期的展望:

高頻度経済データの経済物理解析の有用性を実証する。

1: 日本国内の企業約100万社間の取引関係のデータ

2: スーパーマーケットやコンビニにおける小売販売の詳細な高頻度データ

3: 為替や株式市場における高頻度取引データ

今年度の目標

地球シミュレータを使う初年度であり、地球シミュレータ用プログラムの開発を行う。日本国内の事実上全ての企業、約100万社の取引関係の隣接行列の分析によって、現実の企業の取引ネットワークの全体構造を初めて描き出す。

1) 全国ネットワークを特徴付ける指数の計算

100万社の距離行列の成分を計算する

隣接行列のN乗の計算

推移的閉方

オーソリティ度、ハブ度の解析

コミュニティ解析、地域別企業成長率マップ

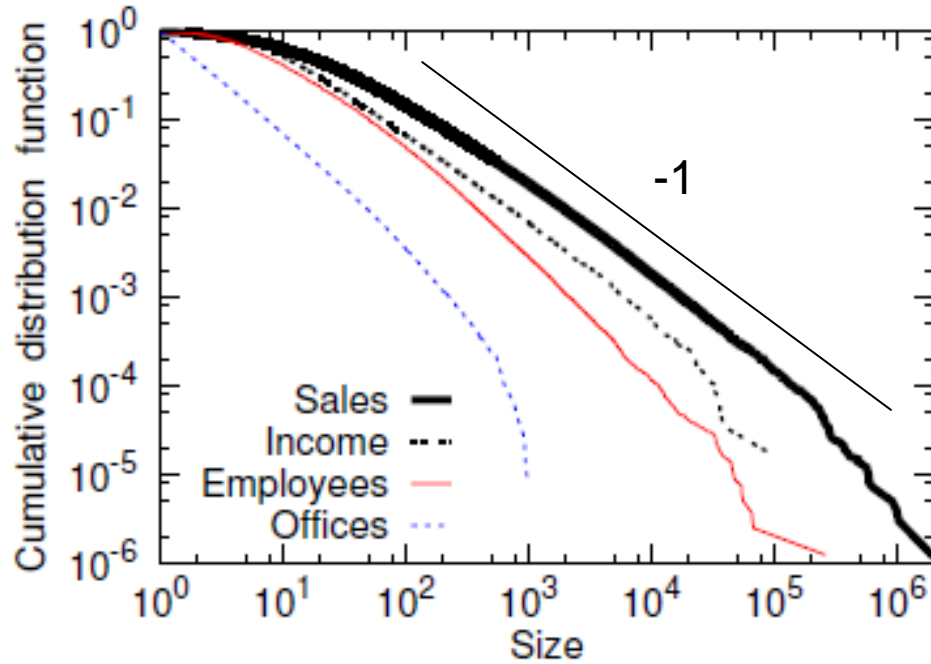
くりこんだ視点での成長率の統計性

2) 企業成長のモデルの開発(次年度のシミュレーションを目指して)

解析に用いたデータ: 経済産業研究所提供、東京商工リサーチ社調査(2005年度)による日本企業約100万社(ペーパーカンパニーを除くほぼ全ての企業)。

財務(3期分)、他社との取引関係の情報(仕入れ、販売、株の所有関係)、他社との親子関係情報など。取引先指定上限24社。

日本の企業のサイズ分布



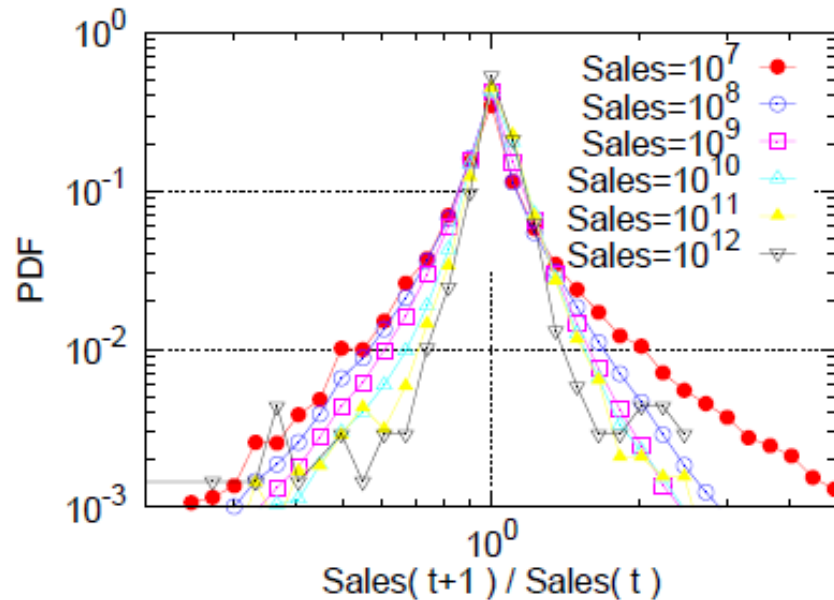
破片の分布と類似

- 売り上げ
- 所得
- 従業員数
- 支店数

どれで見てもベキ分布

スケールフリー=フラクタル性

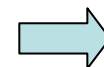
(ジップの法則)が成立している



成長率の分布

大きな企業ほど成長率のゆらぎは小さい

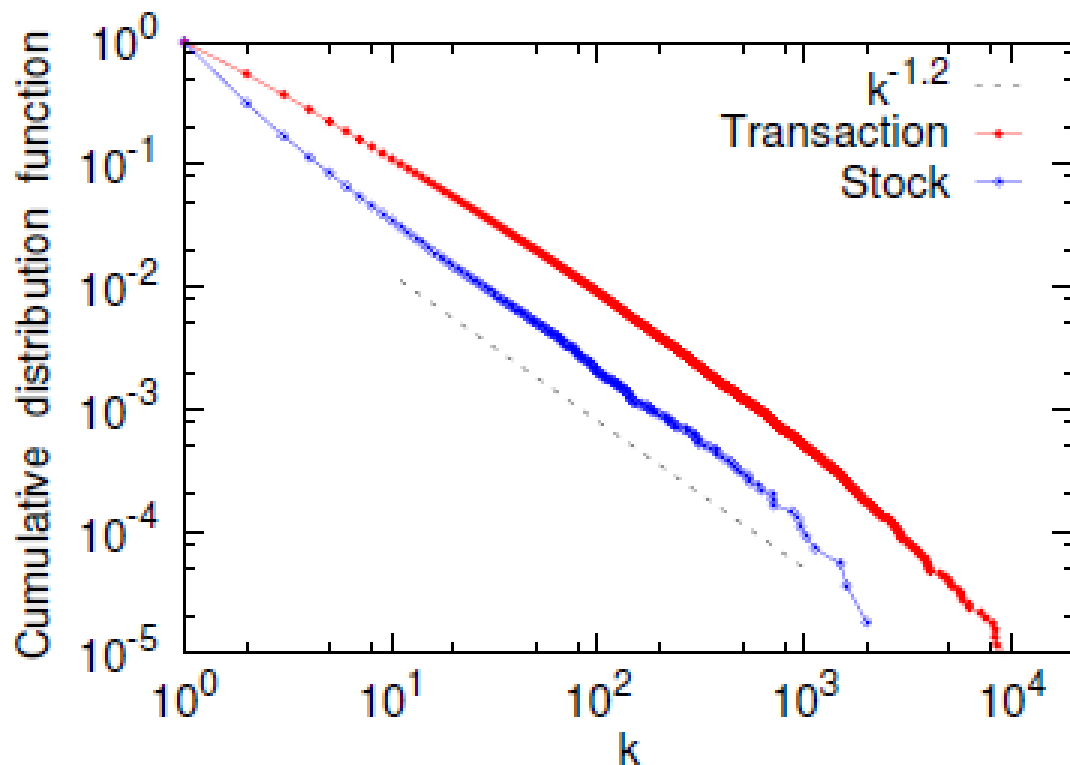
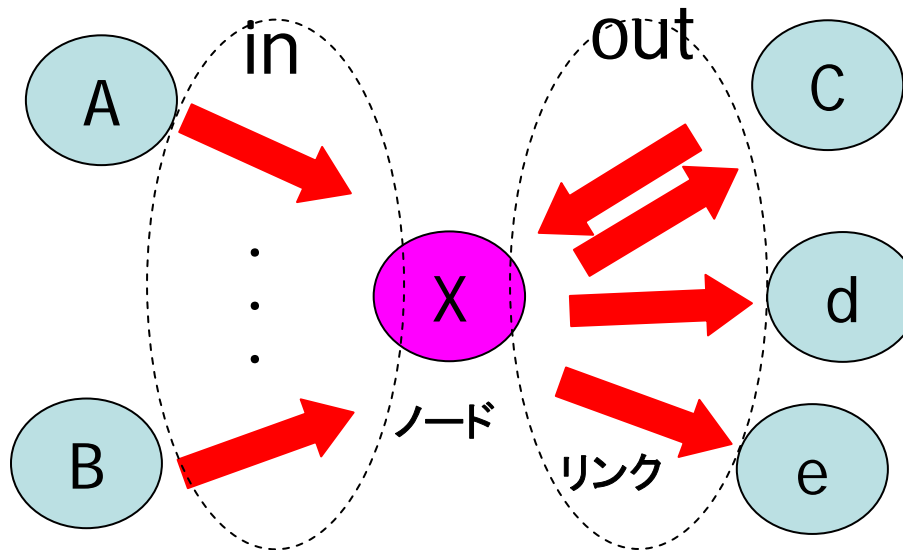
分布関数は正規分布よりもすその大きな分布



企業の統計物理モデルの提案

取引のネットワーク構造

お金の流れ
=物の流れの逆



リンク数の分布

ベキ分布

(スケールフリーネットワーク)

取引ネットワーク

株所有のネットワーク

ハブ度オーソリティ度の導入

Webページ検索において、Webページの重要度を判定するための指標。

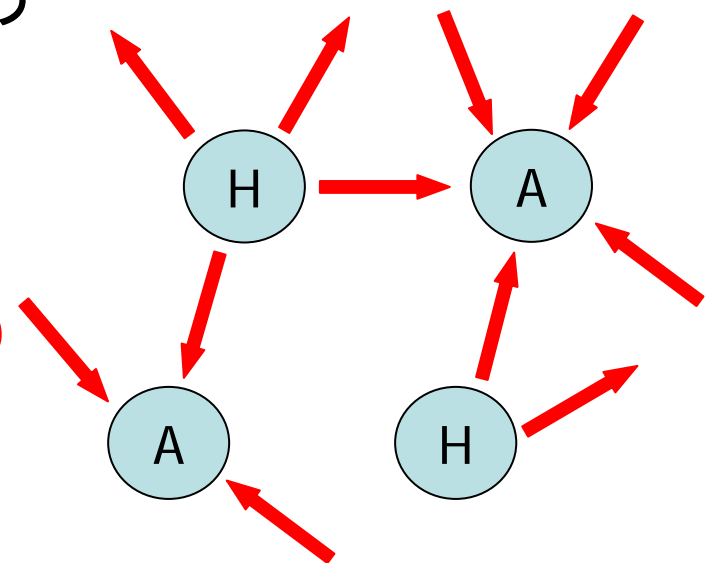
- **オーソリティ度**：他企業からどのくらい取引先として指名を受けているか(お金が入ってくるか)という指標。
- **ハブ度**：他企業をどれくらい多く取引先として指名しているか(お金を出しているか)という指標。

オーソリティ度とハブ度は互いに強化し合う関係にある。

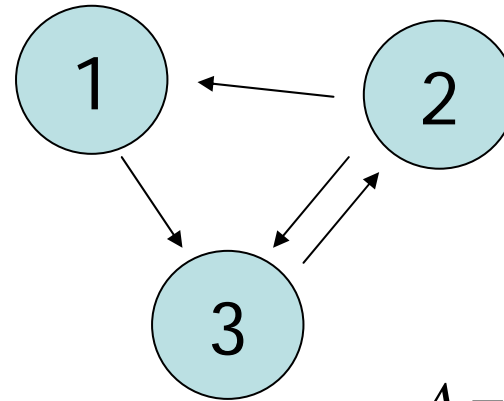
数値の高いハブ(よいハブ)は、数値の高いオーソリティを指名している。

数値の高いオーソリティ(よいオーソリティ)は、数値の高いハブに指名されている。

単に次数の多いものではない！！



N社の企業ネットワークにおいて企業*i*→企業*j*というリンクがある時(お金の流れがある時)、 $A(i,j)=1$ 、そうでないとき0とするような $N \times N$ 行列を隣接行列と呼ぶ。



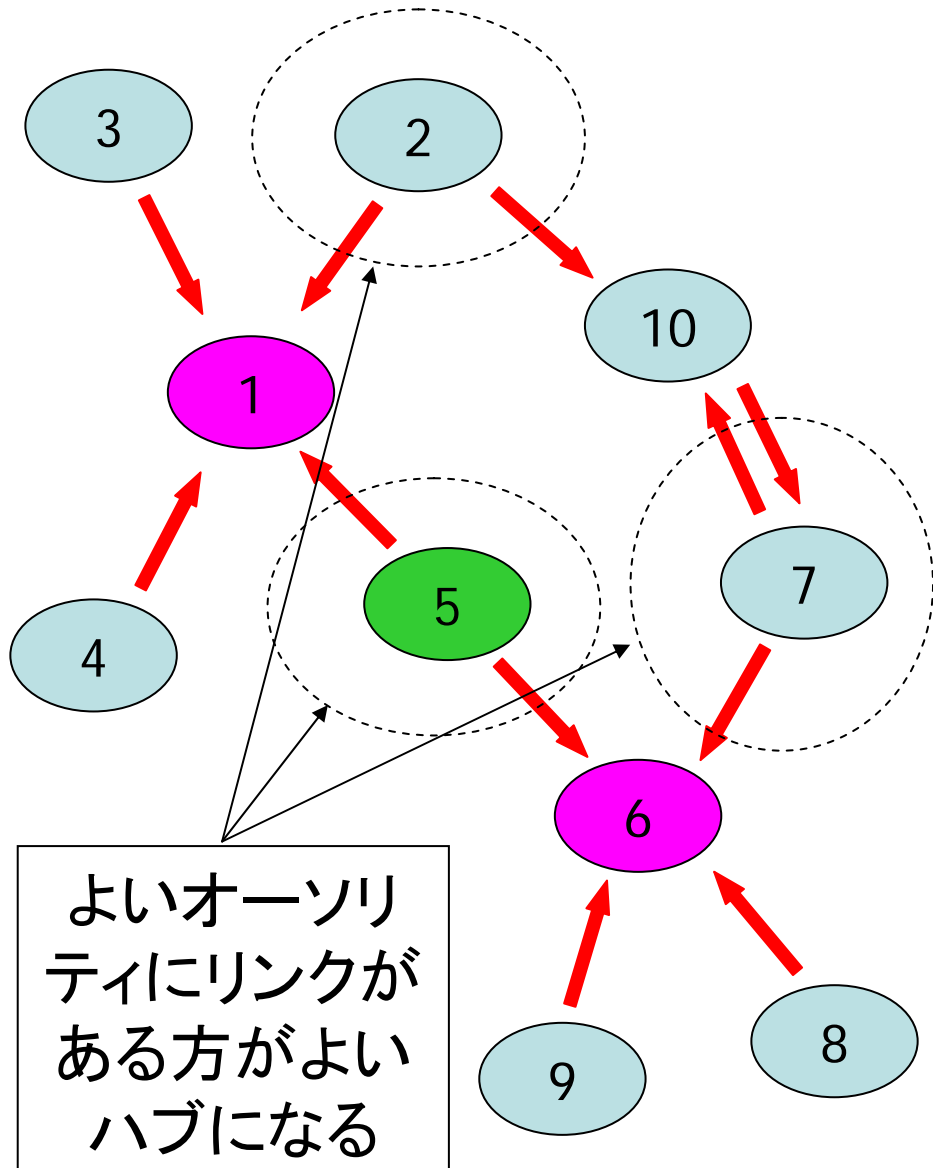
$$A = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

●オーソリティ度、ハブ度を求めるアルゴリズム

適当なオーソリティ度 $\mathbf{X}(t)$ とハブ度 $\mathbf{Y}(t)$ から始め、以下の様な操作を値が収束するまで繰り返す。

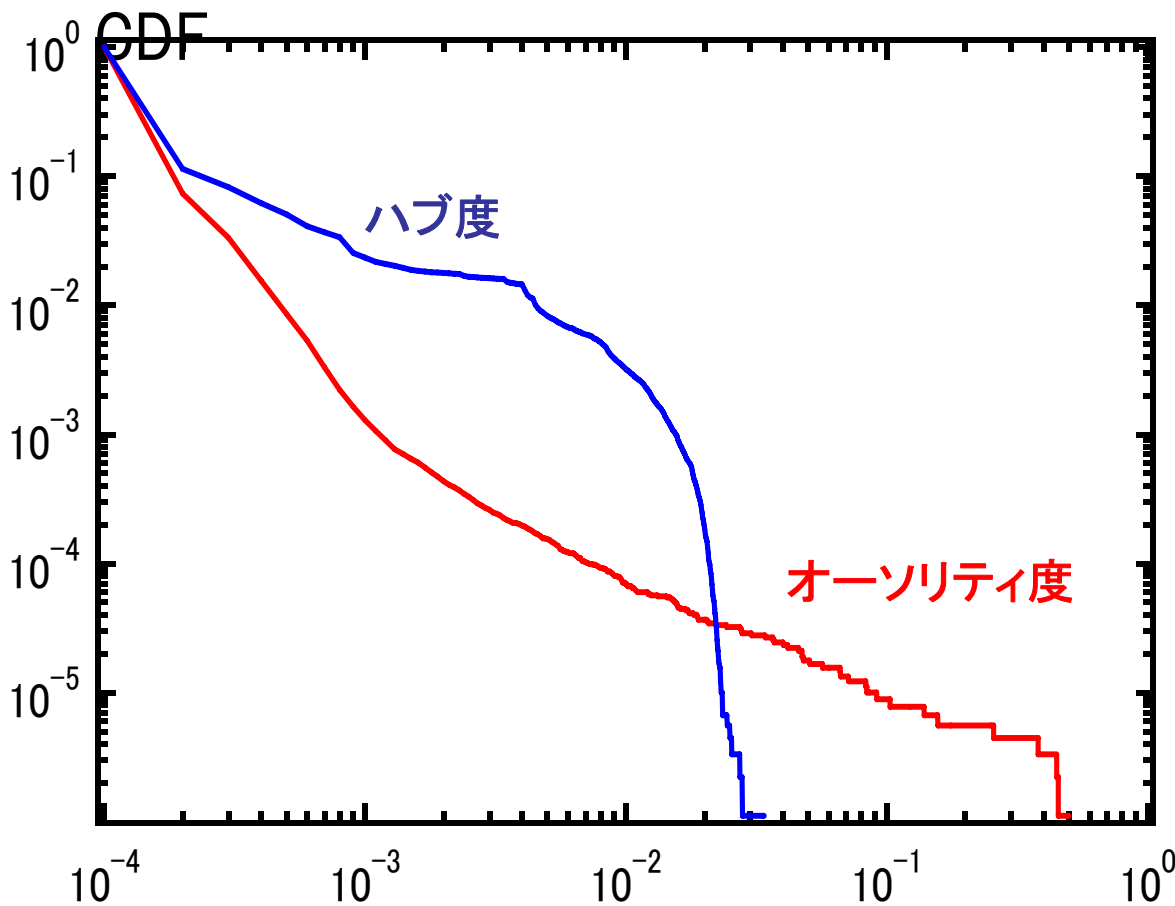
$$\begin{cases} \mathbf{X}(t+1) = \mathbf{A}^T \mathbf{A} \mathbf{X}(t) \\ \mathbf{Y}(t+1) = \mathbf{A} \mathbf{A}^T \mathbf{Y}(t) \end{cases}$$

計算例



番号	オーソリ ティ度	ハブ度
1	0.657	0.001
2	0.001	0.435
3	0.001	0.279
4	0.001	0.279
5	0.001	0.557
6	0.657	0.001
7	0.002	0.435
8	0.001	0.279
9	0.001	0.279
10	0.369	0.001

オーソリティ度とハブ度の分布



オーソリティ度はベキ分布に近いロングテールを持つ分布

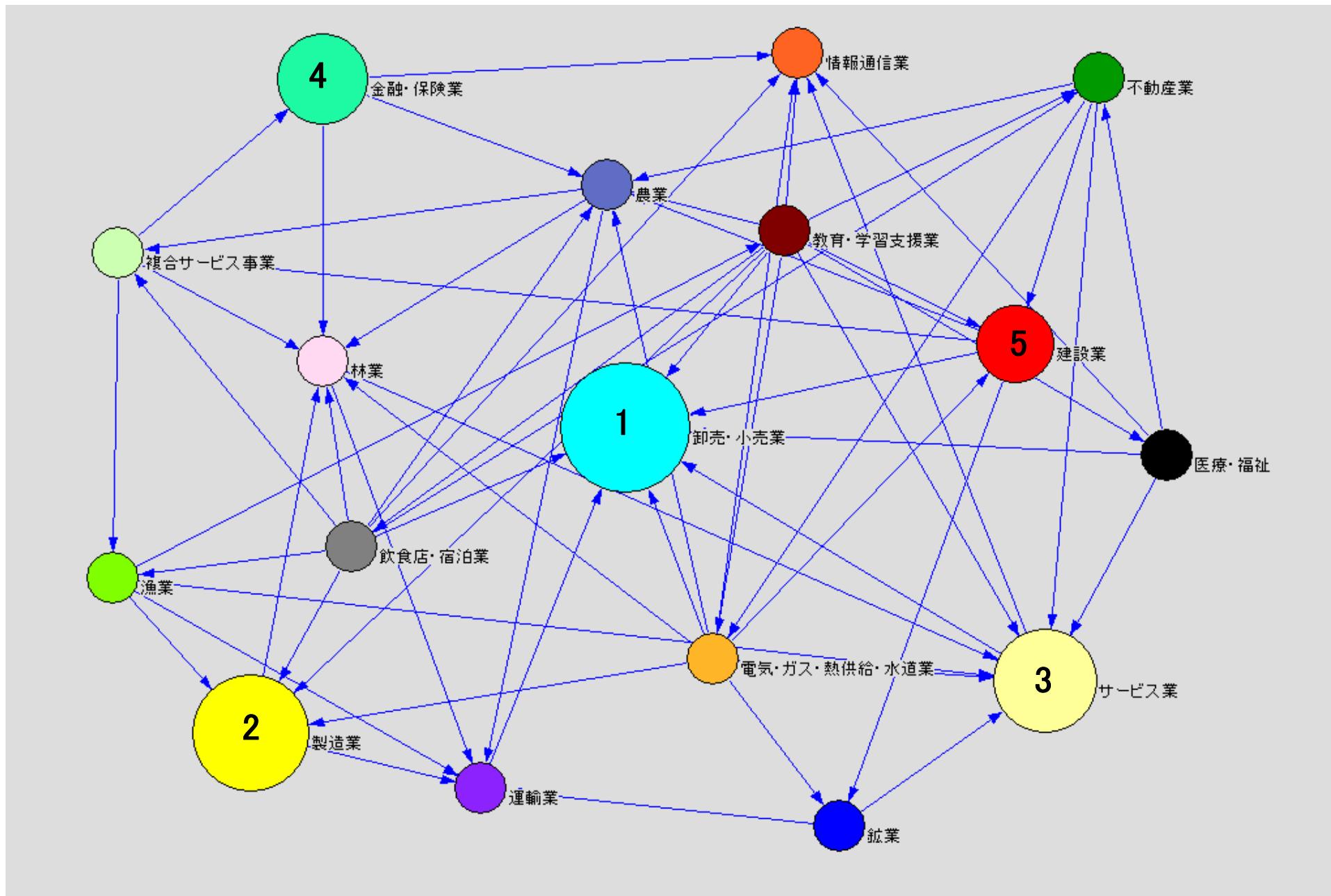
ハブ度はカットオフが入ったようなベキ分布で近似できる

オーソリティ度の上位は建築業が大半を占める

ハブ度の上位は、卸売業やサービス業が多い

順位	大分類番号	業種名	全企業の総売上に対して各業種の占める割合
1	10	卸売・小売業	0.315
2	6	製造業	0.264
3	17	サービス業	0.0931
4	11	金融・保険業	0.0794
5	5	建設業	0.0771
6	9	運輸業	0.0362
7	8	情報通信業	0.0349
8	16	複合サービス業	0.0310
9	12	不動産業	0.0198

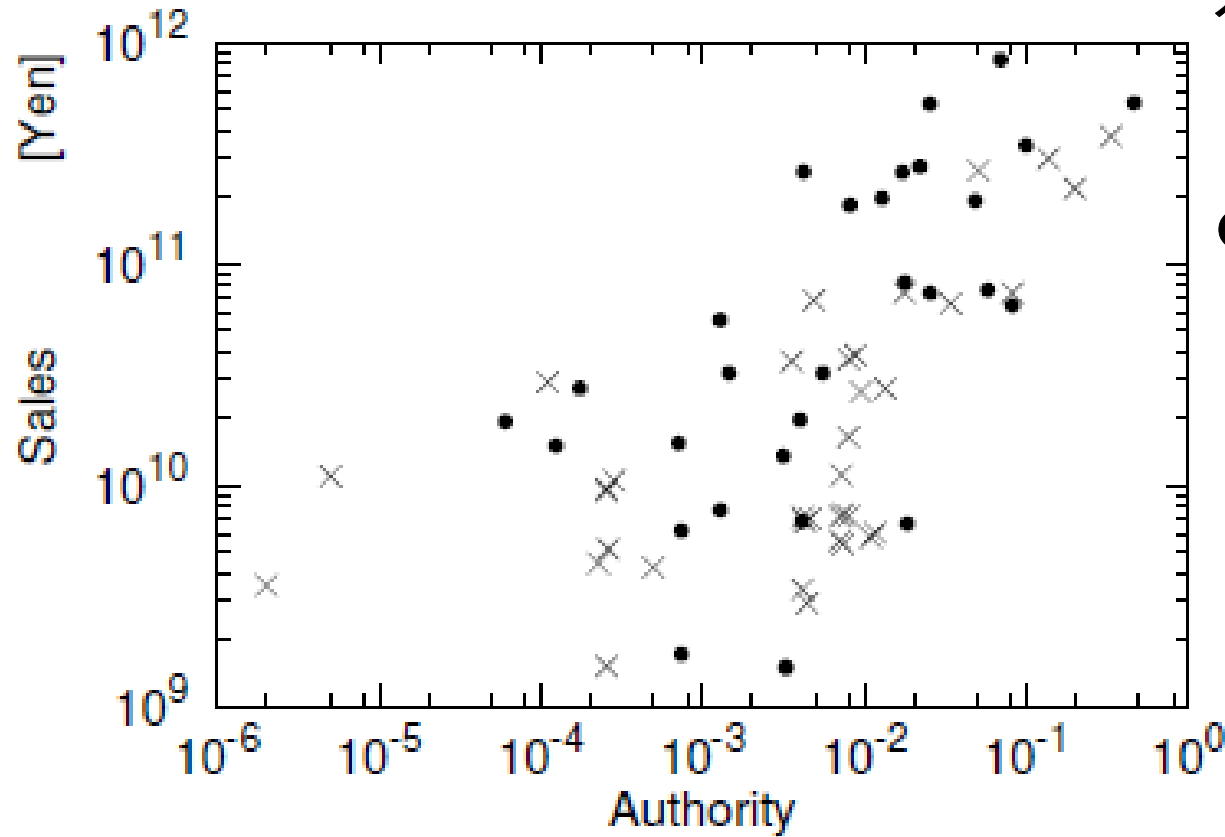
10	14	医療・福祉	0.0172
11	7	電気・ガス・熱供給・水道業	0.0144
12	13	飲食店・宿泊業	0.00854
13	15	教育・学習支援業	0.00533
14	4	鉱業	0.00164
15	1	農業	0.00112
16	3	漁業	0.000586
17	18	公務	0.000160
18	2	林業	0.000108



日本全体における業種間のネットワーク(17分類)構造

リンクの本数:60・売上上位5業種のノードを少し大きくしている

オーソリティ度と売り上げの相関



同一業種内で株の持合のネットワーク構造に対してオーソリティ度を計算

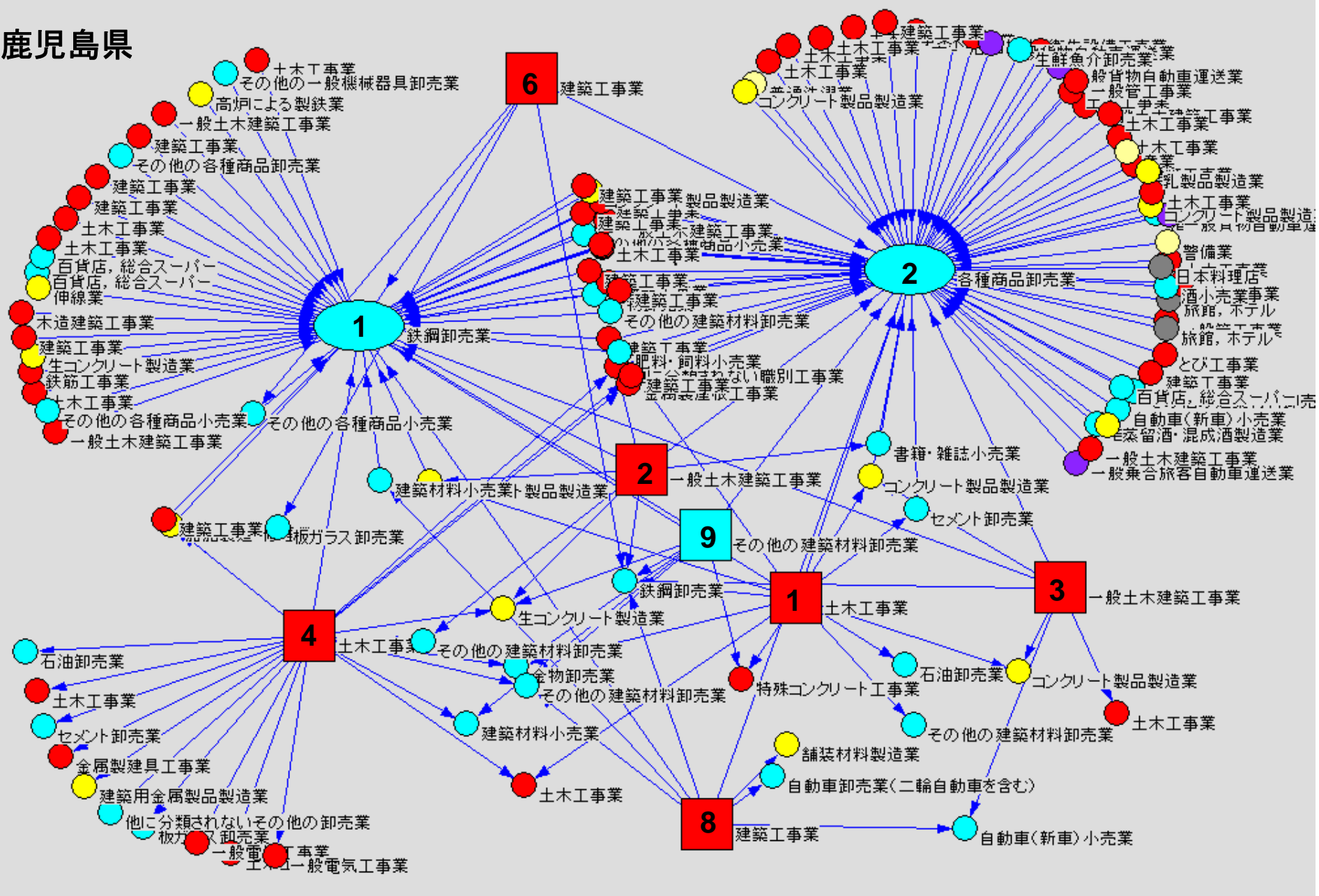
× 放送業

● 非鉄金属製造業

同一業種内で比較するとオーソリティ度が高いほど売り上げも高くなる傾向がある

オーソリティ度やハブ度という量は企業のネットワーク構造の評価に有効

鹿児島県



○ :オーソリティ上位2社
□ :ハブ度上位10 ○:それ以外

赤 :建設業 黄 :製造業 水色 :卸売・小売業

なぜ地球シミュレータが必要か？

$(A_{ij})^n$ 100万行x100万列の行列のベキ乗を計算する必要あり

100万行x100万列の行列の成分数は、 10^{12} 個 \doteq 1TB(テラバイト)
計算を実行するには、メモリーの大きさが数TBは必要。

地球シミュレータを利用すればこの計算は実行可能

この量から様々なネットワークを特徴付ける量が計算できる