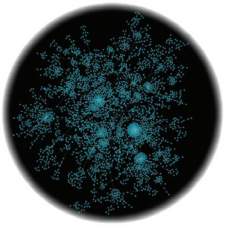
グラフエントロピーによる中心性指標のロバスト性推定
— ノードの中心性指標のノイズに対する順位不変性の推定 —

OSAKA UNIVERSITY
情報科学研究科　情報ネットワーク学専攻　情報流通プラットフォーム講座　博士前期課程　石野 正典

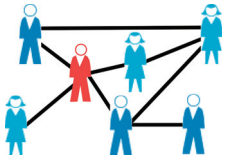
背景

複雑ネットワーク分析



- 実世界に存在する巨大で複雑なネットワークが持つ特徴や性質をグラフ理論を用いて分析
- 複雑ネットワークの例:
  - ソーシャルネットワーク (社会ネットワーク)
  - インターネット
  - 脳や神経細胞などの生物系ネットワーク
  - インフラ網 (送電網、交通網 など)
- 複雑ネットワーク分析では、人や計算機などのエンティティをノード(頂点)、それらの間の関係をリンク (辺) で表現したグラフを用いて分析を行う

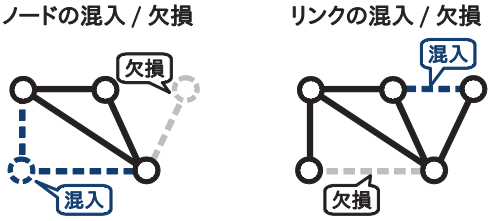
ノードの中心性指標



- ノードが他のノードに及ぼす影響度を測る指標
  - 組織内で重要な役割を持つ人物 の推定や送電網の脆弱性評価などに広く利用されている
  - 数種類の定義が存在
    - 例: 次数中心性、近接中心性、媒介中心性、固有ベクトル中心性

複雑ネットワーク分析おけるノイズ

- 複雑ネットワーク分析において用いられるグラフには、通常様々な種類のノイズが含まれている



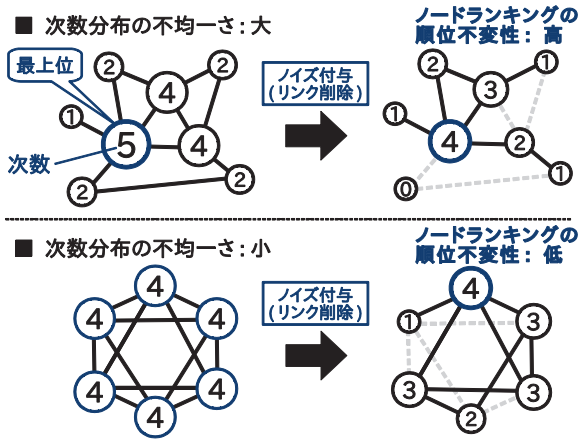
- グラフに含まれるノイズにより、ノードの中心性指標も影響を受ける
  - 中心性指標の高いノードを正しく推定することが困難になる

ノードの中心性指標のノイズに対するロバスト性(耐性) を知りたい

- 中心性指標のロバスト性: ノイズに対するノードランキングの安定度

次数分布の不均一さと中心性指標のロバスト性との関係

- 次数分布が不均一 (均一) なグラフほど、ノードの中心性指標のノイズに対するロバスト性は高い (低い) 傾向を示すと予想される



次数分布の不均一さを知ることは、ノードの中心性指標のノイズに対するロバスト性の推定に有用であると予想される

- グラフの次数分布の不均一さを定量化する指標として、本研究ではグラフエントロピーに着目する

研究の目的

- グラフエントロピーが、ノードの中心性指標のノイズに対するロバスト性を推定するのにどの程度有用であるかを実験により明らかにする
  - ノイズの種類: ランダムなノード / リンクの追加・削除
  - 中心性指標の種類: 近接、媒介、次数、固有ベクトル中心性

グラフエントロピーと、グラフにノイズを加える前後での中心性指標に基づくノードランキングの一致度との関係を実験により調査する

Dehmer のグラフエントロピー

- グラフ $G = (V, E)$ 上での抽象情報量 (G に関する集合 S から正実数 $\mathbb{R}_+$ への関数 $f : S \to \mathbb{R}_+$) に基づく確率 p^f を用いたグラフエントロピー
- S をノード集合 V 、関数 f をノードの次数とするとき、ノード v_i の確率 $p^f(v_i)$ は、以下の式で定義される

$$p^f(v_i) = \frac{f(v_i)}{\sum_{j=1}^{|V|} f(v_j)}$$

- グラフ G の Dehmer のグラフエントロピー $D(G)$ は、以下の式で定義される

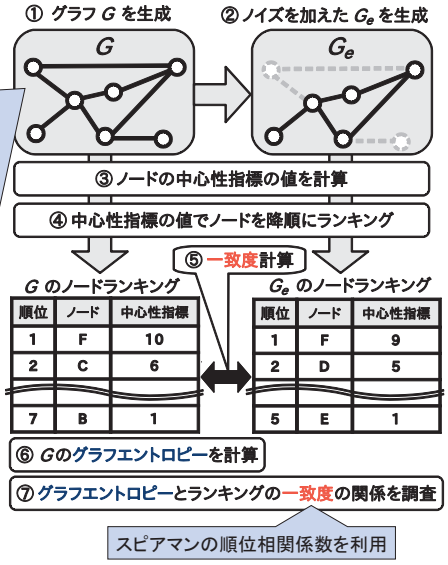
$$D(G) = -\sum_{i=1}^{|V|} p^f(v_i)\log(p^f(v_i))$$

実験方法

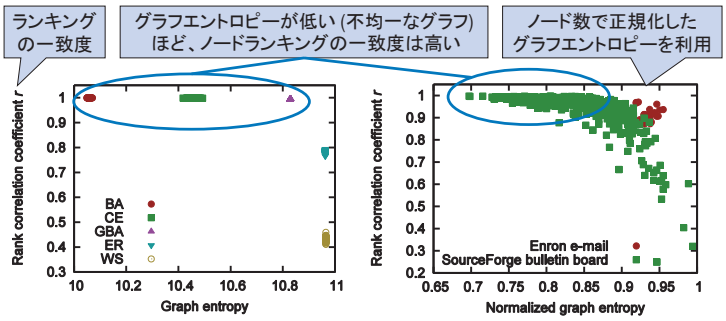
ネットワーク生成モデルにより生成したグラフとソーシャルネットワークをグラフ G として使用

- Barabási – Albert (BA) model
- Community Emergence (CE) model
- Generalized model Barabási – Albert (GBA) model
- Erdös – Rényi (ER) model
- Watts – Strogatz (WS) model

- Enron e-mail (企業における社員間の e-mail の送受信履歴から構築したソーシャルネットワーク)
- SourceForge bulletin board (オンライン掲示板の返信履歴から構築したソーシャルネットワーク)



グラフエントロピーと中心性に基づくノードランキングの一致度との関係



(a) ネットワーク生成モデルにより生成されたグラフ　(b) 実ネットワーク (2種類のソーシャルネットワーク)

図: Dehmer のグラフエントロピーと次数中心性に基づくノードランキングの一致度 r の関係　(ノイズの種類: リンク削除、ノイズレベル: 0.2)

- グラフエントロピーと次数中心性に基づくノードランキングの一致度 r との間に相関が存在することを確認した
- ノイズの種類としてノード / リンクの追加・削除の 4 種類を、ロバスト性の評価を行う中心性指標として近接、媒介、次数および固有ベクトルの 4 種類をそれぞれ用いたとき、おおむね同様の結果を示すことを確認した

グラフエントロピーは、ノードの中心性指標のノイズに対するロバスト性の推定に有用である

今後の課題

- グラフエントロピーの、ランダムでないノイズに対する中心性指標のロバスト性の推定に関する有用性の調査
- 複数種類のノイズがグラフに含まれた際の、グラフエントロピーの、ノイズに対する中心性指標のロバスト性の推定に関する有用性の調査

関連発表文献

[1] 石野 正典, 津川 翔, 大崎 博之, 長谷川 亨, “グラフエントロピーによる中心性指標のロバスト性推定 — ノードの中心性指標のノイズに対する順位不変性の推定 —,” 電子情報通信学会技術研究報告 (CQ2013-101), pp. 71-76, Mar. 2014.