

Osaka University Advanced Network Architecture Research Group  
http://www.anarg.jp/

## 残存次数のエントロピーを用いた トポロジー構造の多様性の評価

シンルー  
舘 璐\*、荒川 伸一\*、村田 正幸\*

\*大阪大学 大学院情報科学研究科

1

Osaka University

### 研究背景

- インターネットの利用形態の多様化
  - 多種多様なプロトコルが複雑に干渉
  - ネットワークの観測情報に基づく最適化設計では対応が困難

↓

**適応性や拡張性を持つネットワークトポロジーの構築が必要**

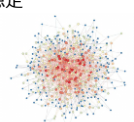
- 適応性や拡張性を保ちつつ、リンクやノードの増強が行える新たなネットワーク設計手法が重要

2


Osaka University

### 研究の目的

- トポロジーが有する構造の多様性を測る指標の評価**
  - 適応性、拡張性の高いネットワーク設計への適用を想定
  - 構造の多様性
    - トポロジーの一部のノード集合を取り出した際に、その集合の接続形態が多様であること
  - 構造の多様性が高い
    - リンクやノードの増強などの環境変化から影響を受けにくい
    - 適応性、拡張性が高い
  - 構造の多様性が低い
    - 現時点での環境に対して最適化されているため、環境変化により性能低下
    - 適応性、拡張性が低い



多様性が高い



多様性が低い

3

Osaka University

### 構造の多様性を測る指標としての相互情報量

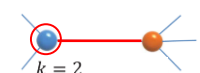
- 情報理論における相互情報量
  - 確率変数Yを知ることにより得られる確率変数Xの情報量
  - $I = H(X) - H_c(X|Y)$ 
    - $H(X)$ : Xのエントロピー
      - Xの不確かさの程度
    - $H_c(X|Y)$ : YのもとでのXの条件付きエントロピー
      - Yの値を知った時のXの値についての不確かさの程度
    - I: 相互情報量
      - 不確かさの減少分
- トポロジーが有する構造の多様性の測定に応用
  - 確率変数Y → トポロジー構造の一部
  - 確率変数X → 残りのトポロジー構造
    - トポロジー構造の一部を知った時、残りのトポロジー構造が不確かな場合、構造は多様

4

Osaka University

### 相互情報量とトポロジー

- 残存次数の相互情報量により、実ネットワークを分析[11]
  - 残存次数  $k$ : リンクを一本取り除いたときに、その片方に接続されていたノードの残り次数
  - リンクの両端にあるノードの次数関係に着目



$k = 2$

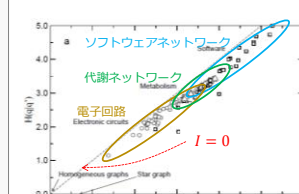
[11] R. Solé and S. Valverde, "Information theory of complex networks: On evolution and architectural constraints," *Complex networks*, vol. 650, pp. 189-207, Aug. 2004.

5

Osaka University

### 様々なネットワークの相互情報量[11]

- 人工ネットワーク、生物ネットワークにおける残存次数の相互情報量を算出



Network type	N	$\langle k \rangle$	$I(q)$	$H(q)$	$H_c(q q)$	r
<b>Technological networks</b>						
<b>人工ネットワーク</b>						
Software 1	168	2.81	1.19	3.04	1.85	-0.39
Software 2	159	4.19	1.03	3.99	2.97	-0.41
Internet AS	3200	3.56	0.50	4.77	4.27	-0.22
Software 3	1993	5.00	0.30	4.82	4.51	-0.08
Circuit TV	320	3.17	0.23	1.37	1.14	0.010
Circuit EC05	899	4.14	0.15	2.98	2.82	-0.15
Software linux	5285	4.29	0.12	4.47	4.35	-0.06
Powergrid	4941	2.67	0.06	3.01	2.95	0.003
<b>Biological networks</b>						
<b>生物ネットワーク</b>						
Silwood park	154	4.75	0.94	4.09	3.14	-0.31
Yihau estuary	134	8.67	0.53	4.74	4.21	-0.24
p53 subnetwork	139	5.09	0.46	4.00	3.54	-0.24
Metabolic map	1174	4.84	0.39	3.58	3.19	-0.17
Neural net (C.elegans)	297	14.5	0.37	5.12	4.74	-0.16
Metabolic map	821	4.70	0.37	3.46	3.09	-0.18
Romanian syntax	5916	5.65	0.31	5.45	5.14	-0.18
Proteome map	1458	2.67	0.24	3.85	3.61	-0.21
<b>Theoretical systems</b>						
<b>生成モデル</b>						
Star graph	17	1.88	1.00	1.00	0.00	-1.00
Barabási-Albert	3000	3.98	0.25	4.12	3.85	-0.078
Erdős-Rényi	300	6.82	0.06	3.31	3.25	-0.005
Module E-R	500	10.3	0.04	3.67	3.62	-0.001

[11] R. Solé and S. Valverde, "Information theory of complex networks: On evolution and architectural constraints," *Complex networks*, vol. 650, pp. 189-207, Aug. 2004.

6

Osaka University 7

### 相互情報量を用いたルーターレベルトポロジーの評価

- ルーターレベルトポロジーは設計理念に基づいて設計される
  - 規則性があり、多様性が低いはず
- ルーターレベルトポロジー<sup>[12]</sup>の多くは相互情報量が大きい
  - 技術的・物理的な制約を考慮した設計に起因
- Verio 社の相互情報量が小さい
  - 小規模な地域ISPの買収を繰り返して規模を拡大したことに起因
    - さまざまな設計指針を内包

ISPのルーターレベルトポロジー	確率的に生成されたトポロジー
Telstra Sprint AT&T Level3 Verio BA Random	
ノード数	329 467 523 623 839 523 523
リンク数	615 1280 1304 5298 1885 1304 1304
H	4.24 4.74 4.46 6.04 4.65 4.24 3.22
Hc	3.11 3.84 3.58 5.42 4.32 3.98 3.15
I	1.13 0.9 0.88 0.63 0.33 0.26 0.07

相互情報量  $I$  が小さい: 構造が多様  
 相互情報量  $I$  が大きい: 構造に規則性が出る  
 $H, H_c$  がトポロジーの性質に与える影響を調査

[12] N. Spring, R. Mahajan, D. Wetherall, and T. Anderson, "Measuring ISP topologies with rocketfuel," IEEE/ACM Transactions on Networking, vol. 12, pp. 2-16, Feb. 2004.

Osaka University 8

### 残存次数のエントロピーと条件付きエントロピー

- 残存次数の相互情報量  $I$  <sup>[11]</sup>
  - $I = H - H_c$
- 残存次数分布  $q(k)$ 
  - 残存次数が  $k$  である確率  $q(k)$
- 残存次数のエントロピー  $H$ 
  - 次数が多様であるトポロジーは  $H$  が大きい
  - $H = -\sum_{k=1}^K q(k) \log(q(k))$
  - トポロジーの次数の不均質性
- 残存次数の条件付きエントロピー  $H_c$ 
  - $H_c = -\sum_{k=1}^K \sum_{k'=1}^K q(k') \pi(k|k') \log(\pi(k|k'))$
  - $\pi(k|k')$ : 残存次数  $k'$  を持つノードと接続されているノードの残存次数が  $k$  である条件付き確率

リンクに接続している片方のノードの次数が与えられたときの、もう片方のノードの次数の不確かさが小さいと  $H_c$  が小さい

次数が均質な正規グラフは  $H$  が0

Abilene-inspired topology <sup>[5]</sup>  
 $I = 1.02 (H = 3.27, H_c = 2.25)$

[5] L. Li, D. Alderson, W. Willinger, and J. Doyle, "A first-principles approach to understanding the Internet's router-level topology," ACM SIGCOMM Computer Communication Review, vol. 34, pp. 3-14, Oct. 2004.

Osaka University 9

### 残存次数のエントロピー $H$ がトポロジーの性質に与える影響

- 平均次数一定の下、残存次数のエントロピーが異なるトポロジーを生成し、平均ホップ長・次数分布を比較
- エントロピー  $H$ 、条件付きエントロピー  $H_c$  が目標値に最も近いトポロジーを生成
  - 初期トポロジー
    - BA モデルで生成 (523ノード、1304リンク)
  - ランダムにリンクを張り替える
    - 平均次数一定
  - 焼きなまし法を用いてポテンシャル関数  $U(G)$  を最小化
    - $U(G) = \sqrt{(H - H(G))^2 + (H_c - H_c(G))^2}$
    - $H(G), H_c(G)$ : 生成過程のトポロジー  $G$  のエントロピー、条件付きエントロピー

Osaka University 10

### 残存次数のエントロピー $H$ と平均ホップ長・次数分布

- エントロピー  $H = 3$  を超えて増加すると、平均ホップ長が減少
  - エントロピー  $H = 4$  付近で次数分布がべき則に近づく

$H = H_c$  となる目標値を与えて生成したトポロジーを用いて評価

ルーターレベルトポロジーは  $H = 4.5$  前後であるため、その時の  $H_c$  がトポロジーに与える影響を調査

Osaka University 11

### 残存次数の条件付きエントロピー $H_c$ がトポロジーの性質に与える影響

- エントロピー  $H$  が大きい値をとるトポロジーで、残存次数の条件付きエントロピー  $H_c$  が異なるトポロジーを生成し、構造の多様性を比較
- エントロピーを大きい値に固定し、条件付きエントロピーが目標値に最も近いトポロジーを生成
  - 初期トポロジー
    - BA モデルで生成 (523ノード、1304リンク) :  $H = 4.24$
  - リンクの張り替え
    - 次数分布を維持した張り替え<sup>[15]</sup>
    - エントロピー  $H$  一定
  - 焼きなまし法を用いてポテンシャル関数  $U^I(G)$  を最小化
    - $U^I(G) = |I - I(G)|$  ( $I(G)$ : 生成過程のトポロジー  $G$  の相互情報量)
    - 条件付きエントロピー  $H_c$  の指定は相互情報量  $I$  の指定に等しい ( $I = H - H_c$ )

[15] P. Mahadevan, D. Krioukov, K. Fall, and A. Vahdat, "Systematic topology analysis and generation using degree correlations," in ACM SIGCOMM Computer Communication Review, vol. 36, pp. 135-146, Oct. 2006.

Osaka University 12

### 相互情報量を最小化したトポロジー

- 残存次数のエントロピー一定の下、相互情報量を最小化したトポロジー
  - 残存次数の条件付きエントロピーの最大化に等しい

$I = 0$  を与えて、生成したトポロジー

構造が多様

- 接続先ノードの残存次数のばらつきが大きい

残存次数が最大のノードの接続先ノードの残存次数分布

残存次数が最小のノードの接続先ノードの残存次数分布

$I(G) = 0.12 (H = 4.24, H_c = 4.13)$

Osaka University 13

### 相互情報量を最大化したトポロジー

- 残存次数のエントロピー一定の下、相互情報量を最大化したトポロジー
- 残存次数の条件付きエントロピーの最小化に等しい

$I = 3$  を与えて、生成したトポロジー

**構造に規則性出現**

- 接続先ノードの残存次数のばらつきが小さい

残存次数最大のノードの接続先ノードの残存次数分布

残存次数が最小のノードの接続先ノードの残存次数分布

$I(G) = 2.70$   
( $H = 4.24, H_c = 1.54$ )

The figure shows a network graph with nodes colored by residual degree. Below it are two histograms. The left histogram shows the distribution of residual degrees for nodes with the highest residual degree, and the right histogram shows the distribution for nodes with the lowest residual degree. Both histograms show a sharp peak at a residual degree of 3, indicating a regular structure.

Osaka University 14

### まとめ

- トポロジーが有する構造の多様性を測る指標の評価
  - 残存次数の相互情報量を用いてルーターレベルトポロジーを評価
    - 設計理念があるため、規則性出現するという予測と合致
  - 残存次数のエントロピー、条件付きエントロピーがトポロジーの性質に与える影響の調査
    - 残存次数のエントロピーが同じとき
      - 残存次数の相互情報量が大いとき、構造が多様
      - 残存次数の相互情報量が小さいとき、構造に規則性出現

**トポロジーが有する構造の多様性を測る指標として相互情報量は有用**

- 今後の課題
  - 相互情報量とネットワーク性能の関係を評価
  - 適応性、拡張性の高いネットワーク設計への応用