

生物ネットワークの使い方

細胞内コンポーネントの相互作用記述

- 代謝ネットワーク
- タンパク質ネットワーク
- 遺伝子ネットワーク

調和振動子

- 概日リズム、蔵本モデル

感染症ネットワーク、パーコレーション

- 相転移

Omics integration

Literature and others

STRING Database (von Mering 2003 *Nature*)

Sequence information

Homologous interaction (Saeed, Deane 2008 *Bioinform*)
Homologous proteins interact similarly.

Visual interaction network

AraNet, 100 million links (for *Arabidopsis thaliana*)

<http://www.functionalnet.org/>
(Lee et al 2010 *Nat Biotech*)

Movie file
<http://www.functionalnet.org/aranet/movie.html>

<http://atted.jp/>
(Obayashi et al. PCP)
Association link from 1500 gene chips

問題点

生物学におけるパスウェイの根拠が不明瞭

Variational method

- Maximize entrainability under constant regularity

Lagrange multiplier λ

$$\mathcal{L}[U] = \mathcal{E}[U] - \lambda \mathcal{V}_T[U]$$

Entrainability $\mathcal{E}[U]$ Regularity $\mathcal{V}_T[U]$

Optimal phase-response curve (PRC)

$$iPRC U_1(\phi) = \frac{\pi^2}{T^2 \lambda} \frac{p(\phi - \psi_M) - p(\phi - \psi_m) \partial F_1(\phi; \rho)}{Q_1(\phi)^2}$$

input signal ϕ related to light entrainment mechanism ρ

本グループの目的

- 数学的「ネットワーク」が、生命科学に真に役立っている例の発見。
- ネットワークが適用できそうな生命現象の発見。
- それらを「文章」として残す。
(統数研の提言として、または論文として。)
- 1月の会合までに文章のドラフトをつくり、再度相談する。

ネットワークモジュール性を測る

- コミュニティ内の辺密度が高く、コミュニティ間の辺が疎であれば良い分割という前提

$$Q = \frac{1}{2E} \sum_{ij} \left[A_{ij} - \frac{k_i k_j}{2E} \right] \delta(c_i, c_j)$$

隣接行列 A_{ij} (頂点 i と j が同じグループに属しているなら) $\delta(c_i, c_j)$ (そうでなければ)

グループのアサインを変えて Q を最大化する。 \leftarrow 任意の次数列で構成されるランダムグラフの結合確率

Low Q \leftarrow \rightarrow High Q

治りにくいがんのシグナルネットワークは高いモジュール性を示す

(Takemoto & Kihara, 2013)

Network Modularity: Low \leftarrow High

Patient Survival Rate: High (more curable) \leftarrow Low (less curable)

ネットワークのモジュール性に注目することで、より効果的な治療戦略が立てられるかもしれない。

Functional Category

(Guimerà & Amaral, Nature, 2005)

適当なアルゴリズム/生物知識でグラフクラスタリング

モジュール内におけるハブらしさ $z_i = \frac{\sigma_{K_i S_i}}{k_i - K_i S_i}$

Participation coefficient, P_i

$$P_i = 1 - \sum_{s=1}^{N_M} \left(\frac{k_{is}}{k_i} \right)$$

モジュール間のコネクタらしさ

異なる役割のハブを見つけることができる

単にハブを狙えば良いという訳ではない。

異なるモジュールを繋げるハブと、モジュール内のハブが存在する。

Party hub: same time and space

Party hub: same time and space

Party hub: different time and/or space

Party hub: same time and space

Party hub: different time and/or space

Characteristic path length (CPL) vs Hubs removed

Party (blue), Date (red), Date (-SPC24) (green)

PLoS Comput Biol 6, e1000817 (2010)

生物の役割の分類ができる

Generalist or Specialist? どの生物をモニタリングするのが効率的?

Module hubs, Network hubs

Peripherals, Connectors

Among module connectivity, c_i

(Olesen et al., 2007)

Q値の問題点

- Resolution limit 解像度限界 (PNAS 104, 36 (2007))
 - 小さなコミュニティを見逃してしまうことがある。
 - 疎なネットワークな場合がおおよそこれに当てはまる。
 - コミュニティ内の次数の合計 k が $< \sqrt{2E}$

m -ノード完全グラフ

$$E = 10 + 10m(m-1)/2$$

$$K = m(m-1)$$

$m > 4$ のとき、それぞれの完全グラフがコミュニティとして認識される。

モジュラリティのより良い定義が必要

何故、圏論？ (1/2)

- 圏とは
 - データ
 - 対象の集まり: A, B, C, \dots
 - 射の集まり: $f: A \rightarrow B, g: B \rightarrow C, \dots$
 - 規則
 - 射は合成でき、結合律を満たす
 - $f: A \rightarrow B, g: B \rightarrow C$ に対して $g \circ f: A \rightarrow C$ が定まる
 - $(h \circ g) \circ f = h \circ (g \circ f)$
 - 各対象は恒等射を持つ
 - $\text{id}_A: A \rightarrow A$
 - $\text{foi}_A = f, \text{id}_A \circ f = f$

何故、圏論？ (2/2)

- 圏 = 有向ネットワーク + α
 - (生物、社会、...) システムを圏で表現
 - Rosen (1958), Ehresmann & Vanbreemersch (2007)
 - しかし、圏の規則(例えば、射が合成できる = 推移性)を満たすようなシステムの側面しか扱えない
- 圏 = 科学の新しい言語 (description)
 - Spivak, 2014, *Category Theory for the Sciences*, MIT Press
 - "I intend to show that category theory is incredibly efficient as a language for experimental design patterns, introducing formality while remaining flexible"
- 圏 = 未知の構造・モデルを発見するための処方箋 (prescription)
 - 「普遍性」
 - 圏論を通じて頂点や矢印の意味を考えることでネットワークの新しい構造やモデルを提案できないか？

圏論による「動的」在り方の表現

相互作用 モノ → モノ

「相互作用は「はたらき」間のインターフェイス」

L_0 「はたらき」(始点側)、「はたらき」(終点側) インターフェイス 表現 M_0

例: 神経ネットワーク 矢印

左Kan拡張 表現 $M \Rightarrow$ 変換 L

「はたらき」 インターフェイス

様々な表現 M, M', \dots グラフでなくともよい

M_0 (もしくは、変換 L_0) は、「モノを「はたらき」と考える」というアイデアのもっとも単純な表現

「動的」在り方の表現における普遍性

表現 M_0 (変換 L_0) は、「モノを「はたらき」と考える」というアイデアのもっとも単純な表現

素朴に考えると インターフェイス = 空集合 恒等表現

「相互作用は「はたらき」間のインターフェイス」に注目すると...

定義 (Haruna, 2013). 与えられた有向ネットワークに対して、「相互作用」は「はたらき」間のインターフェイスであることを表す矢印集合上の写像 ϕ_{ab} が任意の表現 M ごとに定まる。 ϕ_{ab} は矢印集合の分割を引き起こすが、 M_0 に対応する ϕ_{ab} による分割が最も細かい。

ネットワークの二つの在り方と圏論的対称性

在り方	「静的」 	「動的」
変換	矢印を頂点に変換 R_0 $\text{Hom}(G, R_0(H)) \cong \text{Hom}(L_0(G), H)$	頂点を矢印に変換 L_0 $\text{Hom}(L_0(G), H) \cong \text{Hom}(G, R_0(H))$
ネットワーク構造	有向経路 ネットワーク上の伝達に関与	側方経路 ネットワークのコヒーレンスに関与

圏論的普遍性
あらゆる「はたらき」の表現に共通

生物ネットワーク上のトレードオフ関係

矢印 f の有向媒介中心性 (DBC)

$$DBC_f = \frac{1}{\sum_{g: g \neq f} (d_{gh} + 1)} \sum_{g: g \neq f} \frac{m_{fg}}{N_{gh}}$$

ネットワーク上の伝達に関する重要度 (Freeman, 1977)

矢印 f の側方媒介中心性 (LBC)

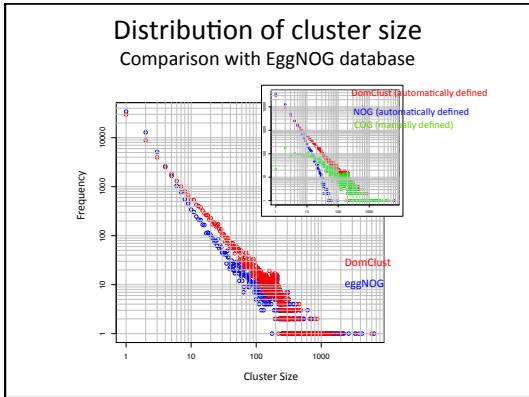
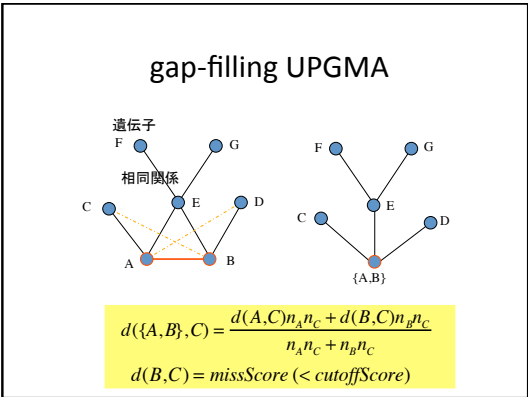
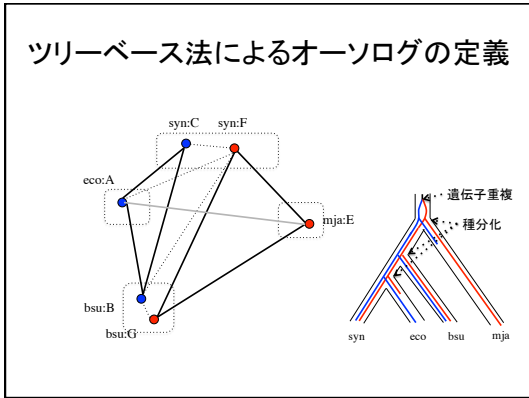
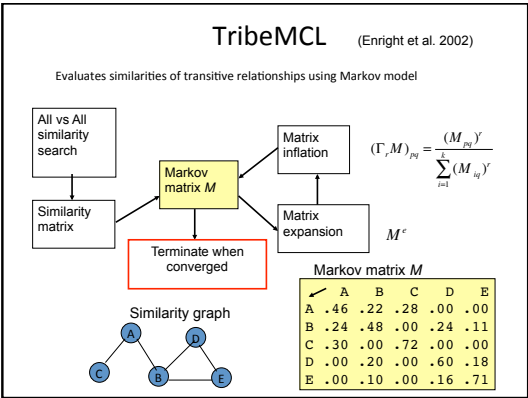
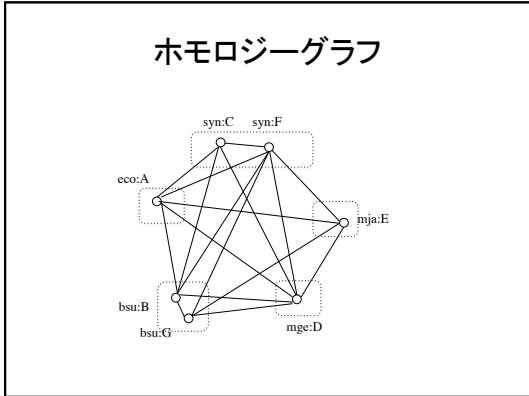
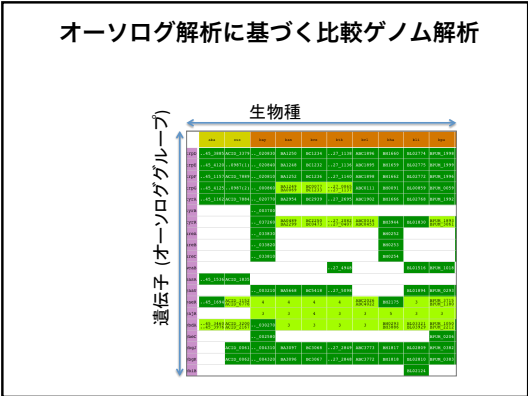
$$LBC_f = \frac{1}{\sum_{g: g \neq f} (m_{gh} + 1)} \sum_{g: g \neq f} \frac{m_{fg}}{m_{gh}}$$

ネットワークのコヒーレンスに関する重要度

生態系ネットワーク 神経ネットワーク 遺伝子ネットワーク

トレードオフ

ネットワーク上の伝達とコヒーレンスに関する分業



Metapopulation SIS dynamics: deterministic reaction-diffusion equation

Change in states of particles at node i ($1 \leq i, j \leq N$)

$$\begin{aligned} \partial_t \rho_{S,i} &= \mu \rho_{I,i} - \frac{\beta_i}{\rho_i} \rho_{S,i} \rho_{I,i} - D_S \sum_j L_{ij} \rho_{S,j}, \\ \partial_t \rho_{I,i} &= -\mu \rho_{I,i} + \frac{\beta_i}{\rho_i} \rho_{S,i} \rho_{I,i} - D_I \sum_j L_{ij} \rho_{I,j}. \end{aligned}$$

recovery infection diffusion

$\beta_i \equiv \beta \rho_i^\gamma$: infection rate at node i ($\gamma \geq 0$)

L : random-walk Laplacian

$$L_{ij} = \begin{cases} 1 & (i = j), \\ -A_{ij}/k_j & (i \neq j). \end{cases} \quad (A: \text{adjacency matrix})$$

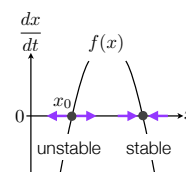
Linear stability of a fixed point

Original equation

$$\frac{d}{dt} x = f(x) \quad x \in \mathbb{R}^N \quad f: \mathbb{R}^N \mapsto \mathbb{R}^N$$

Linearized equation at $x = x_0$

$$\frac{d}{dt} \delta x = J(x_0) \delta x \quad (J(x_0))_{ij} = \left. \frac{\partial f_i}{\partial x_j} \right|_{x=x_0}$$



Formal solution

$$\delta x(t) = \exp(tJ(x_0)) \delta x_0 = \sum_i c_i e^{\lambda_i t} u_i$$

Fixed point $x = x_0$ is linearly unstable
 $\Leftrightarrow \max_{1 \leq i \leq N} \text{Re } \lambda_i (J(x_0)) > 0$

Procedure of analysis

Goal

Derive the endemic threshold β_c (or its upper bound)

Method

Disease-free equilibrium: the fixed point of the reaction-diffusion equations

$$(\rho_{S,i}^*, \rho_{I,i}^*) = \left(\frac{k_i}{\langle k \rangle} \rho, 0 \right) \quad (1 \leq \forall i \leq N)$$

This fixed point is (linearly) unstable \Leftrightarrow endemic equilibrium arises

$$\Leftrightarrow \lambda_{\max}(\beta, \mu, \dots) > 0$$

$\lambda_{\max} = 0$ is solvable in terms of $\beta \rightarrow \beta_c = \dots$ (exact threshold)
 Otherwise $\rightarrow \beta_c < \dots$ (upper bound)

10

An improved upper bound

$$\text{Let } x_i = \begin{cases} 0 & (k_i < k_c), \\ \sqrt{k_i/K_c} & (k_i \geq k_c). \end{cases} \quad K_c \equiv \sum_{j: k_j \geq k_c} k_j$$

total number of links between nodes with $k_i \geq k_c$

$$\begin{aligned} \lambda_{\max}(J^{(2)}) &\geq \max_{k_c} \left[\beta \left(\frac{k_c}{\langle k \rangle} \rho \right)^\gamma + D_I \frac{2E_c}{K_c} \right] - (\mu + D_I) \\ \Rightarrow \beta_c &< \min_{k_c} \left\{ \left[\mu + \left(1 - \frac{2E_c}{K_c} \right) D_I \right] \left(\frac{k_c}{\langle k \rangle} \rho \right)^{-\gamma} \right\} \end{aligned}$$

Implication

More links between large-degree nodes \rightarrow smaller upper bound of β_c

19

Summary

- Metapopulation network + SIS epidemic dynamics
- Target: **endemic threshold** β_c for arbitrary networks
- Upper bound of β_c based on HMF approximation: $\beta_c < (\mu + D_I) \left(\frac{k_{\max}}{\langle k \rangle} \right)^{-\gamma}$
 \rightarrow valid for **arbitrary networks**
- An improvement of upper bound: $\beta_c < \min_{k_c} \left\{ \left[\mu + \left(1 - \frac{2E_c}{K_c} \right) D_I \right] \left(\frac{k_c}{\langle k \rangle} \rho \right)^{-\gamma} \right\}$
 \rightarrow **rich-club networks**: smaller upper bound of β_c
- Considerable improvement with realistic population-dependency γ

Journal reference:

Taro Takaguchi and Renaud Lambiotte,
 "Sufficient conditions of endemic threshold on metapopulation networks"
 Journal of Theoretical Biology **380**, 134-143 (2015).

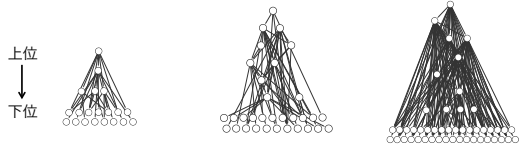
Full-text available upon request: t_takaguchi@nii.ac.jp

23

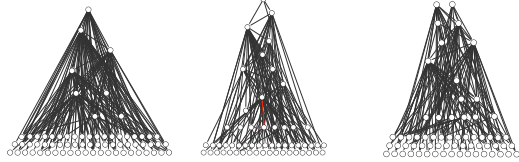
しばしば突き当たること・・・

- community, centrality の "良さ" をどう主張するか?
- ネットワーク全体を取り込む意義?
 (e.g., 次数で OK?)
- temporal net ならでの解析とは?
- 固有値 etc の解析 \rightarrow 構造の言葉へ還元できるのか?

アリの順位行動ネットワーク

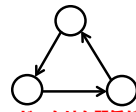


Directed Acyclic Graph (DAG)



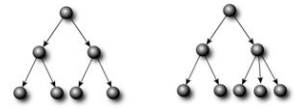
Shimoji, Abe, Tsuji, Masuda (2014)

Directed Acyclic Graph (DAG)

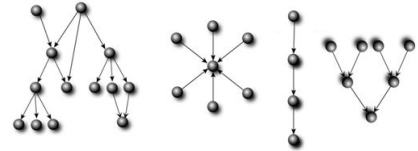


じゃんけん関係がない

⇔ 順番が決まっている

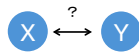
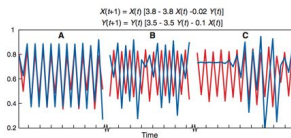


様々な DAG



因果検出

時系列間の因果性検出メソッド



無相関でも因果性がある場合あり

相関ありでも因果性がない場合もある (擬似相関)

Granger causality (Granger, 1969)

- 線形な系に適用
- 予測の良さに基づくので擬似相関も出る

CCM (Convergent Cross Mapping)

- 非線形な系に適用
- 擬似相関も区別

Detecting Causality in Complex Ecosystems

George Sugihara,^{1*} Robert May,² Hao Ye,³ Chih-hao Hsieh,⁴ Ethan Deyle,¹ Michael Fogarty,⁵ Stephan Munch²

Science, 2012

CCM

2つの時系列 X, Y

YからXに因果がある場合

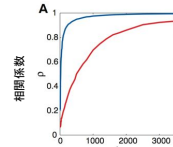
$$x(t) = (X(t), X(t-1), X(t-2))$$

の最近傍点を

$$x(u) = (X(u), X(u-1), X(u-2))$$

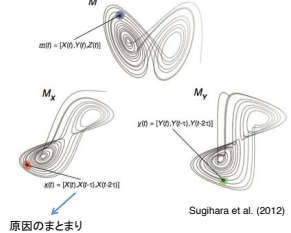
y(t) と y(u) は近傍 (似ている)

Y(t) と Y(u) の相関係数を計算



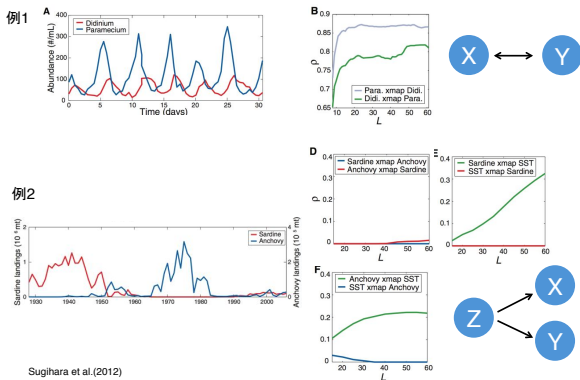
アトラクターの再構成に使う時系列の長さ

埋め込みの定理 (Takens, 1981)



Sugihara et al. (2012)

CCMの結果の例



Sugihara et al. (2012)

まとめ

- ネットワークから生命の理論がつかれるか？
 - 機能・適応性・進化可能性・頑健性
 - ネットワークからしかわからない現象はあるか？
- ダイナミクスとどう絡めるか？
 - 解析する手法も出てきたが