

第 12 回 デジタルサービス・プラットフォーム技術研究会

フェイクニュースとその訂正記事の相互作用
から考えるトイレットペーパー買い占め事件

会田 雅樹

東京都立大学

2023 年 3 月 17 日

背景

ソーシャルネットワーキングサービス (SNS) の社会生活への浸透
⇒ オンラインユーザダイナミクスの影響が実社会の活動にも及ぶ

- 利点

- 欲しい情報が簡単に入手可能
- 個人の情報発信力の強化

- 欠点

- 風評被害, ネットいじめ, 国際的なテロ活動の活性化など
- 社会の分断の加速

特に, フェイクニュースは実社会に大きな影響を与える

フェイクニュースによる社会の分断の例

今日のオンラインソーシャルネットワークでフェイクニュースは深刻さを増している。

- 2016年のアメリカ大統領選挙の期間中、フェイクニュースがアメリカ社会の分断を加速
- トランプ氏に関連した SNS 投稿の約 70%がフェイクニュースに分類される情報^[1]
- フェイクニュースの悪影響を緩和することは緊急の課題

— ★ — ★ — ★ —

[1] A. Guess, B. Nyhan, and J. Reifler, "Selective exposure to misinformation: Evidence from the consumption of fake news during the 2016 U.S. presidential campaign," *European Research Council*, 2018.

フェイクニュースの直観的な対策

フェイクニュース対策として「正しい情報を積極的に宣伝する」ことは有効か？

- 直観的には、正しい情報がフェイクニュースを無効化すると期待できる
- しかしながら、現実はそれほど単純ではない
⇒ 訂正情報が期待通りの効果を発揮しなかった事例が日本で報告されている [2],[3]
- 詳細は次スライド以降...

— ★ — ★ — ★ —

[2] 鳥海不二夫, “データから見るフェイクニュース,” プラットフォームサービスに関する研究会, 総務省, April 2021.

https://www.soumu.go.jp/main_content/000745175.pdf

[3] 日経ビジネスオンライン,

<https://business.nikkei.com/atcl/gen/19/00002/022801115/>

トイレットペーパー買い占め事例 (1)

事件の顛末:

- 2020年2月下旬のCOVID-19流行初期に、トイレットペーパーが品薄の恐れがあるという主張がSNSに掲載された。主張は、中国のトイレットペーパーの生産地域がCOVID-19によって悪影響を受けたため、トイレットペーパーの不足が差し迫っていたというもの
- 実際には日本には十分な在庫と生産能力があり、通常の需要では在庫不足の発生はありえない状況だった。したがって、この投稿は事実と反するフェイクニュース
- この時点で、このフェイクニュース自体はあまり注目されておらず、トイレットペーパーの買いだめは発生していない
- その後、このフェイクニュースに対応するため、「トイレットペーパーは不足しておらず、先の投稿はフェイクニュースである」とした訂正情報が発信された
(続く)

トイレットペーパー買い占め事例 (2)

事件の顛末 (続き):

- 訂正情報を受け取った多くのユーザは「他の多くの人々がトイレットペーパーを買い求めるはず」と考え、その結果「トイレットペーパーが品薄になるかもしれない」との危機感を煽る結果となった
- 更に、訂正情報によって元のフェイクニュースに注目が集まり、その結果、フェイクニュースの拡散が促進された
- その結果、皮肉にも、フェイクニュースを訂正する情報がきっかけで、トイレットペーパーの買いだめが発生し、品薄状態が現実になった

この事件は「正しい情報の発信が必ずしも良い結果につながるとは限らない」という点で、フェイクニュースに対抗することの難しさ再認識させるもの

フェイクニュースの効果的な対策に向けて

- 訂正情報を広めることが、必ずしも期待される効果を生み出すとは限らない
- 寧ろ、訂正情報の発信が状況を悪化させる可能性がある
- その原因は、フェイクニュースと訂正情報の間の相互作用にある
- 適切なフェイクニュース対策技術のためには、まずこの相互作用の理解が不可欠
- 既存のフェイクニュースの拡散モデルは、この相互作用を考慮していない

関連研究

フェイクニュース拡散の既存モデルの分類：

- ユーザ自身の特性に基づくモデル [4],[5]
- ユーザを取り巻く環境に基づくモデル [6],[7],[8].

— ★ — ★ — ★ —

[4] D.C. Brody and D.M. Meier, “How to model fake news,” *arXiv*, arXiv:1809.00964, 2018.

[5] L. Cheng, R. Guo, K. Shu, H. Liu, “Causal understanding of fake news dissemination on social media,” *arXiv*, arXiv:2010.10580, 2020.

[6] P. Törnberg, “Echo chambers and viral misinformation: Modeling fake news as complex contagion,” *PLOS ONE*, 2018..

[7] H. Peng, A. Nematzadeh, D.M. Romero, and E. Ferrara, “Network modularity controls the spread of information diffusion,” *Physical Review E*, vol. 102, 052316, 2020.

[8] Q.F. Lotito, D. Zanella, and P. Casari, “Realistic aspects of simulation models for fake news epidemics over social networks” *Future Internet*, vol. 13, no. 3, 76, 2021.

本研究の目的

- フェイクニュースとその訂正情報との相互作用をモデル化すること
- フェイクニュースに影響されるユーザのクラスタの発生メカニズムを理解すること
- 効果的なフェイクニュース対策の指針を得ること

本研究の基本的な立場

- フェイクニュースを 知っている or 知らない
- フェイクニュースと訂正情報を両方知っている前提で
フェイクニュースを信じている vs 訂正情報を信じている
- フェイクニュースを「知っている or 信じている」かどうかは
問題ではなく、ユーザの行動内容が重要
フェイクニュースに影響を受けている vs 影響を受けていない

本研究は三番目の立場：

トイレットペーパーを買いだめした人も、必ずしも元のフェイクニュースを信じているわけではなく、品薄状態が予想されたための行動であったはず

反応拡散系

- 簡単な例：二次元座標 (x, y) で表される平面に二種類の対象が分布
- 時刻 t におけるそれらの密度関数をそれぞれ $a(x, y; t)$, $h(x, y; t)$ とする
- 反応拡散方程式：

$$\frac{\partial a(x, y; t)}{\partial t} = f(a, h; t) + \kappa_a \Delta a(x, y; t)$$

$$\frac{\partial h(x, y; t)}{\partial t} = g(a, h; t) + \kappa_h \Delta h(x, y; t)$$

- $f(a, h; t)$ と $g(a, h; t)$: 反応項 (二種類の対象の相互作用を記述)
- $\kappa_a \Delta a(x, y; t)$, と $\kappa_h \Delta h(x, y; t)$: 拡散項 (二種類の対象の空間的な拡散を記述)
- κ_a と κ_h : 拡散係数 (拡散の強さを記述)

活性因子・抑制因子系

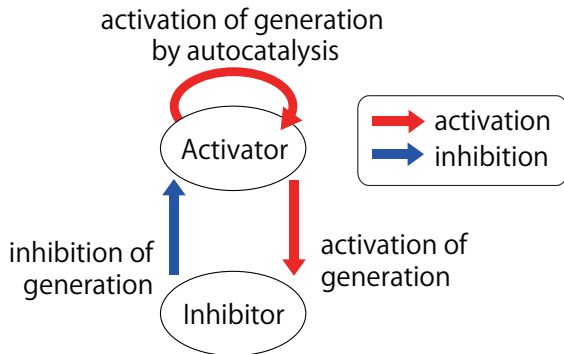
- 活性因子・抑制因子モデルは反応項を以下のように与えた反応拡散モデルの一種

$$\begin{aligned}f(a, h; t) &= c_a a(x, y; t) - b h(x, y; t) - \mu_a a(x, y; t) + \rho_a, \\g(a, h; t) &= c_h a(x, y; t) - \mu_h h(x, y; t) + \rho_h,\end{aligned}$$

- ここで $c_a, c_h, b, \mu_a, \mu_h, \rho_a, \rho_h$ は正の定数
- $a(x, y; t)$ は活性因子, $h(x, y; t)$ は抑制因子の密度関数
- 反応項の働きは以下の通り：
 - 活性因子の存在は, その場所での活性因子と抑制因子の両方の生成を促進する
 - 抑制因子の存在は, その場所での活性因子の生成を阻害する

反応項の働き

- 反応項によって引き起こされる活性因子と抑制因子の関係

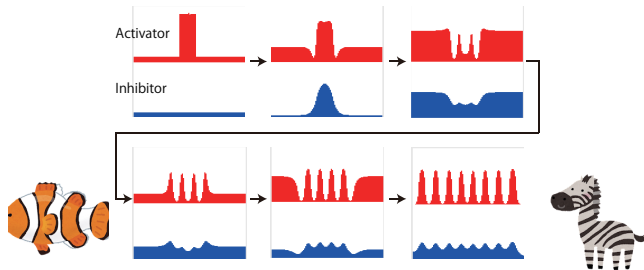


拡散係数の大小関係とチューリングパターン

- 抑制因子は活性因子よりも速く空間的に拡散する

$$\kappa_h > \kappa_a,$$

- チューリングパターン:



活性因子・抑制因子モデルにより発生するチューリングパターン

- シマウマの模様，熱帯魚の模様など
- なぜ体表の細胞は自分が何色に変化すべきか分かるのか？

フェイクニュースと訂正情報の相互作用 (1)

フェイクニュースと訂正情報の相互作用をモデル化するための前提：

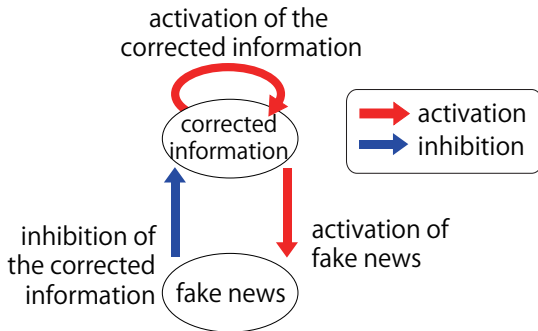
- 一般に、フェイクニュースはユーザの興味を惹き付ける情報を扱うので、通常のニュースよりも拡散が速い [9]
- 訂正情報の広がりには、その本来の目的に反して、トイレットペーパーが多くの人に買われるかもしれないというユーザーの不安を煽る効果がある。このため、訂正情報は本来の意図を伝えるだけでなく、フェイクニュースにも注目を集める効果を持つ、
- フェイクニュースが広まると、フェイクニュースの内容が知られるようになり、それを否定する訂正情報を広めることが難しくなる。

— ★ — ★ — ★ —

[9] S. Vosoughi, D. Roy, and S. Aral, “The spread of true and false news online,” *Science*, vol. 359, issue 6380, p. 1146, 2018.

フェイクニュースと訂正情報の相互作用 (2)

- フェイクニュースと訂正情報の間の相互作用は、活性因子・抑制因子モデルで記述できる。



フェイクニュースと訂正情報の相互作用

活性因子 = 訂正情報 & 抑制因子 = フェイクニュース

ネットワーク上の活性因子・抑制因子系

- $a_i(t)$ と $h_i(t)$: それぞれ時刻 t におけるノード i ($i = 1, 2, \dots, n$) の活性因子と抑制因子の量
- 活性因子と抑制因子に対応する二つの状態ベクトルを導入 :

$$\begin{aligned} \mathbf{a}(t) &:= (a_1(t), a_2(t), \dots, a_n(t))^\top, \\ \mathbf{h}(t) &:= (h_1(t), h_2(t), \dots, h_n(t))^\top, \end{aligned}$$

- ネットワーク上の活性因子・抑制因子モデル :

$$\begin{aligned} \frac{\partial}{\partial t} \mathbf{a}(t) &:= \mathbf{F}(\mathbf{a}(t), \mathbf{h}(t)) - \kappa_a \mathbf{L} \mathbf{a}(t), \\ \frac{\partial}{\partial t} \mathbf{h}(t) &:= \mathbf{G}(\mathbf{a}(t), \mathbf{h}(t)) - \kappa_h \mathbf{L} \mathbf{h}(t). \end{aligned}$$

- $\mathbf{F}(\mathbf{a}(t), \mathbf{h}(t))$ と $\mathbf{G}(\mathbf{a}(t), \mathbf{h}(t))$ は反応項を表すベクトル値関数
- $-\kappa_a \mathbf{L} \mathbf{a}(t)$ と $-\kappa_h \mathbf{L} \mathbf{h}(t)$ は拡散項

反応項と局所相互作用

- 反応項の詳細：

$$F(\mathbf{a}(t), \mathbf{h}(t)) := C_a \mathbf{a}(t) - B \mathbf{h}(t) - M_a \mathbf{a}(t) + R_a \mathbf{1},$$

$$G(\mathbf{a}(t), \mathbf{h}(t)) := C_h \mathbf{a}(t) - M_h \mathbf{h}(t) + R_h \mathbf{1}.$$

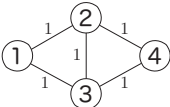
- $C_a, C_h, D, M_a, M_h, R_a, R_h$ は全て対角行列で与える.
- 反応項に現れる行列が全て対角行列 = 反応項が局所相互作用を表す
- $\mathbf{1}$ は成分が全て 1 の列ベクトル.

$$\mathbf{1} = (1, \dots, 1)^\top$$

拡散項とラプラシアン行列

- 拡散項はラプラシアン行列を用いて記述される
- ラプラシアン行列の定義：次数行列と隣接行列の差

$$L := D - A$$



$$D = \begin{pmatrix} 2 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \quad A = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad L = \begin{pmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{pmatrix}$$

ラプラシアン行列の例

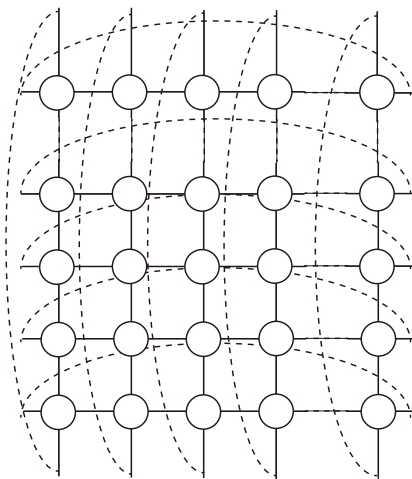
- フェイクニュースの拡散は訂正情報より速いという仮定から

$$\kappa_h > \kappa_a$$

数値評価の目的

- ネットワーク構造には一切の偏りが無い理想的な条件下で、フェイクニュースに影響されるユーザーのクラスタ構造が発生する可能性があるか？
- これを確認するために、非常に高い対称性を持つネットワークモデルを用いて評価
- 偏りがなく、高い対称性を持つネットワークモデル：
 - 100×100 格子状ネットワークトポロジー
 - トーラス状の周期境界条件
- ネットワーク上のどのノードも同じ条件

ネットワークモデル



トーラス境界を持つ 100×100 格子ネットワーク

反応項のモデル

- 反応項の詳細：

$$F(a(t), h(t)) := C_a a(t) - B h(t) - M_a a(t) + R_a \mathbf{1},$$

$$G(a(t), h(t)) := C_h a(t) - M_h h(t) + R_h \mathbf{1}.$$

- $C_a, C_h, D, M_a, M_h, R_a, R_h$ は全て対角行列で以下のように与えられる

$$C_a := c_a I, \quad B := b I, \quad M_a := \mu_a I,$$

$$C_h := c_h I, \quad M_h := \mu_h I, \quad R_a := \rho_a I, \quad R_h := \rho_h I,$$

ここで、 I は $n \times n$ 単位行列。

- 偏りが無いネットワーク構造にするため、全てのノードに共通のパラメータ値を設定。

数値例: パラメータ設定

- リンクの重み：全て 1.
- 活性因子・抑制因子モデルのパラメータ設定

c_a	s_a	μ_a	ρ_a	κ_a	c_h	μ_h	ρ_h	κ_h
0.08	0.08	0.03	0.0015	0.02	0.11	0.12	0.0005	0.5

- フェイクニュースの拡散の速さを反映して $\kappa_h > \kappa_a$
- 実験は、大阪大学・近藤研究室の「反応拡散系シミュレータ」を利用^[10]

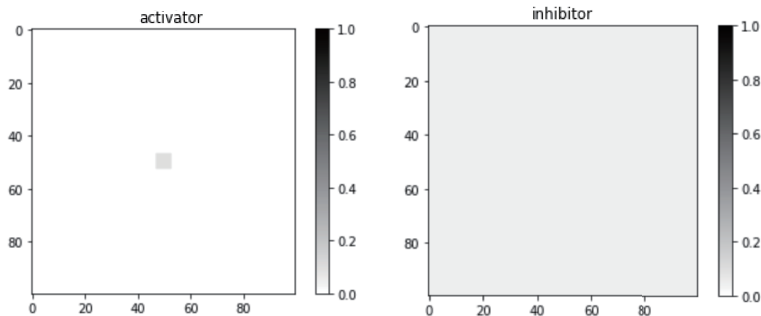
— ★ — ★ — ★ —

[10] *Reaction-Diffusion System Simulator*, Kondo Laboratory, Osaka University.

www.fbs.osaka-u.ac.jp/labs/skondo/simulators/rd_new.html

初期条件 (1)

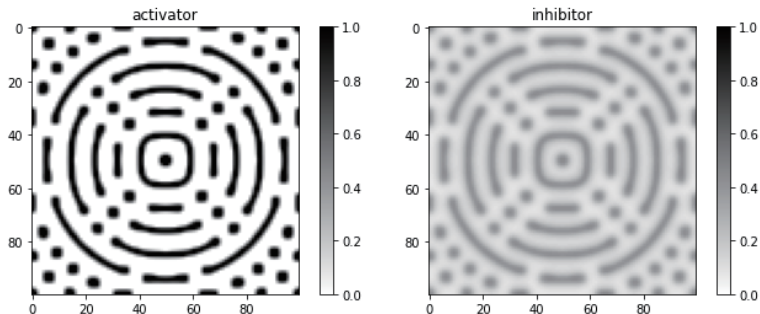
- フェイクニュースが先に蔓延していて、訂正情報が後から発信



濃い色：値が大きい
薄い色：値が小さい

数値例 (1): 評価結果

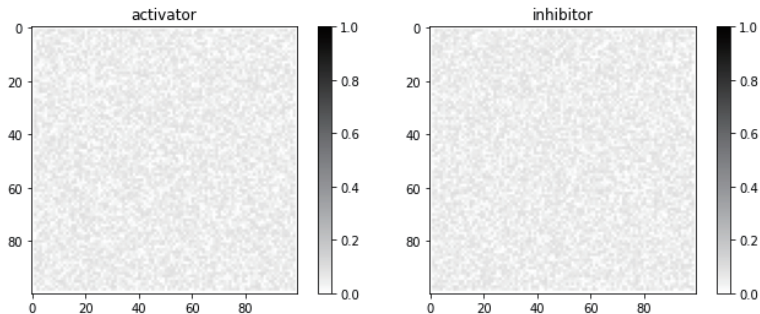
- フェイクニュースと訂正情報の空間分布



フェイクニュースと訂正情報の定常分布

初期条件 (2)

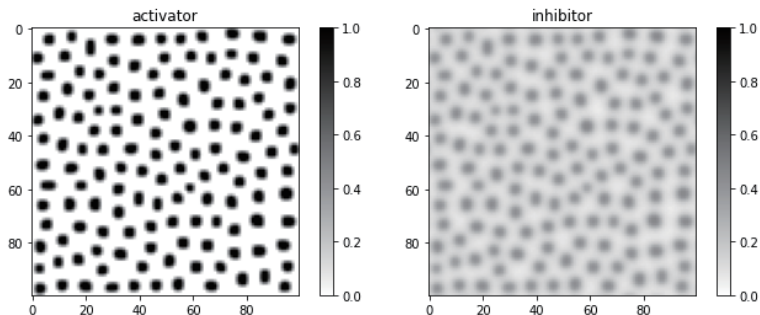
- マスコミ等の影響でフェイクニュースと訂正情報がランダムに周知



濃い色：値が大きい
薄い色：値が小さい

数値例 (2): 評価結果

- フェイクニュースと訂正情報の空間分布



フェイクニュースと訂正情報の定常分布

フェイクニュース対策

フェイクニュースの影響が大きなユーザークラスターの発生を防ぐ方法

⇒ チューリングパターンが発生しない条件を実現すること

- 例えば、拡散係数の大きさの関係を逆転、または差の減少
- この条件は、訂正情報の拡散の速度を上げることに対応

Turing 不安定性の条件 (1) : 拡散がないときの安定性

- 拡散がないとき

$$\frac{\partial}{\partial t} a_i(t) = f(a_i(t), h_i(t)), \quad \frac{\partial}{\partial t} h_i(t) = g(a_i(t), h_i(t)).$$

- 安定点

$$\frac{\partial \bar{a}}{\partial t} = f(\bar{a}, \bar{h}) = 0, \quad \frac{\partial \bar{h}}{\partial t} = g(\bar{a}, \bar{h}) = 0,$$

の周りで線形化.

- 以下の行列の固有値が二つとも負でなければならない.

$$\frac{\partial}{\partial t} \begin{pmatrix} \Delta a \\ \Delta h \end{pmatrix} = \begin{bmatrix} \frac{\partial f(\bar{a}, \bar{h})}{\partial a} & \frac{\partial f(\bar{a}, \bar{h})}{\partial h} \\ \frac{\partial g(\bar{a}, \bar{h})}{\partial a} & \frac{\partial g(\bar{a}, \bar{h})}{\partial h} \end{bmatrix} \begin{pmatrix} \Delta a \\ \Delta h \end{pmatrix}.$$

Turing 不安定性の条件 (1) : (続き)

- 拡散がないときに安定する条件

$$\frac{\partial f(\bar{a}, \bar{h})}{\partial a} + \frac{\partial g(\bar{a}, \bar{h})}{\partial h} < 0, \quad (1)$$

$$\frac{\partial f(\bar{a}, \bar{h})}{\partial a} \frac{\partial g(\bar{a}, \bar{h})}{\partial h} - \frac{\partial f(\bar{a}, \bar{h})}{\partial h} \frac{\partial g(\bar{a}, \bar{h})}{\partial a} > 0. \quad (2)$$

Turing 不安定性の条件 (2) : 拡散があるときの不安定性

- 拡散項のラプラシアン行列を対角化

$${}^t P L P = \Lambda := \text{diag}(\lambda_1, \dots, \lambda_n).$$

- それに合わせて時間発展方程式を変換

$$\begin{aligned} \frac{\partial}{\partial t} ({}^t P \mathbf{a}(t)) &= F({}^t P \mathbf{a}(t), {}^t P \mathbf{h}(t)) - \kappa_a \Lambda ({}^t P \mathbf{a}(t)), \\ \frac{\partial}{\partial t} ({}^t P \mathbf{h}(t)) &= G({}^t P \mathbf{a}(t), {}^t P \mathbf{h}(t)) - \kappa_h \Lambda ({}^t P \mathbf{a}(t)). \end{aligned}$$

- 変換された状態ベクトル

$$\begin{aligned} ({}^t P \mathbf{a}(t)) &:= ({}^t \tilde{a}_1(t), \dots, \tilde{a}_n(t)), \\ ({}^t P \mathbf{h}(t)) &:= ({}^t \tilde{h}_1(t), \dots, \tilde{h}_n(t)), \end{aligned}$$

Turing 不安定性の条件 (2) : (続き)

- 成分ごとの方程式

$$\begin{aligned}\frac{\partial}{\partial t} \tilde{a}_i(t) &= f(\tilde{a}_i(t), \tilde{h}_i(t)) - \kappa_a \lambda_i \tilde{a}_i(t), \\ \frac{\partial}{\partial t} \tilde{h}_i(t) &= g(\tilde{a}_i(t), \tilde{h}_i(t)) - \kappa_h \lambda_i \tilde{h}_i(t).\end{aligned}$$

- 一様解の周りで線形化

$$\frac{\partial}{\partial t} \begin{pmatrix} \Delta \tilde{a}_i \\ \Delta \tilde{h}_i \end{pmatrix} = \begin{bmatrix} \frac{\partial f(\tilde{a}_i, \tilde{h}_i)}{\partial a} - \kappa_a \lambda_i & \frac{\partial f(\tilde{a}_i, \tilde{h}_i)}{\partial h} \\ \frac{\partial g(\tilde{a}_i, \tilde{h}_i)}{\partial a} & \frac{\partial g(\tilde{a}_i, \tilde{h}_i)}{\partial h} - \kappa_h \lambda_i \end{bmatrix} \begin{pmatrix} \Delta \tilde{a}_i \\ \Delta \tilde{h}_i \end{pmatrix}.$$

- 上記に現れる行列の固有値の少なくとも一方が正。

Turing 不安定性の条件 (2) : (続き 2)

- 拡散のあるときの不安定性から得られる条件

$$\kappa_h \frac{\partial f(\tilde{a}, \tilde{h})}{\partial a} + \kappa_a \frac{\partial g(\tilde{a}, \tilde{h})}{\partial h} > 0, \quad (3)$$

$$\left(\kappa_h \frac{\partial f(\tilde{a}, \tilde{h})}{\partial a} + \kappa_a \frac{\partial g(\tilde{a}, \tilde{h})}{\partial h} \right)^2 - \kappa_a \kappa_h \left(\frac{\partial f(\tilde{a}, \tilde{h})}{\partial a} \frac{\partial g(\tilde{a}, \tilde{h})}{\partial h} - \frac{\partial f(\tilde{a}, \tilde{h})}{\partial h} \frac{\partial g(\tilde{a}, \tilde{h})}{\partial a} \right) > 0. \quad (4)$$

- その他の情報

$$\frac{\partial f}{\partial a} = c_a - \mu_a, \quad \frac{\partial f}{\partial h} = -s_a, \quad \frac{\partial g}{\partial a} = c_h, \quad \frac{\partial g}{\partial h} = -\mu_h.$$

Turing パターンが発生しない条件

- Turing パターンが発生しないための訂正情報の拡散係数 κ_a の範囲

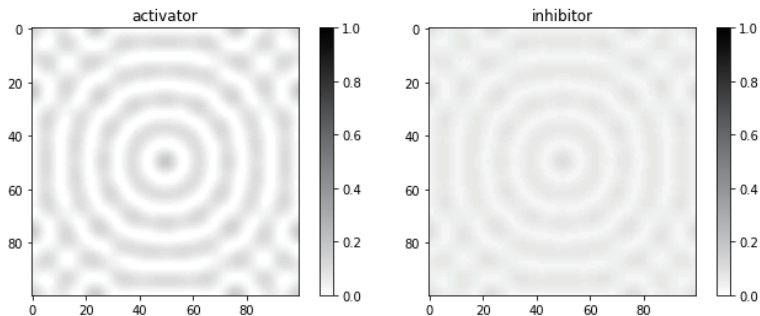
$$\kappa_a \geq \kappa_h \frac{1}{\mu_h^2} \left[- (c_a - \mu_a) \mu_h + 2 s_a c_h - 2 \sqrt{s_a^2 c_h^2 - s_a c_h (c_a - \mu_a) \mu_h} \right].$$

- 今回の条件では

$$\kappa_a \simeq 0.058. \quad (5)$$

フェイクニュース対策の数値例 (1-1)

- フェイクニュースと訂正情報の空間分布

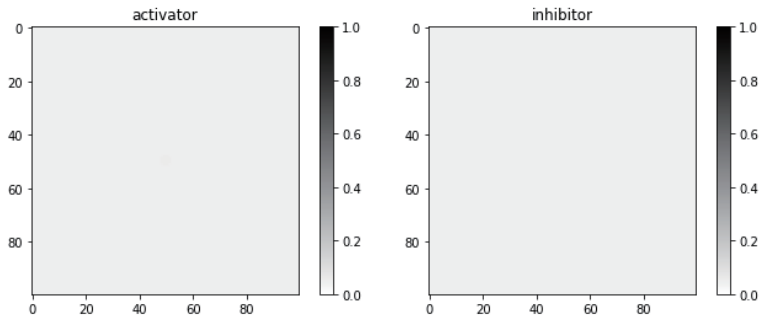


The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of CENTER with

$$\kappa_a = 0.057$$

フェイクニュース対策の数値例 (1-2)

- フェイクニュースと訂正情報の空間分布

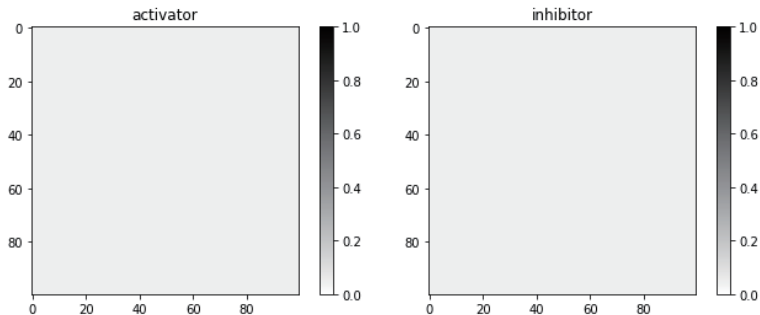


The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of CENTER with

$$\kappa_a = 0.058$$

フェイクニュース対策の数値例 (1-3)

● フェイクニュースと訂正情報の空間分布

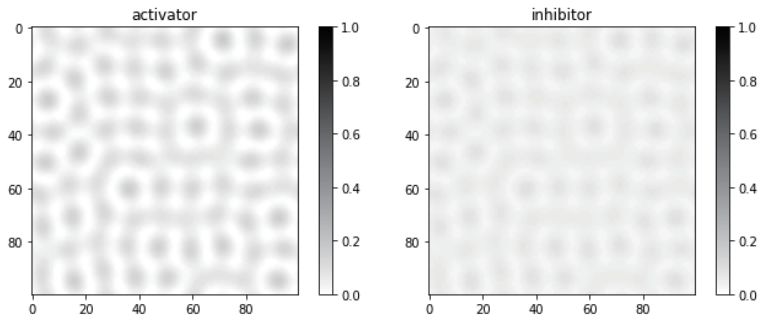


The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of CENTER with

$$\kappa_a = 0.059$$

フェイクニュース対策の数値例 (2-1)

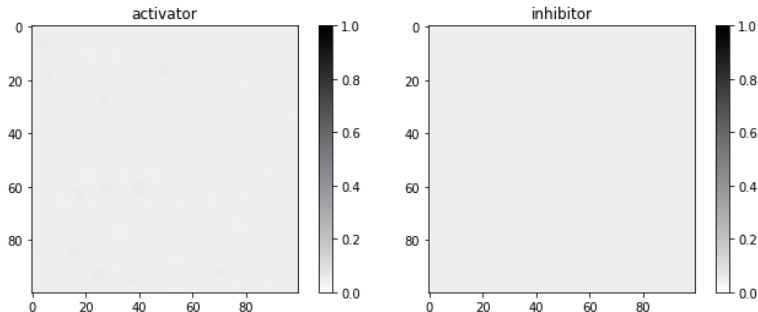
● フェイクニュースと訂正情報の空間分布



The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of RANDOM with $\kappa_a = 0.057$

フェイクニュース対策の数値例 (2-2)

● フェイクニュースと訂正情報の空間分布

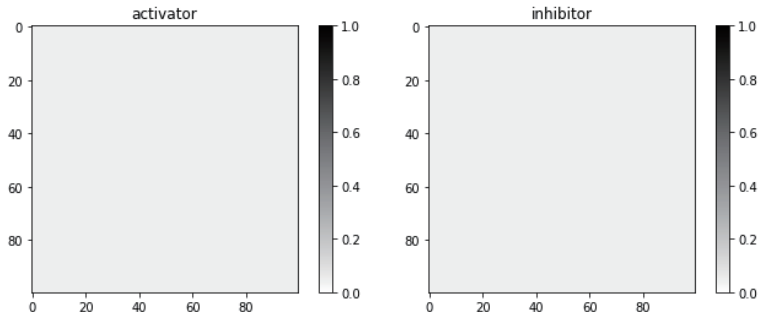


The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of RANDOM with

$$\kappa_a = 0.058$$

フェイクニュース対策の数値例 (2-3)

- フェイクニュースと訂正情報の空間分布



The distributions of the strengths of the correction (left) and the fake news (right) for the initial condition of RANDOM with

$$\kappa_a = 0.059$$

結論

- フェイクニュースを修正するために発信された訂正情報は、必ずしも期待される効果をもたらすとは限らず、逆にフェイクニュースの影響を悪化させる可能性
- この現象を理解するには、フェイクニュースとその訂正情報との相互作用の理解が必要
- 活性因子・抑制因子モデルに基づく相互作用モデルを提案し、フェイクニュースと訂正情報の両方の特性と相互作用を適切に再現
- 提案モデルは、フェイクニュースに強く影響されたユーザークラスターの生成を説明することが可能
- ユーザークラスターはチューリングパターンとして現れるので、チューリングパターンが発生しない条件が実現すれば、ユーザークラスターは発生せず
⇒ フェイクニュース対策につながる可能性
- その例として、訂正情報の拡散の速度を上げることが対応

今後の課題

- よりユーザのクラスタが発生しやすい、偏りのあるネットワークモデルで、ユーザクラスタの発生条件を検討
- より一般的な条件でユーザクラスタの発生を抑える条件の検討
- 上記条件を実現するための、具体的なネットワークオペレーションの検討とその効果の検証