

# 認知バイアスを考慮した意見形成メカニズムのモデル化

○小竹祥生 高橋真吾 (早稲田大学)

## Modeling of Opinion Formation Mechanism Considering Cognitive Bias

\* Y. Kotake and S. Takahashi (University of Waseda)

**概要**— SNS上の虚偽情報を正しいと思い込んだまま何かしらの意思決定をすることは自分や社会にとって望ましくない結果を導いてしまい、現実世界に悪影響を及ぼす。虚偽情報が拡散される理由は内容の信憑性よりもどれだけ共感できるかを人は重視しているためである。また共感を重視することで、確認バイアスや単純接触効果といった認知バイアスに陥りやすくなる。本研究では上記のことを踏まえ、SNS上での意見交換をオピニオンダイナミクスのモデルとして解釈し、そのうえで各エージェントの意思決定に認知バイアスを組み込み、モデルを構築する。エージェントの行動をシナリオ分析することで、エージェントタイプごとの意見形成やネットワーク形成の違いを比較分析する。

**キーワード:** オピニオンダイナミクス, 認知バイアス, SNS

## 1 研究背景と目的

### 1.1 研究背景

SNS等で見ず知らずの他人が発した情報の真偽を見分けることは非常に困難なことである。そのため、虚偽情報とは知らずにその情報が正しいと思い込み、知らず知らずに拡散してしまうというケースが見受けられる。また、虚偽情報を正しいとしたうえで意思決定を下してしまうため、本来自分や社会にとって望ましい選択とは異なった選択をしてしまい、望ましくない結論や結果を導いてしまうことから現実世界にも悪影響を及ぼしている。

近年、フェイクニュースといったワードがメディアで取り上げられているのにもかかわらず、なぜ人々はいまだに虚偽情報を拡散してしまうのかは疑問である。これは総務省のデータ<sup>1)</sup>からもわかるように、情報の信憑性よりも内容にどれだけ共感できるかを重視して情報拡散をしているユーザーの方が多いためである。したがって、人々は情報の信憑性の確認を軽視しがちであることから、虚偽情報の拡散に一役買ってしまったと言える。また、共感に基づいた意思決定では知的活動にしかるべき注意や努力を向けることはないため、認知バイアスによる影響を受けやすくなる。

### 1.2 従来研究

#### 1.2.1 確認バイアス

人はある言明を理解しようとする際など、信じたがるバイアスを備えている。そのため自分の信念を肯定してくれる証拠を意図的に探す傾向にある。このような認知傾向を確認バイアス<sup>2)</sup>という。人は認知的なショートカットを利用することで、適応的な意思決定ができるようになるというメリットがある。しかしながら、自分の意見や価値観にあった情報ばかりにアクセスすることになり、結果として自分の視野を狭めてしまうということが問題となる。

#### 1.2.2 エコーチェンバー

エコーチェンバーとはソーシャルメディア上で、自分と似た興味関心を持つユーザーを大量にフォローすることで、結果的に同じようなニュースや情報ばかりが流通する閉じた情報環境に陥り、自分と似たような意見ばかりが返ってくるようになる状況のことである<sup>3)</sup>。特に先述した確認バイアスに陥っているユーザー

が多いと、ネットワーク環境がエコーチェンバーになりやすくなる。何度も同じ情報に接触し、価値観の近い人とのみの交流が繰り返されることによって、認知バイアスが強化されるだけでなく、閉じた情報環境に陥っていることから訂正情報などが外部から入ってきにくくなることで、虚偽情報が訂正されることなくネットワーク内で拡散される可能性が高くなってしまふ。ネットワークが分断されることにより、意見の対立や社会の分断をも強めてしまふ。

#### 1.2.3 単純接触効果

単純接触効果とはある対象に反復して接触することで、その対象への好意度が高まる現象のこと<sup>4)</sup>である。単純接触効果に影響を及ぼす要因としては(a)刺激変数(b)呈示変数(c)測定変数(d)参加者変数の4つがあげられる<sup>5)</sup>。刺激変数とはSNSを例に考えると、SNSにあげている情報が文字か画像か、動画か、またそれらの情報は複雑かといった情報の質のことを表しており、その情報が新奇で複雑であればあるほど効果的であるとされている。呈示変数とは情報に触れる回数、またその順序、時間などを表現している。一般的に表示回数が多く、表示内容をよく混ぜて呈示するほど効果的であるとされている。測定変数は評価尺度や刺激を与えた後に評価するまでの間隔を表している。最後に参加者変数とは個々のパーソナリティなどのことである。単純接触効果自体は性別や動物の種類にかかわらず影響力があるとされている。

本研究では呈示変数を対象としてモデル化していく。最適な呈示割合は厳密にはわかっていないが、全呈示数の50~100%あると効果的であると言われている。単純接触効果では「変化」があることによって強く影響力が出ることから50~100%の間に単純接触効果を最大にする割合がある(情報内容の対比ができるため)。また、変化の新奇性も影響力を持つことから、一人が何度も同じ情報を呈示するよりも、多くの人が一度ずつ同じ情報を呈示した方が効果的である<sup>6)</sup>。

#### 1.2.4 オピニオンダイナミクス

オピニオンダイナミクスとは意見形成とそれに至るまでの過程を研究する分野のことである。行動モデルのパラメータの違いやモデル自体の違いによって、意見の収束値や収束数がどのように変化するかを調査することが主な目的である。意見を連続的と捉える場合であれば、意見を数直線上の値として表現する。連続

的に意見を捉えるモデルも大きく分けると二つの意見形成の方法が存在し、Deffuant-Weisbuch<sup>7)</sup>のモデルと Hegselmann-Krause<sup>8)</sup>のモデルである。大きな集団における個別の相互作用は前者のモデル、大きな会議などグループでの相互作用は後者のモデルが表現として優れていると言える<sup>9)</sup>。

この意見形成のモデルに SNS の特徴を入れたモデルが存在する。笹原<sup>10)</sup>らはオピニオンダイナミクスのモデルをベースにニュースフィード(タイムライン)という概念を導入した。それに加え、エージェント同士のリンクの繋ぎ変えを考慮したことで SNSらしさを表現した。エージェントに蓄積された他人の投稿(意見)と相互作用することで意見更新をする。時刻  $t$  においてランダムに選ばれたエージェント  $i$  はニュースフィード上において自分の意見と近い意見(自分の意見とある投稿の意見値の距離が閾値以下)と相互作用する。意見更新をしたエージェント  $i$  は確率  $p$  でスクリーン上から意見を選択し、情報拡散する。一方、確率  $1-p$  で自分の意見を隣接するエージェントに対して発信する。

$$I_e(o, m) = \begin{cases} 1 & \text{if } |o - m| < \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

$$o_i(t+1) = o_i(t) + \mu \frac{\sum_{j=1}^l I_e(o_i(t), m_j)(m_j - o_i(t))}{\sum_{j=1}^l I_e(o_i(t), m_j)}$$

$o$ : 意見  $m$ : スクリーン上の投稿  $\varepsilon$ : 閾値  $\mu$ : 社会的影響力

オピニオンダイナミクスに追加する形でリンクの繋ぎ変えも考慮されている。したがって、従来型のオピニオンダイナミクスモデルで利用されている環境モデルのような静的なネットワーク上での相互作用ではなく、動的なネットワーク上での相互作用について分析することになる。具体的にはエージェント  $i$  は確率  $q$  で自分の意見と一致しないエージェントをスクリーン(ニュースフィード)上から選択し、リンクを切断する。そして、全エージェントからランダム、スクリーン上の投稿者、自分と似たような意見を投稿している投稿者を状況に応じて選択し、リンク接続をする。その結果、どのようなネットワーク構造になるかを分析している。結果として、 $\varepsilon=0.4$  前後で意見の二極化が観測され、リンクの繋ぎ変えをすることで意見とネットワークの分極を表現している。このような状態は先述したエコーチェンバーの状況を表現している。

### 1.3 研究目的

本研究の目的はフェイクニュースが広がる原因となっている認知バイアスをオピニオンダイナミクスのモデルに取り入れることである。それに加え、SNS におけるエージェントの行動傾向をシナリオ分析することである。エージェントが意見の投稿/拡散、リンクの接続/切断、などの基準が変化した場合に、意見構造とネットワーク構造にどのような影響を与えるかを分析する。また、エージェントタイプごとの差についても言及する

## 2 モデル

### 2.1 タイムラインモデル

エージェントはスクリーン上の意見と相互作用することで自分の意見を更新していく。表示される投稿は先入れ先出しであり、古い投稿から削除される。行動フローを見てもわかるように 1 ステップ 1 投稿であるため、複数の投稿が同時にスクリーンを埋めることはない。スクリーンの長さの初期値は 10 とする。これは最大 10 投稿と相互作用したうえで意見を更新するということを表している。

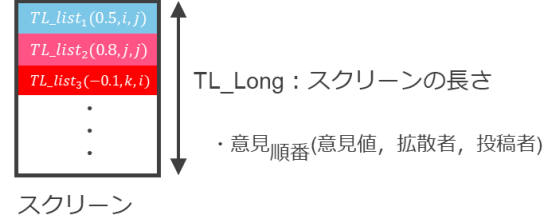


Fig. 1: スクリーン構造

### 2.2 意見投稿/拡散モデル

時刻  $t$  においてランダムに選ばれたエージェント  $i$  は確率  $p$  で意見を投稿(自分の意見を隣接するエージェントに伝達)、 $1-p$  で拡散(スクリーン上のある意見を隣接エージェントに伝達)。拡散する意見はスクリーン上にある投稿を等確率で選択する。

### 2.3 意見更新モデル

意見更新モデルは先行研究をベースとし、状況に応じて、エージェント間の信頼度を考慮できるように拡張しておく。エージェント  $i$  の隣接エージェント  $j$  は意見を受け取った後、その意見を参考に自分の意見値を更新する。オピニオンダイナミクスの考え方をを用いていることに加え、人は虚偽情報を虚偽情報と認識したうえで意見更新や拡散をしていないことから、本研究では情報の質は考慮しない。次節で 3 種類の閾値更新モデルについて説明するが、これらすべてのエージェントはこの節のモデルに従って、意見更新をする。

$$-1 \leq o_i \leq 1, \quad 0 \leq \varepsilon_i \leq 1$$

$$I_e(o, m, \varepsilon) = \begin{cases} 1 & \text{if } |o - m| < \varepsilon \\ 0 & \text{otherwise} \end{cases}$$

$$o_i(t+1) = o_i(t) + \mu \frac{\sum_{j=1}^l (\alpha_j + \beta_j) I_e(o_i(t), m_j, \varepsilon_i)(m_j - o_i(t))}{2 \sum_{j=1}^l I_e(o_i(t), m_j, \varepsilon_i)}$$

$o_i$ : エージェント  $i$  の意見値  $m_j$ :

スクリーン上の  $j$  番目の投稿

$\varepsilon_i$ : エージェント  $i$  の閾値  $\mu$ : 社会的影響力  $t$ : 時刻

$\alpha_j$ : スクリーン上の  $j$  番目の意見の拡散者に対する信頼度

$\beta_j$ : スクリーン上の  $j$  番目の意見の投稿者に対する信頼度

### 2.4 閾値更新モデル

#### 2.4.1 典型的エージェント

それぞれ閾値  $\varepsilon$  を持つが時間経過によってその値は変化しない。初期値がランダムであれば、それだけ多様な閾値を持った集団が表現できることになる

#### 2.4.2 確証バイアスエージェント

確証バイアスに陥ることで情報取得の視野が狭まるという側面を閾値が動的に変化するモデルを構築することで表現する。確証バイアスを表現するにあたり、タイムライン上に自分の意見に肯定的( $|o - m| < \varepsilon$ )な

意見が多ければ多いほど、また、肯定的な意見が自分の意見と近くなれば近くなるほど、自分は正しいと認識し、バイアスが強化されやすくなるとする。すなわち、閾値の幅が小さくなりやすいということである。そこでエージェント*i*の隣接エージェント*k*は意見更新が終わった後に以下の式に従い閾値を更新する。Zは閾値内の投稿の意見値の平均と自分の意見の差を線形に定義した関数である。

if  $positivepost\_num_k(t) > TL\_Long/2$ :

$$\varepsilon_k(t+1) = \varepsilon_k(t) - \delta \frac{positivepost\_num_k(t)}{TL\_Long} Z$$

otherwise:  $\varepsilon_k(t+1) = \varepsilon_k(t)$

if  $\varepsilon_k(t+1) < 0$ :  $\varepsilon_k(t+1) = 0$

TL\_Long: スクリーンの長さ  $\delta$ : バイアスパラメータ

$positivepost\_num_k$ :

*k* がスクリーン上で肯定的な意見と感じた投稿数

$$Z = - \left| o_k(t) - \frac{\sum_{h \in A} m_h(t)}{positivepost\_num_k(t)} \right| / 2 + 1$$

### 2.4.3 単純接触効果エージェント

本研究では呈示者の多様性と呈示割合をモデル化する。また、自分と似ていない意見がスクリーン上に表示されたことに対する反応は呈示者の多様度が高く、呈示割合が0.5~1の間に収まると強く出るとする。呈示割合が最大となる関数はFig. 2のように定義する。

if  $positivepost\_num_k(t) \leq TL\_Long/2$ :

$$\varepsilon_k(t+1) = \varepsilon_k(t) + \delta \frac{posters\_num_k(t)}{TL\_Long} \times f\left(\frac{1 - positivepost\_num_k(t)}{TL\_Long}\right)$$

Otherwise:  $\varepsilon_k(t+1) = \varepsilon_k(t)$

if  $\varepsilon_i(t+1) > 1$ :  $\varepsilon_i(t+1) = 1$

$posters\_num_k$ : スクリーン上投稿者数

$\theta$ : 単純接触効果

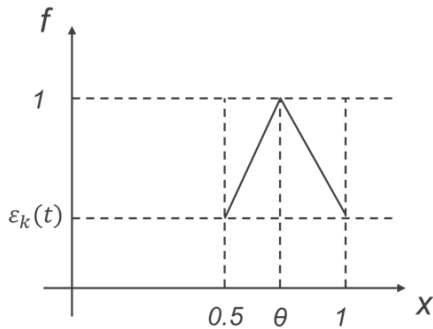


Fig. 2: 呈示割合による影響度関数

## 3 静的ネットワークにおけるエージェントタイプごとの比較

### 3.1 シミュレーションフロー

意見を発信/拡散するエージェント、リンクを接続するエージェント、リンクを切断するエージェントの3種類である。従来研究とは異なり、リンクの接続するエージェントと切断するエージェントが異なることから、各エージェントの次数は保存されない。

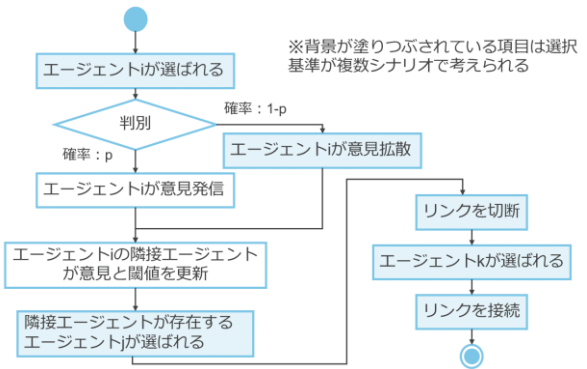


Fig. 3: シミュレーションフロー

### 3.2 初期値

オピニオンダイナミクスに限らず、様々なシミュレーションは初期値の違いによる影響を受ける。ただ、本研究では初期値の影響ではなく、シミュレーション中に生じる、誰が何を意見発信するかなどのランダム性のみを考慮したい。そこで、まずシミュレーションの前にネットワーク (Fig. 5)、閾値、意見 (Fig. 4)、スクリーン上の意見をランダムに初期生成し、その値を今後のシミュレーションでは初期値として用いる。ネットワーク構造はSNSを模していることから有向のランダムグラフを利用する。

Table 1: 変数表 (黄: 試行内で変化)

変数名	値	変数名	値
ノード数	100	グラフ	(有向)ランダムグラフ
エッジ数	500	(初期)スクリーンの中身	ランダム
バイアスパラメータ	0.01	(初期)エージェント意見	ランダム
社会的影響力	0.25	(初期)エージェント閾値	ランダム
タイムラインの長さ	10	意見発信者	ランダム
単純接触効果	0.75	意見発信内容	ランダム
エージェント間信頼度	1	リンク接続者	ランダム
投稿/拡散割合	0.5	リンク切断者	ランダム

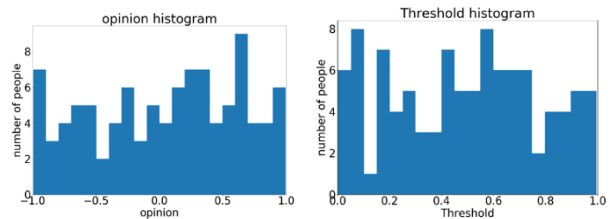


Fig. 4: 初期意見分布(左)初期閾値分布(右)

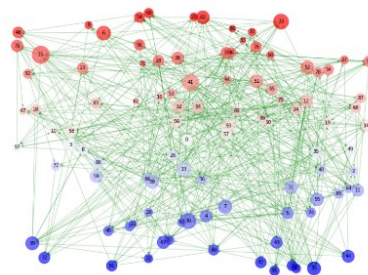


Fig. 5: 初期ネットワーク

### 3.3 エージェントタイプ比較

3章ではネットワークの繋ぎ替えは行わない。したがって、従来のオピニオンダイナミクスのモデルと同様の環境で分析を行う。3タイプのエージェントをネットワーク上に配置してそれぞれの意見形成の違いを分析する。

典型的エージェントは基本的に多くのエージェントが一つの意見にまとまろうとする傾向がみられる。Fig. 7からもわかるように意見がまとまろうとしないエージェントの特徴は閾値が小さいことである (Table 2)。また、その閾値が小さいエージェント同士のネットワークが疎であることから、シミュレーションステップを重ねても、小クラスタの意見が一つにまとまる可能性は低いと解釈した。

最終意見のヒストグラムを確認すると0付近に収束しているとは言えない (Fig. 6)。これはランダム生成した初期意見において0以上の意見のほうが相対的に多かったためであると考えられる。一度意見がまとまっても収束しないさまは他人の意見を参考にしよく吟味しながら自分の意見を決定していると意味的に解釈できる。

確認バイアスエージェントの方が意見の分極が起きやすい。場合によっては初期値の偏りの影響を受けずに負の意見値の方が大きいクラスタを形成する場合もあった。確認バイアスエージェントの意見形成では一度束になってからばらつかないことから自分の信念に対する確信度合いが強くなることで迷いがなくなり、速い段階で自分の意見を固めたと解釈することができる。

単純接触効果エージェントはほぼ一つの値に収束する。最終意見分布において-1付近に少し意見が残るがこれは大きなクラスタから1以上離れてしまっているため相互作用できないためである。意見が極端なエージェントはあまりにも違う意見を許容できないことを表現している。意見収束も典型的エージェントに比べると速いことが特徴的である。

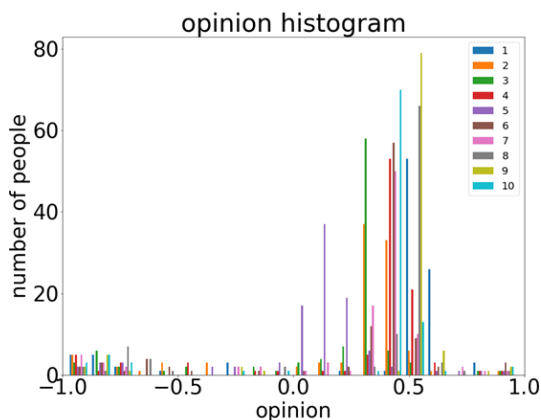


Fig. 6: 典型的エージェントの最終意見分布

Table 2: 代表的な外れ値の閾値 (試行 5)

エージェント番号	閾値	エージェント番号	閾値	エージェント番号	閾値
15	0.080258	17	0.211945	44	0.215497
48	0.116386	21	0.05154	67	0.002111
62	0.090436	40	0.293297	72	0.163261
99	0.058467	42	0.012309	81	0.045304

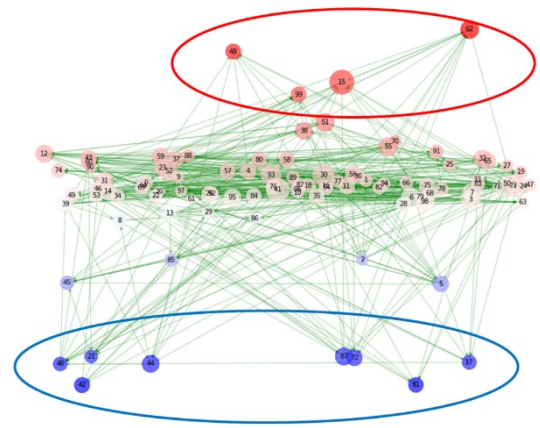


Fig. 7: ネットワーク (試行 5)

### 3.4 パラメータ感度分析

バイアスパラメータ  $\delta$  を変化させることでどのような変化が生じるかを検証した。パラメータ値が小さいほど、単位時間当たりの閾値の変化量が小さくなることから、確認バイアスエージェント、単純接触効果エージェントともに、典型的エージェントの閾値をランダム配置した場合と似たような意見変化をした。また逆に  $\delta$  を大きくすると確認バイアスエージェントのみの方は意見が様々な収束値にばらつき、単純接触効果エージェントでは意見がきれいに一つに収束した。

単純接触効果エージェントの最適割合  $\theta$  を 0.5 から 1 の間で動かした場合、0.5 の場合、最終意見値は 0.25~0.5、1.0 の場合は 0~0.25 の間に収まる場合がほとんどであることから、分布が動く傾向が見られた。

## 4 動的ネットワークにおけるエージェントタイプごとの比較とシナリオ分析

### 4.1 リンク更新基準

どのような基準でリンクの更新を決めるかはひとそれぞれである。そこでリンク接続基準はランダム、フォロワーが多いエージェント、意見が近いエージェント、友人の友人エージェントの4通り、切断基準はランダム、投稿頻度が多いエージェント、意見が離れているエージェントの3通りを組み合わせることでどのような意見やネットワークの変化がみられるかを分析する。

#### 4.1.1 リンク切断基準：ランダム

リンクの切断においては隣接エージェントがいるエージェントの中からランダムで選んだが、リンク接続に関しては全エージェントから完全ランダムで選択している。したがって、友人の友人をフォローするシナリオであっても、場合によっては友人の友人をすべてフォローしている可能性はある。このような場合はランダムにフォローする相手を決定している。そのため、友人の友人シナリオではあまりランダムシナリオと差が出なかったと考えられる (Fig. 8&11)。

入次数が多いエージェントを優先的にフォローする場合、最終意見分布にばらつきがみられる (Fig. 9)。これは、ある特定のエージェントの入次数が大きくなり、影響力が増すためである。また、意見間距離の近さを基準とした場合 (Fig. 10)、意見の収束は速くなる

(Fig. 12). これはスクリーンが効率的に許容範囲内の意見で埋め尽くされるからである. 実際, 自分のスクリーン上には許容範囲内の意見が10また0といった結果になった.

確証バイアスエージェントにおいて, 接続基準がランダム場合, 静的ネットワークの分析で見られたような明らかな意見の分極は見られず (Fig. 13), 少数クラスタは意見としてなんとなく近い状態にとどまっている. これはランダムなリンクの切り替えが効率的な意見交換を阻害したことに加え, 閾値の収束の速さによる影響が強く出たためであると考えられる. 意見の近さシナリオの場合は複数の意見クラスタになっている場合も観測された (Fig. 14). 典型的エージェントよりも1クラスタに多くの要素が含まれている.

単純接触効果エージェントの意見は分かれることなくある値に収束する. ただ, 各シナリオで比較すると他のエージェントタイプと同様のばらつき方を見せていた.

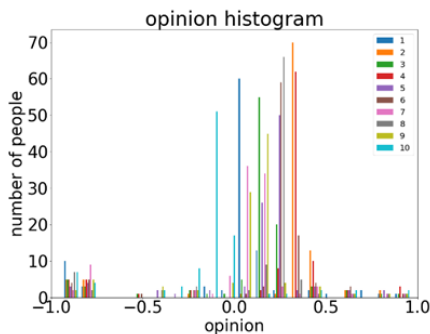


Fig. 8: 典型的エージェントの最終意見分布 (ランダム)

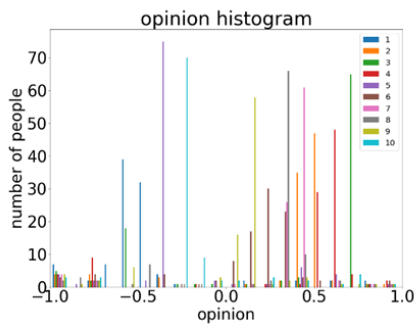


Fig. 9: 典型的エージェントの最終意見分布 (次数優先)

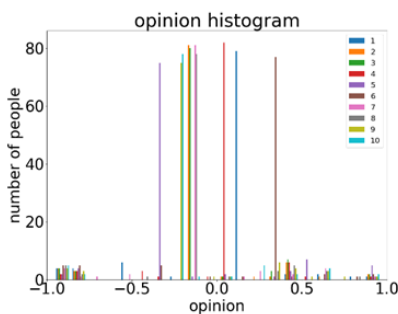


Fig. 10: 典型的エージェントの最終意見分布 (意見の近さ優先)

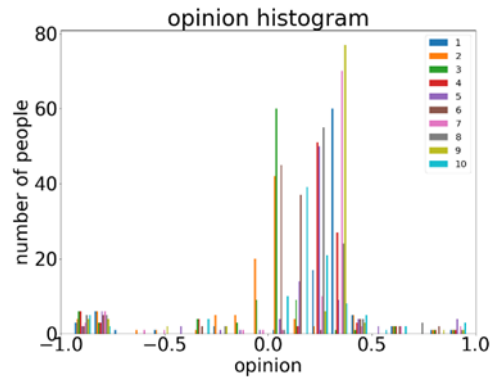


Fig. 11: 典型的エージェントの最終意見分布 (友人の友人優先)

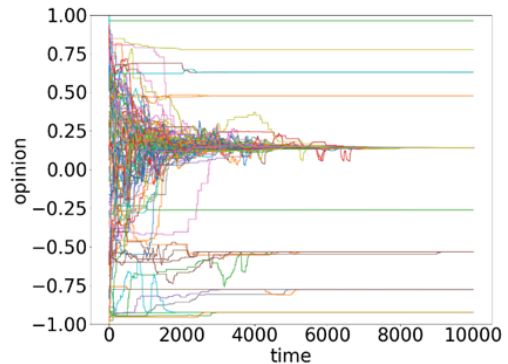


Fig. 12: 確証バイアスエージェントの意見変化 (意見の近さ優先: 試行1)

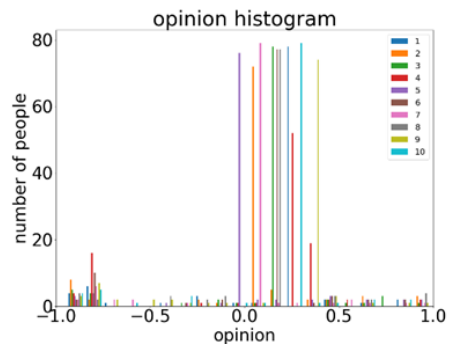


Fig. 13: 確証バイアスエージェントの最終意見分布 (ランダム)

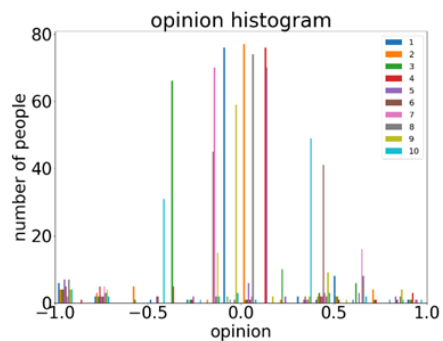


Fig. 14: 確証バイアスエージェントの最終意見分布 (意見の近さ優先)

#### 4.1.2 リンク切断基準：投稿頻度

リンク切断基準がランダムな場合と最終意見分布に大きな差は見受けられなかった。これは投稿頻度の差を考慮するとはいえ、誰が意見発信するか自体がランダムであるため、実質ランダムとあまり差が生じなかったと考えられる。入次数の大きいエージェントを積極的にフォローしていく場合はそもそも試行ごとにばらつきが大きいのであまり差がないと判断した。

#### 4.1.3 リンク切断基準：意見間距離の遠さ

最終意見分布のみを比較するとあまり他と差がない様に見えるが意見変化の推移とネットワーク概形には差が出ている。リンク切断基準がランダムな場合よりも収束に至るまでの意見のばらつきや蛇行がみられなくなった。意見の遠いエージェントを優先的にリンク削除することで、結果として意見の収束が速くなることがわかる。意見の近いエージェントを優先的にフォローする場合において基本的に一つの意見に収束するが1試行のみ意見とネットワークの分断が見られた (Fig. 15)。

確認バイアスエージェントの場合、意見間距離が近いエージェントを優先的にフォローした場合の分布は顕著に異なる (Fig. 16)。相対的に意見が分断されるだけでなくネットワークも分断されやすくなっていることがわかる (Fig. 17)。本研究における初期値においては閾値が変化しない典型的エージェントよりも確認バイアスエージェントの方が本質的にエコーチェンバー現象を表現できていると言うことができる。

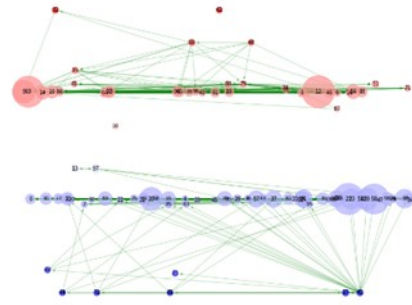


Fig. 17: 確認バイアスエージェントのネットワーク (意見の近さ優先：試行1)

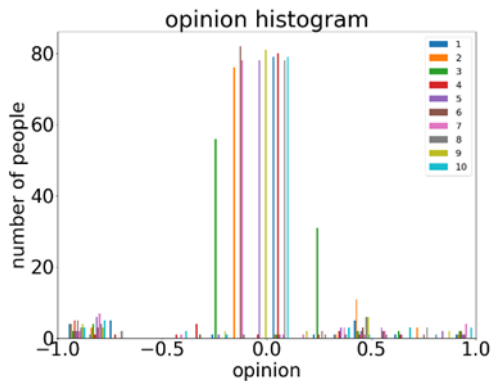


Fig. 15: 典型的エージェントの最終意見分布 (意見の近さ優先)

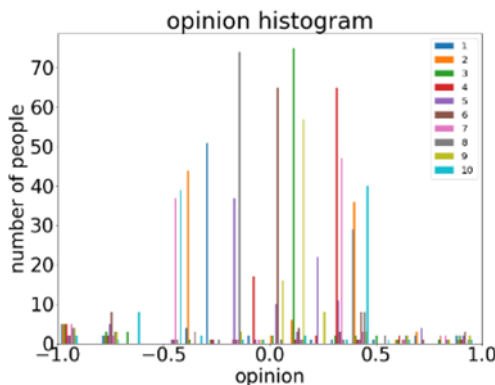


Fig. 16: 確認バイアスエージェントの最終意見分布 (意見の近さ優先)

## 4.2 リンク更新者

どのような特徴を持ったエージェントがリンクの更新をしやすいかの優先度をつけることで、意見形成やネットワーク形成に変化が生まれるかをシナリオ分析する。シナリオはランダム、閾値が小さいエージェントが優先的、意見の絶対値が大きいエージェントが優先的 (意見が極端)、出次数が大きいエージェントが優先的 (フォローしている人が多い) の4通りを分析する。ここで言う優先的とはある特徴を持ったエージェントが確率的に選ばれやすい (積極的に行動しやすい) ということである。

明らかに最終意見分布が異なるのは出次数が大きいエージェントが積極的に行動した場合である。これはフォローしてる人が多い人ほど積極的にリンクの繋ぎ替えをするというシナリオである。結果としてフォローされているけどフォローしていないエージェントが多く生まれる。これにより、多くのエージェントは意見更新ができず、またフォローしている人が多いエージェントの意見はなかなか収束せず、振動するという現象が生じる (Fig. 18)。また、基本的に他のシナリオでは確認バイアスエージェントの閾値はほとんどが0に収束するがこのシナリオの場合はそうではない (Fig. 19)。意見の更新ができないことから、閾値の更新もできないためである。

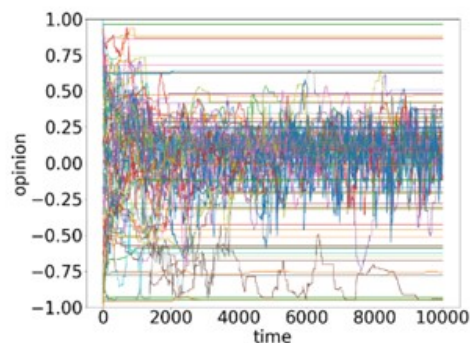


Fig. 18: 典型的エージェントの意見変化 (出次数優先：試行10)

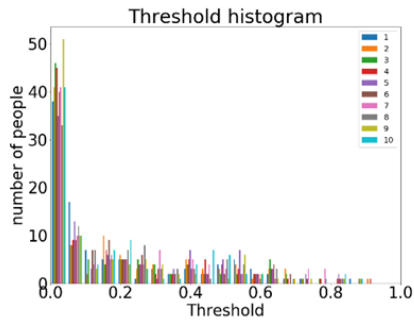


Fig. 19: 確証バイアスエージェントの最終閾値分布  
(出次数優先)

#### 4.3 意見投稿／拡散割合

自分の意見の発信と意見の拡散どちらを優先的に行うかを分析する。典型的エージェントの場合、あまり最終意見分布に差は見られないが、時間経過が浅い段階においての意見のばらつきが多い試行が拡散優先の場合の方が多くみられた (Fig. 20&21)。これはスクリーン上の意見の初期値がランダムであるため、自分の意見と遠い意見が多いと周囲の意見とあまり似ていない意見を発信し、収束を遅くするためであると考えられる。

確証バイアスエージェントは拡散行動を優先させることで、場合によっては初期値の意見の偏りとは逆に大きな意見のクラスタが形成するパターンが見られた (Fig. 22)。したがって、それだけ、序盤の意見形成が最終意見分布に影響を与えるということである。

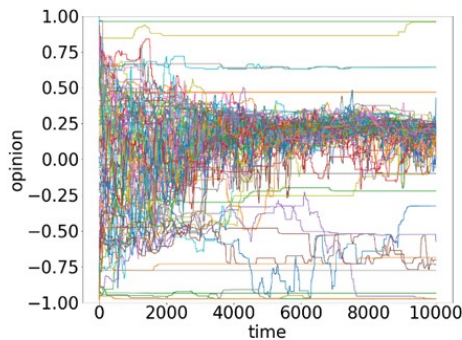


Fig. 20: 典型的エージェントの意見変化  
(拡散行動優先：試行 2)

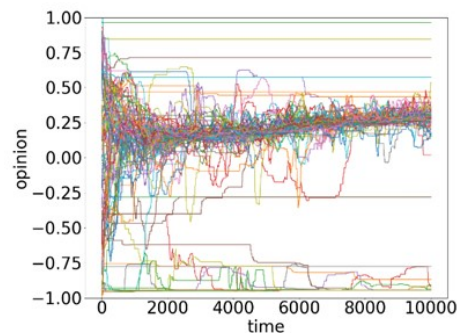


Fig. 21: 典型的エージェントの意見変化  
(投稿行動優先：試行 4)

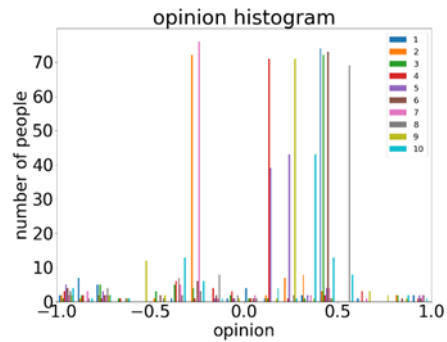


Fig. 22: 確証バイアスエージェントの最終意見分布  
(拡散行動優先)

#### 4.4 意見投稿／拡散者

どのような特徴を持ったエージェントが意見を発信しやすいかの優先度をつけることで、意見形成やネットワーク形成に変化が生まれるかをシナリオ分析する。シナリオはランダム、閾値が小さいエージェントが優先的、意見の絶対値が大きいエージェントが優先的(意見が極端)、出次数が大きいエージェントが優先的(フォローしている人が多い)の4通りを分析する。

意見の絶対値が大きいエージェントが優先的に発言することでエージェントタイプによらず、意見が二つに分かれやすくなることがわかった (Fig. 23~25)。確証バイアスエージェントが一番その影響を強く受け、単純接触効果はわずかではあるが影響を受けた。ただ、意見がきれいに分かれたとしても、ネットワークの更新基準がランダムであるため、ネットワークの分断は生じなかった。

出次数が大きいエージェントが優先の場合はあまりランダムと大きな差が生じなかった。これは意見発信にかかわるのは入次数であるからである。また、閾値が小さいエージェントが優先的に発言することで最終意見分布はランダムと大きな差はなかったが、閾値が小さいということは意見的に孤立している可能性も高いということであり、ランダムよりも意見が収束するまでのばらつきが大きくなっている試行が多く見られた。

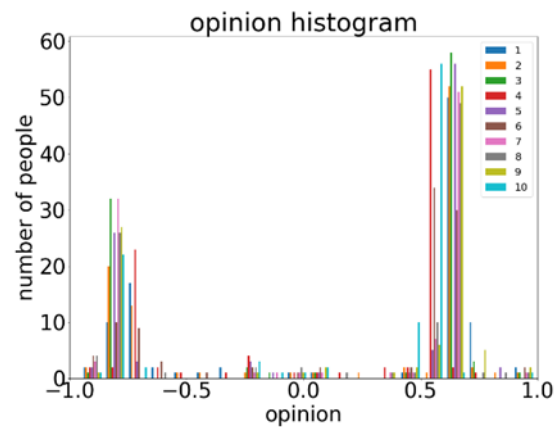


Fig. 23: 典型的エージェントの最終意見分布  
(極端な意見を持つ人優先)

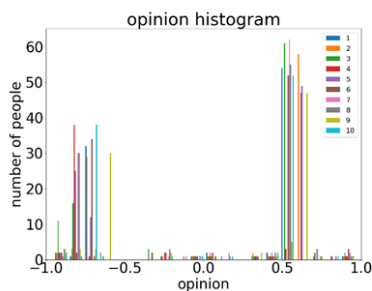


Fig. 24: 確証バイアスエージェントの最終意見分布 (極端な意見を持つ人優先)

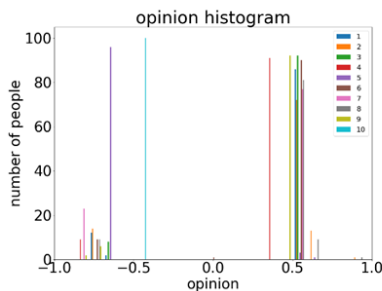


Fig. 25: 単純接触効果エージェントの最終意見分布 (極端な意見を持つ人優先)

#### 4.5 拡散意見内容

人は内容の信憑性よりもどれだけ内容に共感できるかが拡散の基準となっていた。そこで、拡散する意見の内容をランダムに加え、自分の意見と近い意見、友人の信頼度が近い意見の3種類でシナリオ分析をする。エージェント間信頼度を前節まで1としてきたが、この節ではランダムとして再設定する。

典型的エージェントと単純接触効果エージェントではあまりシナリオ間での差が見られなかった。これは信頼度をランダムとして取り入れたが、他の行動のランダム要素と相殺されて、信頼度自体のランダム性が強く反映されなかったためであると考えられる。また意見の近さ優先で意見拡散した場合も結局、自分の意見を発信している状況とほぼ似ているためあまり差が出なかったと考えられる。

ただ、確証バイアスエージェントの場合は若干意見分布に差が見られた(Fig. 26&27)。他の節の分析からもわかるようにランダムでも、意見が分断する可能性はある。しかし、それが相対的に多く結果として出ていることから、本研究の設定においては信頼度の影響を確証バイアスのエージェントは受けたと言える。

#### 5 結論と今後の課題

本研究では確証バイアスと単純接触効果をオピニオンドYNAMICS上で表現できる形にモデル化した。今後、各試行において特徴的な行動をしていたエージェントを追跡することで、本研究では得ることができなかった知見を得ることができている可能性がある。また、モデルをより精緻なものにするためには認知バイアスに人が陥っていく変化速度や変化量をターゲットとした実証的な実験がさらに求められる。

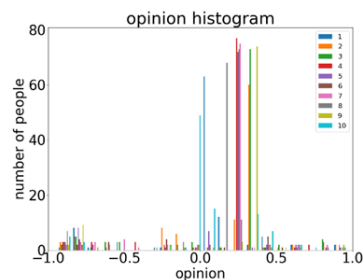


Fig. 26: 確証バイアスエージェントの最終意見分布 (ランダム)

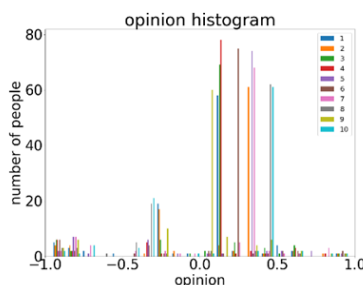


Fig. 27: 確証バイアスエージェントの最終意見分布 (信頼度優先)

#### 参考文献

- 1) <http://www.soumu.go.jp/johotsusintokei/whitepaper/ja/h27/html/nc242250.html>
- 2) ダニエル・カーネマン：ファスト&スローあなたの意思はどのように決まるか？上(邦題)，早川書房 (2012)
- 3) 笹原和俊：フェイクニュースを科学する 拡散するデマ、陰謀論、プロパガンダのしくみ，化学同人 (2018)
- 4) Zajonc, R. B. : Attitudinal effects of mere exposure, *Journal of Personality and Social Psychology*, 9, 1/27 (1968)
- 5) Bornstein, R. F. : Exposure and affect : Overview and meta-analysis of research, 1968-1987, *Psychological Bulletin*, 106, 265/289 (1989)
- 6) 川上直秋：単純接触効果と無意識—われわれの好意はどこから来るのか，エモーションスタディーズ，第1巻第1号，81/86 (2015)
- 7) Guillaume Deffuant, David Neau, Frederic Amblard, and Gerard Weisbuch : Mixing beliefs among interacting agents, *Advances in Complex Systems*, Vol.03No.01n04, 87/98 (2000)
- 8) Rainer Hegselmann, Ulrich Krause : Opinion Dynamics and Bounded Confidence Models, Analysis, and Simulation, *Journal of Artificial Societies and Social Simulation (JASSS)*, vol.5 no. 3 (2002)
- 9) Alina Sirbu, Vittorio Loreto, Vito D.P.Servedio, and Francesca Tria : Opinion Dynamics: Models, Extensions and External Effects, *Participatory Sensing, Opinions and Collective Awareness*, Springer International Publishing, 363/401 (2017)
- 10) Kazutoshi Sasahara, Wen Chen, Hao Peng, Giovanni Luca Ciampaglia, Alessandro Flammini, and Filippo Menzer : On the Inevitability of Online Echo Chambers, arXiv:1905.03919v2 [cs.CY] (2019)