

ソーシャルポルノ仮説の提案とその観測に向けて

Proposal of "Social Porn" Hypothesis and Pre-analysis towards its Observation

榎 剛史 *1,*2

Takeshi Sakaki

鳥海 不二夫 *2

Fujio Toriumi

*1株式会社ホットリンク

Hottolink, Inc.

*2東京大学

The University of Tokyo

In recent years, the negative aspects of information creation by individuals, such as fake news, flaming and echo chamber phenomena, have been paid attention. We propose a hypothesis "social porn" as one of the causes of these negative phenomena. We define "social porn" as "information that a user belonging to a specific community wants to diffuse and share instantly".

In this paper, we define the scale of user reaction time as preparation for observation of social porn. We examine the difference of user reaction time distribution for some tweets. As a result, analysis results suggest that the distribution of user response times may differ between tweets randomly extracted and tweets spreaded by specific community.

1. はじめに

2016年、オックスフォード大学出版局は、2016年の「Word of the year」に「Post-truth」を選定した。彼らによれば、「Post-truth」とは、「世論を形成する際に、客観的な事実よりも、むしろ感情や個人的信条へのアピールの方がより影響力があるような状況」と定義されている。この語は、フェイクニュースや炎上、エコーチェンバー現象が日常的に発生するようになった現在の社会状況を端的に表した語として、2016年～2017年にかけて注目を浴びた*1。

ソーシャルメディアの登場により、中小規模メディアや個人が情報を発信し、それを流通させるのが容易になった。それらは、ソーシャルメディアが普及し始めた当初はその正の側面としてクローズアップされてきた。しかし、現在では、「Post-Truth」の言葉が端的に表すように、その負の側面に注目が集まっている。例えば、2016年の大統領選挙では、フェイクニュースが世論形成に大きな影響を及ぼしたとされている。フェイクニュース業者の多くは、PV数に基づいて支払われる広告プログラムに参加し、そこから利益を得るためにフェイクニュースを発信・拡散させたことを認めている。エコーチェンバー現象は、様々な事例について観測されている。その多くは、より衝撃的な内容を伝達することを競いあった結果、より過激な意見が創出されるようになったと言われている*2。

これらの多くの分析は、主に情報の発信者側に着目して、その要因や発生過程を明らかにしようと試みている。では、受信者に着目すると、どうであろうか？彼らは、いや、我々はどうのような行動原理によりエコー・チェンバーを引き起こし、フェイクニュースを拡散させ、炎上をさらに加速させてしまうのであろうか？「Post-truth」の定義にあるように、多くの人は「感情や個人的信条へのアピール」に影響を受けてしまっているのではあろうか？本論文では、ここで「ソーシャルポルノ(Social Porn)」という概念を提唱する。「ソーシャルポルノ」とは「社会欲求(Social needs)を充足する目的で消費・拡散され

るコンテンツ」と定義する。ここでの「porn」は「Inspiration porn」や「Patriot Porn」と同様の用法であり、また「ソーシャルポルノ」自体はそれらを包含する概念である。

ソーシャルポルノの存在を仮定すると、前述のソーシャルメディアにおける負の側面、ネガティブな現象は容易に説明することができる。フェイクニュースにせよ、炎上にせよ、エコー・チェンバー現象にせよ、ある個人の持つ意見や感情を強化する内容のコンテンツが配信された時に、その個人の「社会欲求」が刺激されてしまうため、真贋や情報の適切性、社会的な反響を考慮せずに脊髄反射的にその情報を拡散(再配信)してしまうと考えられる。本来であれば、情報の真贋や適切性を確認した上で情報を拡散することが望ましいが、そのような「世の中に正しい/適切な情報を伝えよう」という欲求は、マズローの欲求段階で言えば、より高次の「承認(尊重)の欲求」や「自己実現の欲求」に属するものと考えられる。マズローによれば、低次の欲求が充足されていない限り、それよりも上位の欲求を満たすことは難しいとされている。結果として、一定規模以上の個人の「社会欲求」を刺激するようなコンテンツは大規模に拡散され、フェイクニュースや炎上、エコー・チェンバー現象を引き起こす、と考えられる。つまり、これらのネガティブな現象は、あくまで個人が脊髄反射的な情報拡散が大規模に行われた結果、多様な形で顕在化したものであり、それらの現象に共通する根源的な要因としてソーシャルポルノが存在すると思われる。

このように人間の社会欲求を充足することを目的として作られたコンテンツーソーシャルポルノという概念を仮定することで、近年ソーシャルメディア上で引き起こされている社会現象を説明することが可能となる。しかし、もちろんこの概念は仮説に過ぎない。そこで本論文では、実際にソーシャルメディアデータを分析し、「ソーシャルポルノ」が実際に存在することを検証していきたい。

まずは「ソーシャルポルノ」を下記のように、より具体的に定義する。

ソーシャルポルノとは、特定のコミュニティに属するユーザが「脊髄反射的に拡散・共有してしまいたくなる情報」である

このように定義するのは下記のような理由である。ここで

連絡先: 榎 剛史, 株式会社ホットリンク, 東京都千代田区富士見 1-3-11 デュープレックスビズ5F

*1 <https://en.oxforddictionaries.com/word-of-the-year/word-of-the-year-2016>

*2 https://en.wikipedia.org/wiki/Post-truth_politics

のコミュニティとは、共通の趣味や感心・社会的な意見を持つ一定規模の集団である。ソーシャルポルノが問題を引き起こす場合は、一定規模以上のユーザが反応したばあいに限られるので、本稿で扱うソーシャルポルノは、一定規模のコミュニティの構成員に共通して訴求するものとした。脊髄反射的にとしたのは、「内容を精査しない」ことを前提とするためである。冷静に吟味してから、真贋や適切性を確かめてから拡散・共有するような情報は「ポルノ」とはいえない。本稿ではあくまで、「目にした瞬間に拡散したくなるような情報」を「ポルノ」と見なす。また、ソーシャルポルノの内容自体もコミュニティや個人によって異なると考えられる。通常の性的な意味での「ポルノ」と同様に人々の欲求に立脚するものであるため、ソーシャルポルノにも多様な趣味・志向があると考えられる。

本研究において、我々はまずはソーシャルポルノを観測し、定量的に評価することを最終的な目標としている。本稿では、その準備段階として、ランダムに抽出された投稿と特定のコミュニティに属するユーザによる投稿について、ユーザの反応に違いがあるかどうかの検証を行う。ここでは特にソーシャルポルノの「脊髄反射的に拡散・共有し」という定義から、ユーザの投稿への反応時間、すなわち「ある情報を受け取ってから、その投稿を拡散するまでに要した時間」に着目する。つまり、特定のコミュニティに属するユーザによる拡散が、ランダムに抽出された拡散よりも短い反応時間で行われているかどうかを分析する。

本研究では、情報取得に容易性から Twitter のデータを対象とする。

2. 着想

我々は、人工知能学会表紙問題の炎上に端を発し、2014 年から今まで Web 上、特にソーシャルメディアを対象とした炎上現象の分析手法の研究・開発ならびに、開発した手法による事例分析を数多く行ってきた [鳥海 14b, 鳥海 14a, 鳥海 17]。扱ってきた分析事例は多岐にわたっている。例えば、社会現象 (アイスバケツチャレンジ) や社会問題 (マクドナルド異物購入問題, オリンピックエンブレム盗作問題), 科学に関するトピック (STAP 細胞問題, Facebook による不適切な心理学実験, 人工知能学会表紙問題), 自然災害 (御嶽山噴火, 鬼怒川洪水), などである。これらの事例分析においては、トピックごとのエントロピを算出し、「そのトピックが幅広いユーザに語られているのか?あるいは特定の偏ったユーザによって語られているのか?」によって評価を行ってきた。具体的には下記のような手順である [鳥海 17]。

1. ソーシャルメディア上のネットワーク関係を元にユーザ・投稿をそれぞれ Modularity Optimization の手法によりクラスタリング [Blondel 08], ユーザコミュニティとトピッククラスタを構築する
2. 各トピック i についてユーザコミュニティの分布からトピックエントロピ H_i を算出する

$$H_i = - \sum_{c \in C} P(c) \log P(c) \quad (1)$$

(C はコミュニティ集合, c はコミュニティ, $P(c)$ はあるツイートがコミュニティ c のメンバーによっておこなされた確率)

3. 各トピッククラスタのトピックエントロピが高いトピックを「多様なユーザコミュニティによって幅広く語られ

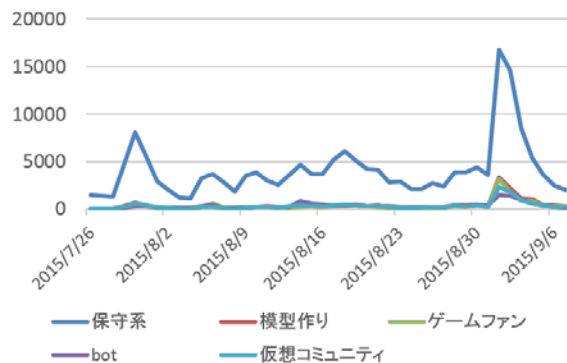


図 1: オリンピックエンブレム問題 コミュニティ別投稿数推移 (x 軸:日付, y 軸:投稿数)

ていたトピック」, 低いトピックを「特定のユーザコミュニティによって偏って語られていたトピック」とみなし, それぞれの内容を俯瞰する

このような事例分析を行う中で、複数の炎上事例に共通して出現するトピックの傾向に着目した。低いエントロピを有するトピックつまり、特定のコミュニティにおいて語られているトピックについては、共通の傾向があることが推測された。例えば、特定の政治思想を持ついくつかのコミュニティは、表 1 のように複数の炎上事例において、その事例に言及しつつ、事例とは直接関係のない政治的な批判意見や陰謀論を拡散しているのが見受けられた^{*3}。特に陰謀論のトピックについては日本のみならず英語圏においても観測されることがわかった。また、トピックに関連する大半の投稿が一つのコミュニティによって投稿される場合もある。図 1 はオリンピックエンブレム問題における投稿数推移をコミュニティ別に表現したものである。図 1 より、特定の政治思想に関連する特徴語を持つグループによる投稿が大半を占めている。

この観測事実から、我々は特定のコミュニティにおいて、好んで拡散されるトピックに類似性があるのではないかと、という仮説を持つに至った。前述の事例では、わかりやすい事例として特定の政治思想に興味を持つコミュニティを取り上げたが、それ以外に多様なコミュニティにおいても、そのように好んで拡散されるトピックがあるのではないかと考えた。それらのように、特定のコミュニティにおいて好んで拡散されるトピックを「ソーシャルポルノ」の一種として捉え、それらを観測することを試みる。

3. 方法論

Social Porn を検出するための方法論について述べる。まずは、あるユーザがある情報を受け取ってから情報を再発信 (拡散) するまでに要した時間: ユーザ反応時間を定義する。

次に、定義したユーザ反応時間に基づいて、各投稿のユーザ反応時間の分布を作成し、その統計量を算出する。

3.1 ユーザ反応時間の定義

ソーシャルポルノとは、ユーザが「脊髄反射的に拡散・共有してしまいたくなる投稿」であると定義した。ここでの「脊髄反射的」とは、内容を精査しないことを意味する。つまり、

^{*3} 倫理的な配慮から、検索できないように文意が変わらないように文言に修正を加えている

表 1: 複数炎上事例に出現する類似トピック

事例	タイプ	Teewts
STAP 細胞	無関連な批判	STAP 細胞の件で日本の学術界の信頼がとか言い出す人は理解できない。 原発事故で失われた信頼がいつ回復したというのか
アイスバケツ	無関連な批判	【韓国経済崩壊】韓国企業の売名行為に世界中で非難殺到!慈善活動「アイスバケツ #韓国 #アイスバケツチャレンジ #韓国経済崩壊 #嫌韓
STAP 細胞	陰謀論	「STAP 細胞問題」の裏で集団的自衛権の行使容認に向けた準備が進んでいる。 安全保障調査会の会長によれば、「安保法制懸」は近々結論を出すらしい。
Facebook 実験	陰謀論	As I predicted, Facebook's psychology experiment was directly linked to the Pentagon's plan for civil unrest.

ソーシャルポルノを抽出するためには「内容を精査していない」を可能な限り確実に観測できることが望ましい。

そこで、我々は、対象とする情報をリツイートの中に絞る、ユーザが拡散に要した時間が極めて短い情報に着目することとした。なぜなら、ユーザが情報を受け取ってからそれを拡散するまでに要した時間が数十秒程度であれば、ユーザがその情報を精査していないことは明らかである。本稿では、この、ユーザ u_j がある情報 i_k を受け取ってからそれを拡散するまでに要した時間をユーザ反応時間 $T_R(u_j, i_k)$ と定義する。

ユーザ反応時間 $T_R(u)$ を測定するためには、ユーザがその情報を受け取った時刻（以下、情報受信時刻）を計測する必要がある。ユーザが Twitter 内で情報を受け取るリツイートであれば、情報受信時刻をある程度推測することができる。一方、リツイート以外では、ユーザが他のメディアから情報を受け取っている可能性も少なくなく、そのような場合には情報受信時刻を推測することは困難である。よって、本稿では、リツイートのみを分析対象とする。ここで、ユーザ u_j が情報 i_k を受信した時刻（情報受信時刻）を $t_r(u_j, i_k)$ 、ユーザ u_j が情報 i_k を送信（拡散）した時刻（情報送信時刻）を $t_s(u_j, i_k)$ と表す。

$t_r(u_j, i_k)$: 情報受信時刻

$t_s(u_j, i_k)$: 情報送信時刻

$$T_R(u_j, i_k) = t_r(u_j, i_k) - t_s(u_j, i_k)$$

あるツイート i_k について $T_R(u_j, i_k)$ を算出することを考える。 $t_s(u_j, i_k)$ は、ツイート i_k がユーザ u_j によって、リツイートされた時刻であるため、データから明示的に取得可能である。一方、 $t_r(u_j, i_k)$ は、データから明示的に取得するのは困難である。なぜなら、1. Twitter の仕様上あるツイートがユーザの TL(タイムライン) 上にいつ出現したかを知ることが難しく、2. またいつユーザが TL をチェックしたかを知ることが困難である。ためである。そこで $t_r(u_j, i_k)$ を推測する必要がある。

ここでは 2 つの推測方法を提案する。1 つ目は、ユーザ u_j がフォローしているユーザのうち、一番最初に情報 i_k をツイートしたユーザの i_k の送信時刻を $t_r(u_j, i_k)$ として用いる手法である。これをアプローチ 1 と呼ぶ。

2 つ目は、ユーザ u_j が情報 i_k を送信する前に送信したツイートのうち、最も新しいツイートの送信時刻を $t_r(u_j, i_k)$ とする手法である。こちらをアプローチ 2 と呼ぶ。

今回は、データ取得の制約上、アプローチ 2 を用いることとする。

3.2 ツイートごとの反応時間分布

ソーシャルポルノの定義は、「脊髄反射的に拡散・共有してしまう投稿のことである。もしもこの仮説が正しかった場合、

表 2: D_{random} のユーザ反応時間分布 統計量

ID	RT 数	最頻値	平均値	中央値	標準偏差
1-1	151	1.0E+05	6.8E+05	1.5E+05	1.1E+06
1-2	1204	10	6.5E+04	1.6E+03	3.3E+05
1-3	551	100	2.3E+05	3.2E+03	1.3E+06
1-4	55	100	4.7E+04	5.7E+03	1.1E+05
1-5	452	10	4.6E+04	4.3E+02	3.9E+05
1-6	558	10	2.8E+04	9.8E+02	1.3E+05
1-7	6681	10	6.9E+04	2.0E+03	3.3E+05
1-8	179	1	2.0E+05	6.0E+02	7.5E+05

ソーシャルポルノにあたる投稿と、それ以外の投稿を比較した場合に、それらを拡散したユーザ群の反応時間の分布が異なることが想定される。より具体的には、ソーシャルポルノにあたる投稿の方が、ユーザ群の反応時間が短いことが想定される。そこで本稿の分析においては、分析対象としたツイート群について、それらのツイートを RT したユーザとその反応時間の分布を作成し、その統計的な指標（平均値、中央値、最頻値、標準偏差）を比較することとした。

4. 分析

4.1 データセット

今回は、比較対象として 2 つのツイート群を用意した。1 つのツイート群は、一般的なツイートに対するユーザの反応時間分布の傾向を見るためのデータセットである。今回は、著者のタイムラインから 2018 年 3 月 6 日・7 日に投稿されたツイートで、リツイート数が 100 を超えるツイートをランダムに 8 件抽出した。こちらを D_{random} とする。

もう一方のツイート群は、ソーシャルポルノらしい投稿に対するユーザの反応時間分布の傾向を見るためのデータセットである。今回は、特定の政治思想に関連した 2 ちゃんねるまとめサイトのアカウントから、2018 年 3 月 8 日 0 時において、直前に投稿された 8 件を抽出した^{*4}。こちらを D_{matome} とする。

今回は、本格的な分析を行う前の試行的な分析であるため、小規模なデータセットを用いることとする。また、リツイート数やいいね数は 2018 年 3 月 8 日 15 時時点のものとする。

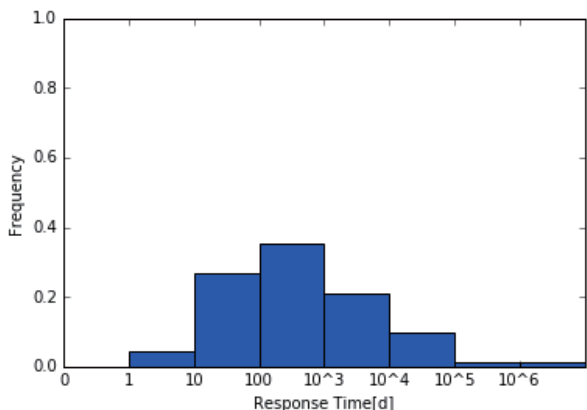
4.2 分析結果

2 つのデータセットにおいて、前述の通りアプローチ 2 を用いてツイート・ユーザごとのユーザ反応時間を算出する。得られたユーザ反応時間に基づいて、ツイートごとのユーザ反応時

*4 倫理的な配慮から、当該アカウントについては付せるものとする

表 3: D_{matome} のユーザ反応時間分布 統計量

ID	RT 数	最頻値	平均値	中央値	標準偏差
2-1	239	10	3.5E+04	3.5E+02	1.4E+05
2-2	352	10	5.2E+04	4.5E+02	3.4E+05
2-3	408	10	8.5E+04	8.0E+02	3.2E+05
2-4	117	10	3.4E+04	2.0E+02	1.2E+05
2-5	429	10	3.0E+04	4.5E+02	1.2E+05
2-6	248	10	3.3E+04	5.9E+02	1.4E+05
2-7	360	100	2.9E+04	1.1E+03	1.2E+05
2-8	72	10	1.2E+04	1.7E+02	5.5E+04

図 2: ユーザ反応時間分布 D_{random} 1-3(x 軸:反応時間, y 軸:頻度)

間分布の統計量を算出した。その結果を表 2,3 に示す。なお、最頻値は、 10^i が $10^i - 10^{i+1}$ を表している。2つの表において、最頻値と中央値については、 D_{matome} の方が、低い傾向が見受けられる。

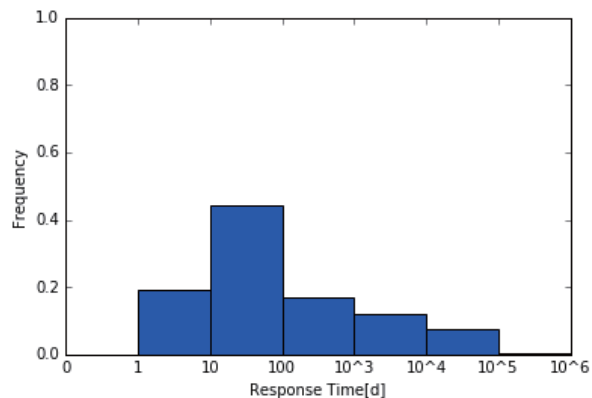
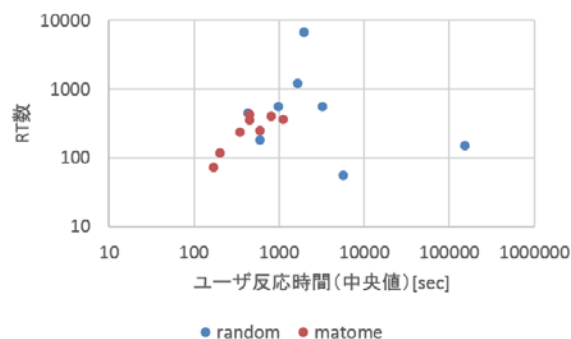
図 2, 3 に代表的なツイート 1-3, 2-5 におけるユーザ反応時間のヒストグラムを示す。2つの図において、 D_{random} に属するツイート 1-3 のがやや対照的なグラフになっているのに対し、 D_{matome} に属するツイート 2-5 の反応時間の分布は、より短い時間の方向に偏っていることが見受けられる。

最後に、各ツイートのユーザ反応時間とリツイート数の関係性をプロットした散布図を図 4 に示す。図 4 より、 D_{random} よりも D_{matome} の方が短い時間で反応している傾向が見受けられる。

ただし、いずれの分析においても、今回用いたデータセットが極めて小規模であるため、有意な傾向が見えたとはいえない。あくまで仮説を指示する可能性がある傾向が示唆されたに過ぎない。今後は、データセットを大規模化し、今回の分析で示唆された内容を有意に示せるような結果を得たいと考えている。

5. おわりに

本稿では、ソーシャルボルノという仮説を提案し、それによりソーシャルメディア上の情報拡散における様々な現象を説明できることを述べた。分析においては、「脊髄反射的に拡散する」というソーシャルボルノの定義に基づき、ユーザが情報を受け取ってから再発信（拡散）するまでの時間:ユーザ反応時間に注目し、一般的な投稿と「ソーシャルボルノ」的な投稿に

図 3: ユーザ反応時間分布 D_{matome} 2-5(x 軸:反応時間, y 軸:頻度)図 4: ユーザ反応時間分布 D_{matome} 2-5(x 軸:反応時間, y 軸:RT 数)

ユーザの反応時間に差があるのかを分析した。結果として、仮説を指示する可能性が示唆された。今後はデータセットを大規模化し、ソーシャルボルノの存在の立証およびそのフェイクニュースや炎上への影響について明らかにしていきたい。

参考文献

- [Blondel 08] Blondel, V. D., Guillaume, J.-L., Lambiotte, R., and Lefebvre, E.: Fast unfolding of communities in large networks, *Journal of Statistical Mechanics: Theory and Experiment*, Vol. 2008, No. 10 (2008)
- [鳥海 14a] 鳥海不二夫, 榎剛史: STAP 問題に対するソーシャルメディアにおける反応の分析, 第 7 回 Web とデータベースに関するフォーラム WebDBForum2014 (2014)
- [鳥海 14b] 鳥海不二夫, 榎剛史, 岡崎直観: 「人工知能」の表紙に関するツイートの分析・続報, 第 4 回 Web インテリジェンスとインタラクション研究会 ARG (2014)
- [鳥海 17] 鳥海不二夫, 榎剛史: バースト現象におけるトピック分析, *情報処理学会論文誌*, Vol. 58, No. 6, pp. 1287-1299 (2017)