Vol.2017-DPS-172 No.10
Vol.2017-SPT-26 No.10
Vol.2017-EIP-78 No.10
2017/11/29

デジタル・ゲリマンダーと SNS 選挙干渉の抑制方法の工学的検討

金子 格 †

東京工芸大学†

概要：ICT を利用して選挙干渉を行ういわゆるデジタル・ゲリマンダーの可能性が論じられている．2016 年には実際に米国で SNS を使った外国による選挙への干渉を試みたと思われる事例が明らかになった．主要国の選挙に他国が干渉を試みたと疑われる事例は大きな話題になったが，大国が相互に影響を及ぼそうとすることは国際社会の現実であり驚くべきこととはいえないかもしれない．しかし今回はたまたま他の事情により通常より念入りな調査が行われ明らかになっただけだとも思われる．すなわち今回とられたのと同じ手法は日常知られることなく企業間競争や，その他の目的で利用されている可能性がある．そしてそうした操作が実際に効果的でありかつ蔓延していることが明らかになればネット社会における wisdom of cloud の信頼性を著しく損なうことが懸念される．一方で SNS が本来もつ個人の自由な情報交換としての特性を棄損しないためには，安易な規制も控えるべきだろう．
そこで本論では SNS に類似した様々なシステムにおいて，すでに取られている手法を修正し組み合わせるなどして，SNS が本来もつ自由な情報交換手段という特徴を失うことなく効果的に取り得る対策がないかを検討する．

A consideration of technological tools for the suppression of digital gerrymandering and SNS interference to the election

ITARU KANEKO†

1. はじめに

ビッグデータ，人工知能，SNS など ICT の発達により「広告」を含む大衆に影響力を及ぼす方法が多様化している．選挙干渉手段にも多様化がみられ，既存の選挙法による規制が想定しないような選挙干渉が可能ではないかという疑いがある．そのような干渉行為の有無や効果がわかりにくく不透明化していることは，選挙の公正性を保ちより適切な結果を得る上で問題と考えられる．

ICT を利用した選挙干渉には様々なタイプがあり呼称も確立していないが，ここではデジタル・ゲリマンダーと SNS 選挙干渉と呼ぶこととし，その手法と防御手段について論ずる．これらの選挙干渉を技術的手段，法的規制，社会的リテラシーの向上のいずれか一つで抑制することは難しいがそれらを組み合わせて抑制することが考えられる．

2. モデル化

2.1. ゲリマンダーの数量的モデル

工学的に検討するため，まず問題の数量的モデルがあると有用である．ここでは，最も単純なモデル化を提案する．

2.2. 地理的ゲリマンダーのモデル

まず地理的ゲリマンダーを図 1 で説明する．この例は定員 2 の選挙区における選挙区分けの操作を行うことで当選者数を増やす状況を示している．(a)では□派と×派がそれぞれ一人ずつ当選する．この場合×が勝利したエリアでは×の支持者にかなり余裕がある．区割りを調整し，この×の支持者をもう一つの選挙区に多く配置されるようにすれば，(b)のようにもう一つの選挙区でも×派が勝利する．

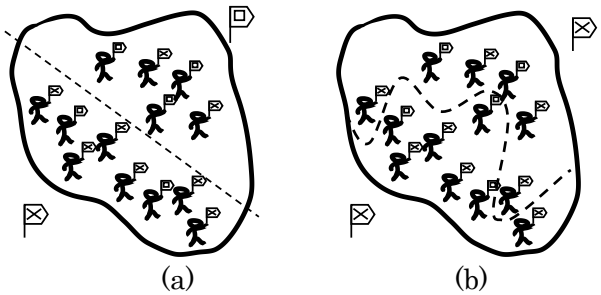


図 1 地理的ゲリマンダーにおける区割り操作.

□派と×派の選挙戦において，(a)の区割りでは双方が 1 つずつの選挙区で勝利するが(b)では 2 つの選挙区で×派が勝利することができる.

† 東京工芸大学
Tokyo Polytechnic University

ⓒ 2017 Information Processing Society of Japan

Vol.2017-DPS-172 No.10
Vol.2017-SPT-26 No.10
Vol.2017-EIP-78 No.10
2017/11/29

いま，□派と×派がC個の選挙区でそれぞれ一人ずつ候補をたてて争っているとすれば，その状況は以下のように行列を使って簡単に表現できる．

有権者，選挙区，候補をそれぞれ，
K人の有権者，$1 \le k \le K$
B個の選挙区，$1 \le b \le B$
×派=1，□派=2 として
C=2 種類の政党を$1 \le c \le 2$
で，表す．

有権者はいずれか一つの選挙区に属するので，有権者kが選挙区bに属する場合を$s_{kb} = 1$，そうでない場合$s_{kb} = 0$とすると，選挙区割はK行B列の行列で表すことができる．

選挙区割 ： $S = (s_{kb})$

有権者kが政党cを支持してる場合を$v_{ck} = 1$，そうでない場合$v_{ck} = 0$とすると，有権者の政党支持傾向も 2 行 K 列の行列で表すことができる．

正当支持傾向:$V = (v_{ck})$

各選挙区の得票数Wは以下の簡単な式で表すことができる．

得票数: $W = (w_{cb}) = VS$

ここでw_{cb}は候補cの選挙区bにおける得票数になる．

各□派($c = 1$)の得票数が×派($c = 2$)の得票数より大きければ当選となるから，全選挙区における□派の当選者数は以下の式で表現できる．H_Oは Heaviside の階段関数で$W_{1,b} > W_{2,b}$ の場合に 1，そうでない場合に 0 とり，その合計は□派の当選者数となる．(同点の場合当選ではないとする)．
□派の当選者数: N

$$N = \sum_{b=1}^{B} H_O(W_{1,b} - W_{2,b})$$

地理的ゲリマンダーは，このように比較的簡単にモデル化できる．

工学的にはこのようにして定式化されたモデルの「最適解」を求めたり，操作を困難にしたりすることが研究テーマとなるだろう．

2.3. デジタル・ゲリマンダーの定義

ディジタルゲリゲリマンダの定義は文献により異なっている．これらを整理して湯淺は①コンピュータ技術を使って恣意的な選挙区割り（地理的ゲリマンダーの高度化），②統計的データ分析（ビッグデータ分析）を用いて選挙区割以外の方法により投票結果にバイアスをかけること，③SNS などでメッセージの伝達にバイアスをかけることによって誘導を行うこと，という 3 つの類型を示した[1]．

①は，コンピュータ技術を使いはするが，モデルとしては地理的ゲリマンダーと同一でよい．②，③の地理的ゲリマンダーとの関係は後で SNS による選挙操作として述べるが，②は後述のように「投票効果のない選挙区」を含めることにより地理的ゲリマンダーの一種としてモデル化可能であると考えられる．③の SNS の誘導の作用は地理的ゲリマンダーのモデルとの関係は薄いが，後述の echo chamber の研究との関連が高いと思われる．

2.4. 類型①と類型②の関係- SNS 干渉の効果

湯淺による類型②では SNS 等により投票行動を選択的に促すことにより選挙に影響を与える可能性が生じる．類型②と類型①は一見異なるがどのような関係にあるだろう．そして SNS による干渉はどのような効果を生じ得るだろう．

まず SNS 機能を使って特定候補を直接応援する場合はここでは除外する．なぜならもしも，SNS に介入できるだれかが特定の候補者に直接的に有利・不利になる介入をすれば，たとえば特定の候補者を支持するメッセージのみを伝達したり妨害したりすれば, SNS を通信手段として用いた単純な選挙応援である．そのようなケースをゲリマンダーに含めてもよいかもしれないが，特に今までの選挙活動と異なるメカニズムが生まれるわけではない．

そうではなく，デジタル・ゲリマンダーの類型②にあたる場合，すなわち有権者がどの候補を支持するかという選択には影響を与えずに，投票を行うか行わないかという決定のみに関与する場合を考える．

まず，これを候補者の応援，妨害と単純にとらえることはできない．なぜならメッセージは候補への支持や投票を直接含んでいないからである．したがって選挙結果への影響が不透明であり法的な規制もむずかしい．

次に後者の方法による選挙介入を行いたい場合，どのようにメッセージの対象を選べば効果を得られるだろうか．今選挙区b_0の有権者のほとんどはすでにある候補を支持しているとする．選挙区b_1の有権者kの選挙区をb_0に移動したとすると，b_1の結果に影響を与えうる．一方kが投票に参加しないように誘導してそれが成功した場合も選

ⓒ 2017 Information Processing Society of Japan

Vol.2017-DPS-172 No.10
Vol.2017-SPT-26 No.10
Vol.2017-EIP-78 No.10
2017/11/29

挙区b_1の結果に与える影響は同じであろう．有権者kを当選結果に寄与しない別の選挙区b_0に割り当てることの効果と同等である．

つまり①の地理的ゲリマンダーに，新たな有権者を加えても影響がない仮想的な「投票効果のない選挙区」を加えるよう拡張すれば，その仮想的な選挙区を加えて選挙区割の最適化を行った場合と，同じ効果を持つ．したがって SNS による後者の方法によるこのような干渉の効果は①の類型と全く同じではないが，その拡張モデルとして扱い得ると考えられる．

3. 抑制方法の工学的手段の考察

つぎに，デジタルゲリマンダーや SNS による選挙干渉の抑制方法について工学的に取りえる手段を考察する

3.1. 推薦システムの信頼性推定手法の適用

今日のネット販売では利用者が商品の推薦記事を提供する推薦システム(recommender system)が必須である．そのようなシステムにおいて，推奨対象者自らがバイアスのあるメッセージ(公平な評価より高い評価を自身に与えたり，他者に低い評価を与えたりするメッセージ)を投稿することを防ぐ必要がある．そうした投稿が増えると推薦システムへの信頼が損なわれてしまう．評価の信頼性を向上するには大きく「異常値」を検出し除外する方法と，評価の「多様性」を高める手法がある．

アカウントで発信者が特定できれば大量投稿の検出は容易だ．しかし実際には偽アカウントが利用されうるので，アカウントによる発信者特定はできないという前提でこれらの手法が研究されている．たとえば Vargas 等による推奨記事の統計的特徴を利用して推奨者の新規性，信頼性，多様性を評価する手法を示した報告[3] はこの分野では最も引用件数が多い研究報告である．またこの手法は投稿者の多様性を担保するが，同時に偏った投稿を大量に行うという意味で「異常」な投稿者を自動的に排除する効果も期待できる．

デジタル・ゲリマンダーを意図して SNS に投稿を行う場合，結果に影響を及ぼすためには意図的な投稿を大量に投稿する必要があるだろう．大量の異常なメッセージを検出，除去する技術はデジタル・ゲリマンダーの有効な抑制手段の一つになると考えられる．

異常値を除去することは選挙に保守的なバイアスをかけることにならないだろうか．異常値を制限すべきなのは投票へ行くようにという勧奨だけで，「候補 A に投票してください」というメッセージは適法な選挙活動であり全く制限されない．デジタル・ゲリマンダーの問題はステルス性(だれが何のためにやっているかが見えないこと)と，費用や実施時期の監視ができないことである．投票の勧奨に関する投稿でありかつ異常な投稿のみを検出，抑制することは，選挙運動に保守的なバイアスをかける恐れなくデジタル・ゲリマンダーを効果的に抑制する手段となりそうである．

3.2. Echo Chamber の推定手段の適用

SNS の研究において Echo chamber とは特定の意見を支持する集団内でお互いの間でのみコミュニケーションが行われることを意味する．

Echo chamber という言葉は音響の分野でもともと用いられていた．音楽演奏などに心地のよい残響を響かせるための部屋を意味する．2000 年ごろから SNS における現象にもこの言葉が使われ始めた．SNS のフィルター機能により促進され，利用者を泡の中に取り込んでしまうように見えるためフィルターバブルとも呼ばれる．

Colleon 等は 2009 年の 4.6 億件のツイートを分析し，民主党支持者と共和党支持者が内輪のみでメッセージ交換をしているかを確認した[4] ．民主党支持者は民主党の支持者へのフォローが 82%，共和党支持者へのフォローが 12%だったが，共和党支持者は民主党支持者のフォローが 76%，民主党支持者へのフォローが 23%であり，この分析では極端な Echo chamber は発生していないことを示す結果となった．

一方笹原もツイートデータベースを分析した[5] が，この場合は同一集団内のつながりの強さが見いだされ，笹原はツイートの連携の強いグループをソーシャルグラフ化した．またツイート数とユーザ数の関係を分析し，グラフからは 1%程度の小数のユーザが半分以上のツイートを行っている状況が示唆された．

Echo chamber が発生している場合，利用者は多数の閉じたグループにわかれ，それらはごく一部の利用者が大部分のメッセージを発信することで結合している．もし，投票行動を促すメッセージをあるグループで大量にリツイートされるよう調整して発信すれば，グループ内で効率的に反復拡散し，グループ単位で投票行動に影響を与えることが可能となる．こうして Echo chamber は類型②のデジタル・ゲリマンダーを効率的，効果的に実行する手段となりえる．

この時選ばれる Echo chamber 自体は必ずしも特定の候補を支持している必要はない． Echo chamber の支持者が目標政党に対立候補の支持者であっても利用できる．なぜなら，もしその Echo chamber のメンバーが当選させたい政党を

ⓒ 2017 Information Processing Society of Japan

Vol.2017-DPS-172 No.10
Vol.2017-SPT-26 No.10
Vol.2017-EIP-78 No.10
2017/11/29

支持する傾向があれば「投票に行くように」，そしてその逆であれば「投票に行かないように」誘導すればよい．

このように Echo chamber はデジタル・ゲリマンダーを行うために容易かつ効果的な手段であり，その研究はデジタル・ゲリマンダーの利用や抑制に役立つと思われる．

Echo chamber 自体は利用者本人の選択であり，Echo chamber 自体を規制することは自己決定権の裁量範囲である．Echo chamber はフィルターバブルとも呼ばれており，その形成には SNS の運営も関与している．そこで Echo chamber を利用した過度に偏った誘導を「自己決定権」との関係で板倉は論じているが，このアプローチで効果的な法規制を行うことは困難ではないかと考えられる．

一方で，Echo chamber の客観的指標は工学的に示しうる．利用者がどの程度 Echo chamber 囲まれているか，Echo chamber に囲まれているとした場合にどのような属性になっているかを利用者自身が確認することはプライバシー権の行使として認めることが可能ではないかと考えられる．これを利用者が自主的に確認可能とすることで，利用者自身の Echo chamber を利用した誘導に対する耐性を高められるかもしれない．もしそのような効果があれば Echo chamber の検出や客観指標の導出手法が Echo chamber による誘導の防止にある程度効果を発揮すると期待できる．

3.3. ゲリマンダーに耐性がある選挙制度

Erdélyi は様々な選挙制度と選挙操作について調べた．特に選挙方法毎に，選挙操作への頑健性を細かく調査している．Erdélyi はマルチエージェント研究の立場から分析をしている．まずゲリマンダーの選挙制度の単純な関係として，ゲリマンダーは当選に結び付かないいわゆる死に票を再配分する操作であるから，大選挙区や比例代表では抑制され，小選挙区で敏感であることは理解しやすい．しかし Erdélyi 等は詳細な理論的分析を展開し，縮退投票(Fallback Voting)がゲリマンダーに対し強い耐性があることを示した[6] ．

Erdélyi が例示するモデルは 2 段階投票で，ゲリマンダーの目的は有権者を 2 グループに分けた時に意図的にそれぞれのグループが選ぶ候補を制御できるかという問題である．Erdélyi は，そのような操作が NP 困難(NP hard)であることを「意図的な操作が実際上は不可能である」ことから「頑健である」と定義して分析を行っている．多くの選挙方式を比較し，その中で Fallback Voting が区割り操作に対し計算量的に頑健(computationaly registant)であることを示している．

4. まとめ

デジタル・ゲリマンダーによる選挙介入は可視性がない(だれが何を目標に行っているか，そもそもその影響を受けているか否かがわかりにくい)ことが不安を誘う．アシモフの「ファウンデーション」シリーズ[7] に登場する「心理歴史学」のように政治が何者かに思いのままに操られているのでは，という不安である．

デジタル・ゲリマンダーを好ましくないと考える意見はありえるが，だからといってデジタル・ゲリマンダーを行う行為自体を検出し禁止しようとすると別の問題を生じる．それは SNS において特定の発言を検出したり禁止したりする手段を意味し，それでは SNS が持っている自由に能率的に情報を交換するという長所を大きく棄損してしまう．そのような方法での解決には，当然強い抵抗があると予想される．

一方，有権者自身からみてデジタル・ゲリマンダーは自身の選択権を奪う行為である．有権者は当然デジタル・ゲリマンダーの影響から逃れたいという意志を持つと期待できる．したがって技術的手段でデジタル・ゲリマンダーそのものを制約するのではなく，有権者がデジタル・ゲリマンダーの影響を自ら発見し回避する手段を提供すれば，発言の取捨選択は有権者自身の判断で行うことができるから，SNS の情報交換手段としての自由度を失うことなくデジタル・ゲリマンダーの抑制を行うことが期待できる．

一連の関連研究を見る限り，デジタル・ゲリマンダーを抑制する手段は多い．したがってステルス性の完全な解消は難しくても，デジタル・ゲリマンダーをできるかぎり検出可能とするとは，おそらく可能と考えられる．

一方地理的ゲリマンダーに関する研究成果は，選挙区割りの操作の影響の最小化につなげられる可能性があるが，同時に選挙の公平性を確認したり，より公平な選挙を実現したりする指針が得られるという面でも有益と考えられる．

デジタル・ゲリマンダーをただちに完全に抑制する特効薬はない．しかし利用者が望めばそれを抑制する工学的手段や情報提供は，SNS のシステム運営者やボランティア的な団体により可能であると考えられる．SNS は発展途上のメディアであり，弊害を認識しつつ，その抑制手段のトライアンドエラーを積み重ねて行くことで，より信頼できるものになっていくことを期待する．

ⓒ 2017 Information Processing Society of Japan

Vol.2017-DPS-172 No.10
Vol.2017-SPT-26 No.10
Vol.2017-EIP-78 No.10
2017/11/29

[参考文献]

[1] 湯淺 墾道, ディジタルゲリマンダの法規制の可能性, 情報処理学会, 情報処理, vol. 58, no.12 (2017)
[2] 板倉陽一郎, ディジタルゲリマンダとプライバシー，自己決定権，情報処理学会, 情報処理, vol. 58, no.12 (2017)
[3] Saúl Vargas , Pablo Castells, " Rank and relevance in novelty and diversity metrics for recommender systems,” ACM Proc. RecSys'11, 2011
[4] E Colleoni, A Rozza, A Arvidsson, "Echo chamber or public sphere? Predicting political orientation and measuring political homophily in Twitter using big data,” Journal of Communication, 2014 - Wiley Online Library, 2014
[5] 笹原 和俊,「連想ネットワークによる集合注意の可視化」, 人工知能学会論文誌, 論文 ID: B-MDF02,2015
[6] Gábor Erdélyi, Lena Piras, Jörg Rothe, “The complexity of voter partition in Bucklin and fallback voting: solving three open problemsm,” Proc. AAMAS’11,Vol.2,p.837-844
[7] アシモフ,「ファウンデーション」, ハヤカワ文庫 SF, (1984)

ⓒ 2017 Information Processing Society of Japan