

ヒット現象の数理モデルによる AKB 総選挙予測 Prediction of "General Election" for AKB using Mathematical Model of Hit Phenomena

太田 奨[†] 石井 晃[‡]
Sho Ota Akira Ishii

1. はじめに

AKB48 は「会いに行けるアイドル」をコンセプトに秋本康がプロデュースし 2005 年に結成されたアイドルグループである。2009 年頃から徐々にテレビ番組、CM 等への出演が増加し現在では「国民的アイドル」の地位を確立した。この地位を確立させるべく従来のアイドルグループとは異なる手法でファンの心を掴んできた。この手法は AKB 商法といわれるもので CD を購入の特典として AKB 総選挙や握手会の参加券などを同封するということが 1 つ挙げられる。AKB 総選挙とはシングル CD のセンターポジションをかけて AKB グループ内で選挙を行い、順位付けをする選挙のことである。ファンはこの選挙で自分の応援しているメンバー（通称押しメン）のために大量の CD を購入するものたちがたくさん出現し問題となった。

そんな中、近年 Twitter、Facebook、mixi、Blog など手軽に思想を書きめる媒体や YouTube やニコニコ動画など自由に動画が投稿できる媒体（ソーシャルメディア）が急速に発展し普及してきた。そして現在、ソーシャルメディア上には多くの人々の思想があふれている。また、石井・吉田の提唱したヒット現象の数理モデルは先行研究により、映画、地域イベント、音楽など様々なジャンルに応用して使えることも分かっている[1][2]。

これらのことからソーシャルメディアから AKB48 についての書き込みを抽出し、数理モデルを用いて分析することにより、AKB 選挙の当選傾向がわかるのではないかと考えた。そして今回の研究はヒット現象の数理モデルを国政選挙に応用して行く第一歩としてクチコミ数・票数が多く分析しやすい AKB 総選挙の当選傾向をヒット現象の数理モデルを用いて探っていく。

2. ヒット現象の数理モデル

2.1 ヒット現象とは

ヒット現象とは、あるコンテンツ(音楽や映画等)が消費者の支持を得て、CD 等の販売数や観客動員数を爆発的に増やし、ピーク後には速やかに減衰する現象である[3]。このヒット現象の発生には、広告・宣伝効果のほかに 2 種類のコミュニケーションが大きく関わっている。今回扱う AKB 総選挙についても中間速報日と開票日付近のブログ書き込み数に関してヒット現象が起こっている。

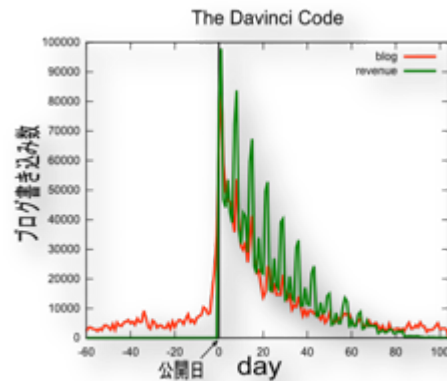


図 1「ダ・ヴィンチ・コード」

図 1 は映画「ダ・ヴィンチ・コード」の観客動員数とブログ書き込み数に定数倍したものをグラフにしたものである。図 1 から分かるように観客動員数とブログの書き込み数には比例関係が見られた。この結果からブログの書き込み数を実際のデータの代わりに使用することが可能だと分かったので、映画分野におけるヒット現象の研究が積極的に行われた。今回の研究に関しては、投票意欲の変化を実際に表したデータというのは現代社会には存在しないため、ブログの書き込み数を投票意欲の数と見なすことでシミュレーションを行うこととする。

2.2 二種類のコミュニケーション

2.1 で述べたようにヒット現象の発生には 2 種類のコミュニケーションが大きく関わっている。その 2 種類のコミュニケーションとは直接コミュニケーションと間接コミュニケーションである。

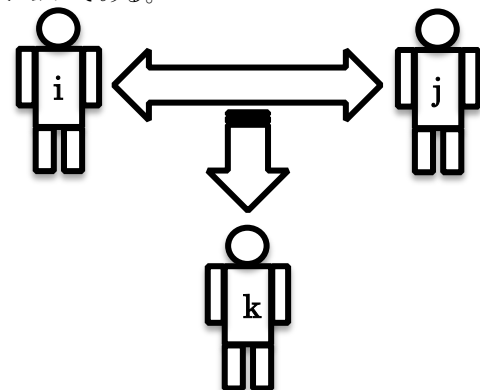


図 2 コミュニケーション関係図

直接コミュニケーションとはあるコンテンツについて j が i に直接的な会話のやり取りをすることで、投票意欲を刺激することである。

[†] 鳥取大学大学院 工学研究科 機械宇宙工学専攻 応用数理工学

[‡] 鳥取大学大学院 工学研究科 機械宇宙工学専攻 応用数理工学 教授

このやりとりがコンテンツに対していい (positive な) 内容であれば投票意欲が上昇し、悪い (negative な) 内容であれば低下する。

間接コミュニケーションとはあるコンテンツについての i と j 会話を第3者の k が聞き影響を受けること行為である。

直接コミュニケーションと同様に良い内容の (positive な) 会話を聞けば投票意欲が上昇し、悪い (negative な) 内容の会話を聞けば低下するが、直接コミュニケーションより購買意欲に大きな影響を与えることがわかっている。近年ではインターネットや SNS などの発達に伴い、コミュニケーションが行われる頻度が急激に増加している。直接コミュニケーションでいえば電子メールや、SNS のメッセージ機能などが上げられる。間接コミュニケーションでいえば Twitter の Retweet などが上げられる。Twitter や掲示板、Blog のコメント欄などでおこる炎上や東日本大震災の Twitter による情報の拡散[4]を想像すれば、間接コミュニケーションがヒット現象に大きく関わっていることが容易に分かるだろう。

2.3 ヒット現象の数理モデル

「購入意欲」には宣伝・広告と、二種類のコミュニケーションが影響を与えると考えられる。この二種類のコミュニケーションをモデル化していく。

まず、直接コミュニケーションについてである。購入した人からの情報は、それまでに購入した人の人数に比例する i 番目の人が j 番目の人との直接のコミュニケーションにより情報を得、購入する気になる確立を D_{ij} と定義すると、購入した j 番目の人からの情報により i 番目の人が購入する確立は、

$$\sum_{j \neq i}^N D_{ij} I_j(t) \quad (2.1)$$

と表わされる。

次に間接コミュニケーションについてである。間接コミュニケーションでは、購入した人同士の情報は購入した人で作る対の数に比例する。購入した人同士の情報交換は、口込みの場合もあればブログやネットの掲示板の場合もあると考えられ、その情報量の総数が問題となる。購入した人同士 (j 番目の人と k 番目の人) のコミュニケーションの影響から i 番目の人が購入する確立は、

$$\sum_j \sum_k P_{ijk} I_j(t) I_k(t) \quad (2.2)$$

と表わされる。ここで j, k は i を含まない。係数 P_{ijk} はその商品に対する評価であり、購入した人の評判が良いほど、この係数は大きな値となる。逆に購入した人に評判が悪いと、この係数は負となる。

ヒット現象の個人による投票意欲の時間変化を表わす基本方程式は、2.1、2.2 で定義したコミュニケーションの項と、宣伝広告の効果を外場 $A(t)$ として表した項を足し合わせたものである。よってヒットの基本方程式は、

$$\frac{dI_i(t)}{dt} = A(t) + \sum_{j \neq i}^N D_{ij} I_j(t) + \sum_j \sum_k P_{ijk} I_j(t) I_k(t) \quad (2.3)$$

と表わされる。[1]

また $A(t)$ について

$$A(t) = \sum_{\xi} Cadv_{\xi} A_{\xi}(t) \quad (2.4)$$

と拡張する。これで各 $A_{\xi}(t)$ に TV 露出やネットニュース、2ちゃんねる Facebook 広告などを別々に入力し、各 $Cadv_{\xi}$ を最適化することで、各広告メディアの効果の強さを推定することができる。

(2.1) で定義したヒット現象の数理モデルは、投票意欲が広告宣伝と二種類のコミュニケーションによって変化することを表す。つまり選挙でいう意欲変化の激しい浮動票が誰に投票されたのかを判断する1つの指標になると考えられる。

2.4 誤差の指標 (R-factor)

R-factor は、電子線解析の研究において実際の電子線のデータと実験によるデータを比較し、精度を検討するためにイギリスの理論物理学者 J.B.Pendry によって提案された式である[5]。

本研究では、以上の J.B.Pendry の R-factor をもとに次のように定義して用いる。

$$R = \frac{\sum_i \int (f(i) - g(i))^2 dE}{\sum_i \int ((f(i))^2 + (g(i))^2) dE} \quad (2.8)$$

ここで、本研究では $f(i)$ はある映画に対するブログの書き込み件数の値、 $g(i)$ はある映画のヒット現象の数理モデルによるシミュレーション結果の値を用いる。

ヒット現象のシミュレーションは以後この R-factor の値を精度の判定基準と定義し、0 に近ければ近いほど高精度となる。

3. 分析対象と結果

3.1 分析対象について

分析対象者は第3回 AKB 総選挙の上位4名について表1に、分析対象イベントは表2にまとめた。

表1 分析対象者情報

分析対象	大島 優子	渡辺 麻友	柏木 由紀	指原 莉乃	篠田 麻里子
中間順位	1	3	2	4	6
最終順位	1	2	3	4	5
最終票数	108837	72574	71076	67339	67017

表2 分析対象イベント

分析イベント	分析開始日	イベント日	分析最終日
AKB総選挙中間速報	2012.5.15	2012.5.23	2012.5.31
AKB総選挙开票	2012.6.1	2012.6.6	2012.6.10

表1の最終票数を見ると分かるが、大島優子はダントツの1位だが、2,3位と4,5位は非常に僅差である。

今回は分析対象者の名前が含まれるブログ上の書き込みを抽出し、各イベントの書き込み数の変化をシミュレーションする。

4. 結果と考察

まず、シミュレーション結果を図3-6に示す。

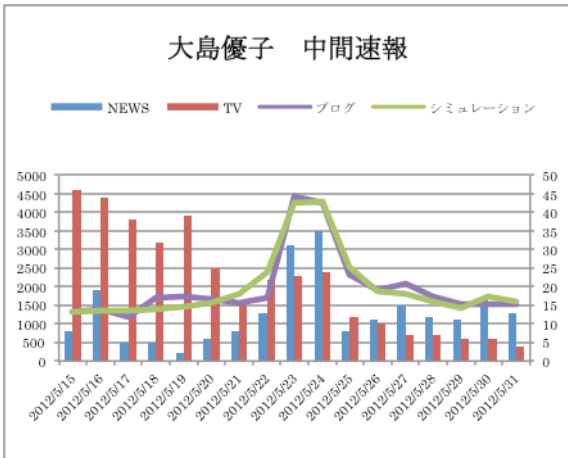


図3 大島優子 AKB総選挙中間速報 グラフ

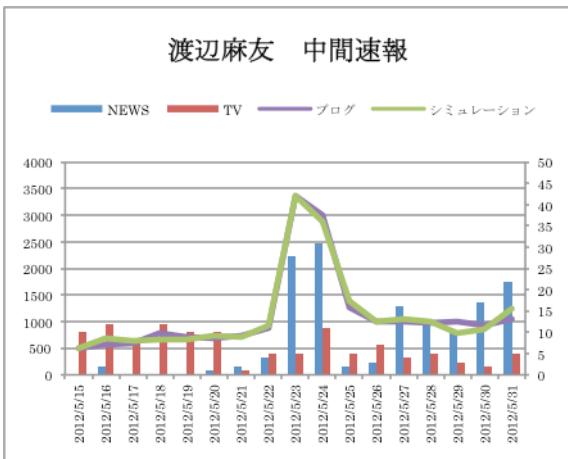


図4 渡辺麻友 AKB総選挙中間速報 グラフ

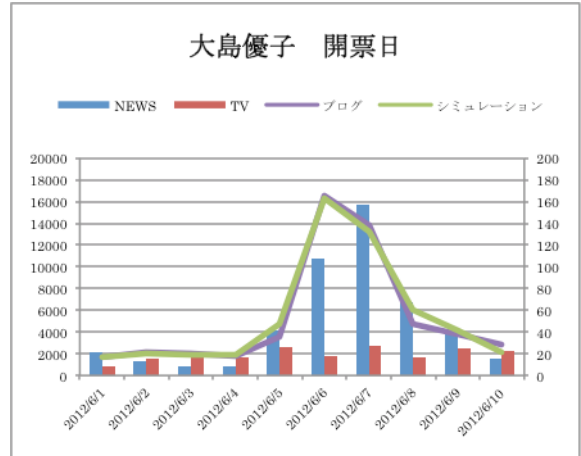


図5 大島優子 AKB総選挙开票日 グラフ

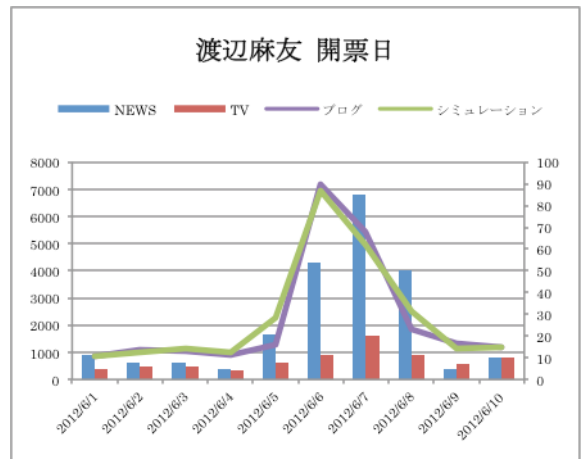


図6 渡辺麻友 AKB総選挙开票日 グラフ

各パラメータの数字 1,2 はそれぞれピーク前、ピーク後、Cadv は1回の広告露出回数に対する効果の割合、tv はテレビ露出、news はネットニュース(ネット上のニュース記事)、D は直接コミュニケーション、P は間接コミュニケーションをそれぞれ表している。

今回行ったすべてシミュレーションの誤差が3%以下であり、精度が確保されている。

また、図5,6,7に各イベントについてコミュニケーションの強さをグラフにしてまとめた。

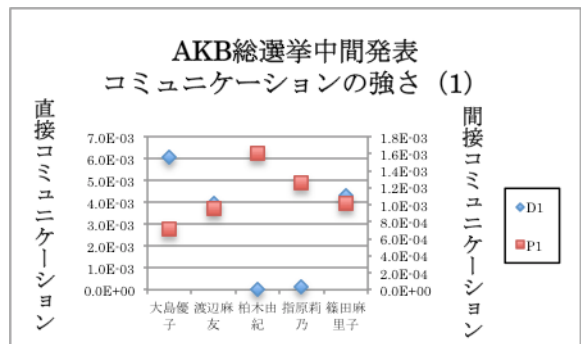


図7 AKB総選挙中間発表 ピーク前 コミュニケーションの強さ

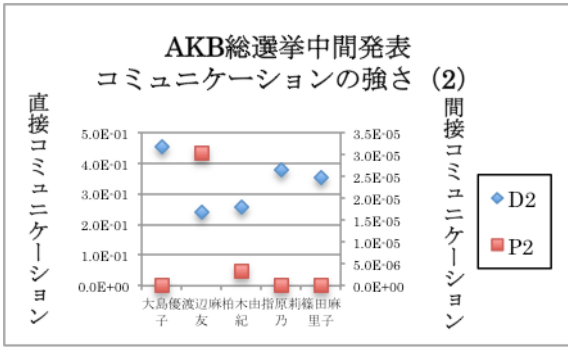


図 8 AKB 総選挙中間発表 ピーク後 コミュニケーションの強さ

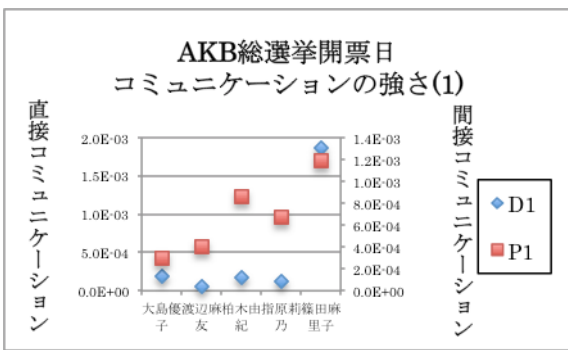


図 9 AKB 総選挙開票日 ピーク前 コミュニケーションの強さ

図 7,8,9 を見るとコミュニケーションの強さは順位と一致していないことが分かる。これは、意欲の変化は投票に置き換えて考えると浮動票の推移と見なせることができるだろう。また、直接コミュニケーションにも関係性は見られないが、図 7,9 (書き込み数の上昇部分) に着目すると中間発表と開票日の間接コミュニケーションがほぼ比例していることが分かる。したがって、①中間発表からの番狂わせが上位ではほぼ起こらない②中間発表のシミュレーション結果から投票日の書き込み数の急増を再現できるということが分かる。①について具体的にいうと今回の大島優子の獲得票数の中に閉める割合は固定票が大半で浮動票が少ないということが間接コミュニケーションの強さから分かる。

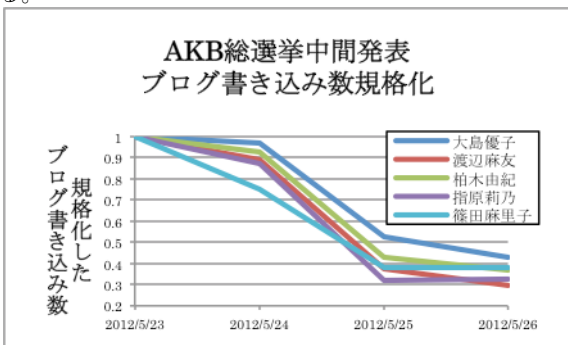


図 10 AKB 総選挙中間発表 ブログ書き込み数規格化

また、図 10 のように中間発表日の書き込み数に併せて規格化すると、減衰の仕方が緩い順と当選順位がほぼ同じということも分かる。

5. まとめ

- 1) 中間発表と開票日のコミュニケーションに比例関係が見られた。
- 2) 中間発表後の書き込み減少割合が少ないものが当選順位の上位にくる傾向がある。

という 2 つが分かった。

このようにヒット現象の数理モデルは投票意欲の変化を見ることが十分に可能な理論であるといえる理論だ。正確な固定票数さえ分かれば今後精度の高い順位予測をすることが可能となるはずだろう。さらにいえば、国政選挙における当選予測もソーシャルメディアから書き込みを抽出できれば可能となるだろう。

謝辞

「クチコミ@係長」の使用を快く許可して下さった株式会社ホットリンク、テレビ露出データの使用を快く許可して下さった株式会社エム・データに心から感謝申し上げます。

参考文献

- [1] 吉田 敦彦, 石井 晃, 新垣 久史, “大ヒットの方程式”, ディスカヴァー・トゥエンティワン, (2010).
- [2] Akira Ishii, Hisashi Arakaki, Naoya Matsuda, Sanae Umemura, Tamiko Urushidani, Naoya Yamagata, Narihiko Yoshida. (2012). The 'hit' phenomenon: a mathematical model of human dynamics interactions as a stochastic process.
- [3] 吉田 敦彦, “ヒット学”, ダイヤモンド社, (2005).
- [4] 篠田 孝祐, 榎 剛史, 鳥海 不二夫, 風間 一洋, 栗原 聡, 野田 五十樹, 松尾 豊, “東日本大震災時における Twitter の活用状況とコミュニケーション構造の分析”, 知能と情報 Vol.25 No.1 pp.598-608, (02/2013)
- [5] J.B.Pendry. (1980). ReliabilityFactorsforLEEDCalculations. J.Phys.