

検索行動の数理モデルを用いた季節性イベントの解析

○岡野のぞみ 東大将 石井晃 (鳥取大学)

Analysis of seasonal events using mathematical model of search behavior

*N. Okano, M. Higashi and A. Ishii (Tottori University)

概要— インターネット上の人々の関心を表す検索行動を、社会物理学の立場で理論として数理モデル化した検索行動の数理モデルを用いて Google Trend から検索行動を解析すると、マスメディア広告などの影響の他、ブログや Twitter の影響も定量的に評価することができる。本研究では日付が限定される季節性イベントを例にとって、そのイベントまでの盛り上がりイベント終了後の評判の終息を検索行動の数理モデルで解析する。

キーワード: クリスマス, ハロウィン, バレンタイン, 恵方巻き

1 序論

インターネット上の検索行動は、往々にして興味ある事柄、特に新たに目に入った新しい事柄で興味を持つ事を検索すると考えら得るので、ネット上の検索行動はその人の興味関心の移り変わりを表すと考えられる。Google Trend のようなサイトでの検索のデータは、世の中の人々の多くが何に興味を持って検索するかを表すデータになっていると考えられる。ブログや Twitter などソーシャルメディアの書込も社会の人々の興味関心を表す指標になるが、世の中には興味関心があっても、それをすぐにソーシャルメディアには書かない人も少なくない。大きな社会的事件があったとして、それについてソーシャルメディアに書く人がいる一方で、関心を持ちつつも何も書かない人の方が多い。そうした書かない人々も、関心がある話題についてインターネット上で検索するであろう。従って、人々の検索行動を解析することは、ソーシャルメディアへの書込とはまた違った意味で、世の中の人々の興味関心を解析することになるはずであり、いわばインターネットの世界の中で「声なき声」を聴くことになると思われる。

そうした人々の検索行動をヒット現象の数理モデル^{1, 2, 3)}にならう形で社会物理学の考え方で数理モデルにまとめ、Google Trend で実測される人々の検索数の時系列的な変化を再現できる数理モデル、検索行動の数理モデルが提案されている^{4, 5, 6, 7)}。先行研究でも検索数を人々の興味関心の指標とし、SIR モデルを用いて情報の拡散過程を説明するモデルが構築されているが⁸⁾、新たに提案されたモデルはより具体的なマスメディア露出やブログ、Twitter の影響を含めての人々の検索行動の動力学を記述するモデルである。それにより、今までからは読み取れなかった潜在層の人々の動きや、及ぼされる外的影響から新たな知見が得られると考えられる。

2 理論

検索行動の数理モデルはヒット現象の数理モデルにならう形で、社会の人々による検索行動の動力学を記述する新しい数理モデルである。その数理モデルでは人々の検索行動に、Blog や Twitter も影響を与えると考える。なお、検索数の測定データは Google Trends を使うが、Google Trends で用いている検索数のデータは 1 から 100 に規格化されたものであり、以下、ここでは関心度と呼ぶことにする。(Google はこれを人気

度と呼んでいる)

2.1 ヒット現象の数理モデル

ヒット現象の数理モデルでは、社会における人々の 1 人 1 人が抱く興味・関心を定量化して方程式にしている。ある人「i さん」が抱く興味・関心を $I_i(t)$ と定義し、興味・関心 ($I_i(t)$) を掻き立てる要員として、

1. メディアによる影響
2. 会話による影響
3. 噂による影響

の 3 つがあるとする。会話や LINE などによる直接的なやりとりによる影響のことを「直接コミュニケーション」とよび、街中の噂やソーシャルメディア上でのやり取りなどで影響を受けることを「間接コミュニケーション」と呼ぶ。それらについて興味・関心の時間的変化を追う微分方程式を立て、数理モデル化すると、以下の形で表せる^{1, 2)}。

$$\frac{dI_i(t)}{dt} = \sum_{\xi} c_{\xi} A_{\xi}(t) + \sum_{j \neq i} D_{ij} I_j(t) + \sum_j \sum_k P_{ijk} I_j(t) I_k(t) \quad (1)$$

ここで D_{ij} は直接コミュニケーションの強さを表す係数、 P_{ijk} は間接コミュニケーションの強さを表す係数である。広告宣伝の影響は外力と考える。ある話題に関する日毎のテレビでの露出秒数や、ネットニュースの件数を A_{ξ} とし、その係数を c_{ξ} とする。 ξ はメディアの種類を表す添え字である。

簡単化のために平均場近似を行う。社会全体の構成員の数を N 人とし、社会全体で平均化された人々の意欲・関心を $I(t)$ として、以下で定義する。

$$I(t) = \frac{1}{N} \sum_i I_i(t) \quad (2)$$

この平均場近似を用いると、 $I(t)$ に従う方程式は次のようになる。

$$\frac{dI(t)}{dt} = \sum_{\xi} c_{\xi} A_{\xi}(t) + DI(t) + PI(t)^2 \quad (3)$$

導出の詳細は¹⁾を参照されたい。
 実際に計算する際には左辺の微分の箇所を、

$$\frac{\Delta I(t)}{\Delta t} = \sum_{\xi} c_{\xi} A_{\xi}(t) + DI(t) + PI(t)^2 \quad (4)$$

と表す。本研究では1日1日の動向をこのヒット現象の数理モデルで探るので、 $\Delta t = 1[\text{日}]$ として、

$$\Delta I(t) = \sum_{\xi} c_{\xi} A_{\xi}(t) + DI(t) + PI(t)^2 \quad (5)$$

となる。つまり実際にはこの $\Delta I(t)$ で計算する。

2.2 検索行動の数理モデル

本研究では検索行動の数理モデル^{4, 5, 6, 7)}を用い、TV露出以外にブログやTwitter当該の話題の検索数も考えて次のように数理モデルをたてる。検索行動の数理モデルは、数学的構造はほぼヒット現象の数理モデルと同じであり、人々の興味関心の指標である $I(t)$ をGoogle Trendsの関心度とし、従来のヒット現象のモデルの広告宣伝の項には T_V 、ネットニュースの影響しか加味していなかったが、第5項のBlogの影響、第6項のTwitterの影響を追加した。つまり、外的な影響にBlogとTwitterが加わる形である。そして、ここでの $I(t)$ は書込数、投稿数ではなく、時刻 t における検索数である。

$$\begin{aligned} \frac{dI(t)}{dt} = & C_{TV} A_{TV}(t) \\ & + C_{NetNews} A_{NetNews}(t) + C_{Twitter} A_{Twitter}(t) \\ & + C_{blog} A_{blog}(t) + (D - a)I(t) + PI^2(t) \end{aligned} \quad (6)$$

2.3 精度の計算方法

また、係数をフィッティングの際に精度を表す指標として、2本の曲線の一致度を測る指標であるR_factorを用いる^{10, 11)}。本研究では以下の式を用いる。

$$R = \frac{\sum_i (f(i) - g(i))^2}{\sum_i ((f(i))^2 + (g(i))^2)} \quad (7)$$

ここで、例えば本研究では $f(i)$ はターゲット(Google Trendsで測定した検索数など)に対する関心度の件数、 $g(i)$ は検索行動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。

R は、 $0 \leq R \leq 1$ の範囲で、値が小さいほどフィッティングの精度が良いと言える。

3 題材の選択

本研究では日付が限定されるイベントについての検索を調べる。それは1日、あるいは多くても数日のイベント事についての検索は、そのイベントに向けての人々の関心の盛り上がり、イベント終了後の余韻とその終息がGoogle Trendsで測る検索数の時系列的な変化からはっきりと読み取れるからである。

日付が限定されて、かつ関心が高くBlogやTwitterへの書込も多く、検索数も多いイベントとして、以下のイベントを選んだ。

1. 新年のカウントダウン

2. バレンタインデー

3. ハロウィン

4. クリスマス

5. 恵方巻き

6. 土用の丑の日の鰻

土用の丑の日の鰻を除いては、いずれも日付は限定されている。特にカウントダウンは、正確に言えば旧年の12月31日23時59分から新年の1月1日0時0分までの一瞬である。土用の丑の日の鰻については1週間程度幅があると考えた。データはTV露出データ、ネット露出データ、Blog書込数、Twitter書込数、そしてGoogle Trendsによる検索数を用い、イベント前1ヶ月をイベント前の期間、イベント後の1ヶ月をイベント後の期間とした。土用の丑の日については、その前後1週間を両方で重ねて解析している。新年のカウントダウンの場合、クリスマス後にならないと人々の関心が高まらない事から、12月26日から31日をイベント前、1月1日から7日をイベント後の期間とした。測定とそれによる計算は、いずれも2016年のものである。

4 計算

まず、検索行動の数理モデルがどれほど正確にGoogle Trendsの検索数の時系列を再現しているかの例を、2016年のハロウィンの場合を例に取って、図1、2、3、4に示す。図からわかるように、計算した検索数と実際にGoogle Trendsで測定した検索数との一致は非常にいい。以下ではこうして精度良く計算した結果、実測を説明する数値として計算から得られた係数 C_{blog} と $C_{twitter}$ を、そのイベント前あるいはイベント後の検索がBlogに影響される強さ、Twitterに影響される強さの測定値と考える。

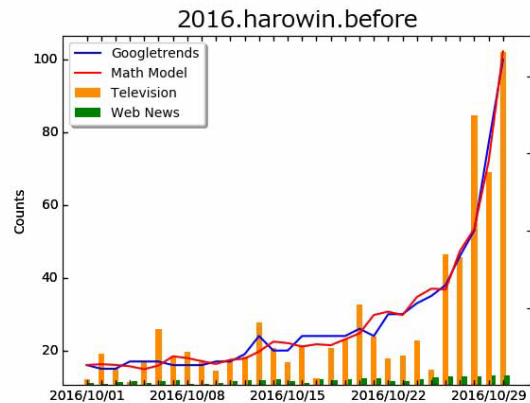


Fig. 1: 2016年のハロウィンの前についての実測と計算。棒グラフは日毎のTV露出時間とネット露出の頻度。

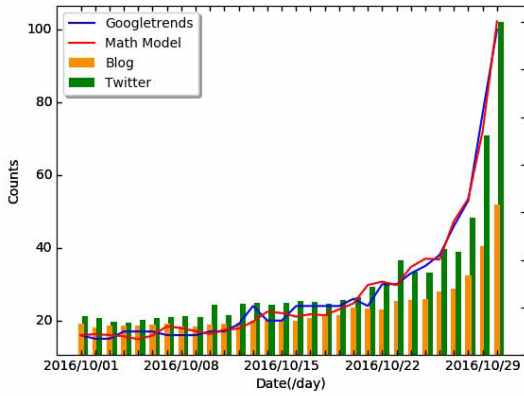


Fig. 2: 2016 年のハロウィンの前についての実測と計算。棒グラフは日毎の Blog と Twitter の投稿数。

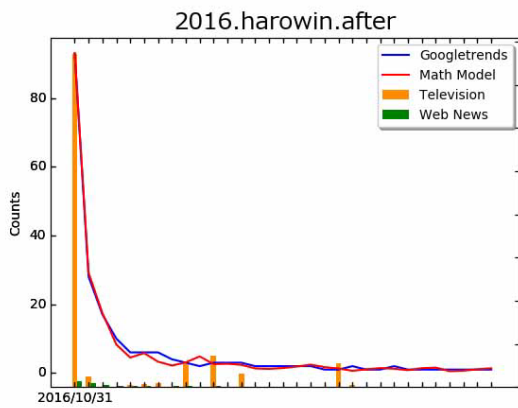


Fig. 3: 2016 年のハロウィンの後についての実測と計算。棒グラフは日毎の TV 露出時間とネット露出の頻度。

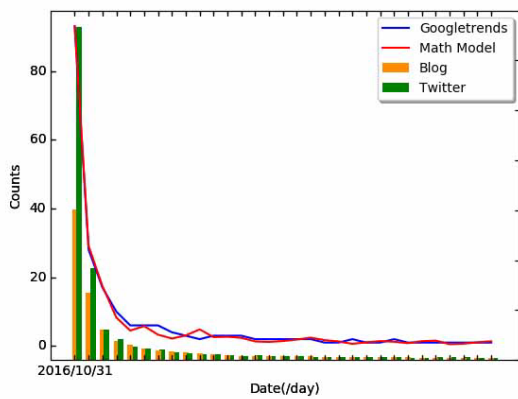


Fig. 4: 2016 年のハロウィンの後についての実測と計算。棒グラフは日毎の Blog と Twitter の投稿数。

次に、実測値と合わせて得られた C_{blog} と $C_{twitter}$ の数値を、Fig.5にクリスマス、Fig.6にハロウィン、Fig.7のバレンタインデー、Fig.8に新年のカウントダウン、Fig.9に恵方巻き、Fig.10に土用の丑の日の鰻について

それぞれ示す。

この結果を見ると、1つの傾向が浮かび上がってくる。クリスマス、ハロウィン、バレンタイン、カウントダウンについては、イベント前は Twitter に影響された検索が多く、イベント後は Blog に影響された検索が強いのである。逆に恵方巻きと土用の丑の日の鰻については、イベント前は Blog に影響された検索が強く、イベント後は Twitter に影響された検索が強い。

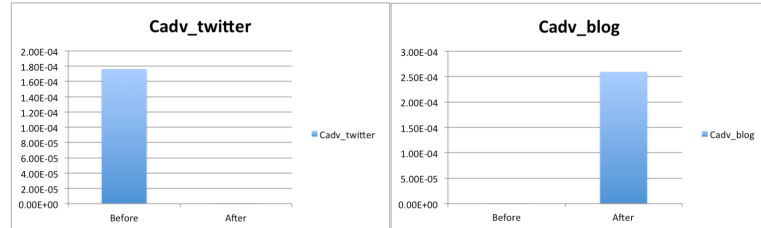


Fig. 5: 2016 年のクリスマスについて、イベント前1ヶ月とイベント後1ヶ月の C_{blog} と $C_{twitter}$ の数値を示す。

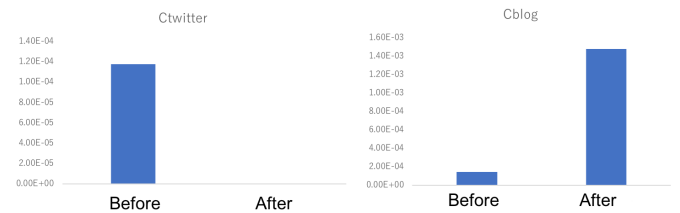


Fig. 6: 2016 年のハロウィンについて、イベント前1ヶ月とイベント後1ヶ月の C_{blog} と $C_{twitter}$ の数値を示す。

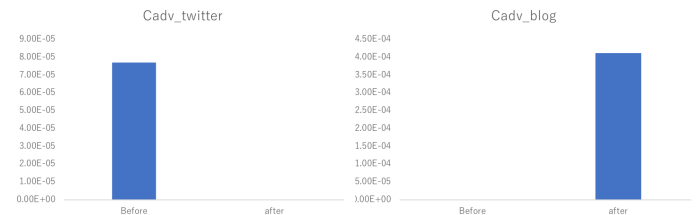


Fig. 7: 2016 年のバレンタインについて、イベント前1ヶ月とイベント後1ヶ月の C_{blog} と $C_{twitter}$ の数値を示す。

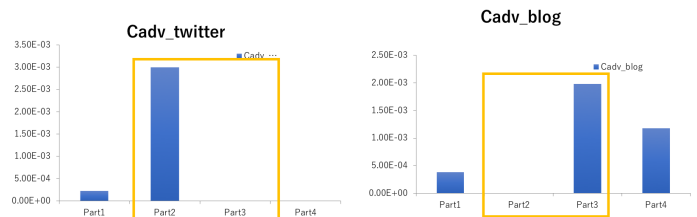


Fig. 8: 2016 年のカウントダウンについて、イベント6日間とイベント後7日間の C_{blog} と $C_{twitter}$ の数値を示す。

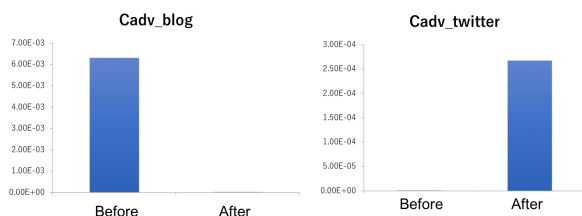


Fig. 9: 2016年の恵方巻きについて、イベント前1ヶ月とイベント後1ヶ月の C_{blog} と $C_{twitter}$ の数値を示す。

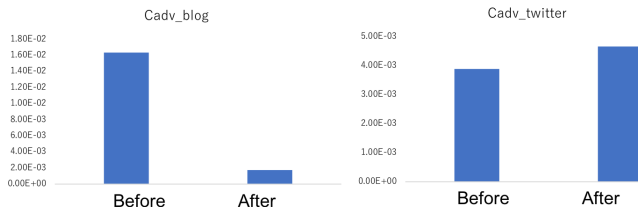


Fig. 10: 2016年の土用の丑の日の鰻について、イベント前1ヶ月とイベント後1ヶ月の C_{blog} と $C_{twitter}$ の数値を示す。

5 考察

計算結果から、クリスマス、ハロウィン、バレンタイン、カウントダウンについては、イベント前はTwitterに影響された検索が多く、イベント後はBlogに影響された検索が強い。また、恵方巻きと土用の丑の日の鰻については、イベント前はBlogに影響された検索が強く、イベント後はTwitterに影響された検索が強い。

食べ物に関係したイベントと、そうでないイベントで傾向が別れている。これはイベントの種類や特徴に依存して傾向がわかれていると思われるが、まずは2つの傾向にきちんと分けられることで、検索行動の数理モデルが人々の検索行動の特徴を捉えているのではないかと考察できる。

次に、なぜ傾向がわかれたかであるが、そのキーはイベント毎の性質に依存すると考えられる。例えばハロウィンの場合、検索する人々はどんな仮装をすべきかに関心があるはずである。その年その年ごとに社会風俗や流行、時事ネタで仮装のネタは変わってくる。また、どこのハロウィンに参加するかの情報も重要である。渋谷のスクランブル交差点付近のハロウィンの馬鹿騒ぎが有名であるが、ハロウィンイベントは東京都内だけでもいろんなお店で企画されており、そのどこに行くかの選択も重要であろう。これらのアップデートな情報を人々はTwitterを頼りに調べていることが、 $C_{twitter}$ の係数がイベント前に多いことで理解できる。

同様に、クリスマス場合はどのレストランでクリスマスディナーを食べるか。あるいはクリスマスパーティに参加するかが考慮すべき重要な事であり、この種の情報をTwitterに頼って集めているのであろう。バレンタインデーについてはどんなチョコレートを買うかの情報が事前に必要である。新年のカウントダウンは、どこの店、どこのイベントでカウントダウンに参加するかが重要であり、これに関する情報をTwitterで集めていると考えられる。つまり、これら4つのイベントについては、事前に演出とかが必要であり、そ

の情報がTwitterに影響されて検索されていると思われる。そして、イベントが終わった後は、イベントのことをどう書かれたかをBlogを見て検索しているのかと思われる。

食べ物である恵方巻きと鰻は、特に食材や料理方法が目新しく変わったりしないので、事前に演出の情報を収集するなどの必要はない。そのため、上記の4イベントと傾向が違ったと考えられる。

このことから、事前にTwitterで情報収集しているイベントには、そこにTwitterを用いたキャンペーンを行うことで、商品に購入を促すなど、マーケティングのチャンスがあることになるとと思われる。

6 結論

検索行動の数理モデルを使って、クリスマス、ハロウィン、バレンタイン、カウントダウン、恵方巻き、土用の丑の日の鰻について調べた。その結果、クリスマス、ハロウィン、バレンタイン、カウントダウンは事前にはTwitterに影響された検索が多く、逆に事後はBlogに影響された検索が多い。食べ物である恵方巻きと土用の丑の日の鰻についてはその傾向が逆であることがわかった。これは、事前に演出や場所設定が必要なイベントについては事前にTwitterに影響されることが推察される。

参考文献

- 1) Akira Ishii, Hisashi Arakaki, Naoya Matsuda, Sanae Umemura, Tamiko Urushidani, Naoya Yamagata and Narihiko Yoshida. "The 'hit' phenomenon: a mathematical model of human dynamics interactions as a stochastic process". *New Journal of Physics* 14(2012) 063018 (22pp)
- 2) Ishii A, "Analysis and Predictions of Social Phenomena via social media using Social Physics method", *A.Ishii, INFORMATION* 20 (2017) 7073/7086
- 3) A Ishii and Y Kawahata, "Sociophysics Analysis of the dynamics of peoples' interests in society" *Frontiers in Physics*, 08 October 2018 — <https://doi.org/10.3389/fphy.2018.00089>
- 4) A Ishii, T Wakabayashi, N Okano and Y Kawahata, "Research on social media writing and search behavior about seasonal topics using sociophysics approach" *Proceedings of WMSCI2018* 21-24
- 5) N Okano, M Higashi, T Wakabayashi, Y Kawahata and A Ishii, "Analysis of seasonal events on social media and internet search using sociophysics model" *Proceedings of ICNAAM2018*
- 6) Y Kawahata, N Okano, M Higashi, T Wakabayashi and A Ishii, "The Influence of Social Media Writing on Online Search Behavior for Seasonal Topics: The Sociophysics Approach" *Proceeding of IEEE Big-Data2018* 4326-4332
- 7) Nozomi Okano, Masaru Higashi and Akira Ishii, "The Influence of Social Media Writing on Online Search Behavior for Seasonal Events: The Sociophysics Approach", *Proceedings of The 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES2018)* 45-49
- 8) 大知正直・長濱憲・榎剛史・森純一郎・坂田一郎『口コミ指数による事例類型化に基づく複数メディアのヒット前の露出を先行指標とした情報拡散過程の分析』*広報研究* 11 (2016) 35/50
- 9) Pendry JB, 1980, "Reliability factors for LEED calculations", *J. Phys.C13* 937/944 (1980)
- 10) 漆谷たみこ『GRPを用いたヒット現象の数理モデルのパラメータ考察』(鳥取大学応用数理工学科 卒業論文) (2010)