

ヒット現象の数理モデルを用いた季節性話題の解析

鳥取大学 群馬大学* 東大将, 川畑泰子*, 岡野のぞみ, 石井晃

1 背景

近年、公衆ネットワークの普及と個人で多様なメディアに接続可能なデバイスが流通し、ソーシャルメディアの利用者も増加する傾向にある。そして、SNS 等を用いて現実世界の人々の行動・思考がオンライン上に投稿され、多くの人と情報を共有することで、互いに影響を及ぼし合っている。つまり、オンライン上を解析することで、現実世界の現象を分析することができることが示唆される研究が多く進んでいる。

オンライン上を分析するため、ツイッター、ブログの書き込み件数、Google の検索の件数、テレビの露出度などのデータを分析することでヒット商品等の要因を見つけ、他の商品等のヒットを予測することが可能であると仮説を立て、鳥取大学の石井晃・デジタルハリウッド大学大学院の吉田就彦らによって提案されたヒット現象の数理モデル [1] を用いた研究が行われてきた。過去の研究には、映画や選挙、音楽配信などについてヒット現象の数理モデルを用いた研究が行われてきた [2, 3]。

今回解析を行う季節性話題について先行研究がされている [4, 5, 6, 7, 8, 9]。季節性話題とは、お花見や花火、クリスマスといった日本全国で認識されている季節に関係する季節性を示唆する行事のことを本研究では指す。季節性の行事に行くにあたって、人々はその行事の情報を求めて Twitter や Blog、TV、ネットから情報を知ることによってスムーズに検索行動に移れる。また、季節性を示唆する行事にあまり関心がなくてもふと行事関連の記事を何かの媒体で見ることによって興味関心が芽生えることも考えられる。

季節性話題は時期が限定されているためにその話題について検索する行動が増えていく状況と、時期がすぎると関心が終息していく状況の測定がしやすい。そこで、毎年訪れる季節性話題をターゲットにすることで、多くのデータを取ることができ、人々の興味・関心の動きが見えてくるのではないかと考えられる。

2 理論

本実験の研究対象である季節性話題は、季節性を示唆する行事の前後で Twitter や Blog に投稿することが予測される。なので、石井・吉田らによって提案されたヒット現象の数理モデルを季節性話題に応用できると考えた [1, 2]。

2.1 検索行動の数理モデル

検索の数理モデルでは、あるトピックについて関心・意欲を駆り立てる要因として以下の 3 つが考えられる。

- 「宣伝広告の影響」
- 「友人からの影響」
- 「街中での噂」

会話やチャットなどによる直接的なやりとりによる影響を「直接コミュニケーション」と呼び、それに対して街中の噂であったり、ネット検索で目に止まった掲示板や Blog 上のやりとりなどに影響されたものを「間接コミュニケーション」と呼ぶ。それらについて興味・意欲の時間変化を追う微分方程式を立てるというという方法で数理モデル化する。ヒット現象の数理モデルによる社会の中の 1 人の人の興味・意欲の方程式は以下のようになる。

$$\frac{dI_i(t)}{dt} = CA(t) + \sum_{j \neq n}^N D_{ij} I_j(t) + \sum_j \sum_k P_{ijk} I_j(t) I_k(t) \quad (1)$$

(1) 式について、右辺第 1 項が広告宣伝によって影響された外力の項、第 2 項が会話やチャットの中での会話から影響され興味を持つ人の項（直接コミュニケーション）、第 3 項が噂や Blog などに影響されて興味を持つ項（間接コミュニケーション）である。

簡単化のために平均場近似を行う。これは、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関心・意欲を $I(t)$ として、以下で定義する。

$$I = \frac{1}{N} \sum_j I_j(t) \quad (2)$$

導出の詳細は [2, 3] を参照されたい。

これを用いると、(1) 式は以下のように簡単化される。

$$\frac{dI_i(t)}{dt} = CA(t) + DI(t) + PI(t)^2 \quad (3)$$

また、本研究では季節性を示唆する行事の後にはもうその行事に対して行動を起こさないと考え、行動する者の行動意欲の減衰のモデルを追加し、

$$\frac{dI_i(t)}{dt} = CA(t) + DI(t) + PI(t)^2 - aI(t) \quad (4)$$

とする。

つまり、季節性を示唆する行事の前のデータ分析では (3) 式を使用し、季節性を示唆する行事の後のデータ分析では (4) 式を使用した。

2.2 検索行動の数理モデル

また、本研究では、文献 [10] で提出された検索行動の数理モデルを用いる。検索行動の数理モデルでは Blog や Twitter での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。

$$\begin{aligned} \frac{dI(t)}{dt} = & c_{TVi} A_{TVi}(t) \\ & + c_{NetNewsi} A_{NetNewsi}(t) \\ & + c_{Blogi} A_{Blogi}(t) \\ & + c_{Twitteri} A_{Twitteri}(t) + DI(t) + PI(t)^2 - aI(t) \end{aligned} \quad (5)$$

このモデルでは人々の興味関心の指数である $I(t)$ を GoogleTrends の関心度とし、従来のヒット現象の数理モデルの広告宣伝の項には、Tv, 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 の影響を追加している。

本研究では、この数理モデルで分析を進める。

2.3 計算精度

本研究では、フィッティングの際に精度を示す指標として R_factor を用いる [11]。本研究では以下の式を用いる。

$$R = \frac{\sum_i (f(i) - g(i))^2}{\sum_i ((f(i))^2 + (g(i))^2)} \quad (6)$$

ここで、本研究では $f(i)$ はある項に対するブログの書き込み件数の値、 $g(i)$ はヒット現象の数理モデルを用いての計算結果の値を用いる。この R_factor の値が小さいほど高い計算精度であると考えることができる。

3 解析方法

3.1 データ取得方法

本研究では、「クチコミ@係長」と「GoogleTrends」というウェブサイトを用いて以下のデータを取得した。

- Twitter の書き込み件数 (1/10 のサンプリング)
- Blog の書き込み件数
- ネットニュースの件数
- テレビの露出件数 (秒数)
- Google での検索件数 (データ取得期間の最大件数を 100 とする規格化)

クチコミ@係長とは、株式会社ホットリンクが運営するソーシャルメディア分析サイトで、取得したいデータのワードを入力し、期間や詳細条件を設定することで日ごとの Blog 件数や Twitter 件数 (1/10 サンプリング) などを取得できる [7]。

Twitter のデータ取得設定では、Twitter のデータ内にリツイートが入るとスパイクが大量に発生してし、ノイズの原因になってしまう。そのため「メンション」「オーガニックツイート」のみにする。Blog のデータ取得設定では、フィルターを「無し」にしてしまうとサイトに 2ch などのリンクを貼ったサイトも含んでしまう。

また、「強」にするとデータが絞られすぎてしまう。よって「弱」に設定する。

ちなみに、「メンション」とは Twitter でユーザーに向けてつぶやいているツイートで、「オーガニック」とは有料ではない通常のツイートのことをである。

GoogleTrends とは、キーワード やトピックの検索回数トレンドを確認できる Web ベースの ツール である [8]。



図 1: クチコミ@係長



図 2: GoogleTrends

3.2 パラメータ

数理モデルで用いるパラメータを紹介する。パラメータとその説明を表 1 に表す。

表 1: パラメータ一覧

D	直接コミュニケーションの強さ
P	間接コミュニケーションの強さ
Cadv-t	テレビの影響の強さ
Cadv-n	ネットニュースから受ける影響の強さ
Cadv-blog	Blog から受ける影響の強さ
Cadv-twitter	Twitter から受ける影響の強さ

直接コミュニケーションは直接友人や知人との会話によって影響される度合いを示し、間接コミュニケーションは噂としてどう伝わっているのかの強さの度合いを示している。外的影響はその話題に対する TV・ネット・Blog・Twitter によって人々が反応した度合いを示している。

3.3 題材の選択

本研究では以下のワードでデータを取り、各行事を計算をした。

- 「エイプリルフール」
- 「クリスマス」
- 「センター試験」
- 「バレンタイン」

上記の行事を選択した理由としては、季節性を示唆する行事の前後での関心・意欲の推移が見やすいと考えたからだ。また、毎年訪れる行事であるため、人々の関心が高く、Blog や Twitter への書き込み数が多いのも選択した理由の 1 つである。

なお、計算期間は before を行事の約 1 ヶ月前からイベント日前日までとし、after を行事の日からその後約 1 ヶ月先とし、過去三年分のデータを使用した。

4 計算結果

4.1 エイプリルフール

図 3 をフィッティング例として示す。また季節性を示唆する行事の前後で比較した各パラメータ 3 年分を図 4、図 5、図 6、図 7、図 8、図 9 に示す。

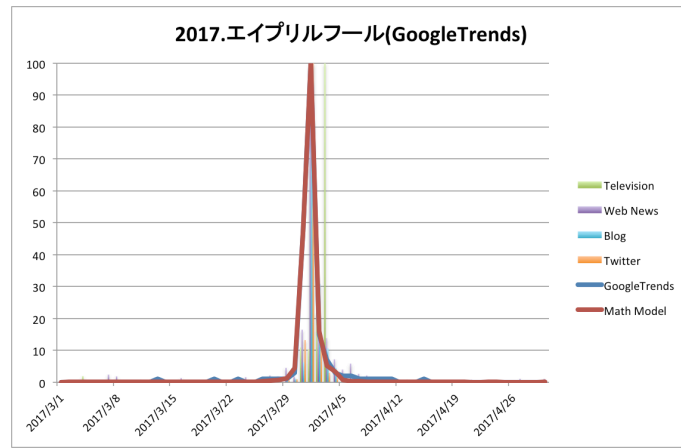


図 3: エイプリルフル.2017 GoogleTrends 計算結果

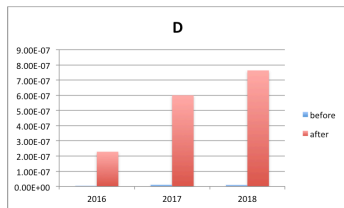


図 4: 直接コミュニケーション

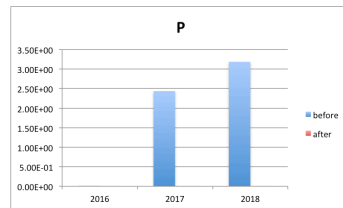


図 5: 間接コミュニケーション

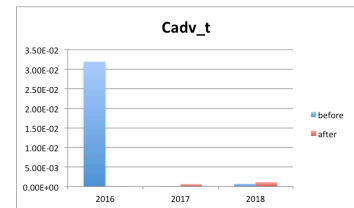


図 6: テレビによる外的影響

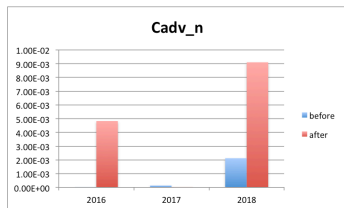


図 7: ネットニュースによる外的影響

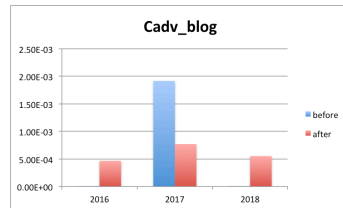


図 8: Blog による外的影響

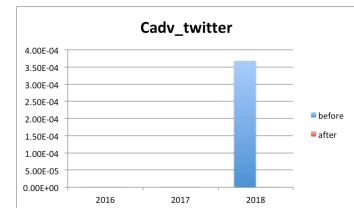


図 9: Twitter による外的影響

4.2 クリスマス

図 10 をフィッティング例として示す。また季節性を示唆する行事の前後で比較した各パラメータ 3 年分を図 11、図 12、図 13、図 14、図 15、図 16 に示す。

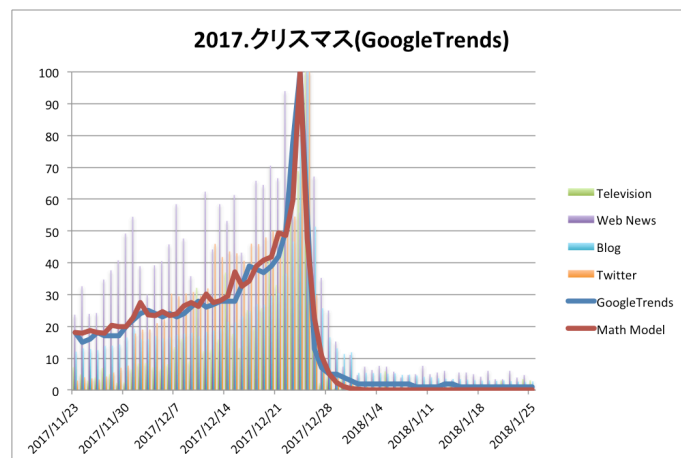


図 10: クリスマス.2017 GoogleTrends 計算結果

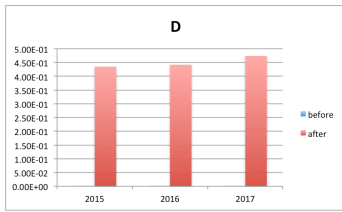


図 11: 直接コミュニケーション

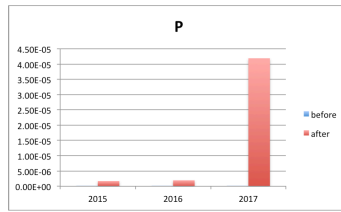


図 12: 間接コミュニケーション

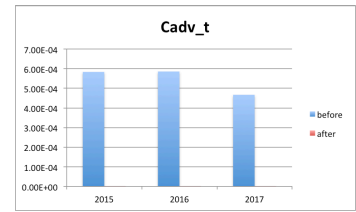


図 13: テレビによる外的影響

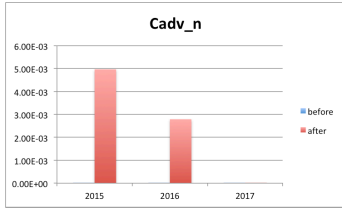


図 14: ネットニュースによる外的影響

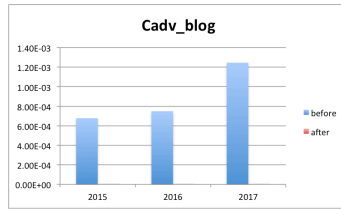


図 15: Blog による外的影響

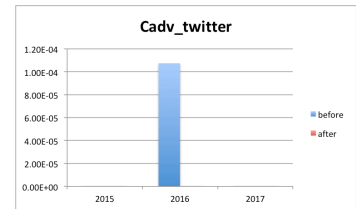


図 16: Twitter による外的影響

4.3 センター試験

図 17 をフィッティング例として示す。また季節性を示唆する行事の前後で比較した各パラメータ 3 年分を図 18、図 19、図 20、図 21、図 22、図 23 に示す。

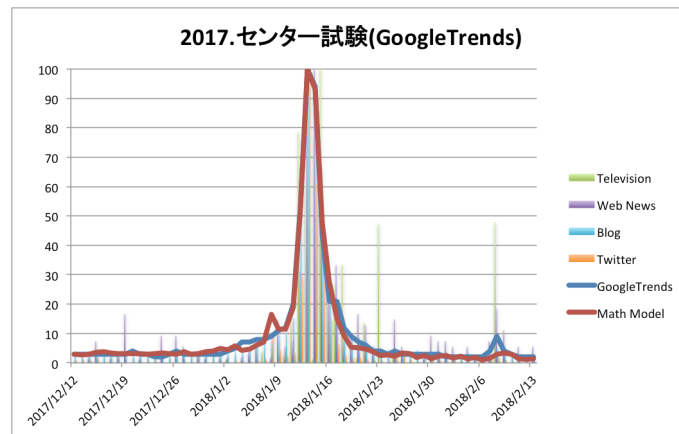


図 17: センター試験.2017 GoogleTrends 計算結果

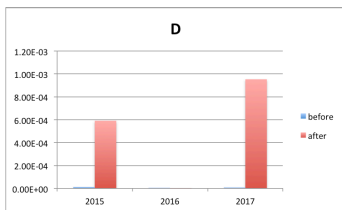


図 18: 直接コミュニケーション

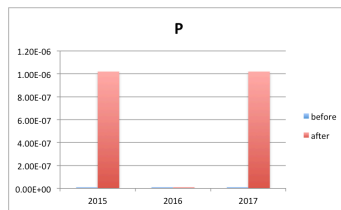


図 19: 間接コミュニケーション

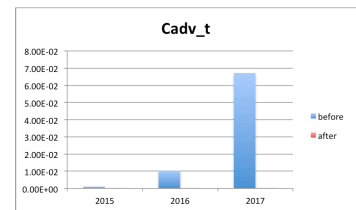


図 20: テレビによる外的影響

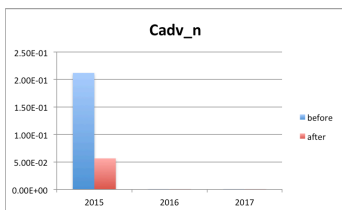


図 21: ネットニュースによる外的影響

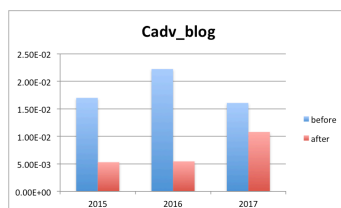


図 22: Blog による外的影響

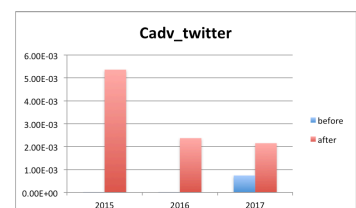


図 23: Twitter による外的影響

4.4 ホワイトデー

図 24 をフィッティング例として示す。また季節性を示唆する行事の前後で比較した各パラメータ 3 年分を図 25、図 26、図 27、図 28、図 29、図 30 に示す。

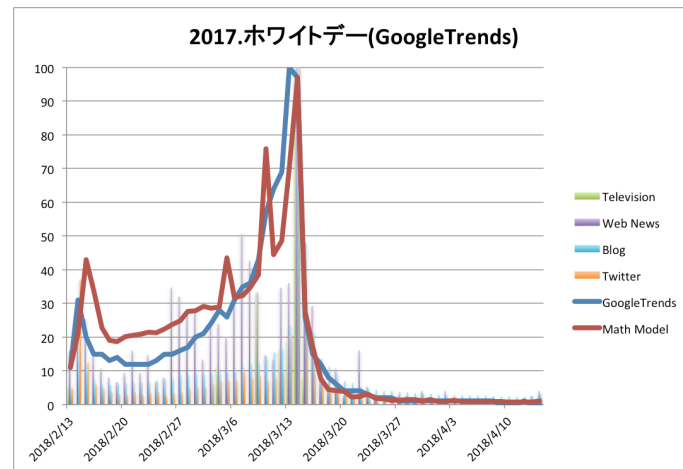


図 24: ホワイトデー.2017 GoogleTrends 計算結果

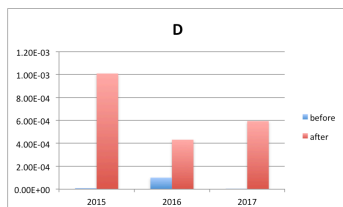


図 25: 直接コミュニケーション

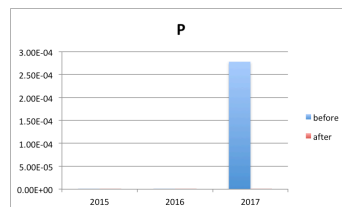


図 26: 間接コミュニケーション

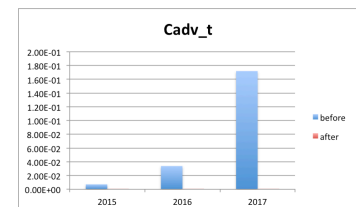


図 27: テレビによる外的影響

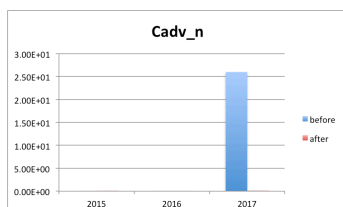


図 28: ネットニュースによる外的影響

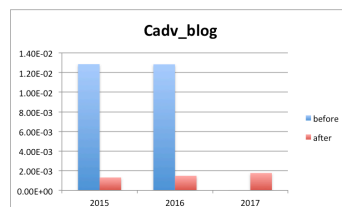


図 29: Blog による外的影響

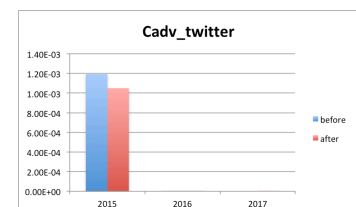


図 30: Twitter による外的影響

5 考察

エイプリルフル

直接コミュニケーション D は行事の日以降に高い値が観測された。このことから、行事の後に話題になっている。これは、エイプリルフルが終わったあと、どのような面白いネタがあったか日常会話で出てきていて、気になって検索行動に移った人がいることを示唆している。間接コミュニケーション P は D とは反対に行事の前に高い値が観測された。これは、毎年様々な話題が出てくるので、今年はどういう話題が出てくるか事前にちょっとした会話に出ていると推察される。

TV、ネットによる外的影響は行事の日から高い値が観測された。これは、テレビとネットの露出は行事の日以降に多くなっていることから、行事の後の特集番組や掲示板を見たことで関心が集まったのではないかと考えられる。また、Blog による外的影響も行事の日から高い値を示した。おそらく、イベント後のブログの投稿から興味関心を得ていると言えるだろう。また、Twitter による影響で行事の前に高い値を示していることもそういった行事の前のツイートによるものだと示唆できる。

クリスマス

直接コミュニケーション D は行事の日以降に高い値を示した。これは、今年どういったことが行われたのが気になって検索したと考えられる。また、元データを見ると、Twitter・Blog は行事の前に書き込み件数が多いことがわかる。そういった事前の Twitter や Blog の発信によって検索行動を助長しているのため、Bog や Twitter からの影響が行事の前に出ていると推察される。間接コミュニケーション P は行事の後から高い値が観測された。これは、クリスマスの後ににどういったことをしたか少し会話が出たと考える。

TV による影響は行事の前に高い値を示しているが、ネットによる影響は行事の後から高い値を示している。TV ではクリスマスに向けておすすめスポットやおすすめスイーツなどの特番が検索行動に結びつく人々の興味関心を高めているのだろう。ネットでは、行事の後から様々なイベントを取り上げるためだと考えられる。

センター試験

直接コミュニケーション D は行事の日から高い値を示しているのは、センター試験というものは当日までどう行っただけの問題が出るかわからないため、試験終了後に大きな話題になっているのだと考えられる。また、間接コミュニケーションは GoogleTrends は行事の日から高い値を示している。これは、行事が終わったあとどういったものだったのか気になっていることがわかる。

Twitter による影響が行事の後高い値を示しているのは、行事が終了した後どういった問題が出たかというツイートによって興味関心を高められたと推察される。また、Blog による影響が行事の前に高い値を示すのは、センター試験前に今年の試験はどういった問題が出るかというネット内でのニュースによる影響だと考えられる。

TV による影響が行事の日以降に高い値を示しているのは、TV によるセンター試験の情報からセンター試験について少しずつ意識し始めていることを示唆している。ネットニュースによる影響は、行事の後に高い値を示しているこのことから、検索意欲行動に走る人々は行事の後どうだったか気になっていることが考えられる。

ホワイトデー

直接コミュニケーション D は行事の日から高い値を示していることから、行事の後に話題となって検索していることがわかる。また、間接コミュニケーション P は行事の日から高い値が観測された。これは、ホワイトデー以降に会話になっていることがわかる。

TV による影響が行事の前に高い値を示したのは、1ヶ月前にバレンタインがあり、バレンタインが終わるとすぐハロウィンに向けての TV 番組があるためだと考えられる。実際に TV の露出度は行事の前に多い。また、ネットによる影響は行事の日から高い値を示した。これは、行事の後にどういったものが今年流行したかといったニュースを見たことで影響を受けているのではないかと考えられる。

また、Twitter と Blog による影響は行事の前に高い値が観測された。これはバレンタインが終わると同時に投稿されるものによる影響で検索行動に結びつく人々の興味関心を高めているのではと考えられる。

6 まとめ

本研究では、従来のヒット現象ではなく、GoogleTrends を用いた検索行動の数理モデルを使い、行事の前後で減衰項を使い分けすることで季節性に関する人々の検索行動および潜在層の興味関心に対する各メディアがどのように影響を受けているか見ることができた。以下で詳しく記述していく。

直接コミュニケーション D、間接コミュニケーション P の計算結果を見ていく。「エイプリルフール」「ホワイトデー」は D が行事の後に高く、P は行事の前に高い値が観測されている。このことからサプライズ性があることが言える。「クリスマス」は D が行事の前に高く、P は行事の後に高い値が観測されている。このことから、行事の前から注目されており、行事が終わった後には何かしらの噂が拡散されていることがわかる。つまり、「クリスマス」を通して様々な情報が拡散されていることわかる。

TV、ネットの影響についての結果を見ていく。「クリスマス」「バレンタイン」「ホワイトデー」は TV、ネットニュースともに行事の前に露出が多いが、ネットニュースによる影響が行事の日以降に受けているのは、今年どういったものが流行ったか気になっていることを示している。また、行事の日以降に TV によって刺激され検索行動に動いていることが考えられる。

「センター試験」は行事の前後両方に TV とネットニュースの露出が散らばっている。これはイベント性質上当然だと考えられる。そして、TV とネットニュース両方とも試験前に影響を受けている。これは、TV やネットが試験が近づいてくるにつれて少しずつ話題になることで、その番組や記事に影響を受けていると考えられる。「エイプリルフール」は TV、ネットニュース両方とも行事の日以降に露出が集中しており、検索行動を起こす人々はそういった記事に影響を与えられている。

Blog による影響は、「エイプリル」は行事の日以降に受けている。これは、行事の日の後に投稿されたものを見て検索していることが推測される。対して「クリスマス」「センター試験」「ホワイトデー」は行事の前に受けている。これは、行事の前に事前にその行事の情報を Blog で見ることで検索行動の意欲に結びついていると考えられる。

Twitter による影響は、「センター試験」のみ行事の日以降に受けている。これは上記でも述べているようにセンター試験は機密性が高いため、行事が終わると一斉に Twitter で問題についての投稿がされる。そういった投稿を見て検索しているのだと考えられる。その他のイベントは行事より前に影響を受けている。これは、事前に行事に関する情報を Twitter で集めていると推察される。

計算結果と考察から、季節性のある話題はその行事の前に事前に情報を得ようとする動きが見られた。また、行事が終わった後にも興味関心が短時間ではあるが向けられていることがわかった。

参考文献

- [1] 石井晃、吉田就彦「ヒット現象の数理モデル」鳥取大学工学部研究報告第 36 号,p71-80,2005
- [2] 吉田就彦・石井晃・新垣久史「大ヒットの方程式 ソーシャルメディアのクチコミ効果を数式化する」ディスカヴァー・トゥエンティワン社,2010
- [3] A.Isii, H.Arakawaki, N.Matsuda, S.Umemura, T.Urushidani, N.Yamagata and N.Yoshda : "The 'hit' phenomenon :a mathematical model of human dynamic Interactionsas stochastic processs" New Journal of Physics 14(2012) 063018 (2013)
- [4] 石井晃、少林俊道、岡野のぞみ、川畑泰子：「ソーシャルメディア上の季節性のある書込の解析」 第二回計算社会科学ワークショップ，2017.
- [5] 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
- [6] 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 印刷中
- [7] 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 BigData2018 4326-4332
- [8] 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
- [9] 岡野のぞみ、東大将、石井晃「検索行動の数理モデルを用いた季節性イベントの解析」計測自動制御学会社会システム部会第 18 回社会システム部会研究会会議録 (2019)
- [10] 石井晃、芦田昇、川畑泰子「インターネット上の検索行動の数理モデル」第 3 2 回人工知能学会全国大会論文集 1E2-05 2018 4 ページ
- [11] 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
- [12] 「クチコミ@係長」 < <https://kakaricho.jp/login> >
- [13] 「GoogleTrends」 < <https://trends.google.co.jp/trends/?geo=JP> >