

Twitter を用いた Web 上の投稿にもとづくヒット現象解析によるパフォーマーの TV 主演時の分析の手法の検討

Study of analysis of technique during the performer of TV starring by hit phenomenon analysis based on the post on the Web using the Twitter

川畑 泰子† 源田 悦夫‡ 石井 晃‡
Yasuko Kawahata Etsuo Genda Akira Ishii

1. 序

社会には音楽や演劇, 美術, 写真・工芸, 文学といった人間の表現活動が多く存在し, それらの表現をする人々が存在する. 表現をする者すべてを本論ではパフォーマーと定義すると, パフォーマー自身を含めた種々のプロモーション活動に関して, 直接的あるいは間接的に提供される情報として, テレビやラジオ新聞, 雑誌などのマスメディアやチラシや街頭告知などの小規模な活動によるばかりでなく, 公衆ネットワークを介した従来のプロモーション活動にはなかった即時性と双方向性を持つ SNS(ソーシャルメディア) による情報も重要なプロモーションの要素になってきている. 最近では Web 上での広告や, SNS を活用しパフォーマー自身やプロモーターの興行側 から, 公演や出版などの情報を, 直接ファンや一般に Web 等で発信することが増えている. さらにファンとファン, ファンと一般など様々な経路によって, これらの情報が提供されていることが増えていることも Web 上から観測ができるようになった. さらにパフォーマーが一般人に情報を発信する機会も盛んとなっている. Web 上での情報をカウントしたり, 内容分析したりすることで定量的なデータを観測することができるようになった. 実際の劇団のプロモーションを行う活動の中で分析を行った結果, 実際にパフォーマーに関する情報を発信するマネジメントを行う層とその情報を積極的に受け取るファン層, またこれからファンになる可能性がある一般人の層が芸術文化の存在する社会には存在すると仮定できた. パフォーマーとともに主に情報を発信するプロモーターとパフォーマーの情報を積極的に送受信するファンの間での情報のやり取りが Web 上でのパフォーマーに関する情報を増やし, 活動の方針のヒントにできた. また舞台公演やコンサート興行, テレビ出演などあらゆるメディアを通じた活動がパフォーマーに関する情報を Web 上で増加させる傾向もあるとわかった. 地方の劇団でもこのような結果を得ることができることに着眼し, さらに Web 上のパフォーマーに関する情報を扱ってある特定のジャンルで活躍するパフォーマーでも本手法の適応を行うことで, Web 上の情報を活用したパフォーマーの今後の活動の方針に役立つ知見を得るための手法を提案したいと考えた. そこで, 本論では TV とネット番組にあるパフォーマーが出演した際の分析の検討を考えた. 本論では, Twitter のデータの取得先として, Topsy を使用した. Topsy では, 2006 年 11 月より Twitter へ書き込まれた公開データを収集し, データベース化をして, 検索ツールとして顧客に提供している. 本研究では Topsy を使用して, 検索単語を含む日ごとのブログ記事を「Tweet」と呼び,

その指定する期間の出現頻度の時系列を CSV データの出力または, 目視で確認を行いながら抽出をした. 本研究で主に扱うデータで, 検索語を含む Tweet(Retweet 数と Reply 数を含む)数の時間刻みの出現頻度時系列の検索単語を含む書き込み数は, データベースに保存されている記事の本文中か記事の表題に検索単語を一度でも含むかどうかで決定している. そのため, 同一記事に複数回, 検索単語を含んでも, 1 と数える. 他方, 同一のアカウントが, 同日に, 検索単語を含む書き込みを複数投稿した場合は, 別々に検索対象となるため, 複数に数え上げる. 半角文字で「a」と検索しても全角文字で「A」としても同じ結果となる.

2. ヒット現象に関する考え方

一般に映画やレコードなどのヒットとは発表したものが多くの人の支持を受けることである. 図 1. 1 は情報の過渡現象を示した例であり, 図で示しているイベント 1~3 のようにコンサートやテレビ放送に関する情報が人による情報伝達やマスメディアなどを媒体として, 情報が急速に広まる過程を, Web 上の投稿数から捉えその数量の増加や減衰を経て, 均衡に至る過渡現象について示したものである. ここでは Twitter における投稿数(Tweet, Retweet, Reply)の数量に関して取り上げる.



図 1. 1 Web 上の投稿数の過渡現象と均衡 (本研究では Twitter のデータを利用するため, Tweet (投稿) 数と表記している.)

また, 従来の研究成果では映画の公開初日を起点にしてブログの書き込み件数と興行収入は類似してピークになり, その後急速に減衰している事例[1-3]のように劇団が舞台公演を開催した際のオフィシャル Web ページの閲覧数も

† 日本学術振興会 特別研究員 PD, 東京大学 大学院 情報理工学系研究科数理情報学専攻

‡ 九州大学 大学院 芸術工学府

‡ 鳥取大学 大学院 工学研究科

映画の事例の際とは異なるが、舞台公演時にピークが来た後に減衰しているといった事例もあり、これらの Web 上における投稿数や閲覧数の過渡現象と平衡状態をヒット現象として捉え分析が行われてきた。本研究では、プロモーターからの多量な情報の広報活動による社会現象にまでつながる規模のヒット現象だけでなく、特に Web 上のパフォーマーに関する情報に注目することで、パフォーマーへの社会の関心の度合いを計測できると考えた。パフォーマーに関する情報を Web 上でヒットさせるにはファン集団の中にいる周辺の人々の関心を動かすことができるかが重要となる。ヒットが起こる過程でどのような要素が人々の関心にアプローチしたかを考えることが、ヒット現象を解析する上で重要である。そこで、Tweet 数を抽出する調査期間内に行われる、コンサートやテレビ放映によるパフォーマーに関する話題(コンサートの名前、パフォーマーの名前、コンサートの内容の名称、テレビのタイトル)等が、ファンやファンの周辺からの人々の話題になることによる増減に注目した。ファンの支持を得ることにより、Web 上でのパフォーマーを指すキーワードが増え、ピークを過ぎると急速に減衰していく現象が多くの事例においても見られた。例えば、アニメの番組タイトルに関するヒット現象の分析結果[4]、日本のガールズグループの「総選挙」に関する分析結果[5]、オンライン音楽の楽曲タイトルのヒット現象に関する分析結果[6]やテレビドラマのタイトルに関するヒット現象の分析結果[7]が挙げられる。

3.1 数理モデルの基本方程式

パフォーマーに関する情報に関して触れる層の人々が N 人であるとして、パフォーマーの情報に触れる割合を $P_{contents}$ 、さらにパフォーマーに関する情報に触れ、行動の判断を行い、ネット上で書き込みやいいね!の行動をする人の割合を T_{action} とし、ファンの一人がパフォーマーのコンテンツに関して 1 つの関心からわき起こる行動をしたと考えると、パフォーマーが話題となる総数は、 $P_{contents}T_{action}N$

(3. 1)

である。時間の単位が 1 日であるとすれば、 Δt の時間でパフォーマーに関して影響を受けて行動が起きる回数、 ΔI は、

$$\Delta I = P_{contents}T_{action}N\Delta t$$

(3. 2)

となる。これから、仮定する状況を微分として扱うことができれば、このパフォーマーの情報を含む Tweet 数の上昇は単純な微分方程式とできる。

$$\frac{dI}{dt} = P_{contents}T_{action}N$$

(3. 3)

となる。ここで注意すべきはパフォーマーの総話題数の時間変化、つまりパフォーマーに対するファンの人々の関心の総数を考えている。

次に、一人ひとりのパフォーマーに関する情報に関して触れる層について取り上げると、パフォーマーに関する情報に関して触れる層の人々が N 人が存在すると考えて、 p 番目の人の関心の度合いを $I_p(t)$ とすると式(3. 3)は、

$$\frac{dI_p(t)}{dt} = P_{contents}T_{action} \quad (3. 4)$$

となる。なお、この段階では一人ひとりの個性は無視している。

3.2. Twitter 上における直接情報と間接情報

本論ではパフォーマーに関する情報を含む Tweet 上の情報の性質を二つに分けて考察することでヒット現象における人々の関心の動きを解析する。Twitter 上でパフォーマーに関する情報を含む話題の一つ目の動きはパフォーマーに対して関心をもっている人と直接的に Tweet をすることによって情報を発信するというような動きである。

具体的に、このような動きの話題の上り方を本研究では「直接情報」とする。実際に通常の Tweet をすることを本論では指す。ロコミの 2 つ目の動きはパフォーマーの情報を知る人同士が会話している内容を見たり聞いたりして間接的に情報を発信するというような動きであり、このような形態のロコミを本研究では「間接情報」とする。また、ここでは Twitter 上での情報に関して考えるので「間接情報」は Twitter 上で見たり聞いたりした広告や人の話題、パフォーマーに関する情報の公開や発表の前に配布される広告や噂から新しい情報を得て、Retweet や引用 Retweet, Reply を本論では指す。

直接情報とは人から人への直接的なコンサートの閲覧や情報の伝達から受ける影響、 p との直接情報によって影響を受ける t のパフォーマーの関心の度合いのことである。Twitter 上での直接情報については、パフォーマーに関する情報に直接触れて Tweet をした人からの情報は、それまでにパフォーマーを話題にした人の人数に比例するので、情報を受ける t 番目の人の関心の度合いを $I_t(t)$ と定義できる。

そして p 番目の人が t 番目の人との直接情報により、情報を得て更に情報を拡散する確率を D_{pt} と定義すると、情報を得た p 番目の人からの情報により t 番目の人が更に話題にする(通常の Tweet をする)確率は、

$$\sum_{t \neq p}^N D_{pt} I_t(t) \quad (3. 5)$$

と表される。

一方、間接情報とは第三者同士の会話を耳にしたり、パフォーマーに関連する情報などから間接的に受ける影響、 p と t の会話をたまたま耳にしたときに i のパフォーマーに対する関心の度合いが受ける影響のことである。

Twitter 上で考える際は、 p と t 、 u は第三者とする。この場合における間接情報では、パフォーマーに関して Tweet をみてさらに Retweet や Reply をした人同士の情報は Retweet や Reply をした人で作る対の数に比例するので、情報を受ける t 番目の人の関心の度合いをそれぞれ、 $I_t(t)$ と $I_u(t)$ と定義できる。

話題にした人同士の情報交換は、Twitter や Facebook、掲示板の場合があると考える。つまり、情報の総数が間

題となる。話題にした同士 (p 番目の人と t 番目の人) の情報の影響から u 番目の人が更に話題にする (Retweet, 引用 Retweet, Reply をする) 確率は,

$$\sum_t \sum_u P_{ptu} I_t(t) I_u(t) \quad (3.6)$$

と表される。ここで t, u は p を含まない。係数 P_{ptu} はパフォーマーに対するキーワード数が増える確率であり、話題にした人の内容が良ければ、この係数は大きな値となる。逆に話題にした人の内容が悪いと、この係数は負になると仮定する。

3.3. マスメディアの広告力の検討

本研究で扱う数理モデルの基本方程式は

$$\frac{dl_p(t)}{dt} = CadVA(t) + \sum_{t \neq p}^N D_{pt} I_t(t) + \sum_t \sum_u P_{ptu} I_t(t) I_u(t) \quad (3.7)$$

と表せる。パフォーマーに関する関心の度合いが盛り上がった際の個人の関心の度合いの時間変化を表す基本方程式は、(3.5), (3.6) で定義した情報の項と、式(3.4)で定義した項とコンサートやテレビ放送等のメディアによる広告の効果をもつにまとめて $A(t)$ として表した項を足し合わせる。

$A(t)$ は入力をするコンサートやニュース、テレビの件数を入力する値のことを指す。本論では、主に Twitter の書き込みをパフォーマーにおける関心とみなし分析を行う。

また、本研究では複数の広告メディアからの影響を受けることを考慮するため

$$A(t) = \sum_{\xi} Cadv_{\xi} A_{\xi}(t) \quad (3.7.1)$$

と拡張をする。これで各 $A_{\xi}(t)$ に TV 露出やネットニュース、コンサート出演の影響を別々に入力し、乱数による最適化で各 $Cadv_{\xi}$ 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。

ここで(3.7)より、外力 $A(t)$ は過去の宣伝やパフォーマー自身の影響力などもある程度あると仮定して次のように定義をする。

$$A(t) = \int_{-\infty}^t a(t) e^{at} dt \quad (3.7.2)$$

(3.7.1) とし、積分公式からその最初の 3 項は以下のようになる。

$$A(t) = \varepsilon(t)[a(t) + e^{-a} adv(t-1) + e^{-2a} adv(t-2)] \quad (3.7.3)$$

$$\varepsilon(t) = Cadv \begin{cases} e^{before|t-t_{open}|} & (\text{情報開示前}) \\ e^{-after|t-t_{open}|} & (\text{情報開示後}) \end{cases} \quad (3.7.4)$$

実際の計算では、 $a(t)$ にコンサート、ニュース、テレビ放送、メディア露出の件数を入力する。そして、 $Cadv$ はプロモーターのパフォーマーに対する Web 上の情報の広告効果の強さのパラメータであり、次のように設定した。
 $Cadv$ パフォーマーの活動における Web 上の関心 (コンサート、イベント後の影響)

$Cadv$ はコンサート、ニュース、テレビの報道数 $a(t)$ が多かった場合の関心は高くなると仮定する。従来の研究においては、テレビの場合は視聴率、広告出稿費、メディアへの露出した実際の秒数などの値が入力され、コンサートにおいては、観客動員数やチケットの売り上げなど P 値が入力として代入される [1-7]。

実際のプロモーション活動への応用や分析を行う際は、もちろん取り扱うパフォーマーに関する具体的な数量を取り扱うことが好ましいことを前提とする。

しかし本研究の意図として、パフォーマーの 1 つの活動が Web 上に与える影響を考察できるかを試行するためにここでは、コンサート、ニュース、イベント、テレビ放映の件数をすべて重みつけせずに 1 件でカウントを行う。

また本論で $a(t)$ への入力に使用するコンサートの件数、テレビ放送の件数、ニュースの件数に関しては、ケーススタディの対象のパフォーマーのオフィシャルページから分析時期の範囲内のデータを取得して行っている。4 章のケーススタディ 1 に関しては参考文献 [79-82] から取得を行っており、ケーススタディ 2 に関しては参考文献 [83-85] からの取得を行っている。5 章に関しては「C」のオフィシャルサイトを参照されたい。

以上が、本論文で用いるヒット現象の数理モデルの基本となる方程式であり、ファンにおける 1 人の人の、特定のパフォーマーに対する興味度の度合いの時間的な変動を記述する。これをパフォーマーのファン全体の人々で平均すると以下のような式になる。

$$\frac{dI(t)}{dt} = \sum_{\xi} Cadv_{\xi} A_{\xi}(t) + DI(t) + PI^2(t) \quad (3.8)$$

平均の操作の具体的式展開とフィッティングに関しては論文 [3] を参照されたい。

本研究では Twitter 上でのパフォーマーの活動における広告の効果の定量的に分析するにあたって、書き込み数や閲覧数をパフォーマーに対する関心と考えてフィッティングを行い、結果からパフォーマーの話題の急上昇の要因や傾向を探り、実際のパフォーマーの宣伝のため

の戦略やパフォーマーに興味がある人々の動きに関して分析を行った。

4.1. ケーススタディの目的

本章では、実際に観測を行ったパフォーマーの活動の内容を理解していることを加え、さらに、パフォーマーの活動においてどの要素が重きにおかれるべきか、テレビ出演があった時はどのキーワードが実際にファン層の中でヒットになっているかファンの関心(ヒットするパフォーマーの関連キーワード数)はどこに向かっているかわかるだろうか分析が必要であると。本研究の手法を日本国内の実際に活動しているパフォーマーに対しても適応可能であるかのケーススタディとしての知見を得るために主に音楽コンサートを盛んに行ってきたパフォーマーとして「C」を本研究の調査対象とした。選定理由は、本研究の意図としてパフォーマーの活動に伴って Web 上での情報を取得する上で、一度も活動休止や解散を 20 年以上行っておらず、継続して活動を続けている。そのため、Web 上でのパフォーマーの情報を分析する上でそれまでの過去のパフォーマーの影響力を正確に数理モデルに反映して分析ができると考えた。

ヒット現象の数理モデルで Twitter の観測ができるパフォーマーに関する情報の分析として、ケーススタディ 2 テレビ放映時における関心の分析を行うことができる対象として 201E 年 F 月の「C」に関する調査のために Topsy を利用し、日本語圏内の実測の Tweet 数の取得とオフィシャルサイトによる情報(ニュース、コンサート、雑誌への露出件数)を取得して分析を行った。日本語圏内に限定した理由としては、パフォーマーの名前は英語表記であるが、英語圏内の Tweet の情報を取得することで、国外からのパフォーマーのキーワードを含む Tweet 数やその他のノイズにあたる Tweet 数までいれてしまうことが問題となる。さらに、今回の分析の事例は国外での活動を範疇にいれていないため、日本語圏内のみとした。本分析で使用したデータはリツイートやリプライの数は流動的であるため 2014 年 10 月 11 日の時点までに観測したデータを元とした。

4.2. 分析する活動とキーワードの選定

まず、TV 出演時のパフォーマーに関する分析を行うパフォーマーの活動とキーワードの選別をして分析を行うことにした。

4.2.1. 形態素解析とは

形態素とは、言語学で使われる専門用語で「意味の最小の単位」とし、文の形態素の単位に分割することを形態素解析という。形態素解析の象徴として、文を形態素の単位に分割すると同時に、各形態素の品詞を特定することができる。

ここでは、ケーススタディ 1 での結果を踏まえ、調査期間内の「C」を含む Tweet の本文データを取得し、統計解析の代表的なツールである「R」と「RMecab」、さらに名詞の頻出頻度や関係性を明らかにするために「KHCoder」を利用した。

4.2.2. KHCoder とは

KHCoder とは、テキストマイニングのためのフリーウェアである。樋口らが、新聞記事・インタビュー記録など、社会調査によって得られる様々なテキスト型データを計量的に分析するために制作されたものである。

4.2.3 解析

ここでは、調査期間内のパフォーマー名「C」を含む Tweet の本文データを元に、形態素解析を行った。

そこで、パフォーマーが TV に出演した時期の・201E 年 F/1-末日の「C」を含む Tweet の本文データ抽出条件としては、以下のルールを設定した。

- 1、抽出品詞は名詞のみ
- 2、上位 10 位まで
- 3、非自立、数、接尾、自立を除く

4.2.4 分析すべきキーワードの検討

インターネット上で公開されているテキスト分析用フリーウェアの KHCoder をダウンロードして、分析ツールとして使用した。KHCoder のアプリケーションとあわせて提供されているマニュアル及びチュートリアルに示された手順に則り、分析用データの準備から分析までを実施した。以下に基本的な流れ(1)を示す。

(1)・201E 年 F/1-末日の「C」を含む Tweet の本文データ

上記のデータを KHCoder が扱えるようにテキストデータにそれぞれ入力をした。前処理として、RT (Retweet を示す記号) とパフォーマー本人の Tweet とアカウント名を除外して取り込みを行った。

なお、電子メディアを用いた文章、日記の特徴といえる絵文字と記号については、KHCoder の分析対象外であり、またその絵文字がどのような感情表現で用いられているかを類推することは困難であったため、感情表現を表す単語をあてはめずに絵文字に置換した。このほか、前処理結果を参考に KHCoder が適切な単語抽出単位を判断することを支援する目的で「てにをは」の追加、計算記号の文字への置き換え、数値単位の追加をあわせて実施した。この前処理と修正のプロセスを 3 回実施して、最終的な分析データとして確定した。

KHCoder による上位 10 語として抽出された言葉を表 4.1 に示す。

そこでテレビ出演を主に行っていた F 月「C」を含む Tweet 数の本文データを取得して最も頻度が高かった名詞上位 10 位までを表 4.1 に取り上げた。

ランキング	F月のキーワード(名詞)
1	「E」
2	「C」を含むキーワード
3	「C」を含むキーワード
4	決定
5	#番組名「D」
6	リリース
7	新曲
8	ツアー
9	解禁
10	公開

表 4.1
調査期間内の関心のパラメータ表
F月の「C」を含む Tweet 数の本文データからの
キーワード(名詞)ランキング

キーワードと表 4.1, そして「C」のオフィシャルサイトを見る限り特に Twitter の話題数がヒットする傾向が強かった時期である①テレビ番組の放映の時期の分析を進めることにした。

テレビ番組の放映の時期のキーワードとしては放送期間中, 表 4.1 でも高い頻度で出現していた番組タイトルである番組名「D」を対象とした。

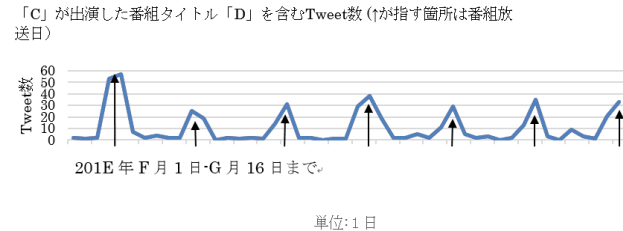


図 4.1
番組名「D」を含む Tweet 数の推移。
図 4.1 と「C」のオフィシャルサイトを比較してもわかるようにテレビ番組の放送日に番組名「D」を含む Tweet 数はヒット現象を起こしていることから, さらに実際の放送がされた前後の時間の番組に対する関心を測ることで, 放送されたメディア(ローカルテレビ, インターネットのテレビ)の放送された際のそれぞれの関心の差も導き出せるだろうと仮定し, 分析をした。

4.3. テレビ番組出演における関心

本分析では, 調査期間内(201E年F月)のテレビ放映時間前後 8 時間の番組タイトルである「D」を含む Tweet 数を取得した。表 4.2 からわかるように放映時期の「C」を含む Tweet 数の中にキーワード「D」やテレビ放送の内容に関する Tweet が多かったことに着目した。さらに, 表 5.6 からわかるようにテレビ放映時間に話題がピークに来ていることに着目し, テレビ番組自体の広告効果を数理モデルで分析をした。

時間	201E/F/AE ローカルテレビ	201E/F/AG インターネットテレビ
18:00:00	4	10
20:00:00	12	32
22:00:00	4	14
0:00:00	24	14
2:00:00	70	5

表 4.2
調査期間内のテレビ放映時間前後 1 2 時間の番組タイトル「D」を含む Tweet 数

表 4.2 を見る限り, 番組タイトルの「D」を含む Tweet 数は放映のタイミングに上昇する。つまり, ある一定の数のファンが反応をして, Tweet 数が上がることがわかった。そこでローカルテレビの放映時間(0:00:00~2:00:00)とインターネットテレビの放映時間(20:00:00~22:00:00)タイミングを 1 として計算を行った。

ここでの, 最終的に数理モデルで算出されるパラメータに関しては以下の表 4.3 のように設定をして, 解析を行った。

Cadv_TV	広告効果の割合(テレビ)
D	直接情報
P	間接情報

表 4.3 フィッティングを行うパラメータ

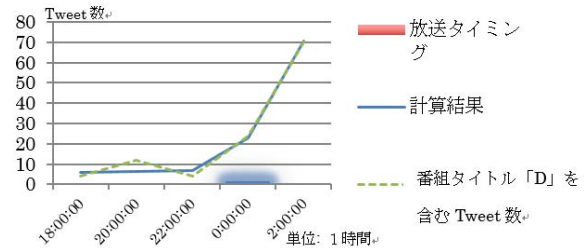


図 4.2
緑の点線(「D」を含む Tweet 数(201E/F/AE))
青の曲線 ((計算結果(201E/F/AE))
赤の縦棒 (放映時間(0:00:00~2:00:00)のタイミング (201E/F/AE))

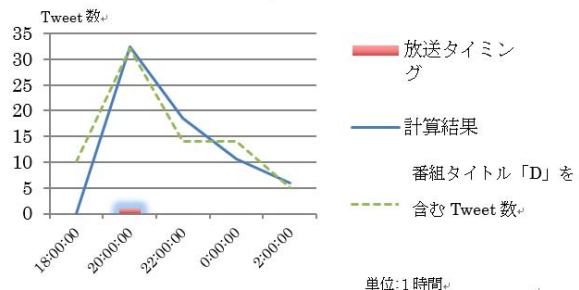


図 4.3
緑の点線(「D」を含む Tweet 数(201E/F/AG))
青の曲線 ((計算結果(201E/F/AG))
赤の縦棒 (放映時間(0:00:00~2:00:00)のタイミング ((201E/F/AG))

調査期間内の関心	201E/F/AG (ローカルテレビ)	201E/F/AG (インターネットテレビ)
Cadv	15.4	32.4

表 4.4

調査期間内の関心のパラメータ表

図 4. 2 と図 4. 3 の計算の結果を見る限り、テレビの放映のタイミングを入力として計算をしても実際の Tweet 数とかなり類似した結果が得られた。しかし、表 4.4 から明らかなように 201E/F/AG のインターネットテレビで放映した際の関心の割合が一番高く、その他のテレビでの放映の際には大差がない結果となった。

5. 考察

番組タイトル「D」を含む Tweet をした地域別の割合

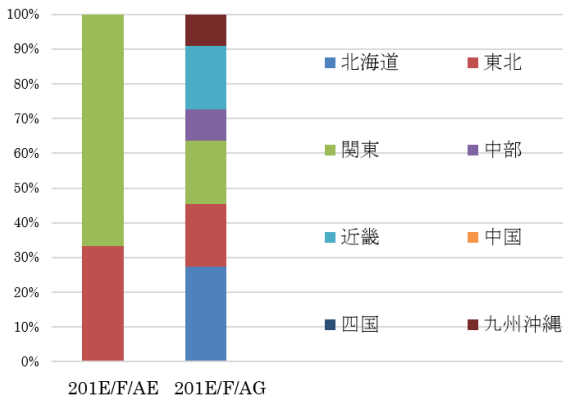


図 5. 1

番組タイトル「D」を含む Tweet をしたファンの地域別の割合

(放送日:F/AE, F/AG)

表 5. 4 における Web 上の広告効果の結果を踏まえた上で図 5. 1 を見ると、ローカルテレビで放映が行われた F/AE に関しては主に関東のファンが Tweet をした割合が高いのがわかった。しかし、インターネットテレビで放映された番組に関しては関東も多くを占めるが近畿や北海道のファンもやや多く反応しているのがわかった。インターネットテレビで同じ内容の番組の放映を行った場合は全国的に閲覧が可能である点からも全国のファンが番組に対して反応をしていたことが図 5. 1 からわかり、表 4. 4 の結果を踏まえて考えると、インターネットテレビでのパフォーマーが「D」に出演したことは全国のファンに情報が届き、反響も最も大きかったと考察ができた。

番組タイトル「D」を含む Tweet 数（直接情報と間接情報）の性質の割合 (放送日:F/AE, F/AG)

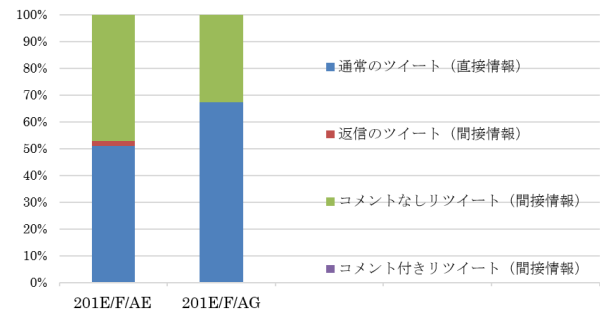


図 5. 2

番組タイトル「D」を含む Tweet 数（直接情報と間接情報）の性質の割合 (放送日:F/AE, F/AG)

番組タイトル「D」を含む Tweet 数（直接情報と間接情報）の性質の割合 (放送日:F/AE, F/AG)

期間内から予測される直接情報Dと間接情報P

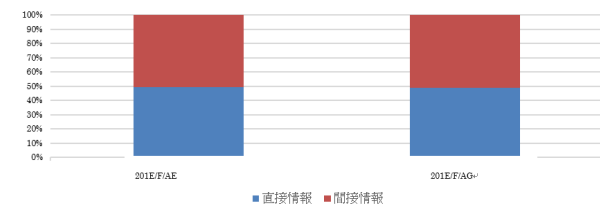


図 5. 3

ケーススタディ 2 における期間内から予測される直接情報 D と間接情報 P

図 5. 2 は計算結果によって算出された直接情報と間接情報のパラメータであり、示している割合は 2 つを足した場合である。図 5. 2 と図 5. 3 を比較すると放送日である F/AE に関しては直接情報にあたる Tweet 数と間接情報にあたる Tweet 数の割合の比としてかなり近い比率であるとみることができた。つまり、ケーススタディ 2 に関しては Web 上のパフォーマーにおける情報の直接情報と間接情報の予測はうまくいったと考えることができた。以上の表 4. 4 と図 5. 1 から図 5. 3 からわかるまででわかったことはインターネットテレビは全国ネットで見られることとローカルテレビは近隣地域でしか見られないことがパラメータの違いとしても結果として現れたと考えられる。つまり、今回ケーススタディの対象として扱ったパフォーマーの場合、全国的にファンがいることが明らかなことから、テレビ放映時は全国的に見られた方が「C」にとっては 1 日単位の関心の割合は大きいと言えた。しかし、逆に言うと全国的に番組とパフォーマー自体に需要があることがわかるため、ローカルで放映したテレビ番組の需要も全国的にあるとわかった。

6. 今後の課題

実際に Twitter を利用した安価な広告が成功したかどうかをチケットの売り上げ高の推移を見ることによる分析、またパフォーマーを今まで知らなかった観客への興味を引き寄せる要素として Twitter は効果があったか舞台公演後に回収したアンケートや Twitter の口コミ自体のテキストマイニングを利用してより細かく考察する必要がある、それに伴い数理モデルの拡張が必要となるだろう。

また、今回のケーススタディ内のプロモーション活動の分野で定量化できなかったテレビ放映以外の広告の影響や過去からのパフォーマーに関する人気の影響、部分の数量の取得と分析も必要となる。また、今回のケーススタディで取り扱ったパフォーマーに関する結果で得られた知見をもとに実際に現場で得られるはずの実測のデータを使用しながら、本格的に様々なパフォーマーに関するプロモーションの応用に関して取り組む必要がある。

今回の数理モデルの分析においては、ケーススタディであったため、 $a(t)$ には実際の観客動員数や視聴率などの具体的な数量を入れずに重みづけをしていない数量を入力として入れたが、さらに今回の研究成果から得た結果より、本来必要である数量化データから得た情報を使いと実際のチケットの利益との関係性を得ながらのさらなる検討と実験も必要となると考える。さらに、Twitter上では話題にはならないパフォーマーの認知度との関係性も明らかにしなければならず、実際にTwitterで頻繁につぶやく年齢層や男女比、地域差なども発生しているはずなので、これらのさらに厳密に検討を行った上で、あらゆる実測の数量やテキストマイニングからの知見を得た上での解析の積み重ねが必要となると考える。

上記のようなさらなる解析を経て、実際のコンテンツの受け取り手の刺激（視覚効果、音響効果、メッセージ内容等）の計算結果への影響の分析など、行動心理学の新たな解析手法へとつながる分析も行っていきたい。

今回の解析では、間接情報と直接情報の計測方法の検討も必要となり、Twitterの場合やその他のSNSの場合の検討なども必要となる。その他にも検討すべきパフォーマーがあることを考慮にいたした上でパフォーマーのファン層のTwitterの利用率なども考慮にいて、別のケーススタディからの知見を得ながら、Web上のパフォーマーに関する情報のジャンル別や活動別のパターンの知見が得られればと考えた。

参考文献

- [1] Yasuko, Kawahata, Etsuo Genda, Chinami Hara, and Akira Ishii. "Analysis of Local Concerts Using Facebook Adapting the Mathematical Model of Hit Phenomena.", *Soft Computing in Machine Learning*, Springer International Publishing, (2013):53-58.
- [2] Yasuko, Kawahata, Etsuo Genda, Chinami Hara, and Akira Ishii. "Promotion Theater Company adopting a mathematical model of the Hit Phenomenon.", *Advanced Science and Technology Letters*, Vol. 35, *Cloud and Super Computing 2013*, (2013):62-65.
- [3] A. Ishii, H. Arakaki, N. Matsuda, S. Umemura, T. Urushidani, N. Yamagata, and N. Yoshida, "The 'hit' phenomenon: a mathematical model of human dynamics interactions as a stochastic process," *New Journal of Physics*, vol. 14 (2012):063018.
- [4] A. Ishii, K. Furuta, T. Oka, H. Koguchi and K. Uchiyama, "Mathematical model of hit phenomena as a theory for collective motion of human mind in societies", in *IOS press Ebook: Intelligent Decision Technologies* edited by Rui Neves-Silva, Junzo Watada, Gloria Phillips-Wren, Lakhmi C. Jain, Robert J.

Howlett, " ", *Frontiers in Artificial Intelligence and Applications* 255 (2013):267 - 276

[5] Ishii A, Ota S, Koguchi H and Uchiyama K, "Quantitative analysis of social popularity of entertainments using mathematical model for hit phenomena for Japanese pop girl group AKB48", the proceedings of the 2013 International Conference on Biometrics and Kansei Engineering(ICBAKE2013):143-147

[6] A. Ishii, H. Fujimoto, W. Fukumoto, H. Koguchi, and K. Uchiyama, "Mathematical model for the hit phenomenon as a stochastic process of interactions of human dynamics and its application to move and online

music market", presentation in WEHIA2012 in Paris.

[7] A. Ishii, S. Ota, T. Tanimura, A. Kitao, H. Arakaki, K. Uchiyama and T. Usui, "Mathematical Theory for Hit Phenomena as a tool to analyze social phenomena using social networks", *Proceedings of AAMAS2014*, (ACM Digital Library, New York 2014) in press.