

人々の関心度推移モデルの構築

グループ演習 4班

原田礼朗 福井大樹 安田健一郎 呉迪 (学籍番号順)
アドバイザー教員 遠藤靖典

1 はじめに

我々は日々、TV・新聞・インターネット等から様々な情報を得ている。その情報は人々の日常生活に多大な影響を及すが、人々の情報に対する関心は報道のされ方・報道量に大きく左右される。例を挙げるまでもなく、時として行き過ぎた報道は不正確な情報伝達のみならず、その事項に対する適切な対処法を人々に誤らせる。つまりメディアの報道は、社会事象に対する誤った認識を人々の中に形成する可能性があり、この問題はリスク認知とも大きく関係してくる。

これらに関連して、これまで大規模な社会調査によって社会問題の存在と規模が明らかにされてきた[?]。一方、今日では多くの情報がインターネット上に流通しており、これらの情報を用いた定量的な調査が可能になりつつある。この調査手法の一つとして、ブログと呼ばれる情報発信形態を利用することが挙げられる。ブログとは、個人や数人のグループで運営され、日々更新される日記的なウェブサイトの総称である。ブログには、人々が現在関心を持っている事例が記述されており、ブログを分析することで、人々の関心度(本研究では、何らかの方法で定量的に得られる、当該事例に対して関心を持つ人・記事の数とする。)の規模を把握できると考えられる。福原ら[?]の研究では、ブログ記事を用いて社会的関心の動向分析を行っており、ブログを分析することで人々の社会問題についての反応や関心を容易に把握できること、社会的関心にはいくつかのパターンが存在することを明らかにしている。また中桐ら[?]は、社会的ブームにおける人々の状態の変化を、微分方程式モデルを構築することで説明している。

本研究では、ウェブ上の事例に関するブログ数を人々の関心度と位置づけ、事例に対する関心度の推移を調査し、その推移の分類を行う。次に、関心度の推移を数理モデルとして構築する。そして事例に対する関心度の推移を報道量から予測する手法について検討する。

2 データ収集方法

本章では、関心度を表す指標として位置づけたブログ数と報道事例の報道量の実データの収集方法について述べる。

2.1 関心度の推移を示す実データの収集方法

事例に対する関心度の推移を示す実データには、YAHOO! ブログ検索[?]による特定のキーワードに関するブログのエントリ数の推移を用いる。その他にも関心度の推移を示すデータは mixi キーワードランキングや同様のブログ検索サイトなどから収集することも可能だが、数値データ収集の容易さ、データを収集可能な期間などから YAHOO! ブログ検索を用いることとした。例として YAHOO! ブログ検索による、キーワード「裁判員制度」を含むブログのエントリ数の推移を図?? に示す。

2.2 報道量の推移を示す実データの収集方法

事例に対する報道量の推移のデータを収集する方法として、Google News Archive Search[?] を利用する。「裁判員制度」に関する報道量の推移を図?? に示す。

3 関心度推移の分類

本章では、関心度の共通項を調べるために、それぞれの事例ごとに関心度推移の分類を行った。社会的関心を web 上のブログ数として定義した場合に、その時系列変化にはいくつかのパターンが存在する。福原らの研究では、社会的関心のパターンには周期型、漸次増加型、突発型、関心持続型、その他の 5 つの分類があるとされている[?]。

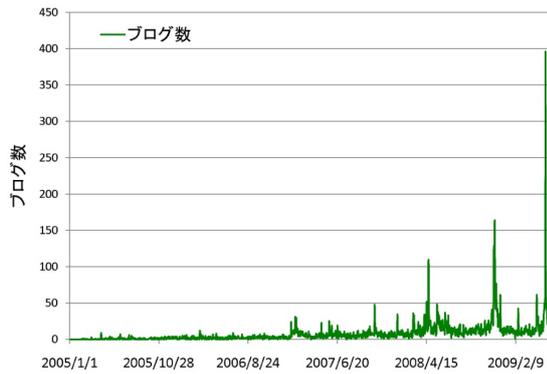


図 1: 「裁判員制度」の関心度の推移



図 2: 「裁判員制度」の報道量の推移

本研究では、社会的関心の増加の仕方と減少の仕方に分けて、以下の4つのパターンを提案する。

- 漸次増加型：徐々に人々の関心が高まる
- 突発型：突発的に人々の関心が高まる
- 関心持続型：人々の関心が持続する
- 関心収束型：人々の関心がすぐに収束する

実際のニュースで報道された事例の関心度の分類結果を図??に示す。分類結果より、ニュースで報道される事例はカテゴリーごとに、図??のように共通の推移パターンを持つことが確認できる。

また「四川省地震」「チベット問題」「裁判員制度」、「大分県汚職事件」の4つの事例は図??に当てはまらず、本来関心が収束するはずが、関心持続型になっている。この原因については、

- 四川省地震：他の災害事例に比べて死者数が多い、北京オリンピックと同時期に発生

	漸次増加型	突発型
関心持続型	トリノオリンピック 北京オリンピック ワールドカップドイツ大会 裁判員制度 郵政民営化(保留) GM 崖の上のポニョ 鳥インフルエンザ	四川省地震 新型インフルエンザ チベット問題 世界同時株安 金融不安 おくりびと 大分県 教育委員会
関心収束型	第1回ワールドベースボールクラシック 第2回ワールドベースボールクラシック 洞爺湖サミット開催 リーマン・ショック タスボ 北朝鮮 ミサイル 冥王星 食品偽装(保留)	スマトラ沖地震 ハビスタン地震 地震半島地震 ペルー沖地震 安徳内閣 康生内閣 小沢内閣 機器修繕改正案 草なぎ 逮捕 マイケル 死 サザンオールスターズ活動休止 福知山線 脱線事故 中央大 教授 刺殺 北朝鮮 核実験 足利事件 分学優等 インド ムンバイ 同時テロ 秋葉原 無差別殺傷事件 iPhone 盗田 宇宙 タスボ 日本 ノーベル賞 日本人 4人

図 3: 事例ごとの分類

	漸次増加型	突発型
関心持続型	スポーツ 映画 病気	映画 病気
関心収束型	スポーツ 生活 政治関連	災害 政治関連 芸能 エンターテインメント 事件・事故 時事関連

図 4: 事例ごとのカテゴリー別の分類結果

- チベット問題：北京オリンピックと同時期
- 裁判員制度：一般の人々にとって身近な問題
- 大分県汚職事件：一連の汚職事件の皮切りとなった事件

が考えられる。このように人々の関心度は複数の事例が関わることで、大きく変化する。本研究では上記のような4つの事例を除いて、図??の分類を参考に関心度の予測を行う。

4 関心度推移モデル

本章では、人々の関心度推移を予測するために関心度の変動を表す数理モデルを提案し、各パラメータと関心度の推移パターンとの関連について考察する。

4.1 関心度推移モデルの提案

先行研究として、中桐らにより社会的ブームの人々の変動を表す微分方程式モデルが考案されている [?]。本研究では、上記のモデルを参考として、ある事例が時間 t において報道がされたとき、 $z(t)$ を報道量、 $y_1(t)$ を報道事例を知らない人々の人数、 $y_2(t)$ を報道事例に対して興味を持った人々（関心を持つ人々）の人数、 $y_3(t)$ を報道事例に対して興味を持たない人々（無関心な人々）の人数と定義して、人々の関心の変化を表すモデルを微分方程式を用いて以下のように表す。

$$y_1'(t) = -b_1 z(t) y_1(t) - b_2 z(t) y_1(t) \quad (1)$$

$$y_2'(t) = b_1 z(t) y_1(t) + b_3 z(t) y_3(t) - b_4 y_2(t) \quad (2)$$

$$y_3'(t) = b_2 z(t) y_1(t) + b_4 y_2(t) - b_3 z(t) y_3(t) \quad (3)$$

定数 $b_1 \sim b_4$ について、 b_1 は関心を持たせる度合い、 b_2 は関心を持たせない度合い、 b_3 は関心が再び起こる度合い、 b_4 は関心が薄れていく度合いと解釈することができる。

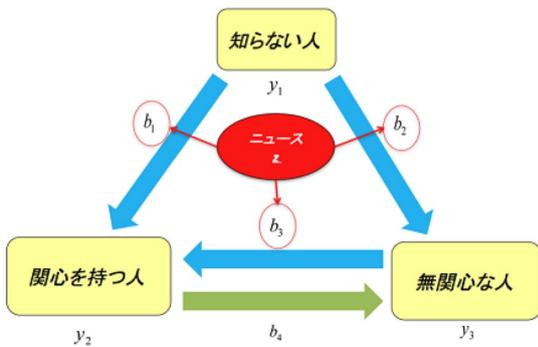


図 5: 関心度の変動イメージ図

4.2 各パラメータと関心度の推移パターンの関連について

yahoo ブログの件数は 2009 年 4 月 29 日で 200 万人となっているため、

$$y_1(t) + y_2(t) + y_3(t) = 2000000 \quad (4)$$

として数値計算を行った [?]。計算方法は、4 次のルンゲ・クッタ法を用いた。以下に実際の「中越沖地震」の報道を例にとった挙動を示す。

一般的に関心を持つ人たちがブログにその内容を記述すると考えて、 $y_2 =$ ブログ数と考える。図??と図??の比較から、このモデルが関心度を推測するのに受



図 6: 「中越沖 地震」の報道量 $z(t)$ の時系列変化

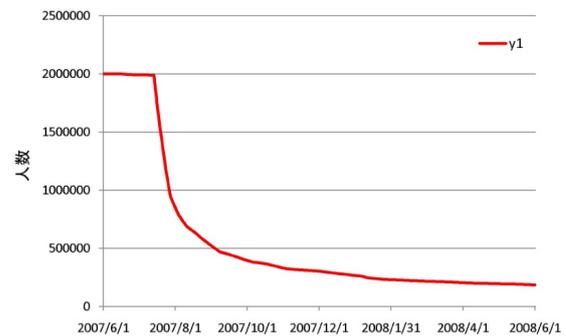


図 7: $b_1 = 0.012$, $b_2 = 0.988$, $b_3 = 0.0001$, $b_4 = 0.9$ のときの $y_1(t)$ の時系列変化

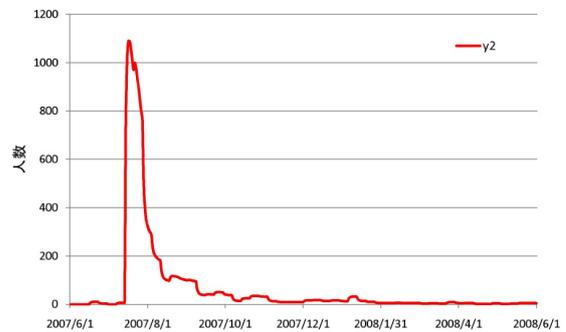


図 8: $b_1 = 0.012$, $b_2 = 0.988$, $b_3 = 0.0001$, $b_4 = 0.9$ のときの $y_2(t)$ の時系列変化

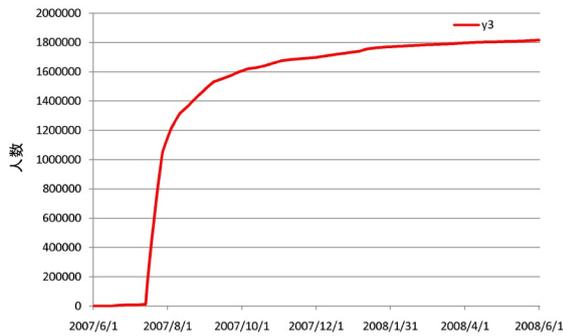


図 9: $b_1 = 0.012$, $b_2 = 0.988$, $b_3 = 0.0001$, $b_4 = 0.9$ のときの $y_3(t)$ の時系列変化

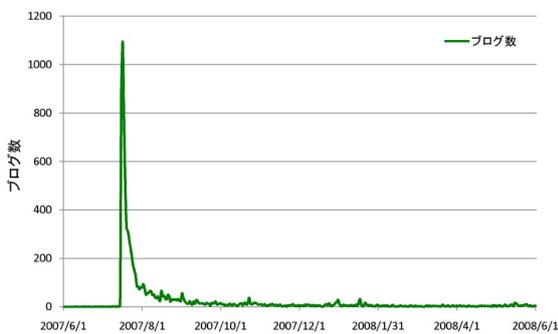


図 10: 「中越沖 地震」の yahoo ブログ検索による関心度推移

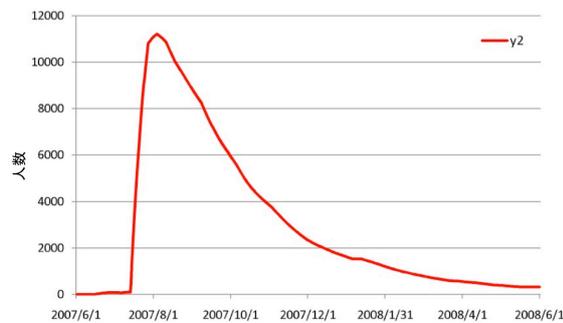


図 11: $b_1 = 0.012$, $b_2 = 0.988$, $b_3 = 0.0001$, $b_4 = 0.02$ のときの $y_2(t)$ の時系列変化

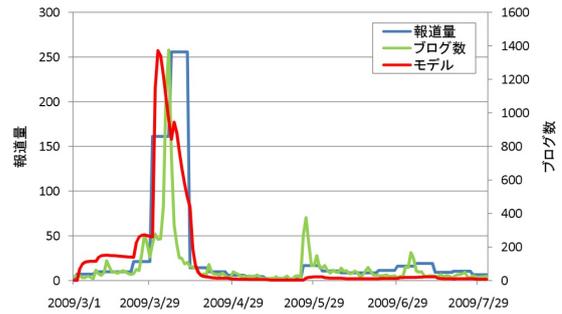


図 12: 「北朝鮮 ミサイル」の報道量, ブログ数, $y_2(t)$ の時系列変化

当であることが確認できる．さらに，パラメータの値を変化させたときの挙動を図??に示す．

図??は関心収束型の波形を示しているのに対して，図??の関心は持続していることが確認できる．以上から， b_3 が大きく， b_4 が小さいときに関心持続型を表し， b_3 が小さく， b_4 が大きいときには関心収束型を表すと考察できる．

次に「北朝鮮 ミサイル」の報道を例にとった挙動を図??に示す．

ブログ数は漸次増加型になっており，モデルの $y_2(t)$ は漸次増加型を表していることが確認できる．関心度の増加の挙動については報道量の推移に大きく依存しており，報道が突発的であれば，関心度は突発型になり，報道が漸次増加であれば，関心度は漸次増加型になる．

5 関心度推移モデルを用いた予測

本章では，前章に述べた関心度推移モデルを用いて関心度の推移の予測するための手法を述べる．最初に，現在進行中の対象事例を選び，YAHOO! ブログ検索と Google News Archive Search を用いてブログ数と報道量の実データを抽出する．次に，図??の分類結果にもとづいて，事例のカテゴリ別に関心が収束するか持続するかを判別するために b_3 , b_4 のパラメータを決定し，実際のブログ数とモデルの最大値が一致するように b_1 , b_2 を決定する．その際に，報道がされたときにその事例を知る割合は一定であり，報道量に完全に依存すると仮定して， $b_1 + b_2 = 1$ とする．最後に今後の報道量を予測し，モデルに入力して関心度推移の予測を試みる．将来の報道量については，収束すると仮定して今後の報道量を対数近似したものと事例ごとに今後起こるであろうことを予測したものを与えた．

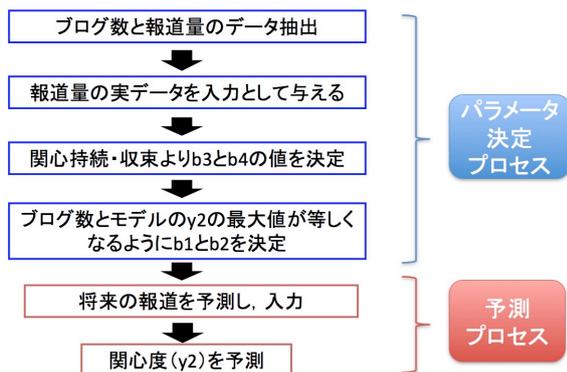


図 13: 関心度推移予測の流れ

6 対象事例の予測数値例

本章では、2009年4月～9月にかけて注目を浴びた対象事例について、関心度の数理モデルを用いた関心度推移の将来予測を行う。以下に対象とした事例と予測結果を記す。

6.1 キーワード「新型インフルエンザ」

6.1.1 概要

新型インフルエンザとは、新たに人から人に伝染する能力を有することとなったウイルスを病原体とするインフルエンザであり、2009年4月下旬ごろから世界中で流行し始めた。

6.1.2 主な出来事・報道

- 4月：メキシコでの感染を皮切りに世界中で感染が広がる。
- 5月8日：日本で初の感染が確認される。
- 6月12日：WHO（世界保健機関）がフェーズ6への引き上げを宣言。
- 8月15日：沖縄県で新型インフルエンザによる日本で初の死者が確認される。
- 8月29日：WHO（世界保健機関）が新型インフルエンザによる死者数が2185人であることが発表される。

6.1.3 関心度推移モデルのパラメータの決定

キーワード「新型インフルエンザ」について関心度推移の予測を行う。病気のカテゴリに属することから関心持続になり、関心度のピーク値がYAHOO!ブログ検索の結果にあうようなパラメータを設定する。YAHOO!ブログ検索によるキーワード「新型インフルエンザ」には、2つのピークが存在しているが、今回は1つめのピークに合わせることにし、パラメータの値は $b_1 = 0.0085$ 、 $b_2 = 0.991$ 、 $b_3 = 0.0004$ 、 $b_4 = 0.9$ となった。図??に2009年4月1日～2009年8月22日までの1日ごとの報道量の推移、YAHOO!ブログ検索による関心度の推移、モデルを用いた同定結果を示す。図??から

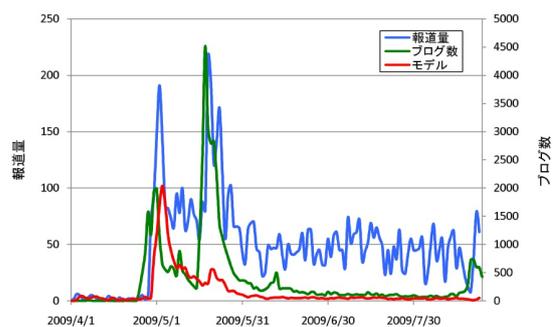


図 14: 「新型インフルエンザ」の関心度の同定結果

わかるように片方のピーク値に合わせると、もう一方が合わないことが確認できる。事例「新型インフルエンザ」の関心度推移については、本研究で提案した関心度推移モデルではうまく表現できない。関心度推移モデルに当てはまらない原因は、マスコミによる報道内容や報道量が過剰すぎるなどが原因として挙げられる。過剰な報道が人々は関心を高め、本来の関心度推移とは違った挙動をしたと考えられる。これは、YAHOO!ブログ検索の結果からも分かる。キーワード「新型インフルエンザ」の最大記事数は、4500以上になっている。この記事数は他の事例に類を見ない多さであり、過剰な報道が人々の関心に大きく影響していることがうかがえる。

6.1.4 関心度推移の予測

新型インフルエンザの関心度推移を予測するために、本論文では3つのシナリオを提案する。まず最初に新型インフルエンザに対する報道が収束した場合のシナリオである。既存の報道量から対数近似を用いて報道の時系列変化を予測した。予測した関心度推移の結果を図??に示す。YAHOO!ブログ検索の結果では、8月

後半に関心が収束するような挙動はみられないことから、このような収束の推移をたどるのは考えにくい。

次に、現実にもつシナリオである。厚生労働省の発表によると流行のピークは10月上旬で、12月上旬には終息すると予想がされている。そこで、10月上旬に報道量のピークがくるような関数を報道量に与えた。結果を図??に示す。

最後のシナリオとして、2つ目のシナリオと関連して、10月に起こる報道が今までの報道量の中で最大になるように与えた。関心度推移の予測結果を図??に示す。図??、図??より、再び報道が生ずれば関心度も再び生ずること、報道量の増加にしたがって関心度も増加することが予測できる。関心持続型のパラメータにおいては、どんな報道量を与えても1回目のピーク値を超えることがないことが確認できる。この結果から、本研究のモデルにおいては、人々の関心が持続すると一旦収束した関心を再び起こすのは難しいということが考えられる。

以上の結果より、新型インフルエンザの事例に関して、最初の報道が人々の関心度に多くの影響を与えたため、感染率の上がる10月に関心度が上がらず、多くの人々がリスク認知を誤り、感染の対策を怠ってしまうという予測をすることができる。

6.2 キーワード「裁判員制度」

6.2.1 概要

裁判員制度とは、国民が刑事裁判において裁判員として参加し、被告人が有罪かどうか、有罪の場合どのような刑にするかを裁判官とともに審理する日本の司法・裁判制度のことである [?]。

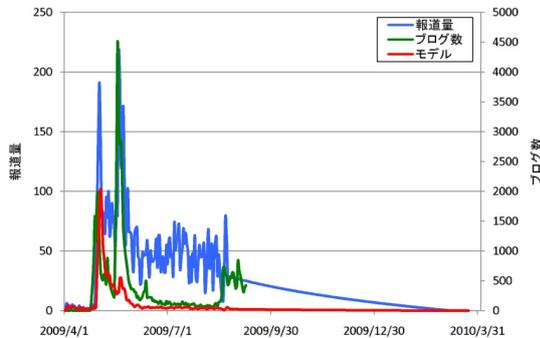


図 15: 「新型インフルエンザ」の関心度推移予測 1

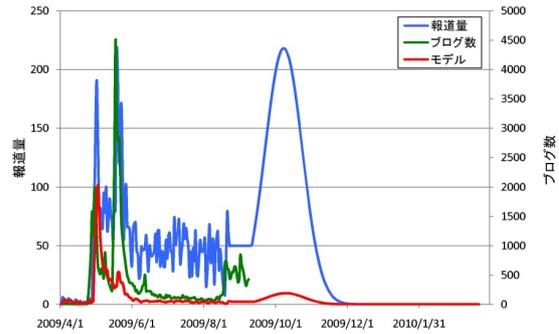


図 16: 「新型インフルエンザ」の関心度推移予測 2

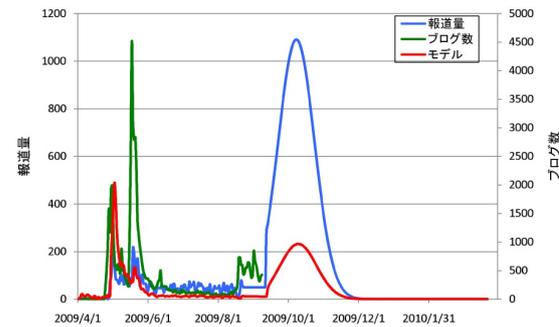


図 17: 「新型インフルエンザ」の関心度推移予測 3

6.2.2 主な出来事・報道

- 2004年5月21日:「裁判員の参加する刑事裁判に関する法律」成立。
- 2009年5月21日:「裁判員の参加する刑事裁判に関する法律」施行。
- 2009年8月3日:裁判員の参加する最初の公判が始まる。
- 2009年8月10日:裁判員の参加する2例目の公判が始まる。

6.2.3 関心度推移モデルのパラメータの決定

裁判員制度はすでにほとんどの人に知られている事例であり、関心度推移モデルの初期値は $y_1 = 200000$, $y_2 = 11$, $y_3 = 1800000 - 11$ とする。全体の人数はYahoo ブログのブログ数 200万, y_2 はモデルによる近似の開始時点でのブログ数としている。パラメータを調整した結果が図??であり、またこのときのパラメータは $b_1 = 0.001$, $b_2 = 0.999$, $b_3 = 0.015$, $b_4 = 4.2$ で

ある。報道量と関心度にずれがあり、モデルによる近似は精度のよくないものとなっている。特に報道の最初のピークの部分において、ブログ数はピークとなっていない。この時期は各地裁ごとに裁判員裁判の対象となる事件が決まったり、裁判員の選任手続きが行われており、報道が局所的で全国的に関心を引くことがなかったためと考えられる。

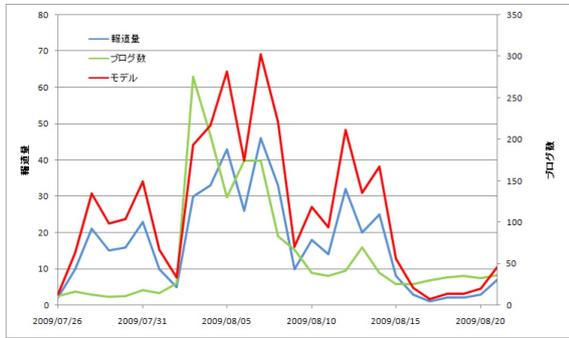


図 18: 「裁判員制度」の関心度の同定結果

6.2.4 関心度推移の予測

実際の関心度の推移を近似したモデルによって、今後の報道量によって関心度がどのように推移するかを予測する。まずピークから最後の報道量まで対数近似し、近似曲線上の値を今後の報道量とした場合が図??である。関心度（関心の強い人 y_2 ）はすぐに 0 に収束していることがわかる。次に、日程が決まっている公判に対して、公判中の期間に報道を発生させた場合が図??である。判決が決まる最終日に報道が少し大きくなり、また裁判員制度への慣れによって徐々に報道は少なくなるという仮定で報道量を決定している。この場合、公判期間、特に後半最終日に当たる日の関心度が増加するが、公判を重ねるごとに関心度が上昇は少なくなる様子が見られる。

6.3 キーワード「酒井法子」

6.3.1 概要

ここでは、2009年8月8日に覚せい剤取締法違反容疑で女優の酒井法子が逮捕された事件を対象とする。

6.3.2 主な出来事・報道

- 2009年8月3日：夫が覚せい剤取締法違反（所持）容疑で現行犯逮捕される。

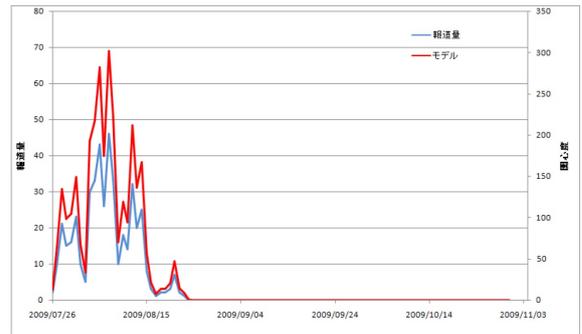


図 19: 「裁判員制度」の関心度の推移予測 1

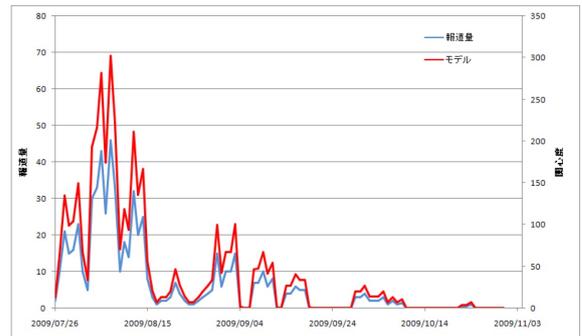


図 20: 「裁判員制度」の関心度の推移予測 2

- 2009年8月8日：警視庁富坂分庁舎に出頭、覚せい剤取締法違反（所持）容疑で逮捕される。
- 2009年8月22日：覚せい剤取締法違反（使用）容疑で再逮捕。
- 2009年8月28日：覚せい剤取締法違反（所持）の罪で起訴される。所属事務所を解雇される。

6.3.3 関心度推移モデルのパラメータの決定

キーワード「酒井法子」について関心度推移の予測を行う。芸能のカテゴリに属することから関心収束型になると予想され、関心度のピーク値がYAHOO!ブログ検索の結果にあうようなパラメータを設定した。図??に2009年7月27日～8月21日までの1日ごとの報道量の推移、yahooブログによる関心度の推移、モデルを用いた同定結果を示す。このときのパラメータは $b_1 = 0.032$, $b_2 = 0.968$, $b_3 = 0.0001$, $b_4 = 2$ となった。モデルの推移は報道量に大きく依存するため完全に一致させることは難しいが、他の事例よりも比較的精度の良い結果となった。このような関心収束型

の事例は、関心持続型と比較してモデルを一致させやすい傾向にあった。すなわち、関心収束型に代表される芸能・政治関連等の事例に対する人々の関心は、そのほとんどが報道量によって左右されているといえる。

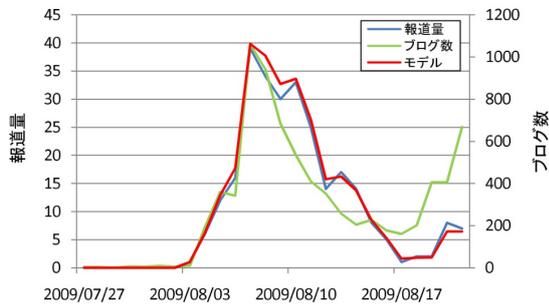


図 21: 「酒井法子」の関心度の同定結果

6.3.4 関心度推移の予測

ここでは2つのシナリオで予測した。1つ目はこのまま報道がほとんどされずに収束するシナリオである。報道量のピークから最後の報道量までを対数近似し、近似曲線上の値を今後の報道量とした場合の関心度の推移を図??に示す。関心度はすぐに0に収束していることがわかる。

2つ目は将来的に起きると予想される出来事から報道量を予測した。酒井法子に関する大きな事象としては、2009年10月21日に夫である高相祐一容疑者の初公判があり、26日に酒井容疑者の初公判が予定されているため、ここに報道のピークがくると予想される。これに基づいて報道量を入力した結果が図??である。

報道量が増えると当然ながら関心度も増加した。しかし、10月26日に1度目のピークと同量の報道量を入力しても、関心度は1度目のピーク程上昇しなかった。この理由としては、報道が再び起きてもすでにこの事件を知っている人が多く、また報道によって再び関心を持つ人が少ないためであると考えられる。

7 おわりに

本研究では、はじめに人々の関心度の推移パターンを分類した。関心度推移には、いくつかのパターンが存在し、事例のカテゴリごとに共通のパターンがあることが分かった。次に、関心度推移モデルを提案した。関心度推移の分類結果と実データの最大値から、パラメータを決定し、過去に終息した事例を用いて、モデルの妥当性について検証した。最後にモデルを使って、

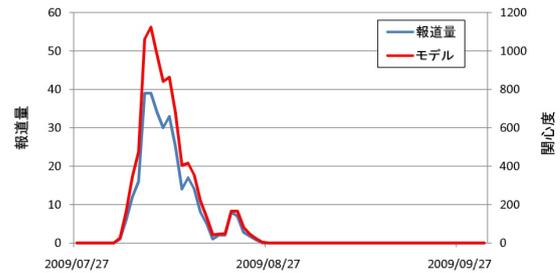


図 22: 「酒井法子」の関心度の推移予測 1

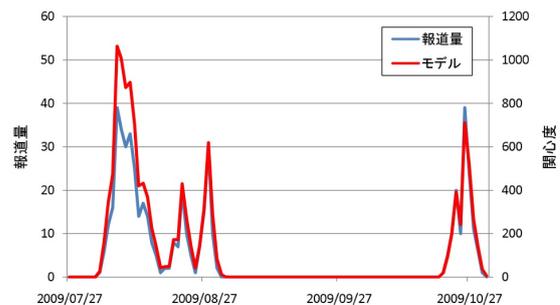


図 23: 「酒井法子」の関心度の推移予測 2

関心度の予測を行った。報道量に基づいて、人々の関心度を推移する数理モデルはこれまで存在せず、ますます増大するメディア報道の社会的影響を無視できない現代においては、本研究の意義は非常に大きいと言える。

今後の課題として、まず関心度の分類が挙げられる。関心度の収束パターンを本研究では、関心持続型と関心収束型を提案したが、この2つのパターンの明確な定義は既存研究では、存在しておらず今回の研究でも明確な判別方法を見つけることはできなかった。持続と収束の原理をより深く考察して、2つのはっきりとした境界線について議論を深める必要がある。次に、今回提案した関心度推移モデルの改善が挙げられる。本研究の関心度推移モデルでは、実際のデータと一致しない事例がいくつか存在した。これは、社会現象や自然現象は非線形要素や確率的要素を含むので、今回提案したような線形のモデルでは表現するのが困難である。そこで、例えば確率の概念を導入した数理モデルによって関心度のより適切な推移予測をすることが可能になると思われる。

参考文献

- [1] 森岡清志, 'ガイドブック 社会調査', 株式会社 日本評論社, 1998.

- [2] 福原 知宏, 村山敏泰, 西田豊明, “ウェブブログ記事を用いた社会事象に関する関心動向の分析”, 人工知能学会第 6 回 AI 若手の集い, 2005.
- [3] 中桐裕子, 栗田治, “社会的なブームの微分方程式モデル”, 日本オペレーションズ・リサーチ学会和文論文誌, Vol.47, pp.83-105, 2004.
- [4] YAHOO! ブログ検索,
<http://blogsearch.yahoo.co.jp/>
- [5] Google News Archive Search ,
<http://news.google.com/archivesearch/>
- [6] 裁判員制度 ,
<http://www.saibanin.courts.go.jp/>