

甲子園開催期間中の Twitter 投稿群から観測する 白熱度の時系列変動

本木 悠介 岩井 将行

東京電機大学 未来科学部 情報メディア学科

{yusuke, iwai}@cps.im.dendai.ac.jp

概要 情報社会の中で発展している SNS の一つである Twitter に存在する情報量は膨大であり、適切に利用することによって有用な情報収集手段となりうる。本研究では、Twitter 上のイベントに関する開催期間中のツイートを収集する。次いで、ツイートを言語解析し、イベントが白熱した要素の特定、その要素の白熱度合いの時系列変動をクラスタリング分析により観測することを目的とした。本研究で対象とするイベントは、第 100 回全国高校野球選手権大会(通称:甲子園)とした。最初に、甲子園の開催期間中に投稿されたツイートを収集した。次いで、甲子園を白熱させた要素を、収集した各ツイートに対して形態素解析を行うことによって特定した。その結果、甲子園を白熱させたのは秋田県立金足農業高等学校ということが分かった。また、形態素解析を行なった各ツイートからポジティブ・ネガティブ値、注目値を算出した。それらデータを日毎に区切り、クラスタリング分析を行う。集めたデータとクラスタリング分析を行った結果の推移を用いて甲子園の白熱度の時系列変動を観測した。最後に、観測によって得られた知見を踏まえた考察と今後の展望について述べる。

キーワード Twitter, 自然言語処理, クラスタリング, イベント分析, 時系列変動

1 はじめに

昨今の情報社会における中で、SNS は大きく発展している。その中で生まれたサービスは数多く、Twitter や Facebook, Instagram などのサービスがあり、それぞれのサービスが様々な用途に利用されている。それら SNS の中でも Twitter を利用する人は多く、日本においてのアクティブユーザは約 4500 万人に上る[1]。また、Twitter 社の発表によると全世界で 1 日に 5 億件のツイートが投稿されている。非常に多くの情報が投稿される Twitter では自らが参加しているイベントの感想や情報を載せる人も多くいる。一方で、様々な情報が発信されているが故に、ある特定のイベントの情報を観測し、解析するのは難しい。そこで、本研究では、対象イベントを全国高校野球選手権大会(以降、甲子園という。)とし、甲子園の開催期間中のツイートを収集する。次いで、ツイートを言語解析し、そのイベントが白熱した要素の特定を行う。最後に、クラスタリング分析や各種データからその要素の白熱度合いの時系列変動を観測することを目的とする。

Twitter 上にある情報の中でも、イベントに関するツイートを題材として行われた研究は多くある。例えば、旅行者が必要としている、興味を持っていると思えるイベントの情報を提供するシステムに関する研究がある[2]。この研究は雑誌などに情報があまり載っていないローカルイベントの情報を Twitter 上のツイートを参考にし

て、ユーザに提供するといったものである。また、過去のイベントに関するツイートを収集し、時空間クラスタリングを行うことによってイベントの開催時間、開催場所、イベントの内容に関する情報の抽出を行うといった研究もある[3]。この他にも、本研究と同様の試みをしている研究がある[4, 5]。Shen, C.氏らの[4]はスポーツ観戦や学術会議といったイベントの参加者が Twitter に投稿したツイート情報を基としたイベント集約アプローチを提案するもので、イベントを解析するという点において本研究と関連があると考えられる。Ifrim, G.氏らの[5]は様々なツイートからニュースとしての価値のあるイベントを検出するシステムの提案を階層的クラスタリングや時間依存性の N-gram などの手法を用いて行うものである。これらの関連研究は本研究と比較して、Twitter 上の情報は膨大であり、適切に利用するのは難しいとし、独自の解析をもって情報処理を行なっているという点においてつながりが深いと言える。一方で、本研究は特定のイベントを白熱させた要素を特定し、その要素を含んだツイート群の時系列変動をクラスタリング分析、観測するという点において他にない独自のものである。

2 観測データの準備

本章ではイベントの白熱要素の時系列変動を観測するために必要なデータの準備として、観測データとなるツイートの取得や白熱要素の推定、それに対する比較要素の選定を行い、最後にクラスタリング分析のためのデータ加工を行う。

表 1 日毎に集まったツイート件数

日付	件数	日付	件数
8/5	109,357	8/14	44,529
8/6	67,781	8/15	40,770
8/7	37,542	8/16	36,108
8/8	33,813	8/17	45,432
8/9	36,074	8/18	79,250
8/10	33,523	8/19	31,279
8/11	51,570	8/20	101,991
8/12	78,902	8/21	216,819
8/13	44,894		

2.1 データ概要

本研究に使用するツイートデータは Twitter で「甲子園」と検索し得られたツイートの中で、リツイートを除外し、甲子園開催中にツイートされたデータを使用する。具体的な開催期間は 2018/8/5 から 2018/8/21 であり、開催期間中に集まったデータの総数は 1,089,634 件である。日毎に収集したツイート件数の詳細を表 1 に示す。各ツイートから取得してくるデータのフォーマットは、ツイート ID、ユーザ名、投稿日時、テキスト本文、「いいね」数、リツイート数の 6 種類である。ツイート ID、ユーザ名はツイートの識別に利用する。投稿日時はクラスタリング分析時の時系列判断として利用する。テキスト本文は言語解析のデータとして利用し、具体的に白熱した要素の抽出、文章中の形容詞と固有名詞の抽出、文章のポジティブ・ネガティブ値(以降、PN 値という。)の算出に利用する。なお、PN 値はクラスタリング分析時の白熱度の要素の一部として利用する。「いいね」数、リツイート数は個々のツイートがどれほど注目されているかの指標とする注目値(以降、Attention 値という。)の算出に使い、クラスタリング分析時の要素として利用する。

2.2 データ収集方法

一般的に、ツイートデータを扱う際は TwitterAPI を利用することが主流であるが、TwitterAPI の検索機能である SearchAPI では制限により、一週間以上前の期間を遡ってツイートを取得することができない。そこで、一週間以上前のツイートを取得制限に関係なく取得することを可能とした、Jefferson-Henrique 氏が公開している GetOldTweets-python を利用し、期間中のツイートを収集する。

2.3 対象とする白熱要素の推定

本研究において重要となる甲子園を白熱させた要素の特定には MeCab を使った形態素解析を利用する。まず、収集したツイート全てに対し単語の分かち書きを行う。この際に利用する辞書は「mecab-ipadic-NEologd」

表 2 ツイート群の固有名詞のカウン上位 6 件

固有名詞	件数	固有名詞	件数
高校野球	115,714	熱闘甲子園	58,289
金足農業	73,133	回	52,694
大阪桐蔭	67,870	金足	42,520

という辞書を採用する。これはオープンソース・ソフトウェアとして公開されている固有名詞に強い辞書である。この辞書を採用した理由は、甲子園といったイベントに対するツイートは学校名や人名、地方名といった固有名詞を多く含んでおり、多くの固有名詞を登録している「mecab-ipadic-NEologd」を利用するのが適切と判断したからである。この他にも、MeCab は形容詞の抽出やその他単語の統計を取るのに利用する。まず、この辞書により分かち書きされた単語を集計し、カウントする。その結果、収集したツイートの中で多く見られた固有名詞が甲子園を白熱させた要素となりうると考えた。実際に集計された結果の上位 6 件を表 2 に示す。この結果から甲子園を白熱させた要素として「秋田県立金足農業高等学校(以降、金足農業という。)」が考えられる。

2.4 比較要素の選定

「金足農業」が実際に甲子園を白熱させた要素と言えるのか確認するために比較する別の要素の選定を行う。比較する要素は、本大会において「金足農業」と決勝で戦った「大阪桐蔭中学校・高等学校(以降、大阪桐蔭という。)」と「猛暑関連」の二つの要素とする。

比較する要素として、「大阪桐蔭」を選んだ理由は「金足農業」と違い、「大阪桐蔭」は当初から優勝候補としてある程度の注目を集めていた。よって、この要素と比較することにより、「金足農業」がどれだけ注目度が当初より大きく変動したかを確認できると考えた。

続いて、「猛暑関連」を選んだ理由は本大会中、常に一定の関心を集めており、それがツイートにも反映されていると考えたからである。よって、この要素と比較することにより、「金足農業」の話題性がどれほど大きいものであったか確認することができるからである。

2.5 クラスタリング用のデータ加工

2.3 節から、甲子園を白熱された要素として「金足農業」が考えられるため、この要素による白熱度の時系列変動を観測するためのクラスタリング用のデータ加工を行う。本研究において行うクラスタリング分析で用いる軸の要素は PN 値と Attention 値の 2 軸である。以下に、各軸の値の算出方法について説明する。

2.5.1 PN 値の算出

各ツイートの PN 値の算出に利用するモデルは、高村大也氏らの単語感情極性対応表[6]を用いる。本研

表 3 本手法から外した単語感情極性対応表：名詞例

名詞	PN 値	名詞	PN 値
高校	-0.143	投手	-0.504
野球	-0.860	野手	-0.315
選手	-0.254	試合	-0.549
チーム	-0.353	バッター	-0.326

究では[6]の中から動詞、形容詞、副詞の3種類のいずれかに対応する単語に対してPN値を与え、ツイート文のPN値を算出する。PN値を算出する数式は、ツイート文内の各単語のPN値の平均とする。なお、今回使用するツイート内で頻出する名詞をPN値の参考にするのは適切でない判断したため名詞をPN値算出のモデルとして利用しない。そのような判断をすることとなった単語の一部例を表3に示す。どの単語もマイナス値が大きく、これら単語によりツイート文のPN値の平均がマイナスに大きく振れてしまうのを防ぐため本研究において名詞はPN値の判定に用いない。

2.5.2 Attention 値の算出

各ツイートの Attention 値の算出には「いいね」とリツイートの数に加え、ツイートの文字数を要素として算出する。また、Attention 値の算出にはロジスティック方程式を用いる。ロジスティック方程式を用いる理由は外れ値に対応するためである。Twitter ではしばしば、「いいね」や「リツイート」を大きく稼ぐツイートが見られる。しかし、そのようなツイートは極一部で、多くのツイートは「いいね」や「リツイート」が少数にとどまるものが多い。そのような状況では外れ値だけが大きく突出してしまい、適切な Attention 値の算出ができないと判断した。このことから、ロジスティック方程式を用いるのが適切と考えた。

まず、ツイートの文字数を得点化するために $K = 1, b = 5, c = 0.05, x$ を文字数として(1)の数式に当てはめる。次いで、 $K = 1, b = 8, c = 0.05, x$ を「いいね」数とリツイート数、先ほど算出した値を足し合わせた値として再度、(1)の数式に当てはめる。この数式を経て算出された値 y を Attention 値とする。

$$y = \frac{K}{1+be^{-cx}} \quad (1)$$

3 白熱要素の特定と観測

2章で行った工程により得られた結果から「金足農業」が甲子園を白熱させた要素であるという確認を行い、「金足農業」に対する群衆の感情推移を観測する。次いで、クラスタリング分析と各種データを用いて「金足農業」の白熱度の時系列変動を観測する。

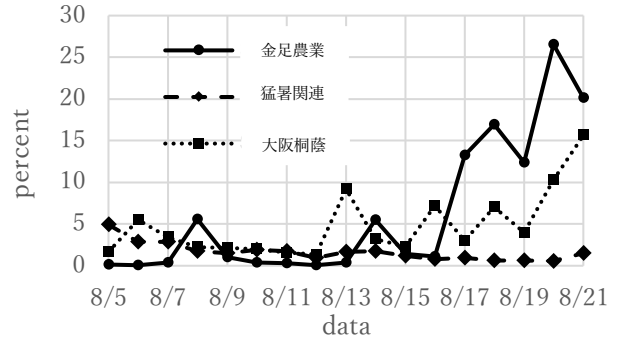


図 1 日毎の各要素のツイート占有率

表 4 「金足農業」のイベント

日付	ツイート占有率	イベント
8/8	5.6%	一回戦(鹿児島実業高校)
8/14	5.5%	二回戦(大垣日本大学高校)
8/17	13.2%	三回戦(横浜高校)
8/18	16.9%	準々決勝(近江高校)
8/19	12.4%	休養日
8/20	26.5%	準決勝(日本大学第三高校)
8/21	20.1%	決勝(大阪桐蔭高校)

3.1 白熱要素を含むツイートの占有率の変動

「金足農業」が甲子園開催期間中のツイート群をどれだけ占有していたかを確認し、その理由を特定する。また、他の要素より大きく注目されたかを 2.4 節で挙げた要素と比較しながら確認する。それにより「金足農業」が甲子園を白熱させた要素であるという確認を行う。最初に、今回収集したツイート群の中で、「金足農業」と 2.4 節で挙げた「大阪桐蔭」、「猛暑関連」の 3 つの要素がそれぞれ日毎にどれだけツイートを占有していたかを図 1 に示す。「大阪桐蔭」に関してはツイート文内に「大阪桐蔭、桐蔭」を含むツイートを対象とし、「猛暑関連」に関してはツイート文内に「猛暑、暑い、暑さ、酷暑、炎天下」を含むツイートを対象とする。図 1 から、8/8, 8/14, 8/17 以降の「金足農業」、「金足」を含むツイートの占有率が高くなっていることが分かる。

続いて、その日にあったイベントと、ツイート占有率を記した表を表 4 に示す。表 4 から、基本的に「金足農業」が注目されるのは試合があった日であることが分かる。特に準々決勝を突破して、ベスト 4 入りが決まってから大きく注目を集めている。

続いて、他の要素のツイート占有率と比べて「金足農業」が他の要素より多く注目されていたかを確認する。図 1 に表れた結果から、2 つのことが分かる。

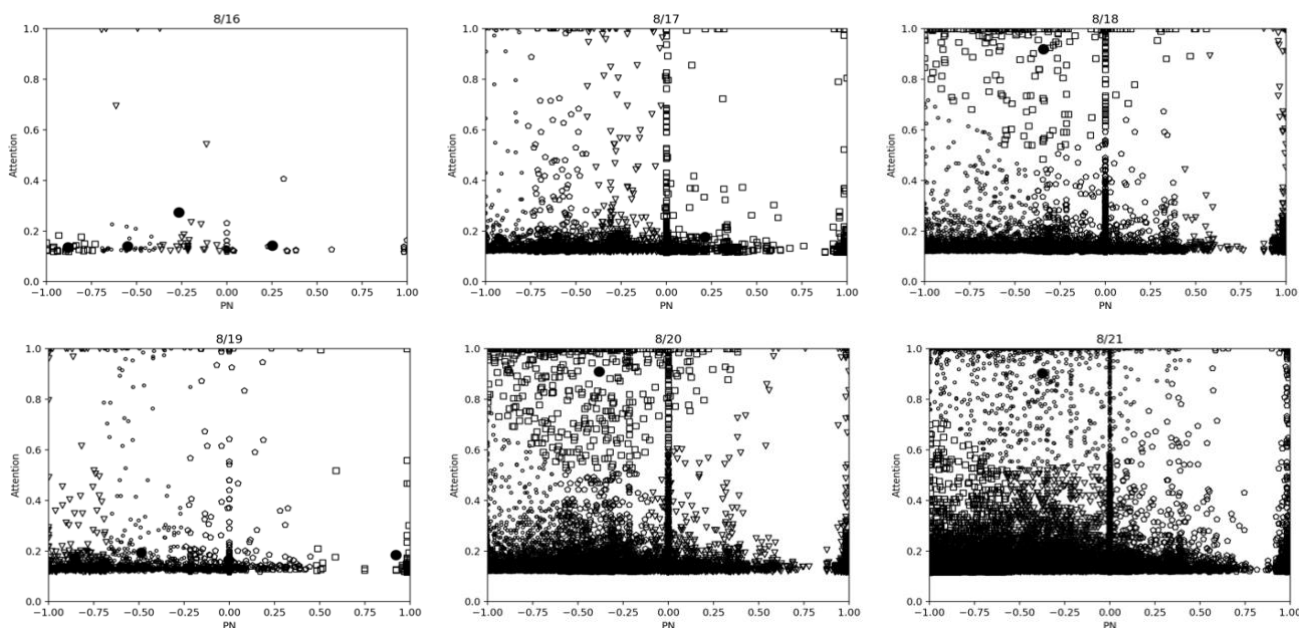


図 2 金足農業に関する日毎のクラスタリング図

- (A)「金足農業」は「大阪桐蔭」と同様、勝ち進むにつれて注目を集めているが、より多くの注目を集めている。8/16 以前は「大阪桐蔭」の方が注目を集めているが、それ以降は「金足農業」が大きな差をつけて注目を集めている。
- (B)「金足農業」は「猛暑関連」と比べて大きな話題性を持っている。「猛暑関連」は全日程を通して、ツイート占有率が 5%以内にとどまっている。それに対して、「金足農業」は序盤「猛暑関連」よりも占有率が低いにも関わらず、後半は占有率を大きく突き放している。

3.2 白熱要素のクラスタリング分析

ここまでの観測で、「金足農業」が甲子園を白熱させた要素として考えられることが分かった。続いて、「金足農業」を含むツイートをクラスタリング分析することによって、白熱要素の時系列変動を観測する。分析を行う条件として、「金足農業」、「金足」を含むツイートを利用し、日毎にクラスタリングを行う。クラスタリング手法には k-means 法を用いる。k-means 法を採用した理由は、日毎に件数が大きく変わるデータ群の中でもこちらが指定したクラスタ数でのクラスタリングを行うことで、クラスタの変動を捉えやすくするためである。また、本研究において用いるデータの数は膨大であり、階層型クラスタリング手法では樹形図が巨大になり、結果が不明瞭になる恐れがあると考えられる。その点において非階層型である k-means 法が適切であると考えたからである。以上が、クラスタリング手法に k-means 法を選んだ理由である。

k-means 法を用いて「金足農業」に関するツイート群を日毎にクラスタリング分析した結果の一部を図 2 に示

す。なお、指定したクラスタ数は 4 個であり、各クラスタの重心に黒点を表示している。このクラスタ図の時系列変動から三つのことが観測できる。

- (C)「金足農業」が勝ち進むにつれ、注目度の高いツイート群によるクラスタが発生した。8/16 以前は試合日を除き、各クラスタの特徴は PN 値によるところが大きく、Attention 値による差は小さい。しかし、8/17 以降から徐々に Attention 値の大きいツイートが増え始め、8/18 以降は Attention 値の高いツイート群によるクラスタが発生している。
- (D)8/17 以降から大きく注目されているが、8/19 で一旦、白熱度が沈んでいる。この日は甲子園が休養日で試合はなく、全体のツイート数も減少している。このことから、クラスタに持続性はなく、その後のイベントの内容によっては消滅しうるとことを示唆している。
- (E)甲子園が終盤に近づくにつれて PN 値がポジティブ寄りのツイートが増えている。今回のツイートデータでは全日程において PN 値がネガティブ寄りのツイートの割合が多い。その中でも終盤はポジティブの値が大きいツイートが見られる。

3.3 白熱要素に対する群衆の感情推移

3.2 節において「金足農業」の時系列変動を観測した。続いて、注目が集まる過程で群衆の感情がどのように推移したかを観測する。感情の判定は「金足農業」を含むツイート文内の形容詞を抜き出してカウントし、日毎に統計を出すことによって行う。図 2 に表した同様の日にちである 8/16 から 8/21 までに観測した形容詞の上位 5 件を表 5 に示す。この結果から二つのことが分かる。

表 5 日毎の形容詞の統計

日付	形容詞
8/16	ほしい:7, 欲しい:7, 強い:7, いい:6, すごい:3
8/17	すごい:200, いい:164, 凄い:124, 素晴らしい:67, 嬉しい:62
8/18	すごい:505, いい:416, 素晴らしい:175, 良い:149, 嬉しい:126
8/19	いい:99, すごい:65, 凄い:41, 良い:28, 素晴らしい:23
8/20	いい:1134, すごい:896, 凄い:458, 良い:280, おお:248
8/21	いい:1810, すごい:1195, 素晴らしい:924, 強い:711, 凄い:666

表 6 日毎の固有名詞の変動

日付	固有名詞
8/16	横浜:32, 秋田:27, 吉田:24, 東北:21, 金農:12
8/17	神奈川:55, 滋賀:20, 板川:15, 吉田君:14, 佐々木希:14
8/18	生駒:56, 高橋:54, 走塁:51, 吉田君:51, 桑田:51
8/19	四乃森蒼紫:58, 桑田:38, 横浜:34, 下関:22, 桑田真澄:21
8/20	柿木:56, 吉田君:54, 海堂:52, 北海道:47, 聖秀:46
8/21	生駒:52, 佐々木:50, 湘北:50, 海堂:45, 菅原:45

(F)「金足農業」について好意的な感情を抱いた人が増えている。甲子園が開催された直後から 8/16 のように件数は少なくとも群衆の感情は好意的な方に傾いていた。そして、白熱度が高まった 8/17 以降は「嬉しい、素晴らしい、すごい」などの件数が多く集まっていることが分かる。

(G)終盤にマイナスな感情が発生している。表 5 には出ていないが、8/21 に 56 件、8/20 にも 47 件の「悪い」という単語が集まっている。このような直接的なマイナスワードが多く見られるのはこの二日間のみであった。

3.4 白熱要素を含むツイートに見られた固有名詞

「金足農業」のどのような要素が甲子園を白熱させる要因となったのかを解析するために「金足農業」を含むツイート内に現れる固有名詞を日毎にカウントし、統計を出す。なお、対象とする固有名詞はカウントが 60 件までのものとする。理由として、カウントが 60 件を超える固

有名詞には変動がなく、変動を観測するためのデータとして利用できないと判断したからである。日程は表 5 と同様にし、観測した固有名詞の上位 5 件を表 6 に示す。この結果から二つのことがわかる。

(H)人名が多くカウントされている。人名の中でも「生駒、佐々木(希)、桑田(真澄)」は芸能人の名前である。また、「吉田、板川、高橋、柿木、菅原」は「金足農業」と対戦校を含む選手名である。

(I)漫画内の高校名が出てきている。「海堂」は漫画「MAJOR」の主人公のライバル校の名前である。また、「聖秀」は主人公の所属する高校名である。この他にも、「湘北」は漫画「SLAM DUNK」の主人公の所属する高校名である。これらの単語を含んだツイートを解析した結果、ツイートを投稿した人たちは「金足農業」を主人公の高校と重ねていることが分かった。

4 考察

3 章でデータを観測して得た知見から「金足農業」の特性を考察する。また時系列変動に伴い、なぜ甲子園を白熱させたのか、どのように注目されたのか考察する。

4.1 ツイート占有率の時系列変動

3.1 節で「金足農業」が比較した他の要素より注目されたのは、そもそも試合を勝ち上がったという点もあるが、誰も予想していなかったという意外性によるものが大きいと考える。このような考えに至った理由は「大阪桐蔭」も勝ち進むにつれ注目を集めているが、最終的に「金足農業」は決勝戦で戦った「大阪桐蔭」より多く注目を集めたからである。「大阪桐蔭」はもともと優勝候補ということもあり、ある程度の注目はされていた一方で、「金足農業」は事前の注目はあまりなかった。この事前の注目度の違いが最終的な差を生んだと考える。また「猛暑関連」は甲子園開催前から一定の話題性を持ち続けており、開催期間中も低い水準ではあったが、その傾向は変わらなかった。そのため、他の要素が伸びるにつれてツイートの占有率が低下し、注目度はあまり伸びていない。

このことから、「金足農業」は他の要素と違い、事前に話題に上がっていなかった要素の台頭という意外性から多くの注目を集めたと考える。

4.2 クラスターの時系列変動

3.2 節のクラスタ図の時系列変動と 3.3 節の感情推移から「金足農業」がどのように注目されたかを考察する。まず、3.2 節の(C)から 8/18 以降は Attention 値の高いツイート群によるクラスタが発生していることが分かった。これは「金足農業」に対して関心を持つ人が増えるとともに「金足農業」のツイートに関心を持つ人が増えていることを示唆していると考えられる。理由として本研究において Attention 値は「いいね」やリツイートの数を基にしたもの

である。つまり「金足農業」のツイートに関心を持つ人が増えればそれだけ、「いいね」やリツイートが増えると考えたからである。また、ポジティブ寄りのツイートが増えていることも分かった。これは 3.3 節での表 5 で観測できたように甲子園が終盤に近づくにつれて、群衆が好意的な感情を抱いていることが主因である。つまり、群衆は「金足農業」に対して好意的な感情を持ちながら注目していたと考える。

このことから、「金足農業」は Twitter 上で強い拡散性を持ちながら、群衆から好意的な感情を向けられて注目されたと言える。

4.3 白熱要素の特性

3.3 節の感情推移と 3.4 節の固有名詞から「金足農業」の特性を考察する。3.3 節の(G)から 8/20 と 8/21 には「悪い」というマイナスワード現れたことが分かった。この単語を含むツイートを解析した結果、「悪い」というのは県立高校で頑張っている「金足農業」に対して私立高校で他県から選手を集めている「大阪桐蔭」はずるい等と揶揄する意見に対して否定的な意味を込めて書いた単語であるということが分かった。もし決勝戦が私立強豪校同士の試合であれば、そのような「大阪桐蔭」に対する批判的な意見は出ず、それに対する否定的な意見も出なかったと考えられる。また、3.4 節の(I)から群衆の中の一部は「金足農業」に対して、漫画の主人公たちの姿を重ねている。これは有名ではなかった「金足農業」が強豪校である「大阪桐蔭」に挑むという展開から漫画の主人公たちの姿を重ね、一種の英雄視をしているのではないかと考える。

このことから「金足農業」の持つ「県立高校」や「事前の注目度が低い」といった特性が甲子園を白熱させた要素となったと言える。また、漫画の主人公のような経緯を辿ったことも甲子園を白熱させた要素となった重要な特性の一つであると考えられる。しかし、その特性が「金足農業」を正義とし、「大阪桐蔭」を悪のような風潮を作り、間接的ではあるが、負の意見を生み出している。

5 将来展望

本研究で扱ったモデルは過去に起きたイベントであり、結果や反響があらかじめ分かっているものであった。本研究の将来的な発展は未来に起きるイベントに対する白熱要素の予測が考えられる。イベントを予測するという研究の性質上、モデルデータを用意することが必須である。つまり、扱うモデルとしては過去に数回行われているイベントが望ましい。そのデータを基にして白熱する要素の特徴やその要素が群衆に与える影響を予測することが主目的となると考える。また、3.2 節の(D)で述べたように白熱要素についてのクラスタの中で有用なクラスタが消滅することがあり得る。未来のイベントを予測

するにはこのようなイベント内の何らかの事象が白熱要素に与える影響も予測することが重要である。このような未来に起きるイベントを予測することは企業のマーケティング戦略において非常に重要な要素と言える。白熱要素をいち早く察知し、その要素を生かした販売戦略をとる。また、群衆の感情を分析し、市場の感情の流れに沿ったイメージ戦略を図る。これらの点から企業のマーケティング戦略において重要であると考えられる。

以上より、未来のイベントを予測することは重要であり、それを実現することが本研究の発展として望ましい。

6 おわりに

本研究では、甲子園をモデルとし、言語解析とクラスタリング分析を用いて様々な要素から白熱度の時系列変動を観測した。最初に白熱要素の推定を行い、他の要素の比較することで白熱要素を確認した。次いで、クラスタリング分析やその他データの時系列変動の観測を行った。それらの観測の結果から「金足農業」の特性やなぜ甲子園を白熱させたのか、どのように注目されたのかを考察した。甲子園などのスポーツの試合のイベントは性質上、勝ち進めば注目を集め、イベントを白熱させるのが自然である。しかし、本研究の観測と考察の結果、「金足農業」が甲子園を白熱させたのは勝敗だけでなく、群心理の推移の特徴によることが分析できた。

参考文献

- [1] Twitter Japan, <https://twitter.com/TwitterJP/status/23671036758958080>, 2017.
- [2] 今井美希, 榎美紀, 小口正人: 訪日外国人の趣向に合わせたイベント情報配信のためのソーシャルストリームの解析, 第 80 回全国大会講演論文集, Vol. 2018, No. 1, pp. 381-382, 2018.
- [3] 坂本朋鴻, 永井秀利, 中村貞吾: 時空間的局所性を考慮した twitter からのイベント情報抽出, 電気関係学会九州支部連合大会講演論文集, pp. 163-163, 2015.
- [4] Shen, C., Liu, F., Weng, F., et al.: A participant-based approach for event summarization using twitter streams, Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, pp. 1152-1162, 2013.
- [5] Ifrim, G., Shi, B., and Brigadir, I.: Event detection in twitter using aggressive filtering and hierarchical tweet clustering, Second Workshop on Social News on the Web (SNOW), Seoul, Korea, 8 April 2014. ACM, 2014.
- [6] 高村大也, 乾孝司, 奥村学: スピンモデルによる単語の感情極性抽出, 情報処理学会論文誌ジャーナル, Vol. 47, No. 02, pp. 627-637, 2006.