

## twitter を用いたリアルタイム実況システムの提案

三輪 将吾† 小高 知宏† 黒岩 丈介† 白井 治彦‡ 諏訪 いずみ†  
 †福井大学工学研究科 ‡福井大学工学部

## 1. はじめに

近年 twitter に代表されるソーシャルメディアの普及により個人がインターネット上に発信する情報量が増加してきている。リアルタイム性の高いソーシャルメディアでの情報は、どこで何が起きているのかを迅速に察知するための手掛かりとして利用することが可能であると考えられる [1]。

本研究では twitter でユーザーが投稿する実況 tweet を用いてプロサッカーの試合のリアルタイム実況を可能にするための手法を検討する。ユーザの tweet 投稿から実況文出力までをリアルタイムで行うことで迅速に試合での出来事を把握することができる。結果や内容を把握できるサッカーの試合を対象とし、実際の試合結果と比較することで実況文についての評価が可能であり、災害時等の状況把握のために活用するシステムとしての発展を期待することができる。今回は 実況 tweet 取得手法、また取得した tweet の特徴について考察する。また取得した tweet をどのように処理し、出力すべきか提案する。

## 2. 実現手法

本研究では twitter 上でユーザーの投稿する試合関連 tweet 取得し、プロサッカーの試合のリアルタイム実況を行うためのシステム作成を目指す。

そのため twitter からサッカー観戦者が試合時間内に得点や、選手交代などの出来事に関して発言している tweet をリアルタイムで取得する必要がある。twitter 上では日々、実況 tweet と呼ばれるような投稿が行われている。実況 tweet とは、スポーツの TV 放送や実際の試合を観覧しながらリアルタイムに試合内容について実況する投稿や内容に関する意見を含む tweet のことである [2]。本研究ではサッカーの試合に関する実況 tweet の取得から出力までをリアルタイムで行い、試合を通して状況を把握できるシステムを作成を目指す。

twitter 上から必要なデータの取得を行うために twitter API を利用する。本研究では twitter Streaming API の status/filter を用いて tweet とタイムスタンプ

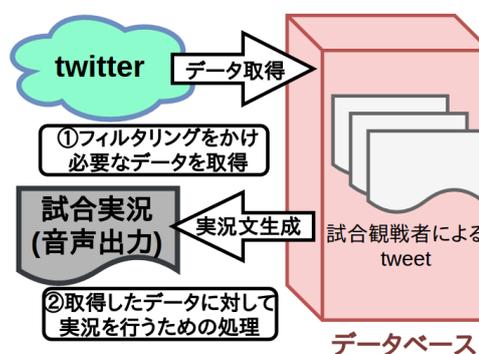


図 1: システム構成

を取得する。この API の特徴は、様々な条件をつけて、ストリーミング形式でリアルタイムに tweet の取得ができるという点である。そのため今回のようなイベント実況システムを構築するのに適した API である。この Streaming API を PHP 上にて実装し、必要なデータの取得を行う。実際に実況 tweet を取得するためのいくつかの条件を与えて tweet の取得を行い、必要な取得条件を検討する。

取得した実況 tweet はテキスト読み上げソフトを用いることで、音声による出力を可能にする。これにより手の離せない状況など、様々なシチュエーションで利用されることが期待できる。音声出力を可能にするために Open JTalk という日本語のテキスト読み上げエンジンを利用する。今回はこれを実装し、実況のリアルタイム性への影響について検証し、取得したデータに対してどのような処理が必要かを検討する。

本研究で目指す試合観戦者の実況 tweet 取得から音声出力を行うまでのシステムの一連の流れをについて図 1 に示す。

## 3. 実況 tweet の取得条件

試合観戦者の実況 tweet を取得するための条件として与えたものは次の通りである。

- 1 日本語 tweet のみ取得
- 2 #チーム名 を含む tweet の取得
- 3 リプライ、リツイートを排除
- 4 URL を含む tweet を排除

Proposal for realtime commentary system using twitter  
 †Shogo Miwa †Tomohiro Odaka †Josuke Kuroiwa  
 †Haruhiko Shirai †Izumi Suwa  
 †Graduate School of Engineering, University of Fukui  
 †Faculty of Engineering, University of Fukui

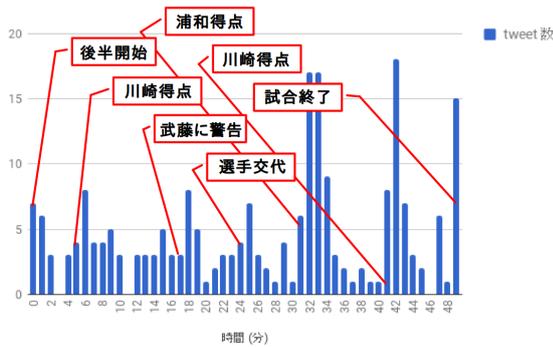


図 2: tweet 数の変位

本システムで対象となる tweet は日本語 tweet のみとなるため 1 の条件をかける。またマッチアップしている両チームのハッシュタグが付いている tweet のみ取得することで試合観戦者の tweet を取得する。リプライには試合に関する情報が乏しいと考え排除する。また、リツイートは複数同一内容の tweet が取得される可能性があるため排除する。URL 付きの tweet については、広告のような tweet が多いため除外する。これらの条件で取得した tweet はハッシュタグ以下を削除した形で、テキストファイルに逐次記述されるよう作成した。

#### 4. 実験

実際に 3 章で述べた条件で J リーグの試合について tweet の取得実験を行い、得られたデータの特徴について調査した。8 月 23 日に行われた浦和レッズと川崎フロンターレの試合を対象とし、前半の時間帯に 100 件、後半の時間帯に 230 件の tweet を取得した。図 2 は対象とした試合 (後半) の時間に取得した tweet 数の 1 分ごとの変位と、試合での出来事の関係を表すグラフを示すものである。得点や選手交代など試合中の重要な出来事が発生してから約 1 分間 tweet 数の上昇が見られた。実際に得点があった時間帯での出力例を図 3 に示す。このように武藤という選手が得点した事が把握できる内容となっていた。

また取得した実況 tweet の音声出力についての実験を行う。2 章で述べた日本語のテキスト読み上げエンジン Open JTalk を用いて、取得した実況 tweet をリアルタイム性を確保して音声出力できるかどうかについて検証を行う。音声出力の実験結果については当日示すものとする。

#### 5. 考察

試合観戦者の実況 tweet の取得を行うための条件を付けてデータの取得を行った。今回与えた条件で試合観戦者の実況 tweet を取得することが可能であった。

```

20:33:42:オフサイドっぽかったけど、武藤ゴール！ギリギリだった？(^^)/
20:33:39:武藤キック————(°▽°)————!!
20:33:35:きたあ——！むと——！
20:33:31:ラファもそろそろ不調から脱出してほしいな
20:33:28:武藤のゴールで1点返した！
20:33:07:武藤雄樹ゴール
20:33:07:武藤きめた！！！！オフサイドじゃない！
20:33:04:武藤ゴール！！アウェーゴール1点目
20:33:02:武藤キック————(°▽°)————!!
20:33:01:ナイス武藤！！
    
```

図 3: 得点時の出力例

今回与えた条件以外にも、不要な記号等を排除して取得することで必要な情報のみ取得できるようにする必要がある。取得した tweet についてはそのまま出力するだけでその試合での得点や選手交代といった重要な出来事を捉えることが可能であった。しかし、試合の出来事について触れられていない tweet や、感情的な tweet が多数見られた。このような不要と考えられる tweet を排除して出力するアルゴリズムが必要となってくると考えられる。今回対象とした試合は比較的注目度の低い試合であったため、データ量が膨大であった際、実況 tweet の制御を行う必要性も考えられる。具体的には重複した内容の tweet を出力させないようなアルゴリズムが必要となる。

音声出力では時間あたりに出力するテキスト量を制御しなければリアルタイム性が失われてしまうため、そういった点でも出力するテキスト量の制御が今後必要と考えられる。

#### 6. まとめ

災害時等の状況把握を行うシステムの作成を念頭に置いて、サッカーの試合観戦者の実況 tweet から試合実況を行うシステムの提案を行った。試合観戦者の実況 tweet を取得することで、その試合での重要な出来事を捉えることが可能であった。取得した実況 tweet はテキスト読み上げエンジンにより音声出力を行うことで様々な場面での利用が期待できる。今後は取得した tweet に対してどのような処理が必要となるか検証し実装することが課題である。

#### 参考文献

- [1] 齊藤 令, 寺田 実. Twitter の実況ツイートを利用したタイムライン上のネタバレ情報検知. 第 78 回全国大会講演論文集 2016.1 (2016): 539-540.
- [2] 小林 尊志, 野田 雅文, 出口 大輔, 高橋 友和, 井手 一郎, 村瀬 洋. Twitter の実況書き込みを利用したスポーツ映像の要約. 電子情報通信学会技術研究報告. IE, 画像工学, vol.110, no.456, pp.165-169, 2011-02-28.