

異メディア間双方向情報支援システムの提案

細川 侑士†

安井 豪基†

河合 由起子†

秋山豊和†

若宮 翔子‡

角谷 和俊‡

†京都産業大学

‡兵庫県立大学

1 はじめに

SNS の普及により、実世界で発生したイベントに対して、発生した場所から携帯端末を用いてリアルタイムに情報を発信するユーザーが増加している。これら大量のツイート情報を分析する研究が活発に行われており、情報を集約し可視化する研究 [1] や、災害情報に対する集約可視化 [2]、イベント状況の推移を可視化する研究 [3] などがあげられる。しかし、これら既存研究では大量情報の大局的な状況把握は効率的にできるが、リアルタイムにツイートユーザーとそれ意外のメディアを利用してユーザー (Web, TV, Facebook 等) 間での情報発信・取得は困難であり、網羅性が低い。

そこで、本研究では、ツイートユーザーと、Web 閲覧ユーザー間でリアルタイムに情報支援可能なシステムを提案する。具体的には、発信されたツイートの実空間の位置とその内容から場所の特徴を抽出し、また、Web コンテンツの内容からも場所の特徴を抽出する。抽出したこれらの場所情報を対応付けることで、発信したツイートユーザーと Web 閲覧ユーザーとを結びつけ、それらのリアルタイム双方向通信を可能にする。

本論文では、位置情報に基づくツイートおよび Web ページ関連付け手法ならびにリアルタイム通信システム構築について述べ、実装したシステム動作を考察する。

2 提案手法

本研究は、ツイートユーザーと Web 閲覧ユーザーが位置情報に基づきリアルタイムに通信可能なシステムの構築を目指す。図 1 にシステムの概要を示す。ツイートを発信すると、ツイート閲覧者だけでなく、ツイートの発生場所の位置情報と内容に基づき、場所名 (形態素解析結果から抽出した地名) を抽出する。また、Web ページからも同様に場所名を抽出し、tf・idf 値の高い単語を検出する。これら抽出した場所名に基づき、関連するツイートと Web ページを対応付けることで、その

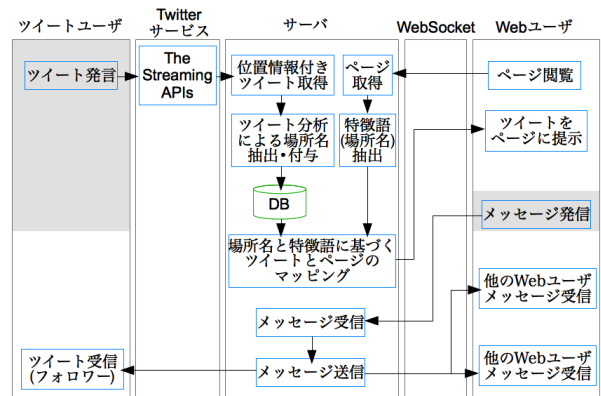


図 1: システム構成図

ページ上にツイートがリアルタイムに提示される。Web 閲覧ユーザーにとっては提示されたツイートを閲覧することで、場所に関する現状把握ができる。また、Web 閲覧ユーザーは提示されているツイートユーザーに対してメッセージを送信できる。送信されたメッセージはツイートユーザーのツイート画面に提示される。なお、ツイートユーザーは本サーバからメッセージを受信する際には、本サービスのフォロワーとなっている必要がある^{*1}、Web 閲覧ユーザーは提案システムとなるアドオンを用いる必要がある。

2.1 ストリーミングツイートデータ取得

本論文では、位置情報に基づく問合せを目的としており、ページとツイートを位置情報に基づき関連付ける。そのためまず、指定地域から重複を除いた緯度経度情報を含むストリーミングツイートを The Streaming APIs を用いて取得する。次に、取得した各ツイートの緯度経度を場所名へ変換する。同時にツイートを形態素解析し、名詞 (地名) となる単語も取得する。

以上より、ツイートユーザー id, アイコン画像 URL, 緯度, 経度, 場所名, ツイート内容, 取得時刻を一定時間管理する。

A proposal of a cross-media interactive communication system for SNS and Web users
 †Yuji Hosokawa †Gouki Yasui †Yukiko Kawai †Toyokazu Akiyama
 ‡Shoko Wakamiya ‡Kazutoshi Sumiya
 †Kyoto Sangyo University
 ‡University of Hyogo

^{*1}<http://twitter.com/@RtQAService>

2.2 ツイートの緯度経度と内容に基づく場所名抽出

前節より取得したストリーミングツイートに対して位置情報に基づいた内容判定を行い、ページと関連付ける。ツイートが発信された場所名と関連するかをツイートの内容から判定することで、ツイート発生場所と関係性の低いツイートの除去を行う。

位置情報に基づいたツイート内容判定法は、まず、取得したツイート t の位置情報より、半径 d 内に存在する一定時間内のツイート n 個を取得する。次に、下記の式よりツイート t に出現する各単語 i のツイートに出現する頻度を抽出し、その平均値を算出する。

$$\sum_{i=1}^m \frac{\text{単語 } i \text{ が出現するツイート数}}{n} \times \frac{1}{m}$$

m はツイート t に出現する単語総数である。最後に、閾値以上のツイート t を位置情報に基づいたツイートとして決定する。

2.3 場所名に基づくページとツイートの対応付け

Web ユーザの Web ページ閲覧から Web ページを取得し、そのページの特徴語(場所名)を抽出する。また、ツイートユーザの位置情報付きツイートを The Streaming APIs を用いて取得し、そのツイートを分析することによる場所名の抽出および、付与を行う。なお、DB には取得したツイートおよび抽出した場所名を格納する。これらのツイートと Web ページを場所名に基づきマッチングし、同一場所名同士を対応付ける。

3 実装

ストリーミングツイート取得は、指定地域から重複を除いた緯度経度情報を含むストリーミングツイートを The Streaming APIs version 1.1^{*1}を用いて取得した。また、1 度以上異なる南西および北東を指定することで囲まれた矩形領域のストリーミングツイートを取得できる。取得したストリーミングツイートの緯度経度から、Google Place API version 3^{*2}を用いて、半径 dm 以内の場所名を取得した。

図2に実装したプロトタイプによる Web 閲覧ユーザ側のインタフェースを示す。スカイツリーの公式サイトトップページと、ストリーミングツイート(トップページの左横)の提示結果である。ユーザはスカイツリー近辺でリアルタイムに発信されたスカイツリーに関するツイート内容を閲覧することが可能である。なお、これらツイートを各ページごとにログファイルとして保存することで、過去のツイートも閲覧可能であ

^{*1}<https://dev.twitter.com/docs/streaming-apis>

^{*2}<https://developers.google.com/places/>

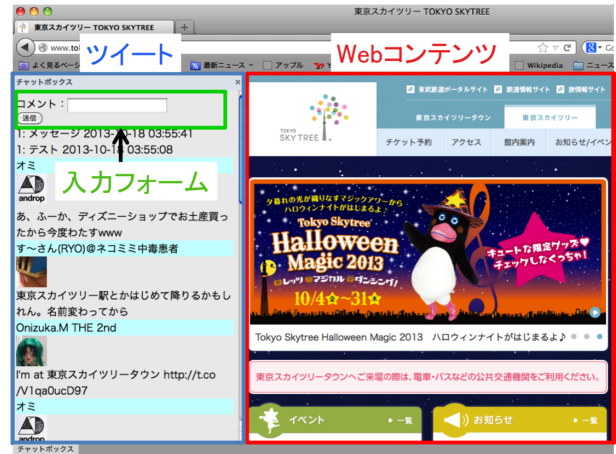


図2: Web 閲覧インタフェース

る。Web 閲覧ユーザは、入力ボックスにメッセージを入力し、他の Web 閲覧ユーザに一斉送信可能である。また、任意のツイートを選択することで、特定のツイートユーザへのみメッセージ送信可能である。プロトタイプでのストリーミングツイートのメッセージ提示および Web 閲覧ユーザのメッセージ提示にかかる時間は数秒程度であった。これにより、少数のユーザ間ではあるが、リアルタイムに通信できることを確認できた。

4 まとめ

本研究では、ツイートユーザと、発信されたツイートの場所と関連する Web ページを閲覧しているユーザ間におけるリアルタイム双方向情報支援システムを構築し、シームレスな情報発信・取得が可能であることを確認した。今後、さらに大量のツイート・Web 閲覧ユーザに対する検証ならびに TV 等の他メディアへの応用を検討する。

謝辞

本研究の一部は、総務省戦略的情報通信研究開発制度(SCOPE)の一環として実装されたものである。ここに記して謝意を表す。

参考文献

- [1] Leah Hoffmann, Looking back at big data Communications of the ACM, Vol.56 Issue.4, pp.21-23, April 2013.
- [2] 小出明弘, 斉藤和巳, 鳥海不二夫, 風間一洋, 松尾真人, デンドログラムを用いた震災時における Mention・Hashtag ツイートの利用傾向分析. WebDB フォーラム 2012
- [3] Kyoung-Sook Kim, Ryong Lee, Koji Zetsu: mTrend: discovery of topic movements on geo-microblogging messages. GIS 2011: 529-532