

防災関連情報集約システムにおける 関連表示を用いた情報見逃し防止手法の提案

安田 大誠[†]吉野 孝[†][†]和歌山大学

1 はじめに

災害時の情報収集は、災害の状況把握を目的に行われる。収集する情報には、災害規模、安否情報、警報や避難関連情報、交通運行状況などが挙げられる。災害時における情報発信手段は多様化しており、災害時に人々が利用する情報収集手段として、Web サービスがマスコミに次いで利用されている [1]。国や地方自治体は、Web サイトを通じて、様々な防災関連情報を提供しているが、近年では地方自治体が Twitter を活用した防災関連情報の収集・発信を行っている [2]。このように、インターネット上で提供される防災関連情報には様々な形態がある。

我々はこれまでに、分散した防災関連情報を集約し、個人の属性によって必要な情報をカスタマイズできる防災関連情報集約システム (以下、先行研究) を開発してきた [3]。しかし、先行研究においてはシステム側がユーザに表示コンテンツを選ばせている。ゆえに、コンテンツの自由度が低く、突発的な事例に対する情報収集が困難であった。例として、2017 年 10 月の台風 21 号の影響により樽井駅～尾崎駅間¹⁾ で橋脚が陥没し、線路が歪んだ南海本線を挙げる。この時、利用者にバスの振り替え輸送の案内がなされたが、バスの運行に関する情報は公式ページ上で即座に公開されず、利用者間でニュースや Twitter を利用して情報を収集するしかなかった。また、この出来事から、「南海本線」という語句は災害と関連する語句になったと考えられる。

そこで、本稿ではインターネット上で提供される防災関連情報から、従来災害と関連することのなかった語句を抽出・表示する。さらに、ユーザに検索を促すことで、大量に散在する情報の中での重要な情報が見逃されないようになることを目指す。

2 関連研究

橋本らは新聞記事に対してテキストマイニングの手法を応用し、定量的な解析システムを開発した [4]。これは、得られた文書集合に対して階層的クラスタリングを施し、クラスタ間の語彙使用の類似性に基づく構造化を行う。そして、クラスタを要約するキーワード及び関係主体を抽出することによって、内容を効率的かつ経済的に捉えられることを示した。本研究では、収集した記事の内容把握だけに留まらず、関連する情報をユーザが認知するところまでを目的としている。

佐藤らは危機管理分野における言語資料からのキー

ワード自動検出手法を提案している [5]。これは、発信・収集された言語資料から、メタデータとして付与すべき情報内容を表すキーワードを自動的・客観的に生成する。この手法を 2004 年の新潟県中越地震に関するウェブニュースに適用した結果、発災から発信されたニュース記事の特徴変化の概況をキーワードを単位で捉えることができた。本研究では、言語資料ではなくリアルタイム性の高い情報源で語句抽出を行う点、およびユーザの情報見逃し防止に重きを置いている点が異なる。

3 情報見逃し防止手法

本研究はインターネット上で提供される防災関連情報から、従来災害と関連することのなかった語句を抽出・表示する。特に、ユーザに検索を促すことで、大量に散在する情報の中での重要な情報が見逃されないようになることを目指す。本研究の具体的な内容を以下に示す。

(1) ニュース記事の収集

ニュース記事を複数新聞社から RSS 収集する。収集対象は、「社会」または「地域」カテゴリの記事とする。これは、社会的・地域的な記事が多く、防災情報が含まれる可能性が高いと判断したためである。

(2) 形態素解析による単語の抽出

取得した記事に対して、形態素解析を行い、一般名詞と固有名詞を抽出する。本研究では、形態素解析器として MeCab²⁾ を用いる。MeCab の辞書には mecab-ipadic-NEologd を用いた。

(3) 記事のトピック解析

記事毎に抽出された単語をまとめ、トピック解析を行う。トピックの解析には潜在的ディリクレ配分法 (Latent Dirichlet Allocation, 以下 LDA) [6] と呼ばれる確率的トピックモデルを用いる。LDA は 1 つの文書に対して複数のトピックが存在し、さらに、1 つのトピックに対して複数の単語が存在すると想定したトピックモデルである。本研究では、収集した記事の単語の集合を用いて LDA モデルを作成し、記事のトピック分布と、トピックの単語分布を求める。

(4) トピック毎の画面表示

(3) でトピック毎に分類された単語をシステムの画面に表示する。これにより、画面出力された単語の中に「台風」や「土砂崩れ」、「床下浸水」などの災害語句が存在すれば、同一トピックに存在する単語は災害関連語句とみなすことができる。

Proposal of Overlooked Prevention Method Using Providing Relevant Information in Disaster Prevention Related Information Aggregation System

Taisei Yasuda[†] Takashi Yoshino[†]

[†]Wakayama University

¹⁾ 南海電気鉄道南海本線の駅。樽井駅：大阪府泉南市樽井 5 丁目 41 番 1 号、尾崎駅：大阪府阪南市尾崎町 95 番地の 1

²⁾ MeCab : <http://taku910.github.io/mecab/>

(5) 防災情報収集用検索ツールとしての利用

ユーザに、表示された単語による Web 検索を促し、個々の関心を持った防災関連情報の詳細を認知させる。

4 防災関連情報集約システム

4.1 システム概要

本システムは、Web ブラウザ上で動作するアプリケーションとして、ユーザの防災関連情報に対する見逃し防止を行うことを想定する。また、先行研究と同じく、平常時と非常時の両方の使用を想定する。平常時には、同様の災害が身近な地域で起こった場合のために、ユーザが遠く離れた地域の災害情報を認知し予備知識の蓄積・事前学習に用いる。また、それに伴い防災意識の向上を図る。非常時には、情報収集の補助システムとして防災行動に役立てる。

4.2 システム構成

システム構成を図 1 に示す。本システムは、各ユーザの Web クライアントと、情報の取得や解析、保存を行うサーバから構成される。サーバ側では、一定時間ごとにニュース記事収集と、単語の抽出、トピック解析、データベースへの保存を行う。データベースは MongoDB³⁾ を利用した。そして、ユーザの要求に沿った情報をクライアント側に送信し、Web ブラウザに表示する。図 2 にシステムの画面例を示す。ユーザは、表示された単語の中から検索したいものを選択し、ドラッグアンドドロップで入力と検索を行う。

4.3 システムの機能

システムの機能として「災害関連語句のリスト表示機能」と「検索補助機能」がある。以下に具体的な内容を示す。

災害関連語句のリスト表示機能

サーバ上で、3 章 (1)~(3) の項目を定期的に行う。保存・更新された災害関連語句は、ユーザのアクセス時に参照され、リスト形式でブラウザに表示される。

検索補助機能

リスト表示された語句を上部の検索フォームにドラッグアンドドロップすると、Web での検索が行われる。この機能により、ユーザがその語句を見るだけに留まらず、詳細情報のスムーズな収集を支援することを想定している。また、検索結果のページにおいて更なる関連情報をユーザが閲覧でき、より情報見逃しの機会が減少する。

5 おわりに

本稿では、防災関連情報集約システムにおける関連表示を用いた見逃し防止手法の提案を行った。さらに、提案手法を用いたシステム構築を行い、概要と機能について述べた。収集したニュース記事から従来災害と関連することのなかった語句を抽出・表示し、ユーザの検索を促すことで、ユーザが防災関連情報を見逃すことを防ぎ、自主的な情報収集のきっかけとなることを目指している。今後はシステムの有用性の評価実験を行う。

³⁾ MongoDB : <https://www.mongodb.com/>

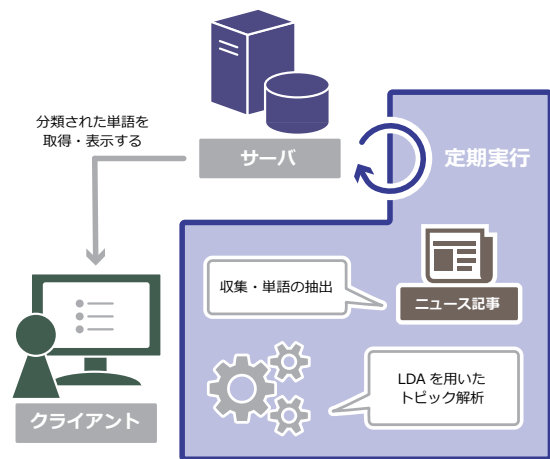


図 1: システム構成



図 2: システムの画面例

謝辞

本研究の一部は、JSPS 科研費基盤研究 (A) (25242037) および和歌山大学平成 24-27 年度独創的研究支援プロジェクトの補助を受けた。

参考文献

- [1] 本條晴一郎, 遊橋裕泰: 災害に強い情報社会—東日本大震災とモバイル・コミュニケーション—, NTT 出版株式会社 (2013).
- [2] 櫻井美穂子, 國領二郎: 自治体 ICT ネットワーキング—3.11 後の災害対応・情報発信・教育・地域活性化—, 慶應義塾大学出版会株式会社 (2012).
- [3] 吉野孝, 田藤千弘, 谷岡遼太: Web 防災袋: カスタマイズ可能な防災関連情報集約システム, 情報処理学会, マルチメディア, 分散, 協調とモバイル (DICOM2016) シンポジウム, pp.480-489 (2016).
- [4] 橋本泰一, 村上浩司, 乾孝司, 内海和夫, 石川正道: 文書クラスタリングによるトピック抽出および課題発見, 社会技術研究論文集, Vol.5, pp.216-226 (2008).
- [5] 佐藤翔輔, 林春男, 牧紀男, 井ノ口宗成: TFIDF/TF 指標を用いた危機管理分野における言語資料体からのキーワード自動検出手法の開発—2004 年新潟県中越地震災害を取り上げたウェブニュースへの適用事例—, 地域安全学会論文集, No.8, pp.367-376 (2006).
- [6] David M. Blei, Andrew Y. Ng, Michael I. Jordan : Latent Dirichlet Allocation, Journal of Machine Learning Research, 3, pp.993-1022 (2003).