2ZC-3

ニュース討論SNSの提案と実装、評価

喜安 伸彦† 角田 博保†
電気通信大学大学院 情報理工学研究科 情報・通信工学専攻†

1 はじめに

本研究では、Web上のニュースとそれについて書かれた記事を同一の話題（トピック）毎に自動的に分類、整頓を行い、それを複数ユーザで共有し議論を行うことでニュースへの理解を深めることを目的としたSNSを実装し、評価する。

現在、メディアの大小に関わらず多くのニュース配信サービスが存在している。しかし、提供される記事は記者により異なる視点を持って書かれているため、その記事を読んで受ける印象や得られる知見は記事によって異なってくる。情報が偏らないようにするためには読者は同様のニュースに対する別メディアの記事を読むなどする必要があるが、それには非常に手間がかかるのが現状である。

またニュースは第一報が報道された後、状況に変化があるたびに新しい記事が作成されるが、日々更新されるニュースに対して第一報から最新の報道までを俯瞰的に把握するのに適した仕組みが無いのが現状である。

本研究では時系列と言語処理を軸にニュース記事をトピック毎に分類、整頓し、その上でユーザ間が議論を行うことでニュースに対する理解を深め知見を拡げることを目的としたシステムを提案、実装、評価する。

2 研究背景

メディアから提供されるニュースは記者の持つ視点から書かれており、その書かれ方により読者のニュースに対する受け取り方は変わってくる。偏った視点を持たないためには記事について自分で調べたり、他のニュースソースによる記事を見比べる必要があるが、それには非常に手間がかかる。

また現在、大手ニュースメディアが提供しているニュースサイトの各記事にはFacebookやTwitterなどの上に記事の感想を書き込めるようなリンク（ソーシャルリンク）が埋め込まれていたり、また記事自体にコメント欄が設けられていたりするが、それらの議論の場は同様の話題であっても新しく更新される記事は新たなWebページに書かれるため、それに伴いソーシャルリンクやコメント欄における議論は断絶され、継続した議論にならないのが現状である。

さらに多くの場合、ニュースに含まれる問題は複数のトピックにまたがっており、１つのトピックのみに注目していては問題の全体像を把握しづらい。トピックを１つの文脈として追いつつ、他の話題や問題と結びつけて考えるための仕組みが現状では十分でない。

3 システム概要

以上の背景を踏まえ、本研究ではニュースへの興味と理解が深まることを目的とし、メディアから随時提供されるニュース記事を系統立てて分類し、それに対する議論を支援するシステムを実現する。トピックと時系列で整理されたニュース記事を俯瞰することでニュースの全体像を理解し、併設された掲示板で議論を行うことで記事間に存在するであろう記者の視点の違いを緩和する(図1)。また議論の中で現出するキーワードを基に他のトピックへのリンクを生成し、ユーザが他のトピックを含めた知見を広めることを支援する。

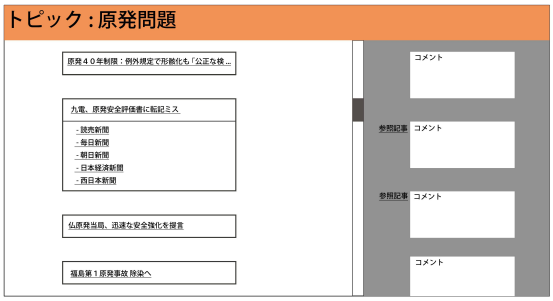

図 1:システム概観
分類されたトピックと掲示板

ニュース記事の取得にはGoogle News[1]を使用した。本サービスはあらかじめ同一の内容を扱う記事を分類し、それぞれをひとつのニュースとして複数のニュースをまとめてRSSによる配信を行っている。本システムではGoogle Newsから配信されるニュースの集合を話題（トピック）毎に分類し、ユーザに議論の場として提示する。本システムにおけるトピック、ニュース、記事の関係は図２の通りである。

ユーザはシステムから提示されるトピックから興味のあるトピックを選択し、以後そのトピックをtwitter[2]のタイムラインのように閲覧することができる。本稿ではこれをtwitterと同様に「トピックをフォローする」と呼ぶことにする。

Presenting an Overview of News Articles Classified by Topic and Supporting to Discuss on it.

† Nobuhiko KIYASU, and Hiroyasu KAKUDA, Department of Communication Engineering and Informatics, Graduate School of Informatics and Engineering, The University of Electro-Communications

Copyright ©2012 Information Processing Society of Japan.
All Rights Reserved.

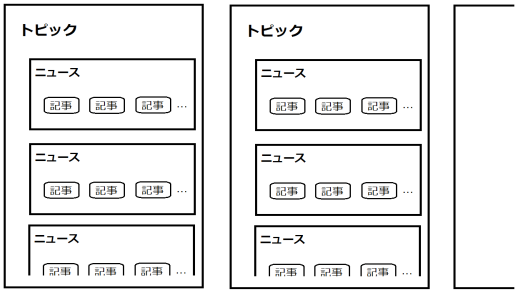

図 2: トピック、ニュース、記事の関係

3.1 ニュース記事の分類

ニュース記事の分類について、Uramoto ら[3]はニュースタイトルの形態素解析の結果から名詞と動詞のみを記事同士の類似度を測定するために使用した。本研究でも同様に形態素を単語ベクトルとして扱い、ニュース記事の分類を行うことにする。

分類されたニュース記事は「トピック」として扱われ、それぞれに掲示板が併設される。ユーザはその掲示板上で議論を行う。

またトピック毎に分類されたニュースについて、後にトピックを俯瞰する際に内容が重複するようなニュースはトピック全体の理解の妨げとなるため、定期的にシステムにより整理され、非表示となる。

また、この項で算出された記事タイトル内の形態素に対する重み付けは、次項のトピック間へのリンクを作成する際に再度使用する。

3.2 トピック間の関連づけ

ユーザのニュースに対する理解を深め、また知見を広める支援を行う。 ユーザは各トピックに設置された掲示板上で議論を行うが、あるトピック内の掲示板に、他のトピックと関連があると思われる（他のトピックにとって重みの大きい形態素を持つ）書き込みがされた時、その書き込みには関連のあるトピックへのリンクが自動的に付与される。(図 3)これによりユーザがフォローしていないトピックに対するナビゲーションが行われ、ニュースに対する広い知見を得るためのサポートを行う。

3.3 議論の支援

それぞれのトピック内掲示板における書き込みについて、トピック内における特徴量の変化を利用して書き込まれた内容とそれに関連すると思われる記事、および他の書き込みを前項と同様に関連付ける。

これにより質問とそれに対する回答の重複を防ぐことでユーザの手間を減らし、また該当する書き込みと関係の無いユーザに対しても、掲示板を整理し可読性を高めることで、その掲示板の含まれるトピックについて理解をする助けとなることを目的とする。

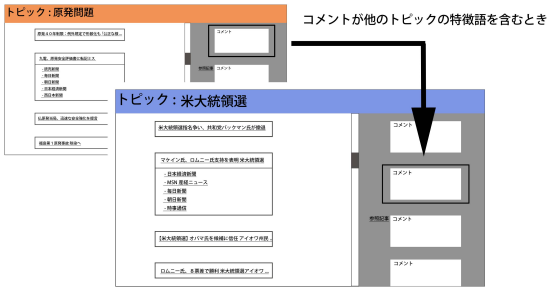

図 3:書き込み内容に関係のあるトピックへのリンク

4 実装

システムの実装には、Google News からのニュースの取得、解析とトピック毎への分類、解析に用いるパラメータの保存に PHP を用い、ユーザに提示し、閲覧、書き込みといったユーザが操作する部分を javascript を用いて実装した。ニュースの解析のための形態素解析には mecab[4]を利用した。

5 評価

実験として本システムを実際にユーザが使用し、インタビューを通じてニュース記事の分類の精度について、また議論とトピック間のリンクによるニュースの理解における貢献度について評価する。

参考文献

[1] Google ニュース, Google Inc, <http://news.google.co.jp>.

[2] Twitter, Twitter Inc, <https://twitter.com>.

[3] Naohiko Uramoto, Koichi Takeda. A Method for Relating Mutiple Newspaper Articles by Using Graphs , and Its Application to Webcasting, Proc. *the 17th international conference on Conputational linguistics* .

[4] Taku Kudo, 2008, Mecab: Yet Another Part-of-Speech and Morphological Analyzer, <http://mecab.sourceforge.net/>.

Copyright ©2012 Information Processing Society of Japan.
All Rights Reserved.