

Twitter への擬似犯罪発言抑止における リスト組み合わせ方式の提案

【 研究系卒研 】

0823005 青柳 翔

(指導教員 速水 治夫 教授)

1. はじめに

Twitter は 2009 年に流行し、今もユーザーを増やしているサービスである。これはツイートと呼ばれる短いテキストを投稿するサービスである^[1]。

この Twitter におけるテキストの投稿は、文字数以外に制限がない。近年ではその仕様から、いたずらで犯罪に関連する投稿を行うユーザーが後を絶たない^[2]。この投稿によって、炎上や、周囲からの評判を悪くするという事態が起きている。

そこで、本研究では Twitter において犯罪に関連する投稿を規制することで、ユーザーが自身の安全を守ることが可能となるシステムを提案する。

2. 試作システム概要

本システムは、PHP、MySQL、Apache、TwitterAPI、形態素解析エンジンの MeCab を利用し、必要に応じてユーザーのツイートに規制をかけるシステムである。この規制はツイートを形態素解析した結果と、単語がリストで登録されたデータベースを比較することで行われる。また、システム構成図を図 1 に示す。

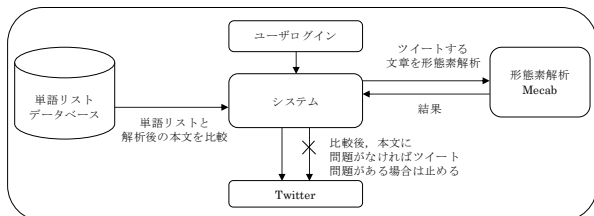


図 1 システム構成

2.1 規制手段

規制する手段には、単語 1 つで規制するブラックリスト方式と、単語同士の一定の組み合わせで規制するオートマトン方式がある。しかし、どちらの方法にも精度や、メンテナンスにおける手間など、欠点がある。そこで本研究ではブラックリスト方式を改良したリスト組み合わせ方式を提案した。

2.2 リスト組み合わせ方式

この方式は、ブラックリスト方式に補助のリストを追加することで文章を分類し、必要に応じて規制する手段である。これは一定の単語同士の組み合わせでなく、分類された複数のリストの単語の組み合わせで規制するため、オートマトン方式と比較する

と精度はやや下がるものの、幅広い文章を比較できる。新語の追加については、リストに種類別に追加してだけでリスト間の組み合わせのパターンが増えるため、容易に行える。本システムでは、精度が高く、メンテナンスも容易であるリスト組み合わせ方式を用いる。

2.3 システム動作

このシステムは OAuth 認証をし、その認証後にアプリケーションに許可を出すことで利用できる。主な機能であるツイートの規制は、システムからツイートをすることで、リスト組み合わせ方式により必要に応じて動作する。

3. 結果

Twitter を利用している 6 名にシステムを利用してもらい、5 段階評価と自由回答の評価実験を行った。以下に評価実験の結果の一部を抜粋して表 1 に示す。

表 1 評価結果

評価内容	1	2	3	4	5
犯罪に関連するようなツイートをシステムが止めてくれたか	0人	1人	3人	1人	1人
犯罪に関連しない通常のツイートを発する際に弊害はないか	0人	0人	0人	2人	4人

評価実験の結果、「犯罪に関連するようなツイートをシステムが止めてくれたか」の項目には平均 3.7 点、「犯罪に関連しない通常のツイートを発する際に弊害はないか」の項目には平均 4.7 点と高い評価を得ることができた。

4. おわりに

評価結果から、システムを利用することで、通常のツイートに弊害を与えずに危険なツイートを止めることができた。しかし、クライアント機能に関しては、改善の余地がある。今後はシステムの改善とさらなる機能の追加を行いたい。

参考文献

- [1] 今更聞けない Twitter の常識: Twitter とは 国内で再流行、一般化の兆しも
<http://www.itmedia.co.jp/news/articles/0907/28/news011.html>
- [2] なんちゃって鶯 半世紀 recollection 「バカ発見機」 Twitter の暴露・炎上騒動まとめ
<http://uguisu.skr.jp/recollection/twitter2.html>