

SNS 投稿に含まれる個人情報指摘による情報リテラシー教育 A method of learning support for information literacy by the notice of personal information in SNS messages

児玉 周平 我妻 了 井上 拳斗 岩田 一
Shuhei Kodama Ryou Azuma Kento Inoue Hajime Iwata

1. はじめに

近年、スマートフォンなどの普及により Social Networking Service (以下 SNS) の利用率は年々増加傾向にある。SNS とは、趣味や嗜好などに共通の観点を持つ人達同士や、これまでの様々な人間関係をネットワーク上で結びつける会員制のサービスのことであり、Web サイトや専用アプリケーションから閲覧や投稿ができることも、広く利用される要因となっている。

SNS 利用者の年齢層は幅広く、2016年の内閣府調査[1]によると青少年のスマートフォン・携帯電話の所持利用状況は平成 22 年から平成 24 年に約 5.9%の増加している。しかし、SNS による人間関係の結びつきが強くなる一方で、SNS に起因するトラブルも多くなっている。SNS は気軽に使用できる反面、SNS 利用時に注意すべき情報モラルやマナーなど、学校や家庭で期待される情報リテラシー教育は追いついておらず、十分ではない。

そこで本研究では、主に青少年ユーザを対象として、これから SNS に投稿する発言の中および、過去の発言において不適切な発言や個人情報の特定に至る内容など、各種の問題点を検出して注意を促すシステムを提案する。学校の授業時間などでは十分な理解が難しい SNS の情報リテラシー教育について、実際に SNS を利用する中で学び学習することができるようになる。

これまでに、発言テキストについて形態素解析を行い危険性の有無を判断することと、発言に含まれる画像へのリンクを検出しユーザに注意をうながすシステムの提案を行った[2]。本研究では、発言中に含まれる不注意な個人情報の発信に関しても注意を促す手法を提案する。

2. SNS 利用の現状と問題点

SNS は広く利用されているが、提供されるサービスの内容によって、使用する年齢層に違いがある。日本で広く使用されている代表的な SNS として、Twitter・Facebook・LINE が挙げられる。Twitter の特徴として 140 文字の投稿制限により気軽に短文を発言できる、リツイートにより簡単に発言を共有することができる、などが挙げられる。利用者層は主に 10 代から 20 代が多い。次に Facebook は実名登録制であるが、公開範囲を設定することができる。LINE は登録した相手同士とのチャットや通話が主な機能であり、登録の際には電話番号認証を必要とされる。主にユーザの連絡帳に登録されているユーザなど、狭いコミュニティで使用される。

このように SNS は広く用いられているが、ユーザがインターネット上に不適切または不謹慎な発言を投稿したこと

により、インターネット上で注目が集まる、炎上という現象が問題となっている。発言が気軽にできることから、現実世界における自分が認識することが可能な範囲内でのみ発言が見られているとの判断により不適切な発言を行いやすい傾向にある。

炎上が拡散するプロセスとして、まず不適切または不謹慎な発言が無いかを探し出す人により、リツイートによる拡散や、掲示板およびまとめサイトなどに問題となる発言が取り上げられる。コミュニティサイトなどで炎上が話題となり、炎上を起こしたユーザの過去投稿や他のユーザとの発言のやりとりを追跡することにより、個人情報となる写真などが見つかると個人情報の特定へとつながる。

ユーザの意思に反して、第三者による故意、また、過失によって個人情報が渡ることにより、発言の内容がもとで本人の本名や住所、学校名会社名などの個人情報だけでなく、関連する友人や本来関係の無い第三者の情報まで流出する事態になることもある。その結果、関係の無い第三者が炎上に巻き込まれることもあり、個人情報を安易に広めないことの重要度は増している。しかし、不特定多数の人物に閲覧されている可能性があるのにも関わらず、実名を使用している割合が全体の 3 割強存在する[3]。

3. システム構成

本研究における SNS 投稿に対する注意を促すシステムの構成を図 1 に示す。本システムの対象とする SNS は Twitter とした。理由として、Twitter にはリツイート機能により発言が共有されるまでの時間が非常に短時間であること、短文の投稿が気軽にできることから実際のトラブル事例が多数報告されていること、社会経験が浅く情報リテラシー教育が特に重要となる 10 代から 20 代が主な利用者層であること、等が挙げられる。また、多くの環境で動作が可能であることから、Java を用いて実装を行った。

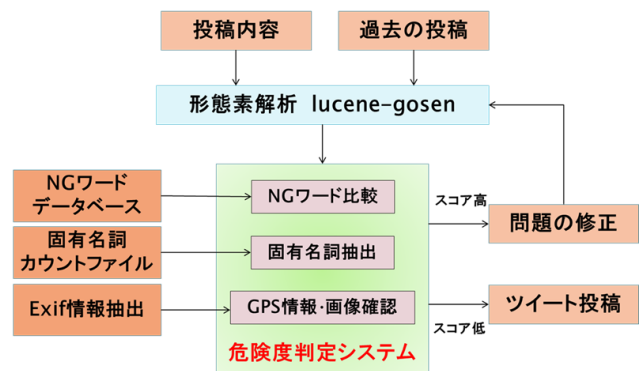


図 1. システム構成図

† 神奈川工科大学 Kanagawa Institute of Technology

本システムは、Twitter に対する投稿を行うためのクライアント機能と、投稿内容についての危険度を判定する機能からなる。本システム内でユーザのアカウント情報を保持することにより、新規の投稿だけでなく本システムを使用していない過去の発言に関しても、合わせて危険度の判定を行うことができる。

本システムにおける投稿内容の危険度判定は、NG ワード比較、固有名詞抽出、GPS 情報および画像情報、の 3 項目について評価する。

3.1.1 NG ワード比較

新規の発言および過去の発言のテキストデータに対して形態素解析を行い、語句、語句の活用前の基本形、語句の品詞の組み合わせとして、システムに読み込む。形態素解析には Java で動作するライブラリである、`lucene-gosen` を使用する[4]。

形態素解析後の発言内容の基本形と、システムの内部に保存されている NG ワードリストを、NG ワードリストと投稿内容の文字列を一致するものがあれば、その単語数に応じて危険度が高いと判定する。

3.1.2 固有名詞抽出

発言に含まれる固有名詞をカウントする。固有名詞は人名や地域などが対象となるため、固有名詞が文章に含まれている場合は個人情報が含まれる可能性が高い。

本システムでは、過去の発言も対象とすることで、固有名詞が単独で出現する場合は特定が難しくても、複数の固有名詞が連続で出現する場合は個人を特定する要素となる。そのため、カウントの結果連続して固有名詞を含む発言が多い場合は、危険度が高いと判定する。

3.1.3 GPS 情報および画像情報の判断

Twitter の発言投稿時に GPS 情報が含まれているかを判断する。GPS 情報はユーザの位置情報を正確に特定することが可能なため、自宅などで GPS 情報が付加された投稿を行うと自宅が特定される恐れがある。本システムでは、位置情報が添付されたツイート削除も可能とする。

また、Twitter の投稿時に画像ファイルに対するリンクを含めることができるため、ユーザが文章を投稿する際に画像へのリンクが含まれているかを確認する。投稿に画像リンクが存在する場合は、画像の内容に社会的に問題があるものが含まれていないか、確認と警告を行う。

また、画像データには撮影した際の情報を追加して保存が可能な Exif (Exchangeable image file format)形式が存在する[5]。撮影時刻、撮影機種、GPS 情報などが含まれるため、Exif 情報を付与した状態でネット上に写真を投稿すると個人情報が流出する可能性があるため、本システムでは Exif 情報について判断する。

Exif 情報の抽出には Java の外部ライブラリとして公開されている `metadata-extractor` を利用する[6]。このライブラリを用いることにより、画像に付与されている Exif 情報をすべて抽出して読み取ることが可能である。Exif 情報は機種名などの基本情報、撮影した際の詳細設定、座標情報がディレクトリとして細分化されている。本システムでは、位置情報が公開されてしまうことへの注意勧告を行うため、

GPS 情報のみを抽出する。画像から抽出した GPS 情報の例を、図 2 に示す。

```

-----
TagType : 1
TagName : GPS Latitude Ref
Description : N
DirectoryName : GPS
TagTypeHex : 0x0001
-----
TagType : 2
TagName : GPS Latitude
Description : 35° 29' 10.66"
DirectoryName : GPS
TagTypeHex : 0x0002
-----

```

図 2. GPS 情報抽出結果の例

また、座標情報からピンポイントで位置情報を特定することが可能なことをユーザに認知させるため、GPS 情報から得られた座標情報を、GoogleMap 上に表示する GPS 情報の結果を GoogleMapURL に渡し、自動的に Web ブラウザを展開することにより、ユーザの現在位置情報が含まれていることに対する警告を行うことができる。GoogleMap 上への表示結果の例を図 3 に示す。また、Twitter の投稿および画像に対して位置情報が付加されていることをユーザが確認した際に、システムで Twitter 発言の位置情報の削除をすることが可能である。

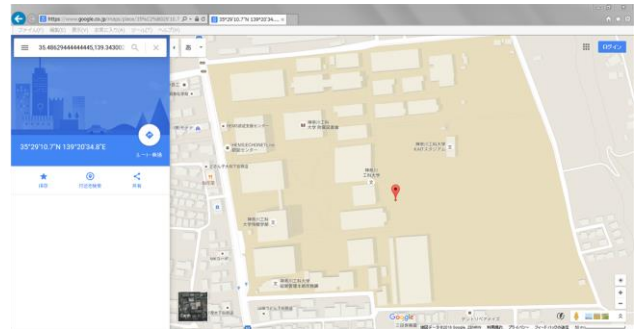


図 3. GoogleMap に対する位置情報の表示例

4. まとめ

本研究により、SNS での不適切な発言、不用意な投稿での自分の発言の危険性を認知させることで、情報リテラシー学習を行うことができる。今後の課題として、画像解析のライブラリを使用して画像に問題となる要素が無いか確認すること、Twitter 以外の他サービスへの対応、発言の影響力を可視化すること、などが挙げられる。

参考文献

- [1] 内閣府, 青少年のインターネット利用環境実態調査平成 27 年度
- [2] 我妻了, 井上拳斗, 岩田一, "Twitter 発言の危険度判定による情報倫理学習支援システム", 第 14 回情報科学技術フォーラム (FIT2015), 2015 年
- [3] 小林直樹, ネット炎上対策の教科書, 双葉社, 2014 年
- [4] `lucene-gosen`, <https://code.google.com/p/lucene-gosen/>
- [5] JEITA, "デジタルスチルカメラ用画像ファイルフォーマット規格 Exif 2.3", 2013 年.
- [6] GitHub `metadata-extractor`, <https://github.com/drewnoakes/metadata-extractor/wiki>