

## SNS から得られる情報を用いたアカウント到達可能性算出

吉國綺乃<sup>†</sup> 渡辺知恵美<sup>‡</sup> 小林一郎<sup>†</sup><sup>†</sup> お茶の水女子大学大学院人間文化創成科学研究科理学専攻情報科学コース<sup>‡</sup> 筑波大学システム情報系情報工学域

## 1 はじめに

近年、ソーシャルネットワークサービス（以下 SNS とする）が普及し、それに伴い世界中で利用者が増加している。SNS は多くの人とコミュニケーションをとる手段として有効である一方、その利用によるトラブルが原因となり利用者の個人情報取得されるという事例が発生している。

SNS の利用において個人情報の取得を防ぐためにも、自身が投稿した内容やプロフィールがどの程度のプライバシーリスクになっているかを把握する必要がある。我々は先行研究において、SNS におけるプライバシーリスクの指標としてアカウント到達可能性を提案している [1]。

利用者が直感的にプライバシーリスクに気付くために、アカウント到達可能性算出により得られた情報がわかりやすく提示される必要がある。そこで本稿ではそれらの情報の提示手法の検討を行う。アカウント到達可能性を算出する際、得られたアカウント同士の類似度を算出する。類似度はプロフィールや投稿内容の情報を用いて算出しており、本稿ではここで得られた情報を可視化することで、アカウント同士の類似性の比較をより直感的に行うことを目的とする。

## 2 アカウント到達可能性

アカウント到達可能性 (Account Reachability) とは、SNS において攻撃者が対象ユーザの既知のアカウントから別の所有アカウントを見つけ出す可能性を表す [1]。たとえば、ある利用者が二つの異なる SNS のアカウント  $s_1$ 、 $s_2$  をそれぞれ持っているとする。また攻撃者は利用者の SNS アカウントのうち、 $s_1$  のみしか知らないとする。攻撃者は  $s_1$  の情報をもとにして、まだ知らないアカウントである  $s_2$  をさまざまな手法を通し

て見つけ出そうとする。ここで攻撃者は  $s_1$  のプロフィールや投稿内容から  $s_1$  のキーワードを抽出し、検索エンジンなどを用いて検索を行い、 $s_2$  になりうるアカウントの候補を取得する手法をとったとする。このとき、取得した候補アカウントそれぞれと  $s_1$  から取得したキーワードをもとに  $s_1$  との類似度をはかり、 $s_2$  が  $s_1$  のアカウントであると特定していく。この可能性がアカウント到達可能性である。アカウント  $s_1$  から別のアカウント  $s_2$  を見つけ出す可能性は以下に表される。

$$AR(s_1 \rightarrow s_2) = \max_{q \in Q} (AR(s_1, s_2, q))$$

$$Q = \text{GenQueries}(s_1.\text{prof}, s_1.\text{msg}).$$

$$AR(s_1 \rightarrow s_2, q) =$$

$$\text{Match}(s_2, \text{Cand}(q)) * \frac{\text{Score}(s_1, s_2)}{\sum_{c \in \text{Cand}(q)} \text{Score}(s_1, c)}$$

$$\text{Match}(s_2) = \begin{cases} 1 & \text{if } s_2 \in \text{Cand}(q) \\ 0 & \text{else} \end{cases}.$$

ここで、 $\text{GenQueries}(s_1.\text{prof}, s_1.\text{msg})$  は  $s_1$  のプロフィールや投稿内容から、 $s_2$  のアカウントを見つけ出すためのクエリを生成する式である。 $Q$  は生成されたクエリの集合であり、 $\text{Cand}(q)$  はクエリ  $q$  で得られた  $s_1$  の別アカウントの候補アカウントの集合である。 $\text{Score}(s_1, c)$  は  $s_1$  と候補アカウント  $c$  との類似度を表す。

## 3 アカウント類似性の可視化

アカウント到達可能性は未知のアカウントを見つけ出す可能性であり、利用者は得られた数値が危険な値であるのか判断が難しいことが課題であった。また数値だけでは今後どのように対策すべきかも分かりづらく、より有効な提示手法が必要である。

そこで本稿では、アカウント到達可能性によりアカウントが特定されたと仮定し、それぞれのアカウント同士の類似性を可視化し利用者に提示することを検討する。我々は  $\text{Score}$  の算出過程に得られた情報に着目した。それぞれの SNS で得られた情報を可視化することで、ふたつの SNS に共通しているところ、また異なるところの比較が可能となる。

先行研究 [1][2][3] ではそれぞれ  $\text{Score}$  の算出にプロフィール情報、地域情報を用いた。これらの情報の可視化を検討する。本稿では、これらの情報を用いたアカウ

## A Computational Model of Account Reachability based on the Information from SNS

<sup>†</sup>Ayano YOSHIKUNI(yoshikuni.ayano@is.ocha.ac.jp),<sup>‡</sup>Chiemi WATANABE(chiemi@cs.tsukuba.ac.jp)<sup>†</sup>Ichiro KOBAYASHI(koba@is.ocha.ac.jp)<sup>†</sup>Advanced Sciences, Graduate School of Humanities and Sciences, Ochanomizu University, 2-1-1 Ohtsuka Bunkyo-ku Tokyo 112-8610 Japan<sup>‡</sup>Faculty of Engineering, Information and Systems, Tsukuba University 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-0006, Japan

ント到達可能性の類似度算出の根拠となるデータの可視化手法について述べる。なお、実装では Twitter\* と Facebook† を対象の SNS とし、Twitter と Facebook のアカウントの類似性を可視化する。

表 1 アカウント間類似度指標

類似度算出指標	Facebookからの情報抽出	Twitterからの情報抽出	算出方法	可視化手法
検索エンジンランキング[1]	プロフィールに登場する単語	検索ランキング順位	Facebookのキーワードで検索した結果のrankの逆数	-
行動範囲 [2]	プロフィールと投稿内容に登場する地名	投稿内容に登場する地名	共通する地名の割合	登場する地域を強調表示
利用者の興味[3]	プロフィール (Favorite情報)	フォローアカウントの投稿内容の重要語	登場する語間の関連度	登場する語間の類似度をグラフ表示
プロフィール情報	プロフィールに登場する単語	プロフィールに登場する単語	共通する単語の割合	登場する語間の類似度をグラフ表示

これまでに定義してきたアカウント間類似度指標を表 1 に示す。先行研究 [1] では類似度指標は検索エンジンの結果を用いランクの逆数を用いた。先行研究 [2] では投稿内容に登場する地名を抽出し、その共通度を類似度とした。先行研究 [3] では利用者の興味を表す語の共起度を類似度とした。また単純にプロフィールに登場する単語の共通度を類似度とする方法も考えられる。アカウント到達可能性を算出する際にはこれらの類似度算出手法のいずれかを用いることを想定する。そこで利用した類似度算出法に合わせて可視化を行うこととする。

本稿では特に地域情報の可視化について述べる。地域情報は個人を特定するために重要な情報のひとつである。各 SNS で得られた地域名を地図上にマッピングすることで視覚的に行動範囲が見られるようになり、文字情報より直感的に理解できると考えられる。

#### 4 地域情報を用いた類似性の可視化の具体例

被験者の SNS 情報から得られた地域を地図上にマッピングした結果を元に、提案手法がプライバシーリスクの意識向上に有効であるか考察する。地図は Google Maps‡ を用いた。取得する地域情報は、Twitter では被験者の最新ツイート 2000 件から抽出した地域名のうち出現頻度上位 5 件を取得する。地域名の抽出には形態素解析器 MeCab§ を用いた。Facebook からはプロフィールに公開している地域名 (居住地、出身地、スポット等) を取得した。それぞれ取得した地域を各 SNS のアイコンと Google Maps API で公開されているサークルを描画する機能を用いて行動範囲を描画する。ここでは 1 名の被験者の結果を図 1, 2 に示す。

この被験者のアカウント到達可能性は 0.17 である。この値は先行研究 [2] の手法で求めている。値を見るとあまり特定される危険性は感じられない。しかし図 1, 2 を見ると類似性は高く、特定される危険性の高さがわかる。図 1 をみると、それぞれの SNS で似た分布が得ら



図 1 地域情報提示結果



図 2 地域情報提示結果 詳細

れた。関東近辺を拡大した図が図 2 である。それぞれ重複しており、ふたつの SNS の使い分けができていないことがわかる。また地図上にマッピングすることで行動範囲がわかりやすくなった。

本稿では被験者 1 名分の結果のみ示したため、今後より多くの被験者で検証を行い、本提示手法がプライバシーリスクの意識向上に有効であるか検証していく。

#### 5 おわりに

SNS の普及により、利用者の個人情報取得される事例が多く発生している。これらのプライバシーリスクに利用者が気がつくためにアカウント到達可能性を提案している。そこで我々はふたつの SNS から得られた情報をわかりやすく提示することで、利用者のプライバシーリスクに対する意識が向上することを目的とし、その提示手法を検討した。本稿では地域情報を提示する手法の検討を行った。地図上に情報をマッピングすることで文字ではわからなかった行動範囲などがわかりやすくなり、自身では気が付かなかった情報を得ることができた。

今後、被験者を増やし検討と考察を行っていく。また、地域情報以外の情報の提示手法も検討を行う。具体的には、SNS 上での活動時間の可視化、それぞれの SNS の友人の可視化など検討している。

#### 参考文献

- [1] Ayano YOSHIKUNI, and Chiemi WATANABE "Account Reachability : A Measure of Privacy Risk for Exposure of a User's Multiple SNS Accounts", ii-WAS2013
- [2] 吉國綺乃, 渡辺知恵美, 小林一郎: SNS の投稿内容に含まれる地域情報を用いたアカウント到達可能性算出モデルの検討, DEIM Forum 2014
- [3] 吉國綺乃, 渡辺知恵美, 小林一郎: SNS の投稿内容に含まれる興味に着目したアカウント到達可能性算出モデルの検討, Web インテリジェンスとインタラクション研究会, 2015

\* <https://twitter.com/>

† <http://www.facebook.com>

‡ <https://maps.google.co.jp/>

§ <http://mecab.googlecode.com/svn/trunk/mecab/doc/index.html>