

## Twitter 投稿文脈による個人情報の推定と個人情報漏洩の防止方法の提案と試作

松下翔太<sup>†1</sup> 東剛秀<sup>†1</sup> 田中貴章<sup>†1</sup>  
杉田裕次郎<sup>†2</sup> 白沢竜馬<sup>†2</sup> 亀澤健太<sup>†2</sup>  
山之上卓<sup>†3</sup> 下園幸一<sup>†3</sup> 小田謙太郎<sup>†3</sup>

ミニブログが注目を集めている。ミニブログは気軽であるがゆえにユーザの意図している以上のレベルで個人情報が流出していることが考えられる。本論文ではミニブログの代表である Twitter の投稿記事から個人情報を推定する方法を挙げ、ユーザが自身のプライバシーを制御する方法を提案する。個人情報の推定方法としてデータマイニング手法を適用し、Twitter ユーザ向けのアンケートの実施及び試作を行い、推定する属性と投稿時刻には相関があることを明らかにした。これにより投稿文脈を手がかりとしてユーザ属性を推定する一つの基準を立てることができた。

## Estimating Personal Information on Twitter and Its Privacy Filter Application

Shota Matsushita<sup>†1</sup>, Takahide Higashi<sup>†1</sup>, Takaaki Tanaka<sup>†1</sup>,  
Yujirou Sugita<sup>†2</sup>, Ryoma Shirasawa<sup>†2</sup>, Kenta Kamezawa<sup>†2</sup>,  
Takashi Yamanoue<sup>†3</sup>, Kouichi Shimozono<sup>†3</sup>  
and Kentaro Oda<sup>†3</sup>

As mini blogs become popular, it is very conceivable that its user personal information exposed at the level more than they are intending it. In this paper, we propose how to estimate personal information on Twitter to control the user's privacy. We conducted questionnaire from voluntary Twitter users to collect ground truth, and applied data mining techniques to estimate personal information. As a result, we showed that user attribute and posting time are correlated. Thus, the correlation could be used for estimating personal information and its filter for his privacy protection.

### 1. はじめに

近年、ブログや SNS といったインターネットサービスに加えて、ミニブログが注目を集めている。日本では Twitter, nowa, Timelog 等のミニブログが多く利用されているが、とりわけ Twitter は日本からのユニークユーザ数が mixi を上回る<sup>1)</sup>など、人気の程が伺える。ミニブログが多くの人に利用されるようになった理由の一つに「気軽さ」が挙げられる。Twitter では一度に投稿できる文字数は最大 140 文字と短いため、書きこむ前に内容を遂行し、時間をかける必要がほとんどない。そのためユーザはあまり深く考えずに投稿することができる。しかしこの気軽さ故に、本人の意図している以上のレベルで個人情報が流出していると考えられる。既に海外では、Twitter で現在外出中と投稿した男性が空き巣の被害に遭ったと訴えている<sup>2)</sup>。

この問題を解決するために、本稿の前半では、先行研究<sup>3)</sup>で述べているように、Twitter を例に挙げ、ユーザの望むプライバシー情報量と外部から観測できるプライバシー情報のギャップについて言及し、ユーザが自身のプライバシーを制御するために Twitter の投稿記事と、その投稿時刻、位置情報、送受信者、投稿手段、投稿アプリケーション等の文脈情報を用い、教師付き学習を行うことで、ユーザの個人情報を推定する方法を提案する。学習と評価にはアンケートを用いる。そして、推定した情報をユーザにわかりやすく提示し、投稿の再考を促す方法を提案する。

後半では、Twitter ユーザ向けに行った自身のプロフィールに関するアンケートの結果を示す。性別や年齢といったプロフィール別にグループを分け、グループごとに Twitter API を利用し、投稿時間や投稿記事の取得を行った。取得したつぶやきをグループごとに分類して解析していくと、ユーザのプロフィールと相関関係のあるものが見つかった。これにより Twitter ユーザに対して、アンケートから判明したプロフィールと相関関係のあるものを含む投稿を示すことができれば、ユーザの意図しないレベルでの個人情報漏洩を防ぐことが可能になると考えている。

最後に今後の課題を示す。

### 2. 問題提起：プライバシー情報のギャップ

ユーザが望んでいるプライバシー情報量は個人差があり、住んでいる県までは公開してもいいが、市町村までは公開したくないという人もいれば、完全にどの地域に住

<sup>†1</sup> 鹿児島大学工学部情報工学科  
Information Engineering, Kagoshima University

<sup>†2</sup> 鹿児島大学大学院理工学研究科  
Graduate School of Science and Engineering, Kagoshima University

<sup>†3</sup> 鹿児島大学学術情報基盤センター  
Computing and Communications Center, Kagoshima University

んでいるのか秘密にしておきたいという人もいる。しかし、ユーザが認可している公開してよい情報と、実際に公開されている情報にはギャップがあり、実際に公開されている情報と第三者の解析によって得られる情報にもギャップがある。前者は投稿記事を非公開にするところを誤って公開する設定にしていたなどといったユーザ側のミスによって起こるギャップである。後者は第三者が Twitter API を利用したり、投稿記事に自然言語処理を施し、解析したりすることによって生まれるギャップである。図 1 に、解析前・後で得られると思われる情報を示す。得られる情報の違いはギャップの存在を示している。Twitter のような緩いコミュニケーションサービス<sup>4)</sup>ではこのギャップが見落とされがちである。

本人が意図している以上のレベルで個人情報が外部に知られてしまうおそれがあることはプライバシー保護の意味で問題がある。このギャップを埋めること、ひいてはユーザが自身のプライバシーを制御することが重要である。

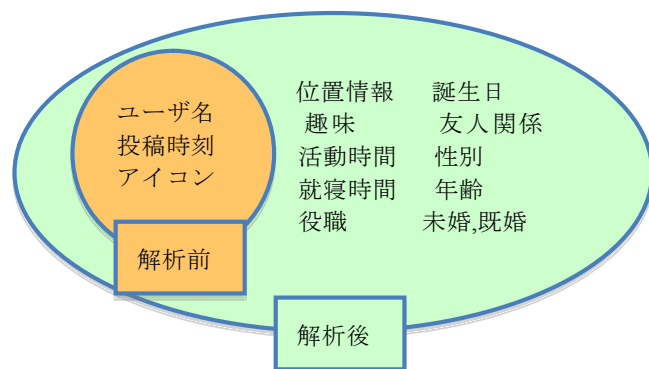


図 1 Twitter 解析前・後で得られる情報の違い

### 3. アプローチ

#### 3.1 投稿文脈による個人情報の推定

推定する対象の個人情報を性別、年齢、身分、収入、趣味、位置情報といった属性に分類する。これらを推定する元となる情報を、投稿記事以外に、投稿時刻、位置情報、送受信者、投稿手段、投稿アプリケーションとする。これらは Twitter API を利用することで簡単に取得可能である。

例えば、「@hoge hoge モバゲーやるといいよ」という投稿記事が、「携帯電話から」

「20 時 15 分」に「位置情報」(緯度: 31.59 経度: 130.55) 付きで投稿されたとする。この投稿文章から、投稿者は「モバゲー」を知っていることが推定できる。加えてこの投稿の「投稿アプリケーション」から、投稿者は携帯電話を所持しており、位置情報から、鹿児島に住んでおり、投稿時刻が就学時間帯外ということから 18 歳以下で、送信先より @hoge hoge と関係がある人物であるということが推定できる。この例のように投稿記事のみでは推定しにくい情報を、投稿文脈で補う形により精度の高い個人情報推定が可能になるのではないかと考えている。実際に時間等を考慮せずに位置履歴のみでユーザの属性を推定する研究<sup>5)</sup>では、SVM で学習を行ない、一定の成果が認められている。

本論文では、学習と評価のためにアンケート結果を用いる。アンケート内容は Twitter ユーザ向けに、先程分類した自身のプロフィール(個人情報)を回答してもらうもので、web 上に特設ページを設けるという形を取る。

回答してもらったユーザ全員分の投稿記事や投稿時刻を Twitter API を用いて収集したのち、アンケート結果を元に、性別なら男のグループ、女のグループ、年齢なら〇〇歳以下のグループといった具合に属性のごとのグループに分ける。

ユーザのプロフィールを推定するにおいて、推定する元となる情報のうちどれが手がかりになるかを明らかにするために、表 1 のように分類したグループと推定する元となる情報に相関関係があるかどうかをそれぞれ調べる。相関が見つかったら、その特徴を学習していく。

プロフィールを推定するには、推定する対象のユーザの投稿記事や投稿時刻を Twitter API から取得し、特徴を見つけて学習元データからユーザ属性を逆算すればよい。

表 1 推定する属性のグループと推定する元となる情報の相関関係イメージ図

	投稿記事	投稿時刻	位置情報	送受信者	投稿手段
性別 (男, 女)	相関あり			相関あり	
年齢 (〇〇歳以下…)	相関あり	相関あり			相関あり
身分 (社長, 一般社員…)			相関あり		
年収 (〇〇万円以下…)	相関あり	相関あり			
位置情報 (〇〇県…)			相関あり		

#### 3.2 個人情報漏洩の防止方法

個人情報漏洩の防止方法として考えられるのは、ユーザ自身が書きこむ量を減らすことや、第三者による解析を混乱させるために投稿時間や位置情報をでたらめなものにする等が考えられるが、前者はユーザに負担がかかり、後者はコミュニケーションツールとしての Twitter の利用に支障が出る恐れがある。そのため、ユーザには

普段と同じように Twitter を利用でき、個人情報推定につながるような投稿をしそうになった場合、もしくは過去にそのような投稿をしていた場合、警告を発するようなシステムが望ましいと考える。

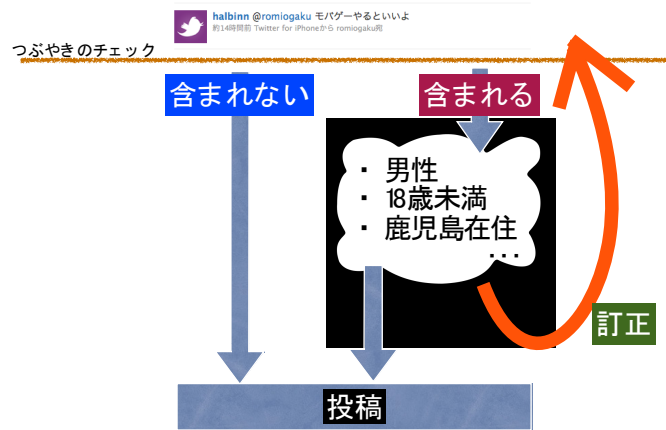


図 2 投稿の再考を促すシステムのイメージ

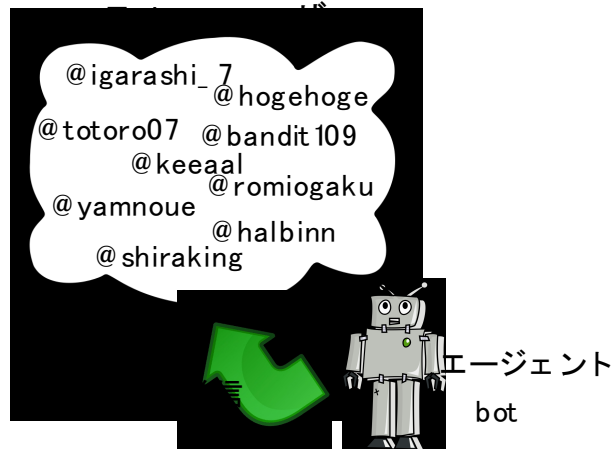


図 3 エージェント bot によるメッセージ送信のイメージ

例えば図 2 のように、ユーザが投稿しようとするつぶやきを事前にチェックし、学習データを参考にして個人情報推定に深く関わっている語や特定のパターンの投稿が含まれている場合は、その語やパターンと、それから推測されるユーザ属性を提示すると共に、本当に投稿してよいかの確認ダイアログを表示する。既に入力済みでしまった記事にも同様にチェックし、削除できるようにする。最後に処理後の投稿記事から個人属性をもう一度推測し、実際の個人属性と一致しなくなるまで操作を繰り返すことで、ユーザの意志による自身のプライバシーの制御を実現する。この機能を、つぶやく、フォローするといった基本機能を持った Twitter クライアントに追加すれば、普段どおりの Twitter の利用が可能になる。

または図 3 のようにユーザが過去に投稿したつぶやきを取得し、個人情報が推定しやすくなっていないかを監視し、投稿に特定のパターンが見つかった場合はそのユーザにメッセージを送信するエージェント bot を作成する方法も効果的であると考えられる。

これらのように、最終的にはユーザの手で自身の望むプライバシーのレベルまで調節できる方法を提案する。

#### 4. Twitter ユーザ対象のアンケート

##### 4.1 概要

前項の表 1 で示した、ユーザの属性グループごとの投稿記事や投稿時刻等の特徴を調べるため、Twitter ユーザ対象のアンケートを実施した。アンケートの概要は以下のとおりである。

- (1) アンケート対象者  
日本のすべての Twitter ユーザ。
- (2) 実施期間  
2010 年 10 月以降。2011 年 1 月現在も継続中。
- (3) 実施方法  
Twitter と連動してアンケートの募集、宣伝、集計が容易にできる web サービス<sup>6)</sup>を利用してアンケート回答用ページを作成した。また、アンケート回答用ページとは別に、本研究の研究目的、データの使用方法等を記述した解説ページも作成し、回答用ページと相互リンクを貼り、記述内容に同意した人のみに回答してもらった。
- (4) 募集方法  
本研究用の Twitter アカウントを取得し、自動フォロー bot を作成し、アンケートの協力を呼びかけた。
- (5) 質問事項とその選択肢  
表 2 に示す。趣味のみ複数選択可の形式にした。

表 2 アンケートの質問事項とその選択肢

質問事項	選択肢
性別	男, 女
年齢	10歳未満, 10~14歳, 15~19歳, 20~24歳, 25~29歳, 30~34歳, 35~39歳, 40~44歳, 45~49歳, 50~54歳, 55~59歳, 60歳以上
職業	事務系会社員, 企画・調査系会社員, 経営・管理職, 営業・販売系会社員, 技術系(ソフトウェア, ネットワーク)会社員, 技術系(電気, 電子, 機械)会社員, 技術系(素材, 食品, 医療)会社員, 技術系(建築, 土木)会社員, 公務員, 医療関連職, 教育関連職, 金融関連職, クリエイティブ系, 自営業, 自由業, 小学生, 中学生, 高校生, 大学生・大学院生, パート・アルバイト, 専業主婦, 無職, その他
役職	会長・社長, 役員, 部長, 課長, 係長・主任, 一般社員・職員, それ以外(派遣, 契約), 働いていない
年収	300万円未満, 300~500万円未満, 500~700万円未満, 700~1000万円未満, 1000万円以上, 収入なし, 答えたくない
都道府県	(47都道府県)
既婚・未婚	未婚(恋人なし), 未婚(恋人あり), 既婚(子供なし), 既婚(子供あり)
休日の過ごし方	家にいるほうが多い 外にいるほうが多い
一日でどのくらいパソコンを利用するか	ほとんど利用しない, ~3時間未満, 3時間~6時間未満, 6時間~9時間未満, 9時間~12時間未満, 12時間~
昼食や夕食を友人と食べることが多いか	はい, いいえ
喫煙するか	はい, いいえ
趣味	スポーツ, コンサート・ライブ等, 車・バイク, パチンコ, 芸術鑑賞, ショッピング, アニメ・漫画, ゲーム, 旅行・観光, カラオケ, 何かの収集, ブログ更新(mixi等SNSの日記を含む), まとめブログ閲覧
自分はリア充だと思えるか	はい, いいえ, リア充の意味がわからない
Twitter ユーザ名	(自分の Twitter ユーザ名)

質問事項に関しては、よりプライベートな情報が得られるように、性別や年齢といった項目の他に年収や趣味、リア充かどうかといった項目を入れた。なお、リア充とは現実の生活が充実している人達の事を指すインターネットネットスラングで、2007年頃からブログや Twitter で流行した。その定義は明確化されていないが、一般的には主に恋人がいる人を指すことが多い。

集計結果は csv ファイルで出力した。

#### 4.2 アンケート結果

2011年1月時点で男性70人、女性24人、計94人分の回答結果が得られた。また、TwitterAPIによって得られたつぶやき数は、男性が93,236件、女性が34,820件、合計128,056件であった。Twitterの年齢別利用者層は、アスキー総合研究所の「Twitter利用実態調査」(2009年12月)<sup>7)</sup>によると、20代が最も利用率が高く、平均年齢は35.7歳であったとされている。今回のアンケートに回答した人の年齢の分布を図4に示す。20~24歳が最も多いが、平均年齢としては、Twitter全体の平均よりやや低いことがわかる。

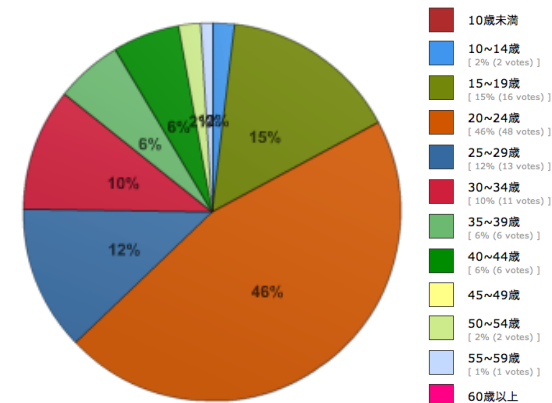


図4 アンケートの年齢結果

#### 4.3 分析と考察

アンケート結果から回答者をグループに分け、TwitterAPIを用いて投稿日時が新しいものからひとりにつき最大3200件のつぶやきを取得した。今回はつぶやきの投稿時間とグループの相関関係について調査した。そのうち「自分はリア充だと思うか」の質問において、リア充だと思うと回答したグループとリア充だと思わないと回答したグループで、投稿時間にどのような違いが見られるのかを解析した。

投稿時間を 0 時台から 23 時台の 24 時間帯に分け、時間帯ごとのつぶやき数を全員分カウントした。つぶやき数には個人差があるので、正規化が必要がある。

Twitter ユーザ  $i$  のつぶやき数を  $T(i)$ 、つぶやきを取得できたユーザ人数を  $n$  とすると、平均つぶやき数  $T_{ave} = \sum_i T(i)/n$  である。まず、今回の結果では  $T_{ave}$  は約 1362、標準偏差は約 1180 であったので、 $1362-1180=182$  以下のつぶやきしか取得できなかったユーザのデータは除外した。

次に、各時間帯ごとの正規化したつぶやき数を求める。  $h$  時台のつぶやき数を  $T_h(i)$ 、  $h$  時台の正規化つぶやき数  $T_{hn}(i)$  を、  $T_{hn}(i) = T_h(i) * \frac{T_{ave}}{T(i)}$  とした。リア充だと思えるグループ、リア充だと思わないグループで時間帯ごとの平均正規化つぶやき数とその標準偏差を求めた。

図 5 にグラフ化した結果を示す。横軸は時間、縦軸は平均正規化つぶやき数である。標準偏差は対応する色の縦区間で表している。

平均正規化つぶやき数の時間推移を見ていくと、朝 8 時頃から 20 時頃では、リア充だと思える人のほうが多くつぶやいているが、21 時頃から翌 7 時頃までは、逆にリア充だと思わない人のほうが多くつぶやいていることがわかる。

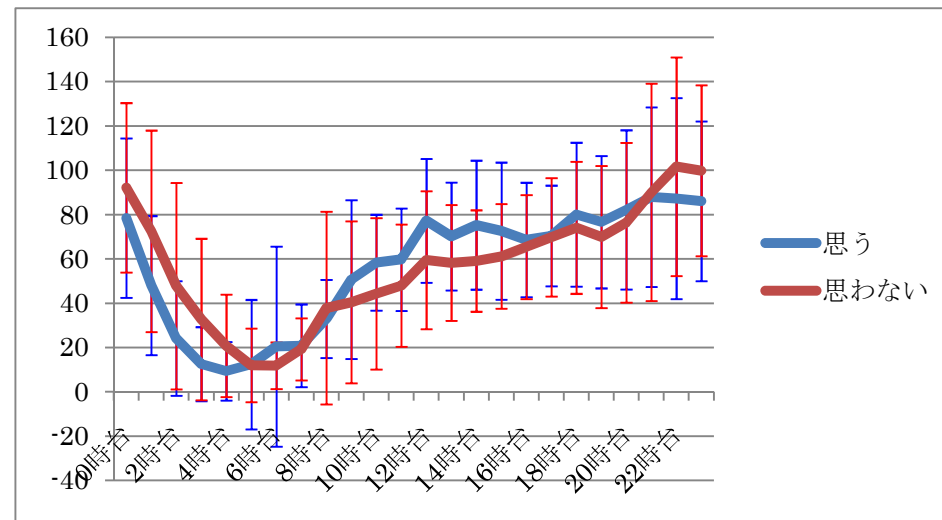


図 5 リア充だと思えるグループ、リア充だと思わないグループの平均正規化つぶやき数の時間推移

分布の形を見ると、重なっている部分もあるが、12 時台や 22 時台など、グループによってつぶやき数の差が大きい時間帯では分布の重ならない部分が比較的はっきりとしており、ばらつきが逆転しているところも殆ど無い。少なくともこの平均正規化つぶやき数の時間帯推移の形は信用できるということが言える。

投稿時間の解析だけで、リア充か否かの判定をする一つの基準ができた。投稿時間以外にも、投稿記事や位置情報、投稿アプリケーション等でも同様に解析を行い、グループごとの特徴を見つけることで、より詳細なユーザ属性推定が可能になると考える。

## 5. 試作

現在投稿時間の解析に続いて、投稿記事に対して bag-of-words による文書分類を行っている。データマイニングツールである weka を用いて単語の重み付けや分類を行っており、様々な手法を用いて推定する属性のグループとの相関関係について解析を行なっている。

## 6. 今後の課題

現在も実施中であるアンケートの回答者数をいかに増やし、解析の精度を向上させるかが重要である。また、アンケート回答者に偏りが出ないように、アンケートの募集を呼びかける自動フォローbot のフォロー先の指定は、ランダムではなくこちらからある程度目星を付ける工夫が必要であると考えている。

今回は投稿時間の解析に留まったが、より詳細なユーザ属性を推定するには他の条件での解析と合わせて総合的に判断しなければならない。いかに効率よくグループごとの特徴を見つけるかが今後の課題である。

## 7. おわりに

本稿では、Twitter において本人の意図している以上のレベルで個人情報が漏洩していることの現状に対する問題提起と、それに対するアプローチとして個人情報を推定する方法と、個人情報の漏洩を防止するためにユーザの手で自身の望むプライバシーのレベルまで調節できるシステムを提案した。Twitter ユーザを対象に実施したアンケートの結果から回答者のグループ分けを行い、そのうち自分がリア充だと思えるグループとリア充だと思わないグループそれぞれに対して投稿時間との相関関係を調査した。その結果、リア充だと思えるグループが多くつぶやく時間帯と、リア充だと思わないグループが多くつぶやく時間帯が存在していることが明らかになった。

これによって投稿時間のみでその人のプライベートな趣向 (= 個人情報) を推定す

る為のひとつの基準ができた。

**謝辞** 本研究のアンケートにご協力頂いた皆様に、謹んで感謝の意を表する。

### 参考文献

- 1) Nielsen Online NetView 2010年4月 家庭と職場のPCからのアクセス,  
[http://www.netratings.co.jp/hot\\_off/archives/NNR06012010.htm](http://www.netratings.co.jp/hot_off/archives/NNR06012010.htm)
- 2) <http://www.azcentral.com/offbeat/articles/2009/06/11/20090611mr-twitter0614.html>
- 3) 松下翔太, 小田謙太郎, 下菌幸一 “Twitter 投稿文脈による個人情報推定と個人情報漏洩の防止方法の提案”, 第63回電気関係学会九州支部連合大会 (2010.9)
- 4) 佐藤浩史, 山崎敬広, 高橋紀之 “位置情報サービスにおけるプライバシー保護の為にコンテンツ情報の希釈化”, 電子情報通信学会技術研究報告 Vol. 108, No.289, pp. 37-41, (2008.11)
- 5) 松尾豊, 岡崎直観, 中村嘉志, 中村拓一, 橋田浩一, 中島秀之, “位置履歴からのユーザ属性の推定”, 情報処理学会論文誌 Vol.48, No. 6, pp.2106-2117, (2007.6)
- 6) twitpoll, <http://twitpoll.com/>
- 7) アスキー総研ニュース, <http://asciimw.jp/info/release/pdf/20091228.pdf>