

マイクロブログアカウントの偽装を判定するシステムの提案

A System of Identifying Suspicious Microblog Accounts

中才 恵太郎†

Keitaro Nakasai

角田 雅照†

Masateru Tsunoda

1. はじめに

近年、代表的なマイクロブログサービスである Twitter が広く利用されている。個人に加え、企業や著名人も Twitter を利用しており、宣伝のために効果的に利用している。ただし、Twitter の普及に伴い、著名人のアカウントのなりすまし（他者があたかも本人のようなアカウントを作成して公開すること）も増加しており、問題となっている。女性アイドルや、女性声優の偽アカウントが特に多い傾向がある。これらのなりすましアカウントは、Google の検索結果でも上位に現れる場合が多くある。Twitter 社はなりすましを防止するため、サービス上で本人であることを保証する、認証済みアカウントを設けている。

ただし、全ての著名人が認証済みアカウントを利用しているわけではなく、著名人であっても通常のアカウントで Twitter を利用している場合もある。この場合、著名人自身の公式サイトやブログに自身のアカウントをリンクすることで本物であることを証明している場合もある [2]。この場合でも、全ての著名人が自身の公式サイトを持っているわけではなく、公式サイトを持っている場合でもアカウントをリンクしているとは限らない。

このため、著名人のアカウントをフォローしようと試みるユーザは、アカウントが本物かどうかを十分に検討する必要がある。Twitter には SNS (Social Networking Service) の機能が実装されており、アカウント同士でつながりを持つことができる。これがフォロー機能と呼ばれるものであり、あるアカウントをフォローしている他アカウントをフォロワーと呼ぶ。偽物のアカウントをフォローすることにより、誤った情報を入手してしまう可能性もある。

そこで本研究では、マイクロブログで最も普及している Twitter に着目し、著名人のアカウントがなりすましであるかどうかを判定するシステムを作ることを目的とする。そのために、Twitter アカウントのなりすましを判定するアルゴリズムを提案するとともに、アルゴリズムの有効性を確かめるために、簡易な実験を行う。

これまで Twitter アカウントのフォロワーが BOT、非アクティブ、アクティブであるかを判定する Web サービス [1]などは公開されている。また Twitter 上でのアカウント間の繋がりに着目した研究がいくつか存在する [3][4]。ただし Twitter アカウントが著名人本人であるかを判定する Web サービスは、我々の知る限り公開されておらず、判定するためのアルゴリズムを提案した研究も存在しない。

2. 提案手法

2.1. 著名人の Twitter アカウントの特徴

多くの著名人の Twitter アカウントでは、フォロワー数は少ないが、フォロワー数が多いという特徴がある。フォ

ローするアカウントには類似の職業に従事している、直接の知り合いが多く、相互にフォローしていることが多い。Twitter サービスでは、数多くのアカウントをフォローすると、タイムラインと呼ばれる画面にフォローしているアカウント全てのツイート（入力内容）が表示される。このため、著名人のアカウントでない場合でもフォロワー数は抑えられることが多い。これに対し、著名人のフォロワー数は特に多い傾向がある。これは著名人の情報を知りたい多くのユーザがフォローするためである。

2.2. 提案アルゴリズム

本研究では、著名人の Twitter アカウントが本人であるかどうかを確認するために、対象アカウントのフォロワーとフォロワーの関係を利用する。著名人の Twitter アカウントのフォローは前節で述べたとおり、類似の職業に従事している知り合いが多く、相互フォローが多い。また、著名人は Twitter アカウントのフォローに慎重であり、認証済みアカウントを持った著名人がフォローするアカウントは信頼性が比較的高いと考えられる。

そこで、本手法では、認証済みアカウントに相互フォローされているかどうかに基づき、Twitter アカウントのなりすましを判定する。具体的には、提案するアルゴリズムでは、以下の手順でアカウントの偽装を判断する。

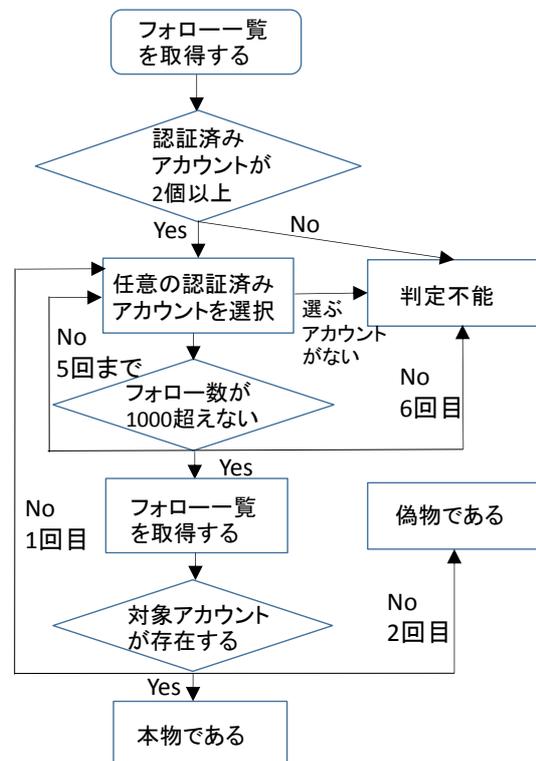


図1 提案する判定アルゴリズム

表 1 実験に用いたアカウント(2015/7/23 現在)

アカウント名	本人かどうか
@rino_sashihara	本人でない可能性が高い
@oshima_yuko_	本人でない可能性が高い
@mimori_suzuko	認証済みではないが本人
@tokui_sorangley	認証済みではないが本人

表 2 本人でないと思われるアカウントの概要 (2015/7/23 現在)

アカウント名	フォロワー	フォロー	開設日
@rino_sashihara	14	287273	2010/12
@oshima_yuko_	20	18779	2010/10



図 2 なりすましの可能性があるアカウントの一例 (https://twitter.com/rino_sashihara)

フローチャートを図 1 に示す。

1. 判別対象の Twitter アカウントの、フォロワー一覧を取得する。
2. フォロワー一覧から認証済みアカウントが 2 個以上あるかどうかを確かめ、存在しない場合は判定不能とする。
3. 2 個以上存在する場合、任意の認証済みアカウントを選択する。
 - i. フォロワー数が 1,000 を超えない場合、フォロワー一覧を取得し、対象アカウントが存在するか確認する。
 - ii. フォロワー数が 1,000 を超える場合、認証済みアカウントを選び直し、同様の操作を行う。
 - iii. 選ぶ認証済みアカウントがなくなった場合と、6 回この操作を行った場合、判定不能とする。
4. 対象アカウントが存在した場合、本人のアカウントと判定する。対象アカウントが存在しない場合、3 の手順を再び行う。対象アカウントが存在しない場合は偽物の可能性があるとして判定する。

相互フォローを調べるツール[5]は存在するが、アカウントの偽装を判定する機能は実装されていない。このサービスでは、フォロワーまたはフォローが 10,000 人以上ある場合、正常に動作しない。これは Twitter API の使用制限によるものだと考えられる。

著名人の Twitter アカウントはフォロワー数が 10,000 を超えることがしばしば発生するため、Twitter API を用いてアカウントの偽装を判定する際、Twitter API の呼び出し回数を抑える必要がある。提案アルゴリズムでは著名人の

フォロワー数は少ない傾向にあることに着目し、フォロワー一覧の中から相互フォローしているかどうかを確かめている。

3. 実験

提案したアルゴリズムが有効であるか確かめるための、簡単な評価実験を行った。2 個の本人でない可能性が高いアカウント（本人である認証済みアカウントが別に存在しているため偽アカウントと考えられる）と、2 個の認証済みアカウントではない、本人であるアカウント（公式ブログからリンクが存在する）に対してアルゴリズムを適応した。各アカウントを表 1 に示す。

実験で用いたなりすましの可能性が高いアカウントの情報を表 2 に示す。どちらのアカウントも開設日が 4 年以上前である（長い間、偽アカウントだと知られていない）であるに関わらず、フォロワーが 1 万人を超えており、まるで本人であるかのように見える。また、Google の検索結果でもかなり上位に表れる。ただし、実際には偽アカウントであるため、なりすましの判別が必要となる。

以降では、各アカウントに対する判定方法とその結果を述べる。

@rino_sashihara (図 2, なりすまし) に対する判定：

1. フォロワー一覧を見ると、認証済みアカウントが 12 個であるため判定可能である。
2. 最初に表示される認証済みアカウント @AyakaUmeda のフォロワー数は 425 個 (1,000 個を超えていない) であるため、フォロワー一覧を取得する。
3. フォロワー一覧に @rino_sashihara のアカウントは存在しなかった。
4. 次に表示される認証済みアカウント @Kobayashikana48 のフォロワー数は 135 個であるため、フォロワー一覧を取得する。
5. フォロワー一覧を見ると、@rino_sashihara のアカウントは存在しなかった。

よって、提案アルゴリズムに基づき、@rino_sashihara はなりすましであると判定できた。

@oshima_yuko_ (なりすまし) に対する判定：

1. フォロワー一覧を見ると、認証済みアカウントが 3 個であるため判定可能である。
2. 最初に表示される認証済みアカウント @kojiharunyan のフォロワー数は 118 個であるため、フォロワー一覧を取得する。
3. フォロワー一覧を見ると @oshima_yuko_ のアカウントは存在しなかった。
4. 次に表示される認証済みアカウントである @tomococo73 のフォロワー数は 51 個であるため、フォロワー一覧を取得する。
5. フォロワー一覧を見ると @oshima_yuko_ のアカウントは存在しなかった。

よって @oshima_yuko_ の場合も、提案アルゴリズムに基づきなりすましである可能性が高いと判定できた。

@mimori_suzuko (本人) に対する判定：

1. フォロワー一覧を見ると、認証済みアカウントは 31 個であるため判定可能である。
2. 最初に表示される認証済みアカウントである @ma_shi_ron のフォロワー数は 87 個であるため、フォロワー一覧を取得する。
3. フォロワー一覧を見ると @mimori_suzuko が存在する。

よって、提案するアルゴリズムにより、このアカウントは本物であると判定できた。

@tokui_sorangley に対する判定：

1. フォロワー一覧を見ると、認証済みアカウントは 40 個であるため判定可能である。
2. 最初に表示される認証済みアカウントである @tenshinmukai のフォロワー数は 721 個であるため、フォロワー一覧を取得する。
3. フォロワー一覧を見ると @tokui_sorangley が存在する。

よって、このアカウントも、提案するアルゴリズムにより本物であると判定できた。本人であるか怪しいアカウント 2 個と、認証済みでないが本人であるアカウント 2 個を正しく判定できたため、提案するアルゴリズムはある程度の妥当性を持つと考えられる。

4. おわりに

本研究では、マイクロブログで広く普及している Twitter に着目し、アカウントの偽装を判定するシステムを提案した。提案システムでは、疑わしいアカウントのフォロワー一覧に着目して判定するため、Twitter API の制限を回避することができる。今後は、実際にシステムを構築し、多くの著名人の Twitter アカウントを利用して偽装を判断し、本手法の信頼性を確かめることである。また、現在のアルゴリズムでは判定できないアカウントも多く存在すると考えられるため、Twitter API の制限を考慮してアルゴリズムを改善し、判別精度を高めることも課題のひとつである。

参考文献

- [1] Fake Follower Check, <http://fakers.statuspeople.com/>
- [2] 剛力彩芽 official web site, <http://www.oscarpro.co.jp/talent/goriki/>
- [3] 小出明弘, 斉藤和巳, 風間一洋, 鳥海不二夫: ネットワーク分析による Twitter ユーザのフォロワー形成に関する一考察, 情報処理学会論文誌 数理モデル化と応用, No.6, Vol.2, pp.164-173 (2013).
- [4] 西村章宏, 土方嘉徳, 三輪祥太郎, 西田正吾: 一般ユーザの観点に基づく Twitter からの人物関係の可視化と事例の考察, 情報処理学会論文誌, Vol.56, No.3, pp.972-982 (2015).
- [5] twitter 片思いチェッカー, <http://kataomoi.058.jp/>