

平成 27 年度 修士論文



Twitter におけるフォロワーマーケットの実 態調査とフェイクアカウントの抽出方法

Understanding Follower Market of Twitter and Detecting Fake
Accounts

指導教員 森 達哉 准教授

早稲田大学基幹理工学研究科情報理工通信専攻

学籍番号 5114F059-1

竹越健斗
2016年2月1日

概要

Twitter を始めとするマイクロブローキングサービスは世界中で利用されている。ユーザーが サービスを利用する目的は多岐に渡るが、中には個人の自己表現や政治活動等の用途で利用する ケースがある。このような「宣伝用途」のニーズに呼応してリツイートやフォロワーを購入できる 「フォロワー マーケット」が登場した。フォロワーマーケットのサービスはソーシャルネット ワークにおける評判情報 を人工的に作り出してしまうため、悪意あるツイートの大量拡散や社 会不安を煽るような情報操作に悪用されてしまうリスクがある。本論文では Twitter を対象とし て国内外のフォロワーマーケットを対象とした実態調査を行い、フェイクアカウントに固有な いくつかの特徴を明らかにした。また実態調査の 結果を元にマーケット内で購入できるフォロ ワー（以下フェイクアカウント）の特徴分析を行い、フェイクアカウントの自動抽出を試みた。そ の結果、約 1,200 万の未知アカウントから約 4,500 のコミュニティ に分類される約 32 万のフェ イクアカウントが抽出された。

目次

第 1 章	序論	8
1.1	研究背景	8
1.2	研究目的, 貢献.....	10
1.3	研究に関する倫理	10
1.4	論文の構成	10
第 2 章	フォロワーマーケット	12
2.1	フォロワーマーケットの概要	12
2.2	フェイクアカウントの脅威	13
2.3	フォロワーマーケットに関する関連研究	14
2.4	Twitter 社側の対策	15
第 3 章	データセット	17
3.1	データセットの概要	17
3.1.1	正規アカウント	17
3.1.2	フォロワーマーケットのフェイクアカウント	17
3.1.3	未知のアカウント集団.....	18
第 4 章	国内マーケットの調査	20
4.1	国内マーケットの実態調査	20
4.1.1	国内マーケットの概要.....	20
4.1.2	followee の特徴	21
4.1.3	フォロワーの特徴.....	21
4.1.4	フォロワー・被フォロワーのバランス	22
4.1.5	klout score の特徴.....	22
4.1.6	活動停止時間の特徴	24
第 5 章	海外マーケットの調査	25
5.1	海外マーケットの実態調査	25

5.1.1	海外マーケットの概要.....	25
5.1.2	followee の特徴	26
5.1.3	フォロワーの特徴.....	26
5.1.4	フォロワー・被フォロワーのバランス.....	27
5.1.5	klout score の特徴.....	28
5.1.6	活動停止時間の特徴	29
第 6 章	フェイクアカウント抽出手法.....	30
6.1	フェイクアカウント抽出手法の概要.....	30
6.1.1	ffrate フィルタリング.....	30
6.1.2	機械学習によるフィルタリング	30
6.2	Jaccard 係数	31
第 7 章	提案手法の評価結果.....	33
7.1	ffrate フィルタリングの考察	33
7.2	SVM によるフィルタリング	33
7.3	Jaccard 係数によるコミュニティ検出	34
7.3.1	Low Jaccard 係数の考察.....	34
7.3.2	Middle Jaccard 係数の考察.....	35
7.3.3	High Jaccard 係数の考察	35
第 8 章	まとめ.....	38
第 9 章	研究業績.....	39
謝辞	40
参考文献	41
付録 A	本研究に関する倫理的問題に関して.....	43
A.1	倫理的な見解	43

目次

2.1	ウェブサイト型のフォロワーマーケットの例	13
2.2	twitter 内におけるフォロワーマーケットの商売活動の例	13
2.3	プロモーションの例	14
4.1	フェイクアカウントの followee 数	21
4.2	フェイクアカウントの follower 数	22
4.3	フェイクアカウントの ffrate	23
4.4	フェイクアカウントの klout score	23
4.5	フェイクアカウントの活動停止時間	24
5.1	フェイクアカウントの followee 数	27
5.2	フェイクアカウントの follower 数	27
5.3	フェイクアカウントの ffrate	28
5.4	フェイクアカウントの klout score	28
5.5	フェイクアカウントの活動停止時間	29
6.1	提案手法のフローチャート	32
7.1	Jaccard 係数の分布	34
7.2	フェイクアカウントのペアと考えられる Middle Jaccard の例	35
7.3	正規ユーザーのペアと考えられる Middle Jaccard の例	35
7.4	同一コミュニティ内のペア (=0.6) の例	36
7.5	同一コミュニティ内のペア (=1.0) の例	36

表目次

1.1	主要な SNS とユーザー数	8
3.1	国内フォロワーマーケットの例 . 価格はそれぞれ 1,000 フォロワー (円) および 1,000 リツイート (円)	18
3.2	海外フォロワーマーケットの例 . 価格はそれぞれ 1,000 フォロワー (USD) お よび 1,000 リツイート (USD)	18
3.3	使用データの概要	19
4.1	国内フォロワーマーケットのサービス調査結果	20
5.1	国内フォロワーマーケットの例	25
5.2	海外フォロワーマーケットの購入履歴	26
6.1	10-fold Cross Validation による SVM の訓練結果	31
7.1	ffrate フィルタリングの実行結果	33
7.2	SVM による分類結果	34
7.3	同一コミュニティ (size=10) の例	37

第 1 章 序論

本章では本論文の背景と概要を示す．1.1 節では本研究の背景について詳しく述べ，1.2 節では本研究は果たす目的について述べる．また 1.3 節にて本論文の構成について述べる．

1.1 研究背景

近年インターネットの普及に伴い，多くの人々が Twitter や Facebook といったソーシャルネットワークサービス（SNS）を利用するようになった．また現在全世界では以下の主要な SNS が利用されている [1]．

表 1.1 主要な SNS とユーザー数

SNS サイト	アクティブユーザー数
Facebook	13 億 5000 万人
Twitter	2 億 8400 万人
LINE	5 億 6000 万人
Instagram	3 億人
Pinterest	7000 万人

このように多くの人々が SNS を利用している．また Twitter や Facebook に代表されるマイクロブローギングサービスの台頭により，今まで疎遠であった友人と長く繋がることを可能にした．さらにこれらのサービスは共通の趣味を持つ者同士，或いはビジネスパートナーを世界中から集めることを容易にできるようにした．

しかし一方でこれらマイクロブローギングサービスのユーザー数増加につけ込み，悪意のあるマーチャントが出現するようになった．このマーチャントはウェブサイトを紹介して，或いは

オークションサイトを介してフォロワーの増加やリツイート代行を販売している．またこのマーチャントは大きく3つのユーザー群をターゲットにしている．

1つ目のユーザー群はマルウェア配布サイトやフィッシングサイト URL を含むツイートを発信するユーザー群である [2]．悪性 URL を正規のユーザーにアクセスしてもらうためにはある程度の信頼を予め獲得する必要がある．その信頼を獲得するために、マーチャントから見せかけのフォロワーを購入したり、リツイートの代行を申請することで、あたかも見かけ上信頼出来るユーザーを作り上げている．さらにこれらのユーザーはハッシュタグを用いることで、そのハッシュタグに興味をもつ正規ユーザー群から信頼を得ることを実際に行っている [2]．このように悪意のあるユーザーは自らのアカウントの信頼度を上げるために、マーチャントから見かけ上のフォロワーを大量に購入し、さらに人気のハッシュタグを利用することで、巧みに正規ユーザーを騙し被害を与えている．

2つ目は小さなビジネスを展開している会社、或いは経営者、有名人になりたいと考えているユーザー群である [3]．これらのユーザーはソーシャルネットワークサービス上で知名度を上げることで、今後のビジネス展開や仕事の増加に繋がていきたいと考えている人たちである．その第一歩としてマーチャントからフォロワーを購入することで見かけ上非常に有名なユーザーであることを装うことができ、その結果正規ユーザーの目に止まることを目標としている．同様にリツイートを使用することで正規ユーザーの目に止まることを目的とする場合もある．このように今後大きなビジネスを展開したいユーザーや今後著名な有名人になりたいと考えているユーザーが初期投資としてマーチャントからフォロワーやリツイート代行を購入することも考えられる．

3つ目は政治関連のイベントで影響力を持ちたいと考えているユーザー群である [4]．近年選挙やデモといった政治的イベントの際に Twitter や Facebook といったマイクロブローキングサービスを利用することで、参加を募ったり指示者を集めたりすることが世界中で広がりつつある．ユーザーによっては自身の信頼度を高めるためにフォロワーやリツイート代行を購入することで、見かけ上の信頼度をあげることを考える者がいる．これによりあたかも世間一般的に信頼されているユーザーを作り出すことができ、結果的に支持団体のない正規ユーザーからの支持を得ることができる可能性を増やしている．

このように大きく3つのユーザー群を対象にマーチャントはフォロワーやリツイート代行を販売している．2011年の調査によると [5] Twitter 内に存在するアクティブユーザーのうち、約3%のユーザーがマーチャントが販売しているフェイクアカウントであると考えてられている．さらに Facebook においてもアクティブユーザーのうち1.5%が同様にマーチャントが販売しているフェイクアカウントとされている．このように現在ウェブ上に存在するフェイクアカウントは単にソーシャルネットワークサービスの問題だけではなく、ウェブサービス全体に関わる大きな問題となってきた．

1.2 研究目的，貢献

本研究は上述の背景をもとに，実在する国内・海外のフォロワーマーケットの実態調査を行う．また，実態調査を元に得られたフェイクアカウントの特徴を利用して，未知のアカウント集団からフェイクアカウントを抽出する方法を確立する．本研究の主要な貢献は下記のとおりである．

- 国内外のフォロワーマーケットに対する実態調査により，フェイクアカウントに固有な特徴を明らかにした．
- 実態調査により明らかにした特徴を用いることで約 1,260 万の未知のアカウント集団アカウントからフェイクアカウントおよびそれらが形成するコミュニティを発見可能であることを実証した．

1.3 研究に関する倫理

本研究ではフォロワーマーケットからフェイクアカウントを購入し，その特徴を分析した．またデータセットとしてアカウント情報を大量に収集した．このような手法に関する研究倫理の基本的な考え方として，本研究では J. Song らによる先行研究 [6] を参考に以下の点に留意することで倫理上の問題に対応した．倫理的な見解について詳しくは巻末にて述べる．

1.4 論文の構成

本論文は以下の構成となっている．

第一章 序論

はじめに本研究の背景を述べる．さらにその背景をもとに本研究の目的を述べる．

第二章 フォロワーマーケットの概要

フォロワーマーケットの概要と先人達が行ってきた先行研究について述べる．

第三章 データセット

本研究で使ったデータセットについて述べる．

第四章 国内マーケットの調査

国内マーケットの実態について述べる．

第五章 海外マーケットの調査

海外マーケットの実態について述べる．

第六章 フェイクアカウント抽出手法

本研究のフェイクアカウント抽出手法を示す．

第七章 提案手法評価結果と考察

提案手法を未知のアカウント集団に対して適応した結果と考察について述べる．

第八章 まとめ

本研究のまとめと今後の展望について述べる．

第九章 研究業績

研究業績の一覧を示す．

第 2 章 フォロワーマーケット

第 2 章ではフォロワーマーケットの概要，脅威となる可能性，また Twitter 社側の対策について述べる．

2.1 フォロワーマーケットの概要

フォロワーマーケットは現在世界中で存在しており，特に Twitter や Facebook を始めとしたマイクロブローギングサービスを対象とするマーケットは非常に大きなフォロワー産業となっている．また Twitter におけるフォロワーマーケットは次の二つに大きく大別できる．

- 顧客が業者にお金を支払いフォロワーを購入する．
- 顧客がフォロワーに一定のお金を提示し，その金額に同意した働き手がフォローする．

前者は主にウェブサイト上においてある単価数のフォロワーに対してお金を支払うことにより，後日顧客のアカウントにフォロワーが付随する．一方で後者はウェブサイト上で自身が欲しい量のフォロワーに関して，一定の単価を指定し，その金額に納得した働き手が顧客のアカウントをフォローする仕組みである．これらの仕組みは年々大きくなり，現在 Twitter におけるフォロワーマーケットは全世界で数百万ドルに及ぶと言われている [7]．ここで一般的なウェブサイト型フォロワーマーケットを図 2.1 で示す．このようなフォロワーマーケットは検索エンジンにおいて "buy twitter follower" と検索するだけで，日本のみならず世界中の購入可能なウェブサイトを見つけることができる．またこれらのウェブサイト経営者は自身のフォロワーマーケットの顧客を増やす目的で Twitter 内においても商売活動を行っている．その例を図 2.2 に示す．この図 2.2 に示したようにフォロワーマーケットの規模が多いほどその宣伝活動規模も大きくなることが知られている [3]．

ここで Twitter において新規にアカウントを登録する際には CAPTCHA を解読する必要や登録元の IP アドレスが同じ場合 Twitter 社側でブロックされる可能性がある．しかしフォロワーマーケットの経営者はこれらの防衛策を巧みに掻い潜る．例えば CAPTCHA 解読に関しては既に自動で解読するツールが存在しており [8]，IP アドレスの問題に関しては Twitter 社側の IP ブラックリストに入らないよう身元の IP アドレスを登録毎に詐称している．このようにフォロ

ワーマーケットの経営者達は巧みに Twitter の防衛策を潜り抜け、大量のフェイクアカウントを世界中で生成している。さらに現在これらのフェイクアカウント作成から販売までの流れは 1 つのビジネスの流れとして成立している [9]。

Devumi Social Media Marketing

TWITTER YOUTUBE SOUNDCLOUD ALL SERVICES

Buy Twitter Followers Now

Get High-Quality Followers, Automatic Retweets & Sponsored @Mentions

1k Follower Plan	2.5k Follower Plan	5k Follower Plan	10k Follower Plan
\$12	\$24	\$42	\$74
1,000 Twitter Followers	2,500 Twitter Followers	5,000 Twitter Followers	10,000 Twitter Followers
2 ~ 3 Days	3 ~ 5 Days	5 ~ 7 Days	5 ~ 7 Days
Order Now	Order Now	Order Now	Order Now

100% Risk-Free! 1-Year Retention Guarantee

AMERICAN EXPRESS VISA MasterCard DISCOVER JCB Diners Club PayPal Bitcoin amazon

MONEY BACK GUARANTEE TOP RATED

図 2.1 ウェブサイト型のフォロワーマーケットの例

Quick way to develop a credible online business, **buy Real Twitter Follower** at bit.ly/1HG9MQV

20% OFF Get 20 % OFF on your first purchase*
Coupon Code: **FOFF20**

FollowerSale.com

Facebook Likes Instagram Services Twitter Followers

FAME SERVICE
Automatic Retweets & Favorites to your each Tweet

We will start instagram and Facebook services on October 2015
Get Automatic Retweets and Favorites to your tweets with Fame Service

Toll Free Available 9pm to 9pm CDT
888-991-3035
followersale

図 2.2 twitter 内におけるフォロワーマーケットの商売活動の例

2.2 フェイクアカウントの脅威

2.1 節で述べたように現在 Twitter 内には大量のフェイクアカウントが存在している。一方で Twitter 社がフェイクアカウントの作成を禁止しているものの [10]，このフェイクアカウント単体では悪影響はほとんどない。しかし 1 度フェイクアカウントが顧客の元に売られた場合に脅

威が生じる恐れがある。フォロワーを購入した顧客がフィッシングサイトやマルウェア配布サイトを含んだツイートを発信すると、多くの正規ユーザーはその URL にアクセスしてしまう。これは顧客のメタデータが第一印象では信用できししまうことや、短縮 URL に自動変換されることが原因である [5]。また選挙やデモといった政治イベントの際ソーシャルネットワークサービスを利用する人々や団体も世界中で増えてきている。その際フェイクアカウントを大量に購入し、見かけ上の名声を高め、正規ユーザーからの賛成も得た場合、本来の世論とは違う考えが台頭してしまう可能性が考えられる [4]。このようにフォロワーマーケットにより多くの正規ユーザーが常に脅威にさらされている。

一方 Twitter 社側もこれらの脅威に多くの被害を受けている。その代表例としてサービス低下による広告収入の減少が挙げられる [11]。Twitter 社の最も大きな収入は広告収入とされている。またその広告主は Twitter のタイムライン上にプロモーションとして出現し、プロモーションがアクセスされた場合やリツイートされた場合などに Twitter 社側に広告料が入る。このプロモーションの例を図 2.3 に示す。ここで Twitter 内にフォロワーマーケットで販売されるフェイクアカウントが蔓延した場合、サービスの品質は大きく低下する。また正規ユーザーが脅威に晒される機会が増えた場合プロモーションサイトへのアクセスレートも下がることが予想される。このようにフォロワーマーケットの台頭により、Twitter 社側のサービスの継続が不可能になる可能性も考えられる。



図 2.3 プロモーションの例

2.3 フォロワーマーケットに関する関連研究

前節で述べたように Twitter におけるフォロワーマーケットの台頭は大きな脅威と成りうる可能性がある。このため先人達は様々な方法でこれらのマーケットを検出し、凍結する方法を示

しており，本節ではその例を順に示していく．

1 つ目は Stringhini ら [9] の研究である．彼らは 4 つのオンラインソーシャルネットワークからログイン情報，アカウント，タイムスタンプを抽出し，その関係性からそれぞれのサービス内から悪質なコミュニティを検出する方法を指名している．さらに彼らはラベル付きの正解データを元に未知のアカウント集団から悪性のコミュニティを検出することにも挑戦している．

二つ目も Stringhini ら [3] の研究である．1 つ目の先行研究と異なり，Twitter のフォロワーマーケットに特化した研究となっている．彼らはフォロワーマーケットで実際に購入する顧客を Twitter 内から検出する方法を示している．対象としてフォロワーマーケットは正規ユーザーが不正アクセス，またはアカウントに対する認証を許可することにより乗っ取られてしまったアカウントがお金を支払った顧客をフォローするマーケットである．彼らは乗っ取られたアカウントの特徴を使うことで，実際に Twitter 内に存在する顧客と疑われるアカウントを数週間監視し，そのアカウントが顧客か否かを検出する手法を提案している．

3 つ目は Thomas ら [12] の研究である．彼らはフォロワーマーケットで販売されているフェイクアカウントを大量に購入し，その特徴を元に検出する方法を提案している．彼らの研究では非常に多くのマーチャントから長期間購入しており，フォロワーマーケット全体を網羅している研究である．また Twitter 社とも共同でフォロワーマーケットの研究をしており，提案手法により検出されたアカウントの凍結にも一役買っている研究となっている．

4 つ目は Song ら [6] の研究である．彼らはフォロワーマーケットの中で，お金を支払いフォロワーを購入するもの，フォロワーが欲しいアカウント同士のコミュニティである crowdturfing service の二つのマーケットに注目している．彼らはこの二つのマーケットに存在するフォロワーを実際にトラッキングし，その特徴を抽出することに着眼点を置いている．また crowdturfing service ではクリックサービスも提供されており，そのサービスを利用するユーザーやどのような URL に対するクリック要求が多いのかといった分野にも着眼点が置かれている．

2.4 Twitter 社側の対策

Twitter 社は正規ユーザーの脅威となるフェイクアカウントを現在規則として禁止している [10]．またフェイクアカウント作成に対しては IP ブロックリストや CAPTCHA を用いることで，自動生成にある程度のフィルタリングをしている [12]．一方フェイクアカウントがボットのように自動的にツイートを発信することや，制限なく自動的に正規アカウントをフォローすることができないよう，Twitter 社側では API 制限を設けている [13]．使用回数については使用する API に依存するが，15 分に 1 度 API 制限がリセットされる仕様になっている．また API 生成時においても電話番号認証を求められ，大量の API キー生成対策は施されている．さらに前節の関連研究で述べたように，大学などの学術機関と共同でフォロワーマーケット対策に乗り出しているケースもある [12]．このように Twitter 社側でも様々な手段を講じることでフォ

ロワーマーケットの撲滅に取り組んでいる。

第 3 章 データセット

第 3 章では使用したデータセットの概要について述べる．それぞれのデータセット数は表 3.3 として示す．

3.1 データセットの概要

3.1.1 正規アカウント

本節では収集した正規アカウントを説明する．正規アカウントを集めるにあたり，フォロワーランキングサイト [14] を用いた．このウェブサイトではフォロワー数に応じてアカウントのランキングが公開されており，上位ほど多くのフォロワーを所持し，下位ほど少ないフォロワーを所持している．本研究ではランキング下位 70,000 位から 483 アカウントを抽出し，正規アカウントとした．これらのアカウントは少なくとも 151 のフォロワーを所持していた．下位から正規アカウントを抽出した理由としては上位アカウントは著名人のアカウントであるため，一般の正規アカウントからかけ離れていると考えられるからである．すなわち下位ユーザーであればある程度の正規性を保証でき，かつ一般のアカウントに近い挙動を示すと想定する．

3.1.2 フォロワーマーケットのフェイクアカウント

本研究ではフェイクアカウントとして，国内のフェイクアカウント，海外のフェイクアカウントを収集した．本研究では表 3.1 に示す 2 つのマーケットを対象とし，それぞれ site1，site2 と呼称する．また，site1，site2 からそれぞれ 2,000，1,000 フォロワーを購入した．フォロワーの単価は site1 と site2 で約 2 倍の差がある．これはフォロワーの品質の差異を反映しており，単価が高いほどより見かけが実際のアカウントに近い傾向がある．

海外マーケットの調査にあたり，その信頼性をランキング化しているウェブサイト [17] を参照した．このウェブサイトではランキングが上位であるほどフェイクアカウントが正規アカウントに近い傾向がある．表 3.2 に本研究で利用する 9 つのフォロワーマーケットを示す．

海外におけるフォロワーマーケットは全体的に日本のマーケットに比べて単価が低い傾向がある．これは世界における購入希望者が多いため，単価も下がっていると考えられる．

表 3.1 国内フォロワーマーケットの例．価格はそれぞれ 1,000 フォロワー (円) および 1,000 リツイート (円) ．

マーケット	フォロワー価格	リツイート価格
site1 [15]	900	1,900
site2 [16]	4,000	2,000

表 3.2 海外フォロワーマーケットの例．価格はそれぞれ 1,000 フォロワー (USD) および 1,000 リツイート (USD) ．

マーケット	フォロワー価格	リツイート価格
site3 [18]	9	9
site4 [19]	12	69
site5 [20]	19	19
site6 [21]	12	19
site7 [22]	9	21
site8 [23]	10	30
site9 [24]	5	5
site10 [25]	5	5
site11 [26]	9	21

またフェイクアカウント購入希望者のニーズに対応して，アメリカベースのフェイクアカウント，アラビアベースのフェイクアカウントと購入者に選択肢を与えるウェブサイトも存在した．さらにリツイートに関して見てみると，海外マーケットは日本のマーケットに比べて高い傾向がある．高いレート of マーケットの場合は1ヶ月間かけてゆっくりとリツイート数を増やしていくことや，複数のツイートに対してリツイートを施すことなど多くの選択肢が与えられているため高いと考えられる．一方比較的低いレート of マーケットの場合，1度に1つのツイートに対してリツイートをを行うことや，そのリツイートしたアカウントが正規ユーザーとかけ離れている場合が多く，高いレートに比べ品質が劣る傾向が見られる．本研究では海外マーケット調査のために上記の9つのマーケットからそれぞれ約1,000フォロワーを購入し，それぞれのマーケットを site3～site11 と名付けた．

3.1.3 未知のアカウント集団

本節では提案手法の有用性を示すために使用した未知のアカウント集団について述べる．本研究では未知のアカウント集団を有名人，或いは著名人をフォローしているアカウントを無作為に抽出することで作り出している．様々な著名人を利用している理由として，多くの著名人や有名人のアカウントは望んでいない場合でもフェイクアカウントが followee の数を増やす目的でフォローしている可能性が高いためである．また著名なアカウントの場合相互フォローを推奨している場合も多く，その目的でフェイクアカウントがフォローしている可能性も存在す

る．本研究では無作為に著名人や有名人のアカウントのフォロワーをクロールし，12,605,718の未知のアカウント集団を抽出した．

表 3.3 使用データの概要

データセット名	アカウント数
良性データ	483
悪性データ	12,058
未知のアカウント集団	1260,5718

第 4 章 国内マーケットの調査

第 4 章では国内フォロワーマーケットから購入したフェイクアカウントの実態調査結果について述べる。

4.1 国内マーケットの実態調査

4.1.1 国内マーケットの概要

国内マーケットの実態調査にあたり，第 3 章で記述した site1，site2 から約 1000 フォロワーずつ購入した．またフォロワーマーケットの検索方法にはフリーランスやウェブブラウザに検索クエリを投げる方法などあるが，今回ウェブブラウザをに検索クエリを投げることで複数のフォロワーマーケットを見つけることができた．それぞれのマーケットが提供しているサービスの調査結果を表 4.1 に示す．

表 4.1 国内フォロワーマーケットのサービス調査結果

マーケット URI	1000 フォロワーの単価	1000 リツイート単価	Twitter アカウント購入の可否	その他のサービスの可否	購入者数
site1 [15]	900 円	1900 円	可能	可能	300/month
site2 [16]	4000 円	2000 円	可能	可能	N/A

本研究では 2 つの国内マーケットをクエリを投げることで見つけることができた．表 4.1 の Twitter アカウントの購入の可否については，フォロワーの購入ではなく，アカウントそのものが購入できるか否かを表している．またその他のサービスの可否については"Facebook"や"Instagram"といったその他のマイクロブロギングサービスにおけるフォロワーの購入やアカウントの購入ができるか否かを表している．さらに購入者数では site1 では月 300 件程度の申し込みが様々なユーザーから依頼されると書かれていた．一方 site2 に関して購入者数は述べられていなかったが，site2 へのアクセスは世界中からされており，国内のみならず世界中で日本のフェイクアカウントが必要とされていることが分かった．

4.1.2 followee の特徴

本次節でフォロワーマーケットのフェイクアカウントが所持している followee の特徴について述べる．前節で述べたように本研究では日本に存在する 2 つのマーケットからフェイクアカウントを購入した．その followee 数を次の図 4.1 として示した．

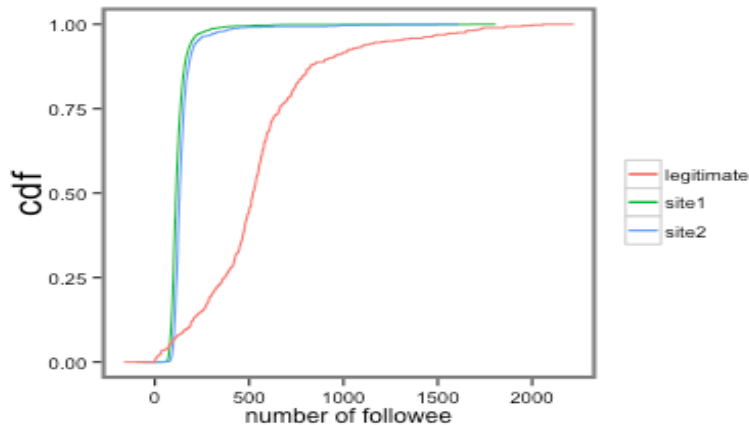


図 4.1 フェイクアカウントの followee 数

このように followee 数に関しては 500 付近にほぼ全てのフェイクアカウントが収束する．これは [10] にも示されているように、急激な followee 数の増加や follower 数に対する極端な多さはアカウント凍結の原因に成りうる可能性があるため、followee 数に限界があると考えられる．またマーケットの経営者もある程度人間味が感じられれば、それ以上人間味を持たせる必要がなく、必要以上の労力をかけることを避けているためにこのような特徴が生じると考えられる．

4.1.3 フォロワーの特徴

本次節でフェイクアカウントが所持しているフォロワーの特徴について述べる．その followee 数を次の図 4.2 として示した．

followee 数に比べフェイクアカウントのフォロワー数は非常に少なく、最大でも 50 程度である．これはフェイクアカウントに対してフォローリクエストを申請するアカウントが少ないためと考えられる．フェイクアカウント同士で相互にフォローすることも可能であるが、前節で述べたような followee 数増加の制限により、そのような運用はなされていないとみられる．また中央値にも正規ユーザー、フェイクアカウントの間には大きな差が見られ、正規ユーザー、site1、site2 の順に 470、4、3 となった．

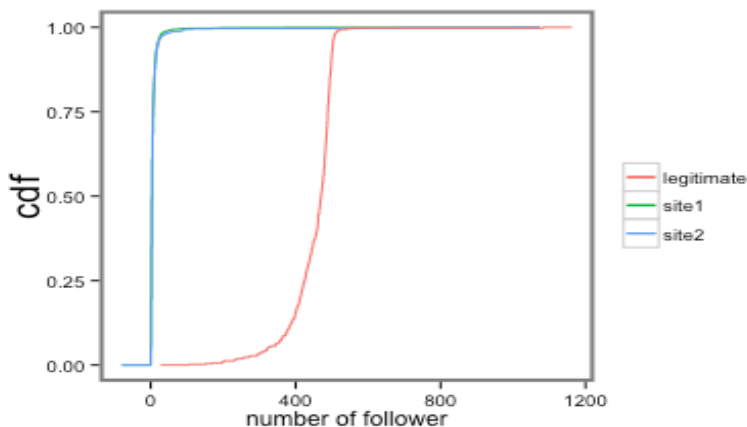


図 4.2 フェイクアカウントの follower 数

4.1.4 フォロー・被フォローのバランス

本節ではフェイクアカウントの特徴として、フォロー・被フォローのバランスを分析する。尺度として ffrate (followee-follower-rate) を定義し、以下のように計算する。

$$\text{ffrate} = \frac{\mathbf{R}(\text{follower})}{\mathbf{R}(\text{followee}) + \mathbf{R}(\text{follower})} \quad (4.1)$$

分母はあるアカウントが所持している follower と followee の和を取っており、分子の follower との割り算の結果を ffrate をしている。この ffrate により、次の二つのことが考えられる。

- ffrate が 0 に近づく時、 $\text{followee} \gg \text{follower}$ となりフェイクアカウントの可能性が高い。
- ffrate が 1 に近づく時、 $\text{follower} \gg \text{followee}$ となり正規ユーザーの可能性が高くなる。

以下にフェイクアカウントが所持している followee と follower から算出した ffrate を累積分布関数として示した。

図 4.3 からわかるように、両マーケット共に ffrate=0.2 付近で全て収束している。このことから多くのフェイクアカウントは正規ユーザーと懸け離れた ffrate を持っていることが分かる。またこの結果からもマーケットの経営者は必要以上にフェイクアカウントに対して人間味を持たせる労力を掛けないということが考えられる。また中央にも特徴は顕著に見られ順に 0.46, 0.03, 0.02 となった。

4.1.5 klout score の特徴

本次節では klout score について述べる。klout score とはソーシャルネットワーク上の影響力を数値として表すものであり、0~100 の間で表され、影響力が大きければ 100 に近づき、影響力が小さければ 0 に近づく。またこの klout score は 35 種類以上の変数を「ネットワークの規

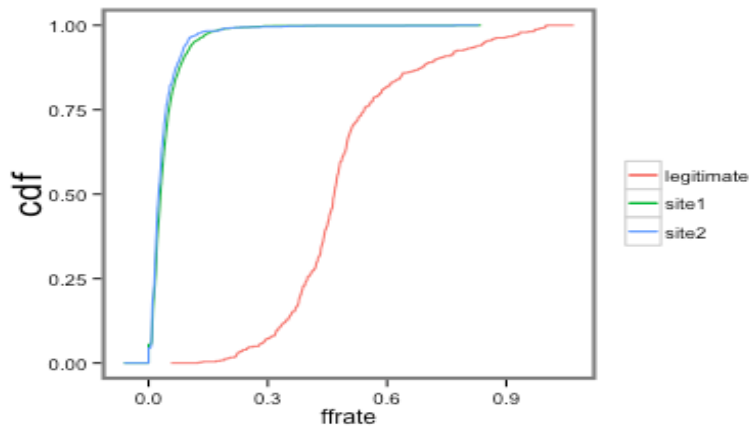


図 4.3 フェイクアカウントの ffrate

模」,「コンテンツのバイラル実績」,「ネットワークの潜在能力」といった3つの指標から算出されている [27] . ここでフェイクアカウントの klout score を図 4.4 として示した .

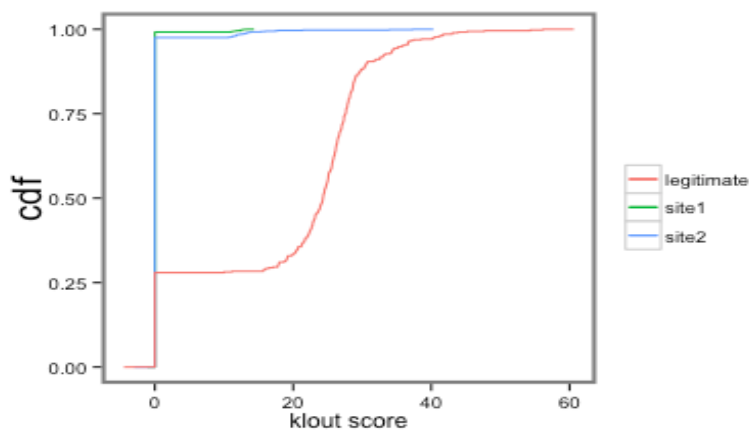


図 4.4 フェイクアカウントの klout score

このように両マーケット共に klout score は 20 に満たず収束している . 一般に klout score のアベレージは 40 程度とされており [28], マーケットに存在するフェイクアカウント自体は非常に影響力の小さいアカウントであると考えることができる . また本研究では実際にフォロワーマーケットからフェイクアカウントを購入したが , フェイクアカウントを購入したアカウント自体の klout score も 0.0 となった . このことからフェイクアカウント自体を大量に購入しても影響力は変わらず , あくまで見かけ上の影響力が増すだけということが分かった . それぞれの中央値は 24 , 0.0 , 0.0 となった .

4.1.6 活動停止時間の特徴

本節ではアカウントのアクティビティを特徴づける指標として、最終活動時刻からどの程度の時間、活動を停止しているかを分析する。活動停止時間を $\tau = t - m$ と定義する。ここで t はデータ取得時の時刻を表し、 m はフェイクアカウントが最後に発信したツイートの時刻を表す。またフェイクアカウントが1度もツイートを発信していない場合はフェイクアカウントが生成された時刻を m とする。活動停止時間 τ の分布を図 4.5 に示す。

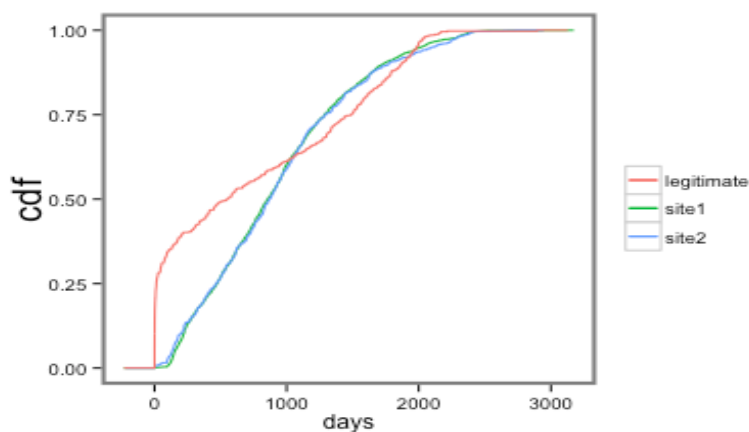


図 4.5 フェイクアカウントの活動停止時間

横軸は活動停止時間の日数を掲示している。図に示されるようにフェイクアカウントの大多数はアカウント生成時にわずかなツイートをしたのち、その後は活動をしない。したがって活動停止時間は大きい。一方僅かながら活動停止時間が非常に小さいフェイクアカウントが存在する。これは比較的新しいフェイクアカウントのみを販売している業者である可能性があり、マーケットの品質を保つためと考えられる。またそれぞれの中央値は 549, 856, 871 となった。

第 5 章 海外マーケットの調査

第 5 章では海外マーケットの調査結果について述べる．章内の構成は第 4 章と同じである．

5.1 海外マーケットの実態調査

5.1.1 海外マーケットの概要

海外マーケットの調査にあたり, 今回海外マーケットのランキングを載せているウェブサイト
を参照した [17]．このウェブサイトはフォロワーマーケットの信頼性を元にランキング付けさ
れており, ランキング上位であるほどフォロワーが実際の正規アカウントに近い傾向がある．以
下に実態調査のために利用したフォロワーマーケットを表 5.1 として示す．

表 5.1 国内フォロワーマーケットの例

マーケット URI	1000 フォロワーの単価	100 リツイート単価	Twitter アカウント購入の可否	その他のサービスの可否	購入者数
site3 [18]	9 ドル	9 ドル	不可	可能	N/A
site4 [19]	12 ドル	69 ドル	不可	可能	N/A
site5 [20]	19 ドル	19 ドル	不可	可能	N/A
site6 [21]	12 ドル	19 ドル	不可	不可	N/A
site7 [22]	9 ドル	21 ドル	不可	可能	N/A
site8 [23]	10 ドル	30 ドル	不可	可能	N/A
site9 [24]	5 ドル	5 ドル	不可	可能	N/A
site10 [25]	5 ドル	5 ドル	不可	可能	N/A
site11 [26]	9 ドル	21 ドル	不可	可能	N/A

海外におけるフォロワーマーケットは全体的に日本のマーケットに比べて単価が低い傾向があ
る．これは世界における購入希望者が多いため, 単価も下がっていると考えられる．ま
た世界中のフェイクアカウント購入希望者のニーズに対応して, アメリカベースのフェイクア
カウント, アラビアベースのフェイクアカウントと購入者に選択肢を与えるサイトも存在した．さ
らにリツイートに関して見てみると, 海外マーケットは日本に比べて高い傾向がある．比較的高
いレートの場合は 1 ヶ月間かけてゆっくりとリツイート数を増やしていくことや, 複数のツイート

に対してリツイートを施すことなど、多くの選択肢が与えられているため、高いと考えられる。一方比較的低いレートのマーケットの場合、1 度に 1 つのツイートに対してリツイートを行うことや、そのリツイートしたアカウントが正規ユーザーとかけ離れている場合が多く、高いレートに比べ品質が劣る傾向が見られる。

本研究では海外マーケットのために上図の 9 つのマーケットからそれぞれ表 ?? 相当のフェイクアカウントを購入した。また今後のためにそれぞれ購入したマーケットを site3 ~ site11 と名付ける。さらにグラフをマーケットの特徴を元に 2 つに分けて示している。

表 5.2 海外フォロワーマーケットの購入履歴

マーケット名	購入単価
site3	9 ドル相当
site4	12 ドル相当
site5	19 ドル相当
site6	12 ドル相当
site7	19 ドル相当
site8	10 ドル相当
site9	5 ドル相当
site10	5 ドル相当
site11	9 ドル相当

5.1.2 followee の特徴

本次節でフォロワーマーケットのフェイクアカウントが所持している followee の特徴について述べる。前節で述べたように本研究では海外の 9 つのマーケットからフェイクアカウントを購入した。その followee 数を図 5.1 として示した。

海外フォロワーマーケットのフェイクアカウント followee 数は日本のフェイクアカウントに比べ多い傾向があり、1 番少ない site3 のフェイクアカウントについても最大値が 1,000 程度となり、日本のフェイクアカウントの 2 倍程度所持している。また最も多い followee を所有しているフェイクアカウントは 10,000 程度と非常に大きな傾向が見られる。このように海外のフェイクアカウントの followee 数は凍結される上限値に近い傾向が見られる。

5.1.3 フォロワーの特徴

本次節で海外フォロワーマーケットのフェイクアカウントが所持しているフォロワーの特徴について述べる。そのフォロワー数を次の図 5.2 として示した。

海外フェイクアカウントのフォロワー数は非常に少なく、フォロワー数が 100 人程度でほぼ

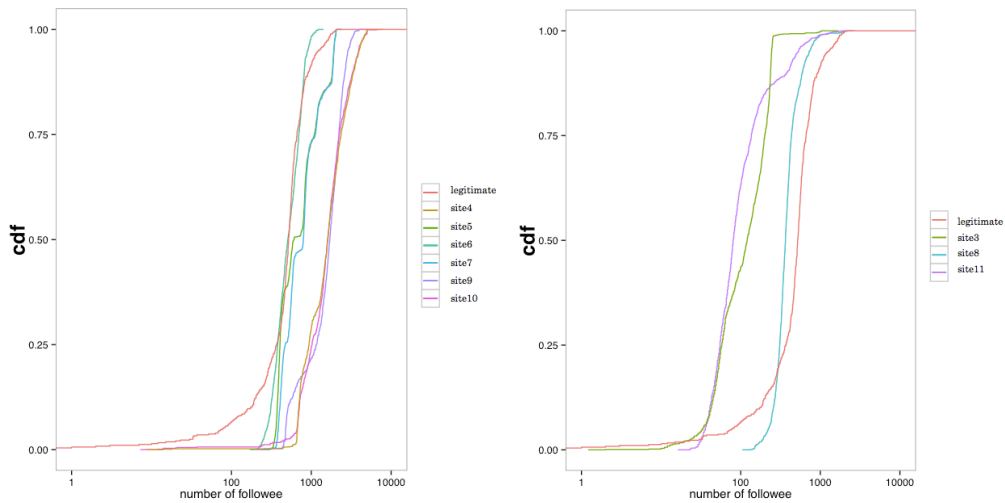


図 5.1 フェイクアカウントの followee 数

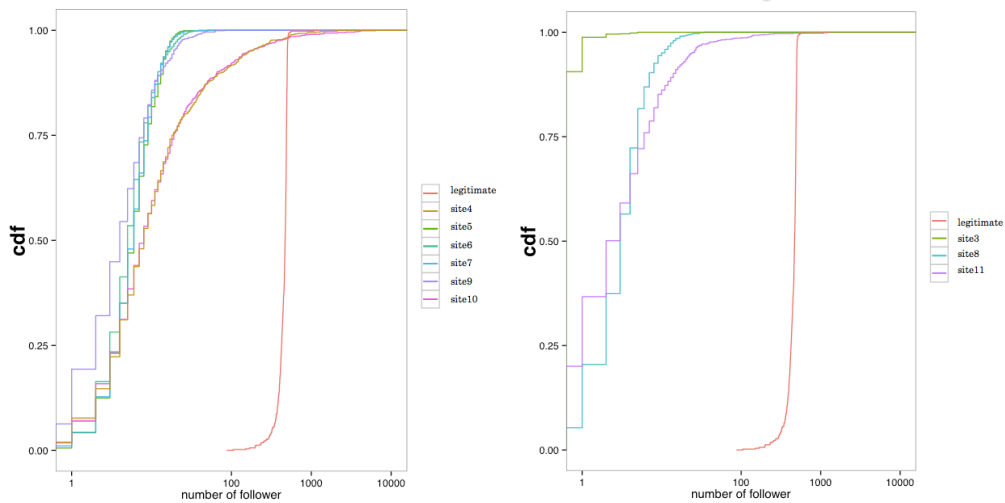


図 5.2 フェイクアカウントの follower 数

収束している．このことから国内マーケット同様フェイクアカウント同士の相互フォロー等で follower 数を増やすといった方法を取っていないことが推測できる．また中央値にも正規ユーザー，フェイクアカウントの間には大きな差が見られ，正規ユーザー，site3～site11の順に 470，3，6，5，6，4，2，8，8，0 となった．

5.1.4 フォロー・被フォローのバランス

本次節では海外フェイクアカウントのフォロー・被フォローのバランスに関して述べる．フォロー・被フォローのバランスについての定義は前節で説明したものと同じである．以下にフェイクアカウントが所持している followee と follower から算出した ffrate を累積分布関数として示した．

図 5.3 からわかるように，海外フェイクアカウントは $followee \gg follower$ となっている傾

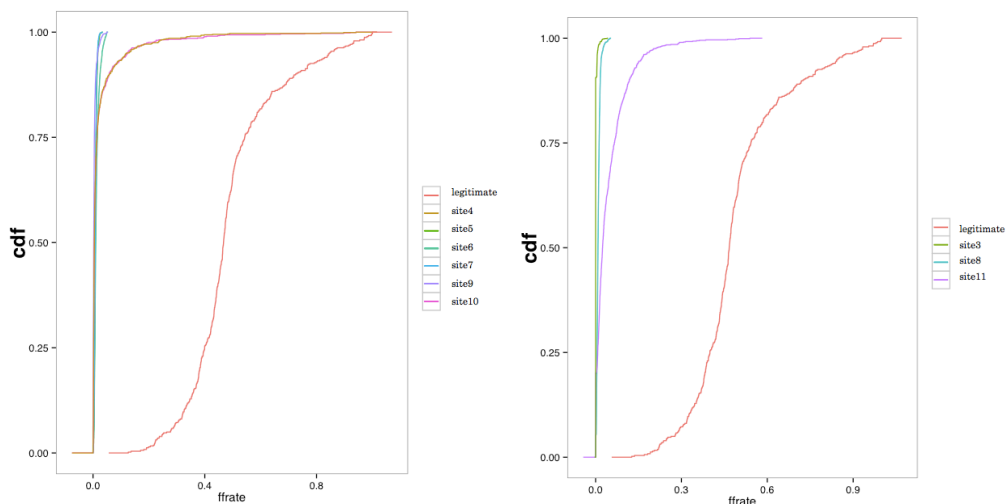


図 5.3 フェイクアカウントの ffrate

向が強いことから，ffrate も非常に小さい傾向が見られた．これは海外フェイクアカウントが所持している followee 数の多さが原因と考えられる．またこの特徴は中央値にも顕著に表れ，site3～site11 の順に 0.008，0.0，0.0，0.0，0.0，0.024，0.0045，0.0045，0.0 となった．

5.1.5 klout score の特徴

本次節では海外フェイクアカウントの klout score について述べる．klout score の定義は国内マーケットと同じであり，以下にその分布を示した．

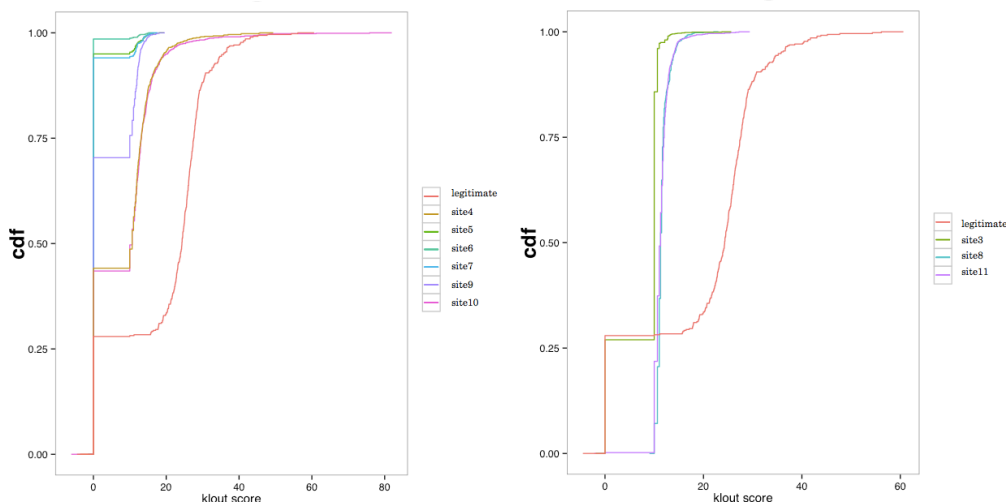


図 5.4 フェイクアカウントの klout score

海外フェイクアカウントも国内マーケット同様に Klout score が 10～15 程度となりソーシャルネットワーク上の影響力が低いアカウントが多数占めている．このように多くの followee を所持し見かけ上の正規ユーザーを作り出しているが，実際の影響力はほとんど皆無となってい

る．また図 5.4 から業者毎にアカウントのメンテナンス状態が異なることが分かる．具体的に site5 のようなほとんど影響力のないフェイクアカウントを作り出している業者もあれば，site3 のようにある程度正規ユーザーに近いアカウントを作り出している業者が存在する．業者毎のメンテナンス特徴は中央値からも読み取れる．site3 ~ site11 の順に 11.2 , 0.0 , 0.0 , 0.0 , 0.0 , 11.28 , 10.84 , 10.64 , 10.0 であった．

5.1.6 活動停止時間の特徴

本次節では活動停止時間について述べる．活動停止時間の定義は国内マーケットと同様である．以下にその分布を示した．

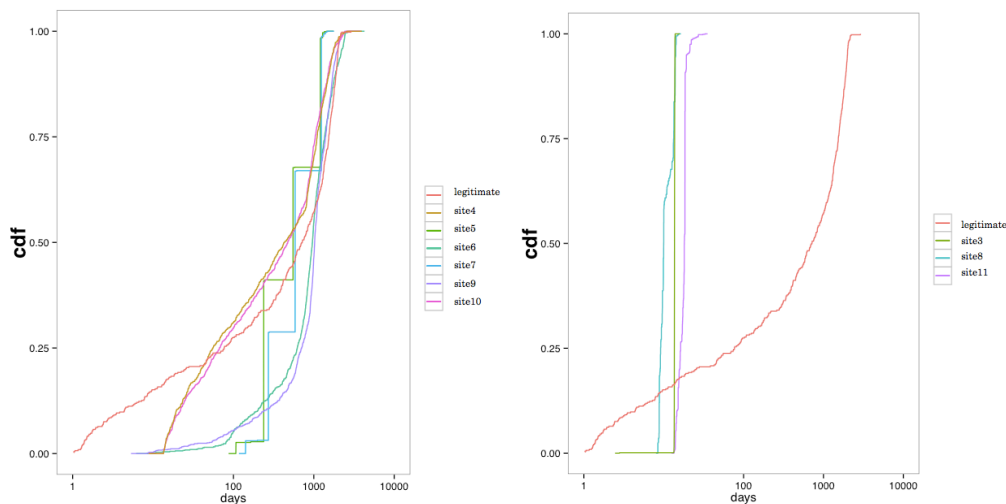


図 5.5 フェイクアカウントの活動停止時間

海外マーケットは．マーケットごとに異なる特徴が見られる．例えば site4 や site6 のように累積分布関数が階段状になる場合，ある期間に大量のフェイクアカウントを作り出し，同時に活動をさせている可能性が推測できる．一方 site3 の場合，活動停止時間が正規ユーザーと比較しても短いアカウントが存在することが分かる．このように海外マーケットの場合も Klout score 同様に業者毎にフェイクアカウントのメンテナンス状態が異なり，その結果差が生じることが分かった．この差に関しては中央値を見ても明らかであり，site3 ~ site11 の順に 9.9 , 550 , 960 , 585 , 1022 , 18.11 , 473.35 , 423.90 , 13.62 となった．

第 6 章 フェイクアカウント抽出手法

第 6 章では第 5 章で得た特徴を元にフェイクアカウント抽出手法を示す。

6.1 フェイクアカウント抽出手法の概要

6.1.1 ffrate フィルタリング

始めに未知のアカウント集団から，正規のアカウントである可能性が高いアカウントを除去する．前節で示したようにフェイクアカウントの ffrate は正規アカウントと比較して大幅に低くなる特徴を利用する．経験的に ffrate の閾値として 0.05 を採用した．この数値は実際のフェイクアカウントにおける中央値よりもやや大きな数値であり，この段階でフィルタリングされるアカウントの多くはフェイクアカウントであることが期待される．

6.1.2 機械学習によるフィルタリング

次に機械学習を用いてフェイクアカウントと正規アカウントを弁別する．前述したようにこの時点で残っているアカウントの多くは正規アカウントではないことが想定されるが，残りのアカウントを精度良く弁別することが目的である．本研究では機械学習アルゴリズムとして Support Vector Machine (SVM) [29] を採用した．3.1.1 節，3.1.2 節で説明したはじめに正規アカウントおよびフェイクアカウントのデータを教師データとして用い，SVM の分類器を訓練する．特徴量として followee 数，フォロワー数，ffrate，Klout score，活動停止時間の 5 つを採用した．SVM の訓練では 10-fold Cross validation によりを用い，パラメタ空間をグリッドサーチすることによって最適なパラメータ値を抽出した．その結果を表 6.1 として示す．FPR，FNR とともに低い数値を得ている．ここで FPR は正規アカウントを誤ってフェイクアカウントと判定した率，FNR はフェイクアカウントを誤って正規アカウントと判定した率である．本研究では 1 アカウントのみ正規と誤判定されたフェイクアカウントが存在したが，これは実際に目を見た結果，正規であるアカウントが高いことが判明した．すなわち正規のアカウントが本研究で使用したアカウントにフォロワーとしてついてしまった可能性が高い．以降の実験ではこのアカウントを排除した上で再度 SVM を訓練して実験を行った．

表 6.1 10-fold Cross Validation による SVM の訓練結果

最適パラメタ値	FPR	FNR
$C = 100, \gamma = 0.0001$	0.00	0.01

6.2 Jaccard 係数

本提案手法ではコミュニティ検出手法の際に Jaccard 係数を用いた．この Jaccard 係数の定義を以下に記す．

$$J(A, B) = \frac{|A(\text{followee}) \cap B(\text{followee})|}{|A(\text{followee}) \cup B(\text{followee})|}$$

A, B はあるアカウントを表している．また A(followee), B(followee) はそれぞれのアカウントが所持している followee を表している．それぞれの followee の共通部分を和集合で割ることで，Jaccard 係数を算出しており， $0.0 \leq J(A, B) \leq 1.0$ の範囲で変動する．この Jaccard 係数を用いることで共通の followee が多い場合，そのペアのアカウントは共通のコミュニティが高いことが推測できる．

前述の機械学習フィルタリングによってフェイクアカウントと判定されたアカウントを対象に Jaccard 係数を用いることでコミュニティ検出を試みる．本研究の制限として，フェイクアカウントと判定されたアカウント全ての followee を調べることは API 制限の関係で困難であったため，検出したフェイクアカウントから約 1 割をランダムサンプリングし，コミュニティの検出を試みた．

最後に本研究のフェイクアカウント抽出手法概要を図 6.1 として示した．

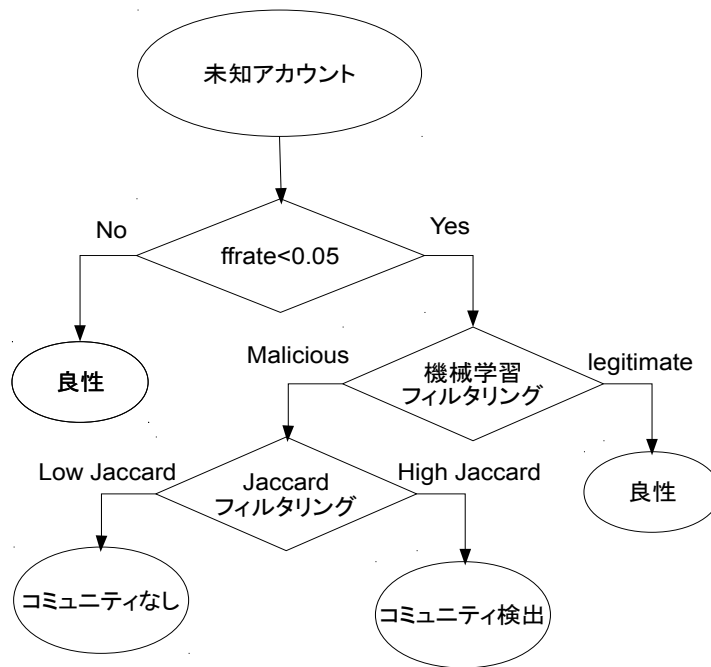


図 6.1 提案手法のフローチャート

第 7 章 提案手法の評価結果

第 7 章では提案手法の評価と考察について述べる．

7.1 ffrate フィルタリングの考察

本提案手法では大規模な未知のアカウント集団から明らかに正規ユーザーを減らす目的で ffrate フィルタリングをデータセットに対して施した．ffrate フィルタリング後のデータセット数を表 7.1 として示した．このように ffrate フィルタリングを用いることでデータ数を予め削減することで効率良くデータの解析を進めることができた．またフィルタリングされ，良性とみなされたデータについては，フォロワーマーケットの実態調査により明らかにした ffrate を参考に閾値を広く取っているため，フェイクアカウントが残っている可能性は低いと推察できる．

表 7.1 ffrate フィルタリングの実行結果

データセット名	総データ数	ffrate フィルタリング後のデータセット数	削減率
未知のアカウント集団	12,600,000	319,985	98%

7.2 SVM によるフィルタリング

データセットを ffrate により削減したのち，本研究では SVM により悪性，良性のアカウントを分類した．その結果を表 7.2 として示した．このように ffrate フィルタリング後に分類器にかけることで僅かながら良性アカウントを抽出することができた．この結果からも ffrate が大きくフィルタリングの役割を果たしていることが推察できる．また良性と判定されたアカウントの中にはややフェイクアカウントと疑われるアカウントも紛れ込んでいた．これは SVM の特徴量評価の際，followee 数，klout score，time span の 3 つが訓練データ内の良性アカウントに近い性質をもっていたためと考えられる．またこの特徴をもつアカウントは複数良性と誤判定されたが，全てトップ画像，スクリーンネームといった部分がかなり似たものとなっており，良性と評価されたアカウントからもフェイクアカウントのコミュニティと考えられる集団を発

見することができた．

表 7.2 SVM による分類結果

未知のアカウント数	ffrate 良性アカウント数	悪性アカウント数
319,985	334	319,651

7.3 Jaccard 係数によるコミュニティ検出

最後に Jaccard 係数を用いてフェイクアカウントのコミュニティを検出した結果を示す．アカウントのペアに対して Jaccard 係数を算出した結果を図 7.1 に示す．図より大多数のペアは低い Jaccard 係数であることがわかる．以下では Jaccard 係数の大小に応じて観測されたアカウントを考察する．

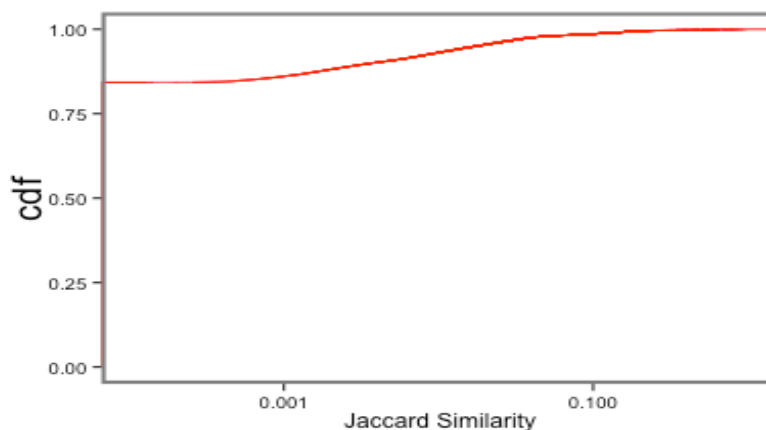


図 7.1 Jaccard 係数の分布

7.3.1 Low Jaccard 係数の考察

本節では Low Jaccard 係数について考察する．Low Jaccard 係数は $J(A, B) < 0.2$ までの範囲とする．この範囲に入るアカウントのペアはほとんど関係のないコミュニティに属していると推察できる．また分類器では悪性と判定されたが，正規ユーザーのサブアカウントと考えられるアカウントを排除することができる．このようにフェイクアカウントに近い良性アカウントを検出することができる．

一方で異なるコミュニティに属するフェイクアカウントを検出できる可能性が存在する．この考察に関しては断定できるアカウントのペアを見つけることはできなかったが，可能性があるという意味で記述しておく．

7.3.2 Middle Jaccard 係数の考察

本節では Middle Jaccard 係数について考察する．Low Jaccard 係数は $0.2 \leq J(A, B) < 0.5$ までの範囲とする．この範囲に入るアカウントのペアは悪性アカウントのペア，良性アカウントのペア両方が混在していると考えられる．前者に関しては既に同一コミュニティのフェイクアカウントは同じアカウントをフォローする傾向が高いことが特徴として分かっている．一方後者に関してはあるジャンルの著名なアカウントのみをフォローしている正規ユーザーのサブアカウントの場合，Jaccard 係数は比較的高くなる可能性が考えられる．このように多くの悪性アカウントコミュニティを検出することができる一方，一部の良性アカウントのコミュニティを検出できることが分かった．以下に前者，後者の例をそれぞれ図 7.2，図 7.3 と示した．



図 7.2 フェイクアカウントのペアと考えられる Middle Jaccard の例



図 7.3 正規ユーザーのペアと考えられる Middle Jaccard の例

7.3.3 High Jaccard 係数の考察

本節では High Jaccard 係数について考察する．Low Jaccard 係数は $J(A, B) \geq 0.5$ までの範囲とする．この範囲に存在するアカウントのペアは全体のペア数に比べて非常に少ないが，フェイクアカウントのコミュニティを検出できると考えられる．本研究では生成日が最も古いアカウントを起点として Jaccard 係数が 0.5 以上のアカウントが 10 個以上存在した場合，それらのアカウント群を 1 つのコミュニティとしてみなす．この条件で重複を排除してコミュニティを検出した結果，4,432 個の異なるコミュニティが検出された．その内大きさが最大のコミュニティは 716 のアカウントが存在した．これらの検出されたコミュニティ内のアカウントを 100 ペアサンプリングし，Twitter のアカウントにアクセスして分析した結果，実際にフェイクアカウントであることが確認できた．網羅的な調査は出来ていないが，抽出したコ

コミュニティは高い確率で同じ業者によるオペレーションで生成されたフェイクアカウントであることが予想される．また同一コミュニティ内においても Jaccard 係数の違いが見られる．これは発信しているツイート内容が類似しているほど Jaccard 係数も 1 に近づくことが確認できた．ここで同一コミュニティ内で Jaccard 係数が異なる例を 7.4, 7.5 として示す．



図 7.4 同一コミュニティ内のペア (=0.6) の例



図 7.5 同一コミュニティ内のペア (=1.0) の例

一方で本手法では正規アカウントのコミュニティを誤ってフェイクアカウントのコミュニティとして検知してしまう可能性も存在する．これはある集団のみをフォローするアカウントがコミュニティを形成するケースである．今回手動で確認した範囲内ではこのような事例は発見することはできなかったが，可能性としては存在し得ることを付記しておく．以下に同一コミュニティの例を図 7.3 として示す．

表 7.3 同一コミュニティ (size=10) の例

起点アカウント ID	ペアアカウント ID	Jaccard 係数
208880837	357494311	1.0
208880837	2778056594	1.0
208880837	149749030	1.0
208880837	190353856	1.0
208880837	1076420935	1.0
208880837	155768961	1.0
208880837	144335378	1.0
208880837	1185121447	1.0
208880837	489521162	1.0
208880837	1232877540	1.0

第 8 章 まとめ

本論文では Twitter におけるフォロワーマーケットの実態調査を行い、その調査結果をもとに機械学習、Jaccard 係数を利用することでフェイクアカウントのコミュニティを検出できることを示した。

先に行った実態調査ではフェイクアカウントを国外、海外から約 12,000 アカウント購入し正規ユーザーとの差が顕著に表れやすい followee 数、フォロワー数、ffrate、klout score、活動停止時間の調査を行った。結果としてフェイクアカウントのフォロワー数は極端に少なくその結果 ffrate にも正規ユーザーとの差が顕著に表れた。また klout score や活動停止時間に関しては購入したマーケット毎に数値が異なり、マーケット毎に品質管理が異なる特徴が見られた。

次に行った未知のアカウント集団からフェイクアカウントのコミュニティを検出する方法について総括する、先に行った実態調査を元にした特徴量を使った機械学習と ffrate フィルタリングを使うことで、約 30 万の悪性と見られるアカウントを抽出した。さらにこの集団からランダムに 30,000 アカウント抽出し、全てのペアに関する Jaccard 係数を求めた。その結果 $J(A, B) \geq 0.5$ となる高い Jaccard 係数のコミュニティを 4,432 個見つけ、実際に自分の目でアカウントのペアを確認した結果あるゲームやアニメに関する宣伝を行うフェイクアカウントのペアであることを確認した。このように実態調査を元にした特徴量を元に Jaccard 係数を組み合わせることで未知のアカウント集団からフェイクアカウントのコミュニティを検出できることを示した。また未知のアカウント集団からフェイクアカウントを見つけ出す研究は本研究が初めてであり、新規性がある。

最後に今後の課題を述べる。1 つめは SVM による正規アカウント抽出量が少なく、その結果悪性と思われるアカウント集団に正規アカウントが混ざっている可能性があるという点である。これに関しては今後さらなる実態調査を元にした特徴量抽出が必要である。また今回フェイクアカウントが持つ特徴が限りなく正規ユーザーに近い場合、本提案手法ではそのようなフェイクアカウントを検出できない。今回の実態調査では正規ユーザーに近いアカウントを見つけることができなかったが、可能性があるとして示しておく。

第 9 章 研究業績

竹越健斗，孫博，森達哉，Twitter におけるフォロワーマーケットの実態調査 とフェイクアカウントの抽出方法，Symposium on Cryptography and Information Security 2016

謝辞

多くの研究に関するご相談にのっていただいた森達哉准教授に感謝申し上げます。また研究に関する相談、議論に付き合っていた博士後期課程に在籍の孫博さんに感謝致します。さらにディスカッション等に協力してくださった森研究室の皆様にも感謝します。

参考文献

- [1] "ソーシャルメディアのデータまとめ一覧". <http://gaiax-socialmedialab.jp/socialmedia/368>.
- [2] C.Grier, K.Thomas, V.Paxson, and M.Zhang. "the underground on 140 characters or less". In *Proceedings of the 17th ACM conference on Computer and communications security*, pp. 27–37, 2010.
- [3] G.Stringhini, G.Wang, M.Egele, C.Kruegel, G.Vigna, H.Zheng, and Ben Y. Zhao. "follow the green: Growth and dynamics in twitter follower markets". In *City Labs Workshop SocInfo2014*, 2014.
- [4] K.Thomas, C.Grier, and V.Paxson. "adapting social spam infrastructure for political censorship". In *5th USENIX Workshop on Large-Scale Exploit and Emergent Threats*, 2012.
- [5] K.Thomas, C.Grier, V.Paxson, and D.Song. "suspended accounts in retrospect: An analysis of twitter spam". In *Proceedings of the 2011 ACM SIGCOMM conference on Internet measurement conference*, pp. 243–258, 2011.
- [6] J.Song, S.Lee, and J.Kim. "crowdtarget: Taget-based detection of crowdturfing in online social network". In *Proceedings of the 22nd ACM SIGSAC Conference on Computer and Communications Security*, pp. 793–804, 2015.
- [7] PERLROTH N. "fake twitter followers become multimillion-dolloar business", 2013.
- [8] M.Motoyama, K.Levchenko, C.Kanich, D.McCoy, G.M.Voelker, and S.Savage. "captchas-understanding captcha-solving services in an economic context". In *proceedings of the 19th USENIX conference on Security*, p. 28, 2010.
- [9] G.Stringhini, P.Mourlanne, G.Jacob, M.Egele, C.Kruegel, and G.Vigna. "evilcohort: detecting communities of malicious accounts on online services". In *Proceedings of the 24th USENIX Conference on Security Symposium*, pp. 563–578, 2015.
- [10] "twitter のルール.". <https://support.twitter.com/articles/253501>.
- [11] "ツイッターの収益とは何か？ 何で儲けているのか？". <https://devumi.com/twitter-followers/>.
- [12] K.Thomas, D.McCoy, C.Grier, A.Kolcz, and V.Paxson. "trafficking fraudulent account: The role of the underground market in twitter spam and abuse". In *proceedings of the 22nd USENIX conference on Security*, pp. 195–210, 2013.

- [13] "api rate limits per user or per application". <https://dev.twitter.com/rest/public/rate-limiting>.
- [14] "twitter フォロワー総合ランキング". http://meyou.jp/ranking/follower_allcat.
- [15] "twitters". twitters.com.
- [16] "twitter フォロワー". [twitter フォロワー.jp](http://twitter-follower.jp).
- [17] "buy twitter followers reviews". <http://buyfollowersguide.com/where-to-buy/>.
- [18] "coincrack". <https://coincrack.com>.
- [19] "devumi". <https://devumi.com/twitter-followers/>.
- [20] "fast followers". <https://fastfollowerz.com>.
- [21] "twitterboost". <https://twitterboost.com>.
- [22] "twitterfollowertrend". <https://twitterfollowertrend.com>.
- [23] "social media marketing that really works". <http://audiencegain.com/>.
- [24] "buy twitter followers". <http://buy1000followers.co/>.
- [25] "followers social". <http://buyfollowerssocial.com/>.
- [26] "twitter followers trend". <http://twitterfollowerstrend.com>.
- [27] "klout スコアの現状". <https://www.google.co.jp/search?q=Klout>
- [28] "what is klout and what does my klout score mean". <http://rawdigital.training/what-is-klout-how-raise-klout-score/>.
- [29] C.C.Chang and C.J.Lin. "libsvm: A library for support vector machines". *Journal of ACM Transactions on Intelligent Systems and Technology*, Vol. 4, , 1991.

付録 A 本研究に関する倫理的問題に関して

A.1 倫理的な見解

本研究では実際にフォロワーマーケットからフェイクアカウントを購入し、その特徴を抽出することを試みている。また実際に未知のアカウント集団をデータセットとして大量に集めている。この際本研究では倫理的な問題が生じる可能性がある。本研究では J.Song ら [6] の関連研究を参考に以下の点に注意することで倫理的な問題に対応する。

1 つ目は実際にフェイクアカウントを購入することに関する問題である。フェイクアカウントを実際に購入することで実際に見かけ上影響力の高いフェイクアカウントを作ることが可能になる。本研究ではこのフェイクアカウントの管理を厳重にし、フェイクアカウント購入後は一切の活動をしないことでこの問題に対応した。

2 つ目は未知のアカウント集団を手にいれる際生じる問題である。データを収集することである程度個人情報を特定できてしまう恐れがある。この問題に関しては個人を特定する可能性があるものに関しては使用せず、また外部に対してデータを提供することは一切しないことでこの問題に対応した。