

活動時間帯と活動量を考慮した Twitter での つながり構築支援手法とつながり構築支援システム

井上 翔太^{1,a)} 櫛山 淳雄¹

概要: 本論文では Twitter 上でのつながり構築支援を目的とし、その構築支援手法の提案とシステム構築を行った。Twitter 上でのつながり構築は主に人間関係や趣味などのつながりから推薦されてきたが、人間関係や趣味でつながりが強い相手であっても、Twitter の利用頻度や投稿量が違いすぎるとコミュニケーションがうまく行えない場合がある。そこで本研究では、活動時間帯、活動量を用いたつながり構築を行えば、よりコミュニケーションをとりやすくなるのではと考え、活動時間帯、活動量からつながり構築支援手法を提案し、システム構築を行った。

A Method and Its Support System for Supporting Connection Building on Twitter Considering Usage Frequency and Time Zone

SHOTA INOUE^{1,a)} ATSUO HAZEYAMA¹

Abstract: This paper proposes a method for supporting connection building on Twitter and its support system. Traditional recommendation systems for supporting connection building on Twitter were based on human relations and/or hobbies. However, even if a caller and a recipient have a strong connection in hobbies and/or human relations, there is a case in which communications can not be done well in case of large difference in the number of their posts, i.e., usage frequency and in the time zone of their usage. In this paper, we focused on usage frequency and time zone of users for supporting connection building on Twitter.

1. はじめに

スマートフォンの普及によって Social Networking Service(以下 SNS) の使い方に大きな変化が起きている。携帯端末を利用することによって、場所を選ばず SNS を利用できる、その更新頻度は上がっている。それに伴い、ユーザー間の交流も増え、より強固なつながりを構築することが可能になった。一般的な SNS ではまず知人、友人などにつながり、その中でコミュニケーションをとることが多い。さらに SNS を利用する上で必要となるのが、有益な人物とつながりを築くことである。

近年 SNS の一種である Facebook, Twitter[1] のような、SNS に近い形のマイクロブログが急速に普及している。またそれに伴い Facebook, Twitter 上でのつながり構築支援

に関する研究も盛んに行われている [2][3][4]。Twitter は匿名性が高く、また tweet(自分のタイムラインに投稿すること) とフォロー (相手のタイムラインを見られる状態にすること) が主な機能となっている。Twitter は面識の有無ではなく“おもしろそう”などの観点からフォローする傾向が見られる。Twitter でのコミュニティ構築において、その基準となる相手の情報は、tweet とプロフィールであるが、それらは文字制限があり、それだけでユーザー情報を知ることが困難である。そのため、自分にとって有益な人物を探すことが困難になっている。

既存の研究ではユーザーの発言内容からユーザーの趣味や属性を抽出し、それらを用いてマッチングする手法が多く見られる。しかし実際に Twitter を利用したとき、その活動時間帯が異なっているために、自分の見ているタイムラインにフォローしたユーザーが現れずユーザーの tweet を見る機会が少なくなってしまうたり、tweet の投稿数に大

¹ 東京学芸大学
Tokyo Gakugei University
^{a)} m123302g@st.u-gakugei.ac.jp

きな差があるために、ユーザーの tweet が他のユーザーの tweet に埋もれてしまったり、逆にフォローしたユーザーによって自分のタイムラインが埋められてしまったりする場合があります。そのため、ユーザー間で交流が行えなかったり、ストレスを感じてしまうことがある。その問題を解決するには似通った活動のユーザーを探すことが必要となると考えた。

本研究ではまず、実際に Twitter 上でつぶやいているユーザーの tweet を取得し、その中でお互いへの reply が多いユーザー同士において活動時間帯、活動量の類似性があるかの検証を行った。次に検証結果を用いたマッチング手法を提案する。これによりユーザーがシステムを用いて他のユーザーをフォローする際に、よりコミュニケーションをとりやすいユーザーが推薦されることが期待できる。

2. Twitter

Twitter(ツイッター)とは、個々のユーザーが140文字以内で“つぶやき(tweet)”を投稿するサービスである。

ユーザーは tweet と呼ばれる140文字までの短い文章を自身のタイムライン上に時系列順に表示することができる。ユーザーは自分の tweet を閲覧することが可能で、さらに他のユーザーを“フォロー”することによって他のユーザーの tweet も、自分の tweet と同時に閲覧することが可能となる。この2つの tweet が表示される場所を“ホームタイムライン”と呼ぶ。このとき、表示される tweet は投稿された時系列にそって表示され、新しい tweet が投稿されるごとにホームタイムラインは更新されていく。

tweet には“ひとりごと”としての tweet と、特定のユーザーに向けてつぶやかれる“reply”がある。reply は tweet の先頭に@+ユーザーネームをつけ、つぶやかれる。そうすると@をつけられてつぶやかれたユーザーは通知により他ユーザーから reply を受け取ったことがわかるようになっている。このような機能により、Twitter は単なる“ひとりごと”をつぶやく場ではなく、ユーザー同士の交流の場として多く活用されている。

また、ユーザーは他のユーザーをフォローすることにより他のユーザーの tweet を自分のタイムライン上に表示させることができる。

ユーザーは自分の興味関心や目的に応じて、フォローするユーザーをプロフィールや tweet 内容から絞り込む。しかし Twitter 上には2億以上のアカウントが存在しているため、その中から興味のあるユーザーを探し、フォローの可否を見定めることは非常に困難と言える。

そこでユーザーが2億以上のアカウントの中から自分の興味関心に見合ったユーザーを絞り込むために様々な情報を用いる。

たとえば

- Twitter 公式ページ上に表示されるおすすめユーザー

- 公式 RT などによって回ってくる他ユーザー
- 自分をフォローしているユーザー
- 外部ツールなどによる推薦
- 実生活でのつながりや口コミである。

さらにユーザーは他ユーザーをフォローする際に“情報取得目的”と、“コミュニケーション目的”の2種類の目的に沿ってユーザーをフォローする傾向が見られる。

フォローするユーザーを探す際に、ユーザーの興味関心に近いユーザーをプロフィールやユーザーの tweet 取得などにより推定する方法や、フォロー・フォロワー関係などからユーザーの人間関係により近い位置にいるユーザーを推薦する方法がとられている。

3. 関連研究

関連研究として tweet 内容などのユーザー情報を主に使った Twitter での人物推薦に関する研究についていくつか述べる。

さらに関連研究を表1にまとめる。

3.1 Twitter 公式推薦システム

Twitter にはすでに“おすすめユーザー”という形で推薦を行う機能が備えられている。その推薦方法は以下の2通りである。

- “芸能”“音楽”などのジャンル別著名人、企業をカテゴリ分けし、主に登録間もない初心者向けに推薦する方法。

これらの著名人、企業とは Twitter Inc. の認定を受けたアカウントとなる。著名人のプライベートな tweet やお知らせ、企業からのお知らせや広報情報などの最新ニュースをタイムライン上で見ることができるようになる。主に情報収集を中心とした使い方のユーザーに対し有効な推薦と考えられる。またアカウントによっては一般のユーザーからの reply に返信する著名人、企業アカウントもあり、コミュニケーションを目的とするユーザーへも有用と考えられる。

- ユーザーのフォロー情報に基づいて推薦する方法。

これはユーザーのフォロー情報に基づいて推薦されていると考えられる。推薦アルゴリズムは明らかにされていないが、フォローやフォロワーの多くがフォローを行っているユーザーを優先的に推薦している。この場合著名人のアカウントなど多くの人にフォローされやすいアカウントが優先的に出てきてしまうため、自分と関係性の高そうなユーザーを推薦するという本来の目的が果たせていない場面が多々見られる。

表 1 既存手法

	カテゴリ	フォロー関係	reply	利用形態
Twitter 公式	△	△	—	—
白木原	○	—	—	—
北村	—	○	○	—
渡部	○	—	—	—
Golder	—	○	—	—

○:考慮 △:一部考慮 -:考慮せず

3.2 カテゴリ分けでの推薦

白木原ら [5] は Twitter の発言で“流行語”を多く、早くつぶやいているユーザーを抽出し、分野（音楽や映画などのカテゴリ）に敏感な人なのかを判断した。利用者は気になるカテゴリに対応する流行語を多く発言するユーザーを見つけられるようになった。

3.3 メンションの有無

北村ら [6] はフォロー情報だけでなく、メンションの有無に着目し、フォロワー推薦を行った。これはフォロー・フォロワー関係はあるが、実際のコミュニケーション（メンションによる tweet の送受信）の有無によって推薦対象を決定している。

3.4 お気に入りによる属性

渡部ら [7] は、お気に入り登録（＝ふぁぼられ）で多く“ふぁぼられた”ユーザーを取得し、そのユーザーの分野分けを行った。特定の分野において“ふぁぼられ”が多いユーザーはその分野での評価を高く位置づけた。被推薦ユーザーは自分がどのような分野に興味があるかを選択すると、分野と紐づけられた評価の高いユーザーが推薦されるシステムを作った。

3.5 Twitter のフォロー関係を用いた研究

Golder[8]らは、Twitterにおいて、ユーザーから見て様々なフォロー関係のつながり方をしている他のユーザーを推薦しその結果を比較している。その中で有効性のあるつながり、有効性のないつながりがあることをあきらかにした。この研究ではユーザー同士のフォロー・フォロワー関係のみで推薦し、tweet の中身は用いていない。

4. 研究アプローチ

4.1 活動時間帯・活動量

Twitter は 140 文字という制限のなかで“つぶやく”。スマートフォンの普及によってつでも、どこからでも利用できるようになった。それに従い、投稿数や投稿時間帯はユーザーによってばらつきが生まれる。

Twitter でつぶやかれた tweet は時系列で並べられタイムラインに表示される。tweet には投稿された日時がつけられる。気になる tweet があつたとしても気づくのが遅れ

てしまうと、reply を送りにくくなってしまう。そのようなリアルタイム性が Twitter にはある。他人の tweet に対して reply を送る場合、できるだけ tweet と近い時間に送る必要がある。そのため自分が Twitter をよく見ている時間帯に同じように Twitter を利用しているユーザーをフォローすることにより、より早く興味関心のある tweet を見つけることができると考える。

また Twitter では複数のユーザーをフォローすることが考えられる。複数のユーザーの tweet が 1 つのタイムライン上に表示されるとき、例えば、1 日に 100 件つぶやくユーザー A、1 日に 5 件つぶやくユーザー B を同時にフォローすると、ホームタイムラインはユーザー A が埋めてしまい、B の tweet は埋もれてしまう。これでは B の tweet を見つけることは難しく、B とつながりを築きにくい。つまり tweet 数が多すぎるとそのユーザーの tweet 自体は目につきやすくなるが、他のユーザーの tweet を埋もれさせてしまう。逆に tweet 数が少ないと他のユーザーの tweet に埋もれて目につきにくくなってしまう。

したがって、コミュニケーション目的でユーザーを探したい場合、他のユーザーと自分の tweet する時間帯と投稿量は近い必要があると考えた。

そこで活動時間帯、活動量の 2 点においてユーザーの類似性を考える。本研究において活動時間帯は各曜日ごと 1 時間ごとに 24 分割した、1 時間あたりの tweet 数の平均とし、活動量は、1 日あたりの tweet 数の平均とする。活動時間帯、活動量を合わせて利用形態とする。

また本研究における“活動”とはユーザーが Twitter で tweet を投稿している状態のことを示す。タイムラインを閲覧しているだけなどの状態は含まない。

さらに本研究ではコミュニケーションを目的としたユーザー推薦を扱う。フォローの目的に挙げた“情報取得を目的”としたユーザーの推薦は考えないこととする。

5. 予備実験

予備実験として、ユーザーの tweet を収集し、reply を多く送っているユーザー同士は活動時間帯、活動量に関係性があるかを検証する。

5.1 データ収集

tweet の取得は TwitterAPI を利用し 1 人のユーザーから最大 3000 件取得する。

ユーザー間の reply の多さと活動時間帯の関係性を調べるため予備実験を行う。元となるユーザー（親）の tweet を取得後、その中で reply を送っているユーザー（子）をリストアップし、そのユーザーの tweet も同様に 3000 件取得した（※ tweet 数が 3000 件に満たない場合は取得できる最大数を取得）。

今回取得したデータを表 2 に表す。

表 2 予備実験取得データ数

	アカウント数
ユーザー (親)	11
ユーザー (子)	1433

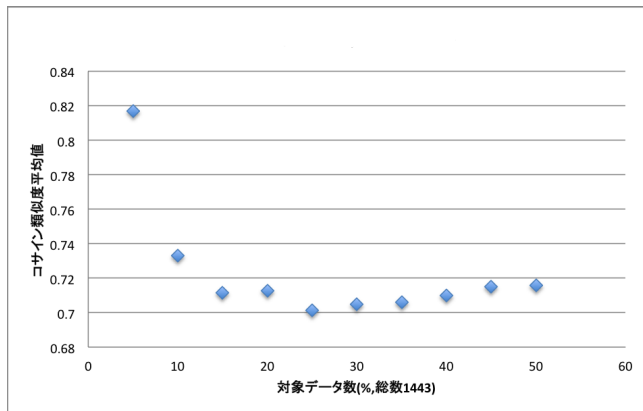


図 1 活動時間帯による類似度

次に取得した tweet から各ユーザーの活動時間帯を計算する。

その後、親ユーザーの各時間における tweet 数の平均と子ユーザーの tweet 数の平均の類似度を算出する。

5.2 活動時間帯による類似度

活動時間帯の類似度 R_c はコサイン類似度を用いて算出する。コサイン類似度は以下の式より求める。

$(x_i, y_i) (i = 1, 2, 3, \dots, n)$ が与えられたとき

$$R_c = \frac{\sum_{i=1}^n x_i y_i}{\sqrt{\sum_{i=1}^n x_i^2} \sqrt{\sum_{i=1}^n y_i^2}}$$

である。

予備実験で x_i, y_i はそれぞれ、親ユーザーの時間帯 i における tweet 投稿数、子ユーザーの時間帯 i における tweet 投稿数とする。

図 1 にコサイン類似度による類似度の平均値を表す。x 軸は親から子への reply の多い順 (よりコミュニケーションをとっている順) に並べ、上位から一定割合で区切ったときのデータ数を示している。y 軸はコサイン類似度の平均値となっている。

5.3 活動量による類似度

活動量による類似度を算出するため活動量による類似度 R_t を

$$R_t = \frac{U_B}{U_A}$$

とする。 U_A, U_B はそれぞれ U_A : 親ユーザー 1 週間分の tweet 平均総数、 U_B : 子ユーザーの 1 週間分の tweet 平均総

数である。図 2 に親子ユーザーの活動量の類似度を示す。

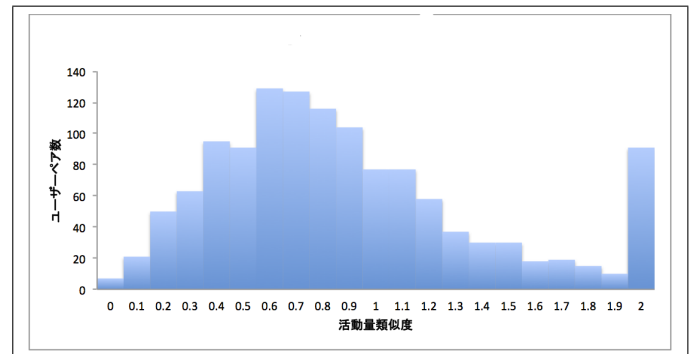


図 2 活動量による類似度

5.4 予備実験考察

予備実験の結果から活動時間帯においては今回取得したデータの中で、親から子への reply の割合が多いほどコサイン類似度で高い値が得られた。また活動量においても、reply を割合の多い親と子の活動量に関係性があることがわかった。

6. 提案手法

予備実験からユーザー間の reply の量と、活動時間帯、活動量に関係性があることがわかった。本研究ではユーザーの tweet の投稿時間による“活動時間帯”と投稿件数による“活動量”によるマッチングを行う。

6.1 既存手法と提案手法

既存手法においては Twitter の利用形態を考慮したユーザー推薦が行われていないため、本手法ではユーザーの利用形態を考慮してマッチング手法を考える。

6.2 類似度

本研究では

- tweet の活動時間帯での類似度
- tweet の活動量による類似度

の 2 つの点において一定以上の値をとったユーザー同士を推薦する。

最終的な推薦は (活動時間帯の類似度) と (活動量による類似度) において設定したしきい値の間を満たしたユーザー同士に行う。式は以下に設定する。

$$R_c \geq T_C \text{ かつ } T_{R_1} \geq R_t \geq T_{R_2}$$

T_C は活動時間帯の類似度から設定するしきい値、 T_{R_1}, T_{R_2} は活動量から設定するしきい値である。

7. つながり構築支援システム

提案手法を用いたマッチングシステムは、ユーザーが自分と同じような活動量、活動時間帯で Twitter を利用しているユーザーを探すシステムである。

7.1 TwitterAPI

本システムでは TwitterAPI 1.1 を用いてユーザー情報の取得を行っている。TwitterAPI は 2013 年 6 月 11 日より API 1 を廃止し API 1.1 に移行した。これにより Twitter 上の情報取得方法についての変更が行われた。システム作成においてもこの変更点が関係する点があるためここで解説する。

TwitterAPI を用いて取得できるデータには様々なものがある。例えば tweet 情報であれば、1 つの tweet 内に、tweet の作成日時、tweet 内容、tweet 固有 ID、Twitter クライアント情報、reply 先ユーザー ID、reply 先テキスト ID、つぶやいた本人のユーザー情報、場所や Retweet 回数、favorite 回数、メンションが含まれていた場合そのメンション先のユーザー情報、などが含まれる。ユーザー情報の中にはユーザーネーム、ID、フォローフォロワー数、Twitter を始めた日付、全 tweet 数などが記録されている。そのように 1 つの tweet から多くの情報を得ることができるようになっている。

これら取得できるデータは TwitterDeveloper[9] や、twitter-sandbox[10] において確認することができる。

本研究では tweet に付加されている情報のうち、“created at”、“id”、“in reply status id” の情報を用いて、その tweet が“いつ”、“誰によって”、“誰に向けて”、つぶやかれたかを判別した。

また既存の多くのシステムではユーザー情報を多く取得する場合、ユーザーごと個別にアカウントの OAuth 認証を行い、ユーザー本人のアカウントから情報を取得していた。これはアカウントごと (IP アドレスごと) に API のレートリミットが定められているため、1 つのアカウント (例えばシステムのアカウント) から、他のユーザー情報を取得しようとするときにリミットに達してしまうためである。しかし 1.1 から、アプリケーション単位でのレートリミットになったため、複数のアプリ登録によってリミットによる制限を回避できるようになった。

これを用い利用ユーザーがアプリの認証を行わずにシステムを使えるように設計した。

7.2 システム概要

予備実験で得られた類似度の式を用いて活動時間帯と活動量を考慮したつながり構築支援システムを作成した。システム構成を図 3 に示す。

本システムでのユーザー、システムの動作は以下の通りである。

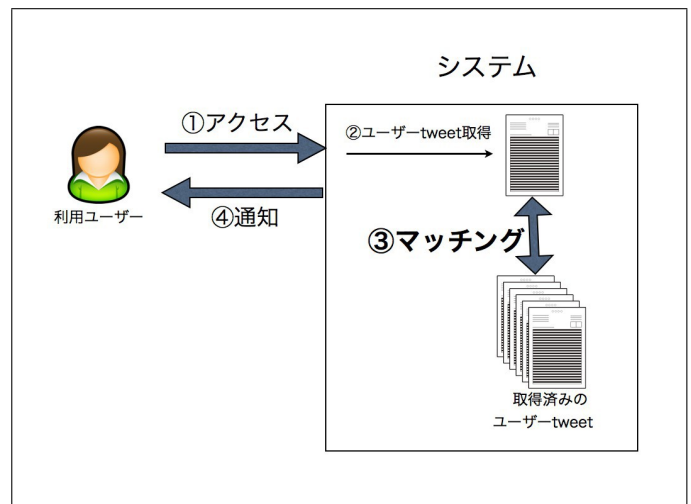


図 3 システム構成

【STEP 1】ユーザーはシステムにアクセスし、自分の TwitterID を入力する。

【STEP 2】システムはユーザーが入力した ID の tweet を Twitter から取得できる最大件数取得する。

【STEP 3】システムは取得したユーザー tweet とデータベース上にあるマッチング用ユーザー tweet との類似度計算を定期的に行う。

【STEP 4】しきい値を満たすユーザーペアが見つかった場合 Twitter 上のシステムアカウントから reply 機能を用いて通知する。

また、ユーザー利用画面はユーザーが ID を入力後【登録】ボタンを押すと、ユーザー情報の詳細が表示される。

システムからユーザーへの通知は reply を用いる。

ユーザー利用画面を図 4 に示す。

システムからユーザーへの通知画面を図 5 に示す。

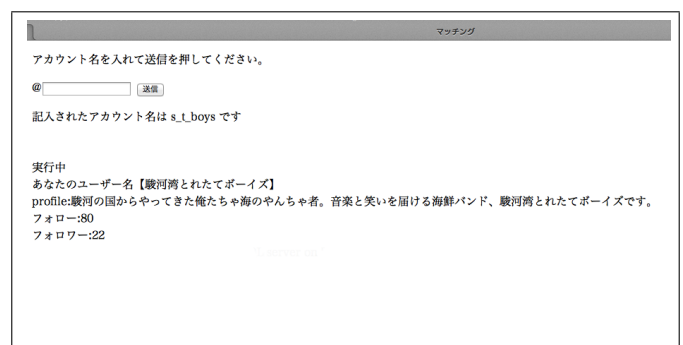


図 4 ユーザー利用画面

本システムではユーザー tweet 群の量が多いほど多くのユーザーとのマッチングを行い、より類似度の高いユーザー



図 5 システムからユーザーへの通知画面

ザーを推薦することができるため、常時 Twitter 上でユーザーの tweet を収集する。

8. まとめ

本研究では Twitter での活動量と活動時間帯を考慮したコミュニケーションを目的とした、つながり構築支援手法を検討した。

予備実験から、すでにつながりがあり、その中でもよくコミュニケーションをとっているユーザー同士には活動時間帯、活動量において関係性があることがわかった。

今後は活動時間帯、活動量を用いたシステムでのマッチングを実際に行い、ユーザーによる効果検証と、既存の手法との比較を行っていく。

参考文献

- [1] Twitter : 公式: <https://twitter.com>
- [2] 桑原雄, 稲垣陽一, 草野奉章, 中島伸介, 張建偉: マイクロブログを対象としたユーザ特性分析に基づく類似ユーザの発見および推薦方式, 情報処理学会研究報告データベース・システム研究会報告, pp. 1-3 (2009)
- [3] 岡本大輝, 豊田正史, 喜連川優: マイクロブログにおける対話ネットワークと投稿内容を併用したユーザ推薦に関する一考察, 情報処理学会研究報告データベース・システム研究会報告, pp. 1-5 (2013)
- [4] 田村政人, 小林亜樹: Twitter における会話しやすいユーザの推薦手法, 第 75 回全国大会講演論文集, pp. 605-607 (2013)
- [5] 白木原渉, 大石哲也, 越村三幸, 藤田博, 長谷川隆三: Twitter の流行語発言者の抽出に基づくフォロワー推薦システムの開発, 情報処理学会創立 50 周年記念 (第 72 回) 全国大会 5S-4 (2010)
- [6] 北村太一, 小川裕樹, 諏訪博彦, 太田敏澄: コミュニケーションに着目した Twitter フォロワーユーザ推薦, The 26th annual conference of the Japanese society for artificial intelligence (2012)
- [7] 渡部将太, 宮森恒: Twitter User Recomender ~Twitter のお気に入り機能を用いたおすすめユーザー推薦システム~, DEIM Forum 2012 B3-4 (2012)
- [8] Scott A. Golder, and S. Yardi: Structural Predictors of Tie Formation in Twitter Transitivity and Mutuality,

Proc. of 2010 IEEE Second International Conference on Social Computing, pp. 88-95 (2010)

- [9] Developers : <https://dev.twitter.com>
- [10] twitSandbox : <http://twitsandbox.com>