

目に留まるツイート投稿タイミング推薦手法

村上 雄亮[†] 小林 亜樹[†]

[†] 工学院大学工学部 〒163-8677 東京新宿区西新宿 1-24-2
E-mail: †c509107@ns.kogakuin.ac.jp, ††aki@cc.kogakuin.ac.jp

あらまし Twitter を広告配信的に利用される際の問題は見逃しである。そこで本研究では、ユーザのツイート閲覧確率を時間帯ごとに推定することで、閲覧可能性の高い時間帯を指し示すことで適切なツイート投稿タイミングを示すことができる手法を提案する。具体的には、ユーザは閲覧、ツイートの二つのアクションをし、各アクションでは一定件数のツイートを閲覧するものとモデル化し、時間帯ごとの期待閲覧密度を求める。この際、直接観測できないユーザタイムライン速度ならびに閲覧密度は、観測可能事象から推定する手法も提案する。推定値より投稿時間帯を推薦することで、目に留まるツイートを実現する。実験的に効果を測定した。

キーワード Twitter, ツイートタイミング, 閲覧密度

Tweet Timing Estimation to Catch a User's Eye

Yusuke MURAKAMI[†] and Aki KOBAYASHI[†]

[†] Faculty of Engineering, Kogakuin University
Nishishinjuku 1-24-2,shinjuku,Tokyo,163-8677 Japan
E-mail: †c509107@ns.kogakuin.ac.jp, ††aki@cc.kogakuin.ac.jp

1. 背景と目的

Twitter が普及していくにつれ、企業の製品広告や、イベントの告知などの広告配信的に利用されることが増え、プロモツイートという公式の広告ツイートサービスも開始された。プロモツイートは有料である代わりに、ユーザの目に留まるような工夫がされている。具体的には、プロモツイートの投稿時刻に関わらず、タイムラインと呼ばれる投稿されたツイートのリストの最新の場所に表示されるようになっていく。

しかし、プロモツイートをしないユーザも広告配信的にツイートをを行うことがある。その際問題になるのがツイートの見逃しである。プロモツイートでない通常のツイートは時系列順に表示されるため、古いツイートはある程度遡って閲覧するしかなく、他のツイートに埋もれ、見逃されてしまう可能性がある。

そこで本研究では、ツイートを見せたい対象ユーザの時間帯別の閲覧確率を算出し、閲覧確率が高い時間帯を見逃されにくい投稿タイミングとし、その時間帯を推薦する手法を提案する。

竹村ら [1] は、ツイートの内容から「今見るべき」、「後でよい」、「もう見なくてよい」という分類を行い、ユーザに対してどのツイートを優先してみるべきかの判定をしている。古賀ら [2] はツイートに含まれる日付表現から、未来のイベントの

情報などを抽出し、ユーザに対して再通知を行うシステムを作成している。これらの手法はツイートに含まれる情報から、分類をしユーザにツイートを見やすくする、認知させる工夫をし、ツイートを閲覧するユーザ側からツイートの見逃しを予防する手法といえる。

これらは過去の情報を整理してユーザに提示する手法であり、閲覧するユーザがツイートを見逃したくない時に使用するシステムである。本研究は投稿ユーザがツイートをみてほしい時に使用するものであり、これらの手法とは利用するユーザと目的が異なる。

また田中ら [3] の、ツイッター特有のユーザ関係を考慮し、不快なユーザの発言を一時的に見づらくすることで、タイムライン閲覧時の不快感を軽減する手法のような研究もあるが、ツイートを見づらくする点は、ツイートをみてもらいたい本研究とは正反対の位置づけであるといえる。

松尾ら [4] は、ツイッター利用者の興味は、ツイッター上の話題と関係していると仮定し、ツイート内容から興味を抽出し、それに合った広告ツイートを配信し閲覧数の最大化を試みるシステムを提案した。

広告ツイートを投稿するユーザが利用する点、ツイートをみてほしいという点は本研究と目的が同じであるが、松尾らの手法が広告ツイートの種類を推薦するのに対し、本手法はツイ

トの投稿タイミングを推薦するという点で異なる。

また、タイミングというキーワードに関しては、林ら [5] が、ユーザが携帯電話でメール機能を利用する際の、状況別使用率のアンケート結果から、情報通知タイミングの考察を行っている。林らは、「待ち合わせしているとき」のような実際のユーザの行動をもとにタイミングを考察しているのに対し、本研究では Twitter 上のユーザの行動からタイミングを考察している。

2. 提案手法

広告配信的にツイートを投稿しているユーザは、閲覧側のユーザにそのツイートを見逃されてしまうことが問題であり、見逃したユーザも欲しい情報を見逃してしまった場合不利益になる。しかし、閲覧側ユーザの行動をあらかじめ知っておくことは不可能であるが、過去の情報から閲覧側ユーザの行動を数値として表し予測を行うことは可能である。

したがって、本手法では閲覧側ユーザの過去の行動履歴から、どのタイミングで、どの程度ツイッターを閲覧しているかを数値として表し、それをもとに投稿側ユーザに対してツイートを投稿するタイミングを推薦する。

ユーザ固有のツイート閲覧確率を時間帯ごとに推定し、閲覧可能性の高い時間帯を指し示すことで適切なツイート投稿タイミングを示す法を提案する。本手法ではユーザのモデル化を行い、ユーザ u の時間帯ごとのアクション密度 f とタイムライン速度 s から、期待ツイート閲覧確率 r を算出し、 r の結果から目に留まるツイートを投稿するタイミングを推薦する。

以下に提案手法の大まかな処理の流れを示す。

- (1) Twitter ユーザ u のモデル化を行う。
- (2) u のアクション密度 f を推定する。
- (3) u タイムライン速度 s を推定する。
- (4) 期待閲覧ツイート確率 r を算出する。
- (5) r から適切なツイート投稿タイミングを推薦する。

2.1 Twitter

Twitter は^(注1)140文字以内の短文を投稿できる情報サービスである。投稿された短文をツイートといい、大きく分けて以下の4つの要素で構成されている。

2.1.1 フォロワー

Twitter は他のアカウントを選択することで、そのアカウントのツイートを閲覧することができる。この選択をフォローといい、自分の好きなアカウントをフォローすることで、自分のみたいツイートだけを閲覧することができるようになる。逆に自身のアカウントをフォローしているアカウントをフォロワーと呼ぶ。当然フォロワーには自身のツイートを閲覧することができる。

2.1.2 アカウント

大きく分けて、通常アカウントと鍵付きアカウントの2つが存在している。通常アカウントは誰からでもフォローすることができ、ツイートを閲覧することができる。一方鍵付きアカウントをフォローする場合は鍵付きアカウントの認証を必要とする。

する。

2.1.3 ツイート

ツイートの種類は以下の3つに分けることができる。

- ツイート
- 返信ツイート
- リツイート

ツイートは投稿欄に書いた文章をそのまま投稿する通常のツイートである。返信ツイートはあるツイートに対して返信を行うツイートであり、返信先はそのツイートの投稿者になる。リツイートは他人のツイートをもう一度自分が投稿し直すツイートである。そのため自身のフォロワーのタイムラインには元のツイートを投稿したアカウントをフォローしていなくても、そのツイートが表示される。

2.1.4 タイムライン

ツイートは時系列順に並べられ、そのリストはタイムラインと呼ばれる。タイムラインは以下の3種類がある。

- パブリックタイムライン
- ユーザタイムライン
- ホームタイムライン

ここでツイートの項で記した3つの種類をそれぞれ、ツイート T 、返信ツイート M 、リツイート R とする。

パブリックタイムラインは非公開設定を除く全ユーザの最新20件の T 、 R で構成される。ユーザタイムラインは、自身の T 、返信ツイート M 、リツイート R で構成される。ホームタイムラインは、自身の T 、 M 、 R の他に、自身がフォローしているユーザの T 、 R で構成される。基本的に M は表示されないが、自身に対する M 、フォローしているユーザ間の M のみ表示される。

T 、 M 、 R はユーザ固有であるためユーザ特性 u を持つ。

例えばユーザ X とユーザ群 Y が存在し、 X が Y をフォローしている時、 X のユーザタイムライン $U.t$ は

$$U.t(X) = T(X) + M(X) + R(X) \quad (1)$$

と表せられ、 X のホームタイムライン $H.t$ は

$$H.t(X) = \begin{cases} U.t(X) + U.t(Y) & \text{特定のMを含むとき} \\ U.(X) + T(Y) + R(Y) & \text{otherwise} \end{cases} \quad (2)$$

と表すことができる。

2.2 TwitterAPI

Twitter は API を公開しており、タイムラインの取得や更新等が可能である。API の種類によっては OAuth 認証が必要であり、利用するためには開発者用のサイトでアプリケーション登録をし、そのアプリケーションを利用するユーザからアプリケーション連携を許可してもらう必要がある。

2.3 ユーザのモデル化

あるユーザ u が twitter にログインした際に、画面にはホームタイムラインが表示されると仮定する。

ユーザ u が行う操作はタイムラインの閲覧、ツイートを投稿、リツイート、返信ツイート、他のアカウントをフォローする、ツイートをお気に入り登録する等がある。

(注1) : <https://twitter.com/>

リツイート、返信ツイートは自身が見たツイートにしかできない、フォローやお気に入りも、ツイートの内容から判断すると仮定すると、これらの操作はタイムラインの閲覧後に行われること考えられる。

したがって、ユーザの操作はタイムラインの閲覧とツイート投稿の二つに分類することができ、これらをユーザのアクションと呼ぶ。

またユーザはユーザ固有の頻度で Twitter を利用すると考えられる。Twitter を利用する際、行う操作は閲覧かツイート投稿の二つのアクションどちらかである。1 時間毎のアクションの回数をアクション密度 f [アクション回数/時間] とし、ある時間 h のアクション密度は $f(h)$ と表される。また時間帯内においては一定間隔でアクションをすると仮定する。さらに $f(h)$ は、閲覧のアクション $F_v(h)$ 、ツイートのアクション $F_t(h)$ の二つに分けることができる。

ユーザは、閲覧時にその時点のホームタイムラインの最新ツイートを含む、一定の遡り閲覧件数 w [件/回] を遡って読むものとする。閲覧時にデバイス、サービス依存の画面上に表示される件数は読むことが期待される。また、そこからある程度遡る操作は可能であると考えられ、それはその操作を行うユーザが相当数存在することを示唆する。遡り閲覧件数は、閲覧時のユーザの状況によって大きく異なることが予想される。しかし、遡り閲覧件数やその変化状況についての定量的な報告はないため、本論文では一定件数であるものとしてモデル化する。

したがって、本手法で扱うユーザ u はパラメータとして $F_v(h)$ 、 $F_t(h)$ の二つのアクションをパラメータとして持つとモデル化する。また閲覧ツイート数 w もパラメータとして持つが、これは一定数として扱っている。

2.4 期待閲覧ツイート確率 r

期待閲覧ツイート確率 r はホームタイムラインに流れるツイートを、ユーザ u がどの程度閲覧しているかを単位時間ごとに表している。

ユーザが一回のアクションごとに閲覧するツイート数を c とし、ユーザはアクション密度 f でアクションすると仮定すると、ある時間 h における c はアクション密度 f と一定閲覧数 w を用いて

$$c(h) = wf(h) \quad (3)$$

と表せる。

さらに、 h 時におけるホームタイムライン上に流れるツイート数を 1 時間当たりで計算し、タイムライン速度 $s(h)$ ツイート件数/時間 と表す。

ここで $c(h)$ 、 $s(h)$ の例を挙げると、ユーザ u が過去一週間で 12 時台に合計 10 件のツイートをし、 $w = 20$ であるとき、 $c(h)$ は

$$c(12) = 20 * 10 \quad (4)$$

となり、ユーザ u の 12 時台のホームタイムラインに 100 件のツイートが流れていたとすると、 $s(h)$ は

$$s(12) = 100[\text{ツイート件数/時間}] \quad (5)$$

と表すことができる。

$c(h)$ と $s(h)$ から r は式 (6) で表され、 $c(h)$ が $s(h)$ より大きい時、ユーザ u はタイムラインに流れる全てのツイートを閲覧していることになるため、 $r(h)$ は上限値として 1 をとり、式 (6) と表せる。

$$r(h) = \begin{cases} \frac{c(h)}{s(h)} & s(h) \geq c(h) \text{ のとき} \\ 1 & \text{otherwise} \end{cases} \quad (6)$$

この期待閲覧ツイート確率 $r(h)$ の数値の高い時間帯を適切なツイートタイミングとしてユーザに推薦する。また、アクション密度 f 、タイムライン速度 s はモデル化や API の制限より、直接観測することができず、推定を行う必要がある。3.2、3.3、3.4 で推定の手法について述べる。

3. 各パラメータの推定

3.1 API の制限

図 1 のように第三者が取得できる情報には制限がある。モデル化を行ったアクションに含まれる「閲覧回数」は API を用いて取得することができない。また、タイムライン速度を求める際に必要なホームタイムラインも認証が必要となり、ツイートを見てもらいたいユーザに対して手間をかけさせることは本研究の目的と異なる。そのためこれらの値については推定を行う。

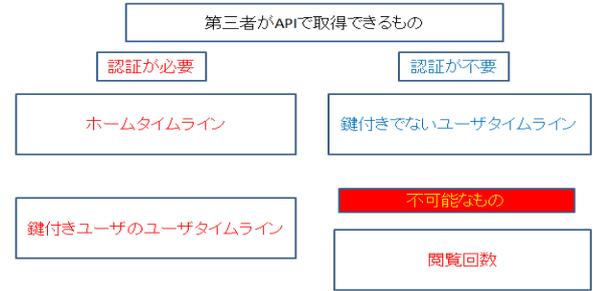


図 1 API で取得できるもの

3.2 アクション密度 f の推定

アクション密度 f はユーザ u の 1 時間当たりのアクションの回数を表したもので、ある時間帯 h におけるユーザ u のアクション密度は $f(h)$ [回/時間] と表すことができる。

モデル化でのアクションは二種類あるが、実際には API で提供されているのはツイートを取得することのみであり、Twitter を閲覧したという情報は取得することができない。そのためアクション密度は推定する必要がある。本手法で行ったモデル化では、図の通り、ツイートを投稿するアクションの際にもタイムラインを閲覧していることになるため、ツイートを投稿するアクションからアクション密度を推定する。

したがって本手法ではユーザのアクションの内、閲覧についてもツイートの情報から推定するものとし推定アクション密度を $\hat{f}(h)$ [回/時間] とする。

$\hat{f}(u, h)$ は API を用いて取得した u のユーザタイムラインの

各ツイートの投稿時間を参照し1時間ごとにツイートを振り分けることで得られる。

本手法で行ったモデル化では、ツイートアクション集合 F_t は、閲覧アクション集合 F_v に含まれることになる。したがって F_t と F_v の関係から \hat{f} の推定精度 p は式 (7) に示される。

$$\begin{cases} F_t \subseteq F_v \\ 0 \leq p = \frac{|F_t|}{|F_v|} \leq 1 \end{cases} \quad (7)$$

この結果からアクション密度 $f(h)$ と推定アクション密度 $\hat{f}(h)$ の関係は式 (5) のようになる。

$$\hat{f}(h) = pf(h) \quad (8)$$

3.3 タイムライン速度 s の推定

1時間当たりのタイムラインに流れるツイート数をタイムライン速度 s [ツイート/時間] と表すことができる。あるユーザ u の時間帯 h のタイムライン速度を得るためには、 u が普段閲覧しているホームタイムラインを取得すればよい。しかし、図1に示す通り、第三者がホームタイムラインを取得するには u の認証が必要であり、ユーザ u に認証して貰うのは本研究の目的と異なる。そのため u のホームタイムラインを推定する必要がある。

2.1.4 で述べたとおり、ホームタイムラインはユーザタイムラインの集合である。

したがって、 u のホームタイムラインを得るには u のフォローしているアカウントのユーザタイムラインを取得すればよい。

実装上では、`friends()` 関数でユーザ u のフォローしているアカウントのリストを取得し、次にリスト内のアカウントのユーザタイムラインを `user_timeline()` 関数で取得する。その際 u のホームタイムラインに表示されない返信ツイートを取り除く。これにより本来 u のホームタイムラインに表示されるツイートのみを抽出することが可能である。

しかし、今回の実装では返信ツイートをすべて取り除いているため、その分の誤差が生じてしまい、また、鍵付きアカウントのユーザタイムラインは外部から取得することが出来ないため完全なホームタイムラインを取得する事はできない。これらの誤差については本研究では扱わない。

この推定から推定タイムライン速度 $\hat{s}(h)$ を得る。

3.4 推定期待閲覧ツイート確率 $\hat{r}(h)$

3.2, 3.3 から、時間帯 h の期待閲覧ツイート確率の推定値 $\hat{r}(h)$ は式9のように表せる。

$$\hat{r}(h) = \begin{cases} \frac{\hat{c}(h)}{\hat{s}(h)} & \hat{s}(h) \geq \hat{c}(h) \text{ のとき} \\ 1 & \text{otherwise} \end{cases} \quad (9)$$

各パラメータを推定するためには過去数週間分の情報から計算することが望ましい。この期間に関しては、生活リズムに大きな変化を及ぼしそうな数カ月単位ではなく、あまり変化のない一週間から数週間程度が妥当だと考えられるが、期間に関する最適な数値の決定は本稿では扱わない。

4. 実験

3.4 で述べた推定期待閲覧ツイート率 \hat{r} は、ユーザの目に留まる時間帯を示しているのかを実験を通じて評価する。

実験ツイートを複数の時間帯に複数回投稿し、その実験ツイートに協力ユーザに返信をしてもらい、その回数を各時間ごと、ユーザ毎に分類する。この結果は、本来得ることの出来ないユーザの閲覧アクションである。この閲覧アクションと、本手法によって得られたユーザの推定期待閲覧ツイート確率 \hat{r} を比べ、 \hat{r} は実際のユーザのアクションとどのような関係があるかを評価する。

4.1 実験方法

実験用アカウントを協力者である大学生10人にフォローしてもらった。

ある時間に実験用アカウントから実験ツイートを投稿し、協力者にはそのツイートを閲覧した際には返信をもらった。返信してもらう際の条件は協力ユーザが実際に目に留まったと感じた時ということと、ユーザが実験ツイートを古いツイートと感じない時とした。

実際の時間別の返信件数と、それに対応する \hat{r} の値を比べ、評価する。実験は21日間行い表1に時間帯別実験ツイートの件数を示した。また一回のアクションでユーザが閲覧するツイート数 w は20件として、一日毎に過去一週間分のデータから \hat{r} を算出し、さらに実験期間中の全て \hat{r} の平均値を評価した。

表2は協力してもらったユーザの基本情報である。

表1 ツイート時刻と回数

| 投稿時間 [時](UTC) | ツイート回数 [件] |
|---------------|------------|
| 3 | 6 |
| 6 | 10 |
| 10 | 7 |
| 13 | 10 |
| 15 | 10 |
| 18 | 9 |
| 20 | 10 |
| 23 | 12 |

表2 協力ユーザの基本データ

| ユーザ | フォロー数 | フォロワー数 | Twitter 利用歴 (日) | 総ツイート数 |
|-----|-------|--------|-----------------|--------|
| a | 59 | 55 | 817 | 1973 |
| b | 43 | 53 | 1281 | 3336 |
| c | 67 | 57 | 684 | 536 |
| d | 77 | 43 | 1110 | 5739 |
| e | 96 | 55 | 731 | 2266 |
| f | 35 | 26 | 461 | 940 |
| g | 165 | 153 | 943 | 5123 |
| h | 8 | 5 | 230 | 39 |
| i | 331 | 349 | 815 | 3193 |
| j | 180 | 90 | 674 | 2069 |

4.2 閲覧状況評価

実験のデータを集計し、表にしたものが表3, 4, 5, 6である。表5, 6から図3のような推定値と実際のユーザの行動を比べたグラフが得られた。

表3 時間, ユーザ別推定アクション密度 \hat{f} [回/時]

| 時間 | \hat{f} | a | b | c | d | e | f | g | h | i | j |
|----|-----------|-----|-----|-----|------|-----|-----|-----|-----|-----|------|
| 3 | | 2.4 | 0.6 | 0.1 | 4.4 | 0.4 | 2.2 | 2.4 | 0.5 | 1.3 | 8.4 |
| 6 | | 1.8 | 1.5 | 0.0 | 2.7 | 1.1 | 1.0 | 0.7 | 0.6 | 3.6 | 3.1 |
| 10 | | 1.3 | 0.0 | 0.0 | 13.1 | 0.3 | 0.4 | 0.7 | 0.0 | 1.7 | 3.3 |
| 13 | | 2.2 | 1.1 | 0.4 | 30.6 | 1.1 | 1.6 | 0.2 | 0.0 | 4.3 | 5.0 |
| 15 | | 4.3 | 0.8 | 0.3 | 11.1 | 0.3 | 0.2 | 3.6 | 0.8 | 6.5 | 14.9 |
| 18 | | 0.0 | 0.3 | 0.0 | 6.5 | 1.1 | 0.0 | 0.0 | 0.0 | 0.3 | 2.8 |
| 20 | | 0.0 | 0.0 | 0.0 | 0.3 | 0.3 | 0.3 | 0.0 | 0.0 | 0.3 | 0.0 |
| 23 | | 2.7 | 1.8 | 0.0 | 0.6 | 0.6 | 0.0 | 0.4 | 0.4 | 0.0 | 0.0 |

表4 時間, ユーザ別タイムライン速度 \hat{s} [件/時]

| 時間 | \hat{s} | a | b | c | d | e | f | g | h | i | j |
|----|-----------|-----|-------|----|-----|-----|----|-----|----|-----|-----|
| 3 | | 180 | 27 | 43 | 120 | 247 | 62 | 131 | 41 | 121 | 287 |
| 6 | | 198 | 18 | 43 | 126 | 198 | 51 | 130 | 19 | 139 | 271 |
| 10 | | 208 | 21 | 46 | 173 | 203 | 60 | 134 | 24 | 108 | 287 |
| 13 | | 228 | 24 | 72 | 233 | 277 | 54 | 156 | 22 | 140 | 376 |
| 15 | | 217 | 18 | 53 | 224 | 249 | 41 | 123 | 19 | 203 | 346 |
| 18 | | 129 | 12 | 10 | 45 | 108 | 6 | 58 | 6 | 86 | 256 |
| 20 | | 99 | 8 | 9 | 37 | 49 | 26 | 32 | 3 | 49 | 264 |
| 23 | | 150 | 14.65 | 38 | 69 | 158 | 41 | 71 | 13 | 59 | 329 |

表5 時間, ユーザ別推定期待閲覧ツイート確率 \hat{r}

| 時間 | \hat{r} | a | b | c | d | e | f | g | h | i | j |
|----|-----------|------|------|------|------|------|------|------|------|------|------|
| 3 | | 0.27 | 0.46 | 0.04 | 0.79 | 0.03 | 0.72 | 0.31 | 0.23 | 0.21 | 0.58 |
| 6 | | 0.22 | 1.32 | 0.00 | 0.38 | 0.10 | 0.35 | 0.11 | 0.56 | 0.51 | 0.24 |
| 10 | | 0.12 | 0.00 | 0.00 | 1.63 | 0.03 | 0.11 | 0.09 | 0.00 | 0.29 | 0.23 |
| 13 | | 0.24 | 0.79 | 0.12 | 2.69 | 0.08 | 0.59 | 0.03 | 0.00 | 0.62 | 0.26 |
| 15 | | 0.38 | 0.52 | 0.15 | 1.03 | 0.02 | 0.08 | 0.58 | 1.06 | 0.66 | 0.85 |
| 18 | | 0.00 | 0.82 | 0.00 | 2.79 | 0.20 | 0.00 | 0.00 | 0.00 | 0.07 | 0.22 |
| 20 | | 0.00 | 0.00 | 0.00 | 0.18 | 0.08 | 0.26 | 0.00 | 0.00 | 0.10 | 0.00 |
| 23 | | 0.32 | 3.23 | 0.00 | 0.15 | 0.07 | 0.00 | 0.11 | 0.55 | 0.00 | 0.00 |

表6 時間, ユーザ別返信率 m

| 時間 | m | a | b | c | d | e | f | g | h | i | j |
|----|-----|------|------|------|------|------|------|------|------|------|------|
| 3 | | 0.38 | 0.52 | 0.20 | 0.61 | 0.06 | 0.14 | 0.51 | 0.31 | 0.11 | 0.61 |
| 6 | | 0.45 | 1.00 | 0.00 | 0.56 | 0.09 | 0.00 | 0.16 | 0.99 | 0.49 | 0.58 |
| 10 | | 0.24 | 0.00 | 0.00 | 1.00 | 0.07 | 0.00 | 0.09 | 0.00 | 0.63 | 0.23 |
| 13 | | 0.58 | 1.00 | 0.22 | 1.00 | 0.04 | 0.38 | 0.12 | 0.17 | 0.62 | 0.09 |
| 15 | | 0.72 | 1.00 | 0.11 | 1.00 | 0.04 | 0.30 | 0.45 | 1.00 | 0.56 | 0.90 |
| 18 | | 0.17 | 0.00 | 0.00 | 1.00 | 0.31 | 0.00 | 0.00 | 0.00 | 0.13 | 0.98 |
| 20 | | 0.29 | 0.00 | 0.00 | 0.19 | 0.17 | 0.00 | 0.00 | 0.00 | 0.00 | 0.00 |
| 23 | | 0.26 | 1.00 | 0.00 | 0.14 | 0.10 | 0.05 | 0.19 | 1.00 | 0.00 | 0.00 |

図3の横軸は、ユーザの過去14日分の推定期待閲覧ツイート確率 \hat{r} の平均値である。縦軸は実験ツイート回数に対する返信ツイート回数の割合で返信率 m と表している。 m は式(10)

のように表される。

$$m = \frac{\text{返信回数}}{\text{実験ツイート数}} \quad (10)$$

例えば、図2の一番右上の点に着目すると、この点は返信率が1、つまり全ての実験ツイートに対して返信を行っていることを表している。さらにこの点においては、本手法で算出した推定期待閲覧ツイート確率 \hat{r} も1であり、本手法で算出した確率 \hat{r} が1の時、実際のユーザの返信する確率も1であったということが分かる。

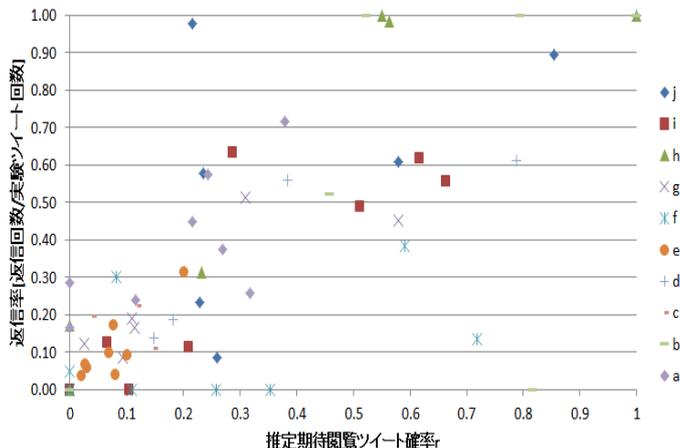


図2 推定期待閲覧ツイート確率 \hat{r} と返信率の関係

推定期待閲覧ツイート確率 \hat{r} の値と実際の返信率 m に着目すると、全体的には \hat{r} が低い場合は m も低く、 \hat{r} が高い場合は m も高い傾向が確認できた。

また図3のように、 \hat{r} が理想の値をとる場合の補助線を加えると、多くの点が線の上側に位置することが分かる。これはほとんどの時間において、ユーザはツイート投稿以外に閲覧のみをしていることが多いことを表している。しかし、ツイートをほとんどしない状態で、極端に閲覧数だけが低い点は、ほとんど見られなかった。

これらから、推定期待閲覧ツイート確率 \hat{r} はユーザがツイートを閲覧している確率を指し示していることが分かった。

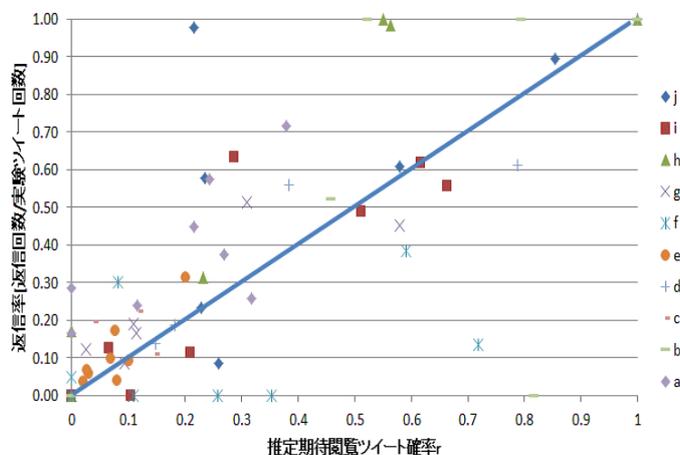


図3 推定期待閲覧ツイート確率 \hat{r} と返信率の関係 (補助線あり)

4.3 モデル化や推定による誤差

4.3.1 モデル化による誤差

本手法のモデル化では、アクションは一定間隔で行っているとして定義しているため、実際にはある時間帯に2つの投稿を連続で行った場合でも、そのユーザは30分に1回、1時間に計2回ツイートをしたとみなしている。しかし、実際には1回のアクションと呼べるような短い間隔でアクションを行っている場合も考えられる。そのため、そのようなアクション状況においては、実際より多くのツイートを閲覧していると判断してしまう。

また、一定閲覧数 w を 20 として実験を行ったが、この値が妥当かは今回の実験からは読み取ることができなかった。今回の実験において、ユーザによっては数時間前の実験ツイートに返信を行ったユーザも居た。そのため実際のユーザは一回のアクションでもっと多くのツイートを閲覧していることも考えられる。

さらに、図4のような場合は、 \hat{r} は上限値1をとるとしているが、実際にはその時間帯より前の時間帯のツイートを閲覧していると考えられる。

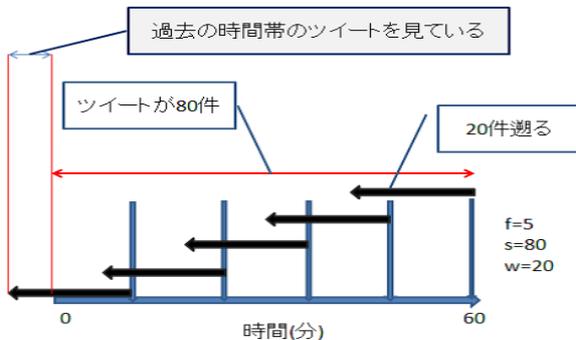


図4 モデル化による誤差

4.3.2 推定による誤差

3.の項でもふれたが、推定値には誤差が含まれている。特にアクション密度に関しては誤差を含んでいると考えられる。図3を見ると、 \hat{r} より大きい確率で返信を行っている点が多くみられる。このことから、ユーザは閲覧アクションのみをしていることが多いことが分かる。

4.4 結論

提案手法の推定値から算出した推定期待閲覧ツイート確率 \hat{r} は、実際にユーザがツイッターを閲覧している可能性が高い時間を指し示していることを確認できた。しかし、本手法で行ったモデル化では、全てのユーザは一律に一回のアクションにおいて一定数 w [件/回] のツイートを閲覧するとしたが、実際には w の値はユーザ固有のものであり、今回の実験で用いた値よりも大きい可能性が高いと考えられる。また、本手法ではツイートを全く行わないユーザに関しては、推定値を計算することができず、ツイート投稿タイミングを推薦することができない。

5. おわりに

Twitter を広告配信的に利用する際の問題はユーザに見逃されてしまうことである。そこで本研究では、ユーザのツイート

閲覧確率を時間帯ごとに推定し、閲覧可能性の高い時間帯を指し示すことで、適切な tweet 投稿タイミングを示すことができる手法を提案した。

具体的には、ユーザは閲覧、ツイートの二つのアクションをし、各アクションでは一定件数のツイートを閲覧するものとモデル化し、時間帯ごとの期待閲覧密度を求めた。この際、直接観測できないユーザタイムライン速度ならびに閲覧密度は、観測可能事象から推定する手法も提案した。実験から推定値の効果を測定し、推定値は実際のユーザの行動には相関があると確認できた。

推定期待閲覧ツイート確率 \hat{r} の値が大きい時間帯をユーザの目に留まるツイート投稿タイミングとして推薦する。

しかし、モデル化やパラメータの推定による誤差の改善、さらには最適なデータ収集期間の決定や、複数ユーザへの対応が今後の課題である。

文献

- [1] 竹村 光, 田島 敬史, “価値の時間依存性に基づくマイクロブログ記事の分類”, DEIM Forum 2012F3-1
- [2] 古賀健士郎, 庄司功樹, 小林亜樹, “タイミングを考慮した将来イベント tweet の再通知”, 第 74 回情処大 3P-4, No. 1, pp. 709-71 (2012-3)
- [3] 田中優美, 伊藤久洋, “Twitter におけるタイムライン閲覧時の不快感軽減の試み”, 全国大会講演論文集 2011(1), 167-169, 2011-03-02
- [4] 松尾潤, 川村秀憲, 鈴木恵二, “Twitter のコメント分析による広告の推薦”, 情報処理学会研究報告. ICS, [知能と複雑系] 2011-ICS-162(4), 1-6, 2011-03-01
- [5] 林智天, 川原圭博, 田村大, 森川博之, 青山友紀, “コンテキスト適応型コンテンツ配信における情報通知タイミングに関する検討”, 電子情報通信学会総合大会講演論文集 2005 年 通信 (1), 727, 2005-03-07