

# トレンドツイートにおけるリツイートの分析

黒滝 朝陽<sup>†</sup> 須鎗 弘樹<sup>†</sup>  
<sup>†</sup> 千葉大学大学院融合科学研究科

## 1. はじめに

近年,ソーシャルネットサービスが人気を博している.なかでも,Twitter におけるリツイート(以下,RT)は,情報伝播の手段として,広く使われている.鳥海ら[1]は,2011 年に起きた東日本大震災において情報共有手段であった RT を長期間分析し,震災発生效后,RT が短期間に短いスパンで行われるようになり,Twitter が積極的に活用されるようになったことを明らかにした.そこで,本論文では,短期間において爆発的に RT されたトレンドツイートに着目し, Twitter がさらに普及した最近において, RT が行われるタイミングはどう変化したのか分析する.

## 2. 研究手法

本研究では,Twitter API を用いて,ツイートを収集した.ツイートを収集する際には,120 時間以内に 10,000RT 以上されたツイートに限定して行った.

以下,モデル化の手順である.

1. ある元ツイートからリツイートされるまでの時間 $x(\text{min})$ の時間分布  $X: \{x_1, x_2, \dots, x_n\}$ とする.
2. データ解析環境 R における[mclust]パッケージで混合クラス数  $K$  を決定する.この際, 図 1 のように BIC (バイズ情報量基準)の数値が高い  $K$  を採用する.
3. 対数正規分布のパラメータ平均 $\mu$ ,分散 $\sigma^2$ ,混合比 $\xi$ の初期値を乱数で決定する.
4. EMアルゴリズムを用いて,混合対数正規分布を推定し, $X$ におけるパラメータの最尤値を推定する.

E-step: パラメータ $\mu, \sigma^2$ また,あるクラスの混合比 $\xi$ を用いて,あるツイート $x_i$ が  $k$  番目のクラスに所属する確率を

$$z_{ik} = \frac{\xi_k f(x_i | \mu_k, \sigma_k^2)}{\sum_{j=1}^K \xi_j f(x_i | \mu_j, \sigma_j^2)}$$

とし,その値を既知として,対数尤度の平均値 $Q(x)$ は,

$$Q(x) = \sum_{i=1}^N \sum_{k=1}^K z_{ik} \log\{\xi_k f(x_i | \mu_k, \sigma_k^2)\}$$

となる.

M-step:  $Q(x)$ の微分から得られる尤度方程式から最尤パラメータを求め, $\mu_k, \sigma_k^2, \xi_k$ を更新する.

5. E-step と M-step を交互に $Q(x)$ の変化量が閾値  $L=1.0 \times 10^{-6}$ より小さくなったとき,  $Q(x)$ は,収束したものとみなし,収束するまで更新を繰り返す.

## 3. 結果と分析

図 2 は,パラメータの対数尤度遷移図である.所々対数尤度値が大きくなるタイミングがあり,収束値に近づいていることがわかる.図 3 がその結果,混合対数正規分布を用いて

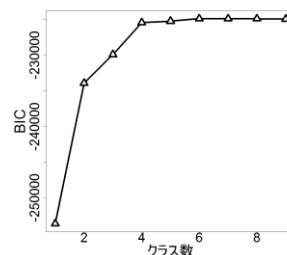


図 1. クラス推定結果

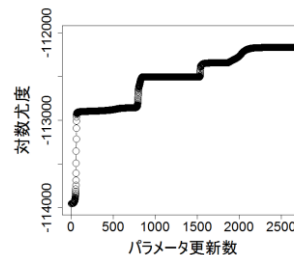


図 2. 対数尤度遷移図

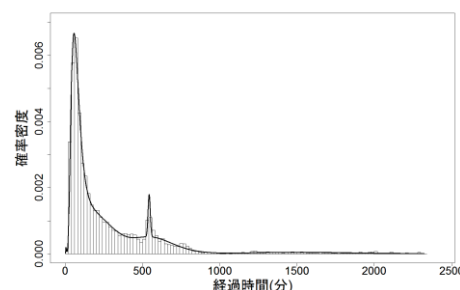


図 3. 混合対数正規分布でモデル化

モデル化できたグラフである.ここで,RT の最大ピークが現れる時刻を表す「最頻値」を分析し,RT がどのタイミングで一番されているのかを調べてみる. [1]によると,東日本大震災前はツイートされてから 10,000 秒以上後に,震災後は 2000~8000 秒後に RT が最もよくされ,震災前後で RT されるタイミングが変化したことがわかった. また,トレンドツイートにおいては,早い場合で 180 秒,平均をとると 2560 秒後に最大ピークを迎えるという結果になった.ツイートの詳細を見ると,芸能人アカウントなど 600 秒以内に最大ピークを迎えるツイートが半分近くあり,震災発生效后よりも最近では,さらに短時間で RT の最大ピークを迎えるように変化したと言えるだろう.これは, Twitter という情報共有ツールがさらに主流になり,ユーザ数が増加し続けていることに起因すると思われる.

## 4. まとめと今後の課題

短期間に多数 RT されたツイートデータを混合対数正規分布を用いてモデル化することができた.また,最近において,震災発生效后の期間よりも RT されるタイミングが早くなっていることが分かった.今後,データ数を増やし,より多くのピークを持つ RT に対するモデル化を目指していきたい.

## 参考文献

- [1] 鳥海不二夫,篠田孝祐,榊剛史,風間一洋,栗原聡,野田五十樹.東日本大震災時におけるリツイートの分析.情報処理学会研究報告,Vol.2012-ICS-168 No.3 2012