

# 日中時系列ニュース・ブログにおけるトピックモデルの推定と分析

胡 碩<sup>†</sup> 鄭 立儀<sup>†</sup> 高橋 佑介<sup>†</sup> 小池 大地<sup>†</sup> 牧田 健作<sup>†</sup>  
宇津呂武仁<sup>††</sup> 吉岡 真治<sup>†††</sup>

<sup>†</sup> 筑波大学大学院システム情報工学研究科 知能機能システム専攻  
<sup>††</sup> 筑波大学システム情報系 知能機能工学域  
<sup>†††</sup> 北海道大学大学院情報科学研究科 コンピュータサイエンス専攻

**あらまし** 本研究では、日本語および中国語の時系列ニュースおよびブログを対象として、教師なしの(時系列)トピックモデルを適用し、時系列のトピックの推移を推定することにより、二言語間あるいはニュース・ブログの間でのトピックの対応の有無、および、関心事項の差異を分析するアプローチをとる。特に、本論文では、日中時系列ニュースの間でトピック対応を同定する方式についてその詳細および評価結果を述べる。また、中国語において、多様なニュースサイトにおいてバーストするトピックを中心として、中国語ブログの収集を行い、中国語におけるニュースとブログの間での関心事項の違いや意見の有無について分析する。

**キーワード** ニュース, ブログ, トピックモデル, 時系列分析

## Estimation and Analysis of Topic Models in Time Series Japanese / Chinese News and Blogs

Shuo HU<sup>†</sup>, Liyi ZHENG<sup>†</sup>, Yusuke TAKAHASHI<sup>†</sup>, Daichi KOIKE<sup>†</sup>, Kensaku MAKITA<sup>†</sup>, Takehito  
UTSURO<sup>††</sup>, and Masaharu YOSHIOKA<sup>†††</sup>

<sup>†</sup> Grad. Sch. of Systems and Information Engineering, University of Tsukuba  
<sup>††</sup> Faculty of Engineering, Information and Systems, University of Tsukuba  
<sup>†††</sup> Grad. Sch. of Information Science and Technology, Hokkaido University

**Abstract** This paper studies issues regarding topic modeling of information flow in multilingual news and blog streams. If someone wants to find differences in the topics of Japanese and Chinese news / blog, it is usually necessary for him/her to carefully watch every article in Japanese and Chinese news streams at every moment. In such a situation, topic models such as LDA (Latent Dirichlet Allocation) and DTM (dynamic topic model) are quite effective in estimating distribution of topics over a document collection such as articles in a news / blog stream. Based on the results of estimating distribution of topics in Japanese / Chinese news streams, this paper proposes how to analyze cross-lingual alignment of topics in time series Japanese / Chinese news streams. This paper also proposes how to analyze differences of concerns and opinions among various Chinese news sites as well as blog hosts.

**Key words** news, blog, topic model, time series analysis

### 1. はじめに

ウェブ上の世界を始めとして、膨大な情報が溢れ、いわゆる情報爆発が起こっている。ウェブ上のニュースやブログの上に限っても、同様に多くの情報が流れている。これらのことを背景にして、時系列に沿って、情報を集約したり俯瞰的に把握するための技術が注目されている。例えば、本研究で利用した統計的トピックモデルのように、文書集合における主要なトピッ

クを推定する技術が確立されてきた。

トピックモデルにおいては、文書が生成される背景において、潜在的に複数のトピックが寄与していることを想定し、文書の生成尤度を高めるようにモデルのパラメータを訓練する。トピックモデルの一種である潜在的ディリクレ配分法 (LDA, Latent Dirichlet Allocation) [1] や DTM(dynamic topic model) [2] は、与えられた文書集合から、文書ごとのトピックの確率分布と、トピックごとの語の確率分布を学習する。

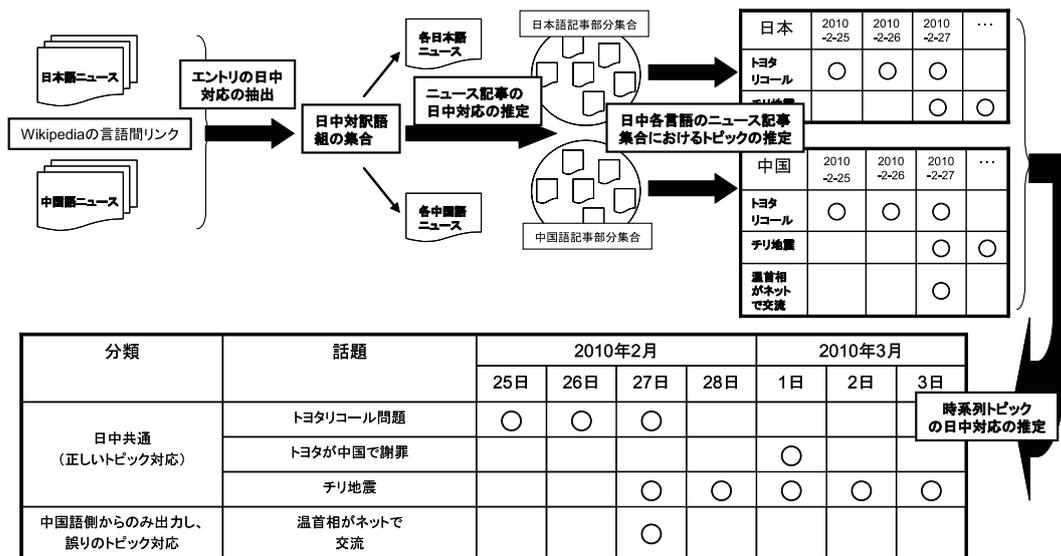


図1 日中時系列ニュースにおけるトピックモデルの推定と二言語間対応の同定の流れ

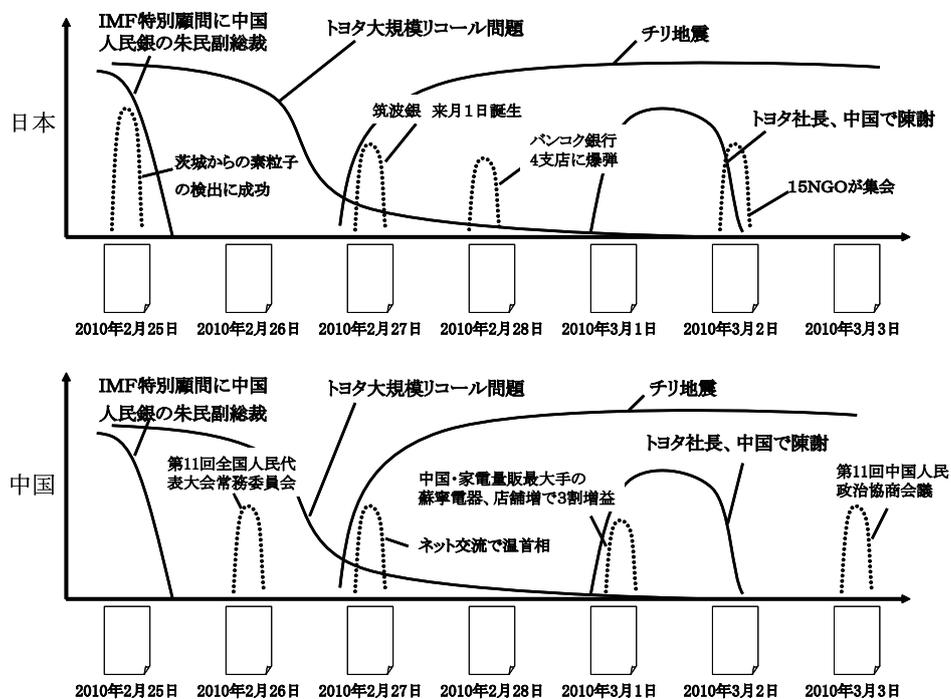


図2 日中時系列ニュースにおけるトピックの推定事例

本論文では、日本語および中国語の時系列ニュースおよびブログを対象として、教師なしの(時系列)トピックモデルを適用し、時系列のトピックの推移を推定することにより、二言語間あるいはニュース・ブログの間でのトピックの対応の有無、および、関心事項の差異を分析するアプローチをとる。特に、本論文では、日中時系列ニュースの間でトピック対応を同定する方式についてその詳細および評価結果を述べる。また、中国語において、多様なニュースサイトにおいてバーストするトピックを中心として、中国語ブログの収集を行い、中国語におけるニュースとブログの間での関心事項の違いや意見の有無について分析する。

以上の研究内容のうち、まず、3. においては、図1の流れに

示すように、日本語および中国語の二言語の時系列ニュースを対象として、各日において、DTMによってトピックの分布を推定する。そして、時系列に沿って継続的に報道されるトピックに対して、日中間でトピックの対応をとる手法を提案する。日中間でトピックの対応をとる際には、Wikipediaの言語間リンクを用いる。日本ニュースと中国ニュースの間では、あるトピックに関する報道期間が異なると、この話題についての関心度が異なる可能性があると言える。例えば、あるトピックについて、日本における報道日数が三日間であるのに対して、中国での報道日数が五日間であるならば、日本においてより、中国においてのほうが関心度は高いと言える。本研究では、このように、日中の時系列ニュースにおけるトピックの二言語間対応

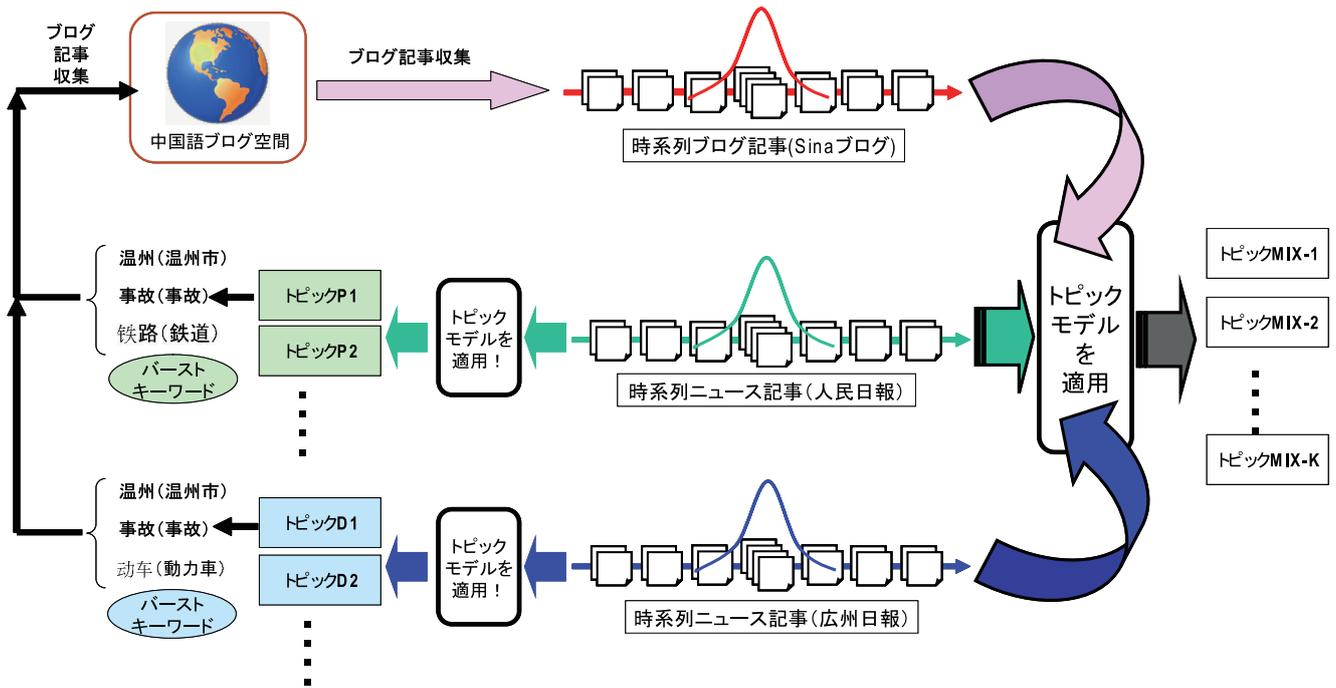


図3 トピックモデルを利用した時系列の中国語ニュース(人民日報・広州日報)・ブログの比較対照分析の流れ

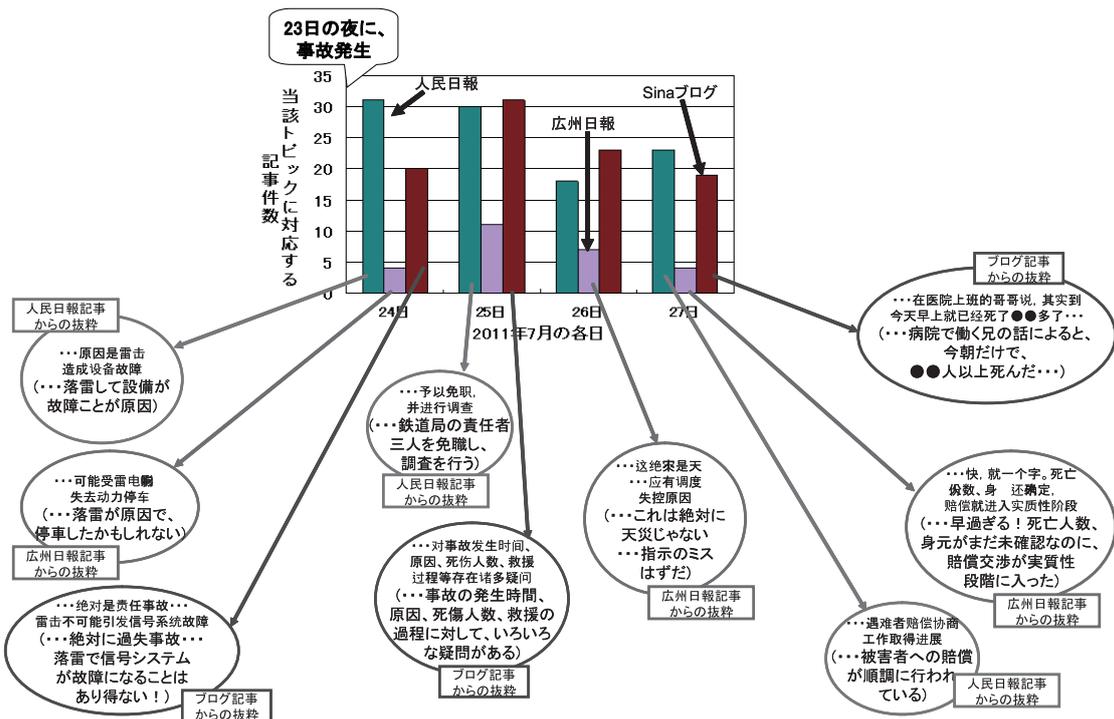


図4 中国語ニュース(人民日報・広州日報)・ブログにおける報道内容・関心・意見の比較対照分析結果の例

を分析することにより、日中間の関心や意見の差異を検出するための基盤技術を確立する。本研究により日中間でトピックの対応を推定した事例を図2に示す。

次に、4.においては、中国語において、報道姿勢、報道内容、政治的立場の異なる多様なニュースサイト、および、ブログを対象として、関心事項の違いや意見の有無について分析する。

具体的には、図3の流れに示すように、まず、各新聞社ごとに時系列のニュース記事を収集し、LDAを適用してトピックの分布を推定する。次に、ニュースにおいてパーストするトピックを中心として、中国語ブログの収集を行った後、ニュース記事およびブログ記事を全て混合した記事集合に対して再度LDAを適用し、新聞社・ブログといった情報源間のトピック分布の

異なりを分析する。分析結果の事例を図 4 に示す。

## 2. トピックのモデル化

### 2.1 トピックモデル

本研究では、トピックモデルとして、潜在的ディリクレ配分法 (LDA, Latent Dirichlet Allocation), および, DTM(dynamic topic model) を用いる。これらのトピックモデルにおいては、語  $w$  の列によって表現される時間情報を含んだ文書の集合と、トピック数  $K$  を入力とし、各単位時間について、各トピック  $z_n (n = 1, \dots, K)$  における語  $w$  の確率分布  $p(w|z_n) (w \in V)$ , 及び、各文書  $b$  におけるトピック  $z_n$  の確率分布  $p(z_n|b) (n = 1, \dots, K)$  を推定する。ここで、 $V$  は文書中に出現する語の集合である。なお、DTM は、LDA とは異なり、文書集合中の時系列情報を考慮しているため、日付等の単位時間を超えて同一トピックを追跡可能である。

本論文では、 $p(w|z_n) (w \in V)$ , 及び、 $p(z_n|b) (n = 1, \dots, K)$  の推定においては、LDA によりトピックモデルを推定するツール GibbsLDA++<sup>(注1)</sup>, および、DTM によりトピックモデルを推定するツール<sup>(注2)</sup> を用いた。ハイパーパラメータ  $\alpha$ ,  $\beta$  と、トピック数  $K$  は、LDA においては、 $\alpha = 2.5$ ,  $\beta = 0.1$ ,  $K = 20$  とした。DTM においては、予備実験を通して調整を行い、 $\alpha = 0.01$ ,  $\beta = 0.1$ ,  $K = 10$  とした。

### 2.2 文書とトピックの対応付け

本研究では、一日ごとに、各トピックに対してニュース記事を一対一で割り当てること、トピックごとのニュース記事集合の要素数を測ることとした。

ある日における文書集合を  $D$ , トピック数を  $K$ , 一つの文書を  $d (d \in D)$  とすると、トピック  $z_n (n = 1, \dots, K)$  のニュース記事集合  $D(z_n)$  は以下の式で表される。

$$D(z_n) = \{d \in D | z_n = \underset{z_u (u=1, \dots, K)}{\operatorname{argmax}} p(z_u|d)\}$$

これはつまり、文書  $d$  におけるトピックの分布において、文書  $d$  に確率が最大のトピックを割り当てていることになる。

## 3. 日中時系列ニュースにおけるトピックの二言語間対応の推定

### 3.1 ニュース記事の日中対応の推定

ニュース記事の日中対応推定においては、一日の単位で、日本語ニュース記事と中国語ニュース記事との間で、共有日中対訳語組数<sup>(注3)</sup>を求めて、一定の値  $\theta_{JC}$  (本稿では、 $\theta_{JC}$  を 10 に設定した) 以上の共有日中対訳語組数を持つ日本語ニュース記事と中国語ニュース記事の組に対して、ニュース記事の日中対

(注1) : <http://gibbslda.sourceforge.net/>

(注2) : <http://www.cs.princeton.edu/~blei/topicmodeling.html>

(注3) : 中国語の簡体字ニュース記事中に出現する簡体字キーワードに対して、Wikipedia の言語間リンクを用いてエントリの日中対応を抽出し、日中対訳語組の集合を作成する。ただし、中国語版 Wikipedia においては、簡体字キーワードは、繁体字のエンタイトルを持つエンタリ中のリダイレクトとして登録されているため、これを利用する。2009 年 6 月 1 日 ~ 2010 年 5 月 31 日の 1 年分の 157,945 日本語ニュース記事、および、204,595 中国語ニュース記事を利用することにより、78,519 組の日中対訳語組を抽出した。

応を推定する。日中ニュース記事の間で、共有される対訳語組の集合の大きさ  $N_{JC}(d_J, d_C)$  を以下のように定義する。

$$N_{JC}(d_J, d_C) = |\{(J, S) \in JSW \mid J \text{ は } d_J \text{ 中に出現する. } S \text{ は } d_C \text{ 中に出現する.}\}|$$

ここで、 $d_J$  は日本語ニュース記事である。 $d_C$  は中国語ニュース記事である。 $\theta_{JC}$  以上の共有日中対訳語組数を持つ日中ニュース記事組の集合  $DD_{JC}(\theta_{JC})$  および  $DD_{CJ}(\theta_{JC})$  を以下の式で定義する。

$$DD_{JC}(\theta_{JC}) = \left\{ \langle d_J, d_C \rangle \mid N_{JC}(d_J, d_C) \geq \theta_{JC}, \right. \\ \left. d_C = \underset{d'_C}{\operatorname{argmax}} N_{JC}(d_J, d'_C) \right\}$$

$$DD_{CJ}(\theta_{JC}) = \left\{ \langle d_J, d_C \rangle \mid N_{JC}(d_J, d_C) \geq \theta_{JC}, \right. \\ \left. d_J = \underset{d'_J}{\operatorname{argmax}} N_{JC}(d'_J, d_C) \right\}$$

### 3.2 トピックの日中対応の推定

本節では、まず、前節で作成した日中ニュース記事組の集合  $DD_{JC}$  もしくは  $DD_{CJ}$  中に含まれる日本語記事  $d_J$  および  $d_C$  のみを対象として、日中それぞれの言語で DTM によりトピック推定を行い、 $i$  番目の日における日本語トピック集合  $TT_J^i$  および中国語トピック集合  $TT_C^i$  を求める。そして、 $TT_J^i$  と  $TT_C^i$  の間で、以下の手順により、トピック組の対応を推定する。

まず、日本語トピック  $t_J \in TT_J^i$  に対して、 $P(t_J|d_J) \geq \theta_t$  ( $\theta_t = 0.6$ ) という条件を満たす記事  $d_J$  を集める。同様に、中国語トピック  $t_C \in TT_C^i$  に対して、 $P(t_C|d_C) \geq \theta_t$  ( $\theta_t = 0.6$ ) という条件を満たす記事  $d_C$  を集める。そして、日本語トピック  $t_J$  と中国語トピック  $t_C$  の間で  $P(t_J|d_J) \geq \theta_t$  および  $P(t_C|d_C) \geq \theta_t$  を満たす記事組  $\langle d_J, d_C \rangle$  のうち、

$DD_{JC}(\theta_{JC})$  または  $DD_{CJ}(\theta_{JC})$  に含まれる記事組を抽出し、その要素数を  $M_{JC}(t_J, t_C, \theta_t, \theta_{JC})$  とする。そして、以下の式により、日本語トピック集合  $TT_J^i$  中のトピックのうち、中国語トピック  $t_C$  に対応するものを同定する。同様に、中国語トピック集合  $TT_C^i$  中のトピックのうち、日本語トピック  $t_J$  に対応するものを同定する。

$$TA_J(t_C, TT_J^i, \theta_t, \theta_{JC}) = \begin{cases} \text{出力なし} & \left( \max_{t_J \in TT_J^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) = 1 \right) \\ \operatorname{argmax}_{t_J \in TT_J^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) & \left( \max_{t_J \in TT_J^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) \geq 2 \right) \end{cases}$$

$$TA_C(t_J, TT_C^i, \theta_t, \theta_{JC}) = \begin{cases} \text{出力なし} & \left( \max_{t_C \in TT_C^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) = 1 \right) \\ \operatorname{argmax}_{t_C \in TT_C^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) & \left( \max_{t_C \in TT_C^i} M_{JC}(t_J, t_C, \theta_t, \theta_{JC}) \geq 2 \right) \end{cases}$$

### 3.3 時系列トピックの日中対応の推定

時系列トピックの日中対応の推定においては、ある連続期

表1 各日におけるトピックの日中対応の出力およびトピック単位での評価結果

分類	話題	2010年2月				2010年3月			
		25日	26日	27日	28日	1日	2日	3日	4日~23日
日中・中日双方向 でトピック対応を検出 (正しいトピック対応)	トヨタリコール 問題	○	○	○					日中間の トピック対応 は検出せず
	トヨタが 中国で謝罪					○			
	チリ地震			○	○	○	○	○	
	IMF 特別顧問に 中国人民銀の副総裁	○							
中日方向 にのみ トピック対応 を検出	温首相が ネットで交流			× (中国語側のみ 適切なトピックを推定. 中国語 10 記事が 日本語 1 記事に対応)					
日中・中日双方向 でトピック対応を検出 (誤りのトピック対応)	自国経済について のニュース	× (全ての日において、トピック対応を検出)							
トピック単位での評価結果:		日本語トピック	80.0% (4/5)	中国語トピック	66.7% (4/6)				

間における日本語と中国語の時系列ニュースにおいて、複数の日に渡るトピックを日中間で対応付ける。連続期間の各日において、DTM モデルで生成した日本語トピック集合の列を  $Q_J = TT_J^1, TT_J^2, \dots, TT_J^n$  とし、中国語トピック集合の列を  $Q_C = TT_C^1, TT_C^2, \dots, TT_C^n$  とする。  $TT_J^i$  は 1 番目の日の日本語トピック集合である。時系列トピックの日中対応の推定は、以下のように行った。

- $i$  番目の日において、任意の日本語トピック  $t_J \in TT_J^i$  に対して、同日において対応付けた中国語トピック  $TA_C(t_J, TT_C^i, \theta_t, \theta_{JC})$  を推定する。
- 同様に、任意の中国語トピック  $t_C \in TT_C^i$  に対して、同日において対応付けた日本語トピック  $TA_J(t_C, TT_J^i, \theta_t, \theta_{JC})$  を推定する。

以上の手順を、 $i = 1, \dots, n$  について行うことにより、連続期間の全体を通して時系列トピックの日中対応を推定する。

### 3.4 分析および評価

ニュース記事の日中対応の推定は、2010年2月25日から、3月23日までの一ヶ月間の読売新聞<sup>(注4)</sup>、日経新聞<sup>(注5)</sup>、および朝日新聞<sup>(注6)</sup>の三紙の12,288ニュース記事、および、人民日報<sup>(注7)</sup>の22,049ニュース記事を対象として行った。この一ヶ月において、3.1節の手法により、日中二言語間で対応が付けられた日本語ニュース記事数は791件、中国語ニュース記事数は1,361件であった。2010年2月25日から、3月3日までの一週間のニュース記事の評価対象として、ニュース記事の日中対応推定結果の評価を行った。この一週間での日本語ニュース記事数は3,278件、中国語ニュース記事数は6,110件であったが、日本語から中国語へのニュース記事対応付け精度は63.2% (247記事中156記事が正解)、中国語から日本語へのニュース記事対応付け精度は53.3% (364記事中194記事が正解) となった。こ

の結果において、大きく精度を下げている要因は、日中両国における自国経済についてのニュースである。これらのニュースは、日中両国間において共通の内容とはならないにも関わらず、数値や経済用語等が共通に出現することが要因となって、日中間において誤った記事対応として多数出力されている。実際に、これらの自国経済についてのニュース記事を除外した後算出した日中間記事対応の精度は、日本語から中国語へのニュース記事対応付け精度が88.1% (177記事中156記事が正解)、中国語から日本語へのニュース記事対応付け精度は86.6% (224記事中194記事が正解) に改善した。これらの自国経済ニュースは、年間を通じて定常的に報道されるため、今後は、トピックのバースト [3] を検出する手法 [4] を併用することにより、除去することが適切であると考えられる。

次に、日中間で対応が付けられた日本語ニュース記事791記事、中国語ニュース記事1,361記事を対象としてDTMを適用し生成したトピックに対して、日中間でトピックの対応の推定を行った結果の抜粋を表1に示す。この結果から分かるように、全ての日において、自国経済についてのニュースに相当するトピックが日中間で対応付けられており、トピック単位ではこれらはすべて誤りの対応となった。また、今回評価対象とした一ヶ月間において、日中間でトピック対応を検出したのは最初の一週間だけであった。実際に、日中両言語において共通の大きな話題が観測されたのは最初の一週間だけであった。この一週間において、中国でのみ盛り上がったトピック「温首相がネットで交流」については、中国語側からのみトピック対応が検出されたが、誤りのトピック対応と判定した。その他の4種類のトピック対応については、適切なトピック対応を出力した。

## 4. トピックモデルを利用した中国語ニュース (人民日報・広州日報) とブログの比較対照分析

### 4.1 分析手順の概要

本節では、図3の手順に沿って、中国語において、報道姿勢、報道内容、政治的立場の異なる多様なニュースサイト、および、

(注4) : <http://www.yomiuri.co.jp/>

(注5) : <http://www.nikkei.com/>

(注6) : <http://www.asahi.com/>

(注7) : <http://www.people.com.cn/>

ブログを対象として、関心事項の違いや意見の有無について分析する。具体的には、ニュースサイトとしては、中国政府の立場に最も近い人民日報、および、中国政府の立場から比較的離れた位置にある広州日報<sup>(注8)</sup>をとりあげる。一方、中国語ブログホストとしては、Sina ブログホスト<sup>(注9)</sup>を対象とする。

まず、2011年7月18日から31日の二週間に渡って、人民日報5,937記事、広州日報1,446記事を収集した。次に、各新聞社の記事集合に対して独立にLDAを適用した。その結果、それぞれ20個のトピックのうち、ある程度記事集合の内容のまとまりが確認できたトピックとして、人民日報からは、

中国列車事故、英国電話盗聴事件、ノルウェーテロ事件、アメリカ軍関連、社会犯罪全般、消費生活関連

の6トピックが、一方、広州日報からは、

中国列車事故、ノルウェーテロ事件、社会犯罪全般、消費生活関連、スポーツ全般

の5トピックが、それぞれ得られた。次に、これらのトピックに特徴的なキーワードを手動で選定し、人民日報からは26キーワードを、広州日報からは21キーワードを、それぞれ、ブログ記事の収集の際のクエリとした。これらのクエリを含む中国語ブログ記事の収集においては、Yahoo! Search BOSS API<sup>(注10)</sup>を利用し、Sina ブログホストのドメインを対象としてブログ記事の収集を行った。その結果、収集されたブログ記事のうち、2011年7月18日から31日の二週間の間に投稿されたものは367記事となった。この367ブログ記事、人民日報5,937記事、広州日報1,446記事を混合した記事集合に対してLDAを適用したところ、20個のトピックのうち、ある程度記事集合の内容のまとまりが確認できたトピックは

中国列車事故、英国電話盗聴事件、ノルウェーテロ事件、消費生活関連

の4トピックとなった。このうち、「中国列車事故」に関する記事数は、人民日報212記事、広州日報56記事、Sina ブログ159記事であり、他のトピックと比較して、十分な数のブログ記事が収集できたため、次節ではこのトピックを対象として、人民日報、広州日報、Sina ブログの比較対照分析を行う。

#### 4.2 分析結果

当該の列車事故は7月23日に発生し、主として、24日から28日の5日間の記事数がバースト傾向であった。そこで、記事内容の変遷と情報源の間の対比が象徴的であった24日から27日の4日間について、当該トピックに対応付けられた記事数の推移、および、主要な記事の内容の抜粋を図4に示す。この図から分かるように、人民日報の記事の内容が相対的に保守的な内容で占められるのに対して、広州日報の記事においては、当局に対する懐疑的な論調の内容が散見された。さらに、Sina ブ

ログにおいては、事故発生直後から、当局の発表に対する疑念の声が観測され、公共のメディアでは報道されないような口コミ情報も散見された。

なお、今回の分析においては、ブログ記事収集の際に用いたキーワードを手動で選定したが、今後は、このキーワード選定の過程を自動化する方式について取り組む予定である。

## 5. 関連研究

文献[5]は、各日において独立に推定されたLDAのトピックを時系列方向に繋げる枠組みを提案している。一方、時系列方向のトピックの繋がりを理論的にモデル化したトピックモデルとして、On-line LDA [6]などがある。また、文献[7]においては、4.の手法と同様に、震災に関するニュース記事・ブログ記事を収集し混合した文書集合に対してトピックモデルを適用し、ニュース・ブログの間での話題の相関、および、時系列での話題の変遷の様子を分析している。

## 6. おわりに

本論文では、日中二言語の時系列ニュースを対象として、各日において、DTMによってトピックの分布を推定した。そして、時系列に沿って継続的に報道されるトピックに対して、日中間でトピックの対応をとる手法を提案し、その有効性を示した。また、中国語において、多様なニュースサイトにおいてバーストするトピックを中心として、中国語ブログの収集を行い、中国語におけるニュースとブログの間での関心事項の違いや意見の有無について分析した。今後は、文献[8]の成果等をふまえて、日中両国のブログにおいて共通に関心を持たれている事項について、意見の違いの分析を行う方式を確立する。

## 文 献

- [1] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent Dirichlet allocation. *Journal of Machine Learning Research*, Vol. 3, pp. 993–1022, 2003.
- [2] D. M. Blei and J. D. Lafferty. Dynamic topic models. In *Proc. 23rd ICML*, pp. 113–120, 2006.
- [3] J. Kleinberg. Bursty and hierarchical structure in streams. In *Proc. 8th SIGKDD*, pp. 91–101, 2002.
- [4] 高橋佑介, 横本大輔, 宇津呂武仁, 吉岡真治, 河田容英, 神門典子, 福原知宏, 中川裕志, 清田陽司. 時系列トピックモデルにおけるバーストの同定. 第4回 DEIM フォーラム論文集, 2012.
- [5] 芹澤翠, 小林一郎. 文書内のトピック数を考慮したトピック追跡の試み. 言語処理学会第18回年次大会論文集, pp. 1196–1199, 2012.
- [6] L. AlSumait, D. Bardara, and C. Domeniconi. On-Line LDA: Adaptive topic models for mining text streams with applications to topic detection and tracking. In *Proc. 8th ICDM*, pp. 3–12, 2008.
- [7] 小池大地, 横本大輔, 牧田健作, 鈴木浩子, 宇津呂武仁, 河田容英, 吉岡真治, 神門典子, 福原知宏, 中川裕志, 清田陽司, 関洋平. ニュース・ブログにおける話題の相関と変遷の分析 — 震災に関する話題を例題として —. 第4回 DEIM フォーラム論文集, 2012.
- [8] 中崎寛之, 川場真理子, 横本大輔, 宇津呂武仁, 福原知宏. 多言語 Wikipedia エントリを知識源とする特定トピックの日英ブログサイト検索と日英対照ブログ分析. *人工知能学会論文誌*, Vol. 25, No. 5, pp. 613–622, 2010.

(注8) : <http://www.dayoo.com/>

(注9) : <http://blog.sina.com.cn/>

(注10) : <http://developer.yahoo.com/search/boss/>