Vol.2014-MPS-100 No.3
2014/9/25

複数のトピックの時間的依存関係を考慮した時系列トピックモデル

佐々木 謙太朗[1,a)]　吉川 大弘[1]　古橋 武[1]

概要：本稿では，時系列文書における複数の話題の時間的依存関係を考慮したトピックモデルを提案する．ニュース記事やブログ記事，SNS におけるユーザの投稿などといった時系列文書において，各トピックは互いに依存し合いながら時間と共に変化し，ある時刻において発生/消滅/結合/分離することがある．しかし既存のモデルでは，各トピックが独立に発展していくと仮定しており，これらの変化の一部分しか捉えることができない．そこで本稿では，ある時刻におけるトピックが，一時刻前のすべてのトピックに依存し得ると仮定した新しいトピックモデルを提案する．提案モデルは，時系列文書中のトピックの発生/消滅/結合/分離すべてを捉えることができ，かつ従来のモデルと比較して，より適切に時系列文書をモデル化できることを示す．

キーワード：トピックモデル，時間発展，時系列文書

Time Series Topic Model Considering Dependence to Multiple Topics

Sasaki Kentaro[1,a)]　Yoshikawa Tomohiro[1]　Furuhashi Takeshi[1]

Abstract: This pater proposes a topic model that considers dependence to multiple topics. In time series documents such as news, blog articles, and SNS user posts, topics evolve with depending on one another, and they can die, be born, merge, or split at any time. The conventional models cannot model the evolution of all of the above aspects because they assume that each topic depends on only one previous topic. In this paper, we propose a new topic model which assumes that a topic can depend on previous multiple topics. This paper shows that the proposed topic model can capture the topic evolution of death, birth, merger, and split and can model time series documents more adequately than the conventional models.

Keywords: Topic Model, Time Evolution, Time Series Document

1. はじめに

近年，Web の発展と共に，ニュース記事やブログ記事，SNS におけるユーザの投稿など，時系列的な文書が大量に生成されるようになった．これらの文書の内容をすべて把握することは困難であるため，いつどのような事が話題になり，それがどのように発展したかを追跡することを目的とした研究が数多く報告されている．それらの中でも，時系列トピックモデルに関する研究が近年注目され，また成果を挙げている [1], [2], [4], [10], [14], [15]．ここでトピックモデルとは，bag-of-words を仮定した，文書の生成過程を確率的にモデル化した言語モデルである．代表的なトピックモデルとしては，Latent Dirichlet Allocation (LDA) がある [5]．LDA では，各文書がそれぞれ固有のトピック比率を持ち，その比率に従ってトピックが生成され，さらにそれらのトピックに固有の単語分布に従って各単語が生成されることをモデル化している．また時系列トピックモデルとは，時間発展を考慮したトピックモデルであり，トピックの比率やトピックの内容に相当する単語分布が，時間に依存すると仮定したものである．時系列トピックモデルを用いることにより，時間の経過に伴う文書集合中のト

[1] 名古屋大学大学院工学研究科
[a)] sasaki@cmplx.cse.nagoya-u.ac.jp

ピックの発展を追跡することが可能である．

時系列文書におけるトピックは，互いに依存し合いながら時間と共に発展していく．例えば，ニュース記事などにおいて，書き手が政治に関する事柄を書く際，それまでの政治的動向だけでなく，経済や社会の動向も考慮する場合が考えられる．また逆に，法律の改正といった政治的動向があった場合，それが経済や社会にどのような影響を与えるかについての記事が書かれることもある．このように，時間の経過と共に，あるトピックに別のトピックが結合したり，分離して複数のトピックへと発展したりすることがある．また，次第に話題にされなくなり消滅するトピックもあれば，地震のような突発的な出来事に関するトピックが同時多発的に発生したりすることも考えられる．しかし既存のモデルの多くは，ある時刻におけるトピック k が，その前の時刻におけるトピック k にのみ依存すると仮定している [1], [2], [4], [15]．この仮定の上では，各トピックは独立に発展していくことになり，実際のトピックの結合や分離といった発展を捉えることができない．

本稿では，トピックの発生と消滅を捉えることが可能である Temporal Dirichlet Process Mixture (TDPM) を拡張した，Multi-dependent Temporal Dirichlet Process Mixture (MdTDPM)，およびその学習方法を提案し，上述の問題に対処する．提案モデルは，複数のトピックが互いに依存し合いながら，時間と共に発展していくと仮定することで，実際のトピックの発展をより適切に捉えることができると期待できる．また，提案モデルを用いることで，時系列文書において発生/消滅/結合/分離を含めたトピックの追跡が可能となる．実際のニュース記事を用いた実験により，提案手法が既存の手法と比べて，より適切に時系列文書のモデル化が可能であることを示す．

2. 従来モデル

2.1 Dirichlet Process Mixture Model

本節では，代表的なノンパラメトリックベイズモデルであり，提案モデルのベースでもある Dirichlet Process Mixture Model (DPM) について説明する．DPM は，Dirichlet Process (DP) を事前分布に導入した無限混合モデルであり，データの複雑さに応じて自動的にトピック数を推定することが可能である．ここで，DP は確率分布に対する分布であり [6]，基底分布 G_0 と集中度パラメータ γ によって定義される．離散確率分布 G が DP に従う時，$G \sim DP(\gamma, G_0)$ と表記する．集中度パラメータ γ が大きいほど，得られる確率分布 G は基底分布 G_0 に近い離散分布となる．

DP の構成法としては，主に Stick-Breaking Process[9] と Chinese Resutaurant Process (CRP)[3] の二つがあるが，後述する学習方法であるギブスサンプリング [8] との相性のよさから，本稿では CRP を用いる．CRP において，1 から i-1 番目のデータ $x_{1:i-1}$ が属するトピック $z_{1:i-1}$ が決まっているとき，i 番目のデータ x_i が属するトピック z_i は以下の確率に従う．

$$P(z_i = k | z_{1:i-1}, \gamma) = \frac{m_k}{i + \gamma - 1} \tag{1}$$

$$P(z_i = k_{new} | z_{1:i-1}, \gamma) = \frac{\gamma}{i + \gamma - 1} \tag{2}$$

ここで，m_k はトピック k に属するデータの数であり，k_{new} は $z_{1:i-1}$ の中で一度も出現しない新しいトピックである．CRP は，各データを客，各トピックをテーブルと見立てて，中華料理店において新しく入店した客は，既に m_k 人の客が座っているテーブルに $m_k/(i+\gamma-1)$ の確率で着席し，誰も座っていないテーブルに $\gamma/(i+\gamma-1)$ の確率で着席することに例えられる．

CRP を用いると，DPM における M 個のデータ生成過程は，$\boldsymbol{z} = z_{1:M}$ として以下のように表現することができる．

(i) $\boldsymbol{z} \sim \mathrm{CRP}(\gamma)$
(ii) $\phi_k | G_0 \sim G_0$
(iii) for $i = 1, ..., M$, $x_i \sim p(x | \phi_{z_i})$

2.2 Temporal Dirichlet Process Mixture

DPM では，すべてのデータは交換可能であると仮定している．しかし，このような仮定は時間と共に変化するような時系列文書や音楽，ビデオデータに適していない．これに対して Ahmed らは，DPM を拡張して，データを一定の期間ごとに分割し，同一期間内ではデータは交換可能であるが，異なる期間ではデータの交換はできないと仮定した Temporal Dirichlet Process Mixture (TDPM) を提案している [1]．これにより，TDPM はトピックの発生・消滅を扱うことのできるモデルとなっている．

TDPM では，CRP の代わりに Recurrent Chinese Restaurant Process (RCRP) を用いる．RCRP は，CRP を拡張してトピックの人気度の時間依存性を考慮できるようにしたもので，時刻 t における i 番目のデータ $x_{t,i}$ が属するトピック $z_{t,i}$ は以下の確率に従う．

$$P(z_{t,i} = k \in I_t \cup I_{t-1} | \boldsymbol{z}_{t-1}, z_{t,1:i-1}, \gamma) = \frac{m_{t-1,k} + m_{t,k}}{M_{t-1} + i + \gamma - 1} \tag{3}$$

$$P(z_{t,i} = k \notin I_t \cup I_{t-1} | \boldsymbol{z}_{t-1}, z_{t,1:i-1}, \gamma) = \frac{\gamma}{M_{t-1} + i + \gamma - 1} \tag{4}$$

ここで，$m_{t,k}$ は時刻 t においてトピック k に属するデータの数，M_t は時刻 t におけるデータの総数，$\boldsymbol{z}_t$ は時刻 t におけるトピックの系列を表す．ただし，$k \notin I_t$ ならば $m_{t,k} = 0$，$k \notin I_{t-1}$ ならば $m_{t-1,k} = 0$ とする．また，I_t は時刻 t において少なくとも一つのデータが属するトピックの集合である．RCRP では，データ $x_{t,i}$ がトピック k に属する確率は，現在の時刻におけるトピック k の人気度 $m_{t,k}$

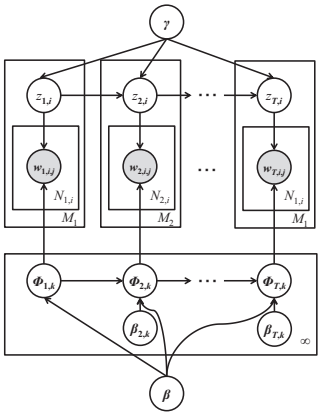


図 1 言語モデルに拡張した TDPM のグラフィカルモデル

だけでなく，一時刻前の人気度 $m_{t-1,k}$ も影響する．すなわち，一時刻前で人気だったトピックは，現在の時刻においても人気になりやすい．

RCRP を用いると，TDPM における時刻 t でのデータの生成過程は，以下のように表現することができる．

(i) $\boldsymbol{z}_t \sim \mathrm{RCRP}(\gamma, \boldsymbol{z}_{t-1})$
(ii) for $k \in I_t$,
$\phi_{t,k}|\phi_{t-1,k} \sim P(\phi_{t,k}|\phi_{t-1,k})$ if $k \in I_{t-1}$ or
$\phi_{t,k}|G_0 \sim G_0(\phi_{t,k})$ if $k \notin I_{t-1}$
(iii) for $i = 1, ..., N_t$,
$x_{t,i} \sim p(x|\phi_{t,z_i})$

TDPM では，$k \in I_{t-1}$ のとき，データを生成する分布 $\phi_{t,k}$ が一時刻前の分布 $\phi_{t-1,k}$ に依存しているように，トピックの人気度の時間依存性だけでなく，トピック内容の時間依存性も考慮している．また，$k \in I_t$ かつ $k \notin I_{t-1}$ のとき，時刻 t においてトピック k が発生したとみなすことができ，$k \notin I_t$ かつ $k \in I_{t-1}$ のとき，時刻 t においてトピック k が消滅したとみなすことができる．

3. 提案手法

本章では，TDPM を拡張し，ある時刻におけるトピックが一時刻前の複数のトピックに依存するという仮定を加えた新しいモデル MdTDPM と，その学習方法を提案する．

3.1 TDPM の言語モデルへの拡張

初めに，時系列文書をモデル化するために，TDPM を言語モデルに拡張することを考える．時刻 t における i 番目の文書 $d_{t,i}$（単語数 $N_{t,i}$) を，その文書が含む単語の集合 $\boldsymbol{w}_{t,i} = \{w_{t,i,j}\}_{j=1}^{N_{t,i}}$ によって表す．各文書にはそれぞれ一つのトピック $z_{t,i}$ が割り当てられ，そのトピックに対応する単語分布 $\phi_{t,z_{t,i}}$ に従って各単語 $w_{t,i,j}$ が生成されるとする．岩田らの手法 [16] の考え方に基づき，単語分布 $\phi_{t,z_{t,i}}$ は，$z_{t,i} = k \in I_{t-1}$ のとき，精度（分散の逆数）が $\beta_{t,k}$，平均が一時刻前の単語分布の推定値 $\hat{\phi}_{t-1,k}$ である以下のディリクレ分布に従って生成されるとする．

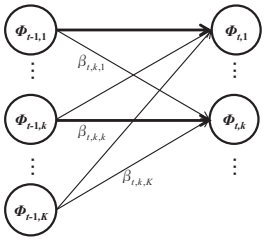


図 2 MdTDPM におけるトピックの依存関係

$$P(\phi_{t,k\in I_{t-1}}|\hat{\phi}_{t-1,k}, \beta_{t,k}) \propto \prod_v \phi_{t,k,v}^{\beta_{t,k}\hat{\phi}_{t-1,k,v}-1} \quad (5)$$

ここで，$\phi_{t,k,v}$ は，$\phi_{t,k}$ における単語 v の生成確率である．$\beta_{t,k}$ は，一時刻前のトピックへの依存度と考えることができ，これが大きいほど $\phi_{t-1,k}$ に近い分布が生成されやすくなる．$z_{t,i} = k \notin I_{t-1}$ のときは，ハイパーパラメータを β とするディリクレ分布に従って生成されるとする．

$$\phi_{t,k} \sim \mathrm{Dirichlet}(\beta) \quad (6)$$

以上の仮定に基づいて，TDPM を言語モデルに拡張したときのグラフィカルモデルを図 1 に示す．また，時刻 t における文書の生成過程は以下のように表現することができる．

(i) $\boldsymbol{z}_t \sim \mathrm{RCRP}(\gamma, \boldsymbol{z}_{t-1})$
(ii) for each topic $k \in I_t$,
$\phi_{t,k} \sim$Dirichlet$(\beta_{t,k}\hat{\phi}_{t-1,k})$ if $k \in I_{t-1}$or
$\phi_{t,k} \sim$Dirichlet(β) if $k \notin I_{t-1}$
(iii) for each document $i = 1, ..., M_t$,
for each word $j = 1, ..., N_{t,i}$,
$w_{t,i,j} \sim$Multinomial$(\phi_{z_{t,i}})$

3.2 MdTDPM

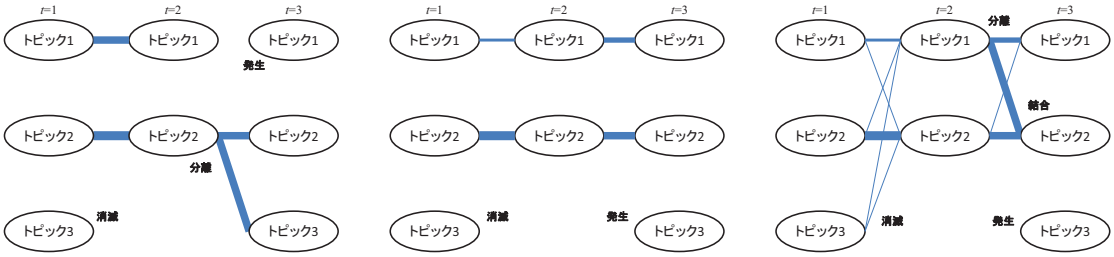
MdTDPM では，複数のトピック間の依存関係を考慮するために，$k \in I_{t-1}$ のとき，上述の単語分布 $\phi_{t,k}$ が，一時刻前のトピックの単語分布 $\{\phi_{t-1,k'}\}_{k'=1}^K$ の重み付き和をハイパーパラメータとする，以下のディリクレ分布から生成されると仮定する．

$$\phi_{t,k} \sim \mathrm{Dirichlet}(\sum_{k'} \beta_{t,k,k'}\hat{\phi}_{t-1,k'}) \quad (7)$$

ここで $\beta_{t,k,k'}$ は，時刻 t におけるトピック k の，一時刻前のトピック k' への依存度を表し，これが大きいほど依存度が高いことを示している．

3.3 MdTDPM の学習

MdTDPM の学習には，確率的 EM アルゴリズムを用いる [13], [15], [16]．トピックの依存度 $\beta_{t,k,k'}$ および単語分布の推定値 $\hat{\phi}_{t,k'}$ は，確率的 EM アルゴリズムを用いるこ



(a) cDPM (b) TDPM (c) MdTDPM

図 3 トピックの時間発展の比較

とで逐次推定することができる．確率的 EM アルゴリズムでは，ギブスサンプリングによる潜在トピックの推定 [8] と，尤度最大化によるトピックの依存度の推定 [7] を交互に繰り返す．

ギブスサンプリングにおいて，時刻 t における i 番目の文書 $d_{t,i}$ に割り当てるトピック $z_{t,i}$ のサンプリングは，以下の式に従って行う．

$$P(z_{t,i}|\boldsymbol{z}_{t\backslash i}, \boldsymbol{z}_{t-1}, \boldsymbol{W}_t, \hat{\boldsymbol{\Phi}}_{t-1}, \boldsymbol{\beta}_t) \quad (8)$$
$$\propto P(z_{t,i}=k|\boldsymbol{z}_{t\backslash i})P(\boldsymbol{w}_{t,i}|z_{t,i}=k, \boldsymbol{W}_{t\backslash i}, \hat{\boldsymbol{\Phi}}_{t-1}, \boldsymbol{\beta}_t)$$

ここで，K_{t-1} を時刻 $t-1$ におけるトピックの数として，$\hat{\boldsymbol{\Phi}}_{t-1} = \left\{\hat{\phi}_{t-1,k}\right\}_{k=1}^{K_{t-1}}$，$\boldsymbol{\beta}_t = \{\beta_{t,k}\}_{k=1}^{K_{t-1}}$，$\boldsymbol{W}_t = \{\boldsymbol{w}_{t,i}\}_{i=1}^{M_t}$ である．また，$\backslash i$ は i 番目の文書を除くことを表している．依存度 $\boldsymbol{\beta}_t$ は，不動点反復法を用いて完全尤度 $P(\boldsymbol{W}_t, \boldsymbol{z}_t|...)$ を最大化することによって推定する．

4. 関連研究

提案モデルや TDPM と同様に，DPM を時系列に拡張したモデルとして，Evolutionary Net-Cluster Generative Model (Evo-NetClus)[11] と continuous Dirichlet Process Mixture (cDPM)[10] がある．Evo-NetClus はネットワークの時間発展を解析するためのモデルであり，時系列文書には直接適用できない．一方 cDPM は，Evo-NetClus をベースにして Twitter におけるトピックの時間発展を追跡できるようにしたモデルであり，Twitter 以外のニュース記事等の時系列文書にも適用可能である．cDPM では，時刻 t におけるトピック k の単語分布 $\phi_{t,k}$ は，一時刻前のある一つの単語分布 $\phi_{t-1,k'}$ に依存するか，あるいは時間に依存せずに β をハイパーパラメータとするディリクレ分布から生成されることを仮定している．一時刻前のトピックに依存する場合，その依存度は語彙数を V として $V\beta$ で与えられる．すなわち，$\phi_{t,k}$ は以下のディリクレ分布に従って生成される．

$$\phi_{t,k} \sim \text{Dirichlet}(V\beta\hat{\phi}_{t-1,k'}) \quad (9)$$

図 3 に，cDPM，TDPM，MdTDPM それぞれにおけるトピックの発展のイメージを示す．cDPM は，トピックの発生/消滅/分離を捉えることができるものの，依存度は一定である．一方 TDPM は，トピックの分離を捉えることはできないが，依存度は学習によって自動的に推定される．これらに対して提案モデルである MdTDPM は，一時刻前の複数のトピックへの依存を考慮しているため，トピックの発生/消滅/分離に加えて結合も考慮することができる．また，それぞれのトピックへの依存度は TDPM 同様，学習によって自動的に推定される．なお MdTDPM において，トピックの消滅は，そのトピックに対する次の時刻のトピックの依存度がすべて低い場合に，結合については，あるトピックが一時刻前の複数のトピックと依存度が高い場合に，分離については，あるトピックに対して次の時刻の複数のトピックの依存度が高い場合にそれぞれ起きたとみなすことができる．

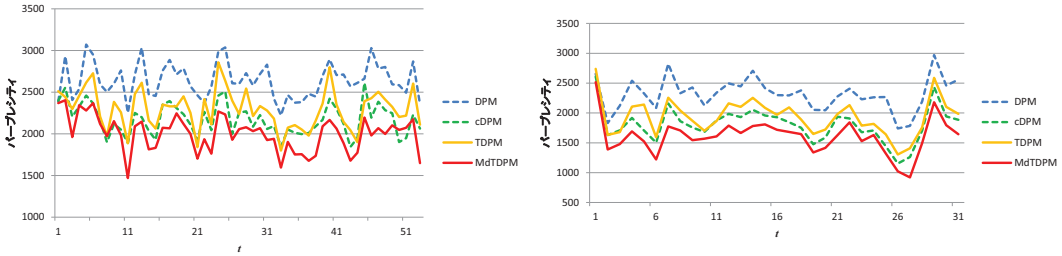
表 1 データ全体におけるパープレシティの平均値 (() 内は標準偏差)

データセット	DPM	cDPM	TDPM	MdTDPM
読売新聞	2636(30.5)	2194(33.0)	2297.8(51.6)	**2001.6**(39.0)
毎日新聞	2318.8(29.6)	1805.7(22.7)	1937.4(42.4)	**1608.4**(17.5)

5. 実験

実際のニュース記事を対象として，提案モデルに対する評価実験を行った．本実験では，読売新聞と毎日新聞の二つのニュース記事データセットを用いた．読売新聞はニュースサイト「YOMIURI ONLINE（読売新聞）」[*1] における 2014 年 4 月 25 日から 2014 年 6 月 16 日までの 4,373 件，毎日新聞はニュースサイト「毎日新聞のニュース・情報サイト」[*2] における 2014 年 7 月 15 日から 8 月 14 日までの 2,880 件のニュース記事を用いた．前処理として，各データセット中のニュース記事を形態素解析して名詞だけを抽出し，さらに出現回数が 5 回未満の単語と stop words を取り除いた．前処理の結果，読売新聞データでは語彙の種類数が 6,169，総単語数が 114,863，毎日新聞では語彙の種類数が 7,167，総単語数が 155,760 となった．

[*1] http://www.yomiuri.co.jp/
[*2] http://mainichi.jp/



(a) 読売新聞　　(b) 毎日新聞

図 4　各時刻のパープレキシティの平均値

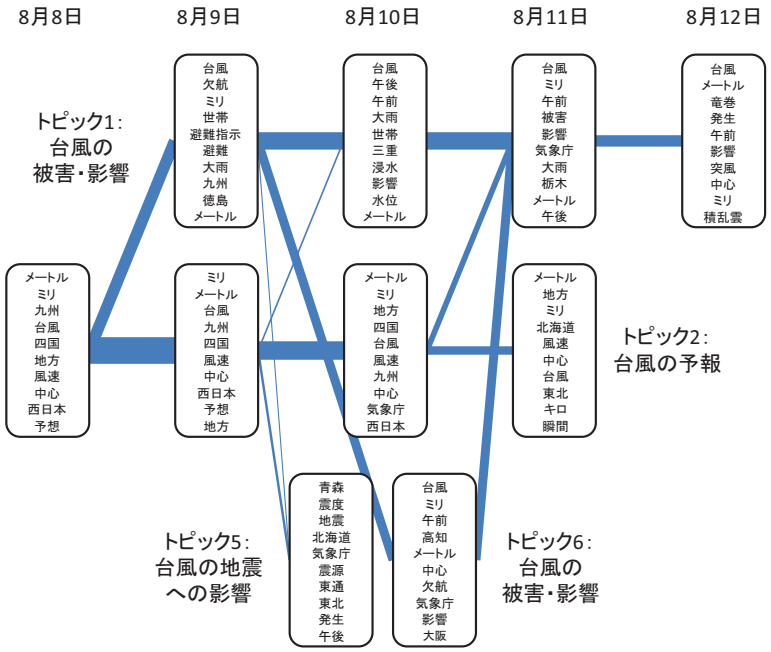


図 5　MdTDPM による台風 11 号に関するトピックの追跡．各時刻の各トピックにおいて確率の高い上位 10 語を示している．また，リンクは各トピックの依存関係を表しており，リンクが太いほど一方のトピックが一時刻前のもう一方のトピックに強く依存していることを示している．

パープレキシティを用いて，提案モデルの性能を従来モデルと比較評価した．パープレキシティは，言語モデルの評価において一般的に用いられる指標であり，学習によって得られたモデルが，テストデータに含まれる単語 $\boldsymbol{W}_{test}$ をどれだけ予測できるかを表す．

$$perplexity = \exp\Big(-\frac{\log P(\boldsymbol{W}_{test})}{N_{test}}\Big) \qquad (10)$$

式 (10) において，N_{test} はテストデータに含まれる単語の総数である．パープレキシティが低いほど，モデルの予測性能が高いことを示している．比較する従来手法としては，DPM，cDPM，TDPM の三つを用いた．ただし DPM は TDPM と同様に，3.1 節に示す手順により言語モデルに拡張したものを用いた．すべてのモデルに共通するハイパーパラメータ γ は 1，β は 0.1 とした．TDPM における依存度 $\beta_{t,k}$ の初期値は 100 とし，MdTDPM における依存度 $\beta_{t,k,k'}$ の初期値は，$k = k'$ のとき 100，それ以外では β とした．時間の単位は 1 日とし，各時刻における文書中の単語をランダムに 9:1 に分割し，90% を学習に用い，残り 10% をテストデータとしてパープレキシティの算出に用いた．また，全時刻でのパープレシティの平均値を算出し，データ全体のパープレキシティとした．これを 10 回試行したときのデータ全体のパープレシティの平均値を表 1 に，各時刻のパープレシティの平均値を図 4 に示す．提案モデルの計算時間は，1 試行あたり平均で 2321.5s（毎日新聞データ），2722.6s（読売新聞データ）であった．

表 1 より，いずれのデータセットにおいても，MdTDPM の性能が従来モデルを上回っていることがわかる．このことから，提案モデルがニュース記事中のトピックの時間発展をより適切にモデル化できているといえる．DPM の性

能が他のすべてのモデルと比較して劣るのは，時間発展を考慮していないためであり．また，cDPM が TDPM よりも性能がよいのは，トピックの分離を考慮しているためであると考えられる．一方で，MdTDPM が cDPM，TDPM の両方と比較して性能が改善しているのは，複数のトピックの依存関係を考慮することで，時間変化に伴うトピックの発展を柔軟に追跡できたためだと考えられる．また，図 4 より，MdTDPM は TDPM と比較して，ほぼすべての時刻においてパープレキシティが低くなっていることがわかる．MdTDPM と TDPM の違いは，トピックが独立に発展するか，互いに依存し合いながら発展するかであるため，このことからも複数のトピックへの依存を考慮することが有用であるといえる．

図 5 に，MdTDPM による平成 26 年台風第 11 号 [*3] に関するトピックの追跡例を示す．用いたデータは毎日新聞データであり，上述のパラメータで MdTDPM を適用した．台風 11 号は 2014 年 7 月 29 日に発生した台風であり，8 月 2 日頃から急速に勢力を強め，7 日午後に沖縄県の南大東島に最も接近した．この頃からニュース記事においても台風に関する情報が書かれ始め，それに伴って図 5 に示すトピック 2 が生成された．8 月 9 日には台風からの雨雲によって日本の広い範囲で大雨となり，トピック 2 から分離したトピック 1 がそれに関する話題を捉えられている．さらに，8 月 10 日には台風が高知県に上陸し，同日午後に日本海へ抜けて北上した．この日にはさらにトピックが分離し，大きな被害を受けた高知や大阪に関するトピック 6 が生成された．また，この日発生した青森を震源地とする地震に関する記事においても，地震に加えて台風の雨の影響による地盤の緩みに関して言及されており，この地震に関するトピック 5 が台風に関するトピックに弱く依存していることが捉えられていた．翌日には台風が過ぎ去ったこともあり，トピック 6 は再びトピック 1 と結合した．8 月 12 日には，前日のうちに台風が温帯低気圧に変わったために，トピック 2 は消滅したが，依然として台風の被害に関する記事は多く，トピック 1 はしばらく残り続けていた．このように，提案モデルにより，分離や結合，他のトピックへの影響なども含めたトピック追跡が可能である．

6. おわりに

本稿では，あるトピックが一時刻前の複数のトピックに依存すると仮定した時系列トピックモデルを提案し，その学習方法を示した．実際のニュース記事を用いた実験により，提案モデルが従来モデルよりも適切にトピックの発展をモデル化できることを示した．台風の話題を例に，提案モデルが発生/消滅/結合/分離に加えて，他のトピックへの影響まで考慮したトピック追跡が可能であることを示した．

今後の課題としては，Hierarchical Dirichlet Processes (HDP) [12] を用いて一つの文書の背景に複数のトピックを仮定できるようにすることが挙げられる．また，本稿で示した実験でのデータサイズであれば大きな問題とはならないが，提案モデルでは複数のトピックへの依存度を推定しているために，従来モデルよりも計算時間がかかってしまうという問題点がある．これに対しては，学習の途中で依存度の低いトピックへのリンクを断ち切る方法などを検討していく予定である．

[*3] http://ja.wikipedia.org/wiki/平成 26 年台風第 11 号

参考文献

[1] Ahmed, A. and Xing, E. P.: Dynamic Non-Parametric Mixture Models and the Recurrent Chinese Restaurant Process: with Applications to Evolutionary Clustering, *Proc. SDM'08* (2008).
[2] Ahmed, A. and Xing, E. P.: Timeline: A dynamic hierarchical Dirichlet process model for recovering birth/death and evolution of topics in text stream, *Proc. UAI'10* (2010).
[3] Aldous, D.: Exchangeability and Related Topics, *In École d'été de probabilités de Saint-Flour,XIII*, Berlin, Springer, pp. 1–198 (1985).
[4] Blei, D. M. and Lafferty, J. D.: Dynamic topic models, *Proc. ICML'06*, ACM, pp. 113–120 (2006).
[5] Blei, D. M., Ng, A. Y. and Jordan, M. I.: Latent dirichlet allocation, *the Journal of machine Learning research*, Vol. 3, pp. 993–1022 (2003).
[6] Ferguson, T. S.: A Bayesian analysis of some nonparametric problems, *The annals of statistics*, Vol. 1, No. 2, pp. 209–230 (1973).
[7] Minka, T.: Estimating a Dirichlet distribution, Technical report, MIT (2000).
[8] Neal, R. M.: Markov chain sampling methods for Dirichlet process mixture models, *Journal of computational and graphical statistics*, Vol. 9, No. 2, pp. 249–265 (2000).
[9] Sethuraman, J.: A constructive definition of Dirichlet priors, *Statistica Sinica*, Vol. 4, pp. 639–650 (1994).
[10] Si, J., Mukherjee, A., Liu, B., Li, Q., Li, H. and Deng, X.: Exploiting Topic based Twitter Sentiment for Stock Prediction, *Proc. ACL'13*, pp. 24–29 (2013).
[11] Sun, Y., Tang, J., Han, J., Gupta, M. and Zhao, B.: Community evolution detection in dynamic heterogeneous information networks, *Proc. MLG'10*, ACM, pp. 137–146 (2010).
[12] Teh, Y. W., Jordan, M. I., Blei, M. J. and Blei, D. M.: Hierarchical dirichlet processes, *Journal of the american statistical association*, Vol. 101, No. 476 (2006).
[13] Wallach, H. M.: Topic modeling: beyond bag-of-words, *Proc. ICML'06*, ACM, pp. 977–984 (2006).
[14] Wei, X., Sun, J. and Wang, X.: Dynamic Mixture Models for Multiple Time-Series, *Proc. IJCAI'07*, pp. 2909–2914 (2007).
[15] 岩田具治，山田武士，櫻井保志，上田修功：オンライン学習可能な多重スケールでの時間発展を考慮したトピックモデル，情報論的学習理論テクニカルレポート 2009 (2009).
[16] 岩田具治，渡部晋治，山田武士，上田修功：購買行動解析のためのトピック追跡モデル，電子情報通信学会論文誌，Vol. 93, No. 6, pp. 978–987 (2010).