Tracking the Evolution of Social Emotions: A Time-Aware Topic Modeling Perspective

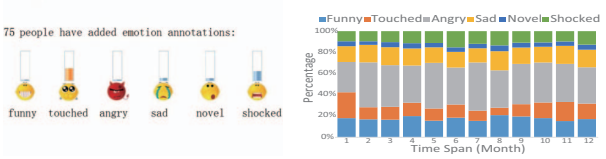
Chen Zhu[1], Hengshu Zhu[1*], Yong Ge[2], Enhong Chen[1], Qi Liu[1]

[1]University of Science and Technology of China, [2]University of North Carolina at Charlotte
[1]{zhuzhen, zhs}@mail.ustc.edu.cn, [1]{cheneh, qiliuql}@ustc.edu.cn, [2]yong.ge@uncc.edu

***Abstract*—Many of today's online news websites have enabled users to specify different types of emotions (e.g., angry and shocked) they have after reading news. Compared with traditional user feedbacks such as comments and ratings, these specific emotion annotations are more accurate for expressing users' personal emotions. In this paper, we propose to exploit these users' emotion annotations for online news in order to track the evolution of emotions, which plays an important role in various online services. A critical challenge is how to model emotions with respect to time spans. To this end, we propose a time-aware topic modeling perspective for solving this problem. Specifically, we first develop a model named *emotion-Topic over Time* (eToT), in which we represent the topics of news as a Beta distribution over time and a multinomial distribution over emotions. While eToT can uncover the latent relationship among news, emotion and time directly, it cannot capture the dynamics of topics. Therefore, we further develop another model named *emotion-based Dynamic Topic Model* (eDTM), where we explore the state space model for tracking the dynamics of topics. In addition, we demonstrate that both eToT and eDTM could enable several potential applications, such as emotion prediction, emotion-based news recommendations and emotion anomaly detections. Finally, we validate the proposed models with extensive experiments with a real-world data set.**

I. Introduction

With the increasing prosperity of Web 2.0, people are encouraged to have various social interactions on the websites. A recent development trend of online news websites, such as Yahoo! and Sina, is to allow readers to specify different types of emotions (e.g., angry and shocked) after reading news. Compared with traditional users' feedbacks (e.g., reviews, tags), such specific emotion annotations are more accurate for expressing users' personal emotions. For example, Figure 1(a) shows an example of users' aggregated emotions at Sina News [1], which has six different kinds of emotions. Each user can choose one emotion, which most accurately reflects his impression after reading, to annotate a piece of news. If we collectively look at and analyze all user emotions from the news website, we may be able to get a good picture about the overall emotions of online social media users, namely *social emotions*. The social emotions often vary with respect to the topics of news and times, and thus have the intrinsic dynamics. For example, Figure 1(b) shows the distribution of aggregated social emotions with respect to different time spans in our data set collected from Sina News. We can observe different distributions of emotions along the time, which indicates that the social emotions evolve over time. In fact, such evolution of social emotions is inherently driven by the dynamic topics of news at different times. Capturing such dynamic characteristics of social emotions is critically important for the successful development of various social services, such as social opinion monitoring and social event detections. In the literature, there are recent studies about social emotion related problems. For example, some works focus on sentiment analysis [24], [30], social emotion analysis [6], [13], [18] of online documents and user emotion modeling [3]. However, few of them have paid attention to the dynamic characteristics of social emotions.



(a) (b)

Fig. 1. (a) Examples of different user emotions. (b) The distribution of social emotions with respect to news in different time spans.

To this end, in this paper we propose to exploit the users' emotion annotations from online news to track the evolution of social emotions. A critical challenge is how to model emotions with respect to time spans. Along this line, we propose a time-aware topic modeling perspective for solving this problem. Specifically, we first develop a model named *emotion-Topic Over Time* (eToT), where we represent each topic of news with a Beta distribution over time and a multinomial distribution over emotions. In this model, the process of modeling topics of news is affected not only by the word co-occurrences but also the emotions and time. Similar as the popular Latent Dirichlet Allocation (LDA) [5], news topics in eToT are static which do not change with respect to different time spans. However, while some researchers have revealed that the latent topics of documents may evolve as time unfolds [4], eToT cannot capture the dynamics of topics. Therefore, we further propose another model, namely *emotion-based Dynamic Topic Model* (eDTM), to capture the dynamics of news topics. In eDTM, we first divide all news into different segmentations with respect to their timestamps, and then implement topic modeling for each of segmentations. Models learned from different segmentations are linked together by the Markov state space model. Indeed, both eToT and eDTM can model the evolution of social emotion effectively and help us understand the semantic relationships between social emotion and news topics in different views. In addition, we demonstrate that both eToT and eDTM are general models and could enable several potential applications of social media, such as emotion prediction, emotion-based news recommendations and emotion anomaly detections. Specifically, the contributions of this paper can be summarized as follows.

*Corresponding Author
[1]http://news.sina.com.cn/

1550-4786/14 $31.00 © 2014 IEEE
DOI 10.1109/ICDM.2014.121

• First, we provide a comprehensive study for tracking the evolution of social emotions by exploiting the user emotion annotations from online news. Indeed, this study is critically vital for the successful development of various social services.
• Second, we propose two novel topic models, i.e., eToT and eDTM, for solving the problem of tracking the evolution of social emotion. Particularly, the proposed models can effectively model the emotions, news and time from different views. Also, we introduce some novel emotion-based applications enabled by the proposed models.
• Third, we carry out extensive experiments with a real-world data set which was collected from Sina News to evaluate the proposed models. As shown in our experiments, the proposed models are both effective for modeling social emotions and other enabled applications.

The remainder of this paper is organized as follows. In section 2, we briefly introduce some related works of this paper. Section 3 introduces the details of our two novel topic models. In section 4 we make some discussion about our models and introduce their potential applications. Section 5 reports the experimental results. Finally, we conclude this paper in section 6.

II. Related Work

The related works can be grouped into two categories: sentiment analysis and topic model. We will give brief discussions on these related works and the relationships between them and the proposed models in this paper.

A. Sentiment Analysis

In the field of sentiment analysis, existing researches mainly focus on two aspects, namely sentiment classification and sentiment information extraction. Among them, since sentiment extraction aims to extract units that are related to emotion at sentence level or paragraph level, it is less relevant to our scenario. Therefore, in the following, we only discuss the works on sentiment classification [13], [16], [18], [25], [27], [31].

Classification is a very fundamental problem in machine learning [25], [33], [34]. Traditionally, sentiment classification is often formalized as a classification task, and thus could be addressed by the classification approaches directly. Recently, researchers have proposed a lot of new approaches for sentiment classification and begun to study the readers' emotions. For instance, in [17] two ranking methods were presented to rank readers' emotions. Lin *et al.* [18] also studied the emotions by which the readers of these articles are triggered, and the authors mainly focus on feature selection. Kozareva *et al.* [13] designed an approach for classifying headline emotion based on the information collected from the World Wide Web. Besides, emoticon, as a kind of emotion label, has also been used for sentiment classification. For example, Zhao *et al.* [32] built a system called MoodLens, in which 95 emoticons are mapped into four categories of sentiments for the sentiment classification of Chinese tweets in Weibo. Liu *et al.* [20] treated the emoticons as labeled data and integrated it and labeled data into a same sentiment classification framework.

We can find that most of existing works aim to study the sentiment directly from the documents (words). Since there is a consensus that a document is composed of topics, it is also reasonable to treat sentiment about a document as a combination of sentiment on these topics. Therefore, in this paper we focus on analyzing sentiment from the perspective of the sentiment (emotion) of the topics. There have been some works about the emotion with respect to topics. For instance, Mei *et al.* [22] modeled texts through a mixture of topic model and sentiment model. Titov and McDonald [27] proposed a Multi-Aspect Sentiment model consisting of an extension of LDA, Multi-Grain LDA, and sentiment predictors for sentiment summarization. Similarly, Lin and He [16] also extended LDA to detect sentiment and topics. However, none of them considers the temporal effect which is very useful when analyzing the relationship between topics and emotions.

B. Topic Models

Topic model is a powerful tool for analyzing texts or other types of discrete data. Many different kinds of topic models [4], [5], [21], [29] have been proposed, among which Latent Dirichlet Allocation (LDA) [5] is most popular. Actually, the models proposed in this paper can be also traced back to LDA. The basic assumption of LDA is that each document can be treated as a mixture of topics and the generation of words in the document is attributable to one of those topics.

However, LDA is a static model that does not consider the influence of time, so it is inappropriate to be adopted in many applications, including social emotion analysis. To this end, several topic models have examined topics and their changes with respect to time variation. For instance, Topics over Time (ToT) [29] model associates each topic with a Beta distribution over time to directly capture the topic changes over time. The relative simplicity of ToT makes it easy to inject its ideas to other topic models. Different from ToT, Dynamic Topic Model (DTM) [4] uses state space model on the natural parameters to model the evolution of topic and the variational approximations depending on Kalman filters and wavelet regression. DTM requires that time must be discretized, and thus how to determine the length of time span is an important problem. Wang *et al.* [28] solved this problem by replacing state space model with Brownian motion which is actually acontinuous generalization of state space model. Another way to solve this problem was proposed in [11], which considers multi-scale time span to track the topic changes. Topic Tracking Model (TTM) [9] is another dynamic topic model for tracking consumer purchase behaviors. Compared with DTM, it simplifies the model by replacing Gaussian distribution with Dirichlet distribution so that the parameters could be estimated by Gibbs sampling which is simpler compared with variational methods.

In summary, there are two main approaches to model the evolution of topics over time: treating time as an attribute of topics, e.g., ToT; using state space model or Brownian motion on natural parameters, e.g., DTM. However, none of existing works could effectively model the relationship among topics, emotions and time simultaneously.

III. Modeling Social Emotion

In this section, firstly, some preliminaries and backgrounds would be introduced. Then, we explain the details of our time-

aware topic modeling solutions, namely eToT and eDTM, for analyzing social emotion.

A. Preliminaries and Backgrounds

Recently, many of todays online news websites have enabled users to specify their emotions after reading news. Therefore, for each piece of news, we can get not only the textual content and timestamp but also the emotions that were annotated by users. For example, Figure 1(a) illustrates the emotion annotations from 75 users, and these emotions are classified into six categories, i.e., "*Funny*", "*Touched*", "*Angry*", "*Sad*", "*Novel*", and "*Shocked*". As shown in Figure 1(a), all emotion annotations of a piece of news from different users are usually aggregated together.

For social emotion modeling, intuitively, users' emotion about a document (e.g., news) is mainly decided by their emotion about the topics of the document, which could be extracted by topic model. Indeed, topic model is an unsupervised method to effectively extract topics from the texts based on the co-occurrence of words. Though some existing topic models aim to model the evolution of topics, none of them could help analyze social emotions with respect to temporal information, which this paper focuses on. Considering that different topics lead to different user emotions, e.g., some topics tend to depress people but some other topics may be very affecting, it is reasonable to associate each topic with a distribution with respect to emotion. In this way, we can model the evolution of social emotions by introducing an emotion and time generation layer into the topic models, in which the discovery of topics is influenced by the social emotions, the co-occurring of words, and also the temporal information.

Formally, we assume that a document d is a bag of words $\mathbf{w}_d$, size of which is N_d, and $\mathbf{w}_d$ is from a vocabulary containing V words. Each document also has a timestamp t and an observed multinomial distribution over emotions $\mathbf{e}$. Thus, we can use a set of triples $\mathbf{D} = \{(\mathbf{w}_0, t_0, \mathbf{e}_0), ..., (\mathbf{w}_D, t_D, \mathbf{e}_D)\}$ to represent a collection of D documents. Next, we will explain our solutions to model the evolution of social emotion from the corpus $\mathbf{D}$.

B. emotion-Topic Over Time

In this subsection, we would propose a novel topic model named emotion-Topic Over Time (eToT) to directly uncover the latent relationship among documents, emotion, and time. Compared with LDA, each topic discovered by eToT also has a distribution with respect to time and a distribution with respect to emotion. That is, besides understanding the content of a topic, we can know the people's impression of this topic (i.e., social emotion) and when this topic emerges.

We first introduce the way to connect social emotions with topics. It has been introduced briefly above that we assume the users' emotions about a specific document come from their emotions about the topics of this document. Thus, each topic also has a latent distribution with respect to emotion in eToT. Note that, the emotions are discrete and the number of each kind of emotions could be observed for any documents (news). That is, given a news article, we could easily get the distribution over emotions $\mathbf{e}$ with respect to this news. Here, $\mathbf{e}$ is observed, while the emotion distributions on topics η, which determine $\mathbf{e}$, are latent. Specifically, we assume this latent distribution follows the Dirichlet distribution, i.e., each topic k has a Dirichlet distribution η_k with respect to emotion. Dirichlet distribution is a family of continuous multivariate probability distributions over simplex of which the summation is equal to one. Thus, it is the best choice to model the distribution of emotions. Given a topic k, the probability of a observed emotion distribution $\mathbf{e}$ could be calculated as follows:

$$P(\mathbf{e}|\eta_k) = \frac{\Gamma(\sum_{l=1}^{E} \eta_{k,l})}{\prod_{l=1}^{E} \Gamma(\eta_{k,l})} \cdot \prod_{l=1}^{E} e_l^{\eta_{k,l}-1}, \quad (1)$$

where the inference for parameter η_k will be shown later, and $\mathbf{e}$ is the vector of proportion of emotions. E is the number of categories of emotions and is also the size of vector η_k. $\Gamma(\cdot)$ is gammma function.

Similarly, we also directly associate time with topics, i.e., each topic k has a latent distribution ψ_k with respect to time. Here, because Beta distribution could behave versatile shapes, we select it as ψ_k. Therefore, we need to normalize the timestamp into a range from 0 to 1 firstly. Then, given a topic k, the probability of a observed timestamp t could be calculated as follows:

$$P(t|\psi_k) = \frac{\Gamma(\psi_{k,1}+\psi_{k,2})}{\Gamma(\psi_{k,1})\Gamma(\psi_{k,2})} \cdot (1-t)^{\psi_{k,1}-1} t^{\psi_{k,2}-1}, \quad (2)$$

where the inference for parameter ψ_k will also be shown later.

Note that, for simplifying the process of parameter estimation, we generate emotions and timestamp with respect to each word token. That is, we assume all of word tokens in one document share the same emotion distribution $\mathbf{e}$ and the same timestamp t with the document. In summary, the corresponding graphical model of eToT is shown in Figure 2, and its parameterizations are

$$\begin{aligned}
\theta_d|\alpha &\sim \text{Dirichlet}(\alpha),\\
\phi_k|\beta &\sim \text{Dirichlet}(\beta),\\
z_{d,i}|\theta_d &\sim \text{Multinomial}(\theta_d),\\
w_{d,i}|\phi_{z_{d,i}} &\sim \text{Multinomial}(\phi_{z_{d,i}}),\\
t_{d,i}|\psi_{z_{d,i}} &\sim \text{Beta}(\psi_{z_{d,i}}),\\
\mathbf{e}_{d,i}|\eta_{z_{d,i}} &\sim \text{Dirichlet}(\eta_{z_{d,i}}).
\end{aligned}$$

Thus, the generative process for a document d is as follows:

1) Draw a multinomial distribution θ_d over topics from a Dirichlet prior α;
2) For a word token i in document d:
   a) Draw a topic $z_{d,i}$ from Mult(θ_d);
   b) Draw a word $w_{d,i}$ from Mult($\phi_{z_{d,i}}$);
   c) Draw a timestamp $t_{d,i}$ from Beta($\psi_{z_{d,i}}$);
   d) Draw a multinomial distribution $e_{d,i}$ over emotions from Dirichlet($\eta_{z_{d,i}}$);

We employ Gibbs sampling to estimate parameters, i.e., to "invert" the generative process and "generate" latent variables from given observations. For simplicity, we estimate ψ and η once per iteration of Gibbs sampling. The joint distribution of topic z, time t, word w, and emotion e is as follows:

$$\begin{aligned}
P(\mathbf{w}, \mathbf{t}, \mathbf{z}, \mathbf{e}|\alpha, \beta, \eta, \psi) = \prod_{z=1}^{K} \frac{\Delta(\mathbf{n}_z+\beta)}{\Delta(\beta)} \prod_{d=1}^{D} \frac{\Delta(\mathbf{m}_d+\alpha)}{\Delta(\alpha)} \cdot \\
\prod_{d=1}^{D} \prod_{i=1}^{N_d} (P(t_{d,i}|\psi_{z_{d,i}}) P(\mathbf{e}_{d,i}|\eta_{z_{d,i}})).
\end{aligned} \quad (3)$$

In Gibbs Sampling procedure, what we need to calculate is the conditional distribution $P(z_{d,i} =$

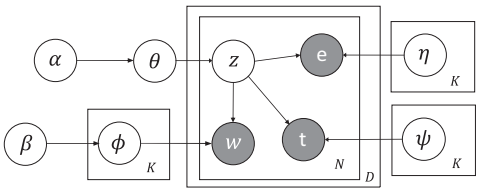


Fig. 2. The graphical representation of eToT.

$k|\mathbf{w}, \mathbf{t}, \mathbf{z}_{\neg d,i}, \mathbf{e}, \alpha, \beta, \Psi, \eta)$ where $\mathbf{z}_{\neg d,i}$ means the topics for all of the word tokens except $w_{d,i}$. Because of conjugation, the conditional distribution in each iteration can be written as:

$$P(z_{d,i} = k|\mathbf{w}, \mathbf{t}, \mathbf{z}_{\neg d,i}, \mathbf{e}, \alpha, \beta, \Psi, \eta) \propto \frac{m_{d,k,\neg_{d,i}} + \alpha_k}{(\sum_{z=1}^{K} m_{d,z,\neg_{d,i}} + \alpha_z) - 1} \cdot \frac{n_{k,w_{d,i},\neg_{d,i}} + \beta_{w_{d,i}}}{(\sum_{v=1}^{V} n_{k,v,\neg_{d,i}} + \beta_v) - 1} \cdot \frac{(1-t_{d,i})^{\psi_{k,1}-1}(t_{d,i})^{\psi_{k,2}-1}}{B(\psi_{k,1}, \psi_{k,2})} \cdot \left(\frac{\Gamma(\sum_{l=1}^{E} \eta_{k,l})}{\prod_{l=1}^{E} \Gamma(\eta_{k,l})} \cdot \prod_{l=1}^{E} e_{d,i,l}^{\eta_{k,l}-1}\right), \tag{4}$$

$m_{d,k}$ is the number of word tokens in document d that are assigned to topic k, $n_{k,v}$ is the number of word v assigned to topic k, and $e_{d,i,l}$ is the l-th emotion proportion of the i-th word token in document d. After each iteration of Gibbs sampling, we update ψ and η as

$$\hat{\psi_{k,1}} = \bar{t_k}\left(\frac{\bar{t_k}(1-\bar{t_k})}{S^{(t)}{}_k^2} - 1\right), \tag{5}$$

$$\hat{\psi_{k,2}} = (1-\bar{t_k})\left(\frac{\bar{t_k}(1-\bar{t_k})}{S^{(t)}{}_k^2} - 1\right), \tag{6}$$

$$\hat{\eta_{k,l}} = \bar{e_{k,l}}\left(\frac{\bar{e_{k,1}}(1-\bar{e_{k,1}})}{S^{(e)}{}_{k,1}^2} - 1\right), \tag{7}$$

where $\bar{t_k}$ and $S^{(t)}{}_k^2$ are the sample mean and the biased sample variance of the timestamps of word tokens belonging to topic k, respectively. Similarly, $\bar{e_{k,l}}$ and $S^{(e)}{}_{k,l}^2$ are the sample mean and the biased sample variance of the i-th emotion of word tokens belonging to topic k. After the iterations, $\theta_{d,k}$ and $\phi_{k,i}$ can be evaluated as follows:

$$\hat{\theta_{d,k}} = \frac{m_{d,k} + \alpha_k}{\sum_{z=1}^{K}(m_{d,z} + \alpha_z)}, \tag{8}$$

$$\hat{\phi_{k,i}} = \frac{n_{k,i} + \beta_i}{\sum_{v=1}^{V}(n_{k,v} + \beta_v)}. \tag{9}$$

By eToT, we could extract topics and some additional information about them from a collection of texts. Though the topics are constant, we get their distributions with respect to time ψ and discrete distributions with respect to emotions η. By these distributions, we can understand the relationships among emotions, topics, and time. And the temporal information could help us to discover those emergencies effectively [29], which will also be shown in experiments well. At last, note that, the generation process of our eToT model is similar to the ToT model [29]. Indeed, in eToT, topic discovery is affected by all of the effects from word co-occurrences, emotion, and temporal information. while ToT only considers the influence from words and time. So ToT could not be used for sentiment analysis and modeling social emotions.

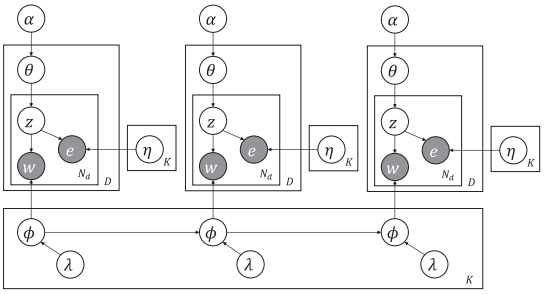


Fig. 3. The graphical representation of eDTM.

C. emotion-based Dynamic Topic Model

Although eToT can uncover the latent relationship among documents, emotion and time, the constant topics learned by eToT could not capture the dynamics of news topics. To this end, in this subsection we propose another model, emotion-based Dynamic Topic Model (eDTM), to solve this problem.

Different from eToT, eDTM discovers topics by the content and emotion information in each time span, and the topics in different time spans are chained by state space model. In this way, eDTM could learn the distribution over words and the distribution with respect to emotion for each topic in different time spans. Thus, the evolution of news topics is demonstrated clearly and directly by eDTM. Here, we assume the distribution over words of a topic should evolve smoothly, while emotion may not. Actually, the study on social science has shown that the popular mind is weird [1], [14]. Sometimes, it is bigoted, conservative, and imperious. At other times, it shows some totally different characteristics such as impulse, fickleness, and irritability. Meanwhile, the detection of abnormal change of social emotion is one of applications could be derived from our models. Thus, we just chain topics depending on their distributions over words. Due to the state space model, we need firstly divide texts into different time spans by timestamps. Then eDTM models the documents in time span t with a topic model which has K topics, where these topics evolve from the topics associated with the time span $t-1$. For simplicity, we use Dirichlet distribution to chain the neighboring topic models on the parameters of the multinomial distributions of topics over word tokens ϕ. We could also employ Gibbs sampling to evaluate these parameters. Specifically, the sequential structure is

$$P(\phi_{t,k}|\phi_{t-1,k}, \lambda_{t,k}) \propto \prod_{v=1}^{V} \phi_{t,k,v}^{\lambda_{t,k}\phi_{t-1,k,v}-1}, \tag{10}$$

where λ is a parameter to control the influence of the topic model in time span $t-1$ on the topic model in time span t.

Since the state space model requires the discretization of time, we need to modify some notations in this subsection. In time span t, the document collection is $\mathbf{D}_t$ and the size of it is D_t. Given a document d, the set of word tokens is $\{w_{t,d,i}\}$ and $|\{w_{t,d,i}\}|$ is $N_{t,d}$. We represent the emotion of document d (word tokens in d) by an observed multinomial distribution $\mathbf{e}_{t,d}$. $\theta_{t,d}$ is the topic distributions in document d. $\phi_{t,k}$ and $\eta_{t,k}$ are the parameters for Dirichlet distributions with respect to words and emotions for the topic k, respectively. E is the

number of emotions. Finally, the graphical model of eDTM is shown in Figure 3, and the eDTM parameterizations are

$$\begin{aligned}
\theta_{t,d}|\alpha_t &\sim \text{Dirichlet}(\alpha_t),\\
\phi_{t,k}|\phi_{t-1,k} &\sim \text{Dirichlet}(\lambda_{t,k}\cdot\phi_{t-1,k}),\\
z_{t,d,i}|\theta_{t,d} &\sim \text{Multinomial}(\theta_{t,d}),\\
w_{t,d,i}|\phi_{t,z_{t,d,i}} &\sim \text{Multinomial}(\phi_{t,z_{t,d,i}}),\\
\mathbf{e}_{t,d,i}|\eta_{t,z_{t,d,i}} &\sim \text{Dirichlet}(\eta_{t,z_{t,d,i}}).
\end{aligned}$$

The generative process for a document d in time span t is:

1) Draw a multinomial distribution $\theta_{t,d}$ over topics from a Dirichlet prior α_t;
2) Draw parameters of topic-word multinomial distributions $\phi_{t,k}$ from $Dirichlet(\lambda_{t,k}\cdot\phi_{t-1,k})$;
3) For a word token i in document d in time span t:
   a) Draw a topic $z_{t,d,i}$ from Mult($\theta_{t,d}$);
   b) Draw a word $w_{t,d,i}$ from Mult($\phi_{t,z_{t,d,i}}$);
   c) Draw a multinomial distribution $e_{t,d,i}$ over emotions from Dirichlet($\eta_{t,z_{t,d,i}}$);

We infer the topic assignments by Gibbs sampling. The joint distribution of topic z, word w and emotion e is as follows:

$$\begin{aligned}
&P(\mathbf{w}_t,\mathbf{z}_t,\mathbf{e}_t|\alpha_t,\phi_{t-1},\lambda_t,\eta_t)\\
&\quad= P(\mathbf{w}_t|\mathbf{z}_t,\phi)p(\mathbf{e}_t|\eta,\mathbf{z}_t)P(\mathbf{z}_t|\alpha_t)P(\phi_t|\phi_{t-1},\lambda_t)\\
&\quad= \prod_{z=1}^{K}\frac{\Delta(\mathbf{n}_{t,z}+\phi_{t-1,z}\cdot\lambda_{t,z})}{\Delta(\phi_{t-1,z}\cdot\lambda_{t,z})}\prod_{d=1}^{D_t}\frac{\Delta(\mathbf{m}_{t,d}+\alpha_t)}{\Delta(\alpha_t)}\\
&\qquad\cdot\prod_{d=1}^{D_t}\prod_{i=1}^{N_d}P(\mathbf{e}_{t,d,i}|\eta_{t,z_{t,d,i}}),
\end{aligned}\tag{11}$$

$$P(\mathbf{e}_{t,d,i}|\eta_{t,z_{t,d,i}}) = \frac{\Gamma(\sum_{l=1}^{E}\eta_{t,z_{t,d,i},l})}{\prod_{l=1}^{E}\Gamma(\eta_{t,z_{t,d,i},l})}\cdot\prod_{l=1}^{E}e_{t,d,i,l}^{\eta_{t,z_{t,d,i},l}-1},\tag{12}$$

where $\mathbf{n}_{t,k}$ is a V-dimensional vector, where each entry represents the number of each word assigned to topic k in time span t and $\mathbf{m}_{t,d}$ is a K-dimensional vector representing the number of word tokens assigned to each topic in text d of time span t. K is the number of topics.

During Gibbs sampling, we need to calculate the full conditional distribution of assigning word token i on each topic, e.g., topic k, and this is

$$\begin{aligned}
&P(z_{t,d,i}=k|\mathbf{w}_t,\mathbf{z}_{t,\neg_{d,i}},\mathbf{e}_t,\alpha_t,\phi_{t-1},\lambda_t,\eta_t) = P(\mathbf{e}_{t,d,i}|\eta_{t,k})\cdot\\
&\frac{n_{t,k,w_{t,d,i},\neg_{d,i}}+\phi_{t-1,k,w_{t,d,i}}\lambda_{t,k}}{\sum_{v=1}^{V}(n_{t,k,v,\neg_{d,i}})+\lambda_{t,k}}\frac{m_{t,d,k,\neg_{d,i}}+\alpha_{t,k}}{\sum_{z=1}^{K}(m_{t,d,z,\neg_{d,i}}+\alpha_{t,z})}.
\end{aligned}\tag{13}$$

Similar to that in eToT, we update η after each iteration of Gibbs sampling by the following equation:

$$\hat{\eta_{t,k,l}} = \bar{e_{t,k,l}}(\frac{\bar{e_{t,k,1}}(1-\bar{e_{t,k,1}})}{{S^{(e)}}^2_{t,k,1}}-1),\tag{14}$$

where $\bar{e_{t,k,l}}$ and ${S^{(e)}}^2_{t,k,l}$ are the sample mean and the biased sample variance of the l-th emotion of word tokens belonging to the topic k in time span t, respectively. After enough iterations, $\theta_{t,d,k}$ and $\phi_{t,k,i}$ can be evaluated as follows:

$$\hat{\theta_{t,d,k}} = \frac{m_{t,d,k}+\alpha_{t,k}}{\sum_{z=1}^{K}(m_{t,d,z}+\alpha_{t,z})},\tag{15}$$

$$\hat{\phi_{t,k,i}} = \frac{n_{t,k,i}+\lambda_{t,k}\phi_{t-1,k,i}}{\sum_{v=1}^{V}(n_{t,k,v}+\lambda_{t,k}\phi_{t-1,k,v})},\tag{16}$$

the parameter λ controls the process of evolution, and it could be fixed or estimated through maximizing the joint distribution by the following equation [10], [23]:

$$\lambda_{t,k}\leftarrow\lambda_{t,k}\cdot\frac{\sum_{v=1}^{V}\phi_{t-1,k,v}\cdot\Lambda_{t,k,v}}{\Psi(\sum_{v=1}^{V}n_{t,k,v}+\lambda_{t,k})-\Psi(\lambda_{t,k})},\tag{17}$$

$$\Lambda_{t,k,v} = \Psi(n_{t,k,v}+\lambda_{t,k}\cdot\phi_{t-1,k,v})-\Psi(\lambda_{t,k}\cdot\phi_{t-1,k,v}),\tag{18}$$

where $\Psi(\cdot)$ is actually the derivative of the logarithm of the gamma function.

Different from eToT model which outputs the constant topics, eDTM can get the distribution over words ϕ and distribution over emotions η for topics in any time span. Because of the state space model, for a given topic, eDTM could unveil the evolution of words and emotions directly. Note that, the traditional DTM [4] also uses state space model to measure the topic evolution over time. However, there are some significant differences between DTM and eDTM. The most important one is that the extraction of topics in DTM could not take account of the influence of social emotion into consideration. Thus, DTM cannot be directly adopted for sentiment analysis. In addition, DTM uses to Gaussian distribution to model the evolution, which means it has to use the variational method to estimate parameters. In eDTM, we choose the Dirichlet distribution as an alternative, which make a relatively easy way available.

IV. Discussion and Applications

In this section, we provide some discussion and potential applications about our novel models.

A. Model Discussion

Indeed, both eToT and eDTM can model the evolution of social emotions effectively, however, they uncover the semantic relationships between social emotion and news topics from different views. To be specific, from the experiments, we have observed that some words, which represent people's names or addresses in some social events, have high generation probabilities in the topics learned by eToT. For example, a top ranked word in one of the topics learned by eToT is the name of a doctor, who received unjustified treatment and attracted lots of public attentions. Therefore, the two emotions with highest generation probabilities in this topic are "*Angry*" and "*Sad*", which reflect people's sympathy for the doctor. It implies that eToT has the ability to find those emergencies and could reflect the popular incline to them. Differently, eDTM can discover the hidden semantics from another dynamic view. For example, we find the top ranked words in a topic learned by eDTM always include "Police", "Hospital" and "Investigation", which may indicate the semantics of crime. Therefore, at the beginning, the two emotions with highest generation probabilities in this topic are "*Angry*" and "*Shocked*". However, there are some interesting changes of the generation probability of emotion during the evolution of this topic. As time passes by, the emotion "*Sad*" became more and more important and finally exceeded "*Shocked*" in the topic, which may indicate people had changed their focus from the crime events to the victims.

B. Model Applications

Indeed, many applications can be derived from our novel models, i.e., eToT and eDTM. In the following, we demonstrate three motivating examples including the emotion prediction, the emotion-based news recommendation as well as the emotion anomaly detection.

Emotion Prediction. Emotion prediction is a classic problem in sentiment analysis, which has great application value in both industry and academia [3], [18]. Indeed, our models could be also leveraged for solving this problem. Specifically, given a document (e.g., a news) d, the goal is to predict the emotion e^* with the highest generation probability, that is,

$$e^* = \arg\max_e P(e|d). \tag{19}$$

The probability $P(e|d)$ can be computed by

$$P(e|d) \propto \sum_z \prod_{w \in d} P(w|z)P(e|z)P(z), \tag{20}$$

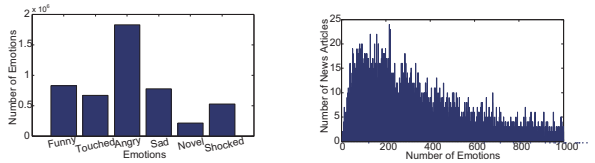
where $P(w|z)$, $P(e|z)$ and $P(z)$ can be learned during the training of models.

Emotion-based News Recommendation. The existing news recommender systems usually recommend news according to the content similarity between news and user preferences [7], [19], most of which neglect the impact of user emotions. However, users often have different emotion preferences during reading news. For example, some people may like news about funny stories that make them happy, while some people may like the shocked news for knowing the society. To this end, we first estimate the emotion preferences of a user u by calculating the probability $P(e|d_u)$, where d_u is the set of historical news read by u. Indeed, the probability $P(e|d_u)$ can be computed in the similar way of Equation 20. After that, the emotion preference $P(e|d_u)$ can be integrated into many state-of-the-art approaches [12], [15] for news recommendation. For example, we can compute the user similarity by Cosine similarity or KL divergence with the user preferences, and leverage user-based collaborative filtering approach to recommend news.

Emotion Anomaly Detection. Anomaly Detection is an important task and detecting the emotion anomaly has attracted lots of researchers' attentions [24], [26]. Although some existing works can find the static anomaly of social emotions, few of them could detect the anomalous changes of social emotions with respect to the social events. To this end, with the eDTM model, we could directly capture the changes of social emotion during different time spans. To be specific, by calculating the Cosine similarity or the KL divergence of emotion distributions, i.e., η in eDTM, between two neighboring time spans, we could measure the similarity of social emotions. After that, if the similarity is less (e.g., Cosine similarity) or larger (e.g., KL divergence) than a pre-defined threshold ϵ, we believe that there are some anomalies during the two time spans, e.g., some significant events happened.

V. Experimental Results

In this section, we evaluate our models, namely eToT and eDTM, with extensive experiments on a real-world data set.



(a) (b)

Fig. 4. The distribution of (a) the number of annotations on different emotions, and (b) the number of news articles with respect to different number of emotion annotations.

A. The Experimental Data

The experimental data were collected from the Society subsection of Sina News [2], which is one of the biggest online news websites in China, from August 21, 2012 to November 11, 2013. In this website, readers are allowed to choose one of the emotions, including "*Funny*", "*Touched*", "*Angry*", "*Sad*", "*Novel*", and "*Shocked*", to annotate the news after reading. In our data set, all the news articles were collected from the official top chart of each emotion every day, which guarantees there are enough emotion annotations for model training. Specifically, the data set contains 7,504 news articles with 4,844,594 emotion annotations. In particular, we have made this data set publicly available for research purpose [3]. To guarantee the modeling performance, all stop words are removed. Note that, our models were trained with original Chinese news articles, and all experimental results were manually translated into English for facilitating demonstration.

Figures 4 shows some statistics of our data set. Specifically, Figure 4(a) shows the distribution of the number of annotations on different emotions. We can observe that "*Angry*" is the most popular emotion, and other emotions have relatively even distribution, which may indicate people are more likely to show their emotions when they are angry. Moreover, Figure 4(b) shows the distribution of the number of news articles with respect to different number of emotion annotations. We can find that most of the news articles have 50 to 600 annotations.

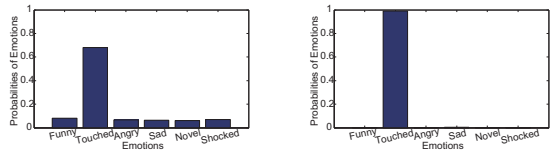
B. Performance Evaluation of eToT

In this subsection, we show the overall performance of eToT. Specifically, the hyperparameters α and β were empirically set to $50/K$ and 0.01 according to [8], and the topic number K was set to 50 by the perplexity based approach in [2]. Meanwhile, to guarantee the convergence of Gibbs sampling, all results were obtained after 500 iterations.

As we introduced above, eToT can discover the latent connections among time, emotions, and news topics. Therefore, we carefully studied 6 randomly selected latent topics learned by eToT. Specifically, the distributions of different emotions in these topics, i.e., $P(e|z)$, are shown in Figure 5. Moreover, the distributions with respect to different time spans in these topics, i.e., $P(t|z)$, are demonstrated in Figure 6. From these results, we can observe that the topics #1, #2, #3 and #5 have similar distributions of emotions, where the generation probability of emotion "*Touched*" is highest. However, these distributions over different time spans are totally different. For example, the distribution of topic #2 is relatively smooth. Therefore, we can argue that time has a significant impact on

[2]http://news.sina.com.cn/society/

[3]http://emotiondata.sinaapp.com/



(a) topic #1 (b) topic #2

(c) topic #3 (d) topic #4

(e) topic #5 (f) topic #6

Fig. 5. The distributions of different emotions in 6 different news topics learned by eToT.

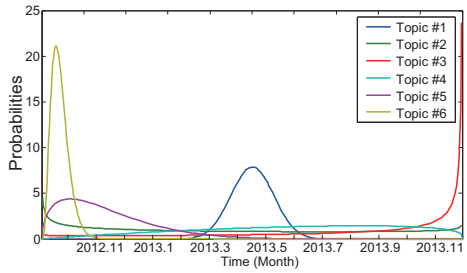


Fig. 6. The distributions of 6 topics learned by eToT with respect to different time spans.

the topic generation of eToT. It also indicates the importance of time in the process of analyzing social emotions.

Furthermore, we inspect the topics in the view of words and social events. Specifically, Table I shows the top 10 ranked words in different learned topics. We can observe that there are some special words, which are associated with some social events closely, in topic #1, #3 and #5. For example, the "lottery-ticket" and "lost" in topic #1 can help us target a social event happened on April, 2013, which is about *a woman returned a lottery with a huge bonus to its owner*. Similarly, in topic #3, "Xueying Hu" and "Suying" are names of two people who always help others. Therefore, the most important social emotion in both topic #1 and #3 is "*Touched*". Particularly, we find that words in topic #4 are insipid, and the corresponding distribution of emotions is even, which may indicate that this topic is about background.

C. Performance Evaluation of eDTM

In this subsection, we study the overall performance of eDTM. For training eDTM, we separated the data by month (one time span), and the parameter α was set to $50/K$ according to [8] firstly. With the perplexity based approach in [2], the topic number K was set to 20. In eDTM, we can treat the $\lambda_t \phi_{t-1}$ as the prior for ϕ_t. However, there is

TABLE I. THE TOP 10 RANKED WORDS IN 6 DIFFERENT TOPICS LEARNED BY ETOT.

	topic #1		topic #2		topic #3	
1	lottery	0.00462	shed	0.00797	book	0.00768
2	lost	0.00446	fierce	0.00311	X. Hu[a]	0.00593
3	inside	0.00413	immensely	0.00305	stall	0.00346
4	immensely	0.00407	money	0.00285	rape	0.00345
5	small	0.00382	the old	0.00278	silence	0.00332
6	hospital	0.00339	leave	0.00264	realm	0.00322
7	yuan	0.00335	vanish	0.00262	rubbish	0.00286
8	money	0.00331	small	0.00257	Suying	0.00267
9	make	0.00331	vein	0.00247	burglar	0.00267
10	leave	0.00326	lucky	0.00244	heal	0.00257
	topic #4		**topic #5**		**topic #6**	
1	immensely	0.00393	H. Li[b]	0.00404	go	0.00409
2	able	0.00280	Z. Wang[c]	0.00321	hospital	0.00396
3	cat	0.00217	X. Tian[d]	0.00298	intend	0.00327
4	discover	0.00208	leave	0.00288	already	0.00302
5	act	0.00206	sudoku	0.00288	yuan	0.00300
6	friend	0.00205	lonely	0.00247	edit	0.00293
7	nowadays	0.00202	money	0.00245	never	0.00291
8	inside	0.00202	sultan	0.00219	make	0.00291
9	dog	0.00197	woman	0.00215	immensely	0.00284
10	leave	0.00187	immensely	0.00214	official	0.00270

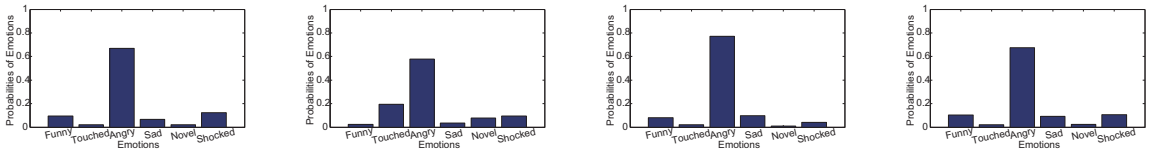
[a]Xueying Hu [b]Hunjun Li [c]Zhanfu Wangm [d]Xinbin Tian

no prior for ϕ_1. For simplifying the process of parameter estimation, we empirically set $\phi_0 = 0.01$ and $\lambda_1 = 1$ in our experiments, and estimated the λ by maximizing the joint distribution. To guarantee convergence, we implemented 500 iterations in Gibbs sampling during model training.

Figure 7 and Figure 8 show the distributions of social emotions in two randomly selected topics, which are learned by eDTM, with respect to different time spans. From these figures we can observe that the social emotions consistently evolve with the evaluation of topics, which clearly validates the motivation of eDTM. Furthermore, Table II and Table III show the top 10 ranked words in the two topics during different time spans. From these results, we find that the words in topics learned by eDTM are more common than those in topics learned by eToT. It may be because the constraints of the state space model in eDTM, thus it cannot uncover special events effectively. While eDTM could unveil the evolution of words and emotions about special kinds of events directly. Particularly, we can observe that there are always words related to crime, such as "judiciary" and "police force", in topic #1 during different time spans. These results may indicate that the social emotions evolve steadily with the evolution of the news topics. Meanwhile, in topic #1 and #2, there are 5 and 6 words (i.e., in bold) are ranked in top 10 during all of time spans, respectively. It indicates the change of words during neighboring time spans is relatively smooth and the evolution of topics captured by eDTM is stable.

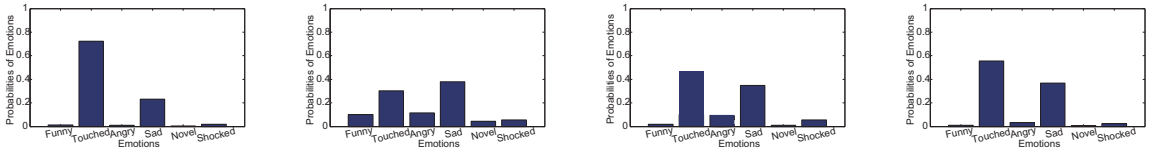
D. Evaluation of Model Application

In this subsection, we evaluate the proposed models by emotion prediction, which is one of the potential applications introduced in Section IV. Besides our two novel models, we select one state-of-the-art model, namely *Emotion-Topic Model* (ETM), proposed in [3], and a classification model, *Maximum Entropy Model* (Maxent), as the baselines. Here we treat the emotion prediction as a multi-classification problem and use all of above models to calculate the posterior probability of each emotion given a news article. Specifically, each model can generate a ranked list of emotions by calculating the probability $P(e|d)$. Thus we can measure the ranking



(a) December, 2012 (b) January, 2013 (c) February, 2013 (d) March, 2013

Fig. 7. The distribution of social emotions in topic #1 learned by eDTM with respect to different time spans.



(a) March, 2013 (b) April, 2013 (c) May, 2013 (d) June, 2013

Fig. 8. The distribution of social emotions in topic #2 learned by eDTM with respect to different time spans.

TABLE II. THE TOP 10 RANKED WORDS IN TOPIC #1 LEARNED BY EDTM DURING DECEMBER, 2012 TO MARCH, 2013

	December, 2012		January, 2013		February, 2013		March, 2013	
1	**thousand**	0.00392	**yuan**	0.00324	**yuan**	0.00374	**yuan**	0.00386
2	**yuan**	0.00363	**thousand**	0.00314	demonstrate	0.00321	**thousand**	0.00342
3	**police**	0.00343	**called**	0.00314	**man**	0.00319	**called**	0.00316
4	leave	0.00333	judiciary	0.00308	**called**	0.00309	demonstrate	0.00282
5	girl	0.00311	police force	0.00301	money	0.00282	process	0.00269
6	friend	0.00308	demonstrate	0.00281	friend	0.00274	**man**	0.00267
7	**called**	0.00293	money	0.00277	company	0.00272	money	0.00266
8	discover	0.00291	**man**	0.00272	**police**	0.00271	discover	0.00264
9	judiciary	0.00283	leave	0.00265	**thousand**	0.00267	company	0.00261
10	**man**	0.00272	**police**	0.00257	police force	0.00260	**police**	0.00260

TABLE III. THE TOP RANKED 10 WORDS IN TOPIC #2 LEARNED BY EDTM DURING MARCH, 2013 TO JUNE, 2013.

	April, 2013		May, 2013		June, 2013		July, 2013	
1	**old man**	0.00571	**leave**	0.00434	**leave**	0.00433	**son**	0.00461
2	**son**	0.00471	**son**	0.00424	**son**	0.00423	**leave**	0.00418
3	**leave**	0.00431	**old man**	0.00403	**old man**	0.00402	**old man**	0.00377
4	**father**	0.00371	**immensely**	0.00344	**immensely**	0.00343	**immensely**	0.00351
5	**mother**	0.00335	**father**	0.00333	**father**	0.00332	inside	0.00345
6	**immensely**	0.00332	inside	0.00322	job	0.00321	**father**	0.00298
7	inside	0.00321	**mother**	0.00304	**mother**	0.00303	mom	0.00294
8	money	0.00320	money	0.00297	hospital	0.00296	make	0.00287
9	hospital	0.00303	hospital	0.00292	good	0.00292	money	0.00270
10	regularly	0.00282	job	0.00289	mom	0.00258	**mother**	0.00268

performance of each model by the popular metric Normalized Discounted Cumulative Gain (NDCG). Indeed, NDCG shows how well the rank order of a given ranked list returned by an approach with a cut-off rank N is. The larger value of NDCG indicates the better ranking performance. Specifically, the Discounted Cumulative Gain (DCG) of a ranked list for a given document can be calculated by $DCG@N = \sum_{i=1}^{N} \frac{2^{rel}{}_i - 1}{log_2^{i+1}}$, where rel_i is the score of the i-th emotion. Here we set $rel_i = C_i/C$, where C_i is the number of the i-th emotion and C is the number of all of emotions. NDCG is the DCG normalized by IDCG, which is the DCG value of the ideal ranking list of the returned results and $NDCG@N = \frac{DCG@N}{IDCG@N}$.

In our experiments we used the 5-fold cross validation to evaluate each model. Table IV illustrates the $NDCG@N$ performance of three different models. From the results, we can observe that both of eToT and eDTM consistently outperform ETM with respect to different N, which clearly validates the importance of time for emotion modeling. However, the performance of eDTM is comparable with that of ETM. Moreover, it is interesting that Maxent has better prediction performance than eDTM, although it is still worse than eToT. We think it may be because that the training data of eDTM is insufficient. To be more specific, since eDTM needs to separate training data into several time spans (e.g., month in our experiments), the training data for each topic model is limited. Moreover, since Maxent cannot capture the latent semantics between emotion and text, it has worse prediction performance than eToT.

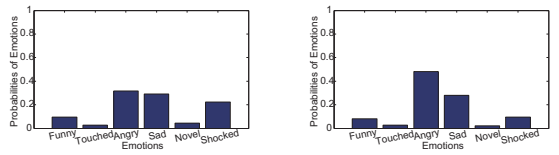
E. Case Study

In this subsection, we carefully study some topics learned by our models, and check whether they can capture the latent semantics behind social emotions.

1) Case Study for eToT: Here we study the latent semantics behind emotion "*Angry*", which can be captured by eToT.

TABLE IV. THE NDCG@N PERFORMANCE OF DIFFERENT MODELS.

	eToT	eDTM	ETM	Maxent
NDCG@1	**0.75497**	0.58638	0.56167	0.67981
NDCG@2	**0.81362**	0.67006	0.65794	0.76535
NDCG@3	**0.83923**	0.73482	0.73405	0.79810
NDCG@4	**0.86471**	0.78007	0.77139	0.82482
NDCG@5	**0.88353**	0.81118	0.80115	0.82482
NDCG@6	**0.90073**	0.83085	0.82229	0.84818



(a) topic #7 (b) topic #8

(c) topic #9 (d) topic #10

Fig. 9. Topics learned by eToT, where the emotion "*Angry*" has the highest generation probability.

Figure 9 shows 4 topics learned by eToT, where the emotion "*Angry*" has the highest generation probability. Moreover, the top ranked words in these topics are shown in Table V and the distributions with respect to different time spans are shown in Figure 10. We can find that these topics always contain words like "Police force", "crime", "hospital" and "investigate", which may indicate the semantic of crime and violence. Particularly, "Ruixi Yang", which is the top ranked word in topic #9, is the name of a murderer who killed six men and still claimed he was not regretful because he used to be bullied. Therefore, it is reasonable that "*Angry*" is the emotion with highest generation probability in these topics. Besides, the case came to court at April, 2013. It matches the distribution of topic #9 with respect to different time spans well, which implies the importance of temporal information when topic discovery.

2) Case Study for eDTM: Firstly, let us take the topic #1 learned by eDTM as an example, of which the evolution is shown in Figure 7 and Table II. As mentioned above, the top 10 ranked words imply that the topic is likely about crime. We can observe that "*Angry*" is always the most representative emotion in topic #1 during different time spans. These results

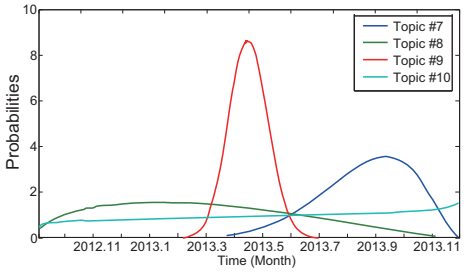


Fig. 10. The distributions of topics learned by eToT with respect to different time spans, where the emotion Angry has the highest generation probability.

TABLE V. THE TOP 10 RANKED WORDS IN TOPICS LEARNED BY ETOT, WHERE THE EMOTION "*Angry*" HAS THE HIGHEST GENERATION PROBABILITY.

	topic #7		topic #8	
1	police force	0.00534	hospital	0.00405
2	police	0.00495	called	0.00297
3	discover	0.00414	leave	0.00277
4	happen	0.00388	immensely	0.00272
5	daughter	0.00383	police force	0.00261
6	husband	0.00378	demonstrate	0.00258
7	hospital	0.00347	daughter	0.00258
8	investigate	0.00299	regularly	0.00255
9	regularly	0.00286	at once	0.00245
10	man	0.00278	discover	0.00244
	topic #9		topic #10	
1	Ruixi Yang	0.00536	police force	0.00515
2	mother	0.00482	discover	0.00407
3	son	0.00451	man	0.00377
4	leave	0.00393	yuan	0.00341
5	man	0.00361	called	0.00338
6	discover	0.00361	demonstrate	0.00311
7	wife	0.00339	investigate	0.00300
8	police force	0.00326	bank	0.00268
9	police	0.00321	crime	0.00262
10	hospital	0.00312	leave	0.00249

TABLE VI. THE NEWS ABOUT TOPICS EIGHT BETWEEN MARCH, 2013 AND MAY, 2013.

March, 2013
E-pal donated for a young ill girl voluntarily.
A bus driver is loyal and devoted to the last.
April, 2013
The post office calls for donation for Ya'an with the slogan "here no MeiMei Guo".
A older man insisted on doing good for 14 years.
May, 2013
The girl hurt when combating the blaze passed away and the remaining donation has been return.
The CEO of Alibaba are praised by people for his benefaction.

may indicate that the social emotions evolve steadily with the evolution of the news topics.

Furthermore, we inspect another topic #2 learned by eDTM, of which the evolution is shown in Figure 8 and Table III. Although the emotion "*Touched*" always has the highest generation probability in this topic during different time spans, the probability of "*Sad*" and "*Angry*" increase abnormally after April 2013. To explain the reasons behind this evolution of social emotion, we manually checked some representative news of this topic during March, 2013 to May, 2013, which are shown in Table VI. Indeed, many representative news during this period are about good Samaritans, thus the social emotion has a trend of "*Touched*". Furthermore, a terrible earthquake occurred in Ya'an, China, on April 20, 2013, and many organizations began to call for donations. The news about the earthquake may result in the generation of emotion "*Sad*". Particularly, "Meimei Guo" in the first news on April, 2013 is a key person related to the corruption of Red Cross China, which is the largest humanitarian organization in China. Although there was no evidence to prove that the corruption also existed in the disaster relief for Ya'an, the news about "Meimei Guo" may result in the social emotion "*Angry*".

VI. Conclusions

In this paper, we studied the problem of exploiting the user emotion annotations from online news to track the evolution of social emotions. We proposed two novel time-aware topic models, namely eToT and eDTM, for building connections between news topics and social emotions. Specifically, in eToT, the news topics are associated with a Beta distribution over time and a multinomial distribution over emotions. In eDTM, the state space model is leveraged for tracking the dynamics of news topics. Furthermore, we demonstrated some potential applications enabled by these two novel models, such as emotion prediction, emotion-based news recommendation and emotion anomaly detection. Finally, the extensive experiments on a real-world data set clearly demonstrate the effectiveness of our models.

VII. ACKNOWLEDGMENTS

This research was partially supported by grants from the National Science Foundation for Distinguished Young Scholars of China (Grant No. 61325010), the National High Technology Research and Development Program of China (Grant No. 2014AA015203), the Fundamental Research Funds for the Central Universities of China (Grant No. WK0110000042), NIH R21AA02395-01, and the Anhui Provincial Natural Science Foundation (Grant No. 1408085QF110).

References

[1] D. Abrams and M. A. Hogg. *Social identifications: A social psychology of intergroup relations and group processes*. Routledge, 2012.

[2] L. Azzopardi, M. Girolami, and K. van Risjbergen. Investigating the relationship between language model perplexity and ir precision-recall measures. In *Proceedings of the 26th annual international ACM SIGIR conference on Research and development in informaion retrieval*, pages 369–370. ACM, 2003.

[3] S. Bao, L. Zhang, R. Yan, Z. Su, D. Han, and Y. Yu. Joint emotion-topic modeling for social affective text mining. In *ICDM'09*, pages 699–704, 2009.

[4] D. M. Blei and J. D. Lafferty. Dynamic topic models. In *Proceedings of the 23rd international conference on Machine learning*, pages 113–120. ACM, 2006.

[5] D. M. Blei, A. Y. Ng, and M. I. Jordan. Latent dirichlet allocation. *the Journal of machine Learning research*, 3:993–1022, 2003.

[6] F.-R. Chaumartin. Upar7: A knowledge-based system for headline sentiment tagging. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 422–425. Association for Computational Linguistics, 2007.

[7] A. S. Das, M. Datar, A. Garg, and S. Rajaram. Google news personalization: scalable online collaborative filtering. In *Proceedings of the 16th international conference on World Wide Web*, pages 271–280. ACM, 2007.

[8] G. Heinrich. Parameter estimation for text analysis. *Web: http://www. arbylon. net/publications/text-est. pdf*, 2005.

[9] T. Iwata, S. Watanabe, T. Yamada, and N. Ueda. Topic tracking model for analyzing consumer purchase behavior. In *IJCAI*, volume 9, pages 1427–1432, 2009.

[10] T. Iwata, S. Watanabe, T. Yamada, and N. Ueda. Topic tracking model for analyzing consumer purchase behavior. In *IJCAI*, volume 9, pages 1427–1432, 2009.

[11] T. Iwata, T. Yamada, Y. Sakurai, and N. Ueda. Sequential modeling of topic dynamics with multiple timescales. *ACM Transactions on Knowledge Discovery from Data (TKDD)*, 5(4):19, 2012.

[12] Y. Koren. Collaborative filtering with temporal dynamics. *Communications of the ACM*, 53(4):89–97, 2010.

[13] Z. Kozareva, B. Navarro, S. Vázquez, and A. Montoyo. Ua-zbsa: a headline emotion classification through web information. In *Proceedings of the 4th International Workshop on Semantic Evaluations*, pages 334–337. Association for Computational Linguistics, 2007.

[14] G. Le Bon. *The crowd: A study of the popular mind*. Macmillan, 1897.

[15] B. Li, X. Zhu, R. Li, C. Zhang, X. Xue, and X. Wu. Cross-domain collaborative filtering over time. In *Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Three*, pages 2293–2298. AAAI Press, 2011.

[16] C. Lin and Y. He. Joint sentiment/topic model for sentiment analysis. In *Proceedings of the 18th ACM conference on Information and knowledge management*, pages 375–384. ACM, 2009.

[17] K. H.-Y. Lin and H.-H. Chen. Ranking reader emotions using pairwise loss minimization and emotional distribution regression. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, pages 136–144. Association for Computational Linguistics, 2008.

[18] K. H.-Y. Lin, C. Yang, and H.-H. Chen. What emotions do news articles trigger in their readers? In *Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 733–734. ACM, 2007.

[19] J. Liu, P. Dolan, and E. R. Pedersen. Personalized news recommendation based on click behavior. In *Proceedings of the 15th international conference on Intelligent user interfaces*, pages 31–40. ACM, 2010.

[20] K.-L. Liu, W.-J. Li, and M. Guo. Emoticon smoothed language models for twitter sentiment analysis. In *AAAI*, 2012.

[21] Q. Liu, Y. Ge, Z. Li, E. Chen, and H. Xiong. Personalized travel package recommendation. In *Data Mining (ICDM), 2011 IEEE 11th International Conference on*, pages 407–416. IEEE, 2011.

[22] Q. Mei, X. Ling, M. Wondra, H. Su, and C. Zhai. Topic sentiment mixture: modeling facets and opinions in weblogs. In *Proceedings of the 16th international conference on World Wide Web*, pages 171–180. ACM, 2007.

[23] T. Minka. Estimating a dirichlet distribution, 2000.

[24] G. Mishne and M. de Rijke. Capturing global mood levels using blog posts. In *AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs*, pages 145–152, 2006.

[25] B. Pang, L. Lee, and S. Vaithyanathan. Thumbs up?: sentiment classification using machine learning techniques. In *Proceedings of the ACL-02 conference on Empirical methods in natural language processing-Volume 10*, pages 79–86, 2002.

[26] H. Tang, S. Tan, and X. Cheng. A survey on sentiment detection of reviews. *Expert Systems with Applications*, 36(7):10760–10773, 2009.

[27] I. Titov and R. McDonald. A joint model of text and aspect ratings for sentiment summarization. *Urbana*, 51:61801, 2008.

[28] C. Wang, D. Blei, and D. Heckerman. Continuous time dynamic topic models. *arXiv preprint arXiv:1206.3298*, 2012.

[29] X. Wang and A. McCallum. Topics over time: a non-markov continuous-time model of topical trends. In *Proceedings of the 12th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 424–433. ACM, 2006.

[30] C. Yang, K. H.-Y. Lin, and H.-H. Chen. Building emotion lexicon from weblog corpora. In *Proceedings of the 45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions*, pages 133–136. Association for Computational Linguistics, 2007.

[31] H. Yu and V. Hatzivassiloglou. Towards answering opinion questions: Separating facts from opinions and identifying the polarity of opinion sentences. In *Proceedings of the 2003 conference on Empirical methods in natural language processing*, pages 129–136, 2003.

[32] J. Zhao, L. Dong, J. Wu, and K. Xu. Moodlens: an emoticon-based sentiment analysis system for chinese tweets. In *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 1528–1531. ACM, 2012.

[33] H. Zhu, E. Chen, H. Xiong, H. Cao, and J. Tian. Mobile app classification with enriched contextual information. *IEEE Transactions on Mobile Computing*, 13(7):1550–1563, 2014.

[34] H. Zhu, H. Xiong, Y. Ge, and E. Chen. Mobile app recommendations with security and privacy awareness. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, KDD, pages 951–960, New York, NY, USA, 2014. ACM.