IEEE Access
Multidisciplinary : Rapid Review : Open Access Journal

Received July 20, 2019, accepted August 9, 2019, date of publication August 13, 2019, date of current version August 28, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2935155

Personalized Travel Recommendation Based on Sentiment-Aware Multimodal Topic Model

XI SHAO[1,2], **(Member, IEEE), GUIJIN TANG**[1], **AND BING-KUN BAO**[1]
[1]College of Communication and Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China
[2]National Laboratory of Pattern Recognition, Institute of Automation, CAS, Beijing 100190, China

Corresponding author: Xi Shao (shaoxi@njupt.edu.cn)

This work was supported in part by the National Science Foundation of China under Grant 61872199, Grant 61872424, and Grant 61772287, in part by the Key University Science Research Project of Jiangsu Province under Grant 18KJA510004, and in part by the Nanjing University of Posts and Telecommunications Support Funding under Grant NY218001.

ABSTRACT In this paper, we try to solve the personalized travel recommendation problem by exploiting the multi-modal data available from the real world social media, and a probabilistic graph model so called Sentiment-aware Multi-modal Topic Model (SMTM) is proposed to mine the latent semantics of the multi-modal data on the online travel website. Distinguished from previous approaches, our proposed approach try to mine the topics from tourist and attraction domains separately for disclosing semantics for tourist topics and attraction themes. In addition, we analyze tourist's sentiments on attractions to further obtain the tourist's attitude over attractions and recommend the attraction with proper sentiment on the related attraction themes accordingly. Based on the proposed SMTM model, the documents in tourist domain and in attraction domain can be compared with each other after they were projected into the mutual topic space, and this latent space projection scheme can be further applied to two personalized traveling recommendations, that is, the single platform traveling recommendation and the inter-platform traveling recommendation. Evaluation results based on the real world online travel website have shown the improved performance of our method.

INDEX TERMS Tourism recommendation, multi-modality, topic model, sentiment analysis.

I. INTRODUCTION

Many social networks surge up with the arising of Web 2.0, leading to the tremendous online propagation of the User Generated Content (UGC), which distributes over multi-networks. Exploiting and aggregating user generated data from online network rises up as a solution towards complete and timely semantic modeling to improve the performance of multimedia based applications, such as searching, annotating, recommendation and advertising. Among all these applications, intelligent travel recommendation is one of the most attractive applications for researchers because it is closely related to people's everyday life. According to the statistics conducted by World Travel & Tourism Council, more and more travel companies provide on line services, and people, especially younger generation prefers to check the travel website for the attraction selection before they plan to visit. For instance, TripAdvisor (https://www.tripadvisor.in) is one of the most popular on-line travel websites where people can share their options and sentiments about the attractions they visited. However, due to the rapid development of these travel websites, a large amount of unorganized information hinders users from quickly and efficiently finding the desired tourist attractions. Moreover, in order to increase the profit, travel companies have to understand the preferences of different travelers and provide them with more attractive suggestions. Therefore, the expected demand of travel recommendation service will increase substantially.

In general, there are two applications in a typical travel recommendation system: attraction recommendation and potential tourist recommendation. For the personalized attraction recommendation, it contains recommended attraction information for the destination when a given user is planning a trip; for example, answering the question such as I want to go to New York, what are the must-see places there? Potential tourist recommendations take into account the traveler's personal preference in order to provide more appropriate traveler lists that match the attractive theme. Based on

The associate editor coordinating the review of this article and approving it for publication was Dian Tjondronegoro.

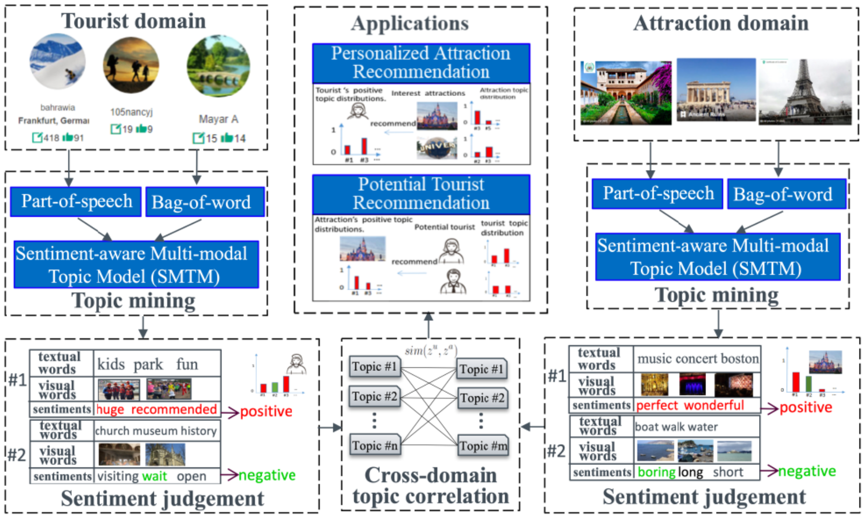

FIGURE 1. Framework of sentiment-aware multi-modal recommendation.

above observation, a successful scheme for tourist attraction recommendation should consider three crucial factors. The first one is the "tourist preferences", which can be obtained from those places he/she visited, followed, commented and liked. The second one is the "attraction themes", which are the types of experience that tourists can get through visiting. For example, the theme of British Museum seems most likely to be "historical" over "entertainment". Of course, an attraction could have multiple themes. The last one is the "sentiment on a theme of the attraction", which measures the option of the attraction on a certain theme from the perspective of the visitors. Based on above analysis, the task of personalized attraction recommendation can be decomposed into excavating the visitor's preferences, then selecting the proper attraction theme, and finally returning the attraction with the positive sentiments of the selected theme.

Some researchers have been devoted to the work of personalizing the travel recommendation. However, existing methods either mainly focused on single modality such as texts, images, etc., leaving the other modality out of consideration [1], [2], or directly researched on tourist topics rather than tapping the preferences of tourists with analyzing the options of a tourist over the attractions [3], [4], and their limitations mainly lie in two aspects: firstly, their performance is not very satisfactory since only one modality is not exact and informative enough to characterize the tourists' preferences and the themes of the attractions; secondly, ignoring tourists' perceptions of the theme, it is likely that some of the recommended attractions are not what tourists really expect. Therefore, there is an increasing demand of utilizing multi-modal tourism data to do tourist attraction recommendation by considering tourist perception, attraction theme, and sentiment on the attraction.

There are three challenges on multi-modal tourist attraction recommendation task. The first one is how to handle the multi-modal data from tourist part and attraction part, which include attraction scene images, textual descriptions, comments and sentiments. The second challenge is how to find the latent semantic spaces for attraction part and tourist part. i.e., the comments are the shared information from the tourist part and the attraction part. Besides these, tourists' part also provide self-introductions, linked articles, and so on, while attraction part also has its own description. In addition, there are always different sentiments on every topic. On tourist part, the sentiments usually indicate their preferences. On attraction part, the sentiments can be positive or negative. Therefore, we need to regard tourist and attraction as two domains. The third challenge is how to link the semantic space from both domains for recommendation.

To handle above mentioned challenges, we proposed a Sentiment-ware Multi-modal Topic Model (SMTM) to discover the relationship between the tourist domain and attraction domain. Fig.1 shows the framework of our proposed scheme. As the figure shows, firstly, we consider the tourist domain and attraction domain separately. Then we represent visual modality by bad-of-word and employ part-of-speech toolbox to classify the textual words into sentiment words and non-sentiment words. After that, we propose SMTM model to mine topics. In the third step, we judge the sentiment of each topic, and then correlate the topics from two domains. Based on the proposed SMTM, we develop a travel recommendation framework, in which the documents in tourist domain and in attraction domain can be compared with each other after they were projected into the mutual topic space, and this latent space projection scheme can be further applied to two applications, that is, the single platform traveling recommendation and the inter-platform traveling recommendation.

The contributions of this paper are summarized as follows:

-We propose a SMTM model which takes into account of three preliminary factors in traveling recommendation problem and the advantages of SMTM model include: (1) the multi-modal data both for tourist domain and attraction domain is fully exploited for better semantics disclosing; (2) the topics in tourists domain and the themes in attractions domain are separately modeled for better disclosing relationships in corresponding semantic space; (3) the tourist sentiments on topics are studied to obtain traveler's opinion.

-We propose a SMTM based traveling recommendation framework which employs mutual document projection for tourist domain and for attraction domain into a latent semantic space, and this latent space projection scheme can be further applied to two personalized recommendations, that is, the single platform traveling recommendation and the inter-platform traveling recommendation.

The rest of this paper is organized as follows. In Section II, we briefly reviewed the related work to travel recommendation. Section III details our proposed SMTM model for travel recommendations. Section IV introduces the proposed mutual latent semantic space projection scheme based on SMTM model and with applying to two applications, namely, single-platform personalized travel recommendation and inter-platform personalized travel recommendation. The experimental results are reported and analyzed in Section V, which is followed by our conclusion in Section VI.

II. RELATED WORK

In this section, two groups of existing related research work are reviewed. The first group introduces the related work on topic modeling with applications on social media analysis, and the second group focuses on personalized travel recommendations.

A. PROBABILISTIC TOPIC MODEL AND ITS EXTENSION

Probabilistic Topic Models (PTM) are statistical algorithms whose aim is to discover the latent semantic structures in large archives of documents. A more comprehensive survey of Probabilistic Topic Models can be found in [5], and it has been successfully applied to many fields (e.g. text, images, music and videos) for various tasks such as classification [6], information retrieval [7] and recommendation [8].

A number of topic models have been proposed in the literature. The Latent Dirichlet Allocation (LDA) [9] model is one of the broadly studied PTMs as it possesses totally generative semantics. Rosen-Zvi *et al*. [10] extended LDA with the inclusion of metadata variables into the model, and it introduced the Author-Topic (AT) model to incorporate the author attribute by modifying LDA's assumption that authors, not documents, are a multinomial distribution over the topics. The Author-Recipient-Topic (ART) model, proposed by McCallum *et al*. [11], extended the idea further by building a topic distribution for every author recipient pair. These generative models assumed the metadata is generated by the hidden topics and the topics are word distributions as well as distributions over the metadata variables. In order to adapt the LDA model to textual, auditory and visual modalities, some variants of topic models were proposed, such as multimodal-LDA [12] and correspondence LDA [6], which employed a set of shared latent variables to explicitly model images and annotated text to capture semantic correlations between the data of two modalities. However, these works focused on topic mining with single modality or multiple modalities without incorporating user's sentiments over the topics, which will result in the impropriate recommendation results, that is, some of the recommended documents are not what the user really needs.

In order to incorporate the subjective emotion in the corpus for improving the recommendation performance, some researchers worked on disclosing sentiment/opinion with topic models [13]–[16]. In [13], Mei *et al*. proposed the Topic Sentiment Mixture model which embeds the topic and sentiment in a unified framework to reveal the latent topic aspects in a Weblog collection and their corresponding opinion. However they only considered the limited but fixed states on opinion spaces like ''negative, positive, neutral''. A domain independent topic-sentiment model, so called Joint Multigain Topic Sentiment was proposed by Alam *et al*. [16], in which review-specific elements and ratable aspects were modeled by global and local topics, respectively, and thereby eliminating the requirement for manually probing for the sentiment categories. To improve the scalability of the topic-sentiment model, Titov and Ryan [17] proposed a Multi-Aspect Sentiment model which can identify the relevant aspects for a rated entity and extracted all textual semantics associated with those aspects. Fang *et al*. [18] presented an opinion mining approach to disclose the opinions of the individual perspectives on the topic. Recently, there is an increasing tendency for research works on topic modelling with the inclusion of multi-modal data and sentiments. In [15], A Multi-modal Joint Sentiment Topic Model was proposed for weakly supervised sentiment analysis on texts and emoticons in microblogging, which applies LDA to simultaneously analyze sentiments and topics hidden in microblog messages. Fang *et al*. [19] designed the Multi-modal Aspect-opinion Model to find the correlations between textual and visual modalities by considering both user-generated images and textual documents. In [20], a multi-modal multi-view topic opinion mining model was proposed for social event analysis from multiple collection sources. Different from the above approaches based on the single domain, our approach considers topic spaces in the tourist domain and the attraction domain respectively, and the two topic spaces are correlated and associated for better semantics disclosing.

B. PERSONALIZED TRAVEL RECOMMENDATION

Recently, the high demand for recommendation systems has led to a boom in research in this field. Generally speaking, according to the data source used in the recommendation system, the travel recommendation approaches can fall into two categories: GPS trajectory based approach and travelogues based approach. GPS trajectory based approach mainly utilized the GPS data obtained by the receivers to infer the attraction preferences of the traveler. Averjanova *et al.* [21] proposed a map-based conversational mobile recommender system by integrating GPS data and electronic map technologies to support users with some personalized recommendations. Carolis *et al.* [22] employed an electronic map for outlining the location of interests and generating comparative descriptions to support users in choosing the attraction to visit. Ricci and Nguyen [23] proposed a more sophisticated on-tour support system, so called MobyRek, to recommend tourism products to on-the-move travelers. The main problem for GPS trajectory based approach is that the GPS data are not always available. To solve this problem, some researchers considered obtaining the trip-related knowledge from the user generated contents. Crandall *et al.* [24] proposed to estimate the location of a photo on a large scale photo database by combining tags and visual features from the photo. Arase *et al.* [25] categorized the geotagged photos into six trip patterns and travelers can browse typical photos of six categories to decide where to go. Wang *et al.* [26] proposed to model the traveler's flexible interests with location-aware user mobility modelling. Although they considered the tourist preferences, they did not analyze the theme from attraction domain. In order to facilitate travel planning, some researches focused on analyzing the theme of the attraction [3], [19], [27], [28]. The approaches mentioned above have considered latent semantic distribution from either tourist domain or attraction domain, but the majority of these methods are lack of analysis of sentiment on themes of attractions. Shen *et al.* [29] presented a personalized travel recommendation scheme by utilizing traveler's interaction with the system and the heterogeneous travel information. This approach took into account the sentiments of the attraction themes and exploits multi-modal data for topic modelling, but failed to decompose the overall opinion for an attraction into opinions of the travelers who actually visited it. In [30], we proposed to incorporate the tourist sentiments on topic model to retain the tourist preferences, however, ignoring to fully exploit the latent relations between topics in tourist domain and attraction domain significantly limited the performance of the travel recommendation. The extensions of this paper include (1) extending the SMTM model for personalized recommendation framework with mutual latent topic space projection in tourist domain and attraction domain, (2) extending the recommendation framework with incorporating the topic-oriented sentiments factors, which can be obtained by decomposing the overall opinion for a document over the topics. (3) extending the SMTM based recommendation for the inter-platform personalized application and evaluating them through a large-scale real-world dataset from two platforms.

III. MULTI-MODAL TOPIC MODEL

In this section we introduce the proposed SMTM model, which is proposed based on the assumption that tourists have

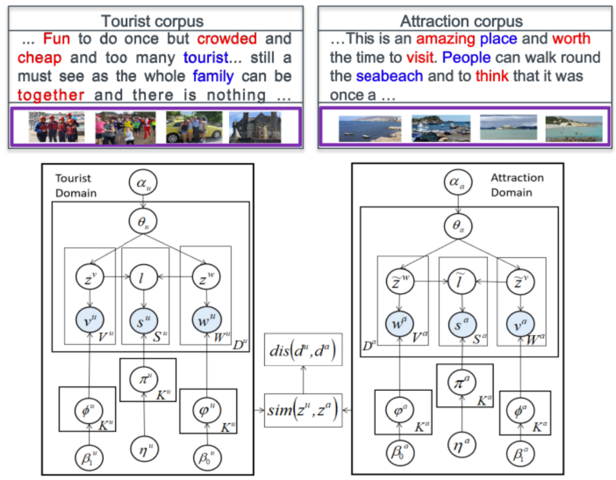

FIGURE 2. **Representation of Sentiment-aware Multi-modal Topic Model.**

traveling preferences specific to topics related to the theme of attraction and the consistent subjective options over the attractions they want to visit. In light of this, we present SMTM with diagram shown in Fig.2, and Table I summarizes the relevant notions used. To simulate generation process of the observed tourists' posts and the attraction documents, which are generated by different users from different perspective, we firstly split the corpus on the platform into tourist domain and attraction domain. Then we perform semantic analysis for documents in two domains respectively, and associate them in some semantic spaces. Finally we try to capture the correlation between tourist topic space and attraction theme spaces for recommendation.

A. PROBLEM DEFINITION

According to above discussion, we can define the problem of SMTM as follows.

Definition 1: Supposing that we are given a collection of tourists' posts on a platform denoted as $U = \{d_1^u, \ldots, d_{D^u}^u\}$ and a collection of the attraction posts on the social platform denoted as $A = \{d_1^a, \ldots, d_{D^a}^a\}$, where $d_i^u = \{U_i^W, U_i^V, U_i^S\}$ are composed of three components: tourist textual component U^W, tourist visual component U^V, and tourist sentiment component U^S, and $d_i^a = \{A_i^W, A_i^V, A_i^S\}$ are composed of three components: attraction textual component A^W, attraction visual component A^V, and attraction sentiment component A^S, we are supposed to estimate the following parameters from a given dataset:

1) Tourists topic space parameter φ^u, ϕ^u and attraction theme space parameter φ^a, ϕ^a;
2) Corresponding sentiment space parameter π^u and π^a;
3) The topic distribution of tourist domain document θ^u and theme distribution of attraction domain document θ^a;
4) The correlation between tourist and attraction topic spaces $sim(z^u, z^a)$, for $z^u \in \{1, 2, \ldots, K^u\}$ and $z^a \in \{1, 2, \ldots, K^a\}$, which can be measured by all the similarities between tourist topics and attraction themes.

TABLE 1. The key notations of the proposed SMTM model.

Notations	Description
$\boldsymbol{U},\boldsymbol{A}$	tourist document set, attraction document set
D^U,D^A	number of tourist documents and attraction documents
K^U,K^A	number of tourist topics and attraction topics
U^W, U^V, U^S	textual word vocabulary, visual word vocabulary and sentiment word vocabulary in tourist document set
A^W, A^V, A^S	textual word vocabulary, visual word vocabulary and sentiment word vocabulary in attraction document set
W^u,V^u,S^u,W^a,V^a,S^a	number of word in U^W, U^V, U^S, A^W, A^V, A^S
φ^u,ϕ^u,π^u	the multinomial distributions over textual words, visual words and sentiment words for tourist topics
φ^a, ϕ^a,π^a	the multinomial distributions over textual words, visual words and sentiment words for attraction topics
θ^u_d, θ^u_a	the multinomial distributions over topics for tourists and attractions
z^w, z^v, l	topic assignment for textual word, visual word, and sentiment word on tourist domain
$\tilde{z}^w, \tilde{z}^v, \tilde{l}$	topic assignment for textual word, visual word, and sentiment word on attraction domain
$\alpha^u, \alpha^a, \eta^u, \eta^a$	Dirichlet priors to multinomial distribution θ^u,θ^a, π^u,π^a
$\beta^u_0, \beta^a_0, \beta^u_1, \beta^a_1$	Dirichlet priors to multinomial distribution φ^u,φ^a, ϕ^u,ϕ^a
$\mathbf{w}_{-i}, \mathbf{z}_{-i}, \mathbf{v}_{-i}, \mathbf{s}_{-i}, \mathbf{l}_{-i}$	vector values of **w**, **z**, **v**, **s**, **l** on all the other dimensions except i

B. TOPIC AND SENTIMENT ANALYSIS ON TOURIST DOMAIN

The aim of topic and sentiment analysis is to obtain the traveler's preferences from tourist corpus which is composed of texts and images. We assume that there are K^u topics in tourist domain. With the proposed SMTM model, the document topic distributions parameter θ^u controls the topic structure in a tourist document, while the textual, visual and sentiment words are generated from the textual multinomial distribution φ^u, visual multinomial distribution ϕ^u and sentiment multinomial distribution π^u conditioned on the corresponding topics, respectively. We can summarized the whole generative process of a document d^u in tourist domain with SMTM model as follows.

1) For each tourist topic $z^u \in \{1, 2, \cdots, K^u\}$ including textual topic z^w and visual topic z^v, draw a multinomial distribution over topic words, $\phi^u \sim Dir(\beta^u_0)$ and $\varphi^u \sim Dir(\beta^u_1)$.
2) For each tourist topic $z^u \in \{1, 2, \cdots, K^u\}$, draw a multinomial sentiment word distribution $\pi^u \sim \mathrm{Dir}(\eta^u)$.
3) For each document d^u:
   (a) Draw a multinomial distribution $\sim^u_d Dir(\alpha^u)$ for document.
   (b) For each textual word w in document d^u: draw a topic $z^w_d \sim Multi(\theta^u_d)$, a textual word $w \sim Multi(\varphi^u_{z^w_x})$.
   (c) For each visual word v in document d^u: draw a topic $z^v_d \sim Multi(\theta^u_d)$, a visual word $v \sim Multi(\phi^u_{z^v_d})$.
   (d) For each sentiment word s in document d^u: draw a topic assignment $l \sim Uniform(z^u_1, z^u_2, \ldots, z^u_{k^u})$, a sentiment word $s \sim Multi(\pi^u_l)$.

The tourist domain SMTM can be inferred by Gibbs sampling [31]. There are three set of latent variables in the model: the textual topic assignment $\mathbf{z}^w$, the visual topic assignment $\mathbf{z}^v$ and the sentiment distribution l. The Gibbs sampler generates posterior samples by sweeping through each variable to sample from its conditional distribution with the remaining variables fixed to their current values. The update rules for latent variables $\mathbf{z}^w$, $\mathbf{z}^v$ and $\mathbf{l}$ are as follows:

$$p(z^w_i = k^u|\mathbf{w}, \mathbf{z}^w_{-i}) \propto \frac{n^u_{kd,-i} + \alpha^u}{\sum_{k=1}^{K^u} n^u_{kd,-i} + K^u\alpha^u} \times \frac{n^u_{wk,-i} + \beta^u_0}{\sum_{w=1}^{W^u} n^u_{wk,-i} + W^u\beta^u_0} \tag{1}$$

$$p(z^v_i = k^u|\mathbf{v}, \mathbf{z}^v_{-i}) \propto \frac{n^u_{kd,-i} + \alpha^u}{\sum_{k=1}^{K^u} n^u_{kd,-i} + K^u\alpha^u} \times \frac{n^u_{vk,-i} + \beta^u_1}{\sum_{v=1}^{V^u} n^u_{vk,-i} + V^u\beta^u_1} \tag{2}$$

$$p(l_i = m^u|\mathbf{s}, \mathbf{l}_{-i}) \propto \frac{n^u_{sm,-i} + \eta^u}{\sum_{s=1}^{S^u} n^u_{sm,-i} + S^u\eta^u} \times \frac{n^u_{md}}{N^u_{kd}} \tag{3}$$

where the subscript $-i$ means a counting variable that excludes the i-th word index in the corpus. $n^u_{kd,-i}$ is the times of words for topic k^u being generated from document d^u except the current assignment. $n^u_{wk,-i}$ denotes the times of word w being generated from topic k^u except the current assignment. $n^u_{uk,-i}, n^u_{sm,-i}, n^u_{md}$ is similarly defined. n^u_{kd} means the times of all topic words in document d^u. After model inference, we can estimate the parameters of SMTM model in tourist domain as follows:

$$\theta^u_{kd} = \frac{n^u_{kd} + \alpha^u}{\sum_{k=1}^{K^u} n^u_{kd} + K^u\alpha^u}, \quad \varphi^u_{wk} = \frac{n^u_{wk} + \beta^u_0}{\sum_{w=1}^{W^u} + W^u\beta^u_0}$$
$$\phi^u_{vk} = \frac{n^u_{vk} + \beta^u_1}{\sum_{v=1}^{V^u} + V^u\beta^u_1}, \quad \pi^u_{sm} = \frac{n^u_{sm} + \eta^u}{\sum_{s=1}^{S^u} n^u_{sm} + S^u\eta^u}. \tag{4}$$

We can perform sentiment analysis with SentiWordNet [32], a lexical toolbox for sentiment computation. It is employed to compute the sentimental value (ranges from -1 to 1) to every sentiment word. The closer its value is to 1, the more likely it is to be positive, otherwise to be negative. Then the sentiment score of tourist topic k can be calculated as:

$$Q^u(k) = \frac{1}{2}\left[\sum_{w=1}^{Nwk} p(w|z^w = k) \cdot Q_w + \sum_{s=1}^{Nsk} p(s|l^s = k) \cdot Q_s\right]. \tag{5}$$

In this equation, Q_s and Q_w are the sentiment scores of a sentiment word s and a topic word w, respectively. $Q^u(k)$ is the overall sentiment score of the k-th tourist topic.

C. THEME AND SENTIMENT ANALYSIS ON ATTRACTION DOMAIN

The purpose of theme modeling on attraction domain is to identify the latent semantic structure for attraction on a traveling-related platform, while sentiment mining on attraction domain is to analyze traveler's sentiments over the themes of attraction. Analog to the tourist domain topic modeling, the similar generative process can be employed to generate the text and visual content in attraction domain. So we only write down the following key formulas. The update rules for latent variables in attraction domain are as follows:

$$p(\widetilde{z}_i^w = k^a|\mathbf{w},\widetilde{\mathbf{z}}_{-i}^w) \propto \frac{n_{kd,-i}^a + \alpha^a}{\sum_{k=1}^{K^a} n_{kd,-i}^a + K^a\alpha^a} \times \frac{n_{wk,-i}^a + \beta_0^a}{\sum_{w=1}^{W^a} n_{wk,-i}^a + W^a\beta_0^a} \tag{6}$$

$$p(\widetilde{z}_i^v = k^a|\mathbf{v},\widetilde{\mathbf{z}}_{-i}^v) \propto \frac{n_{kd,-i}^a + \alpha^a}{\sum_{k=1}^{K^a} n_{kd,-i}^a + K^a\alpha^a} \times \frac{n_{vk,-i}^a + \beta_1^a}{\sum_{v=1}^{V^a} n_{vk,-i}^a + V^a\beta_1^a} \tag{7}$$

$$p(\widetilde{l}_i = m^a|\mathbf{s},\widetilde{\mathbf{l}}_{-i}) \propto \frac{n_{sm,-i}^a + \eta^a}{\sum_{s=1}^{S^a} n_{sm,-i}^a + S^a\eta^a} \times \frac{n_{md}^a}{N_{kd}^a}. \tag{8}$$

After sampling, the corresponding parameters for attraction domain can be estimated as follows:

$$\theta_{kd}^a = \frac{n_{kd}^a + \alpha^a}{\sum_{k=1}^{K^a} n_{kd}^a + K^a\alpha^a}, \quad \varphi_{wk}^a = \frac{n_{wk}^a + \beta_0^a}{\sum_{w=1}^{W^a} + W^a\beta_0^a}$$
$$\phi_{vk}^a = \frac{n_{vk}^a + \beta_1^a}{\sum_{v=1}^{V^a} + V^a\beta_1^a}, \quad \pi_{sm}^a = \frac{n_{sm}^a + \eta^a}{\sum_{s=1}^{S^a} n_{sm}^a + S^a\eta^a} \tag{9}$$

Similarly, the sentiment score of k-th attraction theme is:

$$Q^a(k) = \frac{1}{2}\left[\sum_{w=1}^{Nwk} p(w|\widetilde{z}^w = k)\cdot Q_w + \sum_{s=1}^{Nsk} p(s|\widetilde{l}^s = k)\cdot Q_s\right]. \tag{10}$$

D. CORRELATION BETWEEN TOURIST TOPIC SPACE AND ATTRACTION THEME SPACE

The similarities between tourist topics and attraction themes should be calculated before they can be compared. Inspired by [33], we employ symmetric Kullback-Leibler distance to measure the similarity of two spaces which is defined by:

$$sim(z^u, z^a) = \frac{1}{\sum_i p(i|z^u)\log\frac{p(i|z^u)}{p(i|z^a)} + \sum_i p(i|z^a)\log\frac{p(i|z^a)}{p(i|z^u)}} \tag{11}$$

where i indexes the word which occurs in both domains.

IV. APPLICATIONS

In this section, we introduce how to apply the proposed SMTM to two recommendation applications: that is, the attraction recommendation and the potential tourist recommendation.

A. PERSONALIZED ATTRACTION RECOMMENDATION

For a tourist query d_j^u and an attraction d_i^a in $A = \{d_1^a \ldots, d_n^a\}$, the multinomial distribution of tourist topic and attraction theme are $\theta_{d_j}^u$ and $\theta_{d_i}^a$, respectively. Then the tourist and attraction space $\mathbf{z}^u$ and $\mathbf{z}^a$ can be obtained from the proposed SMTM model. To recommend a proper attraction document to a tourist, we need to project the attraction document from the attraction domain to the tourist domain, taking into account both the topic space projection and the sentiment factor projection. The topic space projection from the attraction domain to the tourist domain can be evaluated by the distribution on j-th tourist topic z_j^u given by i-th attraction document d_i^a:

$$p(z_j^u|d_i^a) = \frac{\sum_m p(z_m^a|d_i^a)sim(z_m^a, z_j^u)}{\sum_n\sum_m p(z_m^a|d_i^a)sim(z_m^a, z_n^u)}. \tag{12}$$

Taking into account the sentimental factors from the both domain, we employ the following equation to project the sentiment score of the i-th attraction document to the k-th topic in tourist domain:

$$Q^{ua}(k, i) = \sum_m Q^a(m, i)w_{mk}p(z_m^a|d_i^a) \tag{13}$$

where $Q^a(m, i)$ denotes the sentiment score for the m-th attraction theme in the i-th attraction document, which can be obtained by Eq.(10), while w_{mk} denotes normalized weight for this projection:

$$w_{mk} = \frac{sim(z_m^a, z_k^u)}{\sum_n sim(z_n^a, z_k^u)}. \tag{14}$$

The distance between d_i^u and each document of attraction documents set $\boldsymbol{A}$ can be calculated and ranked by following equation:

$$dis(d_j^u, d_i^a) = \sum_n \sqrt{Q^u(n,j)Q^{ua}(n,i) \times \left[p(z_n^u|d_j^u) - p(z_n^u|d_i^a)\right]^2} \tag{15}$$

where $Q^u(n, j)$ denotes the sentiment score for the n-th tourist topic in the j-th tourist document, which can be calculated by Eq.(5).

B. POTENTIAL TOURIST RECOMMENDATION

Potential tourist recommendation is similar to the interest attraction recommendation. Specifically, given an attraction d_i^a and a tourist d_j^u in set $U = \{d_1^u, \ldots, d_n^u\}$. The theme and topic distribution $\theta_{d_i}^a, \theta_{d_j}^u$ and the topic and theme space $\mathbf{z}^a$, $\mathbf{z}^u$ can be learned by proposed model. To recommend the tourists to the attraction, we need to project the tourist document from the tourist domain to the attraction domain, taking into account both the topic space projection and the sentiment factor projection.

The topic space projection from the tourist domain to the attraction domain can be evaluated by the distribution on i-th attraction theme z_i^a given by j-th tourist document d_j^u:

$$p(z_i^a|d_j^u) = \frac{\sum_m p(z_m^u|d_j^u)sim(z_m^u, z_i^a)}{\sum_n \sum_m p(z_m^u|d_j^u)sim(z_m^u, z_n^a)}. \tag{16}$$

The sentiment score projection of the j-th tourist document to the k-th topic in attraction domain can be obtained by:

$$Q^{au}(k,j) = \sum_m Q^u(m,j)w_{mk}p(z_m^u|d_j^u). \tag{17}$$

Then the distance between d_i^a and each tourist document from U can be calculated by follows:

$$dis(d_i^a, d_j^u) = \sum_n \sqrt{Q^a(n,i)Q^{au}(n,j) \times \left[p(z_n^a|d_j^u) - p(z_n^a|d_i^a)\right]^2}. \tag{18}$$

The recommended tourists are ranked in descending order and *top-k* of the potential tourists are recommended to travel attraction d_i^a.

V. EXPERIMENTS

In this section, we perform some experiments to evaluate our proposed SMTM model and to compare with the state-of-the-art topic models for the online travel platform recommendation. We firstly elaborate how we obtain the experimental dataset and describe the implementation details. Then we define the evaluation strategies and evaluate our proposed SMTM model for topic mining on a real world online travel website. Finally, we validate the performance of the proposed recommendation scheme in the single platform and the inter-platform applications.

A. EXPERIMENTAL SETTINGS

We constructed two datasets for evaluation, one for single platform validation and the other for the inter-platform validation. The single platform dataset I was constructed from TripAdvisor, where we collected multi-modality data from tourist and attraction domains respectively. In dataset I, we have 459,180 textual comments or descriptions, and 43,964 images from 14,648 tourist documents, while 392,680 textual comments and 26,182 images from 8,724 attraction documents, with each at least 20 comments and 1 image. The dataset II was constructed from another online travel website, Trip (https://www.trip.com/), where we only collected attraction domain data. The dataset II includes 293,847 textual comments and 19,492 images from 6,513 attractions.

We employ the SIFT-Bow features which contain 968 visual words to represent the visual content of each image. With the similar assumptions used in [18], [20], all the nouns in the document are extracted as text words, and adjectives, verbs, and adverbs are extracted as sentiment words. All the textual words can be classified by Part-of-Speech, a word tagging function provided by Stanford NLP toolkits6 [34].

In our experiment, we set Dirichlet hyper parameters with $\beta_0^u = \beta_1^u = \beta_0^a = \beta_0^a = 0.02$ and symmetric priors with $\alpha^u = \alpha^a = 50/K$, $\eta^u = \eta^a = 0.01$. Each time, we sample the model for 200 Gibbs sampling iterations, and the first 50 iterations were ignored to remove the random initialization effect.

B. EVALUATION OF SMTM

To evaluate the proposed model, we resort to the perplexity as the measurement metric, which can be used to measure the generalization ability of a probability model, and the lower the perplexity value is, the better generalizability the topic model has. The perplexity value for a set of test documents D_t can be defined as follows:

$$perplexity(D_t) = \exp(-\frac{\sum_{d \in D_t} \log p(\mathbf{w}_d, \mathbf{v}_d, \mathbf{o}_d)}{\sum_{d \in D_t} (N_{w,d} + N_{v,d} + N_{o,d})}) \tag{19}$$

where $\mathbf{w}_d$, $\mathbf{v}_d$, $\mathbf{o}_d$ represent the textual word vector, visual word vector and sentiment vector of the test document d, respectively, and $p(\mathbf{w}_d, \mathbf{v}_d, \mathbf{o}_d) = p(\mathbf{w}_d) \cdot p(\mathbf{v}_d) \cdot p(\mathbf{o}_d|\mathbf{w}_d, \mathbf{v}_d)$.

In our experiment, the dataset I is divided into two parts respectively: 80% are randomly selected for training and the remaining 20% are used for testing. We choose the following baselines for performance comparison:

- LDA [6]: This model only depends on the text modality and take into account the tourist domain and attraction domain as whole without distinguishing them.
- Multimodal LDA (MMLDA) [12]: This model extends LDA by embedding two modality of textual and visual contents for latent topic disclosing.
- Topic-Sentiment (TS) [13]: It models topics and sentiments for a document in a unified framework but on a single domain.

Fig.3 illustrates perplexity values of different topic number and different models for the test set. Fig. 3(a) and Fig. 3(b) illustrate the perplexity of SMTM model with the different Gibbs sampling iterations in tourist domain and attraction domain, respectively. Firstly, we can see from the figure that as the number of Gibbs sampling iterations increases, the value of perplexity decreases, and it tends to stabilize after 100 iterations. Secondly, we can see from the figure that large topic number reaches low perplexity, but the perplexity is stable when the topics number is about 100. Therefore, we can choose the desired topic number $K = 100$ in our experiment. Fig. 3(c) and Fig. 3(d) show the perplexity values of different models varied with topic number in tourist domains and attraction domain respectively. From the figure, we can observe that as the number of topics varies, LDA get the highest perplexity value among all four approaches, which means the worst generalization ability. It is probably because LDA only models text modality but ignoring the other data source. The MMLDA and TS model achieve better than LDA, probably due to the additional dependencies of visual or

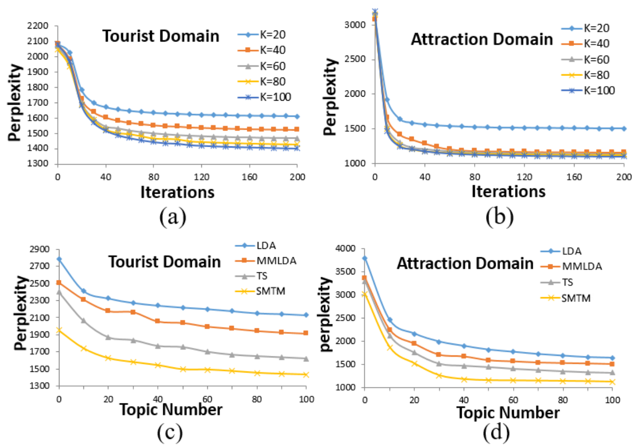

FIGURE 3. **Perplexity of different topic numbers and different models.**

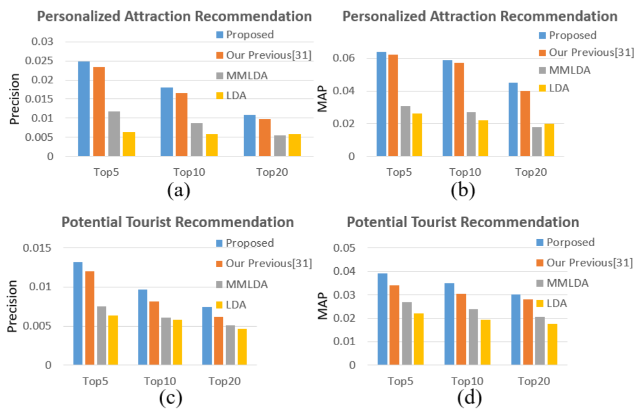

FIGURE 4. **Precision and MAP of two single platform recommendations.**

sentiment information. Our proposed SMTM model outperforms three baselines both on tourist domain and attraction domain.

C. EVALUATION OF RECOMMENDATION DERIVED FROM SMTM

The proposed SMTM can be used in many potential applications based on the mutual latent space mapping introduced in Section IV for the tourist domain and the attraction domain. In this paper, we apply the SMTM to two personalized applications, that is, the single platform recommendations and the inter-platform recommendations.

To evaluate the performance of the single platform recommendations, two test sets are created from dataset I. The first test set includes 1,261 tourists who have visited at least 15 attractions from the dataset I, and the second test set includes a total of 2,411 tourist destinations, which have been visited by at least 15 tourists, from the dataset I. Once the model is created, the formula derived in Section IV was employed to make attraction recommendation and potential tourist recommendation.

1) *Evaluation Methodology*: 24 subjects volunteered for the evaluation, and they are university students, including 12 males and 12 females. Their ages range from 22 to 30. We asked them to label the returned recommendation list. For each query, subjects need to judge whether each of returned result was relevant to the query. If more than 12 subjects thought it is relevant to the query, then this returned result can be annotated with label 1, and 0 otherwise.
2) *Evaluation Metrics*: Since Precision and Mean Average Precision (MAP) are two commonly used metrics to evaluate the performance of the information retrieval task, we can employ them to measure the performance of proposed recommendation schemes. For a given $q \in Q$, *Precision* $@n$ is defined as:

$$precission@n = \frac{\sum_{k=1}^{n} r(k)}{n} \tag{20}$$

MAP $@n$ is the mean of average precision scores over the test queries set Q and is defined as:

$$MAP@n = \frac{1}{|Q|}\sum_{q=1}^{|Q|}\frac{\sum_{k=1}^{n} Precision@qk \cdot r(qk)}{\sum_{k=1}^{n} r(qk)} \tag{21}$$

where $r(k)$ is the relevant level to the query at position k and $r(qk)$ is the relevant level at position $k$for the query q. n is the truncation level in the returned results.

We report the *Precision* $@n$ and *MAP* $@n$ for two single platform recommendations when n is selected as 5, 10 and 20. Fig. 4(a) and Fig. 4(b) show the performance of personalized attraction recommendation with different approaches. We can see from the figure that LDA performs worst as it is lack of the capability for effectively modeling the multi-modality topics and sentiments for the documents. The MMLDA achieves better performance than LDA approach since it can capture the semantic consistency between different modalities, which indicates that incorporating other modality information can improve the performance of recommendation scheme. Our previous method proposed in [30] performs better than LDA and MMLDA, which suggests that embedding the textual data, visual data, and sentiments into a unified framework can improve latent structure disclosing capabilities for social media documents and further help to achieve better recommendation. We notice that our proposed method achieves best performance, which indicates that mutual latent semantic mapping both for topic space and sentiment space in tourist domain and attraction domain can capture the essential relations in these two domains, and can further help to improve the recommendation performance significantly. Similar results can be observed from Fig. 4(c) and Fig. 4(d), which illustrates the performance of potential tourist recommendation. In summary, with combining topic-oriented and sentiment-oriented analysis, our proposed approach achieves best compared with all the baseline approaches.

The purpose of the inter-platform recommendation is to recommend single platform users with items/services on

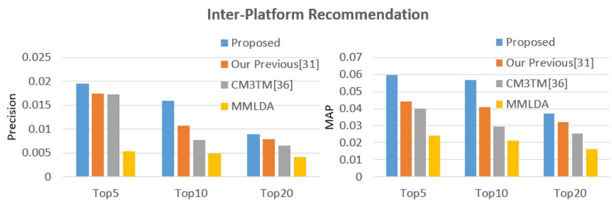

FIGURE 5. **Precision and MAP of inter-platform recommendationsf.**

another platform according to users' preferences, and it is considered as the solution to the so-called cold-start problem in the recommendation system. According to our proposed SMTM model, which models semantic topics from tourist domain and attraction domain separately, it is easy to extend the SMTM model for the inter-platform traveling recommendation with modeling the tourist domain and attraction domain from two different platforms separately. Then, we can employ the mutual topic space projection scheme introduced in Section IV to perform inter-platform recommendation. In our experiment, the tourist domain data was selected from dataset I, which is same as the tourist domain data selected in single platform validation, while the attraction domain data was selected from dataset II, which contains a total of 2,738 tourist destinations, with each visited by at least 15 tourists.

After the model is created, for an input tourist query in TripAdvisor platform, we can recommend the attractions in Trip platform to him. We followed the methodology described above to ask all the subjects to judge whether each of returned attraction is relevant to the tourist query, and to annotate the returned results with relevance to the query. *Precision @n* and *MAP @n* defined above are employed as the metric to evaluate the performance of inter-platform recommendation. Fig. 5 reports the performance comparison for inter-platform travel recommendations. We compared our proposed approach with our previous method proposed in [30], the state-of-the-art inter-platform recommendation scheme proposed in [35] and MMLDA based approach. It can be seen that our proposed method achieves best performance as it shows in the single platform validation, which indicates that including the sentiment factor in the topic modeling can help to improve the recommendation performance.

VI. CONCLUSION AND FUTURE WORK

In this paper, a SMTM is proposed to solve the travel recommendation problem by jointly modeling attraction theme, tourist preference and sentiment of the attractions. We evaluated SMTM using real-world datasets and benchmarked against the state-of-the-art topic models. Our experimental results show that SMTM is able to model the multimodal tourist topics and attraction themes with corresponding sentiments from two separated semantic spaces. The proposed SMTM can be further applied for the traveling recommendation applications, such as attraction recommendation, tourist recommendation and so on. In future work, we will extend our proposed model by incorporating the social relationships (e.g. friendship) to correlate two domains for improving the performance of the model.

REFERENCES

[1] G. Adomavicius and A. Tuzhilin, "Toward the next generation of recommender systems: A survey of the state-of-the-art and possible extensions," *IEEE Trans. Knowl. Data Eng.*, vol. 17, no. 6, pp. 734–749, Jun. 2005.
[2] J. Borràs, A. Moreno, and A. Valls, "Intelligent tourism recommender systems: A survey," *Expert Syst. Appl.*, vol. 41, no. 16, pp. 7370–7389, 2014.
[3] F. Leal, H. González-Vélez, B. Malheiro, and J. C. Burguillo, "Semantic profiling and destination recommendation based on crowd-sourced tourist reviews," in *Proc. Int. Symp. Distrib. Comput. Artif. Intell.* Cham, Switzerland: Springer, 2017, pp. 140–147.
[4] D. Yang, D. Zhang, Z. Yu, and Z. Wang, "A sentiment-enhanced personalized location recommendation system," in *Proc. 24th ACM Conf. Hypertext Social Media*, 2013, pp. 119–128.
[5] D. Blei, L. Carin, and D. Dunson, "Probabilistic topic models [A focus on graphical model design and applications to document and image analysis]," *IEEE Signal Process. Mag.*, vol. 27, no. 6, pp. 55–65, Nov. 2010.
[6] D. Blei and M. I. Jordan, "Modeling annotated data," in *Proc. 26th Annu. Int. ACM SIGIR Conf. Res. Develop. Informaion Retr.*, 2003, pp. 127–134.
[7] H. Matthew, D. Blei, and P. Cook, *Content-Based Musical Similarity Computation using the Hierarchical Dirichlet Process*. ISMIR, 2008.
[8] C. Wang and D. Blei, "Collaborative topic modeling for recommending scientific articles," in *Proc. 17th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 448–456.
[9] D. M. Blei, A. Y. Ng, and M. I. Jordan, "Latent Dirichlet allocation," *J. Mach. Learn. Res.*, vol. 3, pp. 993–1022, Mar. 2003.
[10] M. Rosen-Zvi, T. Griffiths, M. Steyvers, and P. Smyth, "The author-topic model for authors and documents," in *Proc. 20th Conf. Uncertainty Artif. Intell.* Arlington, VA, USA: AUAI Press, 2004.
[11] A. McCallum, X. Wang, and A. Corrada-Emmanuel, "Topic and role discovery in social networks with experiments on Enron and academic Email," *J. Artif. Intell. Res.*, vol. 30, pp. 249–272, Oct. 2007.
[12] K. Barnard, P. Duygulu, D. Forsyth, N. de Freitas, D. M. Blei, and M. I. Jordan, "Matching words and pictures," *J. Mach. Learn. Res.*, vol. 3, pp. 1107–1135, Mar. 2003.
[13] Q. Mei, X. Ling, M. Wondra, H. Su, and C. X. Zhai, "Topic sentiment mixture: Modeling facets and opinions in weblogs," in *Proc. 16th Int. Conf. World Wide Web*, 2007, pp. 171–180.
[14] C. Lin, Y. He, R. Everson, and S. Ruger, "Weakly supervised joint sentiment-topic detection from text," *IEEE Trans. Knowl. Data Eng.*, vol. 24, no. 6, pp. 1134–1145, Jun. 2012.
[15] F. Huang, S. Zhang, J. Zhang, and G. Yu, "Multimodal learning for topic sentiment analysis in microblogging," *Neurocomputing*, vol. 253, pp. 144–153, Aug. 2017.
[16] M. H. Alam, W.-J. Ryu, and S. K. Lee, "Joint multi-grain topic sentiment: Modeling semantic aspects for online reviews," *Inf. Sci.*, vol. 339, pp. 206–223, Apr. 2016.
[17] I. Titov and R. McDonald, "A joint model of text and aspect ratings for sentiment summarization," in *Proc. ACL*, vol. 08, 2008, pp. 308–316.
[18] Y. Fang, L. Si, N. Somasundaram, and Z. Yu, "Mining contrastive opinions on political texts using cross-perspective topic model," in *Proc. 5th ACM Int. Conf. Web Search Data Mining*, 2012, pp. 63–72.
[19] Q. Fang, C. Xu, J. Sang, M. S. Hossain, and G. Muhammad, "Word-of-mouth understanding: Entity-centric multimodal aspect-opinion mining in social media," *IEEE Trans. Multimedia*, vol. 17, no. 12, pp. 2281–2296, Dec. 2015.
[20] S. Qian, T. Zhang, C. Xu, and J. Shao, "Multi-modal event topic model for social event analysis," *IEEE Trans. Multimedia*, vol. 18, no. 2, pp. 233–246, Feb. 2016.
[21] O. Averjanova, F. Ricci, and Q. N. Nguyen, "Map-based interaction with a conversational mobile recommender system," in *Proc. 2nd Int. Conf. Mobile Ubiquitous Comput., Syst., Services Technol.*, Sep./Oct. 2008, pp. 212–218.

[22] B. De Carolis, N. Novielli, V. L. Plantamura, and E. Gentile, ''Generating comparative descriptions of places of interest in the tourism domain,'' in *Proc. 3rd ACM Conf. Recommender Syst.*, 2009, pp. 277–280.

[23] F. Ricci and Q. N. Nguyen, ''Mobyrek: A conversational recommender system for on-the-move travelers,'' *Destination Recommendation Systems: Behavioral Foundations and Applications*. Wallingford, U.K.: CABI, 2006, pp. 281–294.

[24] D. J. Crandall, L. Backstrom, D. Huttenlocher, and J. Kleinberg, ''Mapping the world's photos,'' in *Proc. 18th Int. Conf. World Wide Web*, 2009, pp. 761–770.

[25] Y. Arase, X. Xie, T. Hara, and S. Nishio, ''Mining people's trips from large scale geo-tagged photos,'' in *Proc. 18th ACM Int. Conf. Multimedia*, 2010, pp. 133–142.

[26] H. Wang, Y. Fu, Q. Wang, H. Yin, C. Du, and H. Xiong, ''A location-sentiment-aware recommender system for both home-town and out-of-town users,'' in *Proc. 23rd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2017, pp. 1135–1143.

[27] C. Huang, Q. Wang, D. Yang, and F. Xu, ''Topic mining of tourist attractions based on a seasonal context aware LDA model,'' *Intell. Data Anal.*, vol. 22, no. 2, pp. 383–405, 2018.

[28] B.-K. Bao, C. Xu, W. Min, and M. S. Hossain, ''Cross-platform emerging topic detection and elaboration from multimedia streams,'' *ACM Trans. Multimedia Comput., Commun., Appl.*, vol. 11, no. 4, p. 54, 2015.

[29] J. Shen, C. Deng, and X. Gao, ''Attraction recommendation: Towards personalized tourism via collective intelligence,'' *Neurocomputing*, vol. 173, pp. 789–798, Jan. 2016.

[30] J. Wang, B.-K. Bao, and C. Xu, ''Sentiment-aware multi-modal recommendation on tourist attractions,'' in *Proc. Int. Conf. Multimedia Modeling*. Cham, Switzerland: Springer, 2019, pp. 3–16.

[31] C. Andrieu, N. De Freitas, A. Doucet, and M. I. Jordan, ''An introduction to MCMC for machine learning,'' *Mach. Learn.*, vol. 50, no. 1, pp. 5–43, Jan. 2003.

[32] S. Haccianella, ''3.0: An enhanced lexical resource for sentiment analysis and opinion mining,'' in *Proc. 7th Conf. Int. Lang. Resour. Eval.*, 2010.

[33] D. Olszewski, ''Employing Kullback-leibler divergence and latent Dirichlet allocation for fraud detection in telecommunications,'' *Intell. Data Anal.*, vol. 16, no. 3, pp. 467–485, 2012.

[34] [Online]. Available: http://nlp.stanford.edu/software/index.shtml

[35] W. Min, B.-K. Bao, C. Xu, and M. S. Hossain, ''Cross-platform multimodal topic modeling for personalized inter-platform recommendation,'' *IEEE Trans. Multimedia*, vol. 17, no. 10, pp. 1787–1801, Oct. 2015.