

# ORec: An Opinion-Based Point-of-Interest Recommendation Framework

Jia-Dong Zhang  
Dept. of Computer Science  
City University of Hong Kong  
jzhang26-  
c@my.cityu.edu.hk

Chi-Yin Chow  
Dept. of Computer Science  
City University of Hong Kong  
chiychow@cityu.edu.hk

Yu Zheng  
Microsoft Research  
Beijing, China  
yuzheng@microsoft.com

## ABSTRACT

As location-based social networks (LBSNs) rapidly grow, it is a timely topic to study how to recommend users with interesting locations, known as *points-of-interest* (POIs). Most existing POI recommendation techniques only employ the check-in data of users in LBSNs to learn their preferences on POIs by assuming a user's check-in frequency to a POI explicitly reflects the level of her preference on the POI. However, in reality users usually visit POIs only once, so the users' check-ins may not be sufficient to derive their preferences using their check-in frequencies only. Actually, the preferences of users are exactly implied in their opinions in text-based tips commenting on POIs. In this paper, we propose an opinion-based POI recommendation framework called ORec to take full advantage of the user opinions on POIs expressed as tips. In ORec, there are two main challenges: (i) detecting the polarities of tips (positive, neutral or negative), and (ii) integrating them with check-in data including social links between users and geographical information of POIs. To address these two challenges, (1) we develop a supervised aspect-dependent approach to detect the polarity of a tip, and (2) we devise a method to fuse *tip polarities* with *social links* and *geographical information* into a unified POI recommendation framework. Finally, we conduct a comprehensive performance evaluation for ORec using two large-scale real data sets collected from Foursquare and Yelp. Experimental results show that ORec achieves significantly superior polarity detection and POI recommendation accuracy compared to other state-of-the-art polarity detection and POI recommendation techniques.

## Categories and Subject Descriptors

H.3.3 [Information Storage and Retrieval]: Information Search and Retrieval—*Information Filtering*

## Keywords

Location-based social networks; point-of-interest recommendations; opinion mining; polarity detection; fusion

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from Permissions@acm.org.

CIKM '15, October 19 - 23, 2015, Melbourne, VIC, Australia  
© 2015 ACM. ISBN 978-1-4503-3794-6/15/10 ...\$15.00  
DOI: <http://dx.doi.org/10.1145/2806416.2806516>.

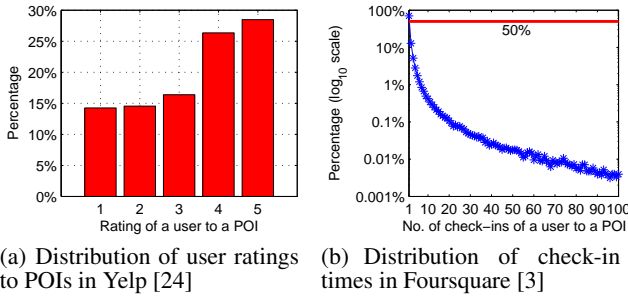
## 1. INTRODUCTION

In location-based social networks (LBSNs) as shown in Figure 1, users establish social links with others, check in some interesting locations, known as *points-of-interest* (POIs), e.g., restaurants, stores and museums, and post tips to express their opinions about various aspects of POIs, e.g., atmosphere, price and service. With the rapid growth of LBSNs, e.g., Foursquare and Yelp, it is prevalent and important to recommend users with their preferred POIs. POI recommendations not only help users explore new places and enrich their life but also enable companies to launch advertisements to potential customers and improve business profits.



Figure 1: A location-based social network (LBSN)

To make POI recommendations, most recent methods [11, 23, 26, 29, 30, 31, 33] only employ users' historical check-in data to learn their preferences for POIs, i.e., these methods assume the check-in frequency of a user to a POI explicitly reflects the user's preference level for the POI. However, the check-in data of users may not be sufficient to indicate their preferences, especially when a user checks in a POI only once, since the user may like the POI very much or conversely. For example, by calculating the distribution of ratings in the five-star scale given by users to POIs from the Yelp Challenge data set [24] as shown in Figure 2(a), it is observed that the attitude of users towards POIs is not always positive (i.e., about 55% ratings > 3) and it is highly possibly negative (i.e., about 30% ratings < 3). One may argue that it is intuitive to assume that a user should have positive preference for a POI with more than one check-in. Unfortunately, in LBSNs users usually check in POIs only once. For instance, Figure 2(b) depicts the distribution of the number of check-ins of a user to a POI in a large-scale check-in data set [3] from Foursquare, in which more than 50% places have been checked in only once by the same user. As a result, merely utilizing check-in data as preference models usually introduces biases to the user preference level.



**Figure 2: (a) Only about 55% ratings are positive ( $> 3$ ); (b) Above 50% places have been checked in only once by the same user.**

In fact, users’ preferences are exactly reflected in their tips commenting on POIs, in which they express their opinions about various aspects of POIs. For example, consider a typical tip written by a user regarding a restaurant: “I went here last night. I loved the place with good atmosphere. The taste is not bad. The food has high quality but with a little high price.” In this tip, the user has expressed different opinions to a variety of aspects of the POI, like “good atmosphere”, “not bad taste”, “high quality” and “high price”. Therefore, the user’s actual preference for the POI can be determined by detecting the tip’s overall sentiment polarity: positive, neutral or negative, based on the opinions expressed in the tip. To the best of our knowledge, only the study [22] recommends POIs to users through utilizing the polarities of tips determined by a *unsupervised aspect-independent* method, in which the sentiment value of an opinion word is obtained by looking it up in sentiment lexicons (e.g., SentiWordNet) and the polarity of a tip is derived from the sum of sentiment values of all opinion words in the tip. Nevertheless, the unsupervised aspect-independent method cannot accurately detect the polarities of tips because of two main reasons. (1) The method uses *the absolute sentiment value of an opinion word independently of its modifying aspect* instead of using *the sentiment orientation of the opinion word depending on the modifying aspect*. As an example, for the two opinion phrases of “high quality” and “high price” in the aforementioned tip, the opinion word “high” in the two phrases has the same sentiment value by querying a lexicon, but it has completely different sentiment orientations: “high” modifying the aspect “quality” is positive while “high” modifying the aspect “price” is negative. (2) The method utilizes the simple summation function to combine different sentiment values rather than learning a sophisticated combination function from data.

In this paper, we propose a *supervised aspect-dependent Opinion-based POI Recommendation framework (ORec)* that derives the sentiment orientations of opinion words depending on aspects in tips and leverages them to predict the polarities of tips for enhancing the quality of POI recommendations. In ORec, there are two main challenges. The first challenge is to accurately detect the polarities of tips. Naturally, the polarity detection problem can be considered as the classical classification problem: training classification models on tips labeled with polarities and using the classification models to predict the polarity of a new tip. A straightforward solution is to apply the text classification methods that view a text as a bag of words and transform the text into a high-dimensional vector, each dimension of which corresponds to a word (e.g., sentiment words or frequent words [12, 17]) in a vocabulary. Unfortunately, these methods result in highly sparse word vectors and low polarity detection accuracy, because the vocabulary is usually very large compared to the text. To this end, we develop an approach

that only considers *the aspects of POIs in their tips* rather than *raw words in a large vocabulary* as dimensions of vectors to reduce the number of dimensions for tips. Further, to decrease the dimension of aspect vectors and reduce the ambiguity between aspects, the proposed approach groups the aspects into clusters. Finally, a classification model is trained on the cluster-based vector representations of tips and exploited to predict the polarity of a new tip.

The other challenge is to combine the tip polarities with traditional check-in data (including social links and geographical information) to improve the POI recommendation quality. In terms of the current studies [23, 27], the check-in behaviors of users to POIs are significantly influenced by *others with social links* and *nearby visited POIs*, since users with social links are more likely to share common interests and physical interactions are required for users to check in POIs. For instance, friends often go to some places like restaurants together or a user may visit POIs recommended by friends. Moreover, users tend to visit POIs close to their homes or offices and also may be interested in exploring the nearby places of their visited POIs. How can we effectively fuse tip polarities with social links and geographical information into a unified recommendation framework? This problem is still open. For this purpose, we design a method to integrate *tip polarities of users regarding POIs* with *social links among users* and *geographical information of POIs* into a unified POI recommendation framework.

In general, our contributions can be summarized below.

- To the best of our knowledge, this is the first study to propose a POI recommendation framework based on users’ opinions to the aspects of POIs in tips. In this framework, we develop a supervised aspect-dependent approach to accurately detect the polarities of tips. (Section 4)
- To improve the quality of POI recommendations, we design a method to fuse the influences of three different types of information in LBSNs including *tip polarities*, *social links* and *geographical information*. (Section 5)
- We conduct extensive experiments to evaluate the performance of ORec using two large-scale real data sets from Foursquare and Yelp. Experimental results show that: (1) ORec outperforms the classical text classification methods studied in [12, 17] and the unsupervised aspect-independent method [22] in terms of *polarity detection accuracy*. (2) ORec achieves better *recommendation accuracy* than state-of-the-art POI recommendation techniques. (Sections 6 and 7)

**Organization of this paper:** Section 2 highlights related work. Section 3 defines the research problems and presents an overview of the ORec framework. We present the proposed supervised aspect-dependent polarity detection approach in Section 4 and the developed POI recommendation method in Section 5. In Sections 6 and 7, we describe our experiment settings and analyze the performance of ORec, respectively. Section 8 concludes this paper.

## 2. RELATED WORK

**POI recommendations using check-in data only.** Most current studies (e.g., [1, 3, 11, 13, 14, 23, 26, 27, 29, 30, 31, 33]) only utilize the check-in data to learn the preference of users on POIs. Specifically, they employ the social links between users to derive the user similarity as an input of collaborative filtering techniques, based on the fact that friends are more likely to share common interests. They also exploit the geographical information of POIs to estimate the distribution of distances between visited POIs, due to the fact that if a POI is closer to the POIs visited by a user, it is more

likely to be visited by the user. For instance, Liu et al. [14] used the instance-level and region-level geographical neighborhood characteristics, other researchers in [13, 23] fitted the distance among visited POIs as a power-law distribution for all users, and the authors of [27, 28, 29, 32] personalized the distance distribution for each user. However, these methods assume that the check-in frequencies of users to POIs directly reflects the preference levels, which may not be true in reality. For example, when a user checks in a POI only once, the user may like the POI very much or conversely.

**POI recommendations using textual information.** There are a few works that use the textual information of users commenting on POIs for location recommendations. For example, the work [10] employs the tags of POIs to provide interpretable representations for latent topics extracted from check-in data, while the study [35] represents each POI as a vector of words in its tips and utilizes the word vector to derive the similarity between POIs as a regularized term of tensor models over check-in data. More sophisticatedly, the method in [25] applies the well-known latent Dirichlet allocation (LDA) over tags of POIs or posts of users to mine the topic profiles for users and POIs, in which each topic is a distribution over words; then the topic profiles and topic distributions are exploited to determine the preference score of users to POIs. However, these studies do not take into account the opinion or sentiment of users on POIs. To the best of our knowledge, only the study [22] extracts opinion words from tips and employ them to determine the polarity of the tips based on the unsupervised aspect-independent method.

**Polarity detection of texts.** Existing polarity detection techniques generally fall into two categories: *supervised methods* and *unsupervised methods*. (1) *Supervised methods*. This category learns classification models from labeled texts and uses them to classify a new text into positive or negative based on supervised machine learning methods, such as support vector machines [12, 17], decision trees and naive Bayes [17, 21]. Nevertheless, the supervised methods often regard a feature as a word in a large vocabulary which causes highly sparse word vectors in texts and low polarity detection accuracy. (2) *Unsupervised methods*. The unsupervised methods [17, 22] search sentiment values of opinion words in a text based on sentiment lexicons and aggregate these values to determine the text’s polarity. In particular, the works [6, 7] make use of the emoticon sentiment lexicon to increase the accuracy of polarity classification for social media services such as Twitter and Facebook. However, these unsupervised polarity detection methods consider the sentiment values of opinion words independently of their modifying aspects and apply simple aggregation functions to combine different sentiment values. As discussed in Section 1, they cannot accurately detect the polarity of tips, since the actual sentiment orientation of an opinion word strongly relies on its modifying aspect and a more sophisticated combination function should be learned from data.

**Aspect-based opinion mining based on aspect extraction and clustering.** There are also plenty of studies that mine the opinions from text at aspects. These studies can be classified into two main categories: *topic models* and *language rules*. (1) *Topic models*. A topic model is a generative model for texts: texts are mixtures of topics and each topic is a probability distribution of words. First, topic models (e.g., probabilistic latent semantic analysis and LDA) can be applied to aspect extraction by regarding each topic as an aspect, in which different words expressing the same aspect can be automatically grouped together [15, 34]. Further, to associate aspects with opinions, the topic models can be extended by adding a sentiment layer to detect aspect and opinion words simultaneously from texts [2, 9, 16]. However, topic models leverage *word co-occurrences among texts* rather than *word semantic meanings*

to identify topics and word distributions; they are only able to find rough aspects instead of fine-grained aspects. (2) *Language rules*. The language-rule-based methods utilize the grammatical relations between aspects and opinion words to induce extraction rules. Most works [5, 18, 19] employ the modifying relation of opinion words and aspects to extract aspects and associate them with opinion words. Although the language-rule-based methods can discover fine-grained aspects, they need an additional step to group these fine-grained aspects into clusters, usually based on the synonym relation of aspects. In this paper, we utilize not only the synonym relation, but also the hypernymy and meronymy relations among aspects to cluster them into groups.

### 3. PROBLEM AND OVERVIEW

We first define basic concepts for this paper.

- **Aspect.** An aspect is an attribute or feature of a POI, e.g., “atmosphere”, “price” and “quality” for a restaurant.
- **Opinion word.** An opinion word usually indicates a sentiment orientation with a sentiment value from lexicons, e.g., “good”, “bad” and “high”. Note that the sentiment orientation of “high” is not that intuitive because it depends on the aspect modified by “high”.
- **Opinion phrase.** An opinion phrase  $\langle a, o \rangle$  is a pair of aspect  $a$  and opinion word  $o$ , where the opinion word is used to express the sentiment orientation towards the aspect, e.g.,  $\langle atmosphere, good \rangle$  and  $\langle price, high \rangle$ . Note that  $o$  is also used to indicate the sentiment value of an opinion word in terms of lexicons for simplicity and the specific meaning of  $o$  (i.e., “opinion word” or “sentiment value”) can be easily determined based on its context.
- **Tip.** A tip is a sequence of words describing a user’s opinions on various aspects of a POI and consists of a set of opinion phrases, i.e.,  $tip = \{ \langle a_i, o_i \rangle \}$ .
- **Polarity of tip.** The polarity of a user’s tip is the overall attitude of the user towards a POI and usually is classified into three cases: positive, neutral and negative. The set of polarities are denoted as  $P = \{ positive, neutral, negative \}$ .

**Problem definitions.** The two problems addressed in this paper are defined as follows:

- **Problem 1: Polarity detection.** Given a training set of tips  $D = \{ \langle tip^j, p^j \rangle \}$ , where each  $tip^j$  is labeled with a class, i.e., polarity  $p^j \in P$ , the goal is to predict the polarity  $p \in P$  of a new  $tip$ . (Section 4)
- **Problem 2: Opinion-based POI recommendation.** Given all users’ *tip polarities* and a certain user  $u$ ’s friends with *social links*  $F(u)$  and *geographical information* of check-in POIs  $L_u = \{ l_1, l_2, \dots, l_n \}$ , the goal is to predict the preference score  $s_{u,l}$  of user  $u$  regarding a new POI  $l$  (i.e.,  $l \notin L_u$ ) and then return the top- $m$  POIs with the highest score  $s_{u,l}$  for  $u$ . (Section 5)

**Overview of ORec.** Figure 3 demonstrates the overview of ORec, including two major parts: *polarity detection* and *POI recommendations*. (1) The *polarity detection* part generates the polarities of tips for the *POI recommendations* part. This part has three steps: (a) extracting opinion phrases (pairs of aspects and opinions) from tips, (b) grouping aspects into clusters and aggregating opinions of a tip into a sentiment vector in the cluster space,

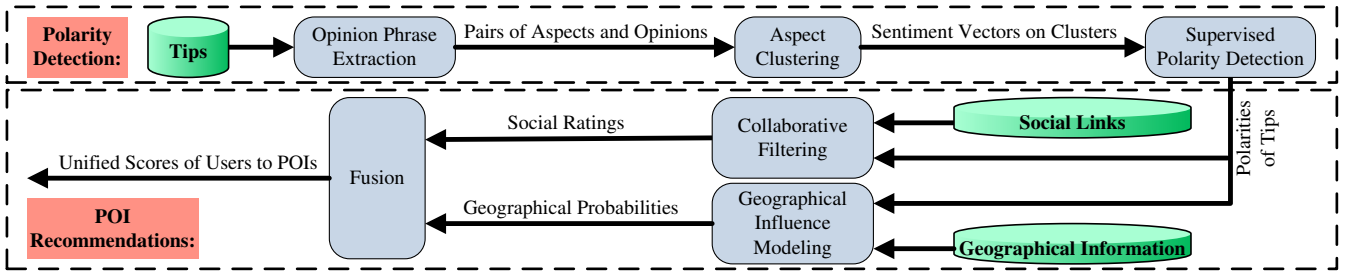


Figure 3: The overview of ORec

and (c) training a classification model in tips labeled with polarities and predicting the polarity of any new tips (without polarity labels). (2) The *POI recommendations* part exploits the tip polarity with social links and geographical information to estimate the score of a user to a new POI in order to recommend POIs for the user. This part also has three steps: (a) using collaborative filtering techniques to combine tip polarities with social links to estimate the social rating of users to new POIs, (b) modeling geographical influence by integrating tip polarities and geographical information of POIs to predict the geographical probability of users visiting new POIs, and (c) fusing the social rating and geographical probability into a unified score for POI recommendations.

## 4. POLARITY DETECTION

In general, our proposed supervised aspect-dependent approach has **three main phases for polarity detection**: (1) Aspects are extracted through nearby opinion words based on sentiment lexicons; accordingly the aspect vectors store the sentiment values of these opinion words. (2) The obtained aspects are grouped into adaptive  $K$  clusters based on the dissimilarities of aspects to decrease the dimension of aspect vectors and reduce the ambiguity between aspects for tips; by this end, each tip can be represented as a  $K$ -dimensional vector, each dimension for a cluster. (3) A classification model is trained on the cluster-based vector representations of tips with labeled polarities and employed to predict the polarity of any new tips.

### 4.1 Phase 1: Opinion Phrase Extraction

We can acquire opinion phrases consisting of pairs of aspects and opinion words through two steps.

**Step 1: Tip preprocessing.** Using the Stanford natural language parsers [20], a tip is firstly tokenized into words and split into a set of sentences. Then, each word in a sentence is identified with the part-of-speech (e.g., “good” is an adjective, “price” is a noun, “love” is a verb, etc.) and is lemmatized to a base form in order to reduce its inflectional and derivational forms (e.g., “loves”, “loving” and “loved” are lemmatized to the same base form “love”). Finally, the grammatical dependencies between words in a sentence are determined by parsing the grammatical structure of the sentence. For example, the sentence of “the taste is not bad” contains a number of grammatical dependencies, such as “nsubj(bad, taste)”, “neg(bad, not)”, and so on, in which “nsubj” and “neg” are a certain grammatical dependency defined in [20].

**Step 2: Opinion phrase extraction.** One may consider some kinds of grammatical dependencies as opinion phrases, e.g., “nsubj(bad, taste)” denoting an opinion phrase  $\langle taste, bad \rangle$ . However, “nsubj” is unnecessary to represent an opinion phrase. For instance, the sentence of “I loved the place with good atmosphere” also contains “nsubj(love, I)” that is not an opinion phrase because of the lack of an aspect. Thus, it is hard to select a set of gram-

matical dependencies as opinion phrases in advance. To this end, we design a method to discover opinion phrases from grammatical dependencies of a sentence based on sentiment lexicons.

It is observed that an opinion phrase always includes an opinion word with a sentiment value by looking it up in lexicons, e.g., SentiWordNet. Thus, opinion words can be used as the indicators of opinion phrases. Further, given a grammatical dependency with an opinion word, it is regarded as an opinion phrase if the other word in the grammatical dependency is a noun that is required by an aspect. Finally, given an opinion phrase  $\langle a, o \rangle$  in a sentence, if the negation dependency “neg( $o$ , not)” exists in the same sentence, the opinion phrase  $\langle a, o \rangle$  will be reversed into opinion phrase  $\langle a, not\ o \rangle$  in order to truly reflect the opinion of a user to an aspect of a POI. As an example, considering the sentence of “the taste is not bad” again, it contains the opinion phrase  $\langle taste, bad \rangle$  with the negation dependency “neg(bad, not)”; subsequently  $\langle taste, bad \rangle$  is reversed into  $\langle taste, not\ bad \rangle$ .

In sum, a tip can be represented as a set of opinion phrases consisting of aspect  $a_i$  and opinion word  $o_i$ , i.e.,  $tip = \{\langle a_i, o_i \rangle\}$ , embedded with the negation information if any. Moreover, according to sentiment lexicons, we can obtain a sentiment value for each opinion word  $o_i$ . Hereafter,  $o_i$  also indicates its corresponding sentiment value for simplicity.

### 4.2 Phase 2: Aspect Clustering

In tips, users may write different words to represent the same aspect of POIs, e.g., “ambiance”, “ambiance” or “environment” instead of “atmosphere”. Thus, aspects should be grouped into clusters based on their dissimilarities in order to decrease the dimension of vectors of tips represented in the aspect space and increase the correlation among tips that is helpful for training classification models. The grouping process includes three key steps.

**Step 1: Aspect distance calculation.** To cluster aspects, the dissimilarities between aspects need to be collected in advance. The dissimilarity can be any valid distance metric, most commonly Euclidean distance, Manhattan distance or Minkowski distance. However, these distance metrics are defined over the Cartesian coordinates of clustered data but the aspects of POIs, i.e., linguistic words, do not have the intrinsic Cartesian coordinates. Hence, we define the aspect dissimilarity or distance based on the semantic relationships of words, including synonymy linking words that have similar meanings (e.g., “atmosphere” and “environment”), hypernymy referring to a hierarchical relationship between words (e.g., “food” is a hypernym of “sandwich”), and meronymy referring to a part/whole relationship (e.g., “paper” is a meronym of “book”). Specifically, based on the dictionary WordNet, where words are linked via semantic relationships to form a semantic network, the distance between aspects is considered as the length of the shortest path between corresponding words in the semantic network.

**Step 2: Grouping aspects into clusters.** Give a set of aspects from all tips and their distance, here we aim to group the aspects



into clusters, so that aspects within a cluster have small distance in comparison to one another but are very far from the aspects in other clusters. To achieve this task, one may employ the popular clustering techniques, like  $K$ -means or  $K$ -medoids. Nonetheless, they are not adaptable to the aspect clustering problem, since the mean of an aspect cluster is not defined for  $K$ -means, and normally we cannot group aspects into a small number of clusters due to their large dissimilarities from each other, so  $K$  is the same order-of-magnitude with the total number of aspects, which leads to high computation cost for  $K$ -medoids.

Therefore, we apply the farthest-point clustering algorithm [4] due to its efficiency and without the requirement of computing mean values for clusters. The primitive farthest-point clustering algorithm discovers a predefined number  $K$  of clusters. We utilize this method to find **an adaptive number  $K$  of clusters**, in which  $K$  is determined by a threshold of distance that is the allowed maximum distance from an aspect to its cluster center. Given a set of aspects  $A = \{a_i\}$  from all tips in Section 4.1, distance matrix *distance* between aspects in Step 1, and maximum distance threshold  $d_{max}$ , the clustering process is described below:

1. The algorithm initially selects an arbitrary aspect in  $A$  as the first cluster center and adds it to the center set  $C$ .
2. In the  $k$ -th iteration:
  - For each aspect  $a_i \in (A - C)$ , the algorithm computes the minimum distance from  $a_i$  to the center set  $C$ :
$$dist_{min}(a_i, C) = \min_{c \in C} distance(a_i, c). \quad (1)$$
  - The farthest-point  $a_k^*$  is determined by:
$$a_k^* = \arg \max_{a_i \in (A - C)} dist_{min}(a_i, C). \quad (2)$$
  - If  $dist_{min}(a_k^*, C) > d_{max}$ , the algorithm considers  $a_k^*$  as a new cluster center, adds it into  $C$  and continues the iteration process; otherwise, it terminates.
3. Aspect  $a_i \in A$  is assigned to its nearest cluster  $c_k \in C$ .

**Step 3: Opinion aggregation.** After the aspects have been grouped into  $K$  clusters, a *tip* =  $\{\langle a_i, o_i \rangle\}$  (i.e., a set of opinion phrases consisting of aspect  $a_i$  and sentiment value  $o_i$ ) can be transformed into a vector  $tip\langle o_{c_1}, o_{c_2}, \dots, o_{c_K} \rangle$  represented in the cluster space  $C$ , in which aggregated sentiment value  $o_{c_k}$  is the average of  $o_i$  for all  $a_i$  that are assigned to cluster center  $c_k \in C$  ( $k = 1, 2, \dots, K$ ), i.e.,

$$o_{c_k} = average_{a_i \text{ assigned to } c_k} (o_i). \quad (3)$$

Note that: for a tip, if there is no aspect  $a_i$  assigned to a certain cluster center  $c_k$ , the aggregated sentiment value  $o_{c_k}$  is missing. Actually, this is the most common case, because the number of aspects described in a short tip is far less than the large number  $K$  of clusters generated from all tips.

**Example.** Figure 4 depicts an example to illustrate the opinion phrase extraction phase and aspect clustering phase. Figure 4(a) gives a set of clusters of aspects from all tips, in which the aspect ‘‘atmosphere’’ has been assigned to cluster  $c_1$ , ‘‘taste’’ to  $c_2$ , ‘‘quality’’ to  $c_3$ , ..., and ‘‘price’’ to  $c_{100}$ . Given a new tip shown in Figure 4(b), four opinion phrases are extracted  $tip = \{\langle atmosphere, good \rangle, \langle taste, not\ bad \rangle, \langle quality, high \rangle, \langle price, high \rangle\}$ . Then, *tip* is aggregated into a cluster-based vector  $tip\langle good, not\ bad, high, ?, \dots, ?, high \rangle$  in Figure 4(c). **Note that:** The tip’s sentiment values on clusters  $o_{c_1}$ ,  $o_{c_2}$ ,  $o_{c_3}$  and  $o_{c_{100}}$  can

$c_1$	$c_2$	$c_3$	...	$c_{100}$
atmosphere	taste	quality	...	price
ambiance	savor	trait	...	charge
ambience	feel	particularity	...	cost
environment	sense	idiosyncrasy	...	fee

(a) Clusters of aspects from all tips

**A tip:** I went here last night. I loved the place with **good atmosphere**. The **taste** is **not bad**. The food has **high quality** but with a little **high price**.  
 $\Rightarrow$   
 $tip = \{\langle atmosphere, good \rangle, \langle taste, not\ bad \rangle, \langle quality, high \rangle, \langle price, high \rangle\}$

(b) Opinion phrases extracted from a tip

$tip\langle o_{c_1}, o_{c_2}, o_{c_3}, \dots, o_{c_{100}} \rangle = \langle good, not\ bad, high, ?, \dots, ?, high \rangle$

(c) Cluster-based vector representation of the tip

**Figure 4: Opinion aggregation on clusters**

be determined based on opinion words ‘‘good’’, ‘‘bad’’ and ‘‘high’’ in a lexicon, but the sentiment values on clusters  $o_{c_4}, \dots, o_{c_{99}}$  are missing.

Hereafter, we reduce  $tip\langle o_{c_1}, o_{c_2}, \dots, o_{c_K} \rangle$  to  $tip\langle o_1, o_2, \dots, o_K \rangle$  for the sake of presentation, where each of the subscripts  $(1, 2, \dots, K)$  indicates its corresponding cluster.

### 4.3 Phase 3: Supervised Polarity Detection

The goal of the supervised polarity detection approach is to predict polarity  $p \in P = \{positive, neutral, negative\}$  for a new *tip*, given a training set of tips  $D = \{\langle tip^j, p^j \rangle\}$ , where each  $tip^j$  is labeled with a class  $p^j \in P$ . By the opinion phrase extraction and aspect clustering phases described in Sections 4.1 and 4.2, respectively, a tip has been represented as a vector of sentiment values regarding each cluster, i.e.,  $tip^j\langle o_1^j, o_2^j, \dots, o_K^j \rangle$  in order to facilitate polarity detection. As aforementioned,  $tip^j\langle o_1^j, o_2^j, \dots, o_K^j \rangle$  contains a large number of missing values since a tip is often short with only a few sentences. Thus, it is required to replace the missing values with the most probable value based on statistics to apply most classification methods, such as decision trees and support vector machines. Nevertheless, filling in a large number of values will result in severely biased classification models. Hence, we employ another important classification approach, i.e., naive Bayes (NB) [8] that can intrinsically handle missing values without any replacement.

The NB classifier is a special form of Bayesian networks and relies on the class conditional independence assumption that the predictive attributes (clusters) are conditionally independent given the class (polarity). Despite the simplifying assumption, experimental results on real-world data have repeatedly shown that the NB classifier is competitive with much more sophisticated classification methods in terms of classification accuracy. Further, the assumption leads to very efficient algorithms for both *predicting* and *training* that are required for processing large-scale Web data. We will discuss the two main steps for the NB classifier.

**Step 1: Predicting of naive Bayesian classifier.** Specifically, the NB classifier predicts that a new *tip* belongs to the polarity  $p^*$  having the highest posterior probability,

$$p^* = \arg \max_{p \in P} \Pr(p|tip), \quad (4)$$

where  $P = \{positive, neutral, negative\}$ . And we have

$$\Pr(p|tip) = \Pr(p) \Pr(tip|p) / \Pr(tip), \quad (5)$$

in which  $\Pr(tip)$  is constant for all polarity  $p \in P$ , so only  $\Pr(p)$  and  $\Pr(tip|p)$  need to be computed based on the given training set of tips  $D = \{\langle tip^j, p^j \rangle\}$ .

**Step 2: Training of naive Bayesian classifier.**  $\Pr(p)$  is the class prior probability and is estimated by

$$\Pr(p) = |\{p^j | p^j = p \wedge p^j \in D\}| / |D|, \quad (6)$$

where the numerator is the number of training tips with polarity  $p^j = p$  in  $D$  and the denominator is the total number of training tips in  $D$ .

$\Pr(\text{tip}|p)$  is first factorized based on the class conditional independence assumption that the values of the attributes in  $\text{tip}(o_1, o_2, \dots, o_K)$  are conditionally independent of one another given the class label  $p$ :

$$\Pr(\text{tip}|p) = \prod_{k=1}^K \Pr(o_k|p). \quad (7)$$

$\Pr(o_k|p)$  is the class conditional probability density function for the  $k$ -th cluster and is computed by kernel density estimation with the most general Gaussian kernels using training set  $D = \{\langle \text{tip}^j, p^j \rangle\} = \{\langle o_1^j, o_2^j, \dots, o_K^j, p^j \rangle\}$ :

$$\Pr(o_k|p) = \frac{1}{N} \sum_{\{o_k^j | p^j = p\}} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(o_k - o_k^j)^2}{2\sigma^2}}, \quad (8)$$

where  $o_k^j$  ranges over the sentiment values of the  $k$ -th cluster in training tips with polarity  $p^j = p$ ,  $N$  is the number of  $o_k^j$  with  $p^j = p$ , i.e.,  $N = |\{o_k^j | p^j = p\}|$ , and  $\sigma = 1/\sqrt{N}$ .

## 5. OPINION-BASED RECOMMENDATION

After the polarity detection part in Section 4, each tip is associated with a polarity, i.e., positive, neutral or negative. In this section, the polarities of tips are integrated with other important information sources of LBSNs (i.e., the social links between users and geographical information of POIs) to make better POI recommendations for users.

### 5.1 Fusion with Social Links

To use polarities of tips for POI recommendations, they need to be mapped into values that indicate users' ratings to **visited** POIs. As usual, we adopt a five-star rating scale widely used in online e-business Web sites, e.g., Amazon and Yelp. Formally, given a user with polarity  $p_{u,l} \in P = \{\text{positive}, \text{neutral}, \text{negative}\}$  from a tip regarding POI  $l$ , the rating  $r_{u,l}$  of user  $u$  to **visited** POI  $l$  is defined as follows:

$$r_{u,l} = \begin{cases} 5, & p_{u,l} = \text{positive}; \\ 3, & p_{u,l} = \text{neutral}; \\ 1, & p_{u,l} = \text{negative}. \end{cases} \quad (9)$$

Therefore, we can collect a user-POI rating matrix  $R = \{r_{u,l}\}$  for all tips of users to POIs. Note that most entries in  $R$  are unknown, since users have only visited a very small proportion of POIs and accordingly tips only contain a little part of interaction between users and POIs.

Here the key task is to estimate the social rating of a user to an **unvisited** POI via existing ratings in  $R$ . To achieve this task, we apply the collaborative filtering technique on  $R$ . Formally, given a certain entry  $r_{u,l} = 0$  (i.e., user  $u$  has not visited POI  $l$ ), the social rating  $\hat{r}_{u,l}$  of user  $u$  to **unvisited** POI  $l$  can be predicted by

$$\hat{r}_{u,l} = \frac{\sum_{u'} \text{sim}(u, u') \cdot r_{u',l}}{\sum_{u'} \text{sim}(u, u')}, \quad (10)$$

where  $r_{u',l}$  denotes the known rating mapped from the tip polarity of user  $u'$  concerning POI  $l$  using Equation (9), and  $\text{sim}(u, u')$  is

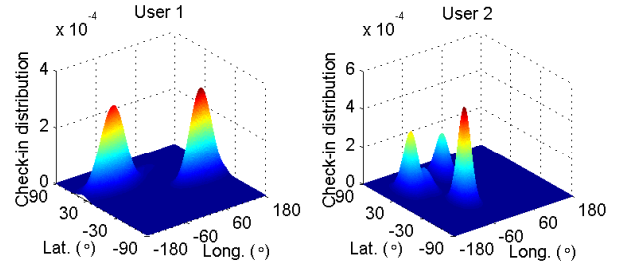


Figure 5: Personal check-in distributions

the similarity measure between users  $u$  and  $u'$ . As friends often share more common interests,  $\text{sim}(u, u')$  is usually defined by social links, e.g.,

$$\text{sim}(u, u') = \begin{cases} 1, & u \text{ and } u' \text{ have the social link;} \\ 0, & \text{Otherwise.} \end{cases} \quad (11)$$

Note that other similarity functions using social links or/and location histories can also be applied here.

### 5.2 Fusion with Geographical Information

The geographical information (geographic coordinates) of spatial POIs included in check-in data plays a significant role in users' check-in behaviors, since POIs are distinct from other non-spatial items, such as books, music and movies in conventional recommendation systems and physical interactions are required for users to visit POIs. For example, users tend to visit POIs close to their homes or offices and also may be interested in exploring the nearby places of their visited POIs. To this end, we exploit the geographical information of POIs to learn a geographical preference model for each user that can be used to predict the geographical probability of the user visiting a new POI.

Concretely, we model the personal check-in distribution for a user using her historical check-in POIs based on kernel density estimation with Gaussian kernels as well, in which however, each check-in of the user to a POI is weighed with the polarity of the tip written by the user to the POI. Formally, given user  $u$ 's set of visited POIs  $L_u = \{l_1, l_2, \dots, l_n\}$ , the probability of  $u$  visiting a new POI  $l$  is computed by:

$$\Pr(l|L_u) = \frac{\sum_{i=1}^n w_{u,l_i} \cdot e^{-\frac{\|l-l_i\|^2}{2\sigma^2}}}{2\pi\sigma^2 \sum_{i=1}^n w_{u,l_i}}, \quad (12)$$

where  $\sigma = 1/\sqrt{n}$ ,  $l_i = (\text{lat}_i, \text{lon}_i)$  is a two-dimensional vector with the latitude ( $\text{lat}_i$ ) and longitude ( $\text{lon}_i$ ) coordinates, and  $\|\cdot\|$  denotes the Euclidean norm. More importantly,  $w_{u,l_i}$  is the normalized weight of  $u$ 's check-in on POI  $l_i$  and is computed in terms of the tip polarity  $p_{u,l_i}$  of user  $u$  to  $l_i$ :

$$w_{u,l_i} = \begin{cases} 1, & p_{u,l_i} = \text{positive}; \\ 0.5, & p_{u,l_i} = \text{neutral}; \\ 0, & p_{u,l_i} = \text{negative}; \end{cases} \quad (13)$$

in which intuitively the check-in with the positive tip ( $p_{u,l_i} = \text{positive}$ ) has the positive influence on the user's check-in behavior whereas the check-in with the negative tip ( $p_{u,l_i} = \text{negative}$ ) has the negative influence on the user's check-in behavior. Otherwise, the influence is neutral.

**Example.** Figure 5 depicts two users' check-in distributions over the latitude and longitude coordinates, estimated through Equation (12), which model their geographical preferences on POIs: User 1 usually checks in POIs in two areas while User 2 checks in POIs in three main areas.

**Table 1: Statistics of the Foursquare data set**

	LA	NYC
Number of users	30,208	47,240
Number of POIs	142,798	203,765
Number of tips	244,861	388,594
Number of social links	349,985	810,672
User-POI matrix density	$5.68 \times 10^{-5}$	$4.04 \times 10^{-5}$

### 5.3 Combining Social-Geographical Influences

Finally, the social rating and geographical probability of user  $u$  to POI  $l$  in Equations (10) and (12) are combined into a unified score  $s_{u,l}$  based on the robust product rule [27]:

$$s_{u,l} = \hat{r}_{u,l} \cdot \Pr(l|L_u). \quad (14)$$

The top- $m$  POIs  $l$  having the highest score  $s_{u,l}$  are recommended to user  $u$ .

## 6. EVALUATION SETTINGS

This section describes our experiment settings for evaluating the performance of ORec.

### 6.1 Two Publicly Available Real Data Sets

**Yelp Challenge data set [24] for polarity detection.** This data set has 113,993 tips labeled by ratings for 15,585 POIs from various categories including Food, Nightlife, etc. We split the tips into different subsets according to their categories and evaluate polarity detection methods on each subset using cross-validation.

**Foursquare data set [1] for POI recommendations.** The Foursquare data set includes two subsets of POIs in Los Angeles (LA) and New York City (NYC). The statistics of the two data sets are shown in Table 1. The polarity of a tip is determined based on the classification model learned from the Yelp Challenge data set, because the tips in Foursquare and Yelp are essentially the same, i.e., short texts. We also evaluate POI recommendation methods using cross-validation.

### 6.2 Evaluated Techniques

**Polarity detection group.** This group includes the widely used polarity detection techniques for tweets and tips: (1) the classical text classification methods such as support vector machines (SVM) [12], decision trees (DT) and naive Bayes (NB) [17]; (2) the unsupervised aspect-independent (UAI) method [22]; and (3) the supervised aspect-dependent approach (ORec) in Section 4.

**POI recommendation group.** This group incorporates the state-of-the-art POI recommendation techniques that are the most related to our work. Most of them measure the similarity between users by using social links as the input of the collaborative filtering method. **The difference lies in the way they utilize the geographical information of POIs or whether they apply textual information.**

- IReMF: This technique exploits two levels of geographical neighborhood characteristics by integrating the item-based collaborative filtering with matrix factorization [14].
- PD: This technique estimates a geographical power-law distribution for all users and integrates it with the social user-based collaborative filtering method [23].
- iGSLR: This technique estimates a geographical nonparametric distribution for each user and integrates it with the social user-based collaborative filtering method [27].

- LCARS: This technique builds a location-content-aware recommender system using the check-in data and textual information based on the LDA topic model [25].
- UAI: Besides the social links of users and geographical information of POIs, this technique uses tip polarities from the unsupervised aspect-independent method [22].
- ORec: The proposed method integrates tip polarities with social links and geographical information into a unified recommendation framework, as described in Section 5.

### 6.3 Performance Metrics

For **polarity detection**, given a class (positive, neutral or negative), the precision and recall are defined as follows:

$$\text{precision} = \frac{\text{No. of tips correctly predicted to the class}}{\text{No. of tips predicted to the class}},$$

$$\text{recall} = \frac{\text{No. of tips correctly predicted to the class}}{\text{No. of tips actually belonging to the class}}.$$

The reported precision and recall are averaged on the three classes.

For **POI recommendations**, a discovered POI is defined as a POI that is recommended to and visited by a target user. Similarly, the precision and recall are defined as follows:

$$\text{precision} = \frac{\text{No. of discovered POIs}}{\text{No. of recommended POIs for the user: } m}.$$

$$\text{recall} = \frac{\text{No. of discovered POIs}}{\text{No. of POIs actually visited by the user}}.$$

### 6.4 Parameter Settings

**Polarity detection.** In the aspect clustering step in Section 4.2, the aspect distance is normalized into the range of [0, 1], where 0 indicates two aspects are completely identical, e.g., “ambiance” and “ambiance”, while 1 denotes two aspect with the largest difference. In the clustering algorithm, the allowed maximum distance from an aspect to its cluster center is set to  $d_{max} = 0.2$ . **Note that the number  $K$  of clusters is adaptive to  $d_{max}$ .**

**POI recommendations.** We examine the effect of various numbers of recommended POIs for users (top- $m$  from 1 to 10) and numbers of visited POIs of users in the training set ( $n$  from 1 to 10) on recommendation accuracy. Note that:  $m$  and  $n$  are set to a relative small number, because a large number  $m$  may not be helpful for users and the number  $n$  in the training set is small.

## 7. EXPERIMENTAL RESULTS

This section analyzes the polarity detection accuracy in Section 7.1 and POI recommendation accuracy in Section 7.2.

### 7.1 Polarity Detection Accuracy

**General trends and important findings.** Figure 6 compares the polarity detection accuracy of ORec, SVM, DT, NB and UAI on the Yelp challenge data sets from the top-10 active categories, e.g., Food, Nightlife, Restaurants, etc.

**SVM, DT and NB.** As the classical classification techniques for short texts, e.g., tweets or tips, the three methods still show the low precision and recall around 40%. The reason is that they consider a tip as a bag of words and transform the tip into a highly sparse word vector with a large number of missing values, so the accuracy of polarity detection is deteriorated.

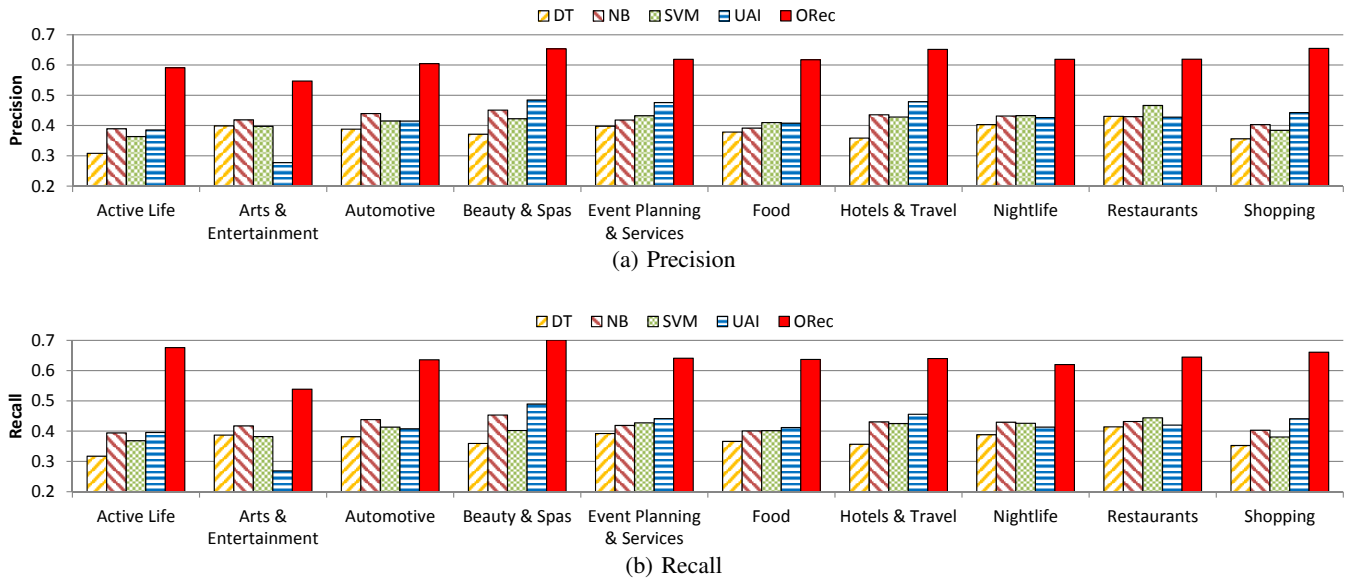


Figure 6: Polarity detection accuracy of ORec in comparison to SVM, DT, NB and UAI

**UAI.** By using sentiment values of opinion words in tips in terms of sentiment lexicons, UAI usually records better precision and recall of polarity detection in comparison to SVM, DT and NB, except for the categories of Arts & Entertainment, Automotive, Nightlife and Restaurants. However, the improvement is considerably limited due to its intrinsic limitation: it simply sums up all the sentiment values in a tip as the polarity of the tip. In other words, UAI implicitly assumes that the sentiment orientation of an opinion word is independent of its modifying aspect, which is at variance with objective reality. For example, in opinion phrases of “high quality” and “high price”, the opinion word “high” has distinct sentiment orientation: the former “high” modifying the aspect “quality” is positive whereas the latter “high” modifying the aspect “price” is negative.

**ORec.** Our ORec exhibits the best polarity detection accuracy in terms of precision and recall in all categories and accomplishes 20% absolute improvement and 50% relative improvement on the performance of SVM, DT, NB and UAI. The reason for the promising results is fourfold: (1) ORec associates each opinion word with its modifying aspect by extracting opinion phrases from tips and determining the sentiment orientation of an opinion word depending on the associated aspect, rather than using the implicit independence assumption in UAI. (2) ORec uses the supervised machine learning method to integrate the sentiment orientations of opinion words, instead of using the unsupervised summation function to combine the absolute sentiment values of opinion words in UAI. (3) To reduce the ambiguity and dimension of word vectors representing tips in SVM, DT and NB, ORec only takes aspects as dimensions and groups aspects into clusters according to their dissimilarity, which significantly decreases the sparsity in the cluster-based vector representations of tips. (4) ORec avoids replacing the missing values with default value that will bring severe biases to underlying classification models. Instead, it leverages the important classification approach, i.e., naive Bayes to naturally deal with the missing values without any replacement. These promising results verify the superiority of the proposed supervised aspect-dependent polarity detection approach over the classical classification techniques SVM, DT and NB and the unsupervised aspect-independent method UAI.

## 7.2 POI Recommendation Accuracy

**Low accuracy due to pretty low density.** It is worth emphasizing that, unlike the polarity detection accuracy, the accuracy of all POI recommendation techniques for LBSNs is usually not high, because the density of a user-POI interaction matrix is pretty low. For example, the reported maximum precision is 0.06 over a data set with  $2.72 \times 10^{-4}$  density in [23]. Even worse, the two data sets used in our experiments have a lower density,  $5.68 \times 10^{-5}$  in the LA data set and  $4.04 \times 10^{-5}$  in the NYC data set (Table 1), so the relatively low precision and recall values are common and reasonable in the experiments. Thus, we focus on the relative accuracy of ORec compared to the state-of-the-art POI recommendation techniques: IRenMF, PD, iGSLR, LCARS and UAI. We expect that ORec can improve the accuracy as more tips and check-in activities are recorded, as shown in Figure 7.

**General trends and important findings.** Here we compare the POI recommendation accuracy of ORec, IRenMF, PD, iGSLR, LCARS and UAI with various numbers of visited POIs of users in the training set (Figure 7) and numbers of recommended POIs for users (Figure 8).

**IRenMF, PD, iGSLR.** (1) Based on the combination of item-based collaborative filtering and matrix factorization techniques, IRenMF intensively models the influence of geographical neighborhood characteristics of POIs on users’ check-in behaviors in order to learn users’ geographical preferences. However, IRenMF does not consider other information, e.g., social links and textual tips. As a result, it performs the worst in most cases in terms of precision and recall. (2) PD models the distance between every pair of POIs visited by the same user as a power-law distribution for all users, and integrates it with the social links based on the user-based collaborative filtering method. Accordingly, PD improves the POI recommendation performance in comparison to IRenMF, but it still suffers from the limitation of the universal distance distribution for all users. (3) iGSLR also utilizes the social links based on the user-based collaborative filtering method, but it personalizes the geographical influence through modeling an individual distance distribution for each user. Consequently, iGSLR further enhances the recommendation precision and recall. (4) Nonetheless, the three methods (IRenMF, PD and iGSLR) only exploit check-



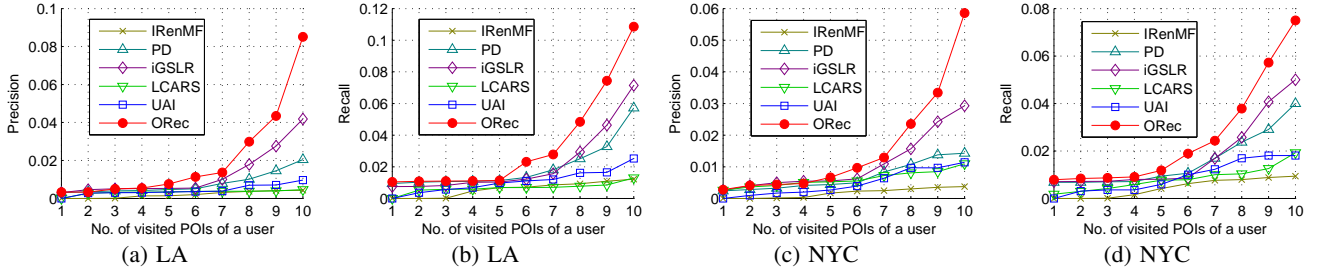


Figure 7: Effect of numbers of visited POIs of users in the training set on recommendation accuracy

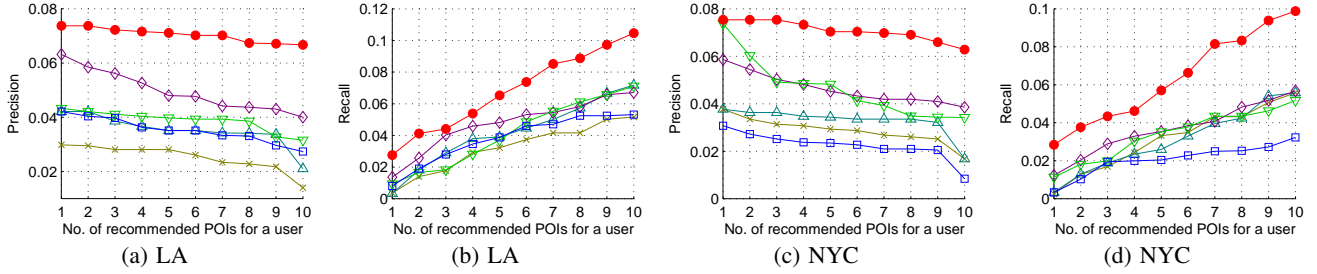


Figure 8: Effect of numbers of recommended POIs for users on recommendation accuracy (with the same legend as in Figure 7)

in data **without tips** by supposing that the check-in frequency of a user to a POI directly reflects the preference of the user to the POI, which may not be in accordance with the reality. For example, in most cases that users only check in POIs once, it is hard to deduce the true preference level of users to POIs. Subsequently, merely utilizing check-in data often generates biased user preference models that deteriorate POI recommendation accuracy.

**LCARS.** By using both check-in data and textual information in tips based on the LDA topic model, LCARS still performs poorly for the cold-start users, i.e., the users who have checked in only a few POIs, as depicted in Figure 7, but it can enhance the overall accuracy (averaged on all users) respecting different numbers of recommended POIs in comparison to IReNMF and PD, as depicted in Figure 8. Our explanation is that LCARS utilizes *topic distributions of users* rather than *polarities of tips* and the accurate topic distributions require much more textual information.

**UAI.** UAI adopts the unsupervised aspect-independent method to detect the polarity of tips and combines the acquired **tip polarities** with check-in data including **social links** and **geographical information** to make POI recommendations. In general, although UAI has better performance than IReNMF, it shows worse precision and recall than PD, especially iGSLR. The reason is that UAI inherits the intrinsic limitation of the unsupervised aspect-independent polarity detection method, as mentioned in Section 7.1.

**ORec.** ORec exploits the supervised aspect-dependent approach to predict the polarities of tips and integrates them with social links and geographical information. Accordingly, ORec exhibits the best precision and recall. In particular, ORec achieves the significant improvement compared to the second best technique, i.e., iGSLR. We attribute the promising results to two reasons: (1) The underlying polarity detection method can accurately predict the polarity of a tip as depicted in Figure 6, which guarantees that ORec is able to take full advantage of opinions indicated in tips of users for POI recommendations. (2) ORec seamlessly integrates tip polarities with social links and geographical information. **It is worth emphasizing two important points:** (i) The difference between ORec and iGSLR lies in whether they apply tips or not. ORec with tips

greatly enhances the recommendation quality against iGSLR without tips, which shows that the textual information is relevant and useful for POI recommendations. (ii) The only difference between ORec and UAI is how they utilize the tips. ORec using the sophisticated supervised aspect-dependent method is much better than UAI using the simple unsupervised aspect-independent method, which shows the effectiveness and novelty of the proposed method in this paper.

**LA vs. NYC.** In general, the precision and recall of ORec on the LA data set is a little higher than that in the NYC data set, because the density of the former data set is larger than that of the latter data set, as shown in Table 1. We believe that our ORec will perform better as more data including tips and check-ins are collected.

**Effect of the number of visited POIs of users in the training set.** Figure 7 depicts the result of the recommendation accuracy as the number of visited POIs of users increases in the training set. As expected, the precision and recall incline accordingly, since ORec can learn users' preferences on POIs more accurately with more check-in data and tips, as they visit more POIs. Interestingly, as for the small number of visited POIs, i.e., less than five, iGSLR is competitive to ORec. However, when the number of visited POIs of users is larger than five, ORec significantly outperforms iGSLR, since the precision and recall of ORec increase dramatically. This fully shows the superiority of exploiting users' opinions implied in tips for POI recommendations.

**Effect of the number of recommended POIs for users.** Figure 8 depicts the recommendation accuracy by varying the number of recommended POIs for users. The recall gradually gets higher with the raise of the number of recommended POIs, since by returning more POIs for users, it is always able to discover more POIs that users would like to visit. However, the precision generally becomes lower as the number of recommended POIs increases. Our explanation is that some recommended POIs are less possible to be liked by users due to their lower preferences; the recommendation techniques return the top- $m$  POIs based on the estimated visiting score, for example, the second recommended POI has the lower visiting score than the first one.

## 8. CONCLUSION AND FUTURE WORK

This paper proposes an opinion-based POI recommendation framework ORec to overcome the two challenges: detecting the polarities of tips and integrating them with social links and geographical information. First, ORec exploits the supervised aspect-dependent approach to learn a classification model in the cluster space of aspects to predict polarities of tips. Further, ORec seamlessly fuses *tip polarities* with *social links* and *geographical information*. Finally, experimental results on two large-scale Foursquare and Yelp data sets show that ORec achieves significantly better performance than the state-of-the-art polarity detection and POI recommendation techniques. In the future, we plan to study two directions to extend ORec: (a) how to derive weights for sentiment values of opinion words in tips, and (b) how to build regression models for polarity detection instead of classification models.

## 9. ACKNOWLEDGEMENTS

The authors were supported by Guangdong Natural Science Foundation of China under Grant S2013010012363.

## 10. REFERENCES

- [1] J. Bao, Y. Zheng, and M. F. Mokbel. Location-based and preference-aware recommendation using sparse geo-social networking data. In *ACM SIGSPATIAL*, 2012.
- [2] R. Dong, M. Schaal, M. P. O'Mahony, and B. Smyth. Topic extraction from online reviews for classification and recommendation. In *IJCAI*, 2013.
- [3] H. Gao, J. Tang, and H. Liu. gSCorr: Modeling geo-social correlations for new check-ins on location-based social networks. In *ACM CIKM*, 2012.
- [4] T. Gonzalez. Clustering to minimize the maximum intercluster distance. *Theoretical Computer Science*, 38:293–306, 1985.
- [5] Z. Hai, K. Chang, and G. Cong. One seed to find them all: mining opinion features via association. In *ACM CIKM*, 2012.
- [6] A. Hogenboom, D. Bal, F. Frasincar, M. Bal, F. de Jong, and U. Kaymak. Exploiting emoticons in sentiment analysis. In *ACM SAC*, 2013.
- [7] X. Hu, J. Tang, H. Gao, and H. Liu. Unsupervised sentiment analysis with emotional signals. In *WWW*, 2013.
- [8] G. H. John and P. Langley. Estimating continuous distributions in Bayesian classifiers. In *UAI*, 1995.
- [9] S. Kim, J. Zhang, Z. Chen, A. Oh, and S. Liu. A hierarchical aspect-sentiment model for online reviews. In *AAAI*, 2013.
- [10] T. Kurashima, T. Iwata, T. Hoshida, N. Takaya, and K. Fujimura. Geo topic model: Joint modeling of user's activity area and interests for location recommendation. In *ACM WSDM*, 2013.
- [11] D. Lian, C. Zhao, X. Xie, G. Sun, E. Chen, and Y. Rui. GeoMF: Joint geographical modeling and matrix factorization for point-of-interest recommendation. In *ACM KDD*, 2014.
- [12] S. Liu, F. Li, F. Li, X. Cheng, and H. Shen. Adaptive co-training SVM for sentiment classification on tweets. In *ACM CIKM*, 2013.
- [13] X. Liu, Y. Liu, K. Aberer, and C. Miao. Personalized point-of-interest recommendation by mining users' preference transition. In *ACM CIKM*, 2013.
- [14] Y. Liu, W. Wei, A. Sun, and C. Miao. Exploiting geographical neighborhood characteristics for location recommendation. In *ACM CIKM*, 2014.
- [15] J. McAuley and J. Leskovec. Hidden factors and hidden topics: Understanding rating dimensions with review text. In *ACM RecSys*, 2013.
- [16] S. Moghaddam and M. Ester. The FLDA model for aspect-based opinion mining: addressing the cold start problem. In *WWW*, 2013.
- [17] F. Moraes, M. Vasconcelos, P. Prado, D. Dalip, J. M. Almeida, and M. Gonçalves. Polarity detection of Foursquare tips. In *Social Informatics*. Springer, 2013.
- [18] A. Mukherjee and B. Liu. Aspect extraction through semi-supervised modeling. In *ACL*, 2012.
- [19] K. Reschke, A. Vogel, and D. Jurafsky. Generating recommendation dialogs by extracting information from user reviews. In *ACL*, 2013.
- [20] R. Socher, J. Bauer, C. D. Manning, and A. Y. Ng. Parsing with compositional vector grammars. In *ACL*, 2013.
- [21] S. Wang and C. D. Manning. Baselines and bigrams: Simple, good sentiment and topic classification. In *ACL*, 2012.
- [22] D. Yang, D. Zhang, Z. Yu, and Z. Wang. A sentiment-enhanced personalized location recommendation system. In *ACM HT*, 2013.
- [23] M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee. Exploiting geographical influence for collaborative point-of-interest recommendation. In *ACM SIGIR*, 2011.
- [24] Yelp. Challenge Data Set [accessed 25-april-2014]. [http://www.yelp.com/dataset\\_challenge](http://www.yelp.com/dataset_challenge), 2014.
- [25] H. Yin, B. Cui, Y. Sun, Z. Hu, and L. Chen. LCARS: A spatial item recommender system. *ACM TOIS*, 32(3):11:1–11:37, 2014.
- [26] J. J.-C. Ying, W.-N. Kuo, V. S. Tseng, and E. H.-C. Lu. Mining user check-in behavior with a random walk for urban point-of-interest recommendations. *ACM TIST*, 5(3):40:1–40:26, 2014.
- [27] J.-D. Zhang and C.-Y. Chow. iGSLR: Personalized geo-social location recommendation - a kernel density estimation approach. In *ACM SIGSPATIAL*, 2013.
- [28] J.-D. Zhang and C.-Y. Chow. CoRe: Exploiting the personalized influence of two-dimensional geographic coordinates for location recommendations. *Information Sciences*, 293:163–181, 2015.
- [29] J.-D. Zhang and C.-Y. Chow. GeoSoCa: Exploiting geographical, social and categorical correlations for point-of-interest recommendations. In *ACM SIGIR*, 2015.
- [30] J.-D. Zhang and C.-Y. Chow. Spatiotemporal sequential influence modeling for location recommendations: A gravity-based approach. *ACM TIST*, *accepted*, 2015.
- [31] J.-D. Zhang and C.-Y. Chow. TICRec: A probabilistic framework to utilize temporal influence correlations for time-aware location recommendations. *IEEE TSC*, *accepted*, 2015.
- [32] J.-D. Zhang, C.-Y. Chow, and Y. Li. iGeoRec: A personalized and efficient geographical location recommendation framework. *IEEE TSC*, *accepted*, 2014.
- [33] J.-D. Zhang, C.-Y. Chow, and Y. Li. LORE: Exploiting sequential influence for location recommendations. In *ACM SIGSPATIAL*, 2014.
- [34] X. Zhang, J. Cheng, T. Yuan, B. Niu, and H. Lu. TopRec: Domain-specific recommendation through community topic mining in social network. In *WWW*, 2013.
- [35] Y.-L. Zhao, L. Nie, X. Wang, and T.-S. Chua. Personalized recommendations of locally interesting venues to tourists via cross-region community matching. *ACM TIST*, 5(3):50:1–50:26, 2014.