Knowledge-Based Systems 123 (2017) 31–40

Contents lists available at ScienceDirect

Knowledge-Based Systems

journal homepage: www.elsevier.com/locate/knosys

Towards social-aware interesting place finding in social sensing applications

Chao Huang, Dong Wang*, Brian Mann

Department of Computer Science and Engineering, University of Notre Dame Notre Dame, IN 46556, United States

ARTICLE INFO

Article history:
Received 23 July 2016
Revised 27 January 2017
Accepted 1 February 2017
Available online 14 February 2017

Keywords:
Interesting place finding
Social dependency
Social sensing
Crowdsourcing
Expectation maximization

ABSTRACT

This paper develops a principled approach to accurately identify interesting places in a city through social sensing applications. Social sensing has emerged as a new application paradigm, where a crowd of social sources (humans or devices on their behalf) collectively contribute a large amount of observations about the physical world. This paper studies an *interesting place finding* problem, in which the goal is to correctly identify the interesting places in a city. Important challenges exist in solving this problem: (i) the interestingness of a place is not only related to the number of users who visit it, but also depends upon the travel experience of the visiting users; (ii) the user's social connections could directly affect their visiting behavior and the interestingness judgment of a given place. In this paper, we develop a new *Social-aware Interesting Place Finding Plus (SIPF+)* approach that addresses the above challenges by explicitly incorporating both the user's *travel experience* and *social relationship* into a rigorous analytical framework. The SIPF+ scheme can find interesting places not typically identified by traditional travel websites (e.g., TripAdvisor, Expedia). We compare our solution with state-of-the-art baselines using two real-world datasets collected from location-based social network services and verified the effectiveness of our approach.

1. Introduction

This paper develops a principled approach to accurately identify interesting places in a city through social sensing applications. This work is motivated by the emergence of social sensing as a new application paradigm of collecting observations about the physical world from social sources (humans or devices on their behalf) [1]. This paradigm is enabled by a few recent technical trends: (i) the proliferation of smart devices (e.g., smartphones) owned by average individuals; (ii) the ubiquitous coverage of wireless communication (e.g., 4G, WiFi, WiMax); (iii) the advent of online social media (e.g., Twitter, Foursquare, Facebook). For example, common citizens can now easily use a Location-Based Social Network (LBSN) service (e.g., Foursquare) on their mobile phones to upload the "check-in" points of the places they visit in a city. Alternatively, a group of drivers may use a smartphone app to report traffic conditions (e.g., congestion, accidents, etc.) they experience in a given area. In this paper, we focus on an *interesting place finding* problem, where the goal is to correctly identify the interesting places in a city where people may have strong interest in visiting (e.g., parks, museums, historic sites, scenic trails, etc.). The results of this work can be used to develop future travel recommendation systems, mobile guidance applications, and user travel experience sharing applications that explore the power of social sensing data contributed by common citizens [2,3]. For example, the results can help people find more interesting places in a city, design a better route for their travels, and share their travel experience with other users in a timely fashion.

Previous studies have adopted social sensing (in some cases referred to as crowdsourcing) to solve the interesting place finding problem. The main idea behind current solutions is to automatically infer the locations of interesting places in a given region (e.g., a city) from the check-in points or GPS traces that users share when using location-aware applications [4]. The advantages of using crowdsourcing methods compared to the traditional methods (e.g., search engine, travel websites) are threefold. *First*, the cost of data collection using crowdsourcing is low since the location data of users is already made available through the location based services (e.g., LBSN) [5]. *Second*, the interestingness of a place may change over time and the crowdsourcing methods can track such changes by analyzing the most recent trajectory data uploaded by the crowd [6]. *Third*, the crowdsourcing traces normally have a better spatial-temporal coverage of the interesting places, as the crowd is naturally distributed across the region [7]. Table 1 also shows some examples of interesting places in the cities of Chicago and San Francisco that are either not recommended by traditional

* Corresponding author.
E-mail address: dwang5@nd.edu (D. Wang).

http://dx.doi.org/10.1016/j.knosys.2017.02.006
0950-7051/

Table 1
Interesting places that are missed by traditional travel websites.

Chicago	San Francisco
North Avenue Beach	Conservatory of Flowers
Museum of Contemporary Art	Japanese Tea Garden
Oriental Theatre	Aquarium of the Bay
The Peggy Notebaert Nature Museum	Yerba Buena Gardens
Music Box Theatre	Mission Dolores Park

travel websites (e.g., TripAdvisor, Expedia, and CityPass) or have very low recommendation rankings on those websites. However, those places are identified as very interesting places by many people who visited them in person and shared their experience on social media (e.g., Twitter and Foursquare). In this paper, we develop a new social sensing based scheme that is able to find such interesting places that cannot easily be identified by traditional travel websites.

While previous studies in information retrieval [7,8] , data mining [4,9], and social sensing [10,11] have made significant efforts to address the interesting place finding problem, two important limitations remain in current solutions. *First*, the current techniques are mostly heuristic-based and make strong assumptions when they handle users in the problem. For example, they either assume all users have exactly the same travel experience[1] or the correlation between a user's travel experience and the number of places he/she visited is simply linear [8]. However, these assumptions cannot be applied in real-world scenarios where the relationship between a user's travel experience and the number of places he/she visited is *nonlinear* [13]. The interesting place finding problem becomes more challenging when neither the user's travel experience nor the interestingness of a place is known *a priori* [14]. Hence, we need to develop a new framework that can accurately model both the user's travel experience and the interestingness of places based on the social sensing data observed. *Second*, the social connections between users could easily affect their visiting behavior and the judgment on the interestingness of places they visited. For example, a group of colleagues who work in the same company are more likely to visit the same building every day; however, the building of their company may not necessarily be interesting to the general public. Unfortunately, current interesting place finding techniques completely ignore the impact of a user's social dependency, which can easily lead to suboptimal solutions, as we observed in our experiments.

In this paper, we develop a Social-aware Interesting Place Finding (SIPF+) scheme that addresses the above limitations by explicitly incorporating both the user's *travel experience* and *social dependency* into a Maximum Likelihood Estimation (MLE) framework. In particular, a principled, unsupervised learning approach based on Expectation Maximization (EM) is developed to jointly estimate both the user's travel experience and the interestingness of a place without prior knowledge on either. We evaluate SIPF+ using two real-world datasets collected from location-based social network services. The evaluation results show that our approach significantly outperforms the state-of-the-art baselines by correctly identifying more interesting places in a city while minimizing the number of false positives. The results of this paper are important because they allow social sensing applications to accurately identify interesting places by taking into account the user's travel experience and social dependency under a principled framework. To summarize, the contributions of this work are as follows:

- To the best of our knowledge, we are among the first to develop a principled, unsupervised learning framework that allows us to derive an optimal solution (in the sense of maximum likelihood estimation) for the social-aware interesting place finding problem.
- We explicitly consider both the user's travel experience and social dependency in the interesting place finding solution.
- Our MLE solution handles the nonlinear relationship between the user's travel experience and the interestingness of places.
- We perform extensive experiments on two real-world datasets, comparing the performance of our scheme to that of the state-of-the-art baselines.

A preliminary version of this work has been published in [11]. We refer to the previous version as SIPF. The current paper is a significant extension of the previous work in the following aspects. First, we extend our previous model in [11] by addressing more complex social dependency (i.e., arbitrary user dependency graph that includes cycles) between users and improve the interesting place finding results (Section 4). Second, we compare our scheme with more state-of-the-art baselines from Point of Interests (POI) recommendation systems and added more evaluation metrics (e.g., mean average precision) to evaluate the performance of our scheme (Section 5.2). Third, we perform a new set of experiments on a second city (i.e., Chicago) to further evaluate the robustness of our scheme (Section 5.3). Finally, we add a few real-world examples to demonstrate that our algorithms can identify interesting places more accurately than other baselines (Section 5.3).

2. Related work

There exists a good amount of work on the topic of Points of Interests (POI) recommendation [8,15–18]. For example, Zhang et al. developed a kernel density estimation method (i.e., iGSLR) to infer the POI based on the geographical proximity [15] and social connections between users and the categorical information of places (i.e., GeoSoCa) [16]. Zheng et al. proposed an iterative approach (i.e., HITS) to explore users' travel experience and discover the interesting places in a simple linear way by using the GPS trajectories generated by users [8]. A geo-topic model (i.e., GTM) was proposed to estimate interesting places by learning the user's activity area and various features of locations [17]. Hu et al. [18] proposed a comprehensive model (i.e., STT) that explicitly considered the geographical influence and temporal activity patterns in POI recommendations. However, most of the above solutions used supervised learning approaches for personalized POI recommendation, which requires a significant amount of labeled data to train their models. In contrast, this paper develops an unsupervised approach to address the interesting place finding problem that requires no training data.

Natural Language Processing (NLP) has received a significant amount of attention with the advent of Social Web and online social media, in particular. Computational Intelligence models and approaches have been developed to achieve a deeper understanding of natural languages by leveraging semantic features and context that are not explicitly expressed in the text [19–21]. For example, Gangemi et al. [20] developed a new model to perform holder and topic detection in opinion sentences based on the neo-Davidsonian assumption. Lau et al. [21] proposed the design, development and evaluation of a weakly-supervised cybercriminal network mining approach to facilitate cybercrime forensics. In contrast, this paper focuses on structured and easy-to-process check-in data (i.e., GPS coordinates) extracted from LBSN. However, the NLP approaches can be integrated with our tool to mitigate data sparsity problems in social sensing where people might only use text to describe the location they visited rather than share the actual GPS coordinates. NLP techniques can be used to reliably transfer the unstructured, human-generated text information to structured,

[1] The travel experience of a user is highly correlated with the user's ability to find interesting places in a city [12].