# An Images-Textual Hybrid Recommender System For Vacation Rental

Pai Chet Ng, James She, Ming Cheung
Department of Electronic and Computer Engineering
Hong Kong University of Science and Technology
Clear Water Bay, Hong Kong
Email: {pcng, eejames, cpming}@ust.hk

Alexander Cebulla
Department of Mechanical and Process Engineering
ETH Zurich
Rämistrasse 101, Switzerland
Email: acebulla@student.ethz.ch

***Abstract*—To look for the specific vacation rental that suits ones personal preferences can be time consuming due to the wealth of information available on the Internet. Often collaborative filtering is used to help people narrow down their search results. However, most of the methods are solely based on the textual data which might insufficient to capture comprehensive details about the accommodation that suits individuals' preferences. The visual effects of the images, on the other hands, might reveal hidden users' preferences which cannot be told through the text. In this paper, an images-textual hybrid recommender system is proposed to enhance the preferable vacation accommodation prediction by leveraging the strength of both data collected from users' traveling histories. The proposed recommender system is demonstrated on the Airbnb dataset for all the advertised accommodation offers in Hong Kong. Around 1 million images features are extracted from a total of 110572 accommodations' images for similarity calculation. Rooms description and review scores are collected through a custom built web crawler program, the review scores are used as a reference to filter out the low quality accommodation prior to the implementation of the proposed recommender system. The proposed hybrid recommender system achieves better recommendations with an average precision score of 36.23%, which shows a 26.44% improvement compared to the baseline, which has a mean precision score of 9.79%.**

***Index Terms*—multimedia big data, recommender system, vacation rental, image features, textual description**

## I. Introduction

In recent years, the 'sharing economy'concept has become a threat to some traditional industries [1], for example, the vacation rental services social networking site (SNS). Through SNS, hosts can advertise their vacant rooms for rent, and thus generate extra income. Users can browse the advertised rooms to find the one that is best suited to their preferences. As an alternative choice to traditional hotel service, vacation rental SNS is attractive as it features an online-to-offline (O2O) interaction, as shown in Fig. 1, where hosts and users can interact with each other online prior to their vacation visit. They can have further offline interactions during their stay which is favorable for culture exchange. Finally, hosts and users can write a review to each other after the visit.

One popular website that offers P2P vacation rental services is Airbnb, which has over 2,000,000 listings in over 190 countries across the world [2] [3]. As of today, Airbnb provides a user-friendly filter for users to filter-out their non-preferences while searching for their preferable accommodation. With the filter, users can set the price range, select their preferable bed-type and amenities. This text-based filter enabled users to locate their preferable accommodation in a shorter time. However, as more and more room listings added each day, the text-based filter alone is insufficient to meet the query requests. Frustration may arise when there are too much choices, which may lead to difficulties in locating the prefer accommodation.

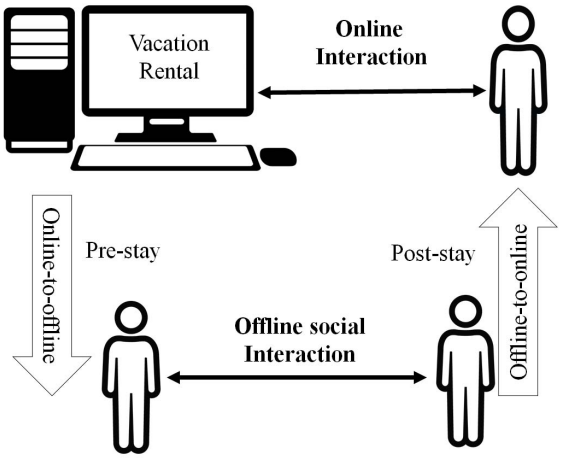


Fig. 1. Vacation rental social network site features an online-to-offline interaction

In this work, a web crawler program is built to collect all the Hong Kong vacation rental advertised on AirBnB [2], as well as the users' travel histories. Using the collected data, a hybrid recommender system is proposed. First, the cosine similarities between the images of the accommodations from users' travel histories and the potential accommodations are computed. Second, the system will filters out the non-preferences accommodations automatically according to the past filter setting inputted by the users.

The rest of the paper is organized as follows: Section 2 reviews related work for item-based recommendation; Section 3 illustrates the used methodology; Section 4 describes the experimental setup and the used dataset; Section 5 discusses the results; and Section 6 concludes the paper and presents future challenges.

978-1-5090-2179-6/16 $31.00 © 2016 IEEE
DOI 10.1109/BigMM.2016.71

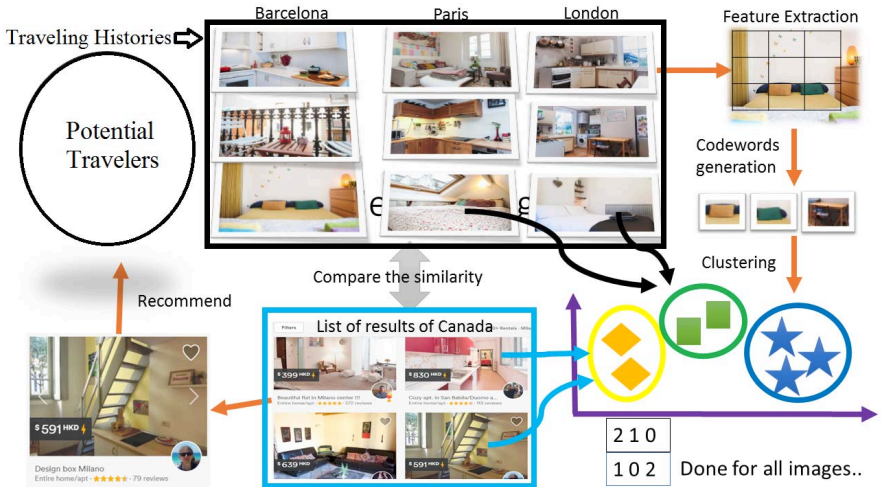


Fig. 2. Overview over the used methodology for vacation rental recommendation

## II. Related works

Online search engines have undoubtedly become a useful tool for users to find accommodations, while they plan for their vacation [4]. Most people will start their search by first specifying the destination. Then the description of the room, nearby attractions and the proximity of public transport [5] will also taken into consideration. However, there are two problems with this approach: The search query is limited by text, and users have to input their search manually.

Wth vacation rental SNS, more complex queries [6] can be obtained through O2O interaction. Furthermore, the unstructured query results are inherently more dynamic. For example, the accommodation reviews, which can influence the users'final decision in vacation planning. [7] proposed a context-aware recommender system based on such reviews and ratings for hotels. [8] presented a novel weighted algorithm to mine the review text. These selected reviews are then used to recommend hotels.

The systems described above work only on textual data. However, as the saying goes, "A picture is worth a thousand word" and images analyzing might better characterize users' traveling preferences. Using a context-aware, agent-based methodology, in [9] a system that can predict users' thinking while they are viewing images during online shopping, is built. In [10] the shape, color, and textures of shoe images are extracted. Based on these features, a shoe recommendation system with high accuracy is build. In [11] a better friendship recommendation is presented. From the information extracted from the images shared by users, the interests of these users can be estimated.

## III. Methodology

### A. Similarity between accommodations based on BoF

Based on the idea that the users' preferable accommodations are captured from their visited accommodations, the similarity between the visited and a potentially recommended accommodations are calculated. However, one must take into consideration that a user might have been previously visited some accommodations which he or she disliked. Therefore, these accommodations should not be taken into account. To do so, the accommodations with a low review score are first filtered out before compute for the images-based similarities.

As shown in Fig. 2, each images are represented with a Bag of Features (BoF) [12], which describe an image with a vector. The vector can then be used to calculate similarities between images. To create a BoF the following steps are necessary:

1) *Feature Extraction:*
   For feature extraction, the images are overlay with a regular grid, and for each grid point the histogram of oriented gradients (HoG) descriptor are calculated.
2) *Codebook Generation:*
   To generate the codebook, first the extracted feature descriptors are clustered with k-means. The set of cluster centers is then used.
3) *Representing an image with BoF:*
   The features of each images are assigned to the entries in the codebook. It is then counted how many features were assigned to each entry.

Next, the cosine similarities, $S_c$, between a single image (respectively BoF) of the visited and the potentially recommended accommodations are computed.

$$S_c = \frac{\vec{a}^T \cdot \vec{b}}{|\vec{a}||\vec{b}|} \tag{1}$$

where $\vec{a}^T$ is images features that reflect users' traveling histories, and $\vec{b}$ is the images features of reference accommodation to be predicted.

The process is then repeated for all images of the visited accommodations, and then the potentially recommended accommodations are then ranked based on these similarities and the top $N$ accommodations are recommended.

TABLE I
DETAILS OF THE SCRAPED DATASET

| Items | Total |
|---|---|
| advertised room | 305 |
| visited guests | 1090 |
| images | 110572 |
| room descriptions | 161806 |

### B. Similarity between accommodations based on textual descriptions

Besides specifying the vacation destination, most of the vacation rental SNS often offer a range of filters to allow users narrow down their search. The filter settings inputted by users manually in the past generally reflect their travels' preferences. Therefore, to measure the text-based similarity, Jaccard similarity, $S_j$, are used:

$$S_j = \frac{|A_{av} \cap A_{pr}|}{|A_{av}| + 1}, \tag{2}$$

where $A_{av}$ and $A_{pr}$ are the sets of filter descriptions of an visited and a potentially recommended accommodation, respectively. Since a potentially recommended accommodation should not be penalized for providing more amenities, only the cardinality of $A_{av}$ is considered in the denominator.

### C. Hybrid Recommender System

The information obtain from both methods above might be different, both also reflects users' traveling preferences, but maybe in different angle. One should understand that not all preferences can be captured by textual data alone, and most of the time the images have the power to tell the hidden information which cannot be captured by text simply.

As both methods have their own strength, an images-textual hybrid recommender system is proposed. This is a multi-modal recommendation approach by fusion the similarity results obtained from both images and textual description. The top $\frac{N}{2}$ from previous 2 methods are combined to achieve a better recommendation.

## IV. EXPERIMENTAL SETUP

### A. Dataset

A web crawler program is built to scrape the users' traveling histories on Airbnb from the 22nd November 2015 to the 30th November 2015. The crawler program started by scraping an initial list of all the advertised accommodations in Hong Kong. The scraped data contains a total of 305 vacation accommodations. The details of 1090 users who have been to Hong Kong are collected, as shown in Table I. The scraping process is repeated for the other cities where these 1090 users are also have been to.

### B. Dataset Processing

There are four csv files generated during the scraping: initial room list, host-user relationship, room description and links to room images. The room images file consists of a list of image links scraped from the room list that users have been to. The images were processed in parallel on 10 virtual machines. In total 110572 images are downloaded and over 1 million features are extracted from them. After generating the codebook, experiment are conducted to recommend the accommodations which have a minimum review score of 20. The experiments are repeated by increasing the minimum review score to 90 with a step of 10 each increment.

To evaluate the recommender system, one of the accommodations users had been to was removed as ground truth. All the evaluated accommodations are ranked based on the methods described in Section III, and the top $N$ accommodations are recommended, in Airbnb case, $N$ is chosen to be 18.

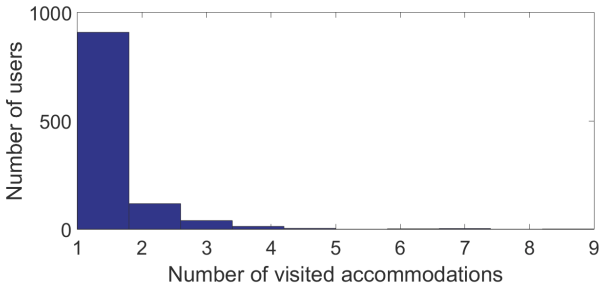


Fig. 3. The trip record of the users that have been to Hong Kong

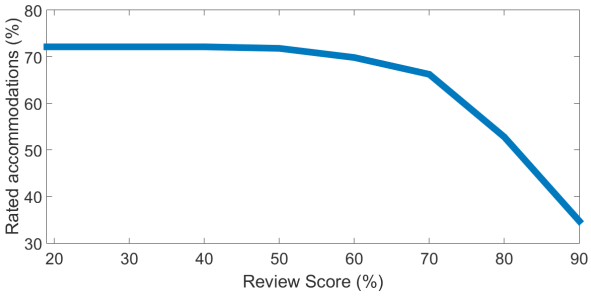


Fig. 4. Percentage of review score with respect to the ratio of rated accommodations over total accommodations advertised in Hong Kong

## V. EXPERIMENTAL RESULTS

### A. Results

The host-user relationship is a directed graph relationship, where the degree distribution follows a power law distribution, as shown in Fig. 3. Among the 1090 users, there are 912 first time users, and 115 second time users. Only 63 users are frequent travelers, who have completed more than 2 trips using Airbnb. The highest trip record ever completed by a user is 9.

The review score indicates the users' experiences with the accommodation during their stays. Fig. 4 shows the percentage of accommodations which have a review score better then the minimally required one. From the graph, it is observed that more than 70% of the accommodations have a review score less than 70%. Only a 45% of the accommodations have a review score of 80% and above. As review scores are an important factor, which users will consider when they are planning

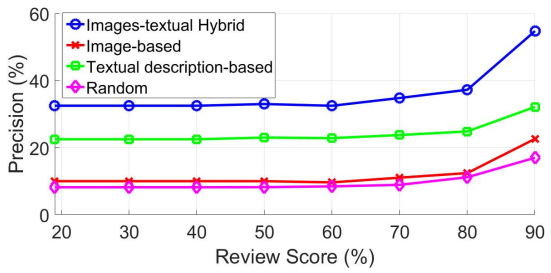


Fig. 5. Statistical analysis based on precision

for their next trip, they have been taken into consideration. In this, accommodations with a low review score were filtered out to increase the quality of the recommender system.

Recall and precision are the two statistical measures that have been used to analyze the implemented recommender systems. Intuitively, recall defines the number of accommodations that can be predicted with respect to the total number of accommodations visited, whereas precision describes how good the prediction is based on the number of predictable accommodations. Fig. 5 compares the precision for all the implemented recommender methods by varing the review score from 20% to 90%. The graphs have clearly depict the performance of the proposed hybrid recommender system.

### *B. Discussion*

The wealth of information shared on vacation rental SNS allows users to acquire sufficient information to plan for their vacation trip. One of them are review texts submitted by the users. From the experiments, it is shown that the higher the required review score, the better the precision. However, the recall are degraded since more and more accommodations are filtered out.

Users' traveling preferences can be estimated by analyzing the images and textual data collected from users' traveling history. The visual effect of images helps to understand the type of accommodations that will attract the users at the first sight. Whereas the textual description estimates the accommodations' conditions that users will consider while planning for their vacation. Neither images or textual data give a comprehensive information of traveling preferences. As both methods suffer from their own limitation, some information can be captured by this method but miss out by the other. In view of this, a hybrid method is proposed to leverage the strengths of both methods. Experiments have proven that a better recommendation can be achieved with the hybrid recommender system.

## VI. Conclusion

Vacation rental SNS has become the best substitution for the hotel room due to its affordable rental and scalable O2O interaction platform. Accommodation reviews, descriptions and images available on the website are the major factors that influences users' decision while planning for their vacation. However, the abundance of these information has make vacation rental searching a daunting task. To enhance the searching efficiency, three methods were suggested in this work: images-based cosine similarity calculation, textual description-based jaccard similarity calculation, and the fusion of both prior mentioned methods. It was shown that recommendation based solely on image similarities or textual description-based similarities are possible. However, it seems that the traveling preferences of users could not be captured completely using either one. The images-textual hybrid recommender system achieves a better result as this method is able to capture the information miss out by either image or textual data. However, this method is only applicable for users who have histories, hence for future work, it might be interesting to tackle the cold-start problem for first time users who do not have any travel history.

## Acknowledgment

This work was supported by HKUST-NIE Social Media Laboratory.

## References

[1] W. J. Penn, Joanna". (2015) Uber, airbnb and consequences of the sharing economy: Research roundup. [Online]. Available: http://journalistsresource.org/studies/economics/business/airbnb-lyft-uber-bike-share-sharing-economy-research-roundup

[2] Airbnb. (2015) About us. [Online]. Available: https://www.airbnb.com/about/about-us

[3] G. Zervas, D. Proserpio, and J. Byers, "The rise of the sharing economy: Estimating the impact of airbnb on the hotel industry," *Boston U. School of Management Research Paper*, no. 2013-16, 2014.

[4] D. R. Fesenmaier, Z. Xiang, B. Pan, and R. Law, "A framework of search engine use for travel planning," *Journal of Travel Research*, p. 0047287510385466, 2010.

[5] B. Pan, S. W. Litvin, and T. E. O'Donnell, "Understanding accommodation search query formulation: The first step in puttingheads in beds'," *Journal of Vacation Marketing*, vol. 13, no. 4, pp. 371–381, 2007.

[6] N. B. Ellison *et al.*, "Social network sites: Definition, history, and scholarship," *Journal of Computer-Mediated Communication*, vol. 13, no. 1, pp. 210–230, 2007.

[7] N. Hariri, Y. Zheng, B. Mobasher, and R. Burke, "Context-aware recommendation based on review mining," *General Co-Chairs*, p. 27, 2011.

[8] A. Levi, O. Mokryn, C. Diot, and N. Taft, "Finding a needle in a haystack of reviews: cold start context-based hotel recommender system," in *Proceedings of the sixth ACM conference on Recommender systems*. ACM, 2012, pp. 115–122.

[9] O. B. Kwon, "i know what you need to buy: context-aware multimedia-based recommendation system," *Expert Systems with Applications*, vol. 25, no. 3, pp. 387–400, 2003.

[10] N. Saxena, V. Khosla, A. Venkataraman, N. Khosla, and V. Venkataraman, "Building an image-based shoe recommendation system," 2013.

[11] M. Cheung and J. She, "Bag-of-features tagging approach for a better recommendation with social big data," in *Proceedings of the 4th International Conference on Advances in Information Mining and Management (IMMM'14)*, 2014, pp. 83–88.

[12] E. Nowak, F. Jurie, and B. Triggs, "Sampling strategies for bag-of-features image classification," in *Computer Vision–ECCV 2006*. Springer, 2006, pp. 490–503.