Deducing Trip Related Information from Flickr

Adrian Popescu
Institut Télécom/TELECOM Bretagne
Technopôle Brest-Iroise
29238 Plouzané
+33 229001435
adrian.popescu@telecom-bretagne.ee

Gregory Grefenstette
Exalead
10 Place de la Madeleine
75008 Paris
+33 155352766
gregory.grefenstette@exalead.com

ABSTRACT
Uploading tourist photos is a popular activity on photo sharing platforms. These photographs and their associated metadata (tags, geo-tags, and temporal information) should be useful for mining information about the sites visited. However, user-supplied metadata are often noisy and efficient filtering methods are needed before extracting useful knowledge. We focus here on exploiting temporal information, associated with tourist sites that appear in Flickr. From automatically filtered sets of geo-tagged photos, we deduce answers to questions like "how long does it take to visit a tourist attraction?" or "what can I visit in one day in this city?" Our method is evaluated and validated by comparing the automatically obtained visit duration times to manual estimations.

Categories and Subject Descriptors
H.2.8 [**Database Management**]: Database applications – data mining

General Terms
Algorithms, Experimentation, Human Factors.

Keywords
Tourist sites, visit times, geographical gazetteer, Flickr, georeferencing, text mining, image mining.

1. INTRODUCTION
E-tourism services, such as HomeAndAbroad [2], present the users with the possibility of planning their trips interactively. Given a destination from a list of around 100 cities, the users can compose thematic trips (from a closed list of themes like Family Time, Local Culture or History Buff) and the system then composes a trip based on the user's choices. Visit duration times are one key component of planning trips. But the duration of time one spends visiting a site is highly variable, as is reflected in the wide estimations provided on sites. For example, published estimations of visit times can vary from 30 minutes to 8 hours for the *Louvre* or from 30 minutes to 3 hours for *Notre Dame de Paris* [2]. These estimates available on e-tourism sites are either provided by field experts or supplied by users. Manual construction of such tourist information has a cost, and a priori estimates can be inexact, so we found it interesting to explore ways of automating the mining of objective data available in user contributed data. In this paper, we exploit an automatically constituted gazetteer [3] and the sets of geo-tagged and time-stamped images in order to extract information about trips described by Flickr photos and to deduce visiting durations for attractions in four major cities.

Information extraction from geo-tagged documents is a known problem but, to the best of our knowledge, past research was not concerned with discovering visit duration times. [4] extracted a geographic and events database from Flickr metadata using a burst analysis technique by which 85% of the automatically mined place names were correct and were also approximately situated. In [3], we extracted a geographical gazetteer from Panoramio and Wikipedia metadata. Each site was named (90% accuracy), categorized as to type of site (92% accuracy) and situated (60% of the elements within 200 m from their actual locations). [1] analyzed the tourist flows in the Province of Florence, Italy, based on a corpus of geo-annotated Flickr photos and their results contribute to understanding how people travel.

2. METADATA FILTERING AND STRUCTURING
Here, we attempt to see if we can deduce supplemental information, the typical visit duration of a tourist site, from a combination of user-supplied noisy metadata and the content of a geographical gazetteer. First, using the Flickr API, we gathered all the metadata (geo-tags, tags, title, textual description, owner etc.) from photos from 2006 onwards found in a square of length 13 km around four popular tourist destinations: London (400,000 images), New York (250,000), Paris (150,000) and San Francisco (130,000). We sort this data by user, and then by day. We retain the set of pictures for a given user on a given day, if (i) there are N>=20 pictures, (ii) if the time stamp between the first image and the Nth image is more than one hour, and (iii) if the unique sort of the concatenated annotations of each of the N pictures yielded more than N/3 different textual annotations. This third constraint is useful since many tourists use bulk uploads with generic tags for all the pictures. Here, we only select pictures tagged by a discriminating tourist who individually tags photos. These three criteria, then, are used to simulate someone who took a variety of pictures over an extended period on a given day (the likely behavior of a tourist), and who individually tagged the photos. At this stage, we have a number of photos such as a tourist would take during a day of visiting, with their timestamps.

Next, we extracted time related site information for the photos in the following way. Using a large coverage geographical gazetteer, Gazetiki [3], which contains tourist sites names, site types, and their GPS coordinates, we consider photos annotated with the site

WWW 2009, April 20–24, 2009, Madrid, Spain.
ACM 978-1-60558-487-4/09/04.

name, taken near its coordinates by a user on a given day, as descriptive for that tourist site. "Near" is defined as 1 km for buildings and other constructions, and 5 km for parks and other large visitor attractions and it is used for selecting only pictures annotated with a site name that were actually taken around the site's coordinates. Then for each tourist site name in the day's photos, we calculate its earliest and latest timestamp. We consider that the user was "visiting" the site if (i) the interval between the earliest and latest timestamp is at least ten minutes inclusive, and (ii) there are at least 5 images manually tagged with that site name within that interval, and (iii) the interval between any two images is not longer than 20 minutes or 5 times the average time between photos tagged with the name (whichever is smaller). These three additional constraints try to capture the activity of someone present at a site for while and taking a number of pictures that the tourist will later tag with the site name (i.e., "visiting" the tourist site), and regularly taking pictures of the tourist attraction during the visit. After this step, for each site name meeting these criteria, we have the length of the tourist's visit.

The resulting individual trips can be reused to propose to a new user that they visit the same tourist sites that someone else has already visited in a given time interval. From many individual behaviors we can thus estimate usual behavior, by averaging all the times found for each visitor attraction, and automatically add these estimations of visit times to the tourist sites descriptions in the gazetteer. A visit duration estimation, similarly to that on HomeAndAbroad, as well as minimum, maximum and average visit times are associated to each tourist site examined. One of the prime interests of this extraction and deduction method is its low cost and easy applicability to any tourist site found in Flickr, or any other geo-tagged photographic collection.

3. EVALUATION AND FUTURE WORK

In order to evaluate our estimates, we compare the times we found with a list of visit times found on HomeAndAbroad, which are manual estimates. We retained visitor attractions for which we extracted at least three different user visits (42 in London, 38 in New York, 18 in Paris and 33 in San Francisco). For each attraction, we plotted the minimum, the average and the maximum visit time compared to the minimum and maximum visit times proposed by HomeAndAbroad [2]. The intersection between the sites, computed using a strict chain matching procedure, identified by our approach and those in HomeAndAbroad is 53 visitor attractions (figure 1).

The results in Figure 1 show that the average visit times obtained from Flickr (FLICKR AVG – yellow) are situated between the HAA MIN and HAA MAX estimations in a large majority of cases and these results seem to validate our method for automatically deducing visit times from Flickr. The FLICKR AVG is closer to HAA MIN for high values of HAA MIN, indicating that tourists tend to spend less time than expected visiting sites to which human experts assign a long minimum visit time. The maximum visit times extracted from Flickr have a large variation compared to HAA MAX values and are, with few exceptions bigger than the corresponding HAA MIN values. Interestingly, the FLICKR MIN values are significantly smaller than corresponding HAA MIN values and indicate that people often visit sites more quickly than estimated by human advisors. This last result can also be explained by the fact that HomeAndAbroad experts consider that a visit always includes an interior tour of the attraction whereas Flickr visits are often limited to an exterior visits.

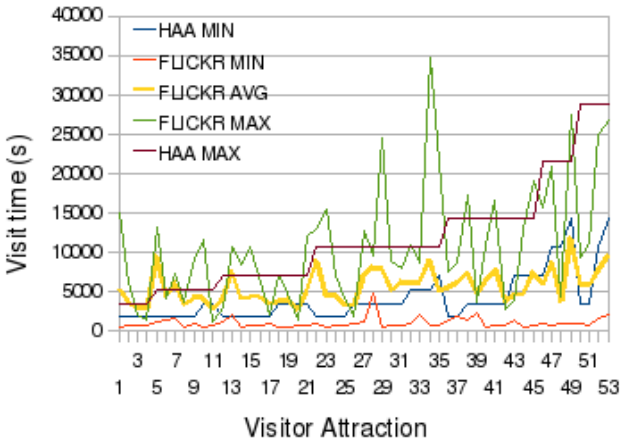

Figure 1. Comparison of Flickr based visit times – minimum (FLICKR MIN), average (FLICKR AVG) and maximum (FLICKR MAX) - to HomeAndAbroad minimum (HAA MIN) and maximum visit times (HAA MAX).

The main limitation of our method comes from the number of places that are sufficiently annotated in Flickr, though existing metadata are already considerable (more than 50 million geo-tagged photos) and their volume is constantly growing (around three million new photos each month). Our results show that visit times can be accurately extracted from user contributed data and we feel that the accuracy of the extraction, as well as its coverage, are likely to improve with the enrichment of Flickr metadata. We plan on using an indoor/outdoor image classifier in order to separate visits including inside views of the sites from exterior visits. A major future work direction will be the development of an interactive e-tourism application based on automatically extracted tourist information which will integrate such visit times in order to recommend trips according to user preferences.

4. ACKNOWLEDGMENTS

This research is part of Georama, a French research project funded by ANR.

5. REFERENCES

[1] Girardin, F., Dal Fiore, F., Blat, J., Ratti, C. Understanding of Tourist Dynamics from Explicitly Disclosed Location Information. In Proc. of the *4th International Symposium on LBS and Telecartography (*Hong-Kong, China, 2007).

[2] Homeandabroad – http://homeandabroad.com

[3] Popescu, A., Grefenstette, G., Moëllic, P.-A. Gazetiki: automatic construction of a geographical gazetteer. In Proc. of JCDL 2008 (Pittsburgh, PA, June 2008).

[4] Rattenbury, T., Good, N., Naaman, M. 2007. Towards Automatic Extraction of Event and Place Semantics from Flickr Tags. *In Proc. of SIGIR 2007* (Amsterdam, The Netherlands, July 2007).